

UNIVERSITÀ DEGLI STUDI DI TORINO

---

Scuola di Scienze della Natura  
Dipartimento di Fisica

Laurea Magistrale in Fisica Nucleare e Subnucleare

**Measurement of the prompt and  
feed-down contributions to the  
 $D^+$ -meson production with the  
ALICE experiment at LHC**



**Supervisor:**

Stefania Beolè

**Candidate:**

Fabrizio Gerosa

**Co-supervisors:**

Francesco Prino  
Silvia Masciocchi

**Examiner:**

Ernesto Migliore

---

ANNO ACCADEMICO 2015-2016



## Abstract

The main topic of this thesis is the measurement of the  $D^+$ -meson production in p-Pb collisions at a centre of mass energy per nucleon pair  $\sqrt{s_{NN}} = 5.02$  TeV with A Large Ion Collider Experiment (ALICE) at the Large Hadron Collider (LHC). The ALICE experimental programme is focused on the study of the Quark-Gluon Plasma (QGP), a phase of nuclear matter in which quarks and gluons are deconfined, that can be created in the laboratory through ultra-relativistic heavy-ion collisions.

Open heavy-flavoured mesons (particles made of one heavy quark, i.e. charm or beauty, and one lighter quark) are a unique tool to study and characterise the properties of the deconfined medium, since the heavy quarks are produced in the initial stages of the collision, and therefore they experience the whole system evolution losing energy in the QGP, both through elastic scatterings with the medium constituents and gluons radiation.

Since charm and beauty quarks have a different mass, they lose energy differently in the QGP and therefore it is crucial to be able to distinguish prompt  $D$  mesons (which derive directly from the hadronisation of a charm quark or from the decay of excited open charm and charmonium states) from  $D$  mesons coming from  $B$ -hadron decays. The analysis presented in this thesis is pursued in order to develop new strategies to measure the fraction of prompt  $D^+$ , alternative to the one based on theoretical calculations. The study of proton-nucleus collisions, in which the QGP formation is not expected, can provide an important measurement of the Cold Nuclear Matter effects, such us the modification of the Parton Distribution Functions in a nuclear environment, which cannot be disentangled from hot medium effects in nucleus-nucleus collisions.

In the first Chapter of this thesis I provide an introduction to the Quark-Gluon Plasma and to the heavy-ion physics, presenting the main experimental observables and measurements. In Chapter two a detailed description of the heavy-flavour production in proton-proton, proton-nucleus and nucleus-nucleus is given, with a particular emphasis on the Cold Nuclear Matter and Hot Medium effects. The third Chapter is dedicated to the description of the ALICE apparatus and its main characteristics, focusing in particular on the sub-detectors used for the  $D^+$ -mesons analysis. In Chapter four I explain the procedure to reconstruct the  $D^+$  candidates in the purely hadronic decay channel  $D^+ \rightarrow K^- \pi^+ \pi^+$  and the signal extraction in the p-Pb data sample at  $\sqrt{s_{NN}} = 5.02$  TeV collected in 2013 by the ALICE Collaboration. In this chapter I also present the results published by ALICE using the standard approach. In Chapter five the first and principal data-driven method for the measurement of the fraction of prompt  $D^+$ - mesons is presented. It consists

on an unbinned log-likelihood fit to the distributions of the candidate impact parameter (distance of closest approach between the reconstructed flight line of the D-meson and the primary vertex). This method exploits the difference between the distribution of prompt  $D^+$ , which is smeared around zero only due to resolution effects, and the distribution of secondary  $D^+$ , which is wider because they do not originate from the primary vertex. Chapter six is dedicated to an alternative analysis, based on the usage of different sets of selection criteria on the D-meson decay topology, which can discriminate primary from secondary D-mesons. Finally, in Chapter seven, I focus on the data analysis using a different vertex-reconstruction approach with respect to the one used for the results published until now by the Collaboration, based on the Kalman filter algorithm.

Even though the first two methods are at the moment limited by statistics, they show a really good agreement with the standard result obtained by perturbative QCD calculations and they open up the possibility to measure the prompt D-meson production cross section using a completely data-driven approach. On the other hand, the Kalman filter based method allows one to reduce significantly the contribution of D-mesons coming from B decays and therefore the systematic uncertainty induced by the pQCD calculation, thanks to the possibility to set constraints on the decay topology. Finally the results obtained with all the methods developed were compared among each other and found to be nicely in agreement within the experimental uncertainties.

# Contents

<b>1 Quark-Gluon Plasma</b>	1
1.1 Quantum Chromodynamics . . . . .	1
1.1.1 The QCD Lagrangian . . . . .	2
1.1.2 Confinement and asymptotic freedom . . . . .	3
1.2 Deconfinement and QCD phase diagram . . . . .	5
1.2.1 M.I.T. Bag Model . . . . .	5
1.2.2 Lattice QCD . . . . .	9
1.3 Ultra-relativistic heavy-ion collisions . . . . .	11
1.3.1 Collision evolution . . . . .	12
1.3.2 The Glauber Model . . . . .	13
1.4 Experimental observables . . . . .	16
1.4.1 Soft probes . . . . .	16
1.4.2 Hard probes . . . . .	21
<b>2 Open heavy-flavour physics</b>	25
2.1 Heavy-quark production in pp collisions . . . . .	25
2.2 Heavy-quark production in p-A and A-A collisions . . . . .	27
2.2.1 Cold Nuclear Matter effects . . . . .	28
2.2.2 Hot Medium effects . . . . .	31
<b>3 A Large Ion Collider Experiment</b>	35
3.1 The Large Hadron Collider . . . . .	35
3.2 The ALICE detector . . . . .	37
3.2.1 Inner Tracking System (ITS) . . . . .	38
3.2.2 Time Projection Chamber (TPC) . . . . .	41
3.2.3 Time of Flight (TOF) . . . . .	42
3.3 ALICE trigger and data acquisition . . . . .	43
3.3.1 Central Trigger Processor . . . . .	43
3.3.2 Data AcQuisition System and High-Level Trigger . . . . .	43
3.4 ALICE performance . . . . .	44
3.4.1 Track and vertex reconstruction in the central barrel . . . . .	44
3.4.2 Particle IDentification (PID) . . . . .	46

<b>4 Prompt <math>D^+</math>-meson reconstruction strategy</b>	49
4.1 $D^+$ candidate reconstruction and selection . . . . .	50
4.1.1 p-Pb data sample and event selection . . . . .	51
4.1.2 Track selection . . . . .	51
4.1.3 $D^+$ reconstruction in the fiducial acceptance region . . . . .	52
4.1.4 Topological selections . . . . .	53
4.1.5 Particle Identification . . . . .	57
4.1.6 Raw yield extraction . . . . .	58
4.2 Prompt and Feed-down $D^+$ . . . . .	59
4.2.1 Theory-driven methods for the $D^+$ prompt fraction evaluation	59
4.3 The Monte Carlo simulation . . . . .	61
4.3.1 Efficiency and Acceptance . . . . .	61
4.4 Results from the standard analysis . . . . .	62
<b>5 Impact-parameter fit method for the prompt <math>D^+</math> fraction determination</b>	65
5.1 Analysis strategy . . . . .	66
5.2 Prefit on the impact-parameter distributions . . . . .	67
5.2.1 MC impact-parameter distributions for prompt $D^+$ . . . . .	67
5.2.2 MC impact-parameter distributions for feed-down $D^+$ . . . . .	68
5.2.3 Impact-parameter distributions for the background . . . . .	70
5.3 Unbinned log-likelihood fits on data . . . . .	72
5.3.1 Binned vs. unbinned fit . . . . .	72
5.3.2 Impact-parameter fits . . . . .	73
5.4 Systematic uncertainties . . . . .	75
5.4.1 Systematic uncertainty due to the impact-parameter shapes	77
5.4.2 Systematic uncertainty due to the assumed $p_T$ shape . . . . .	81
5.4.3 Systematic uncertainty induced by the uncertainty on the signal yield $S$ . . . . .	83
5.4.4 Consistency of the fitting procedure . . . . .	84
5.4.5 Summary of the systematic uncertainties . . . . .	88
5.5 Results . . . . .	88
5.6 Perspective of an improved measurement with the upgraded ITS .	91
<b>6 Data-driven cut-variation method for the <math>D^+</math> measurement</b>	93
6.1 Cut-variation feed-down subtraction method . . . . .	93
6.1.1 Analytic minimisation . . . . .	95
6.1.2 Incentre minimisation . . . . .	97
6.2 Strategy for the selection of topological cuts . . . . .	99
6.2.1 Topological variables . . . . .	100
6.2.2 Sets of topological cuts . . . . .	102
6.3 Efficiency . . . . .	104

6.4	Raw yield extraction . . . . .	105
6.5	Systematic uncertainties . . . . .	107
6.5.1	Systematic uncertainty due to the selection efficiency . . . . .	108
6.5.2	Systematic uncertainty due to the raw yield extraction . . . . .	112
6.5.3	Systematic uncertainty due to the $p_T$ shape . . . . .	114
6.6	Results . . . . .	115
<b>7</b>	<b>Prompt <math>D^+</math> reconstruction with the Kalman filter vertexing package</b>	<b>119</b>
7.1	The Kalman filter algorithm . . . . .	120
7.1.1	Reconstruction of particle decays with the KFParticle package	121
7.1.2	Fit constraints . . . . .	121
7.2	Reconstruction performances of the KFParticle package for prompt $D^+$ -mesons . . . . .	122
7.2.1	Prompt $D^+$ -mesons at the decay vertex . . . . .	123
7.2.2	Prompt $D^+$ -mesons at the production vertex . . . . .	124
7.2.3	Prompt $D^+$ -meson decay length . . . . .	126
7.2.4	Prompt $D^+$ -meson $\chi^2$ and probability . . . . .	127
7.2.5	Primary vertex determination . . . . .	128
7.3	Prompt $D^+$ production measurement with the KFParticle package	129
7.3.1	Topological selections . . . . .	129
7.3.2	Tuning of the $D^+$ -meson $\chi^2/ndf$ selection . . . . .	131
7.3.3	Raw yield extraction . . . . .	135
7.4	Systematic uncertainties . . . . .	136
7.4.1	Systematic uncertainty due to the raw yield extraction . . . . .	137
7.4.2	Systematic uncertainty due to the topological selection efficiency . . . . .	138
7.4.3	Systematic uncertainty due to the PID efficiency . . . . .	139
7.4.4	Systematic uncertainty due to the generated $p_T$ shape . . . . .	140
7.4.5	Systematic uncertainty due to the feed-down $D^+$ -meson subtraction . . . . .	141
7.5	Results . . . . .	144
<b>8</b>	<b>Conclusions</b>	<b>147</b>
<b>A</b>	<b>Additional figures</b>	<b>149</b>
A.1	Fits to the MC impact-parameter distributions of prompt $D^+$ -mesons	150
A.2	Fits to the MC true impact-parameter distributions of feed-down $D^+$ -mesons . . . . .	151
A.3	MC impact-parameter distributions of feed-down $D^+$ -mesons . . . . .	152
A.4	Fits to the impact-parameter distributions of the combinatorial background . . . . .	153

A.5	Cut-variation analysis plots . . . . .	154
A.5.1	Mass fits prompt-enhanced set . . . . .	154
A.5.2	Mass fits mixed set . . . . .	155
A.5.3	Mass fits feed-down-enhanced set . . . . .	156
A.5.4	Incentre determination . . . . .	157
A.6	Kalman filter analysis invariant-mass fits . . . . .	158
<b>Bibliography</b>		161
<b>Acknowledgements</b>		167

# List of Figures

1.1	Running of the strong coupling constant $\alpha_s$ [3]. . . . .	4
1.2	The QCD phase diagram. . . . .	9
1.3	Left: various paths from the point $(x_a, t_a)$ to the point $(x_b, t_b)$ in the discretised space-time. Right: the pressure normalised by $T^4$ as a function of the temperature from lQCD calculations, varying the number of time points $N_t$ in the lattice [10]. . . . .	10
1.4	Left: the impact parameter of the nucleus-nucleus collision $b$ in the transverse plane. Right: participants and spectators nucleons in a nucleus-nucleus collision. . . . .	11
1.5	Evolution of a heavy-ion collision represented in a space-time diagram. . . . .	12
1.6	Schematic representation of two nuclei before a collision together with the coordinate system for the nucleons. . . . .	14
1.7	Left: total multiplicity as a function of the number of participant nucleons measured by ALICE in Pb-Pb collisions at $\sqrt{s_{NN}} = 2.76$ TeV compared to the PHOBOS measurement for Au-Au collisions at $\sqrt{s_{NN}} = 200$ GeV [13]. Right: particle density as a function of the number of participants measured by ALICE in pp collisions at $\sqrt{s} = 2.76$ TeV, p-Pb collisions at $\sqrt{s_{NN}} = 5.02$ TeV and Pb-Pb collisions at $\sqrt{s_{NN}} = 2.76$ TeV and $\sqrt{s_{NN}} = 5.02$ TeV [14]. . . . .	17
1.8	Particle density for nucleus-nucleus, proton-nucleus and proton-proton collisions as a function of the centre of mass energy measured by several experiments [14]. . . . .	18
1.9	Left: fits to the $p_T$ -integrated abundances for different hadron species in Pb-Pb collisions at $\sqrt{s_{NN}} = 2.76$ TeV measured by ALICE [15]. Right: QCD phase diagram and data from a chemical freeze-out analysis in nucleus-nucleus collisions at various beam energies. The data points closest to the $T$ -axis are from the highest collision energies [16]. . . . .	19
1.10	Transverse momentum spectra for pions and protons in different centrality classes in Pb-Pb collisions at $\sqrt{s_{NN}} = 2.76$ TeV measured by ALICE [17]. . . . .	20

1.11	Left: sketch of a non-central nucleus-nucleus collisions with the almond-shaped interaction volume. Right: the elliptic flow in different centrality classes in Pb-Pb collisions at $\sqrt{s_{NN}} = 2.76$ TeV measured by ALICE [19]. . . . .	21
1.12	Left: the nuclear modification factor of charged particles measured by ALICE in the most central Pb-Pb collisions at $\sqrt{s_{NN}} = 2.76$ TeV compared to the result from CMS and the theoretical models [21]. Right: angular correlation between high $p_T$ hadrons measured by STAR in pp, d-Au and Au-Au collisions [22]. . . . .	22
1.13	Left: CMS dimuon invariant-mass spectrum in the $\Upsilon(1S)$ , $\Upsilon(2S)$ and $\Upsilon(3S)$ bottomonium states region, in PbPb collisions at $\sqrt{s_{NN}} = 2.76$ TeV, compared to the rescaled result obtained in pp collisions at $\sqrt{s_{NN}} = 7$ TeV [25]. Right: the nuclear modification factor for the $J/\psi$ as a function of $p_T$ measured by ALICE in central Pb-Pb collisions at $\sqrt{s_{NN}} = 2.76$ TeV compared to the PHENIX result in central Au-Au collision at $\sqrt{s_{NN}} = 200$ GeV [26]. . . . .	24
2.1	Left: $p_T$ -differential cross section for prompt $D^+$ -meson production measured by ALICE in pp collisions at $\sqrt{s} = 7$ TeV compared with the FONLL and GM-VFNS calculations [30]. Right: The non-prompt $J/\psi$ $p_T$ -differential cross section measured by LHCb in pp collisions at $\sqrt{s} = 7$ TeV compared to $J/\psi$ from $b$ production as predicted by FONLL [31]. . . . .	26
2.2	Measurements of the nuclear modification of the PDFs performed by several experiments in a wide range of Bjorken $x$ . . . . .	29
2.3	Left: nuclear modification factor $R_{pPb}$ for prompt D-mesons (average of $D^0$ , $D^+$ and $D^{*+}$ ) as a function of the transverse momentum measured with the ALICE detector in p-Pb collisions at $\sqrt{s_{NN}} = 5.02$ TeV compared to model calculations [36]. Right: $p_T$ -integrated nuclear modification factor $R_{pPb}$ for non-prompt $J/\psi$ as a function of rapidity measured by LHCb in p-Pb collisions at $\sqrt{s_{NN}} = 5.02$ TeV [41]. . . . .	31
2.4	Left: prompt D-meson (average of $D^0$ , $D^+$ and $D^{*+}$ ) nuclear modification factor as a function of $p_T$ in the 10% most central Pb-Pb collisions measured by ALICE at $\sqrt{s_{NN}} = 2.76$ TeV compared to $D^0 R_{AA}$ measured by STAR in Au-Au collisions at $\sqrt{s_{NN}} = 200$ GeV [45]. Right: prompt D-meson $R_{AA}$ as a function of the number of participants measured by ALICE compared to $R_{AA}$ of non-prompt $J/\psi$ measured by CMS in Pb-Pb collisions at $\sqrt{s_{NN}} = 2.76$ TeV [44].	33
3.1	Accelerator complex at CERN. . . . .	36
3.2	The ALICE schematic layout. . . . .	37

3.3	Layout of the Inner Tracking System. . . . .	39
3.4	Layout of a SDD sensor. . . . .	41
3.5	Layout of the ALICE Time Projection Chamber. . . . .	42
3.6	Left: comparison between the TPC-ITS matching efficiency, defined as the fraction of TPC tracks with a prolongation in the ITS, in data and in the simulation for Pb-Pb collisions at $\sqrt{s_{NN}} = 2.76$ TeV, for different requirements on ITS points. Right: impact-parameter resolution in the $xy$ -plane for primary particles in pp, p-Pb and Pb-Pb collisions. . . . .	45
3.7	Left: specific energy loss ( $dE/dx$ ) in the TPC as a function of the particle momentum in p-Pb collisions at $\sqrt{s_{NN}} = 5.02$ TeV. The lines show the parametrisation of the expected mean energy loss. Right: velocity $\beta$ as measured with the TOF detector as a function of momentum for particles reaching the TOF in p-Pb collisions at $\sqrt{s_{NN}} = 5.02$ TeV. . . . .	47
4.1	Sketch of the $D^+$ hadronic decay $D^+ \rightarrow K^- \pi^+ \pi^+$ and some of the reconstructed topological quantities associated. . . . .	50
4.2	Rapidity vs transverse momentum distribution of the reconstructed $D^+$ -mesons. The fiducial acceptance region is defined as $ y_{D^+}  < y_{fid}(p_T)$ . . . . .	52
4.3	Left: reconstructed decay length distributions for $D^+$ in three different $p_T$ intervals from Monte Carlo simulations. Right: comparison between the distributions of the normalised decay length in the transverse plane for signal (prompt and feed-down $D^+$ ) and background. . . . .	53
4.4	Left: comparison of the distribution of the reconstructed cosine of the pointing angle between signal (prompt and feed-down $D^+$ ) and background. Right: distributions of the reconstructed cosine of the pointing angle in the transverse plane for the $D^+$ in three different $p_T$ intervals. . . . .	54
4.5	Comparison between the distributions of reconstructed $\sigma_{vertex}$ (left) and transverse momentum of the kaon (right) of signal (prompt and feed-down $D^+$ ) and background candidates. . . . .	55
4.6	Invariant-mass fits in the transverse momentum intervals $2 < p_T < 3$ GeV/c, $4 < p_T < 5$ GeV/c and $12 < p_T < 16$ GeV/c obtained applying the topological cuts reported in table 4.2 and the PID selection. . . . .	58
4.7	$D^+$ prompt fraction $f_{prompt}$ evaluated with the theory-driven methods $f_c$ and $N_b$ for the set of topological cuts reported in table 4.2.	60

4.8	Left: the reconstruction efficiency as a function of $p_T$ for prompt and feed-down $D^+$ obtained applying the topological cuts in table 4.2 and the PID. Right: the acceptance factor for prompt $D^+$ -mesons as a function of $p_T$ . . . . .	62
4.9	Prompt $D^+$ $p_T$ -differential production cross section measured by the ALICE Collaboration in p-Pb collisions at $\sqrt{s_{NN}} = 5.02$ TeV compared to the pp reference cross section scaled by the Pb mass number $A = 208$ [36]. . . . .	63
5.1	Sketch of the decay topology for prompt (right) and feed-down (left) $D^+$ with particular emphasis on the true impact parameter. . . . .	65
5.2	Examples of fit to the simulated impact-parameter distributions for prompt $D^+$ -mesons in two different $p_T$ intervals. . . . .	67
5.3	Examples of fit to the true generated impact-parameter distributions for feed-down $D^+$ -mesons in two different $p_T$ intervals. . . . .	68
5.4	Examples of reconstructed impact-parameter distributions for feed-down $D^+$ -mesons in two different $p_T$ intervals. The function superimposed is the convolution between the true impact parameter function for feed-down $D^+$ (eq. 5.2) and the resolution term (eq. 5.1). . . . .	69
5.5	Left: invariant-mass intervals considered for the parametrisation of the impact-parameter distribution for the background ( <i>side-band regions</i> ) in the transverse momentum interval $4 < p_T < 5$ GeV/c. Right: resulting background impact-parameter distribution. . . . .	70
5.6	Comparison between the impact-parameter distributions for the combinatorial background obtained from the side-bands region, with (red) and without (blue) cut on normalised decay length $xy$ . In both cases all the other topological selections reported in table 4.2 are applied. . . . .	71
5.7	Examples of fit to the impact-parameter distributions for the combinatorial background obtained from the side-band regions, in two different $p_T$ intervals. . . . .	72
5.8	Left: invariant-mass range considered for the impact-parameter distribution final fit in the transverse momentum interval $4 < p_T < 5$ GeV/c. Right: resulting impact-parameter distribution. . . . .	74
5.9	Fits to the $D^+$ impact-parameter distributions in the $xy$ -plane in the seven transverse momentum intervals of the analysis. The curves are the fit functions for prompt $D^+$ (green), feed-down $D^+$ (blue), background (magenta) and the sum of all contributions (red). . . . .	76

5.10 Left: comparison between the $\sigma_{prompt}$ parameter as a function of $p_T$ obtained from the impact-parameter fit on data and the one obtained by fitting the simulated distributions for prompt $D^+$ in the prefit phase. Right: ratio between the $D^+$ prompt fraction $f_{prompt}$ obtained by fixing and letting free $\sigma_{prompt}$ in the fit on data as a function of $p_T$ . . . . .	77
5.11 Ratio of $f_{prompt}$ as a function of $p_T$ obtained varying the fit range with respect to the central value. In the left plot all the fit ranges both in the fit and the prefit phase are changed coherently, while in the right one the impact-parameter range for the background is fixed to [1000,1000] $\mu\text{m}$ . . . . .	78
5.12 Left: example of fit to the background distribution in the $p_T$ interval $2 < p_T < 3$ GeV/c with the symmetric and asymmetric functions. Right: ratio of $f_{prompt}$ obtained with the symmetric and the asymmetric fit functions for the background. . . . .	79
5.13 Left: the five invariant-mass intervals in the side-bands region tested to evaluate the systematic uncertainty for the combinatorial background parametrisation in the $p_T$ interval $4 < p_T < 5$ GeV/c. Right: The ratio of $f_{prompt}$ with respect to the reference value obtained with the mass region $4\sigma <  M - M_{peak}  < 15\sigma$ for the background parametrisation as a function of $p_T$ . . . . .	80
5.14 $p_T$ -spectrum of generated $D^+$ -mesons (left) and B-mesons decaying to $D^+$ (right) from the Pythia simulation compared to FONLL calculations. . . . .	82
5.15 Top: comparison between the prompt and feed-down $D^+$ impact-parameter distributions with and without applying the weight factor used to reproduce the FONLL $p_T$ spectrum for prompt $D^+$ -mesons and B-mesons which decay to $D^+$ respectively, in two $p_T$ intervals of the analysis. Bottom: ratio between the reweighted and not reweighted impact-parameter distributions. . . . .	82
5.16 Left: ratio of $f_{prompt}$ obtained fixing the signal parameter $S$ to its value plus or minus its statistical and systematic uncertainty, with respect to the central value. Right: relative variation of the $p_T$ -differential cross section due to the variation of the signal in the raw yield and in $f_{prompt}$ within its statistical and systematic uncertainties. . . . .	83
5.17 Left: four invariant-mass windows tested for the stability checks in the $p_T$ interval $4 < p_T < 5$ GeV/c. Right: ratio of $f_{prompt}$ with respect to the reference value obtained with $ M - M_{peak}  < 2\sigma$ as a function of $p_T$ . . . . .	84
5.18 Mean (left) and RMS (right) of the residual distributions for the $D^+$ prompt fraction evaluated with the MC closure test, for four different input values of $f_{prompt}$ . . . . .	85

5.19 Left: pulls of $f_{prompt}$ distributions (left) and mean value of $\sigma_{prompt}$ (right) evaluated with the MC closure test, for four different input values of $f_{prompt}$ . . . . .	86
5.20 $D^+$ prompt fraction obtained with the impact-parameter fit method, compared to the FONLL-based method ( $N_b$ and $f_c$ combined). The vertical bars represent the statistical uncertainty, while the boxes the systematic one. . . . .	88
5.21 Examples of impact-parameter fits (left column) and resulting prompt fraction as a function of $p_T$ (right column) for the three different $D$ -meson species analysed: $D^0$ (top), $D^+$ (middle), $D^{+*}$ (bottom) [1]. The arrow in the interval $1 < p_T < 2$ GeV/c for the $D^0$ -meson represents the minimum value within a 95% confidence level. . . . .	90
5.22 Left: Layout of the upgraded ALICE ITS. Right: improvement of the impact-parameter resolution in the transverse plane expected with the upgraded ITS [66]. . . . .	91
5.23 Left: simulated impact-parameter distributions for prompt and feed-down $D^0$ obtained with the current and upgraded ITS configurations in the interval $2 < p_T < 3$ GeV/c [67]. Right: $D^0$ impact-parameter resolution as a function of $p_T$ . . . . .	92
6.1 Schematic representation of the selection of the $D^+$ candidates for two different sets of topological cuts. . . . .	94
6.2 Three lines corresponding to different sets of topological cuts in the interval $6 < p_T < 8$ GeV/c. Their intersections define a triangle, whose incentre is used to evaluate the corrected yields of prompt and feed-down $D^+$ -mesons. . . . .	97
6.3 Left: distribution of the generated incentres with the toy MC simulation superimposed to the lines corresponding to three different sets of cuts in the interval $6 < p_T < 8$ GeV/c. The dotted lines represent the maximum effect of the uncertainty on the line parameters within $1\sigma$ . Right: projection of the distribution of the generated incentres on the $N_{prompt}$ -axis. . . . .	98
6.4 Comparison between the prompt and feed-down $D^+$ distributions from MC simulations of the decay length in the $xy$ -plane (left) in the interval $6 < p_T < 8$ GeV/c and the normalised decay length in the $xy$ -plane (right) for $2 < p_T < 3$ GeV/c. . . . .	100
6.5 Comparison between the distributions of the cosine of the pointing angle in the $xy$ -plane in the interval $12 < p_T < 16$ GeV/c (left) and the maximum standardised difference between measured and expected track impact parameter (right) in the interval $5 < p_T < 6$ GeV/c for prompt and feed-down $D^+$ from MC simulations. . . . .	101

6.6	Sketch of the $D^+$ decay topology with the quantities used to compute the expected daughter track impact parameter. . . . .	102
6.7	Comparison between the efficiencies for prompt and feed-down $D^+$ for the three sets of topological cuts used in the cut-variation analysis. . . . .	104
6.8	Example of invariant-mass fits for the three different sets of topological cuts in the transverse momentum interval $4 < p_T < 5 \text{ GeV}/c$ . . . . .	106
6.9	Comparison between the Gaussian width (left) and mean (right) obtained by fitting the invariant-mass spectra of the candidates in data and MC for the mixed cut set. The peak position is also compared to the PDG value of the mass of the $D^+$ -meson. . . . .	107
6.10	Left: ratio of the corrected yields for prompt $D^+$ as a function of $p_T$ obtained with the analytic minimisation with respect to the central value, for the cut variations on the prompt-enhanced set. Right: distributions of the ratios of the corrected yield in the interval $5 < p_T < 6 \text{ GeV}/c$ for the variations on the three cut sets separately (red, blue, green) and for all the trials together (black). . . . .	108
6.11	Left: ratio of the corrected yields for feed-down $D^+$ as a function of $p_T$ obtained with the incentre minimisation with respect to the central value, for the feed-down-enhanced set variations. Right: distributions of the corrected yield variations in the interval $6 < p_T < 8 \text{ GeV}/c$ for the three cut sets separately (red, blue, green) and for all the trials together (black), obtained removing the bad trial, corresponding to the variation of the cut on the maximum normalised residual of the daughter impact parameter. . . . .	111
6.12	Prompt $D^+$ corrected yield obtain with the incentre minimisation method as a function of the trial number (left) and the resulting distribution (right) for the $p_T$ interval $3 < p_T < 4 \text{ GeV}/c$ . . . . .	112
6.13	Feed-down $D^+$ corrected yield obtained with the analytic minimisation method as a function of the trial number (left) and the resulting distribution (right) for the $p_T$ interval $5 < p_T < 6 \text{ GeV}/c$ . . . . .	113
6.14	Left: ratio of the efficiency for prompt and feed-down $D^+$ -mesons obtained using the FONLL and Pythia $p_T$ shapes, for the prompt-enhanced and the feed-down-enhanced cut sets. Right: ratio of the corrected yields obtained with the analytic minimisation, using the efficiencies computed with the FONLL and Pythia $p_T$ shapes. . . . .	115
6.15	Top: prompt $D^+$ production cross section measured in p-Pb collisions at $\sqrt{s_{NN}} = 5.02 \text{ TeV}$ using the cut-variation method, compared to the result obtained with the standard approach published by the ALICE Collaboration [36]. The vertical bars represent the statistical errors, while the boxes the systematic uncertainties. Bottom: ratio of the measured cross section with the cut-variation approach and the published result. . . . .	116

6.16 Feed-down $D^+$ production cross section measured in p-Pb collisions at $\sqrt{s_{NN}} = 5.02$ TeV using the cut-variation method. The vertical bars represent the statistical errors, while the boxes the systematic uncertainties. . . . .	117
7.1 Example of Gaussian fit to the residual (left) and the pull (right) distributions of the $z$ coordinate of the prompt $D^+$ decay vertex in the interval $2 < p_T < 3$ GeV/c obtained with the KFParticle package. . . . .	122
7.2 Resolution (left) and pulls (right) of the $x$ and $z$ coordinates of the prompt $D^+$ decay vertex position as a function of $p_T$ . . . . .	124
7.3 Resolution (left) and pulls (right) of the $x$ and $z$ coordinates of the prompt $D^+$ production vertex position as a function of $p_T$ . . . . .	125
7.4 Resolution (left) and pulls (right) of the impact parameter in the transverse plane for prompt $D^+$ -mesons as a function of $p_T$ . . . . .	125
7.5 Resolution (left) and pulls (right) of the decay length in the transverse plane of the prompt $D^+$ -mesons as a function of $p_T$ . . . . .	126
7.6 $\chi^2$ (left) and probability (right) distributions for the prompt $D^+$ -mesons reconstructed with the Kalman filter algorithm, with and without applying the topological constraint. . . . .	127
7.7 Resolution (left) and bias (right) of the $x$ and $z$ coordinates of the reconstructed primary-vertex position as a function of the the number of tracks used for the vertex determination in events with at least a $D^+$ -meson, for different cases of tracks selected for the vertex determination. . . . .	128
7.8 Comparison between the distributions of $\chi^2/ndf$ of prompt and feed-down $D^+$ -mesons after applying the topological constraint at the primary vertex from MC simulations, in two different $p_T$ intervals of the analysis. . . . .	130
7.9 Values of the extracted raw yield in the $p_T$ bins of the analysis for different $\chi^2/ndf$ cut values (left) and relative variation with respect to the values obtained without applying the $\chi^2/ndf$ cut (right). . . . .	132
7.10 Relative variation of the efficiencies for prompt (left) and feed-down (right) $D^+$ obtained for the different $\chi^2/ndf$ cuts with respect to those obtained without applying the $\chi^2/ndf$ cut. . . . .	132
7.11 Left: $D^+$ prompt fraction evaluated with the FONLL-based methods for the different $\chi^2/ndf$ cuts. Right: relative variation of the prompt $D^+$ $p_T$ -differential cross section obtained with the different $\chi^2/ndf$ cut values with respect to the central value. . . . .	133
7.12 Invariant-mass fits in the transverse momentum intervals $1 < p_T < 2$ GeV/c, $5 < p_T < 6$ GeV/c and $16 < p_T < 24$ GeV/c obtained applying the topological cuts reported in table 7.1 and the PID selection. . . . .	134

7.13 Comparison between the peak width (left) and mean (right) obtained by fitting the invariant-mass distributions obtained from data and simulation. The peak position is also compared to the PDG value of the $D^+$ mass. . . . .	136
7.14 Raw yield obtained via the fit to the invariant-mass distributions and the bin counting method (left) and the Gaussian width (right) as a function of the trial number for $D^+$ candidates reconstructed with the KFParticle package with $5 < p_T < 6 \text{ GeV}/c$ . . . . .	137
7.15 Left: ratio of the $p_T$ -differential cross section obtained varying the cut on the normalised decay length $xy$ relative to the central value as a function of $p_T$ . Right: distribution of the ratio of the $p_T$ -differential cross section varying all the topological cuts in the interval $7 < p_T < 8 \text{ GeV}/c$ . . . . .	139
7.16 Left: comparison between the fit to the invariant-mass distribution obtained with (blue) and without (red) applying the PID selection in the interval $4 < p_T < 5 \text{ GeV}/c$ . The Gaussian width is fixed to the value obtained from the MC simulation. Right: ratio of the $p_T$ -differential cross section of prompt $D^+$ obtained with and without the PID selection. . . . .	140
7.17 Left: Ratio of the efficiencies for prompt $D^+$ obtained with the FONLL and Pythia $p_T$ shapes. Right: ratio of the fraction of prompt $D^+$ calculated using the efficiencies for prompt and feed-down $D^+$ obtained with the FONLL and Pythia $p_T$ shapes. . . . .	141
7.18 Comparison between the fraction of prompt $D^+$ obtained using the KFParticle package for the reconstruction of the $D^+$ candidates, applying the topological selections reported in table 7.1, and with the topological selections used in the standard analysis. $f_{prompt}$ is computed with the FONLL-based methods in both cases. . . . .	142
7.19 Comparison between the efficiencies of prompt and feed-down $D^+$ (left) and the $D^+$ prompt fraction (right) obtained applying the selection on the $\chi^2/ndf$ and on the maximum difference between the measured and expected impact parameter of the daughter tracks. . . . .	143
7.20 Left: scatter plot of the maximum difference between the measured and expected impact parameter of the daughter tracks vs. the $\chi^2/ndf$ for the $D^+$ candidates with $3 < p_T < 4 \text{ GeV}/c$ . Right: ratio between the $p_T$ -differential cross section of prompt $D^+$ obtained applying the cut on the maximum difference between the measured and expected impact parameter of the daughter tracks and the one obtained applying the $\chi^2/ndf$ cut. . . . .	144

7.21 Comparison between the $p_T$ -differential production cross section of prompt $D^+$ -mesons in p-Pb collisions at $\sqrt{s_{NN}} = 5.02$ TeV obtained using the KFParticle package for the reconstruction of the $D^+$ candidates compared to the published result [36]. . . . .	145
8.1 Comparison among the $p_T$ -differential production cross section of prompt $D^+$ -mesons in p-Pb collisions at $\sqrt{s_{NN}} = 5.02$ TeV obtained with all the analyses discussed in this thesis and the one published by the ALICE Collaboration [36]. . . . .	148

# List of Tables

3.1	Dimensions of the Inner Tracking System. . . . .	40
4.1	D-mesons hadronic decays measured by the ALICE Collaboration. The Characteristics reported are from [56]. . . . .	49
4.2	Summary table of the $D^+$ topological cuts optimised for the published result [36]. The optimisation procedure is explained in [59]. . . . .	56
4.3	PID response for the TPC and TOF detectors in different $p_T$ intervals of the tracks. . . . .	57
5.1	Raw yields in the invariant-mass range $ M - M_{peak}  < 2\sigma$ with the statistical error in the seven $p_T$ intervals of the analysis. The systematic uncertainties on the yield extraction are taken from the values estimated for the standard analysis [59]. . . . .	75
5.2	Summary of the systematic uncertainty for the data-driven impact-parameter analysis for the evaluation of the fraction of prompt $D^+$ -mesons. . . . .	87
5.3	Comparison between the total uncertainty on the $D^+$ prompt fraction evaluated with the impact-parameter fit method and the FONLL-based method. . . . .	89
6.1	Summary table of the three sets of topological cuts used for the $D^+$ cut-variation analysis. . . . .	103
6.2	Summary table of the raw yields and the significance for the three sets of topological cuts. . . . .	106
6.3	Estimated uncertainties of the prompt $D^+$ corrected yields for the single cut sets, the sum in quadrature of the three contributions ( <i>Sum</i> ) and from all the trials together ( <i>Tot</i> ). . . . .	109
6.4	Comparison between the systematic uncertainty due to the cut efficiency for prompt $D^+$ estimated considering all the variations and excluding the variations of the mixed cut set. . . . .	110

6.5	Estimated uncertainties on the feed-down $D^+$ corrected yields for the single cut sets, the sum in quadrature of the three contributions ( <i>Sum</i> ) and from all the trials together ( <i>Tot</i> ). . . . .	111
6.6	Systematic uncertainty on the corrected yields due to the raw yield extraction. . . . .	114
6.7	Systematic uncertainty on the corrected yields due to the $p_T$ shape of the generated $D^+$ -mesons. . . . .	114
7.1	Summary table of the topological cuts optimised for the analysis performed using the Kalman filter algorithm for the reconstruction of $D^+$ -mesons. . . . .	130
7.2	Raw yield, signal-to-background ratio and significance obtained from the fit to the invariant-mass distributions obtained by reconstructing the $D^+$ candidates with the KFParticle package. . . . .	135
7.3	Summary of the systematic uncertainty for the analysis performed reconstructing the $D^+$ candidates with the KFParticle package. . . . .	141
7.4	Comparison between the systematic uncertainty on the $D^+$ prompt fraction obtained using the KFParticle package for the reconstruction of the $D^+$ candidates and the one obtained in the standard analysis. . . . .	144

# 1

## Quark-Gluon Plasma

### 1.1 Quantum Chromodynamics

In the Standard Model the strong interaction is described by Quantum Chromodynamics (QCD), which is a non-Abelian quantum gauge field theory, based on the invariance under local  $SU(3)_c$  group transformations.

The choice of  $SU(3)_c$  is dictated by the hypothesis of the existence of a new degree of freedom of matter called colour, which can assume three different states: red, blue and green. The introduction of colour was necessary to explain the propriety of symmetry (or anti-symmetry) of the hadron wave-functions. In fact, according to the static quark model, the hadrons are made of elementary spin- $\frac{1}{2}$  constituents called quarks, and they are divided in two categories: mesons, which are composed by a quark and an anti-quark, and baryons which are made of three quarks. Therefore the meson wave-function takes the form

$$\psi_{q\bar{q}} = \psi_{q\bar{q},space}(x) \cdot \psi_{q\bar{q},spin} \cdot \psi_{q\bar{q},flavour}, \quad (1.1)$$

while the baryon one

$$\psi_{qqq} = \psi_{qqq,space}(x) \cdot \psi_{qqq,spin} \cdot \psi_{qqq,flavour}. \quad (1.2)$$

These wave-functions must be completely symmetric for mesons, which are bosons (particles with integer spin), and completely anti-symmetric for baryons, which are fermions (particles with half-integer spin). During the '60s it was realised that the

product of these space, spin and flavour wave-functions was symmetric for baryons too. In fact, considering for example the  $\Delta^{++}$  state, we obtain that:

- it is made of three u quarks, therefore is flavour symmetric
- it has  $J^P = \frac{3}{2}^+$ , then has zero orbital momentum (spatially symmetric), and has a symmetric  $S = \frac{3}{2}$  spin wave-function.

The consequence of that was the introduction of a colour wave-function,  $\psi_{colour}$ , which has to be symmetric for mesons and anti-symmetric for baryons.

### 1.1.1 The QCD Lagrangian

The conservation of the colour charge in a local  $SU(3)_c$  gauge theory leads to the introduction of eight mass-less vector bosons, called gluons, and the definition of a covariant derivative

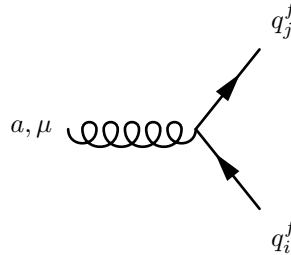
$$D_{ij}^\mu = \partial^\mu \delta_{ij} - ig_s \left( \frac{\lambda_{ij}^a}{2} \right) G_a^\mu(x), \quad (1.3)$$

where  $G_a^\mu(x)$  are the gluon's fields,  $\lambda_{ij}^a$  the Gell-Mann matrices (generators of  $SU(3)$ ) and  $g_s$  is the coupling constant for the strong interaction.

Therefore we obtain the following Lagrangian [2]

$$L_{QCD} = \sum_f \bar{q}_i^f(x) (i \not{D}_{ij}(x) - m_f \delta_{ij}) q_j^f(x) - \frac{1}{4} G_a^{\mu\nu}(x) G_{\mu\nu}^a(x). \quad (1.4)$$

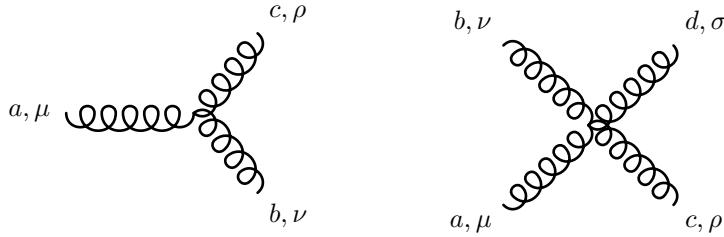
The first term of the Lagrangian describes the quark kinematics and the interaction between quarks and gluons, producing “QED-like” vertex, as shown in the following diagram



The second term describes the kinematics and the dynamics of the gluons, in particular if we make explicit the tensor

$$G_{\mu\nu}^a(x) = \partial_\mu G_\nu^a(x) - \partial_\nu G_\mu^a(x) + ig_s f_{abc} G_\mu^b(x) G_\nu^c(x), \quad (1.5)$$

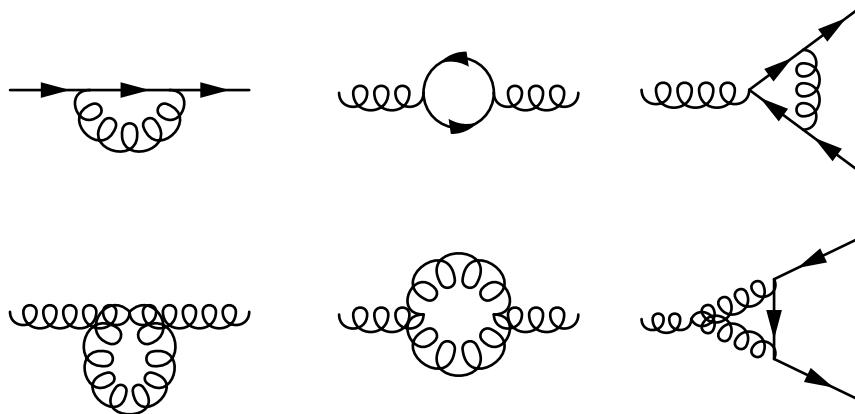
we find that the non-Abelian term ( $ig_s f_{abc} G_\mu^b(x) G_\nu^c(x)$ , where  $f_{abc}$  are the structure constants of SU(3)) produces self-interactions among gluons, which, in terms of Feynman's diagram, means the production of the following vertices



In fact, in contrast to the vector boson of QED, the photon, which does not carry the electromagnetic charge, the gluons carry a colour and an anti-colour, so they can interact not only with quarks, but also among themselves.

### 1.1.2 Confinement and asymptotic freedom

In order to renormalise a quantum field theory, higher orders of the perturbative expansion must be taken into account to compensate divergent terms. The QCD first order correction ( $O(\alpha_s^2)$ ) includes self-energy corrections, vacuum polarisation and vertices corrections, as shown in the following diagrams



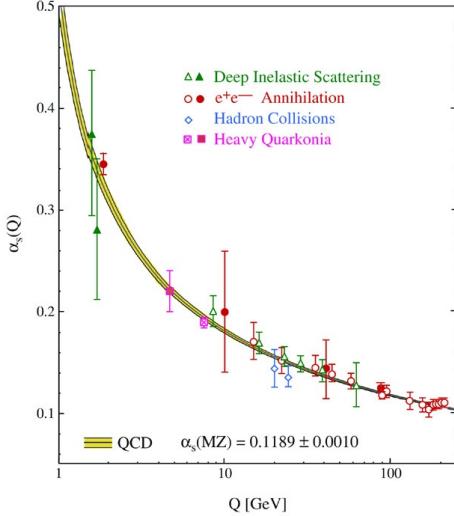


Figure 1.1. Running of the strong coupling constant  $\alpha_s$  [3].

One of the consequences of these corrections is the energy dependence of the strong coupling constant [3], expressed by the following scaling rule

$$\alpha_s(Q^2) = \frac{\alpha_s(\mu^2)}{1 + \frac{\alpha_s(\mu^2)}{12\pi}(33 - 2n_f)\ln(Q^2/\mu^2)} \quad (1.6)$$

where  $\alpha_s = g_s^2/4\pi$  is the strong coupling constant,  $\mu$  is the mass scale of the renormalisation and  $n_f$  is the number of quark families, which is supposed to be equal to three. Therefore  $\alpha_s$  decreases with increasing  $Q^2$  (figure 1.1).

In order to remove the mass scale parameter  $\mu$ , a dimensional parameter  $\Lambda_{QCD}$  is introduced, which represents the energy scale at which  $\alpha_s(Q^2)$  diverges to infinity. Therefore the equation 1.6 can be rewritten as

$$\alpha_s(Q^2) = \frac{12\pi}{(33 - 2n_f)\ln(Q^2/\Lambda_{QCD}^2)}. \quad (1.7)$$

The behaviour of the coupling constant yields the measured value of  $\alpha_s$  at  $\mu^2 = M_Z^2$  if  $\Lambda_{QCD} \sim 200$  MeV.

In the region of  $Q \lesssim 1$  GeV the coupling constant becomes very ‘strong’, therefore it is not possible to apply a perturbative approach. Additionally for decreasing  $Q^2$

(increasing distances)  $\alpha_s$  diverges. This means that quarks are strongly bound in hadrons and they cannot be separated. This phenomenon is called *confinement*.

On the other hand, for large values of  $Q^2$ , the coupling constant becomes very small, therefore in this energy region the hadrons constituents can be considered free and weakly interacting; for this reason these values of  $\alpha_s$  correspond to the *asymptotic freedom*. Moreover in this region, because of the small value of  $\alpha_s$ , it is possible to perform perturbative calculations.

## 1.2 Deconfinement and QCD phase diagram

### 1.2.1 M.I.T. Bag Model

The M.I.T. Bag Model [4] is a phenomenological model which describes the confinement of quarks and gluons inside hadrons. It simply assumes that quarks are mass-less particles, free to move inside a finite region of space, called *bag* [5], and the confinement into this bag is caused by an external pressure  $B$ , which compensate the internal pressure due to their kinetic energy. This external pressure is the effect of the non-perturbative QCD which does not allow quarks to move outside the bag. Considering the free Dirac equation for a mass-less fermion

$$i\gamma^\mu \partial_\mu \psi = \gamma^\mu p_\mu \psi = 0 \quad (1.8)$$

where  $\gamma^\mu$  are the Dirac matrices and  $p^\mu$  the four-momentum of the particle. Writing the four-spinor  $\psi$  in terms of two 2-dimensional spinors

$$\psi = \begin{pmatrix} \psi_+ \\ \psi_- \end{pmatrix}, \quad (1.9)$$

we end up with the following solutions

$$\begin{aligned} \psi_+ &= N e^{-ip_0 t} j_0(p_0 r) \chi_+ \\ \psi_- &= N e^{-ip_0 t} (\vec{\sigma} \cdot \hat{r}) j_1(p_0 r) \chi_-, \end{aligned} \quad (1.10)$$

where  $N$  is a normalisation factor,  $r$  the radius,  $t$  the time,  $j_i(p_0 r)$  are the spherical Bessel functions,  $\sigma^i$  the Pauli matrices and  $\chi_\pm$  two normalised bispinors. The confinement can be then expressed imposing the scalar current equal to zero on the bag surface ( $r = R$ ), which leads to

$$\bar{\psi} \psi|_{r=R} = \psi^\dagger \gamma^0 \psi|_{r=R} = 0 \quad \Leftrightarrow \quad p_0 R = 2.04. \quad (1.11)$$

From this relation we can obtain the kinetic energy of  $N$  quarks inside a bag with radius  $R$

$$E_{kin} = N p_0 = \frac{N 2.04}{R}. \quad (1.12)$$

As stated above the kinetic energy has to be compensated by an external pressure, which in case of spherical bag, can be expressed as

$$E = \frac{N2.04}{R} + \frac{4\pi}{3} R^3 B, \quad (1.13)$$

where  $B$  is the bag pressure. The dimension  $R$  of the bag corresponds to the minimum value of the energy of the system

$$dE/dR = -\frac{N2.04}{R^2} + 4\pi R^2 B = 0. \quad (1.14)$$

If we assume that the dimension of the bag corresponds to the dimension of an hadron ( $R \sim 0.8$  fm) and we consider the number of valence quarks ( $N = 3$ ), it is possible to deduce the external pressure of the bag

$$B^{1/4} = \left( \frac{3 \times 2.04}{4\pi} \hbar c \right)^{1/4} \frac{1}{0.8 \text{ fm}} \simeq \frac{206 \text{ MeV}}{(\hbar c)^{3/4}}. \quad (1.15)$$

According to this model the confinement of quarks into hadrons is the result of the balance between the internal and external pressure. Nevertheless this balance can be broken if the internal pressure outward directed increases. In this case a new state of matter is expected, in which quarks and gluons can move in a larger volume and the colour degrees of freedom are no more confined into colourless hadrons. This deconfined state of matter has been named *Quark-Gluon Plasma* (QGP) [6]. This transition can occur at large temperature or densities.

### QGP at high temperature

One way to produce the QGP is to increase the temperature above a certain critical temperature  $T_c$ ; in this case the state of matter is called “hot” QGP. According to the Cosmological Standard Model this “hot” QGP was the state of matter of the Universe between  $10^{-12}$  (electroweak symmetry breaking) and  $10^{-6}$  seconds after the Big Bang. This is also the situation expected to be reproduced at RHIC and LHC, in ultra-relativistic heavy-ion collisions, as will be described in section 1.3.

To evaluate the critical temperature for the phase transition we consider a system of mass-less and non interacting quarks and gluons in thermal equilibrium in a finite volume  $V$ . We also assume that the total baryon number of the system is zero (*Bjorken regime*), which means that there are the same amount of quarks and anti-quarks. Under these hypotheses we can estimate the pressure (which depends from on the temperature), and therefore evaluate  $T_c$ . Using the Fermi-Dirac statistic distribution we first evaluate the density of quarks ( $n_q$ ), which is

$$n_q = \frac{4\pi g_q}{2\pi^3} \int_0^\infty \frac{p^2 dp}{e^{(p-\mu_q)/T} + 1}, \quad (1.16)$$

where  $g_q = N_c N_s N_f$  are the degrees of freedom of quarks,  $T$  is the temperature and  $\mu_q$  the chemical potential, which represents the energy necessary to add a particle with given quantum numbers to the system. Similarly we can obtain the density of anti-quarks ( $n_{\bar{q}}$ )

$$n_{\bar{q}} = \frac{4\pi g_{\bar{q}}}{2\pi^3} \int_{-\infty}^0 \left\{ 1 - \frac{p^2 dp}{e^{(p-\mu_{\bar{q}})/T} + 1} \right\} = \frac{4\pi g_q}{2\pi^3} \int_0^\infty \frac{p^2 dp}{e^{(p+\mu_q)/T} + 1}, \quad (1.17)$$

which differs from  $n_q$  only by the sign between  $p$  and  $\mu_q$  in the exponential. Using the hypothesis of the *Bjorken regime* ( $n_q = n_{\bar{q}}$ ) we obtain  $\mu_q = 0$ . From this expression we can therefore calculate the energy density for quarks (and anti-quarks)

$$\epsilon_{q(\bar{q})} = \frac{g_{q(\bar{q})}}{2\pi^2} \int_0^\infty \frac{p^3 dp}{e^{p/T} + 1} = \frac{g_{q(\bar{q})}}{2\pi} T^4 \Gamma(4) (1 - 2^{-3}) \zeta(4) = \frac{7}{8} \frac{g_{q(\bar{q})} \pi^2}{30} T^4. \quad (1.18)$$

The energy density of gluons can be obtained as well, using the Bose-Einstein statistical distribution instead of the Fermi-Dirac one

$$\epsilon_g = \frac{g_g}{2\pi^2} \int_0^\infty \frac{p^3 dp}{e^{p/T} - 1} = \frac{g_g}{2\pi} T^4 \Gamma(4) \zeta(4) = \frac{g_g \pi^2}{30} T^4. \quad (1.19)$$

Finally, using the relation between the pressure and the energy density

$$P = \epsilon/3, \quad (1.20)$$

we obtain the total pressure

$$P = P_q + P_{\bar{q}} + P_g = \left\{ \frac{7}{8} (g_q + g_{\bar{q}}) + g_g \right\} \frac{\pi^2}{90} T^4. \quad (1.21)$$

The critical temperature is therefore the temperature at which the internal pressure caused by the kinetic energy of quarks and gluons is equal to the bag pressure (equation 1.15)

$$T_c = \left\{ \left[ \frac{7}{8} (g_q + g_{\bar{q}}) + g_g \right] \right\}^{-1/4} B^{1/4}. \quad (1.22)$$

The degrees of freedom of the gluons take into account the two possible polarisations (as shown in section 1.1.1 the gluons are mass-less bosons), while for the quarks (and anti-quarks) we consider the colour, the spin and the three light flavours (up, down and strange). Therefore  $g_g = 8 \times 2 = 16$  and  $g_q = g_{\bar{q}} = 3 \times 2 \times 3 = 18$ . Then the critical temperature results to be  $T_c \simeq 145$  MeV, while the critical energy density  $\epsilon_c \simeq 1$  GeV/fm<sup>3</sup>. If the matter inside the bag is brought to a higher temperature than  $T_c$ , then the bag pressure is not able to confine it, and therefore there is a transition, from hadronic matter to Quark-Gluon Plasma.

## QGP by compression

Increasing the temperature is not the only way to obtain a transition to QGP, but it is also possible to produce a deconfined state increasing the baryonic density. This condition is expected to be met in the core of neutron stars and it is called “cold” QGP.

In this case we can assume the limit case of  $T = 0$  and very high baryonic density. In this situation the contribution of gluons and anti-quark is negligible. Because of the Pauli principle, all the energy states up to  $\mu_q$  are occupied and the pressure is

$$P = \frac{\epsilon_q}{3} = \frac{g_q}{6\pi^2} \int_0^{\mu_q} p^3 dp = \frac{g_q}{24\pi^2} \mu_q^4. \quad (1.23)$$

Using the condition  $P = B$  we can extract the critical chemical potential, which results

$$\mu_q^c = \left( \frac{24\pi^2 B}{g_q} \right)^{1/4}. \quad (1.24)$$

Finally, we can evaluate the critical baryon density

$$n_B^c = \frac{1}{3} (n_q^c)^3 = \frac{g_q}{6\pi^2} \int_0^{\mu_q^c} p^2 dp = \frac{g_q}{18\pi^2} \mu_q^3 = \frac{4}{3} \left( \frac{g_q}{24\pi^2} \right)^{1/4} B^{3/4}. \quad (1.25)$$

If we consider  $B \simeq 206$  MeV and  $g_q = 18$ , the critical baryon density is  $n_B^c = 0.72$  fm $^{-3}$ , which is around five times larger than the nucleus density. If the baryon density is above the critical value, quarks and gluons are not bound into hadrons, and the QGP formation occurs.

## QCD phase diagram

The situations shown in the previous sections define two limits: in the first case the phase transition occurs when the temperature is very high and the net baryon density is zero, while the second one when the temperature is zero and the baryon density is approximately five times the equilibrium nuclear matter density. For a system in between these two limits there is a pressure arising from the thermal motion and from the degeneracy of the fermion gas, therefore there are intermediate values of baryon density and temperature at which the nuclear matter becomes deconfined. It is then possible to define a phase diagram of the transition [7], as schematically shown in figure 1.2.

Nevertheless it is important to notice that the Bag Model can provide a qualitative explanation for the deconfinement, but for more quantitative calculations and predictions the so-called *Lattice QCD* is needed. A brief explanation of lQCD will be provided in the next section.

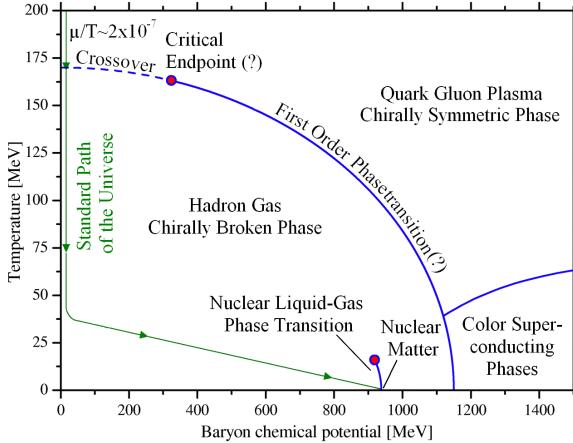


Figure 1.2. The QCD phase diagram.

### 1.2.2 Lattice QCD

The Lattice QCD is a numerical technique which allows to perform calculations in a non-perturbative regime, based on the approximation of the continuous space-time by a discrete lattice of points. Such a lattice will introduce a minimum distance, namely the lattice spacing  $a$  between neighbouring points. Since two points can not be closer than  $a$ , there is now a corresponding maximum momentum, which is the theory *cut-off* that avoids ultraviolet divergences that occur in the perturbative QCD.

The starting point of this theory is the Feynman path integrals approach to classical quantum mechanics, which assumes that a system evolves from the initial state to the final state through all the possible paths, with a probability weighted by the action evaluated along the considered path. The transition amplitude from the state  $(x_a, t_a)$  to the state  $(x_b, t_b)$  is therefore

$$A[(x_a, t_a) \rightarrow (x_b, t_b)] = \langle x_b | e^{-iH(t_b - t_a)} | x_a \rangle = \sum_{\text{paths}} e^{iS_M}, \quad (1.26)$$

where  $H$  is the Hamiltonian of the system, and  $S_M$  is the action evaluated in the Minkowsky space.

Taking now the continuum limit we obtain an integration over space

$$\sum_{\text{paths}} e^{iS_M} \rightarrow \int \prod_{i=1}^{n_t-1} dx_i e^{iS_M}, \quad (1.27)$$

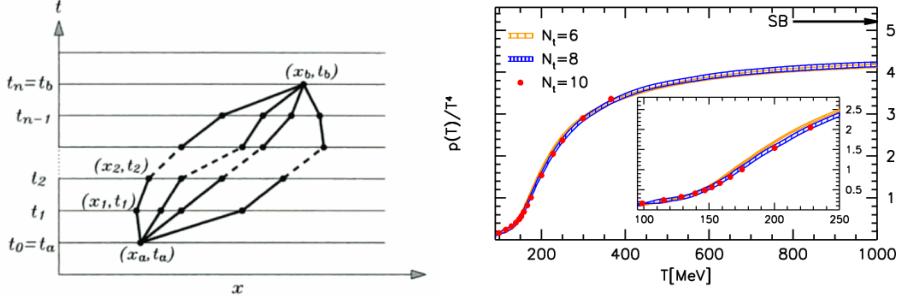


Figure 1.3. Left: various paths from the point  $(x_a, t_a)$  to the point  $(x_b, t_b)$  in the discretised space-time. Right: the pressure normalised by  $T^4$  as a function of the temperature from lQCD calculations, varying the number of time points  $N_t$  in the lattice [10].

where  $\prod_{i=1}^{n_t-1}$  is the product of all the time coordinates between  $t_a$  and  $t_b$ .

We notice that the amplitude shows a formal analogy with the partition function for a single particle in quantum statistical mechanics  $Z$ , which is

$$Z = \sum_a \langle x_a | e^{-\beta H} | x_a \rangle. \quad (1.28)$$

In order to complete this analogy we have to apply the *Wick rotation*, which means to introduce an imaginary time coordinate  $\tau = it$ , with  $\tau_a = 0$  and  $\tau_b = \beta$ . With this transformation the Minkowsky action changes into the Euclidean one

$$\begin{aligned} iS_M|_{t=-i\tau} &= i \int L dt = i \int \left\{ \frac{1}{2} m \left( \frac{dx}{dt} \right)^2 - V(x) \right\} dt = \\ &= - \int \left\{ \frac{1}{2} m \left( \frac{dx}{d\tau} \right)^2 + V(x) \right\} d\tau \equiv -S_E. \end{aligned} \quad (1.29)$$

Then we have to consider the periodic boundary condition  $x(\tau_a) = x(\tau_b)$ , and add the integration on  $d\tau_a = d\tau_b = d\tau_{n_t}$  to equation (1.27). Finally we can rewrite the partition function in terms of the path integral

$$Z = \int \mathcal{D}[x] e^{-S_E}, \quad (1.30)$$

where  $\mathcal{D}[x]$  is a compact notation for  $\prod_{i=1}^{n_t} dx_i$ .

What we found for a single particle can be extended to the quantum field theory and, in particular, to QCD. More information about the lQCD can be found in [8].

The partition function  $Z$  can be used to provide many physical observables and it is usually evaluated using Monte Carlo simulation techniques. The limit of this approach is essentially the large amount of computational resources needed, which fixes the limit of the minimum lattice spacing  $a$  that can be reached. For example, the right panel of figure 1.3 shows the pressure normalised by  $T^4$  as a function of the temperature obtained with lQCD calculations. The behaviour of  $p/T^4$  around  $T \sim 150$  MeV indicates an increase of the number of degrees of freedom because of the phase transition, from hadronic matter to QGP. However the Stefan-Boltzmann limit for the ideal gas is never reached (arrow in the plot). The Lattice QCD can also provide the order of the phase transition, which depends on the treatment of the quark masses. In particular, in the pure gauge approximation (in which only the gluonic degrees of freedom are considered) or in the case of massless quarks, the transition is expected to be of the first order. Assuming the scenario of three quark flavours with non zero masses and  $m_s > m_{u,d}$ , lQCD predicts a cross-over between the two phases above a critical temperature (corresponding to a critical point as shown in figure 1.2), while below this critical temperature the transition is expected to be of the first order.

### 1.3 Ultra-relativistic heavy-ion collisions

As discussed in the previous sections, the way to produce the deconfined state of QGP is to reach extreme conditions of temperature and energy density. It is also needed to create a spatially extended and long-lived system, in order to use the thermodynamics and fluidodynamics variables to describe its evolution. The experimental approach which fulfils all these requirements is achieved by colliding

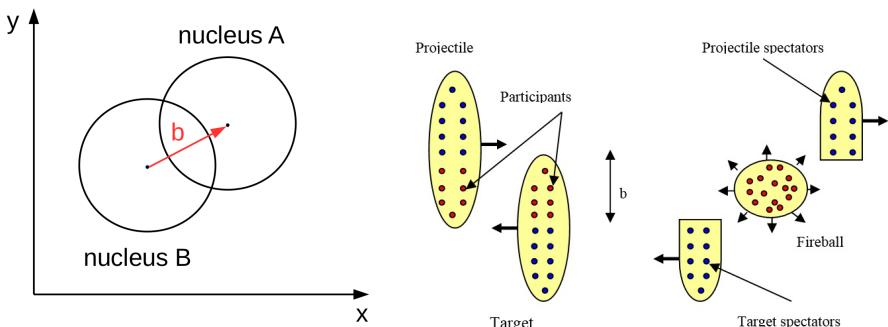


Figure 1.4. Left: the impact parameter of the nucleus-nucleus collision  $b$  in the transverse plane. Right: participants and spectators nucleons in a nucleus-nucleus collision.

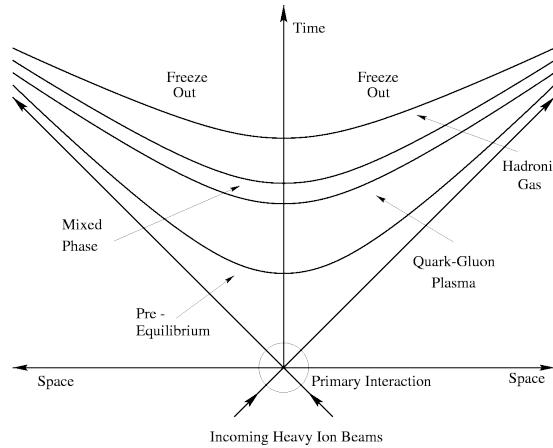


Figure 1.5. Evolution of a heavy-ion collision represented in a space-time diagram.

heavy nuclei at very high energies. Heavy-ion collisions are characterised by the mass number of the colliding nuclei, the centre-of-mass energy, which is usually expressed in energy per nucleon pair ( $\sqrt{s_{NN}}$ ), and the *centrality* of the collision, which is related to the impact parameter  $b$ , which is the vector connecting the centres of the two nuclei in the transverse plane, as shown in the left panel of figure 1.4.

### 1.3.1 Collision evolution

The collision between two heavy nuclei is an event with a complex space-time evolution and can be described considering several phases.

First of all it is possible to distinguish two different regimes. If the centre-of-mass energy is of the order of  $\sqrt{s_{NN}} \simeq 10$  GeV, the nucleons are stopped in the collision region, then the net baryon density is high. This condition is called *stopping regime*. With increasing beam energy also the rapidity of the beams increase and less nucleons are stopped. At high  $\sqrt{s_{NN}}$ , all the nucleons move away from the collision region, and a system is produced with high energy density and baryon number close to zero. In this case we reach the *Bjorken regime* or *transparency regime* [11]. If we now assume to satisfy this *Bjorken* condition, we can describe the evolution of the system with the following steps:

- **Collision:** the collision take place in a very short time, e.g. at the RHIC energy ( $\sqrt{s_{NN}} = 200$  GeV/c)  $\tau_{coll} = \frac{2R}{\gamma} \sim 0.1$  fm/c.
- **Pre-equilibrium:** after a time  $\tau_f > \tau_{coll}$  secondary particles are produced.

In the mid-rapidity region, where the net baryon density is zero, according to the *Bjorken regime*, it is possible to evaluate the energy density using the following relation

$$\epsilon_{BJ} = \frac{\langle m_T \rangle}{\tau_f A} \frac{dN}{dy}(\tau_f) \quad (1.31)$$

where  $A$  is transverse surface of the collision region,  $m_T$  is the transverse mass defined as

$$m_T = \sqrt{m^2 + pt^2}, \quad (1.32)$$

$dN/dy$  the particle multiplicity per unit of rapidity and  $\tau_f$  the formation time. Assuming  $\tau_f \sim 1$  fm/c and considering the  $dN/dy$  values measured at the centre-of-mass energy of RHIC, the energy density obtained is well above the critical value obtained in section 1.2.1. In particular at the RHIC energies it results to be  $\epsilon_{BJ} \sim 5.4$  GeV/fm<sup>3</sup>, while at LHC  $\sim 15$  GeV/fm<sup>3</sup>. A more precise evaluation of the formation time can be performed using the Heisenberg's uncertainty principle  $\tau_f \simeq \hbar/\langle m_T \rangle$ , where the transverse mass can be experimentally obtained from the measurement of

$$\langle m_T \rangle = \frac{dE_T/dy}{dN/dy} \quad (1.33)$$

In this case at the energies of RHIC the formation time results smaller than 1 fm/c, and then the energy density even higher.

- **QGP formation:** after the production of the secondary particles, during the expansion, the system reaches the thermal equilibrium through binary scatterings and the QGP forms. From hydrodynamic studies the time needed to reach the thermal equilibrium is evaluated to be  $0.6 \leq \tau_{eq} \leq 1$  fm/c at RHIC.
- **Hadronisation:** when the temperature decreases below the critical value for the phase transition the QGP cannot longer exist and the system hadronises. A gas of interacting hadrons is formed, which still expands.
- **Freeze-out:** below a certain temperature the hadrons cease to interact. First the inelastic interactions cease (*chemical freeze-out*), so the relative abundances of species are fixed, then also the elastic ones (*thermal freeze-out*) cease and the momentum spectra are defined.

### 1.3.2 The Glauber Model

The Glauber Model [12] allows one to describe the geometry of the collision and calculate useful variables such as the interaction probability, the number of participant nucleons ( $N_{part}$ ) and the number of binary nucleon-nucleon collisions ( $N_{coll}$ ), in a semi-classical approximation. The assumptions of the model are the following:

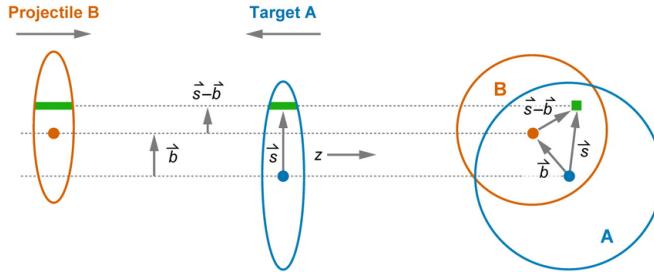


Figure 1.6. Schematic representation of two nuclei before a collision together with the coordinate system for the nucleons.

- The collision between two nuclei is described as incoherent superposition of nucleon-nucleon interactions.
- The nucleons inside the nuclei are indistinguishable (protons from neutrons) and point-like.
- The nucleus and the nucleons inside it are not deflected in the interaction.
- The cross section for the elementary interaction between two nucleons does not vary through all the collision.

In addition the Glauber Model needs two input quantities:

- The nucleon density distribution, which can be parametrised with different functions. One possibility is to use the *Wood-Saxon* (2 parameters Fermi) distribution, which well approximate spherical nuclei, e.g.  $^{115}In$  or  $^{208}Pb$ . Its expression is

$$\rho(r) = \frac{\rho_0}{1 + e^{\frac{r-r_0}{\delta}}}, \quad (1.34)$$

where  $\rho_0$  is the nuclear density in the core,  $r_0$  is the nuclear radius (the distance at which the nuclear density is equal to  $\rho_0/2$ ) and  $\delta$  is the skin depth. The interval  $[r_0 - 2\delta, r_0 + 2\delta]$  defines the region where the density decrease from 90% to 10% of the nuclear density  $\rho_0$ .

- The inelastic nucleon-nucleon cross section  $\sigma_{inel}$ .

Considering the collision between the nucleus A (with  $A$  nucleons) and the nucleus B (with  $B$  nucleons), we can place the origin of the coordinate system in the centre of A, and express the position on the transverse plane of the nucleons inside A with the vector  $\vec{s}$ . The same position with respect to the centre of the nucleus B is then given by  $\vec{b} - \vec{s}$ , as shown in figure 1.6. The probability to find a nucleon of the nucleus A at the transverse coordinate  $\vec{s}$  is therefore given by the *nuclear thickness function*

$$T_A(\vec{s}) = \int_{-\infty}^{+\infty} dz_A \rho_A(\vec{s}, z_i^A). \quad (1.35)$$

The probability to find one nucleon of A and one of B in the same infinitesimal surface element  $d^2 s$  and that they interact is then

$$dp(\vec{b}) = d^2 s \cdot T_A(\vec{s}) \cdot T_B(\vec{b} - \vec{s}) \cdot \sigma_{inel}. \quad (1.36)$$

Integrating the latter expression on the transverse coordinate  $\vec{s}$  we obtain the probability to have an elementary collision between two nucleons

$$p(\vec{b}) = \sigma_{inel} \int d^2 s \cdot T_A(\vec{s}) \cdot T_B(\vec{b} - \vec{s}) = \sigma_{inel} T_{AB}(\vec{b}), \quad (1.37)$$

where  $T_{AB}(\vec{b})$  is the *nuclear overlap function*.

The probability to have  $n$  nucleon-nucleon collisions during a nucleus-nucleus interaction at a given impact parameter  $b$  is then given by the binomial law

$$P_n(\vec{b}) = \binom{AB}{n} \left[ \sigma_{inel} T_{AB}(\vec{b}) \right]^n \left[ 1 - \sigma_{inel} T_{AB}(\vec{b}) \right]^{AB-n}. \quad (1.38)$$

The interaction probability between two nuclei is given by the probability to have at least one elementary interaction between two nucleons:

$$P_{AB}(\vec{b}) = 1 - P_0(\vec{b}) = 1 - \left[ 1 - \sigma_{inel} T_{AB}(\vec{b}) \right]^{AB}. \quad (1.39)$$

Considering the (1.38) we can estimate also the average number of elementary collisions  $N_{coll}$  as the mean value of the binomial distribution

$$N_{coll}(\vec{b}) = AB \sigma_{inel} T_{AB}(\vec{b}). \quad (1.40)$$

Furthermore, with similar calculations, it is possible to estimate the number of nucleons involved in the collision (*participants*) and the number of nucleons which do not interact (*spectators*). A schematic view of the participant and spectator nucleons is shown in the right panel of figure 1.4.

## 1.4 Experimental observables

There are many physical observables which are affected by the formation of the deconfined medium and therefore represent a very useful tool in order to understand if the QGP is formed and, if it is the case, characterise it. In particular we can distinguish between three categories:

- **Electromagnetic probes:** these are photons or lepton pairs, which are created by the black body radiation of the *fireball* in the early stages of the collision. The measurement of thermal photons is very difficult because of the large background of photons produced in successive phases (e.g. from  $\pi^0$  and Dalitz decays). The study of electromagnetic probes can provide a direct information from the reaction zone and its anisotropy, since they cannot interact with the QGP.
- **Soft probes:** the soft processes include all the effects and the particles produced in the final stages of the collision. The majority of the total number of particles are produced in soft processes.
- **Hard probes:** with hard probes we refer to particles originating in partonic scatterings with large transferred momentum, which correspond to short time. Therefore the particles produced in these processes are sensitive to all the phases of the system evolution and can interact with the medium eventually created.

In the next sections the soft and the hard probes will be discussed in some more details.

### 1.4.1 Soft probes

In this section some of the main observables in the soft sector together with the results obtained at the SPS, RHIC and LHC are briefly presented.

#### Particle multiplicity

As discussed in the previous section the energy density can be estimated from the particle multiplicity (see Bjorken formula 1.31). Therefore a measurement of this quantity can help to understand whether the QGP has been formed or not in the collisions. In order to apply the Bjorken formula the charged particle density in the region of mid-rapidity

$$\frac{dN_{ch}}{dy} \Big|_{y=0} \quad \text{or similarly} \quad \frac{dN_{ch}}{d\eta} \Big|_{\eta=0} \quad (1.41)$$

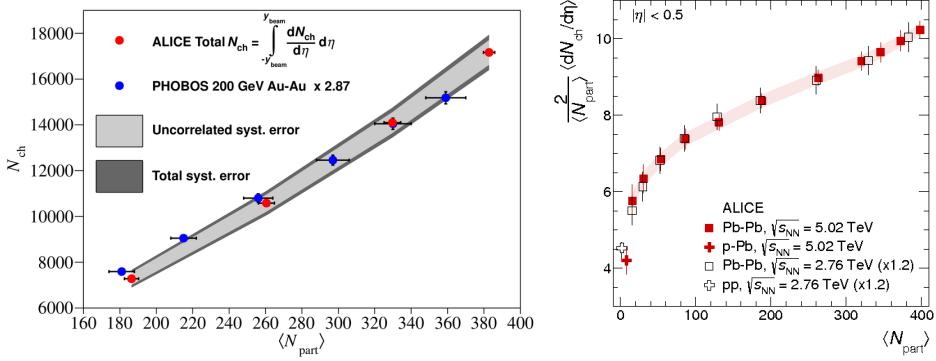


Figure 1.7. Left: total multiplicity as a function of the number of participant nucleons measured by ALICE in Pb-Pb collisions at  $\sqrt{s_{NN}} = 2.76$  TeV compared to the PHOBOS measurement for Au-Au collisions at  $\sqrt{s_{NN}} = 200$  GeV [13]. Right: particle density as a function of the number of participants measured by ALICE in pp collisions at  $\sqrt{s} = 2.76$  TeV, p-Pb collisions at  $\sqrt{s_{NN}} = 5.02$  TeV and Pb-Pb collisions at  $\sqrt{s_{NN}} = 2.76$  TeV and  $\sqrt{s_{NN}} = 5.02$  TeV [14].

and the total multiplicity

$$N_{ch} = \int \frac{dN}{d\eta} d\eta \quad (1.42)$$

are introduced. From the experimental observations in proton-nucleus collisions, the total multiplicity was expected to scale with the number of participant nucleons, therefore it is usually expressed normalised to the number of participant nucleon pairs. In the left plot of figure 1.7 it is possible to see that the measurements of total multiplicity by ALICE and PHOBOS [13] confirm at LHC and RHIC energies this scaling rule also for nucleus-nucleus collisions, which is no longer true for the particle density  $dN/d\eta$  at mid-rapidity (right panel) [14]. The multiplicity increases with the centre of mass energy of the collision. In figure 1.8 is shown the dependence of the particle density  $dN/d\eta$  per participant pair on the collision energy [14]. From this plot it is possible to notice that the slope is different for nucleus-nucleus collisions with respect to proton-proton ones.

### Multiplicity of identified particles

The measurement of identified hadron yields can provide an important estimation of the chemical freeze-out temperature  $T_{ch}$ , the baryon-chemical potential  $\mu_b$  and the volume of the deconfined matter  $V$ . Those parameters can be in fact evaluated by comparing the measured abundances with the predictions of the statistical

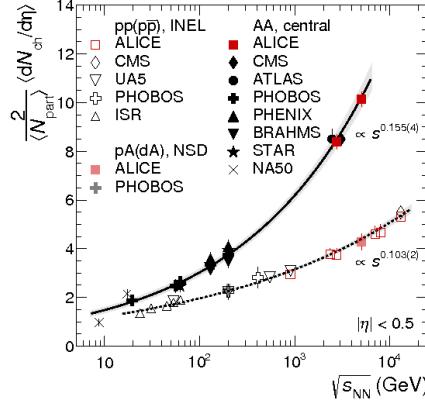


Figure 1.8. Particle density for nucleus-nucleus, proton-nucleus and proton-proton collisions as a function of the centre of mass energy measured by several experiments [14].

hadronisation model, which assumes that the system is in the thermal and chemical equilibrium at the chemical freeze-out. In particular, considering the system exchanging energy and particles with the outside (*Grand Canonical Ensemble*), the abundance of a defined species  $i$  is

$$N_i(T, V, \mu_i) = \frac{g_i V}{2\pi^2} \int_0^\infty \frac{p^2 dp}{e^{(E - \mu_i)/T} \pm 1} \quad (1.43)$$

where  $g_i$  are the degrees of freedom of the species  $i$ . In the left plot of figure 1.9 the fit to the total number of particles for different species in Pb-Pb collisions at  $\sqrt{s_{NN}} = 2.76$  TeV [15] is shown. In the right panel of the same figure is shown the comparison of the chemical freeze out temperatures obtained at different collision energies and the phase transition curve [16]. For values of  $\mu_B$  lower than  $\sim 400$  MeV, corresponding to the centre-of-mass energy of  $\sqrt{s_{NN}} \simeq 8-10$  GeV (SPS),  $T_{ch}$  is close to the expected phase boundary and therefore one can conclude that the chemical freeze-out happens right after the transition from QGP to hadronic matter.

### Transverse momentum particle spectra and collective flows

After the chemical freeze-out the species abundances are fixed, but the particles can still interact among themselves through elastic scatterings. Only after the thermal freeze-out all the interactions cease and therefore also the momentum spectra of the particles become fixed. If the system was in thermal equilibrium,

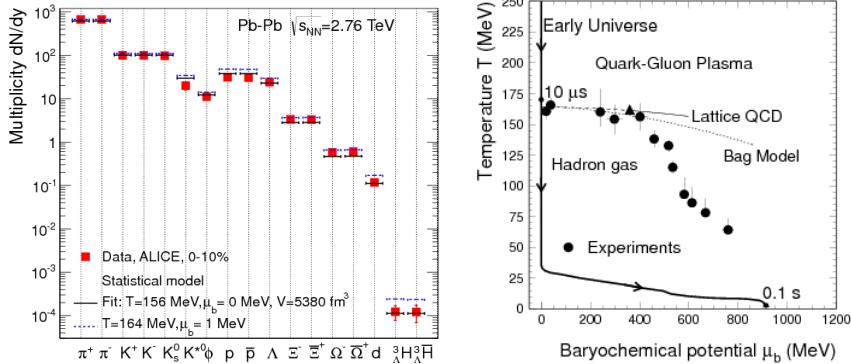


Figure 1.9. Left: fits to the  $p_T$ -integrated abundances for different hadron species in Pb-Pb collisions at  $\sqrt{s_{NN}} = 2.76$  TeV measured by ALICE [15]. Right: QCD phase diagram and data from a chemical freeze-out analysis in nucleus-nucleus collisions at various beam energies. The data points closest to the  $T$ -axis are from the highest collision energies [16].

then the transverse momentum spectrum at low  $p_T$  is described by the Maxwell-Boltzmann distribution, where the parameter  $T_{fo}$  represents the thermal freeze-out temperature

$$\frac{1}{pt} \frac{dN}{dp_T} \propto e^{-\sqrt{m^2 + pt^2}/T_{fo}}. \quad (1.44)$$

This is the case of proton-proton collisions, while in Pb-Pb collisions, where the QGP is formed, has been observed a different slope parameter  $T_{slope}$  for particles with different masses [17]. This is the effect of a collective motion in the transverse plane, called *radial flow*, as superimposed to the thermal random motion due to the internal pressure generated by the fireball. Therefore the slope parameter is the sum of the thermal freeze-out temperature and the radial flow thermos:

$$T_{slope} = T_{fo} + \frac{1}{2} m \langle v_T \rangle^2, \quad (1.45)$$

where  $\langle v_T \rangle$  is the collective transverse velocity of the medium. This effect is shown in figure 1.10, where the transverse momentum spectra in different centrality classes for pions and protons are compared. The fits have been performed with a Blast-wave function, which is not a simple Maxwell-Boltzmann distribution, but it takes also into account the expansion of the thermal source [18]. Another important collective motion, which has been observed in nucleus-nucleus collisions, is the

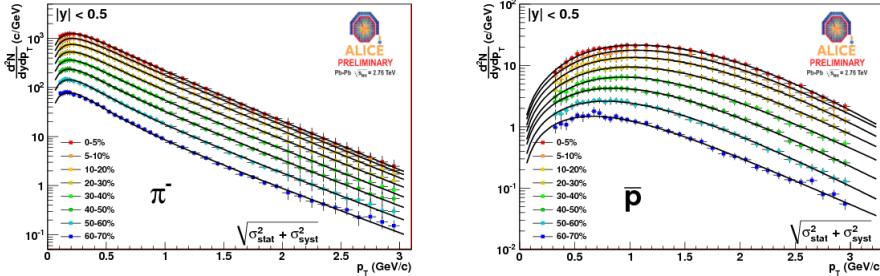


Figure 1.10. Transverse momentum spectra for pions and protons in different centrality classes in Pb-Pb collisions at  $\sqrt{s_{NN}} = 2.76$  TeV measured by ALICE [17].

*anisotropic flow*. This is established especially in non-central collisions, because of the almond shape of the interaction volume (see left panel of figure 1.11). The initial spacial anisotropy is converted in transverse momentum anisotropy due to anisotropic pressure gradients and therefore it can be measured through the azimuthal angle distributions of the particles in the final state. The anisotropy is characterised via a Fourier series expansion of the azimuthal angle with respect to the plane defined by the impact parameter and the beam direction (*reaction plane*)

$$\frac{dN}{d(\varphi - \Psi_{RP})} = \frac{N_0}{2\pi} \sum_{n=0}^{\infty} 2v_n \cos[n(\varphi - \Psi_{RP})], \quad (1.46)$$

where  $v_n$  are the Fourier coefficients defined as

$$v_n = \langle \cos[n(\varphi - \Psi_{RP})] \rangle. \quad (1.47)$$

In particular the  $v_2$  coefficient (*elliptic flow*) is expected to be larger than zero, since it represents the anisotropy between the parallel and orthogonal directions with respect to the impact parameter vector. In the right panel of figure 1.11 is shown the  $v_2$  measurement as a function of  $p_T$  for charged particles in different centrality classes performed by ALICE in Pb-Pb collisions at  $\sqrt{s_{NN}} = 2.76$  TeV [19]. The conversion between the initial geometrical anisotropy to the final momentum anisotropy is very sensitive to viscosity of the QGP. In particular, the comparison of the measured  $v_2$  with the theoretical models leads to a value of the shear viscosity to entropy ratio of about  $\eta/s \approx 0.2$  for both RHIC and LHC energies, which correspond to a nearly perfect fluid, with a hint for an incomplete thermalisation or dissipative effects in the hadronic phase [20].

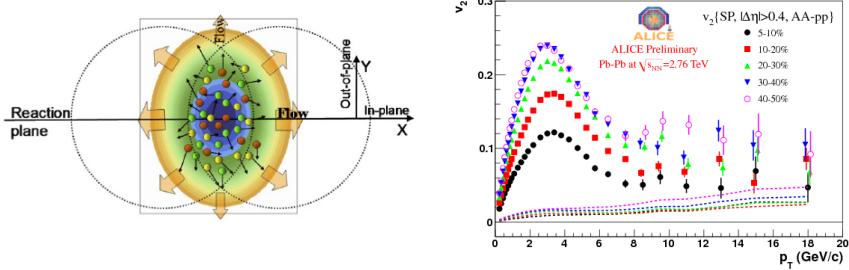


Figure 1.11. Left: sketch of a non-central nucleus-nucleus collisions with the almond-shaped interaction volume. Right: the elliptic flow in different centrality classes in Pb-Pb collisions at  $\sqrt{s_{\text{NN}}} = 2.76$  TeV measured by ALICE [19].

### 1.4.2 Hard probes

In this section I will discuss the *jet-quenching* and the *quarkonia suppression*, which represent two of the most important experimental observables in the sector of the hard probes. Moreover the next chapter is dedicated to the *open heavy-flavour* production, which is the main topic of this thesis.

#### High- $p_T$ hadrons and jet quenching

In heavy ion collisions, the production of particles in hard processes is expected to scale with the number of binary collisions  $N_{\text{coll}}$  (*binary scaling*), in fact

$$P_{AB}^{\text{hard}}(b) = 1 - \left[ 1 - \sigma_{\text{hard}} T_{AB}(b) \right]^{AB} \approx AB \sigma_{\text{hard}} T_{AB}(b) \propto \sigma_{\text{hard}} N_{\text{coll}}, \quad (1.48)$$

where, as compared to equation 1.39, we have substituted the hard cross section  $\sigma_{\text{hard}}$  to the total inelastic cross section  $\sigma_{\text{inel}}$ . Therefore we expect to obtain the  $p_T$ -spectra in A-A collisions  $dN_{AA}/dp_T$  simply rescaling those obtained in pp collisions  $dN_{pp}/dp_T$ . However, a quark or a gluon traversing the QGP loses part of its energy interacting with the partons of the deconfined medium. As a consequence, the  $p_T$ -spectrum of the hadrons in A-A collisions is shifted to lower values of the transverse momentum, resulting to a break of the scaling rule mentioned above. The variable commonly used to test the binary scaling is the *nuclear modification factor*, defined as

$$R_{AA}(p_T) = \frac{1}{\langle N_{\text{coll}} \rangle} \frac{dN_{AA}/dp_T}{dN_{pp}/dp_T}. \quad (1.49)$$

It is important to notice that the  $R_{AA}$  is only sensitive to the modification of the  $p_T$  spectra, while the binary scaling can be satisfied at the level of  $p_T$ -integrated

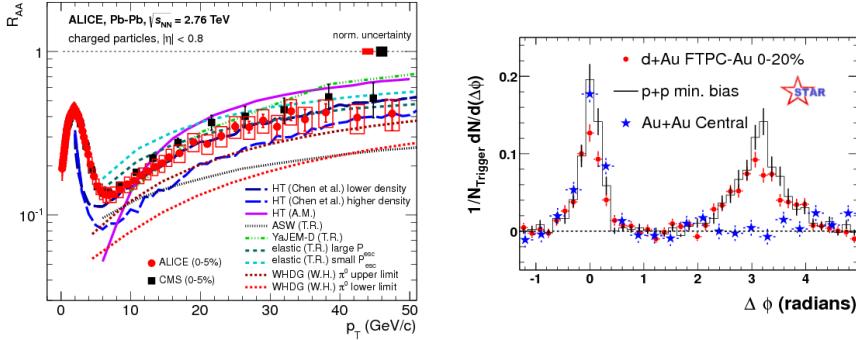


Figure 1.12. Left: the nuclear modification factor of charged particles measured by ALICE in the most central Pb-Pb collisions at  $\sqrt{s_{NN}} = 2.76$  TeV compared to the result from CMS and the theoretical models [21]. Right: angular correlation between high  $p_T$  hadrons measured by STAR in pp, d-Au and Au-Au collisions [22].

yields, even if  $R_{AA}(p_T) \neq 1$ . In the left panel of figure 1.12, the  $R_{AA}$  of charged particles measured by ALICE in the most central (0-5%) Pb-Pb collisions at  $\sqrt{s_{NN}}$  measured by ALICE and CMS is shown. The two measurements are well in agreement within the uncertainties and both show a suppression for  $p_T \gtrsim 1$  GeV/c due to the energy loss. For lower values of the transverse momentum the hadrons are produced via soft processes and therefore these considerations cannot be applied. In addition to the leading hadrons, it is also interesting to study the *jets*, which are collimated sprays of hadrons, originating from the fragmentation of a quark or a gluon with high transverse momentum created in the collision. Interesting measurements connected to the jets can be performed through the azimuthal correlations of high- $p_T$  hadrons with respect to the one with the highest transverse momentum (*trigger particle*). The hadrons associated to the trigger particle are selected setting a minimum  $p_T$  threshold. In proton-proton collisions, for leading order processes, we expect hadrons coming from two back-to-back jets, therefore the angular distribution should have two peaks, one centred at  $\Delta\varphi = 0$  (*near-side peak*), and another peak around  $\Delta\varphi = \pi$  (*away-side peak*). In central heavy-ion collisions, the away-side peak can be strongly suppressed, depending on the selected  $p_T$  threshold of the associated hadrons. This observation can be understood in terms of travelled distance in the deconfined medium. In fact, if the original partons are produced on the surface of the fireball, the one which is directed outward produces the jet which is observed at  $\Delta\varphi = 0$ , while the other one loses a large amount of its energy traversing the QGP and hence does not contribute to the away-side peak. STAR was the first experiment that measured this effect [22],

which is a consequence of the softening of the jets (*jet quenching*). In figure 1.12 is shown the azimuthal angle correlation between high- $p_T$  hadrons in pp, d-Au and Au-Au collisions. The energy imbalance between the two back-to-back jets due to the different path through the QGP can also be measured via the di-jet asymmetry factor

$$A_J = \frac{E_{T,1} - E_{T,2}}{E_{T,1} + E_{T,2}}, \quad (1.50)$$

where  $E_{T,1}$  is the transverse energy of the jet with the highest energy of the event (*leading jet*) and  $E_{T,2}$  refers to the highest transverse energy jet with  $|\Delta\varphi| > \pi/2$  (*sub-leading jet*). The ATLAS and CMS Collaborations found that in peripheral Pb-Pb collisions  $A_J$  is similar to the one in pp, while increasing the centrality more events with a large value of the di-jet asymmetry are observed [23] [24].

### Quarkonium production

In the collision the  $c$  and  $b$  quarks must be produced in hard scattering processes together with the relative anti-quark, since there is no net charm or bottom charge in the initial state. A fraction of these  $q\bar{q}$  pairs forms bound states, named *quarkonia*. The strength of the binding interaction between two quark and the anti-quark which form the quarkonium state can be evaluated using the *Cornell potential*

$$V_{QCD}(r) = -\frac{4}{3} \frac{\alpha_s(r)}{r} + kr \quad (1.51)$$

where  $\alpha_s$  is the strong coupling constant,  $r$  the distance between the quarks and  $k$  a constant called *string tension*. Due to their large mass,  $c$  and  $b$  quarks are created in the initial phase of the collision, therefore, if the QGP is created, they can interact with it and we should observe a modification of the formation of the quarkonia. In particular, since in the QGP quarks and gluons are not bound in hadrons, we can expect a screening between the  $q$  and the  $\bar{q}$  due to the colour charge present in the deconfined medium. This phenomenon is called *Debye screening* and has the effect to reduce the potential, as described by the following formula:

$$V_{QCD}(r) = -\frac{4}{3} \frac{\alpha_s(r)}{r} e^{-r/\lambda_D}, \quad (1.52)$$

where  $\lambda_D$  (*Debye length*) represents the maximum distance at which the partons can still interact. In this scenario, because of the screening, it becomes less probable for the heavy  $q\bar{q}$  quarks to bind together in a quarkonium state. Therefore the charmonium and bottomonium states should be suppressed in nucleus-nucleus collisions with respect to proton-proton ones. Since Debye length  $\lambda_D$  depends on the temperature of the medium, for different temperatures, different states are suppressed. In particular the excited states, which are less strongly bound and

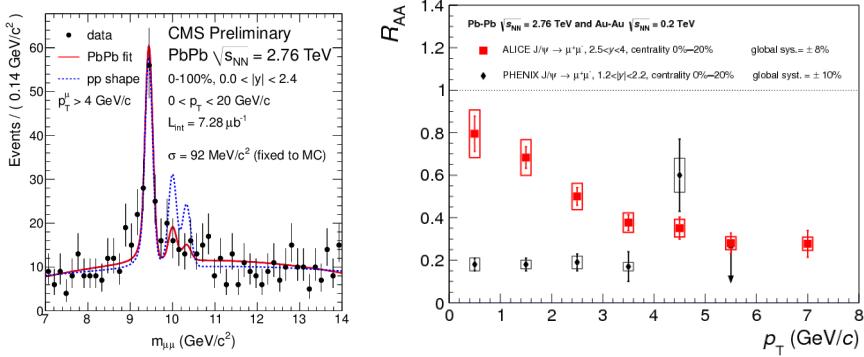


Figure 1.13. Left: CMS dimuon invariant-mass spectrum in the  $\Upsilon(1S)$ ,  $\Upsilon(2S)$  and  $\Upsilon(3S)$  bottomonium states region, in  $\text{PbPb}$  collisions at  $\sqrt{s_{NN}} = 2.76 \text{ TeV}$ , compared to the rescaled result obtained in  $\text{pp}$  collisions at  $\sqrt{s_{NN}} = 7 \text{ TeV}$  [25]. Right: the nuclear modification factor for the  $\text{J}/\psi$  as a function of  $p_T$  measured by ALICE in central  $\text{Pb-Pb}$  collisions at  $\sqrt{s_{NN}} = 2.76 \text{ TeV}$  compared to the PHENIX result in central  $\text{Au-Au}$  collision at  $\sqrt{s_{NN}} = 200 \text{ GeV}$  [26].

therefore the distance between the  $q$  and the  $\bar{q}$  is larger, are suppressed for lower temperatures. An example of this effect is shown in the left panel of figure 1.13, where the rescaled lineshape of the dimuon invariant-mass fit obtained in  $\text{pp}$  collisions is superimposed to the spectrum measured in nucleus-nucleus collisions. It is clear from this picture how the mass peaks, corresponding to the bottomonium states  $\Upsilon(2S)$  and  $\Upsilon(3S)$ , are strongly suppressed in  $\text{Pb-Pb}$  collisions compared to proton-proton collisions. Therefore the measurement of the quarkonia suppression can also provide information on the temperature of the QGP. Furthermore the heavy quarks screened by the coloured medium can hadronise outside the medium or, they can combine with another quark in the medium itself. In particular, if the number of heavy quarks produced is high, they have a not negligible probability to meet a corripective anti-quark and to form a quarkonium state. This phenomenon, called *recombination* is expected to happen for the charmonium much more at LHC, where the multiplicity of  $c$ -quarks in  $\text{Pb-Pb}$  collisions is  $\sim 200$ , with respect to RHIC where  $\sim 10$  charm quarks per  $\text{Au-Au}$  collision are created. In the right panel of figure 1.13 is shown the comparison between the  $\text{J}/\psi$  nuclear modification factor as a function of  $p_T$  measured by PHENIX in  $\text{Au-Au}$  collisions at  $\sqrt{s_{NN}} = 200 \text{ GeV}$  with the one measured by ALICE in  $\text{Pb-Pb}$  collisions at  $\sqrt{s_{NN}} = 2.76 \text{ TeV}$  [26]. At high  $p_T$  the two measurements become compatible within the experimental uncertainties, while at low  $p_T$  the ALICE measurement shows a significantly lower suppression, because of the larger recombination.

# 2

## Open heavy-flavour physics

The production of open heavy-flavoured hadrons (particles made of one charm or beauty quark and other lighter quarks (e.g. D or B mesons and  $\Lambda_c$  or  $\Lambda_b$  baryons) is a unique tool in high energy physics to study many aspects of the strong interaction. First of all, because of their large mass, their production can be always in principle described by a perturbative theory. Therefore their measurement in proton-proton collisions can be a benchmark for the pQCD and can provide constraints for the theory parameters. In nucleus-nucleus collision they give us an important opportunity to probe hot dense matter, through the energy-loss mechanism. Finally, in proton-nucleus interactions, where the formation of the QGP is not expected, open-heavy flavours measurement can provide information on Cold Nuclear Matter effects.

### 2.1 Heavy-quark production in pp collisions

Heavy quarks are always produced in scattering processes with large momentum transfer ( $Q^2 \gtrsim 4m_q^2$ ). Since the strong coupling constant is much smaller than unity in this  $Q^2$  region (see section 1.1.2), the elementary partonic cross section  $\sigma_{ij \rightarrow q\bar{q}}$  can be calculated with a perturbative expansion. The heavy-flavoured hadron production cross section in pp collisions  $\sigma_{pp \rightarrow HX}$  can be evaluated with the following

factorisation formula

$$\sigma_{pp \rightarrow HX} = \sum_{ij=g,q,\bar{q}} f_i(x_i, Q^2) f_j(x_j, Q^2) \otimes \sigma_{ij \rightarrow q\bar{q}} \otimes D_{q \rightarrow H}(z_q, Q^2), \quad (2.1)$$

where  $f(x_{i,j}, Q^2)$  are the parton distribution functions (PDFs) for the two interacting partons in the colliding protons and  $D_{q \rightarrow H}(z_q, Q^2)$  is the fragmentation function of the quark  $q$  into the hadron  $H$ . Both these functions describe non-perturbative processes and therefore cannot be calculated with pQCD.

The parton distribution function represents the probability to find a parton  $i$  carrying a fraction  $x_i$  of momentum of the proton. If the partons were non interacting point-like Dirac particles inside the protons, this probability should depend only on the parton type and the momentum fraction  $x$  (*Bjorken scaling*). If we consider the interaction between partons (especially the gluon emission), moreover we find that the PDFs depend also on the transferred momentum  $Q^2$ . Hence the parton distribution functions have to be measured at a certain energy scale  $Q_0^2$ , and then their evolution with  $Q^2$  can be calculated with the so-called *DGLAP* equations (further details can be found in [27], [28] and [29]).

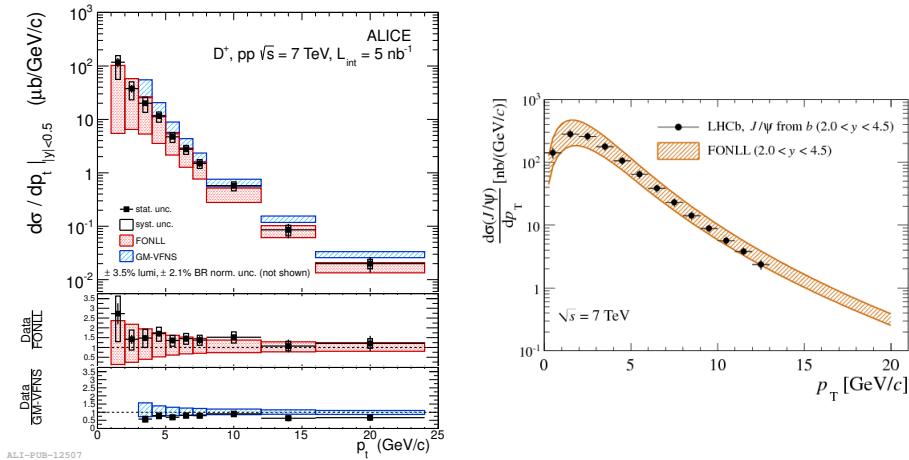


Figure 2.1. Left:  $p_T$ -differential cross section for prompt  $D^+$ -meson production measured by ALICE in pp collisions at  $\sqrt{s} = 7$  TeV compared with the FONLL and GM-VFNS calculations [30]. Right: The non-prompt  $J/\psi$   $p_T$ -differential cross section measured by LHCb in pp collisions at  $\sqrt{s} = 7$  TeV compared to  $J/\psi$  from  $b$  production as predicted by FONLL [31].

The fragmentation function describes the probability of hadronisation of the quark  $q$  into the hadron  $H$ , which carry a fraction of momentum of the quark  $z_q = p_H/p_q$ . As well as the PDFs, also the fragmentation functions are obtained from the experiment, in particular in  $e^+e^-$  collisions, in which all the process before the hadronisation is perturbative, and therefore precisely calculated. Also for the fragmentation function case the evolution with  $Q^2$  is evaluated with the DGLAP equations. These calculations describe well the experimental data, especially for the beauty hadrons. An example is the measurement of the cross section of  $J/\psi$  from  $b$  decays (non-prompt  $J/\psi$ ) in pp collisions at  $\sqrt{s} = 7$  TeV performed by LHCb [31]. In the right panel of figure 2.1 it is possible to see the comparison of the data with the pQCD prediction (FONLL). For the charmed hadrons the experimental results are also in agreement with the FONLL prediction, however all the measurements performed at Tevatron, RHIC and LHC show that the data are systematically on the upper edge of the theoretical prediction. An example of this behaviour is the  $p_T$ -differential cross section of prompt  $D^+$ -mesons ( $D^+$  mesons coming directly from the hadronisation of a  $c$ -quark) measured in pp collisions at  $\sqrt{s} = 7$  TeV by ALICE [30] (left panel of figure 2.1).

## 2.2 Heavy-quark production in p-A and A-A collisions

In heavy-ion collisions the production mechanism of heavy quarks is the same as in proton-proton collisions, since they are predominantly created in hard partonic scatterings, because of their large mass, in fact the thermal production in the QGP is negligible at the temperatures reached at RHIC and LHC. Therefore, neglecting medium effects, the hard particles production in heavy-ion collisions is expected to scale with the number of binary collisions, as explained in 1.4.2. However, as for the high- $p_T$  hadrons, considering the  $p_T$  spectrum, the binary scaling is broken by the interactions with the nuclear matter, which can be divided in two main categories:

- Cold Nuclear Matter (CNM) effects
- Hot Medium effects

The first category includes all the initial and final state effects which are not related to the formation of the QGP, but only to the presence of the nuclear matter, such as the modification of the PDFs in the nuclei with respect to free nucleons. On the contrary, the hot medium effects are related to the interaction with the deconfined medium, therefore they can be only final state effects. These modifications can be investigated and quantified with the nuclear modification factor, which was already defined in section 1.4.2. Here I report two different definitions, for A-A

and p-A collisions, which are equivalent to the one in equation 1.49. In case of nucleus-nucleus collisions the  $R_{AA}$  can be expressed with the following formula:

$$R_{AA} = \frac{1}{\langle T_{AA} \rangle} \frac{dN_{AA}/dp_T}{d\sigma_{pp}/dp_T}, \quad (2.2)$$

where  $dN_{AA}/dp_T$  is the  $p_T$ -differential corrected yield measured in  $A - A$  collisions,  $\langle T_{AA} \rangle$  the average nuclear overlap function and  $d\sigma_{pp}/dp_T$  the  $p_T$ -differential cross section measured in pp collisions. The same quantity, computed for p-A collisions can be written as

$$R_{pA} = \frac{1}{\langle T_{pA} \rangle} \frac{dN_{pA}/dp_T}{d\sigma_{pp}/dp_T} = \frac{1}{A} \frac{d\sigma_{pA}/dp_T}{d\sigma_{pp}/dp_T}, \quad (2.3)$$

where  $\sigma_{pA}/dp_T$  is the  $p_T$ -differential cross section measured in p-A collisions and  $A$  the mass number of the colliding nuclei. If there are no modifications due to initial or final state effects then, because of the binary scaling, we should observe a nuclear modification factor equal to 1 in the full  $p_T$  range, otherwise it should be different. In particular, the various effects lead to different modifications in different  $p_T$  regions, as explained in the following sections.

### 2.2.1 Cold Nuclear Matter effects

In this section I will provide a brief explanation of initial-state CNM effects, which include the nuclear modification of the PDFs, the *Cronin enhancement* and the initial-state energy loss in Cold Nuclear Matter (the energy-loss mechanism will be described in the following section, since is the same for heavy quarks travelling through the QGP). However, experimental observations might suggest also the presence of final-state effects in p-Pb collisions at LHC, due to the interaction of the produced heavy-flavoured hadrons with the particles produced in the collision, for which some kind of collective behaviour could be established.

#### Nuclear modification of the Parton Distribution Functions

If the nucleus were a simple collection of almost free nucleons, the only modification expected for the PDFs would be due to the Fermi motion of the protons and neutrons inside the nucleus, which leads to a modification of the distribution of the fraction of the nucleon momentum  $x$  carried by the partons. However, as observed for the first time by the European Muon Collaboration (EMC) in 1982 in muon-nuclei Deep Inelastic Scatterings experiments, there are other effects that should be taken into account. In order to quantify the modifications of the PDFs in bound nucleons it is possible to compute the ratio between the DIS cross section measured in experiments involving nuclei and the one with protons (or deuterium),

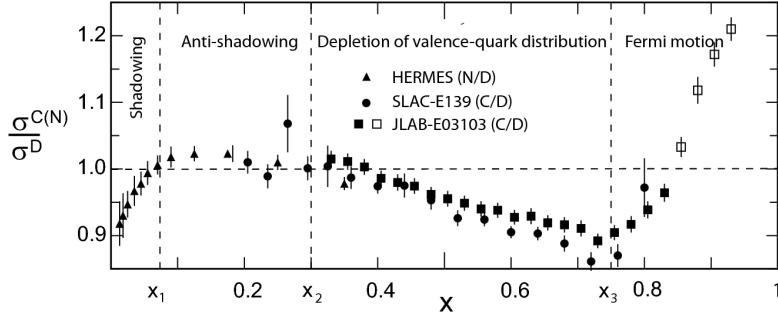


Figure 2.2. Measurements of the nuclear modification of the PDFs performed by several experiments in a wide range of Bjorken  $x$ .

normalised by the mass number  $A$  of the nucleus. This ratio can be equivalently expressed using the structure functions  $F_2(x, Q^2)$ , instead of the cross sections  $\sigma_{DIS}(x, Q^2)$

$$R^A(x, Q^2) = \frac{\sigma_{DIS}^A(x, Q^2)}{A\sigma_{DIS}^{nucleon}(x, Q^2)} = \frac{F_2^A(x, Q^2)}{AF_2^{nucleon}(x, Q^2)} \quad (2.4)$$

This quantity has been measured by many experiments for several nuclei and different intervals of *Bjorken*  $x$ . As shown in figure 2.2.1 it is possible to distinguish four different regions of  $x$ :

- For  $x \lesssim 0.1$  a significant suppression is observed, which increases with the mass number of the nucleus. This interval is called *shadowing region* [32]. This effect can be explained by a phenomenological model (*Color Glass Condensate*), which predicts a saturation of the partonic density at very low  $x$ , which leads to gluon fusion processes.
- The interval  $0.1 \lesssim x \lesssim 0.3$  identifies *the anti-shadowing region* where the ratio is slightly above the unity.
- For  $0.3 \lesssim x \lesssim 0.8$  the ratio is lower than unity, reaching a minimum around  $x \simeq 0.7$ . This suppression is called *EMC effect* [33].
- For  $0.8 \lesssim x \lesssim 1$  the ratio is above unity because of the Fermi motion.

Both the EMC effect and the shadowing increase with the mass number  $A$  or the density of the nucleus considered, while they almost do not show any dependence from the transferred momentum  $Q^2$ .

This modification affects the charm and bottom production in collisions involving nuclei, as it can be easily deduced from the factorisation theorem (see equation 2.1). In particular, at LHC energies, the heavy-flavour production is expected to be influenced mostly by the shadowing, since, at central rapidity,  $c\bar{c}$  and  $b\bar{b}$  pairs are produced in hard scatterings with values of Bjorken  $x$  around  $\sim 10^{-4}$  for charm and  $\sim 10^{-3}$  for beauty, increasing at high transverse momentum [34].

### Cronin enhancement

The *Cronin enhancement*, which was observed for the first time in p-A experiments at Fermilab [35], consists in an increase of the nuclear modification factor above unity, due to a shift of the  $p_T$ -spectrum toward higher values of transverse momentum. The interpretation of this effect is based on the idea that the partons inside the projectile particle go through multiple elastic collisions with the constituents of the target nucleus (*random walk*), before the hard scattering. These elastic interactions transfer an initial transverse momentum  $k_T$  to the partons, which is responsible of the  $p_T$ -spectrum shift for the particles produced in hard scatterings, such as heavy-quarks.

### Nuclear modification factor in p-A collisions

Since the QGP formation is not expected in proton-nucleus collisions, their study is important to disentangle the effects of Quark-Gluon Plasma from those due to Cold Nuclear Matter, and to provide essential input to the understanding of the modifications observed in nucleus-nucleus collisions. An example of measurement in p-A collisions is shown in the left panel of figure 2.3. The nuclear modification factor of D-mesons in the rapidity range  $-0.96 < y_{cms} < 0.04$  measured in p-Pb collisions at  $\sqrt{s_{NN}} = 5.02$  TeV by the ALICE experiment [36] can give a hint of an enhancement in the  $p_T$  interval between 3 and 6 GeV/c, which can be attributed to the Cronin effect. However the effect is small and the experimental data are compatible with unity within the uncertainties. The measurement is well reproduced by predictions which consider only initial-state effect, based either on next-to-leading order (NLO) pQCD calculations [37] including the EPS09 modification of the CTEQ6M PDFs [38], or on the Color Glass Condensate [39]. Also calculations which include Cold Nuclear Matter energy loss, nuclear shadowing and  $k_T$ -broadening describe well the data [40]. Another interesting measurement is the  $p_T$ -integrated  $R_{pPb}$  for  $J/\psi$  coming from B-hadrons decays as a function of rapidity [41], performed by LHCb (right panel of figure 2.3). In this case the experimental data are compatible with unity at backward rapidity (Pb-going direction), while they show a modest suppression at forward rapidity (p-going direction). This measurement is well described by NLO pQCD calculations with the EPS09 modification of the PDFs as shown in the figure. The suppression in the p-going direction

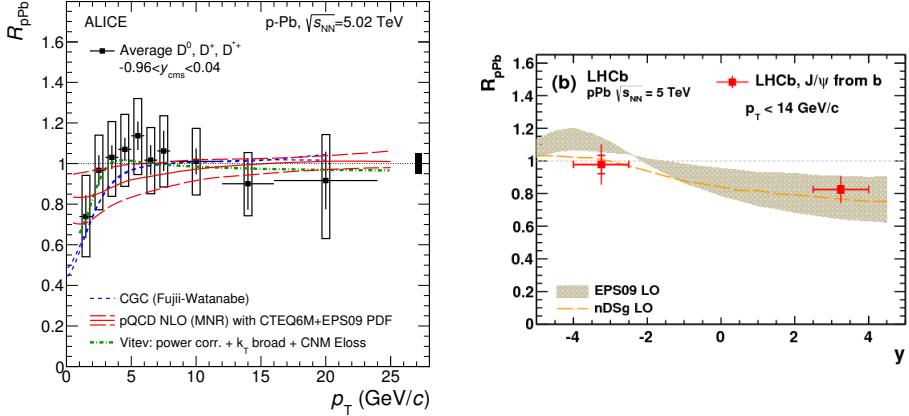


Figure 2.3. Left: nuclear modification factor  $R_{pPb}$  for prompt D-mesons (average of  $D^0, D^+$  and  $D^{+*}$ ) as a function of the transverse momentum measured with the ALICE detector in  $p$ - $Pb$  collisions at  $\sqrt{s_{NN}} = 5.02$  TeV compared to model calculations [36]. Right:  $p_T$ -integrated nuclear modification factor  $R_{pPb}$  for non-prompt  $J/\psi$  as a function of rapidity measured by LHCb in  $p$ - $Pb$  collisions at  $\sqrt{s_{NN}} = 5.02$  TeV [41].

is explained with the shadowing effect, since at forward rapidity smaller values of  $x$  in the Pb nucleus are probed. In both cases the results show that the Cold Nuclear Matter effects are small, especially at high transverse momentum.

### 2.2.2 Hot Medium effects

As described in section 2.1, because of their large mass, heavy quarks are exclusively produced via hard scatterings and therefore, according to the Heisenberg's uncertainty principle, they are created in the initial phase of the collision, before the QGP formation:

$$\tau_{prod} \sim \frac{\hbar}{m_{b,c}} \sim 0.05 - 0.1 \text{ fm/c.} \quad (2.5)$$

Hence they subsequently interact with the constituents of the deconfined medium, and the modification of their dynamical properties is a powerful tool to investigate the Quark-Gluon Plasma.

### In-medium energy loss

Heavy quarks traversing a deconfined medium lose energy via elastic scatterings with the partons of the medium (*collisional energy loss*) and gluon radiation (*Gluonsstrahlung*). The first mechanism is similar to the ionisation energy loss of charged particles in ordinary matter via electromagnetic interaction, while the second one is the analogous in QCD of the QED *Bremsstrahlung*. For a high energy parton the dominant mechanism is expected to be the gluon radiation, which can be described by the so-called BDMPS model [42]. According to this theoretical model, the mean energy loss

$$\langle \Delta E \rangle \propto \alpha_s C_R \hat{q} L^2 \quad (2.6)$$

is proportional to the square of the distance  $L$  traversed in the medium, to the strong coupling constant  $\alpha_s$ , to the colour coupling factor  $C_R$  (which is equal to 3 for gluons and equal to 4/3 for quarks) and to a parameter  $\hat{q}$  called transport coefficient, which depends on the gluon density (and therefore the energy density) of the medium. Moreover the *dead cone effect* [43] predicts that the energy loss decreases with increasing mass of the travelling parton, since the gluon emission is forbidden at angles smaller than

$$\theta_c = M_q/E_q, \quad (2.7)$$

where  $M_q$  and  $E_q$  are respectively the mass and the energy of the quark.

### Heavy-quark thermalisation

The thermalisation of light quarks in the QGP is reached via multiple scatterings after a time of about  $0.6 - 1$  fm/c at RHIC energies and even earlier at LHC. On the contrary, the thermal equilibrium of heavy quarks needs a longer time to be established, due to their larger mass, and eventually it can also not be reached. In particular, it is more likely expected to be reached for charm and beauty quarks with low transverse momentum, since they have to lose less energy in the medium to thermalise. If the thermal equilibrium is completely reached, then the heavy quarks are subject to the same pressure gradients which act on light quarks, as explained in 1.4.1. Concerning that, the nuclear modification factor at low  $p_T$  is a sensitive observable for the radial flow and can be an interesting tool to investigate the thermalisation of heavy quarks.

### Hadronisation via coalescence

The heavy quarks which are produced in the core of the fireball could not have sufficient energy to emerge from the medium before hadronising. In this case, they can hadronise through a different mechanism, called *coalescence*, which basically

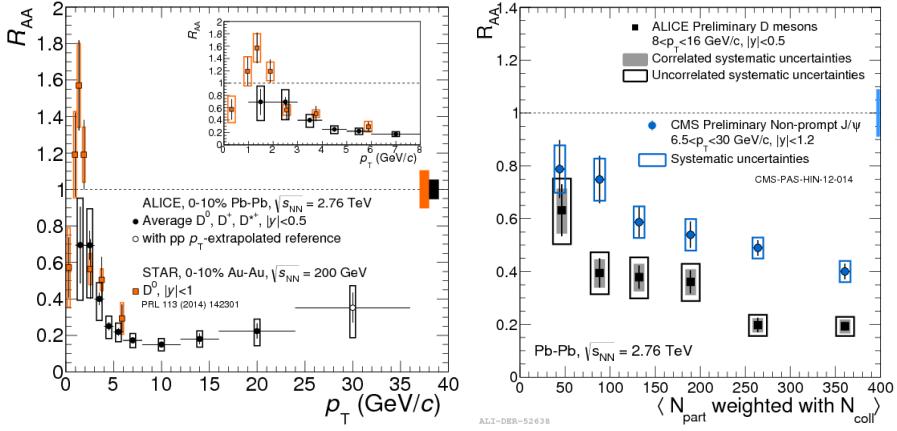


Figure 2.4. Left: prompt D-meson (average of  $D^0, D^+$  and  $D^{*+}$ ) nuclear modification factor as a function of  $p_T$  in the 10% most central Pb-Pb collisions measured by ALICE at  $\sqrt{s_{NN}} = 2.76$  TeV compared to  $D^0 R_{AA}$  measured by STAR in Au-Au collisions at  $\sqrt{s_{NN}} = 200$  GeV [45]. Right: prompt D-meson  $R_{AA}$  as a function of the number of participants measured by ALICE compared to  $R_{AA}$  of non-prompt  $J/\psi$  measured by CMS in Pb-Pb collisions at  $\sqrt{s_{NN}} = 2.76$  TeV [44].

consists in the recombination with a light quark of the medium to form a heavy-flavoured hadron. The main difference of this mechanism with respect to the fragmentation described in section 2.1 is that, in this case, two low momentum partons form a higher momentum hadron, while in the fragmentation the final hadron has a fraction of the initial parton momentum  $z_q = p_H/p_q$ . In particular, it happens when two partons are close in the phase space and therefore have a significant overlap of their wave functions. The recombination picture becomes even more important in the presence of a thermalised medium, in which collective motions increase the probability to find a light quark close in space and with similar momentum with respect to the heavy quark. This alternative hadronisation method modifies the probability for the heavy quark to hadronise into a specific hadron  $H$  and therefore the production mechanism of charmed and beauty hadrons.

### Nuclear modification factor in A-A collisions

The nuclear modification factor is an observable sensitive to the energy-loss mechanism, the radial flow and the shadowing effect which occur in A-A collisions. For example the left panel of figure 2.4 shows the prompt D-meson  $R_{AA}$  (average of  $D^0, D^+$  and  $D^{*+}$ ) as a function of  $p_T$  in the 10% most central Pb-Pb collisions

at  $\sqrt{s_{NN}} = 2.76$  TeV measured by ALICE [45] together with the STAR measurement for  $D^0$  mesons in central Au-Au collisions at  $\sqrt{s_{NN}} = 200$  GeV [46]. For  $p_T > 2 - 3$  GeV/c the two measurements are compatible within uncertainties and show a large suppression. On the contrary at low transverse momentum the STAR measurement shows a bump-like structure, which is less evident at LHC energies. This behaviour could be explained by the combination of radial flow and hadronisation via recombination described in the previous paragraphs. Note, also, that the initial-state effects presented in section 2.2.1, and in particular shadowing, are different at the two collision energies. In the right panel of the same figure the D-meson  $R_{AA}$  measured by ALICE for  $8 < p_T < 16$  GeV/c as a function of centrality expressed in terms of number of participants is compared with the nuclear modification factor of non-prompt  $J/\psi$  with  $p_T > 6.5$  GeV/c measured at the same centre-of-mass energy by CMS [44]. These results indicate a stronger suppression of charm hadrons than beauty hadrons in central Pb-Pb collisions, as expected from the predicted quark-mass dependence of energy loss discussed in the previous section.

# 3

## A Large Ion Collider Experiment

### 3.1 The Large Hadron Collider

The Large Hadron Collider (LHC) is a two-ring superconducting hadron accelerator and collider based at the European particle physics laboratory CERN [47], near Geneva in Switzerland. The LHC is the world largest circular particle accelerator with a 26.7 kilometer circumference housed in a 3.8-meter-wide tunnel buried 50 to 175 meters underground. It was constructed to accelerate mainly protons and Pb ions, which have to be pre-accelerated by smaller accelerators before to be injected in the LHC. In figure 3.1 it is possible to see a schematical representation of the full accelerator complex at CERN.

Protons, obtained from hydrogen atoms, begin the chain in the Linear Accelerator 2 (LINAC2), where they reach an energy of 50 MeV. Subsequently, they are injected in the Proton Synchrotron Booster (PS Booster), which accelerate them to 1.4 GeV, then to the Proton Synchrotron, where they are accelerated up to 25 GeV. Finally they are sent to the Super Proton Synchrotron (SPS), where they are accelerated to 450 GeV, and then they are ready to be injected into the LHC, where they reach their final energy. The LHC proton collision design luminosity is  $\mathcal{L} = 10^{34} \text{ cm}^{-2} \text{ s}^{-1}$  at a centre-of-mass energy of  $\sqrt{s} = 14 \text{ TeV}$ . The values of the centre-of-mass energy provided until now are 5.02, 7, 8 and 14 TeV. Pb ions follow a different path. First electrons are stripped from lead into  $^{208}_{82}\text{Pb}^{29+}$  ions and accelerated in the Linear Accelerator 3 (LINAC3) to an energy of 4.2 MeV per

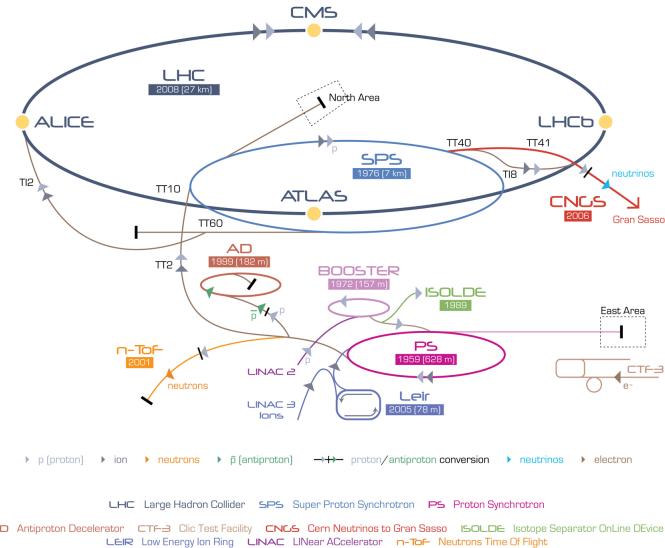


Figure 3.1. Accelerator complex at CERN.

nucleon. Subsequently, Pb ions are further ionised (up to 54+) passing through a carbon foil. Then, they are accelerated into the Low Energy Ion Ring (LEIR), where they reach the energy of 72 MeV per nucleon, and then they enter into the PS which accelerate them up to 5.9 GeV. The final step before injecting the ions into the LHC is again the SPS, where the ions reach 177 GeV and get fully ionised with a second foil. The Pb-Pb collision design luminosity is  $\mathcal{L} = 10^{27} \text{ cm}^{-2} \text{ s}^{-1}$  at a centre-of-mass energy per nucleon of  $\sqrt{s_{\text{NN}}} = 5.52 \text{ TeV}$ , while the energies studied until now are 2.76 and 5.02 TeV. The LHC provides the beams for the four largest experiments at CERN, which are placed on four of the eight beam intersections, as illustrated in figure 3.1. They are:

- A Large Ion Collider Experiment (ALICE): A detector designed to study the physics of strongly interacting matter at extreme energy densities, such as those reached in Pb-Pb collisions, where the formation of a new phase of matter, the Quark-Gluon Plasma (chapter 1), is expected. ALICE is also studying proton-proton and proton-Pb collisions as a comparison with Pb-Pb collisions.
- Compact Muon Solenoid (CMS): multi-purpose particle detector studying a wide range of physics, including the Higgs boson and the search for supersymmetry and dark matter.

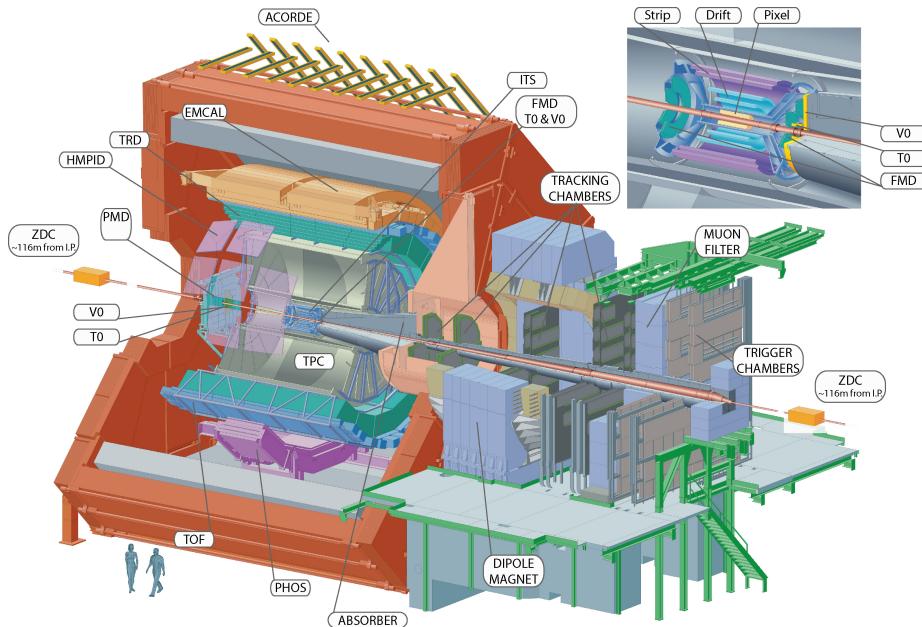


Figure 3.2. The ALICE schematic layout.

- A Toroidal LHC Apparatus (ATLAS): multi-purpose experiment as CMS, but with complementary detection techniques.
- LHC beauty (LHCb): experiment designed to investigate the physics of beauty quark, especially the CP violation in the B-meson sector, and to study the matter-antimatter asymmetry.

## 3.2 The ALICE detector

ALICE is a general-purpose detector for the study of ultra-relativistic heavy-ion collisions [48]. It is designed to address the physics of strongly interacting matter and the QGP at extreme values of energy density and temperature. The physics programme includes not only Pb-Pb collisions, but also p-Pb collisions, in order to study Cold Nuclear Matter effects, and proton-proton collisions to provide a reference data for the heavy-ion programme. The ALICE detector can reconstruct charged particles down to very low momenta (about 100 MeV/c for pions) in an environment characterised by a large charged particle multiplicity. In addition, a

key feature of ALICE is the possibility to identify particles in a wide momentum range, using the information provided by different particle identification detectors. Before to describe the ALICE apparatus, it is necessary to introduce the coordinate system used in the ALICE experiment, which is a right-handed Cartesian system. The  $z$  axis identifies the beam direction, pointing to the ATLAS experiment (see figure 3.1), while the orthogonal plane is defined by the  $x$  and  $y$  coordinates. In particular the  $x$  axis is aligned with the horizontal and points toward the accelerator centre, while the  $y$  axis is oriented upward. Furthermore, the polar angle  $\theta$  is defined with respect to the  $z$  direction and  $\phi$ , the azimuthal angle, increases counter-clockwise starting from the  $x$  axis towards the CMS side. The ALICE detector (figure 3.2) is 26 m long, 16 m high, 16 m wide and weights more than 10000 tons. It consists of a central barrel part, which reconstructs hadrons, electrons and photons, and a forward muon spectrometer. The central part covers polar angles from  $45^\circ$  to  $135^\circ$  (which in terms of pseudorapidity means  $|\eta| < 0.9$ ). It is embedded in a large solenoid magnet reused from the L3 experiment at LEP, which generates a magnetic field of  $B = 0.5$  T parallel to the beam direction. From the inside out, the central barrel composed of the Inner Tracking System (ITS), which is made of two silicon pixel detector layers (SPD), two silicon drift detector layers (SDD) and two silicon strip detector layers (SSD), a large cylindrical Time Projection Chamber (TPC), the Transition Radiation Detector (TRD), the Time Of Flight (TOF) detector, the Ring Imaging Cherenkov (HMPID) detector, and two electromagnetic calorimeters (PHOS and EMCAL). All these detectors cover the full azimuth, except for PHOS, HMPID and EMCAL.

The forward muon arm covers polar angles from  $2^\circ$  to  $9^\circ$  (which in terms of pseudorapidity means  $-4 < \eta < -2.5$ ) and the full azimuth. It consists of a complex arrangement of absorbers, a large dipole magnet and fourteen planes of tracking and triggering Resistive Plate Chambers (RPC).

ALICE includes also several smaller detectors for global event characterisation and triggering (ZDC, PMD, FMD, T0, V0A and V0C) and an array of scintillators (ACORDE) located on top of L3 magnet used to trigger on cosmic rays. In the next sections the more detailed description of the main sub-detectors used in the analysis presented in this thesis is provided.

### 3.2.1 Inner Tracking System (ITS)

The ITS [49] is the first sub-detector reached by the particles originating from the collisions. The main tasks of this detector are:

- The localisation of the primary vertex with a resolution better than  $100\ \mu\text{m}$ .
- The reconstruction of the secondary vertices from the decays of D and B mesons and hyperons.

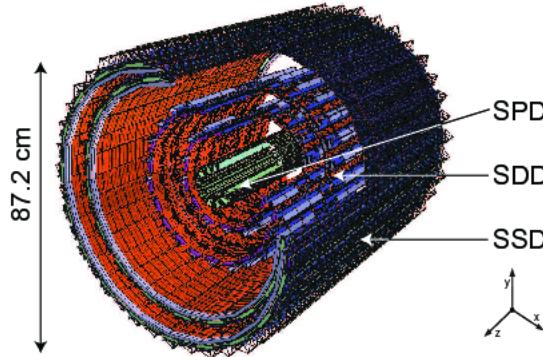


Figure 3.3. Layout of the Inner Tracking System.

- The tracking and identification of the particles with momentum below 200 MeV/c.
- The improvement of the momentum and the angle resolution for particles reconstructed by the TPC.

Because of the high particle density in heavy-ion collisions, and in order to achieve the required impact-parameter (distance of closest approach between the tracks and the primary vertex) resolution, Silicon Pixel Detectors (SPD), which has a spatial resolution of about 12 (100)  $\mu\text{m}$  in the  $r\varphi$  ( $z$ ) direction, have been chosen for the innermost two layers, Silicon Drift Detectors (SDD) for the following two layers and Silicon Strip Detectors (SSD) for the last two layers. The four outer layers have analogue readout and therefore can be used for Particle IDentification (PID) via  $dE/dx$  measurement. The dimensions of each sub-detector are summarised in table 3.1. Since the momentum and impact-parameter resolutions for low-momentum particles are dominated by multiple scattering effects in the detector material, the amount of material in the active volume has been minimised, resulting to an amount of single detector effective thickness of about 0.4% of  $X_0$ . The total material encountered, including the support structures and the shields, by a perpendicular track crossing the ITS is around 7.2% of  $X_0$ .

### Silicon Pixel Detectors (SPD)

The Silicon Pixel Detector is a fundamental element for the determination of the position of the primary vertex as well as for the measurement of the impact parameter of secondary tracks originating from the weak decays of strange, charm

Layer	Sub-detector	r (cm)	z (cm)	Active Area (m <sup>2</sup> )	Channels
1	SPD	3.9	14.1	0.07	3276800
2	SPD	7.6	14.1	0.14	6553000
3	SDD	15.0	22.2	0.042	43008
4	SDD	23.9	29.7	0.89	90112
5	SSD	38.0	43.1	2.20	1148928
6	SSD	43.0	48.9	2.80	1459200

Table 3.1. Dimensions of the Inner Tracking System.

and beauty particles. Thanks to its fast response, it can also provide a fast trigger signal. The SPD is based on hybrid silicon pixels, consisting of a two-dimensional matrix (sensor ladder) of reverse-biased silicon detector diodes bump-bonded to the readout chips. The readout is binary, which means that the digital output changes when the signal is above a set threshold. The sensor matrix includes 256 x 160 cells, corresponding to a sensor area of 12.8 mm ( $r\varphi$ ) x 70.7 mm ( $z$ ). The thickness of the sensor is 200  $\mu\text{m}$  which is the smallest that can be achieved with an affordable yield in standard processes. In total the SPD has 240 ladders and 1200 readout chips, for a total of  $9.8 \cdot 10^6$  cells. It has an intrinsic spatial resolution of about 12  $\mu\text{m}$  in  $r\varphi$  direction and 100  $\mu\text{m}$  in  $z$  direction, and a two-track resolution of about 100  $\mu\text{m}$  in  $r\varphi$  direction and 850  $\mu\text{m}$  in  $z$  direction.

### Silicon Drift Detectors (SDD)

The ALICE Silicon Drift Detectors have very good multi-track capability and provide two out of the four  $dE/dx$  samples needed for the ITS particle identification. The sensors were produced from a very homogeneous high-resistivity (3 k $\Omega\text{cm}$ ) Neutron Transmutation Doped (NTD) silicon. They have a thickness of 300  $\mu\text{m}$  and a sensitive area of 70.17 mm ( $r\varphi$ ) x 75.26 mm ( $z$ ), which is split into two drift regions by the central cathode strip to which a bias of 1.8 kV is applied. In each drift region on both detector surface there are 291 cathode strips, with 120  $\mu\text{m}$  pitch, which completely deplete the detector volume and generate a drift field of about 500 V/m parallel to the surface. Each drift region has 256 anodes with 294  $\mu\text{m}$  pitch, where the electrons generated by the crossing charged particles are collected. The layout of a sensor is shown in figure 3.4. The  $r\varphi$  coordinate of the crossing particle is obtained from the drift time of the electrons with respect to the trigger time. The drift speed depends on temperature ( $v_{drift} \propto T^{-2.4}$ ), therefore for each drift region there are 33 MOS charge injectors to monitor the drift velocity during data acquisition.

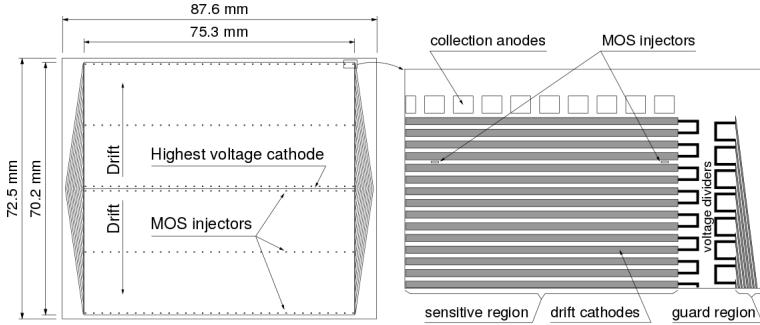


Figure 3.4. Layout of a SDD sensor.

The space precision of the SDD is better than  $38 \mu\text{m}$  in  $r\varphi$  direction and  $25 \mu\text{m}$  in  $z$  direction (it grow up to  $60 \mu\text{m}$  close to the anodes). The double-track resolution is estimated to be around  $200 \mu\text{m}$  for  $r\varphi$  direction and  $600 \mu\text{m}$  for  $z$  direction.

### Silicon Strip Detectors (SSD)

The Silicon Strip Detector is crucial for the prolongation of tracks from the TPC to the ITS. SSDs also provide  $dE/dx$  information to assist PID for low-momentum particles. The basic SSD module is a double-sided strip detector. The sensors are  $300 \mu\text{m}$  thick and have 768 strips on each side with  $95 \mu\text{m}$  pitch. The stereo angle between the strips of the two sides is 35 mrad, which is a compromise between the best resolution on  $r\varphi$  direction and the minimisation of ambiguous measurement resulting from high particle densities. The spatial resolution is about  $20 \mu\text{m}$  in  $r\varphi$  direction and  $800 \mu\text{m}$  in  $z$  direction.

### 3.2.2 Time Projection Chamber (TPC)

The Time Projection Chamber [50] is the main tracking detector of the central barrel and it is optimised to provide charged particle momentum measurements with good two-track separation and particle identification via  $dE/dx$ . In addition data from central barrel can be used to generate a fast online High-Level Trigger (HLT) for the selection of low cross section signals.

The TPC, as shown in figure 3.5, is cylindrical; it covers the full azimuthal angle, having an active radial range from about 85 cm to 250 cm and overall length along the beam direction of 500 cm. It covers the pseudorapidity range  $|\eta| < 0.9$  for tracks with full radial track length (matches in ITS, TRD and TOF),

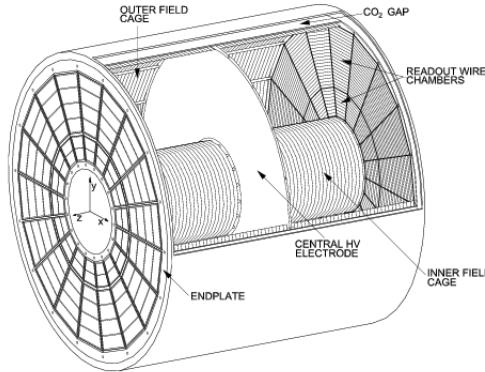


Figure 3.5. Layout of the ALICE Time Projection Chamber.

and up to  $|\eta| = 1.5$  for reduced track length. The TPC also allows charged particle reconstruction in a wide range of transverse momentum, from low  $p_T$  of about 0.1 GeV/c up to 100 GeV/c. Each end plate is divided in 18 trapezoidal sectors, where multi-wire proportional chambers are mounted. The field cage is based on a design with a central high-voltage electrode and two opposite axial resistive potential dividers, which create a highly uniform electrostatic field in the common gas volume. The choice of the gas is strongly dependent on the requirements imposed by the extreme environment of Pb-Pb collisions. In fact the ALICE TPC has been designed for a maximum multiplicity  $dN_{ch}/dy = 8000$ , resulting in 20000 charged primary and secondary tracks in the acceptance, an unprecedented track density for a TPC. Therefore, the drift gas Ne/CO<sub>2</sub>/N<sub>2</sub> was initially chosen and used during Run I data taking (2009–2013). This gas mixture is optimised for drift speed, low diffusion, low radiation length and hence low multiple scattering, small space-charge effect and ageing and stability effects. The N<sub>2</sub> mixture improves the quenching and allows higher maximum gas gains.

Thanks to this design, the detector provides more than 90% efficiency in track finding, with a resolution up to about 2% in charge particle momentum measurement at  $p_T \approx 10$  GeV/c. The relative  $dE/dx$  resolution was measured to be 6% for particle crossing the entire detector.

### 3.2.3 Time of Flight (TOF)

The Time of Flight detector [51] is a large area detector arranged around the TRD, it covers the same pseudorapidity range of ITS, TPC and TRD ( $|\eta| < 0.9$ ). Its main purpose is the particle identification in the intermediate momentum range,

up to 2.5 GeV/c for pion-kaon separation and up to 4 GeV/c for proton-kaon separation.

The TOF is made of Multi-gap Resistive Plate Chambers (MRPC) which cover an area of about  $160 \text{ m}^2$  filled with a gas mixture composed of 90% of  $\text{C}_2\text{H}_2\text{F}_4$ , 5% of  $\text{C}_4\text{H}_1\text{O}$  and 5% of  $\text{SF}_6$ . Each multi-gap is made of 10 layer double-stack strip detectors with a time resolution of about 40 ps. The time of flight for each particle is given together with the information of the start time of the event,  $t_0$  ( $t_{\text{flight}} = t_{\text{hit}} - t_0$ ). The start time  $t_0$  is determined with the T0 detector or using the particle arrival times at the TOF. The overall TOF resolution is around 85 ps in Pb-Pb events and 100 ps in proton-proton collisions, where there is a larger uncertainty on the determination of the event time zero.

## 3.3 ALICE trigger and data acquisition

### 3.3.1 Central Trigger Processor

The ALICE Central Trigger Processor (CPT) [52] collects and manages all the trigger inputs coming from all the detectors and provides a trigger signal to the readout detectors in case of fulfilled trigger conditions. The main features of the CPT are the capability to combine many detectors, which are busy for widely different periods, and to perform optimised trigger selections for different running modes (Pb-Pb, p-Pb and pp), varying by almost two orders of magnitude in counting rate. The first level trigger signal, called L0 and made of 24 inputs, arrives 1.2  $\mu\text{s}$  after the collision. The signals from the detectors which are not fast enough for L0 trigger are collected in the level 1 trigger signal (24 L1 inputs), that is sent after 6.5  $\mu\text{s}$ . Finally a third level of trigger (12 L2 inputs) is sent after 88  $\mu\text{s}$ , which is the maximum drift time of electrons in the TPC, in order to guarantee a past-future protection that looks for other events of requested type in the time window around the collision under investigation. This feature helps the rejection of the pile-up events in the TPC.

The *minimum-bias* (MB) trigger which will be used in the analysis presented in this thesis is the one used in p-Pb collisions, which requires the signal from both V0A and V0C detectors (see figure 3.2).

### 3.3.2 Data AcQuisition System and High-Level Trigger

The Data AcQuisition (DAQ) system of ALICE must operate in really challenging conditions, starting from the high interaction in rate proton-proton collisions to the huge amount of data collected in Pb-Pb collisions (up to 4 Gb/s).

If all the trigger requirements are fulfilled, the data are sent to the Detector Data

Link (DDL) and then to the Local Data Concentrators, a farm of computers which builds sub-events from the event fragments they receive from the front-end electronics. Finally, the complete event is built by the Global Data Collectors (GDC), which send it to the storage facilities if it is accepted by the High Level Trigger (HLT). The HLT can accept or reject an event after performing an online analysis, which includes a fast track and primary vertex reconstruction. Then, if the event is accepted, the High Level Trigger compress the data and transfer them to the CERN computer centre, where they are registered on tape.

## 3.4 ALICE performance

### 3.4.1 Track and vertex reconstruction in the central barrel

In ALICE the tracking procedure starts with the conversion of the detector hits in *clusters*, characterised by several quantities, such as position, signal amplitude and signal time, and their associated errors. Then, a first evaluation of the primary vertex is performed using the clusters in the two innermost layers of the ITS (SPD). It is determined as a space point to which a maximum number of tracklets (lines defined by pairs of clusters, one cluster in each SPD layer) converge (*Vertexer3D*). In low multiplicity events, if it is not possible to find a single point of convergence, the algorithm performs a one-dimensional minimisation to reconstruct the position of the interaction point along the beam direction ( $z$  axis) by using the distances of closest approach of the tracklets to the average position of the beam in the transverse plane (*VertexerZ*). After this procedure the tracks are reconstructed and fitted in TPC and ITS using the Kalman Filter algorithm [53] and subsequently prolonged to the other central barrel detectors and fitted again. The procedure can be summarised in the following steps:

- The track seeds are build first from the clusters in the two outermost pad rows of the TPC and the primary vertex. Then, with three clusters and without the vertex constraint. Then, the seeds are propagated inward and updated with every nearest cluster found by the Kalman Filter algorithm, until the inner radius of the TPC is reached.
- The tracks reconstructed in the TPC are matched to the hits in the outermost SSD layer and become the seed for the track finding in the ITS. The seed is then propagated inward and updated at each ITS layer attached to all the clusters within a proximity cut. Therefore, for each TPC track, a corresponding tree of track hypotheses in the ITS is produced. The candidates are then selected according to their  $\chi^2$ . In the left panel of figure 3.6 is shown the comparison between the fraction of tracks in TPC which have a prolongation in the ITS (*matching efficiency*) in data and in the simulation

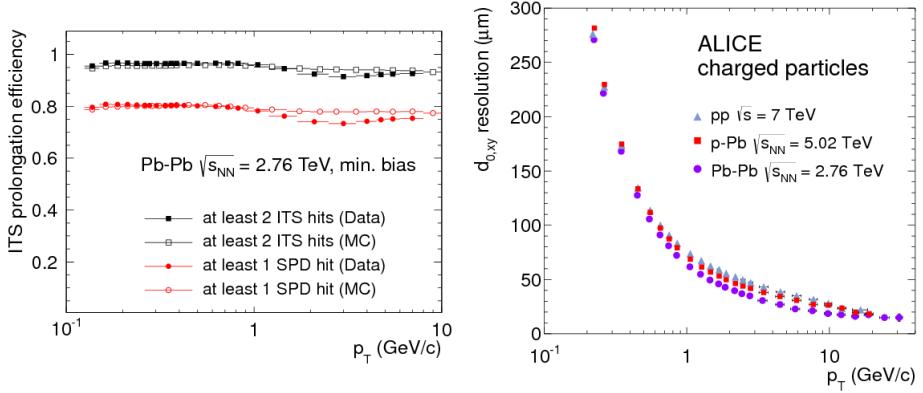


Figure 3.6. Left: comparison between the TPC-ITS matching efficiency, defined as the fraction of TPC tracks with a prolongation in the ITS, in data and in the simulation for Pb-Pb collisions at  $\sqrt{s_{NN}} = 2.76$  TeV, for different requirements on ITS points. Right: impact-parameter resolution in the  $xy$ -plane for primary particles in pp, p-Pb and Pb-Pb collisions.

for Pb-Pb collisions at  $\sqrt{s_{NN}} = 2.76$  TeV, for different requirements on ITS points [54].

- The hits in the ITS not attached to tracks propagated from the TPC are used to perform an ITS stand-alone reconstruction, which allows to obtain a better efficiency for tracks with low transverse momentum and to reconstruct tracks of particles traversing dead zones of the TPC or which decay before entering in the TPC.
- The tracks obtained are then back-propagated in the outward direction in order to match the clusters in the TRD and TOF, and the signals in EM-Cal, PHOS and HMPID. Finally, the tracks are re-fitted inward using the previously found clusters. The parameters of the track state vector and its covariance matrix are calculated.

After this procedure, if it is possible, the primary vertex is re-evaluated using the reconstructed tracks. The determination of the primary vertex with the tracks instead of the SPD tracklets allows to improve significantly the resolution, which is fundamental for analysis of charm and beauty hadrons, whose decay vertices are typically displaced by hundreds of microns from the primary vertex.

The right panel of figure 3.6 shows the resolution of the distance of closest approach to the primary vertex (impact parameter) in the transverse plane for charged tracks in pp, p-Pb and Pb-Pb collisions [54]. From this picture it is possible to notice

that the resolution is slightly better in p-Pb and even more in Pb-Pb collisions as compared to pp interactions, because the reconstruction on the primary vertex position improves with increasing multiplicity of charged particles produced in the collision.

The reconstructed data are stored in files containing all the physical information needed for the analysis and for checking the quality of the reconstruction for each sub-detector in a format called ESD (*Event Summary Data*). In order to reduce the size of the files and the computational time needed for the analysis, the AOD (*Analysis Object Data*) files, which contains only the information needed for the analysis, are produced. The conversion from ESD to AOD is called *filtering*. In addition, the AODs also contain some initial step of the analysis. For example, for the D-meson analysis, the AODs already contain the D-meson candidates, which are combinations of tracks coming from a secondary vertex, selected with loose criteria.

### 3.4.2 Particle IDentification (PID)

One of the distinctive features of the ALICE detector is the excellent particle identification capability, which can be performed in a wide range of momentum, exploiting different techniques. In this section I only discuss the subsystem for identifying charged hadrons used in the analysis described in this thesis, which are the Time Projection Chamber and the Time of Flight.

#### PID with the Time Projection Chamber

The PID with the Time Projection Chamber is performed by simultaneously measuring the specific energy loss (via the measurement of the charge deposited on up to 159 pad rows), the charge and the momentum of each particle traversing the detector gas. A truncated mean of the  $dE/dx$  is calculated using the 65% of the lowest-amplitude clusters. The particle identification is based on the comparison of the measured  $dE/dx$  and the expected  $dE/dx$  for a specific particle species with a certain momentum  $p$ . The relative  $dE/dx$  resolution achieved is about of the 6% for particle crossing the entire detector. In the left panel of figure 3.7 is shown the specific energy loss measured with the TPC in p-Pb collisions as a function of the track momentum, with the Bethe-Bloch curves for electrons, pions, kaons, protons, deuterium and tritium superimposed. The TPC provides a better particle identification for  $p_T \lesssim 0.7$  GeV/c, where the largest separation is achieved (see figure 3.7), however in the relativistic rise region ( $2 \lesssim p_T \lesssim 20$  GeV/c) a statistical approach allows to extract the relative abundances.

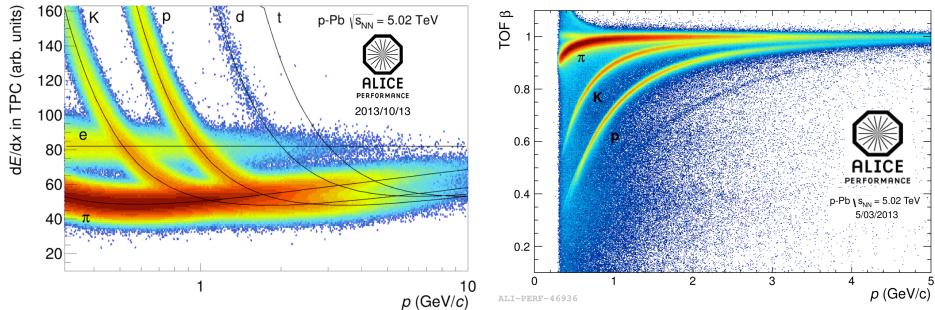


Figure 3.7. Left: specific energy loss ( $dE/dx$ ) in the TPC as a function of the particle momentum in p-Pb collisions at  $\sqrt{s_{NN}} = 5.02$  TeV. The lines show the parametrisation of the expected mean energy loss. Right: velocity  $\beta$  as measured with the TOF detector as a function of momentum for particles reaching the TOF in p-Pb collisions at  $\sqrt{s_{NN}} = 5.02$  TeV.

### PID with the Time of Flight

The Time of Flight is a dedicated PID detector, which measures the arrival time of particles. In order to obtain the flight time of the particles the start time for the TOF detector can be provided either by the T0 detector, which consists of two arrays of Cherenkov counters (T0A and T0C), positioned at opposite sides of the interaction point (see figure 3.2) or via a combinatorial analysis of the particle arrival times at the TOF detector. The overall TOF resolution, including the uncertainty on the start time of the event, and the tracking and momentum resolution contributions, is up to  $\sim 85$  ps for high-multiplicity p-Pb collisions [54]. As shown in the right panel of figure 3.7, where the particle velocity is plotted as a function of momentum, the TOF detector provides a good proton-kaon separation up to  $p_T \simeq 4$  GeV, while for larger values it is not possible to achieve a satisfactory separation among the hadron species.



# 4

## Prompt D<sup>+</sup>-meson reconstruction strategy

This Chapter is dedicated to the description of the strategy for the reconstruction of the D<sup>+</sup>-mesons in p-Pb collisions at  $\sqrt{s_{NN}} = 5.02$  TeV, which is adopted for both the standard analysis published by the ALICE Collaboration [36] and the analyses presented in this thesis. The D<sup>+</sup>-mesons are reconstructed in the hadronic decay channel  $D^+ \rightarrow K^-\pi^+\pi^+$ , which has a branching ratio of 9.13%, as reported in table 4.1. D<sup>+</sup>-mesons can be *prompt*, if they come directly from the hadronisation of a *c*-quark or from the decay of a excited open charm or charmonium state, or *feed-down*, if they originate from the decay of hadrons containing *b*-quarks. In the first section of this Chapter both the contributions are considered as *signal* against the *background* due to combinatorial background, while starting from the second section the two contributions to the D<sup>+</sup> measured yield are treated separately. Most

Meson	Mass (MeV/c <sup>2</sup> )	decay channel	BR(%)	<i>cτ</i> (μm)
D <sup>+</sup>	$1869.61 \pm 0.09$	$K^-\pi^+\pi^+$	$9.13 \pm 0.19$	311.8
D <sup>0</sup>	$1865.84 \pm 0.05$	$K^-\pi^+$	$3.88 \pm 0.05$	122.9
D <sup>*+</sup>	$2010.27 \pm 0.05$	$D^0\pi^+ \rightarrow K^-\pi^+\pi^+$	$67.7 \pm 0.5$	strong decay
D <sub>s</sub> <sup>+</sup>	$1969.30 \pm 0.10$	$\phi\pi^+ \rightarrow K^+K^-\pi^+$	$2.24 \pm 0.10$	149.9

Table 4.1. D-mesons hadronic decays measured by the ALICE Collaboration. The Characteristics reported are from [56].

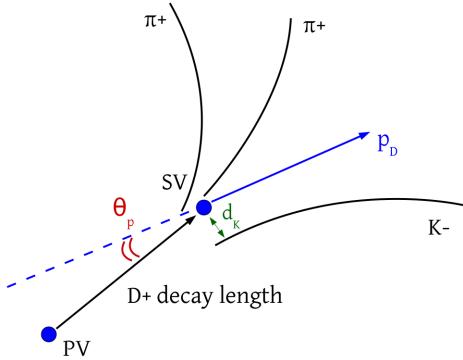


Figure 4.1. Sketch of the  $D^+$  hadronic decay  $D^+ \rightarrow K^- \pi^+ \pi^+$  and some of the reconstructed topological quantities associated.

of the procedures described in the following sections were already implemented in the analysis framework of ALICE, however for the analysis presented in Chapter 7 I re-implemented them. This was done because of the different file format needed for the analysis: while the studies presented in Chapter 5 and 6 were performed on the AODs (see section 3.4.1), which already contain the information of the  $D^+$  candidates, for the analysis in Chapter 7 the ESDs were needed because of the different vertexing package used.

## 4.1 $D^+$ candidate reconstruction and selection

$D^+$ -mesons (and their charge conjugates) are reconstructed in the hadronic decay channel  $D^+ \rightarrow K^- \pi^+ \pi^+$ , using an analysis based on the selection of fully reconstructed decay topologies displaced from the primary interaction vertex. This is possible thanks to the relatively high branching ratio ( $BR = 9.13\%$ ) of this decay and the mean proper decay length of the  $D^+$  ( $c\tau = 312 \mu\text{m}$ ), which allows one to resolve the decay vertex from the interaction vertex. The main  $D^+ \rightarrow K^- \pi^+ \pi^+$  characteristics are summarised in table 4.1, together with the other  $D$ -mesons measured by ALICE for comparison. Even if the branching ratios were recently updated by the PDG (see [57]), I reported the values from [56], since the analyses which I will present are performed using these values in order to be able to compare the results with the published ones.

### 4.1.1 p-Pb data sample and event selection

The analysis is performed on a data sample of p-Pb collisions at  $\sqrt{s_{NN}} = 5.02$  TeV collected in 2013 with the ALICE detector. The events were selected online using a Minimum Bias trigger which requires signals in both V0 detectors, as explained in section 3.3.1. Furthermore, only the events which pass a selection, called *physics selection*, based on the timing information of V0 and ZDC detectors are considered for the analysis. This selection removes background events due to the interaction of the beams with the residual gas molecules present in the vacuum tube.

As discussed in section 3.4.1, the interaction vertex can be reconstructed using global (ITS-TPC) tracks or only SPD tracklets. Considering that the  $D^+$  decay length is of the order of hundred of microns, it is crucial to have a primary vertex reconstructed as precisely as possible. Therefore, only the vertices reconstructed with global tracks are kept into account. In order to have a uniform acceptance for tracks in the ITS and TPC detectors, all the events with the reconstructed primary vertex position along the beam direction ( $z_{vtx}$ ) more than 10 cm away from the nominal interaction point position are rejected. Finally, the events with at least two reconstructed interaction vertices having a minimum distance of 0.8 cm between each other and the lower multiplicity vertex with minimum 5 contributors, are identified as pile-up events and rejected. This selection rejects less than 1% of the events passing the physics selections, as expected for the interaction rate of the Minimum Bias p-Pb data taking, which was of about 10 kHz. After all these selections the number of events available results to be of about 100 million, which corresponds to an integrated luminosity of about  $L_{int} = N_{MB}/\sigma_{MB}^{pPb} = 48.6 \mu b^{-1}$ , where  $\sigma_{MB}^{pPb}$  is the Minimum Bias cross section measured with van der Meer scans [58] ( $\sigma_{MB}^{pPb} = 2.09 \pm 3.5\%$  b).

### 4.1.2 Track selection

The  $D^+$ -meson candidates are built by combining three ITS-TPC tracks which fulfil several kinematical and quality requirements. First of all, only the tracks which have been successfully fitted by the Kalman filter algorithm during the final inward refit step described in section 3.4.1 are selected. Then tracks are required to have at least 70 associated space points in the TPC (out of a maximum of 159) and 2 hits in the ITS (out of six possible), of which at least one has to be in one of the two SPD layers. This latter requirement is really important for the heavy-flavour measurements, since it provides optimal impact-parameter and secondary-vertex resolution, thanks to the presence of a measured point close to the primary and secondary vertices. Furthermore the tracks are selected by requiring a fit quality of  $\chi^2/ndf < 2$  in the TPC. Finally only the tracks with  $|\eta| < 0.8$  and  $p_T > 0.3$  GeV/c are taken into account for the analysis.

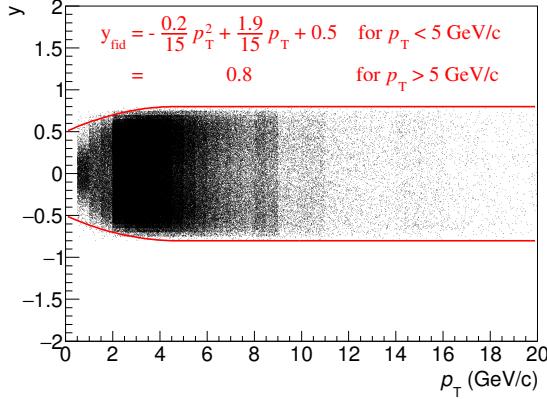


Figure 4.2. Rapidity vs transverse momentum distribution of the reconstructed  $D^+$ -mesons. The fiducial acceptance region is defined as  $|y_{D+}| < y_{fid}(p_T)$ .

#### 4.1.3 $D^+$ reconstruction in the fiducial acceptance region

The  $D^\pm$ -meson candidates are reconstructed combining triplets of good tracks with the proper charge sign (two positive tracks and one negative track for the  $D^+$  and the opposite for the  $D^-$ ). For each candidate the momentum and the invariant mass are calculated starting from the energy and momentum of the measured tracks

$$\begin{cases} \vec{p} = (p_K^x + p_{\pi_1}^x + p_{\pi_2}^x, p_K^y + p_{\pi_1}^y + p_{\pi_2}^y, p_K^z + p_{\pi_1}^z + p_{\pi_2}^z) \\ M_{K\pi\pi} = \sqrt{(E_K + E_{\pi_1} + E_{\pi_2})^2 - (\vec{p}_K + \vec{p}_{\pi_1} + \vec{p}_{\pi_2})^2} \end{cases} \quad (4.1)$$

where

$$E_{K,\pi} = \sqrt{m_{K,\pi}^2 + \vec{p}_{K,\pi}^2}. \quad (4.2)$$

The mass of the kaon  $m_K$  is assigned to the track with the opposite charge sign with respect to the  $D^+$  candidate and the mass of the pion  $m_\pi$  to the other two tracks. For each candidate the associated rapidity is calculated using the reconstructed quantities as

$$y_{D^+} = \ln\left(\frac{E + p_z}{E - p_z}\right), \quad (4.3)$$

where  $E$  is the energy of the candidate computed using its momentum component and the  $D^+$  mass from the PDG. Only the  $D^+$  mesons which are in a fiducial acceptance region are considered for the analysis. This is needed to exclude the  $y$  regions

where the acceptance steeply decreases due to the single track selections ( $|\eta| < 0.8$  and  $p_T > 0.3$  GeV/c). The fiducial acceptance region is  $p_T$ -dependent, since  $D^+$ -mesons with low transverse momentum have a lower acceptance, because of the larger dispersion of the daughter tracks around the  $D^+$  flight direction. Therefore the rapidity values accepted are those within  $|y_{D^+}| < y_{fid}(p_T)$ , where  $y_{fid}(p_T)$  increases from 0.5 for  $p_T = 0$  up to 0.8 for  $p_T > 5$  GeV/c, as shown in figure 4.2.

#### 4.1.4 Topological selections

The  $D^+$  candidates are dominated by the background because of the large number of track combinations. Therefore it is necessary to apply geometrical selections on the displaced decay vertex topology, in order to improve the signal over background ratio and the statistical significance of the signal. First of all, for this purpose, the decay vertex position is computed performing a minimisation of the distance among the tracks of the selected triplet, weighted according their uncertainties. Then other geometrical and kinematical quantities, which can discriminate the  $D^+$  signal from the combinatorial background, are exploited.

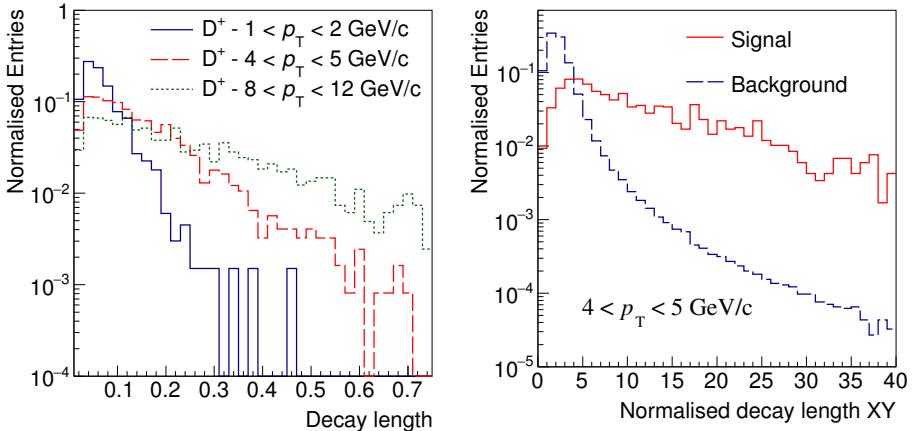


Figure 4.3. Left: reconstructed decay length distributions for  $D^+$  in three different  $p_T$  intervals from Monte Carlo simulations. Right: comparison between the distributions of the normalised decay length in the transverse plane for signal (prompt and feed-down  $D^+$ ) and background.

## Decay length

One of the most effective topological variables which can discriminate between the  $D^+$  signal and the background is the distance between the primary and the secondary vertices, indicated as decay length, since on average it is larger for real  $D^+$  than for the background candidates. Note that the distance between the reconstructed secondary and primary vertices is an approximation of the real decay length because of the curvature of the  $D^+$  trajectory in the magnetic field, however it leads only to a negligible effect considering that the mean proper decay length of the  $D^+$  is of the order of few hundred microns. The decay length distribution of real  $D^+$  depends on the  $D^+$  momentum because of the Lorentz boost, as illustrated in the left plot of figure 4.3, where the decay length distributions of  $D^+$  in three different transverse momentum interval are compared from Monte Carlo simulations. The decay length is also evaluated in the transverse plane, since the detector resolution is better in the  $x$  and  $y$  coordinates with respect to the  $z$  coordinate, as discussed in section 3.2.1. In particular, the decay length in the  $xy$ -plane divided by its error (*normalised decay length*  $xy$ ) is used in the selection of the  $D^+$  candidates because it can provide a significant separation between signal and background, as shown in the right panel of figure 4.3. In this plot the signal distribution, which includes the prompt and the feed-down contributions, is obtained from the MC

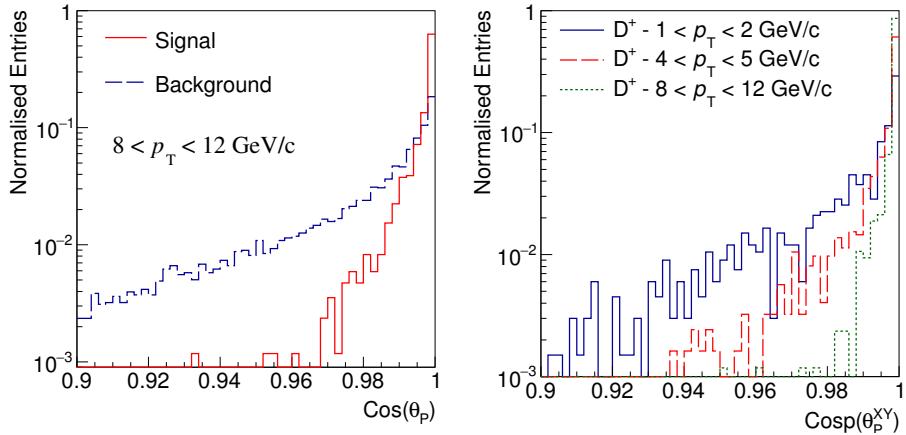


Figure 4.4. Left: comparison of the distribution of the reconstructed cosine of the pointing angle between signal (prompt and feed-down  $D^+$ ) and background. Right: distributions of the reconstructed cosine of the pointing angle in the transverse plane for the  $D^+$  in three different  $p_T$  intervals.

simulation, while the background distribution from the data, using candidates in the invariant-mass region in the side-bands of the  $D^+$  peak.

### Cosine of the pointing angle

The *pointing angle* ( $\theta_P$ ) is defined as the angle between the direction of the reconstructed momentum of the  $D^+$  candidate and the line connecting the primary and secondary vertices, as illustrated in the sketch of the  $D^+$  decay topology of figure 4.1. Neglecting the experimental resolution on the vertices and the momentum, this variable is exactly 0 for real prompt  $D^+$  and close to 0 for real feed-down  $D^+$ . Therefore, the distribution of the reconstructed cosine of the pointing angle is peaked at one for the signal and it is wider for the background. In the left plot of figure 4.4, the signal and background distributions of the reconstructed cosine of the pointing angle are shown for candidates with  $8 < p_T < 12$  GeV/c. As the decay length, the  $\cos(\theta_P)$  is evaluated also in the transverse plane, because of the better resolution achieved. In the right panel of figure 4.4 it is possible to see how the distribution of  $\cos(\theta_P^{xy})$  of signal candidates becomes narrower with increasing  $D^+$  transverse momentum.

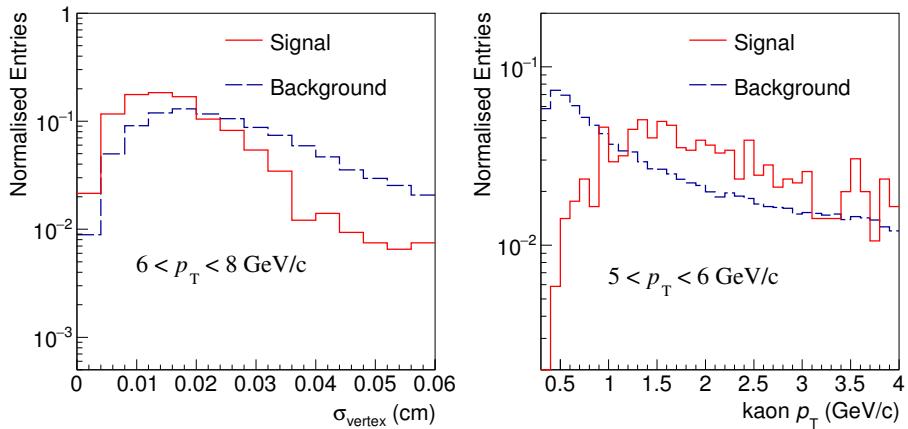


Figure 4.5. Comparison between the distributions of reconstructed  $\sigma_{\text{vertex}}$  (left) and transverse momentum of the kaon (right) of signal (prompt and feed-down  $D^+$ ) and background candidates.

### Daughter-track dispersion around the secondary vertex

Another quantity which can be helpful in the rejection of the background is the so called *sigma vertex*, defined as

$$\sigma_{vertex} = \sqrt{d_K^2 + d_{\pi 1}^2 + d_{\pi 2}^2}, \quad (4.4)$$

where  $d_i$  are the distances of closest approach between the daughter tracks and the reconstructed secondary vertex. For real  $D^+$ -mesons this variable should be exactly 0, however it assumes non-zero values because of the tracking and vertexing resolution. On the contrary, for combinatorial background the three considered tracks do not come from a real  $D^+$  decay vertex, and hence we expect in average a larger displacement and therefore a larger value of the  $\sigma_{vertex}$ , as confirmed by the distributions for signal and background shown in the left panel of figure 4.5.

### Daughter transverse momentum

The transverse momentum of  $D^+$  decay tracks is on average larger than that of background. Therefore, a lower cut on this quantity can help to exclude a significant amount of background, keeping the majority of the signal. An example is shown in the right plot of figure 4.5, where the  $p_T$  distributions of the kaons for signal and background candidates with  $5 < p_T < 6$  GeV/c are compared.

### Topological cut values used in the published result

In table 4.2 I report all the values for the cuts on the described geometrical and kinematical quantities which have been applied for the standard analysis, published by the ALICE Collaboration. The choice of these cut values has been performed

$p_T$ (GeV/c)	[1,2]	[2,8]	[8,9]	[9,10]	[10,11]	[11,12]	[12,14]	[14,16]	[16,24]
$p_T^K$ (GeV/c) >	0.3	0.3	0.3	0.3	0.3	0.3	0.3	0.3	0.3
$p_T^\pi$ (GeV/c) >	0.3	0.35	0.35	0.35	0.35	0.35	0.35	0.35	0.35
Decay length (cm) >	0.02	0.04	0.04	0.04	0.04	0.04	0.10	0.10	0.15
Norm. $L_{xy}$ >	9	8	8	8	8	6	6	9	5
$\cos(\theta_P)$ >	0.990	0.990	0.990	0.990	0.990	0.990	0.990	0.990	0.990
$\cos(\theta_P^{xy})$ >	0.995	0.990	0.990	0.990	0.990	0.990	0.990	0.990	0.990
$\sigma_{vertex}$ (cm) <	0.030	0.030	0.035	0.035	0.035	0.070	0.070	0.090	0.030

Table 4.2. Summary table of the  $D^+$  topological cuts optimised for the published result [36]. The optimisation procedure is explained in [59].

taking into account the stability of the raw yield extraction, the value of the efficiency obtained and the fraction of prompt  $D^+$  selected. Further details about the optimisation procedure for the cut values can be found in [59].

#### 4.1.5 Particle Identification

The rejection of the background can be further improved by applying particle identification (PID) selection on the daughter tracks. As already mentioned in section 3.4.2 the specific energy loss in the TPC and the time of flight measured with the TOF detector have been used for this purpose. For each track of the triplet, the degree of agreement with a particle hypothesis (pion, kaon or proton) is evaluated considering the number of sigmas by which the measure value deviates from the expected one. Depending on the value of  $n\sigma$  in TOF and TPC, the track can be considered *identified*, *compatible* or *rejected* as a particular particle species. The values used to obtain the response for the published result and for the analyses presented in the next Chapters are summarised in table 4.3. The  $D^+$  candidates are then selected according to the following requirements:

- None of the tracks has to be positively identified as a proton and simultaneously rejected as kaon or pion by either TOF or TPC.
- No more than one track has to be identified as a kaon and rejected as pion by either TOF and TPC.
- At most two tracks can be rejected as kaon by either TOF and TPC.
- The tracks with the same charge sign with respect to the  $D^+$  candidate have to be at least compatible with the pion hypothesis for both TPC and TOF.

TPC	$p_T^{trk} < 0.6 \text{ GeV}/c$	$0.6 < p_T^{trk} < 0.8 \text{ GeV}/c$	$p_T^{trk} > 0.8 \text{ GeV}/c$
$n\sigma_{TPC} < 1$	identified	identified	compatible
$1 < n\sigma_{TPC} < 2$	identified	compatible	compatible
$2 < n\sigma_{TPC} < 3$	compatible	compatible	compatible
$n\sigma_{TPC} > 3$	rejected	rejected	rejected

TOF	$p_T^{trk} < 1.5 \text{ GeV}/c$	$p_T^{trk} > 1.5 \text{ GeV}/c$
$n\sigma_{TOF} < 3$	identified	compatible
$n\sigma_{TOF} > 3$	rejected	rejected

Table 4.3. PID response for the TPC and TOF detectors in different  $p_T$  intervals of the tracks.

- The track with opposite charge sign with respect to the  $D^+$  candidate has to be at least compatible with the kaon hypothesis for both TPC and TOF.

If the reconstructed transverse momentum of the candidate is smaller than 2 GeV/c, the last two selections require the tracks to be positively identified instead of just being compatible with the particle hypothesis. This stronger requirement is needed because of the large amount of background, which is widely dominant in this  $p_T$  region.

#### 4.1.6 Raw yield extraction

In figure 4.6 are shown three example of fit performed to the invariant-mass distributions of the  $D^+$ -meson candidates (and their charge conjugates) in three different  $D^+$  transverse momentum intervals, obtained applying the topological cuts in table 4.2 and the particle identification. The signal is parametrised with a Gaussian, while the background with an exponential function. The stability of the fit parameters was checked by comparing the Gaussian mean and width with the values obtained by fitting the MC distributions. As it can be seen in figure 4.6 the Gaussian width increases with increasing  $p_T$ , as expected from the transverse-momentum resolution. Thanks to the topological cuts, the signal-to-background ratio and the

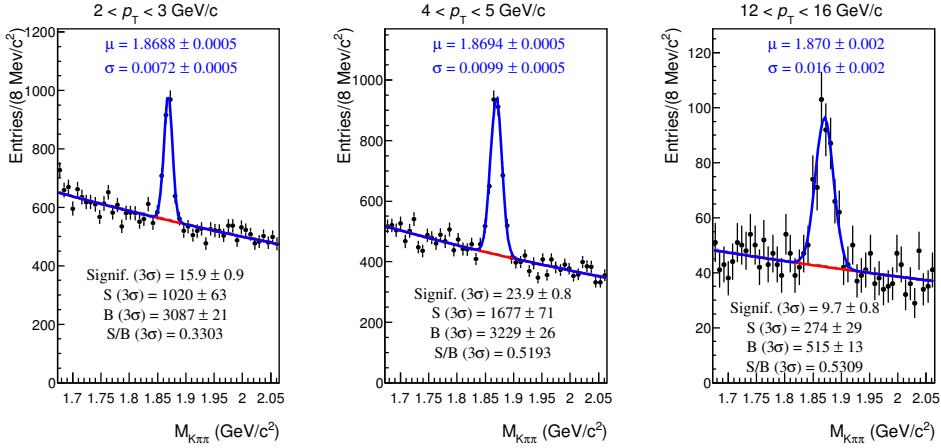


Figure 4.6. Invariant-mass fits in the transverse momentum intervals  $2 < p_T < 3$  GeV/c,  $4 < p_T < 5$  GeV/c and  $12 < p_T < 16$  GeV/c obtained applying the topological cuts reported in table 4.2 and the PID selection.

significance, defined as

$$\text{significance} = \frac{S}{\sqrt{S + B}}, \quad (4.5)$$

are maximised. The best significance is achieved at intermediate values of transverse momentum, while it is lower at low  $p_T$  because of the large combinatorial background and at high  $p_T$  because of the poorer statistics.

## 4.2 Prompt and Feed-down $D^+$

As discussed in section 2.2.2 the interaction with the medium is different for charm and beauty quarks, because of their different masses. Therefore in the  $D^+$ -meson measurement one should separate the *prompt* contribution ( $D^+$  originated directly from the hadronisation of a  $c$ -quark, or from the decay of an excited open charm or charmonium state) from the *feed-down* contribution ( $D^+$  coming from the decay of a B-hadron). This can be done with two different approaches: the first one, which was used for the standard analysis [36], is a *theory-driven* method, which consists on the evaluation of the fraction of  $D^+$  prompt making use of pQCD calculations and of the prompt and feed-down  $D^+$  efficiencies from Monte Carlo Simulations. On the other hand it is also possible to apply *data-driven* methods that exploit the different shapes of the distributions of quantities related to the decay topology, which can discriminate between prompt and feed-down  $D^+$ . In particular the development of data-driven methods for the measurement of the  $D^+$  prompt fraction is the main topic of thesis and will be presented in Chapter 5 and 6.

### 4.2.1 Theory-driven methods for the $D^+$ prompt fraction evaluation

The prompt fraction of  $D^+$ -mesons  $f_{\text{prompt}}$  can be calculated using two different theory-driven methods, both based on FONLL pQCD calculations. The usage of FONLL calculations is supported by the observation (see section 2.1) that they describe the  $p_T$ -differential cross section measured in pp collisions. According to the first approach (called  $N_b$  method), the prompt fraction is defined as

$$f_{\text{prompt}} = 1 - A \cdot \left( \frac{d^2\sigma}{dp_T dy} \right)_{\text{feed-down}}^{\text{FONLL}} \cdot R_{pPb}^{\text{feed-down}}(p_T) \times \\ \times \frac{(Acc \times \epsilon)_{\text{feed-down}} \cdot \Delta y \cdot \Delta p_T \cdot BR \cdot L_{int}}{N_{\text{raw}}^{D^\pm}/2}, \quad (4.6)$$

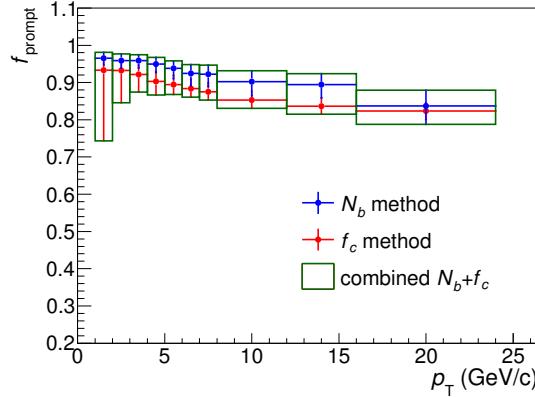


Figure 4.7.  $D^+$  prompt fraction  $f_{prompt}$  evaluated with the theory-driven methods  $f_c$  and  $N_b$  for the set of topological cuts reported in table 4.2.

where  $A$  is the mass number of the Pb nucleus,  $(Acc \times \epsilon)_{feed-down}$  is the product of acceptance and efficiency for feed-down  $D^+$  evaluated with the MC simulation described in the next section,  $\Delta p_T$  and  $\Delta y$  are the widths of the transverse momentum and rapidity intervals,  $BR$  is the branching ratio of the considered decay channel,  $L_{int}$  the integrated luminosity and  $N_{raw}^{D^\pm}$  is the raw yield extracted from the invariant-mass fits described above. The feed-down  $D^+$  cross section  $\left( d^2\sigma/dp_T dy \right)$  is evaluated from the B-meson cross section in pp collisions at  $\sqrt{s} = 5.02$  TeV estimated with FONLL and subsequently using the EvtGen package [60] to describe the  $B \rightarrow D + X$  kinematics. The nuclear modification factor  $R_{pPb}^{feed-down}(p_T)$  is assumed equal to the one for prompt  $D^+$ , on the basis of calculations including initial state effects, since it was never measured.

The alternative method for the computation of  $f_{prompt}$ , called  $f_c$ , needs as input also the FONLL differential cross section, the nuclear modification factor and the acceptance times efficiency for prompt  $D^+$ -mesons:

$$f_{prompt} = \left( 1 + \frac{\frac{d^2\sigma}{dp_T dy}_{feed-down}^{FONLL} \cdot (Acc \times \epsilon)_{feed-down} \cdot R_{pPb}^{feed-down}(p_T)}{\frac{d^2\sigma}{dp_T dy}_{prompt}^{FONLL} \cdot (Acc \times \epsilon)_{prompt} \cdot R_{pPb}^{prompt}(p_T)} \right)^{-1}. \quad (4.7)$$

It is important to notice that the prompt fraction depends on the  $D^+$  transverse momentum because of the different shape of the FONLL  $p_T$ -differential cross section of prompt and feed-down  $D^+$  and the topological selections which lead to a  $p_T$ -dependent efficiency. For the central value of  $f_{prompt}$ , the  $N_b$  method is used, in light of the fact that the measured B-meson cross-section is in good agreement with FONLL, lying close to the central value of the theoretical calculation, while the D-meson one is found to lie on the upper edge of the FONLL uncertainty band (see section 2.1). The  $f_c$  method is kept in order to evaluate a systematic uncertainty, as shown in figure 4.7. The other main sources of systematics of  $f_{prompt}$  come from the perturbative calculation of the cross sections and the assumption on nuclear modification factor for feed-down  $D^+$ , which is varied in the range  $0.9 < R_{pPb}^{feed-down}/R_{pPb}^{prompt} < 1.3$  in order to evaluate an uncertainty.

## 4.3 The Monte Carlo simulation

The Monte Carlo simulation is needed to evaluate the efficiency and acceptance factors used to correct the measured raw yields and it is also used to obtain the distribution templates for the analysis presented in Chapter 5.

The MC sample used is based on pp events generated with Pythia v6.4.21 with the Perugia-0 [61] tune with at least a  $c\bar{c}$  or a  $b\bar{b}$  pair in each event. In order to make the multiplicity distribution in the MC more realistic, a number of binary collisions is extracted from a Glauber MC simulation of p-Pb collisions and, if larger than 1, a HIJING [62] p-Pb event is added as underlying event to the Pythia one. After the particle generation, the GEANT3 transport package is used to simulate the interaction with the detector. The detector response is modelled using AliRoot [63], the software adopted by the ALICE offline project. AliRoot is an object-oriented code based on Root [64], a software specifically designed for the data analysis of high energy physics experiments. AliRoot provides the packages to perform event generation, detector simulation event reconstruction and data analysis. In particular, for the generation of the MC sample used in this analysis, the description of the detector was designed in order to match the experimental conditions for all the sub-systems during the 2013 p-Pb data taking period.

### 4.3.1 Efficiency and Acceptance

The efficiency and acceptance factors for the raw yield correction are evaluated with the MC sample described above. In particular the efficiency is defined as the number of reconstructed  $D^+$ , after all the topological and PID selections, divided by the number of generated  $D^+ \rightarrow K^- \pi^+ \pi^+$  decays with the three daughters which fulfil the acceptance requirements of  $|\eta| < 0.9$  and  $p_T > 0.1$  GeV/c for events with the  $z$ -coordinate of the generated primary vertex within 10 cm from the nominal

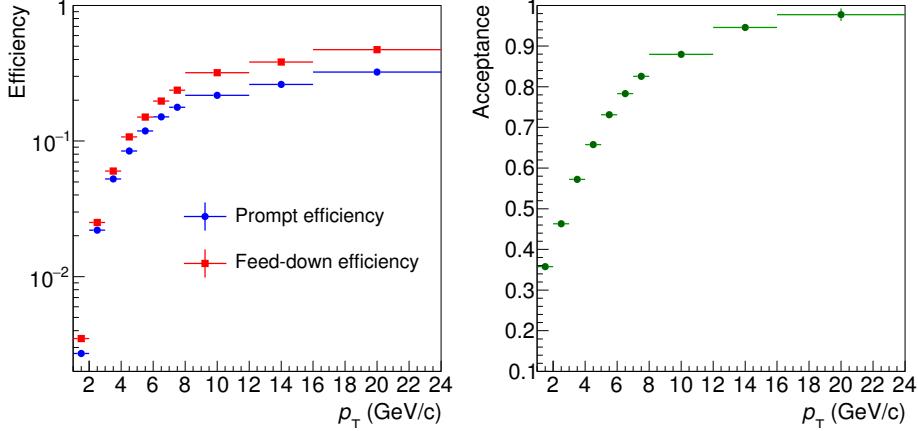


Figure 4.8. Left: the reconstruction efficiency as a function of  $p_T$  for prompt and feed-down  $D^+$  obtained applying the topological cuts in table 4.2 and the PID. Right: the acceptance factor for prompt  $D^+$ -mesons as a function of  $p_T$ .

interaction point. Furthermore, the acceptance is evaluated as the ratio between the number of the generated  $D^+ \rightarrow K^-\pi^+\pi^+$  with the three daughters in the acceptance and the  $D^+ \rightarrow K^-\pi^+\pi^+$  generated in the rapidity range  $|y| < y_{fid}$ , for events with  $|Z_{GenVertex}| < 10$  cm. In figure 4.8 the efficiencies for prompt and feed-down  $D^+$  obtained applying the topological cuts reported in table 4.2 and the PID selection (left) and the acceptance (right) are shown as a function of the transverse momentum. The efficiency for  $D^+$  coming from B-hadron decays is higher than that of prompt  $D^+$  because of the position of their decay vertex, which is on average more displaced from the primary vertex with respect to the one for prompt  $D^+$ .

## 4.4 Results from the standard analysis

In this section I present the result obtained with the theory-driven method for the evaluation of the  $D^+$  prompt fraction, which is published by the ALICE Collaboration [36]. The prompt  $D^+$   $p_T$ -differential cross section was measured applying all the event, track, topological and PID selections described in this Chapter, starting from the raw yield  $N_{raw}^{D^\pm}|_{|y| < y_{fid}}$  and correcting for the acceptance and efficiencies

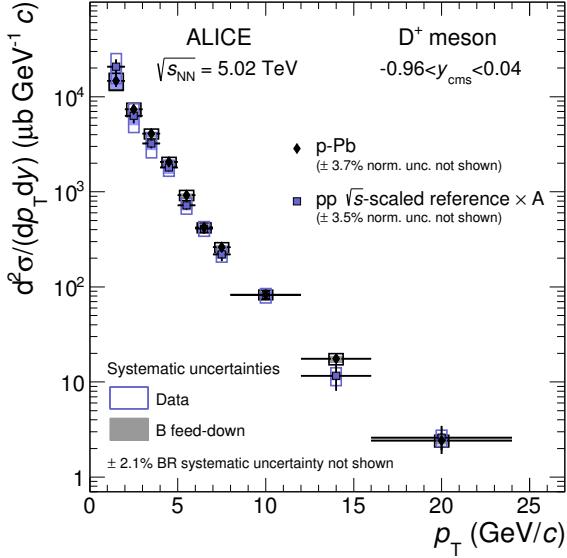


Figure 4.9. Prompt  $D^+$   $p_T$ -differential production cross section measured by the ALICE Collaboration in p-Pb collisions at  $\sqrt{s_{NN}} = 5.02 \text{ TeV}$  compared to the pp reference cross section scaled by the Pb mass number  $A = 208$  [36].

factors (see [59] for more details). In particular the cross section is computed as:

$$\frac{d\sigma_{prompt}^{D^+}}{dp_T} \Big|_{|y|<0.5} = \frac{1}{2} \cdot \frac{f_{prompt} \cdot N_{raw}^{D^\pm}|_{|y|<y_{fid}}}{\Delta p_T \Delta y} \cdot \frac{1}{(Acc \times \epsilon)_{prompt}(p_T)} \cdot \frac{\sigma_{MB}^{pPb}}{BR \cdot N_{ev}}. \quad (4.8)$$

In the formula above the factor  $1/2$  takes into account the fact that the raw yields include particles and anti-particles,  $f_{prompt}$  is evaluated with the  $N_b$  method, and  $BR$  is the branching ratio of the decay channel  $D^+ \rightarrow K^- \pi^+ \pi^+$ . The width of the fiducial acceptance coverage  $\Delta y = 2y_{fid}$  is common for all the transverse momentum intervals  $\Delta p_T$  measured. The  $\sigma_{MB}^{pPb}$  is the Minimum Bias trigger cross section from van der Meer scans and  $N_{ev}$  is the number of accepted events. This number has to include the events with the primary vertex which fulfills the condition  $|Z_{vertex}| < 10 \text{ cm}$  but is not reconstructed. This is necessary because the efficiency and acceptance factors take into account all the events with  $|Z_{GenVertex}| < 10 \text{ cm}$  independently if the primary vertex is reconstructed or not in the event. Therefore

the number of events for normalisation is computed as

$$\begin{aligned}
 N_{ev} &= N_{RecoVert}(|Z_{RecoVert}| < 10 \text{ cm}) + N_{NoVert}(|Z_{RecoVert}| < 10 \text{ cm}) = \\
 &= N_{RecoVert}(|Z_{RecoVert}| < 10 \text{ cm}) + N_{RecoVert}^{Tot} - N_{NoVert}(|Z_{RecoVert}| > 10 \text{ cm}) = \\
 &= N_{RecoVert}(|Z_{RecoVert}| < 10 \text{ cm}) + N_{RecoVert}^{Tot} - N_{NoVert}^{Tot} \frac{N_{NoVert}(|Z_{RecoVert}| > 10 \text{ cm})}{N_{RecoVert}^{Tot}}.
 \end{aligned} \tag{4.9}$$

The final result is shown in figure 4.9, compared to the pp reference cross section scaled by the Pb mass number  $A = 208$ . Since a measurement of the  $p_T$ -differential cross section at the same centre-of-mass energy was not available, the one measured at  $\sqrt{s} = 7$  TeV was used, scaling it to  $\sqrt{s} = 5.02$  TeV according to the FONLL calculations. One of the main source of uncertainty for this measurement, especially at low  $p_T$ , comes from the evaluation of the prompt fraction with the FONLL-based methods described in section 4.2.1 (shaded box in the plot). This is the motivation for the studies presented in the next Chapters, since the development of data-driven feed-down subtraction approaches aims to reduce these uncertainties and to avoid the usage of theoretical calculations in the  $D^+$  prompt fraction evaluation.

# 5

## Impact-parameter fit method for the prompt $D^+$ fraction determination

In this Chapter I will present the main data-driven method that I developed for the evaluation of  $D^+$  prompt fraction. The main advantage of this technique is that it does not rely on any theoretical assumption, unlike the FONLL-based methods presented in section 4.7. It also aims to reduce the large systematic uncertainty of the theory-driven approach, especially at low transverse momentum. This first data-driven method exploits the different shapes of the impact-parameter

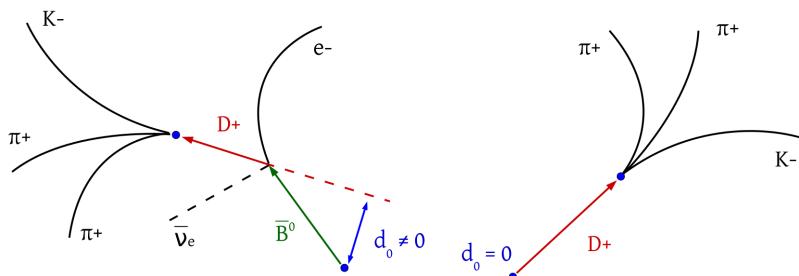


Figure 5.1. Sketch of the decay topology for prompt (right) and feed-down (left)  $D^+$  with particular emphasis on the true impact parameter.

distributions for prompt and feed-down  $D^+$ . The impact parameter is defined as the distance of closest approach between the prolonged momentum direction of the  $D^+$ -meson and the primary vertex. Since the prompt  $D^+$ -mesons come from the primary vertex, their real impact parameter is by definition exactly zero. Therefore, we can expect the measured distribution to be centred at zero and smeared for resolution effects. On the other hand, the impact parameter for feed-down  $D^+$  is in general different from zero and it depends on the decay length of the B-hadron and its relative direction with respect to the  $D^+$ -meson. In figure 5.1 is illustrated an example of decay and the corrispective true impact parameter for both prompt and feed-down  $D^+$ -mesons. Beside that, there is an alternative data-driven method, that will be presented in Chapter 6.

## 5.1 Analysis strategy

Before to start with the description of the analysis, I would like to introduce the strategy adopted. First and foremost, all the selections applied (event, track quality, topological cuts and PID) are the same as those used for the standard analysis (see Chapter 4). This is very important in order to be able to compare the two results, since the prompt fraction depends on the topological cuts applied.

Then, the procedure is divided in two main steps:

- A prefit phase, which consists in the parametrisation of the impact-parameter distributions in the transverse plane ( $d_0^{xy}$ ) for the three different types of candidates that contribute to the measured distribution: prompt  $D^+$ , feed-down  $D^+$  and background. For the prompt and feed-down  $D^+$  the distributions are obtained from the Monte Carlo sample described in section 4.3. The distribution for the combinatorial background is directly taken from the data sample (see section 4.1.1), in the invariant-mass region aside the  $D^+$  peak (see section 5.2.3). Each contribution is fitted with a specific function, which is called *template*.
- An unbinned log-likelihood fit to the impact-parameter distributions in the data using the templates obtained in the previous step.

All the parameters obtained in the prefit phase are fixed in the final fit on data, except for the detector-resolution term which is kept free in order to compensate a possible not perfect description of the impact-parameter resolution in the MC simulation (a detailed description will be provided in section 5.2.1).

This method requires a certain amount of statistics in order to obtain a stable fit result, therefore the transverse momentum range considered for this analysis is narrower ( $2 < p_T < 16$  GeV/c instead of  $1 < p_T < 24$  GeV/c) with respect to the one used for the standard analysis obtained with the theory-driven approach, presented in section 4.9.

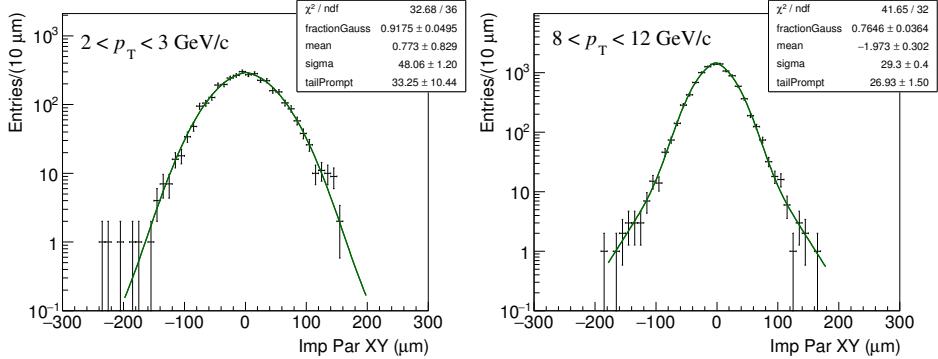


Figure 5.2. Examples of fit to the simulated impact-parameter distributions for prompt  $D^+$ -mesons in two different  $p_T$  intervals.

## 5.2 Prefit on the impact-parameter distributions

In this section I will present the prefits performed to the impact-parameter distributions in the transverse plane for prompt and feed-down  $D^+$ -mesons and for the combinatorial background.

### 5.2.1 MC impact-parameter distributions for prompt $D^+$

The impact-parameter distributions in the transverse plane for prompt  $D^+$  were obtained from the Monte Carlo simulation described in section 4.3, in seven  $p_T$  intervals in the range  $2 < p_T < 16$  GeV/c. Since the generated impact parameter for prompt  $D^+$ -mesons is by definition zero, the distributions from the reconstruction of the MC data (*simulated distributions*) are approximately Gaussian due to the finite track and primary vertex resolution of the detector. However, as shown in figure 5.2, they clearly present non-Gaussian tails, which are more prominent at higher  $D^+$  transverse momentum. For this reason, a simple Gaussian function would not give a proper description of the distribution and therefore the fit function used is the sum of a Gaussian and a symmetric exponential function

$$\begin{aligned}
 F^{prompt}(d_0^{xy}) &= F^{g+exp}(d_0^{xy}) = \\
 &= N \cdot \left\{ \frac{f_{gaus}^{prompt}}{\sqrt{2\pi}\sigma_{prompt}} e^{-\frac{(d_0^{xy}-\mu_{prompt})^2}{2\sigma_{prompt}^2}} + \frac{1-f_{gaus}^{prompt}}{2\lambda_{prompt}} e^{-\frac{|d_0^{xy}-\mu_{prompt}|}{\lambda_{prompt}}} \right\}, \tag{5.1}
 \end{aligned}$$

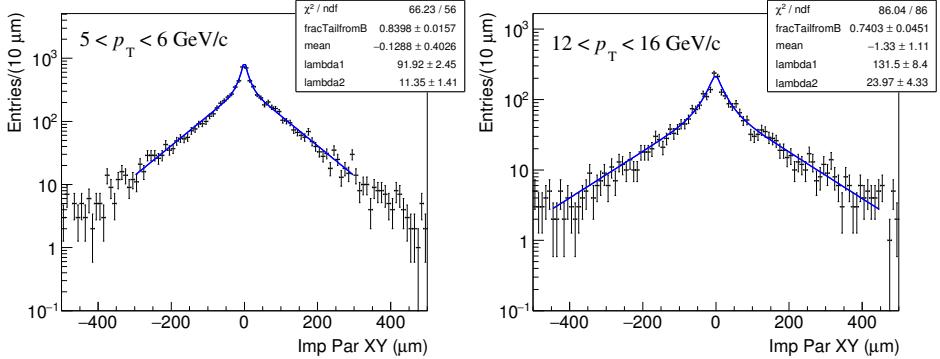


Figure 5.3. Examples of fit to the true generated impact-parameter distributions for feed-down  $D^+$ -mesons in two different  $p_T$  intervals.

where  $N$  is a normalisation factor which is fixed to the integral of the histogram,  $f_{\text{gaus}}^{\text{prompt}}$  represents the fraction of the integral contained in the Gaussian function,  $\mu_{\text{prompt}}$  is the common mean value for the two contributions,  $\lambda_{\text{prompt}}$  is the slope of the exponential function and  $\sigma_{\text{prompt}}$  is the width of the Gaussian function, which represents the resolution of the detector. As shown in the two examples in figure 5.2, the  $\chi^2$  values show that the chosen functional form gives a reasonably good description of the distributions. Furthermore, it is possible to notice how the Gaussian width decreases with increasing  $D^+$  transverse momentum, due to the better impact-parameter resolution on high  $p_T$  tracks (right panel of figure 3.6). Also the  $f_{\text{gaus}}^{\text{prompt}}$  parameter decreases with increasing  $p_T$ , because of the higher contribution in the exponential. All the other fits to the impact-parameter distributions for prompt  $D^+$  in the  $p_T$  intervals considered for the analysis can be found in the appendix A.1.

### 5.2.2 MC impact-parameter distributions for feed-down $D^+$

The true impact parameter for the feed-down  $D^+$ -mesons is different from zero because in general the momentum of the  $D^+$  is not parallel to the one of the B-hadron, as shown in the left panel of figure 5.1. In particular it depends on the decay length of the B-hadron and the angle between the directions of the two particles. For this reason the impact-parameter distribution of feed-down  $D^+$  is more difficult to parametrise with respect to the one of prompt  $D^+$ . In order to achieve a satisfactory description of the distribution, two steps were performed. First, the true impact-parameter distribution in the  $xy$ -plane for feed-down  $D^+$  from the Monte Carlo simulation in the seven  $p_T$  intervals of the analysis, applying

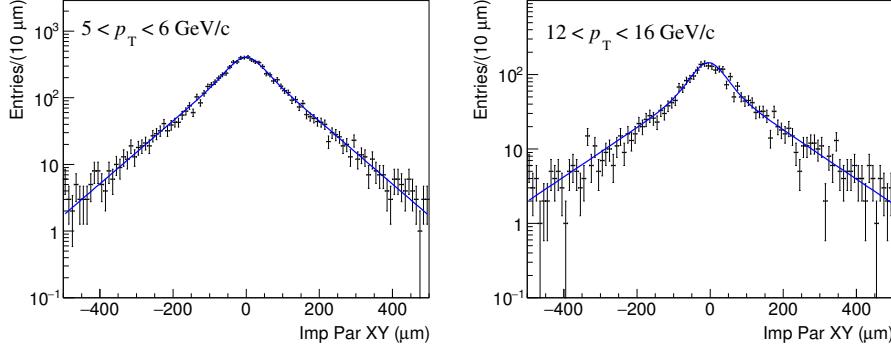


Figure 5.4. Examples of reconstructed impact-parameter distributions for feed-down  $D^+$ -mesons in two different  $p_T$  intervals. The function superimposed is the convolution between the true impact parameter function for feed-down  $D^+$  (eq. 5.2) and the resolution term (eq. 5.1).

all the selections described in Chapter 4, was fitted with a double exponential function, symmetrised with respect to its mean value

$$F_{true}^{feed-down}(d_0^{xy}) = N \cdot \left\{ \frac{f_{\lambda_1}^{FD}}{2\lambda_1^{FD}} e^{-\frac{|d_0^{xy} - \mu_{FD}|}{\lambda_1^{FD}}} + \frac{1 - f_{\lambda_1}^{FD}}{2\lambda_2^{FD}} e^{-\frac{|d_0^{xy} - \mu_{FD}|}{\lambda_2^{FD}}} \right\}, \quad (5.2)$$

where  $N$  is again fixed to the integral of the histogram,  $\lambda_1^{FD}$  and  $\lambda_2^{FD}$  are the slopes of the two symmetrised exponential functions,  $f_{\lambda_1}^{FD}$  is the fraction of integral contained in the first function and  $\mu_{FD}$  is the common mean value. The impact-parameter range in the fit increases with the transverse momentum, starting from  $[-250, 250] \mu\text{m}$  in the interval  $2 < p_T < 3 \text{ GeV}/c$  up to  $[-450, 450] \mu\text{m}$  in  $12 < p_T < 16 \text{ GeV}/c$ , in order to get a good description of the distribution in the steepest region. Two examples of fit to the true impact-parameter distribution for feed-down  $D^+$  are presented in figure 5.3. The reasonable agreement between the functional form chosen and the distributions is confirmed by the  $\chi^2$  values obtained. The second step for the parametrisation consisted in a convolution of the function obtained for the feed-down  $D^+$  true impact-parameter distribution  $F_{true}^{feed-down}(d_0^{xy})$  and the  $F_{prompt}(d_0^{xy})$  function, which represents the effect of the detector resolution, for each  $p_T$  interval

$$F_{reco}^{feed-down}(d_0^{xy}) = N \cdot \int_{d_0^{xy,min}}^{d_0^{xy,max}} F_{true}^{feed-down}(d_0^{xy'}) F_{prompt}(d_0^{xy} - d_0^{xy'}) dd_0^{xy'}, \quad (5.3)$$

where the limits in the convolution were set in order to cover the impact-parameter range between  $[-1000,1000]$   $\mu\text{m}$ . In figure 5.4 the simulated impact-parameter distributions in the same  $p_T$  intervals are shown, with the function obtained with the convolution superimposed. All the other plots can be found in the appendix A.2.

### 5.2.3 Impact-parameter distributions for the background

Unlike the impact-parameter distributions for prompt and feed-down  $D^+$ , which were modelled from the MC simulation, the distributions for the combinatorial background were obtained directly from the Minimum Bias p-Pb data sample described in section 4.1.1, selecting two mass intervals on the two sides of the  $D^+$  peak, where there is no signal. The invariant-mass intervals considered, called *side-band regions*, are defined by:

$$4\sigma < |M - M_{peak}| < 15\sigma, \quad (5.4)$$

where  $M_{peak}$  and  $\sigma$  are respectively the mean value and the width of the Gaussian function used to describe the signal in the invariant-mass fit. The lower limit of  $4\sigma$  allows to remove most of the signal (99.99%) from the impact-parameter distribution and at the same time to select a mass region close to the peak, in order to get a reasonable description of the background under the  $D^+$  peak. The mass interval is symmetric with respect to the mass of the  $D^+$  to keep more statistics for the parametrisation and to average on a possible invariant-mass dependence of the

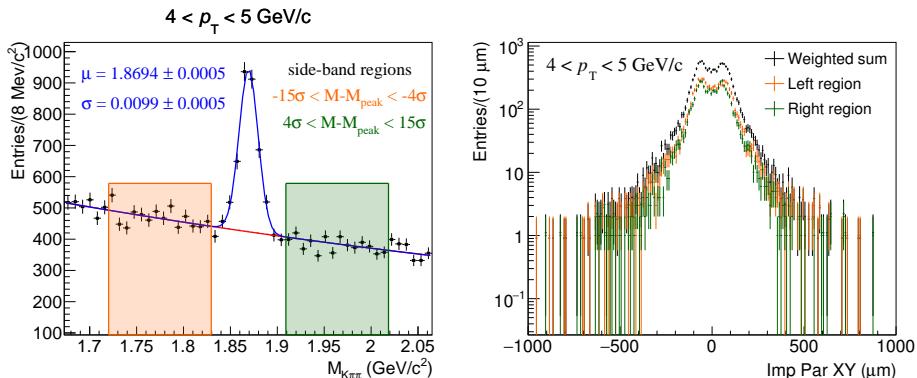


Figure 5.5. Left: invariant-mass intervals considered for the parametrisation of the impact-parameter distribution for the background (*side-band regions*) in the transverse momentum interval  $4 < p_T < 5$  GeV/c. Right: resulting background impact-parameter distribution.

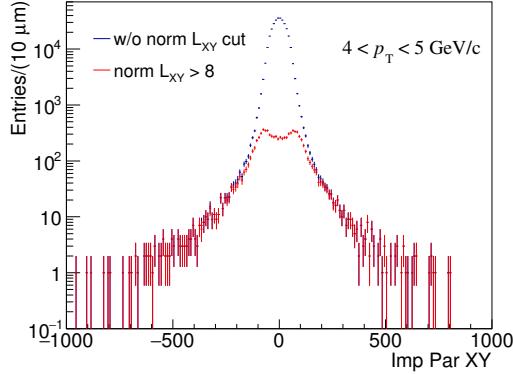


Figure 5.6. Comparison between the impact-parameter distributions for the combinatorial background obtained from the side-bands region, with (red) and without (blue) cut on normalised decay length  $xy$ . In both cases all the other topological selections reported in table 4.2 are applied.

impact-parameter distribution. In figure 5.5 the side-band regions selected (left) and the resulting background impact-parameter distributions in the  $p_T$  interval  $4 < p_T < 5 \text{ GeV}/c$  (right) are shown. The distributions obtained from the two side-band regions are combined with a weighted sum, according to their integrals. The double-peak structure is induced by the topological selections applied, especially by the cut on the normalised decay length in the  $xy$ -plane, as shown in figure 5.6. The red distribution is obtained from the side-bands with the normalised decay length cut, while the blue one without. Both the distributions are obtained applying all the other topological selections summarised in table 4.2. Because of this double-peak shape, a sum of two Gaussian functions with exponential tails is used for the fit

$$F^{bkg}(d_0^{xy}) = N \cdot \left\{ F_1^{gaus+exp}(d_0^{xy}; f_{gaus}^{bkg}, \mu_1^{bkg}, \sigma_{bkg}, \lambda_{bkg}) + \right. \\ \left. + F_2^{gaus+exp}(d_0^{xy}; f_{gaus}^{bkg}, \mu_2^{bkg}, \sigma_{bkg}, \lambda_{bkg}) \right\}. \quad (5.5)$$

In order to avoid as much as possible statistical fluctuations, the two  $F^{gaus+exp}$  functions were forced to have all the parameters equal, except for the two mean values,  $\mu_1$  and  $\mu_2$ . The common parameters are hence the following:  $f_{gaus}^{bkg}$  is the fraction of integral contained in each Gaussian function,  $\sigma_{bkg}$  is the width of the Gaussians and  $\lambda_{bkg}$  the slope of the exponential contributions. The  $\chi^2$  values

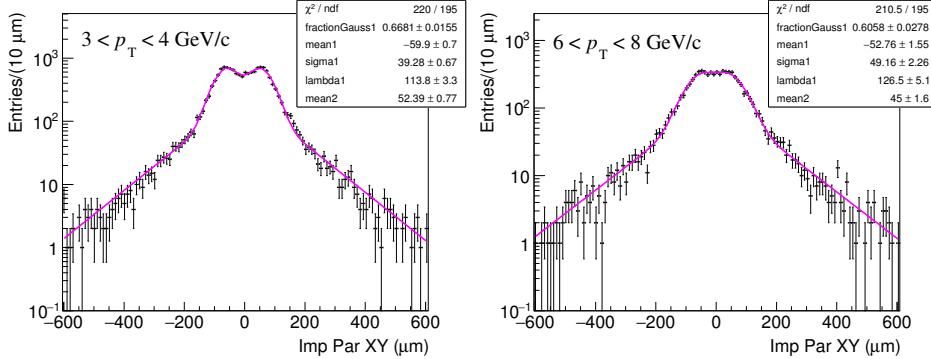


Figure 5.7. Examples of fit to the impact-parameter distributions for the combinatorial background obtained from the side-band regions, in two different  $p_T$  intervals.

obtained confirm that this functional form gives an acceptable description of the distributions. From the two fit examples shown in figure 5.7 is it possible to notice how the double-peak structure becomes less pronounced with increasing  $p_T$ . All the other fits to the impact-parameter distributions in the transverse plane for the background are reported in the appendix A.4.

### 5.3 Unbinned log-likelihood fits on data

In this section I will describe the procedure used to fit the impact-parameter distributions in the transverse plane for  $D^+$  candidates in the Minimum Bias data sample for p-Pb collisions at  $\sqrt{s_{NN}} = 5.02 \text{ TeV}$ , to extract the prompt fraction. First a brief motivation for using the log-likelihood unbinned fit technique is provided, then the fitting procedure is described.

#### 5.3.1 Binned vs. unbinned fit

Given a set of experimental observables  $\vec{x} = (x_1, \dots, x_n)$  described by the probability density function  $f(\vec{x}; \vec{\theta})$ , which depends on the set of parameters  $\vec{\theta} = (\theta_1, \dots, \theta_m)$ , the likelihood function is equal to the probability function evaluated with fixed  $\vec{x}$ , depending on  $\vec{\theta}$

$$\mathcal{L}(\vec{\theta}; \vec{x}) = f(\vec{x}; \vec{\theta}). \quad (5.6)$$

Both the *binned* and *unbinned* fit techniques are performed finding the set of parameters  $\vec{\theta}$  which maximise the likelihood function, or more conveniently the logarithm

of  $\mathcal{L}$ . The difference between the two fit methods is in the treatment of the data, which leads a different definition of the likelihood function itself.

Let assume  $x$  to be a single variable and  $\mathbf{x} = (x_1, \dots, x_N)$  a sample of  $N$  independent measurements of  $x$ . In case of unbinned fit the log-likelihood is defined as

$$\ln\mathcal{L}(\vec{\theta}; \mathbf{x}) = \sum_{i=1}^N \ln f(x_i; \vec{\theta}), \quad (5.7)$$

where  $f(x_i, \vec{\theta})$  is the probability density function evaluated in the point  $x_i$  and depending on the parameters  $\vec{\theta}$ . On the other hand, the binned fit is performed by first filling an histogram with the measurements  $\mathbf{x}$ , and therefore the log-likelihood is modified as follow

$$\ln\mathcal{L}(\vec{\theta}) = \sum_{i=1}^{N_{bins}} n_i \ln \nu_i(\vec{\theta}) \quad (5.8)$$

where  $N_{bins}$  is the number of bins of the histogram and  $n_i$  is the number of entries in the  $i$ -th bin. In this case the expected probability for the value  $x_i$  does not enter directly in the formula, which instead contains the expected number of entries in the  $i$ -th bin, given by

$$\nu_i(\vec{\theta}) = N \int_{x_i^{min}}^{x_i^{max}} f(x, \vec{\theta}) dx. \quad (5.9)$$

Therefore the binned fit becomes convenient in terms of computing time when the sample of measurement  $\mathbf{x}$  is large, while it results less precise in case of low statistics, because of the few entries in the bins of the histogram.

The unbinned technique was chosen for the fits on data because the impact-parameter distributions in the transverse plane show a small number of entries in the tails, where the shape for prompt and feed-down  $D^+$ -mesons are more different playing a crucial role for the proper separation of the two contributions.

### 5.3.2 Impact-parameter fits

The impact-parameter distributions for the final fit were obtained from the data, applying all the selections described in Chapter 4, for the seven  $p_T$  intervals of the analysis. Furthermore, the invariant-mass range considered for the  $D^+$  candidates is

$$|M - M_{peak}| < 2\sigma \quad (5.10)$$

in order to keep as much high as possible the signal-to-background ratio, while keeping the majority of the signal (95.5%). In figure 5.8 the selected invariant-mass range for the  $4 < p_T < 5$  GeV/c interval (left) and the corresponding impact-parameter distribution obtained (right) are shown. As already mentioned the fit

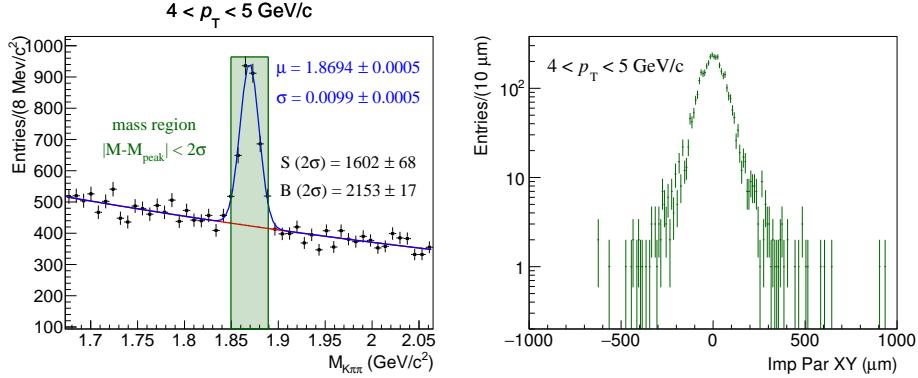


Figure 5.8. Left: invariant-mass range considered for the impact-parameter distribution final fit in the transverse momentum interval  $4 < p_T < 5 \text{ GeV}/c$ . Right: resulting impact-parameter distribution.

was not performed to the histogram, but with the unbinned technique, and therefore to a ROOT structure called *tree*, which allows to store the data without binning them. The fitting function used is

$$F(d_0^{xy}) = N \cdot \left\{ \frac{S}{N} \left[ f_{prompt} F^{prompt}(d_0^{xy}; \sigma_{prompt}) + (1 - f_{prompt}) F^{feed-down}(d_0^{xy}; \sigma_{prompt}) \right] + \frac{N - S}{N} F^{bkg}(d_0^{xy}) \right\}, \quad (5.11)$$

where the normalisation factor  $N$  is fixed to the number of entries in the tree and  $S$  is the parameter which represent the amount of signal, fixed to the value extracted with the fit to the invariant-mass distribution, integrating the signal fit function over the selected invariant-mass range (see figure 5.8). The values  $S$  obtained with this procedure are reported in table 5.1.  $F^{prompt}(d_0^{xy})$ ,  $F^{feed-down}(d_0^{xy})$  and  $F^{bkg}(d_0^{xy})$  are the functions for prompt  $D^+$ -mesons, feed-down  $D^+$ -mesons and the combinatorial background respectively, obtained in the prefit phase (see section 5.2). All the parameters for each contribution are fixed to the values obtained from the prefits, except for the Gaussian width of the detector-resolution term  $\sigma_{prompt}$ , which is kept free to vary within 20% around the value obtained by fitting the MC distribution for prompt  $D^+$  to compensate a possible not perfect description of the impact-parameter resolution in the Monte Carlo simulation. None of the values obtained from the fit reaches the 20% limit. In section 5.4.1 the values obtained from the data and the MC simulation are compared in order to evaluate a possible systematic uncertainty. The only other free parameter in the fit function is the fraction of prompt  $D^+$ -mesons  $f_{prompt}$ , which is the quantity that we want

$p_T$ (GeV/c)	$S(2\sigma)$	$\sigma_S/S(2\sigma)$ stat.	$\sigma_S/S$ syst.
[2,3]	$975 \pm 60$	0.06	0.08
[3,4]	$1734 \pm 72$	0.04	0.05
[4,5]	$1602 \pm 68$	0.04	0.05
[5,6]	$1122 \pm 52$	0.05	0.05
[6,8]	$1297 \pm 55$	0.04	0.05
[8,12]	$985 \pm 53$	0.05	0.05
[12,16]	$263 \pm 27$	0.10	0.08

Table 5.1. Raw yields in the invariant-mass range  $|M - M_{peak}| < 2\sigma$  with the statistical error in the seven  $p_T$  intervals of the analysis. The systematic uncertainties on the yield extraction are taken from the values estimated for the standard analysis [59].

to measure. The impact-parameter fit range gradually increases with increasing transverse momentum to better account for the broadening of the distributions. Figure 5.9 shows the resulting fits in the transverse momentum intervals of the analysis. In this figure the histogram and the curves, rescaled according to the integral of the histogram, are depicted in order to provide a visual representation of the unbinned fit.

## 5.4 Systematic uncertainties

In this section the various sources of systematic uncertainty that affect the fraction of prompt  $D^+$  extracted from the fit of the impact-parameter distributions in the  $xy$ -plane are presented. They include:

1. Systematic uncertainty due to the imperfections in the distributions taken from the simulations and the side-bands of the signal peak, in particular:
  - a. Shapes for prompt  $D^+$ , feed-down  $D^+$  and background impact-parameter distributions.
  - b. Shape of generated  $p_T$  distribution of prompt  $D^+$  and  $B$  mesons decaying to  $D^+$ .
2. Systematic uncertainty induced by the uncertainty on the signal yield  $S$ .
3. Systematic uncertainty evaluated by checking the consistency of the procedure.

Several tests have been performed in order to quantify these sources of systematic effects. Concerning the first category, tests on the degree of agreement of the

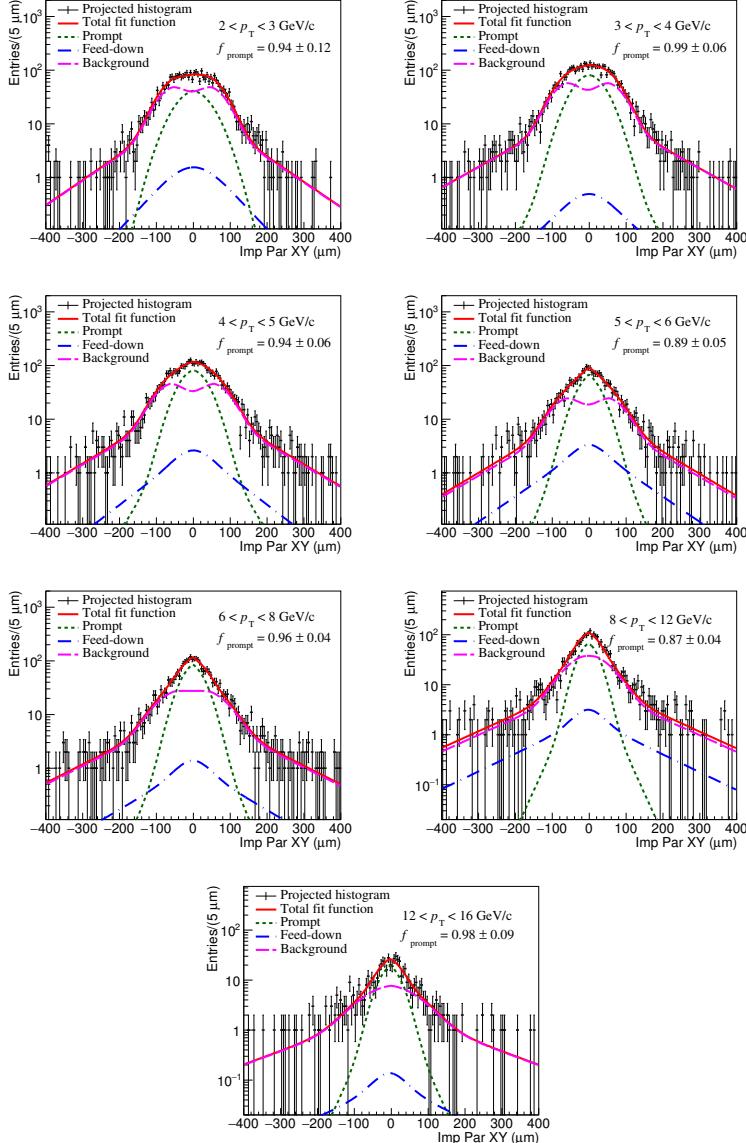


Figure 5.9. Fits to the  $D^+$  impact-parameter distributions in the  $xy$ -plane in the seven transverse momentum intervals of the analysis. The curves are the fit functions for prompt  $D^+$  (green), feed-down  $D^+$  (blue), background (magenta) and the sum of all contributions (red).

resolution term in data and MC, the impact-parameter fitting range and the parametrisation of the background distribution from the side-band regions have been performed. To take into account the discrepancy between the  $p_T$ -shape in data and MC, the impact-parameter distributions of prompt and feed-down  $D^+$  used in the prefit phase have been reweighted in order to match the FONLL  $p_T$ -shape. The error induced by the signal yield was quantified by changing the  $S$  parameter within  $1\sigma$  in the fitting function. Finally, to check the consistency of the procedure the invariant-mass window used to obtain the impact-parameter distributions was changed and a MC closure test was performed. All these tests will be described in more detail in the following subsections.

### 5.4.1 Systematic uncertainty due to the impact-parameter shapes

#### Test on the degree of agreement between the resolution term in data and MC

In order to test the degree of agreement between the detector resolution in data and MC, all the fits have been repeated fixing  $\sigma_{prompt}$  to the values obtained from the

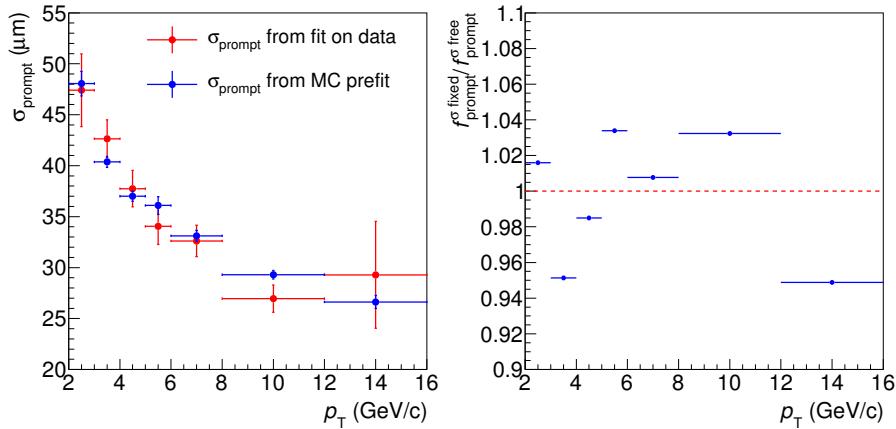


Figure 5.10. Left: comparison between the  $\sigma_{prompt}$  parameter as a function of  $p_T$  obtained from the impact-parameter fit on data and the one obtained by fitting the simulated distributions for prompt  $D^+$  in the prefit phase. Right: ratio between the  $D^+$  prompt fraction  $f_{prompt}$  obtained by fixing and letting free  $\sigma_{prompt}$  in the fit on data as a function of  $p_T$ .

simulated impact-parameter distributions for prompt  $D^+$  in the prefit phase. To estimate the uncertainty, the ratio of the  $f_{prompt}$  values obtained with and without fixing the  $\sigma_{prompt}$  parameter was computed. As shown in the right panel of figure 5.10 this ratio fluctuates around unity without any particular trend, therefore no systematic uncertainty due to the detector-resolution term is assigned. This observation is also confirmed by the comparison between the values of  $\sigma_{prompt}$  obtained by fitting the data and in the prefit phase on the simulated sample reported in the left panel of figure 5.10, which shows that all the values agree within the statistical uncertainty.

### Variation of the fitting range

Another test performed to quantify the systematic uncertainty due to the shape of the impact-parameter distribution templates consisted in varying the impact-parameter fitting range. The prefit and the fit phases were repeated varying the impact-parameter ranges from  $[-200,200]$   $\mu\text{m}$ , up to  $[-1000,1000]$   $\mu\text{m}$  coherently. Then, in order to disentangle the effect due to the background distribution from the others, the same procedure was repeated fixing the fitting range of the background in the prefit phase to the value of  $[-1000,1000]$   $\mu\text{m}$  and varying only the ranges for the prompt and feed-down components. Figure 5.11 shows the ratio of the prompt

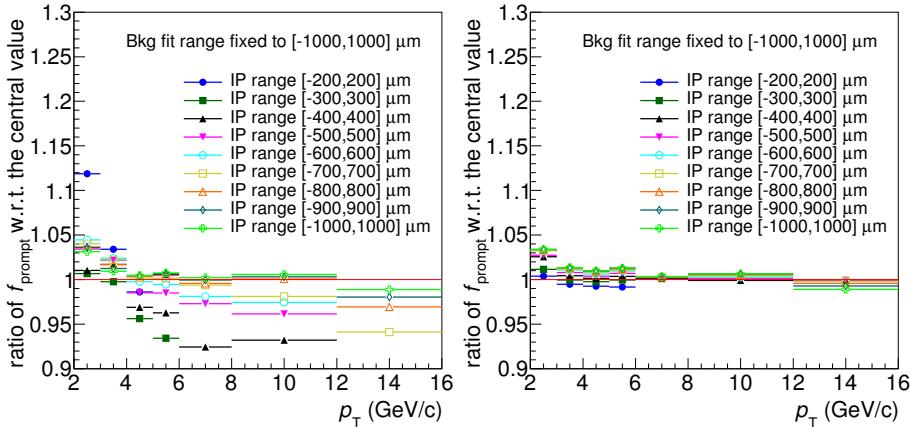


Figure 5.11. Ratio of  $f_{prompt}$  as a function of  $p_T$  obtained varying the fit range with respect to the central value. In the left plot all the fit ranges both in the fit and the prefit phase are changed coherently, while in the right one the impact-parameter range for the background is fixed to  $[1000,1000]$   $\mu\text{m}$ .

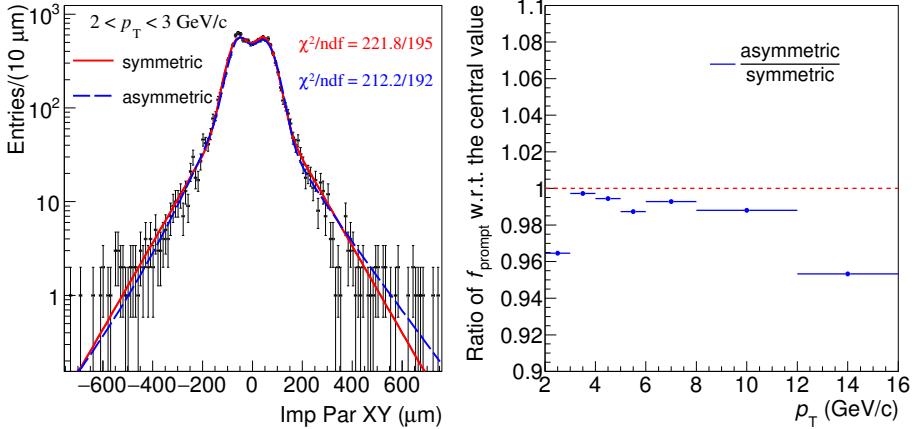


Figure 5.12. Left: example of fit to the background distribution in the  $p_{\text{T}}$  interval  $2 < p_{\text{T}} < 3 \text{ GeV/c}$  with the symmetric and asymmetric functions. Right: ratio of  $f_{\text{prompt}}$  obtained with the symmetric and the asymmetric fit functions for the background.

fraction obtained for each trial with respect to the central value, in the two cases. From these plots it is clear that there is a systematic trend and that the larger source of uncertainty comes from the variation of the fit range of the background distribution in the prefit phase. It should be also pointed out that the narrower impact-parameter ranges tested do not allow to describe properly the background distribution, which is much broader, especially at high  $p_{\text{T}}$ . For this reason, the first two fitting ranges were not considered for the transverse momentum interval  $6 < p_{\text{T}} < 12 \text{ GeV/c}$ , and the first three for  $12 < p_{\text{T}} < 16 \text{ GeV/c}$ . The systematic uncertainties were assigned by evaluating the RMS of the trials shown in the left panel of figure 5.11 for each  $p_{\text{T}}$  bin, which results to be of about 2% in the full transverse momentum range.

### Uncertainty on the determination of the background distribution

For the evaluation of the systematic uncertainty due to the parametrisation of the combinatorial background distributions, two different sources have been taken into account:

- Systematic uncertainty due to the fit function chosen.
- Systematic uncertainty due to the invariant-mass range in the side-band regions used to obtain the impact-parameter distribution for the background.

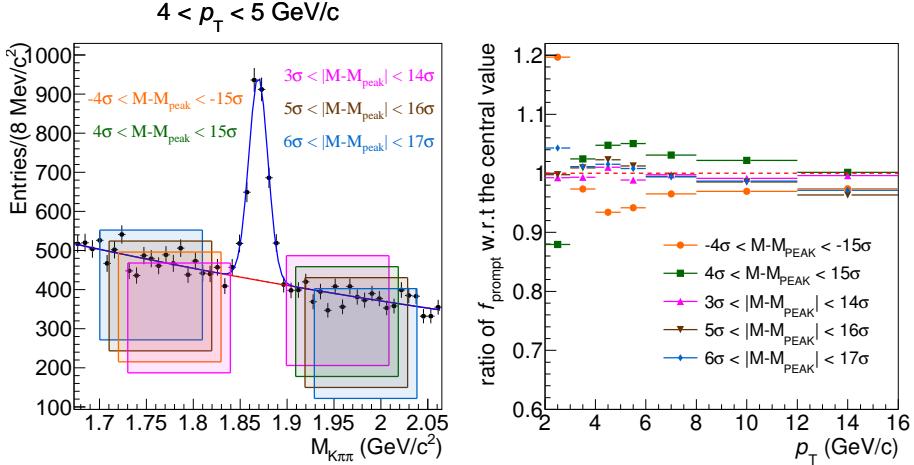


Figure 5.13. Left: the five invariant-mass intervals in the side-bands region tested to evaluate the systematic uncertainty for the combinatorial background parametrisation in the  $p_T$  interval  $4 < p_T < 5 \text{ GeV}/c$ . Right: The ratio of  $f_{\text{prompt}}$  with respect to the reference value obtained with the mass region  $4\sigma < |M - M_{\text{peak}}| < 15\sigma$  for the background parametrisation as a function of  $p_T$ .

In order to take into account the possibility that the background distribution is not symmetric with respect to its central value, the prefit phase and the final fits have been repeated with a slightly different and asymmetric fit function for the background contribution. This alternative function is indeed the same used as default (see equation 5.5), but with all the parameters of the two Gaussian functions with exponential tails free and independent. The left plot of figure 5.12 shows the comparison between the two different fit functions obtained for the  $p_T$  interval  $2 < p_T < 3 \text{ GeV}/c$ . The two curves result to be similar and both give an acceptable value of the  $\chi^2$ . In the right panel of the same plot, the relative variation between the prompt fraction of  $D^+$ -mesons obtained in the two cases is shown. Based on this result, a 4% uncertainty for  $2 < p_T < 3 \text{ GeV}/c$  and 1% in the remaining  $p_T$  intervals was assigned. In particular, in the  $p_T$  range  $12 < p_T < 16 \text{ GeV}/c$  an uncertainty smaller than the observed variation of  $f_{\text{prompt}}$  was assigned because of the bad description of the distribution obtained with the asymmetric function.

As mentioned in section 5.2.3 the impact-parameter distribution for the combinatorial background was obtained from the  $D^+$  candidates in the side-band regions. However, the choice of the invariant-mass range used to define these regions is arbitrary and can induce some systematic effects. Therefore, the fitting procedure

was repeated changing the limits of the side-band regions for the parametrisation of the background. In particular two tests have been performed:

- Selecting the same invariant-mass windows used for the central value, but considering only the left or the right side of the  $D^+$  peak.
- Selecting different invariant-mass ranges symmetric with respect to the  $D^+$  peak.

In the right panel of figure 5.13 all the different ranges considered in the side-band regions are displayed. In the right panel the ratio of  $f_{prompt}$  with respect to the central value is shown as a function of the transverse momentum. The largest variations are observed for the two trials performed selecting only the left or the right invariant-mass intervals. However these two tests are extreme, since the choice of two ranges symmetric around the peak was done to average on a possible invariant-mass dependence of the impact-parameter distributions and therefore to get a better estimation of the background under the  $D^+$  peak. For this reason, a systematic uncertainty of 6% in the first  $p_T$  interval and 3% for  $3 < p_T < 16$  GeV/c was assigned, instead of the full spread.

### 5.4.2 Systematic uncertainty due to the assumed $p_T$ shape

Since the shape of the impact-parameter distributions in the  $xy$ -plane depends on the transverse momentum, and the distributions for prompt and feed-down are modelled starting from the MC distributions in  $p_T$  intervals of finite size, it is important to use reasonable shapes for the generated the  $p_T$  distribution of  $D^+$ -mesons and B-mesons which decay to  $D^+$  in the simulation. However, as shown in figure 5.14, the shape of the generated  $p_T$  spectrum in the MC using the MC Pythia generator is slightly different with respect to the FONLL one, which describes reasonably well the measurements (see section 2.1). This discrepancy could generate a bias, especially in the transverse momentum intervals where the  $p_T$  shape in FONLL and MC is more different and in the wider  $p_T$  bins of the analysis. Therefore, the MC impact-parameter distributions for prompt (feed-down)  $D^+$  have been reweighted, according to the  $D^+$   $p_T$  using as weight the ratio between the  $p_T$  spectrum of  $D^+$ -mesons (B-mesons) predicted by FONLL and that from the MC simulation. Figure 5.15 shows that the difference between the impact-parameter distributions obtained with and without applying the weights, for prompt and feed-down  $D^+$ , is very small. As a consequence, the variation of  $f_{prompt}$  obtained with the reweighted impact-parameter distributions results to be of the order of  $\sim 0.1\%$ , which is completely negligible with respect to the other sources of systematic uncertainty.

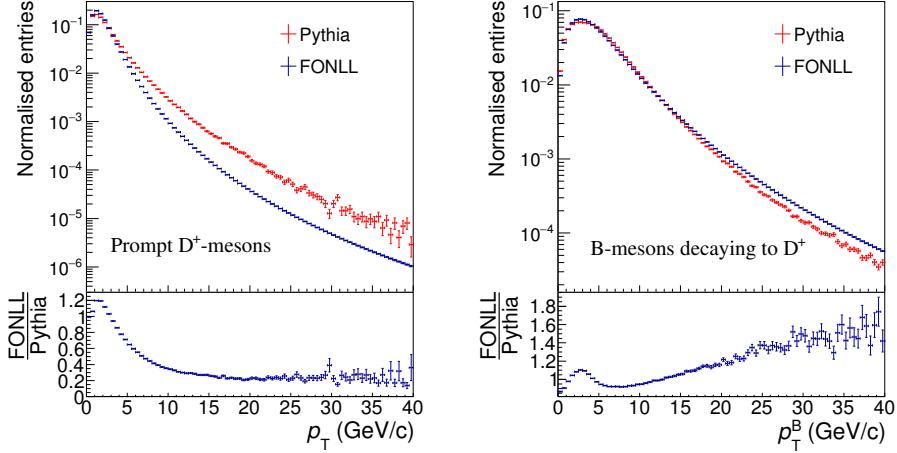


Figure 5.14.  $p_T$ -spectrum of generated  $D^+$ -mesons (left) and B-mesons decaying to  $D^+$  (right) from the Pythia simulation compared to FONLL calculations.

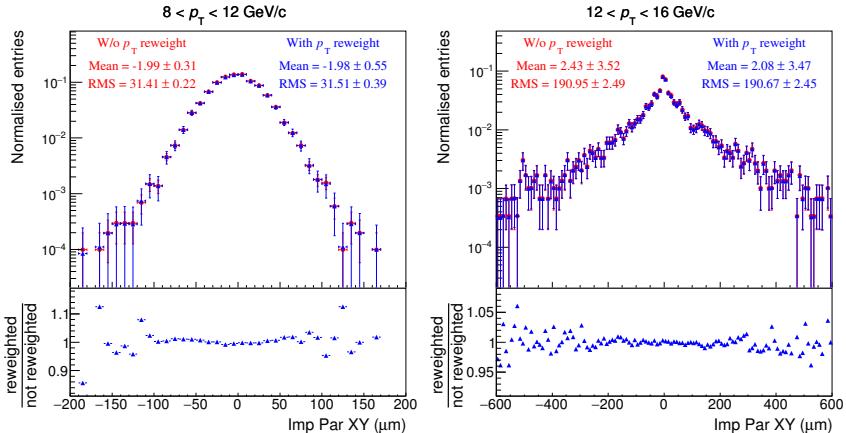


Figure 5.15. Top: comparison between the prompt and feed-down  $D^+$  impact-parameter distributions with and without applying the weight factor used to reproduce the FONLL  $p_T$  spectrum for prompt  $D^+$ -mesons and B-mesons which decay to  $D^+$  respectively, in two  $p_T$  intervals of the analysis. Bottom: ratio between the reweighted and not reweighted impact-parameter distributions.

### 5.4.3 Systematic uncertainty induced by the uncertainty on the signal yield $S$

In the impact-parameter fit on data, the amount of signal  $S$ , which is a parameter of the fit function 5.11, is fixed to the value obtained by integrating the Gaussian term from the fit to the invariant-mass distribution in a range of  $2\sigma$  around the mean. However, this value is affected by a statistical and systematic error, which has to be taken into account in the evaluation of the  $D^+$  prompt fraction. Therefore, the fits were repeated fixing the  $S$  parameter to the value of the raw yield plus or minus its uncertainty, which is reported in table 5.1. The systematic uncertainty for the raw yield extraction is taken from the values of the standard analysis published by ALICE [36], which were evaluated in [59], while the total uncertainty is computed summing in quadrature the statistical and systematic errors

$$\sigma_S = \sqrt{\sigma_S^2(\text{stat}) + \sigma_S^2(\text{syst})}. \quad (5.12)$$

As shown in the left panel of figure 5.16, an increase of the signal corresponds to a decrease of the  $D^+$  prompt fraction. Since  $f_{\text{prompt}}$  is computed for the calculation of the  $p_T$ -differential cross section, which is proportional to both the prompt fraction and the raw yield (see equation 4.8), we can expect a partial compensation of this

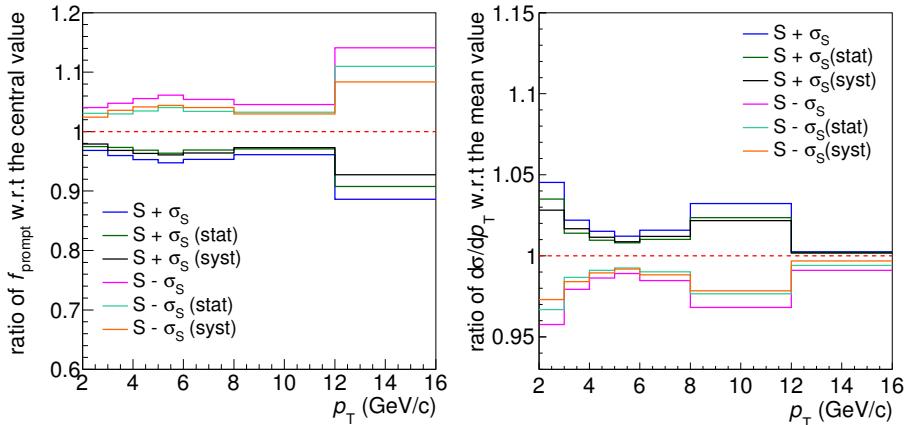


Figure 5.16. Left: ratio of  $f_{\text{prompt}}$  obtained fixing the signal parameter  $S$  to its value plus or minus its statistical and systematic uncertainty, with respect to the central value. Right: relative variation of the  $p_T$ -differential cross section due to the variation of the signal in the raw yield and in  $f_{\text{prompt}}$  within its statistical and systematic uncertainties.

uncertainty in the cross section calculation. This is indeed the case, as shown in the right panel of figure 5.16, where the relative variation of the  $p_T$ -differential cross section calculated using the formula 4.8 is reported. The cross section values were computed using consistently in the calculations the values  $S$ ,  $S + \sigma_S$  and  $S - \sigma_S$  for the raw yield and for the determination of  $f_{prompt}$  in the impact-parameter fits. Not to double count the error on the signal extraction in the computation of the cross section, the uncertainty is assigned on the basis of the  $d\sigma/dp_T$  variation, shown in the right panel of figure 5.16. Therefore, a 4% in the transverse momentum interval  $2 < p_T < 3$  GeV/c and 2% in the remaining  $p_T$  range was assigned.

#### 5.4.4 Consistency of the fitting procedure

##### Variation of the invariant-mass window

As explained above, the  $f_{prompt}$  result was obtained selecting  $2\sigma$  invariant-mass region around the peak, to build the candidate impact-parameter distribution and to extract the amount of signal  $S$  from the fit to the invariant-mass distributions (see figure 5.8). The stability of the result was tested by changing the number of  $\sigma$  around the  $D^+$  peak. In particular the impact-parameter fits have been repeated on the distributions obtained with four different invariant-mass windows, from a minimum of  $1\sigma$  to a maximum of  $3\sigma$ , as shown in the left plot of figure 5.17 for

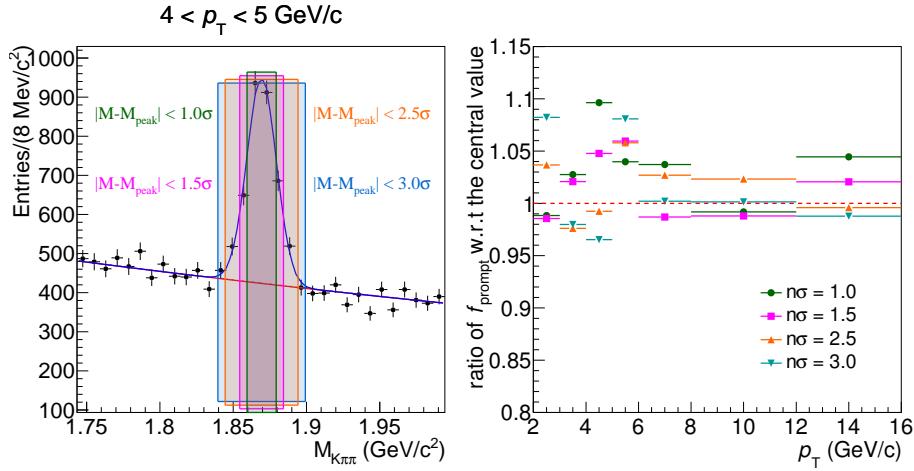


Figure 5.17. Left: four invariant-mass windows tested for the stability checks in the  $p_T$  interval  $4 < p_T < 5$  GeV/c. Right: ratio of  $f_{prompt}$  with respect to the reference value obtained with  $|M - M_{peak}| < 2\sigma$  as a function of  $p_T$ .

the  $p_T$  interval  $4 < p_T < 5 \text{ GeV}/c$ . The estimation of the systematic effect was performed computing the ratio between each value of  $f_{prompt}$  and the reference value obtained using the mass window  $|M - M_{peak}| < 2\sigma$ . The right plot of figure 5.17 shows that this ratio fluctuates around unity without any clear trend, therefore no systematic uncertainty has been assigned based on this test.

### Monte Carlo closure test

In this section I will present the Monte Carlo closure test, which was performed to verify the consistency and stability of the fitting procedure using a fast simulation tool. The aim of this study is to check if the fits can reproduce well the input values of  $f_{prompt}$ , to be sure not to have biases in the measurements. The closure test is divided in two steps:

- Simulation of the entries in the tree in the same  $p_T$  intervals of the analysis, starting from the MC impact-parameter distributions for prompt and feed-down  $D^+$  and from the distributions obtained from the sideband region for the background, which are the same used in the prefit phase.
- Unbinned fit of the simulated impact-parameter distributions, applying the procedure described in section 5.3.

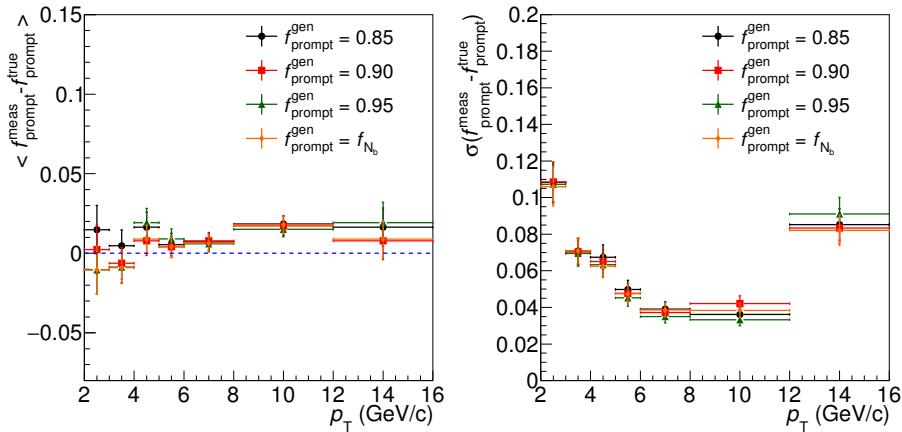


Figure 5.18. Mean (left) and RMS (right) of the residual distributions for the  $D^+$  prompt fraction evaluated with the MC closure test, for four different input values of  $f_{prompt}$ .

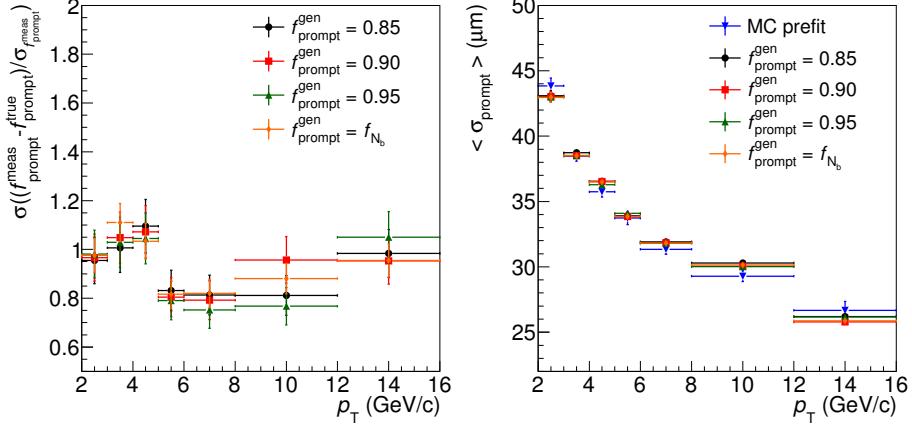


Figure 5.19. Left: pulls of  $f_{prompt}$  distributions (left) and mean value of  $\sigma_{prompt}$  (right) evaluated with the MC closure test, for four different input values of  $f_{prompt}$ .

The number of entries for each contribution is given by three input parameters ( $N$ ,  $S$  and  $f_{prompt}^{gen}$ ), as follow:

- $N_{bkg} = N - S$ , the number of entries extracted from the background distribution.
- $N_{prompt} = f_{prompt}^{gen} \times S$ , the number of entries extracted from the prompt  $D^+$  distribution.
- $N_{feed-down} = (1 - f_{prompt}^{gen}) \times S$ , the number of entries extracted from the feed-down  $D^+$  distribution.

The  $N$  and  $S$  parameters were fixed to the values in the data, in order to simulate the same statistics that is available for the analysis, while for the input prompt fraction three different values have been tested:  $f_{prompt}^{gen} = 0.85, 0.90, 0.95$  and, in addition, a fourth test was done fixing  $f_{prompt}^{gen}$  to the values obtained with the theory-driven  $N_b$  method (see section 4.2.1). The resulting distributions were then fitted following the same procedure and using the same fit functions (see equation 5.11) with  $f_{prompt}$  and the detector-resolution term  $\sigma_{prompt}$  as free parameters. The simulation and fitting procedure were repeated 50 times. In figure 5.18 the mean value (left) and the RMS (right) of the residual distributions for  $f_{prompt}$  are shown as a function of the transverse momentum, for all the four different input values of the prompt fraction. The mean value is compatible with zero within the

statistical uncertainty in the almost the full  $p_T$  range, which means that the recovered value of  $f_{prompt}$  reproduces reasonably well the value set as input. However, the mean values of the difference seem to be systematically slightly higher than zero. Therefore a 1% systematic uncertainty is assigned to the measurement, as reported in table 5.2. In order to test whether the error on the single measurement is well estimated, the pull distributions are also checked in the  $p_T$  intervals of the analysis. The pull is defined as the Gaussian width of the distribution of the difference between the  $D^+$  prompt fraction obtained from the fit and the one set as input, divided by the error on the measured prompt fraction:

$$Pull(p_T) = \sigma \left( \frac{f_{prompt}^{meas} - f_{prompt}^{gen}}{\sigma_{f_{prompt}^{meas}}} \right) (p_T). \quad (5.13)$$

Since the difference between the input and measured  $f_{prompt}$  should be covered by the statistical errors, the pull distribution is expected to be a normal Gaussian distribution with zero mean and width equal to unity. As shown in the left panel of figure 5.19 the pulls obtained stay between 0.8 and 1.1, indicating that the statistical error is reasonably well estimated. Finally, the right panel of the same figure shows the mean of  $\sigma_{prompt}$  values obtained for each input prompt fraction tested, compared with the value obtained from the MC prefit phase. Since the fitted distributions are simulated starting from the MC templates for prompt and feed-down  $D^+$ -mesons, the two values of this parameter should be compatible, indeed, they result to be in agreement within the uncertainties.

$p_T$ (GeV/c)	[2,3]	[3,4]	[4,5]	[5,6]	[6,8]	[8,12]	[12,16]
$\sigma_{prompt}$	negl.	negl.	negl.	negl.	negl.	negl.	negl.
Fit range	2%	2%	2%	2%	2%	2%	2%
Bkg fit function	4%	1%	1%	1%	1%	1%	1%
Side-band regions	6%	3%	3%	3%	3%	3%	3%
$p_T$ shape	negl.	negl.	negl.	negl.	negl.	negl.	negl.
$S$ parameter	4%	2%	2%	2%	2%	2%	2%
Mass window	negl.	negl.	negl.	negl.	negl.	negl.	negl.
MC closure test	1%	1%	1%	1%	1%	1%	1%
total syst.	9%	4%	4%	4%	4%	4%	4%

Table 5.2. Summary of the systematic uncertainty for the data-driven impact-parameter analysis for the evaluation of the fraction of prompt  $D^+$ -mesons.

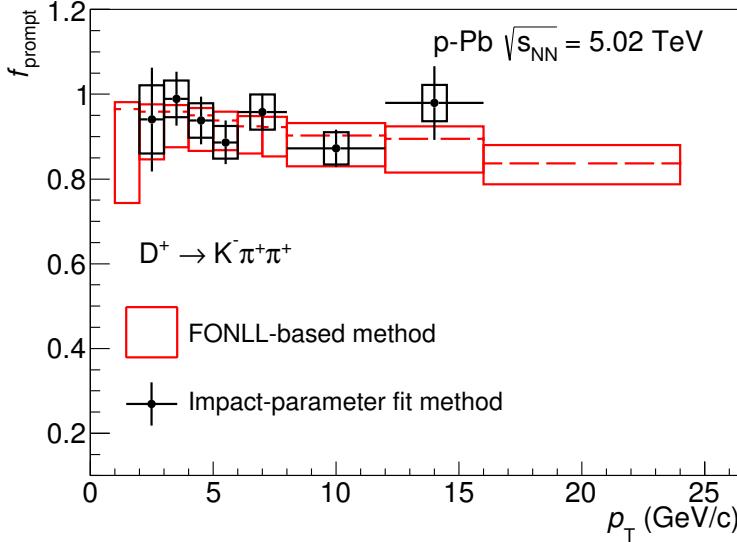


Figure 5.20.  $D^+$  prompt fraction obtained with the impact-parameter fit method, compared to the FONLL-based method ( $N_b$  and  $f_c$  combined). The vertical bars represent the statistical uncertainty, while the boxes the systematic one.

#### 5.4.5 Summary of the systematic uncertainties

The total systematic uncertainty was evaluated by considering all the tests described above. The systematic uncertainty due to the assumed shapes of the impact-parameter distributions was found to be the dominant source, especially considering the tests performed on the impact-parameter distributions of the background. To compute the total systematic uncertainty, all the variations observed performing the tests were treated as uncorrelated and therefore summed in quadrature, as reported in table 5.2.

### 5.5 Results

The results for  $f_{prompt}$  obtained with the impact-parameter fit method in the seven  $p_T$  intervals are found to be compatible with the one obtained with the FONLL-based method within the uncertainties in the full  $p_T$  range, as shown in figure 5.20. The vertical bars represent the statistical error, while the boxes the systematic one. In some  $p_T$  bins the upper edge of the error bars exceeds unity,

because the statistical uncertainty obtained by the fitting procedure is symmetric with respect to the central value, which lies generally very close to  $f_{prompt}=1$ . This is a consequence of the fitting procedure used to estimate the relative amount of the three components (prompt D<sup>+</sup>, feed-down D<sup>+</sup> and background). Note that, the fits were performed without limiting  $f_{prompt}$  at 1, since a bias in the result was observed with the fast simulation tool if this limit was set. Clearly, a fraction of prompt D<sup>+</sup> in the raw yield larger than unity does not have a physical meaning. One option for representing graphically the result could have been to sum in quadrature the statistical and systematic uncertainties and draw the points with a total uncertainty bar, which on the upper side could be truncated at one. The reason why it was chosen not to represent the total uncertainty in the picture, was to provide information about the different contribution of the statistical and systematic uncertainty to the  $f_{prompt}$  uncertainty. It is indeed possible to notice that the statistical uncertainty is larger than the systematic one in each  $p_T$  interval, which means that this analysis will greatly benefit starting in the LHC Run II, which has already started in 2015, because of the larger data samples expected. Moreover, since  $f_{prompt}$  depends on the topological cuts applied for the selection of the D<sup>+</sup> candidates, it is not a physical observable in itself, but it is an intermediate quantity needed to obtain the  $p_T$ -differential production cross section of prompt D<sup>+</sup>-mesons. The total uncertainty reported in table 5.3 was computed treating the statistical and systematic uncertainties as uncorrelated and therefore summed in quadrature. The fraction of prompt D<sup>+</sup> evaluated by fitting the impact-parameter distributions has similar uncertainties in the intermediate  $p_T$  intervals ( $3 < p_T < 12$  GeV/c) and larger uncertainties in the first and in the last analysed  $p_T$  intervals, with respect to the theory-driven method. Even if this method does not allow to improve significantly the measurement with the current statistics, it provides an important confirmation of the result obtained with the FONLL-based method, which depends on theoretical pQCD calculations and the assumption on the  $R_{pPb}$  of feed-down D<sup>+</sup>-meson. Furthermore, the  $f_{prompt}$  measurement with this method will improve in the future, because of the larger data samples expected in the Run2

$p_T$ (GeV/c)	[2,3]	[3,4]	[4,5]	[5,6]	[6,7]	[7,8]	[8,12]	[12,16]
Theory-driven (%)	+2 -12	+2 -9	+2 -9	+2 -8	+3 -7	+3 -7	+3 -8	+3 -9
Imp. par. fit (%)	$\pm 16$	$\pm 7$	$\pm 7$	$\pm 7$	$\pm 6$		$\pm 6$	$\pm 10$

Table 5.3. Comparison between the total uncertainty on the D<sup>+</sup> prompt fraction evaluated with the impact-parameter fit method and the FONLL-based method.

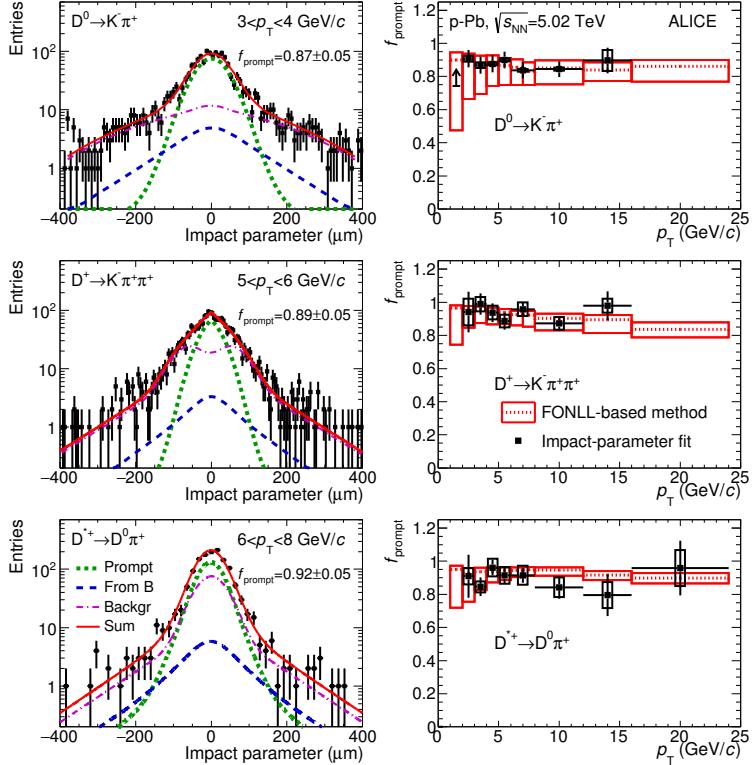


Figure 5.21. Examples of impact-parameter fits (left column) and resulting prompt fraction as a function of  $p_T$  (right column) for the three different D-meson species analysed:  $D^0$  (top),  $D^+$  (middle),  $D^{*+}$  (bottom) [1]. The arrow in the interval  $1 < p_T < 2$  GeV/c for the  $D^0$ -meson represents the minimum value within a 95% confidence level.

and even more with the upgrade of the Inner Tracking System planned for 2019, as will be briefly discussed in the next section. The results of this work have been published as a cross-check of the standard approach by the ALICE Collaboration [1]. Figure 5.21, taken from [1], shows an example of fit (left column) and the resulting  $f_{prompt}$  as a function of  $p_T$  (right column) for the three D-meson species analysed:  $D^0$  (top),  $D^+$  (middle) and  $D^{*+}$  (bottom). The prompt fraction evaluated with the impact-parameter fit method was found to be compatible with the FONLL-based estimation within the uncertainties for the other D-meson species as well. In particular, for the  $D^0$  the data-driven approach provides a more precise

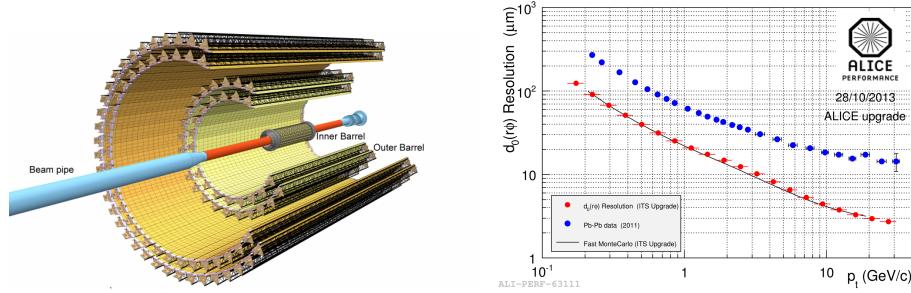


Figure 5.22. Left: Layout of the upgraded ALICE ITS. Right: improvement of the impact-parameter resolution in the transverse plane expected with the upgraded ITS [66].

estimation of  $f_{prompt}$  in the transverse momentum range  $2 < p_T < 12$  GeV/c as compared to the FONLL-based approach, while for  $1 < p_T < 2$  GeV/c only a lower limit at a 95% confidence level could be estimated, given the poor accuracy of the impact-parameter fit.

## 5.6 Perspective of an improved measurement with the upgraded ITS

The ALICE Collaboration is preparing a major upgrade of the experimental apparatus, planned for installation in the second long LHC shutdown in 2019, which will also involve the Inner Tracking System. The main goals of the upgrade of the ALICE ITS are an improved reconstruction of primary and secondary vertices and an improved performance in the tracking of low-momentum particles. In order to achieve these targets, the following changes will be applied:

- Replace the six ITS layers of the current setup with seven concentric barrel layers of silicon pixel detectors with radii between 22 mm for the innermost and 400 mm for the outermost layer, as shown in the left panel of figure 5.22.
- Reduce the material budget down to 0.3%  $X_0$  in the innermost and  $\sim 0.9\%$   $X_0$  in the outer and middle layers compared to 1.14%  $X_0$  per layer in the current ALICE Silicon Pixel Detector.
- Use sensors with a pixel size of  $30 \mu\text{m} \times 30 \mu\text{m}$ , compared to  $50 \mu\text{m} \times 425 \mu\text{m}$  in the current pixel detector.

The central part of the upgraded ALICE ITS is the pixel chip. The most challenging requirements for the detector design are the high spatial resolution and the

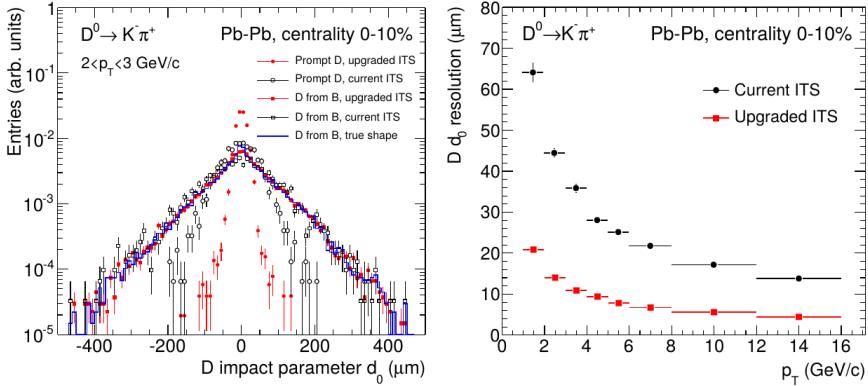


Figure 5.23. Left: simulated impact-parameter distributions for prompt and feed-down  $D^0$  obtained with the current and upgraded ITS configurations in the interval  $2 < p_T < 3 \text{ GeV}/c$  [67]. Right:  $D^0$  impact-parameter resolution as a function of  $p_T$

extremely low material budget, in particular for the inner layers. These requirements, together with the modest radiation environment compared to the other LHC trackers, led to the choice of Monolithic Active Pixel Sensors (MAPS) [65]. As already mentioned, the impact-parameter resolution will greatly benefit from this upgrade. The right panel of figure 5.22 shows the comparison between the current impact-parameter resolution in the  $xy$ -plane (blue points) and the expected one (red points) with the upgraded ITS, obtained with a MC simulation, in the 10% most central Pb-Pb collisions [66]. The improvement of the impact-parameter resolution in the  $xy$ -plane, which will reach a factor  $\sim 3$ , will allow to achieve a better separation between the prompt and feed-down  $D^+$  contributions, which, together with the higher statistics expected, will result in a more precise estimation of the prompt fraction with the impact-parameter fit method. Tests of the performances of the upgraded ITS for the measurement of the prompt fraction with the impact-parameter fit method have been carried out for  $D^0$  mesons [67]. In the left panel of figure 5.23 the simulated impact-parameter distributions for prompt and feed-down  $D^0$  with  $2 < p_T < 3 \text{ GeV}/c$ , obtained with the current and upgraded ITS configurations, are compared. In the right panel of the same figure the comparison between the impact-parameter resolution of  $D^0$  with the two different configurations is shown; as already mentioned, with the upgraded ITS the resolution will be improved up to a factor 3.

# 6

## Data-driven cut-variation method for the $D^+$ measurement

In this Chapter I will present an alternative data-driven method for the measurement of the  $p_T$ -differential production cross section of prompt and feed-down  $D^+$ -mesons in the decay channel  $D^+ \rightarrow K^-\pi^+\pi^+$ , which is based on the different decay chain of the two types of  $D^+$ . Indeed, for feed-down  $D^+$  there are two decay vertices, one of the  $D^+$  and the other of the B-hadron from which the  $D^+$  originate. The presence of such vertices implies a difference in the distributions of some reconstructed variables between the prompt and feed-down  $D^+$ -mesons. As the impact-parameter fit method presented in Chapter 5, this approach has the advantage of not depending on theoretical calculations and on any assumption on the feed-down  $D^+$ -meson  $R_{pPb}$ . The first section of this Chapter is dedicated to the description of the method and the strategy adopted, while in the second part the analysis of the Minimum Bias p-Pb data sample described in 4.1.1 is discussed. Finally, in the last section of this Chapter, the results obtained are compared with the standard analysis.

### 6.1 Cut-variation feed-down subtraction method

As explained in section 4, the  $D^+$ -meson candidates reconstructed in the decay channel  $D^+ \rightarrow K^-\pi^+\pi^+$  include prompt  $D^+$ , feed-down  $D^+$  and the combinatorial

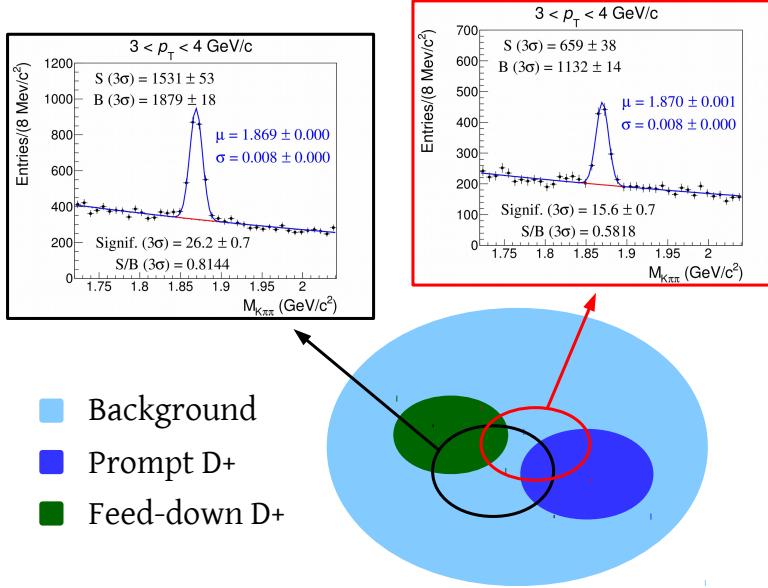


Figure 6.1. Schematic representation of the selection of the  $D^+$  candidates for two different sets of topological cuts.

background. Due to the different efficiency of the applied selections for prompt and feed-down candidates, for each set of topological cuts, a different amount of prompt and feed-down  $D^+$  is selected, as illustrated in figure 6.1. This sketch shows how the extracted raw yield in the same  $p_T$  interval changes for two different sets of topological cuts. Hence we can define a equation, which relates the extracted raw yield  $Y_i$  and the reconstruction efficiencies for prompt  $D^+$  ( $\epsilon_i^{prompt}$ ) and feed-down  $D^+$  ( $\epsilon_i^{feed-down}$ ), for each set of topological cuts  $i$ :

$$\epsilon_i^{prompt} \cdot N_{prompt} + \epsilon_i^{feed-down} \cdot N_{feed-down} = Y_i, \quad (6.1)$$

where  $N_{prompt}$  and  $N_{feed-down}$  are the corrected yields of prompt and feed-down  $D^+$  respectively, which we want to determine in order to compute their production cross sections. However, it is not possible to obtain the corrected yields from a single equation (i.e. a single set of cuts): at least two are needed. The basic idea of the cut-variation method is therefore to define  $n$  different sets of topological cuts

in order to obtain a system of  $n$  equations

$$\begin{cases} \epsilon_1^{prompt} \cdot N_{prompt} + \epsilon_1^{feed-down} \cdot N_{feed-down} = Y_1 \\ .. \\ .. \\ \epsilon_n^{prompt} \cdot N_{prompt} + \epsilon_n^{feed-down} \cdot N_{feed-down} = Y_n \end{cases} \quad (6.2)$$

and determine  $N_{prompt}$  and  $N_{feed-down}$ . The system can be exactly solved in the case of two equations, otherwise it is likely to be overdetermined. Furthermore, since the extracted raw yields and the efficiencies are affected by statistical and systematic uncertainties, a number of sets of topological cuts larger than two would lead to a better estimation of the corrected yields. For this reason, a minimisation procedure is needed in order to obtain an approximated solution of the system. In particular for this analysis two different approaches have been used.

### 6.1.1 Analytic minimisation

The first method used in the analysis to get an approximated solution of the system 6.2 is based on the minimisation of the  $\chi^2$ . In order to give a general description of this method, the matrix formalism for  $n$  generic sets of topological cuts, corresponding to  $n$  equations involving the measured raw yields and the efficiencies from simulations, will be used. The quantities involved in the minimisation are the following:

- The column vector of raw yields  $Y_i$

$$\mathbf{Y} = \begin{pmatrix} Y_1 \\ Y_2 \\ .. \\ Y_n \end{pmatrix}. \quad (6.3)$$

- The efficiency matrix for prompt and feed-down  $D^+$

$$\boldsymbol{\epsilon} = \begin{pmatrix} \epsilon_1^{prompt} & \epsilon_1^{feed-down} \\ \epsilon_2^{prompt} & \epsilon_2^{feed-down} \\ .. & .. \\ \epsilon_n^{prompt} & \epsilon_n^{feed-down} \end{pmatrix}. \quad (6.4)$$

- The corrected yields vector

$$\mathbf{N} = \begin{pmatrix} N_{prompt} \\ N_{feed-down} \end{pmatrix}. \quad (6.5)$$

Hence, the system of equation 6.2, can be rewritten in a more compact way:

$$\boldsymbol{\epsilon} \times \mathbf{N} - \mathbf{Y} = \mathbf{0}, \quad (6.6)$$

where  $\mathbf{0}$  is the  $n$ -dimensional null vector. As already mentioned, for  $n > 2$  the system is likely to be overdetermined, and therefore an additional quantity  $\boldsymbol{\delta}$  is needed. In particular,  $\boldsymbol{\delta}$  is the *residual vector*, which coincides with the null vector  $\mathbf{0}$  in case of a determined system, while is different otherwise. Therefore, the latter equation becomes

$$\begin{pmatrix} \epsilon_1^{prompt} & \epsilon_1^{feed-down} \\ \epsilon_2^{prompt} & \epsilon_2^{feed-down} \\ .. & .. \\ \epsilon_n^{prompt} & \epsilon_n^{feed-down} \end{pmatrix} \times \begin{pmatrix} N_{prompt} \\ N_{feed-down} \end{pmatrix} - \begin{pmatrix} Y_1 \\ Y_2 \\ .. \\ Y_n \end{pmatrix} = \begin{pmatrix} \delta_1 \\ \delta_2 \\ .. \\ \delta_n \end{pmatrix}. \quad (6.7)$$

The best estimation of the corrected yields is obtained by minimising the  $\chi^2$ , defined as

$$\chi^2 = \boldsymbol{\delta}^T \mathbf{C}^{-1} \boldsymbol{\delta}, \quad (6.8)$$

where  $\boldsymbol{\delta}^T$  is the row vector of residuals and  $\mathbf{C}$  the weighting matrix

$$\mathbf{C} = \begin{pmatrix} \sigma_1^2 & & & \\ & \sigma_2^2 & & \\ & & \ddots & \\ & & & \sigma_n^2 \end{pmatrix}, \quad (6.9)$$

which is defined from the uncertainties inherent to each equation and assumed to be diagonal. This is clearly an approximation, however the correlation among the  $n$  sets of topological cuts is hard to quantify. For this reason only three as much as possible different sets of cuts will be considered for the analysis, as will be described in section 6.2. The weights  $\sigma_i^2$  are calculated starting from the statistical errors on the raw yields and the efficiencies, according to the following relation:

$$\sigma_i^2 = \sigma_{Y_i}^2 + N_{prompt} \cdot \sigma_{\epsilon_i^{prompt}}^2 + N_{feed-down} \cdot \sigma_{\epsilon_i^{feed-down}}^2. \quad (6.10)$$

Since the corrected yields are not yet determined at this stage, the weights are evaluated interactively: in the first step  $N_{prompt}$  and  $N_{feed-down}$  are set to zero and therefore only the uncertainty on the raw yields is taken into account, while from the second iteration the weights are computed using the value of the corrected yields obtained in the previous step. Only few iterations are needed to obtain a stable result, since the statistical error on the efficiencies is typically much smaller

than the one on the raw yield. Then, the minimisation of the  $\chi^2$  with respect to the corrected yields leads to

$$\mathbf{N} = \mathbf{Cov}(\mathbf{N})\boldsymbol{\epsilon}^T \mathbf{C}^{-1} \mathbf{Y}, \quad (6.11)$$

where  $\mathbf{Cov}(\mathbf{N}) = (\boldsymbol{\epsilon}^T \mathbf{C}^{-1} \boldsymbol{\epsilon})^{-1}$  is the covariance matrix which contains the statistical uncertainties on the corrected yields.

### 6.1.2 Incentre minimisation

Besides the analytic minimisation, another method based on a geometrical approach has been used for the analysis. Considering equation 6.1, each set of cuts can be drawn as a line in the  $(N_{prompt}, N_{feed-down})$  plane:

$$N_{prompt} = -\frac{\epsilon_i^{feed-down}}{\epsilon_i^{prompt}} \cdot N_{feed-down} + \frac{Y_i}{\epsilon_i^{prompt}}. \quad (6.12)$$

The slope of the line is given by the ratio between the feed-down and prompt D<sup>+</sup> efficiencies, while the constant term is given by the ratio of the raw yield and the efficiency of prompt D<sup>+</sup>. Considering now three different sets of cuts, the three corresponding lines should intersect in one single point, whose coordinates correspond to the corrected yields. However, because of the experimental uncertainties on the raw yield extraction and on the efficiency determination, the intersections

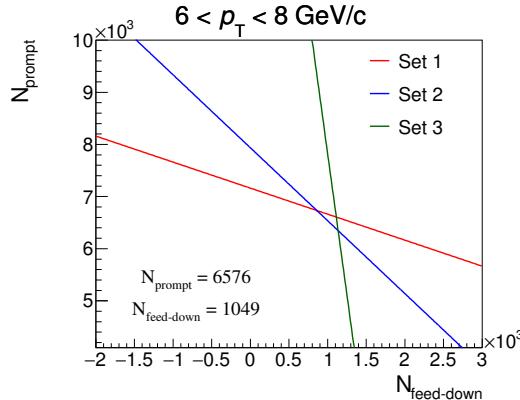


Figure 6.2. Three lines corresponding to different sets of topological cuts in the interval  $6 < p_T < 8$  GeV/c. Their intersections define a triangle, whose incentre is used to evaluate the corrected yields of prompt and feed-down D<sup>+</sup>-mesons.

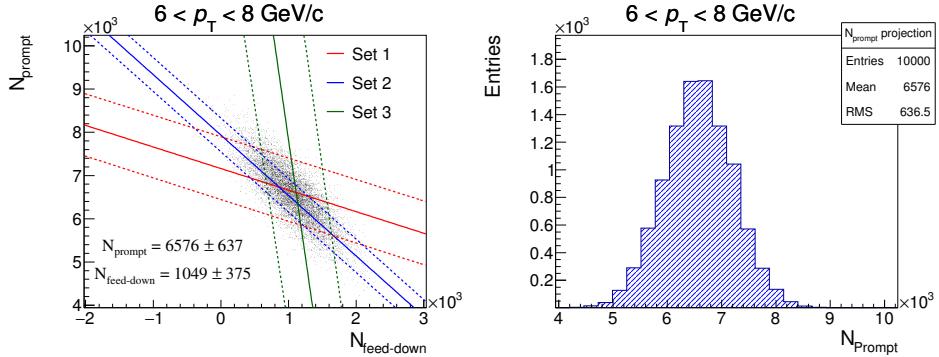


Figure 6.3. Left: distribution of the generated incentres with the toy MC simulation superimposed to the lines corresponding to three different sets of cuts in the interval  $6 < p_T < 8 \text{ GeV}/c$ . The dotted lines represent the maximum effect of the uncertainty on the line parameters within  $1\sigma$ . Right: projection of the distribution of the generated incentres on the  $N_{\text{prompt}}$ -axis.

of all the possible pairs of lines will not coincide among each other. In particular, in the case of three sets of cuts, the number of intersections becomes equal three and the intersection points define a triangle. The corrected yields are then defined as the coordinates of the incentre of the triangle, which is the crossing point of the internal angle bisectors and is thus the point equidistant from all the triangle's sides.

### Estimation of the statistical uncertainty on the incentre coordinates

The uncertainty on the slope of the lines depends on the uncertainty on the efficiencies for prompt and feed-down  $D^+$ -mesons, while the uncertainty on the constant term, which implies a shift of the line, is dominated by the raw yield uncertainty, which is typically larger than the one on the efficiency of prompt  $D^+$ . The statistical uncertainty on the coordinates of the incentre, which represent the corrected yields of prompt and feed-down  $D^+$ , is propagated from the errors on the line parameters with a toy Monte Carlo simulation. It consists in generating several times the parameters of the lines sampling them from a Gaussian distribution, with mean equal to the value of the considered parameter and sigma equal to its uncertainty. For each generation, the coordinates of the incentre are computed and a two-dimensional histogram is filled. In the left panel of figure 6.3 an example of incentre distribution obtained with the toy MC simulation in the transverse momentum interval  $6 < p_T < 8 \text{ GeV}/c$  is shown together with the lines used for the incentre determination, which correspond to three different sets of topological cuts.

The dotted lines are obtained setting the line parameters to their value plus or minus their uncertainties, in order to give a visual representation of the uncertainty on the determination of the lines. Note that the dominant uncertainty is the one on the constant term of the lines, as confirmed by the fact that the dotted lines are almost parallel to the centred ones. Finally, the uncertainties on the incentre coordinates are evaluated as the RMS of the one-dimensional distributions, obtained projecting the two-dimensional distribution of the incentres on the corresponding axis. In the right panel of figure 6.3 the projection on the  $N_{prompt}$ -axis of the distribution of the incentre coordinates, which was used to assign the statistical uncertainty on the corrected yield for prompt  $D^+$ , is shown.

## 6.2 Strategy for the selection of topological cuts

As described in the previous section, the weighting matrix used for the analytic minimisation is assumed to be diagonal (see equation 6.9) and therefore possible correlations among the sets of topological cuts are neglected. Furthermore, considering the incentre method, if two or more sets are similar (and therefore correlated), the corresponding lines would be almost parallel, resulting in a poor precision of the incentre calculation, because of the large overlap of the regions defined by the dotted lines in figure 6.3. For this reason, the sets of topological cuts have to be chosen in order to generate sub-samples of  $D^+$  candidates as much as possible disjoint. Moreover, a large number of sets would produce several sub-samples of  $D^+$  candidates, which would have a small amount of entries because of the necessity to keep all the sets as much as possible uncorrelated. Hence, only three cut sets were defined for the analysis:

- **Enhanced prompt contribution:** this first set of topological cuts is defined maximising the prompt  $D^+$  contribution. Since the efficiency is typically higher for feed-down  $D^+$ , the aim is to achieve a similar or slightly higher efficiency for the prompt contribution with respect to the feed-down one.
- **Mixed:** the goal of this set of cuts is to maximise the amount of signal extracted from the invariant-mass distributions. For this reason, this set is similar to the one defined for the standard analysis, which was optimised to obtain a large signal-to-background ratio and a large significance.
- **Enhanced feed-down contribution:** the third set of cuts aims to maximise the feed-down  $D^+$  contribution. Even if the maximisation of the efficiency for feed-down  $D^+$  can be easily achieved, the complication for this set of topological cuts is to obtain a reasonable yield extraction, since the feed-down  $D^+$  are the minority of the candidates.

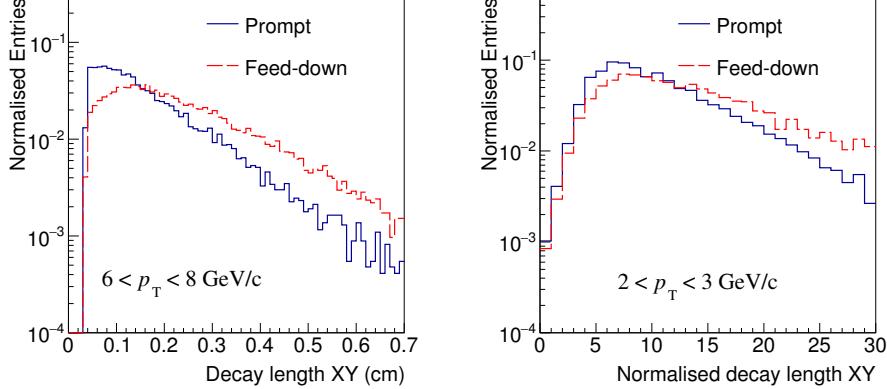


Figure 6.4. Comparison between the prompt and feed-down  $D^+$  distributions from MC simulations of the decay length in the  $xy$ -plane (left) in the interval  $6 < p_T < 8 \text{ GeV}/c$  and the normalised decay length in the  $xy$ -plane (right) for  $2 < p_T < 3 \text{ GeV}/c$ .

### 6.2.1 Topological variables

In order to obtain the three different cut sets described above, the variables used in the selection must have a significant separation power between the prompt and the feed-down contributions. The quantities described in the following paragraphs are used.

#### Decay length in the $xy$ -plane

As already discussed in section 4.1.4, the decay length and the normalised decay length can discriminate between the signal and background candidates (see right panel of figure 4.3). However, this quantity can also be used to separate the prompt from the feed-down contribution, since the  $D^+$ -mesons which originate from B-hadron decays are on average more displaced from the interaction point. Furthermore, the decay length in the transverse plane would allow us to achieve a better separation, because of the better resolution with respect to the  $z$ -coordinate. An example of this propriety is illustrated in figure 6.4, where the comparison between the distributions of prompt and feed-down  $D^+$  from MC simulations is shown for the decay length in the  $xy$ -plane in  $6 < p_T < 8 \text{ GeV}/c$  (left) and for the normalised decay length in the  $xy$ -plane in the interval  $2 < p_T < 3 \text{ GeV}/c$  (right).

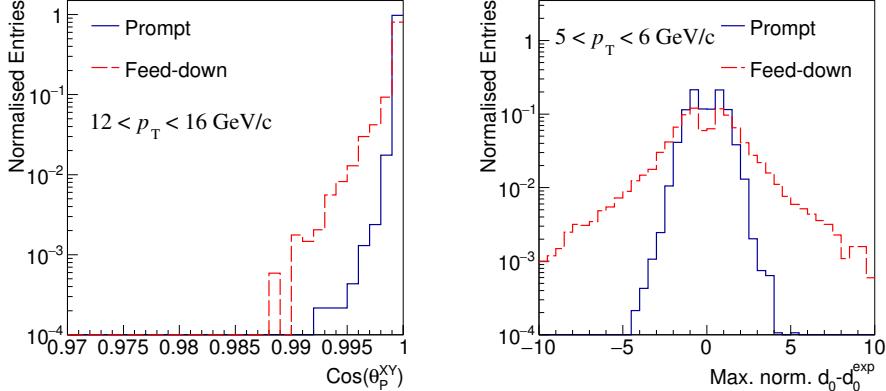


Figure 6.5. Comparison between the distributions of the cosine of the pointing angle in the  $xy$ -plane in the interval  $12 < p_T < 16 \text{ GeV}/c$  (left) and the maximum standardised difference between measured and expected track impact parameter (right) in the interval  $5 < p_T < 6 \text{ GeV}/c$  for prompt and feed-down D<sup>+</sup> from MC simulations.

### Cosine of the pointing angle in the $xy$ -plane

The momentum vector of the feed-down D<sup>+</sup>-mesons in general is not parallel to the line which connects the primary vertex and the D<sup>+</sup> decay vertex; their relative orientation depends on the direction of the D<sup>+</sup>-meson with respect to that of the parent B-hadron. As a result, the distribution of the cosine of the pointing angle for feed-down D<sup>+</sup> is less peaked at 1 than the one of prompt D<sup>+</sup>. In the left panel of figure 6.5 the distributions of  $\cos(\theta_P^{xy})$  for prompt and feed-down D<sup>+</sup> in the transverse momentum interval  $12 < p_T < 16 \text{ GeV}/c$  from MC simulations are compared. In order to achieve a better separation of the two contributions, only the  $xy$ -projection of the cosine of the pointing angle are exploited, because of the better resolution with respect to the  $z$ -coordinate.

### Normalised residual of the daughter impact parameter

Another quantity which can discriminate the prompt from the feed-down contribution is the variable defined as the difference between the measured and the expected impact parameter of the daughter tracks, divided by its uncertainty. In particular,

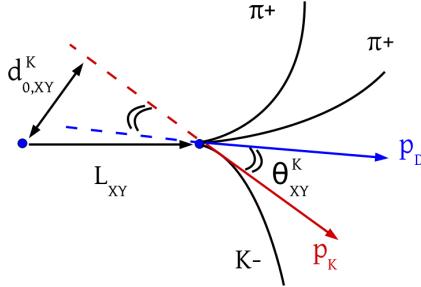


Figure 6.6. Sketch of the  $D^+$  decay topology with the quantities used to compute the expected daughter track impact parameter.

the maximum of the three daughter tracks is used as selection variable:

$$\max n\sigma_{res} = \max \left\{ \frac{d_{0,xy}^{exp} - d_{0,xy}^{meas}}{\sqrt{\sigma_{d_{0,XY}^{exp}}^2 + \sigma_{d_{0,XY}^{meas}}^2}} \right\}, \quad (6.13)$$

where  $d_{0,xy}^{meas}$  is the measured value of the daughter impact parameter in the  $xy$ -plane and  $\sigma_{d_{0,XY}^{meas}}$  its error. The expected value of the impact parameter for each daughter is computed as

$$d_{0,xy}^{exp} \simeq L_{xy} \sin(\theta_{xy}^{tr}), \quad (6.14)$$

where  $L_{xy}$  is the  $D^+$  decay length, i.e. the distance between the primary and secondary vertices, and  $\theta_{xy}^{tr}$  is the angle between the reconstructed momenta of the considered daughter and the  $D^+$ -meson in the  $xy$ -plane, as shown in the sketch in figure 6.6. The uncertainty on  $d_{0,xy}^{exp}$  is calculated neglecting the error on  $\sin(\theta_{xy}^{tr})$ :

$$\sigma_{d_{0,xy}^{exp}} = \sigma_{L_{xy}} \sin(\theta_{xy}^{tr}). \quad (6.15)$$

This variable provides a significant separation between the two contributions, since it is on average larger for feed-down  $D^+$ , as shown in the right panel of figure 6.5, where the distributions of prompt and feed-down  $D^+$  from MC simulations are compared for the  $p_T$  interval  $5 < p_T < 6$  GeV/c.

### 6.2.2 Sets of topological cuts

For all the sets of cuts, in order to remove as much as possible the combinatorial background, some common selections are applied. These cuts are the same as those in the standard analysis, which are summarised in table 4.2. Furthermore,

all the event, track quality and PID selections described in chapter 4 are used. The differences concern the cuts applied on the variables listed above, which are different from those applied in the standard analysis, in order to obtain the three sets of topological cuts described in section 6.2. The values of the cuts used for this purpose are reported in table 6.1.

For the first set an upper cut to the decay length and the normalised decay length in the transverse plane has been applied to reduce the feed-down contribution. However, in order to limit the background, also a lower cut has been applied to these variables. Furthermore, a lower cut to the cosine of the pointing angle in the  $xy$ -plane and a symmetric cut with respect to zero on the maximum normalised difference between the measured and expected track impact parameter, has been applied to remove the tails, where the feed-down contribution is dominant.

The second set of topological cuts, as already mentioned, was chosen similar to the one of the standard analysis, with lower cuts on the decay length, the normalised decay length and the cosine of the pointing angle in the transverse plane to reject the combinatorial background.

For the third set, tighter cuts on the minimum decay length and normalised decay length have been applied to reduce the prompt contribution. For the same purpose, only the tail of the distribution of the maximum normalised difference between the measured and expected track impact parameter was taken into account. In particular, only the left side was kept for technical reasons. Finally, the cut on the cosine of the pointing angle was kept at a not too tight value, aimed at rejecting the background. An upper cut on this variable to reduce the prompt contribution was not applied because it implies a significant loss of signal, which does not allow

	$p_T$ (GeV/c)	[2,4]	[4,6]	[6,8]	[8,12]	[12,16]
Set 1	$L_{xy}$ (cm)	[0.02,0.14]	[0.02,0.15]	[0.02,0.15]	[0.04,0.15]	[0.04,0.25]
	Norm. $L_{xy}$	[7,14]	[4,16]	[4,16]	[4,16]	[4,22]
	$Cos(\theta_P^{xy})$	>0.998	>0.998	>0.998	>0.998	>0.998
	$max n\sigma_{res}$	[-1.5,1.5]	[-1.5,1.5]	[-1.5,1.5]	[-1.5,1.5]	[-2.0,2.0]
Set 2	$L_{xy}$ (cm)	>0.05	>0.05	>0.05	>0.10	>0.10
	Norm. $L_{XY}$	>9	>9	>9	>6	>6
	$Cos(\theta_P^{xy})$	>0.990	>0.990	>0.990	>0.990	>0.990
Set 3	$L_{xy}$ (cm)	>0.08	>0.12	>0.20	>0.20	>0.20
	Norm. $L_{xy}$	>9	>12	>20	>20	>20
	$Cos(\theta_P^{xy})$	>0.990	>0.990	>0.990	>0.990	>0.990
	$max n\sigma_{res}$	<-1.5	<-1.5	<-1.5	<-1.5	<-1.0

Table 6.1. Summary table of the three sets of topological cuts used for the  $D^+$  cut-variation analysis.

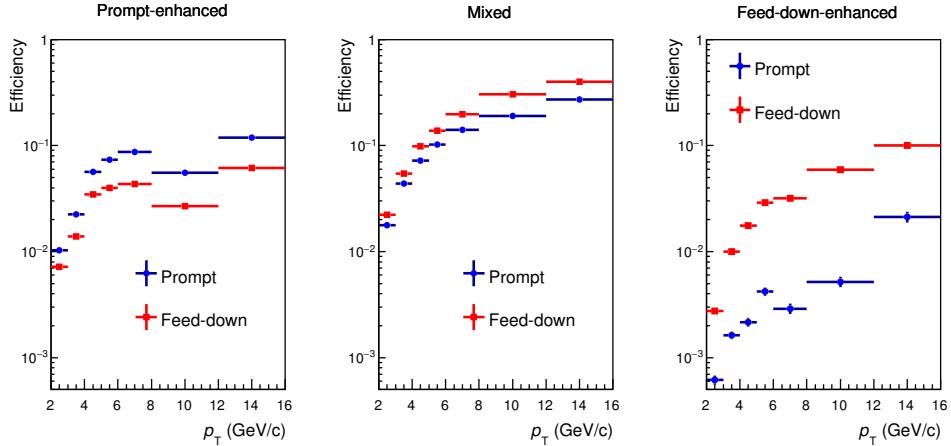


Figure 6.7. Comparison between the efficiencies for prompt and feed-down  $D^+$  for the three sets of topological cuts used in the cut-variation analysis.

to extract the raw yields.

The  $p_T$  range considered for this analysis is  $2 < p_T < 16$  GeV/c. For the first and the last  $p_T$  intervals of the standard analysis was not possible to extract the signal with the three set of cuts, due to the high combinatorial background at low  $p_T$  and the low statistics at high  $p_T$ . The  $p_T$  intervals are the same as those in the impact-parameter fit analysis.

### 6.3 Efficiency

The efficiencies for prompt and feed-down  $D^+$ -mesons are obtained from the MC simulation, as explained in 4.3.1. However an additional consideration must be taken into account, namely the fact that the efficiency is computed in  $p_T$  intervals of finite size and therefore it depends on the generated  $p_T$  distributions of  $D^+$ -mesons and B-hadrons which decay into  $D^+$ . Hence, in order to avoid a bias due to a different  $p_T$  shape of  $D^+$  in data and MC, the  $D^+$ -meson spectrum in the simulation was reweighted to reproduce the FONLL  $p_T$  shape at  $\sqrt{s_{NN}} = 5.02$  TeV. For the feed-down  $D^+$  the weights have been computed in order to match the FONLL  $p_T$  spectrum of the  $D^+$  from B-hadrons decays. The comparison between the Pythia and FONLL  $D^+$ -meson and B-hadron  $p_T$  shapes and their ratio, which was used to reweight the efficiencies, is shown in figure 5.14.

The resulting efficiencies for the prompt-enhanced cut set are shown in the left panel of figure 6.7. The efficiency for prompt  $D^+$  is higher than that for feed-down  $D^+$  in the full  $p_T$  range, by a factor  $\sim 1.5 - 2$ . For this set of topological cuts the values of the efficiency for prompt (feed-down)  $D^+$  increases from about 1% (0.7%) up to a maximum of about 12% (6%), with increasing  $D^+$  transverse momentum. The values of the efficiency for the second set, shown in the middle panel of figure 6.7, are close to the ones of the standard analysis (see section 4.3.1), since the applied selections are similar. In particular, the efficiency for feed-down  $D^+$  varies from  $\sim 2\%$  for  $2 < p_T < 3$  GeV/c up to about 40% in the highest  $p_T$  bin of the analysis. It results to be slightly higher than the prompt  $D^+$  efficiency, which reaches  $\sim 30\%$  for  $12 < p_T < 16$  GeV/c. The efficiencies obtained for the mixed set of cuts are  $\sim 2 - 3$  times higher than those obtained for the prompt-enhanced cut set.

Finally, in the right panel of figure 6.7 the efficiencies for the cut set used to enhance the feed-down contribution are shown. For this set of cuts the prompt contribution is strongly suppressed, as confirmed by its efficiency, which varies from about 0.06% at low  $p_T$  to 2% at high  $p_T$ . The feed-down  $D^+$  efficiency is lower than the one in the second cut set, however it is higher than the one for prompt  $D^+$ -mesons by up to a factor  $\sim 10$ .

## 6.4 Raw yield extraction

The raw yields were obtained by fitting the candidate invariant-mass distributions obtained applying the three different sets of topological cuts summarised in table 6.1, in addition to all the common selections (event, track quality, PID and common topological cuts), as explained in section 6.2.2. The fitting function is the sum of an exponential function which describes the background and a Gaussian function for the signal. In figure 6.8 an example of invariant-mass fits in the interval  $4 < p_T < 5$  GeV/c for the prompt-enhanced set (left), mixed set (middle) and feed-down-enhanced set (right) is shown. The most critical fits are those to the invariant-mass distributions obtained applying the third cut set, used to enhance the feed-down contribution, because of the low statistics. In particular, the statistical fluctuations for this set lead to a very unstable result of the fit in all the  $p_T$  intervals if the Gaussian width is left as a free parameter. For this reason, the Gaussian width was fixed in each set to the values obtained in the simulation with the mixed cut set, which is the most stable, since it was optimised to obtain a good significance. This was justified by comparing the Gaussian width obtained by fitting the invariant-mass spectra in the data to those in the MC simulation. The peak width was found to be compatible within two sigmas with the MC values in all the considered  $p_T$  intervals, except for the bin  $4 < p_T < 5$  GeV/c, as shown in the left panel of figure 6.9. This discrepancy was taken into account in the evaluation of the

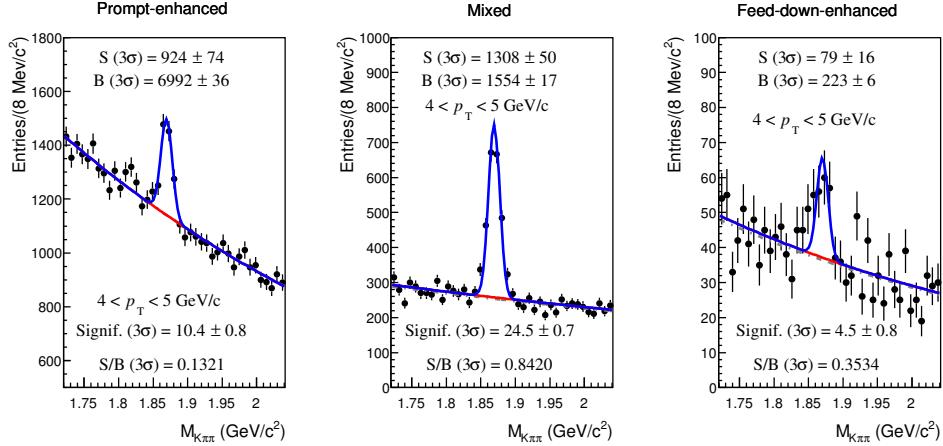


Figure 6.8. Example of invariant-mass fits for the three different sets of topological cuts in the transverse momentum interval  $4 < p_T < 5 \text{ GeV}/c$ .

systematic uncertainties, described in section 6.5.2. Also the peak position was found to be compatible within two sigmas with the MC and with the PDG value of the  $D^+$  mass, except for the first  $p_T$  bin, as shown in the right plot of figure 6.9 for the second set of cuts. For each cut set, the  $D^+$  raw yields are obtained integrating the signal fit function over the full mass range. The extracted raw yields together with the significance for each cut set are reported in table 6.2, while all the invariant-mass fits can be found in the appendix A.5.1.

$p_T$ (GeV/c)	Signal			Significance ( $3\sigma$ )		
	Set 1	Set 2	Set 3	Set 1	Set 2	Set 3
[2,3]	$479 \pm 36$	$897 \pm 43$	$40 \pm 14$	$11.8 \pm 0.8$	$18.7 \pm 0.7$	$2.2 \pm 0.7$
[3,4]	$659 \pm 38$	$1531 \pm 53$	$74 \pm 20$	$15.6 \pm 0.7$	$26.2 \pm 0.7$	$3.1 \pm 0.8$
[4,5]	$924 \pm 74$	$1308 \pm 50$	$79 \pm 16$	$10.4 \pm 0.8$	$24.5 \pm 0.7$	$4.5 \pm 0.8$
[5,6]	$588 \pm 57$	$999 \pm 41$	$68 \pm 14$	$8.7 \pm 0.8$	$21.8 \pm 0.7$	$4.3 \pm 0.8$
[6,8]	$633 \pm 56$	$1126 \pm 44$	$42 \pm 11$	$9.9 \pm 0.8$	$23.3 \pm 0.7$	$4.1 \pm 0.7$
[8,12]	$211 \pm 29$	$787 \pm 29$	$53 \pm 12$	$6.7 \pm 0.8$	$18.6 \pm 0.7$	$4.0 \pm 0.8$
[12,16]	$95 \pm 17$	$242 \pm 24$	$27 \pm 8$	$5.3 \pm 0.8$	$9.6 \pm 0.8$	$2.8 \pm 0.7$

Table 6.2. Summary table of the raw yields and the significance for the three sets of topological cuts.

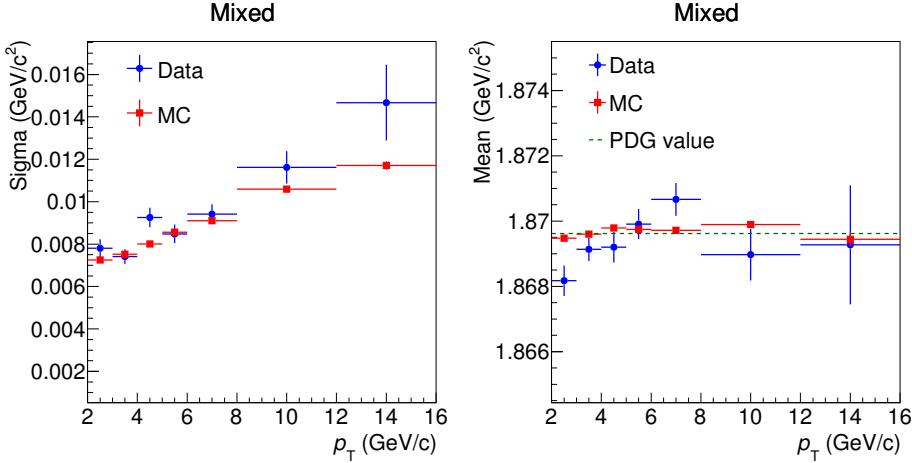


Figure 6.9. Comparison between the Gaussian width (left) and mean (right) obtained by fitting the invariant-mass spectra of the candidates in data and MC for the mixed cut set. The peak position is also compared to the PDG value of the mass of the  $D^+$ -meson.

## 6.5 Systematic uncertainties

In this section, the sources of systematic uncertainties on the determination of the corrected yields with the cut-variation analysis are discussed. The systematics were studied for both the analytic and the incentre minimisation methods, evaluating the influence of each source via the variation of the corrected yields. The three contributions, which have been tested, are:

1. Systematic uncertainty due to the efficiency of the selection criteria for the sets of topological cuts.
2. Systematic uncertainty due to the raw yield extraction.
3. Systematic uncertainty due to the generated  $p_T$  shape of  $D^+$  and B mesons.

Furthermore, four additional sources have been taken into account for the evaluation of the total systematic uncertainty on the cross section measurement:

4. Systematic uncertainty due to the tracking efficiency.
5. Systematic uncertainty due to the PID selection efficiency.
6. Systematic uncertainty on the integrated luminosity.

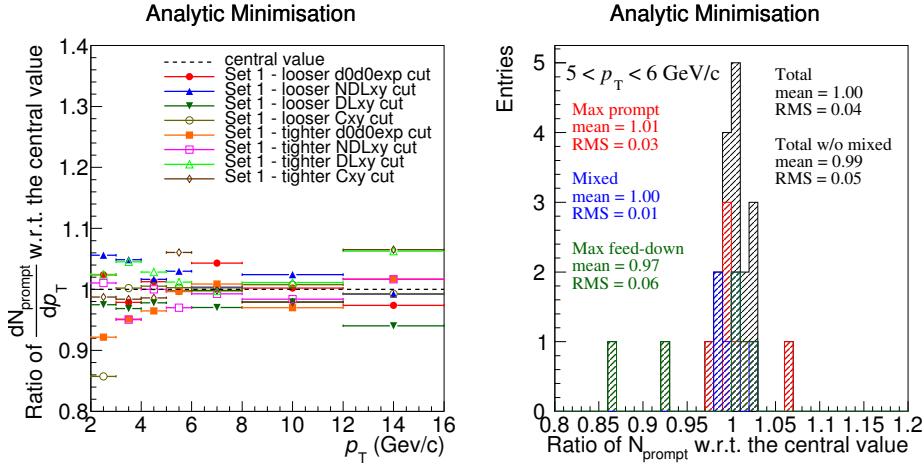


Figure 6.10. Left: ratio of the corrected yields for prompt  $D^+$  as a function of  $p_T$  obtained with the analytic minimisation with respect to the central value, for the cut variations on the prompt-enhanced set. Right: distributions of the ratios of the corrected yield in the interval  $5 < p_T < 6 \text{ GeV}/c$  for the variations on the three cut sets separately (red, blue, green) and for all the trials together (black).

## 7. Systematic uncertainty on the $D^+ \rightarrow K^- \pi^+ \pi^+$ branching ratio.

The first two contributions are the same as for the standard analysis. In particular, the uncertainty on the Particle Identification was found to be negligible, while the one on the tracking efficiency was of 3% for each measured track, resulting in a 9% independent on  $p_T$ . The uncertainty on the integrated luminosity is due to the uncertainty on the Minimum Bias trigger cross section, which was estimated to be of 3.5% in [58], and to the fluctuation of the order of 1% observed in the run-by-run integrated luminosity. Therefore, it results to be 3.7%. The uncertainty on the branching ratio is 2.1% [56].

### 6.5.1 Systematic uncertainty due to the selection efficiency

If the distributions of the variables used for the topological selections were exactly reproduced in the Monte Carlo simulation, the result would be independent from the cut values used in the analysis. However, a possible imperfect description of these variables in the simulation could lead to a bias in the result. For this reason, the minimisation procedure was repeated varying the selection criteria for each cut set. In particular, the corrected yields have been recomputed, adopting a lower and a tighter cut value for each variable, separately for each set. For example, the

$p_T$ (GeV/c)	Prompt D <sup>+</sup> corrected yields									
	Analytic					Incentre				
	Set 1	Set 2	Set 3	Sum	Tot	Set 1	Set 2	Set 3	Sum	Tot
[2,3]	8%	2%	9%	12%	7%	5%	2%	11%	12%	7%
[3,4]	5%	2%	5%	7%	4%	7%	2%	5%	9%	6%
[4,5]	2%	1%	3%	4%	3%	3%	0%	4%	5%	2%
[5,6]	4%	1%	9%	10%	4%	6%	0%	6%	8%	4%
[6,8]	2%	2%	8%	8%	5%	4%	0%	4%	6%	4%
[8,12]	3%	2%	6%	7%	3%	6%	0%	3%	7%	3%
[12,16]	5%	1%	4%	6%	3%	6%	1%	2%	6%	4%

Table 6.3. Estimated uncertainties of the prompt D<sup>+</sup> corrected yields for the single cut sets, the sum in quadrature of the three contributions (*Sum*) and from all the trials together (*Tot*).

cut on the cosine of the pointing angle in the  $xy$ -plane for the prompt-enhanced set was changed to 0.995 and 0.999, while the central value was 0.998. It is important to notice that the cut values adopted for the variations cannot be too different from the central values (see table 6.1), in order to keep the three sets of topological cuts as much as possible uncorrelated. In the left panel of figure 6.10 the ratio between the corrected yields for prompt D<sup>+</sup>-mesons for each cut-variation with respect to central value are shown, for the variations applied to the prompt-enhanced set. The procedure was repeated for the other two sets of topological cuts used in the analysis and for both the minimisation methods. The results in figure 6.10 are obtained with the analytic minimisation described in section 6.1.1. The resulting variations have been treated with two different approaches:

- Considering the variations of the three sets of topological cuts (prompt-enhanced, mixed and feed-down-enhanced) as uncorrelated.
- Considering all the variations together.

In the first case the systematic uncertainty was assigned adding in quadrature the three contributions, evaluated as the sum of the RMS and the shift of the mean, with respect to unity, of the distribution of the ratios of the results from the modified set of topological cuts to the central one. In the second case, the uncertainty is estimated considering the RMS and the shift of the distribution of all the ratios obtained from the variations on the three sets of cuts (prompt-enhanced, mixed and feed-down-enhanced) together. The resulting values are summarised in table 6.3, where the column labelled *Sum* corresponds to the uncorrelated case, while the one labelled *Tot* corresponds to the second case. It is possible to notice how

	$p_T$ (GeV/c)	[2,3]	[3,4]	[4,5]	[5,6]	[6,8]	[8,12]	[12,16]
Tot	analytic	7%	4%	3%	4%	5%	3%	3%
	incentre	7%	6%	2%	4%	4%	3%	4%
W/o mixed	analytic	9%	4%	4%	6%	6%	4%	4%
	incentre	8%	7%	4%	5%	4%	4%	5%

Table 6.4. Comparison between the systematic uncertainty due to the cut efficiency for prompt  $D^+$  estimated considering all the variations and excluding the variations of the mixed cut set.

the values obtained considering all the trials together are systematically smaller than the ones evaluated considering the variations of three cut sets as uncorrelated. This is due to a partial compensation of the effect on the corrected yields due to the variation on the three cut sets, which leads to an overestimation of the uncertainty if the three variations are considered uncorrelated. In the right panel of figure 6.10 an example of this effect is shown for the transverse momentum interval  $5 < p_T < 6$  GeV/c: the distribution of the ratios for the first cut set (red) is shifted toward values higher than unity, while the one for the third cut set (green) toward lower values. Therefore, the total distribution results to be centred at one. Furthermore, from the values reported in table 6.3, it is possible to notice how the mixed cut set is much more stable than other two in the full  $p_T$  range, for both the analytic minimisation and the incentre methods. This suggests that the result is not particular sensitive to the mixed set of cuts. As a consequence, including the ratios from the variations of the mixed cut set in a distribution of all trials (black histogram in the right panel of figure 6.10) could result in an artificial reduction of the RMS and in an underestimation of the uncertainty. For this reason, the systematic uncertainty on the prompt  $D^+$  corrected yields was assigned considering the RMS plus the shift of the distributions without considering the variations of mixed cut set. The comparison between the uncertainties obtained considering the variations of all the three sets and those obtained removing the mixed one are reported in table 6.4.

The same procedure was done for the corrected yields for feed-down  $D^+$ -mesons. In this case the result was found to be more unstable, since the determination of the feed-down contribution mostly depends on the third cut set, which it is the most critical in terms of signal extraction. Therefore, the usage of tighter cut values for the feed-down-enhanced set typically leads to a bad raw yield extraction, which is reflected in the determination of the corrected yield. For this reason, the trials corresponding to a bad fit result were excluded, in order not to double count the systematic uncertainty on the raw yield extraction and to limit as much as possible the contribution due to statistical fluctuations. In the left panel of figure

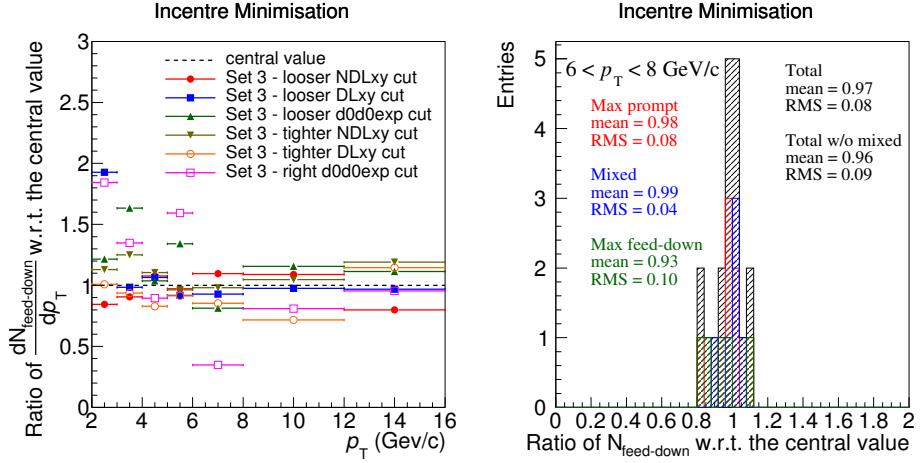


Figure 6.11. Left: ratio of the corrected yields for feed-down  $D^+$  as a function of  $p_T$  obtained with the incentre minimisation with respect to the central value, for the feed-down-enhanced set variations. Right: distributions of the corrected yield variations in the interval  $6 < p_T < 8 \text{ GeV}/c$  for the three cut sets separately (red, blue, green) and for all the trials together (black), obtained removing the bad trial, corresponding to the variation of the cut on the maximum normalised residual of the daughter impact parameter.

$p_T$ (GeV/c)	Feed-down $D^+$ corrected yields									
	Analytic					Incentre				
	Set 1	Set 2	Set 3	Sum	Tot	Set 1	Set 2	Set 3	Sum	Tot
[2,3]	14%	10%	29%	34%	26%	27%	27%	36%	52%	25%
[3,4]	26%	21%	34%	48%	28%	19%	28%	35%	49%	23%
[4,5]	10%	5%	10%	15%	10%	7%	6%	11%	14%	9%
[5,6]	12%	2%	38%	40%	20%	9%	5%	32%	34%	21%
[6,8]	10%	5%	17%	20%	11%	7%	4%	18%	20%	10%
[8,12]	10%	6%	19%	22%	11%	5%	4%	22%	23%	11%
[12,16]	15%	7%	16%	23%	12%	12%	8%	17%	22%	12%

Table 6.5. Estimated uncertainties on the feed-down  $D^+$  corrected yields for the single cut sets, the sum in quadrature of the three contributions (*Sum*) and from all the trials together (*Tot*).

[6.11](#) the ratio of the corrected yields for the feed-down contribution with respect to the central value is shown for the variations on the feed-down-enhanced cut set. In the right plot of the same figure the distributions of the ratios in the interval  $6 < p_T < 8$  GeV/c are shown after excluding the bad trial corresponding to the variation of the cut on the maximum difference between the measured and the expected daughter impact parameter for the feed-down-enhanced set. All the values obtained considering the three sets of topological cuts separately and together are reported in table [6.5](#). As expected, the largest contribution is given by the variations of the feed-down-enhanced set but, unlike the values obtained for the prompt  $D^+$  corrected yields, the uncertainty induced by the mixed set is not negligible. Therefore, in case of the corrected yields for the feed-down  $D^+$ -mesons, the uncertainty is assigned considering the distributions of all the variations, without excluding those on the mixed set of cuts.

### 6.5.2 Systematic uncertainty due to the raw yield extraction

The systematic uncertainty on the yield extraction was evaluated by changing the configuration of the invariant-mass fits for all the three sets of topological cuts and performing the minimisation procedure and the incentre calculation for each trial. In particular, the fits have been repeated setting all the possible combinations of the following variations:

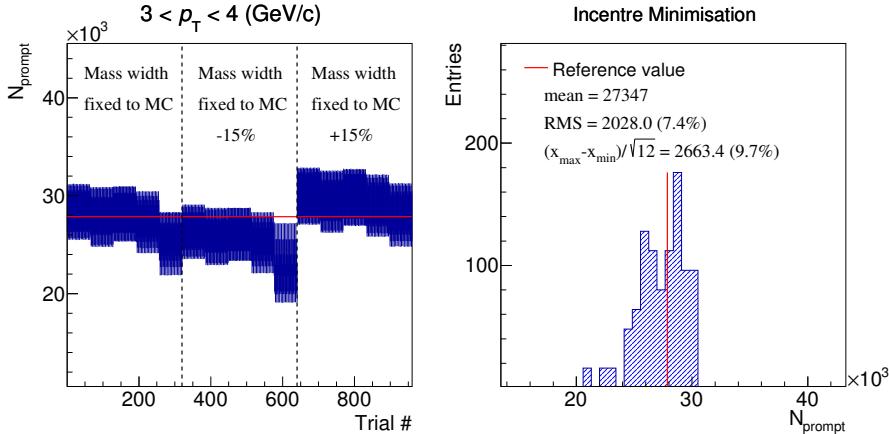


Figure 6.12. Prompt  $D^+$  corrected yield obtain with the incentre minimisation method as a function of the trial number (left) and the resulting distribution (right) for the  $p_T$  interval  $3 < p_T < 4$  GeV/c.

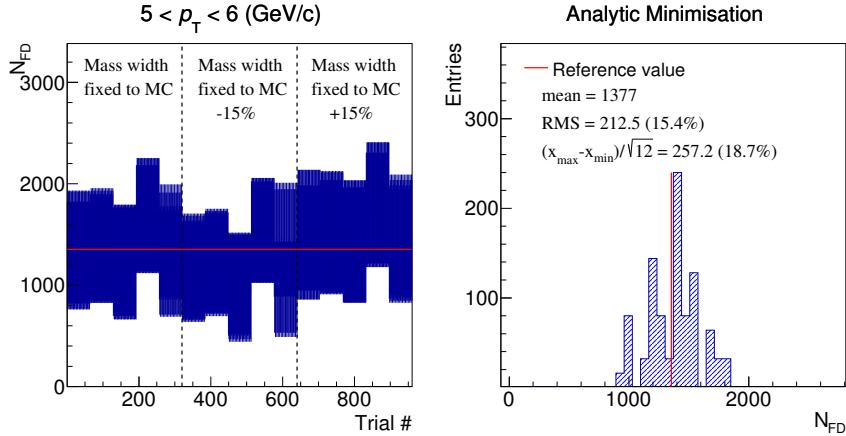


Figure 6.13. Feed-down  $D^+$  corrected yield obtained with the analytic minimisation method as a function of the trial number (left) and the resulting distribution (right) for the  $p_T$  interval  $5 < p_T < 6$  GeV/c.

- The Gaussian width of the signal was fixed to the value obtained from the MC and varying it by  $\pm 15\%$ , which is the maximum discrepancy between the values obtained by fitting the invariant-mass distributions from data and MC simulations.
- The Gaussian mean of the signal was first let free and then fixed to the PDG value of the  $D^+$ -meson mass.
- The bin width of the invariant-mass distributions was changed from 4 to 16 MeV/c<sup>2</sup>.
- The range of the invariant-mass fit has been varied using 4 different lower and upper limits.
- Two different functions have been used for the parametrisation of the background (exponential and parabolic).

The systematic uncertainty was estimated considering the RMS of the distributions of the corrected yields obtained with the analytic minimisation and the incentre methods. Since the resulting distributions are not Gaussian for all the  $p_T$  intervals considered in the analysis, also the variance assuming a uniform distribution (i.e. the full spread divided by  $\sqrt{12}$ ) was evaluated. In figures 6.12 and 6.13 are shown two examples of variation of the prompt and feed-down corrected yields are shown for the intervals  $3 < p_T < 4$  GeV/c and  $5 < p_T < 6$  GeV/c respectively. In the left

	$p_T$ (GeV/c)	[2,3]	[3,4]	[4,5]	[5,6]	[6,8]	[8,12]	[12,16]
Prompt $D^+$	analytic	12%	8%	8%	8%	9%	9%	15%
	incentre	12%	8%	8%	8%	10%	12%	15%
Feed-down $D^+$	analytic	26%	20%	20%	16%	10%	11%	20%
	incentre	30%	20%	20%	18%	11%	11%	20%

Table 6.6. Systematic uncertainty on the corrected yields due to the raw yield extraction.

panel the variation of the corrected yields is shown as a function of the trial number, while in the right plot the resulting distributions are shown. For the prompt  $D^+$  case the largest effect is due to the variation of the Gaussian width, however all the trials are compatible among each other within the statistical uncertainty. For the feed-down corrected yield the statistical uncertainties are dominant, while the effect due to the variation of the Gaussian width is less prominent with respect to the prompt  $D^+$  case. In table 6.6 are reported the values of systematic uncertainty assigned for both the prompt and feed-down corrected yields with the analytic minimisation and the incentre methods for the  $p_T$  intervals of the analysis.

### 6.5.3 Systematic uncertainty due to the $p_T$ shape

The shape of the generated  $D^+$   $p_T$  distribution affects the efficiency, as discussed in section 6.3. Therefore a systematic uncertainty was estimated from the relative variation of the corrected yields obtained using the efficiencies computed with the FONLL and Pythia  $p_T$  shapes. In the left panel of figure 6.14 the ratio of the efficiencies for prompt and feed-down  $D^+$ -mesons obtained with the two different  $p_T$  shapes is shown for the prompt-enhanced and feed-down-enhanced cut sets, while in the right panel the ratio of the corrected yields computed with the analytic minimisation, using the efficiencies obtained with the two different  $p_T$  shapes is displayed. In table 6.7 the values assigned as systematic uncertainty on the basis

	$p_T$ (GeV/c)	[2,3]	[3,4]	[4,5]	[5,6]	[6,8]	[8,12]	[12,16]
Prompt $D^+$	analytic	1%	1%	1%	1%	1%	2%	2%
	incentre	1%	1%	1%	1%	1%	4%	2%
Feed-down $D^+$	analytic	1%	1%	1%	1%	1%	5%	5%
	incentre	1%	1%	1%	1%	1%	5%	5%

Table 6.7. Systematic uncertainty on the corrected yields due to the  $p_T$  shape of the generated  $D^+$ -mesons.

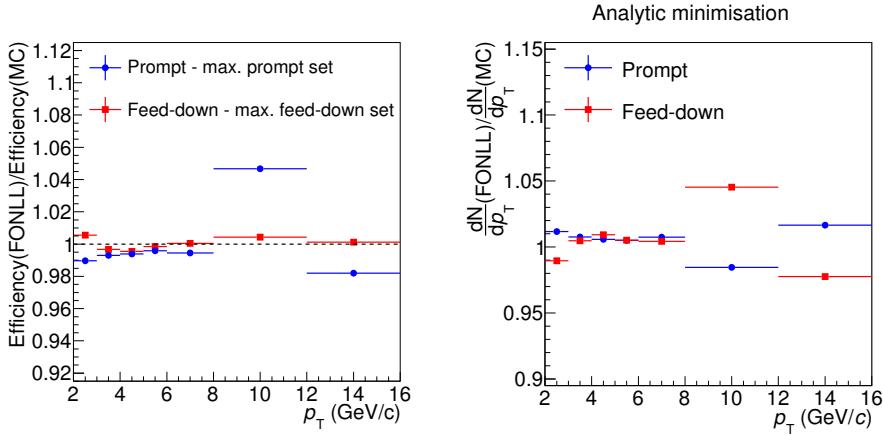


Figure 6.14. Left: ratio of the efficiency for prompt and feed-down  $D^+$ -mesons obtained using the FONLL and Pythia  $p_T$  shapes, for the prompt-enhanced and the feed-down-enhanced cut sets. Right: ratio of the corrected yields obtained with the analytic minimisation, using the efficiencies computed with the FONLL and Pythia  $p_T$  shapes.

of the corrected yield variations are reported.

## 6.6 Results

After the minimisation procedure, the  $p_T$ -differential cross section was calculated as:

$$\left. \frac{d\sigma_X^{D^+}}{dp_T} \right|_{|y| < 0.5} = \frac{1}{2} \cdot \frac{N_X^{D^\pm}|_{|y| < y_{fid}}}{\Delta p_T \Delta y \cdot (Acc)_X} \cdot \frac{\sigma_{MB}^{pPb}}{BR \cdot N_{ev}}, \quad (6.16)$$

where  $X$  indicates the prompt or the feed-down production, and  $N_X^{D^\pm}$  are the corresponding corrected yields computed with the two methods described in section 6.1. Like in the formula used in the standard analysis (see equation 4.8), the factor 1/2 takes into account the fact that the yields  $N_X^{D^\pm}$  include both particles and anti-particles,  $\Delta y$  is the rapidity coverage, which is  $|y| < y_{fid}$  with  $y_{fid}$  depending on  $p_T$ ,  $\Delta p_T$  is the width of the  $p_T$  interval,  $\sigma_{MB}^{pPb}$  is the Minimum Bias trigger cross section and  $N_{ev}$  is the number of events computed as reported in equation 4.9. Note that in equation 6.16 the efficiency factor does not appear, because the yields  $N_X^{D^\pm}$  are already corrected for the efficiency and for the fraction of

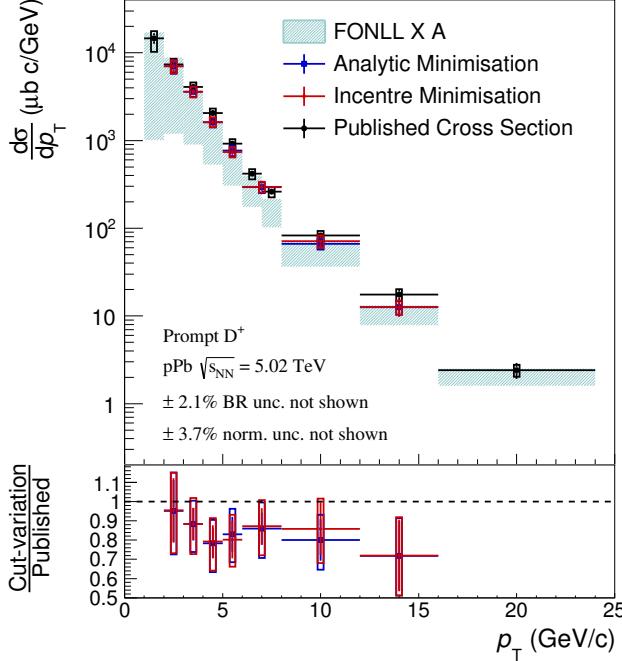


Figure 6.15. Top: prompt  $D^+$  production cross section measured in p-Pb collisions at  $\sqrt{s_{NN}} = 5.02 \text{ TeV}$  using the cut-variation method, compared to the result obtained with the standard approach published by the ALICE Collaboration [36]. The vertical bars represent the statistical errors, while the boxes the systematic uncertainties. Bottom: ratio of the measured cross section with the cut-variation approach and the published result.

prompt (feed-down)  $D^+$ . The resulting  $p_T$ -differential cross section of prompt  $D^+$ -mesons obtained with the cut-variation method is shown in figure 6.15, compared with the published result obtained with the standard approach and to the FONLL prediction at  $\sqrt{s_{NN}} = 5.02 \text{ TeV}$  multiplied by the lead-nucleus mass number, which should describe the data in case of binary scaling. The measured cross sections lie on the upper edge of the FONLL uncertainty band, as it was also observed in pp collisions at LHC energies (see section 2.1). This observation is consistent with the observation of D-mesons  $R_{pPb}$  compatible with unity in the measured  $p_T$  range, which indicates small Cold Nuclear Matter effects. In the bottom panel it is possible to see the ratio between the result obtained with the

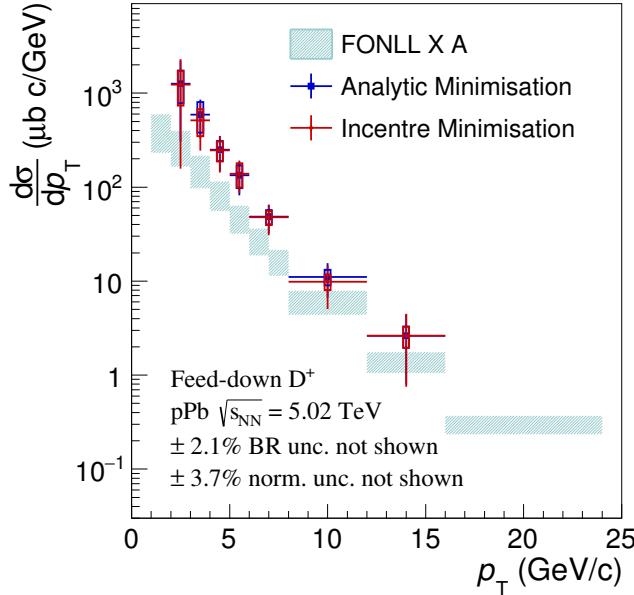


Figure 6.16. Feed-down  $D^+$  production cross section measured in p-Pb collisions at  $\sqrt{s_{NN}} = 5.02$  TeV using the cut-variation method. The vertical bars represent the statistical errors, while the boxes the systematic uncertainties.

cut-variation method and the standard approach, which uses FONLL calculations for the feed-down subtraction. The statistical and the systematic uncertainties have been treated as uncorrelated in the ratio, except for the systematic uncertainties due to the tracking efficiency, the integrated luminosity and the branching ratio, which have been considered as fully correlated. The results obtained with the analytic and the incentre minimisation methods are very similar and they are both compatible within uncertainties with the published one, although systematically lower by about 20%. This effect could be due to the tight selection criteria needed to obtain three different sets of topological cuts, which can create a bias in case of a not perfect description of the distributions of the selection variables in the simulation. Finally, in figure 6.16 the  $p_T$ -differential cross section of feed-down  $D^+$ -mesons obtained with the cut-variation analysis is shown. It is compared to the FONLL prediction at  $\sqrt{s_{NN}} = 5.02$  TeV multiplied by the lead-nucleus mass number. The FONLL prediction, which describes well the beauty production in pp collisions (see section 2.1), should describe well also the data in p-Pb collisions, since the Cold

Nuclear Matter effects are negligible. Despite that, the result obtained with the cut-variation analysis is significantly higher than the FONLL prediction, especially at low  $p_T$ . This can be related to the 20% effect observed in the prompt  $D^+$  differential cross section: note that a small relative deficit in the prompt yields can imply a large relative increase in the feed-down yields because the feed-down contribution is smaller by a factor of 8-9 as compared to the prompt one. The observation of such discrepancy seems to indicate an underestimation of the systematic uncertainties. It could be the consequence of the fact that the feed-down corrected yields are determined mainly by the feed-down-enhanced cut set, which is characterised by very tight selections. In particular, the efficiencies for prompt and feed-down  $D^+$  for this set of cuts are really low, the raw yield extraction is critical because of the low statistics and the systematic uncertainty due to the cut selection efficiency was difficult to evaluate, because of the necessity to keep the three sets of cuts as much as possible uncorrelated. Therefore, the measurement of the feed-down  $D^+$ -mesons with the cut-variation method cannot be considered reliable with the current level of statistics, while it could be improved in the future, because of the larger data sample expected in the LHC Run2 and Run3.

# 7

## Prompt D<sup>+</sup> reconstruction with the Kalman filter vertexing package

In this Chapter I will present the measurement of the  $p_T$ -differential production cross section of prompt D<sup>+</sup>-mesons in p-Pb collisions at  $\sqrt{s_{NN}} = 5.02$  TeV, performed using an alternative vertex-reconstruction approach with respect to the one implemented in the ALICE analysis framework, based on the Kalman filter algorithm. The software used for this analysis (*KFParticle* package [68]) is under development for the CBM experiment, which will be one of the major scientific pillars of the future Facility for Antiproton and Ion Research (FAIR) in Darmstadt [69]. However, because of its interesting features for the analyses which involve displaced decay vertices, such as open heavy-flavoured or strange hadrons, it is also planned to include this software in the ALICE framework. At the moment it is only possible to use it to analyse the ALICE data at the GSI (*Gesellschaft für Schwerionenforschung*) laboratory, where the *KFParticle* package has been included as an external library in the software.

In the first section of this Chapter a brief introduction to the Kalman filter algorithm and, in particular, to the application for the particle-decay reconstruction is provided. In the second section I will present a test of the performances of the *KFParticle* package for the prompt D<sup>+</sup>-meson reconstruction in the MC simulation described in section 4.3. Finally, the analysis performed reconstructing the D<sup>+</sup>-mesons with the *KFParticle* package will be presented and the measured  $p_T$ -differential cross section of prompt D<sup>+</sup>-mesons will be compared to the published

result.

## 7.1 The Kalman filter algorithm

The Kalman filter is an algorithm which provides a precise estimation  $\mathbf{r}$  and its covariance matrix  $\mathbf{C}$  of an unknown state vector  $\mathbf{r}^t$ , using a series of  $n$  measurements  $\mathbf{m}_k$ , with  $k = 1, \dots, n$ , containing statistical noise. The algorithm starts with an approximation of the vector, which is subsequently updated adding one measurement after another. In general the vector  $\mathbf{r}^t$  is changed after each iteration. However, in order to provide a simplified description of the algorithm, I assume  $\mathbf{r}^t$  identical for each determination and the measurement  $\mathbf{m}_k$  linearly dependent on  $\mathbf{r}^t$ :

$$\mathbf{m}_k = H_k \mathbf{r}^t + \boldsymbol{\eta}_k, \quad (7.1)$$

where the matrix  $H_k$  is the model of measurement and  $\boldsymbol{\eta}_k$  is the statistical error on the  $k$ -th measurement  $\mathbf{m}_k$ , which is assumed to be unbiased with a known covariance matrix  $V_k$ :

$$\begin{aligned} \langle \boldsymbol{\eta}_k \rangle &= 0, \\ \langle \boldsymbol{\eta}_k \cdot \boldsymbol{\eta}_k^T \rangle &= V_k. \end{aligned} \quad (7.2)$$

Then the fitting algorithm is divided in two steps:

- **Initialisation:** the state vector is approximated to  $\mathbf{r}_0$  and its covariance matrix is set to  $\mathbf{C} = \mathbb{1} \cdot N$ , where  $N$  is a large number.
- **Filtration:** for each  $k$ -th measurement  $\mathbf{m}_k$ , starting from the values obtained in the  $(k-1)$ -th step, the following quantities are calculated:

$$\begin{aligned} S_k &= (V_k + H_k \mathbf{C}_{k-1} H_k^T)^{-1}, \\ K_k &= \mathbf{C}_{k-1} H_k^T S_k, \\ \zeta_k &= \mathbf{m}_k - H_k \mathbf{r}_{k-1}, \\ \mathbf{r}_k &= \mathbf{r}_{k-1} + K_k \zeta_k, \\ \mathbf{C}_k &= \mathbf{C}_{k-1} - K_k H_k \mathbf{C}_{k-1}, \\ \chi_k^2 &= \chi_{k-1}^2 + \zeta_k^T S_k \zeta_k. \end{aligned} \quad (7.3)$$

In the above formula  $K_k$  is the gain matrix,  $\zeta_k$  the residuals vector and  $\chi_k^2$  the total chi square of the obtained estimation  $\mathbf{r}_k$  from the  $k$ -th measurement.

The best estimation  $\mathbf{r}_n$  of the state vector  $\mathbf{r}^t$  together with the covariance matrix  $\mathbf{C}_n$  is obtained after the filtration of the last measurement. The same procedure can be performed of measurements non-linearly dependent on  $\mathbf{r}^t$ ; in this case one speaks of *extended* Kalman filter.

### 7.1.1 Reconstruction of particle decays with the KFParticle package

The KFParticle package provides an estimation of the parameters of the decayed particles [70], starting from the daughter tracks and their covariance matrices. Furthermore, the algorithm is improved with respect to the standard Kalman filter described in the section above and it takes into account the correlation among the  $k$ -th measurement  $\mathbf{m}_k$  and the previous measurements  $\mathbf{m}_{i < k}$ . The state vector used for the parametrisation of the particle is

$$\mathbf{x}^T = (x, y, z, p_x, p_y, p_z, E, s), \quad (7.4)$$

where  $\mathbf{v} = (x, y, z)$  are the spatial coordinates of the point where the particle is built (the decay vertex),  $\mathbf{p} = (p_x, p_y, p_z)$  is the particle momentum in the considered spatial point,  $E$  is the energy and  $s$  is a parameter called *signed path*, which is defined as the length of the particle trajectory normalised to the momentum. Therefore,  $s$  is non zero only in case of the particle is computed using the production vertex information (see next section). By parametrising the trajectory in this way, the particle with the corresponding covariance matrix can be transported to each point of its trajectory.

### 7.1.2 Fit constraints

The precision of the decay vertex parameters obtained in the fit can be improved by taking into account additional information known about the decayed particle, for example the production vertex or the mass. These assumptions can be used to constrain the parameters of the state vector. In particular I focus on the *topological constraint*, which will be used in the analysis presented in the next sections.

#### Topological constraint

The topological constraint consists in requiring the alignment of a decayed particle with the already known production vertex  $\mathbf{v}_{pv}$ . Considering the simplified case of a neutral particle, the trajectory which connects the primary vertex (from which the particle originates) to the decay vertex is a straight line:

$$\mathbf{v} - s \cdot \mathbf{p} = \mathbf{v}_{pv}. \quad (7.5)$$

This latter equation can be used to fix a constraint to the parameters of the state vector (see equation 7.4) and force the reconstructed particle to point to the production vertex. The primary vertex errors can also be taken into account, for example by adding the primary vertex parameters into the state vector [70]. The same procedure can be repeated for charged particles in a magnetic field, but in

this case the algorithm takes into account the curvature of the particle trajectory. A possible application of the topological constraint can be the separation of the prompt from the feed-down contribution in the  $D^+$ -meson raw yield by evaluating the  $\chi^2$  of the particle after applying the topological constraint. If the  $D$ -meson is prompt, it is produced at the primary vertex and therefore the resulting  $\chi^2$  value divided by the number of degrees of freedom should be on average close to 1. Since the feed-down  $D^+$ -mesons do not originate from the primary vertex, imposing the topological constraint at the primary vertex should imply on average a larger value for their  $\chi^2$ . This is the basic idea of the analysis presented in this Chapter. In the next section the performance of the reconstruction of the prompt  $D^+$ -mesons using the Kalman filter algorithm is tested with the MC sample described in section 4.3, while in section 7.3 this analysis method is applied to the measurement of the  $p_T$ -differential production cross section of prompt  $D^+$ -mesons.

## 7.2 Reconstruction performances of the KFParticle package for prompt $D^+$ -mesons

The performance of the reconstruction quantities of prompt  $D^+$ -mesons from the decays  $D^+ \rightarrow K^- \pi^+ \pi^+$  using the KFParticle package was tested in the MC sample described in section 4.3, applying the event and track quality selections discussed in Chapter 4. The checks have been performed selecting only prompt  $D^+$  in the MC simulation, in order to test the topological constraint at the primary vertex, which, in case of prompt  $D^+$ , coincides with the production vertex. The results were also compared to the values obtained with the  $D$ -meson vertex-reconstruction algorithm

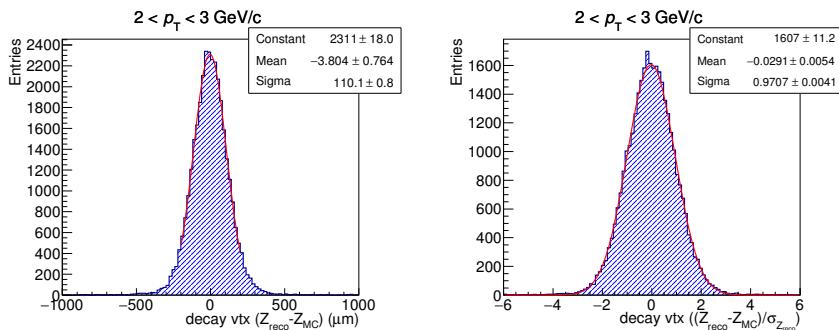


Figure 7.1. Example of Gaussian fit to the residual (left) and the pull (right) distributions of the  $z$  coordinate of the prompt  $D^+$  decay vertex in the interval  $2 < p_T < 3$  GeV/c obtained with the KFParticle package.

used as default in the ALICE Collaboration. In particular, the precision of the determination of the geometrical quantities related to the displaced decay vertex topology was quantified by computing the resolution, defined as the Gaussian width obtained by fitting the residual distributions of a given observable  $X$  in a  $p_T$  interval:

$$Res(p_T) = \sigma(X_{reco} - X_{MC})(p_T), \quad (7.6)$$

where  $X_{reco}$  is the reconstructed value of the variable  $X$ , while  $X_{MC}$  is the value known from the MC truth. To check whether the uncertainty on the variable  $X$  is properly estimated, the pulls are evaluated:

$$Pull(p_T) = \sigma \left( \frac{X_{reco} - X_{MC}}{\sigma_{X_{reco}}} \right)(p_T). \quad (7.7)$$

The pull distribution is expected to be Gaussian with zero mean and width equal to unity, if the difference between the reconstructed and the MC values is covered by the statistical error  $\sigma_{X_{reco}}$  on the reconstructed observable  $X$ . In figure 7.1 an example of Gaussian fit to the residual (left) and the pull (right) distributions of the  $z$  coordinate of the  $D^+$  decay vertex is shown for the  $D^+$  transverse momentum interval  $2 < p_T < 3$  GeV/c.

### 7.2.1 Prompt $D^+$ -mesons at the decay vertex

The left panel of figure 7.2 shows the resolution on the  $D^+$  decay vertex position (only  $x$  and  $z$  coordinates, since the  $y$  coordinate is equivalent to the  $x$  one) as a function of  $p_T$ , using the KFParticle package (labelled *KF*), with and without the topological constraint at the reconstructed primary vertex, and the standard algorithm used in ALICE (labelled *AliVertexer*). When the topological constraint is applied, the particle trajectory is recomputed and its parameters are stored at the production vertex. Therefore it is necessary to transport the reconstructed particle to the decay vertex. Therefore, the parameters of the particle has been transported to the point on its trajectory closest to the reconstructed decay vertex, which was stored before applying the constraint. The residuals and the pulls are then evaluated computing the difference between the coordinates of the position obtained after the transport and the coordinates of the true decay-vertex position, known from the MC truth. As expected, the resolution is worse for the  $z$  coordinate rather than the  $x$  coordinate, for all the reconstruction algorithms tested. The resolution obtained with the Kalman filter without constraint is compatible with the one obtained using the standard vertex algorithm. When the topological constraint is applied, an improvement in the low (full)  $p_T$  region is observed for the  $x$  ( $z$ ) coordinate. This observation confirms the improved reconstruction of prompt  $D^+$  decay vertices achieved by applying the topological constraint, however it cannot be used in the analysis, since it is not possible to separate a priori the

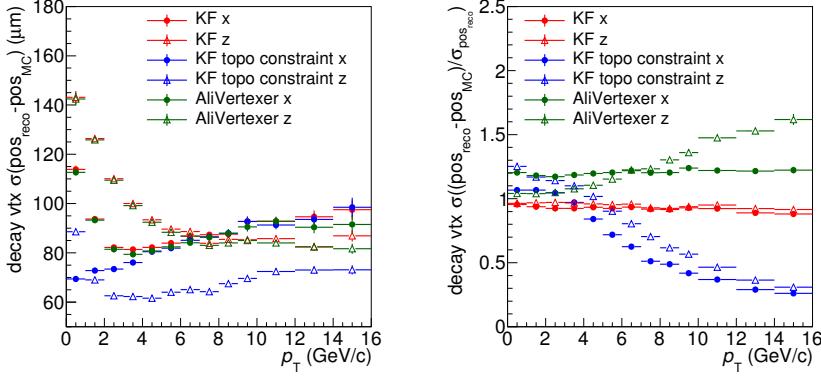


Figure 7.2. Resolution (left) and pulls (right) of the  $x$  and  $z$  coordinates of the prompt  $D^+$  decay vertex position as a function of  $p_T$ .

prompt from the feed-down  $D^+$ -mesons. The values of the pulls obtained with the standard vertexer are slightly larger than unity, indicating an underestimation of the errors of about 20% (up to 50% for the  $z$  coordinate). The pulls obtained with the Kalman filter are close to unity in case of no constraint, while they decrease up to 0.2 if the topological constraint is applied. This effect is due to the fact that the errors on the reconstructed trajectory of the particle and the secondary vertex previously stored are strongly correlated. However, as already mentioned, the reconstructed position of the secondary vertex and its error should not be used after applying the topological constraint.

### 7.2.2 Prompt $D^+$ -mesons at the production vertex

The resolution and the pulls have been studied also for the position of the prompt  $D^+$  at the production vertex. In case of topological constraint the particle parameters are already stored at the production vertex, while, if the constraint is not applied, the particle must be transported to the point of closest approach to the reconstructed primary vertex. In the latter case, the uncertainties on the reconstructed primary vertex are taken into account. Figure 7.3 shows the resolution (left) and the pulls (right) of the  $x$  and  $z$  coordinates of the production vertex position in both cases. The resolution is better for the  $x$  coordinate and it is improved for both the coordinates after applying the topological constraint due to the fact that the particle trajectory is forced to pass closer to the primary vertex. The pulls are close to unity for the  $z$  coordinate, while 10%-20% lower for the  $x$

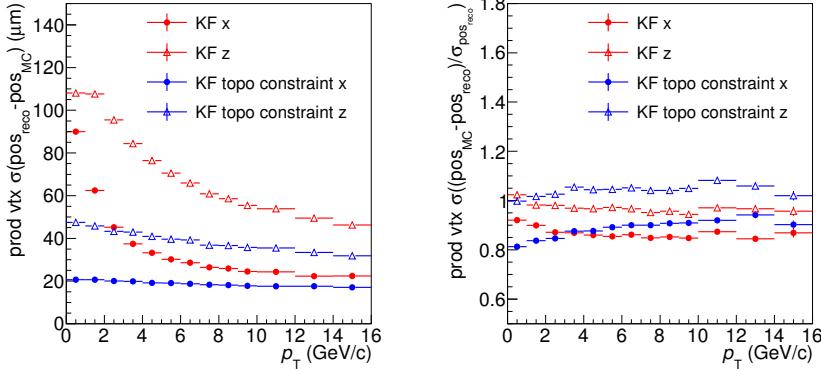


Figure 7.3. Resolution (left) and pulls (right) of the  $x$  and  $z$  coordinates of the prompt  $D^+$  production vertex position as a function of  $p_{\text{T}}$ .

coordinate. It is important to note that, in this case, when the topological constraint is applied, the pulls are still properly evaluated. A quantity related to the position at the production vertex is the impact parameter, which it is computed as the distance of closest approach of the  $D^+$  trajectory to the reconstructed primary

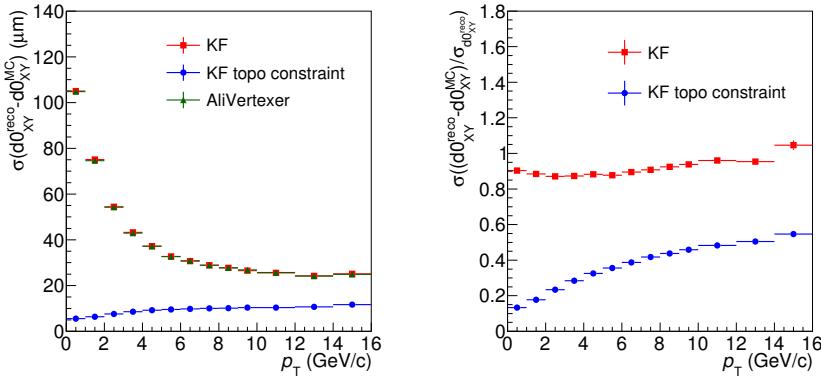


Figure 7.4. Resolution (left) and pulls (right) of the impact parameter in the transverse plane for prompt  $D^+$ -mesons as a function of  $p_{\text{T}}$ .

vertex. The true value of the  $D^+$  impact parameter is zero, since the test is performed only on prompt  $D^+$ -mesons. In figure 7.4 the resolution (left) and the pulls (right) are shown for the impact parameter in the transverse plane obtained with the Kalman filter, compared to those achieved with the standard algorithm. Even if, with the standard algorithm, the impact parameter is computed by approximating the prolongation of the  $D^+$  trajectory with a straight line, the resolution (left panel of figure 7.4) is similar to the one achieved with the Kalman filter without constraints, which takes into account the curvature of the  $D^+$  track induced by the magnetic field. This is a consequence of the short decay length of the  $D^+$ -mesons, for which the curvature becomes negligible. The impact parameter resolution is improved as expected when the topological constraint is applied. The right panel of figure 7.4 shows that the pulls of the impact parameter in the transverse plane obtained with the Kalman filter without constraints are close to unity, while they indicate a significant overestimation of the errors in case of topological constraint.

### 7.2.3 Prompt $D^+$ -meson decay length

The decay length is computed differently by the standard algorithm in the ALICE software and the KFParticle package. In the first case it is calculated as the distance between the reconstructed decay vertex of the  $D^+$ -meson and the reconstructed primary vertex, and therefore it is approximated as a straight line, while in the second case it is given by the  $s$  parameter, multiplied by the momentum of the particle (see section 7.1.1). In the left panel of figure 7.5 the resolution of the decay length in the transverse plane calculated with the two different approaches

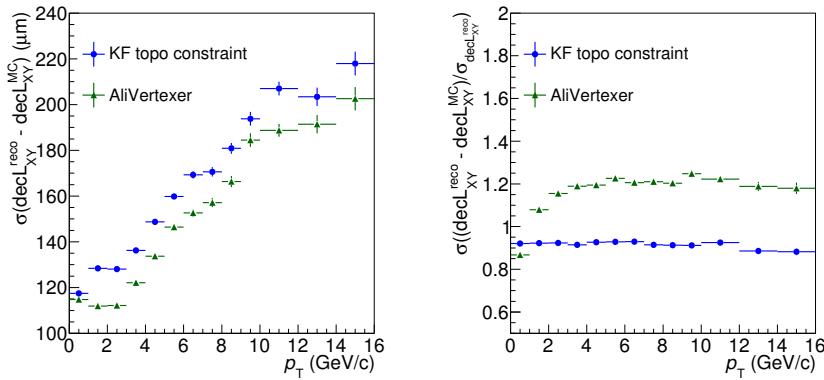


Figure 7.5. Resolution (left) and pulls (right) of the decay length in the transverse plane of the prompt  $D^+$ -mesons as a function of  $p_T$ .

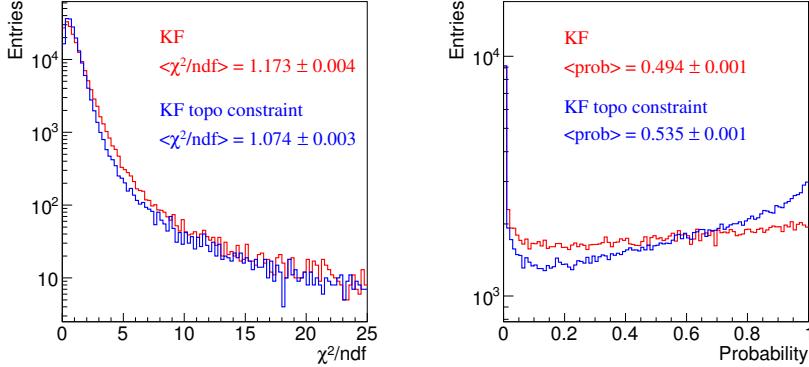


Figure 7.6.  $\chi^2$  (left) and probability (right) distributions for the prompt  $D^+$ -mesons reconstructed with the Kalman filter algorithm, with and without applying the topological constraint.

is shown as a function of  $p_T$ . The resolution seems to be worse considering the reconstruction of Kalman filter, however the MC value is computed as the distance between the MC primary and secondary vertices, neglecting the track curvature, since the information of the decay length is not stored in the MC truth. In the right panel of the same figure the pulls are plotted. For the standard algorithm they suggest an underestimation of the error by about 20%, while for the Kalman filter an overestimation by about 10%. This feature will result in a shift to lower values of the normalised decay length distribution reconstructed with the Kalman filter algorithm with respect to the standard one.

#### 7.2.4 Prompt $D^+$ -meson $\chi^2$ and probability

The possibility to evaluate the  $\chi^2$  of the reconstructed decayed particle is the feature of the Kalman filter that, in the next section, will be exploited for the reduction of the feed-down contribution. However, it is important to check whether the  $\chi^2$  is well evaluated for prompt  $D^+$ , especially when the topological constraint is applied, before applying a selection on this variable. Figure 7.6 shows the distribution of the  $\chi^2$  divided by the number of degrees of freedom (left) and the associated probability (right). As expected for the case of good estimation of the errors, the mean value of the  $\chi^2/\text{ndf}$  distribution is close to unity and the probability distribution is approximately flat, both with and without applying the topological constraint.

### 7.2.5 Primary vertex determination

Another feature of the Kalman filter algorithm is the possibility to add or remove measurements to the evaluation of the state vector. In the case of the primary vertex it results in adding or removing tracks from the fit. In particular, the tracks of the  $D^+$  daughters (which do not originate from the primary vertex) can be removed from the primary vertex evaluation. Furthermore, it is also possible to add the reconstructed  $D^+$  itself, as an additional measurement. In figure 7.7 the Gaussian widths (resolution) and the means (bias) of the residual distributions of the  $x$  and  $z$  coordinates of the primary vertex position are shown as a function of the number of tracks used to determine the primary-vertex position, for events with at least a  $D^+$ -meson. The resolution (left panel) is better for the  $x$  coordinate with respect to the  $z$  coordinate. Furthermore, it seems to be worse if the daughter tracks are removed, but this is only due to the fact that a lower number of tracks contribute to its determination. On the contrary, the resolution is slightly improved if the reconstructed track of the  $D^+$  is added to the determination of the primary vertex using the Kalman filter algorithm, since the prompt  $D^+$ -mesons originate from the primary interaction vertex. However, this feature cannot be used in the analysis, since it is not possible to identify which  $D^+$ -mesons are prompt or feed-down. Despite that, the  $\chi^2$  of the primary vertex, after adding the  $D^+$  to the fit, can be used as an additional information to discriminate between the two contributions. In fact, since the feed-down  $D^+$  trajectories in general do not point

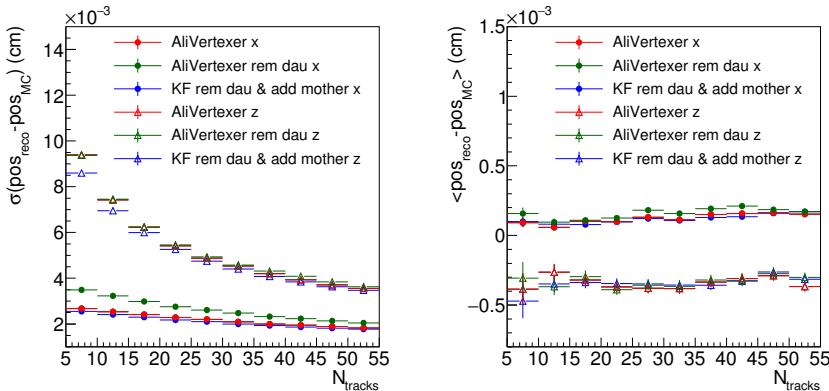


Figure 7.7. Resolution (left) and bias (right) of the  $x$  and  $z$  coordinates of the reconstructed primary-vertex position as a function of the the number of tracks used for the vertex determination in events with at least a  $D^+$ -meson, for different cases of tracks selected for the vertex determination.

at the primary vertex, adding them in the fit would cause a deterioration of the  $\chi^2$ . Including the daughter tracks in the primary vertex determination is expected to create a bias in the primary vertex position and consequently in the  $D^+$  flight direction. This is the reason to exclude them from the fit. However, the  $D^+$  direction is random with respect to the  $(x, y, z)$  reference frame and therefore the mean values of the residual distributions for the single coordinates do not change after removing the daughter tracks or adding the prompt  $D^+$ , as illustrated in the right plot of figure 7.7. In particular, for all the tests performed, a small bias is observed, which is of the order of  $1 - 2 \mu\text{m}$  in the  $x$  direction and about  $3 - 4 \mu\text{m}$  for the  $z$  coordinate.

## 7.3 Prompt $D^+$ production measurement with the KFParticle package

The  $p_{\text{T}}$ -differential production cross section of prompt  $D^+$ -mesons was measured in the p-Pb data sample at  $\sqrt{s_{NN}} = 5.02 \text{ TeV}$  described in Chapter 4, using the KFParticle package for the reconstruction of the  $D^+ \rightarrow K^- \pi^+ \pi^+$  candidates. The transverse momentum range covered is  $1 < p_{\text{T}} < 24 \text{ GeV}/c$  and the  $p_{\text{T}}$ -intervals are the same as in the standard analysis, published by the ALICE Collaboration [36]. For this analysis, all the event, track quality and PID selections used in the standard analysis were applied. As explained in the next section, the selection criteria on the topological variables were slightly changed with respect to those reported in table 4.2 and a selection on the  $\chi^2/\text{ndf}$  of the reconstructed  $D^+$ -mesons with the Kalman filter algorithm, evaluated after applying the topological constraint, was added, in order to reduce the feed-down contribution. As already anticipated in Chapter 4, the code for the latter selections and the  $D^+$  candidate reconstruction has been re-implemented, since it was necessary to use the ESDs instead of the AODs. Furthermore, the available sample on the GSI farm was slightly smaller than the full Minimum Bias sample available for grid analysis on the AODs: the number of analysed events is of the order of  $80 \cdot 10^6$ .

### 7.3.1 Topological selections

All the quantities related to the  $D^+$  decay topology exploited in the analysis were computed by reconstructing the  $D^+$  candidates with the KFParticle package, without applying any constraint, except for the decay length and the  $\chi^2$  of the candidate, for which the topological constraint at the primary vertex was used. In the first case the topological constraint was applied because the Kalman filter algorithm needs the information of the production vertex to compute the decay length, while in the second it was applied to reduce the feed-down  $D^+$  contribution. Figure 7.8 shows the comparison between the distributions of the  $\chi^2/\text{ndf}$  of prompt

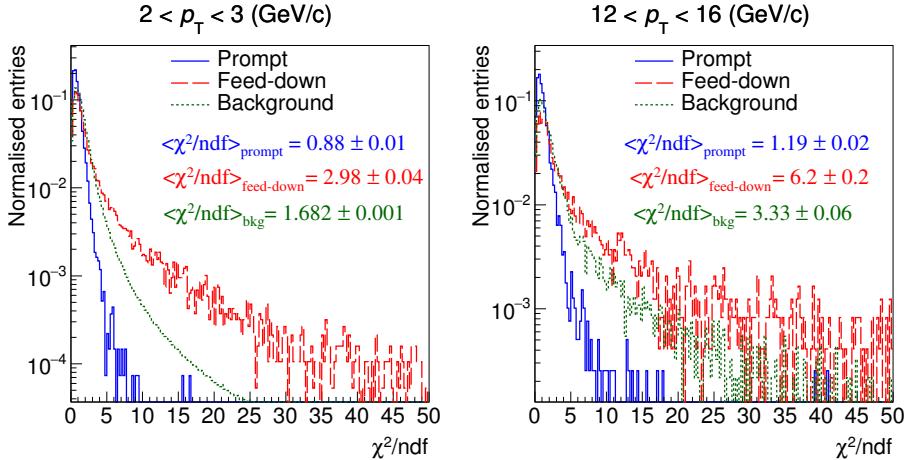


Figure 7.8. Comparison between the distributions of  $\chi^2/ndf$  of prompt and feed-down  $D^+$ -mesons after applying the topological constraint at the primary vertex from MC simulations, in two different  $p_T$  intervals of the analysis.

$D^+$  and feed-down  $D^+$ , obtained from MC simulations, and the distribution of the combinatorial background, obtained from data using the  $D^+$  candidates in the invariant-mass region in the side-bands of the  $D^+$  peak, in two different  $p_T$  intervals of the analysis. The mean value of  $\chi^2/ndf$  is close to unity for prompt  $D^+$ , while it is larger for feed-down  $D^+$ , for which the  $\chi^2/ndf$  distribution is wider. The difference between the two distributions increases with increasing  $D^+$  transverse momentum. The distribution of the background is wider than that of prompt  $D^+$ ,

$p_T$ (GeV/c)	[1,2]	[2,3]	[3,8]	[8,12]	[12,16]	[16,24]
$p_T^{dau}$ (GeV/c) >	0.3	0.3	0.3	0.3	0.3	0.3
Norm. $L_{xy}$ >	6	5	5	5	6	6
$\cos(\theta_P) >$	0.985	0.985	0.985	0.985	0.985	0.985
$\sigma_{vertex}$ (cm) <	0.030	0.030	0.034	0.050	0.050	0.050
$\chi^2/ndf <$	2.5	3.0	3.0	3.5	3.5	4.0

Table 7.1. Summary table of the topological cuts optimised for the analysis performed using the Kalman filter algorithm for the reconstruction of  $D^+$ -mesons.

but narrower with respect to the one of feed-down  $D^+$ , since most of the combinatorial background is given by tracks that come from the primary vertex. In order to exploit as much as possible the selection on the  $\chi^2/ndf$ , the other selections have been slightly released, as reported in table 7.1. In particular, the cuts on the cosine of the pointing angle and on the normalised decay length in the transverse plane were set to looser values with respect to those reported in table 4.2, while the selections on  $\cos(\theta_P^{xy})$  and the decay length were not applied. The cut on the  $\chi^2/ndf$  was tuned taking into account the efficiency for prompt and feed-down  $D^+$ , the fraction of prompt  $D^+$  and the possible bias on the  $p_T$ -differential cross section induced by the discrepancy between the distributions of  $\chi^2/ndf$  of prompt  $D^+$  in data and MC, as discussed in detail in the next section.

### 7.3.2 Tuning of the $D^+$ -meson $\chi^2/ndf$ selection

In this section I present the studies performed to tune the cut on the  $D^+$  candidate  $\chi^2/ndf$  after applying the topological constraint, which guided the choice of the values reported in table 7.1 and allowed us to test the stability of the result. In particular, several cut values have been tested, varying those reported in table 7.1 from a minimum of  $-1.5$  to a maximum of  $+1.5$ .

#### Raw yield variation

The raw yield was extracted in the  $p_T$  intervals of the analysis by fitting the invariant-mass distributions with the sum of a Gaussian function for the signal (prompt and feed-down  $D^+$ ) and an exponential function for the combinatorial background. For these tests the Gaussian width was fixed to the values obtained from the MC simulations in order to avoid as much as possible the effect of statistical fluctuations. The values of the raw yields in the  $p_T$  intervals of the analysis for each trial are shown in the left panel of figure 7.9, while in the right panel the relative variation with respect to the values obtained without applying the cut on  $\chi^2/ndf$  is shown. The amount of signal (prompt and feed-down  $D^+$ ) loss applying the default  $\chi^2/ndf$  cut (*central set*) is of the order of about 5-10% (red points), while for the tightest cut tested the raw yields decrease by up to 30%. Therefore, the central cut on  $\chi^2/ndf$ , reported in table 7.1, allows one to keep the majority of the signal and hence it does not lead to a significant increase of the statistical uncertainty.

#### Efficiency variation

In figure 7.10 the relative variation of the efficiencies for prompt (left) and feed-down (right)  $D^+$  are shown for the different  $\chi^2/ndf$  cuts, with respect to the efficiency obtained without applying the  $\chi^2/ndf$  cut. The decrease of the efficiency

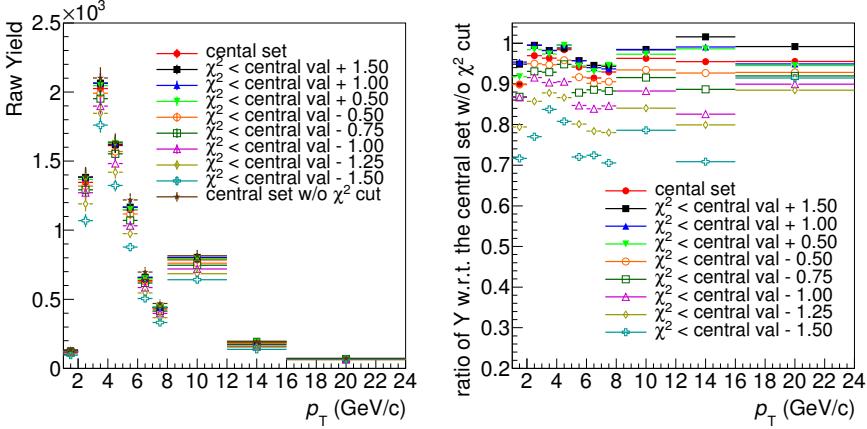


Figure 7.9. Values of the extracted raw yield in the  $p_T$  bins of the analysis for different  $\chi^2/ndf$  cut values (left) and relative variation with respect to the values obtained without applying the  $\chi^2/ndf$  cut (right).

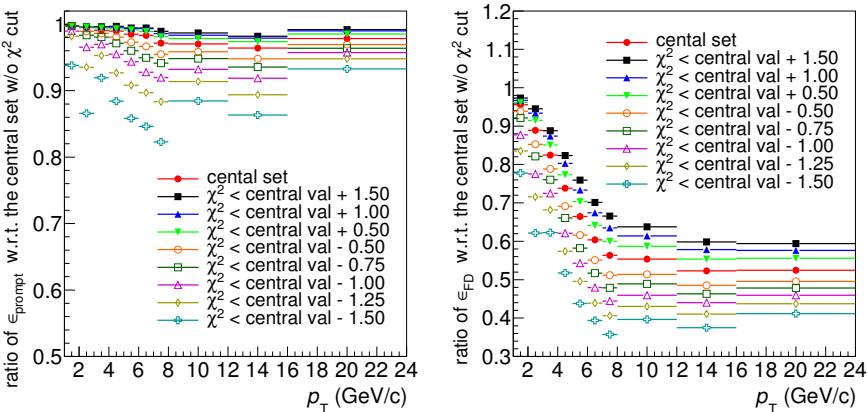


Figure 7.10. Relative variation of the efficiencies for prompt (left) and feed-down (right)  $D^+$  obtained for the different  $\chi^2/ndf$  cuts with respect to those obtained without applying the  $\chi^2/ndf$  cut.

is lower than 5% (20%) for prompt  $D^+$  and up to 45% (60%) for feed-down  $D^+$  mesons for the central (tightest) value of the  $\chi^2/ndf$  cut. These variations show that the rejection of the feed-down contribution achieved with the central value of the  $\chi^2/ndf$  cut is similar to that obtained with the tightest cut value tested, but it allows to keep more signal, as discussed in the previous section and confirmed by the results on the prompt  $D^+$  efficiency. Furthermore, the variation of the efficiency for feed-down  $D^+$  is strongly dependent on the  $D^+$  transverse momentum, because of the shape of the  $\chi^2$  distribution, which becomes wider with increasing  $D^+$  transverse momentum. These studies show that the cut on the  $\chi^2/ndf$  reduce the feed-down contribution in the  $D^+$  raw yield, as expected.

### $D^+$ prompt fraction variation

The prompt fraction in the  $D^+$  raw yield has been evaluated with the FONLL-based methods,  $N_b$  and  $f_c$ , described in section 4.2.1. The central value is given by the  $N_b$  method, while the  $f_c$  method is used to evaluate the systematic uncertainty. The resulting prompt fraction is larger than 0.9 for all the  $\chi^2/ndf$  cut values tested and slightly dependent on the  $D^+$   $p_T$ . As illustrated in the left panel of figure 7.11, the reduction of the feed-down contribution achieved with the cut on the  $\chi^2/ndf$  is significant at high  $p_T$ , while negligible at low  $p_T$ . The variation of the  $D^+$  prompt

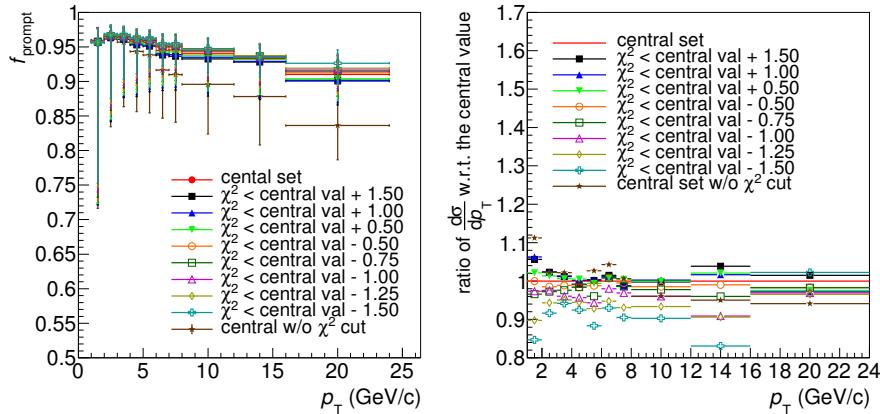


Figure 7.11. Left:  $D^+$  prompt fraction evaluated with the FONLL-based methods for the different  $\chi^2/ndf$  cuts. Right: relative variation of the prompt  $D^+$   $p_T$ -differential cross section obtained with the different  $\chi^2/ndf$  cut values with respect to the central value.

fraction among the different cut values tested is small. Note that the reduced feed-down contribution provides also a reduction of the uncertainties on the prompt  $D^+$  cross section due to the feed-down subtraction.

### $p_T$ -differential cross section variation

If the distributions of  $\chi^2/ndf$  were equal in data and MC, then the prompt  $D^+$   $p_T$ -differential cross section should not vary changing the value of the  $\chi^2/ndf$  cut. The right panel of figure 7.11 shows the ratio of the  $p_T$ -differential production cross section of prompt  $D^+$ , computed using equation 4.8, for the different  $\chi^2/ndf$  cut values with respect to the one obtained with the central  $\chi^2/ndf$  cut. From this plot it is possible to see that the  $p_T$ -differential cross section is systematically lower for tighter  $\chi^2/ndf$  cuts, indicating a possible discrepancy between the data and MC simulation in the region where the distributions are steeper ( $\chi^2/ndf \lesssim 2$ ). This effect was taken into account in the evaluation of the systematic uncertainties (see section 7.4.2). However, for the central values chosen, this bias is not observed or, in any case, it is very small.

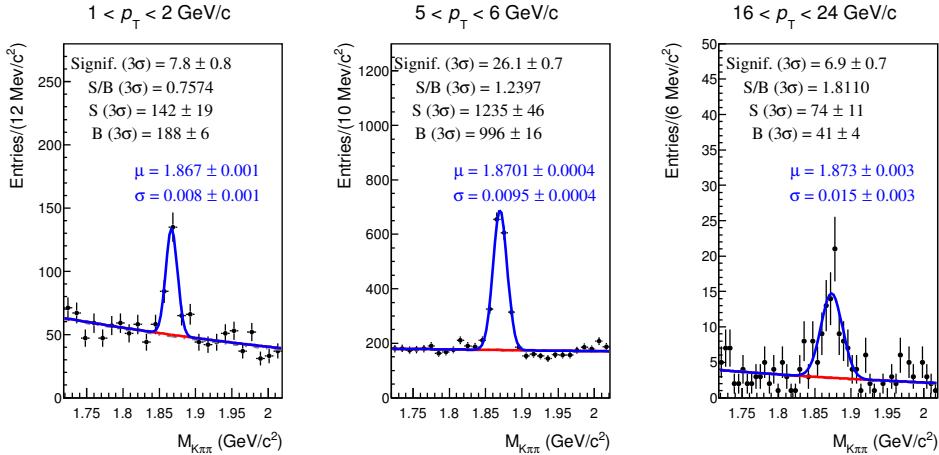


Figure 7.12. Invariant-mass fits in the transverse momentum intervals  $1 < p_T < 2$  GeV/c,  $5 < p_T < 6$  GeV/c and  $16 < p_T < 24$  GeV/c obtained applying the topological cuts reported in table 7.1 and the PID selection.

### 7.3.3 Raw yield extraction

In order to compute the central value of the  $p_T$ -differential cross section, the raw yields have been extracted from the invariant-mass distributions obtained applying the topological cuts summarised in table 7.1 and the PID selection on the daughter tracks. The fitting function is composed of a Gaussian function for the signal (prompt and feed-down  $D^+$ ) and an exponential function for the background. Both the Gaussian mean and width are left free in the fit. In figure 7.12 three examples of invariant-mass fits for different  $p_T$  intervals are shown, while the others can be found in the appendix A.6. In figure 7.13 the values of the Gaussian width (left) and mean (right) are compared to those obtained fitting the invariant-mass distributions of the signal from Monte Carlo simulations. The peak position is also compared to the PDG value of the  $D^+$  mass. Concerning the peak width, the values are all compatible with the simulation within 2 sigmas, except for the  $5 < p_T < 6$  GeV/c and  $8 < p_T < 12$  GeV/c. For the mean, only the value obtained in the interval  $4 < p_T < 5$  GeV/c is not in agreement within the uncertainty. The discrepancy between the values of the parameters extracted from data and simulation will be taken into account in the evaluation of the systematic uncertainties. In the lowest and the highest  $p_T$  intervals the statistical significance is lower with respect to the other  $p_T$  bins. At low  $p_T$  this is due to the strong PID strategy applied necessary to extract the signal because of the large combinatorial background, while at high  $p_T$  because of the low statistics. The extracted raw yield, the signal-to-background ratio and the significance within  $3\sigma$  obtained in all the  $p_T$  intervals of the analysis are reported in table 7.2.

$p_T$ (GeV/c)	Signal	$S/B(3\sigma)$	Significance( $3\sigma$ )
[1,2]	$142 \pm 19$	0.7574	$7.8 \pm 0.8$
[2,3]	$1432 \pm 76$	0.3376	$19.0 \pm 0.9$
[3,4]	$2249 \pm 85$	0.4970	$27.3 \pm 0.9$
[4,5]	$1842 \pm 65$	0.7851	$28.5 \pm 0.8$
[5,6]	$1235 \pm 46$	1.2397	$26.1 \pm 0.7$
[6,7]	$735 \pm 35$	1.3602	$20.6 \pm 0.7$
[7,8]	$503 \pm 26$	1.7449	$17.9 \pm 0.6$
[8,12]	$910 \pm 38$	1.4563	$23.2 \pm 0.7$
[12,16]	$201 \pm 16$	2.1726	$11.7 \pm 0.6$
[16,24]	$74 \pm 11$	1.8110	$6.9 \pm 0.7$

Table 7.2. Raw yield, signal-to-background ratio and significance obtained from the fit to the invariant-mass distributions obtained by reconstructing the  $D^+$  candidates with the KFParticle package.

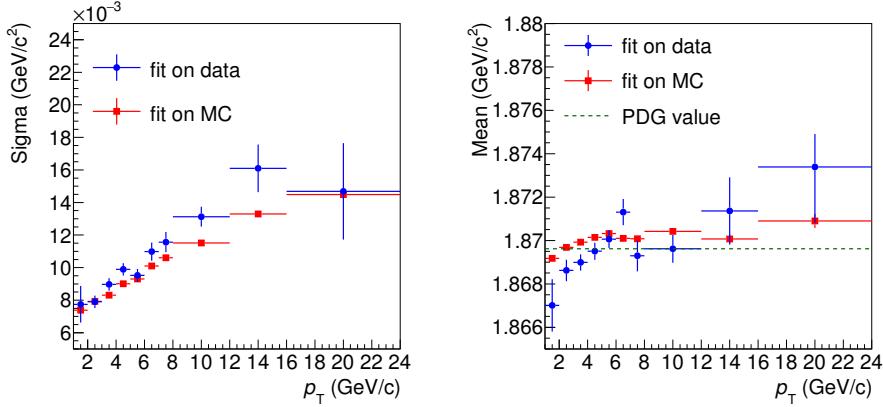


Figure 7.13. Comparison between the peak width (left) and mean (right) obtained by fitting the invariant-mass distributions obtained from data and simulation. The peak position is also compared to the PDG value of the  $D^+$  mass.

## 7.4 Systematic uncertainties

The sources of systematic uncertainties evaluated for this analysis are the following:

1. Systematic uncertainty due to the raw yield extraction.
2. Systematic uncertainty due to the topological selection efficiency.
3. Systematic uncertainty due to the PID selection efficiency.
4. Systematic uncertainty due to the generated  $p_T$  shape of  $D^+$  and  $B$  mesons.

Moreover, for the computation of the  $p_T$ -differential production cross section of prompt  $D^+$ , other four sources have been considered:

1. Systematic uncertainty due to the beauty feed-down subtraction.
2. Systematic uncertainty due to the tracking efficiency.
3. Systematic uncertainty on the integrated luminosity.
4. Systematic uncertainty on the  $D^+ \rightarrow K^- \pi^+ \pi^+$  branching ratio.

In particular, the systematic uncertainty due to the beauty feed-down subtraction is smaller than that in the standard analysis, thanks to the selection on the

$\chi^2/ndf$  of the reconstructed  $D^+$  candidates, evaluated after applying the topological constraint at the primary vertex, which reduces the feed-down contribution. The comparison between the two results will be presented in section 7.4.5. For the latter three sources the assigned systematic uncertainties are the same as those in the data-driven cut-variation analysis: 9% systematic uncertainty due to the tracking efficiency, 3.7% due to the integrated luminosity and 2.1% due to the branching ratio.

### 7.4.1 Systematic uncertainty due to the raw yield extraction

The systematic uncertainty due to the raw yield extraction was evaluated with a multi-trial approach, repeating the fits to the invariant-mass distributions, considering all the possible combinations of the following fit configurations:

- The Gaussian width was first left free and then fixed to the value obtained from the MC simulation.
- The bin width of the invariant-mass distributions was changed from 4 to 16 MeV/c<sup>2</sup>.
- The range of the invariant-mass fit has been varied using 6 different lower and upper limits.

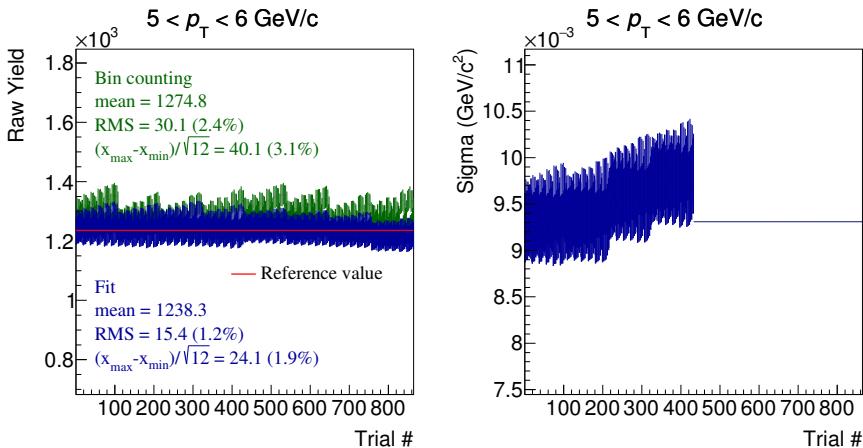


Figure 7.14. Raw yield obtained via the fit to the invariant-mass distributions and the bin counting method (left) and the Gaussian width (right) as a function of the trial number for  $D^+$  candidates reconstructed with the KFParticle package with  $5 < p_T < 6$  GeV/c.

- Three different functions have been used for the parametrisation of the background (exponential, linear and parabolic).

In addition, for each trial, the amount of signal was extracted from the invariant-mass distributions using an alternative method called *bin counting*, which consists in counting the entries in the invariant-mass histogram within  $3\sigma$  around the peak centre after subtracting the background estimated via the fit. The bin counting method is used to test the effect of the line shape of the signal. In figure 7.14 the raw yield extracted by integrating the Gaussian function and with the bin counting method (left) and the Gaussian width (right) are shown as a function of the trial number in the interval  $5 < p_T < 6$  GeV/c. The values of Gaussian width without statistical error in the right plot correspond to the trials in which the sigma was fixed to the value obtained in the simulation. The systematic uncertainty has been assigned considering the RMS of the distributions of the raw yield extracted with the two methods. In case of a not Gaussian distribution, also the variance assuming a flat distribution (i.e. difference between the maximum and minimum values divided by  $\sqrt{12}$ ) was evaluated. Finally, since the values obtained with the bin counting are systematically slightly higher than those extracted from the fit, also the difference between the mean values of the two distributions was taken into account. The values assigned as systematic uncertainty are reported in table 7.3.

#### 7.4.2 Systematic uncertainty due to the topological selection efficiency

As already mentioned, an imperfect description of the distributions of the variables used for the topological selections in the Monte Carlo simulation would induce a bias in the efficiency and consequently in the  $p_T$ -differential cross section. To evaluate this possible systematic effect, the analysis was repeated varying the cut values. In particular, for each variable, several looser and tighter selections have been applied, leaving the others unchanged. In the left panel of figure 7.15 the ratio of the  $p_T$ -differential cross section with respect to the central value, obtained varying the cut on the normalised decay length is shown. Also the variations of the cut on the  $\chi^2/ndf$  of the reconstructed  $D^+$  candidates after applying the topological constraint (see figure 7.11) were considered for the evaluation of the systematic uncertainty. The systematic uncertainty was assigned considering the RMS and shifts with respect to unity of the mean of the distributions of the ratios to the central value obtained with all the variations. In the right panel of figure 7.15 an example for the interval  $7 < p_T < 8$  GeV/c is shown. The systematic uncertainty assigned decreases from 10% at low  $p_T$  to 5% at high  $p_T$ . All the values are summarised in table 7.3.

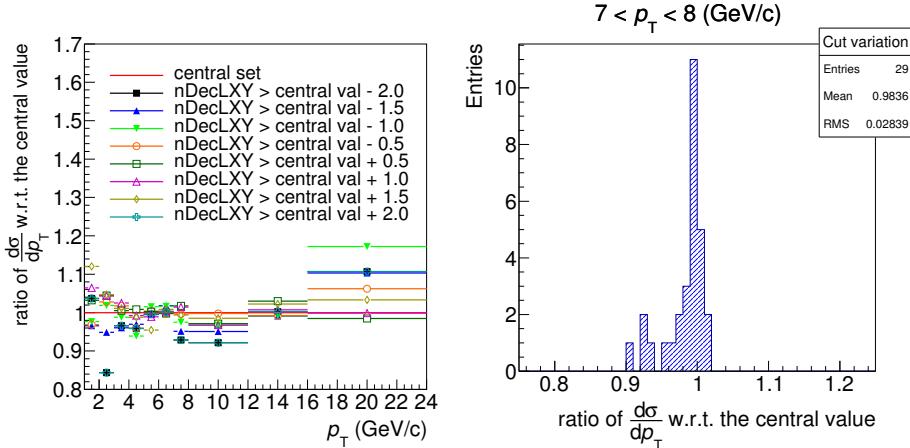


Figure 7.15. Left: ratio of the  $p_T$ -differential cross section obtained varying the cut on the normalised decay length  $xy$  relative to the central value as a function of  $p_T$ . Right: distribution of the ratio of the  $p_T$ -differential cross section varying all the topological cuts in the interval  $7 < p_T < 8$  GeV/c.

### 7.4.3 Systematic uncertainty due to the PID efficiency

Since the code for the PID selection was re-implemented for this analysis, the possible systematic uncertainty induced by the PID selection efficiency was re-evaluated as a cross-check. In the left panel of figure 7.16 an example of fit to the invariant-mass distribution for  $D^+$  candidates with  $4 < p_T < 5$  GeV/c obtained with and without the PID selections is shown. All the topological selections reported in table 7.1 are applied in both cases. In order to avoid as much as possible statistical fluctuations, for this test the Gaussian width in the fit has been fixed to the value obtained by fitting the invariant-mass distribution from the MC simulation. The loss of signal when the PID selection is applied relative to the case without PID is of about 6-10%, while the amount of combinatorial background is reduced by more than a factor 2-3 in all the  $p_T$  intervals except for  $1 < p_T < 2$  GeV/c. In this latter case, because of the stronger PID strategy adopted, the loss of signal is larger than 50%, while the background is reduced by a factor 30. The right panel of figure 7.16 shows the ratio between the  $p_T$ -differential cross sections obtained with and without the PID selection. Since the  $D^+$  candidates obtained with the PID selection are a sub-sample of those obtained without using the PID, the error on the ratio is assumed to be binomial. The ratio fluctuates around unity without any trend and therefore no systematic uncertainty has been assigned for the PID selection efficiency, as in the standard analysis.

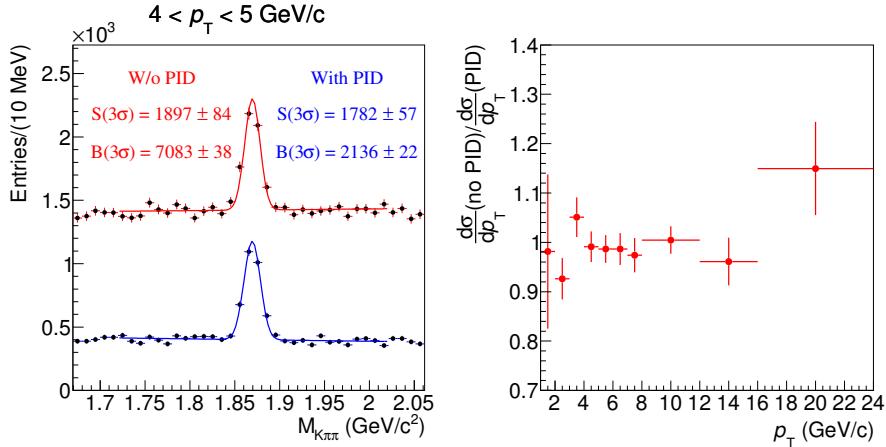


Figure 7.16. Left: comparison between the fit to the invariant-mass distribution obtained with (blue) and without (red) applying the PID selection in the interval  $4 < p_T < 5$  GeV/c. The Gaussian width is fixed to the value obtained from the MC simulation. Right: ratio of the  $p_T$ -differential cross section of prompt  $D^+$  obtained with and without the PID selection.

#### 7.4.4 Systematic uncertainty due to the generated $p_T$ shape

To compute the efficiency of prompt  $D^+$ -mesons, the generated distribution of prompt  $D^+$  has been reweighted in order to reproduce the shape of the FONLL  $p_T$  spectrum. That was done also for the efficiency of feed-down  $D^+$ , since it is needed for the computation of the fraction of prompt  $D^+$  with the  $N_b$  and  $f_c$  methods (see section 4.2.1). The weights used to reproduce the FONLL  $p_T$  spectrum for feed-down  $D^+$  have been obtained considering those of  $D^+$  from B-meson decays, as explained in section 5.4.2. On the other hand, the efficiencies obtained with the  $p_T$  shapes of prompt and feed-down  $D^+$  generated with Pythia have been considered for the evaluation of the systematic uncertainties. In the left panel of figure 7.17 the ratio between the efficiencies for prompt  $D^+$  obtained with the two different  $p_T$  shapes is shown. The effect is in general small and, in particular, it is larger at high  $p_T$ , because the  $p_T$  intervals considered are wider, and in the first bin, where the efficiency varies steeply with the  $D^+$  transverse momentum. The right panel of figure 7.17 shows the ratio between the fraction of prompt  $D^+$  evaluated with the  $N_b$  method, using the efficiencies for prompt and feed-down  $D^+$  obtained with the two different  $p_T$  shapes. The small variation observed indicates that the effect of the generated  $p_T$  shape of  $D^+$  from B-meson decays on the prompt  $D^+$  cross section is negligible. The values assigned as systematic uncertainty are reported in

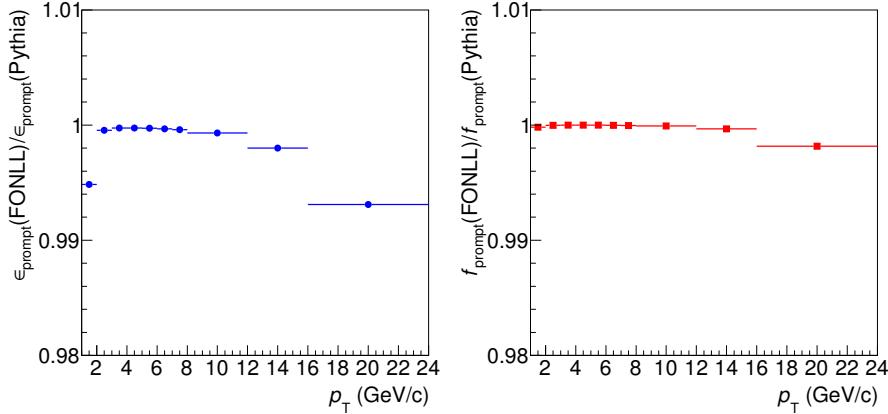


Figure 7.17. Left: Ratio of the efficiencies for prompt  $D^+$  obtained with the FONLL and Pythia  $p_T$  shapes. Right: ratio of the fraction of prompt  $D^+$  calculated using the efficiencies for prompt and feed-down  $D^+$  obtained with the FONLL and Pythia  $p_T$  shapes.

table 7.3.

#### 7.4.5 Systematic uncertainty due to the feed-down $D^+$ -meson subtraction

The fraction of prompt  $D^+$  obtained using the KFParticle for the reconstruction of the  $D^+$  candidates and applying the selections reported in table 7.1 is reported in figure 7.18, compared to the one obtained by applying the cuts used in the published result (see table 4.2). Both of them have been computed using the  $N_b$  method

$p_T$ (GeV/c)	[1,2]	[2,3]	[3,4]	[4,5]	[5,6]	[6,7]	[7,8]	[8,12]	[12,16]	[16,24]
raw yield	10%	8%	5%	5%	5%	5%	5%	5%	8%	8%
cut var	10%	10%	6%	6%	6%	5%	5%	5%	5%	5%
PID	negl.	negl.	negl.							
$p_T$ shape	1%	negl.	1%	1%						

Table 7.3. Summary of the systematic uncertainty for the analysis performed reconstructing the  $D^+$  candidates with the KFParticle package.

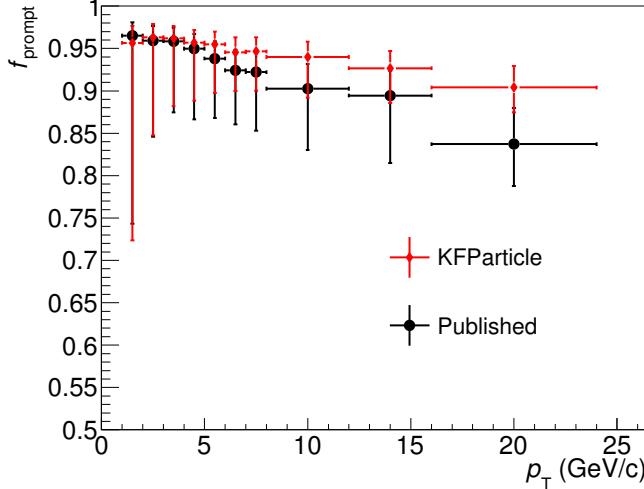


Figure 7.18. Comparison between the fraction of prompt  $D^+$  obtained using the KFParticle package for the reconstruction of the  $D^+$  candidates, applying the topological selections reported in table 7.1, and with the topological selections used in the standard analysis.  $f_{prompt}$  is computed with the FONLL-based methods in both cases.

for the central value and the  $f_c$  method to estimate a systematic uncertainty (in addition to the uncertainty due to the variation of the mass of the charm quark  $m_c$  and of the factorisation and renormalisation scales in the FONLL calculation), as explained in section 4.2.1. The cut on  $\chi^2/ndf$  after the application of the topological constraint at the primary vertex allowed to achieve a larger fraction of prompt  $D^+$  in the raw yield especially at high  $p_T$ , as shown in figure 7.1. The larger value of  $f_{prompt}$  provides also a reduction of the systematic uncertainty, mainly because for a fraction of prompt  $D^+$  close to unity, the values estimated with the  $N_b$  and  $f_c$  methods become more similar. The comparison between the systematic uncertainties on  $f_{prompt}$  for the two different sets of topological selections are reported in table 7.4.

#### Comparison to the selection on the maximum difference between the measured and expected impact parameter of the daughter tracks

As discussed in section 6.2.1, another powerful variable that allows one to discriminate between prompt and feed-down  $D^+$ -mesons is the maximum difference

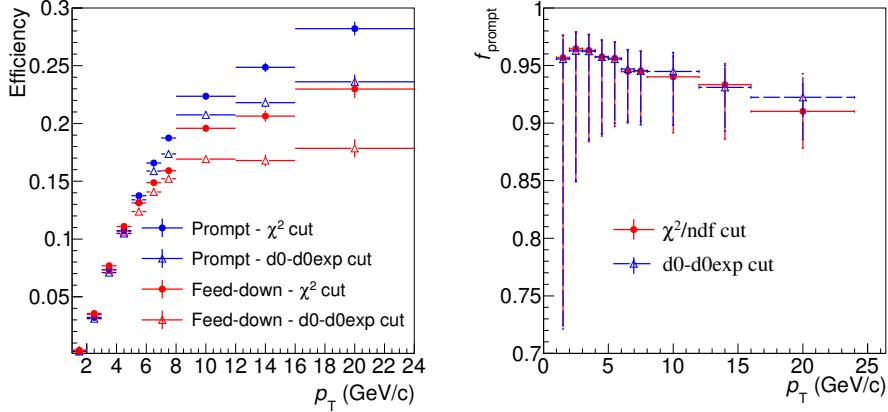


Figure 7.19. Comparison between the efficiencies of prompt and feed-down  $D^+$  (left) and the  $D^+$  prompt fraction (right) obtained applying the selection on the  $\chi^2/\text{ndf}$  and on the maximum difference between the measured and expected impact parameter of the daughter tracks.

between the measured and expected impact parameter of the daughter tracks. In order to compare the effect of a selection on this variable with the one obtained with the cut on the  $\chi^2/\text{ndf}$  of the reconstructed  $D^+$  candidate after the application of the topological constraint, the  $p_T$ -differential cross section has been recomputed replacing the cut on the  $\chi^2/\text{ndf}$  with that on the maximum difference between the measured and expected track impact parameter, set to  $[-2.5\sigma, 2.5\sigma]$  for all the  $p_T$  intervals of the analysis. In figure 7.19 the efficiencies of prompt and feed-down  $D^+$  (left) and the  $D^+$  prompt fraction (right) obtained with the two different selections are compared. The prompt fraction is very similar in the two cases, while the efficiency is slightly higher for both prompt and feed-down  $D^+$ -mesons if the cut on  $\chi^2/\text{ndf}$  is applied. The similar result obtained is justified by the correlation between the two quantities, as confirmed by the scatter plot of the maximum difference between the measured and expected impact parameter of the daughter tracks as a function of the  $\chi^2/\text{ndf}$  of the  $D^+$  candidates with  $3 < p_T < 4$  GeV/c, which is reported in the left panel of figure 7.20. The right panel of the same figure shows the ratio between the  $p_T$ -differential cross section of prompt  $D^+$  obtained applying the two different cuts. This ratio fluctuates around unity without any particular trend, indicating that the application of one selection does not introduce a bias with respect to the result obtained applying the other one.

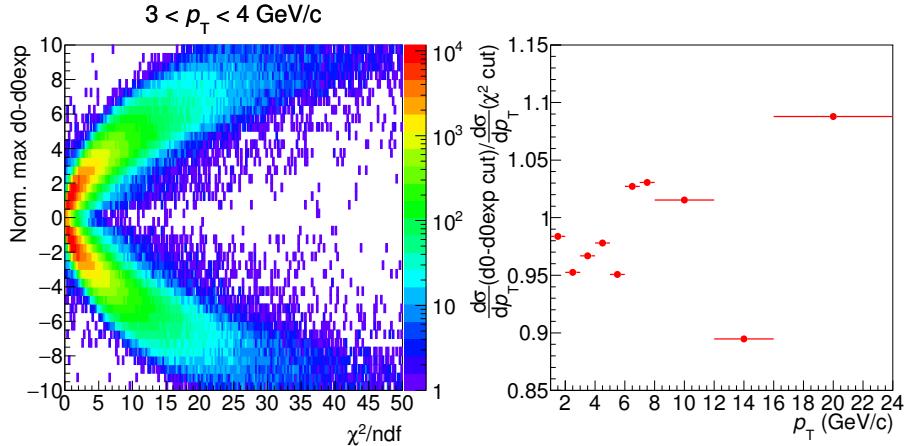


Figure 7.20. Left: scatter plot of the maximum difference between the measured and expected impact parameter of the daughter tracks vs. the  $\chi^2/ndf$  for the  $D^+$  candidates with  $3 < p_T < 4$  GeV/c. Right: ratio between the  $p_T$ -differential cross section of prompt  $D^+$  obtained applying the cut on the maximum difference between the measured and expected impact parameter of the daughter tracks and the one obtained applying the  $\chi^2/ndf$  cut.

## 7.5 Results

The main goal of this study was to achieve an improved subtraction of the feed-down  $D^+$ -mesons in the extracted raw yields. Hence, the result was compared to the one obtained applying the topological selections used in the standard analysis. Furthermore, since the selection of the  $D^+$  candidates was performed with a different vertex-reconstruction algorithm, the  $p_T$ -differential cross section was

$p_T$ (GeV/c)	[1,2]	[2,3]	[3,4]	[4,5]	[5,6]	[6,7]	[7,8]	[8,12]	[12,16]	[16,24]
Published (%)	+2	+2	+2	+2	+2	+3	+3	+3	+3	+5
	-22	-12	-9	-9	-8	-7	-7	-8	-9	-6
KFParticle (%)	+2	+2	+2	+2	+2	+2	+2	+2	+2	+3
	-24	-12	-8	-7	-6	-5	-5	-5	-5	-3

Table 7.4. Comparison between the systematic uncertainty on the  $D^+$  prompt fraction obtained using the KFParticle package for the reconstruction of the  $D^+$  candidates and the one obtained in the standard analysis.

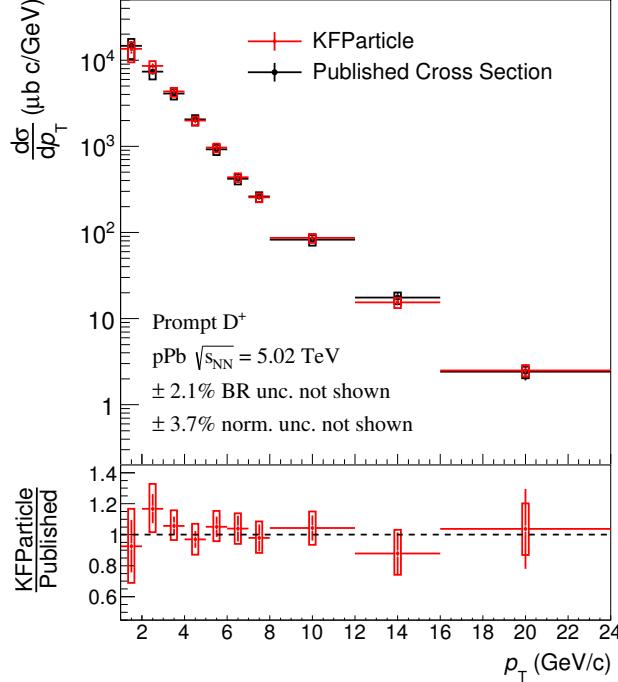


Figure 7.21. Comparison between the  $p_T$ -differential production cross section of prompt  $D^+$ -mesons in p-Pb collisions at  $\sqrt{s_{NN}} = 5.02 \text{ TeV}$  obtained using the KFParticle package for the reconstruction of the  $D^+$  candidates compared to the published result [36].

compared to the published result, in order to check the consistency of the analysis. The  $p_T$ -differential production cross section of  $D^+$  mesons in p-Pb collisions at  $\sqrt{s_{NN}} = 5.02 \text{ TeV}$  obtained using the KFParticle package for the reconstruction of the  $D^+$  candidates is shown in figure 7.21, compared to the one obtained with the standard approach, which was published by the ALICE Collaboration [36]. The central values have been computed using the formula reported in equation 4.8. The vertical bars represent the statistical uncertainty, while the boxes the systematic uncertainty, which was computed considering all the sources presented in section 7.4 as uncorrelated and therefore summed in quadrature. In the bottom panel of figure 7.21 the ratio between the result obtained using the vertex-reconstruction algorithm based on the Kalman filter and the published one is shown. The statistical

uncertainty have been treated as uncorrelated in the ratio, although some degree of correlation is present since most of the selections applied were similar. However, the estimation of the correlation between the two results were beyond the scope of this study. Also the systematic uncertainties have been treated as uncorrelated in the ratio, except for the systematic uncertainties due to the tracking efficiency, the beauty feed-down subtraction, the integrated luminosity and the branching ratio, which have been considered as fully correlated. Some degree of correlation is also present among the systematic uncertainty due to the topological selection efficiency and the raw yield extraction, since the sub-samples of  $D^+$  candidates selected for the two analysis are not completely disjoint. Therefore, the uncertainties on the ratio shown in the bottom panel of figure 7.21 are slightly overestimated. Despite that, the ratio is found to fluctuate around unity, showing that the two results are in agreement between each other and no systematic difference is observed between the two results.

# 8

## Conclusions

This thesis was focused on the measurement of the production of prompt  $D^+$ -mesons in p-Pb collisions at  $\sqrt{s_{NN}} = 5.02$  TeV with the ALICE experiment at the LHC. All the analyses presented have been performed exploiting the full reconstruction of the  $D^+$ -mesons in the purely hadronic decay channel  $D^+ \rightarrow K^-\pi^+\pi^+$ . In particular, three different techniques for the measurement of the prompt and feed-down contributions to the  $D^+$ -meson production have been developed. They consists in:

- An unbinned log-likelihood fit to the  $D^+$  candidates impact-parameter distributions.
- The usage of different sets of selection criteria on the  $D^+$  decay topology.
- The reconstruction of the  $D^+$  candidates performed with a different vertex-reconstruction algorithm based on the Kalman Filter, followed by a selection on the  $\chi^2/ndf$  after applying a topological constraint on the  $D^+$  production point.

They all exploit the larger displacement of the feed-down  $D^+$ -mesons from the primary vertex, with respect to the prompt  $D^+$  mesons. In figure 8.1, the  $p_T$ -differential production cross section of  $D^+$ -mesons obtained with the three different methods is compared to the published one. The vertical bars represent the statistical uncertainties, while the boxes the systematic ones. All the results are

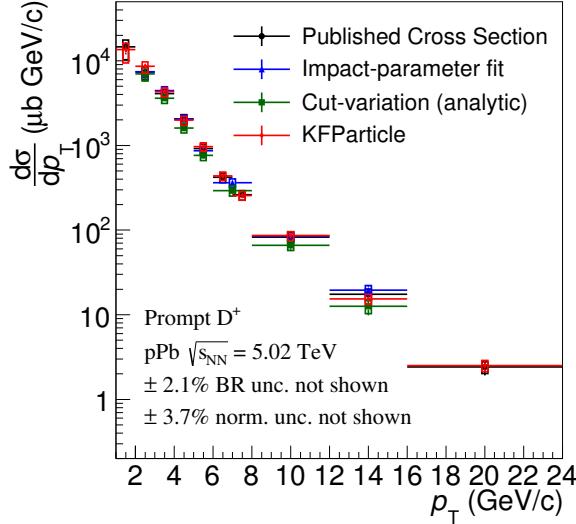


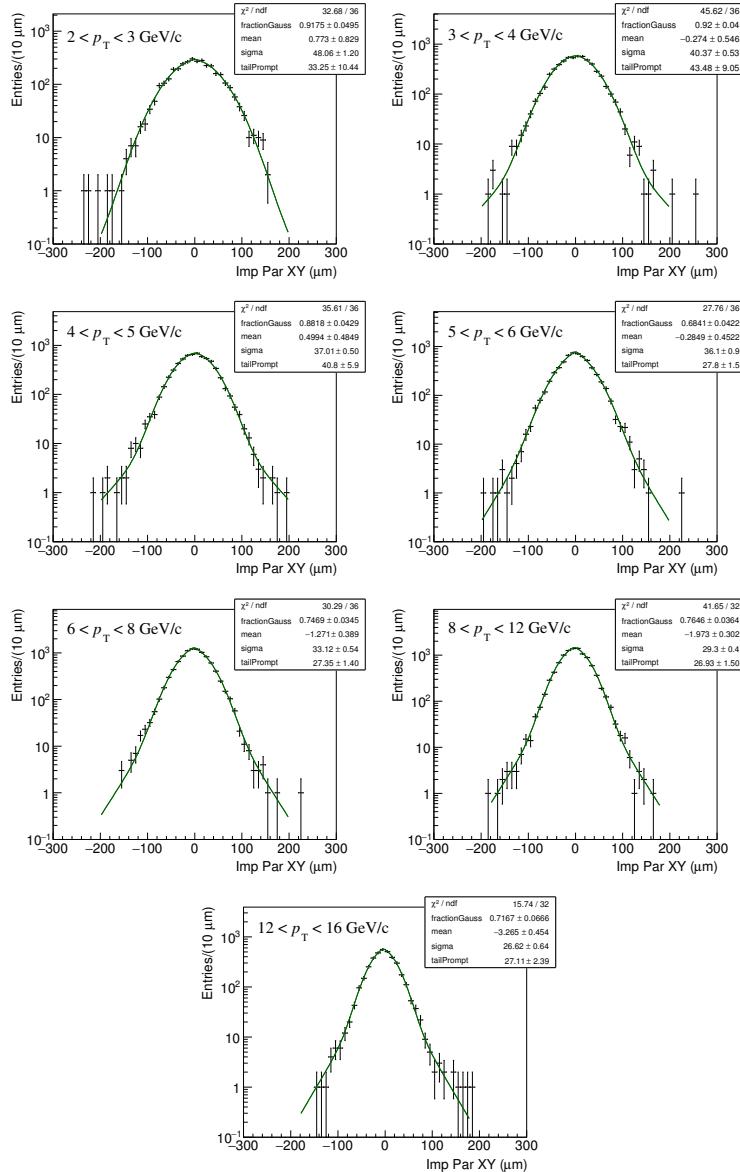
Figure 8.1. Comparison among the  $p_T$ -differential production cross section of prompt  $D^+$ -mesons in  $p$ -Pb collisions at  $\sqrt{s_{NN}} = 5.02 \text{ TeV}$  obtained with all the analyses discussed in this thesis and the one published by the ALICE Collaboration [36].

found to be compatible among each other within the experimental uncertainties. The first two methods have the advantage to be fully data-driven and therefore not to rely on perturbative QCD calculations for the determination of the fraction of prompt  $D^+$ . These measurements will be improved starting from the Run2 at LHC thanks to larger data samples expected and even more with the upgrade of the Inner Tracking System of ALICE, planned for the year 2019, thanks to the improved vertex-position and impact-parameter resolutions. On the other hand, the Kalman-filter based analysis allowed to reduce significantly the feed-down  $D^+$  contribution in the raw yield. Furthermore, there is the possibility to exploit the  $\chi^2/ndf$  obtained after applying the topological constraint for data-driven analyses, such as the cut-variation method presented in Chapter 6. Finally, the vertex-reconstruction package based on the Kalman filter could also be used for other analyses, such as the measurement of strange hadrons (e.g.  $K_s^0$ ,  $\Lambda$  and  $\Xi$ ) or hypernuclei, which have a larger mean proper decay length (e.g.  $(c\tau)_\Lambda = 7.89 \text{ cm}$  [56]) and therefore, if they are charged, the curvature of their trajectory due to the magnetic field cannot be neglected.

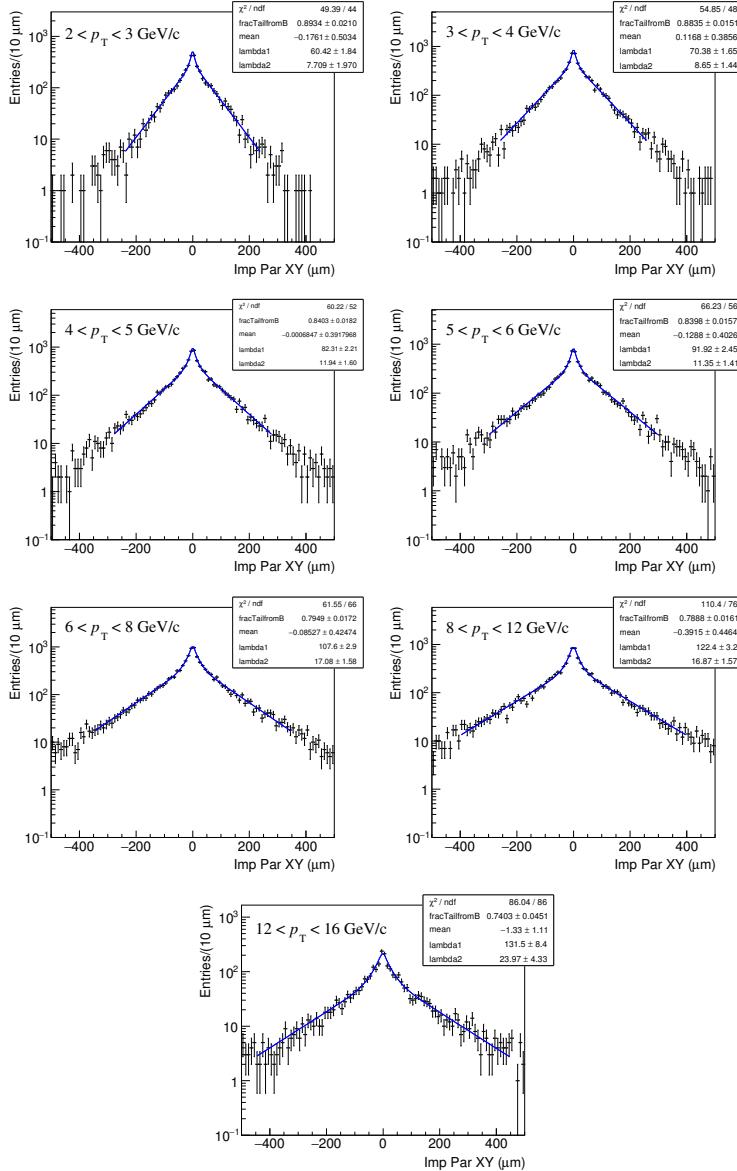
# A

Additional figures

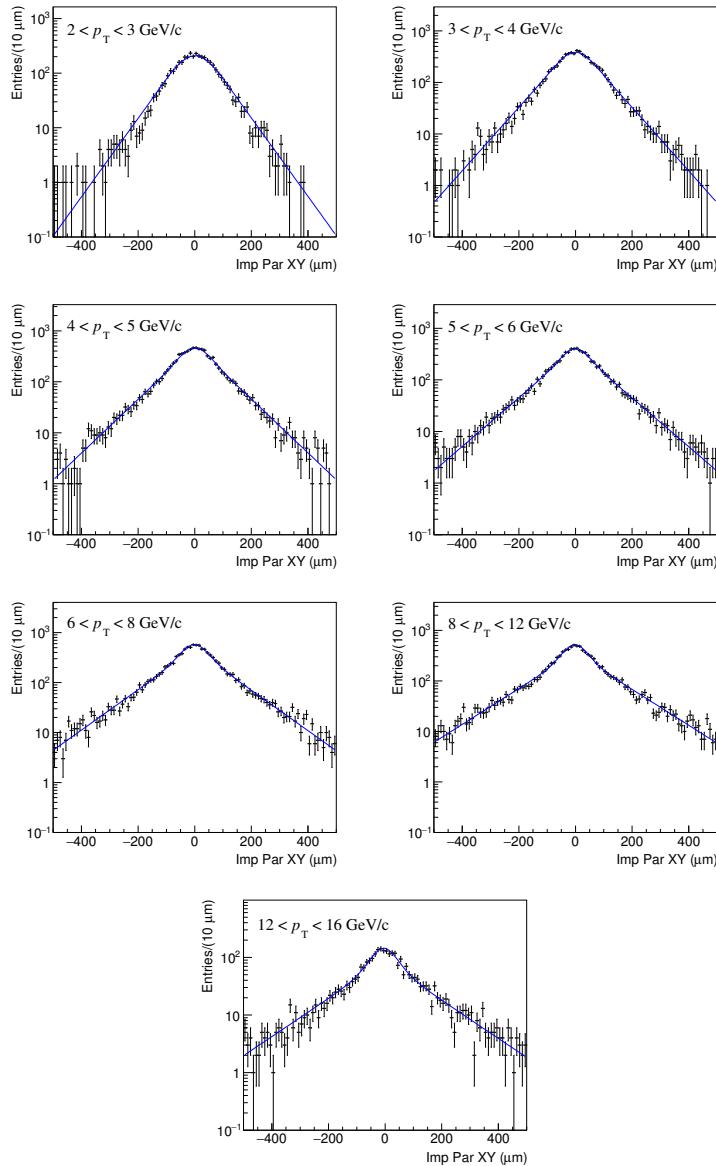
## A.1 Fits to the MC impact-parameter distributions of prompt D<sup>+</sup>-mesons



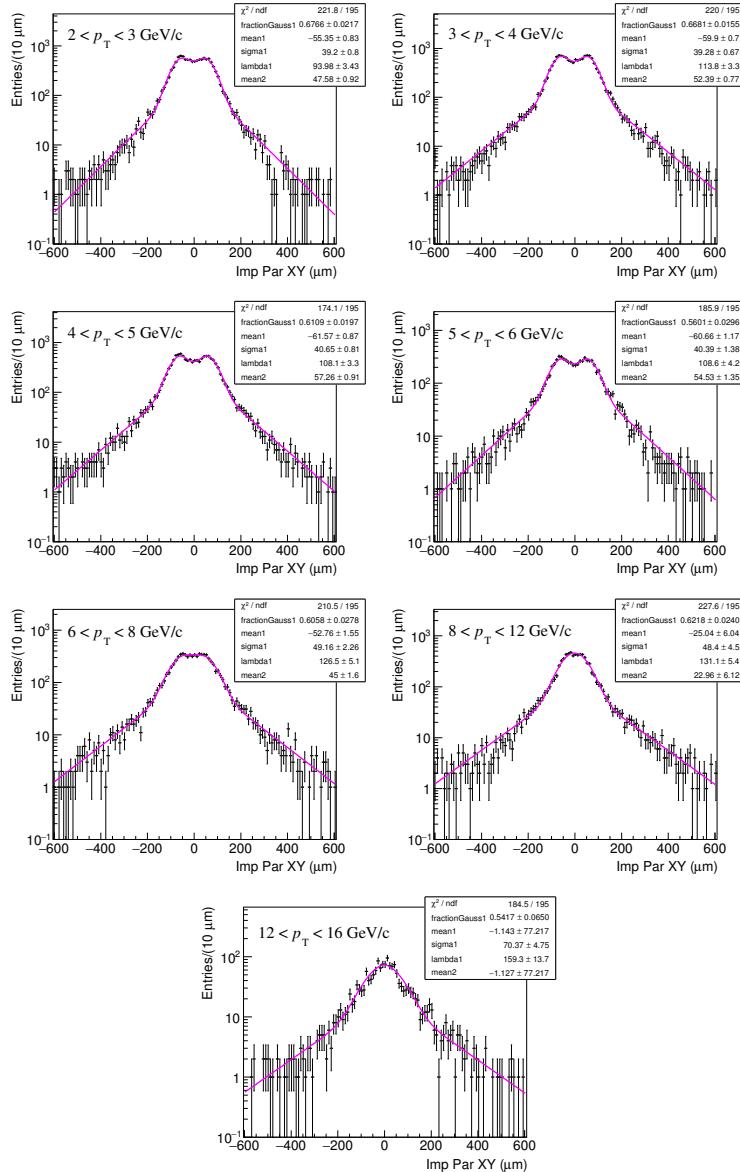
## A.2 Fits to the MC true impact-parameter distributions of feed-down $D^+$ -mesons



### A.3 MC impact-parameter distributions of feed-down D<sup>+</sup>-mesons

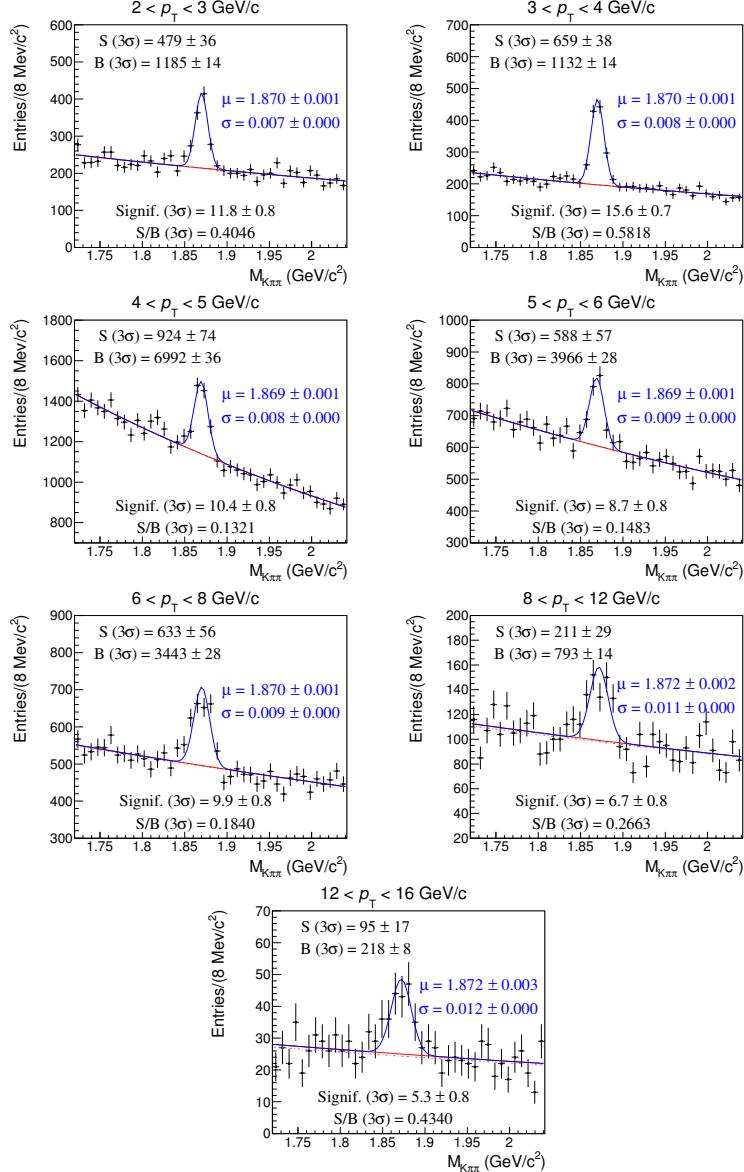


## A.4 Fits to the impact-parameter distributions of the combinatorial background

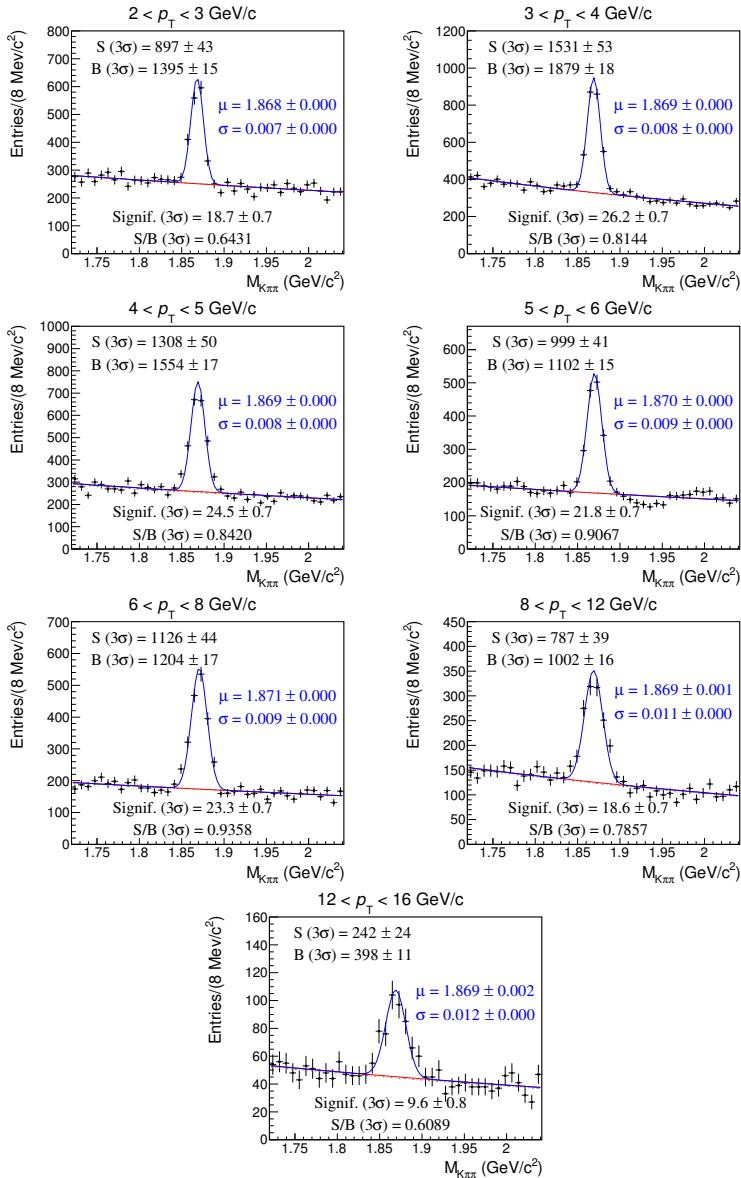


## A.5 Cut-variation analysis plots

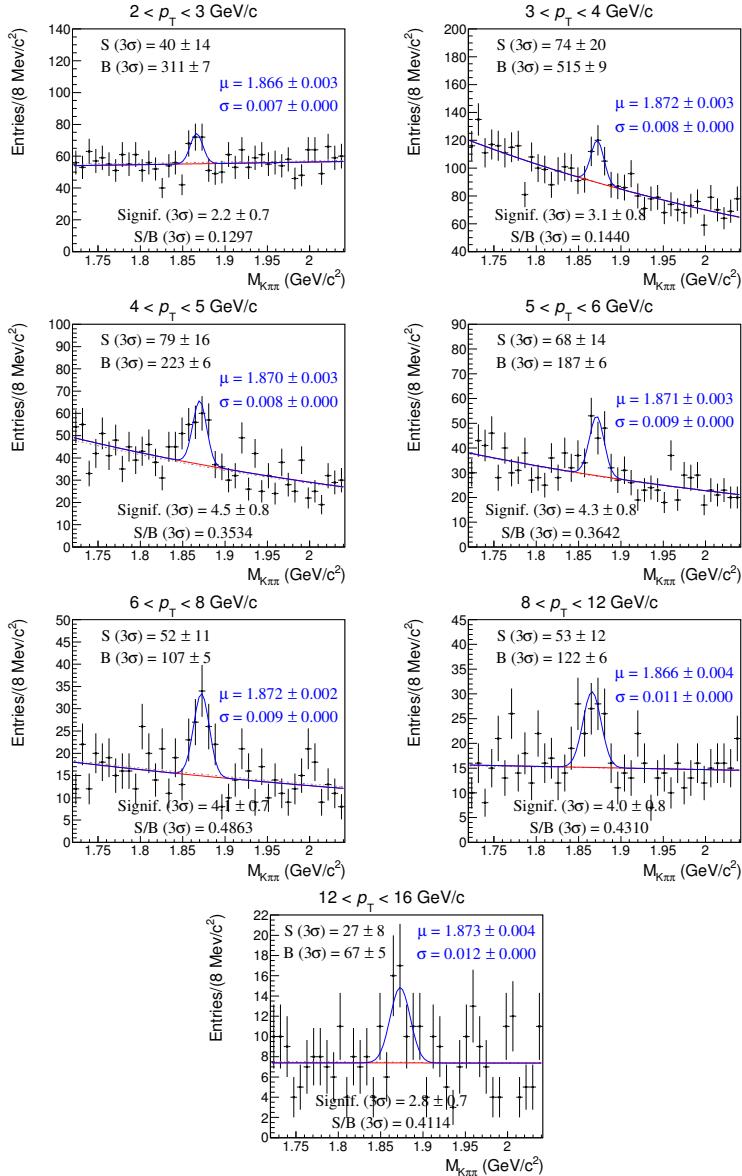
### A.5.1 Mass fits prompt-enhanced set



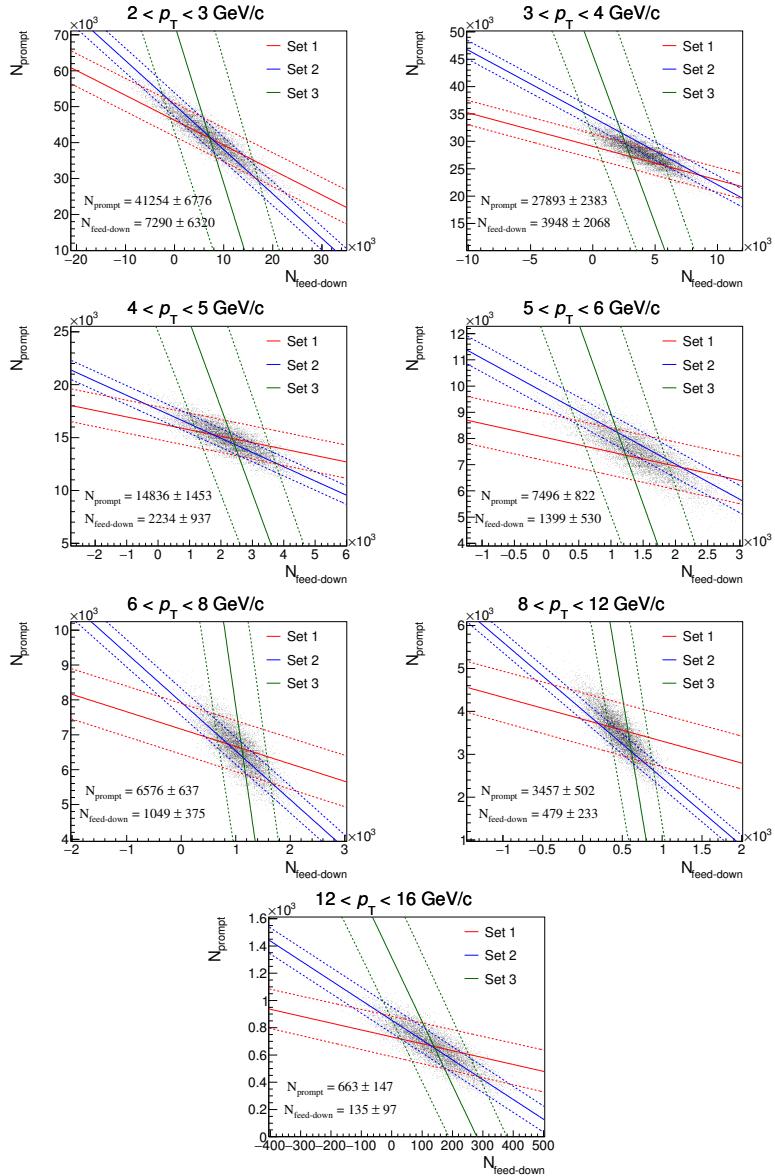
### A.5.2 Mass fits mixed set



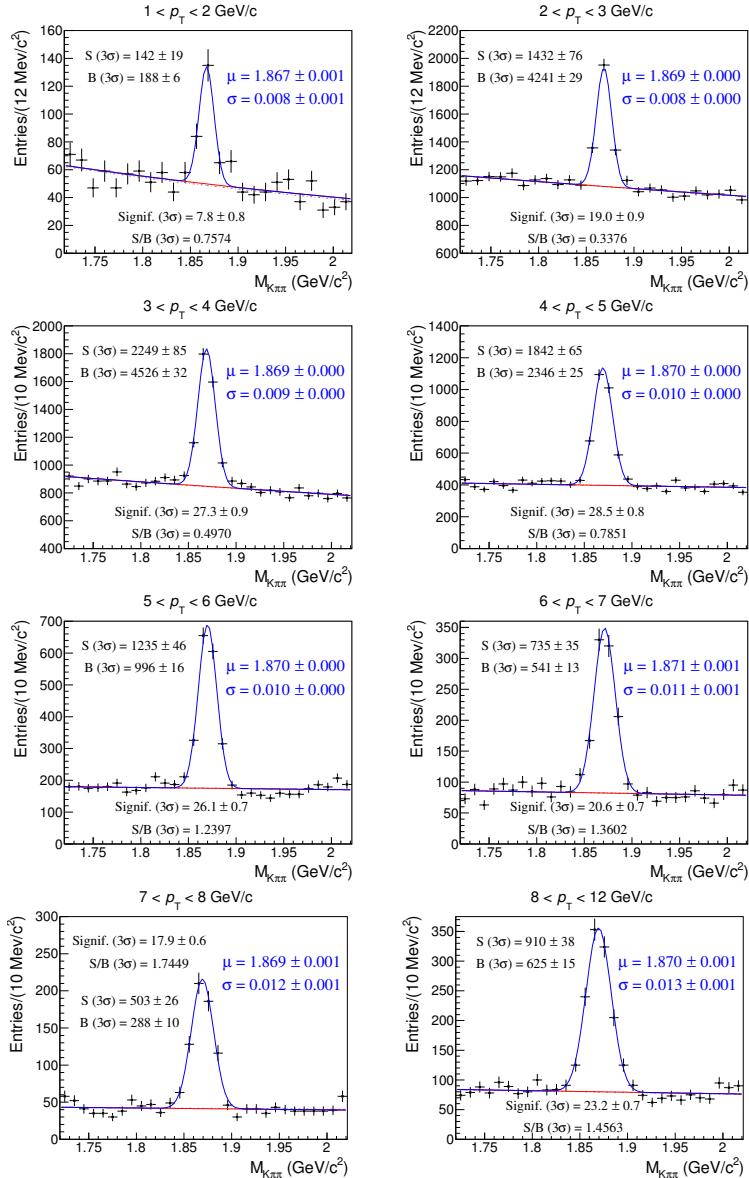
### A.5.3 Mass fits feed-down-enhanced set



#### A.5.4 Incentre determination

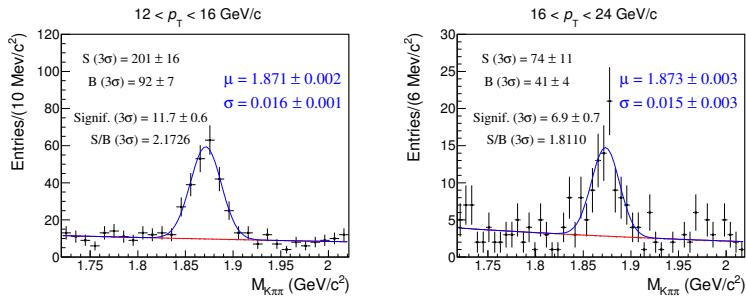


## A.6 Kalman filter analysis invariant-mass fits



### A.6 – Kalman filter analysis invariant-mass fits

---





## Bibliography

- [1] J. Adam *et al.* [ALICE Collaboration], *D-meson production in p-Pb collisions at  $\sqrt{s_{NN}} = 5.02 \text{ TeV}$  and in pp collisions at  $\sqrt{s} = 7 \text{ TeV}$* , arXiv:1605.07569 [nucl-ex].
- [2] Ian J.R. Aitchinson and Anthony J.G. Hey, *Gauge theories in particle physics*, 4th ed. (CRC Press, 2012).
- [3] S. Bethke, *Experimental tests of asymptotic freedom*, Prog. Part. Nucl. Phys. **58**, 351 (2013) [hep-ex/0606035].
- [4] K. Johnson, *The M.I.T. Bag Model*, Acta Phys. Polon. B **6** (1975) 865.
- [5] A. Chodos, R. L. Jaffe, K. Johnson, C. B. Thorn and V. F. Weisskopf, *A New Extended Model of Hadrons*, Phys. Rev. D **9** (1974) 3471.
- [6] K. Yagi, T. Hatsuda and Y. Miake, *Quark-gluon plasma: From big bang to little bang*, Camb. Monogr. Part. Phys. Nucl. Phys. Cosmol. **23** (2005) 1.
- [7] S. Hands, *The phase diagram of QCD*, Contemp.Phys.42:209-225 (2001).
- [8] C. Wong, *Introduction to High-Energy Heavy-Ion Collisions*, World Scientific Publishing, (1994).
- [9] F. Karsch, *Lattice QCD at high temperature and density*, Lect. Notes Phys. **583** (2002) 209 [hep-lat/0106019].
- [10] S. Borsanyi, G. Endrodi, Z. Fodor, A. Jakovac, S. D. Katz, S. Krieg, C. Ratti and K. K. Szabo, *The QCD equation of state with dynamical quarks*, JHEP **1011** (2010) 077 doi:10.1007/JHEP11(2010)077 [arXiv:1007.2580 [hep-lat]].
- [11] J. D. Bjorken, *Highly Relativistic Nucleus-Nucleus Collisions: The Central Rapidity Region*, Phys. Rev. D **27** (1983) 140. doi:10.1103/PhysRevD.27.140.
- [12] 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 doi:10.1146/annurev.nucl.57.090506.123020 [nucl-ex/0701025].
- [13] E. Abbas *et al.* [ALICE Collaboration], *Centrality dependence of the pseudo-rapidity density distribution for charged particles in Pb-Pb collisions at  $\sqrt{s_{NN}} = 2.76 \text{ TeV}$* , Phys. Lett. B **726** (2013) 610 doi:10.1016/j.physletb.2013.09.022 [arXiv:1304.0347 [nucl-ex]].
- [14] J. Adam *et al.* [ALICE Collaboration], *Centrality dependence of the charged-particle multiplicity density at mid-rapidity in Pb-Pb collisions at  $\sqrt{s_{NN}} = 5.02 \text{ TeV}$* , arXiv:1512.06104 [nucl-ex].

- [15] J. Stachel, A. Andronic, P. Braun-Munzinger and K. Redlich, *Confronting LHC data with the statistical hadronisation model*, J. Phys. Conf. Ser. **509** (2014) 012019 doi:10.1088/1742-6596/509/1/012019 [arXiv:1311.4662 [nucl-th]].
- [16] P. Braun-Munzinger and J. Wambach, *The Phase Diagram of Strongly-Interacting Matter*, Rev. Mod. Phys. **81** (2009) 1031 doi:10.1103/RevModPhys.81.1031 [arXiv:0801.4256 [hep-ph]].
- [17] R. Pregheonella [ALICE Collaboration], *Spectra of identified hadrons with the ALICE detector in pp and Pb-Pb collisions at the LHC*, arXiv:1203.5904 [hep-ex].
- [18] 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 doi:10.1103/PhysRevC.48.2462 [nucl-th/9307020].
- [19] R. Snellings [ALICE Collaboration], *Anisotropic flow at the LHC measured with the ALICE detector*, J. Phys. G **38** (2011) 124013 doi:10.1088/0954-3899/38/12/124013 [arXiv:1106.6284 [nucl-ex]].
- [20] C. Shen, U. Heinz, P. Huovinen and H. Song, *Radial and elliptic flow in Pb+Pb collisions at the Large Hadron Collider from viscous hydrodynamic*, Phys. Rev. C **84** (2011) 044903 doi:10.1103/PhysRevC.84.044903 [arXiv:1105.3226 [nucl-th]].
- [21] B. Abelev *et al.* [ALICE Collaboration], *Centrality Dependence of Charged Particle Production at Large Transverse Momentum in Pb-Pb Collisions at  $\sqrt{s_{NN}} = 2.76$  TeV*, Phys. Lett. B **720** (2013) 52 doi:10.1016/j.physletb.2013.01.051 [arXiv:1208.2711 [hep-ex]].
- [22] J. Adams *et al.* [STAR Collaboration], *Direct observation of dijets in central Au+Au collisions at  $\sqrt{s_{NN}} = 200$  GeV*, Phys. Rev. Lett. **97** (2006) 162301 doi:10.1103/PhysRevLett.97.162301 [nucl-ex/0604018].
- [23] G. Aad *et al.* [ATLAS Collaboration], *Observation of a Centrality-Dependent Dijet Asymmetry in Lead-Lead Collisions at  $\sqrt{s_{NN}} = 2.77$  TeV with the ATLAS Detector at the LHC*, Phys. Rev. Lett. **105** (2010) 252303 doi:10.1103/PhysRevLett.105.252303 [arXiv:1011.6182 [hep-ex]].
- [24] S. Salur [CMS Collaboration], *Dijet imbalance in 2.76-TeV PbPb collisions in CMS*, J. Phys. Conf. Ser. **389** (2012) 012003. doi:10.1088/1742-6596/389/1/012003.
- [25] [CMS Collaboration], *Quarkonium production in PbPb collisions*, CMS-PAS-HIN-10-006.
- [26] B. B. Abelev *et al.* [ALICE Collaboration], *Centrality, rapidity and transverse momentum dependence of  $J/\psi$  suppression in Pb-Pb collisions at  $\sqrt{s_{NN}}=2.76$  TeV*, Phys. Lett. B **734** (2014) 314 doi:10.1016/j.physletb.2014.05.064 [arXiv:1311.0214 [nucl-ex]].
- [27] V. N. Gribov and L. N. Lipatov, *Deep inelastic e p scattering in perturbation theory*, Sov. J. Nucl. Phys. **15** (1972) 438 [Yad. Fiz. **15** (1972) 781].

- [28] G. Altarelli and G. Parisi, *Asymptotic Freedom in Parton Language*, Nucl. Phys. B **126** (1977) 298, doi:10.1016/0550-3213(77)90384-4.
- [29] Y. L. Dokshitzer, *Calculation of the Structure Functions for Deep Inelastic Scattering and  $e^+ e^-$  Annihilation by Perturbation Theory in Quantum Chromodynamics*, Sov. Phys. JETP **46** (1977) 641 [Zh. Eksp. Teor. Fiz. **73** (1977) 1216].
- [30] B. Abelev *et al.* [ALICE Collaboration], *Measurement of charm production at central rapidity in proton-proton collisions at  $\sqrt{s} = 7$  TeV*, JHEP **1201** (2012) 128 doi:10.1007/JHEP01(2012)128 [arXiv:1111.1553 [hep-ex]].
- [31] R. Aaij *et al.* [LHCb Collaboration], *Measurement of  $J/\psi$  production in  $pp$  collisions at  $\sqrt{s} = 7$  TeV*, Eur. Phys. J. C **71** (2011) 1645 doi:10.1140/epjc/s10052-011-1645-y [arXiv:1103.0423 [hep-ex]].
- [32] N. Armesto, *Nuclear shadowing*, J. Phys. G **32** (2006) R367 doi:10.1088/0954-3899/32/11/R01 [hep-ph/0604108].
- [33] K. Rith, *Present Status of the EMC effect*, Subnucl. Ser. **51** (2015) 431 [arXiv:1402.5000 [hep-ex]].
- [34] N. Carrer and A. Dainese, *Charm and beauty production at LHC*, hep-ph/0311225.
- [35] D. Antreasyan, J. W. Cronin, H. J. Frisch, M. J. Shochet, L. Kluberg, P. A. Piroue and R. L. Sumner, *Production of Hadrons at Large Transverse Momentum in 200-GeV, 300-GeV and 400-GeV  $p-p$  and  $p-n$  Collisions*, Phys. Rev. D **19** (1979) 764, doi:10.1103/PhysRevD.19.764.
- [36] B. B. Abelev *et al.* [ALICE Collaboration], *Measurement of prompt  $D$ -meson production in  $p - Pb$  collisions at  $\sqrt{s_{NN}} = 5.02$  TeV*, Phys. Rev. Lett. **113** (2014) no.23, 232301 doi:10.1103/PhysRevLett.113.232301 [arXiv:1405.3452 [nucl-ex]].
- [37] M. L. Mangano, P. Nason and G. Ridolfi, *Heavy quark correlations in hadron collisions at next-to-leading order*, Nucl. Phys. B **373** (1992) 295, doi:10.1016/0550-3213(92)90435-E.
- [38] K. J. Eskola, H. Paukkunen and C. A. Salgado, *EPS09: A New Generation of NLO and LO Nuclear Parton Distribution Functions*, JHEP **0904** (2009) 065, doi:10.1088/1126-6708/2009/04/065, [arXiv:0902.4154 [hep-ph]].
- [39] D. Stump, J. Huston, J. Pumplin, W. K. Tung, H. L. Lai, S. Kuhlmann and J. F. Owens, *Inclusive jet production, parton distributions, and the search for new physics*, JHEP **0310** (2003) 046, doi:10.1088/1126-6708/2003/10/046, [hep-ph/0303013].
- [40] H. Fujii and K. Watanabe, *Heavy quark pair production in high energy  $pA$  collisions: Open heavy flavors*, Nucl. Phys. A **920** (2013) 78, doi:10.1016/j.nuclphysa.2013.10.006, [arXiv:1308.1258 [hep-ph]].
- [41] R. Aaij *et al.* [LHCb Collaboration], *Study of  $J/\psi$  production and cold nuclear matter effects in  $pPb$  collisions at  $\sqrt{s_{NN}} = 5$  TeV*, JHEP **1402** (2014) 072, doi:10.1007/JHEP02(2014)072, [arXiv:1308.6729 [nucl-ex]].

- [42] 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, doi:10.1016/S0550-3213(96)00581-0, [hep-ph/9608322].
- [43] M. Djordjevic and M. Gyulassy, *Where is the charm quark energy loss at RHIC?*, Phys. Lett. B **560** (2003) 37, doi:10.1016/S0370-2693(03)00327-7, [nucl-th/0302069].
- [44] A. Rossi [ALICE Collaboration], *Measurement of  $D$ -meson production in  $pp$ ,  $p$ - $Pb$  and  $Pb$ - $Pb$  collisions with ALICE at the LHC*, Nucl. Phys. A **932** (2014) 51, doi:10.1016/j.nuclphysa.2014.09.051, [arXiv:1409.4001 [hep-ex]].
- [45] J. Adam *et al.* [ALICE Collaboration], *Transverse momentum dependence of  $D$ -meson production in  $Pb$ - $Pb$  collisions at  $\sqrt{s_{NN}} = 2.76$  TeV*, JHEP **1603** (2016) 081, doi:10.1007/JHEP03(2016)081, [arXiv:1509.06888 [nucl-ex]].
- [46] L. Adamczyk *et al.* [STAR Collaboration], *Observation of  $D^0$  Meson Nuclear Modifications in  $Au+Au$  Collisions at  $\sqrt{s_{NN}} = 200$  GeV*, Phys. Rev. Lett. **113** (2014) no.14, 142301, doi:10.1103/PhysRevLett.113.142301, [arXiv:1404.6185 [nucl-ex]].
- [47] L. Evans and P. Bryant, *LHC Machine*, JINST **3** (2008) S08001.
- [48] K. Aamodt *et al.* [ALICE Collaboration], *The ALICE experiment at the CERN LHC*, JINST **3** (2008) S08002.
- [49] G. J. L. Nooren, *The inner tracking system of ALICE*, Int. J. Mod. Phys. E **16** (2007) 2445.
- [50] H. R. Schmidt [ALICE Collaboration], *The ALICE TPC: Status and perspectives*, J. Phys. Conf. Ser. **230** (2010) 012023.
- [51] A. Alici [ALICE Collaboration], *The MRPC-based ALICE Time-Of-Flight detector: status and performance*, Nucl. Instrum. Meth. A **706** (2013) 29, doi:10.1016/j.nima.2012.05.004, [arXiv:1203.5976 [physics.ins-det]].
- [52] ALICE collaboration, *ALICE trigger data-acquisition high-level trigger and control system: Technical Design Report*, CERN-LHCC-2003-062, <http://cdsweb.cern.ch/record/684651>.
- [53] R. Fruhwirth, *Application of Kalman filtering to track and vertex fitting*, Nucl. Instrum. Meth. A **262** (1987) 444, doi:10.1016/0168-9002(87)90887-4.
- [54] B. B. Abelev *et al.* [ALICE Collaboration], *Performance of the ALICE Experiment at the CERN LHC*, Int. J. Mod. Phys. A **29** (2014) 1430044, [arXiv:1402.4476 [nucl-ex]].
- [55] W. Blum, W. Riegler, L. Rolandi, *Particle Detection with Drift Chambers*, 2nd ed. (Springer-Verlag, 2008).
- [56] K. A. Olive *et al.* [Particle Data Group Collaboration], *Review of Particle Physics*, Chin. Phys. C **38** (2014) 090001, doi:10.1088/1674-1137/38/9/090001.
- [57] K. A. Olive *et al.* [Particle Data Group Collaboration], *Review of Particle Physics (2015 update)*, <http://pdg.lbl.gov>

- [58] B. B. Abelev *et al.* [ALICE Collaboration], *Measurement of visible cross sections in proton-lead collisions at  $\sqrt{s_{NN}} = 5.02$  TeV in van der Meer scans with the ALICE detector*, JINST **9** (2014) no.11, P11003, doi:10.1088/1748-0221/9/11/P11003, [arXiv:1405.1849 [nucl-ex]].
- [59] R. Russo, *Measurement of  $D^+$  meson production in  $p$ -Pb collisions with the ALICE detector*, arXiv:1511.04380 [nucl-ex].
- [60] D. J. Lange, *The EvtGen particle decay simulation package*, Nucl. Instrum. Meth. A **462** (2001) 152, doi:10.1016/S0168-9002(01)00089-4.
- [61] P. Z. Skands, *The Perugia Tunes*, arXiv:0905.3418 [hep-ph].
- [62] 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** (1991) 3501, doi:10.1103/PhysRevD.44.3501.
- [63] The ALICE offline framework, *AliRoot*, <http://aliweb.cern.ch/Offline/>.
- [64] R. Brun, F. Rademakers, Nucl. Instrum. Meth. A389 (1997) 81-86, *ROOT*, <https://root.cern.ch/>.
- [65] ALICE Collaboration, *Conceptual Design Report for the Upgrade of the ALICE ITS*, CERN-LHCC-2012-005, LHCC-G-159, 2012.
- [66] S. Siddhanta [ALICE Collaboration], *The upgrade of the ALICE Inner Tracking System*, Nuovo Cim. C **38** (2015) no.1, 57, doi:10.1393/ncc/i2015-15057-6.
- [67] B. Abelev *et al.* [ALICE Collaboration], *Technical Design Report for the Upgrade of the ALICE Inner Tracking System*, J. Phys. G **41** (2014) 087002, doi:10.1088/0954-3899/41/8/087002.
- [68] I. Kisel, I. Kulakov and M. Zyzak, *Standalone first level event selection package for the CBM experiment*, doi:10.1109/RTC.2012.6418385.
- [69] P. Senger, *The Compressed Baryonic Matter Experiment at FAIR*, JPS Conf. Proc. **8** (2015) 022001, doi:10.7566/JPSCP.8.022001.
- [70] S. Gorbunov, I. Kisel, *Secondary vertex fit based on the Kalman filter*, CBM-SOFT-note-2007-003 (2007).



## Acknowledgements

Sembra ieri che ho iniziato il percorso che mi ha portato alla scrittura di questa tesi e, se guardo indietro, mi sembra incredibile pensare a quante cose abbia imparato e quante altre, invece, debba ancora impararne. Per questo mi sento di ringraziare tante persone che, in modi diversi, mi hanno aiutato, supportato e mi hanno accompagnato durante questi anni.

Per prima cosa vorrei esprimere un sincero ringraziamento alla mia famiglia. A mia madre Mariateresa e mio padre Franco, che mi hanno permesso di studiare ciò che più mi piace e hanno sempre creduto in me, e a mio fratello Roberto, che mi ha aiutato a comprendere un punto di vista diverso dal mio, dal momento che (purtroppo!) è ingegnere. Vorrei poi ringraziare Giulia e Riccardo, che mi hanno aiutato a superare i momenti più difficili e faticosi sostenendomi e, grazie ai quali, questi anni sono stati più piacevoli di quanto potessi sperare. Grazie!

Un speciale ringraziamento va ai miei supervisor di Torino, Stefania Beolè e Francesco Prino, che mi hanno insegnato tantissimo, facendomi scoprire l'affascinante mondo della ricerca e, soprattutto, hanno sempre trovato la pazienza di ascoltare le mie domande e darmi spiegazioni sempre precise ed esaustive. Ringrazio anche Elena Bruna, che con grande disponibilità, mi ha aiutato sia a Torino che al GSI, e Riccardo Russo, che per primo mi ha seguito durante il lavoro di tesi. Allo stesso modo voglio ringraziare Silvia Masciocchi e Ralf Averbeck, i miei supervisor durante i periodi trascorsi al GSI, i quali mi hanno sempre fatto sentire come a casa e mi hanno fatto comprendere quanto sia importante e proficua la collaborazione a livello internazionale. Grazie anche a Maksym che durante il periodo al GSI, con gran pazienza, mi ha spiegato come utilizzare il codice da lui scritto, e a Lukas (che non ho ancora avuto il piacere di conoscere di persona), che con il suo lavoro mi è stato di grande aiuto per il progetto svolto al GSI.

Vorrei poi dire grazie a tutti coloro che hanno reso le giornate di lavoro più piacevoli, lavorando al mio fianco. In particolare, grazie agli amici e colleghi di Torino e Darmstadt:

- Ai miei compagni di tesi magistrale, Luca e Matteo, che anche loro, come me, si laureano quest'anno.
- Ad Anastasia, Stefano, Ivan e Maximiliano che hanno sempre trovato un momento per aiutarmi e condividere con me la loro maggiore esperienza.

---

*Bibliography*

---

- Agli amici del GSI: Tonatiuh, Jan, Tatiana, Edgar, Andrea, Mariella, Stefan, Julius, Lilit, Jaap e Lukas, per avermi reso più piacevole il soggiorno a Darmstadt.

Vorrei inoltre ringraziare tutti i membri che non ho ancora nominato del gruppo di ALICE di Torino, in particolare il Prof. Massimo Masera e Stefania Bufalino, e del gruppo di ALICE del GSI, per la disponibilità e la gentilezza che in ogni occasione non hanno mancato di dimostrarmi. Infine ringrazio tutti i membri del gruppo di lavoro chiamato *D2H*, con cui ho avuto la possibilità di confrontarmi e che mi hanno aiutato a crescere ed imparare molto.

GRAZIE!