

CERN-EP-2023-018
14 February 2023

Measurement of the non-prompt D-meson fraction as a function of multiplicity in proton–proton collisions at $\sqrt{s} = 13$ TeV

ALICE Collaboration

Abstract

The fractions of non-prompt (i.e. originating from beauty-hadron decays) D^0 and D^+ mesons with respect to the inclusive yield are measured as a function of the charged-particle multiplicity in proton–proton collisions at a centre-of-mass energy of $\sqrt{s} = 13$ TeV with the ALICE detector at the LHC. The results are reported in intervals of transverse momentum (p_T) and integrated in the range $1 < p_T < 24$ GeV/ c . The fraction of non-prompt D^0 and D^+ mesons is found to increase slightly as a function of p_T in all the measured multiplicity intervals, while no significant dependence on the charged-particle multiplicity is observed. In order to investigate the production and hadronisation mechanisms of charm and beauty quarks, the results are compared to PYTHIA 8 as well as EPOS 3 and EPOS 4 Monte Carlo simulations, and to calculations based on the colour glass condensate including three-pomeron fusion.

arXiv:2302.07783v1 [nucl-ex] 15 Feb 2023

1 Introduction

Measurements of the production of hadrons containing heavy quarks, i.e. charm or beauty, in proton–proton (pp) collisions provide an important test of quantum chromodynamics (QCD) calculations. Several measurements of charm- and beauty-hadron production were carried out in pp collisions by the ALICE [1–11], ATLAS [12–16], CMS [17–22], and LHCb [23–31] experiments at the LHC, and by the STAR experiment at RHIC [32]. The measured D- and B-meson production cross sections are generally compatible within uncertainties with theoretical predictions based on the factorisation approach, which describe them as the convolution of the parton distribution functions (PDFs), the partonic cross section calculated with perturbative QCD (pQCD) calculations, and the fragmentation functions (FFs). Calculations of the partonic cross sections are nowadays available at next-to-leading-order accuracy (like k_T -factorisation [33–35]) or next-to-leading-order with next-to-leading logarithm resummation (like FONLL [36–38] and GM-VFNS [39–44]). The FFs are typically constrained from measurements carried out in e^+e^- or ep collisions [45], under the assumption that the hadronisation of heavy quarks into hadrons is a universal process independent of the colliding system. However, measurements of baryons containing heavy quarks at hadronic colliders showed an enhancement of the baryon-to-meson yield ratios relative to the values measured at e^+e^- colliders [30, 46], challenging the assumption of the universality of the fragmentation across different collision systems. Monte Carlo (MC) generators that implement the transition from the heavy quark to the hadron via string fragmentation (as PYTHIA 8 [47] with the Monash-13 [48] tune) or cluster hadronisation (such as HERWIG 7 [49]), in which the heavy-quark fragmentation is tuned to e^+e^- and ep measurements, cannot reproduce the baryon-to-meson yield ratios measured in pp collisions. When including the colour reconnection mechanism beyond the leading colour (CR-BLC) approximation in PYTHIA 8 [50], which introduces new colour-reconnection topologies that fragment into baryons, a much better agreement with data is obtained [8–10]. However, in case of charm baryons with strange-quark content, a significant discrepancy still remains, suggesting that additional effects should be introduced in order to have a complete description of the hadronisation processes [6, 7, 11].

Given that the production of heavy quarks occurs in initial hard partonic scattering processes while the production of light particles in the underlying event is dominated by soft processes, the measurement of heavy-flavour hadron production as a function of the charged-particle multiplicity has the potential to give insights into the interplay between the soft and hard mechanisms in particle production. In particular, multi-parton interactions (MPI) [51, 52], i.e. several hard partonic interactions occurring in a single pp collision, influence the production of light quarks and gluons, affecting the total event multiplicity, as well as the production of heavy quarks. In addition, high-multiplicity events allow one to test the heavy-flavour hadron production at small Bjorken- x , i.e. a kinematic region where the density of low-momentum gluons in the colliding protons is very high and is expected to reach saturation, which otherwise would require significantly larger energies [53]. A faster-than-linear increase has been observed in the production of prompt D mesons, as well as that of inclusive, prompt, and non-prompt (from beauty-hadron decays) J/ψ mesons at midrapidity as a function of the charged-particle multiplicity in pp collisions [54, 55]. A linear increase was instead observed in the measurement of J/ψ mesons at forward rapidity, if a pseudorapidity gap is introduced between the J/ψ mesons and the multiplicity estimator [56]. This behaviour is described by several MC generators including MPI, such as PYTHIA 8 [47] and EPOS 3 [57], and models based on initial-state effects such as the colour glass condensate (CGC) with the three-pomeron fusion mechanism [53].

It is also important to note that the charged-particle densities reached in high-multiplicity pp collisions at LHC energies are comparable with those measured in peripheral heavy-ion collisions. Measurements in high-multiplicity pp collisions showed features that resemble those associated with the formation of a colour-deconfined state of the matter called quark–gluon plasma [58, 59] in heavy-ion collisions [60–62]. In this context, one of the most interesting effects is the modification of the hadronisation mechanism.

Model calculations based on statistical hadronisation [63], which evaluate the population of hadron states according to statistical weights governed by the masses of the hadrons and a universal temperature, or hadronisation via coalescence [64, 65], namely recombination of partons close in phase space, predict an enhancement of the baryon-to-meson and strange-to-nonstrange yield ratios as a function of the charged-particle multiplicity. Recently, the ALICE Collaboration observed a multiplicity dependence of the transverse momentum (p_T) differential Λ_c^+/D^0 ratio, smoothly evolving from pp to Pb–Pb collisions. The same quantity measured p_T integrated was found not to vary significantly as a function of the charged-particle multiplicity. No modification of the D_s^+/D^0 ratio with increasing multiplicity was measured in pp collisions [66, 67]. Conversely, in the beauty sector, the LHCb Collaboration found evidence of an increase of the B_s^0/B^0 production ratio with the multiplicity, in case of charged-particle multiplicity estimated with tracks in the same pseudorapidity interval of the B mesons [68], while no measurements of beauty-baryon production as a function of charged-particle multiplicity are available. Finally, the fraction of $\chi_{c1}(3872)$ and $\psi(2S)$ states promptly produced at the collision vertex was found by the LHCb Collaboration to decrease as charged-particle multiplicity increases [69]. This suppression is interpreted as a consequence of the heavy-quark breakup via interactions with comoving hadrons [70, 71].

In this article, the first measurement of the fraction of D^0 and D^+ mesons originating from beauty-hadron decays ($f_{\text{non-prompt}}$) at midrapidity ($|y| < 0.5$) is reported as a function of the charged-particle multiplicity in pp collisions at $\sqrt{s} = 13$ TeV. In addition, the ratio between the fraction measured in different multiplicity classes divided by the one measured in the multiplicity-integrated sample is presented. The experimental apparatus and the multiplicity determination are described in Section 2. The measurement of $f_{\text{non-prompt}}$ in six transverse momentum intervals and integrated in $1 < p_T < 24$ GeV/ c is described in Section 3, while the evaluation of the systematic uncertainties is discussed in Section 4. Finally, the results are presented and compared to model calculations in Section 5.

2 Experimental apparatus and data sample

The ALICE apparatus is composed of several detectors for particle reconstruction and identification at midrapidity, embedded in a large solenoidal magnet that provides a magnetic field of $B = 0.5$ T parallel to the beams. It also includes a forward muon spectrometer ($-4 < \eta < -2.5$) and a set of forward and backward detectors for triggering and event characterisation. A comprehensive description of the ALICE detector and its performance is reported in Refs. [72, 73].

The Inner Tracking System (ITS), consisting of six cylindrical layers of silicon detectors, allows for a precise reconstruction of primary and secondary vertices, and it is used for tracking. The Time Projection Chamber (TPC) provides up to 159 space points to reconstruct the charged-particle trajectory, and provides particle identification (PID) via the measurement of the specific ionisation energy loss dE/dx of charged particles. The Time-Of-Flight detector (TOF) extends the PID capability by measuring the flight time of charged particles from the interaction point to the TOF. These detectors cover the full azimuth in the pseudorapidity interval $|\eta| < 0.9$. The V0 detector arrays, covering the intervals $2.8 < \eta < 5.1$ (V0A) and $-3.7 < \eta < -1.7$ (V0C), are used for triggering purposes and event multiplicity measurements.

The data used for this analysis are from pp collisions at $\sqrt{s} = 13$ TeV collected in 2016, 2017, and 2018. A minimum-bias (MB) trigger was used, based on coincident signals in V0A and V0C. To enrich the data sample in the highest multiplicity regions, a high-multiplicity trigger based on a minimum threshold for the V0 amplitudes (HMV0) was used as well. The data sample collected with such a trigger corresponds to the 0.17% highest-multiplicity events out of all inelastic collisions with at least one charged particle in the pseudorapidity range $|\eta| < 1$ (denoted as $\text{INEL} > 0$). Offline selections were applied to remove background from beam–gas collisions, as described in Ref. [74]. Events with multiple reconstructed primary vertices were rejected. The remaining pile-up events were at a percent level and, therefore, did not affect the present analysis. Only the events with a primary vertex reconstructed within $|z_{\text{vtx}}| < 10$ cm

Table 1: Summary of the multiplicity event classes at forward rapidity expressed in percentiles of the VOM signal amplitude ($p_{\text{VOM}}(\%)$). The average charged-particle densities $\langle dN_{\text{ch}}/d\eta \rangle_{|\eta| < 0.5}$ at midrapidity are shown, together with the value corresponding to the multiplicity-integrated class.

Multiplicity interval	$\langle dN_{\text{ch}}/d\eta \rangle_{ \eta < 0.5}$
[30, 100]%	4.41 ± 0.05
[0.1, 30]%	13.81 ± 0.14
[0, 0.1]%	31.53 ± 0.38
INEL > 0	6.93 ± 0.09

from the nominal interaction point along the beam-line direction were considered for the analysis. To select events in the INEL > 0 class, at least one track segment reconstructed with the first two ITS layers (denoted as tracklet) within the pseudorapidity region $|\eta| < 1$ was required. After these selections, the integrated luminosities are about $\mathcal{L}_{\text{int}} \approx 32 \text{ nb}^{-1}$ for the MB triggered events, and $\mathcal{L}_{\text{int}} \approx 7.7 \text{ pb}^{-1}$ for the HMOV0 triggered events [66]. The event multiplicity was determined in the forward rapidity region, exploiting the sum of signal amplitudes in the VOA and VOC scintillators, VOM, and defining its percentile distribution, p_{VOM} . Low p_{VOM} values represent high-multiplicity events. The definition of the mean multiplicity density ($\langle dN_{\text{ch}}/d\eta \rangle_{|\eta| < 0.5}$) of charged-primary particles at midrapidity is given in Ref. [75]. It was obtained by converting the measured event multiplicities as described in Ref. [74]. Table 1 summarises the multiplicity event classes at forward rapidity used in this analysis ($p_{\text{VOM}}(\%)$) and the corresponding values for $\langle dN_{\text{ch}}/d\eta \rangle_{|\eta| < 0.5}$, together with the corresponding value for the multiplicity-integrated class [74].

Monte Carlo simulations were utilised in the analysis mainly for the machine-learning training, and to obtain the correction factors for the limited detector acceptance as well as the reconstruction and selection efficiencies. They were obtained by simulating pp collisions with the PYTHIA 8.243 event generator [47, 76] (Monash-13 tune [48]). In order to enrich the simulated data samples of prompt and non-prompt D mesons, either a $c\bar{c}$ or $b\bar{b}$ quark pair was required in each simulated PYTHIA pp event and D mesons were forced to decay into the hadronic channels of interest for the analysis. The generated particles were transported through the apparatus by using the GEANT3 package [77].

3 Data analysis

The D^0 and D^+ mesons and their charge conjugates were reconstructed via the hadronic decay channels $D^0 \rightarrow K^-\pi^+$, with branching ratio $\text{BR} = (3.947 \pm 0.030)\%$, and $D^+ \rightarrow K^-\pi^+\pi^+$, with $\text{BR} = (9.38 \pm 0.16)\%$ [78]. D-meson candidates were built by combining pairs or triplets of tracks with the proper charge signs, each track with $p_{\text{T}} > 0.3 \text{ GeV}/c$, $|y| > 0.8$, at least 70 out of a maximum of 159 crossed TPC pad rows, a minimum number of two hits (out of six) in the ITS, with at least one in either of the two innermost layers, and a track fit quality $\chi^2/\text{ndf} < 2$ in the TPC. These track-selection criteria reduce the D-meson acceptance in rapidity, which falls steeply to zero for $|y| > 0.5$ at low p_{T} and for $|y| > 0.8$ at $p_{\text{T}} > 5 \text{ GeV}/c$. Thus, a fiducial acceptance selection $|y| < y_{\text{fid}}(p_{\text{T}})$, was applied to grant a uniform acceptance inside the rapidity range considered. The $y_{\text{fid}}(p_{\text{T}})$ value was defined as a second-order polynomial function, increasing from 0.5 to 0.8 in the transverse-momentum range $0 < p_{\text{T}} < 5 \text{ GeV}/c$, and as a constant term, $y_{\text{fid}} = 0.8$, for $p_{\text{T}} > 5 \text{ GeV}/c$ [2].

To suppress the large combinatorial background and to separate at the same time the contribution of prompt and non-prompt D mesons, a machine-learning approach with multi-class classification, based on Boosted Decision Trees (BDT) provided by the XGBOOST [79, 80] library was adopted. Signal samples of prompt and non-prompt D mesons for the BDT training were obtained from PYTHIA 8 simulations as described in Sec. 2. The background samples were obtained from candidates in the

sideband region in the data, i.e. in the interval $5\sigma < |\Delta M| < 9\sigma$ of the invariant mass distribution, where ΔM is the deviation between the invariant mass of the candidate and the mean of a Gaussian function describing the signal peak and σ is the Gaussian width. The training procedures are the same as reported in Ref. [2]. Before the training, loose kinematic and topological selections were applied to the D-meson candidates. The D-meson candidate information used for training the BDT models was mainly based on the displacement of the tracks from the primary vertex, the impact parameter of the D-meson daughter tracks, the distance between the D-meson decay vertex and the primary vertex, the cosine of the pointing angle between the D-meson candidate line of flight (the vector connecting the primary and secondary vertices) and its reconstructed momentum vector, and the PID information of the decay tracks. Independent BDTs were trained for each D-meson species and p_T interval in the multiplicity-integrated sample. Subsequently, the BDTs were applied to the real data sample in which the type of candidate is unknown. The BDT outputs are related to the candidate probability to be a non-prompt D meson or combinatorial background. Selections on the BDT outputs were optimised to obtain a high non-prompt D-meson fraction while maintaining a reliable signal extraction (with statistical significance larger than 5).

The signal extraction was performed in each p_T and multiplicity interval via a binned maximum-likelihood fit to the candidate invariant-mass distribution. The raw yields could be extracted in the transverse momentum interval $1 < p_T < 24$ GeV/ c and in six subranges, for both D^0 and D^+ mesons. A Gaussian function and an exponential function were used to describe the signal peak and the background distribution, respectively. To improve the stability of the fits, the widths of the D-meson signal peaks were fixed to the values extracted from data samples dominated by prompt candidates, given the naturally higher abundance of prompt compared to non-prompt D mesons. In addition, for the D^0 meson, the contribution of signal candidates to the invariant-mass distribution with the wrong decay-particle mass assignment (reflections) was included in the fit. It was parameterised by fitting the invariant-mass distribution of reflections with a double Gaussian function, and normalised according to the reflection-to-signal ratio from the PYTHIA 8 simulations. The contribution of reflections to the raw yield is about 0.5%–4%, increasing with increasing p_T . Examples of invariant-mass fits with different contribution of signal from beauty-hadron decays in the $2 < p_T < 4$ GeV/ c interval for the lowest multiplicity class and in the $1 < p_T < 24$ GeV/ c interval for the highest multiplicity class are shown in Fig. 1 and Fig. 2 for D^0 and D^+ mesons, respectively. Based on the selections on the BDT outputs, samples dominated by non-prompt (prompt) candidates were selected by requiring low probability for a candidate to be combinatorial background and a high (low) probability to be non-prompt. The invariant-mass fits from non-prompt (prompt) enhanced samples are shown in each right (left) panel, indicating the corresponding selection applied on the BDT output score related to the probability to be a non-prompt D meson.

In each p_T and multiplicity interval, the fraction of non-prompt D mesons, $f_{\text{non-prompt}}$, was estimated by sampling the raw yield with different BDT selections related to the candidate probability of being a non-prompt D meson. In this way, a set of raw yields Y_i (index i refers to a given selection on the BDT output) with different contributions from prompt and non-prompt D mesons was obtained. These raw yields are related to the corrected yields of prompt (N_{prompt}) and non-prompt ($N_{\text{non-prompt}}$) D mesons via the acceptance-times-efficiency ($\text{Acc} \times \varepsilon$) factors as

$$(\text{Acc} \times \varepsilon)_i^{\text{prompt}} \times N_{\text{prompt}} + (\text{Acc} \times \varepsilon)_i^{\text{non-prompt}} \times N_{\text{non-prompt}} - Y_i = \delta_i. \quad (1)$$

In the above equation, the δ_i term represents a residual originating from the uncertainties on Y_i , $(\text{Acc} \times \varepsilon)_i^{\text{non-prompt}}$, and $(\text{Acc} \times \varepsilon)_i^{\text{prompt}}$. The $\text{Acc} \times \varepsilon$ factors were obtained from MC simulations as described in Sec. 2. They are different for prompt and non-prompt D mesons due to the different decay topology. Since the resolution of the reconstructed primary vertex depends on the multiplicity, the simulated events were weighted to reproduce the charged-particle multiplicity distribution measured in data for events that contain D-meson candidates having an invariant mass compatible with the one of the signal. After that,

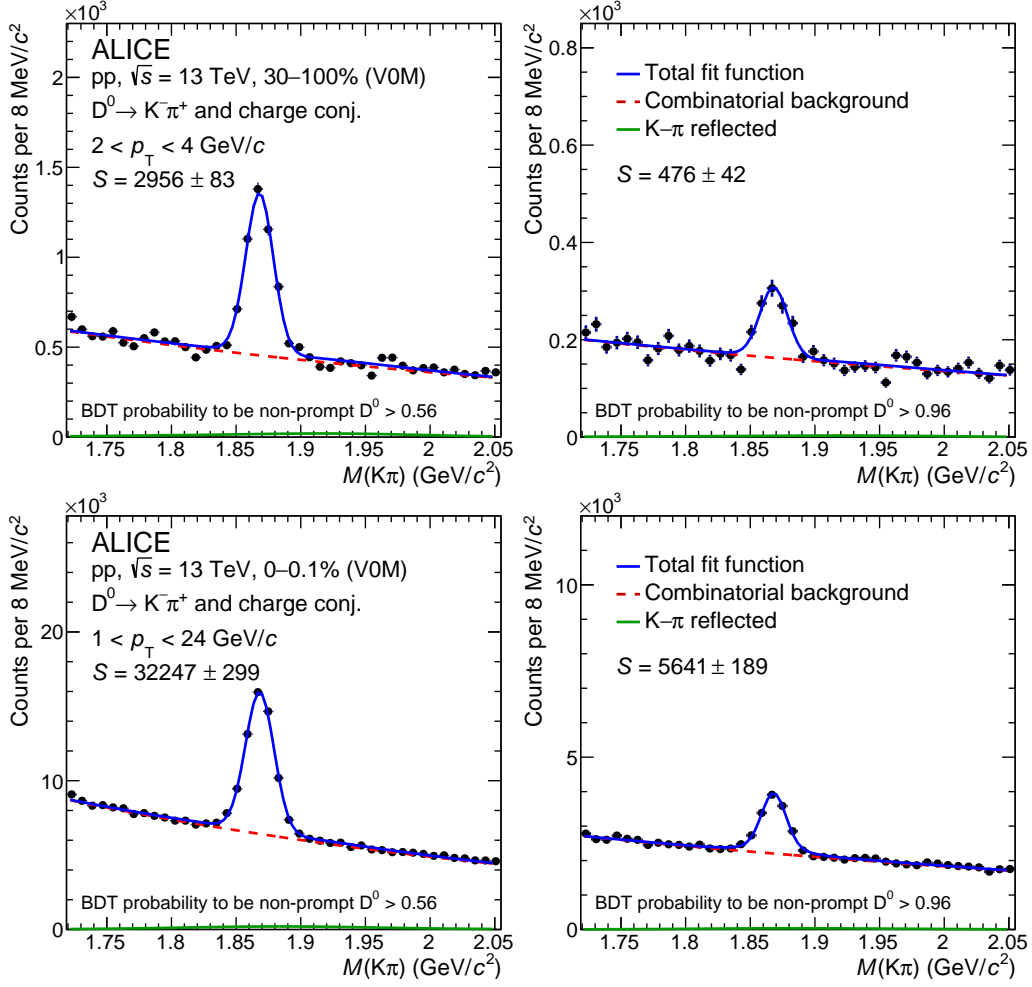


Figure 1: Invariant-mass distribution of D^0 candidates and their charge conjugates in selected p_T and multiplicity intervals. The blue solid curves show the total fit function and the red dashed curves show the combinatorial-background contribution. The green solid lines represent the reflection contribution. The raw-yield (S) values are reported together with their statistical uncertainties resulting from the fit. Top row: D^0 mesons in the $2 < p_T < 4$ GeV/c interval for the low multiplicity class. Bottom row: D^0 mesons in the $1 < p_T < 24$ GeV/c interval for the high multiplicity class. The corresponding BDT probability minimum threshold for the candidate selection is reported. The left (right) column corresponds to the prompt (non-prompt) D^0 meson candidates dominated sample.

the $\text{Acc} \times \varepsilon$ factors were computed for each BDT selection for prompt and non-prompt D mesons within the fiducial acceptance region. In the case of the number of sets of selections $n \geq 2$, a χ^2 function can be defined based on the set of equations of Eq. 1, which can be minimised to obtain N_{prompt} and $N_{\text{non-prompt}}$. More details can be found in Ref. [2]. The $N_{\text{non-prompt}}$ and N_{prompt} values can be used to calculate the corrected fraction of non-prompt D mesons as follows

$$f_{\text{non-prompt}} = \frac{N_{\text{non-prompt}}}{N_{\text{non-prompt}} + N_{\text{prompt}}}. \quad (2)$$

In addition, the ratio between the fraction of non-prompt D mesons measured in each multiplicity interval and the one measured in the INEL > 0 class of events, $f_{\text{non-prompt}}^{\text{mult}}/f_{\text{non-prompt}}^{\text{INEL}>0}$, was computed in multiplicity and p_T intervals in order to investigate the modification of the non-prompt fraction with respect to the one measured in the multiplicity-integrated sample.

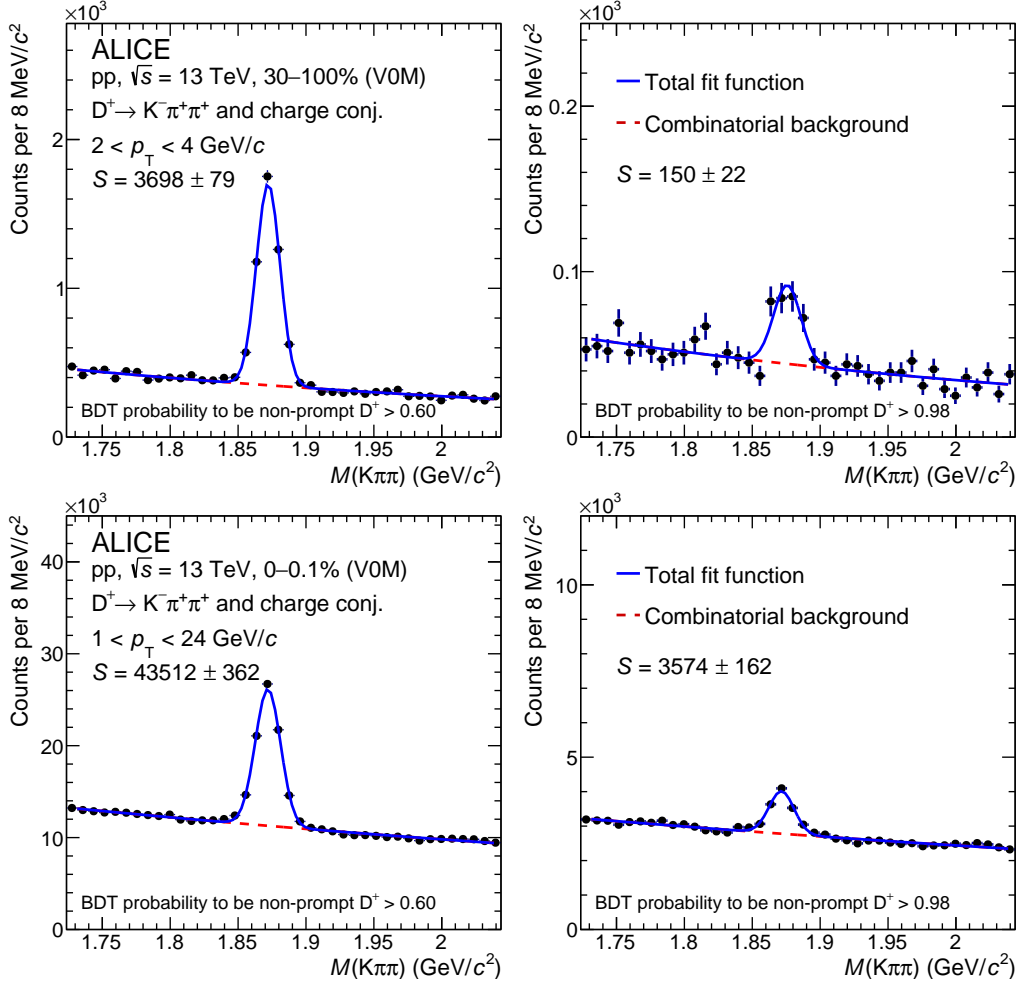


Figure 2: Invariant-mass distribution of D^+ candidates and their charge conjugates in selected p_T and multiplicity intervals. The blue solid curves show the total fit function and the red dashed curves show the combinatorial-background contribution. The raw-yield (S) values are reported together with their statistical uncertainties resulting from the fit. Top row: D^+ mesons in the $2 < p_T < 4$ GeV/ c interval for the low multiplicity class. Bottom row: D^+ mesons in the $1 < p_T < 24$ GeV/ c interval for the high multiplicity class. The corresponding BDT probability minimum threshold for the candidate selection is reported. The left (right) column corresponds to the prompt (non-prompt) D^+ meson candidates dominated sample.

Figure 3 shows an example of the raw-yield distribution as a function of the BDT-based selection used in the χ^2 -minimisation procedure for D^0 (top panels) and D^+ (bottom panels) mesons in the transverse-momentum intervals $2 < p_T < 4$ GeV/ c and $1 < p_T < 24$ GeV/ c for the low-multiplicity and high-multiplicity classes of events, respectively. The raw yield decreases with increasing minimum threshold for the probability to be a non-prompt D meson, corresponding to an increasing non-prompt D fraction. Note also that the raw yields used in this procedure are largely correlated among each other, implying that adjacent data points are expected to fluctuate in the same direction. The prompt and non-prompt components of the raw yields for each BDT-based selection obtained from the χ^2 -minimisation procedure as $(\text{Acc} \times \epsilon)_i^{\text{prompt}} \times N_{\text{prompt}}$ and $(\text{Acc} \times \epsilon)_i^{\text{non-prompt}} \times N_{\text{non-prompt}}$, are reported as the red and blue distributions, and their sum is represented by the green histogram.

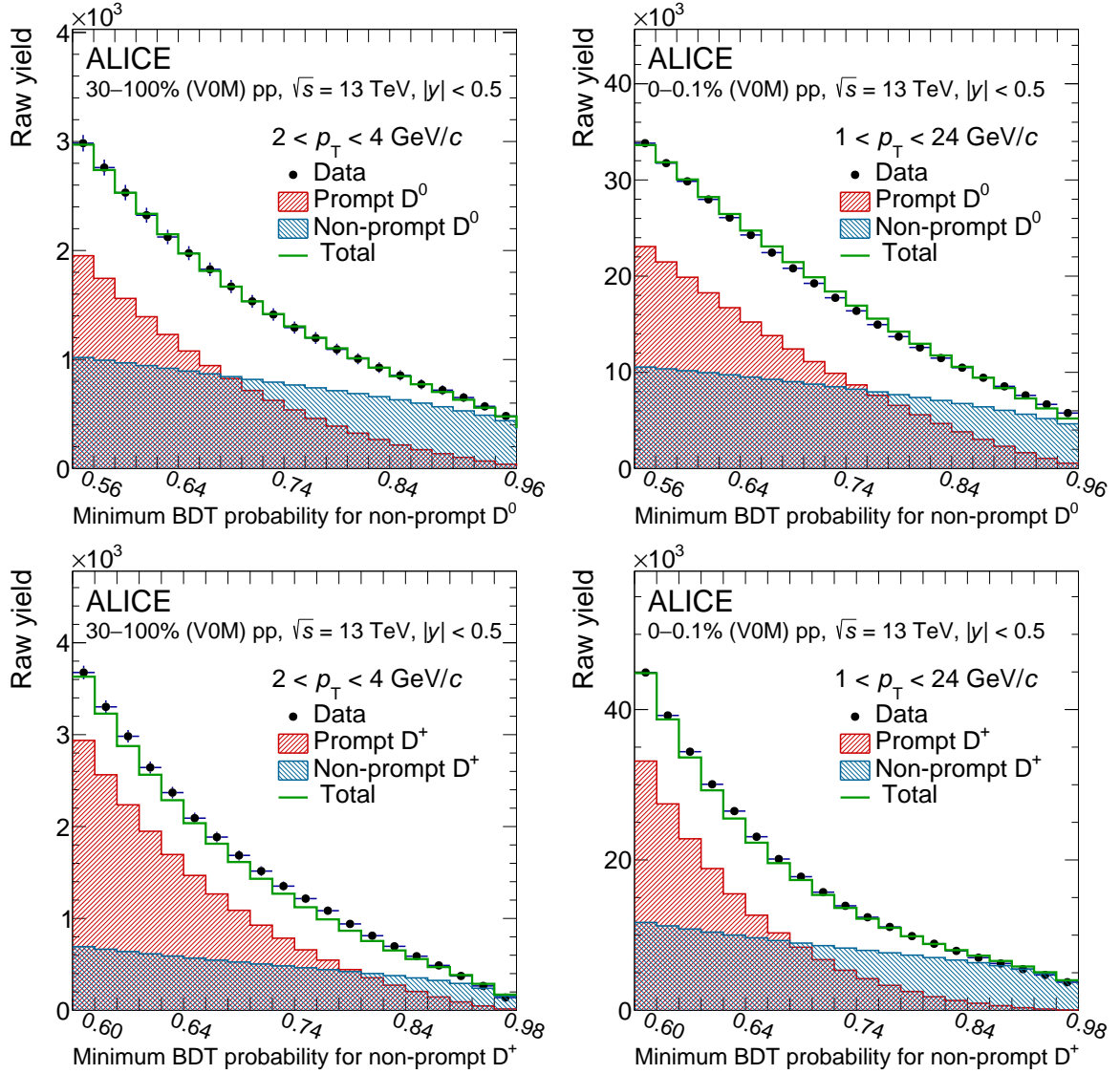


Figure 3: Examples of raw-yield distribution as a function of the BDT-based selection employed in the χ^2 -minimisation procedure adopted for the determination of $f_{\text{non-prompt}}$ of D mesons. Top row: D^0 mesons in low multiplicity (left) and high multiplicity (right) classes. Bottom row: D^+ mesons in low multiplicity (left) and high multiplicity (right) classes.

4 Systematic uncertainties

The values of systematic uncertainty on the non-prompt D-meson fraction were estimated with procedures similar to those described in Refs. [2, 66]. They include the uncertainties on (i) the raw-yield extraction from the invariant-mass distributions; (ii) the selection efficiency estimation; (iii) the dependency of the efficiency on the charged-particle multiplicity; and (iv) the D-meson p_T shape in the simulation. The estimated values of the systematic uncertainties for some representative p_T intervals of D^0 and D^+ mesons are summarised in Table 2.

The systematic uncertainty of the raw-yield extraction was evaluated by repeating the fits to the invariant-mass distribution varying the fit range and the functional form of the background and signal fit functions. To further test the sensitivity to the line shape of the signal, a bin-counting method, in which the signal yield was obtained by integrating the background-subtracted invariant-mass distribution within the $\pm 3\sigma$

Table 2: Summary of the relative systematic uncertainties on the non-prompt D^0 , D^+ -meson fractions in various p_T and multiplicity intervals.

Meson p_T (GeV/c)	D^0		D^+		D^0		D^+	
	2–4	8–12	2–4	8–12	1–2	12–24	1–2	12–24
	$f_{\text{non-prompt}}^{30-100\%}$				$f_{\text{non-prompt}}^{0-0.1\%}$			
Raw-yield extraction	2%	2%	3%	3%	5%	2%	6%	6%
Efficiency estimation	2%	2%	4%	5%	6%	3%	5%	5%
MC multiplicity distribution	2%	1%	4%	0%	0%	0%	0%	0%
MC D-meson p_T distribution	6%	3%	3%	1%	9%	3%	8%	3%
	$f_{\text{non-prompt}}^{30-100\%} / f_{\text{non-prompt}}^{\text{INEL}>0}$				$f_{\text{non-prompt}}^{0-0.1\%} / f_{\text{non-prompt}}^{\text{INEL}>0}$			
Raw-yield extraction	2%	2%	4%	4%	4%	2%	7%	5%
Efficiency estimation	2%	3%	4%	5%	4%	2%	4%	3%
MC multiplicity distribution	0%	2%	0%	0%	1%	0%	4%	0%
MC D-meson p_T distribution	4%	2%	5%	1%	0%	0%	1%	1%

region relative to the peak position, was used. In the case of D^0 mesons, an additional contribution due to signal reflections in the invariant-mass distribution was estimated by varying the normalisation and the shape of the templates used for the reflections in the invariant-mass fits. The systematic uncertainty was defined as the RMS of the distribution of the resulting $f_{\text{non-prompt}}$ obtained from all these variations and ranges from 2% to 6% depending on the D-meson species, multiplicity, and p_T interval.

The systematic uncertainty of the selection-efficiency determination, arising from possible imperfections of the description of the decay topologies or the detector resolution in the simulation, was estimated by using alternative sets of BDT-output selections for the procedure described in Section 3. In particular, stricter and looser selections were tested, as well as different combinations of selections adopted to define the system of equations described in Eq. 1. A systematic uncertainty ranging from 2% to 6% was assigned.

To estimate the systematic uncertainty on the sensitivity of the efficiency on the charged-particle multiplicity, due to the multiplicity dependence of the primary-vertex reconstruction resolution, the distribution of the number of tracklets in the MC simulation for each VOM class of events was weighted using the one obtained in the real data considering events containing a D-meson candidate, without requiring the invariant-mass region selection. The resulting effect on the $f_{\text{non-prompt}}$ estimation ranges from 0% to 4%.

The systematic uncertainty on the efficiency calculation due to a possible difference between the real and simulated D-meson transverse-momentum distributions was estimated by evaluating the efficiency after reweighting the p_T shape from the PYTHIA 8 generator to match the one from FONLL calculations, in addition to the reweighting of the multiplicity distribution mentioned above. The weights were applied to the p_T distributions of prompt D mesons and to the parent beauty-hadron p_T distributions in case of non-prompt D mesons. The assigned uncertainty ranges from 1% to 9%.

The aforementioned sources of systematic uncertainty were assumed to be uncorrelated among each other. The total systematic uncertainty is defined as the square root of the quadratic sum of the estimated values in each p_T and multiplicity interval. In order to assess the correlation between the systematic uncertainties on $f_{\text{non-prompt}}$ in the different multiplicity intervals with respect to the one in the $\text{INEL} > 0$

sample, the effect of the variations and the estimation of the uncertainties were directly evaluated on the ratio $f_{\text{non-prompt}}^{\text{mult}}/f_{\text{non-prompt}}^{\text{INEL}>0}$.

5 Results

The measured fractions of D-mesons originating from beauty-hadron decays, $f_{\text{non-prompt}}$, in pp collisions at $\sqrt{s} = 13$ TeV are shown in Fig. 4 as a function of p_T . The results are reported in different panels for D^0 (left) and D^+ (right) mesons and for the INEL > 0 class (top panels) and the three multiplicity classes of events (lower panels). The statistical and total systematic uncertainties are shown by vertical error bars and boxes, respectively. In all the event classes and for both D^0 and D^+ mesons, $f_{\text{non-prompt}}$ increases with p_T from 5%–7% to about 10%. This increase is motivated by the harder p_T distribution of beauty hadrons compared to the charm ones, which is only partly compensated by the $b \rightarrow D + X$ decay kinematics [2, 38]. The fraction of non-prompt D^0 mesons is slightly larger than that of D^+ mesons, as a consequence of the different branching ratios of B mesons with a D^0 or D^+ meson in the final state, and of the different charm-quark fragmentation fractions for the prompt D-meson production. This increasing trend is expected from pQCD calculations, as shown in Ref. [44]. Measurements are compared to predictions from the PYTHIA 8 [47, 81] and EPOS [57, 82] event generators. PYTHIA 8 simulations were obtained using the standard Monash 2013 tune [48] as well as with colour reconnection settings beyond-leading-colour approximation (CR-BLC Mode 0, 2, and 3) [50], and with colour ropes [81] using PYTHIA version 8.307. Both the versions 3.448 and 4.0.0 of the EPOS MC generator were tested. In EPOS 4, parallel partonic scatterings based on the S-matrix theory are implemented, leading to the factorisation of the hard and soft scales, particularly important for heavy quarks. This factorisation allows the computation of the PDFs within the EPOS framework itself. In both EPOS versions, the hadronisation is implemented via a microcanonical ensemble, consisting in the decay of plasma droplets which conserves energy, momentum, and flavour. The EPOS predictions presented in this paper do not include a hydrodynamic expansion of the system. However, the results were found not to significantly change if the latter is included. Following what was done for data, all PYTHIA 8 and EPOS simulations were selected according to percentiles of the INEL > 0 cross section based on the charged-particle multiplicity counts in the ALICE V0A and V0C acceptance. While all models qualitatively reproduce the increase of $f_{\text{non-prompt}}$ with increasing p_T , EPOS significantly underpredicts $f_{\text{non-prompt}}$ of D^+ mesons and D^0 mesons in the INEL > 0 and in the lowest multiplicity classes of events, by up to a factor of two. Moreover, EPOS 3 predicts a slightly stronger multiplicity dependence compared to EPOS 4. On the other hand, PYTHIA 8 is generally closer to data but overpredicts $f_{\text{non-prompt}}$ by approximately 20–30%. No significant difference in the various PYTHIA 8 settings tested in this work is observed, with the exception of the ‘CR-BLC Mode 3’ setting (see Ref. [50] for a detailed description), which predicts a lower D^+ and D^0 non-prompt fraction especially in the two highest multiplicity intervals, providing a better description of the data.

The ratio of the D-meson non-prompt fractions in the multiplicity classes relative to that in the INEL > 0 class, $f_{\text{non-prompt}}^{\text{mult}}/f_{\text{non-prompt}}^{\text{INEL}>0}$, is shown in Fig. 5 as a function of transverse momentum for the three multiplicity classes. This double ratio isolates the relative variation of $f_{\text{non-prompt}}$ as a function of the charged particle multiplicity from absolute scaling factors. The double ratio of D^0 and D^+ was found to be compatible for all the multiplicity classes as expected. In order to improve the statistical precision, the average D^0 and D^+ $f_{\text{non-prompt}}^{\text{mult}}/f_{\text{non-prompt}}^{\text{INEL}>0}$ was computed. The average was computed using the inverse of the quadratic sum of the relative statistical and uncorrelated systematic uncertainties as weights. The systematic uncertainties were propagated through the averaging procedure considering the contributions from the raw-yield extraction and the selection efficiency as uncorrelated, while the other sources as fully correlated between the two D-meson species. In all multiplicity classes, the measured ratio is compatible with unity within uncertainties. This finding suggests similar production mechanisms of charm and beauty quarks as a function of multiplicity. The expectation obtained with

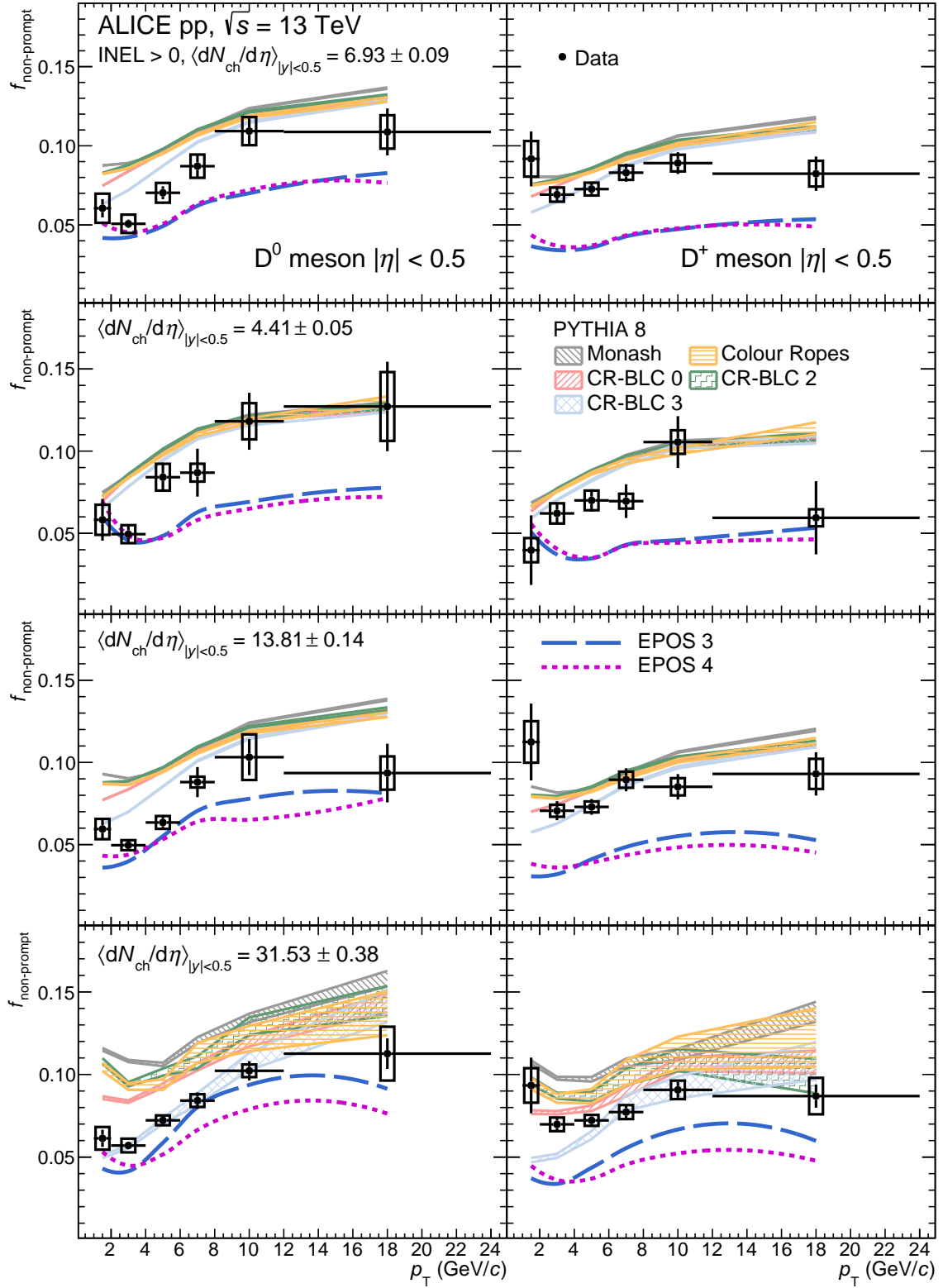


Figure 4: Fractions of non-prompt D^0 (left column) and D^+ (right column) mesons as a function of p_T for the INEL > 0 class and the three multiplicity classes of events in pp collisions at $\sqrt{s} = 13$ TeV. The measurements are compared with the predictions obtained with PYTHIA 8 [50] and EPOS [57] event generators.

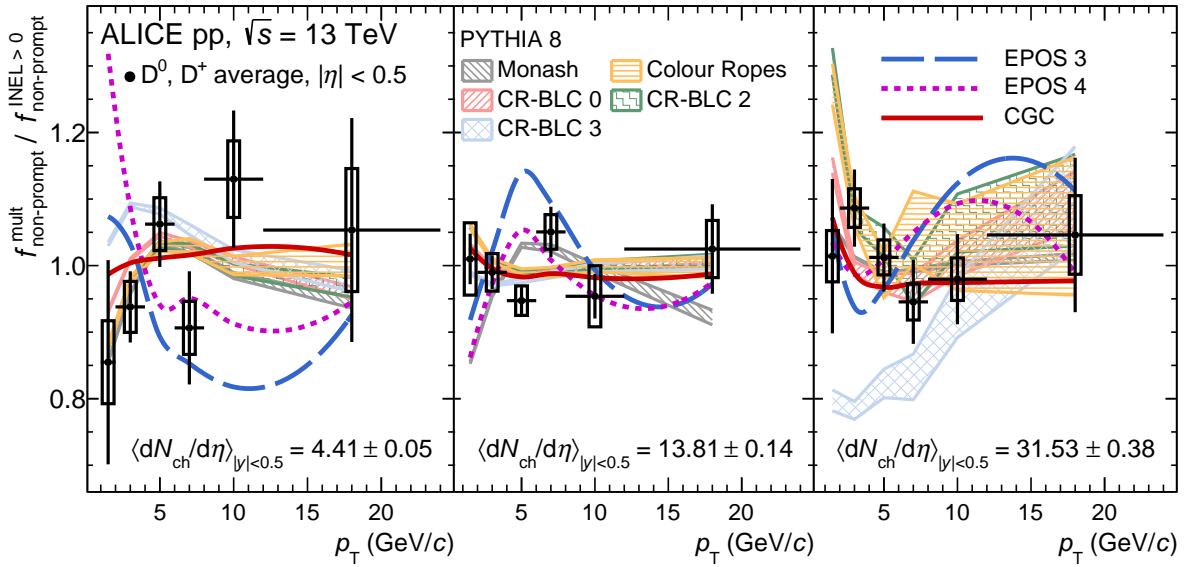


Figure 5: Average fractions of non-prompt D^0 and D^+ mesons as a function of p_T for different multiplicity intervals normalised to the one measured in the INEL > 0 class of pp collisions at $\sqrt{s} = 13$ TeV. The measurements are compared with the predictions obtained with PYTHIA 8 [50] and EPOS [57] event generators and the CGC model.

EPOS 3 shows a modification of the p_T spectrum different for charm and beauty hadrons due to their different mass, which is not supported by the measurement. A qualitatively similar but weaker behaviour is obtained with EPOS 4, which is more in agreement with the data. All the PYTHIA 8 configurations reproduce the measurements within the uncertainties, indicating a small influence of the hadronisation in the multiplicity dependence, except for the CR-BLC Mode 3 setting, which underestimates the data at low p_T in the high-multiplicity class of events. The data points are further compared to a CGC model that includes the three-pomeron exchange mechanism [53]. In this model, the transition from the beauty quark to the charm hadron is modelled in a single step using $f(b \rightarrow H_c)$ fragmentation functions measured in e^+e^- collisions [83]. Even though these fragmentation functions were shown to be unable to reproduce the measured cross sections of non-prompt D mesons in previous studies [2], the CGC predictions indicate a very weak multiplicity dependence of the double ratio and are consistent with the data within uncertainties.

The specific multiplicity dependence of $f_{\text{non-prompt}}^{\text{mult}}/f_{\text{non-prompt}}^{\text{INEL}>0}$ can be studied in more detail by plotting the values obtained in each individual transverse momentum interval as a function of the charged-particle multiplicity density normalised to the value corresponding to the INEL > 0 class of events, as shown in Fig. 6. In all p_T intervals, the average D^0 and D^+ $f_{\text{non-prompt}}^{\text{mult}}/f_{\text{non-prompt}}^{\text{INEL}>0}$ ratio is found to be compatible with unity, indicating a weak (if any) dependence of $f_{\text{non-prompt}}$ with the charged-particle multiplicity. Comparisons with models reveal that the EPOS event generator predicts a multiplicity dependence at intermediate transverse momentum ($4 < p_T < 6$ GeV/c) which is ruled out by the data. At lower and higher p_T , instead, EPOS predicts a milder charged-particle multiplicity dependence and is hence closer to the data. Moreover, the multiplicity-independence of CGC predictions is also consistent with the data. Finally, most PYTHIA 8 predictions are consistent with the data, with the notable exception of CR-BLC Mode 3 results, in which the double ratio is shown to decrease with multiplicity. This behaviour can be understood by isolating the double ratio for D mesons originating from beauty-meson and beauty-baryon decays in each of the specific PYTHIA 8 configurations being used, as represented in Fig. 7. While in all cases the $f_{\text{non-prompt}}^{\text{mult}}/f_{\text{non-prompt}}^{\text{INEL}>0}$ ratio from beauty baryons increases systematically with multiplicity, the Mode 3 setting results in a decrease of this double ratio for D mesons originating

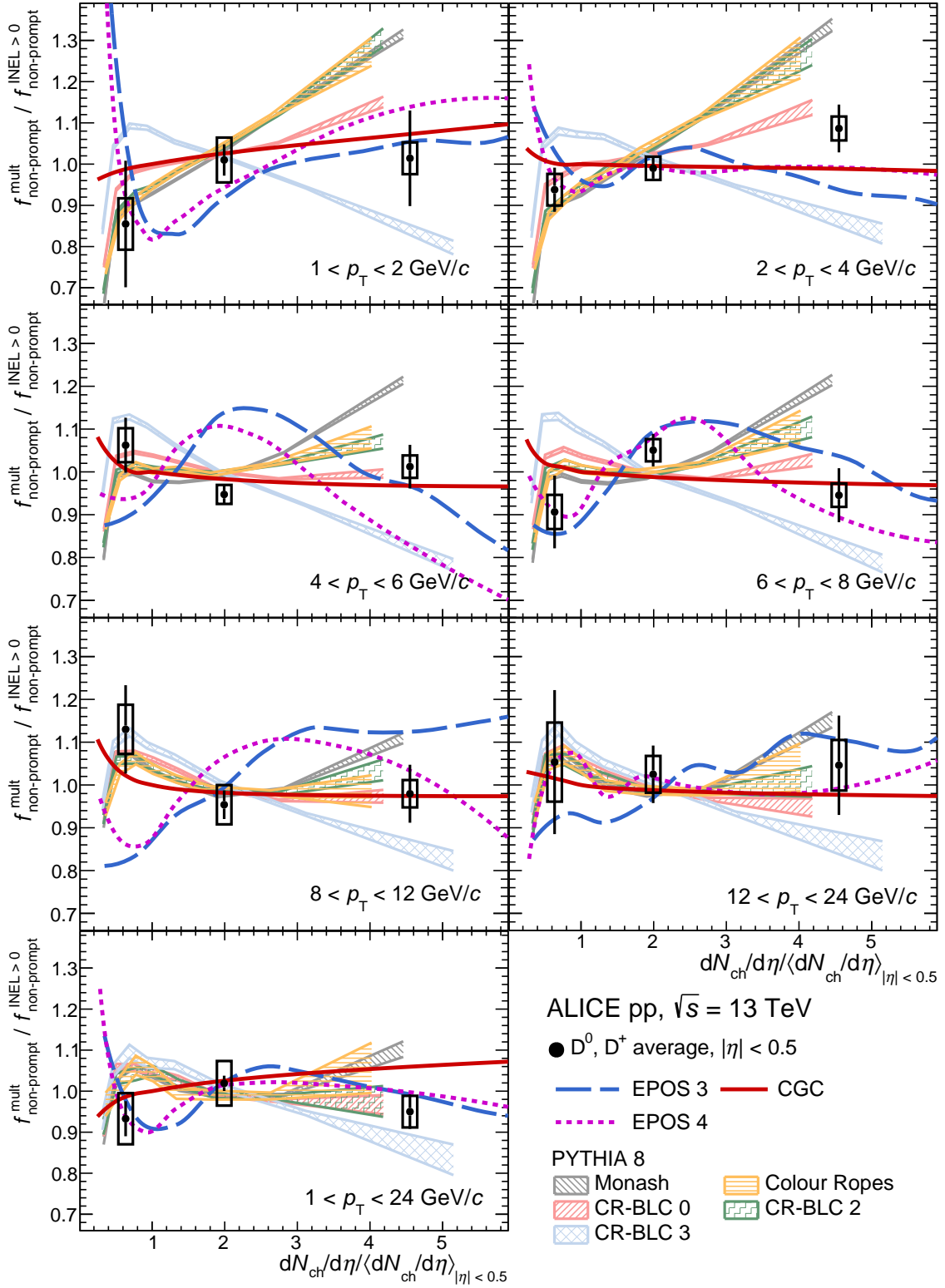


Figure 6: Average fractions of non-prompt D^0 and D^+ mesons as a function of multiplicity, both normalised to the value corresponding to the $\text{INEL} > 0$ class, for pp collisions at $\sqrt{s} = 13$ TeV in different p_T intervals and integrated in $1 < p_T < 24$ GeV/c. The measurements are compared with predictions obtained with the PYTHIA 8 [50] and EPOS [57] event generators and the CGC model [53].

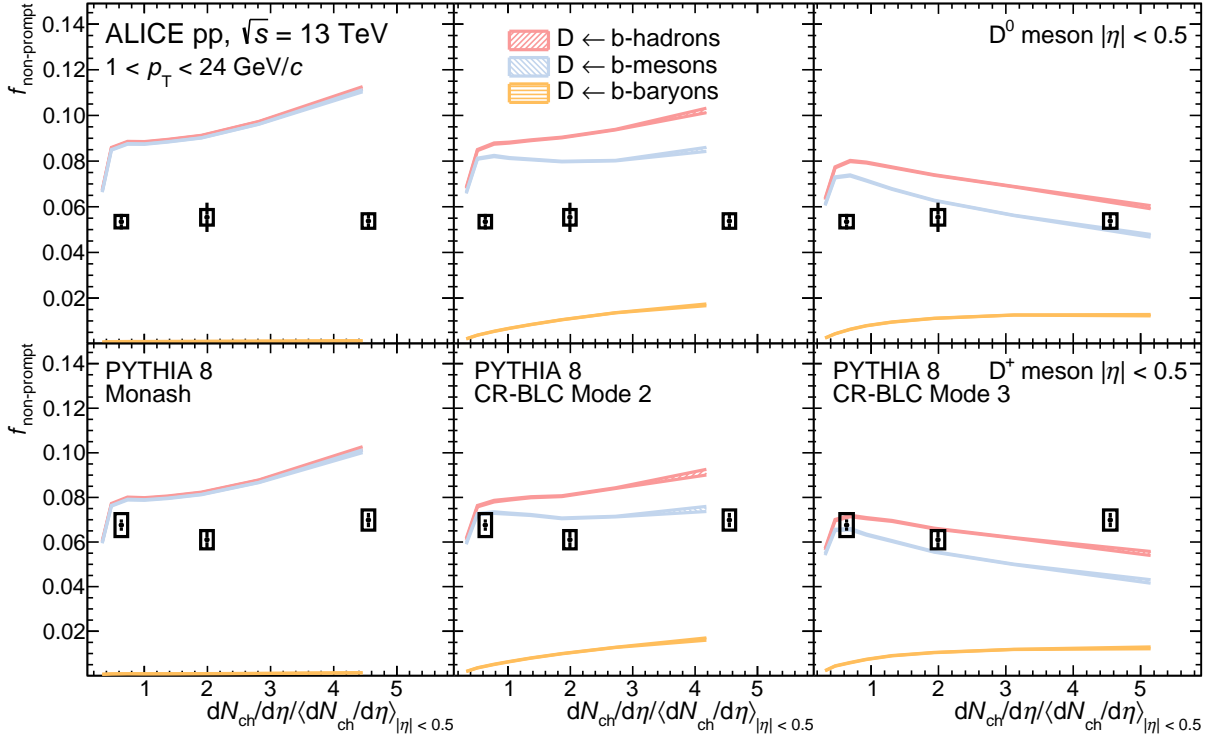


Figure 7: Fractions of non-prompt D^0 (first row) and D^+ (second row) mesons in $1 < p_T < 24$ GeV/c as a function of multiplicity for pp collisions at $\sqrt{s} = 13$ TeV compared with predictions obtained with the PYTHIA 8 [50] event generator. The contributions from beauty meson and baryon decays in PYTHIA 8 are displayed separately.

from B-meson decays. More specifically, a clean MC-only test can be performed with the beyond-leading-colour tunes by calculating the ratio of baryons and mesons at hadronisation time in PYTHIA 8 as a function of multiplicity in each model, as depicted in Fig. 8. Notably, CR-BLC Mode 3 differs from other PYTHIA 8 predictions due to the fact that, in that case, beauty quarks produce significantly more baryons, and charm quarks produce fewer baryons than in other cases. Consequently, the fraction of non-prompt D mesons decreases with the multiplicity as a combination of two effects. On the one side, charm quarks hadronise more to D mesons, increasing the prompt contribution to the D-meson production and, on the other side, beauty quarks will tend towards being contained in baryons, which in turn will feed preferentially into charm baryons such as the Λ_c^+ baryon. This strong preference towards beauty baryons is not favoured by current ALICE data, which essentially rules out the CR-BLC Mode 3 dynamics in favour of models in which $f_{\text{non-prompt}}$ tends to either remain constant or increase slightly with multiplicity. Future studies of meson and baryon production in the beauty sector as a function of the charged-particle multiplicity will allow for firmer conclusions.

6 Summary

The fractions of the D^0 and D^+ mesons originating from beauty-hadron decays, $f_{\text{non-prompt}}$, were measured at midrapidity ($|y| < 0.5$) in pp collisions at $\sqrt{s} = 13$ TeV in events with at least a charged particle at midrapidity (INEL > 0 class of events) and as a function of charged-particle multiplicity and transverse momentum. Events with different charged-particle multiplicities were selected as percentiles of the INEL > 0 cross section based on the charged-particle multiplicity counts in the ALICE V0A and V0C at forward and backward rapidity. The D^+ and D^0 $f_{\text{non-prompt}}$ were observed to slightly increase from about 5%–7% for $1 < p_T < 3$ GeV/c to about 10% for $8 < p_T < 24$ GeV/c. The ratios $f_{\text{non-prompt}}^{\text{mult}}/f_{\text{non-prompt}}^{\text{INEL}>0}$

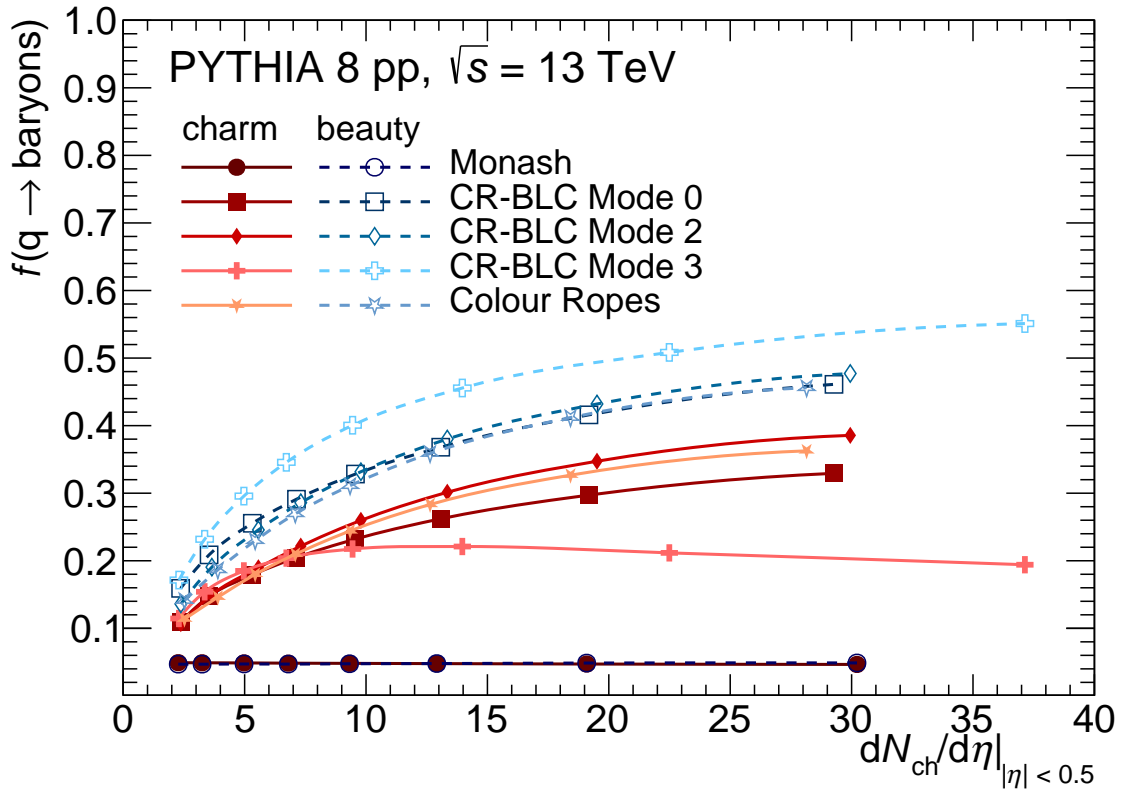


Figure 8: Fraction of charm and beauty quarks hadronising to baryons as a function of the charged particle multiplicity at midrapidity in PYTHIA 8 [50] simulations with different tunes.

are compatible with unity both as a function of p_T and charged-particle multiplicity, suggesting either no or only a mild multiplicity dependence. This finding suggests a similar production mechanism of charm and beauty quarks as a function of multiplicity.

The measured $f_{\text{non-prompt}}$ values are compared to predictions obtained with different MC generators. The EPOS 3 and EPOS 4 generators tend to underestimate the measurements, while PYTHIA 8 with different tunes, including the colour reconnection mechanism beyond leading colour approximation and colour ropes, slightly overestimates the data. The variation of $f_{\text{non-prompt}}$ with multiplicity is satisfactorily described by the MC simulations except for the $4 < p_T < 6$ GeV/ c interval, where the EPOS generator predicts a significant increase. In all the considered p_T intervals, the CR-BLC Mode 3 tune of PYTHIA 8 foresees a decrease at high multiplicity. In that tune, this decrease with increasing multiplicity is motivated by an interplay between an increased fraction of charm quarks hadronising into mesons and an increased fraction of beauty quarks hadronising into baryons and is not favoured by data. The ratio $f_{\text{non-prompt}}^{\text{mult}}/f_{\text{non-prompt}}^{\text{INEL}>0}$ is also described well by the CGC model.

The measurements presented in this paper provide an important test for production and hadronisation models in the charm and beauty sectors, and they pave the way for future studies of beauty-hadron production in pp collisions as a function of the charged-particle multiplicity.

References

- [1] 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** (2019) 388, arXiv:1901.07979 [nucl-ex].

- [2] **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 [nucl-ex].
- [3] **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].
- [4] **ALICE** Collaboration, S. Acharya *et al.*, “Measurement of electrons from semileptonic heavy-flavour hadron decays at midrapidity in pp and Pb–Pb collisions at $\sqrt{s_{NN}} = 5.02$ TeV”, *Phys. Lett. B* **804** (2020) 135377, arXiv:1910.09110 [nucl-ex].
- [5] **ALICE** Collaboration, S. Acharya *et al.*, “Production of muons from heavy-flavour hadron decays in pp collisions at $\sqrt{s} = 5.02$ TeV”, *JHEP* **09** (2019) 008, arXiv:1905.07207 [nucl-ex].
- [6] **ALICE** Collaboration, S. Acharya *et al.*, “Measurement of the production cross section of prompt Ξ_c^0 baryons at midrapidity in pp collisions at $\sqrt{s} = 5.02$ TeV”, *JHEP* **10** (2021) 159, arXiv:2105.05616 [nucl-ex].
- [7] **ALICE** Collaboration, S. Acharya *et al.*, “Measurement of the Cross Sections of Ξ_c^0 and Ξ_c^+ Baryons and of the Branching-Fraction Ratio $BR(\Xi_c^0 \rightarrow \Xi^- e^+ \nu_e)/BR(\Xi_c^0 \rightarrow \Xi^- \pi^+)$ in pp collisions at 13 TeV”, *Phys. Rev. Lett.* **127** (2021) 272001, arXiv:2105.05187 [nucl-ex].
- [8] **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** (2021) 202301, arXiv:2011.06078 [nucl-ex].
- [9] **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** (2021) 054905, arXiv:2011.06079 [nucl-ex].
- [10] **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** (2022) 012001, arXiv:2106.08278 [hep-ex].
- [11] **ALICE** Collaboration, “First measurement of Ω_c^0 production in pp collisions at $\sqrt{s} = 13$ TeV”, arXiv:2205.13993 [nucl-ex].
- [12] **ATLAS** Collaboration, G. Aad *et al.*, “Measurement of the b-hadron production cross section using decays to $D^* \mu^- X$ final states in pp collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector”, *Nucl. Phys. B* **864** (2012) 341–381, arXiv:1206.3122 [hep-ex].
- [13] **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 [hep-ex].
- [14] **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 [hep-ex].
- [15] **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** (2015) 262001, arXiv:1507.08925 [hep-ex].
- [16] **ATLAS** Collaboration, M. Aaboud *et al.*, “Measurement of the relative B_c^\pm/B^\pm production cross section with the ATLAS detector at $\sqrt{s} = 8$ TeV”, *Phys. Rev. D* **104** (2021) 012010, arXiv:1912.02672 [hep-ex].

- [17] **CMS** Collaboration, S. Chatrchyan *et al.*, “Measurement of the cross section for production of bb^- bar X , decaying to muons in pp collisions at $\sqrt{s} = 7$ TeV”, *JHEP* **06** (2012) 110, arXiv:1203.3458 [hep-ex].
- [18] **CMS** Collaboration, A. M. Sirunyan *et al.*, “Nuclear modification factor of D^0 mesons in PbPb collisions at $\sqrt{s_{NN}} = 5.02$ TeV”, *Phys. Lett. B* **782** (2018) 474–496, arXiv:1708.04962 [nucl-ex].
- [19] **CMS** Collaboration, A. M. Sirunyan *et al.*, “Production of Λ_c^+ baryons in proton-proton and lead-lead collisions at $\sqrt{s_{NN}} = 5.02$ TeV”, *Phys. Lett. B* **803** (2020) 135328, arXiv:1906.03322 [hep-ex].
- [20] **CMS** Collaboration, V. Khachatryan *et al.*, “Measurement of the total and differential inclusive B^+ hadron cross sections in pp collisions at $\sqrt{s} = 13$ TeV”, *Phys. Lett. B* **771** (2017) 435–456, arXiv:1609.00873 [hep-ex].
- [21] **CMS** Collaboration, A. M. Sirunyan *et al.*, “Studies of Beauty Suppression via Nonprompt D^0 Mesons in Pb-Pb Collisions at $Q^2 = 4$ GeV²”, *Phys. Rev. Lett.* **123** (2019) 022001, arXiv:1810.11102 [hep-ex].
- [22] **CMS** Collaboration, V. Khachatryan *et al.*, “Measurement of the ratio of the production cross sections times branching fractions of $B_c^\pm \rightarrow J/\psi\pi^\pm$ and $B^\pm \rightarrow J/\psi K^\pm$ and $\mathcal{B}(B_c^\pm \rightarrow J/\psi\pi^\pm\pi^\mp)/\mathcal{B}(B_c^\pm \rightarrow J/\psi\pi^\pm)$ in pp collisions at $\sqrt{s} = 7$ TeV”, *JHEP* **01** (2015) 063, arXiv:1410.5729 [hep-ex].
- [23] **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].
- [24] **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)].
- [25] **LHCb** Collaboration, R. Aaij *et al.*, “Measurement of B_c^+ production in proton-proton collisions at $\sqrt{s} = 8$ TeV”, *Phys. Rev. Lett.* **114** (2015) 132001, arXiv:1411.2943 [hep-ex].
- [26] **LHCb** Collaboration, R. Aaij *et al.*, “Study of the production of Λ_b^0 and \bar{B}^0 hadrons in pp collisions and first measurement of the $\Lambda_b^0 \rightarrow J/\psi p K^-$ branching fraction”, *Chin. Phys. C* **40** (2016) 011001, arXiv:1509.00292 [hep-ex].
- [27] **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** (2017) 052002, arXiv:1612.05140 [hep-ex]. [Erratum: *Phys. Rev. Lett.* **119**, (2017) 169901].
- [28] **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].
- [29] **LHCb** Collaboration, R. Aaij *et al.*, “Measurement of the mass and production rate of Ξ_b^- baryons”, *Phys. Rev. D* **99** (2019) 052006, arXiv:1901.07075 [hep-ex].
- [30] **LHCb** Collaboration, R. Aaij *et al.*, “Measurement of b hadron fractions in 13 TeV pp collisions”, *Phys. Rev. D* **100** (2019) 031102, arXiv:1902.06794 [hep-ex].
- [31] **LHCb** Collaboration, R. Aaij *et al.*, “Measurement of Ξ_{cc}^{++} production in pp collisions at $\sqrt{s} = 13$ TeV”, *Chin. Phys. C* **44** (2020) 022001, arXiv:1910.11316 [hep-ex].

- [32] **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].
- [33] R. Maciula and A. Szczurek, “Open charm production at the LHC - k_T -factorization approach”, *Phys. Rev. D* **87** (2013) 094022, arXiv:1301.3033 [hep-ph].
- [34] R. Maciula and A. Szczurek, “Production of Λ_c baryons at the LHC within the k_T -factorization approach and independent parton fragmentation picture”, *Phys. Rev. D* **98** (2018) 014016, arXiv:1803.05807 [hep-ph].
- [35] B. Guiot, “Heavy-quark production with k_T -factorization: The importance of the sea-quark distribution”, *Phys. Rev. D* **99** (2019) 074006, arXiv:1812.02156 [hep-ph].
- [36] M. Cacciari, M. Greco, and P. Nason, “The p_T spectrum in heavy flavor hadroproduction”, *JHEP* **05** (1998) 007, arXiv:hep-ph/9803400.
- [37] M. Cacciari, S. Frixione, and P. Nason, “The p_T spectrum in heavy flavor photoproduction”, *JHEP* **03** (2001) 006, arXiv:hep-ph/0102134.
- [38] M. Cacciari *et al.*, “Theoretical predictions for charm and bottom production at the LHC”, *JHEP* **10** (2012) 137, arXiv:1205.6344 [hep-ph].
- [39] B. Kniehl, G. Kramer, I. Schienbein, and H. Spiesberger, “Inclusive $D^{*\pm}$ production in $p\bar{p}$ collisions with massive charm quarks”, *Phys. Rev. D* **71** (2005) 014018, arXiv:hep-ph/0410289.
- [40] B. Kniehl, G. Kramer, I. Schienbein, and H. Spiesberger, “Inclusive Charmed-Meson Production at the CERN LHC”, *Eur. Phys. J. C* **72** (2012) 2082, arXiv:1202.0439 [hep-ph].
- [41] M. Benzke *et al.*, “Prompt neutrinos from atmospheric charm in the general-mass variable-flavour-number scheme”, *JHEP* **12** (2017) 021, arXiv:1705.10386 [hep-ph].
- [42] 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].
- [43] I. Helenius and H. Paukkunen, “Revisiting the D-meson hadroproduction in general-mass variable flavour number scheme”, *JHEP* **05** (2018) 196, arXiv:1804.03557 [hep-ph].
- [44] 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].
- [45] 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.
- [46] **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** (2022) L011103, arXiv:2105.06335 [nucl-ex].
- [47] T. Sjöstrand *et al.*, “An introduction to PYTHIA 8.2”, *Comput. Phys. Commun.* **191** (2015) 159–177, arXiv:1410.3012 [hep-ph].
- [48] P. Skands, S. Carrazza, and J. Rojo, “Tuning PYTHIA 8.1: the Monash 2013 Tune”, *Eur. Phys. J. C* **74** (2014) 3024, arXiv:1404.5630 [hep-ph].
- [49] J. Bellm *et al.*, “Herwig 7.0/Herwig++ 3.0 release note”, *Eur. Phys. J. C* **76** (2016) 196, arXiv:1512.01178 [hep-ph].

- [50] J. R. Christiansen and P. Z. Skands, “String Formation Beyond Leading Colour”, *JHEP* **08** (2015) 003, arXiv:1505.01681 [hep-ph].
- [51] T. Sjostrand and M. van Zijl, “A Multiple Interaction Model for the Event Structure in Hadron Collisions”, *Phys. Rev. D* **36** (1987) 2019.
- [52] P. Bartalini and L. Fano, eds., *Proceedings, 1st International Workshop on Multiple Partonic Interactions at the LHC (MPI08): Perugia, Italy, October 27-31, 2008*. DESY, Hamburg, 2009. arXiv:1003.4220 [hep-ex].
- [53] I. Schmidt and M. Siddikov, “Production mechanisms of open-heavy flavor mesons”, *Phys. Rev. D* **101** (2020) 094020, arXiv:2003.13768 [hep-ph].
- [54] ALICE Collaboration, J. Adam *et al.*, “Measurement of charm and beauty production at central rapidity versus charged-particle multiplicity in proton-proton collisions at $\sqrt{s} = 7$ TeV”, *JHEP* **09** (2015) 148, arXiv:1505.00664 [nucl-ex].
- [55] 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].
- [56] ALICE Collaboration, S. Acharya *et al.*, “Forward rapidity J/ψ production as a function of charged-particle multiplicity in pp collisions at $\sqrt{s} = 5.02$ and 13 TeV”, *JHEP* **06** (2022) 015, arXiv:2112.09433 [nucl-ex].
- [57] K. Werner, B. Guiot, I. Karpenko, and T. Pierog, “Analysing radial flow features in p-Pb and p-p collisions at several TeV by studying identified particle production in EPOS3”, *Phys. Rev. C* **89** (2014) 064903, arXiv:1312.1233 [nucl-th].
- [58] P. Braun-Munzinger, V. Koch, T. Schäfer, and J. Stachel, “Properties of hot and dense matter from relativistic heavy ion collisions”, *Phys. Rept.* **621** (2016) 76–126, arXiv:1510.00442 [nucl-th].
- [59] ALICE Collaboration, “The ALICE experiment – A journey through QCD”, arXiv:2211.04384 [nucl-ex].
- [60] CMS Collaboration, V. Khachatryan *et al.*, “Observation of Long-Range Near-Side Angular Correlations in Proton-Proton Collisions at the LHC”, *JHEP* **09** (2010) 091, arXiv:1009.4122 [hep-ex].
- [61] ATLAS Collaboration, G. Aad *et al.*, “Observation of Long-Range Elliptic Azimuthal Anisotropies in $\sqrt{s} = 13$ and 2.76 TeV pp Collisions with the ATLAS Detector”, *Phys. Rev. Lett.* **116** (2016) 172301, arXiv:1509.04776 [hep-ex].
- [62] 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].
- [63] Y. Chen and M. He, “Charged-particle multiplicity dependence of charm-baryon-to-meson ratio in high-energy proton-proton collisions”, *Phys. Lett. B* **815** (2021) 136144, arXiv:2011.14328 [hep-ph].
- [64] V. Minissale, S. Plumari, and V. Greco, “Charm hadrons in pp collisions at LHC energy within a coalescence plus fragmentation approach”, *Phys. Lett. B* **821** (2021) 136622, arXiv:2012.12001 [hep-ph].

- [65] S. Plumari, V. Minissale, S. K. Das, G. Coci, and V. Greco, “Charmed Hadrons from Coalescence plus Fragmentation in relativistic nucleus-nucleus collisions at RHIC and LHC”, *Eur. Phys. J. C* **78** (2018) 348, arXiv:1712.00730 [hep-ph].
- [66] ALICE Collaboration, S. Acharya *et al.*, “Observation of a multiplicity dependence in the p_T -differential charm baryon-to-meson ratios in proton-proton collisions at $\sqrt{s} = 13$ TeV”, *Phys. Lett. B* **829** (2022) 137065, arXiv:2111.11948 [nucl-ex].
- [67] ALICE Collaboration, S. Acharya *et al.*, “Constraining hadronization mechanisms with Λ_c^+/D^0 production ratios in Pb-Pb collisions at $\sqrt{s_{NN}} = 5.02$ TeV”, arXiv:2112.08156 [nucl-ex].
- [68] LHCb Collaboration, “Evidence for modification of b quark hadronization in high-multiplicity pp collisions at $\sqrt{s} = 13$ TeV”, arXiv:2204.13042 [hep-ex].
- [69] LHCb Collaboration, R. Aaij *et al.*, “Observation of Multiplicity Dependent Prompt $\chi_{c1}(3872)$ and $\psi(2S)$ Production in pp Collisions”, *Phys. Rev. Lett.* **126** (2021) 092001, arXiv:2009.06619 [hep-ex].
- [70] A. Esposito, E. G. Ferreira, A. Pilloni, A. D. Polosa, and C. A. Salgado, “The nature of $X(3872)$ from high-multiplicity pp collisions”, *Eur. Phys. J. C* **81** (2021) 669, arXiv:2006.15044 [hep-ph].
- [71] E. Braaten, L.-P. He, K. Ingles, and J. Jiang, “Production of $X(3872)$ at High Multiplicity”, *Phys. Rev. D* **103** (2021) L071901, arXiv:2012.13499 [hep-ph].
- [72] 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].
- [73] ALICE Collaboration, K. Aamodt *et al.*, “The ALICE experiment at the CERN LHC”, *JINST* **3** (2008) S08002.
- [74] 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** (2021) 630, arXiv:2009.09434 [nucl-ex].
- [75] ALICE Collaboration, S. Acharya *et al.*, “ALICE 2016-2017-2018 luminosity determination for pp collisions at $\sqrt{s} = 13$ TeV”, <https://cds.cern.ch/record/2776672/>.
- [76] T. Sjostrand, S. Mrenna, and P. Z. Skands, “PYTHIA 6.4 Physics and Manual”, *JHEP* **05** (2006) 026, arXiv:hep-ph/0603175.
- [77] R. Brun *et al.*, *GEANT: Detector Description and Simulation Tool; Oct 1994*. CERN Program Library. CERN, Geneva, 1993. <http://cds.cern.ch/record/1082634>. Long Writeup W5013.
- [78] Particle Data Group Collaboration, R. L. Workman *et al.*, “Review of Particle Physics”, *PTEP* **2022** (2022) 083C01.
- [79] T. Chen and C. Guestrin, “Xgboost: A scalable tree boosting system”, *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (2016) 785–794, arXiv:1603.02754 [cs.LG].
- [80] L. Barioglio, F. Catalano, M. Concas, P. Fecchio, F. Grosa, F. Mazzaschi, and M. Puccio, “hipe4ml/hipe4ml”, July, 2021. <https://doi.org/10.5281/zenodo.5070132>.
- [81] C. Bierlich *et al.*, “A comprehensive guide to the physics and usage of PYTHIA 8.3”, arXiv:2203.11601 [hep-ph].

- [82] K. Werner, “On a deep connection between factorization and saturation: new insight into modeling high-energy proton-proton and nucleus-nucleus scattering in the EPOS4 framework”, arXiv:2301.12517 [hep-ph].
- [83] 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].