



Politecnico
di Torino

ScuDo

Scuola di Dottorato ~ Doctoral School

WHAT YOU ARE, TAKES YOU FAR

Doctoral Dissertation

Doctoral Program in High Energy Physics (36th cycle)

Heavy quark study via prompt and non-prompt strange D-meson measurements in pp and Pb–Pb collisions with ALICE at the LHC

By

Stefano Politanò

Supervisor(s):

Prof. S. Bufalino, Supervisor

Prof. F. Prino, Co-Supervisor

Doctoral Examination Committee:

Prof. V. Greco, Referee, University of Catania

Dr. B. Audurier, Referee, Departement de Physique Nucleaire-CEA, Saclay

Politecnico di Torino

2024



Declaration

I hereby declare that, the contents and organization of this dissertation constitute my own original work and does not compromise in any way the rights of third parties, including those relating to the security of personal data.

Stefano Politanò
2024

* This dissertation is presented in partial fulfillment of the requirements for **Ph.D. degree** in the Graduate School of Politecnico di Torino (ScuDo).

A mia madre, Luisa.

Abstract

This Thesis focuses on the study of the strange D-mesons production in proton-proton (pp) and lead-lead (Pb–Pb) collisions employing the data collected by the ALICE experiment at the CERN Large Hadron Collider. The study of the production of D_s^+ mesons in small collision systems at LHC energies is extremely important to test the predictions based on pQCD calculations and to provide a benchmark for the heavy-ion measurements where, given the high energy densities and temperatures, the nuclear matter is expected to undergo a phase transition to a colour-deconfined state called Quark-Gluon Plasma (QGP). Considering the small distance travelled by the D_s^+ meson inside the ALICE apparatus ($\sim 151 \mu\text{m}$), to perform the analyses presented in this Thesis the D_s^+ mesons were reconstructed in their hadronic decay channel $D_s^+ \rightarrow \phi \pi^+ \rightarrow K^+ K^- \pi^+$ from their decay products. This is made possible by exploiting the excellent spatial resolution of the ALICE detector, which allows us to separate the D_s^+ meson decay vertex from the interaction point of the LHC beams. In this Thesis, multiclass classification algorithms are employed to select the D_s^+ meson candidates. This approach not only allows for optimal separation between the signal and the background, mainly derived from the wrong association of triplets of charged tracks, but it also provides a powerful tool to discriminate between the D_s^+ mesons originated from the charm quark hadronisation, i.e. prompt, and from the decay of beauty hadrons, i.e. non-prompt. The latter consists of $\sim 5\%$ of the D_s^+ meson produced in Nature, making the study of their production a challenging task. To this end, a Boosted Decision Tree algorithm is adopted, given its robustness, its efficient performance, and the minimal number of input variables required for the algorithm training. In this Thesis, the measurement of prompt and non-prompt D_s^+ meson p_T -differential production cross sections in pp collisions at $\sqrt{s} = 13 \text{ TeV}$ are presented. The measurements are in agreement with the predictions based on pQCD calculations. In addition, by measuring the relative production of strange and non-strange D-mesons, the fragmentation fractions of charm and beauty quarks were measured. Furthermore, a preliminary measurement of the production of excited D_s^+ meson states, the D_{s1}^+ and the D_{s2}^{*+} mesons, is presented for the first time at the LHC energies. Moreover, the first measurement of non-prompt D_s^+ meson production in central and peripheral Pb–Pb

collisions at $\sqrt{s_{\text{NN}}} = 5.02$ TeV is presented. The measurement is compared to the one of prompt D_s^+ mesons and non-prompt D^0 mesons and the ratios of the nuclear modification factor are calculated. In central collisions, where the QGP-induced effects are expected to be more pronounced, the ratios exhibit a hint of enhancement in the $4 < p_T < 12$ GeV/ c interval, with a statistical significance of 1.6 and 1.7σ , respectively. Finally, a preliminary measurement of the D_s^+ meson production in pp collisions at $\sqrt{s} = 13.6$ TeV obtained using the data collected by the ALICE experiment in the current data-taking period, Run 3, is presented.

Contents

1	Quarks, gluons, and open heavy flavour	1
1.1	Quantum Chromodynamics	1
1.2	Quark-Gluon Plasma	3
1.3	Heavy-Ion collisions	5
1.3.1	Collision evolution	7
1.3.2	The geometry of the collision: Glauber Model	10
1.4	Properties of the QGP	13
1.4.1	Charged particle multiplicity	14
1.4.2	Identified hadron abundances: the chemical freeze-out temperature	15
1.4.3	Strangeness enhancement	17
1.4.4	Thermal freeze-out temperature and radial flow	20
1.4.5	Azimuthal anisotropy: elliptic flow	22
1.4.6	High- p_T hadrons and jet quenching	25
1.4.7	Quarkonium suppression	28
1.5	Open heavy-flavour hadrons	29
1.5.1	Open heavy-flavour production in pp collision	30
1.5.2	Cold nuclear matter effects	32
1.5.3	Interaction of heavy quarks with the QGP	35
1.6	Azimuthal anisotropies	39
1.7	Hadronisation	42

2	ALICE: A Large Ion Collider Experiment	44
2.1	The Large Hadron Collider	44
2.2	The ALICE experiment	46
2.2.1	The Inner Tracking System (ITS)	48
2.2.2	The Time Projection Chamber (TPC)	51
2.2.3	The Time of Flight (TOF)	53
2.2.4	V0 detector	54
2.2.5	T0 detector	55
2.2.6	The Zero Degree Calorimeter (ZDC)	56
2.3	The ALICE offline framework	56
2.4	The centrality determination in Pb–Pb collisions	58
2.5	Track and vertex reconstruction	60
2.5.1	Secondary vertex reconstruction	64
2.6	Particle identification	65
2.6.1	ITS	65
2.6.2	TPC	66
2.6.3	TOF	67
3	Machine Learning	69
3.1	Machine Learning in a nutshell	69
3.1.1	Receiver Operating Characteristic	73
3.1.2	Cross Validation	74
3.1.3	Hyperparameters Optimisation	75
3.1.4	Feature Importance	77
3.2	Boosted Decision Tree	79
3.2.1	XGBoost	80
3.3	Software tools for Machine Learning analyses	82

4	Strange D meson reconstruction in pp and Pb–Pb collision	84
4.1	Data sample and event selection	86
4.1.1	Data taking and trigger conditions	86
4.1.2	Offline event selection	87
4.2	Strange D meson reconstruction	89
4.2.1	D_s^+ meson secondary vertex reconstruction	89
4.2.2	D_{s1}^+ and D_{s2}^{*+} resonances reconstruction	92
4.2.3	Track-quality selections	93
4.2.4	Topological and kinematic selections	94
4.2.5	Particle identification selections	101
4.3	Monte Carlo simulations	103
4.3.1	Event generation and particle transport	103
4.3.2	Acceptance and efficiency corrections	105
5	Strange D-meson production in pp collisions at $\sqrt{s} = 13$ TeV	107
5.1	Machine learning for the D_s^+ meson signal extraction	110
5.1.1	Training, testing, and application sets	111
5.1.2	Training features and BDT performance evaluation	113
5.1.3	BDT-based selection optimisation	116
5.2	Prompt and non-prompt D_s^+ meson production in pp collisions at $\sqrt{s} = 13$ TeV	119
5.2.1	Raw-yield extraction	119
5.2.2	Efficiency and acceptance corrections	122
5.2.3	Prompt and non-prompt D_s^+ meson fraction	123
5.3	Systematic uncertainties	129
5.3.1	Systematic on the raw-yield extraction	131
5.3.2	MC-Data discrepancy	132
5.4	Prompt and non-prompt D_s^+ meson results	138
5.4.1	Prompt D-meson p_T -differential cross sections	138

5.4.2	Energy dependence of prompt D meson cross section ratios	140
5.4.3	Rapidity dependence of prompt D-meson ratios	147
5.4.4	Charm-quark production in pp collisions at $\sqrt{s} = 13$ TeV	150
5.4.5	Non-prompt D-meson p_T -differential cross sections	154
5.4.6	Non-prompt D meson cross section ratios	156
5.4.7	Beauty-quark production in pp collisions at $\sqrt{s} = 13$ TeV	159
5.5	D_{s1}^+ and D_{s2}^{*+} meson production in pp collisions at $\sqrt{s} = 13$ TeV	164
5.5.1	D_{s1}^+ and D_{s2}^{*+} candidate selection	164
5.5.2	Raw yield extraction	166
5.5.3	Efficiency and acceptance correction	167
5.5.4	Non-prompt D-meson fraction estimation	171
5.5.5	Systematic uncertainties	172
5.5.6	Branching ratio	176
5.5.7	Results	176
6	Strange D-meson production in Pb–Pb collisions at $\sqrt{s_{NN}} = 5.02$ TeV	180
6.1	Machine-learning for the D_s^+ meson signal extraction	182
6.1.1	Training, testing, and application sets	182
6.1.2	BDT-based selection optimisation	184
6.2	Non-prompt D_s^+ meson production in Pb–Pb collisions at $\sqrt{s_{NN}} = 5.02$ TeV	187
6.2.1	Raw-yield extraction	187
6.2.2	Acceptance-times-efficiency corrections	189
6.2.3	Non-prompt D_s^+ meson fraction	192
6.2.4	Proton-proton reference	194
6.3	Systematic uncertainties	196
6.3.1	Systematic uncertainty on the raw-yield extraction	197
6.3.2	Systematic uncertainty on the track-reconstruction efficiency	199
6.3.3	Systematic uncertainty on the non-prompt fraction estimation	199

6.3.4	Systematic uncertainty on the D-meson selection efficiency and PID efficiency	200
6.3.5	Systematic uncertainty on the D-meson simulated p_T shape	200
6.3.6	Systematic uncertainty on the B-meson relative abundances	201
6.4	Results	202
7	Conclusions and outlook	211
7.1	Conclusions	211
7.2	Outlook	213
	List of Tables	220
	List of Figures	223
	References	238

Chapter 1

Quarks, gluons, and open heavy flavour

Developed within the last century, the *Standard Model* is the most widely accepted theoretical framework in particle physics to describe the electromagnetic, weak, and strong nuclear forces, as well as the elementary particles that interact through these forces. The strong interaction, which governs the behaviour of quarks and gluons, is described by the *Quantum Chromodynamics* (QCD). The strong interaction is responsible for binding quarks and gluons together to form protons, neutrons, and other composite particles called *hadrons*. The main features of the theory are presented in Section 1.1. According to QCD calculation on the lattice, there exists a phase transition in strongly interacting matter from the hadronic phase to a deconfined state known as the *Quark-Gluon Plasma* (QGP) in presence of high temperature and energy density conditions. This unique state characterised the early universe approximately 10 microseconds after the Big Bang and can be reproduced in high energy heavy-ion collisions. The properties the QGP are discussed in Section 1.2, while in Section 1.3 the main characteristics of heavy-ion collisions are introduced. The chapter concludes with Section 1.4, highlighting the key signatures of QGP formation, and with Section 1.5 dedicated to the measurements of heavy-flavour hadrons, which are fundamental probes of the QGP.

1.1 Quantum Chromodynamics

The strong interaction is described by *Quantum Chromodynamics* (QCD), which is a non-Abelian quantum gauge field theory invariant under local $SU(3)$ colour-group transformations [1]. The Lagrangian symmetry under $SU(3)_c$ transformations introduces a conserved

charge called *colour*. The QCD Lagrangian density is given by [2]:

$$\mathcal{L}_{QCD} = \sum_f \bar{q}_f^i [i\gamma^\mu (D_\mu)_{ij} - m_f \delta_{ij}] q_f^j - \frac{1}{4} G_{\mu\nu}^a G_a^{\mu\nu}. \quad (1.1)$$

The first term of the Lagrangian density describes the free quark field. The q_f and m_f symbols refer to the quark field of *flavour* f , which can assume six distinct values (up, down, strange, charm, beauty, and top), and its corresponding mass, respectively. The quark field is represented by a Dirac spinor, which is a four-component complex vector. The index i refers to the colour of the quark, which can assume three different values (red, green, and blue). γ^μ are the Dirac matrices, δ_{ij} is the Dirac delta, and D_μ is the covariant derivative. The covariant derivative is defined as:

$$D_\mu = \partial_\mu - ig_s \frac{\lambda_a}{2} A_\mu^a, \quad (1.2)$$

where the intensity of the interaction is represented by the adimensional constant g_s that relates to the *strong coupling constant* (α_s) by $g_s = \sqrt{4\pi\alpha_s}$. λ_a are the eight generators of the $SU(3)_c$ group, called *Gell-Mann matrices*. The gluon field is represented by A_μ^a , where $a = 1, 2, \dots, 8$. The kinematics and dynamics of gluons are described in the second term of Eq. 1.1 by the gluon field tensor $G_{\mu\nu}^a$, which is defined as:

$$G_{\mu\nu}^a = \partial_\mu A_\nu^a - \partial_\nu A_\mu^a + f_{bc}^a A_\mu^b A_\nu^c, \quad (1.3)$$

where the structure constant of the $SU(3)_c$ group is denoted as f_{bc}^a . The non-Abelian nature of the $SU(3)_c$ group is a fundamental aspect of QCD. Unlike the U(1) symmetry of the Quantum Electrodynamics (QED), where the structure constant is null and the gauge boson (photon) does not interact with itself, the non-Abelian nature of the $SU(3)_c$ group introduces non-null structure constants and self-interactions of gluons are permitted. Hence, in addition to the quark-gluon vertices, gluons can interact with each other through three- and four-gluon vertices. The existence of these vertices has a huge impact on the phenomenology of the strong interaction. Even though the gluons are massless, like the photon, the strong interaction is a short-range force, in contrast to the long-range electromagnetic force. Figure 1.1 shows the Feynman diagrams allowed in QCD.

The interactions among gluon fields themselves give rise to a phenomenon known as *anti-screening* in colour interactions. Consequently, the QCD coupling constant strongly

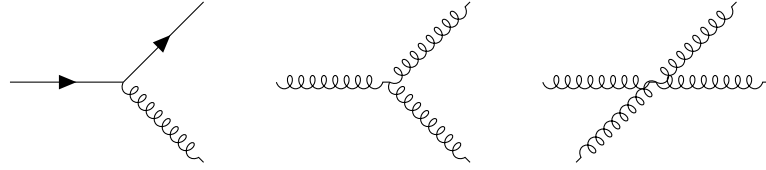


Fig. 1.1 Quark-gluon and gluon-only vertices in QCD.

depends on the transferred momentum $\mathcal{Q}$ [3, 4]:

$$\alpha_s(\mathcal{Q}^2) = \frac{\alpha_s(\mu^2)}{1 + \frac{\alpha_s(\mu^2)}{12\pi} (33 - 2n_f) \ln(\mathcal{Q}^2/\mu^2)}. \quad (1.4)$$

The variable n_f represents the number of flavours that contribute to 1-loop corrections (i.e. those for which $m_f < |\mathcal{Q}|$), while μ represents the renormalisation mass scale, which is not determined by the theory itself. The dependence of α_s on $\mathcal{Q}^2$ was experimentally determined as reported in Fig. 1.2. At high values of $\mathcal{Q}$, corresponding to short distances, $\alpha_s(\mathcal{Q}^2)$ decreases, allowing quarks to be treated as nearly free particles. This regime is called *asymptotic freedom* and permits a perturbative description known as *perturbative QCD* (pQCD). On the other hand, at low values of $\mathcal{Q}$ ($\mathcal{Q} \lesssim 1 \text{ GeV}/c$), the strong coupling constant becomes larger and larger, strongly constraining quarks within hadrons. In this regime, referred to as *confinement*, a perturbative treatment is not applicable. Nonetheless, alternative theoretical methods can be employed to study the strong interaction in the non-perturbative regime of QCD. One approach is to compute the QCD Lagrangian's Green's functions on a discrete space-time grid, as done in *lattice QCD* (lQCD) [5]. This theory utilizes a discrete space-time structure represented by a complex-time hypercubic lattice, with lattice sites occupied by quarks and linked by gluons. The space-time discretisation offers two main advantages. Firstly, it helps to avoid the ultraviolet divergences that occur in perturbative QCD calculations by introducing a finite step-size for the lattice. Secondly, it enables the rewriting of the partition function of the system in the form of path integrals by introducing an imaginary time coordinate. The predictions for various physical phenomena are obtained in the limit of the lattice spacing a approaching zero, allowing for a continuum limit.

1.2 Quark-Gluon Plasma

To explore the phase diagram of the strongly interacting matter, lQCD calculations can be employed, examining its behaviour in relation to temperature T and *baryon-chemical potential* μ_B . The baryon-chemical potential is defined as the energy required to increase the

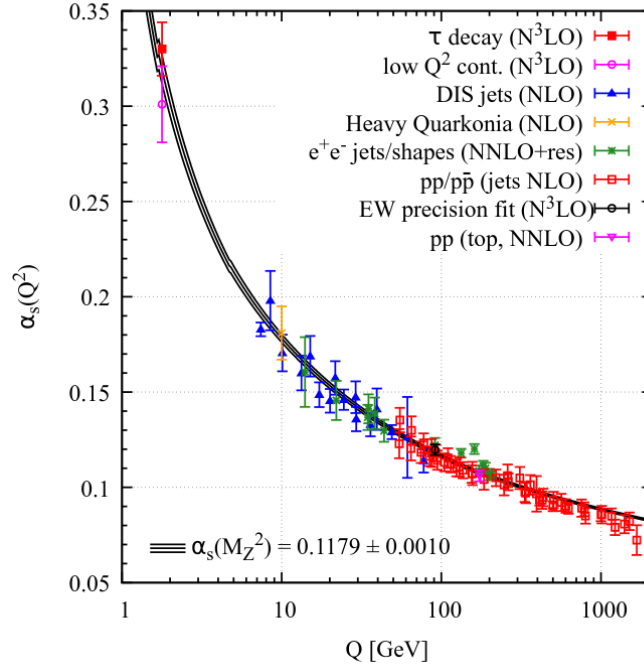


Fig. 1.2 Summary of measurements of α_s as a function of the momentum transfer Q . The respective degree of QCD perturbation theory used in the extraction of α_s is indicated in brackets (NLO: next-to-leading order; NNLO: next-to-next-to-leading order; NNLO+res.: NNLO matched to a resummed calculation; N³LO: next-to-NNLO). Figure from Ref. [6].

total baryon number by one unit and is proportional to the net baryon density. A schematic view of the QCD phase diagram is reported in Fig. 1.3.

Ordinary nuclear matter is characterised by $T \approx 0$ MeV and $\mu_B \approx 1$ GeV. As we move towards higher values of μ_B , corresponding to a system pressure exceeding 10^{32} Pa, nuclear matter can be described as a degenerate gas of neutrons. This phase is believed to form in the core of neutron stars [7]. Further increasing the pressure to $\mu_B \approx 10^{35}$ Pa, the internal repulsion within the degenerate gas succumbs to the external pressure, leading to the transformation of the hadronic gas into a low-temperature state consisting of quarks and gluons usually referred as *colour superconducting state* [8]. Instead, at temperatures $T \approx 100$ MeV, the ordinary nuclear matter converts to a hadron gas. At higher temperatures the hadron gas undergoes a transition to a colour deconfined state called *Quark-Gluon Plasma* (QGP). This state is characterised by partonic degrees of freedom, i.e. quarks and gluons, which are no longer confined within hadrons. According to the Cosmological Standard Model, a hot QGP state is believed to have existed in the early stages of the primordial Universe, shortly after the Big Bang, lasting for approximately 10^{-6} seconds [9, 10]. Indeed, in the regime of low baryon chemical potential ($\mu_B \approx 0$ GeV), the transition from confined to

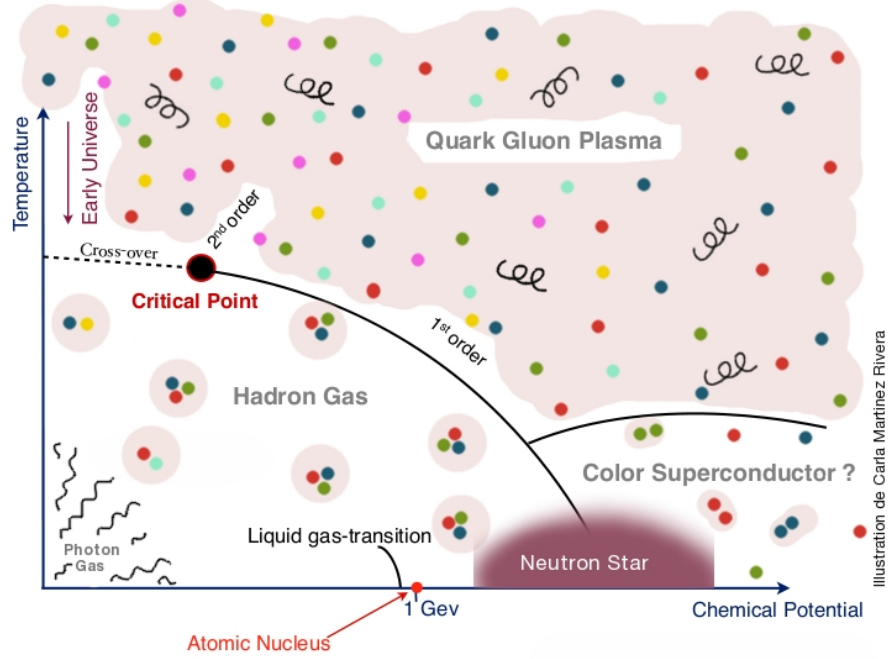


Fig. 1.3 Layout of the QCD phase diagram. Figure from Ref. [13].

deconfined nuclear matter is expected to occur in the form of a cross-over, at temperatures around $154 - 158$ MeV (approximately $1.8 - 2.0 \times 10^{12}$ K) corresponding to an energy density of $\epsilon \approx 0.42$ GeV/fm³ [11, 12]. As the baryon chemical potential increases to higher values, a first-order phase transition into QGP is expected to occur, while the coexistence with the cross-over at low μ_B implies the existence of a critical point. According to lQCD calculations, the properties of the phase transition depend on the number of quark flavours and their masses. In particular, it is possible to link the deconfinement process to the value of the quark masses and, in the limit of vanishing quark masses, the transition is expected to be a second-order phase transition. On the other hand, for finite quark masses, the transition is expected to be a cross-over. The transition from the hadronic phase to the QGP phase is characterised by the restoration of the *chiral symmetry*, which is spontaneously broken in the low-temperature regime [12]. The chiral symmetry is a fundamental symmetry of the QCD Lagrangian.

1.3 Heavy-Ion collisions

In the context of heavy-ion collisions, such as lead (Pb) nuclei, the large number of nucleon-nucleon interactions in the collision results in a significant deposit of energy in the interaction

region, referred to as the *fireball*. The high energy densities and temperatures reached in the initial stages of the collision often exceed the thresholds required for the formation of a QGP phase making this kind of experiment crucial for the study of the QGP phase diagram.

As mentioned in Section 1.2, ultrarelativistic heavy-ion collisions create the necessary thermodynamic conditions for the formation of QGP. Originally conceived in the second half of the last century, heavy-ion experiments have been performed at several facilities around the world, such as the *Super Proton Synchrotron* (SPS) at European Organisation for Nuclear Research (CERN), the *Alternating Gradient Synchrotron* (AGS) at the *Brookhaven National Laboratory* (BNL), the *Relativistic Heavy Ion Collider* (RHIC) at BNL, and the *Large Hadron Collider* (LHC) at CERN. The first heavy-ion experiments were conducted at the SPS and the AGS. These experiments involved fixed-target setups, operating at a center-of-mass energy per nucleon-pair ($\sqrt{s_{NN}}$) ranging from 7.7 to 17.3 GeV. In this energy regime, the colliding nucleons come to a stop within the collision region, resulting in a high μ_B at midrapidity ranging between 200 and 500 MeV. This particular condition is commonly referred to as the *stopping regime*. The stopping regime is of significant interest to facilities participating in the *Beam Energy Scan* (BES) [14] program at RHIC and the *Compressed Baryonic Matter* (CBM) [15] experiment at the *Facility for Antiproton and Ion Research* (FAIR). These experiments aim to explore the QCD phase diagram in the region of high baryon density, with a specific focus on searching for clear signatures of the QCD critical point.

To investigate the region of $\mu_B \approx 0$ MeV, it is necessary to increase the collision energy. This is accomplished in collider experiments by colliding beams of heavy ions, notably lead (Pb) and gold (Au). Experiments at colliders are performed at RHIC, with top energy of $\sqrt{s_{NN}} = 200$ GeV in Au–Au collisions, and at the LHC, where the highest centre-of-mass energy ever reached in heavy-ion collisions, $\sqrt{s_{NN}} = 5.4$ TeV, was achieved in xenon-xenon collisions. In this scenario, the colliding nuclei do not stop within the collision region but instead pass through each other, resulting in a low net-baryon density at midrapidity. During the collision, the nucleons lose a fraction of their kinetic energy, which corresponds to a rapidity (y^1) loss of about two units [16]. This loss is however small compared to the rapidity of the beams. As a consequence, the baryon density in the central region of the collision is

¹The rapidity y of a particle with energy E and longitudinal momentum p_z is defined as:

$$y := \frac{1}{2} \ln \left(\frac{E + p_z}{E - p_z} \right). \quad (1.5)$$

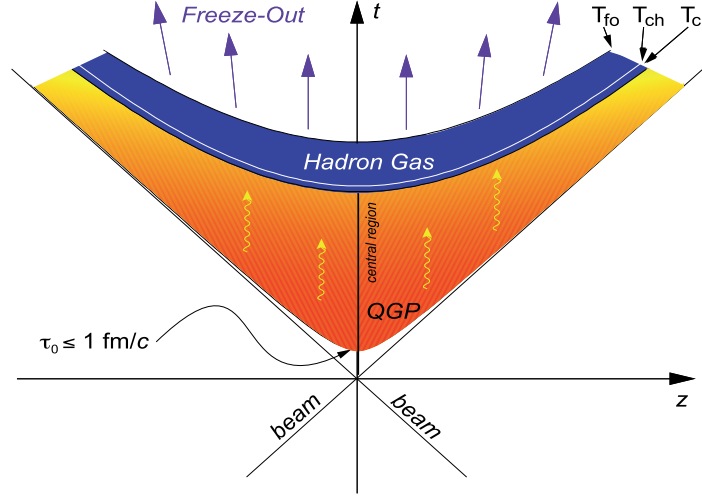


Fig. 1.4 Space-time diagram of longitudinal QGP evolution. Figure from Ref. [18].

depleted. This particular regime is commonly referred to as *Bjorken regime* or *transparency regime* [17].

1.3.1 Collision evolution

In the collision between two nuclei a complex system of particles is created, which undergoes a space-time evolution. J. D. Bjorken proposed a simplistic model to describe the space-time evolution of the system in the context of the Landau hydrodynamic model [13, 17]. Under the assumption of the existence of a plateau in the rapidity distribution of the produced particles at midrapidity and to reach the condition of transparency regime, the evolution can be described as follows:

- I. **Collision:** the Lorentz-contracted nuclei (*pancake*) collide for a very short time computed as:

$$\tau_{\text{cross}} = \frac{2R_N}{\gamma},$$

where R_N is the radius of the colliding nucleus ($R_N \sim 1.124 \times A^{1/3} \simeq 7.1$ fm for Pb nuclei) and γ is the Lorentz factor ($\gamma \simeq 2.5 \times 10^3$ at $\sqrt{s_{\text{NN}}} = 5.02$ TeV). At LHC energies, the crossing time is around 5.7×10^{-3} fm/c. This time lap is much smaller than the timescale of the strong interaction $\tau_{\text{strong}} = 1/\Lambda_{\text{QCD}} \simeq 1/200$ MeV $\simeq 1$ fm/c. Hence, particles generated by the strong interaction between the partons of colliding nucleons, are created once the nuclei have already crossed each other.

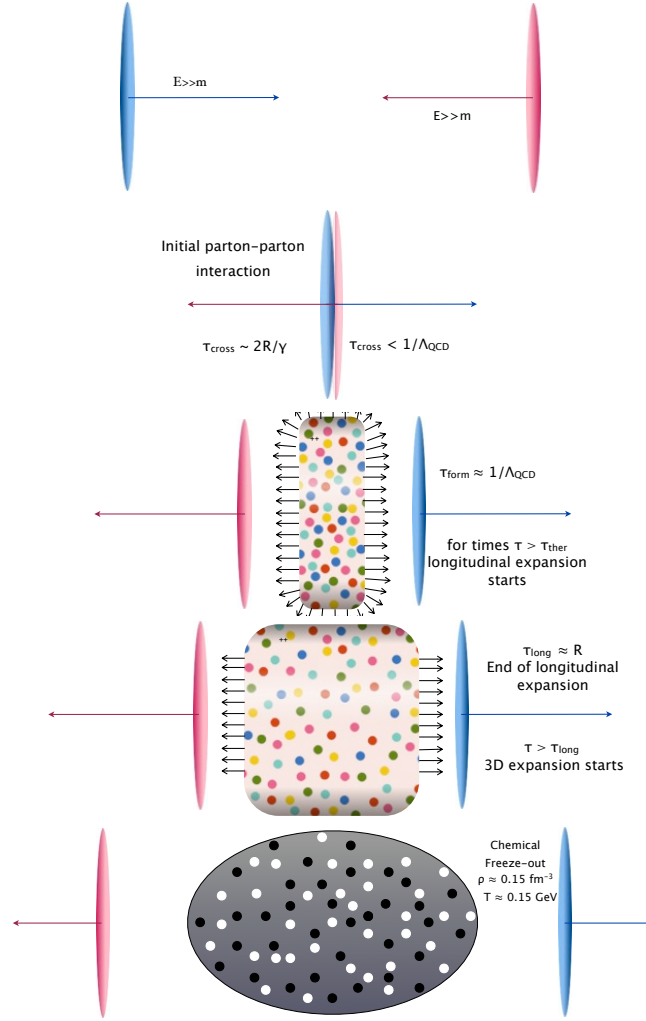


Fig. 1.5 Space-time evolution of a heavy ion collision. Figure from Ref. [13].

II. **Pre-equilibrium:** the produced particles at the moment of their formation ($\sim 1 \text{ fm}/c$) at a proper time τ_f are not in thermal equilibrium in its first phases and the system is characterised by a high energy density and a low baryon density. In the mid-rapidity region ($|y| < 0.5$), the net baryon density is close to zero for collisions at RHIC and LHC energies, satisfying the condition of the Bjorken regime. The system energy density can be estimated as:

$$\varepsilon = \frac{1}{\pi R_N^2 \tau_f} \frac{dE_T}{dy} \Big|_{y=0}, \quad (1.6)$$

where τ_f is the formation time, πR_N^2 is the transverse area, and E_T represents the transverse energy of the particles produced in the collision². Assuming $\tau_f = 1$ fm/c and considering the highest energies achieved at RHIC and LHC, the estimated energy density is $\varepsilon \sim 5 - 12$ GeV/fm³ [19]. This value is significantly higher than the pseudo-critical energy density discussed in Sec. 1.2.

- III. **QGP (Quark-Gluon Plasma)**: the system of interacting particles reaches the thermal equilibrium at a time τ_{eq} , since the energy density is higher than the pseudo-critical energy density. The system ... QGP. The QGP evolves according to the laws of relativistic hydrodynamics [20] expanding due to the pressure gradients and cooling down. According to the hydrodynamical models, the thermal equilibrium is reached at a proper time $0.6 < \tau_{eq} < 1$ fm/c. At the LHC energies, the QGP phase is expected to last ~ 10 fm/c.
- IV. **Hadron gas**: while the system expands, the temperature gradually decreases until it drops below the pseudo-critical value. At this point, the colour charges are no longer free and the QGP undergoes a transition to a hadronic system, where partons are confined into hadrons. These newly formed hadrons contribute to an expanding hadron gas. In situations where a first-order transition takes place, the system enters a mixed phase where confined hadronic matter coexists with the QGP.
- V. **Freeze-out**: the hadron gas continues to expand and cool down. At a given moment, inelastic scatterings cease due to the low particle density, and the relative abundance of different hadronic species is fixed. This moment is referred to as *chemical freeze-out*. The chemical freeze-out temperature is estimated to be $T_{ch} \sim 157$ MeV [21]. When the particle density decreases to the point where even elastic collisions cease, the momentum spectra of the hadrons are also fixed. This moment is referred to as *kinetic freeze-out* (or *thermal freeze-out*). The thermal freeze-out temperature is estimated to be $T_{kin} \sim 100$ MeV [21]. The hadrons then continue to expand freely until they reach the detector.

Figure 1.4 and Figure 1.5 show a schematic representation of the collision evolution in the longitudinal direction based on the Bjorken model.

²The transverse energy of a particle of mass m and transverse momentum p_T is defined as $E_T := \sqrt{m^2 + p_T^2}$.

1.3.2 The geometry of the collision: Glauber Model

Originally developed in the late 1950s, the Glauber model [22, 23] is a semi-classical model used to describe the initial stages of heavy-ion collisions. In this model, nucleus-nucleus collisions are interpreted as a superposition of incoherent binary nucleon-nucleon collisions. The Glauber model operates under the so-called *optical limit* assuming nuclei are constituted of point-like and non-interacting nucleons. In its first formulation, no distinction was made between the different nucleon species inside the nucleus, meaning that protons and neutrons are indistinguishable. In the optical limit, the nuclei are also assumed to have sufficiently large momenta to prevent significant deflections after the collision and force the nucleons inside them to travel in straight lines. The Glauber model relies on two crucial physical inputs that cannot be derived from theory:

- i. The nuclear matter density distribution inside the nuclei ($\rho(r)$). The Wood-Saxon distribution is a commonly used parametrisation that captures the essential features of this distribution as:

$$\rho(r) = \frac{\rho_0}{1 + \exp\left(\frac{r-R}{d}\right)},$$

where ρ_0 is the nuclear matter density at the center of the nucleus, R is the nuclear radius, and d is the diffuseness parameter.

- ii. The nucleon-nucleon differential cross section ($\sigma_{\text{inel}}^{AB}$). This quantity represents the probability of inelastic scattering between two nucleons and depends on the energy scale of the collision. Since the cross section involves processes with low momentum transfer, it is not feasible to calculate it using pQCD. Therefore, the measured inelastic nucleon-nucleon cross section is employed as an input, serving as the sole source of non-trivial beam-energy dependence for the predictions of the Glauber model.

The Glauber model gives access to several quantities related to the collision process. One such quantity is the number of *participant nucleons*, denoted as N_{part} . It represents the number of nucleons that have participated in the collision by undergoing at least one interaction with a nucleon of the other nucleus. Another important quantity is the total number of *binary nucleon-nucleon collisions*, denoted as $N_{\text{coll}}(\vec{b})$, where $\vec{b}$ is the *impact parameter*. The impact parameter is defined as the perpendicular distance between the centres of the colliding nuclei in the transverse plane of the collision. It determines the *centrality* of the collision and influences the number of binary collisions that occur. Nucleus-nucleus collisions can be classified as either *central* or *peripheral*. Central collisions have a very low impact parameter ($|\vec{b}| \sim 0$), resulting in a nearly complete superposition of the nuclei in the

transverse plane, whether peripheral collisions have a high impact parameter ($|\vec{b}| \sim R_A + R_B$), resulting in a small overlap region between the nuclei. The centrality of a collision is usually expressed as a percentage of the total inelastic cross section $\sigma_{\text{inel}}^{AB}$, where the most central collisions have the highest percentage. More details on centrality determination are given in Sec. 2.4.

Figure 1.6 schematically reports a nucleus-nucleus collision from both the side and frontal views. The coordinate system is centered at the A nucleus. The relative position of nucleons in the two nuclei is described by the impact parameter $\vec{b}$. The positions of nucleons in the A and B nuclei are represented by the vectors $\vec{s}$ and $\vec{s} - \vec{b}$, respectively. According to this coordinate system, the probability of finding a nucleon in the A nucleus at position $\vec{s}$ is given by the *nuclear thickness function* $T_A(\vec{s})$ defined as:

$$T_A(\vec{s}) = \int \rho_A(\vec{s}, z) dz_A, \quad (1.7)$$

where $\rho_A(\vec{s}, z)$ ³ is the nuclear matter density distribution inside the nucleus A and z_A is the coordinate along the beam axis. Analogously, the probability to find a nucleon in the B nucleus at position $\vec{s} - \vec{b}$ is given by $T_B(\vec{s} - \vec{b})$. Hence, the probability to find a nucleon in the A nucleus at position $\vec{s}$ and a nucleon in the B nucleus at position $\vec{s} - \vec{b}$ is given by the *nuclear overlap function* $T_{AB}(\vec{b})$:

$$T_{AB}(\vec{b}) = \int T_A(\vec{s}) T_B(\vec{s} - \vec{b}) d^2s. \quad (1.8)$$

The nuclear overlap function has the dimension of an area and is proportional to the probability of a nucleon-nucleon collision given by:

$$P_{\text{coll}}(\vec{b}) = \sigma_{\text{inel}}^{AB} T_{AB}(\vec{b}). \quad (1.9)$$

Given the probability of a nucleon-nucleon collision, the probability of having n nucleon-nucleon collisions is given by the binomial distribution:

$$P_{n,\text{coll}}(\vec{b}) = \binom{AB}{n} \left[P_{\text{coll}}(\vec{b}) \right]^n \left[1 - P_{\text{coll}}(\vec{b}) \right]^{AB-n}, \quad (1.10)$$

³The nuclear matter density distribution inside the nucleus A can be redefined as $\rho_A(\vec{s}, z) \equiv \rho(\vec{s}, z)/A$. This is done in order to have a normalised density distribution. The same is done for the nucleus B.

where AB is the total number of nucleon pairs in the two nuclei and $1 - P_{\text{coll}}(\vec{b})$ is the probability of no nucleon-nucleon collision. The average number of nucleon-nucleon collisions is naturally given by the first moment of the binomial distribution:

$$\langle N_{\text{coll}}(\vec{b}) \rangle = nP_{n,\text{coll}}(\vec{b}) = ABP_{\text{coll}}(\vec{b}) = \sigma_{\text{inel}}^{AB} T_{AB}(\vec{b}). \quad (1.11)$$

Analogously, the number of participant nucleons N_{part} is defined as the number of nucleons in the two nuclei that have interacted:

$$\begin{aligned} N_{\text{part}}(\vec{b}) &= A \int T_A(\vec{s}) P_{\text{inter}}^A(\vec{b}) d^2s \\ &+ B \int T_B(\vec{s} - \vec{b}) P_{\text{inter}}^B(\vec{b}) d^2s. \end{aligned} \quad (1.12)$$

Starting from Eq. 1.10, it is possible to derive the probability of having at least one nucleon-nucleon collision, which implies that the two nuclei have interacted, as⁴:

$$P_{\text{inter}}^{AB}(\vec{b}) = 1 - P_{0,\text{coll}}(\vec{b}) = 1 - \left[1 - P_{\text{coll}}(\vec{b}) \right]^{AB}. \quad (1.13)$$

In case of *hard scattering processes*, i.e. processes characterised by a large momentum transfer, Eq. 1.13 can be rewritten by expanding the binomial term according to the Taylor series as:

$$\begin{aligned} P_{\text{hard}}^{AB}(\vec{b}) &= 1 - \left[1 - \sigma_{\text{hard}}^{AB} T_{AB}(\vec{b}) \right]^{AB} \\ &= 1 - \left[1 - AB \sigma_{\text{hard}}^{AB} T_{AB}(\vec{b}) \right] \\ &\approx \sigma_{\text{hard}}^{AB} \langle N_{\text{coll}}(\vec{b}) \rangle. \end{aligned} \quad (1.14)$$

The probability of having at least one hard scattering process is proportional to the average number of nucleon-nucleon collisions. This proportional law is referred to as *binary scaling*. According to the binary scaling, the particle yield in a hard scattering process in a nucleus-nucleus (A–A) collision is expected to be proportional to the one measured in a proton-proton (pp) collision scaled by the average number of nucleon-nucleon collisions.

Despite the ability of the Glauber model to successfully predict many quantities related to the geometry of the collision, it fails in describing quantities that are discrete in Nature.

⁴This probability is given by the complement of the probability of no nucleon-nucleon collision.

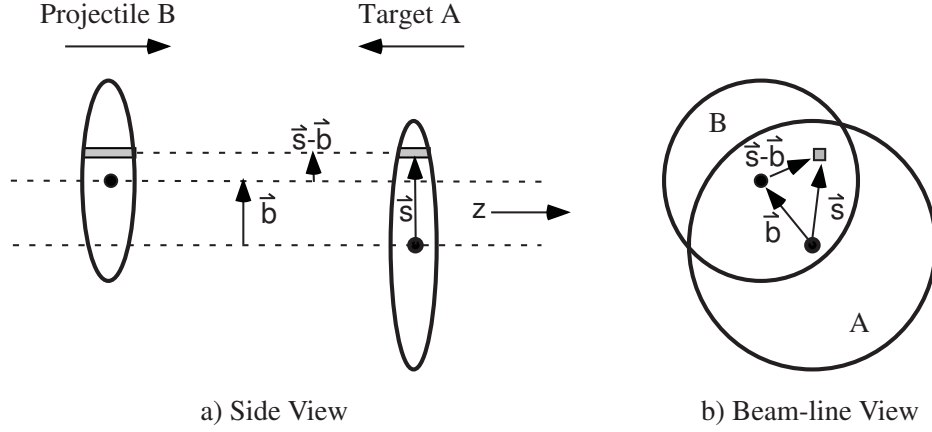


Fig. 1.6 Schematic representation of two colliding nuclei in the Glauber model description. The coordinate system for the nucleons is reported. Figure from Ref. [22].

This limitation rises from the fact that continuous distributions are employed in the model, neglecting possible event-by-event fluctuations. One example is the position of the nucleons inside the colliding nuclei. Event-by-event fluctuations are addressed in Glauber Monte Carlo simulations [22]. The position of the nucleons inside the colliding nuclei is sampled according to the nuclear density distribution and a random impact parameter is picked from the distribution $\frac{d\sigma}{db} \propto 2\pi b$. A nucleon-nucleon collision is considered to have occurred if the distance in the orthogonal plane between the two nucleons d satisfies the following condition:

$$d < \sqrt{\sigma_{\text{inel}}^{AB}/\pi}. \quad (1.15)$$

The event-by-event fluctuations of N_{part} and N_{coll} are obtained by averaging over a multitude of Glauber Monte Carlo simulations. Figure 1.7 shows the energy density profile of the collision obtained from a Glauber Monte Carlo simulation.

1.4 Properties of the QGP

A selection of key observables and experimental findings that have contributed to our current understanding of the properties of the QGP medium generated in ultrarelativistic heavy-ion collisions is presented in this Section.

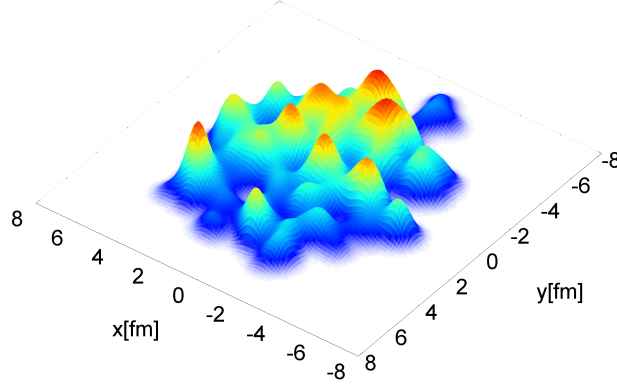


Fig. 1.7 Initial energy density in the transverse plane in Glauber Monte Carlo simulation. Figure from Ref. [24].

1.4.1 Charged particle multiplicity

The *charged particle multiplicity* N_{ch} is defined as the number of charged particles produced in a collision. The charged particle multiplicity is commonly expressed per unit of pseudorapidity η ⁵ and is denoted as $dN_{\text{ch}}/d\eta$. The charged particle multiplicity at midrapidity serves as a global observable sensitive to the energy density during the pre-equilibrium phase of the collision and offers insights into the transition to the QGP phase. To properly compare the results obtained in different collision systems by taking into account the size of the colliding systems, the charged particle multiplicity is typically normalised by the number of participant pairs $\langle N_{\text{part}} \rangle / 2$ calculated with the MC-Glauber (refer to Sec. 1.3.2).

The left panel of Fig. 1.8 illustrates the charged particle multiplicity at midrapidity as a function of the center-of-mass energy per nucleon pair $\sqrt{s_{\text{NN}}}$, as measured by ALICE in various colliding systems [25]. The charged particle density exhibits a smooth increase with $\sqrt{s_{\text{NN}}}$ for both proton-proton (pp) and nucleus-nucleus (A–A) collisions. The dependence of

⁵The pseudorapidity, denoted as η , is defined as:

$$\eta := -\ln \left[\tan \left(\frac{\theta}{2} \right) \right], \quad (1.16)$$

where θ is the polar angle of the particle relative to the beam axis (refer to Section 2.2). The pseudorapidity is approximately equal to the rapidity in the limit of $m \ll p$, where m is the particle mass and p is the particle momentum:

$$y = \frac{1}{2} \ln \left(\frac{E + p_z}{E - p_z} \right) \rightarrow \frac{1}{2} \ln \left(\frac{1 + \cos \theta}{1 - \cos \theta} \right) = -\ln \left[\tan \left(\frac{\theta}{2} \right) \right] = \eta.$$

Unlike the rapidity y , η is not Lorentz invariant under longitudinal boosts and does not require a priori knowledge of the particle species nor energy.

$\frac{2}{\langle N_{\text{part}} \rangle} \langle \frac{dN_{\text{ch}}}{d\eta} \rangle$ on $\sqrt{s_{\text{NN}}}$ follows a power-law behaviour of the form $a \cdot s^b$, with $b = 0.155 \pm 0.004$ and $b = 0.103 \pm 0.002$ for A–A and pp collisions, respectively. This indicates a steeper rise in the charged particle multiplicity as a function of $\sqrt{s_{\text{NN}}}$ in nucleus-nucleus collisions compared to proton-proton collisions. Additionally, the results obtained in p–Pb and d–Au collisions align with the curve for pp collisions, suggesting that the significant increase observed in A–A collisions can not be exclusively attributed to the multiple collisions experienced by the participants. The dependence of $\frac{2}{\langle N_{\text{part}} \rangle} \langle \frac{dN_{\text{ch}}}{d\eta} \rangle$ on $\langle N_{\text{part}} \rangle$ measured by ALICE in pp, p–Pb, and Pb–Pb collisions is shown in the right panel of Fig. 1.8. According to the *wounded nucleon* model [26], the production of particles in processes characterised by a low momentum transfer (*soft processes*), which are more frequent than the hard ones, depends on the number of participant nucleons rather than the number of binary collisions ($N_{\text{ch}} \sim N_{\text{part}}$). Although one would expect a flat trend of $\frac{2}{\langle N_{\text{part}} \rangle} \langle \frac{dN_{\text{ch}}}{d\eta} \rangle$ as a function of $\langle N_{\text{part}} \rangle$, the charged particle multiplicity unexpectedly exhibits a strong decreasing trend with $\langle N_{\text{part}} \rangle$ moving from central collisions, i.e. high $\langle N_{\text{part}} \rangle$, to peripheral collisions, i.e. low $\langle N_{\text{part}} \rangle$, with a slope of 1.8. Note that, in the case of soft probes, the Wounded Nucleon model predicts their production to depend on N_{part} and, hence, no dependence on the charged particle multiplicity is expected. This trend is consistent over different colliding systems and center-of-mass energies. In the *Colour Glass Condensate* (CGC) model [27] the production of soft particles is connected to the presence of low-momentum gluons in the initial stage of the collision. In this framework, the gluon density within the colliding nuclei is described by the Balitsky-Fadin-Kuraev-Lipatov (BFKL) equation [28] and is expected to increase significantly at low momentum fractions x until it reaches a certain *saturation scale* Q_s^2 . The saturation scale Q_s^2 limits the maximum phase space occupancy of the gluons and, once the saturation scale is reached, the system is treated as a color condensate. Limiting the number of initial low-momentum gluons effectively restricts the number of the produced partons and, consequently, the number of particles produced in the collisions. Predictions based on CGC models successfully describe the experimental measurements [29–31].

1.4.2 Identified hadron abundances: the chemical freeze-out temperature

As explained in Section 1.3.1, the evolution of the heavy-ion collision when the transition temperature is reached, the system enters the hadronisation phase. The microscopic description of the non-perturbative process in which countless quarks and gluons combine to form thousands of final state hadrons is extremely intricate. The relative abundances of the different hadron species produced in hadronic collisions are found to be well described by the

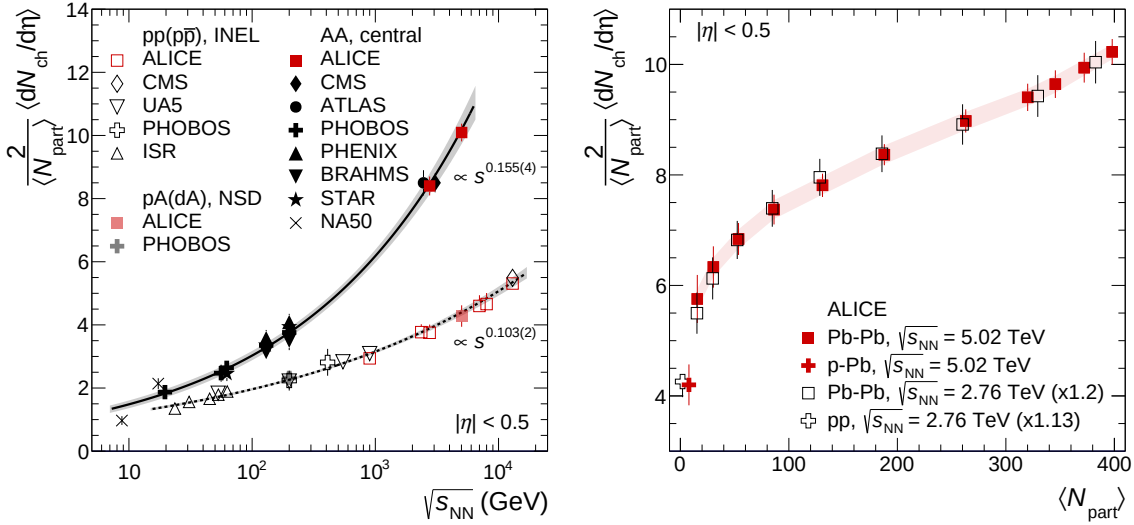


Fig. 1.8 Charged particle multiplicity at midrapidity as a function of the center-of-mass energy per nucleon pair $\sqrt{s_{\text{NN}}}$ (left) and the number of participant pairs $\langle N_{\text{part}} \rangle$ (right) for various colliding systems. Figures from Ref. [25].

Statistical Hadronisation Model (SHM), which is based on the assumption that the system is in thermal and chemical equilibrium at the time of chemical freeze-out. In particular, the resulting hadrons are treated as a hadron resonance gas, i.e. a gas of non-interacting hadrons and resonances, described by a grand-canonical ensemble. In the grand-canonical ensemble description, the energy and the quantum numbers (baryon number, isospin, and strangeness) of the system are conserved globally⁶.

The SHM can calculate the relative abundance N_i of the different hadron species i in terms of the temperature T , the volume V , and the chemical potential μ_i as:

$$N_i(T, V, \mu_i) = \frac{g_i V}{2\pi^2} \int_0^\infty \frac{p^2 dp}{e^{(E - \mu_i)/T} \pm 1}. \quad (1.17)$$

The degrees of freedom of the i -th species are indicated with g_i . The chemical potential $\mu_i = \mu_B B + \mu_S S + \mu_{I_3} I_3$ is defined for each particle species as the sum of the quantum numbers multiplied by their respective chemical potential. The $\pm$ sign is needed to discriminate between the Bose-Einstein and the Fermi-Dirac statistics for bosons and fermions, respectively. Derivations from the grand-canonical description can be applied by introducing additional parameters to account for possible incomplete thermalisation for the strange (γ_s)

⁶In the case of proton-proton and proton-nucleus collisions, the quantum numbers and the energy are conserved locally according to the canonical and micro-canonical ensembles, respectively.

quarks. In particular, in the case of pp collisions, the γ_s parameter is found to be significantly smaller than unity.

The yields of different identified hadron species measured by the ALICE Collaboration in the 10% most central Pb–Pb collisions at $\sqrt{s_{\text{NN}}} = 5.02$ TeV are presented in Fig. 1.9. These data are compared to predictions from various implementations of the SHM, including THERMUS [32], GSI-Heidelberg [33], and SHARE [34]. These models make different assumptions regarding the equilibrium condition and/or conservation laws at the chemical freeze-out stage. By performing a simultaneous fit to the measured yields, these models extract the values of the chemical freeze-out temperature T_{ch} , the volume V , and the baryochemical potential μ_B , which are treated as free parameters. Remarkably, these models successfully describe the measured particle abundances over a wide range of nine orders of magnitude, with a resulting reduced χ^2 ranging between 3.8 and 5.3. However, significant discrepancies are found in the case of excited K mesons, which decouple from the system at different times and are therefore excluded from the fits, as well as for protons. Note that the proton asymmetry has been studied in detail and different interpretations have been proposed [35–37]. All the models consistently determine a chemical freeze-out temperature of approximately $T_{\text{ch}} = 153$ MeV. This value closely aligns with the anticipated critical temperature of the phase transition from the hadronic to the deconfined state, indicating that the chemical freeze-out likely occurs shortly thereafter.

1.4.3 Strangeness enhancement

The NA35 [38], WA97 [39], and NA57 [40] Collaborations in the 1990s provided the first evidence of enhanced production of hadrons containing strange quarks in A–A collisions compared to those observed in pp collisions. This phenomenon was called *strangeness enhancement*. The degree of enhancement was found to correlate with the number of strange quarks contained in the measured hadrons. These findings aligned with the prediction made by Rafelski and Müller, who proposed the strangeness enhancement as a characteristic signal of deconfinement [41]. In colliding nuclei, there are no valence strange quarks present, thus any observed strange quark must be produced either during the collision or within the QGP itself. At temperatures below the critical temperature of the phase transition, the production of strange quarks is thermally suppressed because the effective mass of $s\bar{s}$ pairs is around $1 \text{ GeV}/c^2$ and exceeds the temperature. However, when the system reaches the critical temperature and the system is in a QGP, the mass of the strange quark decreases approximately to $100 \text{ MeV}/c^2$ due to the restoration of chiral symmetry (refer to Sec. 1.2). Consequently, an abundant thermal production of strange-antistrange quark pairs, which

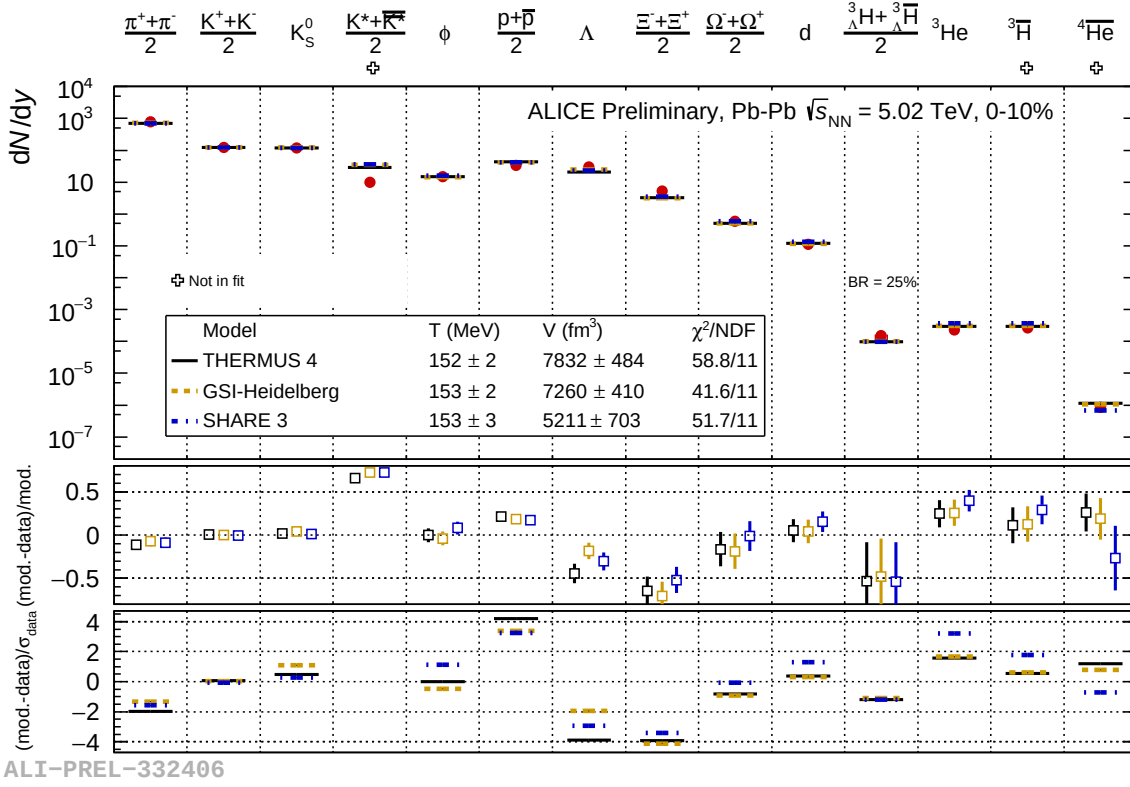


Fig. 1.9 Simultaneous thermal fit to the identified-particle yields measured by the ALICE Collaboration in the 10% most central Pb–Pb collisions at $\sqrt{s_{\text{NN}}} = 5.02$ TeV. The results obtained with the THERMUS [32], GSI [33], and SHARE [34] models are reported. Figure from ALICE Figure repository ©.

have a mass of the order of the temperature of the medium, is expected through gluon-fusion processes ($g\bar{g} \rightarrow s\bar{s}$) within the QGP, facilitated by the high density of gluons. This leads to an increase of the relative abundance of strange particles.

An alternative explanation for the strangeness enhancement within SHM is known as the lifting of the *canonical suppression* [42]. According to this explanation, the strangeness enhancement is not a consequence of the transition to the deconfined phase, but rather a result of the exact local conservation of the quantum numbers in small collision systems as required in the canonical ensemble. In the canonical ensemble, the strangeness conservation is enforced, leading to a suppression of strange quark production as compared to the grand canonical ensemble. This effect is expected to be significant in small systems, like pp and pA collisions, where the number of strange quarks is small. In contrast, in larger systems such as A–A collisions, the grand canonical ensemble provides a good description of the medium and the constraints of strangeness conservation are less severe, resulting in a more abundant strangeness production.

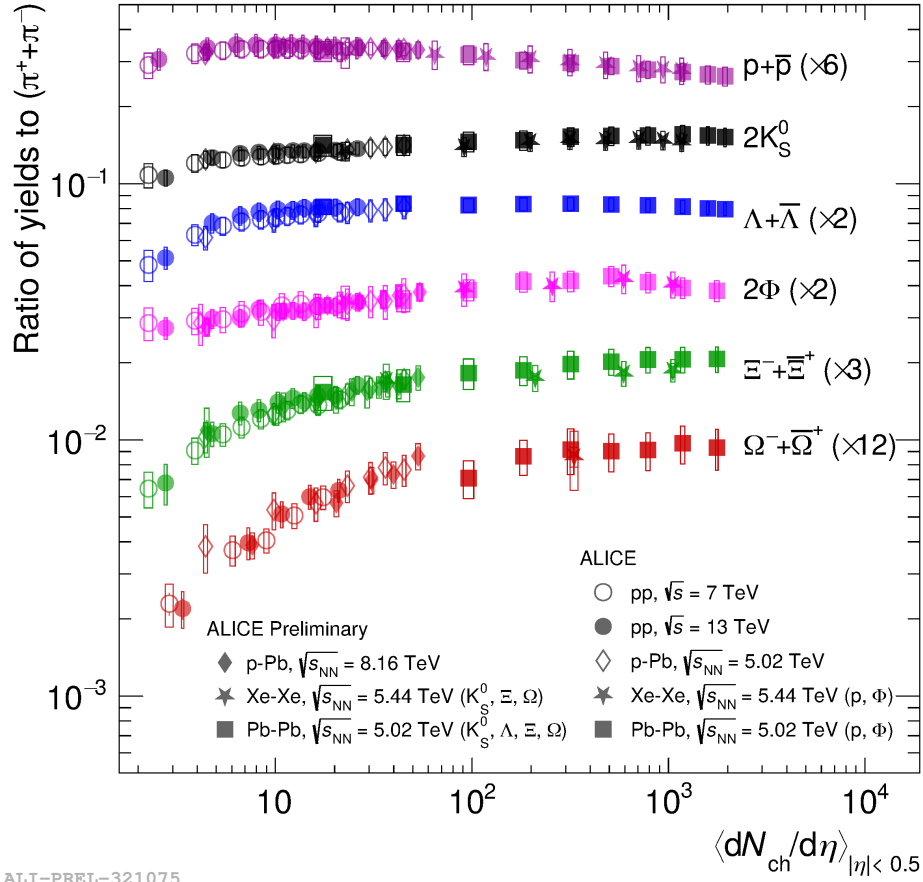


Fig. 1.10 Sum of particle-antiparticle production yields relative to that of $(\pi^+ + \pi^-)$ for non-strange ($p + \bar{p}$), strange ($2K_S^0$, $\Lambda + \bar{\Lambda}$), and multi-strange (2ϕ , $\Xi^- + \bar{\Xi}^+$, and $\Omega^- + \bar{\Omega}^+$) hadrons as a function of the average charged particle multiplicity at midrapidity, measured by the ALICE Collaboration across various colliding system sizes and energy scales. Figure from the ALICE Figure repository ©.

An enhanced production of strange hadrons was also observed in heavy-ion collisions at RHIC [43] and LHC [44]. Figure 1.10 illustrates the sum of particle-antiparticle production yields relative to that of $(\pi^+ + \pi^-)$ for non-strange ($p + \bar{p}$), strange ($2K_S^0$, $\Lambda + \bar{\Lambda}$), and multi-strange (2ϕ , $\Xi^- + \bar{\Xi}^+$, and $\Omega^- + \bar{\Omega}^+$) hadrons as a function of the average charged particle multiplicity at midrapidity, measured by the ALICE Collaboration across various colliding system sizes and energy scales [45]. The measurements show a smooth increase moving from small charged particle multiplicity, pp collisions, to high charged-particle multiplicity, as those achieved in central Pb–Pb collisions. In particular, the steepness of the increase is found to be proportional to the strangeness content of the hadron. No significant dependence on the collision energy is observed, indicating the strangeness enhancement to be mainly a consequence of the final charged-particle multiplicity.

1.4.4 Thermal freeze-out temperature and radial flow

At the chemical freeze-out, the particle abundances are fixed. However, interactions between particles persist through elastic scatterings. Only at the thermal freeze-out these interactions finally cease, shaping the momentum spectra of primary particles. Assuming the fireball to be a static thermal source at the equilibrium, the low- p_T interval of the transverse-momentum spectra of the produced particles can be elegantly described by a Maxwell-Boltzmann distribution⁷. The slope parameter T_{fo} assumes the role of the thermal freeze-out temperature:

$$\frac{1}{p_T} \frac{dN}{dp_T} = \frac{1}{m_T} \frac{dN}{dm_T} \propto \exp\left(-\frac{m_T}{T_{fo}}\right), \quad (1.18)$$

where m_T is the transverse mass⁸. Accordingly, the transverse-mass spectra of different particle species conform to the exponential law described by Eq. 1.18, sharing the same slope parameter T_{fo} . This phenomenon is commonly known as m_T scaling [46] and is observed to hold remarkably well at low p_T in small colliding systems [47]. However, in A–A collisions, deviations from the m_T scaling become apparent. These deviations arise from the presence of pressure gradients within the fireball, resulting in the collective expansion of the system. As a consequence, the particles produced in the collision not only undergo random thermal motion, akin to pp collisions, but also participate in the collective flow propelled by the expanding fireball in the transverse plane. This collective motion is commonly referred to as *radial flow*. The slope parameter T_{fo} is then reinterpreted to incorporate the effects of radial flow as:

$$T_{fo} = T_{kin} + \frac{m}{2} \langle \beta_T \rangle^2. \quad (1.19)$$

The second term in the Eq. 1.19 represents the kinetic energy associated with the radial flow, with $\langle \beta_T \rangle$ denoting the average transverse velocity of the particles. As a result, the average

⁷The Maxwell-Boltzmann description is derived by the following Taylor expansion:

$$E \frac{dN}{d^3p} = \frac{g_i V E}{(2\pi)^3} \frac{1}{\exp\left(\frac{E - \mu_i}{T}\right) \pm 1} \approx \frac{g_i V E}{(2\pi)^3} \exp\left(-\frac{E - \mu_i}{T}\right),$$

where g_i is the degeneracy factor, E is the particle energy, μ_i is the chemical potential, and T is the temperature. The approximation is valid for $E \gg T$.

⁸The transverse mass m_T is defined as $m_T = \sqrt{p_T^2 + m^2}$, where m is the particle mass and p_T is the transverse momentum of the particle. Hence, the following relation holds:

$$\frac{dm_T}{dp_T} = \frac{p_T}{\sqrt{p_T^2 + m^2}} = \frac{p_T}{m_T} \rightarrow m_T dm_T = p_T dp_T$$

p_T of the particles increases with their mass. Additionally, the $\langle\beta_T\rangle$ parameter is not uniform across the transverse plane. Moving from the central region towards the periphery of the fireball, the pressure gradients progressively intensify, resulting in a stronger radial flow. Consequently, the m_T spectra of particles emitted at the outer boundaries of the fireball exhibit a steeper slope compared to those produced in the central region of the interaction zone. Specifically, the dependence of $\langle\beta_T\rangle$ on the transverse radius r of the fireball is parametrised as follows:

$$\beta_T(r) = \beta_s \left(\frac{r}{R} \right)^n, \quad (1.20)$$

where R represents the maximum radius of the fireball, β_s denotes the maximum radial flow velocity at the surface of the fireball, and the parameter n corresponds to the exponent of the *flow velocity profile* and is commonly found to be around 1 – 2 in the literature [48]. The resulting p_T spectra of different particle species can be accurately described by the Boltzmann-Gibbs Blast-Wave function [48]. This function incorporates the effects of thermalisation and radial expansion of the fireball, enabling calculations of the p_T spectra based on parameters such as temperature, radial velocity, and flow profile. Specifically, particles are presumed to be produced in a fluid element within a static thermal source characterised by a temperature T_{kin} and subsequently undergo a radial boost with the velocity $\beta_T(r)$ of the fluid cell. According to the Blast-Wave model, the observed particle spectrum arises from the sum of spectra originating from individual thermal sources possessing distinct radial velocities. It is important to notice that the Blast-Wave is an approximation, the production spectrum is properly described by hydrodynamic calculations in the QGP phase and by including the hadronisation process.

The ALICE Collaboration measured the correlation between $\langle\beta_T\rangle$ and T_{kin} in Pb–Pb collisions at $\sqrt{s_{\text{NN}}} = 2.76$ [49] and $\sqrt{s_{\text{NN}}} = 5.02$ TeV [50]. The results are shown in Fig. 1.11 for different centrality classes. The values of $\langle\beta_T\rangle$ and T_{kin} are obtained from the fit to the p_T spectra of pions, kaons, and protons. The average transverse velocity $\langle\beta_T\rangle$ is found to increase with the collision centrality, reaching a maximum value of $\langle\beta_T\rangle = 0.663 \pm 0.003$ in the 5% most centrality class. On the other hand, the thermal freeze-out temperature T_{kin} is found to decrease with the collision centrality, reaching a minimum value of $T_{\text{kin}} = 0.090 \pm 0.003$ GeV in the 5% most central collisions. These outcomes indicate a more rapid expansion of the fireball as the collisions centrality increases, owing to the larger system size, and a delayed thermal freeze-out at lower temperatures in central collisions. Notably, no significant discrepancies were observed between different collision energies at LHC.

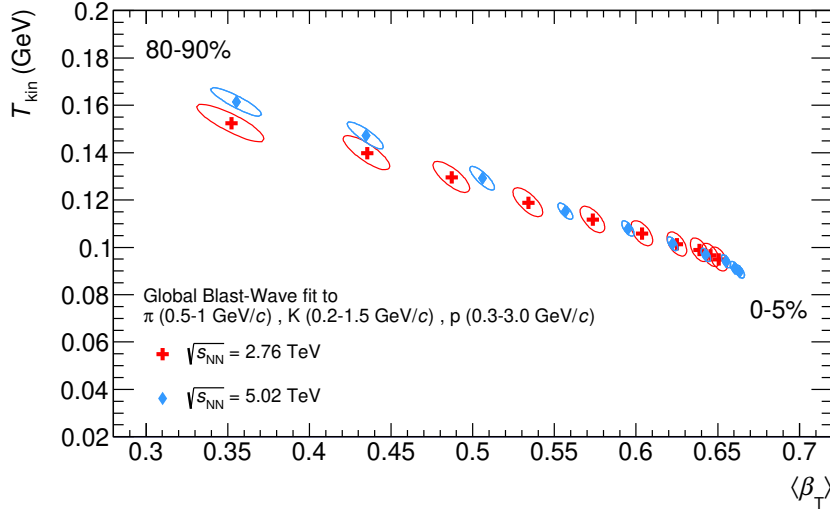


Fig. 1.11 Average expansion velocity and kinetic freeze-out temperature (T_{kin}) obtained from the simultaneous Boltzmann-Gibbs Blast-Wave fit to pions, kaons, and protons spectra measured in Pb–Pb collisions at $\sqrt{s_{\text{NN}}} = 5.02$ and 2.76 TeV. The correlated 1σ uncertainties, comprehensive of statistical and systematic uncertainty, are shown as ellipses. Figure from Ref. [50]

1.4.5 Azimuthal anisotropy: elliptic flow

As mentioned earlier in Section 1.3.2, the initial geometry of the collision is characterised by the impact parameter $\vec{b}$. Figure 1.12 shows a schematic representation of the collision geometry for a peripheral heavy-ion collision. The highlighted golden region represents the overlapping area of the incoming nuclei, which deviates from a spherical shape expected in central collisions and instead presents an almond-like shape. Consequently, due to this initial relative geometrical anisotropy in the plane defined by the impact parameter $\vec{b}$ and the beam axis, commonly known as the *reaction plane* Ψ_{RP} , the pressure gradients within the fireball exhibit non-uniformity in the transverse plane. The initial geometrical anisotropy gives rise to an azimuthal anisotropy of the momentum distribution of the produced particles. The presence of azimuthal anisotropy in the momentum distributions of the produced particles is a non-ambiguous signature of collective behaviour. Measurements of the azimuthal anisotropy of the final state particles can be used to constrain the initial conditions of the collision and the transport properties of the QGP, such as the shear viscosity to entropy density ratio η/s and the bulk viscosity to entropy density ratio ζ/s [51, 52]. The azimuthal anisotropy is studied by measuring the azimuthal distribution of the produced particles with respect to the

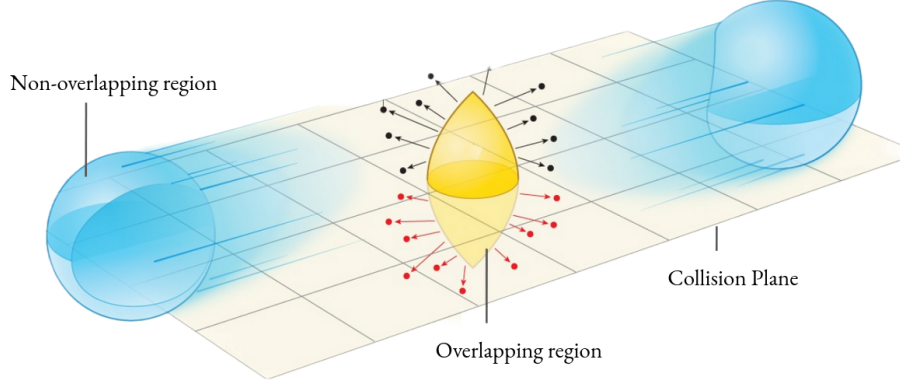


Fig. 1.12 Schematic view of a peripheral heavy-ion collision.

reaction plane and decomposing it into a Fourier series [53]:

$$E \frac{d^3N}{dp^3} = \frac{1}{2\pi} \frac{d^2N}{p_T dp_T dy} \left(1 + 2 \sum_{n=1}^{\infty} v_n \cos[n(\phi - \Psi_n)] \right), \quad (1.21)$$

where ϕ is the azimuthal angle of the particle, and Ψ_n is the symmetry plane of the n -th harmonic, which is an estimator of Ψ_{RP} . The v_n coefficient is defined as:

$$v_n = \langle \cos[n(\phi - \Psi_n)] \rangle, \quad (1.22)$$

and represents the n -th order Fourier coefficient⁹. The brackets in Eq. 1.22 denote an average over all particles and all tracks produced in the events. In peripheral collisions, the dominant term of the Fourier series is the second-order coefficient v_2 , also referred to as *elliptic flow*. A positive value of v_2 indicates that the produced particles are preferentially emitted in the reaction plane (in-plane) than in the orthogonal direction (out-of-plane). The higher-order coefficients, *triangular flow* (v_3) and *quadrangular flow* (v_4), are mainly driven by fluctuations of the distributions of participant nucleons in the nuclei in the initial state and by non-linear response of the medium, respectively [54–57].

The ALICE Collaboration measured the elliptic flow of identified charged particles in Pb–Pb collisions at $\sqrt{s_{NN}} = 5.02$ TeV in several centrality intervals [58]. The results are shown in Fig. 1.13. The v_2 is found to increase with decreasing centrality, as expected the initial geometrical anisotropy is more pronounced in more peripheral collisions providing fur-

⁹The first term of Eq. 1.21 is the $n = 0$ term of the Fourier series representing an isotropic distribution.

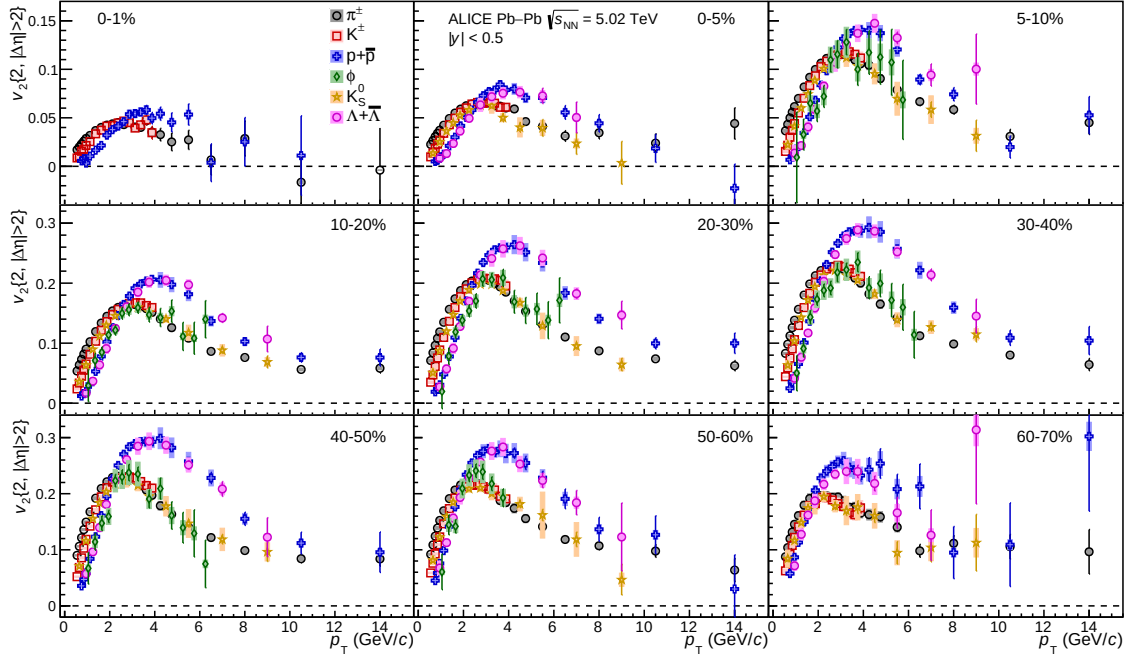


Fig. 1.13 Elliptic flow of identified charged particles as a function of p_T in Pb–Pb collisions at $\sqrt{s_{NN}} = 5.02$ TeV for different centrality intervals. The error bars and boxes represent the statistical and systematic uncertainties, respectively. Figure from [58].

ther evidence that collision geometry primarily determines the generation of flow coefficients. However, in very peripheral collisions (i.e. 50 – 60% and 60 – 70%), the shorter system lifetime and its lower density limit the development of v_2 . For $p_T < 2$ GeV/c, the elliptic flow of $\pi^\pm$, $K^\pm$, $(p + \bar{p})$, ϕ , K_S^0 , and $\Lambda + \bar{\Lambda}$ is found to be mass-ordered, with lighter particles exhibiting a larger elliptic flow than heavier particles. The mass-ordering phenomenon is attributed to the interplay with radial flow (refer to Section 1.4.4), which pushes higher-mass particles towards higher momenta. In the intermediate- p_T region ($3 < p_T < 8 - 10$ GeV/c), two distinct trends emerge depending on whether the particles are mesons or baryons. In particular, a larger v_2 is observed for baryons. This observation further supports the understanding that hadron formation occurs through the combination of multiple quarks from the medium, with each quark contributing to the overall v_2 of the resulting hadronic state. Note that this is also an indication that the flow originate at partonic level, rather than at the hadronic one.

1.4.6 High- p_T hadrons and jet quenching

High- p_T partons are mainly generated in hard scattering processes occurring in the initial stages of the collision. The partons can produce jets, focused spray of hadrons originating from the fragmentation of the high-energy parton. At leading order, these partons are produced in pairs that are back-to-back in the transverse plane of the collision, forming what is known as *di-jets*. The parton shower generated by the jet produced in the QGP, interact with the medium constituents while traversing the medium. The interactions affect the properties of the jet. These phenomena are commonly referred to as *jet quenching* [59]. One of these effects is the energy-loss of the produced hadrons. The magnitude of the energy loss depends on the path length traveled by the jet within the QGP. The energy-loss of the partons propagating through the QGP can be also studied by measuring the *nuclear modification factor* R_{AA} , defined as the ratio of the p_T -differential yield measured in heavy-ion collisions to the yield measured in pp collisions scaled by the average number of binary nucleon-nucleon collisions $\langle N_{\text{coll}} \rangle$ in the considered centrality class:

$$R_{AA} = \frac{1}{\langle N_{\text{coll}} \rangle} \frac{d^2 N^{AA} / dy dp_T}{d^2 N^{pp} / dy dp_T}. \quad (1.23)$$

Equation 1.23 is derived based on the concept of *binary scaling* of the hard scattering cross section in heavy-ion collisions. Under this assumption, in absence of nuclear effects, the R_{AA} is expected to be at unity for particles produced in hard processes, i.e. high p_T hadrons, heavy flavour hadrons, and jets. The energy loss in the QGP is expected to manifest as a suppression of the yield of high p_T hadrons and jets relative to pp collisions and therefore as $R_{AA} < 1$ at high p_T . Additionally, other effects unrelated to the presence of the QGP, known as *cold nuclear matter* (CNM) effects, can affect the R_{AA} with the expected flat trend. These effects can be studied by examining the nuclear modification factor in pA collisions and are deeply discussed in Sec. 1.5.2.

The effect of the jet quenching is also studied by measuring the two-particle azimuthal distribution correlation of the produced hadrons. Jets produced near at the edges of the fireball and directed towards the outer regions of the collision volume produce a *trigger hadron*, i.e. a leading particle used to the fine a reference ϕ axis called *near-side* ($\Delta\phi = 0$). The trigger hadron experience a shorter path length and is less suppressed. On the other hand, the hadrons produced in the parton shower jets directed towards the inner regions of the collision volume experience a longer path length and undergo a larger suppression. These hadrons are referred to as *associated particles* and are used to define the region of the azimuthal angle $\Delta\phi = \pi$, called *away-side*. The suppression of the away-side peak

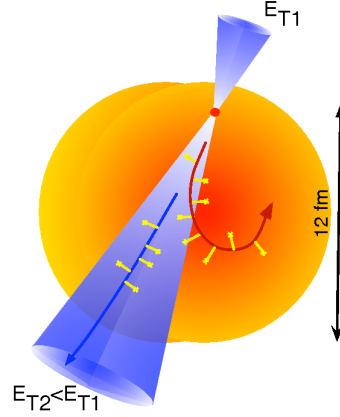


Fig. 1.14 Schematic representation of the di-jet production and jet quenching in heavy-ion collisions. Figure from Ref. [60].

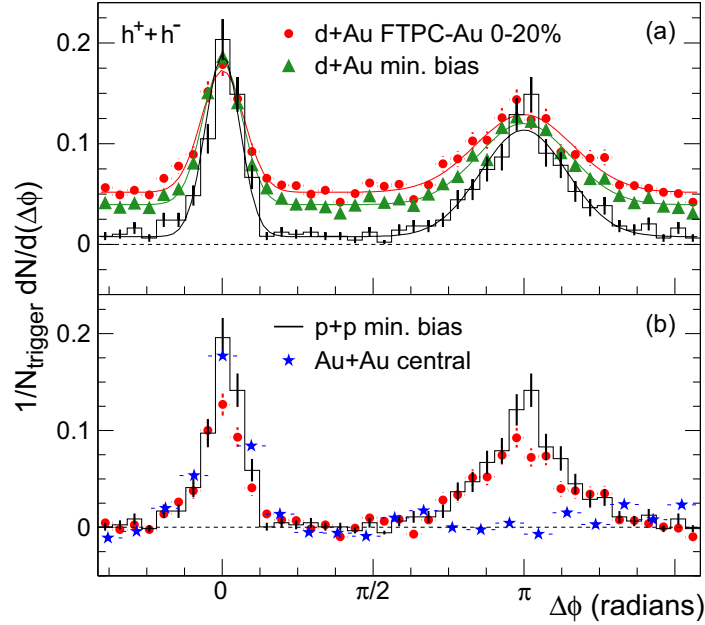


Fig. 1.15 Top: Two-particle azimuthal correlations in pp and d–Au (minimum bias and central) collisions at $\sqrt{s_{NN}} = 200$ GeV measured by the STAR Collaboration. Bottom: Comparison of pedestal-subtracted yields in pp, d–Au, and Au–Au. Figure from Ref. [59].

is a clear signature of jet quenching. A schematic representation of this phenomenon is depicted in Fig. 1.14. The results from the STAR Collaboration in pp, d–Au, and Au–Au collisions at $\sqrt{s_{NN}} = 200$ GeV [59] reported in Fig. 1.15 show a striking suppression of the away-side peak in central heavy-ion collisions. This suppression is not observed in pp and d–Au collisions, where an extended QGP medium is not expected to be created.

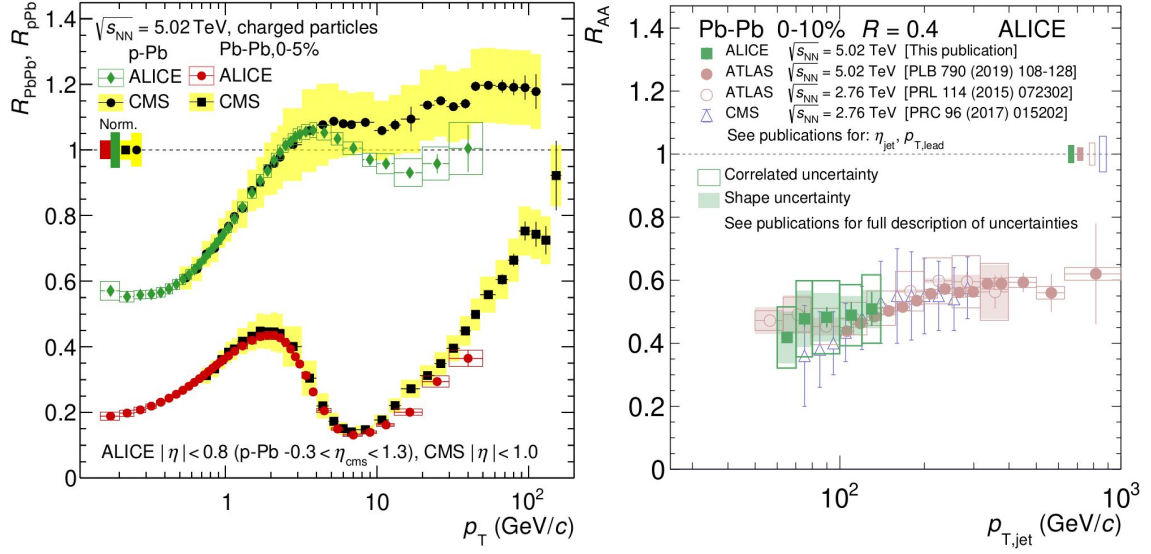


Fig. 1.16 Left: charged particle R_{AA} as a function of p_T measured by the ALICE [61] and CMS [62, 63] Collaborations in p-Pb and central (0-5%) Pb-Pb collisions at $\sqrt{s_{NN}} = 5.02$ TeV. Figure from Ref. [61]. Right: compilation of charged jet R_{AA} in central Pb-Pb collisions measured by the ALICE [67], ATLAS [68–70], and CMS [71, 72] Collaborations. Figure from Ref. [67].

The left panel of Figure 1.16 illustrates the R_{AA} of charged particles as a function of p_T , as measured by the ALICE [61] and CMS [62, 63] Collaborations in p-Pb and central (0–5%) Pb-Pb collisions at $\sqrt{s_{NN}} = 5.02$ TeV. The R_{AA} is strongly suppressed in the $p_T < 1$ GeV/c interval. In this p_T region particle production is dominated by the soft processes, which do not scale as $\langle N_{coll} \rangle$ and therefore the binary scaling assumption does not hold. Then, it reaches a maximum around 3 GeV/c. In p-Pb collisions the nuclear modification factor R_{p-Pb} is compatible with 1 for $p_T > 2$ GeV/c confirming that the suppression observed in Pb-Pb collisions is a final-state effect due to energy loss in the QGP. In addition, the nuclear modification factor of electroweak probes (i.e. photons, Z and W bosons), which do not carry any colour charge and do not interact with the QGP, is compatible with unity in Pb-Pb collisions [64–66]. For momenta higher than 10 GeV/c, a smooth growth towards unity is observed in Pb-Pb collisions, while this is not the case for jets. This is reported in the right panel of Fig. 1.16, which displays a compilation of nuclear modification factor of jets in central Pb-Pb collisions measured by the ALICE [67], ATLAS [68–70], and CMS [71, 72] Collaborations. The jet R_{AA} reaches a plateau approximately at 0.5 at high p_T . The different trend indicates that in Pb-Pb collisions a portion of the energy disperses outside the jet cone, leading to a reduced nuclear modification factor for jets, which falls below unity.

1.4.7 Quarkonium suppression

The investigation of *quarkonium* states, which are bound states of $c\bar{c}$ and $b\bar{b}$ pairs (such as the J/ψ the $\Upsilon(1S)$ mesons), offers noteworthy insights into characterisation of the QGP medium [73]. Specifically, the suppression of quarkonium states in heavy-ion collisions, initially postulated by Matsui and Satz due to Debye screening of the potential between quark-antiquark pairs [74], is a powerful probe of the deconfinement and the medium temperature. The binding energy of quarkonium in vacuum can be determined by employing the Cornell potential:

$$V(r) = kr - \frac{\alpha_s}{r}, \quad (1.24)$$

where r and k represent the distance and the string tension between the quarks, respectively, and α a Coulomb-like constant. The confinement of the quark pairs in the quarkonium states is due to the first term of Eq. 1.24, which is dominant at large distances. Immersing the quarkonium state in a deconfined medium, as the QGP, the linear term of the potential cancels by definition and a colour-screening term is introduced to take into account the interaction between the quarkonium and the medium constituents. The potential can be expressed as:

$$V(r) = -\frac{\alpha}{r} e^{-r/\lambda_D}, \quad (1.25)$$

where λ_D is the *Debye screening length* and represents the distance at which the potential is screened. The Debye screening length is inversely proportional to the temperature of the medium and, considering that the radius of the quarkonium state is inversely proportional to its binding energy, the suppression of quarkonium states is expected to increase with the temperature reached by the medium. In addition, excited quarkonium states (e.g. $\psi(2S)$, $\Upsilon(1S)$, $\Upsilon(2S)$, and $\Upsilon(3S)$) have smaller binding energies than the ground states (J/ψ and $\Upsilon(1S)$), and are therefore more easily dissociated in the QGP. Hence, a *sequential suppression* of quarkonium states is expected in heavy-ion collisions. The dissociation of quarkonium in heavy-ion collisions result in a reduction of their production relative to pp collisions which can be quantified by the R_{AA} observable (refer to Sec. 1.4.6).

The left panel of Fig. 1.17 shows the R_{AA} of the $\Upsilon(1S)$, $\Upsilon(2S)$, and $\Upsilon(3S)$ as a function of the centrality of the collisions expressed in terms of number of participant nucleons $\langle N_{\text{part}} \rangle$ measured by the CMS Collaboration in Pb–Pb collisions at $\sqrt{s_{NN}} = 5.02$ TeV [75]. The results indicate a clear ordered suppression of the bottomonium states from the $\Upsilon(1S)$ to the $\Upsilon(3S)$ from low- to high-multiplicity intervals. This is consistent with the sequential suppression picture. The right panel of Fig. 1.17 shows the R_{AA} of the J/ψ as a function of $\langle N_{\text{part}} \rangle$ measured by the ALICE Collaboration in Pb–Pb collisions at $\sqrt{s_{NN}} = 2.76$ TeV [76],

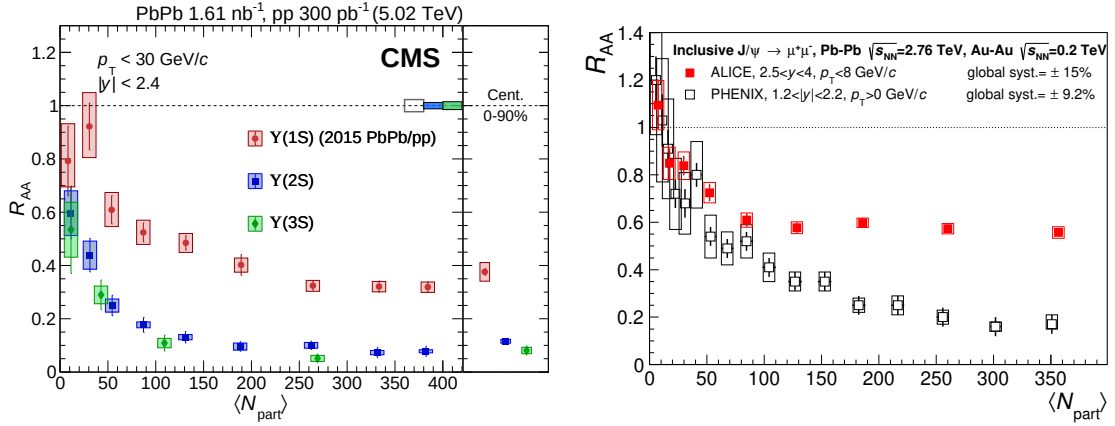


Fig. 1.17 Left: Measured R_{AA} for Υ states as functions of number of participant nucleons $\langle N_{part} \rangle$ in Pb–Pb collisions at $\sqrt{s_{NN}} = 5.02$ TeV including the centrality-integrated bin (on the right). The vertical lines correspond to statistical uncertainties, while the boxes represent the systematic uncertainties. Figure from Ref. [75]. Right: J/ψ -meson R_{AA} as a function of $\langle N_{part} \rangle$ in Pb–Pb collisions at $\sqrt{s_{NN}} = 2.76$ TeV and in Au–Au collisions at $\sqrt{s_{NN}} = 200$ GeV. The vertical bars represent the statistical uncertainties, while the boxes represent the systematic uncertainties. Figure from Ref. [76].

and by the PHENIX Collaboration in Au–Au collisions at $\sqrt{s_{NN}} = 200$ GeV [77]. The results indicate a similar suppression of the J/ψ at the two collision energies in the low-multiplicity interval. However, for more central events, the suppression of the J/ψ is observed to be almost three times larger at RHIC with respect to LHC. This phenomenon is referred to as *quarkonium regeneration*. Through this mechanism, uncorrelated charm and anti-charm quarks within the medium can statistically recombine to form the charmonium state. The magnitude of this effect is proportional to the number of $c\bar{c}$ pairs produced in the collision. The number of $c\bar{c}$ pairs is expected to be 10 times higher at LHC than at RHIC, therefore a larger regeneration is expected¹⁰.

1.5 Open heavy-flavour hadrons

Due to their large masses, $m_c \simeq 1.3$ GeV/ c^2 and $m_b \simeq 4.2$ GeV/ c^2 , heavy quarks are predominantly produced in the initial stages of the collision through hard scatterings processes, allowing them to probe the full evolution of the collision and the QGP medium [78]. Accurate measurements of heavy-flavour hadron production, like D and B mesons in heavy-ion collisions provide insights into the energy loss process and the mechanism of formation

¹⁰The quarkonium regeneration would affect the production of $b\bar{b}$ states too. However, even at the LHC energies, the production of beauty quarks is limited and the magnitude of this effect results negligible.

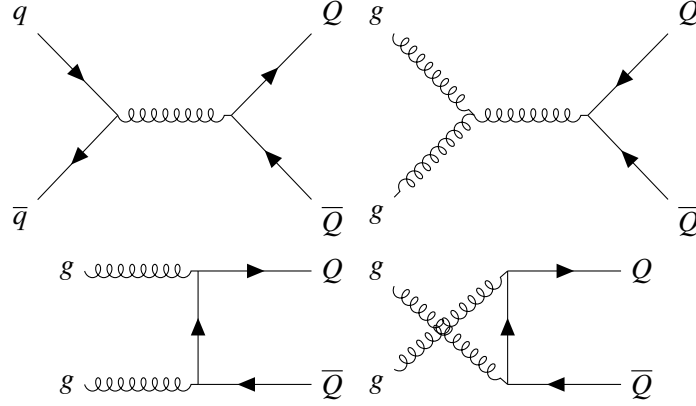


Fig. 1.18 Feynman diagrams of leading order quark-antiquark annihilation and gluon fusion for the production of $Q\bar{Q}$ pairs.

of these particles. The study of heavy-flavour hadron production in pp collisions helps validate pQCD calculations and serves as a reference for heavy-ion collisions. In this Section, the production and the hadronisation mechanisms of open heavy-flavour hadrons and their in-medium energy loss are discussed. A more detailed description of the measurements performed in heavy-ion collisions can be found in Chapter 6.

1.5.1 Open heavy-flavour production in pp collision

Perturbative QCD calculations of heavy-flavour hadron production can be tested in pp collisions. The perturbative approach is permitted by the large $\mathcal{Q}^2$ required for the production of the $c\bar{c}$ and $b\bar{b}$ pairs ($\mathcal{Q}^2 \gtrsim 4m_Q^2$), which implies a small value of the α_s parameter. The processes contributing to the production of heavy quarks in pp collisions at leading order are the annihilation of quark-antiquark pairs ($q\bar{q} \rightarrow Q\bar{Q}$) and the gluon fusion ($gg \rightarrow Q\bar{Q}$) [79]. The Feynman diagrams of these processes are depicted in Fig. 1.18. However, higher-order processes introduce a dependency on the energy scale chosen for the renormalisation of the ultraviolet divergences, usually referred to as *renormalisation scale* μ_R .

The production cross section of heavy-flavour hadrons (H_Q) can be described in pQCD calculations employing a *factorisation theorem* [80], as:

$$\begin{aligned} \sigma_{pp \rightarrow H_Q X} = & \sum_{a,b=q,\bar{q},g} f_a(x_a, \mathcal{Q}^2) f_b(x_b, \mathcal{Q}^2) \\ & \otimes \sigma_{ab \rightarrow Q\bar{Q}} \\ & \otimes D_{Q \rightarrow H_Q}(z, \mathcal{Q}^2). \end{aligned} \quad (1.26)$$

The first term of Eq. 1.26 represents the *parton distribution functions* (PDF) of the partons inside the colliding protons, which determines the parton probability to inherit a certain fraction x of the proton momentum. The second term defines the *hard scattering cross section*, describing the production of the $Q\bar{Q}$ pairs from the interaction of initial partons, which can be calculated as a perturbative series in α_s . The term $D_{Q \rightarrow H_Q}(z, \mathcal{Q}^2)$ is the *fragmentation function* (FF), which quantifies the probability of newly produced quarks forming a hadron H_Q with a fraction z of the momentum of Q . The value of the FFs integrated over z provides the total probability of a quark to hadronise into a given hadron species H_Q and is commonly referred to as *fragmentation fraction*. PDFs and FFs can not be calculated using a pQCD approach as they involve non-perturbative processes. They are typically parameterised on measurements from deep-inelastic scattering [81, 82] and e^+e^- measurements [83, 84], respectively. This approach assumes the universality of the fragmentation function across collision systems. PDFs strongly depend on the momentum transfer $\mathcal{Q}^2$, thus the parametrisation is performed at a given $\mathcal{Q}_0^2$ and evolved in second step to the desired $\mathcal{Q}^2$ scale by means of the Dokshitzer-Gribov-Lipatov-Altarelli-Parisi (DGLAP) equations [85]. Usually, both the *normalisation scale* (μ_R) and the *factorisation scale* (μ_F), which determine the energy scales for the validity of the perturbative regime, are chosen to be of the same order of the momentum transfer of the hard process $\mu_F \sim \mu_R \equiv \sqrt{m_Q^2 + p_{T,Q}^2}$.

State-of-the-art pQCD calculations that employ the collinear factorisation approach, such as FONLL [86, 87] and GM-VFNS [88, 89] calculations, provide predictions for the production cross sections of open heavy flavour hadrons. These calculations achieve *next-to-leading order* (NLO) accuracy while incorporating the resummation of *next-to-leading logarithms*. The two calculations differ on the choice of the PDFs [81, 82], of the fragmentation function, of the heavy quark mass, and in the way the calculations in the low p_T region, performed at fixed order, are matched with the ones at high p_T .

These choices contribute to the uncertainty on the theoretical predictions, which also include the variation of the renormalisation and factorisation scales. Other calculations at the *leading order* (LO) approximation and including higher-order corrections, implement the factorisation in terms of the partonic transverse momentum k_T , the k_T -factorisation approach [90–92]. In addition, MC event generators such as PYTHIA [93] and HERWIG [94] can be employed to describe the production of heavy-flavour hadrons. These MC generators offer a more comprehensive representation of the final state compared to pQCD calculations. They incorporate soft and multi-parton interactions, parton showers in the initial and final states, and descriptions of hadron decays. However, they only implement LO processes precisely and approximate higher-order effects. Event generators like MC@NLO [95] and POWHEG [96] provide next-to-leading order accuracy, predicting differential observables

and jets. They can be effectively combined with parton showers generated by PYTHIA and HERWIG.

The left panel of Fig. 1.19 shows the p_T -differential production cross section of D^+ mesons originating directly from the charm-quark hadronisation or from decays of excited charm-hadron states (*prompt*) measured at mid-rapidity in pp collisions at $\sqrt{s} = 13$ TeV [97]. The measurements are compared with the predictions from FONLL, GM-VFNS, and k_T -factorisation calculations. The measurements lie on the upper-edge of the predictions from FONLL, while the k_T -factorisation calculation is in good agreement with the data in the low- and intermediate- p_T region, but it tends to overestimate the data at high p_T . The central values of the calculation in the GM-VFNS framework slightly underestimate the data in the low- p_T region and overestimate the data at high p_T . The right panel of Fig. 1.19 shows the p_T -differential production cross section of D^+ mesons originating from the decay of a beauty hadron (*non-prompt*) at the same collision energy. The data are in agreement with FONLL calculations coupled with PYTHIA 8 to describe the decay of B mesons and the one from TAMU [98], which is based on the predictions from FONLL for the beauty hadron p_T -differential cross sections but adopt a statistical hadronisation approach to describe the beauty quark fragmentation fractions.

1.5.2 Cold nuclear matter effects

As anticipated in Sec. 1.4.6, there are effects, unrelated to the formation of the QGP medium, which may occur in p–Pb and Pb–Pb collisions that can affect the yield and the momentum distribution of the produced particles. They are usually referred to as *cold nuclear matter* (CNM). At LHC energies, the predominant effect is the *modification of the PDFs* [99]. If nuclei were merely a simple collection of independent nucleons, the expected modification in the PDFs would arise only from the Fermi motion of nucleons inside the nucleus [100]. This motion influences the distribution of the momentum fraction x carried by the partons. However, *deep-inelastic scattering* (DIS) experiments conducted with pA, dA, and A–A collisions have revealed additional effects that vary depending on the range of x (*Bjorken x*) under consideration. The degree of modification in the PDFs is typically measured by the ratio [101]:

$$R(x, \mathcal{Q}^2) = \frac{F_2^A(x, \mathcal{Q}^2)}{AF_2^N(x, \mathcal{Q}^2)}, \quad (1.27)$$

where $F_2^A(x, \mathcal{Q}^2)$ and $F_2^N(x, \mathcal{Q}^2)$ are the structure functions of the nucleus, with A nucleons, and of the nucleon, respectively. Figure 1.20 illustrates the modulation of the ratio $R(x, \mathcal{Q}^2)$

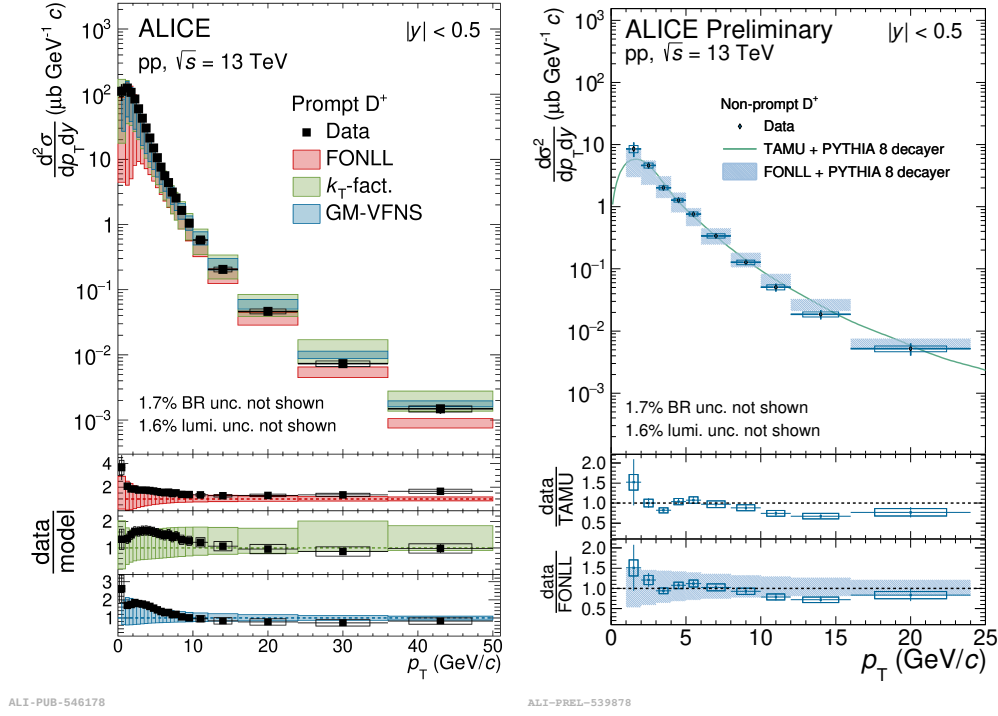


Fig. 1.19 Left: Comparison of the FONLL, GM-VFNS, and k_T -factorisation calculations with the measured p_T -differential production cross section of prompt D^+ mesons in pp collisions at $\sqrt{s} = 13$ TeV. Figure from Ref. [97]. Right: Comparison of the FONLL and TAMU calculations with the measured p_T -differential production cross section of non-prompt D^+ mesons in pp collisions at $\sqrt{s} = 13$ TeV. Figure from ALICE Figure repository ©.

as a function of x . In case of no CNM effects, the ratio $R(x, \mathcal{Q}^2)$ would be equal to unity in the full x region. For $x \lesssim 0.05$, a phenomenon known as *shadowing* occurs [102, 103], leading to a strong suppression of the PDFs. The spatial overlap among low- x partons from different nucleons is assumed to cause this effect. According to the *Colour Glass Condensate* (CGC) model [27], an effective field theory for QCD at low x , the gluon density steeply increases at low x saturating the phase space occupancy. At $0.05 \lesssim x \lesssim 0.3$ the PDFs are enhanced with respect to the free nucleon case, a phenomenon known as *anti-shadowing*. This effect is attributed to the depletion of the low- x gluon density which requires a slight increase of the PDFs at intermediate x to ensure the momentum conservation. At $0.3 \lesssim x \lesssim 0.8$ a second suppression region is found, and the ratio reaches a minimum at $x \simeq 0.7$. This phenomenon is the so-called *EMC effect* [104]. It is attributed to the modification of the PDFs of valence quark due to the nuclear environment. At higher values of x the Fermi motion causes the increase of the ratio. Analogously to the utilisation of free-proton PDFs in theoretical calculations for pp collisions, the nuclear PDFs are parameterised through comprehensive fits encompassing various measurements in DIS and of Drell-Yan processes [105, 106]. The

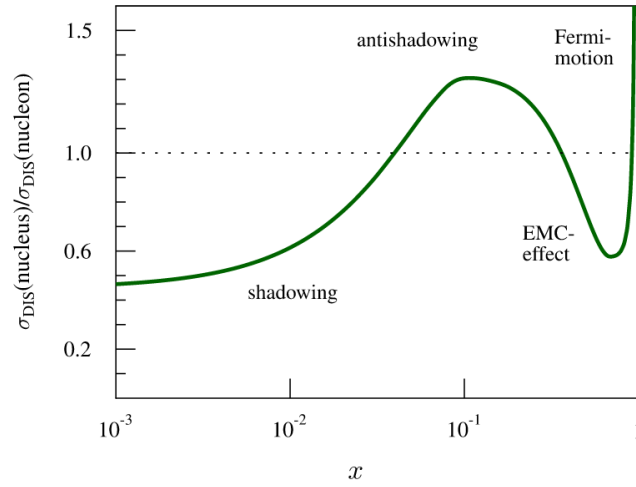


Fig. 1.20 Illustration of the modification of the nucleon PDFs in a nucleus observed in deep inelastic scattering experiments. Figure from Ref. [101].

production of heavy-flavor hadrons in heavy-ion collisions at LHC energies is primarily affected by the shadowing effect. Specifically, the formation of $c\bar{c}$ and $b\bar{b}$ quark-pairs at midrapidity arises from hard scatterings involving partons with Bjorken- x values around 10^{-4} for charm quarks and 10^{-3} for beauty quarks, which increase as the p_T increase [107]. Consequently, the yield of heavy flavors, normalised to the number of nucleon-nucleon collisions, is suppressed in comparison to pp interactions, particularly in the low- p_T region.

Another important component of CNM effects is the *Cronin enhancement* [108] which is a broadening of the p_T -distribution of the produced particles compared to that in pp collisions. The Cronin enhancement was explained as due to the multiple scattering of partons within the colliding nuclei before the hard scattering process takes place. The transverse momentum transfer during these scatterings introduces an additional transverse kick to the partons participating in the hard scattering process (k_T broadening). Consequently, the resulting particles in pA and A–A collisions exhibit an increased p_T distribution compared to pp collisions. As discussed in Sec. 1.4.4, an analogous modification of the p_T distribution is observed in A–A collisions [109–111]. This effect is attributed to the radial flow. Interestingly, collective-like behaviours have been observed in pA collisions, raising the possibility of the establishing of flow in these collisions as well which would produce an alternative explanation of the modification of the p_T spectra in pA collisions known as *Cronin effect*. However, the debate on the origin of the observed phenomena is still open.

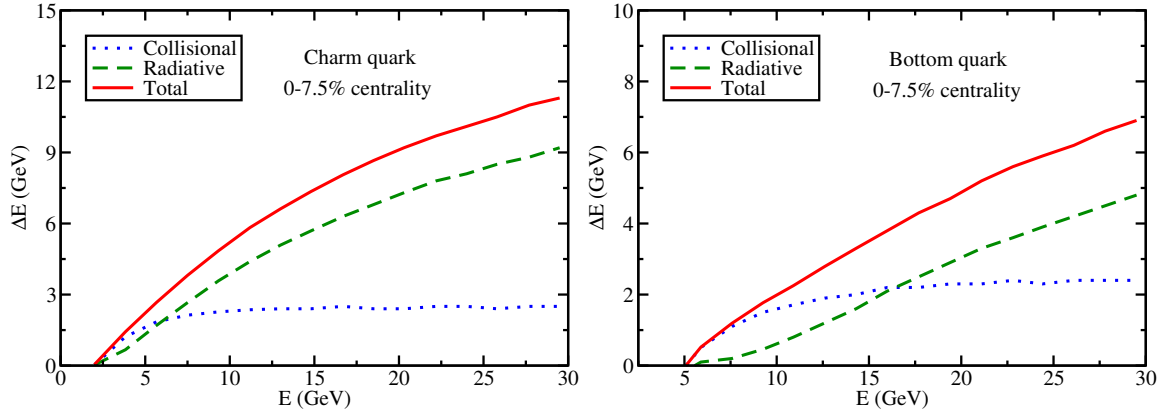


Fig. 1.21 Collisional and radiative energy loss of charm (left) and beauty (right) quarks in the QGP as a function of their initial energy in central Pb–Pb collisions at $\sqrt{s_{\text{NN}}} = 2.76$ TeV, as predicted by the model outlined in Ref. [113]. Figure from Ref. [113].

1.5.3 Interaction of heavy quarks with the QGP

Heavy quarks traversing through the QGP interact with the constituents of the medium, engaging in multiple elastic and inelastic (i.e. gluon radiation) processes. These interactions play a significant role in the dynamics of partons in the medium. The exchanged momentum during these scatterings is typically small compared to the mass of charm and beauty quarks. Consequently, the propagation of heavy quarks in the medium, whose constituents work as scattering centres, can be likened to a Brownian motion characterised by numerous small momentum kicks [78]. The interactions with the constituents of the QGP bring low- and intermediate- p_T heavy quarks closer to thermal equilibrium with the medium, effectively transferring the collective flow of the system to them. The thermalisation time of heavy quarks in the medium is comparable to or even longer than the lifetime of the QGP [112]. As a result, the momentum spectra of heavy-flavour hadrons conserve a memory of their interaction history, thereby providing valuable information about the strength of their coupling to the medium. For energetic (high momentum) quarks, the main effect of the interactions is the in-medium energy loss. Figure 1.21 illustrates the collisional and radiative energy loss ΔE of charm (left) and beauty (right) quarks in the QGP as a function of their initial energy E , as predicted by the model outlined in Ref. [113]. In central (0 – 7.5%) Pb–Pb collisions at $\sqrt{s_{\text{NN}}} = 2.76$ TeV, these theoretical calculations indicate that collisional energy loss is the prevailing mechanism for low- p_T heavy quarks. Instead, at high transverse momentum, the largest portion of energy loss occurs through radiative processes (i.e. medium-induced gluon radiation).

Collisional processes

The effect of the multiple elastic collisions between the heavy quarks, traversing the medium, and the medium constituents results in an average energy loss ΔE_{coll} , which depends on the mass of the quark [114, 115]. In particular, ΔE_{coll} is higher for charm quarks ($m_c \sim 1.5 \text{ GeV}/c^2$) than for beauty quarks ($m_b \sim 4.2 \text{ GeV}/c^2$) and the p_T range at which the radiative processes become dominant is higher for beauty quarks. The collisional energy loss can be expressed as [116]:

$$\langle \Delta E_{\text{coll}} \rangle \approx \frac{1}{\sigma T} \int t \frac{d\sigma}{dt} dt, \quad (1.28)$$

where σ refers to the parton-parton cross section, T is the temperature of the medium and t is the momentum transfer in the collision. In particular, $d\sigma/dt$ describes the probability of the parton to interact with the medium constituents per unit of momentum. Assuming the dominant process contributing to the cross section is:

$$\frac{d\sigma}{dt} \approx \frac{4\pi C_i \alpha_s^2}{t^2}, \quad (1.29)$$

where C_i is the colour factor and α_s is the strong coupling constant, then, in the limit in which the initial energy of the quark (E_Q) is much larger than its mass, $(m_Q) E_Q \gg m_Q^2 T$, $\langle \Delta E_{\text{coll}} \rangle$ is expected to increase logarithmically with E_Q and linearly with the thickness of the traversed medium. Nevertheless, at low E_Q , the exchange of energy and momentum between the heavy quarks and the medium leads the quarks towards thermalisation with the surrounding medium. In this scenario, the radial and anisotropic flow of the system are transferred to the charm and beauty quarks. The dynamics and space-temporal evolution of heavy quarks can be formally described using the Boltzmann equation in the transport formalism [78]. This equation provides a framework to understand the behaviour of heavy quarks in the medium, taking into account their interactions with the QGP constituents and the resulting energy exchanges and thermalisation processes.

$$\left[\frac{\partial}{\partial t} + \frac{\vec{p}}{E_Q} \frac{\partial}{\partial \vec{x}} + \vec{F} \frac{\partial}{\partial \vec{p}} \right] f_Q(\vec{x}, \vec{p}, t) = C[f_Q] \quad (1.30)$$

The left-hand side of Eq. 1.30 describes the evolution of the distribution function $f_Q(\vec{x}, \vec{p}, t)$ of heavy quarks in the phase space, where $\vec{x}$ and $\vec{p}$ are the spatial and momentum coordinates, respectively. The first term on the left-hand side of Eq. 1.30 describes the time evolution of the distribution function; the second term describes the spatial evolution of the distribution function; and the third term describes the evolution of the distribution function in momentum space by means of the force $\vec{F}$, which includes the electromagnetic and strong interactions.

The right-hand side of Eq. 1.30 describes the collisional processes, which are modelled using the Boltzmann collision term $C[f_Q]$, including the parton scattering amplitudes. Assuming the heavy quarks to follow a Brownian motion, i.e. the momentum transfer during the scattering processes is small compared to the mass of the quark, Eq. 1.30 can be approximated by the Fokker-Planck equation:

$$\frac{\partial}{\partial t} f_Q(t, \vec{p}) = \frac{\partial}{\partial p_i} \{A_i f_Q(t, \vec{p}) + \frac{\partial}{\partial p_j} [B_{ij}(\vec{p}) f_Q(t, \vec{p})]\}, \quad (1.31)$$

where A_i and B_{ij} are the drift and diffusion coefficients, respectively. They both describe the evolution of the distribution function in momentum space per unit time. The drift coefficient A_i provides information on the friction experienced by the quark in the medium, while the diffusion coefficient B_{ij} describes the average momentum broadening. In the non-relativistic limit, the Fokker-Planck equation can be solved analytically, yielding the Langevin equation:

$$\frac{\partial}{\partial t} f_Q(t, p) = \gamma \frac{\partial}{\partial p_i} [p_i f_Q(t, p)] + D_p \nabla_{\vec{p}} f_Q(t, p), \quad (1.32)$$

where $\gamma \equiv A_i(\vec{p})$ is the drag coefficient and $D_p \equiv B_{ij}(\vec{p})$ is the momentum-diffusion coefficient. The friction coefficient γ is the inverse of the thermalisation time τ_Q^{eq} of heavy quarks. Furthermore, the drag coefficient is related to the spatial-diffusion coefficient D_s by the relation:

$$D_s = \frac{T}{m_Q \gamma} = \frac{T}{m_Q^2} \tau_Q^{\text{eq}}. \quad (1.33)$$

At LHC energies, the thermalisation time of charm and beauty quarks is expected to be of the order of 5 – 15 fm/c which is compatible with the lifetime of the QGP phase [117, 78, 117].

Radiative processes

The radiation of gluons is the main process contributing to the energy loss of energetic heavy quarks in the QGP. Different calculations of the average radiative energy loss $\langle \Delta E_{\text{rad}} \rangle$ have been proposed, employing various assumptions. For instance, the energy loss can be calculated using the BDMPS formalism [118, 119], which is based on the assumption that the parton traversing the medium interacts with the medium constituents which are treated as static interaction points. The parton is assumed to undergo multiple soft scattering processes acquiring a transverse momentum k_T at each scattering. When the transverse momentum acquired by the parton is large enough, the parton can decohere a gluon from its wave

function. The average energy loss of the parton is then given by:

$$\langle \Delta E_{\text{rad}} \rangle \propto \alpha_s C_R L^2 \hat{q}, \quad (1.34)$$

where C_R is the Casimir operator of the representation of the parton, L is the length of the medium traversed by the parton, α_s is the strong coupling constant, and $\hat{q} = \langle k_T^2 \rangle / L$ is the transport coefficient of the medium. The Casimir operator is defined as $C_R = (N_c^2 - 1)/(2N_c)$ for the fundamental representation and $C_R = N_c$ for the adjoint representation, where N_c is the number of colours [120]. In particular, for a quark-gluon interaction $C_R = 4/3$, while for gluon-gluon interactions $C_R = 3$. In the BDMPS formalism the energy loss does not depend on the initial parton energy, but rather on the properties of the medium and the length travelled by the parton inside it. Additionally, the average energy loss for a gluon is expected to be 9/4 times larger than the one of a quark due to the different C_R parameters. In other words, particles produced in gluon-initiated jets, such as light-flavour hadrons, are expected to suffer more from parton energy loss and thus to be more suppressed with respect to heavy-flavour hadrons. On top of this, the gluon radiation for different quark flavours does not coincide. The emission of a gluon is forbidden in the angular region $\Theta < \Theta_c \equiv m/E$, meaning that the larger the quark mass, the smaller the angular region where gluon radiation is allowed, thus reducing the rate of gluon radiation. This phenomenon is known as the *dead-cone effect* [121, 122]. As a consequence, the energy loss is expected to be mass-ordered $\langle \Delta E_{\text{rad}}(b) \rangle < \langle \Delta E_{\text{rad}}(c) \rangle < \langle \Delta E_{\text{rad}}(u, d, s) \rangle < \langle \Delta E_{\text{rad}}(g) \rangle$. This effect is not only expected in presence of a QGP medium in heavy-ion collisions, but it is also observed in the parton shower of jets in vacuum in pp collisions. Only recently, the first direct observation of the dead-cone effect have been performed by the ALICE Collaboration [121].

The effect of the energy loss on heavy-flavour hadrons can be studied by measuring the nuclear modification factor R_{AA} of heavy-flavour hadrons in Pb–Pb collisions (see Sec. 1.16). The left panel of Fig. 1.22 shows the R_{AA} of prompt D^0 , D^+ , and D^{*+} mesons as a function of p_T in central (0 – 10%), semi-central (30 – 50%), and peripheral (60 – 80%) Pb–Pb collisions, and in p–Pb collisions at $\sqrt{s_{NN}} = 5.02$ TeV [117]. The R_{AA} of the D mesons is compatible with unity in the measured p_T range in p–Pb collisions. However, it progressively decreases from unity moving from peripheral to central collisions, where the suppression factor is around 5. The suppression is particularly evident for $p_T > 6$ GeV/c, where the radiative processes are dominant. The right panel of Fig. 1.22 shows the R_{AA} of prompt (i.e. originating from charm quarks hadronisation) and non-prompt (i.e. originating from beauty-hadrons decay) D^0 mesons in central (0 – 10%) Pb–Pb collisions at $\sqrt{s_{NN}} = 5.02$ TeV as a function of p_T . The non-prompt D^0 R_{AA} is systematically higher than the prompt D^0 one

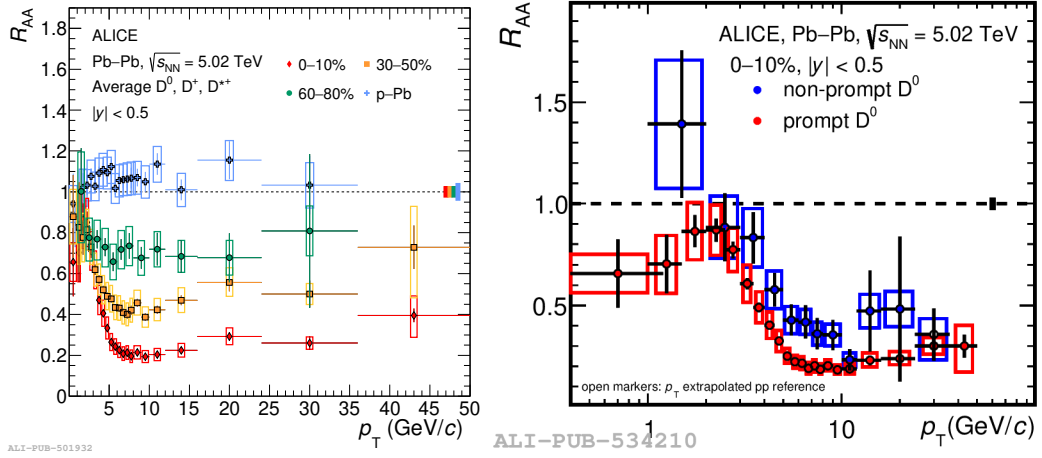


Fig. 1.22 Left: average R_{AA} of prompt D^0 , D^+ , and D^{*+} mesons as a function of p_T in central (0 – 10%), semi-central (30 – 50%), and peripheral (60 – 80%) Pb–Pb collisions, and p–Pb collisions at $\sqrt{s_{NN}} = 5.02$ TeV. Figure from Ref. [117]. Left: R_{AA} of prompt and non-prompt D^0 mesons in central (0 – 10%) Pb–Pb collisions at $\sqrt{s_{NN}} = 5.02$ TeV as a function of p_T . Figure from Ref. [123].

starting from $p_T > 5$ GeV/c, as expected by the dead-cone effect causing a smaller energy loss of beauty quarks as compared to charm quarks [123].

1.6 Azimuthal anisotropies

As anticipated in Sec. 1.4.5, the azimuthal anisotropies in the momentum distribution of the particles produced in non-central heavy-ion collisions are a powerful tool to study the properties of the QGP. The initial geometrical anisotropy of the system translates to an anisotropy in the momentum space of the particles produced in the collision, which undergo a collective expansion. However, being produced in the early stages of the collision, heavy quarks are expected to be less affected by the collective expansion of the medium with respect to their lighter counterparts, which compose the bulk of the system. Indeed, heavy quarks are not expected to be in thermal equilibrium with the light partons at the QGP-formation time. Nevertheless, by interacting with the medium constituents as discussed in the previous Sections, the heavy quarks inherit (part of) the collective motion of the medium which manifest under the form of a non-vanishing elliptic or triangular flow [124, 125], in addition to a modification of the p_T spectra which impacts on the R_{AA} at low and intermediate p_T ($p_T < 4 - 5$ GeV/c). Measurements of the azimuthal anisotropies of heavy-flavour hadrons can therefore provide information on the transport properties of the QGP complementing the information that can be extracted from R_{AA} .

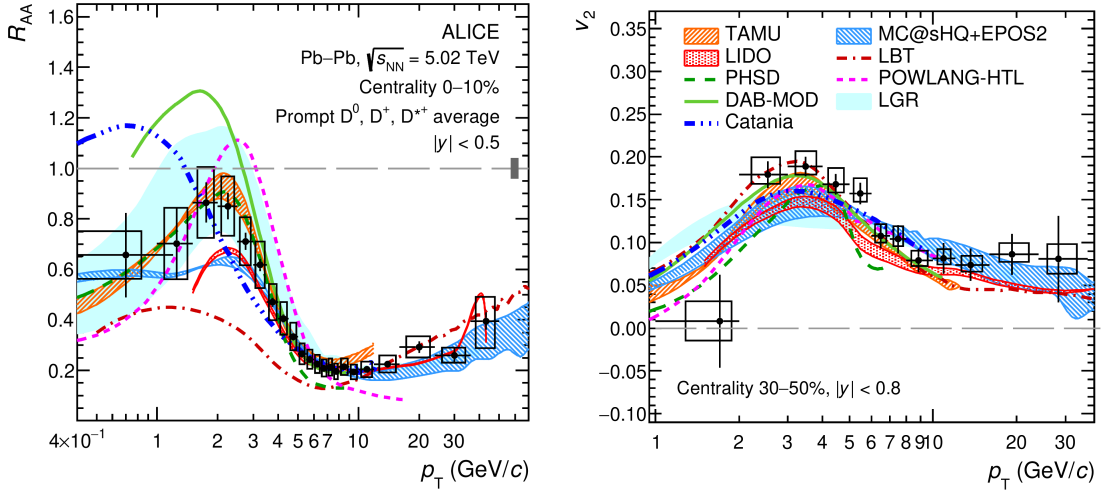


Fig. 1.23 Left: Average R_{AA} of prompt D^0 , D^+ , and D^{*+} mesons as a function of p_T in central (0–10%) Pb–Pb collisions at $\sqrt{s_{NN}} = 5.02$ TeV compared with different transport models. Figure from Ref. [117]. Right: Average v_2 of prompt D^0 , D^+ , and D^{*+} mesons as a function of p_T in semi-central (30–50%) Pb–Pb collisions at $\sqrt{s_{NN}} = 5.02$ TeV compared with different transport models. Figure from Ref. [117].

Figure 1.23 shows the average R_{AA} and v_2 of prompt D^0 , D^+ , and D^{*+} mesons measured in central (0–10%) and semi-central (30–50%) Pb–Pb collisions at $\sqrt{s_{NN}} = 5.02$ TeV as a function of p_T [117]. The measurements are compared with different transport-model predictions. These models implement the transport of the heavy quarks either employing the Langevin equation (TAMU [126], POWLANG-HTL [127], DAB-MOD [128], and LGR [129]) or the Boltzmann equation (Catania [130], MC@sHQ+EPOS2 [125], LBT [131], and LIDO [132]). Both the elastic and radiative energy loss process are considered in the model description, with the exceptions for TAMU, POWLANG-HTL, PHSD, and Catania which include only the collisional process (which are dominant at low p_T). The hadronisation of heavy quarks is permitted via two competing mechanisms: the coalescence, discussed in Sec. 1.7, and the fragmentation. The models that better describe¹¹ the data use values of the charm quark diffusion coefficient $2\pi T D_s$ between 1.5 and 7 at the critical temperature $T_c = 155$ MeV corresponding to a charm-quark thermalisation time of about 3–8 fm/c.

Further insights into the heavy flavour transport properties can be obtained by comparing the v_2 of charm and beauty hadrons. Figure 1.24 shows the v_2 (and v_3) of the prompt and non-prompt D^0 mesons in central (0–10%) and semi-central (10–30% and 30–50%) Pb–Pb

¹¹The model-to-data agreement is estimated by computing the reduced χ^2 . Only model with a $\chi^2_{R_{AA}} < 5$ and $\chi^2_{v_2} < 2$ are then considered.

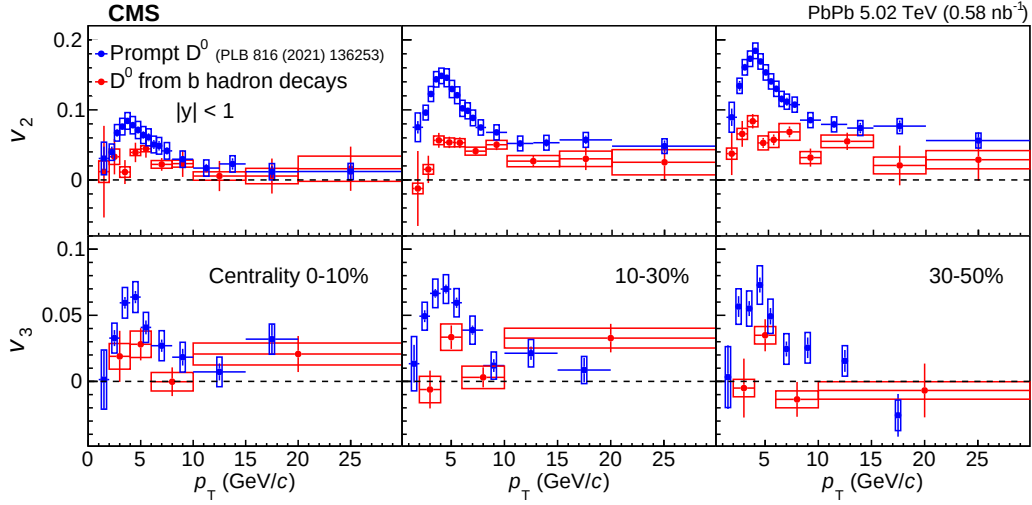


Fig. 1.24 Elliptic, v_2 (upper panels), and the triangular, v_3 (lower panels), flow of prompt [134] and non-prompt [133] D^0 mesons as functions of p_T in central (0 – 10%), semi-central (10 – 30% and 30 – 50%) Pb–Pb collisions at $\sqrt{s_{NN}} = 5.02$ TeV measured by the CMS Collaboration. Figure from Ref. [133].

collisions at $\sqrt{s_{NN}} = 5.02$ TeV as a function of p_T measured by the CMS Collaboration [133]. The v_2 of prompt D^0 mesons is systematically higher than the non-prompt D^0 one in all the three centrality intervals. This suggests that beauty quarks participate less to the collective expansion of the medium with respect to the charm quarks. The thermalisation time is expected to scale with the mass of the quark, thus the thermalisation time of beauty quarks is expected to be approximately 3 times longer than the one of charm quarks meaning that it is unlikely that beauty quarks completely thermalise in the QGP. For what concerns the triangular flow, the current experimental uncertainties do not allow for a significant separation between prompt and non-prompt D^0 .

Furthermore, more quantitative conclusion can be drawn by exploiting the predictive power of data-driven models based on Bayesian inference. The left panel of Fig. 1.25 shows the heavy-quark spatial diffusion coefficient as a function of T in units of T_c as estimated from the current measurements of R_{AA} and v_2 performed by the ALICE [135, 136] and CMS [137–139] Collaborations in central and semicentral Pb–Pb collisions. The predicted range of the $2\pi D_s$ parameter of heavy quark starts to be significantly constrained, especially considering the limited amount of measurements used as input in the case of the beauty quark. The upgrade of the ALICE experiment in view of the LHC Run 3, including the installation of a completely new inner tracking system, together with the expected increase of the integrated luminosity (10 nb^{-1} [140, 141]), will allow for a significant improvement

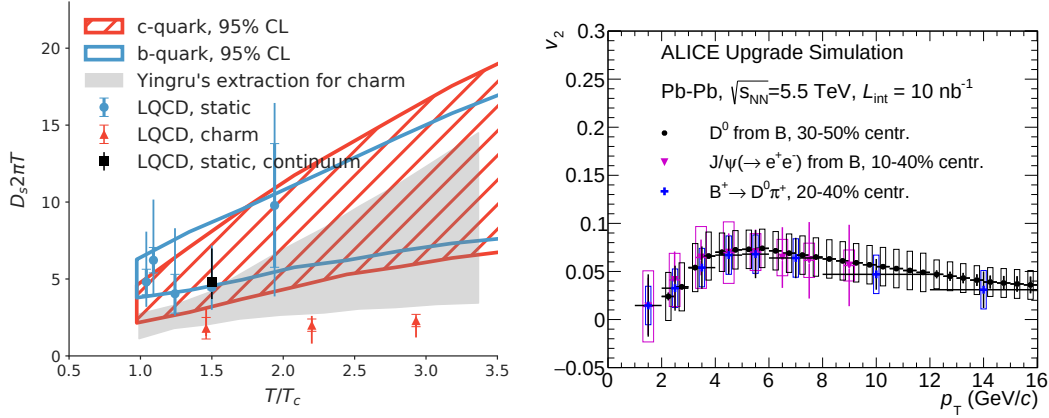


Fig. 1.25 Left: Average v_2 of prompt D^0 , D^+ , and D^{*+} mesons as a function of p_T in central Pb–Pb collisions at $\sqrt{s_{NN}} = 5.02$ TeV compared with different transport models. Figure from Ref. [142]. Right: Perspectives v_2 measurements of non-prompt D^0 , non-prompt J/ψ , and $B^+(\rightarrow D^0)$ mesons by the ALICE experiment in Pb–Pb collisions at $\sqrt{s_{NN}} = 5.5$ TeV in Run 3. Figure from Ref. [141].

of the precision of the measurements, such as the extension of the v_2 of beauty hadrons via the non-prompt D^0 mesons measurements down to very low p_T ($p_T \sim 1$ GeV/c) and fully-reconstructed B-meson decays at midrapidity [140, 141], as shown in the right panel of Fig. 1.25.

1.7 Hadronisation

The process leading a system of quarks to form the hadronic final state is called *hadronisation*. Two competing hadronisation mechanisms can occur at the hadronisation of the QGP: the *fragmentation* and the *coalescence* (or *recombination*). The former is a mechanism that describes the hadronisation of the parton showers in vacuum (such as the case of jets produced in pp collisions) and consists of a high- p_T parton emitting a spray of lower-momentum hadrons. This process is expected to dominate for high p_T partons, which escape from the QGP and hadronise via fragmentation in the vacuum. A microscopic description of the fragmentation mechanism relies on the *Lund string model* [143–145]. In the Lund string model only the linear term of the Cornell potential (see Eq. 1.24) is considered, meaning that the potential increases when quark and antiquark are taken apart from each other. Energy is thus accumulated in the *string* (i.e. a colour full tube) which connects the two quarks. In this scenario the hadronisation takes place when the potential is sufficiently high to form a new quark-antiquark pair breaking the colour string. The PYTHIA event generator, one of the most used MC generators in the field, adopts the Lund string

model to describe the hadronisation process. Even though the Lund string model is able to describe the hadronisation process some limitations were observed in pp collisions at LHC energies, especially for baryon production [146, 147]. Notably, the ALICE Collaboration recently measured the production of Λ_c^+ baryons in pp collisions at $\sqrt{s} = 13$ TeV [146] down to very low p_T , finding a significant discrepancy for the Λ_c^+/D^0 yield ratio with respect to the PYTHIA predictions. Several explanations have been proposed to explain this discrepancy, which mainly includes the presence of a colour reconnection mechanism. The hadronisation process is expected to be radically different in presence of the QGP medium. At the moment when the transition temperature to the deconfined phase is reached, a bulk of deconfined partons, which have thermalised with the medium, is formed. This makes unnecessary the production of additional quark-antiquark pairs from string breaking. Instead, it is energetically convenient for the partons from the bulk, sharing common phase space coordinates (i.e. position and velocity), to combine into hadrons [47, 130, 148, 127, 132]. Indeed, the recombination mechanism is typical of heavy-ion collisions and consists of two or three quarks recombining among themselves in the medium to form a higher- p_T hadron. This process is expected to dominate at low p_T and is supported by the measurements of the v_2 of identified particles, p_T -differential baryon-to-meson ratios, as well as the R_{AA} heavy-flavour hadrons and J/ψ . Various implementations of this mechanism are available, most of them rely on the Wigner-function formalism [149].

The study of the production of strange heavy-flavour mesons can help to shed light on the description of the hadronisation mechanism of charm and beauty quarks. These measurements complement the one of non-strange D mesons in pp collisions and contribute to the understanding of the fragmentation of heavy quarks by means of the fragmentation fraction of strange-to-nonstrange mesons $\frac{f_s}{f_u+f_d}$. In Pb–Pb collisions, the production of strange mesons is expected to be enhanced with respect to the one of non-strange mesons due to the effect of the strangeness enhancement (See Sec. 1.4.3). In particular, measuring the production of non-prompt D_s^+ mesons is crucial to reach a complete description of the energy loss and hadronisation mechanisms of beauty quarks in the QGP.

Chapter 2

ALICE: A Large Ion Collider Experiment

In order to have a comprehensive understanding of the measurements presented in this Thesis, a description of the experimental framework is necessary. In this Chapter the CERN Large Hadron Collider (LHC) facility is presented together with a detailed description of the ALICE experiment setup and the ALICE offline framework in Section 2.1, Section 2.2, and Section 2.3, respectively. The centrality determination procedure in Pb–Pb collision is described in Section 2.4. Finally, a detailed description of the ALICE track reconstruction, primary and secondary vertex determination, and particle identification capabilities is given in Section 2.5 and Section 2.6, respectively. In addition, a description of the ALICE upgrades planned for the LHC Run 3 and Run 4 are given in Chapter 7.

2.1 The Large Hadron Collider

The *Conseil Européen pour la Recherche Nucléaire* (CERN) is the European Organisation for Nuclear Research. CERN officially opened on 29 September 1954 when the twelve founding members signed the CERN Convention. Nowadays, the CERN community accounts for 23 Member States, over 110 nationalities, and more than 17,000 scientists [150]. Italy is one of the founding members and, in the specific case of the ALICE experiment, over 200 Italian scientists contribute to the Collaboration [151]. CERN owes its remarkable achievements to the vast community of scientists and researchers who have contributed to revolutionary discoveries such as the identification of the W and Z bosons in 1983 [152], the foundation of the World Wide Web in 1989 [153], and the discovery of the Higgs boson in 2012 [154, 155].

However, what makes the CERN complex in Geneva so unique nowadays is the presence of the largest and most powerful particle accelerator in the world, the *Large Hadron Collider* (LHC).

The LHC is a 26.7-kilometre-long circular accelerator located in a tunnel underground the French-Swiss border near Geneva. The tunnel, that was used to host the *Large Electron-Positron Collider* (LEP), is about 3 metres in diameter and is located at a depth which ranges between 50 and 175 metres below the surface. In the LHC era, and differently from what was done at LEP, protons and ions, notably lead ions ($^{208}\text{Pb}^{82+}$), are injected and accelerated in the LHC reaching a maximum centre-of-mass energy of $\sqrt{s} = 14$ TeV and $\sqrt{s_{\text{NN}}} = 5.5$ TeV, respectively¹. The LHC injection scheme, shown in Figure 2.1, involves several machines that are used to collect the particle beams and increase their energy until it reaches the required energy for the LHC ring. Collecting, storing, and accelerating protons and ions is a very complex task, and it requires rather different steps for protons and ions. Protons are firstly collected by stripping the electrons from a hydrogen gas with an electric field and secondly grouped into bunches. The *Linear accelerator 2* (Linac 2), the first linear accelerator in the LHC injection chain, accelerates the protons to 50 MeV. The protons are then injected into the *Proton Synchrotron Booster* (PSB), which accelerates them to 1.4 GeV, and then into the *Proton Synchrotron* (PS) and the *Super Proton Synchrotron* (SPS), reaching 25 GeV and 450 GeV, respectively. The protons are finally injected in the LHC, where they are accelerated up to a maximum energy of 7 TeV. In the LHC, the beams are forced to follow the circular path of the ring by strong magnetic fields (~ 8 T) produced by superconducting dipole magnets cooled with liquid helium. Other magnets are used to steer the beams in the interaction points. Heavy ions instead are collected by stripping the electrons from a gas obtained by heating a small metal target before being accelerated to 4.5 MeV per nucleon by the Linac 3. The ions then travel through the *Low Energy Ion Ring* (LEIR) where they are grouped into bunches and accelerated from 4.2 MeV to 72 MeV. From there, the ions are injected into the PS and will follow the same path as the protons.

Inside the LHC ring, the bunches circulate in vacuum-filled pipes in opposite directions until they collide. Collisions occur at specifically designed interaction points along the LHC

¹Proton and Pb ions top energy are different due to the *rigidity* (R) parameter:

$$R = \frac{p}{q} = r_{\text{LHC}} B_{\text{LHC}},$$

where p is the momentum, q is the charge, r_{LHC} is the radius of the LHC ring, and B_{LHC} is the magnetic field of the LHC ring ($B_{\text{LHC}} = 8.33$ T). As the rigidity is a constant parameter and the charge of the proton is $q = 1$, the proton rigidity is $R = p$. On the other hand, the charge of the Pb ion is $q = 82$, so the Pb ion rigidity is $R = \frac{p}{82}$. The Pb ion rigidity is therefore 82 times smaller than the proton rigidity.

ring. The four main interaction points coincide with the location of the four major LHC experiments:

- ALICE (A Large Ion Collider Experiment). ALICE is a general purpose² experiment designed to study the properties of the QGP created in heavy-ion collisions. The ALICE experiment and its purpose are described in detail in Section 2.2.
- ATLAS (A Toroidal LHC ApparatuS) and CMS (Compact Muon Solenoid). ATLAS and CMS experiments are general purpose experiments designed to study the Higgs boson, the production of the top quark, and to investigate physics *Beyond the Standard Model* (BSM).
- LHCb (LHC-beauty). LHCb was conceived to study the physics of beauty quarks and *CP* violation.

2.2 The ALICE experiment

Among the major experiments at the LHC, ALICE [157] is rather unique both in terms of physics goals and in terms of detector design. ALICE is the only experiment at the LHC that is mainly dedicated to the study of ultrarelativistic heavy-ion collisions. In head-on collisions of heavy ions at LHC energy, the strongly interacting matter created in the collision is heated up to temperatures 10^5 times higher than the temperature of the core of the Sun. The hadronic matter melts into its constituent quarks and gluons, eventually forming a QGP state (refer to Chapter 1). The aim of the ALICE experiment is to study the properties of the QGP. Indeed, to indirectly explore the medium properties (i.e. temperature, energy density, viscosity, etc.), the production yields of different particle species and their momentum and angular distributions are measured. This exploration helps scientists to understand the characteristics of QGP more effectively.

Firstly conceived in 1989 as a heavy-ion experiment to be installed at the LHC, the ALICE setup was designed to effectively embrace and address emerging and exciting physics prospects. To achieve this goal, the ALICE detector was designed to be able to perform measurements of charged particles down to very low momentum (about 0.1 GeV/c), and in presence of a very high multiplicity of particles such as that expected in central Pb–Pb collisions where the impact of the deconfined phase is expected to be more evident. The

¹General purpose experiments are able to measure several signals of potential interest, even if their relevance may only become apparent later during the data taking.

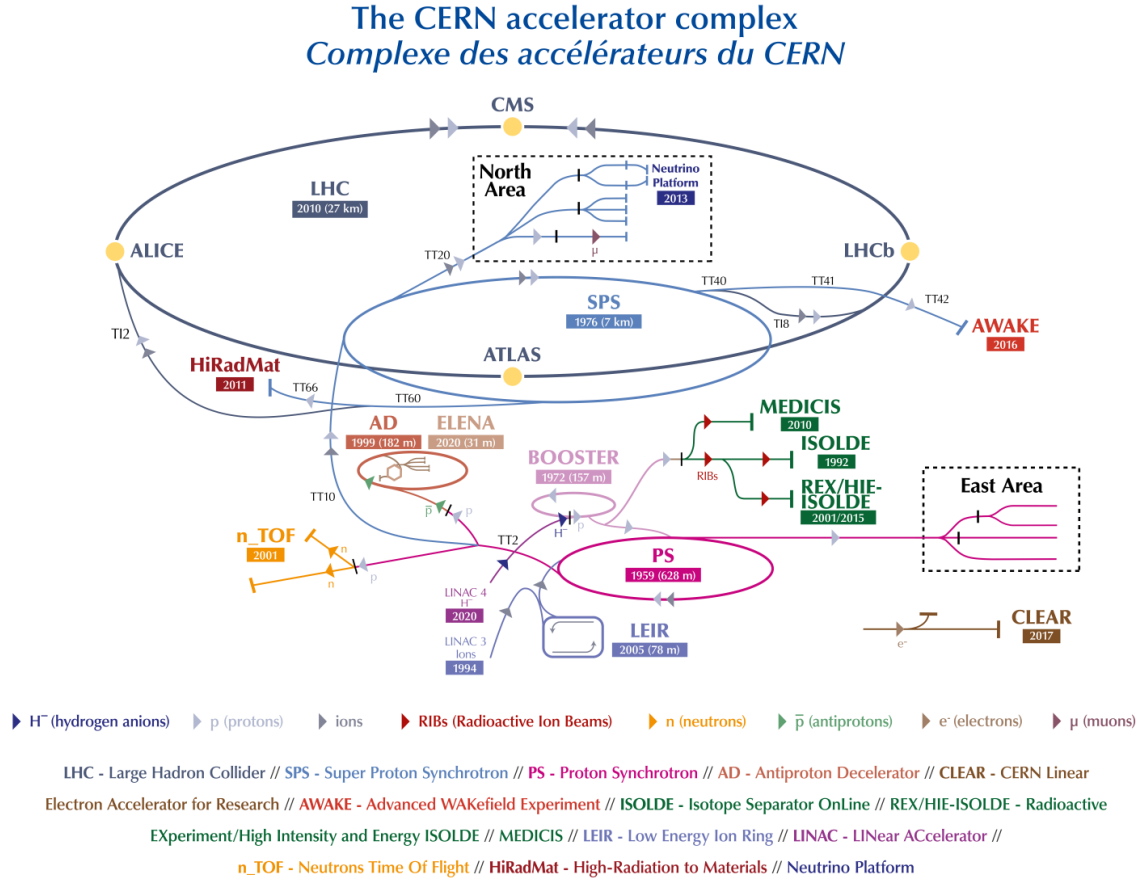


Fig. 2.1 The LHC injection scheme. Figure from Ref. [156].

initially foreseen charged particle multiplicity (dN_{ch}/dy) at mid-rapidity in central Pb–Pb collisions was between $dN_{\text{ch}}/dy = 2000$ and $dN_{\text{ch}}/dy = 8000$. In addition, ALICE was designed to have excellent *particle identification* (PID) capabilities with a high efficiency in a wide momentum range up to $20 \text{ GeV}/c$ (refer to Sec. 2.6).

The ALICE apparatus is made of a large detector of about $16 \times 16 \times 26 \text{ m}^3$ and $10,000 \text{ t}$, which is located at the LHC interaction point 2 (IP2). The ALICE experiment is composed of several sub-detectors, each of them designed to perform a specific task. Two main groups of sub-detectors can be identified according to their position inside the apparatus: the *central barrel* and the *forward muon spectrometer*. The central barrel includes the sub-detectors positioned around the nominal interaction point and that cover the pseudorapidity range $|\eta| < 0.9$. Starting from the beam axis and moving outwards, the sub-detectors of the central barrel are the *Inner Tracking System* (ITS, see Sec. 2.2.1), which is the silicon tracker of ALICE; the *Time Projection Chamber* (TPC, see Sec. 2.2.2) which provides tracking and

particle identification information; the *Transition Radiation Detector* (TRD) that is used for electron identification and trigger; the *Time Of Flight* (TOF, see Sec. 2.2.3) that provides PID information; the *High Momentum Particle Identification Detector* (HMPID), which is dedicated to the identification of charged particles; the *Electromagnetic Calorimeter* (EMCal), the *Photon Spectrometer* (PHOS), and the *Di-jet Calorimeter* (DCal), which are dedicated to the measurement of energy deposition of photons and electrons. The L3 solenoid magnet encloses the central barrel and provides a magnetic field of 0.5 T along the beam axis. The ITS, TPC, TRD, and TOF cover the full azimuthal angle (2π), while the other detectors of the central barrel have a smaller geometrical acceptance. On top of it, an array of scintillators (ACORDE) is used to trigger atmospheric muons, which are used in cosmic-ray studies and alignment purposes. The *forward muon spectrometer* is devoted to the tracking and identification of muons in the pseudorapidity range $-4 < \eta < -2.5$. It is composed of a set of absorbers and tracking stations and a dipole magnet. Additionally, a set of small detectors (*Zero Degree Calorimeter* (ZDC), *Photon Multiplicity Detector* (PMD), *Forward Muon Detector* (FMD), T0, V0A and V0C) focusing on global event characterisation and trigger are located at forward and backward rapidity.

A schematic lateral and frontal view of ALICE is shown in Figure 2.2. The convention for the coordinate system in ALICE is a right-handed Cartesian system with the z-axis along the beam pipe pointing towards the ATLAS experiment, the x-axis pointing towards the center of the LHC ring, and the y-axis pointing upwards. The azimuthal angle ϕ is measured in the transverse plane (xy) from the x-axis, and the polar angle θ is measured from the z-axis.

2.2.1 The Inner Tracking System (ITS)

The *Inner Tracking System* (ITS) [159, 160] is a cylindrical-shaped tracking detector. Figure 2.3 displays the layout of the ITS. It encircles the interaction region with its six layers of silicon detectors spanning a radius ranging from 3.9 cm to 43.0 cm, and covering the pseudo-rapidity range $|\eta| < 0.9$. The layers are characterised by different technologies: the *Silicon Pixel Detectors* (SPD), the *Silicon Drift Detectors* (SDD), and the *Silicon Strip Detector* (SSD). The main properties of each layer are summarised in Table 2.1.

Silicon Pixel Detector

The *Silicon Pixel Detectors* (SPD) compose the two innermost layers of the ITS. The SPD sensors are based on a matrix of 256×160 pixels (corresponding to a sensor area of 12.8 mm

THE ALICE DETECTOR

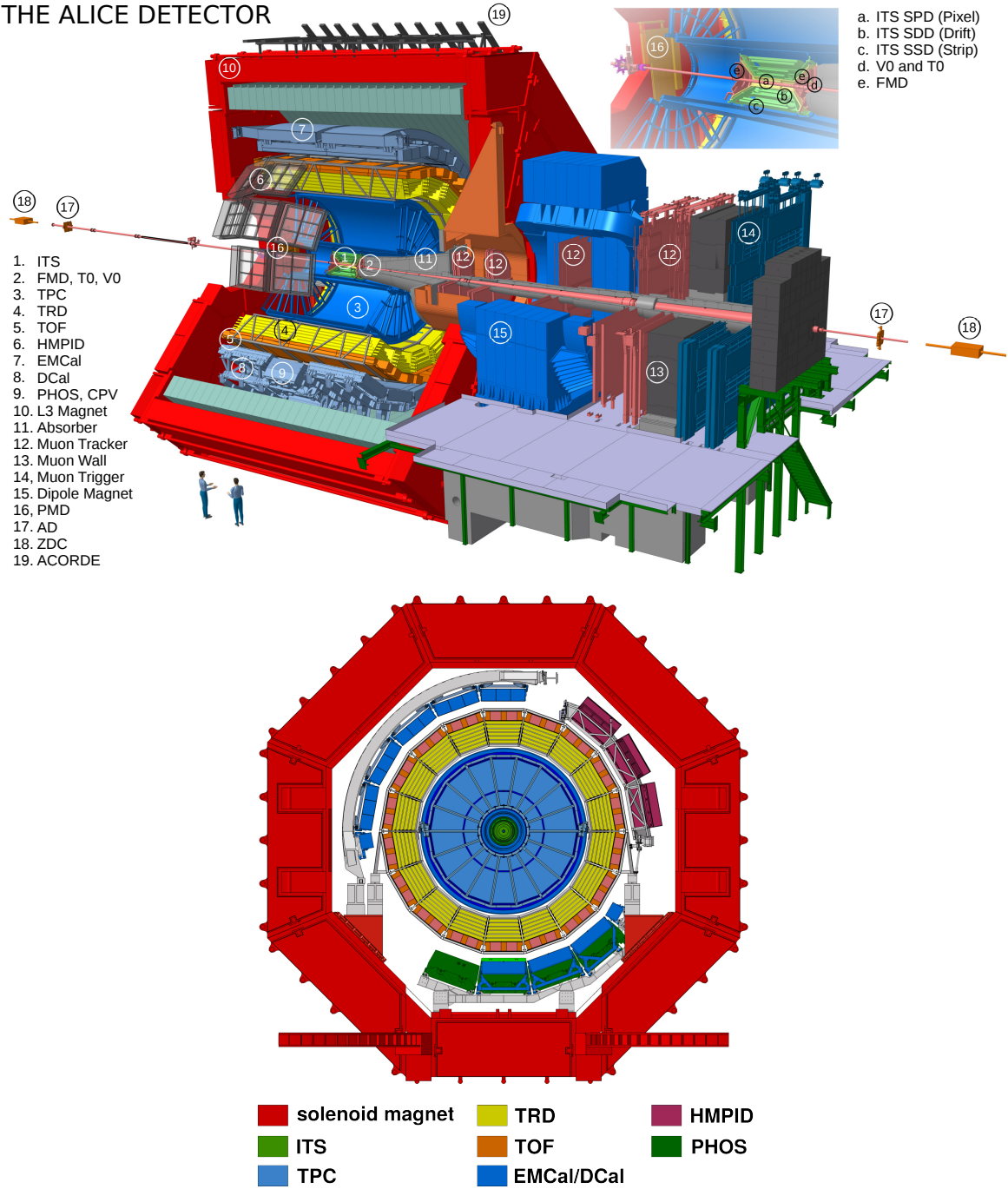


Fig. 2.2 Top: Lateral view of the ALICE experiment. Figure from the ALICE figure repository ©. Bottom: Frontal view of the ALICE experiment. Figure from Ref. [158]

$(r\phi) \times 70.7 \text{ mm } (z)$) arranged in a five-cell two-dimensional sensor with pixel size of $50 \text{ } \mu\text{m } (r\phi) \times 425 \text{ } \mu\text{m } (z)$. Two matrices are joined together along the longitudinal direction to create a *half-stave* that has a length of 141.6 mm. When two of these half-staves are mounted

head-to-head along the z direction on a carbon-fibre support, they form a *stave*. The SPD detector includes ten azimuthal sectors, each with two staves in the innermost layer and four staves in the outermost layer. The two layers have a distance of about 3.9 cm and 7.6 cm from the beam axis, with pseudorapidity acceptance of $|\eta| < 2.0$ and $|\eta| < 1.4$, respectively. The assembly of staves on sectors is designed for complete coverage in the azimuthal angle. The pixel matrix has a design spatial resolution of $\sim 12 \mu\text{m}$ in the radial direction and $\sim 100 \mu\text{m}$ along z . The SPD detector plays a crucial role in producing high-precision points for charged particle tracking close to the interaction point, which allow for an accurate determination of the position of the point of interaction (*primary vertex*) between the two incoming beams, as well as of decay vertices (*secondary vertices*). It is also crucial for a precise measurement of the track *impact parameter*, defined as the distance of closest approach between the track trajectory and primary vertex, which is fundamental for identifying secondary tracks originating from weak decays of strange, charm, and beauty particles. Additionally, due to its fast response time, it can provide a fast trigger signal.

Silicon Drift Detector

The third and fourth layers of ITS are composed of *Silicon Drift Detectors* (SDD). The primary unit of the SDD is a module divided into two drift regions by a central anode strip which produces a drift field of $\simeq 500 \text{ V/cm}$. The electrons, produced via ionisation by the interaction with a charged particle traversing the module, drift towards the anodes accommodated on both sides of the module. The SDD detector excels in its multi-track capability and offers two out of the four $\text{rdE}/\text{d}x$ samples required for the ITS PID by exploiting the total charge collected by the anodes as a proxy of the energy deposited in the detector by the crossing particle. The SDD sensors are $300 \mu\text{m}$ thick and have a sensitive area of $70.17 \text{ mm} (r\phi) \times 75.26 \text{ mm} (z)$. By measuring the drift time of the electrons produced by the crossing particles, with respect to the trigger time, it is possible to obtain the $r\phi$ coordinate, while the z coordinate is determined from the anodes collecting the drifting electrons. These modules are mounted on linear structure called *ladders*: the inner layer, placed at a radial distance of 15 cm from the beam pipe is made of 14 ladders with six modules each, while the outer one, located at 24 cm from the beam pipe, is made of 22 ladders with 8 modules each.

Silicon Strip Detector

The outermost two layers of the ITS are composed of *Silicon Strip Detector* (SSD). The SSDs are made of double-sided strip sensors where the bottom and top sides correspond to the p- and n-junctions of a silicon diode, respectively. Each sensor has 768 strips on each side with a pitch of 95 μm and a stereo angle of 35 mrad between the two sides, which is a compromise between the best resolution on $r\phi$ direction and the minimisation of ambiguities derived from the high particle densities expected in Pb–Pb collisions. The modules are arranged on ladders: the inner layer (Layer 5) contains 34 ladders, each featuring a linear array of 22 modules along the beam direction, while the outer layer (Layer 6) has 38 ladders made up of 25 modules each. To ensure complete coverage in η , the modules on the ladders have small overlaps and are staggered by 600 μm in the radial direction. Each neighboring ladder is installed at one of two slightly different radii, allowing for full azimuth coverage. The spatial resolution in the $r\phi$ direction is of $\sim 20 \mu\text{m}$ and of $\sim 830 \mu\text{m}$ in the direction along the beam.

2.2.2 The Time Projection Chamber (TPC)

The main tracking detector of the ALICE experiment is the *Time Projection Chamber* (TPC). Its purpose is to reconstruct the trajectories of charged particles measuring their momenta and identify them. The TPC has full azimuthal acceptance, covering a pseudorapidity range of $|\eta| < 0.9$ for particles that can reach the TOF detector, and $|\eta| < 1.5$ for slower particles. As shown in Fig. 2.4, the TPC has a cylindrical shape with an internal radius of about 85 cm and an external radius of 247 cm. It spans a length of 510 cm along the beam direction, making it as today the largest TPC on any experimental apparatus. The total active volume of the TPC is 88 m³. The TPC is divided into two drift regions by a

Table 2.1 Main properties of the ITS layers [159].

Layer	Technology	r (cm)	$\pm z$ (cm)	Material budget ($\%X_0$)	Spatial resolution $r\phi \times z (\mu\text{m}^2)$
1	SPD	3.9	14.1	1.14	12×100
2	SPD	7.6	14.1	1.14	12×100
3	SDD	15.0	22.2	1.13	35×25
4	SDD	23.9	29.7	1.26	35×25
5	SSD	38.0	43.1	0.83	20×830
6	SSD	43.0	48.9	0.86	20×830

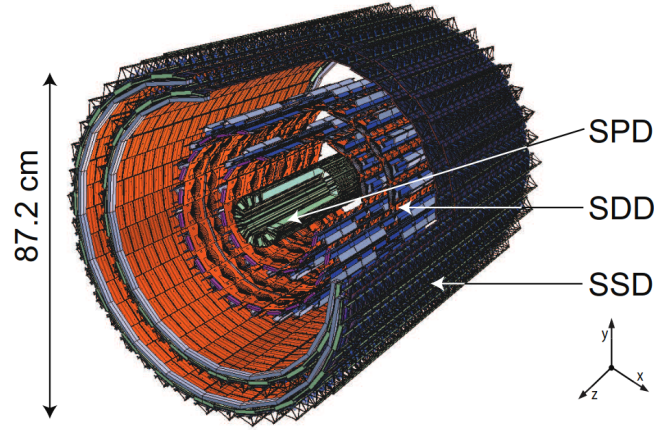


Fig. 2.3 Layout of the ALICE ITS. Figure from Ref. [160].

central cathode kept at a high negative potential $V \simeq 100$ kV. It generates a constant electric field $E \simeq 400$ V/cm with the help of an external field cage. The internal volume is split into two halves filled with a Ne (or Ar)/CO₂/N₂ mixture (90/10/5) with a maximum drift time for electrons of 90 μ s. Proper thermal stability ($\Delta T \leq 0.1$ K) is necessary because this mixture exhibits strong velocity dependence on temperature. Electrons resulting from gas ionisation caused by traversing charged particles undergo drift motion towards the cylinder end caps segmented into 18 trapezoidal sectors, each divided into an *inner* and *outer chamber* region equipped with a Multi-Wire Proportional Chamber (MWPC). The right panel of Fig. 2.4 shows a schematic view of one of the sectors of the TPC. The two MWPC together count 159 readout pad rows (63 rows of which are in the inner region), where cathodes collect the avalanche produced by primary electrons. This allows the reconstruction of the x-y projection of the particle trajectory. In addition, the drift time provides the spatial information along the beam direction measured using a buffered collected charge with a fixed frequency. The TPC detector ensures tridimensional charged particle track reconstruction with 1% transverse momentum resolution for low-momentum particles ($p_T \lesssim 1$ GeV/c) with a magnetic field of 0.5 T, growing up to $\sim 3.5\%$ for $p_T \sim 100$ GeV/c. Besides tracking, the TPC identifies charged particles by measuring their specific detector energy loss per unit length (dE/dx) while passing through the gas using the amplitude of the MWPC signal. Additional information on the PID capabilities of the TPC are reported in Sec. 2.6.

¹The gas mixture was changed during Run 2 data taking moving from Ne to Ar (2016 and 2018) mixture. The gas was changed back to the Neon mixture in 2017 due to large phase-charge distortions observed with the Argon mixture in TPC

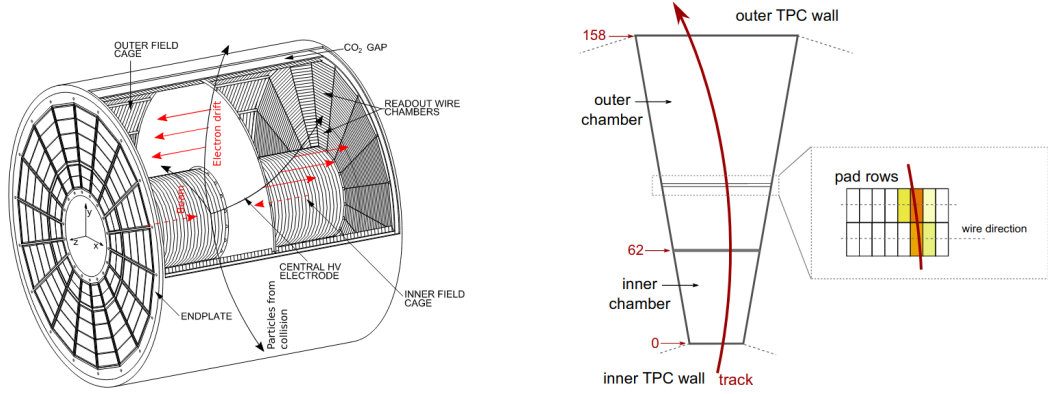


Fig. 2.4 Left: Layout of the ALICE TPC. Figure from Ref. [161]. Right: Azimuthal section of TPC detector. Figure from Ref. [162].

2.2.3 The Time of Flight (TOF)

The *Time of Flight* (TOF) [163] detector, like other central barrel detectors, is a cylindrical detector with a radius of 370 – 399 cm covering the full azimuth and the pseudorapidity interval $|\eta| < 0.9$. The TOF is made up of *Multi-gap Resistive-Plate Chambers* (MRPCs) arranged in a double-stack configuration. Each MRPC strip has an active area of $120 \times 7.4 \text{ cm}^2$ and an intrinsic timing resolution of approximately $\sim 40 \text{ ps}$. MRPCs are composed of two cathodes and a central anode, the space between them is filled with ten glass plates stacked with a $250 \text{ }\mu\text{m}$ gap and a non-flammable freon-rich gas mixture. An electric field of the order of a few kV/cm is maintained throughout the active volume, hence whenever a charged particle interacts with the gas in the active volume, producing ionisation charges, an electron avalanche is generated in each vacuum space between two plates. The formed avalanches are stopped by the plates, but the electrodes remain transparent to fast signals induced on them by the movement of ionisation electrons, resulting in a sum of signals generated in every vacuum space between two plates. Strips are oriented transversally to the beam direction and tilted to minimize signal sharing among adjacent pads. To cover the entire cylinder along the z -direction, five modules of varying lengths are deployed, superimposing junction regions with dead areas of other detectors. The central module is 117 cm long, the intermediate ones are 137 cm, and the external ones are 177 cm. Three modules from a linear structure called *supermodule*. The barrel comprises 18 supermodules arranged to form a cylindrical geometry. The TOF detector serves as a dedicated PID system complementary to the TPC identifying charged particles in the intermediate momentum range, i.e., up to 3 GeV/ c for pions and kaons, and up to 5 GeV/ c for protons. More information on the TOF

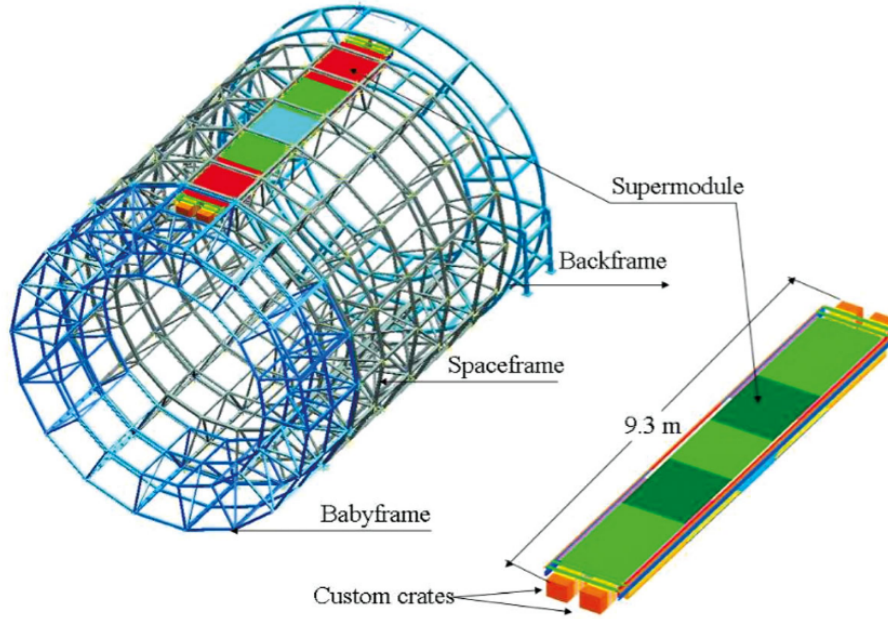


Fig. 2.5 Left: Layout of the ALICE TOF detector. Figure from Ref. [163].

PID principle of operation and performance can be found in Section 2.2.3. Figure 2.5 shows the layout of the TOF detector.

2.2.4 V0 detector

The ALICE V0 detector [164] is composed of two scintillator arrays, which are asymmetrically positioned with respect to the nominal interaction point position, the V0A and the V0C. Each V0 plate is made of 32 scintillator tiles arranged in four rings, each with eight tiles. Scintillators are made of plastic materials that emit light when they are excited by charged particles passing through them. The light is detected by *photomultiplier* tubes (PMTs), which convert light into an electrical signal. The V0 detector covers a wide pseudorapidity interval $-3.7 < \eta < -1.7$ (V0C) and $2.8 < \eta < 5.1$ (V0A) complementary to the detectors of the central barrel. The V0 detector serves multiple functions, such as providing minimum-bias (MB) and centrality triggers for the experiment. It provides timing information that is used to separate beam-beam interactions from background events like beam-gas interactions. In Pb–Pb collisions, the V0 detector is employed to determine the collision centrality and event plane. Figure 2.6 shows the layout of the V0 detector and Table 2.2 lists the main characteristics of the V0 detector.

Table 2.2 Geometric properties of V0A and V0C [164].

Ring	V0A			V0C		
	$\eta_{\min}/\eta_{\max}$	$r_{\min}/r_{\max}$ (cm)/(cm)	z (cm)	$\eta_{\min}/\eta_{\max}$	$r_{\min}/r_{\max}$ (cm)/(cm)	z (cm)
0	5.1/4.5	4.3/7.5	329	-3.7/-3.2	4.5/7.1	-86
1	4.5/3.9	7.7/13.7	329	-3.2/-2.7	7.3/11.7	-87
2	3.9/3.4	13.9/22.6	329	-2.7/-2.2	11.9/19.3	-88
3	3.4/2.8	22.8/41.2	329	-2.2/-1.7	19.5/32.0	-88

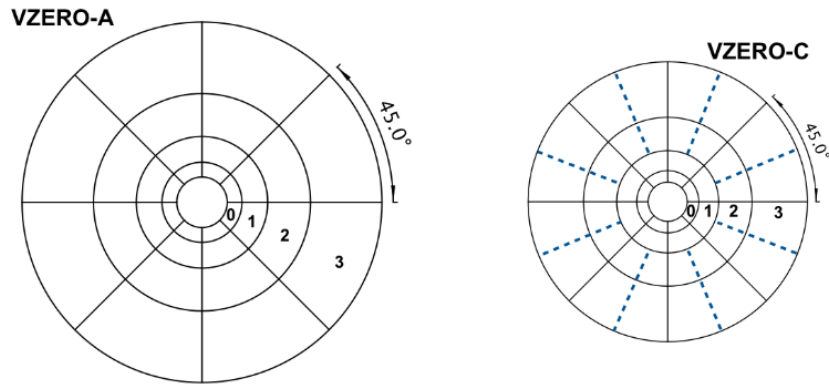


Fig. 2.6 Layout of the ALICE V0 detector plates. Figure from Ref. [157].

2.2.5 T0 detector

The *T0 detector* [164] is composed of two arrays of Cherenkov counters, each composed of a quartz radiator coupled with a PMT. These arrays are strategically located on opposite sides of the Interaction Point (IP), and they cover a pseudorapidity range of $-3.28 < \eta < -2.97$ (T0C) and $4.61 < \eta < 4.92$ (T0A), respectively. Although the coverage of the T0 detector is much narrower than the one of the V0 detector, it is significantly faster. Indeed, the primary purpose of the T0 detector is to provide fast timing signals to ensure that the TOF receives a collision time reference, necessary to perform PID as will be discussed in Sec. 2.6. The time resolution, estimated as the difference in time between TOA and TOC, is around ~ 40 ps and ~ 25 ps in pp and in Pb–Pb collisions, respectively. In addition to this, the T0 detector also provides a fast evaluation of the multiplicity of the collision with the possibility to generate MB and multiplicity triggers, similarly to the V0.

2.2.6 The Zero Degree Calorimeter (ZDC)

The ALICE *Zero Degree Calorimeter* (ZDC) [165] is a subdetector system incorporating six distinct calorimeters including two electromagnetic (ZEM), two proton (ZP), and two neutron (ZN) calorimeters. These sub-detectors are made of quartz fiber spaghetti calorimeters with fibers parallel to the beam axis which provide the necessary radiation hardness and the energy resolution of the order of a few percent. The ZP and ZN calorimeters are located at 112.5 m from the interaction point along the beam direction on either side of the IP. The ZNs are placed between the two beam pipes, while the ZPs are placed outside the outgoing beam pipe. The positioning of the neutron calorimeter is chosen such that neutrons emitted at forward rapidity, whose trajectory is not affected by the LHC magnetic field, will reach the ZN. The proton calorimeter is positioned such that positively charged particles (with different rigidity as compared to the beam particles) deflected by the LHC magnetic field will be detected. The main purpose of ZN and ZP is to veto the background events from beam-gas interactions and beam-halo particles. In addition, ZN and ZP are used to determine the centrality of the collision by measuring the energy deposited by spectator nucleons. The ZEM calorimeters are placed in proximity to the interaction point. These components are used to differentiate between central and peripheral Pb–Pb collisions by measuring the energy deposited by π^0 and photons emitted at forward rapidity [166]. The more central the collision, the larger will be the energy deposited in these ZEM calorimeters. This is crucial to resolve ambiguities in the centrality determination with ZP and ZN induced by the presence of nuclear fragments in the spectators.

2.3 The ALICE offline framework

Since its start in 2008, ALICE has accumulated more than 160 petabytes (PB) of raw data, a few terabytes of which have been used for the measurements presented in this Thesis. Such a gargantuan amount of data necessitates a robust and efficient framework to handle it effectively. This need gave rise to the *World LHC Computing Grid Project* (WLCG) or, as it is more commonly called, the *Grid*. Approved by the CERN Council in the early 2000s, WLCG is a global partnership comprising approximately 170 computing centres across over 40 nations, linking national and international grid infrastructures. The infrastructure accounts for 1.4 million computer cores and 1.5 exabytes (EB) of storage, making it one of the largest distributed computing infrastructures worldwide (WW). Over 12,000 scientists WW have

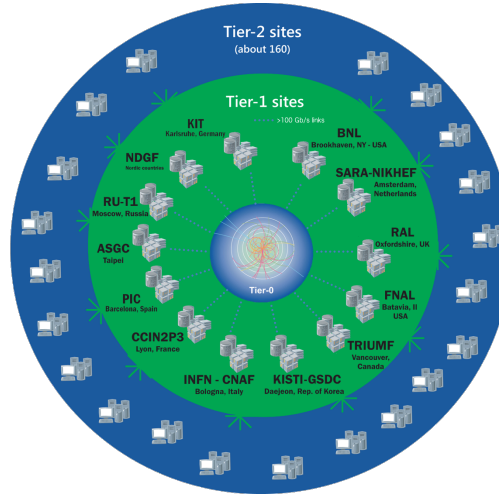


Fig. 2.7 Schematic representation of WLCG hierarchy. Figure from Ref. [168].

nearly real-time access to LHC data, performing 2 million analysis tasks per day using this powerful tool, regardless of their geographical location.

Besides the storage and the backup of the raw data, this infrastructure is instrumental in reconstructing and analysing the LHC data, which amount to almost 1 EB, and in processing Monte Carlo simulations. The WLCG is based on *Models of Networked Analysis at Regional Centres for the LHC Experiments* (MONARC) [167] and organised into three Tiers, following a hierarchical structure. The first Tier, known as Tier 0, is located at the CERN Computing Centre and at the computing centre at Wigner Research Centre for Physics in Budapest. Tier 0 serves as the primary entry point for all the data collected by the LHC experiments, including ALICE. The second Tier, known as Tier 1, consists of several centres spread around the globe (one of which is located within the Italian National Institute of Nuclear Physics (INFN) computing centre in Bologna) that are responsible for creating replicas of the raw and reconstructed data received from Tier 0. Replicas ensure the preservation of the information thus avoiding possible leaks. Additionally, the Tier 1 centres play a pivotal role in the reconstruction of the data. Finally, there are Tier 2 centres, which have small storage capabilities, but they provide essential computing resources for Monte Carlo simulations, and data analysis. The Italian community involved in the ALICE experiments coordinates four of the Tier-2 centres and one of these is located in the Torino INFN computing centre. Figure 2.7 shows a schematic representation of the WLCG hierarchy.

The analysis framework is as important as any detector for a HEP experiment. The large amount of data collected by ALICE in the last data-taking periods (Run 2) requires a complete set of tools in order to process and analyse the reconstructed events. The

ALICE Environment (AliEn) allows the analysers to access the data stored in the WLCG infrastructure, submit jobs to the computing centres, and monitor the status of the analyses and simulations. The information provided by the ALICE sub-detectors after the reconstruction of the raw data is stored in *Event Summary Data* (ESD) files. These files contain the reconstructed tracks, vertices, detector signals, and more information about the collision properties. The information contained in the ESDs are mainly intended for calibration and validation purposes, which are not needed for the final physics analysis. Therefore, the ESD files are filtered into *Analysis Oriented Data* (AOD) files, which are the final data structures used for the analysis. The data format of the AOD files is based on ROOT [169], which is a C++-based object-oriented framework for data analysis and storage. ROOT is a powerful tool that allows the user to store, manipulate, and ultimately handle and analyse data in a flexible way. The ALICE offline framework is based on ROOT, and consists of two main parts: *AliRoot* [170] and *AliPhysics* [171]. AliRoot provides a set of software classes and macros for the alignment, calibration, and reconstruction of the data collected by the experiment. It also allows performing Monte Carlo (MC) simulations by interfacing with event generators, such as PYTHIA 6 [93] and PYTHIA 8 [172] for pp collisions, and HIJING [173] for heavy-ion collisions. The transport of the particles through the detector is performed by GEANT3 [174] and GEANT4 [175]. With AliRoot and GEANT3/4 simulation a detailed description of the detector conditions (gain, dead channels, etc.), of the beam conditions, and their evolution with time is provided. On the other hand AliPhysics provides the tools for the analysis of the data, and it is constantly updated with new features by the analysers of the Collaboration.

2.4 The centrality determination in Pb–Pb collisions

The collisions of extended objects, such as heavy-ions, are classified according to their impact parameter b , which is the distance between the centres of the two colliding nuclei in the transverse plane. Even though this quantity is not directly accessible, according to the Glauber Model (refer to Sec. 1.3.2) b is linked to the charged particle multiplicity N_{ch} and the number of spectator nucleons, which are measurable. The centrality can be expressed as the fraction of the total hadronic cross section σ_{inel} corresponding to a given impact parameter b :

$$c(b) = \frac{\int_0^b \frac{d\sigma}{db'} db'}{\int_0^\infty \frac{d\sigma}{db'} db'} = \frac{1}{\sigma_{inel}} \int_0^b \frac{d\sigma}{db'} db', \quad (2.1)$$

where σ_{inel} is the total hadronic cross section. The integral starts from $b = 0$, i.e. head-on (central) collisions, which corresponds to 0% centrality, and ends at $b = \infty$, i.e. peripheral collisions, which corresponds to 100% centrality. In ALICE, the centrality is determined by measuring the energy deposited in the amplitude of the signal registered by the V0, which is proportional to N_{ch} . Alternatively, it can be determined by measuring the energy deposited in the ZDC (E_{ZDC}) and the ZED calorimeters.

Under the assumption that charged particle multiplicity N_{ch} (E_{ZDC}) increases (decreases) monotonically with the event centrality, Eq. 2.1 can be rewritten as:

$$c(b) = \frac{1}{\sigma_{\text{inel}}} \int_0^b \frac{d\sigma}{db'} db' \approx \frac{1}{\sigma_{\text{inel}}} \int_{N_{ch}}^{\infty} \frac{d\sigma}{dN'_{ch}} dN'_{ch} = \frac{1}{\sigma_{\text{inel}}} \int_0^{E_{ZDC}} \frac{d\sigma}{dE'_{ZDC}} dE'_{ZDC}, \quad (2.2)$$

In Eq. 2.2, σ_{inel} and $d\sigma$ can be replaced with the total number of collisions N_{ev} , corrected for the trigger efficiency and the non-hadronic interaction background, and the number of collisions within a certain event multiplicity/energy deposit, respectively.

The results presented in this Thesis are based on the centrality determination using the V0 detector. In particular, the percentile of the hadronic cross section is estimated by integrating the distribution of the sum of the measured V0 amplitudes (V0A and V0C) starting from the highest multiplicity events and integrating towards lower multiplicities down to a minimum V0 amplitude, which is called *anchor point*. This anchor point corresponds to 90% of the hadronic cross section and establishes the overall scale of centrality. The anchor is chosen to exclude events with very low V0 amplitudes, which are heavily contaminated by electromagnetic processes and suffer from trigger inefficiencies. Figure 2.8 shows the distribution of the sum of the signals registered by the V0 detectors of the ALICE experiment in Pb–Pb collisions at $\sqrt{s_{NN}} = 2.76$ TeV [166]. The V0-signal distribution is fitted with the function:

$$g(n) = P_{\mu,k}(n) \times [fN_{\text{part}} + (1-f)N_{\text{coll}}]. \quad (2.3)$$

The first term describes the particle multiplicity produced in a single nucleon-nucleon collision in terms of the *Negative Binomial Distribution* (NBD):

$$P_{\mu,k}(n) = \frac{\Gamma(n+k)}{\Gamma(n+1)\Gamma(k)} \times \frac{(\mu/k)^n}{(1+\mu/k)^{n+k}}, \quad (2.4)$$

which is the probability of a single emitting source, called *ancestor*, to produce n particles. In Eq. 2.4, μ and k are the mean and the dispersion of the distribution, respectively. The second term in Eq. 2.3 parametrises the number of ancestors. The linear combination expressed in Eq. 2.3 is inspired by two-component models [176, 177] that decompose the particle

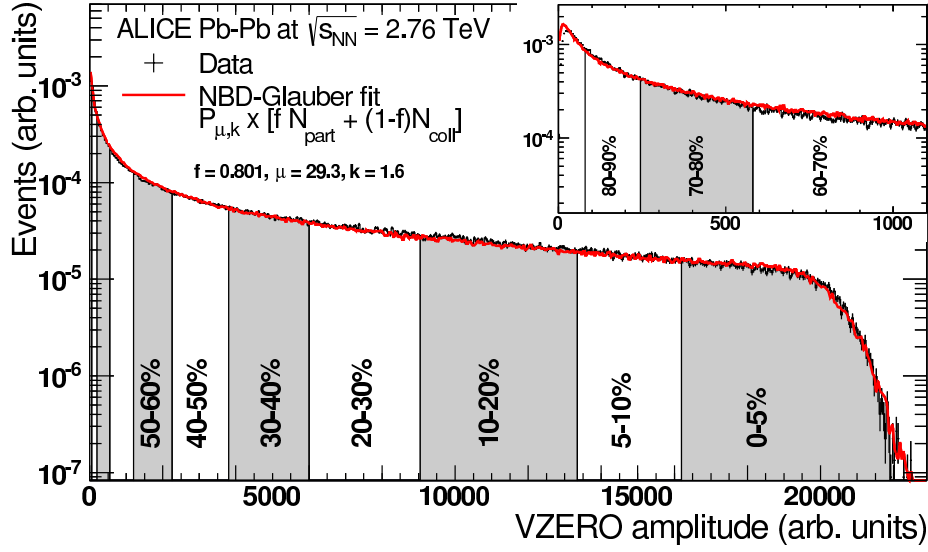


Fig. 2.8 Distribution of the sum of signals registered by the V0 scintillators of the ALICE experiment in Pb-Pb collisions at $\sqrt{s_{NN}} = 2.76$ TeV. Figure from Ref. [166].

produced in a nucleus-nucleus collisions in soft process, in which particle production is proportional to the number of participants N_{part} (refer to Sec. 1.4.6), and hard processes, which scale with the number of binary collisions N_{coll} (refer to Sec. 1.3.2). With this technique the resolution on the centrality determination ranges from 0.5% in central events to 2% in peripheral events [166].

2.5 Track and vertex reconstruction

The reconstruction of the tracks and primary vertices in ALICE follows a well-defined workflow [178]. First, the detector hits are converted into *clusters* (i.e. group of hits produced by the same crossing particle) containing information about position, signal amplitude, and timing. This is done separately for each detector. Next, a first determination of the interaction vertex position is carried out as the space point to which a maximum number of *tracklets*, i.e. track segments connecting pairs of clusters in the two SPD layers, converge.

The track finding and fitting process is then performed in three steps, following an *inward-outward-inward* scheme, as shown in Figure 2.9. In the TPC, the track search begins at a large radius. Initially, track seeds are formed using two TPC clusters in the outermost pad rows and the vertex point. Then, additional seeds are constructed using three clusters without considering the vertex constraint. These seeds are then iteratively propagated inward and updated at each step by incorporating the nearest cluster using the Kalman filter

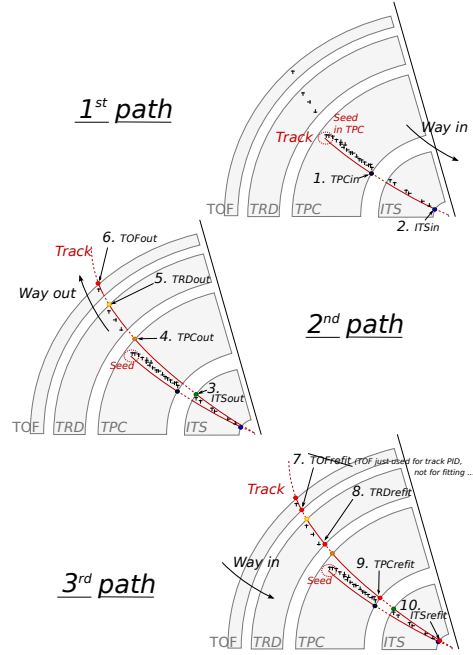


Fig. 2.9 Workflow of the track reconstruction in ALICE. Figure from Ref. [179].

technique [180]. As clusters can be shared among different seeds, the same physical track can be reconstructed multiple times. To address this, a specialised algorithm is employed to identify pairs of tracks that share a certain fraction of clusters (between 25% and 50%). The track with lower quality is subsequently rejected, ensuring each track is unique. The process continues until the reconstructed tracks reach the inner radius of the TPC, and only tracks with at least 20 clusters (out of a maximum of 159) are considered.

The reconstructed TPC tracks are then propagated inward to reach the outermost ITS layer and act as starting points for track reconstruction in the ITS. The track candidates are propagated inward and undergo updates at each ITS layer by considering nearby clusters that meet a proximity cutoff. Consequently, each TPC track generates a collection of track hypotheses within the ITS. This tracking step is performed in two stages, with and without a vertex constraint, as in the TPC. In the end, only high-quality track candidates are kept, according to the χ^2 resulting from the Kalman-filter procedure. The left and right panels of Figure 2.10 show the track-reconstruction efficiency in the TPC and the prolongation efficiency of tracks from the TPC to the ITS as a function of p_T in pp and Pb–Pb collisions (0–5% and 80–90% centrality classes) at $\sqrt{s} = 8$ TeV and $\sqrt{s_{NN}} = 2.76$ TeV, respectively. Besides a steep fall to zero for $p_T < 0.5$ GeV/c, due to modification of the helicoidal trajectory of the particle caused by the multiple scattering in the detector material, in the left panel of Fig. 2.10 the TPC tracking efficiency is almost constant as a function of p_T . The ITS

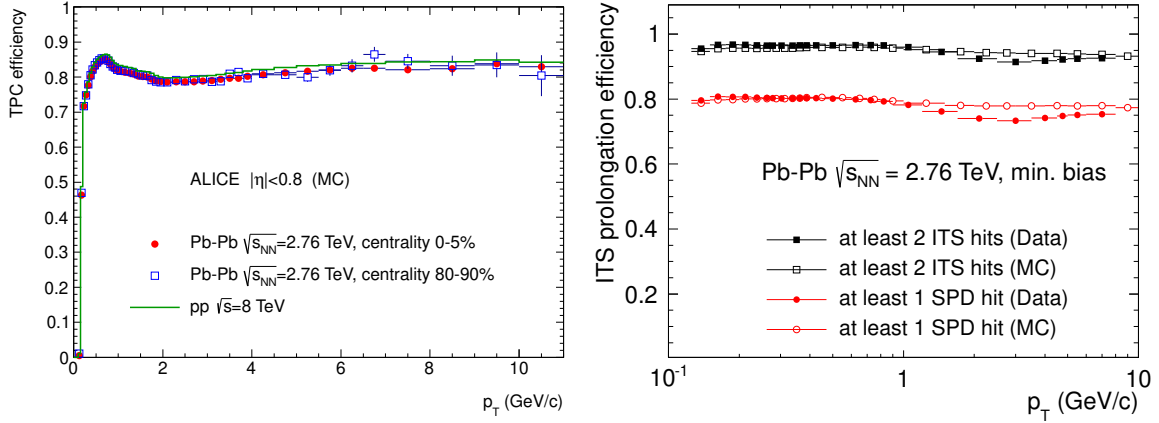


Fig. 2.10 Left: track-reconstruction efficiency in the TPC as a function of p_T in pp collisions at $\sqrt{s} = 8$ TeV. Figure from Ref. [178]. Right: prolongation efficiency of tracks from the TPC to the ITS as a function of p_T in Pb-Pb collisions at $\sqrt{s_{NN}} = 2.76$ TeV. Figure from Ref. [178].

matching efficiency depends on the requirement on the ITS points attached to the track mainly due to the presence of inactive modules in the different ITS layers. The TPC-ITS matching efficiency can be compared between data and MC simulations to study how the MC reproduce the tracking performance of the data.

In particular, in the low p_T region a good agreement is found between data and MC. At higher p_T the data points are slightly below the MC predictions. This is due to the fact that the MC simulation does not completely reproduce misalignment effects and the relative abundances of primary and secondary particles. At the analyses level, this discrepancy is addressed as described in Refs. [181–184].

After the reconstruction in the ITS is complete, all tracks are propagated to their closest point of approach to the preliminary interaction vertex. From there, the second matching step, which is an outward propagation, begins. The tracks are refitted in the outward direction using the Kalman filter and the clusters obtained from the previous stage. As the tracks reach the radii of the TRD and TOF detectors, an attempt is made to match them with the clusters in these detectors. There is also a matching process with the signals from the HMPID and calorimeters. Importantly, the propagation to detectors located beyond the TPC radius does not update the kinematic track parameters, but it is used to attach to the track the information needed for particle identification.

Finally, in the third step, all tracks are propagated from the outer radius of the TPC towards the interaction point. A final refitting is performed, which allows for the computation of the final track parameters and covariance matrix.

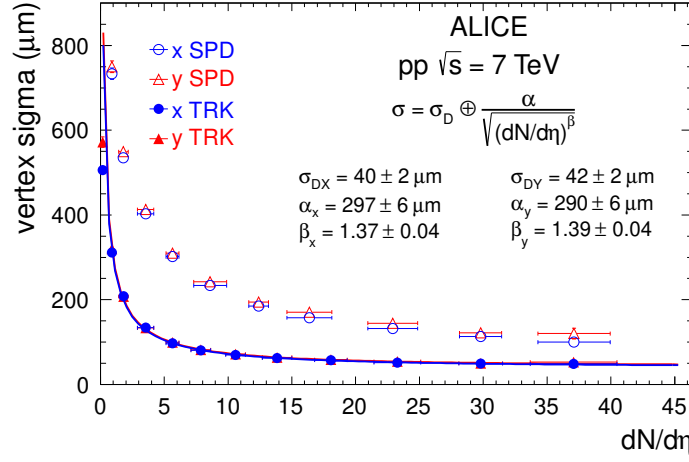


Fig. 2.11 Comparison of the width of the distribution of primary vertex positions reconstructed in the transverse plane from global tracks (named as TRK) and SPD tracklets (named as SPD) as a function of the charged-particle multiplicity per unit of pseudorapidity in pp collisions at $\sqrt{s} = 7$ TeV. Figure from Ref. [178].

After the tracking is completed, the position of the primary vertex is recalculated using the TPC+ITS tracks, also known as *global tracks*. Similar to the method used with SPD tracklets, global tracks are extrapolated to the point closest to the measured beam-line position, followed by vertex finding and fitting. The tracks are then propagated to the reconstructed vertex position. Figure 2.11 shows a comparison of the width of the distribution of primary vertex positions in the transverse plane reconstructed from global tracks and SPD tracklets as a function of the charged-particle multiplicity per unit of pseudorapidity in pp collisions at $\sqrt{s} = 7$ TeV [178]. The width of the distribution can be expressed in terms of the extension of the beam spot (σ_D), which mainly depends on the shape of the bunches, and the vertex resolution which depends on the number of tracks (tracklets) used to compute the vertex and is parametrised as $\alpha / \sqrt{(dN_{ch}/d\eta)^\beta}$. An improvement of the resolution by a factor of 2.5 is observed when using global tracks instead of SPD tracklets for the calculation of the primary vertex. The vertex resolution is around 200 μm for low-multiplicity ($dN/d\eta \approx 2$) pp collisions, 50 μm for high-multiplicity ($dN/d\eta \approx 30 - 40$) pp collisions, and 10 μm for Pb–Pb collisions [178].

The left panel of Fig. 2.12 displays the resolution of the projection in the transverse plane of the impact parameter to the primary vertex for identified ITS-TPC tracks in pp, p–Pb, and Pb–Pb collisions at $\sqrt{s} = 7$, $\sqrt{s_{NN}} = 5.02$, and $\sqrt{s_{NN}} = 2.76$ TeV, respectively. A clear improvement of the resolution is observed moving from small to large colliding systems thanks to the better resolution on the primary vertex position guaranteed by the higher multiplicities. The right panel of the same Figure shows the $1/p_T$ resolution

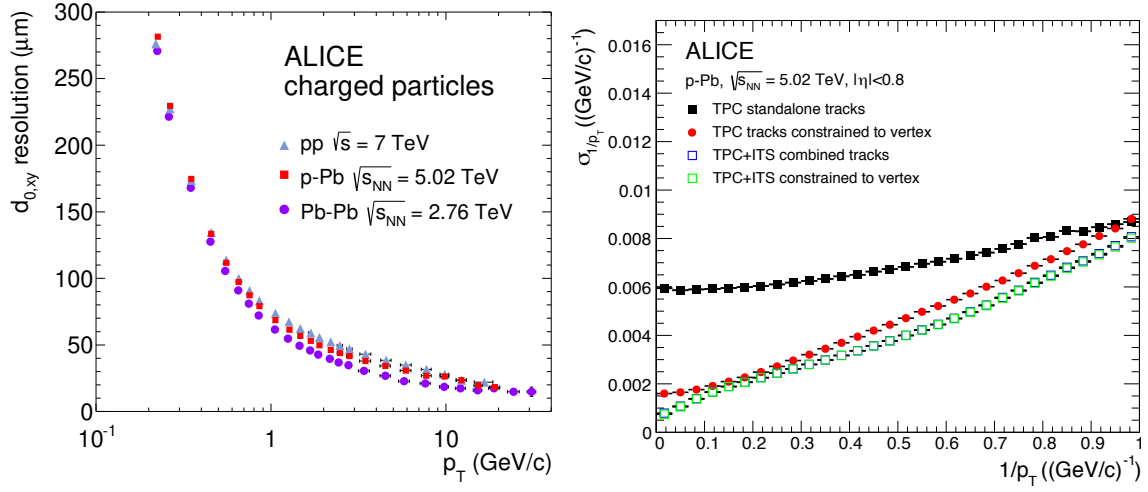


Fig. 2.12 Left: resolution of the projection in the transverse plane of the impact parameter to the primary vertex for identified ITS-TPC tracks in pp, p-Pb, and Pb-Pb collisions at $\sqrt{s} = 7$, $\sqrt{s_{NN}} = 5.02$, and $\sqrt{s_{NN}} = 2.76$ TeV, respectively. Right: $1/p_T$ resolution for TPC-standalone and ITS-TPC-combined tracks, together with the case of constrained primary vertex, in p-Pb collisions at $\sqrt{s_{NN}} = 5.02$ TeV. Both figures are taken from Ref. [178].

for TPC-standalone and global tracks, together with the case of tracks constrained primary vertex as described before, measured in p-Pb collisions at $\sqrt{s_{NN}} = 5.02$ TeV. This resolution is taken from the covariance matrix of the track, and it represents the error on the track curvature estimated from the Kalman filter procedure. The inverse- p_T resolution is linked to the relative p_T resolution via:

$$\frac{\sigma_{1/p_T}}{1/p_T} = \frac{\sigma_{p_T}}{p_T}. \quad (2.5)$$

The p_T resolution for global tracks is better than 1% at $p_T = 1$ GeV/c and degrades to about 2% at $p_T = 10$ GeV/c.

2.5.1 Secondary vertex reconstruction

The reconstruction of the decay vertices of the candidate D mesons is performed at the analysis level. The procedure is based on the minimisation of the distance D between the tracks forming the candidate D-meson decay topology and the secondary vertex. The tracks are approximated as straight lines in the vicinity of the primary vertex. The distance D is defined as:

$$D = \sqrt{\sum_{i=1}^{N_{trk}} \left[\left(\frac{x_i - x_0}{\sigma_{x_i}} \right)^2 + \left(\frac{y_i - y_0}{\sigma_{y_i}} \right)^2 + \left(\frac{z_i - z_0}{\sigma_{z_i}} \right)^2 \right]}, \quad (2.6)$$

where the coordinates of the *Point of Closest Approach* (PCA) of the track i to the secondary vertex are (x_i, y_i, z_i) , and the uncertainties on these coordinates are σ_{x_i} , σ_{y_i} , and σ_{z_i} . The coordinates of the secondary vertex are (x_0, y_0, z_0) . The uncertainties on the track coordinates are obtained from the covariance matrix of the track parameters. The secondary vertex is reconstructed by minimising the distance D with respect to the coordinates of the secondary vertex.

The momentum of the particle originating from the secondary vertex is calculated as the sum of the momenta of the tracks associated to the secondary vertex calculated at their respective PCAs to the secondary vertex. The uncertainty on the momentum of the particle is calculated as the sum in quadrature of the uncertainties on the momenta of the tracks. In the specific case of D_s^+ mesons, the resolution on the secondary vertex position is about $100 \mu\text{m}$ for x , y , and z coordinates slightly depending on the particle p_T .

2.6 Particle identification

The ALICE experiment features unique particle identification (PID) capabilities thanks to the combination of the information provided by the ITS, the TPC, and the TOF detectors. Each of these detectors provides a different PID technique, which is based on the measurement of the specific energy loss (dE/dx) in the ITS and in the TPC, and on the measurement of the velocity of the particle with the TOF detector. These PID techniques are complementary to each other, and they are used in combination to achieve the best possible PID performance over a wide range of particle momentum. In addition, the HMPID detector is used to extend the PID capabilities to high p_T by precisely measure the Cherenkov angle of the photons emitted by the particles traversing the detector. Electrons are identified also with the TRD detector, which exploits the radiation emitted by the particles traversing layers of different dielectric materials, and by measuring the energy released in the PHOS and EMCal calorimeters and comparing it with their momentum. The PID performance and the operational details of ITS, TPC, and TOF, which are employed in the analyses presented in this Thesis, are discussed in the following sections. In addition, the application at the analyses level of the PID information provided by these detectors is discussed in Chapter 5.

2.6.1 ITS

In the ITS the PID is based on the measurement of the specific ionisation energy-loss produced by the charged particles traversing the detector. The analogic read-out of the

SDD and SSD layers is able to provide a measurement of the collected charge. To ensure a good PID performance, only tracks with at least 3 clusters in the SDD and SSD layers are considered. The $\langle dE/dx \rangle$ is calculated as the truncated mean of the $\langle dE/dx \rangle$ values measured on each layer. The PID information provided by the ITS is extremely useful for low- p_T particles, which may not reach the TPC or the TOF. The dE/dx measured by the ITS as a function of the momentum in Pb–Pb collisions at $\sqrt{s_{NN}} = 5.02$ TeV is shown in the left panel of Fig. 2.13. A parametrisation of the expected specific energy-loss for electrons, pions, kaons, and protons are shown as solid lines, and a good separation between the pion and the kaon (proton) bands is observed for momenta below 500 (900) MeV/c.

2.6.2 TPC

The PID information provided by the TPC is crucial to identify particles in a wide momentum range. The PID in the TPC is based on the measurement of the dE/dx produced by the charged particles traversing the detector. The specific energy-loss is estimated as the truncated mean of the distribution of 65% lowest-amplitude signals for each track. This procedure is necessary to remove the contribution of the δ -rays, which are produced by the ionisation of the gas molecules by the primary ionisation electrons. Once the momentum p of the track is reconstructed and its dE/dx has been measured, the PID is performed by comparing the measured specific energy-loss with the expected one for each particle species at momentum p . The expected specific energy-loss, which is described by the Bethe-Bloch formula, is parametrised with a data-driven approach as:

$$f(\beta\gamma) = \frac{P_1}{\beta^{P_4}} \left[P_2 - \beta^{P_4} - \ln \left(P_3 + \frac{1}{\beta\gamma} \right) \right]. \quad (2.7)$$

The parameters $P_1 - P_5$ are free parameters, which are determined in the fitting procedure, β is the velocity of the particle, and $\gamma = \frac{1}{\sqrt{1-\beta^2}}$ is the Lorentz factor deduced from the momentum of the track $p = \beta\gamma m$. The PID performance of the TPC is shown in the right panel of Fig. 2.13 as a function of the momentum in Pb–Pb collisions at $\sqrt{s_{NN}} = 5.02$ TeV. The central values of the expected specific energy-loss for each particle species (i.e. electrons, pions, kaons, protons, deuterons, tritons, and ^3He) are shown as solid lines. The largest separation power is achieved for momenta $p < 0.8$ GeV/c for pions and kaons, and $p < 1.5$ GeV/c for kaons and protons. In the higher- p range, the separation power is reduced due to the saturation of the specific energy-loss in presence of the so-called *relativistic rise*. However, PID information can still be inferred via statistical methods, such as multi-Gaussian fits.

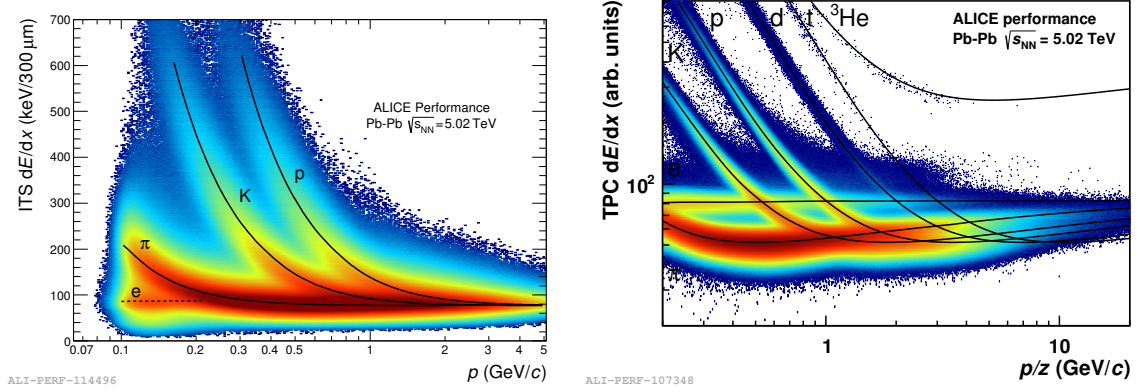


Fig. 2.13 Specific energy-loss measured by the ITS (left) and the TPC (right) as a function of the momentum (p) and rigidity (p/z) in Pb–Pb collisions at $\sqrt{s_{\text{NN}}} = 5.02$ TeV. The central values of the expected specific energy-loss for each particle species (i.e. electrons, pions, kaons, protons, deuterons, tritons, and ^3He) are shown as solid lines. Figures from the ALICE Figure repository ©.

2.6.3 TOF

Unlike the ITS and the TPC, the TOF detector, as its name suggests, does not rely on the measurement of the energy to determine the mass of a particle, but rather on its time-of-flight. Given a particle with momentum p and mass m , the time-of-flight is calculated as:

$$t = \frac{L}{\beta c} = \frac{LE}{pc^2} = \frac{L}{c} \sqrt{1 + \frac{m^2 c^4}{p^2 c^2}}, \quad (2.8)$$

where p is experimentally reconstructed following the procedure described in Sec. 2.5, L is the distance between the primary vertex and the TOF detector, and c is the speed of light. The PID is performed by comparing the measured time-of-flight with the expected one for each particle of momentum p . Hence, the measurement of the time-of-flight requires a starting time, which is provided by the T0 signal if present, or by a combinatorial analysis of the arrival times of the particles in the TOF detector. In case both the T0 and the TOF are present, the starting time is provided by the weighted average of the two measurements. The β parameter is calculated as $\beta = L/tc$, where L is the distance between the primary vertex and the TOF detector, and t is the time-of-flight measured by the TOF. The PID performance of the TOF is shown in Fig. 2.14 as a function of the momentum, measured by the TPC, in Pb–Pb collisions at $\sqrt{s_{\text{NN}}} = 5.02$ TeV. The TOF detector is able to provide a good PID performance in the momentum range $0.5 < p < 4$ GeV/c for pions, kaons, and protons. The separation power is reduced at higher momenta due to the saturation of the β parameter. However, the TOF detector is able to provide a good PID performance also at

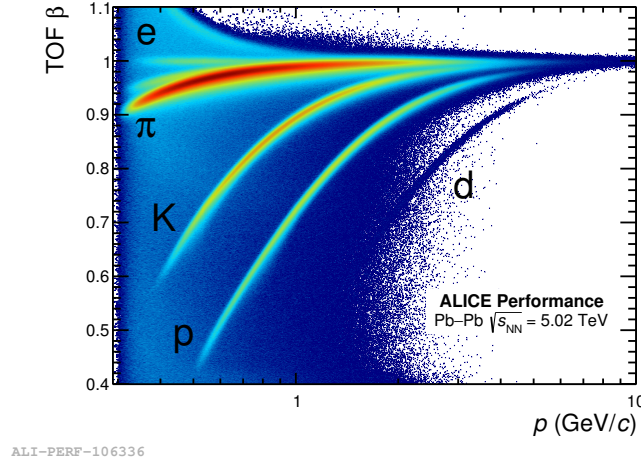


Fig. 2.14 The β parameter, measured by the TOF, as a function of p , measured by the TPC, in Pb–Pb collisions at $\sqrt{s_{NN}} = 5.02$ TeV. Figure from the ALICE Figure repository ©.

higher momenta thanks to the combination of the information provided by the TPC and the TOF detectors. Even though the distributions of the β parameter for pions, kaons, protons, and deuterons are clearly distinct, a huge background due to wrongly associated tracks is present. The overall TOF resolution, comprehensive of the contribution of electronics, calibrations, start-time uncertainty, and momentum resolution, is about 56 ps for pions with a momentum of 1.5 GeV/c in Pb–Pb collisions. The largest separation power is achieved for momenta $p < 3$ GeV/c for pions and kaons, and $p < 5$ GeV/c for kaons and protons.

Chapter 3

Machine Learning

In this Chapter, the basic concepts of *Machine Learning* (ML) are introduced. The Chapter is divided in three sections. In Sec. 3.1, the ML technique applied to the problem of *regression* and *classification* is introduced. In Sec. 3.2, the *Boosted Decision Tree* (BDT) algorithm, employed for the analyses presented later in the Thesis, and its main features are presented. Finally, Section 3.3 briefly presents the interface between the BDT algorithm and the ROOT framework in ALICE.

3.1 Machine Learning in a nutshell

Machine Learning (ML) represents a specialised branch of computer science that delves into the study of algorithm design capable of autonomous learning from data. The core objective of ML is to empower computers with the ability to learn automatically and adapt their actions accordingly, miming the human brain response. These algorithms find widespread application across diverse fields, encompassing computer vision, speech recognition, and search engines [185, 186].

Over the last decade, ML techniques have emerged as an advantageous tool in particle physics, driving remarkable advancements in this field. The performance achieved at the data analysis level with respect to the traditional linear selections, but also their capability to improve the performance of the detectors, have been key factors contributing to this success [187–190]. The ML techniques are particularly effective when dealing with either *classification* or *regression* problems, which are daily challenges faced by particle physicists. The classification problem is encountered in this Thesis to classify D_s^+ -meson candidates,

distinguishing between background and true physical signals. A quite common classification problem consists in classifying instances into two classes and serves as a fundamental demonstration of the working principles of ML algorithms. For this purpose, in this Chapter, an illustrative example of D_s^+ meson classification based on binary classification is presented. However, it is important to highlight that to perform the measurements, presented in this Thesis, a multiclass classification algorithm, capable of handling more than two classes, was utilised. The details of the analysis procedure for the multiclass classification algorithm are presented in Chapter 5. The regression problem regards the capability to predict a continuous numerical outcome based on one or more input variables. For these ML-based analyses, a *Boosted Decision Tree* (BDT) algorithm was employed, whose features are presented in Sec. 3.2.

Both the classification and regression problem fall in the category of *supervised learning* algorithms [191]. Supervised learning algorithms follow a structured process encompassing three distinct phases: *training*, *testing* and *validation*, and *application*. During the training phase, the algorithm is presented with a set of n examples called the *training set* (A), defined as:

$$A = \{(\vec{x}_1, y_1), (\vec{x}_2, y_2), \dots, (\vec{x}_n, y_n)\} . \quad (3.1)$$

Each example in A is composed of a set of *features* $\vec{x}_i$, which represent the input variables given to the algorithm (e.g. the physical information of the example), and the corresponding true output value y_i . In the case of classification, the output value consists of the class label, while for regression problems the output is the value of the function under study for that specific instance. Assuming the existence of a function f , stochastic or deterministic, that maps the features to the output value, the relation between the feature and the output value can be expressed as:

$$y = f(\vec{x}) . \quad (3.2)$$

The primary goal of the training phase is to approximate the mapping function f with an *objective function* to be used as a predictor in the following phases. The objective function can be defined as:

$$\text{obj}(\vec{\theta}) = L(\hat{y} - y, \vec{\theta}) + \Omega(\vec{\theta}) , \quad (3.3)$$

where $\vec{\theta}$ are the internal parameters, or *weights*, of the ML algorithm. The objective function is composed of two terms: the *loss function* $L(\hat{y} - y, \vec{\theta})$ and the *regularisation term* $\Omega(\vec{\theta})$. The loss function term measures the discrepancy between the predicted true value, $\hat{y}$, and the actual output value, y . Several loss functions are available (e.g. Mean Squared Error, Binary Cross-entropy, Huber class, etc.) [192] each of them with its own advantages and disadvantages. In order to better approximate f , the algorithm aims to minimise the loss

function by adjusting its weights by going through an iterative optimisation procedure. This operation can be implemented using different optimisation algorithms (e.g. Gradient Descent, Stochastic Gradient Descent, etc.), which are characterised by different convergence speeds and computational costs. The regularisation term controls the model complexity and prevents the algorithm from *overfitting* or *underfitting* the training set. The overfitting occurs when the algorithm *memorises* the training set due to excessive complexity, resulting in a poor generalisation to new examples. In other words, the algorithm is only capable of predicting the output value of the examples in A , but it fails in the case of unknown examples. On the other hand, underfitting occurs when the algorithm is not able to reproduce the training set because it is too simple, resulting in a poor performance on the training set itself. In this case, the algorithm is not capable of learning the feature-output value relation. In the best-case scenario, the algorithm learns some general patterns from the training set, which can be applied to new examples.

Once the training process is complete, the algorithm is able to make predictions for new examples based on its understanding of the feature-output value relation acquired from the training set. Before applying the trained algorithm, it is essential to assess its performance. This evaluation must be performed using a completely new and independent dataset, called *test set*. Using the same training set for assessment can lead to an inflated performance since the algorithm may be subjected to underfitting or overfitting, giving a misleading impression of its effectiveness. As for the training set, also in the test set the true output value y is known. This allows for an unbiased evaluation of the algorithm ability to generalise to unknown data. A detailed description of the method employed to assess the performance of the ML algorithm is reported in Sec. 3.1.1.

At this stage, an additional set of examples, called *validation set*, is used to optimise the *hyperparameters* of the algorithm. The hyperparameters are parameters that can not be optimised by the algorithms themselves during the training phase, but rather set by the analyser before the training process begins. The hyperparameter optimisation is discussed in Sec. 3.1.3 and some relevant examples of hyperparameters in the context of the analyses presented in this Thesis are reported in Sec. 3.2.1.

Finally, the algorithm is applied to the data set of interest, the so-called *application set*, where the true output value is unknown. In the case of binary classification, the output value will be a single numerical value (*score*) representing the probability of the instance belonging to the reference class. While, in the case of multiclass classification, the ML output will be a set of numerical values, one for each class, representing the probability of the instance belonging to a specific class. High output scores indicate a greater probability of the instance

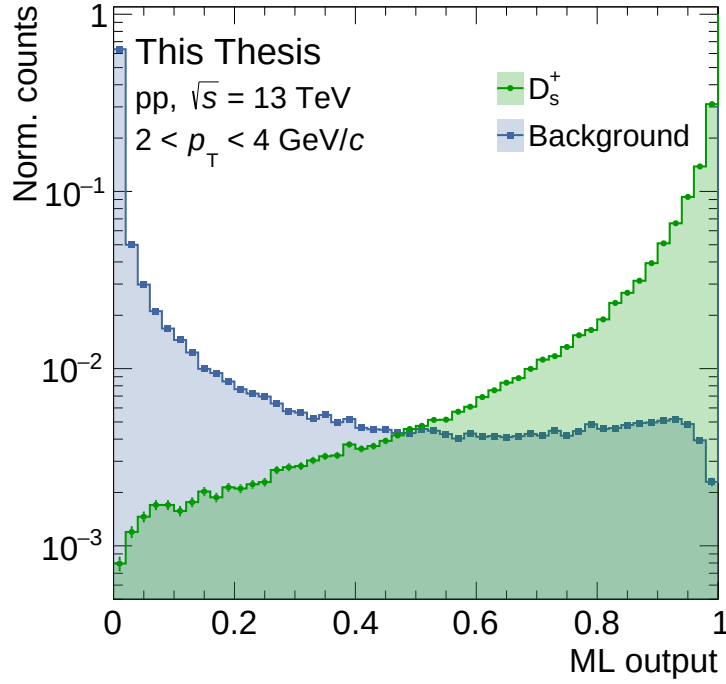


Fig. 3.1 Output score distribution for the signal (green) and background (blue) classes in the context of the D_s^+ meson production in pp collisions at midrapidity at $\sqrt{s} = 13$ TeV. The $2 < p_T < 4$ GeV/c interval is considered.

belonging to that class. In order to categorise instances as belonging to a particular class, a threshold value must be established. Instances with a score below or over this threshold will be rejected or accepted depending on the type of the selection. The optimisation of this threshold value in the context of the D_s^+ meson analyses is discussed in Sec. 5.1.3 and Sec. 6.1.2 for pp and Pb–Pb analyses, respectively. In the case of regression instead, the output value will directly be the predicted value of the function under study for that instance.

Figure 3.1 shows the output score distribution in the case of a binary classification problem: the selection of D_s^+ -meson candidates from the sample of events collected by the ALICE experiment at midrapidity in pp collisions at $\sqrt{s} = 13$ TeV in the $2 < p_T < 4$ GeV/c interval. The ML score distribution for the signal class is shown in green and, as expected, is peaked at higher values of the score compared to the background class, shown in blue. The background class is peaked around zero, indicating that the training-testing-validation process was successful, and the algorithm is now capable of distinguishing between the two classes.

Table 3.1 Confusion matrix for a binary classification problem.

		Predicted Class	
		<i>Positive</i>	<i>Negative</i>
Actual Class	<i>Positive</i>	TP	FP
	<i>Negative</i>	FN	TN

3.1.1 Receiver Operating Characteristic

A numerical metric is usually employed to assess the reliability and prognostic capabilities of the ML algorithm. Even though several metrics are suited to this purpose, the most common metric for binary classification problems is the *Area Under the Receiver Operating Characteristic curve* (ROC AUC) [193]. This metric is particularly useful in the case of imbalanced datasets, where the number of instances belonging to the positive class (P) is much smaller than the number of instances belonging to the negative class (N). This condition is often encountered in particle physics analyses, where the number of signal events is much smaller than the number of background events. When considering a set of examples from the positive class, instances correctly classified by the algorithm as members of P are considered as *true positive* (TP), while those misclassified are referred to as *false negatives* (FN). On the other hand, instances truly belonging to the negative class but classified as positive are *false positives* (FP), and those correctly classified as negative are *true negatives* (TN). Table 3.1 exemplifies the resulting *confusion matrix*.

The ROC curve is a visual representation of the trade-off between the *true positive rate* (TPR) and the *false positive rate* (FPR) for different threshold values. The TPR and FPR can be estimated from the elements of the confusion matrix. The TPR is defined as the ratio between the number of true positives and the sum of true positives and false negatives:

$$\text{TPR} = \frac{\text{TP}}{\text{TP} + \text{FN}}. \quad (3.4)$$

The TPR is also known as *recall* or *sensitivity* and, in case of a binary classification problem, indicates the *First class efficiency*. The FPR, instead, is defined as the ratio between the number of false positives and the sum of false positives and true negatives:

$$\text{FPR} = \frac{\text{FP}}{\text{FP} + \text{TN}}. \quad (3.5)$$

The FPR is also known as *fall-out* or *false alarm rate* and indicates the *Second class efficiency*.

The ROC AUC is then defined as the area under the ROC curve. The ROC AUC is a metric that ranges between 0 and 1 by definition, where a value of 0 indicates poor performance, while a value of 1 indicates perfect performance. The left panel of Fig. 3.2 shows the comparison of the ROC curve of the binary-classification algorithm employed in the study of the production of the D_s^+ meson in pp collisions at $\sqrt{s} = 13$ TeV with the ROC curve of a random classifier. The random classifier assigns a random class to each instance, hence its ROC curve corresponds to the bisector of the first quadrant, which is a straight line with a ROC AUC of 0.5. Instead, the ROC curve of the binary classification algorithm is above the ROC curve of the random classifier, indicating a better performance. The ROC AUC obtained with the trained and tested algorithm is around 0.95. A small decrease of the ROC curve between the training, reported in red, and the test set, reported in blue, indicates a slight worsening of the performance. However, the difference between the two values is small, suggesting that the algorithm is not overfitting the training set, otherwise a significantly lower ROC curve would be found for the test, and is capable of generalising to new examples.

The ROC AUC is natively defined for binary classification problems. However, it can be extended to multiclass classification problems [194]. Two approaches are possible. The first approach is to compute the ROC AUC for each class separately, considering the class of interest as the positive class and the remaining classes as the negative class (*One-vs-Rest* approach). The second approach is to compute the ROC AUC for each pair of classes, considering one class i as the positive class and the other class j as the negative class (*One-vs-One* approach). Typically, the $AUC_{i,j}$ does not coincide with the $AUC_{j,i}$, since a different output score distribution is considered to compute the TPR and the FPR. The final ROC AUC is then obtained by averaging the ROC AUCs obtained for each class. Unlike the *One-vs-Rest* approach, the *One-vs-One* approach does not suffer from the class imbalance problem, hence it is the chosen approach for the multiclass classification analyses presented in this Thesis. The right panel of Fig. 3.2 shows the ROC curve of the multiclass classification algorithm employed in the study of the production of the D_s^+ meson in pp collisions at $\sqrt{s} = 13$ TeV.

3.1.2 Cross Validation

As aforementioned, the validation of the algorithm performance is a crucial step in the ML workflow. The validation process is typically performed using a dedicated validation set, which is independent of the training set. However, in the case of small datasets, there are not enough examples to split the dataset into training, validation, and test sets, or the validation set may not be representative of the full dataset. In this case, the validation process may lead to an overestimation of the algorithm performance. Also, in the case of small datasets,

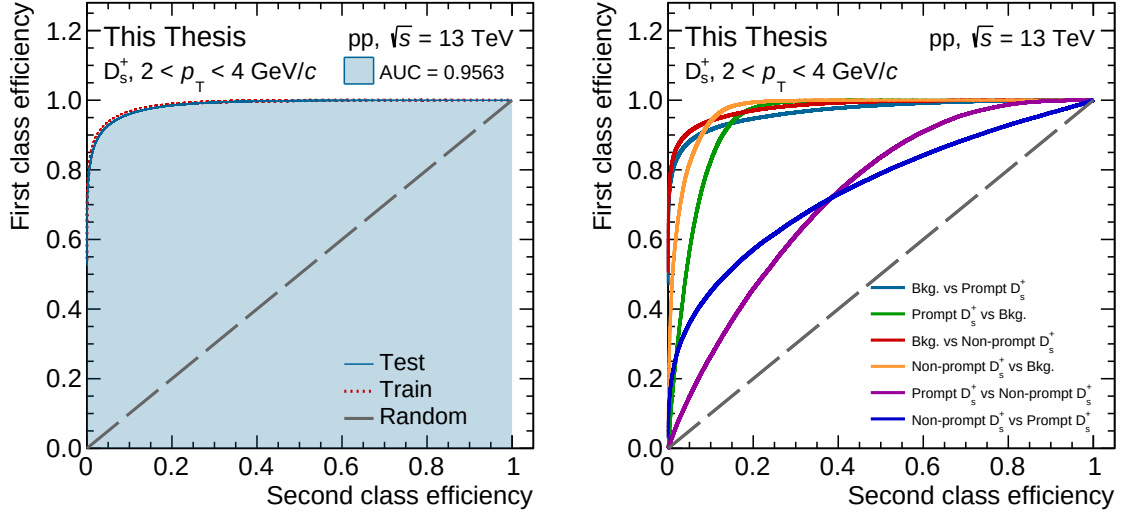


Fig. 3.2 Left: ROC curve of the binary-classification algorithm employed in the study of the production of the D_s^+ mesons in pp collisions at $\sqrt{s} = 13$ TeV during the training and test phase. The $2 < p_T < 4$ GeV/c interval is considered. The ROC curve of the random classifier is also shown for comparison. Right: ROC curve of the multiclass classification algorithm employed in the study of the production of the D_s^+ meson in pp collisions at $\sqrt{s} = 13$ TeV. The $2 < p_T < 4$ GeV/c interval is considered.

the validation set may not be large enough to properly optimise the hyperparameters of the algorithm. In order to overcome these issues, the *cross-validation* technique can be employed [193]. The cross-validation consists of dividing the training set into k subsets of equal size, called *folds*. Then, the algorithm is trained k times, each time using a different fold as the validation set. The remaining $k - 1$ folds are used as a training set, adopting a specific hyperparameter configuration. The final performance is then obtained by averaging the performance obtained for each fold. This operation is repeated for each hyperparameter configuration the analyser wants to test. The hyperparameter configuration that minimise the objective function is then chosen as the optimal configuration. Typically, the number of folds is set to 5 or 10 [193]. This allows for a good trade-off between the accuracy of the final model and the computational cost, both increasing with the number of folds.

3.1.3 Hyperparameters Optimisation

Hyperparameters are essential settings that are external to the ML algorithm but significantly impact its performance. Unlike the internal parameters, which are optimised during the training, hyperparameters are set before the training begins and are not further updated. Hence, the optimisation of the hyperparameter has to be included in the ML chain to ensure

the best performance of the algorithm. The process of hyperparameter optimisation heavily relies on the complexity of the ML algorithm and the available computational resources. Various approaches can be utilised, with the most popular ones being the *grid search* [195], *random search* [195], and *Bayesian optimisation* [196, 197].

In the grid search method, a predefined grid of hyperparameter values is defined, and the algorithm performance is evaluated for each combination of hyperparameter values. The random search, on the other hand, involves randomly sampling the hyperparameter space and evaluating the algorithm performance for each configuration. In both cases, the hyperparameter setup with the best performance is then chosen as the optimal configuration. The Bayesian optimisation approach consists of modeling the objective function as a posterior probabilistic surrogate function, denoted as $p(\text{obj}|X)$, where X is the history vector containing each previously-tested hyperparameter configuration and their corresponding value of the objective function. Typically, the posterior probabilistic surrogate function is modelled using multivariate Gaussian distributions. Starting from X , providing information on the past iterations, and an *acquisition function* $a(\theta)$, which quantifies the utility of evaluating the objective function for a specific hyperparameter configuration, the next set of hyperparameters to be evaluated is selected. This operation minimises the number of objective function evaluations and explores only the most promising regions of the hyperparameter space. The posterior probabilistic surrogate function is continually updated with each new evaluation of the objective function.

Figure 3.3 illustrates the optimisation procedure followed by the three approaches in the case of two hyperparameters [198]. The top and left plots show the dependence of the loss function on the first and second hyperparameter, respectively. The colour gradient of the dots in the Central plot indicates the order of the evaluated hyperparameter configurations, from black to yellow to orange. The grid search (left-panel) is the least efficient approach as it does not learn from previous evaluations of the objective function, making it slow to converge to the optimal hyperparameter configuration. However, in a more general context, it can be computationally less expensive than other methods as it only requires the evaluation of the objective function for the hyperparameter configurations defined in the grid set by the user, which are typically a restricted number. The random search (central panel) strikes a good balance between efficiency and computational cost when the number of hyperparameters is low, as in the case reported in Fig. 3.3. Though, it becomes less efficient with an increasing number of hyperparameters to be optimised, which is the case for the analyses presented in this Thesis. Finally, the Bayesian optimisation (right-panel) is the most efficient approach as it learns from previous evaluations of the objective function pointing the search toward the global minimum of the objective function in the hyperparameter space in a few iterations.

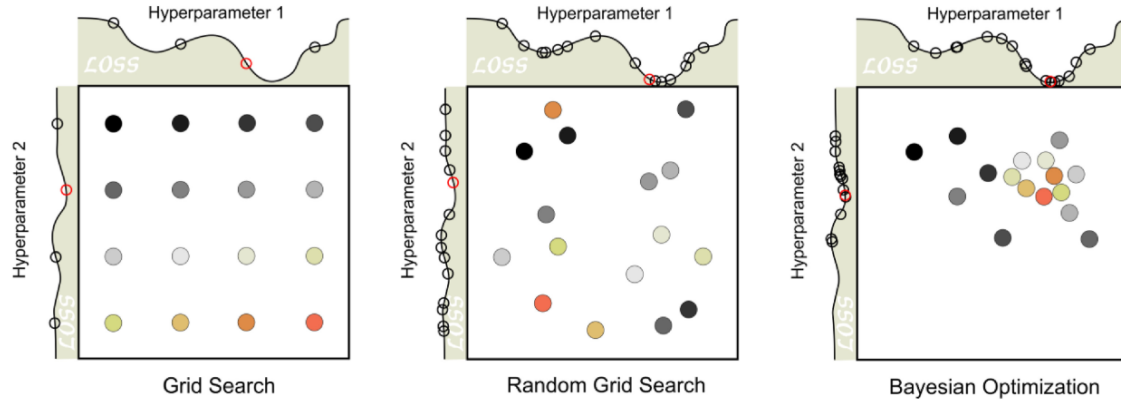


Fig. 3.3 Hyperparameters optimisation for the grid search (left-panel), the random search (central-panel), and the Bayesian optimisation (right-panel). The dependence of the loss function on the two hyperparameters is shown on the top and on the left of each plot. The colour gradient of the points, from black to “yellowish” to orange indicates the order of the search. Figure from Ref. [198].

The Bayesian optimisation approach was adopted for the ML algorithm employed in the analyses presented in this Thesis.

3.1.4 Feature Importance

The model explainability and accountability are essential aspects of ML algorithms. Even though their performance is often superior to traditional linear selections in particle physics studies, the ML output is typically more difficult to interpret and this is often a major drawback. In other words, the analyser is not able to properly infer the feature-output value relation learned by the algorithm, leading to an incomplete comprehension of the physics behind the output score and the subsequent selection criteria. A way to overcome this issue is to investigate the importance of the features provided to the algorithm, also known as *feature importance* [199]. Understanding which features are the most important to the ML algorithm to make a certain decision can help the analyser to interpret the outcome of the algorithm and to improve the performance of the algorithm itself. Also, in the case of physics analyses, the feature importance can be used as a cross-check to ensure that the algorithm predilection for a specific feature is matched by the physics of the problem under study.

The feature importance can be evaluated using different approaches. For instance, one could evaluate the feature importance by considering the number of times a specific feature is employed by the algorithm to split the training set. The most influencing features will have a higher number of occurrences. Another approach is to evaluate the feature importance by

considering the impact of a certain feature on the objective function. The impact of a feature can be evaluated by considering the number of times a specific feature is employed by the algorithm to modify the score of a candidate. An alternative approach is the *permutation importance* [200]. The permutation importance is evaluated by interchanging the values of the features in the validation set and measuring the impact on the objective function. The permutation importance is then defined as the difference between the objective function before and after the exchange. The idea behind this technique is that the most influencing feature will have a higher impact on the objective function.

One of the most common approaches is provided by the *SHapely Additive exPlanations* (SHAP) [201, 202]. The SHAP approach is based on the *Shapley values* [203], which are used in game theory to quantify the contribution of each player to the total payoff of a coalition game. The SHAP approach is based on the assumption that the output score of the ML algorithm can be expressed as a linear combination of the feature values. The SHAP values are computed by iteratively computing the output score including an additional feature at each iteration. Subsequently, the SHAP values are computed by considering all the possible combinations of features and averaging the contribution of each feature to the output score. The SHAP values are then normalised by the standard deviation of the output score. The SHAP values can be used to compute the feature importance. The SHAP approach is particularly useful in the case of complex ML algorithms where the feature importance can not be easily evaluated using other approaches. In addition, SHAP value gives access to the feature contribution to the output score for each instance in the validation set. This allows for a better understanding of the feature-output value relation learned by the algorithm. For the measurements presented in this Thesis, the SHAP approach was employed to evaluate the feature importance of the ML algorithm.

A beeswarm-style SHAP summary plot is reported in Fig. 3.4. The features provided to the model are reported on the left of the plot according to their importance. Namely, the most important features are reported on the top of the plot, while the least important features are reported on the bottom of the plot. For every example employed in the hyperparameter optimisation, a dot is added to each feature distribution. The position of the dots on the x-axis corresponds to the impact that the feature has on the final output score measured by SHAP. In the case of the most effective features, the dots will lie on the right-hand side of the plot. Dots sharing the same x position for the same feature will pile up to show the density of that region. Finally, the colour of the dots corresponds to the value of the feature. The colour scale is reported on the right of the plot. The colour scale is symmetric around zero, where the red colour corresponds to the highest value of the feature, while the blue

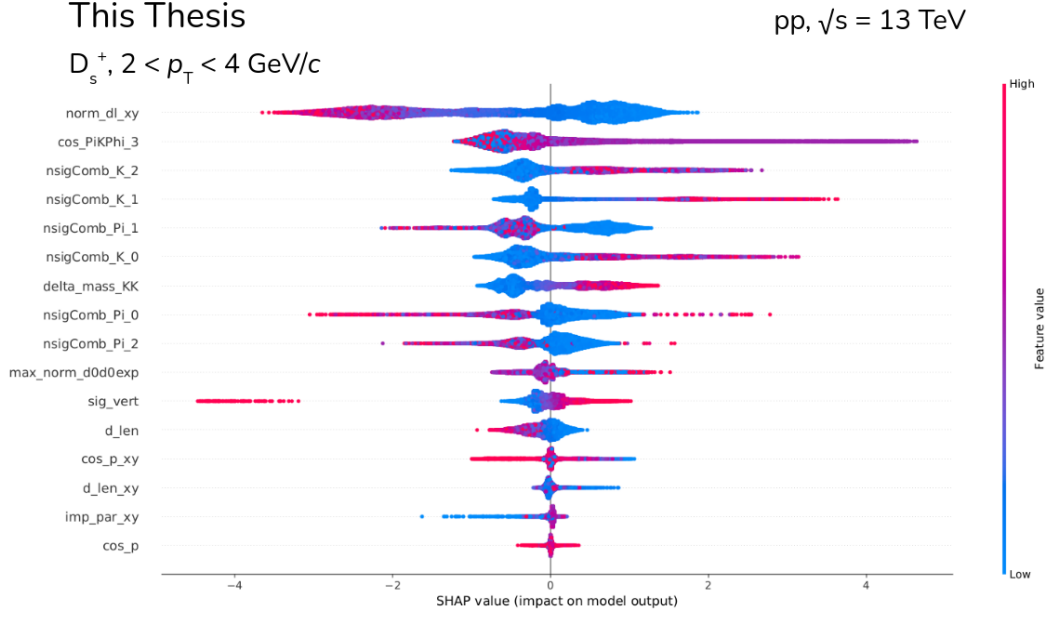


Fig. 3.4 SHAP feature importance plot for the D_s^+ meson analyses in pp collisions at $\sqrt{s} = 13$ TeV. The $2 < p_T < 4$ GeV/c interval is considered.

colour corresponds to the lowest value of the feature. A complete description of the features employed in the analyses presented in this Thesis is reported in Chapter 5.

3.2 Boosted Decision Tree

The *Boosted Decision Tree* (BDT) [200, 204] is a popular supervised learning algorithm employed for regression and classification problems. The *Decision Tree* (DT) [205] serves as the fundamental component of BDTs. Introduced in the 1980s [206], DTs expand upon linear-selection-based analyses by accommodating events that may fail particular selection criteria. As a significant fraction of candidates lacks all the distinguishing characteristics of either signal or background classes, finding a single selection criterion for effective separation becomes challenging. To address this, DT algorithms consist of a series of binary-output selections, known as *nodes*. The nodes are organised in a hierarchical structure, where the initial nodes, known as *roots*, are associated with the input features, and the final nodes, known as *leaves*, are associated with the final output score. Each node compares a single feature of the input candidate to a specified threshold value. Instead of simply accepting or rejecting the candidate, the DT assigns a positive or negative score based on the outcome of

the comparison. This score is propagated to the next node, where a new selection criterion is applied to a different feature. The more effective the selection criterion, the higher will be the score assigned to the candidate and the separation between the classes. The process is iteratively repeated until a predefined stopping criterion is met or all the instances are classified.

Similarly to other supervised ML algorithms, DTs have to follow the training-testing-validation process. During the training phase, the DT architecture, including the threshold values and node order, is optimised to maximise the separation between the classes, in the case of a classification problem, or minimise the discrepancy between the predicted and expected value of the function under study, in the case of regression problems. The sequential logic of DT and its interpretable output make it more appealing than complex ML algorithms, where inferring the logic behind the output is less straightforward and the black-box nature of the algorithm is unavoidable. However, their capability to learn complex relations between features and output value is limited, which is why DTs are often referred to as *weak learners*. Increasing the depth of the DT, representing the number of nodes it comprises, allows the algorithm to enhance its complexity, though this makes the DT more prone to overfitting. To overcome this limitation, a technique called *boosting* is employed.

Boosting involves training a series of weak learners, as small-depth DTs, and combining them into a single strong learner, the BDT. The DTs are trained sequentially, with each one using the training set and the output of the previous DT as input. The final output of the BDT is obtained by combining the output of all the DT of the ensemble. A very popular implementation of the boosting technique is the *AdaBoost* algorithm [207], which is based on the *gradient boosting* technique [208]. The gradient boosting technique involves training each DT on the *residuals* of the previous DT. The residuals are defined as the difference between the true output value and the predicted output value of the previous DT. The residuals are then used as the new true output value for the next DT. This process is repeated until the residuals are minimised. The BDT is a powerful algorithm capable of learning complex relations between features and output value, while being less prone to overfitting than a deep DT. For these reasons, the BDT is the algorithm employed for the analyses presented in this Thesis. Figure 3.5 shows the schematic representation of a BDT with two DTs.

3.2.1 XGBoost

In this Thesis, the BDT implementation from the *Extreme Gradient Boosting* (XGBoost) [210] library is employed for the analyses. This open-source package has achieved state-of-the-art

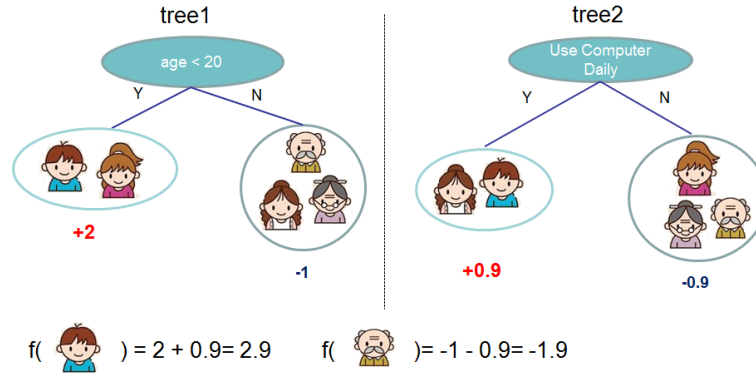


Fig. 3.5 Schematic representation of a BDT. Figure from Ref. [209].

results in numerous machine-learning competitions. Additionally, XGBoost supports parallel execution on CPUs and GPU, leading to reduced training times. This BDT is based on the gradient boosting method presented before and is characterised by high flexibility and scalability. The XGBoost algorithm is based on the *CART* (Classification and Regression Tree) algorithm [206], which is a DT algorithm capable of handling both classification and regression problems. The algorithm provided by XGBoost has many hyperparameters that can be tuned to control the BDT model complexity and improve the performance for the problem of interest. The full list of available hyperparameters, together with their default values and a description of their meaning, can be found in the XGBoost documentation [209]. The hyperparameters considered during the BDT optimisation in this Thesis are:

- **max_depth**: the maximum depth of the tree. Increasing this value makes the model more complex and likely to overfit. The default value is around 4;
- **learning_rate**: the learning rate of the algorithm is a parameter that controls the step size at each iteration of the gradient boosting process. The default value is around 0.08;
- **n_estimators**: the number of trees to be used in the BDT. The default value is around 1000;
- **min_child_weight**: the minimum sum of instance weight needed in a node to be split. This parameter sets the stopping criterion in the tree growth. Increasing this value makes the model more conservative. The default value is around 0.90;
- **sub_sample**: the fraction of the training set used to train each tree. The default value is around 0.90;

- `colsample_bytree`: the fraction of features used to train each tree. The default value is 1;
- `n_jobs`: the number of parallel threads used to train the BDT. The default value is 1;
- `tree_method`: the algorithm used to build the trees. The default value is `hist` which is a fast approximate algorithm based on the histogram of the features;

3.3 Software tools for Machine Learning analyses

The *Toolkit for Multivariate Data Analysis* (TMVA) package [211] in the ROOT framework, based on C++, provides an implementation of the BDT algorithm. Although TMVA seamlessly integrates machine-learning tools with the ALICE experiment analysis software (refer to Sec 2.3) a different approach has been adopted in this Thesis, utilising the Python ecosystem. The idea behind this choice is that Python is a dominant programming language in the machine-learning community, offering a vast array of documentation and advanced techniques, surpassing what TMVA provides. In order to bridge the gap between the ALICE analysis framework and the ML tools, the ROOT bindings for Python [212] are employed in this Thesis to be able to use the ROOT methods in Pythonic scripts. In addition, the proper handling of the ALICE data, the AODs, which are ROOT objects, in the Python code is ensured by the `uproot` package [213].

The ML workflow is implemented in `hipec4ml` [214]¹, developed and maintained by the ALICE group in Turin. The `hipec4ml` package includes a variety of common ML packages which allows for the complete implementation of the ML workflow. The `hipec4ml` package is based on:

- `bayesian-optimisation` [196], which is a Python package for hyperparameters optimisation;
- `matplotlib` [215], which is a comprehensive library for creating static, animated, and interactive visualisations in Python;
- `numpy` [216], which is a fundamental package for scientific computing in Python;
- `pandas` [217, 218], which is fast, powerful, flexible and easy-to-use open source data analysis and manipulation tool;

¹<https://github.com/hipec4ml/hipec4ml>

- `scikit-learn` [192, 219], which is one of the most popular ML packages in Python;
- `scipy` [220], which is a Python package for scientific computing;
- `shap` [202], which is a Python package for model explainability;
- `uproot` [213], which is a Python package for reading and writing ROOT files;
- `xgboost` [210, 209], which is a Python package implementing the BDT algorithm.

Finally, the `DmesonAnalysis` repository is employed to perform the analyses described later in this Thesis. The `DmesonAnalysis` repository is a collection of Python and C++ scripts developed by the ALICE analysers focused on the study of D mesons. The repository is based on the `hipec4ml` package and ROOT, and provides a complete implementation of the analysis workflow, from the data handling to the final results. The repository is available on GitHub².

²<https://github.com/DmesonAnalysers/DmesonAnalysis>

Chapter 4

Strange D meson reconstruction in pp and Pb–Pb collision

The D_s^+ mesons are composite particles encompassing a charm ($m_c \sim 1.3 \text{ GeV}/c$ [6]) and an anti-strange quark ($m_{\bar{s}} \sim 0.5 \text{ GeV}/c$ [6]). They are characterised by a mean proper decay length $c\tau \simeq 150\mu\text{m}$ [6], preventing them from reaching the innermost layer of the ALICE detector. Consequently, they cannot be directly detected. Instead, their measurements rely on the exclusive reconstruction of their decay chain by means of the detection of their decay products. In particular, for what concerns the measurements presented in this Thesis, the D_s^+ mesons are measured by fully reconstructing the topology of the hadronic decay channel $D_s^+ \rightarrow \phi \pi^+ \rightarrow K^+ K^- \pi^+$. Despite the relatively low frequency of the D_s^+ meson decay into this specific channel, referred to as *branching ratio* $\text{BR} = (2.22 \pm 0.06)\%$ [6], which includes the BR of the $\phi \rightarrow K^+ K^-$ decay that is $\text{BR} = (49.1 \pm 0.5)\%$ [6], this decay channel is chosen for the presence of an intermediate resonant state. This state is particularly useful at reconstruction and analysis level, and to exploit the complete reconstruction of the decay products.

The D_s^+ meson is the weakly-decaying ground state of the charm-strange meson family. However, it is not the only charm-strange meson. Over the last decade, charm meson spectroscopy has undergone a resurgence due to the experimental observation of numerous excited charm and charm-strange states and exotic bound states [221–224]. So far, the measurements performed at the LHC energy mainly focused on the characterisation of these states, i.e. their mass, spin properties, and their lifetime. Within the scope of this Thesis, an explorative study of the $D_{s1}^+(2536)$ ($J^P = 1^+$) and $D_{s2}^{*+}(2573)$ ($J^P = 2^+$) meson production is undertaken, and it is described in Chapter 5. The study of the production of these particles

offers a novel approach to investigate the charm quark hadronisation mechanism. Due to their rather short lifetime ($\tau(D_{s1}^+) \sim 219$ fm [6] and $\tau(D_{s2}^{*+}) \sim 12$ fm [6]), they also allow to inspect the hadronic phase of heavy-ion collisions [225, 47, 226, 227]. The existence of this phase in the evolution of high-multiplicity pp and p–Pb collisions is still under debate and these measurements could shed light on this open point [45, 47].

Furthermore, some of these states are potential contenders for *exotic states*, making this analysis instrumental in elucidating their underlying characteristics. These resonant states, just like the D_s^+ meson, decay before encountering the innermost layer of the ALICE detector. Therefore, the reconstruction strategy is also based on the reconstruction of their decay products. The D_{s1}^+ and D_{s2}^{*+} mesons are reconstructed in the hadronic decay channels $D_{s1}^+ \rightarrow D^{*+}K_S^0$ and $D_{s2}^{*+} \rightarrow D^+K_S^0$, respectively. The D^{*+} and D^+ mesons, in turn, are reconstructed in the hadronic decay channels $D^{*+} \rightarrow D^0(\rightarrow K^-\pi^+)\pi^+$ (BR = $(2.67 \pm 0.02)\%$ [6]) and $D^+ \rightarrow K^-\pi^+\pi^+$ (BR = $(9.38 \pm 0.16)\%$ [6]), via an invariant mass analysis of fully reconstructed decay topologies. The K_S^0 are reconstructed in their weak decay channel $K_S^0 \rightarrow \pi^+\pi^-$ (BR = $(69.20 \pm 0.05)\%$ [6]) following the procedure described in Ref. [228].

The analysis strategy relies on the creation of D-meson candidates from the combination of triplets, in the case of D^+ and D_s^+ mesons, or couples, in the case of D^0 mesons, of reconstructed tracks with appropriate charge sign. For instance, the number of charged particles produced in central Pb–Pb collisions [47] per unit of η at midrapidity is of the order of 2000, leading to numerous possible combinations of tracks. Consequently, a large amount of D-meson candidates originate from wrongly-associated charged tracks. This kind of D-meson candidate is referred to as *combinatorial background*¹. To be able to measure, with reasonable statistical significance, the rare processes leading to the production of the D mesons, a series of selection criteria are applied at the reconstruction and analysis level. At reconstruction level, beyond making use of the particle identification capabilities of the ALICE experiment to select the proper decay products, a crucial selection for heavy-flavour measurements arises from the excellent spatial resolution of the ALICE detector [178, 229]. Hence, the D-meson decay vertex (namely *secondary vertex*), can be reconstructed and selections on its displacement from the collision point (*primary vertex*) position can be applied. This allows for a clear rejection of the background, but it also permits the distinction between prompt, i.e. originated from charm quark hadronisation and the decay of excited-

¹Due to its combinatorial nature, the number of combinatorial background candidates depends on the number of possible combinations of the reconstructed tracks. Thus, the total number of combinatorial background candidates is expected to increase with the number of tracks employed to build the D-meson candidate (commonly referred to as 2- or 3-prongs). For instance, a larger combinatorial background is observed for D^+ , reconstructed in a 3-prong decay channel, compared to D^0 , reconstructed in a 2-prong decay channel.

charm states, and non-prompt, i.e. originated from beauty hadron decays, D mesons. This distinction is possible due to the greater displacement of non-prompt D_s^+ mesons from the primary vertex compared to prompt ones. The difference is due to the distance travelled by the parent beauty hadrons, which have a mean decay length of about $c\tau \simeq 500\mu\text{m}$ [6] depending on the particle species.

In addition, to effectively suppress the background from the large number of combinations of charged tracks, machine-learning-based selection criteria are implemented at the analysis level for the D_s^+ , D^+ , and D^0 mesons. These criteria are based on the *decay topology*, the *kinematics* of the D-meson candidate, and the PID of the decay products. In the case of the D^{*+} -meson reconstruction the selections are applied on the D^0 meson decay topology. The choice of the selection criteria is described in detail in the following Chapters.

The data samples collected by the ALICE experiment and employed in the analyses presented in this Thesis are described in Secs. 4.1.1 and 4.1.2, respectively, together with the criteria of the event selection. The reconstruction strategy of strange-D-meson candidates is presented in detail in Sec. 4.2. The track-quality selection criteria, the geometrical and kinematic variables, and the PID strategy adopted in the analyses are described in Secs. 4.2.3, 4.2.4, and 4.2.5, respectively. Finally, the Monte Carlo samples employed for training and testing phases of the machine learning algorithm, and the estimation of the reconstruction efficiency of the strange D mesons are described in Sec. 4.3.

4.1 Data sample and event selection

4.1.1 Data taking and trigger conditions

The data analysed in this Thesis were collected by the ALICE detector during the LHC Run 2, in 2016, 2017, and 2018 with pp collision at centre-of-mass energy $\sqrt{s} = 13$ TeV and in 2018 during the Pb–Pb run at a centre-of-mass energy per nucleon pair of $\sqrt{s_{\text{NN}}} = 5.02$ TeV. All data samples for the D_s^+ meson study were collected with the minimum-bias (MBand) trigger, which requires a signal in the V0A and V0C detectors in coincidence with the bunch crossing time (refer to Chapter 2). For the study of the production of the strange-D-meson resonances as a function of the charged-particle multiplicity in pp collisions, the data samples were enriched with a *V0-triggered high-multiplicity* (HMV0) trigger based on a minimum selection of V0 signal amplitudes. In the case of the production of D_s^+ mesons in central and semicentral Pb–Pb collisions, the *centrality* trigger (CENT, SEMI) that are based on the V0 signal amplitude were adopted to enhance the number of central and semicentral events,

respectively [230, 178]. In particular, different thresholds on the summed-up signals of the V0 detectors were imposed depending on the centrality, and then the coincidence of the two sides was required [231].

4.1.2 Offline event selection

The presence of machine-induced background events stemming from beam-gas interactions and collisions with multiple interaction vertices in the same bunch crossing (*pileup*) was effectively eliminated by employing several event selection criteria. In Pb–Pb collisions, the timing information derived from the V0 and ZDC detectors was utilised to reject beam-gas interaction, out-of-bunch pileup, and collisions involving satellite bunches. In the case of pp collisions, the same purpose was achieved by leveraging the timing information from the V0 detector and the correlation between the number of hits and tracklets within the SPD detector. As explained in detail in Sec. 2.5, the reconstruction of the primary vertex can be accomplished using either global tracks or SPD tracklets. Given that the mean proper decay length of D_s^+ mesons is approximately $151\mu\text{m}$ [6], a precise determination of the primary vertex position is of paramount importance for efficient selections based on the decay topology. As a result, only events with a primary vertex reconstructed through global tracks (refer to Chapter 2) were employed for the analyses presented in this Thesis. However, this choice comes at the expense of reduced efficiency in vertex reconstruction for low-multiplicity pp events. It is noteworthy that low-multiplicity events devoid of a reconstructed vertex using global tracks are highly improbable sources of D mesons at midrapidity decaying in the relevant hadronic channels.

To ensure pseudorapidity coverage of $|\eta| < 0.8$ for all ITS layers, selected events were required to possess a reconstructed primary vertex within ± 10 cm from the nominal position along the beam axis ($|z_{\text{vtx}}| \leq 10$ cm), which corresponds to the z axis in the ALICE reference frame (refer to Chapter 2). This selection not only enhances the geometrical uniformity but also helps to filter out residual satellite collisions surviving the timing selection described above. The occurrence of multiple primary vertices is used to reject events with the pileup of multiple collisions within the same bunch crossing within the SPD readout time in pp collisions. This effect is due to the accumulation of multiple interaction vertices occurring within the same bunch crossing, referred to as *in-bunch pileup events*, or in bunch-crossing separated by the one that fired the trigger by less than the SPD readout time of 300 ns. In the Pb–Pb sample these sources of pileup are negligible due to the low interaction rate and the filling scheme adopted in the data taking (with a bunch spacing ≥ 100 ns). In pp collisions pileup events are recognised with an algorithm based on the SPD tracks unlinked to the

primary vertex, which is chosen as the space point to which the highest number of tracklets point, as described in detail in Sec. 2.5. Should another vertex with a minimum of five contributors be identified beyond ± 0.8 cm from the main vertex, the event is identified as pileup and subsequently removed. The substantial pileup occurring during the TPC readout time, which is about $90\mu\text{s}$, is addressed at the analysis level. Further selections on the centrality were applied for each specific analysis and will be discussed in the dedicated Chapters.

The event selection efficiency should be considered in the calculation of the number of events to be used to normalise the D-meson yield. In particular, one should account for those collisions which occurred within $|z_{\text{vtx}}| \leq 10$ cm but for which the primary vertex was not reconstructed. These events are rejected by the event selection, but they need to be considered for a correct normalisation. The correct normalisation is²:

$$\begin{aligned}
 N_{\text{ev}}^{\text{norm}} &= N_{\text{ev}}^{\text{reco vtx}} \Big|_{|z_{\text{vtx}}| < 10\text{cm}} + N_{\text{ev}}^{\text{no vtx}} \Big|_{|z_{\text{vtx}}| < 10\text{cm}} \\
 &= N_{\text{ev}}^{\text{reco vtx}} \Big|_{|z_{\text{vtx}}| < 10\text{cm}} + N_{\text{ev}}^{\text{no vtx}} - N_{\text{ev}}^{\text{no vtx}} \Big|_{|z_{\text{vtx}}| < 10\text{cm}} \\
 &= N_{\text{ev}}^{\text{reco vtx}} \Big|_{|z_{\text{vtx}}| < 10\text{cm}} + N_{\text{ev}}^{\text{no vtx}} - N_{\text{ev}}^{\text{no vtx}} \times N_{\text{ev}}^{\text{reco vtx}} \Big|_{|z_{\text{vtx}}| > 10\text{cm}} / N_{\text{ev}}^{\text{reco vtx}} .
 \end{aligned} \tag{4.1}$$

In Pb–Pb collisions, an additional selection on the centrality of the collision was applied, based on the amplitude of the signal measured by the V0 detectors, as described in detail in Sec. 2.4.

Table 4.1 summarises the data samples used in this analysis, together with the corresponding integrated luminosity ($\mathcal{L}_{\text{int}}$) and number of events (N_{ev}), after the event selections described above. The visible cross sections for the MB trigger were measured with Van der Meer scans and correspond to $\sigma_{\text{vis}}^{\text{pp}} = (57.95 \pm 0.016)$ mb [232] and $\sigma_{\text{vis}}^{\text{Pb–Pb}} = (7.67 \pm 0.25)$ b [233] for pp collisions at $\sqrt{s} = 13$ TeV and Pb–Pb collisions at $\sqrt{s_{\text{NN}}} = 5.02$ TeV, respectively.

²The second term derives from the assumption that the primary vertex reconstruction efficiency is independent of the z_{vtx} position.

Table 4.1 Summary of the data samples used for the analyses presented in this Thesis. The values of the integrated luminosity ($\mathcal{L}_{\text{int}}$) are calculated using the MB trigger cross sections taken from Refs. [232, 233]. †: due to the *trigger turn-on curves* [234], the calculation of the integrated luminosity is not straightforward. The values reported here are computed considering only the data periods for which the triggers are fully efficient in the multiplicity region of interest.

System	Year	Trigger	$\sqrt{s}$ or $\sqrt{s_{\text{NN}}}$ (TeV)	$\mathcal{L}_{\text{int}}$ (μb^{-1})	N_{ev}
pp	2016–2018	MBand	13	31900 ± 500	1.6×10^9
pp	2016–2018	HMV0	13	$0.8 \times 10^{6\dagger}$	0.4×10^9
Pb–Pb	2018	SEMI	5.02	55.5 ± 0.2	0.9×10^8
Pb–Pb	2018	CENT	5.02	130.5 ± 0.5	1.0×10^8

4.2 Strange D meson reconstruction

4.2.1 D_s^+ meson secondary vertex reconstruction

The D_s^+ mesons (and their charge conjugates) are reconstructed in their hadronic decay channel $D_s^+ \rightarrow \phi \pi^+ \rightarrow K^+ K^- \pi^+$ (BR = $(2.22 \pm 0.06)\%$ [6]). This channel is chosen also to exploit the presence of the intermediate resonant state, the ϕ -meson, which notably facilitates the D_s^+ -meson candidate selection at the reconstruction and analysis level³.

The analysis strategy relies on the selection of fully reconstructed decay-vertex topologies, which are displaced with respect to the primary vertex position. An invariant mass analysis of the decay products is then performed to extract the signal raw yield, which is subsequently corrected for the geometrical acceptance and reconstruction and selection efficiency obtained from simulated events. The prompt D_s^+ mesons typically travel for about a few hundred microns before decaying, $c\tau \simeq (151 \pm 1)\mu\text{m}$ [6]. Furthermore, the decay vertex of non-prompt D_s^+ mesons will be even more displaced from the primary vertex position compared to the prompt component due to the distance travelled by the beauty-hadron parents h_b , which have $c\tau \simeq 500\mu\text{m}$ [6]. The main characteristics of D_s^+ mesons and the beauty hadrons of interest are reported in Table 4.2.

According to the fragmentation fractions of b quarks into b-hadron species, $f \rightarrow h_b$, measured in $Z^0 \rightarrow b\bar{b}$ decays [6] and the beauty-hadron BRs from the PDG [6] or from

³A fraction of D_s^+ mesons decay in the same final state of the chosen hadronic decay but via a different resonance. In particular, the $D_s^+ \rightarrow \bar{K}^{*0}(892)K \rightarrow K^+ K^- \pi^+$ (BR = $(2.58 \pm 0.08)\%$ [6]) and the $D_s^+ \rightarrow f_0(980)\pi^+ \rightarrow K^+ K^- \pi^+$ (BR = $(1.14 \pm 0.31)\%$) decay channels are expected to contribute the most [6]. The contamination from these channels is taken into account later on in the analyses, as discussed in the Sec. 4.2.4 and Ref. [181, 235].

Table 4.2 Main characteristics of the charm and beauty hadrons of interest. The branching ratios (BR) are taken from Ref. [6]. (†: the BR of the hadronic decay channels of the D-meson resonances are not measured, they are calculated at the analyses level as described in Chapter 5. *: the BR of the hadronic decay channels of the B-meson is different for particle and antiparticle. For antiparticle are reported in parentheses.)

Particle	Mass (MeV/c)	$c\tau$ (μm)	Decay channel	BR (%)	Daughter decay channel
D^0	1864.83 ± 0.05	123 ± 1	$D^0 \rightarrow K^- \pi^+$	3.88 ± 0.04	—
D^+	1869.61 ± 0.05	313 ± 2	$D^+ \rightarrow K^- \pi^+ \pi^+$	9.38 ± 0.16	—
D^{*+}	2010.26 ± 0.05	123 ± 2	$D^{*+} \rightarrow D^0 \pi^+$	67.7 ± 0.5	$D^0 \rightarrow K^- \pi^+$
D_s^+	1968.34 ± 0.07	151 ± 1	$D_s^+ \rightarrow \phi \pi^+ \rightarrow K^+ K^- \pi^+$	2.22 ± 0.06	—
D_{s1}^+	2535.21 ± 0.28	219 ± 3	$D_{s1}^+ \rightarrow D^{*+} K_S^0$	n.a.†	$D^{*+} \rightarrow D^0 \pi^+$
D_{s2}^{*+}	2569.10 ± 0.80	12.7 ± 0.3	$D_{s2}^{*+} \rightarrow D^+ K_S^0$	n.a.†	$D^+ \rightarrow K^- \pi^+ \pi^+$
B^0	5279.65 ± 0.12	455 ± 1	$B^0 \rightarrow D_s^+ X$	$10.3^{+1.1}_{-1.8} (< 2.6)^*$	—
B^+	5279.34 ± 0.12	457 ± 3	$B^+ \rightarrow D_s^+ X$	$7.9^{+1.4}_{-1.3} (1.1^{+0.4}_{-0.3})^*$	—
B_s^0	5366.88 ± 0.14	454 ± 3	$B_s^0 \rightarrow D_s^+ X$	93 ± 25	—
Λ_b^0	5619.6 ± 0.17	441 ± 3	$\Lambda_b^0 \rightarrow \Lambda_c^+ D_s^-$	1.10 ± 0.10	—

PYTHIA 8 event [172], which are typically employed for normalisation purposes in beauty hadron measurements at LHC [236, 237, 139, 238, 239], around half of non-prompt D_s^+ mesons originate from non-strange B-meson decays (i.e. B^0 and B^+) while the other half is produced from decays of B_s^0 mesons. The contribution due to Λ_b^0 -baryon decays is negligible. Figure 4.1 shows a sketch of the D_s^+ meson decay topology in the case of prompt and non-prompt D_s^+ mesons.

The D_s^+ -meson candidates are built by combining triplets of tracks with the proper charge sign, i.e. $(+, -, +)$ or $(-, +, -)$ for D_s^+ and D_s^- , respectively. The decay vertex position is defined as the point that minimises the distance among the three tracks, following the fit procedure explained in Sec. 2.5.1. It is important to note that, due to the ϕ short lifetime ($\tau \simeq 1.6 \times 10^{-22}$ s [6]), the decay vertex of the ϕ meson is indistinguishable from the decay vertex of the D_s^+ meson. The momentum of the reconstructed D_s^+ is obtained as the sum of the momenta of the three daughter tracks, $\vec{p}_i$, evaluated in the space point along their trajectories closest to the decay vertex. The invariant mass of the D_s^+ -meson candidates is calculated as:

$$M_{\text{inv}}^2 = \left(\sum_{i=1}^{N_{\text{prongs}}} \sqrt{m_i^2 + \|\vec{p}_i\|^2} \right)^2 - \sum_{i=1}^{N_{\text{prongs}}} \|\vec{p}_i\|^2, \quad (4.2)$$

where m_i is the adopted mass hypothesis for the decay track. In the case of the $D_s^+ \rightarrow \phi \pi^+ \rightarrow K^+ K^- \pi^+$ decay channel, the adopted mass hypothesis are the pion mass m_π and the kaon mass m_K . For the track with an opposite charge compared to the D_s^+ -meson candidate under study, the only possible mass hypothesis is the kaon one. Instead, for the two tracks having

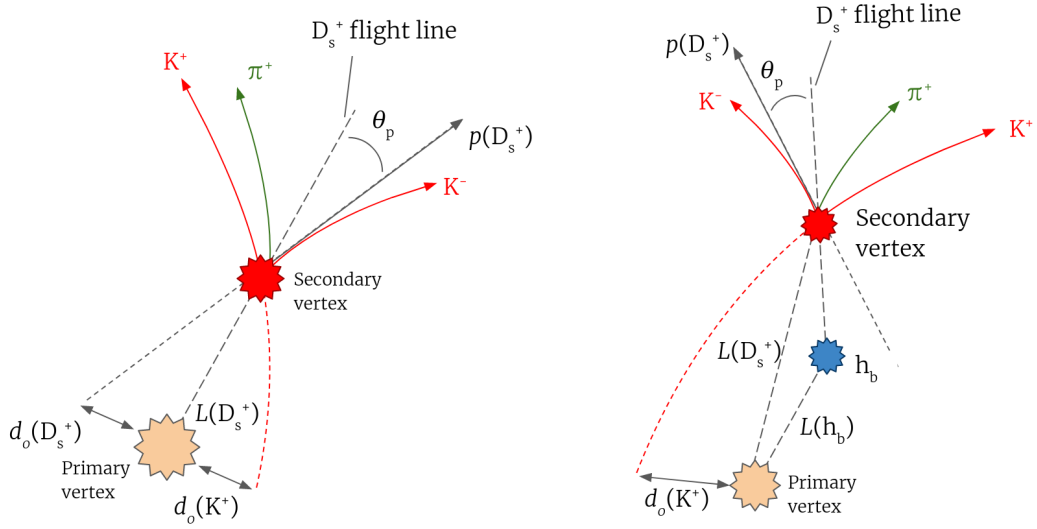


Fig. 4.1 Sketch of the D_s^+ meson decay topology in the case of prompt (left) and non-prompt (right) D_s^+ mesons. Relative distances and angles on the figure are purely schematic.

the same charge sign, two possibilities are available by interchanging the pion and kaon hypotheses. In this way two mass hypotheses are calculated for each D_s^+ -meson candidate. For signal candidates, i.e. triplets of tracks originating from the decay of a D_s^+ meson, having the wrong decay-particle mass assignment (called *reflections*) only the selections on the PID of the decay tracks and the ϕ -meson mass allow to reduce the contamination from reflections to a negligible level.

On average, each event involving pp collisions leads to the reconstruction of approximately 3 D_s^+ -meson candidates in the $|y| < 0.8$ range after the application of preliminary filtering criteria to the data. In the case of central (semicentral) Pb–Pb collisions, the corresponding values are 10^4 (10^3). To enhance the signal-over-background ratio and achieve a reasonable statistical significance in the measurements, severe analyses selections must be applied. In particular, three kinds of selection are usually employed: selection on the daughter tracks, selection on the D_s^+ candidate decay topology and kinematics, and particle identification of the decay tracks. All three categories will be discussed in more detail in the following Sections.

4.2.2 D_{s1}^+ and D_{s2}^{*+} resonances reconstruction

The D_{s1}^+ and D_{s2}^{*+} resonances are reconstructed via the $D_{s1}^+ \rightarrow D^{*+} K_S^0$ and $D_{s2}^{*+} \rightarrow D^+ K_S^0$. The BR of the hadronic decay channels of the D_{s1}^+ and D_{s2}^{*+} resonances are not measured. Therefore, as will be discussed in the following Chapters, various assumptions and theoretical predictions are employed to compute the BR of the chosen decay channels [6]. Unlike the case of the D_s^+ , where the candidates are built by combining triplets of tracks, the reconstruction of the D_{s1}^+ and D_{s2}^{*+} resonances starts from reconstructing the K_S^0 candidate in the $K_S^0 \rightarrow \pi^+ \pi^-$ decay (BR = $(69.20 \pm 0.05)\%$) [6], also called *V0-decay*. It consists of a neutral decay vertex largely displaced from the primary vertex ($c\tau \simeq 269$ mm [6]). The K_S^0 candidate is built by combining two opposite-sign tracks fulfilling specific track quality selections, as the ones presented in Ref. [228]. Due to their long lifetime, the K_S^0 decay at a radial distance from the beam line outside the inner layers of the ALICE ITS. Therefore, the daughter tracks of the K_S^0 decay are not required to have at least one hit in the SPD layers of ITS, unlike the case of the D^+ and D_s^+ meson daughters. The large displacement from the primary vertex enables a clear reconstruction of the K_S^0 without the necessity of further tight selections.

The reconstructed K_S^0 is then paired with a D-meson candidate (D^+ or D^{*+}) to build the desired resonant state. Differently from the D_s^+ meson analysis, no secondary decay vertex is reconstructed for the resonances due to their short lifetime ($c\tau \simeq 219$ fm and $c\tau \simeq 12$ fm [6] for the D_{s1}^+ and D_{s2}^{*+} , respectively), which makes it impossible to distinguish the secondary vertex from the pp interaction point. In addition, no further selection at the analysis level is applied to the resonances. The D-meson species produced in the resonance decay are reconstructed in the hadronic decay channels $D^{*+} \rightarrow D^0 (\rightarrow K^- \pi^+) \pi^+$ (BR = $(2.67 \pm 0.02)\%$ [6]) and $D^+ \rightarrow K^- \pi^+ \pi^+$ (BR = $(9.38 \pm 0.16)\%$ [6]) following a procedure similar to the one described above for the D_s^+ meson.

In the case of the D^+ and D^{*+} there is a large combinatorial background due to uncorrelated tracks as for the D_s^+ . The reduction of the background and the separation of prompt and the non-prompt signal components are based on the reconstruction and selection of secondary-vertex topologies that have significant separation from the primary vertex. In the case of the D^{*+} , the selections are applied on the displaced decay vertex topology of the D^0 produced in the decay. Then, the D^0 meson is paired with a low-momentum (*soft*) π^+ and, due to the short lifetime of the D^{*+} meson, no secondary decay vertex is built. The D^0 and D^+ candidates are reconstructed by pairing tracks with the appropriate charge-sign combination passing the selections described in the following Sections. Similarly to the D_s^+ meson analysis, the D^0 and D^+ meson analyses require tight selections to reach the desired

signal-over-background ratio and statistical significance of the measurements; for this reason a machine-learning-based approach is employed.

4.2.3 Track-quality selections

The tracks employed to build the heavy-flavour candidates are required to fulfil several quality criteria. It is worth mentioning that the selection criteria used for the D_s^+ (D^0 , D^+) daughter track are different from those used for the soft pion track produced in the D^{*+} decay and from the ones applied to the K_S^0 -decay tracks in the reconstruction of D_{s1}^+ and D_{s2}^{*+} resonances. In pp (Pb–Pb) collisions, only tracks with $p_T > 0.3$ GeV/c ($p_T > 0.6$ GeV/c and $p_T > 0.5$ GeV/c, in central and semicentral collisions, respectively) and absolute pseudorapidity $|\eta| < 0.8$ were chosen for the D-meson candidate selection. For central (semicentral) Pb–Pb collisions, a criterion of $|d_0| > 60$ (25) μm was introduced for tracks with $p_T < 2$ GeV/c to select tracks displaced from the primary vertex and reduce the processing time used to produce the candidates. In addition, only tracks that exhibit successful fit in both the TPC and the ITS according to the procedure explained in Sec. 2.5, crossing a minimum of 70 pad rows (out of 159) in the TPC, were considered. A χ^2/ndf value lower than 2 (1.25 for Pb–Pb⁴ collisions.) is requested for the momentum fit in the TPC, where ndf corresponds to the number of degrees of freedom employed in the tracking process. Moreover, a condition is imposed on the ratio between the number of crossed rows in the TPC and the number of findable clusters (pad rows that are theoretically feasible clusters based on the geometry of the track) which should exceed 0.8. To ensure a suitable balance between efficiency and impact-parameter resolution, both of which are pivotal for heavy-flavour hadron measurements, at least two hits in the ITS, out of a maximum of six, with one of them situated within the first two layers, are required for each track. In the case of Pb–Pb collisions, a further requirement of at least 50 clusters available for particle identification in the TPC was applied to reject tracks at the TPC sector edges which are imperfectly tailored in the simulations.

For the soft pion candidate of the D^{*+} decay, the same track quality criteria described above were employed with some exceptions. In particular, only tracks with $p_T > 0.05$ MeV/c were considered and, while the requirement of at least 0.8 crossed rows over findable clusters in the TPC was maintained, the requirement of a successful refit in the TPC at the track reconstruction stage (refer to Chapter 2) was dismissed, since the soft pion momentum is

⁴The difference in the χ^2/ndf value between pp and Pb–Pb collisions is due to a different parametrisation of the uncertainties on the TPC clusters. In particular, the parametrisation adopted for the Pb–Pb data sample collected in 2018 leads to a smaller width of the χ^2/ndf with respect to the pp one. The values reported above were chosen to ensure a similar track rejection [240]

typically too low to cross the full TPC. Also, no specific requirement on the layer position of the two ITS hits was imposed. Regarding the selection of tracks composing the K_S^0 candidates, the ITS successful fit and ITS cluster requirements were released. The same TPC track-quality selections, described above for the D-meson daughter tracks, were used. For tracks meeting the aforementioned prerequisites, the p_T resolution in pp collisions was determined to be superior to 1% for $p_T > 1$ GeV/ c and around 2% for $p_T > 10$ GeV/ c . The impact-parameter resolution for tracks with $p_T > 1$ GeV/ c is about 75 μm , as depicted in Fig. 2.12.

4.2.4 Topological and kinematic selections

As a result of the track selection criteria described above, the detector acceptance for D mesons is reduced and drops steeply to zero at $|\eta| > 0.8$ for high p_T values ($p_T > 5$ GeV/ c). A fiducial acceptance criterion was imposed on the candidate rapidity, denoted as $|y| < y_{\text{fid}}(p_T)$. The $y_{\text{fid}}(p_T)$ factor was defined as:

$$y_{\text{fid}}(p_T) = \begin{cases} 0.5 + \frac{1.9}{15} p_T - \frac{0.2}{15} p_T^2 & \text{for } p_T < 5 \text{ GeV}/c \\ 0.8 & \text{for } p_T > 5 \text{ GeV}/c \end{cases}$$

Figure 4.2 shows the $y_{\text{fid}}(p_T)$ selection as a function of p_T for the D_s^+ mesons in pp collisions at $\sqrt{s} = 13$ TeV.

The selection of the D-meson candidates (i.e. D_s^+ , D^+ , and D^0) is based on a machine learning (ML) approach. As explained in detail in Chapter 3, supervised ML algorithms, like the ones adopted in this Thesis, are firstly trained to "learn" how to properly discriminate the signal from the combinatorial background. The variables provided to the algorithm during the training are based on the decay vertex topology, the decay kinematics, and the PID information discussed in Sec. 4.2.5. The complete list of topological and kinematic variables is reported below. For the most significant variables, the distribution of the prompt D_s^+ , non-prompt D_s^+ , and combinatorial background candidates is shown in the Figures. Similar distributions are observed for D^+ and D^0 . The prompt and non-prompt signal examples used in the training are taken from MC simulations, while the background examples are extracted from candidates in the invariant mass distribution outside the signal region (*sidebands*) in the data. This region is defined as the invariant mass (ΔM) intervals $|\Delta M| > 5\sigma$ with respect to the nominal value of the D^0 mass for the D^0 meson (i.e. $M < 1.80$ GeV/ c^2 and $M > 1.95$

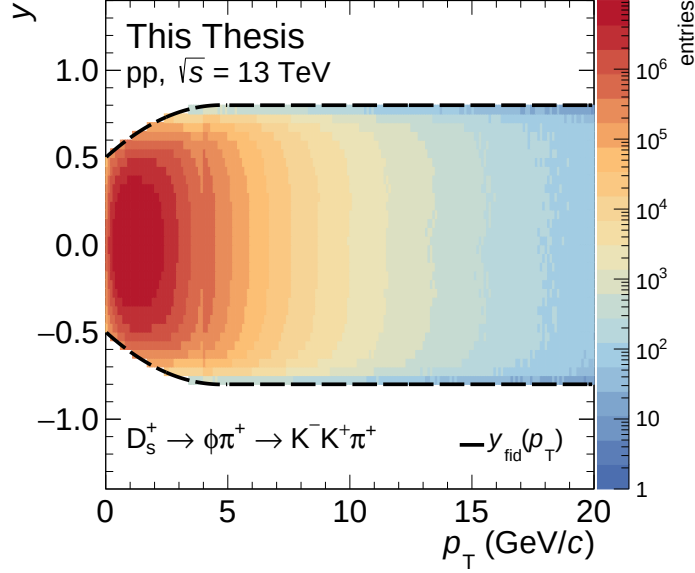


Fig. 4.2 Rapidity distribution versus transverse momentum for D_s^+ mesons reconstructed in pp collisions at $\sqrt{s} = 13$ TeV. The fiducial acceptance region is defined by $|y| < y_{\text{fid}}(p_T)$.

GeV/c^2) and the D^+ mass for D^+ (i.e. $M < 1.82 \text{ GeV}/c^2$ and $M > 1.92 \text{ GeV}/c^2$). Similarly, for the D_s^+ meson a $\Delta M > 5\sigma$ with respect to the nominal value of the D_s^+ mass is applied. However, in this case a second selection of $\Delta M < 5$ with respect to the nominal value of the D^+ mass is applied (i.e. the resulting sidebands-region consists of the $M < 1.83 \text{ GeV}/c^2$ and $2.01 > M \text{ GeV}/c^2$ interval). This is done to reject the contribution from D^+ meson decaying in the same decay channel (refer to Chapter 5). It is important to notice that the relative abundance of signal and background candidates, as well as the capability of the analysis selections to discriminate between the two, strongly depend on the p_T of the D meson. Consequently, the training of the ML algorithm and the ML-based selection are performed in defined p_T interval.

- **Decay length (L, L_{xy})**

The decay length is defined as the spatial distance between the primary and the secondary vertex, measured either in three-dimensional space (L) or projected onto the transverse plane (L_{xy}). The L_{xy} variable is particularly important due to the higher resolution achievable in the x and y coordinates compared to the z coordinate (refer to Fig. 2.2). The influence of the trajectory curvature, resulting from the magnetic field effect, can be neglected for heavy-flavour hadrons due to their relatively short decay length, typically of the order of few hundred micrometers, as reported in Table 4.2.

Figure 4.3 (left panel) shows the decay length distribution for D_s^+ -mesons candidates in pp collisions at $\sqrt{s} = 13$ TeV in the $2 < p_T < 4$ GeV/c interval. The majority of the background candidates originate from the erroneous combination of tracks produced in the primary vertex (i.e. combinatorial background), hence they are expected to have a decay length compatible with zero. The non-zero displacement in the case of the combinatorial background candidates is solely attributed to the limited detector resolution. On the other hand, the prompt D_s^+ mesons are expected to have a larger decay length compared to the combinatorial background candidates due to the distance travelled by the D_s^+ mesons before decaying.

Since for non-prompt D_s^+ -meson the decay length accounts also for the distance travelled by the parent beauty hadrons, the L distribution is on average shifted towards higher values compared to prompt D_s^+ and combinatorial background candidates. The decay length has a harder distribution for high-momentum heavy-flavour hadrons because of the Lorentz boost.

- **Normalised decay length in the transverse plane ($L_{xy}/\sigma_{L_{xy}}$)**

It is defined as the projection of the decay length in the transverse plane L_{xy} divided by its uncertainty $\sigma_{L_{xy}}$. This variable is particularly powerful since it provides better separation between signal and background compared to the nominal decay length variables introduced above, especially at low p_T . In particular, the $L_{xy}/\sigma_{L_{xy}}$ allows for an effective selection of those candidates for which the decay vertex reconstruction was extremely precise, hence with a very low $\sigma_{L_{xy}}$.

The right panel of Fig. 4.3 displays the distribution of the normalised decay length in the transverse plane for D_s^+ -mesons candidates in pp collisions at $\sqrt{s} = 13$ TeV in the $2 < p_T < 4$ GeV/c interval. Compared to the left panel of the same Figure, the separation among the three classes of candidates is more pronounced.

- **Cosine of the pointing angle ($\cos(\theta_P)$, $\cos(\theta_P^{xy})$)**

The pointing angle θ_P is the angle between the line connecting the primary vertex to the reconstructed decay vertex and the reconstructed momentum of the D meson. To exploit the better resolution in the transverse plane, also its projection θ_P^{xy} is employed. For prompt D-meson signals, the cosine of the pointing angle is expected to be equal to unity, assuming perfect vertex reconstruction. On the other hand, since non-prompt D-meson candidates originate from the beauty-hadron decay vertex, and not from the primary vertex, the cosine of pointing angle is expected to be slightly lower than unity. This is due to the fact that the D_s^+ meson momentum does not point exactly to the primary vertex, but has still a large degree of pointing because the D_s^+ meson

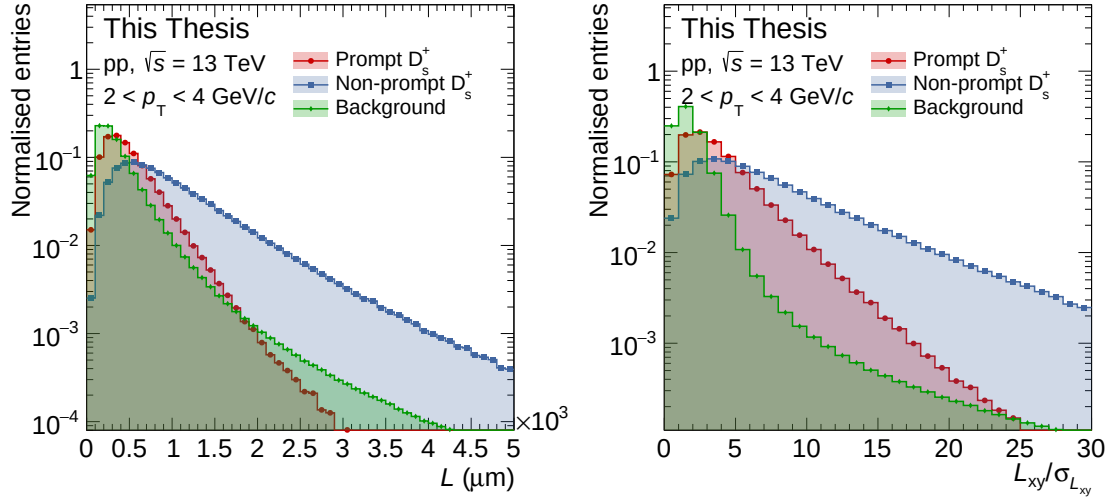


Fig. 4.3 Decay length (left panel) and normalised decay length in the transverse plane (right panel) distributions for D_s^+ mesons in pp collisions at $\sqrt{s} = 13$ TeV in $2 < p_T < 4$ GeV/c. The distributions are shown for prompt D_s^+ mesons (red), non-prompt D_s^+ mesons (blue), and combinatorial background (green). For the prompt and non-prompt D_s^+ mesons, the distributions are taken from Monte Carlo simulations, while for the combinatorial background they are taken from data excluding the signal region in the invariant mass distribution.

inherits a large fraction of the beauty-hadron momentum in the decay. In both cases, the imperfect experimental resolution smears out the distribution towards values lower than unity. Instead, the tracks used to build the combinatorial background are not expected to originate from a common decay vertex, hence the cosine of pointing angle is not expected to point to the primary vertex (or to any other preferential space point, as well). Thus, the distribution of the cosine of the pointing angle is expected to be flatter. In addition, at higher p_T the pointing angle for the signal candidates is expected to be narrower due to the Lorentz boost favouring the separation of the signal from the combinatorial background candidates.

The left panel of Fig. 4.4 shows the distribution of the cosine of the pointing angle in the transverse plane for D_s^+ -meson candidates in pp collisions at $\sqrt{s} = 13$ TeV in the $2 < p_T < 4$ GeV/c interval.

- **Track dispersion at the decay vertex (σ_{vtx})**

The σ_{vtx} parameter is defined as the square root of the sum of the squared distances of closest approach between the decay tracks, or *prongs*, and the reconstructed decay

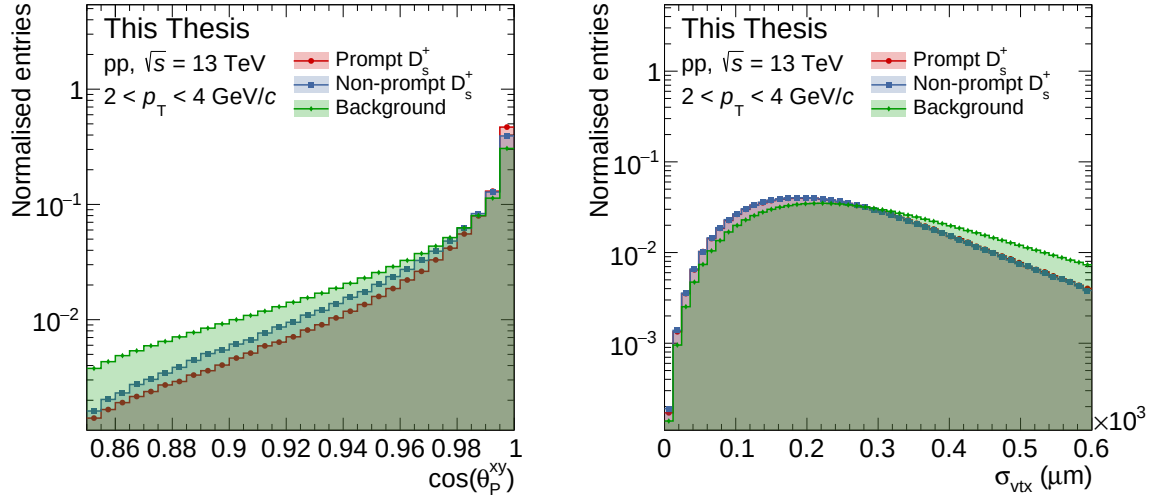


Fig. 4.4 Distributions of the cosine of the pointing angle (left panel) and track dispersion at the decay vertex (right panel) for D_s^+ -meson candidates in pp collisions at $\sqrt{s} = 13$ TeV in $2 < p_T < 4$ GeV/c. The distributions are shown for prompt D_s^+ mesons (red), non-prompt D_s^+ mesons (blue), and combinatorial background (green). For the prompt and non-prompt D_s^+ mesons, the distributions are taken from Monte Carlo simulations, while for the combinatorial background they are taken from data excluding the signal region in the invariant mass distribution.

vertex and is computed as:

$$\sigma_{\text{vtx}} = \sqrt{\sum_{i=1}^{N_{\text{prongs}}} d_i^2}. \quad (4.3)$$

True signal candidates are expected to exhibit a σ_{vtx} equal to zero, neglecting resolution effects, unlike the background candidates that may include tracks from different decay vertices or low-quality tracks which are expected to show higher σ_{vtx} values.

The right panel of Fig. 4.4 displays the σ_{vtx} distribution for D_s^+ -meson candidates in pp collisions at $\sqrt{s} = 13$ TeV in the $2 < p_T < 4$ GeV/c interval.

- **Impact parameter in the transverse plane (d_0^{xy})**

The impact parameter in the transverse plane is defined as the distance of closest approach between the primary vertex and the reconstructed flight line of the D-meson candidate, projected in the transverse plane. Besides rejecting the background, which is expected to exhibit high impact parameter values because the candidate momentum does not point to the primary vertex, this variable is extremely useful for distinguishing between prompt and non-prompt D mesons. Indeed, while prompt D mesons are

expected to have an impact parameter equal to zero, neglecting the experimental resolution, non-prompt D mesons show a harder distribution.

The left panel of Fig. 4.5 shows the distribution of the D_s^+ -meson impact parameter in the transverse plane for the D_s^+ -meson candidates in pp collisions at $\sqrt{s} = 13$ TeV in the $2 < p_T < 4$ GeV/c interval.

- **Maximum normalised difference between the measured and the expected track impact parameters in the transverse plane (D_0^{xy})**

The selection is applied to the track exhibiting the largest discrepancy $\Delta d_{0,i}^{xy}$ between the reconstructed impact parameter in the transverse plane, $d_{0,i}^{xy}(\text{reco})$, and its expected value, $d_{0,i}^{xy}(\text{exp})$, divided by the uncertainty, σ_Δ . Hence, the D_0^{xy} variable is computed as:

$$D_0^{xy} = \max_{i=1}^3 \{ |\Delta d_{0,i}^{xy}| \} \times \text{sign}(d_{0,max}^{xy}), \quad (4.4)$$

where

$$\Delta d_{0,i}^{xy} = \frac{d_{0,i}^{xy}(\text{reco}) - d_{0,i}^{xy}(\text{exp})}{\sigma_\Delta} \quad (4.5)$$

and $\sigma_\Delta = \sqrt{\sigma_{d_{0,i}^{xy}(\text{exp})}^2 + \sigma_{d_{0,i}^{xy}(\text{reco})}^2}$ is the uncertainty on the $\Delta d_{0,i}^{xy}$. For each track the expected impact parameter is computed from the D-meson L_{xy} and the angle θ_i^{xy} between the decay-track and the reconstructed D-meson candidate momentum in the transverse plane as $d_{0,i}^{xy}(\text{exp}) \simeq L_{xy} \sin(\theta_i^{xy})$, under the assumption of linear trajectories. The uncertainty on the expected impact parameter $\sigma_{d_{0,i}^{xy}(\text{exp})}^2$ is computed as $\sigma_{d_{0,i}^{xy}(\text{exp})} = \sigma_{L_{xy}} \sin(\theta_i^{xy})$ assuming no uncertainty on θ_i^{xy} . Similarly to the one on the D_s^+ impact parameter, this selection helps not only to reject the combinatorial background, which exhibits larger D_0^{xy} values compared to the true signal, but also to discriminate between prompt and non-prompt D_s^+ mesons.

The right panel of Fig. 4.5 shows the D_0^{xy} distribution for D_s^+ -meson candidates in pp collisions at $\sqrt{s} = 13$ TeV in the $2 < p_T < 4$ GeV/c interval.

- **Difference between KK invariant mass and ϕ meson mass (ΔM)**

As mentioned above, one of the reasons for choosing the hadronic decay of the D_s^+ via the ϕ meson is that this specific channel enables the exploitation of the narrow width of the ϕ resonance and, among the other possible channels featuring an intermediate state, provides the best discrimination between signal and background [241, 182]. In particular, a selection is applied on the absolute discrepancy between the reconstructed invariant mass of the K^+K^- pair and the ϕ -meson mass ($M(\phi) = (1019.461 \pm 0.016)$ MeV/c [6]). Similarly, a selection on the $\Delta M(K^-\pi^+)$

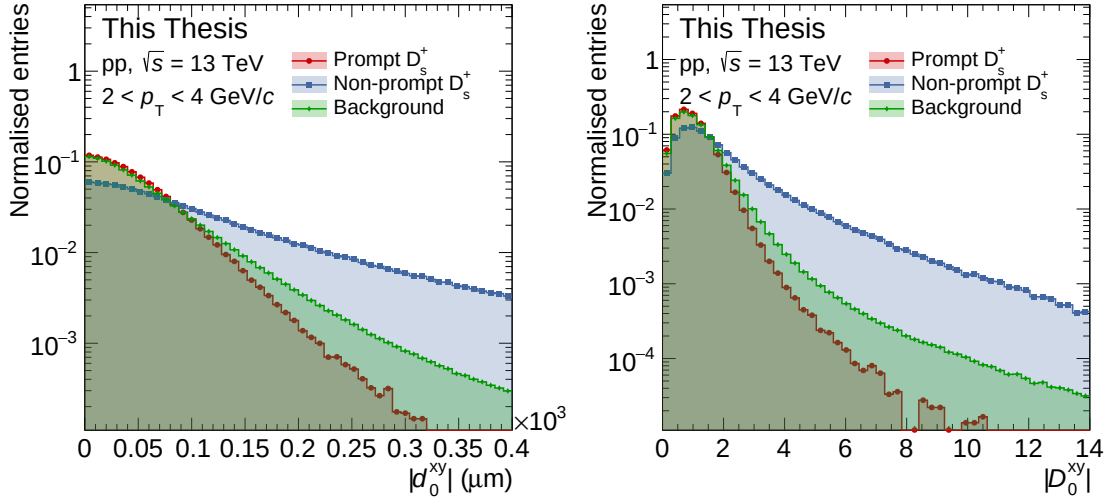


Fig. 4.5 Impact parameter in the transverse plane (left) and D_0^{xy} (right) distributions for D_s^+ mesons in pp collisions at $\sqrt{s} = 13$ TeV in $2 < p_T < 4$ GeV/c. The distributions are shown for prompt D_s^+ mesons (red), non-prompt D_s^+ mesons (blue), and combinatorial background (green). For the prompt and non-prompt D_s^+ mesons, the distributions are taken from Monte Carlo simulations, while for the combinatorial background they are taken from data excluding the signal region in the invariant mass distribution.

with respect to the D^0 mass and $\Delta M(\pi^+\pi^-)$ with respect to the mass of the K_S^0 is applied to reject the background in the case of the D^{*+} and K_S^0 meson candidates in the D_{s1}^+ and D_{s2}^{*+} resonances analyses.

In the case of the D_s^+ -meson candidates, two possible K^+K^- pairs can be built for each triplet of tracks, depending on the mass hypothesis assigned to the decay tracks. For both hypotheses, the K^+K^- invariant mass is computed and compared with the ϕ -meson mass. The smallest absolute difference is chosen as ΔM value. Therefore, this selection is extremely helpful not only in rejecting the combinatorial background, which is expected to show flat ΔM distribution, while the signal is expected to be peaked at zero, but also in rejecting the reflections, i.e. D mesons with the wrongly-assigned mass hypothesis (refer to Sec. 4.2).

The reflections have an invariant mass distribution that is different from that of the combinatorial background and can bias the yield extraction from the invariant mass fits by inflating the width of the signal peak. The ΔM selection, together with the PID selection presented in Sec. 4.2.5, is extremely important to ensure a proper signal extraction. The relative contribution of the reflected candidates to the D_s^+ signal in the invariant mass interval of $\pm 5\sigma$ around the D_s^+ mass peak is around 6% without PID

and ΔM selections, while it is around 0.1% after the application of these selections in pp collisions at $\sqrt{s} = 5.02$ TeV [235].

In addition, the selection on the ΔM variable is useful to reject the contribution of D mesons decaying in the same $K^+K^-\pi^+$ final state via intermediate resonances different from the one under study. In the case of the $D_s^+ \rightarrow \phi\pi^+ \rightarrow K^+K^-\pi^+$ decay channel, the largest contributions are expected to come from the $D_s^+ \rightarrow \bar{K}^{*0}(892)K \rightarrow K^+K^-\pi^+$ (BR = $(2.58 \pm 0.08)\%$) and $D_s^+ \rightarrow f_0(980)\pi^+ \rightarrow K^+K^-\pi^+$ (BR = $(1.14 \pm 0.31)\%$) [6] decay channels, given the similar mass of the ϕ , $\bar{K}^{*0}(892)$, and $f_0(980)$ mesons. The contamination arising from the $\bar{K}^{*0}(892)$ decay channel is expected to be negligible (less than 1%) according to recent efficiency studies [235]. It is more difficult to assess the magnitude of the contamination from the $f_0(980)$ decay channel due to the large uncertainty on the $f_0(980)$ mass and width according to the PDG [6]. Based on a preliminary measurement of the ALICE Collaboration, that measured $M(f_0(980)) \simeq 973$ MeV/ c^2 and $\Gamma(f_0(980)) \simeq 44$ MeV/ c^2 [242], a relatively small contamination ($\sim 2.9\%$ [235]) is expected.

The $\Delta M(K^+K^-)$ distribution is shown in the left panel of Fig. 4.6 for D_s^+ -meson candidates in pp collisions at $\sqrt{s} = 13$ TeV in the $2 < p_T < 4$ GeV/ c interval.

- **Cosine cubed of one of the D_s^+ -decay angles ($\cos^3(\theta'(K))$)**

In the decay of the (pseudoscalar) D_s^+ meson to the (vector) ϕ meson analysed the (pseudoscalar) pion, the spin of the ϕ meson is necessarily aligned with its direction of motion relative to the D_s^+ -meson flight line. It can be demonstrated that the distribution of the cosine of the angle $\theta'(K)$ between one of the kaons and the pion in the K^+K^- rest frame follows a $\cos^2(\theta'(K))$ shape, implying a flat distribution for $\cos^3(\theta'(K))$. In contrast, the $\cos(\theta'(K))$ distribution of the combinatorial background is flat and its $\cos^3(\theta'(K))$ distribution peaks at zero [243].

The $\cos^3(\theta'(K))$ distribution is shown in Fig. 4.6 for D_s^+ -meson candidates in pp collisions at $\sqrt{s} = 13$ TeV in the $2 < p_T < 4$ GeV/ c interval.

4.2.5 Particle identification selections

As explained in detail in Chapter 2, the signal measured by TPC and TOF can be used to associate the reconstructed tracks to a specific particle species. A track is considered to be compatible with a given mass hypothesis (electron, muon, pion, kaon, proton, deuteron, etc.) depending on the discrepancy between the measured $S_{\text{meas},(\text{TPC},\text{TOF})}$ and the expected $S_{\text{exp},(\text{TPC},\text{TOF})}$ signal S in the TPC or TOF. This is expressed in units of experimental resolution

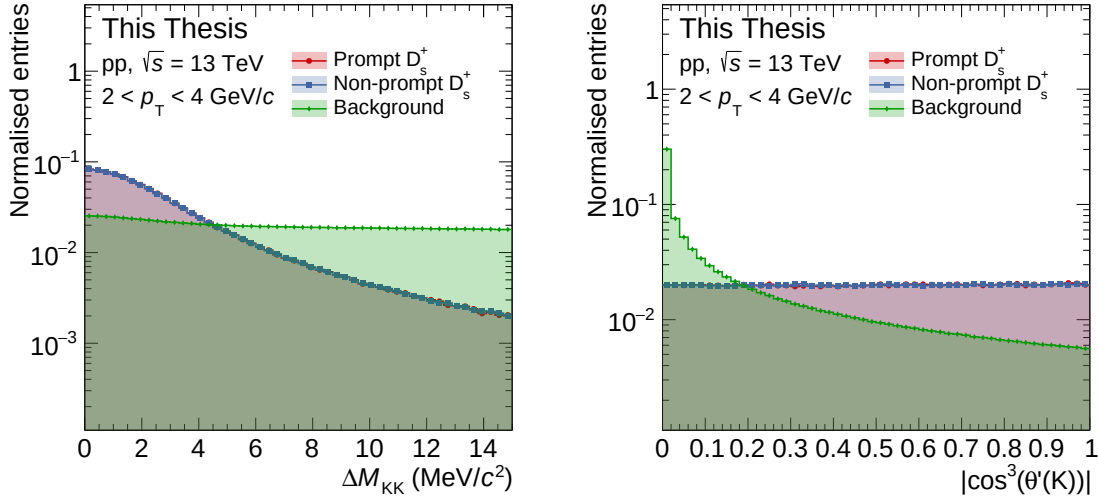


Fig. 4.6 Distributions of the absolute difference between the K^+K^- invariant mass and the PDG world average for the ϕ meson (left) and $\cos^3(\theta'(K))$ (right) for D_s^+ mesons in pp collisions at $\sqrt{s} = 13$ TeV in $2 < p_T < 4$ GeV/c. The distributions are shown for prompt D_s^+ mesons (red), non-prompt D_s^+ mesons (blue), and combinatorial background (green). For the prompt and non-prompt D_s^+ mesons, the distributions are taken from Monte Carlo simulations, while for the combinatorial background they are taken from data excluding the signal region in the invariant mass distribution.

σ as:

$$n\sigma_{\text{TPC,TOF}}^{\pi,K} = \frac{|S_{\text{meas, (TPC,TOF)}} - S_{\text{exp, (TPC,TOF)}}|}{\sigma_{\text{TPC,TOF}}} . \quad (4.6)$$

As described in detail in Sec. 2.6, the expected signal is computed considering a parametrisation of the Bethe-Bloch formula describing the energy loss versus rigidity for the TPC signal and the velocity versus momentum for the TOF signal [178].

The PID selections are based on the combination of the TPC and TOF PID information. In particular, the PID information in the D_s^+ -meson analysis is two-fold. First, a conservative "prefiltering" selection is applied to reject the most probable background candidates while preserving the signal. The simulated candidates passing the PID preselection constitute the MC sample employed in the ML part of the analyses. After that, the PID information is provided to the ML algorithm.

The conservative preselections are rectangular selections applied to the $n\sigma$ of TPC and TOF. For both the TPC and TOF detectors, a response value of -1, 0, or +1 is associated to each mass hypothesis. A response value equal to -1 is assigned if $n\sigma$ is larger than 3. For the TPC, an additional comparison with a threshold value ($n\sigma_{\text{TPC}}^{\text{max}}$) depending on the track momentum and the chosen PID strategy, either *conservative* or *strong*, is performed. If the $n\sigma$

value is lower than the threshold value, a response value of +1 is assigned, while a response value of 0 is assigned if the $n\sigma$ value is larger than the threshold value but still smaller than 3. In the case of conservative PID, a response value equal to 0 is assigned for tracks with $n\sigma$ lower than 3 and p_T higher than 1.5 GeV/c. A response value of +1 is assigned otherwise. This approach points to reject those tracks that do not fulfil the PID requirements in the TPC in a transverse momentum interval where the discrimination power of the TOF is not optimal. Track missing the TOF PID information (up to 30% – 40% for $p_T \lesssim 1$ GeV/c [235]) were assigned a response value of 0 but still considered to be compatible with the mass hypotheses; those tracks are considered in the analyses to avoid loss in the efficiency and are identified as *TPC-only* tracks. If the sum of the scores of the TPC and TOF detectors is larger or equal to 0 (or +1, depending on the adopted PID strategy), the mass hypothesis is accepted.

In the specific case of the D_s^+ mesons, the track with the charge opposite to that of the D_s^+ candidate is necessarily a kaon. Hence, candidates for which the track with the opposite charge is not compatible with the kaon mass hypothesis are rejected. Furthermore, a D_s^+ -meson candidate is rejected in case one of the tracks is not compatible with either the kaon or the pion mass hypothesis.

For the ML, the $n\sigma$ values from the TPC and TOF detectors are combined in a single variable, denoted as $n\sigma^{\pi,K}$, and defined as:

$$n\sigma^{\pi,K} = \begin{cases} \frac{1}{\sqrt{2}} \sqrt{(n\sigma_{\text{TPC}}^{\pi,K})^2 + (n\sigma_{\text{TOF}}^{\pi,K})^2} & \text{TPC PID only} \\ n\sigma_{\text{TPC}}^{\pi,K} & \text{TOF PID only} \\ n\sigma_{\text{TOF}}^{\pi,K} & \text{TOF PID only} \end{cases}$$

In this way, the number of features associated with the PID information is kept low (6), and candidates missing part of the information are properly handled.

4.3 Monte Carlo simulations

4.3.1 Event generation and particle transport

Specific Monte Carlo (MC) simulations were employed within the analysis procedure, e.g. to train the ML algorithm and optimise the ML-based selections and computing acceptance and efficiency corrections. In the context of pp collisions, the MC simulations were produced

with event generator PYTHIA 8.243 [172] with the Monash-2013 [244] tune. To enrich the sample of heavy-flavour hadrons, only events containing a $c\bar{c}$ or a $b\bar{b}$ pair in the $|y| < 1$ interval were considered. In addition, the generated D^0 , D^+ , and D_s^+ mesons were forced to decay in the decay channel of interest. For the analyses of non-prompt D_s^+ mesons, the decay of the beauty hadron (namely the mother) was not forced, to not introduce biases in the estimation of the non-prompt D_s^+ selection and reconstruction efficiencies. Notably, different MC samples were employed to train, test, and apply the ML algorithms in order to avoid biases (refer to Chapter 3). Using the same MC sample for training and testing the ML algorithm and for the final application of the ML-based selections can lead to an overestimation of the performance of the ML algorithm and inflate the estimation of the selection efficiency.

In case of heavy-ion collisions, the HIJING 1.36 generator [173] was used to simulate the underlying Pb–Pb event while PYTHIA 8 was used to inject pp events containing D meson signals using the same approach as for pp collisions. The number of PYTHIA events injected per HIJING event was tuned to not significantly affect the detector occupancy, which could lead to a deterioration of the tracking efficiency, and to ensure a reasonable computing time. For this reason, the number of PYTHIA events injected per HIJING event was set to 90 for the central collisions and 40 for the semicentral collisions.

In particular, the number of PYTHIA events injected per HIJING event increased with increasing collision centrality. For all collision systems, the generated particles were propagated through the ALICE apparatus using the GEANT3 transport package [174]. The simulation encompasses a detailed and complete description of the pp collisions starting from the beam conditions and their evolution during the data taking, the apparatus geometry, and the detector response in terms of resolution, noise, and dead channels. The simulation and reconstruction of MC data were performed within the AliRoot framework.

Potential discrepancies between the data and MC simulations concerning the topological and kinematic properties of interest in this Thesis may emerge. Such differences could be attributed to residual misalignment in the ITS detector and imperfect modelling of the detector response (in particular the space-charge distortions⁵ in the TPC, the tracks nearby the TPC sector edges, and the resolution on the variables used for PID) within the MC simulations. To ensure a satisfactory agreement between data and MC simulations, a dedicated process was applied to fine-tune the parameters and uncertainties related to single tracks in the MC

⁵The space-charge distortions consist of a modification of the electric field in the drift volume of the TPC due to the accumulation of positive ions, produced by the ionisation of the gas by the charged particles, and their motion (*ion backflow*) towards the centre of the chamber. This effect is particularly relevant in the high interaction rate conditions as the one delivered by LHC during the Run 3 data-taking period.

in addition to the track-quality selections described above. This adjustment is of paramount importance for analyses necessitating precision in single-track impact parameters at the micron level, as is the case of this Thesis. This procedure involves adjusting the impact parameter of each reconstructed track in the simulation to align it with the impact parameter resolution observed in the data. Implementing this correction minimises residual systematic bias in the efficiency estimation. A detailed description of this procedure is reported in Ref. [184]. Similarly, the $n\sigma$ PID variables were adjusted in MC simulations to match the distributions observed in the data.

4.3.2 Acceptance and efficiency corrections

The efficiency (ϵ) and acceptance (Acc) correction factors were derived from the aforementioned MC simulations. These corrections are necessary for determining the production yield of D_s^+ , D^+ , D^{*+} , and K_S^0 mesons at midrapidity ($|y| < 0.5$). This process corrects the raw yield obtained from the candidate reconstruction and subsequent application of topological, kinematic, and particle identification (PID) selections.

The efficiency was evaluated as a function of p_T as:

$$\epsilon(p_T) = \frac{N_{\text{reco}}(p_T) \Big|_{|y| < y_{\text{fid}}(p_T)}}{N_{\text{gen}}^{\text{acc}}(p_T) \Big|_{|y| < y_{\text{fid}}(p_T)}}, \quad (4.7)$$

where the numerator indicates the number of reconstructed mesons obtained after the application of selections. The denominator represents the number of generated mesons in the defined rapidity range $|y| < y_{\text{fid}}(p_T)$, decaying within the specific decay channel, and having all the daughter particles meeting the requirements of $p_T > 0.1$ GeV/ c and $|\eta| < 0.8$. Only events with a generated primary vertex positioned within ± 10 cm from the nominal interaction point along the z-direction are considered. The efficiency depends on the D-meson p_T . The p_T dependence is induced by the track reconstruction efficiency and by the selection criteria employed to reject the background.

The acceptance correction factor is computed as:

$$\text{Acc}(p_T) = \frac{N_{\text{gen}}^{\text{acc}}(p_T) \Big|_{|y| < y_{\text{fid}}(p_T)}}{N_{\text{gen}}(p_T) \Big|_{|y| < y_{\text{fid}}(p_T)}}, \quad (4.8)$$

where the denominator is the number of generated mesons in the fiducial rapidity region and decaying in the desired decay channel, and originating from events featuring a generated primary vertex with $|z| < 10$ cm.

An additional correction factor, denoted as $c_{\Delta y}$, is introduced to normalise the D-meson yield to one unit of rapidity, thus accounting for the rapidity coverage of the measurements in $|y| < y_{\text{fid}}(p_T)$. The $c_{\Delta y}$ factor was taken from PYTHIA simulations, and it is $c_{\Delta y} \simeq y_{\text{fid}}/0.5$ since the rapidity distribution of D-mesons in PYTHIA is nearly flat for $|y| < 0.8$. The $c_{\Delta y}$ factor was evaluated using FONLL pQCD calculations [245, 181], which accurately describe the rapidity dependence of D-meson cross sections in pp collisions, and found to be fully consistent with the one extracted from PYTHIA. The product of the $\varepsilon(p_T)$, the acceptance, and the rapidity coverage factors is finally used to correct the measured yields, as described in Chapter 5.

Chapter 5

Strange D-meson production in pp collisions at $\sqrt{s} = 13$ TeV

The measurement of the production cross sections for prompt and non-prompt D_s^+ mesons in pp collisions at LHC energies is a fundamental test for pQCD calculations [87]. Several pQCD-based calculations have been developed to describe the production of heavy-flavour hadrons in pp collisions, each of them implementing the factorisation as collinear factorisation, in terms of transferred momentum [80], or exploiting the transverse momentum of the parton, referred to as k_T [246, 90, 91]. These methods allow for the factorisation of the perturbative and non-perturbative terms in the cross section. In the collinear factorisation approach, the p_T -differential production cross section is computed as the convolution of three ingredients: (i) the parton distribution function (PDF) of quarks and gluons inside the incoming protons, (ii) the partonic cross section describing the hard scattering probability, which is computed with a perturbative approach, and (iii) the fragmentation function (FF) that parametrises the hadronisation of the heavy quark into the heavy-flavour hadron. The non-perturbative ingredients, i.e. the PDFs and the FFs, are parametrised on data as discussed in Chapter 1. Measurements of the p_T -differential production cross section of prompt and non-prompt D_s^+ mesons are an important benchmark for these calculations; in particular, they allow us to test the *universality* of the FFs. Indeed, the FFs are tuned on measurements in e^+e^- collisions and assumed to be independent of the collision system and the energy of the collision.

Typical calculations based on collinear factorisation are the GM-VFNS [88, 89] and the FONLL [86, 87] calculations. Both calculations provide a next-to-leading order (NLO)

accuracy with all-order resummation of next-to-leading logarithms. For the k_T -factorisation approach, the calculations can go beyond leading-order approximation in α_s [90–92].

Although these models successfully describe the production cross sections of heavy-flavour hadrons in pp collisions at several collision energies [181, 239, 241] and for different rapidity intervals [247, 248, 245], the uncertainty associated to their predictions is still larger than the experimental one. The main contributions to the theoretical uncertainty come from: the choice of the energy scales for the validity of the perturbative regime adopted in the factorisation approach, called *normalisation* (μ_R) and *factorisation* (μ_F) scales; the assumption on the heavy-quark mass (m_Q); and the uncertainties of the set of PDFs adopted in the calculation. As discussed in detail in Ref. [249], specific sensitivity to the PDF uncertainties can be obtained by computing the ratio of the p_T -differential D meson production at different energies scales and at different rapidities. In the ratio, the contribution related to μ_R and μ_F , and the choice of m_Q are subdominant, allowing for a better constraint of the PDFs [97]. In addition, the measurements of the ratios of the p_T -differential cross sections of different hadron species allow for a direct study of the hadronisation mechanisms of heavy quarks. In particular, ratios of strange-to-non-strange D-meson yields (D_s^+/D^0 and D_s^+/D^+ for prompt and non-prompt mesons) allow us to investigate the *fragmentation fractions* of charm and beauty quarks into strange (f_s) and non-strange (f_u, f_d) mesons. Although the assumption of the FF universality is so far supported by experiments in the meson sector [239], in the baryon sector the fragmentation fractions are found to be different in pp and e^+e^- collisions both for charm and beauty quarks [250]. These findings highlight the necessity to extend the current measurements to different collision systems and energy scales.

The p_T -differential cross section of (non-)prompt D_s^+ mesons is calculated as:

$$\left. \frac{d^2\sigma_{(\text{non-})\text{prompt}}^D}{dp_T dy} \right|_{|y|<0.5} = \frac{1}{2} \frac{1}{\Delta p_T} \frac{f_{(\text{non-})\text{prompt}}^{\text{raw}} N_{|y|<y_{\text{fid}}}^{D+\bar{D}}}{c_{\Delta y}(\text{Acc} \times \epsilon)_{(\text{non-})\text{prompt}}} \cdot \frac{1}{\text{BR}} \cdot \frac{1}{\mathcal{L}_{\text{int}}} . \quad (5.1)$$

The (non-)prompt D_s^+ -meson candidates are selected via machine-learning-based selection criteria, as described in Sec. 5.1. The term $N^{D+\bar{D}}_{|y|<y_{\text{fid}}}$ represents the (non-)prompt D_s^+ -meson raw-yields, which are obtained from the fitting procedure to the invariant-mass distribution of the D_s^+ candidates passing the selections, as discussed in detail in Sec. 5.2.1. The raw-yields are divided by 2 to average the yields of particles and antiparticles and corrected by the acceptance-times-efficiency term presented in Sec. 5.2.2. The efficiency is different for prompt and non-prompt mesons, due to their different decay topology. An additional correction is applied to consider the fraction of prompt ($f_{\text{prompt}}^{\text{raw}}$) and non-prompt

($f_{\text{non-prompt}}^{\text{raw}}$) D_s^+ mesons in the selected sample. The computation of these fractions is described in detail in Sec. 5.2.3. Finally, the production cross section in each p_T interval is derived by dividing by the width of the analysed p_T interval (Δp_T), the branching ratio for the decay channel used for signal reconstruction (BR) (refer to Sec. 4.2), and the integrated luminosity ($\mathcal{L}_{\text{int}}$). The procedure adopted to estimate the systematic uncertainties associated to the measurement of the p_T -differential cross section of prompt and non-prompt D_s^+ mesons is reported in Sec. 5.3. The results related to the prompt D_s^+ meson production were published in Ref. [97], while the results related to the non-prompt D_s^+ meson production, although already approved by the ALICE Collaboration, are reported here for the first time and will be published in the near future.

The D_s^+ meson, or F meson as it was called when first discovered in 1983 [251], is the ground state of the widely populated charm-strange meson family reported in Fig. 5.1. Indeed, over the past decade, there has been a revival in the field of charm meson spectroscopy. This resurgence can be attributed to the discovery of numerous excited charm and charm-strange states by experiments such as BABAR and Belle at B-Factories [252–256], and by the CLEO experiment [257, 258]. In the past few years, the LHCb experiment at the LHC has demonstrated its capability to observe these states and to accurately determine their properties [222, 221, 223, 224]. Consequently, there has been a significant interest from theoreticians in assigning spectroscopic properties to these new states by comparing their predictions with the experimental measurements [259–267]. This gave rise to a plethora of QCD-based models, all of which focusing on the spectroscopy of charm-meson states; at the same time there have been less studies of the production mechanisms of these states. However, the production of excited charm meson states is a fundamental test for QCD in the non-perturbative regime and could provide complementary information to spectroscopy. In addition, due to their short lifetime (refer to Chapter 4) some of the excited mesons are expected to decay in the hadronic phase of the collision, enabling the study of this intermediate phase also in small collision systems as discussed in Chapter 4 and Ref. [47]. In Sec. 5.5, the first measurement of the production of the D_{s1}^+ and D_{s2}^{*+} mesons in pp collisions at $\sqrt{s} = 13$ TeV at the LHC as a function of the charged particle multiplicity is presented.

The p_T -integrated production yields of prompt D_{s1}^+ and D_{s2}^{*+} mesons is computed as:

$$\frac{d^2 N^{\text{D}_{\text{reso.}}}}{dp_T dy} = \frac{f_{\text{prompt}} \cdot N_{\text{raw}}^{\text{D}_{\text{reso.}}} \cdot \epsilon^{\text{trigger}}}{2 \cdot \Delta p_T \cdot c_{\Delta y} \cdot (\text{Acc} \times \epsilon)_{\text{prompt}} \cdot \text{BR}(D_{\text{daugh.}}) \cdot \text{BR}(V0) \cdot N_{\text{ev}}} \cdot \quad (5.2)$$

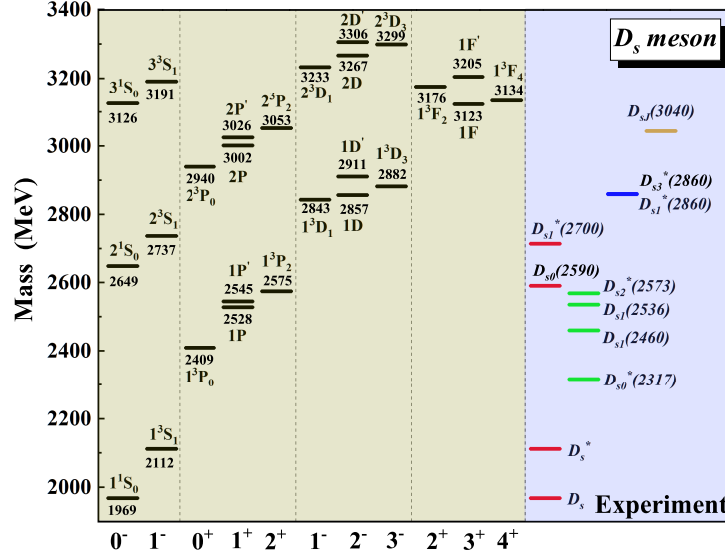


Fig. 5.1 Predicted spectroscopy of the strange D-meson states compared to the experimental measurements. Figure from Ref. [268].

The $\text{BR}(D_{\text{daugh.}}) \cdot \text{BR}(V0)$ term represents the product between the branching ratios of the decay channel $D^{*+} \rightarrow D^0(\rightarrow K^-\pi^+)\pi^+$ ($\text{BR} = (2.67 \pm 0.02)\%$ [6]) or $D^+ \rightarrow K^-\pi^+\pi^+$ ($\text{BR} = (9.38 \pm 0.16)\%$ [6]), depending on the resonance under study, times the one of the $K_S^0 \rightarrow \pi^+\pi^-$ ($\text{BR} = (69.20 \pm 0.05)\%$ [6]) decay. The $\epsilon^{\text{trigger}}$ term instead indicates the trigger efficiency of the multiplicity class considered in the measurement computed as the ratio of the number of event passing the event selections discussed in Sec. 4.1.2 but not selected by the triggers and the total number of selected events estimated in PYTHIA 8 and GEANT 3 simulations, as described in detail in Ref. [269], and corresponding to 0.92 and 1 for the low- and high-multiplicity intervals, respectively [270]. The selection criteria adopted in the analysis, as well as the procedure employed for the estimation of the ingredients of Eq. 5.2, are described in detail in Sec. 5.5.

5.1 Machine learning for the D_s^+ meson signal extraction

In the analysis of the production of prompt and non-prompt D_s^+ mesons, a machine learning approach is adopted for the selection of the D_s^+ candidates. As discussed in detail in Chapter 3, ML algorithms are a powerful tool to optimise the selection criteria to reject

the background candidates while preserving the signal. A detailed comparison between the ML-based approach and the traditional linear-selection method is reported in Ref. [240] and demonstrates a significant improvement in terms of precision of the results obtained with ML-based selections. In the context of this Thesis, a multiclass BDT algorithm is employed to classify D_s^+ -meson candidates among three possible categories: prompt D_s^+ mesons, non-prompt D_s^+ mesons, and background. A similar strategy is adopted for the selection of the D^+ and D^0 mesons produced in the decay of D_{s1}^+ and D_{s2}^{*+} mesons in the measurement of the production of the excited charm-strange meson states.

5.1.1 Training, testing, and application sets

The datasets used for training the BDT and evaluating its performance encompass both signal candidates, composed of prompt and non-prompt D_s^+ -meson candidates, and background candidates. The background candidates were extracted from the "sideband" region of the D-meson candidate invariant-mass distributions in data. The sideband regions are defined as the invariant mass intervals $M > 5\sigma$ with respect to the nominal value of the D_s^+ mass and $M < 5\sigma$ with respect to the nominal value of the D^+ mass (i.e. $M < 1.83 \text{ GeV}/c^2$ and $2.01 > M \text{ GeV}/c^2$). The sideband regions are chosen to reject the signal coming from D_s^+ mesons decaying in the $D_s^+ \rightarrow \phi\pi^+ \rightarrow K^+K^-\pi^+$ decay channel as well as D^+ mesons decaying in the same decay channel ($\text{BR} = 2.69_{-0.08}^{+0.07} \times 10^{-3}$ [6]). If not excluded, the latter contribution would bias the training procedure and ultimately worsen the separation power of the BDT. The choice of using real data in the training sample is aimed at minimising the effects of potential shortcomings deriving from an imperfect description of real data in the simulation. The accuracy of modelling the combinatorial background heavily depends on how well the MC generator reproduces several aspects of the collision from the physics of the collision (e.g. charged particle multiplicity, kinematic distributions of the produced particles, relative abundance of different particle species, etc.) to the description of the experimental setup (e.g. detector geometry, material budget, detector response, occupancy and dead channels, etc.). Instead, for the signal candidates were obtained from a dedicated MC simulation using the PYTHIA 8 event generator [172], as described in detail in Sec. 4.3.2. In this case, the MC was used because it allows to tag the candidates originating from true D_s^+ decay and separating prompt and non-prompt D_s^+ mesons.

The candidates composing the training sample were required to satisfy the loose kinematic and topological selections reported in Table. 5.1, in addition to the single-track selection criteria detailed in Sec. 4.2.3. They were also subject to a conservative PID strategy, as described in Sec. 4.2.5. The training of the BDT algorithm was conducted in broader p_T

Table 5.1 Topological and kinematic pre-selections, applied to compose the training sample, for the D_s^+ analysis in pp collisions at $\sqrt{s} = 13$ TeV.

p_T (GeV/c)	[1,50]
decay length (cm)	> 0.08
$\cos(\theta_p)$	> 0.85
$\cos(\theta_p^{xy})$	> 0.85
σ_{vertex} (cm)	< 0.08
$M_{\text{inv}}^\phi - M_{\text{PDG}}^\phi$ (MeV/c ²)	< 15

Table 5.2 Number of signal (prompt and non-prompt D_s^+) and background candidates used to build the training and test sets.

p_T interval (GeV/c)	prompt D_s^+	non-prompt D_s^+	combinatorial background
1-2	$\sim 160 \cdot 10^3$	$\sim 360 \cdot 10^3$	$\sim 150 \cdot 10^5$
2-4	$\sim 140 \cdot 10^3$	$\sim 350 \cdot 10^3$	$\sim 130 \cdot 10^5$
4-6	$\sim 790 \cdot 10^3$	$\sim 370 \cdot 10^3$	$\sim 174 \cdot 10^3$
6-12	$\sim 356 \cdot 10^3$	$\sim 456 \cdot 10^3$	$\sim 203 \cdot 10^3$
12-50	$\sim 468 \cdot 10^3$	$\sim 371 \cdot 10^3$	$\sim 210 \cdot 10^3$

intervals compared to those used in the analysis. In this way, the training sample was sufficiently large to ensure proper training for each class. This is particularly important at high p_T where the number of D_s^+ -meson examples is typically lower compared to low p_T and the risk of an extremely imbalanced composition of the training sample is higher. All the prompt and non-prompt D_s^+ meson examples available from the dedicated MC production were considered in the training process. To provide a training sample with a balanced composition of background and signal examples, the number of background candidates was roughly matched to that of the signal candidates at low p_T . Conversely, at higher p_T , where the number of candidates in the sideband regions is lower compared to the low- p_T region, fewer background candidates were considered in the training compared to the low p_T intervals to avoid eventual biases in the final measurement. Following a widely accepted approach, the training dataset was split into two subsets: the training set, comprising 80% of the full training sample, and the test set, containing the remaining 20% of available candidates, which is used to assess the quality of the predictions of the trained algorithms. The number of signal and background candidates used for the training and test sets are summarised in Tab. 5.2.

5.1.2 Training features and BDT performance evaluation

The variables employed in the BDT selections and training are the ones introduced in Secs. 4.2.4 and 4.2.5. It is really important to investigate the linear correlation among each variable pair in order to spot possible correlations between the chosen training variables and the D_s^+ meson mass and transverse momentum. The last two quantities are excluded from the training since they are strictly related to the possibility of extracting the raw yield from the invariant-mass distribution and to the p_T -differential production cross section of the D_s^+ mesons, therefore a selection based on these variables would bias the measurement. If a strong correlation is found between one of the training variables and the p_T or the mass of the candidate, the BDT may learn to discriminate the signal from the background by exploiting the correlation with the D_s^+ meson mass and transverse momentum, rather than the physical properties of the signal and background. This effect could eventually modify the shape of the invariant mass distribution introducing a bias in the measurement. The correlation coefficient between two variables is defined as the Pearson correlation coefficient ρ [271]. The ρ coefficient ranges from -1 to 1, where 1 indicates a total positive linear correlation, 0 corresponds to a non-linear correlation, and -1 indicates a total negative linear correlation.

Figure 5.2 shows the ρ coefficient for all pairs of training variables, and the candidate invariant mass and transverse momentum for the prompt D_s^+ , non-prompt D_s^+ , and the combinatorial background examples in the $2 < p_T < 4$ GeV/ c interval in the training sample. The correlation matrices show that the training variables are not correlated with the D_s^+ meson mass or transverse momentum. Consequently, it is unlikely that the BDT selection would modify the invariant-mass distribution of the selected candidates and ultimately introduce a bias in the measurement. In addition, the correlation matrices show that variables sharing similar physical information are positively correlated, as expected.

As described in detail in Chapter 3, within the training-test-validation procedure it is extremely important to optimise the internal parameters of the BDT, called hyperparameters. To this end, the Bayesian optimisation method was employed (refer to Sec. 3.1.3), and a 5-fold cross-validation was conducted to assess the performance of the BDT with the specific set of hyperparameters. An example of the ROC AUC obtained for the prompt D_s^+ meson measurement in the $4 < p_T < 6$ GeV/ c interval is reported in Fig. 5.3. The ROC AUC values for the various models are all close to unity indicating a good performance of the BDT. The optimised hyperparameters for the lowest and highest p_T intervals of the training set for the prompt and non-prompt D_s^+ meson production study are reported in Table 6.3.

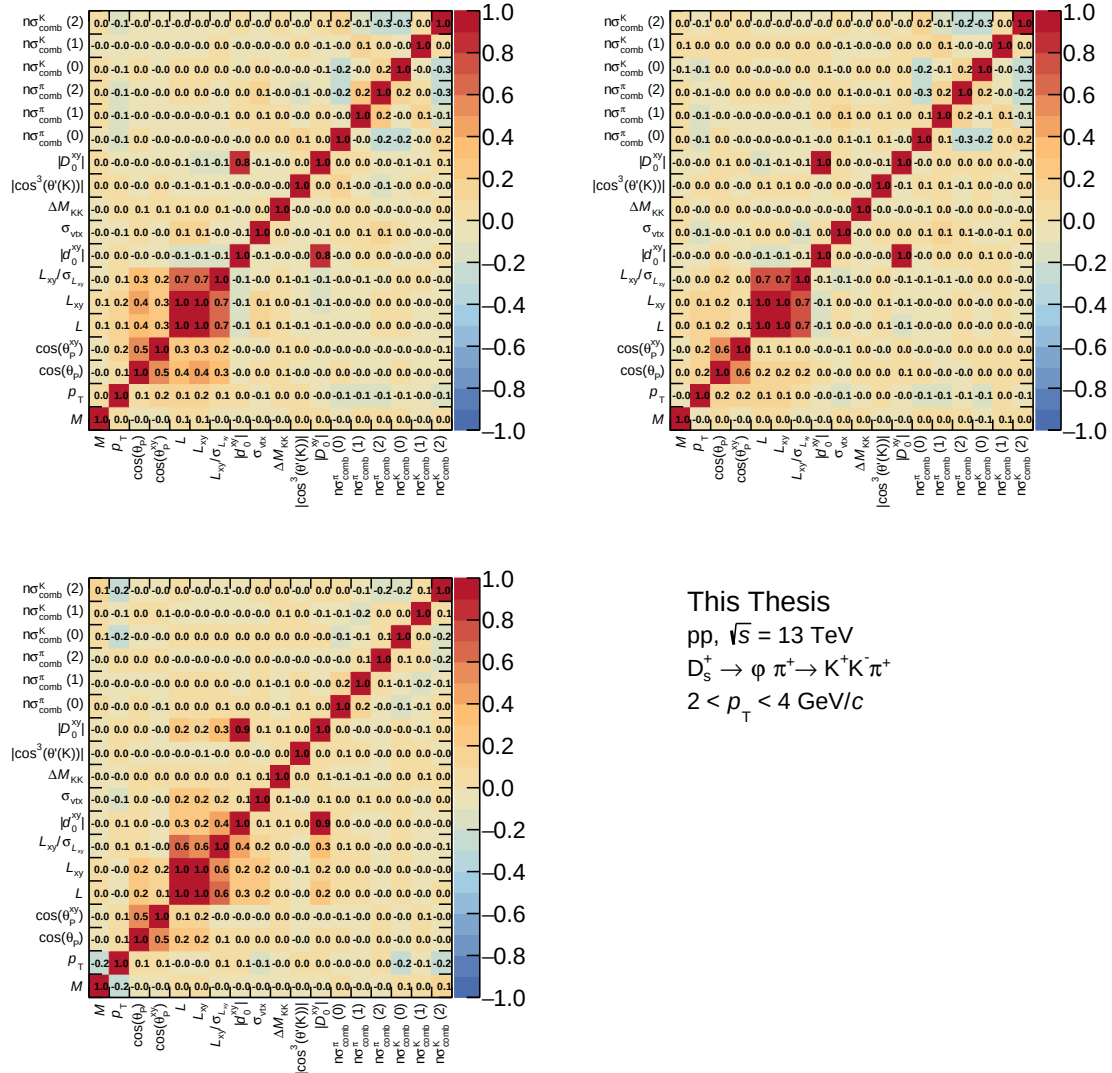
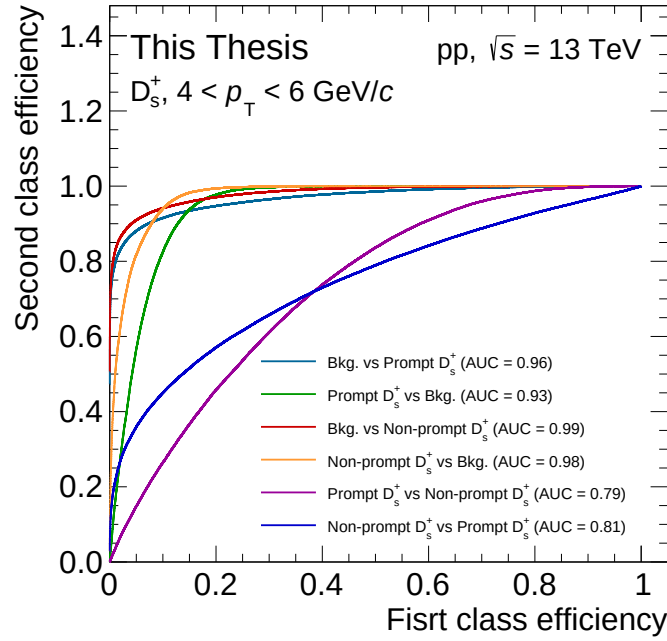


Fig. 5.2 Correlation matrices among the training variables and the D_s^+ -meson invariant mass and transverse momentum for the prompt D_s^+ (top-left), non-prompt D_s^+ (top-right), and the combinatorial background (bottom-left) examples in the $2 < p_T < 4$ GeV/c interval in the training sample.

After the training and test, the BDT can be finally applied to the dataset under analysis. In this application, the BDT assigns to each candidate of the dataset three different output scores (one for each class) representing the probability of the candidate belonging to each of the three classes. In the case of the analyses reported in this Chapter, the BDT output scores indicate the probability of the candidate being a prompt D_s^+ meson, a non-prompt D_s^+ meson, or background. The BDT scores sum up to unity by definition, hence one out of the three scores is completely determined from the values of the other two.

Fig. 5.3 ROC AUC obtained for the prompt D_s^+ meson test set in the $4 < p_T < 6$ GeV/c interval.Table 5.3 Hyperparameter configuration for the p_T bins considered in the training for the D_s^+ mesons.

Hyperparameter	p_T interval (GeV/c)					
	1 – 2	2 – 4	4 – 6	6 – 8	8 – 12	12 – 50
n_estimators	747	511	511	511	772	492
learning_rate	0.04	0.027	0.027	0.027	0.081	0.036
max_depth	3	4	4	4	3	3
min_child_weight	9.5	8.0	8.0	8.0	4.5	2.3
subsample	0.84	0.90	0.90	0.97	0.82	0.87
colsample_bytree	0.98	0.90	0.90	0.91	0.83	0.87
tree_method	hist	hist	hist	hist	hist	hist

The distribution of the BDT output scores obtained in the prompt-enhanced D_s^+ study in the $2 < p_T < 4$ GeV/c interval is shown in Fig. 5.4. As expected, due to the larger displacement from the primary vertex (refer to Chapter 4), the non-prompt D_s^+ mesons are more easily separated from the background as compared to the prompt ones and consequently receive high scores related to the hypothesis of being non-prompt rather than the other hypotheses. Analogously, background candidates are typically composed of tracks originating from the primary vertex and have much shorter distance between the primary and secondary vertices compared to signal candidates. Hence, they reach higher scores for the background class compared to the other possibilities. The prompt D_s^+ mesons are on average more displaced

than the background ones, though they are less displaced than non-prompt D_s^+ mesons. Consequently, the BDT struggles to distinguish prompt from non-prompt D_s^+ mesons, more than simply distinguishing signal from background. Indeed, the prompt D_s^+ mesons receive intermediate scores for the non-prompt hypothesis and cannot get the maximum score for the prompt hypothesis. This is also related to the xgboost implementation of the multiclass BDT where a series of BDT algorithms, in the same number of the classes under consideration, are trained. Each one of these algorithms handles a single one-vs-rest classification task. The final output scores are generated by these different models and then properly normalised. The score associated with the prompt D_s^+ -meson hypothesis is derived from a classification that distinguishes prompt candidates (one class) from non-prompt D_s^+ and combinatorial background candidates (the rest of the classes). This combination of background and non-prompt D_s^+ -meson candidates, which exhibit distinct characteristics, poses a greater challenge for discrimination compared to other combinations [240].

The distributions of the BDT output scores are employed to define the selection criteria for D_s^+ meson candidates. The selection is performed by applying a threshold to two of the BDT output score distributions, namely the one associated to the prompt or non-prompt D_s^+ -meson hypotheses and to the one associated to the probability to be background. The thresholds are optimised following the procedure described in detail in the next section.

5.1.3 BDT-based selection optimisation

The choice of the threshold value to be imposed to the BDT output score is a crucial step of the analysis and has to be carefully studied. Given that for each D_s^+ -meson candidate the three BDT output scores collectively sum up to 1, by enforcing a criterion on two scores, out of three, a selection is also applied to the third score by definition. Indeed, at the analysis level, only two BDT-based selections are simultaneously applied. The first selection is typically applied to the probability of the candidate being background, while the second selection is applied to the probability of the candidate being either a prompt or non-prompt D_s^+ meson, depending on the analysis goal. The first selection rejects the large contribution of the background, while the second one enhances the contribution of prompt (*prompt-enhanced*) or non-prompt (*non-prompt-enhanced*) D_s^+ mesons in the selected sample.

To optimise both the BDT-based selections, the two thresholds are iteratively and independently varied to explore a wide spectrum of configurations. For each configuration, a set of physical variables, such as the expected signal ($S_{\text{exp.}}$), the expected background ($B_{\text{exp.}}$), the expected signal-to-background ratio and significance ($S_{\text{exp.}}/\sqrt{S_{\text{exp.}} + B_{\text{exp.}}}$), and the prompt

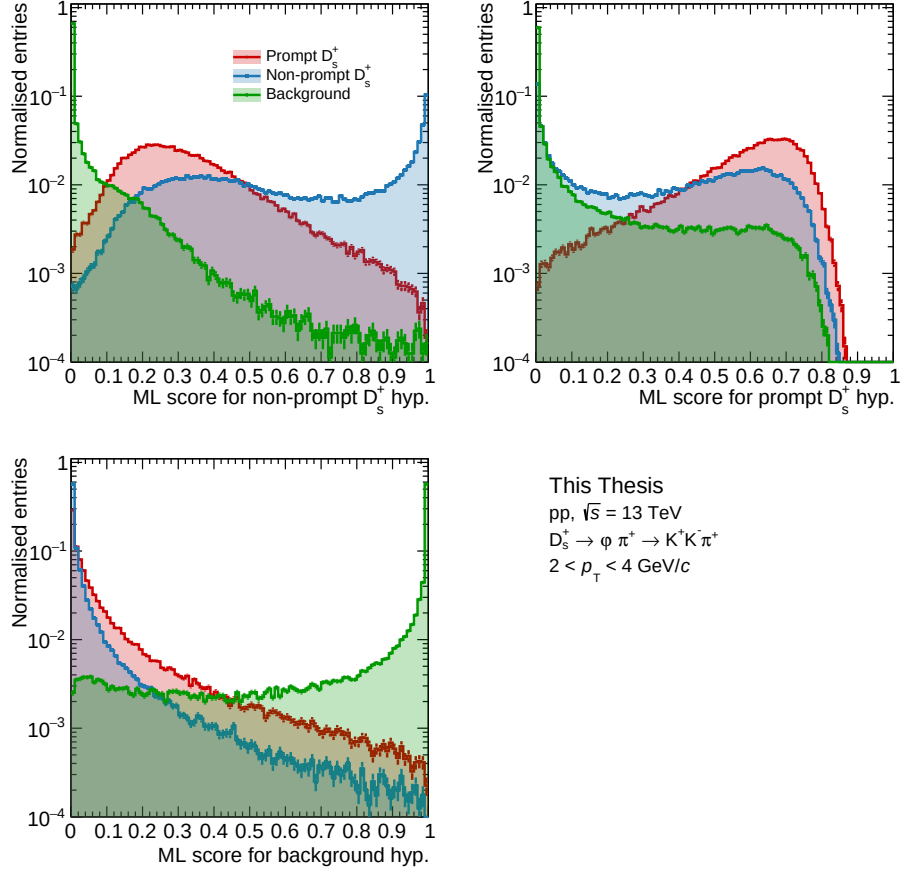


Fig. 5.4 Distribution of the BDT output scores obtained in the D_s^+ meson training sample in the $2 < p_T < 4$ GeV/c interval for the non-prompt D_s^+ (top-left), prompt D_s^+ (top-right), and background (bottom-left) hypotheses.

and non-prompt D_s^+ -meson efficiencies and fractions, are computed as explained below. This procedure is repeated separately for each p_T interval considered in the analyses.

The expected signal yields, normalised per event, are computed exploiting the FONLL predictions [86] as:

$$S_{\text{exp.}} = \left. \frac{d^2\sigma^{D_s^+}}{dp_T dy} \right|_{\text{FONLL}} \times \frac{2(\text{Acc} \times \varepsilon)^{(\text{non-})\text{prompt}} c_{\Delta y} \Delta p_T \text{BR} \mathcal{L}_{\text{int}}}{f_{(\text{non-})\text{prompt}}}, \quad (5.3)$$

where $d^2\sigma/(dp_T dy)|_{\text{FONLL}}$ is the FONLL [86, 87] prediction for the D_s^+ meson p_T -differential production cross section, $\mathcal{L}_{\text{int}}$ is the integrated luminosity, $\varepsilon^{(\text{non-})\text{prompt}}$ is the (non)-prompt D_s^+ -meson selection efficiency, BR is the branching ratio of the D_s^+ meson decay channel under consideration, $c_{\Delta y}$ is the correction factor for the rapidity interval under consideration,

Table 5.4 Fragmentation fractions of beauty quarks into beauty hadrons measured in $Z^0 \rightarrow b\bar{b}$ decays [274] and in pp collisions at $\sqrt{s} = 1.96$ TeV [6].

Beauty hadron	Fraction at Z (%)	Fraction at $p\bar{p}$ (%)
B^0, B^+	40.8 ± 0.7	34.4 ± 2.1
B_s^0	10.0 ± 0.8	11.5 ± 1.3
Λ_b^0	8.4 ± 1.1	19.8 ± 4.6

and Δp_T is the width of the p_T interval. In the case of prompt D_s^+ mesons, no FONLL predictions are currently available, thus the cross section is computed following the procedure described in Ref. [240, 234] using the charm-quark cross section predicted by FONLL, the charm quark fragmentation functions from Ref. [272], and the fragmentation fractions into D_s^+ mesons from e^+e^- collisions by the ALEPH experiment [273] under the assumption of the universality of the fragmentation fraction. In the case of non-prompt D_s^+ mesons, the FONLL predictions are computed using the beauty-hadron cross section from FONLL, while the decay kinematics of beauty-hadrons to D_s^+ mesons are computed using PYTHIA 8 [172]. The fragmentation fractions of beauty quarks into beauty hadrons ($f_b \rightarrow h_b$) measured in $Z^0 \rightarrow b\bar{b}$ decays [274] were used to weight the contribution from the different beauty-hadron species. The values of the beauty-quark fragmentation fractions for several beauty-hadron species measured in $Z^0 \rightarrow b\bar{b}$ decays [274] and in pp collisions at $\sqrt{s} = 1.96$ TeV [6] are reported in Table 5.4, and are widely employed by the ATLAS, CMS, and LHCb Collaborations to normalise the B-meson measurements [236, 237, 139].

The expected background yields, normalised per event, were estimated by fitting the sidebands of the invariant-mass distributions of the candidates in the data with an exponential function. The expected background yields are then computed by integrating the fitted function in the invariant-mass interval of the signal region, defined as a 3σ interval around the world-average D_s^+ meson mass [6]. The σ parameter, accounting for the invariant mass resolution, is obtained by the fit to the invariant-mass distribution of the D_s^+ candidates in the MC simulation with a Gaussian function. To avoid correlations between statistical fluctuations in the estimated significance and the measurement, only a fraction of about 10% of the available data was used in the computation of the expected background. Finally, the expected background value was scaled to match the full data sample collected by the experiment.

The $(\text{Acc} \times \epsilon)$ term in Eq. 5.3 represents the product of the geometrical acceptance and the selection efficiency of the prompt or non-prompt D_s^+ mesons which is computed by using a different MC sample with respect to the one employed for the BDT training. The MC sample used for the computation of the $(\text{Acc} \times \epsilon)$ contribution is the same used for the

computation of the $(\text{Acc} \times \varepsilon)$ correction factors described in Sec. 4.3.2. The fraction of prompt D_s^+ mesons, f_{prompt} , is computed as:

$$f_{\text{prompt}} = \left[1 + \frac{(\text{Acc} \times \varepsilon)^{\text{non-prompt}} \frac{d^2\sigma}{dp_T dy} \Big|_{\text{FONLL+PYTHIA 8}}}{(\text{Acc} \times \varepsilon)^{\text{prompt}} \frac{d^2\sigma}{dp_T dy} \Big|_{\text{FONLL}}} \right]^{-1}. \quad (5.4)$$

The fraction of non-prompt D_s^+ mesons, $f_{\text{non-prompt}}$, is computed as $f_{\text{non-prompt}} = 1 - f_{\text{prompt}}$.

The statistical significance is a proxy of the inverse of the statistical uncertainty on the signal yield, and therefore the higher the significance the smaller the uncertainty. On the other hand, high efficiency and a high prompt (or non-prompt) D_s^+ fraction ensure a high sensitivity to the production of prompt or non-prompt D_s^+ mesons and a smaller sensitivity to possible imperfections in the MC simulations affecting the $(\text{Acc} \times \varepsilon)$ factors. In case of loose BDT-base selections, the significance will be low but, at the same time, the efficiency will be high. Conversely, in case of tight BDT-based selections, the significance will be high but the efficiency and prompt or non-prompt D_s^+ fraction will be low. The optimal selection is the one that provides a large expected statistical significance, while ensuring a sufficiently high selection efficiency and prompt (or non-prompt) D_s^+ fraction.

Figure 5.5 displays the expected significance, the expected signal-to-background ratio, and the non-prompt fraction as a function of the BDT output score thresholds for the probability to be prompt and background in the $2 < p_T < 4$ GeV/c interval. The chosen working point is represented by a black marker in the Figure.

5.2 Prompt and non-prompt D_s^+ meson production in pp collisions at $\sqrt{s} = 13$ TeV

5.2.1 Raw-yield extraction

The raw yields of D_s^+ mesons, calculated as the sum of particles and antiparticles, are measured in the range $1 < p_T < 36$ GeV/c for prompt mesons and $1 < p_T < 24$ GeV/c for non-prompt mesons in the p_T intervals reported in Tables 5.5 and 5.6, respectively. The raw yields are extracted performing a binned maximum-likelihood fit of the invariant-mass distributions of the D_s^+ -meson candidates passing the BDT selection criteria optimised according to the procedure presented above. The fit operation is performed separately for the prompt- and non-prompt-enhanced samples. The signal peak is parametrised with a Gaussian

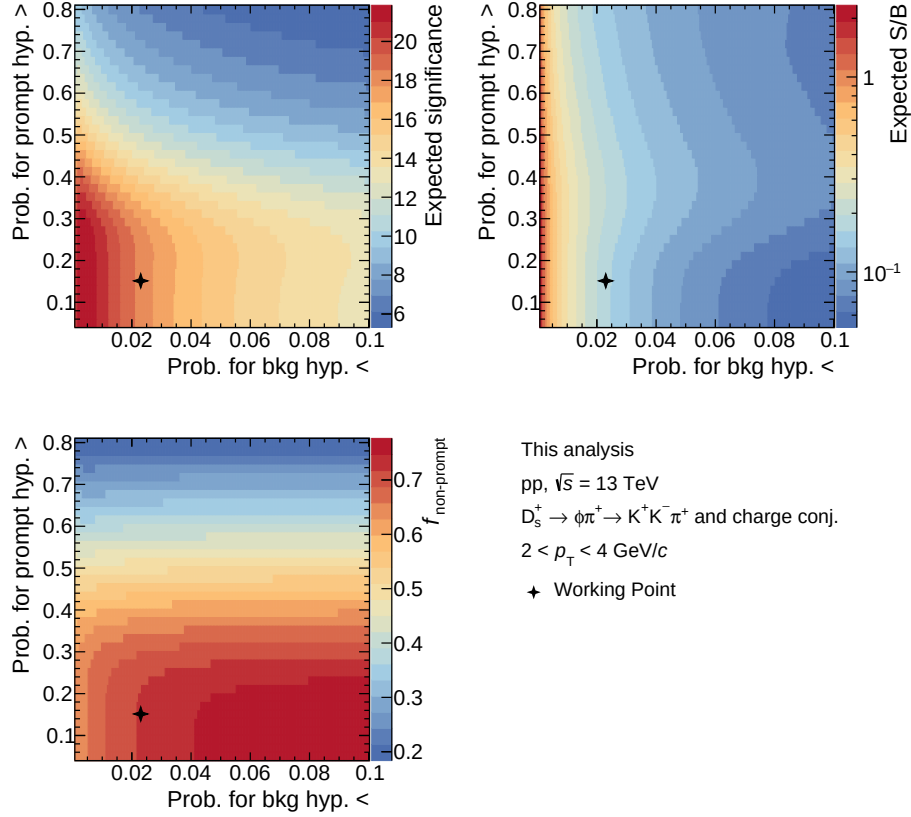


Fig. 5.5 Expected significance (top-left), expected signal-to-background ratio (top-right), and non-prompt fraction (bottom) as a function of the BDT output score thresholds for the probability to be prompt and background in the $2 < p_T < 4$ GeV/c interval.

function. In the case of the prompt-enhanced sample, the Gaussian width is a free parameter in the fit. Instead, in the case of the non-prompt-enhanced sample where the number of D_s^+ candidates is naturally rather limited¹, the Gaussian width is fixed to the one extracted from the fits to the prompt-enhanced sample. This is done to improve the stability of the fit thanks to the higher statistical significance of the prompt-enhanced sample compared to the non-prompt-enhanced one. The shape of the background is modelled with an exponential function.

Given that D^+ mesons decay in the $D^+ \rightarrow \phi \pi^+ \rightarrow K^+ K^- \pi^+$ channel ($\text{BR} = 2.69^{+0.07}_{-0.08} \times 10^{-3}$ [6]), an independent Gaussian function is adopted to model the D^+ peak in the invariant-mass distribution. The width of the D^+ peak ($\sigma(D^+)_{\text{data}}$) is fixed to the width of the D_s^+ peak ($\sigma(D_s^+)_{\text{data}}$) scaled by the ratio of the D^+ and D_s^+ meson widths in the MC simulation ($\sigma(D^+)_{\text{MC}}/\sigma(D_s^+)_{\text{MC}}$). This choice was made under the assumption that the relative

¹In Nature the abundance of prompt D_s^+ corresponds to $\simeq 95\%$ of the produced D_s^+ mesons.

Table 5.5 Raw yields, signal-to-background ratios, and significance of prompt D_s^+ mesons in pp collisions at $\sqrt{s} = 13$ TeV. A 3σ interval around the D_s^+ peak is used to extract the values.

p_T interval (GeV/c)	Raw-yield (3σ)	S/B (3σ)	Significance (3σ)
1-2	251 ± 46	0.14	5.6 ± 1.0
2-3	1014 ± 72	0.25	14.3 ± 0.9
3-4	1484 ± 75	0.44	21.5 ± 0.9
4-5	1726 ± 73	0.52	24.3 ± 0.9
5-6	1276 ± 55	0.76	23.5 ± 0.8
6-8	1582 ± 53	1.6	31.3 ± 0.7
8-12	1238 ± 59	0.67	22.3 ± 0.9
12-16	268 ± 23	1.18	12.0 ± 0.8
16-24	152 ± 18	0.66	7.8 ± 0.8
24-36	45 ± 9	1.087	4.8 ± 0.7

Table 5.6 Raw yields, signal-to-background ratios, and significance of non-prompt D_s^+ mesons in pp collisions at $\sqrt{s} = 13$ TeV. A 3σ interval around the D_s^+ peak is used to extract the values.

p_T interval (GeV/c)	Raw yield (3σ)	S/B (3σ)	Significance (3σ)
2-4	201 ± 21	0.62	8.7 ± 0.7
4-6	370 ± 24	1.56	15.0 ± 0.7
6-8	302 ± 20	2.56	14.7 ± 0.6
8-12	203 ± 116	2.75	12.2 ± 0.6
12-24	46 ± 8	2.42	5.7 ± 0.6

variation of the mass resolution of the D^+ and D_s^+ mesons is the same in data and MC simulation. Indeed, the typical value of the $\sigma(D^+)_{MC}/\sigma(D_s^+)_{MC}$ ratio is 0.9 independently of the p_T interval, as shown in Ref. [235]. In fact, the width of the D^+ peak is expected to be similar to the one of the D_s^+ meson since the D^+ mesons have a slightly lower mass compared to the D_s^+ mesons, leading to a lower Q -value of the decay. At the same time, the BR for the D^+ is slightly larger than the one of the D_s^+ mesons. Tables 5.5 and 5.6 report the values of the extracted raw yield, S/B , and significance for the prompt-enhanced and non-prompt-enhanced samples in the p_T intervals of the analyses.

The invariant-mass spectra are shown in Fig. 5.6 and Fig. 5.7, together with the fit functions and the resulting values for the D_s^+ Gaussian mean (μ) and width (σ), the raw yield, and the $S/B(3\sigma)$ for the prompt- and non-prompt-enhanced sample in pp collisions at $\sqrt{s} = 13$ TeV at midrapidity ($|y| < 0.5$). The data points are reported as black dots, while the total and the background fit functions are depicted as a blue solid and an orange dashed lines, respectively. The statistical significance, estimated within a 3σ interval around the D_s^+

peak, ranges between 5 and 30 (5 and 15) for the prompt- (non-prompt-) enhanced results leading to a statistical uncertainty on the raw yield of about 20% at low p_T and 3 – 6% at high p_T . Notably, using BDT-based selections instead of rectangular selections leads to an improvement in the signal extraction in terms of statistical significance, as reported in Ref. [234]. This is partly due to the fact that the BDT exploits the correlation among the training variables to better discriminate the signal from the background, but also to the fact that it is much easier to optimise the selection criteria for the two BDT probabilities, rather than having to deal with all the variables used in the rectangular selections. The systematic uncertainty on the raw yield is discussed in Sec. 5.3.1.

The values of the peak position (Gaussian mean) in data are found to agree with the world-average value of the D_s^+ meson mass [6] and with those from the simulations within the statistical uncertainties, though a systematic shift towards higher values is observed. The shift, which is common to all the D meson species measured in the Run 2 data taking, is attributed to the residual misalignment of the ITS detector which affects the reconstruction of the D-meson daughter momentum and, ultimately, the calculated invariant mass of the D_s^+ mesons. The width of the D_s^+ meson invariant-mass peak is found to be compatible between data and the MC simulation in all p_T intervals. The value of the peak width increases with increasing p_T of the D_s^+ meson due to the dependence of the momentum resolution of the daughter tracks.

5.2.2 Efficiency and acceptance corrections

As anticipated in Sec. 4.3.2, the $(\text{Acc} \times \epsilon)$ factor was computed from MC simulations. In order to avoid potential biases, different MC samples were employed for this purpose compared to those used for training and testing the BDT algorithms. The $(\text{Acc} \times \epsilon)$ correction factors for the prompt- and non-prompt-enhanced D_s^+ meson analyses as a function of p_T , inclusive of the rapidity-coverage correction, are presented in the left and right panels of Fig. 5.8, respectively.

The application of the BDT selection effectively enhances the contribution of either the prompt or non-prompt component, depending on the analysis goal. It is important to highlight that, in the case of analyses based on rectangular selections, the contribution of the non-prompt component in the prompt-enhanced sample was much higher than the one obtained with the BDT-based selection. Indeed, by employing more sophisticated analysis techniques is possible to reduce the contamination of the non-prompt component in the prompt-enhanced sample. Notably, the non-prompt D_s^+ meson acceptance-times-efficiency

in the prompt-enhanced sample is higher than that of prompt D_s^+ mesons in the non-prompt-enhanced sample. This difference arises from the larger displacement from the primary vertex of the non-prompt D_s^+ mesons decay vertex with respect to the prompt ones. In addition, in the case of the prompt-enhanced sample, looser selection criteria are sufficient to achieve reasonable values of statistical significance and efficiency due to the larger abundance of prompt mesons as compared to the non-prompt ones. Hence, there is no strong motivation to tighten the selection criteria, which would result in lower efficiency, to further suppress the non-prompt component. This is not the case for the non-prompt enhanced sample, where tight selection criteria are unavoidable to reduce the prompt contribution in order to measure the non-prompt component. In both cases, an increasing trend of efficiency is observed with increasing p_T . This effect is attributed to the presence of a larger Lorentz boost, which enhances the separation between the primary and the secondary vertices facilitating the selection of the signal candidate.

5.2.3 Prompt and non-prompt D_s^+ meson fraction

The raw yields extracted following the procedure described above need to be corrected for the residual contamination of prompt D_s^+ mesons in the non-prompt enhanced sample and vice versa. The fraction of prompt and non-prompt D_s^+ mesons $f_{(\text{non-})\text{prompt}}$ is estimated with either a theory- [181, 182, 117] or a data-driven method [239, 97, 123]. The theory-driven method exploits the FONLL predictions [86] of the D_s^+ meson production cross section and the $(\text{Acc} \times \epsilon)$ factors from the MC simulation. The data-driven method instead combines the raw yields and the $(\text{Acc} \times \epsilon)$ factors obtained for several BDT selection thresholds tuned to progressively vary the prompt and non-prompt D_s^+ meson contributions to the selected sample. This requirement restricts the p_T range where the data-driven approach can be applied in the case of the prompt-enhanced sample. This is due to the fact that the rejection of the non-prompt D_s^+ component in the prompt-enhanced sample is extremely challenging, especially in the low- p_T region, where the displacement of the non-prompt D_s^+ mesons decay vertex from the primary vertex is higher to the one of the prompt D_s^+ mesons and, consequently, the separation provided by the BDT is less effective. In addition, in the case of the prompt-enhanced sample the fraction of prompt D_s^+ mesons is close to unity, hence the correction based on the theory-driven method to subtract the non-prompt component is expected to be small. Hence, the theory-driven method is employed for the calculation of f_{prompt} in the case of the prompt-enhanced sample, while the data-driven method is employed to estimate $f_{\text{non-prompt}}$ in the case of the non-prompt-enhanced sample. The consistency of

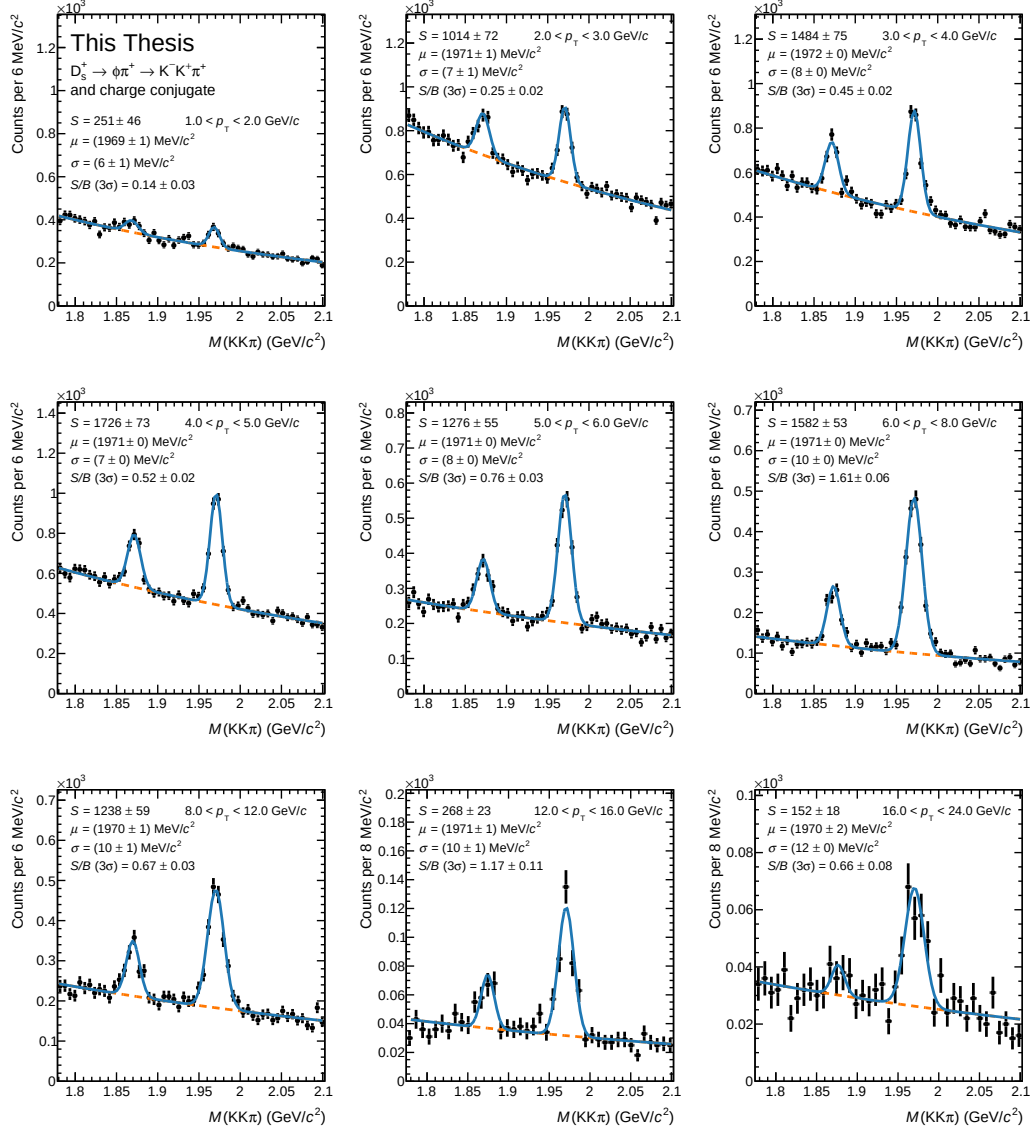


Fig. 5.6 Invariant-mass distributions of prompt D_s^+ -meson candidates in pp collisions at $\sqrt{s} = 13$ TeV in the nine p_T intervals of the analysis. The blue solid line represents the result of the fit to the data. The orange dashed line represents the background component of the fit.

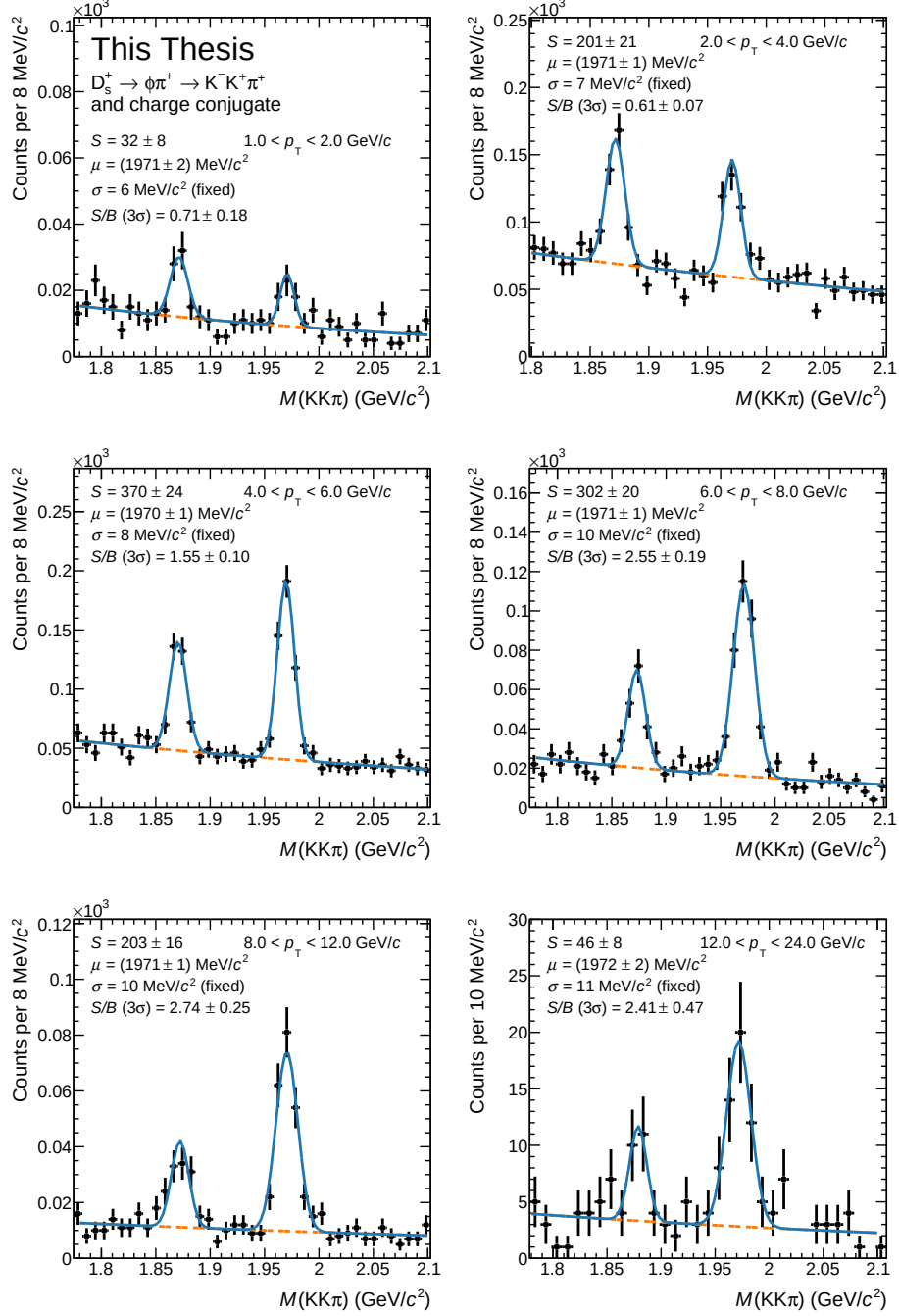


Fig. 5.7 Invariant-mass distributions of non-prompt D_s^+ -meson candidates in pp collisions at $\sqrt{s} = 13$ TeV in the six p_T intervals of the analysis. The blue solid line represents the result of the fit to the data. The orange dashed line represents the background component of the fit.

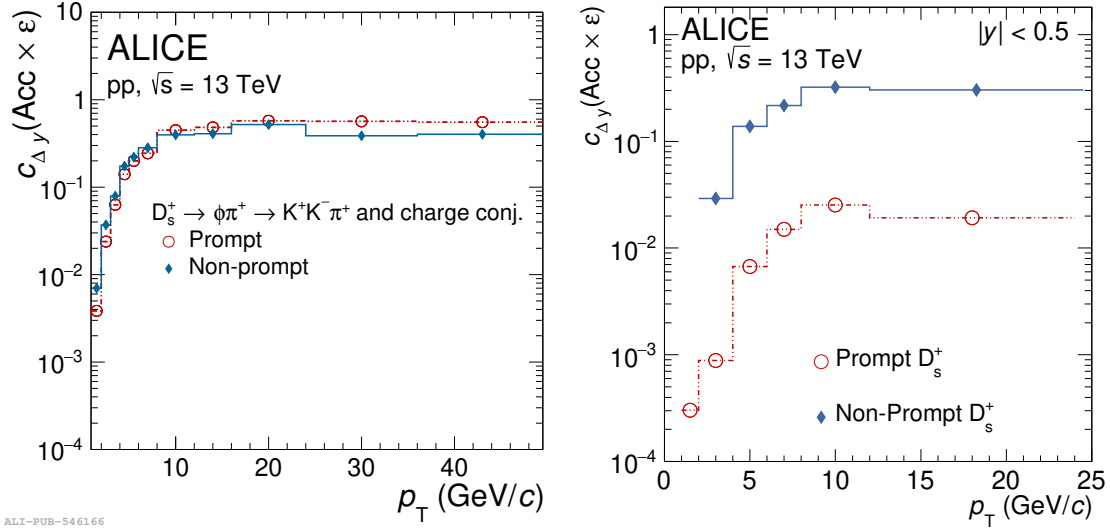


Fig. 5.8 Acceptance-times-efficiency correction factors for prompt and non-prompt D_s^+ mesons as functions of p_T , inclusive of the rapidity-coverage correction, for the prompt-(left) and non-prompt-enhanced (right) sample.

the two methods was tested in pp collisions at $\sqrt{s} = 5.02$ TeV for the prompt D_s^+ meson fraction as reported in Refs. [239, 240].

Theory-driven method

In the theory-driven method the prompt D_s^+ meson fraction is computed as:

$$f_{\text{prompt}} = 1 - \frac{d^2\sigma_{D_s^+}^{\text{non-prompt}}}{dp_T dy} \Big|_{\text{FONLL+PYTHIA 8}} \times \frac{(\text{Acc} \times \epsilon)_{\text{non-prompt}} c_{\Delta y} \Delta p_T \text{BR} \mathcal{L}_{\text{int}}}{\frac{1}{2} \times N_{\text{raw}}(D_s^+)}, \quad (5.5)$$

where the cross section for non-prompt D_s^+ mesons is obtained by combining the beauty-hadron cross section from FONLL with the decay kinematic of the beauty-hadron into the D meson state from PYTHIA 8. The $(\text{Acc} \times \epsilon)$ term is determined via MC simulation as described above and $N_{\text{raw}}(D_s^+)$ corresponds to the raw yield of D_s^+ mesons extracted from the fitting procedure. The prompt D_s^+ meson fraction is computed in all p_T intervals of the analyses and is found to vary between 0.81 and 0.90 depending on the p_T interval. Note that, the fraction of prompt D_s^+ mesons in the low- p_T intervals are slightly lower than the natural ones since the BDT selections end up including a high contribution of non-prompt D_s^+ mesons, as shown in Fig. 5.8.

Data-driven method

In the data-driven approach a set of n BDT selection criteria, characterised by a different mixture of prompt and non-prompt D_s^+ mesons, is firstly defined. In particular, the threshold value on the probability to be background, optimised in the first step of the analysis, remains unchanged, while the threshold value on the probability to be prompt or non-prompt D_s^+ meson is varied. For each selection i , the extracted raw yield (Y_i) and the corresponding $(\text{Acc} \times \epsilon)$ factors ($(\text{Acc} \times \epsilon)_i^{\text{prompt}}, (\text{Acc} \times \epsilon)_i^{\text{non-prompt}}$) are computed. A relation between the i -th raw yield, the $(\text{Acc} \times \epsilon)$ factors, and the *true* prompt and non-prompt D_s^+ meson yields ($N_{\text{prompt}}, N_{\text{non-prompt}}$)² can be defined as:

$$Y_i = (\text{Acc} \times \epsilon)_i^{\text{prompt}} \times N_{\text{prompt}} + (\text{Acc} \times \epsilon)_i^{\text{non-prompt}} \times N_{\text{non-prompt}} . \quad (5.6)$$

The relation in Eq. 5.6 can be extended to a system of equations by including all the selection criteria in the set:

$$\begin{cases} Y_1 = (\text{Acc} \times \epsilon)_1^{\text{prompt}} \times N_{\text{prompt}} + (\text{Acc} \times \epsilon)_1^{\text{non-prompt}} \times N_{\text{non-prompt}} \\ Y_2 = (\text{Acc} \times \epsilon)_2^{\text{prompt}} \times N_{\text{prompt}} + (\text{Acc} \times \epsilon)_2^{\text{non-prompt}} \times N_{\text{non-prompt}} \\ \vdots \\ Y_n = (\text{Acc} \times \epsilon)_n^{\text{prompt}} \times N_{\text{prompt}} + (\text{Acc} \times \epsilon)_n^{\text{non-prompt}} \times N_{\text{non-prompt}} \end{cases} . \quad (5.7)$$

For $n = 2$, the system can be solved exactly to obtain a unique solution for the "true" prompt and non-prompt D_s^+ meson yields in the sample. In the case of $n > 2$ though, the system is over-constrained and the solution in terms of N_{prompt} and $N_{\text{non-prompt}}$ is obtained via a minimisation method. For simplicity, the system of equations is rewritten in the matrix form as:

$$\begin{pmatrix} (\text{Acc} \times \epsilon)_1^{\text{prompt}} & (\text{Acc} \times \epsilon)_1^{\text{non-prompt}} \\ (\text{Acc} \times \epsilon)_2^{\text{prompt}} & (\text{Acc} \times \epsilon)_2^{\text{non-prompt}} \\ \vdots & \vdots \\ (\text{Acc} \times \epsilon)_n^{\text{prompt}} & (\text{Acc} \times \epsilon)_n^{\text{non-prompt}} \end{pmatrix} \begin{pmatrix} N_{\text{prompt}} \\ N_{\text{non-prompt}} \end{pmatrix} - \begin{pmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_n \end{pmatrix} = \begin{pmatrix} \delta_1 \\ \delta_2 \\ \vdots \\ \delta_n \end{pmatrix} , \quad (5.8)$$

where δ_i are the residuals arising from the inexact solution of the i -th equation in the system due to the uncertainty on the raw yield and $(\text{Acc} \times \epsilon)$ factors. In the specific case of $n = 2$, the vector $\boldsymbol{\delta}$ coincides with the null vector. The best estimation of the true prompt and

²It is important to notice that, unlike the extracted raw yields and the $(\text{Acc} \times \epsilon)$ factor, the true prompt and non-prompt D_s^+ meson yields are not affected by the choice of BDT selection.

non-prompt yields is obtained by minimising the χ^2 , defined as:

$$\chi^2 = \boldsymbol{\delta}^T \mathbf{C}^{-1} \boldsymbol{\delta} . \quad (5.9)$$

The $\mathbf{C}$ term is the covariance matrix of the raw yields and $(\text{Acc} \times \varepsilon)$ factors, computed as:

$$\mathbf{C} = \begin{pmatrix} \sigma_1^2 & \sigma_{1,2} & \dots & & \\ \sigma_{2,1} & \sigma_2^2 & & & \\ \vdots & & \ddots & & \vdots \\ & & & \sigma_{n-1}^2 & \sigma_{n-1,n} \\ \dots & & & \sigma_{n,n-1} & \sigma_n^2 \end{pmatrix} . \quad (5.10)$$

The diagonal elements of the covariance matrix are computed as the sum of the variance of the raw yield and the variance of the prompt and non-prompt $(\text{Acc} \times \varepsilon)$ factors weighted by the square of the N_{prompt} and $N_{\text{non-prompt}}$ yields, respectively. It is important to notice that, at this stage, the final value of the true prompt and non-prompt yields are unknown. Hence, they are initially set to zero, and are updated at each iteration of the minimisation process. The off-diagonal elements are computed as $\sigma_{i,j} = \rho_{i,j} \sigma_i \sigma_j$, where $\rho_{i,j}$ is the correlation coefficient between the i -th and j -th selection estimated as $\rho_{i,j} = \sigma_i / \sigma_j$ assuming the i -th selection to fully contain the j -th one. This assumption is justified by the fact that, to vary the mixture of prompt and non-prompt, the thresholds on the probability score follow a monotonic increase from low to high values with increasing i . Thus, the i -th selection is completely included in the $i - 1$ -th, by definition. The minimisation procedure goes on until a predefined convergence criterion is met. Few iterations are typically sufficient to reach the convergence.

Once the minimisation procedure is completed, it is then possible to compute the corrected yields as³:

³The relation described in Eq. 5.11 is obtained from the following derivation:

$$\begin{aligned} \frac{\partial \chi^2}{\partial N} &= \boldsymbol{\varepsilon}^T \mathbf{C}^{-1} (\boldsymbol{\varepsilon} \mathbf{N} - \mathbf{Y}) + (\boldsymbol{\varepsilon} \mathbf{N} - \mathbf{Y})^T \mathbf{C}^{-1} \boldsymbol{\varepsilon}^T \mathbb{I} \\ &= \boldsymbol{\varepsilon}^T \mathbf{C}^{-1} \boldsymbol{\varepsilon} \mathbf{N} - \boldsymbol{\varepsilon}^T \mathbf{C}^{-1} \mathbf{Y} + \mathbf{N}^T \boldsymbol{\varepsilon}^T \mathbf{C}^{-1} \boldsymbol{\varepsilon} \mathbb{I} - \mathbf{Y}^T \mathbf{C}^{-1} \boldsymbol{\varepsilon} \mathbb{I} \\ &= \boldsymbol{\varepsilon}^T \mathbf{C}^{-1} \boldsymbol{\varepsilon} \mathbf{N} - \boldsymbol{\varepsilon}^T \mathbf{C}^{-1} \mathbf{Y} + \mathbb{I}^T \boldsymbol{\varepsilon}^T (\mathbf{C}^{-1})^T \boldsymbol{\varepsilon} \mathbf{N} - \mathbb{I}^T \boldsymbol{\varepsilon}^T (\mathbf{C}^{-1})^T \boldsymbol{\varepsilon} \mathbf{N} \\ &= 2(\boldsymbol{\varepsilon}^T \mathbf{C}^{-1} \boldsymbol{\varepsilon} \mathbf{N} - \boldsymbol{\varepsilon}^T \mathbf{C}^{-1} \mathbf{Y}) \end{aligned}$$

In the third line of the derivation, the symmetry of the $\mathbf{C}$ matrix is exploited. Imposing the condition $\partial \chi^2 / \partial N = 0$ leads to the relation in Eq. 5.11:

$$\frac{\partial \chi^2}{\partial N} = 0 \Rightarrow \mathbf{N} = \text{Cov}(\mathbf{N}) \boldsymbol{\varepsilon}^T \mathbf{C}^{-1} \mathbf{Y} ,$$

$$\mathbf{N} = \mathbf{Cov}(\mathbf{N})\boldsymbol{\varepsilon}^T \mathbf{C}^{-1} \mathbf{Y}, \quad (5.11)$$

where $\mathbf{N}$ is the vector of the corrected yields and $\mathbf{Cov}(\mathbf{N})$ is the covariance matrix of the corrected yields containing the statistical uncertainties of the corrected yields and the correlations among them. Given an arbitrary BDT selection k , not necessarily included in the set of selections used to compute the corrected yields, the non-prompt D_s^+ meson fractions are computed in each p_T interval as:

$$f_{\text{non-prompt}}^k = \frac{(\text{Acc} \times \varepsilon)_{\text{non-prompt}}^k \times N_{\text{non-prompt}}}{(\text{Acc} \times \varepsilon)_{\text{prompt}}^k \times N_{\text{prompt}} + (\text{Acc} \times \varepsilon)_{\text{non-prompt}}^k \times N_{\text{non-prompt}}}. \quad (5.12)$$

Figure 5.9 shows the evolution of the prompt and non-prompt $(\text{Acc} \times \varepsilon)$ terms as a function of the BDT-selection criteria (top-left), the correlation matrix of the chosen BDT-selections (top-right), the contribution of the prompt and non-prompt component to the extracted raw-yields (bottom-left), and the prompt and non-prompt D_s^+ meson fraction as a function of p_T (bottom-right) in the non-prompt-enhanced sample. The results reported in the figure have been obtained for pp collisions at $\sqrt{s} = 13$ TeV in the $2 < p_T < 4$ GeV/ c interval. As expected, the non-prompt $(\text{Acc} \times \varepsilon)$ exceeds the prompt one by a factor 10 – 20 depending on the BDT selection. The contribution of the non-prompt component progressively increases moving from loose BDT-selection criteria to tighter ones. Considering the results obtained from the minimisation, the $f_{\text{non-prompt}}$ fraction corresponding to the set of BDT selections adopted for the analysis is computed exploiting the relation in Eq. 5.12. The non-prompt fraction obtained in the non-prompt enhanced sample with the reference BDT-selection criteria is found to be around 0.7 in all the p_T intervals of the analysis.

5.3 Systematic uncertainties

The following sources of systematic uncertainty were considered for the prompt and non-prompt D_s^+ -meson production cross sections: (i) raw-yield extraction, (ii) track-reconstruction efficiency, (iii) PID efficiency, (iv) non-prompt fraction estimation, (v) D-meson selection efficiency, and (vi) D-meson simulated p_T shape.

The summary of the systematic uncertainty for the prompt and non-prompt D_s^+ measurements are reported in Table 5.7 and 5.8, respectively.

where $\mathbf{Cov}(\mathbf{N}) = (\boldsymbol{\varepsilon}^T \mathbf{C}^{-1} \boldsymbol{\varepsilon})^{-1}$ is the corrected yield covariance matrix. To invert the matrix $\boldsymbol{\varepsilon}^T \mathbf{C}^{-1} \boldsymbol{\varepsilon}$, the Cholesky decomposition is employed [275].

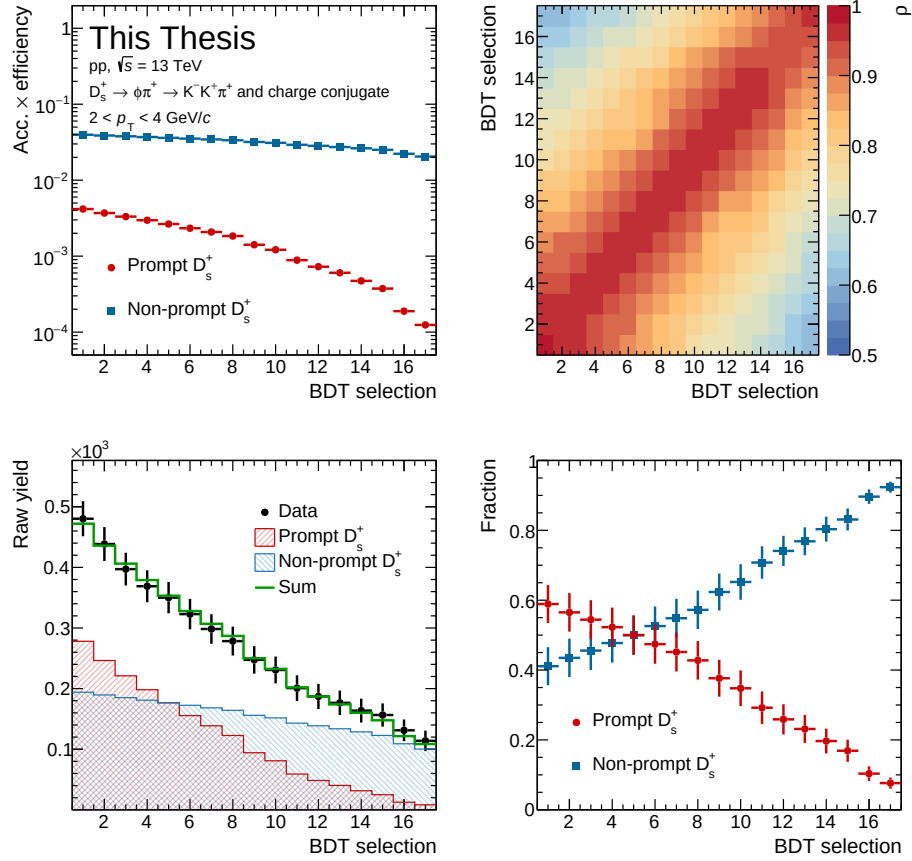


Fig. 5.9 Top-left: prompt and non-prompt ($\text{Acc} \times \epsilon$) factors as a function of the BDT selection criteria. Top-right: correlation matrix of the BDT selections expressed in terms of the correlation coefficient ρ . Bottom-left: contribution of the prompt and non-prompt component to the extracted raw-yields as a function of the BDT selection criteria. Bottom-right: prompt and non-prompt D_s^+ meson fraction as a function of the BDT selection criteria. The results are obtained in the $2 < p_T < 4$ GeV/c interval in the non-prompt-enhanced sample in pp collisions at $\sqrt{s} = 13$ TeV.

Table 5.7 Relative systematic uncertainties of the measured cross section of prompt D_s^+ mesons at midrapidity in pp collisions at $\sqrt{s} = 13$ TeV. Uncertainties found to be negligible are indicated as such (negl.).

	p_T interval (GeV/c)				
	1 – 2	2 – 3	3 – 4	4 – 5	5 – 6
Signal extraction	8%	4%	2%	2%	2%
Tracking efficiency	4%	4%	5%	6%	7%
PID efficiency	negl.	negl.	negl.	negl.	negl.
Non-prompt fraction	6%	6%	5%	4%	
Selection efficiency	10%	3%	3%	3%	3%
MC p_T shape	1%	1%	1%	negl.	negl.
Branching ratio	2.7%				
Luminosity	1.6%				
	p_T interval (GeV/c)				
	6 – 8	8 – 12	12 – 16	16 – 24	24 – 36
Signal extraction	2%	2%	4%	5%	5%
Tracking efficiency	7%	7%	8%	8%	8%
PID efficiency	negl.	negl.	negl.	negl.	negl.
Non-prompt fraction	4%	3%	3%	3%	2%
Selection efficiency	3%	3%	3%	3%	3%
MC p_T shape	negl.	negl.	negl.	negl.	negl.
Branching ratio	2.7%				
Luminosity	1.6%				

5.3.1 Systematic on the raw-yield extraction

The systematic uncertainty on the raw-yield extraction was assessed by adopting several fit configurations varying the background fit function (linear, exponential, and parabolic), the upper and lower fit limits, and the binning of the invariant-mass distribution. In case of non-prompt D_s^+ meson measurements, the Gaussian width of the D_s^+ peak was varied within the uncertainty range of the D_s^+ meson width in the prompt-enhanced sample. In case of prompt D_s^+ meson measurements, the Gaussian width of the D_s^+ peak was left free to vary as in the reference fit configuration. To exclude non-reliable fit configurations, a quality selection requiring a $\chi^2/\text{ndf} < 2$ was applied. The systematic uncertainty was computed considering the mean and the RMS of the distribution of the raw-yield values obtained with the different fit configurations. In particular, the uncertainty was defined as the sum in quadrature of the shift of the mean with respect to the raw yield value obtained with the

Table 5.8 Relative systematic uncertainties of the measured cross section of non-prompt D_s^+ mesons at midrapidity in pp collisions at $\sqrt{s} = 13$ TeV. Uncertainties found to be negligible are indicated as such (negl.).

	p_T interval (GeV/c)				
	2 – 4	4 – 6	6 – 8	8 – 12	12 – 24
Signal extraction	3%	3%	3%	3%	5%
Tracking efficiency	5%	6%	7%	8%	8%
PID efficiency	negl.	negl.	negl.	negl.	negl.
Non-prompt fraction	4%	2%	2%	2%	2%
Selection efficiency	3%	3%	3%	3%	3%
MC p_T shape	3%	3%	2%	2%	2%
Branching ratio	2.7%				
Luminosity	1.6%				

reference fit configuration and the RMS of the distribution. Figure 5.10 shows the distribution of the raw-yields obtained with the various fit configurations for the prompt-enhanced (left panel) and non-prompt-enhanced (right panel) samples in pp collisions at $\sqrt{s} = 13$ TeV in the $2 < p_T < 3$ and $2 < p_T < 4$ GeV/c intervals, respectively.

The sensitivity to the line shape of the D_s^+ -meson peak was tested by comparing the raw-yield values obtained from the fits with those obtained by counting the candidates in the invariant mass region of the signal after subtracting the background estimated from a fit to the sidebands. This method is referred to as *Bin Counting* (BC) method. The BC method was performed considering $\pm 3\sigma$ and $\pm 5\sigma$ intervals around the D_s^+ meson peak, respectively. The results of the BC method were solely employed as a cross check and were not included in the systematic uncertainty estimation.

5.3.2 MC-Data discrepancy

The sources of systematic (ii)–(vi) originate from possible differences between the data and the simulation due to an imperfect modelling of the particle interactions, or to an inaccurate description of the detector response and alignment, and of the underlying physics processes in the MC simulation.

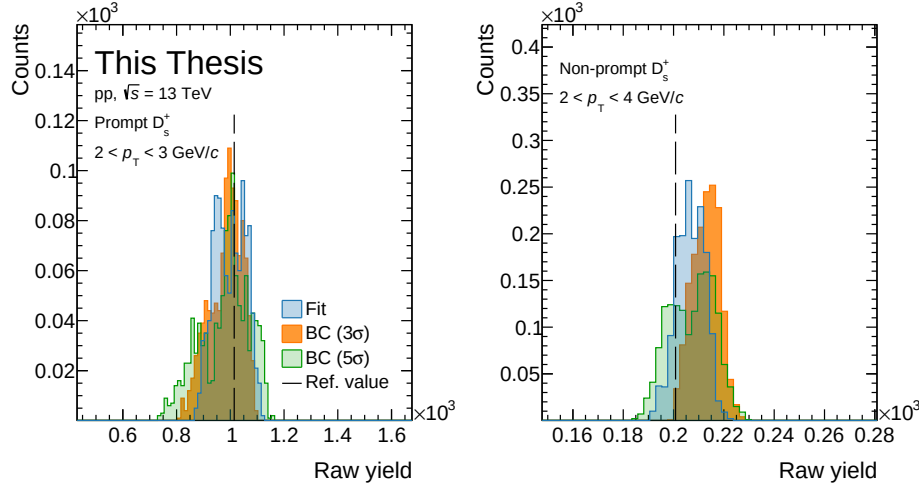


Fig. 5.10 Distribution of the raw-yield values obtained with the different fit configurations for the prompt- (left) and non-prompt-enhanced (right) samples in pp collisions at $\sqrt{s} = 13$ TeV in the $2 < p_T < 3$ and $2 < p_T < 4$ GeV/c intervals, respectively.

Systematic on the track-reconstruction efficiency

The systematic uncertainty associated with the track reconstruction efficiency takes into account potential differences between data and MC simulation in both the TPC-ITS track prolongation efficiency and the selection efficiency governed by track-quality criteria in the TPC (refer to Chapter 4). The systematic uncertainties per single track were estimated by modifying the track-quality selection criteria: the number of crossed rows in the TPC, which depends on the reconstructed p_T of the track as $120 - 5/(p_T[\text{GeV}/c])$, the requirement of a number of TPC clusters greater than 0.65 times the TPC crossed-row number, and the ratio of crossed rows to findable clusters in the TPC, set to 0.8 as the default value. Varying these selections has an impact on the measured cross-section for D_s^+ mesons, regardless of their p_T , of approximately 1.5%.

The TPC-to-ITS prolongation efficiency was calculated as the ratio of tracks successfully fitted with the Kalman filter [180] in both the TPC and ITS, with at least one hit in the SPD layers, to the number of reconstructed tracks successfully fitted in the TPC. This was computed separately for data and MC simulations. The associated uncertainty was determined by comparing the prolongation efficiency from MC simulations to that observed in data. The ITS-TPC matching efficiency is expected to be higher for primary particles [276], namely those particles produced in the beam collision or coming from decays of charm and beauty hadrons with a proper decay length of order of tens-hundred of microns, than for secondary particles produced in the interactions with the detector material or in decays of strange

hadrons, which typically exhibit a mean proper decay length of several centimetres. However, since this separation is not possible in data, an inclusive ITS-TPC matching efficiency is used in the comparison with the MC. The relative abundances of primary and secondary particles are not accurately reproduced in the MC simulations based on the PYTHIA 8 event generator and the GEANT 3 [174] transport package. Consequently, the ITS-TPC matching efficiency in the MC simulation has to be reweighted so that the relative abundance of primary and secondary particles in the MC simulation matches the measured ones. The relative proportions of primary and secondary particles in the analysed data sample, for each p_T interval, are estimated through a template fit of the single-track impact parameter d_0^{xy} distribution, as discussed in detail in Ref. [184].

The individual per-track systematic uncertainties on the TPC-ITS track prolongation efficiency were propagated to the D_s^+ mesons through the kinematics of the decay by considering the correlation of the p_T of the D_s^+ -meson particle and its decay products. The systematic uncertainties on the single track were considered as fully correlated and summed linearly. Finally, the systematic uncertainty related to the track reconstruction efficiency is computed as the sum in quadrature of the uncertainty associated to the track quality selections and the one associated with the ITS-TPC matching efficiency. The left panel of Fig. 5.11 illustrates the relative systematic uncertainty on the track reconstruction efficiency as a function of p_T in pp collisions at $\sqrt{s} = 13$ TeV. The assigned systematic uncertainty ranges between 4% and 8%, depending on the p_T interval, in both the prompt-enhanced and non-prompt-enhanced samples, as expected.

PID efficiency

The systematic uncertainty of the PID selection efficiency for the D_s^+ -meson candidates was estimated by comparing the PID selection efficiencies in the data and in the MC simulation for pions and kaons. A pure sample of pions was isolated considering decays of strange hadrons (K_S^0 , Λ), while a pure sample of kaons was isolated in the TPC (TOF) by applying strict selections on the TOF (TPC) information. The resulting per-track uncertainty was then propagated to the D_s^+ -meson candidates using their decay kinematics. Considering the conservative PID approach described in Chapter 4, no significant difference was observed between data and MC simulation. The systematic uncertainty on the PID efficiency was therefore considered negligible for both the prompt- and non-prompt-enhanced sample.

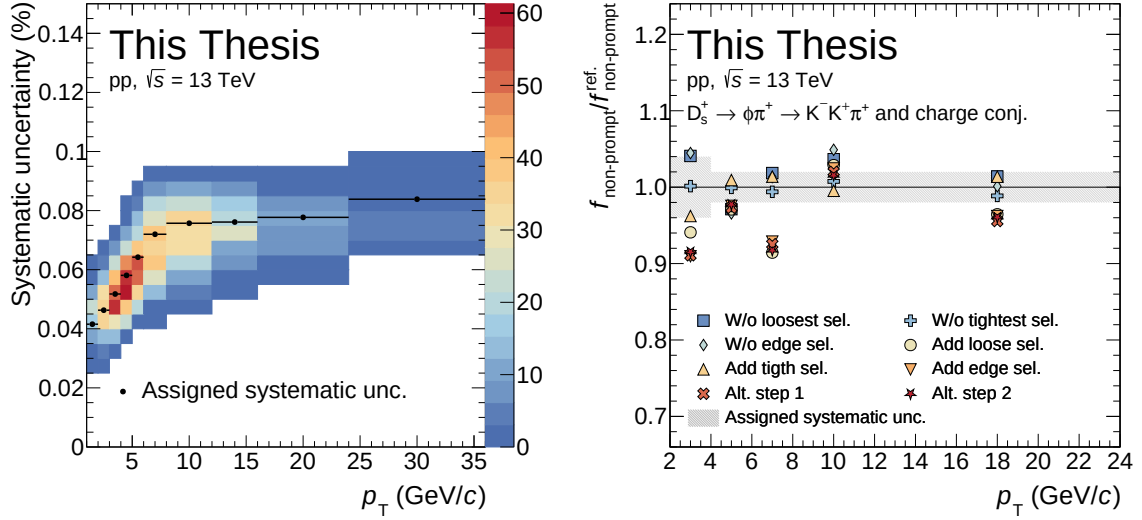


Fig. 5.11 Left: Relative systematic uncertainty on the track reconstruction efficiency as a function of p_T in pp collisions at $\sqrt{s} = 13$ TeV for the prompt-enhanced sample. Right: Relative systematic uncertainty on the non-prompt D_s^+ meson fraction as a function of p_T in pp collisions at $\sqrt{s} = 13$ TeV for the non-prompt-enhanced sample.

Prompt and non-prompt fraction estimation

The procedure followed to determine the systematic uncertainty of the prompt and non-prompt fractions depends on the method employed to estimate the fractions. In the case of the theory-driven method, the systematic uncertainty was evaluated by considering the uncertainty on the FONLL predictions. The FONLL uncertainties are computed by varying the parameters employed in the calculations, namely the mass of the charm and beauty quarks (varied between 1.3 and 1.7 GeV/ c^2 for charm and between 4.5 and 5 GeV/ c^2 for beauty), the normalisation and factorisation scales ($0.5 < \mu_{F,R}/\mu_0 < 2$, where $\mu_0 = \sqrt{p_T^2 + m_b^2}$ is the default scale adopted in the FONLL calculations). In addition, the uncertainty on the chosen set of PDF (CTEQ6.6 [81]) [87] has been also considered.

In the case of the data-driven method, the systematic uncertainty was estimated by varying the set of BDT selections included in the data-driven method described above. The set of BDT selections was varied either to include more severe and looser selections on the probability to be a non-prompt D_s^+ meson, or to exclude the tightest and loosest selections on this probability with respect to the reference set of selections, as well as by adopting a smaller or larger step size among the BDT selections. The non-prompt D-meson fraction was then computed separately for each set and the systematic uncertainty was ultimately assigned considering the RMS of the resulting distribution and the shift of the mean with respect to

the reference set. The right panel of Fig. 5.11 shows the ratio of the non-prompt D-meson fraction obtained with the different sets of the BDT selections to the one obtained for the reference set in the case of the non-prompt-enhanced sample in pp collisions at $\sqrt{s} = 13$ TeV together with the assigned systematic uncertainty.

D_s^+ meson selection efficiency

The systematic uncertainty on the D_s^+ -meson selection efficiency was estimated by varying the BDT selections employed to select candidates. The BDT selection on the probability to be background and the probability to be prompt (or non-prompt, depending on the goal of the measurement) were varied, independently and then simultaneously, for a total of a hundred variations. The variations were performed up to relative variation of the prompt (or non-prompt) efficiency-times-acceptance factor of $\pm 30\%$, with respect to the reference value. To exclude problematic cases or too extreme variations, only the selections providing a statistical significance of at least 3 and an efficiency-times-acceptance factor between 0.5 and 2.5 times the default values were taken into account in the evaluation of the systematic uncertainty. The uncertainty was assigned considering the RMS in the distribution of the corrected yields derived from the BDT selections and the shift of its mean with respect to the reference value.

The top-left panel of Figure 5.12 shows the distribution of the statistical significance obtained with the different configurations of the BDT selections for the prompt-enhanced sample in pp collisions at $\sqrt{s} = 13$ TeV in the $2 < p_T < 3$ GeV/ c interval. The top-right panel displays the ratio of the efficiency-times-acceptance factor obtained from the different BDT selections relative to the reference one. The bottom-left panel shows the ratio of the (non-)prompt fraction obtained from the different BDT selections to the reference one. The bottom-right panel displays the distribution of the ratio between the corrected yield obtained from the different BDT selections and the reference one. The assigned systematic uncertainty is reported as a coloured band.

Generated p_T shape

The calculation of the $(\text{Acc} \times \varepsilon)$ factor presented in Chapter 4 can be influenced by differences between the p_T distributions of the generated D_s^+ mesons in the simulation and those in the data. This difference introduces a systematic bias in the estimate of the efficiencies in the p_T intervals used in the analysis. To assess the magnitude of this systematic uncertainty the p_T distributions of the prompt D_s^+ mesons in the PYTHIA 8 simulations were weighted

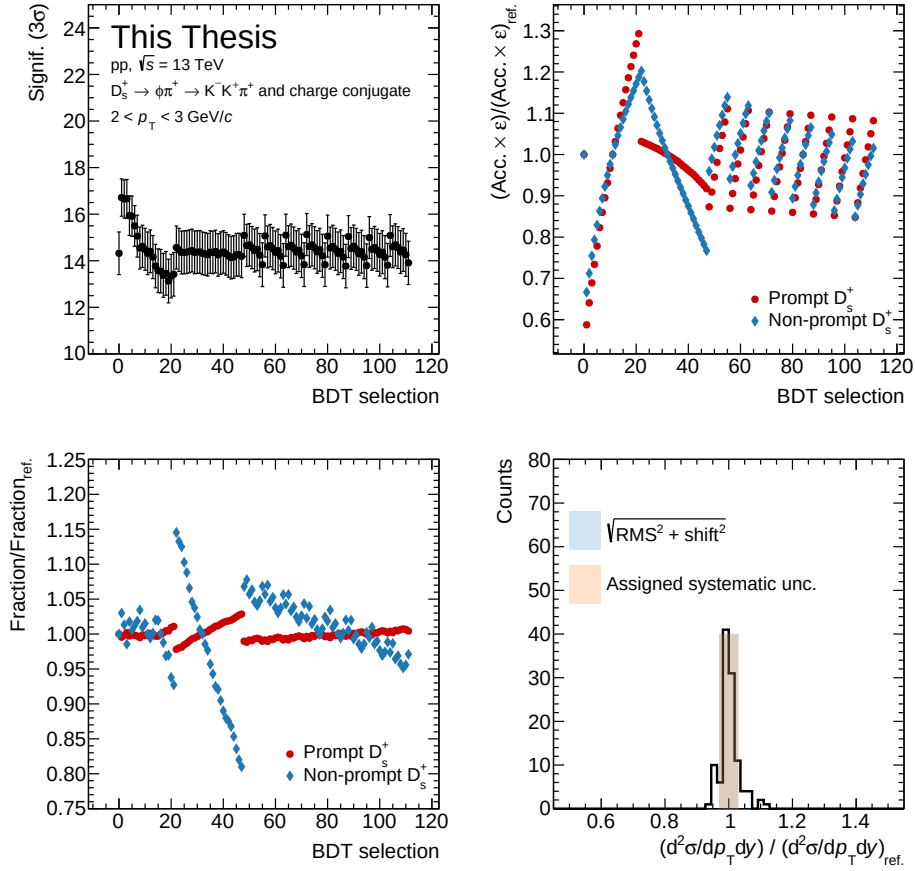


Fig. 5.12 Distribution obtained from the variation of the BDT selections with respect to the reference one for the statistical significance (top-left panel), the ratio of the efficiency-times-acceptance factor (top-right panel), the ratio of the (non-)prompt fraction (bottom-left panel), and the corrected yield (bottom-right panel). The distributions refer to the prompt-enhanced sample in pp collisions at $\sqrt{s} = 13$ TeV in the $2 < p_T < 3$ GeV/c interval.

to match the spectra derived from FONLL calculations. In the case of non-prompt D_s^+ mesons, the p_T distributions of the beauty-hadron parent were weighted. The $(\text{Acc} \times \epsilon)$ factor and the raw non-prompt D_s^+ meson fraction were recomputed after the application of the weights and compared to the reference case, obtained from PYTHIA 8 simulations. The difference between the two results was assigned as a systematic uncertainty in the case of the prompt-enhanced sample. Instead, in the case of the non-prompt enhanced sample the prompt and non-prompt $(\text{Acc} \times \epsilon)$ factors enter both in the computation of the non-prompt D_s^+ fraction with the data-driven method and, consequently, in the estimate of the production cross section. Therefore, in this case the systematic uncertainty was assigned considering the final effect on the non-prompt D_s^+ meson production cross section.

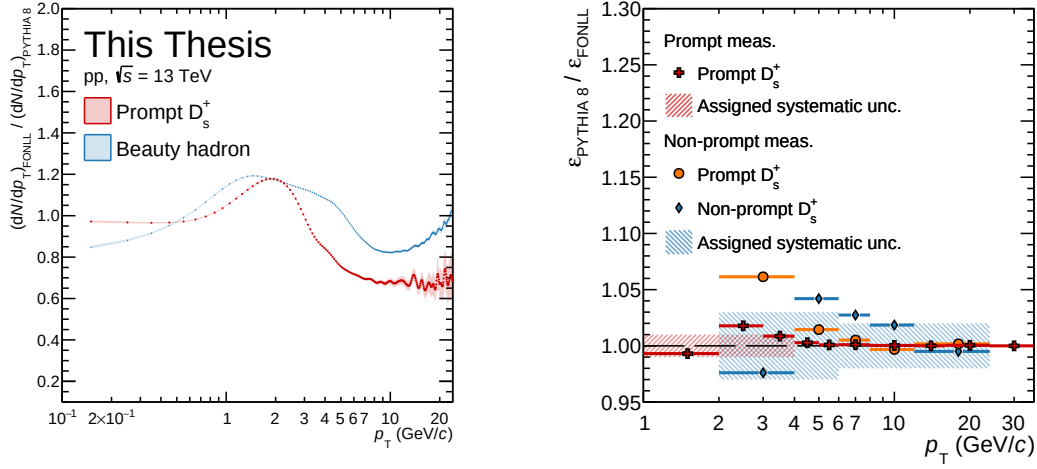


Fig. 5.13 Left: Ratio of the p_T spectra for prompt D_s^+ and the beauty hadrons between FONLL and PYTHIA 8. Right: Ratio between the $(\text{Acc} \times \epsilon)$ correction factor obtained with the FONLL p_T weights over the reference for the prompt- and non-prompt-enhanced sample in pp collisions at $\sqrt{s} = 13$ TeV together with the assigned systematic.

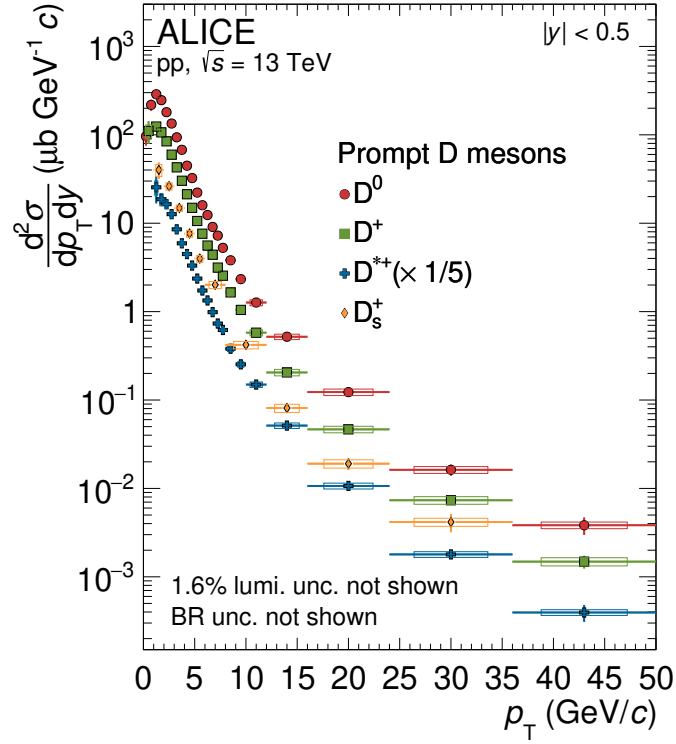
The left panel of Fig. 5.13 shows the ratio of the p_T spectra for prompt D_s^+ and the beauty hadrons between FONLL and PYTHIA 8. The right panel displays the ratio between the $(\text{Acc} \times \epsilon)$ correction factor obtained with the FONLL p_T weights and the reference one obtained with PYTHIA 8 for the prompt- and non-prompt-enhanced sample in pp collisions at $\sqrt{s} = 13$ TeV together with the assigned systematic uncertainty.

In addition, the efficiency of non-prompt D_s^+ mesons originating from the decay of different beauty-hadron species depends on the $c\tau$ of the parent beauty hadron and its decay kinematics. For this reason, an additional source of systematic was considered for the measurement of the non-prompt D_s^+ -meson production in Pb–Pb collisions, as discussed in detail in Chapter 6.

5.4 Prompt and non-prompt D_s^+ meson results

5.4.1 Prompt D-meson p_T -differential cross sections

The p_T -differential production cross sections of prompt D^0 , D^+ , D^{*+} , and D_s^+ mesons at midrapidity ($|y| < 0.5$) in pp collisions at a centre-of-mass energy of $\sqrt{s} = 13$ TeV [97] are presented in Fig. 5.14. Hereafter, the statistical uncertainties are represented by vertical lines while systematic uncertainties are shown as empty boxes. The systematic uncertainty does



ALI-PUB-546170

Fig. 5.14 The p_T -differential production cross sections of prompt D^0 , D^+ , D^{*+} , and D_s^+ mesons at midrapidity ($|y| < 0.5$) in pp collisions at a centre-of-mass energy of $\sqrt{s} = 13$ TeV. The vertical bars and boxes report the statistical and systematic uncertainties, respectively. The systematic uncertainties reported in the plots do not include the contributions of the luminosity and the branching ratio, which are reported separately. The measured prompt D^{*+} -meson production cross section was scaled by a factor of 1/5 for better visibility reasons. Figure from Ref. [97].

not include the one relative to the BR of the analysed decay and the one on the estimation of the integrated luminosity. The differential production cross sections of D_s^+ expand the momentum coverage of the measurements and provide a finer p_T binning thanks to the higher integrated luminosity collected by ALICE in pp collisions at $\sqrt{s} = 13$ TeV compared to lower energies.

Figure 5.15 compares the measured cross sections with predictions from various pQCD calculations. The calculations are: FONLL [86, 87] (except for D_s^+ mesons), GM-VFNS [88, 89], and a calculation based on the k_T -factorisation approach [92, 246]. The theoretical uncertainties for these predictions are shown as shaded boxes. As aforementioned, in the case of FONLL the theoretical uncertainty accounts for variations of the μ_F and μ_R scales, the charm-quark mass, and includes the uncertainty on the employed PDFs, which in this case are based on CTEQ6.6 [81]. The fragmentation fractions employed in the FONLL calculations

are derived from Ref. [277]. The GM-VFNS calculations follow a similar configuration to Ref. [239], using CTEQ6.6 PDFs [106] as default. The factorisation and renormalisation scales correspond to those used in Ref. [181]. The calculations are performed in the same p_T intervals of the measurements to facilitate a straightforward comparison with data. The only exception regards the p_T region near to 0 GeV/c, where, at this collision energy, numerical instabilities arise in the GM-VFNS calculations and the predictions are not reliable [92].

The k_T -factorisation framework employs a variable-flavour-number scheme and improves the factorisation scheme used in previous predictions [239, 181] for the estimation of the unintegrated PDFs and, ultimately, the production cross section. The variable-flavour-number scheme was shown to be highly efficient by resumming to all orders some large logarithms $\ln(p_T^2/m_Q^2)$ thanks to the heavy-quark distribution function, as discussed in detail in Ref. [92]. It is also demonstrated that flavour-excitation processes play a crucial role in the description of the charm-quark production. The theoretical uncertainties were estimated by varying the factorisation scale.

Similarly to what observed at lower collision energies and over different rapidity intervals, pQCD calculations implementing the factorisation approach succeed in describing the measured D-meson production cross sections within the uncertainties [181, 239, 247, 248, 245]. The measurements align with the upper edge of the FONLL calculations up to $p_T = 36$ GeV/c, and are compatible with the upper edge of the calculations within about 2σ in $36 < p_T < 50$ GeV/c. In the GM-VFNS framework, the central values of the calculations slightly underestimate (overestimate) the data in the low- (high-) p_T region. The predictions based on the k_T -factorisation exhibit a good agreement with the data for $p_T > 10$ GeV/c, while at lower p_T the central values of the calculations underestimate the data.

5.4.2 Energy dependence of prompt D meson cross section ratios

The prompt D^+/D^0 , D^{*+}/D^0 , D_s^+/D^0 , D_s^+/D^+ , and $D_s^+/(D^0+D^+)$ yield ratios as a function of p_T at midrapidity ($|y| < 0.5$) in pp collisions at $\sqrt{s} = 13$ TeV are shown in Fig. 5.16. The systematic uncertainties were propagated in the ratio as fully correlated among the different D-meson species, except for the systematic related to the raw-yield extraction and the selection efficiency which were considered as uncorrelated. The results obtained at midrapidity at $\sqrt{s} = 13$ TeV [97] are compared to the ones obtained in the same rapidity range at $\sqrt{s} = 7$ TeV [182] and $\sqrt{s} = 5.02$ TeV [239] by the ALICE Collaboration. The results are fully compatible among the three energies considering the experimental uncertainties. This observation suggests that there are common fragmentation functions for charm quarks,

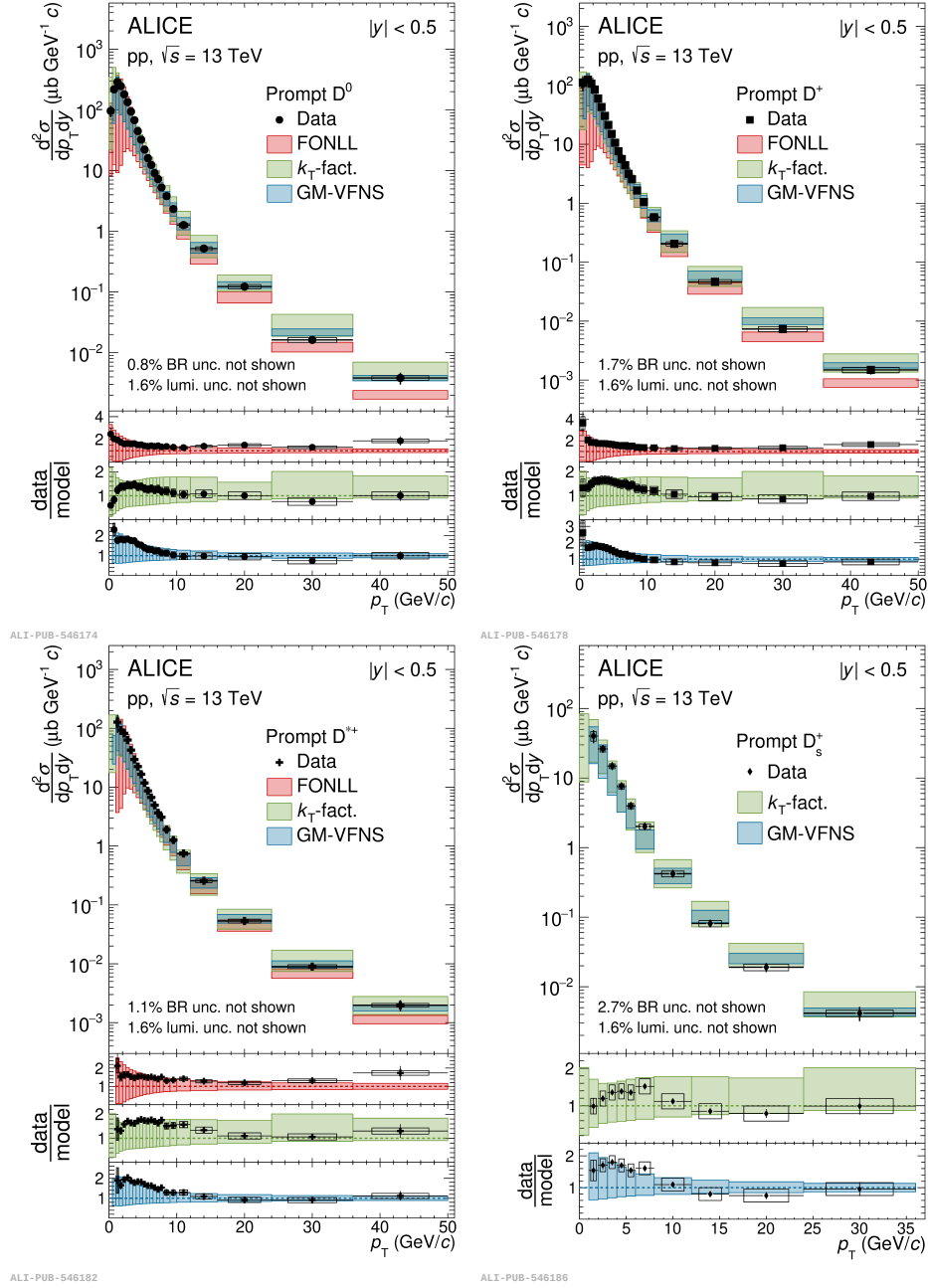


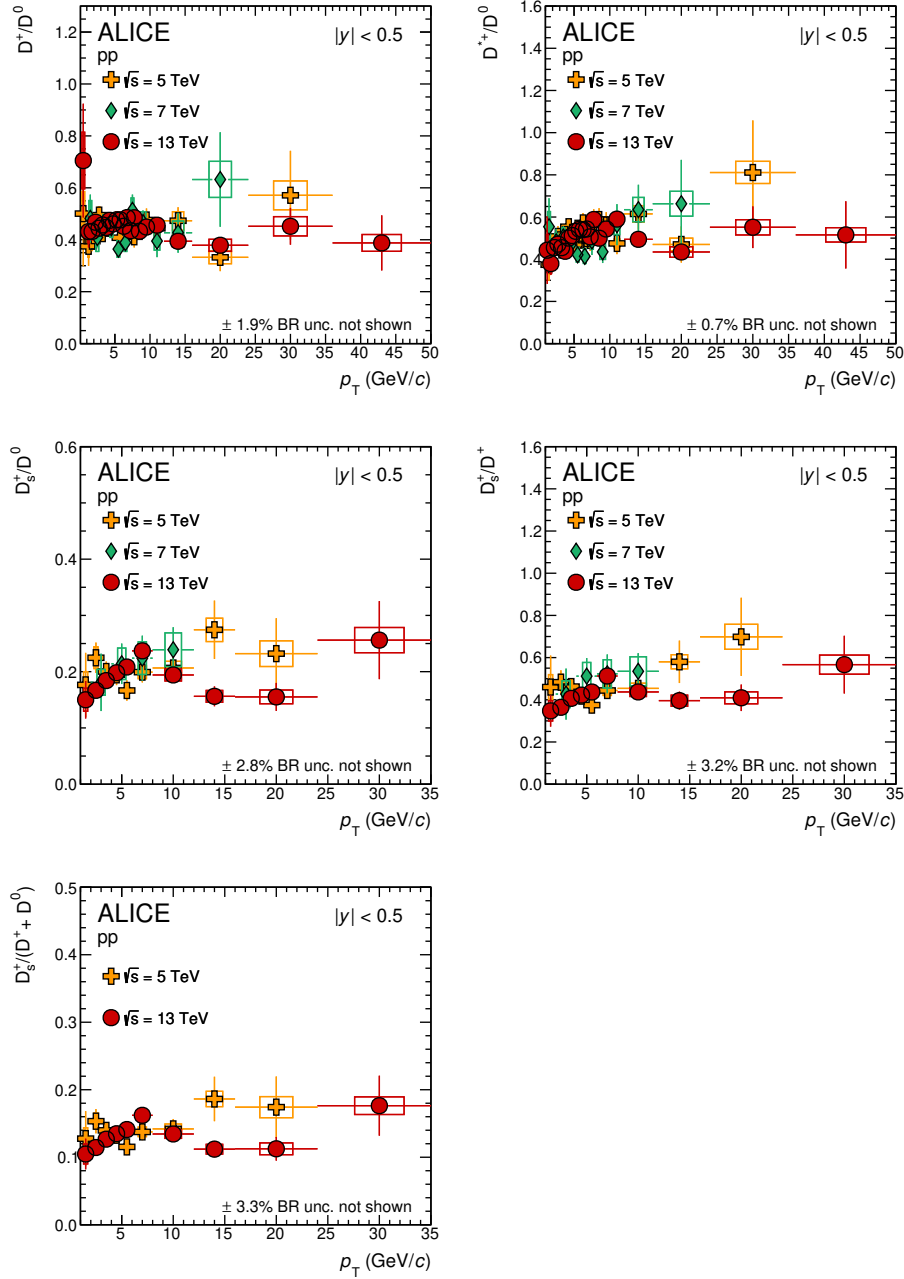
Fig. 5.15 p_T -differential production cross sections for prompt D mesons in comparison with pQCD calculations: FONLL [86, 87] (except for D_s^+ mesons), GM-VFNS framework [88, 89], and a calculation based on the k_T -factorisation approach [92, 246]. The theoretical uncertainties for these predictions are shown as shaded boxes. The vertical bars and boxes report the statistical and systematic uncertainties, respectively. The systematic uncertainties shown in the plots do not include the contributions of the luminosity and the branching ratio, which are reported separately. Figure from Ref. [97].

regardless of whether they lead to pseudoscalar or vector mesons or mesons with or without strange quark content, across different LHC energies. A hint of increase in the region $p_T < 8$ GeV/c is observed for the D_s^+/D^0 , D_s^+/D^+ , and $D_s^+/(D^0 + D^+)$ ratios. Conversely, for the remaining cases, the ratios of p_T -differential cross sections do not exhibit any significant dependence on the p_T of the D mesons, given the experimental uncertainties.

The relationship between the production mechanism of D-mesons and the centre-of-mass energy of the pp collisions can be further explored by computing the p_T -differential ratio of the D-meson cross sections at $\sqrt{s} = 13$ TeV [97] to those at $\sqrt{s} = 5.02$ TeV [239]. Figure 5.17 displays the 13-to-5.02 TeV ratio for prompt D^0 , D^+ , D^{*+} , and D_s^+ mesons at midrapidity ($|y| < 0.5$) in pp collisions [97]. The systematic uncertainties were propagated in the ratio as fully uncorrelated between the results at the two energies, except for those associated with the (non-)prompt-fraction estimation and the branching ratio, which were considered correlated and cancel out in the ratio. The ratios exhibit a common enhancement with increasing p_T due to the hardening of the p_T -differential cross section at higher $\sqrt{s}$ values. This behaviour resembles the one observed for the ratios of the cross section at $\sqrt{s} = 7$ and 5.02 TeV [181], where a good agreement with pQCD calculations was found. Figure 5.18 displays the comparison between the energy ratio for the D^0 meson, for which a FONLL prediction is available, and the predictions from FONLL. A good agreement below one standard deviation is observed.

To further test the dependence of the D-meson production on the collision energy, the p_T -integrated D-meson ratios measured at $\sqrt{s} = 13$ TeV were compared to those measured at $\sqrt{s} = 5.02$ TeV. In the case of D^0 and D^+ mesons, whose p_T -differential cross sections measurements reach $p_T = 0$ GeV/c, the p_T -integrated cross section is equal to the visible cross section being the contribution of the production for $p_T > 50$ GeV/c negligible given the steeply falling shape of the p_T spectrum. The p_T -integrated cross section was calculated by integrating the p_T -differential results over the measured p_T interval. All the systematic uncertainties were propagated assuming them as fully correlated over p_T , except the one associated with the signal extraction, which is assumed to be uncorrelated.

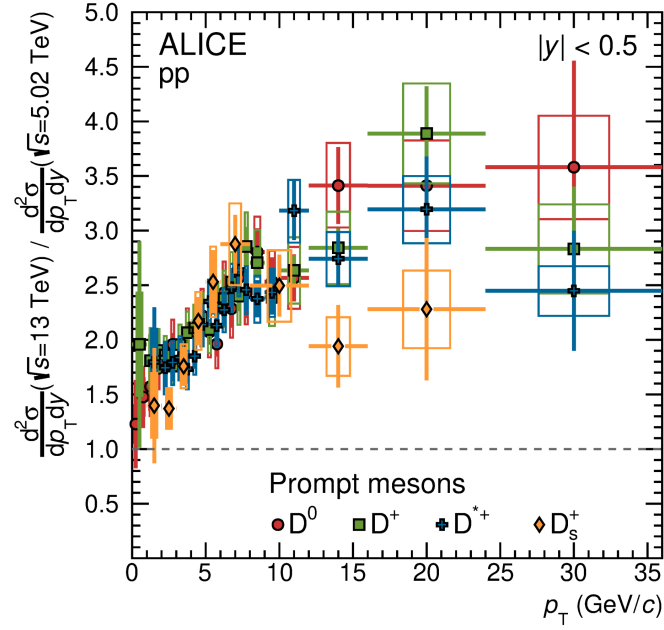
For the D_s^+ meson, for which the measurements stop at $p_T = 1$ GeV/c, the total cross section was obtained by scaling the visible cross section by an extrapolation factor. The procedure to compute the extrapolation factor was the same one adopted in Refs. [239, 181, 97]. This factor was calculated as the ratio of the total cross section estimated with FONLL in the full p_T range to the one in the p_T (and y) interval covered by the measurement. Since, as anticipated in Sec. 5.1.3, FONLL calculations are not available for the prompt D_s^+ mesons, the extrapolation factor was computed by combining the charm-quark p_T



ALI-PUB-546194

Fig. 5.16 Ratios of production cross sections as a function of p_T of prompt D^+/D^0 , D^{*+}/D^0 , D_s^+/D^0 , D_s^+/D^+ , and $D_s^+/(D^0 + D^+)$ mesons at midrapidity ($|y| < 0.5$) in pp collisions at $\sqrt{s} = 13$ TeV [97], $\sqrt{s} = 7$ TeV [182], and $\sqrt{s} = 5.02$ TeV [239]. The vertical bars and boxes report the statistical and systematic uncertainties, respectively. Figure from Ref. [97].

differential cross section, the fragmentation fractions $f(c \rightarrow D_s^+)$ and $f(c \rightarrow D_s^{*+})$ from ALEPH measurements [273], and the charm-quark fragmentation functions from Ref. [272]. The systematic uncertainties on the extrapolation factor include the uncertainties on the PDFs,



ALI - PUB - 546206

Fig. 5.17 Ratios of production cross sections as a function of p_T of prompt D^0 , D^+ , D^{*+} , and D_s^+ mesons at midrapidity ($|y| < 0.5$) in pp collisions at $\sqrt{s} = 13$ TeV [97] to those at $\sqrt{s} = 5.02$ TeV [239]. The vertical bars and boxes report the statistical and systematic uncertainties, respectively. Figure from Ref. [97].

the variation of the charm-quark mass, and the values of the renormalisation and factorisation scales for the FONLL calculations. For the D_s^+ meson, the extrapolation factor results to be equal to $1.21^{+0.26}_{-0.10}$.

The p_T -integrated cross section for D^0 , D^+ , and D_s^+ mesons in pp collisions at $\sqrt{s} = 13$ TeV are reported in Table 5.9. Table 5.10 reports the comparison of the p_T -integrated D-meson ratios at $\sqrt{s} = 13$ TeV and $\sqrt{s} = 5.02$ TeV. In the ratio, the systematic uncertainties associated with the tracking efficiency, the luminosity estimation, and the non-prompt fraction estimation were treated as correlated among the different D-meson species. The other systematic uncertainties were considered uncorrelated. The p_T -integrated D-meson ratios at $\sqrt{s} = 13$ TeV are compatible with the ones measured at $\sqrt{s} = 5.02$ TeV within the current experimental uncertainties. In analogy with the p_T -differential ratios, also in the case of the p_T -integrated D-meson ratios no significant dependence on the collision energy is observed.

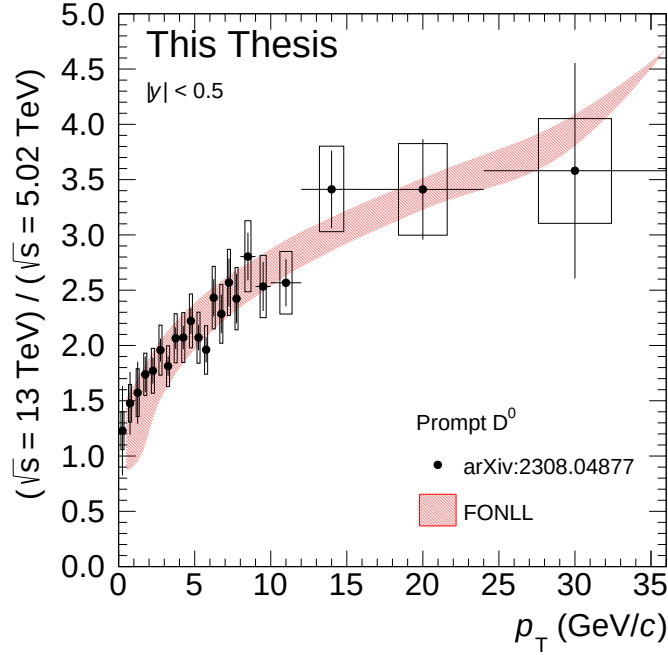


Fig. 5.18 Ratio of the p_T -differential production cross sections of prompt D^0 mesons at midrapidity ($|y| < 0.5$) in pp collisions at $\sqrt{s} = 13$ TeV [97] to those at $\sqrt{s} = 5.02$ TeV [239] compared to the FONLL prediction [86, 87].

Table 5.9 $d\sigma/dy|_{|y|<0.5}$ of D^0 , D^+ , and D_s^+ in pp collisions at $\sqrt{s} = 13$ TeV [97]. These results are obtained by integrating the measured p_T -differential cross sections at midrapidity and extrapolating down to $p_T = 0$ where necessary.

$d\sigma/dy _{ y <0.5} (\mu\text{b}), p_T > 0$	
D^0	749 ± 27 (stat.) $^{+48}_{-50}$ (syst.) ± 12 (lumi.) ± 6 (BR)
D^+	375 ± 32 (stat.) $^{+35}_{-35}$ (syst.) ± 6 (lumi.) ± 6 (BR)
D_s^+	120 ± 11 (stat.) $^{+12}_{-13}$ (syst.) $^{+25}_{-10}$ (extrap.) ± 2 (lumi.) ± 3 (BR)

Strange to-non strange charm-meson production ratio

The strange to non-strange D-meson production cross section ratio presented above enables the study of the charm-quark fragmentation fraction. The fragmentation fractions, indicated with f_x , represent the probability that a charm quark hadronises with a quark of a different flavour x . Experimentally the ratio of the fragmentation fractions is measured as the ratio of the production cross section of a heavy-flavour hadron with a strange quark content (f_s) to the one of hadrons without strange quarks ($f_u + f_d$). In the case of D mesons, this is done by computing the $D_s^+/(D^0 + D^+)$ ratio. This computation is performed under the assumption

Table 5.10 Ratios of the measured production cross sections of prompt D mesons for $p_T > 0$ at midrapidity ($|y| < 0.5$) in pp collisions at $\sqrt{s} = 5.02$ TeV [239] and $\sqrt{s} = 13$ TeV [97].

pp , $\sqrt{s} = 5.02$ TeV [239]	
D^+ / D^0	0.442 ± 0.055 (stat.) ± 0.033 (syst.) ± 0.008 (BR)
D_s^+ / D^0	0.186 ± 0.028 (stat.) ± 0.015 (syst.) $^{+0.051}_{-0.018}$ (extrap.) ± 0.007 (BR)
D_s^+ / D^+	0.419 ± 0.078 (stat.) ± 0.041 (syst.) $^{+0.116}_{-0.040}$ (extrap.) ± 0.017 (BR)
$D_s^+ / (D^0 + D^+)$	0.128 ± 0.020 (stat.) ± 0.010 (syst.) $^{+0.035}_{-0.012}$ (extrap.) ± 0.005 (BR)
pp, $\sqrt{s} = 13$ TeV [97]	
D^+ / D^0	0.500 ± 0.047 (stat.) ± 0.033 (syst.) ± 0.009 (BR)
D_s^+ / D^0	0.160 ± 0.015 (stat.) ± 0.012 (syst.) $^{+0.034}_{-0.013}$ (extrap.) ± 0.004 (BR)
D_s^+ / D^+	0.319 ± 0.039 (stat.) ± 0.030 (syst.) $^{+0.068}_{-0.024}$ (extrap.) ± 0.010 (BR)
$D_s^+ / (D^0 + D^+)$	0.106 ± 0.010 (stat.) ± 0.008 (syst.) $^{+0.023}_{-0.009}$ (extrap.) ± 0.003 (BR)

that the contribution of excited D_s^+ mesons, like the D_{s1}^+ and D_{s2}^{*+} , which exchange their strangeness in the decay is negligible.

The $f_s/(f_u + f_d)$ ratio can be measured by fitting the p_T -differential $D_s^+/(D^0 + D^+)$ ratio with a constant function, assuming the ratio to be flat in p_T , as done in Ref. [239] for the data at $\sqrt{s} = 5.02$ TeV. However, thanks the higher granularity and the improved precision of the measurements achieved in pp collisions at $\sqrt{s} = 13$ TeV compared to that at $\sqrt{s} = 5.02$ TeV, a hint of increasing trend as a function of p_T is observed in the low- p_T region. For this reason, the constant-fit approach is no longer justified, instead the $f_s/(f_u + f_d)$ was measured by integrating the p_T -differential $D_s^+/(D^0 + D^+)$ ratio over the measured p_T interval (i.e. $p_T < 1$ GeV/c). The systematic uncertainties were propagated as described in the previous Section for the p_T -integrated ratios.

The value of the $f_s/(f_u + f_d)$ ratio measured in pp collisions at $\sqrt{s} = 13$ TeV is:

$$\left(\frac{f_s}{f_u + f_d} \right)_{\text{charm}} = 0.116 \pm 0.011(\text{stat.}) \pm 0.009(\text{syst.}) \pm 0.003(\text{BR}) \quad (5.13)$$

and is reported in Fig. 5.19 together with a compilation of previous measurements performed in different collision systems and energies by ALICE [239, 97, 182], H1 [278], ZEUS [279], and ATLAS [280] Collaborations. The $f_s/(f_u + f_d)$ ratio is strictly related to the strangeness

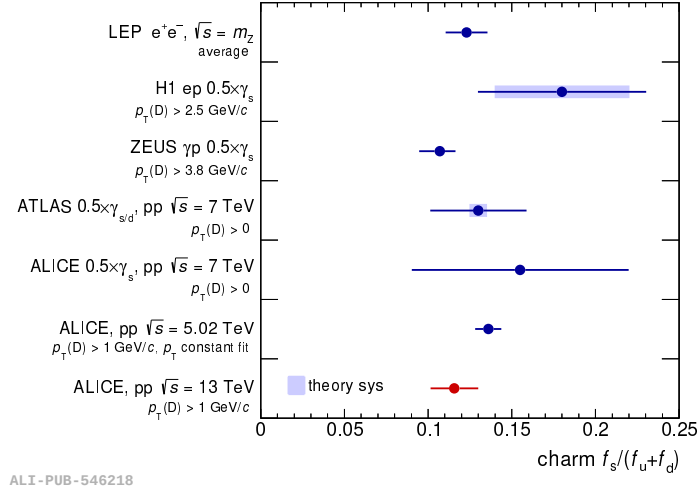


Fig. 5.19 Compilation of the $f_s/(f_u + f_d)$ factor measured in different collision systems and energies by ALICE [239, 97, 182], H1 [278], ZEUS [279], ATLAS [280], and LEP [277] Collaborations. The horizontal bars indicate the total experimental uncertainty, while the shaded boxes represent the theoretical systematic uncertainty. For the measurement of H1, the theoretical uncertainty encompasses the uncertainty on the branching ratio and the model dependencies on the acceptance determination, while for the ATLAS measurement, it consists of the extrapolation uncertainty to the full phase space. Figure from Ref. [97].

suppression factor γ_s [281–283], which is defined as $\gamma_s = 2f_s/(f_u + f_d)$. The γ_s , or *Wroblewski factor*, is a key quantity in the study of the charm-quark hadronisation mechanism in the context of the statistical hadronisation model (refer to Chapter 1), hence some measurements reported in Fig. 5.19 are reported as $\gamma_s \times 0.5$. The horizontal bars indicate the total experimental uncertainty, while the shaded boxes represent the theoretical systematic uncertainty. For the measurement of H1, the theoretical uncertainty encompasses the uncertainty on the branching ratio and the model dependencies on the acceptance determination, while for the ATLAS measurement it consists of the extrapolation uncertainty to the full phase space. The results are compatible among each other and are in agreement with the average of measurements at LEP [277]. This finding reinforces the observation that the ratio of fragmentation fractions of charm quarks to D mesons is independent of the collision system and energy.

5.4.3 Rapidity dependence of prompt D-meson ratios

The rapidity dependence of the D-meson production in pp collisions can be also explored. To this end, the ratio between the production cross sections of the four D-meson species measured by the ALICE Collaboration at midrapidity [97] and the ones measured by the

LHCb Collaboration at forward rapidity [248] in pp collisions at $\sqrt{s} = 13$ TeV is considered. The ratios are shown in Fig. 5.17 for three different rapidity intervals at forward rapidity (top-left panel: $2 < y < 2.5$, top-right panel: $3 < y < 3.5$, bottom-left panel: $4 < y < 4.5$) out of the five measured by the LHCb Collaboration were considered. The uncertainties were propagated in the ratios as completely uncorrelated, except for the BR uncertainties which cancels out in the ratio. The ratios of mid to forward rapidity cross sections show an increasing p_T which becomes more pronounced at larger rapidity. In particular, the magnitude of the increase is larger for the $4 < y < 4.5$ interval compared to the other two rapidity intervals. This behaviour is attributed to the softening of the D-meson p_T spectrum at forward rapidity, and it is correctly described by pQCD calculations, as shown in Refs. [182, 181].

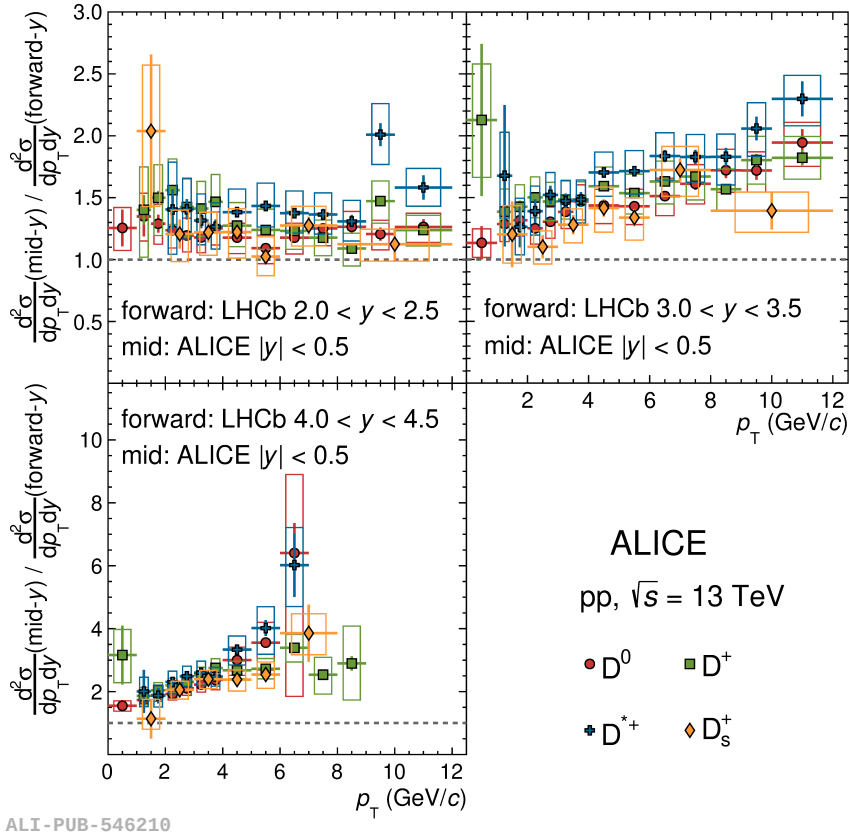


Fig. 5.20 Ratios of the production cross sections as a function of p_T of prompt D^0 , D^+ , D^{*+} , and D_s^+ mesons measured at midrapidity ($|y| < 0.5$) by ALICE [97] to the one measured by LHCb [248] in three rapidity intervals at forward rapidity ($2 < y < 2.5$, $3 < y < 3.5$, and $4 < y < 4.5$) in pp collisions at $\sqrt{s} = 13$ TeV. Figure from Ref. [97].

Ratio of the D^0 meson at $\sqrt{s} = 13$ and 5.02 TeV and different rapidity intervals

As discussed in Refs. [249, 284], precise measurements of D-meson cross section ratios at different centre-of-mass energy and rapidity intervals can significantly constrain the uncertainty on the PDFs in FONLL calculations, particularly at small values of Bjorken- x (10^{-4} to 10^{-5}) where the contribution of gluons is expected to be dominant. The possibility to constrain those uncertainties is due to the fact that the other sources of uncertainty in the FONLL calculations, which relate to the choice of the charm-quark mass and the factorisation and normalisation scales, partially cancel out in the ratio. These double ratio, ρ , was computed only for the D^0 meson considering the measurements performed by the ALICE Collaboration [181, 97] at midrapidity and the LHCb Collaboration [147, 245] at forward rapidity. The D^0 meson was chosen among the others for the following reasons: (i) the D^0 measurement covers the largest p_T interval; (ii) the D^0 measurement is one of the most precise measurements; (iii) given the results shown in Fig. 5.20, a similar trend is expected among the different D-meson species. Hence, the D^0 meson is considered as representative of all the other D-meson species. The double ratio ρ is defined as:

$$\rho = \left[\frac{\frac{d^2\sigma(D^0)_{\text{mid-}y}^{13 \text{ TeV}}}{dp_T dy}}{\frac{d^2\sigma(D^0)_{\text{fwd-}y}^{13 \text{ TeV}}}{dp_T dy}} \right] / \left[\frac{\frac{d^2\sigma(D^0)_{\text{mid-}y}^{5.02 \text{ TeV}}}{dp_T dy}}{\frac{d^2\sigma(D^0)_{\text{fwd-}y}^{5.02 \text{ TeV}}}{dp_T dy}} \right]. \quad (5.14)$$

The top row of Fig. 5.21 displays the p_T -differential ratios between the prompt D^0 -meson production cross section at midrapidity and that at forward rapidity (left panel: $2 < y < 2.5$, middle panel: $3 < y < 3.5$, right panel: $4 < y < 4.5$), separately in pp collisions at $\sqrt{s} = 5.02$ TeV and $\sqrt{s} = 13$ TeV. The results at the two energies are consistent within uncertainties when considering the forward rapidity interval $2 < y < 2.5$. However, at more forward rapidity, the ratio at $\sqrt{s} = 5.02$ TeV systematically exceeds the one at $\sqrt{s} = 13$ TeV energy, suggesting that the hardening of the D^0 p_T -spectrum from forward y to mid y is $\sqrt{s}$ dependent. This behavior reflects the different Bjorken- x values probed by measuring a charm hadron in various rapidity intervals at different center-of-mass energies. At low p_T , these values range from $x \simeq 10^{-4}$ at midrapidity to $x \simeq 10^{-6}$ at $y = 4.5$.

The middle and bottom rows of Fig. 5.21 show the comparison of measured ρ with the predictions based on FONLL calculations employing the CTEQ6.6 [81] and NNPDF30 [285] PDF sets, respectively. The different contributions to the total theoretical uncertainty are shown. They include the value assumed for the charm-quark mass, which was varied between 1.3 and 1.5 $\text{GeV}c^2$, the factorisation and normalisation scales, treated as discussed

in Ref. [286], and the uncertainty on the chosen set of PDFs. The same variations were considered for all the cross sections involved in the calculations of the ratio, and the uncertainty band reported in the Figure corresponds to the envelope of the calculated double ratios. The uncertainty on the PDFs were provided directly by the FONLL authors following the procedure listed in Ref. [81] and Ref. [285] for CTEQ6.6 and NNPDF30, respectively.

The predictions successfully describe the p_T dependence of ρ . In the case of CTEQ6.6, the uncertainties associated with the theoretical predictions are significantly lower than the experimental values for $p_T > 2$ GeV/ c , whereas in the lower- p_T region they are compatible. The dominant contribution originates from the μ_R and μ_F scales. Considering the NNPDF30, the major source of uncertainty is the one on the PDFs especially in the $p_T < 2$ GeV/ c region. Even though the predictions are in agreement with the measurements, the theoretical uncertainties are in general larger than the experimental ones by a factor 2 – 3. These findings demonstrate that considering the current experimental uncertainty, the double ratio ρ represents a powerful tool to quantitatively constrain the gluon PDFs in the low- x region. The p_T range in which the uncertainty on the PDFs is compatible with the one on the scales increase moving from the $2 < y < 2.5$ interval, where the maximum p_T is around 2 GeV/ c , to $4 < y < 4.5$, where it reaches $p_T = 6$ GeV/ c .

A decreasing trend of ρ as a function of p_T is predicted by both the FONLL configurations and is expected to be more significant moving to more forward rapidities. The overall theoretical uncertainty is found to be smaller compared to the experimental one in the most forward rapidity interval. At $4 < y < 4.5$, the high p_T would allow us probing the gluon PDFs in the Bjorken- $x \sim 0.2$, a region where the gluon PDFs are currently not strongly constrained by the measurements. However, this regime is currently not explored due to a lack of measurements, highlighting the need to extend the p_T and y coverage of the D-meson measurements in pp collisions at the LHC, as will be discussed in Chapter 7.

5.4.4 Charm-quark production in pp collisions at $\sqrt{s} = 13$ TeV

Charm-quark fragmentation fractions

The measurements of the production of the prompt D-mesons, together with the measurements of the charm-hadron states (J/ψ , Λ_c^+ , Ξ_c^0 , Ξ_c^+ , and $\Sigma_c^{0,+,++}$) presented in Ref. [97], allow for the measurement of others charm-quark fragmentation fractions $f(c \rightarrow h_c)$ which describe the probability of a charm quark to hadronise into specific charm-hadron species. To compute $f(c \rightarrow h_c)$, the p_T -differential cross sections of every charm-hadron state was

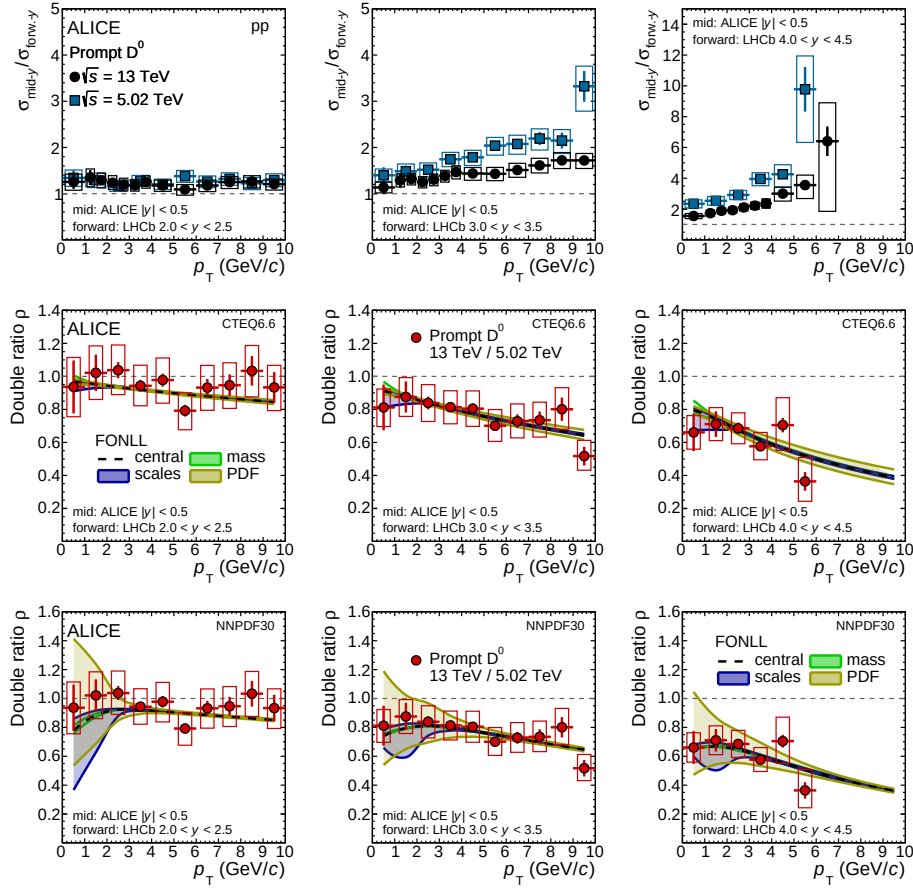


Fig. 5.21 Ratio of the p_T -differential production cross sections of prompt D^0 mesons at midrapidity ($|y| < 0.5$) measured by the ALICE Collaboration [97, 181] to the one measured by the LHCb Collaboration [147, 245] in three rapidity intervals at forward rapidity (left panel: $2 < y < 2.5$, middle panel: $3 < y < 3.5$, right panel: $4 < y < 4.5$) in pp collisions at $\sqrt{s} = 13$ TeV and $\sqrt{s} = 5.02$ TeV. The double ratio ρ is compared to the predictions based on FONLL calculations employing the CTEQ6.6 [81] and NNPDF30 [285] PDF sets in the middle and bottom rows, respectively. Figure from Ref. [97].

integrated in p_T , following the procedure described above for the D mesons, and normalised to the sum of the p_T -integrated production cross section of all the charm-hadron states. More details regarding the integration of the charm-baryon p_T -differential cross section are given in Ref. [97]. The D^{*+} meson and the $\Sigma_c^{0,+,++}$ baryons were excluded from the sum at the denominator given that they strongly decay into the charm-hadron states already considered in the sum, namely the D^0 , the D^+ , and the Λ_c^+ . The systematic uncertainty related to the tracking and the prompt fraction correction were propagated as fully correlated among all the species. The systematic uncertainties on the signal extraction, as well as the statistical uncertainty, were propagated as fully uncorrelated. The uncertainty on the extrapolation factor

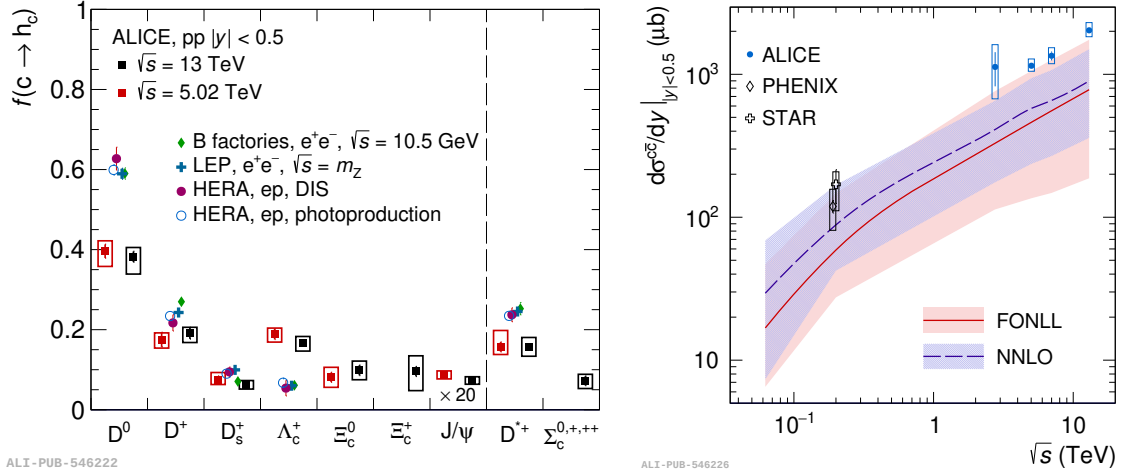


Fig. 5.22 Left: Charm-quark fragmentation fractions measured in pp collisions at $\sqrt{s} = 13$ TeV [97] and $\sqrt{s} = 5.02$ TeV [97] compared to the measurement performed in e^+e^- and ep collisions [287]. Right: Total production cross section of $c\bar{c}$ pairs at midrapidity ($|y| < 0.5$) as a function of $\sqrt{s}$ measured by the ALICE Collaboration at $\sqrt{s} = 5.02$ TeV, $\sqrt{s} = 7$ TeV, and $\sqrt{s} = 13$ TeV [291, 97, 182], and by the STAR [292] and PHENIX [293] Collaborations at $\sqrt{s} = 200$ GeV. Figure from Ref. [97].

was assumed to be partially correlated depending on the strategy adopted in its computation. An additional systematic uncertainty relative to the contribution of the Ω_c^0 was considered, as described in detail in Ref. [97].

The left panel of Fig. 5.22 shows the charm-quark fragmentation fraction measured in pp collisions at midrapidity ($|y| < 0.5$) at $\sqrt{s} = 13$ TeV [97] and $\sqrt{s} = 5.02$ TeV compared to the measurements performed in e^+e^- at LEP and B factories and in ep collisions at HERA [287]. The measured charm-quark fragmentation fractions into D mesons do not show any significant energy dependence within the current experimental uncertainty at LHC energies and are compatible with the measurements performed at lepton colliders. This result reinforces the conclusion that, while the universality of charm fragmentation fractions holds true in the meson sector. The baryon production is enhanced at the LHC with respect to lepton-lepton and lepton-hadron colliders indicating a modification of the hadronisation mechanisms at play in the parton-rich environment produced in pp collisions, regardless of the centre-of-mass energy. The difference between the meson and baryon sector is currently under debate and different models were proposed for the enhancement of the baryon production in pp collisions at the LHC energies [288–290].

Total production cross section of $c\bar{c}$ at midrapidity

The total production cross section of $c\bar{c}$ pairs at midrapidity in pp collisions at $\sqrt{s} = 13$ TeV was measured following a procedure similar to the one described above for the charm-quark fragmentation fractions. The $c\bar{c}$ cross section was computed as the sum of the p_T -integrated cross sections of all the ground state D-meson and charm-baryon states. The $c\bar{c}$ cross section was corrected to account for the different rapidity distribution of the single charm quark compared to the charm hadrons as prescribed in Refs. [239, 182, 97]. The correction factor and its uncertainty were evaluated by computing the ratio of the two rapidity distributions predicted by FONLL calculations in the rapidity range of the measurements and comparing it with the predictions from PYTHIA 8. The ratios were found to be close to unity excluding any significant difference between the two rapidity distributions. A 2% uncertainty was assigned considering the difference between the FONLL and PYTHIA 8 predictions given that the correction factor was found to be equal to 1.021 ± 0.004 . A second correction factor was applied to account for possible differences between the rapidity distribution of the single charm quark and the $c\bar{c}$ pairs. According to NLO pQCD calculations based on POWHEG [96], this correction was evaluated to be 1.036. An uncertainty of 3% was assigned to this correction factor by independently varying the renormalisation and factorisation scales by a factor of two in the calculations and by considering different sets of PDF [82, 294–297]. Finally, a third correction for the Ω_c^0 baryon contribution was applied, as described in Ref. [97]. The systematic uncertainty related to the tracking and the prompt fraction correction were propagated as fully correlated among all the species, while the other sources of systematic uncertainty was treated as uncorrelated. The systematic associated with luminosity and BR are displaced separately.

The resulting total production cross section of $c\bar{c}$ pairs at midrapidity in pp collisions at $\sqrt{s} = 13$ TeV [97] is reported in Table 5.11, together with the analogous result obtained in pp collision at $\sqrt{s} = 5.02$ TeV at midrapidity [97]. The right panel of Fig. 5.22 displays the $c\bar{c}$ cross section at midrapidity ($|y| < 0.5$) as a function of $\sqrt{s}$ measured by the ALICE Collaborations at $\sqrt{s} = 5.02$ TeV, $\sqrt{s} = 7$ TeV, and $\sqrt{s} = 13$ TeV [291, 97, 182], and STAR [292] and PHENIX [293] Collaborations at $\sqrt{s} = 200$ GeV. The measurements confirm the increasing trend versus $\sqrt{s}$ predicted by FONLL and NNLO predictions [298, 299]. The results obtained at RHIC are compatible with the upper-edge of the theoretical predictions, whereas the measurements performed at LHC are systematically slightly underestimated by the calculations and are above the upper edge of the predictions. Considering the current experimental uncertainties, the results at the LHC are compatible with the upper edge of the predictions within 1.4σ (1σ) in pp collisions at $\sqrt{s} = 13$ TeV ($\sqrt{s} = 5.02$ TeV). These

Table 5.11 Total production cross section of $c\bar{c}$ pairs at midrapidity ($|y| < 0.5$) in pp collisions at $\sqrt{s} = 13$ TeV [97] and $\sqrt{s} = 5.02$ TeV [97]. The results are obtained by integrating the measured p_T -differential cross section at midrapidity and extrapolating down to $p_T = 0$ where necessary.

$\sqrt{s}$ (TeV)	$\sigma_{c\bar{c}}^{\text{mid-}y}$ (μb)
5.02	1148 ± 43 (stat.) $^{+62}_{-65}$ (syst.) $^{+98}_{-36}$ (extrap.) ± 43 (BR) ± 24 (lumi.) ± 41 (y)
13	2031 ± 61 (stat.) $^{+135}_{-141}$ (syst.) $^{+196}_{-63}$ (extrap.) ± 97 (BR) ± 33 (lumi.) ± 73 (y)

measurements are crucial to constrain either the uncertainty related to the normalisation and factorisation scales, or on the parton distribution functions.

5.4.5 Non-prompt D-meson p_T -differential cross sections

The non-prompt D-meson cross section in pp collisions at $\sqrt{s} = 13$ TeV was measured following the same procedure employed for the measurement of the prompt D-meson cross section presented above. In this case, a non-prompt-enhanced sample was selected at the stage of the optimisation of the analysis selections. The raw yields were then corrected using the $(\text{Acc} \times \epsilon)$ factor of non-prompt D_s^+ and the $f_{\text{non-prompt}}$ described in Sec. 5.2.3 and Sec. 5.2.2, respectively. The left panel of Fig. 5.23 shows the p_T -differential production cross sections of non-prompt and prompt D^0 , D^+ , and D_s^+ mesons at midrapidity ($|y| < 0.5$). The prompt D-meson cross sections are larger than the non-prompt ones owing to the lower mass of charm with respect to beauty quarks, which makes charm-hadron production more abundant as compared to beauty. The right panel of Fig. 5.23 shows the ratio of the non-prompt to prompt D-meson cross sections for D^0 (top panel), D^+ (middle panel), and D_s^+ (bottom panel) mesons at both $\sqrt{s} = 5.02$ TeV [239] and 13 TeV. In the ratios, the systematic uncertainties were propagated as fully uncorrelated between prompt and non-prompt D mesons, except for the ones associated with the tracking efficiency, the luminosity, and the branching ratio which are fully correlated. The production of prompt D mesons exceeds by a factor of about 20 the non-prompt D mesons one at low- p_T and by a factor of about 10 at $p_T \simeq 10$ GeV/c. The rising trend is due to the harder p_T spectrum of beauty hadrons (h_b) with respect to that of prompt charm mesons. Note that the central values of the ratios of the D_s^+ meson are systematically higher than the ones of the D^0 and D^+ mesons indicating a larger contribution of beauty-hadron decaying into D_s^+ mesons, like in the case of the B_s^0 . This trend is consistent within the experimental uncertainties with the measurements at $\sqrt{s} = 5.02$ TeV. The ratios shown in Fig. 5.23 are linked to the pure non-prompt D-meson fraction in pp collisions at $\sqrt{s} = 13$ TeV which can be computed as the ratio between the non-prompt

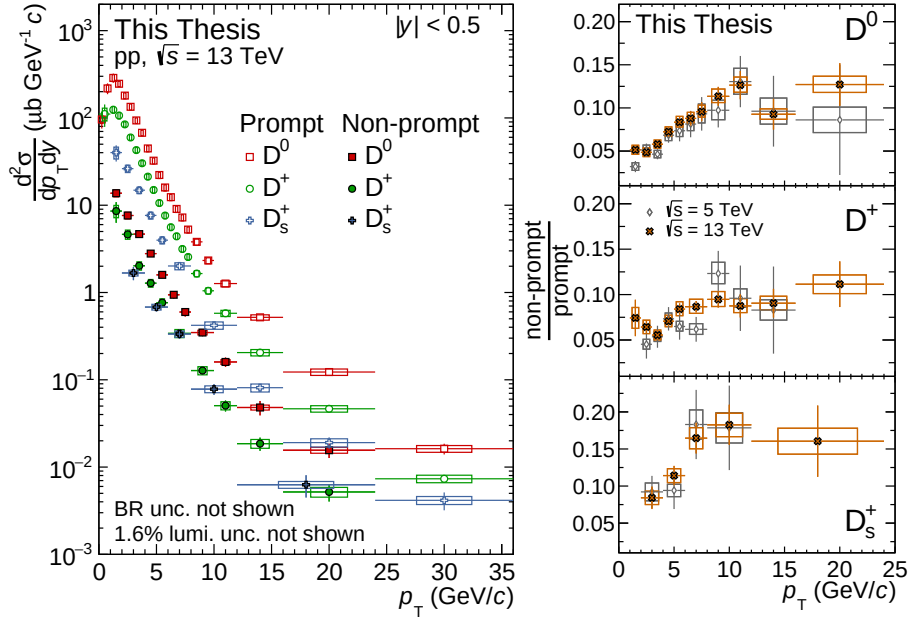


Fig. 5.23 Left: p_T -differential production cross sections of prompt [97] and non-prompt D^0 , D^+ , and D_s^+ mesons at midrapidity ($|y| < 0.5$) in pp collisions at $\sqrt{s} = 13$ TeV. Right: ratios between the production cross sections of non-prompt and prompt D^0 (top), D^+ (middle), and D_s^+ (bottom) mesons at $\sqrt{s} = 5.02$ TeV [239] and $\sqrt{s} = 13$ TeV. The vertical bars and boxes represent the statistical and systematic uncertainties, respectively.

D-meson cross section and the sum of the non-prompt and prompt D-meson cross sections. The pure non-prompt D meson fraction computed as the non-prompt-to-prompt ratios is found to align with the measurements from ALICE [300] where the non-prompt D meson fraction was measured as a function of p_T and in several multiplicity intervals in pp collision at $\sqrt{s} = 13$ TeV.

In analogy with the prompt D mesons, the p_T differential cross section of the non-prompt D_s^+ mesons are compared with pQCD calculations in Fig 5.24. In addition to the predictions from FONLL [86, 87] and GM-VFNS [89, 88, 301] calculations presented above, which employ the beauty-quark fragmentation fraction $f(b \rightarrow h_b)$ from e^+e^- measurements [302], also the predictions from a calculation based on the TAMU model [303] were considered. The latter calculation adopts the p_T -differential beauty-quark cross section and the same fragmentation functions employed by FONLL and a statistical hadronisation approach for the fragmentation fraction $f(b \rightarrow h_b)$. The chemical equilibrium among beauty-hadron species is assumed within the statistical model and the relative abundances of different hadron species are determined solely by the mass of the beauty-hadron and a universal hadronisation temperature ($T_H = 170$ MeV). In the case of FONLL and TAMU predictions,

the beauty hadron cross section is convoluted with the decay kinematics and branching ratios of $h_b \rightarrow h_c + X$ obtained using PYTHIA 8 [244] simulations. The GM-VFNS model implements the transition from the beauty quark to the observed D-meson state via a two-step approach encompassing the hadronisation of the beauty quark into the beauty hadron ($b \rightarrow h_b$) and its decay into the measured D meson ($h_b \rightarrow D + X$). An alternative single-step-transition method exists ($b \rightarrow D + X$), which significantly underestimates the predictions derived from the former approach and the measured cross sections, as already observed in Ref. [239]. For these reasons, the results obtained with this method are not shown. The FONLL and GM-VFNS uncertainties includes the same sources discussed in the previous Sections, while no uncertainty estimation is available for the TAMU model predictions.

The FONLL calculations describe the measured non-prompt D_s^+ -meson production cross sections within uncertainties. In particular, the data are consistent with the central values of the FONLL+PYTHIA 8 predictions. The TAMU predictions tend to overestimate the yield of non-prompt D_s^+ mesons, though the absence of theoretical uncertainty does not allow for any firm conclusions. The GM-VFNS calculations underestimate the measurements in the low- p_T region ($p_T < 8$ GeV/c), whereas at higher p_T a better agreement is found.

5.4.6 Non-prompt D meson cross section ratios

To investigate the beauty quark hadronisation mechanisms the relative production of different species of non-prompt D-mesons was measured. The left panel of Fig. 5.25 shows the prompt [97] and non-prompt p_T -differential $D_s^+/(D^+ + D^0)$ ratio. The procedure described above for the charm sector was adopted to propagate the systematic uncertainty in the ratio. Unlike the prompt $D_s^+/(D^+ + D^0)$ ratio, the non-prompt one does not show any significant dependence on p_T given the current experimental uncertainties. The results are shown to be fully in line with the measurements performed by the ALICE Collaboration at $\sqrt{s} = 5.02$ TeV [239] indicating that, regardless the collision energy, the fragmentation fraction of charm and beauty quarks into D mesons stays the same. However, as already observed in the charm sector, also in the beauty sector the fragmentation fraction into baryons strongly depends on the collision system, as observed by the non-prompt $\Lambda_c^+/(D^0)$ ratio measured by the ALICE Collaboration [238] and by the $\Lambda_b^0/(B^0 + B^+)$ ratio measured by the LHCb Collaboration [147].

The right panel of Fig. 5.25 displays the comparison of the measured prompt and non-prompt $D_s^+/(D^+ + D^0)$ ratio with the predictions based on FONLL+PYTHIA 8⁴. The

⁴No FONLL prediction is currently available for the D_s^+ meson.

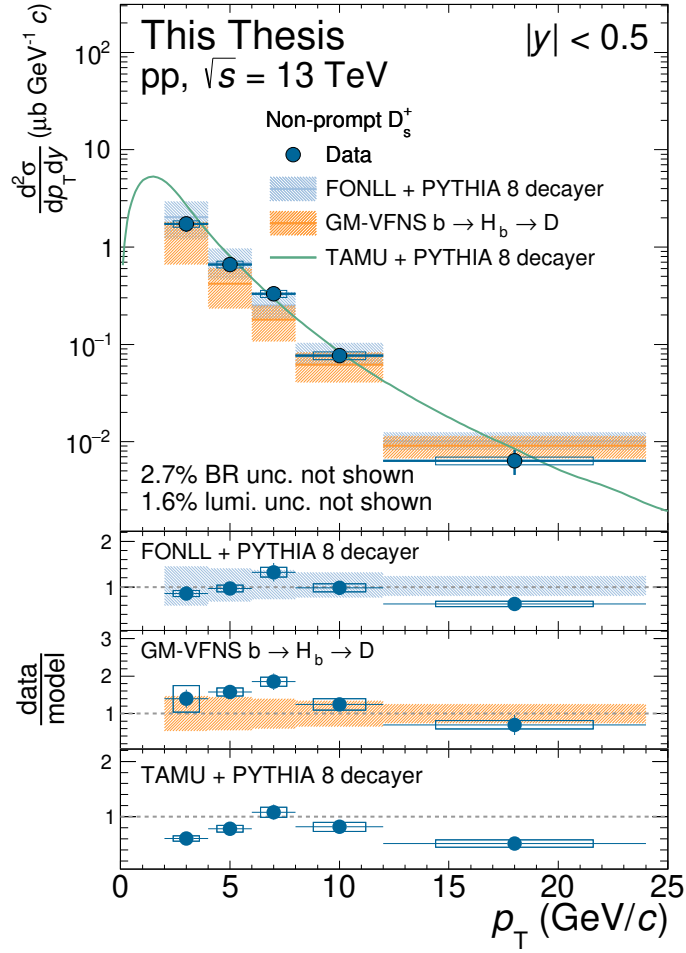


Fig. 5.24 p_T -differential production cross sections of non-prompt D_s^+ mesons at midrapidity ($|y| < 0.5$) in pp collisions at $\sqrt{s} = 13$ TeV compared with the predictions from FONLL + PYTHIA 8 [86, 87], GM-VFNS [89, 88], and TAMU + PYTHIA 8 [98] calculations. The lower panels report the data-to-model ratios. The vertical bars and boxes represent the statistical and systematic uncertainties, respectively.

theoretical predictions, which include a large amount of D_s^+ mesons originating from the decay of non-strange B mesons, are in agreement with the measurements.

Energy dependence of non-prompt D meson cross section ratios

To further explore the relationship between the non-prompt D-meson production and the centre-of-mass energy of the collision system, the ratio of the p_T -differential non-prompt D-meson cross sections measured at $\sqrt{s} = 13$ TeV [97] to the ones measured at $\sqrt{s} = 5.02$ TeV [239] was computed for D^0 , D^+ , and D_s^+ mesons. The systematic uncertainties

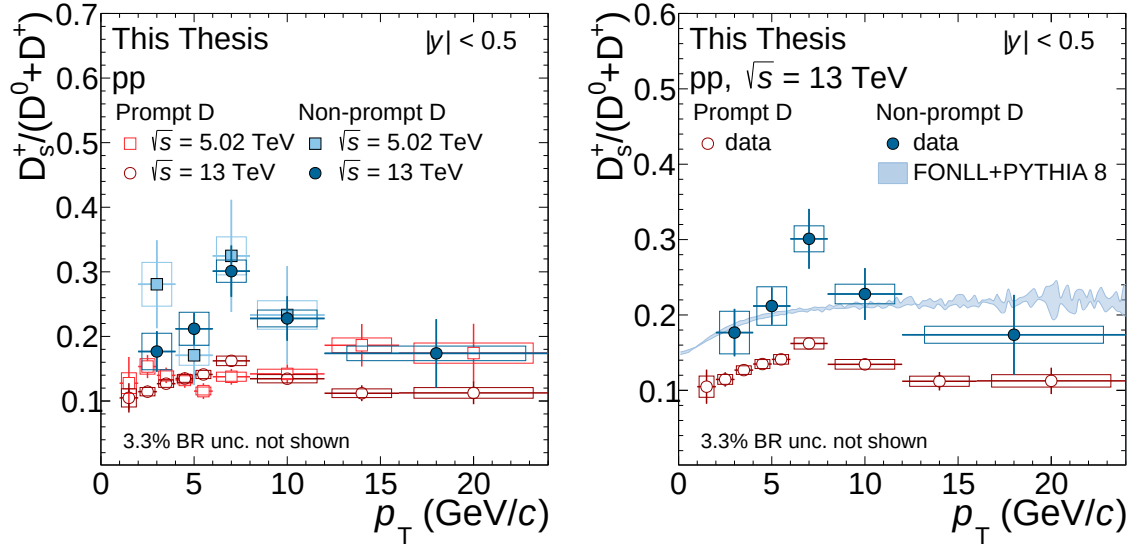


Fig. 5.25 Left: prompt and non-prompt $D_s^+/(D^+ + D^0)$ ratios as a function of p_T in pp collisions at $\sqrt{s} = 5.02$ [239] and $\sqrt{s} = 13$ [97] TeV. Right: prompt [97] and non-prompt $D_s^+/(D^+ + D^0)$ ratios as a function of p_T in pp collisions at $\sqrt{s} = 13$ TeV compared with FONLL+PYTHIA 8 predictions. The vertical bars and empty boxes report the statistical and systematic uncertainties, respectively.

were propagated as uncorrelated between the two collision energies, except for those related to the non-prompt fraction estimation and the branching ratio, which were treated as fully correlated.

The left panel of Fig. 5.26 shows the ratio for the different D-meson species. The ratios are compatible among each other within the experimental uncertainties and show a hint of a common increasing trend as a function of p_T , which is compatible to what observed in the charm sector in Fig. 5.17 and in the ratio for B^+ mesons in the rapidity interval $2.0 < y < 4.5$ measured by the LHCb Collaboration [304]. This is an expected consequence of the hardening of the heavy-quark p_T spectrum with increasing $\sqrt{s}$. The right panel of the same Figure shows the comparison between the measured 13-to-5 TeV ratio for non-prompt D^0 mesons and the FONLL+PYTHIA 8 predictions. In analogy with the charm sector, the non-prompt D^0 meson was chosen among the other D-meson states due to the wider p_T coverage of its measurement and the higher precision. The theoretical predictions describe the increasing trend as a function of p_T and are in agreement with the measurement within the experimental uncertainty.

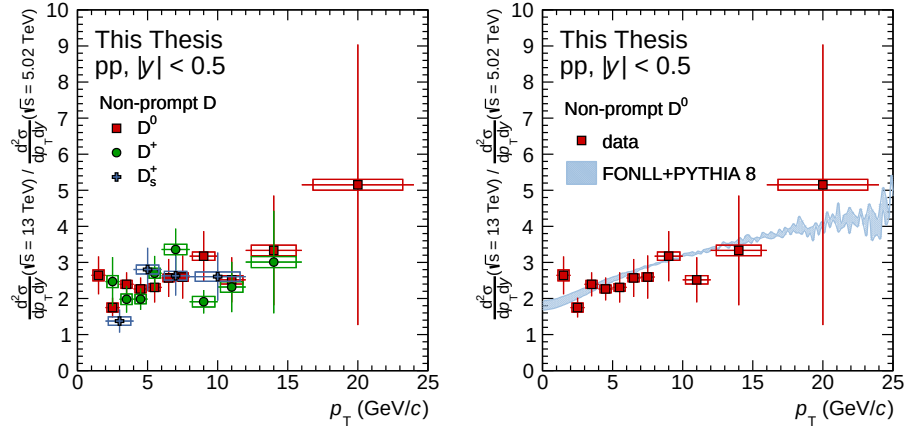


Fig. 5.26 Ratios of non-prompt D-meson production cross sections at $\sqrt{s} = 13$ and $\sqrt{s} = 5.02$ TeV [239] as a function of p_T . Left: comparison of measurements for D^0 , D^+ , and D_s^+ mesons. Right: comparison of D^0 results with FONLL+PYTHIA 8 calculations. The vertical bars and boxes report the statistical and systematic uncertainties, respectively.

5.4.7 Beauty-quark production in pp collisions at $\sqrt{s} = 13$ TeV

Beauty-quark fragmentation fractions

The measurement of the fragmentation fractions of beauty quarks into D mesons is extremely important to understand the hadronisation mechanism in the heavy-flavour sector. It can be derived from the strange to non-strange D meson ratio, similarly to what was done for the charm-quark fragmentation fraction. However, in the beauty sector a slightly different procedure is adopted compared to the one presented above for the charm quark. The procedure is indeed similar to the one discussed in Ref. [239]. Given that a significant fraction of the non-prompt D_s^+ mesons originates from decays of non-strange B mesons, a correction factor, $\alpha_{\text{corr.}}^{\text{FONLL+PYTHIA 8}}$, is applied to scale the p_T -differential non-prompt $D_s^+/(D^0 + D^+)$ ratio. The correction factor is calculated from the FONLL+PYTHIA 8 predictions as:

$$\alpha_{\text{corr.}}^{\text{FONLL+PYTHIA 8}} = \left[\frac{N(D_s^+ \leftarrow B_s^0)}{N(D_s^+ \leftarrow h_b)} \times \frac{N(D^0, D^+ \leftarrow h_b)}{N(D^0, D^+ \leftarrow B^{0,+})} \right]^{\text{FONLL+PYTHIA 8}}, \quad (5.15)$$

where $N(D_s^+ \leftarrow B_s^0)$ is the number of D_s^+ mesons produced in the decay of B_s^0 mesons, $N(D^0, D^+ \leftarrow B^{0,+})$ is the number of non-strange D mesons originating from non-strange B-meson decays, $N(D_s^+ \leftarrow h_b)$ is the number of D_s^+ from beauty-hadron decays, and $N(D^0, D^+ \leftarrow h_b)$ is the number of D^0 and D^+ mesons originating from beauty-hadron decays. The $\alpha_{\text{corr.}}^{\text{FONLL+PYTHIA 8}}$ term is found to be close to 0.5 in the full p_T range of the measurement. Given that the majority of the non-prompt D^0 and D^+ originate from

non-strange B mesons, the obtained value of $\alpha_{\text{corr.}}$ indicates that approximately 50% of the non-prompt D_s^+ mesons originate from decays of strange B mesons. Note that the value of $\alpha_{\text{corr.}}$ is consistent with the one reported in Ref. [239] for pp collisions at $\sqrt{s} = 5.02$ TeV. The uncertainty on $\alpha_{\text{corr.}}^{\text{FONLL+PYTHIA 8}}$ was estimated by varying the beauty-quark fragmentation fractions ($f(b \rightarrow h_b)$) and the branching ratios of the $h_b \rightarrow D + X$ decays in PYTHIA 8. The alternative set of ($f(b \rightarrow h_b)$) was taken from the one measured in $p\bar{p}$ collisions reported in Ref. [6]. The branching ratios instead were varied by weighting the ones implemented in PYTHIA 8 to match the ones measured Ref. [6]. In Ref. [239] it was verified the correction factor obtained with the EvtGen package [305] to describe the decay process was consistent with the one computed with PYTHIA 8. The uncertainty on the correction factor $\alpha_{\text{corr.}}^{\text{FONLL+PYTHIA 8}}$ is considered in the computation of the uncertainty on the beauty quark $f_s/(f_u + f_d)$, which includes the uncertainty on the $D_s^+/(D^0 + D^+)$ ratio.

Unlike what was done in the charm sector, the computation of the beauty-quark fragmentation fractions is performed via the constant-fit procedure. This choice was driven by the fact that, considering the current uncertainty, the non-prompt $D_s^+/(D^0 + D^+)$ ratio is compatible with a flat distribution as a function of p_T . In addition, this procedure leads to a lower uncertainty compared to the one adopted in charm sector. The resulting beauty-quark fragmentation fraction measured in pp collisions at midrapidity ($|y| < 0.5$) at $\sqrt{s} = 13$ TeV is:

$$\left(\frac{f_s}{f_u + f_d} \right)_{\text{beauty}} = 0.114 \pm 0.016(\text{stat.}) \pm 0.010(\text{syst.}) \pm 0.004(\text{BR}) \pm 0.004(\text{extrap.}) . \quad (5.16)$$

where "stat" ("syst") denotes the statistical (systematic) uncertainty, "BR" is the uncertainty on the BR of the considered decays.

Figure 5.27 compares the measured fragmentation fraction ratio ($f_s/(f_u + f_d)$) with previous measurements in pp and $p\bar{p}$ collisions at different energies from the CDF [306], ALICE [239], ATLAS [307], and LHCb [308, 147] Collaborations. The ATLAS measurement of f_s/f_d was divided by a factor of two assuming isospin symmetry for the u and d quarks (i.e. $f_u = f_d$). The beauty-quark fragmentation fractions extracted from the measurements in pp ($p\bar{p}$) collisions are compatible with those computed by the HFLAV Collaboration [309] in e^+e^- collisions at LEP and with the outcome of PYTHIA 8 simulations with the Monash-13 tune [244] which are shown as a blue dashed line in Fig. 5.27. Similarly to the ATLAS measurement, also the HFLAV results were scaled by a factor of two. A significant improvement in the precision of the measurement is observed owing to the larger data sample collected by ALICE at $\sqrt{s} = 13$ TeV compared to lower collision energies. Notably, the fragmentation

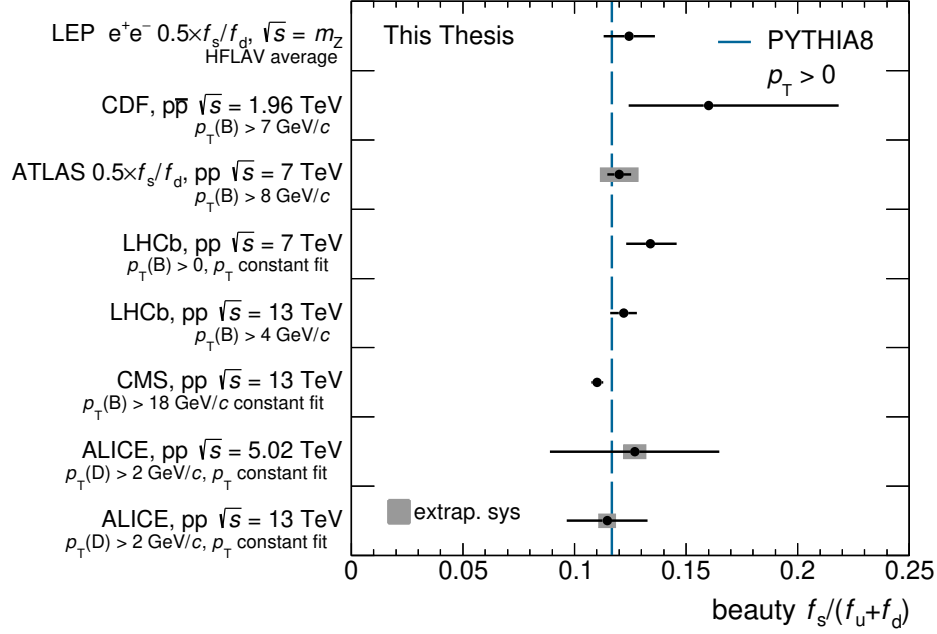


Fig. 5.27 Beauty-quark fragmentation-fraction ratio $f_s/(f_u + f_d)$ calculated from non-prompt D-meson measurements in pp collisions at $\sqrt{s} = 13$ TeV compared with previous measurements from CDF [306], ALICE [239], ATLAS [307], CMS [310], and LHCb [308, 147], with the average of LEP measurements [309], and with the results of PYTHIA 8 simulations with Monash-13 tune [244].

fraction ratio $f_s/(f_u + f_d)$ for beauty quarks presented in this Thesis is compatible with the one measured for the charm quark at a centre-of-mass energies of 13 [97] and 5.02 [239] TeV, and with the predicted value of the Wroblewski factor (refer to Sec. 5.4.2) for light-flavour particles in the statistical hadronisation model, which is of the order of 0.1 [282].

Total production cross section of $b\bar{b}$ at midrapidity

The visible cross sections of the different non-prompt charm-hadron species were obtained by integrating the p_T -differential cross section in the measured p_T interval, as done for the prompt D mesons. In this procedure, the statistical uncertainty and the systematic uncertainty on the raw-yield extraction were treated as uncorrelated among the different p_T intervals, while all the other sources were considered as correlated. Since the measurements do not reach $p_T = 0$, they are extrapolated in the unmeasured region. The extrapolation factor was computed as the ratio of the beauty-quark production cross section and the visible cross sections of the measured non-prompt charm-hadron states in the p_T range of the analyses, as predicted by FONLL+PYTHIA 8, as:

$$\alpha_{\text{extrap.}}^{\text{bb}} = \frac{d\sigma_{\text{bb}}/dy|_{|y|<0.5}^{\text{FONLL}}}{d\sigma_{\text{hc}}/dy|_{|y|<0.5}^{\text{FONLL+PYTHIA 8}}(p_{\text{T}}^{\text{min}} < p_{\text{T}} < p_{\text{T}}^{\text{max}})} . \quad (5.17)$$

The systematic uncertainty of the extrapolation factor was determined by considering the uncertainty associated with the FONLL calculations, the uncertainties on the fragmentation fractions of the beauty quark, and the uncertainties on the BR of the decay channel, as was done for the estimation of the beauty-quark fragmentation fractions described in Sec. 5.4.7. Two correction factors were considered in the calculations to account for the different rapidity distributions of beauty quarks and beauty hadrons and the different rapidity distributions of $\text{b}\bar{\text{b}}$ pairs and beauty quarks. The first correction factor was estimated to be unity within the relevant rapidity range, based on FONLL calculations. An uncertainty of 1% was derived from the deviation between FONLL and PYTHIA 8. The second correction factor was computed from the rapidity distributions of b quarks and $\text{b}\bar{\text{b}}$ pairs obtained with POWHEG [311] simulations. The resulting value is 1.06 ± 0.01 within the range $|y| < 0.5$, where the uncertainty was derived by varying the factorisation and renormalisation scales in the POWHEG calculation.

The production cross section of $\text{b}\bar{\text{b}}$ pairs per unit of rapidity at midrapidity was computed starting from the cross section of each non-prompt charm hadron species. The production cross section of $\text{b}\bar{\text{b}}$ pairs was found to equal to $74.2 \pm 2.7(\text{stat}) \pm 5.5(\text{syst}) \pm 1.2(\text{lumi}) \pm 0.6(\text{BR})_{-0.5}^{+12.9}(\text{extrap}) \mu\text{b}$ for the D^0 , $85.2 \pm 10.9(\text{stat}) \pm 10.6(\text{syst}) \pm 1.4(\text{lumi}) \pm 1.5(\text{BR})_{-0.6}^{+33.7}(\text{extrap}) \mu\text{b}$ for the D^+ , and $62.0 \pm 6.5(\text{stat}) \pm 4.4(\text{syst}) \pm 1.0(\text{lumi}) \pm 1.7(\text{BR})_{-5.0}^{+5.5}(\text{extrap}) \mu\text{b}$ for the D_s^+ . Then the final estimate was computed by averaging the obtained values, adopting the inverse of the quadratic sum of the absolute statistical and uncorrelated systematic uncertainties as weights. The measured $\text{b}\bar{\text{b}}$ production cross section per unit of rapidity at midrapidity in pp collisions at $\sqrt{s} = 13$ TeV is:

$$\left. \frac{d\sigma}{dy} \right|_{|y|<0.5}^{\text{pp}, \sqrt{s}=13 \text{ TeV}} = 75.2 \pm 3.2(\text{stat.}) \pm 5.2(\text{syst.})_{-3.2}^{+12.3}(\text{extrap.}) \mu\text{b} . \quad (5.18)$$

The left panel of Fig. 5.28 displays a compilation of $\text{b}\bar{\text{b}}$ cross section measurements in pp collisions from the ALICE [312–317] and PHENIX [318] Collaborations, and in $\text{p}\bar{\text{p}}$ from the CDF [319] and UA1 [320] Collaborations, as a function of $\sqrt{s}$. The measurement of the total beauty-quark production cross section presented in this Thesis is fully compatible with the one estimated from the measurements of non-prompt D^0 mesons non-prompt Λ_c^+ baryons at midrapidity ($|y| < 0.5$) [321], and with the one obtained from the measurement of non-prompt J/ψ ($|y| < 0.5$) [312] performed by the ALICE Collaboration at the same collision energy.

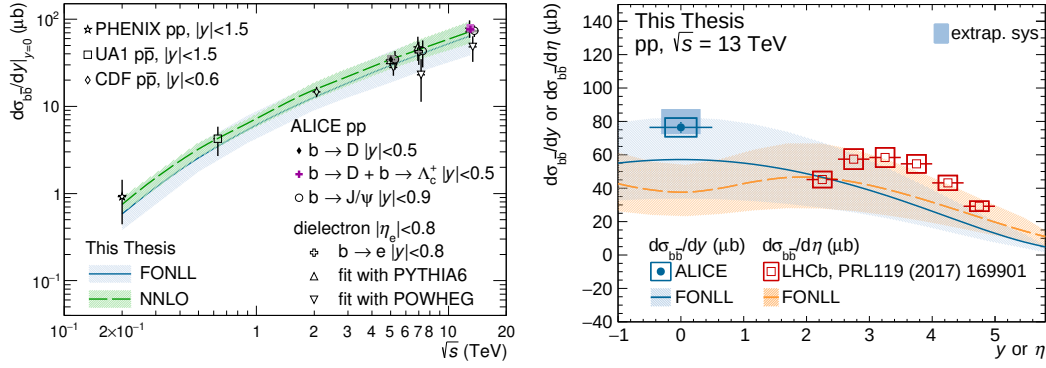


Fig. 5.28 Left: beauty-quark production cross section per unit of rapidity at midrapidity, as a function of the centre-of-mass energy measured in pp and $p\bar{p}$ collisions by the ALICE [312–317], PHENIX [318], CDF [319] and UA1 [320] Collaborations. The solid and dashed lines, accompanied by a shaded band, represent the central values and the associated uncertainties predicted by FONLL [86, 323, 87] and NNLO [324] calculations, respectively. Right: beauty-quark production cross section per unit of (pseudo)rapidity as a function of (pseudo)rapidity measured by the ALICE and LHCb [322] Collaborations. The solid and dashed lines, accompanied by a shaded band, represent the central values and the associated uncertainties predicted by FONLL [86, 323, 87]. The vertical bars and boxes report the statistical and systematic uncertainties, respectively.

The experimental results are compared with the predictions from FONLL and NNLO calculations. The measured $b\bar{b}$ cross sections and their trend with $\sqrt{s}$ are compatible with the predictions from pQCD calculations within the theoretical uncertainties, which are smaller in the case of NNLO calculations. The central values of the NNLO calculations are found to be closer to the data, as expected by the higher perturbative accuracy. Finally, the dependence of the $b\bar{b}$ production cross section on the rapidity interval was investigated. The right panel of Fig. 5.28 reports the $b\bar{b}$ production cross section at midrapidity presented in this work along with the measurements performed in pseudorapidity intervals by the LHCb Collaboration at forward rapidity [322]. The experimental results are compared to FONLL predictions, which are shown both for $d\sigma_{b\bar{b}}/dy$ and $d\sigma_{b\bar{b}}/d\eta$, to match the observables reported by ALICE and LHCb, respectively. The measured $b\bar{b}$ cross sections generally lie close to the upper edge of the FONLL theoretical uncertainty band, except for the LHCb data point in $2 < \eta < 2.5$ which is compatible with the central value of the predictions.

The measurements of non-prompt D mesons revealed to be a key tool to investigate the production of beauty quarks in pp collisions at midrapidity ($|y| < 0.5$) at low p_T and to compare the dependence of the beauty production as a function of p_T and y between data and model calculations. These results, together with the analogous ones obtained in the charm sector, will be used to further constrain pQCD calculations and hadronisation models.

5.5 D_{s1}^+ and D_{s2}^{*+} meson production in pp collisions at $\sqrt{s} = 13$ TeV

The p_T -integrated production yields of prompt D_{s1}^+ and D_{s2}^{*+} mesons is computed as:

$$\frac{d^2N(D_{\text{reso}})}{dp_T dy} = \frac{f_{\text{prompt}} \times N_{\text{raw}}^{\text{Dreso}} \times \epsilon^{\text{trigger}}}{2 \times \Delta p_T \times c_{\Delta y} \times (\text{Acc} \times \epsilon)_{\text{prompt}} \times \text{BR}(D_{\text{daugh.}}) \times \text{BR}(V0) \times N_{\text{ev}}}, \quad (5.19)$$

where $N_{\text{raw}}^{\text{Dreso}}$ is the raw yield of the resonances extracted as described in Sec. 5.5.2 and f_{prompt} is the p_T -integrated prompt fraction computed as described in Sec. 5.5.4, respectively. As before, factor 2 takes into account that the raw-yields are measured for the sum of particles and antiparticles, while the corrected yields are given for their average. Δp_T is the width of the analysed p_T interval ($2 < p_T < 24$ GeV/c), $c_{\Delta y}$ is the correction factor for the rapidity coverage, $(\text{Acc} \times \epsilon)_{\text{prompt}}$ is the p_T -integrated acceptance-times-efficiency correction for prompt D_{s1}^+ and D_{s2}^{*+} mesons computed as described in Sec. 5.5.3, $\epsilon^{\text{trigger}}$ is the trigger efficiency associated with the MB ($\epsilon^{\text{trigger}} = 0.92$, taken from Ref [270]) and HM ($\epsilon^{\text{trigger}} = 1$) multiplicity classes of events, $\text{BR}(D_{\text{daugh.}}) \times \text{BR}(V0)$ is the product of the branching ratios of the D meson and the V0 produced in the resonance decay taken from Ref. [6], and N_{ev} is the number of events.

5.5.1 D_{s1}^+ and D_{s2}^{*+} candidate selection

The D_{s1}^+ and D_{s2}^{*+} mesons were reconstructed via the $D_{s1}^+ \rightarrow D^{*+} K_S^0$ and $D_{s2}^{*+} \rightarrow D^+ K_S^0$ decay channels [6], respectively. To reconstruct the excited strange D-meson states, the K_S^0 candidates were combined with the appropriate non-strange D-meson candidate. The K_S^0 , thanks to their large decay length, can be selected with high purity by applying standard linear selections on their displaced decay-vertex topology, as the ones reported in Ref. [325]. Given the large contribution of the background and the necessity to separate the prompt from the non-prompt D mesons, the selection of the D-meson candidates requires a more sophisticated analysis procedure to achieve reliable statistical significance. The D^+ and the D^0 candidates were selected following a procedure similar to the one presented in this Chapter for the D_s^+ mesons based on optimised selections on the BDT output score.

Figure 5.29 shows the correlation between the invariant-mass distribution of the selected K_S^0 candidates ($M(\pi\pi)$) and those of selected D^+ ($M(K\pi\pi)$) or D^{*+} ($\Delta M = M(K\pi\pi) -$

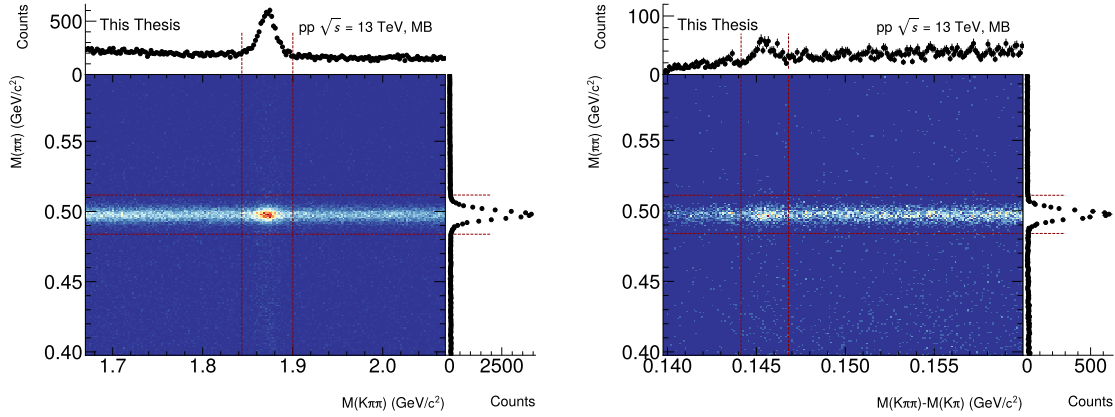


Fig. 5.29 Left: correlation between the invariant-mass distribution of the selected K_S^0 candidates ($M(\pi\pi)$) and those of selected D^+ ($M(K\pi\pi)$) candidates used to build D_{s2}^{*+} candidates in the MB sample. Right: correlation between the invariant-mass distribution of the selected K_S^0 candidates ($M(\pi\pi)$) and those of selected D^{*+} ($\Delta M = M(K\pi\pi) - M(K\pi)$) candidates used to build D_{s1}^+ candidates in the MB sample. The projection on each axis is also shown. The red dotted lines delimit the invariant-mass region of the signal selected to build the D_{s2}^{*+} and D_{s1}^+ candidates.

$M(K\pi)$ candidates used to build the excited D_s^+ -meson candidates, respectively. The results reported in Fig. 5.29 are extracted from the sample of D mesons and V0 candidates selected from MB events. Similar distributions are found in the HM sample. The red dotted lines delimit the invariant-mass region of the signal selected to build the D_{s2}^{*+} and D_{s1}^+ candidates. These regions were defined by fitting the invariant-mass distributions of D-meson and K_S^0 candidates using the flarefly fitting package [326]. This package consists of a wrap of several python packages employed in the field of high-energy physics and was developed to simplify the fits to the invariant-mass distribution in the case of particular shapes, such as the ones of the resonances, and of unbinned maximum likelihood fits. For D^+ mesons, the fitting function was composed of an exponential term for the background and a Gaussian term for the signal peak. For D^{*+} mesons, the background was parametrised with an exponential times a power law function, $p_0\sqrt{\Delta M - M_\pi}e^{p_1(\Delta M - M_\pi)}$, where p_0 and p_1 are free parameters, and M_π is the mass of the charged pion. The signal region was then defined as three times the width of the D-meson invariant-mass peak. The invariant-mass distribution of K_S^0 candidates was fitted with a sum of two Gaussian functions to describe the signal and a constant term for the background. In this case, the signal region was defined as $|M(\pi\pi) - M(K_S^0)| < 3\sigma_{\text{eff}}$. The resolution parameter in this case was defined as $\sigma_{\text{eff}} = f_1\sigma_1 + (1 - f_1)\sigma_2$, where σ_1 and σ_2 are the widths of the two Gaussian functions, and f_1 is the fraction of integral belonging to the first Gaussian function. The D_{s1}^+ - and D_{s2}^{*+} - meson candidates built from the signal region indicated above were all kept in the analyses, and no additional selection was applied.

Given the minimum p_T requirement for D^+ and D^{*+} candidates ($p_T > 1$ GeV/c) there are essentially no D_{s1}^+ and D_{s2}^{*+} candidates below $p_T = 2$ GeV/c. In addition, given the limitation on the amount of available data and the rarity of these states, almost no candidates were found for $p_T > 24$ GeV/c. Therefore, a kinematic region of $2 < p_T < 24$ GeV/c was chosen for the measurement of these two excited states. An attempt was made to perform a p_T -differential measurement in both the multiplicity intervals, but the statistical significance was too low to achieve a reliable result. Hence, it was decided to perform the measurement integrated over the p_T range of $2 < p_T < 24$ GeV/c.

5.5.2 Raw yield extraction

The D_{s1}^+ and D_{s2}^{*+} raw yields in the $2 < p_T < 24$ GeV/c interval for the MB and HM samples were extracted with a fit to the invariant-mass distribution differences $M(D^{*+}K_S^0) - M(D^{*+})$ and $M(D^+K_S^0) - M(D^+)$, respectively. The fitting function is composed of a Voigtian function for the signal peak [327] and a background term. The Voigtian function, defined as the convolution of a Cauchy distribution and a Gaussian distribution, was chosen given the width of the D_{s1}^+ and D_{s2}^{*+} resonances ($\Gamma(D_{s1}^+) = 0.92 \pm 0.05$ MeV, $\Gamma(D_{s2}^{*+}) = 16.9 \pm 0.7$ MeV [6]), which is comparable with the resolution on the invariant mass (especially in the case of the D_{s1}^+ state). The width of the resonances was fixed to the PDG value in the fits [6]. The background distribution of D_{s1}^+ candidates was described with a exponential times a power law function, $p_0 \sqrt{\Delta M - M_{K_S^0}} e^{p_1(\Delta M - M_{K_S^0})}$ given that the D_{s1}^+ peak is close to the $D^{*+}K_S^0$ threshold, while the one of D_{s2}^{*+} candidates with a polynomial function of the first order.

Figure 5.30 shows the fits to the invariant-mass distributions for the D_{s1}^+ and D_{s2}^{*+} candidates in the top and bottom rows, respectively. In the left panels, the fit to the MB sample is shown, while in the right one, the fit to the HM sample is shown. Given the non-Gaussian shape of the signal, like in the analysis of the D_s^+ meson where the significance is estimated within a 3σ interval around the signal peak, in the case of the resonances the significance is estimated considering the *Half Width Half Maximum* (HWHM) of the Voigtian peak. In particular, the fiducial intervals of invariant mass defined for the computation of the statistical significance and the signal-to-background ratios are $0.515 < \Delta M(D_{s1}^+) < 0.535$ GeV/c and $0.650 < \Delta M(D_{s2}^{*+}) < 0.750$ GeV/c. The obtained statistical significance is 3.0 (6.6) for the D_{s1}^+ and 6.8 (5.2) for the D_{s2}^{*+} in the MB (HM) sample.

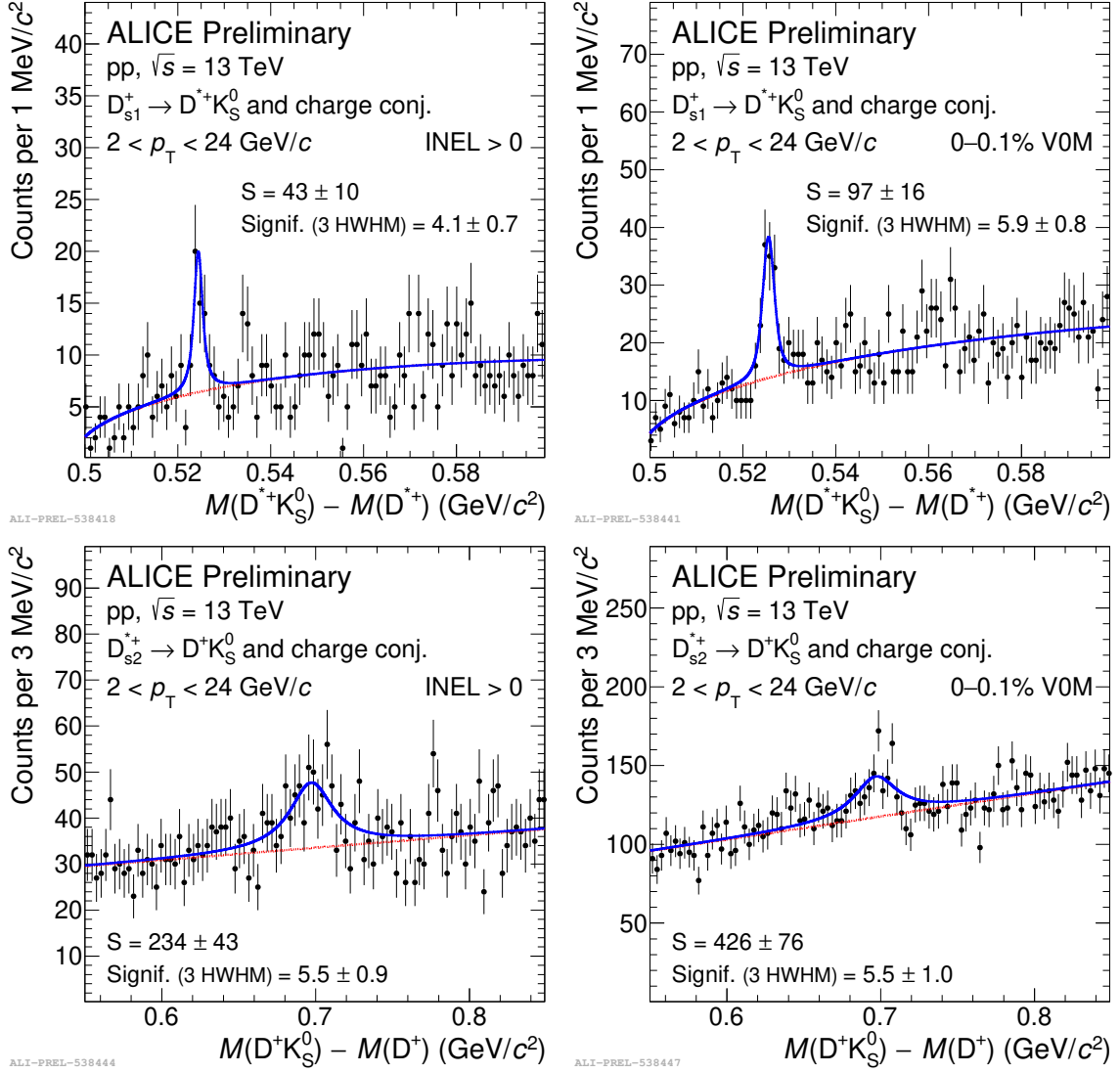


Fig. 5.30 Invariant-mass distributions of $D^{*+}K_S^0$ (top) and $D^+K_S^0$ (bottom) candidates for the MB (left) and HM (right) samples. The red solid line represents the fit to the data. The blue dashed line represents the background contribution.

5.5.3 Efficiency and acceptance correction

The D_{s1}^+ and D_{s2}^{*+} states are not included in the simulation based on PYTHIA 8. In particular, in the native PYTHIA simulation only the non-prompt component of this excited state is included, while the prompt one is suppressed. There are however specific PYTHIA tunes that can be used to include excited meson states and they are currently under study for the purposes of these analyses. Therefore, the efficiency-times-acceptance correction factor

was computed with a different procedure compared to that described in Sec. 5.2.2 for the D_s^+ mesons. Several alternatives were considered.

A possible solution would be to inject $c\bar{c}$ pairs in the simulation on top of the PYTHIA 8 generated event, similarly to what is done for the Pb–Pb collisions (refer to Chapter 4), and force the charm quark to hadronise in the D_{s1}^+ or D_{s2}^{*+} state. However, this approach would require a dedicated simulation campaign, which would be extremely expensive in terms of computing resources, and, more importantly, the remaining anti-charm quark would hadronise into a charm hadron whose decay products would inevitably influence the reconstruction of the primary vertex position. This would introduce a bias in the event topology and, ultimately, in the reconstruction of the excited D-meson states.

Another solution would be to replace the generated prompt D^+ and D^{*+} mesons in the PYTHIA 8 simulation with the D_{s1}^+ and D_{s2}^{*+} mesons, respectively. This replacement should be done at the generation level before the particles are transported through the apparatus with GEANT 3. Technically, this can be done in the AliRoot code after the PYTHIA event is generated and the D mesons are made decay. Therefore, in addition to the replacement of D^+ and D^{*+} with the excited charm-strange meson states, one needs also to make the D_{s1}^+ and D_{s2}^{*+} decay and to remove the decay products of the substituted D-mesons in the generated event. In this way, the bias described above would not be present and the kinematic and the relative abundances of the generated charm hadrons would be reliable. Even though this solution is currently under development, a third procedure, described below, was adopted for the results presented in this Thesis.

The $(\text{Acc} \times \epsilon)$ term of D_{s1}^+ and D_{s2}^{*+} was estimated by propagating the $\text{Acc} \times \epsilon$ factors of the hadronic decay products (i.e. D^+ and K_S^0 , and D^{*+} and K_S^0 , respectively) to the mother particle exploiting the kinematics of the decay predicted by the simulation. This procedure is made possible considering that no selection criteria are directly applied to the D_{s1}^+ or D_{s2}^{*+} states, and therefore their reconstruction efficiency is determined by the product of the efficiencies for reconstructing their two decay products. The acceptance-times-efficiency factors of prompt and non-prompt D^+ and D^{*+} , and K_S^0 mesons were obtained with the PYTHIA 8 event generator with the Monash-13 tune [244] following the same procedure presented above for the D_s^+ meson. The $\text{Acc.} \times \epsilon$ term for D^{*+} (D^+) and K_S^0 are computed as the ratio between the number of candidates passing the analysis selection criteria and the generated ones as a function of p_T , y , and ϕ . The resulting efficiencies can be organised in the form of three-dimensional efficiency maps, one for the prompt D-meson, one for the non-prompt D meson, and one for the K_S^0 . The right panel of Fig. 5.31 shows an example of efficiency map obtained for prompt D^+ in the MB sample.

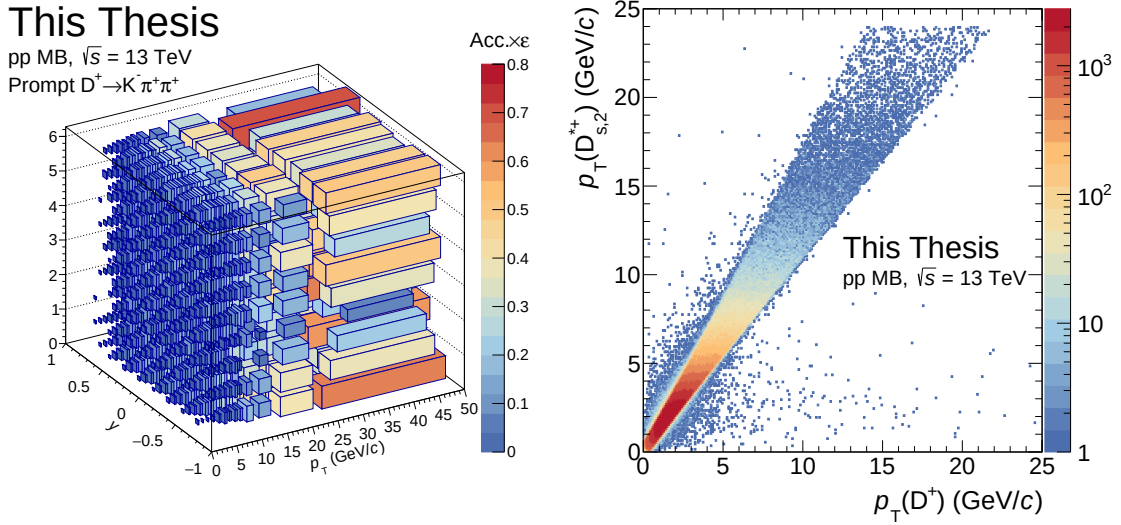


Fig. 5.31 Left: efficiency map of prompt D^+ mesons in MB pp collisions at $\sqrt{s} = 13$ TeV. Right: correlation between the p_T of the D^+ and D_{s2}^{*+} mesons.

The $\text{Acc} \times \epsilon$ of the D_s^+ excited states is computed via an *MC-sampling* method [328]. For this purpose, a fast MC simulation using the PYTHIA 8 decayer was developed. The D_{s1}^+ and D_{s2}^{*+} resonances were generated using the measured p_T distribution of D_s^+ mesons, and they were made decay with the PYTHIA 8 decayer in the desired decay channels. An example of the decay kinematics of the D_{s2}^{*+} into D^+ mesons is shown in the left panel of Fig. 5.31.

The $(\text{Acc} \times \epsilon)$ terms of the D^+ or D^{*+} meson daughter and of the K_S^0 were retrieved from their efficiency map according to their p_T , y , and ϕ . Hence, each generated D_{s1}^+ or D_{s2}^{*+} was accepted or rejected by extracting a random number in the range $[0, 1]$ for each daughter and by comparing it with the corresponding reconstruction and selection efficiency. The $(\text{Acc} \times \epsilon)$ for each p_T interval was computed as the ratio of the number of accepted resonances over the total number of generated resonances in each p_T interval. The uncertainty associated with the efficiency term obtained with this method strongly depends on the number of generated resonances.

To validate the approach based on the daughter-particle efficiency maps and ensure that the calculated efficiency for D_{s1}^+ and D_{s2}^{*+} is reliable, the same procedure was applied to the calculation of the D^{*+} meson efficiencies, which is a resonance like the D_{s1}^+ and D_{s2}^{*+} states and decays into a D^0 and a π^+ . The $(\text{Acc} \times \epsilon)$ term obtained using the D^0 and π^+ efficiency maps as a function of p_T , y , and ϕ was compared to the one obtained with the standard procedure adopted in the D_s^+ analysis based on full simulations of the D^{*+} meson

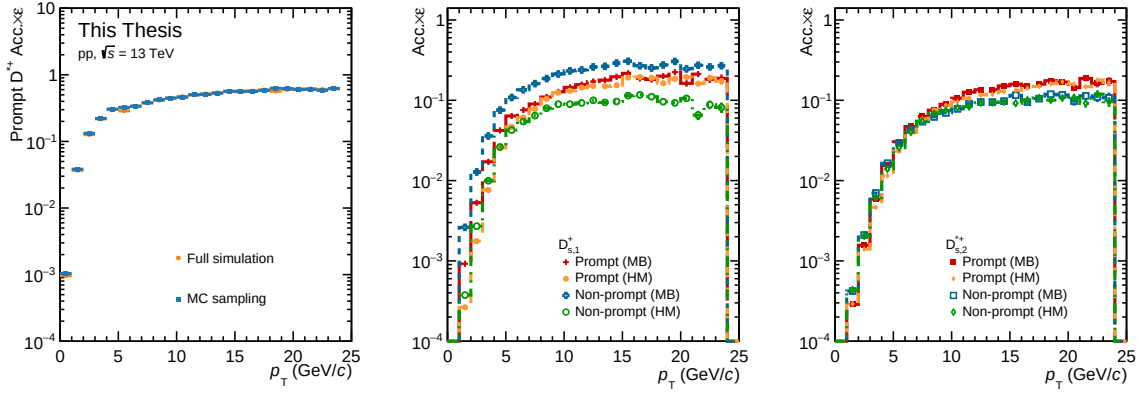


Fig. 5.32 Left: comparison between the $(\text{Acc} \times \epsilon)$ factor of prompt D^{*+} mesons computed with the standard procedure adopted in the D_s^+ analysis and the MC-sampling method. Central: comparison between the $(\text{Acc} \times \epsilon)$ factor of prompt and non-prompt D_{s1}^+ mesons in MB and HM pp collisions at $\sqrt{s} = 13$ TeV. Right: comparison between the $(\text{Acc} \times \epsilon)$ factor of prompt and non-prompt D_{s2}^{*+} mesons in MB and HM pp collisions at $\sqrt{s} = 13$ TeV.

decay. The results of this check are reported in the left panel of Fig. 5.32. The efficiencies obtained with the two methods are fully compatible in the p_T range of the analyses, though some small fluctuations are present, especially in the low- p_T region where the selection on the soft pion is not well reproduced in the fast simulation. The right panel of Fig. 5.32 shows the $\text{Acc} \times \epsilon$ factor as a function of p_T in MB and HM samples for prompt and non-prompt D_{s1}^+ and D_{s2}^{*+} , respectively. It is possible to notice that the efficiencies measured in the MB sample are systematically higher than the ones from the HM sample. This is due to the fact that the D_{s1}^+ and D_{s2}^{*+} resonances are more likely to be produced in the HM sample, where the average p_T of the charm quarks is higher. This is attributed to the application of more stringent selections in the HM sample compared to the MB one. Indeed, this choice is motivated by the presence of a larger combinatorial background in the HM sample. In addition, the D_{s1}^+ efficiency term is found to be systematically higher than the one of the D_{s2}^{*+} meson. This is due to the fact that the selection applied to select the D-meson daughter of the D_{s1}^+ , the D^{*+} , are less severe than the ones applied to the D-meson daughter of the D_{s2}^{*+} , the D^+ . In particular, in the case of the D^{*+} a threshold is applied only to the BDT output score of the background class, while in the case of the D^+ a selection is also applied to the output score related to the probability to be prompt. This also explains the fact that, for the D_{s1}^+ , the efficiency of the non-prompt component is much larger than the one observed for the D_{s2}^{*+} in both multiplicity intervals.

To obtain the value of $\text{Acc} \times \epsilon$ used to correct the measured raw yields, the $\text{Acc} \times \epsilon$ were computed in the $2 < p_T < 24$ GeV/c interval.

The reconstruction and selection efficiencies of the heavy-flavour hadrons, especially the ones with a secondary vertex reconstructed, depend on the multiplicity of charged particles produced in the collisions, since the primary vertex resolution improves with increasing multiplicity [107]. As a consequence, the resolution of the topological variables that make use of the primary vertex position (track impact parameter, particle decay length, and pointing angle) improves as well with increasing multiplicity, introducing a multiplicity dependence in the selection efficiency. It is therefore important that the multiplicity distributions in the MC samples reproduce the multiplicity distribution in data. Given that the multiplicity in the central barrel is not properly reproduced in the simulations, the distributions of the number of tracklets (N_{trkl} , refer to Sec. 2.5) in the MC simulations were weighted to match the distributions from data. The weights were obtained by dividing the N_{trkl} distribution obtained in data for the selected events by the one obtained in the MC simulation. Three different multiplicity weights are available, namely those obtained by considering: all the events, the events containing at least a D-meson candidate, and the events containing at least a D-meson candidate in the invariant-mass region of the signal. The third one was used as default, while the others were considered to estimate the systematic uncertainties as described in Sec. 5.5.5.

5.5.4 Non-prompt D-meson fraction estimation

The measurement of the production of D_{s1}^+ and D_{s2}^{*+} resonances requires the estimation of the fraction of promptly-produced excited states in order to subtract the contribution of beauty-hadron decays. Given the limited number of signal candidates, the application of the data-driven method described in Sec. 5.2.3 can not be directly applied to the D_{s1}^+ and D_{s2}^{*+} resonances. It can be applied to the D mesons produced in the decay of the excited states. However, an additional correction f_{corr} is needed to account for the fact that the D_{s1}^+ and D_{s2}^{*+} resonances contain a strange quark. Indeed, as shown in Fig. 5.25, the non-prompt strange-to-nonstrange D-meson ratio is larger than the prompt one, indicating that the non-prompt D mesons are more likely to contain a strange quark compared to the prompt ones. In particular, the correction factor was computed as the double p_T -integrated ratio between the D_s^+ over D production cross sections:

$$f_{\text{corr}} = \frac{D_s^+}{D^0 + D^+} \Big|_{\text{non-prompt}} / \frac{D_s^+}{D^0 + D^+} \Big|_{\text{prompt}} = 1.97 \pm 0.53$$

Finally, to obtain the D_{s1}^+ and D_{s2}^{*+} prompt fractions from the one of the D^{*+} and D^+ daughters, the same propagation strategy based on the decay kinematics used for the

Table 5.12 Systematic uncertainties for the measurement of prompt D_{s1}^+ and D_{s2}^{*+} corrected yield.

	D_{s1}^+		D_{s2}^{*+}	
	MB	HM	MB	HM
Multiplicity interval				
Raw-yield extraction	5%	6%	7%	14%
Selection efficiency	10%	10%	10%	10%
Non-prompt fraction	7%	4%	4%	4%
Tracking efficiency	6%	6%	6.5%	6.5%
MC p_T shape	10%	12%	13%	13%
Multiplicity MC distribution	1%	2%	1%	negl.
$BR(D_{\text{daug.}}) \times BR(K_S^0)$	1.1%		1.7%	

efficiency is applied, via the following relation:

$$\begin{aligned}
 f_{\text{non-prompt}}(D_{sJ}) &= \frac{N_{\text{non-prompt}}}{N_{\text{prompt}} + N_{\text{non-prompt}}} \bigg|_{D_{sJ}} (p_T(D_{sJ})) \\
 &= \frac{f_{\text{corr}} \times N_{\text{non-prompt}}}{N_{\text{prompt}} + f_{\text{corr}} \times N_{\text{non-prompt}}} \bigg|_{D_{\text{daug.}}} (p_T(D_{\text{daug.}}))
 \end{aligned} \tag{5.20}$$

The prompt fraction can be then obtained as $f_{\text{prompt}}(D_{sJ}) = 1 - f_{\text{non-prompt}}(D_{sJ})$. The p_T dependence of the non-prompt fraction is similar for both the resonance and the D meson since most of the p_T of the resonance is inherited by the D meson daughter. No noticeable p_T dependence is observed.

5.5.5 Systematic uncertainties

The full list of systematic uncertainty sources considered in the measurement of the D_{s1}^+ and D_{s2}^{*+} mesons is reported in Tab. 5.12. The systematic uncertainties were estimated following the procedures described below.

Raw yield extraction

The systematic uncertainty associated with the raw-yield extraction of the resonances was determined following the same procedure adopted for the D_s^+ meson and described in Sec. 5.3.1. The fit procedure was repeated varying the fit configurations. In particular, the Γ parameter of the resonances in the Voigtian function was varied considering the uncertainty from Ref. [6]. It is important to note that the Gaussian σ was not varied from the default configuration since it is significantly smaller than the Γ parameter. The signal fit function is the appropriate function to describe the resonance states, hence it was not varied. The same quality criteria as the ones described in Sec. 5.3.1 were applied to exclude non-reliable fit results. The systematic uncertainty was assigned considering the RMS of the raw-yield distribution for all the trials and the shift of its mean with respect to the default value. The bin counting method described in Sec. 5.3.1 for the D_s^+ meson was not employed. The assigned systematic uncertainties are of 14% to 18% for the D_{s2}^{*+} and 5% to 6% for the D_{s1}^+ in the MB and HM samples, respectively.

Given the unexpectedly large systematic uncertainty observed for the D_{s2}^{*+} resonance compared to the D_{s1}^+ and the D_s^+ mesons, an additional check was conducted to evaluate the possible effect of the limited amount of data on the systematic estimation. Considering the fit functions from the default configuration as the "truth", i.e. the configuration which describes signal and background distributions, several invariant-mass distributions were produced sampling the fit functions and increasing the number of entries in the invariant-mass distribution using a multiplication factor ranging between 1 – 100. Then, for every invariant-mass distribution generated in the previous step, the estimation of the systematic uncertainties was repeated. This procedure was repeated iteratively varying the multiplication factor. For every multiplication factor, the ratio between the RMS and the mean of the resulting yield distribution is computed. To perform this check, the quality criteria on the fit results were relaxed to prevent the rejection of a significant portion of the trial configurations, particularly when dealing with large multiplication factors. If the original systematic uncertainty is actually inflated by statistical fluctuations, then the average of the RMS/mean ratio distribution would strongly depend on the number of generated events N_{ev} according to $f(N_{ev}) = p_0 + \frac{p_1}{\sqrt{N_{ev}}}$. In the limit where N_{ev} tends to infinity, the remaining contribution is purely systematic (p_0). Following this procedure, the resulting systematic was estimated to be 7% and 14% for the MB and HM samples for the D_{s2}^{*+} resonance, respectively. The results for the RMS/mean ratio and the fit with $p_0 + \frac{p_1}{\sqrt{N_{ev}}}$ from the HM sample are reported in the left panel of Fig. 5.33. Note that the resulting values of systematics

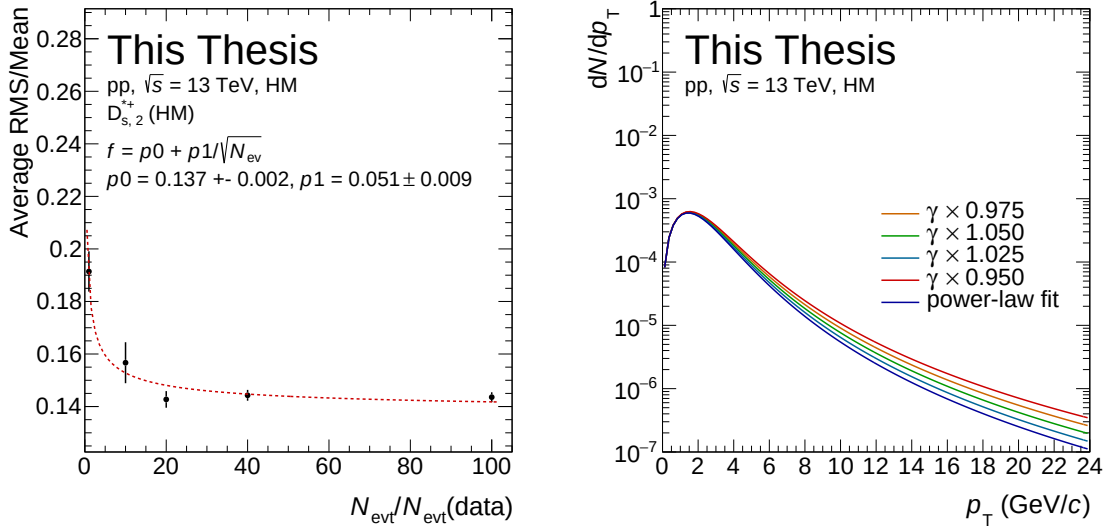


Fig. 5.33 Left: RMS/mean ratio of the systematic uncertainty as a function of the multiplication factor of the data sample for the D_{s2}^{*+} resonance in the HM sample. Right: alternative p_T -shapes for systematic uncertainty estimation.

are significantly smaller than the values of 14% and 18% that were initially estimated from the systematic studies.

Selection efficiency

As discussed above, a systematic uncertainty is assigned to account for possible imperfections in the description of the variables used in the training of the ML models in the simulations. This source of systematic uncertainty was studied by repeating the analysis procedure and varying the applied selection criteria based on the ML output scores, for the D mesons, and the V0 topology, for the K_S^0 . In particular, a scan of the ML outputs for the probability to be combinatorial background and to be a prompt D meson was performed by defining 5 looser and 5 tighter selections for each of these two scores. Additionally, 5 looser and 5 tighter selections on the $\cos(\theta_P)$, and decay radius of the K_S^0 , and on the DCA of the K_S^0 daughters were tested. To exclude problematic cases or too extreme variations, only the sets of selections with a statistical significance of at least 3, a variation of the efficiencies between 0.5 and 2.5 times the default values were taken into account in the evaluation of the uncertainty. The magnitude of the systematic uncertainty was assigned considering the RMS and the shift of the corrected yields compared to the reference case. A 10% of systematic uncertainty was estimated for both the D_{s1}^{*+} and D_{s2}^{*+} mesons in the MB and HM events.

Generated p_T distributions

The systematic uncertainty due to the generated p_T -shape of the resonances in the MC simulations was evaluated by repeating the efficiency estimation varying the p_T -shape. In particular, for the reference case the p_T shape was defined from a power-law fit performed to the p_T distribution of the D_s^+ meson p_T -differential cross section. To test softer and harder p_T shapes, the γ parameter of the power-law fit ($f(p_T) = Np_T/(1 + (p_T/\alpha)^\beta)^\gamma$) was varied symmetrically around the reference value. An example of the variation of the p_T shape is shown in the right panel of Fig. 5.33. The p_T shape also has a small impact on the estimated prompt fraction. Hence, the systematic uncertainty was evaluated directly on the corrected yield, which includes both the efficiency and the prompt fraction.

MC multiplicity distribution

The D^+ and D^{*+} reconstruction and selection efficiency depends on the charged-particle multiplicity of the event, which affects the primary-vertex resolution. Since the efficiencies are included in the prompt fraction evaluation, also this fraction could depend on the multiplicity distribution, even though its dependence is found to be negligible. Thus, to estimate a systematic uncertainty related to the multiplicity distribution in the simulation the corrected yields were recomputed by varying the weights used to match the multiplicity distribution in the simulations to the measured one. As anticipated, three different multiplicity weights were considered: the ones including all the events, those obtained considering events containing at least a D-meson candidate, and those including events containing at least a D-meson candidate in the invariant-mass region of the signal. The third weight was used as default, while the first two were adopted for the systematic uncertainty evaluation. The assigned systematic uncertainties are obtained by considering the deviation between the results obtained with different weights, the effect was found to be small, corresponding to 1% – 2% for the D_{s1}^+ and D_{s2}^{*+} mesons in the MB and HM events.

Track-reconstruction efficiency

The systematic uncertainty on the track-reconstruction efficiency accounts for possible discrepancies between data and MC in the TPC-ITS tracks prolongation efficiency, the description of the ALICE detector material budget, and the selection efficiency due to track-quality criteria in the TPC. The systematic uncertainties for D_{s1}^+ and D_{s2}^{*+} mesons are

obtained by propagating those of the decay daughters using the decay kinematics, similarly to what was done in the case of the D_s^+ meson measurements.

The systematic-uncertainty values for the decay daughters are inherited from previous analyses in pp collisions at $\sqrt{s} = 13$ TeV. In the case of K_S^0 mesons, they account for the track-quality criteria in the TPC and the material-budget description, while TPC-ITS prolongation efficiency is not considered because no ITS points are required in the K_S^0 daughter selection. The values for D^+ and D^{*+} account for the discrepancies in the ITS-TPC prolongation efficiency and the track-quality criteria in the TPC. The contribution due to the material-budget description is considered negligible in this case. The resulting uncertainties are 6% and 6.5% for the D_{s1}^+ and D_{s2}^{*+} mesons in the MB and HM events, respectively.

5.5.6 Branching ratio

The excited D_s^+ -meson states were reconstructed in the $D_{s1}^+ \rightarrow D^{*+} K_S^0$ and $D_{s2}^{*+} \rightarrow D^+ K_S^0$ hadronic decay channels for which branching ratios are not measured. Therefore, as anticipated in Chapter 4, the branching ratios of these channels were not used to correct the measured yields. The measurement is corrected for the BRs of $K_S^0 \rightarrow \pi^+ \pi^-$ and either $D^{*+} \rightarrow D^0(\rightarrow K^- \pi^+) \pi^+$ or $D^+ \rightarrow K^- \pi^+ \pi^+$, depending on the resonance. Thus, the relative uncertainties on these branching ratios were taken from Ref. [6] and propagated, resulting in an uncertainty of 1.1% and 1.7% for D_{s1}^+ and D_{s2}^{*+} , respectively.

5.5.7 Results

Figure 5.34 reports the first measurement of prompt D_{s1}^+ and D_{s2}^{*+} yields in the $2 < p_T < 24$ GeV/c interval as a function of charged particle multiplicity ($\langle dN_{ch}/d\eta \rangle$) in pp collisions at $\sqrt{s} = 13$ TeV at LHC. The resonance yields are compared to the yields of the D_s^+ meson in the $2 < p_T < 24$ GeV/c interval measured by ALICE and reported in Refs. [270, 97, 234]. The statistical and systematic uncertainties are depicted as vertical bars and empty boxes, respectively. It is important to notice that, given the lack of measurements of the BR of the D_{s1}^+ and D_{s2}^{*+} resonances, the yields of the resonances are not corrected for the BR of the decay channel under study. However, a high branching ratio is expected for the D_{s1}^+ and D_{s2}^{*+} resonances, as discussed in the following, compared to the one of the D_s^+ meson. The production of the ground state exceeds the one of the excited states, and the yield increases with increasing multiplicity of particles production in the collision.

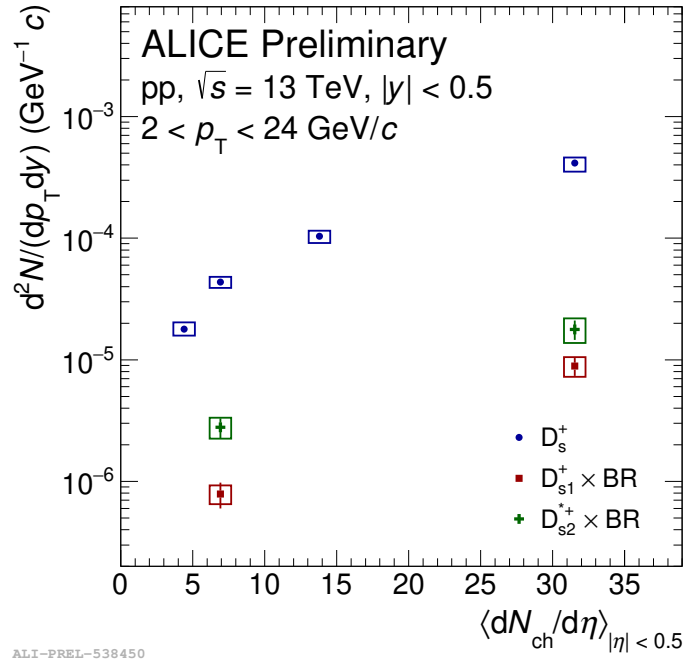


Fig. 5.34 p_T -integrated yields (in $2 < p_T < 24$ GeV/c) times the BR of the analysed decay channel of prompt D_{s1}^+ and D_{s2}^{*+} mesons compared with the p_T -integrated yields of prompt D_s^+ mesons from Ref. [270].

In addition to the computation of prompt D_{s1}^+ and D_{s2}^{*+} meson production yields, the results have been scaled by the p_T -integrated yields of the D_s^+ meson. This approach cancels out the dependence on strangeness and charm corrections in theoretical calculations, making the ratios more convenient for comparisons to theory. The yield ratios are shown as a function of the charged-particle multiplicity in Fig. 5.35. The ratios are compared with the predictions based on the SHM developed by He and Rapp [11] and SHM for charm hadrons (SHMc) of the GSI-Heidelberg group [329], scaled by the branching ratios of the resonance decay channels. The theoretical predictions were performed for $p_T > 0$. As anticipated, there are no available measurements of the BR of the resonance decay channels. Hence, the values of the BR and its uncertainty were assigned in the theory predictions considering the results presented in Ref. [330] where the BRs of these decay channel were calculated in the *Relativised Quark Model* (RQM) framework [331]. The RQM model, which has been found to successfully describe most known mesons and has proven to be a useful guide to understanding newly found states [332, 224, 222, 333], predicts a BR of the $D_{s1}^+ \rightarrow D^* K$ of 95.3%. However, the RQM model does not distinguish between the two possible charge states of the D^* meson. Therefore, given the ratio between the two charge states reported in Ref. [6], $\text{BR}(D_{s1}^+ \rightarrow D^{*+} K^0)/\text{BR}(D_{s1}^+ \rightarrow D^{*0} K^+) = 0.85 \pm 0.12$ and assuming that 50%

of the K^0 are K_S^0 , the following value is obtained:

$$\begin{aligned} \text{BR}(D_{s1}^+ \rightarrow D^{*+} K_S^0) &= \frac{1}{2} \times \text{BR}(D_{s1}^+ \rightarrow D^{*+} K^0) \\ &= \frac{95.3\%}{1 + \frac{\text{BR}(D_{s1}^+ \rightarrow D^{*0} K^+)}{\text{BR}(D_{s1}^+ \rightarrow D^{*+} K_S^0)}} = (22 \pm 2)\% . \end{aligned} \quad (5.21)$$

Similarly, the BR of the $D_{s2}^{*+} \rightarrow DK$ decay is expected to be equal to 93.4%. In this case though, no information about the ratio between the two D-meson charge states (i.e. D^0 and D^+) is provided, therefore, assuming them to be equally abundant, the following value is obtained:

$$\text{BR}(D_{s2}^{*+} \rightarrow D^+ K_S^0) = 23.35\% . \quad (5.22)$$

The latest assumption on the charged states is supported by the fact that the D^0 and D^+ masses are very close to each other hence it is likely that the two states appear in basically equal fractions. In the case of the D_{s1}^+ , no significant dependence of the yield ratio D_{s1}^+/D_s^+ ratios is observed on the average charged-particle multiplicity. However, for the D_{s2}^{*+} , a hint of enhancement is observed at low multiplicity compared to high multiplicity, which could be attributed to the different lifetimes of the resonances. The D_{s1}^+ has a larger lifetime compared to the D_{s2}^{*+} ($\tau(D_{s1}^+) \sim 219$ fm [6] and $\tau(D_{s2}^{*+}) \sim 12$ fm [6]). This means that the D_{s1}^+ are more likely to decay after the phase where hadronic scattering effects can affect the resonance reconstruction by altering the momenta of the decay products, unlike the D_{s2}^{*+} . A similar behaviour is observed in the light-flavour sector [47]. In particular, the yield ratios ρ^0/π , K^*/K , and Λ^*/Λ resonances decrease with increasing averaged charged particle multiplicity, while ϕ/K ratio remains approximately constant. This is attributed to the ϕ meson lifetime, one order of magnitude larger than that of the $\bar{K}^{*0}(892)$ resonance, which causes ϕ mesons to decay after the hadronic phase. The theoretical calculations do not predict a multiplicity dependence because they do not include hadronic rescattering effects. They are found to be consistent with the measured ratio within $1 - 2(2 - 3) \sigma^5$ in the D_{s1}^+ (D_{s2}^{*+}) case.

This kind of measurement are a novel at LHC and provides an important tool to test the hadronisation mechanisms of charm quarks and the interactions of charm hadrons within the hadronic phase of the collision. The measurements discussed above represent a headstone for future measurements in pp and Pb–Pb collisions during the next data taking period at the LHC.

⁵The σ here indicates the sum in quadrature of statistical and systematic uncertainties.

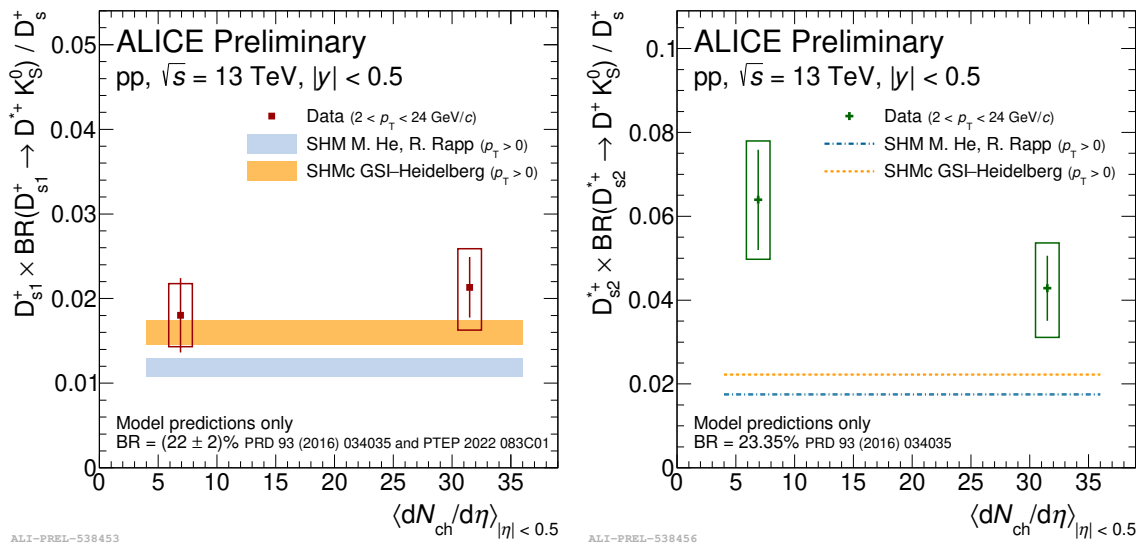


Fig. 5.35 Left: p_T -integrated yield ratios of prompt D_{s1}^+ mesons over prompt D_s^+ mesons as a function of the average charged-particle multiplicity ($\langle dN_{ch}/d\eta \rangle$) in pp collisions at $\sqrt{s} = 13$ TeV. Right: p_T -integrated yield ratios of prompt D_{s2}^{*+} mesons over prompt D_s^+ mesons as a function of the average charged-particle multiplicity ($\langle dN_{ch}/d\eta \rangle$) in pp collisions at $\sqrt{s} = 13$ TeV. The vertical bars and the empty boxes represent the statistical and systematic uncertainties, respectively. The shaded boxes represent the systematic uncertainties on the branching ratios. The predictions based on the SHM [11] and SHMc [329] are shown as solid and dashed lines, respectively.

Chapter 6

Strange D-meson production in Pb–Pb collisions at $\sqrt{s_{\text{NN}}} = 5.02 \text{ TeV}$

At temperatures exceeding the pseudo-critical value of $T \sim 154 - 158 \text{ MeV}$ at vanishing baryon density, the strongly-interacting matter is predicted to undergo a phase transition to a state known as Quark-Gluon Plasma (QGP) characterised by the deconfinement of partons from the hadrons. In ultrarelativistic heavy-ion collisions at the LHC, the thermodynamic conditions necessary for the production of the QGP medium are reached. The lifetime of the QGP medium in the collisions is typically of the order of $10 \text{ fm}/c$, preventing it from being directly detected. The QGP properties can be inferred from the study of the final-state particles produced in the collision which are sensitive to the presence of a deconfined phase within the system evolution. Owing to their large masses, heavy quarks, i.e. charm and beauty quarks, are predominantly produced in the first stages of the collision in hard parton scattering processes. These processes occur on shorter timescales compared to the QGP formation ($0.5 - 1 \text{ fm}/c$). In addition, due to their large masses the thermal production of heavy quarks in the QGP medium, as well as their production via string-breaking mechanisms are negligible. Hence, heavy quarks are expected to maintain their *identity* throughout the entire evolution of the medium created in the collision. This makes heavy quarks and, subsequently, heavy-flavour hadrons exceptionally sensitive probes to investigate the characteristics of the QGP medium.

Traversing the expanding, hot, and dense QGP medium, heavy quarks lose energy interacting with the medium constituents. The interactions encompass both elastic and inelastic (mainly gluon radiation) processes. For energetic quarks, these interactions result in in-medium energy loss. For lower momentum heavy quarks, the energy-momentum

exchanges with the medium constituents may lead to the thermalisation of heavy quarks in the QGP. Hence, the interactions with the medium modify the energy spectrum of the observed heavy hadrons. In addition, the presence of the QGP medium modifies the hadronisation mechanism of heavy quarks. In particular, heavy quarks are expected to hadronise via two competing mechanisms in heavy-ion collisions. The first mechanism is the fragmentation, dominant in the $p_T > 6 - 8 \text{ GeV}/c$ interval, and the second one is the recombination (or coalescence), expected to be more relevant in the low- p_T region. The former consists of high-momentum heavy quarks fragmenting into some lower-momentum hadrons in a vacuum-like procedure, while in the recombination process, the hadrons are formed by the combination, or coalescence, of heavy quarks with other quarks from the QGP medium sharing the same phase-space coordinates (i.e. position and velocity) in the medium. A detailed description of the hadronisation mechanisms is provided in Sec. 1.7.

The combined effect of the modification of the hadronisation mechanism and the abundance of $s\bar{s}$ quark pairs thermally produced in heavy-ion collisions compared to pp collisions (*strangeness enhancement*), leads to an expected enhanced production of charm- and beauty-strange hadrons compared to non-strange ones. Measurements of the p_T -differential yields of prompt and non-prompt D_s^+ mesons in Pb–Pb collisions can provide insight into the hadronisation mechanism of charm and beauty quarks in the QGP medium, respectively.

The corrected p_T -differential yields of non-prompt D_s^+ mesons were computed as:

$$\frac{dN(D)_{\text{non-prompt}}}{dp_T} = \frac{1}{2} \frac{f_{\text{non-prompt}}^{\text{raw}}(p_T) \times N^{D+\bar{D},\text{raw}}(p_T)}{c_{\Delta y}(p_T) \times (\text{Acc} \times \varepsilon)(p_T) \times \text{BR} \times N_{\text{ev}}}, \quad (6.1)$$

in analogy with the analyses performed on the pp sample. The measurements described in this Chapter were performed in central (0 – 10%) and semicentral (30 – 50%) Pb–Pb collisions at a centre-of-mass energy per nucleon pair $\sqrt{s_{\text{NN}}} = 5.02 \text{ TeV}$. In particular, the energy loss and the modification of the hadronisation mechanisms can be studied by computing the nuclear modification factor, R_{AA} . The R_{AA} is defined as:

$$R_{\text{AA}} = \frac{1}{\langle T_{\text{AA}} \rangle} \frac{d^2 N^{\text{AA}} / dy dp_T}{d^2 \sigma^{\text{pp}} / dy dp_T}, \quad (6.2)$$

where $d\sigma_{\text{pp}}/dp_T$ is the p_T -differential cross section in pp collisions, and $\langle T_{\text{AA}} \rangle$ is the average of the nuclear overlap function taken from Ref. [334]. In the specific case of heavy quarks and in absence of nuclear or QGP effects, the R_{AA} is expected to be equal to unity independently on the p_T interval. Energy loss or modification of the hadronisation mechanism are expected

to affect the yield and p_T distribution of heavy-flavour hadrons in Pb–Pb collisions compared to pp collisions, and consequently to lead to a deviation of the R_{AA} from unity.

In this chapter the first measurement of the R_{AA} of non-prompt D_s^+ mesons in central (0 – 10%) and semicentral (30 – 50%) Pb–Pb collisions at $\sqrt{s_{\text{NN}}} = 5.02$ TeV is presented. These results were published in Ref. [335]. The non-prompt D_s^+ mesons were selected via machine-learning-based selection criteria similar to what was done in the pp analyses, as described in Sec. 6.1. The D_s^+ mesons yields were extracted following the procedure described in Sec. 6.2.1. The estimation of the efficiency and non-prompt fraction correction is described in Secs. 6.2.2 and 6.2.3, respectively. The pp reference, necessary for the estimation of the R_{AA} , is described in Sec. 6.2.4, while the estimation of the systematic uncertainties and the description of the final results are reported in Sec. 6.3 and Sec. 6.4, respectively.

6.1 Machine-learning for the D_s^+ meson signal extraction

Following the procedure presented in Chapter 5, a machine-learning approach was employed to select the non-prompt D_s^+ meson candidates in Pb–Pb collisions. Given, the larger sample of D_s^+ meson candidates collected in Pb–Pb compared to pp collisions, the inference step was performed on the ALICE Grid computing infrastructure presented in Chapter 2, while the training and test steps were performed locally on a subsample of the dataset. To enhance the contribution of non-prompt D_s^+ mesons while rejecting both prompt D_s^+ mesons and the background, a multiclass classification algorithm was employed.

6.1.1 Training, testing, and application sets

The training and test sets include both the signal, made of prompt and non-prompt D_s^+ examples, and background instances. The signal examples were obtained from MC simulations. The underlying Pb–Pb event was simulated with the HIJING 1.36 event generator [173] and, in each HIJING event, multiple $c\bar{c}$ - and $b\bar{b}$ -quark pairs generated with PYTHIA 8 were injected to produce D_s^+ -meson signals, as described in detail in Sec. 4.3. The background instances instead were extracted from the sidebands of the D_s^+ -meson invariant-mass distribution obtained from data. This strategy was adopted to limit the effect of potential discrepancies between data and MC on the performance of the machine-learning algorithm and allows us to mitigate the effect of possible shortcomings in the description of the background properties by the HIJING simulation.

To reject a large fraction of background candidates while preserving the physical signal, the D_s^+ -meson candidates were required to fulfil the selection criteria reported in Table 6.1 before being fed to the BDT. These selections are applied to the kinematics of the decay particles, on the displaced vertex topology, and the compatibility of the K^+K^- invariant mass with the ϕ meson mass. In addition, the PID selections described in Sec. 4.2.3 were applied at the track level. The training of the BDT models was performed separately in the two centrality intervals in order to avoid possible biases due to the different characteristics of signal and background examples in central and semicentral collisions. Like in the pp analysis, the training was performed in wider p_T intervals with respect to the ones adopted in the measurements. This choice is twofold. On one hand, employing wider p_T intervals in the training phase allows us to increase the number of examples in the regions where the amount of candidates is limited, like at high p_T . On the other hand, it reduces the number of BDT algorithms to be applied to the Grid during the inference step leading to a faster data processing. The p_T intervals employed in the training of the BDT algorithms are reported in Table 6.2 along with the number of signal and background instances employed in the training of the BDT.

Table 6.1 Topological and kinematic pre-selections applied for the D_s^+ analysis in 0–10% centrality class.

Centrality p_T (GeV/c)	0 – 10%		30 – 50%	
	[2,6]	[6,50]	[2,6]	[6,50]
$p_T(\pi, K)$ (MeV/c)	> 0.6	> 0.6	> 0.4	> 0.4
decay length (μm)	> 300	> 200	> 200	> 200
norm. dec. length xy	> 4	> 2	> 4	> 2
$\cos \theta_P$	> 0.97	> 0.90	> 0.96	> 0.90
$\cos \theta_{Pxy}$	> 0.96	> 0.90	> 0.96	> 0.90
σ_{vertex} (cm)	< 0.04	< 0.06	< 0.045	< 0.06
$M_{\text{inv}}^\phi - M_{\text{PDG}}^\phi$ (MeV/c ²)	< 10	< 15	< 15	< 15

For the signal, all the instances in the training set were employed. The number of background examples in the low- p_T region was chosen to be similar to the number of signal instances. At higher p_T , where the overall number of candidates in data is low, fewer background candidates are used to maintain small the fraction of real data employed in the training with respect to the total amount of background candidates available.

The features provided to the BDT during the training are the same topological, kinematic, and PID variables presented in Chapter 4 and employed in the analysis of pp collisions. In the training phase, the hyperparameters of the BDT are optimised with a Bayesian approach. However, to limit the CPU-time consumption at the moment of the application to the Grid,

Table 6.2 Number of signal (prompt and non-prompt D_s^+) and background candidates used to build the training and test sets for the 0 – 10% and 30 – 50% centrality classes.

p_T interval (GeV/c)	prompt D_s^+	non-prompt D_s^+	combinatorial background
0 – 10%			
2 – 4	$\sim 88 \times 10^3$	$\sim 206 \times 10^3$	$\sim 294 \times 10^3$
4 – 6	$\sim 210 \times 10^3$	$\sim 370 \times 10^3$	$\sim 174 \times 10^3$
6 – 12	$\sim 356 \times 10^3$	$\sim 456 \times 10^3$	$\sim 203 \times 10^3$
12 – 50	$\sim 468 \times 10^3$	$\sim 371 \times 10^3$	$\sim 210 \times 10^3$
30 – 50%			
2 – 4	$\sim 336 \times 10^3$	$\sim 535 \times 10^3$	$\sim 435 \times 10^3$
4 – 6	$\sim 265 \times 10^3$	$\sim 386 \times 10^3$	$\sim 196 \times 10^3$
6 – 12	$\sim 365 \times 10^3$	$\sim 453 \times 10^3$	$\sim 654 \times 10^3$
12 – 50	$\sim 406 \times 10^3$	$\sim 332 \times 10^3$	$\sim 150 \times 10^3$

the phase-space considered in the hyperparameter optimisation is limited compared to the one considered for the measurements in pp collisions. This is particularly relevant for the number of estimators composing the final BDT and the maximum depth of the trees which have the largest impact in terms of CPU time. In particular, they were limited to 1200 and 4, respectively. The complete list of the optimised hyperparameters is reported in Table 6.3. The performance of the optimised BDT algorithms is assessed by computing the ROC AUC (refer to Chapter 3). The ROC curves obtained from the test set for the models employed in central Pb–Pb in the $4 < p_T < 6$ GeV/c are reported in Fig. 6.1. The ROC AUC values are found to be close to unity indicating a good performance of the BDT algorithms. Similar results are found in the other p_T intervals and in semicentral collisions.

6.1.2 BDT-based selection optimisation

The procedure utilised to optimise the BDT-based selections is the one described in detail in Chapter 5 and consists of performing a scan of the BDT-score distribution in the p_T intervals of the analysis. In the scan, the threshold value on two of the BDT-scores, namely the ones related to the probability to be and non-prompt D_s^+ , are iteratively varied. For each set of selections, the corresponding expected signal and background yields were computed. In the

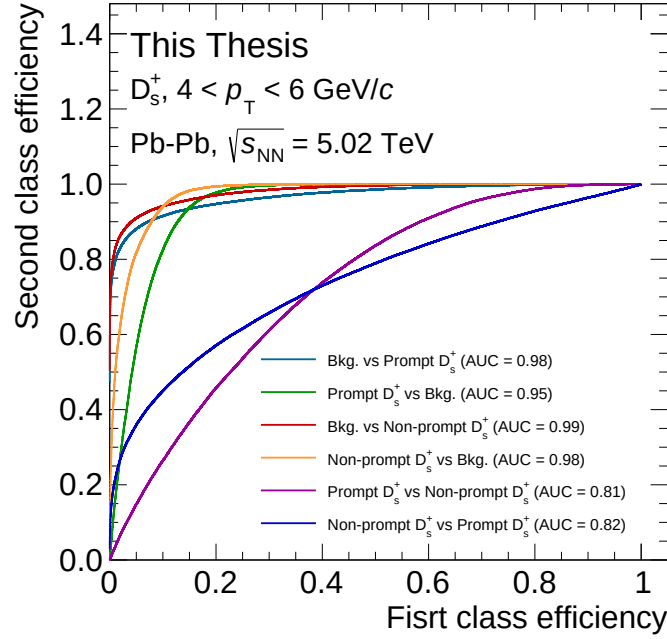


Fig. 6.1 ROC curves obtained from the test set for the models employed in central Pb–Pb collisions in the $4 < p_T < 6$ GeV/ c interval.

case of Pb–Pb collisions the expected signal yield was computed as:

$$\begin{aligned}
 S_{\text{exp}} = & 2 \times (d\sigma/dp_T)_{\text{non-prompt}}^{\text{FONLL}} \times R_{\text{AA}}^{\text{TAMU, non-prompt}} \times \langle T_{\text{AA}} \rangle \\
 & \times \frac{(\text{Acc} \times \varepsilon)_{\text{non-prompt}}}{f_{\text{non-prompt}}} \\
 & \times c_{\Delta y} \times \Delta p_T \times \text{BR} \times N_{\text{ev}} ,
 \end{aligned} \tag{6.3}$$

where the $R_{\text{AA}}^{\text{TAMU, non-prompt}}$ is the R_{AA} of non-prompt D_s^+ mesons predicted by the TAMU model [126] and $\langle T_{\text{AA}} \rangle$ is the average nuclear overlap function for the considered centrality class. The expected background yields, B_{exp} , were estimated with a fit to the sideband regions of the invariant-mass distribution of the D_s^+ -meson candidates. The sideband regions are defined in the same way as in the pp analysis.

The optimum BDT-based selection was chosen as the one that maximises the expected significance ($S_{\text{exp}}/\sqrt{S_{\text{exp}} + B_{\text{exp}}}$) while ensuring a sufficiently high non-prompt fraction. A second set of BDT selection criteria was chosen to maximise the contribution of prompt D_s^+ mesons in order to provide a statistically stable reference in the raw-yield extraction. In addition this second set allows us to cross-check the validity of the method by comparing the results obtained with the prompt-enhanced sample with the published results for prompt D_s^+

Table 6.3 Hyperparameter configuration for the p_{T} bins considered in the training for the D_s^+ mesons in the 0 – 10% and 30 – 50% centrality class, respectively.

0 – 10%	p_{T} interval (GeV/c)			
Hyperparameter	2 – 4	4 – 6	6 – 12	12 – 50
n_estimators	952	952	952	953
learning_rate	0.07	0.09	0.09	0.1
max_depth	4	4	4	4
min_child_weight	8.3	6.9	6.9	7.2
subsample	0.85	0.89	0.89	0.96
colsample_bytree	0.98	0.86	0.86	0.96
tree_method	hist	hist	hist	hist
30 – 50%	p_{T} interval (GeV/c)			
Hyperparameter	2 – 4	4 – 6	6 – 12	12 – 50
n_estimators	719	724	798	798
learning_rate	0.09	0.1	0.1	0.09
max_depth	4	3	4	4
min_child_weight	8.3	6.9	6.9	7.2
subsample	0.87	0.81	0.98	0.84
colsample_bytree	0.96	0.98	0.84	0.85
tree_method	hist	hist	hist	hist

mesons [336]. The procedure was repeated in each p_{T} interval of the analysis in both the centrality classes.

Figure 6.2 shows the expected significance (top-left), the expected signal-to-background ratio (top-right) and the efficiency term (bottom left) for non-prompt D_s^+ mesons as a function of the selection criterion the two on the BDT probabilities. Looking at the top-left panel it is possible to observe that by applying stringent selections on the probability to be background, the expected significance quickly increases. However, the opposite trend is observed for the non-prompt D_s^+ meson efficiency. In addition, the application of a minimum threshold value on the probability to be a non-prompt D_s^+ meson allows us to increase the expected signal-to-background ratio at the cost of a lower efficiency. The optimised set of BDT selections ranges in the 0.005 – 0.02 and 0.6 – 0.8 intervals for the probability to be a background and a non-prompt D_s^+ meson, respectively depending on the p_{T} interval and the centrality class.

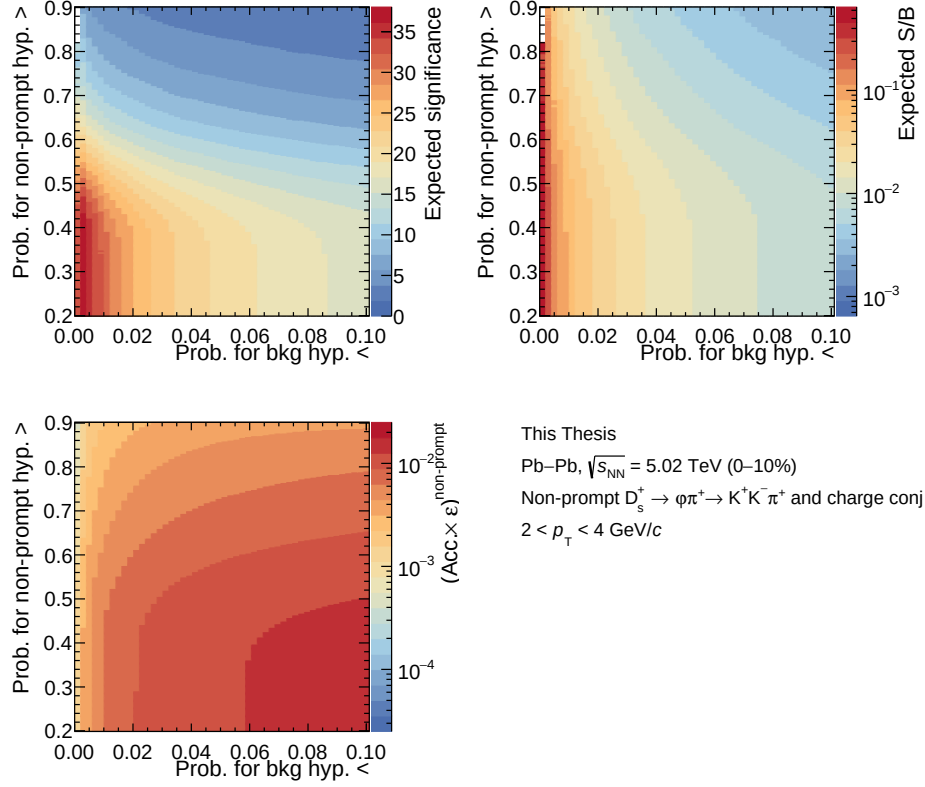


Fig. 6.2 Expected significance (top-left), expected signal-to-background ratio (top-right) and efficiency term (bottom left) for non-prompt D_s^+ mesons obtained in central Pb–Pb collisions in $2 < p_T < 4$ GeV/c interval.

6.2 Non-prompt D_s^+ meson production in Pb–Pb collisions at $\sqrt{s_{NN}} = 5.02$ TeV

6.2.1 Raw-yield extraction

The extraction of the non-prompt D_s^+ meson raw-yields in different p_T intervals was performed by fitting the invariant-mass distributions of the D_s^+ -meson candidates passing the optimised BDT selection criteria. A robust signal extraction, meaning a statistical significance higher than 3, was obtained in several p_T intervals in the ranges $4 < p_T < 36$ GeV/c and $2 < p_T < 24$ GeV/c in central and semicentral collisions, respectively. The different extension of the p_T ranges in the two studied centrality classes is due to the different abundance of signal and background candidates moving from central to peripheral collisions. In particular, as reported in Table 6.4, the signal-to-background ratio is higher in the low- p_T region for semicentral collisions compared to central ones. Despite the more stringent selection criteria

Table 6.4 Raw yields, signal-to-background ratios, and significance of non-prompt D_s^+ mesons in pp collisions at $\sqrt{s} = 13$ TeV.

p_T interval (GeV/c)	Central (0 – 10%)		
	Raw-yield (3σ)	S/B (3σ)	Significance (3σ)
4 – 6	789 ± 66	0.16	10.4 ± 0.8
6 – 8	1726 ± 73	0.12	9.3 ± 0.8
8 – 12	1276 ± 55	0.21	10.5 ± 0.8
12 – 16	1582 ± 53	1.45	10.1 ± 0.7
16 – 24	1238 ± 59	1.60	8.2 ± 0.7
24 – 36	268 ± 23	1.30	4.3 ± 0.8
	Semicentral (30 – 50%)		
	Raw-yield (3σ)	S/B (3σ)	Significance (3σ)
2 – 4	174 ± 38	0.09	3.8 ± 0.8
4 – 6	185 ± 32	0.15	5.0 ± 0.8
6 – 8	220 ± 21	0.84	10.0 ± 0.7
8 – 12	187 ± 18	1.08	9.9 ± 0.7
12 – 16	125 ± 14	1.66	8.8 ± 0.7
16 – 24	268 ± 23	0.85	5.1 ± 0.7

applied in central collisions compared to semicentral, higher raw-yield values are obtained in central collisions due to the larger production of charm quarks, which scales with the number of binary nucleon-nucleon collisions enabling a higher significance and, ultimately, allowing us to extend the measurements to higher p_T .

In the fits to the invariant mass distribution, a Gaussian function was employed to model the signal peak, while an exponential function was used to describe the background. An additional Gaussian function was used to model the D^+ peak resulting from D^+ mesons decaying in the same decay channel under study for the D_s^+ meson. The Gaussian width of the D_s^+ peak was fixed to the one obtained from the prompt-enhanced sample in order to provide a more stable fit. This choice is motivated by the natural larger abundance of prompt D_s^+ mesons compared to non-prompt.

The invariant-mass distributions of D_s^+ meson candidates from the non-prompt enhanced samples in central and semicentral collisions in the different p_T intervals of the analyses are reported together with the fit functions in Figs. 6.3 and 6.4, respectively.

6.2.2 Acceptance-times-efficiency corrections

The computation of the p_T -differential corrected yields of the non-prompt D_s^+ mesons requires the raw-yields to be corrected for the detector acceptance and the reconstruction and selection efficiency, as reported in Eq. 6.1. The acceptance-times-efficiency term is obtained from MC simulation based on the PYTHIA 8 and HIJING 1.36 [173] event generators. As anticipated, different MC samples were used for the calculation of the $(\text{Acc} \times \epsilon)$ term with respect to the ones adopted in the BDT training and test to avoid introducing a bias in the estimation of the efficiency. The p_T distributions of the prompt D_s^+ and the parent B mesons in the simulation were weighted in order to reproduce a realistic p_T distribution. For this purpose, the p_T weights were computed as the ratio between the p_T distributions of D_s^+ and B mesons obtained with FONLL calculations multiplied by the R_{AA} predicted by the TAMU model and those from PYTHIA 8. The R_{AA} of both strange and non-strange B mesons were used to weight the p_T distribution of B^0 , B^+ , and B_s^0 mesons in the simulations since a slightly different shape is predicted by the TAMU model.

The top panel of Fig. 6.5 shows the acceptance-times-efficiency term as a function of p_T for central and semicentral Pb–Pb collision for prompt and non-prompt D_s^+ mesons with the selections employed in the measurements of the non-prompt D_s^+ meson production. The acceptance-times-efficiency term includes also the correction factor for the rapidity interval under consideration ($c_{\Delta y}$). A common increase of the efficiency as a function of p_T is observed for both the centrality intervals in the case of prompt and non-prompt D_s^+ mesons. Indeed, the displacement of the D_s^+ meson decay vertex from the primary vertex is increased by the Lorentz boost at higher p_T favouring the separation of signal and background. The non-prompt D_s^+ mesons display a higher efficiency compared to the prompt ones. This is expected considering that the BDT selections are tuned to enhance the contribution of the non-prompt D_s^+ mesons in the selected sample.

The bottom panel of Fig. 6.5 shows the ratio of the acceptance-times-efficiency term of non-prompt D_s^+ mesons to the prompt ones in the two centrality intervals. A relative difference of a factor of 5 – 20 depending on the p_T and the centrality is observed. In particular, in the 0 – 10% centrality class, the suppression of the prompt component is found to increase with increasing p_T , leading to a higher ratio of the efficiencies at higher p_T . An opposite trend is observed in semicentral collisions. This is due to the fact that less stringent selection criteria are applied compared to central collisions. Hence, a lower suppression of the prompt D_s^+ meson is obtained in the 30 – 50% centrality class, leading to a lower efficiency ratio as a function of p_T .

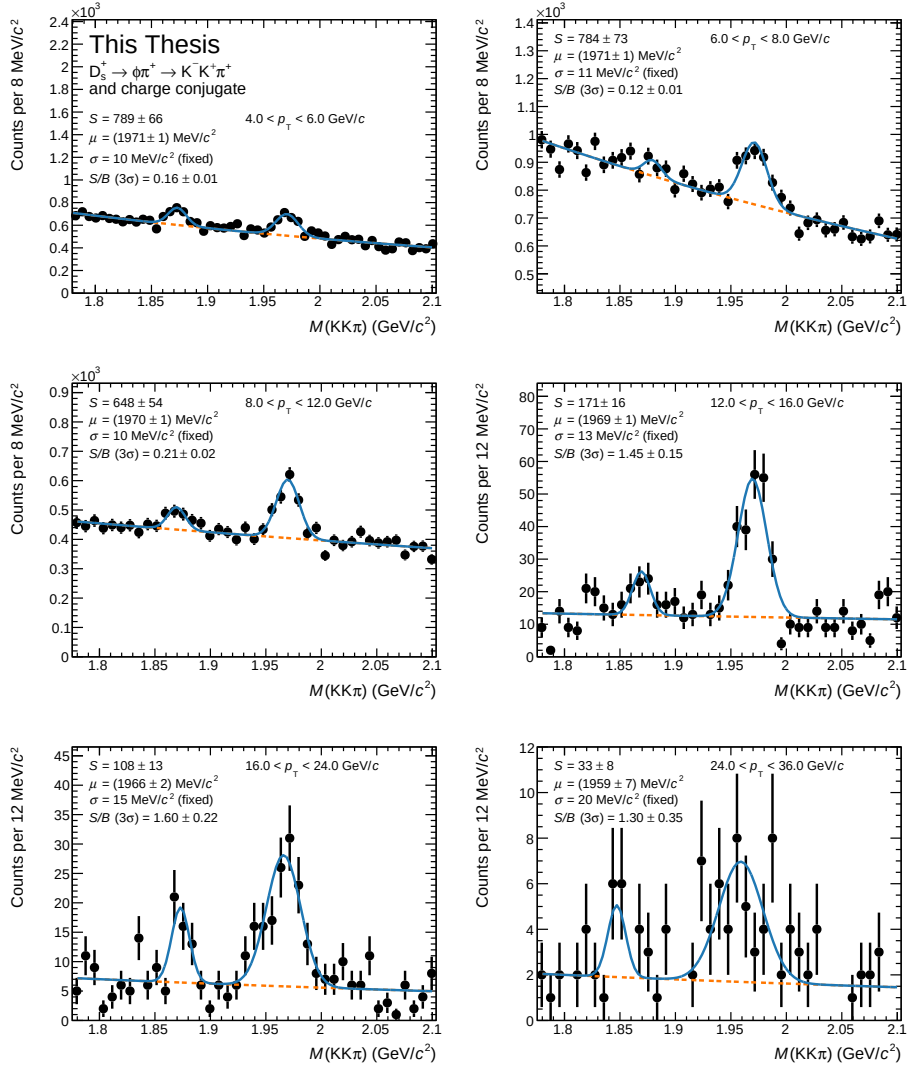


Fig. 6.3 Invariant-mass distribution of non-prompt D_s^+ meson candidates in central (0 – 10%) Pb–Pb collisions at $\sqrt{s_{\text{NN}}} = 5.02$ TeV in the p_T intervals of the analyses. The blue solid line represents the result of the fit to the data. The orange dashed line represents the background component of the fit.

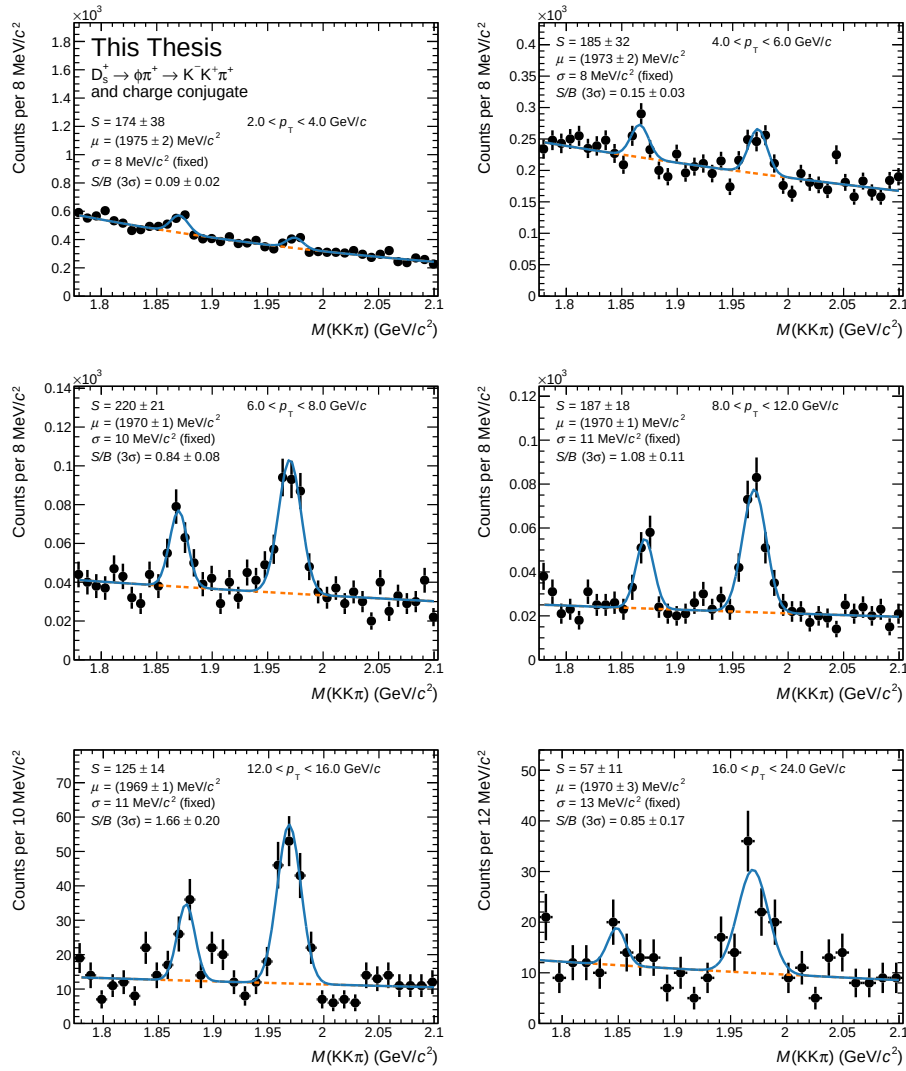


Fig. 6.4 Invariant-mass distribution of non-prompt D_s^+ meson candidates in semicentral (30 – 50%) Pb–Pb collisions at $\sqrt{s_{NN}} = 5.02$ TeV in the p_T intervals of the analyses. The blue solid line represents the result of the fit to the data. The orange dashed line represents the background component of the fit.

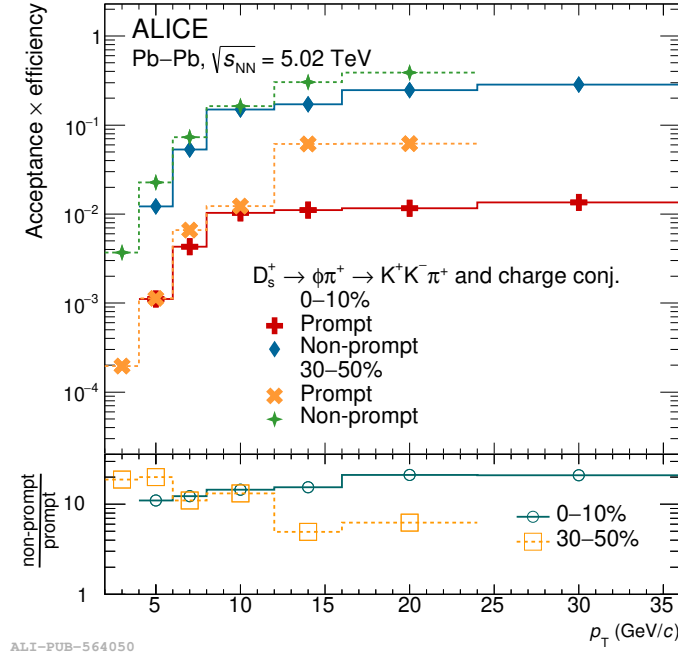


Fig. 6.5 Acceptance-times-efficiency as a function of p_T for central and semicentral Pb–Pb collisions for prompt and non-prompt D_s^+ mesons in the non-prompt enhanced sample. The bottom panel shows the ratio of the efficiency of non-prompt and prompt D_s^+ mesons for the two centrality classes. Figure from Ref. [335].

6.2.3 Non-prompt D_s^+ meson fraction

The non-prompt D_s^+ meson fraction in the extracted raw-yields was estimated with the data-driven method presented in Sec. 5.2.3. This method has the important advantage of being independent of theoretical predictions, although it relies on the acceptance-times-efficiency estimated from MC simulations as described above. In the data-driven method, a set of BDT selection criteria is defined. The selections are chosen to obtain a significant variation of the contribution of prompt and non-prompt D_s^+ mesons in the extracted raw yields. In particular, the threshold value on the BDT score related to the probability of the candidate to be a non-prompt D_s^+ was progressively varied, while keeping the threshold on the probability to be background fixed to the optimised value. Each set of BDT selections is associated with an equation linking the corresponding prompt and non-prompt efficiency terms and the corrected yields of prompt and non-prompt D_s^+ with the extracted raw yields. The set of BDT selections can then be interpreted as a system of equations. The system can be solved in an approximated way in terms of the "true" amount of prompt and non-prompt D_s^+ mesons in the selected sample, which are then used to estimate the non-prompt D_s^+ meson fractions in the raw yields extracted with the reference configuration.

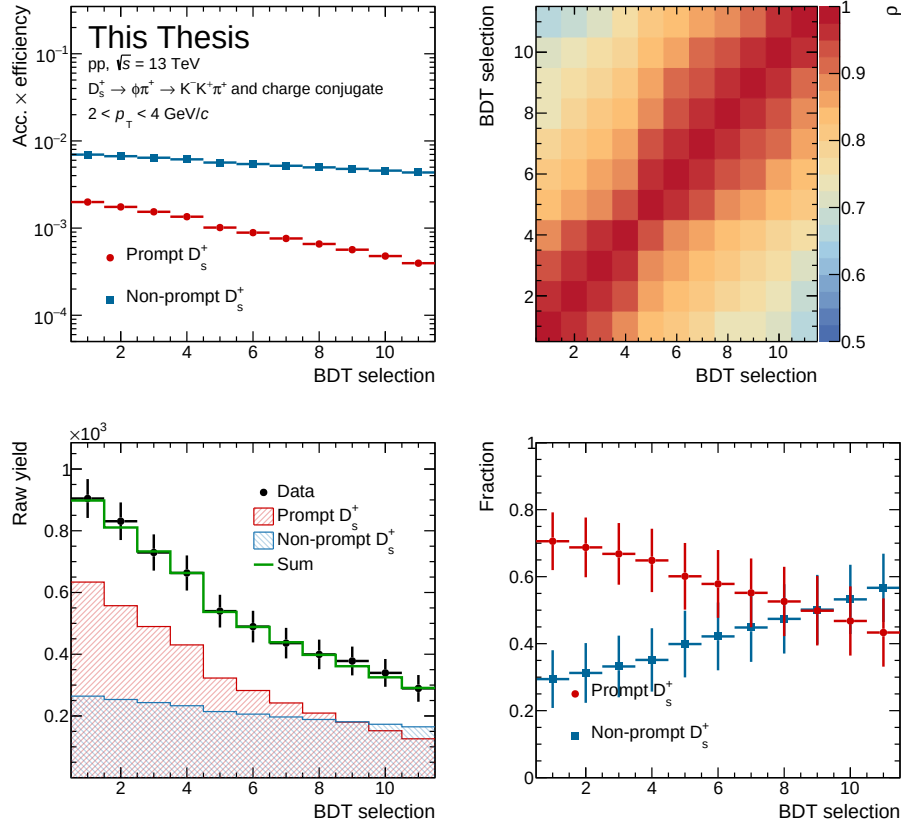


Fig. 6.6 Top-left: $(\text{Acc} \times \epsilon)$ factors for prompt and non-prompt D mesons as a function of the set of selection. Top-right: correlation matrix between pairs of sets of selection. The colour scale represents the correlation coefficient between the two sets of selections. Bottom-left: raw-yield as a function of the cut set with the prompt and non-prompt contributions obtained from the minimisation procedure. Bottom-right: prompt and non-prompt fractions as a function of the cut set.

The results obtained with this technique for semicentral collisions are reported in Fig. 6.6 for the $2 < p_T < 4$ GeV/c interval. The top-left panel displays the $(\text{Acc} \times \epsilon)$ factors for prompt and non-prompt D_s^+ mesons as a function of the set of BDT selections identified by a progressive number, with higher numbers corresponding to tighter selection criteria. The top-right panel displays the correlation matrix between pairs of sets of BDT selections computed as described in Sec. 5.2.3. In the bottom-left panel the values of the raw yield, obtained from the minimisation procedure, are reported for the prompt and non-prompt contributions as a function of the cut set, while the bottom-right panel shows the resulting prompt and non-prompt fractions as a function of the cut set.

Figure 6.7 shows the non-prompt D_s^+ meson fraction as a function of p_T in central and semicentral Pb–Pb collisions at $\sqrt{s_{NN}} = 5.02$ TeV. The statistical and systematic uncertainties

associated with the estimation of the non-prompt D_s^+ meson fraction are reported as vertical bars and boxes, respectively. The systematic uncertainty associated with the estimation of the non-prompt D_s^+ meson fraction was assessed by repeating the minimisation procedure after varying the set of BDT selections taken into account in the system to include or exclude tighter and lower selections. The systematic uncertainty is then assigned considering the variation of the measured p_T -differential non-prompt D_s^+ meson corrected yields obtained with the different non-prompt fractions. The $f_{\text{non-prompt}}$ values are found to vary between 0.6 and 0.8 in the 0 – 10% and 30 – 50% centrality classes. It is important to highlight the fact that the natural abundance of non-prompt D_s^+ mesons consists of 5% at low p_T and 20% for $p_T \sim 30$ GeV/ c . Hence, the optimised BDT selections were crucial to obtain a substantial contribution of non-prompt D_s^+ mesons in the selected sample. The fraction of non-prompt D_s^+ mesons is on average lower in semicentral collision compared to central collisions. This difference is attributed to the looser BDT selections employed in semicentral collisions compared to the ones adopted in central collisions, which leads to a larger contamination of prompt D_s^+ mesons in the extracted raw yields. For central collisions, no significant dependence on the p_T of the D_s^+ meson is observed, while for semicentral collision a decreasing trend with increasing p_T is found. This is attributed to the fact that much looser BDT selections are employed in semicentral collisions compared to the central ones. Hence, the contamination of prompt D_s^+ mesons in the extracted raw yields is larger in semicentral collisions, leading to a lower non-prompt D_s^+ meson fraction.

6.2.4 Proton-proton reference

The measurements of the non-prompt D_s^+ cross section in pp collisions at $\sqrt{s} = 5.02$ TeV [239] were used as reference for the computation of the R_{AA} . Given that the maximum p_T reached in the measurement is 12 GeV/ c , the cross section was extrapolated to higher p_T exploiting FONLL calculations. To this end, the same procedure adopted in Ref. [337] was adopted. Given that no FONLL prediction is available for the D_s^+ meson, the procedure described in detail Sec. 5.1.3 was followed. In particular, the FONLL prediction of the charm-quark cross section was combined with the charm quark fragmentation functions from Ref. [272], and the fragmentation fractions into D_s^+ mesons from e^+e^- collisions [273] assuming them to be independent on the collision system. In particular, the extrapolation was computed as the ratio between the measured p_T -differential cross section and the one predicted by FONLL in the p_T intervals of the measurements. The p_T -differential ratio was then fitted with a constant

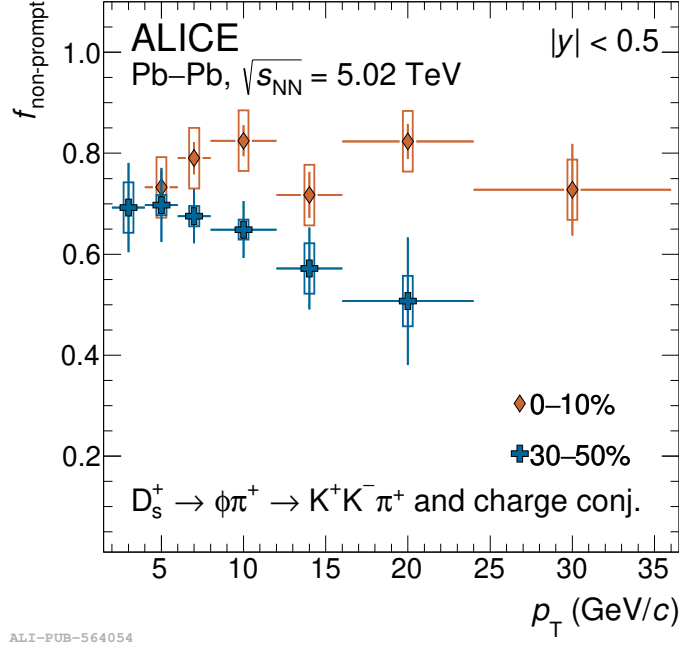


Fig. 6.7 Non-prompt D_s^+ fraction as a function of p_T in central and semicentral Pb–Pb collisions at $\sqrt{s_{NN}} = 5.02$ TeV. The statistical and systematic uncertainties associated with the estimation of the non-prompt D_s^+ meson fraction are reported as vertical bars and boxes, respectively. Figure from Ref. [335].

function. Finally, the cross section in the unmeasured p_T intervals was obtained as:

$$\frac{d\sigma}{dp_T}(\text{extrap}) = \frac{d\sigma}{dp_T}(\text{FONLL}) \times R, \quad (6.4)$$

where R is the parameter of the constant fit to the p_T -differential ratio between data and FONLL in the measured p_T range. It is noteworthy to mention that the data/FONLL ratio is almost independent on the p_T in the measured p_T range, hence the R parameter takes into account the fact that FONLL underestimates the measurements. The same procedure was repeated for the upper and lower edges of the FONLL uncertainties band which are driven by the choice of the normalisation and factorisation scales, the heavy-quark mass, and the uncertainties on the PDFs. The magnitude of the systematic uncertainty associated with the extrapolated cross section was defined by propagating the statistical and systematic uncertainties of the measurement to the scaling factor R . The systematic uncertainties on the measurement were considered as fully correlated in p_T .

The resulting cross section in pp collisions at $\sqrt{s_{NN}} = 5.02$ TeV is reported in Fig. 6.8, where the extrapolated points are depicted with open markers.

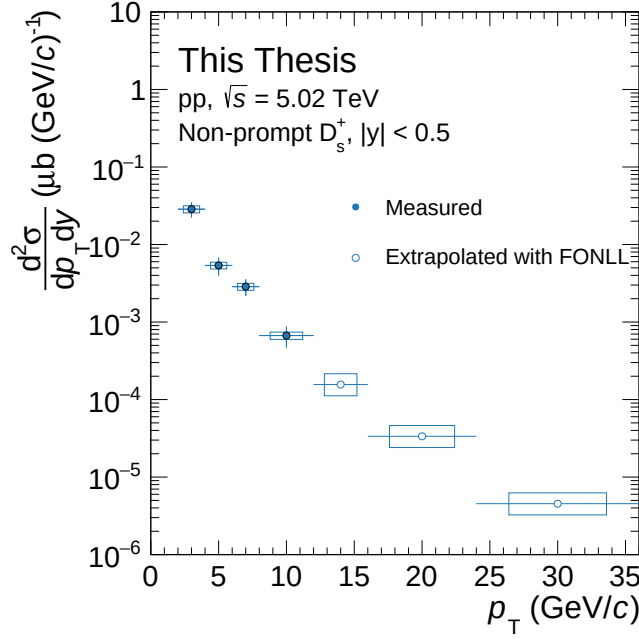


Fig. 6.8 Non-prompt D_s^+ meson pp reference cross section at $\sqrt{s} = 5.02$ TeV at midrapidity. The full markers indicate the measured cross section, while the open markers indicate the extrapolated cross section.

6.3 Systematic uncertainties

The sources of systematic uncertainty affecting the measurements of the non-prompt D_s^+ -meson p_T -differential corrected yields are: (i) raw-yield extraction, (ii) track-reconstruction efficiency, (iii) PID efficiency, (iv) non-prompt fraction estimation, (v) D-meson selection efficiency, (vi) D-meson simulated p_T shape, (vii) B-meson relative abundances. The magnitude of these effects was assessed independently for the two centrality classes and the obtained values are reported in Tables 6.5 and 6.6, respectively. The total systematic uncertainty also includes the contributions from the BR uncertainty [6], and from the determination of the collision centrality, which was found to be lower than 0.1% and around 2% for central and semicentral collisions, respectively, as reported in Ref. [337], and from the luminosity determination, which was found to be 2.1% [233]. An additional source of uncertainty, solely affecting the R_{AA} , comes from the determination of the $\langle T_{AA} \rangle$ which was found to be 0.7% and 1.5% for central and semicentral collisions, respectively [334]. It is worth noticing that the uncertainty associated with the BR cancels out in the computation of the R_{AA} .

6.3.1 Systematic uncertainty on the raw-yield extraction

The estimation of the systematic uncertainty related to the raw-yield extraction was performed by varying the fit configuration (i.e. the background fit function, the upper and lower fit limits, the binning of the invariant-mass distribution, etc.). The Gaussian width of the D_s^+ peak, which was fixed to the value extracted from the prompt-enhanced sample, was varied within the uncertainty range of the D_s^+ meson width in the prompt-enhanced sample. To exclude non-reliable or extreme fit configurations, the same quality selections detailed in Sec. 5.3.1, for the pp measurements, were applied. The systematic uncertainty associated with the raw-yield extraction was estimated considering the sum in quadrature of the RMS of the distribution of the raw-yield values obtained with the different fit configurations and the shift of the mean of this distribution with respect to the reference value. In addition, as a cross-check, the raw-yield values extracted from the fits were compared to those obtained from the Bin Counting method by counting the number of candidates in the invariant-mass region of the signal after the subtraction of the background estimated from a fit to the sidebands.

Table 6.5 Relative systematic uncertainties of the measured non-prompt D_s^+ -meson corrected yields and R_{AA} in central (0 – 10%) Pb–Pb collisions at $\sqrt{s_{NN}} = 5.02$ TeV in the transverse momentum-intervals of the analysis. Uncertainties found to be negligible are indicated as such (negl.).

p_T interval (GeV/c)	Non-prompt D_s^+ meson (0-10%)					
	4 – 6	6 – 8	8 – 12	12 – 16	16 – 24	24 – 36
Signal extraction	5%	5%	5%	5%	5%	10%
Tracking efficiency	13%	13%	13%	12%	10%	8%
PID efficiency	negl.	negl.	negl.	negl.	negl.	negl.
Non-prompt fraction	6%	6%	6%	6%	6%	12%
Selection efficiency	8%	5%	5%	5%	5%	5%
MC p_T shape	10%	8%	8%	2%	2%	2%
B-meson relative abundances	1.0%					
$\langle T_{AA} \rangle$	0.7%					
Branching ratio	3.6%					
Luminosity	2.1%					

Table 6.6 Relative systematic uncertainties of the measured non-prompt D_s^+ -meson corrected yields and R_{AA} in semicentral (30–50%) Pb–Pb collisions at $\sqrt{s_{\text{NN}}} = 5.02$ TeV in the transverse momentum-intervals of the analysis. Uncertainties found to be negligible are indicated as such (negl.).

p_{T} interval (GeV/c)	Non-prompt D_s^+ meson (30-50%)					
	2–4	4–6	6–8	8–12	12–16	16–24
Signal extraction	10%	10%	5%	4%	4%	6%
Tracking efficiency	11%	12%	13%	13%	12%	10%
PID efficiency	negl.	negl.	negl.	negl.	negl.	negl.
Non-prompt fraction	5%	2%	2%	2%	8%	5%
Selection efficiency	10%	5%	5%	5%	5%	5%
MC p_{T} shape	15%	10%	5%	2%	5%	2%
B-meson relative abundances	1.0%					
$\langle T_{\text{AA}} \rangle$	1.5%					
Branching ratio	3.6%					
Luminosity	2.1%					

6.3.2 Systematic uncertainty on the track-reconstruction efficiency

The systematic uncertainty on the track-reconstruction efficiency accounts for possible discrepancies between data and MC in the TPC-ITS track prolongation efficiency and in the selection efficiency due to track-quality criteria in the TPC. The procedure described in Sec. 5.3.2 was employed to estimate this systematic uncertainty. In particular, the track quality selections described in Sec. 4.2.3 were varied by changing the minimum number of crossed rows, the minimum value of the ratio of crossed-over-findable rows in TPC, the number of clusters times the TPC crossed-rows, and the number of available clusters to perform PID in TPC. The per-track systematic uncertainty was assigned to be equal to 4.5% and 3% for the 0 – 10% and 30 – 50% centrality classes, respectively¹.

The TPC-to-ITS prolongation efficiency was studied by comparing the performance in the data and MC simulations. The procedure followed to perform the comparison is the one described in detail in Sec. 5.3.2. The resulting p_T -dependent per-track systematic uncertainty was propagated to the D_s^+ meson considering the decay kinematics and summing linearly the contribution of the tracks. The systematic uncertainty associated with the track-reconstruction efficiency was assigned as the sum in quadrature of the two contributions.

6.3.3 Systematic uncertainty on the non-prompt fraction estimation

The systematic uncertainty associated with the non-prompt fraction estimation was estimated by varying the set of BDT selection criteria utilised in the data-driven method. Namely, the sets of selections were extended or reduced by adding or removing more stringent and/or looser selections. In addition, alternative step sizes among the thresholds on the BDT scores were employed. For each set, the estimation of the non-prompt fraction was repeated. The systematic uncertainty was assigned considering the dispersion of the corrected yield values obtained with the various sets and the shift of the mean with respect to the reference value. Note that the effect of the selection variation was studied on the corrected yields rather than on the non-prompt fraction to account for the variation of the prompt D_s^+ contribution when varying the selection criteria. The procedure was performed separately for central and semicentral collisions and similar values of systematic uncertainty were found.

¹Slightly different values were found for strange and non-strange D mesons in Refs. [117] and [240], respectively. Given the higher precision of the measurement, the former results were adopted in this Thesis.

6.3.4 Systematic uncertainty on the D-meson selection efficiency and PID efficiency

The systematic uncertainty related to the estimation of the selection efficiency of the BDT selection criteria was estimated by varying the applied threshold value on the BDT probability. For each variation, the full analysis chain was repeated and the corresponding corrected yields were computed. The magnitude of the systematic uncertainty was assigned as the sum in quadrature of the RMS of the distribution of the corrected yields and the shift of the mean of this distribution with respect to the reference value. Note that the choice of checking the effect of the variations in the corrected yields rather than on the efficiency is motivated to account for the variation of the prompt D_s^+ meson selection efficiency, which contributes to the estimation of the non-prompt fraction. Hence, it ultimately affects the non-prompt D_s^+ corrected yields.

The BDT selections on the probability to be background and non-prompt D_s^+ were varied independently and simultaneously providing a sample of about 100 configurations. The range in which the selections were varied was chosen to induce a relative variation of the selection efficiency ranging between 0.5 and 2.5. Selections exceeding this range, or not fulfilling minimum significance requirements, were excluded from the estimation of the systematic uncertainty. In this way, problematic or extreme selections do not influence the estimation of the systematic uncertainty. The procedure was performed separately for central and semicentral collisions in each p_T interval of the analysis.

The systematic uncertainty associated with the PID efficiency was estimated by varying the PID selection criteria employed to prepare the data sample for machine-learning purposes. In particular, a pure sample of pions and kaons was selected in either TPC or TOF by applying strict PID selections based on the TOF or TPC PID information, respectively. No significant deviation was observed as a function of p_T for both pions and kaons in the p_T intervals of the analyses. Hence, the systematic uncertainty was considered to be negligible for both the centrality classes, in line with the findings of the pp measurements reported in Sec. 5.3.2 and Refs. [240, 117].

6.3.5 Systematic uncertainty on the D-meson simulated p_T shape

A systematic uncertainty is specifically assigned to account for possible discrepancies between the p_T distribution of B and D mesons used in the MC simulations and the one observed in data. The systematic uncertainty associated with the p_T shape was evaluated by

adopting alternative p_T -shapes compared to the reference one. In particular, the reference p_T shape of prompt D_s^+ and B mesons was obtained by multiplying the p_T distribution of D mesons obtained from FONLL calculations by the R_{AA} predicted by the TAMU model. The TAMU model was chosen since it implements both the in-medium energy loss and the enhanced charm- and beauty-hadron production via the recombination mechanism in a strangeness-rich environment. As an alternative p_T shape the LIDO model [142] was adopted. The main difference between the TAMU and LIDO models derives from the fact that the former includes the enhanced production of B_s^0 mesons, unlike the latter. An additional variation of the shape of the p_T distributions of prompt D_s^+ mesons was included considering the results from Ref. [336]. The p_T shape predicted by FONLL calculation times the R_{AA} predicted by TAMU model was considered as reference. The systematic uncertainty was assigned taking into account the variation of the corrected yield compared to the default case.

6.3.6 Systematic uncertainty on the B-meson relative abundances

Beauty hadrons decay into D_s^+ mesons through different decay channels characterised by different decay kinematics. In addition, different beauty hadron species have different lifetimes and mean proper decay lengths resulting in different topologies for the produced non-prompt D_s^+ . Hence, the selection efficiency of non-prompt D_s^+ mesons can be influenced by the relative abundances of the parent hadron species. An imperfect description in the MC simulation of the beauty-hadron composition, or *hadrochemistry*, might induce a bias in the estimation of the non-prompt D_s^+ meson efficiency. In the case of D_s^+ mesons, all the three ground-state B meson species, namely B^+ ($\tau \simeq 1.638$ ps [6]), B^0 ($\tau \simeq 1.519$ ps [6]), and B_s^0 ($\tau \simeq 1.521$ ps [6]) mesons, significantly contribute to the overall D_s^+ production. Indeed, as discussed in Sec. 5.4.7, half of the D_s^+ mesons originate from the decay of non-strange B mesons, while the other half originates from B_s^0 mesons. Thus, the non-prompt D_s^+ mesons might be significantly affected by the hadrochemistry of the B meson species in the simulations. This can be estimated by looking at the $(\text{Acc} \times \varepsilon)$ factor for non-prompt D_s^+ mesons originating from B^0 , B^+ , and B_s^0 , respectively. The results are shown in Fig. 6.9 and for $p_T > 4$ GeV/c a hierarchy reflecting the lifetimes of the three B mesons is observed.

As discussed in detail in Sec. 5.4.5, the PYTHIA 8 generator quantitatively reproduces the measurements of non-prompt strange and non-strange D_s^+ mesons in pp collisions and, consequently, the relative abundances of the different B meson species. However, in heavy-ion collisions an enhanced thermal production of strange quarks is expected compared to pp collisions, leading to an enhancement of the B_s^0 -meson production compared to the one of B^0 and B^+ mesons. However, since no precise measurements of B_s^0 -meson production in the

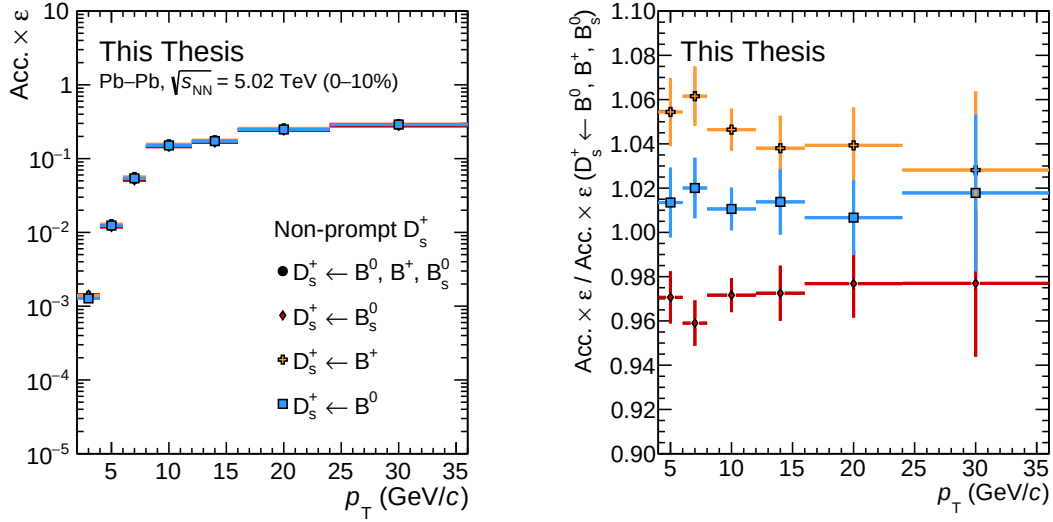


Fig. 6.9 $(\text{Acc} \times \epsilon)$ factors for non-prompt D_s^+ mesons originating from $B^0, B^+,$ and B_s^0 mesons as a function of p_T in central Pb–Pb collisions at $\sqrt{s_{\text{NN}}} = 5.02$ TeV. Right: ratio of the $(\text{Acc} \times \epsilon)$ factors for non-prompt D_s^+ mesons originating from $B^0, B^+,$ and B_s^0 mesons as a function of p_T in central Pb–Pb collisions at $\sqrt{s_{\text{NN}}} = 5.02$ TeV. The statistical and systematic uncertainties are depicted as vertical lines and empty boxes, respectively.

low- p_T region are currently available in heavy-ion collisions, the abundances predicted by PYTHIA 8 were employed as central values. The magnitude of the systematic uncertainty associated with this assumption was assessed considering the enhancement of B_s^0 -meson production predicted by the TAMU model. Practically, this was done by reweighting the B meson abundances predicted by PYTHIA 8 in order to increase by a factor of 2 the contribution of B_s^0 mesons to reproduce the enhancement predicted by the TAMU model. The systematic uncertainty was then assigned considering the effect on the measured corrected yield of non-prompt D_s^+ mesons with and without the weights. The assigned systematic uncertainty is 1% independently of the p_T interval for both the centrality classes. The fact that the effect is small justifies the coarse reweight.

6.4 Results

The p_T -differential production yield of non-prompt D_s^+ mesons at midrapidity ($|y| < 0.5$) in central and semicentral Pb–Pb collisions at $\sqrt{s_{\text{NN}}} = 5.02$ TeV is shown in Fig. 6.10. This is the first measurement ever reported of non-prompt D_s^+ mesons in heavy-ion collisions. The results are compared with the measurements of the prompt D_s^+ , taken from Ref. [336]

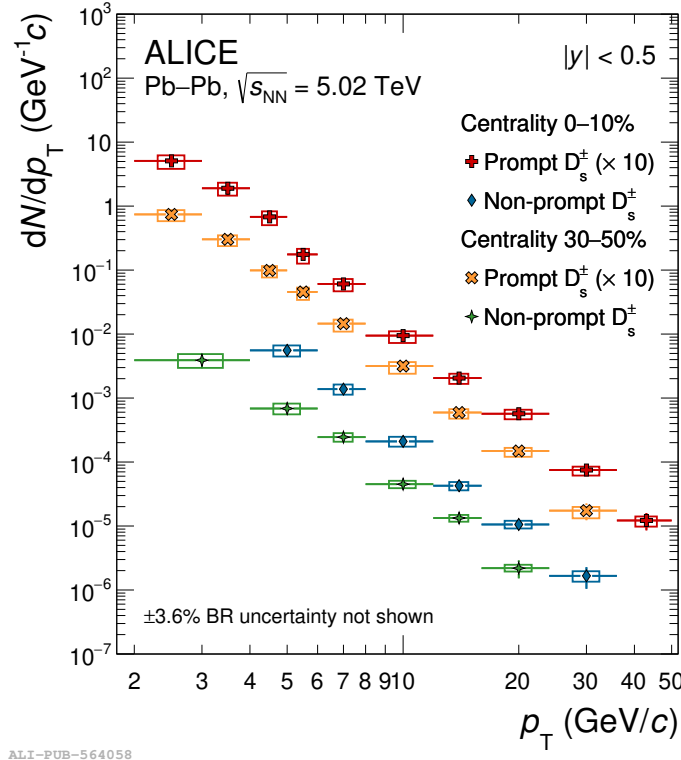


Fig. 6.10 Prompt [336] and non-prompt D_s^+ mesons p_T -differential production yield in central (0 – 10%) and semicentral (30 – 50%) Pb–Pb collisions at $\sqrt{s_{NN}} = 5.02$ TeV. The statistical and systematic uncertainties are depicted as vertical lines and empty boxes, respectively. The prompt D_s^+ meson production yield is scaled by a factor of 10 for visibility reasons. Figure from Ref. [335].

and scaled by a factor of 10 for visibility reasons. Hereafter, the statistical and systematic uncertainties are depicted as vertical lines and empty boxes, respectively. The systematic uncertainty in Fig. 6.10 does not include the uncertainty associated with the BR of the hadronic decay channel under study, which is reported separately in the Figure.

The results for the prompt D_s^+ mesons in both the centrality classes were obtained by the ALICE Collaboration following a procedure similar to the one described in this Thesis. However, in that case, a binary classification BDT algorithm was employed. This approach is justified by the natural larger abundance of prompt D_s^+ mesons compared to non-prompt. To further test the stability of the predictions based on the multiclass BDT algorithm, the production yields extracted with the binary classification were compared to those that I extracted from the prompt-enhanced sample obtained with the multiclass approach in the 0 – 10% centrality class. The results are reported in Fig. 6.11. The values obtained with the two approaches are fully compatible with each other meaning that the multiclassification algorithm can reproduce the published results based on the binary classification.

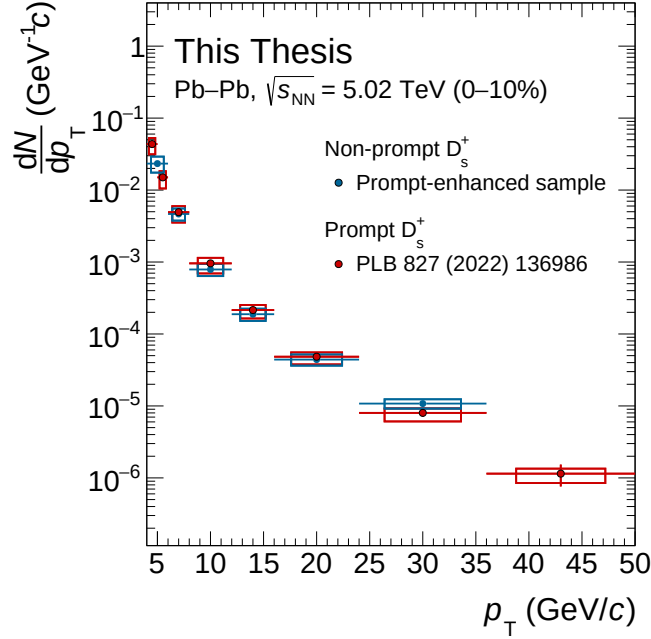


Fig. 6.11 Comparison of the prompt D_s^+ meson p_T -differential production yield in central (0 – 10%) and semicentral (30 – 50%) Pb–Pb collisions at $\sqrt{s_{NN}} = 5.02$ TeV obtained with the prompt-enhanced sample from the analysis method presented in this Thesis and the one from Ref. [336]. The statistical and systematic uncertainties are depicted as vertical lines and empty boxes, respectively.

The effects due to the QGP medium formation in heavy-ion collisions can be investigated by measuring the relative production of different D-meson species along with the prompt and non-prompt contributions to each of them. In particular, measuring the non-prompt-to-prompt production yield ratio allows us to disentangle the effect of the in-medium energy loss on-frequency charm and beauty quarks. Instead, comparing the production yield of non-prompt D_s^+ and non-prompt D^0 is extremely useful to highlight effects related to the modification of the hadronisation mechanism of beauty quarks in the presence of the QGP medium.

Figure 6.12 reports the ratios of the production yields of non-prompt to prompt D_s^+ (left panel) and non-prompt D_s^+ to D^0 (right panel) in central and semicentral Pb–Pb collisions, as well as in pp collisions at the same centre-of-mass energy [239]. The ratio of non-prompt to prompt D_s^+ mesons increases from about 0.05 in the lowest p_T intervals of the measurements to 0.20 at high p_T (> 10 GeV/c). The increasing trend with increasing p_T is compatible with the analogous ratio measured in pp collisions at $\sqrt{s} = 13$ TeV reported in Fig. 5.23. In the p_T range of $4 < p_T < 12$ GeV/c, the non-prompt to prompt D_s^+ meson ratio in the 0 – 10% centrality class also shows a mild enhancement compared to pp collisions, with a

significance of 1.6σ , where σ represents the quadratic sum of the statistical and systematic uncertainties. In the calculation of the number of sigmas, all systematic uncertainties, except those related to raw-yield extraction, are considered to be fully correlated among p_T bins. Even though the current experimental uncertainties do not allow us to draw a definitive conclusion, the observed trend is compatible with a scenario where beauty quarks are less affected by the in-medium energy loss compared to charm quarks. This is expected in energy loss calculations because of the larger mass of beauty quarks compared to charm quarks, which limits the emission angle of the radiated gluons in inelastic scattering processes, as detailed in Sec 1.5.3. Note that also a different contribution of the hadronisation mechanism via recombination between charm and beauty quarks can influence the observed trend. At higher p_T (> 12 GeV/ c), the non-prompt to prompt D_s^+ ratios are found to be compatible in central and peripheral collisions.

The ratio of non-prompt D_s^+ mesons to non-prompt D^0 mesons displays a relatively constant trend around 0.2 within the measured p_T range. The ratios in pp and semicentral Pb–Pb collisions are consistent within the experimental uncertainties. However, there is a mild indication of enhancement at low p_T (< 6 GeV/ c) when comparing central Pb–Pb collisions to pp collisions, with a significance of 1.7σ . The estimation is derived from a weighted average of the non-prompt D_s^+/D^0 ratio values in the $4 < p_T < 12$ GeV/ c range for the 0 – 10% centrality class. The inverse of the squared sum of statistical and p_T -uncorrelated systematic uncertainties are employed as weights in the calculation of the average of the ratios in the p_T intervals of the measurement. All systematic uncertainties, except those related to raw-yield extraction, are considered to be fully correlated in p_T . This hint of a higher non-prompt D_s^+/D^0 yield ratio at low and intermediate p_T aligns with the expectation of enhanced production of strange-beauty mesons in heavy-ion collisions compared to pp collisions. This enhancement is attributed to the hadronisation of beauty quarks via recombination with surrounding quarks within the QGP medium and the large strange-quark abundance in the QGP.

The D_s^+ -meson production yields are employed together with the proton-proton reference and the T_{AA} to compute the nuclear modification factor using the Eq. 6.2. The $\langle T_{AA} \rangle$ for the considered centrality classes is estimated via Glauber-model simulations anchored to the V0 detector signal amplitude, following the procedure described in Sec. 2.4. The $\langle T_{AA} \rangle$ values are (23.26 ± 0.17) mb $^{-1}$ and (3.92 ± 0.06) mb $^{-1}$ for central and semicentral centrality intervals, respectively [334].

The R_{AA} is a powerful and experimentally-accessible observable to assess the effects induced on the particle p_T -differential yields by the production of the QGP medium in Pb–Pb

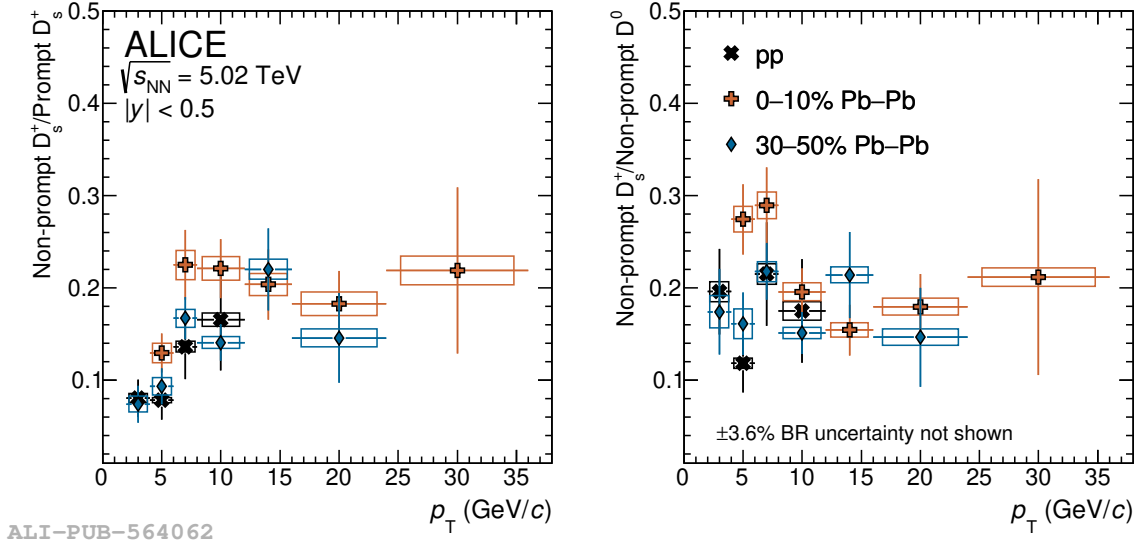


Fig. 6.12 Left: ratio of the production yield of non-prompt to prompt D_s^+ mesons in central (0 – 10%) and semicentral (30 – 50%) Pb–Pb collisions at $\sqrt{s_{NN}} = 5.02$ TeV, and pp collisions at $\sqrt{s} = 5.02$ TeV. Right: ratio of the production yield of non-prompt D_s^+ to non-prompt D^0 mesons in central (0 – 10%) and semicentral (30 – 50%) Pb–Pb collisions at $\sqrt{s_{NN}} = 5.02$ TeV, and pp collisions at $\sqrt{s} = 5.02$ TeV. The statistical and systematic uncertainties are depicted as vertical lines and empty boxes, respectively. Figure from Ref. [335].

collisions. Heavy quarks are mainly produced in hard scattering processes, hence their production is expected to be proportional to the number of binary nucleon-nucleon collisions. Therefore, the R_{AA} is foreseen to be equal to unity for heavy flavour hadrons in the absence of nuclear effects. The production of the QGP can modify this expectation.

The R_{AA} values for prompt and non-prompt D_s^+ mesons are compared in the upper- and lower-left panels of Fig. 6.13 for the 0 – 10% and 30 – 50% centrality classes, respectively. Similarly, the comparison between the nuclear modification factor of non-prompt D_s^+ and D^0 mesons is presented in the right panels of the same figure. The R_{AA} values for prompt and non-prompt D mesons, independently on the species, exhibit a decreasing trend as a function of p_T , reaching a minimum around 0.2 – 0.4 for $p_T \sim 10$ GeV/c in the 0 – 10% and 30 – 50% centrality classes. At low p_T , the R_{AA} progressively increases with decreasing p_T reaching a maximum around unity for $p_T \simeq 2 - 4$ GeV/c. In particular, in the 0 – 10% centrality class for $p_T < 6$ GeV/c, the central values of the non-prompt D_s^+ -meson R_{AA} show a hint of enhancement compared to prompt D_s^+ and non-prompt D^0 , although they remain consistent within the experimental uncertainties. The enhancement in the case of the comparison of prompt and non-prompt D_s^+ mesons is expected to be related to the different in-medium energy loss experimented by charm and beauty quarks due to their

different masses, as predicted by the dead-cone effect. It is worth noticing that, given the different mass of charm and beauty quarks, the p_T interval in which the hadronisation via recombination is expected to dominate is different for the two quark species. In particular, the hadronisation via recombination is expected to cover a larger p_T region for the beauty quarks. In addition, in the case of non-prompt D_s^+ mesons, an additional effect is expected to come from the decay kinematic of the parent B mesons, which differs for the different B-meson species. Instead, the higher non-prompt D_s^+ -meson R_{AA} compared to the non-prompt D^0 R_{AA} at low p_T is associated with the modification of the hadronisation mechanism in the strangeness-rich environment of the QGP. In semicentral collisions, there is no significant separation among the R_{AA} values of prompt D_s^+ , non-prompt D_s^+ , and non-prompt D^0 within the current experimental uncertainties. This can be explained considering the smaller spatial extension of the fireball compared to central collisions and the subsequent lower magnitude of the effects due to the QGP medium.

The measured R_{AA} values are compared to the predictions of the TAMU model [126, 338]. In the TAMU model, the Langevin equation (Eq. 1.32) is employed to describe the transport of the heavy quarks in the QGP including elastic processes. The TAMU model implements the hadronisation via both coalescence and fragmentation. The uncertainty band for prompt D_s^+ mesons arises from the uncertainty associated with the modification of the PDF of charm quarks due to the shadowing effect occurring in nuclei as discussed in detail in Sec. 1.5.2. This effect is not considered for the beauty quarks, hence no uncertainty band is reported. The TAMU model qualitatively describes the trend of prompt and non-prompt D-meson R_{AA} . However, in the case of the non-prompt D_s^+ mesons, the model overestimates the measured R_{AA} .

The QGP-induced effects can be further investigated by looking at the ratio of the R_{AA} of non-prompt D_s^+ mesons to prompt D_s^+ mesons and non-prompt D^0 mesons. The left and right panels of Fig. 6.14 display the ratios of the nuclear modification factors of non-prompt and prompt D_s^+ mesons and non-prompt D_s^+ and D^0 mesons in central and semicentral Pb–Pb collisions, respectively. In the 0 – 10% centrality classes, the non-prompt to prompt D_s^+ R_{AA} ratio shows a hint of enhancement in the p_T region between 4 and 12 GeV/c, with a statistical significance of $1.6\sigma^2$. This R_{AA} ratio confirms the hint of a larger in-medium energy loss experienced by charm quarks when compared to beauty quarks due to their smaller mass. As already pointed out when discussing the particle ratios. In semicentral collisions, no significant indications of a R_{AA} ratio higher than unity are observed. Given the current experimental uncertainties, TAMU predictions are found to describe the results for central

²It is important to notice that the ingredients considered in the corrected yield ratios and the R_{AA} ratios are the same, hence the uncertainties at play are analogous and, ultimately, the σ values.

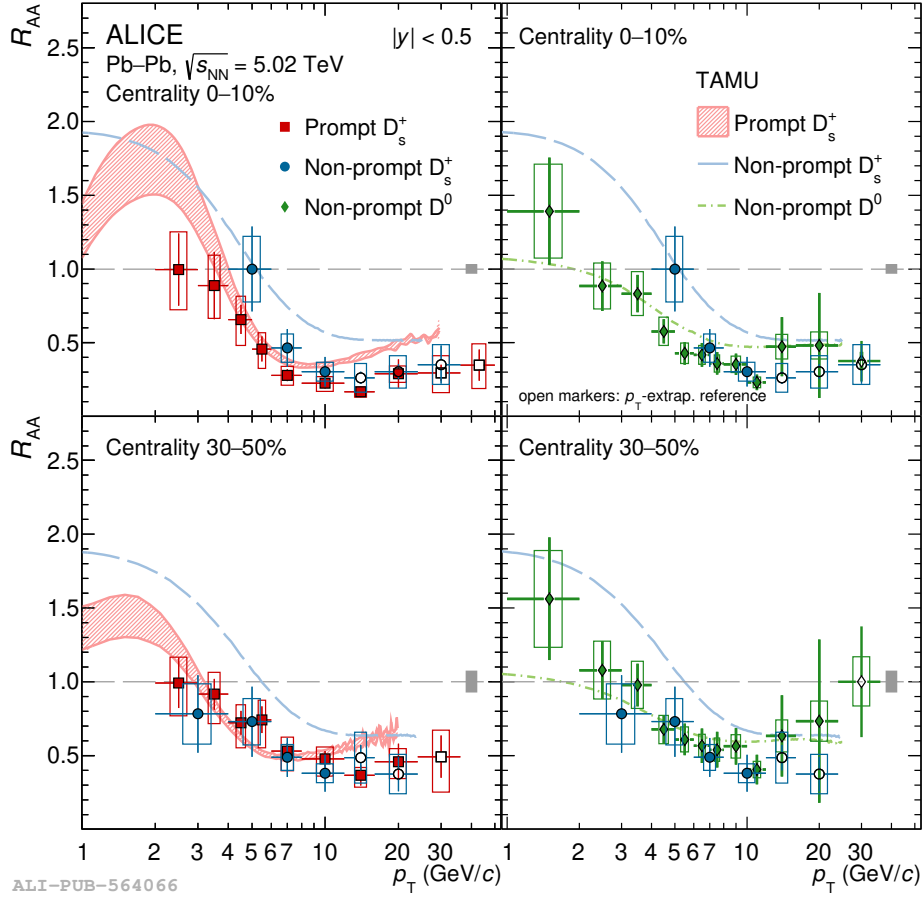


Fig. 6.13 Left: R_{AA} of non-prompt and prompt D_s^+ mesons in central (0 – 10%, top panel) and semicentral (30 – 50%, bottom panel) Pb–Pb collisions at $\sqrt{s_{\text{NN}}} = 5.02$ TeV. Right: R_{AA} of non-prompt D_s^+ and D^0 mesons in central (0 – 10%, top panel) and semicentral (30 – 50%, bottom panel) Pb–Pb collisions at $\sqrt{s_{\text{NN}}} = 5.02$ TeV. The statistical and systematic uncertainties are depicted as vertical lines and empty boxes, respectively. The R_{AA} values are compared to the predictions of the TAMU model [126, 338]. Figure from Ref. [335].

collisions. However, for semicentral collisions, the TAMU model tends to overestimate the measured R_{AA} ratio.

The measurements of the non-prompt D_s^+ to non-prompt D^0 R_{AA} ratio suggest a potential enhancement compared to unity in $4 < p_{\text{T}} < 12$ GeV/c for central collisions, mirroring the observations made for the yield ratio also in terms of significance. In this case, the rise at low p_{T} may be attributed to the abundance of strange quarks thermally produced in the QGP medium and the dominance of the hadronisation process via recombination within this momentum range. This is consistent with the insights obtained from the B_s^0 -to- B^+ yield ratio measured by the CMS Collaboration in Pb–Pb collisions at $\sqrt{s_{\text{NN}}} = 5.02$ TeV [339] and scaled by the analogous ratio measured in pp collisions at $\sqrt{s} = 13$ TeV by the LHCb

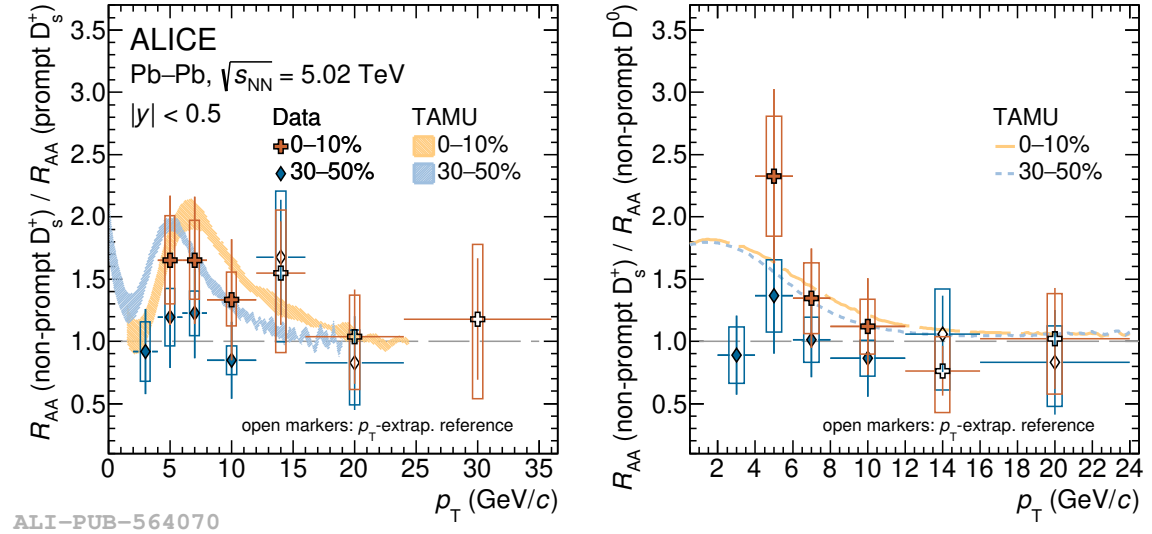


Fig. 6.14 Left: ratio of the production yield of non-prompt to prompt D_s^+ mesons in central (0 – 10%) and semicentral (30 – 50%) Pb–Pb collisions at $\sqrt{s_{NN}} = 5.02$ TeV, and pp collisions at $\sqrt{s} = 5.02$ TeV. Right: ratio of the production yield of non-prompt D_s^+ to D^0 mesons in central (0 – 10%) and semicentral (30 – 50%) Pb–Pb collisions at $\sqrt{s_{NN}} = 5.02$ TeV, and pp collisions at $\sqrt{s} = 5.02$ TeV. The statistical and systematic uncertainties are depicted as vertical lines and empty boxes, respectively. Figure from Ref. [335].

Collaboration [147]. Note that the choice of using the LHCb measurement as a reference is motivated by the fact that the ratio of the fragmentation fraction of beauty quarks into strange and non-strange mesons is not expected to depend on the energy of the collision nor the rapidity interval of the measurements as discussed in Sec. 5.4.7. The comparison between the non-prompt D_s^+ to non-prompt D_s^+ / D^0 ratio obtained in this Thesis and the B_s^0 / B^+ calculated from the measurements of the CMS and LHCb Collaborations is shown in Fig. 6.15. The R_{AA} ratio of non-prompt D_s^+ to non-prompt D^0 mesons obtained in this Thesis exceeds the B_s^0 / B^0 ratio in terms of precision. However, given the experimental uncertainties, no definitive conclusions can be drawn. These findings underscore the exciting opportunity for future research to make increasingly precise measurements in the beauty sector. This will enable us to comprehensively understand the impact of the deconfined medium on the hadronisation process of heavy quarks.

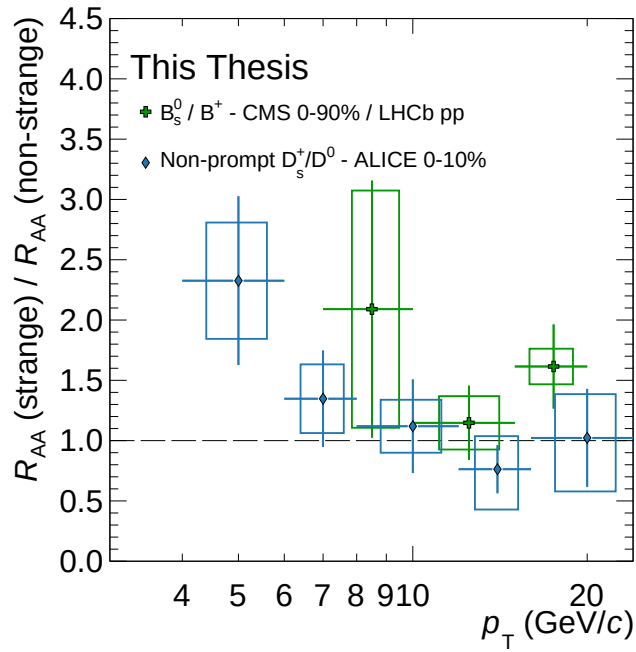


Fig. 6.15 R_{AA} of non-prompt D_s^+ to non-prompt D^0 mesons compared to the analogous ratio of B_s^0 to B^+ mesons measured by the CMS Collaboration in Pb–Pb collisions at $\sqrt{s_{\text{NN}}} = 5.02$ TeV [339] and scaled by the analogous ratio measured in pp collisions at $\sqrt{s} = 13$ TeV by the LHCb Collaboration [147]. The statistical and systematic uncertainties are depicted as vertical lines and empty boxes, respectively.

Chapter 7

Conclusions and outlook

7.1 Conclusions

In this Thesis, the measurements of the production of D_s^+ , D_{s1}^+ , and D_{s2}^{*+} mesons originating from the hadronisation of charm quarks, i.e. prompt [97], and D_s^+ mesons from the decay of beauty hadrons, i.e. non-prompt, at midrapidity ($|y| < 0.5$) in pp collisions at $\sqrt{s} = 13$ TeV and Pb–Pb collisions at $\sqrt{s_{NN}} = 5.02$ TeV [123] performed on the data collected with the ALICE experiment between 2016 and 2018 were discussed. The measurement of prompt and non-prompt D_s^+ mesons plays a key role in understanding charm- and beauty-quark production and hadronisation. In pp collisions, heavy-flavour measurements provide crucial tests for pQCD calculations and important benchmarks for Pb–Pb collisions. Instead, in heavy-ion collisions they are a powerful tool to study the in-medium energy loss of charm and beauty quarks interacting with the medium constituents and the modification of the hadronisation mechanisms of heavy quarks in a hot and dense medium, like the Quark-Gluon Plasma (QGP).

In the measurements discussed in this Thesis, the D_s^+ -meson candidates were reconstructed in the hadronic decay channel $D_s^+ \rightarrow \phi \pi \rightarrow K^- K^+ \pi^+$ exploiting the excellent spatial resolution of the ALICE detectors. The D_s^+ -meson candidates were then selected via a machine-learning multiclass classification algorithm to enhance the contribution of either prompt or non-prompt D_s^+ signals with respect to the combinatorial background and to ultimately study both charm and beauty quarks independently. This approach is more sophisticated than the traditional one based on rectangular selections, thus allowing us to extend the p_T interval of the measurements and achieve a higher precision compared to previously published results.

The measurements of the p_T -differential production cross sections of prompt and non-prompt D_s^+ mesons in pp collisions at $\sqrt{s} = 13$ TeV [97] discussed in Chapter 5 are in agreement with the predictions from pQCD calculations implementing the collinear factorisation approach, like FONLL, GM-VFNS, or the k_T factorisation. This result is an additional indication of the validity of the factorisation approach in the description of heavy-flavour meson production at LHC energies. Furthermore, the universality of the fragmentation fractions, assumed in the factorisation approach, was tested by measuring the relative production of D_s^+ mesons with respect to the one of non-strange D mesons. The measured prompt and non-prompt $D_s^+/(D^0 + D^+)$ p_T -differential ratios are compatible with previous measurements performed at lower energies and in different collision systems in agreement with the assumed universality. Conversely, this assumption is violated for baryons in both the charm and beauty sectors and is currently an open point in the description of heavy-flavour hadronisation. The total production cross section of charm and beauty quarks was measured in pp collisions at $\sqrt{s} = 13$ TeV and found to be in agreement with the predictions from pQCD calculations.

A comparison was made between the measurements reported in this Thesis in pp collisions at midrapidity and the ones performed by the LHCb Collaboration at forward rapidity [248]. The dependence of the production cross section of promptly produced strange and non-strange D mesons on the rapidity range was tested, and a larger production was observed at midrapidity with respect to forward rapidities. In addition, the mid-to-forward p_T -differential ratios measured at two different collision energies ($\sqrt{s} = 5.02$ and 13 TeV) were studied by computing a double ratio. This double ratio is of paramount importance from a theoretical point of view given that the theoretical uncertainties partially cancel out, leaving the uncertainty associated with the PDFs as the dominant one and allowing us to strongly constrain the gluon PDFs for small values of Bjorken x ($10^{-5} - 0^{-4}$). The measured double ratio is in agreement with the theoretical predictions within uncertainties, however, the latter shows larger uncertainties compared to the measurements, which indicates the possibility of further constraining the gluon PDFs at small x with future measurements.

A preliminary measurement of the production of the D_{s1}^+ and D_{s2}^{*+} excited states in pp collisions at $\sqrt{s} = 13$ TeV as a function of the average charged-particle multiplicity of the collision was performed in the context of this Thesis. The measurement is the first of its kind at LHC energies and provides an important tool to test the hadronisation mechanisms and the effect of interactions in the hadronic phase of the collision. In particular, the ratios between the D_{s1}^+ and D_{s2}^{*+} corrected yield and the one of the ground state, the D_s^+ meson, were compared with theoretical calculations based on statistical hadronisation models [11, 329]. In the case of the D_{s1}^+/D_s^+ ratio, the theoretical calculations are in agreement with the

measurements within uncertainties and no significant dependence on the charged particle multiplicity is found. Conversely, the D_{s2}^{*+}/D_s^+ ratio displays a hint of decreasing trend with increasing multiplicity. The measurement also shows a tension with the theoretical predictions of about 2.5 and 1.6σ at low and high multiplicity, respectively. This tension could be explained considering the different lifetime of the resonant states and the interaction of the decay products in a possible hadronic phase of the collision evolution, as already observed in the case of light flavour hadrons [47]. Even though no firm conclusion can be driven considering the current experimental uncertainty, this measurement provides an intriguing insight into possible effects of the hadronic phase of the collision and is the headstone for future measurements in pp and Pb–Pb collisions during the current data-taking period at the LHC, Run 3.

Finally, I presented in this Thesis the first measurement of the non-prompt D_s^+ meson production yields in central and semicentral Pb–Pb collisions at $\sqrt{s_{NN}} = 5.02$ TeV, which was published in Ref. [335]. This measurement provides an important tool to study the modification of the hadronisation mechanism of beauty quarks in the QGP and to investigate the energy-loss induced by the interaction of the beauty quarks with the QGP constituents. The ratios between the nuclear modification factor, R_{AA} , of non-prompt D_s^+ mesons and that of prompt D_s^+ mesons and non-prompt D^0 mesons were computed and compared with theoretical calculations implementing in-medium energy loss along with the modification of the hadronisation process [126], which include the possibility of beauty quarks recombining with the light quarks of the medium (coalescence). In central collisions, where the magnitude of the effects induced by the presence of the QGP is expected to be larger, the ratios show a hint of enhancement in the $4 < p_T < 12$ GeV/ c interval, with a statistical significance of 1.6 and 1.7σ , respectively. Even if the theoretical predictions qualitatively describe the trend of the data, a final conclusion cannot be drawn given the current experimental uncertainties. This measurement is a key observable to study the modification of the hadronisation mechanism of beauty quarks in the QGP and paves the way to future measurements in Pb–Pb collisions during the ongoing LHC Run 3.

7.2 Outlook

During the Long Shutdown 2 (2019 – 2021) of the LHC, various subsystems of the ALICE experiment undergone substantial upgrades [340], aimed at advancing the performance for several measurements of interest in pp and heavy-ion collisions. In particular, these upgrades focused on enhancing the precision for measuring rare QCD probes, especially in heavy-ion

collisions, and extending the measurements towards low p_T region, that is accessible only by ALICE given its low magnetic field compared to the other LHC experiments. In addition, the increased luminosity (peak values reaching $\mathcal{L} = 6 \times 10^{27} \text{ cm}^{-2}\text{s}^{-1}$ in Pb–Pb collisions)¹ and interaction rates ($\sim 50 \text{ kHz}$ in Pb–Pb collisions) expected during Run 3 and Run 4 at the LHC which allow ALICE to collect substantially larger data samples as compared to the previous Runs. The upgrades were therefore designed to obtain improved vertexing and tracking capabilities, higher data acquisition rates, and preserving the excellent PID capabilities of the experiment. These detector upgrades include a new beam pipe, a high-resolution Inner Tracking System (ITS2) [341] based on *ALPIDE* chips, the adoption of GEM detectors for the TPC readout [342], upgrades to the readout electronics of many detectors, the addition of a *Muon Forward Tracker* (MFT) [343], and a comprehensive overhaul of the trigger detectors, data acquisition, and data processing systems [344]. This holistic upgrade is designed to handle the increased data rate efficiently, crucial for managing the expected data volume during the upcoming Runs at the LHC. The upgrades involved the implementation of the Online-Offline (O^2) framework [345], which swiftly transfers data from the detector to the computing system and processes them in real-time alongside data collection. This on-the-fly processing conducts a partial calibration followed by a first reconstruction of the data, and by the replacement of the original raw data with compressed versions. This crucial step ensures that data storage remains manageable and fits within available resources, which is vital for managing the extensive data size expected during the upcoming Runs.

A preliminary observation of D_s^+ meson signals in the hadronic decay channel $D_s^+ \rightarrow \phi \pi^+ \rightarrow K^+ K^- \pi^+$ in pp collisions at $\sqrt{s} = 13.6 \text{ TeV}$ collected in Run 3 was obtained for the $2 < p_T < 3 \text{ GeV}/c$ and $3 < p_T < 4 \text{ GeV}/c$ intervals. The results, reported in Fig. 7.1, show the invariant-mass distribution of the D_s^+ meson candidates reconstructed with the O^2 framework. The results were obtained by analysing $\sim 16 \times 10^9$ pp collisions recorded with a minimum bias trigger. These results were obtained with a preliminary reconstruction including residual misalignments of the detectors, a preliminary status of the TPC calibration, and a first global reconstruction. The candidates were selected via the application of rectangular selection criteria to the kinematic, topological, and PID variables of the reconstructed candidates, described in detail in Sec. 4.2.4. In this way, it was possible to test the effectiveness of the new analysis framework before moving to more sophisticated analyses based on machine-learning algorithms. Similarly to what was observed in Run 2, the invariant-mass distribution displays a second peak at lower invariant-mass values due to the contribution of D^+ mesons decaying

¹In this scenario, the upgraded ALICE apparatus is expected to collect about $\mathcal{L} = 13 \text{ nb}^{-1}$ of Pb–Pb collisions at $\sqrt{s_{NN}} = 5.5 \text{ TeV}$ during Run 3 and Run 4 at the LHC, accompanied by a pp data sample at the same collision energy corresponding to an integrated luminosity of about 6 pb^{-1} , which may be collected within few months at an event readout rate of about 200 kHz .

in the same hadronic decay of the D_s^+ meson. The distributions are normalised to the number of entries. The invariant-mass distributions are fitted with an exponential function describing the line shape of the background and two Gaussian functions to describe the peaks of D_s^+ and D^+ mesons. These results, even if preliminary, are a strong indication of the effectiveness of the O^2 framework and of the success of the detectors upgrade.

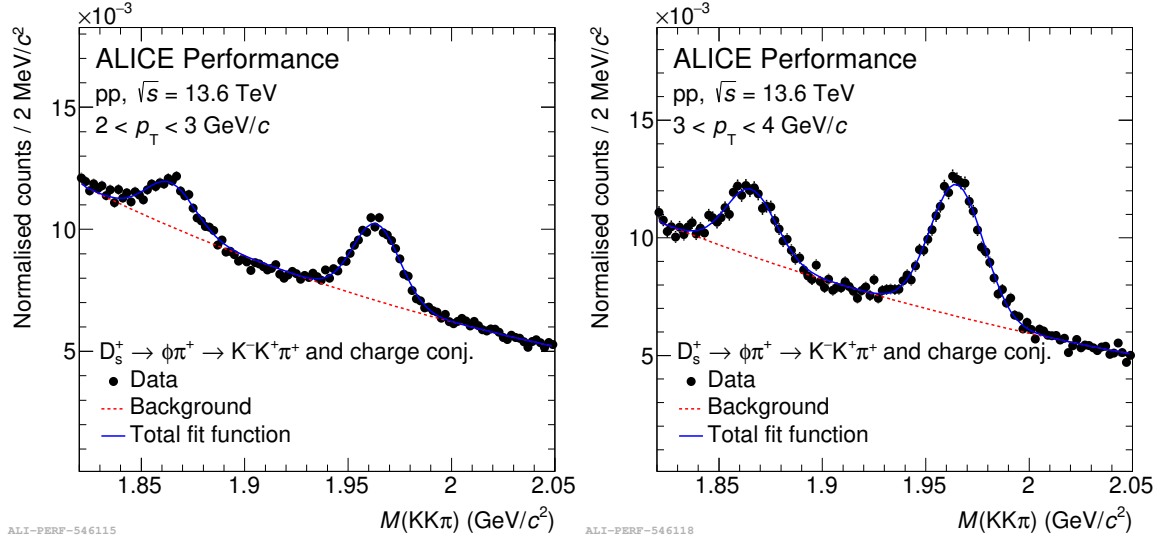


Fig. 7.1 Invariant-mass distribution of D_s^+ meson candidates in the hadronic decay channel $D_s^+ \rightarrow \phi\pi^+ \rightarrow K^+K^-\pi^+$ in pp collisions at $\sqrt{s} = 13.6$ TeV obtained by analysing Run 3 data for the $2 < p_T < 3$ GeV/c (left) and $3 < p_T < 4$ GeV/c (right) intervals. The distributions are normalised to the considered number of entries. Figure from ALICE Figure repository©.

Despite the significant performance enhancements achieved with the ALPIDE chips in terms of signal-to-noise ratio, spatial resolution, readout speed, and material budget, there is still room for further improvements with the technology used for ALPIDE, the *Monolithic Active Pixel Sensors* (MAPS) incorporating the readout electronics within the silicon active sensor. Thanks to the fast technological advancements in this field, it is now possible to produce large-scale (i.e. 14×14 cm²) thin (i.e. 20 – 40 μ m) MAPS and assemble them with a technique called *stitching*. This technological innovation enables the construction of truly cylindrical layers of silicon-only sensors. In this way, the material budget of the detector will be reduced to the bare minimum for a silicon detector ($\sim 0.05\% X_0$, where X_0 is the radiation length of the material). The ALICE experiment is currently investigating the possibility of using this technology for the construction of the three innermost layers of the new ITS detector, called ITS3, the next generation of the ITS detector for the next data-taking period at LHC, Run 4 (2028 – 2031). The ITS3 will comprise an Inner and Outer Barrel, where the latter corresponds to the existing outer barrel of the ITS2 detector. However, the Inner Barrel

of ITS3 will be entirely replaced and will consist of two half-barrels, each composed of three half-layers with a genuinely cylindrical shape, arranged around the beam pipe. In particular, the ITS3 innermost layer will be positioned at a radius of 18 mm allowing for a remarkable improvement in track reconstruction efficiency and impact parameter resolution, especially at low p_T . Comparing ITS2 and ITS3 detectors, the tracking efficiency for charged particles with p_T below 100 MeV/c increases by a factor ranging between 1.2 and 2 with the ITS3, and the impact parameter resolution at all transverse momenta improved by a factor of about 2. Simulation studies were performed to estimate the performance of heavy-flavour hadron measurements using the ITS2 and ITS3 geometries. These studies employed a detailed simulation of the ALICE apparatus with the ITS2 setup. To evaluate the performance of the ITS3 layout, a *Fast Monte Carlo Simulation* (FMCS) method called the *hybrid method* was employed. This approach adjusted the full-simulation ITS2 setup to mimic the ITS3 momentum and impact parameter resolutions by applying a scaling factor determined as the ratio of ITS2 and ITS3 expected resolutions to the track impact parameters and p_T [346, 347]

The expected performance for the measurements of non-prompt D_s^+ mesons with the ALICE ITS2 and ITS3 were compared by analysing a simulated sample of central (0 – 10%) and semicentral (30 – 50%) Pb–Pb collisions at $\sqrt{s_{NN}} = 5.5$ TeV [347]. The simulations used the HIJING 1.36 generator [173] to produce the underlying Pb–Pb events, and PYTHIA 8 for injecting $c\bar{c}$ (or $b\bar{b}$) pairs. GEANT 3 was utilised for the detector response, as described in detail in Ref. [347, 346]. The D_s^+ meson candidates were selected via BDT selection criteria after training the algorithm on a different simulated sample, similar to what was done for the analyses presented in Chapter 6. However, in this case, the instances for the background classes were taken from the simulation given the absence of data. The results reported in the left panel of Fig. 7.2 show perspectives for the measurements of the nuclear modification factor, R_{AA} , of non-prompt D_s^+ mesons as a function of p_T for the ITS2 and ITS3 geometries compared with the TAMU model predictions of the non-prompt D_s^+ and non-prompt D^0 R_{AA} . The central values of the data points correspond to the predictions by the TAMU model scaled to match the measurements discussed in this Thesis for the non-prompt D_s^+ meson. Both the original TAMU curves and those rescaled to match the Run 2 measurements for the non-prompt D^0 [347] and non-prompt D_s^+ mesons are reported. The results obtained with the ITS2 geometry are slightly shifted towards higher p_T values for visibility reasons. The precision of the measurement achieved with the ITS3 geometry is expected to be sufficient to disentangle the R_{AA} of non-prompt D_s^+ and non-prompt D^0 mesons according to the predictions by TAMU.

The right panel of Fig. 7.2 shows the expected precision of the measurement of the elliptic flow, v_2 , of non-prompt D_s^+ mesons with the ITS2 and ITS3 geometries as a function of

p_T . The predictions of prompt and non-prompt D_s^+ meson v_2 from the TAMU model are also reported. For the elliptic flow, the precision expected with the ITS2 detector does not guarantee discrimination between the expected v_2 of prompt and non-prompt D_s^+ mesons. Instead, with the ITS3 detector the precision would be sufficient to appreciate this difference down to $p_T = 1$ GeV/c thanks to a reduction of both statistical and systematic uncertainties.

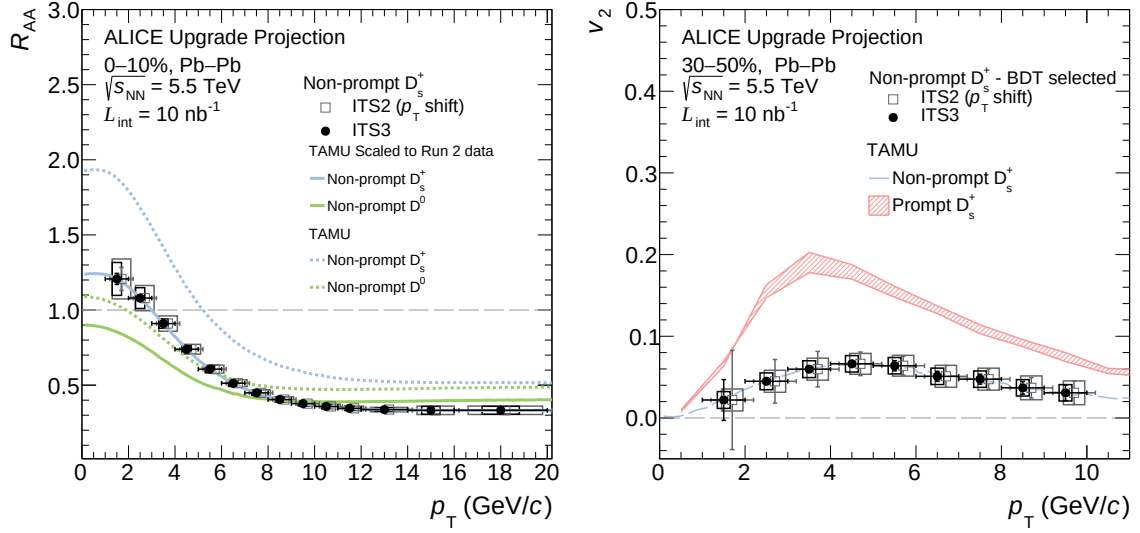


Fig. 7.2 Left: expected nuclear modification factor of non-prompt D_s^+ mesons as a function of p_T , compared to the predictions for non-prompt D_s^+ and D^0 mesons from the TAMU model. Right: the expected elliptic flow of non-prompt D_s^+ mesons as a function of p_T , compared to the predictions for prompt and non-prompt D_s^+ mesons from the TAMU model. In both panels, the vertical bars represent the expected statistical uncertainty on the measurements, while the empty boxes represent the expected systematic uncertainty. Figure from Ref. [347].

Acknowledgements

This Thesis, like any other achievement, is the result of the hard work of many people. I would like to express my gratitude to all of them.

Stefania and Francesco, my supervisors, thank you for giving me the opportunity of the PhD in the first place, for your continuous support and encouragement, and for the time you subtracted from your busy schedule to help me both physics-wise and not. The work we did in these years has been very fruitful, and I look forward to the many exciting projects we will work on in the future.

Fabrizio and Fabio, aka the DmesonAnalysers. Thank you for sharing your knowledge, time, and experience with me. I greatly benefited from our discussions and the time we had together. A special "thank you" goes to Fabrizio, my HF guide. The many discussions we had, and the time you spent reviewing my work, advising me, and always being there for me, were invaluable. To both of you, I am not sure I would have made it without you.

I want to express my gratitude to the ALICE group in Torino. I found a great environment here in Turin and in the "Stanzone" (the ALICE group office) thanks to all of you. A thought goes to Mario, Francesco, Alessandro, and Giacomo with whom I shared many special moments at the conferences we attended. Some of the best moments of my PhD were spent with you. I want to thank the many people I met at CERN during my year abroad. Not only did I have the fantastic opportunity to work side-by-side with some of you in a very exciting environment, but I also had the pleasure of sharing many moments that I will never forget. Especially the many dinners we had (mostly Luigia's), all the coffee breaks at CERN, the numerous discussions we had, and the laughs we shared. Next time I will be at CERN, I will make sure to visit you all.

I want to thank all the people from the D2H PAG and the PWG HF I had the pleasure to work with. They spent a lot of time reviewing, discussing, and improving my work, and, thanks to them, I had the opportunity to present my work on many occasions inside and outside the Collaboration. Personally, this was one of the most rewarding experiences I had

in my PhD. In particular, I want to thank Mattia and Cristina. Thank you guys for your constant support and encouragement and for the many hilarious moments we shared.

I want to thank my family and my friends for being there for me, always. I would not have made it without you. I want to thank my parents, Luisa and Aurelio, and my brother and his wife, Andrea and Michelle, for their unconditional love and support. I am very grateful to them for the many sacrifices they made for me. Thanks to my friends from my hometown, Savigliano, Mattia, Giacomo, Luca, Alex, Davide, and Chiaretta, for being such a great companionship, and to my friends from Florence, Saverio (even if he is an LHCb enthusiast) and Maddalena, for your support and the great moments we shared.

Finally, I want to thank Irene for her love and support. Thank you for always being by my side, for the laughs, and for the many adventures we had together all around the world. I am very grateful to you for the many sacrifices you made for me. I know I can be a pain in the neck sometimes, but I am very lucky to have you. We made it, Irene. I am very proud of us, and I am looking forward to the many adventures we will have in the future.

List of Tables

2.1	Main properties of the ITS layers [159].	51
2.2	Geometric properties of V0A and V0C [164].	55
3.1	Confusion matrix for a binary classification problem.	73
4.1	Summary of the data samples used for the analyses presented in this Thesis. The values of the integrated luminosity ($\mathcal{L}_{\text{int}}$) are calculated using the MB trigger cross sections taken from Refs. [232, 233]. †: due to the <i>trigger turn-on curves</i> [234], the calculation of the integrated luminosity is not straightforward. The values reported here are computed considering only the data periods for which the triggers are fully efficient in the multiplicity region of interest.	89
4.2	Main characteristics of the charm and beauty hadrons of interest. The branching ratios (BR) are taken from Ref. [6]. (†: the BR of the hadronic decay channels of the D-meson resonances are not measured, they are calculated at the analyses level as described in Chapter 5. *: the BR of the hadronic decay channels of the B-meson is different for particle and antiparticle. For antiparticle are reported in parentheses.)	90
5.1	Topological and kinematic pre-selections, applied to compose the training sample, for the D_s^+ analysis in pp collisions at $\sqrt{s} = 13$ TeV.	112
5.2	Number of signal (prompt and non-prompt D_s^+) and background candidates used to build the training and test sets.	112
5.3	Hyperparameter configuration for the p_T bins considered in the training for the D_s^+ mesons.	115

5.4	Fragmentation fractions of beauty quarks into beauty hadrons measured in $Z^0 \rightarrow b\bar{b}$ decays [274] and in pp collisions at $\sqrt{s} = 1.96$ TeV [6].	118
5.5	Raw yields, signal-to-background ratios, and significance of prompt D_s^+ mesons in pp collisions at $\sqrt{s} = 13$ TeV. A 3σ interval around the D_s^+ peak is used to extract the values.	121
5.6	Raw yields, signal-to-background ratios, and significance of non-prompt D_s^+ mesons in pp collisions at $\sqrt{s} = 13$ TeV. A 3σ interval around the D_s^+ peak is used to extract the values.	121
5.7	Relative systematic uncertainties of the measured cross section of prompt D_s^+ mesons at midrapidity in pp collisions at $\sqrt{s} = 13$ TeV. Uncertainties found to be negligible are indicated as such (negl.).	131
5.8	Relative systematic uncertainties of the measured cross section of non-prompt D_s^+ mesons at midrapidity in pp collisions at $\sqrt{s} = 13$ TeV. Uncertainties found to be negligible are indicated as such (negl.).	132
5.9	$d\sigma/dy _{ y <0.5}$ of D^0 , D^+ , and D_s^+ in pp collisions at $\sqrt{s} = 13$ TeV [97]. These results are obtained by integrating the measured p_T -differential cross sections at midrapidity and extrapolating down to $p_T = 0$ where necessary.	145
5.10	Ratios of the measured production cross sections of prompt D mesons for $p_T > 0$ at midrapidity ($ y < 0.5$) in pp collisions at $\sqrt{s} = 5.02$ TeV [239] and $\sqrt{s} = 13$ TeV [97].	146
5.11	Total production cross section of $c\bar{c}$ pairs at midrapidity ($ y < 0.5$) in pp collisions at $\sqrt{s} = 13$ TeV [97] and $\sqrt{s} = 5.02$ TeV [97]. The results are obtained by integrating the measured p_T -differential cross section at midrapidity and extrapolating down to $p_T = 0$ where necessary.	154
5.12	Systematic uncertainties for the measurement of prompt D_{s1}^+ and D_{s2}^{*+} corrected yield.	172
6.1	Topological and kinematic pre-selections applied for the D_s^+ analysis in 0–10% centrality class.	183
6.2	Number of signal (prompt and non-prompt D_s^+) and background candidates used to build the training and test sets for the 0–10% and 30–50% centrality classes.	184

6.3	Hyperparameter configuration for the p_T bins considered in the training for the D_s^+ mesons in the 0 – 10% and 30 – 50% centrality class, respectively. .	186
6.4	Raw yields, signal-to-background ratios, and significance of non-prompt D_s^+ mesons in pp collisions at $\sqrt{s} = 13$ TeV.	188
6.5	Relative systematic uncertainties of the measured non-prompt D_s^+ -meson corrected yields and R_{AA} in central (0 – 10%) Pb–Pb collisions at $\sqrt{s_{NN}} = 5.02$ TeV in the transverse momentum-intervals of the analysis. Uncertainties found to be negligible are indicated as such (negl.).	197
6.6	Relative systematic uncertainties of the measured non-prompt D_s^+ -meson corrected yields and R_{AA} in semicentral (30 – 50%) Pb–Pb collisions at $\sqrt{s_{NN}} = 5.02$ TeV in the transverse momentum-intervals of the analysis. Uncertainties found to be negligible are indicated as such (negl.).	198

List of Figures

1.1	Quark-gluon and gluon-only vertices in QCD.	3
1.2	Summary of measurements of α_s as a function of the momentum transfer $\mathcal{Q}$. The respective degree of QCD perturbation theory used in the extraction of α_s is indicated in brackets (NLO: next- to-leading order; NNLO: next-to-next- to-leading order; NNLO+res.: NNLO matched to a resummed calculation; N3LO: next-to-NNLO). Figure from Ref. [6].	4
1.3	Layout of the QCD phase diagram. Figure from Ref. [13].	5
1.4	Space-time diagram of longitudinal QGP evolution. Figure from Ref. [18]. .	7
1.5	Space-time evolution of a heavy ion collision. Figure from Ref. [13].	8
1.6	Schematic representation of two colliding nuclei in the Glauber model de- scription. The coordinate system for the nucleons is reported. Figure from Ref. [22].	13
1.7	Initial energy density in the transverse plane in Glauber Monte Carlo simula- tion. Figure from Ref. [24].	14
1.8	Charged particle multiplicity at midrapidity as a function of the center-of- mass energy per nucleon pair $\sqrt{s_{NN}}$ (left) and the number of participant pairs $\langle N_{part} \rangle$ (right) for various colliding systems. Figures from Ref. [25].	16
1.9	Simultaneous thermal fit to the identified-particle yields measured by the ALICE Collaboration in the 10% most central Pb–Pb collisions at $\sqrt{s_{NN}} =$ 5.02 TeV. The results obtained with the THERMUS [32], GSI [33], and SHARE [34] models are reported. Figure from ALICE Figure repository ©.	18

- 1.10 Sum of particle-antiparticle production yields relative to that of $(\pi^+ + \pi^-)$ for non-strange ($p + \bar{p}$), strange ($2K_S^0, \Lambda + \bar{\Lambda}$), and multi-strange ($2\phi, \Xi^- + \bar{\Xi}^+$, and $\Omega^- + \bar{\Omega}^+$) hadrons as a function of the average charged particle multiplicity at midrapidity, measured by the ALICE Collaboration across various colliding system sizes and energy scales. Figure from the ALICE Figure repository ©. 19
- 1.11 Average expansion velocity and kinetic freeze-out temperature (T_{kin}) obtained from the simultaneous Boltzmann-Gibbs Blast-Wave fit to pions, kaons, and protons spectra measured in Pb–Pb collisions at $\sqrt{s_{\text{NN}}} = 5.02$ and 2.76 TeV. The correlated 1σ uncertainties, comprehensive of statistical and systematic uncertainty, are shown as ellipses. Figure from Ref. [50] 22
- 1.12 Schematic view of a peripheral heavy-ion collision. 23
- 1.13 Elliptic flow of identified charged particles as a function of p_T in Pb–Pb collisions at $\sqrt{s_{\text{NN}}} = 5.02$ TeV for different centrality intervals. The error bars and boxes represent the statistical and systematic uncertainties, respectively. Figure from [58]. 24
- 1.14 Schematic representation of the di-jet production and jet quenching in heavy-ion collisions. Figure from Ref. [60]. 26
- 1.15 Top: Two-particle azimuthal correlations in pp and d–Au (minimum bias and central) collisions at $\sqrt{s_{\text{NN}}} = 200$ GeV measured by the STAR Collaboration. Bottom: Comparison of pedestal-subtracted yields in pp, d–Au, and Au–Au. Figure from Ref. [59]. 26
- 1.16 Left: charged particle R_{AA} as a function of p_T measured by the ALICE [61] and CMS [62, 63] Collaborations in p–Pb and central (0–5%) Pb–Pb collisions at $\sqrt{s_{\text{NN}}} = 5.02$ TeV. Figure from Ref. [61]. Right: compilation of charged jet R_{AA} in central Pb–Pb collisions measured by the ALICE [67], ATLAS [68–70], and CMS [71, 72] Collaborations. Figure from Ref. [67]. . . 27

- 1.17 Left: Measured R_{AA} for Υ states as functions of number of participant nucleons $\langle N_{\text{part}} \rangle$ in Pb–Pb collisions at $\sqrt{s_{NN}} = 5.02$ TeV including the centrality-integrated bin (on the right). The vertical lines correspond to statistical uncertainties, while the boxes represent the systematic uncertainties. Figure from Ref. [75]. Right: J/ψ -meson R_{AA} as a function of $\langle N_{\text{part}} \rangle$ in Pb–Pb collisions at $\sqrt{s_{NN}} = 2.76$ TeV and in Au–Au collisions at $\sqrt{s_{NN}} = 200$ GeV. The vertical bars represent the statistical uncertainties, while the boxes represent the systematic uncertainties. Figure from Ref. [76]. 29
- 1.18 Feynman diagrams of leading order quark-antiquark annihilation and gluon fusion for the production of $Q\bar{Q}$ pairs. 30
- 1.19 Left: Comparison of the FONLL, GM-VFNS, and k_T -factorisation calculations with the measured p_T -differential production cross section of prompt D^+ mesons in pp collisions at $\sqrt{s} = 13$ TeV. Figure from Ref. [97]. Right: Comparison of the FONLL and TAMU calculations with the measured p_T -differential production cross section of non-prompt D^+ mesons in pp collisions at $\sqrt{s} = 13$ TeV. Figure from ALICE Figure repository ©. 33
- 1.20 Illustration of the modification of the nucleon PDFs in a nucleus observed in deep inelastic scattering experiments. Figure from Ref. [101]. 34
- 1.21 Collisional and radiative energy loss of charm (left) and beauty (right) quarks in the QGP as a function of their initial energy in central Pb–Pb collisions at $\sqrt{s_{NN}} = 2.76$ TeV, as predicted by the model outlined in Ref. [113]. Figure from Ref. [113]. 35
- 1.22 Left: average R_{AA} of prompt D^0 , D^+ , and D^{*+} mesons as a function of p_T in central (0 – 10%), semi-central (30 – 50%), and peripheral (60 – 80%) Pb–Pb collisions, and p–Pb collisions at $\sqrt{s_{NN}} = 5.02$ TeV. Figure from Ref. [117]. Left: R_{AA} of prompt and non-prompt D^0 mesons in central (0 – 10%) Pb–Pb collisions at $\sqrt{s_{NN}} = 5.02$ TeV as a function of p_T . Figure from Ref. [123]. 39
- 1.23 Left: Average R_{AA} of prompt D^0 , D^+ , and D^{*+} mesons as a function of p_T in central (0 – 10%) Pb–Pb collisions at $\sqrt{s_{NN}} = 5.02$ TeV compared with different transport models. Figure from Ref. [117]. Right: Average v_2 of prompt D^0 , D^+ , and D^{*+} mesons as a function of p_T in semicentral (30 – 50%) Pb–Pb collisions at $\sqrt{s_{NN}} = 5.02$ TeV compared with different transport models. Figure from Ref. [117]. 40

1.24	Elliptic, v_2 (upper panels), and the triangular, v_3 (lower panels), flow of prompt [134] and non-prompt [133] D^0 mesons as functions of p_T in central (0 – 10%), semi-central (10 – 30% and 30 – 50%) Pb–Pb collisions at $\sqrt{s_{NN}} = 5.02$ TeV measured by the CMS Collaboration. Figure from Ref. [133].	41
1.25	Left: Average v_2 of prompt D^0 , D^+ , and D^{*+} mesons as a function of p_T in central Pb–Pb collisions at $\sqrt{s_{NN}} = 5.02$ TeV compared with different transport models. Figure from Ref. [142]. Right: Perspectives v_2 measurements of non-prompt D^0 , non-prompt J/ψ , and $B^+(\rightarrow D^0)$ mesons by the ALICE experiment in Pb–Pb collisions at $\sqrt{s_{NN}} = 5.5$ TeV in Run 3. Figure from Ref. [141].	42
2.1	The LHC injection scheme. Figure from Ref. [156].	47
2.2	Top: Lateral view of the ALICE experiment. Figure from the ALICE figure repository ©. Bottom: Frontal view of the ALICE experiment. Figure from Ref. [158]	49
2.3	Layout of the ALICE ITS. Figure from Ref. [160].	52
2.4	Left: Layout of the ALICE TPC. Figure from Ref. [161]. Right: Azimuthal section of TPC detector. Figure from Ref. [162].	53
2.5	Left: Layout of the ALICE TOF detector. Figure from Ref. [163].	54
2.6	Layout of the ALICE V0 detector plates. Figure from Ref. [157].	55
2.7	Schematic representation of WLCG hierarchy. Figure from Ref. [168]. . . .	57
2.8	Distribution of the sum of signals registered by the V0 scintillators of the ALICE experiment in Pb–Pb collisions at $\sqrt{s_{NN}} = 2.76$ TeV. Figure from Ref. [166].	60
2.9	Workflow of the track reconstruction in ALICE. Figure from Ref. [179]. . .	61
2.10	Left: track-reconstruction efficiency in the TPC as a function of p_T in pp collisions at $\sqrt{s} = 8$ TeV. Figure from Ref. [178]. Right: prolongation efficiency of tracks from the TPC to the ITS as a function of p_T in Pb–Pb collisions at $\sqrt{s_{NN}} = 2.76$ TeV. Figure from Ref. [178].	62

- 2.11 Comparison of the width of the distribution of primary vertex positions reconstructed in the transverse plane from global tracks (named as TRK) and SPD tracklets (named as SPD) as a function of the charged-particle multiplicity per unit of pseudorapidity in pp collisions at $\sqrt{s} = 7$ TeV. Figure from Ref. [178]. 63
- 2.12 Left: resolution of the projection in the transverse plane of the impact parameter to the primary vertex for identified ITS-TPC tracks in pp, p-Pb, and Pb-Pb collisions at $\sqrt{s} = 7$, $\sqrt{s_{NN}} = 5.02$, and $\sqrt{s_{NN}} = 2.76$ TeV, respectively. Right: $1/p_T$ resolution for TPC-standalone and ITS-TPC-combined tracks, together with the case of constrained primary vertex, in p-Pb collisions at $\sqrt{s_{NN}} = 5.02$ TeV. Both figures are taken from Ref. [178]. 64
- 2.13 Specific energy-loss measured by the ITS (left) and the TPC (right) as a function of the momentum (p) and rigidity (p/z) in Pb-Pb collisions at $\sqrt{s_{NN}} = 5.02$ TeV. The central values of the expected specific energy-loss for each particle species (i.e. electrons, pions, kaons, protons, deuterons, tritons, and ^3He) are shown as solid lines. Figures from the ALICE Figure repository ©. 67
- 2.14 The β parameter, measured by the TOF, as a function of p , measured by the TPC, in Pb-Pb collisions at $\sqrt{s_{NN}} = 5.02$ TeV. Figure from the ALICE Figure repository ©. 68
- 3.1 Output score distribution for the signal (green) and background (blue) classes in the context of the D_s^+ meson production in pp collisions at midrapidity at $\sqrt{s} = 13$ TeV. The $2 < p_T < 4$ GeV/ c interval is considered. 72
- 3.2 Left: ROC curve of the binary-classification algorithm employed in the study of the production of the D_s^+ mesons in pp collisions at $\sqrt{s} = 13$ TeV during the training and test phase. The $2 < p_T < 4$ GeV/ c interval is considered. The ROC curve of the random classifier is also shown for comparison. Right: ROC curve of the multiclass classification algorithm employed in the study of the production of the D_s^+ meson in pp collisions at $\sqrt{s} = 13$ TeV. The $2 < p_T < 4$ GeV/ c interval is considered. 75

3.3	Hyperparameters optimisation for the grid search (left-panel), the random search (central-panel), and the Bayesian optimisation (right-panel). The dependence of the loss function on the two hyperparameters is shown on the top and on the left of each plot. The colour gradient of the points, from black to “yellowish” to orange indicates the order of the search. Figure from Ref. [198].	77
3.4	SHAP feature importance plot for the D_s^+ meson analyses in pp collisions at $\sqrt{s} = 13$ TeV. The $2 < p_T < 4$ GeV/c interval is considered.	79
3.5	Schematic representation of a BDT. Figure from Ref. [209].	81
4.1	Sketch of the D_s^+ meson decay topology in the case of prompt (left) and non-prompt (right) D_s^+ mesons. Relative distances and angles on the figure are purely schematic.	91
4.2	Rapidity distribution versus transverse momentum for D_s^+ mesons reconstructed in pp collisions at $\sqrt{s} = 13$ TeV. The fiducial acceptance region is defined by $ y < y_{\text{fid}}(p_T)$	95
4.3	Decay length (left panel) and normalised decay length in the transverse plane (right panel) distributions for D_s^+ mesons in pp collisions at $\sqrt{s} = 13$ TeV in $2 < p_T < 4$ GeV/c. The distributions are shown for prompt D_s^+ mesons (red), non-prompt D_s^+ mesons (blue), and combinatorial background (green). For the prompt and non-prompt D_s^+ mesons, the distributions are taken from Monte Carlo simulations, while for the combinatorial background they are taken from data excluding the signal region in the invariant mass distribution.	97
4.4	Distributions of the cosine of the pointing angle (left panel) and track dispersion at the decay vertex (right panel) for D_s^+ -meson candidates in pp collisions at $\sqrt{s} = 13$ TeV in $2 < p_T < 4$ GeV/c. The distributions are shown for prompt D_s^+ mesons (red), non-prompt D_s^+ mesons (blue), and combinatorial background (green). For the prompt and non-prompt D_s^+ mesons, the distributions are taken from Monte Carlo simulations, while for the combinatorial background they are taken from data excluding the signal region in the invariant mass distribution.	98

- 4.5 Impact parameter in the transverse plane (left) and D_0^{xy} (right) distributions for D_s^+ mesons in pp collisions at $\sqrt{s} = 13$ TeV in $2 < p_T < 4$ GeV/c. The distributions are shown for prompt D_s^+ mesons (red), non-prompt D_s^+ mesons (blue), and combinatorial background (green). For the prompt and non-prompt D_s^+ mesons, the distributions are taken from Monte Carlo simulations, while for the combinatorial background they are taken from data excluding the signal region in the invariant mass distribution. 100
- 4.6 Distributions of the absolute difference between the K^+K^- invariant mass and the PDG world average for the ϕ meson (left) and $\cos^3(\theta'(K))$ (right) for D_s^+ mesons in pp collisions at $\sqrt{s} = 13$ TeV in $2 < p_T < 4$ GeV/c. The distributions are shown for prompt D_s^+ mesons (red), non-prompt D_s^+ mesons (blue), and combinatorial background (green). For the prompt and non-prompt D_s^+ mesons, the distributions are taken from Monte Carlo simulations, while for the combinatorial background they are taken from data excluding the signal region in the invariant mass distribution. 102
- 5.1 Predicted spectroscopy of the strange D-meson states compared to the experimental measurements. Figure from Ref. [268]. 110
- 5.2 Correlation matrices among the training variables and the D_s^+ -meson invariant mass and transverse momentum for the prompt D_s^+ (top-left), non-prompt D_s^+ (top-right), and the combinatorial background (bottom-left) examples in the $2 < p_T < 4$ GeV/c interval in the training sample. 114
- 5.3 ROC AUC obtained for the prompt D_s^+ meson test set in the $4 < p_T < 6$ GeV/c interval. 115
- 5.4 Distribution of the BDT output scores obtained in the D_s^+ meson training sample in the $2 < p_T < 4$ GeV/c interval for the non-prompt D_s^+ (top-left), prompt D_s^+ (top-right), and background (bottom-left) hypotheses. 117
- 5.5 Expected significance (top-left), expected signal-to-background ratio (top-right), and non-prompt fraction (bottom) as a function of the BDT output score thresholds for the probability to be prompt and background in the $2 < p_T < 4$ GeV/c interval. 120

- 5.6 Invariant-mass distributions of prompt D_s^+ -meson candidates in pp collisions at $\sqrt{s} = 13$ TeV in the nine p_T intervals of the analysis. The blue solid line represents the result of the fit to the data. The orange dashed line represents the background component of the fit. 124
- 5.7 Invariant-mass distributions of non-prompt D_s^+ -meson candidates in pp collisions at $\sqrt{s} = 13$ TeV in the six p_T intervals of the analysis. The blue solid line represents the result of the fit to the data. The orange dashed line represents the background component of the fit. 125
- 5.8 Acceptance-times-efficiency correction factors for prompt and non-prompt D_s^+ mesons as functions of p_T , inclusive of the rapidity-coverage correction, for the prompt-(left) and non-prompt-enhanced (right) sample. 126
- 5.9 Top-left: prompt and non-prompt ($\text{Acc} \times \epsilon$) factors as a function of the BDT selection criteria. Top-right: correlation matrix of the BDT selections expressed in terms of the correlation coefficient ρ . Bottom-left: contribution of the prompt and non-prompt component to the extracted raw-yields as a function of the BDT selection criteria. Bottom-right: prompt and non-prompt D_s^+ meson fraction as a function of the BDT selection criteria. The results are obtained in the $2 < p_T < 4$ GeV/c interval in the non-prompt-enhanced sample in pp collisions at $\sqrt{s} = 13$ TeV. 130
- 5.10 Distribution of the raw-yield values obtained with the different fit configurations for the prompt- (left) and non-prompt-enhanced (right) samples in pp collisions at $\sqrt{s} = 13$ TeV in the $2 < p_T < 3$ and $2 < p_T < 4$ GeV/c intervals, respectively. 133
- 5.11 Left: Relative systematic uncertainty on the track reconstruction efficiency as a function of p_T in pp collisions at $\sqrt{s} = 13$ TeV for the prompt-enhanced sample. Right: Relative systematic uncertainty on the non-prompt D_s^+ meson fraction as a function of p_T in pp collisions at $\sqrt{s} = 13$ TeV for the non-prompt-enhanced sample. 135
- 5.12 Distribution obtained from the variation of the BDT selections with respect to the reference one for the statistical significance (top-left panel), the ratio of the efficiency-times-acceptance factor (top-right panel), the ratio of the (non-)prompt fraction (bottom-left panel), and the corrected yield (bottom-right panel). The distributions refer to the prompt-enhanced sample in pp collisions at $\sqrt{s} = 13$ TeV in the $2 < p_T < 3$ GeV/c interval. 137

- 5.13 Left: Ratio of the p_T spectra for prompt D_s^+ and the beauty hadrons between FONLL and PYTHIA 8. Right: Ratio between the $(\text{Acc} \times \varepsilon)$ correction factor obtained with the FONLL p_T weights over the reference for the prompt- and non-prompt-enhanced sample in pp collisions at $\sqrt{s} = 13$ TeV together with the assigned systematic. 138
- 5.14 The p_T -differential production cross sections of prompt D^0 , D^+ , D^{*+} , and D_s^+ mesons at midrapidity ($|y| < 0.5$) in pp collisions at a centre-of-mass energy of $\sqrt{s} = 13$ TeV. The vertical bars and boxes report the statistical and systematic uncertainties, respectively. The systematic uncertainties reported in the plots do not include the contributions of the luminosity and the branching ratio, which are reported separately. The measured prompt D^{*+} -meson production cross section was scaled by a factor of 1/5 for better visibility reasons. Figure from Ref. [97]. 139
- 5.15 p_T -differential production cross sections for prompt D mesons in comparison with pQCD calculations: FONLL [86, 87] (except for D_s^+ mesons), GM-VFNS framework [88, 89], and a calculation based on the k_T -factorisation approach [92, 246]. The theoretical uncertainties for these predictions are shown as shaded boxes. The vertical bars and boxes report the statistical and systematic uncertainties, respectively. The systematic uncertainties shown in the plots do not include the contributions of the luminosity and the branching ratio, which are reported separately. Figure from Ref. [97]. 141
- 5.16 Ratios of production cross sections as a function of p_T of prompt D^+/D^0 , D^{*+}/D^0 , D_s^+/D^0 , D_s^+/D^+ , and $D_s^+/(D^0 + D^+)$ mesons at midrapidity ($|y| < 0.5$) in pp collisions at $\sqrt{s} = 13$ TeV [97], $\sqrt{s} = 7$ TeV [182], and $\sqrt{s} = 5.02$ TeV [239]. The vertical bars and boxes report the statistical and systematic uncertainties, respectively. Figure from Ref. [97]. 143
- 5.17 Ratios of production cross sections as a function of p_T of prompt D^0 , D^+ , D^{*+} , and D_s^+ mesons at midrapidity ($|y| < 0.5$) in pp collisions at $\sqrt{s} = 13$ TeV [97] to those at $\sqrt{s} = 5.02$ TeV [239]. The vertical bars and boxes report the statistical and systematic uncertainties, respectively. Figure from Ref. [97]. 144
- 5.18 Ratio of the p_T -differential production cross sections of prompt D^0 mesons at midrapidity ($|y| < 0.5$) in pp collisions at $\sqrt{s} = 13$ TeV [97] to those at $\sqrt{s} = 5.02$ TeV [239] compared to the FONLL prediction [86, 87]. 145

- 5.19 Compilation of the $f_s/(f_u + f_d)$ factor measured in different collision systems and energies by ALICE [239, 97, 182], H1 [278], ZEUS [279], ATLAS [280], and LEP [277] Collaborations. The horizontal bars indicate the total experimental uncertainty, while the shaded boxes represent the theoretical systematic uncertainty. For the measurement of H1, the theoretical uncertainty encompasses the uncertainty on the branching ratio and the model dependencies on the acceptance determination, while for the ATLAS measurement, it consists of the extrapolation uncertainty to the full phase space. Figure from Ref. [97]. 147
- 5.20 Ratios of the production cross sections as a function of p_T of prompt D^0 , D^+ , D^{*+} , and D_s^+ mesons measured at midrapidity ($|y| < 0.5$) by ALICE [97] to the one measured by LHCb [248] in three rapidity intervals at forward rapidity ($2 < y < 2.5$, $3 < y < 3.5$, and $4 < y < 4.5$) in pp collisions at $\sqrt{s} = 13$ TeV. Figure from Ref. [97]. 148
- 5.21 Ratio of the p_T -differential production cross sections of prompt D^0 mesons at midrapidity ($|y| < 0.5$) measured by the ALICE Collaboration [97, 181] to the one measured by the LHCb Collaboration [147, 245] in three rapidity intervals at forward rapidity (left panel: $2 < y < 2.5$, middle panel: $3 < y < 3.5$, right panel: $4 < y < 4.5$) in pp collisions at $\sqrt{s} = 13$ TeV and $\sqrt{s} = 5.02$ TeV. The double ratio ρ is compared to the predictions based on FONLL calculations employing the CTEQ6.6 [81] and NNPDF30 [285] PDF sets in the middle and bottom rows, respectively. Figure from Ref. [97]. 151
- 5.22 Left: Charm-quark fragmentation fractions measured in pp collisions at $\sqrt{s} = 13$ TeV [97] and $\sqrt{s} = 5.02$ TeV [97] compared to the measurement performed in e^+e^- and ep collisions [287]. Right: Total production cross section of $c\bar{c}$ pairs at midrapidity ($|y| < 0.5$) as a function of $\sqrt{s}$ measured by the ALICE Collaboration at $\sqrt{s} = 5.02$ TeV, $\sqrt{s} = 7$ TeV, and $\sqrt{s} = 13$ TeV [291, 97, 182], and by the STAR [292] and PHENIX [293] Collaborations at $\sqrt{s} = 200$ GeV. Figure from Ref. [97]. 152
- 5.23 Left: p_T -differential production cross sections of prompt [97] and non-prompt D^0 , D^+ , and D_s^+ mesons at midrapidity ($|y| < 0.5$) in pp collisions at $\sqrt{s} = 13$ TeV. Right: ratios between the production cross sections of non-prompt and prompt D^0 (top), D^+ (middle), and D_s^+ (bottom) mesons at $\sqrt{s} = 5.02$ TeV [239] and $\sqrt{s} = 13$ TeV. The vertical bars and boxes represent the statistical and systematic uncertainties, respectively. 155

- 5.24 p_T -differential production cross sections of non-prompt D_s^+ mesons at midrapidity ($|y| < 0.5$) in pp collisions at $\sqrt{s} = 13$ TeV compared with the predictions from FONLL + PYTHIA 8 [86, 87], GM-VFNS [89, 88], and TAMU + PYTHIA 8 [98] calculations. The lower panels report the data-to-model ratios. The vertical bars and boxes represent the statistical and systematic uncertainties, respectively. 157
- 5.25 Left: prompt and non-prompt $D_s^+/(D^+ + D^0)$ ratios as a function of p_T in pp collisions at $\sqrt{s} = 5.02$ [239] and $\sqrt{s} = 13$ [97] TeV. Right: prompt [97] and non-prompt $D_s^+/(D^+ + D^0)$ ratios as a function of p_T in pp collisions at $\sqrt{s} = 13$ TeV compared with FONLL+PYTHIA 8 predictions. The vertical bars and empty boxes report the statistical and systematic uncertainties, respectively. 158
- 5.26 Ratios of non-prompt D-meson production cross sections at $\sqrt{s} = 13$ and $\sqrt{s} = 5.02$ TeV [239] as a function of p_T . Left: comparison of measurements for D^0 , D^+ , and D_s^+ mesons. Right: comparison of D^0 results with FONLL+PYTHIA 8 calculations. The vertical bars and boxes report the statistical and systematic uncertainties, respectively. 159
- 5.27 Beauty-quark fragmentation-fraction ratio $f_s/(f_u + f_d)$ calculated from non-prompt D-meson measurements in pp collisions at $\sqrt{s} = 13$ TeV compared with previous measurements from CDF [306], ALICE [239], ATLAS [307], CMS [310], and LHCb [308, 147], with the average of LEP measurements [309], and with the results of PYTHIA 8 simulations with Monash-13 tune [244]. 161
- 5.28 Left: beauty-quark production cross section per unit of rapidity at midrapidity, as a function of the centre-of-mass energy measured in pp and $p\bar{p}$ collisions by the ALICE [312–317], PHENIX [318], CDF [319] and UA1 [320] Collaborations. The solid and dashed lines, accompanied by a shaded band, represent the central values and the associated uncertainties predicted by FONLL [86, 323, 87] and NNLO [324] calculations, respectively. Right: beauty-quark production cross section per unit of (pseudo)rapidity as a function of (pseudo)rapidity measured by the ALICE and LHCb [322] Collaborations. The solid and dashed lines, accompanied by a shaded band, represent the central values and the associated uncertainties predicted by FONLL [86, 323, 87]. The vertical bars and boxes report the statistical and systematic uncertainties, respectively. 163

- 5.29 Left: correlation between the invariant-mass distribution of the selected K_S^0 candidates ($M(\pi\pi)$) and those of selected D^+ ($M(K\pi\pi)$) candidates used to build D_{s2}^{*+} candidates in the MB sample. Right: correlation between the invariant-mass distribution of the selected K_S^0 candidates ($M(\pi\pi)$) and those of selected D^{*+} ($\Delta M = M(K\pi\pi) - M(K\pi)$) candidates used to build D_{s1}^+ candidates in the MB sample. The projection on each axis is also shown. The red dotted lines delimit the invariant-mass region of the signal selected to build the D_{s2}^{*+} and D_{s1}^+ candidates. 165
- 5.30 Invariant-mass distributions of $D^{*+}K_S^0$ (top) and $D^+K_S^0$ (bottom) candidates for the MB (left) and HM (right) samples. The red solid line represents the fit to the data. The blue dashed line represents the background contribution. 167
- 5.31 Left: efficiency map of prompt D^+ mesons in MB pp collisions at $\sqrt{s} = 13$ TeV. Right: correlation between the p_T of the D^+ and D_{s2}^{*+} mesons. . . 169
- 5.32 Left: comparison between the $(\text{Acc} \times \epsilon)$ factor of prompt D^{*+} mesons computed with the standard procedure adopted in the D_s^+ analysis and the MC-sampling method. Central: comparison between the $(\text{Acc} \times \epsilon)$ factor of prompt and non-prompt D_{s1}^+ mesons in MB and HM pp collisions at $\sqrt{s} = 13$ TeV. Right: comparison between the $(\text{Acc} \times \epsilon)$ factor of prompt and non-prompt D_{s2}^{*+} mesons in MB and HM pp collisions at $\sqrt{s} = 13$ TeV. 170
- 5.33 Left: RMS/mean ratio of the systematic uncertainty as a function of the multiplication factor of the data sample for the D_{s2}^{*+} resonance in the HM sample. Right: alternative p_T -shapes for systematic uncertainty estimation. 174
- 5.34 p_T -integrated yields (in $2 < p_T < 24$ GeV/c) times the BR of the analysed decay channel of prompt D_{s1}^+ and D_{s2}^{*+} mesons compared with the p_T -integrated yields of prompt D_s^+ mesons from Ref. [270]. 177
- 5.35 Left: p_T -integrated yield ratios of prompt D_{s1}^+ mesons over prompt D_s^+ mesons as a function of the average charged-particle multiplicity ($\langle dN_{\text{ch}}/d\eta \rangle$) in pp collisions at $\sqrt{s} = 13$ TeV. Right: p_T -integrated yield ratios of prompt D_{s2}^{*+} mesons over prompt D_s^+ mesons as a function of the average charged-particle multiplicity ($\langle dN_{\text{ch}}/d\eta \rangle$) in pp collisions at $\sqrt{s} = 13$ TeV. The vertical bars and the empty boxes represent the statistical and systematic uncertainties, respectively. The shaded boxes represent the systematic uncertainties on the branching ratios. The predictions based on the SHM [11] and SHMc [329] are shown as solid and dashed lines, respectively. 179

- 6.1 ROC curves obtained from the test set for the models employed in central Pb–Pb collisions in the $4 < p_T < 6$ GeV/ c interval. 185
- 6.2 Expected significance (top-left), expected signal-to-background ratio (top-right) and efficiency term (bottom left) for non-prompt D_s^+ mesons obtained in central Pb–Pb collisions in $2 < p_T < 4$ GeV/ c interval. 187
- 6.3 Invariant-mass distribution of non-prompt D_s^+ meson candidates in central (0 – 10%) Pb–Pb collisions at $\sqrt{s_{NN}} = 5.02$ TeV in the p_T intervals of the analyses. The blue solid line represents the result of the fit to the data. The orange dashed line represents the background component of the fit. 190
- 6.4 Invariant-mass distribution of non-prompt D_s^+ meson candidates in semicentral (30 – 50%) Pb–Pb collisions at $\sqrt{s_{NN}} = 5.02$ TeV in the p_T intervals of the analyses. The blue solid line represents the result of the fit to the data. The orange dashed line represents the background component of the fit. . . 191
- 6.5 Acceptance-times-efficiency as a function of p_T for central and semicentral Pb–Pb collisions for prompt and non-prompt D_s^+ mesons in the non-prompt enhanced sample. The bottom panel shows the ratio of the efficiency of non-prompt and prompt D_s^+ mesons for the two centrality classes. Figure from Ref. [335]. 192
- 6.6 Top-left: $(\text{Acc} \times \epsilon)$ factors for prompt and non-prompt D mesons as a function of the set of selection. Top-right: correlation matrix between pairs of sets of selection. The colour scale represents the correlation coefficient between the two sets of selections. Bottom-left: raw-yield as a function of the cut set with the prompt and non-prompt contributions obtained from the minimisation procedure. Bottom-right: prompt and non-prompt fractions as a function of the cut set. 193
- 6.7 Non-prompt D_s^+ fraction as a function of p_T in central and semicentral Pb–Pb collisions at $\sqrt{s_{NN}} = 5.02$ TeV. The statistical and systematic uncertainties associated with the estimation of the non-prompt D_s^+ meson fraction are reported as vertical bars and boxes, respectively. Figure from Ref. [335]. . . 195
- 6.8 Non-prompt D_s^+ meson pp reference cross section at $\sqrt{s} = 5.02$ TeV at midrapidity. The full markers indicate the measured cross section, while the open markers indicate the extrapolated cross section. 196

- 6.9 (Acc $\times$ ϵ) factors for non-prompt D_s^+ mesons originating from B^0 , B^+ , and B_s^0 mesons as a function of p_T in central Pb–Pb collisions at $\sqrt{s_{NN}} = 5.02$ TeV. Right: ratio of the (Acc $\times$ ϵ) factors for non-prompt D_s^+ mesons originating from B^0 , B^+ , and B_s^0 mesons as a function of p_T in central Pb–Pb collisions at $\sqrt{s_{NN}} = 5.02$ TeV. The statistical and systematic uncertainties are depicted as vertical lines and empty boxes, respectively. 202
- 6.10 Prompt [336] and non-prompt D_s^+ mesons p_T -differential production yield in central (0 – 10%) and semicentral (30 – 50%) Pb–Pb collisions at $\sqrt{s_{NN}} = 5.02$ TeV. The statistical and systematic uncertainties are depicted as vertical lines and empty boxes, respectively. The prompt D_s^+ meson production yield is scaled by a factor of 10 for visibility reasons. Figure from Ref. [335]. . . 203
- 6.11 Comparison of the prompt D_s^+ meson p_T -differential production yield in central (0 – 10%) and semicentral (30 – 50%) Pb–Pb collisions at $\sqrt{s_{NN}} = 5.02$ TeV obtained with the prompt-enhanced sample from the analysis method presented in this Thesis and the one from Ref. [336]. The statistical and systematic uncertainties are depicted as vertical lines and empty boxes, respectively. 204
- 6.12 Left: ratio of the production yield of non-prompt to prompt D_s^+ mesons in central (0 – 10%) and semicentral (30 – 50%) Pb–Pb collisions at $\sqrt{s_{NN}} = 5.02$ TeV, and pp collisions at $\sqrt{s} = 5.02$ TeV. Right: ratio of the production yield of non-prompt D_s^+ to non-prompt D^0 mesons in central (0 – 10%) and semicentral (30 – 50%) Pb–Pb collisions at $\sqrt{s_{NN}} = 5.02$ TeV, and pp collisions at $\sqrt{s} = 5.02$ TeV. The statistical and systematic uncertainties are depicted as vertical lines and empty boxes, respectively. Figure from Ref. [335]. 206
- 6.13 Left: R_{AA} of non-prompt and prompt D_s^+ mesons in central (0 – 10%, top panel) and semicentral (30 – 50%, bottom panel) Pb–Pb collisions at $\sqrt{s_{NN}} = 5.02$ TeV. Right: R_{AA} of non-prompt D_s^+ and D^0 mesons in central (0 – 10%, top panel) and semicentral (30 – 50%, bottom panel) Pb–Pb collisions at $\sqrt{s_{NN}} = 5.02$ TeV. The statistical and systematic uncertainties are depicted as vertical lines and empty boxes, respectively. The R_{AA} values are compared to the predictions of the TAMU model [126, 338]. Figure from Ref. [335]. . 208

- 6.14 Left: ratio of the production yield of non-prompt D_s^+ mesons in central (0 – 10%) and semicentral (30 – 50%) Pb–Pb collisions at $\sqrt{s_{NN}} = 5.02$ TeV, and pp collisions at $\sqrt{s} = 5.02$ TeV. Right: ratio of the production yield of non-prompt D_s^+ to D^0 mesons in central (0 – 10%) and semicentral (30 – 50%) Pb–Pb collisions at $\sqrt{s_{NN}} = 5.02$ TeV, and pp collisions at $\sqrt{s} = 5.02$ TeV. The statistical and systematic uncertainties are depicted as vertical lines and empty boxes, respectively. Figure from Ref. [335]. 209
- 6.15 R_{AA} of non-prompt D_s^+ to non-prompt D^0 mesons compared to the analogous ratio of B_s^0 to B^+ mesons measured by the CMS Collaboration in Pb–Pb collisions at $\sqrt{s_{NN}} = 5.02$ TeV [339] and scaled by the analogous ratio measured in pp collisions at $\sqrt{s} = 13$ TeV by the LHCb Collaboration [147]. The statistical and systematic uncertainties are depicted as vertical lines and empty boxes, respectively. 210
- 7.1 Invariant-mass distribution of D_s^+ meson candidates in the hadronic decay channel $D_s^+ \rightarrow \phi \pi^+ \rightarrow K^+ K^- \pi^+$ in pp collisions at $\sqrt{s} = 13.6$ TeV obtained by analysing Run 3 data for the $2 < p_T < 3$ GeV/c (left) and $3 < p_T < 4$ GeV/c (right) intervals. The distributions are normalised to the considered number of entries. Figure from ALICE Figure repository©. 215
- 7.2 Left: expected nuclear modification factor of non-prompt D_s^+ mesons as a function of p_T , compared to the predictions for non-prompt D_s^+ and D^0 mesons from the TAMU model. Right: the expected elliptic flow of non-prompt D_s^+ mesons as a function of p_T , compared to the predictions for prompt and non-prompt D_s^+ mesons from the TAMU model. In both panels, the vertical bars represent the expected statistical uncertainty on the measurements, while the empty boxes represent the expected systematic uncertainty. Figure from Ref. [347]. 217

References

- [1] M. E. Peskin and D. V. Schroeder, *An introduction to quantum field theory*. Westview, Boulder, CO, 1995. <https://cds.cern.ch/record/257493>. Includes exercises.
- [2] T. Cheng and L. Li, *Gauge Theory of Elementary Particle Physics*. Oxford University Press, Oxford, UK, 1984.
- [3] D. J. Gross and F. Wilczek, “Ultraviolet behavior of non-abelian gauge theories”, *Phys. Rev. Lett.* **30** (Jun, 1973) 1343–1346.
- [4] H. D. Politzer, “Reliable perturbative results for strong interactions”, *Phys. Rev. Lett.* **30** (Jun, 1973) 1346–1349.
- [5] K. G. Wilson, “Confinement of quarks”, *Phys. Rev. D* **10** (Oct, 1974) 2445–2459.
- [6] **Particle Data Group** Collaboration, R. L. Workman *et al.*, “Review of Particle Physics”, *PTEP* **2022** (2022) 083C01.
- [7] A. Li, Z. Y. Zhu, E. P. Zhou, J. M. Dong, J. N. Hu, and C. J. Xia, “Neutron star equation of state: Quark mean-field (QMF) modeling and applications”, *JHEAp* **28** (2020) 19–46, [arXiv:2007.05116 \[nucl-th\]](#).
- [8] K. Rajagopal and F. Wilczek, *The Condensed matter physics of QCD*, pp. 2061–2151. 11, 2000. [arXiv:hep-ph/0011333](#).
- [9] L.-M. Wang, R. R. Caldwell, J. P. Ostriker, and P. J. Steinhardt, “Cosmic concordance and quintessence”, *Astrophys. J.* **530** (2000) 17–35, [arXiv:astro-ph/9901388](#).
- [10] **Planck** Collaboration, P. Ade *et al.*, “Planck 2015 results-xiii. cosmological parameters”, *Astronomy & Astrophysics* **594** (2016) A13.
- [11] **HotQCD** Collaboration, A. Bazavov *et al.*, “Chiral crossover in QCD at zero and non-zero chemical potentials”, *Phys. Lett. B* **795** (2019) 15–21, [arXiv:1812.08235 \[hep-lat\]](#).
- [12] S. Borsanyi, Z. Fodor, J. N. Guenther, R. Kara, S. D. Katz, P. Parotto, A. Pasztor, C. Ratti, and K. K. Szabo, “QCD Crossover at Finite Chemical Potential from Lattice Simulations”, *Phys. Rev. Lett.* **125** no. 5, (2020) 052001, [arXiv:2002.02821 \[hep-lat\]](#).
- [13] G. Martinez, “Advances in Quark Gluon Plasma”, 4, 2013. [arXiv:1304.1452 \[nucl-ex\]](#).

- [14] **STAR** Collaboration, G. Odyniec, “Beam Energy Scan Program at RHIC (BES I and BES II) – Probing QCD Phase Diagram with Heavy-Ion Collisions”, *PoS CORFU2018* (2019) 151.
- [15] **CBM** Collaboration, K. Agarwal *et al.*, “The compressed baryonic matter (CBM) experiment at FAIR: physics, status and prospects”, *Physica Scripta* **98** no. 3, (Feb, 2023) 034006.
- [16] I. Arsene and e. a. I.G. Bearden, “Nuclear stopping and rapidity loss in Au–Au collisions at $\sqrt{s_{NN}} = 62.4$ GeV”, *Physics Letters B* **677** no. 5, (2009) 267–271.
- [17] J. D. Bjorken, “Highly relativistic nucleus-nucleus collisions: The central rapidity region”, *Phys. Rev. D* **27** (Jan, 1983) 140–151.
- [18] R. Stock, “Relativistic Nucleus-Nucleus Collisions and the QCD Matter Phase Diagram”, [arXiv:0807.1610 \[nucl-ex\]](#).
- [19] **ALICE** Collaboration, J. Adam *et al.*, “Measurement of transverse energy at midrapidity in Pb–Pb collisions at $\sqrt{s_{NN}} = 2.76$ TeV”, *Phys. Rev. C* **94** no. 3, (2016) 034903, [arXiv:1603.04775 \[nucl-ex\]](#).
- [20] P. Romatschke and U. Romatschke, *Relativistic Fluid Dynamics In and Out of Equilibrium*. Cambridge Monographs on Mathematical Physics. Cambridge University Press, 5, 2019. [arXiv:1712.05815 \[nucl-th\]](#).
- [21] A. Andronic, P. Braun-Munzinger, K. Redlich, and J. Stachel, “Decoding the phase structure of QCD via particle production at high energy”, *Nature* **561** no. 7723, (2018) 321–330, [arXiv:1710.09425 \[nucl-th\]](#).
- [22] M. L. Miller, K. Reygers, S. J. Sanders, and P. Steinberg, “Glauber modeling in high energy nuclear collisions”, *Ann. Rev. Nucl. Part. Sci.* **57** (2007) 205–243, [arXiv:nucl-ex/0701025](#).
- [23] R. J. Glauber, “Lectures on theoretical physics.” In lectures in theoretical physics, edited by w. e. brittin and l. g. dunham (interscience, new york, 1959), vol. 1, p. 315., 1959.
- [24] B. Schenke, P. Tribedy, and R. Venugopalan, “Fluctuating Glasma initial conditions and flow in heavy ion collisions”, *Phys. Rev. Lett.* **108** (2012) 252301, [arXiv:1202.6646 \[nucl-th\]](#).
- [25] **ALICE** Collaboration, J. Adam *et al.*, “Centrality Dependence of the Charged-Particle Multiplicity Density at Midrapidity in Pb–Pb Collisions at $\sqrt{s_{NN}} = 5.02$ TeV”, *Phys. Rev. Lett.* **116** no. 22, (2016) 222302, [arXiv:1512.06104 \[nucl-ex\]](#).
- [26] A. Bialas, M. Bleszynski, and W. Czyz, “Multiplicity Distributions in Nucleus-Nucleus Collisions at High-Energies”, *Nucl. Phys. B* **111** (1976) 461–476.
- [27] F. Gelis, E. Iancu, J. Jalilian-Marian, and R. Venugopalan, “The Color Glass Condensate”, *Ann. Rev. Nucl. Part. Sci.* **60** (2010) 463–489, [arXiv:1002.0333 \[hep-ph\]](#).

- [28] J. R. Andersen and A. Sabio Vera, “Solving the BFKL equation in the next-to-leading approximation”, *Phys. Lett. B* **567** (2003) 116–124, [arXiv:hep-ph/0305236](#).
- [29] J. L. Albacete, “Testing the CGC in proton-lead collisions at the LHC”, *Nucl. Phys. A* **910-911** (2013) 155–162, [arXiv:1209.0336 \[hep-ph\]](#).
- [30] J. L. Albacete, “Cgc and initial state effects in heavy ion collisions”, *Journal of Physics: Conference Series* **270** no. 1, (Jan, 2011) 012052. <https://dx.doi.org/10.1088/1742-6596/270/1/012052>.
- [31] A. Dumitru, D. E. Kharzeev, E. M. Levin, and Y. Nara, “Gluon saturation in pp collisions at energies available at the cern large hadron collider: Predictions for hadron multiplicities”, *Phys. Rev. C* **85** (Apr, 2012) 044920. <https://link.aps.org/doi/10.1103/PhysRevC.85.044920>.
- [32] S. Wheaton and J. Cleymans, “THERMUS: A Thermal model package for ROOT”, *Comput. Phys. Commun.* **180** (2009) 84–106, [arXiv:hep-ph/0407174](#).
- [33] A. Andronic, P. Braun-Munzinger, and J. Stachel, “Hadron production in central nucleus-nucleus collisions at chemical freeze-out”, *Nucl. Phys. A* **772** (2006) 167–199, [arXiv:nucl-th/0511071](#).
- [34] G. Torrieri, S. Steinke, W. Broniowski, W. Florkowski, J. Letessier, and J. Rafelski, “SHARE: Statistical hadronization with resonances”, *Comput. Phys. Commun.* **167** (2005) 229–251, [arXiv:nucl-th/0404083](#).
- [35] Y. Pan and S. Pratt, “Baryon Annihilation in Heavy Ion Collisions”, [arXiv:1210.1577 \[nucl-th\]](#).
- [36] A. Andronic, P. Braun-Munzinger, B. Friman, P. M. Lo, K. Redlich, and J. Stachel, “The thermal proton yield anomaly in Pb–Pb collisions at the LHC and its resolution”, *Phys. Lett. B* **792** (2019) 304–309, [arXiv:1808.03102 \[hep-ph\]](#).
- [37] J. Stachel, A. Andronic, P. Braun-Munzinger, and K. Redlich, “Confronting LHC data with the statistical hadronization model”, *J. Phys. Conf. Ser.* **509** (2014) 012019, [arXiv:1311.4662 \[nucl-th\]](#).
- [38] NA35 Collaboration, J. Baechler *et al.*, “Strangeness enhancement in central S + S collisions at 200-GeV/nucleon”, *Nucl. Phys. A* **525** (1991) 221C–226C.
- [39] WA97 Collaboration, E. Andersen *et al.*, “Strangeness enhancement at mid-rapidity in Pb Pb collisions at 158-A-GeV/c”, *Phys. Lett. B* **449** (1999) 401–406.
- [40] NA57 Collaboration, T. Virgili *et al.*, “Study of the production of strange and multi-strange particles in lead–lead interactions at the CERN SPS: the NA57 experiment”, *Nuclear Physics A* **681** no. 1, (2001) 165–173.
- [41] J. Rafelski and B. Müller, “Strangeness production in the quark-gluon plasma”, *Phys. Rev. Lett.* **48** (Apr, 1982) 1066–1069.
- [42] A. Tounsi and K. Redlich, “Strangeness enhancement and canonical suppression”, [arXiv:hep-ph/0111159](#).

- [43] **STAR** Collaboration, B. I. Abelev *et al.*, “Enhanced strange baryon production in Au + Au collisions compared to p + p at $\sqrt{s_{NN}} = 200$ -GeV”, *Phys. Rev. C* **77** (2008) 044908, [arXiv:0705.2511](#) [nucl-ex].
- [44] **ALICE** Collaboration, B. B. Abelev *et al.*, “Multi-strange baryon production at mid-rapidity in Pb–Pb collisions at $\sqrt{s_{NN}} = 2.76$ TeV”, *Phys. Lett. B* **728** (2014) 216–227, [arXiv:1307.5543](#) [nucl-ex]. [Erratum: *Phys.Lett.B* 734, 409–410 (2014)].
- [45] **ALICE** Collaboration, J. Adam *et al.*, “Enhanced production of multi-strange hadrons in high-multiplicity proton-proton collisions”, *Nature Phys.* **13** (2017) 535–539, [arXiv:1606.07424](#) [nucl-ex].
- [46] G. Gattoff and C.-Y. Wong, “Origin of the soft p_T spectra”, *Phys. Rev. D* **46** (Aug, 1992) 997–1006.
- [47] **ALICE** Collaboration, “The ALICE experiment – A journey through QCD”, [arXiv:2211.04384](#) [nucl-ex].
- [48] E. Schnedermann, J. Sollfrank, and U. W. Heinz, “Thermal phenomenology of hadrons from 200-A/GeV S+S collisions”, *Phys. Rev. C* **48** (1993) 2462–2475, [arXiv:nucl-th/9307020](#).
- [49] **ALICE** Collaboration, B. Abelev *et al.*, “Pion, Kaon, and Proton Production in Central Pb–Pb Collisions at $\sqrt{s_{NN}} = 2.76$ TeV”, *Phys. Rev. Lett.* **109** (2012) 252301, [arXiv:1208.1974](#) [hep-ex].
- [50] **ALICE** Collaboration, S. Acharya *et al.*, “Production of charged pions, kaons, and (anti-)protons in Pb–Pb and inelastic pp collisions at $\sqrt{s_{NN}} = 5.02$ TeV”, *Phys. Rev. C* **101** no. 4, (2020) 044907, [arXiv:1910.07678](#) [nucl-ex].
- [51] S. McDonald, C. Shen, F. Fillion-Gourdeau, S. Jeon, and C. Gale, “Hydrodynamic predictions for Pb+Pb collisions at 5.02 TeV”, *Phys. Rev. C* **95** no. 6, (2017) 064913, [arXiv:1609.02958](#) [hep-ph].
- [52] W. Zhao, H.-j. Xu, and H. Song, “Collective flow in 2.76 A TeV and 5.02 A TeV Pb+Pb collisions”, *Eur. Phys. J. C* **77** no. 9, (2017) 645, [arXiv:1703.10792](#) [nucl-th].
- [53] S. Voloshin and Y. Zhang, “Flow study in relativistic nuclear collisions by Fourier expansion of Azimuthal particle distributions”, *Z. Phys. C* **70** (1996) 665–672, [arXiv:hep-ph/9407282](#).
- [54] B. H. Alver, C. Gombeaud, M. Luzum, and J.-Y. Ollitrault, “Triangular flow in hydrodynamics and transport theory”, *Phys. Rev. C* **82** (2010) 034913, [arXiv:1007.5469](#) [nucl-th].
- [55] D. Teaney and L. Yan, “Non linearities in the harmonic spectrum of heavy ion collisions with ideal and viscous hydrodynamics”, *Phys. Rev. C* **86** (2012) 044908, [arXiv:1206.1905](#) [nucl-th].

- [56] B. Alver and G. Roland, “Collision geometry fluctuations and triangular flow in heavy-ion collisions”, *Phys. Rev. C* **81** (2010) 054905, [arXiv:1003.0194 \[nucl-th\]](#). [Erratum: *Phys.Rev.C* 82, 039903 (2010)].
- [57] **ALICE** Collaboration, S. Acharya *et al.*, “Linear and non-linear flow modes in Pb–Pb collisions at $\sqrt{s_{NN}} = 2.76$ TeV”, *Phys. Lett. B* **773** (2017) 68–80, [arXiv:1705.04377 \[nucl-ex\]](#).
- [58] **ALICE** Collaboration, S. Acharya *et al.*, “Anisotropic flow of identified particles in Pb–Pb collisions at $\sqrt{s_{NN}} = 5.02$ TeV”, *JHEP* **09** (2018) 006, [arXiv:1805.04390 \[nucl-ex\]](#).
- [59] **STAR** Collaboration, J. Adams *et al.*, “Evidence from d + Au measurements for final state suppression of high p_T hadrons in Au+Au collisions at RHIC”, *Phys. Rev. Lett.* **91** (2003) 072304, [arXiv:nucl-ex/0306024](#).
- [60] J. Casalderrey-Solana, J. G. Milhano, and U. Wiedemann, “Jet quenching via jet collimation”, *J. Phys. G* **38** (2011) 124086, [arXiv:1107.1964 \[hep-ph\]](#).
- [61] **ALICE** Collaboration, 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 \[nucl-ex\]](#).
- [62] **CMS** Collaboration, V. Khachatryan *et al.*, “Charged-particle nuclear modification factors in Pb–Pb and pPb collisions at $\sqrt{s_{NN}} = 5.02$ TeV”, *JHEP* **04** (2017) 039, [arXiv:1611.01664 \[nucl-ex\]](#).
- [63] **CMS** Collaboration, V. Khachatryan *et al.*, “Nuclear Effects on the Transverse Momentum Spectra of Charged Particles in p–Pb Collisions at $\sqrt{s_{NN}} = 5.02$ TeV”, *Eur. Phys. J. C* **75** no. 5, (2015) 237, [arXiv:1502.05387 \[nucl-ex\]](#).
- [64] **ALICE** Collaboration, J. Adam *et al.*, “Direct photon production in Pb–Pb collisions at $\sqrt{s_{NN}} = 2.76$ TeV”, *Phys. Lett. B* **754** (2016) 235–248, [arXiv:1509.07324 \[nucl-ex\]](#).
- [65] **ATLAS** Collaboration, G. Aad *et al.*, “Measurement of Z boson Production in Pb+Pb Collisions at $\sqrt{s_{NN}} = 2.76$ TeV with the ATLAS Detector”, *Phys. Rev. Lett.* **110** no. 2, (2013) 022301, [arXiv:1210.6486 \[hep-ex\]](#).
- [66] **CMS** Collaboration, S. Chatrchyan *et al.*, “Study of W Boson Production in Pb–Pb and pp Collisions at $\sqrt{s_{NN}} = 2.76$ TeV”, *Phys. Lett. B* **715** (2012) 66–87, [arXiv:1205.6334 \[nucl-ex\]](#).
- [67] **ALICE** Collaboration, S. Acharya *et al.*, “Measurements of inclusive jet spectra in pp and central Pb–Pb collisions at $\sqrt{s_{NN}} = 5.02$ TeV”, *Phys. Rev. C* **101** no. 3, (2020) 034911, [arXiv:1909.09718 \[nucl-ex\]](#).
- [68] **ATLAS** Collaboration, G. Aad *et al.*, “Measurement of the jet radius and transverse momentum dependence of inclusive jet suppression in lead-lead collisions at $\sqrt{s_{NN}} = 2.76$ TeV with the ATLAS detector”, *Phys. Lett. B* **719** (2013) 220–241, [arXiv:1208.1967 \[hep-ex\]](#).

- [69] **ATLAS** Collaboration, M. Aaboud *et al.*, “Measurement of the nuclear modification factor for inclusive jets in Pb+Pb collisions at $\sqrt{s_{\text{NN}}} = 5.02$ TeV with the ATLAS detector”, *Phys. Lett. B* **790** (2019) 108–128, [arXiv:1805.05635 \[nucl-ex\]](#).
- [70] **ATLAS** Collaboration, M. Aaboud *et al.*, “Measurement of jet fragmentation in Pb+Pb and pp collisions at $\sqrt{s_{\text{NN}}} = 5.02$ TeV with the ATLAS detector”, *Phys. Rev. C* **98** no. 2, (2018) 024908, [arXiv:1805.05424 \[nucl-ex\]](#).
- [71] **CMS** Collaboration, V. Khachatryan *et al.*, “Measurement of inclusive jet cross sections in pp and Pb–Pb collisions at $\sqrt{s_{\text{NN}}} = 2.76$ TeV”, *Phys. Rev. C* **96** no. 1, (2017) 015202, [arXiv:1609.05383 \[nucl-ex\]](#).
- [72] **CMS** Collaboration, S. Chatrchyan *et al.*, “Observation and studies of jet quenching in Pb–Pb collisions at nucleon-nucleon center-of-mass energy = 2.76 TeV”, *Phys. Rev. C* **84** (2011) 024906, [arXiv:1102.1957 \[nucl-ex\]](#).
- [73] A. Mocsy and P. Petreczky, “Can quarkonia survive deconfinement?”, *Phys. Rev. D* **77** (2008) 014501, [arXiv:0705.2559 \[hep-ph\]](#).
- [74] T. Matsui and H. Satz, “ J/ψ suppression by quark-gluon plasma formation”, *Physics Letters B* **178** no. 4, (1986) 416–422.
<https://www.sciencedirect.com/science/article/pii/0370269386914048>.
- [75] **CMS** Collaboration, “Observation of the $\Upsilon(3S)$ meson and sequential suppression of Υ states in Pb–Pb collisions at $\sqrt{s_{\text{NN}}} = 5.02$ TeV”,.
- [76] **ALICE** Collaboration, J. Adam *et al.*, “Differential studies of inclusive J/ψ and $\psi(2S)$ production at forward rapidity in Pb–Pb collisions at $\sqrt{s_{\text{NN}}} = 2.76$ TeV”, *JHEP* **05** (2016) 179, [arXiv:1506.08804 \[nucl-ex\]](#).
- [77] **PHENIX** Collaboration, A. Adare *et al.*, “ J/ψ suppression at forward rapidity in Au+Au collisions at $\sqrt{s_{\text{NN}}} = 200$ GeV”, *Phys. Rev. C* **84** (2011) 054912, [arXiv:1103.6269 \[nucl-ex\]](#).
- [78] F. Prino and R. Rapp, “Open Heavy Flavor in QCD Matter and in Nuclear Collisions”, *J. Phys. G* **43** no. 9, (2016) 093002, [arXiv:1603.00529 \[nucl-ex\]](#).
- [79] M. L. Mangano, “Two lectures on heavy quark production in hadronic collisions”, *Proc. Int. Sch. Phys. Fermi* **137** (1998) 95–137, [arXiv:hep-ph/9711337](#).
- [80] J. C. Collins, D. E. Soper, and G. F. Sterman, “Factorization of Hard Processes in QCD”, *Adv. Ser. Direct. High Energy Phys.* **5** (1989) 1–91, [arXiv:hep-ph/0409313](#).
- [81] J. Pumplin, D. R. Stump, J. Huston, H. L. Lai, P. M. Nadolsky, and W. K. Tung, “New generation of parton distributions with uncertainties from global QCD analysis”, *JHEP* **07** (2002) 012, [arXiv:hep-ph/0201195](#).
- [82] S. Dulat, T.-J. Hou, J. Gao, M. Guzzi, J. Huston, P. Nadolsky, J. Pumplin, C. Schmidt, D. Stump, and C. P. Yuan, “New parton distribution functions from a global analysis of quantum chromodynamics”, *Phys. Rev. D* **93** no. 3, (2016) 033006, [arXiv:1506.07443 \[hep-ph\]](#).

- [83] M. Cacciari, P. Nason, and C. Oleari, “A Study of heavy flavored meson fragmentation functions in e^+e^- annihilation”, *JHEP* **04** (2006) 006, [arXiv:hep-ph/0510032](#).
- [84] V. Kartvelishvili, A. Likhoded, and V. Petrov, “On the fragmentation functions of heavy quarks into hadrons”, *Physics Letters B* **78** no. 5, (1978) 615–617.
- [85] G. Altarelli and G. Parisi, “Asymptotic freedom in parton language”, *Nuclear Physics B* **126** no. 2, (1977) 298–318.
- [86] M. Cacciari, M. Greco, and P. Nason, “The p_T spectrum in heavy flavor hadroproduction”, *JHEP* **05** (1998) 007, [arXiv:hep-ph/9803400](#).
- [87] M. Cacciari, S. Frixione, N. Houdeau, M. L. Mangano, P. Nason, and G. Ridolfi, “Theoretical predictions for charm and bottom production at the LHC”, *JHEP* **10** (2012) 137, [arXiv:1205.6344 \[hep-ph\]](#).
- [88] G. Kramer and H. Spiesberger, “Study of heavy meson production in p–Pb collisions at $\sqrt{s}=5.02$ TeV in the general-mass variable-flavour-number scheme”, *Nucl. Phys. B* **925** (2017) 415–430, [arXiv:1703.04754 \[hep-ph\]](#).
- [89] M. Benzke, M. V. Garzelli, B. Kniehl, G. Kramer, S. Moch, and G. Sigl, “Prompt neutrinos from atmospheric charm in the general-mass variable-flavor-number scheme”, *JHEP* **12** (2017) 021, [arXiv:1705.10386 \[hep-ph\]](#).
- [90] M. Luszczak, R. Maciula, and A. Szczurek, “Nonphotonic electrons at RHIC within k_T -factorization approach and with experimental semileptonic decay functions”, *Phys. Rev. D* **79** (2009) 034009, [arXiv:0807.5044 \[hep-ph\]](#).
- [91] R. Maciula and A. Szczurek, “Open charm production at the LHC - k_T -factorization approach”, *Phys. Rev. D* **87** no. 9, (2013) 094022, [arXiv:1301.3033 \[hep-ph\]](#).
- [92] B. Guiot and A. van Hameren, “D and B-meson production using k_T -factorization calculations in a variable-flavor-number scheme”, *Phys. Rev. D* **104** no. 9, (2021) 094038, [arXiv:2108.06419 \[hep-ph\]](#).
- [93] T. Sjostrand, S. Mrenna, and P. Z. Skands, “PYTHIA 6.4 Physics and Manual”, *JHEP* **05** (2006) 026, [arXiv:hep-ph/0603175](#).
- [94] G. Marchesini, B. Webber, G. Abbiendi, I. Knowles, M. Seymour, and L. Stanco, “Herwig 5.1 - a monte carlo event generator for simulating hadron emission reactions with interfering gluons”, *Computer Physics Communications* **67** no. 3, (1992) 465–508.
- [95] S. Frixione and B. R. Webber, “Matching NLO QCD computations and parton shower simulations”, *JHEP* **06** (2002) 029, [arXiv:hep-ph/0204244](#).
- [96] S. Frixione, G. Ridolfi, and P. Nason, “A positive-weight next-to-leading-order monte carlo for heavy flavour hadroproduction”, *Journal of High Energy Physics* **2007** no. 09, (Sep, 2007) 126–126.

- [97] **ALICE** Collaboration, S. Acharya *et al.*, “Charm production and fragmentation fractions at midrapidity in pp collisions at $\sqrt{s} = 13$ TeV”, [arXiv:2308.04877 \[hep-ex\]](#).
- [98] M. He and R. Rapp, “Bottom hadro-chemistry in high-energy hadronic collisions”, [arXiv:2209.13419 \[hep-ph\]](#).
- [99] M. Arneodo, “Nuclear effects in structure functions”, *Physics Reports* **240** no. 5, (1994) 301–393. <https://www.sciencedirect.com/science/article/pii/0370157394900485>.
- [100] A. Bodek and J. L. Ritchie, “Fermi-motion effects in deep-inelastic lepton scattering from nuclear targets”, *Phys. Rev. D* **23** (Mar, 1981) 1070–1091. <https://link.aps.org/doi/10.1103/PhysRevD.23.1070>.
- [101] H. Paukkunen, “Nuclear pdfs in the beginning of the LHC era”, *Nuclear Physics A* **926** (2014) 24–33. <https://www.sciencedirect.com/science/article/pii/S0375947414000839>. IS2013.
- [102] N. Armesto, “Nuclear shadowing”, *J. Phys. G* **32** (2006) R367–R394, [arXiv:hep-ph/0604108](#).
- [103] **H1** Collaboration, F. D. Aaron *et al.*, “A Precision Measurement of the Inclusive ep Scattering Cross Section at HERA”, *Eur. Phys. J. C* **64** (2009) 561–587, [arXiv:0904.3513 \[hep-ex\]](#).
- [104] “The ratio of the nucleon structure functions f_2^n for iron and deuterium”, *Physics Letters B* **123** no. 3, (1983) 275–278.
- [105] K. J. Eskola, P. Paakkinen, H. Paukkunen, and C. A. Salgado, “EPPS16: Nuclear parton distributions with LHC data”, *Eur. Phys. J. C* **77** no. 3, (2017) 163, [arXiv:1612.05741 \[hep-ph\]](#).
- [106] K. Kovarik *et al.*, “nCTEQ15 - Global analysis of nuclear parton distributions with uncertainties in the CTEQ framework”, *Phys. Rev. D* **93** no. 8, (2016) 085037, [arXiv:1509.00792 \[hep-ph\]](#).
- [107] N. Carrer and A. Dainese, *Charm and beauty production at LHC*, pp. 1768–1783. 2006. [arXiv:hep-ph/0311225](#).
- [108] J. W. Cronin, H. J. Frisch, M. J. Shochet, J. P. Boymond, P. A. Piroué, and R. L. Sumner, “Production of hadrons at large transverse momentum at 200, 300, and 400 GeV”, *Phys. Rev. D* **11** (Jun, 1975) 3105–3123. <https://link.aps.org/doi/10.1103/PhysRevD.11.3105>.
- [109] A. Beraudo, A. De Pace, M. Monteno, M. Nardi, and F. Prino, “Heavy-flavour production in high-energy d-Au and p-Pb collisions”, *JHEP* **03** (2016) 123, [arXiv:1512.05186 \[hep-ph\]](#).
- [110] A. M. Sickles, “Possible Evidence for Radial Flow of Heavy Mesons in d+Au Collisions”, *Phys. Lett. B* **731** (2014) 51–56, [arXiv:1309.6924 \[nucl-th\]](#).

- [111] Y. Xu, S. Cao, M. Nahrgang, W. Ke, G.-Y. Qin, J. Auvinen, and S. A. Bass, “Heavy-flavor dynamics in relativistic p-pb collisions at $\sqrt{s_{NN}}=5.02$ tev”, *Nuclear and Particle Physics Proceedings* **276-278** (2016) 225–228.
<https://www.sciencedirect.com/science/article/pii/S2405601416300840>. 7th International Conference on Hard and Electromagnetic Probes of High Energy Nuclear Collisions.
- [112] G. D. Moore and D. Teaney, “How much do heavy quarks thermalize in a heavy ion collision?”, *Phys. Rev. C* **71** (2005) 064904, [arXiv:hep-ph/0412346](#).
- [113] S. Cao, G.-Y. Qin, and S. A. Bass, “Heavy-quark dynamics and hadronization in ultrarelativistic heavy-ion collisions: Collisional versus radiative energy loss”, *Phys. Rev. C* **88** (2013) 044907, [arXiv:1308.0617 \[nucl-th\]](#).
- [114] H. van Hees, V. Greco, and R. Rapp, “Heavy-quark probes of the quark-gluon plasma at RHIC”, *Phys. Rev. C* **73** (2006) 034913, [arXiv:nucl-th/0508055](#).
- [115] S. Peigne and A. Peshier, “Collisional energy loss of a fast heavy quark in a quark-gluon plasma”, *Phys. Rev. D* **77** (2008) 114017, [arXiv:0802.4364 \[hep-ph\]](#).
- [116] S. Peigne and A. V. Smilga, “Energy losses in a hot plasma revisited”, *Phys. Usp.* **52** (2009) 659–685, [arXiv:0810.5702 \[hep-ph\]](#).
- [117] ALICE Collaboration, S. Acharya *et al.*, “Prompt D^0 , D^+ , and D^{*+} production in Pb–Pb collisions at $\sqrt{s_{NN}} = 5.02$ TeV”, *JHEP* **01** (2022) 174, [arXiv:2110.09420 \[nucl-ex\]](#).
- [118] C. A. Salgado and U. A. Wiedemann, “Calculating quenching weights”, *Phys. Rev. D* **68** (2003) 014008, [arXiv:hep-ph/0302184](#).
- [119] R. Baier, Y. L. Dokshitzer, A. H. Mueller, S. Peigne, and D. Schiff, “Radiative energy loss and $p(T)$ broadening of high-energy partons in nuclei”, *Nucl. Phys. B* **484** (1997) 265–282, [arXiv:hep-ph/9608322](#).
- [120] M. L. Mangano, “Introduction to QCD”, in *1998 European School of High-Energy Physics*, pp. 53–97. 1998.
- [121] ALICE Collaboration, S. Acharya *et al.*, “Direct observation of the dead-cone effect in quantum chromodynamics”, *Nature* **605** no. 7910, (2022) 440–446, [arXiv:2106.05713 \[nucl-ex\]](#). [Erratum: *Nature* 607, E22 (2022)].
- [122] Y. L. Dokshitzer and D. E. Kharzeev, “Heavy quark colorimetry of QCD matter”, *Phys. Lett. B* **519** (2001) 199–206, [arXiv:hep-ph/0106202](#).
- [123] ALICE Collaboration, S. Acharya *et al.*, “Measurement of beauty production via non-prompt D^0 mesons in Pb–Pb collisions at $\sqrt{s_{NN}} = 5.02$ TeV”, *JHEP* **12** (2022) 126, [arXiv:2202.00815 \[nucl-ex\]](#).
- [124] A. Beraudo, A. De Pace, M. Monteno, M. Nardi, and F. Prino, “Development of heavy-flavour flow-harmonics in high-energy nuclear collisions”, *JHEP* **02** (2018) 043, [arXiv:1712.00588 \[hep-ph\]](#).

- [125] M. Nahrgang, J. Aichelin, P. B. Gossiaux, and K. Werner, “Influence of hadronic bound states above T_c on heavy-quark observables in Pb + Pb collisions at the CERN Large Hadron Collider”, *Phys. Rev. C* **89** no. 1, (2014) 014905, [arXiv:1305.6544 \[hep-ph\]](#).
- [126] M. He and R. Rapp, “Hadronization and Charm-Hadron Ratios in Heavy-Ion Collisions”, *Phys. Rev. Lett.* **124** no. 4, (2020) 042301, [arXiv:1905.09216 \[nucl-th\]](#).
- [127] A. Beraudo, A. De Pace, M. Monteno, M. Nardi, and F. Prino, “Heavy flavors in heavy-ion collisions: quenching, flow and correlations”, *Eur. Phys. J. C* **75** no. 3, (2015) 121, [arXiv:1410.6082 \[hep-ph\]](#).
- [128] R. Katz, C. A. G. Prado, J. Noronha-Hostler, J. Noronha, and A. A. P. Suaide, “Sensitivity study with a D and B mesons modular simulation code of heavy flavor RAA and azimuthal anisotropies based on beam energy, initial conditions, hadronization, and suppression mechanisms”, *Phys. Rev. C* **102** no. 2, (2020) 024906, [arXiv:1906.10768 \[nucl-th\]](#).
- [129] S. Li and J. Liao, “Data-driven extraction of heavy quark diffusion in quark-gluon plasma”, *Eur. Phys. J. C* **80** no. 7, (2020) 671, [arXiv:1912.08965 \[hep-ph\]](#).
- [130] F. Scardina, S. K. Das, V. Minissale, S. Plumari, and V. Greco, “Estimating the charm quark diffusion coefficient and thermalization time from D meson spectra at energies available at the BNL Relativistic Heavy Ion Collider and the CERN Large Hadron Collider”, *Phys. Rev. C* **96** no. 4, (2017) 044905, [arXiv:1707.05452 \[nucl-th\]](#).
- [131] S. Cao, T. Luo, G.-Y. Qin, and X.-N. Wang, “Linearized Boltzmann transport model for jet propagation in the quark-gluon plasma: Heavy quark evolution”, *Phys. Rev. C* **94** no. 1, (2016) 014909, [arXiv:1605.06447 \[nucl-th\]](#).
- [132] W. Ke, Y. Xu, and S. A. Bass, “Modified Boltzmann approach for modeling the splitting vertices induced by the hot QCD medium in the deep Landau-Pomeranchuk-Migdal region”, *Phys. Rev. C* **100** no. 6, (2019) 064911, [arXiv:1810.08177 \[nucl-th\]](#).
- [133] CMS Collaboration, “Measurements of azimuthal anisotropy of nonprompt D^0 mesons in Pb–Pb collisions at $\sqrt{s_{NN}} = 5.02$ TeV”, [arXiv:2212.01636 \[nucl-ex\]](#).
- [134] CMS Collaboration, A. M. Sirunyan *et al.*, “Measurement of prompt D^0 and $\bar{D}^0$ meson azimuthal anisotropy and search for strong electric fields in Pb–Pb collisions at $\sqrt{s_{NN}} = 5.02$ TeV”, *Phys. Lett. B* **816** (2021) 136253, [arXiv:2009.12628 \[hep-ex\]](#).
- [135] ALICE Collaboration, S. Acharya *et al.*, “D-meson azimuthal anisotropy in midcentral Pb–Pb collisions at $\sqrt{s_{NN}} = 5.02$ TeV”, *Phys. Rev. Lett.* **120** no. 10, (2018) 102301, [arXiv:1707.01005 \[nucl-ex\]](#).
- [136] ALICE Collaboration, S. Acharya *et al.*, “Measurement of D^0 , D^+ , D^{*+} and D_s^+ production in Pb–Pb collisions at $\sqrt{s_{NN}} = 5.02$ TeV”, *JHEP* **10** (2018) 174, [arXiv:1804.09083 \[nucl-ex\]](#).

- [137] CMS Collaboration, A. M. Sirunyan *et al.*, “Measurement of prompt D^0 meson azimuthal anisotropy in Pb–Pb collisions at $\sqrt{s_{NN}} = 5.02$ TeV”, *Phys. Rev. Lett.* **120** no. 20, (2018) 202301, [arXiv:1708.03497](#) [nucl-ex].
- [138] CMS Collaboration, A. M. Sirunyan *et al.*, “Nuclear modification factor of D^0 mesons in Pb–Pb collisions at $\sqrt{s_{NN}} = 5.02$ TeV”, *Phys. Lett. B* **782** (2018) 474–496, [arXiv:1708.04962](#) [nucl-ex].
- [139] CMS Collaboration, A. M. Sirunyan *et al.*, “Measurement of the $B^\pm$ Meson Nuclear Modification Factor in Pb–Pb Collisions at $\sqrt{s_{NN}} = 5.02$ TeV”, *Phys. Rev. Lett.* **119** no. 15, (2017) 152301, [arXiv:1705.04727](#) [hep-ex].
- [140] ALICE Collaboration, B. Abelev *et al.*, “Technical Design Report for the Upgrade of the ALICE Inner Tracking System”, tech. rep., 2014. <https://cds.cern.ch/record/1625842>.
- [141] Z. Citron *et al.*, “Report from Working Group 5: Future physics opportunities for high-density QCD at the LHC with heavy-ion and proton beams”, *CERN Yellow Rep. Monogr.* **7** (2019) 1159–1410, [arXiv:1812.06772](#) [hep-ph].
- [142] W. Ke, Y. Xu, and S. A. Bass, “Linearized Boltzmann-Langevin model for heavy quark transport in hot and dense QCD matter”, *Phys. Rev. C* **98** no. 6, (2018) 064901, [arXiv:1806.08848](#) [nucl-th].
- [143] B. Andersson, *The Lund model*, vol. 7. Cambridge University Press, 7, 2005.
- [144] B. Andersson, G. Gustafson, G. Ingelman, and T. Sjöstrand, “Parton fragmentation and string dynamics”, *Physics Reports* **97** no. 2, (1983) 31–145.
- [145] S. Ferreres-Solé and T. Sjöstrand, “The space–time structure of hadronization in the Lund model”, *Eur. Phys. J. C* **78** no. 11, (2018) 983, [arXiv:1808.04619](#) [hep-ph].
- [146] ALICE Collaboration, S. Acharya *et al.*, “Measurement of Prompt D^0 , Λ_c^+ , and $\Sigma_c^{0,++}(2455)$ Production in Proton–Proton Collisions at $\sqrt{s} = 13$ TeV”, *Phys. Rev. Lett.* **128** no. 1, (2022) 012001, [arXiv:2106.08278](#) [hep-ex].
- [147] LHCb Collaboration, R. Aaij *et al.*, “Measurement of b hadron fractions in 13 TeV pp collisions”, *Phys. Rev. D* **100** no. 3, (2019) 031102, [arXiv:1902.06794](#) [hep-ex].
- [148] T. Song, H. Berrehrah, D. Cabrera, J. M. Torres-Rincon, L. Tolos, W. Cassing, and E. Bratkovskaya, “Tomography of the Quark-Gluon-Plasma by Charm Quarks”, *Phys. Rev. C* **92** no. 1, (2015) 014910, [arXiv:1503.03039](#) [nucl-th].
- [149] C. B. Dover, U. Heinz, E. Schnedermann, and J. Zimányi, “Covariant coalescence model for relativistically expanding systems”, *Phys. Rev. C* **44** (Oct, 1991) 1636–1654.
- [150] V. Aa., “CERN: Our People.”.
- [151] V. Aa., “ALICE Diversity Office.”.

- [152] L. Di Lella and C. Rubbia, “The Discovery of the W and Z Particles”, *Adv. Ser. Direct. High Energy Phys.* **23** (2015) 137–163.
- [153] T. Berners-Lee, “Information management: A proposal”, 1989.
<https://www.w3.org/History/1989/proposal.html>.
- [154] CMS Collaboration, S. Chatrchyan *et al.*, “Observation of a new boson at a mass of 125 gev with the cms experiment at the lhc”, *Physics Letters B* **716** no. 1, (2012) 30–61. <https://www.sciencedirect.com/science/article/pii/S0370269312008581>.
- [155] ATLAS Collaboration, G. Aad *et al.*, “Observation of a new particle in the search for the standard model higgs boson with the atlas detector at the lhc”, *Physics Letters B* **716** no. 1, (2012) 1–29.
<https://www.sciencedirect.com/science/article/pii/S037026931200857X>.
- [156] E. Lopienska, “The CERN accelerator complex, layout in 2022. Complexe des accélérateurs du CERN en janvier 2022”,. <https://cds.cern.ch/record/2800984>. General Photo.
- [157] ALICE Collaboration, K. Aamodt *et al.*, “The ALICE experiment at the CERN LHC”, *JINST* **3** (2008) S08002.
- [158] ALICE Collaboration, S. Acharya *et al.*, “The ALICE Transition Radiation Detector: construction, operation, and performance”, *Nucl. Instrum. Meth. A* **881** (2018) 88–127, [arXiv:1709.02743](https://arxiv.org/abs/1709.02743) [physics.ins-det].
- [159] ALICE Collaboration, “ALICE Inner Tracking System (ITS): Technical Design Report”,. <https://cds.cern.ch/record/391175>.
- [160] ALICE Collaboration, K. Aamodt *et al.*, “Alignment of the ALICE Inner Tracking System with cosmic-ray tracks”, *JINST* **5** (2010) P03003, [arXiv:1001.0502](https://arxiv.org/abs/1001.0502) [physics.ins-det].
- [161] J. Parkkila, “Quantifying the transport properties of quark-gluon plasma through measurement of higher harmonic flow and their non-linear response”, 2021.
<https://cds.cern.ch/record/2791172>. Presented 19 Nov 2021.
- [162] A. Maire, “Production des baryons multi-étranges au LHC dans les collisions proton-proton avec l’expérience ALICE”, 2011. <https://cds.cern.ch/record/1490315>. Presented 13 Oct 2011.
- [163] ALICE Collaboration, G. Dellacasa *et al.*, “ALICE technical design report of the time-of-flight system (TOF)”,.
- [164] ALICE Collaboration, P. Cortese *et al.*, “ALICE technical design report on forward detectors: FMD, T0 and V0”,. <http://cds.cern.ch/record/781854>.
- [165] ALICE Collaboration, G. Dellacasa *et al.*, “ALICE technical design report of the zero degree calorimeter (ZDC)”,.
- [166] ALICE Collaboration, B. Abelev *et al.*, “Centrality determination of Pb–Pb collisions at $\sqrt{s_{NN}} = 2.76$ TeV with ALICE”, *Phys. Rev. C* **88** no. 4, (2013) 044909, [arXiv:1301.4361](https://arxiv.org/abs/1301.4361) [nucl-ex].

- [167] V. Aa., “**Monarc Simulation Framework**.”.
- [168] V. Aa., “**Worldwide LHC Computing Grid**.”.
- [169] R. Brun and F. Rademakers, “ROOT: An object-oriented data analysis framework”, *Nucl. Instrum. Meth. A* **389** (1997) 81–86.
- [170] V. Aa., “**AliRoot**”, 1998-.
- [171] V. Aa., “**AliPhysics**.”.
- [172] T. Sjöstrand, S. Ask, J. R. Christiansen, R. Corke, N. Desai, P. Ilten, S. Mrenna, S. Prestel, C. O. Rasmussen, and P. Z. Skands, “An introduction to PYTHIA 8.2”, *Comput. Phys. Commun.* **191** (2015) 159–177, [arXiv:1410.3012 \[hep-ph\]](#).
- [173] X.-N. Wang and M. Gyulassy, “hijing: A monte carlo model for multiple jet production in pp, pA, and AA collisions”, *Phys. Rev. D* **44** (Dec, 1991) 3501–3516.
- [174] R. Brun, F. Bruyant, F. Carminati, S. Giani, M. Maire, A. McPherson, G. Patrick, and L. Urban, *GEANT: Detector Description and Simulation Tool; Oct 1994*. CERN Program Library. CERN, Geneva, 1993. <http://cds.cern.ch/record/1082634>. Long Writeup W5013.
- [175] S. A. et al., “Geant4-a simulation toolkit”, *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* **506** no. 3, (2003) 250–303.
- [176] D. Kharzeev, E. Levin, and M. Nardi, “Color glass condensate at the LHC: Hadron multiplicities in pp, pA and AA collisions”, *Nucl. Phys. A* **747** (2005) 609–629, [arXiv:hep-ph/0408050](#).
- [177] W.-T. Deng, X.-N. Wang, and R. Xu, “Hadron production in p+p, p+Pb, and Pb+Pb collisions with the HIJING 2.0 model at energies available at the CERN Large Hadron Collider”, *Phys. Rev. C* **83** (2011) 014915, [arXiv:1008.1841 \[hep-ph\]](#).
- [178] **ALICE** Collaboration, B. B. Abelev *et al.*, “Performance of the ALICE Experiment at the CERN LHC”, *Int. J. Mod. Phys. A* **29** (2014) 1430044, [arXiv:1402.4476 \[nucl-ex\]](#).
- [179] A. Maire, “Track reconstruction principle in ALICE for LHC run I and run II. Principes de reconstruction de traces dans ALICE pour les runs I et II du LHC.” General Photo, 2011.
- [180] R. Frühwirth, “Application of kalman filtering to track and vertex fitting”, *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* **262** no. 2, (1987) 444–450. <https://www.sciencedirect.com/science/article/pii/0168900287908874>.
- [181] **ALICE** Collaboration, S. Acharya *et al.*, “Measurement of D^0 , D^+ , D^{*+} and D_s^+ production in pp collisions at $\sqrt{s} = 5.02$ TeV with ALICE”, *Eur. Phys. J. C* **79** no. 5, (2019) 388, [arXiv:1901.07979 \[nucl-ex\]](#).

- [182] **ALICE** Collaboration, S. Acharya *et al.*, “Measurement of D-meson production at mid-rapidity in pp collisions at $\sqrt{s} = 7$ TeV”, *Eur. Phys. J. C* **77** no. 8, (2017) 550, [arXiv:1702.00766 \[hep-ex\]](#).
- [183] **ALICE** Collaboration, B. Abelev *et al.*, “Measurement of charm production at central rapidity in proton-proton collisions at $\sqrt{s} = 7$ TeV”, *JHEP* **01** (2012) 128, [arXiv:1111.1553 \[hep-ex\]](#).
- [184] M. Faggin, “Measurement of heavy-flavour decay electrons and heavy-flavour baryon production with ALICE experiment at LHC”, 2021. <https://cds.cern.ch/record/2798337>. Presented 16 Dec 2021.
- [185] D. Silver *et al.*, “Mastering the game of Go with deep neural networks and tree search”, *Nature* **529** no. 7587, (2016) 484–489.
- [186] O. Russakovsky *et al.*, “Imagenet large scale visual recognition challenge”,. <http://arxiv.org/abs/1409.0575>.
- [187] A. Radovic *et al.*, “Machine learning at the energy and intensity frontiers of particle physics”, *Nature* **560** (2018) pp. 41–48. <http://cds.cern.ch/record/2647378>.
- [188] D. Guest, K. Cranmer, and D. Whiteson, “Deep learning and its application to lhc physics”, *Annual Review of Nuclear and Particle Science* **68** no. 1, (2018) 161–181, <https://doi.org/10.1146/annurev-nucl-101917-021019>.
- [189] K. Albertsson *et al.*, “Machine Learning in High Energy Physics Community White Paper”, *J. Phys. Conf. Ser.* **1085** no. 2, (2018) 022008, [arXiv:1807.02876 \[physics.comp-ph\]](#).
- [190] J. Lin, M. Freytsis, I. Moulton, and B. Nachman, “Boosting $H \rightarrow b\bar{b}$ with Machine Learning”, *JHEP* **10** (2018) 101, [arXiv:1807.10768 \[hep-ph\]](#).
- [191] I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning*. MIT Press, 2016.
- [192] V. Aa., “**Scikit-learn documentation**”, 2023.
- [193] T. Hastie, R. Tibshirani, and J. Friedman, *The Elements of Statistical Learning*. Springer Series in Statistics. Springer New York Inc., 2001.
- [194] D. Hand and R. Till, “A simple generalisation of the area under the roc curve for multiple class classification problems”, *Hand, The* **45** (11, 2001) 171–186.
- [195] J. Bergstra and Y. Bengio, “Random search for hyper-parameter optimization”, *Journal of Machine Learning Research* **13** no. Feb, (2012) 281–305. <http://www.jmlr.org/papers/v13/bergstra12a.html>.
- [196] J. Snoek, H. Larochelle, and R. P. Adams, “Practical bayesian optimization of machine learning algorithms”, in *Advances in neural information processing systems*, pp. 2951–2959. 2012.
- [197] E. Brochu, V. M. Cora, and N. de Freitas, “A tutorial on bayesian optimization of expensive cost functions, with application to active user modeling and hierarchical reinforcement learning”, 2010.

- [198] D. Passos and P. Mishra, “A tutorial on automatic hyperparameter tuning of deep spectral modelling for regression and classification tasks”, *Chemometrics and Intelligent Laboratory Systems* **223** (2022) .
- [199] I. Guyon and A. Elisseeff, “An introduction to variable and feature selection”, *J. Mach. Learn. Res.* **3** (2003) 1157–1182.
<http://portal.acm.org/citation.cfm?id=944919.944968>.
- [200] L. Breiman, “Random forests”, *Machine Learning* 5–32.
- [201] S. Lundberg, G. Erion, H. Chen, A. DeGrave, J. Prutkin, B. Nair, R. Katz, J. Himmelfarb, N. Bansal, and S.-I. Lee, “From local explanations to global understanding with explainable ai for trees”, *Nature Machine Intelligence* **2** (01, 2020) .
- [202] S. Lundberg and S.-I. Lee, “A unified approach to interpreting model predictions.”, 2017. <http://dblp.uni-trier.de/db/journals/corr/corr1705.html#LundbergL17>.
- [203] A. Roth, “The shapley value: essays in honor of lloyd s. shapley”, 1988.
- [204] S. Kotsiantis, “Decision trees: A recent overview”, *Artificial Intelligence Review* (04, 2013) 1–23.
- [205] J. Furnkranz, *Decision Tree*, pp. 330–335. 2017.
- [206] L. Breiman, *Classification and regression trees*. Routledge, 1984.
- [207] Y. Freund and R. E. Schapire, “A decision-theoretic generalization of on-line learning and an application to boosting”, *Journal of Computer and System Sciences* **55** no. 1, (1997) 119–139.
- [208] J. Friedman, T. Hastie, and R. Tibshirani, “Additive Logistic Regression: a Statistical View of Boosting”, *The Annals of Statistics* **38** no. 2, (2000) .
- [209] V. Aa., “**XGBoost documentation**”, 2023.
- [210] T. Chen and C. Guestrin, “**Xgboost: A scalable tree boosting system**”, in *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, p. 785–794. Association for Computing Machinery, 2016.
- [211] J. Therhaag, “TMVA Toolkit for multivariate data analysis in ROOT”, *PoS ICHEP2010* (2010) 510.
- [212] F. Rademakers *et al.*, “root-project/root: v6.20/04”, June, 2020.
<https://doi.org/10.5281/zenodo.3895855>.
- [213] J. Pivarski *et al.*, “scikit-hep/uproot4: 4.0.11”, June, 2021.
- [214] L. Barioglio, F. Catalano, M. Concas, P. Fecchio, F. Grosa, F. Mazzaschi, and M. Puccio, “hipe4ml/hipe4ml”, Apr., 2022. <https://doi.org/10.5281/zenodo.7014886>.
- [215] J. D. Hunter, “Matplotlib: A 2d graphics environment”, *Computing In Science & Engineering* **9** no. 3, (2007) 90–95.

- [216] C. R. Harris *et al.*, “Array programming with NumPy”, *Nature* **585** no. 7825, (Sep, 2020) 357–362.
- [217] The Pandas development team, “pandas-dev/pandas: Pandas”, Feb., 2020. <https://doi.org/10.5281/zenodo.3509134>.
- [218] Wes McKinney, “Data Structures for Statistical Computing in Python”, in *Proceedings of the 9th Python in Science Conference*, Stéfan van der Walt and Jarrod Millman, eds., pp. 56 – 61. 2010.
- [219] F. Pedregosa *et al.*, “Scikit-learn: Machine learning in python”, [arXiv:1201.0490](https://arxiv.org/abs/1201.0490) [cs.LG].
- [220] P. Virtanen *et al.*, “SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python”, *Nature Methods* **17** (2020) 261–272.
- [221] LHCb Collaboration, R. Aaij *et al.*, “Study of D_{sJ} decays to $D^+K_S^0$ and D^0K^+ final states in pp collisions”, *JHEP* **10** (2012) 151, [arXiv:1207.6016](https://arxiv.org/abs/1207.6016) [hep-ex].
- [222] LHCb Collaboration, R. Aaij *et al.*, “Study of D_J meson decays to $D^+\pi^-$, $D^0\pi^+$ and $D^{*+}\pi^-$ final states in pp collision”, *JHEP* **09** (2013) 145, [arXiv:1307.4556](https://arxiv.org/abs/1307.4556) [hep-ex].
- [223] LHCb Collaboration, R. Aaij *et al.*, “Observation of overlapping spin-1 and spin-3 $\bar{D}^0K^-$ resonances at mass $2.86\text{GeV}/c^2$ ”, *Phys. Rev. Lett.* **113** (2014) 162001, [arXiv:1407.7574](https://arxiv.org/abs/1407.7574) [hep-ex].
- [224] LHCb Collaboration, R. Aaij *et al.*, “Dalitz plot analysis of $B_s^0 \rightarrow \bar{D}^0K^-\pi^+$ decays”, *Phys. Rev. D* **90** no. 7, (2014) 072003, [arXiv:1407.7712](https://arxiv.org/abs/1407.7712) [hep-ex].
- [225] G. Torrieri and J. Rafelski, “Strange hadron resonances as a signature of freezeout dynamics”, *Phys. Lett. B* **509** (2001) 239–245, [arXiv:hep-ph/0103149](https://arxiv.org/abs/hep-ph/0103149).
- [226] C. Bierlich, T. Sjöstrand, and M. Uthm, “Hadronic rescattering in pA and AA collisions”, *Eur. Phys. J. A* **57** no. 7, (2021) 227, [arXiv:2103.09665](https://arxiv.org/abs/2103.09665) [hep-ph].
- [227] ALICE Collaboration, S. Acharya *et al.*, “System size dependence of hadronic rescattering effect at LHC energies”, [arXiv:2308.16115](https://arxiv.org/abs/2308.16115) [nucl-ex].
- [228] ALICE Collaboration, S. Acharya *et al.*, “Production of light-flavor hadrons in pp collisions at $\sqrt{s} = 7$ and $\sqrt{s} = 13$ TeV”, *Eur. Phys. J. C* **81** no. 3, (2021) 256, [arXiv:2005.11120](https://arxiv.org/abs/2005.11120) [nucl-ex].
- [229] E. Bruna, A. Dainese, M. Masera, and F. Prino, “Vertex reconstruction for proton-proton collisions in ALICE”, <https://cds.cern.ch/record/1225497>.
- [230] ALICE Collaboration, E. Abbas *et al.*, “Performance of the ALICE VZERO system”, *JINST* **8** (2013) P10016, [arXiv:1306.3130](https://arxiv.org/abs/1306.3130) [nucl-ex].

- [231] **ALICE** Collaboration, C. W. Fabjan, L. Jirdén, V. Lindestruth, L. Riccati, D. Rorich, P. Van de Vyvre, O. Villalobos Baillie, and H. de Groot, *ALICE trigger data-acquisition high-level trigger and control system: Technical Design Report*. Technical design report. ALICE. CERN, Geneva, 2004. <https://cds.cern.ch/record/684651>.
- [232] **ALICE** Collaboration, “ALICE 2016-2017-2018 luminosity determination for pp collisions at $\sqrt{s} = 13$ TeV”, <https://cds.cern.ch/record/2776672>.
- [233] **ALICE** Collaboration, “ALICE luminosity determination for Pb–Pb collisions at $\sqrt{s_{NN}} = 5.02$ TeV”, [arXiv:2204.10148](https://arxiv.org/abs/2204.10148) [nucl-ex].
- [234] L. A. Vermunt, “Hadronisation of Heavy Quarks; production measurements of heavy-flavour hadrons from small to large collision systems. Hadronisatie van Zware Quarks; productie metingen van zware hadronen in kleine tot grote botsingssystemen”, 2022. <https://cds.cern.ch/record/2810650>. Presented 16 May 2022.
- [235] F. Grosa, “Strange and non-strange D-meson production in pp, p-Pb, and Pb–Pb collisions with ALICE at the LHC”, 2020. <https://cds.cern.ch/record/2713513>. Presented 11 Feb 2020.
- [236] **ATLAS** Collaboration, G. Aad *et al.*, “Measurement of the differential cross-section of B^+ meson production in pp collisions at $\sqrt{s} = 7$ TeV at ATLAS”, *JHEP* **10** (2013) 042, [arXiv:1307.0126](https://arxiv.org/abs/1307.0126) [hep-ex].
- [237] **LHCb** Collaboration, R. Aaij *et al.*, “Measurement of the $B^\pm$ production cross-section in pp collisions at $\sqrt{s} = 7$ TeV”, *JHEP* **04** (2012) 093, [arXiv:1202.4812](https://arxiv.org/abs/1202.4812) [hep-ex].
- [238] **ALICE** Collaboration, S. Acharya *et al.*, “Study of flavor dependence of the baryon-to-meson ratio in proton-proton collisions at $\sqrt{s} = 13$ TeV”, [arXiv:2308.04873](https://arxiv.org/abs/2308.04873) [hep-ex].
- [239] **ALICE** Collaboration, S. Acharya *et al.*, “Measurement of beauty and charm production in pp collisions at $\sqrt{s} = 5.02$ TeV via non-prompt and prompt D mesons”, *JHEP* **05** (2021) 220, [arXiv:2102.13601](https://arxiv.org/abs/2102.13601) [nucl-ex].
- [240] F. Catalano, “Strange D-meson measurements in pp and Pb–Pb collisions exploiting machine learning with ALICE at the LHC”, 2022. <https://cds.cern.ch/record/2810132>. Presented 19 May 2022.
- [241] B. Abelev *et al.*, “ D_s^+ meson production at central rapidity in proton-proton collisions at $\sqrt{s} = 7$ TeV”, *Physics Letters B* **718** no. 2, (2012) 279–294. <https://www.sciencedirect.com/science/article/pii/S0370269312011033>.
- [242] **ALICE** Collaboration, “ $f_0(980)$ production in inelastic pp collisions at $\sqrt{s} = 5.02$ TeV”, [arXiv:2206.06216](https://arxiv.org/abs/2206.06216) [nucl-ex].
- [243] **ZEUS** Collaboration, S. Chekanov *et al.*, “Measurement of charm fragmentation ratios and fractions in photoproduction at HERA”, *Eur. Phys. J. C* **44** (2005) 351–366, [arXiv:hep-ex/0508019](https://arxiv.org/abs/hep-ex/0508019).

- [244] P. Skands, S. Carrazza, and J. Rojo, “Tuning PYTHIA 8.1: the Monash 2013 Tune”, *Eur. Phys. J. C* **74** no. 8, (2014) 3024, [arXiv:1404.5630 \[hep-ph\]](#).
- [245] **LHCb** Collaboration, R. Aaij *et al.*, “Measurements of prompt charm production cross-sections in pp collisions at $\sqrt{s} = 5$ TeV”, *JHEP* **06** (2017) 147, [arXiv:1610.02230 \[hep-ex\]](#).
- [246] S. Catani, M. Ciafaloni, and F. Hautmann, “High energy factorization and small-x heavy flavour production”, *Nuclear Physics B* **366** no. 1, (1991) 135–188.
- [247] **LHCb** Collaboration, R. Aaij *et al.*, “Prompt charm production in pp collisions at $\sqrt{s}=7$ TeV”, *Nucl. Phys. B* **871** (2013) 1–20, [arXiv:1302.2864 \[hep-ex\]](#).
- [248] **LHCb** Collaboration, R. Aaij *et al.*, “Measurements of prompt charm production cross-sections in pp collisions at $\sqrt{s} = 13$ TeV”, *JHEP* **03** (2016) 159, [arXiv:1510.01707 \[hep-ex\]](#). [Erratum: JHEP 09, 013 (2016), Erratum: JHEP 05, 074 (2017)].
- [249] M. Cacciari, M. L. Mangano, and P. Nason, “Gluon PDF constraints from the ratio of forward heavy-quark production at the LHC at $\sqrt{S} = 7$ and 13 TeV”, *Eur. Phys. J. C* **75** no. 12, (2015) 610, [arXiv:1507.06197 \[hep-ph\]](#).
- [250] **ALICE** Collaboration, S. Acharya *et al.*, “Multiplicity dependence of J/ψ production at midrapidity in pp collisions at $\sqrt{s} = 13$ TeV”, *Phys. Lett. B* **810** (2020) 135758, [arXiv:2005.11123 \[nucl-ex\]](#).
- [251] A. Chen, M. Goldberg, N. Horwitz, A. Jawahery, P. Lipari, G. C. Moneti, C. G. Trahern, H. Van Hecke, M. S. Alam, S. E. Csorna, L. Garren, M. D. Mestayer, R. S. Panvini, X. Yi, R. Giles, J. Hassard, M. Hempstead, K. Kinoshita, W. W. MacKay, F. M. Pipkin, R. Wilson, H. Kagan, P. Avery, C. Bebek, K. Berkelman, D. G. Cassel, J. W. DeWire, R. Ehrlich, T. Ferguson, R. Galik, M. G. D. Gilchriese, B. Gittelman, H. Halling, D. L. Hartill, D. Herrup, S. Holzner, M. Ito, J. Kandaswamy, V. Kistiakowsky, D. L. Kreinick, Y. Kubota, N. B. Mistry, F. Morrow, E. Nordberg, M. Ogg, R. Perchonok, R. Plunkett, A. Silverman, P. C. Stein, S. Stone, D. Weber, R. Wilcke, A. J. Sadoff, S. Behrends, K. Chadwick, J. Chauveau, P. Ganci, T. Gentile, J. M. Guida, J. A. Guida, R. Kass, A. C. Melissinos, S. L. Olsen, G. Parkhurst, D. Peterson, R. Poling, C. Rosenfeld, G. Rucinski, P. Tipton, E. H. Thorndike, J. Green, R. G. Hicks, F. Sannes, P. Skubic, A. Snyder, and R. Stone, “Evidence for the f meson at 1970 mev”, *Phys. Rev. Lett.* **51** (Aug, 1983) 634–637. <https://link.aps.org/doi/10.1103/PhysRevLett.51.634>.
- [252] **BaBar, Belle** Collaboration, A. J. Bevan *et al.*, “The Physics of the B Factories”, *Eur. Phys. J. C* **74** (2014) 3026, [arXiv:1406.6311 \[hep-ex\]](#).
- [253] **BaBar** Collaboration, P. del Amo Sanchez *et al.*, “Observation of new resonances decaying to $D\pi$ and $D^*\pi$ in inclusive e^+e^- collisions near $\sqrt{s} = 10.58$ GeV”, *Phys. Rev. D* **82** (2010) 111101, [arXiv:1009.2076 \[hep-ex\]](#).
- [254] **BaBar** Collaboration, B. Aubert *et al.*, “Observation of a New $D(s)$ Meson Decaying to DK at a Mass of $2.86\text{-GeV}/c^2$ ”, *Phys. Rev. Lett.* **97** (2006) 222001, [arXiv:hep-ex/0607082](#).

- [255] **Belle** Collaboration, J. Brodzicka *et al.*, “Observation of a new $D(sJ)$ meson in $B^+ \rightarrow \text{anti-}D^0 D^0 K^+$ decays”, *Phys. Rev. Lett.* **100** (2008) 092001, [arXiv:0707.3491 \[hep-ex\]](#).
- [256] **BaBar** Collaboration, J. P. Lees *et al.*, “Dalitz plot analyses of $B^0 \rightarrow D^- D^0 K^+$ and $B^+ \rightarrow \bar{D}^0 D^0 K^+$ decays”, *Phys. Rev. D* **91** no. 5, (2015) 052002, [arXiv:1412.6751 \[hep-ex\]](#).
- [257] **CLEO** Collaboration, D. Besson *et al.*, “Observation of a narrow resonance of mass 2.46-GeV/ c^{*2} decaying to $D^{*+}(s) \pi^0$ and confirmation of the $D^{*}(sJ)(2317)$ state”, *Phys. Rev. D* **68** (2003) 032002, [arXiv:hep-ex/0305100](#). [Erratum: *Phys.Rev.D* **75**, 119908 (2007)].
- [258] J. Alexander *et al.*, “Production and decay of the $ds1$ (2536)+”, *Physics Letters B* **303** no. 3, (1993) 377–384.
<https://www.sciencedirect.com/science/article/pii/037026939391448V>.
- [259] Z.-G. Wang, “ $D_{s3}^*(2860)$ and $D_{s1}^*(2860)$ as the $1D$ $c\bar{s}$ states”, *Eur. Phys. J. C* **75** no. 1, (2015) 25, [arXiv:1408.6465 \[hep-ph\]](#).
- [260] P. Colangelo, F. De Fazio, and S. Nicotri, “ $D(sJ)$ (2860) resonance and the $s^{*}P(1) = 5/2$ c anti- s (c anti- q) doublet”, *Phys. Lett. B* **642** (2006) 48–52, [arXiv:hep-ph/0607245](#).
- [261] L.-Y. Xiao and X.-H. Zhong, “Strong decays of higher excited heavy-light mesons in a chiral quark model”, *Phys. Rev. D* **90** no. 7, (2014) 074029, [arXiv:1407.7408 \[hep-ph\]](#).
- [262] H.-W. Ke, J.-H. Zhou, and X.-Q. Li, “Study on radiative decays of $D_{sJ}^*(2860)$ and $D_{s1}^*(2710)$ into D_s by means of LFQM”, *Eur. Phys. J. C* **75** no. 1, (2015) 28, [arXiv:1411.0376 \[hep-ph\]](#).
- [263] Q.-T. Song, D.-Y. Chen, X. Liu, and T. Matsuki, “ $D_{s1}^*(2860)$ and $D_{s3}^*(2860)$: candidates for $1D$ charmed-strange mesons”, *Eur. Phys. J. C* **75** no. 1, (2015) 30, [arXiv:1408.0471 \[hep-ph\]](#).
- [264] Q.-T. Song, D.-Y. Chen, X. Liu, and T. Matsuki, “Higher radial and orbital excitations in the charmed meson family”, *Phys. Rev. D* **92** no. 7, (2015) 074011, [arXiv:1503.05728 \[hep-ph\]](#).
- [265] Y. Sun, X. Liu, and T. Matsuki, “Newly observed $D_J(3000)^{+,0}$ and $D_J^*(3000)^0$ as $2P$ states in D meson family”, *Phys. Rev. D* **88** no. 9, (2013) 094020, [arXiv:1309.2203 \[hep-ph\]](#).
- [266] B. Zhang, X. Liu, W.-Z. Deng, and S.-L. Zhu, “ $D_{sJ}(2860)$ and $D_{sJ}(2715)$ ”, *Eur. Phys. J. C* **50** (2007) 617–628, [arXiv:hep-ph/0609013](#).
- [267] G.-L. Yu, Z.-G. Wang, and Z.-Y. Li, “Systematic analysis of the $D_J(2580)$, $D_J^*(2650)$, $D_J(2740)$, $D_J^*(2760)$, $D_J(3000)$ and $D_J^*(3000)$ in D meson family”, *Chin. Phys. C* **39** no. 6, (2015) 063101, [arXiv:1402.5955 \[hep-ph\]](#).

- [268] R.-H. Ni, Q. Li, and X.-H. Zhong, “Mass spectra and strong decays of charmed and charmed-strange mesons”, *Phys. Rev. D* **105** (Mar, 2022) 056006.
<https://link.aps.org/doi/10.1103/PhysRevD.105.056006>.
- [269] ALICE Collaboration, S. Acharya *et al.*, “Pseudorapidity distributions of charged particles as a function of mid- and forward rapidity multiplicities in pp collisions at $\sqrt{s} = 5.02, 7$ and 13 TeV”, *Eur. Phys. J. C* **81** no. 7, (2021) 630, [arXiv:2009.09434](https://arxiv.org/abs/2009.09434) [nucl-ex].
- [270] ALICE Collaboration, S. Acharya *et al.*, “Observation of a multiplicity dependence in the pT-differential charm baryon-to-meson ratios in proton–proton collisions at $\sqrt{s}=13$ TeV”, *Phys. Lett. B* **829** (2022) 137065, [arXiv:2111.11948](https://arxiv.org/abs/2111.11948) [nucl-ex].
- [271] K. Pearson, “Note on regression and inheritance in the case of two parents”, *Proceedings of the Royal Society of London* **58** (1895) 240–242.
- [272] E. Braaten, K.-m. Cheung, S. Fleming, and T. C. Yuan, “Perturbative QCD fragmentation functions as a model for heavy quark fragmentation”, *Phys. Rev. D* **51** (1995) 4819–4829, [arXiv:hep-ph/9409316](https://arxiv.org/abs/hep-ph/9409316).
- [273] ALEPH Collaboration, R. Barate *et al.*, “Study of charm production in Z decays”, *Eur. Phys. J. C* **16** (2000) 597–611, [arXiv:hep-ex/9909032](https://arxiv.org/abs/hep-ex/9909032).
- [274] H.-T. Ding, F. Karsch, and S. Mukherjee, “Thermodynamics of strong-interaction matter from Lattice QCD”, *Int. J. Mod. Phys. E* **24** no. 10, (2015) 1530007, [arXiv:1504.05274](https://arxiv.org/abs/1504.05274) [hep-lat].
- [275] G. H. Golub and C. F. van Loan, *Matrix Computations*. JHU Press, fourth ed., 2013.
<http://www.cs.cornell.edu/cv/GVL4/golubandvanloan.htm>.
- [276] ALICE Collaboration, “The ALICE definition of primary particles”,
<https://cds.cern.ch/record/2270008>.
- [277] L. Gladilin, “Fragmentation fractions of c and b quarks into charmed hadrons at LEP”, *Eur. Phys. J. C* **75** no. 1, (2015) 19, [arXiv:1404.3888](https://arxiv.org/abs/1404.3888) [hep-ex].
- [278] H1 Collaboration, A. Aktas *et al.*, “Inclusive production of D^+ , D^0 , $D^+(s)$ and D^{*+} mesons in deep inelastic scattering at HERA”, *Eur. Phys. J. C* **38** (2005) 447–459, [arXiv:hep-ex/0408149](https://arxiv.org/abs/hep-ex/0408149).
- [279] ZEUS Collaboration, H. Abramowicz *et al.*, “Measurement of charm fragmentation fractions in photoproduction at HERA”, *JHEP* **09** (2013) 058, [arXiv:1306.4862](https://arxiv.org/abs/1306.4862) [hep-ex].
- [280] ATLAS Collaboration, G. Aad *et al.*, “Measurement of $D^{*\pm}$, $D^\pm$ and $D_s^\pm$ meson production cross sections in pp collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector”, *Nucl. Phys. B* **907** (2016) 717–763, [arXiv:1512.02913](https://arxiv.org/abs/1512.02913) [hep-ex].
- [281] F. Becattini, M. Gazdzicki, and J. Sollfrank, “On chemical equilibrium in nuclear collisions”, *Eur. Phys. J. C* **5** (1998) 143–153, [arXiv:hep-ph/9710529](https://arxiv.org/abs/hep-ph/9710529).

- [282] P. Braun-Munzinger, J. Cleymans, H. Oeschler, and K. Redlich, “Maximum relative strangeness content in heavy ion collisions around 30-GeV/A”, *Nucl. Phys. A* **697** (2002) 902–912, [arXiv:hep-ph/0106066](#).
- [283] A. Wroblewski, “On the strange quark suppression factor in high energy collisions”, *Acta Physica Polonica. Series B* **16** no. 4, (1985) 379–392.
- [284] **PROSA** Collaboration, O. Zenaiev *et al.*, “Impact of heavy-flavour production cross sections measured by the LHCb experiment on parton distribution functions at low x”, *Eur. Phys. J. C* **75** no. 8, (2015) 396, [arXiv:1503.04581 \[hep-ph\]](#).
- [285] **NNPDF** Collaboration, R. D. Ball *et al.*, “Parton distributions for the LHC Run II”, *JHEP* **04** (2015) 040, [arXiv:1410.8849 \[hep-ph\]](#).
- [286] R. Averbeck, N. Bastid, Z. C. del Valle, P. Crochet, A. Dainese, and X. Zhang, “Reference Heavy Flavour Cross Sections in pp Collisions at $\sqrt{s} = 2.76$ TeV, using a pQCD-Driven s-Scaling of ALICE Measurements at $\sqrt{s} = 7$ TeV”, [arXiv:1107.3243 \[hep-ph\]](#).
- [287] M. Lisovyi, A. Verbytskyi, and O. Zenaiev, “Combined analysis of charm-quark fragmentation-fraction measurements”, *Eur. Phys. J. C* **76** no. 7, (2016) 397, [arXiv:1509.01061 \[hep-ex\]](#).
- [288] **ALICE** Collaboration, S. Acharya *et al.*, “ Λ_c^+ Production and Baryon-to-Meson Ratios in pp and p-Pb Collisions at $\sqrt{s_{NN}}=5.02$ TeV at the LHC”, *Phys. Rev. Lett.* **127** no. 20, (2021) 202301, [arXiv:2011.06078 \[nucl-ex\]](#).
- [289] **ALICE** Collaboration, S. Acharya *et al.*, “ Λ_c^+ production in pp and in p-Pb collisions at $\sqrt{s_{NN}}=5.02$ TeV”, *Phys. Rev. C* **104** no. 5, (2021) 054905, [arXiv:2011.06079 \[nucl-ex\]](#).
- [290] **ALICE** Collaboration, S. Acharya *et al.*, “First measurement of Λ_c^+ production down to pT=0 in pp and p-Pb collisions at $\sqrt{s_{NN}}=5.02$ TeV”, *Phys. Rev. C* **107** no. 6, (2023) 064901, [arXiv:2211.14032 \[nucl-ex\]](#).
- [291] **ALICE** Collaboration, S. Acharya *et al.*, “Charm-quark fragmentation fractions and production cross section at midrapidity in pp collisions at the LHC”, *Phys. Rev. D* **105** no. 1, (2022) L011103, [arXiv:2105.06335 \[nucl-ex\]](#).
- [292] **STAR** Collaboration, L. Adamczyk *et al.*, “Measurements of D^0 and D^* Production in $p + p$ Collisions at $\sqrt{s} = 200$ GeV”, *Phys. Rev. D* **86** (2012) 072013, [arXiv:1204.4244 \[nucl-ex\]](#).
- [293] **PHENIX** Collaboration, A. Adare *et al.*, “Heavy Quark Production in $p + p$ and Energy Loss and Flow of Heavy Quarks in Au+Au Collisions at $\sqrt{s_{NN}} = 200$ GeV”, *Phys. Rev. C* **84** (2011) 044905, [arXiv:1005.1627 \[nucl-ex\]](#).
- [294] H.-L. Lai, M. Guzzi, J. Huston, Z. Li, P. M. Nadolsky, J. Pumplin, and C. P. Yuan, “New parton distributions for collider physics”, *Phys. Rev. D* **82** (2010) 074024, [arXiv:1007.2241 \[hep-ph\]](#).

- [295] T.-J. Hou *et al.*, “New CTEQ Global Analysis with High Precision Data from the LHC”, [arXiv:1908.11238 \[hep-ph\]](#).
- [296] P. M. Nadolsky, H.-L. Lai, Q.-H. Cao, J. Huston, J. Pumplin, D. Stump, W.-K. Tung, and C. P. Yuan, “Implications of CTEQ global analysis for collider observables”, *Phys. Rev. D* **78** (2008) 013004, [arXiv:0802.0007 \[hep-ph\]](#).
- [297] S. Forte and S. Carrazza, “Parton distribution functions”, [arXiv:2008.12305 \[hep-ph\]](#).
- [298] D. d’Enterria and A. M. Snigirev, “Triple parton scatterings in high-energy proton-proton collisions”, *Phys. Rev. Lett.* **118** no. 12, (2017) 122001, [arXiv:1612.05582 \[hep-ph\]](#).
- [299] D. d’Enterria and A. M. Snigirev, “Triple-parton scatterings in proton–nucleus collisions at high energies”, *Eur. Phys. J. C* **78** no. 5, (2018) 359, [arXiv:1612.08112 \[hep-ph\]](#).
- [300] **ALICE** Collaboration, “Measurement of the non-prompt D-meson fraction as a function of multiplicity in proton–proton collisions at $\sqrt{s} = 13$ TeV”, [arXiv:2302.07783 \[nucl-ex\]](#).
- [301] P. Bolzoni and G. Kramer, “Inclusive charmed-meson production from bottom hadron decays at the LHC”, *J. Phys. G* **41** (2014) 075006, [arXiv:1310.2924 \[hep-ph\]](#).
- [302] T. Kneesch, B. A. Kniehl, G. Kramer, and I. Schienbein, “Charmed-meson fragmentation functions with finite-mass corrections”, *Nucl. Phys. B* **799** (2008) 34–59, [arXiv:0712.0481 \[hep-ph\]](#).
- [303] M. He and R. Rapp, “Bottom hadrochemistry in high-energy hadronic collisions”, *Phys. Rev. Lett.* **131** (Jul, 2023) 012301. <https://link.aps.org/doi/10.1103/PhysRevLett.131.012301>.
- [304] **LHCb** Collaboration, R. Aaij *et al.*, “Measurement of the $B^\pm$ production cross-section in pp collisions at $\sqrt{s} = 7$ and 13 TeV”, *JHEP* **12** (2017) 026, [arXiv:1710.04921 \[hep-ex\]](#).
- [305] D. J. Lange, “The evtgen particle decay simulation package”, *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* **462** no. 1, (2001) 152–155. <https://www.sciencedirect.com/science/article/pii/S0168900201000894>. BEAUTY2000, Proceedings of the 7th Int. Conf. on B-Physics at Hadron Machines.
- [306] **CDF** Collaboration, T. Aaltonen *et al.*, “Measurement of ratios of fragmentation fractions for bottom hadrons in $p\bar{p}$ Collisions at $\sqrt{s} = 1.96$ TeV”, *Phys. Rev. D* **77** (2008) 072003, [arXiv:0801.4375 \[hep-ex\]](#).
- [307] **ATLAS** Collaboration, G. Aad *et al.*, “Determination of the ratio of b -quark fragmentation fractions f_s/f_d in pp collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector”, *Phys. Rev. Lett.* **115** no. 26, (2015) 262001, [arXiv:1507.08925 \[hep-ex\]](#).

- [308] **LHCb** Collaboration, R. Aaij *et al.*, “Measurement of b -hadron production fractions in 7 TeV pp collisions”, *Phys. Rev. D* **85** (2012) 032008, [arXiv:1111.2357 \[hep-ex\]](#).
- [309] **Heavy Flavor Averaging Group**, **HFLAV** Collaboration, Y. S. Amhis *et al.*, “Averages of b -hadron, c -hadron, and τ -lepton properties as of 2021”, *Phys. Rev. D* **107** no. 5, (2023) 052008, [arXiv:2206.07501 \[hep-ex\]](#).
- [310] **CMS Collaboration** Collaboration, A. Tumasyan *et al.*, “Measurement of the dependence of the hadron production fraction ratios f_s/f_u and f_d/f_u on b meson kinematic variables in proton-proton collisions at $\sqrt{s} = 13$ TeV”, *Phys. Rev. Lett.* **131** (Sep, 2023) 121901. <https://link.aps.org/doi/10.1103/PhysRevLett.131.121901>.
- [311] S. Alioli, P. Nason, C. Oleari, and E. Re, “A general framework for implementing NLO calculations in shower Monte Carlo programs: the POWHEG BOX”, *JHEP* **06** (2010) 043, [arXiv:1002.2581 \[hep-ph\]](#).
- [312] **ALICE** Collaboration, S. Acharya *et al.*, “Prompt and non-prompt J/ψ production cross sections at midrapidity in proton-proton collisions at $\sqrt{s} = 5.02$ and 13 TeV”, *JHEP* **03** (2022) 190, [arXiv:2108.02523 \[nucl-ex\]](#).
- [313] **ALICE** Collaboration, B. Abelev *et al.*, “Measurement of prompt J/ψ and beauty hadron production cross sections at mid-rapidity in pp collisions at $\sqrt{s} = 7$ TeV”, *JHEP* **11** (2012) 065, [arXiv:1205.5880 \[hep-ex\]](#).
- [314] **ALICE** Collaboration, B. Abelev *et al.*, “Measurement of electrons from beauty hadron decays in pp collisions at $\sqrt{s} = 7$ TeV”, *Phys. Lett. B* **721** (2013) 13–23, [arXiv:1208.1902 \[hep-ex\]](#). [Erratum: *Phys.Lett.B* 763, 507–509 (2016)].
- [315] **ALICE** Collaboration, S. Acharya *et al.*, “Dielectron production in proton-proton collisions at $\sqrt{s} = 7$ TeV”, *JHEP* **09** (2018) 064, [arXiv:1805.04391 \[hep-ex\]](#).
- [316] **ALICE** Collaboration, S. Acharya *et al.*, “Dielectron and heavy-quark production in inelastic and high-multiplicity proton–proton collisions at $\sqrt{s_{NN}} = 13$ TeV”, *Phys. Lett. B* **788** (2019) 505–518, [arXiv:1805.04407 \[hep-ex\]](#).
- [317] **ALICE** Collaboration, S. Acharya *et al.*, “Dielectron production in proton-proton and proton-lead collisions at $\sqrt{s_{NN}} = 5.02$ TeV”, *Phys. Rev. C* **102** no. 5, (2020) 055204, [arXiv:2005.11995 \[nucl-ex\]](#).
- [318] **PHENIX** Collaboration, A. Adare *et al.*, “Measurement of Bottom versus Charm as a Function of Transverse Momentum with Electron-Hadron Correlations in pp Collisions at $\sqrt{s} = 200$ GeV”, *Phys. Rev. Lett.* **103** (2009) 082002, [arXiv:0903.4851 \[hep-ex\]](#).
- [319] **CDF** Collaboration, D. Acosta *et al.*, “Measurement of the J/ψ meson and b -hadron production cross sections in $p\bar{p}$ collisions at $\sqrt{s} = 1960$ GeV”, *Phys. Rev. D* **71** (2005) 032001, [arXiv:hep-ex/0412071](#).
- [320] **UA1** Collaboration, C. Albajar *et al.*, “Beauty production at the CERN $p\bar{p}$ collider”, *Phys. Lett. B* **256** (1991) 121–128. [Erratum: *Phys.Lett.B* 262, 497 (1991)].

- [321] **ALICE** Collaboration, S. Acharya *et al.*, “Study of flavor dependence of the baryon-to-meson ratio in proton-proton collisions at $\sqrt{s} = 13$ TeV”, *Phys. Rev. D* **108** (Dec, 2023) 112003. <https://link.aps.org/doi/10.1103/PhysRevD.108.112003>.
- [322] **LHCb** Collaboration, R. Aaij *et al.*, “Measurement of the b -quark production cross-section in 7 and 13 TeV pp collisions”, *Phys. Rev. Lett.* **118** no. 5, (2017) 052002, [arXiv:1612.05140](https://arxiv.org/abs/1612.05140) [hep-ex]. [Erratum: *Phys.Rev.Lett.* 119, 169901 (2017)].
- [323] M. Cacciari, S. Frixione, and P. Nason, “The p_T spectrum in heavy flavor photoproduction”, *JHEP* **03** (2001) 006, [arXiv:hep-ph/0102134](https://arxiv.org/abs/hep-ph/0102134).
- [324] S. Catani, S. Devoto, M. Grazzini, S. Kallweit, and J. Mazzitelli, “Bottom-quark production at hadron colliders: fully differential predictions in NNLO QCD”, *JHEP* **03** (2021) 029, [arXiv:2010.11906](https://arxiv.org/abs/2010.11906) [hep-ph].
- [325] **for the ALICE** Collaboration, R. Derradi de Souza, “Identified particle production in pp collisions at $\sqrt{s} = 7$ and 13 TeV measured with ALICE”, *J. Phys.: Conf. Ser.* **779** no. 1, (2017) 012071, [arXiv:1610.02744](https://arxiv.org/abs/1610.02744). <http://cds.cern.ch/record/2223643>. Comments: Proceedings of the Strangeness in Quark Matter Conference 2016. 4 pages, 5 figures.
- [326] “flarefly.” <https://github.com/flarefly/flarefly>.
- [327] R. A. Kycia and S. Jadach, “Relativistic voigt profile for unstable particles in high energy physics”, *Journal of Mathematical Analysis and Applications* **463** no. 2, (2018) 1040–1051. <https://www.sciencedirect.com/science/article/pii/S0022247X18302774>.
- [328] L. Lista, *Statistical Methods for Data Analysis: With Applications in Particle Physics*. Lecture Notes in Physics. Springer International Publishing, 2023. <https://books.google.it/books?id=kJJlwEACAAJ>.
- [329] A. Andronic, P. Braun-Munzinger, M. K. Köhler, A. Mazeliauskas, K. Redlich, J. Stachel, and V. Viskovic, “The multiple-charm hierarchy in the statistical hadronization model”, *JHEP* **07** (2021) 035, [arXiv:2104.12754](https://arxiv.org/abs/2104.12754) [hep-ph].
- [330] S. Godfrey and K. Moats, “Properties of excited charm and charm-strange mesons”, *Phys. Rev. D* **93** (Feb, 2016) 034035. <https://link.aps.org/doi/10.1103/PhysRevD.93.034035>.
- [331] S. Godfrey and N. Isgur, “Mesons in a relativized quark model with chromodynamics”, *Phys. Rev. D* **32** (Jul, 1985) 189–231. <https://link.aps.org/doi/10.1103/PhysRevD.32.189>.
- [332] **LHCb** Collaboration, R. Aaij *et al.*, “First observation and amplitude analysis of the $B^- \rightarrow D^+ K^- \pi^-$ decay”, *Phys. Rev. D* **91** (May, 2015) 092002. <https://link.aps.org/doi/10.1103/PhysRevD.91.092002>.
- [333] **LHCb** Collaboration, R. Aaij *et al.*, “Dalitz plot analysis of $B^0 \rightarrow \bar{d}^0 \pi^+ \pi^-$ decays”, *Phys. Rev. D* **92** (Aug, 2015) 032002. <https://link.aps.org/doi/10.1103/PhysRevD.92.032002>.

- [334] **ALICE** Collaboration, “Centrality determination in heavy ion collisions”,
<https://cds.cern.ch/record/2636623>.
- [335] S. A. et al., “Measurement of beauty-strange meson production in pb–pb collisions at $\sqrt{s_{NN}}=5.02\text{TeV}$ via non-prompt D_s^+ mesons”, *Physics Letters B* **846** (2023) 137561.
<https://www.sciencedirect.com/science/article/pii/S0370269322006955>.
- [336] **ALICE** Collaboration, S. Acharya *et al.*, “Measurement of prompt D_s^+ -meson production and azimuthal anisotropy in Pb–Pb collisions at $\sqrt{s_{NN}}=5.02\text{TeV}$ ”, *Phys. Lett. B* **827** (2022) 136986, [arXiv:2110.10006](https://arxiv.org/abs/2110.10006) [nucl-ex].
- [337] **ALICE** Collaboration, J. Adam *et al.*, “Transverse momentum dependence of D-meson production in Pb–Pb collisions at $\sqrt{s_{NN}} = 2.76\text{ TeV}$ ”, *JHEP* **03** (2016) 081, [arXiv:1509.06888](https://arxiv.org/abs/1509.06888) [nucl-ex].
- [338] M. He, R. J. Fries, and R. Rapp, “Heavy Flavor at the Large Hadron Collider in a Strong Coupling Approach”, *Phys. Lett. B* **735** (2014) 445–450, [arXiv:1401.3817](https://arxiv.org/abs/1401.3817) [nucl-th].
- [339] A. T. et al., “Observation of B_s^0 mesons and measurement of the B_s^0/B^+ yield ratio in pb–pb collisions at image 1 tev”, *Physics Letters B* **829** (2022) 137062.
<https://www.sciencedirect.com/science/article/pii/S0370269322001964>.
- [340] **ALICE** Collaboration, B. e. a. Abelev, “Upgrade of the ALICE Experiment: Letter of Intent”, tech. rep., CERN, Geneva, 2014. <https://cds.cern.ch/record/1475243>.
- [341] **ALICE** Collaboration, B. A. et al, “Technical design report for the upgrade of the alic inner tracking system”, *Journal of Physics G: Nuclear and Particle Physics* **41** no. 8, (Jul, 2014) 087002. <https://dx.doi.org/10.1088/0954-3899/41/8/087002>.
- [342] **ALICE** Collaboration, C. Lippmann, “Upgrade of the ALICE Time Projection Chamber”,.
- [343] “Technical Design Report for the Muon Forward Tracker”, tech. rep., 2015.
<https://cds.cern.ch/record/1981898>.
- [344] **ALICE** Collaboration, P. Antonioli, A. Kluge, and W. Riegler, “Upgrade of the ALICE Readout & Trigger System”, tech. rep., 2013.
- [345] P. Buncic, M. Krzewicki, and P. Vande Vyvre, “Technical Design Report for the Upgrade of the Online-Offline Computing System”, tech. rep., 2015.
<https://cds.cern.ch/record/2011297>.
- [346] L. Musa, “Letter of Intent for an ALICE ITS Upgrade in LS3”, tech. rep., CERN, Geneva, 2019. <https://cds.cern.ch/record/2703140>.
- [347] **ALICE** Collaboration, “Upgrade of the ALICE Inner Tracking System during LS3: study of physics performance”,
<https://cds.cern.ch/record/2868015>.