Fragmentation through Heavy and Light-flavor Measurements with the LHC ALICE Experiment

Zoltán Varga

Ph.D. Thesis

Advisor: Róbert Vértesi, Ph.D.
Dept. of High Energy Experimental Particle and Heavy Ion Physics
HUN-REN Wigner Research Centre for Physics

Consultant: László Szunyogh, Ph.D.
Department of Theoretical Physics
Budapest University of Technology and Economics

Doctoral School of Physical Sciences
Budapest University of Technology and Economics
Budapest, Hungary
2024

Acknowledgments

Here I would like to say thank you to those without whom I could have not been able to finish my PhD thesis.

First and foremost, I am deeply grateful to my advisor Róbert Vértesi. I first met Robi in 2017, when he gave a lecture at my university, which immediately inspired me to work with him, and our early research projects truly laid the foundation of my research interest. Even before he became my advisor, I received so much support from him. He not only taught me how to become a good researcher, but has always been there for me when I needed his help. I've always felt grateful to have an advisor like him.

I must also thank Gergely G. Barnaföldi, my ALICE group leader and research collaborator. I learned a lot from him throughout the years and wrote my very first research article together with him and Robi. His optimistic attitude, encouragement and support helped me to get through some of the hard times in my research career, and I truly appreciate it.

I am also thankful to Laura Havener, with whom I started working in the ALICE Collaboration a few years ago, and taught me a lot about data analysis. Her positive and encouraging personality, combined with her physics insights, had a great impact on me.

Last but not least I am the most grateful to my family: my parents and two sisters, for their unwavering support and belief in me.

The research in this doctoral thesis was supported by the NKFIH grants OTKA FK131979 and K135515, as well as by the 2019-2.1.11-TÉT-2019-00078 and 2021-4.1.2-NEMZ_KI-2024-00034 grants, and by the Wigner Scientific Computing Laboratory (WSCLAB) and the HUN-REN Wigner Cloud. Experimental contributions were facilitated by the Veszter-gombi High Energy Laboratory at the HUN-REN Wigner RCP.

Contents

List of Figures

Chapter 1

Introduction

A few microseconds after the Big Bang, the universe was filled with an extremely hot and dense mixture of particles moving at near light speed. This matter was dominated by quarks and gluons that are subject to the strong force. Under usual circumstances, quarks are confined into hadrons. However, at extreme temperatures and densities, quarks and gluons can enter a deconfined state, forming the so-called quark-gluon plasma (QGP).

To recreate the extreme conditions of matter which existed in the very early universe, we use powerful accelerators to make collisions between protons or heavy-ions, such as gold or lead nuclei. In the heavy-ion collisions hundreds of protons and neutrons in two such nuclei collide into each other to form a minuscule fireball in which everything melts into the quark-gluon plasma. The fireball then starts cooling and the individual quarks and gluons (collectively called partons) recombine into ordinary matter that flies away in all directions until they end up in our detectors. The final state of the collision contains many species of particles such as pions and kaons, which are mesons consisting of a quark-antiquark pair; or protons and neutrons, which are baryons containing three quarks; and even antiprotons and antineutrons, which may combine to form the nuclei of antiatoms as heavy as helium.

Even though heavy-ion collisions are only on the femtoscale, in high-energy physics they are often referred to as large systems, because the volume of the system created in the collision is large enough for the quark-gluon plasma to form. In contrast, we call the system of colliding protons (or antiprotons) small system. In proton–proton collisions, no nuclear effects are present, therefore they serve as a baseline for heavy-ion collisions, while proton–heavy ion collisions are investigated to learn about cold nuclear matter effects. Recently, however, collisions of small systems with high final-state multiplicity became a focus of intensive investigation. With the advent of the Large Hadron Collider (LHC), features have been observed that had been associated with the presence of quark-gluon plasma, even though QGP is not expected to be formed in small systems in a significant volume. These observations include collectivity effects such as long-range

near-side correlations and the asymmetry in the azimuthal distribution of final state particles, also called anisotropic flow. A possible explanation is that quark-gluon plasma can be created in a small volume in a fraction of collisions of small systems. There are, however, alternative explanations of these observed phenomena that do not assume the quark-gluon plasma. Semi-soft vacuum-QCD effects such as multiple-parton interactions are shown to produce signatures of collectivity. The aim of this thesis is to study heavy and light-flavor jet fragmentation and hadronization properties in high-energy proton–proton collisions to shed light on the particle production mechanisms that lead to collective-like behavior in small systems.

One of the main research areas is the study of jets. I characterized differential and integral jet shapes to look for modifications caused by non-trivial quantum chromodynamics (QCD) effects. I also investigated the multiplicity distributions (number of charged final state particles) as a function of the jet transverse momentum. Recent results show that the multiplicity distributions follow a scaling similar to the Koba-Nielsen-Olesen (KNO) scaling, which provides important lessons on jet fragmentation. In my studies, I searched for the KNO-like scaling in simulations for heavy-flavor jets, and I also carried out the first measurement of jet-momentum-dependent jet multiplicity distributions with the ALICE experiment.

Another major topic in my thesis, connected to jet fragmentation, concerns the understanding of heavy-flavor hadroproduction in proton–proton collisions at LHC energies. The production cross section of hadrons can be calculated using the factorization theorem, which usually assumes that the fragmentation functions are universal across different collision systems. Experimental results such as the enhanced production of charm baryons question this assumption. I used a model with color reconnection beyond leading color approximation (CR-BLC) to seek explanation for the charm-baryon enhancement, and proposed new observables for future measurements. I characterized the collision events using different event-activity classifiers, that allow for investigating the connections between the leading QCD processes and the underlying event.

The primary motivation driving my research is to explore the boundary between hard and soft processes occurring in proton–proton collisions, which is a significant yet relatively unexplored area. Keeping these overarching objectives in mind, the thesis is structured as follows.

In Chapter 2 the theoretical background for the study of high-energy hadron collisions is laid out to serve as a basis for understanding the results presented in this thesis. I start with detailing the most important properties of the strong interaction, one of the four fundamental interactions of Nature, which is also the most relevant interaction in our case. After a brief discussion of the phase diagram of the strong interaction and its implications, we are going to turn our attention to the experimental observables that we can measure. Many of these observables are related to collimated sprays of hadrons, called jets. A section is dedicated to the properties and experimental signatures of the QGP. Following

that I discuss the role of heavy-flavor in the analysis of hadron collisions and discuss the event classifiers. To conclude this chapter, I will cover the relevant experimental facilities, detailing the Large Hadron Collider (LHC) and the ALICE apparatus.

In Chapter 3 we dive into the study of jet shapes and their multiplicity dependence by measuring the differential and integral jet shapes as a function of the event multiplicity. My main goal is to show that multiplicity differential jet shape observables are experimentally accessible using data from the LHC Run 3 data taking period, and to motivate such measurements. Not only can we observe a multiplicity dependence, but a characteristic jet size was also found, which is a certain transverse-momentum-dependent radius of the reconstructed jets for which the transverse momentum density is independent of the event multiplicity regardless of the choice of physics models and jet reconstruction algorithms. Finally, I introduce a double ratio of the multiplicity dependent jet shapes, to observe this dependence while minimizing the biases in this measurement.

The scaling properties of jets, which are closely connected to the subject of multiplicity dependence, are discussed in Chapter 4. Motivated by the recently revealed KNO-like scaling inside the jets, which indicates that the KNO scaling is violated by complex vacuum-QCD processes outside the jet development, I investigated the flavor-dependence of this scaling. I demonstrated that using heavy-flavor, the origin of scaling can be pinned to the partonic level. These studies motivate experimental verification that can be carried out using the ALICE detector.

In Chapter 5 I present an ongoing measurement of the multiplicity distributions as a function of the jet momentum, that will allow for the experimental verification of KNO-like scaling within jets. This was motivated by the phenomenological studies discussed in the previous chapter. The technical steps of the analysis such as the evaluation of corrections for detector efficiency and momentum smearing are also presented.

In Chapter 6 I discuss the measurement of azimuthal correlations of heavy-flavor decay electrons in ALICE data. I participated in this analysis mainly by performing detailed Monte-Carlo simulations which serve as input for the corrections in the measurement and the interpretation of the data. Experimental data are compared to the simulation results in proton–proton and proton–lead collisions at a center-of-mass energy of $\sqrt{s} = 5.02\ \mathrm{TeV}$.

Chapter 7 is devoted to my studies on the charm baryon enhancement, one of the hot research topics in recent years. Experimental results coming from multiple LHC experiments imply that the commonly assumed universality of the fragmentation functions (i.e. collision system independence) does not hold, and a so-far not completely explained charm baryon enhancement is observed in proton–proton collisions. Several models aim at explaining these results. I use a model with color reconnection beyond leading color approximation (CR-BLC) to investigate the expected details of the charm-baryon enhancement and to propose new methods of measurements.

Finally, in Chapter 8 I summarize the results and the main conclusions presented in this thesis and provide an outlook for future research.

Chapter 2

High-energy Hadron Collisions

In this chapter I present the theoretical background of high-energy hadron collisions, which is necessary to understand the results presented in this thesis. I introduce the Standard Model (SM) of particle physics in Sec. 2.1, which is the theory that provides the best description of fundamental particles and their interactions. In Sec. 2.2 I discuss the quantum field theory of the strong interaction, quantum chromodynamics (QCD) in detail. Simulating hadron collision events to test QCD predictions is so complex that it requires us to utilize Monte Carlo event generators. After describing these event generators in Sec. 2.3, I turn to one of the most important experimental consequences of theory in Sec. 2.4: the collimated streams of particles created from a hard scattering process, known as jets. This section covers the reconstruction of jets and the jet shape observables that we can measure. Another important consequence of QCD is the existence of the extremely hot and dense matter, the quark-gluon plasma, which is discussed in Sec. 2.5, including its most important properties and experimental signatures. The focus of this thesis is heavy-flavor measurements, I present some of the most important heavy-flavor results in Sec. 2.6. The underlying event affects measurements, therefore I discuss the subject in Sec. 2.7, including the technical details of the event classifiers that I utilize throughout this thesis. Finally, I close the chapter with Sec. 2.8 by describing the Large Hadron Collider at CERN, and the technical details of the ALICE detector.

2.1 The Standard Model

In the Universe there are four known fundamental interactions: the electromagnetic, the weak and the strong forces, and gravity. At high-enough energy scales the electromagnetic and weak interactions can be unified into the electroweak interaction, but so far it is an open problem to unify all the interactions.

The Standard Model (SM) of elementary particle physics describes three of the known fundamental interactions except for gravity (which is neglected in the analysis of hadron

collisions due to its weakness compared to the other forces), and classifies the elementary particles similarly to the periodic table of the elements. These elementary particles are categorized by their spin in two groups: in the units of the reduced Planck constant $\hbar$, the bosons have spin in integer values, while the fermions have half-odd-integer values. According to the spin-statistics theorem, the spin dictates the kind of quantum statistics the particles follow. The quantum-mechanical wave function of a boson is symmetric under particle exchange, while the fermion wave function is anti-symmetric and they obey Pauli's exclusion principle. The former particles follow Bose-Einstein statistics, the latter follow Fermi-Dirac statistics.

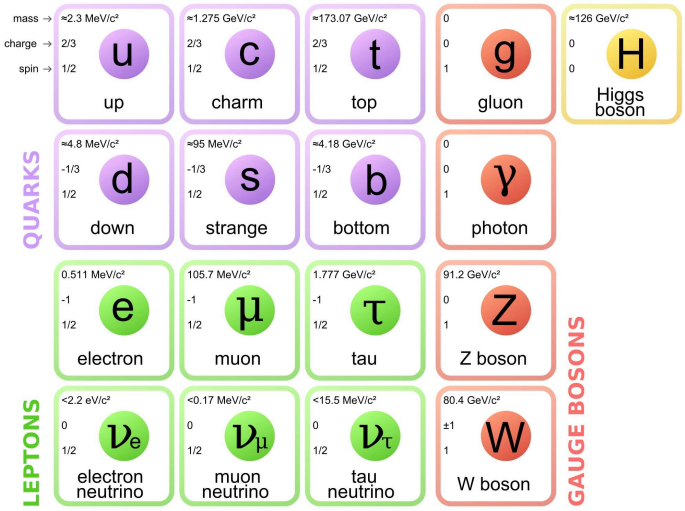

Figure 2.1: The elementary particles of the Standard Model.

The elementary fermions are the 6 quarks and 6 leptons. The leptons themselves consist of the electron, muon and tau particles, and their associated neutrinos. Furthermore, both quarks and leptons can be divided into 3 generations, where each generation has a quark pair and lepton pair (see Fig. 2.1, taken from [1]). The generations reflect the chronology of their discoveries. It is important to note that all fermions have an antiparticle. Quarks form colorless bound states called hadrons, which can be mesons (quark-antiquark pairs), baryons (three quarks) or antibaryons (three antiquarks).

The elementary bosons are categorized into vector (or gauge) bosons with spin 1, and a scalar boson with spin 0. The gauge bosons are associated with the fundamental interactions as their force carrier particles: the photon (γ) with the electromagnetic interaction, the 8 gluons (g) with the strong interaction, and the $W^{\pm}$ and Z^0 bosons with the weak interaction. The only known elementary scalar boson is the Higgs boson (H^0) which is the quantum of the Higgs field that is responsible for the mass generation of the elementary particles through the Higgs mechanism. The mechanism behind the neutrino mass generation is still an open question.

The Standard Model not only categorizes the elementary particles, but describes the dynamics between these fundamental constituents, i.e. how these particles interact with each other. Currently the best description is provided by quantum gauge theories, which are based on symmetry principles. More precisely, we use the Lagrangian formalism generalized for quantum fields, where the form of the Lagrangian density is dictated by its invariance under local gauge transformations. The non-Abelian gauge theory describing the SM is based on the $SU(3)_c \times SU(2)_L \times U(1)_Y$ symmetry, where c denotes the color degrees of freedom, "L" is an abbreviation for "left-handed", and Y stands for hypercharge. The $SU(3)_c$ group is the symmetry group of the strong interaction, while $SU(2)_L \times U(1)_Y$ is the symmetry group describing the electroweak interaction. Before discussing the gauge theory of the strong interaction, we should recall what a quantum field theory is.

In classical point mechanics, we can use Hamilton's principle, which states that the dynamics of a physical system are determined by a variational problem for the Lagrangian L, which is a functional containing all information of a system which is required to determine its equations of motion. In classical field theory, we use instead the Lagrangian density $\mathcal{L}$ (referred to simply as Lagrangian from now on). The action functional S is then defined as

$$S[\psi(x)] = \int_V d^4x \, \mathcal{L}\big(\psi(x), \partial_\mu \psi(x)\big) \,, \tag{2.1}$$

which gives us the equations of motion if we solve the variational problem for the 4D spacetime volume. The action in this example is varied by the $\psi(x)$ field, where x is the spacetime coordinate.

The Lagrangian formalism has the advantage of making the symmetries of the theory more prevalent. One way to go from classical field theories to quantum field theories is to use second quantization, in which the quantum Hamiltonian is assumed to take the form of the classical Hamiltonian, then the Poisson relations of the classical quantities are replaced with commutation relations. The resulting theory is a quantum field theory with the fields becoming operator valued distributions, and unlike non-relativistic quantum mechanics, it can describe both particle creation and annihilation processes.

Another way to quantize a classical field theory is to use the path integral formalism, which is the commonly used approach to quantize gauge theories. A proper action functional $S(\phi, g, m)$ is first chosen, where ϕ denotes all fields, g denotes the coupling constants, and m denotes the masses and possible extra parameters. The expectation value of a quantum mechanical operator is then calculated by integrating over all possible field configurations, weighted by the exponential of the action times the imaginary unit ($\hbar = 1$):

$$\mathcal{O}[\phi] = \frac{\int \mathcal{D}\phi \, \mathcal{O}[\phi] e^{iS[\phi,g,m]}}{\int \mathcal{D}\phi \, e^{iS[\phi,g,m]}} \,. \tag{2.2}$$

2.2 The Strong Interaction

Quantum chromodynamics (QCD) is one of the main pillars of the Standard Model, which describes the strong interaction that couples to color charges and is mediated by eight gluons. The color charge is the QCD analog of the electric charge in electromagnetism (EM), but while the symmetry of the electromagnetic interaction is connected to the Abelian Lie group $U(1)$, the symmetry of QCD is connected to the more complex non-Abelian Lie group $SU(3)$. A significant difference between the force carrier particles of QCD and EM is, that while the photons do not carry electric charge, the gluons carry color charge, therefore they participate in the strong interaction and can interact with each other. This property of the strong force is closely connected to the non-Abelian symmetry of the theory. As an important consequence, the gluons can form bound states with each other, called glueballs.

In the rest of this section, I briefly describe the mathematical description of quantum chromodynamics (QCD) and summarize the most important elements of the theory. For a more detailed discussion I refer to the literature [2–5].

2.2.1 The QCD Lagrangian

The fundamental degrees of freedom of QCD are the fermionic quark fields $\psi(x)$ and bosonic gluon fields $A^a_\mu(x)$ with $a = 1, ..., 8$, where the gluon fields belong to the adjoint representation of the $SU(3)$ Lie group. The full QCD Lagrangian is given by

$$\mathcal{L}_{QCD} = -\frac{1}{4} F^a_{\mu\nu} F^{a\,\mu\nu} + \bar{\psi} \left(i \gamma^\mu \mathcal{D}_\mu - m \right) \psi, \tag{2.3}$$

consisting the pure gauge part and the fermionic part. The quarks have six different flavors, therefore the fermionic part contains six separate terms, one for each flavor. The pure gauge part contains the field strength tensor

$$F^a_{\mu\nu} = \partial_\mu A^a_\nu - \partial_\nu A^a_\mu + g\, f^{abc} A^b_\mu A^c_\nu, \tag{2.4}$$

where g is the coupling constant and f^{abc} is the structure constant of the SU(3) group. The form of the quark-gluon interaction is given by the covariant derivative

$$\mathcal{D}_\mu = \partial_\mu - i\, g\, A^a_\mu \frac{\lambda^a}{2}, \tag{2.5}$$

which has the form to allow comparing the fields at different points of the space-time manifold by parallel transporting the fields while keeping the Lagrangian invariant under local phase transformations of $\Psi(x)$. The $\lambda^a/2$ are the generators of $SU(3)$ in the fundamental representation (the quark fields belong here), where λ^a are the Gell-Mann matrices.

The invariance of the Lagrangian under local gauge transformations of the fields is described by the following equations:

$$\psi(x) \to G(x)\,\psi(x), \tag{2.6}$$

$$\bar{\psi}(x) \to \bar{\psi}(x)\,G^{\dagger}(x), \tag{2.7}$$

$$A_{\mu}(x) \to G(x)\,A_{\mu}(x)\,G^{\dagger}(x) - \frac{i}{g}\,[\partial_{\mu}\,G(x)]\,G^{\dagger}(x). \tag{2.8}$$

In the functional integral formalism, the vacuum expectation values of an n-point operator can be obtained as

$$\langle 0|T\{O_1(x_1)\,...\,O_n(x_n)\}|0\rangle = \frac{\int\,\mathcal{D}\psi\,\mathcal{D}\bar{\psi}\,\mathcal{D}A_{\mu}\,O_1(x_1)\,...\,O_n(x_n)\,e^{\,i\,S[\psi,\,\bar{\psi},\,A_{\mu}]}}{\int\,\mathcal{D}\psi\,\mathcal{D}\bar{\psi}\,\mathcal{D}A_{\mu}\,e^{\,i\,S[\psi,\,\bar{\psi},\,A_{\mu}]}}, \tag{2.9}$$

where we have to integrate over all the possible field configurations to obtain the expectation value.

To resolve the issue of wildly oscillating integrals due to the imaginary unit in the exponent, we proceed to change the metric from Minkowski to Euclidean. This is achieved by the transformation known as Wick-rotation, where we make a rotation in the complex plane by an angle $-\pi/2$ and change the currently used time τ to time t in Euclidean metric by

$$\tau \to -i\,t. \tag{2.10}$$

With the above transformation the action no longer has the imaginary factor and the exponential changes as $e^{iS} \to e^{-S_E}$, where the Euclidean action takes the form

$$S_E = \int d^4x\,\mathcal{L}_E = \int d^4x \left\{ \frac{1}{4} F^a_{\mu\nu} F^a_{\mu\nu} + \bar{\psi}(\gamma^{\mu}\mathcal{D}_{\mu} + m)\psi \right\}. \tag{2.11}$$

The final form of the functional integral then becomes

$$\langle 0|T\{O_1(x_1)...O_n(x_n)\}|0\rangle = \frac{\int\,\mathcal{D}\psi\,\mathcal{D}\bar{\psi}\,\mathcal{D}A_{\mu}\,O_1(x_1)...O_n(x_n)\,e^{-S_E[\psi,\,\bar{\psi},\,A_{\mu}]}}{\int\,\mathcal{D}\psi\,\mathcal{D}\bar{\psi}\,\mathcal{D}A_{\mu}\,e^{-S_E[\psi,\,\bar{\psi},\,A_{\mu}]}}. \tag{2.12}$$

To calculate cross-sections and perturbative QCD (pQCD) processes, we use the functional integral formalism derived from the Euclidean action to obtain the Feynman rules from the QCD Lagrangian. These rules are then applied in the context of Feynman diagrams to compute scattering amplitudes [6].

2.2.2 Properties of the Strong Interaction

Now that we have discussed the form of the QCD Lagrangian, I briefly mention three important properties of the strong interaction, all of which has far-reaching consequences. One of them is the phenomenon known as asymptotic freedom. It states that the (running) coupling of the theory, denoted by α_s, goes to zero at high energies (or short distances). Due to this property, QCD becomes amenable to the methods of perturbation theory at sufficiently high energies. The exact dependence of the strong coupling on the energy scale can be calculated using the renormalization group method.

The second important property of QCD is color confinement, namely that the coupling α_s becomes larger at lower energy scales (or longer distance scales). As a consequence, using the perturbation expansion in the coupling at low energy scales leads to a failure due to the coefficients of the expansion being infinite. In the low-energy regime therefore non-perturbative methods are required, the most successful being lattice gauge theory. An experimental consequence of this property of the running coupling is that when two color charges are being separated, the force between them remains constant while the energy grows linearly, resulting in the breakup of the color string by the production of a quark-antiquark pair. This process repeats as long as sufficient energy is present. At the end, quarks and antiquarks will be confined into colorless states instead of acting as isolated color charges.

The third important property of QCD is the dynamical spontaneous breaking of chiral symmetry, which generates masses for hadrons far above the mass scale of quarks, while making the masses of pseudoscalar mesons exceptionally low.

2.2.3 The QCD Phase Diagram

Part of the quest to understand the properties of strongly-interacting matter is to reveal the structure of the phase diagram of quantum chromodynamics. A phase diagram shows the different phases in which matter can exist and also provides information on the required conditions for the phase transition to occur. The phase diagram of QCD is usually drawn as a 2-dimensional diagram with the temperature on the vertical axis and the baryon chemical potential (or the net baryon density) on the horizontal axis, as it is shown in Fig. 2.2 (taken from [7]). Another important illustration is provided by the Columbia plots in Fig. 2.3 (taken from [8]), which show two possible scenarios of the quark-mass dependence of the hadronic matter to QGP transition.

As it was mentioned earlier, we need to employ non-perturbative methods to fully describe QCD due to the limitations of the perturbative approach. Lattice gauge theory is the most successful non-perturbative theory, in which we discretize spacetime into a 4D lattice and try to solve the QCD equations numerically on a computer using Monte Carlo simulations. The success of this approach is well represented by the ab initio calculation of the

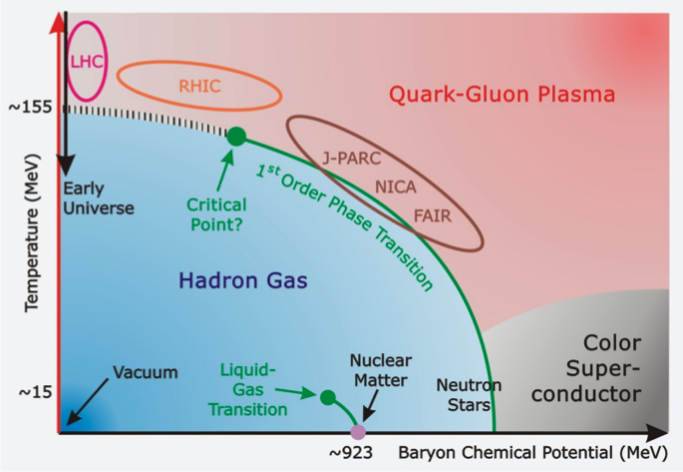

Figure 2.2: A sketch of the QCD phase diagram.

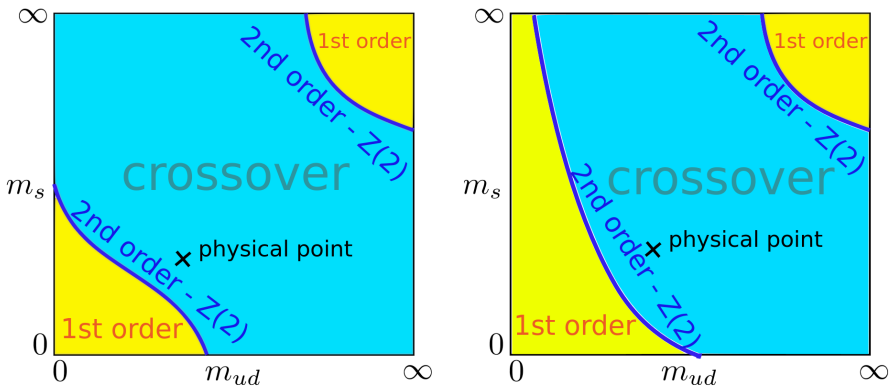

Figure 2.3: Two possible scenarios for the Columbia plot of QCD, which shows how the hadronic matter to QGP transition depends on the quark masses.

light hadron masses shown in Fig. 2.4 [9]. However, ab initio calculations in the domain of non-zero baryon number density are plagued by the infamous sign-problem, in which the numerical evaluation of the QCD integrals fails because of the near-cancellation of the negative and positive contributions to the integrals. There are multiple proposed solutions to alleviate the problem, such as reweighting the integrals, Taylor expansion around zero chemical potential, analytic continuation from purely imaginary chemical potentials, or the complex Langevin method. Since none of these approaches work satisfactorily, only a small region of the phase diagram can be successfully investigated with lattice QCD, which makes experimental measurements even more essential to guide the theory.

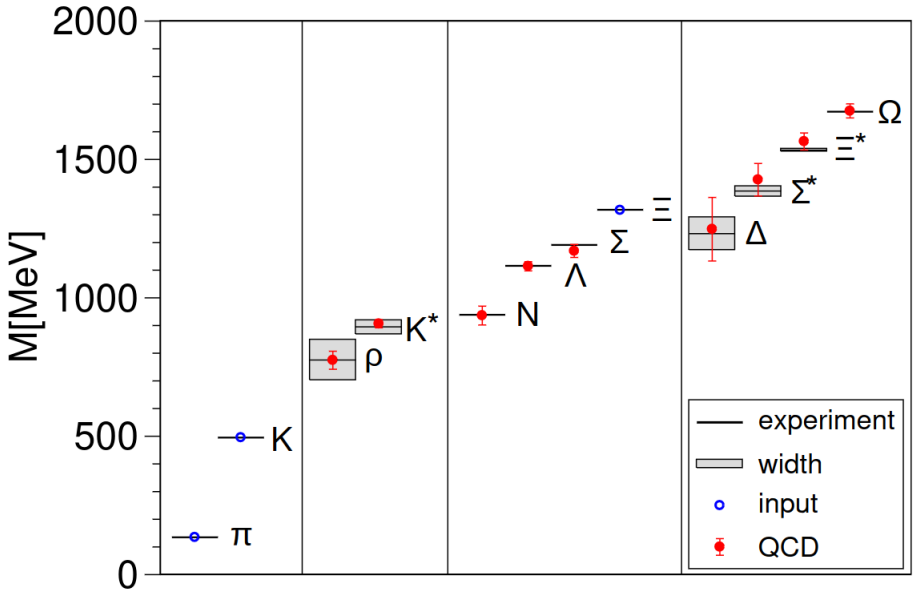

Figure 2.4: The light hadron spectrum of QCD. Horizontal lines and bands are the experimental values with their decay widths. Vertical error bars represent the combined statistical and systematic errors. The π, K and Ξ hadrons were used to set the light quark mass, the strange quark mass and the overall scale, respectively.

2.2.4 Hot & Cold QCD

Research of the strong interaction is sometimes divided into two categories: hot and cold QCD. The former aims to understand the emergence of the fluid behavior of the quark-gluon plasma (see Sec. 2.5) from the theory of QCD. Another important research area is to characterize the temperature and chemical potential dependence of the properties of the quark-gluon plasma.

The latter focuses on obtaining a complete picture of quarks and gluons giving rise to the spin, momentum and mass of the nucleons. Its focus includes the structure of the hadrons, i.e. the description of the bound states of QCD in terms of the quarks and gluons, and the study of hadronization. Finally, it also studies the interactions of hadrons, including the effects due to the color flow in the collision processes.

2.3 Monte Carlo Event Generators

Monte Carlo (MC) event generators utilize pseudo-random sequences of numbers and combine multiple theoretical and phenomenological models for different stages of a collision event to give a full, and best possible description of high-energy collisions. They provide a fully exclusive final state in terms of particles which resembles the real-world experiments as closely as possible, meaning that they generate events with the same probabilities as they occur in nature. Since many of the results in this thesis rely on data simulated by MC event generators, and it also plays an essential role in the analysis of experimental data both in the estimation of systematic uncertainties and to correct for detector effects, I give an overview of MC event generators in this section.

As the name suggests, MC event generators rely on Monte Carlo integration techniques, in which we numerically compute definite (usually) high-dimensional integrals utilizing pseudo-random numbers. Unlike regular integration techniques, we evaluate the integrand at randomly chosen points of the integration space, instead of integrating on a regular grid [10]. A few important advantages of the MC integration are that it works in arbitrarily complex integration regions and converges faster in higher dimensions, the error estimation is simple, multiple quantities can be evaluated at the same time, and a good estimation is achievable with relatively few sample size. A disadvantage is that in low dimensions this method converges slowly.

There are multiple techniques to perform MC integration, one of the most important being the importance sampling method, in which we sample the integrand with such a probability distribution that the most significant contributions are the most likely to be part of the sample [11]. It is important to note that the error convergence of MC simulations goes as $1/\sqrt{N}$ with the sample size N. The error scales the same even if we generalize the simulations to higher dimensions.

The simulation of high-energy collision events requires performing so many integrations that Monte Carlo techniques are the only viable options. Event generators carry out the event simulation in the following phases: hard process, multiple scatterings, parton shower, hadronization, hadronic rescattering, secondary decays. In general we can categorize the MC generators into two classes: general purpose event generators and specialized event generators. Typically we use a combination of both.

The hard processes normally involve $2 \to 2$ (or in some cases $2 \to 3$) scatterings, and many specialized programs are available for the calculation of matrix elements of high-multiplicity final states. To calculate the cross sections, we first need to calculate the matrix element for a given phase space point, and do an integration over the phase space region of interest (depending on the particular cuts). Since the matrix elements can have multiple peaks and singularities, integrating them over the phase space is often the most challenging part.

Some of the most commonly used MC generators are PYTHIA [12], HERWIG [13, 14], SHERPA [15], POWHEG [16–18] and EPOS3 [19]. In the following subsections, I describe some of them in more details, to provide background for the results presented in this thesis.

2.3.1 PYTHIA 8

PYTHIA 8 is a general-purpose event generator for high-energy particle collisions, which was written from scratch in C++ as a successor to PYTHIA 6. Its physics base includes concepts such as soft and hard processes, parton showers, multiple-parton interactions (MPI) and string fragmentation.

PYTHIA models a basic hard scattering process with leading-order pQCD calculations, combined with initial- and final-state radiations as well as MPI at the partonic level. The hadronic final state is produced using Lund string fragmentation, and then secondary decays and rescattering between hadrons are computed, forming the final state of the collision. Since many physical details cannot be derived from first principles due to our limited understanding of Nature, the MC event generators, including PYTHIA, require extra input parameters. Determining these parameters are far from trivial, and a given set of parameters are generally sufficient only for reproducing certain experimental data. A given configuration of these parameters, optimized for reproducing experimental results in certain physical aspects, are called tunes.

The Monash 2013 tune is mainly focused on describing the minimum-bias and underlying event (UE) distributions accurately. The Monash tune uses the NNPDF2.3LO PDF set [20], is specifically configured to both $\mathrm{e^+e^-}$ and $\mathrm{pp}/\mathrm{p\bar{p}}$ data [21]. MonashStar (or CUETP8M1-NNPDF2.3LO) is an underlying event tune based on the Monash tune and was configured to CMS data [22]. The 4C tune is a newer one introduced with PYTHIA version 8.145 [23]. It is based on the tune 2C, but it uses the CTEQ6L1 PDF set [24] and has further changes including a reduced cross section for diffraction and modified multiple-parton interaction parameters to produce a higher and more rapidly increasing charged pseudorapidity plateau for better agreement with some early key LHC numbers [25].

The Angantyr Model

PYTHIA 8 does not natively support collisions involving nuclei, therefore this feature is implemented separately in the Angantyr model, which combines several nucleon–nucleon collisions to build a proton–nucleus (p–A) or nucleus–nucleus (A–A) collision. In this model some modifications are made over the dynamics of pp collisions. The Angantyr model improves the inclusive definition of collision types of the FRITIOF model [26, 27]. In this model a projectile nucleon can interact with several target nucleons where one primary collision looks like a typical pp non-diffractive (ND) collision. However, other target nucleons may also undergo ND collisions with the projectile.

The Angantyr model treats secondary ND collisions as modified single-diffractive (SD) interactions. For every p–A or A–A collision, nucleons are distributed randomly inside a nucleus based on a Glauber formalism similar to the one described in Ref. [28]. This model is able to correctly reproduce final-state observables of heavy-ion collisions, i.e., multiplicity and p_{T} distributions [29]. As explicit hydrodynamic evolution is not incorporated in this model, its predictions serve as a baseline for studying observables sensitive to collective behavior in p–A and A–A systems.

2.3.2 HIJING++

HIJING++ (Heavy Ion Jet INteraction Generator) [30, 31] is the successor of the widely used original HIJING [32, 33], completely rewritten in C++ from scratch. The main motivation during the development was to obtain an event generator that is easily maintainable, extensible and works effectively with high throughput. It contains all the physical models that were also present in HIJING, but removes some of its limitations. It includes native thread based parallelism, an easy-to-use analysis interface and a modular plugin system, which makes room for possible future improvements. An initial tune of the parameters was performed for RHIC and LHC energies.

HIJING++ implements a mechanism for creating the underlying event and QCD effects on the soft-hard boundary differently from PYTHIA. It uses the PYTHIA jet fragmentation, but instead of MPI as implemented in PYTHIA, HIJING++ uses minijet production.

2.3.3 EPOS3

The EPOS3 event generator is largely used for the description of ultra-relativistic heavy-ion collisions. It employs a core-corona description of the fireball produced in these collisions: in its inner part, the ”core”, the quark–gluon plasma is formed, which evolves based on hydrodynamic description, while in the external regions of the ”corona” the partons fragment and hadronize independently. A study of radial flow performed with the EPOS3 event generator in proton–proton collisions at $\sqrt{s} = 7$ TeV [34] has shown that the energy density reached in such collisions is large enough to grant the applicability of the hydrodynamic evolution to the core of the collision.

2.4 Jets

Jets are the collimated sprays of hadrons created from the hard scattering of partons (quarks and gluons) by fragmentation and hadronization. Their existence is a direct consequence of the confinement property of QCD (see Section 2.2.2). Quark and gluon jets radiate proportionally to their color factors, which makes it possible to differentiate between them experimentally. Quark jet fragmentation is also sensitive to the flavor of the initiating quark because of the mass-dependent dead-cone effect [35, 36].

2.4.1 Kinematic variables

A particle with energy E, intrinsic mass m and impulse $\vec{p}$ is denoted by the four-vector $p_\mu = (E, \vec{p}) = (E, p_\mathrm{x}, p_\mathrm{y}, p_\mathrm{z})$, using the convention $c = 1$. The component of the impulse parallel to the beam direction is called the longitudinal component and is denoted by p_L. Since the z-axis of the coordinate system is usually chosen to be parallel to the beam direction, the longitudinal component of the impulse is simply $p_\mathrm{L} = p_\mathrm{z}$. The transverse component of the impulse is $p_\mathrm{T} = \sqrt{p_\mathrm{x}^2 + p_\mathrm{y}^2}$, which falls in the plane vertical to the beam direction.

To describe the relativistic kinematics of a particle in the z direction, we can introduce a Lorentz-additive quantity, rapidity, as $y = \frac{1}{2}\ln\frac{E+p_\mathrm{z}}{E-p_\mathrm{z}}$. In practice we often use the pseudorapidity η, calculated as $\eta = \frac{1}{2}\ln\frac{|\vec{p}|+p_\mathrm{z}}{|\vec{p}|-p_\mathrm{z}} = -\ln[\tan(\frac{\theta}{2})]$, where θ is the angle between the particle and beam axis. In the low-mass or ultra-relativistic limit ($p_\mathrm{L} \gg m$) the pseudorapidity is equivalent to the rapidity. The 3rd degree of freedom is usually expressed by the azimuth angle ϕ, perpendicular to the beam direction.

2.4.2 Jet Reconstruction

Since jets reach the detector as a set of multiple particles, they first need to be reconstructed from collision data, before any of their properties can be determined. In practice the reconstruction is done by utilizing jet reconstruction algorithms, which serve as an interface between theory and experiment [37].

The earliest jet algorithms, such as the SisCone algorithm, suffered from the lack of ”collinear” and ”infrared” safety. Collinear safety means that predictions remain unchanged under collinear (parallel) particle emissions, while infrared safety ensures predictions are unaffected by low-energy (soft) particle emissions, both crucial for reliable theoretical calculations. In the following, I focus on collinear and infrared-safe state-of-the-art sequential recombination algorithms. These algorithms start from a list of particles, and recombine them step-by-step to form the jet. A distance is first defined between two particles (indexed by i and j) as $\Delta R_{ij}^2 = (\phi_i - \phi_j)^2 + (\eta_i - \eta_j)^2$, where ϕ is the azimuth angle and η is the pseudorapidity coordinates of the particle. We also define a p_T-weighted

distance as

$$d_{ij} = \min(p_{\mathrm{T}i}^{2p}, p_{\mathrm{T}j}^{2p}) \frac{\Delta R_{ij}^2}{R^2} \ , \tag{2.13}$$

where R is the resolution parameter, corresponding to the size of the jet. The resolution parameter is typically chosen to be a number between 0.2 and 1.0, depending on the collision system and the physics goals of the particular analysis. Heavy-ion collision analyses tend to use smaller resolution parameters due to the underlying event contribution.

The parameter p in Eq. 2.13 is typically chosen to take one of the following three values: $p = 1$ corresponds to the k_{T} algorithm [38], $p = 0$ corresponds to the Cambridge-Aachen algorithm [39], and $p = -1$ corresponds to the anti-k_{T} algorithm [40]. The jet reconstruction algorithms then work as follows:

- Calculate the distance d_{ij} between every pair of final state particles, and the distance d_{iB} to the beam.
- Find the minimum between d_{ij} and d_{iB}.
- If the minimum is the d_{ij} then combine the ith and jth particles and go back to step 1.
- If the minimum is the d_{iB} then the ith particle is a final state jet and so remove it from the list of particles and go back to Step 1.
- Repeat the above steps until no particles left.

The three main jet algorithms have their advantages and disadvantages, so it is important to understand when to prefer one over the other. The k_T algorithm is usually utilized by theorists, but it is less favored by experimentalists due to the high computational cost (scales with N^3 or $N\log(N)$ in FASTJET [41]). The Cambridge-Aachen algorithm clusters the particles based only on their geometry and has the same disadvantage as the former algorithm by being slow and not producing circular shaped jets. However, an advantage to this algorithm is that it reflects the angular ordering of the hadrons inside the reconstructed jet. Perhaps the most widely used algorithm is the anti-k_T algorithm, which starts the clustering with the most energetic particles and grows the jet around this seed. The advantage of such clustering are the resulting approximately circular shaped jet areas, but the disadvantage is the lower efficiency of capturing the branching processes inside the jets, illustrated in Fig. 2.5 (taken from [40]).

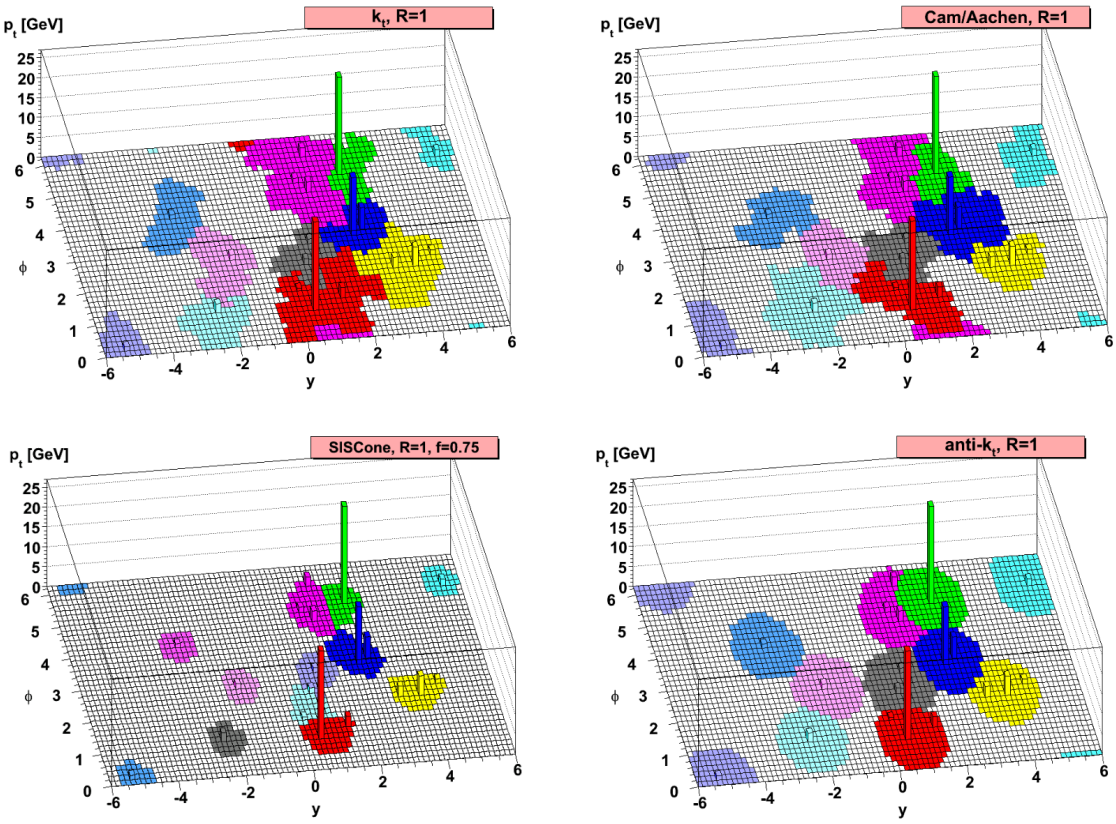

Figure 2.5: Illustration of the resulting jet areas for different jet reconstruction algorithms.

2.4.3 Jet Shapes

More information can be obtained on jet fragmentation by studying the jet substructure. Jet shapes are one such class of observables, which measure the average distribution of energy flow within the jets. Their usefulness involves them being a good tool for testing showering models in MC generators, help discriminate between different underlying event models, and could even help in searches for new physics. An interesting application is to use them to separate quark and gluon jets, due to this observable being sensitive to the quark/gluon mixture.

Two widely used observables to describe the momentum distribution inside the jets are the differential jet shape ρ and the integrated jet shape Ψ (see Fig. 2.6, taken from [42]). The differential jet shape describes the radial transverse momentum distribution inside the jet cone and is defined as follows:

$$\rho(r) = \frac{1}{\delta r}\frac{1}{p_{\mathrm{T}}^{\mathrm{jet}}}\sum_{r_a<r_i<r_b} p_{\mathrm{T}}^{i} \ , \tag{2.14}$$

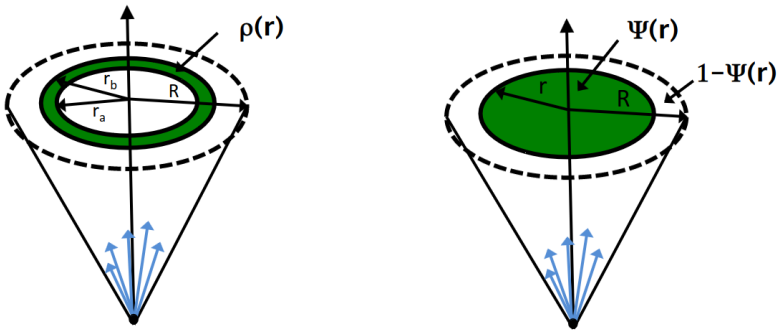

Figure 2.6: Illustration of the differential jet shape (left) and integrated jet shape (right).

where p_{T}^{i} is the transverse momentum of a particle inside a δr wide annulus with an inner radius $r_a = r - \delta r/2$ and an outer radius $r_b = r + \delta r/2$ concentric to the jet axis and $p_{\mathrm{T}}^{\mathrm{jet}}$ is the transverse momentum of the jet (i.e. a four-momentum sum of the jet constituents). The distance of a given particle from the jet axis is given by $r_i = \sqrt{(\phi_i - \phi_{\mathrm{jet}})^2 + (\eta_i - \eta_{\mathrm{jet}})^2}$, where ϕ denotes the azimuth and η denotes the pseudo-rapidity.

The integrated jet shape is defined as the average fraction of the jet transverse momentum inside a cone of radius r concentric to the jet axis and is calculated as

$$\Psi(r) = \frac{1}{p_{\mathrm{T}}^{\mathrm{jet}}} \sum_{r_i < r} p_{\mathrm{T}}^{i} \,, \tag{2.15}$$

where the symbols denote the same quantities as for the differential jet shape. The differential and integral jet shapes are connected through the following relations if there is no discretization:

$$\Psi(r) = \int_0^r \rho(r')dr' \,, \quad \text{and} \quad \Psi(r = R) = \int_0^R \rho(r')dr' = 1 \,, \tag{2.16}$$

where R is the jet resolution parameter, i.e. the size of the jet.

2.5 The Quark–Gluon Plasma

A new phase of strongly interacting matter has long been predicted by both phenomenological considerations and lattice QCD calculations. A phase transition was expected from ordinary hadronic matter to a deconfined phase of strongly interacting quarks and gluon, if the temperature (energy density) is sufficiently large [43–45]. This predicted state of matter, first referred to as the quark-gluon plasma in Ref. [46], is assumed to have existed in the early stages of the universe. In the laboratory we can create this matter by colliding heavy nuclei together in ultra-relativistic heavy-ion collisions. One of the main goals of building large heavy-ion colliders was to find and study the properties of the QGP.

In the early 2000s, a new phase of strongly interacting matter was found at the Relativistic Heavy Ion Collider (RHIC) [47–50]. The initial temperature of this matter was found to be above the limit where hadronic material can exist, proving that QGP comes into existence in high-energy heavy-ion collisions. Contrary to expectations, this matter was found to be strongly coupled (sQGP). In this section I detail the properties of this extremely hot and dense matter, and describe the experimental signatures through which it can be studied, such as particle yields, spectra (momentum distributions) and particle correlations.

2.5.1 Suppression of High-p_{T} Hadrons and Jets

Quarks and gluons do not exist in free form due to color confinement. After the initial interaction of partons during a collision event, they are leaving the vicinity of the interaction point and fragment into hadrons while forming a narrow cone of hadrons, called jets (see Sec. 2.4). The original partons can be studied through these jets that we measure in our detectors. Interactions with large momentum and energy transfer are referred to as hard processes, and in these elementary hard processes most jets are produced in back-to-back pairs.

Since the deconfined phase of QGP is strongly interacting, the created medium of this matter is opaque for partons that carry color charge, which causes the medium traversing jets to lose energy. Statistically the quark-antiquark pair is created asymmetrically in the medium, i.e. not in the center of the medium, therefore one of the jets can escape the medium faster, while the other one of the pair has to traverse the medium. In this case we expect a suppression in the detected jets, which we call jet quenching. An illustration is shown in Fig. 2.7 (taken from [51]).

One of the first important results in the study of the QGP is the observed suppression of hadrons in heavy-ion collisions, which is a clear sign of a strongly interacting QGP. The PHENIX experiment at RHIC measured the spectra of identified hadrons in both pp and Au–Au collisions at the same $\sqrt{s_{\mathrm{NN}}} = 130\ \mathrm{GeV}$ center-of-mass energy per nucleon [52]. The hadron spectra in pp was scaled with the average number of binary (nuclear–nuclear)

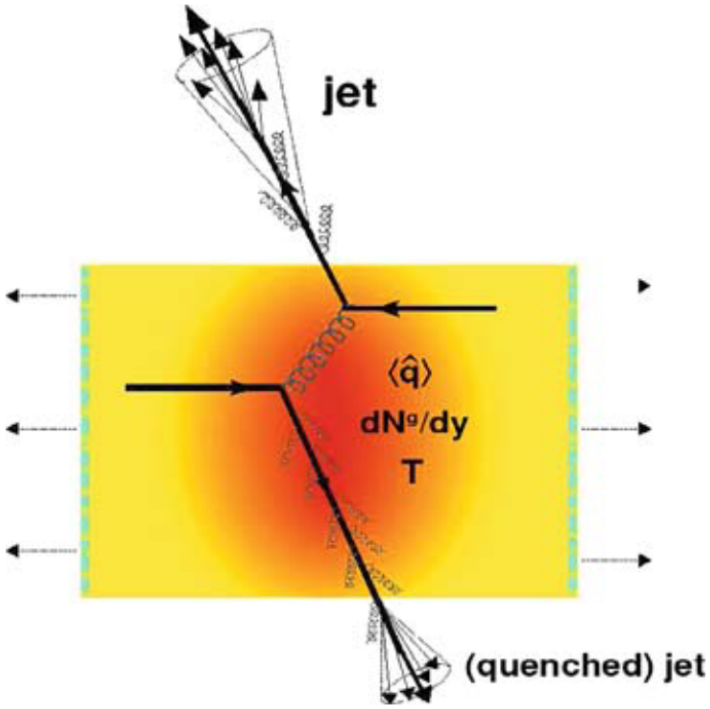

Figure 2.7: Illustration of jet quenching. One quark goes out directly to the vacuum, radiates a few gluons, and hadronizes. The other quark goes through the dense plasma that is formed in the collision, suffers energy loss due to medium-induced gluon radiation, and finally fragments outside into a quenched jet. The plasma is characterized by a transport coefficient $\hat{q}$, temperature T, and gluon density $\mathrm{d}N^{\mathrm{g}}/\mathrm{d}y$.

collisions N_{coll} which happen in a Au–Au collision, then the scaled pp spectra was divided by the Au–Au spectra. This ratio is the nuclear modification factor R_{AA} and any deviation from unity indicates a mechanism that is a feature only of central heavy-ion collisions. The R_{AA} for a hadron is defined as

$$R^{h}_{\mathrm{AA}}(p_{\mathrm{T}},\eta) = \frac{1}{\langle N_{\mathrm{coll}}\rangle}\left(\frac{d^2\sigma_{\mathrm{AA}\to h}}{dp_{\mathrm{T}}d\eta}\right)\left(\frac{d^2\sigma_{\mathrm{pp}\to h}}{dp_{\mathrm{T}}d\eta}\right)^{-1}, \tag{2.17}$$

where σ denotes the cross section of the particular process. A summary of important measurements of R_{AA} for three different center-of-mass energies is shown in Fig. 2.8 [53]. The value of R_{AA} is significantly reduced for both charged and neutral hadrons over a wide p_{T} range. At RHIC energies a strong suppression is clearly visible in the range relevant for pQCD ($p_{\mathrm{T}} \gtrsim 4\ \mathrm{GeV}/c$).

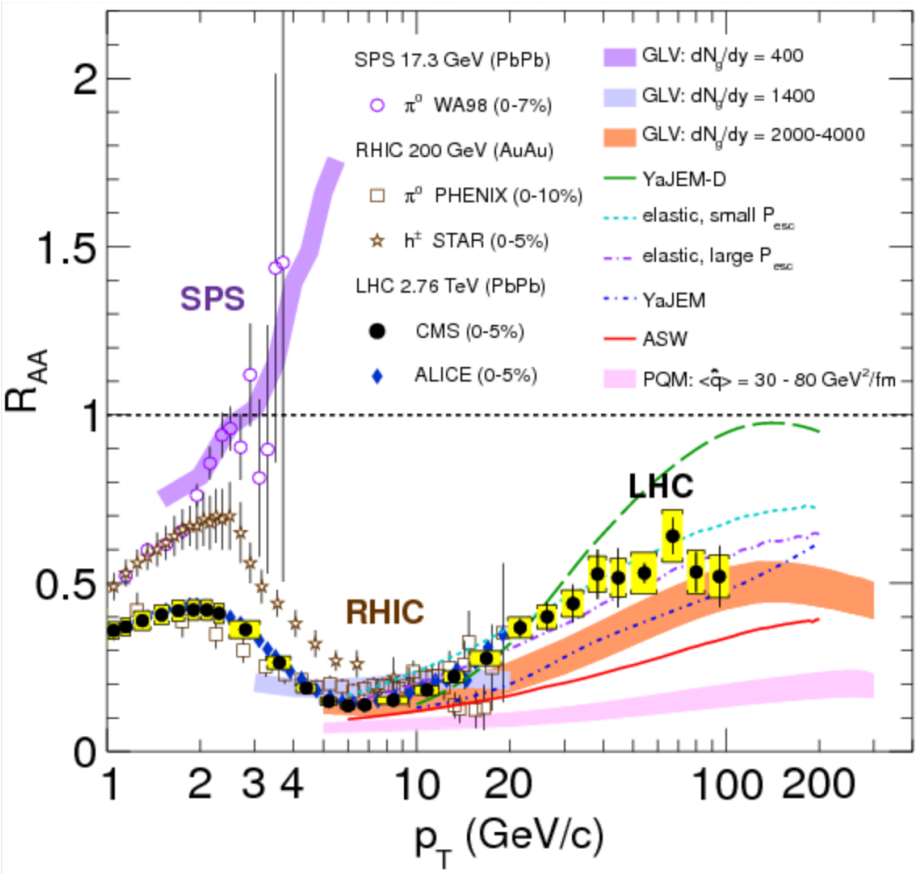

Figure 2.8: Measurements of the nuclear modification factor R_{AA} in central heavy-ion collisions at three different center-of-mass energies, for neutral pions (π^0), charged hadrons ($\mathrm{h}^{\pm}$), and charged particles and compared to several theoretical predictions.

The STAR experiment also measured the angular correlation of the leading (most energetic) hadrons of jets at mid-rapidity in $\sqrt{s_{\mathrm{NN}}} = 200$ GeV Au–Au collisions, shown in Fig. 2.9 (taken from [48]). The angular correlations are compared across different collisions systems, and a clear absence of the back-to-back correlation peak is missing, confirming the predicted pehnomenon of jet quenching. Jet suppression in the high-momentum region of heavy-ion collisions was also reported by the STAR Collaboration [54].

It is an important question whether the jet quenching should be attributed to initial or final state effects. As a supplementary measurement therefore, the PHENIX Collaboration measured the particle production in d–Au at the same $\sqrt{s_{\mathrm{NN}}} = 200$ GeV energy. The same binary scaling was observed as in peripheral Au–Au collisions, but there was no suppression visible [55]. This measurement provided strong evidence that the suppression is in the final state, i.e. a new state of matter is created that is only present in high-energy

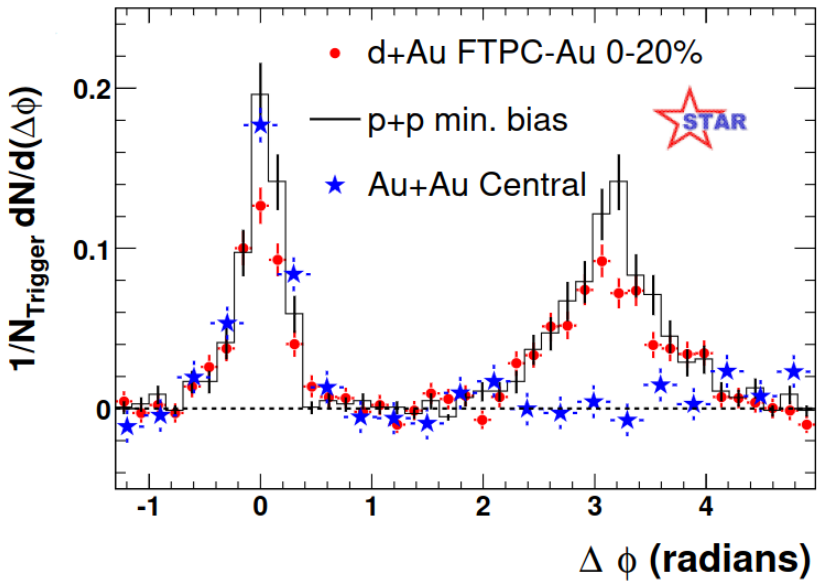

Figure 2.9: High-p_{T} dihadron azimuthal correlation measurement at $\sqrt{s_{\mathrm{NN}}} = 200$ GeV for pp, central d–Au and central Au–Au collisions (background subtracted) from the STAR Collaboration.

collisions of heavy ions, where the system has a sufficiently large size.

As we discussed, the quark-gluon plasma created in heavy-ion collisions is a strongly interacting medium. Photons, however, do not have color charge therefore the QGP is a nearly transparent medium for them. Especially useful are direct photons, i.e. which are produced from sources other than hadronic decays, as their production is sensitive to the early stages of the reaction. The direct photons carry information on the hard process, therefore it can be used as a control measurement for hadron suppression. Such a direct photon measurement was carried out by the PHENIX Collaboration in Au–Au collisions at $\sqrt{s_{\mathrm{NN}}} = 200$ GeV and showed that the nuclear modification factor for photons is consistent with unity in the mid-p_{T} ranges, confirming that the hadron suppression is indeed caused by strongly interacting matter [56]. Here I also mention that measuring the thermal photons can provide information on the temperature of the QGP (see Sec. 2.5.3).

2.5.2 A Perfect Fluid of Quarks

While the observed jet quenching phenomenon at RHIC proved the existence of a new phase of matter, the RHIC experiments uniformly found substantial azimuthal anisotropy in measurements of several particle types, that could only be explained by assuming a strongly coupled QGP that behaves as a liquid rather than like a gas. This observation was

completely against our naive expectations from asymptotic freedom. It became clear that the most important properties of this matter can be only accessed through soft processes: these processes have relatively small momentum transfer, therefore pQCD calculations cannot be used for their analysis.

The multiplicity distribution of the bulk of the particles with lower momenta exhibits a dependency as

$$\frac{dN}{d\phi} \approx 1 + 2\nu_2(p_{\mathrm{T}})\, cos(2\phi) \ , \tag{2.18}$$

where ϕ is the angle with the reaction plane, defined by the (parallel) trajectories of the colliding nuclei, and ν_2 is the so-called elliptic flow parameter. This is a direct hydrodynamical consequence of the elliptic shape of the nucleus overlapping during the initial state of the collision. This pressure-anisotropy of the initial state then results in the anisotropy of the momentum distribution of the final state (see Fig. 2.10, taken from [57]).

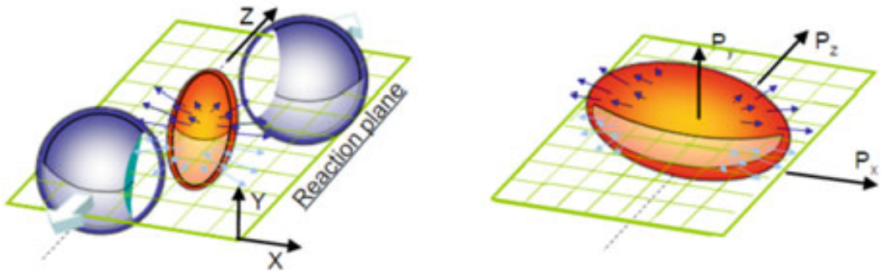

Figure 2.10: Illustration of the elliptic flow. The pressure anisotropy of the initial state shows up as momentum distribution anisotropy in the final state. The expanding ellipsoidal region of hot quark matter has a principal axis perpendicular to the reaction plane.

Later measurements also found that the viscosity over entropy ratio of this strongly-interacting fluid is at least an order of magnitude less than the viscosity of super-fluid helium, which makes it the most perfect fluid that we currently know. Furthermore, a new scaling property was also found when the ν_2 elliptic flow parameter was measured versus the transverse kinetic energy. After rescaling with the number of constituent quarks, all hadrons lay on the same curve, which shows a number-of-constituent-quark (NCQ) scaling behavior of the elliptic flow. This clearly proved that the degrees of freedom are quarks, and the viscosity was measured to be at the quantum limit, hence the QGP is a perfect fluid of deconfined quarks (see Fig. 2.11, taken from [58]).

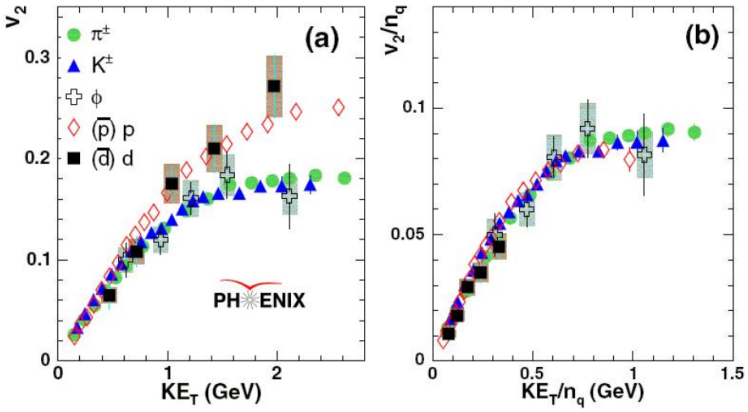

Figure 2.11: Scaling of hadrons in the PHENIX experiment at RHIC. (a) Elliptic flow parameter ν_2 vs. transverse kinetic energy in $\sqrt{s_{\mathrm{NN}}} = 200$ GeV Au–Au collisions. (b) The same data but scaled by the number of constituent quarks n_{q}. The different types of hadrons lay on the same curve.

2.5.3 The Temperature of the QGP

As it was mentioned in the previous section, the RHIC experiments found a new, extremely hot and dense state of matter, which behaves more like a liquid than a gaseous phase. The observation of thermal photons in principle allows the access of the thermal properties, e.g. the determination of the initial temperature of this matter, but the precision of low-p_{T} direct photon measurements is limited due to the large background from the hadronic decay photons. The leptons on the other hand are excellent tools for studying the heavy-ion collisions at ultra-relativistic energies. They do not carry color charge and therefore are unaffected by the strong interaction, which allows them to escape the dense medium without any final state interaction. For example, measuring the dilepton spectra can probe the complete time evolution and dynamics of heavy-ion collisions. Moreover, any source of high-energy photons emit virtual photons which convert into low-mass $\mathrm{e}^+\mathrm{e}^-$ pairs. The virtual photon production then can be related to the direct photon production itself, thus the direct photon yield can be reconstructed from dileption measurements [59].

The PHENIX experiment measured the thermal photon spectrum using virtual photons decaying to electron–positron pairs in central Au–Au collisions. From the slope of the spectrum, the temperature of the QGP could be inferred. The inverse slope parameter is $T = 221 \pm 19(\mathrm{stat}) \pm 19(\mathrm{syst})$ MeV, and based on this the initial temperature could be estimated using hydro-dynamical models to fall between $300 - 600$ MeV, corresponding to 2-4 trillion Kelvins. Since this temperature is substantially higher than the Hagedorn temperature ($T_{\mathrm{H}} \approx 160$ MeV) [60] above which matter cannot exist in a hadronic phase anymore, it is a direct proof to the existence of the strongly interacting QGP.

2.5.4 Collectivity in Small Systems

Collective behavior in ultra-relativistic collisions of heavy ions has long been interpreted as a signature of the strongly coupled quark-gluon plasma. Surprisingly, small collision systems, such as pp and p–Pb collisions, at high center-of-mass energies show similar collective features to those observed in events of heavy-ion collisions with comparable multiplicities. These observations, including long-range near-side correlations and v_n ("flow") coefficients [61, 62] (see Fig. 2.12, taken from [62]), may also question the signatures previously considered as definite signs of the QGP. An intensively researched question in high-energy physics is therefore whether small droplets of quark-gluon plasma can form in collisions of small systems. Whether the collective behavior may be attributed to the presence of a deconfined state in small systems is an open question. However, possible medium-like effects in pp collisions may also question the widely exploited assumption that pp collisions are safe to use as a reference for heavy-ion systems. Recent studies showed that flow patterns may emerge from features different than hydrodynamics. For instance, radial flow in pp collisions may be explained by pure QCD mechanisms such as multiple-parton interactions (MPI) [34]. Alternative explanations also exist for the observed collectivity in small colliding systems, such as calculations based on Color Glass Condensate initial states with Lund fragmentation [63] or non-Abelian Bremsstrahlung [64].

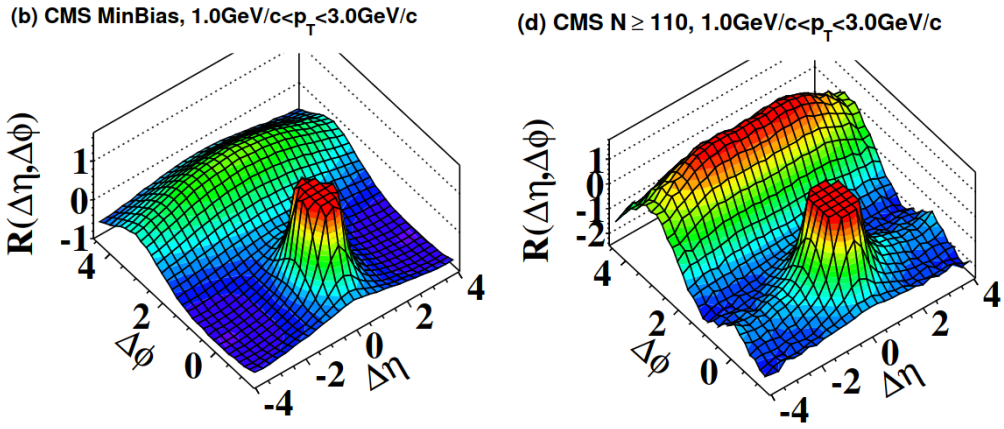

Figure 2.12: Two-dimensional two-particle correlation functions are shown for 7 TeV in pp collisions, for minimum bias events (*left*) and high-multiplicity events (*right*). The sharp near-side peak from jet correlations was cut off in order to better illustrate the structure outside that region.

Recent analyses of pp and p–Pb collisions also show a universal enhancement of heavy-flavor particles, that is usually attributed to MPI and higher gluon radiation associated with short distance production processes [65]. However, we lack the qualitative

understanding of these effects. While we cannot expect to observe direct modification of particle yields by any medium created in collisions of small systems (because of the small volume of such a medium), phenomena that act in the soft-hard transitional regime should in principle pose an effect on hard processes as well. A modification in the shapes of developing jets can in principle be accessible by existing experiments.

In recent years a consensus starts to emerge that the observed collective phenomena could be explained by semi-soft vacuum-QCD effects, such as multiparton-interactions [66] with color-reconnection [34] or minijets (semi-hard partons produced by incoming partons or bremsstrahlung) [67] production.

2.6 Heavy Flavor

The Standard Model has 6 quark flavors (see Sec. 2.1), of which the three heaviest, the charm, beauty and top quarks, are referred to as heavy-flavor quarks. Hadrons which contain heavy quarks are usually called heavy flavor particles. The top quark has the highest mass not only among the quarks, but also among the currently known elementary particles, and because it decays before it could hadronize, it is usually not included in heavy-flavor analyses [68].

2.6.1 Production of Heavy Flavor

There are two mechanisms that generate quark masses with different degrees of importance, depending on the energy scale: bare quark masses (also called naked or current quark masses) are generated by the electroweak symmetry breaking mechanism (Higgs mass), while the constituent quark masses in QCD (QCD mass) are the results of the spontaneous breaking of chiral symmetry. The light quark (u, d, s) masses are strongly affected by the QCD interaction, while the heavy quark (c, b, t) masses are mainly determined by the Higgs mechanism. This interplay in the generation of quark masses is shown in Fig. 2.13 (taken from [69]). Heavy quarks are mostly created in the early stages of the collision, in processes which are accessible with perturbative quantum chromodynamics (pQCD) calculations: a heavy quark can be created from a pair of gluons or light quarks by flavor creation (FLC), a gluon splitting into the quark-antiquark pair (GSP), or through flavor excitation (FLX) [70, 71]. They may also interact in semi-hard processes and participate in the formation of the underlying event [72]. The heavy-quark production is also sensitive to the parton distribution function. Heavy quarks are an ideal probe to study the properties of the hot and dense medium created in high-energy nuclear collisions, because unlike the light quark masses, heavy quark masses are much higher than the initial excitation of the system and are therefore not modified by the surrounding QCD medium.

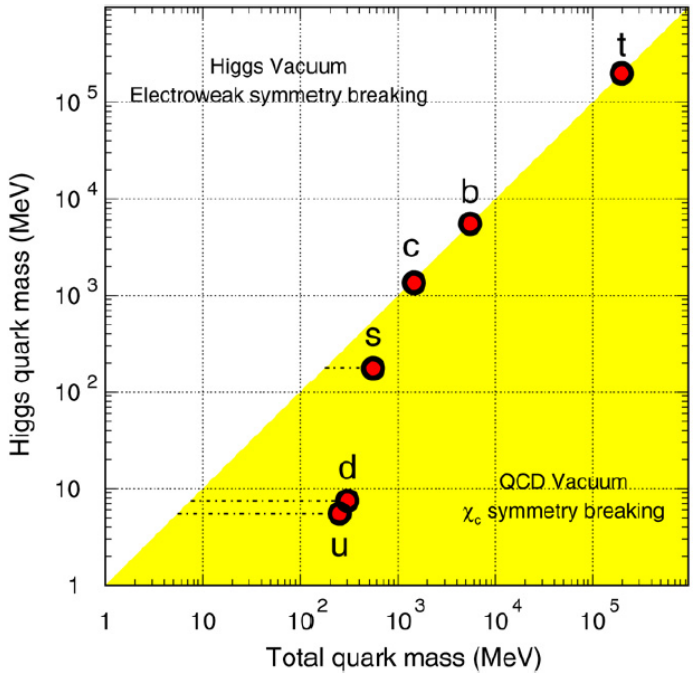

Figure 2.13: Quark masses in the QCD vacuum and the Higgs vacuum. A large fraction of the light quark masses is due to the chiral symmetry breaking of the QCD vacuum.

Studying heavy-flavor quark production serves as a useful tool across multiple collision systems [73–77]. In pp collisions we can test perturbative QCD (pQCD) predictions through the measurement of cross sections. We can also test the parton shower through the dead-cone effect [35, 36] (mass ordering) and Casimir color factor (color charge effect). The measurement of heavy-flavor production also provides means to test fragmentation models. Jet structure may reveal flavor dependence stemming from mass and color charge effects, while the comparison of the charm mesonic and baryonic sector can reveal quark coalescence processes and the effect of color junctions. Another crucial role of heavy/flavor measurements is to provide baseline for similar measurements in p–Pb and Pb–Pb collisions. Collisions of p–Pb systems are useful to isolate the initial state and cold nuclear matter (CNM) effects and can serve as tools for studies of possible collective effects. In Pb–Pb collisions, heavy flavor is produced before the QGP can form ($m \gg T_{\mathrm{QGP}}$), and experiences the full evolution of the medium. They experience energy loss via gluon radiation and collisions, and thus can be used to probe the dynamical and thermal properties of the medium.

The production of heavy-flavor hadrons in high-energy collisions is usually described by the factorization approach, in which the production cross section of the heavy-flavor particles is expressed as a convolution of three independent terms:

$$d\sigma^{\rm hard}_{AB\to C} = \Sigma_{a,b} f_{a/A}(x_a, Q^2) \otimes f_{b/B}(x_b, Q^2) \otimes d\sigma^{\rm hard}_{ab\to c} \otimes D_{c\to C}(z, Q^2)\,, \tag{2.19}$$

where the first two terms are the parton distribution functions (PDF) of the incoming hadrons, the third term is the parton-parton hard scattering cross section, and the last term is the fragmentation function (FF). The hard parton cross sections can be derived using perturbative QCD calculations, as the large mass of the heavy quark establishes the hard scale, allowing production to be computed down to low transverse momentum. However, the PDFs and FFs need to be obtained through measurements. The pQCD factorization is illustrated for dijet production in Fig. 2.14 (taken from [78]).

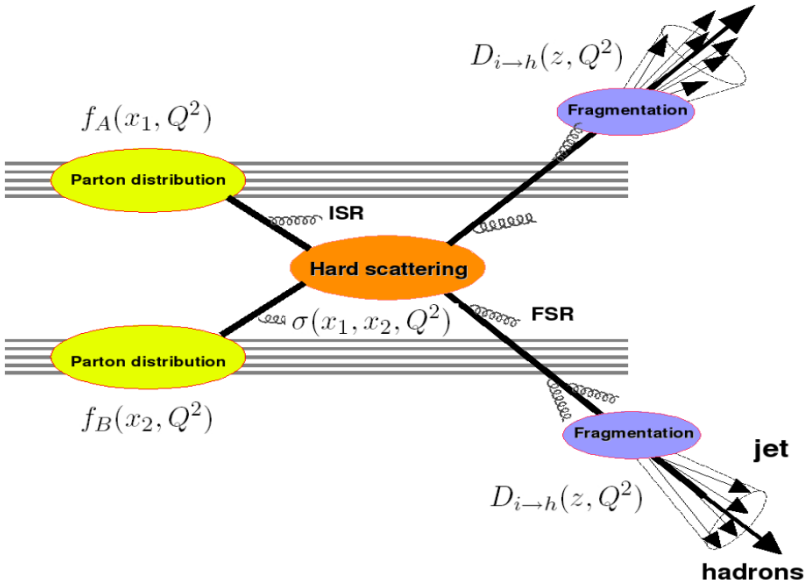

Figure 2.14: Illustration of the pQCD factorization in dijet production in hadronic collisions: $f_{a/A}(x_1, Q^2)$ are the PDFs, $D_{i\to h}(z, Q^2)$ the FFs, and ISR (FSR) represents the initial (final)-state radiation.

The production cross sections of several open heavy-flavor hadrons and of their decay leptons in pp collisions were measured at both mid- and forward-rapidity at the LHC [79–85]. The production cross sections are described by pQCD calculations [86–88] with large theoretical uncertainties.

2.6.2 Heavy-flavor Jets and Correlations

The pattern of the parton shower is influenced by the mass of the emitting parton due to a phenomenon known as the dead-cone effect. The dead-cone effect is a fundamental aspect of all gauge field theories, that causes radiation with energy E off an emitter of mass m to be suppressed at angular scales smaller than m/E relative to the emitter's direction [35]. The dead-cone effect has been observed by the ALICE experiment for the first time [36]. Jet and two-particle angular correlation measurements allow for the flavor-dependent characterization of parton shower and fragmentation. Recent jet substructure measurements allow for the detailed understanding of jet development.

The reconstruction of jets containing heavy-flavor hadrons provides a more direct access to the primary heavy-flavor parton kinematics compared to the inclusive measurements of heavy-flavor hadrons. Heavy-flavor jet measurements allow studying the production and fragmentation effects separately. The ALICE detector has excellent tracking capabilities for low-p_{T} charged particles, which makes it possible to measure heavy-flavor jets down to low transverse momenta. Jets containing a D^0 meson have been measured by the ALICE experiment in proton–proton collisions at several energies [89, 90]. It was found that pQCD-based models generally well describe the data within uncertainties. More recent groomed substructure measurements of D^0 jets show significant differences between charm jets and jets initiated by light quarks and gluons [91]. The first measurement of inclusive charged-particle b-jet p_{T}-differential cross section and the b-jet fraction was reported by the ALICE experiment [92]. Since heavy-flavor jet tagging is challenging at low-momenta, the correlation measurements can be used to provide insight into heavy-flavor jet properties at low transverse momentum.

The typical structure of a two-particle angular correlation distribution of high p_{T} trigger particles with associated charged particles features a "near-side" (NS) peak at coordinates $(\Delta\varphi, \Delta\eta) = (0, 0)$, where $\Delta\varphi$ is the difference in azimuth angle between the trigger and associated particles, and $\Delta\eta$ is their pseudorapidity difference. There is also an "away-side" (AS) peak at $\Delta\varphi = \pi$, which extends over a wide pseudorapidity range. The NS peak is mainly created by particles emerging from the fragmentation of the parton that produced the trigger particle. The AS peak is related to the fragmentation of the other parton produced in the hard scattering. The correlation peaks lie on top of an approximately flat continuum extending over the full $(\Delta\varphi, \Delta\eta)$ range [93]. By varying the p_{T} interval of the trigger and associated particles, the correlation measurements allow the detailed study of jet fragmentation, such as the jet angular profile and the momentum distribution of the particles produced in the fragmentation of the hard parton.

The azimuthal correlation distributions of prompt D mesons and charged particles were measured by the ALICE Collaboration in pp collisions at $\sqrt{s} = 5.02$ TeV, $\sqrt{s} = 7$ TeV, and $\sqrt{s} = 13$ TeV [93–95]. Measuring the correlation distribution between heavy-flavor decay electrons and charged particles provides a substantially larger sample of correlation

pairs, compared to measuring the azimuthal correlations of D mesons and charged particles [93, 94]. The larger sample allows to significantly extend the $p_{\mathrm{T}}^{\mathrm{assoc}}$ range of associated particles to provide a more complete picture of the heavy quark fragmentation. In addition, electrons originating from beauty-hadron decays (b $\rightarrow$ (c $\rightarrow$) e) dominate the heavy-flavor hadron decay electron spectrum ($> 50\%$) at high $p_{\mathrm{T}}^{\mathrm{e}}$ (> 5 GeV/c) [96]. Hence, probing large enough trigger electron transverse momenta enables the study of the correlation function of particles originating from beauty-hadron decays, and provides information on the different correlation structures for charm and beauty quarks. This additional information can be used to further constrain MC simulations. The mentioned advantages have a cost, however, because an additional smearing is introduced in the correlation function, due to the non-zero angle between the direction of the trigger electron and the parent heavy-flavor hadron before its decay. The momentum of the electron can also be further away from the quark momentum compared to that of the parent hadron, due to the kinematics of its decay.

In proton–nucleus (p–A) collisions, several CNM effects can influence the production, fragmentation, and hadronization of heavy quarks. In the initial state, the parton distribution functions (PDFs) are modified in bound nucleons as compared to free nucleons [97–99]. Color-Glass Condensate calculations [100–102] predicts momentum correlations in the initial state, that would impact the angular correlations of the produced heavy-quark pairs. The CGC predicts momentum correlations in the initial state, that would impact the angular correlations of the produced heavy-quark pairs. Partons may undergo multiple scatterings in the nucleus during the initial state and interact in the high-density environment in the final state, especially in high-multiplicity collisions [103, 104]. These effects can be studied by measuring changes in the angular shape or associated-particle peak yields of the angular correlation distributions of heavy-flavor particles with charged hadrons [93, 94]. Measurements of azimuthal correlations of prompt Dmesons and charged hadrons in p–Pb collisions by the ALICE collaboration[93, 94] showed that the near- and away-side peaks are consistent with those in pp collisions in the same kinematic region. Using heavy-flavor decay electrons as trigger particles instead of prompt D mesons allows studying cold-nuclear-matter effects over a wider associated particle $p_{\mathrm{T}}^{\mathrm{assoc}}$ range and investigating their impact on beauty-quark fragmentation and hadronization. Two-particle angular correlations have been extensively used to understand radiation patterns and to probe the medium response to the high-p_{T} parton. A recent measurement of angular correlations between D mesons and charged particles in Au–Au collisions by the STAR Collaboration [105], shows a significant modification of the near-side peak width and associated yield, which increases from peripheral to central collisions. Correlations between electrons from heavy-flavor hadron decays and charged particles measured by PHENIX reveal modifications in the away-side peak yield and width in Au–Au compared to pp collisions [106]. For future heavy-flavor hadron correlation studies in heavy-ion collisions at the LHC, similar measurements in pp and p–Pb collisions are essential to serve as references [107].

In Chapter 6, I present my work on an analysis of the azimuthal correlations of heavy-flavor decay electrons with charged particles in pp and p–Pb collisions with the ALICE experiment [108].

2.6.3 Fragmentation of Charm Baryons

Recent measurements of charm-baryon production at midrapidity in pp collisions are not reproduced by pQCD calculations and event generators adopting a fragmentation model tuned on e^+e^- data [109–114]. In Fig. 2.15, a recent measurement of the ALICE Collaboration shows the cross-section ratios of Λ_c^+/D^0 and $\Sigma_c^{0,+,++}/D^0$, comparing them to model predictions. The cross-section ratio values in pp collisions at $\sqrt{s} = 13$ TeV are compatible with values measured at $\sqrt{s} = 5.02$ TeV within uncertainties. The Λ_c^+/D^0 ratio is observed to decrease with increasing p_T and is, however, significantly larger than the ≈ 0.12 values observed in e^+e^- and ep collisions at several collision energies [115–119], therefore a remarkable difference between the collision systems of pp and e^+e^- is present in the experimental data.

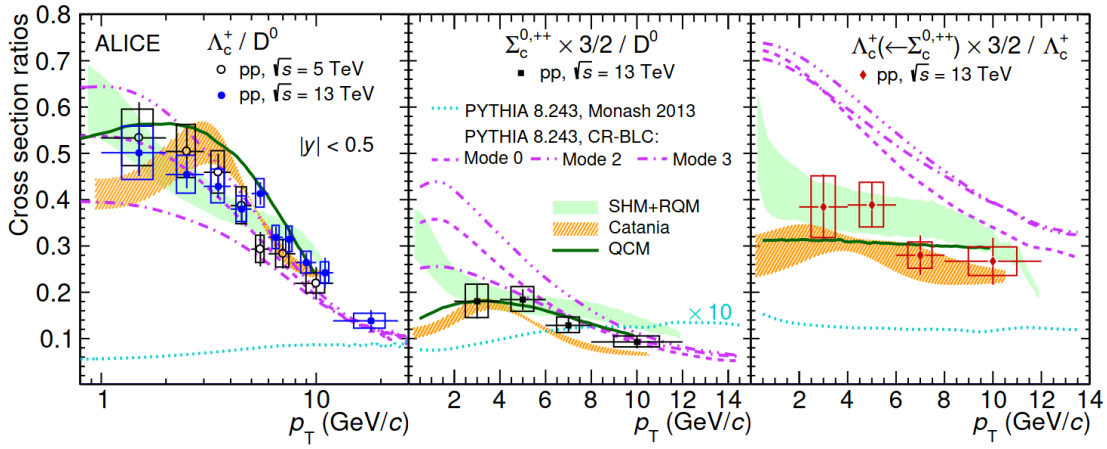

Figure 2.15: Prompt-charm-hadron cross-section ratios in pp at $\sqrt{s} = 13$ TeV: Λ_c^+/D^0 (left), $\Sigma_c^{0,+,++}/D^0$ (middle) and $\Lambda_c^+ \leftarrow \Sigma_c^{0,+,++}/\Lambda_c^+$ (right), compared with model predictions and also data frompp at $\sqrt{s} = 5.02$ TeV [111].

These conclusions are supported by the fact that while simulations performed with the default version of PYTHIA 6.2 reasonably reproduce Belle data [120], the default version of PYTHIA 8.243 (Monash 2013 tune) severely under-predicts ALICE data, despite the very similar modelling of charm fragmentation in the two simulations.

As Fig. 2.15 shows, a better description of these measurements can be obtained by models including hadronization mechanisms such as quark coalescence [121], additional color reconnections among parton fragments [122], or by including enhanced feed-down

from higher-mass charm-baryon states within a statistical hadronization approach [123], where the higher-mass excited charm-baryon states are predicted by the Relativistic Quark Model [124], but not yet measured.

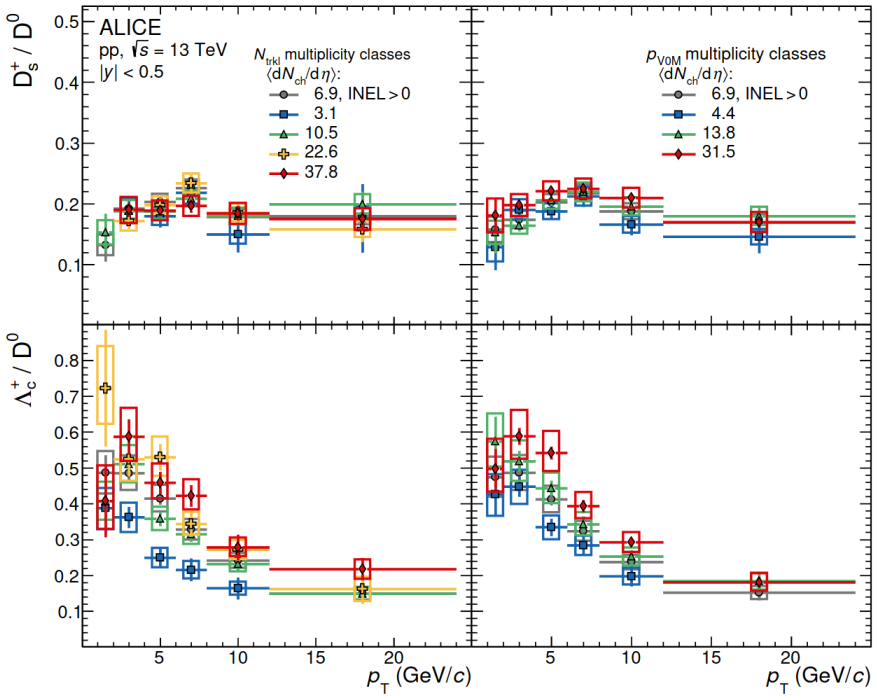

Figure 2.16: The D_{S}^{+}/D^{0} (top) and $\Lambda_{\mathrm{c}}^{+}/\mathrm{D}^{0}$ (bottom) ratios measured in pp collisions at $\sqrt{s}$ $= 13$ TeV for different multiplicity classes at mid- (left) and forward (right) rapidity [112].

The ALICE Collaboration recently conducted a multiplicity-differential study of the $\Lambda_{\mathrm{c}}^{+}/\mathrm{D}^{0}$ and $D_{\mathrm{S}}^{+}/\mathrm{D}^{0}$ cross-section ratios as shown in Fig. 2.16. The results confirm my previous studies of the multiplicity-dependece of the charm-baryon ratios [125] and it opens up an interesting research area for the multiplicity-dependent study of excited charm states. In Chapter 7, I present my work aimed at the understanding of the underlying mechanisms of charm baryon fragmentation.

2.7 Event Activity and the Underlying Event

Hadron collisions are highly complex and it is far from obvious how to study specific particle processes while excluding the background processes that occur during a collision (see illustration in Fig. 2.17, taken from [126]). The process of interest is typically a single parton–parton scattering process at very high transverse momentum (p_{T}) transfer, while the background consists of the remnants of the colliding hadrons that did not participate in

the hard scattering, including the products of additional soft, multiple-parton interactions. Understanding this background contribution is crucial for all physics measurements in the LHC, let that be the study of quark-gluon plasma, Higgs physics or beyond the Standard Model searches, or to increase the precision of our physics measurements, especially at high p_{T}.

The properties of the underlying event currently cannot be derived from first principles in QCD, but there are experimental techniques to measure the underlying event contribution by using event classifiers. One example is the average charged-particle multiplicity measured in the transverse plane, i.e. in the region of phase space orthogonal to the high-p_{T} process. In this section I define the underlying event and discuss the observables that provide a quantitative description.

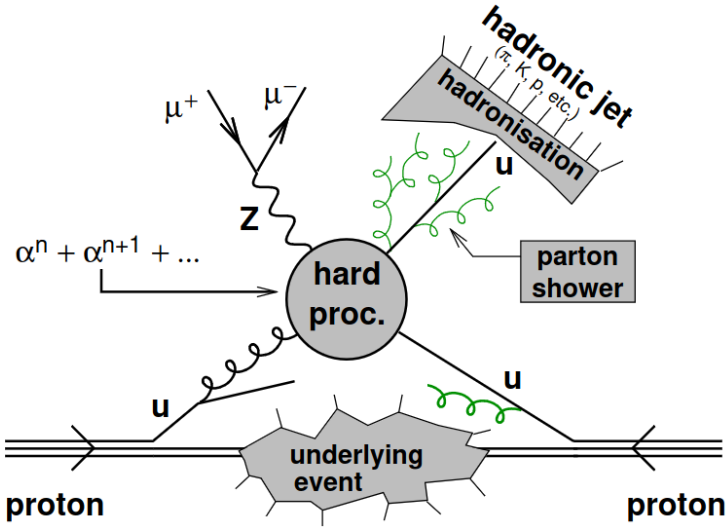

Figure 2.17: Illustration of proton–proton collision with UE included, that leads to a final state consisting of a Z boson and a hard jet.

2.7.1 Underlying Event

The underlying event (UE) was first defined by the CDF collaboration in 2002 [127]. The underlying event consists of all the secondary processes, as well as beam remnants. The UE is usually treated as independent from the leading hard process [128], and it is usually measured in the region far away from the jets. Understanding the UE has been essential in order to understand the background in measurements of hard probes. More recent definitions of the UE also include events with multiple jets [129]. In recent years, differential measurements came into focus that characterize the UE not only as a background compo-

nent, but in terms of its relation to the leading process. This way processes that connect the soft and hard regimes of a collision, such as multi-parton interactions, can be explored.

2.7.2 Classification of Events

Insight into the connection of hard processes and the underlying event can be gained by the differential exploration of events with respect to event-shape variables.

Event Multiplicity Variables

Traditionally the final-state (event) multiplicity $N_{\rm ch}$ is used to categorize events by activity, defined as the number of all charged final state particles in the event in the mid-rapidity acceptance. Since $N_{\rm ch}$ is defined in the same pseudorapidity range as where the charmed-hadron yields are computed from, effects observed via $N_{\rm ch}$ may be influenced by autocorrelation. Therefore we also used the forward multiplicity $N_{\rm fw}$, which we defined as the number of charged particles within the acceptance $2 < |\eta| < 5$.

We characterize the event with event-activity classifiers that are selectively sensitive to the activity in the underlying-event or to that caused by high-momentum jets in the leading process. One such observable is the transverse event-activity classifier $R_{\rm T} \equiv N_{\rm ch}^{\rm trans}/\langle N_{\rm ch}^{\rm trans}\rangle$ [130], where $N_{\rm ch}^{\rm trans}$ is the charged-hadron multiplicity in the transverse region defined the following way. The charged final-state hadron with the highest transverse momentum within the acceptance (the trigger hadron) is selected and the event is accepted only if the trigger hadron has a transverse momentum $p_{\rm T}^{\rm trig} > 5$ GeV/c. $N_{\rm ch}^{\rm trans}$ is then the number of charged final-state hadrons in the transverse side, defined with the azimuth angle relative to the trigger hadron as $\frac{\pi}{3} < |\Delta\phi| < \frac{2\pi}{3}$ within $|\eta| < 1$. Since the trigger hadron most likely comes from a high-momentum jet initiated by the leading process, and the recoil jet is expected to show up at the opposite side, $N_{\rm ch}^{\rm transverse}$ is dominated by hadrons from the underlying event [131]. In models such as PYTHIA that describe events in terms of MPI, $R_{\rm T}$ is strongly correlated with the number of MPIs in an event [130]. Analogously, we defined the near-side cone activity $R_{\rm NC} \equiv N_{\rm ch}^{\text{near-side cone}}/\langle N_{\rm ch}^{\text{near-side cone}}\rangle$ in a narrow cone around the trigger particle, $\sqrt{\Delta\phi^2 + \Delta\eta^2} < 0.5$. As this region is dominated by the fragments of the jet containing the trigger particle, $R_{\rm NC}$ will be primarily determined by the multiplicity of the jet initiated by the leading hard process. In Fig. 2.18 (taken from [131]) we show an illustration of the toward, away, and transverse regions in the azimuthal plane, with respect to the leading particle direction.

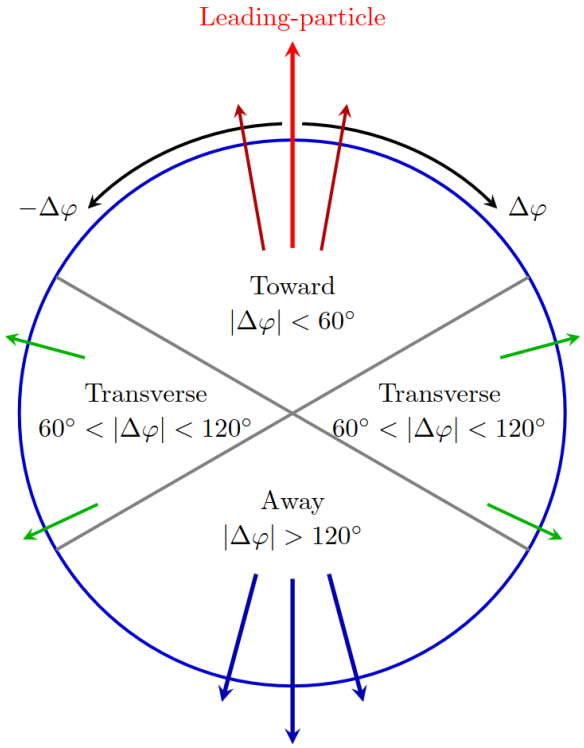

Figure 2.18: Schematic picture of the toward, away, and transverse regions in the azimuthal plane, with respect to the leading particle direction.

Transverse Spherocity

Event classification based on R_{T} and R_{NC} requires a high-transverse-momentum trigger, and thus it restricts the analysis to events containing a hard process. This introduces a bias into the sample and makes it more difficult to accumulate sufficient statistics in an experimental environment. A similar event selection may however be achieved based on the geometrical shape of the event. One such variable is transverse spherocity [132], defined as

$$S_0 \equiv \frac{\pi}{4} \min_{\hat{\mathbf{n}}} \left(\frac{\sum_i |\mathbf{p}_{\mathrm{T},i} \times \hat{\mathbf{n}}|}{\sum_i p_{\mathrm{T},i}} \right) , \tag{2.20}$$

where i runs over all the particles in the acceptance and $\hat{\mathbf{n}}$ is any unit vector in the azimuth plane. The transverse spherocity characterizes the events by jettiness in the azimuthal plane, and it is defined between 0 and 1 by construction: for isotropic events S_0 approaches unity, and for events determined by collimated clusters of final-state particles, S_0 is close to zero. In Fig. 2.19 (taken from [133]), I show an illustration of the transverse spherocity S_0 describing jetty and isotropic events in the transverse plane.

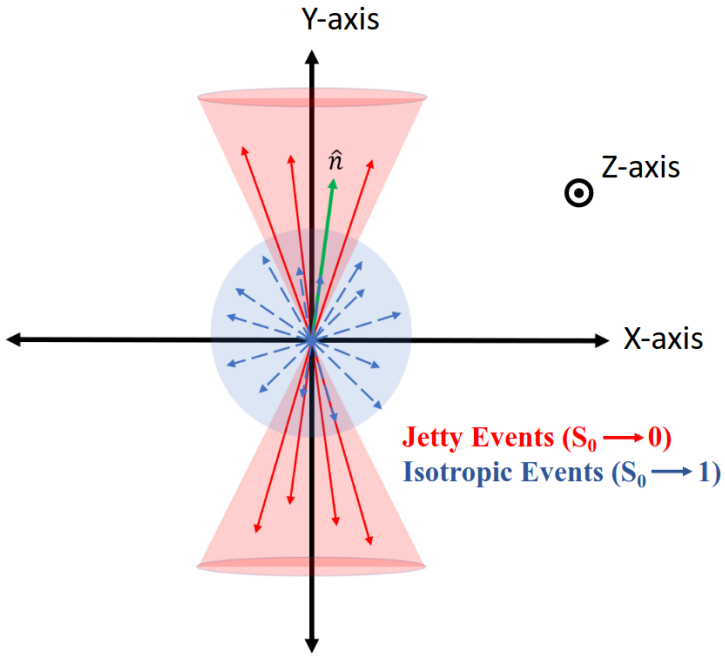

Figure 2.19: Schematic picture of the spherocity S_0 showing jetty and isotropic events in the transverse plane.

Flattenicity

The transverse spherocity S_0 concentrates on the central η range and therefore it is not sensitive to the part of event which expands toward higher η. A new event classifier has been proposed recently to overcome this limitation, the so-called flattenicity [134], which similarly to the spherocity, provides information about the "jettiness" of an event and does not require a trigger. This event quantifier is capable of selecting hedgehog-like events without a characteristic jetty structure in high-multiplicity pp collisions. Not requiring a trigger particle is also advantageous to gain a better statistics, as we do not have to exclude so many events from the analysis.

The flattenicity ρ is calculated by splitting up the η-ϕ plane of an event into equal-sized rectangles, and the average transverse momenta of the particles inside the rectangles is taken. Then the flattenicity is simply the relative standard deviation of the average momentum inside the rectangles according to the formula

$$\rho = \frac{\sigma_{p_T^{\text{cell}}}}{\langle p_T^{\text{cell}} \rangle}, \tag{2.21}$$

where $\sigma_{p_T^{\text{cell}}}$ is the deviation, while the $\langle p_T^{\text{cell}} \rangle$ is the average of the transverse momenta. Flattenicity has been successfully utilized to classify events corresponding to the underlying physics process [135–137]. As an illustration, Fig. 2.20 (taken from [134]) shows

two events with significantly different flattenicity values. As clearly visible, the event with lower ρ value is much more isotropic, while the other one with higher ρ is overall more "jetty".

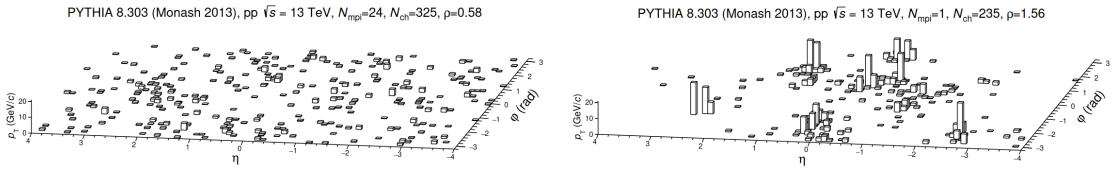

Figure 2.20: Example of two PYTHIA 8.244 events with Monash tune, one with low flattenicity (left) and the other with high flattenicity (right). The event with lower flattenicity looks significantly more isotropic.

2.7.3 KNO Scaling

Final-state multiplicities in small colliding systems are known to follow a negative binomial distribution (NBD) regardless of the exact type of colliding particles over several orders of magnitude of energy ranges [138–140]. Koba, Nielsen and Olesen demonstrated that the event multiplicity distributions can be all collapsed onto a universal scaling curve, this is the Koba–Nielsen–Olesen (KNO) scaling [141, 142]. It has been observed in e^+e^- collisions that the multiplicity distributions at different collision energies can indeed be collapsed into a single distribution, however, found to be violated at higher energies and in more complex, hadronic collision systems [143, 144]. In Fig. 2.21 [145, 146], the KNO scaling is shown in experimental data of $\mathrm{e^+e^-}$ and pp collisions. Neither the origin of the scaling, nor the reason for its breakdown is yet completely understood, although many explanations have been proposed in the past decades [147–151].

At higher center-of-mass energies, where the average final-state multiplicity is higher, semi-hard vacuum-QCD effects such as multiple-parton interactions (MPI) play a significant role. Multiple works proposed that the scaling violation may be caused by these effects. A scenario based on the Lund string model [145, 152] proposes that the overlapping color strings break the scaling. Another work argues that underlying-event activity linked to MPI with color reconnection (CR) is responsible for the violation of the scaling [153]. The are earlier works also proposed that the KNO scaling may be a property of the jet itself [154].

At lower collision energies, since events with jet events have very little background, collision energy can be directly linked to average $p_{\mathrm{T}}^{\mathrm{jet}}$. This suggests that $p_{\mathrm{T}}^{\mathrm{jet}}$ may be a more fundamental scaling variable, and the violation may be explained by the breaking-down of the connection of $\sqrt{s}$ to the average $p_{\mathrm{T}}^{\mathrm{jet}}$ toward higher energies as well as more complex colliding systems.

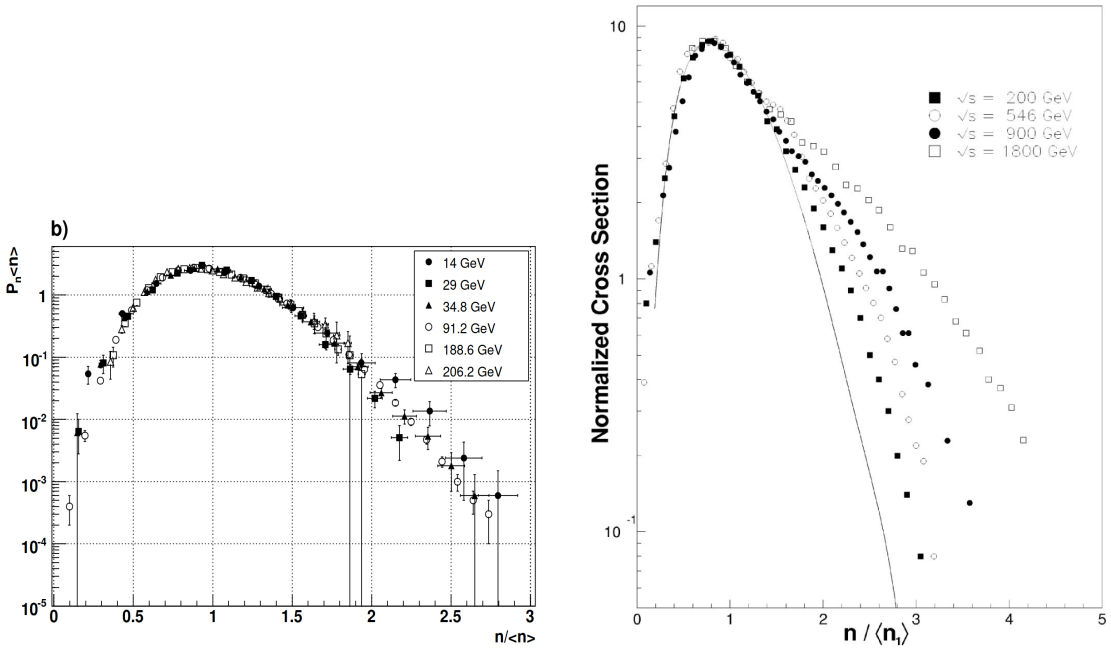

Figure 2.21: Example for the KNO scaling in experimental data where the multiplicity distributions have been normalized with the average both for electron-positron collisions (*left*) and proton–proton collisions (*right*).

2.8 Experimental Facilities

To create QCD matter for the investigation of its properties, we use powerful accelerators to collide particles into each other. Presently, there are two primary facilities dedicated to heavy-ion physics: the Large Hadron Collider (LHC) at CERN in Switzerland, and the Relativistic Heavy Ion Collider (RHIC) at Brookhaven National Laboratory in the United States. Additionally, a major new facility, the Electron-Ion Collider (EIC), is under construction. Its primary goal will be to explore the properties of gluons.

2.8.1 The Large Hadron Collider

The Large Hadron Collider (LHC) is the world's largest and most powerful particle accelerator. It first started up in 2008, and still remains the latest addition to the CERN accelerator complex. The LHC consists of a ring of superconducting magnets with a circumference of 27-kilometer, together with a number of accelerating structures to boost the energy of the particles along the beam pipe.

Inside the accelerator's beam pipes, two high-energy particle beams travel at nearly the speed of light, in opposite directions in separate beam pipes, both tubes kept at ultra-high vacuum. They are guided around the accelerator ring by a strong magnetic field

maintained by superconducting electromagnets, conducting electricity without resistance or loss of energy. To maintain the superconducting state, the electromagnets must be kept at a temperature of $-271.3°$, colder than outer space. To achieve this, the accelerator is connected to a distribution system of liquid helium.

Thousands of magnets of various types and sizes are used to direct the beams around the accelerator, including 1232 dipole magnets (15 meters in length) which bend the beams, and the 392 quadrupole magnets (each 5–7 metres long), that focus the particle beams. Another type of magnet is used to squeeze the particles inside the beams closer together, to increase the chances of a collision between two particles.

All the controls for the accelerator, its services and technical infrastructure are housed under one roof at the CERN Control Centre. From here, the beams inside the LHC are made to collide at four locations around the accelerator ring, corresponding to the positions of the four big particle detector systems: ATLAS, CMS, ALICE and LHCb. An illustration of the CERN accelerator complex is shown in Fig. 2.22.

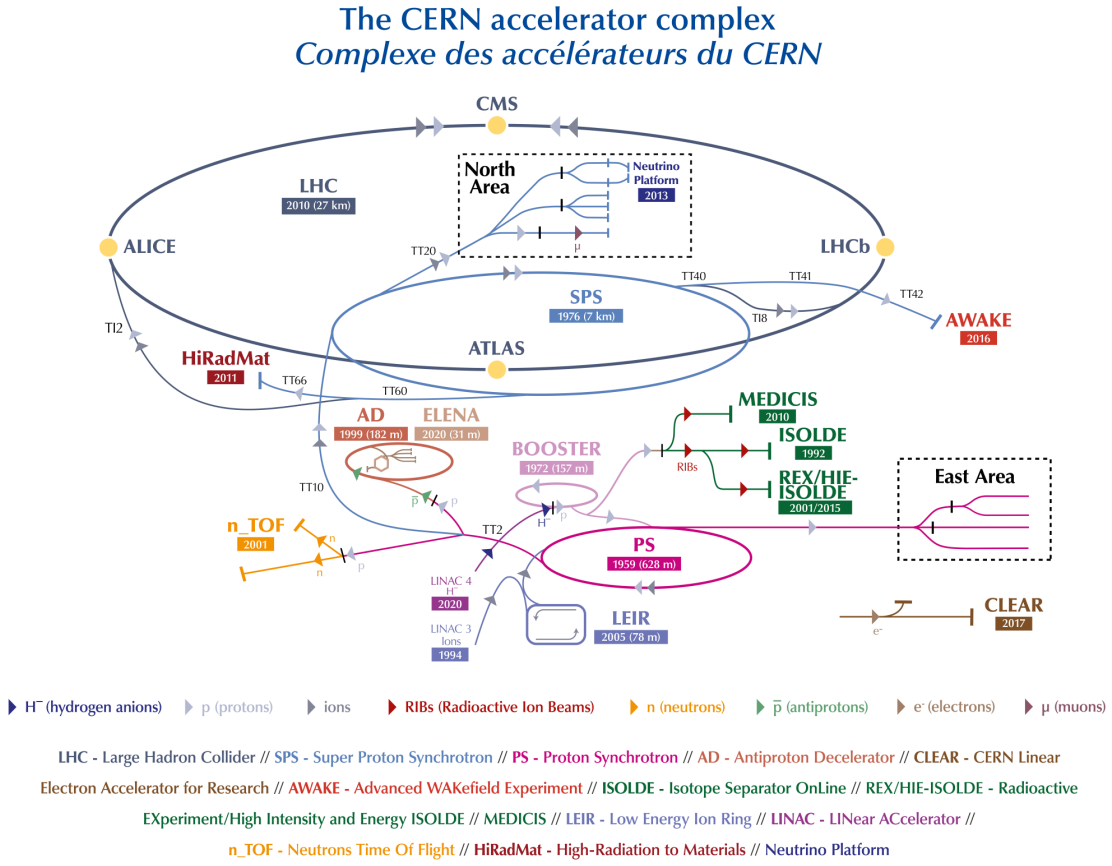

Figure 2.22: The CERN accelerator complex, layout in 2022 [155].

2.8.2 The ALICE Detector

The ALICE detector is specifically designed for studying the strongly interacting matter formed in ultra-relativistic heavy-ion collisions [77]. The experiment carries out a comprehensive study of the hadrons, electrons, muons, and photons produced in these collisions, to unravel the properties of the quark-gluon plasma, which is formed in these extreme conditions (see Sec. 2.5). ALICE also studies pp and p–A collisions both as a comparison with A–A collisions and in their own right. In 2021, ALICE completed a significant upgrade of its detectors to further enhance its capabilities for the data collection period of Run 3 and 4, which is planned to be finished at the end of 2032. At the same time, upgrade plans are being made for ALICE 3, a next-generation experiment for data collection periods of Run 5 and 6 [107]. The experimental results presented in Chap. 5 and Chap. 6 are based on data collected by the ALICE experiment. In the following I briefly describe the ALICE detector system. More detailed information on its setup and performance can be found in Refs. [156, 157].

The ALICE apparatus consists of a central barrel, covering the pseudorapidity region $|\eta| < 0.9$, a muon spectrometer with a $-4 < \eta < -2.5$ coverage, and both forward- and backward-pseudorapidity detectors employed for triggering, background rejection, and event characterization. The central-barrel detectors used in the analyses covered in this thesis are the Inner Tracking System (ITS), the Time Projection Chamber (TPC), and the electromagnetic calorimeters (EMCal and DCal). They are embedded in a large solenoid magnet which provides a maximum magnetic field of $B = 0.5$ T parallel to the beam direction, in order to bend the tracks of charged particles for charge and momentum determination.

Inner Tracking System

Heavy-flavor hadrons decay weakly and have a long lifetime. The mean lifetime for charm hadrons is $\tau_c \approx 150\ \mu m$ and $\tau_b \approx 400\ \mu m$ for beauty hadrons, therefore the position of their decay (the secondary vertex) will be displaced from the position of the collision (primary vertex). One key task of the ITS [158] is to locate the primary vertex with a precision of a tenth of a millimeter. Another functions of the ITS are the reconstruction of the secondary vertices from the decays of D and B mesons and hyperons, the tracking and identification of particles with momentum below 200 MeV/c, and improving the angle and momentum resolution for particles reconstructed by the TPC. It consists of six layers of silicon detectors, with the innermost two composed of Silicon Pixel Detectors (SPD).

Time Projection Chamber

The TPC [159] is the main particle tracking detector of the central barrel. It is a $90m^3$ gaseous chamber capable of three-dimensional reconstruction of charged-particle tracks. The detection is based on the charged particles ionizing the gas atoms along the path they travel inside the gas of the TPC. The liberated electrons then drift towards the end plates of the detector where they are detected. The TPC measures the specific energy loss ($\mathrm{d}E/\mathrm{d}x$) of charged particles in the detector gas. The particles can then be identified using the Bethe-Bloch formula, which describes the average energy loss of charged particles through inelastic Coulomb collisions with the atomic electrons of a medium.

Electromagnetic Calorimeter

The EMCal and DCal detectors [160, 161] are shashlik-type sampling calorimeters consisting of alternate layers of lead absorber and scintillator material. The EMCal covers ranges of $|\eta| < 0.7$ in pseudorapidity and $\Delta\varphi = 107^\circ$ ($80^\circ < \varphi < 187^\circ$) in azimuth. The DCal is located azimuthally opposite the EMCal, with a coverage of $0.22 < |\eta| < 0.7$ and $\Delta\varphi = 60^\circ$ ($260^\circ < \varphi < 320^\circ$) and $|\eta| < 0.7$ and $\Delta\varphi = 7^\circ$ ($320^\circ < \varphi < 327^\circ$). The EMCal and DCal are part of the same detector system, used for electron identification.

V0 Detector

The V0 detector, consisting of two scintillator arrays [162], placed on each side of the interaction point (with pseudorapidity coverage $2.8 < \eta < 5.1$ and $-3.7 < \eta < -1.7$) are utilized for triggering and offline rejection of beam-induced background events. The minimum bias trigger is defined requiring coincident signals in both scintillator arrays of the V0 detector. In p–Pb collisions, the contamination from beam-induced background interactions and electromagnetic interactions can be further removed with the information of the Zero Degree Calorimeters (ZDC) [163], located along the beam line at 112.5 m on both sides of the interaction point. A T0 detector [164], composed of two arrays of quartz Cherenkov counters, covering an acceptance of $4.6 < \eta < 4.9$ and $-3.3 < \eta < -3.0$, is employed to determine the luminosity together with the V0 detector.

The schematics of the ALICE detector is shown in Fig. 2.23 for the Run 2 data taking period, corresponding to the data analyses in Chapters 5 and 6.

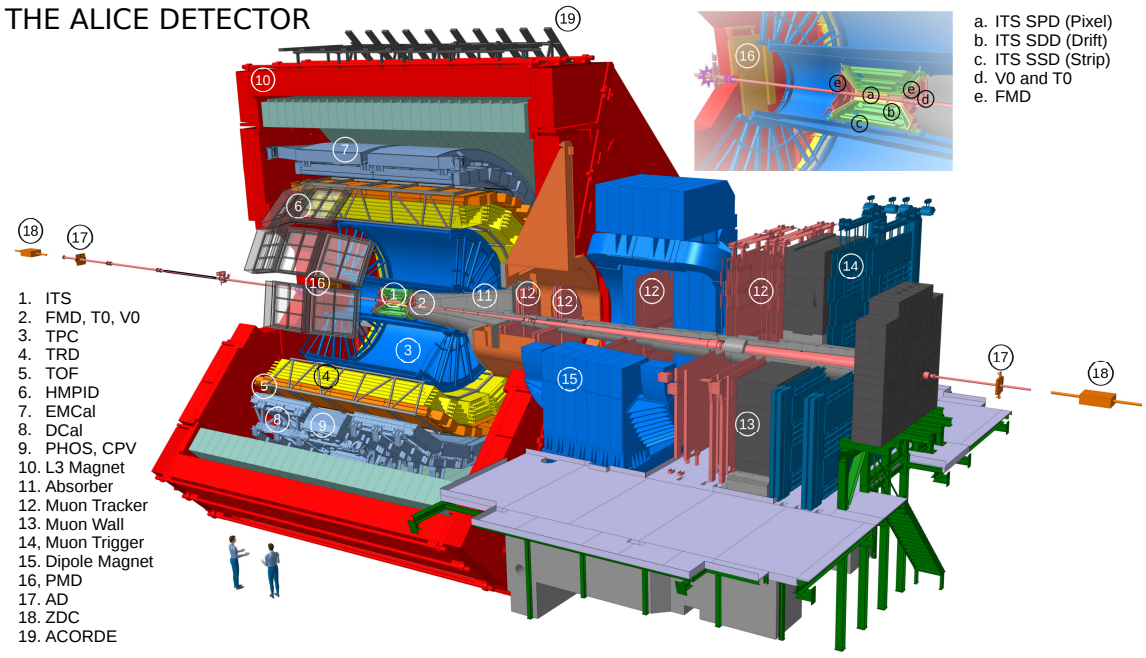

Figure 2.23: Schematics of the ALICE detector during Run 2 [165].

Chapter 3

Multiplicity Dependence of Jet Shapes

Jet profile measurements in hadron colliders have long been suggested as sensitive probes of QCD parton splitting and showering calculations [166–168], and even as an indicator of the quark-gluon plasma (QGP) [168]. A recent study suggests to verify a possible existence of a QGP-droplet by measuring properties of jets in association with a Z-boson in ultra-central pp collisions [169]. Another suggestion is to measure the structures of photon-tagged recoil jets in comparison to inclusive jet. This would allow for the detailed understanding of the color-charge effect in fragmentation as well as the elimination of surface bias in jet quenching [170]. In experiment, jet structure observables with full jet reconstruction have been studied in different collisional systems at HERA, the Tevatron, and the LHC [42, 171–177], among others. Gaining a detailed understanding of the multiplicity dependence of jet structures up to high momenta is particularly important with the recent advent of machine learning classification techniques in jet studies. These techniques rely heavily on modeling parton shower and fragmentation, and their connection to the underlying event and avoiding possible selection biases is essential [178, 179].

In this analysis I extensively studied the multiplicity-dependent jet shapes with the PYTHIA 8 event generator [12], using different tunes and setups to examine the possible effects of multiple-parton interactions (MPI) on jets. I provide predictions for pp collisions at $\sqrt{s} = 7\ \mathrm{TeV}$ to motivate similar, multiplicity-dependent jet structure measurements at the LHC. In models with string hadronization, the recombination of overlapping color strings (color reconnection or CR) influence fragmentation and are also known to produce collective-like patterns such as radial flow [180]. I investigate the effects caused by the choice of the CR scheme within PYTHIA on the simulated jet structures. As a reference point in my investigations, I decided to use a set of jet structure measurements by the CMS experiment at $\sqrt{s} = 7\ \mathrm{TeV}$, carried out in a wide jet momentum range from $15\ \mathrm{GeV}/c$ up to $1000\ \mathrm{GeV}/c$ [42]. A previous CMS study investigated multiplicity-differential jet structures, albeit momentum-inclusively with a $p_{\mathrm{T}}^{\mathrm{jet}} > 5\ \mathrm{GeV}/c$ jet transverse momentum threshold, to understand the influence of the underlying event (UE) on jets [177].

This chapter is organized as follows. In Section 3.1, I describe the analysis method in detail. In Section 3.2, I present the event multiplicity distributions and discuss the differential and integral jet shape measurements. The observation of a characteristic jet size is also presented, and a double ratio of differential jet shapes is proposed to measure the effects of multiplicity. Finally, I summarize my results in Section 3.3.

3.1 Analysis Method

I used the PYTHIA 8.226 [12] event generator to generate random pp collisions at a center-of-mass energy of $\sqrt{s} = 7\ \mathrm{TeV}$. I allowed any hard pQCD process, but in order to decrease simulation time I limited the phase space by requiring a certain minimum invariant transverse momentum $\hat{p_{\mathrm{T}}}$ of the hardest $2 \to 2$ process in an event.

I selected $\hat{p_{\mathrm{T}}} > 5\ \mathrm{GeV}/c$, $\hat{p_{\mathrm{T}}} > 20\ \mathrm{GeV}/c$, $\hat{p_{\mathrm{T}}} > 40\ \mathrm{GeV}/c$ and $\hat{p_{\mathrm{T}}} > 80\ \mathrm{GeV}/c$ for the evaluation of jets with $p_{\mathrm{T}}^{\mathrm{jet}} > 15\ \mathrm{GeV}/c$, $p_{\mathrm{T}}^{\mathrm{jet}} > 50\ \mathrm{GeV}/c$, $p_{\mathrm{T}}^{\mathrm{jet}} > 80\ \mathrm{GeV}/c$, and $p_{\mathrm{T}}^{\mathrm{jet}} > 125\ \mathrm{GeV}/c$, respectively. These cutoffs were determined so that they do not have influence on the shape of the reconstructed $p_{\mathrm{T}}^{\mathrm{jet}}$ spectrum.

I simulated 5 million events for each setting. Besides the default tune Monash 2013 (Monash), I investigated two others, the MonashStar and 4C tunes (see Sec. 2.3.1). Using the Monash tune as a starting point I also did investigations where I changed some settings in PYTHIA to directly study their effect on the jet structure. There are continuously developed models of multiple-parton interactions implemented in PYTHIA [181, 182]. To understand the multiplicity-dependent jet modification by MPI, I used data samples where I switched this effect on and off.

I also studied the different color reconnection schemes provided by PYTHIA, including turning off this feature completely. Color reconnection is a built-in mechanism in PYTHIA that allows interactions between partons originating in MPI and initial/final state radiations, by minimizing color string length. Since this procedure is quite ambiguous, several models are implemented. The original MPI-based scheme used in PYTHIA 8.226 (that I denote CR0 in the followings) relies on the parton shower-like configuration of the beam remnant. In an additional step, it merges the gluons of a lower-p_{T} MPI system with gluons of a higher-p_{T} MPI system. A newer QCD-based scheme [122] (CR1) relies, however, on the full QCD color configuration in the beam remnant. Then the color reconnection is made by minimizing the potential string energy. The QCD color rules are incorporated in the CR to determine the probability that a reconnection is allowed. This model also allows the creation of junction structures. Besides the above-mentioned CR schemes, a so-called gluon move scheme [183] (CR2) has been implemented to PYTHIA recently, in which gluons can be moved from one location to another so as to reduce the total string length.

I carried out a full jet reconstruction including both charged and neutral particles, using three popular algorithms, the anti-k_T [40], k_T [38, 184], and Cambridge-Aachen [39, 185] algorithms, provided by the FASTJET [41] software package. All of them are sequential clustering algorithms, meaning that the closest particle tracks in momentum space are sequentially merged one-by-one according to the minimum of a distance measure between the particle four-momenta. While all three algorithms are infrared and collinear safe, in high-multiplicity environments the clusterization outcomes will be rather different. Anti-k_T is popular because it is only slightly susceptible to pile-up and underlying events, and it clusterizes hard jets into nearly perfect cones with a resolution parameter R even in high-multiplicity events, in accordance with the general image of how a jet should look like. The other two algorithms are more suitable for jet substructure studies but provide jets of irregular shape that are not uniform in area, especially the k_T algorithm, where the area of the jets fluctuates considerably [41].

Similarly to the CMS analysis [42, 177], I selected inclusive jets, with a resolution parameter $R = 0.7$. I considered constituent particles, with a transverse momentum threshold $|p_{\mathrm{T,track}}| > 0.15\ \mathrm{GeV}/c$, at the generator level. My experience matches earlier findings that the detector effects, after corrections, do not change the simulated jet observables significantly [42]. I examined jets in the pseudorapidity window $|\eta_{\mathrm{jet}}| < 1$ and restricted my investigations to the $15\ \mathrm{GeV}/c < p_{\mathrm{T}}^{\mathrm{jet}} < 400\ \mathrm{GeV}/c$ jet momentum range, where multiplicity-differential studies are feasible in the near future.

For the investigation of a possible jet shape modification I analyze the transverse momentum profile of the jets as in [42]. As a first step I showed that my simulations reproduce CMS data [42] within uncertainty throughout this range. I show examples in three different $p_{\mathrm{T}}^{\mathrm{jet}}$ windows in Fig. 3.3. For harder jets, the calculated momentum density distribution gets steeper in the central (small-r) region of the jets, in qualitative accordance with the calculations of [168].

I investigate the jet structure for different charged hadron multiplicity (N_{ch}) classes. Generally, PYTHIA is known to reproduce multiplicities in LHC data with little differences over a broad p_{T} range. Charged hadron multiplicities at mid-rapidity are well reproduced by the 4C tune except for a slight discrepancy at very high N_{ch} values [177].

The CUETP8M1 and Z2* tunes reproduce pion and kaon average p_{T} versus track multiplicities within errors [186]. The D6T and Z2 tunes show a marginal agreement with the CMS jet-multiplicity data, with about 5% higher predictions than the mean values, flat in p_{T} [42]. I use charged hadron multiplicity at mid-rapidity (referred to as multiplicity in the followings for the sake of simplicity), defined as the number of the charged final state particles with $|\eta| < 1$ in a given event.

As my main focus, I carry out a detailed study on the multiplicity dependence of ρ, but I also use Ψ evaluated at fixed radii as it allows for easier experimental comparison.

3.2 Results

In this section I present my results and consider the physical implications. As a first step I compute $\rho(r)$ similarly to Fig. 3.3, but this time I divide up the data into two multiplicity classes, $N_{\rm ch} \leq 50$ and $50 < N_{\rm ch}$, respectively. We see a multiplicity dependence in the jet shapes in Fig. 3.4. Namely, the jets contain a higher fraction of their transverse momentum closer to their axis and a lower fraction further away from their axis in the case of low multiplicity. For high multiplicity the jet shape behaves in the opposite way. This is a trivial, expected multiplicity dependence arising from two reasons. The first one is that event multiplicity is correlated with jet multiplicity, resulting in a higher fraction of narrow jets in low-$N_{\rm ch}$ events. The second reason is the UE background, which affects the jet structure more at higher r values, and its effect is stronger in the case of high-$N_{\rm ch}$ events.

3.2.1 Event-multiplicity Distributions

I show the multiplicity distributions in Fig. 3.1 for the jet momentum window $110\ \mathrm{GeV}/c < p_{\rm T}^{\rm jet} < 125\ \mathrm{GeV}/c$ as an example. As shown in the left panel, distributions of the multiplicity are very similar for the different tunes. However, when considering the multiplicity distribution from different settings of the Monash tune, shown on the right panel, a substantial difference can be seen between the settings with and without MPI or CR. Disabling MPI (and CR, which assumes MPI) causes the distribution to shift toward lower values, while keeping a similar shape.

Disabling CR only, on the other hand, causes the multiplicity distribution to extend toward higher values. This means that care should be taken when one compares distributions with MPI or CR settings on and off, as it may be biased when the chosen multiplicity class is too wide. I note that multiplicity distributions from different color reconnection schemes do not differ significantly.

The $p_{\rm T}^{\rm jet}$ dependence of the mean and RMS values of the multiplicity distribution is compared in Fig. 3.2 for different tunes, as well as for different settings in the case of the Monash tune. The three tunes predict very similar mean and RMS values throughout the $p_{\rm T}^{\rm jet}$ range. While the means of the 4C and Monash tunes overlap, MonashStar predicts slightly lower multiplicities.

The $p_{\rm T}^{\rm jet}$ dependence of multiplicity distributions is a key observable for validating the strength of multiple-parton interaction and color reconnection effects in Monte Carlo models, as both MPI and CR have a grave effect on the distributions. Switching off MPI causes a downward shift of about 15 to 25 in mean $N_{\rm ch}$ at any $p_{\rm T}^{\rm jet}$, or almost a factor of three at lower $p_{\rm T}^{\rm jet}$ values, while switching off CR alone causes a somewhat less drastic increase of about 10 to 20 in mean $N_{\rm ch}$ counts. The three examined color reconnection models provide very similar distributions except for lower $p_{\rm T}^{\rm jet}$ values, where the gluon-moving CR scheme predicts slightly narrower multiplicity distributions and a somewhat

lower mean value.

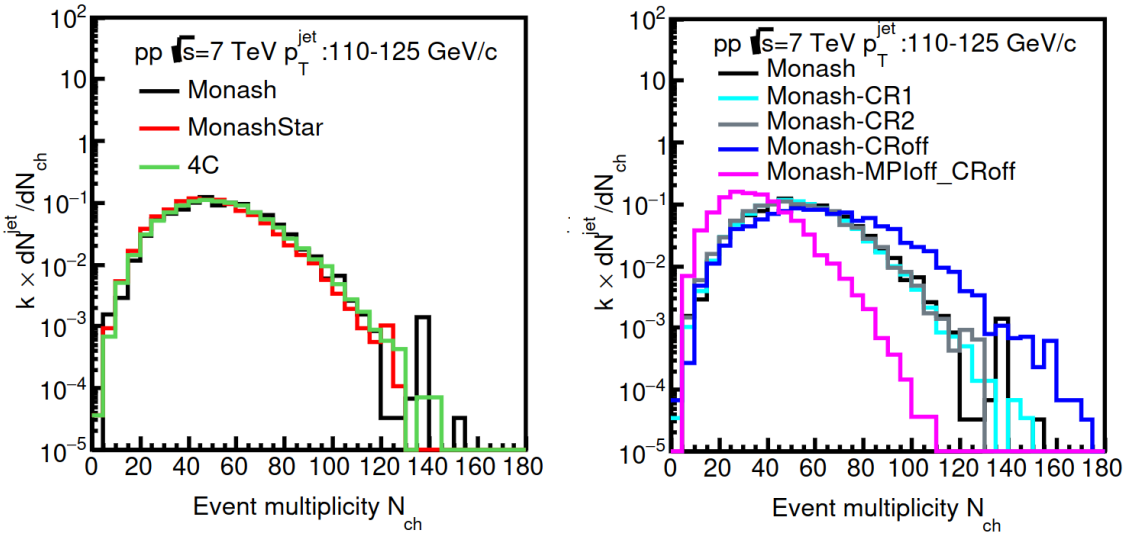

Figure 3.1: Event-multiplicity distributions for jets in the $110\ \mathrm{GeV}/c < p_{\mathrm{T}}^{\mathrm{jet}} < 125\ \mathrm{GeV}/c$ window, compared for the Monash, MonashStar, and 4C tunes (*left*) and for the Monash tune with CR0, CR1 and CR2 settings applied, as well as both CR and MPI turned off (*right*).

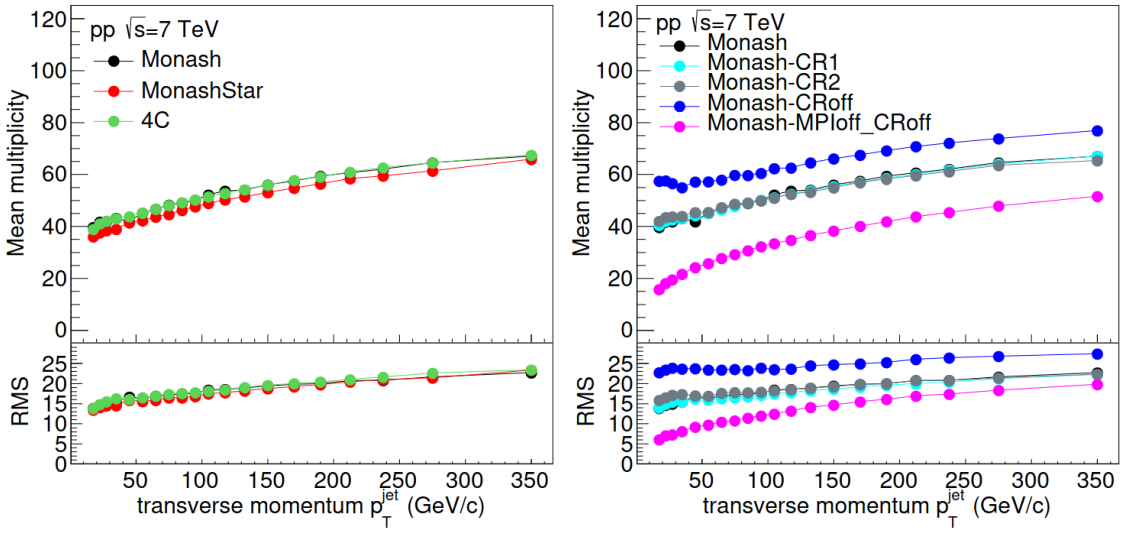

Figure 3.2: Mean and RMS values of the event-multiplicity distributions as a function of the $p_{\mathrm{T}}^{\mathrm{jet}}$, compared for different tunes (*left*) and settings (*right*). The uncertainties of the mean and RMS values are comparable to the symbol size.

3.2.2 Differential Jet Shapes

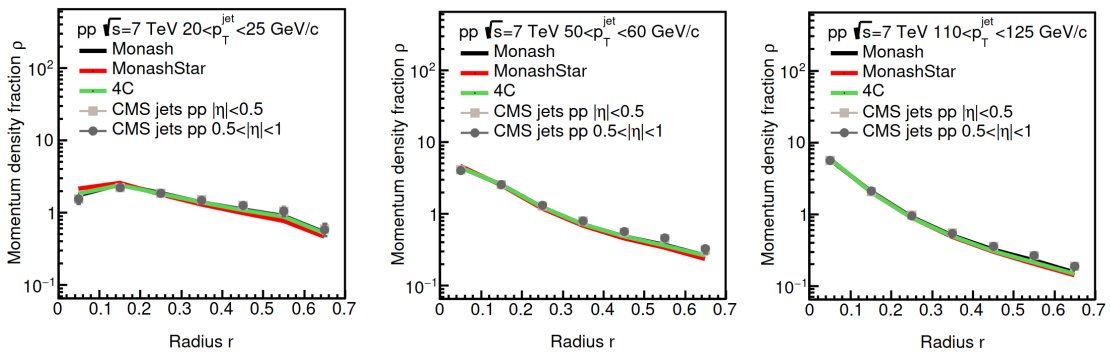

Figure 3.3: Differential jet shape $\rho(r)$ for different PYTHIA tunes compared to measurements of CMS experiment in pp collisions at $\sqrt{s} = 7$ TeV [42], for $20\ \mathrm{GeV}/c < p_{\mathrm{T}}^{\mathrm{jet}} < 25\ \mathrm{GeV}/c$ (*left*), $50\ \mathrm{GeV}/c < p_{\mathrm{T}}^{\mathrm{jet}} < 60\ \mathrm{GeV}/c$ (*center*) and $110\ \mathrm{GeV}/c < p_{\mathrm{T}}^{\mathrm{jet}} < 125\ \mathrm{GeV}/c$ (*right*).

Measurements by the CMS experiment [177] that compare five multiplicity classes within the range $10 < N_{\mathrm{ch}} \leq 140$ and reconstruct jets at momenta $p_{\mathrm{T}}^{\mathrm{jet}} > 5\ \mathrm{GeV}/c$ saw a remarkable difference between low- and high-multiplicity $\rho(r)$ for low r values. The same is observed for relatively low $p_{\mathrm{T}}^{\mathrm{jet}}$ values (3.4 left panel). Dividing $\rho(r)$ for both the high- and low-multiplicity classes with the multiplicity-integrated $\rho_{\mathrm{MI}}(r)$ (no condition on N_{ch}), shown in Fig. 3.5, highlights this trend. The curves are much more apart for low r values at small $p_{\mathrm{T}}^{\mathrm{jet}}$, while there is relatively little difference between different $p_{\mathrm{T}}^{\mathrm{jet}}$ windows for high r values. This suggests that jets in high-multiplicity events contain much more contribution from the soft regime, and soft physics is selected by a choice of lower momentum range.

Jets in low-multiplicity events are on average narrower than in high-multiplicity events, hence the corresponding $\rho_{\mathrm{MI}}(r)$ ratio is above unity, while for high-multiplicity events this ratio is below unity. At high r values, where UE tracks give a non-negligible contribution especially in the high-multiplicity events, the situation is just the opposite. In between, there is a point at a given r value where the two curves intersect each other at unity, meaning that the jets are just average at that radius. In Fig. 3.5, we see three examples in different $p_{\mathrm{T}}^{\mathrm{jet}}$ windows and we can observe that the intersection point has a jet-momentum dependence.

This is not unexpected since harder jets are narrower and UE is significant already at smaller radii. To have a closer look at this behavior, I evaluate $\rho_{\mathrm{MI}}(r)$ in a more refined division of data with seven multiplicity classes in the range $1 \leq N_{\mathrm{ch}} \leq 250$. We find that the curves intersect unity at virtually the same location for a given $p_{\mathrm{T}}^{\mathrm{jet}}$ value. This statement holds even if we compare different PYTHIA tunes and MPI or CR settings, as

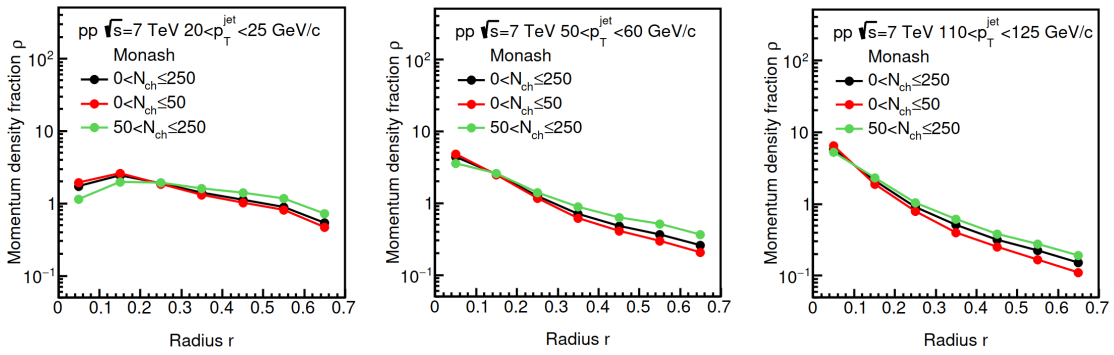

Figure 3.4: Comparing the differential jet shape $\rho(r)$ of multiplicity-integrated (*black*), low-multiplicity ($N_{\mathrm{ch}} \leq 50$, *red*), and high-multiplicity ($N_{\mathrm{ch}} > 50$, *green*) events, for 20 GeV/c $< p_{\mathrm{T}}^{\mathrm{jet}} <$ 25 GeV/c (*left*), 50 GeV/c $< p_{\mathrm{T}}^{\mathrm{jet}} <$ 60 GeV/c (*center*) and 110 GeV/c $< p_{\mathrm{T}}^{\mathrm{jet}} <$ 125 GeV/c (*right*).

shown on the examples in Fig. 3.6 for the Monash and 4C tunes as well as the Monash tune without color reconnection.

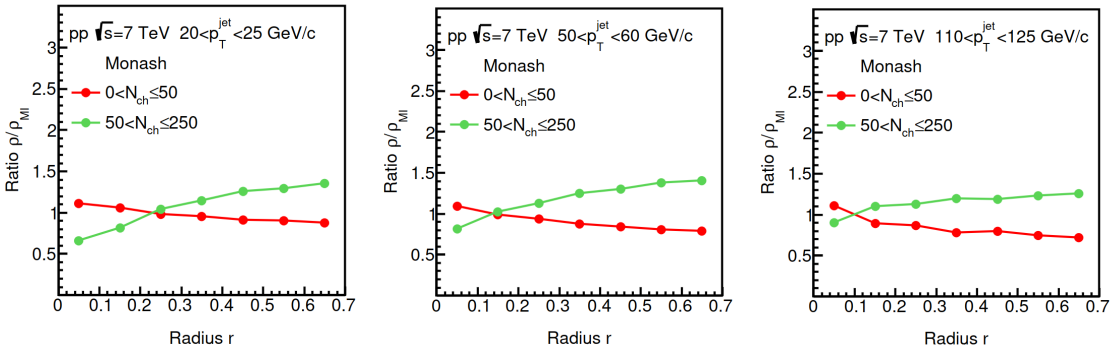

Figure 3.5: Ratio $\rho(r)/\rho_{\mathrm{MI}}(r)$ of differential jet shape in low-multiplicity ($N_{\mathrm{ch}} \leq 50$, *red*) and high-multiplicity events ($N_{\mathrm{ch}} > 50$, *green*) over multiplicity-integrated events, for 20 GeV/c $< p_{\mathrm{T}}^{\mathrm{jet}} <$ 25 GeV/c (*left*), 50 GeV/c $< p_{\mathrm{T}}^{\mathrm{jet}} <$ 60 GeV/c (*center*) and 110 GeV/c $< p_{\mathrm{T}}^{\mathrm{jet}} <$ 125 GeV/c (*right*).

3.2.3 Characteristic Jet Size

In the left and center panels of Fig. 3.7, the r dependence of the intersection radius is plotted with respect to the jet transverse momentum for different tunes, as well as for the different settings of the Monash tune. The intersection radius is computed using a linear interpolation between the two nearest points of $\rho(r)/\rho_{\mathrm{MI}}(r)$, and its uncertainty is estimated by taking both the high- and the low-multiplicity classes, moving the points to the upper and lower edge of their error bars in both cases, and determining the maximum

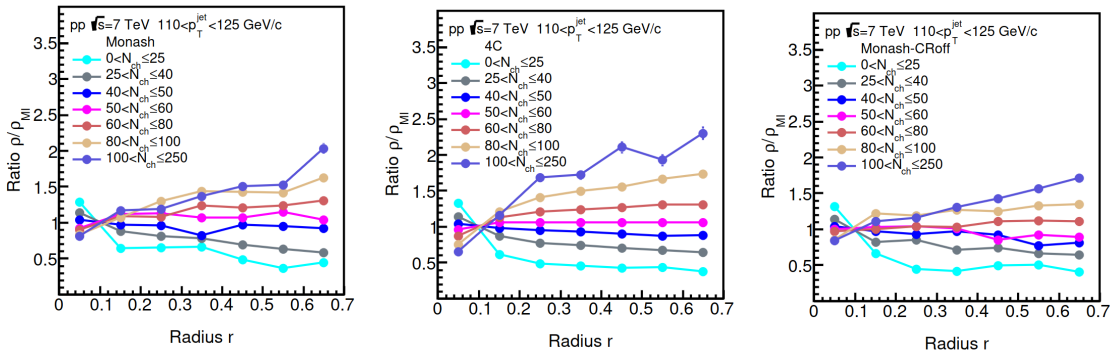

Figure 3.6: Ratio $\rho(r)/\rho_{\mathrm{MI}}(r)$ of differential jet shapes in several multiplicity classes over multiplicity integrated events, for $110\ \mathrm{GeV}/c < p_{\mathrm{T}}^{\mathrm{jet}} < 125\ \mathrm{GeV}/c$. The panels show events generated using the Monash tune (*left*), the 4C tune (*center*), and the Monash tune without CR (*right*).

and minimum values of the intersection radius from these cases. We observe that for all tunes and settings that I tested, the intersection radii are consistent within uncertainties for any chosen $p_{\mathrm{T}}^{\mathrm{jet}}$ value. There is additional uncertainty on the obtained intersection radius stemming from the linear interpolation between finite, $\delta r = 0.1$ wide bins. In order to estimate this, I repeated the analysis with the three tunes in $r = 0.05$ wide bins. While the statistical fluctuations increase, the points move a maximum of 4% upwards or 28% downwards in a strongly correlated manner (see Fig. 3.7).

Nevertheless, the overall shape of the curves remains very similar and statistically consistent between different tunes point-by-point. Therefore, I suggest that the intersection radius $r_{\mathrm{ch}} = r|_{\rho=\rho_{\mathrm{MI}}}$ be considered as a characteristic jet size measure specific for a given jet transverse momentum. I note that this quantity was first referred to as R_{fix} in my publication [187], but it should not be compared to the jet resolution parameter R, which is typically chosen so that most of the jet momentum is contained within the radius R. In contrast, r_{ch} is defined as a radius where the momentum density of the jet from events of any multiplicity is just like in the average jet, and substantial fraction of jet momentum falls toward smaller as well as toward larger radii.

Jet shapes depend on the jet reconstruction algorithm, so I investigated whether the observed stability of the intersection radius can be an artifact of the jet reconstruction algorithm itself. Besides the anti-k_{T} algorithm which I first utilized, I have reprocessed all the data with using the k_{T} and the Cambridge-Aachen algorithms. There is no significant difference beyond the statistics-driven fluctuations between data reconstructed by different clusterization algorithms in any of the tunes or MPI/CR settings. In the right panel of Fig. 3.7, I show a comparison of $r_{\mathrm{ch}}(p_{\mathrm{T}}^{\mathrm{jet}})$ for the Monash tune with the three different jet reconstruction algorithms.

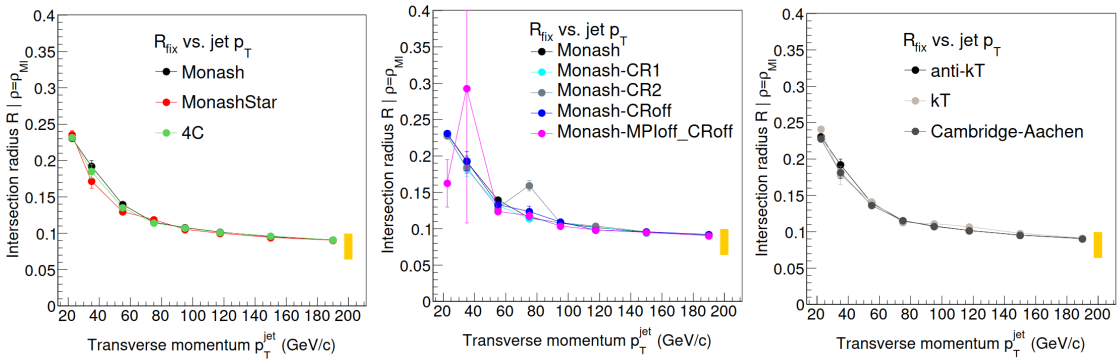

Figure 3.7: Evolution of the characteristic radius r_{ch} as a function of $p_{\mathrm{T}}^{\mathrm{jet}}$, for several PYTHIA tunes (*left*), settings (*center*), and jet reconstruction algorithms for the Monash tune (*right*). The absolute uncertainty arising from the choice of bin width is indicated by the yellow shaded band.

Jets are more collimated with increasing transverse momentum. In a simple picture this can be linked to applying a Lorentz boost, i.e., to the momentum of the initiating parton in the laboratory system. The $p_{\mathrm{T}}^{\mathrm{jet}}$-dependent evolution of r_{ch} may also be explained by Lorentz-boosted high-p_{T} jets (see the illustration in the left panel of Fig. 3.8).

In order to gain an effect-level understanding, I use a simplistic model. I consider particles radiating from a point in a plane with momenta of equal absolute value p_0. I boost these particles along the axis perpendicular to their plane, with a certain momentum p_{boost}. The resulting particles will form a cone around the boost axis in the lab system, representing our "jet". In the right panel of Fig. 3.8, we see that the resulting size of the "jet", R_{cone}, depends on p_{boost} in a qualitatively similar manner to how the intersection radius r_{ch} depends on $p_{\mathrm{T}}^{\mathrm{jet}}$. This attests to the assumption that the universal behavior can, at least partially, be understood by the narrowing by Lorentz-boost of high-p_{T} jets. It is to be noted, however, that this model is a simplistic treatment of jet narrowing, and therefore an exact agreement is not expected. A proper description requires a perturbative QCD-based approach where the narrowing is a consequence of the running of the strong coupling constant α_s with k_{T} [166–168]. Since there is no angular cut-off that would limit the jet sizes at low momenta, R_{cone} is allowed to blow up at low p_{boost} values. Also, one cannot expect real jets to go below a certain size because after certain point the clustering algorithms will be driven by the presence of the UE. This may explain the apparent convergence of the r_{ch} curves to a finite value at high p_{T}. As mentioned before, r_{ch} at high-p_{T} is also influenced by the choice of δr. A particularly interesting question is whether r_{ch} can be generalized to the larger and more complex systems produced in heavy-ion collisions. To see that, one would need to do simulations in heavy-ion collisions and verify the outcome with data. In case r_{ch} is representative of the jet size in heavy-ion collisions, it would provide a handy observable for the exploration of medium modification of jets.

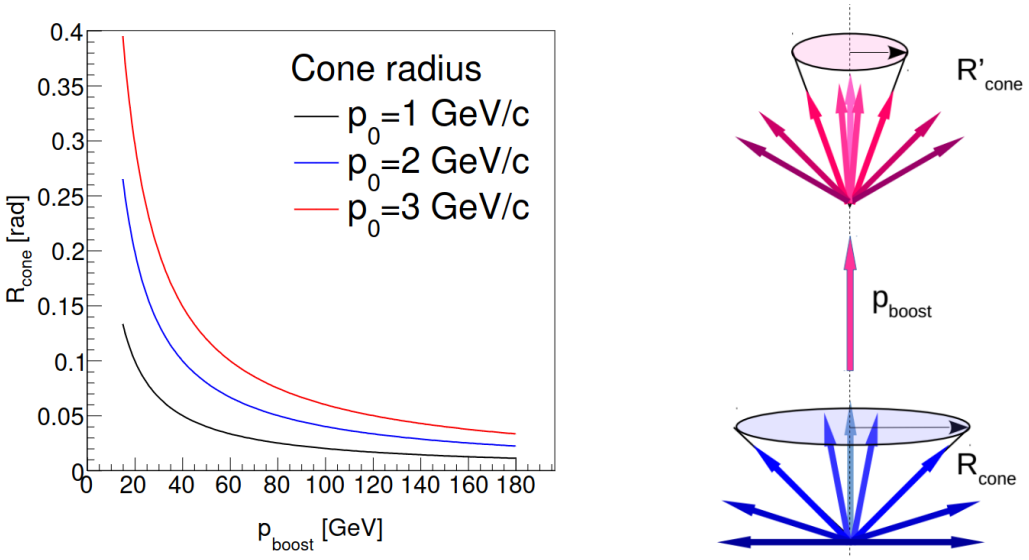

Figure 3.8: Illustration of the jet cone narrowing by Lorentz boost (*left*). Evolution of cone radius based on a simplistic boosted cone model for different p_0 assumptions (*right*).

3.2.4 Momentum Fractions

In Fig. 3.9, the integral jet shape is plotted as a function of the multiplicity, at high momenta within $180\ \mathrm{GeV}/c < p_{\mathrm{T}}^{\mathrm{jet}} < 200\ \mathrm{GeV}/c$. In the left panel, where tunes are compared, there is no observable effect in the integral structure between the tunes Monash, MonashStar, and 4C. I present the effects of different MPI and CR settings on the integral jet structure in the right panel of Fig. 3.9. Different color reconnection schemes do not lead to significant differences, but there is a slight deviation at high N_{ch} values when color reconnection is turned off. However, the lack of MPI causes a significant difference within the same multiplicity class, that grows approximately linearly with N_{ch}, which suggests that the MPI has a strong influence on the jet structure, especially at high N_{ch} values. It is to be noted that the effect is less significant in case of lower $p_{\mathrm{T}}^{\mathrm{jet}}$ windows and in case of larger r values.

At lower multiplicities, MPI and CR cause little difference in the integrated jet shape. The $\Psi(r)$ values at high N_{ch} are lower in the case when MPI is turned off, meaning that the jets are more concentrated in a narrow cone. This can be understood by a higher relative fraction of soft tracks coming from the UE in case when there is no MPI, compared to the MPI case with the same multiplicity where there is a more relevant contribution from tracks which come from the jet itself. Note that the points in Fig. 3.9 are not at the bin centers, but they are placed to represent the weight of the N_{ch} distribution in a given bin, to eliminate the possible bias stemming from different N_{ch} distributions within multiplicity classes.

Understanding the observed dependence of the integrated jet structure on the multiplicity needs further analysis supported by experimental data. The above observation, if compared to real data, may provide a control over the extent of MPI effects. Further studies are needed to identify MPI/CR effects and separate them from the UE, also using other observables that are less sensitive to the UE.

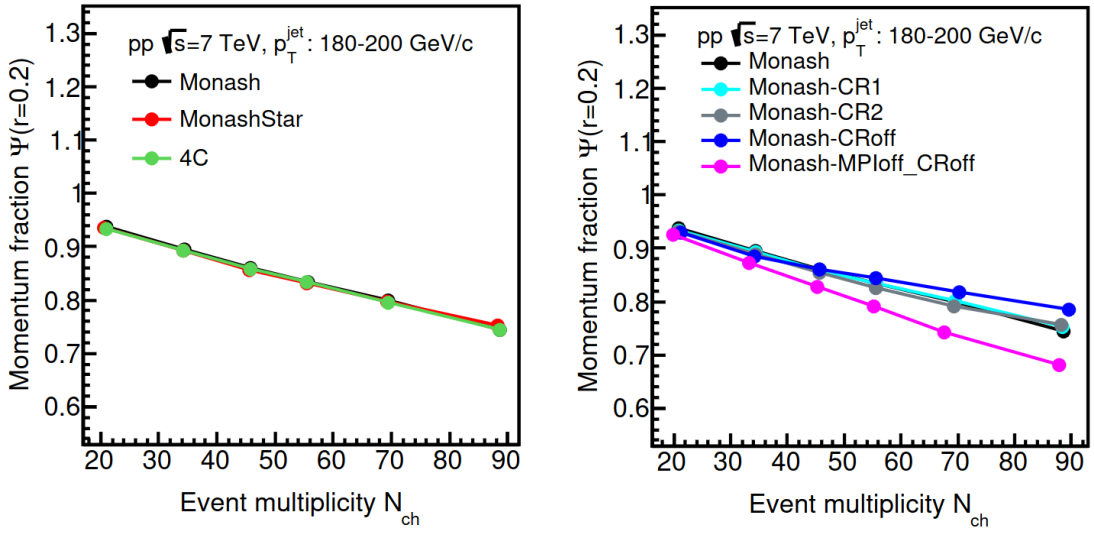

Figure 3.9: Evolution of the integral jet shape $\Psi(r = 0.2)$ as a function of the event multiplicity N_{ch} within $180\ \mathrm{GeV}/c < p_{\mathrm{T}}^{\mathrm{jet}} < 200\ \mathrm{GeV}/c$, with a jet resolution parameter $R = 0.7$. Several PYTHIA tunes (*left*) and settings (*right*) are compared. The points are placed according to the weight of the distribution in each multiplicity class.

3.2.5 Double Ratio of Momentum Densities

In Fig. 3.10 the differential jet structure for various PYTHIA tunes is plotted in a particular $p_{\mathrm{T}}^{\mathrm{jet}}$ window to compare them in the low- and high-multiplicity regions. In the right panel of Fig. 3.10, I take the differential jet shapes for the above-mentioned low- and high-multiplicity classes and divide them with each other to highlight the differences for the different tunes. As expected, jets in low-multiplicity events have a more steeply falling momentum density distribution than the ones in high-multiplicity events, which is also reflected in a falling ratio. However, there are also certain significant differences between the selected tunes that are beyond this trivial effect.

To highlight the differences between the jet structures from different tunes, I compute the double ratio

$$\mathrm{DR}(r) = \frac{\rho_{\mathrm{low}}/\rho_{\mathrm{high}}}{(\rho_{\mathrm{low}}/\rho_{\mathrm{high}})_{\mathrm{ref.tune}}}\ , \tag{3.1}$$

where I divide the former ratio of the high- and low-multiplicity classes with the very same

ratio calculated for the Monash tune. After the trivial effect is gone, a rather sizable effect in the order of a factor of 2 can be seen for both the 4C and the MonashStar tunes, with respect to the Monash as the reference tune. The right panel of Fig. 3.11 shows the same calculations for the 4C tune, for several different choices of high- and low-multiplicity class pairs. In this selected $p_{\mathrm{T}}^{\mathrm{jet}}$ range all of them show similar structures, and generally the effect is larger when the separation in multiplicity is larger. It is very important to note that these curves are derived from statistically independent samples, hence cannot be explained by fluctuations. Since on Fig. 3.10 and 3.11 I calculate ratios of binned data without a bin center correction, I tested its possible effect by decreasing the bin size from $\delta r = 0.1$ to $\delta r = 0.05$. I did not find any difference beyond statistical uncertainties.

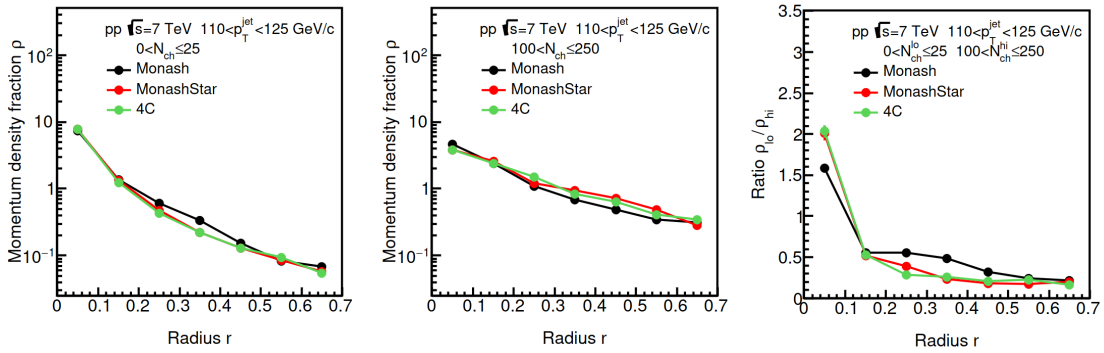

Figure 3.10: Differential jet shape using several PYTHIA tunes, for 110 GeV/c $< p_{\mathrm{T}}^{\mathrm{jet}} <$ 125 GeV/c. The jet shapes from low-multiplicity events ($N_{\mathrm{ch}} \leq 25$, *left*) are compared to high-multiplicity events ($100 < N_{\mathrm{ch}} \leq 250$, *center*), and their ratio computed (*right*).

In order to understand the dependence of the effect on $p_{\mathrm{T}}^{\mathrm{jet}}$, one might wish to describe the deviations for each $p_{\mathrm{T}}^{\mathrm{jet}}$ value with a single number. Therefore I computed the squared sum of the bin-by-bin deviations of the double ratio from the Monash tune, i.e.,

$$\mathrm{RSD} = \sqrt{\sum_{0<r_i<r} (\mathrm{DR}(r_i) - 1)^2} \tag{3.2}$$

at a given $p_{\mathrm{T}}^{\mathrm{jet}}$. In Fig. 3.12, I show the results for different tunes as well as for different selections of multiplicity class pairs. Again, we see a rather parallel behavior of the 4C and MonashStar tunes (in other words, the Monash tune is the one that deviates from these two). The behavior versus $p_{\mathrm{T}}^{\mathrm{jet}}$ is nontrivial with several minima and maxima, and is not easily explained without taking into account peculiar details of each tune. However, one sees again a very strong correlation between curves of different multiplicity selections calculated independently from each other, and that the amplitude strongly depends on the separation between the low- and high-multiplicity classes. Thus we can conclude that the multiplicity-dependent analysis of jet structures in a wide $p_{\mathrm{T}}^{\mathrm{jet}}$ range has the potential of

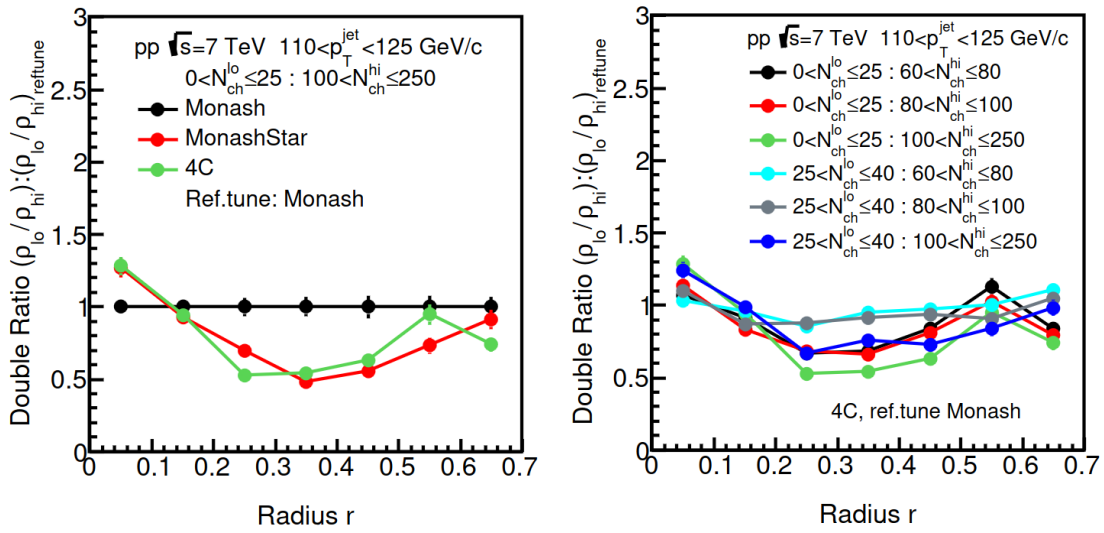

Figure 3.11: The double ratio with Monash as reference tune plotted (*left*). The double ratio of the 4C tune and reference tune plotted for different selections of high- and low-multiplicity classes (*right*).

evaluating the goodness of tunes that otherwise preform equally well in several tests.

3.2.6 Differential Jet Shapes in HIJING++

To verify that the observed characteristic jet size is not a model peculiarity in PYTHIA but a general feature of high-energy collisions, I also computed the differential jet shapes using another MC generator. The HIJING++ was selected for this purpose because it implements a mechanism for creating the underlying event and QCD effects on the soft-hard boundary that is different from PYTHIA (see Section 2.3.2). HIJING++ uses the PYTHIA jet fragmentation, therefore we do not expect any difference during the later stages. Instead of MPI as implemented in PYTHIA, HIJING++ uses minijet production. Differences at lower momenta may arise below the minijet cutoff. In case of the transverse momentum and multiplicity distributions, these effects do not exceed the variation caused by applying different tunes in PYTHIA.

In Fig. 3.13, I show the evolution of the characteristic jet size as a function of $p_{\mathrm{T}}^{\mathrm{jet}}$ using HIJING++, and compare it to the PYTHIA results. In case of HIJING++, the analysis was done using two different PDF sets to observe if there is any difference in the conclusions. The results show quantitatively the same characteristic jet size dependence on $p_{\mathrm{T}}^{\mathrm{jet}}$ within systematic errors.

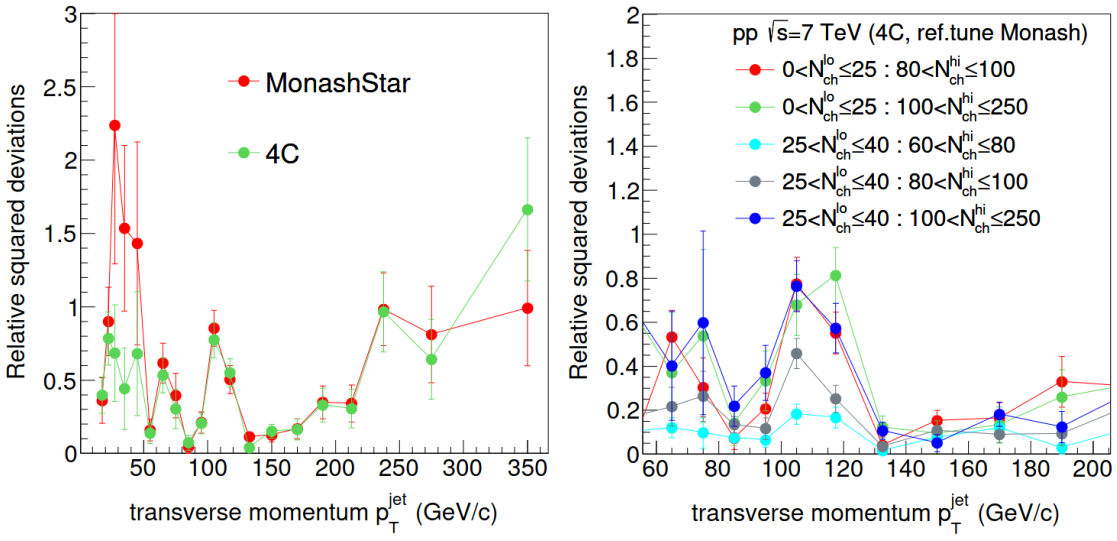

Figure 3.12: Square sum of the bins in the $\mathrm{DR}(r) - 1$ diagram for the 4C (*green*) and MonashStar (*red*) tunes with respect to the Monash reference tune, depending on the $p_{\mathrm{T}}^{\mathrm{jet}}$, for $0 < N_{\mathrm{ch}} \leq 25$ as low-multiplicity, and $80 < N_{\mathrm{ch}} \leq 100$ as high-multiplicity selections (*left*). Square sum of the bins in the $\mathrm{DR}(r) - 1$ diagram for the 4C tunes with respect to the Monash reference tune, depending on the $p_{\mathrm{T}}^{\mathrm{jet}}$, for various low- and high-multiplicity selections, as listed in the legend. (The $p_{\mathrm{T}}^{\mathrm{jet}}$ range is restricted to omit parts with large fluctuations.)

3.2.7 Heavy-flavor Jet Shapes

Jets originating from different flavors undergo different fragmentation due to both the color-charge effect and the dead-cone effect [188]. Therefore, I compared flavor-inclusive jets to heavy-flavor (beauty and charm) jets in the study of jet shapes. It was ensured that heavy flavor comes from the initial stages by only enabling leading order processes in PYTHIA. I also made a comparison to flavor-inclusive jets that contain only leading and subleading jets.

The first thing to note is that the effect of selecting only the leading and subleading jets from events is negligible on the characteristic jet size (see left panel of Fig. 3.14). Although the overall tendency for heavy-flavor is similar to that observed for light flavor, there is also a clear quantitative difference between heavy and light flavors, which points to a different jet structure. The leading beauty jets differ for higher $p_{\mathrm{T}}^{\mathrm{jet}}$ and the leading charm jets for lower $p_{\mathrm{T}}^{\mathrm{jet}}$. This suggests that the interplay between the mass and color-charge effects is non-trivial and needs further investigation. One possibility for that would be a parallel study of the UE and the fragmentation region corresponding to a heavy-flavor trigger, in a similar manner to [153]. From all the above, we can assume that the characteristic jet size is a property of the jets that is associated with the final state.

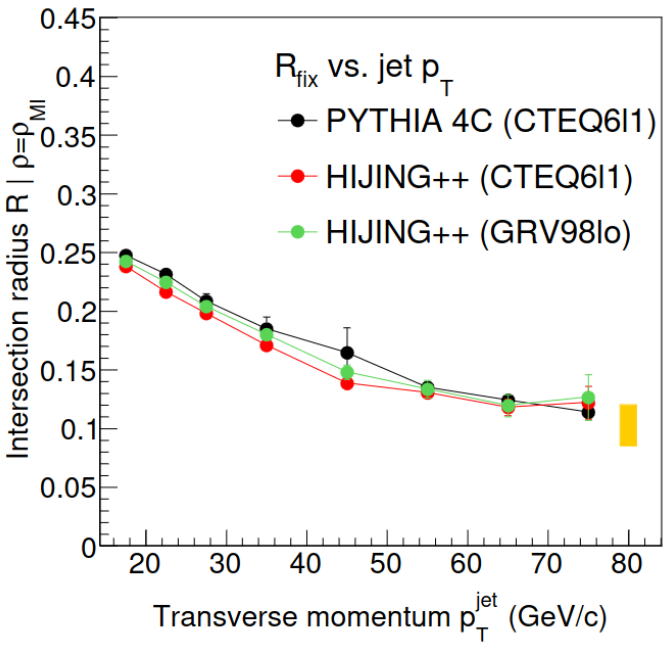

Figure 3.13: Evolution of the characteristic radius r_{ch} as a function of $p_{\mathrm{T}}^{\mathrm{jet}}$ for PYTHIA 4C and for HIJING++ with two different PDFs. The absolute uncertainty arising from the choice of bin width is indicated by the yellow shaded band.

In the right panel of Fig. 3.14, I plotted the integral jet structure at $r = 0.2$ for heavy-flavor jets. Here we can observe that heavy-flavor jets are narrower than inclusive jets on the average. This clearly shows that multiplicity-differential jet structures are sensitive to flavor-dependent fragmentation. However, the effect is interestingly not ordered by mass, since the beauty $\Psi(r = 0.2, N_{\mathrm{ch}})$ curve is between the inclusive and the light jets. It is also to be noted that the ordering is $p_{\mathrm{T}}^{\mathrm{jet}}$-dependent, with a similar trend as the one observed for the characteristic jet size.

3.3 Summary

I performed a novel jet shape analysis in $\sqrt{s} = 7$ TeV pp collisions to explore the multiplicity and $p_{\mathrm{T}}^{\mathrm{jet}}$-dependence of differential and integrated jet structure observables. I used several models implemented in the PYTHIA 8.226 event generator. I demonstrated that the simulations describe CMS data, and I gave predictions for the jet structure observables in several multiplicity classes, over a wide momentum range. I found that there is a given radius r_{ch} where jet momentum density is independent of multiplicity. This radius is insensitive to the choice of simulation settings (choice of tune, presence, and modelling of MPI and CR) within the investigated model class and even of jet clustering algorithms, and its p_{T}-dependence qualitatively follows a Lorentz-boost curve. These observations suggest

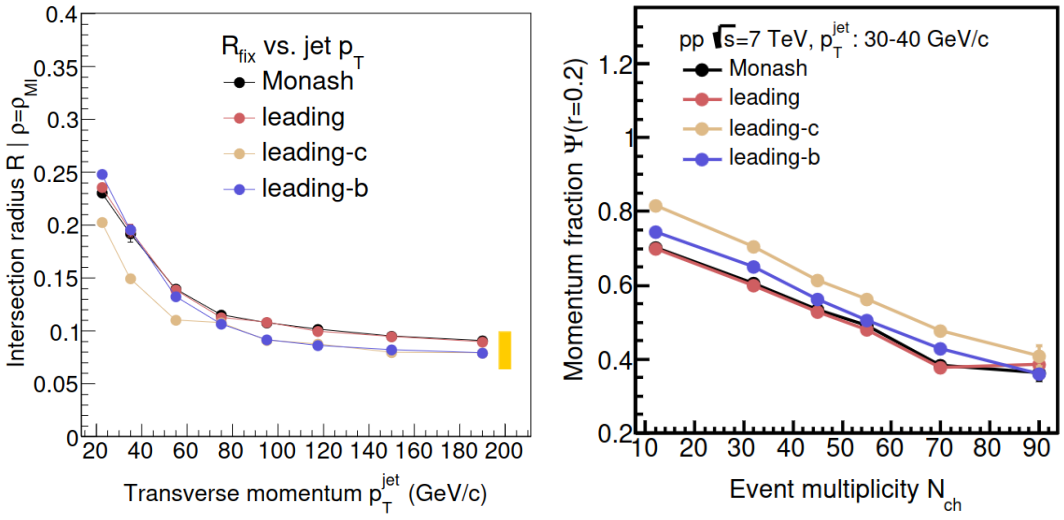

Figure 3.14: The evolution of the characteristic jet size as a function of $p_{\mathrm{T}}^{\mathrm{jet}}$ (*left*) and the integral jet structure at $r = 0.2$ (*right*) for both heavy-flavor and inclusive jets. The absolute uncertainty arising from the choice of bin width is indicated by the yellow shaded band in the *left* panel.

that r_{ch} is an inherent property of jets that is characteristic to the spatial development of the parton shower at a given momentum.

I compared the multiplicity dependence of jet structure variables for three popular PYTHIA tunes as well as different MPI and CR models in several p_{T} bins. I found that the evolution of the differential jet structure $\rho(r)$ with multiplicity significantly differs in several $p_{\mathrm{T}}^{\mathrm{jet}}$ ranges for the Monash, MonashStar, and 4C tunes. The shape of the difference is nontrivial in $p_{\mathrm{T}}^{\mathrm{jet}}$, but persistent through all tested choices of multiplicity selections. With this I demonstrated that the multiplicity-dependent analysis of jet momentum profiles can differentiate among otherwise well-established models. This lack of understanding may have grave consequences on studies based on classification by jet properties. My observation highlights the need of extending multiplicity-dependent jet structure measurements such as in [177] to higher $p_{\mathrm{T}}^{\mathrm{jet}}$ regimes. I also performed a cross-check analysis with HIJING++, using two different PDF sets, which confirmed that the characteristic jet size was not a peculiarity in PYTHIA. These observations suggest that the characteric jet size is an inherent property of the jets and is characteristic to the space-time evolution of the parton shower at a given momentum.

However, the characteristic jet size does depend on the jet flavor. Flavor-dependent jet structure studies may be a way to access mass versus color charge effects that is complimentary to hadron- or jet-production cross-section measurements. My findings motivate further phenomenology studies as well as cross-checks with real data to gain a deeper understanding of flavor-dependent jet fragmentation. Another direction for future research

could be to investigate the effects of MPI on jets without the underlying event. This could be done either by choosing an observable that depends very weakly on the underlying event [189], or by both understanding the underlying event and the fragmentation region [153]. Finally, we also saw that the integrated jet structure variable $\Psi(r = 0.2)$ shows a rather different N_{ch}-dependence when MPI are turned off. This attests to the important role of multiple-parton interactions in higher multiplicity events and the need for their detailed understanding in order to develop accurate models in jet physics.

Chapter 4

Scaling Properties of Jet Structures

In this chapter I present a study on the scaling properties of event multiplicity distributions in proton–proton collisions. First, I introduce the KNO scaling which involves the scaling of the event multiplicity distributions with the collision energy, then I present my measurements on a similar scaling phenomenon, simply referred to as "KNO-like scaling", in which the jet multiplicity distributions scale with the jet transverse momentum.

4.1 KNO-like Scaling in Jets

As we discussed in Section 2.7.3, the Koba–Nielsen–Olesen (KNO) scaling [141, 142] is the collapse of the final-state event multiplicity distributions onto a universal scaling curve. In a recent study, based on proton–proton collision data I simulated, it was found that a KNO-like scaling may be fulfilled within single jets [190]. This indicates that the KNO scaling is violated by complex quantum-chromodynamics (QCD) processes outside the jet development, such as single and double-parton scatterings or softer multiple-parton interactions (MPI).

In this section I present my study of the scaling properties of heavy-flavor (HF) jets compared to an inclusive jet sample. Heavy flavor is mostly produced in hard (large momentum-transfer) processes, in the early stages of the collision event. The most relevant perturbative QCD processes that contribute to the production cross section are leading-order (LO) flavor creation, and next-to-leading order (NLO) gluon splitting as well as flavor excitation [71]. The parton shower and fragmentation of heavy-flavor jets is different from light-flavor jets due to two main reasons: the color charge effect, i.e. heavy flavor jets are initiated by quarks as opposed to light-flavor jets that are mostly gluon-initiated [191]; and the dead-cone effect, meaning that small-angle gluon radiations off a massive parton are forbidden in QCD, and as a consequence, heavy-flavor fragmentation is harder and results in different jet substructures [35, 36, 192, 193].

In this study I modeled both light and heavy-flavor jets using the PYTHIA 8 Monte Carlo event generator [12], and differentiated the jet samples according to the process by which they were created. Whether a KNO-like scaling observed for inclusive jets is retained or violated in heavy-flavor jets, can shed light on the origin of the scaling itself. It can also provide information on the possible mechanisms which are responsible for the violation of the scaling in heavy-flavor jets. The methods I present in this chapter can further be used to gain insight into the flavor-dependent evolution of jets. Future measurements targeted on the scaling of light and heavy-flavor jets can also serve as a validation tool for heavy-flavor production and fragmentation models.

4.2 Analysis Method

I simulated proton–proton collisions at $\sqrt{s} = 7$ TeV center-of-mass energy utilizing the PYTHIA 8 (version 8.226) event generator with the Monash tune and HardQCD settings [12, 194]. PYTHIA 8 is tuned to describe both the fundamental physical observables of the leading hard process and the underlying event, and it is known to reproduce final-state multiplicities well [177, 186]. In PYTHIA 8 the hard parton scatterings and decays are simulated using LO matrix elements (ME). These are amended by initial and final state radiations, which create the parton shower (PS) in perturbative QCD calculations based on Dokshitzer-Gribov-Lipatov-Altarelli-Parisi (DGLAP) splitting kernels [12], as well as soft and hard multiple-parton interactions integrated into a single framework [181]. The hadronic final state is then produced with the Lund string fragmentation model [195].

Using the option in PYTHIA 8 to restrict event generation to certain hard processes, similarly as it was described in Sec. 3.1, I created four different datasets. As the baseline for my study, I created an inclusive-jet sample, in which any hard QCD scattering process was allowed above an appropriately selected value of the minimal momentum transfer in the hardest process ($\hat{p_{\mathrm{T}}}$), depending on the jet transverse momentum ($p_{\mathrm{T}}^{\mathrm{jet}}$), as detailed in [187]. As a next step, I created samples with ME flavor creation, where hard $2 \to 2$ parton scatterings were allowed only with heavy-flavor outgoing partons: $gg \to b\bar{b}(c\bar{c})$ and $q\bar{q} \to b\bar{b}(c\bar{c})$. This provided wide-angle heavy-flavor jets created directly in the leading process of the event. Finally, I created a sample that is dominated by beauty jets from the PS, by allowing only those $2 \to 2$ processes that do not directly create heavy flavor: $gg \to gg$, $gg \to q\bar{q}$, $qg \to qg$, $q\bar{q} \to gg$, $q\bar{q} \to q'\bar{q'}$ (where incoming heavy flavor is allowed, but only light flavor exits), and finally three more processes: $qq' \to qq'$, $q\bar{q'} \to q\bar{q'}$ and $\bar{q}\bar{q'} \to \bar{q}\bar{q'}$ (where outgoing and incoming flavors are the same and q and q' may be of the same flavor) [196]. In this case the heavy quark pair is produced in a later step, e.g. in a $g \to b\bar{b}$ gluon splitting process, typically with smaller opening angles. The heavy quarks then often manifest in the final state as secondary jets besides the leading jet, or may even end up in the same jet.

In all cases, charged-particle jets were clusterized from final-state charged pions, kaons and (anti)protons with $p_{\mathrm{T}} > 0.15$ GeV/c using the anti-k_{T} jet-clustering algorithm [40] with a resolution parameter of $R = 0.7$ in the mid-rapidity range $|\eta| < 1$ and full azimuth coverage. The reconstructed jets were categorized in 20 different $p_{\mathrm{T}}^{\mathrm{jet}}$ ranges, from 15 GeV/c up to 400 GeV/c. In the case of the charm and beauty jet samples, the corresponding heavy quark was required to fall within the cone of the selected jet, similarly to jet-tagging methods that are utilized in experiments [92, 197].

4.3 Results

In Fig. 4.1, I plot the mean and the root mean square (RMS) values of the event multiplicity (N_{ch}) distributions at central pseudorapidity ($|\eta| < 1$), as a function of $p_{\mathrm{T}}^{\mathrm{jet}}$, separately for inclusive jets, beauty jets and charm jets from ME flavor creation as well as for beauty jets from parton shower processes. As one expects, events having jets with a higher $p_{\mathrm{T}}^{\mathrm{jet}}$ contain more final-state hadrons on the average, and the distribution also gets broader toward higher $p_{\mathrm{T}}^{\mathrm{jet}}$. Heavy-flavor jets from ME flavor creation correspond to a lower average multiplicity at a given $p_{\mathrm{T}}^{\mathrm{jet}}$, while heavy-flavor from the parton shower follows the trend of inclusive jets. The difference is especially prominent for higher $p_{\mathrm{T}}^{\mathrm{jet}}$.

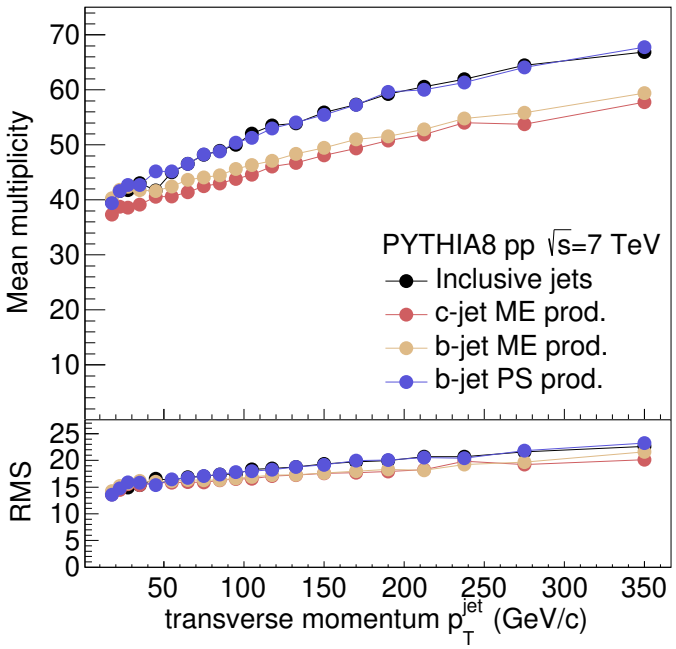

Figure 4.1: The mean (top panel) and RMS values (bottom panel) of the charged-hadron multiplicity distributions at $|\eta| < 1$ for inclusive jets, for charm and beauty jets from ME-level production as well as for beauty jets from production in the PS, as a function of $p_{\mathrm{T}}^{\mathrm{jet}}$.

As a next step, I fitted the multiplicity distributions with a negative binomial distribution (NBD) function in each of the jet transverse momentum ranges,

$$P_{N_{\rm ch}} = \frac{\Gamma(N_{\rm ch}k + a)}{\Gamma(a)\Gamma(N_{\rm ch}k + 1)} p^{N_{\rm ch}k}(1-p)^a, \tag{4.1}$$

where a, k and p are parameters related to the mean and dispersion of the distribution of the multiplicity $N_{\rm ch}$. In Fig. 4.2, I show the multiplicity distributions after all the $p_{\rm T}^{\rm jet}$ ranges have been scaled on top of each other using the NBD fits. The scaling approximately holds for all four jet samples which I investigated. However, for jets containing charm or beauty from flavor creation, the data show minor departures from the NBD fits: the distribution is wider for larger $p_{\rm T}^{\rm jet}$, while narrower for smaller $p_{\rm T}^{\rm jet}$ values.

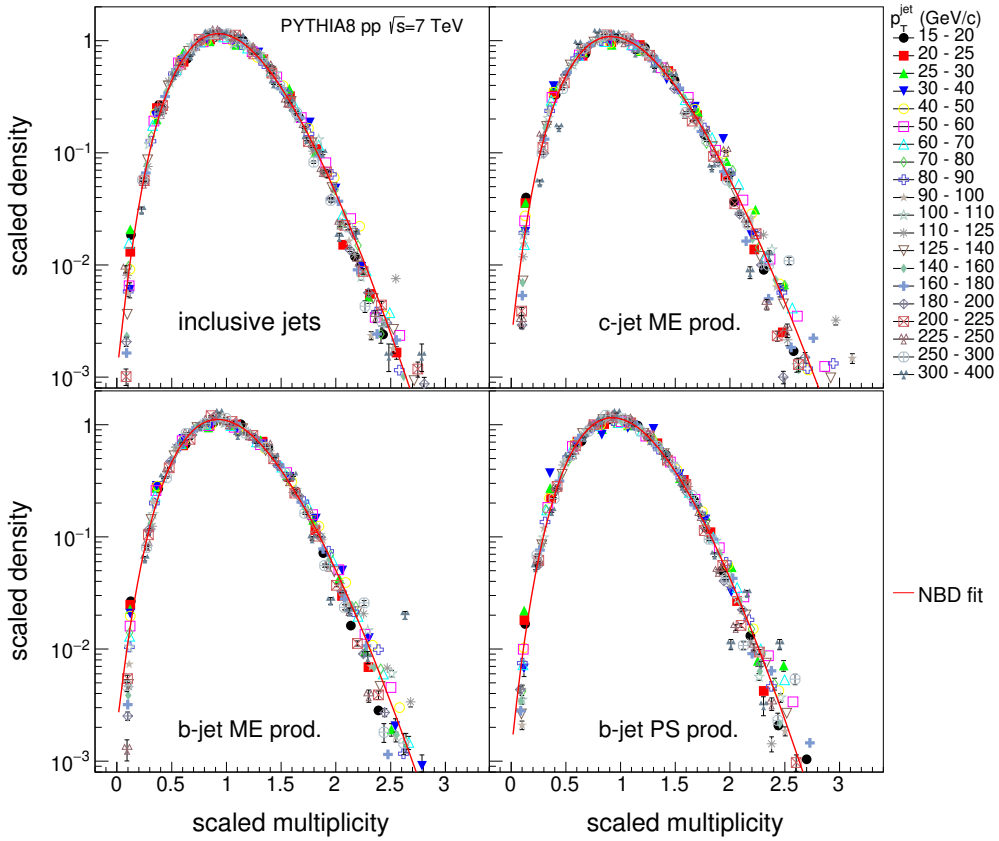

Figure 4.2: Charged-hadron multiplicity distributions with an NBD fit at $|\eta| < 1$, for all $p_{\rm T}^{\rm jet}$ ranges, scaled by the NBD fit means. The four panels from top left to bottom right correspond to inclusive jets, charm and beauty jets from ME flavor creation, and beauty jets from PS production.

To quantify the deviations from the scaling behavior, and also to mitigate the effect of fluctuations, I calculated the higher moments of the multiplicity distributions in a similar manner to [190]. Here the q^{th} moment in a given $p_{\mathrm{T}}^{\mathrm{jet}}$ window is defined as

$$\langle N_{\mathrm{ch}}^{q} \rangle = \sum_{N_{\mathrm{ch}}=1}^{+\infty} P_{N_{\mathrm{ch}}} N_{\mathrm{ch}}^{q}, \tag{4.2}$$

where $P_{N_{\mathrm{ch}}}$ is the probability distribution corresponding to the event multiplicity N_{ch}. If the scaling is fulfilled and the mean of the distribution scales with a factor λ, then it is expected that the q^{th} moment scales with λ^q as

$$\langle N_{\mathrm{ch}}^{q}(p_{\mathrm{T}}^{\mathrm{jet}}) \rangle = \lambda^{q}(p_{\mathrm{T}}^{\mathrm{jet}}) \langle N_{\mathrm{ch}}^{q}(p_0) \rangle, \tag{4.3}$$

where p_0 is chosen so that the scaling factor is $\lambda(p_0) = 1$.

In Fig. 4.3, I plotted on a log-log scale the first nine statistical moments of the multiplicity distributions divided by the order q, as a function of the mean charged-hadron multiplicity $\langle N_{\mathrm{ch}} \rangle$ at $|\eta| < 1$. The four panels correspond to the four jet categories. The linear fits show a similar trend for all four cases, which means that the scaling is present also for heavy-flavor jets.

Fig. 4.4 summarizes the slopes of the fits for the first nine statistical moments, as well as the goodness-of-fit parameter χ^2/NDF. All slopes are around unity within ≈5%. As expected, the goodness of the linear fits is worse for higher moments. The beauty jets coming from parton shower processes follow the same trend as the inclusive jets. On the other hand, heavy-flavor jets coming from matrix element production in the simulations correspond to slope parameters that are slightly but significantly different from unity: for charm jets, slopes of the fits for moments $2 \leq q \leq 6$ tend to be lower than unity, while for beauty jets the slopes for moments $q \geq 7$ are larger than unity. Furthermore, the goodness of the fits for heavy-flavor ME production tends to be worse than for inclusive jets, $\chi^2/\mathrm{NDF} \geq 10$ for any $q \geq 5$. This suggests that the KNO-like scaling originates from the hard parton production, and it is less influenced by the parton shower. The similar patterns of the inclusive jets and the beauty jets from PS also indicate that event multiplicities are not driven by flavor-dependent jet fragmentation processes.

4.4 Summary

In this section I summarize my results on the scaling properties of heavy-flavor jets from different production processes, and compare them to those of inclusive jets. I used PYTHIA 8 simulations to evaluate the charged-hadron event multiplicities at central pseudorapidity, as a function of the charged-particle jet transverse momentum within the range $15 < p_{\mathrm{T}}^{\mathrm{jet}} < 400$ GeV/c.

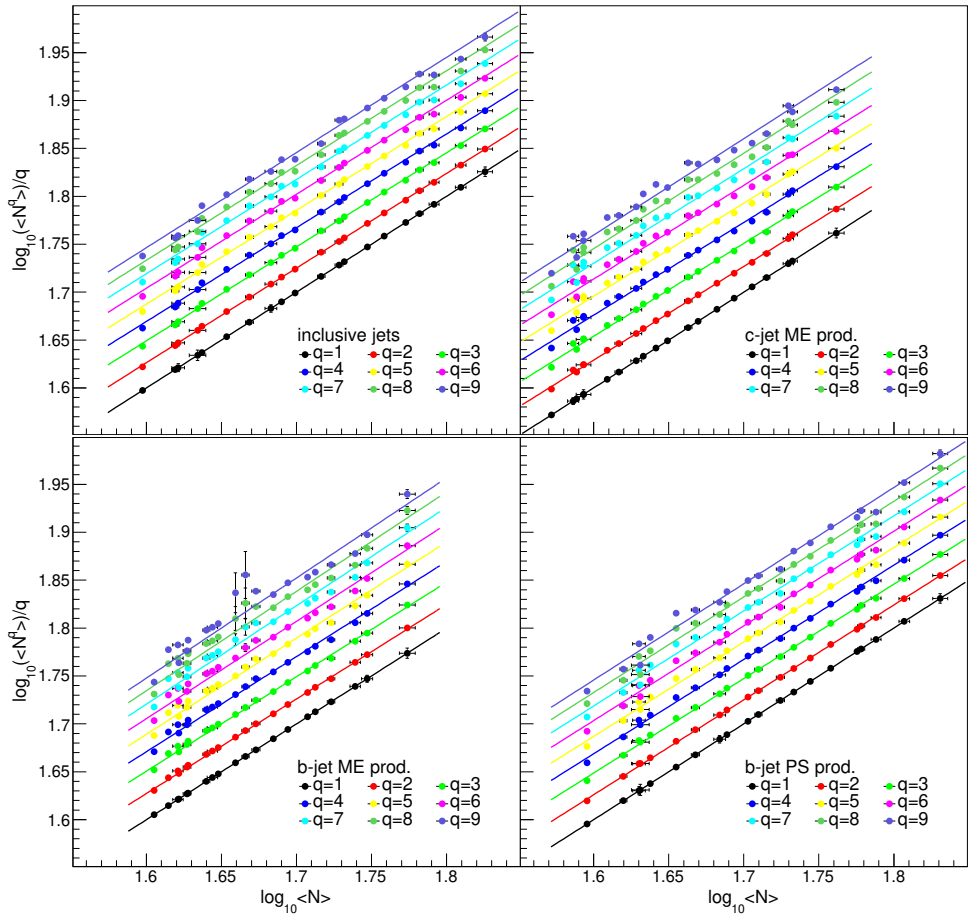

Figure 4.3: The first nine moments of the charged-hadron multiplicity distributions at $|\eta| < 1$, as a function of the average multiplicity corresponding to each $p_{\mathrm{T}}^{\mathrm{jet}}$ range. The four panels are for inclusive jets, charm and beauty jets from ME flavor creation, and beauty jets from PS production. The distributions are normalized by their order q on a log-log scale, and linear fits are applied.

I found that the multiplicity distributions satisfy a KNO-like scaling with $p_{\mathrm{T}}^{\mathrm{jet}}$ for charm and beauty jets similarly to what has been observed for inclusive jets. I note, however, that multiplicity distributions in events with jets initiated by charm and beauty directly from the leading hard process show some departure from the negative binomial shape, depending on the $p_{\mathrm{T}}^{\mathrm{jet}}$. Further analysis of the statistical moments of the multiplicity distributions shows that the scaling is fulfilled within ≈5% throughout the full $p_{\mathrm{T}}^{\mathrm{jet}}$ range, but the deviations are more significant for leading-order heavy flavor creation, especially in the case of beauty. On the other hand, beauty production from the parton shower tends to deviate less from scaling expectations and follows the inclusive-jet trend within uncertainties. I conclude therefore that the KNO-like scaling originates from the parton level of the early stages of

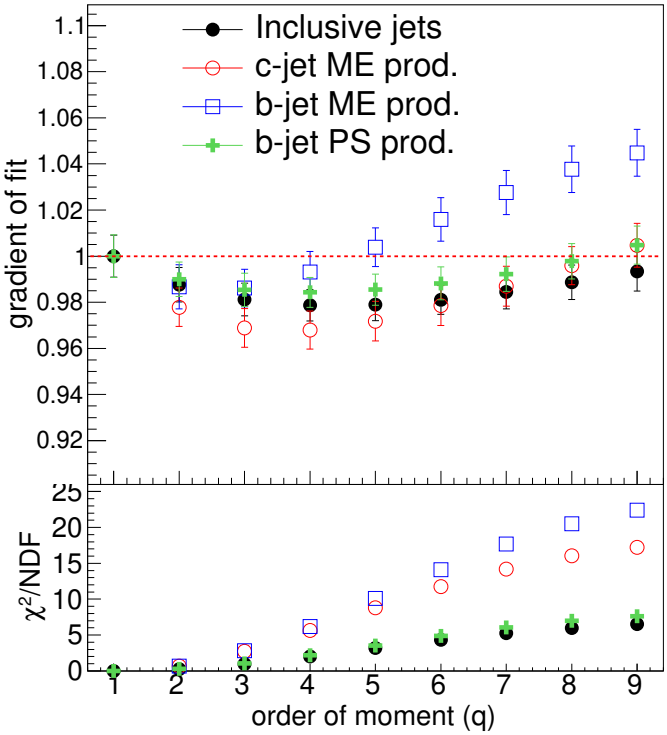

Figure 4.4: The slope parameters (top panel) and the goodness-of-fit parameters χ^2/NDF (bottom panel) of the linear fits for the first nine statistical moments of the multiplicity distributions, for charm and beauty jets from ME production and beauty jets from PS production, compared to that for inclusive jets, as a function of the order of moments of the multiplicity distributions.

the collision, and not from the later stages of parton shower or jet fragmentation.

A good description of hadron multiplicity distributions is a basic requirement for models and it is generally fulfilled by the most widely used event generators. However, multiplicities as a function of the jet momentum for jets tagged with different flavors can provide means to further validate heavy-flavor production and fragmentation models. Also, while event multiplicity is a good proxy for jet multiplicity in case of jets coming from the leading hard process, this is not necessarily the case for jets that come from secondary hard processes or gluon radiation. An interesting extension of the current work in this direction is therefore to evaluate the scaling in terms of the jet multiplicity instead of the event multiplicity, and to see whether in that case the scaling of heavy flavor jets from the parton shower follows light or heavy jets.

Chapter 5

Jet Multiplicity Distributions in pp Collisions at the ALICE Experiment

As we discussed in Section 2.7.3, the Koba–Nielsen–Olesen (KNO) scaling [141, 142], i.e. the collapse of the final-state event multiplicity distributions onto a universal scaling curve, was first proposed to be present in collision events, but was later found to be violated beyond a certain energy threshold and also in hadronic collision systems [143, 144]. In Section 4, I presented my research on the KNO-like scaling in proton–proton collisions utilizing PYTHIA 8 simulations.

Neither the origin of the KNO scaling, nor the mechanism that violates it is yet completely understood, although several scenarios have been suggested [147–151, 198]. A recent study based on proton–proton collision data that I simulated found a KNO-like scaling within the jets, and concluded that the KNO scaling is violated by processes outside the jet development such as single and double-parton scatterings or softer multiple-parton interactions [190]. In a later simulation study, which I presented in Chapter 4, I used heavy-flavor jets to show that the KNO-like scaling originates from the partonic level [193]. These findings, however, are not yet confirmed experimentally.

In this chapter I present the first measurement of the jet multiplicity distributions as a function of jet transverse momentum in pp collisions at $\sqrt{s} = 13\ \mathrm{TeV}$ in the ALICE experiment. I also present the first time measurement of the KNO-like scaling of the measured jet multiplicity distributions. The results will allow for the verification of intra-jet KNO-like scaling or for the observation of its violation.

The KNO scaling within jets further brings up a lot of interesting questions regarding the hadronization and fragmentation processes: What is the exact origin of the scaling? How does it depend on the collision system or the parton the jet originates from? and many more... In this data analysis I hope to provide valuable information to help answer these questions and to motivate future research.

5.1 Experimental Setup and Data Sets

A complete description of the ALICE experimental setup and its performance can be found in [156, 157], and a detailed description was given in Sec. 2.8.2. For particle tracking and identification, the ALICE central barrel detectors located within a solenoid magnet ($B = 0.5$ T) were used. These include the Inner Tracking System (ITS), the Time Projection Chamber (TPC), and the Time-Of-Flight detector (TOF). The ITS tracks charged particles within the pseudorapidity interval $|\eta| < 0.9$ and also helps in the primary and secondary vertex reconstruction. The TPC is the main tracking device. The V0 scintillating detector covers the ranges of $-3.7 < \eta < -1.7$ and $2.8 < \eta < 5.1$ and is used for the online trigger and offline event selection. A minimum-bias (MB) trigger was used to collect the data sample, requiring a signal above a given threshold in both V0 counters. Events were selected offline, using the timing information from the V0 with the correlation between the number of hits and track segments in the first two layers of the ITS, to remove beam-gas interactions. Only events with a primary vertex reconstructed within ± 10 cm from the center of the detector along the beamline were used for the analysis. The analysed data sample consists of about 1.7 billion minimum-bias events from pp collisions at $\sqrt{s} = 13$ TeV corresponding to an integrated luminosity of $L_{\mathrm{int}} = 22.55\,\mathrm{nb}^{-1}$, collected during the years 2016, 2017 and 2018. The Monte Carlo samples are ALICE GEANT3 MC simulations with pp events generated by PYTHIA 8 using Monash tune. These MC samples are pthard-binned jet-jet productions anchored to the 2016, 2017 and 2018 datasets.

5.2 Analysis Method

The analysis procedure is briefly the following. I extract the jet multiplicity distributions as a function of the jet transverse momenta, so that I can investigate whether the KNO-like scaling holds (as in Chapter 4), i.e. the jet multiplicity distributions can be all collapsed onto a universal scaling curve. Before the multiplicity distributions are fitted for the scaling, I correct them for the detector effects (called unfolding), and calculate the systematic uncertainties. The unfolded multiplicity distributions with the systematic errors are then investigated for the scaling by a fitting procedure. I present these steps in the remainder of this chapter.

5.2.1 Event and Track Selection

For my analysis minimum bias events were selected by the ALICE MB trigger, which requires coincidence in the V0A and V0C detectors. Events are accepted only if the vertex position on the beam-axis falls within ($|z_{vertex}| < 10$ cm) of the nominal interaction point (IP), as can be seen in Fig. 5.1.

Tracks are only accepted if they fall within ALICE acceptance, i.e. if they satisfy the $|\eta_{\mathrm{trk}}| < 0.9$ and $p_{\mathrm{T}} > 0.15$ GeV$/c$ constraints. In Fig. 5.2 I show the control plot for the transverse momentum distribution together with the distribution for the number of tracks per event in Fig. 5.3, where the vertical axis for both plots was set to be logarithmic. In Fig. 5.4 and Fig. 5.5 I show the control plots for the η and ϕ distributions of the accepted tracks, respectively.

Since MC samples are pthard-binned jet-jet productions, I had to scale the different samples one by one and merge them. In Fig. 5.6 I show the merged p_{T} hard bin spectrum as a control plot, which shows that the merged distribution looks as expected, i.e. no discontinuities or jumps are visible.

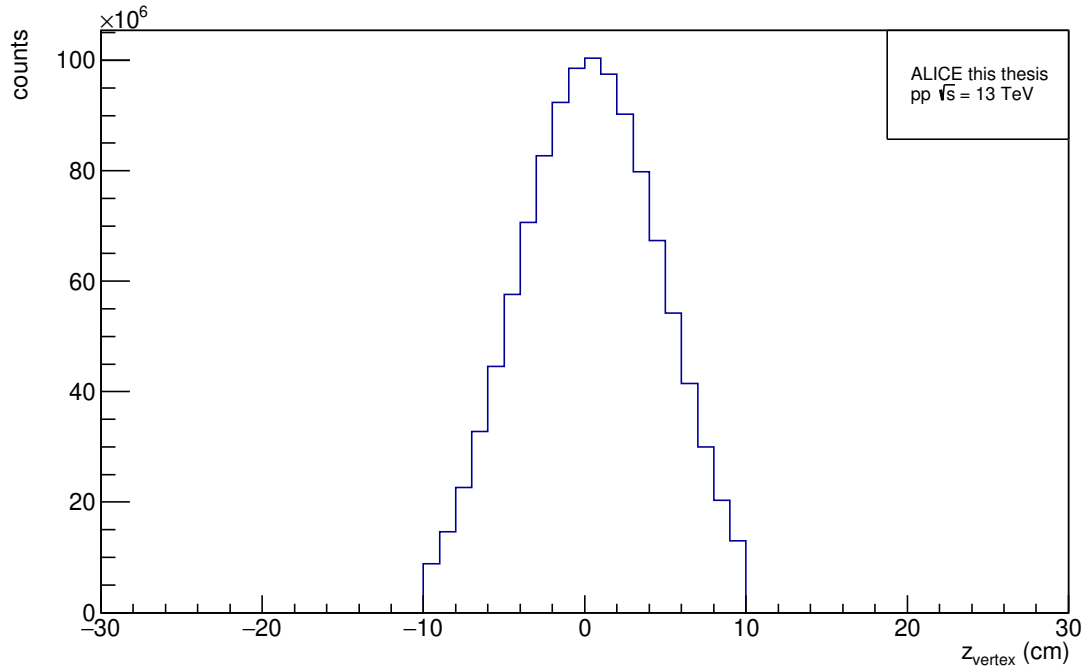

Figure 5.1: The distribution of z_{vertex}.

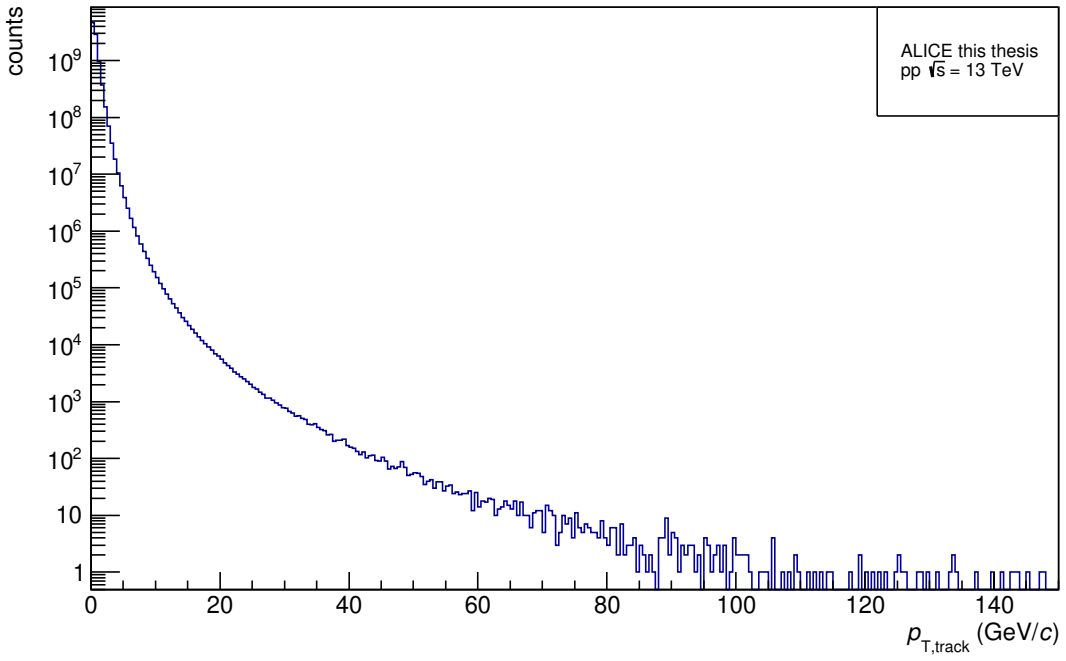

Figure 5.2: The p_{T} distribution of tracks.

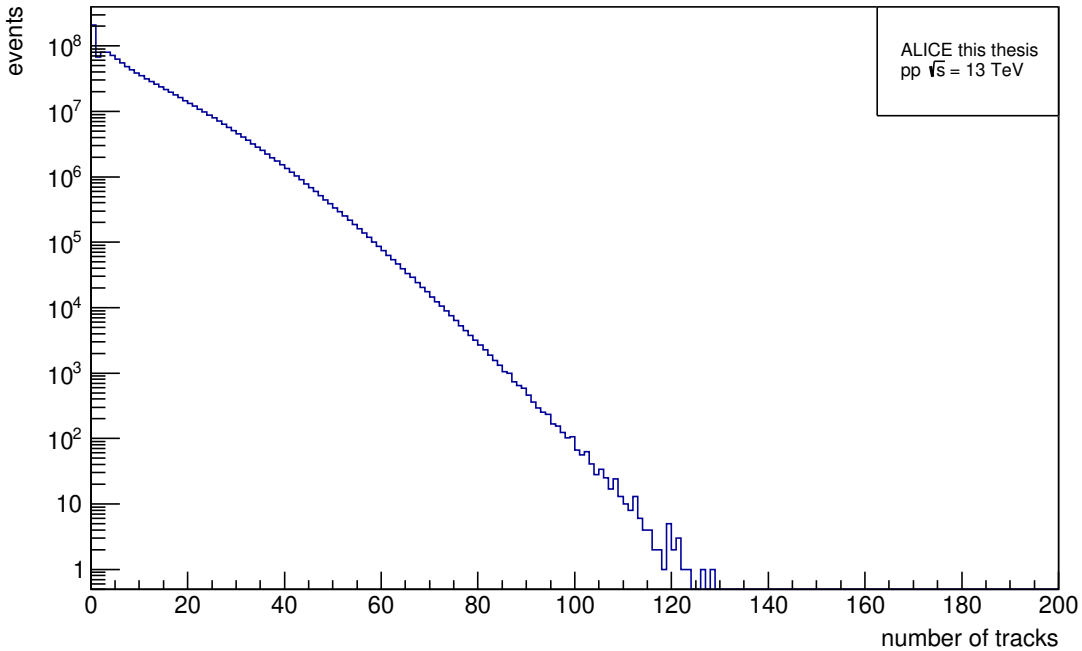

Figure 5.3: The distribution of the number of tracks per event.

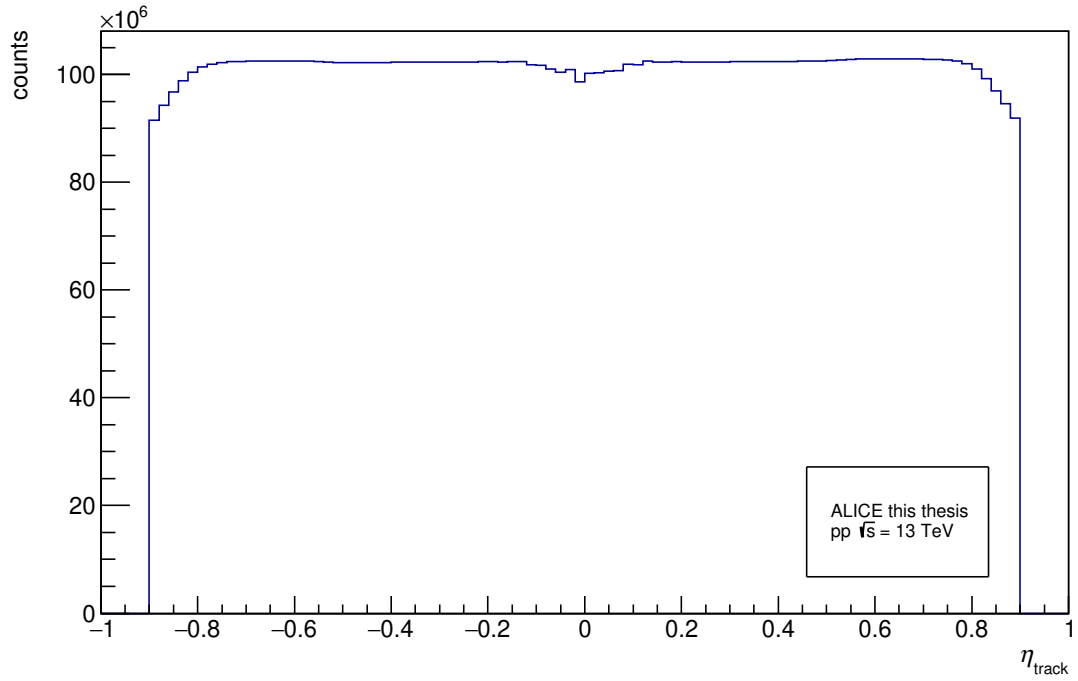

Figure 5.4: The η distribution of tracks.

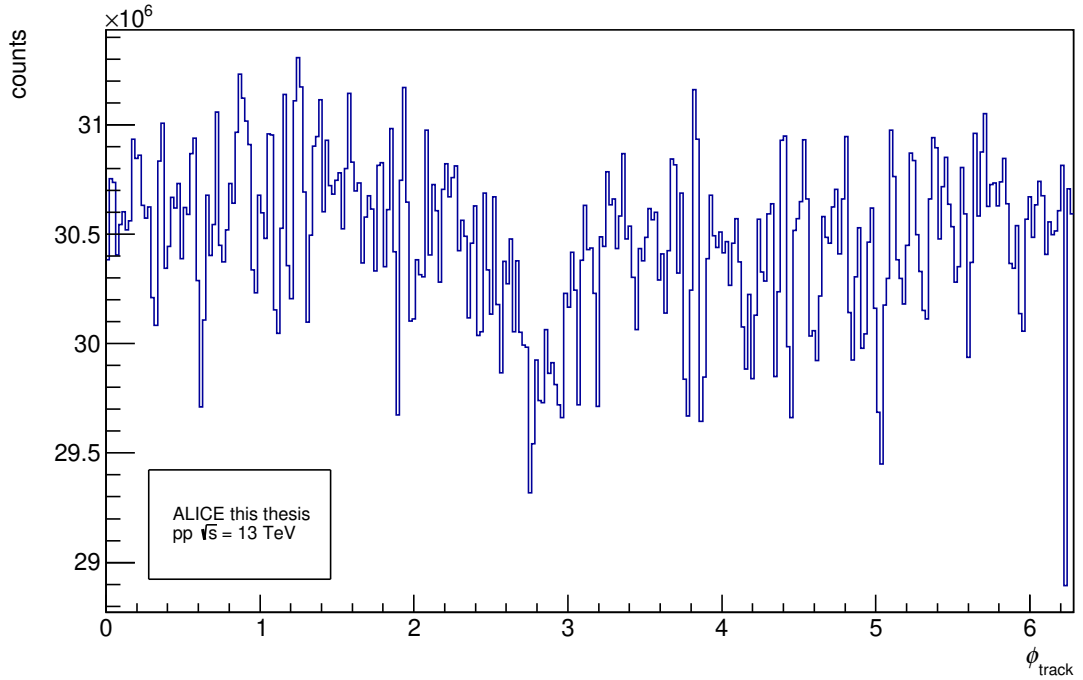

Figure 5.5: The ϕ distribution of tracks.

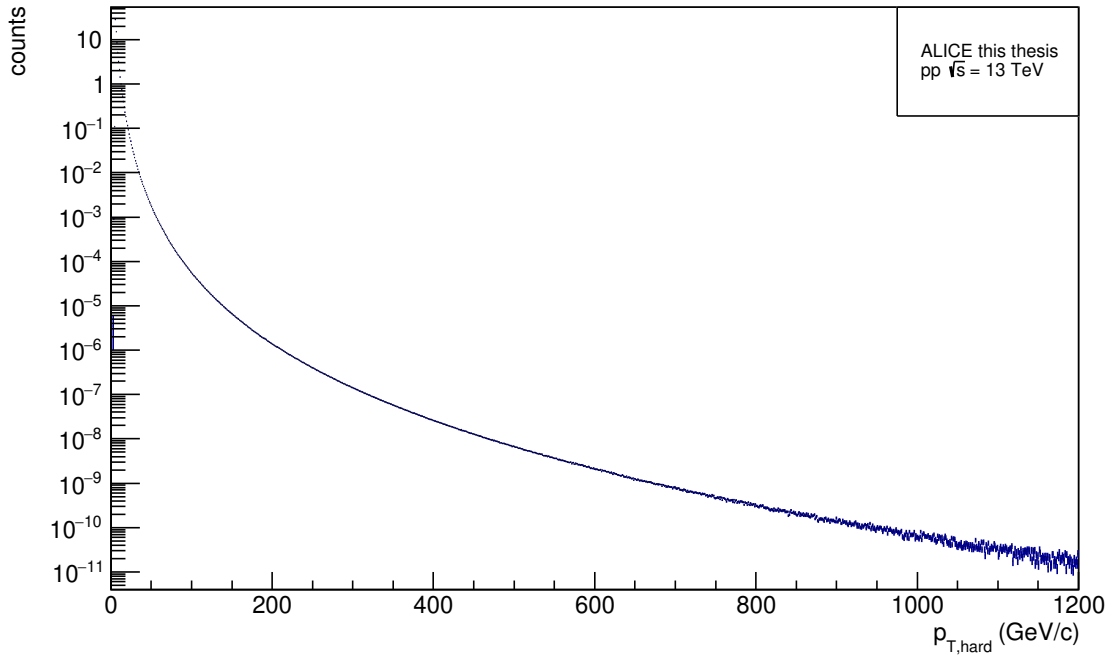

Figure 5.6: The transverse momentum distributions for the p_{T} hard bins scaled and merged together.

5.2.2 Jet Reconstruction

Charged jets with size $R = 0.4$ were reconstructed using the anti-k_{T} algorithm. The minimum charged track p_{T} is 0.15 GeV/c, and the track acceptance is within the TPC acceptance ($|\eta| < 0.9$ and $0 < \phi < 2\pi$). The $p_{\mathrm{T}}^{\mathrm{jet}}$ range is between 5 GeV/c and 140 GeV/c, and the jets must be contained inside the fiducial acceptance of the TPC ($|\eta| < 0.5$ and $0 < \phi < 2\pi$), i.e. the jet axis is at least an R distance away from the edges of the TPC. In Fig. 5.7 and Fig. 5.8, I present the control plots for the most important kinematic distributions of the reconstructed jets.

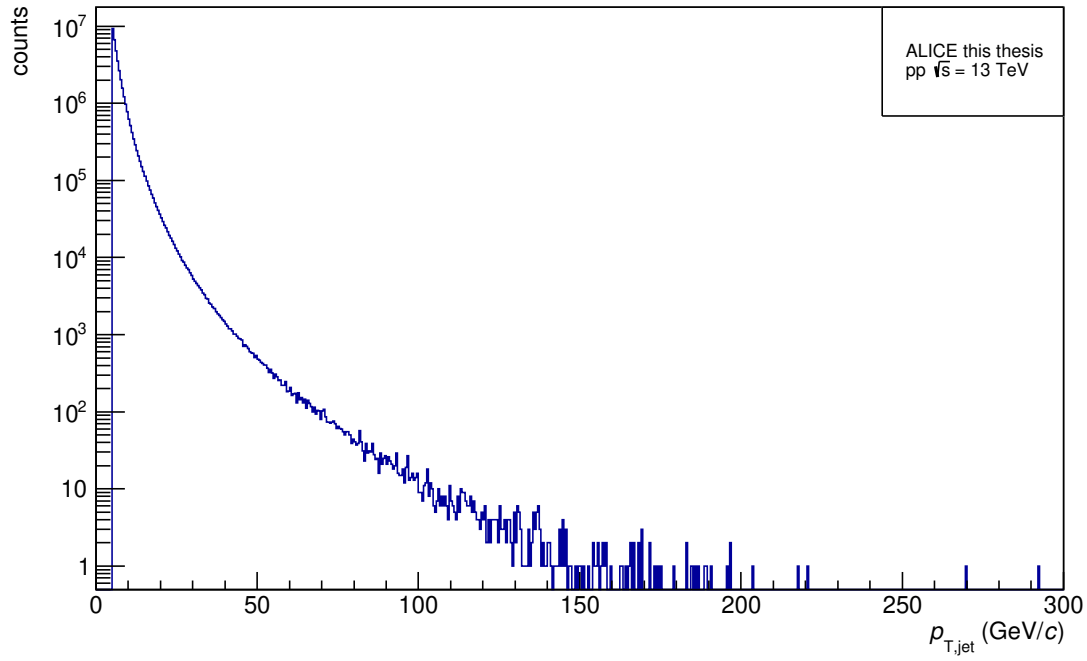

Figure 5.7: The p_{T} distribution of jets.

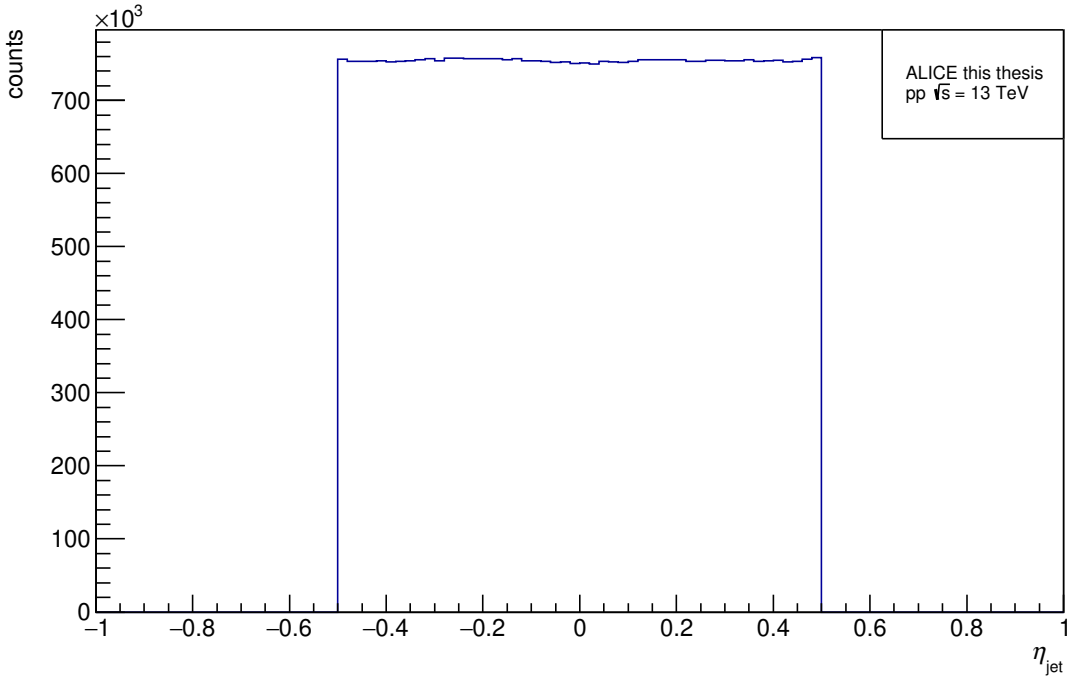

Figure 5.8: The η distribution of jets.

5.3 Unfolding

Measured data is influenced by effects such as detector resolution, inefficiencies and noise as well as background fluctuations, that smear the distributions of the observed physical variables. To correct for these effects, we use the so-called unfolding technique [199]. We use simulations to understand the smearing effects. First, "particle level" or "generator level" samples are produced as the output of Monte Carlo event generators. The generator-level samples are then fed into the GEANT 3 detector and simulation tool [200]. Then a response matrix is constructed, which is used in a deconvolution process to obtain the true physical distributions from the measured ones. I use the RooUnfold package [201] implemented in the ALICE analysis framework to do a 2-dimensional unfolding using the Bayesian method [202] in order to obtain the truth-level jet multiplicity distributions as a function of $p_{\mathrm{T}}^{\mathrm{jet}}$.

5.3.1 Response Matrix

My analysis uses a 2D unfolding, which requires the construction of a 4D response matrix. This matrix was filled with the observables $N_{\mathrm{ch}}^{\mathrm{reco}}$, $p_{\mathrm{T}}^{\mathrm{jet,reco}}$, $N_{\mathrm{ch}}^{\mathrm{true}}$, $p_{\mathrm{T}}^{\mathrm{jet,true}}$, and the following bins were used:

- Detector level jet p_{T} bins: {40, 55, 70, 85, 95, 105, 115, 125, 140} GeV
- Detector level N_{ch} bins: {6, 8, 10, 12, 14, 16, 18, 20, 22, 24}
- Particle level jet p_{T} bins: {25, 40, 55, 70, 85, 95, 105, 115, 125, 140, 200} GeV
- Particle level N_{ch} bins: {4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 50}

Since, as described in Sec. 4, the extracted jet multiplicity distributions have to be fitted with an NBD curve to investivate their scaling properties, we need a fine enough binning in the jet multiplicities. On the other hand, as it will be shown in the closure test (Sec.5.3.4), the statistics are low for high multiplicity jets in the low $p_{\mathrm{T}}^{\mathrm{jet}}$ region, and similarly, statistics are low for the low multiplicity jets in the high $p_{\mathrm{T}}^{\mathrm{jet}}$ region. Consequently, I cannot use the very low $p_{\mathrm{T}}^{\mathrm{jet}}$ regions to extract the multiplicity distributions for the physics, and I also limit the analysis to $p_{\mathrm{T}}^{\mathrm{jet}} < 140$ GeV/c. In Fig. 5.9 I show the 2D distributions of the $p_{\mathrm{T}}^{\mathrm{jet}}$ and in Fig. 5.10 the jet multiplicity for the particle and detector levels, respectively.

After the simultaneous unfolding of the jet multiplicity N_{ch} and jet transverse momentum p_{T}, Bayesian unfolding was chosen and carried out with the help of the RooUnfold package. I perform various checks on the unfolding procedure, including Monte Carlo closure test: a validation technique that ensures that known simulated samples are reproduced after applying the detector effects and then the unfolding procedure. The relevant checks are outlined in the next few sections.

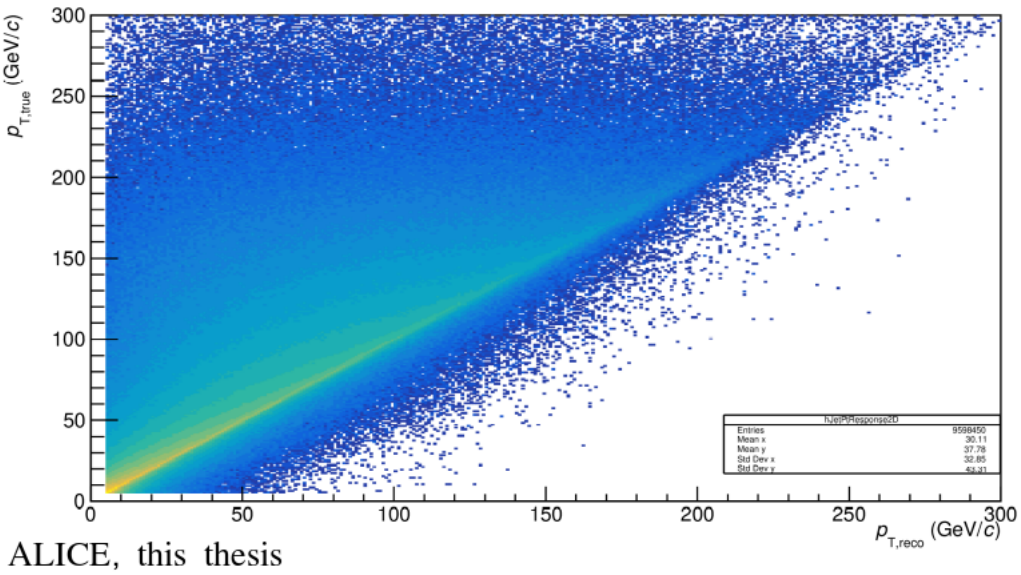

Figure 5.9: The 2D response matrix for detector and particle level $p_{\mathrm{T}}^{\mathrm{jet}}$ values.

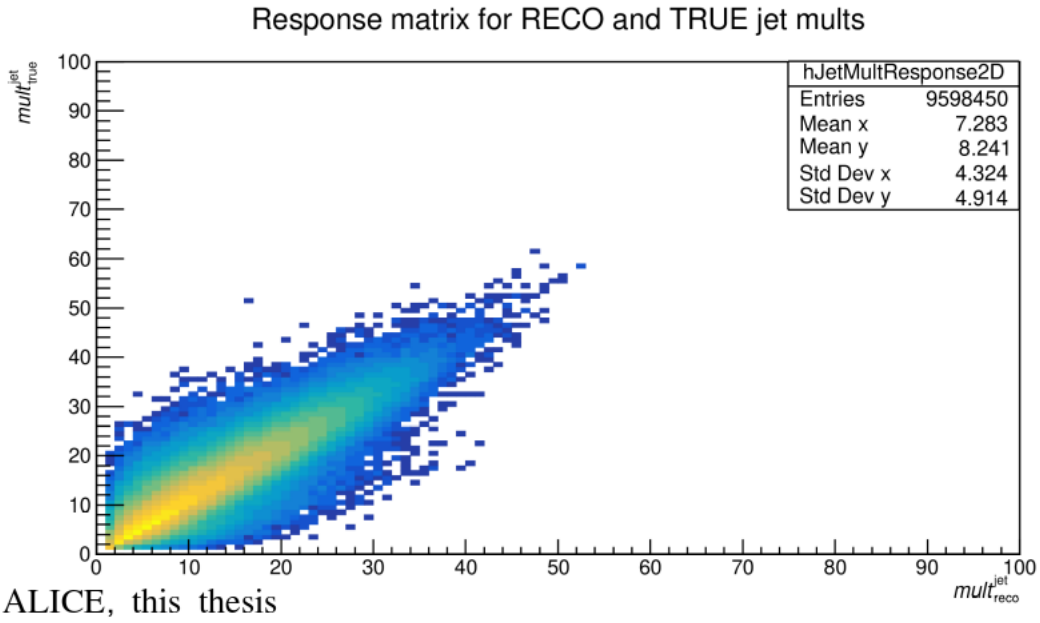

Figure 5.10: The 2D response matrix for detector and particle level jet multiplicity values.

5.3.2 Convergence

The first unfolding check is the convergence with the number of iterations. The unfolded data at each number of iterations is compared to the distribution at the next iteration to ensure that the results converge and do not vary significantly after iterating a certain number of times. I found that the unfolded 2D distribution already converges well after 3 iterations, as it is shown in Fig. 5.11, and the optimal choice of the number of iterations was found to be 3 iterations.

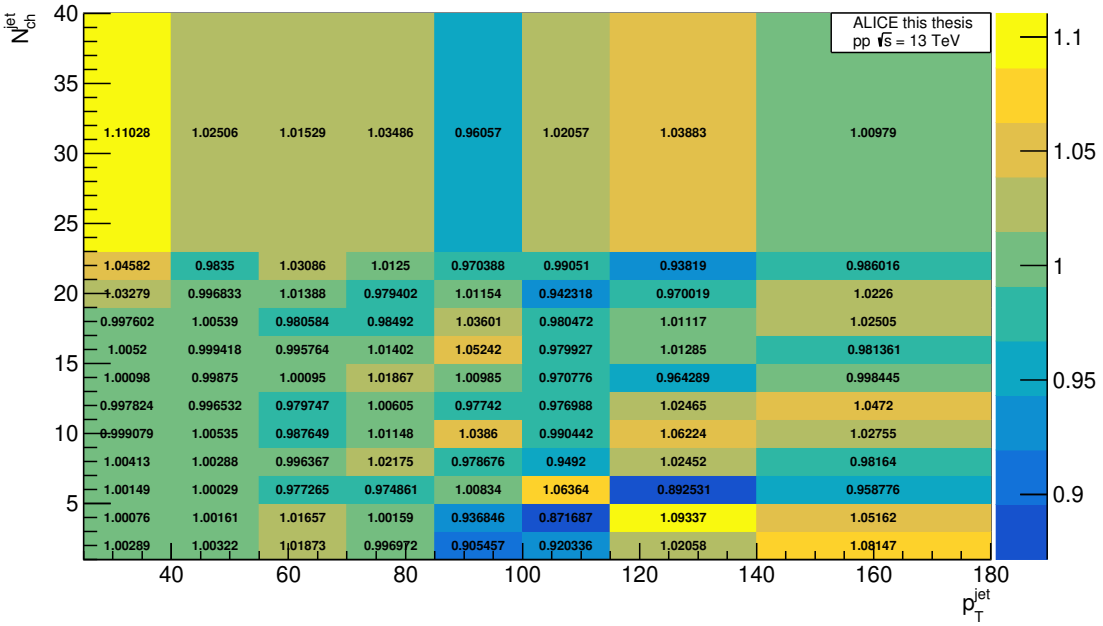

Figure 5.11: To check the convergence of the unfolding, I calculate the ratios of the unfolded 2D distributions at the 3rd and 4th iterations.

5.3.3 Refolding Test

The unfolded data is smeared with the response matrix and compared to the original raw data to check the mathematical consistency of the unfolding process. In Fig. 5.12 I show the 2D distribution of the $p_{\mathrm{T}}^{\mathrm{jet}}$ and jet multiplicity, which was used in the unfolding procedure. In Fig. 5.13 and Fig. 5.14 I show that the refolded and original distributions are in good agreement.

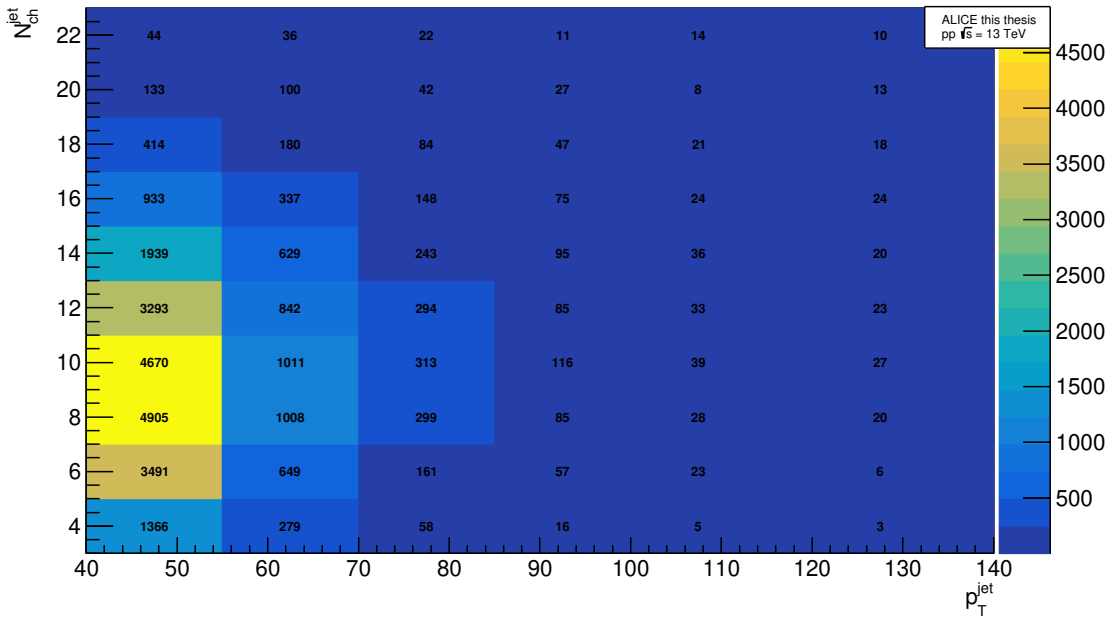

Figure 5.12: The ALICE data of the 2D $p_{\mathrm{T}}^{\mathrm{jet}}$–$N_{\mathrm{ch}}^{\mathrm{jet}}$ distribution used for the unfolding.

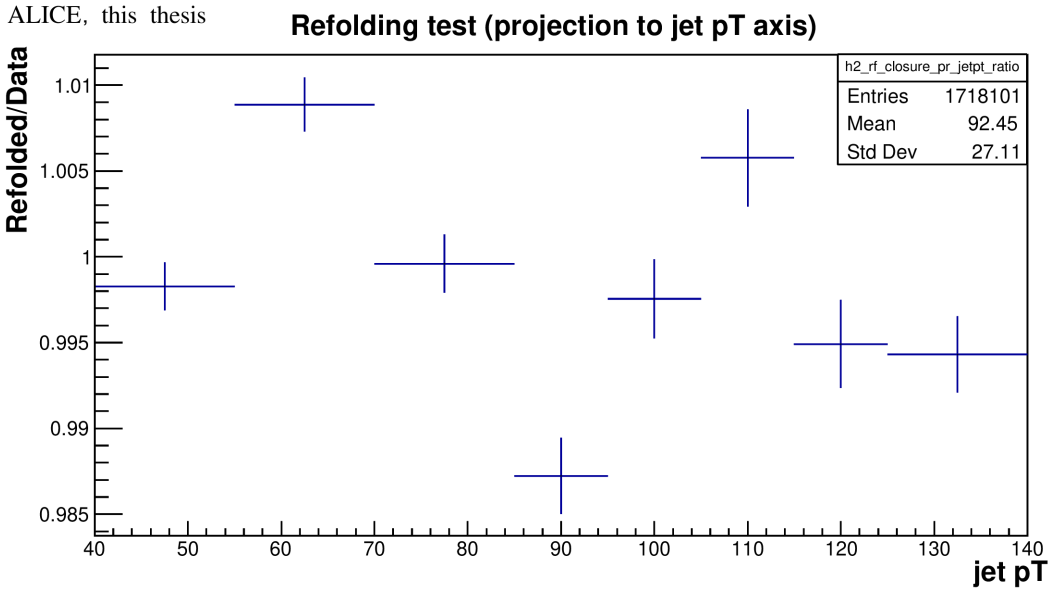

Figure 5.13: The 2D unfolded and refolded distributions are projected on the $p_{\mathrm{T}}^{\mathrm{jet}}$ axis and their ratios taken. The ratios are close to unity.

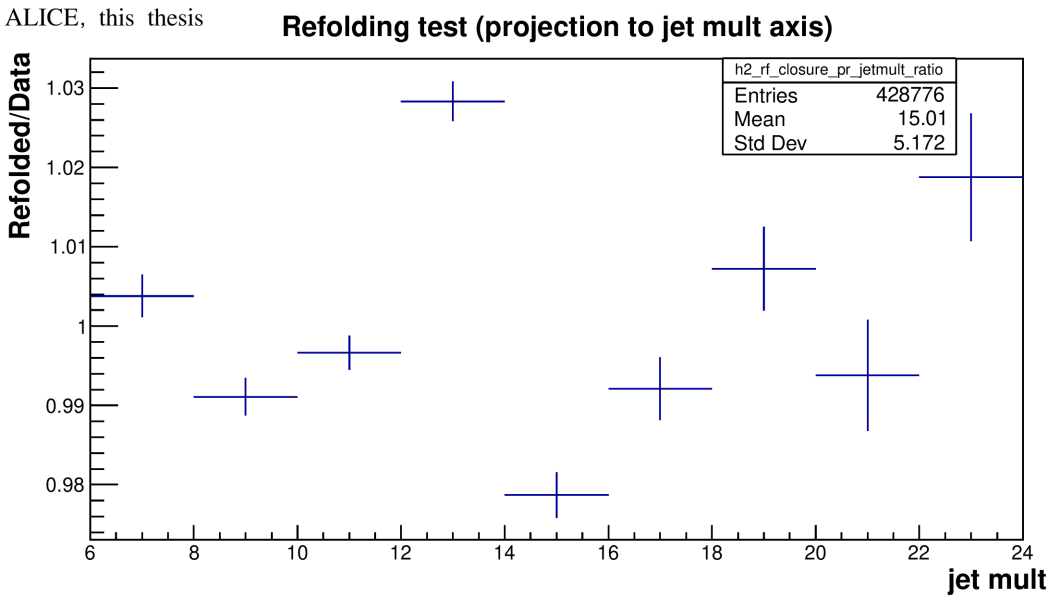

Figure 5.14: The 2D unfolded and refolded distributions are projected on the on the jet multiplicity axis and their ratios taken. The ratios are close to unity.

5.3.4 Closure Checks

A few closure checks are performed using the MC simulation to check the robustness of the unfolding procedure. An important closure test to check the unfolding procedure is the so-called split closure test where the MC sample is split such that 80% fills a response matrix and 20% fills a pseudo-data distribution with corresponding pseudo-true distribution. The pseudo-data is then unfolded as if it was real data, and the result is compared to its underlying prior or the pseudo-true distribution. In Fig. 5.15 and Fig. 5.16 I show the pseudo-data and pseudo-true distributions that we gain after the splitting of the Monte Carlo samples.

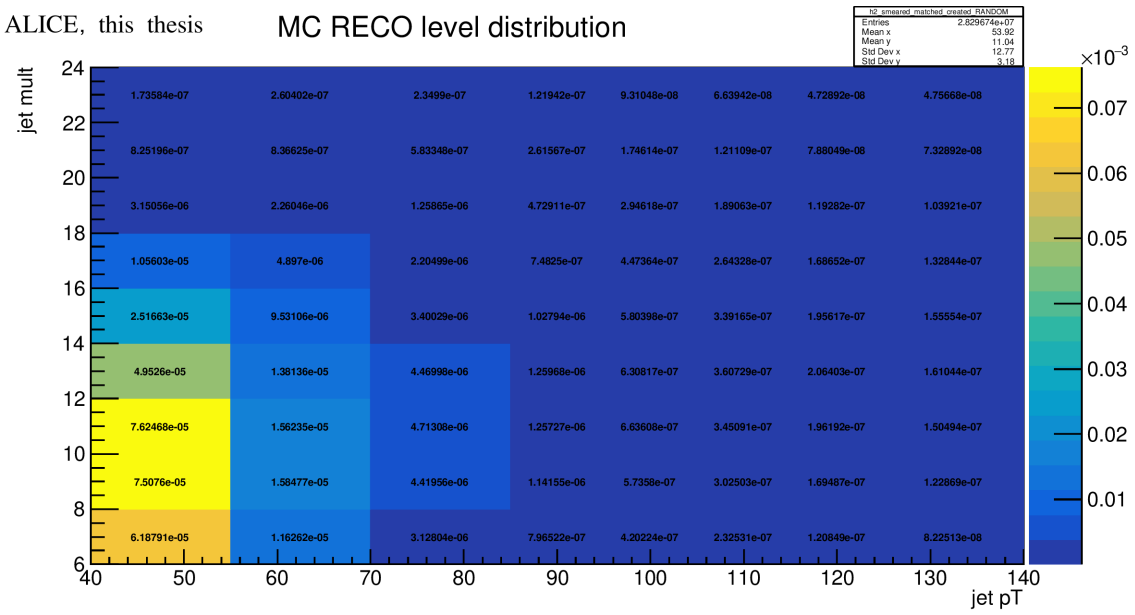

Figure 5.15: The pseudo-data distribution. The distribution was used for the split closure check.

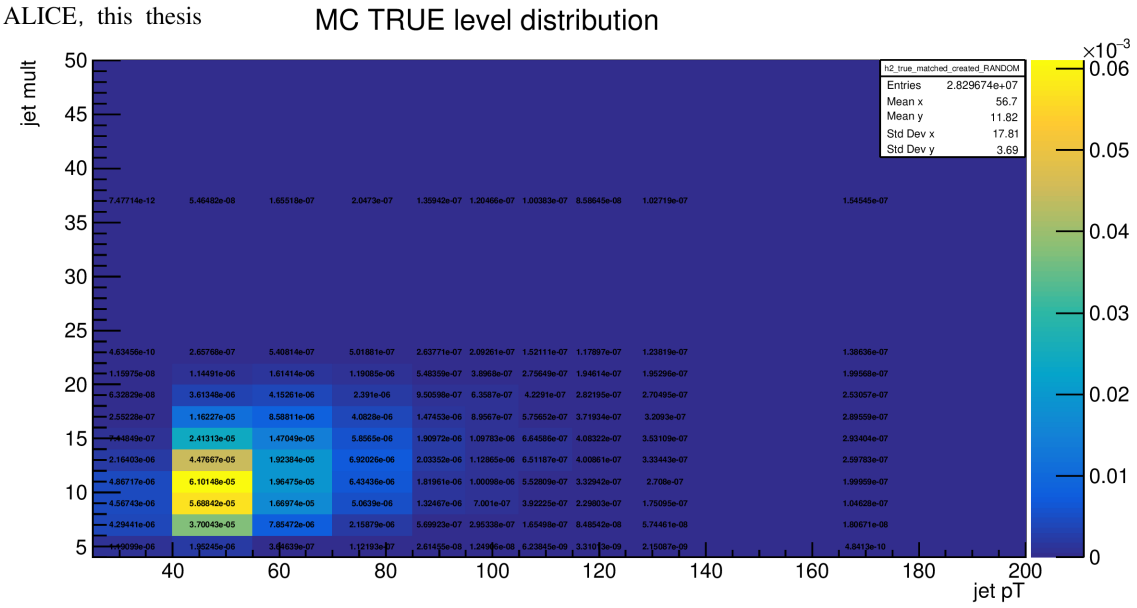

Figure 5.16: The pseudo-true distribution. The distribution was used for the split closure check.

In Fig. 5.17 I show the result of the split closure check, in which I compared the unfolded pseudo-data to the pseudo-true distribution. The comparison is done by calculating and plotting their ratio for each bin. The statistics for high-multiplicity jets at low jet p_{T} is low, similarly the statistics of low-multiplicity jets for higher jet p_{T} is also low.

This results in the closure check not producing good results for these bins. Since these bins are not relevant for the physics that we want to do (i.e. extracting the jet multiplicity distributions and fit them), I instead make sure that the closure is good for the relevant regions.

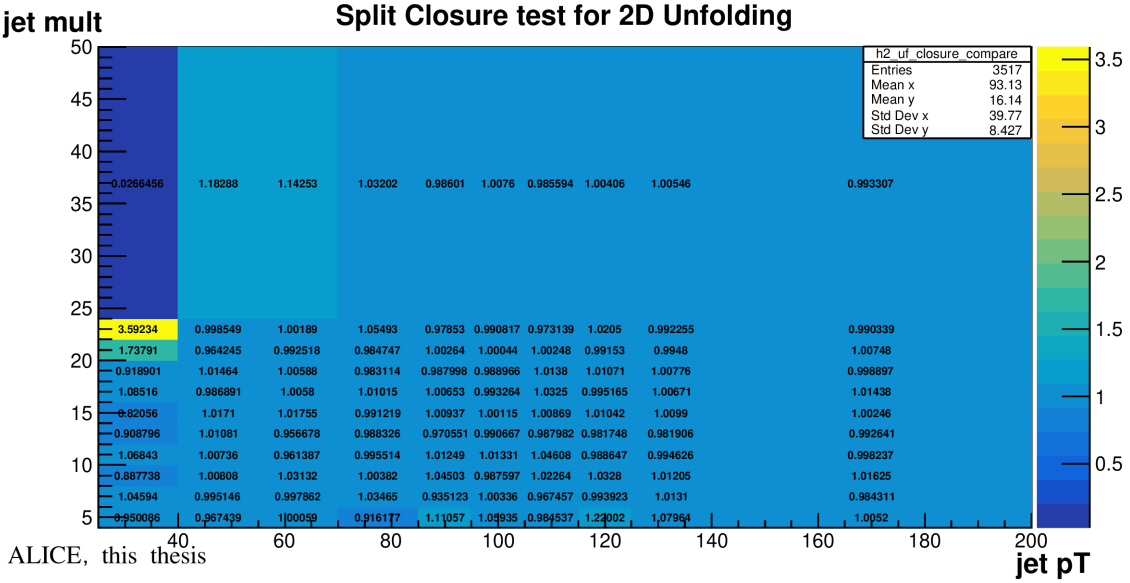

Figure 5.17: The ratio of the unfolded 2D pseudo-data and the pseudo-true distributions. In the regions which are relevant to the physics (i.e. where the jet multiplicity distributions are extracted from), the closure check works well.

5.4 Systematic Uncertainties

The unfolded jet-multiplicity distributions need to be complemented with the systematic uncertainties that affect the measurement. In the following I summarize the different sources and quantify the systematic errors.

The most significant contribution to the overall systematic error is the tracking efficiency uncertainty. For the hybrid track selection, the uncertainty on the efficiency was quantified to be 3%. To estimate the impact at the level of the observables, I degraded the detector-level PYTHIA events by randomly rejecting an extra 3% of the tracks, then repeated the jet reconstruction process on this new MC particle sample and the reconstruction of the response matrix. I then unfolded the jet-multiplicity distributions with this new response matrix instead of the nominal response, and the difference between this result and the nominal unfolded result, obtained with the nominal tracking efficiency, was used to assign a systematic uncertainty to the data.

Another source of systematic error is coming from the choice of regularization. The unfolded result calculated with 1 iteration up and 1 iteration down from the nominal one is compared to the nominal distribution, and the difference was taken as the systematic error.

The unfolding process requires a binning choice in the detector and particle level distributions. I varied the bin boundaries randomly by 5%, while leaving the edges the same. The resulting jet-multiplicity distributions were compared to the nominal one, and the difference was taken as the systematic error.

Finally, the last significant source of systematic error comes from the choice of the prior distribution. The prior represents our initial beliefs about the true distribution of a variable before considering observed data. The initial MC distributions were reweighted and smoothed out as the new choice of prior for the unfolding process. The resulting new unfolded distributions were compare to the nominal distribution and the difference was taken to be the systematic error. In Fig. 5.18 the total systematic errors, coming from all the mentioned sources, were plotted.

5.5 KNO-like Scaling in the Unfolded Jet Multiplicity Distributions

I performed a two-dimensional measurement of the jet multiplicity as a function of $p_{\mathrm{T}}^{\mathrm{jet}}$. I corrected the extracted jet multiplicity distributions for detector effects and projected the distributions for specific $p_{\mathrm{T}}^{\mathrm{jet}}$ ranges (see Sec. 5.3), so I can fit them with an NBD curve and investigate the presence of the KNO-like scaling. In Fig. 5.19, I plot the scaled multiplicity distributions and show that they indeed can be collapsed onto a universal NBD distribution.

In Fig. 5.20, I present the first nine statistical moments of the unfolded jet-multiplicity distributions, calculated similarly as it was done in my simulation analysis [193]. The statistical errors are large for the higher $p_{\mathrm{T}}^{\mathrm{jet}}$ ranges, but we can still make a linear fit. To make the gradients more visible, similarly as in Chapter 4, the gradients are plotted separately for the first nine moments in Fig. 5.21. There is an approximately 20% discrepancy toward higher statistical moments. This might indicate that the KNO-like scaling was seen in simulations only because of the specific multiplicity distributions of the particular applied model, or that the experimentally investigated energy range is not high enough. It is to be noted however, that the current data is not completely corrected: the background subtraction is yet to be implemented. This may influence the observed multiplicity distributions.

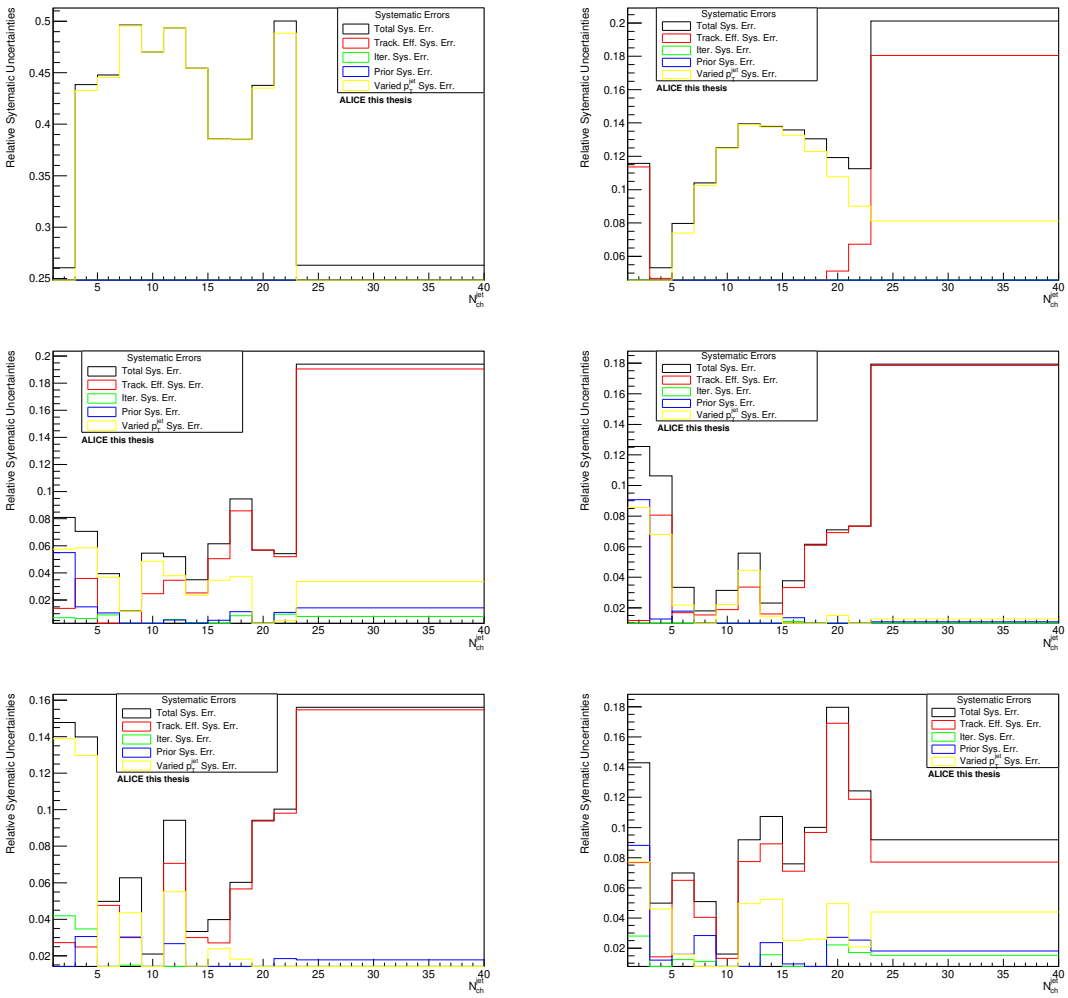

Figure 5.18: The total systematic errors shown for all significant sources, as a function of the jet multiplicity. Every figure is for a different $p_{\mathrm{T}}^{\mathrm{jet}}$ range.

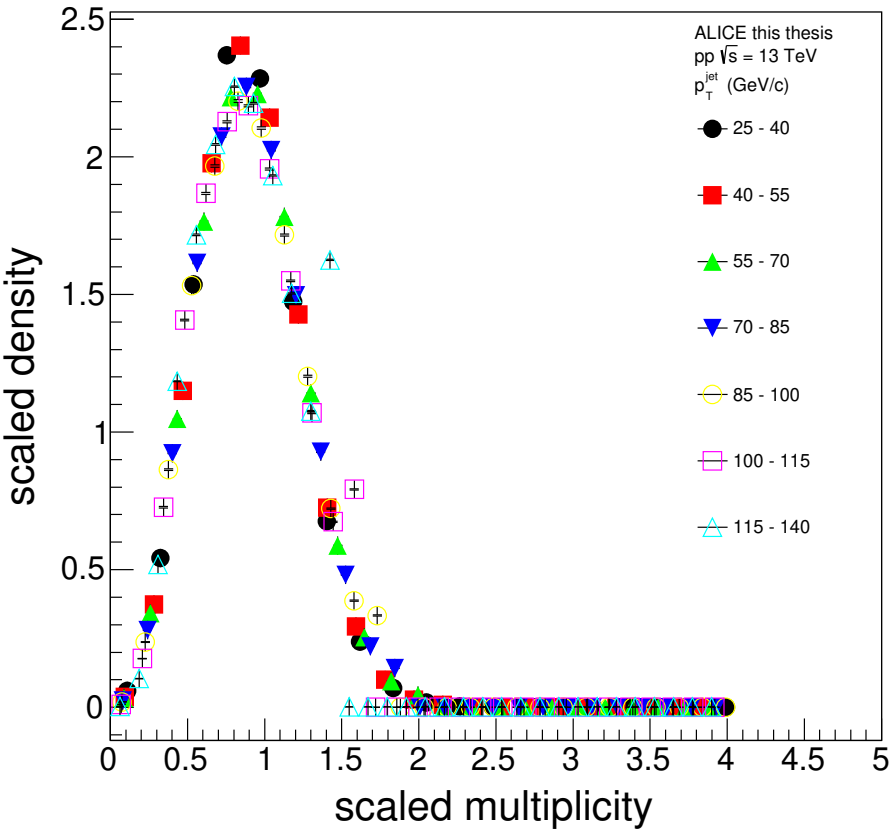

Figure 5.19: The unfolded charged jet multiplicity distributions for different $p_{\mathrm{T}}^{\mathrm{jet}}$ ranges, all collapsed onto a universal scaling curve.

5.6 Summary

In this chapter I presented the measurement of the unfolded jet multiplicity distributions as a function of the jet transverse momenta in pp collisions at $\sqrt{s} = 13$ TeV. I corrected the extracted jet-multiplicity distributions for detector effects and investigated the presence of a KNO-like scaling with similar methods as I did for heavy-flavor jets [193].

I showed that the jet-multiplicity distributions can be scaled onto a universal curve. However, calculating the statistical moments of the jet multiplicity distributions show that the scaling is not perfectly fulfilled in the measured $p_{\mathrm{T}}^{\mathrm{jet}}$ range, and this needs further investigations.

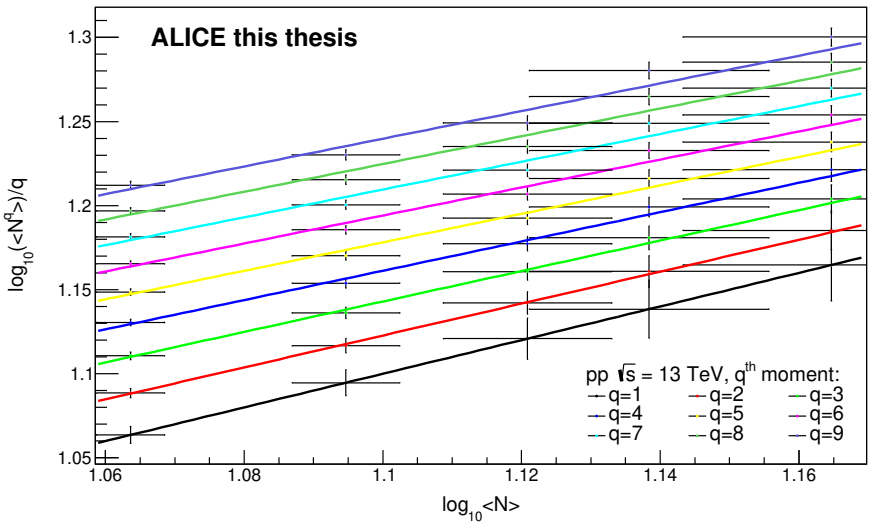

Figure 5.20: The first nine moments of the unfolded jet-multiplicity distributions, as a function of the average jet multiplicity corresponding to each $p_{\mathrm{T}}^{\mathrm{jet}}$ range.

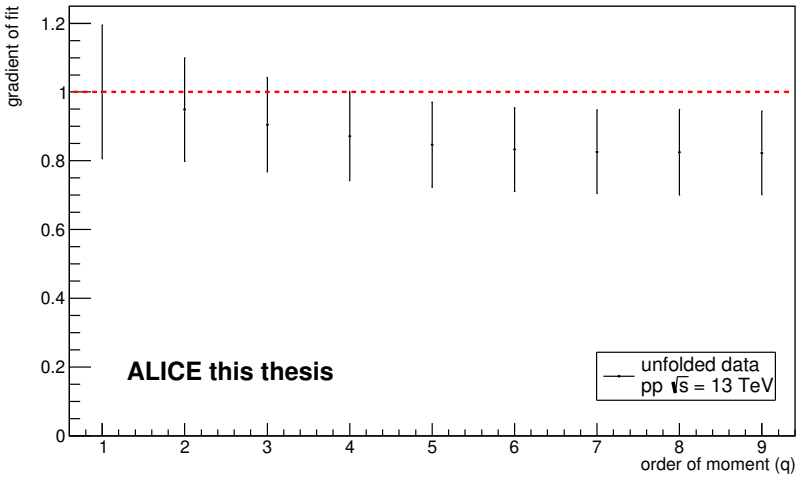

Figure 5.21: The gradients of the linear fits for the first nine statistical moments of the unfolded jet-multiplicity distributions.

Chapter 6

Azimuthal Correlations of Heavy-flavor Decay Electrons

Heavy-flavor measurements serve as an important tool to study QCD as we discussed in Sec. 2.6, and among those the heavy-flavor correlation measurements are particularly useful to study the fragmentation properties of hadrons (see Sec. 2.6.2).

In this chapter I present the ALICE measurements of the azimuthal correlations of electrons from heavy-flavor hadron decays with associated charged particles in pp collisions at $\sqrt{s} = 5.02$ TeV and p–Pb collisions at $\sqrt{s_{\mathrm{NN}}} = 5.02$ TeV [108]. My main contribution to the analysis was the creation of detailed simulations in PYTHIA 8, which were used to compare the near- and away-side peaks of the azimuthal-correlation distributions in both pp and p–Pb collisions to the model predictions. I determined the correlation peak shapes using FONLL pQCD calculations for the modelling of charm and beauty contributions.

The correlation distributions in this analysis were measured for associated charged particles in the range $1 < p_{\mathrm{T}}^{\mathrm{assoc}} < 7$ GeV/c and trigger electrons originating from heavy-flavor hadron decays in the range $4 < p_{\mathrm{T}}^{\mathrm{e}} < 12$ GeV/c. The used associate particle p_{T} range is significantly higher compared to previously published measurements of D-meson correlations with charged particles [94, 95], and the measurements were also compared with Monte Carlo (MC) simulations with different event generators: PYTHIA 6 [203] and PYTHIA 8 [204], HERWIG [14, 205], EPOS [206, 207], and POWHEG coupled with PYTHIA 6 for the parton shower and hadronization (POWHEG + PYTHIA 6) [16, 18]. A substantial difference among the generators was observed, with PYTHIA 8 and POWHEG + PYTHIA 6 providing the best description of the measured observables. These differences can be ascribed to the specific implementation of features such as hard-parton scattering matrix elements, parton showering, hadronization algorithm, and underlying event generation, which affect the correlation functions of heavy-flavor hadrons and charged particles.

The correlation distributions for trigger electrons in the transverse momentum ranges $4 < p_{\mathrm{T}}^{\mathrm{e}} < 7$ GeV/c and $7 < p_{\mathrm{T}}^{\mathrm{e}} < 16$ GeV/c were also measured in order to study the correlation shapes in the kinematic ranges where the electrons are dominantly produced by charm- and beauty-hadron decays, respectively.

This chapter is organized as follows: the data samples are reported and the analysis procedure is described in Sec. 6.1. The systematic uncertainties associated with the measurements are discussed in Sec. 6.2. The simulation details are discussed in Sec. 6.3. The analysis results are presented and discussed in Sec. 6.4. The chapter is briefly summarized in Sec. 6.5.

6.1 Analysis Overview

The analysis in this chapter is presented closely following the paper of the published results [108]. The results were obtained using minimum bias triggered data of pp collisions at $\sqrt{s} = 5.02$ TeV and p–Pb collisions at $\sqrt{s_{\mathrm{NN}}} = 5.02$ TeV, recorded with the ALICE detector during the LHC Run 2 data-taking period. The pile-up events that contain two or more primary vertices were rejected using an algorithm based on the detection of multiple vertices reconstructed from track segments in the SPD. A uniform acceptance of the detectors were obtained by considering only events with a reconstructed primary vertex within ± 10 cm from the center of the detector along the beam line for both pp and p–Pb collisions. The number of selected pp and p–Pb events were about 800M and 546M, respectively, corresponding to integrated luminosities of (16.63 ± 0.32) nb^{-1} [208] and (250 ± 10) μb^{-1} [209].

The measurements of two-particle azimuthal correlations between heavy-flavor hadron decay electrons (trigger) and charged (associated) particles were obtained from the correlation distributions of all identified electrons after subtracting the contribution of those electrons that do not originate from heavy-flavor hadron decays. The effects of the detector inhomogenieties and the limited two-particle acceptance were corrected with the event-mixing technique. The per-trigger correlation distributions were corrected for the efficiency of the associated-particle reconstruction, but they were not corrected for the trigger-electron efficiency, because the efficiency was found to be independent of p_{T} and the correction factor cancels with the per-trigger normalization. The peak yields and widths of the correlation distributions were obtained by applying a fit to the corrected $\Delta\varphi$ distribution. The same analysis technique was used for both pp and p–Pb measurements. Throughout this chapter, the term ”electron” will refer to both electrons and positrons. In the following subsections I briefly present the analysis procedure.

6.1.1 Azimuthal Correlation Distribution and Mixed-event Correction

The two-dimensional correlation distribution, denoted by $C(\Delta\varphi, \Delta\eta)$, was computed as a function of the azimuthal angle difference ($\Delta\varphi$) and pseudorapidity difference ($\Delta\eta$) between a heavy-flavor decay electron and charged particles. The p_{T} interval of the trigger particle was chosen to be $4 < p_{\mathrm{T}}^{\mathrm{e}} < 12$ GeV/c, as well as $4 < p_{\mathrm{T}}^{\mathrm{e}} < 7$ GeV/c and $7 < p_{\mathrm{T}}^{\mathrm{e}} < 16$ GeV/c. For the associated particles five p_{T} intervals were chosen between 1 and 7 GeV/c as $1 < p_{\mathrm{T}}^{\mathrm{assoc}} < 2$ GeV/c, $2 < p_{\mathrm{T}}^{\mathrm{assoc}} < 3$ GeV/c, $3 < p_{\mathrm{T}}^{\mathrm{assoc}} < 4$ GeV/c, $4 < p_{\mathrm{T}}^{\mathrm{assoc}} < 5$ GeV/c, and $5 < p_{\mathrm{T}}^{\mathrm{assoc}} < 7$ GeV/c. For each kinematic range, the correlation distributions were corrected both for the limited pair-acceptance and detector inhomogeneities, using the event-mixing technique [210] as shown in Fig. 6.1.

The mixed-event correlation distribution, $ME(\Delta\varphi, \Delta\eta)$, was obtained by correlating electrons in one event with charged particles from other events that had similar multiplicity and primary-vertex position along the beam direction. The mixed-event distributions have a triangular-like shape as a function of $\Delta\eta$, due to the limited η coverage of the detector, and is approximately flat as a function of $\Delta\varphi$ (see Fig. 6.1). Any non-flatness observed in $\Delta\varphi$ would indicate φ-dependent detector inefficiencies and inhomogeneities. At $(\Delta\varphi, \Delta\eta) \approx (0, 0)$ the trigger and associated particles experience the same detector effects and therefore the per-trigger correlation distribution is not affected. This property was used to obtain the normalization factor β for the mixed event distribution, which is defined as the average number of counts in the $-0.2 < \Delta\varphi < 0.2$ and $-0.07 < \Delta\eta < 0.07$ ranges.

The $\mathrm{d}^2N/(\mathrm{d}\Delta\eta\mathrm{d}\Delta\varphi)$ mixed-event corrected correlation distribution was obtained as the ratio of the correlation distribution from the same event to the mixed event distribution, scaled by the normalization factor β, i.e.

$$\frac{\mathrm{d}^2N}{\mathrm{d}\Delta\eta\mathrm{d}\Delta\varphi} \equiv S(\Delta\eta, \Delta\varphi) = \beta \times \frac{C(\Delta\eta, \Delta\varphi)}{ME(\Delta\eta, \Delta\varphi)} \,, \tag{6.1}$$

where the mixed-event corrected correlation distribution is denoted as $S(\Delta\eta, \Delta\varphi)$.

Due to the limited size of the heavy-flavor decay electron sample, the two-dimensional correlation distribution was subject to significant statistical fluctuations, especially at large $|\Delta\eta|$ values. To enhance the precision of the measurement, a one-dimensional $S(\Delta\varphi)$ distribution was obtained by integrating the mixed-event corrected azimuthal correlation distribution over the $|\Delta\eta| < 1$ range.

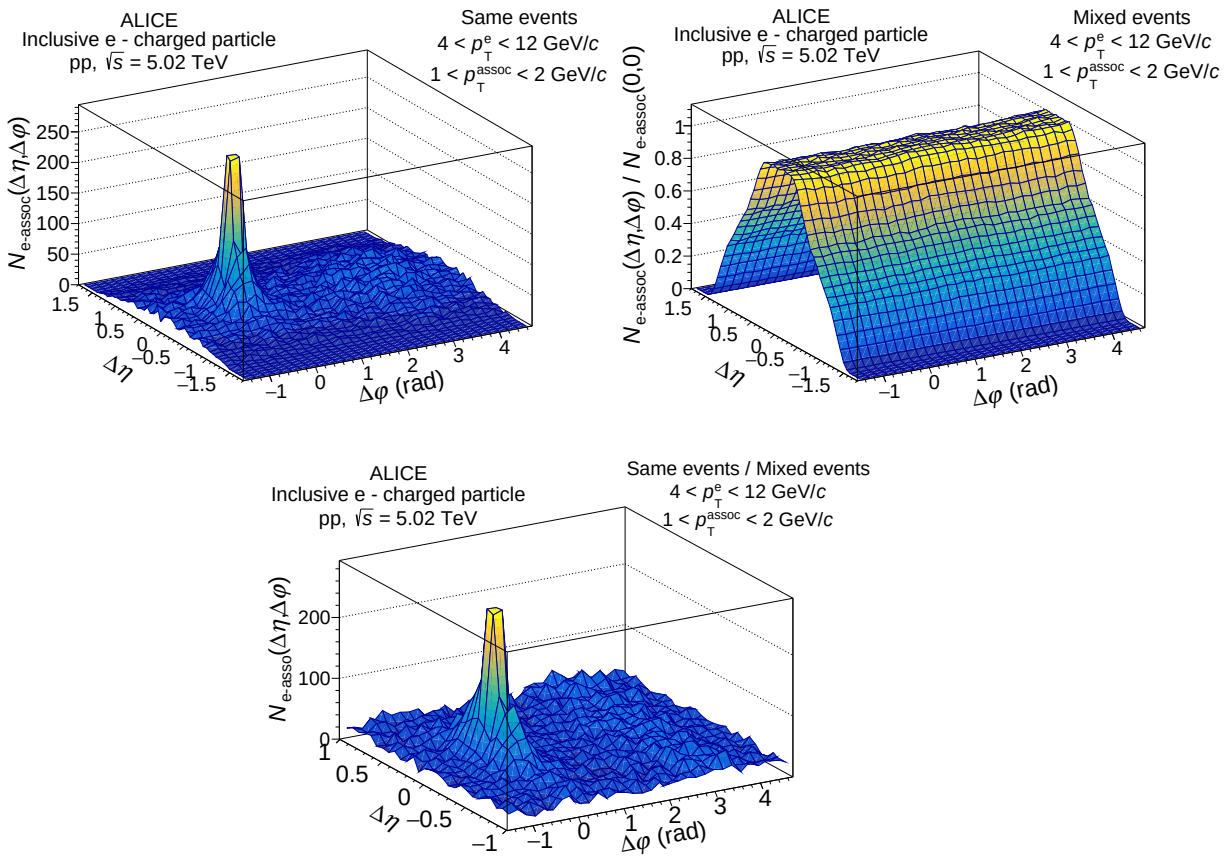

Figure 6.1: Example of measured same-event (*top left*), mixed-event (*top right*), and corrected (*bottom*) correlation distributions in the $4 < p_{\mathrm{T}}^{\mathrm{e}} < 12$ GeV/c and $1 < p_{\mathrm{T}}^{\mathrm{assoc}} < 2$ GeV/c ranges.

6.1.2 Background Subtraction

The selected trigger electron sample was contaminated by hadrons, which was estimated by considering tracks identified as hadrons using $n\sigma_{\mathrm{e}}^{\mathrm{TPC}} < -3.5$. Similarly to the procedure applied in Ref. [211], the energy/momentum (E/p) distribution the contaminating hadrons was scaled to match the E/p distribution of the heavy-flavor electron candidate in the $0.3 < E/p < 0.65$ region, which is outside the signal region. The charged hadron contamination was estimated to be $\approx 1\%$ at p_{T} = 4 GeV/c, increasing to $\approx 12\%$ at 16 GeV/c in both pp and p–Pb collisions. The estimated contamination was then subtracted from the inclusive electron correlation distribution.

The selected electrons come from two sources: signal electrons originating from heavy-flavor hadron decays and background electrons. The main source for background electrons is the Dalitz decay of neutral mesons (π^0 and η) and photon conversions in the detector material, which produce $\mathrm{e}^+\mathrm{e}^-$ pairs with low invariant mass, peaked around zero. The background electrons were identified using the invariant-mass technique [212, 213], in

which each selected electron is paired with a partner of opposite sign, and their invariant mass is calculated. In order to increase the efficiency of finding the partner electrons, they are selected by applying a similar but less strict track-quality and particle-identification criteria than used for selecting the signal electrons [213, 214]. Electron–positron pairs coming from the background have a small invariant mass, while random combinations of electron–positron pairs, such as heavy-flavor decay electrons forming a pair with other electrons, have a wider invariant-mass distribution. This combinatorial contribution was estimated from the invariant-mass distribution of like-sign electron pairs. The $S(\Delta\varphi)$ distributions of both like-sign electron pairs and unlike-sign electron pairs were obtained, and the background contribution was evaluated by subtracting the like-sign electron pair distribution from the unlike-sign electron pair distribution in the invariant mass region below $0.14\ \mathrm{GeV}/c$. The tagging efficiency, i.e. the efficiency of finding the partner electron, was estimated with MC simulations. In the p_{T} ranges used in this analysis the background contribution from other sources, such as decays of J/ψ and kaons, is negligible [212].

The azimuthal correlation distribution of electrons from heavy-flavor hadron decays and charged particles had to be corrected for the inefficiencies in the reconstruction of the associated particles. They also had to be corrected for the contamination of secondary particles in the associated particle sample. The reconstruction efficiency for charged primary particles was obtained using a different MC sample, which had no embedded particles. For pp this sample was generated by PYTHIA 6 [203] and for p–Pb by HIJING [32].

Utilizing the same MC simulations, the amount of contamination from secondary particles [215] was also estimated, and for the p_{T} intervals considered, it was found to be in the range 2–4% in pp collisions and 4–6% in p–Pb collisions. The per-trigger normalization was obtained by dividing the fully-corrected azimuthal-correlation distribution by the number of electrons originating from heavy-flavor hadron decays ($N_{(\mathrm{c,b})\to\mathrm{e}}$), expressed as

$$N^{(\mathrm{c,b})\to\mathrm{e}} = N_{\mathrm{InclE}} - \frac{1}{\varepsilon_{\mathrm{tag}}}[N_{\mathrm{ULS}} - N_{\mathrm{LS}}]\ , \tag{6.2}$$

where N_{InclE} is the number of electrons in the inclusive sample, N_{ULS} is the number of unlike-sign electron pairs, N_{LS} is the like-sign electron pairs, and ϵ_{tag} is the efficiency of finding a partner electron.

6.1.3 Characterization of the Azimuthal Distribution

The measured azimuthal correlation was fitted with two von Mises functions and a constant term b, to quantify its properties:

$$f(\Delta\varphi) = b + Y_{\mathrm{NS}}\frac{e^{\kappa_{\mathrm{NS}}\cos(\Delta\varphi)}}{2\pi I_0(\kappa_{\mathrm{NS}})} + Y_{\mathrm{AS}}\frac{e^{\kappa_{\mathrm{AS}}\cos(\Delta\varphi-\pi)}}{2\pi I_0(\kappa_{\mathrm{AS}})}. \tag{6.3}$$

where the two von Mises functions are used to model circular data, i.e. describe the near-side (NS) and away-side (AS) peaks. The terms κ_{NS} and κ_{AS} in the von Mises function are the measure of concentration of NS and AS peaks, respectively. The $1/\kappa$ parameter is analogous to the variance σ^2, and I_0 is the zeroth-order modified Bessel function evaluated at κ. The parameters Y_{NS} and Y_{AS} denote the integral of the near- and away-side peaks, respectively. By symmetry considerations, the mean of the NS is fixed to $\Delta\varphi = 0$ and the mean of the AS peak is at $\Delta\varphi = \pi$. The constant term b is a free parameter and describes the baseline, the physical minimum of the $\Delta\varphi$ distribution. The width (σ) of the peaks can be calculated by

$$\sigma = \sqrt{-2 \log \frac{I_1(\kappa)}{I_0(\kappa)}}, \tag{6.4}$$

where I_1 is the first-order modified Bessel function evaluated at κ. The per-trigger yields of the NS and AS peaks were obtained by subtracting the baseline b from the distribution and integrating over the bin counts in the ranges $-3\sigma_{\mathrm{NS}} < \Delta\varphi < 3\sigma_{\mathrm{NS}}$ and $-3\sigma_{\mathrm{AS}} < \Delta\varphi - \pi < 3\sigma_{\mathrm{AS}}$, respectively.

6.2 Systematic Uncertainties

There are important systematic uncertainties that affect the $\Delta\varphi$ correlation distributions, the per-trigger NS and AS yields, and widths, which need to be taken into account. These uncertainties are related to the electron-track selection procedures, the identification and subtraction of the hadron contamination, the background-electron subtraction, the associated-particle efficiency correction, the mixed-event correction, and the fitting procedure of the correlation distributions. Each source of the aforementioned systematic uncertainties was estimated separately, by either varying the selection criteria or by using an alternative approach and observing the difference it causes in the results. For each of these variations, their effect on both the NS and AS peak yields and widths were obtained by fitting and subtracting the baseline of the newly obtained correlation distributions and recalculating the observables. The uncertainties were then computed separately for each trigger electron and associated particle p_{T} range.

It was also taken into account that the systematic uncertainties coming from the associated-particle efficiency correction and mixed-event correction are correlated in $\Delta\varphi$. The other sources were considered to be independent. A summary of the systematic uncertainties of the correlation distribution, NS and AS yields and widths for $4 < p_{\mathrm{T}}^{\mathrm{e}} < 12$ GeV/c were reported in detail in Ref. [108].

There are possible biases related to the specific track quality selection for electrons, which were studied by varying the selection criteria [79]. The uncertainty from track selection was found to have a negligible effect on the NS and AS widths.

There is an uncertainty in the electron identification using the TPC and EMCAL signals, which was estimated by the variation of the selection criteria, i.e. changing the parameters $n\sigma_{\mathrm{e}}^{\mathrm{TPC}}$, E/p, and σ^2_{long}. The resulting uncertainties were found to be 2-6% for the AS and NS yields, and 2-7% for the AS and NS widths.

As I mentioned earlier, the contribution from background electrons was estimated by the invariant-mass method. The procedure has a systematic uncertainty that mainly affects the average tagging efficiency. The uncertainty was estimated by varying the track selection criteria of the partner electrons, which includes the minimum p_{T} and the invariant-mass window of the electron–positron pairs.

There is also an uncertainty related to the specific selection of associated particles, which was estimated by the variation of the charged track selection criteria. This included the requirement of registering a hit in one of the two SPD layers of the ITS. The selection on the distance of closest approach was also varied, which affects the secondary particle contamination. Important to note that this uncertainty was considered to be correlated in $\Delta\varphi$.

The limited detector acceptance and local inhomogenieties in the detector have significant effects. These were corrected for by using the mixed-event technique, and the uncertainty was estimated by varying the normalization factor β in the mixed-event technique.

A non-zero flow v_2 of heavy-flavor decay electrons and charged particles can further affect the $\Delta\varphi$ distributions. Since there are no previous measurements of heavy-flavor decay electron v_2 in minimum bias pp and p–Pb collisions, a conservative estimate was obtained from measurements in 0–20% central p–Pb collisions in Ref. [216].

Some of the uncertainties come from the fitting procedure, and several checks were performed to study the stability of the correlation distribution fits. Alternative functions were also used for the fit, for example a Gaussian and a generalized Gaussian were utilized to fit the NS and AS peaks instead of the von Mises function. Alternative fits were also used to fix the baseline value b and study its stability. Instead of the default bin counting procedure, the NS and AS yields were also calculated by integrating the fitting functions in the range $-3\sigma_{\mathrm{NS}} < \Delta\varphi < 3\sigma_{\mathrm{NS}}$ and $-3\sigma_{\mathrm{AS}} < \Delta\varphi - \pi < 3\sigma_{\mathrm{AS}}$. The maximum variation of the above results was taken as the overall systematic uncertainty. The uncertainty coming from the baseline estimation is quoted as an absolute numbers which affects all $\Delta\varphi$ bins by the same value. The uncertainty of the NS and AS peak yields and width for the $4 < p_{\mathrm{T}}^{\mathrm{e}} < 12$ GeV/c range varies by 4–9% for pp and 10–11% for p–Pb collisions, respectively.

The systematic uncertainties from the aforementioned sources were found to be similar for the $4 < p_{\mathrm{T}}^{\mathrm{e}} < 7$ GeV/c and $7 < p_{\mathrm{T}}^{\mathrm{e}} < 16$ GeV/c ranges, in both collision systems, and a detailed summary of the systematic uncertainties is reported in Ref. [108].

6.3 Simulation Details

My main contribution to the analysis was to create detailed simulations in PYTHIA 8 (including using the Angantyr model to simulate p–Pb collisions) to compare the near- and away-side peaks of the azimuthal-correlation distribution in pp and p–Pb collisions to the model predictions. This allowed verifying the implementation of the processes of charm- and beauty-quark production, fragmentation, and hadronization, which have an impact on the observables studied in this analysis.

The models used to compare the measurement in pp collisions are PYTHIA 8 with the Monash tune [21, 204] and EPOS 3.117 [206, 207]. As discussed in Section 2.3.1, the PYTHIA 8 event generator is widely used in particle physics, as it provides an accurate description of high-energy collisions. It is capable of generating both hard and soft interactions, initial and final-state parton showers, particle fragmentation, and multi-partonic interactions. It also incorporates color reconnection mechanisms to rearrange color connections between quarks and gluons during hadronization. The prediction of these models for correlations of D-mesons with charged particles can be found in Refs. [93, 94]. The p–Pb measurements were compared with PYTHIA 8 Angantyr [29] and EPOS 3.117 [206, 207] models.

For PYTHIA 8 simulations, I obtained the correlation distributions for electrons from charm- and beauty-hadron decays separately, and summed them after weighting their relative fractions based on FONLL calculations [88, 217, 218]. In Fig. 6.2 I show the fraction of the beauty decay electrons to the charm decay electrons, which I used for the reweighting.

As PYTHIA 8 does not natively support collisions involving nuclei, this feature is implemented in the Angantyr model, which combines several nucleon–nucleon collisions to build a proton–nucleus (p–A) or nucleus–nucleus (A–A) collision (see Sec. 2.3.1). I used the Angantyr [29] model to simulate ultra-relativistic p–Pb collisions with the PYTHIA 8 event generator. The EPOS3 event generator is largely used for the description of ultra-relativistic heavy-ion collisions, as discussed in Sec. 2.3.3.

In the models, the azimuthal correlation function of trigger electrons from charm- and beauty-hadron decays with charged particles was evaluated using the same prescriptions applied for data analysis in terms of kinematic and particle-species selections. The peak properties of the correlation functions were obtained by following the same approach employed in data, i.e. by fitting the distributions with two von Mises functions and a constant term.

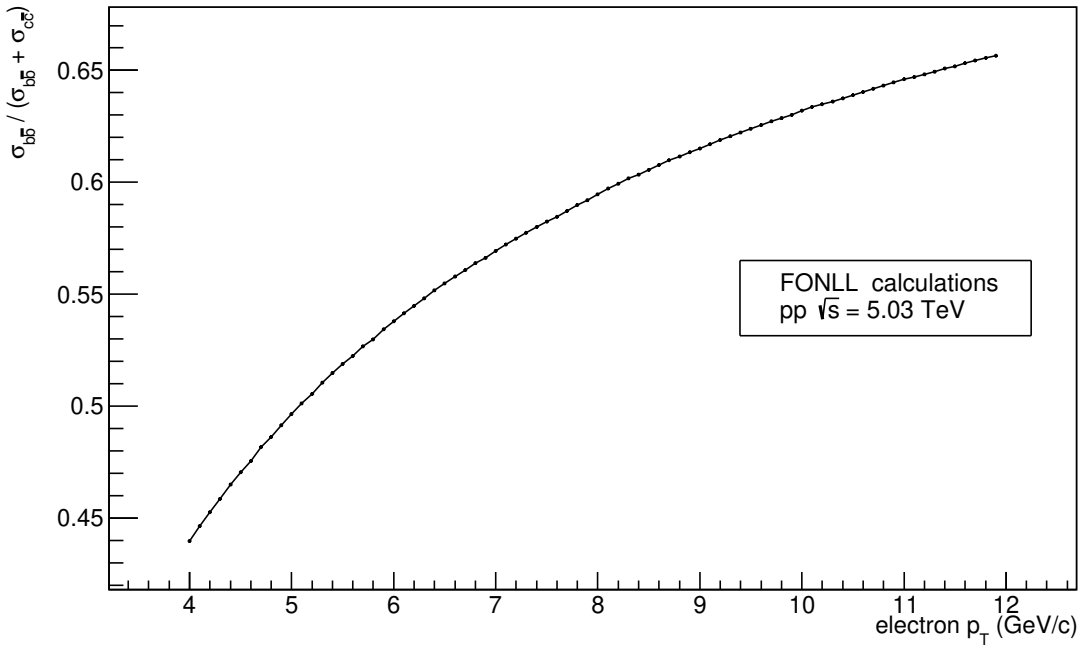

Figure 6.2: The fraction $\sigma_{b\bar{b}}/(\sigma_{b\bar{b}} + \sigma_{c\bar{c}})$ as a function of electron p_{T} which I used for the FONLL reweighting.

6.4 Results

In this section I present the main results of the azimuthal correlations analysis of heavy-flavor decay electrons with charged particles, closely following our published article [108].

6.4.1 Comparison of the Results in pp and p–Pb Collisions

In Fig. 6.3 the azimuthal-correlation distributions are shown for different $p_{\mathrm{T}}^{\mathrm{assoc}}$ ranges. The three top panels present the results for pp collisions, while the bottom three panels contain the p–Pb collision results. The distributions are normalized with the number of heavy-flavor decay electrons. The baseline is fitted with a constant and shown with a green line, together with the absolute systematic uncertainty of the baseline estimation as a solid box at $\Delta\varphi \sim -2$ rad. The correlated systematic uncertainties from the associated particle selection and mixed-event correction are shown for each $p_{\mathrm{T}}^{\mathrm{assoc}}$ range under the plotted curves. It can be seen that the near- and away-side peaks are described well by the von Mises fit function in all of the $p_{\mathrm{T}}^{\mathrm{assoc}}$ ranges. The baseline contribution is higher in p–Pb collisions compared to pp collisions, due to the larger charged-particle multiplicity. However, its absolute value decreases with increasing $p_{\mathrm{T}}^{\mathrm{assoc}}$ for both pp and p–Pb collisions. Since a large fraction of the baseline is coming from the underlying event, the pairs

contributing to it are dominated by low p_{T} particles.

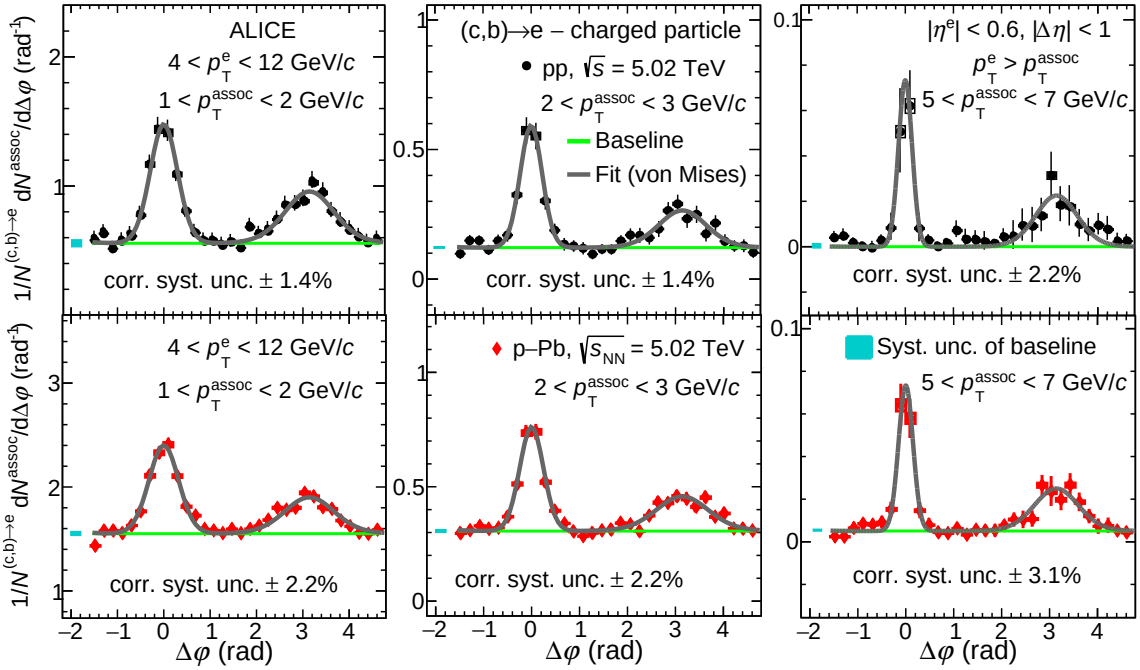

Figure 6.3: The azimuthal-correlation distribution for the $4 < p_{\mathrm{T}}^{\mathrm{e}} < 12$ GeV/c range fitted with a constant function for the baseline and von Mises functions for the NS and AS peaks, for different associated p_{T} ranges. The statistical uncertainties are shown as vertical lines, while the systematic uncertainties are shown as empty boxes.

The baseline-subtracted azimuthal-correlation distributions are compared for pp and p–Pb collisions in Fig. 6.4. The three panels contain results for three different $p_{\mathrm{T}}^{\mathrm{assoc}}$ ranges, while the heavy-flavor decay electron is in the range $4 < p_{\mathrm{T}}^{\mathrm{e}} < 12$ GeV/c. It can be seen that the NS and AS peaks decrease with increasing $p_{\mathrm{T}}^{\mathrm{assoc}}$ for both pp and p–Pb. A tendency for the NS peak to become more collimated with increasing $p_{\mathrm{T}}^{\mathrm{assoc}}$ is also visible. The overall profile of the correlation peaks is consistent in pp and p–Pb collisions within statistical and systematic uncertainties. The results indicate that there are no visible cold-nuclear matter effects impacting heavy-quark fragmentation and hadronization in the measured p_{T} range. This observation is consistent with previous measurements of D-meson correlations with charged particles [93, 94].

In order to quantitatively compare the azimuthal-correlation distributions in pp and p–Pb data, the per-trigger AS and NS peak yields are plotted for both collision systems in Fig. 6.5 in the top panels, while the peak widths are plotted in the bottom panels. These fit parameters are plotted as a function of $p_{\mathrm{T}}^{\mathrm{assoc}}$ in the $4 < p_{\mathrm{T}}^{\mathrm{e}} < 12$ GeV/c electron momentum range. The systematic uncertainties on the ratio consider all sources uncorrelated between pp and p–Pb collisions except for the baseline estimation. The partially correlated uncertainty of the baseline estimation was obtained by using different fit functions.

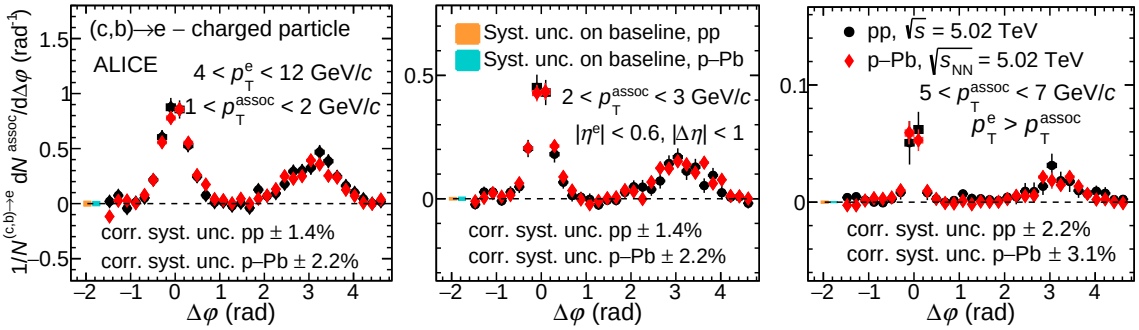

Figure 6.4: The azimuthal-correlation distribution after baseline subtraction for the $4 < p_{\mathrm{T}}^{\mathrm{e}} < 12$ GeV/c range and for different associated p_{T} ranges. The statistical uncertainties are shown as vertical lines, while the systematic uncertainties are shown as empty boxes.

The total uncertainty then was obtained by taking the quadratic sum of the correlated and uncorrelated uncertainties.

Both the NS and AS yields show a trend of decreasing in value as we go for higher $p_{\mathrm{T}}^{\mathrm{assoc}}$. This result is consistent within uncertainties for both pp and p–Pb collision systems for all measured $p_{\mathrm{T}}^{\mathrm{assoc}}$ ranges. The decrease in yields with increasing $p_{\mathrm{T}}^{\mathrm{assoc}}$ can be explained by the heavy quarks on average fragmenting hard into heavy-flavor hadrons. Due to energy conservation, and since a large fraction of the energy is spent on fragmenting into heavy-flavor hadron, it is on average more likely that the associated particles to the decay electron are produced with lower p_{T}. The NS width values also show a trend of decreasing with increasing $p_{\mathrm{T}}^{\mathrm{assoc}}$, with having a value of about 0.3 at $p_{\mathrm{T}}^{\mathrm{assoc}} = 1$ GeV/c. At $p_{\mathrm{T}}^{\mathrm{assoc}} = 6$ GeV/c the NS width is roughly 0.15, having a significance of about 3σ for both pp and p–Pb collision systems. The significance was calculated by taking the difference of the widths in the lowest and highest $p_{\mathrm{T}}^{\mathrm{assoc}}$ ranges, and also taking into account both statistical and systematic uncertainties. The AS widths on the other hand seem to be independent of the $p_{\mathrm{T}}^{\mathrm{assoc}}$ range, and take a value around 0.5. The NS peak distribution is closely connected to the fragmentation of the jet which contains the trigger particle.

The narrowing of the NS width with increasing $p_{\mathrm{T}}^{\mathrm{assoc}}$ indicates that the charged particles with higher p_{T} tend to be closer to the jet-axis, which can be approximated by the direction of the trigger electron. This parallels the higher p_{T} emissions from heavy quarks being more collinear to it than lower p_{T} emissions. The AS peak exhibits a lower sensitivity to the fragmentation of a specific heavy quark, because it can contain particles produced via the fragmentation of heavy quarks originating from processes that are not azimuthally back-to-back, including certain next-to-leading order processes. These processes can have different relative fractions depending on the p_{T} of the heavy quark. For example, considering gluon splitting, the AS peak can also include particles originating from the recoil gluon not directly associated with the heavy quarks produced in the event. Even considering back-to-back processes, the correlation between the transverse momentum of the

heavy-flavor decay electron and the heavy quark on the opposite side in the event, which is responsible for generating the AS peak through fragmentation, is significantly weaker than for the near-side peak.

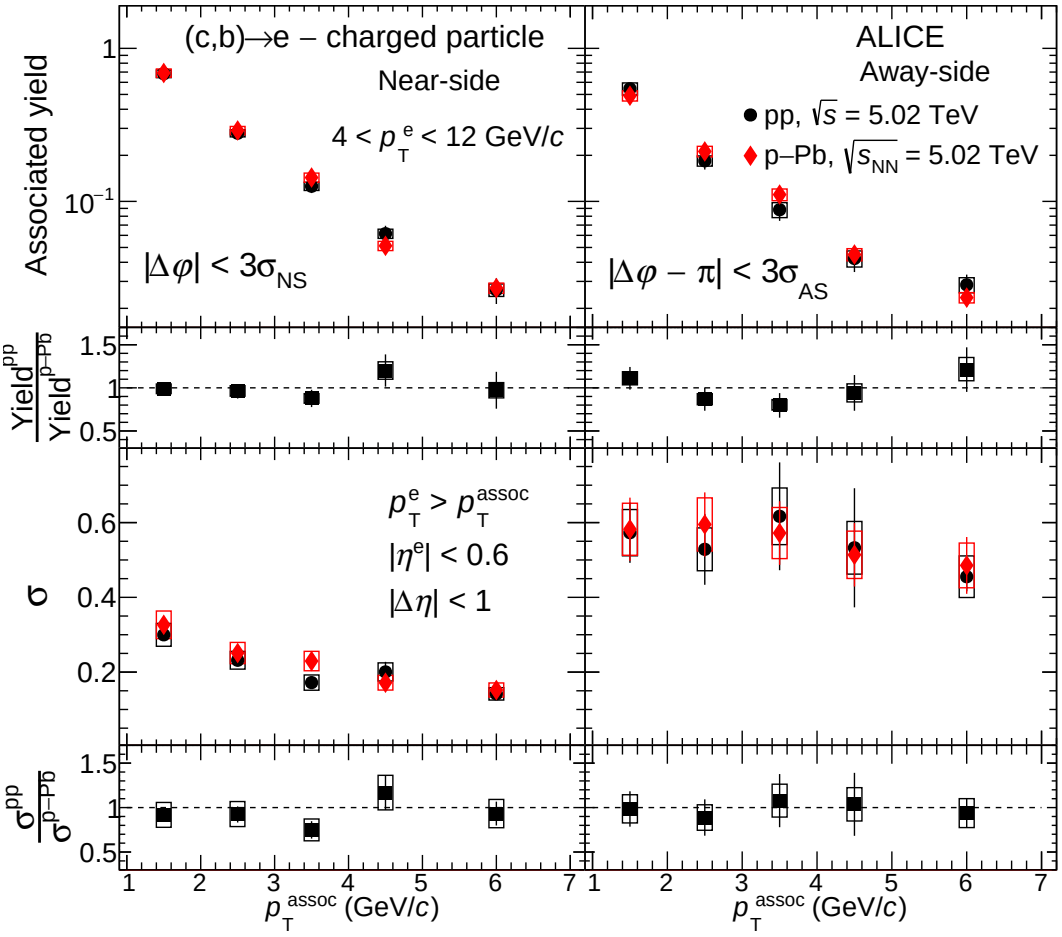

Figure 6.5: The near- and away-side per-trigger yields (*top panels*) and widths (*bottom panels*) as a function of $p_{\mathrm{T}}^{\mathrm{assoc}}$, for $4 < p_{\mathrm{T}}^{\mathrm{e}} < 12$ GeV/c. The ratios between pp and p–Pb yields and widths are also shown. The statistical uncertainties are shown as vertical lines, while the systematic uncertainties are shown as empty boxes.

6.4.2 Comparison with Predictions from MC Event Generators

In Fig. 6.6 the azimuthal-correlation distribution in pp is compared with model predictions after baseline subtraction for $4 < p_{\mathrm{T}}^{\mathrm{e}} < 12$ GeV/c and in three different $p_{\mathrm{T}}^{\mathrm{assoc}}$ ranges. In Fig. 6.7 the same is plotted but for Pb–Pb collisions. The distributions are reflected in the $0 < \Delta\varphi < \pi$ range and the statistical uncertainties are shown as vertical lines, while the systematic uncertainties are shown as empty boxes.

Both the PYTHIA 8 and EPOS3 event generators describe the data well in all the $p_{\mathrm{T}}^{\mathrm{assoc}}$ intervals under investigation. However, the EPOS3 predictions show some deviation from the measured NS and AS peaks in the highest $p_{\mathrm{T}}^{\mathrm{assoc}}$ range, and this deviation from the data is much more pronounced for the p–Pb collision data. In Figs. 6.8 and 6.9 we also compare with the model predictions the extracted fit parameters: the peak yields and widths of the distributions. The PYTHIA 8/Angantyr simulations predict NS and AS yields which decrease with increasing $p_{\mathrm{T}}^{\mathrm{assoc}}$ and this is consistent with the data within statistical and systematic uncertainties.

The NS widths simulated by PYTHIA 8/Angantyr decrease as $p_{\mathrm{T}}^{\mathrm{assoc}}$ increases, which is consistent with the data in both collision systems. On the other hand, the AS widths show a slightly decreasing trend as a function of $p_{\mathrm{T}}^{\mathrm{assoc}}$, which is also consistent with data within statistical and systematic uncertainties in both collision systems. The EPOS3 model describes the data qualitatively in pp collisions within statistical and systematic uncertainties. For p–Pb collisions it overestimates the NS peak yield for high $p_{\mathrm{T}}^{\mathrm{assoc}}$, but the AS peak yield remain consistent with the data within statistical and systematic uncertainties. As it can also be seen on the figures, the EPOS3 event generator overestimates the NS widths and underestimate the AS widths for all the $p_{\mathrm{T}}^{\mathrm{assoc}}$ ranges in both pp and p–Pb collisions.

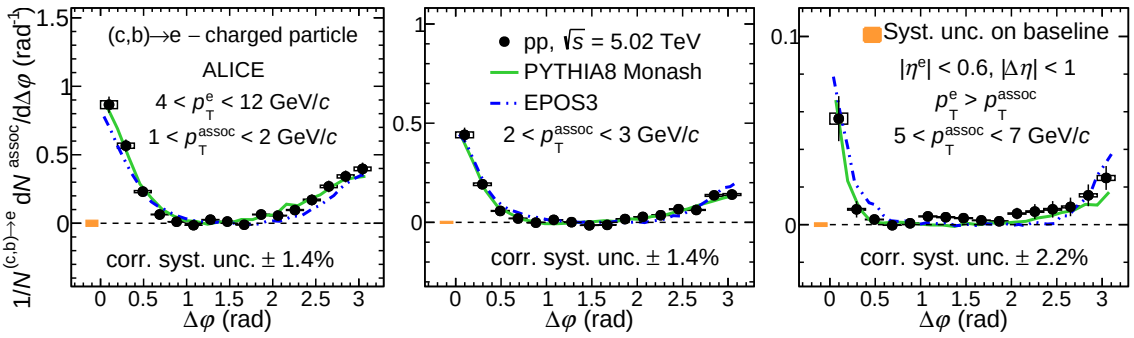

Figure 6.6: The azimuthal-correlation distribution compared with model predictions after baseline subtraction, for $4 < p_{\mathrm{T}}^{\mathrm{e}} < 12$ GeV/c, in different $p_{\mathrm{T}}^{\mathrm{assoc}}$ ranges. The statistical uncertainties are shown as vertical lines, while the systematic uncertainties are shown as empty boxes.

6.4.3 The $p_{\mathrm{T}}^{\mathrm{e}}$-dependence of the Correlation Distribution

The relative fraction of heavy-flavor electrons produced by charm- or beauty decay have a strong p_{T}-dependence [96]. For heavy-flavor decay electrons with $p_{\mathrm{T}}^{\mathrm{e}}$ = 4 GeV/c, about 40% of the yield originates from beauty decays. This ratio increases to about 60-70% for heavy-flavor electrons with $p_{\mathrm{T}}^{\mathrm{e}} > 8$ GeV/c.

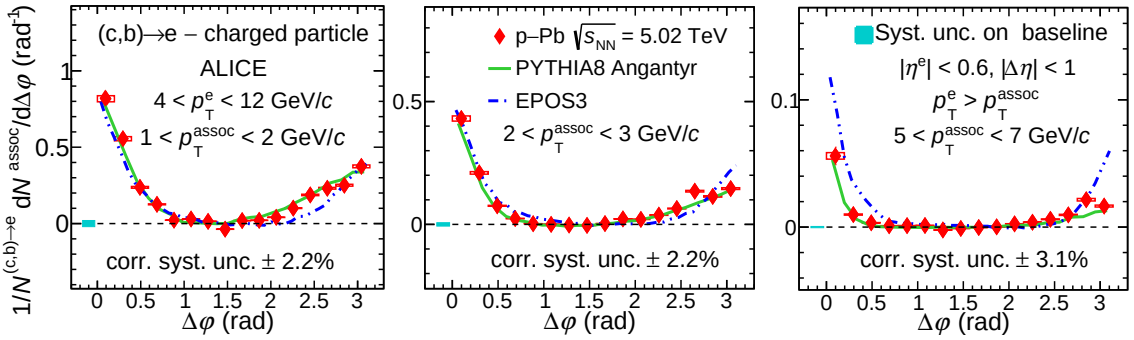

Figure 6.7: The azimuthal-correlation distribution compared with model predictions after baseline subtraction, for $4 < p_{\mathrm{T}}^{\mathrm{e}} < 12$ GeV/c, in different $p_{\mathrm{T}}^{\mathrm{assoc}}$ ranges. The statistical uncertainties are shown as vertical lines, while the systematic uncertainties are shown as empty boxes.

Since the charm and beauty quarks fragment differently, we can expect the azimuthal correlation distributions of the electrons originating from these quarks to have a flavor dependence. For a given quark flavor, it is also expected to have a $p_{\mathrm{T}}^{\mathrm{e}}$-dependence due to the electron-producing partons having different energies and different fraction of LO and NLO processes being involved in their production. Azimuthal correlation distribution measurements in the ranges $4 < p_{\mathrm{T}}^{\mathrm{e}} < 7$ GeV/c and $7 < p_{\mathrm{T}}^{\mathrm{e}} < 16$ GeV/c provide valuable information on fragmentation because the latter $p_{\mathrm{T}}^{\mathrm{e}}$ range is dominated by electrons originating from beauty-hadron decays. The results for the yields and widths of the distributions in these regions are shown in Fig. 6.10 and 6.11, and the ratios between the yields of in the $7 < p_{\mathrm{T}}^{\mathrm{e}} < 16$ GeV/c and $4 < p_{\mathrm{T}}^{\mathrm{e}} < 7$ GeV/c regions is shown in the second row of the figures. The two figures show that the associated yield for both the near-side and away-side peaks are higher in the range where the trigger electron has a higher momentum, and this is true for both collision systems. This is to be expected due to the heavy-flavor quark fragmentation process with higher trigger electron usually producing a larger amount of associated particles.

The near-side peak width values are decreasing with $p_{\mathrm{T}}^{\mathrm{assoc}}$ for both $p_{\mathrm{T}}^{\mathrm{e}}$ ranges. On the other hand, the away-side width values show an approximately flat trend for both $p_{\mathrm{T}}^{\mathrm{e}}$ ranges. It is important to note that the $p_{\mathrm{T}}^{\mathrm{assoc}} < p_{\mathrm{T}}^{\mathrm{e}}$ selection condition introduces a kinematic bias, because measuring the distributions in the $4 < p_{\mathrm{T}}^{\mathrm{e}} < 7$ GeV/c and $4 < p_{\mathrm{T}}^{\mathrm{assoc}} < 7$ GeV/c ranges results in missing some associated particles.

The figures also show the comparison of the measured yield and width values to model calculations in PYTHIA 8/Angantyr and EPOS3. The PYTHIA 8/Angantyr simulations describe the data within uncertainties for both trigger electron ranges. The EPOS3 predictions of the yields are also consistent with data for both trigger electron intervals, however the predicted trend of the widths differ from data. Both MC generators describe the ratio of the yields and widths in both $p_{\mathrm{T}}^{\mathrm{e}}$ ranges well.

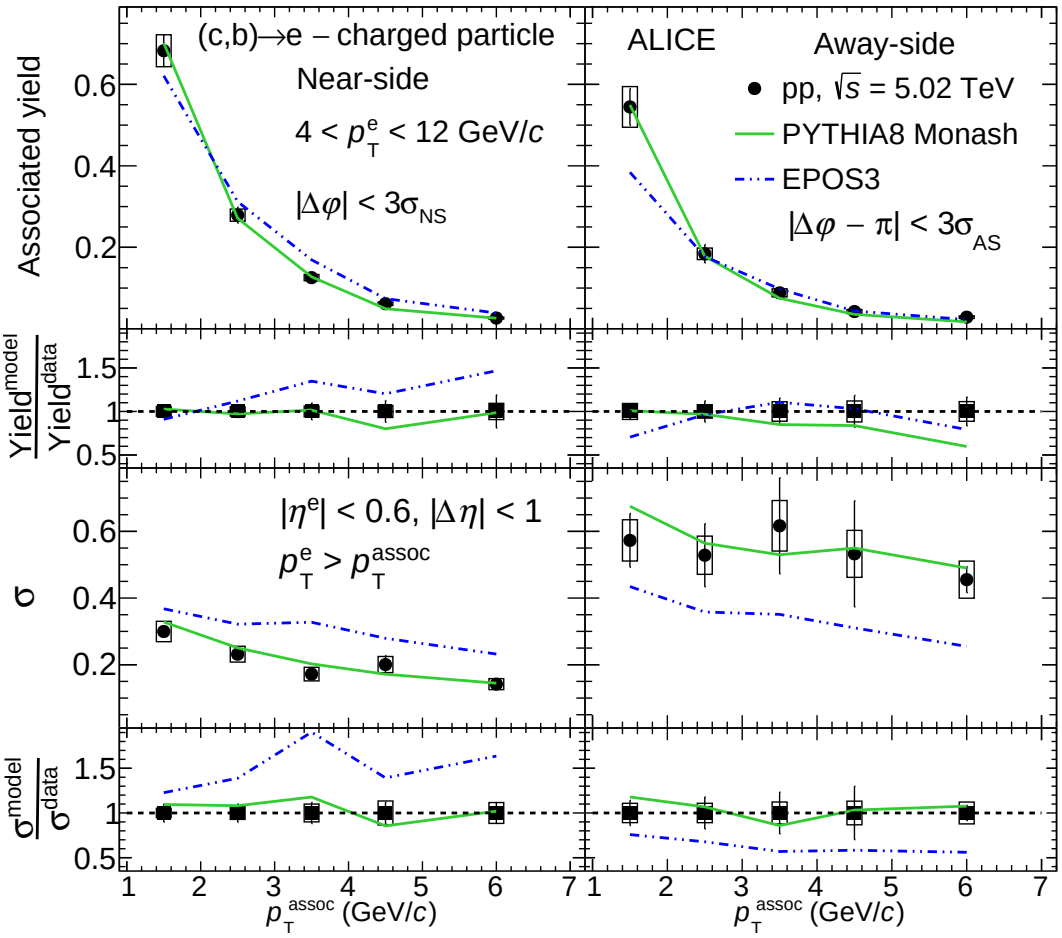

Figure 6.8: The near- and away-side per-trigger yields (*top panels*) and widths (*bottom panels*) as a function of $p_{\mathrm{T}}^{\mathrm{assoc}}$, for $4 < p_{\mathrm{T}}^{\mathrm{e}} < 12$ GeV/c, compared with predictions from PYTHIA 8 Monash tune and EPOS3. The ratios between model predictions and data are also shown. The statistical uncertainties are shown as vertical lines, while the systematic uncertainties are shown as empty boxes.

In order to understand how the different charm and beauty fragmentation affects the observed $p_{\mathrm{T}}^{\mathrm{e}}$-dependence of the distributions, I calculated the correlation distributions in PYTHIA 8 MC simulations, separately for electrons created in charm-hadron decays and electrons created in beauty-hadron decays. Fig. 6.12 shows the near-side and away-side peak yields, as well as the widths for electrons from charm- and beauty-hadron decays. The figure also contains the ratios of these values to the values measured with the combined electron sample. The near-side yields for trigger electrons from beauty-hadron decays are observed to be lower than those from charm-hadron decays, in both $p_{\mathrm{T}}^{\mathrm{e}}$ intervals. This observation can be explained by the harder fragmentation of beauty quarks into beauty hadrons, compared to that of charm quarks, therefore less energy remaining for the production of other particles in the parton shower.

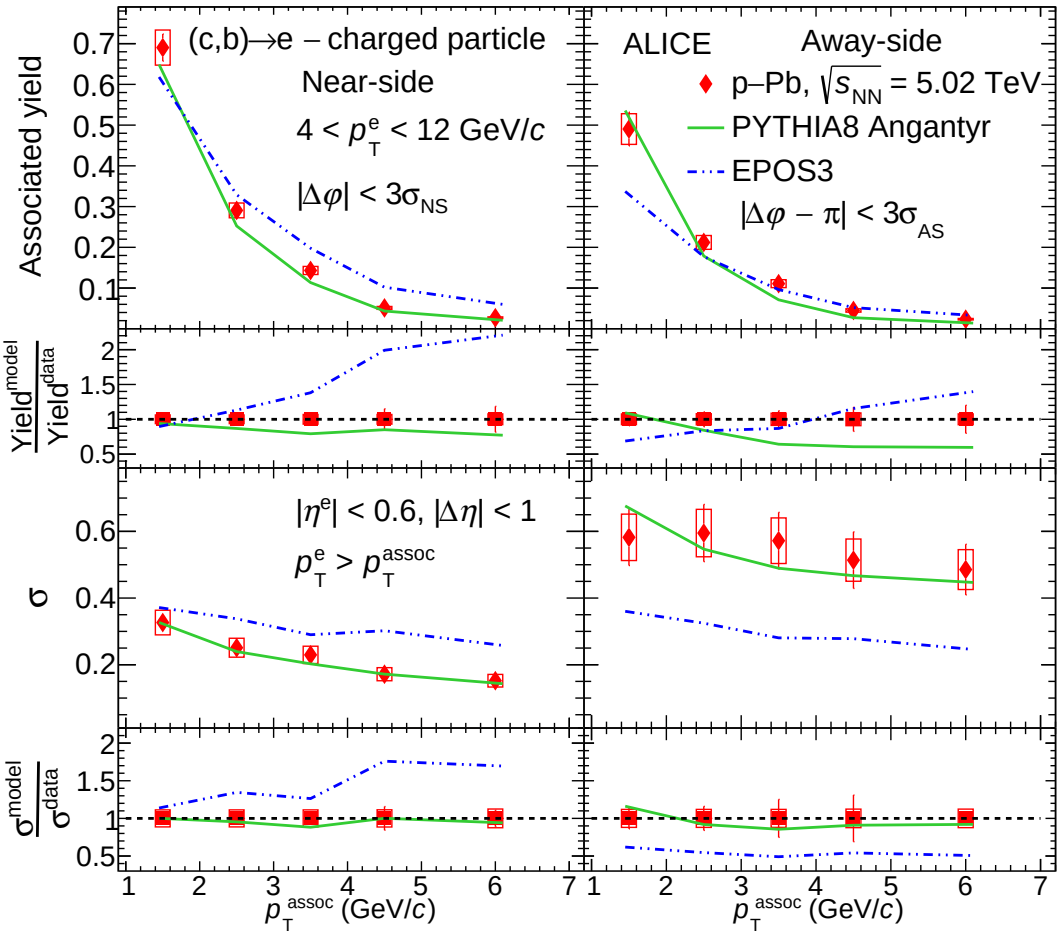

Figure 6.9: The near- and away-side per-trigger yields (*top panels*) and widths (*bottom panels*), as a function of $p_{\mathrm{T}}^{\mathrm{assoc}}$, for $4 < p_{\mathrm{T}}^{\mathrm{e}} < 12$ GeV/c, compared with predictions from PYTHIA 8 Angantyr and EPOS3. The ratios between model predictions and data are also shown. The statistical uncertainties are shown as vertical lines, while the systematic uncertainties are shown as empty boxes.

In the case of the away-side yields, the difference becomes smaller as we go for higher $p_{\mathrm{T}}^{\mathrm{assoc}}$ ranges. In both $p_{\mathrm{T}}^{\mathrm{e}}$ ranges, both the near-side and away-side widths of the distributions decrease with increasing $p_{\mathrm{T}}^{\mathrm{e}}$ for both charm- and beauty-hadron decays, but for electrons from beauty-hadron decays the widths are wider than for electrons from charm-hadron decays. These opposite effects lead to similar measured width values for the two $p_{\mathrm{T}}^{\mathrm{e}}$ ranges in Figs. 6.10 and 6.11.

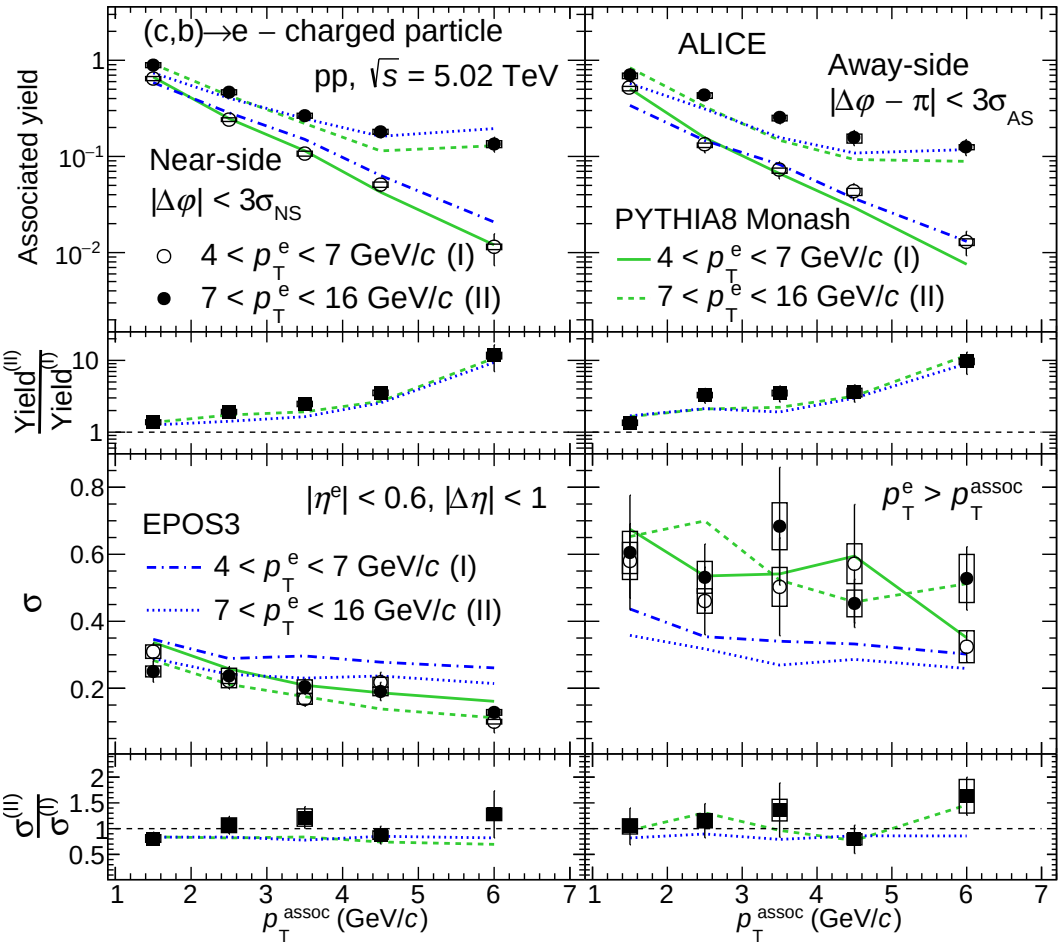

Figure 6.10: The NS and AS per-trigger yields (*top panels*) and widths (*bottom panels*) for two $p_{\mathrm{T}}^{\mathrm{e}}$ ranges: $4 < p_{\mathrm{T}}^{\mathrm{e}} < 7$ GeV/c and $7 < p_{\mathrm{T}}^{\mathrm{e}} < 16$ GeV/c, as a function of $p_{\mathrm{T}}^{\mathrm{assoc}}$. The ratios between the yields of the two regions are also shown. The data is compared with PYTHIA 8 Monash and EPOS3 predictions. The statistical uncertainties are shown as vertical lines, while the systematic uncertainties are shown as empty boxes.

6.5 Summary

In this chapter I reported measurements of azimuthal-correlation functions of heavy-flavor hadron decay electrons with charged particles in pp and p–Pb collisions at $\sqrt{s_{\mathrm{NN}}} = 5.02$ TeV. The correlation distributions were obtained for trigger electrons in the range $4 < p_{\mathrm{T}}^{\mathrm{e}} < 12$ GeV/c, and for different associated particle p_{T} ranges between 1 and 7 GeV/c. Two von Mises functions with a constant was used to fit the azimuthal distributions in order to characterize the near- and away-side peaks.

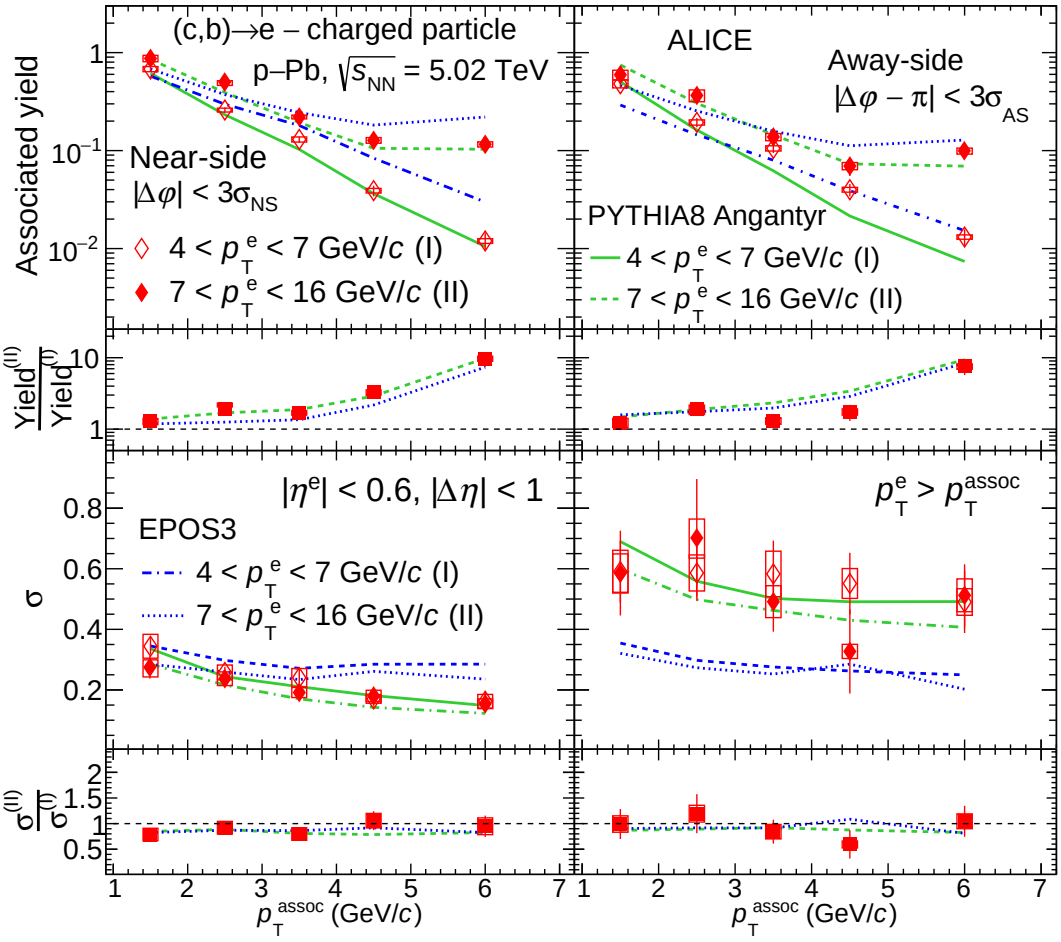

Figure 6.11: The NS and AS per-trigger yields (*top panels*) and widths (*bottom panels*) for two $p_{\mathrm{T}}^{\mathrm{e}}$ ranges: $4 < p_{\mathrm{T}}^{\mathrm{e}} < 7$ GeV/c and $7 < p_{\mathrm{T}}^{\mathrm{e}} < 16$ GeV/c, as a function of $p_{\mathrm{T}}^{\mathrm{assoc}}$. The ratios between the yields and widths of the two regions are also shown. The data is compared with PYTHIA 8 Angantyr and EPOS3 predictions. The statistical uncertainties are shown as vertical lines, while the systematic uncertainties are shown as empty boxes.

In both pp and p–Pb collisions the evolution of the near- and away-side peaks was found to be similar for the whole investigated kinematic range, which means that within the current precision we cannot detect a modification in the heavy-quark fragmentation or hadronization due to cold-nuclear-matter effects, but as the ALICE experiment continues collecting more data, it may be possible in the future. On the other hand, the extracted near- and away-side per-trigger yields and widths as a function of the associated particle p_{T} provide access to the momentum distribution of the particles fragmented from a hard parton, and allow for a differential study of jet substructure. The per-trigger yield as a function of the charged (associate) particle p_{T} was found to have the same trend in both pp and p–Pb collisions, namely it decreases as $p_{\mathrm{T}}^{\mathrm{assoc}}$ increases.

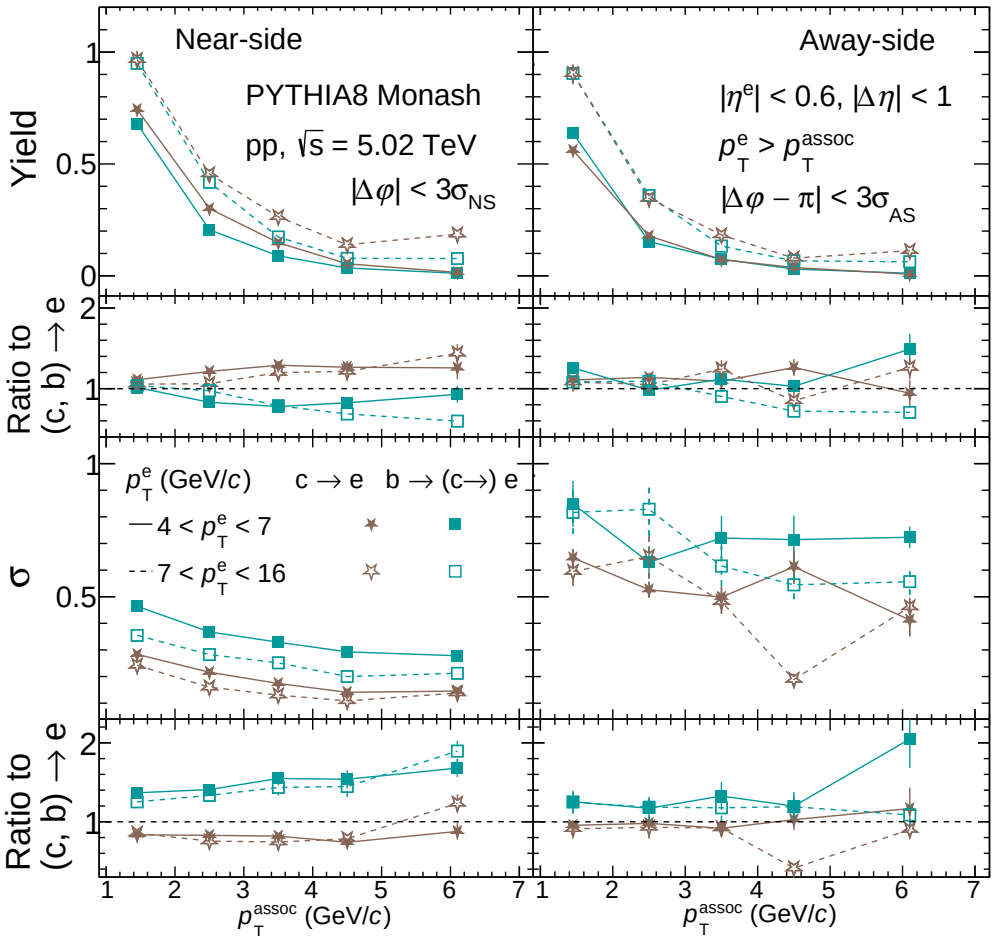

Figure 6.12: The PYTHIA 8 Monash prediction for NS and AS per-trigger yields (*top panels*) and widths (*bottom panels*) in two $p_{\mathrm{T}}^{\mathrm{e}}$ ranges: $4 < p_{\mathrm{T}}^{\mathrm{e}} < 7$ GeV/c and $7 < p_{\mathrm{T}}^{\mathrm{e}} < 16$ GeV/c, as a function of $p_{\mathrm{T}}^{\mathrm{assoc}}$. The ratios to c, b $\rightarrow$ e yields and widths are also shown. The statistical uncertainties are shown as vertical lines, while the systematic uncertainties are shown as empty boxes.

The p_{T}-dependence of the width is also consistent between the two collision systems: the width of the near-side peak decreases with increasing $p_{\mathrm{T}}^{\mathrm{assoc}}$ and the away-side peak width has no pronounced trend as a function of $p_{\mathrm{T}}^{\mathrm{assoc}}$. All observables, including the $\Delta\varphi$ distributions, per-trigger yields, and widths in both pp and p–Pb collisions were compared with predictions from EPOS3 and PYTHIA 8 simulations, where the latter was utilized with the Monash tune to simulate pp collisions, and the Angantyr model to simulate p–Pb collisions. The best description was provided by the PYTHIA 8 simulations, which gave good predictions for the yields and widths of both the near- and away-side peaks.

The relative fraction of electrons originating from charm- and beauty-hadron decays was shown to have a strong p_{T} dependence, and this quality was exploited by studying the correlation distribution in different kinematic regions, such as $4 < p_{\mathrm{T}}^{\mathrm{e}} < 7\ \mathrm{GeV}/c$ and $7 < p_{\mathrm{T}}^{\mathrm{e}} < 16\ \mathrm{GeV}/c$, where the latter range is dominated by the decays of beauty-hadrons.

Comparing the aforementioned trigger p_{T} ranges, the $7 < p_{\mathrm{T}}^{\mathrm{e}} < 16\ \mathrm{GeV}/c$ range has a systematically larger per-trigger yield, because more particles are produced from the fragmentation of the more energetic heavy-quark. This effect is actually stronger than the increased beauty-hadron decays over charmed hadron decays, which according to the PYTHIA 8 studies results in lower correlation peak yields in this region than those of electrons originating from charm quarks.

The PYTHIA 8 studies also shed light on the observation of the near- and away-side widths being similar for both $p_{\mathrm{T}}^{\mathrm{e}}$ ranges. This happens due to competing effects between the larger boost of the initial heavy quark leading to a stronger collimation of the peaks with increasing $p_{\mathrm{T}}^{\mathrm{e}}$ for both charm- and beauty-origin contributions, and the broader peak widths for trigger electrons originating from beauty-hadron decays, whose contribution increases with $p_{\mathrm{T}}^{\mathrm{e}}$.

These results also contribute to future Pb–Pb measurements at the same center-of-mass energy by serving as a reference. Finding modifications in the correlation functions in heavy-ion collisions could provide valuable information on the dynamics of heavy-flavor quarks inside the quark-gluon plasma [107].

Chapter 7

Enhanced Production of Charmed Baryons

Charmed baryon-to-meson ratios are sensitive probes of heavy-quark fragmentation. The production of these heavy-flavor hadrons in high-energy collisions is usually described by the factorization approach (see Sec. 2.6.1). Fragmentation functions have been assumed to be universal across different collision systems. However, the ALICE and CMS experiments at the LHC observed a low-momentum enhancement in the production of charmed Λ_c^+ baryons compared to charmed D^0 mesons in high-energy proton–proton collisions, with respect to model calculations which were tuned for electron-positron collisions [114, 219, 220]. Since the PDF and partonic cross sections cancel in the ratio, this suggests that the universality of charm fragmentation is not fulfilled.

There have been many proposed models which try to explain the observed excess production, such as color reconnection beyond leading color approximation (CR-BLC) [122], quark recombination mechanism (QCM) [221], feed-down from a largely augmented set of (so-far directly unobserved) higher mass charm-baryon states [123], and the Catania model based on a coalescence plus fragmentation approach [222]. Although all these scenarios tend to qualitatively describe the trends observed for the Λ_c^+ production, most of them fall short in explaining the yields and ratios of further charmed baryonic states e.g. $\Xi_c^{0,+}$ and $\Sigma_c^{0,++}$ [223]. Recently published data show that the Λ_c^+ enhancement correlates with event multiplicity [112]. Trends observed in the event multiplicity are reproduced well with models containing color reconnection beyond leading color approximation. This is expected since in this latter scenario, CR is linked to multiple-parton interactions (MPI), which in turn is related to the activity of the underlying event (UE) [130] and thus to the event multiplicity.

In this chapter I present the studies concerning the charm-baryon enhancement using simulations with color reconnection beyond leading color approximation, utilizing multiple event-activity classifiers which are sensitive to the origin of the charm production. The proposed methods can be used in future analysis of experimental data from the ongoing ALICE Run 3 data taking period to achieve a high discriminatory power between the different proposed scenarios.

7.1 Analysis Method

For this analysis I simulated 1 billion pp collisions at $\sqrt{s} = 13$ TeV center-of-mass energy using the PYTHIA 8.303 Monte Carlo (MC) event generator with the Monash tune and softQCD settings [12], with the color reconnection beyond leading color approximation model [122] implemented. The CR-BLC model has three different modes which are based on the Monash tune, established using a broad set of minimum-bias, Drell-Yan and underlying-event data from the LHC to constrain the parameters of the initial-state radiation and multiple parton interactions, combined with data from SPS and the Tevatron LHC to constrain the energy scaling [21].

Mode 2 of the CR-BLC model, which includes time dilation using the boost factor obtained from the final-state mass of the dipoles and requires a causal connection of all dipoles, is known to reproduce the trends in Λ_c^+/D^0 ratios [114, 224] and therefore I chose Mode 2 as the default setting. The Mode 0 results, which use no time-dilation constraints and where the amount of CR is controlled by the invariant mass scale parameter m_0, yield qualitatively similar results to Mode 2 with slightly less enhancement. Finally, Mode 3, which has time dilation constraints but only requires a single causal connection, was found to vastly overestimate the underlying event. Final-state charged particles were selected at mid-rapidity, in the pseudo-rapidity window $|\eta| < 1$ and in the full range of the azimuth angle φ, with a minimum transverse momentum $p_\mathrm{T} > 0.15$ GeV/c. The charmed Λ_c^+ and D^0 hadrons, and their charged conjugates, were selected in the rapidity window $|y| < 0.5$ based on their MC particle numbering scheme codes [225]. In the following I refer to both the particles and their charge conjugates simply as Λ_c^+ and D^0. It must be noted that the feed-down contribution from beauty hadrons is not removed, as its contribution to the ratio is only about 5% to 10% depending on the p_T.

To quantitatively characterize an event, first I used the event multiplicity N_ch, defined as the number of all charged final state particles in the event in the mid-rapidity acceptance defined above. Since N_ch is defined in the same pseudorapidity range where the charmed-hadron yields are computed, effects observed via N_ch may be influenced by autocorrelation. Therefore I also used the forward multiplicity N_fw, which I defined as the number of charged particles within the acceptance $2 < |\eta| < 5$.

While a dependence of Λ_c^+/D^0 on general event-activity observables is already seen in the experiment [112], these experimental results do not carry direct information on whether the observed pattern is governed by charm fragmentation in jets or hadronization in the underlying event. To address this question, I characterized the event with R_{T} and R_{NC}. As described in Sec. 2.7.2, these observables are selectively sensitive to the activity in the underlying-event or to that caused by high-momentum jets in the leading process. One such observable is the transverse event-activity classifier $R_{\mathrm{T}} \equiv N_{\mathrm{ch}}^{\mathrm{trans}}/\langle N_{\mathrm{ch}}^{\mathrm{trans}}\rangle$ [130], where $N_{\mathrm{ch}}^{\mathrm{trans}}$ is the charged-hadron multiplicity in the transverse region defined the following way. The charged final-state hadron with the highest transverse momentum within the acceptance (the trigger hadron) is selected and the event is accepted only if the trigger hadron has a transverse momentum $p_{\mathrm{T}}^{\mathrm{trig}} > 5$ GeV/c. $N_{\mathrm{ch}}^{\mathrm{trans}}$ is then the number of charged final-state hadrons in the transverse side, defined with the azimuth angle relative to the trigger hadron as $\frac{\pi}{3} < |\Delta\phi| < \frac{2\pi}{3}$ within $|\eta| < 1$. Since the trigger hadron most likely comes from a high-momentum jet initiated by the leading process, and the recoil jet is expected to show up at the opposite side, $N_{\mathrm{ch}}^{\mathrm{transverse}}$ is dominated by hadrons from the underlying event [131]. In models such as PYTHIA that describe events in terms of MPI, R_{T} is strongly correlated with the number of MPIs in an event [130]. Analogously, I defined the near-side cone activity $R_{\mathrm{NC}} \equiv N_{\mathrm{ch}}^{\text{near-side cone}}/\langle N_{\mathrm{ch}}^{\text{near-side cone}}\rangle$ in a narrow cone around the trigger particle, $\sqrt{\Delta\phi^2 + \Delta\eta^2} < 0.5$. As this region is dominated by the fragments of the jet containing the trigger particle, R_{NC} will be primarily determined by the multiplicity of the jet initiated by the leading hard process.

To quantify the jettiness and underlying event activity of events without a high-p_{T} trigger particle, I used the transverse spherocity S_0 [132], as defined in Section 2.7.2. The transverse spherocity takes values between 0 and 1. Isotropic events correspond to $S_0 \to 1$, while $S_0 \to 0$ for jetty events with strongly collimated particles. To overcome the limitations of the transverse spherocity, I also used flattenicity [134] to characterize the event. In this analysis I divided the η axis into 10 and the ϕ axis into 8 ranges of equal length, and used 80 flattenicity cells altogether.

While N_{MPI} is not a physically observable quantity, it is closely related to the observed charmed-baryon enhancement in the model class under investigation [125]. Therefore I used N_{MPI} as the most handy quantity to represent the relation of the enhancement to the final-state event activity. In this analysis I propose easily accessible physical observables that are powerful in distinguishing between different model scenarios and can be directly used in future data comparison.

In each event I selected charm hadrons in the central rapidity window $|y| < 0.5$. I analyzed the charmed baryons Λ_c^+, Σ_c^{++} and Σ_c^0 (referred to as $\Sigma_{\mathrm{c}}^{0,++}$), Ξ_c^+ and Ξ_c^0 (referred to as $\Xi_{\mathrm{c}}^{0,+}$), Ω_c^0 and Ω_c^{*0} (referred to as Ω_{c}^0), as well as the charmed D^0 meson. I excluded feed-down from beauty hadrons. In case of Λ_{c}^+ a substantial contribution comes from decays of other charm-baryon (predominantly the $\Sigma_{\mathrm{c}}^{0,++}$) states. Based on event-generator information I evaluated Λ_{c}^+ contributions from direct hadronization, i.e. separately from those that

stem from the decay of $\Sigma_c^{0,++}$ baryons. Since the meson-to-baryon and baryon-to-baryon ratios are sensitive to differences in fragmentation mechanisms without sensitivity to the heavy-flavor production cross section, I used the ratios of the different charmed-baryon to the D^0 meson as well as to the Λ_c^+, as a function of p_T and different event activity classes. The event-activity-class limits, summarized in Table 7.1, were determined to contain roughly similar number of events for each variable under study.

class	#1	#2	#3	#4	#5
N_{ch}	$\leq$15	16–30	31–40	41–50	$\geq$51
N_{fw}	$\leq$45	46–90	91–120	121–150	$\geq$151
R_T	$<$0.5	0.5–1.0	1.0–1.5	1.5–2.0	$>$2.0
R_{NC}	$<$0.5	0.5–1.0	1.0–1.5	1.5–2.0	$>$2.0
S_0	0.0–0.25	0.25–0.45	0.45–0.55	0.55–0.75	0.75–1.0
ρ	0.0–1.0	1.0–1.5	1.5–2.0	2.0–2.5	$>$2.5
N_{MPI}	$\leq$5	6–10	11–13	14–16	$\geq$17

Table 7.1: Definition of event classes for the different event-activity classifiers.

7.2 The Λ_c^+/D^0 Ratio

In Fig. 7.1 I compare the Λ_c^+ to D^0 ratios for different charged event multiplicity bins for CR-BLC Mode 2 for both mid-rapidity (left) and forward multiplicities (center). In the left and center panels of Fig. 7.1 ALICE results from Ref. [112] are also shown for comparison. On the right hand side, the Λ_c^+ to D^0 ratios are shown in different classes of N_{MPI}, extracted from Monte Carlo information. In the left panel it can be seen that the Λ_c^+ yield has a significant dependence on the multiplicity, as it increases with the multiplicity. This qualitatively reproduces the behavior seen in ALICE data [112].

In the center panel one can observe a similar dependence on the forward multiplicity: the Λ_c^+ yield again increases for higher forward multiplicities. In the latter case, correlation between the measured charmed-hadron yields from leading hard processes and the high charged-hadron multiplicity stemming from the same charmed jets is reduced. Since the yields increase in a similar way for both central and forward rapidity, the multiplicity-dependence according to my simulation results is not driven by charm-production in jets. Both trends are well reproduced by the trends observed in N_{MPI} classes, suggesting again that the enhancement is tied to charm production in the underlying event, driven by MPI in the CR-BLC model. Precise multi-differential measurements in both the central and forward rapidity regime from the ongoing and upcoming high-luminosity data-taking periods at the LHC experiments will be essential to shed light on the exact nature of Λ_c^+ enhancement and validate the predictions obtained using the CR-BLC model.

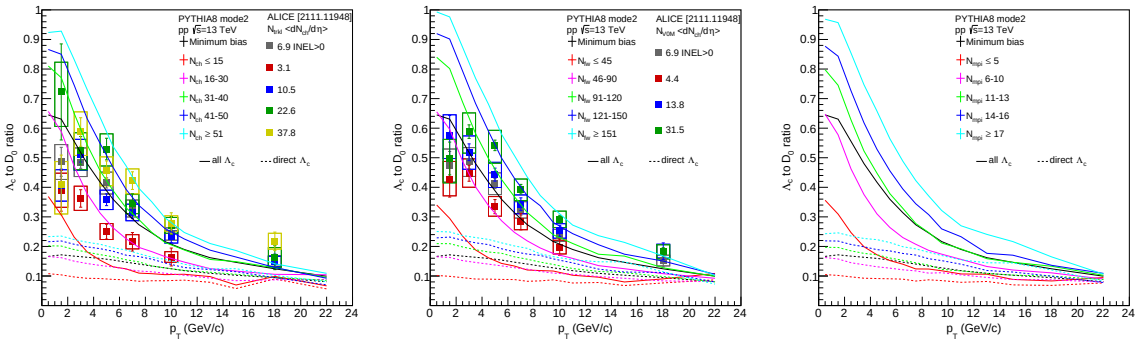

Figure 7.1: Λ_c^+/D^0 ratios from minimum-bias PYTHIA 8 simulations with CR-BLC as a function of p_T, shown as solid lines, for three different event-activity classifiers: charged-hadron multiplicity at mid-rapidity (N_{ch}, *left*), charged-hadron multiplicity at forward-rapidity (N_{fw}, *center*), and number of multiparton-interactions (N_{MPI}, *right*). The results for the N_{ch} and N_{fw} classes are also compared to data from ALICE [112]. The contribution of direct Λ_c^+ production is shown separately as dashed lines in each panel.

In the upper row of Fig. 7.2 I plot the Λ_c^+ to D^0 ratios as a function of p_T for different R_T and R_{NC} classes. In the case of the transverse-side multiplicities, higher R_T values correspond to stronger Λ_c^+ enhancement. On the other hand, for different near-side-cone multiplicities the Λ_c^+ to D^0 ratios remain consistent within fluctuations in the $p_T < 6$ GeV/c range, while there is some difference for $p_T > 6$ GeV/c. We can conclude therefore, that in the model class under investigation the increased Λ_c^+ yield is primarily connected to charm production within the underlying event and not the jet region. This is further highlighted in the bottom panel of Fig. 7.2, where I show the Λ_c^+ to D^0 ratio integrated over the semi-soft (or coalescence) regime $2 < p_T < 6$ GeV/c, plotted as a function of R_T as well as R_{NC}. While the change from small to large R_T values is almost threefold, there is a very slight dependence on R_{NC}, which may be caused by the UE that is not subtracted from the jet region.

In Fig. 7.3 I evaluated the Λ_c^+/D^0 excess in terms of transverse spherocity classes. The presence of high-momentum jets correlates with the event multiplicity. To minimize the effect of this correlation, we observe spherocity in fixed multiplicity intervals. Fig. 7.3 (*left*) shows the Λ_c^+/D^0 ratio for different spherocity bins for events with $N_{ch} > 50$. Jetty events show a significantly smaller Λ_c^+ enhancement than more isotropic events. In the right panel I show the Λ_c^+/D^0 ratios for different S_0 bins, integrated over the coalescence regime and plotted as a function of the N_{ch} classes. It can be observed that for lower multiplicity events the dependence of the ratio on spherocity is weak, while for higher multiplicities the separation between the ratios for low and high S_0 values becomes more significant. This pattern can can be understood by considering that in case of lower-multiplicity events most of the contribution comes from a single process, therefore the separation of the lead-

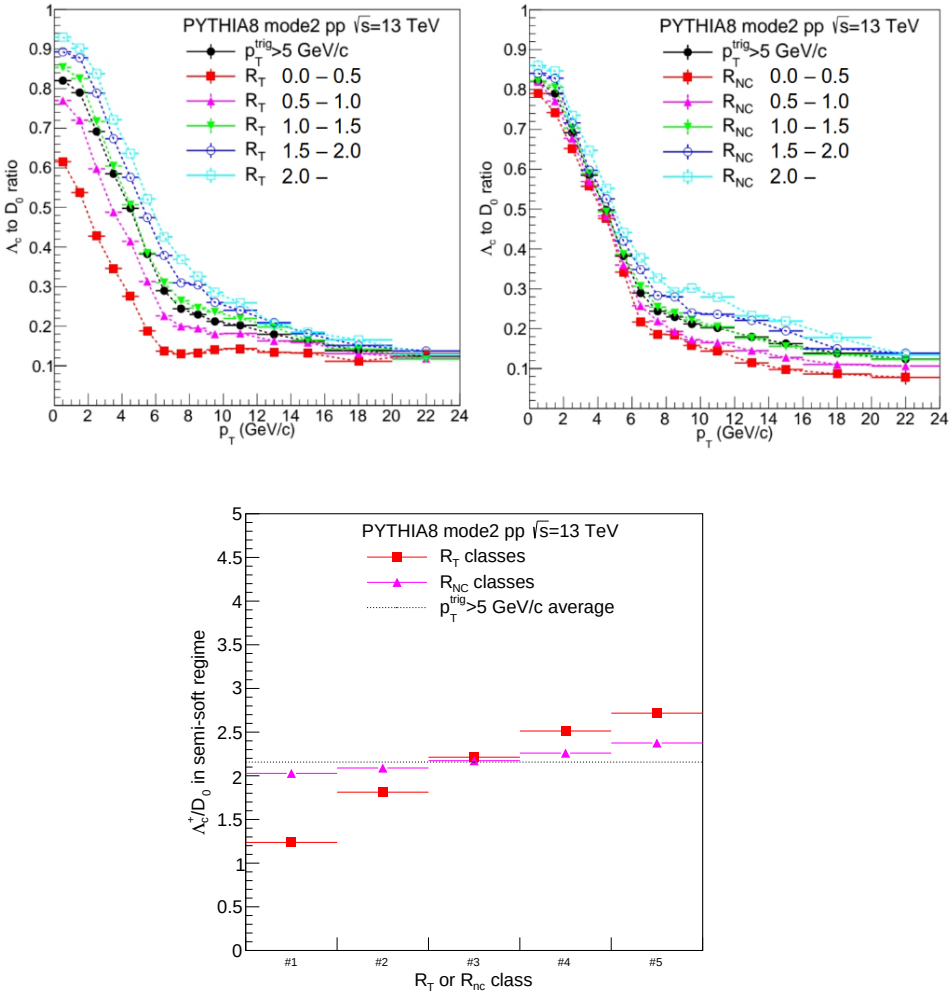

Figure 7.2: Λ_c^+ to D^0 ratios of hadron-triggered events as a function of p_T, for different R_T bins (*top left*) and R_{NC} bins (*top right*), together with multiplicity-inclusive hadron-triggered data, shown for reference. The Λ_c^+ to D^0 ratios were integrated over the coalescence regime $2 < p_T < 6$ GeV/c as a function of R_T and R_{NC} bins for hadron-triggered data (*bottom*). The dashed line represents the average of triggered events.

ing process and the UE is not meaningful. For higher multiplicities, however, both the leading process and the UE contribute to the event, and S_0 characterizes their relative proportion on a statistical basis. The left panel of Fig. 7.4 shows the Λ_c^+/D^0 ratio for different flattenicity classes, while the right panel shows the integrated enhancement for different flattenicity classes for several N_{ch} classes. Flattenicity correlates with the number of MPI and therefore it is strongly bound to the underlying event.

Similarly to transverse spherocity, flattenicity is a more powerful observable than R_T and R_{NC} in the sense that events not containing a high-p_T hadron can also be analyzed. However, the enhancement in the Λ_c^+/D^0 ratio decreases with increasing flattenicity in ev-

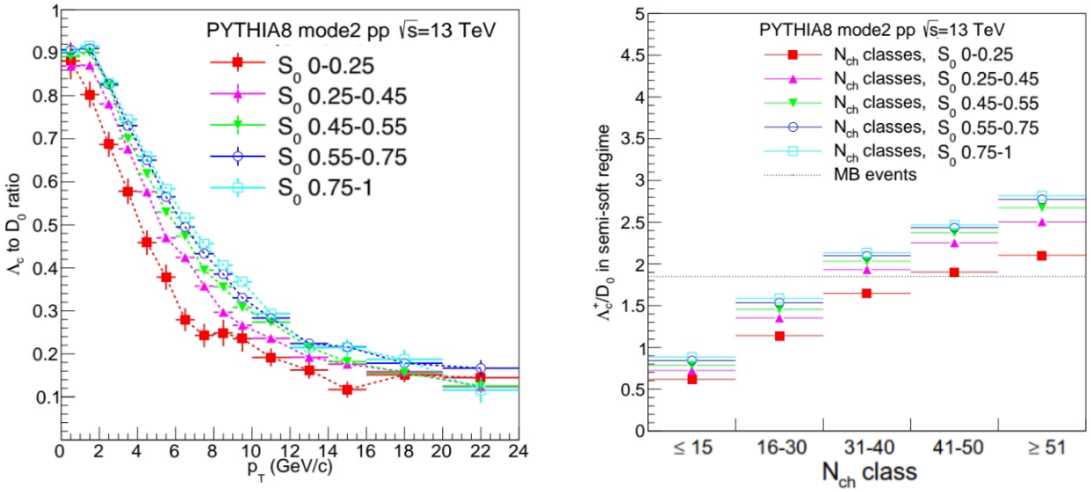

Figure 7.3: Λ_c^+ to D^0 ratios as a function of p_T, for different transverse spherocity classes for $N_{ch} > 50$ (*left*). Λ_c^+ to D^0 ratios integrated over the coalescence regime $2 < p_T < 6$ GeV/c, as a function of N_{ch}, in different spherocity bins (*right*). The dashed line represents the average of minimum-bias events.

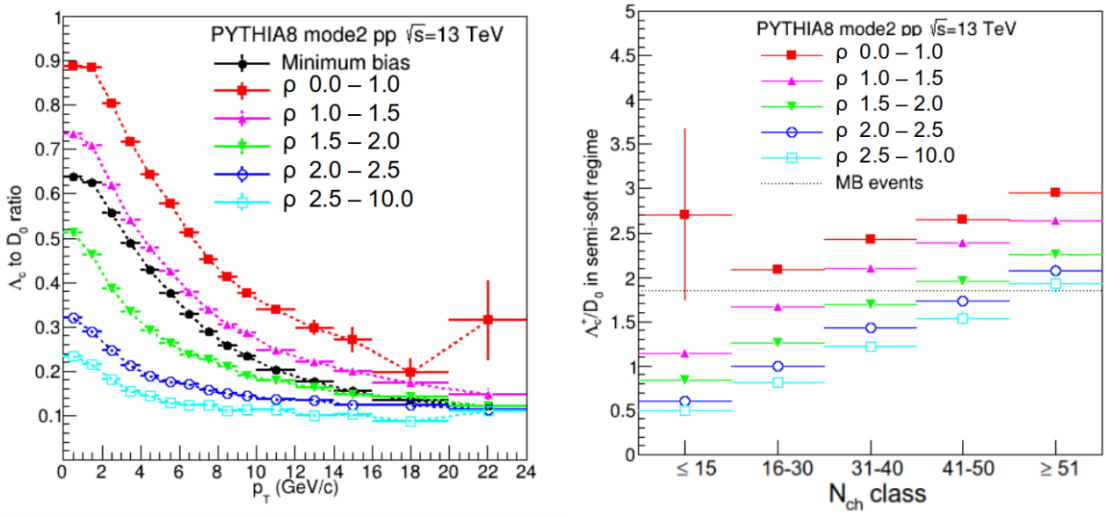

Figure 7.4: Λ_c^+ to D^0 ratios as a function of p_T in different ρ classes (*left*). Λ_c^+ to D^0 ratios integrated over the coalescence regime $2 < p_T < 6$ GeV/c, as a function of N_{ch}, in different flattenicity bins (*right*). The dashed line represents the average of minimum-bias events.

ery N_{ch} class, which makes it a more sensitive observable than the transverse spherocity. Future measurements of the charmed baryon and meson production as a function of flattenicity will therefore provide crucial feedback for models and will therefore play a key role in the understanding of heavy-flavor fragmentation.

7.3 Production of Excited Charm and Charm-strange Baryon States

In Section 7.2 I argued based on my published manuscript [125] that the Λ_c^+/D^0 enhancement with respect to different event-activity classifiers provide sensitive probes that can access the source of the enhanced charmed-baryon production, and thus differentiate between the above scenarios. In the CR-BLC model [122] the excess is linked to the MPI, which is strongly correlated with the underlying-event (UE) activity [130]. I used this model to conduct detailed studies for the Λ_c^+/D^0 ratios in proton–proton collisions at $\sqrt{s}$ =13 TeV. We saw that Λ_c^+/D^0 enhancement strongly depends on the final-state hadron multiplicity both in the central and the forward rapidity region. This behavior is supported by recent ALICE measurements that had been since finalized [112], hinting that the multiplicities taken from forward and central regions show the same enhancement qualities, therefore attesting to the usability of the CR-BLC scenario. I also proposed to observe the Λ_c^+/D^0 ratio in event classes that are determined based on the activity within the underlying event, as well as the multiplicity of the jet caused by the leading process, and concluded that the charmed-baryon enhancement can be observed to depend on the UE activity but not on the activity inside the jet region.

In this section I extend my studies of Λ_c^+ (qqc, $I = 0$) to the charmed baryon states $\Sigma_c^{0,++}$ (qqc, $I = 1$), $\Xi_c^{0,+}$ (qsc) and Ω_c^0 (ssc), to pin down the relative contributions of strangeness and charm in the baryon enhancement, and also address the role of the isospin. While I phrase the observations in terms of MPI using the CR-BLC scenario [122], I provide predictions for observables that are accessible for the experiment already in the LHC Run3 data-taking phase.

The calculations from PYTHIA 8 with CR-BLC reproduce data trends in which the enhancement of the Λ_c^+/D^0 ratio strongly correlates with the event multiplicities. It should be noted that the multiplicity classes in the model calculations do not correspond exactly to the ones from Ref. [112], and in case of forward rapidity the pseudorapidity definition also differs. In the ALICE analysis the multiplicity in the forward rapidity region has been estimated from the percentile distribution of the sum of signal amplitudes in the V0A and V0C scintillators, covering the pseudorapidity regions $2.8 < |\eta| < 5.1$ and $-3.7 < |\eta| < -1.7$, respectively. As mentioned in Sec. 7.2, in the scenario described by PYTHIA 8 with CR-BLC, the low-p_T Λ_c^+/D^0 excess is bound to the underlying-event activity, which is well represented by N_{MPI} in the model. While in this case the very same dependence on N_{MPI} can be observed similarly to the case of the N_{ch} and N_{fw} classes, classification based on observables linked to N_{MPI} can better differentiate between specific scenarios [125].

Besides Λ_c^+ baryons originating directly from the hadronization, there is a significant contribution from decay products of other charmed hadrons (predominantly the $\Sigma_c^{0,++}$ states), that even exceeds direct production at low p_T [111]. These two charm baryonic states differ in their isospin, which may influence the production of these particles depending on the hadronization mechanism. Both contributions are ordered with event activity. However, secondary Λ_c^+ production from $\Sigma_c^{0,++}$ decays dominates the low-p_T range and diminishes toward higher p_T, the contribution of direct Λ_c^+ is relatively flat in p_T, and thus dominates the high-p_T range. In Fig. 7.5 I show the ratios of $\Sigma_c^{0,++}$ yields both to the D^0

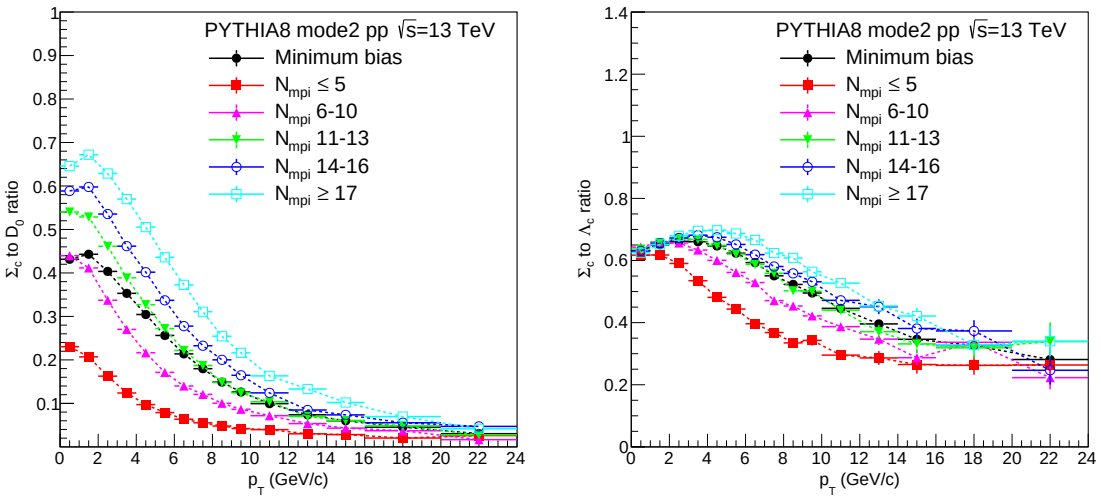

Figure 7.5: $\Sigma_c^{0,++}/D^0$ (*left*) and $\Sigma_c^{0,++}/\Lambda_c^+$ (*right*) ratios from PYTHIA 8 simulations with CR-BLC as a function of p_T, for MB events as well as for different multiparton-interaction (N_{MPI}) classes.

meson and to Λ_c^+, in different N_{MPI} classes. The $\Sigma_c^{0,++}/D^0$ ratio (left) shows the ordering by N_{MPI} (and N_{ch}) the very same way as it is present in the Λ_c^+ from $\Sigma_c^{0,++}$ decays. The differences in trends between Λ_c^+ and $\Sigma_c^{0,++}$ are highlighted in the baryon-to-baryon ratios of Fig. 7.5 (right). Since this ratio is also ordered by N_{MPI}, we can exclude the effect of decay kinematics.

In Fig. 7.6 I show the ratios of the yields of the strange charmed baryon $\Xi_c^{0,+}$ (left panel) and the double-strange charmed baryon Ω_c^0 (right panel) over D^0 as a function of p_T in terms of different N_{MPI} classes. The trends are generally similar to Λ_c^+/D^0 for both $\Xi_c^{0,+}$ and Ω_c^0. It is to be noted however that, while all ratios fall with p_T, the steepness of the trends are different: while in the range p_T = 2 GeV/c to p_T =10 GeV/c the value of Λ_c^+/D^0 falls with a factor of ≈ 3, and $\Sigma_c^{0,++}/D^0$ decreases with about a factor of ≈ 4, this decrease is only about a factor of ≈ 2 in the case of $\Xi_c^{0,+}$ and the Ω_c^0.

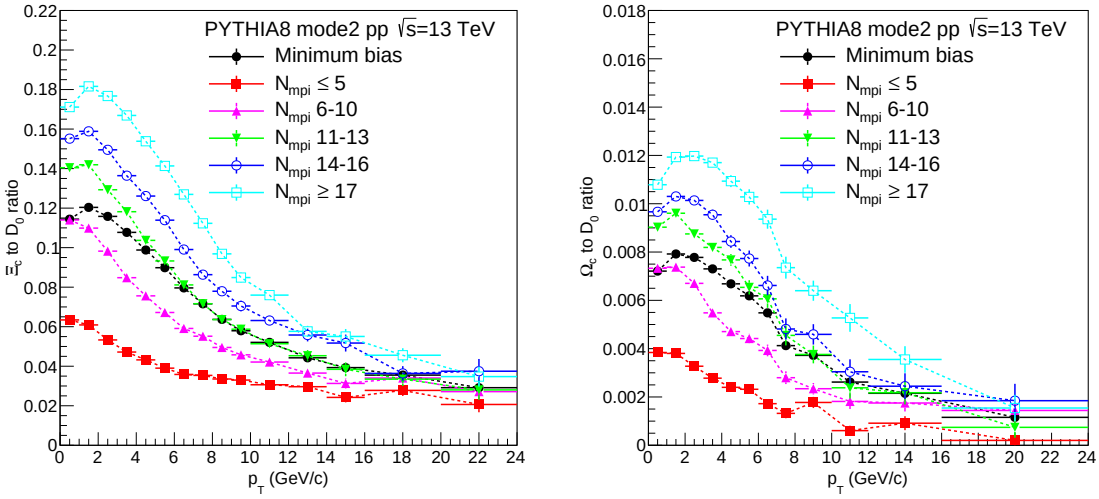

Figure 7.6: $\Xi_c^{0,+}/D^0$ (*left*) and Ω_c^0/D^0 (*right*) ratios from PYTHIA 8 simulations with CR-BLC as a function of p_T, for MB events as well as for different multiparton-interaction (N_{MPI}) classes.

Considering that for the Λ_c^+ there is a significant feed-down from $\Xi_c^{0,+}$ and therefore the result is expected to be a mixture of direct Λ_c^+ and those coming from $\Xi_c^{0,+}$, this pattern can be attributed to the presence or lack of strange content.

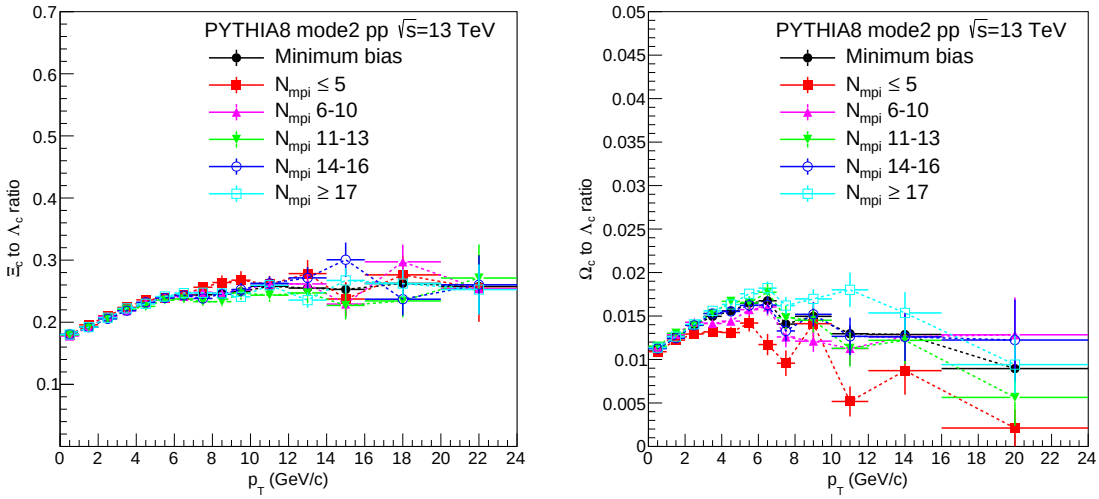

Figure 7.7: $\Xi_c^{0,+}/\Lambda_c^+$ (*left*) and Ω_c^0/Λ_c^+ (*right*) ratios from PYTHIA 8 simulations with CR-BLC as a function of p_T, for MB events as well as for different multiparton-interaction (N_{MPI}) classes.

Fig. 7.7 (left panel) shows the strange-to-non-strange charmed baryon ratio $\Xi_c^{0,+}/\Lambda_c^+$ as a function of p_T for several N_{MPI} classes. Most notable in the figure is that there is no significant N_{MPI} ordering, indicating that the event-activity-dependent production is predominantly connected to the charm content and it is not affected by strangeness content. The slight dependence on p_T can be the consequence of the different masses and an overall relative suppression of strangeness production at low p_T. It is also to be noted that the N_{MPI}-dependent enhancement in the strange baryon-to-meson ratio Λ^0/K_S^0 is predicted by PYTHIA 8 alone, albeit to a smaller extent compared to PYTHIA with CR-BLC. In Fig. 7.7 (right panel) I show the double-strange-to-non-strange charmed baryon ratio Ω_c^0/Λ_c^+ as a function of p_T for several N_{MPI} classes. While for smaller p_T values the trend looks similar to that observed for the $\Xi_c^{0,+}/\Lambda_c^+$, at higher p_T values there is a clear separation between the MPI classes despite the statistical fluctuations. Further studies are needed to decide whether the observed trend can be attributed to the difference in the strange content between $\Xi_c^{0,+}$ and Ω_c^0.

In the following I investigate some distinctive experimental signatures of the enhancement of different charm baryons recapitulated above. To focus on the differences caused by the isospin and strangeness content, we can look at the baryon-to-baryon ratios. Following the method outlined in Sec.7.2 I take the integral of the ratios in the semi-soft (coalescence) regime $2 < p_T < 8$ GeV/c. To characterize the UE and the jettiness of an event, I use a single variable, spherocity, in minimum-bias data, as well as the R_{NC} and R_T variables in events where a high-momentum trigger hadron is present.

In Fig. 7.8 I show the $\Sigma_c^{0,++}/\Lambda_c^+$ (left), $\Xi_c^{0,+}/\Lambda_c^+$ (center) and Ω_c^0/Λ_c^+ (right) ratios integrated over the coalescence regime, in fixed N_{ch} ranges, for different S_0 classes. Using fixed multiplicity windows reduces the bias from the correlation of jet production with multiplicity.

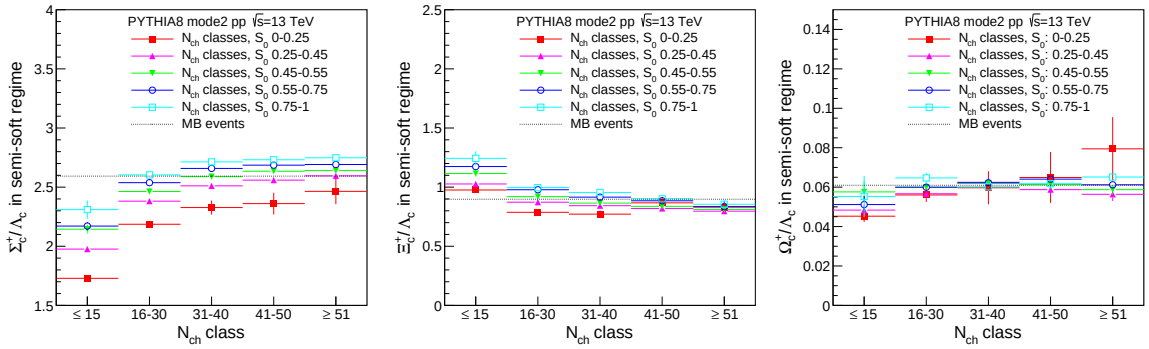

Figure 7.8: Charmed baryon-to-baryon ratios integrated over $2 < p_T < 8$ GeV/c, in fixed N_{ch} ranges, for different S_0 classes (colored curves). $\Sigma_c^{0,++}/\Lambda_c^+$ is shown in the *left*, $\Xi_c^{0,+}/\Lambda_c^+$ in the *center* and Ω_c^0/Λ_c^+ in the *right* panel.

In case of $\Sigma_c^{0,++}/\Lambda_c^+$, the ratio consistently depends on S_0 in all $N_{\rm ch}$ classes. This is consistent with the pattern observed in the case of $N_{\rm MPI}$ and hints that the enhanced charm-baryon production is sensitive to the isospin. On the other hand, strangeness content has only a slight effect in the semi-soft (coalescence) regime. Note that different decay topologies of higher-mass states may have an effect in the lowest $N_{\rm ch}$ range where the UE is small.

In Fig. 7.9 I plot the integrated $\Sigma_c^{0,++}/\Lambda_c^+$ (left), $\Xi_c^{0,+}/\Lambda_c^+$ (center) and Ω_c^0/Λ_c^+ (right) ratios in the $R_{\rm T}$ and $R_{\rm NC}$ classes. While there is virtually no $R_{\rm NC}$ dependence for the

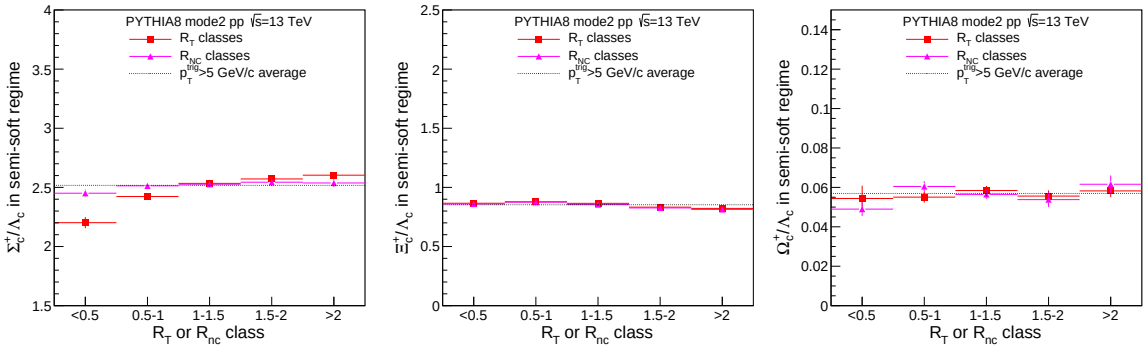

Figure 7.9: Charmed baryon-to-baryon ratios integrated over $2 < p_{\rm T} < 8$ GeV/c, for different $R_{\rm T}$ (red) and $R_{\rm NC}$ (magenta) classes. $\Sigma_c^{0,++}/\Lambda_c^+$ is shown in the *left*, $\Xi_c^{0,+}/\Lambda_c^+$ in the *center* and Ω_c^0/Λ_c^+ in the *right* panel.

$\Sigma_c^{0,++}/\Lambda_c^+$ ratio, the $R_{\rm T}$-dependence is significant. This corroborates the observation that in the PYTHIA 8 CR-BLC model the enhancement is primarily linked to the UE and not to the jet production. In case of the strange to non-strange charmed baryon ratios, no dependence is observed in either case.

7.4 Summary

Recent observations of low-$p_{\rm T}$ enhancement of charmed-baryon production relative to charmed mesons in proton–proton collisions at LHC energies question the universality of charm fragmentation and provide a challenge for our present theoretical models.

In this chapter I demonstrated that appropriately defined event-activity classifiers provide great sensitivity to the charm production mechanisms of the Λ_c^+ baryons in proton–proton collisions at LHC energies. Utilizing PYTHIA 8 simulations with color-reconnection beyond leading color approximation, I found that the $\Lambda_c^+/{\rm D}^0$ yield ratio in hadron-triggered events shows a pronounced dependence on the transverse-event-activity classifier $R_{\rm T}$, but shows no significant dependence on the $R_{\rm NC}$ transverse-event-activity classifier which reflects event activity related to the leading hard process. Considering that in PYTHIA 8, $R_{\rm T}$ is correlated with the number of MPIs of an event, and that by construction $R_{\rm NC}$ reflects

the event activity related to the leading hard process, this provides the means to distinguish between these two sources.

I also showed that the $\Lambda_{\rm c}^{+}/{\rm D}^{0}$ yield ratio depends on the transverse spherocity S_0 in events with sufficiently high final-state multiplicity. Since collision events characterized by low values of transverse spherocity S_0 are more jet-like and events with high S_0 in the same multiplicity class are more isotropic, this variable has sensitivity to whether the excess $\Lambda_{\rm c}^{+}$ production is linked to the underlying event or linked to the jet processes.

I found that the usage of flattenicity, a new quantity to represent multiple-parton interactions, is even more distinctive and helps pin down which type of events the enhancement stems from. These observations therefore show that in the scenario implemented by PYTHIA 8 with CR-BLC Mode 2 the excess $\Lambda_{\rm c}^{+}$ production is primarily linked to the underlying event activity and not to the jet production.

Here I also note that simulations of the low-$p_{\rm T}$ charmed-baryon enhancement as a function of event multiplicity in the central and the forward pseudorapidity regions show very similar trends. Since the multiplicity at forward pseudorapidity is much less connected to jet production than that in the central range due to the gap in the phase space between the two regions, this also suggests that in the CR-BLC scenario the $\Lambda_{\rm c}^{+}/{\rm D}^{0}$ enhancement is driven by mechanisms other than charm-production in the jets.

In the model class considering CR-BLC, $\Lambda_{\rm c}^{+}$ is sensitive to the underlying-event activity. While this sensitivity is present both in directly produced as well as decay $\Lambda_{\rm c}^{+}$ baryons, the excess is dominated by the decay contribution at low $p_{\rm T}$ ($\lesssim 8$ GeV/c), and by the direct production at higher $p_{\rm T}$ ($\gtrsim 16$ GeV/c) values. This is not a consequence of decay kinematics. The comparison of $\Sigma_{\rm c}^{0,++}$ and $\Lambda_{\rm c}^{+}$ enhancement suggests that isospin-dependent effects are present and they are linked to the formation of charm baryons via color junctions. While strangeness enhancement itself is reproduced by models without color junctions, the enhanced production of charm baryons (either with strange or non-strange content) requires color reconnection beyond leading order. In both the strange and non-strange cases, charm baryon enhancement comes from the underlying event and not from the jet region. While the isospin-dependent effects seem to be linked to the formation of charm baryons via color junctions, results on the $\Xi_{\rm c}^{0,+}$ implies that strangeness does not play a strong role in the enhancement of charmed baryons, although the case of the double-strange $\Omega_{\rm c}^{0}$ needs further investigation.

The ongoing Run 3 data taking period at the LHC and beyond will allow for differential measurements of charmed-hadron production with an unprecedented precision. The comparison of $\Lambda_{\rm c}^{+}/{\rm D}^{0}$ ratios in hadron-triggered data as a function of $R_{\rm T}$ and $R_{\rm NC}$, as well as the double-differential evaluation of minimum-bias data as a function of S_0 and $N_{\rm ch}$, and as a function of ρ and $N_{\rm ch}$, opens up the possibility to differentiate between competing scenarios that describe flavor-dependent hadron production in the underlying event and within jets. The LHC Run 3 data taking period will also allow to differentiate between mechanisms of strangeness and charm enhancement.

Chapter 8

Summary

The research presented in this thesis primarily focuses on the exploration of the boundary between hard and soft processes that occur in proton–proton collisions using light- and heavy flavor production. This is a relatively unexplored area despite its significance. One of the specific aims was to study jet fragmentation and hadronization properties in high-energy proton–proton collisions to shed light on the particle production mechanisms that lead to collective-like behavior in small systems, and may help provide a better explanation of observed collective phenomena in high-multiplicity proton–proton collisions, that do not include the quark-gluon plasma. In this chapter I summarize the main results and conclusions, while also providing an outlook on future research.

One of the main areas of my investigations was the study of jets. I measured differential and integral jet shapes in simulations and found modifications caused by non-trivial quantum chromodynamics effects, including a clear multiplicity (number of charged final state particles) dependence, which could be experimentally measured with the methods proposed in this thesis. Based on simulations with different MC event generators, I introduced a characteristic jet size observable r_{ch} that depends only on the $p_{\mathrm{T}}^{\mathrm{jet}}$ but is independent of the jet reconstruction algorithms, parton density functions, and even the choice of simulation parameters such as color reconnection and multiple-parton interactions. Its p_{T}-dependence qualitatively follows a Lorentz-boost curve. In a novel jet shape analysis, I introduced a double ratio of the differential jet shapes to measure the multiplicity dependence with minimizing biases. These observations suggest that r_{ch} is an inherent property of jets that is characteristic to the spatial development of the parton shower at a given momentum.

I compared the multiplicity dependence of jet structure variables for multiple popular PYTHIA 8 tunes and also for different MPI and CR models in several p_{T} bins. I found that the evolution of the differential jet structure $\rho(r)$ with multiplicity significantly differs in several $p_{\mathrm{T}}^{\mathrm{jet}}$ ranges for the Monash, MonashStar, and 4C tunes. The shape of the difference is nontrivial in $p_{\mathrm{T}}^{\mathrm{jet}}$, but persistent through all tested choices of multiplicity selections.

With this I demonstrated that the multiplicity-dependent analysis of jet momentum profiles can differentiate among otherwise well-established models. This lack of understanding may have grave consequences on studies based on classification by jet properties. My observations highlight the need of extending multiplicity-dependent jet structure measurements to higher $p_{\mathrm{T}}^{\mathrm{jet}}$ regimes.

My studies of the integrated jet structure variable Ψ shows a rather different N_{ch} dependence when MPI are turned off. This attests to the important role of multiple-parton interactions in higher multiplicity events and the need for their detailed understanding in order to develop accurate models in jet physics. The results serve as motivation for future experimental measurements.

My investigations also included the multiplicity distributions as a function of the jet transverse momenta. Recent results show that the multiplicity distributions follow a scaling similar to the Koba-Nielsen-Olesen (KNO) scaling, which provides important lessons on jet fragmentation. I presented results on the scaling properties of heavy-flavor jets from different production processes, and compared them to those on inclusive jets. I used PYTHIA 8 simulations to evaluate the charged-hadron event multiplicities at central pseudorapidity, as a function of the charged-particle jet transverse momentum within a wide jet transverse momentum range. I found that the multiplicity distributions satisfy a KNO-like scaling with $p_{\mathrm{T}}^{\mathrm{jet}}$ for both charm and beauty jets similarly to what has been observed for inclusive jets. I also found that the multiplicity distributions in events with jets initiated by charm and beauty directly from the leading hard process show some departure from the negative binomial shape, depending on the $p_{\mathrm{T}}^{\mathrm{jet}}$.

Further analysis of the statistical moments of the multiplicity distributions shows that the scaling is fulfilled within $\approx$5% throughout the full $p_{\mathrm{T}}^{\mathrm{jet}}$ range, but the deviations are more significant for leading-order heavy flavor creation, especially in the case of beauty. On the other hand, beauty production from the parton shower tends to deviate less from scaling expectations and follows the inclusive-jet trend within uncertainties. We conclude therefore that the KNO-like scaling originates from the parton level of the early stages of the collision, and not from the later stages of parton shower or jet fragmentation.

A good description of hadron multiplicity distributions is a basic requirement for models and it is generally fulfilled by the most widely used event generators. However, multiplicities as a function of the jet momentum for jets tagged with different flavors can provide means to further validate heavy-flavor production and fragmentation models. Also, while event multiplicity is a good proxy for jet multiplicity in case of jets coming from the leading hard process, this is not necessarily the case for jets that come from secondary hard processes or gluon radiation. An interesting extension of the current work in this direction could therefore be to evaluate the scaling in terms of the jet multiplicity instead of event multiplicity, and to see whether in that case the scaling of heavy flavor jets from the parton shower follows light or heavy jets.

The search for the KNO-like scaling in simulations for heavy-flavor jets served as a motivation to carry out the first measurement of jet-momentum-dependent jet multiplicity distributions with the ALICE experiment and measure the KNO-like scaling. The jet multiplicity distributions can be all collapsed onto a universal scaling curve, however calculating the first nine moments of the distributions, I found an approximately 20% discrepancy toward higher statistical momenta, which is not very significant considering the statistical errors. This may indicate that the KNO-like scaling was seen in simulations only because of the specific multiplicity distributions of the particular applied model. It is to be noted however, that the current data is not completely corrected: the background subtraction is yet to be implemented. This may influence the observed multiplicity distributions.

In high-energy hadron collisions, heavy quarks (charm and beauty) are mainly produced in hard parton scattering processes. Two-particle angular correlations originating from heavy-flavor particles allow for the characterization of parton shower and fragmentation. The ALICE collaboration measured the heavy-flavor electron-hadron azimuthal correlation distributions between heavy-flavor decay electrons and associated charged particles in proton–proton and p–Pb collisions at $\sqrt{s} = 5.02$ TeV. I created detailed simulations to compare the near- and away-side peaks of the azimuthal-correlation distribution in pp and p–Pb collisions to the model predictions. This allowed verifying the implementation of the processes of charm- and beauty-quark production, fragmentation, and hadronization, which have an impact on the observables studied in this analysis. The correlation structures are fitted with a constant and two von Mises functions to obtain the baseline and the near- and away-side peaks, respectively. The evolution of the near- and away-side peaks of the correlation functions in pp and p–Pb collisions was found to be similar in all the considered kinematic ranges. This suggests that the modification of the fragmentation and hadronization of heavy quarks due to cold-nuclear-matter effects is indistinguishable within the current precision of the measurements.

Another major topic in my thesis, connected to jet fragmentation, concerns the understanding of heavy-flavor hadroproduction in proton–proton collisions at LHC energies. The production cross section of hadrons can be calculated using the factorization theorem, which usually assumes that the fragmentation functions are universal across different collision systems. Recent observations of low-p_{T} enhancement of charmed-baryon production relative to charmed mesons in proton–proton collisions at LHC energies questioned the universality of charm fragmentation and provide a challenge for our present theoretical models.

I used a model with color reconnection beyond leading color (CR-BLC) approximation to seek explanation for the charm-baryon enhancement, and proposed new observables for future measurements. I characterized the collision events using different event-activity classifiers, that allow for investigating the connections between the leading QCD processes and the underlying event. I demonstrated that appropriately defined event-activity classifiers provide great sensitivity to the charm production mechanisms of the Λ_{c}^{+} baryons.

Utilizing PYTHIA 8 simulations with enhanced color-reconnection I found that the $\Lambda_{\rm c}^{+}/{\rm D}^{0}$ yield ratio in hadron-triggered events shows a pronounced dependence on the transverse-event-activity classifier $R_{\rm T}$, but shows no significant dependence on the $R_{\rm NC}$ classifier which reflects event activity related to the leading hard process. Considering that in PYTHIA 8, $R_{\rm T}$ is correlated with the number of MPI of an event, and that by construction $R_{\rm NC}$ reflects the event activity related to the leading hard process, this provides the means to distinguish between these two sources.

I also showed that the $\Lambda_{\rm c}^{+}/{\rm D}^{0}$ yield ratio significantly depends on the transverse spherocity S_0 in events with sufficiently high final-state multiplicity. Since collision events characterized by low values of transverse spherocity S_0 are more jet-like and events with high S_0 in the same multiplicity class are more isotropic, this variable has sensitivity to whether the excess Λ_c^+ production is linked to the underlying event or linked to the jet processes.

I found that the usage of flattenicity, a new quantity to represent multiple-parton interactions, is even more distinctive and helps pin down which type of events the enhancement stems from. These observations therefore show, that in the scenario implemented by PYTHIA 8 with enhanced color reconnection Mode 2, the excess $\Lambda_{\rm c}^{+}$ production is primarily linked to the underlying event activity and not to the jet production.

The comparison of $\Sigma_{\rm c}^{0,++}$ and $\Lambda_{\rm c}^{+}$ enhancement suggests that isospin-dependent effects are present and they are linked to the formation of charm baryons via color junctions. While strangeness enhancement itself is reproduced by models without color junctions, the enhanced production of charm baryons (either with strange or non-strange content) requires color reconnection beyond leading order. In both the strange and non-strange cases, charm baryon enhancement comes from the underlying event and not from the jet region.

While the isospin-dependent effects seem to be linked to the formation of charm baryons via color junctions, results on the $\Xi_{\rm c}^{0,+}$ implies that strangeness does not play a strong role in the enhancement of charmed baryons, although the case of the double-strange $\Omega_{\rm c}^{0}$ needs further investigation.

The ongoing Run 3 data taking period at the LHC and beyond will allow for differential measurements of charmed-hadron production with an unprecedented precision. The comparison of $\Lambda_{\rm c}^{+}/{\rm D}^{0}$ ratios in hadron-triggered data as a function of $R_{\rm T}$ and $R_{\rm NC}$, as well as the double-differential evaluation of minimum-bias data as a function of S_0 and $N_{\rm ch}$, and as a function of ρ and $N_{\rm ch}$, opens up the possibility to differentiate between competing scenarios that describe flavor-dependent hadron production in the underlying event and within jets. The LHC Run 3 data taking period will also allow to differentiate between mechanisms of strangeness and charm enhancement.

Thesis Statements

This thesis presents my research on the fragmentation properties of light and heavy-flavor hadrons, covering jet shapes, multiplicity distribution scaling, azimuthal correlations of heavy-flavor decay electrons and hadrons, and charmed baryon production as a function of the event activity. In the following thesis statements I summarize my novel contributions to each of these subjects.

1. Multiplicity-dependent Jet Structure and Fragmentation

I analyzed the differential and integral jet shapes and performed an event multiplicity-differential study in proton-proton collisions created with the PYTHIA event generator. By utilizing a double ratio of observables to minimize the bias in the multiplicity measurement, I showed that the jet shapes depend on the event multiplicity, which can be used for model differentiation in experimental data. By looking at the multiplicity dependence of the jet shape distributions, I observed a characteristic jet size, which is robust against different parton density functions, jet reconstruction algorithms and multiplicity selections. It qualitatively follows a Lorentz-boost curve, which suggests it being an inherent property of the jets and is characteristic to the space-time evolution of the parton shower at a given momentum [187, 226, 227].

2. Scaling Properties of Jet Structure in Theory and the ALICE Experiment

The Koba-Nielsen-Olesen (KNO) scaling hypothesis is an influential contribution to the analysis of event multiplicities in high-energy particle collisions, according to which the event-multiplicity distributions can be all collapsed onto a universal scaling curve. Recent phenomenological studies suggest that a similar scaling may hold within single jets, if we consider the jet multiplicity as a function of the jet transverse momentum. I conducted an analysis on the KNO-like scaling for heavy-flavor jets in pp collisions at $\sqrt{s} = 13$ TeV using Monte Carlo event generators. I found that the KNO-like scaling stems from the partonic level of the interaction. Motivated by the results, I conducted the first measurement

of the jet multiplicity distributions as a function of jet transverse momentum in pp collisions in the ALICE experiment at $\sqrt{s} = 13$ TeV, which allowed to quantify the KNO-like jet scaling properties and thus help further our understanding of jet fragmentation properties [193, 228].

3. Azimuthal Correlations of Heavy-flavor Decay Electrons with the ALICE Experiment

In high-energy hadron collisions, heavy quarks (charm and beauty) are mainly produced in hard parton scattering processes. Two-particle angular correlations originating from heavy-flavor particles allow for the characterization of parton shower and fragmentation. The ALICE collaboration measured the heavy-flavor electron-hadron azimuthal correlation distributions between heavy-flavor decay electrons and associated charged particles in pp and p–Pb collisions at $\sqrt{s} = 5.02$ TeV. My main contribution to the analysis was to create detailed simulations to compare the near- and away-side peaks of the azimuthal-correlation distribution in pp and p–Pb collisions to the model predictions. This allowed verifying the model implementation of the processes of charm- and beauty-quark production, fragmentation, and hadronization, which have an impact on the observables studied in this analysis. I also determined the correlation peak shape using FONLL pQCD calculations for the modelling of charm and beauty contributions. The correlation structures are fitted with a constant and two von Mises functions to obtain the baseline and the near- and away-side peaks, respectively. The evolution of the near- and away-side peaks of the correlation functions in pp and p–Pb collisions was found to be similar in all the considered kinematic ranges. This suggests that the modification of the fragmentation and hadronization of heavy quarks due to cold-nuclear-matter effects is indistinguishable within the current precision of the measurements [108].

4. Charm-baryon Enhancement and the Role of the Underlying Event

Perturbative quantum chromodynamics (pQCD) calculations have been successful in describing the production of heavy-flavor mesons for several collision energies at the LHC. The usual description relies on the factorization approach, in which the production cross section of heavy-flavor hadrons in the hadronic collisions is calculated as a convolution of the parton density functions (PDFs) of the colliding hadrons, the cross section of the hard-scattering process and the heavy-quark fragmentation function. However, recent experimental results from the CERN LHC show a relative enhancement of charmed baryons compared to the factorization approach expectations based on electron-positron collisions. I utilized the PYTHIA 8 Monte Carlo event generator with color reconnection beyond leading-color approximation and proposed experimental methods based on event-activity classifiers to probe the source of the charmed-baryon enhancement. I concluded, that in

the considered model class the Λ_c^+ enhancement is connected to the underlying event and does not depend significantly on the processes inside the jet region [125, 229].

5. Production of Excited Charm and Charm-strange Baryon States

I extended the studies of the Λ_c^+ baryon enhancement to several different charmed baryons and I also investigated the production of charmed baryons with different isospin and strangeness content, then compared it to both charmed D^0 mesons and to Λ_c^+ baryons in pp collisions at LHC energies. I showed that the isospin of the charmed-baryon state has a strong impact on the enhancement pattern. Using the observables I propose, upcoming high-precision experimental data will be able to differentiate between mechanisms of strangeness and charm enhancement [230, 231].

Bibliography

[1] A.B. Arbuzov. Quantum Field Theory and the Electroweak Standard Model. pages 1–34, 2018. 35 pages. 5

[2] Sidney Coleman. Lectures of Sidney Coleman on Quantum Field Theory. WSP, Hackensack, 12 2018. 7

[3] Michael E. Peskin and Daniel V. Schroeder. An Introduction to quantum field theory. Addison-Wesley, Reading, USA, 1995.

[4] Matthew D. Schwartz. Quantum Field Theory and the Standard Model. Cambridge University Press, 3 2014.

[5] Dezső Horváth and Zoltán Trócsányi. Introduction to Particle Physics. Cambridge Scholars Publishing, Newcastle upon Tyne, UK, 2019. 7

[6] Jeffrey R. Forshaw and D. A. Ross. Quantum Chromodynamics and the Pomeron, volume 9. Oxford University Press, 1998. 8

[7] Hannah Elfner and Berndt Müller. The exploration of hot and dense nuclear matter: introduction to relativistic heavy-ion physics. J. Phys. G, 50(10):103001, 2023. 9

[8] Jana N. Guenther. Overview of the QCD phase diagram: Recent progress from the lattice. Eur. Phys. J. A, 57(4):136, 2021. 9

[9] S. Durr et al. [BMW Collaboration]. Ab-Initio Determination of Light Hadron Masses. Science, 322:1224–1227, 2008. 10

[10] William H. Press, Saul A. Teukolsky, William T. Vetterling, and Brian P. Flannery. Numerical Recipes 3rd Edition: The Art of Scientific Computing. Cambridge University Press, 3 edition, 2007. 12

[11] T. Kloek and H. K. van Dijk. Bayesian estimates of equation system parameters: An application of integration by monte carlo. Econometrica, 46(1):1–19, 1978. 12

[12] Torbjörn Sjöstrand, Stefan Ask, Jesper R. Christiansen, Richard Corke, Nishita Desai, Philip Ilten, Stephen Mrenna, Stefan Prestel, Christine O. Rasmussen, and Peter Z. Skands. An introduction to PYTHIA 8.2. Comput. Phys. Commun., 191:159–177, 2015. 13, 43, 44, 61, 106

[13] M. Bahr et al. Herwig++ Physics and Manual. Eur. Phys. J. C, 58:639–707, 2008. 13

[14] Johannes Bellm et al. Herwig 7.0/Herwig++ 3.0 release note. Eur. Phys. J. C, 76(4):196, 2016. 13, 85

[15] Enrico Bothmann et al. Event Generation with Sherpa 2.2. SciPost Phys., 7(3):034, 2019. 13

[16] Paolo Nason. A New method for combining NLO QCD with shower Monte Carlo algorithms. JHEP, 11:040, 2004. 13, 85

[17] Simone Alioli, Paolo Nason, Carlo Oleari, and Emanuele Re. A general framework for implementing NLO calculations in shower Monte Carlo programs: the POWHEG BOX. JHEP, 06:043, 2010.

[18] Stefano Frixione, Paolo Nason, and Carlo Oleari. Matching NLO QCD computations with Parton Shower simulations: the POWHEG method. JHEP, 11:070, 2007. 13, 85

[19] K. Werner, B. Guiot, Iu. Karpenko, and T. Pierog. Analysing radial flow features in p-Pb and p-p collisions at several TeV by studying identified particle production in EPOS3. Phys. Rev. C, 89(6):064903, 2014. 13

[20] Richard D. Ball, Valerio Bertone, Stefano Carrazza, Luigi Del Debbio, Stefano Forte, Alberto Guffanti, Nathan P. Hartland, and Juan Rojo. Parton distributions with QED corrections. Nucl. Phys. B, 877:290–320, 2013. 13

[21] Peter Skands, Stefano Carrazza, and Juan Rojo. Tuning PYTHIA 8.1: the Monash 2013 Tune. Eur. Phys. J. C, 74(8):3024, 2014. 13, 92, 106

[22] Vardan Khachatryan et al. [CMS Collaboration]. Event generator tunes obtained from underlying event and multiparton scattering measurements. Eur. Phys. J. C, 76(3):155, 2016. 13

[23] Richard Corke and Torbjörn Sjöstrand. Interleaved Parton Showers and Tuning Prospects. JHEP, 03:032, 2011. 13

[24] J. Pumplin, D. R. Stump, J. Huston, H. L. Lai, Pavel M. Nadolsky, and W. K. Tung. New generation of parton distributions with uncertainties from global QCD analysis. JHEP, 07:012, 2002. 13

[25] Andy Buckley et al. General-purpose event generators for LHC physics. Phys. Rept., 504:145–233, 2011. 13

[26] Bo Andersson, G. Gustafson, and B. Nilsson-Almqvist. A Model for Low p_{T} Hadronic Reactions, with Generalizations to Hadron-Nucleus and Nucleus-Nucleus Collisions. Nucl. Phys. B, 281:289–309, 1987. 13

[27] Hong Pi. An Event generator for interactions between hadrons and nuclei: FRITIOF version 7.0. Comput. Phys. Commun., 71:173–192, 1992. 13

[28] M. Alvioli and M. Strikman. Color fluctuation effects in proton-nucleus collisions. Phys. Lett. B, 722:347–354, 2013. 14

[29] Christian Bierlich, Gösta Gustafson, Leif Lönnblad, and Harsh Shah. The Angantyr model for Heavy-Ion Collisions in PYTHIA8. JHEP, 10:134, 2018. 14, 92

[30] Gábor Bíró, Gergely Gábor Barnaföldi, Gábor Papp, Miklos Gyulassy, Péter Lévai, Xin-Nian Wang, and Ben-Wei Zhang. Introducing HIJING++: the Heavy Ion Monte Carlo Generator for the High-Luminosity LHC Era. PoS, Hard-Probes2018:045, 2019. 14

[31] Gábor Papp, Gergely Gábor Barnaföldi, Gábor Bíró, Miklos Gyulassy, Szilveszter Harangozó, Guoyang Ma, Péter Lévai, Xin-Nian Wang, and Ben-Wei Zhang. First Results with HIJING++ on High-energy Heavy Ion Collisions. In 12th International Workshop on High-pT Physics in the RHIC/LHC Era, 5 2018. 14

[32] Xin-Nian Wang and Miklos Gyulassy. HIJING: A Monte Carlo model for multiple jet production in p p, p A and A A collisions. Phys. Rev. D, 44:3501–3516, 1991. 14, 89

[33] Wei-Tian Deng, Xin-Nian Wang, and Rong Xu. Hadron production in p+p, p+Pb, and Pb+Pb collisions with the HIJING 2.0 model at energies available at the CERN Large Hadron Collider. Phys. Rev. C, 83:014915, 2011. 14

[34] Antonio Ortiz, Gyula Bencédi, and Héctor Bello. Revealing the source of the radial flow patterns in proton–proton collisions using hard probes. J. Phys. G, 44(6):065001, 2017. 14, 25, 26

[35] Yuri L. Dokshitzer, Valery A. Khoze, and S. I. Troian. On specific QCD properties of heavy quark fragmentation ('dead cone'). J. Phys. G, 17:1602–1604, 1991. 15, 27, 29, 60

[36] Shreyasi Acharya et al. [ALICE Collaboration]. Direct observation of the dead-cone effect in quantum chromodynamics. Nature, 605(7910):440–446, 2022. 15, 27, 29, 60

[37] Gavin P. Salam. Towards Jetography. Eur. Phys. J. C, 67:637–686, 2010. 15

[38] Stephen D. Ellis and Davison E. Soper. Successive combination jet algorithm for hadron collisions. Phys. Rev. D, 48:3160–3166, 1993. 16, 45

[39] Yuri L. Dokshitzer, G. D. Leder, S. Moretti, and B. R. Webber. Better jet clustering algorithms. JHEP, 08:001, 1997. 16, 45

[40] Matteo Cacciari, Gavin P. Salam, and Gregory Soyez. The anti-k_t jet clustering algorithm. JHEP, 04:063, 2008. 16, 45, 62

[41] Matteo Cacciari, Gavin P. Salam, and Gregory Soyez. FastJet User Manual. Eur. Phys. J. C, 72:1896, 2012. 16, 45

[42] Serguei Chatrchyan et al. [CMS Collaboration]. Shape, Transverse Size, and Charged Hadron Multiplicity of Jets in pp Collisions at 7 TeV. JHEP, 06:160, 2012. 7, 17, 43, 45, 48

[43] Edward V. Shuryak. Quark-Gluon Plasma and Hadronic Production of Leptons, Photons and Psions. Phys. Lett. B, 78:150, 1978. 19

[44] J. D. Bjorken. Highly Relativistic Nucleus-Nucleus Collisions: The Central Rapidity Region. Phys. Rev. D, 27:140–151, 1983.

[45] Frithjof Karsch. Lattice results on QCD thermodynamics. Nucl. Phys. A, 698:199–208, 2002. 19

[46] Edward V. Shuryak. Quantum Chromodynamics and the Theory of Superdense Matter. Phys. Rept., 61:71–158, 1980. 19

[47] K. Adcox et al. [PHENIX Collaboration]. Formation of dense partonic matter in relativistic nucleus-nucleus collisions at RHIC: Experimental evaluation by the PHENIX collaboration. Nucl. Phys. A, 757:184–283, 2005. 19

[48] John Adams et al. [STAR Collaboration]. Experimental and theoretical challenges in the search for the quark gluon plasma: The STAR Collaboration's critical assessment of the evidence from RHIC collisions. Nucl. Phys. A, 757:102–183, 2005. 21

[49] B. B. Back et al. [PHOBOS Collaboration]. The PHOBOS perspective on discoveries at RHIC. Nucl. Phys. A, 757:28–101, 2005.

[50] I. Arsene et al. [BRAHMS Collaboration]. Quark gluon plasma and color glass condensate at RHIC? The Perspective from the BRAHMS experiment. Nucl. Phys. A, 757:1–27, 2005. 19

[51] David d'Enterria and Barbara Betz. High-p(T) hadron suppression and jet quenching. Lect. Notes Phys., 785:285–339, 2010. 19

[52] K. Adcox et al. [PHENIX Collaboration]. Suppression of hadrons with large transverse momentum in central Au+Au collisions at $\sqrt{s_{NN}}$ = 130-GeV. Phys. Rev. Lett., 88:022301, 2002. 19

[53] Serguei Chatrchyan et al. [CMS Collaboration]. Study of High-pT Charged Particle Suppression in PbPb Compared to pp Collisions at $\sqrt{s_{NN}} = 2.76$ TeV. Eur. Phys. J. C, 72:1945, 2012. 20

[54] J. Adams et al. [STAR Collaboration]. Azimuthal anisotropy and correlations at large transverse momenta in p+p and Au+Au collisions at s(NN)**(1/2) = 200-GeV. Phys. Rev. Lett., 93:252301, 2004. 21

[55] S. S. Adler et al. [PHENIX Collaboration]. Absence of suppression in particle production at large transverse momentum in S(NN)**(1/2) = 200-GeV d + Au collisions. Phys. Rev. Lett., 91:072303, 2003. 21

[56] S. S. Adler et al. [PHENIX Collaboration]. Centrality dependence of direct photon production in s(NN)**(1/ 2) = 200-GeV Au + Au collisions. Phys. Rev. Lett., 94:232301, 2005. 22

[57] Madan M. Aggarwal. Elliptic flow in relativistic heavy-ion collisions. In Rajeev K. Puri, Joerg Aichelin, Sakshi Gautam, and Rohit Kumar, editors, Advances in Nuclear Physics, pages 161–188, Singapore, 2021. Springer Singapore. 23

[58] S. Afanasiev et al. [PHENIX Collaboration]. Elliptic flow for phi mesons and (anti)deuterons in Au + Au collisions at s(NN)**(1/2) = 200-GeV. Phys. Rev. Lett., 99:052301, 2007. 23

[59] A. Adare et al. [PHENIX Collaboration]. Enhanced production of direct photons in Au+Au collisions at $\sqrt{s_{NN}} = 200$ GeV and implications for the initial temperature. Phys. Rev. Lett., 104:132301, 2010. 24

[60] R. Hagedorn. Statistical thermodynamics of strong interactions at high-energies. Nuovo Cim. Suppl., 3:147–186, 1965. 24

[61] Li Yan and Jean-Yves Ollitrault. Universal fluctuation-driven eccentricities in proton-proton, proton-nucleus, and nucleus-nucleus collisions. Phys. Rev. Lett., 112:082301, Feb 2014. 25

[62] Vardan Khachatryan et al. [CMS Collaboration]. Observation of Long-Range Near-Side Angular Correlations in Proton-Proton Collisions at the LHC. JHEP, 09:091, 2010. 25

[63] Björn Schenke, Soeren Schlichting, Prithwish Tribedy, and Raju Venugopalan. Mass ordering of spectra from fragmentation of saturated gluon states in high multiplicity proton-proton collisions. Phys. Rev. Lett., 117(16):162301, 2016. 25

[64] M. Gyulassy, P. Levai, I. Vitev, and T. S. Biro. Non-Abelian Bremsstrahlung and Azimuthal Asymmetries in High Energy $p + A$ Reactions. Phys. Rev. D, 90(5):054025, 2014. 25

[65] J. Adam et al. [ALICE Collaboration]. Measurement of D-meson production versus multiplicity in p-Pb collisions at $\sqrt{s_{NN}} = 5.02$ TeV. JHEP, 08:078, 2016. 25

[66] S. Schlichting. Initial state and pre-equilibrium effects in small systems. Nucl. Phys. A, 956:216–221, 2016. 26

[67] K. J. Eskola. Minijets in ultrarelativistic heavy ion collisions at future colliders. Comments Nucl. Part. Phys., 22(4):185–203, 1998. 26

[68] A. Andronic et al. Heavy-flavour and quarkonium production in the LHC era: from proton–proton to heavy-ion collisions. Eur. Phys. J. C, 76(3):107, 2016. 26

[69] X. Zhu, M. Bleicher, S. L. Huang, K. Schweda, Horst Stoecker, N. Xu, and P. Zhuang. D anti-D correlations as a sensitive probe for thermalization in high-energy nuclear collisions. Phys. Lett. B, 647:366–370, 2007. 26

[70] E. Norrbin and T. Sjostrand. Production and hadronization of heavy quarks. Eur. Phys. J. C, 17:137–161, 2000. 26

[71] Philip Ilten, Nicholas L. Rodd, Jesse Thaler, and Mike Williams. Disentangling Heavy Flavor at Colliders. Phys. Rev. D, 96(5):054019, 2017. 26, 60

[72] Róbert Vértesi, Gyula Bencédi, Anett Misák, and Antonio Ortiz. Probing the interaction of semi-hard quarks and gluons with the underlying event in light- and heavy-flavor triggered proton-proton collisions. Eur. Phys. J. A, 57(10):301, 2021. 26

[73] Bernd A. Kniehl, Gustav Kramer, Ingo Schienbein, and Hubert Spiesberger. Finite-mass effects on inclusive B meson hadroproduction. Phys. Rev. D, 77:014011, 2008. 27

[74] M. Cacciari, S. Frixione, M. L. Mangano, P. Nason, and G. Ridolfi. QCD analysis of first b cross-section data at 1.96 TeV. JHEP, 07:033, 2004.

[75] B. A. Kniehl, G. Kramer, I. Schienbein, and H. Spiesberger. Collinear subtractions in hadroproduction of heavy quarks. Eur. Phys. J. C, 41:199–212, 2005.

[76] Matteo Cacciari and Paolo Nason. Charm cross-sections for the Tevatron Run II. JHEP, 09:006, 2003.

[77] [ALICE Collaboration]. The ALICE experiment – A journey through QCD. 11 2022. 27, 40

[78] David d'Enterria. Jet quenching. Landolt-Bornstein, 23:471, 2010. 28

[79] Shreyasi Acharya et al. [ALICE Collaboration]. Measurement of electrons from semileptonic heavy-flavour hadron decays at midrapidity in pp and Pb-Pb collisions at $\sqrt{s_{\rm NN}}$ = 5.02 TeV. Phys. Lett. B, 804:135377, 2020. 28, 90

[80] Albert M Sirunyan et al. [CMS Collaboration]. Nuclear modification factor of $\rm D^0$ mesons in PbPb collisions at $\sqrt{s_{\rm NN}} = 5.02$ TeV. Phys. Lett. B, 782:474–496, 2018.

[81] Roel Aaij et al. [LHCb Collaboration]. Measurements of prompt charm production cross-sections in pp collisions at $\sqrt{s} = 13$ TeV. JHEP, 03:159, 2016. [Erratum: JHEP 09, 013 (2016), Erratum: JHEP 05, 074 (2017)].

[82] Shreyasi Acharya et al. [ALICE Collaboration]. Measurement of $\rm D^0$, $\rm D^+$, $\rm D^{*+}$ and $\rm D_s^+$ production in pp collisions at $\sqrt{s} = 5.02$ TeV with ALICE. Eur. Phys. J. C, 79(5):388, 2019.

[83] Georges Aad et al. [ATLAS Collaboration]. Measurement of the differential cross-section of B^+ meson production in pp collisions at $\sqrt{s}$ = 7 TeV at ATLAS. JHEP, 10:042, 2013.

[84] Serguei Chatrchyan et al. [CMS Collaboration]. Measurement of the B^0 production cross section in $\rm pp$ collisions at $\sqrt{s} = 7$ TeV. Phys. Rev. Lett., 106:252001, 2011.

[85] Vardan Khachatryan et al. [CMS Collaboration]. Measurement of the total and differential inclusive B^+ hadron cross sections in pp collisions at $\sqrt{s}$ = 13 TeV. Phys. Lett. B, 771:435–456, 2017. 28

[86] B. A. Kniehl, G. Kramer, I. Schienbein, and H. Spiesberger. Inclusive Charmed-Meson Production at the CERN LHC. Eur. Phys. J. C, 72:2082, 2012. 28

[87] B. A. Kniehl, G. Kramer, I. Schienbein, and H. Spiesberger. Inclusive B-Meson Production at the LHC in the GM-VFN Scheme. Phys. Rev. D, 84:094026, 2011.

[88] Matteo Cacciari, Stefano Frixione, Nicolas Houdeau, Michelangelo L. Mangano, Paolo Nason, and Giovanni Ridolfi. Theoretical predictions for charm and bottom production at the LHC. JHEP, 10:137, 2012. 28, 92

[89] Shreyasi Acharya et al. [ALICE Collaboration]. Measurement of the production of charm jets tagged with $\rm D^0$ mesons in pp collisions at $\sqrt{\rm s} = 7$ TeV. JHEP, 08:133, 2019. 29

[90] Shreyasi Acharya et al. [ALICE Collaboration]. Measurement of the production of charm jets tagged with $\rm D^0$ mesons in pp collisions at $\sqrt{s}$ = 5.02 and 13 TeV. JHEP, 06:133, 2023. 29

[91] Shreyasi Acharya et al. [ALICE Collaboration]. Measurements of Groomed-Jet Substructure of Charm Jets Tagged by D0 Mesons in Proton-Proton Collisions at s=13 TeV. Phys. Rev. Lett., 131(19):192301, 2023. 29

[92] Shreyasi Acharya et al. [ALICE Collaboration]. Measurement of inclusive charged-particle b-jet production in pp and p-Pb collisions at $\sqrt{s_{\rm NN}}$ = 5.02 TeV. JHEP, 01:178, 2022. 29, 62

[93] Jaroslav Adam et al. [ALICE Collaboration]. Measurement of azimuthal correlations of D mesons and charged particles in pp collisions at $\sqrt{s} = 7$ TeV and p-Pb collisions at $\sqrt{s_{\rm NN}} = 5.02$ TeV. Eur. Phys. J. C, 77(4):245, 2017. 29, 30, 92, 94

[94] Shreyasi Acharya et al. [ALICE Collaboration]. Azimuthal correlations of prompt D mesons with charged particles in pp and p–Pb collisions at $\sqrt{s_{NN}}$ = 5.02 TeV. Eur. Phys. J. C, 80(10):979, 2020. 30, 85, 92, 94

[95] Shreyasi Acharya et al. [ALICE Collaboration]. Investigating charm production and fragmentation via azimuthal correlations of prompt D mesons with charged particles in pp collisions at $\sqrt{\rm s} = \mathbf{13}$ TeV. Eur. Phys. J. C, 82(4):335, 2022. 29, 85

[96] Betty Bezverkhny Abelev et al. [ALICE Collaboration]. Beauty production in pp collisions at $\sqrt{s}$ = 2.76 TeV measured via semi-electronic decays. Phys. Lett. B, 738:97–108, 2014. 30, 97

[97] K. J. Eskola, H. Paukkunen, and C. A. Salgado. EPS09: A New Generation of NLO and LO Nuclear Parton Distribution Functions. JHEP, 04:065, 2009. 30

[98] D. de Florian and R. Sassot. Nuclear parton distributions at next-to-leading order. Phys. Rev., D69:074028, 2004.

[99] M. Hirai, S. Kumano, and T. H. Nagai. Determination of nuclear parton distribution functions and their uncertainties in next-to-leading order. Phys. Rev., C76:065207, 2007. 30

[100] Prithwish Tribedy and Raju Venugopalan. QCD saturation at the LHC: Comparisons of models to p + p and A + A data and predictions for p + Pb collisions. Phys. Lett. B, 710:125–133, 2012. [Erratum: Phys.Lett.B 718, 1154–1154 (2013)]. 30

[101] Javier L. Albacete, Adrian Dumitru, Hirotsugu Fujii, and Yasushi Nara. CGC predictions for p + Pb collisions at the LHC. Nucl. Phys., A897:1–27, 2013.

[102] Amir H. Rezaeian. CGC predictions for p+A collisions at the LHC and signature of QCD saturation. Phys. Lett., B718:1058–1069, 2013. 30

[103] A. Accardi, F. Arleo, W. K. Brooks, David D'Enterria, and V. Muccifora. Parton Propagation and Fragmentation in QCD Matter. Riv. Nuovo Cim., 32(9-10):439–554, 2009. 30

[104] C. A. Salgado et al. Proton-Nucleus Collisions at the LHC: Scientific Opportunities and Requirements. J. Phys. G, 39:015010, 2012. 30

[105] J. Adam et al. [STAR Collaboration]. Measurement of D^0-meson + hadron two-dimensional angular correlations in Au+Au collisions at $\sqrt{s_{\rm NN}}$ = 200 GeV. Phys. Rev. C, 102(1):014905, 2020. 30

[106] A. Adare et al. [PHENIX Collaboration]. Azimuthal correlations of electrons from heavy-flavor decay with hadrons in $p + p$ and Au+Au collisions at $\sqrt{s_{\rm NN}} = 200$ GeV. Phys. Rev. C, 83:044912, 2011. 30

[107] [ALICE Collaboration]. Letter of intent for ALICE 3: A next-generation heavy-ion experiment at the LHC. 11 2022. 30, 40, 104

[108] Shreyasi Acharya et al. [ALICE Collaboration]. Azimuthal correlations of heavy-flavor hadron decay electrons with charged particles in pp and p–Pb collisions at $\sqrt{s_{\rm NN}}$ = 5.02 TeV. Eur. Phys. J. C, 83(8):741, 2023. 31, 85, 86, 90, 91, 93, 123

[109] Shreyasi Acharya et al. [ALICE Collaboration]. First measurement of Λc+ production down to pT=0 in pp and p-Pb collisions at sNN=5.02 TeV. Phys. Rev. C, 107(6):064901, 2023. 31

[110] Shreyasi Acharya et al. [ALICE Collaboration]. First measurement of $\Omega_{\rm c}^0$ production in pp collisions at $\sqrt{s} = 13$ TeV. 5 2022.

[111] Shreyasi Acharya et al. [ALICE Collaboration]. Measurement of Prompt D^0, Λ_c^+, and $\Sigma_c^{0,++}(2455)$ Production in Proton–Proton Collisions at $\sqrt{s}$ = 13 TeV. Phys. Rev. Lett., 128(1):012001, 2022. 6, 31, 113

[112] Shreyasi Acharya et al. [ALICE Collaboration]. Observation of a multiplicity dependence in the pT-differential charm baryon-to-meson ratios in proton–proton collisions at s=13 TeV. Phys. Lett. B, 829:137065, 2022. 6, 11, 32, 105, 107, 108, 109, 112

[113] R Aaij et al. [LHCb Collaboration]. Prompt charm production in pp collisions at sqrt(s)=7 TeV. Nucl. Phys. B, 871:1–20, 2013.

[114] Shreyasi Acharya et al. [ALICE Collaboration]. $\Lambda_{\rm c}^+$ production and baryon-to-meson ratios in pp and p-Pb collisions at $\sqrt{s_{\rm NN}} = 5.02$ TeV at the LHC. 11 2020. 31, 105, 106

[115] Leonid Gladilin. Fragmentation fractions of c and b quarks into charmed hadrons at LEP. Eur. Phys. J. C, 75(1):19, 2015. 31

[116] H. Abramowicz et al. [ZEUS Collaboration]. Measurement of charm fragmentation fractions in photoproduction at HERA. JHEP, 09:058, 2013.

[117] H. Albrecht et al. [ARGUS Collaboration]. Observation of the Charmed Baryon $\Lambda(c)$ in e^+e^- Annihilation at 10-GeV. Phys. Lett. B, 207:109–114, 1988.

[118] P. Avery et al. [CLEO Collaboration]. Inclusive production of the charmed baryon Lambda(c) from e+ e- annihilations at s**(1/2) = 10.55-GeV. Phys. Rev. D, 43:3599–3610, 1991.

[119] H. Albrecht et al. [ARGUS Collaboration]. Inclusive production of D0, D+ and D*+ (2010) mesons in B decays and nonresonant e+ e- annihilation at 10.6-GeV. Z. Phys. C, 52:353–360, 1991. 31

[120] M. Niiyama et al. [BELLE Collaboration]. Production cross sections of hyperons and charmed baryons from e^+e^- annihilation near $\sqrt{s} = 10.52$˜GeV. Phys. Rev. D, 97(7):072005, 2018. 31

[121] L. Ravagli and R. Rapp. Quark Coalescence based on a Transport Equation. Phys. Lett. B, 655:126–131, 2007. 31

[122] Jesper R. Christiansen and Peter Z. Skands. String Formation Beyond Leading Colour. JHEP, 08:003, 2015. 31, 44, 105, 106, 112

[123] Min He and Ralf Rapp. Charm-Baryon Production in Proton-Proton Collisions. Phys. Lett. B, 795:117–121, 2019. 32, 105

[124] D. Ebert, R. N. Faustov, and V. O. Galkin. Spectroscopy and Regge trajectories of heavy baryons in the relativistic quark-diquark picture. Phys. Rev. D, 84:014025, 2011. 32

[125] Zoltán Varga and Róbert Vértesi. The role of the underlying event in the Λ_c^+ enhancement in high-energy pp collisions. J. Phys. G, 49(7):075005, 2022. 32, 107, 112, 124

[126] Jonathan M. Butterworth, Guenther Dissertori, and Gavin P. Salam. Hard Processes in Proton-Proton Collisions at the Large Hadron Collider. Ann. Rev. Nucl. Part. Sci., 62:387–405, 2012. 32

[127] T. Affolder et al. [CDF Collaboration]. Charged Jet Evolution and the Underlying Event in $p\bar{p}$ Collisions at 1.8 TeV. Phys. Rev. D, 65:092002, 2002. 33

[128] J. A. Hanks et al. Method for separating jets and the underlying event in heavy ion collisions at the BNL Relativistic Heavy Ion Collider. Phys. Rev. C, 86:024908, 2012. 33

[129] Gergely Gabor Barnafoldi, Andras G. Agocs, and Peter Levai. Underlying Event Studies for LHC Energies. AIP Conf. Proc., 1348:124–129, 2011. 33

[130] Tim Martin, Peter Skands, and Sinead Farrington. Probing Collective Effects in Hadronisation with the Extremes of the Underlying Event. Eur. Phys. J. C, 76(5):299, 2016. 34, 105, 107, 112

[131] Shreyasi Acharya et al. [ALICE Collaboration]. Underlying Event properties in pp collisions at $\sqrt{s}$ = 13 TeV. JHEP, 04:192, 2020. 34, 107

[132] Antonio Ortiz, Guy Paić, and Eleazar Cuautle. Mid-rapidity charged hadron transverse spherocity in pp collisions simulated with Pythia. Nucl. Phys. A, 941:78–86, 2015. 35, 107

[133] Aditya Nath Mishra, Gergely Gábor Barnaföldi, and Guy Paić. Quantifying the underlying event: investigating angular dependence of multiplicity classes and transverse-momentum spectra in high-energy pp collisions at LHC energies. J. Phys. G, 50(9):095004, 2023. 35

[134] Antonio Ortiz and Guy Paic. A look into the "hedgehog" events in pp collisions. Rev. Mex. Fis. Suppl., 3(4):040911, 2022. 36, 107

[135] Gyula Bencédi et al. [ALICE Collaboration]. Multi-parton interactions in pp collisions using charged-particle flattenicity with ALICE. PoS, EPS-HEP2023:206, 2024. 36

[136] Antonio Ortiz, Arvind Khuntia, Omar Vázquez-Rueda, Sushanta Tripathy, Gyula Bencédi, Suraj Prasad, and Feng Fan. Unveiling the effects of multiple soft partonic interactions in pp collisions at s=13.6 TeV using a new event classifier. Phys. Rev. D, 107(7):076012, 2023.

[137] Aniko Horvath, Eszter Frajna, and Róbert Vértesi. Event-Shape-Dependent Analysis of Charm–Anticharm Azimuthal Correlations in Simulations. Universe, 9(7):308, 2023. 36

[138] C. K. Chew and Y. K. Lim. Charged Particle Multiplicity Distributions in e^+e^- Annihilation and Negative Binomial Distributions. Phys. Lett. B, 163:257–260, 1985. 37

[139] Alberto Giovannini and L. Van Hove. Negative Binomial Multiplicity Distributions in High-Energy Hadron Collisions. Z. Phys. C, 30:391, 1986.

[140] S. Acharya et al. [ALICE Collaboration]. Charged-particle multiplicity distributions over a wide pseudorapidity range in proton-proton collisions at $\sqrt{s} = 0.9$, 7, and 8 TeV. Eur. Phys. J. C, 77(12):852, 2017. 37

[141] Z. Koba, Holger Bech Nielsen, and P. Olesen. Scaling of multiplicity distributions in high-energy hadron collisions. Nucl. Phys. B, 40:317–334, 1972. 37, 60, 67

[142] A. M. Polyakov. A Similarity hypothesis in the strong interactions. 1. Multiple hadron production in e+ e- annihilation. Zh. Eksp. Teor. Fiz., 59:542–552, 1970. 37, 60, 67

[143] G. J. Alner et al. [UA5 Collaboration]. Scaling Violations in Multiplicity Distributions at 200-GeV and 900-GeV. Phys. Lett. B, 167:476–480, 1986. 37, 67

[144] M. Arneodo et al. Comparison of Multiplicity Distributions to the Negative Binomial Distribution in Muon - Proton Scattering. Z. Phys. C, 35:335, 1987. [Erratum: Z.Phys.C 36, 512 (1987)]. 37, 67

[145] V. A. Abramovsky and N. V. Radchenko. KNO scaling in processes of electron-positron annihilation to hadrons. 6 2007. 37

[146] W. D. Walker. Multiparton interactions and hadron structure. Phys. Rev. D, 69:034007, 2004. 37

[147] C. S. Lam and M. A. Walton. A Proposal for the Origin of KNO Scaling. Phys. Lett. B, 140:246–248, 1984. 37, 67

[148] K. Kudo. Study of the Violation of KNO Scaling and the Validity of Modified KNO Scaling Based on the Cluster Model. Prog. Theor. Phys., 74:1281–1289, 1985.

[149] G. J. H. Burgers, R. Hagedorn, and V. Kuvshinov. Multiplicity distributions in high-energy collisions derived from the statistical bootstrap model. Phys. Lett. B, 195:507, 1987.

[150] S. Hegyi. Renormalization group approach to multiparticle density fluctuations. Phys. Lett. B, 411:321–325, 1997.

[151] Sergei G. Matinyan and W. D. Walker. Multiplicity distribution and mechanisms of the high-energy hadron collisions. Phys. Rev. D, 59:034022, 1999. 37, 67

[152] V. A. Abramovsky and O. V. Kancheli. DISTRIBUTION OF SECONDARY HADRON MULTIPLICITY. (IN RUSSIAN). Pisma Zh. Eksp. Teor. Fiz., 31:566–569, 1980. 37

[153] Antonio Ortiz and Lizardo Valencia Palomo. Probing color reconnection with underlying event observables at the LHC energies. Phys. Rev. D, 99(3):034027, 2019. 37, 56, 59

[154] A. Bassetto. KNO SCALING IN QCD JETS AND THE NEGATIVE BINOMIAL DISTRIBUTION. Nucl. Phys. B, 303:703–712, 1988. 37

[155] Ewa Lopienska. The CERN accelerator complex, layout in 2022. Complexe des accélérateurs du CERN en janvier 2022. 2022. General Photo. 6, 39

[156] K. Aamodt et al. [ALICE Collaboration]. The ALICE experiment at the CERN LHC. JINST, 3:S08002, 2008. 40, 68

[157] Betty Bezverkhny Abelev et al. [ALICE Collaboration]. Performance of the ALICE Experiment at the CERN LHC. Int. J. Mod. Phys. A, 29:1430044, 2014. 40, 68

[158] K Aamodt et al. [ALICE Collaboration]. Alignment of the ALICE Inner Tracking System with cosmic-ray tracks. JINST, 5:P03003, 2010. 40

[159] J. Alme et al. The ALICE TPC, a large 3-dimensional tracking device with fast readout for ultra-high multiplicity events. Nucl. Instrum. Meth. A, 622:316–367, 2010. 41

[160] P. Cortese et al. [ALICE Collaboration]. ALICE electromagnetic calorimeter technical design report. 9 2008. 41

[161] J. Allen et al. ALICE DCal: An Addendum to the EMCal Technical Design Report Di-Jet and Hadron-Jet correlation measurements in ALICE. 6 2010. 41

[162] E. Abbas et al. [ALICE Collaboration]. Performance of the ALICE VZERO system. JINST, 8:P10016, 2013. 41

[163] R. Arnaldi et al. The Zero degree calorimeters for the ALICE experiment. Nucl. Instrum. Meth. A, 581:397–401, 2007. [Erratum: Nucl.Instrum.Meth.A 604, 765 (2009)]. 41

[164] Jaroslav Adam et al. [ALICE Collaboration]. Determination of the event collision time with the ALICE detector at the LHC. Eur. Phys. J. Plus, 132(2):99, 2017. 41

[165] Arturo Tauro. ALICE Schematics. General Photo, 2017. 6, 42

[166] M. H. Seymour. Jet shapes in hadron collisions: Higher orders, resummation and hadronization. Nucl. Phys. B, 513:269–300, 1998. 43, 51

[167] Stephen D. Ellis, Zoltan Kunszt, and Davison E. Soper. Jets at hadron colliders at order $\alpha - s^{3:}$ A Look inside. Phys. Rev. Lett., 69:3615–3618, 1992.

[168] Ivan Vitev, Simon Wicks, and Ben-Wei Zhang. A Theory of jet shapes and cross sections: From hadrons to nuclei. JHEP, 11:093, 2008. 43, 45, 51

[169] Michelangelo L. Mangano and Benjamin Nachman. Observables for possible QGP signatures in central pp collisions. Eur. Phys. J. C, 78(4):343, 2018. 43

[170] Yang-Ting Chien and Ivan Vitev. Towards the understanding of jet shapes and cross sections in heavy ion collisions using soft-collinear effective theory. JHEP, 05:023, 2016. 43

[171] C. Adloff et al. [H1 Collaboration]. Measurement of internal jet structure in dijet production in deep inelastic scattering at HERA. Nucl. Phys. B, 545:3–20, 1999. 43

[172] J. Breitweg et al. [ZEUS Collaboration]. Measurement of jet shapes in photoproduction at HERA. Eur. Phys. J. C, 2:61–75, 1998.

[173] F. Abe et al. [CDF Collaboration]. A Measurement of jet shapes in $p\bar{p}$ collisions at $\sqrt{s} = 1.8$ TeV. Phys. Rev. Lett., 70:713–717, 1993.

[174] S. Abachi et al. [D0 Collaboration]. Transverse energy distributions within jets in $p\bar{p}$ collisions at $\sqrt{s} = 1.8$ TeV. Phys. Lett. B, 357:500–508, 1995.

[175] D. Acosta et al. [CDF Collaboration]. Study of jet shapes in inclusive jet production in $p\bar{p}$ collisions at $\sqrt{s} = 1.96$ TeV. Phys. Rev. D, 71:112002, 2005.

[176] G. Aad et al. [ATLAS Collaboration]. Study of Jet Shapes in Inclusive Jet Production in pp Collisions at $\sqrt{s} = 7$ TeV using the ATLAS Detector. Phys. Rev. D, 83:052003, 2011.

[177] Serguei Chatrchyan et al. [CMS Collaboration]. Jet and Underlying Event Properties as a Function of Charged-Particle Multiplicity in Proton–Proton Collisions at $\sqrt{s}$ = 7 TeV. Eur. Phys. J. C, 73(12):2674, 2013. 43, 45, 48, 58, 61

[178] Daniel Guest, Julian Collado, Pierre Baldi, Shih-Chieh Hsu, Gregor Urban, and Daniel Whiteson. Jet Flavor Classification in High-Energy Physics with Deep Neural Networks. Phys. Rev. D, 94(11):112002, 2016. 43

[179] Rüdiger Haake et al. [ALICE Collaboration]. Machine and deep learning techniques in heavy-ion collisions with ALICE. 9 2017. 43

[180] Christian Bierlich, Gösta Gustafson, Leif Lönnblad, and Andrey Tarasov. Effects of Overlapping Strings in pp Collisions. JHEP, 03:148, 2015. 43

[181] Torbjörn Sjöstrand and Maria van Zijl. A Multiple Interaction Model for the Event Structure in Hadron Collisions. Phys. Rev. D, 36:2019, 1987. 44, 61

[182] Torbjörn Sjöstrand. The Development of MPI Modeling in Pythia. Adv. Ser. Direct. High Energy Phys., 29:191–225, 2018. 44

[183] Spyros Argyropoulos and Torbjörn Sjöstrand. Effects of color reconnection on $t\bar{t}$ final states at the LHC. JHEP, 11:043, 2014. 44

[184] Marcela Carena, S. Pokorski, and C. E. M. Wagner. On the unification of couplings in the minimal supersymmetric Standard Model. Nucl. Phys. B, 406:59–89, 1993. 45

[185] M. Wobisch and T. Wengler. Hadronization corrections to jet cross-sections in deep inelastic scattering. In Workshop on Monte Carlo Generators for HERA Physics (Plenary Starting Meeting), pages 270–279, 4 1998. 45

[186] Albert M Sirunyan et al. [CMS Collaboration]. Measurement of charged pion, kaon, and proton production in proton-proton collisions at $\sqrt{s} = 13$ TeV. Phys. Rev. D, 96(11):112003, 2017. 45, 61

[187] Zoltán Varga, Róbert Vértesi, and Gergely Gábor Barnaföldi. Modification of jet structure in high-multiplicity pp collisions due to multiple-parton interactions and observing a multiplicity-independent characteristic jet size. Adv. High Energy Phys., 2019:6731362, 2019. 50, 61, 122

[188] R. Thomas, Burkhard Kampfer, and G. Soff. Gluon emission of heavy quarks: Dead cone effect. Acta Phys. Hung. A, 22:83–91, 2005. 56

[189] Vardan Khachatryan et al. [CMS Collaboration]. Studies of inclusive four-jet production with two b-tagged jets in proton-proton collisions at 7 TeV. Phys. Rev. D, 94(11):112005, 2016. 59

[190] Róbert Vértesi, Antal Gémes, and Gergely Gábor Barnaföldi. Koba-Nielsen-Olesen-like scaling within a jet in proton-proton collisions at LHC energies. Phys. Rev. D, 103(5):L051503, 2021. 60, 64, 67

[191] Liliana Apolinário, Yen-Jie Lee, and Michael Winn. Heavy quarks and jets as probes of the QGP. 3 2022. 60

[192] Vit Kučera et al. [ALICE Collaboration]. Measurements of groomed heavy-flavour jet substructure with ALICE. PoS, HardProbes2020:149, 2021. 60

[193] Zoltán Varga and Róbert Vértesi. The Partonic Origin of Multiplicity Scaling in Heavy and Light Flavor Jets. Symmetry, 14(7):1379, 2022. 60, 67, 81, 83, 123

[194] Richard Corke and Torbjörn Sjöstrand. Interleaved parton showers and tuning prospects. Journal of High Energy Physics, 2011(3), mar 2011. 61

[195] Torbjörn Sjöstrand. The Lund Monte Carlo for Jet Fragmentation. Comput. Phys. Commun., 27:243, 1982. 61

[196] G. C. Strong. Gluon splitting to b-quark pairs in proton-proton collisions at $\sqrt{s} = 8$ TeV with ATLAS. MSc Thesis, 2015. 61

[197] Serguei Chatrchyan et al. [CMS Collaboration]. Identification of b-Quark Jets with the CMS Experiment. JINST, 8:P04013, 2013. 62

[198] S. Hegyi. KNO scaling 30 years later. Nucl. Phys. B Proc. Suppl., 92:122–129, 2001. 67

[199] Volker Blobel. An Unfolding method for high-energy physics experiments. In Conference on Advanced Statistical Techniques in Particle Physics, pages 258–267, 8 2002. 74

[200] René Brun, F. Bruyant, Federico Carminati, Simone Giani, M. Maire, A. McPherson, G. Patrick, and L. Urban. GEANT Detector Description and Simulation Tool. 10 1994. 74

[201] Tim Adye. Unfolding algorithms and tests using RooUnfold. In PHYSTAT 2011, pages 313–318, Geneva, 2011. CERN. 74

[202] G. D'Agostini. A Multidimensional unfolding method based on Bayes' theorem. Nucl. Instrum. Meth. A, 362:487–498, 1995. 74

[203] Torbjorn Sjostrand, Stephen Mrenna, and Peter Z. Skands. PYTHIA 6.4 Physics and Manual. JHEP, 05:026, 2006. 85, 89

[204] Torbjorn Sjostrand, Stephen Mrenna, and Peter Z. Skands. A Brief Introduction to PYTHIA 8.1. Comput. Phys. Commun., 178:852–867, 2008. 85, 92

[205] Borut Paul Kersevan and Elzbieta Richter-Was. The Monte Carlo event generator AcerMC versions 2.0 to 3.8 with interfaces to PYTHIA 6.4, HERWIG 6.5 and ARIADNE 4.1. Comput. Phys. Commun., 184:919–985, 2013. 85

[206] K. Werner, Iu. Karpenko, T. Pierog, M. Bleicher, and K. Mikhailov. Event-by-Event Simulation of the Three-Dimensional Hydrodynamic Evolution from Flux Tube Initial Conditions in Ultrarelativistic Heavy Ion Collisions. Phys. Rev. C, 82:044904, 2010. 85, 92

[207] H. J. Drescher, M. Hladik, S. Ostapchenko, T. Pierog, and K. Werner. Parton based Gribov-Regge theory. Phys. Rept., 350:93–289, 2001. 85, 92

[208] S. Acharya et al. [ALICE Collaboration]. ALICE 2017 luminosity determination for pp collisions at $\sqrt{s}$ = 5 TeV. Technical Report ALICE-PUBLIC-2018-014, Nov 2018. 86

[209] Betty Bezverkhny Abelev et al. [ALICE Collaboration]. Measurement of visible cross sections in proton-lead collisions at $\sqrt{s_{\rm NN}}$ = 5.02 TeV in van der Meer scans with the ALICE detector. JINST, 9(11):P11003, 2014. 86

[210] Betty Bezverkhny Abelev et al. [ALICE Collaboration]. Long-range angular correlations of π, K and p in p-Pb collisions at $\sqrt{s_{\rm NN}}$ = 5.02 TeV. Phys. Lett. B, 726:164–177, 2013. 87

[211] Shreyasi Acharya et al. [ALICE Collaboration]. Measurement of electrons from heavy-flavour hadron decays as a function of multiplicity in p-Pb collisions at $\sqrt{s_{\rm NN}}$ = 5.02 TeV. JHEP, 02:077, 2020. 88

[212] Jaroslav Adam et al. [ALICE Collaboration]. Measurement of electrons from heavy-flavour hadron decays in p-Pb collisions at $\sqrt{s_{\rm NN}} = 5.02$ TeV. Phys. Lett. B, 754:81–93, 2016. 88, 89

[213] Jaroslav Adam et al. [ALICE Collaboration]. Measurement of the production of high-$p_{\rm T}$ electrons from heavy-flavour hadron decays in Pb-Pb collisions at $\sqrt{s_{\rm NN}}$ = 2.76 TeV. Phys. Lett. B, 771:467–481, 2017. 88, 89

[214] Shreyasi Acharya et al. [ALICE Collaboration]. Measurements of low-p$_T$ electrons from semileptonic heavy-flavour hadron decays at mid-rapidity in pp and Pb-Pb collisions at $\sqrt{s_{\rm NN}} = 2.76$ TeV. JHEP, 10:061, 2018. 89

[215] S. Acharya et al. [ALICE Collaboration]. The ALICE definition of primary particles. Technical Report ALICE-PUBLIC-2017-005, Jun 2017. 89

[216] Shreyasi Acharya et al. [ALICE Collaboration]. Azimuthal Anisotropy of Heavy-Flavor Decay Electrons in p-Pb Collisions at $\sqrt{s_{\rm NN}}$ = 5.02 TeV. Phys. Rev. Lett., 122(7):072301, 2019. 91

[217] Matteo Cacciari, Mario Greco, and Paolo Nason. The $p_{\rm T}$ spectrum in heavy flavor hadroproduction. JHEP, 05:007, 1998. 92

[218] Matteo Cacciari, Michelangelo L. Mangano, and Paolo Nason. Gluon PDF constraints from the ratio of forward heavy-quark production at the LHC at $\sqrt{S} = 7$ and 13 TeV. Eur. Phys. J. C, 75(12):610, 2015. 92

[219] Shreyasi Acharya et al. [ALICE Collaboration]. $\Lambda_{\rm c}^+$ production in pp collisions at $\sqrt{s} = 7$ TeV and in p-Pb collisions at $\sqrt{s_{\rm NN}} = 5.02$ TeV. JHEP, 04:108, 2018. 105

[220] Albert M Sirunyan et al. [CMS Collaboration]. Production of $\Lambda_{\rm c}^+$ baryons in proton-proton and lead-lead collisions at $\sqrt{s_{\rm NN}} = 5.02$ TeV. Phys. Lett. B, 803:135328, 2020. 105

[221] Jun Song, Hai-hong Li, and Feng-lan Shao. New feature of low p_T charm quark hadronization in pp collisions at $\sqrt{s} = 7$ TeV. Eur. Phys. J. C, 78(4):344, 2018. 105

[222] Salvatore Plumari, Vincenzo Minissale, Santosh K. Das, G. Coci, and V. Greco. Charmed Hadrons from Coalescence plus Fragmentation in relativistic nucleus-nucleus collisions at RHIC and LHC. Eur. Phys. J. C, 78(4):348, 2018. 105

[223] Shreyasi Acharya et al. [ALICE Collaboration]. Measurement of the cross sections of $\Xi_{\rm c}^0$ and $\Xi_{\rm c}^+$ baryons and branching-fraction ratio BR($\Xi_{\rm c}^0 \rightarrow \Xi^-{\rm e}^+\nu_{\rm e}$)/BR($\Xi_{\rm c}^0 \rightarrow \Xi^-\pi^+$) in pp collisions at 13 TeV. 5 2021. 105

[224] Christopher Hills et al. [ALICE Collaboration]. Charmed-baryon production and hadronization studies with ALICE. PoS, HardProbes2020:079, 2021. 106

[225] P. A. Zyla et al. Review of Particle Physics. PTEP, 2020(8):083C01, 2020. 106

[226] Zoltán Varga, Róbert Vértesi, and Gergely Gábor Barnaföldi. Jet Structure Studies in Small Systems. Universe, 5(5):132, 2019. 122

[227] Zoltán Varga, Róbert Vértesi, and Gergely Gábor Barnaföldi. Multiplicity Dependence of the Jet Structures in pp Collisions at LHC Energies. MDPI Proc., 10(1):3, 2019. 122

[228] Z. Varga. Jet momentum dependence of jet multiplicity in pp collisions at 13 TeV. ALICE Analysis Note, 2024. 123

[229] Zoltán Varga, Anett Misák, and Róbert Vértesi. The role of the underlying event in the charm-baryon enhancement observed in pp collisions at LHC energies. SciPost Phys. Proc., 15:014, 2024. 124

[230] Zoltán Varga, Anett Misák, and Róbert Vértesi. Event-activity-dependent production of strange and non-strange charmed baryons in the enhanced color-reconnection scheme. J. Phys. G, 50(7):075002, 2023. 124

[231] Róbert Vértesi and Zoltán Varga. Connection of event shapes to the heavy-flavor baryon enhancement. Proceedings of the 52nd International Symposium on Multiparticle Dynamics, arXiv:2402.01234, 2024. 124