INTERNATIONAL DOCTORAL
SCHOOL OF THE USC

Imanol
Corredoira Fernández

PhD Thesis

Prompt charged particle production measurements in small collision systems at the LHCb experiment

Santiago de Compostela, 2024

Doctoral Programme in Nuclear and Particles Physics

ESCOLA DE DOUTORAMENTO
INTERNACIONAL DA USC

DOCTORAL THESIS

PROMPT CHARGED PARTICLE PRODUCTION MEASUREMENTS IN SMALL COLLISION SYSTEMS AT THE LHCB EXPERIMENT

Imanol Corredoira Fernández

DOCTORAL PROGRAMME IN NUCLEAR AND PARTICLES PHYSIC

Supervisor/s: Abraham Antonio Gallas Torreira

Ricardo Vázquez Gómez

Tutor: Abraham Antonio Gallas Torreira

SANTIAGO DE COMPOSTELA

2024

Acknowledgements

Aínda que esta tese leva o meu nome, é un traballo colectivo. Intentarei agradecer nestas páxinas a xente que aportou e fixo posible este traballo aínda que é unha tarefa imposible de completar.

Quero comezar agradecendo ao IGFAE como institución, e a toda a xente que aportou o seu traballo a ciencia en Galicia e España porque grazas a estas persoas tiven a oportunidade de facer esta tese na miña terra. Agradecer aos meus directores Abraham e Ricardo que me desen a oportunidade e que me guiasen nos meus comezos. Ademais de agradecer ao resto dos profesores do grupo de LHCb do IGFAE por terme ensinado cousas, tanto personais como científicas, aquí mencionar a Cibrán, Juan, Toño e Marcos. Tamén debo moito ao grupo de teoría de QCD; Nestor, Carlos e especialmente Elena polas horas de charla, explicación e discusión científica que espero que nunca rematen.

Gustaríame mencionar aos estudantes e postdocs cos que compartín esta viaxe, Clara, Lidia, Iván, Victoria e Arnau e especialmente a Sara, Sam e Óscar. Xunto a eles/elas aprendín, ensinei e paseino moi ben este anos. Desexolle aos estudantes que se encontran no longo camiño da tese que non perdan a motivación e recorden sempre porque están facendo esto. Agradecer especialemente a Óscar que me levase da man dende o comezo, non se me ocurre mellor mentor ca él.

Durante este anos viaxei por moitos lugares e coñecín a moita xente. Aprendín que a comunidade científica é aberta, inclusiva e que sempre hai xente disposta a ensinarche cousas de maneira desinteresada. Gustaríame mencionar aquí a xente dos lugares onde pasei mais tempo e onde me ensinaron moito. Primero París, onde Frederic, Emilie, Benjamin e Kara fixeron que me sentise como na casa mentras aprendía moito. Ademais non me quero olvidar de Michael Winn e Jean-Yves Ollitrault. Agardo con ilusión seguir traballando con vós. Despois a xente de Los Alamos: Matt, César, Julie, Hugo, Ming e todo o grupo. Logo de estar alí a miña idea de EE.UU cambiou. Teño que agradecerlles que me fixesen sentir que un se pode divertir facendo ciencia ao mais alto nivel e que todo se pode conseguir. Tamén agradecer a Michal e Susan que me acollesen tan ben na súa casa, síntome moi agradecido.

No eido mais persoal non podo deixar de mencionar a xente de Sarria, dende as amigas de sempre Andrea, Ánxela, Lucía, Ana e Ainara a os amigos de sempre como Flores e Nuria. Ademais de facer especial mención a o Bar Xarope a Asociación Rua da Música por ensinarme a importancia da comunidade e axudarme a canalizar as miñas inquedanzas artísticas. Teño moita sorte de ter crecido rodeado destas persoas. Volver a Sarria sempre me fai recordar o importante que é a comunidade.

Quero continuar agradecendo a xente de Santiago e aredeodores. Primerio aos meus compañeiros de piso: Juan, Ledicia, Mario e Fran pola maravillosa convivencia durante anos. Como non, quero mencionar a Jorge de Santa Comba, a persona máis honorable que coñezo, sempre un amigo e compañeiro exemplar, oxalá che vaia moi ben. Tamén agradecer a Mateo, Andrea Gregores e Alvaro Crego que sigan formando parte das miñas

mellores amizades dende a carreira. Ademais, este último ano tiven a sorte de coñecer a un grupo de mulleres incríbles coas que vivín moitos bos momentos e que vou botar de menos. Refírome a Carmen, Tea, Salomé e Sara. Grazas polas conversacións, as risas e todo o que enseñastedes e me fixestedes vivir. Fóstedes un apoio moi importante este ano.

Para rematar, os agradecementos mais importantes van para a miña familia e para Carolina. Todos son un exemplo de dedicación, paciencia (a que eu non teño) e cariño hacia min que nunca serei capaz de devolver. Só espero que sintades este traballo un pouco voso, porque eu non estaría aquí sin a vosa axuda e exemplo.

Todas estas personas aportaron moito a miña formación durante este anos, non so de forma científica senón tamén persoal, que considero igual de importante.

Abstract

This thesis includes two measurements of prompt charged particle properties from proton-proton and proton-lead collisions at $\sqrt{s_{\rm NN}} = 5.02\,\text{TeV}$ recorded at the LHCb experiment at CERN. One is the multiplicity distribution and the other is the multiplicity dependence of the average transverse momentum. Charged particles are measured in $0.5 < p_{\rm T} < 8.0\,\text{GeV}/c$ and $2.0 < \eta_{LAB} < 4.8$. The pseudorapidity dependence of both measurements in the forward region has been studied for the first time in this thesis.

The thesis is structured as follows. In Chap. 1 is an **introduction** to the topic where the current knowledge and open problems of the field are summarised. Also, the new information provided by this work to the field will be explained. In the Chap. 2, a more profound theoretical description is made. Starting from the standard model and going through strong interaction theory until arriving at the particle production models, saturation and quark-gluon-plasma. Then, the **methodology** of the analysis is detailed. Starting from the description of the dataset used in this analysis (Chap. 4) and explaining the preparation and selection (Chap. 5). The core of the analysis is detailed in Chap. 6 where all the analysis techniques are explained. The systematic uncertainty estimation is done in Chap. 7 and results are shown in Chap. 8. Some **prospects** for future measurements in central ion collisions at LHC-Run5 are done in Chap. 9. Finally, some **discussion** and **conclusions** are provided in Chap. 10. A **summary in Galician** is presented in Chap. A, and the **bibliography** is included at the end of the document.

Key words: Experimental High Energy Physics, LHCb, Heavy Ions Physics, QCD, Inclusive Prompt Charged Particles, Proton-proton Collision, Proton-Lead Collision, Saturation, QGP, Multi Parton Interactions, Run5, EoS, speed of sound

Resumo

Esta tese inclúe dúas medidas de propiedades das partículas cargadas primarias producidas en colisións protón-protón e protón-núcleo. Os datos corresponden a colisións rexistradas no experimento LHCb do CERN a unha enerxía de $\sqrt{s_{\rm NN}} = 5.02\,\text{TeV}$ no centro de masas. A primeira medida corresponde a distribución de probabilidade de partículas cargadas primarias e a segunda a o momento transverso promedio das mesmas en función da multiplicidade da colisión. En ambas medidas as partículas cargadas seleccionadas están no intervalo de momento transverso $0.5 < p_{\rm T} < 8.0\,\text{GeV}/c$ e de pseudorapidez $2.2 < \eta_{LAB} < 4.8$. Ambos observables son estudados en función da pseudorapidez neste traballo.

A tese esta estruturada do seguinte xeito. En Chap. 1 faise unha **introdución** onde se resume o coñecemento actual e algúns problemas abertos arredor do tema da tese. Ademais, indicase o novo coñecemento que esta tese trae ao campo. No Chap. 2 faise unha introdución teórica mais detallada. Dende o modelo estándar da física de partículas pasando pola teoría da interacción forte ata chegar aos modelos de produción de partículas en colisións hadrónicas, a saturación e o quark-gluon-plasma. Logo disto desenrolase a **metodoloxía** nos seguintes capítulos. Comezando pola explicación dos datos utilizados (Chap. 4) e explicando como foron filtrados e preparados para o análise (Chap. 5). Una vez explicados os datos que se usan comeza o desenrolo dos métodos de análise no Chap. 6. Aquí explícanse todos os pasos do **análise de datos**. As **incertezas sistemáticas** do procedemento experimental explícanse nun capítulo dedicado Chap. 7 e os **resultados** móstranse no capítulo Chap. 8. Ademais, as **predicións** para análises similares que se poderán levar a cabo en colisións centrais de núcleos no Upgrade 2 do experimento LHCb mostranse no capítulo Chap. 9. Finalmente, no capítulo Chap. 10 expóñense as **conclusións** da tese xunto con unha **discusión** dos resultados.

El mayor enemigo del descubrimiento no es la ignorancia, es la ilusión del conocimiento

Stephen Hawking

El sabio sincero y de vocación permanece profundamente humano

Santiago Ramón y Cajal

¡Desgraciado del que en presencia de un libro se queda mudo y absorto!

Santiago Ramón y Cajal

INDEX

1

Introduction

The strong nuclear force is a crucial fundamental interaction responsible for the existence of stable nuclear matter. The quantum field theory that models this force, called Quantum Chromodynamics (QCD) [1], [2], [3] gives a framework to compute predictions for cross-sections of processes involving the strong force. However, these predictions are limited by the properties of QCD to the processes involving high transferred momentum in comparison with the Λ_{QCD} parameter (*hard* processes). Those processes involving small transferred momenta (*soft* processes) require some hypothesis or approximations to be computed producing in some cases very different paradigms.

The study of particles produced directly in high-energy hadronic collisions on an event-by-events basis provides information about the particle production process and the inner structure of the colliding objects. As particle production in this regime occurs via QCD without requiring any particular energy scale, particles are produced via hard and (mainly) soft interactions within the same event. This allows the study of soft and hard QCD and its interplay. Moreover, the hadron inner structure and its modification when forming bound nuclei is unknown in the high energy region and no model successfully explains them. These are two reasons, the study of the interplay between soft and hard QCD processes and the study of hadron inner structure and its nuclear modifications, that motivate this thesis. As particle detectors are better when measuring charged particles, this work focuses on prompt charged particles.

The charged particle multiplicity is a common observable when studying particle production in hadronic collisions and it imposes important constraints on the mechanisms of particle production [4]. Usually, the multiplicity distribution, i.e., the probability distribution of obtaining a definite number of produced particles is a quantity that models, usually

available through Monte Carlo event generators, must be able to reproduce. Specifically, the multiplicity distribution is related to the collective or individual character of particle production. The multiplicity distribution follows a Poisson distribution if the final-state particles are produced independently. In this case, the dispersion $D = \sqrt{\langle n^2 \rangle - \langle n \rangle^2}$ is related to the average multiplicity as $D = \sqrt{\langle n \rangle}$. Deviations from a Poisson distribution indicate correlations in particle production as explained in [4]. The origin of this correlation may vary depending on the model. Moreover, measuring the multiplicity distribution of a collision system at a given energy characterises that collision system, which is a statistical system. In other words, allows us to know how rare is a hadronic collision by knowing the produced multiplicity. This quantity has been extensively measured in the central rapidity region by many experiments, for example from hadronic collisions at LHC [5] and from e^+e^- annihilation at LEP [6].

Apart from the multiplicity distribution, there are other quantities related to prompt charged particles that can be measured in hadronic colliders. The multiplicity dependence of the average transverse momentum also carries information about the particle production mechanism and the hadronic inner structure. This quantity has been measured at hadron colliders in $pp(\bar{p})$ collisions in a wide range of energies from $\sqrt{s_{NN}} = 31\,\text{GeV}$ until $\sqrt{s_{NN}} = 13\,\text{TeV}$ and also nuclear collisions as Xe-Xe and Pb-Pb at $\sqrt{s_{NN}} = 5\,\text{TeV}$. These experiments focused on the central rapidity region, had observed an increase in the average transverse momentum as a function of multiplicity. Different theories belonging to different paradigms, including a modification of the hadron structure at high energies or a correlated partonic interaction, tried to explain this data. In this work, the first measurement of this observable in the forward region is presented. Moreover, the first measurement studying the pseudorapidity dependence of this quantity is also shown.

This thesis addresses the mentioned physical open questions by measuring two quantities related to prompt charged particles: the multiplicity distribution and the multiplicity dependence of the average transverse momentum. This measurement is done for proton-proton, proton-lead and lead-proton collisions at the LHC using data collected by the LHCb experiment at $\sqrt{s_{NN}} = 5\,\text{TeV}$. This document is organised as follows. In the Chap. 2, a general introduction of the necessary theoretical background to understand proton-proton and proton-lead collisions is presented. The overview starts from the standard model and explains the topics more related to understanding the thesis. At Chap. 3, the LHCb detector is described in detail. This chapter is especially focused on the subsystems used to record the data used in this work. Then, in Chap. 4 the dataset used is described, and its preparation is detailed in Chap. 5. Analysis techniques and measurement procedures are described in Chap. 6. Systematic uncertainty treatment is detailed in Chap. 7. Finally, results and discussion are enclosed in Chap. 8 and Chap. 10 respectively. Apart from the measurement, some prospective work is done for LHCb's Upgrade 2 considering central PbPb collisions, this is detailed in Chap. 9.

1.1 Objectives and methodology

The thesis aims to provide new measurements for prompt charge particle production in small systems (pp, pPb and Pbp) in the forward region at LHC energies. Moreover, new measurement techniques for multiplicity-dependent analysis are developed and predictions for the LHCb Upgrade II are performed. The first measurement is the multiplicity distribution for pp, pPb and Pbp collisions in the forward region. The second corresponds to the average transverse momentum as a function of multiplicity. Both observables are measured in pseudorapidity bins. Moreover, the measurement technique implemented in this thesis to perform multiplicity-dependent analysis at LHCb is compared with other techniques already implemented in other published analyses. Finally, prospects for PbPb collisions show that LHCb has access to a unique physics program where the equation of the state of quantum chromodynamics can be measured in a wider range of temperatures than in other LHC experiments.

The methodology used in this study makes use of LHCb recorded data at the LHC collider. This data is analysed starting from the reconstructed tracks. The first step is to optimise a selection in order to enhance our signal-to-background ratio. Then detector effects have to be estimated and corrected from the candidate sample. Some of these effects are reconstruction and selection efficiency or signal purity. The correction factor for each effect is estimated from the simulation and corrected with a data-driven correction factor when possible. After that, the two measurements are performed. First, the multiplicity spectra require Bayesian Unfolding in order to correct for the finite size of the bins and bin migration. In order to apply this statistical method a response matrix has to be filled with simulation. Moreover, a closure test has to be fulfilled to make sure that the Unfolding works properly. Regarding the average transverse momentum measurement a proxy is computed in order to estimate event multiplicity from data. The performance and bias of the proxy in the measurement are estimated using a closure test. Finally, after applying all the detector corrections to the measured transverse momentum spectra the average transverse momentum measurement can be produced.

2

Theoretical foundations

This chapter presents an overview of the most relevant theoretical aspects necessary to understand the thesis content. The chapter is organised as follows. Firstly, some basic definitions concerning kinematics at colliders are presented in Sec. 2.1. Secondly, the standard model of particle physics is summarised in Sec. 2.2. Thirdly, the quantum field theory of strong force is discussed in Sec. 2.3. Fourthly, an overview of the physics of high energy hadronic collisions, such as those happening at the LHC, is presented in Sec. 2.4. Finally, General Purpose Monte Carlo Generators are explained describing the differences between each one at Sec. 2.5.

2.1 Coordinate systems and kinematic variables in collider physics

Particles produced at LHC are described using Lorentz-invariant quantities as they have very high energy in comparison with their mass. The description of these particles is based on two four vectors: position $x^\mu = (ct, x, y, z)$ and momentum $p^\mu = (E/c, p_x, p_y, p_z)$ where the transverse momentum magnitude is $p_T = \sqrt{p_x^2 + p_y^2}$. In LHC experiments, the coordinate system in the laboratory reference frame is defined such as the z-axis is aligned with the beam, the x-axis points towards the centre of the LHC and the y-axis points perpendicularly up to the sky as shown in Fig. 2.1. Apart from the Cartesian coordinates, two angles are defined: the azimuthal angle, ϕ, and the polar angle, θ. However, despite its Lorentz invariant properties, the pseudorapidity η is used instead of θ, this variable is defined as follows,

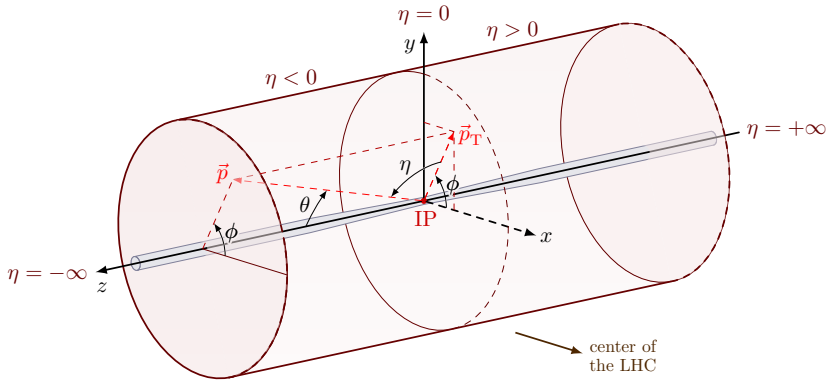


Figure 2.1: Coordinate system of an LHC experiment, where the interaction point is on the centre. The LHCb experiment is located in the $\eta > 0$ region. Draw by the author inspired in the work of I. Neutelings.

$$\eta = -\ln(\tan(\theta/2)). \tag{2.1}$$

Using pseudorapidity the component of the momentum can be defined as follows

$$p_x = \|\vec{p_T}\cos\phi\|\,, p_y = \|\vec{p_T}\sin\phi\|\,, p_x = \|\vec{p_T}\sinh\eta\|\,. \tag{2.2}$$

These variables are used to describe the position of particles with respect to the interaction point (IP) located at x=y=z=0. For high-momentum particles where $E \simeq pc$, pseudorapidity is approximately equal to the rapidity of the particle with respect to the beam defined as

$$y = \frac{1}{2}\ln\frac{E + p_x c}{E - p_z c}. \tag{2.3}$$

Rapidity is a useful quantity because it transforms additively under Lorentz boost, unlike velocity.

2.2 The standard model of particle physics

The standard model of particle physics (SM) is the most fundamental physics theory ever written. It explains the interaction between the fundamental pieces of matter through fundamental forces. This theory is based on the Quantum Field Theory (QFT) mathematical framework and experimental observations. As a consequence, it is able to explain and predict many experimental observables with huge precision.

The SM considers that the most fundamental matter structures are *quarks* and *leptons*. Quarks come in six types called *flavours*: up (u), down (d), charm (c), strange (s),

top (t), and bottom (b) while leptons also consist of six types: electron (e), muon (μ), tau (τ) and its corresponding neutrinos ($\nu_{e,\mu,\tau}$). Quarks and leptons are fermions with spin = 1/2. Moreover, all these particles have antimatter partners with opposite quantum numbers and the same mass. A summary of the fundamental pieces of matter and its properties can be found in Fig. 2.2.

These fundamental pieces of matter interact between themselves through three kinds of forces:

- Strong force: Mediated by the gluon, g.
- Weak force: Mediated by $W^{\pm}$ and Z^0 bosons.
- Electromagnetic force: Mediated by the photon γ.

Each one of these forces is associated with a local gauge symmetry of the SM Lagrangian. In this way, the algebraic structure of the theory explains the properties of the fundamental interactions. In the case of the SM, the algebraic structure is based on the symmetry of the Lagrangian under three Lie groups, SU(3) $\otimes$ SU(2) $\otimes$ U(1). Moreover, the SM includes also the scalar Higgs boson (H). This particle explains how the masses of the mediators ($W^{\pm}, Z^0$) are related to the fundamental parameters of the theory. This is achieved via the Higgs mechanism, which involves the spontaneous breaking of the electroweak symmetry.

Nevertheless, the SM has several limitations, including the charge-parity (CP) violation problem related to the observed matter/antimatter asymmetry in the Universe, the fact that it does not include gravity and the explanation for the fermion masses for example.

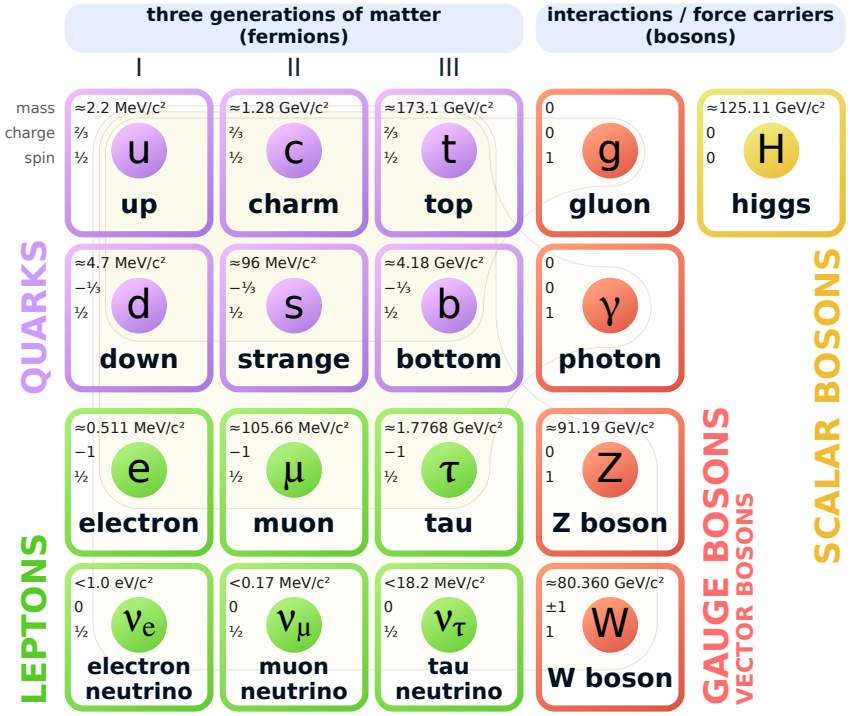


Figure 2.2: Scheme of the standard model of particle physics including all the particles and their properties. Figure taken from [7] under the Creative Commons Attribution 3.0 Unported license.

2.3 The theory of Quantum Chromodynamics

The part of the SM that explains the strong force is known as Quantum Chromodynamics (QCD) [1]. This theory is modelled by the following part of the SM Lagrangian

$$\mathcal{L}_{\mathcal{QCD}} = -\frac{1}{4}\sum_{A=1}^{8} F^{A\mu\nu}F^{A}_{\mu\nu} + \sum_{j=1}^{n_f} \bar{q}_j(i\not{D} - m_j)q_j. \tag{2.4}$$

In this equation, q_j are the quark fields (of n_f different flavours) with mass m_j. Moreover, $\not{D} = D_\mu\gamma^\mu$ are written in the Feynman *slash* notation are corresponds to

$$D_\mu = \partial_\mu - ie_s\sum_A t^A g^A_\mu, \tag{2.5}$$

where γ^μ are the Dirac γ-matrices and D_μ is the covariant derivative. Other ingredients in Eq. 2.4 are the gauge coupling e_s and the gluon fields g^A_μ, with A running from 1 to the number of colours squared (3) minus one; i.e. $N_c^2 - 1 = 8$. Thus, there are eight kinds of gluons. Also, t^A represent 3×3 matrices that correspond to the number of generators of the SU(3) colour group; i.e.: these generators fulfil the algebra $[t^A, t^B] = iC_{ABC}t^C$,

where C_{ABC} are the complete asymmetric structure constant of SU(3). In the first term of Eq. 2.4, $F^A_{\mu\nu}$ is the field tensor, which describes the dynamics of the gluon field, and is given by

$$F^A_{\mu\nu} = \partial_\mu g^A_\nu - \partial_\nu g^A_\mu - e_s C_{ABC} g^B_\mu g^C_\nu . \tag{2.6}$$

Each term in the QCD Lagrangian is related to a physical process, using Feynman diagrams it can show all the processes involving coloured particles (quarks and gluon) under strong force at *tree level* at Fig.2.3.

The physical vertices in QCD include gluon-quark-antiquark, 3-gluon and 4-gluon vertices. The fact that QCD mediators can interact with each other is an important difference between QCD and other sectors of SM. Moreover, the QCD theory shows three salient properties:

- Color Confinement: In QCD, the potential between two quarks with opposite charges is $V_{q\bar{q}}(r) \simeq \frac{-4}{3}\frac{\alpha_s \hbar c}{r} + \kappa r$. This causes the energy to increase until a quark-antiquark pair is spontaneously created, resulting in the formation of a pair of hadrons rather than isolating a colour charge. Although analytically unproven, colour confinement is well-supported by lattice QCD calculations and extensive experimental evidence.

- Asymptotic Freedom: As the energy scale of interactions between quarks and gluons increases (and the corresponding length scale decreases), the strength of these inter-

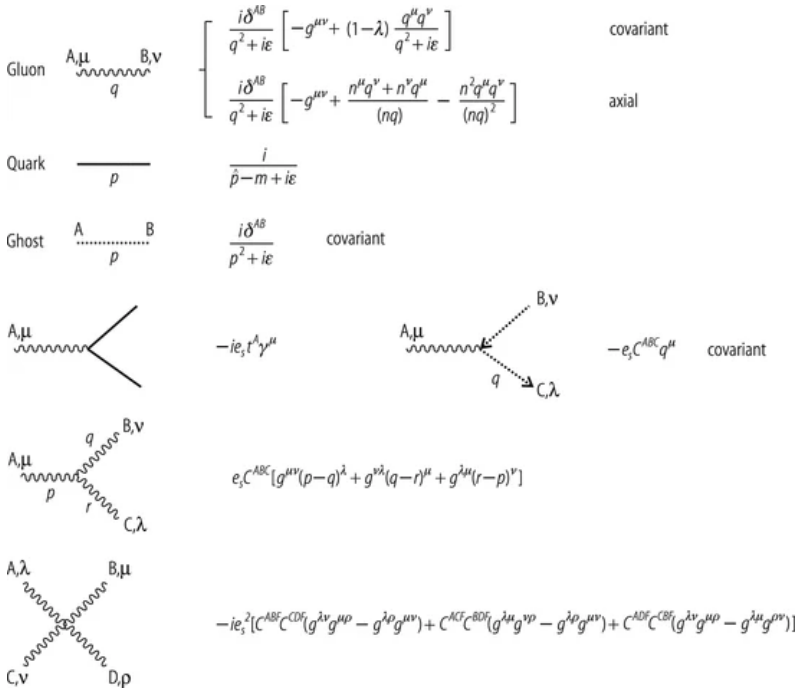


Figure 2.3: Feynman rules for QCD. Quarks are represented by solid lines, curly lines represent gluons, and dotted lines the ghosts. The gauge parameter is denoted by λ. Figure taken from Ref. [8].

actions decreases. This phenomenon, known as asymptotic freedom, was discovered in 1973 by David Gross and Frank Wilczek, and independently by David Politzer. Their work earned them the 2004 Nobel Prize in Physics.

- Chiral Symmetry Breaking: This property involves the spontaneous breaking of chiral symmetry, leading to the generation of hadron masses significantly greater than those of the quarks, and making pseudoscalar mesons exceptionally light. Yoichiro Nambu's work in 1960 elucidated this phenomenon, and he was awarded the 2008 Nobel Prize in Physics for his contributions, which were later confirmed by lattice simulations.

One of the fundamental parameters for a theory of interactions is the coupling constant. This parameter, in QCD represented by e_s (or $\alpha_s = e_s^2/(4\pi)$), provides the strength of the interaction. The behaviour of the QCD coupling constant gives rise to some of the already mentioned main characteristics of this theory; i.e. the asymptotic freedom and confinement. Although denominated *constant*, α_s is not a constant parameter for each physical process because it depends on the transferred four-momentum, q^2, between the interacting particles.

In Fig. 2.4 it can be seen the dependence of α_s with Q as measured in different experiments in the range $1 < Q < 10^3\,\mathrm{GeV}/c$.

It is important to remember that in order to be able to compute observables using perturbation theory the coupling constant has to be $<< 1$. What can be learnt from the experimental measurements shown in Fig. 2.4 is that for large Q, α_s decreases and as a consequence the strong interaction becomes weaker. In the $Q \to \infty$ limit quarks and gluons experiment have no strong interaction, they are free particles. This is a common property of all gauge theories based on a non-commuting group of symmetry and is known as asymptotic freedom. The coupling constant in these theories is called *running coupling constant*. Since the coupling decreases asymptotically, QCD-related quantities can be computed using perturbation theory in this regime. Calculations from perturbative QCD (pQCD) are relevant at leading orders starting typically from $Q \simeq 1 - 2$ GeV/c. The running of the coupling constant can be computed and follows the equation

$$\alpha_{\mathrm{s}}(k^2) \overset{\mathrm{def}}{=} \frac{g_{\mathrm{s}}^2(k^2)}{4\pi} \approx \frac{1}{\beta_0 \ln\left(k^2/\Lambda^2\right)}. \tag{2.7}$$

Where the scale parameter Λ is introduced and defined as

$$\Lambda \equiv \mu^2 e^{-\frac{4\pi}{\beta_0 \alpha_s(\mu)}}. \tag{2.8}$$

The opposite behaviour is observed for low Q, where α_s grows fast as Q decreases. For $Q \lesssim \Lambda$ the coupling constant becomes large and perturbation theory cannot be applied. In this regime, non-perturbative methods (Lattice-QCD or effective field theories) need to be used to investigate the theory. The divergence of α_{s} at low-Q values is related to the fact that quarks are bound to hadrons – quark confinement.

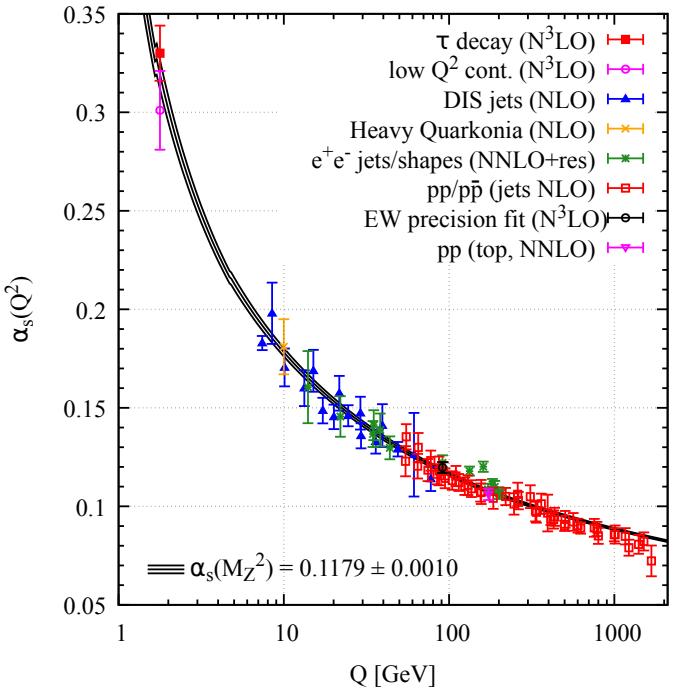


Figure 2.4: Dependence of α_s with the energy scale Q from different measurements [9]. In parenthesis, the precision degree of perturbation theory used to extract α_s is indicated (NLO: next-to-leading order; NNLO: next-to-next-to-leading order; NNLO + res.: NNLO matched to a resummed calculation; N^{3}LO: next-to-NNLO). Figure extracted from Ref. [10] under a Creative Commons 4.0 International (CC BY 4.0) license.

2.3.1 Perturbative QCD

As explained in the previous section, to use perturbation theory (expand the cross-section computation in powers of the constant α_s successfully), the requirement of $\alpha_s \ll 1$ has to be fulfilled. In QCD, this approach is valid for hard processes i.e. when $q^2 \gg \Lambda^2_{\text{QCD}}$. The measurements presented in this thesis involve perturbative and non-perturbative processes, so perturbative calculations will be explained. However, before explaining how perturbative processes are computed in pQCD, it is necessary to explain radiation processes.

For a given parton i (quark or gluon), there is a probability of radiating another parton j represented by the splitting probability functions, P_{ij}. The lifetime of an emission with energy ω and transverse momentum k_T can be approximated as

$$\tau \simeq \frac{\omega}{k_T}. \tag{2.9}$$

From Eq. 2.9 it can be seen that soft and collinear emissions, lead to infrared and collinear divergences ($\propto \frac{1}{w}, \propto \frac{1}{k_T^2}$) and thus, need to be factorised away from the cross

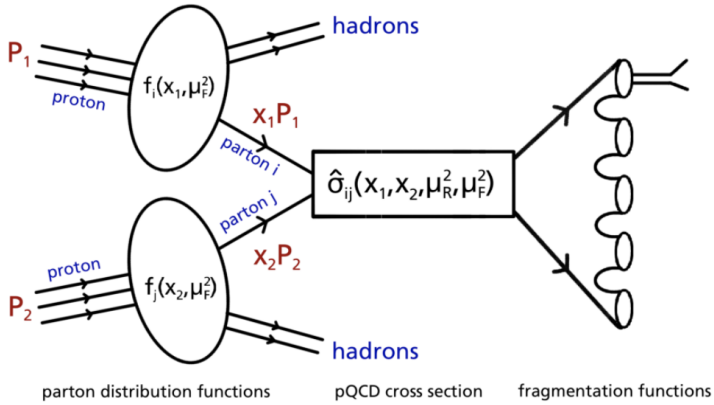


Figure 2.5: Illustration of the factorisation theorem applied to a proton-proton collision. Sketch drawn by T. Poulsen. The momentum of the incident protons is denoted as $P_{1,2}$.

sections and then described using re-summation techniques as is explained in detail in [11]. In QCD the probability of observing an emission with fractional momentum x depends on the scale of the process (or the resolution parameter) Q^2. Depending on where this splitting is produced, QCD radiation is treated in a different way:

- Parton Distribution Functions (PDFs): When the radiation emission is produced by partons forming a hadronic state. The PDFs, named $f_i^A(x, \mu^2)$ represent the probability of finding a parton with a fractional momentum x inside the hadron A at a given resolution scale μ^2.

- Fragmentation Functions (FFs): Noted as D_i^h, represent the probability of a parton i ending up being a colourless hadron h after a number of radiative emissions.

This leads us to the factorisation theorem which is a fundamental part of the cross-section computation for hadronic process. This theorem allows us to factorize the cross section into the PDFs for the protons (f_i and f_j) and the hard scattering partonic cross section σ_{ij} as can be seen in Fig. 2.5. The indices i and j indicate the two interacting partons and the complete mathematical formulation is:

$$\mathrm{d}\sigma^{h_1+h_2\to k+X}(\mu^2, Q^2) = \sum_{i,j,X'} f_{i/h_1}(x_1, \mu_F) \otimes f_{i/h_2}(x_2, \mu_F) \otimes \mathrm{d}\hat{\sigma}^{ij\to k+X'}(\mu_R, \mu_F). \quad (2.10)$$

Here, $f_{i/h_1}(x_1, \mu_F)$ and $f_{i/h_2}(x_2, \mu_F)$ are the parton distribution functions (PDFs), which describe the number density distribution of partons i (j) in a hadron h_1 (h_2) at a momentum fraction x_1 (x_2) and a factorisation scale μ_F. The factorisation scale, μ_F, determines the scale below which the emissions are absorbed into the PDFs. Moreover, μ_R denotes the factorization and the renormalization scales. The momentum fraction x

of a parton within a proton is defined as the fraction of the proton momentum carried by the parton. The quantity $\mathrm{d}\hat{\sigma}^{ij\to k+X'}(\mu_R, \mu_F)$ is the production cross-section from the process where partons i and j from h_1 and h_2 interact and yield the elementary particle k and any additional products X', considering the interaction diagrams up to the targeted order of the prediction. Note that the equation is summed over i, j and X', and therefore all combinations of partons and possible products have to be considered. The dependence with μ is related to the normalisation scale. The differential cross-section of Eq. 2.10 needs to be integrated over the values of x_1 and x_2 that contribute to $\mathrm{d}\hat{\sigma}^{ij\to k+X'}(\mu_R, \mu_F)$ for a given final state with particular kinematics.

However, it should be noticed that Eq. 2.10 computes the cross section for producing a parton k. As it was explained before this parton will involve radiating other partons until it forms a colourless hadron. This process is called hadronization and is ruled by FFs. In order to compute the cross-section of hadron production the following equation has to be used,

$$\mathrm{d}\sigma^{h_1+h_2\to h+X}(\mu_R, \mu_F) = \sum_k \mathrm{d}\sigma^{h_1+h_2\to k+X}(\mu_R, \mu_F) \otimes D_{h/k}(z, \mu_F). \tag{2.11}$$

Here, z describes the momentum fraction carried away by the hadron h from the parent parton k; $D_{h/k}(z, \mu_F)$ is the parton-to-hadron fragmentation function; and μ_F is the fragmentation scale. As with Eq. 2.10, the integration over the contributing x_1 and x_2 must be performed to obtain inclusive cross-sections for particles with specific kinematics. FFs are generally obtained through global analyses of experimental data. A recent review on FFs can be found in Ref. [12]. There are many examples of analyses of fragmentation functions for different hadron species [13–15].

As FFs, PDFs cannot be computed from the QCD Lagrangian so they are determined from global data analyses for deep inelastic lepton-nucleon scattering (DIS) at hadron-electron colliders such as HERA. Although they cannot be calculated from first principles, its Q^2 evolution is perturbatively driven by the Dokshitzer-Gribov-Lipatov-Altarelli-Parisi (DGLAP) equations [16–18]. A recent review of the current status of PDFs can be found in Ref. [19]. As example, Fig. 2.6 shows the parton distribution functions from the CT18 analysis [20] at $Q = 2\,\mathrm{GeV}$ and $Q = 100\,\mathrm{GeV}$. usually, PDFs are displayed as a function of the fractional momentum with respect to the hadron, x, (also called Björken-x).

According to collider kinematics, $x \propto \frac{1}{\sqrt{s}e^y}$, therefore, the partonic composition of ultra-relativistic hadrons is dominated by gluons as can be seen in Fig. 2.6. However, following unitarity principles (Froissart bound), gluons are expected to start recombining and the gluonic content saturates as $x \to 0$. This is actively researched and usually called *saturation*, however, it has not been explicitly measured yet.

Additionally, when nucleons are bound into nuclei their inner structure is modified and PDFs are modified into nPDFs (nuclear Parton Distribution Functions). The extraction of nPDFs is an active field of research, its precise understanding is crucial to understanding both the structure of matter and the QGP effects that will be explained later. A summary of nPDFs can be found in [21] and, as can be seen in Fig.2.7, the nPDF extraction from

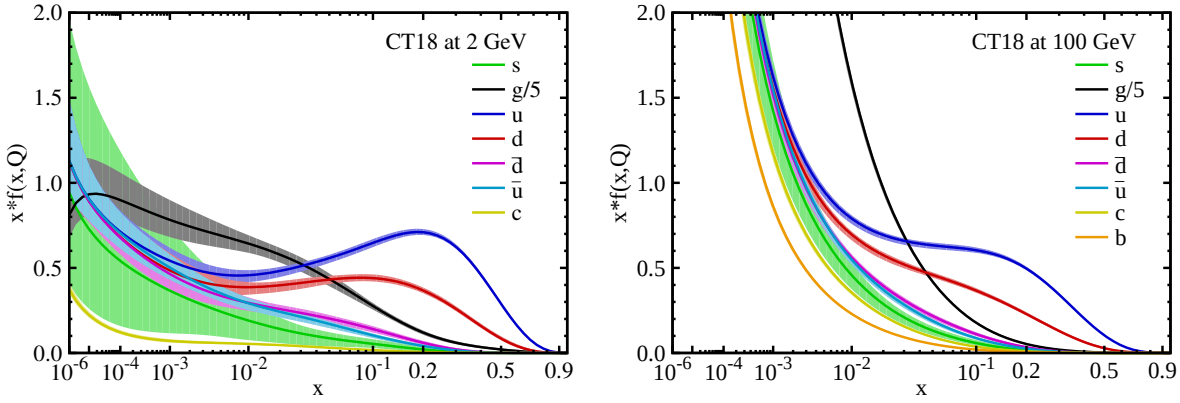


Figure 2.6: PDFs from the CT18 analysis [20] at $Q = 2\,\text{GeV}$ (left) and $Q = 100\,\text{GeV}$ (right). The PDFs consider u, $\overline{u}$, d, $\overline{d}$, $s = \overline{s}$, $b = \overline{b}$ and g. The gluon PDF is scaled as $g(x, Q)/5$ and the charm distribution $c(x, Q)$ is perturbatively generated by evolving from $Q_0 = 1.3$ and $1.4\,\text{GeV}$. The estimated uncertainty is estimated with a band. Figure under Creative Commons Attribution 4.0 International license

data is model-dependent.

2.3.2 The QCD phase diagram

As was already explained, the strength of the QCD interaction depends on the energy scale of the process. Observing the coupling constant evolution with Q in Fig. 2.4 it can be intuited that at high energies the interactions will be so weak that partons are free, this is asymptotic freedom. This state of matter formed by partons in this region of the phase space is called *Quark Gluon Plasma* (QGP). The existence of this new phase of matter in QCD theory was confirmed using the lattice QCD, (see Refs. [23, 24] and references therein).

In QCD, asymptotic freedom and Lagrangian symmetries determine the phases of strongly interacting matter. These phases are shown in Fig. 2.8 as a function of temperature T and baryon chemical potential μ which represents the net number of baryons with respect to anti-baryons. The strongly interacting matter has a rich phase structure, this includes a nuclear liquid phase, a hadronic gas, and the quark-gluon plasma.

When temperature and chemical potential are zero, the interaction between quarks is dominated by large distances and the coupling is large. Thereby quarks and gluon are confined in color singlet hadrons with masses in the order of Λ_{QCD}. Chiral symmetry breaking implies the emergence of Goldstone bosons with the quantum numbers of the generators of the broken axial symmetry $SU(3)_A$. These particles are π, K, and η.

When the temperature is high enough, quark and gluon have momentum $p \sim T >> \Lambda_{\text{QCD}}$. In this regime asymptotic freedom starts to play a role and particles interact weakly forming a plasma of colour charges, this is the quark-gluon plasma. It is important to

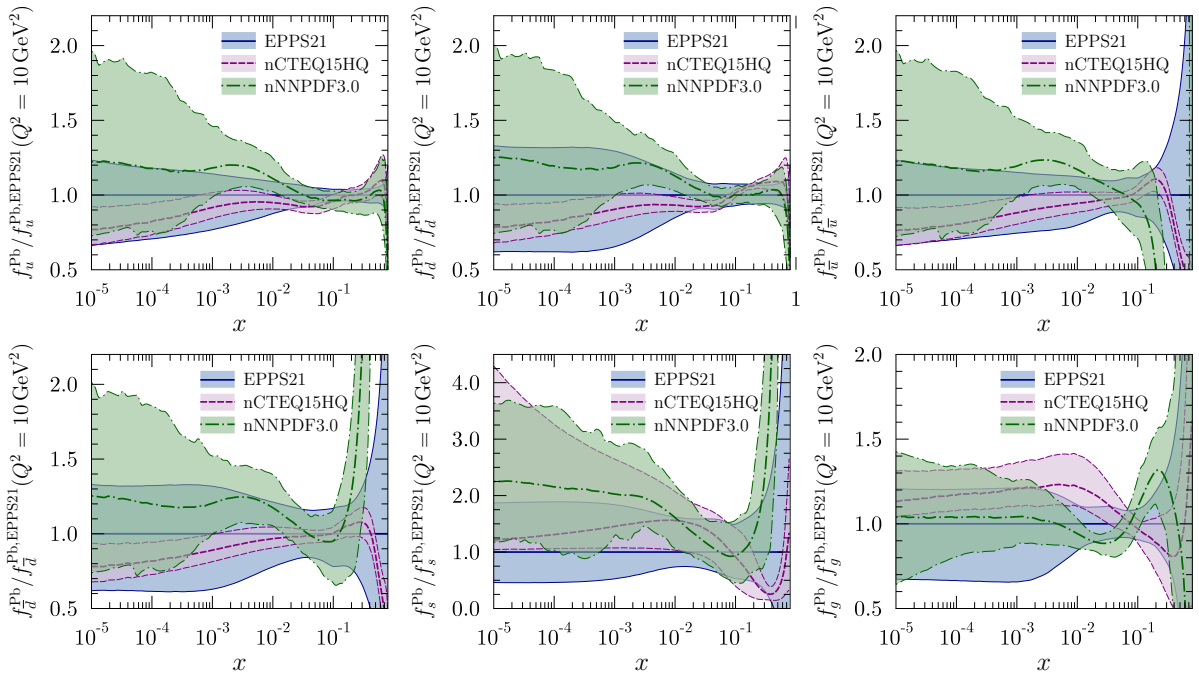


Figure 2.7: Comparison of the ^{208}Pb nuclear modifications resulting from the EPPS21 (full, blue), nCTEQ15HQ (dashed, red) and nNNPDF3.0 (dot-dashed, green) global analyses of nuclear PDFs, i.e. the PDFs of lead divided by the summed PDFs of 82 free protons and 126 free neutrons. Uncertainty bands correspond to 90% CL. Figure extracted from [22] under a Creative Commons Attribution 4.0 International License.

explain that at high temperatures there is a thermal population of mobile colour charges that goes as $n \sim T^3$. These charges screen the interaction at large distances making this phase dominated by short distance interactions. In this plasma phase, matter does not exhibit colour confinement or chiral symmetry breaking. This is usually understood as a phase transition from the low-temperature phase.

From Lattice computations as [25], it is shown that for quark masses, $m_u, m_d = 10$ MeV and $m_s = 120$ MeV, the phase transition is not a first-order phase transition (as it was thought at the beginning) but a gradual phase transition where the degrees of freedom of the system goes from being hadronic to partonic as the temperature rises. This means the transition occurs over a range of temperatures without the sharp behaviour that could be expected from a first-order phase transition. This is a lattice QCD computation that is in agreement with LHC and RHIC observations. The critical temperature is estimated as $T_c \sim 151$ MeV.

When chemical potential increases quarks and gluons are again weakly coupled as in QGP. The difference between cold quark matter and hot QGP is the large density of

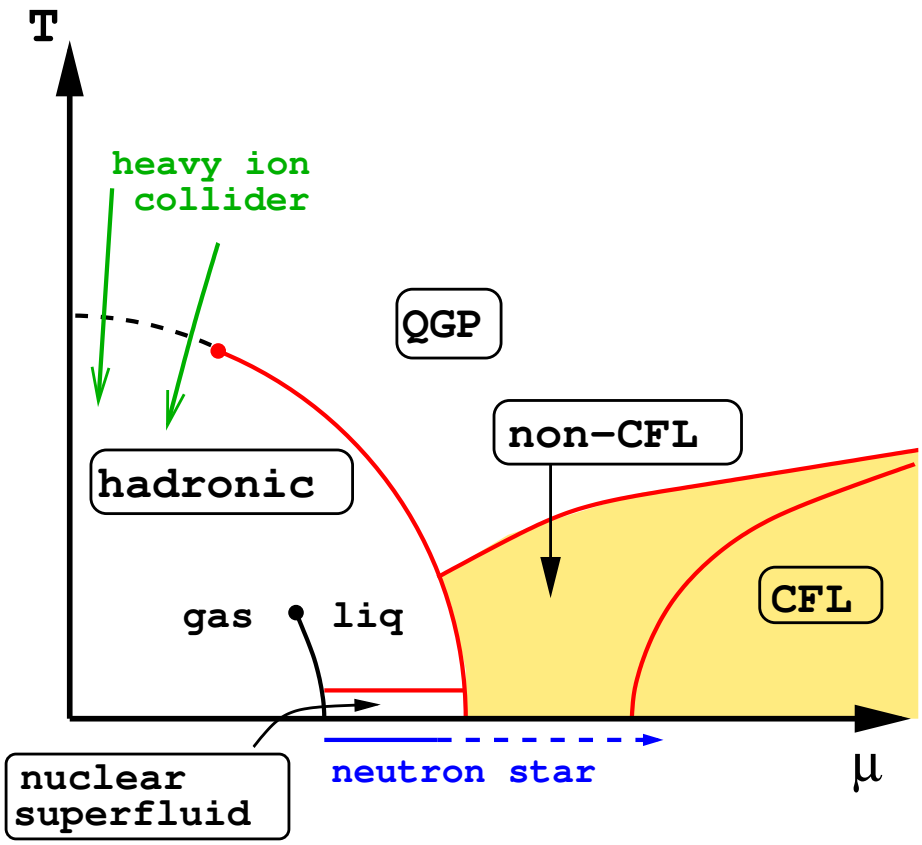


Figure 2.8: Schematic phase diagram of QCD as a function of temperature T and baryon chemical potential μ. QGP refers to the quark-gluon plasma. The CFL (colour-flavour locked) phase is the colour superconducting phase that occurs at an asymptotically large chemical potential. The red and black points denote the critical endpoints of the chiral and nuclear liquid-gas phase transitions, respectively. The dashed line is the chiral pseudo-critical line associated with the crossover transition at low temperature. The green arrows denote the regions of the phase diagram that are being explored by the experimental heavy-ion programs at the LHC and RHIC. This figure was taken from Refs. [23].

states near the quark Fermi surface. Thereby, even weak interactions can cause qualitative changes in the ground state of dense matter. In particular, attractive interactions between quark pairs lead to colour superconductivity and the formation of a $\langle qq \rangle$ condensate. A superconducting phase called the color-flavor locked (CFL) phase is predicted to exist at very high density.

When describing matter, a fundamental equation is the equation of state (EOS). This equation describes the dependence of pressure of the system temperature and chemical

potential $P = P(T, \mu)$. A fundamental quantity that determines the expansion of hot dense matter, and is derived from the EoS, is the speed of sound,

$$c_s^2 = \frac{\partial P}{\partial \varepsilon}\Big|_{s/n_B}, \tag{2.12}$$

where ε is the system energy and the derivative is taken at constant entropy per baryon, s/n_B. The EoS describes how gradients in the energy density profile are translated into pressure gradients. When the number of baryons is zero; i.e. $n_B = 0$, c_s is just a function of temperature. Neglecting the chemical potential is a good approximation in high-energy ion collisions as the nucleons of ultrarelativistic Pb nuclei escape the interaction volume before the plasma develops. Moreover, the high-energy collisions subsequently lead to a sizable baryon production balanced by anti-baryons due to conservation laws.

2.4 High energy hadronic collisions

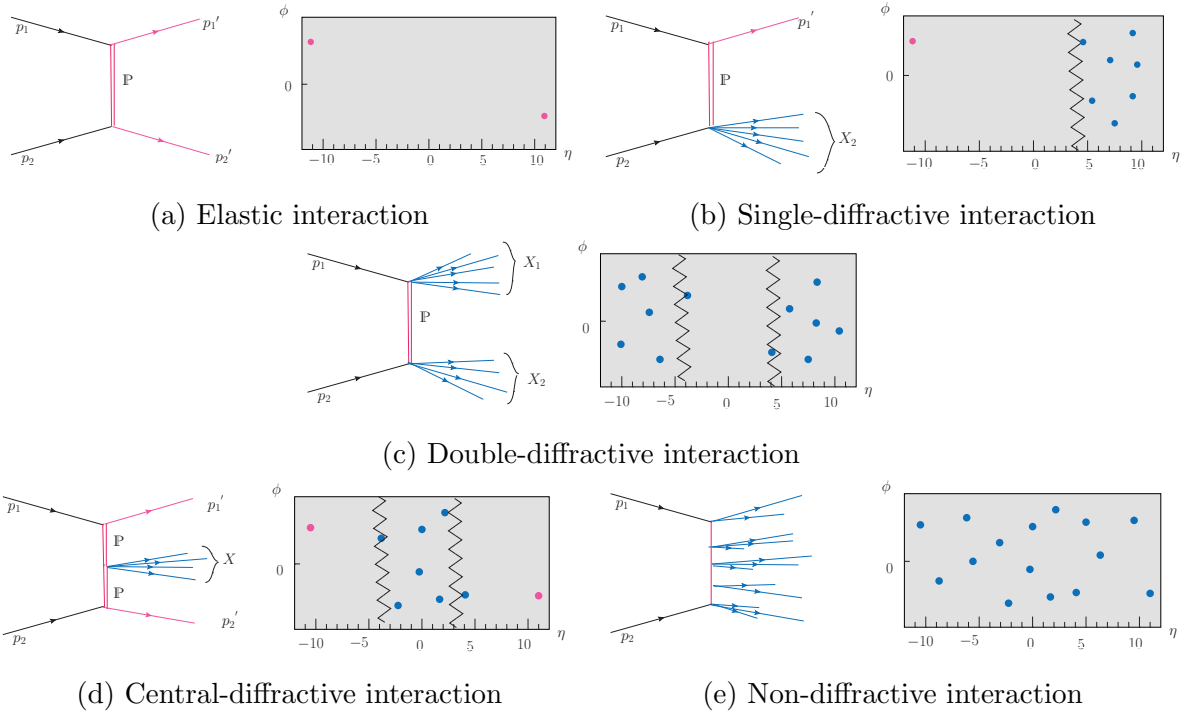
Hadronic matter and QCD properties can be studied in several ways. This work will be focused on high-energy hadronic collisions in colliders. In this kind of study, the chemical potential is approximately zero, $n_B = 0$. The reason is that the spectator nuclei escape from the interaction region when the bulk matter evolves. Moreover, in the interacting region, the number of baryon and anti-baryon production is in equilibrium. This means that the region of the QCD phase diagram from Fig. 2.8 studied in high-energy hadronic collisions corresponds to the vertical axis; i.e. the evolution with temperature.

2.4.1 High energy proton-proton collisions

When two protons, p_1 and p_2, collide, different kinds of processes can occur depending on the final state of the collision as shown in Fig. 2.9.

If the final state of the collision does not involve more particles than the ones in the initial state; i.e. two protons, then the collision is called *elastic* and the process is $p_1 + p_2 \to p_1 + p_2$. Elastic interaction takes place thanks to a Pomeron exchange, which is a colourless QCD state designed with $\mathbb{P}$. The elastic proton-proton cross-section at $\sqrt{s} = 7\,\text{TeV}$ is 25.4 ± 1.1mb, about a fourth of the total cross-section of proton-proton collisions 98.6 ± 2.2mb [27].

When additional particles are produced in the collision the interaction is called *inelastic*. Inelastic processes are dominant with a cross-section of 72.9 ± 1.5mb [28] at $\sqrt{s} = 7\,\text{TeV}$. Inelastic processes are usually classified as diffractive and non-diffractive (ND) depending on whether the interaction is carried on by a Pomeron interaction (diffractive) or not (non-diffractive). Within the diffractive category, several kinds of interactions can be considered as single diffractive (SD), double diffractive (DD) and central diffractive (CD) as can be seen in Fig. 2.9 depending on the topology of the final state particles.



(a) Elastic interaction

(b) Single-diffractive interaction

(c) Double-diffractive interaction

(d) Central-diffractive interaction

(e) Non-diffractive interaction

Figure 2.9: Possible types of interactions in proton-proton collisions. The diagrams on the right indicate the distribution of particles in the final state in the (η, ϕ) plane to represent the presence of large rapidity gaps. Figures taken from Ref. [26].

In a diffractive interaction [29] at least one of the initial protons is dissociated into more final state particles with smaller mass. Typically they can be experimentally recognised by the presence of Large Rapidity Gaps (LRG) in the angular distribution of final state particles. These are regions in the η without any particle. This kind of process is also produced by a Pomeron exchange.

When both protons interact at partonic level they break into multiple final state particles that cover all the angular region. This is denominated as a non-diffractive process. These are presented in Fig. 2.9e. The total proton-proton interaction cross-section can be calculated by adding the interaction cross-sections for all the possible processes as follows,

$$\sigma_{\text{total}} = \sigma_{\text{elastic}} + \sigma_{\text{inel}} = \sigma_{\text{elastic}} + \sigma_{\text{diff}} + \sigma_{\text{ND}} = \sigma_{\text{elastic}} + \sigma_{\text{SD}} + \sigma_{\text{DD}} + \sigma_{\text{CD}} + \sigma_{\text{ND}}. \quad (2.13)$$

In this thesis, no specific study was done in order to separate different proton-proton interactions. Thus, generic collisions will be the object of study in this document.

As it was explained before, two kinds of processes can be defined in QCD in terms of the coupling constant value. Perturbative processes, when $Q^2 >> 1$ GeV/c^2, in collider physics are called "hard" processes, and the non-perturbative ones are called "soft"

processes and $Q^2 << 1$ GeV/c^2. In a non-diffractive proton-proton collision, multiple processes happen simultaneously as shown in Fig. 2.10. The observables studied in this thesis involved both soft and hard processes, so they are sensitive to the interplay between perturbative and non-perturbative QCD.

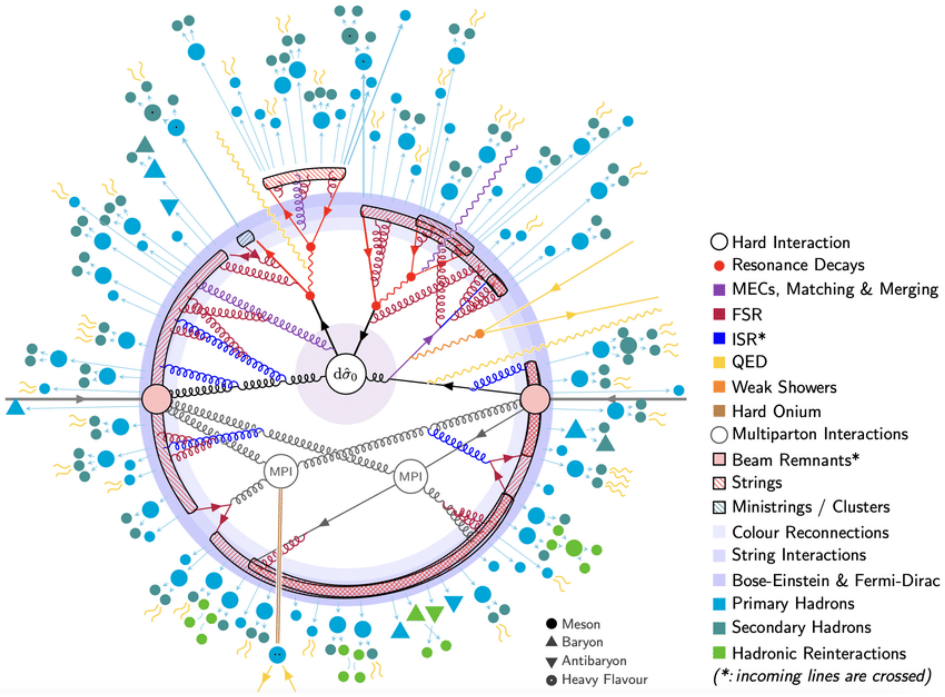


Figure 2.10: Sketch from the PYTHIA8 [30] representation of a proton-proton collision event when all different processes are shown. Figure under a Creative Commons Attribution 4.0 International License.

Two of the interactions shown in Fig. 2.10 are especially important for this thesis:

- **Hard scattering:** Represented by the cross-section $d\sigma_0$. This is the most energetic process and the one that produces the biggest number of final state particles. This process is initialized by a parton with high momentum and can be described by pQCD using the factorization theorem. The final state particles form a shower of hadrons in a definite (η, ϕ) region and are called a jet.

- **MPI:** Multi Parton Interactions. Secondary partonic interactions with smaller energy than the hard scattering could be a semi-hard or soft process. MPIs also produce many final state particles forming, together with the rest of the processes, what is called the Underlying Event (UE). There is a lot of research on the understanding of MPI. One of the main questions regarding MPI physics is whether they are independent of each other or correlated. The average transverse momentum as a function of multiplicity is a key observable for MPI physics.

2.4.2 High energy heavy ion collisions

The ions are Lorentz-contracted objects in a collision involving ultra-relativistic energies. Instead of having a spherical shape, they form a disc of thickness about $14/\gamma$ fm with $\gamma \simeq 2500$ at LHC energies and a transverse size $R_A \simeq A^{1/3}$. In this scenario, it can be assumed that all nucleons initially (before the collisions) propagate along parallel, straight-line trajectories. Moreover, depending on the role of those nucleons in the collisions they can be classified. Interacting nucleons are called participants (N_{part}), the ones not interacting are called spectators and the total number of incoherent nucleon-nucleon collisions is defined as N_{coll} as is shown in Fig. 2.11.

When describing heavy-ion collisions, the characterization of the initial stage of the collisions is crucial. In this sense, the impact parameter (b) is defined as the distance between the centres of the two colliding nuclei. The impact parameter characterizes the centrality of the collision. A central collision is one with a small impact parameter in which the two nuclei collide almost head-on, $b \to 0$; a peripheral collision, on the contrary, is one with a large impact parameter, $0 < b < 2R_A$, and an ultra-peripheral collision happens when the nuclei don't pass through each other, $b > 2R_A$.

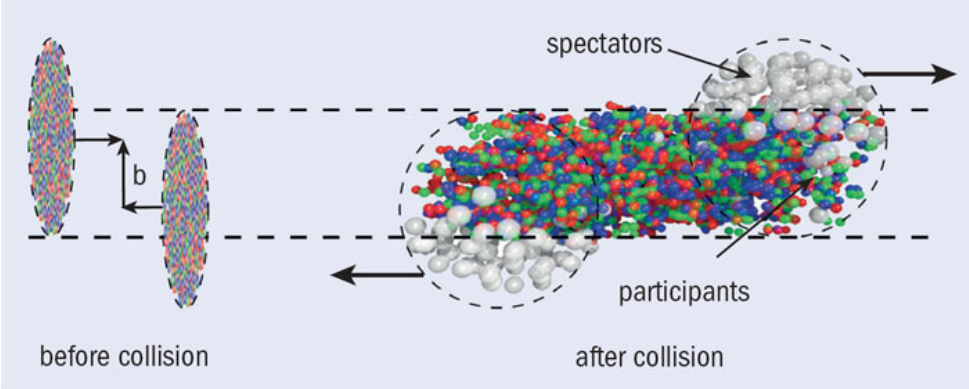


Figure 2.11: Ultra-relativistic heavy ion collision scheme where participant, spectators and impact parameter is shown. Figure taken from [31]

In an ultra-relativistic heavy ion collision with high enough centrality, from a theoretical perspective, the system evolves through several stages which are represented in Fig. 2.12. In this evolution there are three main stages:

1. Early stages: The collision and the brief period of time afterwards. Here the description of the initial stage of the colliding objects is crucial. Possible saturation of the gluon PDFs in the nuclei could be explained by classical field theory in the high occupancy limit; i.e. Color Glass Condensate (CGC). The nuclei wave function controls the evolution of the system in the very early times.

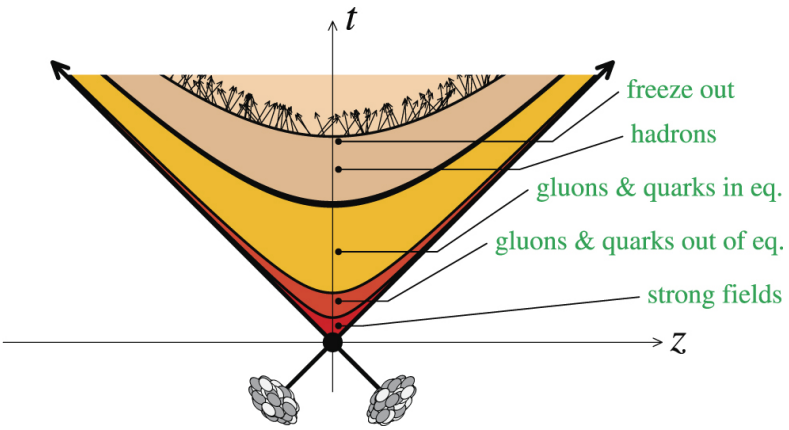


Figure 2.12: Scheme representing the stages of a peripheral and central Ultra-Relativistic Heavy Ion collision, figure extracted from [32] under Creative Commons Attribution (CC BY) licence.

2. QGP: A longer stage where the matter is close to local thermal equilibrium. Here, the temperature is high enough to reach QGP, hence deconfinement. The system is described using relativistic viscous hydrodynamics and partons are the relevant degrees of freedom. Some of the experimental signatures of this stage are collective flow, n_q scaling, quarkonia suppression and R_{AA}.

3. Hadronization: The final stage during which the matter has become too dilute to remain in equilibrium. While the temperature drops partons join forming hadrons that travel towards the detectors.

In general, it is difficult to know from which stage comes the measured behaviour of a given quantity. For this reason, studying the same observable in different collision systems is very important. For example, it is typically assumed that minimum-bias pp and pA collisions do not form QGP. In this sense, these collision systems are used as a baseline to interpret heavy ion collisions where the QGP emergence is assumed. In particular, pA collisions are used to understand what is called *Cold Nuclear Matter* effects (CNM). Those are nuclear effects that do not correspond to high-temperature behaviours (QGP) like for example nPDFs.

However, since the LHC era started, some collective behaviours have been observed in pp and pA (collective flow or strangeness enhancement). This implies that the explained picture may have to change. In particular, it raises the question of whether the initial stage or CNM effects could exhibit collective behaviour that mimics QGP or is it possible that QGP is formed in some pp, pA events. This is one of the main questions in our field that is not solved at the moment. Characterising collective behaviour in small systems (pp, pA) as a function of multiplicity is a useful tool that helps us understand this problem.

2.5 Monte-Carlo generators

In this section, General Purpose Monte Carlo (GPMC) generators will be explained [10]. GPMCs are a fundamental software tool for simulating hadronic collisions. They are widely used together with software detector descriptions like Geant4 [33] to understand how the detector affects measurement and also to study backgrounds and signals for many studies. The main goal of GPMCs is to simulate all physics processes in a hadronic collision, some of them are PYTHIA 8.3 [30], SHERPA 2 [34], HERWIG 7.2 [35] or EPOS-LHC [36].

Since QCD interactions are weak at short distances (below a femtometer), the components of the GPMC that handle short-distance physics are based on perturbation theory. However, at larger distances, where soft hadronic phenomena occur—such as hadronization and the formation of the underlying event in hadron collisions—first-principles computations are currently not feasible. Instead, QCD-inspired models must be used and different GPMCs assume different hypotheses for these processes.

As GPMC are very complex, they have many parameters that can be adjusted using data. This is commonly called *tunning* and can be done within experimental collaboration or directly by the MC authors.

In this thesis, two GPMC are used PYTHIA 8 and EPOS-LHC. Thus, this generator will be briefly explained emphasising the differences in soft physics and MPI, which are the most important parts regarding the observables that will be presented in this work.

2.5.1 PYTHIA 8.3

A pp collision simulated in PYTHIA can be seen in the illustration in Fig. 2.10 and it is structured as follows:

1. The parton kinematics of the beam parton are determined based on the PDFs. The nature of the event is decided (e.g. Z^0 production) and the resonance decay is produced.

2. All the surrounding partonic activity is simulated. This includes the initial and final state parton radiation (ISR and FSR), MPIs, and interaction from the beam particle remnants. The MPIs are based on [37]. However, recent developments have included a more advanced modelling of flavour and colour correlations. This field is quickly evolving. A major unresolved issue is how the colour of all these subsystems is correlated. While there is certainly a correlation due to the colour-singlet nature of the incoming beams, final-state colour rearrangements can alter this picture. These additional effects are essential for accurately describing data, such as the $\langle p_T \rangle (N_{ch})$ measurements which is the objective of this thesis. A straightforward implementation of colour rearrangement can be found in the description of beam remnants in PYTHIA.

3. Hadronisation from partons is produced following the Lund string model. This part is completely non-perturbative and fully phenomenological. Unstable particles decay and a full final-state particle collection is obtained.

Although PYTHIA 8 was typically used to model *pp* collisions, recently an extension to heavy ion collisions was made under the name PYTHIA Angantyr [38].

2.5.2 EPOS-LHC

The MC generator EPOS-LHC [36] manages hadronic collisions in a very different way. Here the approach is not partonic, in the sense that the interaction between hadrons is done coherently. The Monte Carlo generator uses the Parton-Based Gribov Regge Theory where the interaction between hadrons is done by Pomeron interaction as can be seen in Fig. 2.13.

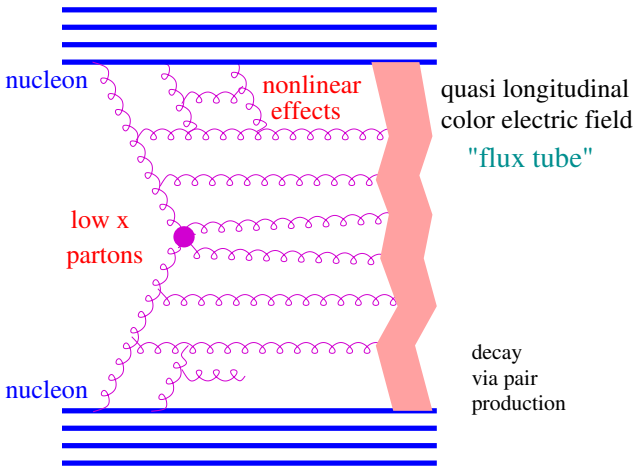


Figure 2.13: Elementary interaction in the EPOS model, figure taken from [36] reproduced with permission under licence RNP/24/AUG/082178.

Each event in EPOS-LHC has two different regions, core and corona, with very different behaviours as can be seen in Fig. 2.14. The starting point is the Pomerons. Some of these flux tubes will form bulk matter that thermalizes and expands collectively, known as the "core." Other segments, which are either near the surface or possess a large transverse momentum, will separate from the bulk matter and appear as hadrons (including jet-hadrons). This phenomenon is referred to as the "corona". Moreover, the core part is evolved using relativistic viscous hydrodynamics expansion. In this sense, EPOS-LHC includes in the core a way to manage collective effects in hadronic collisions, which makes this model especially suitable for minimum bias studies. After hadronization, there occur still hadron-hadron rescatterings via the Ultra-Relativistic Quantum Molecular Dynamics Model (UrQMD) [39].

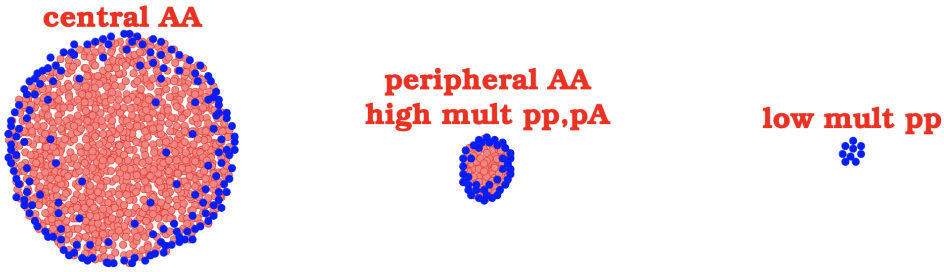


Figure 2.14: Schematic representation of core (red) and corona (blue) in a plane perpendicular to the beam axis. In the left central AA, the middle is a high multiplicity event from a small collision system or a peripheral AA, and on the right a low multiplicity pp collision. The dots are prehadrons in the transverse plane. Figure taken from Fig. [36].

3

The LHCb experiment at the LHC

3.1 The Large Hadron Collider

The Large Hadron Collider (LHC) stands as the particle accelerator with the highest energy in the world, situated within the CERN (Conseil Européen pour la Recherche Nucléaire) accelerator complex proximal to Geneva, Switzerland. This apparatus is located in a 26.7 Km subterranean tunnel, originally constructed during the 1980s to house the Large Electron-Positron (LEP) collider. Situated beneath the surface, between 45 m and 170 m, the tunnel traverses the boundary between Switzerland and France.

The primary objective of the LHC is to facilitate high-energy hadronic collisions at the highest energy and maximum instantaneous luminosity in the world. In its standard configuration, the LHC operates by accelerating two counter-rotating beams of protons within distinct circular rings. The trajectory of these beams is sustained by superconducting magnets generating a magnetic field exceeding 8 T. These magnets are housed within a cryostat containing superfluid helium, maintained at a temperature of 1.9 K to ensure optimal functionality. The refrigeration of these magnets is a pivotal aspect of LHC operation, alongside the imperative maintenance of ultra-high vacuum conditions within the beam tubes. Notably, the maximum laboratory energy achievable per proton stands at 7 TeV, thereby enabling a center-of-mass energy ($\sqrt{s}$) of colliding proton pairs of up to 14 TeV upon collision.

Prior to injection into the LHC, a proton beam undergoes a sequential progression through several preaccelerator stages: LINAC 4, Proton Synchrotron Booster (BOOSTER), Proton Synchrotron (PS), and Super Proton Synchrotron (SPS). At each stage of this

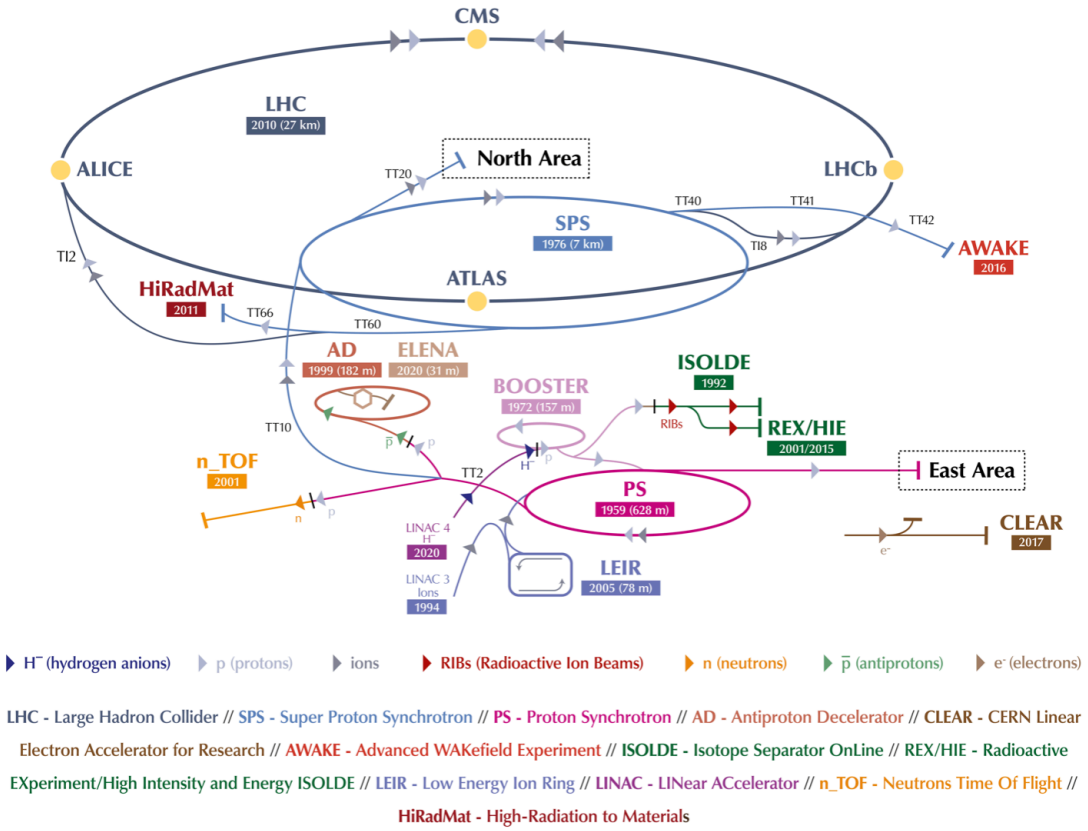


Figure 3.1: Diagram of the CERN accelerator complex in 2019. Note that the new LINAC 4 substitutes the LINAC 2, which was operating during Run 1 and Run 2. Figure extracted from [40] under CERN copyright.

process, the proton beam undergoes incremental energy augmentation, culminating in an energy level of 450 GeV upon reaching the SPS. Subsequently, the beam undergoes division into two streams and enters the LHC, where both beams are accelerated to their final energy thresholds before collision. In addition to protons, the LHC accommodates the acceleration of heavier ions, predominantly ^{208}Pb, albeit also including a brief utilization of ^{129}Xe in 2017. The ion acceleration process parallels that of protons, commencing at the LINAC 3 facility, where ions are injected into the Low Energy Ion Ring (LEIR) before proceeding through the PS along a trajectory analogous to that of protons. An unforeseen operational mode emerged in the LHC's operational paradigm, wherein proton-ion collisions are facilitated by combining proton and ion beams. The logistical flow of protons and ions across the various facilities within CERN's infrastructure can be further elucidated in Fig. 3.1.

Within the beams of the LHC, hadrons are structured into discrete bundles, termed

bunches, and they are systematically positioned along the LHC rings. This alignment of filled and vacant bunches within the LHC is denoted as the filling scheme. In proton operational mode, the interspace between bunches can extend to 25 nanoseconds, yielding a collision frequency of up to 40 MHz. Under these conditions, the peak instantaneous luminosity achievable at the LHC attains a magnitude of 10^{34} cm^{-2} s^{-1}.

Conversely, the delivered peak luminosity during lead-lead runs approximates $\simeq 10^{27}$ cm^{-2} s^{-1}, predominantly attributed to the reduced number of bunches within the LHC in this operational configuration. The beams undergo slight deviation at four specific points along the LHC's trajectory to produce hadronic collisions. Notably, the tunnel features expansive caverns at these junctures, accommodating the primary LHC experiments: ATLAS (A Toroidal LHC Apparatus), CMS (Compact Muon Solenoid), ALICE (A Large Ion Collider Experiment), and LHCb (Large Hadron Collider beauty).

The ATLAS and CMS experiments serve as two exemplary general-purpose detectors (GPDs) with commensurate scientific objectives. Both instruments explore a broad spectrum of physics, ranging from elucidating the properties of the Higgs boson to probing potential manifestations of extra dimensions or particles posited to comprise dark matter. A seminal achievement of these detectors was the landmark discovery of the Higgs boson in July 2012. The detection capabilities of ATLAS and CMS predominantly encompass the central region, corresponding to a narrow span around $\eta = 0$. This strategic positioning optimizes the detection sensitivity for the decay products of unstable heavy particles, typically characterized by substantial transverse momentum relative to the beam axis.

The ALICE experiment is specifically designed to investigate the properties of the quark-gluon plasma by scrutinizing the strong interaction dynamics within this exotic state. As explained in Sec. 2.4.2, achieving the requisite extreme energy densities necessitates the collision of heavy ions. Consequently, the ALICE detector is meticulously tailored for operation in heavy-ion collisions, which entail exceptionally high particle multiplicities (up to 3000 charged particles per event). The geometric acceptance of ALICE is distinctive owing to its detector configuration: a central barrel, covering a pseudorapidity range around $\eta \simeq 0$, is outfitted with instrumentation for particle tracking, identification, and calorimetry. This is complemented by a forward muon spectrometer spanning a region characterized by $2.5 < \eta < 4.0$. Additionally, several auxiliary detectors are strategically positioned for global event characterization and triggering within the forward region.

In contrast, the LHCb experiment primarily focuses on investigating phenomena related to CP violation, flavour symmetry breaking, and the pursuit of new physics beyond the Standard Model via indirect methods. This research domain has yielded numerous noteworthy outcomes, such as recent hints indicating the violation of lepton universality in beauty-quark decays or the first observation of CP violation in the charm sector. Furthermore, LHCb has expanded its research agenda to encompass the study of heavy-ion collisions, as evidenced by the developments outlined in this thesis. Distinguishing it from other detectors at the LHC, LHCb is situated to cover the low-angle or forward region surrounding the interaction point. A comprehensive exposition of the LHCb experiment is provided subsequently.

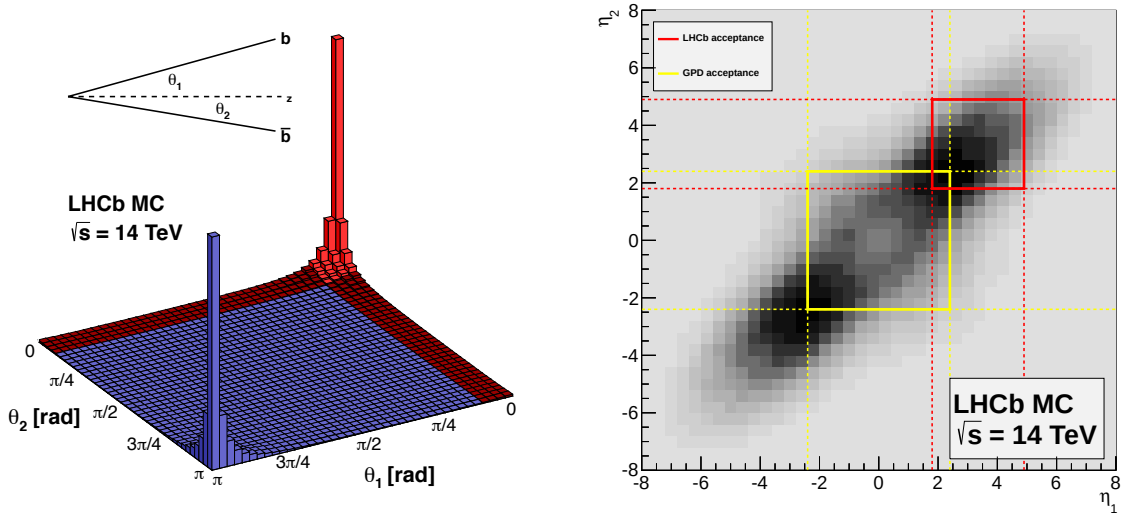


Figure 3.2: Left: Angular distribution of $b\bar{b}$ pair production in pp collisions at $\sqrt{s} = 14$TeV [41]. Right: coverage in pseudorapidity of $b\bar{b}$ pairs for LHCb and a standard GPD, considering an acceptance of $1.8 < \eta < 4.9$ and $|\eta| < 2.4$ respectively. Figures from [41].

3.2 The LHCb experiment

The LHCb experiment is situated at Point 8 of the LHC complex. It occupies the cavern previously utilized by the DELPHI (Detector with Lepton, Photon and Hadron Identification) experiment during the LEP era. Being a forward spectrometer in the forward region, the LHCb detector boasts an angular aperture spanning from 10 mrad to 300 mrad in the horizontal plane and from 10 mrad to 250 mrad in the vertical plane. This angular coverage translates into a pseudorapidity range of $2 < \eta < 5$. This distinctive forward coverage distinguishes the LHCb experiment from other LHC detectors. The strategic rationale for targeting this kinematic region lies in the optimization of acceptance criteria for B-hadrons, which constitute a focal point of the physics agenda at LHCb. Moreover, also the space limitations of the Point 8 cavern played a role in its design. Such hadrons arise from the hadronization processes associated with bottom quarks (b or $\bar{b}$), predominantly generated at low scattering angles in proton-proton collisions at LHC energies. Illustratively, Fig. 3.2 shows the angular distribution of $b\bar{b}$ pair production at a center-of-mass energy of $\sqrt{s} = 14\,\text{TeV}$, simulated with PYTHIA generator.

The detector configuration is tailored to place collisions with low pile-up and low event multiplicity in comparison with other detectors at the LHC. During proton-proton data acquisition, the instantaneous luminosity is regulated to a level ranging between $2-5\times10^{32}\text{cm}^{-2}\text{s}^{-1}$, as depicted in Fig. 3.3, a value notably lower than the one in ATLAS and CMS experiments.

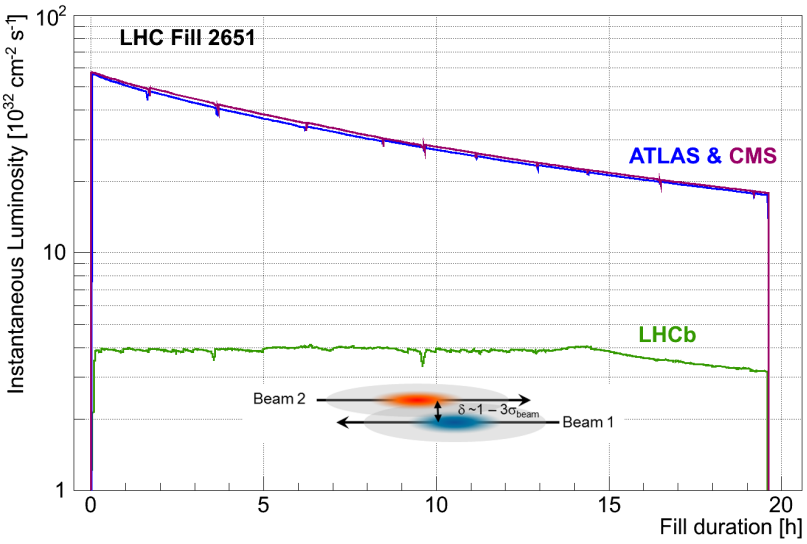


Figure 3.3: Instantaneous luminosity during LHC fill 2651 for ATLAS, CMS and LHCb experiments. Figure from Ref. [42].

This deliberate reduction of the pile-up reduces the number of simultaneous pp interactions. These design characteristics facilitate the separation of primary vertices (PV) and secondary vertices (SV). Moreover, under conditions of reduced pile-up, the average detector occupancy experiences a diminution. This makes certain detector specifications such as granularity and radiation tolerance of some subdetectors less critical.

The core of the LHCb physics program requires the meticulous reconstruction of complex decay chains ending in numerous particles in the final state [43, 44]. Having a high reconstruction efficiency and good particle identification capabilities for individual particles is necessary to be able to measure substantial samples of rare decays. Moreover, exceptional momentum resolution, particularly for charged particles, is imperative for the unambiguous discrimination between distinct resonances decaying into identical final states. Noteworthy examples include the B^0 and B_s^0 resonances [45] and the excited states of χ_{nc} [46, 47]. The degradation of momentum resolution commonly stems from multiple scattering events involving particles and the detector material. As a consequence of this, optimizing momentum resolution entails minimizing the material budget. This fundamental fact permeates the entirety of the LHCb design paradigm and hardware requirements.

In Fig. 3.4 a transverse view of the experiment is detailed, where a representation of the main detection systems is drawn. The different subdetectors from left to right are: the Vertex Locator, the RICH1 (Ring Imaging Cherenkov detector 1), TT (Tracker Turicensis), the magnet, three tracking stations (T1, T2, T3), RICH2, a muon station (M1), the calorimeter system formed by the Scintillating Pad Detector (SPD), the Preshower detector (PrS), the Electromagnetic Calorimeter (ECAL) and the Hadronic Calorimeter

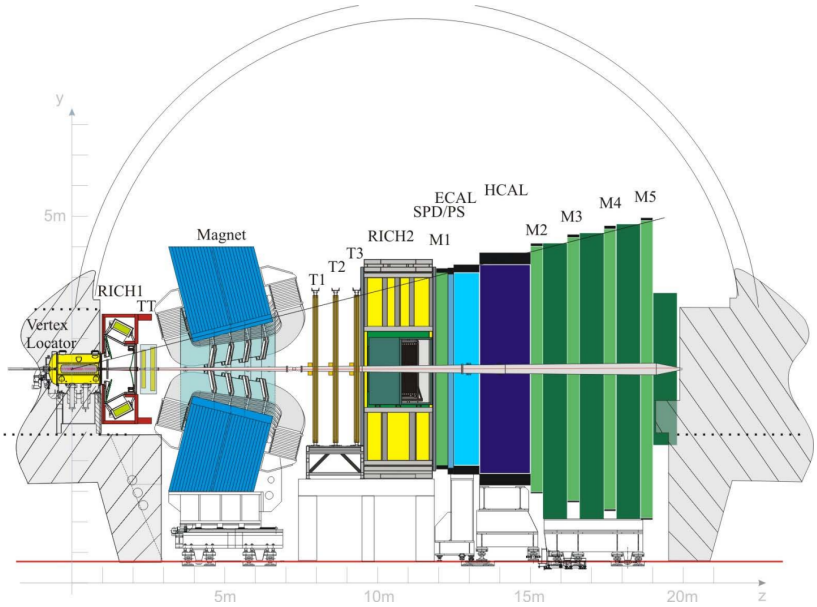


Figure 3.4: General layout of the LHCb detector. Figure from Ref. [48].

(HCAL), and the rest of the muon system stations (M2, M3, M4 and M5).

The LHCb experiment has been engaged since 2013 in the heavy-ion physics program at the LHC. This strategic decision was undertaken to leverage specific attributes of the experiment pertinent to this field. One of the features of the LHCb experiment is the acceptance coverage in the forward region, which goes beyond the scope of other detectors at the LHC. As was explained in the introduction, the Heavy Ions community wants to understand the hadron and nuclei partonic structure. Due to its forward acceptance, LHCb can probe very low x values through different observables, as depicted in Figure [REFERENCE]. While the ALICE detector also covers the forward region, its capabilities are confined to the muon spectrometer and multiplicity counters. These technical constraints limit its physics coverage. Conversely, LHCb has complete hadron reconstruction and identification capabilities in this region. Moreover, it features precise vertexing for disentangling prompt from secondary particle production and an excellent momentum resolution. The LHCb's biggest challenge in the Heavy Ion program is the very high multiplicity events, especially in lead-lead collisions where operational efficiency is restricted to approximately 60% of the most central events due to the saturation of the tracking system. Nonetheless, proton-lead collisions fall well within the multiplicity reach of LHCb.

3.2.1 Vertex Locator

The Vertex Locator (VELO) [49] is a silicon microstrip detector positioned around the beam pipe region. It is the first subdetector that particles traverse in the LHCb

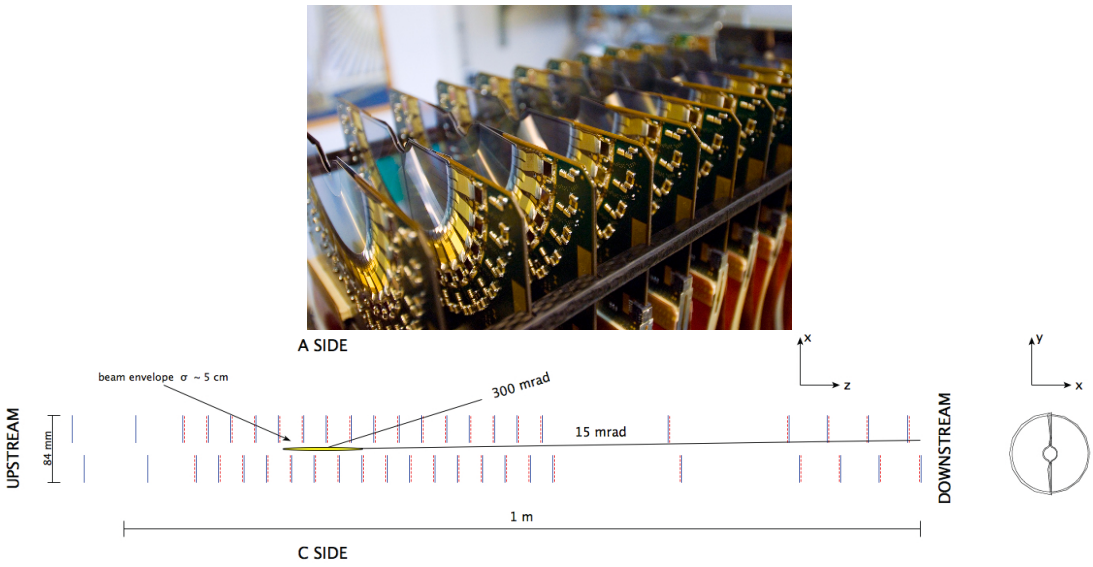
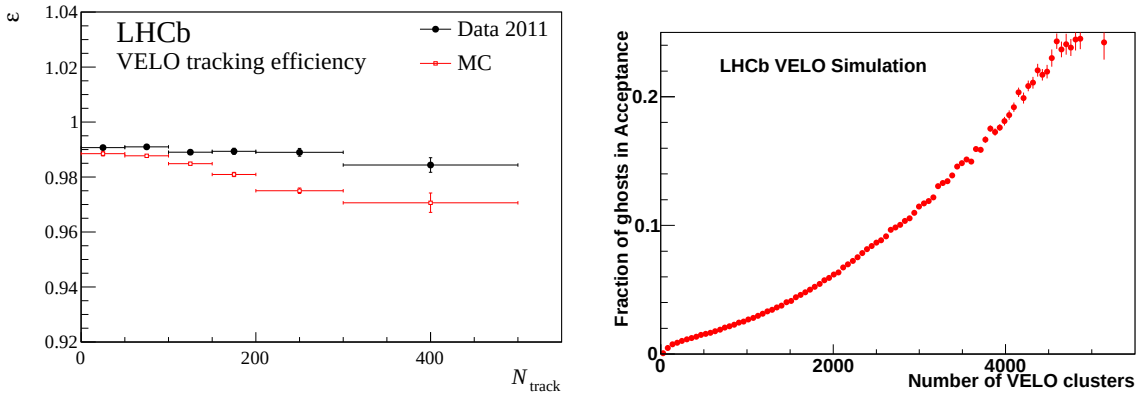


Figure 3.5: Top: photography of one side of the VELO sensors. Bottom: disposition of the VELO sensors along z and cross-section of a VELO sensor in the xy plane.

experiment. It facilitates precise determination of the coordinates of tracks originating from charged particles. This track information enables accurate localization of the PV, representing the point of interaction between hadrons from both colliding beams. Furthermore, the VELO exhibits high spatial resolution, allowing the decay points of unstable particles, such as b- and c-hadrons generated during the collision to be discerned. Despite its precise resolution, the VELO distinguishes particles produced in the primary interaction from those arising from secondary decays, facilitating precise measurements of decay lifetimes and impact parameters of particles.

The VELO contains 21 detection stations, each consisting of two semicircular silicon sensors: one of type R, responsible for measuring the radial coordinate, and one of type ϕ, dedicated to the azimuthal coordinate. These sensors are arranged longitudinally along the beamline, spanning approximately one meter in length, as shown in Fig. 3.5. With a radius of approximately 42 mm and a thickness of around 300 µm, each sensor has an 8 mm aperture at its center where the LHC beam travels through. Type R sensors utilize azimuthal strips to measure the radial coordinate, while type ϕ sensors employ radial strips to gauge the angular coordinate. Charged particles generated in collisions traverse the silicon detectors, inducing the generation of electron-hole pairs; subsequently, the resultant current is captured using specialized electronics. As a result, the direction of the particles can be reconstructed.

During operation, the inner part of the sensors is positioned 7 mm away from the LHC beams, which is smaller than the required aperture during injection. To avoid issues with the beam at injection, the VELO is divided into two retractable halves that close when stable-beam conditions are reached. The VELO sensors are within a secondary



(a) (b)

Figure 3.6: VELO tracking efficiency (left) in data and simulation from 2011 as a function of the number of reconstructed tracks in the event, N_{track}. Error bars show the statistical uncertainty. Fraction of ghost tracks in the VELO (right) as a function of the number of VELO clusters. Figures extracted from Ref. [49].

vacuum, isolated from the primary vacuum where the primary collisions occur in the LHC. The region in contact with the beam, called RF foil, is 0.3 mm thick. The thickness of this section is important to minimize the detrimental effects of multiple scattering while measuring particle momentum. The composition of the VELO material is meticulously characterized in many studies [50] and incorporated into the LHCb simulation framework.

In terms of performance, the sensors exhibit a signal-to-noise ratio of around 20, with a best hit resolution of 4 µm achieved at an optimal track angle. For vertices with 25 tracks, a transverse plane resolution of 13 µm and a longitudinal resolution of 71 µm is obtained. Moreover, an impact parameter resolution of less than 35 µm is achieved for particles with transverse momentum (p_{T}) greater than 1 GeV/c. The track reconstruction efficiency typically exceeds 98% in data, as can be seen in Fig. 3.6a. Another pertinent parameter when characterising a silicon tracker is the fraction of ghost tracks. Those are defined as tracks where less than 70% of the VELO clusters belong to a simulated particle, which means that these tracks are likely to be reconstruction artefacts not belonging to a particle. This fraction is approximately 0.5% for randomly triggered events and 1% for triggered events. The fractions of ghost tracks in the VELO increase with detector occupancy, as shown in Fig. 3.6b.

In Fig. 3.6b the total number of VELO clusters in the event, $N_{\text{VELO}}^{\text{clusters}}$, is the chosen variable to quantify detector occupancy. These performance metrics are based on standard pp collision operations and average events. For proton-lead or lead-lead collisions, where the detector occupancy reaches higher levels, the performance in the parameters that we have discussed decreases.

3.2.1.1 Primary vertex reconstruction

At LHCb, primary vertices are reconstructed using tracks from the VELO [51], i.e. Velo Tracks. Those tracks follow straight lines due to the absence of a magnetic field in the VELO. The reconstruction algorithm examines the distribution of tracks in an event to infer where most tracks originate.

However, optimizing the PV reconstruction has two challenges: high efficiency and a low rate of false positives. By high efficiency, it is understood that a high proportion of generated vertices are reconstructed. False PVs can arise from randomly linking tracks together. This effect increases with the number of tracks in the event.

The primary vertex reconstruction is done following two steps: seeding and fitting. First, the PV candidates are identified by looking for clusters of track trajectories. Second, by using the weighted least squares method the final position of the vertex is determined. Vertices are fitted starting from the highest multiplicity to the lowest. This is done in order to avoid secondary vertices absorbing tracks for the primary vertices. After fitting, we double-check the distance between the new PV and the existing ones to catch any false positives. We repeat these steps until we've identified all the PVs in the event.

3.2.2 Tracking system

One of the main parts of the LHCb experiment is the tracking system. The goal of this subsystem is to identify the trajectories of charged particles and measure their momentum. The track reconstruction is done using information from several subdetectors. Specifically: the VELO (already discussed in Sec. 3.2.1), the magnet, the silicon tracker (ST) and the outer tracker (OT) were the ones used during Run2. In this section, all of them will be summarized.

3.2.2.1 Magnet

In the LHCb experiment, a dipole magnet generates the magnetic field between the upstream and downstream tracking stations. This magnetic field bends the trajectory of charged particles and enables the determination of their momentum coordinates. The total magnetic field produced by the magnet is 4 T m. To explore potential systematic uncertainties and asymmetries, the magnet polarity is regularly switched during data collection into *Magnet Down* and *Magnet Up* polarization.

The magnetic field integral, $\int B\mathrm{d}l$, created by the magnet is known with a relative precision of 10^{-4}. Moreover, the position of the peak magnetic field, B-field, is determined with a precision of a few millimetres. The precise characterisation of the magnetic field properties is crucial in order to achieve the desired momentum resolution. Due to this, precise measurements of the magnetic field are conducted across all regions of the LHCb detector.

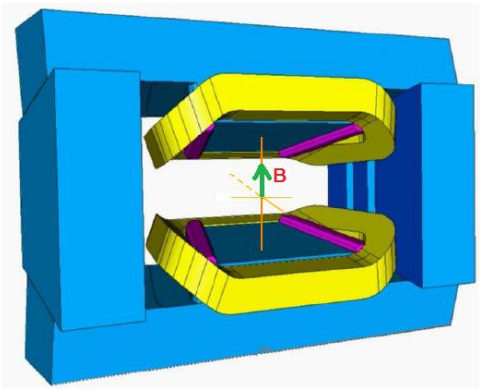


Figure 3.7: Schematic representation of the LHCb experiment dipole magnet and the produced magnetic field.

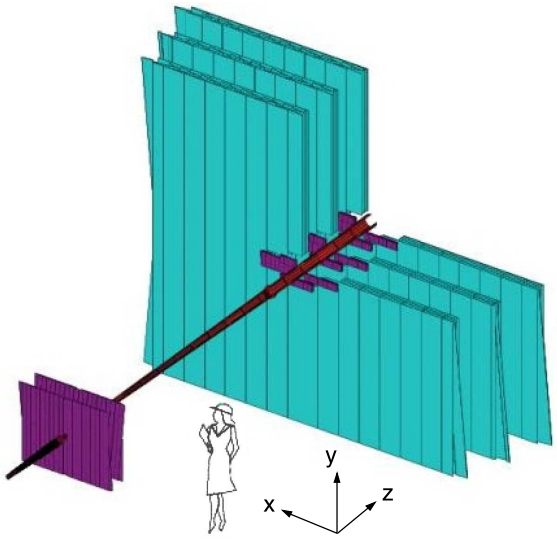


Figure 3.8: Illustration of the tracking stations as located in the beam pipe. The ST stations (TT and IT) are shown in violet. The cyan part corresponds to the OT.

3.2.2.2 Silicon tracker

The Silicon Tracker (ST) is formed by two detectors around the beam pipe: the Tracker Turicensis (TT), positioned upstream of the magnet, and the Inner Tracker (IT) situated within the inner region of tracking stations T1, T2, and T3. Both the TT and IT employ silicon microstrip sensors with a strip pitch of $\simeq 200\,\mu\text{m}$. In Fig. 3.8, an illustration of the tracking stations can be seen.

Positioned immediately before the dipole magnet, the TT constitutes a planar tracking station with dimensions 150 cm in width and 130 cm in height, covering the entire detector acceptance. The primary role of the TT is to provide information to connect tracks coming from the VELO and those in the tracking stations downstream of the

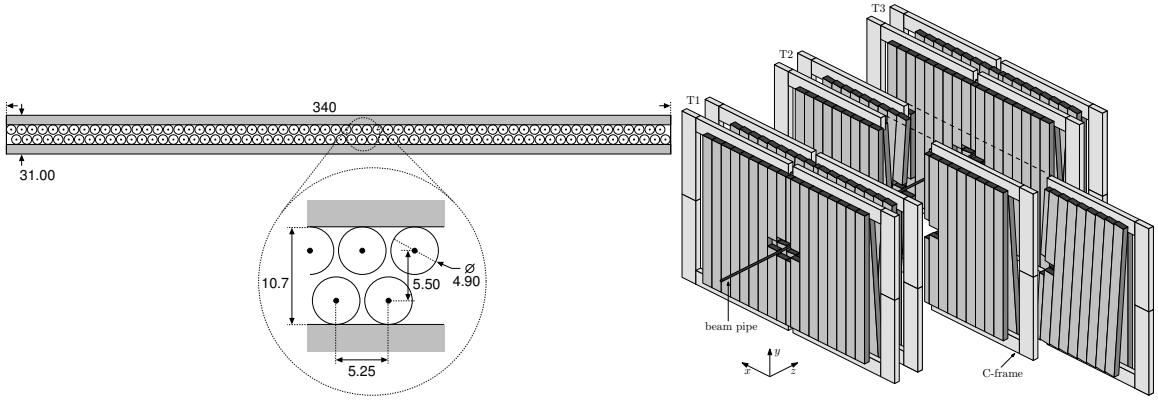


Figure 3.9: Fig. a) corresponds to the OT module cross-sections where the tubes can be seen. Fig. b) shows each station of the OT system.

magnet.

Conversely, the IT is positioned downstream of the magnet and involves three subdetectors placed around the beam pipe T1, T2, and T3. Each IT layer measures 120 cm in width and 40 cm in height. Within the four ST stations, each layer is composed of four detection layers arranged in a stereo configuration. The external layers feature vertical strips, while the internal layers have strips oriented at an angle of $\pm 5°$ with respect to the vertical axis.

3.2.2.3 Outer tracker

The Outer Tracker (OT) is the sensitive area within stations T1-T3 [52, 53] that does not belong to the IT as it is shown in Fig. 3.9. It is a gaseous detector made of straw-tube drift chambers organized into modules. Each module has two staggered layers of drift tubes with inner diameters of 4.9 mm. In each station, these modules are arranged in four layers, mirroring the configuration of the ST layers. Together, they cover a sizable active area measuring $5971 \times 4850\,\mathrm{mm}^2$.

As can be observed at Fig. 3.8 the inner part of the tracking stations is made out of silicon technology and does not use a drift chambers system. This decision was taken because is a region where a high density of particles per angle near is produced. In order to achieve the requirements to operate in this environment, the drift chamber technology cannot be used.

3.2.2.4 Track reconstruction at LHCb

The LHCb reconstruction software system uses information from different tracking detectors. Combining that information with the mapping of the magnetic field it is capable of reconstructing the trajectories and momenta of charged particles. The process of

correlating hits at different detectors and matching them into tracks is done using several algorithms. Track reconstruction within the LHCb experiment is an integral component of the broader event reconstruction application known as Brunel.

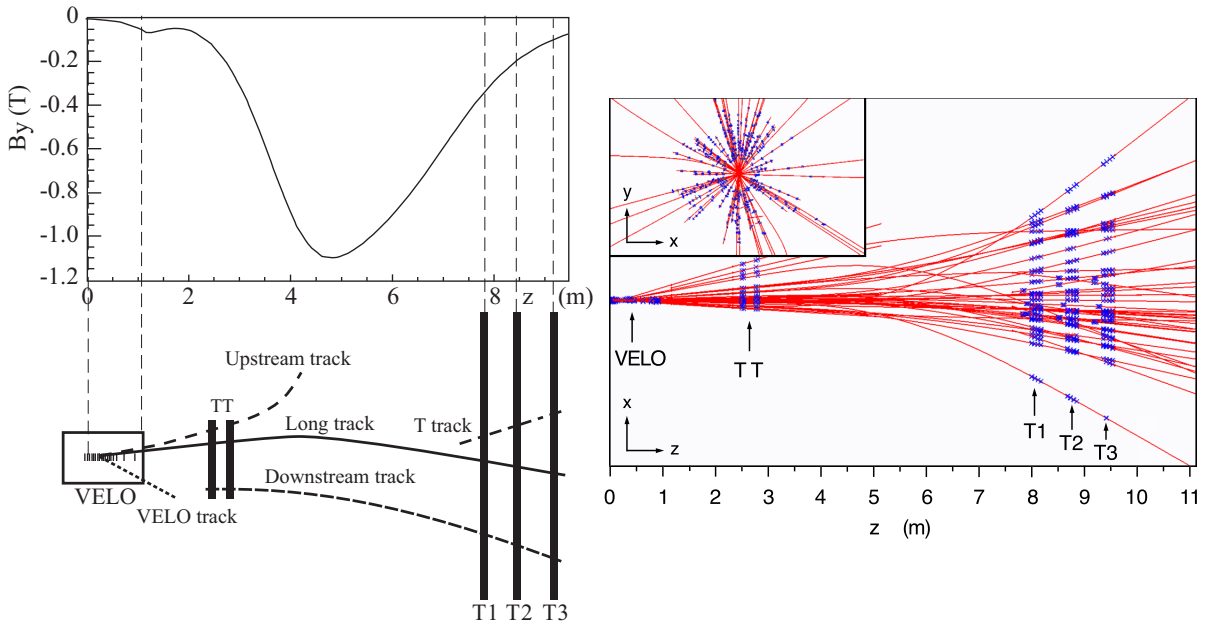
Depending on the origin of the hit used to build the track, they can be classified into:

- **VELO track:** Build from only VELO hits, they don't have hits in the other tracking detectors. They can have a broader angle and be at backward rapidity.
- **T track:** They have only hits in the T stations (T1, T2 and T3).
- **Upstream track:** These tracks are built out of hits in the VELO and TT stations. In general, they have low momentum and they are removed by the magnetic field. For these tracks, the momentum measurement is poor and they are used for RICH1 calibration and to reconstruct some specific decays.
- **Downstream track:** They are made out of hits in the tracking stations after the magnet. They can be used to reconstruct decays of long-lived particles that decay after the VELO like K^0_S and Λ.
- **Long track:** They are the best kind of tracks in terms of momentum resolution and they are used in the majority of the physics analysis within LHCb. They traverse the full tracking system from the VELO to the T-stations, and optionally they can hit the TT.

In Fig. 3.10a we can see a scheme of the different tracks mentioned above. The reconstruction algorithm starts looking for track seeds in the VELO [54] and the T-stations. In these regions, the magnetic field is small as can be seen in Fig. 3.10a. After this step, those track candidates are combined with hits in the T-stations using different algorithms [55, 56]. After long-track candidates have been found, their trajectories are refitted with a Kalman filter [57]. This step considers corrections from multiple scattering and energy loss. The quality of the fit is given by the χ^2 per degree of freedom, χ^2/ndf. This variable can be used later to select the quality of the tracks in a physics analysis. Fig. 3.10b shows the reconstructed tracks in a typical event.

In the LHCb data flow, a DST is the data format that stores event information. Next to the reconstruction stage, tracks are saved into DST files in the `TES["Rec/Track/Best"]` directory. Here TES stands for `Transient Event Storage`. In this directory, only the "best" tracks are stored. This means that only the *highest level* tracks are saved, i.e. if a VELO track is matched with a T-track to form a long track, only the long track will appear.

Apart from tracks corresponding to a real particle trajectory, several track types can be saved into the DST files. For example, because of the distance between VELO and T-stations (more than 5 m), the VELO-track to T-track matching is sensitive to create *fake tracks*. These tracks, sometimes also called *ghost tracks*, do not correspond to a particle trajectory being reconstruction artefacts created by the algorithm. The content of the



(a) (b)

Figure 3.10: In the left there is the LHCb's tracking scheme where the different kinds of tracks provided by the reconstruction software can be seen (bottom) and magnetic field profile (top). On the right side, it is shown the display of the reconstructed tracks for a specific event and the assigned hits in the xz plane for the different tracking detectors. Figures extracted from Ref. [48].

container `TES["Rec/Track/Best"]` will be studied in detail in this document Chap. 5. A critical part of this work is to remove all tracks not corresponding to particle trajectories from our data set.

As with every measurement tool, the LHCb tracking system is imperfect. Two kinds of imperfections are important for the goal of this thesis: fake signals and inefficiencies. The first one was already explained in the paragraph above. By inefficiency, we mean that the tracking system can miss charged particle trajectories and not reconstruct its track. The performance of a tracking detector regarding this can be quantified using *tracking efficiency*. As for other technical capabilities, such as particle identification, the tracking efficiency can be estimated using Monte Carlo simulation. This simulation includes the detector description and is very useful for evaluating the impact of the detector hardware and software on the measurements. However, as the simulation is imperfect and can be incomplete, a data-driven approach is necessary to know precisely about the tracking detector's capabilities.

The data-driven technique used to compute tracking efficiency at LHCb consists of a tag-and-probe technique using $J/\psi \to \mu^+\mu^-$ decays as detailed in [58]. One of the muons is fully reconstructed, the "tag" one, while the "probe" is partially reconstructed. Then,

an invariant mass fit to the di-muon distributions is done to compute J/ψ candidates. Finally, a search for long tracks matched to a partially reconstructed probe is made. Tracks under those criteria are marked as efficient. The final efficiency is obtained with the ratio of efficient candidates over the total number of candidates. This technique measures the tracking efficiency in data and simulated events to correct discrepancies between simulation and data. The results from this calibration procedure are used in the present analysis and are described in Sec. 6.5.

3.2.3 Particle identification system

Apart from detecting a particle's trajectories and measuring its momentum, being able to associate these tracks to particle species is a very important step in a particle detector operation. This is called particle identification (PID). Typically the mass difference between particles is often used as an advantage to make this classification. The identifiable particles at the LHCb experiment are γ, e, μ, $\pi^{\pm}$, $K^{\pm}$ and p. Particle identification is made using information from several subdetectors like the RICH detectors, the calorimeters and the muon system.

The information from these subdetectors is integrated into variables like a combined likelihood (DLL) or an Artificial Neural Network output (ProbNN). These variables are optimised to maximise the identification efficiency and minimise the misidentification rate. The identification efficiency varies between $90 - 100\%$ for different particles. In the next sections, the different subdetectors involved in PID at the LHCb experiment will be briefly explained.

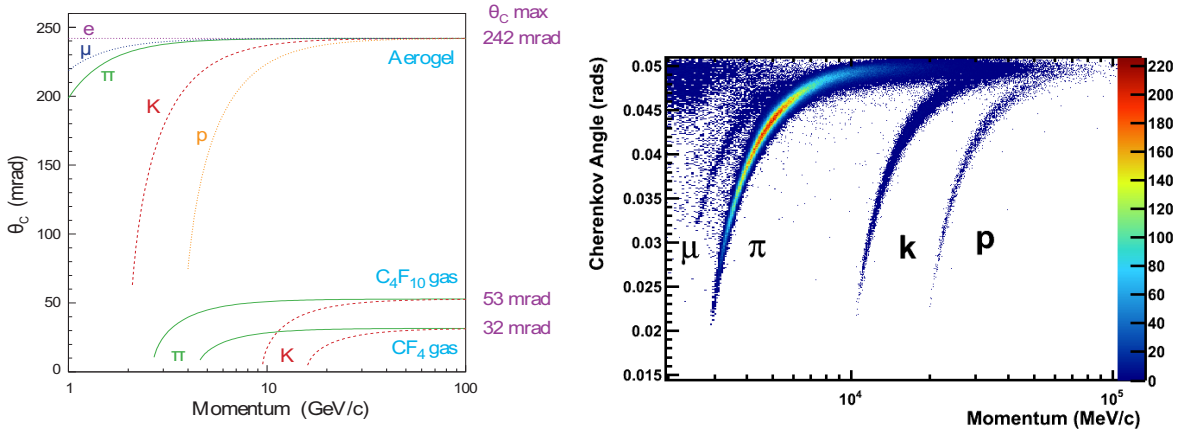
3.2.3.1 RICH system

The first subdetector involved in PID that particles reach is the RICH (Ring Imaging Cherenkov) system [59]. This system involves two detectors: RICH-1, located between the VELO and TT, and RICH-2, between T3 and M1. The main purpose is the identification of charged hadrons; i.e. π, K and p. However, they also contribute to lepton identification; i.e. e and μ in the low p region.

A Cherenkov detector works based on the emission of Cherenkov light rings emitted by particles traversing a medium with refractive index $n > 1$. The identification is achieved by a measurement of the angle of emission, θ_c, of the Cherenkov radiation, which is related to the charged-particle velocity, v, by

$$\cos\theta_c = \frac{c}{nv}, \tag{3.1}$$

where c is the speed of the light. The velocity is estimated by the LHCb tracking system. With the velocity information and using the known n of the RICH radiator gives a corresponding prediction of θ_c that can be compared to the θ_c of the detected Cherenkov photons thus indicating the particle's identity. With its silica aerogel and C_4F_{10} gas radiator, the RICH-1 detector is therefore set up to detect low-momentum



(a)

Figure 3.11: In the left: Reconstructed Cherenkov angles as a function of the track momentum in the different RHIC radiators. In the right θ_c, p and particle type distribution from RICH detector at LHCb. Figures taken from Ref. [48].

particles, while RICH-2 has an acceptance that is limited to the low-angle region where there are mostly high-momentum particles and uses a CF_4 gas radiator. The use of two Cherenkov detectors with different radiators provides wide coverage in the momentum $2 - 100\,\text{GeV}/c$. In Fig. 3.11 we can see the dependence of the Cherenkov angle with the particle momentum for each particle type and radiator. The relation between the Cherenkov angle, θ_c, and the momentum for each particle can be seen in Fig. 3.11a.

Particles produced will travel through the mirrors of RICH1 before passing through the tracking system and have a momentum measurement. To reduce the amount of scattering and not worsen the momentum resolution, RICH-1 uses special lightweight spherical mirrors constructed from a carbon-fibre reinforced polymer (CFRP), instead of using glass. As RICH2 is located after the tracking system and the magnet, glass could be used for its spherical mirrors.

As we already mentioned, the two RICH systems are optimised for different momentum regimes. On the one hand, the RICH1 covers the range $2 - 40\,\text{GeV}/c$ over the full spectrometer angular acceptance of $25 - 300\,\text{mrad}$. Here the acceptance is limited due to the size of the beam pipe upstream of the magnet. On the other hand, the RICH2 system covers the high-momentum region in the range $15 - 100\,\text{GeV}/c$, over the angular range $15 - 120\,\text{mrad}$. As we can note the RICH2 has a coverage closer to the beam. When a particle goes through the radiator medium the photons emitted are focused into ring images on the photon detector, situated outside of the experiment acceptance as can be seen in Fig. 3.12. There a hybrid photodetector (HPD) design for the RICH detector is used to transform the light into electric current. This HPD employs vacuum tubes with a 75 mm active diameter, with a quartz window and multialkali photocathode. The photoelectrons are focused onto a silicon pixel array of pixel size $2.5 \times 2.5\,\text{mm}^2$, using an accelerating

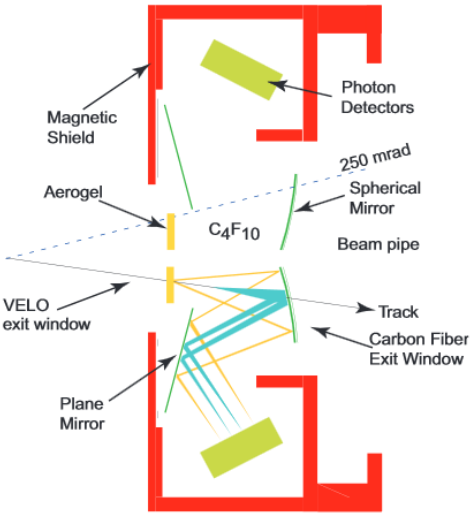


Figure 3.12: Diagram of the RICH1 system including the photodetectors, radiator and mirrors.

voltage of −16kV. A total of 484 HPDs are close-packed to cover the four photodetector planes. An important goal of the RICH design is to have a precise determination of the Cherenkov angle.

3.2.3.2 Calorimeter system

Apart from measuring the particle trajectory and identifying the particle species it is important to measure the particle energy. That is the purpose of the calorimeter systems. To do that the particle is stopped and its energy is absorbed by a medium. Because of that reason, the calorimeter is placed after the tracking and RICH subdetectors in the detector design.

At the LHCb experiment, the calorimeter system [60] is composed of several subdetectors: the Scintillating Pad Detector (SPD), the Pre-Shower Detector (PrS), the Electromagnetic Calorimeter (ECAL) and the Hadron Calorimeter (HCAL). Individually, each one of these subdetectors plays a specific role and as a whole, they enable the identification of γ, e, μ, π as well as the reconstruction of π^0, η mesons offline. This is done by the measurement of their energy and location. As with many other subdetectors in the LHCb experiment, the calorimeter system information is used in the fast hardware trigger L0 [61]. Specifically, the deposited transverse energy E_{T}. Each system and its specific function are summarised in the following.

1. **SPD and PrS:** These detectors are two walls of scintillator pads with a thickness of 15 mm separated by a lead curtain with a thickness corresponding to 2.5 radiation lengths, X_0. The emitted light is collected by wavelength-shifting (WLS) fibres. The

WLS fibre is used to transmit the light to multi-anode photomultipliers (MAPMTs) located at the periphery of the detector. The SPD determines whether particles hitting the calorimeter system are charged or neutral by comparing the deposited energy with a threshold. The PrS indicates the electromagnetic character of the particle (i.e. whether it is an electron if charged, or a photon, if neutral). They are used at the trigger level in association with the ECAL to indicate the presence of electrons, photons, and neutral pions.

2. **ECAL:** This detector is located next to the PrS and is designed to measure showers made of electrons and photons. The ECAL is subdivided into three sections, Inner, Middle and Outer, comprising modules of different granularities. Alternate layers of scintillating material and lead absorber are mounted in a *shashlik* structure. The cell size varies from $4\,\mathrm{x}\,4\,\mathrm{cm}^2$ in the inner part of the detector, to $6\,\mathrm{x}\,6\,\mathrm{cm}^2$ and $12\,\mathrm{x}\,12\,\mathrm{cm}^2$ in the middle and outer parts. The overall detector dimensions are $7.76\,\mathrm{x}\,6.30\,\mathrm{m}^2$ covering an acceptance of $25\,\mathrm{mrad} < \theta_c < 300\,\mathrm{mrad}$ in the horizontal plane and $25\,\mathrm{mrad} < \theta_c < 250\,\mathrm{mrad}$ in the vertical.

   At heavy ion collisions, the ECAL energy per event is used to estimate the impact parameter of the collision through the centrality computation.

3. **HCAL:** It is located next to the ECAL and its internal structure consists of thin iron plates interspaced with scintillating tiles. The HCAL measures the transverse energy of hadrons both at 40 MHz for the L0 trigger and for contributing to the offline PID. It has a thickness is 5.6 interaction lengths.

3.2.3.3 Muon system

In the LHCb experiment, the muon detector system is designed to send binary information to the data acquisition (DAQ) and the hardware processors of the muon trigger (L0MU). This together with the calorimeter trigger constitutes the bulk of the first level trigger. The muon system [62] is composed of five stations M1-M5 placed along the beam axis. Stations M2-M5 are placed after the calorimeters and are interleaved by $80 -$ cm thick iron absorbers to filter muons. Station M1 is located downstream of the RICH2 station and before the calorimeters to improve the p_{T} measurement in the trigger. This avoids the impact of the multiple scattering in the absorbers. Each station is equipped with 276 multi-wire proportional chambers except for the inner part of the first station, subject to the highest radiation, which is equipped with 12 Gas Electron Multiplier detectors. The geometry of the five stations is projective. The transverse dimensions of the stations scale with their distance from the interaction point. The chambers are positioned to form, across the stations, adjacent projective towers pointing to the beam crossing position.

The L0MU trigger processors perform a stand-alone muon track reconstruction which requires finding hits in all the 5 stations and calculating the transverse momentum of the tracks. Muon candidates are accepted if their p_{T} is above a given threshold.

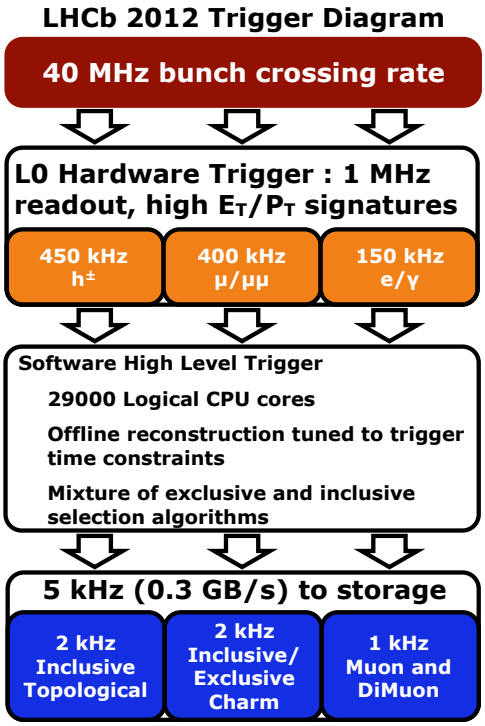


Figure 3.13: Diagram of the trigger system during Run1. Figure extracted from [63].

Muon detection is crucial for the Heavy Ion program because it is the main decay where quarkonia states are measured. The cleaner and more precise is the muon measurement the better will be the separation power and statistics between different quarkonia states.

3.2.4 Trigger system

Once the particles produced at hadronic collisions go through the LHCb detector, the produced information at each subdetector has to be processed. As proton bunches are separated by 25 ns the amount of data produced is huge. That means that data has to be filtered and selected before being saved. The goal of the *trigger* system is to select the interesting events and information in order to fit the technical requirements which constrain the amount of information that can be saved and stored. The LHCb trigger system uses information collected by the different subdetectors to reduce the event rate, from the 40 MHz rate of bunch crossings at the LHC to about $3 - 5$ kHz as can be seen in Fig. 3.13. This reduction implies that only interesting events are fully processed and stored by the experiment. The trigger is composed of two levels: Level-0 (L0) and High-Level Trigger (HLT).

The first stage of the trigger system, L0, is implemented on hardware. The L0 trigger reduces the event rate from the LHC delivered standard of 40 MHz to 1 MHz. It uses

	hadronic	muonic		electromagnetic		
		μ	di-μ	e	γ	π^0
threshold						
p_T/E_T (GeV)	3.5	1.3	$\sum > 1.5$	2.6	2.3	4-4.5
rate(kHz)	700	200		200		

Table 3.1: Summary of the L0 trigger conditions with an event output rate of 1.1 MHz. It is built as a logical OR of the hadronic, the muonic and the electromagnetic trigger decision. About 10% of the events are triggered by more than one L0 trigger. Information extracted from [64].

information from the calorimeter system together with the SPD and PrS, the muon system and the pile-up system of the VELO. The L0 trigger reconstructs and selects particles with high transverse momentum in the muon chambers or with high transverse energy in the calorimeter system. The pile-up veto provides a rejection of events with multiple proton-proton collisions through the reconstruction of all primary vertices in the event. Events with two interactions are identified with an efficiency of 60% and a purity of about 95%.

An L0 hadron trigger decision is taken if the transverse energy of a cluster of hadronic type is above 3.5 GeV. An L0 muon trigger is issued if the muon candidate with the highest p_T is above the threshold of 1.3 GeV or if the sum of the two highest p_T candidates is above 1.5 GeV. The total output rate of the L0 trigger is 1.1 MHz, and about 10% of the events are triggered by more than one L0 trigger type. A summary of the trigger conditions at the L0 level is shown in Tab. 3.1.

After the hardware trigger L0 is passed, a software-based trigger called High-Level Trigger (HLT) is placed. The HLT is run on computing nodes in the Event Filter Farm (EFF). It is a C++ application running on L0 accepted events at a rate of 1.1 MHz. This state of the data processing reduces the event rate from 1 MHz down to 5 kHz using a full event reconstruction and a more exhaustive selection. Because it is software-based, it is more flexible than the L0 and can be tuned to meet the requirements of different physics analyses. The structure of the HLT consists on two steps, i.e. the HLT1 and HLT2.

At the HLT1 stage, the rate is reduced from 1.1 MHz to tens of kHz by applying different algorithm sequences. The first stage, HLT1, carries out a partial event reconstruction and selects potential signal candidates in an inclusive manner.

At a reduced rate of 80 kHz, HLT2 performs a full event reconstruction with only slight modifications compared to the offline reconstruction process. Following this reconstruction, a combination of inclusive and exclusive selections further reduces the trigger rate to 5 kHz, which is then saved for subsequent offline analysis. These rates are average values from the 2012 run of the LHC; in 2011, the output rate from HLT1 was around 40 kHz, and from HLT2, it was 3 kHz.

Finally, to know what trigger conditions were applied in a specific run of data-taking the *Trigger Configuration Key* (TCK) is used. A TCK includes all trigger lines and their

specific threshold, requirements and prescales that were implemented during the detector operation in data-taking. The TCK is unique for each run. The LHCb trigger framework is included in a software package known as Moore.

3.2.5 Simulation and data flow at LHCb

In particle physics, Monte Carlo simulations are crucial in many steps of the research. For example, to perform complex calculations or to estimate the detector impact on a measurement. In the LHCb collaboration, a common tool and software framework is used to produce high-quality and reproducible simulation samples. As geometry and detector properties are encoded in software description using GEANT4, simulation can be used to estimate detector efficiencies, resolutions and more. Simulation computation is often a resource-demanding activity, for this reason, HPC clusters are used.

The LHCb simulation framework called Gauss [65] includes several applications which perform different steps in the simulation chain:

1. **Generation:** This part is in charge of producing the hadronic collisions. As proton-proton collisions are different from ion-proton or ion-ion collisions, the generators are also different. While for proton-proton collisions the Pythia event generator [66] is used, for collisions involving ions the EPOS-LHC [36] generator is used. As some steps in the hadronic physics are not fully understood, generators are complex and include several *tunes*. Those are different implementations or modifications of the same model. Their differences are explained in Sec. 2.5.

2. **Particle evolution:** Here the signal particle is decayed using the EvtGen [67] package.

3. **Detector:** After all final state particles are produced and evolved, they are transported through the detector material using the Geant4 [33, 68] software environment. An accurate detector description is important to have reliable simulations that reproduce well the detector effects. Detector conditions can vary over time, the specific detector conditions used for a specific run are encoded in two keys: the description database (DDDB) and the condition database (CondDB).

After simulating the physics of hadronic collisions, particle evolution and decay and detector interactions, the data processing steps are also simulated. The first step after Gauss is the digitization of the particle signals in the detector using Boole [65]. This step mimics the LHCb's DAQ systems. Then the digitized signals can be processed using the reconstruction software using Moore (see Sec. 3.2.4) as during online data taking. Next to Moore the event reconstruction software Brunel is used to perform event reconstruction. The output file after reconstruction is a `.dst` file. After the reconstruction is complete, some filtering steps are performed to get the final data files. The *Stripping* is the filtering step done by the DaVinci software application to optimise data storage

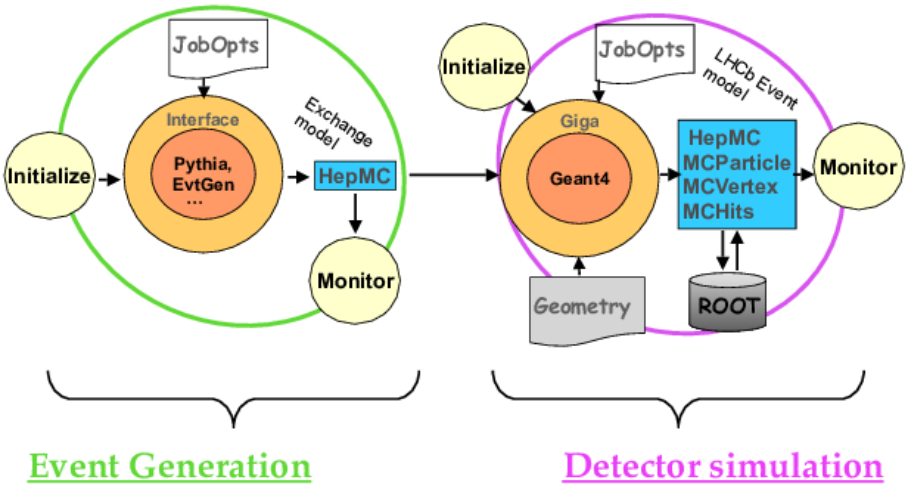


Figure 3.14: Schematic view of the LHCb simulation framework GAUSS. Figure taken from [69].

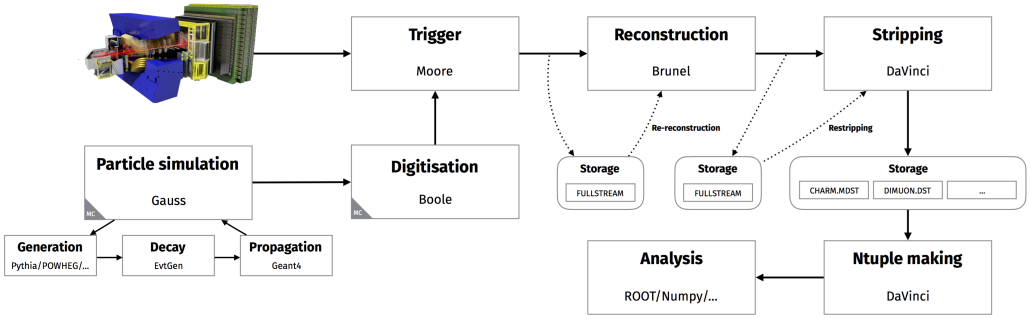


Figure 3.15: Schematic view of the LHCb data flow. Figure taken from [69].

resources. The output of this filtering step is already a `.root` file, which is the data format used by the analysts in general. The description of the datasets used in this analysis, both simulation and data, is addressed in the next chapter.

4

Description of the datasets

In this section, the data sets used in the analysis will be explained. Since the conditions of proton-proton are different from the proton-nucleus samples each configuration will be discussed in a different subsection.

4.1 Proton-lead data

The considered $\sqrt{s_{NN}} = 5.02$ TeV proton-lead (pPb and Pbp) data were taken in January-February 2013. The energy per nucleon of the lead beam and the proton beam was 1.58 TeV and 4.00 TeV, respectively. This means an energy in the nucleon-nucleon centre-of-mass system of $\sqrt{s_{NN}} = 5.02$ TeV. The asymmetry in the beam energies produces a boost of $y_{boost} = 0.465$ in the direction of the proton beam. This data set corresponds to LHC fills in the [3510, 3544] range and amounts to a total integrated luminosity of $\mathcal{L}_{int} \approx 1.6\,\mathrm{nb}^{-1}$. During this data-taking period, the average instantaneous luminosity was $\mathcal{L} = 3 \times 10^{27}\,\mathrm{cm}^{-2}\,\mathrm{s}^{-1}$. Consequently, the average number of visible pPb and Pbp interactions per bunch crossing is $\mu \approx 0.04$, which is a lower value than that of pp collisions. For pPb and Pbp samples, the polarity of the magnet was switched so data sets are divided into Magnet Up (MU) and Magnet Down (MD) configurations.

The hardware trigger (L0) was operated in a pass-through mode redirecting all events to the high-level software trigger (HLT). This analysis makes use of a minimum bias trigger which selects events with at least one reconstructed VELO track. This condition has been implemented in the trigger line `Hlt1MBMicroBiasVelo`. Besides, no particular HLT2 selection is required.

The reconstruction software version for this data-taking was Reco14r1. In this reconstruction, a baseline selection for the tracks is applied consisting of a track quality selection of $\chi^2/\text{ndf} < 3$.

After the data were taken a stripping selection was applied. Minimum-bias lines correspond to the stripping version v20r3. For these lines, the stripping consisted of prescaling events with factors shown in Tab. 4.1.

Sample		ODIN RunNumber	Stripping Line	Prescale
pPb	MD	(136799 - 137045)	StrippingProtonIonMagDownMinBiasLine	0.025
	MU	(136417 - 136758)	StrippingProtonIonMagUpMinBiasLine	0.110
Pbp	MD	(135576 - 136212)	StrippingIonProtonMagDownMinBiasLine	0.08
	MU	(136237 - 136341)	StrippingIonProtonMagUpMinBiasLine	0.09

Table 4.1: Stripping lines and prescale factor for minimum bias data at pPb and Pbp collisions.

TCK	**Comments**
0x00691710	2013 Physics, pPb run, base do not use for data taking
0x006a1710	2013 Physics, pPb run default
0x006b0048	2013 Physics, pPb run. L0 SPD0 line.
0x006c0048	2013 Physics, pPb run. L0 SPD0 line.
0x006d0048	2013 Physics, pPb run. As 0x006c0048 with NoBias prescale 0.05.
0x006e0048	2013 Physics, pPb run. As 0x006c0048 with NoBias prescale 0.05.
0x006e0049	2013 Physics, pPb run. Et>400 in b1gas.
0x006e004a	2013 Physics, pPb run. Et>300 in b1gas.
0x006f004a	2013 Physics, pPb run. As 0x006e004a with Hlt1MBMicroBiasVelo prescale 0.01.

Table 4.2: TCK information from LHCb system implemented in pPb and Pbp 2013 data. The TCKs in red are not considered.

Some data samples are not considered in this analysis due to data quality criteria, as shown in Tab. 4.2. Because of a muon inefficiency problem events with the TCK[1] `0x006a1710` and also runs (135710, 135711, 135900, 135901) are excluded. Due to having a `sumEt` threshold, data corresponding with TCK `0x006e0049` and `0x006e0049a` are not considered either. Data from TCK `0x006f004a` are disregarded because they are strongly prescaled with a factor of 1/100 in the `Hlt1MBMicroBiasVelo` trigger line. Finally, runs (136729 - 136736) from the Pbp MU configuration are neither considered. This decision is based on the displacement of the average y coordinate of the primary vertex (PV) distribution. These decisions were taken following previous analyses using these same data samples [70].

[1] Trigger Configuration Key

4.2 Proton-proton data

The pp data correspond to the November 2015 data taking at $\sqrt{s} = 5\,\text{TeV}$ with a 25 ns bunch crossing spacing. The used data correspond to four different fills: 4638, 4639, 4640, and 4643. During this data-taking period, the magnet polarity was not inverted, so data were only acquired with the MD configuration.

For this data sample, only the `0x0115014e` TCK trigger configuration was enabled during the data-taking. This configuration selects no bias events from collisions of leading bunches (bunch preceded by an empty bunch) using the `Hlt1NoBiasLeadingCrossing` and the `Hlt2NoBiasLeadingCrossing` trigger selections. As opposed to the trigger line used in proton-lead events, no requirements on the event topology are made. Because of the trigger selection applied, events from leading bunch crossings are free from spillover.[2] During this data taking the average value of *visible* interactions per bunch crossing was $\mu = 0.6$.

The reconstruction version for these data is Reco15a. In this reconstruction, a baseline selection for the tracks is applied. Specifically, a cut in ghost probability of $\texttt{GhostP} < 0.4$ and track quality selection of $\chi^2/\text{ndf} < 4$.

4.3 Simulation samples

Together with the data collected by the LHCb detector, Monte Carlo (MC) simulation samples for pPb, Pbp and pp events are used in this analysis. In Tab. 4.3 a summary of the considered simulation samples is presented. The MC samples in this work are generated with minimum bias configuration, which corresponds to event-type = 30000000. In general Pythia minimum-bias is used to simulate pp collisions and EPOS-LHC minimum-bias for proton-nucleus collisions.

A private sample of EPOS-LHC for pp collisions was produced in addition to the Pythia sample. This was done to correct the proton-to-pion and kaon-to-pion fractions in pp collisions. The particle production mechanism is different in both event generators and, as a consequence, the proportions of the different particle species are also different. As explained in Chap. 6, these fractions are not properly modelled in Pythia, resulting in wrong estimations of the charged particles efficiencies. To correct for this, the proportions obtained from EPOS-LHC are used instead which better reproduces previous measurements of these fractions. Distributions of some variables are studied in Sec. 5.3.3 to check the compatibility between data and simulation.

Regarding the proton-nucleus events, data sets were produced with the official LHCb tune of EPOS-LHC [71]. The amount of MU (MD) simulated pPb events is 5208647 (5226592), whereas for Pbp events it is 5165389 (5166396). These are minimum-bias single interaction events.

[2] The spillover corresponds to contamination between two neighbour colliding bunches.

Generator	**Sim version**	**Beams**	**Magnet**	#evts	**Pile-up**
EPOS-LHC	Sim09e	pPb, Pbp	MD, MU	$5 \cdot 10^{6*}$	$1\,\text{int/evt}$
EPOS-LHC	Sim09k	pPb, Pbp	MD, MU	$7 \cdot 10^{6*}$	$3,4\,\text{int/evt}$
Pythia	Sim09d	pp	MD	$1 \cdot 10^{7}$	$\nu = 1.5$
EPOS-LHC	private	pp	MD	$1 \cdot 10^{7}$	$1\,\text{int/evt}$

Table 4.3: Summary of the different simulation data sets available for the analysis.
* Number of requested events per beam configuration and magnet polarity.

4.3.1 Enhanced multiplicity sample

Charged particle multiplicity is an important variable in this analysis and must be carefully studied. Since multiplicity is underestimated in simulation, the agreement with data can be achieved by weighting simulation events. However, for the pPb and Pbp simulation samples produced with EPOS-LHC, the range of the occupancy variables used to estimate the event multiplicity ends at an upper limit below the one observed in the data. This characteristic makes event weighting not viable or produces large uncertainties in the high multiplicity region. To enable the study of high multiplicity events, modified EPOS-LHC simulation samples are generated for pPb (Pbp) configurations, where three (four) generated events are produced per bunch crossing. The event type of this production is `21263010` for D^+ and `23263020` for D^s. The production points of these four events are placed at the same location, see Fig. 4.1, so that only one PV is reconstructed.[3]

[3]These event samples were created for open-charm production studies and contain either a D_s or a D^+ per event, see [72]. Since this analysis is focused on minimum bias events, the open charm decay products are excluded in efficiency and purity determinations.

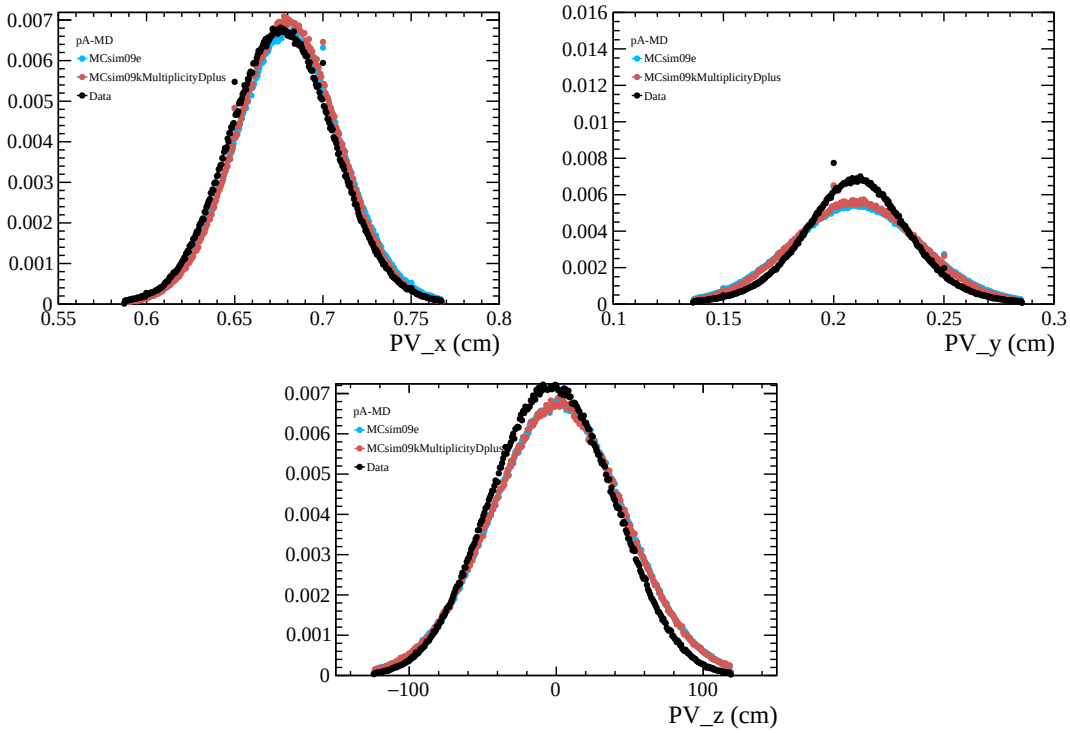


Figure 4.1: Normalised primary vertex position distribution for data and both simulation samples. Blue dots correspond to the standard LHCb simulation and red to the enhanced multiplicity sample.

5

Preparation and selection of the data and simulation

In this section, the event and track selection criteria are explained. The purpose of the selection cuts is to minimise the influence of background contributions while keeping most of the signal.

5.1 Event Selection

At the LHCb experiment, events correspond to particle-bunch crossings determined by the LHC internal clock. Occasionally, not only bunch-bunch events occur but also events with empty bunches. Accordingly, three possibilities arise; empty-empty, bunch-empty, and bunch-bunch events, where bunch indicates that the particle-bunch is populated (non-empty). This analysis only studies minimum-bias bunch-bunch events (those corresponding to the `BunchCrossingType==3` condition in the data samples).

Background collisions: Aside from bunch-bunch collisions, other interactions may occur within the detector:

- *Beam-gas* interactions. It is produced when a beam particle hits a gas molecule nuclei within the VELO vacuum.
- *Beam-splash* interactions. They occur when a beam particle interacts with the detector material.

- *Interactions between a nominal and a previous bunch* (spill-over). These events are mitigated in proton-lead data as the time between bunch crossings is larger than for the *pp* configuration. In *pp* events this background is not present, since only leading bunch crossings are considered.

Some of these processes can be misinterpreted by the trigger resulting in an accepted *background event*. If they occur within the VELO acceptance, a legitimate single primary vertex can be reconstructed, in the same manner as for single interaction collisions. Still, the observed charged particle multiplicity would be contaminated with tracks originating in the parasitic interactions. In addition, *pp* or proton-nucleus collisions can also pile up to one of these spurious interactions.

The interactions from background collisions share two key aspects:

- *Energy:* The energy of this interaction is less than in bunch-bunch collisions.

- *Interaction point:* The average interaction point is displaced from the typical location for non-background events.

This last feature will be exploited to reduce the impact of background collisions in the signal using two strategies: a selection on the number of reconstructed PVs and a selection on the position of these reconstructed PVs.

Background reduction strategies for events:

- Primary vertex condition:

  Data-taking conditions were very different for *pp* and *p*Pb/Pb*p* collisions. In the proton-nucleus collisions, the instantaneous luminosity was significantly lower and, therefore, no more than one interaction per bunch crossing is expected, $\mu \sim 0.04$. However, the multiplicity per interaction on average is higher than for *pp* events. As the requirement for having a reconstructed PV is the existence of at least four converging VELO tracks, this is more likely to be achieved in proton-nucleus due to the higher multiplicity per collision resulting in a higher PV reconstruction efficiency, as seen in Fig. 5.1. In this plot, ε_{PV} is defined as the ratio between the number of reconstructed PVs in simulation over the number of generated PVs in simulation. Moreover, it can occur that some PVs are reconstructed while not being actual primary vertex interactions, and this situation is more likely to happen in proton-nucleus collisions. This motivates the analysis to consider only one PV per event.

  In *pp* events, the instantaneous luminosity was higher. Thus, the probability of having more than one primary vertex per event is much higher than in proton-nucleus, $\mu = 0.6$, as was explained in Chap. 4. However, only single-PV events are also considered for *pp* data because the charged particle multiplicity has to be related to a single interaction to be physically meaningful.

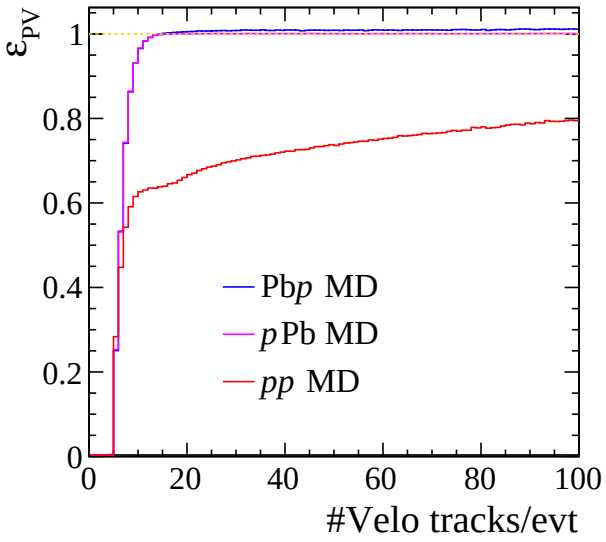


Figure 5.1: Primary vertex reconstruction efficiency as a function of the number of Velo tracks in the event for pPb, Pbp and pp collisions extracted from Ref. [70].

- Luminous region: As the interaction point of background collisions is usually displaced from the distribution for bunch-bunch collisions, a selection requirement in the reconstructed primary vertex location can be used to reduce the background. The strategy to reduce this background is to require the reconstructed primary vertex (PV) exclusively within a fiducial *luminous region*.

  In order to define the *luminous region*, the distribution of each of the PV coordinates, (x, y, z), are fit to a Gaussian function. In the LHCb convention the z coordinate is parallel to the beam pipe and x and y are the transverse coordinates. The *luminous region* is defined as:

$$|x_{\mathrm{PV}} - \mu_x| < 3\sigma_x, \quad |y_{\mathrm{PV}} - \mu_y| < 3\sigma_y, \quad |z_{\mathrm{PV}} - \mu_z| < 3\sigma_z, \tag{5.1}$$

  where $\mu_{x,y,z}$ are the mean values and $\sigma_{x,y,z}$ the standard deviations for each of the Gaussian fits. The values for these parameters are detailed in Tab. 5.1 and the primary vertex positions are shown in Fig. 5.2.

  In Fig. 5.2 we can see how data PVs distributions are not well reproduced outside the luminous region by simulation samples in pPb, Pbp collisions. This mismatch does not appear in pp collisions where data PVs distribution is well modelled by simulation. Moreover, the number of events outside the luminous region in pp data is much smaller. For that reason, we do not apply any cut on the PV position for pp events.

Moreover, the PV reconstruction efficiency decreases for low numbers of reconstructed tracks in the VELO, see Fig. 5.3. This means that we cannot go arbitrarily towards low

	pPb					
	MD			MU		
	x_{PV}	y_{PV}	z_{PV}	x_{PV}	y_{PV}	z_{PV}
μ[mm]	0.677	0.211	-2.7	0.698	0.208	3.83
σ[mm]	0.030	0.025	40.6	0.024	0.023	61.08

	Pbp					
	MD			MU		
	x_{PV}	y_{PV}	z_{PV}	x_{PV}	y_{PV}	z_{PV}
μ[mm]	0.694	0.192	-1.17	0.687	0.201	2.66
σ[mm]	0.025	0.024	38.5	0.025	0.024	60.40

Table 5.1: Results from the Gaussian fits used to define the luminous region.

Dataset	Data			Simulation-MCsim09k		
	BCT*	1 PV	Lumi. Region	BCT	1 PV	Lumi. Region
pPb MD	97.64%	81.92%	80.29%	100%	99.12%	97.27%
pPb MU				100%		
Pbp MD	96.42	79.80%	78.03%	100%	87.83%	86.71%
Pbp MU				100%		

Table 5.2: Fraction of events passing the event selection cuts for each configuration. Percentages are cumulative concerning the previous cut. *BunchCrossingType == 3.

multiplicities without considering the PV reconstruction efficiency. According to Ref. [73], for pp collisions at $\sqrt{s_{NN}} = 7$ TeV, the PV reconstruction efficiency for isolated PVs is $\sim 90\%$ when 10 reconstructed tracks are assigned to the PV. This will correspond to our minimum multiplicity value. Following simulation we can correlate `nVeloTracks` to N_{ch}, see Fig. 5.3, and define a minimum N_{ch} value for our measurement. For this analysis, the minimum considered multiplicity will be $N_{\mathrm{ch}} = 5$ and PV reconstruction efficiency is not computed. For pPb and Pbp, there are no PV reconstruction studies, so we will apply the same cut as for pp. The full event selection criteria are summarised in Tab. 5.3.

	pp	pPb	Pbp
Trigger	`Hlt1NoBiasLeadingCrossing`	`Hlt1MBMicroBiasVelo`	`Hlt1MBMicroBiasVelo`
HLT2	`Hlt2NoBiasLeadingCrossing`	None	None
Bunch crossing type	bunch-bunch	bunch-bunch	bunch-bunch
number of PV	1	1	1
PV position	-	within luminous region	within luminous region

Table 5.3: Summary of the event selection considered for the analysis.

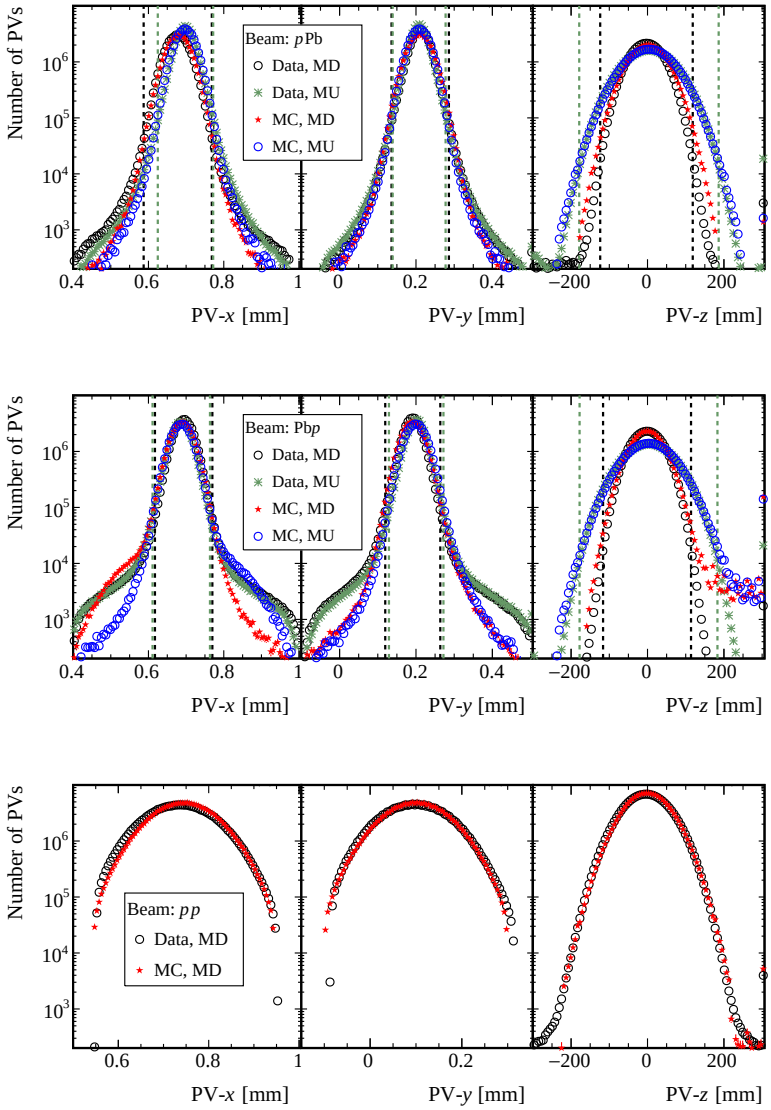


Figure 5.2: Reconstructed PV distribution and luminous region definition extracted from Ref. [70]. More details about this can be found in the same reference.

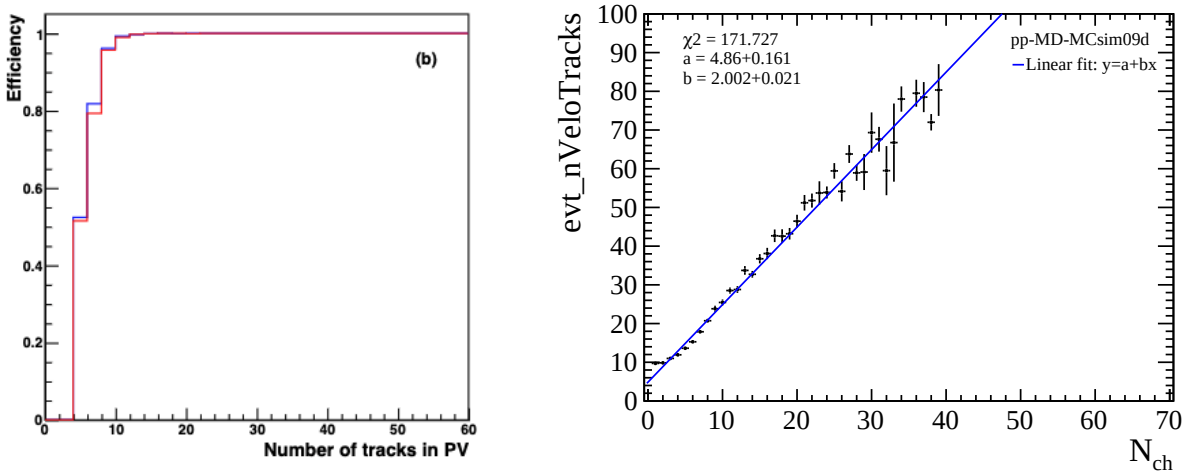


Figure 5.3: PV reconstruction efficiency (left) as a function of the number of VELO tracks assigned to the PV from Ref. [73]. Relation between N_{ch} and `nVeloTracks` from Pythia (right). These events have all event selection cuts in Tab. 5.3, except $nPVs == 1$.

5.2 Candidate selection

The candidates for prompt charged particles in this analysis are reconstructed *long tracks* from the `TES["Rec/Track/Best"]` container of the `.dst` files. According to Ref. [74], these tracks go through all the tracking stations in the experiment, as depicted in Fig. 5.4. These tracks have hits both in the VELO and the T stations, and optionally in the TT stations. As they have hits before and after the magnet, they provide the most precise momentum estimate and, therefore, are adequate tracks for physics analyses. However, there are still several background sources that need to be taken into account when considering long tracks. In this section, those different background sources will be detailed, as well as the strategies used to reduce their contribution.

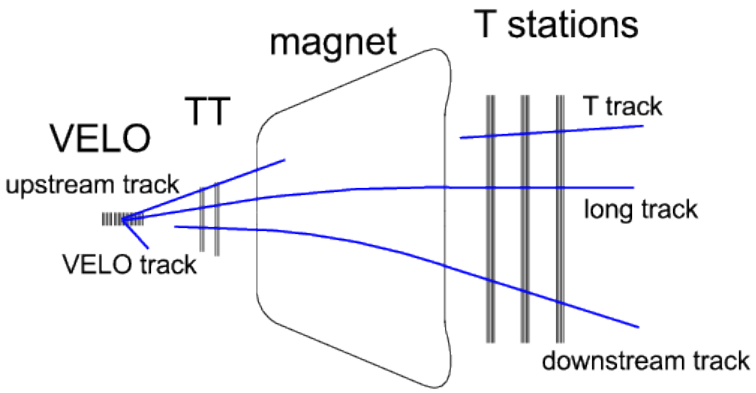


Figure 5.4: Different kinds of tracks that can be reconstructed in the LHCb experiment according to Ref. [75].

5.2.1 Candidates kinematic region

Since long tracks are reconstructed using hits from detectors after the magnet, only tracks with $p > 2$ GeV/c are considered for the analysis because being the ones passing through the magnet material. The LHCb tracking system reconstructs tracks within 10 mrad to 300 mrad in the bending plane, which corresponds to $1.9 < \eta < 5.3$. However, due to the limited acceptance of the TT station, the effective angular coverage is limited to $2 < \eta < 4.8$.

The lower limit in the p_{T} range depends on (p, η), and spans from $p_{\mathrm{T}} = 500$ MeV/c at $\eta = 2$ to $p_{\mathrm{T}} = 20$ MeV/c at $\eta = 4.8$. This will be detailed when discussing the analysis binning in Chap. 6. Since there is an increase in the background candidates and a reduction of the reconstruction efficiency at low p_{T}, a global lower limit of $p_{\mathrm{T}} > 500$ MeV/c is imposed. Moreover, it is the minimum p_{T} value common to all η bins as can be seen in Fig. 5.5. An upper limit is set to $p_{\mathrm{T}} < 8$ GeV/c, motivated by the abundance of

reconstruction artefacts and the limitations from the size of the simulated samples beyond this value.

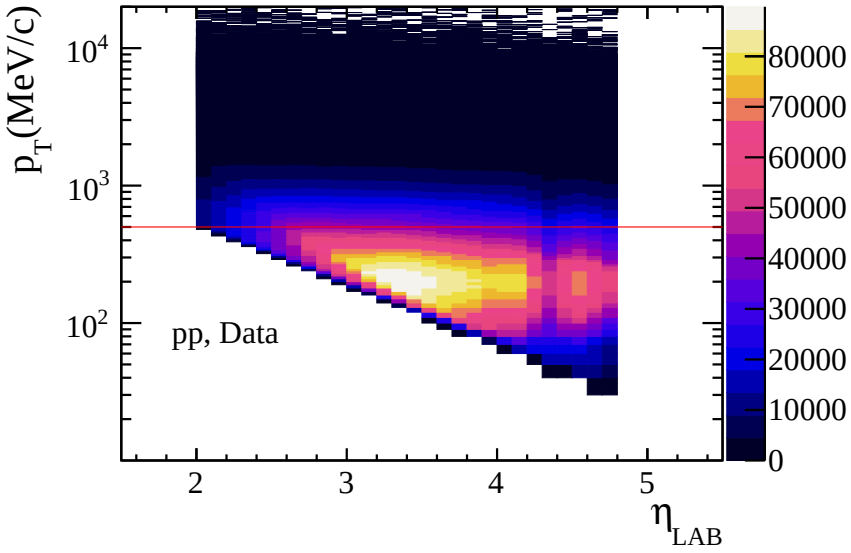


Figure 5.5: Candidates η, p_{T} distribution. The acceptance, $2.0 < \eta_{LAB} < 4.8$ and momentum cuts are applied, $p > 2$ GeV. The horizontal line shows the minimum p_{T} cut applied in this analysis.

5.3 Background candidates

The background is constituted of candidates that do not correspond to a prompt charged particle. Three background sources are identified: fake tracks, clone tracks and secondary particles. This section will be dedicated to discussing each background component and explaining the strategy to reduce and identify its impact on the analysis. There are three background categories.

- **Secondary particles** are particles which do not satisfy the prompt definition of Sec. 6.1. They can be classified into two classes. The first class is composed of particles from decays of long-lived ($\tau > 30$ ps) particles:
  - K_s^0 mesons.
  - Λ baryons.
  - Hyperons: Σ^+, Σ^-, Ξ^0, Ξ^-, Ω^-.
  - Pions and kaons decaying into leptons, especially muons.

Hadronic decays to muons will be treated separately from the other hadronic decays because of having a different final state.

The second class are particles not coming from the primary interaction or subsequent particle decay. Typically, these particles are produced in other interactions such as:

- Electrons from photon conversions (also called γ conversion), $\gamma \to e$.
- Showers from hadronic interactions with the detector material.

The category of particles in simulation is identified with the **IsPrompt** algorithm, which iteratively explores the decay chain backwards to the primordial collision. The algorithm checks three elements:

- If the particle mean lifetime is greater than $\tau > 30$ ps.
- If the particle has no ancestors with a mean lifetime greater than $\tau > 30$ ps.
- If the particle or any of its ancestors do not originate in an interaction with the detector material.

When these three conditions are satisfied a charged particle is classified as prompt. The study of the secondary particles contribution to the signal is done with simulation identifying tracks matched to a non-prompt charged particle. In order to suppress the secondary particle contribution in our candidate sample a dedicated selection in the impact parameter track variable (tr_IP) is used. The impact parameter (IP) of a track is defined as the distance between the track and the PV at the track point of closest approach to the PV, see Ref. [76]. The tracks coming from particles produced in the decay of long-lived particles or produced as a consequence of interactions with the detector material tend to have a larger IP than those from particles produced at the PV. Consequently, cuts on the IP are very effective at excluding non-prompt backgrounds, and maximising the signal content of our candidate sample.

- **Clone tracks** are two tracks produced by the same particle.

  Clone tracks are reconstruction artefacts. As defined in Ref [77], if two tracks share more than 70% of the hits in the VELO and at least 70% of the hits in the T-stations seeding region they will be considered clones (`IsClone` $== True$). Note that this LHCb definition does not take into consideration hits in the TT stations. A selection cut is already applied at the reconstruction level using the Kullback-Leibler (KL) clone distance, which measures the amount of information shared between two reconstructed tracks [78]. The default cut is $KL < 5000$, which significantly reduces this background contribution in the data sample. However, it is still necessary to estimate the clone tracks contribution to the sample after the cut. This is done with simulation by counting the number of tracks matched to the same particle. The track with the highest χ^2/nDoF is considered a clone.

- **Fake tracks.** When a track does not correspond to a particle trajectory, it is considered a *fake track*. As for clone tracks, those are reconstruction artefacts and can appear at different stages of the LHCb experiment data processing. Most of the fake long tracks arise when a particle cannot be tracked through the full detector because it interacts with the material. Other fake long track sources are fake track segments in the T-stations or tracks reconstructed from two different particles in the upstream and downstream parts of the detector.

  To reduce the number of fake tracks, a multivariate classifier called *ghost probability* (`GhostP`) is commonly used in the LHCb experiment, see Ref. [79]. This variable relies on a neural network that considers 22 variables from all tracking detectors to assign a fake probability to each track. In the pp data sample, there is a `GhostP` cut at the reconstruction level, `GhostP` < 0.4, which is not present in the pPb and Pbp data samples. In this analysis, the `GhostP` cut listed in Tab. 5.5 follows the value optimised by other studies using the same data sample, see Ref. [70].

  In Sec. 6.7.2, the remaining fraction of fake tracks in the candidate sample will be estimated from simulation and taken into account as a function of (η, p_{T}, `nVeloTracks`). In simulation, tracks can be identified by matching them with a particle. The matching condition is fulfilled when a track shares at least 70% of its hits with the particle. As a consequence of that, if a track is not matched to a particle it is classified as fake.

5.3.1 Background suppression strategies

As mentioned above, IP and `GhostP` will be used to reduce the contribution of fake tracks and secondary particles in the candidate sample. The selection in both variables is optimised using, as a figure-of-merit (FoM), the product of the significance and the purity of the candidate sample defined as,

$$\mathcal{SP} = \frac{S(g,p)}{\sqrt{S(g,p)+B(g,p)}} \cdot \frac{S(g,p)}{S(g,p)+B(g,p)}. \tag{5.2}$$

Other FoMs were tested in Ref. [70] concluding that $\mathcal{SP}$ is the best one for prompt charged analysis using these data sets. The goal of this optimization is to find an optimal pair of values (g,p) as selection cuts for `GhostP` and IP respectively, i.e., `GhostP` $< g$, $IP < p$. The heatmap of the FoM can be seen in Fig. 5.6 and the final selection values, g and p, in Tab. 5.4.

When two long tracks share a common VELO segment it is quite likely that only one of them represents a prompt charged particle, while the other is background. The background track can correspond to a fake, clone track or a secondary particle. The latter happens when a prompt neutral particle such as K^0_S or Λ decays after VELO. Afterwards,

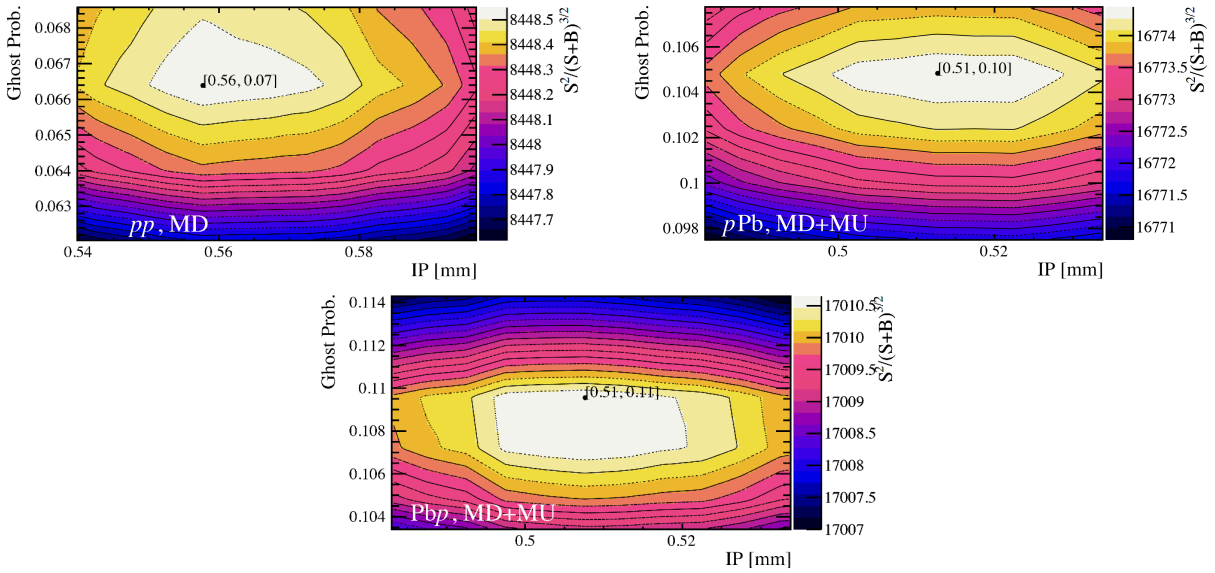


Figure 5.6: Optimization of the selection using $\mathcal{SP}$ figure of merit (FoM) for the different collision modes, pp, pPb, Pbp.

	Pbp	pPb	pp
`GhostP`	0.11	0.10	0.07
IP	0.51	0.51	0.56

Table 5.4: Summary of the optimised values for `GhostP` and IP using the $\mathcal{SP}$ FoM.

its decay products produce a track in the downstream trackers and this track is matched to the VELO segment of a different charged particle.

These background sources will be discussed in detail in Sec. 6.7.2. Making use of the `ITrackUniqueSegmentSelector` algorithm, pairs of long tracks sharing a VELO segment can be flagged. For these tracks the one with the highest χ^2/nDoF, [75], is classified as a *shared VELO* track for both data and simulation. In Ref. [70] the composition of tracks sharing a VELO segment was studied, concluding that they are dominated by fake tracks in $70\% - 80\%$ of the cases. Moreover, in the same reference, the fraction of these tracks in the candidate sample was found to be within $[0.06, 0.2]\%$. As a consequence, tracks sharing a VELO segment are removed from the candidate sample.

5.3.2 Selection summary

Only events with a single reconstructed primary vertex within the luminous region are considered. The optimised selection cuts for IP and `GhostP` indicated in Tab. 5.4 are applied to every long track. Moreover, long tracks flagged with the shared VELO tag are removed. In the following sections, the purity of the sample and selection efficiencies

will be detailed. In general, the purity is above 85% and the selection efficiency is above 80%. Finally, events with less than five prompt-charged particles are not considered in the result because of the limitations in the primary vertex reconstruction efficiency. The full set of selection cuts applied to long tracks and events to produce the final candidate sample is summarised in Tab. 5.5.

		pp (2015)	pPb (2013)	Pbp (2013)
Event selection	Trigger	`Hlt1NoBiasLeadingCrossing`	`Hlt1MBMicroBiasVelo`	`Hlt1MBMicroBiasVelo`
	Bunch crossing type	bunch-bunch	bunch-bunch	bunch-bunch
	number of PV	1	1	1
	PV position	-	within luminous region	within luminous region
Candidate selection	Track location	`TES["Rec/Track/Best"]`	`TES["Rec/Track/Best"]`	`TES["Rec/Track/Best"]`
	Track type	Long	Long	Long
	η	$2 < \eta < 4.8$	$2 < \eta < 4.8$	$2 < \eta < 4.8$
	p	$p > 2$ GeV/c	$p > 2$ GeV/c	$p > 2$ GeV/c
	p_{T}	$0.200 < p_{\mathrm{T}} < 8$ GeV/c	$0.200 < p_{\mathrm{T}} < 8$ GeV/c	$0.200 < p_{\mathrm{T}} < 8$ GeV/c
	IP	0.508 mm	0.508 mm	0.568 mm
	`GhostP`	0.078	0.103	0.109
	`IsClone`	`False`	`False`	`False`

Table 5.5: Summary of the selection considered for the analysis.

5.3.3 Simulation validation

The comparison between data and different simulation samples for the distribution of the tracks p_{T}, η, and p and the event `nVeloTracks` is shown in Figs. 5.7, 5.8, 5.9 and 5.10. This data/simulation comparison is made after applying the event and candidate selection explained in Sec. 5.2, Sec. 5.1. All distributions exhibit reasonable agreement, particularly with the enhanced multiplicity sample, discussed in Chap. 4, except for `nVeloTracks` which will be subsequently corrected as will be detailed in Sec. 5.3.4.

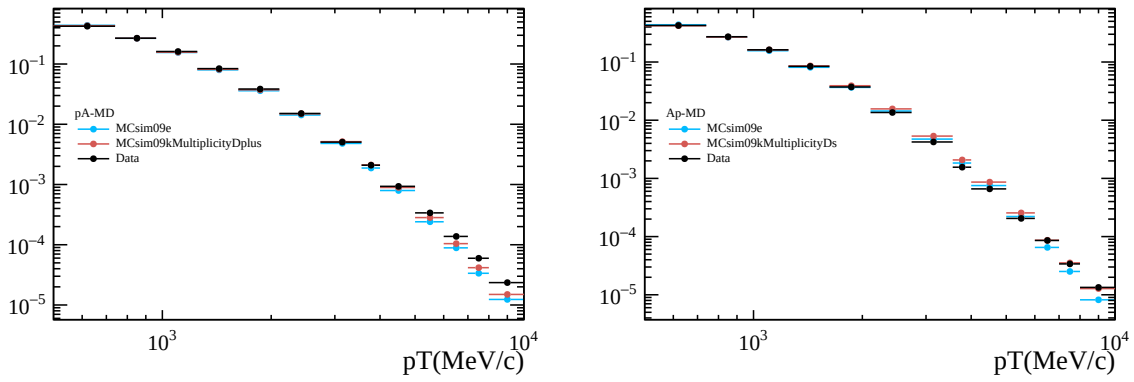


Figure 5.7: Distribution of the track transverse momentum for pPb (left) and Pbp (right) collisions.

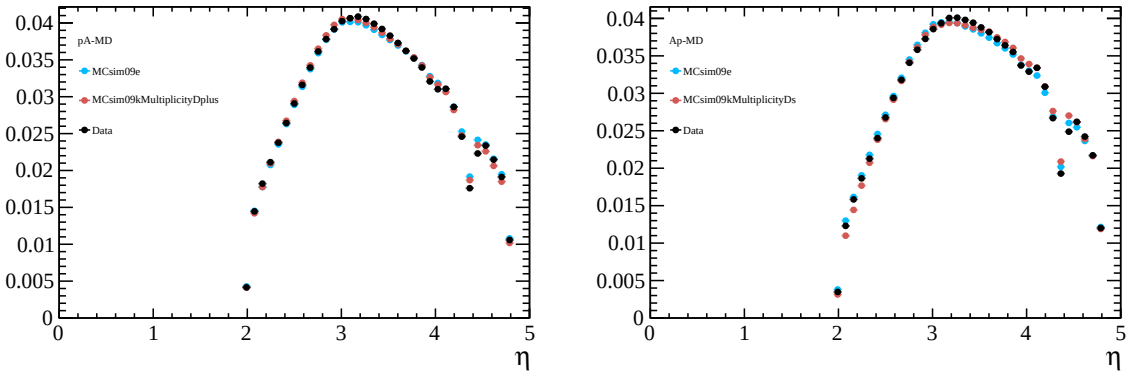


Figure 5.8: Distribution of the track pseudorapidity for pPb (left) and Pbp (right) collisions.

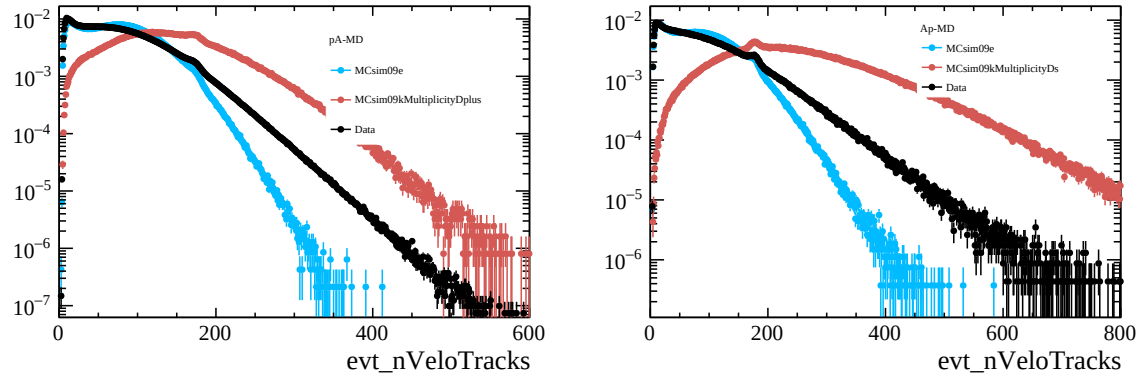


Figure 5.9: Distribution of the total number of `nVeloTracks` for pPb (left) and Pbp (right) collisions.

5.3.4 Detector occupancy

The detector's behaviour depends on its occupancy. The occupancy can be estimated using different subdetectors (VELO, TT, IT, ECAL, etc.) and magnitudes (clusters, tracks, etc.). Certainly, these methods are correlated. The discrepancy in occupancy distributions between data and simulation has already been mentioned. Since a correct occupancy description is crucial to the analysis, a method is applied to weigh the simulation to match data occupancy variable distributions. The weighting is obtained with the *Gradient Boost Reweighting* algorithm described in Ref. [80]. This algorithm uses decision trees and a special loss function called *ReweightLossFunction* with positive and negative weights, enabling the simultaneous weighting of multiple variables and achieving high precision in describing target distributions. In this note, the implementation of the algorithm found in `hep_ml`[1] is used.

[1] https://pypi.org/project/hep-ml/

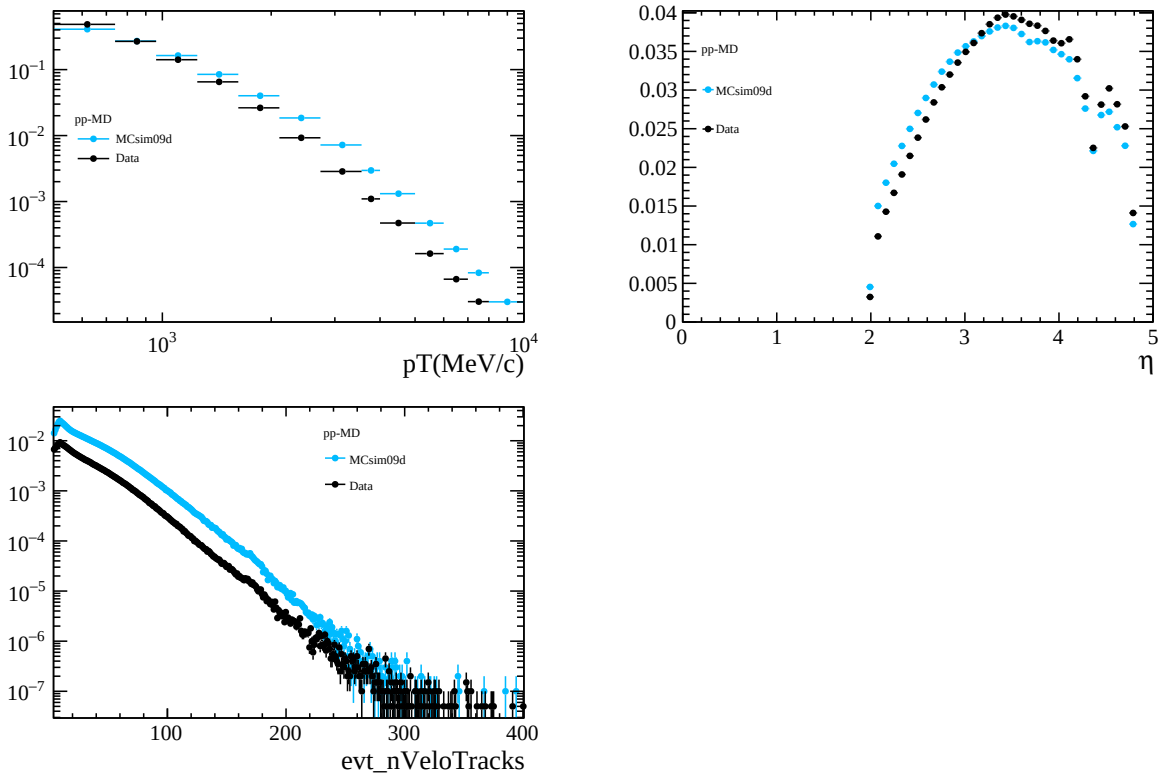


Figure 5.10: Distribution of p_{T}, η and `nVeloTracks` for pp collisions.

At LHCb, there are several variables related to particle multiplicity. Those are called *occupancy* variables and typically correspond to the number of hits, clusters or tracks in one LHCb subdetector. In previous studies Ref. [70] it was demonstrated that in LHCb charged particles studies using minimum-bias proton-nucleus data, the reweighting best performance is achieved considering the number of clusters from every tracking station as the occupancy variable. This means simultaneously computing event weights that account for VELO, TT, IT and OT clusters. In the case of pp collisions, weights are computed again using the same variables as in Ref. [70]. Different sets of variables were compared and the one containing the number of reconstructed tracks[2] and the number of reconstructed PVs was found to be the one that gave a better description of the data.

Some examples of the results from the reweighting algorithm and the output weights are displayed in Fig. 5.11 and Fig. 5.12. The complete set of results can be seen in App. B.2 for all simulation samples and beam collision systems. The data on those plots fulfil the event selection explained in Sec. 5.1 and therefore, the distributions are those of the final sample used in the analysis. Overall, the data show good agreement in the distribution of *occupancy* variables after applying the reweighting procedure.

[2] All reconstructed tracks from the LHCb reconstruction software.

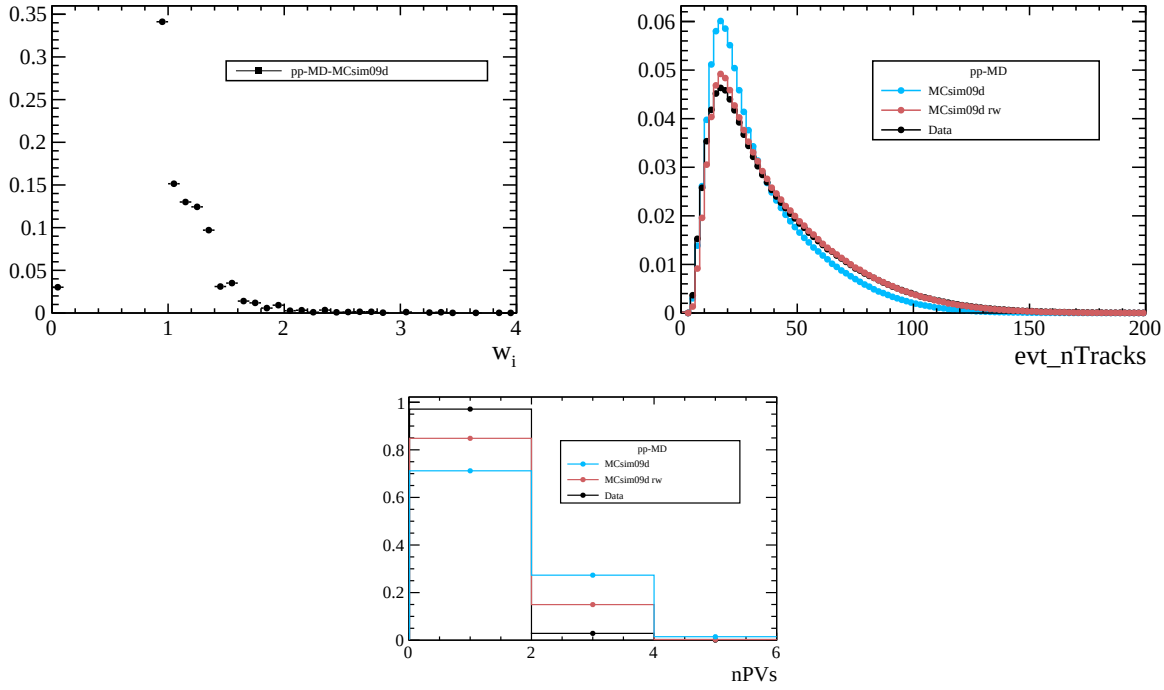


Figure 5.11: The distribution of occupancy weights is shown in the upper left. Reconstructed tracks (top right) and reconstructed number of PVs (bottom) distributions are shown before and after applying the weights for *pp* collisions with the MCsim09d sample.

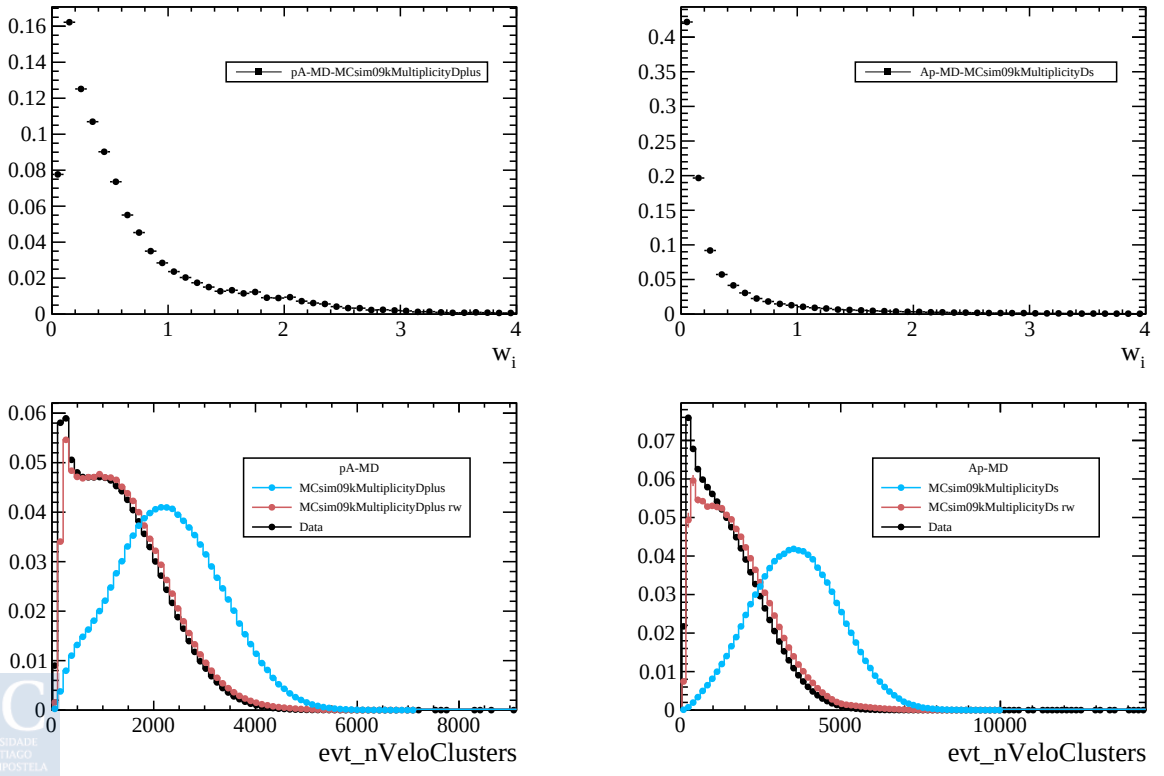


Figure 5.12: Distributions of the weights (top row) and nVeloClusters (bottom row) before and after applying the weights for *p*Pb (left column) and Pb*p* (right column) collisions.

6

Multiplicity distribution and the average transverse momentum

6.1 Observables

The signal in this work are *prompt charged particles*. They are defined by the LHC Minimum Bias and Underlying Event Working Group [81] as:

Hadrons and leptons, with mean lifetime $\tau > 0.3 \cdot 10^{-10}$ s, *produced directly or from decays of shorter-lifetime particles.*

The particles included in this definition are listed in Tab. 6.1.

Three definitions are related to the number of tracks and/or particles relevant to this work.

- $N_{\text{cand}}^{sim/data} \rightarrow$ The number of long **tracks** after the selection explained in Chap. 5.
- $N_{\text{corr}}^{sim/data} \rightarrow$ The result of correcting N_{cand} with efficiencies and purity.
- $N_{\text{ch}}^{sim} \rightarrow$ The number of prompt charged **particles** (at generator level when referring to simulation).

The variables N_{cand} and N_{corr} exist in data (*data*) and simulation (*sim*), while N_{ch} is only accessible from simulation. One of the tasks in this analysis is to measure N_{ch} in data. For that, N_{corr} will be used as a proxy.

Species	**Mean proper lifetime**	
	τ**(s)**	$c\tau$**(cm)**
$\pi^{\pm}$	$2.60 \cdot 10^{-8}$	780
$K^{\pm}$	$1.24 \cdot 10^{-8}$	371
$p^{\pm}$	∞	∞
$e^{\pm}$	∞	∞
$\mu^{\pm}$	$2.19 \cdot 10^{-6}$	65800
$\Xi^{\pm}$	$1.64 \cdot 10^{-10}$	4.91
Σ^{+}	$0.80 \cdot 10^{-10}$	2.40
Σ^{-}	$1.48 \cdot 10^{-10}$	4.434
Ω^{-}	$0.821 \cdot 10^{-10}$	2.461

Table 6.1: Particle species included in the prompt charged particle definition.

This work aims to measure the multiplicity dependence of the average transverse momentum, $\langle p_{\mathrm{T}} \rangle$, of prompt charged particles, defined as

$$\langle p_T \rangle (N_{\mathrm{ch}}) = \frac{\sum_{p_{\mathrm{T}}} \frac{dN_{\mathrm{ch}}}{dp_{\mathrm{T}} d\eta} \cdot p_{\mathrm{T}} \cdot \Delta p_{\mathrm{T}}}{\sum_{p_{\mathrm{T}}} \frac{dN_{\mathrm{ch}}}{dp_{\mathrm{T}} d\eta} \cdot \Delta p_{\mathrm{T}}}. \tag{6.1}$$

Moreover, the multiplicity spectrum, $P(N_{\mathrm{ch}})$, is also measured using Eq. 6.2. Where N_{evt} is the total number of events.

$$P(N_{\mathrm{ch}}) = \frac{1}{N_{evt}^{T}} \frac{dN_{evt}}{dN_{\mathrm{ch}}}. \tag{6.2}$$

As N_{ch} is not directly accessible we will use N_{corr} as a proxy. Then, the bias produced by using a proxy is evaluated and corrected for each observable. In the case of the multiplicity spectra, an unfolding strategy is used to solve this issue.

The first quantity that we need to compute is N_{cand}, which is determined for each event by applying the selection cuts to the long-track sample. Secondly, corrections are applied to account for detector effects like miss-reconstruction or background in the selected sample. These corrections are the reconstruction efficiency, ε_{reco}, the selection efficiency, ε_{sel}, and the purity, P. These factors are computed in $(\eta, p_{\mathrm{T}}, \texttt{nVeloTracks})$ bins, where `nVeloTracks` is the number of tracks in the VELO (VErtex LOcator) for each event. After that, the number of corrected tracks for each event in data and simulation, N_{corr}, is computed using Eq. 6.3, together with the candidates p_{T} spectra in N_{corr} classes.

Moreover, using truth-level information (N_{ch}) and reconstruction level (N_{corr}) from the simulation, we build the response matrix (R). In the case of the $P(N_{\mathrm{ch}})$ measurement, the response matrix is used within a Bayesian Unfolding algorithm to correct the $P(N_{\mathrm{corr}})$ spectrum for bin migration and finite bin width on the corrections. This procedure is validated by performing a closure test as explained in Sec. 6.8. For the $\langle p_{\mathrm{T}} \rangle$ results, the measurement is done as a function of N_{corr} and then a closure test is done to estimate the

bias due to the difference between $N_{\rm corr}$ and $N_{\rm ch}$,

$$N_{\rm corr} = \frac{N_{\rm cand} \cdot \Pi}{\varepsilon_{reco} \cdot \varepsilon_{sel}}. \tag{6.3}$$

Pseudo-rapidity dependent measurements in multiplicity analysis

Recent studies of J/ψ production as a function of multiplicity suggest the possibility of having an auto-correlation bias arising from the multiplicity selection [82, 83]. This means that the results can be biased if both the signal and the multiplicity proxy are measured in the same acceptance. For example, if high multiplicity events are selected in a pseudorapidity range, a preference for events with jets or B-hadron decays within this region can arise. As a consequence, $\langle p_{\rm T} \rangle$ can be higher just because the sample is biased to those events.

To estimate the contribution of this effect the approach is to measure the observable but select the event multiplicity in different acceptance regions. The latter will be used in this note because it is feasible to conduct prompt charged particle analysis. For this reason, $\langle p_{\rm T} \rangle$ will be studied as a function of the multiplicity defined in different regions. As the LHCb detector is fully instrumented in the $\eta \in [2.0, 4.8]$ range, it is possible to make a multiplicity measurement considering different η sub-intervals within the full coverage.

Concerning the η range where $\langle p_{\rm T} \rangle$ and $N_{\rm ch}$ are defined, different combinations can be proposed:

1. $\langle p_{\rm T} \rangle$ is measured in a **sub-η** region where $N_{\rm ch}$ is measured.
   For example, calculating $\langle p_{\rm T} \rangle$ in $\eta \in [2.0, 2.5]$ as a function of $N_{\rm ch}$ within $\eta \in [2.0, 4.8]$.

2. $\langle p_{\rm T} \rangle$ is measured in the **same** η region where $N_{\rm ch}$ is measured.
   For example, calculating $\langle p_{\rm T} \rangle$ in $\eta \in [2.0, 2.5]$ as a function of $N_{\rm ch}$ within $\eta \in [2.0, 2.5]$.

3. $\langle p_{\rm T} \rangle$ is measured in a **different** η region where $N_{\rm ch}$ is measured.
   For example, calculating $\langle p_{\rm T} \rangle$ in $\eta \in [2.0, 2.5]$ as a function of $N_{\rm ch}$ within $\eta \in [3.5, 4.8]$.

The selection and preparation of the data sample used to obtain selected candidates, $N_{\rm cand}$, will be explained in Chap. 5. In Chap. 6 the corrections applied to $N_{\rm cand}$ to obtain $N_{\rm corr}$ are discussed. Those corrections are first estimated from simulation before applying data-driven correction factors to account for discrepancies with data. In Sec. 6.8 the unfolding and its validation are explained, as the last step of data processing. Moreover, in the same section, the computation of $\langle p_{\rm T} \rangle$ and its closure test is done. A discussion of systematic uncertainties is included in Chap. 7. Finally, the analysis results are shown in Chap. 8.

This section shows how the observables are computed. To obtain the final measurements three correction factors are applied to the signal yield: the reconstruction efficiency (ε_{reco}), the selection efficiency (ε_{sel}) and the purity (P). The strategy to compute these factors starts with a determination from simulation, followed by data-driven corrections that improve the agreement between data and simulation. All three corrections depend on (η, p_{T}, `nVeloTracks`).

6.2 Computing the multiplicity distribution

When measuring the multiplicity distribution, $P(N_{\mathrm{ch}})$, several detector effects come into play. Some of them, such as reconstruction efficiencies or background in the signal sample are accounted for computing efficiencies and purity and they are applied using Eq. 6.4. Moreover, finite bin size effects and bin-to-bin migration in the determination of efficiencies can be corrected with Unfolding methods as explained in Ref. [84]. In this section, the procedure to get $P(N_{\mathrm{ch}})$ will be detailed.

The $P(N_{\mathrm{ch}})$ measurement is done in three steps. First, the candidate selection detailed in Chap. 5 is applied for each event and the distribution of the number of candidates is named as $P(N_{\mathrm{cand}})$.

Secondly, the number of candidates is corrected by efficiencies and purity,

$$N_{\mathrm{corr}} = N_{\mathrm{cand}} \frac{\Pi(evt)}{\varepsilon_{reco}(evt) \cdot \varepsilon_{sel}(evt)}. \tag{6.4}$$

where the per-event efficiencies and purity in Eq. 6.4 are computed as

$$C(evt) = \frac{\sum N_{tracks} C^{track}(p_{\mathrm{T}}, \eta)}{\sum N_{tracks}}; \qquad \text{where } C = \{\varepsilon_{reco}, \varepsilon_{sel}, \Pi\}. \tag{6.5}$$

As an illustrative example, the $\varepsilon_{reco}(evt)$ for pp collisions is shown in Fig. 6.1. After this step, we obtain the N_{corr} distribution called $P(N_{\mathrm{corr}})$.

The third and last step in the $P(N_{\mathrm{ch}})$ determination is the Bayesian Unfolding, which is applied to $P(N_{\mathrm{corr}})$ in order to get the final result $P(N_{\mathrm{ch}})$. The unfolding procedure needs a response matrix (R) which relates *detector-level* information (N_{corr}) with *generator-level* information (N_{gen}). We use the ROOT module RooUnfold [85] to build the response matrix and perform the Unfolding. Following a previous analysis on the same quantity [86], the Bayesian Unfolding algorithm implementing the D'Agostini framework [87] is used. The response matrix is filled event by event with (N_{corr}, N_{ch}) pairs from simulation, the response matrices and the corresponding closure-test will be discussed in Sec. 6.8.

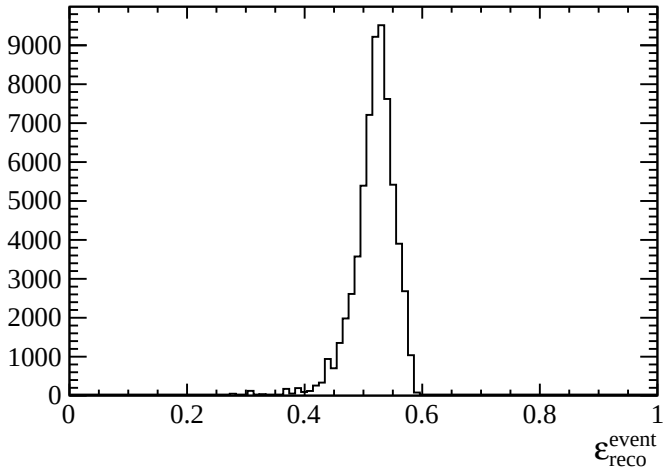


Figure 6.1: Example of reconstruction efficiency distribution per event for pp collisions.

6.3 Computing the average transverse momentum

To compute $\langle p_T \rangle (N_{\text{corr}})$, the charged particle transverse momentum spectrum as a function of p_{T} for corrected candidates,

$$\frac{1}{N_{evt}} \frac{dN_{\text{corr}}}{dp_{\text{T}}} = \frac{1}{N_{evt}} \frac{dN_{\text{cand}}}{dp_{\text{T}}} \frac{\Pi}{\varepsilon_{reco} \cdot \varepsilon_{sel}}, \tag{6.6}$$

is the main ingredient. This quantity is computed using the candidate transverse momentum spectra and applying the already mentioned corrections. From the corrected p_{T} spectrum, the average transverse momentum as a function of corrected candidates, $\langle p_T \rangle (N_{\text{corr}})$, can be computed as

$$\langle p_T \rangle (N_{\text{corr}}) = \frac{\sum_{p_{\text{T}}} \frac{dN_{corr}}{dp_{\text{T}}} \cdot p_{\text{T}} \cdot \Delta p_{\text{T}}}{\sum_{p_{\text{T}}} \frac{dN_{corr}}{dp_{\text{T}}} \cdot \Delta p_{\text{T}}}. \tag{6.7}$$

where Eq. 6.7 is a variation of Eq. 6.1 but selecting the event using N_{corr} instead of N_{ch}.

6.4 Binning scheme

The corrections that will be discussed in this sections are performed considering a $(\eta, p_{\text{T}}, \texttt{nVeloTracks})$ binning. This section will explain in detail the binning proposed for this analysis. As we are binning our simulation sample in several variables, $(\eta, p_{\text{T}}, \texttt{nVeloTracks})$, the number of bins is very high. Thus, the criteria used to define the number of bins is based on keeping the maximum number of bins allowed by the

statistics in the simulation sample.

Transverse momentum

The p_{T} binning scheme considers the exponential decrease of the charged particle multiplicity with this variable. The proposed binning includes 10 bins equally spaced on a logarithmic scale except for the last bin, which is slightly bigger than the others to increase the sample size that it contains,

$$p_{\mathrm{T}} = [0.500, 0.740, 0.961, 1.249, 1.622, 2.107, 2.737, 3.556, 4.619, 6.0, 8.0] \text{ GeV/c}. \quad (6.8)$$

The minimum p_{T} value is proposed due to the LHCb acceptance shown in Fig. 5.5, and the maximum value is proposed because of statistical limitations.

Pseudorapidity

For the pseudorapidity binning, the rapidity boost in proton-nucleus collisions needs to be considered. In the analysed data, the energies per nucleon of the lead beam and the proton beam are 1.58 TeV and 4.00 TeV, respectively, corresponding to a centre-of-mass energy of $\sqrt{s_{\mathrm{NN}}} = 5.02\,\mathrm{TeV}$. The beam energy and the proton and lead nucleus mass asymmetries result in a centre-of-mass boost of $y_{boost} \approx 0.465$ in the direction of the proton beam. Due to this boost, the kinematic acceptance of the LHCb detector in the nucleon-nucleon centre-of-mass system differs in the forward and backward configurations. In order to have a scheme that provides common bins for pPb and Pbp events, allowing to compute forward/backward ratios, pseudorapidity bin widths of $\Delta\eta = 0.5$ are considered. The binning in the laboratory and center-of-mass reference frame for all collision modes are shown in Tab. 6.2.

$\lvert\eta_{cms}\rvert$	$\eta\,(pp)$	$\eta\,(p\mathrm{Pb})$	$\lvert\eta\,(\mathrm{Pb}p)\rvert$
$[1.6, 2.0]$	–	$[2.065, 2.465]$	–
$[2.0, 2.5]$	$[2.0, 2.5]$	$[2.465, 2.965]$	–
$[2.5, 3.0]$	$[2.5, 3.0]$	$[2.965, 3.465]$	$[2.035, 2.535]$
$[3.0, 3.5]$	$[3.0, 3.5]$	$[3.465, 3.965]$	$[2.535, 3.035]$
$[3.5, 4.0]$	$[3.5, 4.0]$	$[3.965, 4.465]$	$[3.035, 3.535]$
$[4.0, 4.3]$	$[4.0, 4.3]$	$[4.465, 4.765]$	$[3.535, 3.835]$
$[4.0, 4.5]$	$[4.0, 4.5]$	–	$[3.535, 4.035]$
$[4.5, 4.8]$	$[4.5, 4.8]$	–	$[4.035, 4.335]$
$[4.8, 5.2]$	–	–	$[4.335, 4.735]$

Table 6.2: Correspondence between the $\lvert\eta_{cms}\rvert$ and η bins for the different considered configurations.

In this analysis, η will correspond to the laboratory frame. Moreover, the η binning used for all collision systems will be the one corresponding to pp in Tab. 6.2.

Multiplicity

In order to measure N_{corr} we need to compute multiplicity-dependent corrections using a detector-based variable. In this analysis `nVeloTracks` will be used to bin all the corrections, which will be later applied to the candidate sample. The number of bins in `nVeloTracks` is the maximum possible that keeps the statistical uncertainties under control. As the `nVeloTracks` distribution for each collision mode is different, the binning is also specific for each collision mode as detailed in Tab. 6.3.

pp	MCsim09d	10	20	30	50	70	90	∞						
pPb	MCsim09k	10	62	72	82	102	142	180	225	∞				
Pbp	MCsim09k	10	80	100	120	170	200	240	300	350	400	500	600	∞

Table 6.3: Binning in `nVeloTracks` used to compute the efficiencies and purity detailed. Each simulation sample has a different binning.

6.5 Reconstruction efficiency

The reconstruction efficiency (ε_{reco}) accounts for non-reconstructed tracks within the LHCb detector. Considering the generated prompt-charged particles definition of Sec. 6.1 (flag `IsPrompt` == `True`), the reconstruction efficiency is estimated using the occupancy weighted simulation sample as the ratio of reconstructed tracks matched to a generated prompt-charged particle over the total number of generated prompt-charged particles. This ratio is corrected with tracking calibration factors, $C_{tracking}$, which accounts for the differences in reconstruction between data and simulation; i.e. $C_{tracking} = \varepsilon^{data}_{tracking}/\varepsilon^{MC}_{tracking}$. These factors are computed by the LHCb tracking group[1] using $J/\psi \to \mu\mu$ decays and comparing the tracking performance in data and simulation. The tracking calibration factors can be seen in Fig. 6.2 and Fig. 6.3.

The reconstruction efficiency depends on the detector occupancy. However, the tracking calibration weights were not obtained as a function of the occupancy. That means that we are assuming that the multiplicity dependency of the reconstruction efficiency in data is well-modelled by simulation. Therefore, we consider the data/simulation discrepancy in the reconstruction efficiency to be independent of the multiplicity.

This assumption is endorsed by Fig. 6.4 where N_{track} means the number of best tracks reconstructed in the LHCb tracking system. According to this study done by the tracking calibration group at LHCb, the multiplicity dependence of the track reconstruction efficiency is well-modelled by simulation within a 1% agreement. For pPb and Pbp, the occupancy is higher, as shown in Fig. 5.9, so the values in our analysis could exceed the range of the study shown in Fig. 6.4. Thus, we need to know which is the range on N_{track} that the Pbp result is covering.

[1] `https://twiki.cern.ch/twiki/bin/viewauth/LHCbInternal/LHCbTrackingEfficiencies`

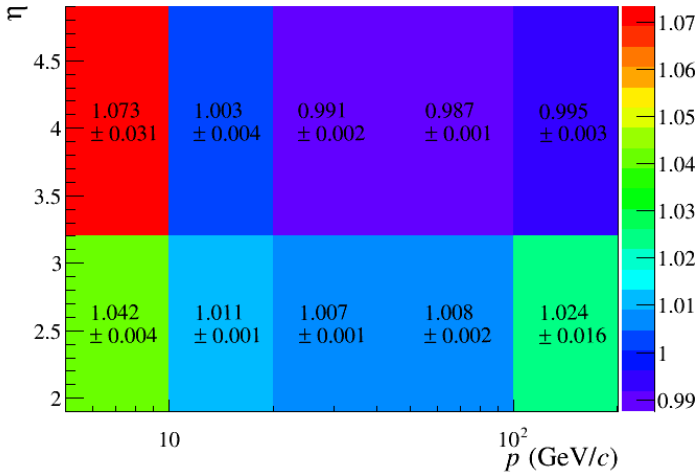


Figure 6.2: Data/MC ratio for reconstruction efficiencies from pp 2012 Data vs MC 2012 (Reco14). This table is used to correct pPb and Pbp data in this analysis.

In order to know whether the majority of the events are within the scope of Fig. 6.4, the profile of the N_{track} vs N_{cand} distribution in Fig. 6.5. According to Fig. 6.5 in the case of pPb the N_{track} range is well within the coverage of the LHCb calibration sample. In the case of Pbp collisions, the occupancy is higher, but still, it is not so far from the limit of 500 NTracks of Fig. 6.4. As a consequence, we assume that the multiplicity dependence of the reconstruction efficiency is well modelled in simulation also for Pbp collisions, and $C_{tracking}$ is independent of the multiplicity in both cases.

Moreover, as explained in Chap. 5, the detector occupancy is correlated to the detector response. For this reason, occupancy dependent weights, w_i, computed as described in Sec. 5.3.4, are applied to the efficiency determination, which is obtained as,

$$\boxed{\varepsilon_{reco} = \frac{\sum N_{reco\ tracks} C_{tracking}(p,\eta)}{\sum N_{gen\ parts}}.} \tag{6.9}$$

6.5.1 Reconstruction efficiency in proton-lead collisions

Since no specific tracking calibration sample exists for the 2013 pPb data set, the 2012 pp data calibration sample is used instead. This sample was chosen because the detector conditions were the same during both data-taking. The values of these tracking calibration factors are shown in Fig. 6.2.

The results of the reconstruction efficiency for pPb and Pbp collisions are shown in Figs. 6.6 and 6.7, as a function of $(\eta, p_{\text{T}}, \texttt{nVeloTracks})$. In general, the reconstruction efficiency decreases with the detector occupancy.

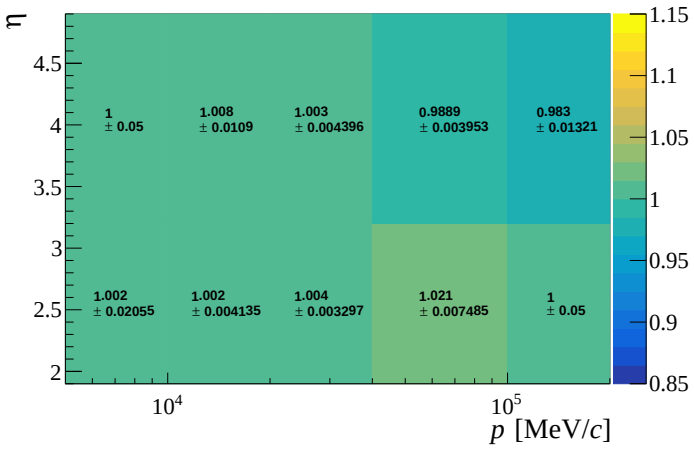


Figure 6.3: Tracking calibration weights for pp data sample

6.5.2 Reconstruction efficiency in proton-proton collisions

The *pp* reconstruction efficiency computation has some particularities that need to be explained in detail. For instance, the discrepancies found in the PYTHIA particle composition of the minimum bias sample require an additional correction [70]. As shown in Fig. 6.8, the kaon over pion fraction measurement from the ALICE collaboration [88] is better reproduced by EPOS-LHC than by PYTHIA.

The method implemented to correct this issue in PYTHIA simulation for *pp* data is the factorization of the reconstruction efficiency into particle species,

$$\varepsilon_{reco} = \sum_{p} f_p \cdot \varepsilon^{p}_{reco}, \tag{6.10}$$

where f_p is the fraction of particle species p computed with EPOS-LHC and ε^{p}_{reco} its reconstruction efficiency computed with PYTHIA using the definition of reconstruction efficiency in Eq. 6.9.

In this factorization, the species considered are π^-, K^-, p, e^-, μ^-, Ξ^-, Σ^-, Σ^+, Ω^-+ *cc*, corresponding with the prompt charged particle definition state in Sec. 6.1. The species fractions will be computed using a private EPOS-LHC sample for *pp* collisions and can be seen in Fig. 6.9 and Fig. 6.10 for PYTHIA and EPOS-LHC, respectively.

The species fractions are not strongly dependent on multiplicity, at least for the main contributions (pions and kaons) as can be seen in B.4 at (Fig. B.8). Because of this, the species fractions, f_p, in Eq. 6.10 are not multiplicity dependent.

To check if the fraction of particles that are not in the factorization list is negligible, their aggregated reconstruction efficiency is compared to the prompt-charged particle efficiency. This comparison is shown in the Appendix. B.4 in Fig. B.10, demonstrating that the effect of unconsidered species is insignificant.

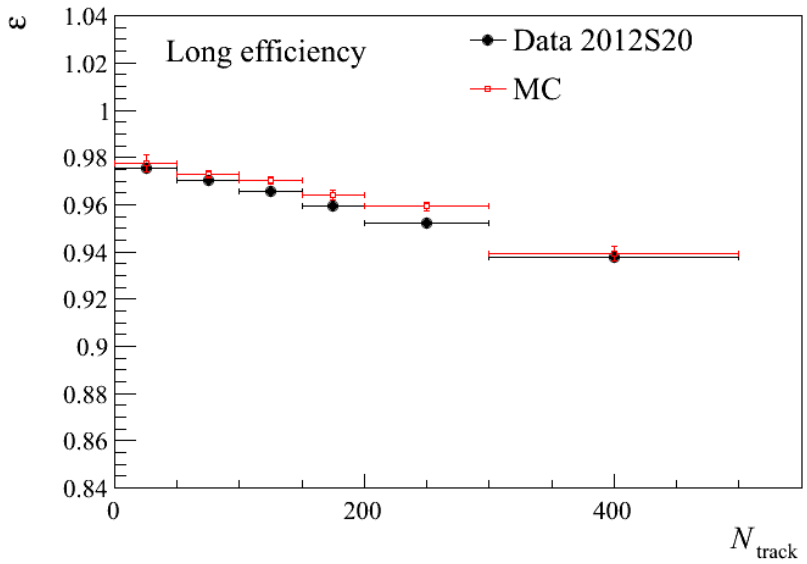


Figure 6.4: Tracking efficiency as a function of N_{track} for data and simulation in pp. Figure extracted from Ref. [75].

Finally, in Fig. B.9 it can be seen the difference in the prompt charged-particle reconstruction efficiency using particle species fractions from EPOS-LHC and Pythia. According to Fig. 6.12 the relative difference is less than 8%, where the high-p_{T} regime gives the largest discrepancies. The final result for pp reconstruction efficiency is shown in Fig. 6.13.

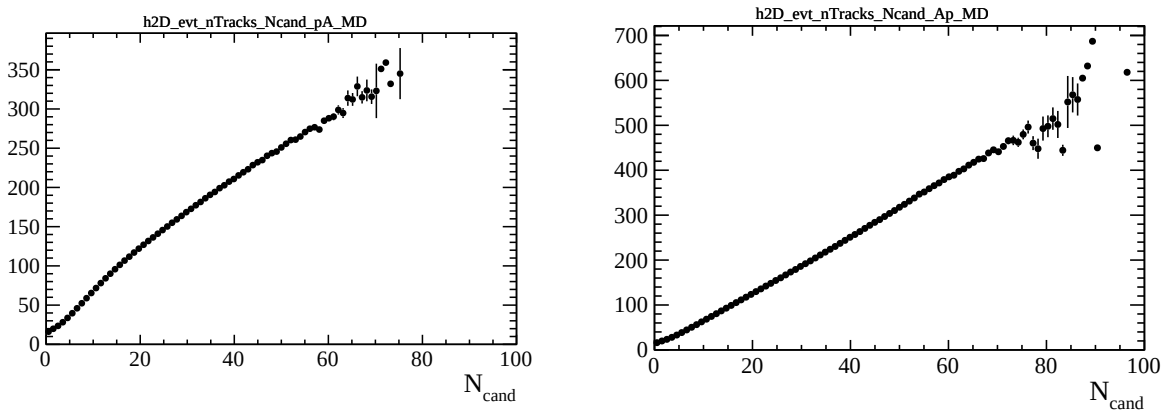


Figure 6.5: Average evt_NTracks vs N_{cand} distribution for pPb and Pbp EPOS-LHC simulation events.

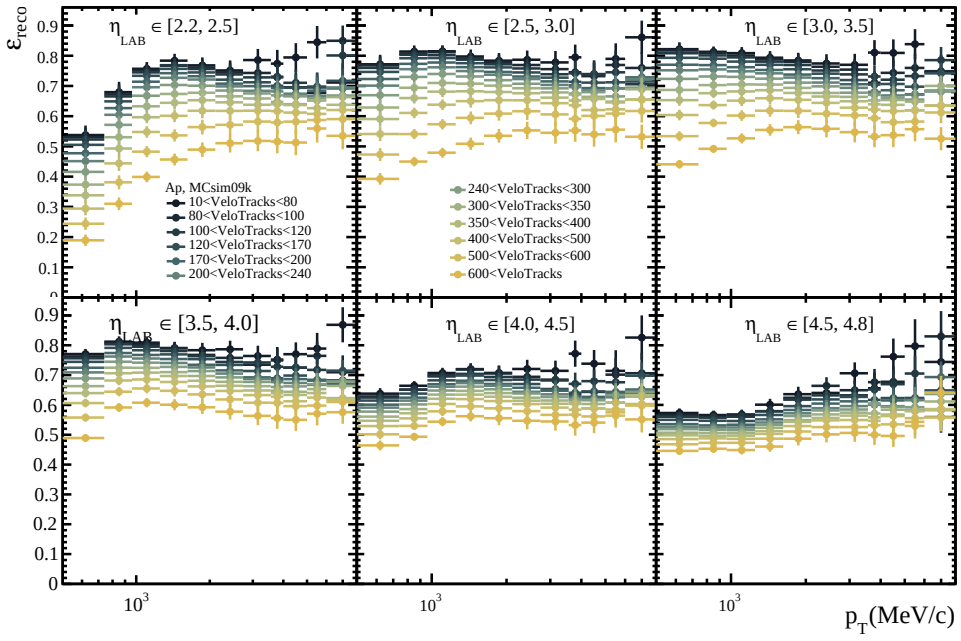


Figure 6.6: Reconstruction efficiency with tracking correction for Pbp events.

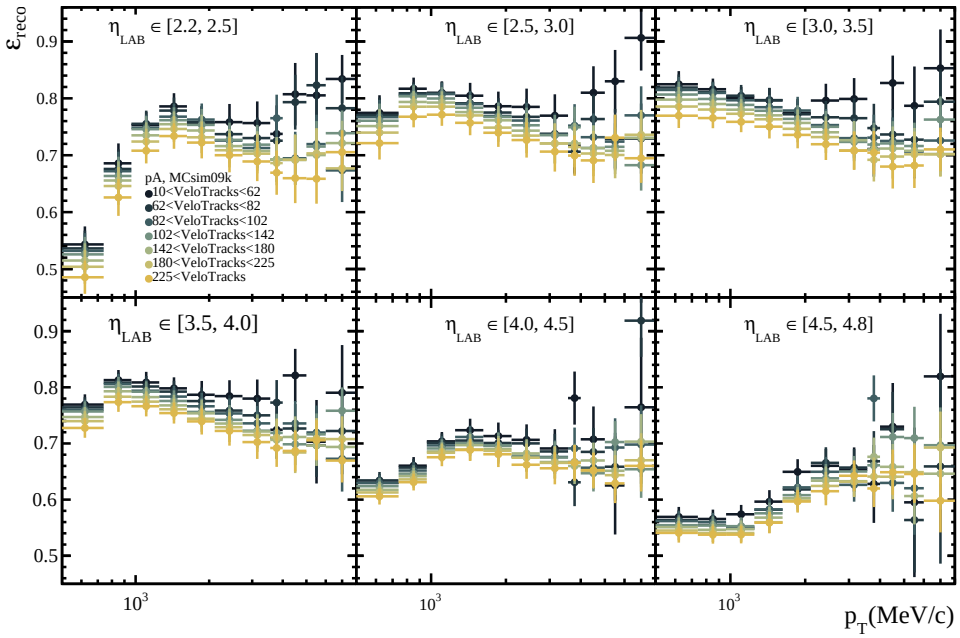


Figure 6.7: Reconstruction efficiency with tracking correction for pPb events.

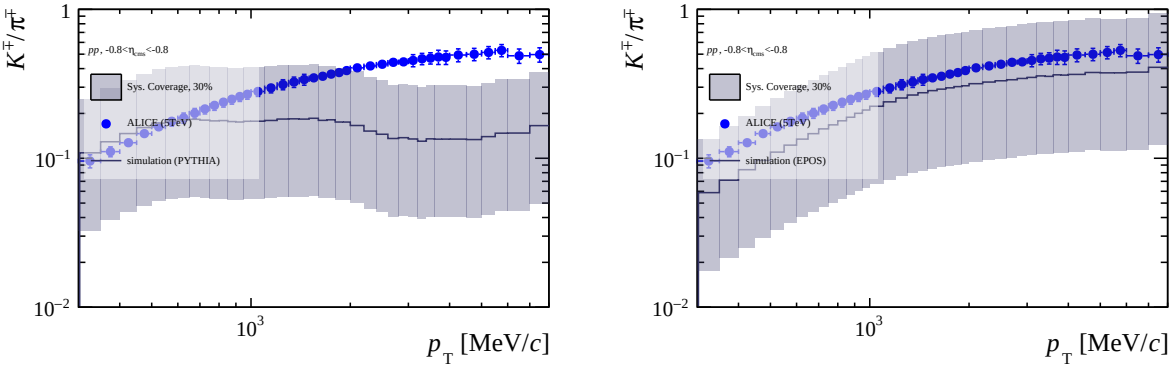


Figure 6.8: Kaon over pion fraction for Pythia (left) and EPOS-LHC (right) in comparison with ALICE measurement

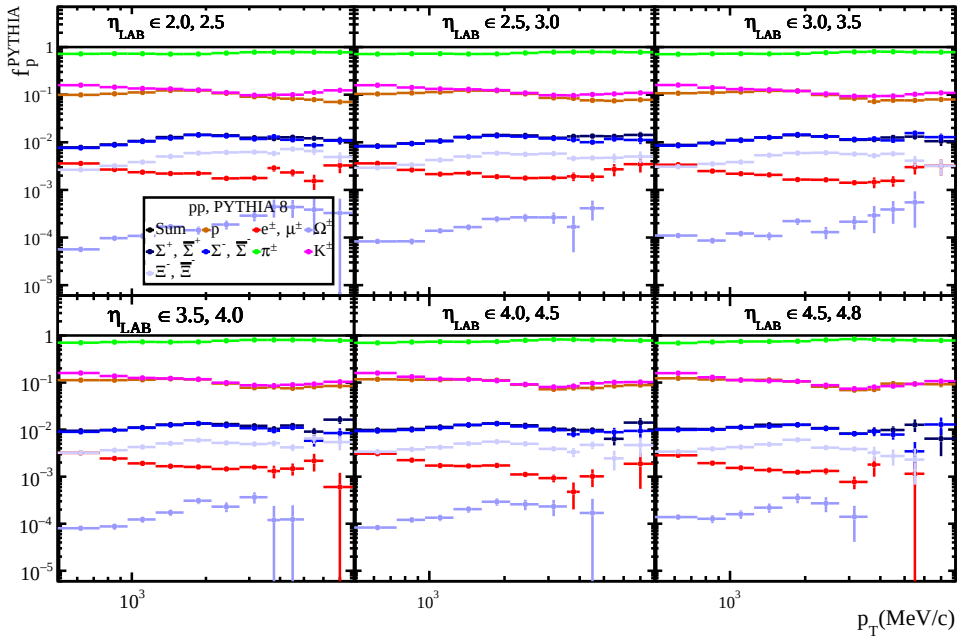


Figure 6.9: Fraction of species abundance in *pp* PYTHIA.

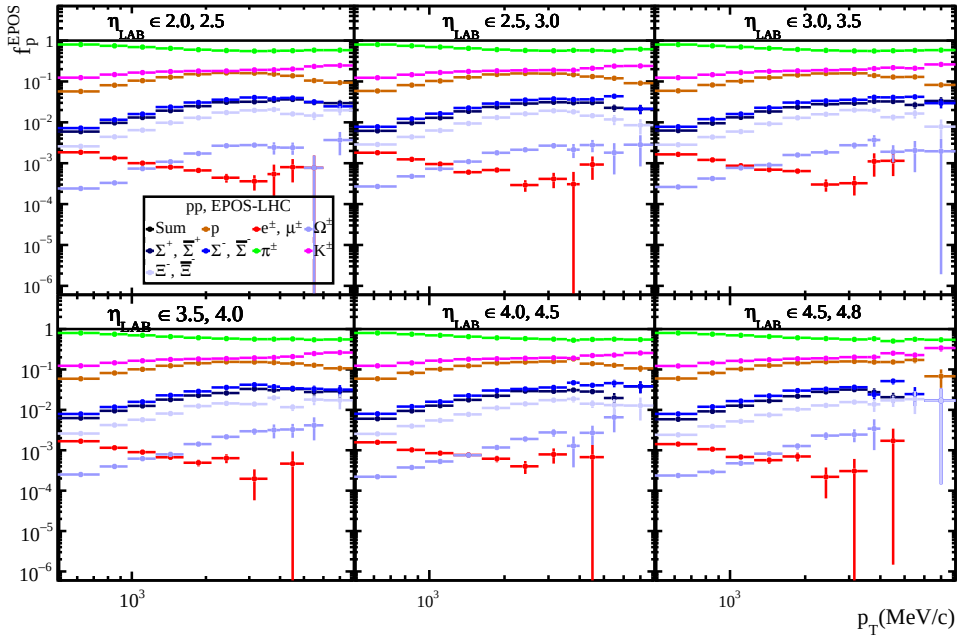


Figure 6.10: Fraction of species abundance in *pp* EPOS-LHC.

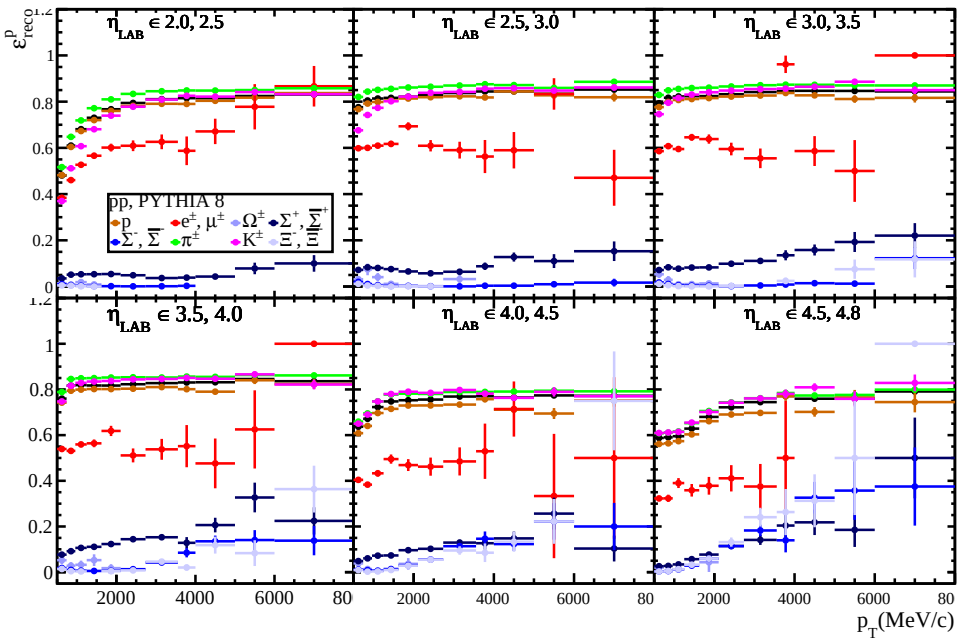


Figure 6.11: Reconstruction efficiency per particle species.

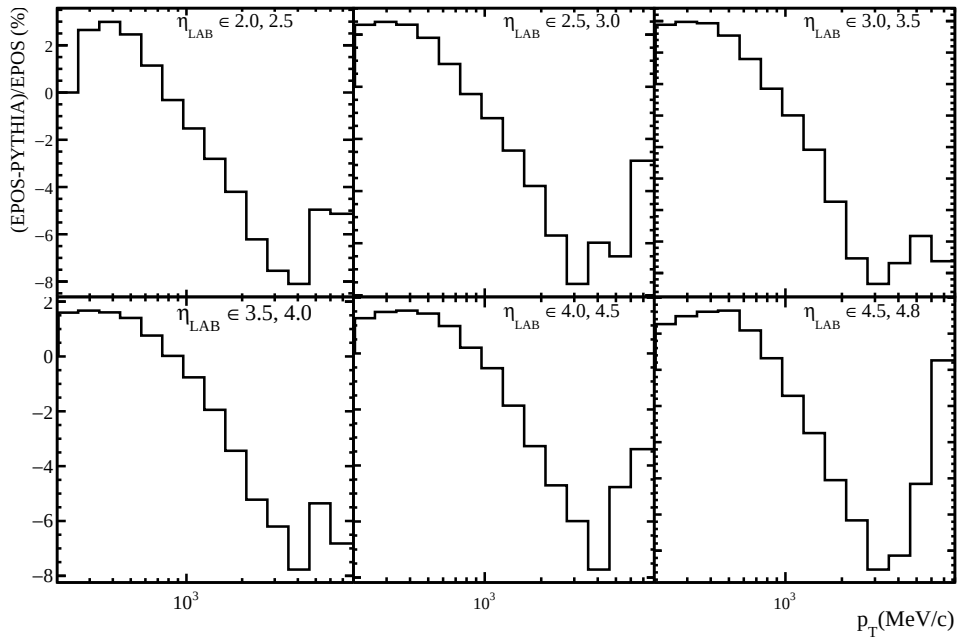


Figure 6.12: Relative difference on the reconstruction efficiency from using Pythia or EPOS-LHC species fractions.

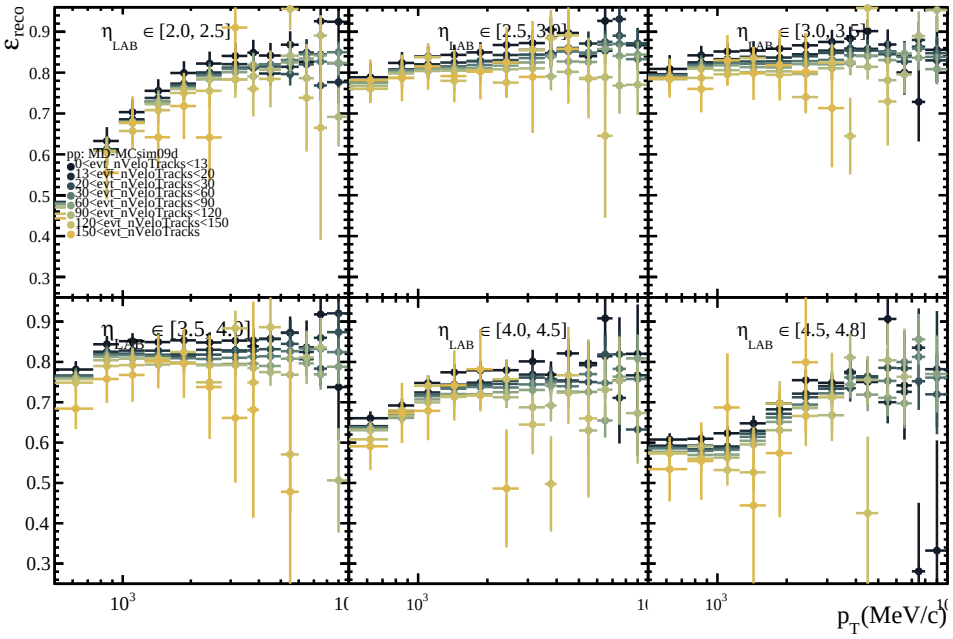


Figure 6.13: Reconstruction efficiency for *pp* events using EPOS-LHC species fractions.

6.6 Selection efficiency

When the selection cuts described in Tab. 5.5 are applied to the sample to remove the background, some signal is also removed. The selection efficiency takes into account this loss of candidates. In order to compute the selection efficiency, simulation-matched tracks, which are reconstructed tracks assigned (matched) to a prompt-charged particle, are used. The selection efficiency is estimated from the occupancy-weighted simulation sample as the number of matched long tracks passing the selection criteria divided by the total number of matched long tracks within the LHCb detector acceptance,

$$\varepsilon_{sel} = \frac{\sum N_{sel\ matched\ tracks}}{\sum N_{matched\ tracks}}. \tag{6.11}$$

The results for the selection efficiency ε_{sel} at different collision systems can be found at Fig. 6.14, Fig. 6.15 and Fig. 6.16.

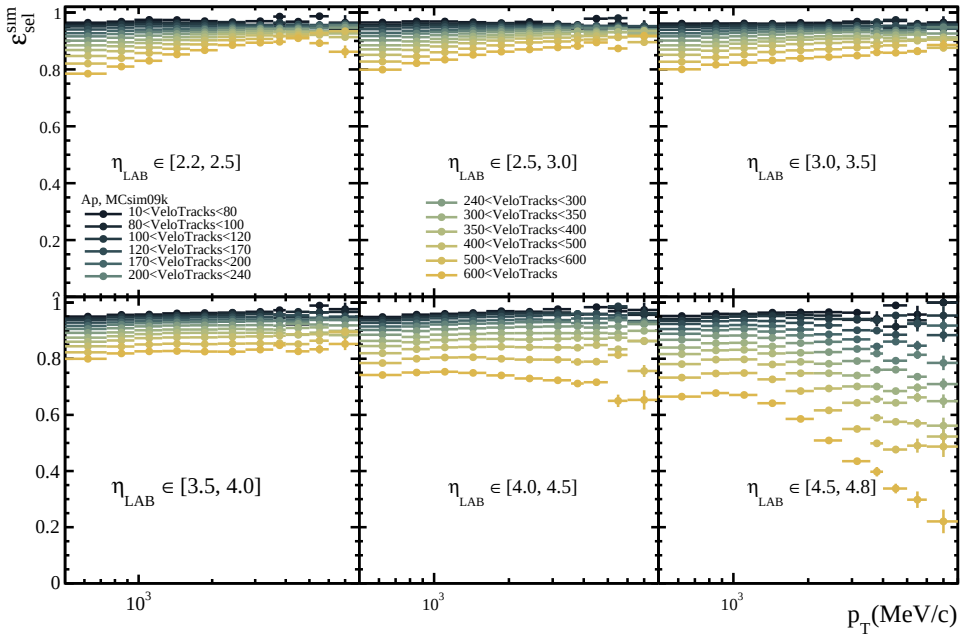


Figure 6.14: Selection efficiency for Pbp collisions.

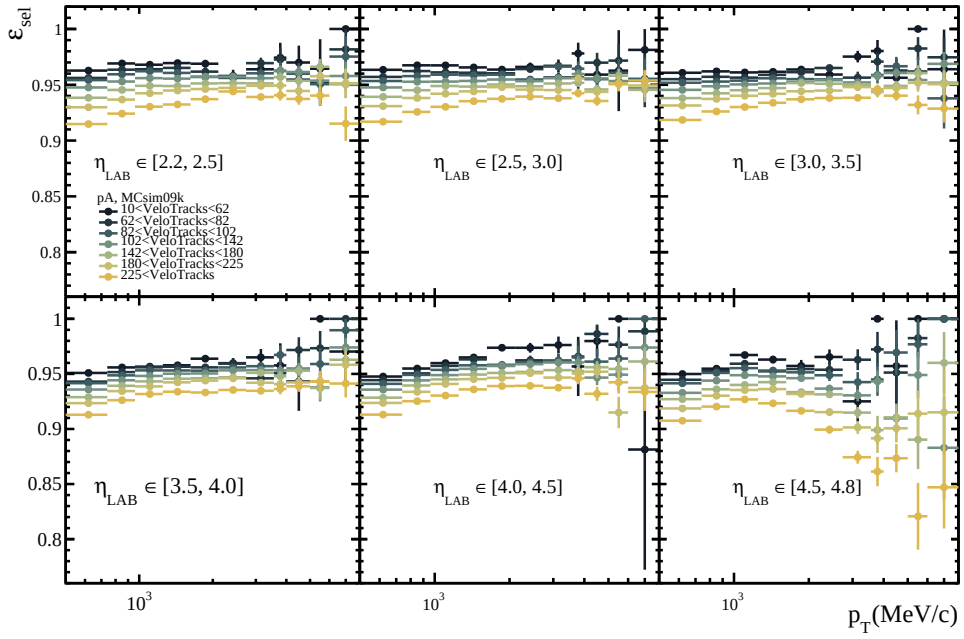


Figure 6.15: Selection efficiency for pPb collisions.

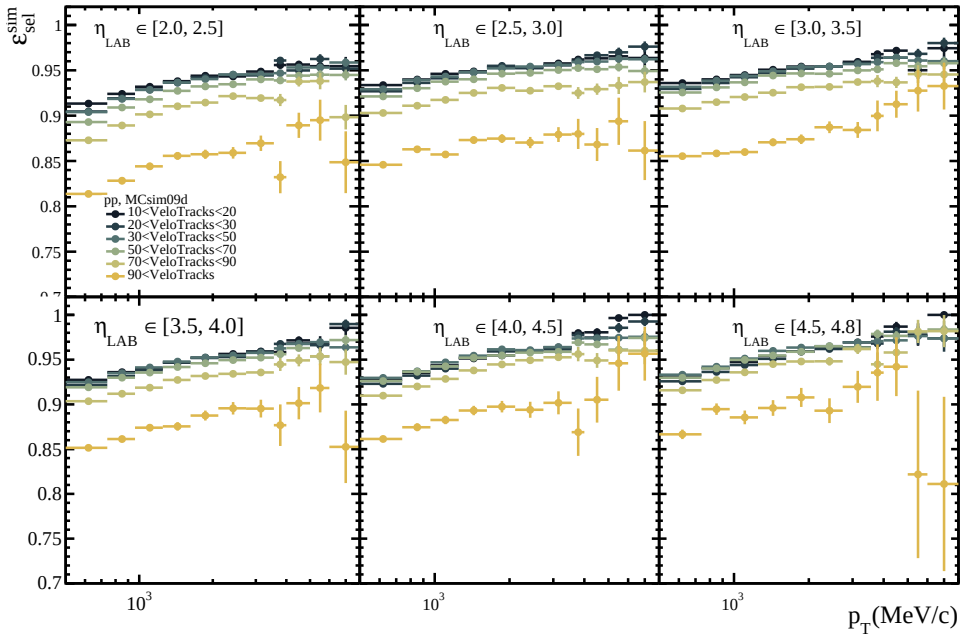


Figure 6.16: Selection efficiency for pp collisions.

6.7 Signal purity

After applying the selection requirements there is still a background contribution in the signal sample. To estimate it we introduced the purity factor, varying from 0 to 1, as

$$\Pi^{sim} = 1 - (f^{sim}_{fake} + f^{sim}_{sec} + f^{sim}_{clone}). \tag{6.12}$$

This section will explain the procedure to compute the purity of the signal sample. The main background sources in this sample are fake tracks, secondary particles and clone tracks. We will denote as f^{sim}_{fake}, f^{sim}_{sec} and f^{sim}_{clone} the fraction of this background sources obtained from simulation. The purity can be computed in simulation by getting the background fractions in the signal sample.

Using Eq. 6.12, an estimation of the signal purity can be computed using simulation. However, the simulation background fractions differ from those in the data. In order to have a more precise computation of these fractions, we employ data-driven correction factors. These factors will be denoted as R_α, where α stands for each background source detailed in Sec. 5.3,

$$R_\alpha = \frac{\frac{1}{N_{\text{cand}}} \sum_i N^{data}_{\alpha,i}}{\frac{1}{\sum_i w_i N_{\text{cand},i}} \sum_i w_i N^{sim}_{\alpha,i}}. \tag{6.13}$$

where the simulation is weighted using the occupancy variables to match data distributions. In this equation, $N^{data}_{\alpha,i}$ ($N^{sim}_{\alpha,i}$) is the number of long tracks in the i-th event of the proxy sample for data (simulation), N_{cand} is the total number of candidates and w_i the occupancy weights for the i-th event. The strategy is to create a high-purity proxy sample in data and simulation for each background source α using dedicated selections. Then, R_α can be computed as a data/simulation ratio of this sample using Eq. 6.13. This procedure will be explained in detail for each background contribution. Moreover, N_{cand} and N_α are computed in $(\eta, p_{\text{T}}, \texttt{nVeloTracks})$ bins. In the $(\eta, p_{\text{T}}, \texttt{nVeloTracks})$ bins where the proxy sample has no candidates R_α is set equal to 1.

Finally, using these correction factors, the background fractions for each background source α can be computed using as,

$$f_\alpha = R_\alpha f^{sim}_\alpha, \tag{6.14}$$

Following this strategy, we assume that the data/simulation ratio of a specific background source is the same in the high-purity sample and in the candidate sample. This assumption will be subject to systematic uncertainties detailed in Chap. 7. The final results of the background fraction data-driven correction factors, R_α, can be seen in Figs. B.20, B.22 and B.21 for fake tracks, in Figs. B.23, B.24, and B.25 for gamma to electron conversion, in Figs. B.26 and B.27 for muon decays and in Figs. B.29 and B.30 for hadronic decays.

Consequently, the purity of the signal sample can be obtained as,

$$\Pi = 1 - (f_{fakes} + f_{sec} + f_{clone}^{sim}). \tag{6.15}$$

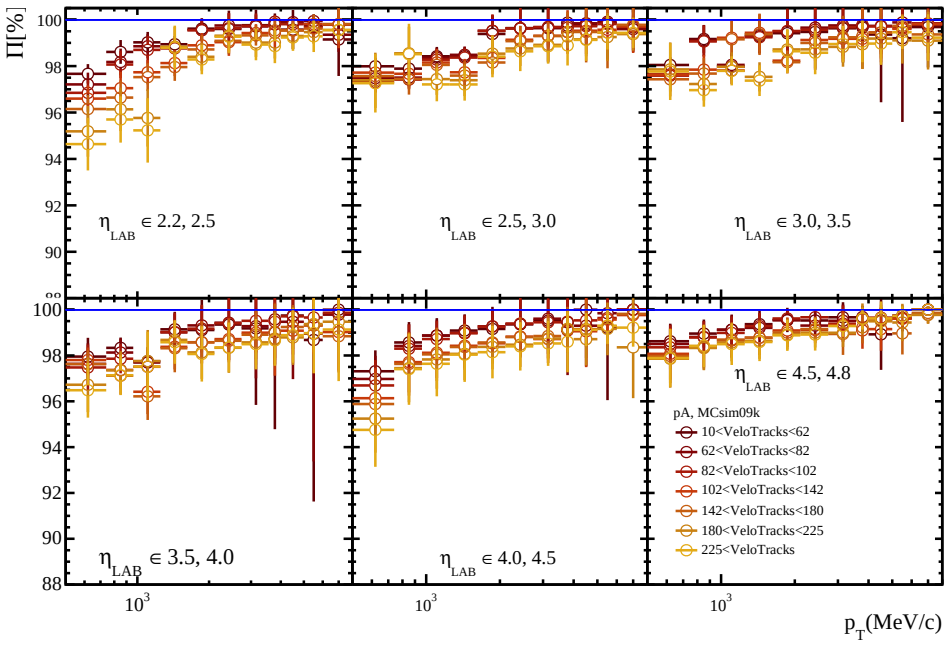


Figure 6.17: Purity for pPb collisions including data-driven correction factors.

6.7.1 Fake Tracks

In the simulation, the fraction of fake tracks is computed using the occupancy-weighted sample with the expression

$$f_{fake}^{sim} = \frac{\sum_i N_{non-matched\,cand}}{\sum_i N_{\text{cand}}}. \tag{6.16}$$

In this equation, $N_{non-matched\,cand}$ is the number of non-matched-to-a-generated-MC-particle tracks and N_{cand} is the number of candidates. Following this definition, a fake track is a reconstructed long track which does not correspond to a generated particle, therefore a fake track is a reconstruction artefact.

In this section, the details to compute the data-driven correction factor for f_{fake}^{sim} are explained. In general, the strategy is the same as for the other data-driven correction factors, R_α, explained at the beginning of this section. We isolate the background source using the specific selection detailed in Tab. 6.4 and we compute the ratio of background fractions between data and simulation using Eq. 6.13.

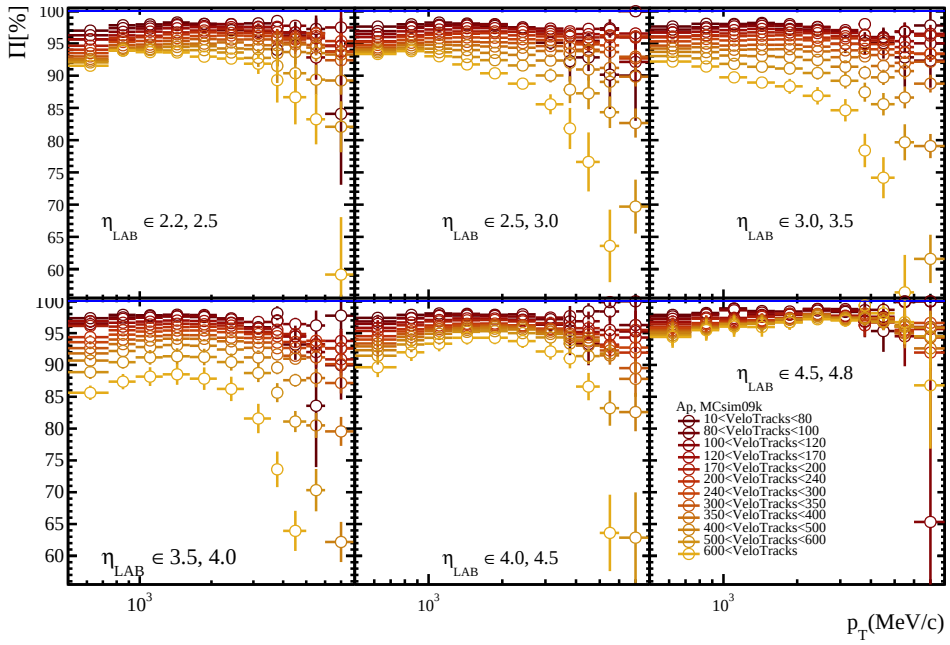


Figure 6.18: Purity for Pbp collisions.

Overall, two different variables will be used to obtain a proxy sample enriched in fake tracks:

- `GhostP`: The fraction of fake tracks increases with `GhostP`. In general, with a `GhostP` > 0.5 requirement the fake track abundance will be 93% for pPb data and 97% for Pbp data. It is worth mentioning that for pp events a `GhostP` < 0.4 cut was applied at the reconstruction level. This cut reduces the fraction of fake tracks but limits the use of this variable as a tool for isolating background samples.

- **Shared VELO**: This variable indicates if the candidate shares the VELO segment with another long track.[2] Samples of candidates with shared VELO tracks are enriched in fake tracks. However, only a small portion of tracks have the shared VELO property, limiting the size of these samples.

For the pPb and Pbp collisions systems, a sample of non-shared-VELO tracks with a high `GhostP` value (between 0.5 and 0.9) is used to define an enriched fake tracks sample. A sample of candidates with shared-VELO will be considered as an alternative proxy to estimate systematic uncertainties, discussed in Sec. 7.2.1. For pp data, shared-VELO tracks (`IsClone` $== True$) with low `GhostP` values (between 0.1 and 0.4) are selected. The summary of the cuts applied to define the proxy enriched in fake tracks can be found

[2] More information about long tracks sharing the VELO segment can be found in Ref. [70]

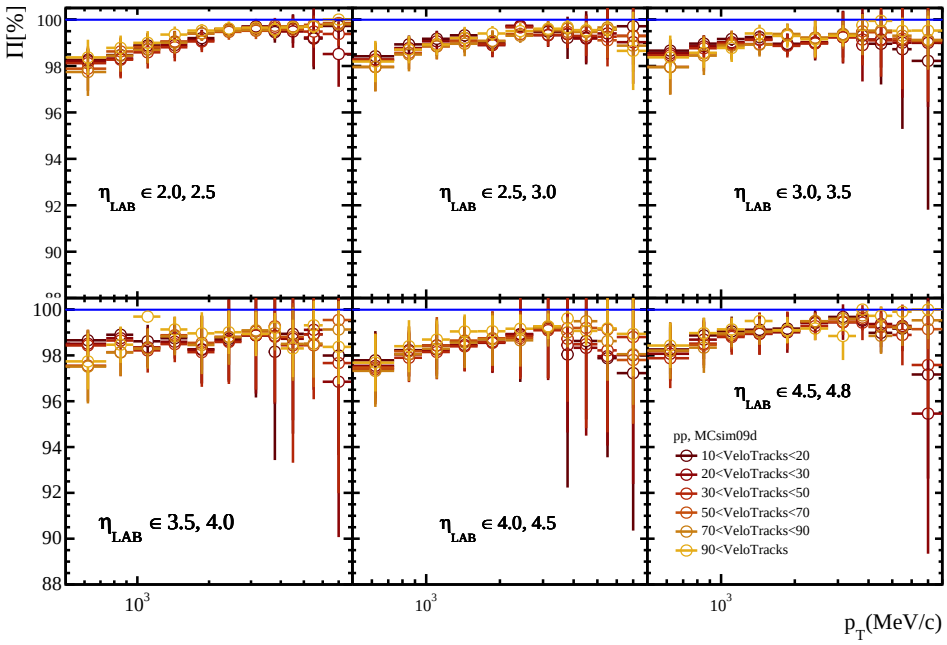


Figure 6.19: Purity for pp collisions.

in Tab. 6.4. Using these samples and Eq. 6.13, the correction factors are computed, as shown in the Appendix B.5. After that, we use Eq. 6.14 to estimate the background fraction. The results are shown in Figs. 6.20, 6.21 and 6.22.

	pp	pPb, Pbp
Track type	Long	Long
η	$2 < \eta < 4.8$	$2 < \eta < 4.8$
p_{T}	$0.5 < p_{\mathrm{T}} < 8$ GeV/c	$0.5 < p_{\mathrm{T}} < 8$ GeV/c
p	$p > 2$ GeV/c	$p > 2$ GeV/c
`GhostP`	$0.1 <$ `GhostP` < 0.4	$0.5 <$ `GhostP` < 0.9
`IsClone`	True	False
IP	< 0.568 mm	< 0.508 mm

Table 6.4: List of applied cuts to define the fake tracks proxy sample for each collision system.

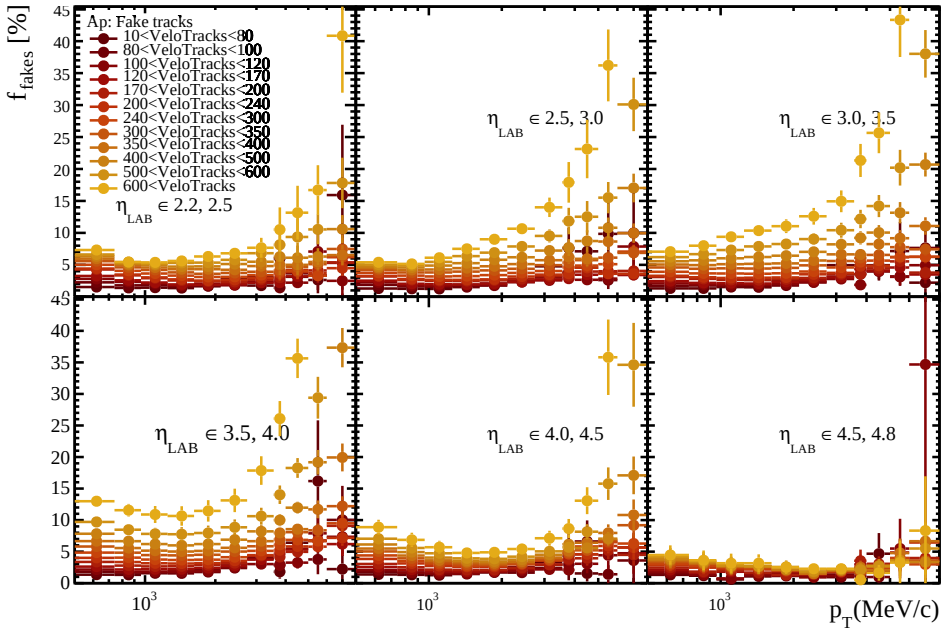


Figure 6.20: Fake tracks fraction for Pbp collisions.

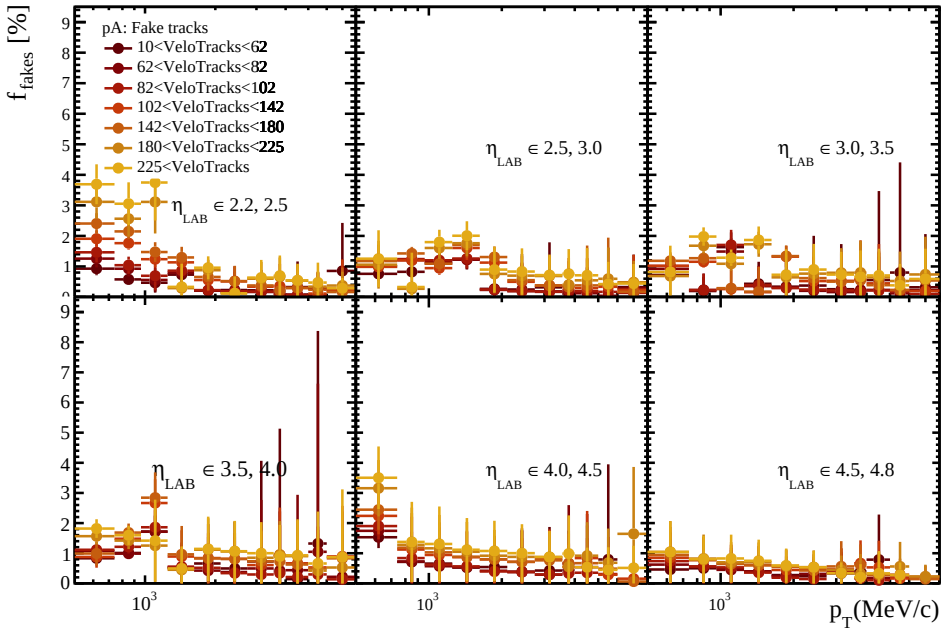


Figure 6.21: Fake tracks fraction for pPb collisions.

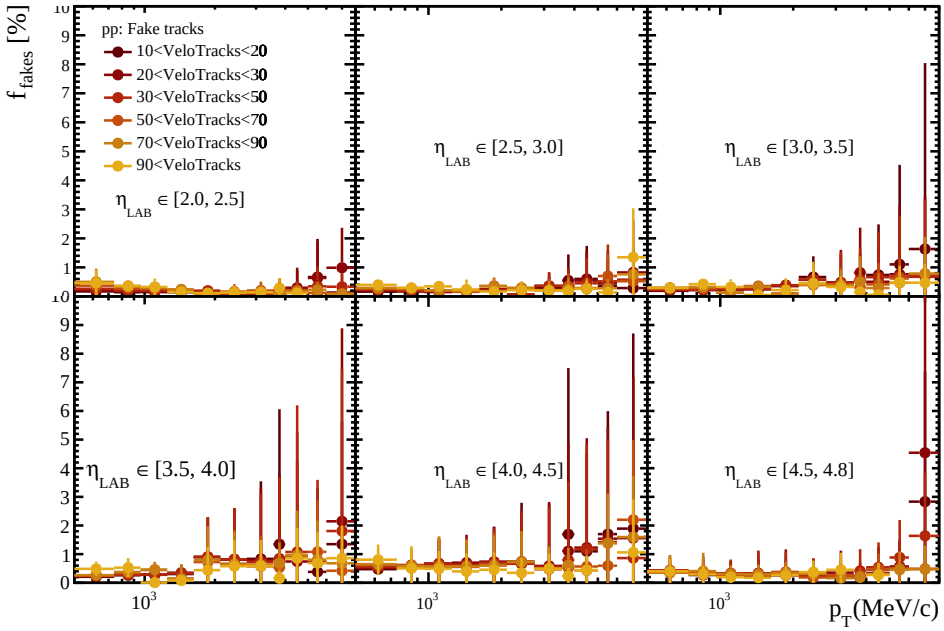


Figure 6.22: Fake tracks fraction for *pp* collisions.

6.7.2 Secondary particles

The fraction of secondary particles is computed from simulation using the occupancy-weighted sample as

$$f_{sec}^{sim} = \frac{\sum_i N_{cand}^{non-prompt}}{\sum_i N_{\text{cand}}} \tag{6.17}$$

where $N_{cand}^{non-prompt}$ is the number of non-prompt candidates. These are tracks that pass selection criteria but do not fulfil the prompt charged particle definition explained in Sec. 6.1. As shown in Fig. 6.25, Fig. 6.23 and Fig. 6.24, the fraction of secondaries is higher in the low-p_{T} region with little η dependence.

If we analyse the secondary particles by its origin, we can identify four different sources:

- $X \to \mu$: **Muons** from particle decays.
- $\gamma \to e$: Converted **photons.**
- **Material:** Particles produced in hadronic interactions with the detector material.
- **Hadronic decays:** Hadrons from decays of prompt particles as $\Lambda, K_s^0, \Sigma^{\pm}, \Xi, \Omega$.

According to these secondary particle sources, the fraction of the secondary particles from the simulation can be factorised as,

$$f_{sec}^{sim} = f_{\gamma}^{sim} + f_{X\to\mu}^{sim} + f_{mat}^{sim} + f_{decays}^{sim}. \tag{6.18}$$

The source fractions in simulation for each collision system are shown in Figs. 6.23, 6.24, and 6.25. In general, converted photons are the dominant component, followed by hadronic decays.

In order to study the differences between the background induced by secondary particles in data and Monte Carlo, we build a proxy sample for each background component using the selection described in Tab. 6.5.

- $DLL_{e\pi} \to$ Electron-pion discrimination PID variable.
  Isolates converted photons background.
- $DLL_{\mu\pi} \to$ Muon-pion discrimination PID variable.
  Isolates background from decays into muons.
- IP $\to$ It isolates particles from hadronic interactions with the material and hadronic decays. Since these two contributions cannot be easily separated in data, they are integrated into a single contribution called $HadDecay$. Therefore, from now f_{mat} and f_{decays} will be substituted for $f_{Had\&Decays}$.

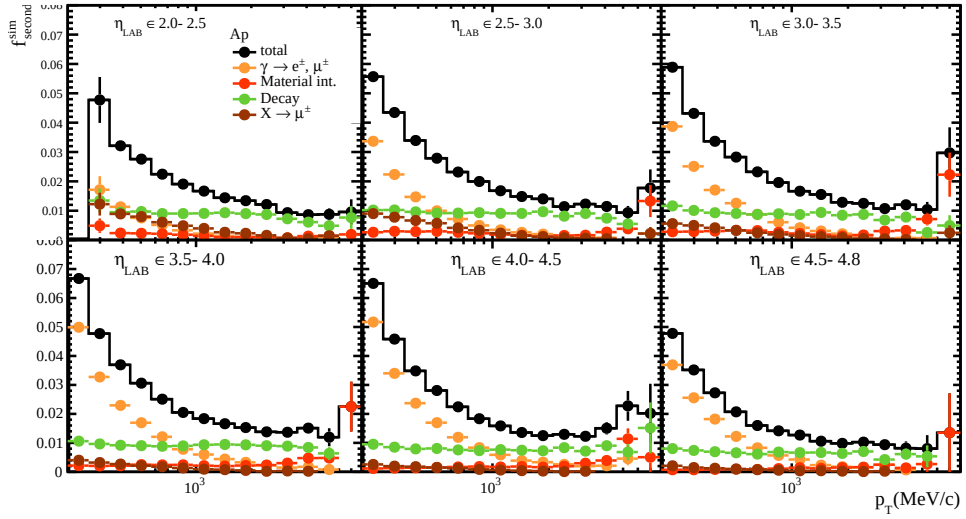


Figure 6.23: Secondary particles fraction per species for Pbp collisions.

To select the DLL and IP values for each (η, p_{T}) bin within the lists defined in Tab. 6.5 we define an estimator as

$$E = \sqrt{u^2_{stat} + u^2_{purity}}, \tag{6.19}$$

where u_{stat} is the statistical uncertainty on R_α and u_{purity} is the systematic uncertainty due to the contamination of the isolation sample computed following the equation

$$u_{purity} = (1 - \Pi)R. \tag{6.20}$$

The selected values for the discrimination variables used to isolate each background contribution are the ones minimizing the E factor on simulation.

Once we have isolated each background contribution, a corresponding data-driven correction factor is computed as

$$R_\alpha = \frac{f^{data}_\alpha}{f^{sim}_\alpha}, \tag{6.21}$$

where α stands for each source of secondary particles. With this factor, we can obtain the final fraction of secondary particles in the signal sample as,

$$f_{sec} = R_\gamma f^{sim}_\gamma + R_{Had\&Decays} f^{sim}_{Had\&Decays} + R_\mu f^{sim}_\mu. \tag{6.22}$$

The results for the gamma conversion data-driven correction are shown in Fig. B.24, Fig. B.23 and Fig. B.25 in the Appendix B.5. The values are within a range from 0.5 to 1.5 with a big difference between (η, p_{T}) bins.

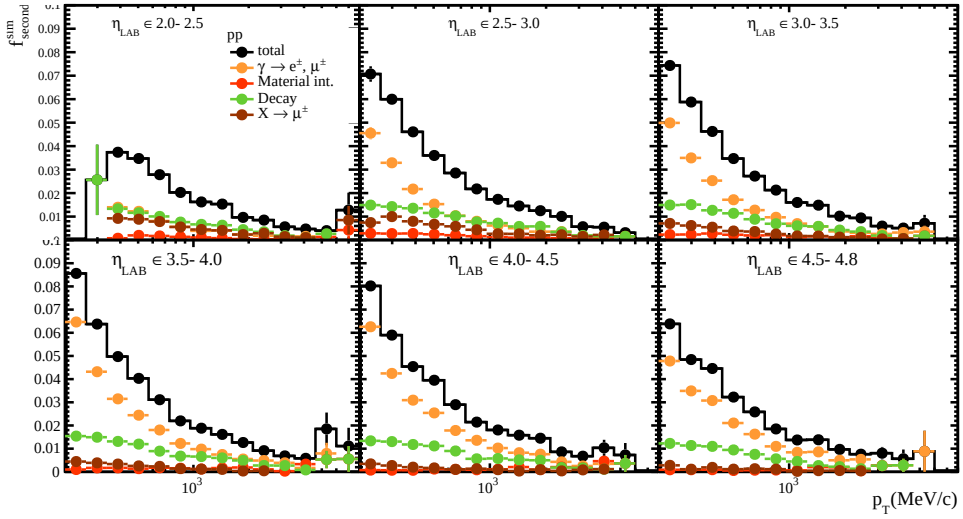


Figure 6.24: Secondary particles fraction per species for *pp* collisions.

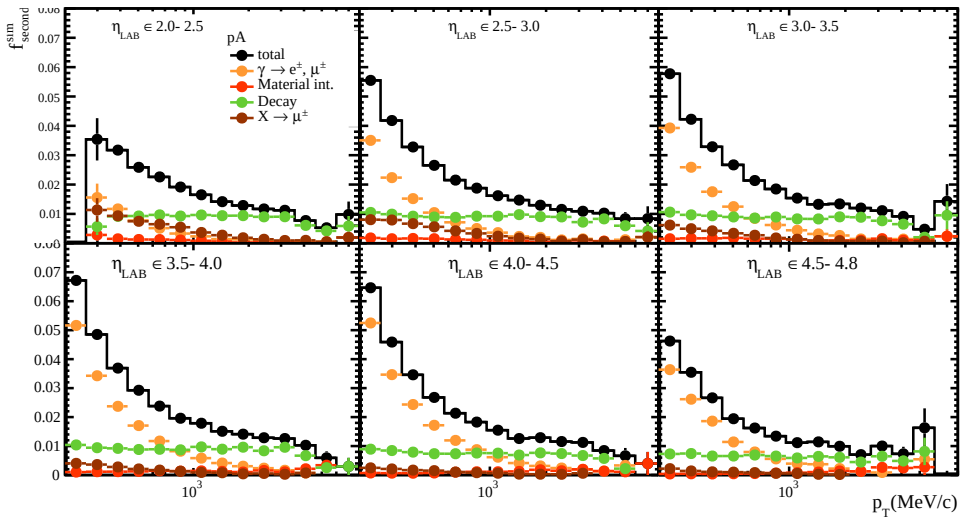


Figure 6.25: Secondary particles fraction per species for *p*Pb collisions.

	pp	pPb	Pbp
Common requirements for all sources			
Track type	Long		
η	$2 < \eta < 4.8$		
p_{T}	$0.5 < p_{\mathrm{T}} < 8$ GeV/c		
p	$p > 2$ GeV/c		
`GhostP`	`GhostP` < 0.078	`GhostP` < 0.103	`GhostP` < 0.109
`IsClone`	False		
R_γ specific requirements			
IP	< 0.368	< 0.348	< 0.348
$DLL_{e\pi}$	$>$[15, 20, 25, 30, 35, 40, 45, 50, 55, 60]		
R_μ specific requirements			
IP	< 0.368	< 0.348	< 0.348
$DLL_{e\pi}$	$<$15		
$DLL_{\mu\pi}$	$>$[15, 20, 25, 30, 35, 40, 45]		
$R_{HadDecay}$ specific requirements			
IP	$> [0.1, 0.15, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8]$ mm		

Table 6.5: List of applied cuts to isolate secondary particles proxy in data for each source and collision system.

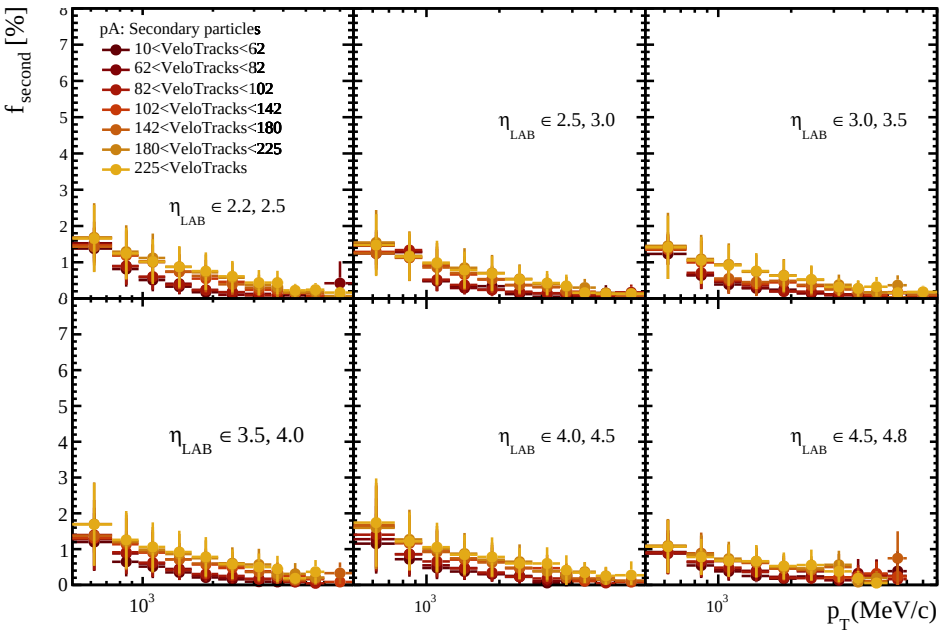


Figure 6.26: f_{sec} for pPb collisions system.

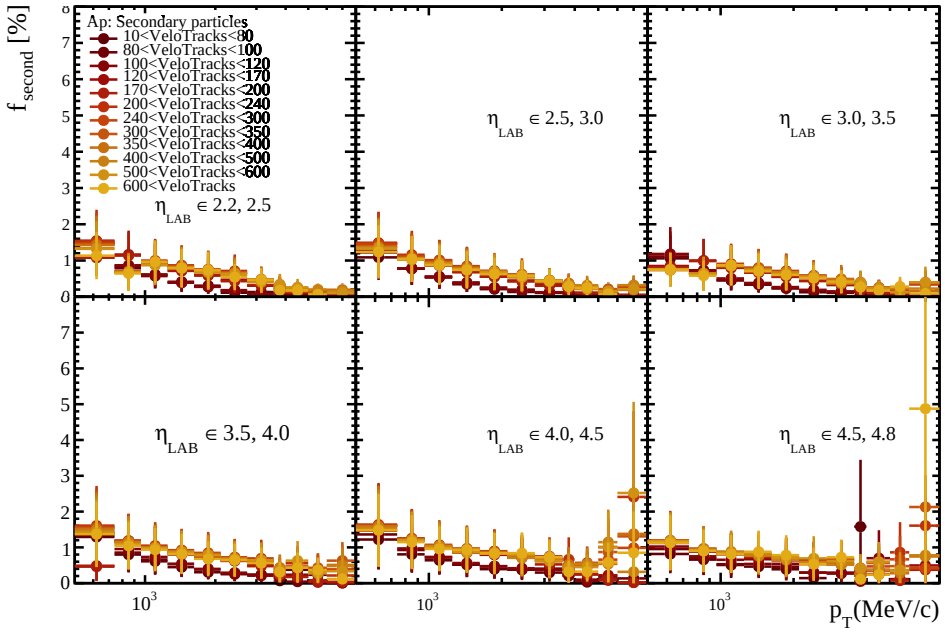


Figure 6.27: f_{sec} for Pbp collisions system.

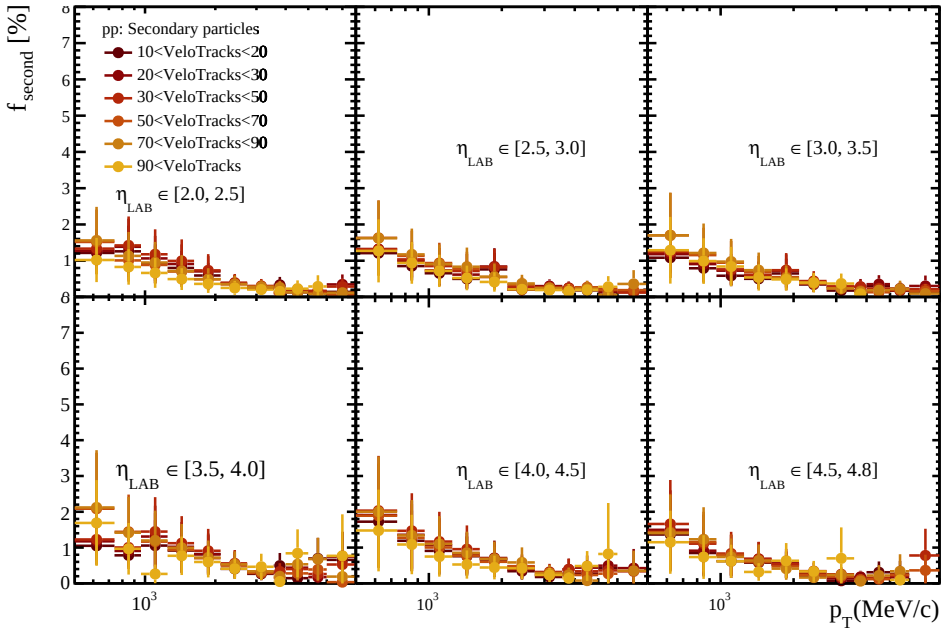


Figure 6.28: f_{sec} for pp collisions system.

6.7.3 Clone tracks

The fraction of clone tracks is computed from simulation using the occupancy-weighted samples according to

$$f_{clone}^{sim} = \frac{\sum_i N_{\text{cand}}^{clone}}{\sum_i N_{\text{cand}}} \tag{6.23}$$

where N_{cand}^{clone} is the number of tracks in the candidate sample that fulfil the clone tracks criteria (`IsClone` $== True$) and N_{cand} is the number of candidates. Results for the different collision systems are shown in Figs. 6.29, 6.30 and 6.31. Since the clone fraction is very small, $f_{clone}^{sim} < 1\%$, no data-driven correction will be computed and this factor will be taken directly from the simulation.

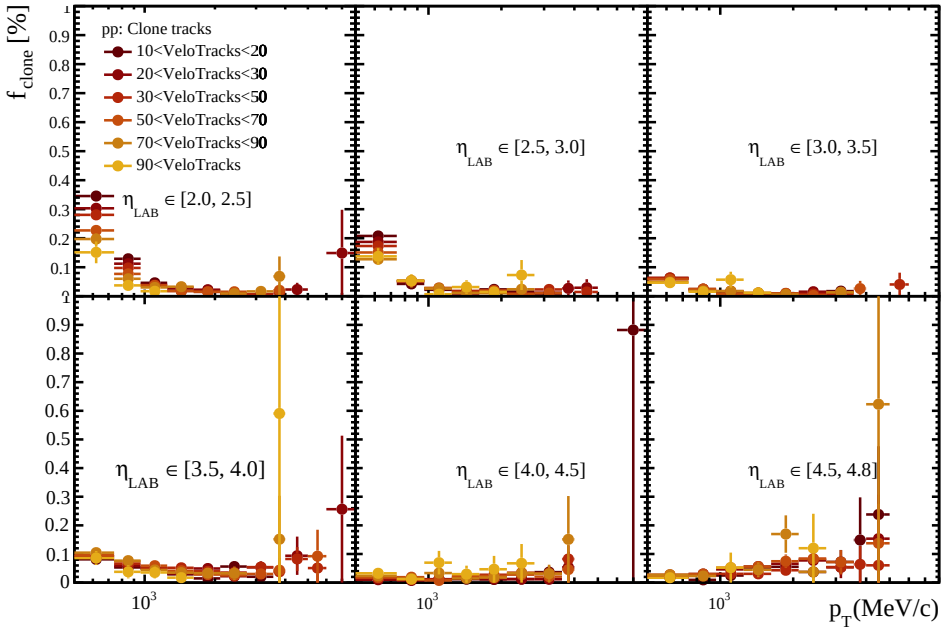


Figure 6.29: Clone tracks fraction for *pp* collisions.

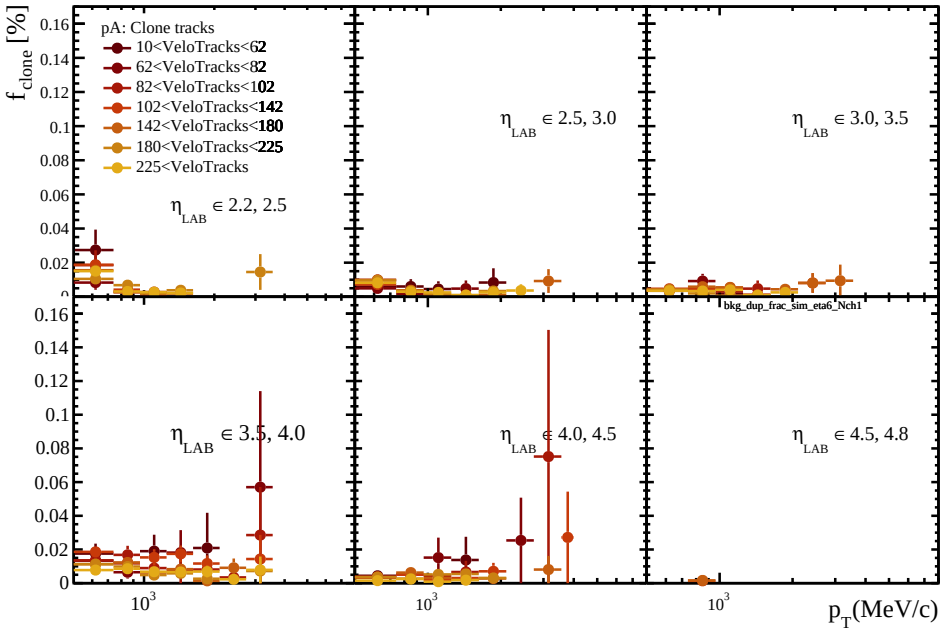


Figure 6.30: Clone tracks fraction for *p*Pb collisions.

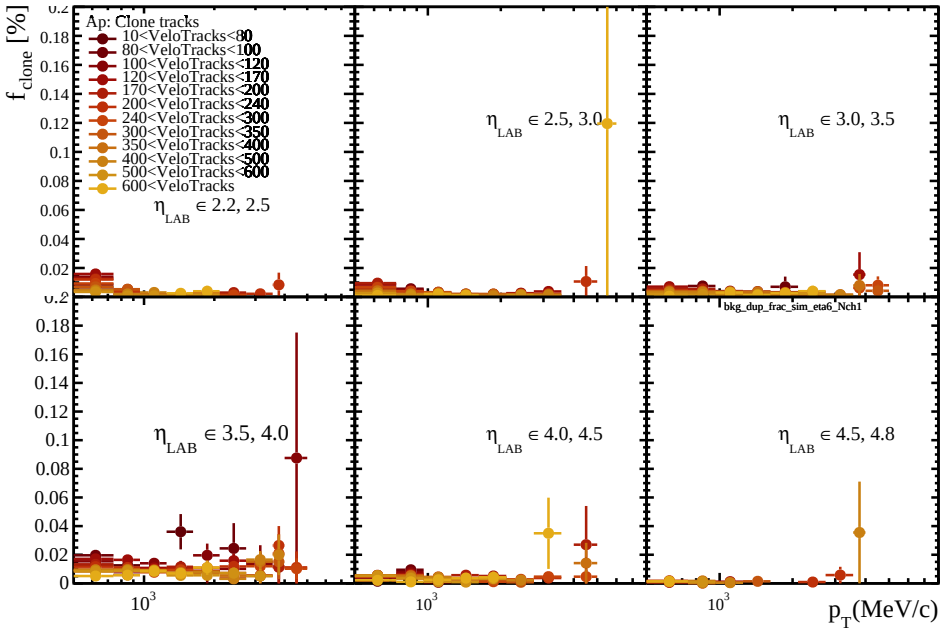


Figure 6.31: Clone tracks fraction for Pb*p* collisions.

6.8 Closure test for the multiplicity distribution

As it was explained in Sec. 6.2 when applying efficiencies and purity correction to the candidate distribution, $P(N_{\text{cand}})$ many detector effects get corrected. This can be seen in Figs. B.32 when the multiplicity distribution is compared with the generator level one after each correction is applied. However, finite bin size and bin migration are still uncorrected from the $P(N_{\text{corr}})$ spectra and this makes $P(N_{\text{corr}})$ different from the generated spectra as can be seen in those figures. In order to correct these effects, a Bayesian Unfolding procedure is applied. In this note, Bayesian Unfolding uses the standard number of iterations in RooUnfold; i.e. $n_{iter} = 3$. No significant change was observed when the number of iterations was changed.

Unfolding matrices can be seen in Figs. 6.32. Unfolding is a delicate and complex procedure that must be carefully cross-checked to avoid model dependency and biases. One fundamental cross-check is the closure test. This is a simulation-based test where we compare the unfolded spectra with the generated ones. If the unfolding is properly implemented, the unfolded spectra should match the generated one exactly. In Fig. 6.33, the closure-test for pPb, Pbp and pp in the full LHCb acceptance are shown. According to this plot, the unfolding algorithm recovers the generated spectrum precisely. It is important to note that the response matrix (R) used to perform the unfolding contains only simulation information and that the corrections do not include data-driven factors.

When the multiplicity spectrum is measured in several pseudorapidity bins, a dedicated response matrix, unfolding and closure test are performed for every η bin. The results can be seen at Sec. B.3. The closure test is fulfilled in all the tests.

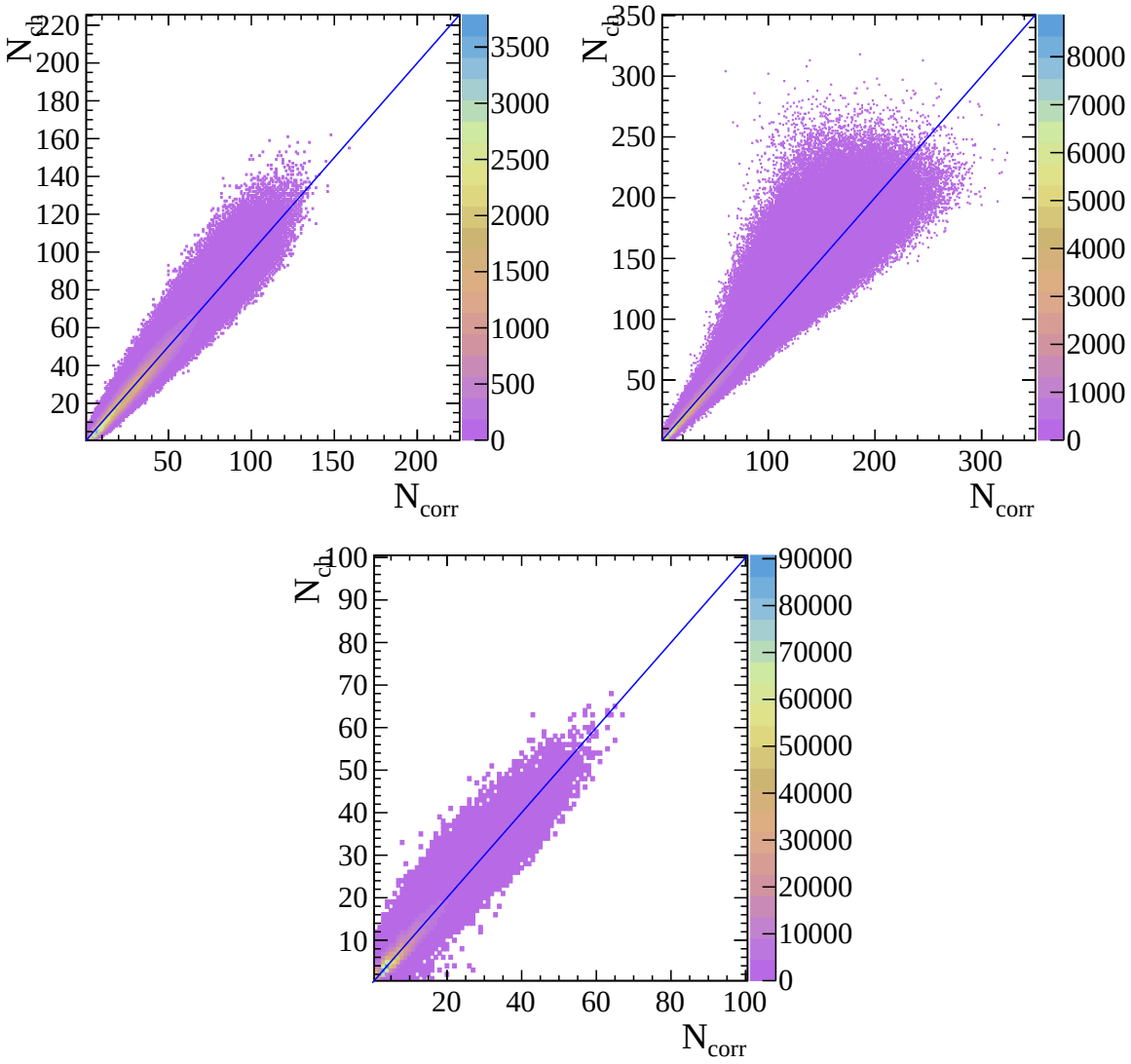


Figure 6.32: Response matrices for *p*Pb (top left), Pb*p* (top right) using MCsim09k simulation and *pp* (bottom) using MCsim09e simulation. The blue line is just a diagonal line to facilitate interpretation.

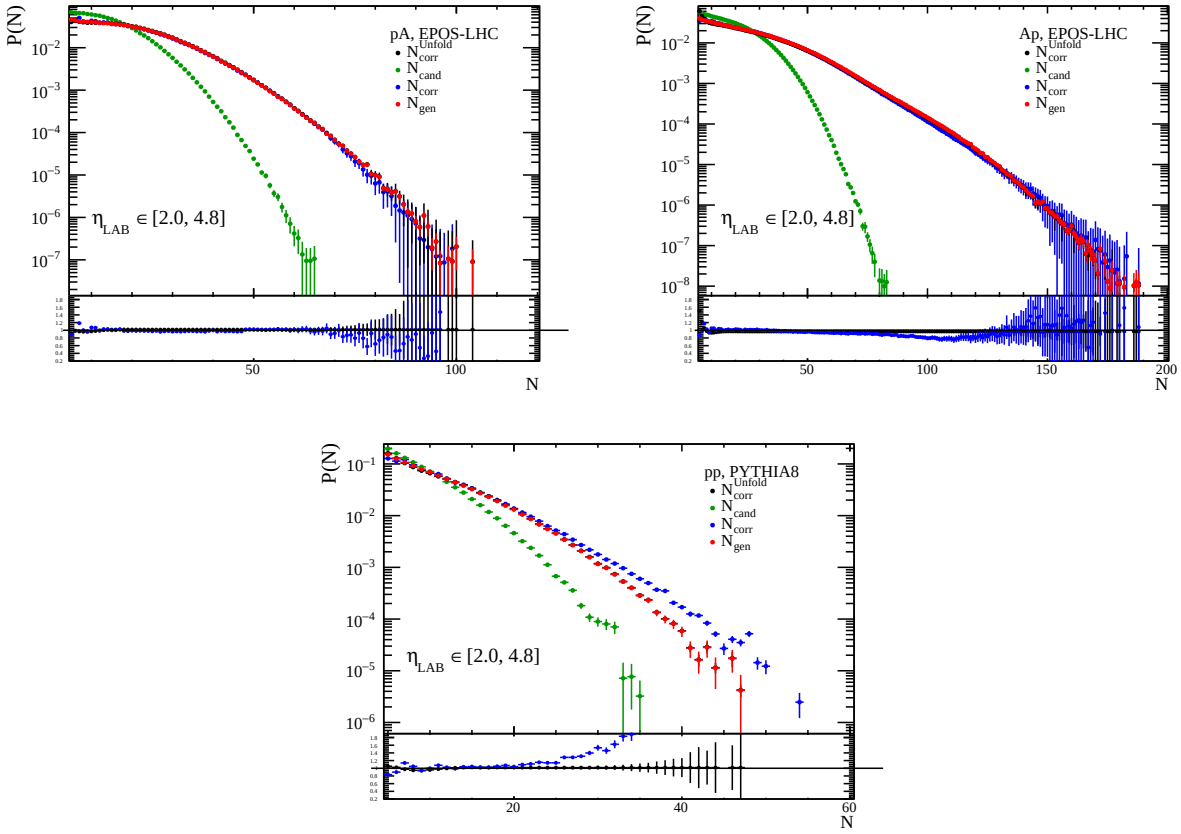


Figure 6.33: Closure test for pPb collisions (left), Pbp collisions (right) and pp collisions (bottom) in the full LHCb acceptance. In green there is the candidates distribution, in blue is the corrected candidates distribution from Eq. 6.3 and in black is the Unfolded corrected candidates distribution. The lower subplot contains the ratio to the generated spectra N_{gen} (red) for the N_{corr}^{Unfold} (black) and N_{corr} (blue) distribution. All distributions are self-normalised to the total number of events.

6.9 Closure test for the average transverse momentum

As we know from the response matrix, the relation between N_{corr} and N_{ch} is not unique, in the sense that tracks in a given N_{corr} class could come from particles originating in several N_{ch} event classes. This means that if we select our events using N_{corr}, our measurement could be biased. This depends on the relation between our multiplicity proxy, N_{corr} in this analysis, and also depends on the trend that our observable shows as a function of multiplicity.[3] To quantify this *event selection bias*, we perform a closure test for $\langle p_{\text{T}} \rangle$. First, a pure simulation-based closure test is performed using N_{ch} for the event selection, and, as can be seen in Fig. 6.34, the agreement is within 2%. In this test, we prove that our corrections remove detector effects up to a 2% precision.

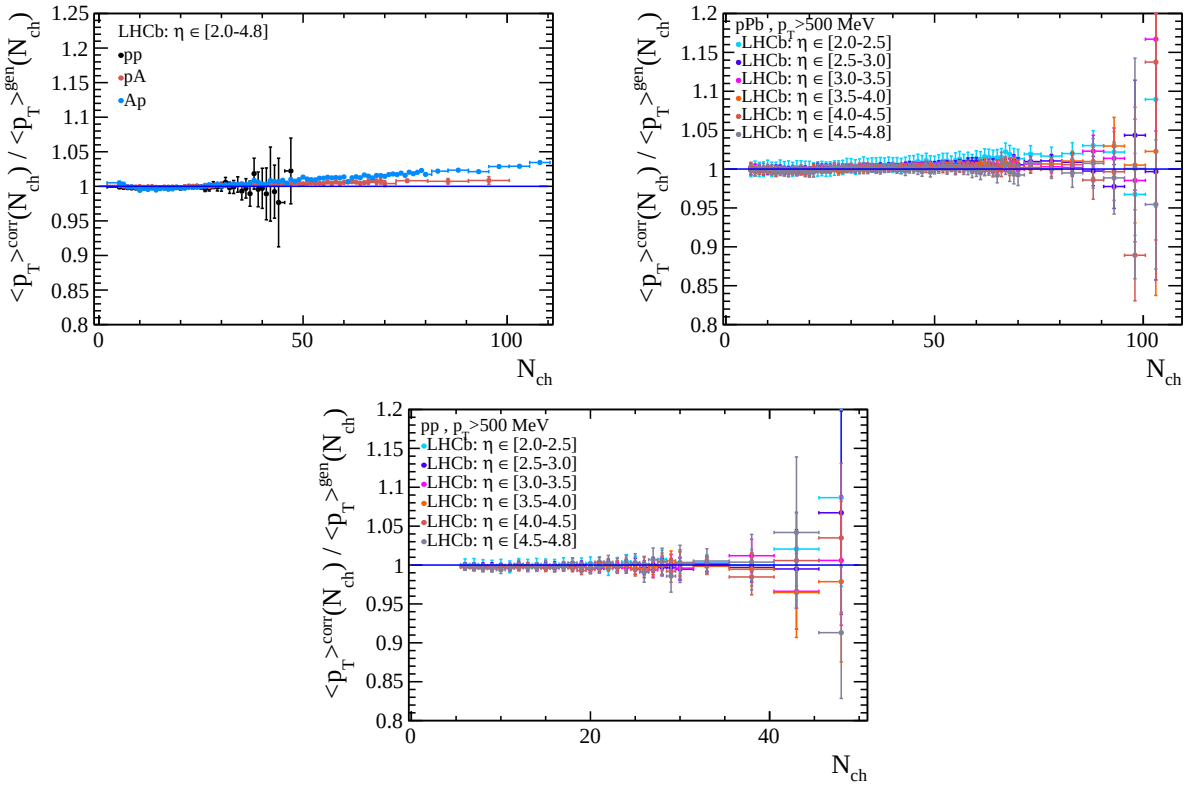


Figure 6.34: Closure test ratio using N_{ch} as event selection variable for pPb and pp simulation data. Plots only show statistical uncertainty. All these plots consider N_{ch} in the full LHCb acceptance.

Then we perform the same test using N_{corr} for the event selection in order to estimate the event-selection bias due to the multiplicity proxy in our measurement. This will be

[3] If the measured observable is constant in multiplicity there is no bias.

analogous to our measurement with data.

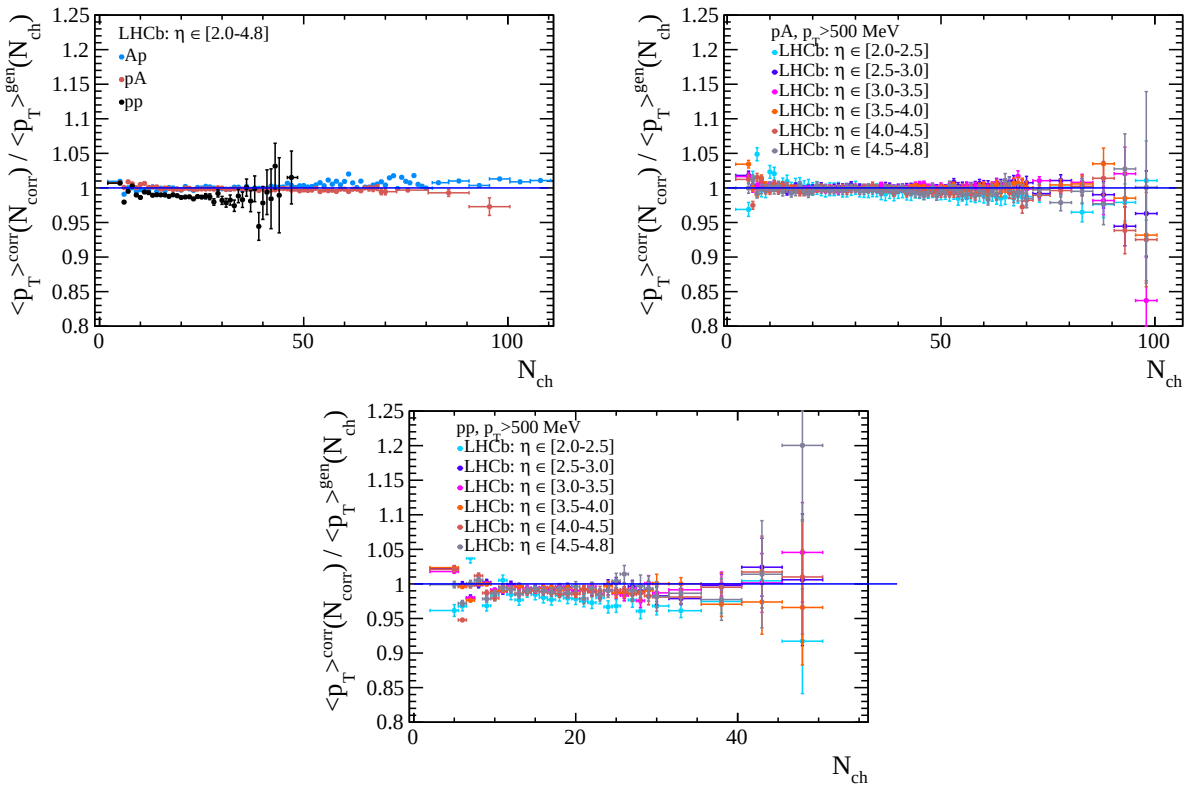


Figure 6.35: Closure test using N_{corr} as events selection for pPb and pp simulation data.

6.10 Initial stage characterisation of hadronic collisions

As was stressed in Chap. 2, one of the main goals of high-energy nuclear collision research is to study the properties of the QGP. In these collisions, a larger transverse overlap of the two incoming ions leads to a larger volume of the created system, resulting in a partial local thermalization with a collective behaviour of the partons. Hence data interpretation of high-energy nuclear collisions requires detailed knowledge of the initial conditions of QGP formation. Moreover, those initial conditions result from the distribution of the collided matter. As a consequence of this, it is crucial to derive from the data the transverse size of the created QGP and its event-by-event fluctuations to be able to interpret experimental observations correctly. In the case of heavy ion (AA) collisions, the impact parameter of the collision (b) is extracted from detector-level (usually calorimeter deposits) information using the Glauber model as explained in [89]. The smaller the impact parameter (the more central the collision) the higher will be the number of charged particles produced. The centrality of the collisions is usually expressed in percentiles of

the total inelastic cross-section. In this sense, 0% indicates a totally central collision ($b = 0\,\text{fm}$) and 100% an ultra-peripheral collision with $b > R_A + R_B$ where R_A and R_B is the radius of the colliding ions.

This can be done in AA collisions because, for a given impact parameter value, the average number of produced charged particles is statistically well-determined. The reason is that the number of nucleon-nucleon collisions is high enough to make event-by-event quantum fluctuations at the sub-nucleonic level negligible (or at least, not the biggest contribution to the total fluctuations). Such internal fluctuations can lead to important changes in the effective collision-by-collision transverse size of the nucleon that are not considered in the Glauber model. For small systems, despite being smaller in the number of average colliding nucleons, sub-nucleonic fluctuations play an important role and make it impossible to use the Glauber model [90], [91]. The number of produced prompt-charged particles, also known as *multiplicity*, is used instead for characterising the initial stage of a hadronic collision involving small systems.

Since the observation of collective effects like long-range correlations [92], [93] or strangeness enhancement [94] in small systems high multiplicity events, multiplicity-dependent analyses are a crucial tool for understanding QCD properties. In the next section, the different methods used in LHCb for measuring multiplicity will be addressed.

6.10.1 Measuring multiplicity at the LHCb experiment

In the last few years, many different multiplicity-dependent analyses have been published in the LHCb collaboration ([95], [96], [97], [98] and [99]). In most of these analyses, multiplicity is estimated using a proxy variable based on the number of tracks reconstructed by the LHCb tracking system. As explained in Chap. 3, at LHCb many different track types are reconstructed. For multiplicity estimation the most used ones are VeloTracks. The caveat of this approach is the difficulty when comparing with other experiment measurements and Monte Carlo-based or phenomenological models. Moreover, the relation between the proxy and the target quantity or other proxies, the number of prompt-charged particles, has never been published.

In this analysis, a different approach has been followed. It is based on measuring the number of prompt-charged particles produced in every event. Moreover, it implements data-driven corrections whenever possible.

As different proxies have different ranges, in order to compare proxies between themselves and with Monte Carlo level multiplicity a self-normalization is usually done. While providing a result for example in a range of $0 - 200$ VeloTracks is difficult to interpret for people outside the experimental collaboration, self-normalised proxies give a better sense of the multiplicity dependence. However, as can be seen in Fig. 6.36 and Fig. 6.37 different proxies have different relations with generator level multiplicity (N_{ch}). This means that **a collision with 2 times the average number of VeloTracks does not correspond to a collision with 2 times the average number of prompt charged particles.** Because of this reason, the interpretation of data is not trivial and Fig. 6.36 is needed

to relate proxies between each other and with generator level multiplicity. Moreover, the non-linear behaviour of the proxy with the generator level multiplicity makes the relation harder to use in a clear way. Something worth mentioning is that Fig. 6.36 is Monte Carlo-based, which means that detector effects not well reproduced by detector simulation could potentially modify conclusions taken from the figure.

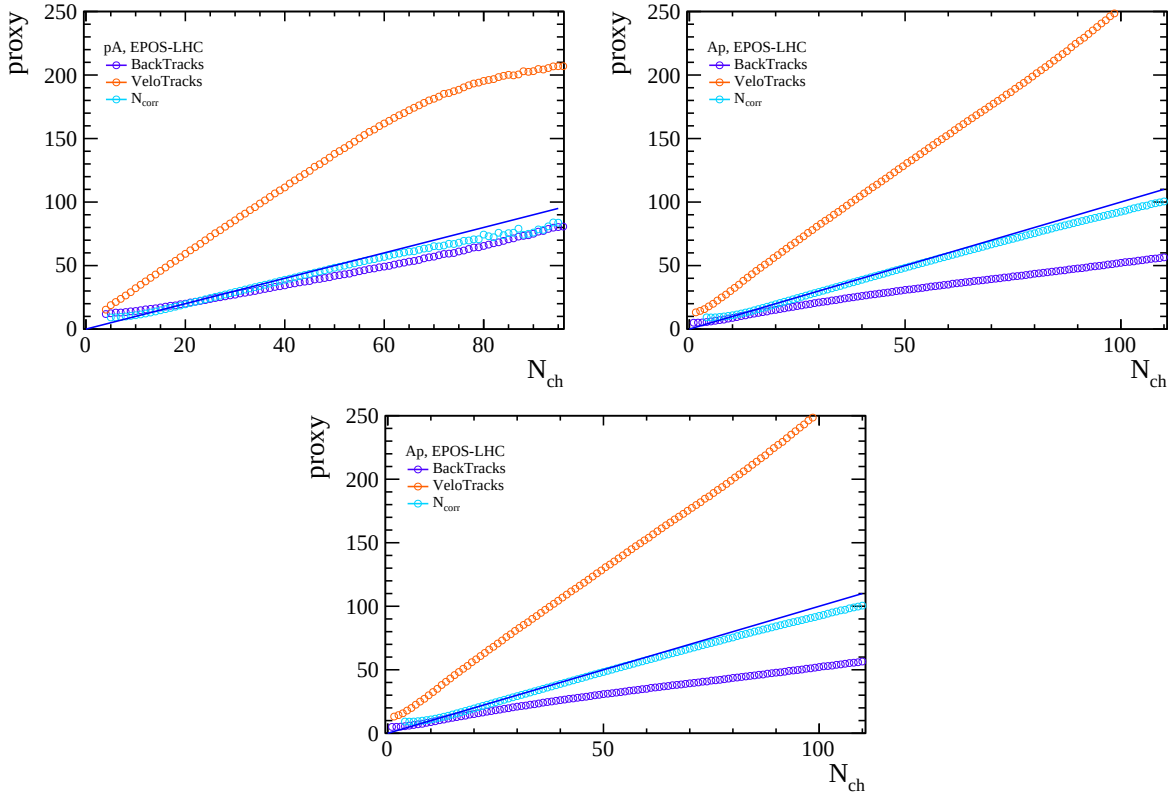


Figure 6.36: Multiplicity proxies as a function of the self-normalised multiplicity at the generator level. The blue line is just a straight line with a slope equal to unity to draw the eye. The left plot corresponds to pPb collisions, the right to Pbp and the bottom for pp.

The multiplicity estimator used in this work, N_{corr}, is very correlated on average to the generator level multiplicity as can be seen in Fig. 6.36 and Fig. 6.37 for all collision systems. Using this variable, the interpretation of the results is easier, and in this case, **the self-normalised multiplicity estimated from the proxy is equal to the target one**. Moreover, as the prompt charged particle definition is the one commonly used in the field [81], it is clearer what we are measuring. It is worth mentioning that the linear relation between N_{corr} and N_{ch} stops when the saturation of the proxy is observed. The hypothesis is that the lack of statistics in this region requires wider bins in a region where efficiencies drop very quickly. More statistics in the Monte Carlo for this region would be needed to solve this.

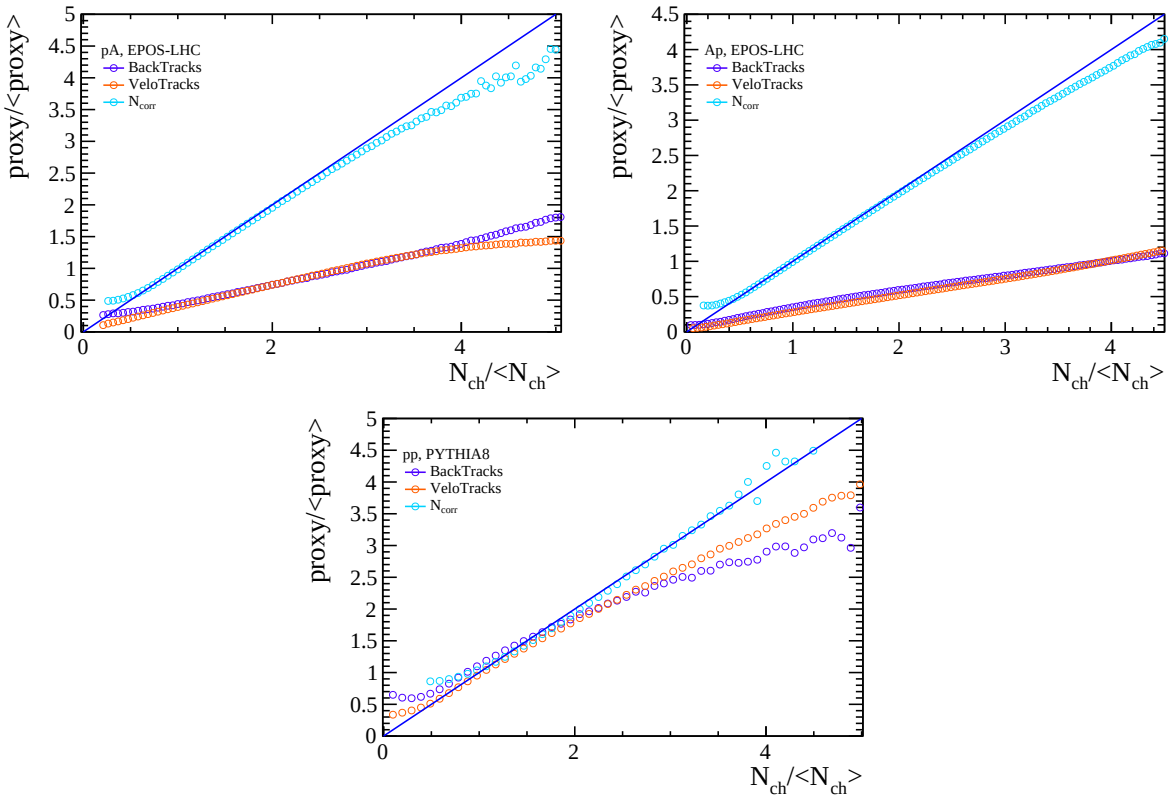


Figure 6.37: Self-normalised multiplicity proxies as a function of the self-normalised multiplicity at the generator level. The blue line is just a straight line with a slope equal to unity to draw the eye. The blue line is just a straight line with a slope equal to unity to draw the eye. The left plot corresponds to pPb collisions, the right to Pbp and the bottom for pp.

However, Fig. 6.36 corresponds to event-averaged values of both, proxy and generator multiplicity. It is worth remembering that there is some spread around the points drawn in Fig. 6.36. This is what we can see in response matrices at Chap. 6. This spreading can be understood by realising that a sharp cut in any multiplicity proxy gives a distribution in N_{ch} as can be seen in Fig. 6.38 and vice-versa. Moreover, distributions from different N_{ch} bins tend to overlap.

This effect produces a bias in our multiplicity-dependent measurement since the multiplicity distribution of our observable is not constant. In order to quantify this effect in the average p_{T} measurement the closure test is used. Here, by changing the event selection between generator level multiplicity and our proxy it is possible to estimate the impact of the multiplicity proxy bias in our measurement. This effect would occur for any multiplicity proxy and the impact of the measurement depends on the measured quantity.

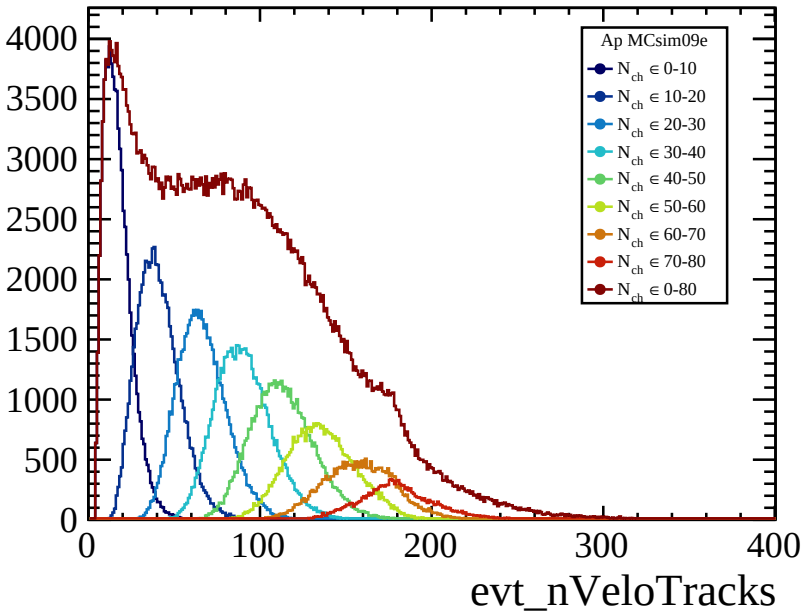


Figure 6.38: Velo track distribution for each N_{ch} cut from EPOS-LHC simulation.

7

Systematic uncertainties

This section will address the estimation of systematic uncertainties in the analysis. First, uncertainties affecting the corrections computed in Chap. 6 will be explained. That includes ε_{reco} and Π, while ε_{sel} only includes statistical uncertainty from simulation. Then, specific uncertainties for each observable will be addressed.

7.1 Reconstruction efficiency

Three uncertainty sources are considered in the reconstruction efficiency calculation:

- Tracking correction factors, $C_{tracking}(p, \eta)$.
- Particle composition of the simulation sample.
- Occupancy weights: According to [70] where a similar approach is followed is considered to a 1%. This is included together with the uncertainty of tracking correction factors.

The final result of the relative systematic uncertainty in the reconstruction efficiency can be seen at Fig. 7.1, Fig. 7.2 and Fig. 7.3 for different collision systems. In this section, the calculation of the different contributions will be detailed.

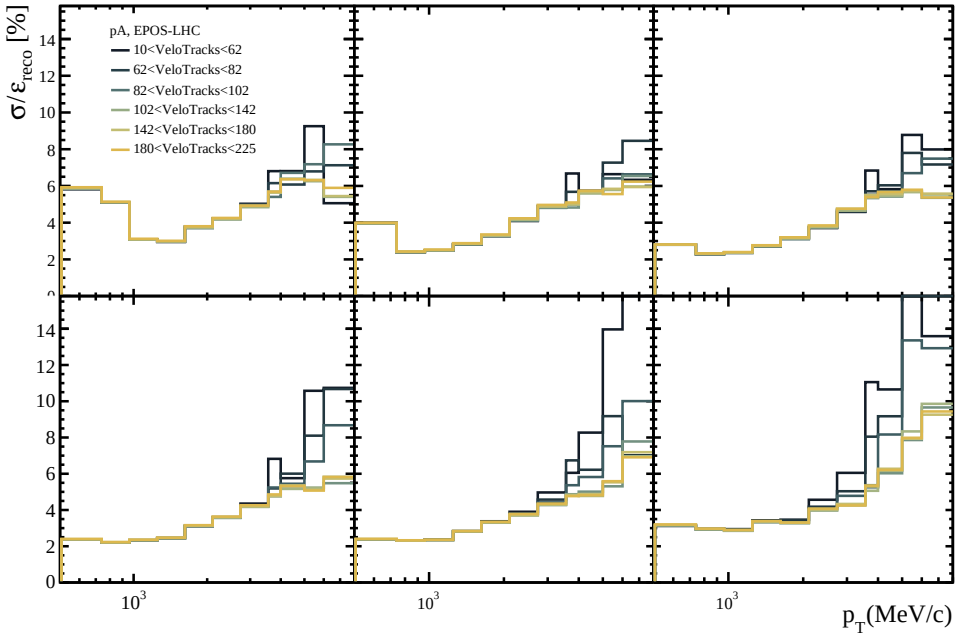


Figure 7.1: Total relative uncertainty in the reconstruction efficiency in pPb collisions.

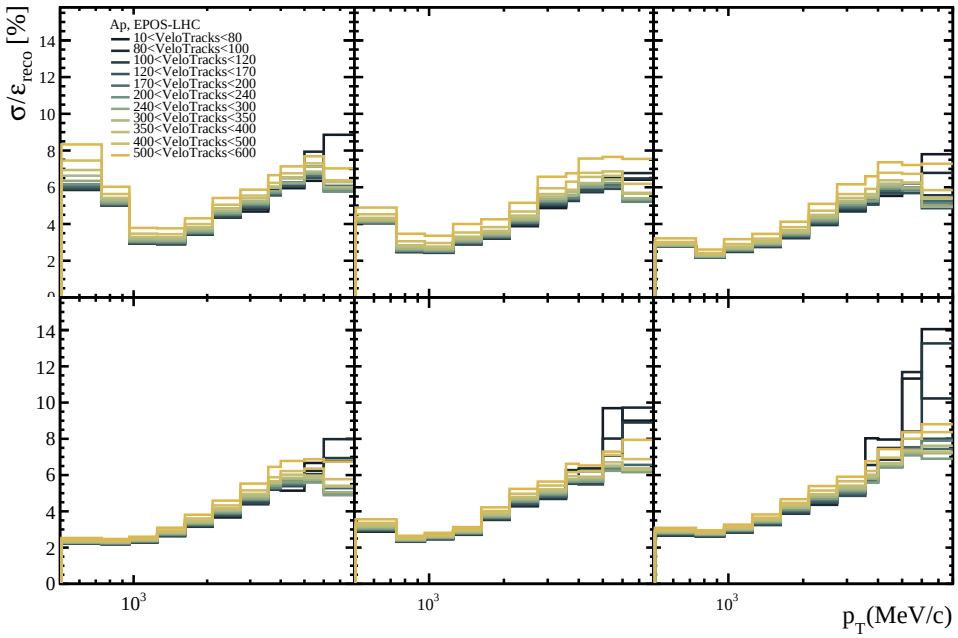


Figure 7.2: Total relative uncertainty in the reconstruction efficiency in Pbp collisions.

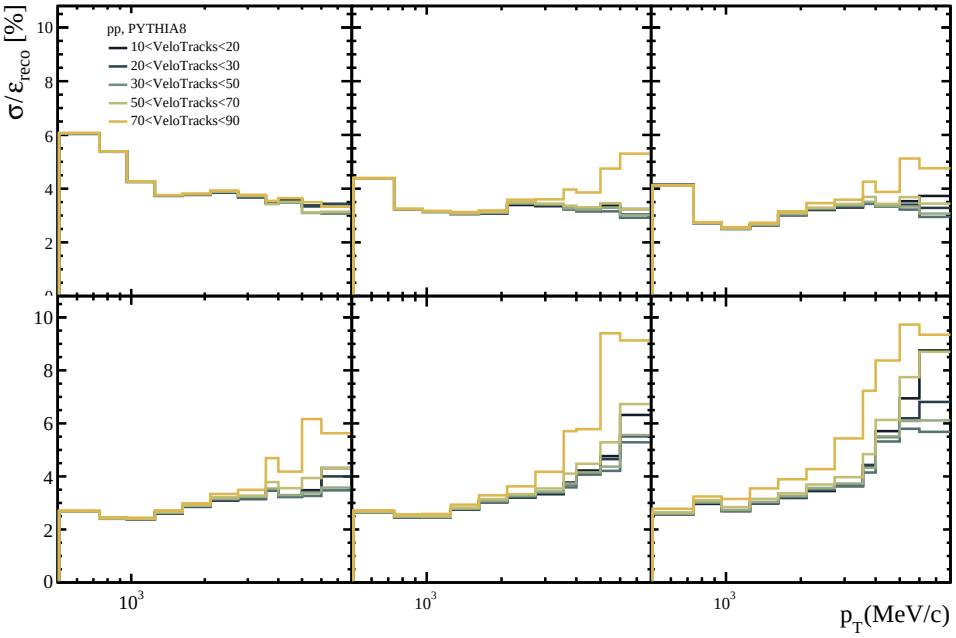


Figure 7.3: Total relative uncertainty in the reconstruction efficiency in pp collisions.

7.1.1 Tracking correction

The correction factors to take into account data/MC discrepancies in the tracking efficiencies have some associated uncertainties. Each track **within the momentum range** $5 < p < 200$ GeV/c, has two uncertainties depending on its (p, η).

- The one shown in Tabs. 6.2, 6.3.
- A global systematic uncertainty of 0.4% (0.8%) for the proton-nucleus (pp) data calibrated with the 2012 (2015) sample.

For tracks **outside the momentum range** of Tabs. 6.2, 6.3, a conservative 5% systematic uncertainty is assigned to take into account the loss of knowledge on the tracking correction for these tracks. Besides, for tracks within and outside the ranges in Tabs. 6.2, 6.3, a 1.4% systematic uncertainty originating from the description of hadronic interactions with the detector material in simulation is added. The propagation of this uncertainty in the calculation of the reconstruction efficiency is done by varying $C_{tracking}(p, \eta)$ by $\pm 1\sigma$, where σ is the quadratic sum of all corresponding uncertainties explained above. The systematic uncertainty on ε_{reco} will be half the difference between these two values.

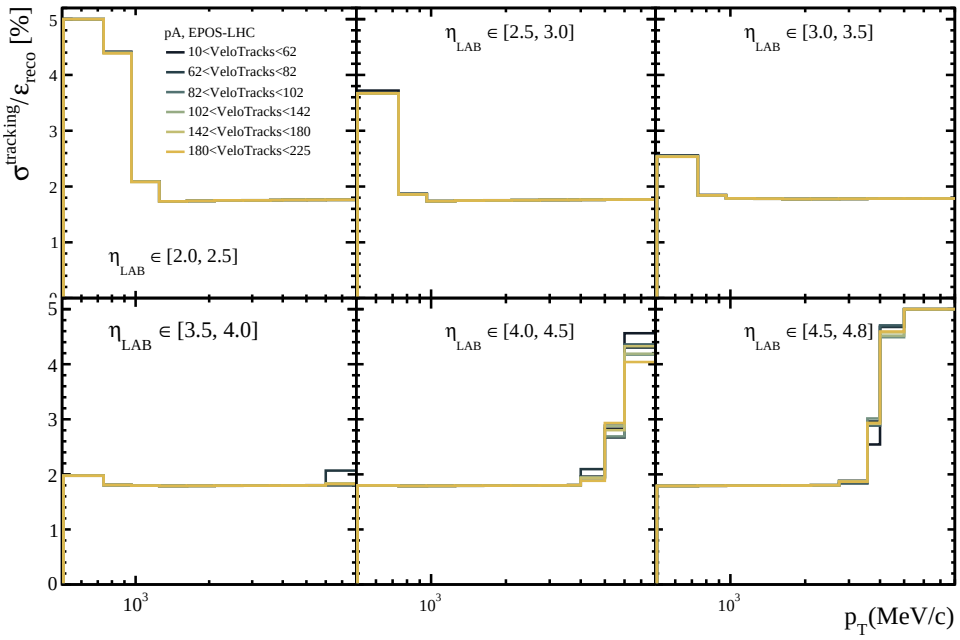


Figure 7.4: Tracking systematic uncertainty and statistical uncertainty in the reconstruction efficiency at pPb collisions.

7.1.2 Particle composition

In Sec. 6.5.2 it was detailed how the species fraction has an impact on the reconstruction efficiency computation from the MC. In that section was proved that EPOS gives a reasonable description of h/π fraction data within a 30%. For this reason, a systematic uncertainty that takes into account a 30% variation on the particle species fraction has to be taken into account. The procedure to compute this uncertainty in the reconstruction efficiency is detailed:

1. MC toy generation: Following a Gaussian distribution with $\mu = N_p^{sim}, \sigma = 0.3 * N_p^{sim}$ a N_p^{toy} value it is generated. This is done 10000 times.

2. For each instance the total number of particles, N_T, and f_p is computed.

3. For each instance, a ε_{reco}^{toy} it is obtained. Finally, this distribution is fitted to a Gaussian distribution where the average value should recover the original ε_{reco}. The systematic uncertainty coming from the particle composition in the reconstruction efficiency will be the width of that distribution.

In Fig. B.33 we can see some examples of the fits to the distribution of ε_{reco} from each toy. The values of this systematic uncertainty go from 1% to 4% as can be seen in Fig. 7.7. The increasing trend as a function of p_{T} is the abundance of K, p and hyperons.

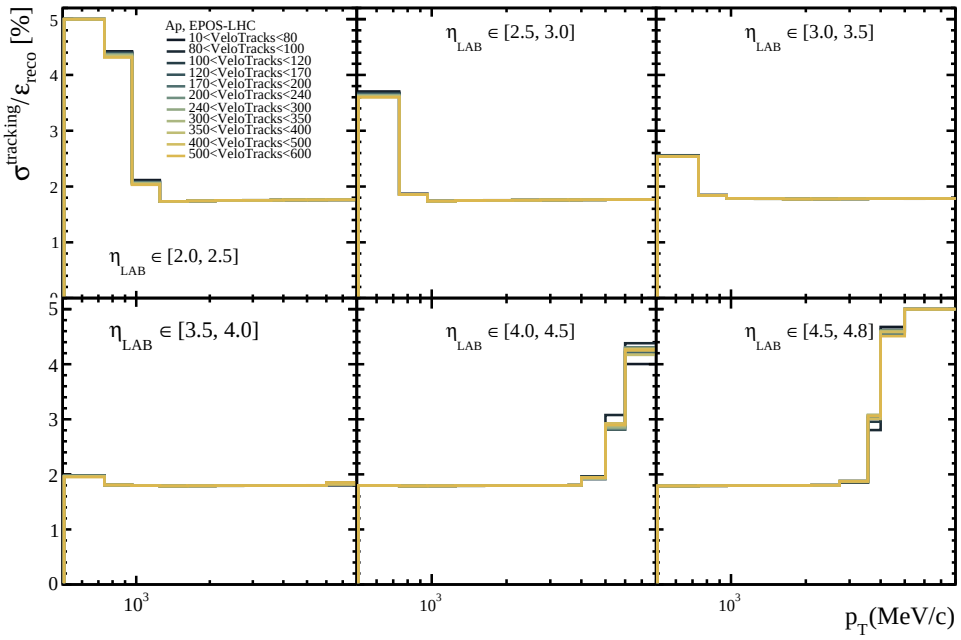


Figure 7.5: Tracking systematic uncertainty and statistical uncertainty in the reconstruction efficiency at Pbp collisions.

Its abundance with respect to pions increases at high-p_{T}, thus the efficiency becomes more sensitive to f_p fluctuations.

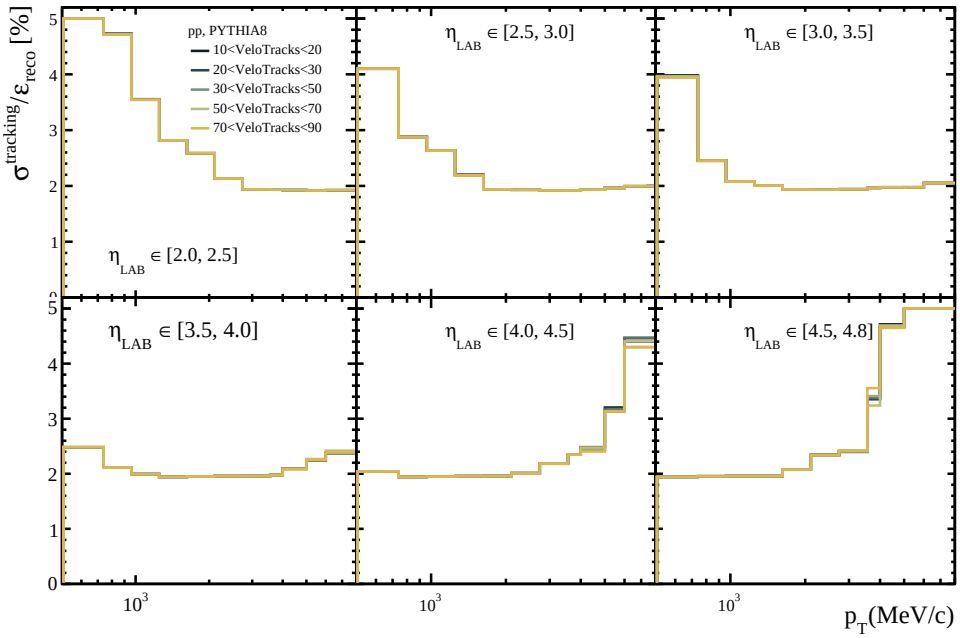


Figure 7.6: Tracking systematic uncertainty and statistical uncertainty in the reconstruction efficiency at *pp* collisions.

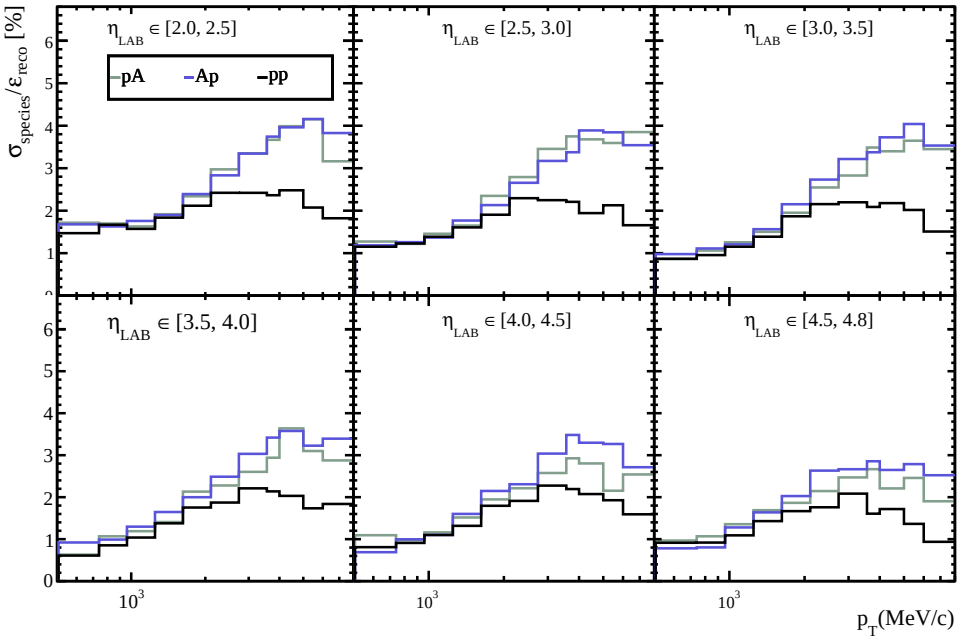


Figure 7.7: Relative systematic uncertainty in the reconstruction efficiency due to the particle composition.

7.2 Background subtraction

This section will detail the uncertainty arising from the purity computation. The uncertainty comes from the data-driven correction factors in R_α. The strategy to obtain the systematic uncertainty for each background source is similar. However, despite some particularities, each background source uncertainty will be explained independently.

7.2.1 Fake tracks

The systematic uncertainty arising from the fake-tracks fraction comes from the data-driven correction factor, R_{fakes}. This quantity has three uncertainty sources:

- Proxy purity
- Statistical uncertainty
- Proxy definition

The proxy purity contribution to the systematic uncertainty is estimated as

$$s^{purity} = (1 - \Pi^{sim}_{fakes})R_{fakes}, \tag{7.1}$$

where Π^{sim}_{fakes} is the purity of the proxy computed in the simulation.

When defining the proxy to compute R_{fakes} several definitions can be used. To compute the systematic uncertainty, an alternative and orthogonal proxy definition is used to compute R^{alter}_{fakes}. The selection cuts used to define the alternative proxy are in Tab. 7.1.

	pp	pA	Ap
Track type	Long	Long	Long
η	$2 < \eta < 4.8$	$2 < \eta < 4.8$	$2 < \eta < 4.8$
p_{T}	$0.5 < p_{\mathrm{T}} < 8$ GeV/c	$0.5 < p_{\mathrm{T}} < 8$ GeV/c	$0.5 < p_{\mathrm{T}} < 8$ GeV/c
p	$p > 2$ GeV	$p > 2$ GeV	$p > 2$ GeV
IP	< 0.508 mm	< 0.508 mm	< 0.568 mm
`GhostP`	$0.375 <$ `GhostP` < 0.4	$0.1 <$ `GhostP` < 0.4	$0.1 <$ `GhostP` < 0.4
IsClone	False	True	True

Table 7.1: Selection cuts used to define the alternative proxy for the R_{fakes} systematic computation.

The same method used to compute the nominal value of R_{fakes} has been followed, and the statistical and systematic uncertainty (due to the proxy purity) is computed.

The systematic uncertainty due to the alternative proxy definition is computed as the weighted standard deviation,

$$s_w = \sqrt{\frac{\sum_{i=1}^{N} w_i (x_i - \overline{x}_w)^2}{\frac{(N'-1)\sum_{i=1}^{N} w_i}{N'}}}, \tag{7.2}$$

where w_i are the weights, in this case $w_i = \frac{1}{\sigma(R^i_{fakes})}$ with $i = alter, nom$. N' is the number of non-zero weights ($N' = 2$ in this case), x_i will be R^{nom}_{fakes} and R^{alter}_{fakes} and $\overline{x}_w$ is the weighted average value between R^{nom}_{fakes} and R^{alter}_{fakes} defined as $\overline{x}_w = \frac{w_{nom} R^{nom}_{fakes} + w_{alter} R^{alter}_{fakes}}{w_{nom} + w_{alter}}$. The relative uncertainty $\sigma(R_{fakes})/R_{fakes}$ is shown in Figs. B.36, B.37.

Then all the R_{fakes} uncertainty contributions are added in quadrature:

$$s(R_{fakes}) = \sqrt{(s_{stat})^2 + (s_{purity})^2 + (s_{alter})^2}. \tag{7.3}$$

Finally, the uncertainty is propagated to the fraction of fakes using uncertainty propagation through the formula

$$\sigma_{f_{fakes}} = \sqrt{(\sigma_{R_{fakes}} f^{sim}_{fakes})^2 + (R_{fakes} \sigma^{sim}_{f_{fakes}})^2}. \tag{7.4}$$

The results for the systematic uncertainty of the R_{fakes} are displayed in Figs. B.38, B.39, B.40.

7.2.2 Secondary particles

The secondary particles background was already discussed in Sec. 6.7.2 where the factorisation was defined as

$$f_{sec} = R_\gamma f^{sim}_\gamma + R_{Had\&Decays} f^{sim}_{Had\&Decays} + R_\mu f^{sim}_\mu. \tag{7.5}$$

The uncertainty of f_{sec} comes from the uncertainty related to the data-driven correction factors, R_α. This uncertainty has two components, statistical and systematic, which are summed in quadrature as

$$u_{R_\alpha} = \sqrt{u^2_{stat} + u^2_{syst}}, \tag{7.6}$$

where α corresponds to each one of the secondary particle sources. The statistical uncertainty accounts for the limited size of the proxy sample, while the systematic uncertainty comes from the purity of the proxy sample. The latter is computed according to

$$u^{syst}_\alpha = (1 - \Pi_\alpha) R_\alpha. \tag{7.7}$$

The statistical uncertainty is propagated from $N^{sim;data}_{proxy}$ to the R-value for the case of γ conversion and μ decays background components, while for the $Had\&Decays$ component we use Eq. 7.8.

$$u^2_{stat} = (u^{sim}_{stat})^2 + (u^{data}_{stat})^2. \tag{7.8}$$

The result for the uncertainty of each contribution of the secondary particles background can be seen in App. B.7. Finally, the uncertainty in the purity can be seen in Figs. 7.8, 7.9.

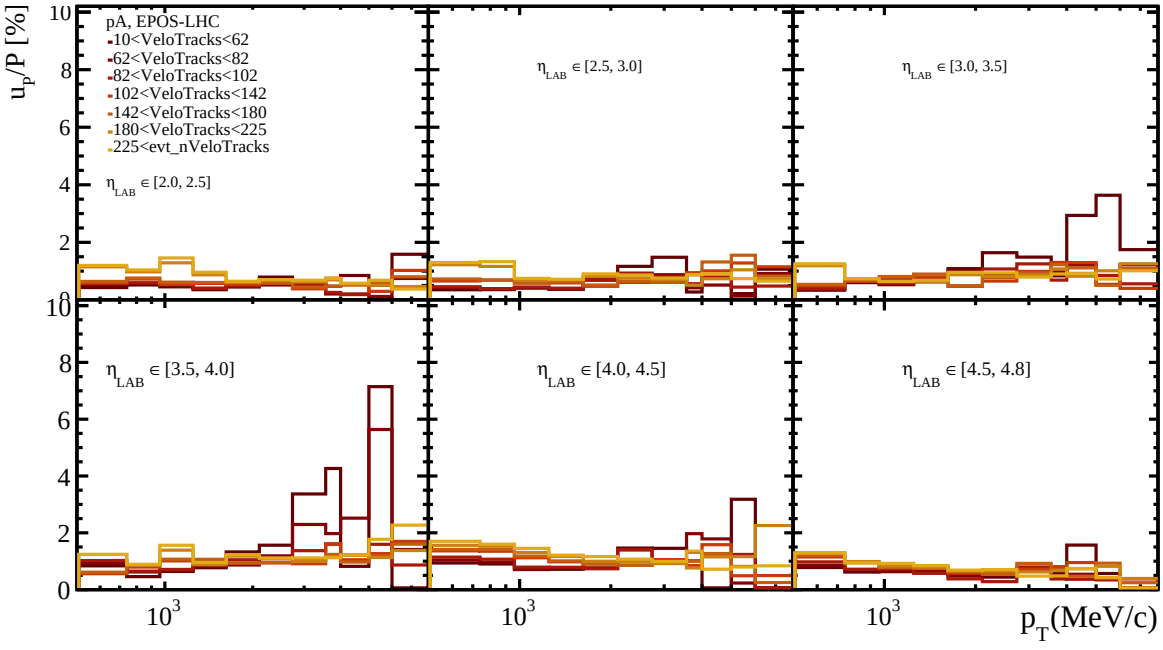


Figure 7.8: Total relative uncertainty for purity at pPb collisions.

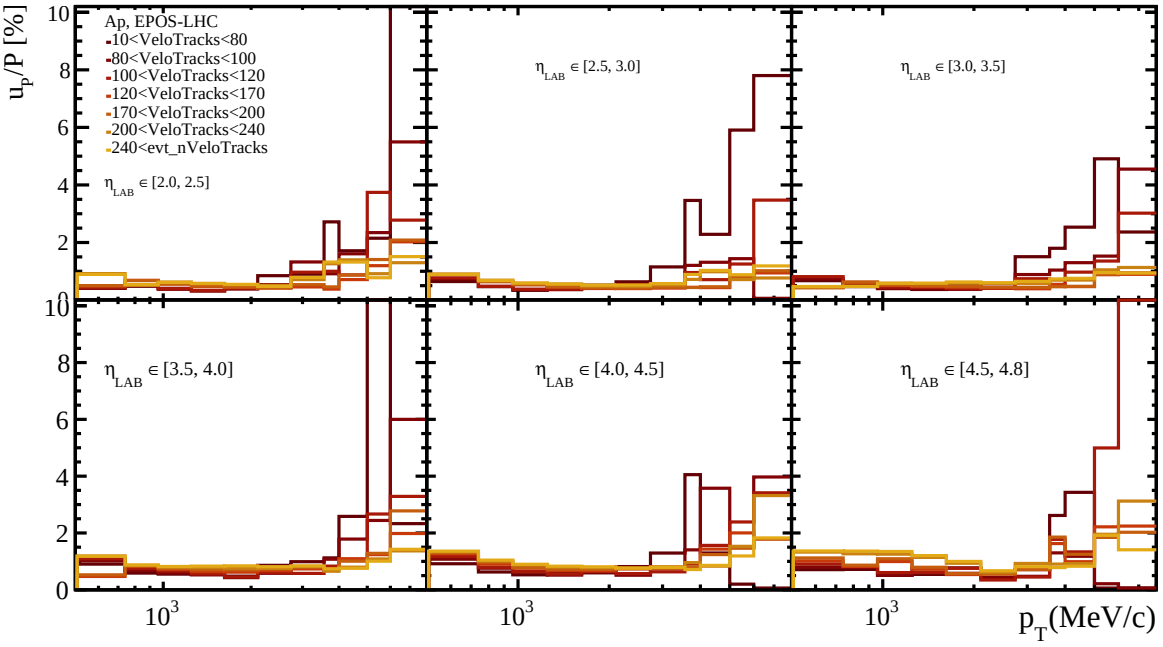


Figure 7.9: Total relative uncertainty for purity at Pbp collisions.

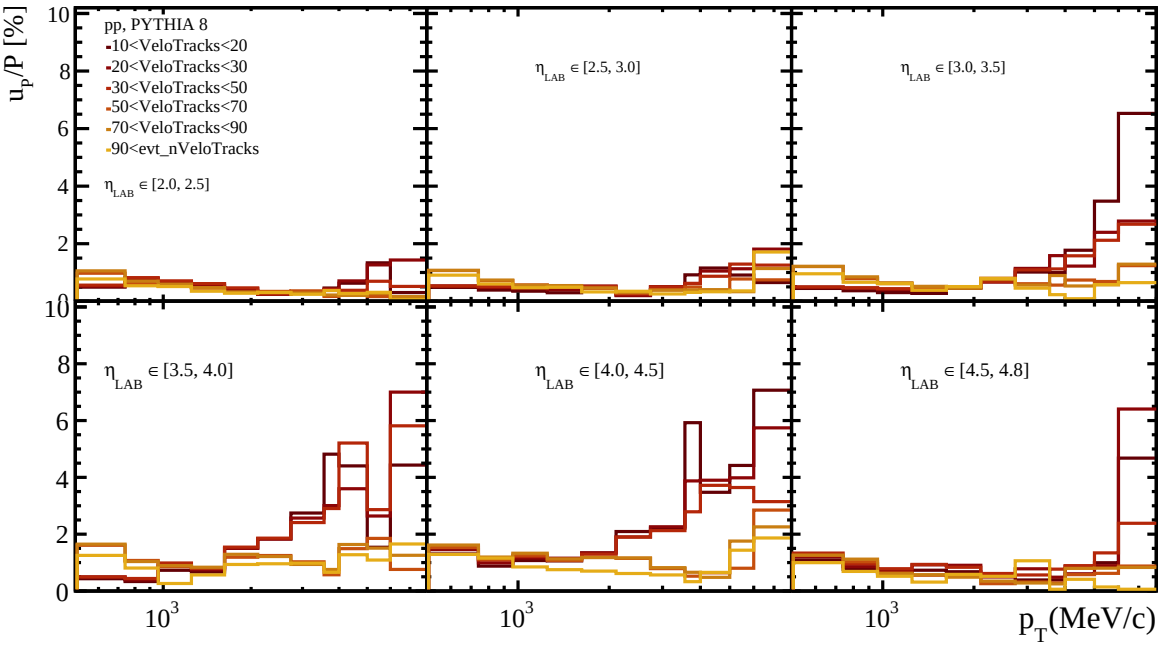


Figure 7.10: Total relative uncertainty for purity at pp collisions.

7.3 Multiplicity distribution: Uncertainty propagation and specific uncertainty sources

The multiplicity spectrum is computed in two steps. First, the N_{corr} (evt) quantity is computed event-by-event using Eq. 6.3. Then, the $P(N_{\text{corr}})$ distribution is unfolded. As a consequence, four uncertainty sources are considered in this measurement, namely :

- Reconstruction efficiency, σ_{reco}
- Selection efficiency, σ_{sel}
- Purity, σ_{Π}
- Statistical, σ_{stat}

First, we generate a random value of each correction following a Gaussian distribution with parameters $\mu = C(evt)$ and $\sigma = \sigma_C(evt)$. We compute an alternative N_{corr} following Eq. 6.3 for each event, as is shown in Fig. 7.11, where C stands for ε_{reco}, P and ε_{sel}.

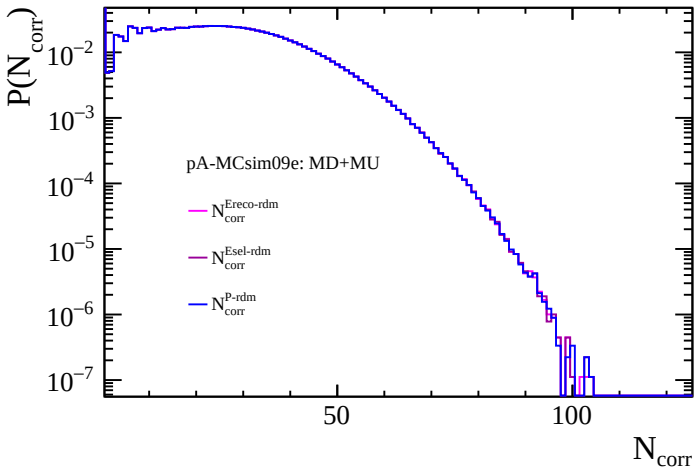


Figure 7.11: Alternative $P(N_{\text{corr}})$ from the random variation of corrections values.

Then, the systematic uncertainty in the $P(N_{\text{corr}})$ distribution from the corrections is computed as half of the difference between the nominal distribution and the alternative ones using Eq. 7.9. The nominal distribution is named $P(N_{\text{corr}})$ and the alternative ones are $P^{C-rdm}(N_{corr})$ where C are the corrections.

$$\sigma_{syst}(C) = \frac{|P(N_{\text{corr}}) - P^{rdm-C}(N_{\text{corr}})|}{2}. \tag{7.9}$$

Finally, the total uncertainty in $P(N_{\text{corr}})$ is computed by adding in quadrature all the sources as is detailed in Eq. 7.10. The statistical uncertainty corresponds to $\sigma_{stat} =$

$\sqrt{\frac{dN_{evt}}{dN_{\rm ch}}}$. The relative systematic uncertainty is shown in Fig. 7.12 together with the total uncertainty, which includes the statistical uncertainty added in quadrature:

$$\sigma(P(N_{\rm corr})) = \sqrt{\sigma^2_{stat} + \sigma^2_{syst}(\varepsilon_{reco}) + \sigma^2_{syst}(\varepsilon_{sel}) + \sigma^2_{syst}(\varepsilon_{\Pi})}. \tag{7.10}$$

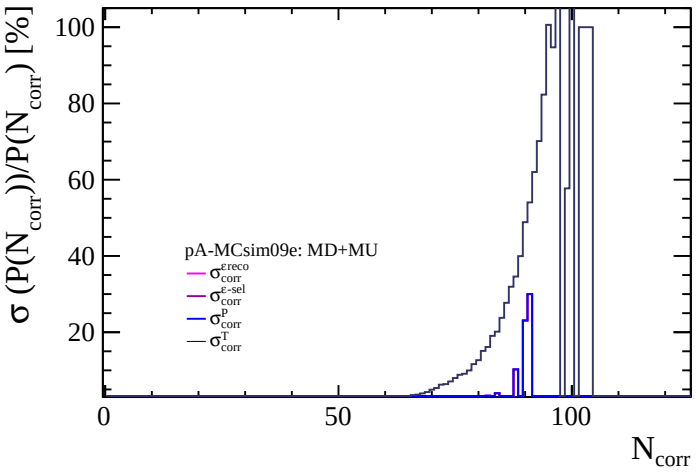


Figure 7.12: Alternative $P(N_{\rm corr})$ from the random variation of corrections values.

The uncertainty propagation of $P(N_{\rm corr})$ uncertainties through the Unfolding procedure is done via the RooUnfold package. According to the documentation [1] uncertainties of the unfolded distribution are estimated with toys using the **IncludeSystematics()** option. It has been checked that the errors of the unfolded distribution are similar to the ones in the input distribution $P(N_{\rm corr})$.

7.4 Average transverse momentum: Uncertainty propagation

The propagation of corrections uncertainties to the $\langle p_{\rm T}\rangle$ result is done in two steps. First, the uncertainty is propagated from the corrections to the $p_{\rm T}$ spectra using the uncertainty propagation on Eq. 6.6 for each $p_{\rm T}$ bin,

$$\sigma_{N_{\rm corr}} = \sqrt{\left(\frac{N_{cand}}{\varepsilon_{reco}\varepsilon_{sel}}\sigma_P\right)^2 + \left(\frac{P}{\varepsilon_{reco}\varepsilon_{sel}}\sigma_{N_{cand}}\right)^2 + \left(\frac{N_{cand}P}{\varepsilon_{reco}\varepsilon_{sel}}\cdot\frac{\sigma_{\varepsilon_{reco}}}{\varepsilon_{reco}}\right)^2 + \left(\frac{N_{cand}P}{\varepsilon_{reco}\varepsilon_{sel}}\cdot\frac{\sigma_{\varepsilon_{sel}}}{\varepsilon_{sel}}\right)^2}. \tag{7.11}$$

[1] https://gitlab.cern.ch/RooUnfold/documentation/-/blob/master/RooUnfold_user_guide.pdf

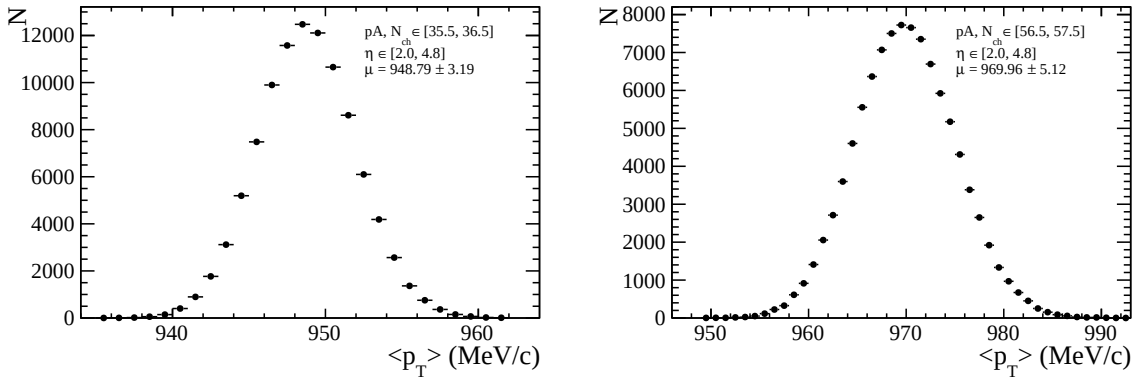


Figure 7.13: Two distributions from the $\langle p_{\mathrm{T}} \rangle$ uncertainty propagation using MC Toys.

The relative uncertainty of the p_{T} distribution can be seen at Fig. B.48, Fig. B.49, Fig. B.50 and Fig. B.47. Secondly, the uncertainty is propagated to the $\langle p_{\mathrm{T}} \rangle$ by using an MC Toy strategy. For each bin of the p_{T} distribution a random value of N_{corr} is generated following a Gaussian distribution with a mean equal to N_{corr} and standard deviation equal to the systematic uncertainty computed using Eq. 7.11. The $\langle p_{\mathrm{T}} \rangle$ of the new distribution is computed. This process is performed 100000 times for each (N_{ch}, η) bin. After all the iterations are done, a distribution of $\langle p_{\mathrm{T}} \rangle$ values is obtained as the example shown in Fig. 7.13. The standard deviation of the $\langle p_{\mathrm{T}} \rangle$ distribution is taken as the uncertainty.

As can be seen in Fig. 7.14 the relative uncertainty in the $\langle p_{\mathrm{T}} \rangle$ measurement in the full LHCb acceptance is smaller than 2%.

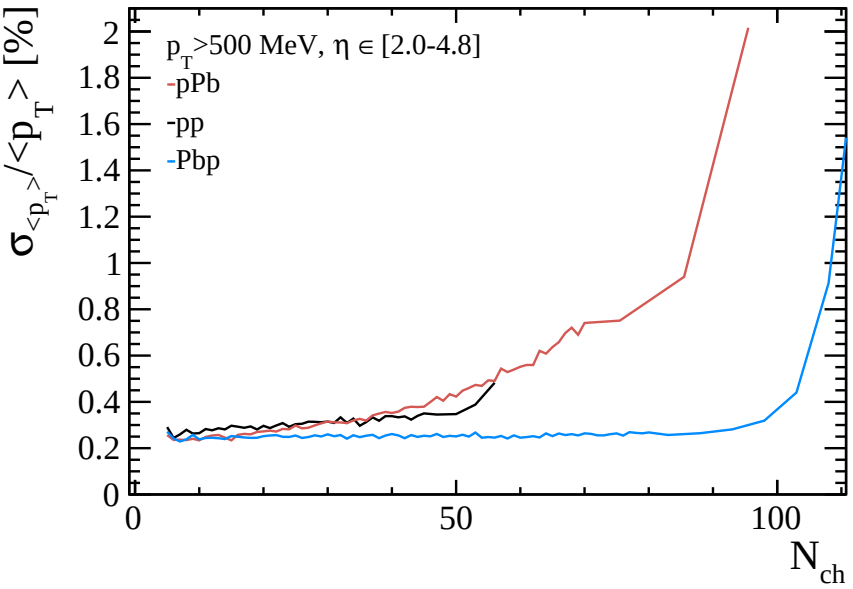


Figure 7.14: Relative uncertainty of the $\langle p_{\mathrm{T}} \rangle$ measurement in all the LHCb acceptance.

8

Results and discussion

In this section, the final results of the thesis are revealed. First, in Sec. 8.1 the results for the multiplicity distribution are shown. In that section both the pseudorapidity integrated result and the pseudorapidity dependent are shown. Secondly, in Sec. 8.2 the results from the integrated and pseudorapidity dependent $\langle p_{\mathrm{T}} \rangle$ measurement are shown. In both sections the results are compared with models and the agreement with other previous measurements is discussed.

8.1 Multiplicity distributions

The final result of the multiplicity distribution for all collision systems in the entire LHCb acceptance is shown in Fig. 8.1. In Fig. 8.4 the pseudorapidity-dependent results can be seen.

The multiplicity distribution results are compared with MC generators in Fig. 8.2 for the integrated η case in all collision systems. The multiplicity range of the comparison with simulation is constrained due to the lack of statistics in the most high multiplicity region. In general, PYTHIA shows more agreement with pp data than EPOS-LHC does with the pPb and Pbp data. Concerning the pPb result, EPOS-LHC underestimates high multiplicity events by a large amount in a systematic way. When observing pp, the PYTHIA generator describes reasonably well the observed trend. However, a 10% difference can be seen at intermediate multiplicities in Fig. 8.2. Finally, for Pbp collisions, the disagreement is similar to the one in pPb collisions except at high multiplicities.

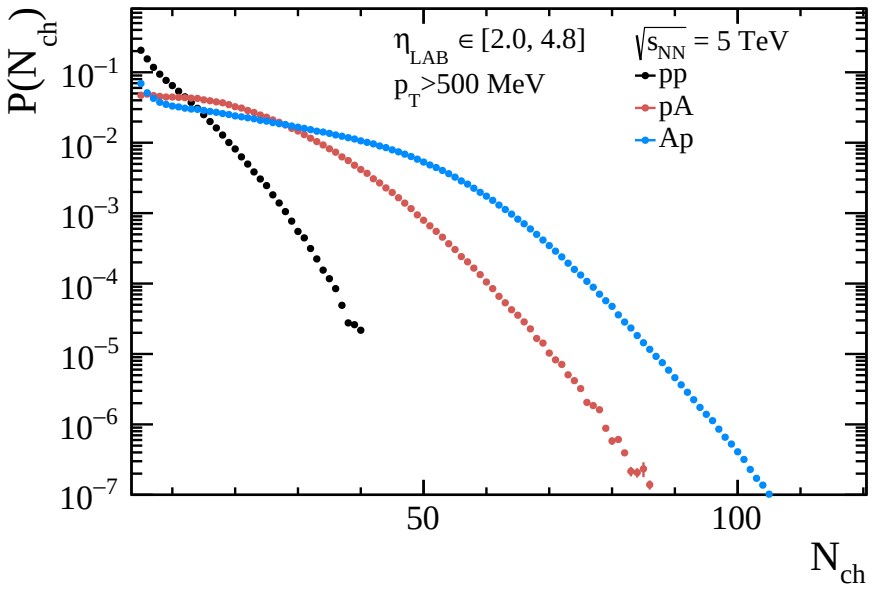


Figure 8.1: Multiplicity distribution result for the three collision systems in the full LHCb acceptance.

The model/data ratio shown in Fig. 8.2 can be compared with the same result from Ref. [5]. The pPb result shows a similar behaviour; i.e. a 20% overestimation from EPOS-LHC at $N_{ch} \sim 30$ followed by a systematic underestimation that grows with multiplicity. Regarding the Pbp result the model/data ratio shows a similar behaviour than for pPb at low and intermediate multiplicity. However, in the high multiplicity range, the underestimation is not so strong as for pPb. For pp, the agreement with simulation is better. However, as seen in Ref. [5] PYTHIA overestimates the multiplicity spectra by 10% in the intermediate multiplicity region and underestimates the result in the high multiplicity region.

The $P(N_{ch})$ results in pseudorapidity bins in Fig. 8.4 can be qualitatively compared with the $dN/d\eta$ measurements from Ref. [100] shown in Fig. 8.3. As seen in the ALICE measurement the pseudorapidity dependence of $dN/d\eta$ is stronger for pPb collisions than for pp collisions. This effect is also visible in Fig. 8.4. Moreover, when comparing the forward (positive η) with the backwards (negative η) in ALICE measurement a stronger trend is observed in the forward region. In the backward region, there is a *plateau* in the pseudorapidity range $\eta_{LAB} \simeq [2.0, 3.5]$ where the $dN/d\eta$ measurement does not change with η. This agrees with our result where the $P(N_{ch})$ distributions in the most central bins of the Pbp result reveal a small difference. As seen in Fig. 8.4 the pseudorapidity dependence of the pPb result (positive η in the ALICE measurement) is stronger than the Pbp case. It is worth saying that the $dN/d\eta$ value from our $P(N_{ch})$ result cannot be directly compared with the ALICE measurement because of the different N_{ch}^{min} requirement.

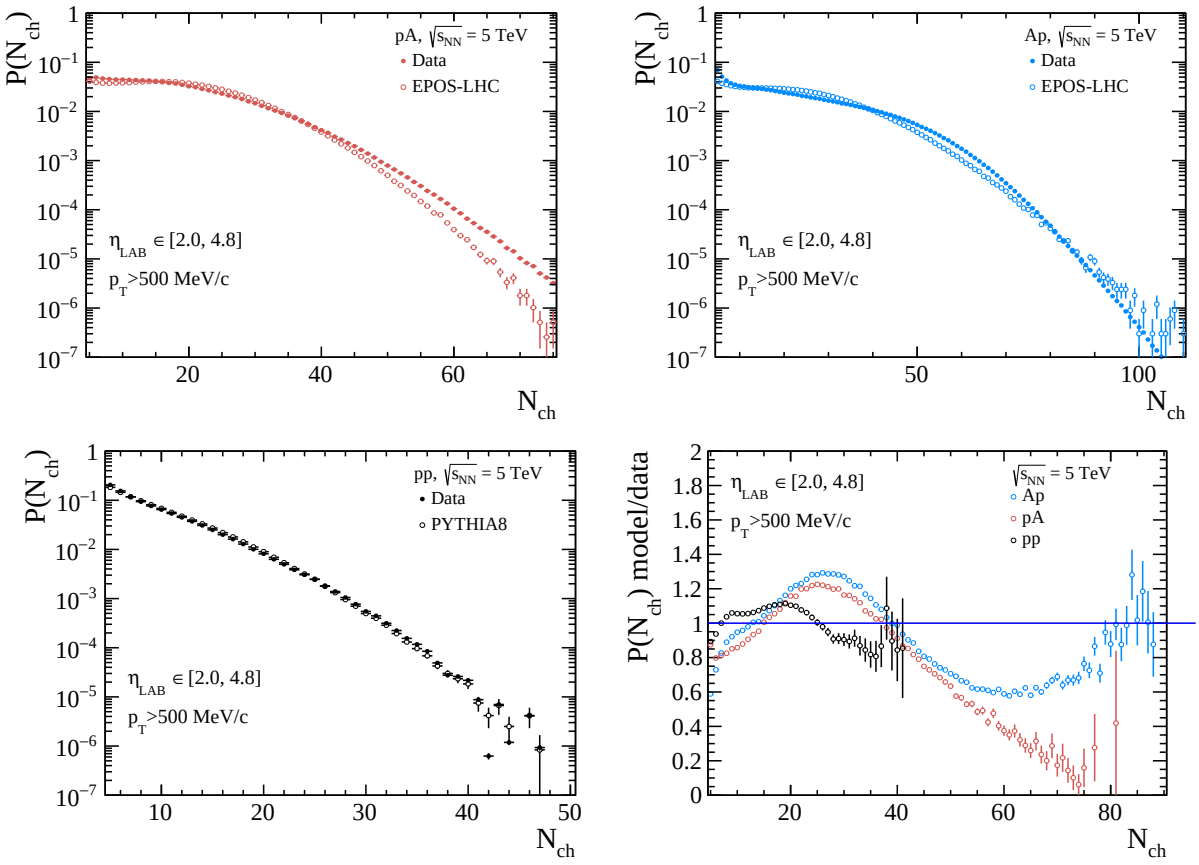


Figure 8.2: Multiplicity distribution comparison with simulation for pPb (top left), Pbp (top right) and pp (bottom left) collisions. The bottom right plot corresponds to the multiplicity distribution model/data ratio for all collision systems in the entire LHCb acceptance.

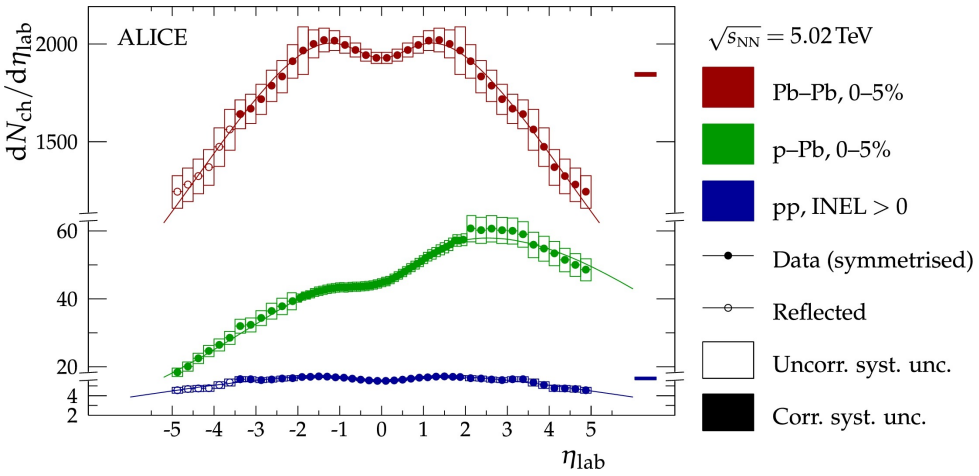


Figure 8.3: $dN/d\eta$ measurement for pp,pPb and $PbPb$ collisions from the ALICE collaboration. Figure extracted from Ref. [100] under Creative Commons CC-BY license.

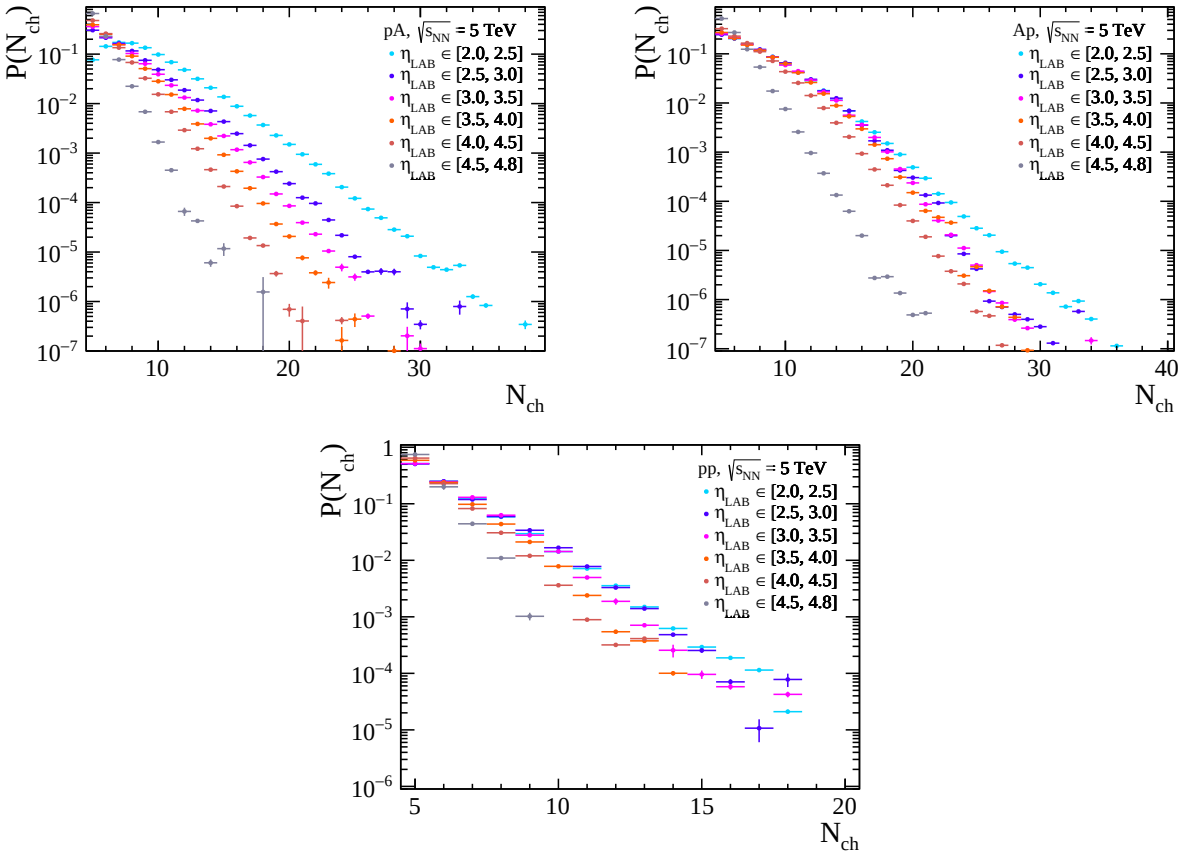


Figure 8.4: Multiplicity distribution results in pseudorapidity bins for pPb (top left), Pbp (top right) and pp (bottom).

8.2 Average transverse momentum

The result for the $\langle p_{\mathrm{T}} \rangle$ as a function of multiplicity integrated in pseudorapidity for all collision systems can be seen at Fig. 8.5. In this result, the most high multiplicity data points from the Pbp result were removed. This will be discussed in the Chap. 10 section, while the result with those data points included can be seen in Sec. B.8. The pseudorapidity-dependent measurement is shown in Fig. 8.7. In the closure test at Fig. 6.35 a systematic deviation of $1-2\%$ has been observed in the most central bin, $\eta_{LAB} \in [2.0, 2.5]$. This biases the $\langle p_{\mathrm{T}} \rangle$ result in that bin and changes the pseudorapidity trend as can be seen in Fig. B.53. For this reason, this bin was not considered. However, as this only affect the $\langle p_{\mathrm{T}} \rangle$ value, the N_{ch} in the X-axis is still $\eta_{LAB} \in [2.0, 4.8]$.

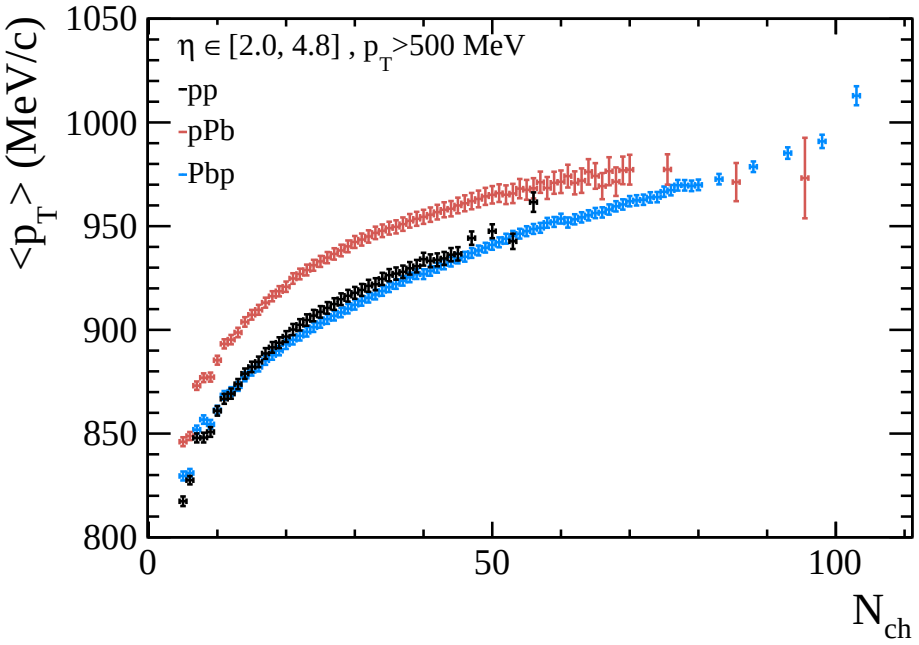


Figure 8.5: Average transverse momentum as a function of multiplicity for Pbp, pPb and pp integrated in pseudorapidity.

In the result, an increase of the $\langle p_{\mathrm{T}} \rangle$ value with multiplicity is observed in all collision systems. This is consistent with previous measurements. However, while the $\langle p_{\mathrm{T}} \rangle$ increasing saturates for pPb at high multiplicities this does not happen in Pbp collisions. Regarding pp data, a hint of saturation is revealed.

According to the most recent measurement published by ALICE collaboration [5]. EPOS-LHC underestimates the $\langle p_{\mathrm{T}} \rangle$ in pPb collisions by a $\simeq 5\%$. Moreover, PYTHIA 8.3 overestimates the $\langle p_{\mathrm{T}} \rangle$ value in pp collisions by a $\simeq 4\%$. This thesis reveals a similar conclusion by looking at the model/data ratio in Fig. 8.6. In that plot, a 10% overestimation of the $\langle p_{\mathrm{T}} \rangle$ is observed in pp collisions. Moreover, in pPb collision, EPOS-LHC underestimated the $\langle p_{\mathrm{T}} \rangle$ value by a 5%. However, although the general trend in pp and

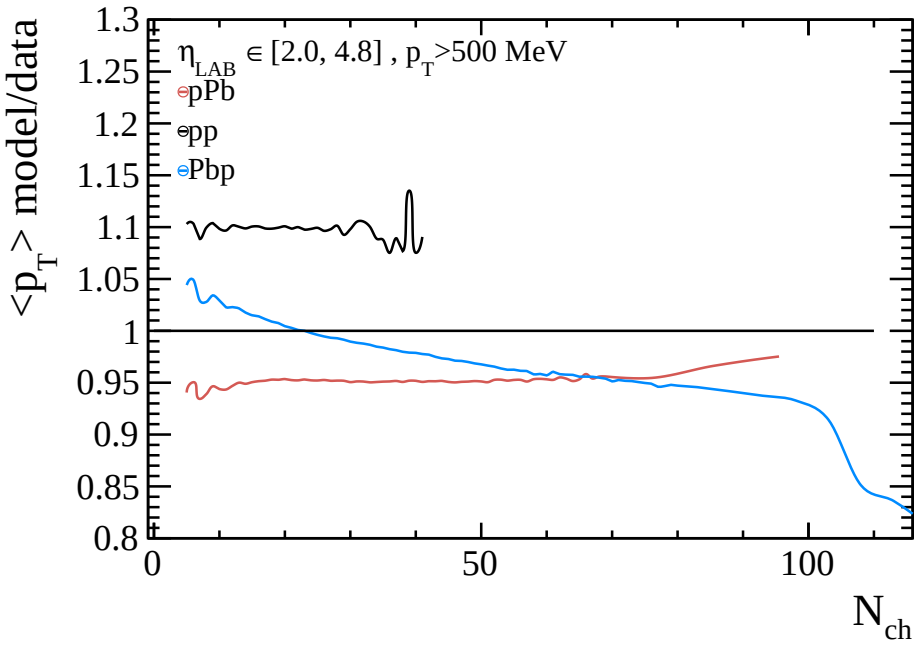


Figure 8.6: Ratio of model predictions to data for all collision systems integrated into pseudorapidity.

pPb collisions is well reproduced by these Monte-Carlo models, in Pbp the trend is not well reproduced. In this collision system, a discrepancy of $5\% - 15\%$ is observed. By looking at the pseudorapidity dependent $\langle p_{\mathrm{T}} \rangle$ measurement in Fig. 8.7 we can obtain more information about the discrepancy in Pbp data.

Concerning the pseudorapidity-dependent measurement, a pseudorapidity decreasing of the $\langle p_{\mathrm{T}} \rangle$ value is observed for all collision systems. This η trend is stronger at high multiplicity events. The result from Pbp data shows that the increasing trend of the η-integrated result comes from the most central region, $\eta_{LAB} \in [2.5, 4.0]$, while for the most forward region, the result achieves a similar value than the pPb one in the high multiplicity limit. This fact reveals that at large η the pPb and Pbp data show the same trend; i.e. an increase of the $\langle p_{\mathrm{T}} \rangle$ followed by a saturation at high multiplicity. However, in the less forward region, a different behaviour in Pbp collisions is observed. This behaviour has never been observed before in the $\langle p_{\mathrm{T}} \rangle$. A possible interpretation concerning the Cronin effect will be discussed in the Chap. 10.

When observing the model/data ratio for $\langle p_{\mathrm{T}} \rangle$ in Fig. 8.8 we see that models describe well the pseudorapidity trend of the $\langle p_{\mathrm{T}} \rangle$ for pPb and pp collisions. In this collision system, the discrepancy is the same for all η bins. However, for Pbp collisions there is a trend change in the $\langle p_{\mathrm{T}} \rangle$ measurement when reducing the pseudorapidity. This change is not reproduced by the EPOS-LHC model.

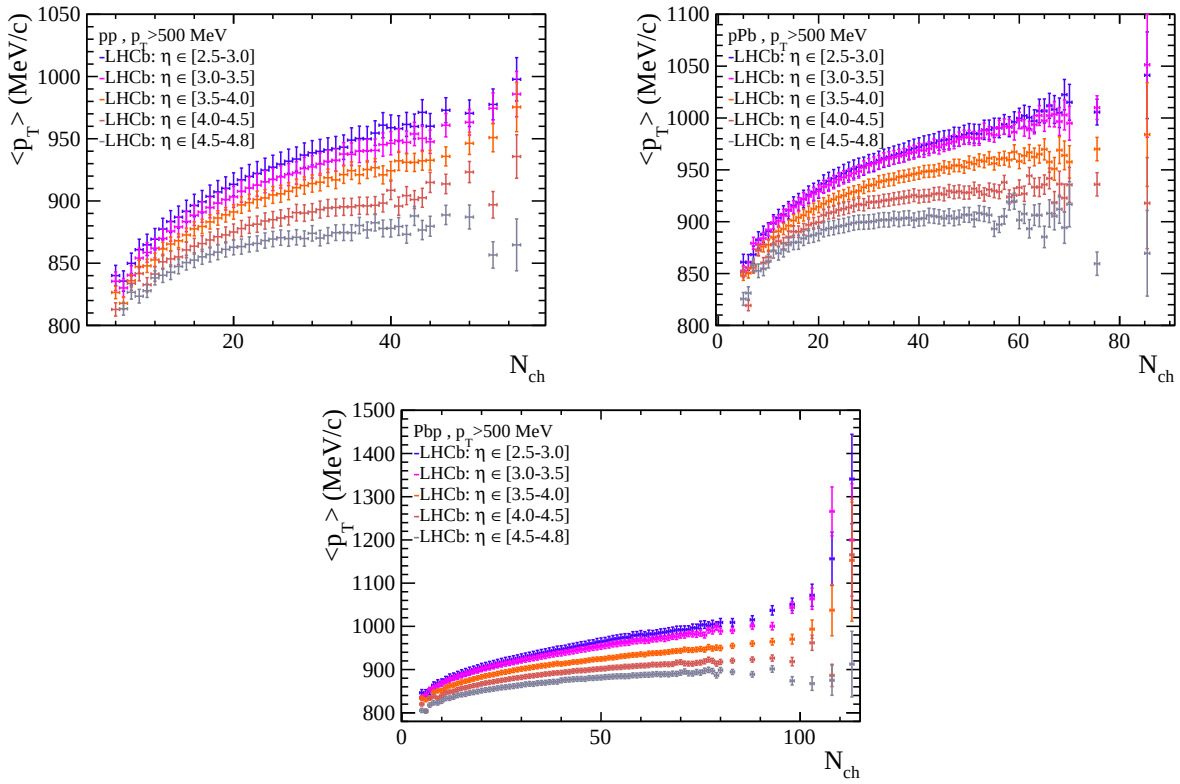


Figure 8.7: Average transverse momentum as a function of multiplicity and pseudorapidity for pp, pPb and Pbp collisions.

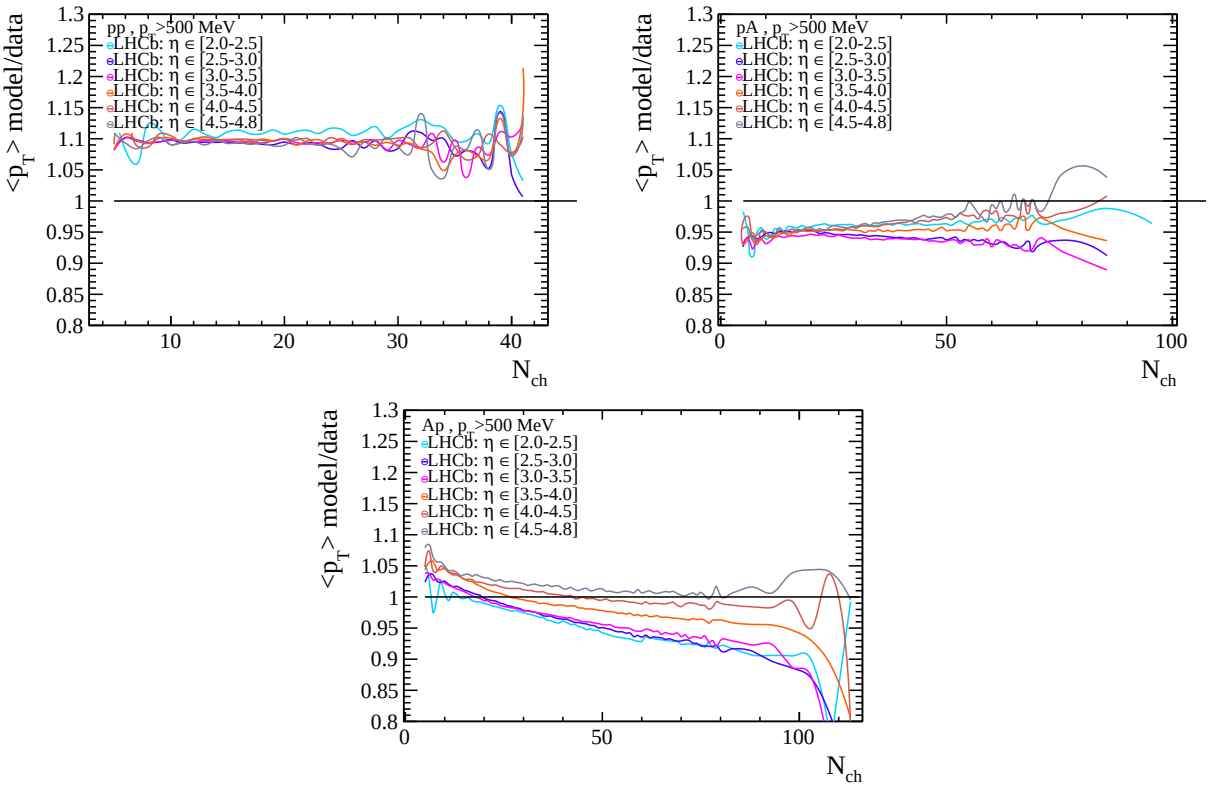


Figure 8.8: Ratio of model predictions to data in pseudorapidity bins. Top left corresponds to *pp*, top right to *p*Pb and bottom plot to Pb*p* collisions.

9

Prospects for Upgrade 2

9.1 The Upgrade 2 at the LHCb experiment

Alongside the implementation of the current upgrade, a second phase (Upgrade II, [101], [102]) is being planned to fully exploit the High Luminosity Large Hadron Collider (HL-LHC) at the LHCb experiment. The goal is to achieve an integrated luminosity of $300\,\text{fb}^{-1}$ for proton-proton (pp) collisions by the end of LHC Run 5. This upgrade aims to handle the anticipated high pile-up of pp collisions (~ 45), requiring extensive development work to operate effectively in a much higher occupancy environment. This advancement will greatly enhance the heavy-ion program by significantly improving performance in nucleus-nucleus collisions across the full centrality range (even the more central collisions). Additionally, dedicated tracking stations are being designed to be placed into the LHCb magnet to provide substantially improved resolution for low-momentum tracks, this subdetector is known as Magnet Stations (MS).

9.2 Temperature dependent measurement of the EoS at LHCb experiment

Thanks to the LHCb Upgrade II a highly efficient low-momentum tracking will be possible in central PbPb collisions. Moreover, the addition of the magnet called Magnet Stations [101] is planned. This subdetector will be focused on very low p_{T} charged particles, down to 50 MeV, increasing the LHCb precision on these measurements. Some of

these measurements are the effective temperature of the Quark-Gluon Plasma, the speed of sound or the viscosity in a wide temperature range. These measurements are crucial to understand the EoS of QCD and will provide valuable insights into the thermodynamics of the early universe.

Bulk physics refers to the study of collective phenomena in ultra-relativistic nucleus-nucleus collisions. These phenomena are sensitive to the nonperturbative QCD interactions between soft particles. These nonperturbative interactions can be used to infer properties of the QGP, such as thermal degrees of freedom or the quantities that characterize the system. Two fundamental quantities are the temperature field $T(x,t)$ and the velocity field $v(x,t)$, being x the position and t time. The relation of other quantities, such as entropy density, with the temperature field, is done via the EoS [25], computed from Lattice QCD. The two main ingredients of this macroscopic description of QCD matter are the hydrodynamic equations and the initial conditions of T and v.

The evolution equations involve the QCD EoS, and, small viscous corrections from transport coefficients (shear and bulk viscosities) of QCD. However, the initial transverse temperature profile modelling (size, shape, fluctuations) is not understood. At the LHC, most theoretical calculations assume longitudinal (Bjorken) symmetry of the initial temperature profile. This means that T is independent of space-time rapidity. However, the rapidity (y) density of prompt charged particles, dN/dy, is not constant [103] which means that longitudinal symmetry is broken. This implies that the final multiplicity is smaller at large rapidities for the same transverse interaction area. Different temperatures can be accessed by solving the hydrodynamic equation for different values of dN/dy. In particular, smaller temperatures at larger rapidity lead to a breaking of Bjorken longitudinal symmetry. This applies to the range covered by the LHCb experiment in collider mode ($2 < \eta < 5$).

Therefore, the unique forward acceptance of the LHCb experiment allows the study of the QCD EoS at different temperatures. In this section, the temperature will be assumed to be the effective temperature, T_{eff}, representing the initial temperature that a uniform fluid at rest would have if it possessed the same amount of energy and entropy as the QGP fluid does in the freeze-out state [1]. Due to longitudinal expansion and cooling, the T_{eff} value is generally lower than the initial temperature of the QGP fluid. Nevertheless, it still characterizes a temperature in the QGP phase. The two quantities studied in this document will be the average transverse momentum as a function of rapidity, $\langle p_{\mathrm{T}} \rangle(y)$, and the speed of the sound squared, $c_s^2(T_{eff})$. Both of these quantities can be calculated from the QCD EoS.

The temperature range in the LHCb acceptance is directly related to the $\langle p_{\mathrm{T}} \rangle(y)$ of charged particles [104]. In the following, the computation of this observable from the EoS and hydrodynamic evolution shown in Fig. 9.1 will be detailed. In order to make a theoretical prediction of $\langle p_{\mathrm{T}} \rangle(y)$, the following steps are done. Starting from the

[1] The point at which the quarks become bound into hadrons.

temperature dependence of the entropy density from EoS, $s(T_{eff})$, the temperature can be related to the average transverse momentum as explained in Ref. [104], $\langle p_{\mathrm{T}} \rangle = 3\,T_{eff}$. Then the particle density is computed using the entropy density via the equation,

$$\frac{dN}{dy} = \frac{dN}{dy}\Big|_{y=0} \cdot \frac{s}{20\,\mathrm{fm}^{-3}}, \tag{9.1}$$

taken from Ref. [104], where the charged particle density at $y = 0$ is taken from Ref. [103]. Finally, the charged particle density can be related to a rapidity range using the ALICE measurement from Ref. [103]. Moreover, the LHCb projection shown in Fig. 9.1 is done using Pythia Angantyr simulation considering a 80% reconstruction efficiency for charged particles.

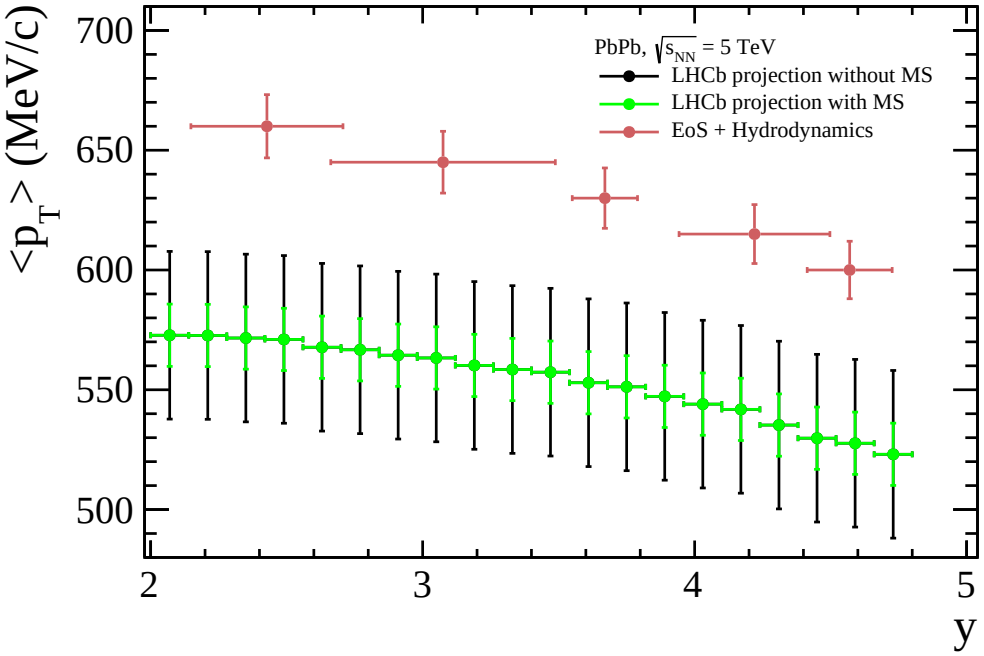


Figure 9.1: Theoretical prediction for $\langle p_{\mathrm{T}} \rangle(y)$ from EoS+Hydro and LHCb projection using PYTHIA Angantyr. The error bars in the EoS+Hydro correspond to a 2% systematic uncertainty. The LHCb projection considers $\mathcal{L}_{\mathrm{int}} = 10\,\mathrm{nb}^{-1}$ collected in PbPb collisions with a centrality range of $[0 - 100]\%$. The error bars on the LHCb projection are the systematic uncertainty from the p_{T} extrapolation.

Following this argument, the LHCb temperature range can be estimated from the $\langle p_{\mathrm{T}} \rangle(y)$ theoretical computation. This estimation gives us a range of $T_{eff} \in [190, 225]$ MeV. The most forward point is not on the plot and had to be extrapolated as the ALICE measurement goes up to $y = 4.67$.

When computing the EoS+Hydro theoretical prediction shown in Fig. 9.1 two uncertainty sources were taken into account. When transforming dN/dy into y using Ref. [103]

the uncertainty from the dN/dy measurement was considered. Moreover, a 2% uncertainty is considered in the $\langle p_{\mathrm{T}} \rangle$ value because of the uncertainty in the temperature determination in the EoS computation from Ref. [25]. From the experimental side, an important systematic uncertainty will be the extrapolation down to $p_{\mathrm{T}} = 0$ using the modified Hagedorn function $\frac{1}{p_{\mathrm{T}}} \frac{d^2N}{dp_{\mathrm{T}}dy} = A(1+p_0/p_{\mathrm{T}})^{-n}$ from Ref. [105]. The $\langle p_{\mathrm{T}} \rangle$ extrapolation uncertainty can be estimated by varying the fitted p_{T} spectra as done in Ref. [106]. In this work we consider as a minimum transverse momentum without MS (with MS), $p_{\mathrm{T}}^{min} = 500\,\mathrm{MeV}$ ($p_{\mathrm{T}}^{min} = 100\,\mathrm{MeV}$). This is a conservative requirement because although this is the minimum value common to the entire LHCb pseudodapidity range, it could be lower in the most forward bins (down to $p_{\mathrm{T}}^{min} = 50\,\mathrm{MeV}$ without MS) [101]. The p_{T} extrapolation uncertainty is estimated by varying the range of measured spectra fitted by the Hagedorn function using Pythia Angantyr simulation. The fit was done varying the p_{T}^{min} in 100 MeV. That corresponds to 500 MeV to 600 MeV in the case of no MS and from 100 MeV to 200 MeV when including MS. Then the systematic uncertainty is estimated as the difference between the $\langle p_{\mathrm{T}} \rangle$ values from each fit. The resulting systematic uncertainty is found to be 35 MeV without MS and 13 MeV with MS. This means that the implementation of low-p_{T} tracking via MS could reduce the systematic uncertainty in a $35-40\%$. According to this estimation, the access to low-p_{T} tracks that the MS could provide would imply a significant reduction of the extrapolation systematic uncertainty. It is worth mentioning that this uncertainty estimation is conservative for the MS case, as we assume 100 MeV as the minimum p_{T} accessible by MS while it could be down to 50 MeV.

Using this T_{eff} range an estimation of the speed of the sound in the QGP, c_s^2, can be determined. In Fig. 9.2, the comparison between the EoS speed of sound calculation and the recent CMS measurement [106] can be seen. The expected LHCb coverage for this measurement is also shown. The estimated uncertainty in the LHCb prediction is the systematic uncertainty from the speed of the sound fit extracted from [106], which is estimated to be 0.010.

Apart from the temperature dependence of the EoS, viscosity coefficients can also be studied in the same temperature range as shown in Fig. 9.2. Viscosity coefficients; i.e. shear (η) and bulk (ξ) viscosity are necessary corrections to ideal hydrodynamics included in state-of-the-art simulations. However, many simulations assume a temperature-independent η/s extracted from measurements. This simplification is used despite studies showing that η/s cannot be constant in QCD matter [107, 108]. As shown in Ref. [109], rapidity differential anisotropic flow measurements in heavy ion collisions can constrain the temperature dependence of η/s of QCD matter. In Upgrade II, LHCb will be able to provide precise measurements of v_2 and v_3 coefficients as a function of pseudorapidity and p_{T} in the forward region. Moreover, thanks to the unique coverage of the PID system at LHCb, an identified particle production measurement can be done. This measurement can be used to perform v_n measurements as a function of rapidity. Hydrodynamics calculations are done as a function of rapidity and its translation to pseudorapidity is not trivial. The $v_n(y)$ measurement would provide valuable information which could be better

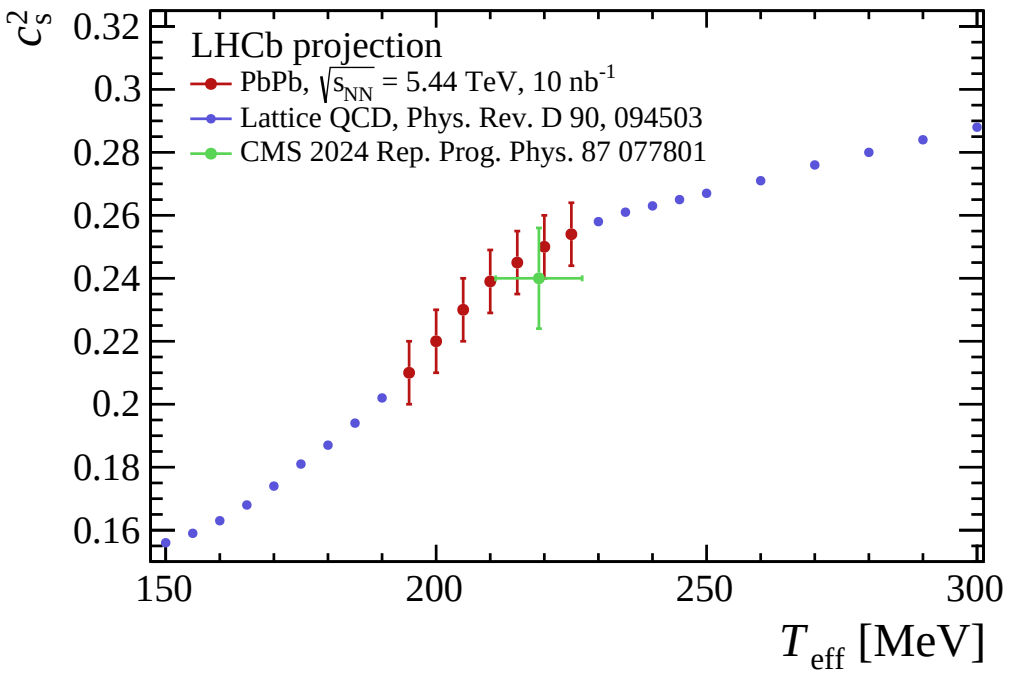


Figure 9.2: Lattice QCD computation for c_s^2 from [25] in the estimated LHCb range compared with the CMS measurement from Ref. [106]. The uncertainty in the LHCb projection corresponds to the systematic uncertainty from the speed of the sound fit.

compared with theory than $v_n(\eta)$ measurements. This constitutes a unique advantage for the LHCb Upgrade II. In addition, lighter collision systems such as OO, XeXe, and KrKr could be used to provide complementary information.

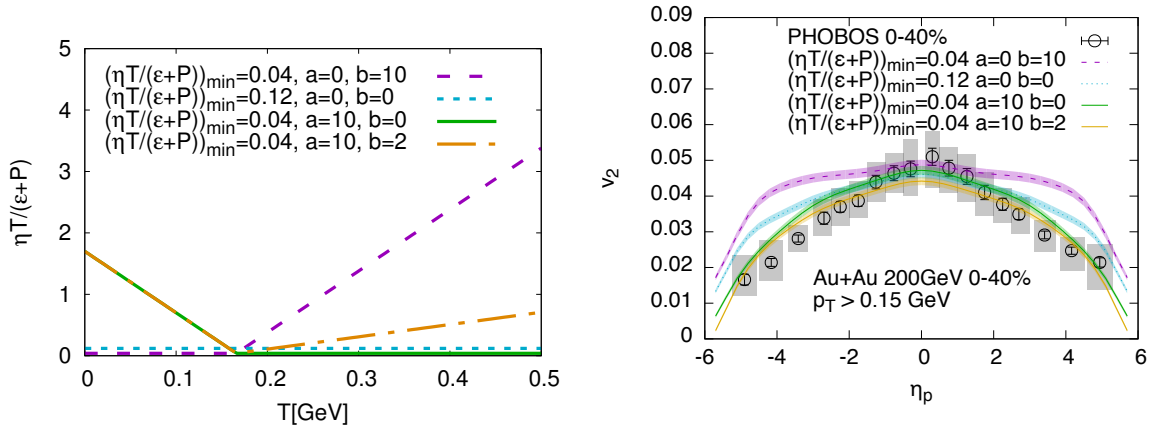


Figure 9.3: (left) Different scenarios for the shear viscosity to entropy ratio as a function of temperature. (right) Resulting v_2 coefficient as a function of the pseudorapidity for each of these scenarios compared with data from PHOBOS [110, 111]. Figures extracted from Ref. [109].

10

Conclusions

In this thesis, the multiplicity distribution and the average transverse momentum multiplicity dependence have been measured in pseudorapidity bins for pp, pPb and Pbp collisions at $\sqrt{s_{NN}} = 5\,\text{TeV}$ in the LHCb experiment. The measurement covers prompt charged particles with $p > 2\,\text{GeV}/c$, $0.5 < p_\text{T} < 8\,\text{GeV}/c$ and η ranges of $-5.3 < \eta < -2.7$ (Pbp), $1.8 < \eta < 4.3$ (pPb) and $2.2 < \eta < 4.8$ (pp), where η is in the nucleon-nucleon centre-of-mass system.

Data recorded by the LHCb experiment for proton-lead collisions in 2013 and for proton-proton collisions during 2015 LHC runs was used in this analysis. Prompt charged particles are defined in simulation using LHC definition [81]. Then, they are measured using long tracks reconstructed by the LHCb tracking system. These tracks are selected using a dedicated set of requirements that optimise the Figure of Merit (FoM), which consists of the product of the significance and the purity. After selection, a set of correction factors are applied to remove detector effects from the signal candidate sample. These corrections are reconstruction efficiency, selection efficiency and background subtraction. A dedicated study is done for each of them as a function of $(p_\text{T}, \eta, \text{nVeloTracks})$. The simulation used in this work is validated by checking that it correctly describes our data. Concerning the occupancy variables where this requirement is not fulfilled, the simulation is fixed by re-weighting. Whenever possible, data-driven corrections are computed for each of the correction factors. Once the corrected signal sample is obtained, the two measurements are performed. The multiplicity distribution requires a Bayesian Unfolding to remove remaining detector effects related to bin migration and finite bin size.

Both measurements are performed for the first time in the LHCb experiment for pPb and Pbp collisions. The average transverse momentum is measured for the first time

simultaneously as a function of pseudorapidity and multiplicity at the LHC.

Apart from the physics results, a method for estimating prompt charged particle multiplicity per event has been developed and implemented in the LHCb software as explained in Chap. 6.10. This method, based on track-by-track correction using simulation, avoids unfolding and allows for data-driven corrections. It was tested by comparing with the generator level multiplicity along with other multiplicity proxies used in the LHCb experiment for the first time. The created proxy has probed to be a good approximation for the generated multiplicity in all collision systems. Additionally, a consistency test called a *closure test* was used to determine whether a given multiplicity estimation method works and how it biases the measurement.

Moreover, in the final part of the thesis, predictions were made concerning the LHCb Upgrade 2. Taking into account the expected capabilities of the upgrade, particularly in PbPb collisions where central collisions could be recorded, an experimental prediction was done. The most relevant calculation is the expected effective temperature range accessible to the LHCb experiment Fig. 9.1. This range was computed using the QCD equation of state (QCD-EoS) [25], the rapidity dependence of the charged particle density [104] and a measurement [103] performed by ALICE collaboration. Based on this calculation, a comparison between LHCb and CMS was made, showing that a wider effective temperature range is accessible in the LHCb detector. Consequently, during Run 5 of the LHCb experiment, a temperature-dependent measurement of the EoS parameters could be performed. The speed of the sound is highlighted as an example in Fig. 9.2. As an OO[1] run is expected to take place in Run3, it would be an excellent opportunity to test these predictions and perform these measurements ahead of Run5.

Several steps in this analysis were improved compared to previous LHCb results such as the Ref [112] in this analysis. For example, the definition of a prompt-charged particle is now aligned with that used in other experiments at the LHC. Overall, the strategy of the multiplicity spectra measurement is based on reducing the impact of Unfolding. Although probed to be necessary [4], unfolding is a complicated method that often works like a black box and can bias results if incorrectly implemented. First, a detailed study of the corrections was conducted in order to remove all the possible detector effects from the track sample. This revealed that a multiplicity-dependent study is relevant, especially for corrections in Pbp collisions. Then, the impact of the Unfolding was evaluated in the closure test in order to understand if it was needed and how much it affected the measurement. Here it was observed that the implemented strategy consisting of applying event-by-event corrections removes an important part of detector effects. The closure test is crucial in order to ensure that the Bayesian Unfolding is correctly implemented. In this document, the closure test gives a perfect agreement for all collision systems.

Regarding the average transverse momentum multiplicity dependence measurement the goal was to have an accurate multiplicity estimator and to assess its bias. In our field, multiplicity-dependent analyses are fundamental tools to study collective effects.

[1]The LHCb detector can record central Oxygen-Oxygen collisions.

However, different analyses measure multiplicity in different ways. In this analysis, some of the different proxies are compared with each other and with the generated values from the simulation. Understanding how those estimators related to the generated distribution helps in the interpretation of the results. Moreover, a multiplicity estimator that has a one-to-one correspondence with the generated multiplicity in a wide range was developed in this thesis. This estimator allows for the inclusion of data-driven corrections, so it is not solely simulation-based. In order to estimate the bias due to our estimator in the multiplicity classification two closure tests were conducted. The tests showed that the estimator reproduced the generated multiplicity with a maximum bias of 5%.

Nevertheless, there are areas of improvement to achieve better results. In the multiplicity spectra results, the cut on the number of primary vertices and the minimum multiplicity cut removes very low multiplicity events that are usually diffractive. Knowing the proportion of these events is important for Monte Carlo generators. Additionally, by removing these events the final results are not comparable to other collaboration results, which extend down to 1 in the multiplicity spectra. It is worth mentioning that other analyses report $dN/d\eta \simeq 5$ (see [5]), so this cut makes the $dN/d\eta$ extracted from our measurement not comparable to analyses without minimum multiplicity cut. Eliminating the minimum multiplicity selection cut and the requirement of having a primary vertex would increase the amount of information obtained from the measurement. Furthermore, although the *truth matching* efficiency was shown to be a very small correction ($\simeq 1\%$) in Ref. [70], it is worth studying this correction as a function of multiplicity. Preliminary studies at LHCb indicate that for high multiplicity events, this correction could be $\simeq 3\%$.

Regarding the average transverse momentum results, the minimum p_{T} requirement could also be relaxed in the forward region bins. Although it is true that $p_{\mathrm{T}} = 500\,\mathrm{MeV}$ is the minimum p_{T} value common to all the η range, in the most forward bins the LHCb detector can reconstruct long tracks down to $p_{\mathrm{T}} = 20\,\mathrm{MeV}$. Thus, an interesting extension of the result could be achieved by decreasing the p_{T} minimum cut in most forward bins to the minimum. This would significantly increase the available statistics.

Comparing our results to the most recent ones provided by the ALICE collaboration in Ref. [5] we can observe several things. First, the results presented here expand in a wider range in $P(N_{ch})$. For example for pPb collisions there result from Ref. [5] covers down to 10^{-4}, while in this thesis, we can access until 10^{-7}. This means that we are accessing less common events, which are more sensitive to rare fluctuations in the nucleon shape. However, although the studied range is wider, the discrepancy with Monte-Carlo models is similar as can be seen in Fig. 10.1. The comparison of pp data with the PYTHIA model shows a $10 - 12\%$ overestimation of the $P(N_{ch})$ value at intermediate multiplicities. Moreover, a 10% underestimation at low multiplicity is observed consistent with the ALICE result. The high multiplicity region is better described in our measurement than in the ALICE results. The reason for this behaviour may correspond to the higher p_{T}^{min} in this study. In general, the smaller the p_{T} the more difficult the modelling is because of the non-perturbative contribution. It is worth nothing that the p_{T}^{min} in Ref. [5] is 150 MeV while in this document is 500 MeV. Regarding the pPb result, a $\sim 20\%$ overestimation of

the $P(N_{ch})$ result was found in this study when compared with EPOS-LHC at intermediate multiplicity. Moreover, a 10% to 15% underestimation of the result is found at low multiplicity. In addition, the comparison with EPOS-LHC shows a systematic underestimation of the $P(N_{ch})$ as the multiplicity increases. These are the same conclusions that show results from Ref. [5]. Finally, although the Pbp result shows a similar discrepancy with EPOS-LHC than pPb until $N_{ch} \simeq 60$, at higher multiplicities the model predicts better the trend.

When comparing our results to those from other experiments as the recent ones from ALICE collaboration from Ref. [5] the p_{T} range has to be taken into account. In particular, the minimum p_{T} has a significant effect on the $\langle p_{\mathrm{T}} \rangle$ measurement, as the low-p_{T} part is where most of the particles are produced. In our study, the minimum p_{T} is larger than the one considered in Ref. [5] which explains the scale difference between both results. However, the qualitative behaviour is consistent, the $\langle p_{\mathrm{T}} \rangle$ grows with multiplicity for all collision systems.

The $\langle p_{\mathrm{T}} \rangle$ results can be compared with the R_{pPb} ones from Ref. [70] shown in Fig. 10.2. Suppression of the low-p_{T} prompt charge particle production is observed in pPb collisions in comparison with pp. Moreover, in the intermediate and high-p_{T} region, the suppression is smaller being almost negligible at $p_{\mathrm{T}} = 8\,\mathrm{GeV}$. Because of this reason, an increase of the $\langle p_{\mathrm{T}} \rangle$ value in pPb collisions could be expected when comparing it to pp data. A similar reasoning can be made for Pbp data. Here the suppression in the interesting p_{T} region (around the average value $\sim 1\,\mathrm{GeV/c}$) is smaller, which means that the expected $\langle p_{\mathrm{T}} \rangle$ value will be more similar to the pp result.

In the introduction of the thesis, it has been mentioned how the hadronic structure and the medium formation can be studied through prompt charged particle measurements. Figure 10.3 shows a specific theoretical calculation, illustrating how the effects of Color Glass Condensate (CGC) and Hydrodynamics produce completely different results for the $\langle p_{\mathrm{T}} \rangle$ (y) trend. The two predictions from the CGC model shown in the plot correspond to two different approaches when modelling particle production within this theory. This study observes that the $\langle p_{\mathrm{T}} \rangle$ in the CGC model increases when moving towards the proton fragmentation (positive η) region due to the increasing saturation momentum of the nucleus, Q_A.

According to the authors, in hydrodynamics, the $\langle p_{\mathrm{T}} \rangle$ is expected to decrease with rapidity. This is in agreement with the calculations made in this thesis for PbPb collisions in Chap. 9. The explanation provided in the paper is that the $\langle p_{\mathrm{T}} \rangle$ the hydrodynamic model comprises two contributions, the thermal motion at the freeze-out and the collective velocity acquired during the expansion. In pPb collisions the matter density is strongly dependent on rapidity showing a larger multiplicity on the lead side. Within the hydrodynamic theory, the flow velocity originates from pressure gradients in the fireball. Moreover, the energy density in the fireball should increase when going to the lead side. This means that a decrease of the $\langle p_{\mathrm{T}} \rangle$ is expected to decrease with the rapidity increasing. Disentangling hadronic structure effects from the medium formation is crucial in small system analyses. This is one of the fundamental goals of this study.

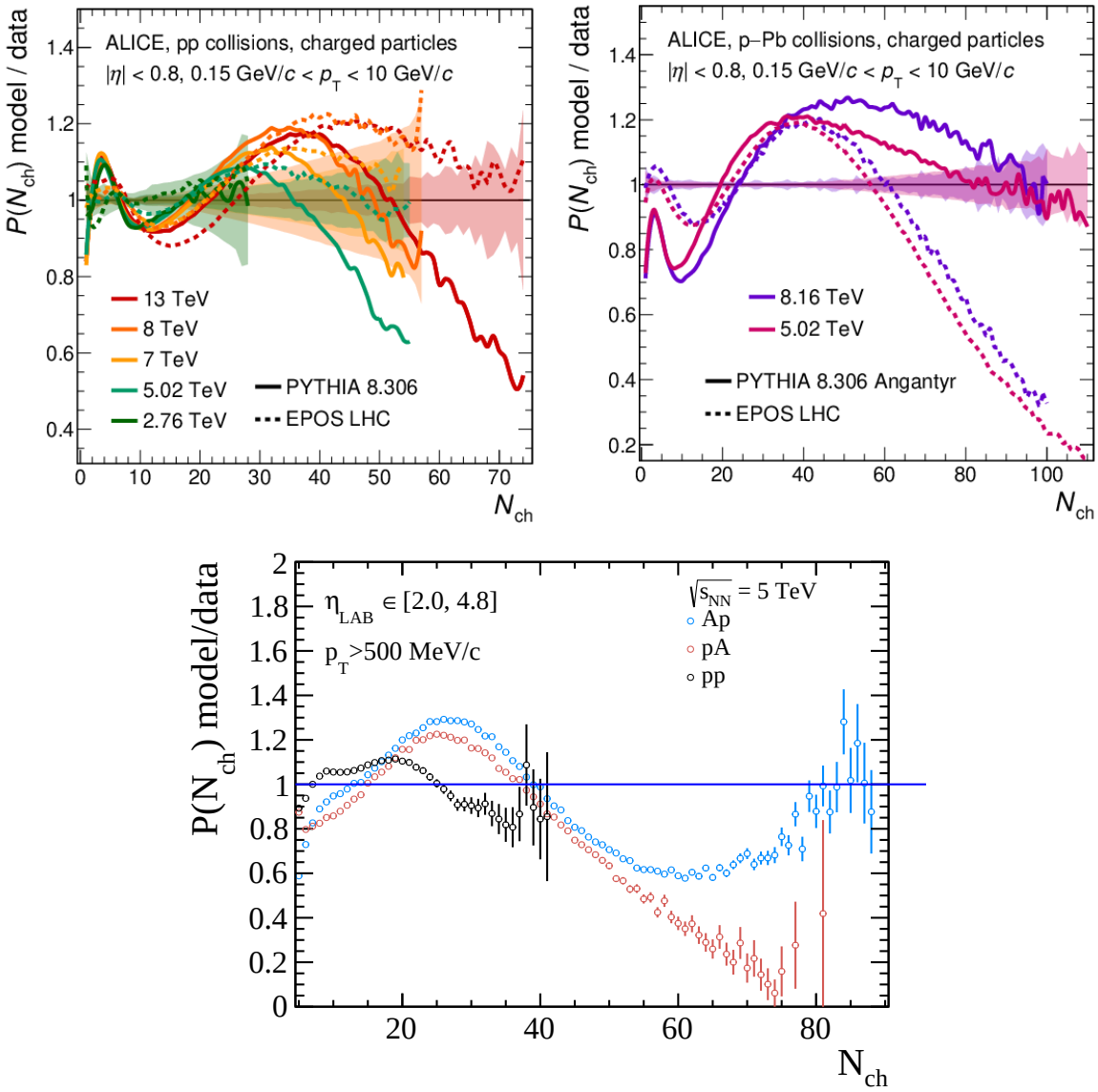


Figure 10.1: Ratio of model predictions to data for pp (top left) and pPb (top right) collisions at various energies from ALICE. Figure taken from Ref. [5] under Creative Commons CC-BY license. The bottom plot corresponds to the model/data ratio from this thesis using LHCb data with a minimum multiplicity of $N_{ch}^{min} = 5$.

However, in Ref. [117], it has been revealed that including parton fragmentation modifies the rapidity dependence of the CGC prediction. Although, after including fragmentation, some CGC-related models like BUW or DHJ could explain a decreasing trend on the $\langle p_{\mathrm{T}} \rangle$ (y), the GBW model still predicts an increasing behaviour that does not explain our data. This suggests that the $\langle p_{\mathrm{T}} \rangle$ (y) still can constrain CGC models. In this context, more theoretical input is needed.

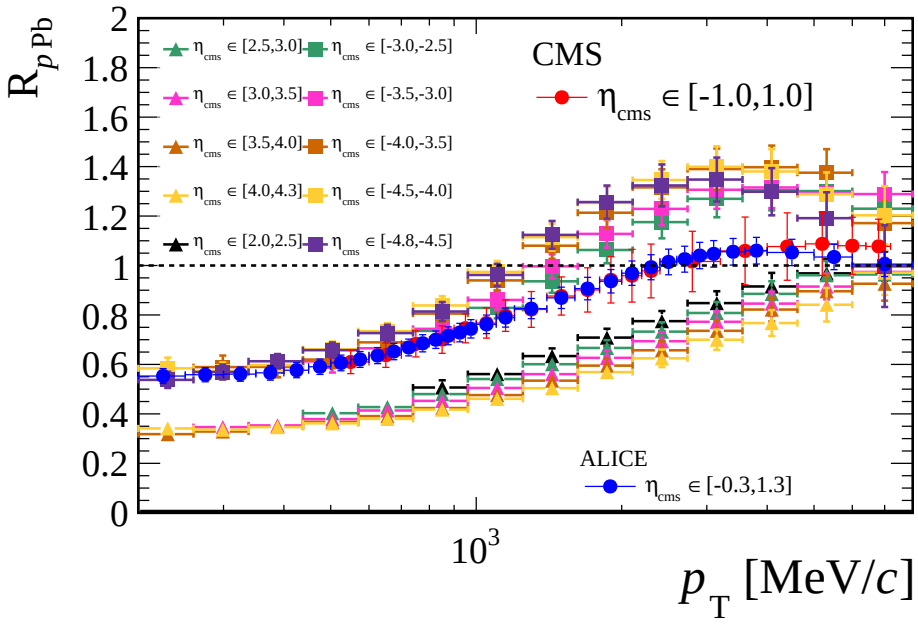


Figure 10.2: Nuclear modification factor for the forward and backward region measured at LHCb [70] and compared with the ALICE [113] and CMS [114, 115] measurements in the central region. Error bars include statistical, systematic and luminosity uncertainties added in quadrature. Figure extracted from Ref. [116] under CC BY-NC-ND 4.0 licence.

An enhancement of particle production at intermediate p_{T}; i.e. $p_{\mathrm{T}} \in [2, 5]\,\mathrm{GeV}/c$ when comparing proton-nucleus with proton-proton collisions was discovered by Cronin et. al. at Ref. [118]. Some explanations of the origin of this effect Ref. [119] revealed that two different regimes are relevant. At low energies, a high-p_{T} particle is produced incoherently on different nucleons, and the Cronin effect is due to soft multiple initial/final state interactions. At LHC energies the nucleons are not resolved, and the interaction with the target is coherent. In that study, the authors predicted a smaller Cronin effect at LHC energies than at RHIC. This prediction is in qualitative agreement with ALICE R_{pPb} measurement from Ref. [120] where a small increase ($\sim$ 10%, almost compatible with zero within uncertainties) of R_{pPb} around $p_{\mathrm{T}} \in [3, 4]\,\mathrm{GeV}/c$ is observed. However, the LHCb measurement shown in Fig. 10.2 from Ref. [70] has revealed a stronger increase of the nuclear modification factor, $R_{pPb} \sim 1.3$, in the backward region. This could prove that the Cronin effect has a strong pseudorapidity dependence. As shown in our study, in Pbp collisions the $\langle p_{\mathrm{T}} \rangle$ result demonstrates an increasing trend that does not saturate as in the pPb and pp case. Moreover, this behaviour arises in the less forward bins, which results in a higher parton x. In this sense, more theoretical input is needed in order to understand the pseudorapidity dependence of the Cronin effect at LHC energies.

Apart from the Hydrodynamics prediction where the $\langle p_{\mathrm{T}} \rangle$ is expected to decrease with rapidity, there have been recently some advances concerning the $\langle p_{\mathrm{T}} \rangle$ multiplicity

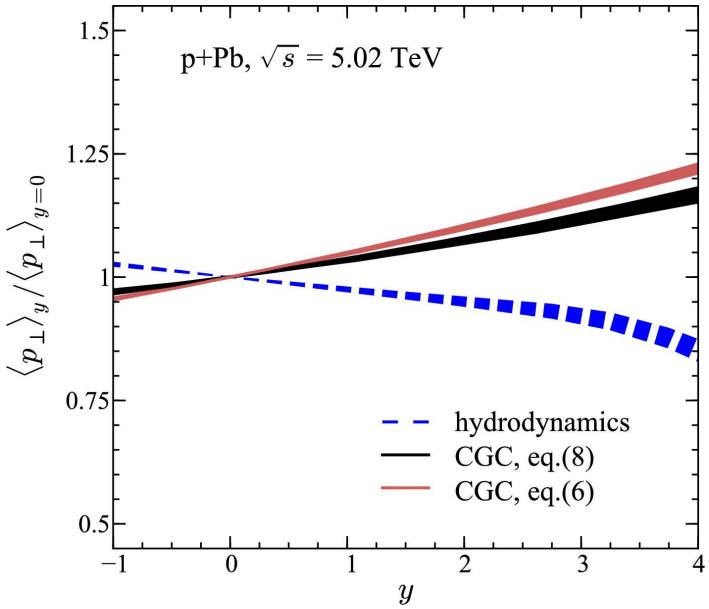


Figure 10.3: Average transverse momentum of produced particles as a function of rapidity, normalised by the average transverse momentum at $y = 0$. Three different centralities are shown. Figure extracted from Ref. 10.3 under Creative Commons CC-BY license.

dependence in high multiplicity PbPb collisions. In Ref. [106] the speed of the sound is measured by fitting the raising trend of $\langle p_{\mathrm{T}} \rangle$ at ultra-central ion collisions, see also Ref. [121]. To date, the increasing of the $\langle p_{\mathrm{T}} \rangle$ at very high multiplicity events has not been explained by any alternative model, nor observed in small collision systems, making it a signature of a deconfined phase of matter. In this study, a rising trend is not observed in pp or pPb collisions while in Pbp the results are compatible with this behaviour. However, since these high multiplicity events are rare, more research is needed to reach a conclusion. Implementing a high multiplicity trigger for pp and pPb collisions could be helpful in order to collect more statistics for those events. Moreover, the Oxygen-Oxygen data taking is an excellent opportunity to study the emergence of hydrodynamic behaviours in intermediate systems (between pPb and PbPb).

A

Resumo da tese

A.1 Motivación teórica e obxectivos da tese

A forza nuclear forte esta descrita pola teoría cuántica de campos denominada Cromodinámica Cuántica (QCD, do inglés "Quantum Chromodynamics"). Esta interacción e a responsable da estabilidade do núcleo atómico e da existencia dos hadróns (como o neutrón e o protón). Os partons, quarks e gluons, son as compoñentes de materia fundamentais desta teoría, e son as pezas que conforman os hadrons e os núcleos atómicos. Co obxetivo de estudar a materia, a súa estructura fundamental e o seu comportamento en rexímenes extremos de densidade de enerxía e temperatura, realizanse colisións hadrónicas a velocidades relativistas e estudanse as suas propiedades.

Nas colisións de hadróns, ao aumentar a enerxía do centro de masas de colisión $\sqrt{s_{NN}}$, supérase a barreira de Coulomb e éntrase no réxime donde as colisión se producen a nivel partónico. É decir, os que coliden entre sí son as compoñentes dos hadrons e non os hadrons como sistema. Este réxime, denominado altas e enerxías, é en xeral o que se estuda no LHC (Large Hadron Collider).

Moitos destes estudos focalízanse no estudo das partículas cargadas primarias que se producen nas colisións protón-protón (pp) e protón-núcleo (pPb). Neste contexto desenrolase esta tese donde se estudan algunhas propiedades das partículas cargadas primarias producidas colisións pp, pPb e Pbp no experimento LHCb do colisiónador LHC situado no CERN.

Nas colisións hadrónicas a altas enerxías a enerxía inicial do sistema transfórmase en materia que sale expulsada do punto de colisión a altas velocidades. O proceso de

formación desta materia, na sua maioria pións, kaons e protóns e un obxeto de estudo da QCD que involucra tamen o coñecemento da estructura interna dos hadrons. Ata agora, non hai un modelo que explique todos os aspectos da produción de partículas primarias nos diferentes sistemas de colisión (dende *pp* ata núcleos pesados). Polo tanto, estudar as propiedades da produción de partículas en sistemas de colisión diferentes axuda a elaborar modelos e a coñecer mellor tanto a estructura interna da materia, como a QCD en sí mesma.

No marco da QCD, os procesos de produción de partículas pódense dividir en dúas categorías atendendo a sua enerxía característica. Os procesos duros (do inglés "hard") teñen enerxías superiores a constante fundamental de QCD, Λ_{QCD}, os procesos blandos (do inglés "soft") teñen unha enerxía característica inferior a Λ_{QCD}. As interacións duras poden describirse usando métodos perturbativos (pQCD), o que quere decir que se poden calcular con moita precisión. Sen embargo, os procesos blandos non se poden calcular usando estes métodos, e hai que recurir a modelos aproximados (effective models) [122,123] ou métodos de calculo numéricos como a Lattice QCD [25]. Estas últimas, as interacións blandas, son as dominantes no canto a produción de partículas cargadas primarias se refire nas enerxías do LHC. Por este motivo é importante estudar as propiedades destas partículas, xa que son relevantes para delimitar os modelos de produción.

Como xa se mencionou, o proceso de produción de partículas nas colisión hadronicas de altas enerxias esta influenciado pola estructura interna dos hadrons. A través do estudo comparativo entre colisións *pp* e *p*Pb podemos ver como os efectos nucleares fríos (CNM, do inglés "Cold Nuclear Matter") modifican a estructura interna dos nucleóns [124,125]. A modificación da estructura interna dos hadrons debido ao feito de formar parte dun núcleo parametrízase normalmente nas funcións de distribución dos partons no núcleo (nPDFs, "nuclear parton distribution functions") [126–128]. O efecto de saturación ocorre cando a densidade de gluóns nos nucleóns aumenta ata chegar a un valor máximo no que para de aumentar e satura. O réxime donde se espera que a saturación ocorra corresponde con valores baixos da fracción de momento dos partóns x e con núcleos cun número elevado de nucleóns A [129]. Por esta razón, comparar as propiedades da produción de partículas primarias nos sistemas *p*Pb e Pb*p* e de especial relevancia xa que se espera que a saturación ocorra en *p*Pb de forma mais pronunciada que en colisións Pb*p* debido ao rango de x ao que se ten acceso. Para entender esto mellor, definimos pseudorapidez definida como:

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

Donde θ e o ángulo da partícula con respecto ao eixo de colisión. Entón, o x dos partons relacionase coa mesma a través da seguinte ecuación,

$$x \approx \frac{m_{\mathrm{T}}}{\sqrt{s_{\mathrm{NN}}}} e^{-\eta}. \tag{A.2}$$

A diferencia no espazo de fases (x, Q^2) que se cubre cos diferentes sistemas de colisión nos diferentes experimentos pode observarse na Fig. A.1.

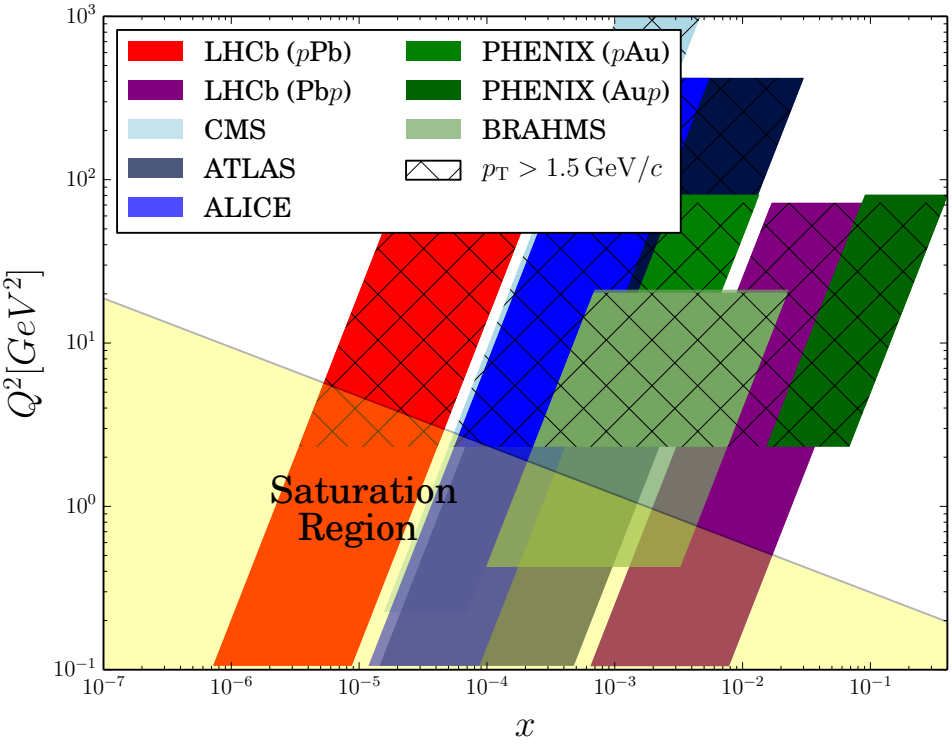


Figure A.1: Kinematic coverage of LHCb and other experiments for pPb or dA collisions in terms x from Eq. A.2 and $q^2 \approx m_{\mathrm{T}}^2$. The kinematic range is taken from measurements of inclusive charged particle production performed at the ALICE [113], CMS [114], ATLAS [130], BRAHMS [131], PHENIX [132] experiments.

Nesta situación, a dinámica de QCD pode describirse coa teoría de campos efectiva do condensado vidrioso de color (CGC, do inglés "colour glass condensate") [133]. Entender estos fenómenos é importante para comprender a interación forte, a estructura da materia nuclear e tamén para interpretar os datos de colisións de ions pesados AA.

O obxetivo desta tese e medir duas propiedades relacionadas coa produción de partículas cargadas primarias, a distribución de multiplicidade, $P(N_{ch})$ e o momento transverso promedio, $\langle p_{\mathrm{T}} \rangle$. No caso de $P(N_{ch})$ estudarase en varios bins de η mentras que no caso do $\langle p_{\mathrm{T}} \rangle$ estudarase a dependencia coa multiplicidade en diferentes bins de η. Ambos observables mediranse en colisións pp, pPb e Pbp. O rango de momento transverso cuberto para as partículas cargadas será de $0.5 < p_{\mathrm{T}} < 8\,\mathrm{GeV}/c$ e o rango de pseudorapidez de $2.0 < \eta_{LAB} < 4.8$.

Como xa se especificou, o sinal neste traballo consiste en partículas cargadas producidas directamente na colisión primaria. A definición específica ven dada polo LHC Minimum Bias and Underlying Event Working Group [81] como:

Hadrons ou leptons, con unha vida media de $\tau > 0.3 \cdot 10^{-10}$ s*, producidos directamente ou dende desintegracións de partículas de vida mais corta.*

Na Tab. A.1 podemos ver unha lista das especies de partículas que caen dentro desta definición.

Especie	**Tempo medio de vida**	
	τ**(s)**	$c\tau$**(cm)**
$\pi^{\pm}$	$2.60 \cdot 10^{-8}$	780
$K^{\pm}$	$1.24 \cdot 10^{-8}$	371
$p^{\pm}$	∞	∞
$e^{\pm}$	∞	∞
$\mu^{\pm}$	$2.19 \cdot 10^{-6}$	65800
$\Xi^{\pm}$	$1.64 \cdot 10^{-10}$	4.91
Σ^{+}	$0.80 \cdot 10^{-10}$	2.40
Σ^{-}	$1.48 \cdot 10^{-10}$	4.434
Ω^{-}	$0.821 \cdot 10^{-10}$	2.461

Table A.1: Especies de partículas incluidas dentro da categoría de partículas cargadas primarias.

A distribución de multiplicidade danos a probabilidade de que unha colisión hadrónica produza un número de partículas concreto N_{ch} e ven definido como,

$$P(N_{\mathrm{ch}}) = \frac{1}{N_{evt}^{T}} \frac{dN_{evt}}{dN_{\mathrm{ch}}}. \tag{A.3}$$

Nesta expresión N_{evt} e o número de eventos, N_{evt}^{T} e o número de eventos totais e N_{ch} e a cantidade de partículas primarias producidas.

O momento transverso promedio para unha multiplicidade dada ven definido pola seguinte expresión,

$$\langle p_T \rangle (N_{\mathrm{ch}}) = \frac{\sum_{p_{\mathrm{T}}} \frac{dN_{\mathrm{ch}}}{dp_{\mathrm{T}} d\eta} \cdot p_{\mathrm{T}} \cdot \Delta p_{\mathrm{T}}}{\sum_{p_{\mathrm{T}}} \frac{dN_{\mathrm{ch}}}{dp_{\mathrm{T}} d\eta} \cdot \Delta p_{\mathrm{T}}}. \tag{A.4}$$

A.2 O experimento LHCb e mostras de datos

O experimento LHCb e un experimento de fisica de partículas de propósito xeral na rexión dianteira descrito en detalle nas Refs. [42, 48]. Este detector está formado por diferentes subsistemas e subdetectores. Para a realización do traballo exposto nesta tese, os subdetectores fundamentais son os involucrados na reconstrución das trazas das partículas cargadas (tracking) e a reconstrución do vértices primario de interación (VP):

- VELO: O VErtex LOcator é un detector de silicio que rodea a rexión de interacción e permite a medida da posición do VP. Ademais é parte fundamental no sistema de reconstrución de trazas xa que reconstrúe un anaco das traxectoria da partículas cargadas.

- Tracking system: Os subdetectores TT, IT e OT son os encargados de proveer da información necesaria para reconstruir as trazas das partículas aos algoritmos de tracking.

- RICH: Dous detectores de anel Cherenkov (RICH1 e RICH2) capaces de discriminar entre distintas especies de partículas cargadas.

Nesta tese, faise uso dos datos de collisions pPb e Pbp recollidos en febreiro do ano 2013. Estos datos contan con unha luminosidade integrada de $42.7 \pm 1.0\,\mu\text{b}^{-1}$ e $38.7 \pm 1.0\,\mu\text{b}^{-1}$, respectivamente, onde as incertezas non están correlacionadas entre as mostras. Os eventos estan filtrados con un sistema de disparo ("trigger") que require unha traza reconstruida no VELO. Sobre estos datos aplicanse outros dous filtros de selecion. O primeiro seleciona os eventos que conteñen un so VP, e o segundo require que este VP estea reconstruído dentro de tres desviacións estándar da media da distribución de VP da mostra total. Esta rexión denomínase rexión luminosa, e correspóndese coa rexión onde se cruzan os feixes. O obxetivo desta selecion e focalizarnos en colisións inelasticas donde so haxa unha colisión por evento, e ademais reducir o fondo debido a colisións co material do detector. A mostra de datos correspondente as colisións pp foi recollida no 2015 e consta dunha luminosidade integrada de $3.49 \pm 0.07\,\text{nb}^{-1}$. Os datos selecciónanse cun sistema de disparo sen nesgo que selecciona os cruces de paquetes do feixe que encabezan un tren. Esta estratexia tamén evita a contaminación procedente de paquetes veciños. Ambas mostras de datos, pPb e pp, corresponden a colisións con unha enerxía no centro de masas de $\sqrt{s_{NN}} = 5\,\text{TeV}$.

Ademais das mostras de datos mencionadas, esta tese utiliza mostras de datos simulados. Estas mostras usanse para estudar a eficiencia da reconstrución, os efectos da selecion e a contribución do fondo a nosa sinal. As mostras de simulacion xeranse utilizando xeradores Monte-Carlo. En concreto, a mostra de pp xerouso con PYTHIA8 [30] mentras que as mostras de pPb e Pbp veñen do xerador EPOS-LHC [36]. Despois de xerar a fisica das colisións, para estudar os efectos experimentais que mencionamos anteriormente e necesario procesar as mostras. As desintegracións de partículas modelizanse con EvtGen [67], mentres que a interacción das partículas co detector e a súa resposta realizase usando Geant4 [33, 68], tal e como se describe no documento Ref. [65].

A.3 Medida da distribución de multiplicidade

O paso mais importante da analise e medir a cantidade de partículas cargadas primarias de acordo coa definición que se detallou anteriormente. O analise utiliza as trazas

longas reconstruidas polo experimento LHCb. Estas trazas son as de mellor calidade do experimento porque deixan sinal en todos os detectores de trazas do experimento. Na mostra de trazas reconstruídas, aparte de partículas cargadas primarias, pódense atopar trazas fondo que non corresponden con partículas. Estas clasifícanse nos seguintes tipos:

- Trazas pantasma. Son defectos da reconstrución que non se corresponden cunha partícula cargada verdadeira. Son especialmente importantes en eventos cunha ocupación do detector alta e a alto p_{T}.
- Trazas clonadas. Tamén son defectos da reconstrución, pero específicamente son parellas de trazas orixinadas por unha única partícula cargada.
- Partículas secundarias. Son trazas producidas por partículas cargadas que non cumpren os requisitos para seren consideradas primarias segundo a definición anterior. Orixínanse principalmente en interaccións de partículas co detector e en desintegracións de partículas primarias.

Co obxetivo de reducir a contribucion do fondo na noso mostra aplicamos unha selecion ao conxunto de trazas. O resumen da selecion pode observarse na Tab. A.2. Ao aplicar a selecion ao número de trazas longas, obtemos o número de candidatos, N_{cand}.

		pp (2015)	pPb (2013)	Pbp (2013)
Event selection	Trigger	`Hlt1NoBiasLeadingCrossing`	`Hlt1MBMicroBiasVelo`	`Hlt1MBMicroBiasVelo`
	Bunch crossing type	bunch-bunch	bunch-bunch	bunch-bunch
	number of PV	1	1	1
	PV position	-	within luminous region	within luminous region
Candidate selection	Track location	`TES["Rec/Track/Best"]`	`TES["Rec/Track/Best"]`	`TES["Rec/Track/Best"]`
	Track type	Long	Long	Long
	η	$2 < \eta < 4.8$	$2 < \eta < 4.8$	$2 < \eta < 4.8$
	p	$p > 2$ GeV/c	$p > 2$ GeV/c	$p > 2$ GeV/c
	p_{T}	$0.200 < p_{\mathrm{T}} < 8$ GeV/c	$0.200 < p_{\mathrm{T}} < 8$ GeV/c	$0.200 < p_{\mathrm{T}} < 8$ GeV/c
	IP	0.508 mm	0.508 mm	0.568 mm
	`GhostP`	0.078	0.103	0.109
	`IsClone`	`False`	`False`	`False`

Table A.2: Resumen da selecion aplicada no analise.

Unha cantidade fundamental nesta tese e o número e partículas cargadas primarias por evento, N_{corr}. Esta cantidade calculase a partir do número de trazas longas reconstruidas no LHCb que pasan unhos filtros de selecion, N_{cand}. A definición e a seguinte,

$$N_{\mathrm{corr}} = \frac{N_{\mathrm{cand}} \cdot \Pi}{\varepsilon_{reco} \cdot \varepsilon_{sel}}. \tag{A.5}$$

Onde Π e a pureza, ε_{reco} a eficiencia de reconstrución e ε_{sel} e eficiencia de selecion. Estos factores, son correcions ao número de candidatos que teñen como obxetivo eliminar

os efectos do detector nos datos para poder obter unha medida fiable. Todos eles se calculan como funcion de $(\eta, p_{\mathrm{T}}, \mathrm{VeloTracks})$.

A eficiencia de reconstrución ten en conta as partículas non reconstruidas polo detector. Esta calculase usando simulacion para logo correxir as diferencias entre datos e simulacion usando trazas de muons da desintegracion $J/\Psi \to \mu^+\mu^-$ no rango $5 < p < 200\,\mathrm{GeV/c}$. O resultado de ε_{reco} para as colisións pPb pode verse na Fig. A.2, os resultados para os demais sistemas de colisión podes verse no Chap.6.

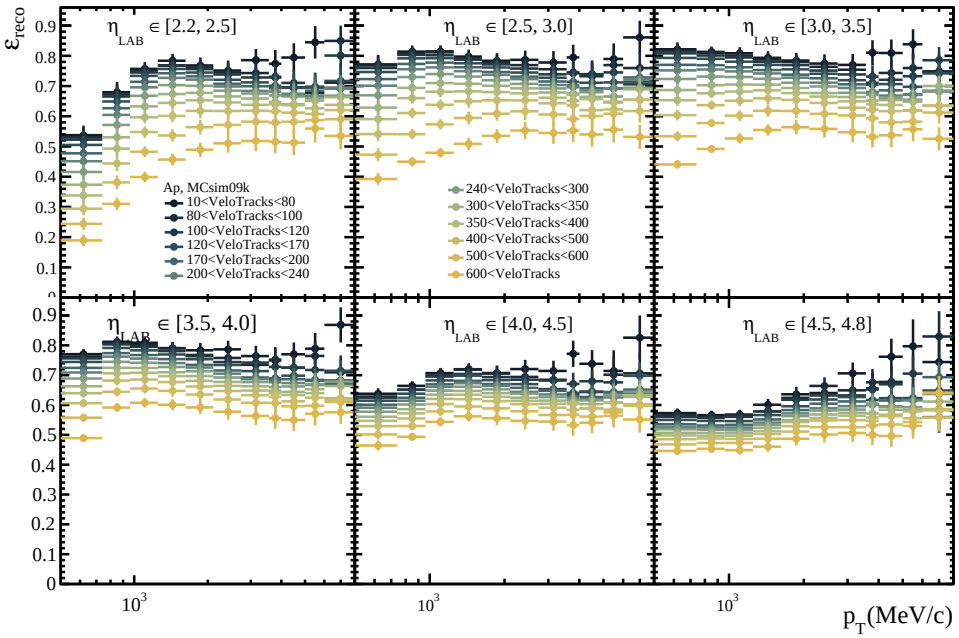


Figure A.2: Eficiencia de reconstrución para datos de colisións Pbp.

A eficiencia de seleccion corrixe pola fraccion de partıculas cargadas primarias eliminadas da mostra de candidatos pola seleccion. Esta estimase usando simulacion, o resultado para pPb pode observarse na Fig. A.3.

Ainda despois de filtrar a nosa mostra de trzas usando a selecion da Tab. A.2 quedan trzas na mostra que non se corresponden ca sinal que queremos medir. Para substraer esta contribucion da nosa mostra estimamos a contribucion do fondo a mostra de sinal, esto denominase pureza da mostra de sinal. O calculo da pureza realizase dividindo o fondo en clases; partículas secundarias (sec), trazas falsas (fake) e trazas duplicadas (clone). Para cada unha destas clases, α, estimase a proporcion de fondo na mostra de sinal usando simulacion, f_α^{sim}. Posteriormente esta fracion corrixese usando unhos factores R_α calculados usando datos a partir da expresión,

$$R_\alpha = \frac{\frac{1}{N_{\mathrm{cand}}} \sum_i N_{\alpha,i}^{data}}{\frac{1}{\sum_i w_i N_{\mathrm{cand},i}} \sum_i w_i N_{\alpha,i}^{sim}}. \tag{A.6}$$

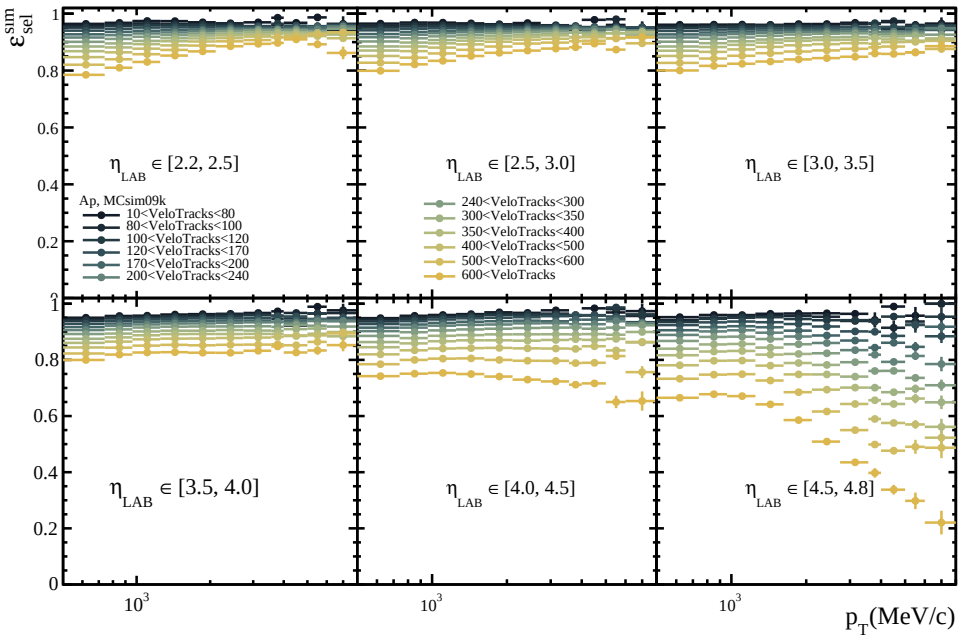


Figure A.3: Eficiencia de seleción para datos de colisions Pbp.

Estos factores corrixen polas diferencias entre a simulacion e os datos e para calculalos e necesario illar cada contribucion ao fondo nos datos. Para isto usamos diferentes algoritmos e propiedades das trazas. Para illar unha mostra de trazas falsas usamos a selecion que se encontra na Tab. A.3.

	pp	pPb, Pbp
Track type	Long	Long
η	$2 < \eta < 4.8$	$2 < \eta < 4.8$
p_{T}	$0.5 < p_{\mathrm{T}} < 8$ GeV/c	$0.5 < p_{\mathrm{T}} < 8$ GeV/c
p	$p > 2$ GeV/c	$p > 2$ GeV/c
`GhostP`	$0.1 <$ `GhostP` < 0.4	$0.5 <$ `GhostP` < 0.9
`IsClone`	True	False
IP	< 0.568 mm	< 0.508 mm

Table A.3: Selecion aplicada para illar unha mostra de trazas falsas nos datos.

Finalmente obtense a fracion de fondo na mostra de sinal para cada clase de fondo, f_α usando a expresión,

$$f_\alpha = R_\alpha f_\alpha^{sim}, \tag{A.7}$$

A pureza definese como

	pp	pPb	Pbp
Common requirements for all sources			
Track type	Long		
η	$2 < \eta < 4.8$		
p_{T}	$0.5 < p_{\mathrm{T}} < 8$ GeV/c		
p	$p > 2$ GeV/c		
`GhostP`	`GhostP` < 0.078	`GhostP` < 0.103	`GhostP` < 0.109
`IsClone`	False		
R_γ specific requirements			
IP	< 0.368	< 0.348	< 0.348
$DLL_{e\pi}$	$>$[15, 20, 25, 30, 35, 40, 45, 50, 55, 60]		
R_μ specific requirements			
IP	< 0.368	< 0.348	< 0.348
$DLL_{e\pi}$	$<$15		
$DLL_{\mu\pi}$	$>$[15, 20, 25, 30, 35, 40, 45]		
$R_{HadDecay}$ specific requirements			
IP	$> [0.1, 0.15, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8]$ mm		

Table A.4: Selecion aplicada para illar os diferentes fondos de partículas secundarias nos datos.

$$\Pi = 1 - (f_{fakes} + f_{sec} + f_{clone}^{sim}). \tag{A.8}$$

Os resultados para a pureza do sinal nas colisións pPb podense observar na Fig. A.4.

Usando a Eq. 6.3 podemos calcular o número de candidatos correxidos por evento. As correcion detalladas anteriormente corrixen os efectos experimentais descritos. Sen embargo, hai outros que non se podes correxir con factores de correcion e requiren de métodos estadisticos. Refirome a os efectos de migracion entre bins, ou do tamaño dos bins no que se calcularos as eficiencias. Para substraer estos efectos utilizamos un Unfolding Bayesiano que utiliza o metodo de D'Agostini. Este metodo utiliza unha matriz de resposta, R, que se construe coa simulacion. Despois, aplicase aos datos usando o algoritmo de D'Agostini implementado no paquete de software RooUnfold. Na Fig. A.5 podemos observar un test de autoconsistencia no que se mostra como a distribución de multiplicidade despois do Unfolding recupera o valor a nivel xerador usando datos de simulacion. Esto quere decir que os efectos de detector que non estaban sendo correxidos agora o estan.

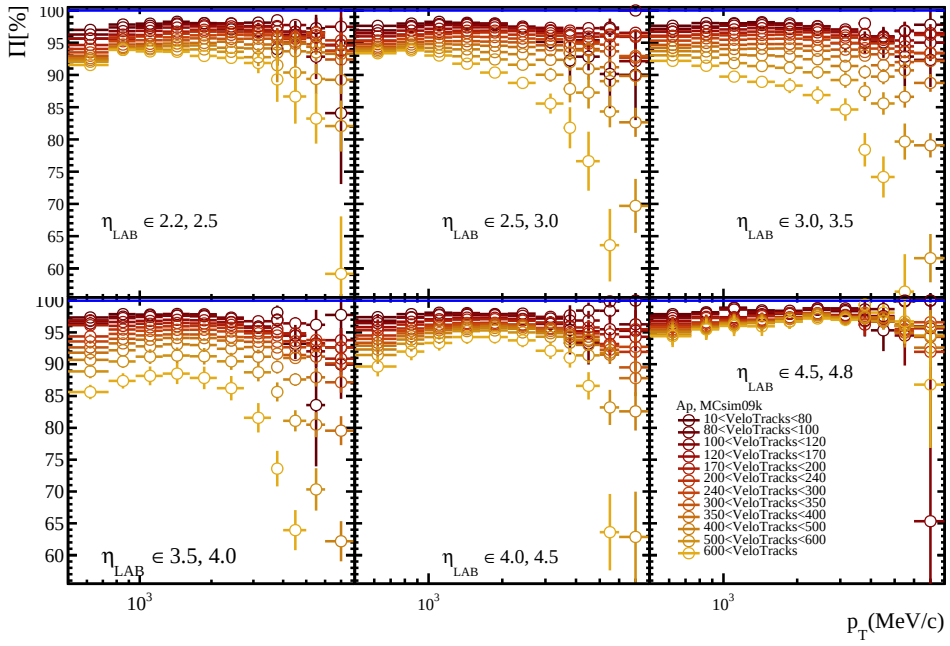


Figure A.4: Pureza da mostra de candidatos das colisions Pbp.

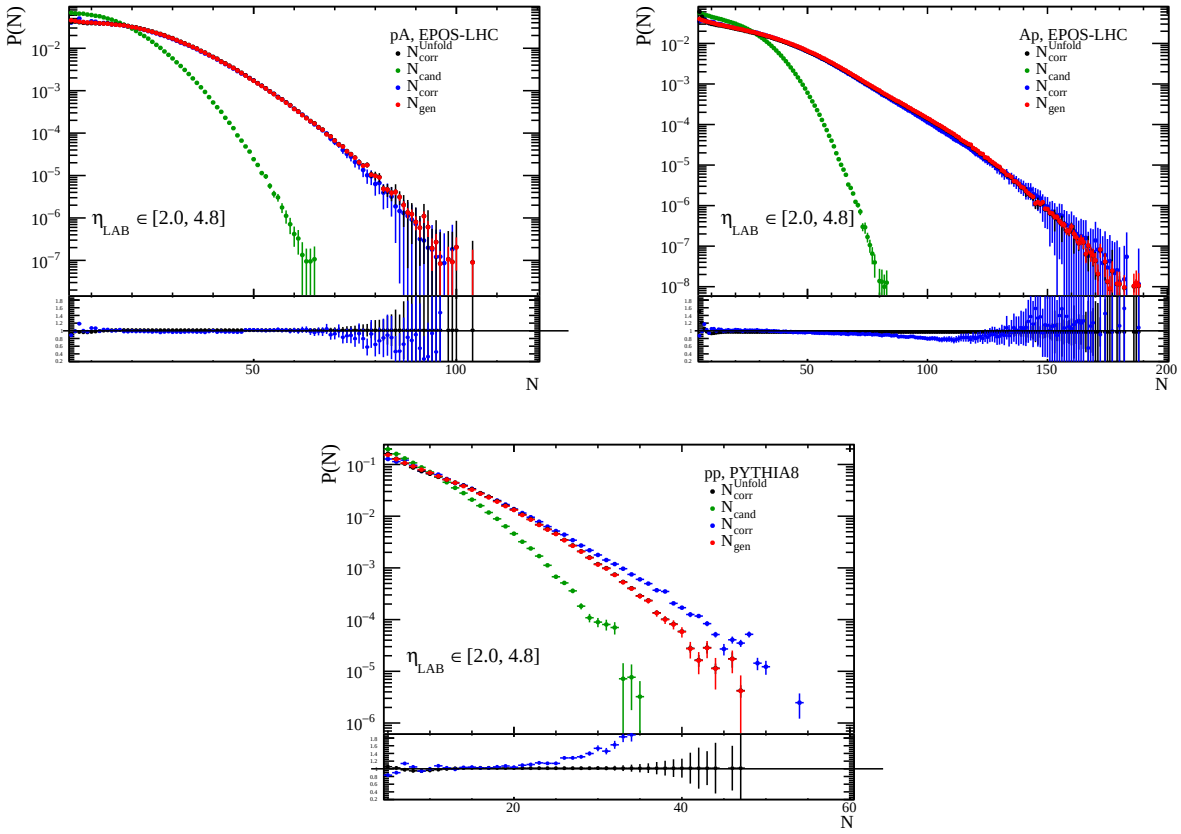


Figure A.5: Closure test usando a mostra de simulación MCsim09kMultiplicityDplus de colisions pPb (esquerda), Pbp (dereita) e pp (abaixo). A gráfica inferior conten a fración con respecto ao espectro xerado N_{gen} (en vermello para a dsitribución N_{corr}^{Unfold} (negro) and N_{corr} (azul) distribution).

A.4 Medida da dependencia coa multiplicidade do momento transverso promedio

Para a medida da dependencia coa multiplicidade do momento transverso promedio usamos o número de candidatos correxidos N_{corr} calculado na seccion anterior como clasificador de eventos. Unha vez clasificados os eventos podemos calcular o $\langle p_{\text{T}} \rangle$ usando a Eq. 6.1 para cada bin de η. Nas Figs. A.6 podese observar un test de autoconsistencia usando datos da simulacion. En dita gráfica comparase o $\langle p_{\text{T}} \rangle$ a nivel xerador co $\langle p_{\text{T}} \rangle$ reconstruido unha vez pasado aplicados todos os pasos do análise.

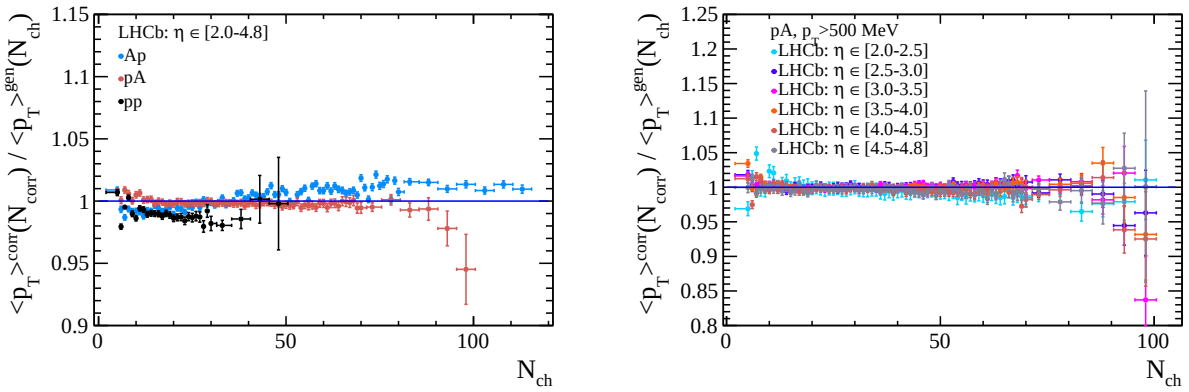


Figure A.6: Test de consistencia para todos os sistemas de colison (esquerda) sin dividir en bins de η. Na dereita amósase un exemplo do mesmo test em colisions pPb subdividindo en rangos de η.

A.5 Resultados e discusión

Nesta seccion mostranse os resultados finais da tese. Primeiro na Sec. A.6 mostranse os resultados da distribucion de multiplicidade, tanto integrada en η como en diferentes bins. Ademais, na Sec. A.6.1 comparanse os resultados obtidos con modelos Monte-Carlo. Despois, na Sec. A.7 ensinanse os resultados para o $\langle p_{\text{T}} \rangle$. Tamen se comparan estoas resultados cos odelos na Sec. A.7.1.

A.6 Distribución de multiplicidade

Os resultados da distribución de multiplicidade en todo o rango de pseudorapidez pódense ver na Fig. A.7. Como cabería esperar, o sistema de colisión mais grande pPb, Pbp produce un número de partículas maior que pp. Ademais, o resultado correspondente a Pbp mostra unha multiplicidade maior que pPb. Esto e esperable xa que estamos a mirar a rexión de fragmentación do núcleo de chumbo, que ten moitas mais nucleóns.

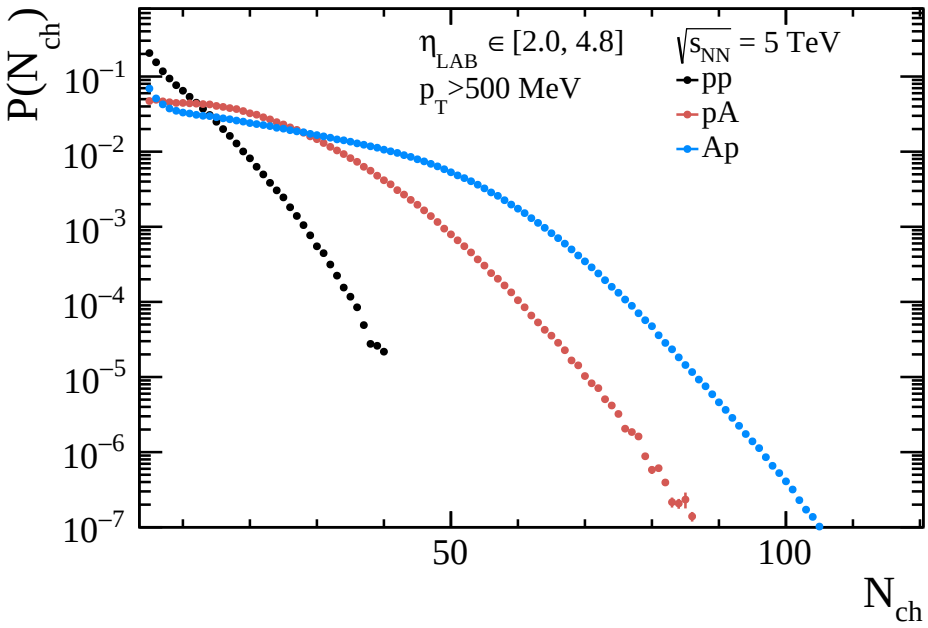


Figure A.7: Distribución de multiplicidade para todos os sistemas de colision integrado en pseudorapidez.

A.6.1 Comparación cos modelos

Nas Fig. A.8 mostránse a comparacións dos resultados cos modelos para o resultado integrado en pseudorapidez. Ademais na Fig. A.9 amósase o cociente entre a predicción e o resultado da tese. Os modelos non describen de xeito adecuado os datos experimentais, como ocorre noutras medidas como por exemplo Ref. [5]. A discrepancia que se observa e similar en todos os sistemas de colision. Os modelos subestiman o resultado a baixa e alta multiplicidade, e sobreestimano en multiplicidades intermedias. A discrepancia co modelo e mais grande no caso das colisions pPb e Pbp. O modelo, EPOS-LHC, discrepa cos datos experimentais da mesama maneira en ambos sistemas de colision ata chegar a multiplicidades altas, $N_{ch} \sim 60$, donde o sistema pPb e o peor reproducido polo modelo.

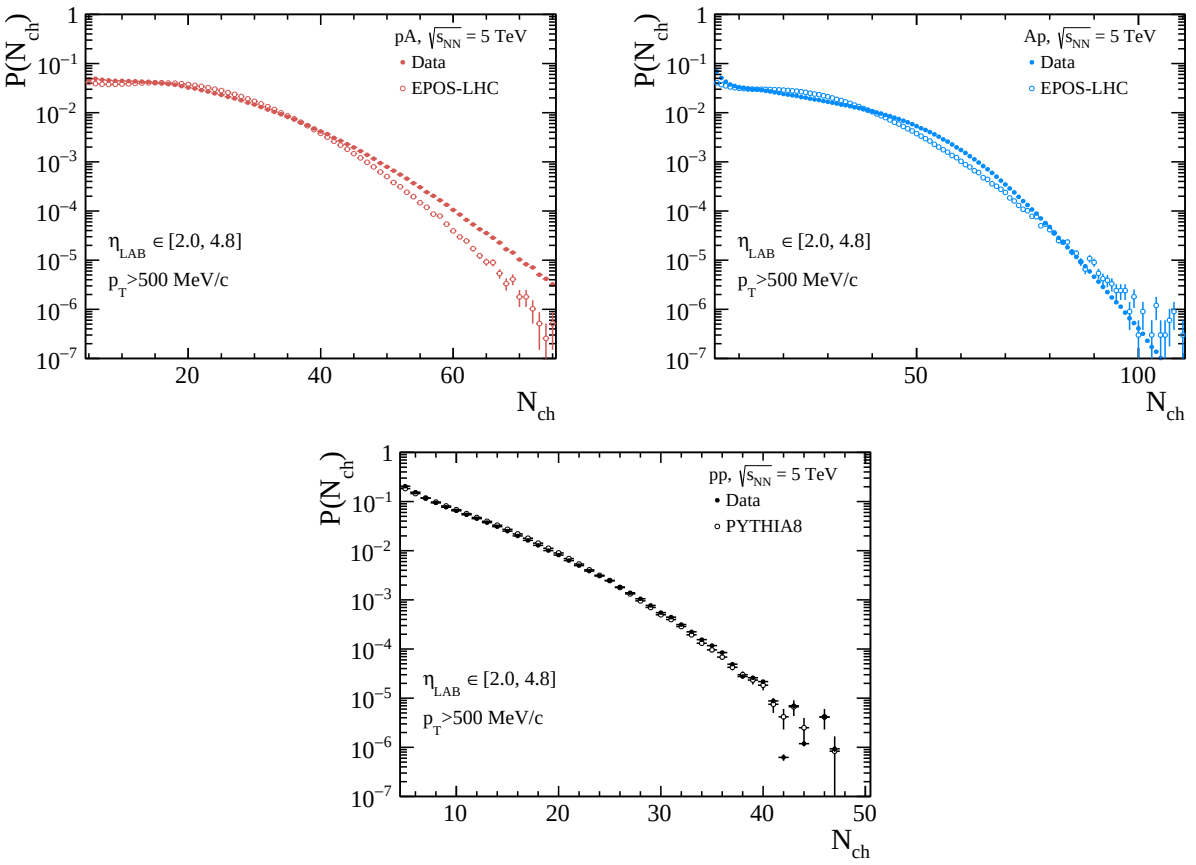


Figure A.8: Comparación cos modelos Monte Carlo dos resultados da distribución de multiplicidade para colisions *p*Pb (arriba esquerda), Pb*p* (arriba dereita) e *pp* (abaixo). Resultado usando todo o rango de pseudorapidez do experimento LHCb.

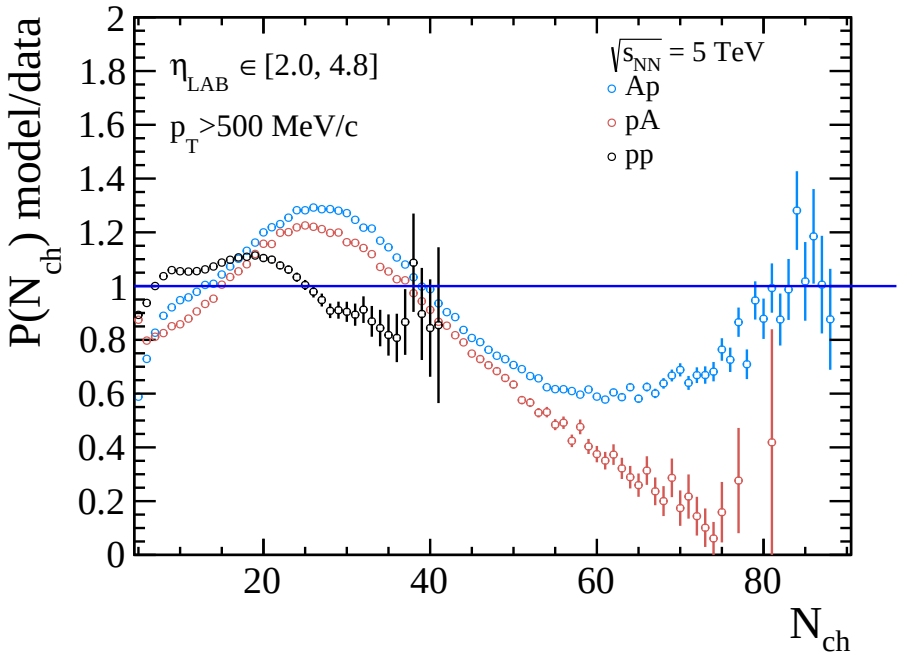


Figure A.9: Cociente da distribucion de multiplicidade entre a predicción do modelo e o resultdo deste traballo. Resultado correspondente a toda a aceptancia de LHCb.

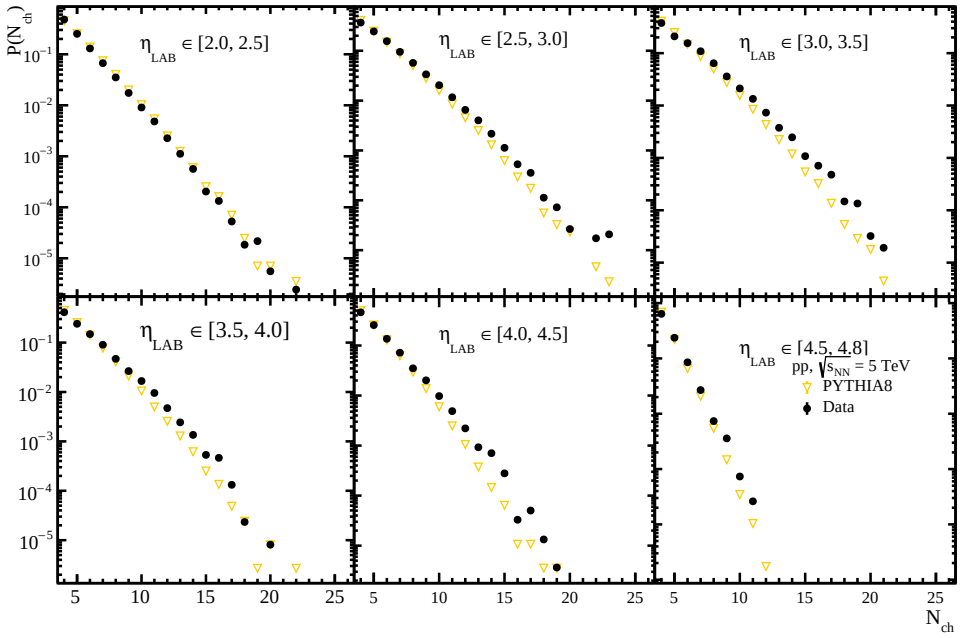


Figure A.10: Distribución de multiplicidade das colisions *pp* en bins de η

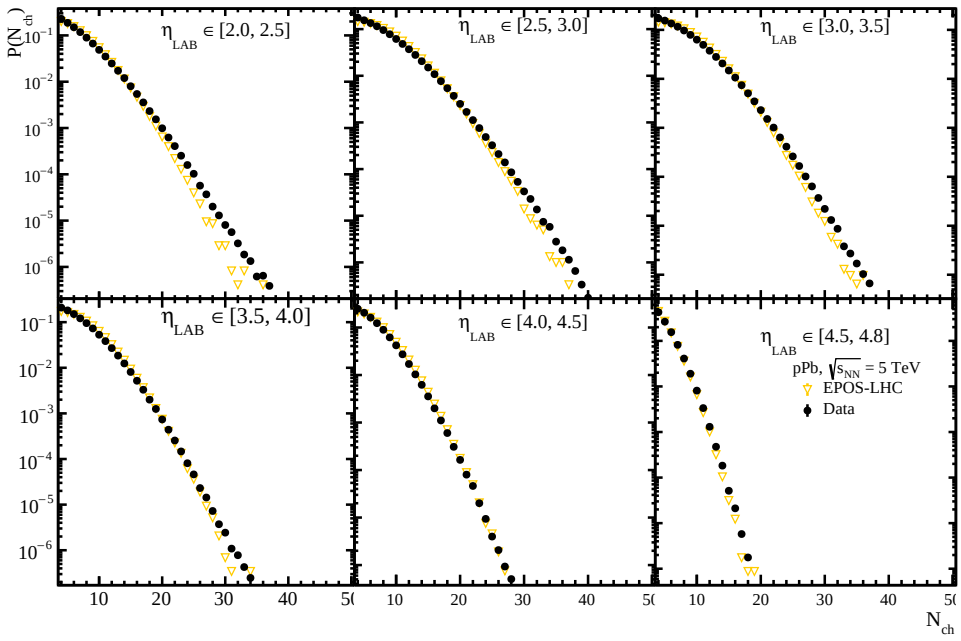


Figure A.11: Distribución de multiplicidade das colisions pPb en bins de η

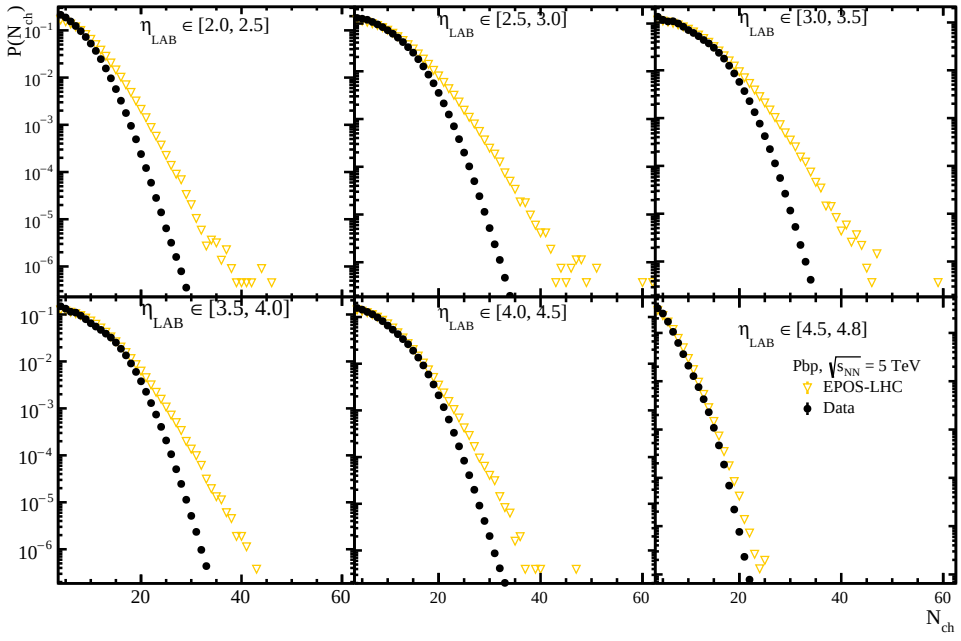


Figure A.12: Distribución de multiplicidade das colisions Pbp en bins de η

A.7 Momento transverso promedio

O resultado integrado en pseudorapidez para o $\langle p_{\mathrm{T}} \rangle$ como funcion da multiplicidade en diferentes sistemas de colision podese ver na Fig. A.13. O resultado en diferentes bins de pseucorapidez podese ver na Fig. A.14. Neste resultado, non se mostra o bin mais central correspondente ao intervalo $\eta_{LAB} \in [2.0, 2.5]$ xa que se observou unha desviación sistemática de un 2% no closure test na Fig. A.6.

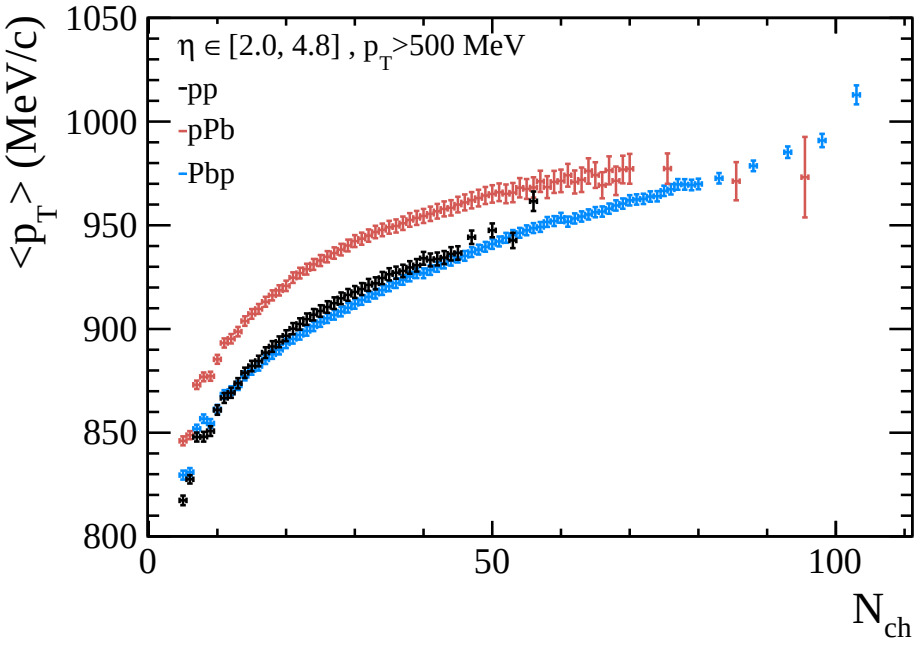


Figure A.13: Momento transverso promedio como funcion da multiplicidade para collisions pPb, Pbp e pp. Resultado integrado en pseudorapidez.

Observase un incremento do $\langle p_{\mathrm{T}} \rangle$ con respecto a multiplcidade en todos os sistemas de colision. Sin embargo, mentras que en pPb e pp o valor satura nun valor maximo aredor de 950 MeV/c, esot parece non ocorrer nos datos das colisions Pbp.

En relación a medida dependende da pseudorapidez observase un decrecemento do valor do $\langle p_{\mathrm{T}} \rangle$ con η. Esta tendencia e maior a altas multiplicidades, sem ebaego a baixas a diferencia entre rangos de η e menor. Nas colisions Pbp observase a mesma tendencia crecente a altas multiplicidades que no resultado integrado. Sen embargo, o incremento do $\langle p_{\mathrm{T}} \rangle$ a altas multiplicidades parece estar limitado a o rango de η mais baixo. Esto quere decir que ocorre nas rexións mais centrais.

A.7.1 Comparacion cos modelos

De acordo cos estudos recents da colaboración ALICE [5], EPOS-LHC subestima o $\langle p_{\mathrm{T}} \rangle$ nas colisions pPb nun $\simeq 5\%$. Ademais, PYTHIA 8.3 sobreestima o valor do $\langle p_{\mathrm{T}} \rangle$

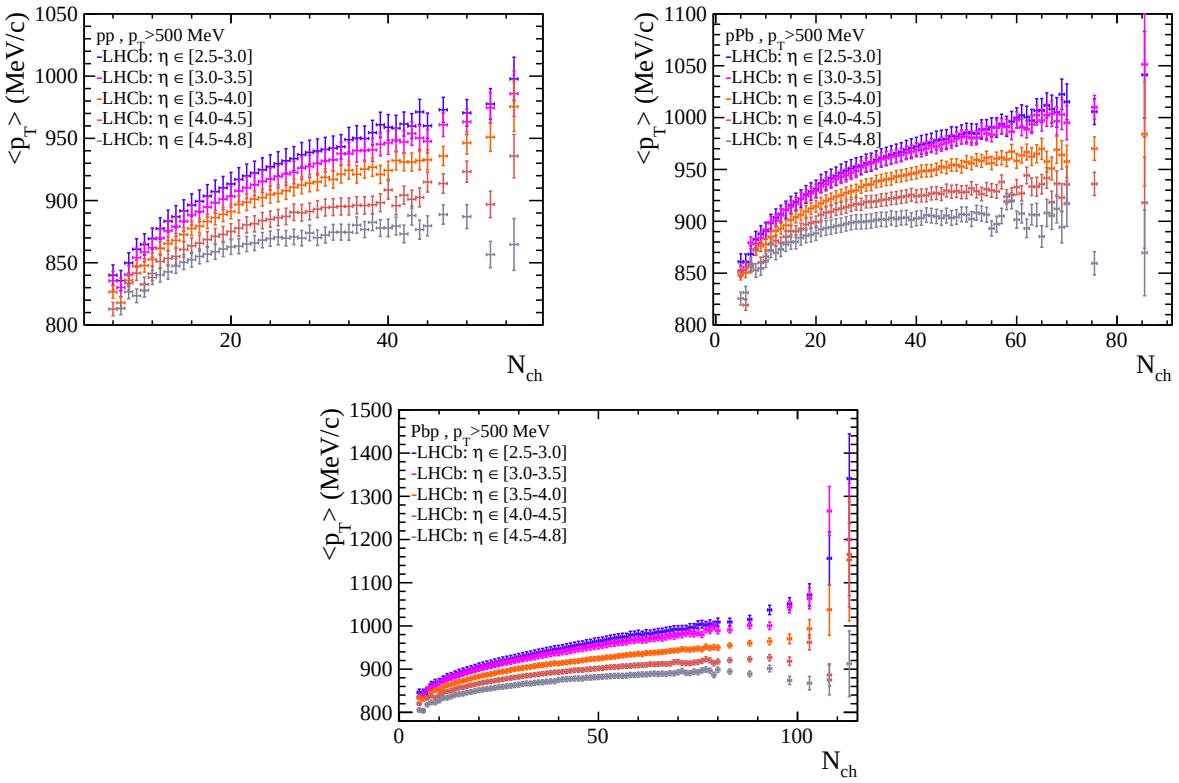


Figure A.14: Resultado do momentto transverso promedio como funcion da multiplicidade en diferentes bins de η. Mostranse os resultados para colisions pp, pPb e Pbp.

nas colisions pp nun $\simeq 4\%$. Segundo os resultados desta tesis Fig. A.15 podemos chegar a conclusions similares. Segundo o que podemos ver na Fig. A.15 EPOS-LHC subestima o $\langle p_{\mathrm{T}} \rangle$ en colisión pPb nun 5%. Ademais nas colisions pp o modelo implementado por PYTHIA sobrestima o $\langle p_{\mathrm{T}} \rangle$ nun 10%. Nestes dous sistemas de colision a tendencia do $\langle p_{\mathrm{T}} \rangle$ esta ben repdocusida polos modelos, a diferencia rapica no valor do $\langle p_{\mathrm{T}} \rangle$. A principal diferenza atopamola nas colisions Pbp donde a tendencia coa multiplicidade non está ben reproducida. Aqui atópanse diferenzas dun $5\% - 15\%$.

A.7.2 Caracterizando o estado inicial nas colisións hadrónicas

Na medida do $\langle p_{\mathrm{T}} \rangle$, utilizamos N_{corr} como estimador da multiplicidade de cada evento. Na literatura diferentes estimadores foron utilizados ao longo dos anos para facer este tipo de medidas. Na Fig. A.16 facemos unha comparacion da relacion coa variable obxetivo, N_{ch}, de diferentes estimadores. Como se pode observar, o unico que e lineal con N_{ch} e N_{corr}. Ademais, podese observar unha desviacion da linealidade a alta multiplicidade. Ainda que estos eventos son raros, para correxir este comportamento no estimador poderíase incrementar a estadistica da simulacion para poder calcular as correcions con un bineado

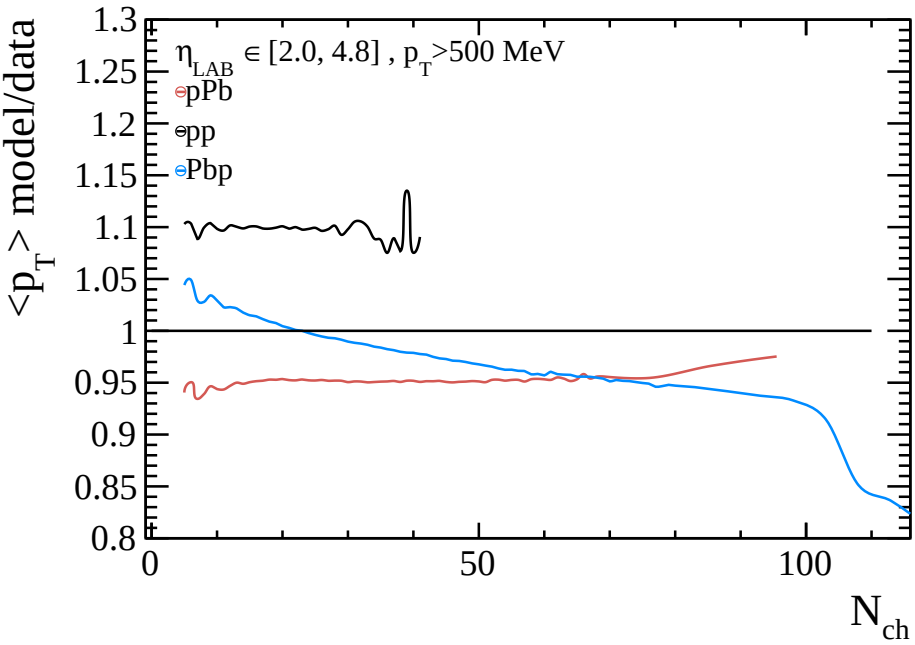


Figure A.15: Ratio do resultado do $\langle p_T \rangle$ cos modelos teóricos. Para pp o modelo e PYHTIA 8, mentras que para pPb e Pbp o modelo e EPOS-LHC.

mais fino a alta multiplicidade.

A.7.3 Proxeccions para o Upgrade 2

Ademais da medidad presentada, e do estudo dos estimadores na caracterizacion do estado inicial nas colisións hadrónicas tamen se realizaron algunhas proxecions para colisións PbPb no LHCb.

O experimento LHCb realizara unha actualizacion dos seus sistemas chamado Upgrade 2. Grazas as melloras que serna introducidas, esperase que o detector sexa capaz de medir colisións centrais de PbPb a enerxias da orde de TeV no Run5 (ano 2035). Neste rexime formase o QGP e podemos estudar a ecuación de estado da QCD no rexime donde os parton son os grados de liberdade do sistema. Nesta tese, fixose un calculo do rango de temperaturas accesible ao experimento LHCb utilizando a EoS calculada con Lattice QCD, medidas experimentais do $dN/d\eta$ e a relacion seguinte

$$\frac{dN}{dy} = \frac{dN}{dy}\Big|_{y=0} \cdot \frac{s}{20\,\mathrm{fm}^{-3}}. \tag{A.9}$$

Grazas a isto, podese relacionar un rango de η con un rango de temperaturas como se pode observar na esquerda da Fig. A.17. Ademais, podemos comparar o potencial do experimento LHCb no canto a media propiedades da EoS nun ranog de temperaturas en

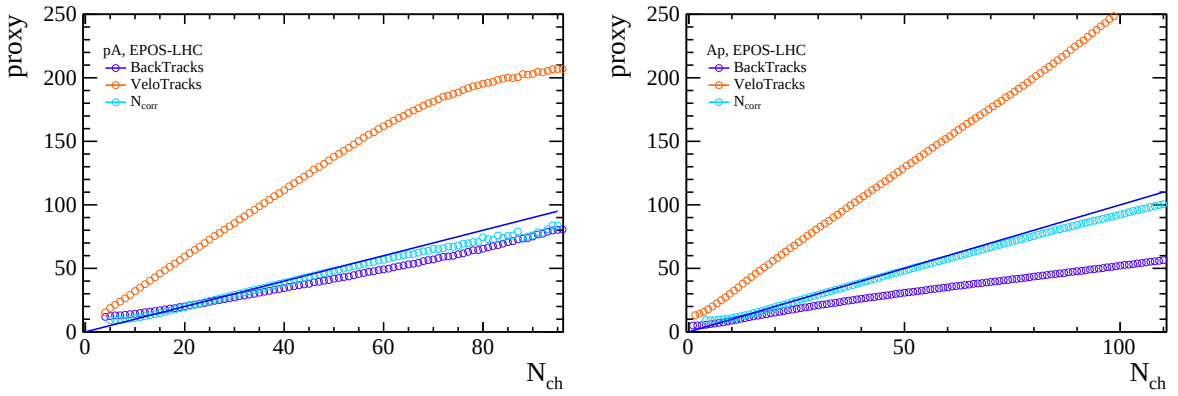


Figure A.16: Comparacion de varios estimadores de multiplicidade normalizados coa multiplicidade a nivel xerador. A linea azul e unha linea recta de pendiente unidade que ten como obxetivo guiar ao ollo. A grafica da esquerda correspondese con colisións pPb e a da dereito con Pbp.

comparacion con outros experimentos. Na dereita da Fig. A.17 pode observarse o rango accesible a LHCb en comparación co acesible a un experimento da rexión central como CMS.

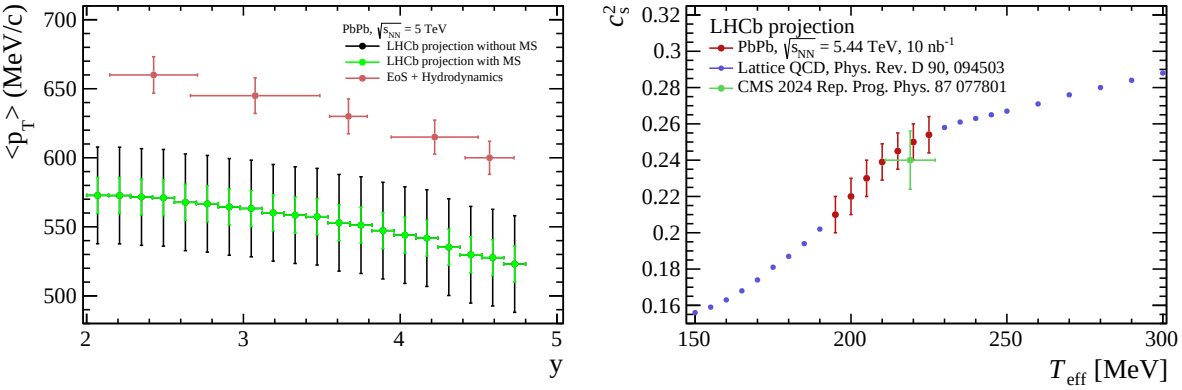


Figure A.17: (left) Predicion teórica para $\langle p_{\mathrm{T}} \rangle(y)$ usando EoS+Hydro.en comparación coa predición experimental. (right) Calculo de Lattice QCD para c_s^2 extraido de [25] no rango de temperaturas estimado para LHCb en comparacion coa medida experimental do experimento CMS [106].

Neste estudo inclúese tamén unha estimación das incertezas asociadas ao cálculo teórico e a medida experimental. Con respecto ao calculo teórico amosado Fig. A.17 esquerda en vermello tense en conta unha incerteza dun 2% correspondente ao cálculo da temperatura na Ref. [25]. Ademais, no eixo X dese mesmo cálculo incluese a interdeza debido a medida experimental do $dN/d\eta$ da Ref. [103]. Despois, na proxeción experimental na mesma figura inclúses a incerteza asociada a extrapolación do espectro de p_{T} ata

$p_{\mathrm{T}} = 0\,\mathrm{MeV}$. Como se detalla na Ref. [101] no Upgrade II esta planificado instalar un detector de trazas no imán. Este subdetector denóminase Magnet Station (MS) e estará especialmente dedicado a reconstruír trazas de baixo momento transverso. O cálculo da incerteza da extrapolación fíxose tendo en conta dous valores de p_{T}^{min}. Por un lado, no caos de contar coas Magnet Stations (MS) o p_{T}^{min} podería ser de 100 MeV. Sen embargo, sen as MS o p_{T}^{min} sería de 500 MeV. Como se pode observar na figura, contar coas MS diminuiría a incerteza nun $35-40\%$ nesta medida. Con respecto a predición experimental da velocidade do sonido, tense en conta a incerteza asociada ao axuste do que se extrae o valor de c_s^2. O valor da incerteza extráese da Ref. [106].

B

Additional figures

B.1 Cross-check: Dependency of simulation occupancy distribution and occupancy binning in the corrections

The candidate p_{T} spectra are corrected using a specific binning for efficiencies and purity. However, for multiplicity this binning is wider than the one used in the final measurement, it is worth studying if the result depends on the binning scheme in this variable that is used in the corrections.

Moreover, we have two simulation samples (MCsim09e and MCsim09k), for pPb and Pbp, where the difference is the multiplicity distribution and binning. The simulation occupancy variables are reweighted to match the data distribution so in principle this wouldn't cause any difference. However, the binning in multiplicity that we use in MCsim09e is different from the one used for MCsim09k. This is because the latter has an event in the high multiplicity region while the other doesn't. As can be seen in Fig. B.1 the difference between using corrections from both simulation samples in the final result is less than 1% for pPb and less than 3% for Pbp data.

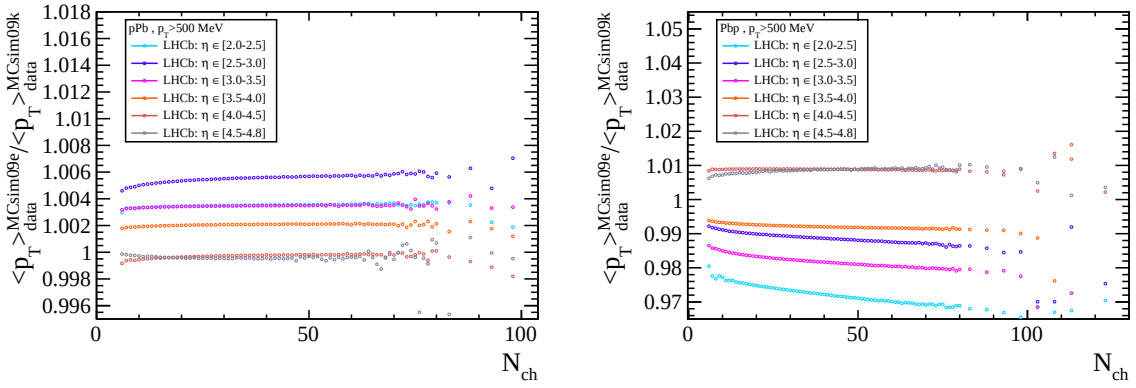


Figure B.1: Ratio of the average transverse momentum result using corrections from MCsim09e simulation and MCsim09k simulation for pPb (left) and Pbp (right) data.

B.2 Occupancy reweighting of simulation samples

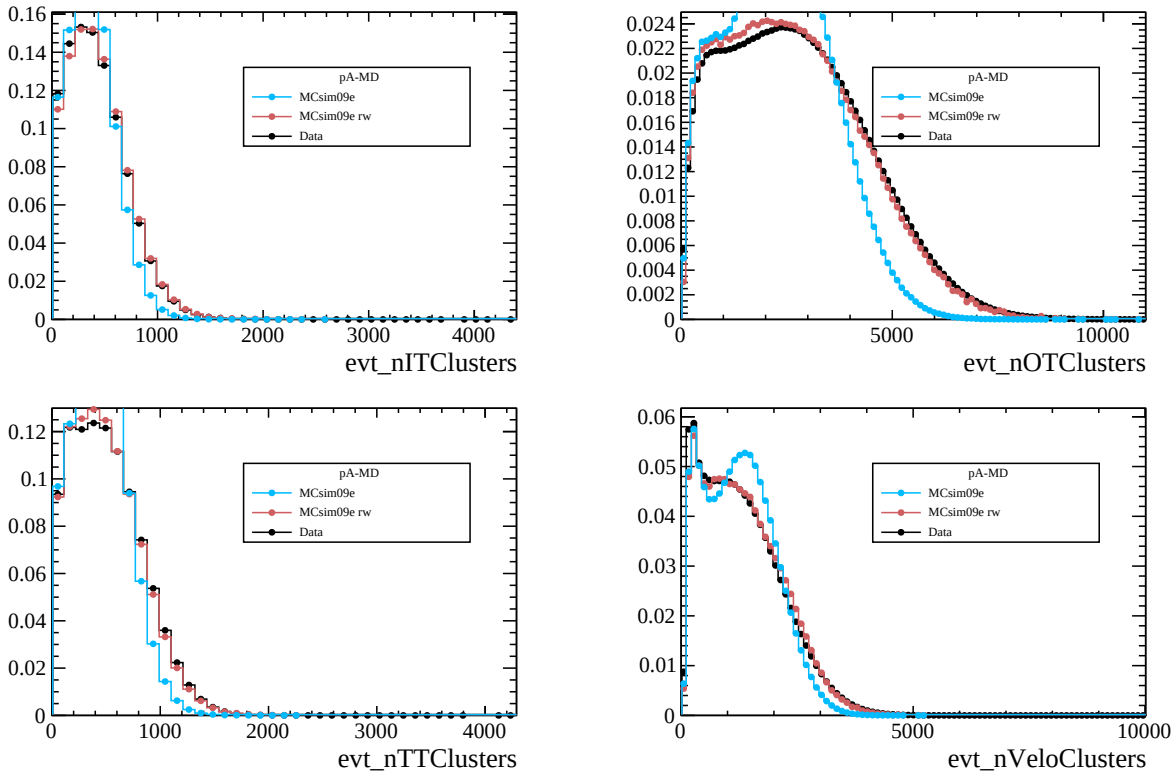


Figure B.2: Occupancy distributions before and after applying the weights for pPb collisions with MCsim09e sample.

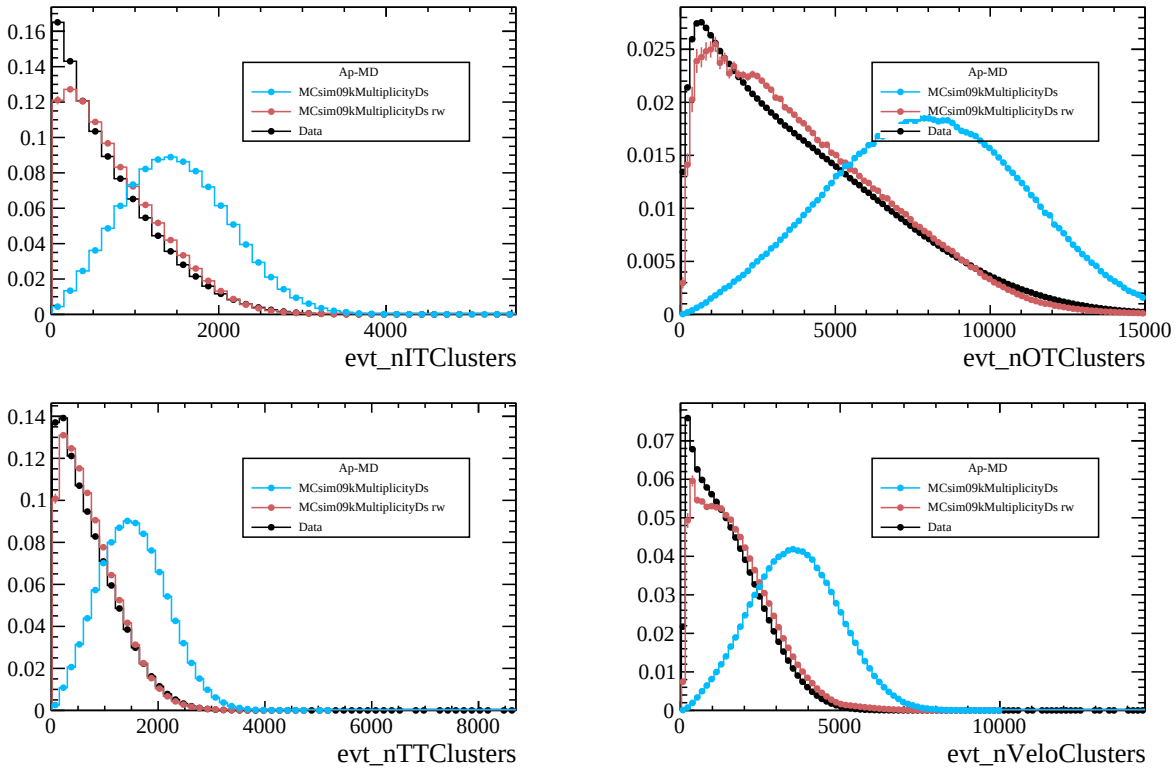


Figure B.3: Occupancy distributions before and after applying the weights for Pbp collisions with MCsim09kMultiplicity sample.

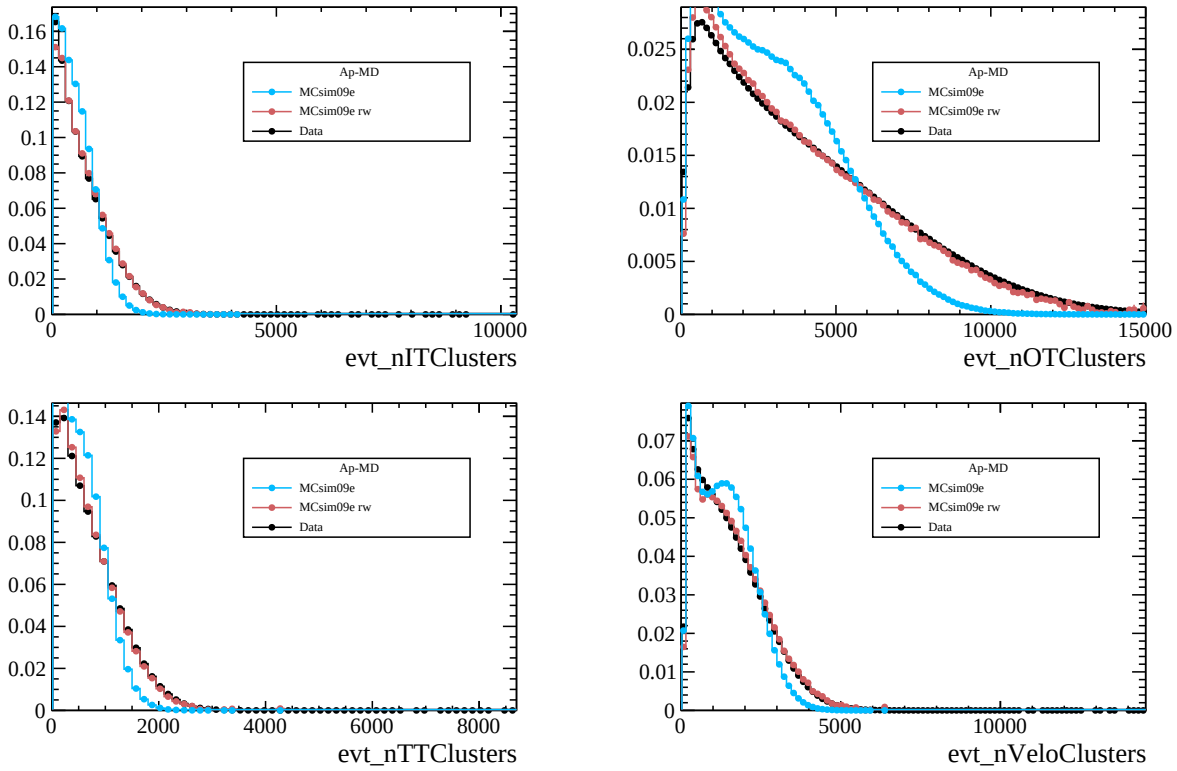


Figure B.4: Occupancy distributions before and after applying the weights for Pbp collisions with MCsim09e sample.

B.3 Closure-test in pseudorapidity bins

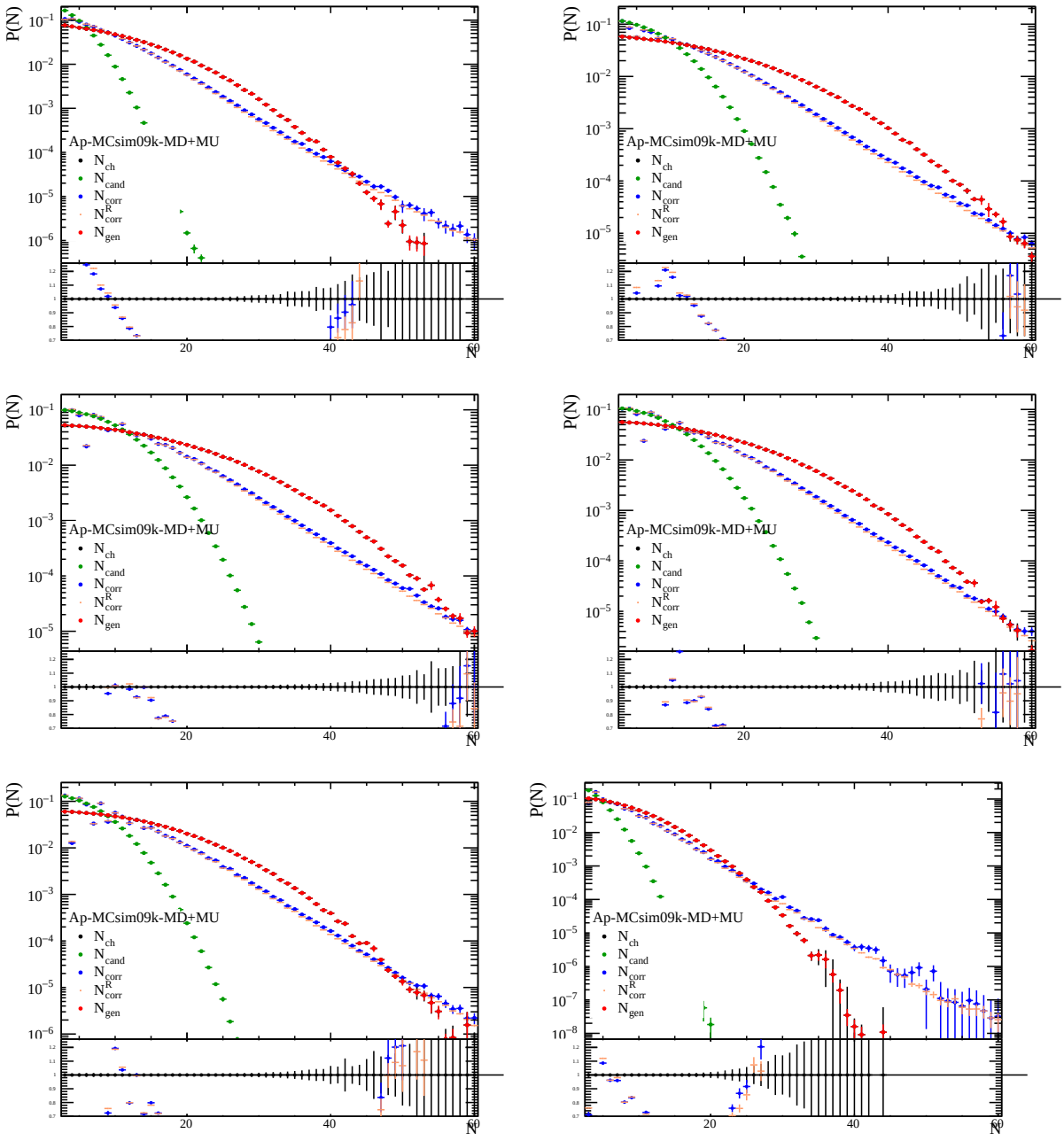


Figure B.5: Closure-test by pseudorapidity bin for the Pbp sample.

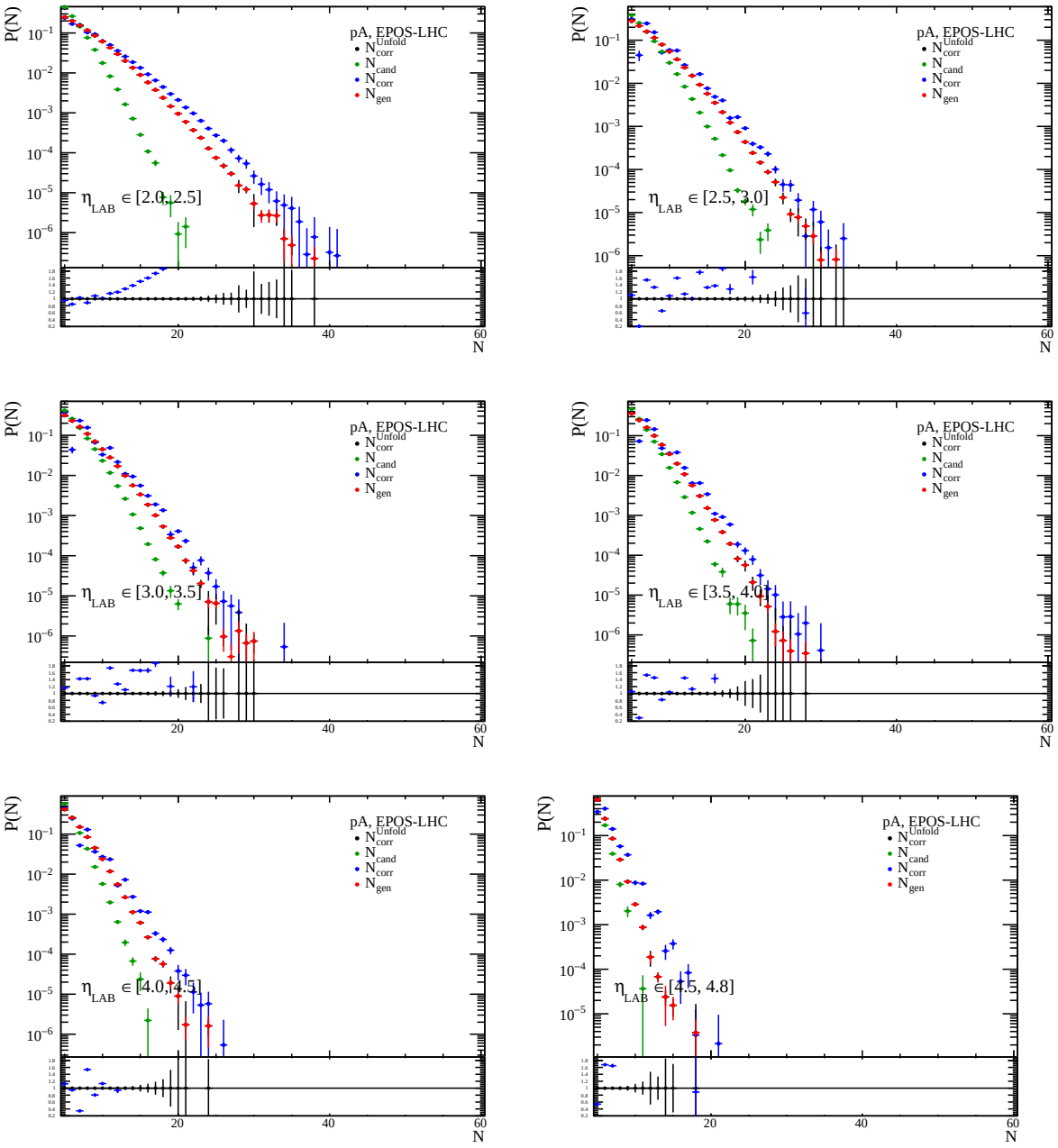


Figure B.6: Closure-test by pseudorapidity bin for the pPb sample.

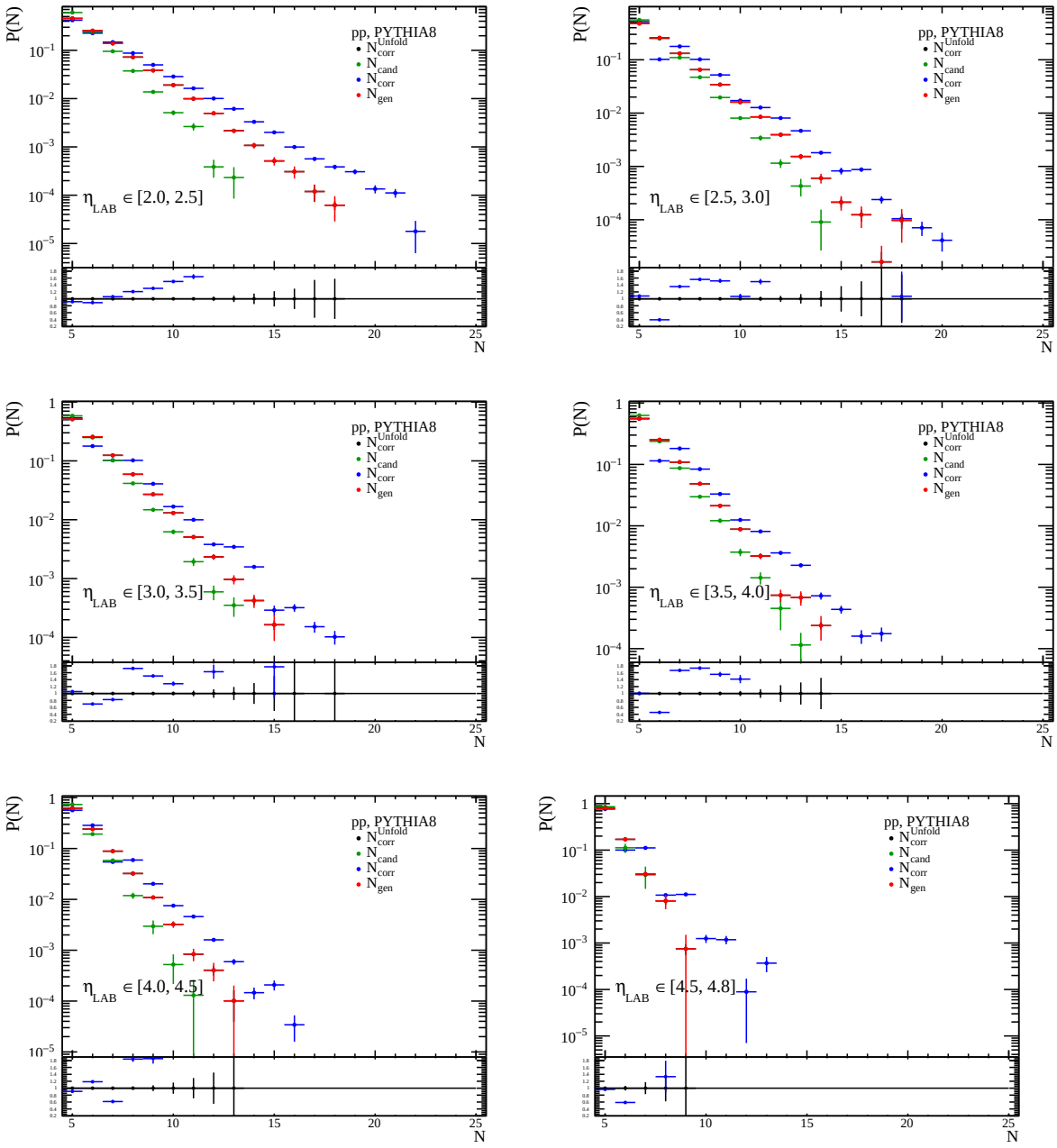


Figure B.7: Closure-test by pseudorapidity bin for the *pp* sample.

B.4 Reconstruction efficiency

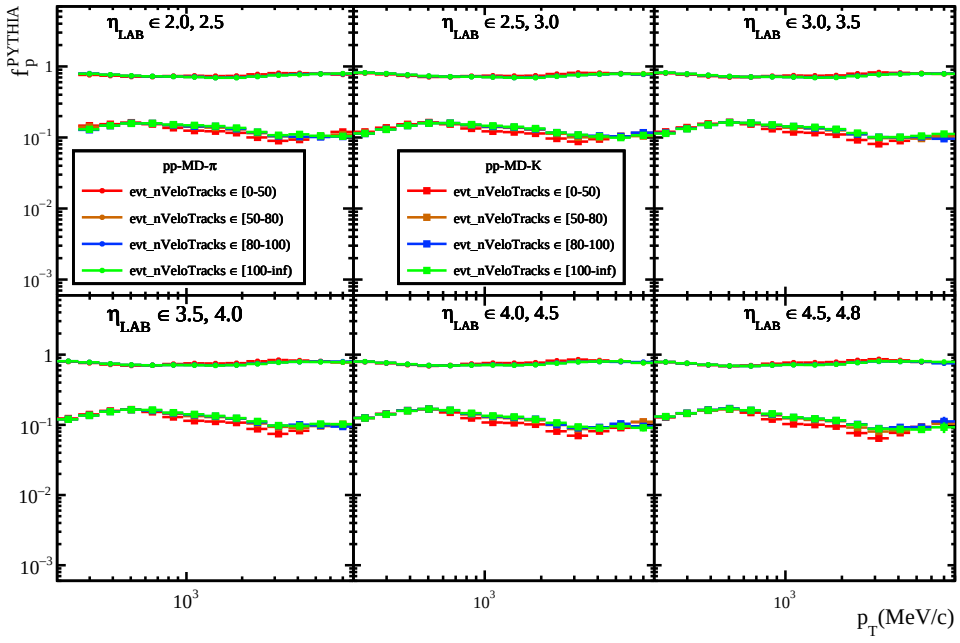


Figure B.8: Fraction of pions (dots) and kaons (squares) as a function of $(p_{\mathrm{T}}, \eta, \texttt{nVeloTracks})$ bins in pp collisions from PYTHIA8.

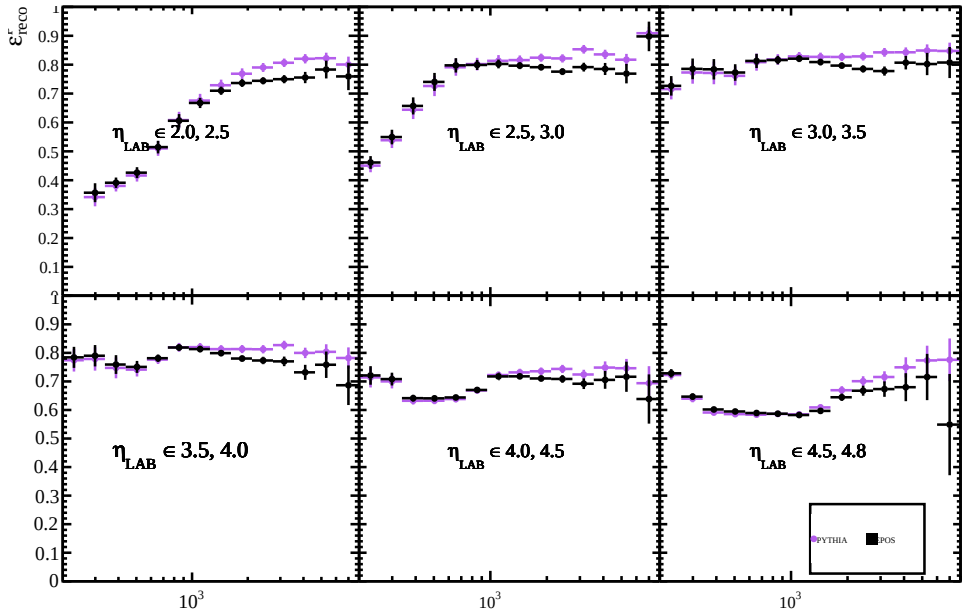


Figure B.9: Prompt charged-particle reconstruction efficiency using a fraction of species abundances from EPOS and PYTHIA8 generators. The one using EPOS species fractions (black) is the one used in this analysis.

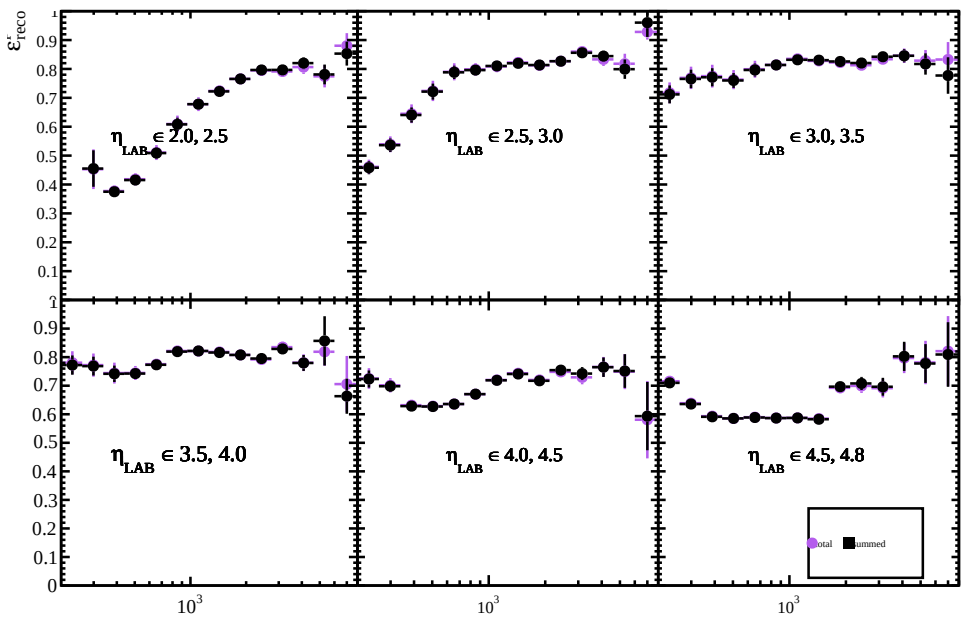


Figure B.10: Cross check. If the sum over all species recovers the total reconstruction efficiency.

B.5 Signal purity

B.5.1 Background fractions for secondary particles per kind

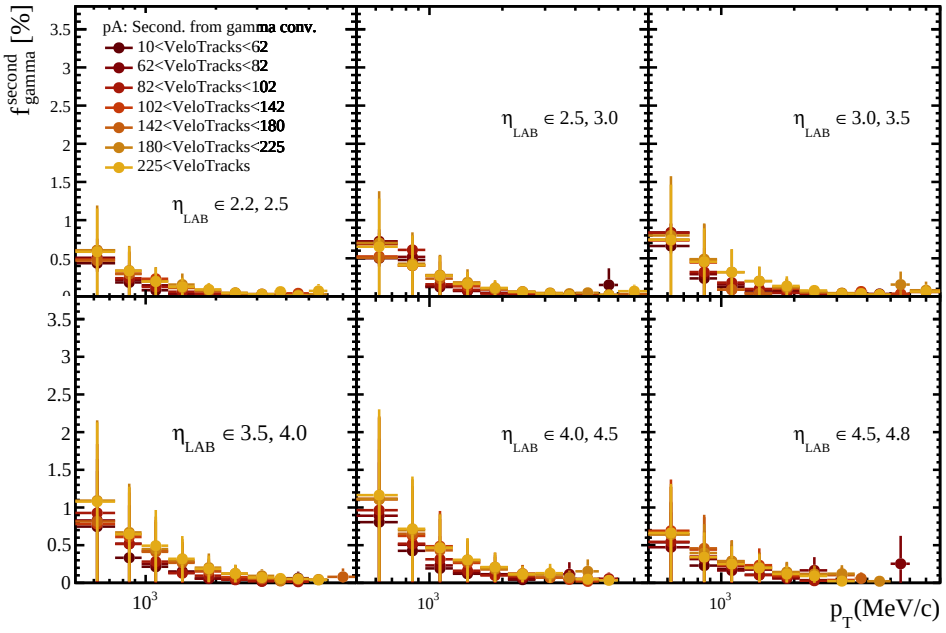


Figure B.11: f_γ for pPb collisions.

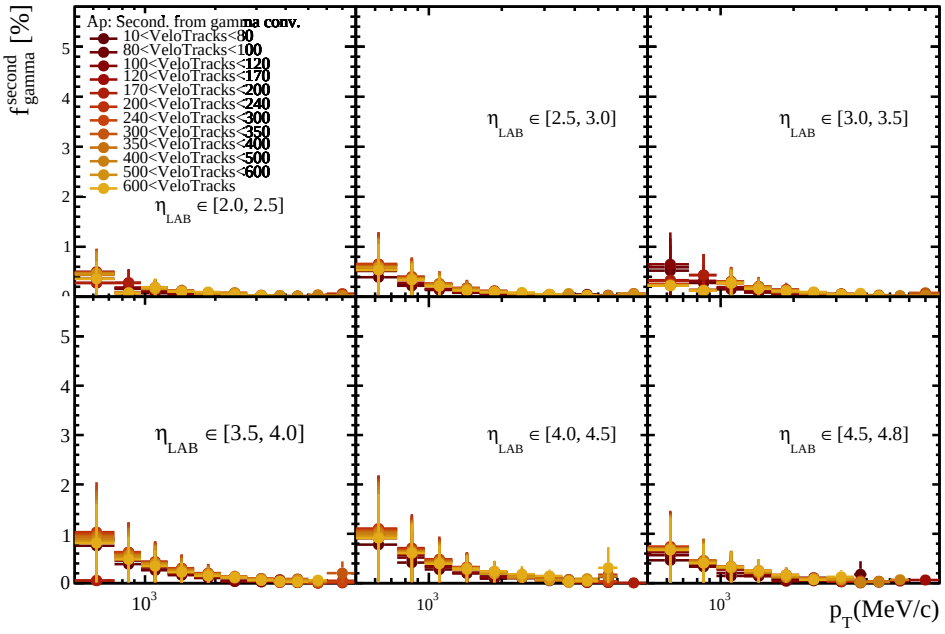


Figure B.12: f_γ for Pbp collisions.

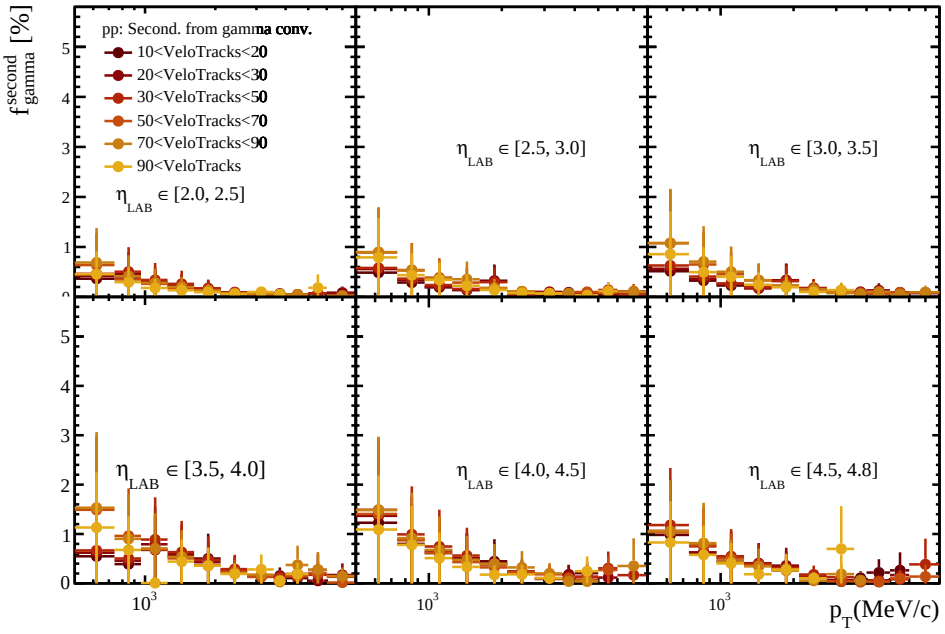


Figure B.13: f_γ for pp collisions.

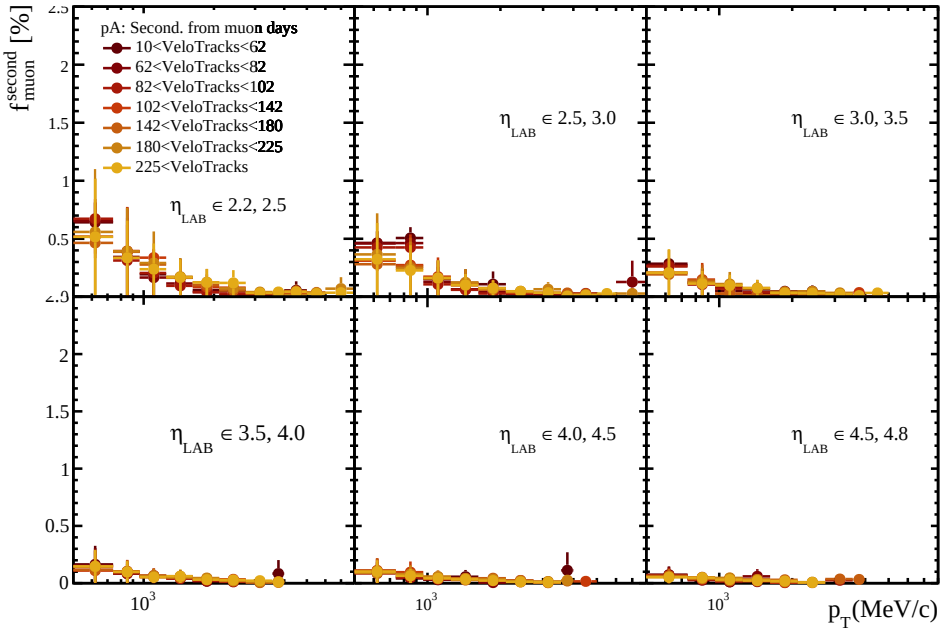


Figure B.14: f_μ for pPb collisions system.

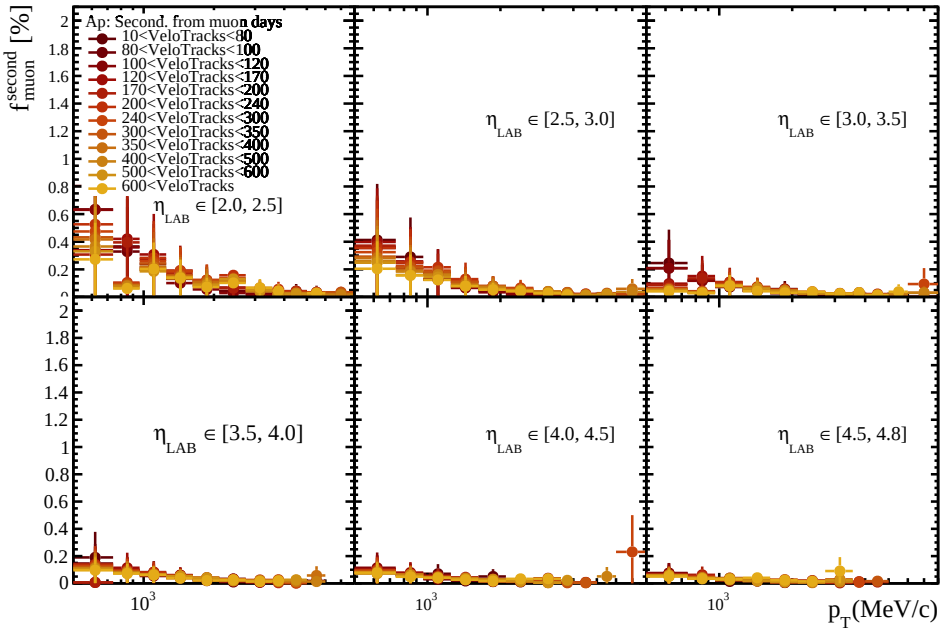


Figure B.15: f_μ for Pbp collisions system.

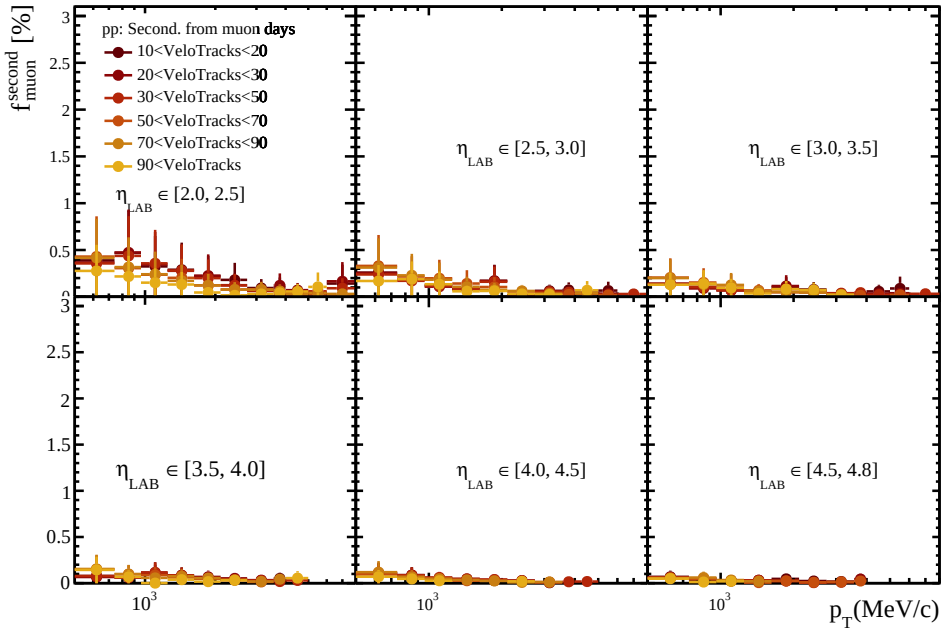


Figure B.16: f_μ for pp collisions system.

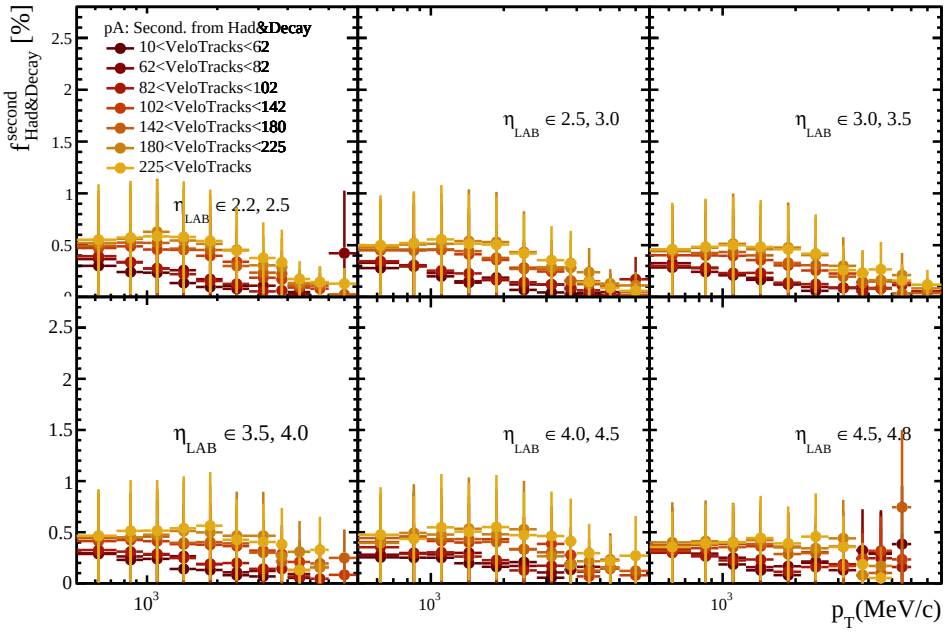


Figure B.17: $f_{Had\&Decays}$ for pPb collisions system.

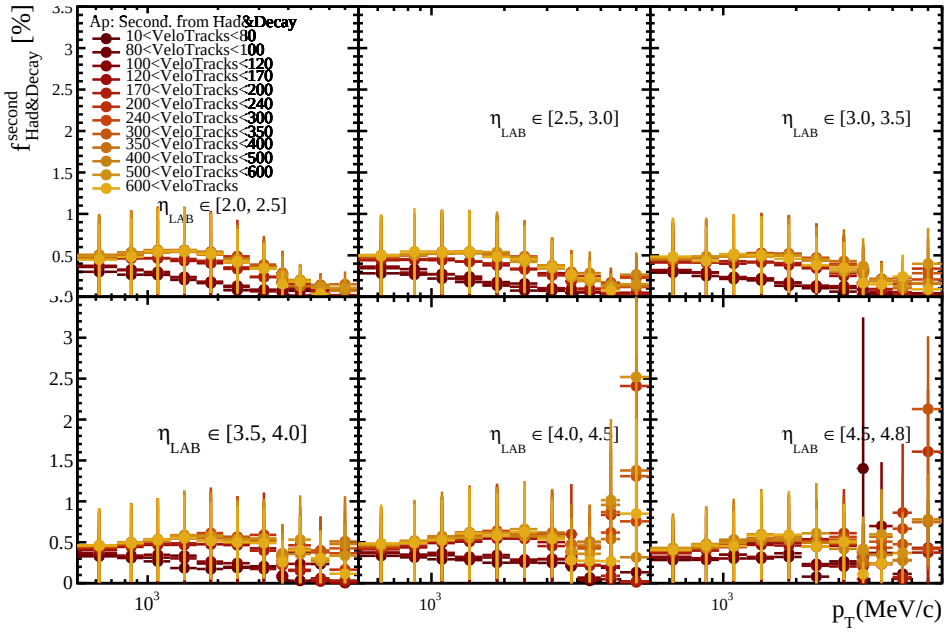


Figure B.18: $f_{Had\&Decays}$ for Pbp collisions system.

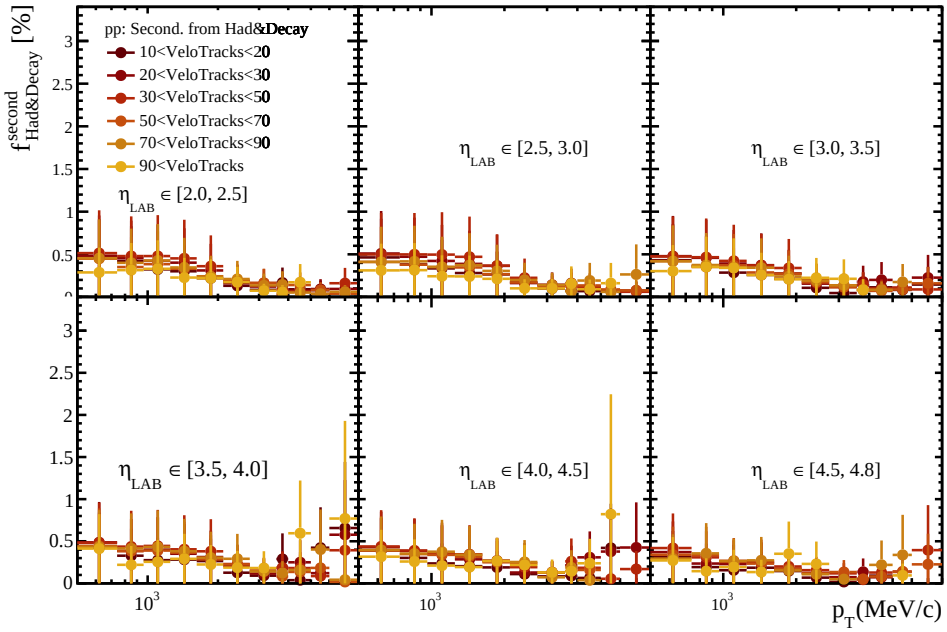


Figure B.19: $f_{Had\&Decays}$ for pp collisions system.

B.5.2 Background data-driven corrections for secondary particles per kind

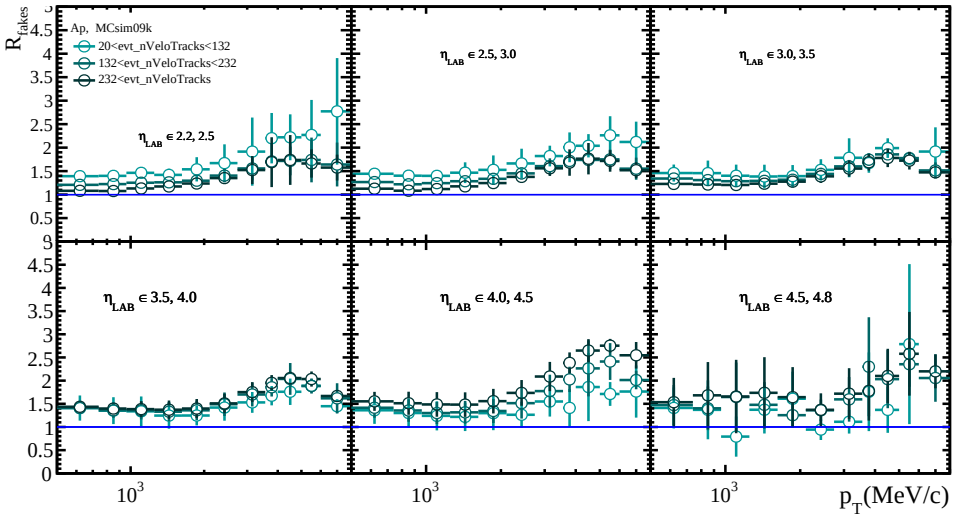


Figure B.20: Data-driven correction factor for fakes in Pbp collisions.

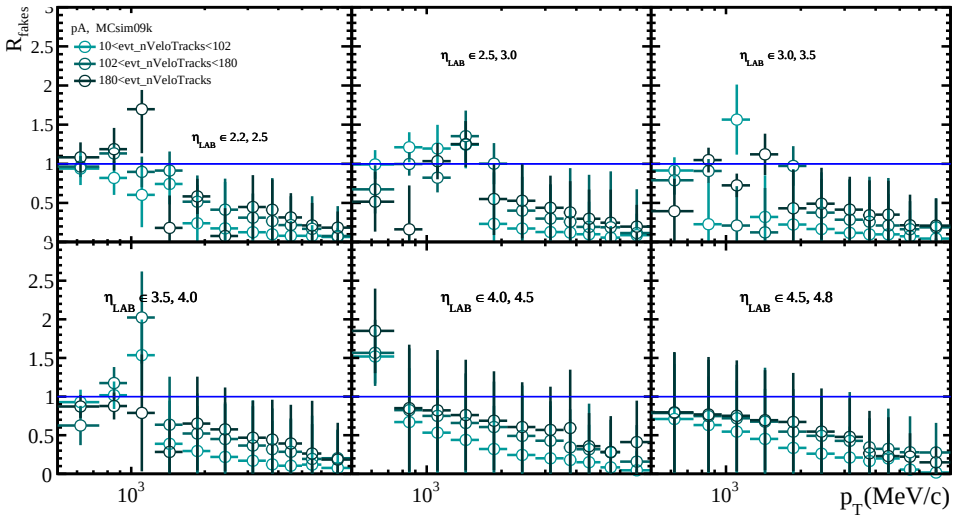


Figure B.21: Data-driven correction factor for fakes in pPb collisions.

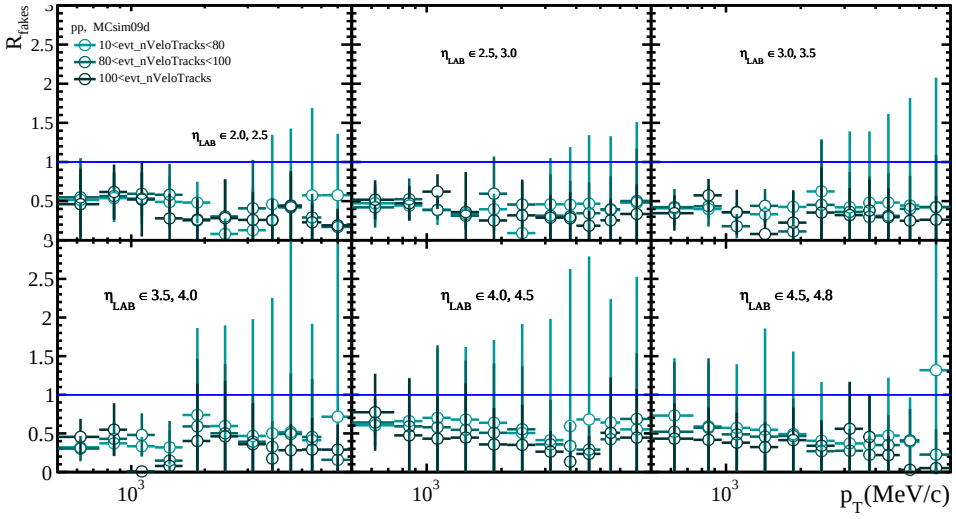


Figure B.22: Data-driven correction factor for fakes in pp collisions.

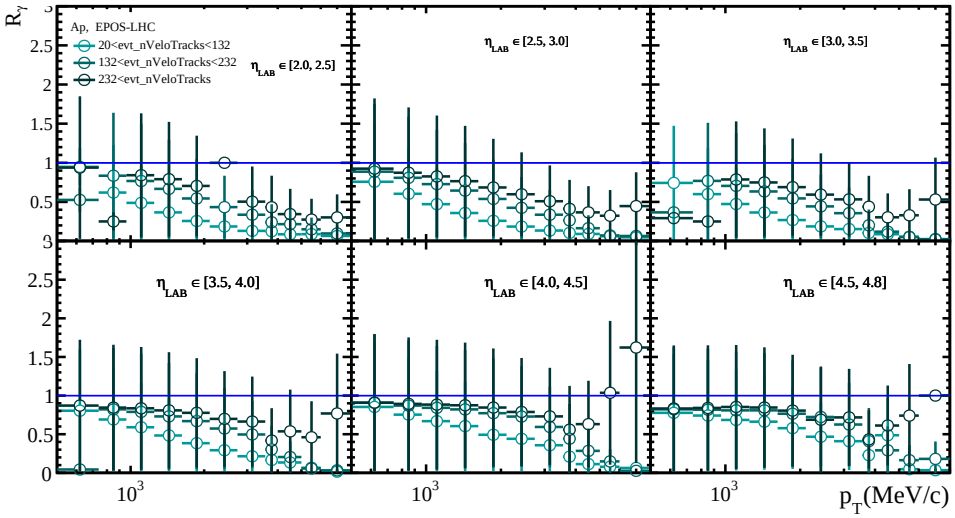


Figure B.23: R_γ for Pbp collisions.

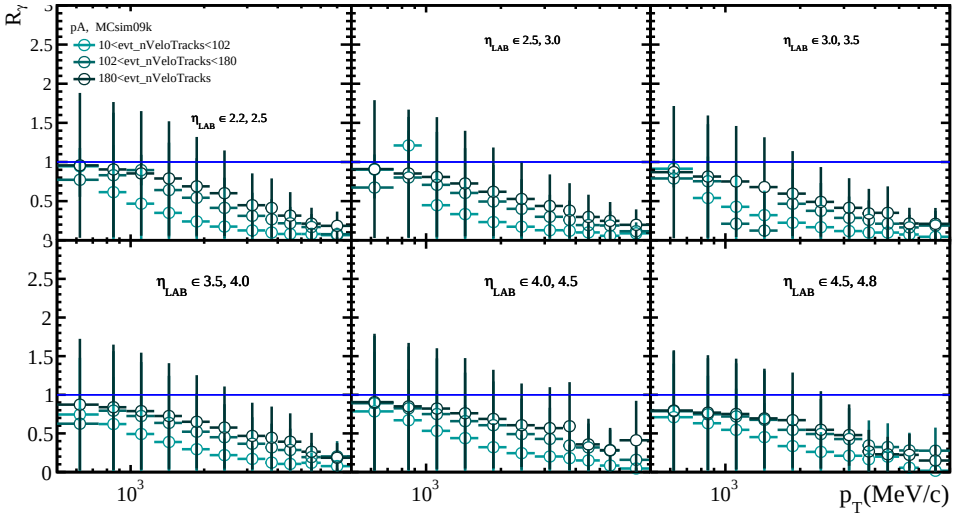


Figure B.24: R_γ for pPb collisions.

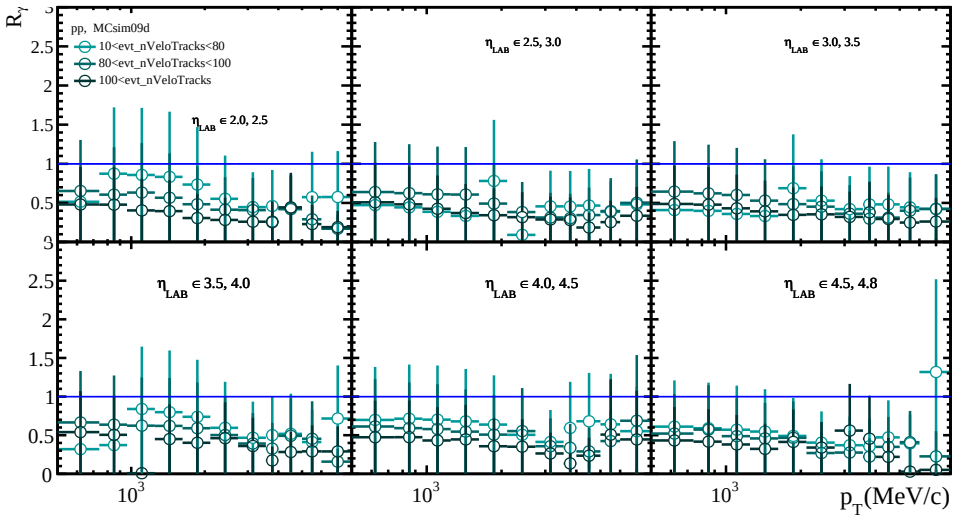


Figure B.25: R_γ for pp collisions.

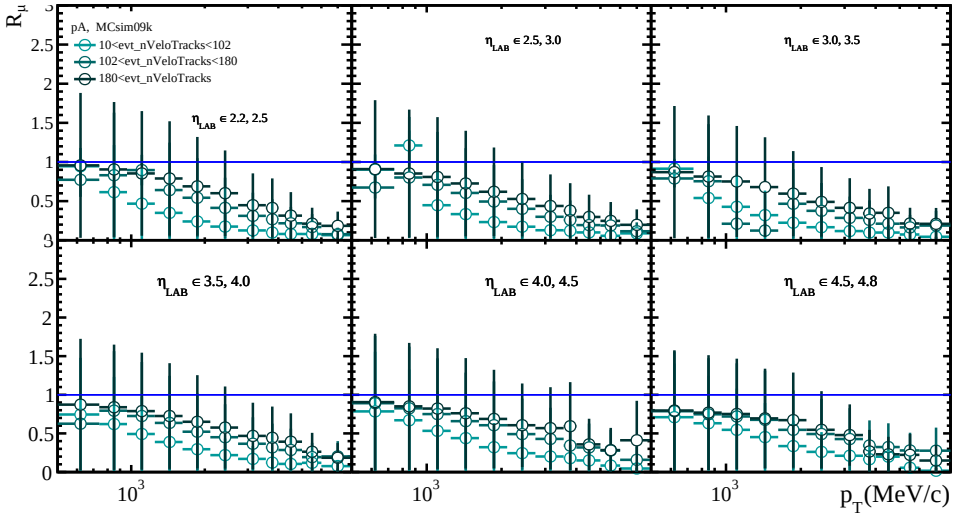


Figure B.26: R_μ for pPb collisions system.

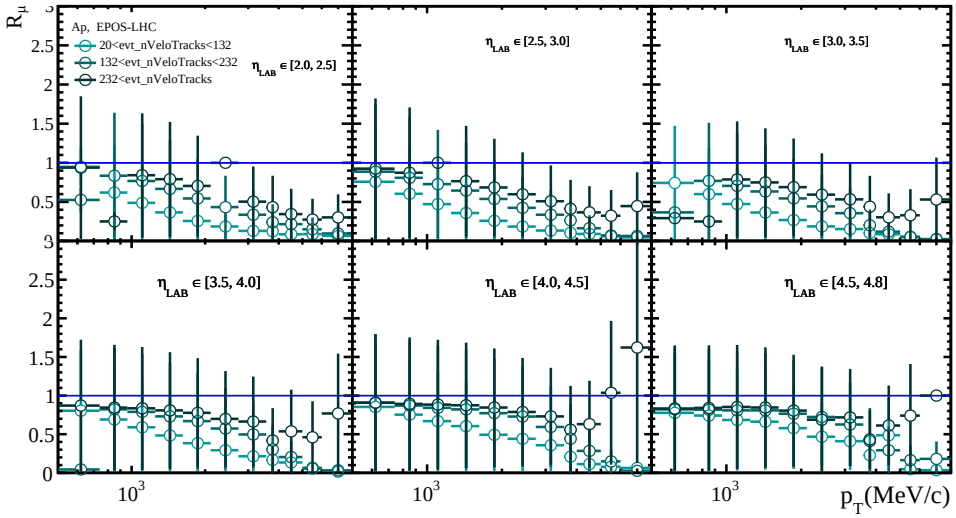


Figure B.27: R_μ for Pbp collisions system.

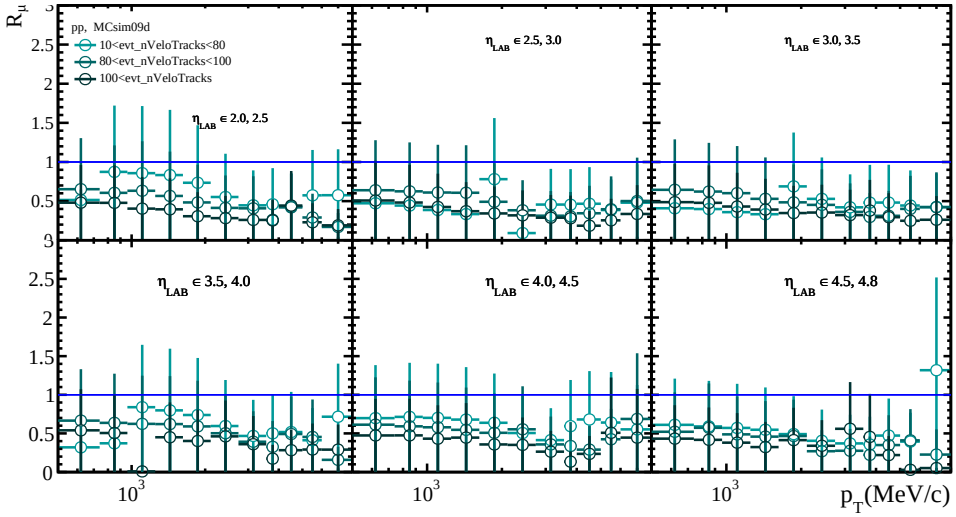


Figure B.28: R_μ for pp collisions system.

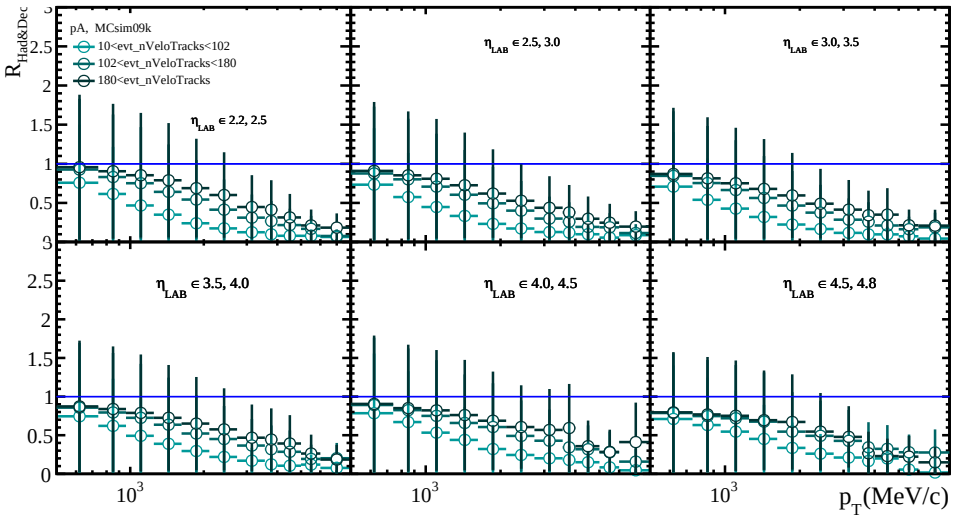


Figure B.29: $R_{Had\&Decays}$ for pPb collisions system.

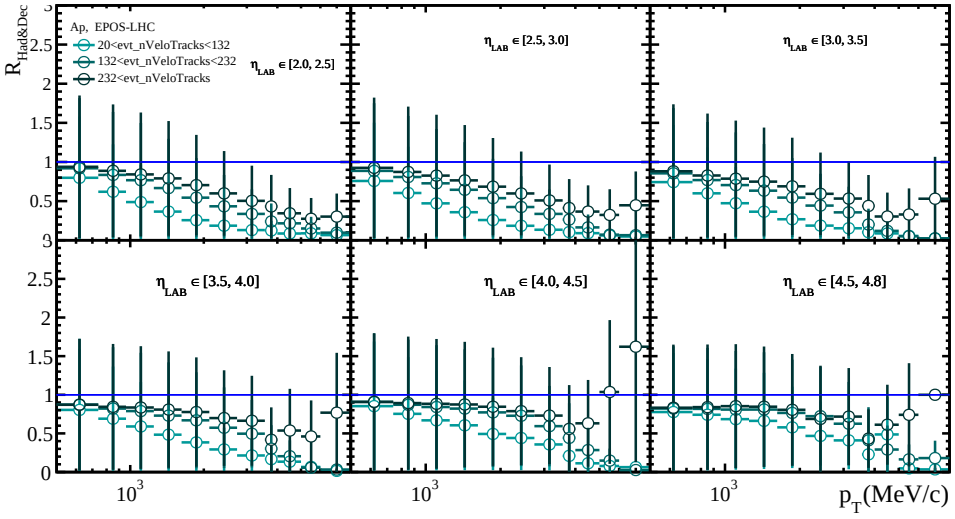


Figure B.30: $R_{Had\&Decays}$ for Pbp collisions system.

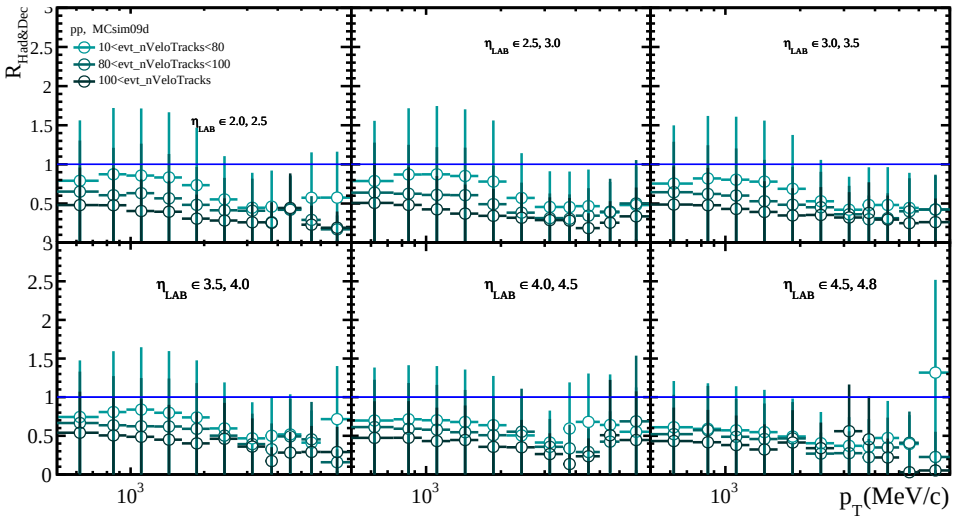


Figure B.31: $R_{Had\&Decays}$ for pp collisions system.

B.6 Effect of each correction in the multiplicity spectra measurement

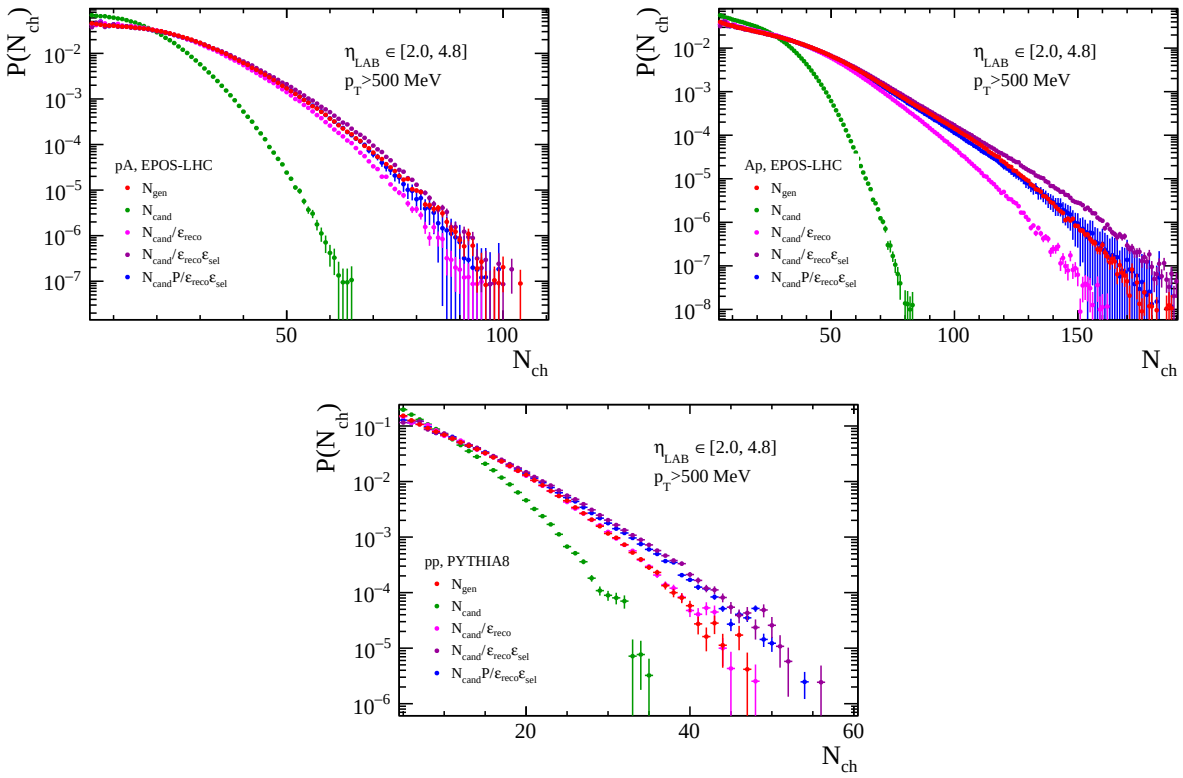


Figure B.32: Effect of the corrections applied to the multiplicity spectra of simulation data. The corrections applied in this plot are simulation-based, they don't include data-driven correction factors. The top left plot corresponds to *p*Pb, Pb*p* is the top right and *pp* is located in the bottom plot.

B.7 Systematic uncertainties

B.7.1 Reconstruction efficiency

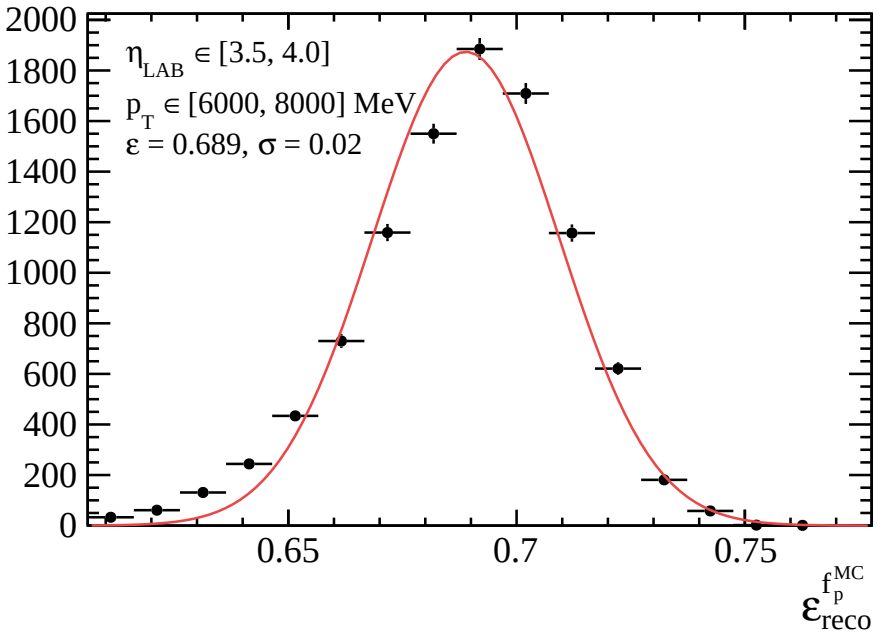


Figure B.33: Example of the fit used to estimate the particle composition systematic for the reconstruction efficiency. This example corresponds to Pbp collisions and the relative uncertainty is 3.04%.

B.7.2 Background subtraction

B.7.2.1 Fake tracks

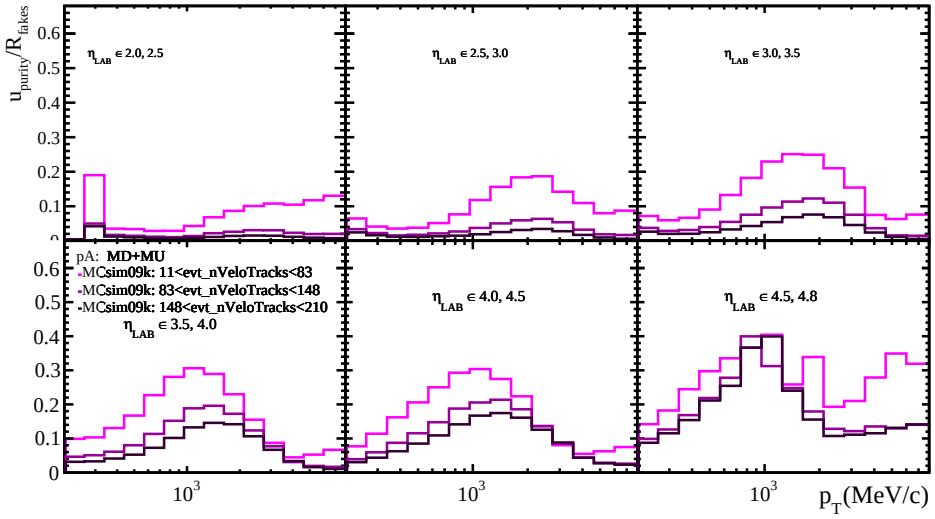


Figure B.34: Relative systematic uncertainty in the R_{fakes} due to the proxy purity in pPb collisions.

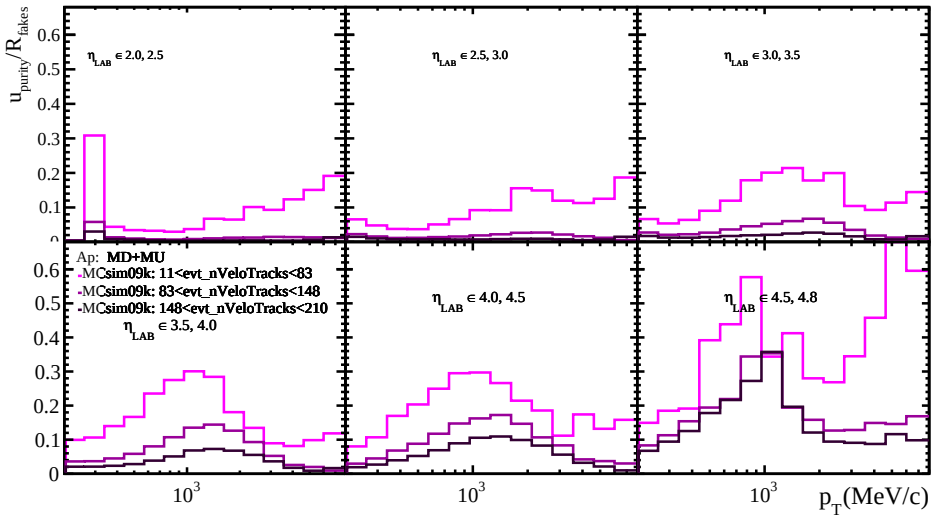


Figure B.35: Relative systematic uncertainty in the R_{fakes} due to the proxy purity in Pbp collisions.

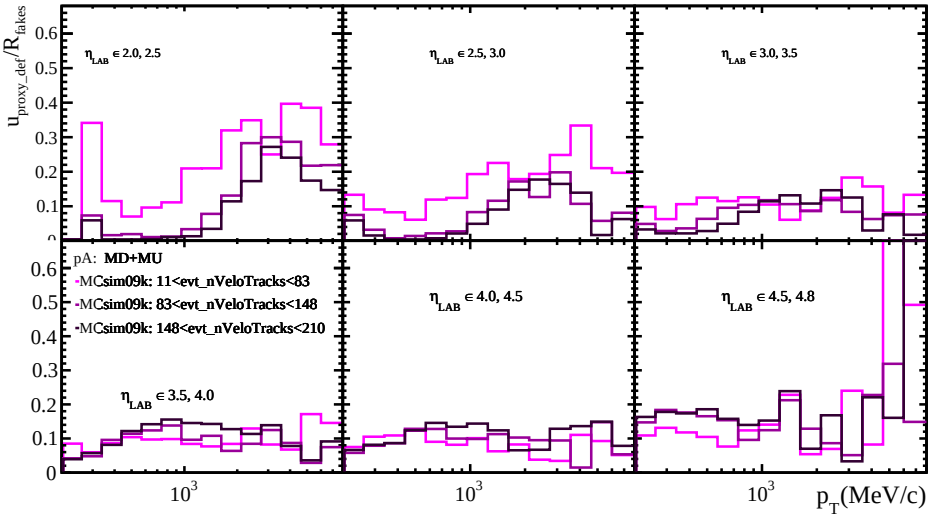


Figure B.36: Systematic uncertainty in the R_{fakes} due to the alternative proxy definition in pPb collisions.

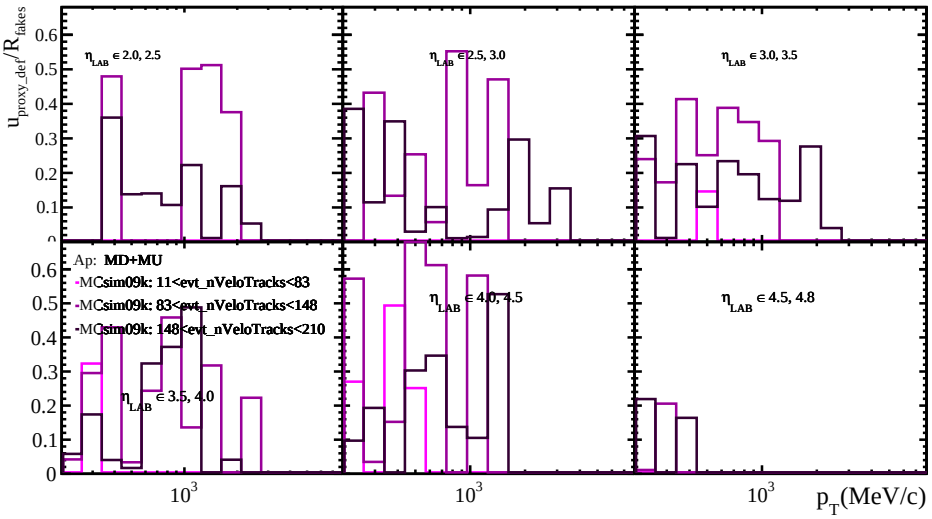


Figure B.37: Systematic uncertainty in the R_{fakes} due to the alternative in the proxy definition in Pbp collisions.

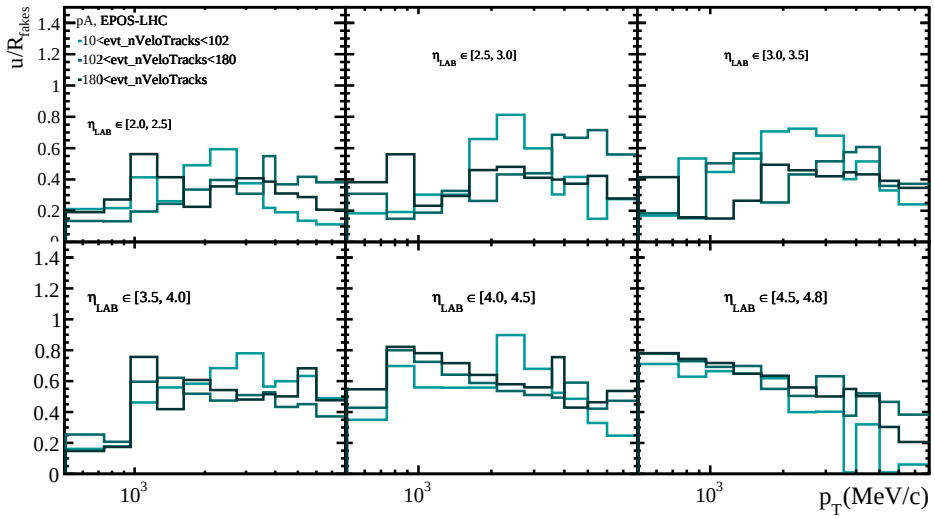


Figure B.38: Total relative uncertainty in the R_{fakes} in pPb collisions.

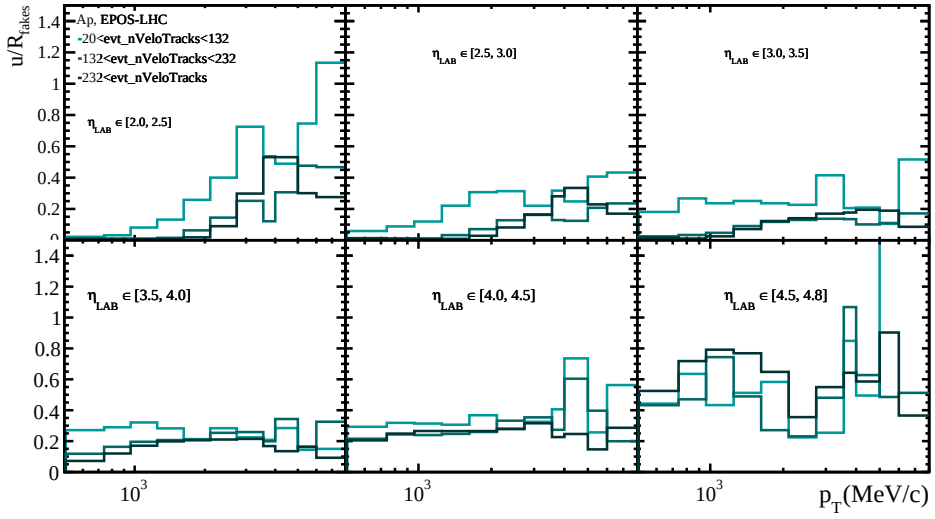


Figure B.39: Total relative uncertainty in the R_{fakes} in Pbp collisions.

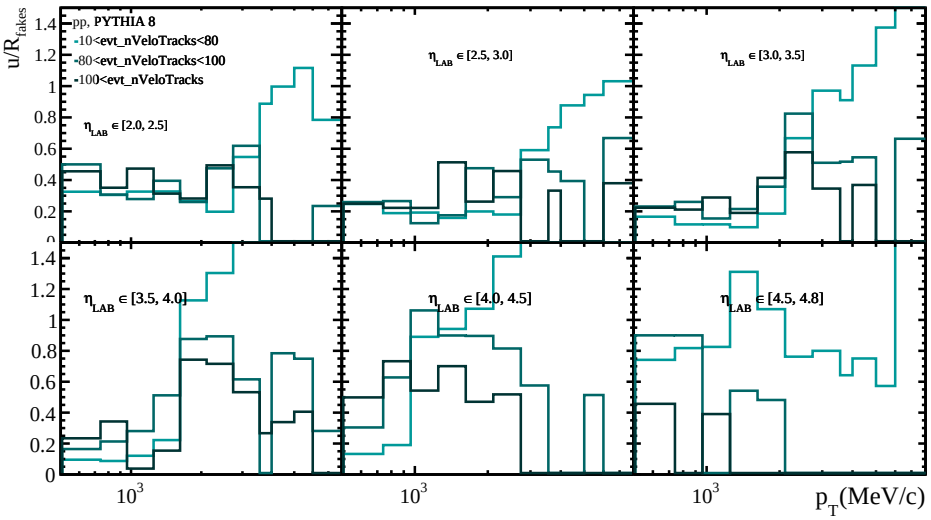


Figure B.40: Total relative uncertainty in the R_{fakes} in pp collisions.

B.7.2.2 Secondary particles

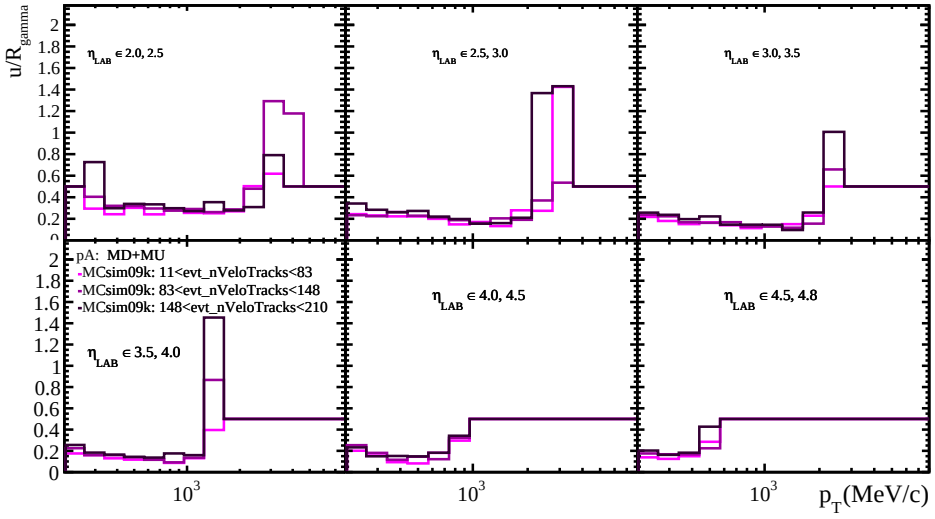


Figure B.41: Total relative uncertainty in the R_γ at pPb collisions.

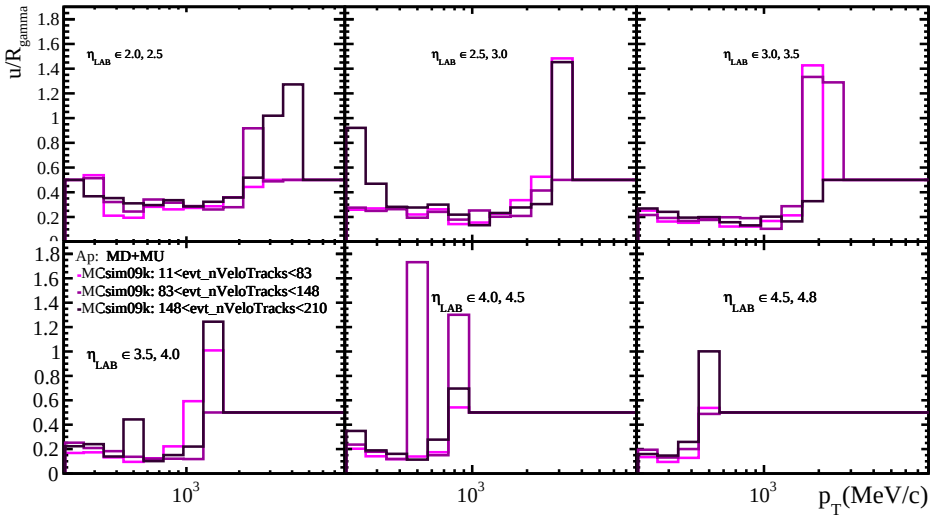


Figure B.42: Total relative uncertainty in the R_γ at Pbp collisions.

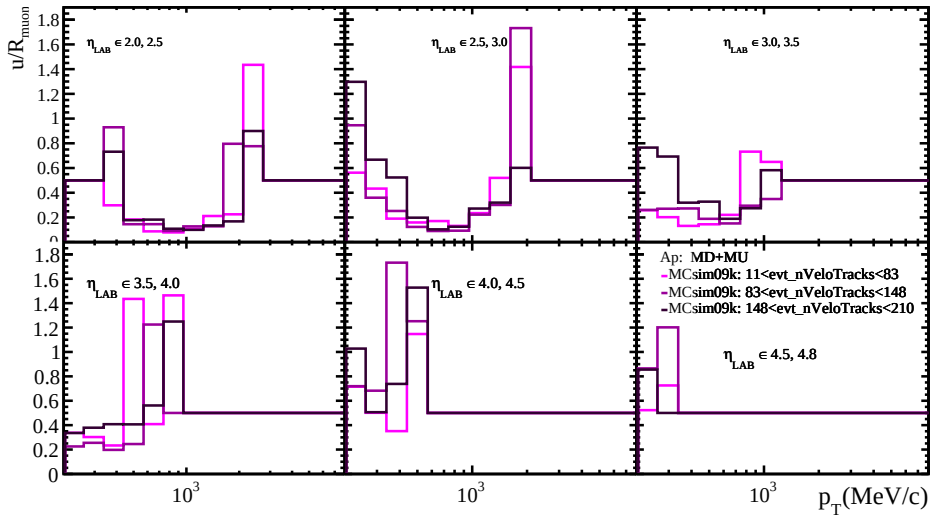


Figure B.43: Total relative uncertainty in the R_μ at pPb collisions.

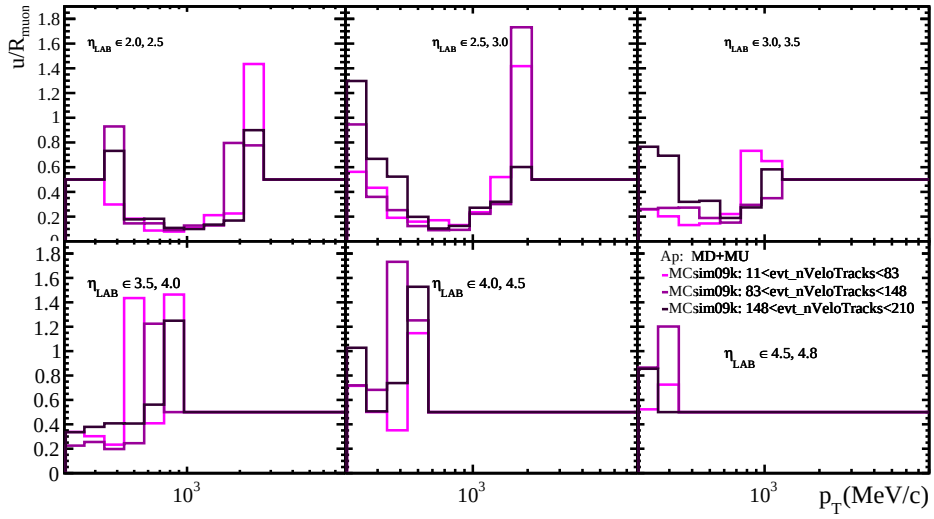


Figure B.44: Total relative uncertainty in the R_μ at Pbp collisions.

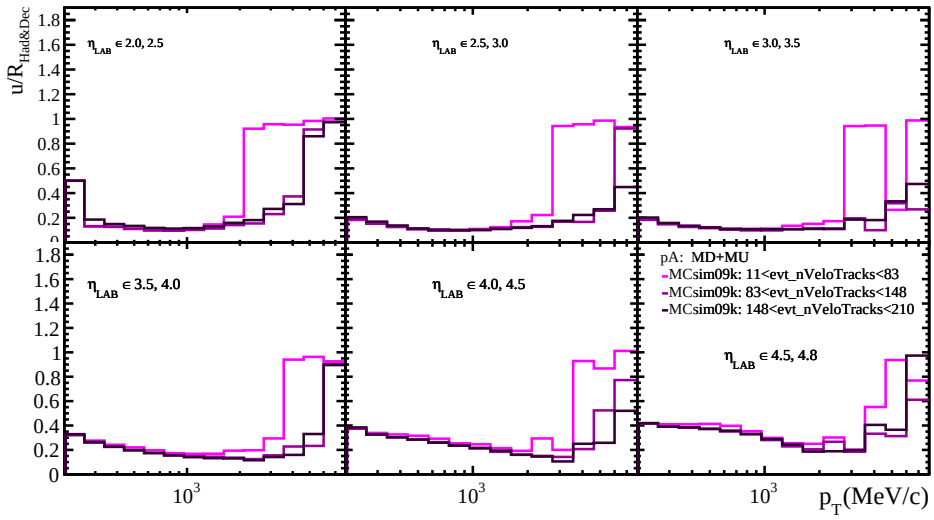


Figure B.45: Total relative uncertainty in the $R_{Had\&Decay}$ at pPb collisions.

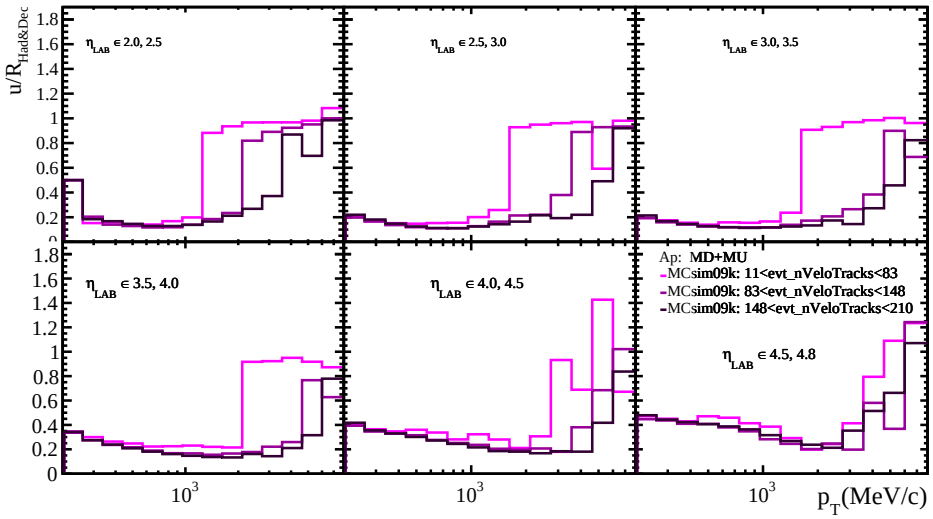


Figure B.46: Total relative uncertainty in the $R_{Had\&Decay}$ at Pbp collisions.

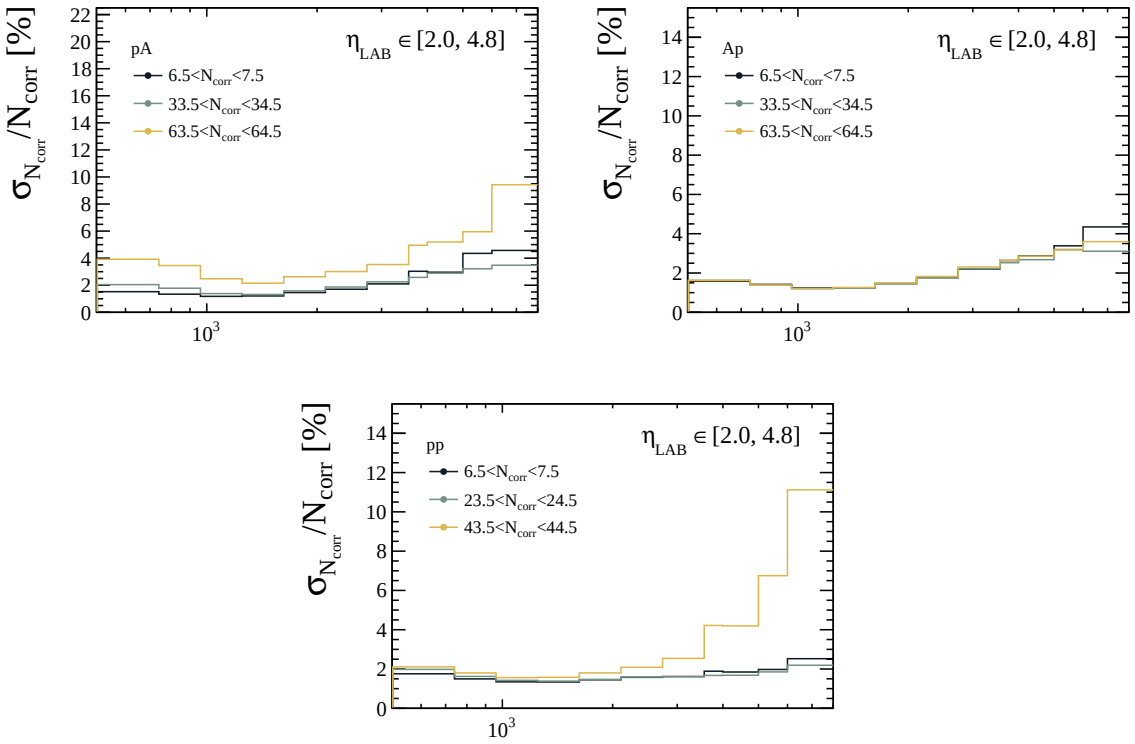


Figure B.47: Relative uncertainty of the p_T spectra in the full LHCb η range. Only some multiplicity bins are displayed.

B.7.3 Transverse momentum spectra uncertainty

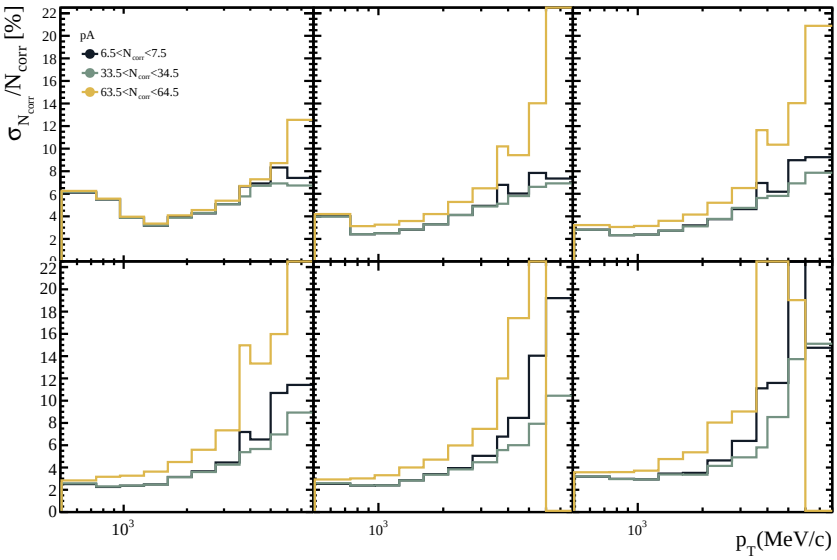


Figure B.48: Relative uncertainty of the p_{T} spectra in η bins for pPb data. Only some multiplicity bins are displayed.

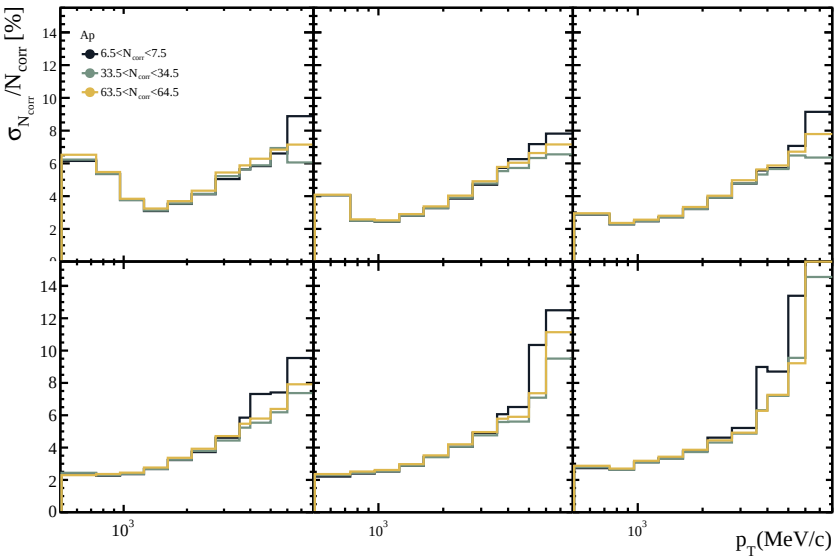


Figure B.49: Relative uncertainty of the p_{T} spectra in η bins for Pbp data. Only some multiplicity bins are displayed.

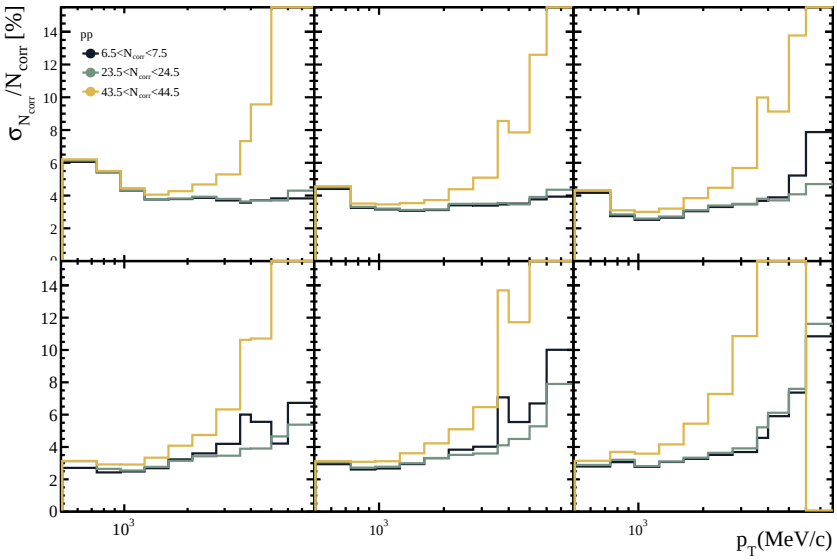


Figure B.50: Relative uncertainty of the p_{T} spectra in η bins for pp data. Only some multiplicity bins are displayed.

B.8 Results

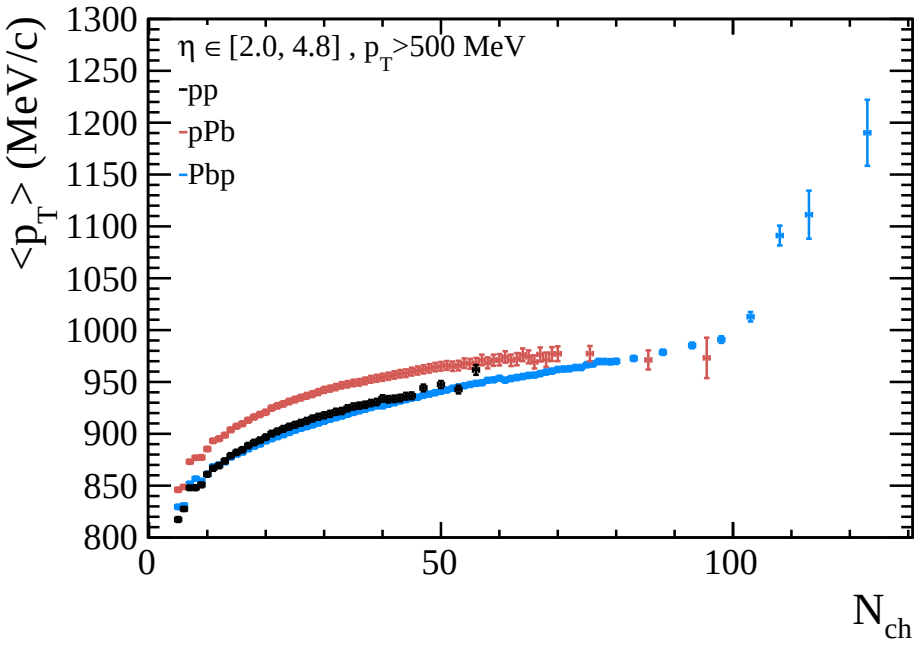


Figure B.51: Average transverse momentum as a function of multiplicity for Pbp, pPb and pp integrated in pseudorapidity using the full range.

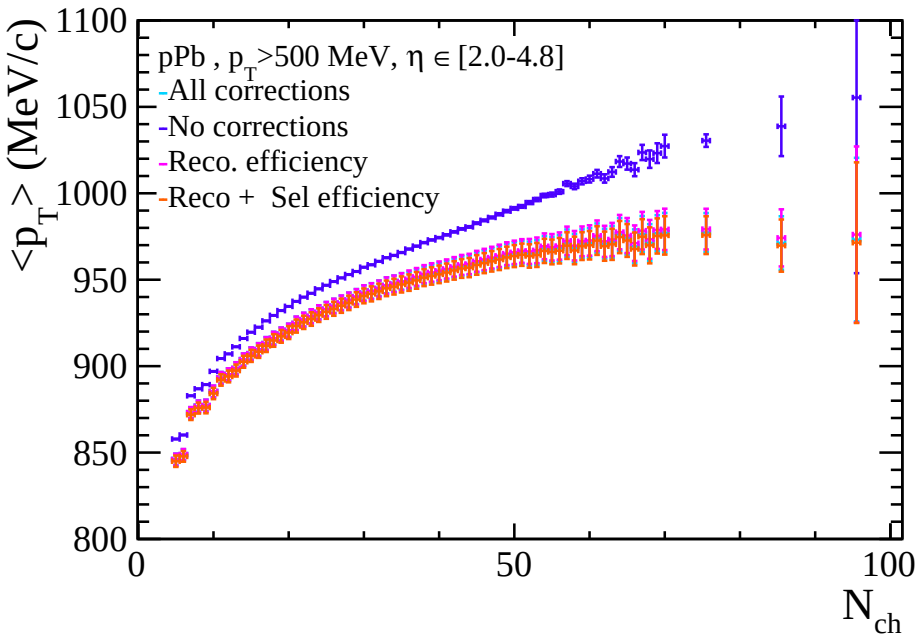


Figure B.52: Average transverse momentum as a function of multiplicity for pPb collisions in the full LHCb range. The effect of each correction step is shown. "No correction" corresponds to the result of candidates (long tracks sample with selection applied). "All corrections" includes ε_{reco}, ε_{sel} and Π.

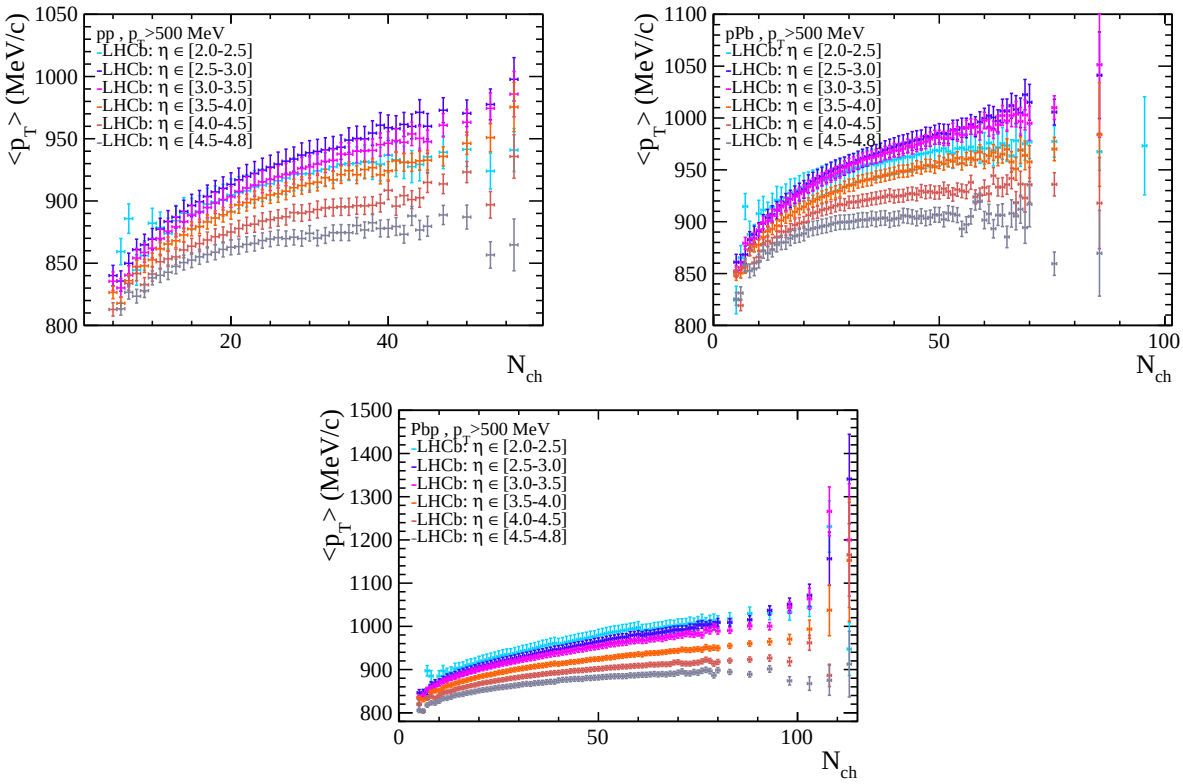


Figure B.53: Average transverse momentum as a function of multiplicity and pseudorapidity for *pp*, *p*Pb and Pb*p* collisions using all the pseudorapidity bins.

List of Figures

List of Tables

Terms of Use for CERN Audiovisual Media

Use of CERN audiovisual media denotes agreement with the following terms:

1. CERN provides the image free of charge for educational and informational use.
2. The image is provided “as-is” and the user shall hold CERN free and harmless in connection with its use.
3. CERN is to be credited as the source of the image.
4. CERN retains copyright in the image. Download and use of the image does not amount to a transfer of intellectual property.
5. The image may not be used in a misleading, inappropriate or offensive manner, in a military context, in advertising or promotion, or in a manner that suggests any kind of endorsement by CERN or its personnel.
6. The image may not be sold, distributed or otherwise made available for use by third parties.
7. The CERN logo is legally protected. CERN’s prior written approval shall be obtained for its use or for the use of any image primarily featuring the logo.
8. CERN reserves the right to alter or delete images without notice.

Please contact us by using the contact form if you have any questions or comments with respect to CERN content; if you are unsure whether your intended use meets these terms; or, if you seek permission for use of the CERN logo or any other use that does not fall within these terms. CERN appreciates your interest in its work. More information about CERN is available at www.cern.ch.

Copyright abuse should be reported using the contact form.

Review of Particle Physics

Author: Beatty, J J; Belousov, V I

Publication: Progress of Theoretical and Experimental Physics

Publisher: Oxford University Press

Date: 2020-08-14

US Government

This article is a work of the United States government. Such works are not entitled to domestic copyright protection under U.S. law and are therefore in the public domain.

This act only applies to U.S. domestic copyright as that is the extent of U.S. federal law. The U.S. government asserts that it can still hold the copyright to those works in other countries https://www.usa.gov/government-works

Editors' Suggestion | Open Access

New CTEQ global analysis of quantum chromodynamics with high-precision data from the LHC

Tie-Jiun Hou, Jun Gao, T. J. Hobbs, Keping Xie, Sayipjamal Dulat, Marco Guzzi, Joey Huston, Pavel Nadolsky, Jon Pumplin, Carl Schmidt, Ibrahim Sitiwaldi, Daniel Stump, and C.-P. Yuan

Phys. Rev. D **103**, 014013 – Published 11 January 2021

ABSTRACT

We present the new parton distribution functions (PDFs) from the CTEQ-TEA collaboration, obtained using a wide variety of high-precision Large Hadron Collider (LHC) data, in addition to the combined

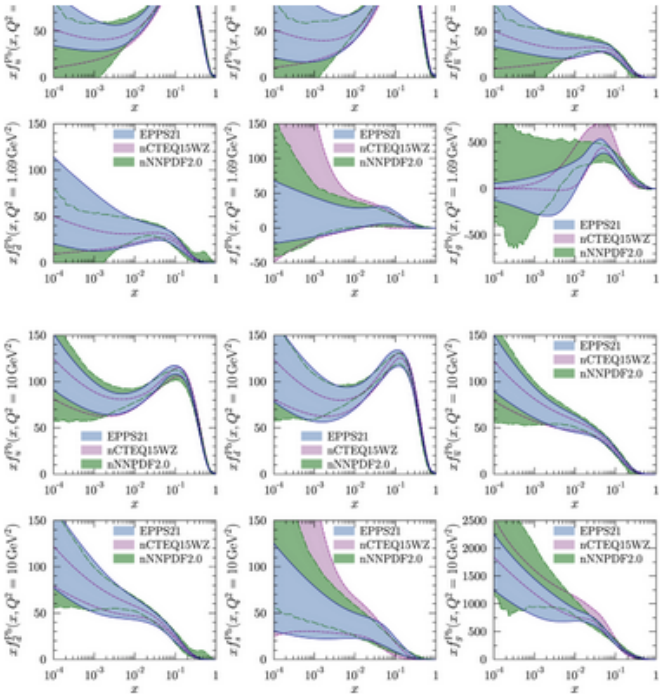
Issue

Vol. 103, Iss. 1 — 1 January 2021

Reuse & Permissions

Rights and permissions

Funded by SCOAP[3]



View in article Full size image

American Physical Society
Reuse and Permissions License

12-Aug-2024

This license agreement between the American Physical Society ("APS") and Imanol Corredoira ("You") consists of your license details and the terms and conditions provided by the American Physical Society and SciPris.

Licensed Content Information

License Number:	**RNP/24/AUG/082178**
License date:	12-Aug-2024
DOI:	10.1103/PhysRevC.92.034906
Title:	EPOS LHC: Test of collective hadronization with data measured at the CERN Large Hadron Collider
Author:	T. Pierog et al.
Publication:	Physical Review C
Publisher:	American Physical Society
Cost:	USD $ 0.00

Request Details

Does your reuse require significant modifications: No

Specify intended distribution locations:	Worldwide
Reuse Category:	Reuse in a thesis/dissertation
Requestor Type:	Student
Items for Reuse:	Figures/Tables
Number of Figure/Tables:	1
Figure/Tables Details:	Figure 2 Elementary interaction in the epos model.
Format for Reuse:	Print and Electronic
Total number of print copies:	Up to 1000

Information about New Publication:

University/Publisher:	University of Santiago de Compostela
Title of dissertation/thesis:	Average transverse momentum and multiplicity distribution measurement of prompt charged particles from proton- proton and proton-nucleus collisions at 5 TeV in the LHCb experiment
Author(s):	Imanol Corredoira Fernandez
Expected completion date:	Nov. 2024

License Requestor Information

Name:	Imanol Corredoira
Affiliation:	Individual
Email Id:	imanolcorredoira.fernandez@usc.es
Country:	Spain

American Physical Society
Reuse and Permissions License

TERMS AND CONDITIONS

The American Physical Society (APS) is pleased to grant the Requestor of this license a non-exclusive, non-transferable permission, limited to Print and Electronic format, provided all criteria outlined below are followed.

1. You must also obtain permission from at least one of the lead authors for each separate work, if you haven't done so already. The author's name and affiliation can be found on the first page of the published Article.
2. For electronic format permissions, Requestor agrees to provide a hyperlink from the reprinted APS material using the source material's DOI on the web page where the work appears. The hyperlink should use the standard DOI resolution URL, http://dx.doi.org/{DOI}. The hyperlink may be embedded in the copyright credit line.
3. For print format permissions, Requestor agrees to print the required copyright credit line on the first page where the material appears: "Reprinted (abstract/excerpt/figure) with permission from [(FULL REFERENCE CITATION) as follows: Author's Names, APS Journal Title, Volume Number, Page Number and Year of Publication.] Copyright (YEAR) by the American Physical Society."
4. Permission granted in this license is for a one-time use and does not include permission for any future editions, updates, databases, formats or other matters. Permission must be sought for any additional use.
5. Use of the material does not and must not imply any endorsement by APS.
6. APS does not imply, purport or intend to grant permission to reuse materials to which it does not hold copyright. It is the requestor's sole responsibility to ensure the licensed material is original to APS and does not contain the copyright of another entity, and that the copyright notice of the figure, photograph, cover or table does not indicate it was reprinted by APS with permission from another source.
7. The permission granted herein is personal to the Requestor for the use specified and is not transferable or assignable without express written permission of APS. This license may not be amended except in writing by APS.
8. You may not alter, edit or modify the material in any manner.
9. You may translate the materials only when translation rights have been granted.
10. APS is not responsible for any errors or omissions due to translation.
11. You may not use the material for promotional, sales, advertising or marketing purposes.
12. The foregoing license shall not take effect unless and until APS or its agent, Aptara, receives payment in full in accordance with Aptara Billing and Payment Terms and Conditions, which are incorporated herein by reference.
13. Should the terms of this license be violated at any time, APS or Aptara may revoke the license with no refund to you and seek relief to the fullest extent of the laws of the USA. Official written notice will be made using the contact information provided with the permission request. Failure to receive such notice will not nullify revocation of the permission.
14. APS reserves all rights not specifically granted herein.
15. This document, including the Aptara Billing and Payment Terms and Conditions, shall be the entire agreement between the parties relating to the subject matter hereof.

American Physical Society
Reuse and Permissions License

12-Aug-2024

This license agreement between the American Physical Society ("APS") and Imanol Corredoira ("You") consists of your license details and the terms and conditions provided by the American Physical Society and SciPris.

Licensed Content Information

License Number:	**RNP/24/AUG/082199**
License date:	12-Aug-2024
DOI:	10.1103/PhysRevLett.116.212301
Title:	Moving Forward to Constrain the Shear Viscosity of QCD Matter
Author:	Gabriel Denicol, Akihiko Monnai, and Björn Schenke
Publication:	Physical Review Letters
Publisher:	American Physical Society
Cost:	USD $ 0.00

Request Details

Does your reuse require significant modifications: No

Specify intended distribution locations:	Worldwide
Reuse Category:	Reuse in a thesis/dissertation
Requestor Type:	Student
Items for Reuse:	Figures/Tables
Number of Figure/Tables:	2
Figure/Tables Details:	Figure 1 The four scenarios of temperature dependent $\eta/T/(\epsilon+P)$ at μB=0. Figure 3 $v2$ of charged hadrons as a function of pseudorapidity for the four different shear viscosity scenarios compar
Format for Reuse:	Print and Electronic
Total number of print copies:	Up to 1000

Information about New Publication:

University/Publisher:	University of Santiago de Compostela
Title of dissertation/thesis:	Average transverse momentum and multiplicity distribution measurement of prompt charged particles from proton- proton and proton-nucleus collisions at 5 TeV in the LHCb experiment
Author(s):	Imanol Corredoira Fernandez
Expected completion date:	Nov. 2024

License Requestor Information

Name:	Imanol Corredoira
Affiliation:	Individual
Email Id:	imanolcorredoira.fernandez@usc.es
Country:	Spain

American Physical Society
Reuse and Permissions License

TERMS AND CONDITIONS

The American Physical Society (APS) is pleased to grant the Requestor of this license a non-exclusive, non-transferable permission, limited to Print and Electronic format, provided all criteria outlined below are followed.

1. You must also obtain permission from at least one of the lead authors for each separate work, if you haven't done so already. The author's name and affiliation can be found on the first page of the published Article.
2. For electronic format permissions, Requestor agrees to provide a hyperlink from the reprinted APS material using the source material's DOI on the web page where the work appears. The hyperlink should use the standard DOI resolution URL, http://dx.doi.org/{DOI}. The hyperlink may be embedded in the copyright credit line.
3. For print format permissions, Requestor agrees to print the required copyright credit line on the first page where the material appears: "Reprinted (abstract/excerpt/figure) with permission from [(FULL REFERENCE CITATION) as follows: Author's Names, APS Journal Title, Volume Number, Page Number and Year of Publication.] Copyright (YEAR) by the American Physical Society."
4. Permission granted in this license is for a one-time use and does not include permission for any future editions, updates, databases, formats or other matters. Permission must be sought for any additional use.
5. Use of the material does not and must not imply any endorsement by APS.
6. APS does not imply, purport or intend to grant permission to reuse materials to which it does not hold copyright. It is the requestor's sole responsibility to ensure the licensed material is original to APS and does not contain the copyright of another entity, and that the copyright notice of the figure, photograph, cover or table does not indicate it was reprinted by APS with permission from another source.
7. The permission granted herein is personal to the Requestor for the use specified and is not transferable or assignable without express written permission of APS. This license may not be amended except in writing by APS.
8. You may not alter, edit or modify the material in any manner.
9. You may translate the materials only when translation rights have been granted.
10. APS is not responsible for any errors or omissions due to translation.
11. You may not use the material for promotional, sales, advertising or marketing purposes.
12. The foregoing license shall not take effect unless and until APS or its agent, Aptara, receives payment in full in accordance with Aptara Billing and Payment Terms and Conditions, which are incorporated herein by reference.
13. Should the terms of this license be violated at any time, APS or Aptara may revoke the license with no refund to you and seek relief to the fullest extent of the laws of the USA. Official written notice will be made using the contact information provided with the permission request. Failure to receive such notice will not nullify revocation of the permission.
14. APS reserves all rights not specifically granted herein.
15. This document, including the Aptara Billing and Payment Terms and Conditions, shall be the entire agreement between the parties relating to the subject matter hereof.

ELSEVIER LICENSE
TERMS AND CONDITIONS

Sep 05, 2024

This Agreement between Imanol Corredoira Fernandez ("You") and Elsevier ("Elsevier") consists of your license details and the terms and conditions provided by Elsevier and Copyright Clearance Center.

License Number	5862460063107
License date	Sep 05, 2024
Licensed Content Publisher	Elsevier
Licensed Content Publication	Physics Reports
Licensed Content Title	Properties of hot and dense matter from relativistic heavy ion collisions
Licensed Content Author	Peter Braun-Munzinger,Volker Koch,Thomas Schäfer,Johanna Stachel
Licensed Content Date	Mar 21, 2016
Licensed Content Volume	621
Licensed Content Issue	n/a
Licensed Content Pages	51
Start Page	76

End Page	126
Type of Use	reuse in a thesis/dissertation
Portion	figures/tables/illustrations
Number of figures/tables/ illustrations	1
Format	both print and electronic
Are you the author of this Elsevier article?	No
Will you be translating?	No
Title of new work	Imanol Corredoira PhD thesis
Institution name	University of Santiago de Compostela
Expected presentation date	Nov 2024
Order reference number	1
Portions	Fig. 1. Schematic phase diagram of QCD as a function of temperature and baryon chemical potential . QGP refers to the quark–gluon plasma
The Requesting Person / Organization to Appear on the License	Imanol Corredoira Fernandez
Requestor Location	University of Santiago de Compostela fernando 3 o santo, 33 santiago de compostela, 15701

Spain
Attn: University of Santiago de Compostela

Publisher Tax ID	GB 494 6272 12
Total	0.00 EUR

Terms and Conditions

INTRODUCTION

1. The publisher for this copyrighted material is Elsevier. By clicking "accept" in connection with completing this licensing transaction, you agree that the following terms and conditions apply to this transaction (along with the Billing and Payment terms and conditions established by Copyright Clearance Center, Inc. ("CCC"), at the time that you opened your RightsLink account and that are available at any time at https://myaccount.copyright.com).

GENERAL TERMS

2. Elsevier hereby grants you permission to reproduce the aforementioned material subject to the terms and conditions indicated.

3. Acknowledgement: If any part of the material to be used (for example, figures) has appeared in our publication with credit or acknowledgement to another source, permission must also be sought from that source. If such permission is not obtained then that material may not be included in your publication/copies. Suitable acknowledgement to the source must be made, either as a footnote or in a reference list at the end of your publication, as follows:

"Reprinted from Publication title, Vol /edition number, Author(s), Title of article / title of chapter, Pages No., Copyright (Year), with permission from Elsevier [OR APPLICABLE SOCIETY COPYRIGHT OWNER]." Also Lancet special credit - "Reprinted from The Lancet, Vol. number, Author(s), Title of article, Pages No., Copyright (Year), with permission from Elsevier."

4. Reproduction of this material is confined to the purpose and/or media for which permission is hereby given. The material may not be reproduced or used in any other way, including use in combination with an artificial intelligence tool (including to train an algorithm, test, process, analyse, generate output and/or develop any form of artificial intelligence tool), or to create any derivative work and/or service (including resulting from the use of artificial intelligence tools).

5. Altering/Modifying Material: Not Permitted. However figures and illustrations may be altered/adapted minimally to serve your work. Any other abbreviations, additions, deletions and/or any other alterations shall be made only with prior written authorization of Elsevier Ltd. (Please contact Elsevier's permissions helpdesk here). No modifications

can be made to any Lancet figures/tables and they must be reproduced in full.

6. If the permission fee for the requested use of our material is waived in this instance, please be advised that your future requests for Elsevier materials may attract a fee.

7. Reservation of Rights: Publisher reserves all rights not specifically granted in the combination of (i) the license details provided by you and accepted in the course of this licensing transaction, (ii) these terms and conditions and (iii) CCC's Billing and Payment terms and conditions.

8. License Contingent Upon Payment: While you may exercise the rights licensed immediately upon issuance of the license at the end of the licensing process for the transaction, provided that you have disclosed complete and accurate details of your proposed use, no license is finally effective unless and until full payment is received from you (either by publisher or by CCC) as provided in CCC's Billing and Payment terms and conditions. If full payment is not received on a timely basis, then any license preliminarily granted shall be deemed automatically revoked and shall be void as if never granted. Further, in the event that you breach any of these terms and conditions or any of CCC's Billing and Payment terms and conditions, the license is automatically revoked and shall be void as if never granted. Use of materials as described in a revoked license, as well as any use of the materials beyond the scope of an unrevoked license, may constitute copyright infringement and publisher reserves the right to take any and all action to protect its copyright in the materials.

9. Warranties: Publisher makes no representations or warranties with respect to the licensed material.

10. Indemnity: You hereby indemnify and agree to hold harmless publisher and CCC, and their respective officers, directors, employees and agents, from and against any and all claims arising out of your use of the licensed material other than as specifically authorized pursuant to this license.

11. No Transfer of License: This license is personal to you and may not be sublicensed, assigned, or transferred by you to any other person without publisher's written permission.

12. No Amendment Except in Writing: This license may not be amended except in a writing signed by both parties (or, in the case of publisher, by CCC on publisher's behalf).

13. Objection to Contrary Terms: Publisher hereby objects to any terms contained in any purchase order, acknowledgment, check endorsement or other writing prepared by you, which terms are inconsistent with these terms and conditions or CCC's Billing and Payment terms and conditions. These terms and conditions, together with CCC's Billing and Payment terms and conditions (which are incorporated herein), comprise the entire agreement between you and publisher (and CCC) concerning this licensing transaction. In the event of any conflict between your obligations established by these terms and conditions and those established by CCC's Billing and Payment terms and conditions, these terms and conditions shall control.

14. Revocation: Elsevier or Copyright Clearance Center may deny the permissions described in this License at their sole discretion, for any reason or no reason, with a full refund payable to you. Notice of such denial will be made using the contact information

provided by you. Failure to receive such notice will not alter or invalidate the denial. In no event will Elsevier or Copyright Clearance Center be responsible or liable for any costs, expenses or damage incurred by you as a result of a denial of your permission request, other than a refund of the amount(s) paid by you to Elsevier and/or Copyright Clearance Center for denied permissions.

LIMITED LICENSE

The following terms and conditions apply only to specific license types:

15. **Translation**: This permission is granted for non-exclusive world **English** rights only unless your license was granted for translation rights. If you licensed translation rights you may only translate this content into the languages you requested. A professional translator must perform all translations and reproduce the content word for word preserving the integrity of the article.

16. **Posting licensed content on any Website**: The following terms and conditions apply as follows: Licensing material from an Elsevier journal: All content posted to the web site must maintain the copyright information line on the bottom of each image; A hyper-text must be included to the Homepage of the journal from which you are licensing at http://www.sciencedirect.com/science/journal/xxxxx or the Elsevier homepage for books at http://www.elsevier.com; Central Storage: This license does not include permission for a scanned version of the material to be stored in a central repository such as that provided by Heron/XanEdu.

Licensing material from an Elsevier book: A hyper-text link must be included to the Elsevier homepage at http://www.elsevier.com . All content posted to the web site must maintain the copyright information line on the bottom of each image.

Posting licensed content on Electronic reserve: In addition to the above the following clauses are applicable: The web site must be password-protected and made available only to bona fide students registered on a relevant course. This permission is granted for 1 year only. You may obtain a new license for future website posting.

17. **For journal authors:** the following clauses are applicable in addition to the above:

Preprints:

A preprint is an author's own write-up of research results and analysis, it has not been peer-reviewed, nor has it had any other value added to it by a publisher (such as formatting, copyright, technical enhancement etc.).

Authors can share their preprints anywhere at any time. Preprints should not be added to or enhanced in any way in order to appear more like, or to substitute for, the final versions of articles however authors can update their preprints on arXiv or RePEc with their Accepted Author Manuscript (see below).

If accepted for publication, we encourage authors to link from the preprint to their formal publication via its DOI. Millions of researchers have access to the formal publications on ScienceDirect, and so links will help users to find, access, cite and use the best available

version. Please note that Cell Press, The Lancet and some society-owned have different preprint policies. Information on these policies is available on the journal homepage.

Accepted Author Manuscripts: An accepted author manuscript is the manuscript of an article that has been accepted for publication and which typically includes author-incorporated changes suggested during submission, peer review and editor-author communications.

Authors can share their accepted author manuscript:

- immediately
  - via their non-commercial person homepage or blog
  - by updating a preprint in arXiv or RePEc with the accepted manuscript
  - via their research institute or institutional repository for internal institutional uses or as part of an invitation-only research collaboration work-group
  - directly by providing copies to their students or to research collaborators for their personal use
  - for private scholarly sharing as part of an invitation-only work group on commercial sites with which Elsevier has an agreement
- After the embargo period
  - via non-commercial hosting platforms such as their institutional repository
  - via commercial sites with which Elsevier has an agreement

In all cases accepted manuscripts should:

- link to the formal publication via its DOI
- bear a CC-BY-NC-ND license - this is easy to do
- if aggregated with other manuscripts, for example in a repository or other site, be shared in alignment with our hosting policy not be added to or enhanced in any way to appear more like, or to substitute for, the published journal article.

Published journal article (JPA): A published journal article (PJA) is the definitive final record of published research that appears or will appear in the journal and embodies all value-adding publishing activities including peer review co-ordination, copy-editing, formatting, (if relevant) pagination and online enrichment.

Policies for sharing publishing journal articles differ for subscription and gold open access articles:

Subscription Articles: If you are an author, please share a link to your article rather than the full-text. Millions of researchers have access to the formal publications on ScienceDirect, and so links will help your users to find, access, cite, and use the best available version.

Theses and dissertations which contain embedded PJAs as part of the formal submission can be posted publicly by the awarding institution with DOI links back to the formal publications on ScienceDirect.

If you are affiliated with a library that subscribes to ScienceDirect you have additional private sharing rights for others' research accessed under that agreement. This includes

use for classroom teaching and internal training at the institution (including use in course packs and courseware programs), and inclusion of the article for grant funding purposes.

Gold Open Access Articles: May be shared according to the author-selected end-user license and should contain a CrossMark logo, the end user license, and a DOI link to the formal publication on ScienceDirect.

Please refer to Elsevier's posting policy for further information.

18. **For book authors** the following clauses are applicable in addition to the above: Authors are permitted to place a brief summary of their work online only. You are not allowed to download and post the published electronic version of your chapter, nor may you scan the printed edition to create an electronic version. **Posting to a repository:** Authors are permitted to post a summary of their chapter only in their institution's repository.

19. **Thesis/Dissertation**: If your license is for use in a thesis/dissertation your thesis may be submitted to your institution in either print or electronic form. Should your thesis be published commercially, please reapply for permission. These requirements include permission for the Library and Archives of Canada to supply single copies, on demand, of the complete thesis and include permission for Proquest/UMI to supply single copies, on demand, of the complete thesis. Should your thesis be published commercially, please reapply for permission. Theses and dissertations which contain embedded PJAs as part of the formal submission can be posted publicly by the awarding institution with DOI links back to the formal publications on ScienceDirect.

Elsevier Open Access Terms and Conditions

You can publish open access with Elsevier in hundreds of open access journals or in nearly 2000 established subscription journals that support open access publishing. Permitted third party re-use of these open access articles is defined by the author's choice of Creative Commons user license. See our open access license policy for more information.

Terms & Conditions applicable to all Open Access articles published with Elsevier:

Any reuse of the article must not represent the author as endorsing the adaptation of the article nor should the article be modified in such a way as to damage the author's honour or reputation. If any changes have been made, such changes must be clearly indicated.

The author(s) must be appropriately credited and we ask that you include the end user license and a DOI link to the formal publication on ScienceDirect.

If any part of the material to be used (for example, figures) has appeared in our publication with credit or acknowledgement to another source it is the responsibility of the user to ensure their reuse complies with the terms and conditions determined by the rights holder.

Additional Terms & Conditions applicable to each Creative Commons user license:

CC BY: The CC-BY license allows users to copy, to create extracts, abstracts and new works from the Article, to alter and revise the Article and to make commercial use of the Article (including reuse and/or resale of the Article by commercial entities), provided the user gives appropriate credit (with a link to the formal publication through the relevant DOI), provides a link to the license, indicates if changes were made and the licensor is not represented as endorsing the use made of the work. The full details of the license are available at http://creativecommons.org/licenses/by/4.0.

CC BY NC SA: The CC BY-NC-SA license allows users to copy, to create extracts, abstracts and new works from the Article, to alter and revise the Article, provided this is not done for commercial purposes, and that the user gives appropriate credit (with a link to the formal publication through the relevant DOI), provides a link to the license, indicates if changes were made and the licensor is not represented as endorsing the use made of the work. Further, any new works must be made available on the same conditions. The full details of the license are available at http://creativecommons.org/licenses/by-nc-sa/4.0.

CC BY NC ND: The CC BY-NC-ND license allows users to copy and distribute the Article, provided this is not done for commercial purposes and further does not permit distribution of the Article if it is changed or edited in any way, and provided the user gives appropriate credit (with a link to the formal publication through the relevant DOI), provides a link to the license, and that the licensor is not represented as endorsing the use made of the work. The full details of the license are available at http://creativecommons.org/licenses/by-nc-nd/4.0. Any commercial reuse of Open Access articles published with a CC BY NC SA or CC BY NC ND license requires permission from Elsevier and will be subject to a fee.

Commercial reuse includes:

- Associating advertising with the full text of the Article
- Charging fees for document delivery or access
- Article aggregation
- Systematic distribution via e-mail lists or share buttons

Posting or linking by commercial companies for use by customers of those companies.

20. **Other Conditions**:

v1.10

Questions? customercare@copyright.com.

REFERENCES

[1] H. D. Politzer, *Reliable perturbative results for strong interactions?*, Phys. Rev. Lett. **30** (1973) 1346. (Cited on pages 1 and 8.)

[2] D. J. Gross and F. Wilczek, *Asymptotically free gauge theories. i*, Phys. Rev. D **8** (1973) 3633. (Cited on page 1.)

[3] M. Y. Han and Y. Nambu, *Three-triplet model with double* SU(3) *symmetry*, Phys. Rev. **139** (1965) B1006. (Cited on page 1.)

[4] J. F. Grosse-Oetringhaus and K. Reygers, *Charged-particle multiplicity in proton–proton collisions*, Journal of Physics G: Nuclear and Particle Physics **37** (2010) 083001. (Cited on pages 1, 2, and 138.)

[5] S. A. et al. *Multiplicity dependence of charged-particle production in pp, p–pb, xe–xe and pb–pb collisions at the lhc*, Physics Letters B **845** (2023) 138110. (Cited on pages 2, 124, 127, 139, 140, 141, 157, 161, and 211.)

[6] D. D. et al, *Measurement of the charged particle multiplicity distribution in hadronic z decays*, Physics Letters B **273** (1991) 181. (Cited on page 2.)

[7] Wikipedia, *Standard model*, . (Cited on page 8.)

[8] Y. V. Kovchegov and E. Levin, *Quantum Chromodynamics at High Energy*, Cambridge Monographs on Particle Physics, Nuclear Physics and Cosmology, Cambridge University Press, 2012. (Cited on pages 9 and 207.)

[9] Particle Data Group, P. A. Zyla *et al.*, *Review of particle physics*, Prog. Theor. Exp. Phys. **2020** (2020) 083C01. (Cited on pages 11 and 207.)

[10] P. D. Group, *Review of Particle Physics*, Progress of Theoretical and Experimental Physics **2020** (2020) 083C01. (Cited on pages 11 and 22.)

[11] J. Campbell, J. Huston, and F. Krauss, *The Black Book of Quantum Chromodynamics : a Primer for the LHC Era*, Oxford University Press, 2018. (Cited on page 12.)

[12] A. Metz and A. Vossen, *Parton Fragmentation Functions*, Prog. Part. Nucl. Phys. **91** (2016) 136, `arXiv:1607.02521`. (Cited on page 13.)

[13] D. de Florian, R. Sassot, and M. Stratmann, *Global analysis of fragmentation functions for pions and kaons and their uncertainties*, Phys. Rev. **D75** (2007) 114010, `arXiv:hep-ph/0703242`. (Cited on page 13.)

[14] D. de Florian *et al.*, *Parton-to-Pion Fragmentation Reloaded*, Phys. Rev. **D91** (2015) 014035, `arXiv:1410.6027`. (Cited on page 13.)

[15] NNPDF collaboration, V. Bertone *et al.*, *Charged hadron fragmentation functions from collider data*, Eur. Phys. J. **C78** (2018) 651, arXiv:1807.03310. (Cited on page 13.)

[16] V. N. Gribov and L. N. Lipatov, *Deep inelastic e p scattering in perturbation theory*, Sov. J. Nucl. Phys. **15** (1972) 438. (Cited on page 13.)

[17] G. Altarelli and G. Parisi, *Asymptotic Freedom in Parton Language*, Nucl. Phys. **B126** (1977) 298. (Cited on page 13.)

[18] Y. L. Dokshitzer, *Calculation of the Structure Functions for Deep Inelastic Scattering and e+ e- Annihilation by Perturbation Theory in Quantum Chromodynamics.*, Sov. Phys. JETP **46** (1977) 641. (Cited on page 13.)

[19] J. Gao, L. Harland-Lang, and J. Rojo, *The Structure of the Proton in the LHC Precision Era*, Phys. Rept. **742** (2018) 1, arXiv:1709.04922. (Cited on page 13.)

[20] T.-J. Hou *et al.*, *New cteq global analysis of quantum chromodynamics with high-precision data from the lhc*, Physical Review D **103** (2021) . (Cited on pages 13, 14, and 207.)

[21] M. Klasen and H. Paukkunen, *Nuclear parton distribution functions after the first decade of lhc data*, Annual Review of Nuclear and Particle Science (2024) . (Cited on page 13.)

[22] K. J. Eskola, P. Paakkinen, H. Paukkunen, and C. A. Salgado, *Epps21: a global qcd analysis of nuclear pdfs*, The European Physical Journal C **82** (2022) . (Cited on pages 15 and 207.)

[23] P. Braun-Munzinger, V. Koch, T. Schäfer, and J. Stachel, *Properties of hot and dense matter from relativistic heavy ion collisions*, Physics Reports **621** (2016) 76–126. (Cited on pages 14 and 16.)

[24] B. Friman *et al.*, eds., *The CBM physics book: Compressed baryonic matter in laboratory experiments*, vol. 814, Springer, 2011. (Cited on page 14.)

[25] HotQCD Collaboration, A. Bazavov *et al.*, *Equation of state in $(2+1)$-flavor qcd*, Phys. Rev. D **90** (2014) 094503. (Cited on pages 15, 132, 134, 135, 138, 146, 164, 211, and 213.)

[26] S. Navin, *Diffraction in pythia*, 2010. (Cited on pages 18 and 207.)

[27] TOTEM collaboration, G. Antchev *et al.*, *Measurement of proton-proton elastic scattering and total cross-section at $\sqrt{s} = 7\,\text{TeV}$*, EPL **101** (2013) 21002. (Cited on page 17.)

[28] TOTEM collaboration, G. Antchev *et al.*, *Luminosity-independent measurements of total, elastic and inelastic cross-sections at* $\sqrt{s} = 7$ *TeV*, EPL **101** (2013) 21004. (Cited on page 17.)

[29] U. Amaldi, M. Jacob, and G. Matthiae, *Diffraction of Hadronic Waves*, Ann. Rev. Nucl. Part. Sci. **26** (1976) 385. (Cited on page 18.)

[30] C. Bierlich *et al.*, *A comprehensive guide to the physics and usage of PYTHIA 8.3*, SciPost Phys. Codebases (2022) 8. (Cited on pages 19, 22, and 149.)

[31] A. Toia, *Heavy ion collisions scheme*, 2013. (Cited on pages 20 and 207.)

[32] F. Gelis, *The early stages of a high energy heavy ion collision*, Journal of Physics: Conference Series **381** (2012) 012021. (Cited on page 21.)

[33] Geant4 collaboration, S. Agostinelli *et al.*, *Geant4: A simulation toolkit*, Nucl. Instrum. Meth. **A506** (2003) 250. (Cited on pages 22, 44, and 149.)

[34] E. Bothmann *et al.*, *Event generation with sherpa 2.2*, SciPost Physics **7** (2019) . (Cited on page 22.)

[35] J. Bellm *et al.*, *Herwig 7.2 release note*, The European Physical Journal C **80** (2020) . (Cited on page 22.)

[36] T. Pierog *et al.*, *Epos lhc: Test of collective hadronization with data measured at the cern large hadron collider*, Phys. Rev. C **92** (2015) 034906. (Cited on pages 22, 23, 24, 44, 149, and 207.)

[37] T. Sjöstrand, *The development of mpi modelling in pythia*, 2017. (Cited on page 22.)

[38] C. Bierlich, G. Gustafson, L. Lönnblad, and H. Shah, *The angantyr model for heavy-ion collisions in pythia8*, Journal of High Energy Physics **2018** (2018) . (Cited on page 23.)

[39] S. A. Bass *et al.*, *Microscopic models for ultrarelativistic heavy ion collisions*, Prog. Part. Nucl. Phys. **41** (1998) 255, `arXiv:nucl-th/9803035`. (Cited on page 23.)

[40] E. Mobs, *The CERN accelerator complex - 2019. Complexe des accélérateurs du CERN - 2019*, `https://cds.cern.ch/record/2684277`, July, 2019. General Photo. (Cited on pages 26 and 207.)

[41] LHCb collaboration, C. Elsässer, $b\bar{b}$ *production angle plots*, https://lhcb.web.cern.ch/lhcb/speakersbureau/html/bb_ProductionAngles.html . (Cited on pages 28 and 207.)

[42] LHCb collaboration, R. Aaij *et al.*, *LHCb detector performance*, Int. J. Mod. Phys. **A30** (2015) 1530022, `arXiv:1412.6352`. (Cited on pages 29, 148, and 207.)

[43] LHCb collaboration, R. Aaij *et al.*, *Measurements of the branching fractions for $B^0_{(s)} \to D_{(s)}\pi\pi\pi$ and $\Lambda^0_b \to \Lambda^+_c\pi\pi\pi$*, Phys. Rev. **D84** (2011) 092001, Erratum ibid. **D85** (2012) 039904, arXiv:1109.6831. (Cited on page 29.)

[44] LHCb collaboration, R. Aaij *et al.*, *Measurement of the ratio of the $\mathcal{B}(B^0 \to D^{*-}\tau^+\nu_\tau)$ and $\mathcal{B}(B^0 \to D^{*-}\mu^+\nu_\mu)$ branching fractions using three-prong τ-lepton decays*, Phys. Rev. Lett. **120** (2018) 171802, arXiv:1708.08856. (Cited on page 29.)

[45] LHCb collaboration, R. Aaij *et al.*, *First evidence for the decay $B^0_s \to \mu^+\mu^-$*, Phys. Rev. Lett. **110** (2013) 021801, arXiv:1211.2674. (Cited on page 29.)

[46] LHCb collaboration, R. Aaij *et al.*, *Measurement of the relative rate of prompt χ_{c0}, χ_{c1} and χ_{c2} production at $\sqrt{s} = 7$ TeV*, JHEP **10** (2013) 115, arXiv:1307.4285. (Cited on page 29.)

[47] LHCb collaboration, R. Aaij *et al.*, *Measurement of prompt-cross-section ratio $\sigma(\chi_{c2})/\sigma(\chi_{c1})$ in pPb collisions at $\sqrt{s_{NN}} = 8.16$ TeV*, arXiv:2103.07349, to appear in Phys. Rev. C. (Cited on page 29.)

[48] LHCb collaboration, A. A. Alves Jr. *et al.*, *The LHCb detector at the LHC*, JINST **3** (2008) S08005. (Cited on pages 30, 37, 39, 148, 207, and 208.)

[49] R. Aaij *et al.*, *Performance of the LHCb Vertex Locator*, JINST **9** (2014) P09007, arXiv:1405.7808. (Cited on pages 30, 32, 207, and 208.)

[50] M. Alexander *et al.*, *Mapping the material in the LHCb vertex locator using secondary hadronic interactions*, JINST **13** (2018) P06008, arXiv:1803.07466. (Cited on page 32.)

[51] M. Kucharczyk, P. Morawski, and M. Witek, *Primary Vertex Reconstruction at LHCb*, LHCb-PUB-2014-044. (Cited on page 33.)

[52] R. Arink *et al.*, *Performance of the LHCb Outer Tracker*, JINST **9** (2014) P01002, arXiv:1311.3893. (Cited on page 35.)

[53] P. d'Argent *et al.*, *Improved performance of the LHCb Outer Tracker in LHC Run 2*, JINST **12** (2017) P11016, arXiv:1708.00819. (Cited on page 35.)

[54] O. Callot, *FastVelo, a fast and efficient pattern recognition package for the Velo*, tech. rep., CERN, Geneva, Jan, 2011. LHCb. (Cited on page 36.)

[55] O. Callot and S. Hansmann-Menzemer, *The Forward Tracking: Algorithm and Performance Studies*, tech. rep., CERN, Geneva, May, 2007. (Cited on page 36.)

[56] M. Needham, *Performance of the Track Matching*, tech. rep., CERN, Geneva, Oct, 2007. (Cited on page 36.)

[57] R. Fruhwirth, *Application of Kalman filtering to track and vertex fitting*, Nucl. Instrum. Meth. A **262** (1987) 444. (Cited on page 36.)

[58] LHCb collaboration, R. Aaij *et al.*, *Measurement of the track reconstruction efficiency at LHCb*, JINST **10** (2015) P02007, arXiv:1408.1251. (Cited on page 37.)

[59] M. Adinolfi *et al.*, *Performance of the LHCb RICH detector at the LHC*, Eur. Phys. J. **C73** (2013) 2431, arXiv:1211.6759. (Cited on page 38.)

[60] R. Aaij *et al.*, , LHCb-DP-2020-001, in preparation. (Cited on page 40.)

[61] R. Aaij *et al.*, *The LHCb trigger and its performance in 2011*, JINST **8** (2013) P04022, arXiv:1211.3055. (Cited on page 40.)

[62] F. Archilli *et al.*, *Performance of the muon identification at LHCb*, JINST **8** (2013) P10020, arXiv:1306.0249. (Cited on page 41.)

[63] LHCb Collaboration, R. Antunes-Nobrega and França-Barbosa, *LHCb trigger system: Technical Design Report*, Technical Design Report LHCb, CERN, Geneva, 2003. revised version number 1 submitted on 2003-09-24 12:12:22. (Cited on pages 42 and 208.)

[64] J. Albrecht, *The lhcb trigger system*, Nuclear Physics B - Proceedings Supplements **187** (2009) 237, Proceedings of the 8th International Conference on Beauty, Charm and Hyperons in Hadronic Interactions. (Cited on pages 43 and 216.)

[65] M. Clemencic *et al.*, *The LHCb simulation application, Gauss: Design, evolution and experience*, J. Phys. Conf. Ser. **331** (2011) 032023. (Cited on pages 44 and 149.)

[66] T. Sjöstrand, S. Mrenna, and P. Skands, *A brief introduction to PYTHIA 8.1*, Comput. Phys. Commun. **178** (2008) 852, arXiv:0710.3820. (Cited on page 44.)

[67] D. J. Lange, *The EvtGen particle decay simulation package*, Nucl. Instrum. Meth. **A462** (2001) 152. (Cited on pages 44 and 149.)

[68] Geant4 collaboration, J. Allison *et al.*, *Geant4 developments and applications*, IEEE Trans. Nucl. Sci. **53** (2006) 270. (Cited on pages 44 and 149.)

[69] M. Clemencic *et al.*, *The lhcb simulation application, gauss: Design, evolution and experience*, Journal of Physics: Conference Series **331** (2011) 032023. (Cited on pages 45 and 208.)

[70] LHCb Collaboration, R. e. a. Aaij, *Measurement of the nuclear modification factor and prompt charged particle production in p-pb and pp collisions at* $\sqrt{s_{NN}} = 5$ TeV, Phys. Rev. Lett. **128** (2022) 142004. (Cited on pages 48, 55, 57, 62, 63, 66, 77, 88, 109, 139, 140, 142, 208, and 212.)

[71] T. Pierog *et al.*, *Epos lhc: Test of collective hadronization with data measured at the cern large hadron collider*, Phys. Rev. C **92** (2015) 034906. (Cited on page 49.)

[72] LHCb Collaboration, R. e. a. Aaij, *Observation of strangeness enhancement with charmed mesons in high-multiplicity pPb collisions at $\sqrt{s_{\rm NN}} = 8.16$ TeV*, Phys. Rev. D **110** (2024) L031105. (Cited on page 50.)

[73] M. Kucharczyk, P. Morawski, and M. Witek, *Primary Vertex Reconstruction at LHCb*, tech. rep., CERN, Geneva, 2014. (Cited on pages 56, 58, and 208.)

[74] LHCbcollaboration, R. Aaij, B. Adeva, and Adinolfi, *LHCb Detector Performance*, Int. J. Mod. Phys. A **30** (2015) 1530022, `arXiv:1412.6352`. (Cited on page 59.)

[75] T. L. collaboration, *Measurement of the track reconstruction efficiency at lhcb*, Journal of Instrumentation **10** (2015) P02007. (Cited on pages 59, 63, 78, 208, and 209.)

[76] R. A. et al, *Performance of the lhcb vertex locator*, Journal of Instrumentation **9** (2014) P09007. (Cited on page 61.)

[77] E. Rodrigues, *Dealing with clones in the tracking*, tech. rep., CERN, Geneva, 2006. (Cited on page 61.)

[78] M. Needham, *Clone Track Identification using the Kullback-Liebler Distance*, tech. rep., CERN, Geneva, 2008. (Cited on page 61.)

[79] M. De Cian, S. Farry, P. Seyfert, and S. Stahl, *Fast neural-net based fake track rejection in the LHCb reconstruction*, tech. rep., CERN, Geneva, 2017. (Cited on page 62.)

[80] A. Rogozhnikov, *Reweighting with boosted decision trees*, Journal of Physics: Conference Series **762** (2016) 012036. (Cited on page 65.)

[81] ALICE Collaboration, A. Collaboration, *The ALICE definition of primary particles*, . (Cited on pages 69, 105, 137, and 147.)

[82] S. G. Weber, A. Dubla, A. Andronic, and A. Morsch, *Elucidating the multiplicity dependence of J/ψ production in proton–proton collisions with pythia8*, The European Physical Journal C **79** (2019) 36. (Cited on page 71.)

[83] The ALICE, S. Acharya and Adamová, *Forward rapidity J/ψ production as a function of charged-particle multiplicity in pp collisions at $\sqrt{s} = 5.02$ and 13 TeV*, Journal of High Energy Physics **2022** (2022) 15. (Cited on page 71.)

[84] J. F. Grosse-Oetringhaus and K. Reygers, *Charged-Particle Multiplicity in Proton-Proton Collisions*, J. Phys. G **37** (2010) 083001, `arXiv:0912.0023`. (Cited on page 72.)

[85] L. Brenner *et al.*, *Comparison of unfolding methods using roofitunfold*, International Journal of Modern Physics A **35** (2020) 2050145, arXiv:https://doi.org/10.1142/S0217751X20501456. (Cited on page 72.)

[86] A. collaboration, *Transverse momentum spectra of charged particles in proton–proton collisions at* $\sqrt{s} = 900\,\text{GeV}$ *with alice at the lhc*, Physics Letters B **693** (2010) 53. (Cited on page 72.)

[87] G. D'Agostini, *Improved iterative bayesian unfolding*, 2010. (Cited on page 72.)

[88] ALICE, J. Adam *et al.*, *Enhanced production of multi-strange hadrons in high-multiplicity proton-proton collisions*, Nature Phys. **13** (2017) 535, arXiv:1606.07424. (Cited on page 77.)

[89] D. Enterria and C. Loizides, *Progress in the glauber model at collider energies*, Annual Review of Nuclear and Particle Science **71** (2021) 315. (Cited on page 103.)

[90] G. Baym, L. L. Frankfurt, and M. Strikman, *Color fluctuations in hadrons and nucleus-nucleus collisions*, Nuclear Physics A **566** (1994) 149. (Cited on page 104.)

[91] L. McLerran and P. Tribedy, *Intrinsic fluctuations of the proton saturation momentum scale in high multiplicity p+p collisions*, Nuclear Physics A **945** (2016) 216. (Cited on page 104.)

[92] CMS, V. Khachatryan *et al.*, *Observation of Long-Range Near-Side Angular Correlations in Proton-Proton Collisions at the LHC*, JHEP **09** (2010) 091, arXiv:1009.4122. (Cited on page 104.)

[93] CMS, *Observation of long-range, near-side angular correlations in ppb collisions at the lhc*, Physics Letters B **718** (2013) 795. (Cited on page 104.)

[94] J. e. a. Adam, *Enhanced production of multi-strange hadrons in high-multiplicity proton–proton collisions*, Nature Physics **13** (2017) 535–539. (Cited on page 104.)

[95] LHCb Collaboration, R. e. a. Aaij, *Observation of multiplicity dependent prompt* $\chi_{c1}(3872)$ *and* $\psi(2s)$ *production in pp collisions*, Phys. Rev. Lett. **126** (2021) 092001. (Cited on page 104.)

[96] LHCb Collaboration, R. e. a. Aaij, *Observation of strangeness enhancement with charmed mesons in high-multiplicity* pPb *collisions at* $\sqrt{s_{\text{NN}}} = 8.16$ TeV, Phys. Rev. D **110** (2024) L031105. (Cited on page 104.)

[97] LHCb Collaboration, R. e. a. Aaij, *Evidence for modification of b quark hadronization in high-multiplicity pp collisions at* $\sqrt{s} = 13$ TeV, Phys. Rev. Lett. **131** (2023) 061901. (Cited on page 104.)

[98] LHCb Collaboration, R. e. a. Aaij, *Enhanced production of Λ_b^0 baryons in high-multiplicity pp collisions at $\sqrt{s} = 13$ TeV*, Phys. Rev. Lett. **132** (2024) 081901. (Cited on page 104.)

[99] R. e. a. Aaij, *Multiplicity dependence of $\sigma\psi(2s)/\sigma j/\psi$ in pp collisions at $\sqrt{s} = 13$ tev*, Journal of High Energy Physics **2024** (2024) . (Cited on page 104.)

[100] S. A. et al, *System-size dependence of the charged-particle pseudorapidity density at $\sqrt{s_{NN}} = 5.02\,\text{TeV}$ for pp, ppb, and pbpb collisions*, Physics Letters B **845** (2023) 137730. (Cited on pages 124, 126, and 211.)

[101] L. Collaboration, *Framework TDR for the LHCb Upgrade II- Opportunities in flavour physics, and beyond, in the HL-LHC era*, tech. rep., CERN, Geneva, 2021. (Cited on pages 131, 134, and 165.)

[102] LHCb, R. Aaij and et al. *Physics case for an LHCb Upgrade II - Opportunities in flavour physics, and beyond, in the HL-LHC era*, tech. rep., CERN, Geneva, 2016. ISBN 978-92-9083-494-6, doi: 10.17181/CERN.QZRZ.R4S6. (Cited on page 131.)

[103] J. Adam, D. Adamová, and M. M. A. et al, *Centrality dependence of the pseudorapidity density distribution for charged particles in pb–pb collisions at $\sqrt{s_{NN}} =$ 5.02 TeV*, Physics Letters B **772** (2017) 567. (Cited on pages 132, 133, 138, and 164.)

[104] F. G. Gardim, G. Giacalone, M. Luzum, and J.-Y. Ollitrault, *Thermodynamics of hot strong-interaction matter from ultrarelativistic nuclear collisions*, Nature Phys. **16** (2020) 615, `arXiv:1908.09728`. (Cited on pages 132, 133, and 138.)

[105] R. Hagedorn, *Multiplicities, pt distributions and the expected hadron → quark-gluon phase transition*, La Rivista del Nuovo Cimento (1978-1999) **6** (1983) 1. (Cited on page 134.)

[106] T. C. Collaboration, *Extracting the speed of sound in quark–gluon plasma with ultra-relativistic lead–lead collisions at the lhc*, Reports on Progress in Physics **87** (2024) 077801. (Cited on pages 134, 135, 143, 164, 165, 211, and 213.)

[107] P. Arnold, G. D. Moore, and L. G. Yaffe, *Transport coefficients in high temperature gauge theories, 2. beyond leading log*, Journal of High Energy Physics **2003** (2003) 051. (Cited on page 134.)

[108] M. Prakash, M. Prakash, R. Venugopalan, and G. Welke, *Non-equilibrium properties of hadronic mixtures*, Physics Reports **227** (1993) 321. (Cited on page 134.)

[109] G. Denicol, A. Monnai, and B. Schenke, *Moving forward to constrain the shear viscosity of qcd matter*, Phys. Rev. Lett. **116** (2016) 212301. (Cited on pages 134, 136, and 211.)

[110] B. B. e. a. Back, *Energy dependence of elliptic flow over a large pseudorapidity range in* Au + Au *collisions at the bnl relativistic heavy ion collider*, Phys. Rev. Lett. **94** (2005) 122303. (Cited on pages 136 and 211.)

[111] PHOBOS Collaboration, B. B. e. a. Back, *Centrality and pseudorapidity dependence of elliptic flow for charged hadrons in au+au collisions at* $\sqrt{s_{NN}} = 200$ GeV, Phys. Rev. C **72** (2005) 051901. (Cited on pages 136 and 211.)

[112] R. Aaij and et al. *Measurement of charged particle multiplicities in pp collisions at* $\sqrt{s} = 7$ *TeV in the forward region*, The European Physical Journal C **72** (2012) . (Cited on page 138.)

[113] ALICE, S. Acharya *et al.*, *Transverse momentum spectra and nuclear modification factors of charged particles in pp, p-Pb and Pb-Pb collisions at the LHC*, JHEP **11** (2018) 013, arXiv:1802.09145. (Cited on pages 142, 147, and 212.)

[114] CMS collaboration, V. Khachatryan *et al.*, *Nuclear Effects on the Transverse Momentum Spectra of Charged Particles in pPb Collisions at* $\sqrt{s_{\rm NN}} = 5.02$ *TeV*, Eur. Phys. J. **C75** (2015) 237, arXiv:1502.05387. (Cited on pages 142, 147, and 212.)

[115] CMS collaboration, V. Khachatryan *et al.*, *Charged-particle nuclear modification factors in PbPb and pPb collisions at* $\sqrt{s_{\rm N\,N}} = 5.02$ *TeV*, JHEP **04** (2017) 039, arXiv:1611.01664. (Cited on pages 142 and 212.)

[116] O. Boente García, *Analysis of charged particle production in proton-nucleus and proton-proton collisions at the LHCb experiment*, PhD thesis, Santiago de Compostela U., 2021. (Cited on pages 142 and 212.)

[117] F. O. Durães, A. V. Giannini, V. P. Gonçalves, and F. S. Navarra, *The color glass condensate and the average transverse momentum in proton-nucleus collisions*, Journal of Physics: Conference Series **736** (2016) 012025. (Cited on page 141.)

[118] J. W. Cronin *et al.*, *Production of hadrons at large transverse momentum at* 200, 300, *and* 400 GeV, Phys. Rev. D **11** (1975) 3105. (Cited on page 142.)

[119] B. Z. Kopeliovich, J. Nemchik, A. Schäfer, and A. V. Tarasov, *Cronin effect in hadron production off nuclei*, Phys. Rev. Lett. **88** (2002) 232303. (Cited on page 142.)

[120] B. e. a. Abelev, *Transverse momentum distribution and nuclear modification factor of charged particles in* pPb *collisions at* $\sqrt{s_{NN}}$=5.02 TeV, Physical Review Letters **110** (2013) . (Cited on page 142.)

[121] F. G. Gardim, A. V. Giannini, and J.-Y. Ollitrault, *Accessing the speed of sound in relativistic ultracentral nucleus-nucleus collisions using the mean transverse momentum*, Phys. Lett. B **856** (2024) 138937, arXiv:2403.06052. (Cited on page 143.)

[122] A. B. Kaidalov and K. A. Ter-Martirosyan, *Multihadron production at high energies in the model of quark gluon strings*, Sov. J. Nucl. Phys. **40** (1984) 135. (Cited on page 146.)

[123] A. Capella, U. Sukhatme, C.-I. Tan, and J. Tran Thanh Van, *Dual parton model*, Physics Reports **236** (1994) 225. (Cited on page 146.)

[124] J. L. Albacete *et al.*, *Predictions for pPb Collisions at* $\sqrt{s_{NN}} = 5\,\text{TeV}$, Int. J. Mod. Phys. **E22** (2013) 1330007, arXiv:1301.3395. (Cited on page 146.)

[125] J. L. Albacete *et al.*, *Predictions for cold nuclear matter effects in pPb collisions at* $\sqrt{s_{NN}} = 8.16\,\text{TeV}$, Nucl. Phys. **A972** (2018) 18, arXiv:1707.09973. (Cited on page 146.)

[126] K. J. Eskola, P. Paakkinen, H. Paukkunen, and C. A. Salgado, *EPPS16: Nuclear parton distributions with LHC data*, Eur. Phys. J. **C77** (2017) 163, arXiv:1612.05741. (Cited on page 146.)

[127] K. Kovarik *et al.*, *nCTEQ15 - Global analysis of nuclear parton distributions with uncertainties in the CTEQ framework*, Phys. Rev. **D93** (2016) 085037, arXiv:1509.00792. (Cited on page 146.)

[128] D. de Florian, R. Sassot, P. Zurita, and M. Stratmann, *Global analysis of nuclear parton distributions*, Phys. Rev. **D85** (2012) 074028, arXiv:1112.6324. (Cited on page 146.)

[129] H. Kowalski, T. Lappi, and R. Venugopalan, *Nuclear enhancement of universal dynamics of high parton densities*, Phys. Rev. Lett. **100** (2008) 022303, arXiv:0705.3047. (Cited on page 146.)

[130] ATLAS, G. Aad *et al.*, *Transverse momentum, rapidity, and centrality dependence of inclusive charged-particle production in* $\sqrt{s_{NN}} = 5.02\,\text{TeV}$ *pPb collisions measured by the ATLAS experiment*, Phys. Lett. B **763** (2016) 313, arXiv:1605.06436. (Cited on pages 147 and 212.)

[131] BRAHMS, I. Arsene *et al.*, *On the evolution of the nuclear modification factors with rapidity and centrality in dAu collisions at* $\sqrt{s_{NN}} = 200\,\text{GeV}$, Phys. Rev. Lett. **93** (2004) 242303, arXiv:nucl-ex/0403005. (Cited on pages 147 and 212.)

[132] PHENIX, C. Aidala *et al.*, *Nuclear modification factor of charged hadrons at forward and backward rapidity in pAl and pAu collisions at* $\sqrt{s_{NN}} = 200\,\text{GeV}$, Phys. Rev. C **101** (2020) 034910, arXiv:1906.09928. (Cited on pages 147 and 212.)

[133] L. D. McLerran and R. Venugopalan, *Gluon distribution functions for very large nuclei at small transverse momentum*, Phys. Rev. **D49** (1994) 3352, arXiv:hep-ph/9311205. (Cited on page 147.)

[134] LHCb Collaboration, P. R. Barbosa-Marinho and Bediaga, *LHCb outer tracker: Technical Design Report*, Technical Design Report LHCb, CERN, Geneva, 2001. (Cited on page 208.)

The study of prompt charged particle production in hadronic collisions is a tool to understand particle production in QCD and the hadronic structure at high energy. Since the observation of collective effects in proton-proton collisions, the formation of a thermal medium in small collision systems is also an open question. In this context, characterising charged particle properties in different systems is crucial for solving these discussions. This thesis includes two measurements of prompt charged particle production from proton-proton and proton-lead collisions at 5 TeV recorded at the LHCb experiment. One is the multiplicity distribution and the other is the multiplicity dependence of the average transverse momentum. These measurements are sensitive to gluon saturation and the formation of a deconfined medium.