

CMS Draft Analysis Note

The content of this note is intended for CMS internal use and distribution only

2019/01/06

Head Id: 409010

Archive Id: 412087M

Archive Date: 2017/06/07

Archive Tag: trunk

Measurement of prompt $D^{*\pm}$, D^{\pm} and $D^0(\bar{D}^0)$ production cross-section in pp collisions at $\sqrt{s} = 13$ TeV with the CMS detector

Fanò L.¹, Mariani V.¹, and Pedrini D.²

¹ Università degli Studi di Perugia and INFN Perugia

² INFN Milano Bicocca

Abstract

The production cross section of the D^{*+} , D^+ and D^0 (and charge conjugates) mesons has been measured with the CMS detector in pp collisions at a center of mass energy of 13 TeV with the data collected in 2016. The differential cross section of the charmed mesons has been measured in bin of transverse momentum and pseudorapidity in the range $4 < p_T < 100$ GeV/c and $|\eta| < 2.1$.

This box is only visible in draft mode. Please make sure the values below make sense.

PDFAuthor: Livio Fano, Valentina Mariani, Daniele Pedrini

PDFTitle: Measurement of prompt D^{*+} , D^+ and D^0 production cross-section in pp collisions at $\sqrt{s} = 13$ TeV with the CMS detector

PDFSubject: CMS

PDFKeywords: CMS, physics, software, computing

Please also verify that the abstract does not use any user defined symbols

Production and decay chain	Branching fraction
$pp \rightarrow D^{*+} X \rightarrow D^0 \pi_{slow}^+ X \rightarrow K^- \pi^+ \pi_{slow}^+ X$	$BR = (2.63 \pm 0.03)\%$
$pp \rightarrow D^0 X \rightarrow K^- \pi^+ X$	$BR = (3.89 \pm 0.04)\%$
$pp \rightarrow D^+ X \rightarrow K^- \pi^+ \pi^+ X$	$BR = (8.98 \pm 0.28)\%$

Table 1: Reconstructed processes and corresponding branching fraction are shown.

1 Introduction

The measurements of charmed mesons production cross section at hadronic collisions provide essential information for the QCD interactions at high energy. Such studies have been well carried out in all the other LHC experiments ([1], [2], [3]) from proton-proton collisions, at different center of mass energy values. CMS published some results analysing the heavy ion collisions ([4], [5]) but nothing have been published from proton-proton collisions. The measurement of the charmed mesons production at higher energies provided by the LHC run-II will provide a new important test of theoretical calculations. Furthermore, since the current calculations suffer from large theoretical uncertainties, the experimental constraints on heavy-quark production cross sections are important for all the physics phenomena for which heavy-quark production is an important background process.

Charmed mesons produced at the collision points, either directly or as decay products of excited charmed resonances (i.e. D^0 coming from D^* decay), are referred to as promptly produced, while charmed mesons produced through the decays of B hadrons are referred to as secondary charm and are considered as a background process. This analysis is focused on the measurement of the cross section for the prompt production of $D^{*\pm}(2010)^+$, D^0 (and \bar{D}^0) and D^\pm (noted below as D^{*+} , D^0 and D^+ implying always both the charged states). The analyzed data corresponds to 36.8 fb^{-1} collected by the CMS experiment in 2016 proton-proton collisions (pp) at a center of mass energy of $\sqrt{s} = 13 \text{ TeV}$.

In this note we present the measurement of the charmed mesons through the production and the decay chain processes shown in Table 1 with the corresponding Branching Ratio (BR) [6]. The differential cross sections for the charmed mesons as a function of p_T and η have been measured according to the following definition:

$$\frac{d\sigma(pp \rightarrow DX)}{dp_T} = \frac{N_i(D \rightarrow f + c.c)}{\Delta p_T \cdot \mathcal{BR}(D \rightarrow f) \cdot \varepsilon_{i,tot}(D \rightarrow f) \cdot k\mathcal{L}} \quad (1)$$

$$\frac{d\sigma(pp \rightarrow DX)}{d\eta} = \frac{N_i(D \rightarrow f + c.c)}{\Delta \eta \cdot \mathcal{BR}(D \rightarrow f) \cdot \varepsilon_{i,tot}(D \rightarrow f) \cdot k\mathcal{L}} \quad (2)$$

where $N_i(D \rightarrow f + c.c)$ is the signal yield of the charmed meson reconstructed in the selected final state for both the charge states in each bin i , Δp_T ($\Delta \eta$) is the p_T (η) bin width, $\mathcal{BR}(D \rightarrow f)$ is the branching ratio of the reconstructed final state, $\varepsilon_{i,tot}(D \rightarrow f)$ is the total efficiency reconstruction of the decay chain evaluated in MC sample and $k\mathcal{L}$ is the delivered luminosity corrected by the trigger prescale factor.

The structure of this document is the following: in Section 2 the data and simulated samples are discussed, followed in Section 3 and Section 4 by the event selection and the reconstruction and selection strategy. In Section 5 we present how to extract the signal and in Section 6 how to calculate the Monte Carlo efficiency. Section 7 and 22 show the secondary decay contamination and the systematic uncertainties respectively. Finally in Section 9 the cross section values are

Datasets	/ZeroBias/Run2016B-23Sep2016-v3/AOD /ZeroBias/Run2016C-23Sep2016-v1/AOD /ZeroBias/Run2016D-23Sep2016-v1/AOD /ZeroBias/Run2016E-23Sep2016-v1/AOD /ZeroBias/Run2016F-23Sep2016-v1/AOD /ZeroBias/Run2016G-23Sep2016-v1/AOD /ZeroBias/Run2016H-PromptReco-v2/AOD /ZeroBias/Run2016H-PromptReco-v3/AOD
JSON file	Cert_271036-284044_13TeV_23Sep2016ReReco_Collisions16_JSON.txt

Table 2: Datasets and json file used for the analysis.

Datasets	Number of Events
DStarToD0Pi_D0KPi_DStarFilter_TuneCUEP8M1_13TeV-pythia8-evtgen/ RunIIFall15DR76-PU25nsData2015v1_76X_mcRun2_asymptotic_v12-v1/AODSIM	58M
MinBias_TuneCUEP8M1_13TeV-pythia8/ RunIIISpring15DR74-NoPU_MCRUN2_74_V8-v3/AODSIM	19M

Table 3: Datasets and number of events for the Monte Carlo samples used in this analysis.

34 shown and we conclude this analysis in the summary section.

35 2 Data and Simulation Samples

36 This analysis is performed with the ZeroBias datasets using the CMSSW version CMSSW8013.
37 In Table 2 the samples and the JSON file are summarized.

38 The Monte Carlo (MC) simulated samples used in this analysis are from the official production.
39 The production configurations, global tags, and the setup for pile-up mixing, are from the
40 official campaigns. A summary of the details for each of the fully simulated samples used in
41 the analysis is shown in Table 3

42 The first MC sample in Table 3, generated with Pythia6 TuneCUEP8M1, is enriched in D^* : a
43 filter for the decay $D^{*+} \rightarrow D^0 \pi_{slow}^+ \rightarrow K^- \pi^+ \pi_{slow}^+$ (and c.c.) has been applied at Gen Level, with
44 an intrinsic cut at $p_{TD^*} > 3.9$ GeV/c. Since the high statistics available, the enriched sample has
45 been used for the efficiency calculation for all the three channels. The second sample in the
46 Table 3, generated with Pythia8 TuneCUEP8M1, contains MinimumBias events and has been
47 used for the contamination evaluation, as will be discussed in section 7.

48 Both of the MC samples have been produced simulating the 2015 data taking conditions; the
49 main difference with respect to the 2016 data taking is the average value of Pile Up (PU). We
50 concluded that PU difference between data and MC is not affecting our analysis, the reasons
51 and additional checks are reported in the Appendix of this note.

52 3 Event Selection

53 An event is selected when passing the ZeroBias trigger stream HLT_ZeroBias_v*, the most in-
54 clusive trigger, chosen because the kinematic of the objects that we want to reconstruct is par-
55 ticularly soft. The efficiency of the trigger evaluated in the Monte Carlo samples is 100 %.
56 The effective luminosity analysed is calculated taking into account the prescale factor of such

trigger using brilcal [7], and it leads to a luminosity value of $29043 \mu\text{b}^{-1}$, out of the 36.8 fb^{-1} collected by CMS.

We select an event if some conditions about the reconstructed vertices are complied. We looked for candidates only in the Primary Vertex (PV) within the RecoVertex collection of each event. Since the vertices are ordered in p_T we selected the first vertex that satisfies the following requests:

- Valid
- Not Fake
- Number Degree of Freedom > 4
- Distance between the z coordinate of the Primary Vertex and the z coordinate of the Beam Spot $|z_{PV} - z_{BS}| < 10 \text{ cm}$

With these requests we evaluated that we select really the first reco vertex in the collection list with a probability of 94%.

In Figure 1 the distributions of the number of all the reconstructed vertices of the events, the track multiplicity and the number of tracks coming from the selected primary vertex are shown for the different 2016 run periods.

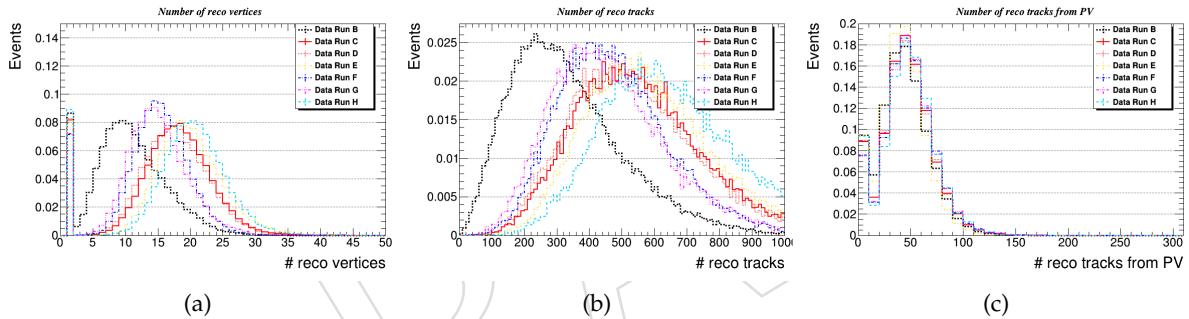


Figure 1: Distributions of: reconstructed vertex (a), tracks multiplicity (b) and number of tracks from the selected PV (c) showed for the different periods of the 2016 data taking.

4 Charmed mesons reconstruction

The strategy used for the reconstruction of the charmed mesons will be deeply discussed below. In the Section 4.1 the tracks selection and the quality requests are shown, commonly to all the final states, while in Section 4.2 all the cuts for the charmed meson candidate reconstruction will be described going into the details for each meson.

4.1 Track selection

Once the primary vertex is selected as explained above, GeneralTracks collection has been used in this analysis for the tracks selection. The basic selection criteria for the tracks used in the mesons reconstruction are listed:

- Track $p_T > 0.5 \text{ GeV}/c$
- Reduced $\chi^2 < 2.5$
- Number of Valid Tracker Hits ≥ 5
- Number of Valid Pixel Hits ≥ 2

- 86 • Impact parameter in the xy plane w.r.t. the PV < 0.1 cm
 87 • Impact parameter in the z direction w.r.t. the PV < 1 cm
 88 In Figure 2 the distributions of these variables before the cuts are shown.

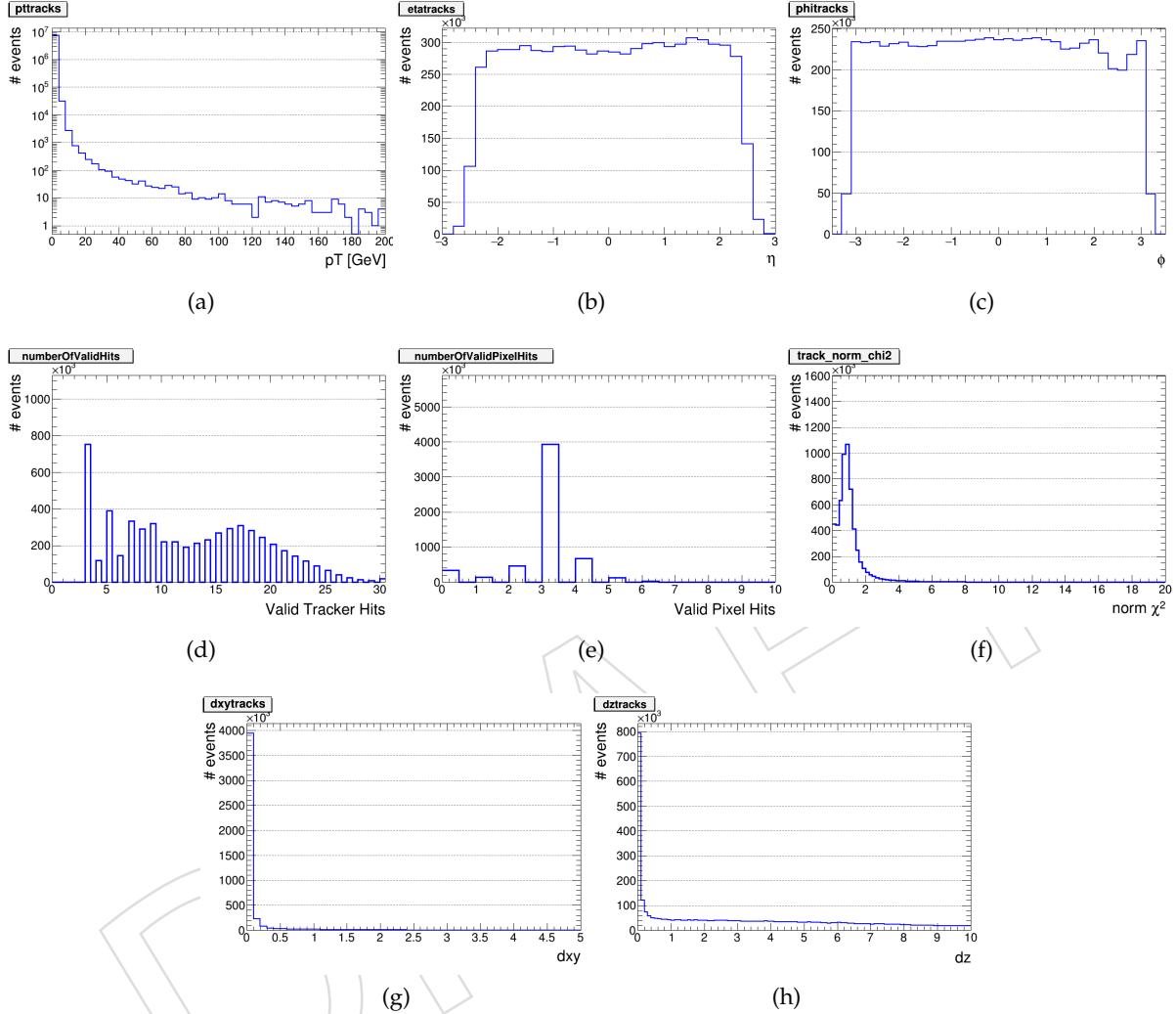


Figure 2: Distributions of GeneralTracks: p_T (a), η (b), ϕ (c), number of valid tracker hits (d), number of valid pixel hits (e), reduced χ^2 (f), d_{xy} w.r.t. the PV (g) and d_z w.r.t. the PV (h).

- 89 For the reconstruction of the D^{*+} decay chain we also had to take in account the additional
 90 track of the pion produced in association with the D^0 . The track selection of such pion, called
 91 *slow*, had to be relaxed with respect to one showed above because of the soft p_T spectrum of
 92 such particle:
- 93 • Track $p_T > 0.3$ GeV/c
 94 • Reduced $\chi^2 < 3$
 95 • Number of Valid Tracker Hits > 2
 96 • Significance of the impact parameter in the xy plane w.r.t. the PV < 3
 97 • Significance of the impact parameter in the z direction w.r.t. the PV < 3
- 98 In Figure 3 the tracks p_T distribution is shown in the very low p_T range ($p_T < 2$ GeV) in order
 99 to show that with our detector we are able to reconstruct even tracks with $p_T \sim 0.3$ GeV.

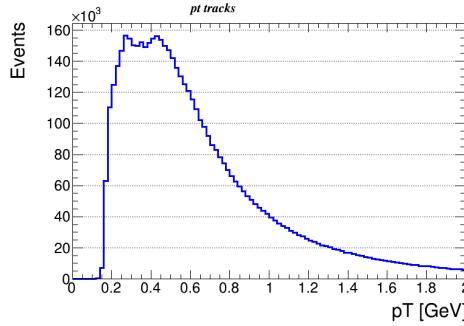


Figure 3: Distributions of tracks in a low p_T range ($p_T < 2$ GeV) to underline the minimum p_T threshold we can reach in offline reconstruction.

4.2 Charmed mesons selection strategy

In this section we describe the strategy used for the open charmed mesons reconstruction. Only the positive sign will be discussed here, the negative was reconstructed with the same strategy but using the charged conjugate. The approach is very similar for all the three mesons, due to the common features between them, but it is important to report the specific cuts applied. Since the charmed mesons have globally a soft p_T distributions (see Figure 4), the main background comes from random tracks combination. It is thus required to use a strategy to achieve a good background rejection.

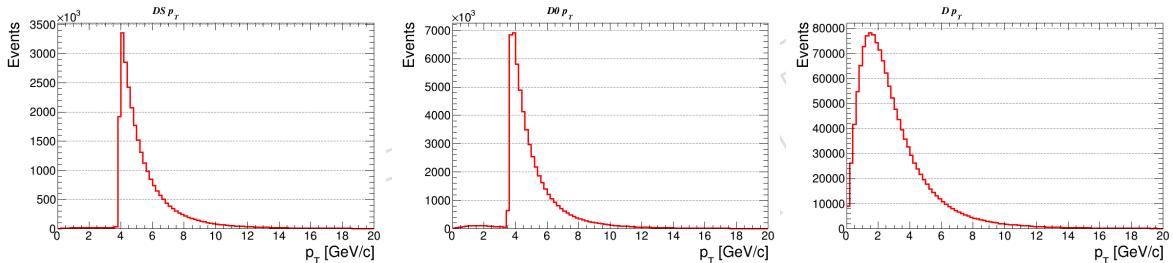


Figure 4: p_T spectra at gen level for D^{*+} (left), D^0 (center) and D^+ (right) mesons. The distributions for D^{*+} and D^0 show a cut at 3.9 GeV/ c because there is an intrinsic cut at the generator level $D_{p_T}^{*+} > 3.9$ GeV/ c .

Both the D^0 and the D^+ mesons have a non-negligible decay length because of the relatively long lifetime, $\tau_{D^0} = (410.1 \pm 1.5) \cdot 10^{-15}$ s and $\tau_{D^+} = (1040 \pm 7) \cdot 10^{-15}$ s. For this reason they generate a secondary vertex (SV) that can be reconstructed from the tracks in the final state. We could thus cut on the decay length, the confidence level of the SV and the direction of the meson with respect to the distance between PV and SV.

All the mesons were reconstructed in the pseudorapidity range $|\eta| < 2.1$ and in a wide transverse momentum range $4 < p_T < 100$ GeV/ c . The explicit strategy is discussed in next sections.

4.2.1 $D^{*\pm}$ meson reconstruction

The D^{*+} was identified by the decay $D^{*+} \rightarrow D^0(\rightarrow K^-\pi^+)\pi_{slow}^+$ (and c.c.). In each event we selected two tracks with opposite charge to combine to get the D^0 candidate and a positive track, with looser requests (specified above), to combine to the D^0 candidates identification. Once the D^0 candidates are reconstructed we associate to them the third track, corresponding to the slow pion, to obtain the D^{*+} meson. We assigned the kaon mass to the negative charged track and the pion mass to the positive one; in order to extract signal we needed to define a

range around the world average value of the D^0 mass. The width of this region was decided by studying the meson invariant mass width in the signal region with a gaussian fit. In Figure 5 the fitted distribution is reported from data.

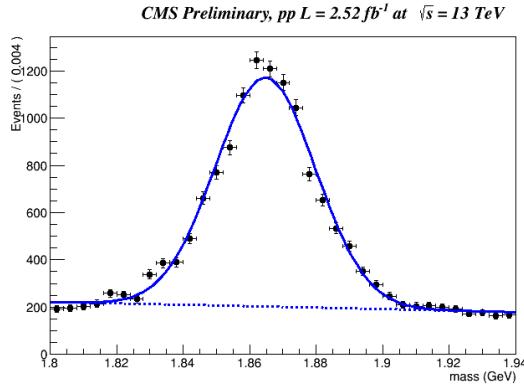


Figure 5: Gaussian fit result on invariant mass distributions for D^0 mass coming from the D^{*+} decay, in signal region $144 < m_{inv}^{D^{*+}} - m_{inv}^{D^0} < 147$ for data at 13 TeV.

From this fit we got a gaussian mean $\mu = 1.8647 \pm 0.0002$ GeV/c and a gaussian width $\sigma = 0.0151 \pm 0.0002$ GeV/c, in agreement with the value foreseen by the PDG (1.86483 ± 0.0005 GeV/c). We studied the stability of such fit in order to figure out any resolution effect of the detector. Thus we repeated the exercise for three different p_T bins and three different η bins. The results are reported in Table 4 and confirm that our global fit gives reasonable results. Only the last pseudorapidity bin is out of range, the resolution in the forward region gets worse as expected, but this is also the less populated bin.

Bin	D^0 width [GeV]	yield	D^0 mass [GeV]
p_T [2-6] GeV	0.01568 ± 0.00003	241688 ± 532	1.8649 ± 0.0001
p_T [6-12] GeV	0.01523 ± 0.00004	255260 ± 648	1.8649 ± 0.0001
p_T [12-24] GeV	0.01510 ± 0.00007	58245 ± 287	1.8649 ± 0.0001
$ \eta < 0.5$	0.01137 ± 0.00003	136930 ± 375	1.8649 ± 0.0001
$0.5 < \eta < 1.5$	0.01536 ± 0.00004	226130 ± 545	1.8649 ± 0.0001
$1.5 < \eta < 2.1$	0.0235 ± 0.0001	74636 ± 320	1.8649 ± 0.0001

Table 4: D^0 width from the fit studied in different p_T and η bins; only the last η bin is out of range but this is also the less populated in term of statistics. This is an expected effect of the detector resolution in the forward region.

We required that the invariant mass of the D^0 candidates had to be within 1.5 standard deviations from the the gaussian fit σ value with respect to the world average value from the PDG, that means 0.023 GeV/c. This cut value represents the best compromise between keeping signal events and rejecting background. A wider range, in fact, would include more signal events but also a huge amount of background events that would make the signal rate extraction very hard since the S/B ratio is very unfavorable for such reconstructed mesons. This is shown in Table 5 where the S/B ratio is reported for different ranges.

Since the S/B values is expected to be even worse than the ones shown in the Table in specific kinematic regions, where the combinatorial is by default higher than signal (low p_T and high $|\eta|$ regions) we decided to be conservative with this cut.

D^0 width cut	S/B
1σ	0.22 ± 0.02
1.5σ	0.18 ± 0.01
2σ	0.16 ± 0.02
2.5σ	0.14 ± 0.03
3σ	0.12 ± 0.02

Table 5: The ratio S/B studied for different D^0 width cut.

142 Since the D^0 meson has a non-negligible decay length it generates a secondary vertex (SV) that
 143 can be reconstructed from the tracks in the final state. We required the existence of a secondary
 144 reconstructed vertex with a confidence level greater than 1%. We can also study the direction of
 145 the D^0 candidates with respect to the distance between PV and SV, defining

$$\cos\Phi = \frac{p_x \cdot \Delta_x + p_y \cdot \Delta_y + p_z \cdot \Delta_z}{|L| \cdot |p|}, \quad (3)$$

146 where p_i are the candidate's momentum components, Δ_i the components of the displacement
 147 between PV and SV, $|L|$ the distance module and $|p|$ the D^0 momentum module. We required
 148 $\cos\Phi > 0.99$, in order to select only candidates that move from PV to SV, not viceversa, and that
 149 are well aligned, removing more than half of the combinatorial background. We also required
 150 the separation of the two vertices, primary and secondary, cutting on the significance of the
 151 distance module L , $L/\sigma_L > 3$. Finally we selected all the D^0 candidates with $p_T > 3$ GeV/c
 152 passing through all these cuts and we combined them with the additional third positive track
 153 assigning to that the pion mass. At the end we select the D^* candidate that minimise the
 154 quantity $\Delta m = m(K\pi\pi_{slow}) - m(K\pi)$ within the upper limit $\Delta m < 0.160$ GeV.

155 For this first decay reconstruction description we report in Figure 6 the distributions of all the
 156 variables used in the selection strategy in order, before the application of the cut on them.
 157 From Figure 6 (a), the distribution of the D^0 candidates mass, it is clear that the combinatorial
 158 between all the selected tracks completely hide the signal, so a tight selection is needed to min-
 159 imise the background coming from the random association. Figure 6 (b) shows the distribution
 160 of the SV Confidence Level, with our cut $CL > 1\%$ we just avoid the peak around 0, since then
 161 the distribution is flat; looking at the sub-Figure (c) instead it is evident the power of the cut
 162 applied at $\cos\Phi > 0.99$, not only to avoid candidates moving in the wrong direction (from SV
 163 to PV) but also to select only candidates that are really well aligned with respect to the PV-SV
 164 distance. In Figure 6 (d) the L/σ_L is shown, with our cut at 3 we are selecting events in the
 165 tail of the distribution, but it is needed to drop down the combinatorial background. Finally
 166 the two Figures in (e) and (f) show the p_T distribution of the D^0 and D^{*+} candidates respec-
 167 tively to underline the soft momentum of such objects and, in Figure (g), the peak of the ΔM
 168 distribution, where the maximum threshold at 0.160 GeV/ c^2 is needed to well appreciate the
 169 shape.

170 In Table 6 the yields and the cut efficiency are reported, evaluated on MC, taking as starting
 171 point all the events in the tracks combinatorial.

172 After the cuts application, in Figure 7 the Δm distributions are shown in different p_T bins.
 173 From lower to higher transverse momentum values the signal peak becomes more and more
 174 clear, with a higher S/B , because the combinatorial contribution drops down, but the statistics
 175 decrease as well.

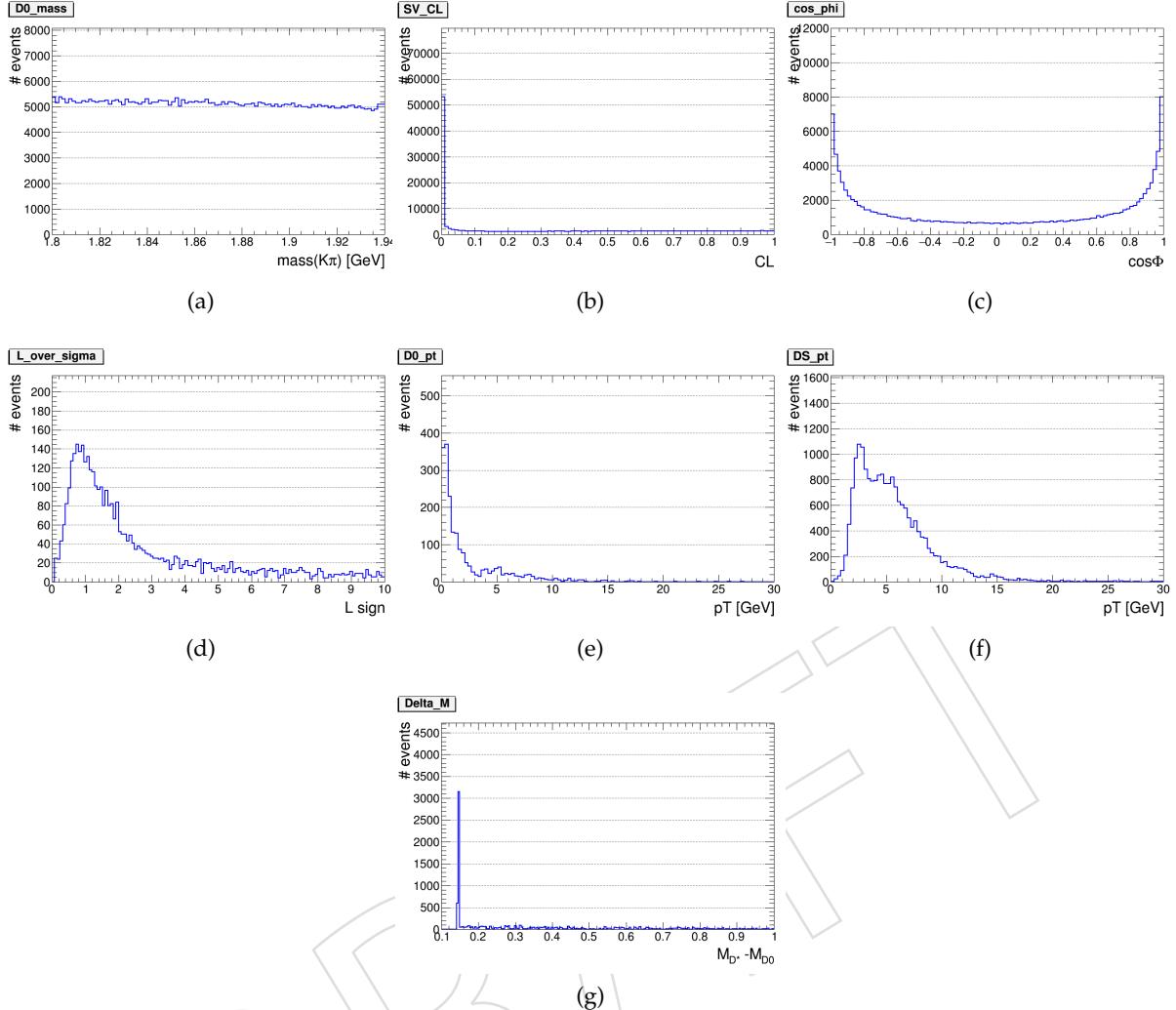


Figure 6: Distributions of the variables used in the decay reconstruction, showed in order before the application of the cut: D^0 candidates mass (a), SV Confidence Level (b), $\cos\Phi$ for the pointing (c), L significance (d), D^0 p_T (e), D^{*+} p_T (f), ΔM (g)

Cut	Yield	Efficiency
Track combinatorial(total charge 0)	$2.16 \cdot 10^{10}$	100 %
D^0 mass	$3.64 \cdot 10^8$	1.7 %
SV Confidence Level	$2.60 \cdot 10^8$	1.2 %
Pointing ($\cos\phi$)	9920720	0.046 %
$L\sigma$	4073736	0.019 %
D^0 p_T	1317024	0.0061 %
ΔM	656360	0.0030 %

Table 6: Cut efficiency, taking as starting point the tracks combinatorial, for the D^{*+} chain reconstructed in the MC sample.

176 4.2.2 D^0 and \bar{D}^0 meson reconstruction

177 The D^0 was identified by the decay $D^0 \rightarrow K^- \pi^+$ (and c.c.). As already said, we considered
178 prompt D^0 both the mesons coming directly from the p-p interaction and the mesons coming

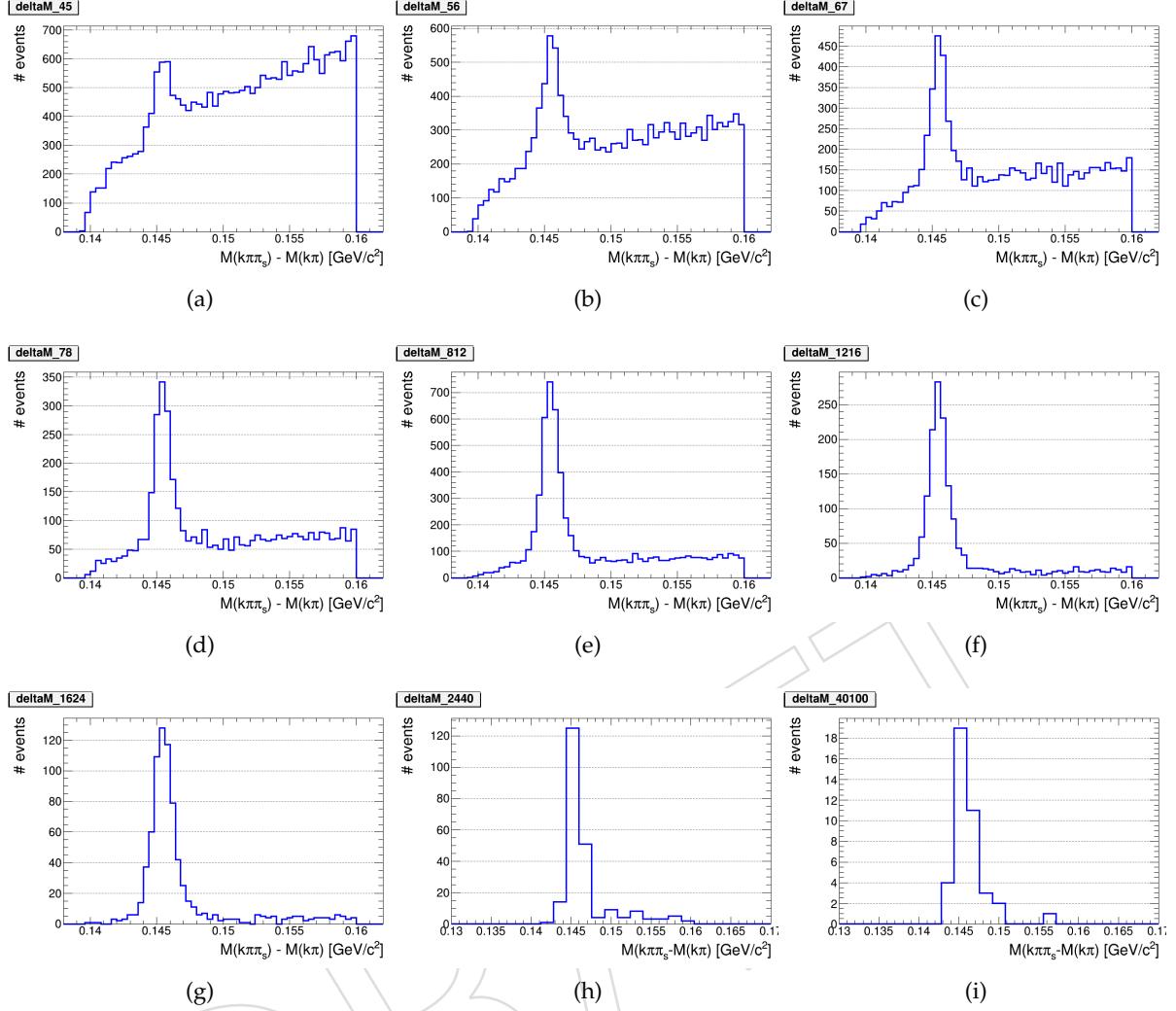


Figure 7: Distributions of $\Delta m = m(K\pi\pi_{slow}) - m(K\pi)$ variable in different p_T bins: 4-5 GeV/c (a), 5-6 GeV/c (b), 6-7 GeV/c (c), 7-8 GeV/c (d), 8-12 GeV/c (e), 12-16 GeV/c (f), 16-24 GeV/c (g), 24-40 GeV/c (h) and 40-100 GeV/c (i).

from the decay of an excited charmed resonance. In the enriched MC we could compute, at gen level, that about 2/3 of the D^0 mesons come from a D^* decay; this value changes to $\sim 24\%$ in the MinBias MC sample without any filter applied.

The strategy used for the reconstruction of this decay is almost the same described in section 4.2.1, with few differences that will be discussed below. The first one is the p_T threshold for the selected tracks: in order to minimize the combinatorial we asked to the tracks to have a p_T greater than $0.8 \text{ GeV}/c$. For this decay reconstruction we couldn't apply the same tight cut on the D^0 invariant mass used for the D^* reconstruction, because a wider range is needed to fit the shape for the signal yield estimation. We just asked to the D^0 candidates to have an invariant mass within $0.1 \text{ GeV}/c^2$ with respect to the world average value ($M_{D^0} = 1.86484 \pm 0.00017 \text{ GeV}/c^2$). The requirements on the reconstructed secondary vertex are the same described in section 4.2.1 for both the confidence level and the pointing of the D^0 candidates direction. Since we needed to cut down the large combinatorial contribution we required the L significance to be greater than 5.

In Figure 8 the D^0 candidates mass distributions are shown in different p_T bins. From lower

194 to higher transverse momentum values the signal peak becomes more and more clear, with a
 195 higher S/B , because the combinatorial contribution drops down, but the statistics decrease as
 196 well.

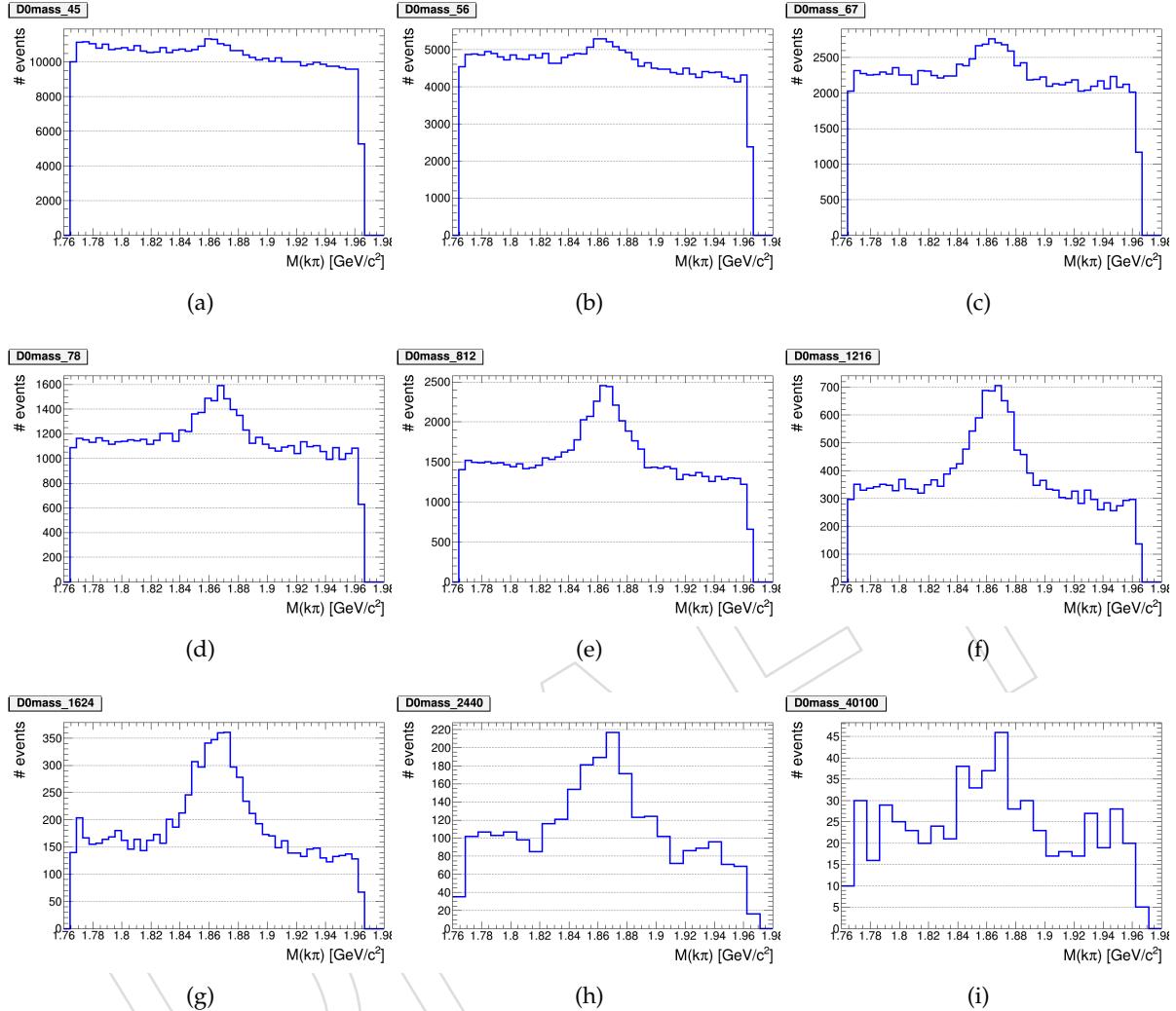


Figure 8: Distributions of $m(K\pi)$ variable in different p_T bins: 4-5 GeV/c (a), 5-6 GeV/c (b), 6-7 GeV/c (c), 7-8 GeV/c (d), 8-12 GeV/c (e), 12-16 GeV/c (f), 16-24 GeV/c (g), 24-40 GeV/c (h) and 40-100 GeV/c (i).

197 4.2.3 D^\pm meson reconstruction

198 The D^+ was identified by the decay $D^+ \rightarrow K^-\pi^+\pi^+$ (and c.c.). We selected three tracks from
 199 the track collection following the requests specified in section 4.1, but we asked the p_T mini-
 200 mum threshold to be 0.7 GeV/c . The global sign of the tracks had to be positive: we assigned
 201 the kaon mass to the negative charged track and the pion mass to the two positive tracks.

202 As for the D^0 reconstruction, the strategy used is almost the same described in section 4.2.1,
 203 with few differences. Again we had to set a larger window for the D^+ invariant mass: we just
 204 asked to the D^+ candidates to have an invariant mass within $0.1 \text{ GeV}/c^2$ with respect to the
 205 world average value ($M_{D^+} = 1.8696 \pm 0.0002 \text{ GeV}/c^2$). The requirements on the reconstructed
 206 secondary vertex were the same of the D^{*+} reconstruction for both the confidence level
 207 and the pointing of the candidates direction. Since the D^+ has a greater lifetime than the D^0 we

208 required the L significance to be greater than 10.

209 In Figure 9 the D^+ candidates mass distributions are shown in different p_T bins. From lower
210 to higher transverse momentum values the signal peak becomes more and more clear, with a
211 higher S/B , because the combinatorial contribution drops down, but the statistics decrease as
212 well.

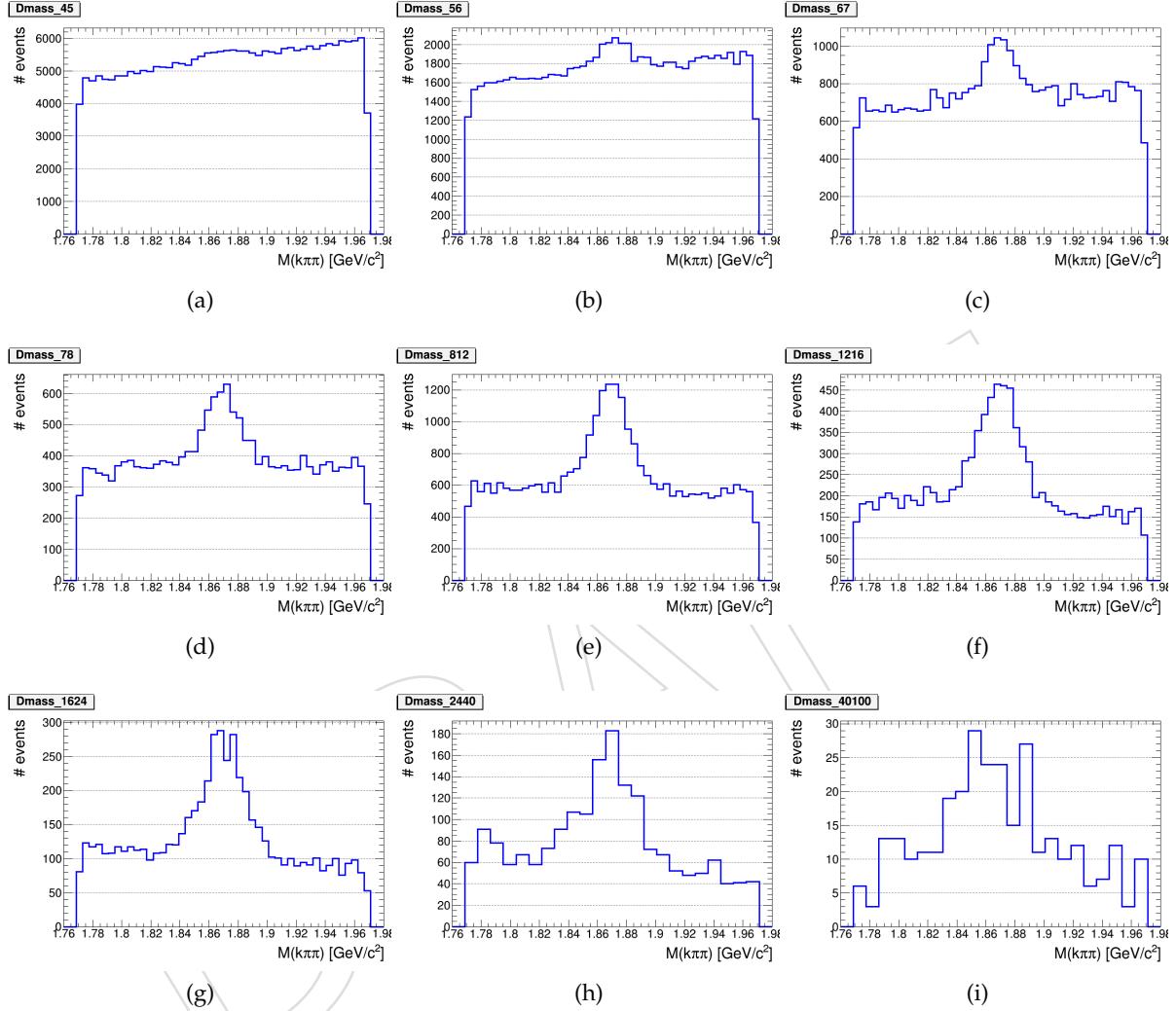


Figure 9: Distributions of $m(K\pi\pi)$ variable in different p_T bins: 4-5 GeV/c (a), 5-6 GeV/c (b), 6-7 GeV/c (c), 7-8 GeV/c (d), 8-12 GeV/c (e), 12-16 GeV/c (f), 16-24 GeV/c (g), 24-40 GeV/c (h) and 40-100 GeV/c (i).

213 The cut summary, specific for each meson, is shown in Table 7.

214 5 Signal Yield Determination

215 In this section the signal yield determination for each charmed meson is described in detail.
216 We determined signal yields in data and MC samples using an unbinned maximum-likelihood
217 fit to the invariant mass distributions for each decay mode (the ΔM distribution is considered
218 for the D^{*+}). For the signal peak we fitted two overlapping gaussians with a common mean
219 value while for the background shape we used different functions according to the decay.

Variables	D^{*+}	D^0	D^+
Tracks: $p_T \text{ min GeV}/c$	0.5 (0.3 for the π_{slow})	0.8	0.7
Tracks: reduced χ^2	< 2.5 (3 for the π_{slow})	< 2.5	< 2.5
Tracks: number of Valid Tracker Hits	≥ 5 (> 2 for the π_{slow})	≥ 5	≥ 5
Tracks: number of Valid Pixel Hits	≥ 2 (none for the π_{slow})	≥ 2	≥ 2
Tracks: IP_{xy} w.r.t. PT	< 0.1 cm (signif < 3 for the π_{slow})	< 0.1 cm	< 0.1 cm
Tracks: IP_z w.r.t. PT	< 1 cm (signif < 3 for the π_{slow})	< 1 cm	< 1 cm
Mass range $ M_{D_{\text{cand}}^0} - M_{D_{\text{PDG}}^0} [\text{GeV}/c^2]$	< 0.023	< 0.10	< 0.10
SV Confidence Level	> 1%	> 1%	> 1%
Pointing, $\cos\Phi$	> 0.99	> 0.99	> 0.99
L significance: L/σ_L	> 3	> 5	> 10

Table 7: List of all the cuts applied to each meson candidate

220 All the following Figures and the signal yields reported in the following tables are referred to
 221 the sum of D^{*+} (D^0 and D^+) and charge conjugation decays contributions.

222 5.1 Signal Yield Determination for D^{*+} meson

223 We fitted the ΔM distributions; for the signal we used two gaussians with a common mean

$$g_1 = \frac{1}{\sigma_1 \sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma_1^2}} \quad g_2 = \frac{1}{\sigma_2 \sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma_2^2}} \quad (4)$$

224 while for background we used a phenomenological threshold function given by

$$f = \left(1 - e^{-\frac{\Delta M - m_0}{p_0}}\right) \left(\frac{\Delta M}{m_0}\right)^{p_1} + p_2 \left(\frac{\Delta M}{m_0} - 1\right) \quad (5)$$

225 where m_0 is the endpoint, taken to be the pion mass and $p_{0,1,2}$ are free parameters.

226 Fit results are shown in Figure 10 for different p_T bins, in the whole pseudorapidity range
 227 $|\eta| < 2.1$ for the data sample and the signal yields provided by the fits with the statistical
 228 errors are reported in Table 8.

229 The quality of a fit can be quantified by the pull distribution, defined as

$$\text{pull}(N_{\text{sig}}) = \frac{N_{\text{sig}}^{\text{fit}} - N_{\text{sig}}^{\text{true}}}{\sigma_N^{\text{fit}}} \quad (6)$$

230 By definition the values are expected to be centered in 0 within the uncertainties; deviations
 231 from 0 correspond to bin not well described by the fit. Figure 11 shows the pull distributions
 232 related to each fit showed in Figure 10, the red line represents the ideal pull value.

233 We also estimated the signal yield in bin of η for all the D^{*+} candidate mesons with p_T within
 234 4 and 100 GeV/c. The fitted Δm distributions are shown in Figure 12, the pull distributions

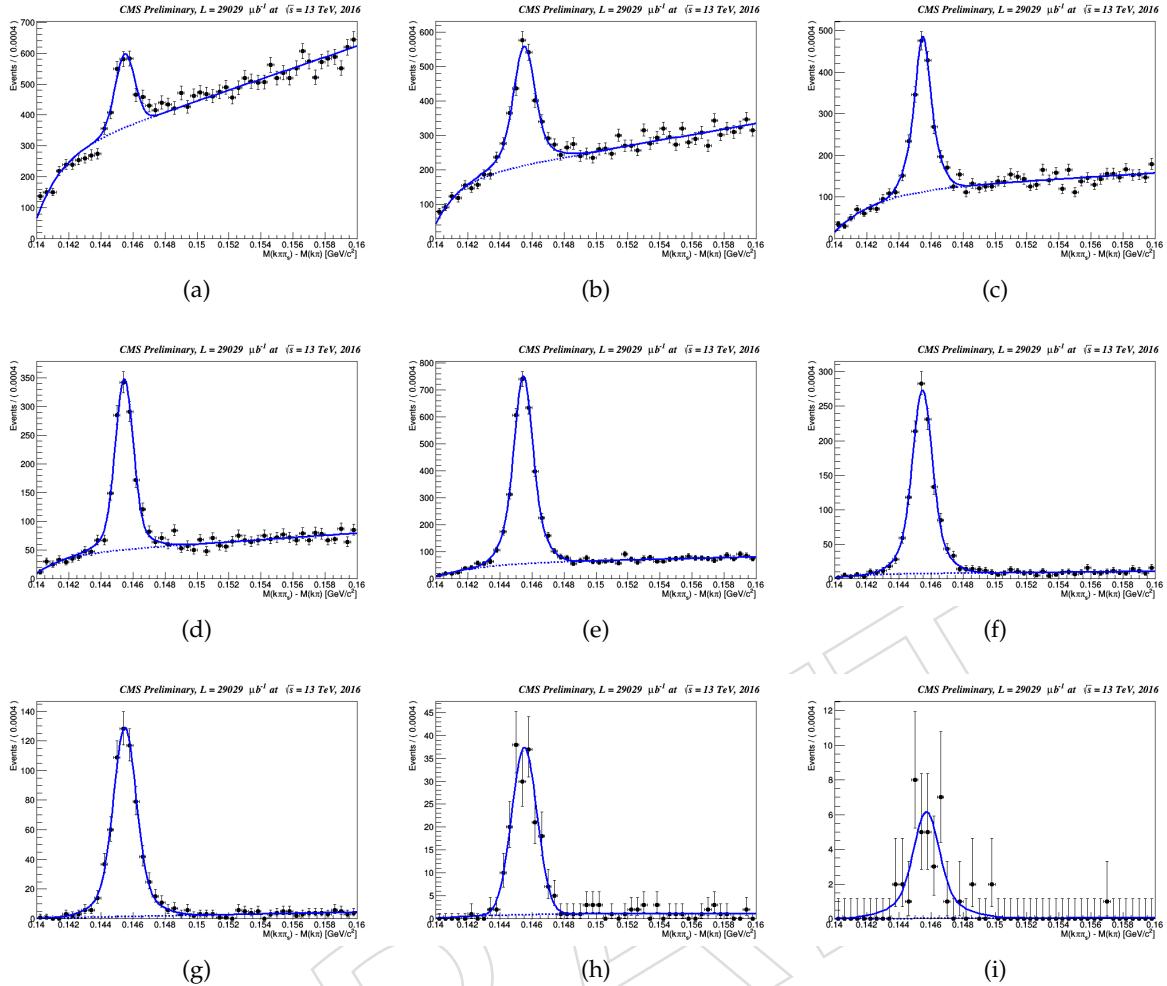


Figure 10: Fit on the $\Delta m = m(K\pi\pi_{slow}) - m(K\pi)$ distributions in different p_T bins with $|\eta| < 2.1$: 4-5 GeV/c (a), 5-6 GeV/c (b), 6-7 GeV/c (c), 7-8 GeV/c (d), 8-12 GeV/c (e), 12-16 GeV/c (f), 16-24 GeV/c (g), 24-40 GeV/c (h) and 40-100 GeV/c (i) integrating in η . The first fit is not perfect because of the huge amount of combinatorial background in such p_T range.

(defined in 6) are shown in Figure 13, with the red line that indicates the ideal value, and the fit results are reported in the Table 9.

5.2 Signal Yield Determination for D^0 meson

We fitted the D^0 invariant mass distributions: for the signal we used two gaussians with a common mean as reported in equation 4, while for background we used a third degree polynomial function where the four coefficients are free parameters.

Fit results are shown in Figure 14 for different p_T bins, in the pseudorapidity range $|\eta| < 2.1$ for the data sample and the signal yields provided by the fits with the statistical errors are reported in the Table 10. In Figure 15 are reported the pull values corresponding to such fits with the red line that indicates the ideal value at 0 according to equation 6.

We also estimated the signal yield in bin of η for all the D^0 candidate mesons with p_T greater than 4 GeV/c . The fitted D^0 invariant mass distributions are shown in Figure 16 and the fit results are reported in the Table 11. In Figure 17 are reported the pull values corresponding to such fits with the red line that indicates the ideal value at 0 according to equation 6.

p_T range [GeV/c]	$ \eta < 2.1$	
	Sign Yield	Uncert
[4-5]	917	94
[5-6]	1555	204
[6-7]	1450	98
[7-8]	1144	75
[8-12]	2962	80
[12-16]	1174	115
[16-24]	647	28
[24-40]	180	10
[40-100]	38	6

Table 8: Signal yields and statistical uncertainties for D^{*+} mesons in p_T bins for $|\eta| < 2.1$.

$ \eta $ range	Sign Yield	Uncert
[0-0.2]	1227	42
[0.2-0.4]	1337	150
[0.4-0.6]	1276	45
[0.6-0.8]	1540	105
[0.8-1]	1283	68
[1-1.2]	1026	88
[1.2-1.4]	857	113
[1.4-1.6]	576	191
[1.6-1.8]	530	84
[1.8-2]	453	38
[2-2.1]	86	31

Table 9: Signal yields and statistical uncertainties for D^{*+} mesons with $4 < p_T < 100$ GeV/c in η bins.

p_T range [GeV/c]	$ \eta < 2.1$	
	Sign Yield	Uncert
[4-5]	5195	704
[5-6]	5478	395
[6-7]	5639	445
[7-8]	3989	497
[8-12]	9752	320
[12-16]	3536	167
[16-24]	2042	122
[24-40]	625	58
[40-100]	91	20

Table 10: Signal yields and statistical uncertainties for D^0 mesons in p_T bins for $|\eta| < 2.1$.

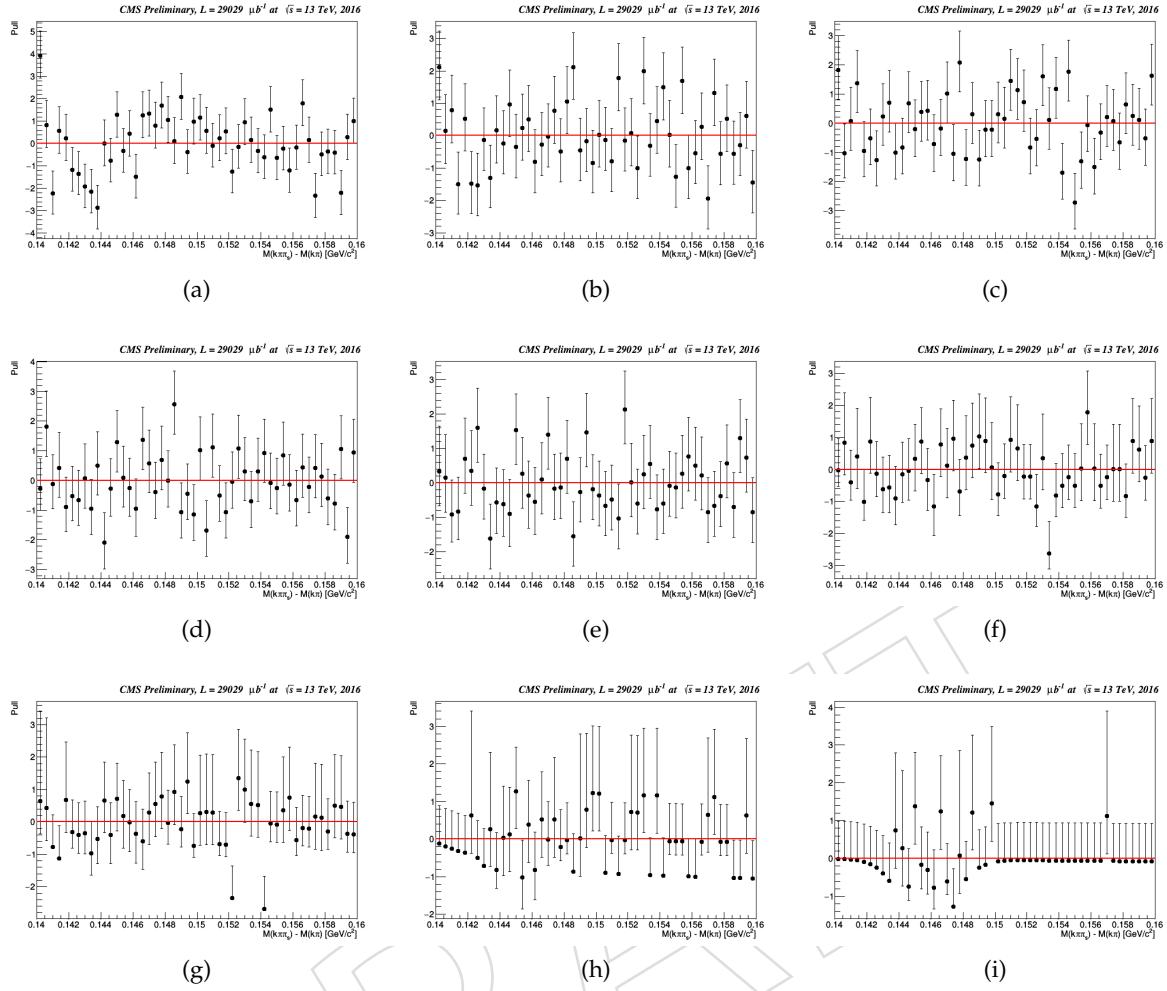


Figure 11: Pull of the fit on the distribution of the $\Delta m = m(K\pi\pi_{slow}) - m(K\pi)$ distributions in different p_T bins with $|\eta| < 2.1$: 4-5 GeV/c (a), 5-6 GeV/c (b), 6-7 GeV/c (c), 7-8 GeV/c (d), 8-12 GeV/c (e), 12-16 GeV/c (f), 16-24 GeV/c (g), 24-40 GeV/c (h) and 40-100 GeV/c (i) integrating in η .

249 5.3 Signal Yield Determination for D^+ meson

250 We fitted the D^+ invariant mass distributions: for the signal we used two gaussians with a
251 common mean as reported in the equation 4, while for background we used a third degree
252 polynomial function where the four coefficients are free parameters.

253 Fit results are shown in Figure 18 for different p_T bins, in the pseudorapidity range $|\eta| < 2.1$ for
254 the data sample and the signal yields provided by the fits with the statistical errors are reported
255 in the Table 12. In Figure 19 are reported the pull values corresponding to such fits with the red
256 line that indicates the ideal value at 0 according to equation 6.

257 We also estimated the signal yield in bin of η for all the D^+ candidate mesons with p_T greater
258 than 4 GeV/c. The fitted D invariant mass distributions are shown in Figure 20 and the fit
259 results are reported in the Table 13. In Figure 21 are reported the pull values corresponding to
260 such fits with the red line that indicates the ideal value at 0 according to equation 6.

$ \eta $ range	Sign Yield	Uncert
[0-0.2]	3636	124
[0.2-0.4]	4146	137
[0.4-0.6]	4281	153
[0.6-0.8]	4588	182
[0.8-1]	3707	214
[1-1.2]	4269	321
[1.2-1.4]	3178	303
[1.4-1.6]	3053	317
[1.6-1.8]	2196	514
[1.8-2]	3096	350
[2-2.1]	1076	238

Table 11: Signal yields and statistical uncertainties for D^0 mesons with $4 < p_T < 100$ GeV/ c in η bins.

p_T range [GeV/ c]	$ \eta < 2.1$	
	Sign Yield	Uncert
[4-5]	5850	1420
[5-6]	3391	455
[6-7]	3618	143
[7-8]	2785	131
[8-12]	6858	179
[12-16]	3047	96
[16-24]	1821	77
[24-40]	628	61
[40-100]	115	14

Table 12: Signal yields and statistical uncertainties for D^+ mesons in p_T bin for $|\eta| < 2.1$.

$ \eta $ range	Sign Yield	Uncert
[0-0.2]	3202	106
[0.2-0.4]	3062	96
[0.4-0.6]	2906	190
[0.6-0.8]	3204	226
[0.8-1]	3040	143
[1-1.2]	3183	161
[1.2-1.4]	2853	89
[1.4-1.6]	2353	251
[1.6-1.8]	2125	205
[1.8-2]	1702	338
[2-2.1]	1077	149

Table 13: Signal yields and statistical uncertainties for D^+ mesons with $4 < p_T < 100$ GeV/ c in η bins.

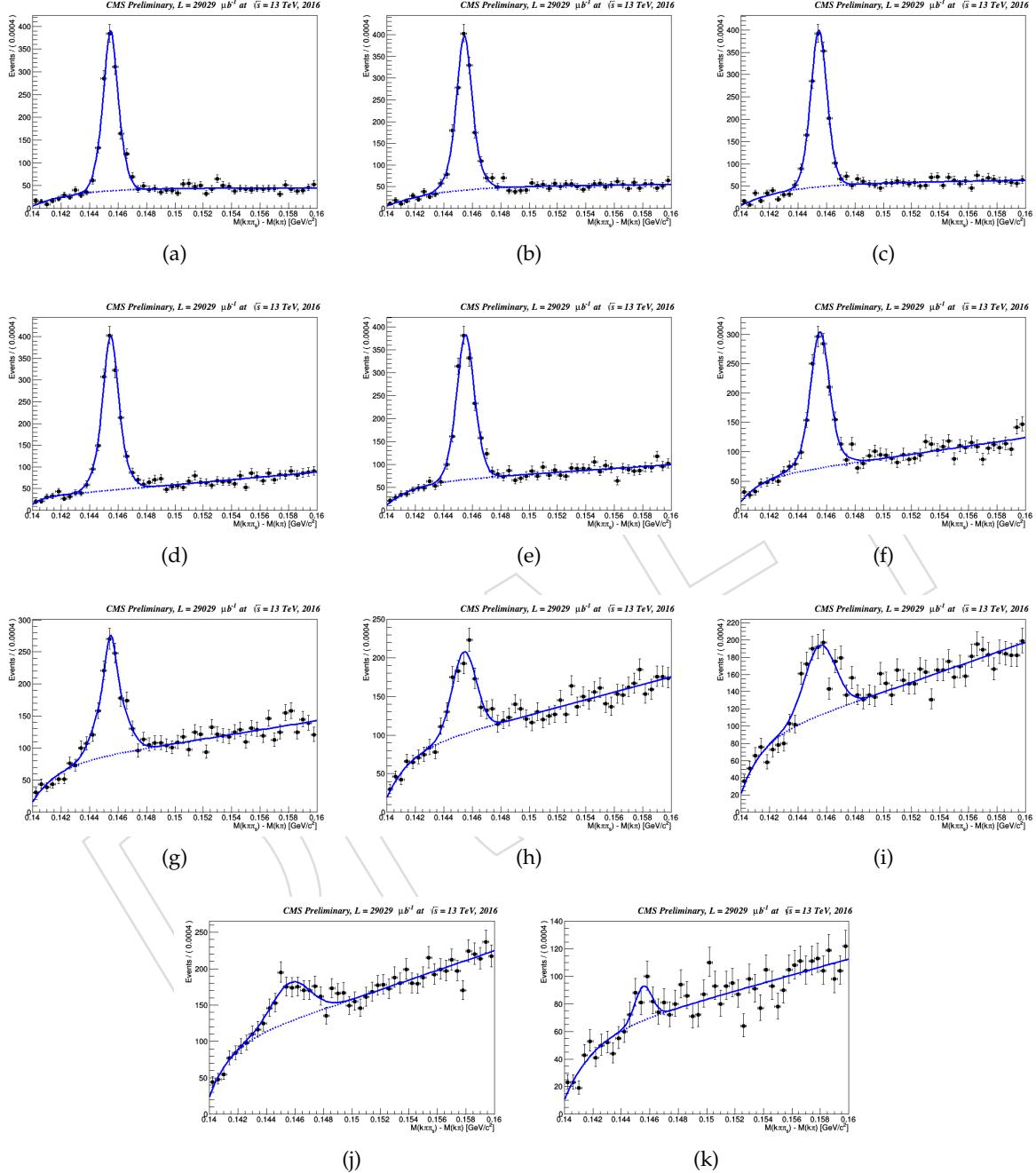


Figure 12: Fit on the $\Delta m = m(K\pi\pi_{slow}) - m(K\pi)$ distributions in different $|\eta|$ bins: 0-0.2 (a), 0.2-0.4 (b), 0.4-0.6 (c), 0.6-0.8 (d), 0.8-1 (e), 1-1.2 (f), 1.2-1.4 (g), 1.4-1.6 (h), 1.6-1.8 (i), 1.8-2 (j), 2-2.1 (k) for D^{*+} mesons with $4 < p_T < 100 \text{ GeV}/c$.

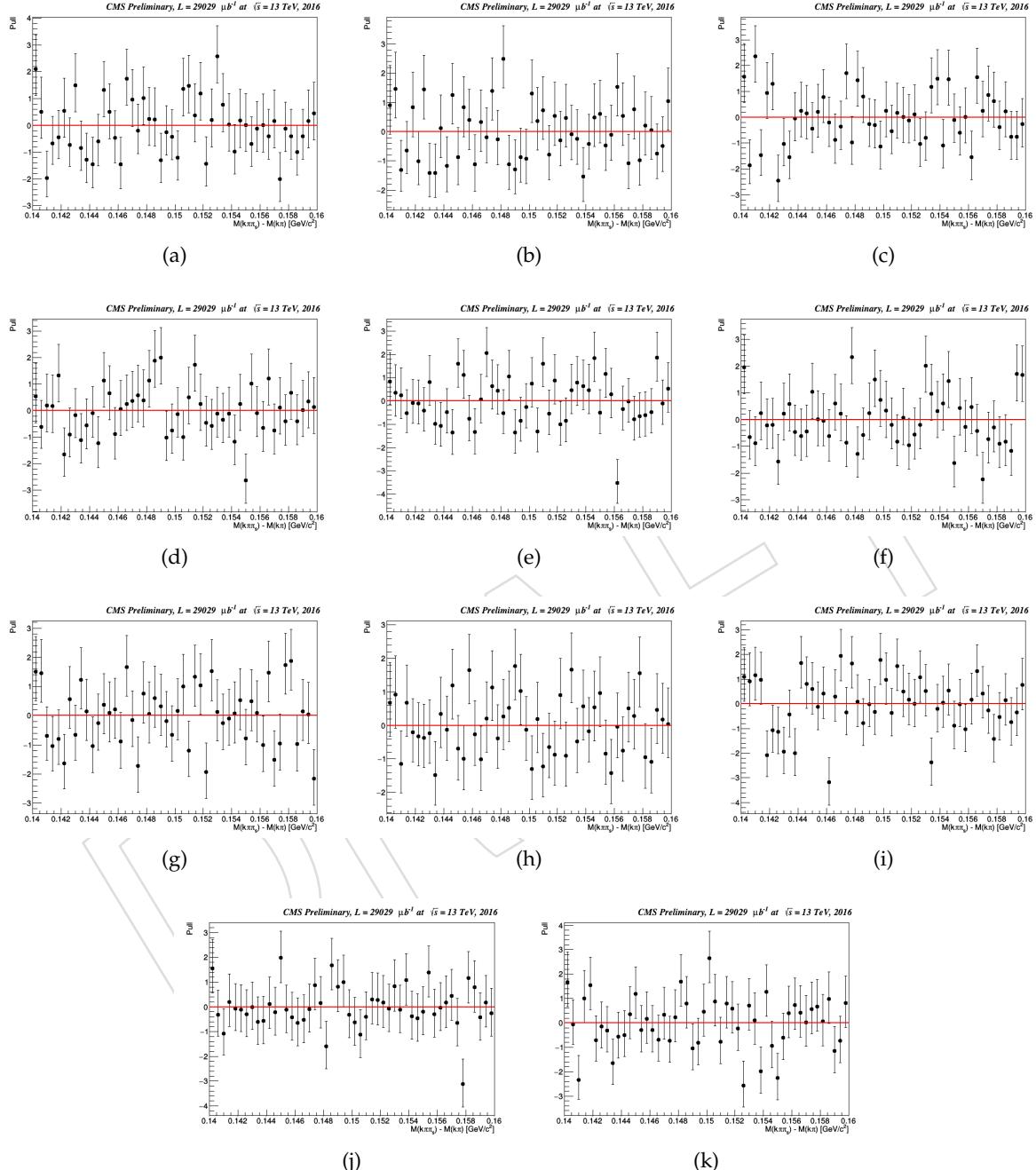


Figure 13: Pull of the fit on the distribution of the $\Delta m = m(K\pi\pi_{slow}) - m(K\pi)$ distributions in different $|\eta|$ bins: 0-0.2 (a), 0.2-0.4 (b), 0.4-0.6 (c), 0.6-0.8 (d), 0.8-1 (e), 1-1.2 (f), 1.2-1.4 (g), 1.4-1.6 (h), 1.6-1.8 (i), 1.8-2 (j), 2-2.1 (k) for D^{*+} mesons with $4 < p_T < 100$ GeV/ c .

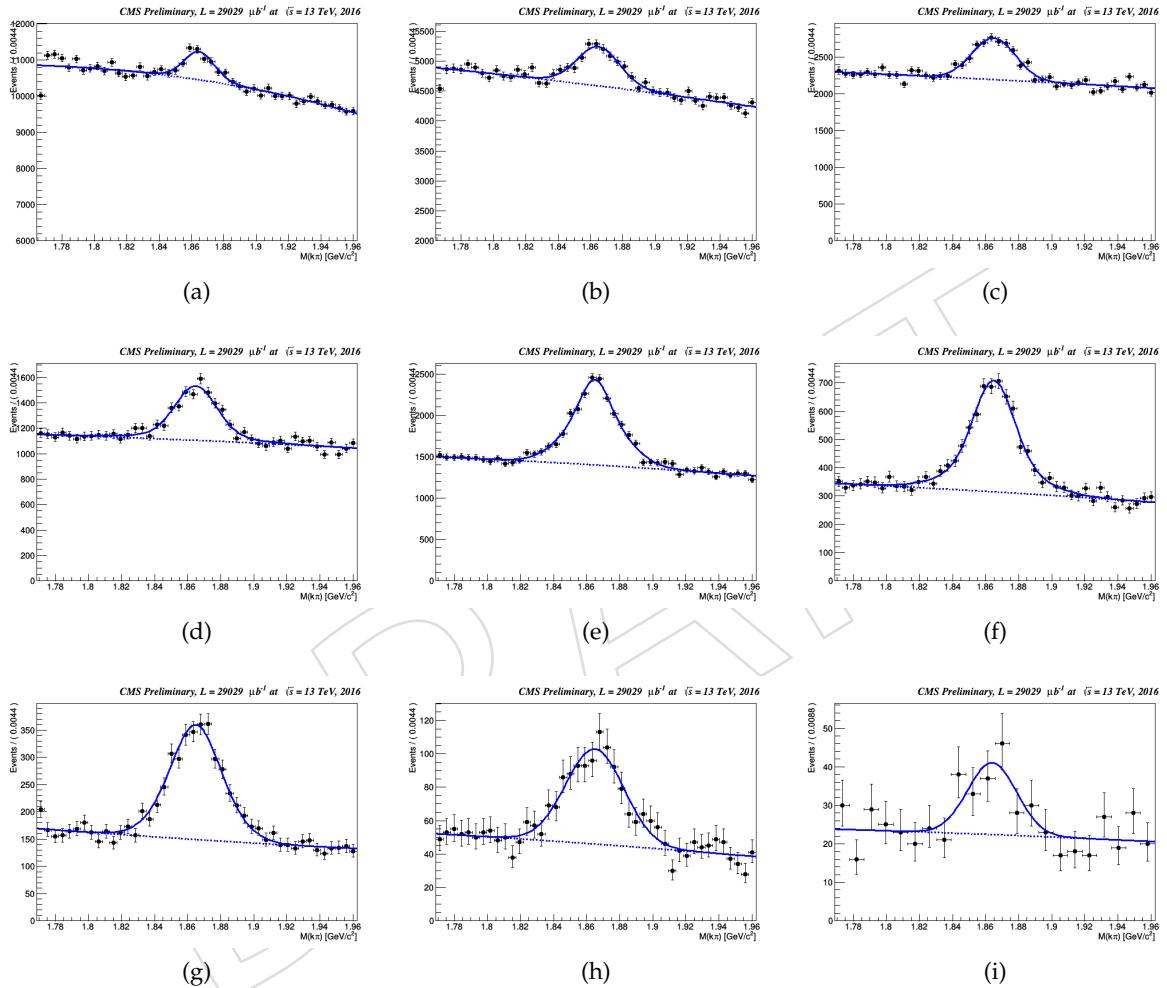


Figure 14: Fit on the $m(K\pi)$ distributions in different p_T bins with $|\eta| < 2.1$: 4-5 GeV/c (a), 5-6 GeV/c (b), 6-7 GeV/c (c), 7-8 GeV/c (d), 8-12 GeV/c (e), 12-16 GeV/c (f), 16-24 GeV/c (g), 24-40 GeV/c (h) and 40-100 GeV/c (i).

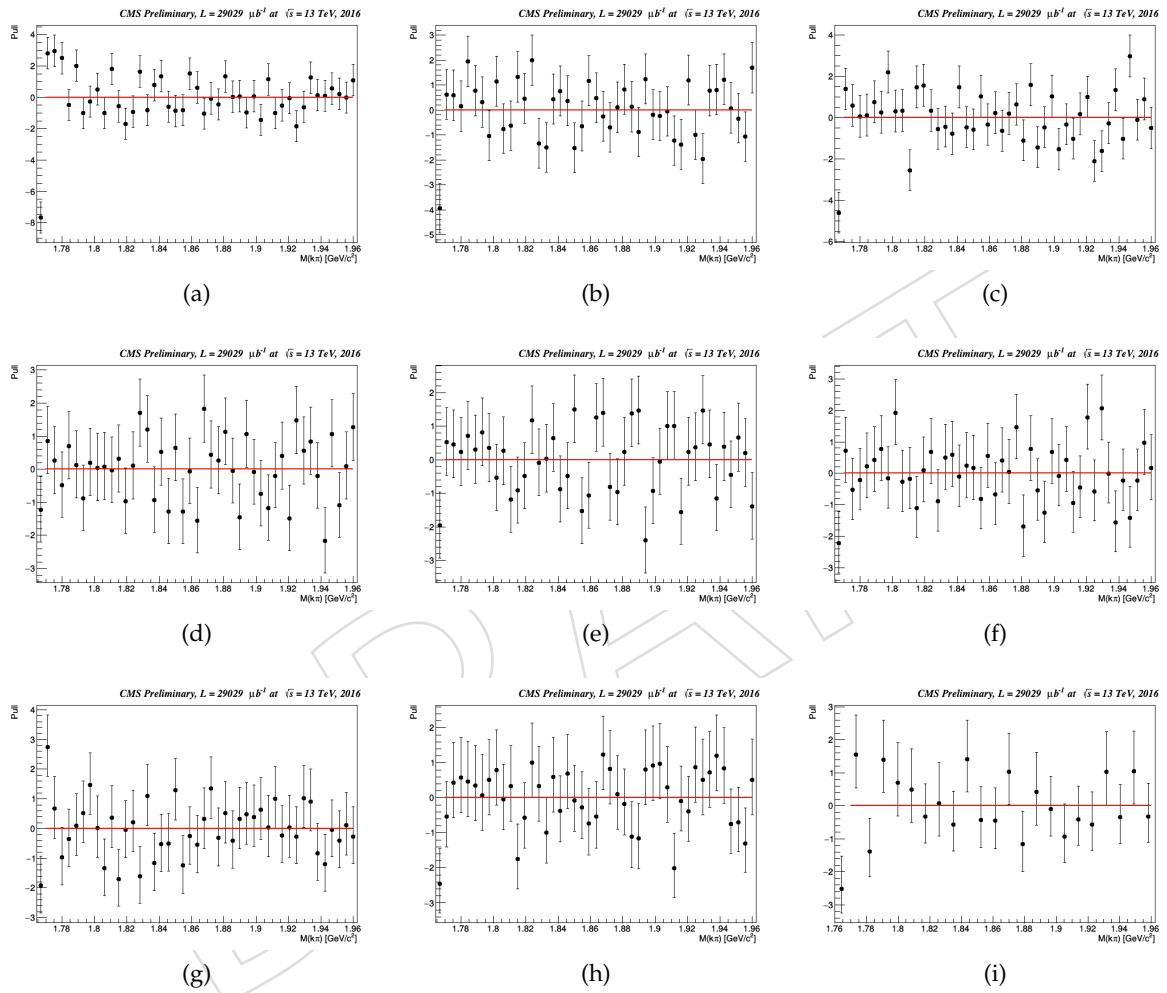


Figure 15: Pull of fit on the $m(K\pi)$ distributions in different p_T bins with $|\eta| < 2.1$: 4-5 GeV/c (a), 5-6 GeV/c (b), 6-7 GeV/c (c), 7-8 GeV/c (d), 8-12 GeV/c (e), 12-16 GeV/c (f), 16-24 GeV/c (g), 24-40 GeV/c (h) and 40-100 GeV/c (i).

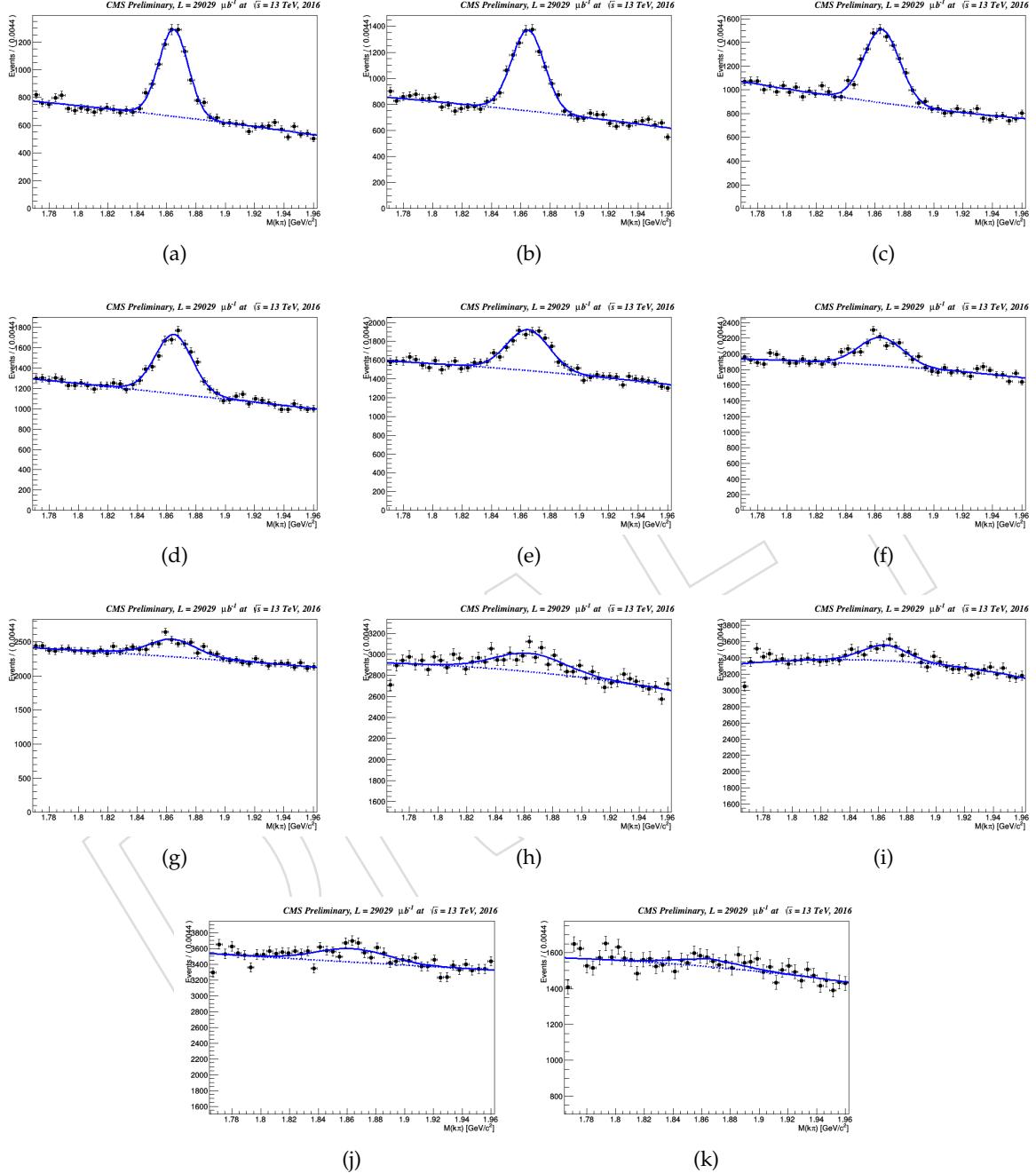


Figure 16: Fit on the $m(K\pi)$ distributions in different $|\eta|$ bins: 0-0.2 (a), 0.2-0.4 (b), 0.4-0.6 (c), 0.6-0.8 (d), 0.8-1 (e), 1-1.2 (f), 1.2-1.4 (g), 1.4-1.6 (h), 1.6-1.8 (i), 1.8-2 (j), 2-2.1 (k) for D^0 mesons with $4 < p_T < 100 \text{ GeV}/c$.

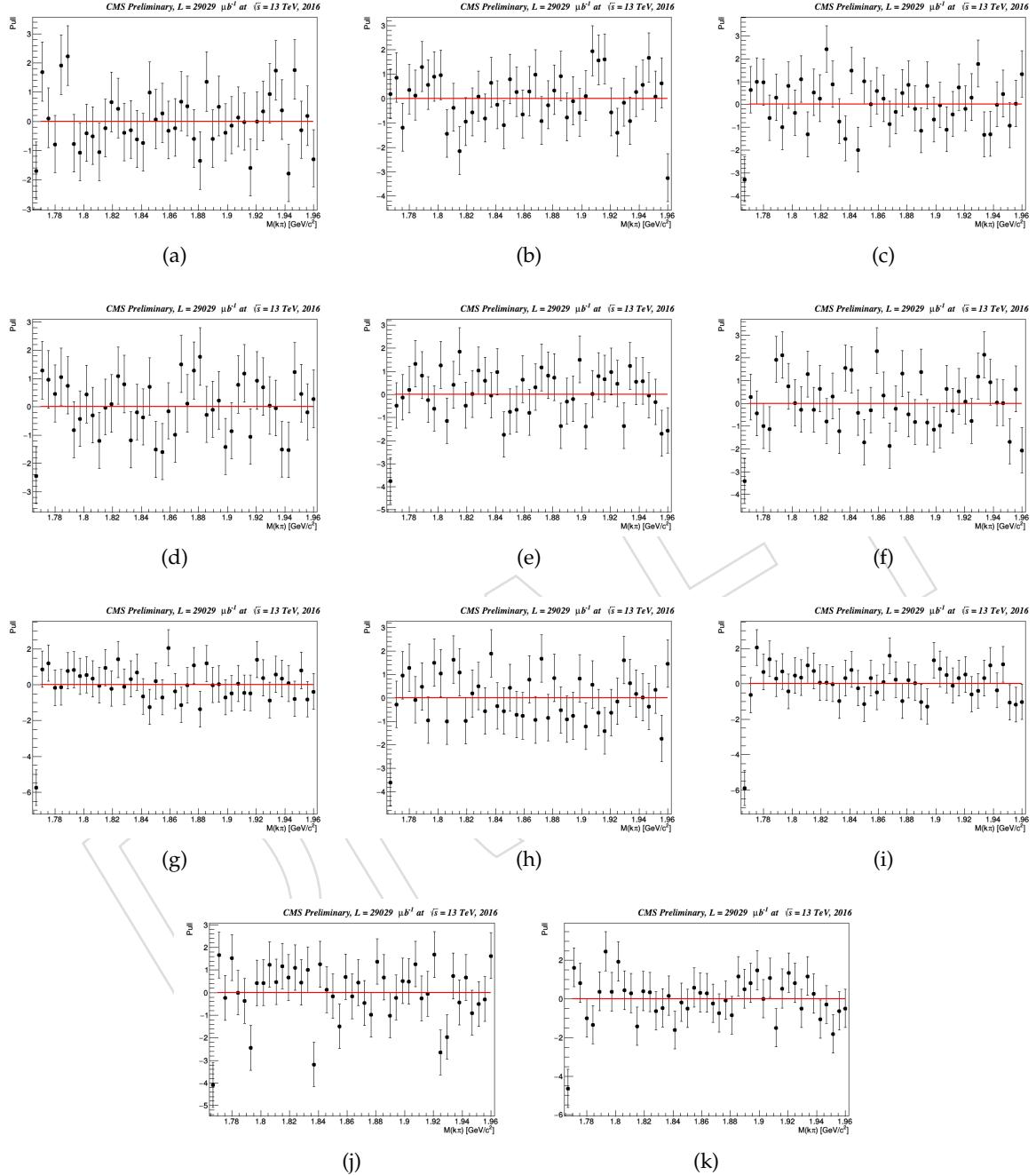


Figure 17: Pull of the fit on the $m(K\pi)$ distributions in different $|\eta|$ bins: 0-0.2 (a), 0.2-0.4 (b), 0.4-0.6 (c), 0.6-0.8 (d), 0.8-1 (e), 1-1.2 (f), 1.2-1.4 (g), 1.4-1.6 (h), 1.6-1.8 (i), 1.8-2 (j), 2-2.1 (k) for D^0 mesons with $4 < p_T < 100 \text{ GeV}/c$.

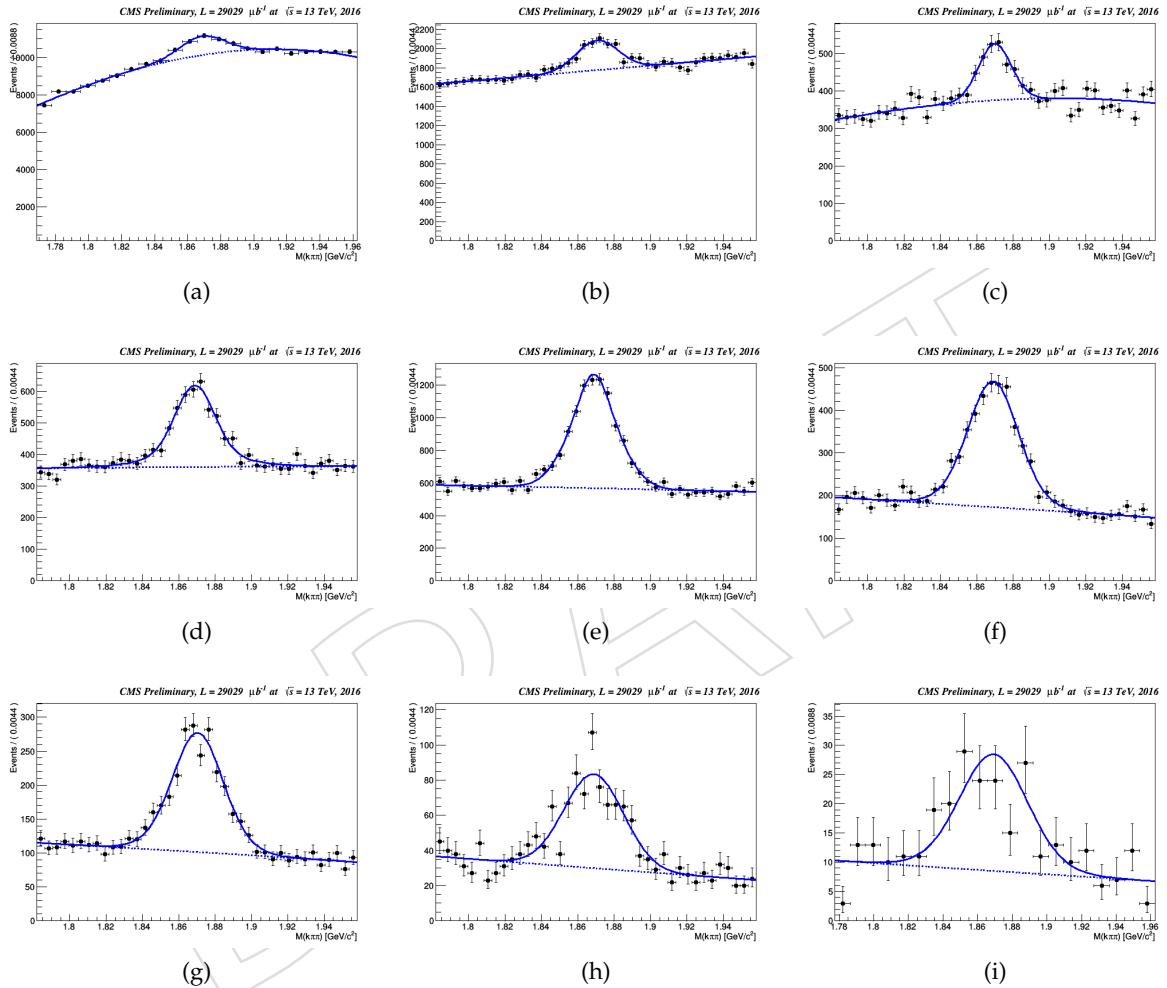


Figure 18: Fit on the $m(K\pi\pi)$ distributions in different p_T bins for $|\eta| < 2.1$: 4-5 GeV/c (a), 5-6 GeV/c (b), 6-7 GeV/c (c), 7-8 GeV/c (d), 8-12 GeV/c (e), 12-16 GeV/c (f), 16-24 GeV/c (g), 24-40 GeV/c (h) and 40-100 GeV/c (i) integrating in η .

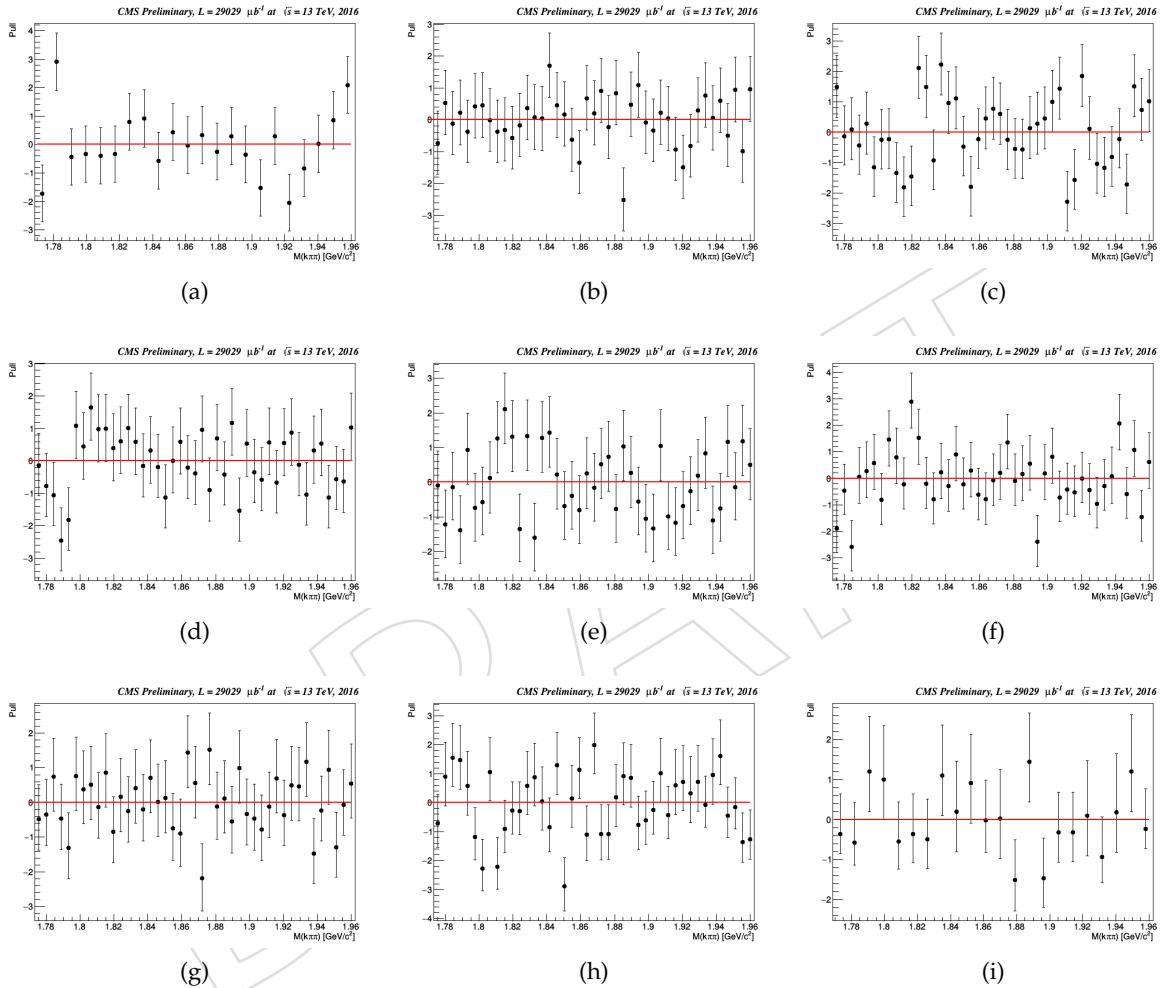


Figure 19: Pull of the fit on the $m(K\pi\pi)$ distributions in different p_T bins for $|\eta| < 2.1$: 4-5 GeV/c (a), 5-6 GeV/c (b), 6-7 GeV/c (c), 7-8 GeV/c (d), 8-12 GeV/c (e), 12-16 GeV/c (f), 16-24 GeV/c (g), 24-40 GeV/c (h) and 40-100 GeV/c (i) integrating in η .

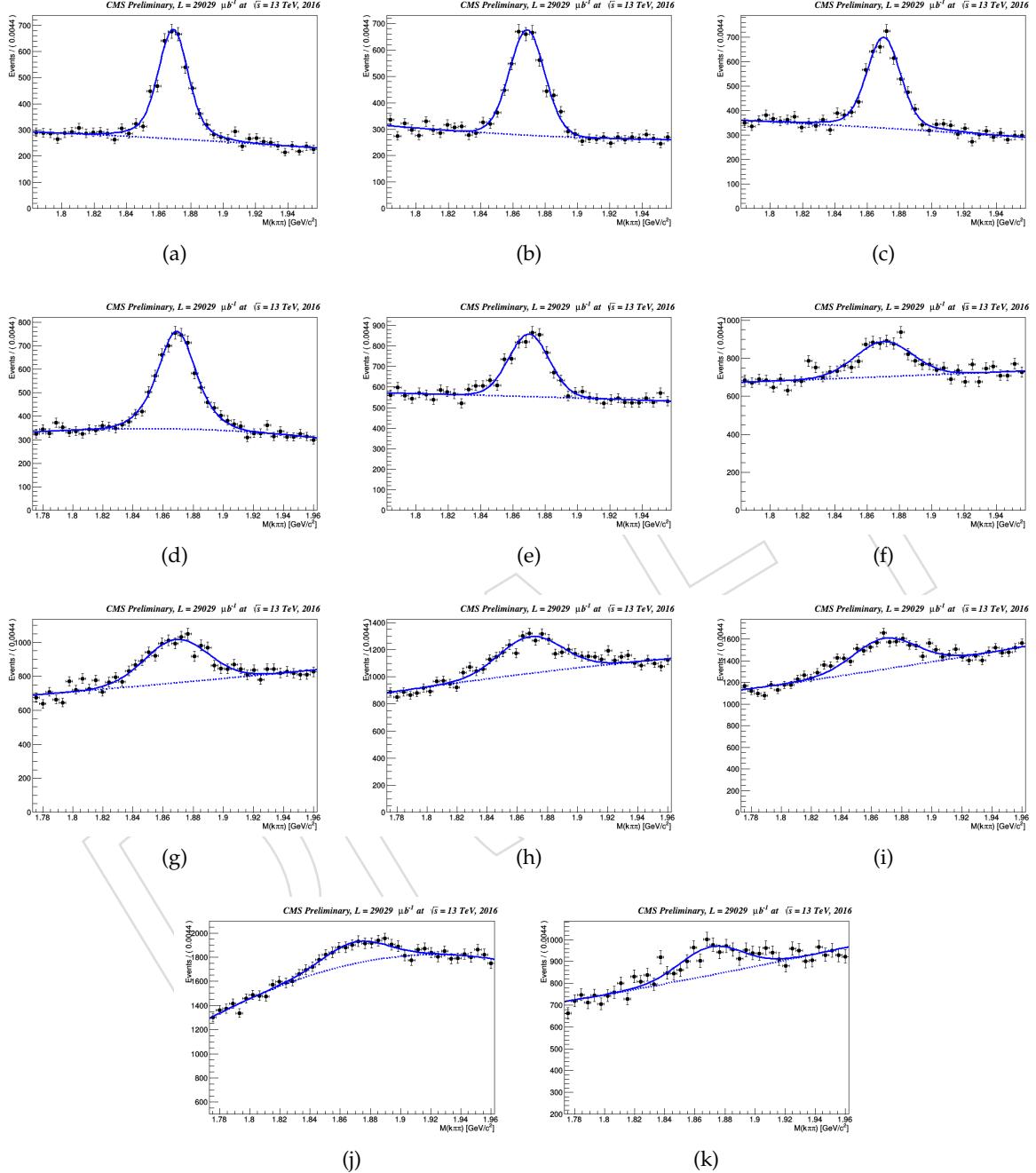


Figure 20: Fit on the $m(K\pi\pi)$ distributions in different $|\eta|$ bins: 0-0.2 (a), 0.2-0.4 (b), 0.4-0.6 (c), 0.6-0.8 (d), 0.8-1 (e), 1-1.2 (f), 1.2-1.4 (g), 1.4-1.6 (h), 1.6-1.8 (i), 1.8-2 (j) and 2-2.1 (k) for D^+ mesons with $4 < p_T < 100 \text{ GeV}/c$.

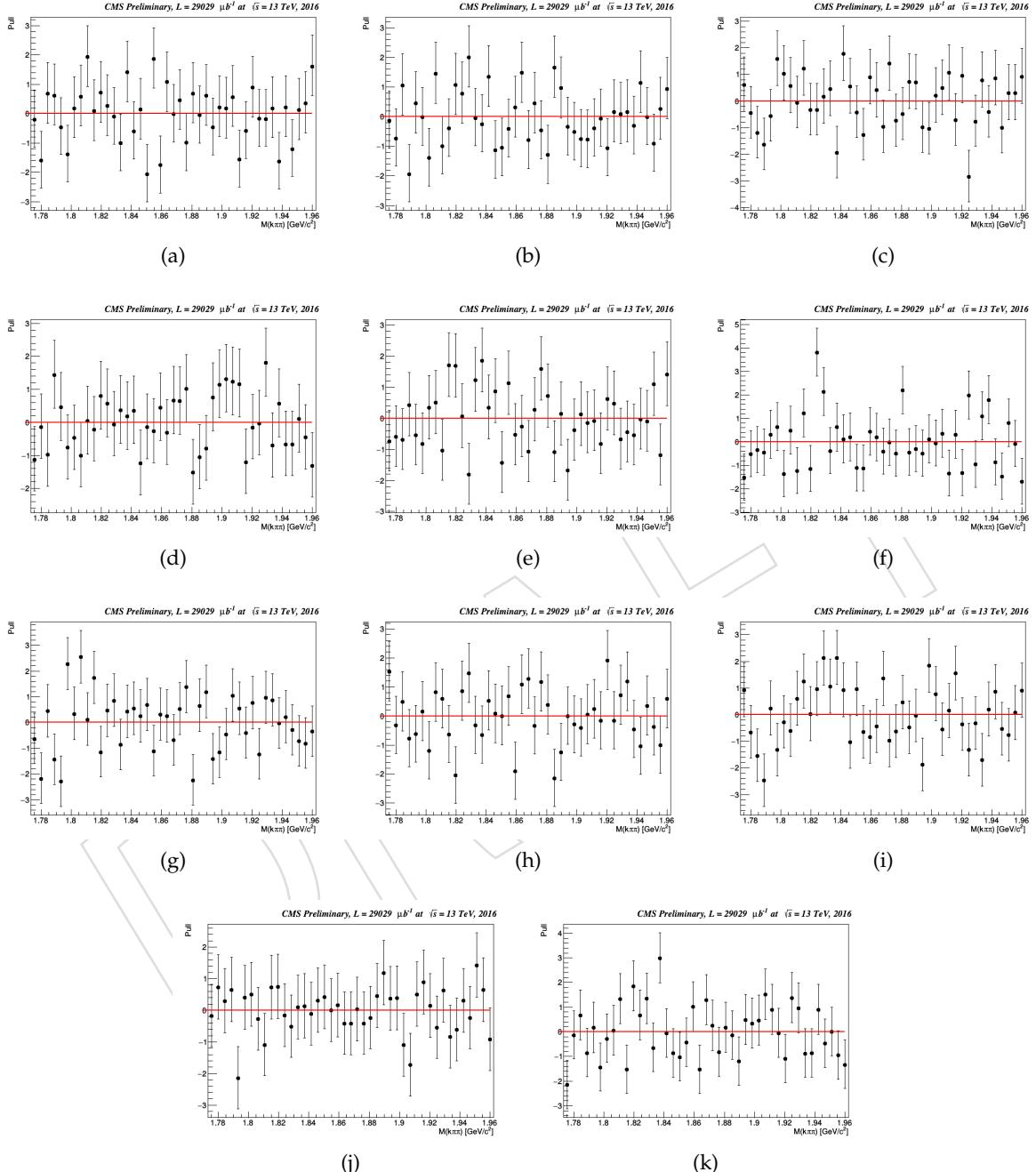


Figure 21: Pull of the fit on the $m(K\pi\pi)$ distributions in different $|\eta|$ bins: 0-0.2 (a), 0.2-0.4 (b), 0.4-0.6 (c), 0.6-0.8 (d), 0.8-1 (e), 1-1.2 (f), 1.2-1.4 (g), 1.4-1.6 (h), 1.6-1.8 (i), 1.8-2 (j) and 2-2.1 (k) for D^+ mesons with $4 < p_T < 100 \text{ GeV}/c$.

261 **5.4 Fit summary**

262 To provide a brief summary and a better global picture of all the fits performed in the p_T and
 263 η bins we report in this section the distributions of the combined width of the two gaussian
 264 system used for the signal peak modelling, the common mean and the signal yield for each
 265 bin.

266 The combined width is defined as $\sigma_{comb} = \sqrt{(f \cdot \sigma_1^2 + (1-f) \cdot \sigma_2^2)}$, where σ_1 and σ_2 are the
 267 width of the two single gaussians and f is the fraction between the two in the combination. The
 268 points are reported in Figure 22 for D^* (left), D^0 (center) and D^+ (right) mesons as a function
 269 of p_T (top) and η (bottom).

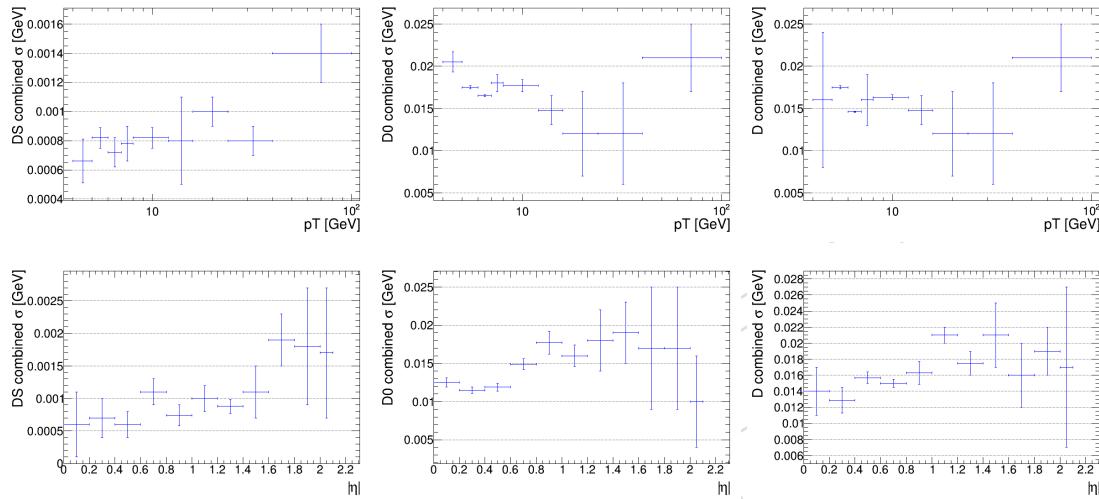


Figure 22: The combined width of the two gaussian system is reported for D^{*+} (left), D^0 (center) and D^+ (right) mesons as a function of p_T (top) and η (bottom).

270 Also the common mean of the two gaussians, that is a free parameter in the fit, has been
 271 scanned. The values and the respective errors are reported in Figure 23 for D^{*+} (left), D^0
 272 (center) and D^+ (right) mesons as a function of p_T (top) and η (bottom); the red line represents
 273 the value foreseen by the PDG for each meson and the box the PDG uncertainty.

274 For the D^* meson we see a global overestimation of the ΔM measurement with respect to the
 275 PDG expectation of about 0.1 MeV, while the central value of the mass of the other two mesons
 276 are in agreement with the PDG expectations within the errors.

277 Finally in Figure 24 the signal yield and the statistical errors coming from the fits are reported
 278 for D^{*+} (left), D^0 (center) and D^+ (right) mesons as a function of p_T (top) and η (bottom).

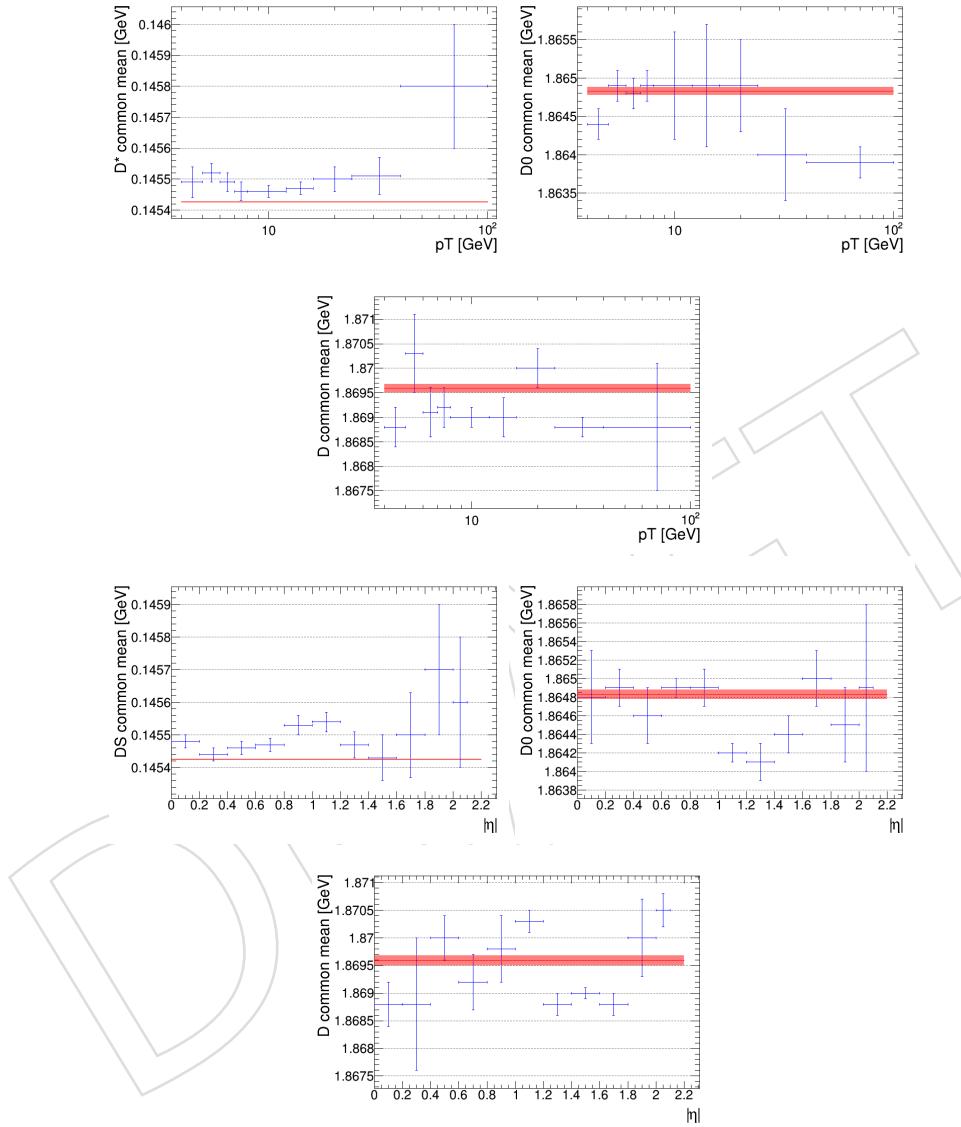


Figure 23: The common mean of the two gaussians, i.e. the central value of the mesons' mass is reported for $D^{\ast+}$ (left), D^0 (center) and D^+ (right) mesons as a function of p_T (top) and η (bottom); the red line represents the value foreseen by the PDG for each meson and the box the PDG uncertainty.

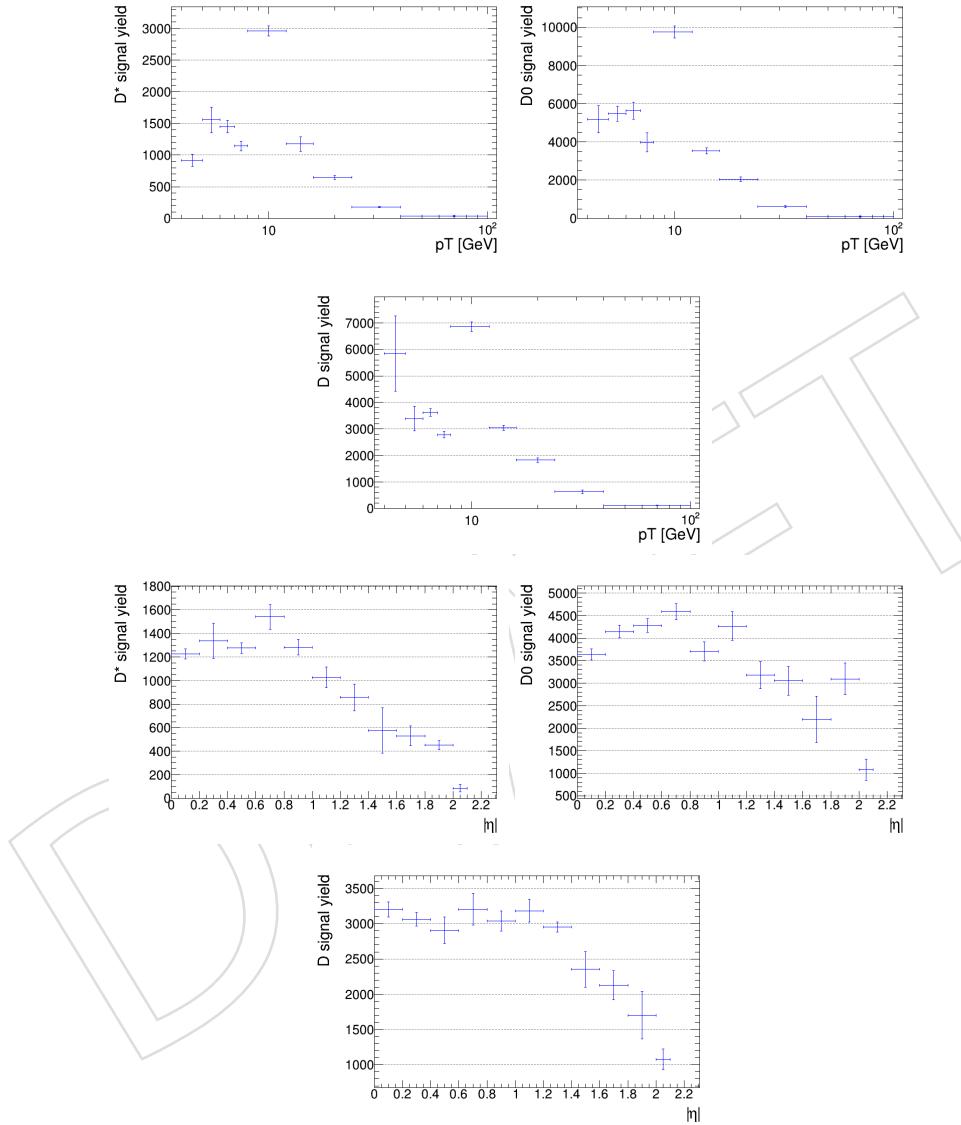


Figure 24: The signal yield provided by the fit is reported for D^{*+} (left), D^0 (center) and D^+ (right) mesons as a function of p_T (top) and η (bottom).

279 6 Efficiency Measurement

280 The efficiency measurements have been performed on the Monte Carlo sample reported in the
 281 first row on Table 3. This sample was enriched in D^* meson decaying in $D^0(\rightarrow K\pi)\pi_{slow}$ but
 282 the statistic was enough to calculate the efficiency for all the three charmed mesons considered
 283 in this analysis.

284 Before going through the detail of the efficiency calculation, in Figure 25 a comparison between
 285 data and the enriched MC is provided. In particular the distributions show the kinematic
 286 variables (p_T , η and ϕ) for all the reconstructed object in the D^{*+} decay mode in the signal
 287 region $0.144 < \Delta M < 0.147 \text{ GeV}/c^2$. The sideband method has been applied to subtract the
 288 combinatorial background events present in such range with the sideband regions defined as
 289 $0.1425 < \Delta M < 0.144 \text{ GeV}/c^2$ and $0.147 < \Delta M < 0.1485 \text{ GeV}/c^2$. Data and Monte Carlo shapes
 290 are in good agreement within the uncertainties.

291 Figure 26 shows then the kinematic distributions for the D^0 and D^+ mesons with the com-
 292 parison between data and the enriched MC. The same method of sideband is used defin-
 293 ing the signal region as $1.84 < M < 1.89 \text{ GeV}/c^2$ and the background regions to subtract
 294 $1.805 < M < 1.83 \text{ GeV}/c^2$ and $1.90 < M < 1.925 \text{ GeV}/c^2$ for both D^0 and D^+ mesons.

295 We used the following method for the efficiency calculation: we first selected at gen level the
 296 events containing the requested decay chain and then we applied the reconstruction strategy
 297 and the acceptance cuts described on Section 4 on such sub-sample. In order to ensure that
 298 the reconstructed event corresponds to the generated one we applied a geometrical matching
 299 between the two mesons. For example, considering the D^{*+} mesons, we can define $\Delta\eta =$
 300 $|\eta_{D^{*+}gen} - \eta_{D^{*+}reco}|$ and $\Delta\phi = |\phi_{D^{*+}gen} - \phi_{D^{*+}reco}|$. Thus the matching is operated on the quan-
 301 tity $\Delta R = \sqrt{\Delta\eta^2 + \Delta\phi^2}$. The distributions of $\Delta\eta$ and $\Delta\phi$ variables are reported in Figure 27.
 302 The matching between generated and reconstructed mesons is considered satisfied if $\Delta R < 0.4$,
 303 the distribution of such variable is shown in Figure 28 (left) with the red line indicating the re-
 304 quest applied; with this cut all the events on the tail are rejected. Figure 28 (right) shows the
 305 $\Delta p_T = |p_{TD^{*+}gen} - p_{TD^{*+}reco}|$ distribution for events that satisfy the matching condition in order
 306 to show that the geometrical matching has a powerful effect also on the kinematic. Most of the
 307 events in fact have a very small Δp_T value, as expected.

308 The efficiency was thus calculated for each p_T and η bin and for both the charged states as:

$$\varepsilon_i = \frac{N_{rec_i}(D \rightarrow f)}{TOT_i(D \rightarrow f)} \quad (7)$$

309 where N_{rec_i} is the number of reconstructed events that pass the matching condition, evaluated
 310 in the same way described in Section 5 fitting the mass distributions, and TOT_i is the number
 311 of generated events in such bin. In the following tables we report the efficiency calculations for
 312 D^{*+} (Table 14, 15), D^0 (Table 16, 17) and D^+ (Table 18, 19) (the c.c. are summed) in all the p_T
 313 and η bins analyzed. The statistical errors quoted by the fit on the signal yield is very close to
 314 \sqrt{N} in most of the p_T and η bins, since the combinatorial background is minimised matching
 315 the reconstructed signal events to the generated ones.

316 In the Appendix has been also reported a stability test in order to figure out how the efficiency
 317 values change with respect to the cut on ΔR . From $\Delta R < 0.4$ to $\Delta R < 0.6$ the efficiency
 318 variations are almost negligible.

p_T range [GeV/c]	$N_i(D \rightarrow f + c.c)$	$TOT_i(D \rightarrow f + c.c)$	ε_i
[4-5]	101970 ± 344	12716660	$(8.02 \pm 0.03) \times 10^{-3}$
[5-6]	143440 ± 390	6212347	$(2.309 \pm 0.006) \times 10^{-2}$
[6-7]	137960 ± 380	3236898	$(4.262 \pm 0.011) \times 10^{-2}$
[7-8]	114150 ± 340	1792115	$(6.37 \pm 0.02) \times 10^{-2}$
[8-12]	264580 ± 523	2352339	$(1.125 \pm 0.002) \times 10^{-1}$
[12-16]	91664 ± 307	459560	$(1.995 \pm 0.006) \times 10^{-1}$
[16-24]	46857 ± 220	183230	$(2.557 \pm 0.010) \times 10^{-1}$
[24-40]	12331 ± 113	41488	$(2.97 \pm 0.02) \times 10^{-1}$
[40-100]	1726 ± 42	5698	$(3.03 \pm 0.06) \times 10^{-1}$

Table 14: Efficiency calculation in p_T bins for D^{*+} mesons in the pseudorapidity range $|\eta| < 2.1$.

$ \eta $ range	$N_i(D \rightarrow f + c.c)$	$TOT_i(D \rightarrow f + c.c)$	ε_i
[0-0.2]	114910 ± 345	2608159	$(4.406 \pm 0.013) \times 10^{-2}$
[0.2-0.4]	117160 ± 349	2613410	$(4.483 \pm 0.013) \times 10^{-2}$
[0.4-0.6]	118620 ± 351	2625104	$(4.519 \pm 0.013) \times 10^{-2}$
[0.6-0.8]	118100 ± 350	2630884	$(4.489 \pm 0.013) \times 10^{-2}$
[0.8-1]	107960 ± 335	2627132	$(4.109 \pm 0.012) \times 10^{-2}$
[1-1.2]	89699 ± 307	2614841	$(3.430 \pm 0.011) \times 10^{-2}$
[1.2-1.4]	72324 ± 276	2589700	$(2.793 \pm 0.010) \times 10^{-2}$
[1.4-1.6]	60902 ± 254	2551391	$(2.387 \pm 0.001) \times 10^{-2}$
[1.6-1.8]	52675 ± 237	2503081	$(2.1044 \pm 0.0009) \times 10^{-2}$
[1.8-2]	44687 ± 219	2441432	$(1.8304 \pm 0.0009) \times 10^{-2}$
[2-2.1]	18418 ± 142	1195203	$(1.541 \pm 0.0011) \times 10^{-2}$

Table 15: Efficiency calculation in η bins for D^{*+} mesons with $4 < p_T < 100$ GeV/c.

p_T range [GeV/c]	$N_i(D \rightarrow f + c.c)$	$TOT_i(D \rightarrow f + c.c)$	ε_i
[4-5]	137110 ± 388	10884790	$(1.260 \pm 0.003) \times 10^{-2}$
[5-6]	126120 ± 380	5182344	$(2.434 \pm 0.007) \times 10^{-2}$
[6-7]	107300 ± 339	2653959	$(4.043 \pm 0.012) \times 10^{-2}$
[7-8]	85539 ± 486	1450841	$(5.90 \pm 0.02) \times 10^{-2}$
[8-12]	193280 ± 467	1876943	$(1.030 \pm 0.002) \times 10^{-1}$
[12-16]	67932 ± 282	362693	$(1.873 \pm 0.007) \times 10^{-1}$
[16-24]	36895 ± 206	143316	$(2.574 \pm 0.012) \times 10^{-1}$
[24-40]	10470 ± 116	32373	$(3.23 \pm 0.03) \times 10^{-1}$
[40-100]	1547 ± 44	4317	$(3.58 \pm 0.07) \times 10^{-1}$

Table 16: Efficiency calculation in p_T bins for D^0 mesons in the pseudorapidity range $|\eta| < 2.1$.

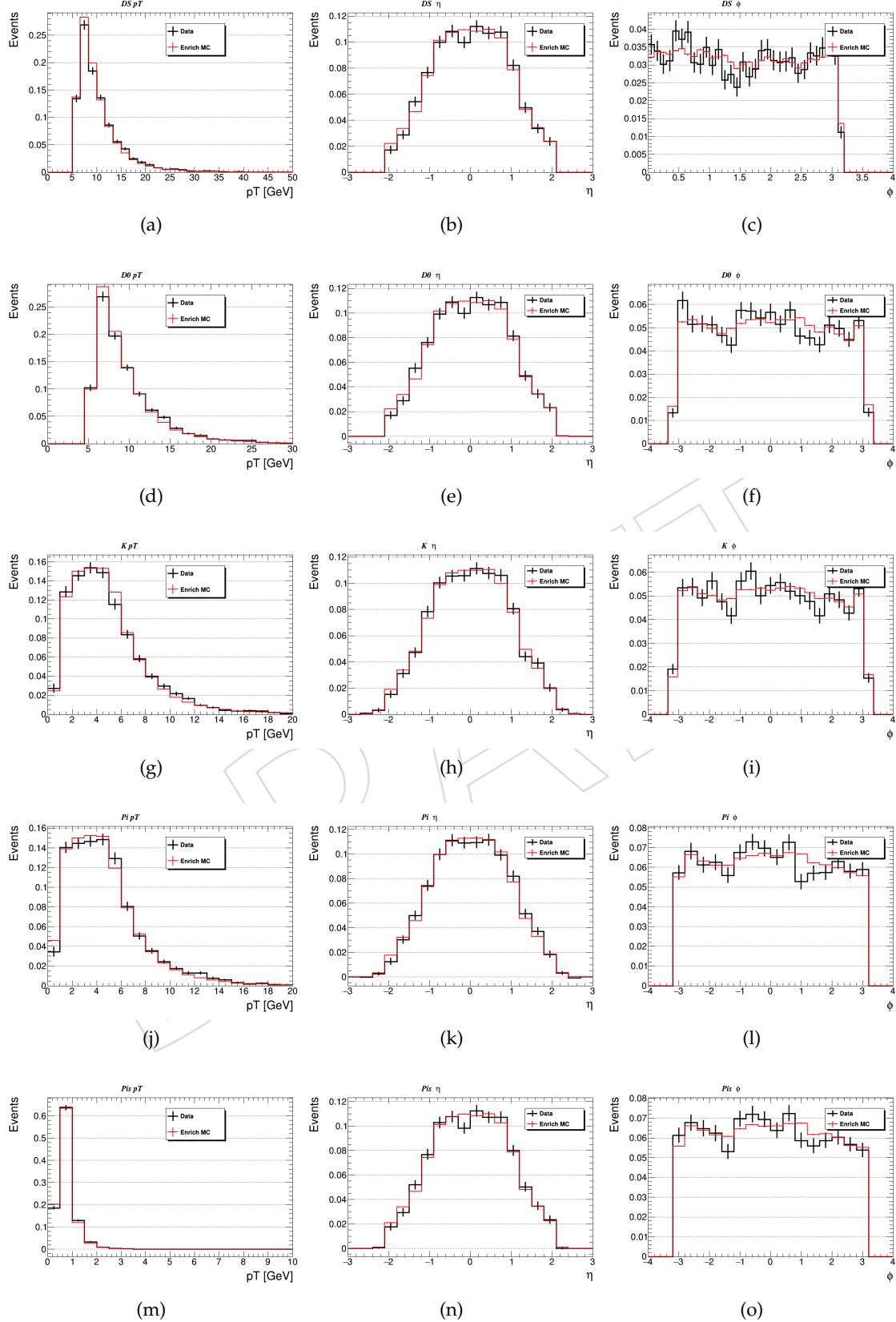


Figure 25: Data-Monte Carlo comparison for the kinematics of the objects reconstructed in the D^{*+} decay. All the distributions are done with events selected in the signal region $0.144 < \Delta M < 0.147 \text{ GeV}/c^2$, background events are subtracted with the sideband method. $D^{*+} p_T$ (a), $D^{*+} \eta$ (b), $D^{*+} \phi$ (c), $D^0 p_T$ (d), $D^0 \eta$ (e), $D^0 \phi$ (f), $K p_T$ (g), $K \eta$ (h), $K \phi$ (i), πp_T (j), $\pi \eta$ (k), $\pi \phi$ (l), $\pi_{slow} p_T$ (m), $\pi_{slow} \eta$ (n) and $\pi_{slow} \phi$ (o).

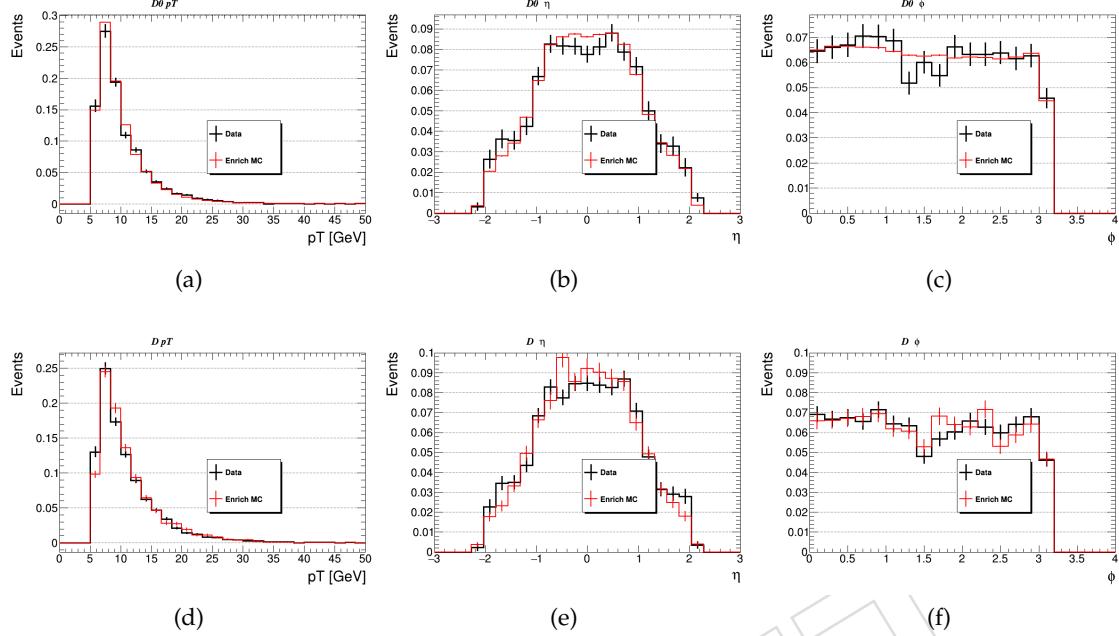


Figure 26: Data-Monte Carlo comparison for the kinematics of the D^0 and D^+ mesons. All the distributions are done with events selected in the signal region $1.84 < M < 1.89 \text{ GeV}/c^2$, background events are subtracted with the sideband method. $D^0 p_T$ (a), $D^0 \eta$ (b), $D^0 \phi$ (c), $D^+ p_T$ (d), $D^+ \eta$ (e), $D^+ \phi$ (f).

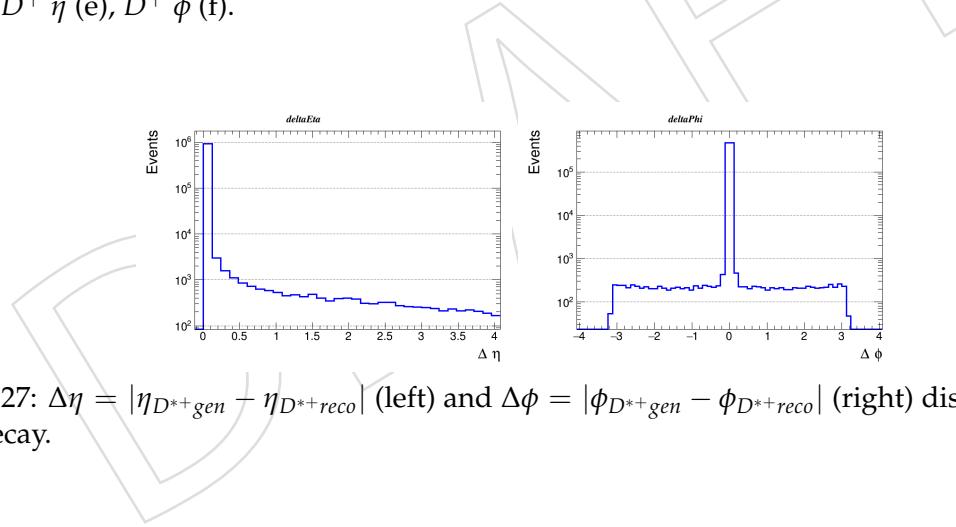


Figure 27: $\Delta\eta = |\eta_{D^{*+}gen} - \eta_{D^{*+}reco}|$ (left) and $\Delta\phi = |\phi_{D^{*+}gen} - \phi_{D^{*+}reco}|$ (right) distributions for D^{*+} decay.

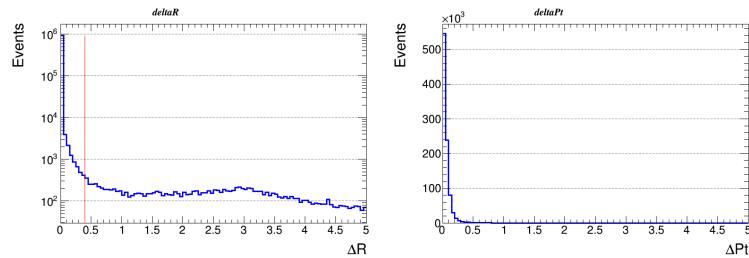


Figure 28: $\Delta R = \sqrt{\Delta\eta^2 + \Delta\phi^2}$ distribution on the left, the red line shows the cut applied for the matching. On the right the distribution of the $\Delta p_T = |p_{TD^{*+}gen} - p_{TD^{*+}reco}|$ shows the kinematic agreement between the generated and reconstructed D^{*+} mesons after the matching.

$ \eta $ range	$N_i(D \rightarrow f + c.c)$	$TOT_i(D \rightarrow f + c.c)$	ε_i
[0-0.2]	91376 ± 303	2200284	$(4.153 \pm 0.014) \times 10^{-2}$
[0.2-0.4]	92521 ± 307	2203023	$(4.200 \pm 0.014) \times 10^{-2}$
[0.4-0.6]	93690 ± 308	2208043	$(4.243 \pm 0.014) \times 10^{-2}$
[0.6-0.8]	93789 ± 310	2208180	$(4.247 \pm 0.014) \times 10^{-2}$
[0.8-1]	87902 ± 270	2201363	$(3.993 \pm 0.013) \times 10^{-2}$
[1-1.2]	78014 ± 293	2188035	$(3.565 \pm 0.013) \times 10^{-2}$
[1.2-1.4]	66781 ± 500	2161141	$(3.090 \pm 0.012) \times 10^{-2}$
[1.4-1.6]	55659 ± 255	2125066	$(2.619 \pm 0.011) \times 10^{-2}$
[1.6-1.8]	49167 ± 100	2080986	$(2.363 \pm 0.011) \times 10^{-2}$
[1.8-2]	40127 ± 173	2024836	$(1.982 \pm 0.010) \times 10^{-2}$
[2-2.1]	17046 ± 143	990618	$(1.721 \pm 0.013) \times 10^{-2}$

Table 17: Efficiency calculation in η bins for D^0 mesons with $4 < p_T < 100$ GeV/c.

p_T range [GeV/c]	$N_i(D \rightarrow f + c.c)$	$TOT_i(D \rightarrow f + c.c)$	ε_i
[4-5]	600 ± 26	63858	$(9.4 \pm 0.4) \times 10^{-3}$
[5-6]	655 ± 11	39977	$(1.64 \pm 0.06) \times 10^{-2}$
[6-7]	722 ± 27	25496	$(2.83 \pm 0.10) \times 10^{-2}$
[7-8]	630 ± 18	16054	$(3.92 \pm 0.15) \times 10^{-2}$
[8-12]	1851 ± 44	25705	$(7.20 \pm 0.16) \times 10^{-2}$
[12-16]	804 ± 34	5983	$(1.34 \pm 0.04) \times 10^{-1}$
[16-24]	546 ± 20	2680	$(2.04 \pm 0.08) \times 10^{-1}$
[24-40]	155 ± 6	629	$(2.46 \pm 0.17) \times 10^{-1}$
[40-100]	21 ± 5	99	$(2.1 \pm 0.4) \times 10^{-1}$

Table 18: Efficiency calculation in p_T bins for D^+ mesons in the pseudorapidity range $|\eta| < 2.1$.

$ \eta $ range	$N_i(D \rightarrow f + c.c)$	$TOT_i(D \rightarrow f + c.c)$	ε_i
[0-0.2]	825 ± 29	17964	$(4.59 \pm 0.16) \times 10^{-2}$
[0.2-0.4]	808 ± 29	17744	$(4.55 \pm 0.16) \times 10^{-2}$
[0.4-0.6]	840 ± 29	17770	$(4.73 \pm 0.16) \times 10^{-2}$
[0.6-0.8]	784 ± 29	17890	$(4.38 \pm 0.15) \times 10^{-2}$
[0.8-1]	673 ± 19	17585	$(3.83 \pm 0.15) \times 10^{-2}$
[1-1.2]	596 ± 23	17692	$(3.37 \pm 0.14) \times 10^{-2}$
[1.2-1.4]	440 ± 17	17202	$(2.56 \pm 0.12) \times 10^{-2}$
[1.4-1.6]	357 ± 11	16825	$(2.12 \pm 0.11) \times 10^{-2}$
[1.6-1.8]	329 ± 8	16432	$(2.00 \pm 0.11) \times 10^{-2}$
[1.8-2]	240 ± 7	15687	$(1.53 \pm 0.10) \times 10^{-2}$
[2-2.1]	117 ± 12	7698	$(1.52 \pm 0.14) \times 10^{-2}$

Table 19: Efficiency calculation in η bins for D^+ mesons with $4 < p_T < 100$ GeV/c.

319 7 Secondary decay contamination

320 Since the goal of the analysis is to measure the prompt open charm cross section production,
 321 it is important to evaluate the contribution coming from the non prompt charmed mesons. It
 322 was already specified that are considered as prompt both the charmed mesons coming directly
 323 from the pp interaction and the mesons coming from the decay of other excited charmed mesons
 324 (i.e. D^0 s from the D^{*+} decay are considered prompt). Thus the main source of contamination
 325 comes from the B-meson decay; we indicate such events as secondary decay contamination.

326 Before discussing in detail the contamination evaluation, in Figure 29 a comparison between
 327 data and the MinimumBias MC used in this step is provided. In particular the distributions
 328 shown are the kinematic variables (p_T , η and ϕ) for all the reconstructed object in the D^{*+}
 329 decay mode in the signal region $0.144 < \Delta M < 0.147 \text{ GeV}/c^2$. The sideband method has
 330 been applied to subtract the combinatorial background events present in such range with the
 331 sideband regions defined as $0.1425 < \Delta M < 0.144 \text{ GeV}/c^2$ and $0.147 < \Delta M < 0.1485 \text{ GeV}/c^2$.
 332 Data and Monte Carlo shapes are almost in a good agreement within the uncertainties.

333 In Figure 30 the kinematic comparison between data and MinimumBias MC for the D^0 and D^+
 334 mesons. The same method of sideband is used defining the signal region as $1.84 < M < 1.89$
 335 GeV/c^2 and the background regions to subtract $1.805 < M < 1.83 \text{ GeV}/c^2$ and $1.90 < M <$
 336 $1.925 \text{ GeV}/c^2$ for both D^0 and D^+ mesons.

337 7.1 Cut based contamination evaluation

338 We tested our strategy on the MinimumBias MC sample reported in row 2 of Table 3 in or-
 339 der to quantify the contamination from secondary decay. Using the gen level information we
 340 identified two sub-samples: prompt charmed mesons and non-prompt charmed mesons. Thus
 341 we applied exactly the same strategy described in 4.2 to both of them, we computed the signal
 342 yields N_{prompt} and N_{sec} as reported in Section 5 and finally the contamination can be evaluated
 343 as

$$contam = \frac{N_{sec}}{N_{prompt} + N_{sec}} \quad (8)$$

344 for each p_T and η bin.

345 In the Tables 20 and 21 the contamination values for the three mesons are reported as a function
 346 of p_T and η bins respectively and the same values are shown in Figures 31 and 32. As usual
 347 both the charged signs are taken into account.

p_T range [GeV/c]	D^{*+} contamination	D^0 contamination	D^+ contamination
[4-5]	0.14 ± 0.02	0.089 ± 0.006	0.099 ± 0.012
[5-6]	0.15 ± 0.02	0.093 ± 0.006	0.10 ± 0.02
[6-7]	0.16 ± 0.03	0.107 ± 0.008	0.106 ± 0.013
[7-8]	0.10 ± 0.03	0.088 ± 0.008	0.091 ± 0.014
[8-12]	0.10 ± 0.02	0.106 ± 0.007	0.108 ± 0.010
[12-16]	0.15 ± 0.04	0.095 ± 0.005	0.12 ± 0.02
[16-24]	0.15 ± 0.05	0.08 ± 0.02	0.16 ± 0.02
[24-40]	0.14 ± 0.09	0.08 ± 0.03	0.07 ± 0.03
[40-100]	0.15 ± 0.18	0.05 ± 0.06	0.09 ± 0.06

Table 20: Contamination computed for the three mesons in p_T bins.

$ \eta $ range	D^{*+} contamination	D^0 contamination	D^+ contamination
[0-0.2]	0.12 ± 0.02	0.07 ± 0.01	0.084 ± 0.012
[0.2-0.4]	0.15 ± 0.03	0.061 ± 0.007	0.093 ± 0.006
[0.4-0.6]	0.12 ± 0.03	0.072 ± 0.010	0.098 ± 0.014
[0.6-0.8]	0.12 ± 0.01	0.086 ± 0.010	0.112 ± 0.013
[0.8-1]	0.08 ± 0.02	0.095 ± 0.011	0.103 ± 0.013
[1-1.2]	0.17 ± 0.04	0.10 ± 0.01	0.133 ± 0.014
[1.2-1.4]	0.13 ± 0.03	0.091 ± 0.013	0.103 ± 0.009
[1.4-1.6]	0.14 ± 0.04	0.13 ± 0.02	0.126 ± 0.014
[1.6-1.8]	0.15 ± 0.04	0.12 ± 0.02	0.12 ± 0.02
[1.8-2]	0.11 ± 0.04	0.15 ± 0.03	0.13 ± 0.02
[2-2.1]	0.18 ± 0.09	0.10 ± 0.04	0.18 ± 0.05

Table 21: Contamination computed for the three mesons in η bins.

348 For the last p_T bin ($[40 - 100] GeV/c$) it was not possible to evaluate the contamination due to
 349 the lack of statistics, so both the value and the respective uncertainty are linearly extrapolated
 350 from all the previous values; only for the D meson we assigned to the last bin the rate of the
 351 penultimate bin since the extrapolation gave a negative number.

352 The contamination values are not negligible in any range and for any reconstructed meson.
 353 This was expected, since the strategy applied, in particular the cut on the decay lifetime that
 354 is needed to discriminate the signal from the combinatorial background, tends to select long-
 355 life mesons. Indeed, it is also interesting to study the contamination rate as a function of the
 356 L/σ_L variable. The cut we apply on the L significance, discussed in the reconstruction strategy
 357 section 4, is one of the more powerful for the combinatorial background reduction but, on the
 358 other hand, selecting mesons with higher L/σ_L could increase the contamination from non
 359 prompt charmed mesons. This is clearly shown in Figure 33: the contamination rate grows
 360 up as the L significance value increases for each charmed meson. Such effect is particularly
 361 evident for the D^{*+} and D^0 mesons, while the increasing seems to be less noticeable for the D^+
 362 meson, that is the one with the longer decay length between the three.

363 Since a strong dependence of the contamination on the L/σ_L is spotted, as expected, the two
 364 most sensitive variable distributions have been compared between data and MC, in addition to
 365 the kinematic distributions reported in Figure 29. On Figure 34 the distributions of $\cos\phi$ and
 366 L/σ_L are thus reported showing a global good agreement, within the statistical uncertainty,
 367 between data and the Minimum Bias Monte Carlo used for the contamination evaluation.

368 The non prompt component of the signal is thus evaluated trusting the results obtained on the
 369 MC sample. The quantities reported in the Tables 20 and 21 will be subtracted from the values
 370 of the visible cross section for each p_T and η bin.

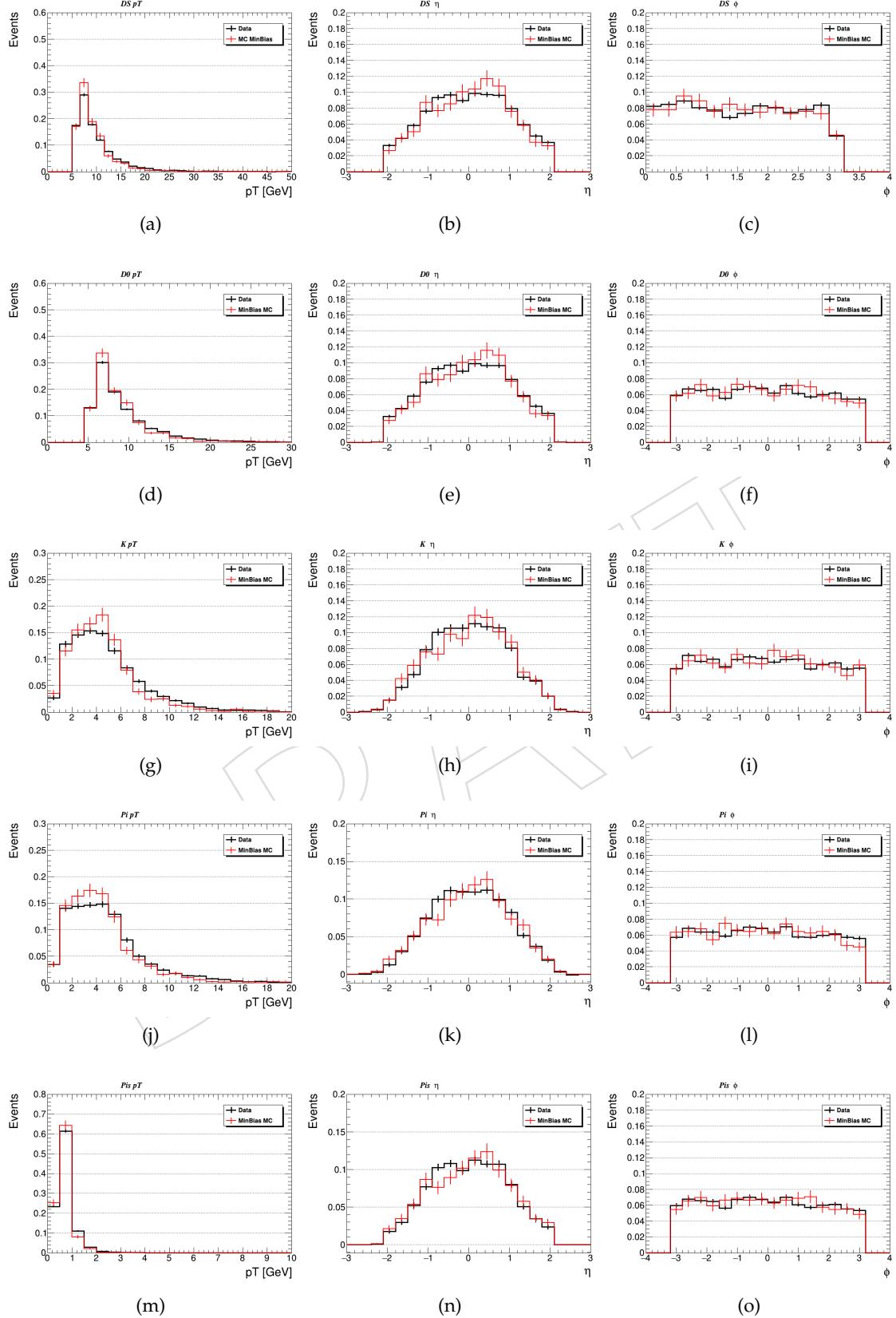


Figure 29: Data-Monte Carlo comparison for the kinematics of the objects reconstructed in the D^{*+} decay. All the distributions are done with events selected in the signal region $0.144 < \Delta M < 0.147 \text{ GeV}/c^2$, background events are subtracted with the sideband method. $D^{*+} p_T$ (a), $D^{*+} \eta$ (b), $D^{*+} \phi$ (c), $D^0 p_T$ (d), $D^0 \eta$ (e), $D^0 \phi$ (f), $K p_T$ (g), $K \eta$ (h), $K \phi$ (i), πp_T (j), $\pi \eta$ (k), $\pi \phi$ (l), $\pi_{slow} p_T$ (m), $\pi_{slow} \eta$ (n) and $\pi_{slow} \phi$ (o).

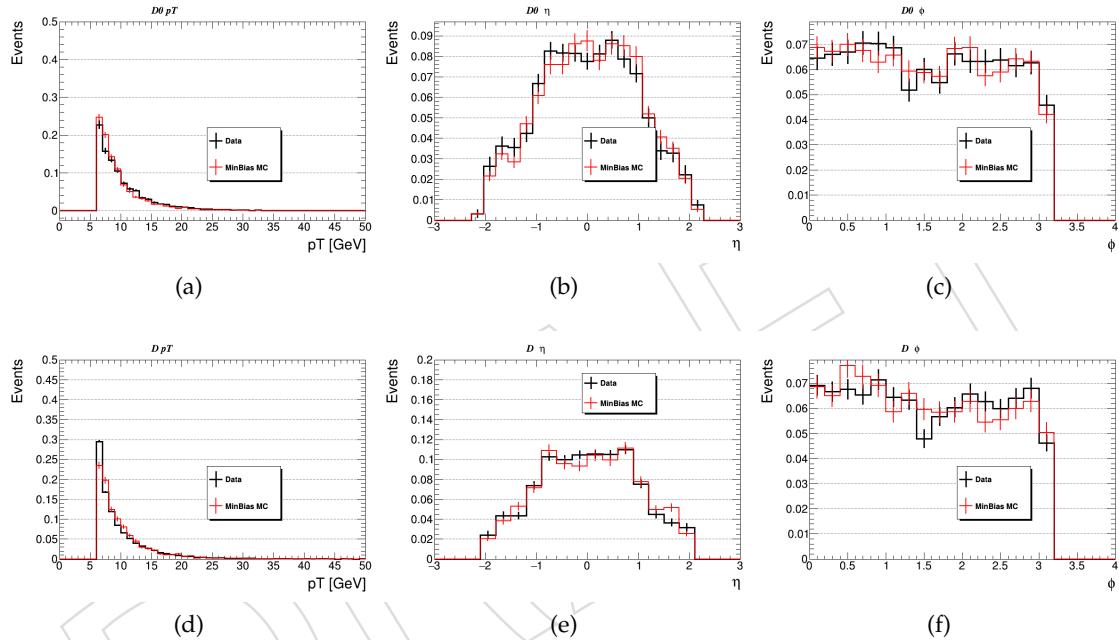


Figure 30: Data-Monte Carlo comparison for the kinematics of the D^0 and D^+ mesons. All the distributions are done with events selected in the signal region $1.84 < M < 1.89 \text{ GeV}/c^2$, background events are subtracted with the sideband method. $D^0 p_T$ (a), $D^0 \eta$ (b), $D^0 \phi$ (c), $D^+ p_T$ (d), $D^+ \eta$ (e), $D^+ \phi$ (f).

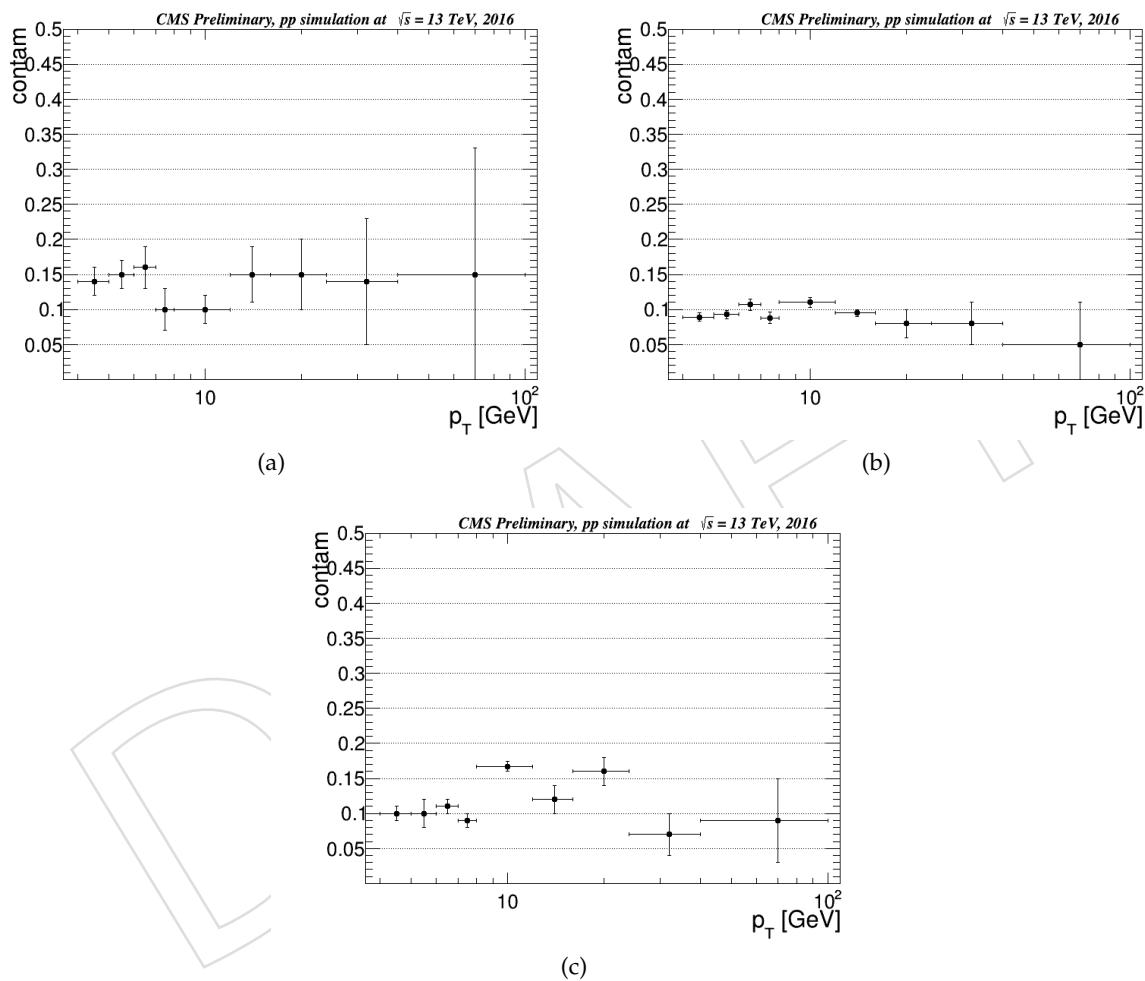


Figure 31: Contamination reported as a function of p_T for: D^{*+} (a), D^0 (b)and D^+ (c) mesons.

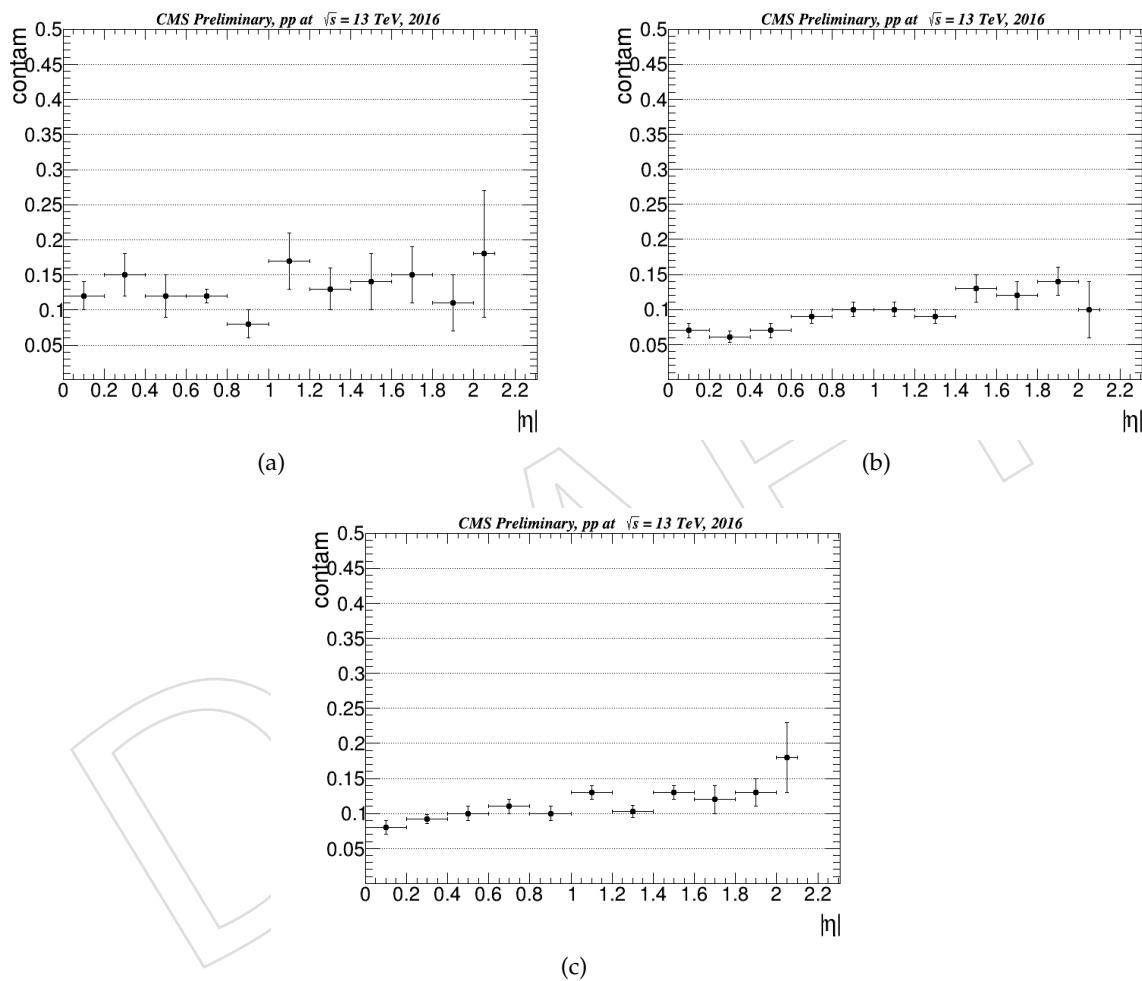


Figure 32: Contamination reported as a function of η for: D^{*+} (a), D^0 (b) and D^+ (c) mesons.

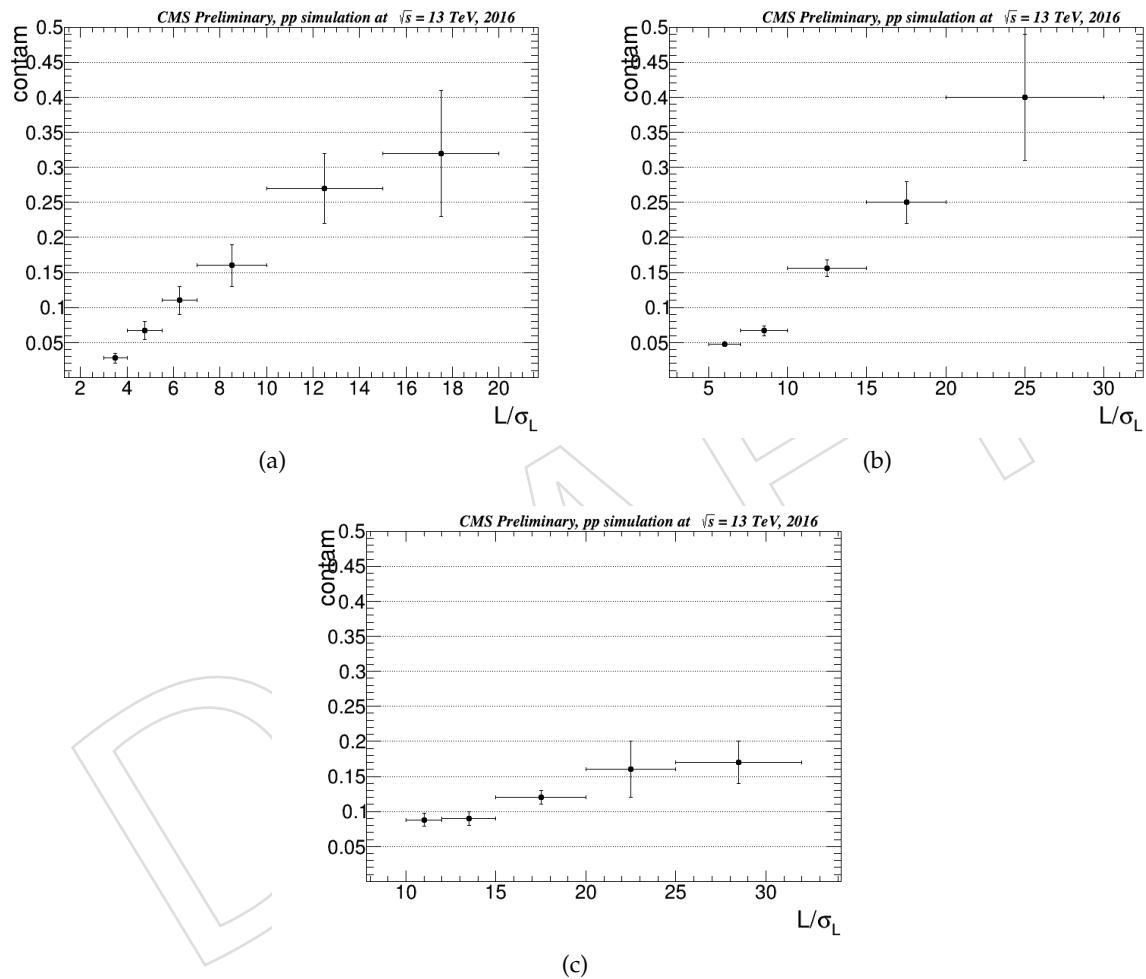


Figure 33: Contamination reported as a function of L/σ_L for: D^{*+} (a), D^0 (b) and D^+ (c) mesons.

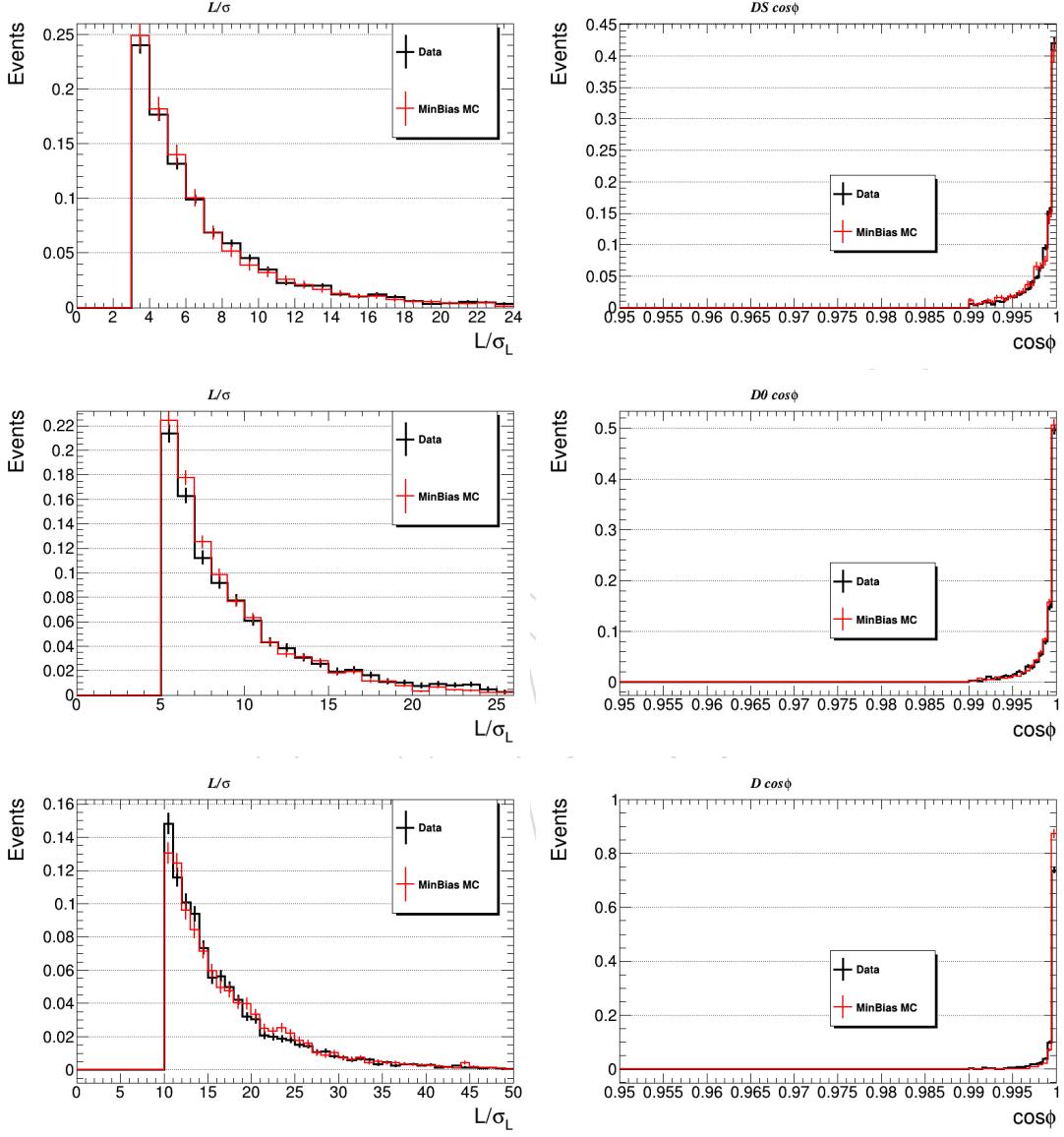


Figure 34: The two most sensitive variables in the D mesons reconstruction are reported here, L/σ on the left and $\cos\phi$ in the right, comparing data (black) to the Minimum Bias Monte Carlo used for the contamination evaluation (red) for $D^{\ast+}$ (top), D^0 (center) and D^+ (bottom); a global good agreement within the statistical uncertainty is shown.

371 8 Systematic Uncertainties

372 Several systematic uncertainty sources have been considered.

373 **Luminosity** The uncertainty related to the luminosity measurement procedure in CMS in 2016
 374 is 2.5% [9], to be propagated.

375 **Branching Ratio** The values of the Branching Ratio of the considered decays are taken from
 376 the PDG, the associated uncertainties are 0.03 % for the D^{*+} , 0.04 % for the D^0 and 0.28
 377 % for the D^+ , to be propagated.

378 **Tracking** A systematic uncertainty has to be assigned to each track reconstructed in the final
 379 state. From [10] we have 2.3 % per track, but since this value contains an uncertainty
 380 term coming from the PDG such component only once per final state and not summed
 381 up for each track. A different consideration is needed for the slow pion in the D^{*+} decay;
 382 since the cuts applied are considerably looser than the other two tracks a lower tracking
 383 efficiency is expected. There are no specific studies available on that, so we estimated
 384 the uncertainty related to the slow pion comparing the yield in data and Monte Carlo,
 385 varying π_{slowpT} and $\pi_{slow\eta}$. A systematic of 5.2 % is thus evaluated for the π_{slow} . Thus,
 386 summing up the uncertainties for each track in the final state, we have 9.4 % for the D^{*+} ,
 387 6.1 % for the D^+ with 3 tracks in the final state and 4.2 % for the D^0 that decays in two
 388 objects.

389 **Mass distribution modelling** As described in section 5 the signal yields have been computed
 390 modelling the resonance peaks with a double gaussian combination in order to take in
 391 account the different resolution effects in various kinematical regions. We evaluated the
 392 systematic related to the signal description using the sum of three gaussians, the fitted
 393 distributions are reported in Figure 35. The corresponding systematic uncertainties of the
 394 yield are 0.6 % for the D^{*+} , 3.4 % for the D^0 and 2.8 % for the D^+ . We also tried to use a
 395 different description for the signal peak, using a Crystal Ball, as shown in Figure 36. We
 396 got as corresponding systematic: 3.0 % for the D^{*+} , 1.1 % for the D^0 and 2.5 % for the D^+ .
 397 The biggest values between these two description has been thus taken as systematic un-
 398 certainty that corresponds to: 3.0 % for the D^{*+} , 3.4 % for the D^0 and 2.8 % for the D^+ .

399 For the combinatorial background description we used a particular function for the ΔM
 400 distribution, defined in equation 5, and a third degree polynomial function for the D^0 and
 401 D^+ mass distributions. For the systematic evaluation we performed the combinatorial fit
 402 using a fourth degree polynomial function for all of the three distributions as shown in
 403 Figure 37.

404 The corresponding systematic uncertainties are 1.2 % for the D^{*+} , 4.8 % for the D^0 and
 405 5.3 % for the D^+ .

406 MC sample statistic

407 We considered as systematic uncertainties the limited statistics of the MC sample used
 408 for the efficiency calculation. Bin by bin we used the uncertainties reported in Tables 14,
 409 15, 16, 17, 18 and 19. Weighted for the bin content we have globally 0.3 % for the D^{*+} ,
 410 0.3 % for the D^0 and 3.5 % for the D^+ (the sample is not enriched in such meson, so the
 411 higher uncertainty is expected).

412 Contamination evaluation

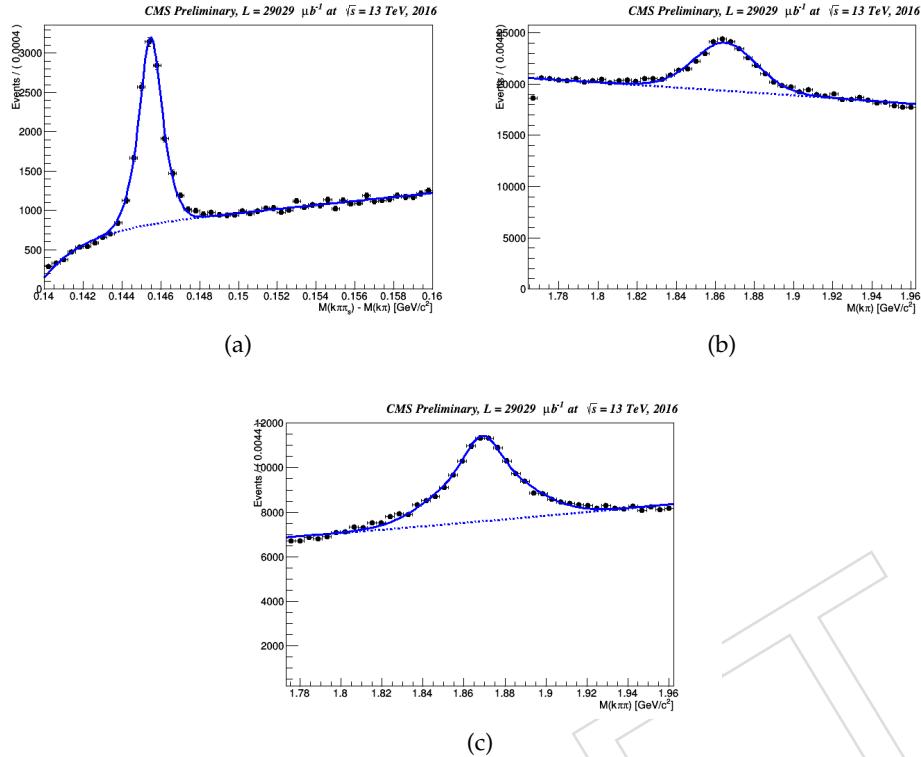


Figure 35: Invariant mass distribution with the signal peak described by the sum of three gaussians fit: D^{*+} (a), D^0 (b) and D^+ (c) mesons.

The systematic uncertainties related to the contamination evaluation have been computed taking into account the error related to the limited statistics of the MC sample used. Weighted for the bin content we have globally 2.9 % for the D^{*+} , 0.8 % for the D^0 and 1.4 % for the D^+ . Since the general good agreement between data and MC showed in Figure 34 we conclude there is no need to add any additional systematic.

Pile Up A section of the Appendix in this document has been dedicated to discuss that our analysis is not sensitive to the different PU scenario between data and MC. A further confirmation for this statement comes from the calculation of the corresponding systematic uncertainties. Calculating the production cross section as reported in Table 33 for the D^* meson at different PU values integrating in p_T and $|\eta|$, we can report these numbers on a plot and fit them with a 0-degree polynomial function, as showed in Figure 38.

The result of such fit is $p_0 = \bar{x} \pm \Delta\bar{x} = 22.24 \pm 1.24$ with a $\chi^2/ndof = 3.031/3$. Using the S-factor method [6] we can thus quantify a systematic uncertainty subtracting the statitical fluctuation. In particular we can define:

$$S = \sqrt{\chi^2/(N-1)} = 1.005 \quad (9)$$

and

$$\Delta\tilde{x} = S \cdot \Delta\bar{x} = 1.25 \quad (10)$$

thus the sistematic error can be calculated as

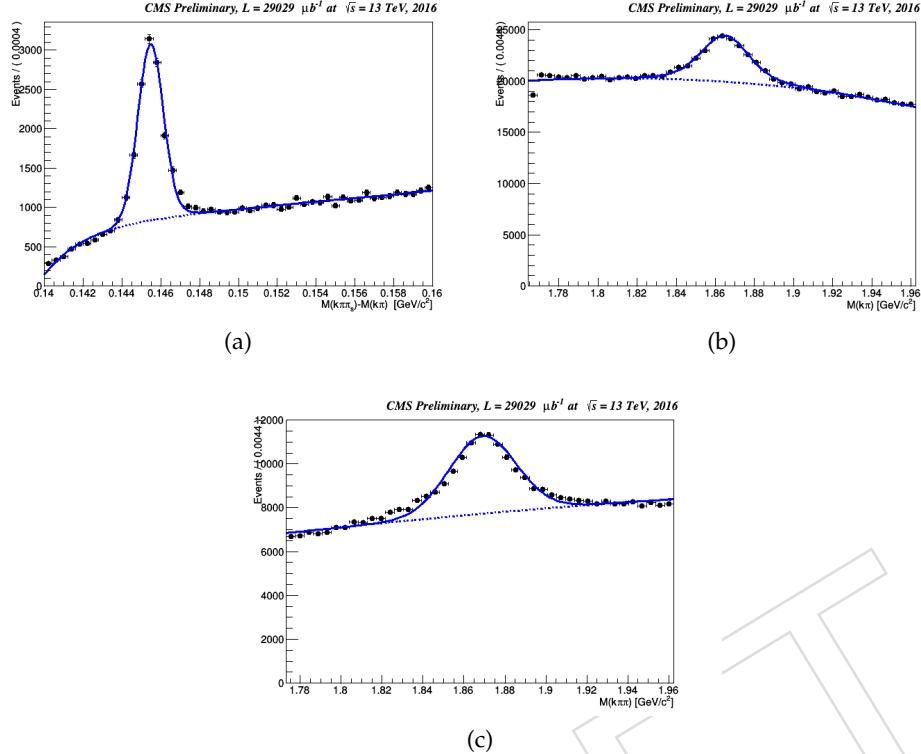


Figure 36: Invariant mass distribution with the signal peak described by a Crystal Ball function fit: D^{*+} (a), D^0 (b) and D^+ (c) mesons.

$$\Delta x_s = \sqrt{\Delta \tilde{x}^2 - \Delta \bar{x}^2} = 0.16 \quad (11)$$

The systematic uncertainty coming from this PU dependency study results to be 0.7%, so we can consider it negligible with respect to the other contributions.

Yield Instability

Studying the PU dependency as reported in the Appendix, has been figured out a yield instability during the different 2016 era of the data taking, as shown in Table 32 and Figure 47. Since this instability is related to the data taking conditions and inefficiencies and is not reproduced in the MC sample we have to take in account this effect as systematic. In particular we know that some inefficiencies affected the strip detector during the first half of data taking and it has been recovered only at the end of run F. This means that runG and runH can be considered as a reference, good quality runs. We followed the same S-method explained at the previous bullet to obtain the systematic uncertainty. Figure 39 shows the D^{*+} signal yields for each 2016 data taking era, normalised to the corresponding era luminosity and the red line represents the 0-degree polynomial fit.

The fit result gives $p_0 = \bar{x} \pm \Delta \bar{x} = 0.35 \pm 0.01$ with a $\chi^2/ndof = 14.78/6$. This brings to $\Delta x_s = 0.012$, that means a 3.5 % of systematic uncertainty.

All the systematic uncertainties are reported in Table 22 for each meson.

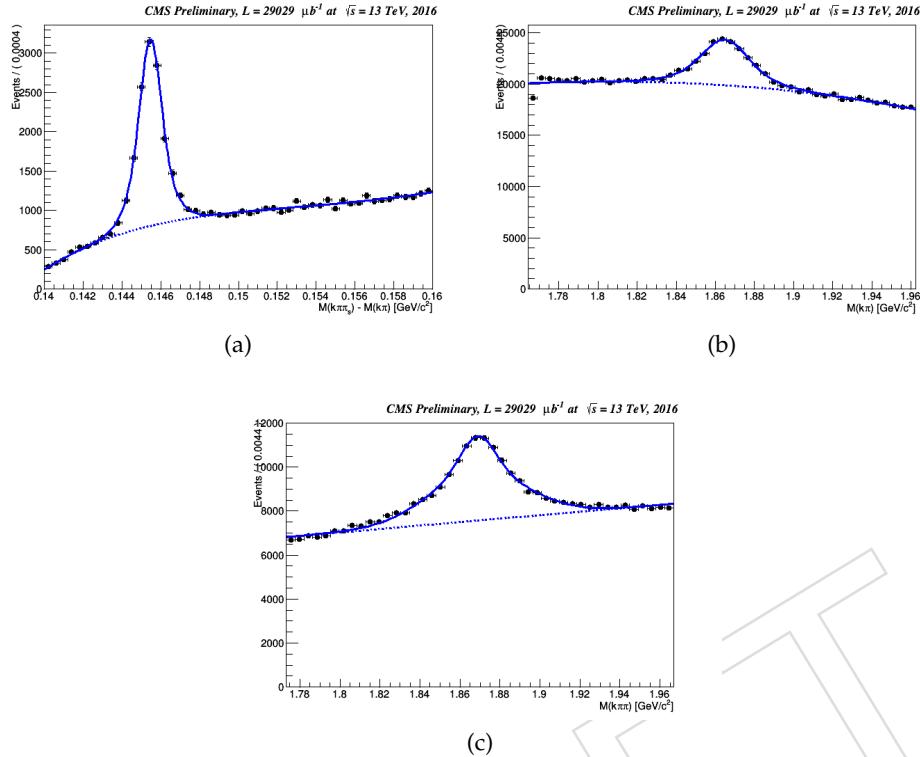


Figure 37: Invariant mass distribution with the combinatorial background described with four degrees polynomial function fit: D^{*+} (a), D^0 (b) and D^+ (c) mesons.

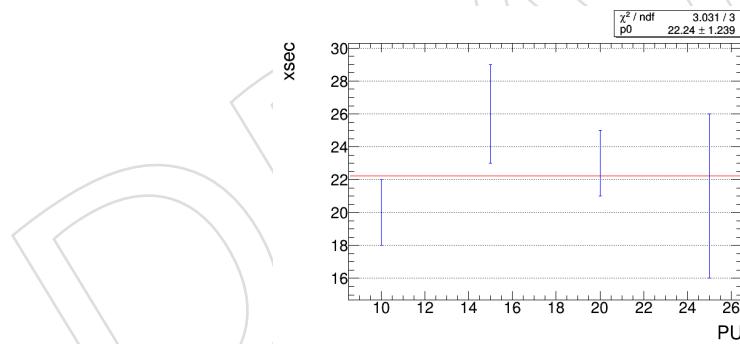


Figure 38: D^{*+} cross section measured for different PU values integrating in p_T and $|\eta|$; the red line shows the 0-degree polynomial function fit.

Systematic source	D^{*+}	D^0	D^+
Lumisection	2.5 %	2.5 %	2.5 %
Branching ratio	0.03 %	0.04 %	0.28 %
Tracking	9.4 %	4.2 %	6.1 %
Signal modeling	3.0 %	3.4 %	2.8 %
Background modeling	1.2 %	4.8 %	5.3 %
Efficiency	0.3 %	0.3 %	3.5 %
Contamination	2.9 %	0.8 %	1.4 %
Yield instability	3.5 %	3.5 %	3.5 %

Table 22: D^{*+} , D^0 and charged D^+ differential cross section in p_T bins with $|\eta| < 2.1$; the first uncertainty is statistics, the second is systematic.

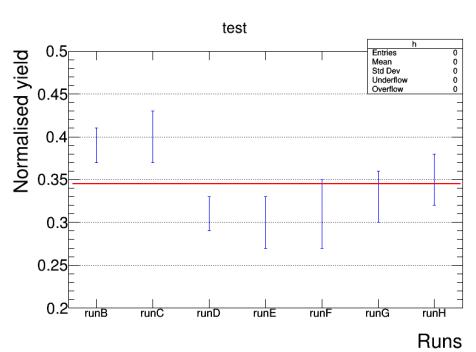


Figure 39: D^* signal yields for each 2016 data taking era normalised to the corresponding era luminosity; the red line shows the 0-degree polynomial function fit.

442 9 Results

443 The differential cross sections for the charmed mesons as a function of p_T and η have been de-
 444 fined as:

$$\frac{d\sigma(pp \rightarrow DX)}{dp_T} = \frac{N_i(D \rightarrow f + c.c)}{\Delta p_T \cdot \mathcal{BR}(D \rightarrow f) \cdot \varepsilon_{i,tot}(D \rightarrow f) \cdot k\mathcal{L}} \quad (12)$$

$$\frac{d\sigma(pp \rightarrow DX)}{d\eta} = \frac{N_i(D \rightarrow f + c.c)}{\Delta\eta \cdot \mathcal{BR}(D \rightarrow f) \cdot \varepsilon_{i,tot}(D \rightarrow f) \cdot k\mathcal{L}} \quad (13)$$

445 where $N_i(D \rightarrow f + c.c)$ is the signal yield of the charmed meson reconstructed in the selected
 446 final state for both the charge states in each bin i , Δp_T ($\Delta\eta$) is the p_T (η) bin width, $\mathcal{BR}(D \rightarrow$
 447 $f)$ is the branching ratio of the reconstructed final state, $\varepsilon_{i,tot}(D \rightarrow f)$ is the total efficiency
 448 reconstruction of the decay chain evaluated in MC sample and $k\mathcal{L}$ is the delivered luminosity
 449 corrected by the trigger prescale factor.

450 In Tables 23 and 24 the differential cross section values are reported for each p_T and η bin re-
 451 spectively. The central values are reported with their uncertainties: the first one is the statistical
 452 error, the second one the square root of the systematics.

p_T range [GeV/c]	$D^{*+} d\sigma/dp_T (\mu b)$	$D^0 d\sigma/dp_T (\mu b)$	$D^+ d\sigma/dp_T (\mu b)$
[4-5]	$127 \pm 6 \pm 14$	$332 \pm 45 \pm 28$	$215 \pm 52 \pm 24$
[5-6]	$75 \pm 10 \pm 8$	$181 \pm 13 \pm 15$	$71 \pm 10 \pm 8$
[6-7]	$37 \pm 3 \pm 4$	$110 \pm 9 \pm 9$	$44 \pm 2 \pm 5$
[7-8]	$21.2 \pm 1.4 \pm 2$	$55 \pm 7 \pm 5$	$24.8 \pm 1.2 \pm 2.7$
[8-12]	$7.8 \pm 0.2 \pm 0.9$	$18.7 \pm 0.6 \pm 1.6$	$8.1 \pm 0.2 \pm 0.8$
[12-16]	$1.64 \pm 0.16 \pm 0.19$	$3.78 \pm 0.18 \pm 0.3$	$1.91 \pm 0.06 \pm 0.21$
[16-24]	$0.352 \pm 0.015 \pm 0.042$	$0.81 \pm 0.05 \pm 0.07$	$0.360 \pm 0.015 \pm 0.04$
[24-40]	$0.043 \pm 0.002 \pm 0.006$	$0.098 \pm 0.009 \pm 0.009$	$0.057 \pm 0.006 \pm 0.008$
[40-100]	$0.0023 \pm 0.0004 \pm 0.0005$	$0.0036 \pm 0.0008 \pm 0.0004$	$0.0032 \pm 0.0004 \pm 0.0007$

Table 23: D^{*+} , D^0 and charged D^+ differential cross section in p_T bins with $|\eta| < 2.1$; the first uncertainty is statistics, the second is systematic.

$ \eta $ range	$D^{*+} d\sigma/d\eta (\mu b)$	$D^0 d\sigma/d\eta (\mu b)$	$D^+ d\sigma/d\eta (\mu b)$
[0-0.2]	91 ± 3 (stat) ± 9 (syst)	180 ± 6 (stat) ± 15 (syst)	62 ± 2 (stat) ± 7 (syst)
[0.2-0.4]	98 ± 9 (stat) ± 9 (syst)	205 ± 7 (stat) ± 18 (syst)	59 ± 2 (stat) ± 6 (syst)
[0.4-0.6]	92 ± 3 (stat) ± 9 (syst)	207 ± 7 (stat) ± 18 (syst)	53 ± 4 (stat) ± 6 (syst)
[0.6-0.8]	112 ± 7 (stat) ± 11 (syst)	219 ± 9 (stat) ± 19 (syst)	62 ± 4 (stat) ± 7 (syst)
[0.8-1]	102 ± 5 (stat) ± 10 (syst)	186 ± 11 (stat) ± 16 (syst)	69 ± 3 (stat) ± 7 (syst)
[1-1.2]	98 ± 7 (stat) ± 9 (syst)	239 ± 18 (stat) ± 20 (syst)	79 ± 4 (stat) ± 9 (syst)
[1.2-1.4]	100 ± 12 (stat) ± 10 (syst)	207 ± 20 (stat) ± 18 (syst)	96 ± 3 (stat) ± 11 (syst)
[1.4-1.6]	79 ± 23 (stat) ± 8 (syst)	224 ± 23 (stat) ± 20 (syst)	93 ± 10 (stat) ± 11 (syst)
[1.6-1.8]	82 ± 11 (stat) ± 8 (syst)	181 ± 42 (stat) ± 16 (syst)	90 ± 9 (stat) ± 10 (syst)
[1.8-2]	81 ± 6 (stat) ± 8 (syst)	256 ± 33 (stat) ± 23 (syst)	93 ± 18 (stat) ± 11 (syst)
[2-2.1]	36 ± 11 (stat) ± 4 (syst)	249 ± 55 (stat) ± 23 (syst)	111 ± 15 (stat) ± 16 (syst)

Table 24: D^{*+} , D^0 and charged D^+ differential cross section in η bins with $4 < p_T < 100$ GeV/c.

453 The differential cross sections for the three mesons as a function of p_T and η are shown in
 454 Figures 40, 41, 42: data points (black) are compared with several Monte Carlo and theoretical

455 predictions. The error bars include both the statistical and systematic uncertainties. The cross-
 456 section values are compared to the predictions of FONLL [11], shown as boxes representing
 457 the upper and lower limit for a given p_T bin, Pythia 6 [12], Pythia 8 (several tunes) [13] and
 458 Powheg [14] prediction. The bottom panel displays the ratio of data to FONLL, PYTHIA and
 459 Powheg predictions including both the statistical and systematic uncertainties.

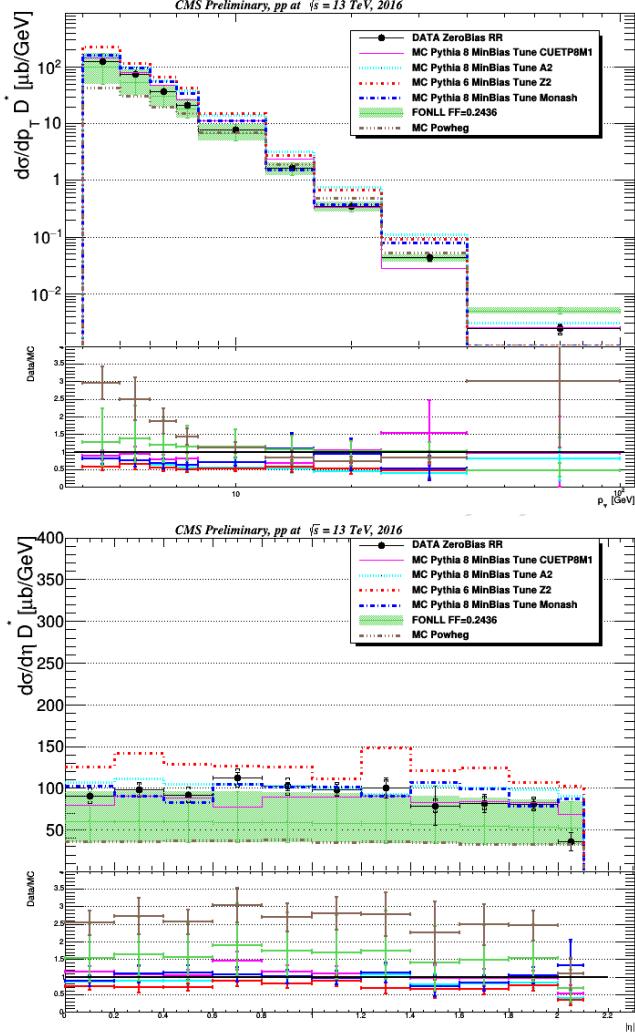


Figure 40: Differential cross section $d\sigma/dp_T$ ($d\sigma/d\eta$) on the left (right) for the D^{*+} meson. Data points (black) are compared with several Monte Carlo and theoretical predictions. On the bottom side the ratio between data and predictions is reported with the corresponding errors. Both the statistics and systematics uncertainties are reported.

460 The agreement with the different MC is fair. There is not a single MC, among those chosen
 461 for the comparison with our data, that describes well all the points. The central values of the
 462 FONLL predictions tend to underestimate data for D^{*+} and D^0 even if they agree within the
 463 large uncertainties. The various tunes of Pythia globally overestimate data; this is particularly
 464 evident for the D^+ mesons where all the tunes of Pythia 8 are systematically over data of about
 465 a factor 2.

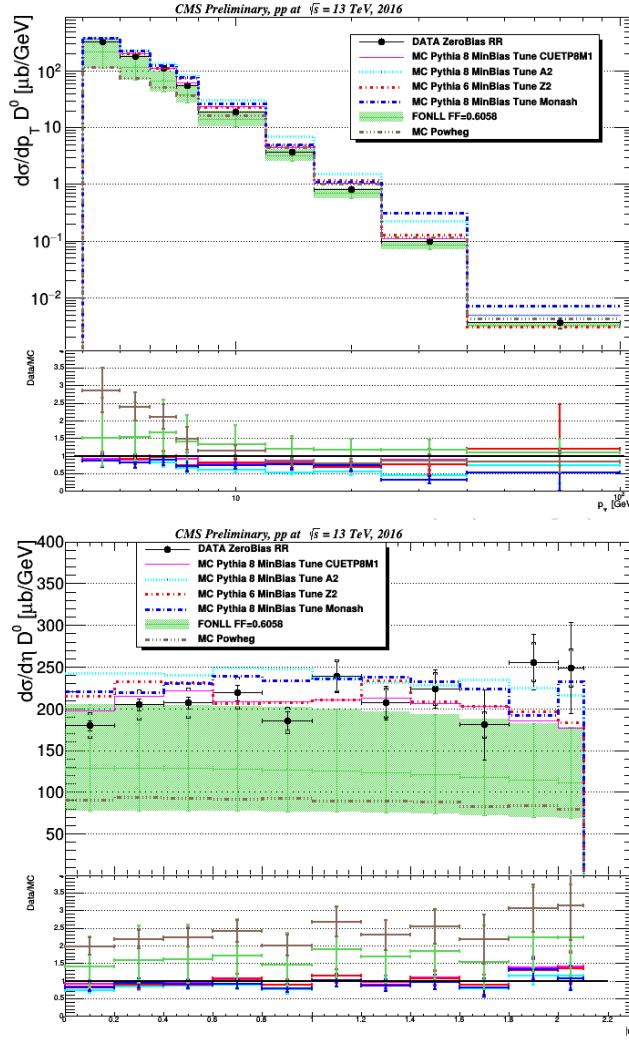


Figure 41: Differential cross section $d\sigma/dp_T$ ($d\sigma/d\eta$) on the left (right) for the D^0 meson. Data points (black) are compared with several Monte Carlo and theoretical predictions. On the bottom side the ratio between data and predictions is reported with the corresponding errors. Both the statistics and systematics uncertainties are reported.

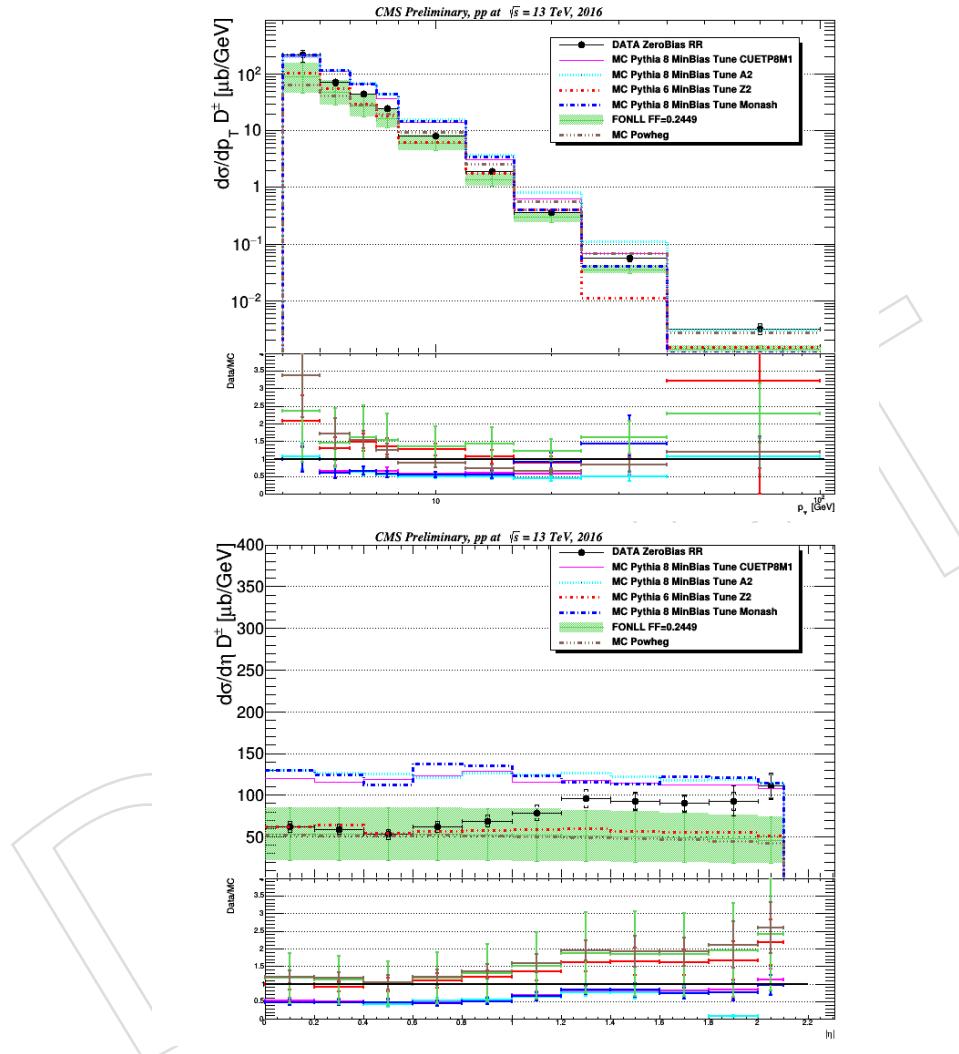


Figure 42: Differential cross section $d\sigma/dp_T$ ($d\sigma/d\eta$) on the left (right) for the charged D^+ meson. Data points (black) are compared with several Monte Carlo and theoretical predictions. On the bottom side the ratio between data and predictions is reported with the corresponding errors. Both the statistics and systematics uncertainties are reported.

466 9.1 Comparison with previous measurements from other experiments

467 As mentioned in the Introduction, all the others LHC Collaborations have performed such
 468 measurement, although at different center of mass energy values and kinematic regions, and it
 469 is interesting to compare the results. The measurement done from ATLAS [3], even if referred
 470 to $\sqrt{s} = 7$ TeV data, is the closest in term of acceptance and kinematic regime. They measured
 471 the production cross section of the $D^{*\pm}$ and D^\pm mesons as a function of p_T and η ; the results
 472 are shown in Figure 43. They also compared data with several Monte Carlo models with a
 473 not-perfect agreement, that becomes more evident in the D^\pm meson distributions although in
 474 this case the data points are in average above the MC prediction of a factor 2-3, included the
 475 FONLL predictions.

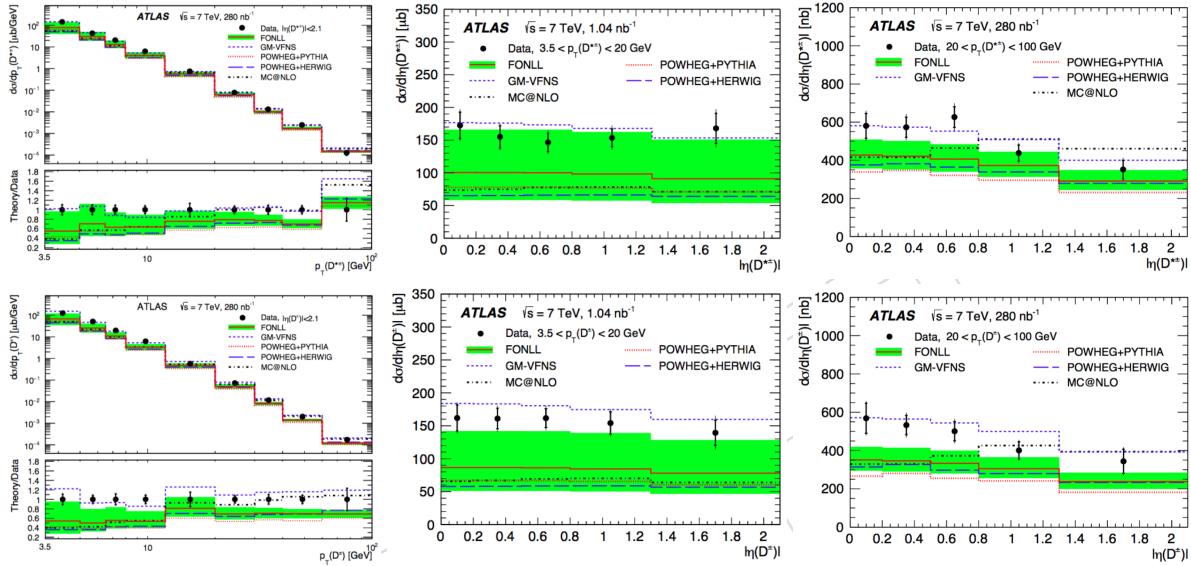


Figure 43: Differential cross section $d\sigma/dp_T$ ($d\sigma/d\eta$) on the left (right) for the D^* meson on the top and the D^\pm on the bottom from the Atlas Collaboration measurements at $\sqrt{s} = 7$ TeV [3].

476 We can roughly compare our results to the results obtained by the ATLAS Collaboration
 477 considering a common p_T bin, $8 < p_T < 12$ GeV/c, for a trivial and more direct comparison. We
 478 have to consider that in their analysis the contamination from secondary decay is not taken into
 479 account, so to make the two values as comparable as possible we have to use our estimations
 480 before the subtraction of the contamination rate. Since the center of mass energy is different for
 481 the two measurements, the values of the cross sections values are clearly not expected to be the
 482 same, but a scale factor, defined as the ratio between the cross sections measured at $\sqrt{s} = 7$ and
 483 13 TeV can be studied. The ratios for D^{*+} and D^+ mesons (the σ_{D^0} has not been measured in
 484 [3]) are reported in table 25; the factor is the same for both the mesons within the uncertainties.
 485 In the same table we also reported the FONLL predictions in the same p_T bin for 7 and 13 TeV
 486 and the respective ratio; the factor obtained by data and the one from FONLL agree within the
 487 large uncertainties.

488 An other measurement at $\sqrt{s} = 7$ TeV was performed by the Alice Collaboration [1], the results
 489 are shown in Figure 44 for D^{*+} , D^0 and D^+ mesons as a function of p_T . From the comparison
 490 between data and Monte Carlo it is clear that the FONLL predictions tend to underestimate
 491 data, reaching the agreement within the large uncertainties.

492 The only other measurement performed at $\sqrt{s} = 13$ TeV comes from the LHCb Collaboration
 493 [2]. Since the different acceptances their results can be considered complementary with respect
 494 to what has been obtained by us. In Figure 45 the production cross section shapes, as a function

	$\sigma_{D^{*+}} (\mu\text{b})$	$\sigma_{D^+} (\mu\text{b})$
CMS (13 TeV)	8.62 ± 0.23	9.19 ± 0.24
ATLAS (7 TeV)	6.34 ± 0.5	6.29 ± 0.56
Ratio 13/7 (CMS and ATLAS data)	1.36 ± 0.11	1.46 ± 0.14
FONLL (13 TeV)	$6.8^{+2.77}_{-1.74}$	$5.94^{+2.38}_{-1.51}$
FONLL (7 TeV)	$3.59^{+1.41}_{-0.92}$	$3.11^{+1.25}_{-0.84}$
Ratio 13/7 FONLL predictions	$1.89^{+1.07}_{-0.69}$	$1.91^{1.08}_{-0.71}$

Table 25: Scale factor between the current 13 TeV analysis and the 7 TeV analysis performed by ATLAS for D^{*+} and D^+ in the unique common bin $8 < p_T < 12 \text{ GeV}/c$ in $|\eta| < 2.1$ and the prediction and scale factor predicted by the FONLL calculations.

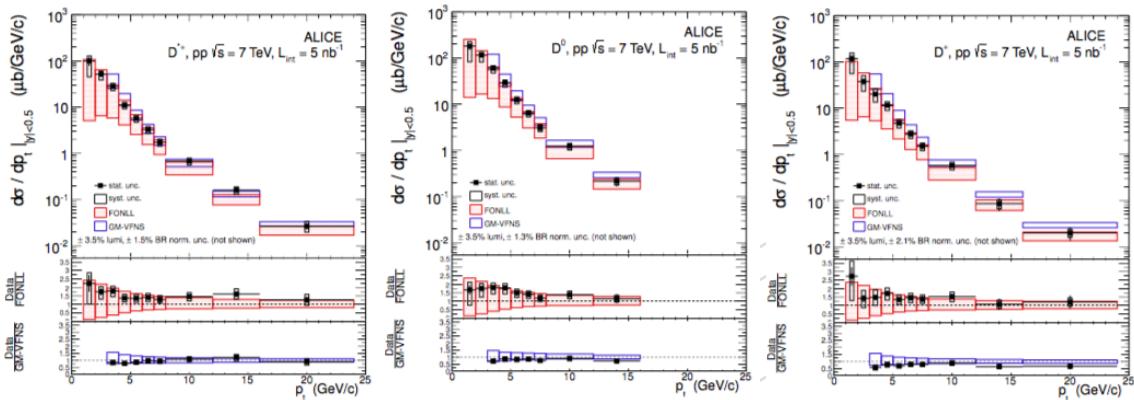


Figure 44: Differential cross section $d\sigma/dp_T$ for the D^{*+} (left), D^0 (center) and D^+ (right) mesons from the Alice Collaboration measurements at $\sqrt{s} = 7 \text{ TeV}$ [1].

of p_T for different rapidity ranges, obtained by LHCb is shown on the right, and the results obtained by us in the pseudorapidity range $|\eta| < 2.1$ are shown on the left (reported here only till $p_T < 16 \text{ GeV}/c$ in order to improve the direct comparison). The easiest way to interpretate these distributions is to consider the CMS shape as an additional row, on top of the others, in the LHCb plot, with $|y| < 2.5$, allowing for a complete coverage of the rapidity acceptance. They could perform a much more detailed scan in p_T and rapidity, reaching values at low p_T that are completely hidden by the combinatorial background in our case. From this rough comparison it has been demonstrated that the results obtained from the two measurements are coherent and in agreement although the different kinematic regions investigated by the experiments. We can conclude that the expected complementarity between the LHCb measurement and the CMS one at $\sqrt{s} = 13 \text{ TeV}$ is satisfied; the agreement of both the measurements to the FONLL predictions within the uncertainties confirms that.

10 Summary

We presented, for the first time in the CMS Collaboration, the measurement of the open charm production cross section using data collected during pp collisions in 2016 at $\sqrt{s} = 13 \text{ TeV}$. The differential cross sections $d\sigma/dp_T$ and $d\sigma/d\eta$ have been measured for the $D^{*\pm}$, D^0 (and \bar{D}^0) and D^\pm in range of $4 < pT < 100 \text{ GeV}/c$ and $|\eta| < 2.1$. A dedicated reconstruction strategy have developed in order to identify the charmed mesons in such a wide kinematic range with very good results. Furthermore a deep study on the contamination coming from no-prompt D mesons (i.e. decaying from B mesons) has been performed using a MC-based strategy. The

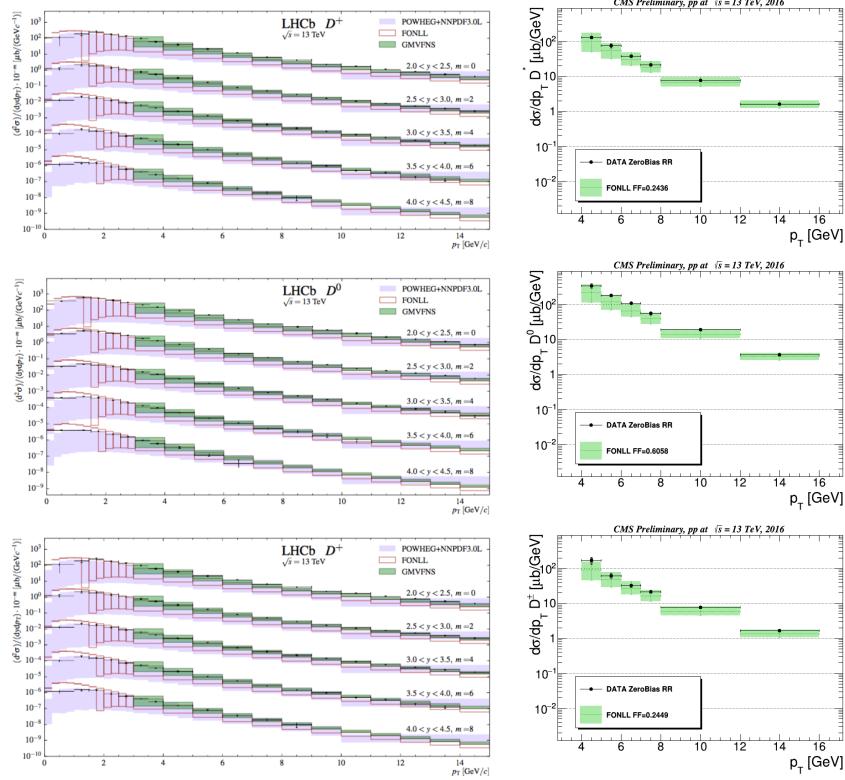


Figure 45: Differential cross section $d\sigma/dp_T$ for the D^{*+} (top), D^0 (middle) and D^+ (bottom) mesons at different rapidity ranges from the LHCb Collaboration measurements at $\sqrt{s} = 13$ TeV [2] on the left, and the respective cross section measurement computed in this analysis on the right in $|\eta| < 2.1$ for $p_T < 16$ GeV/c (only the comparison to FONLL has been reported to make more trivial the comparison).

515 measured cross section values have been compared to different Monte Carlo predictions. The
 516 agreement with the various models can be considered fair, since there is not a specific Monte
 517 Carlo that describes well all the data points in the whole range. This measurement gives an im-
 518 portant contribution in the understanding of the charm production in a hadronic environment,
 519 that is still affected by large uncertainties in the present theoretical models.

520 11 Appendix

521 11.1 ΔR stability check

522 In order to test the stability of the ΔR cut used for the gen-reco matching applied for the
 523 efficiency calculations.

524 The cut applied is $\Delta R < 0.4$, already quite tight looking at the left distribution of Figure 28;
 525 thus for this test two looser cut has been used: $\Delta R < 0.5$ and $\Delta R < 0.6$. The numbers are
 526 reported in Tables 26, 28, 30, 27, 29 and 31 for all the three reconstructed mesons in p_T and η
 527 bins.

p_T [GeV/c]	$\Delta R < 0.5$			$\Delta R < 0.6$		
	N_i	TOT_i	ε_i	N_i	TOT_i	ε_i
[4-5]	102030 ± 340	12716660	$(8.02 \pm 0.03) \times 10^{-3}$	102090 ± 340	12716660	$(8.03 \pm 0.03) \times 10^{-3}$
[5-6]	143540 ± 390	6212347	$(2.310 \pm 0.006) \times 10^{-2}$	143620 ± 388	6212347	$(2.312 \pm 0.006) \times 10^{-2}$
[6-7]	138060 ± 380	3236898	$(4.265 \pm 0.011) \times 10^{-2}$	138140 ± 378	3236898	$(4.268 \pm 0.011) \times 10^{-2}$
[7-8]	114220 ± 340	1792115	$(6.37 \pm 0.02) \times 10^{-2}$	114280 ± 341	1792115	$(6.38 \pm 0.02) \times 10^{-2}$
[8-12]	264730 ± 523	2352339	$(1.125 \pm 0.002) \times 10^{-1}$	264830 ± 520	2352339	$(1.126 \pm 0.002) \times 10^{-1}$
[12-16]	91711 ± 307	459560	$(1.996 \pm 0.006) \times 10^{-1}$	91744 ± 305	459560	$(1.996 \pm 0.006) \times 10^{-1}$
[16-24]	46878 ± 220	183230	$(2.558 \pm 0.010) \times 10^{-1}$	46892 ± 219	183230	$(2.559 \pm 0.010) \times 10^{-1}$
[24-40]	12332 ± 113	41488	$(2.97 \pm 0.02) \times 10^{-1}$	12338 ± 112	41488	$(2.97 \pm 0.02) \times 10^{-1}$
[40-100]	1734 ± 42	5698	$(3.04 \pm 0.06) \times 10^{-1}$	1739 ± 42	5698	$(3.05 \pm 0.06) \times 10^{-1}$

Table 26: Efficiency calculation in p_T bins for D^* mesons in the pseudorapidity range $|\eta| < 2.1$ with different matching conditions: $\Delta R < 0.5$ and $\Delta R < 0.6$.

$ \eta $ range	$\Delta R < 0.5$			$\Delta R < 0.6$		
	N_i	TOT_i	ε_i	N_i	TOT_i	ε_i
[0-0.2]	114990 ± 343	2608159	$(4.409 \pm 0.013) \times 10^{-2}$	115070 ± 343	2608159	$(4.412 \pm 0.013) \times 10^{-2}$
[0.2-0.4]	117240 ± 347	2613410	$(4.486 \pm 0.013) \times 10^{-2}$	117300 ± 347	2613410	$(4.488 \pm 0.013) \times 10^{-2}$
[0.4-0.6]	118700 ± 349	2625104	$(4.522 \pm 0.013) \times 10^{-2}$	118760 ± 349	2625104	$(4.524 \pm 0.013) \times 10^{-2}$
[0.6-0.8]	118190 ± 348	2630884	$(4.492 \pm 0.013) \times 10^{-2}$	118260 ± 348	2630884	$(4.495 \pm 0.013) \times 10^{-2}$
[0.8-1]	108020 ± 333	2627132	$(4.112 \pm 0.012) \times 10^{-2}$	108090 ± 333	2627132	$(4.114 \pm 0.012) \times 10^{-2}$
[1-1.2]	89749 ± 305	2614841	$(3.432 \pm 0.011) \times 10^{-2}$	89799 ± 305	2614841	$(3.434 \pm 0.011) \times 10^{-2}$
[1.2-1.4]	72355 ± 274	2589700	$(2.794 \pm 0.010) \times 10^{-2}$	72394 ± 277	2589700	$(2.795 \pm 0.010) \times 10^{-2}$
[1.4-1.6]	60928 ± 253	2551391	$(2.388 \pm 0.001) \times 10^{-2}$	60943 ± 253	2551391	$(2.389 \pm 0.001) \times 10^{-2}$
[1.6-1.8]	52694 ± 236	2503081	$(2.1050 \pm 0.0009) \times 10^{-2}$	52710 ± 236	2503081	$(2.1060 \pm 0.0009) \times 10^{-2}$
[1.8-2]	44713 ± 218	2441432	$(1.8310 \pm 0.0009) \times 10^{-2}$	44721 ± 218	2441432	$(1.8320 \pm 0.0009) \times 10^{-2}$
[2-2.1]	18422 ± 141	1195203	$(1.541 \pm 0.0011) \times 10^{-2}$	18424 ± 141	1195203	$(1.541 \pm 0.0011) \times 10^{-2}$

Table 27: Efficiency calculation in η bins for D^* mesons with $p_T > 4$ GeV/c with different matching conditions: $\Delta R < 0.5$ and $\Delta R < 0.6$.

528 11.2 Pile Up checks

529 In section 2 the details of the two MC samples used in this analysis have been showed. Both
 530 of the MC samples have been produced simulating the 2015 data taking conditions; the main
 531 difference with respect to the 2016 data taking is the average value of Pile Up. The PU distri-
 532 butions are reported in Figure 46 for the different 2016 runs and the enriched MC.

533 Although the shapes are different we are quite confident that our analysis is not sensible to the
 534 PU for several reasons:

p_T [GeV/c]	$\Delta R < 0.5$			$\Delta R < 0.6$		
	N_i	TOT_i	ε_i	N_i	TOT_i	ε_i
[4-5]	137470 ± 408	10884790	$(1.263 \pm 0.003) \times 10^{-2}$	137660 ± 378	10884790	$(1.265 \pm 0.003) \times 10^{-2}$
[5-6]	126290 ± 380	5182344	$(2.437 \pm 0.007) \times 10^{-2}$	126760 ± 386	5182344	$(2.446 \pm 0.007) \times 10^{-2}$
[6-7]	107380 ± 240	2653959	$(4.046 \pm 0.012) \times 10^{-2}$	107390 ± 353	2653959	$(4.046 \pm 0.012) \times 10^{-2}$
[7-8]	86555 ± 312	1450841	$(5.97 \pm 0.02) \times 10^{-2}$	86576 ± 177	1450841	$(5.97 \pm 0.02) \times 10^{-2}$
[8-12]	193070 ± 540	1876943	$(1.029 \pm 0.002) \times 10^{-1}$	194710 ± 410	1876943	$(1.037 \pm 0.002) \times 10^{-1}$
[12-16]	68010 ± 281	362693	$(1.875 \pm 0.007) \times 10^{-1}$	68089 ± 284	362693	$(1.877 \pm 0.007) \times 10^{-1}$
[16-24]	36900 ± 210	143316	$(2.575 \pm 0.012) \times 10^{-1}$	36902 ± 205	143316	$(2.575 \pm 0.012) \times 10^{-1}$
[24-40]	10484 ± 117	32373	$(3.24 \pm 0.03) \times 10^{-1}$	10513 ± 110	32373	$(3.25 \pm 0.03) \times 10^{-1}$
[40-100]	1605 ± 45	4317	$(3.72 \pm 0.07) \times 10^{-1}$	1607 ± 45	4317	$(3.72 \pm 0.07) \times 10^{-1}$

Table 28: Efficiency calculation in p_T bins for D^0 mesons in the pseudorapidity range $|\eta| < 2.1$ with different matching conditions: $\Delta R < 0.5$ and $\Delta R < 0.6$.

$ \eta $ range	$\Delta R < 0.5$			$\Delta R < 0.6$		
	N_i	TOT_i	ε_i	N_i	TOT_i	ε_i
[0-0.2]	91455 ± 167	2200284	$(4.157 \pm 0.013) \times 10^{-2}$	91621 ± 307	2200284	$(4.164 \pm 0.013) \times 10^{-2}$
[0.2-0.4]	92582 ± 162	2203023	$(4.202 \pm 0.014) \times 10^{-2}$	92615 ± 308	2203023	$(4.204 \pm 0.014) \times 10^{-2}$
[0.4-0.6]	93699 ± 311	2208043	$(4.244 \pm 0.014) \times 10^{-2}$	94056 ± 311	2208043	$(4.260 \pm 0.014) \times 10^{-2}$
[0.6-0.8]	93853 ± 312	2208180	$(4.250 \pm 0.014) \times 10^{-2}$	93884 ± 313	2208180	$(4.252 \pm 0.014) \times 10^{-2}$
[0.8-1]	87923 ± 250	2201363	$(3.994 \pm 0.013) \times 10^{-2}$	88311 ± 301	2201363	$(4.012 \pm 0.013) \times 10^{-2}$
[1-1.2]	78445 ± 303	2188035	$(3.585 \pm 0.013) \times 10^{-2}$	78848 ± 312	2188035	$(3.587 \pm 0.013) \times 10^{-2}$
[1.2-1.4]	66785 ± 288	2161141	$(3.090 \pm 0.012) \times 10^{-2}$	67136 ± 282	2161141	$(3.107 \pm 0.012) \times 10^{-2}$
[1.4-1.6]	56611 ± 299	2125066	$(2.664 \pm 0.011) \times 10^{-2}$	56679 ± 263	2125066	$(2.667 \pm 0.011) \times 10^{-2}$
[1.6-1.8]	49176 ± 348	2080986	$(2.363 \pm 0.011) \times 10^{-2}$	49190 ± 248	2080986	$(2.364 \pm 0.011) \times 10^{-2}$
[1.8-2]	41754 ± 239	2024836	$(2.062 \pm 0.010) \times 10^{-2}$	42107 ± 233	2024836	$(2.080 \pm 0.010) \times 10^{-2}$
[2-2.1]	17314 ± 157	990618	$(1.748 \pm 0.013) \times 10^{-2}$	17498 ± 193	990618	$(1.766 \pm 0.013) \times 10^{-2}$

Table 29: Efficiency calculation in η bins for D^0 mesons with $p_T > 4$ GeV/c with different matching conditions: $\Delta R < 0.5$ and $\Delta R < 0.6$.

p_T [GeV/c]	$\Delta R < 0.5$			$\Delta R < 0.6$		
	N_i	TOT_i	ε_i	N_i	TOT_i	ε_i
[4-5]	599 ± 26	63858	$(9.4 \pm 0.4) \times 10^{-3}$	601 ± 26	63858	$(9.4 \pm 0.4) \times 10^{-3}$
[5-6]	655 ± 16	39977	$(1.64 \pm 0.06) \times 10^{-2}$	703 ± 24	39977	$(1.76 \pm 0.07) \times 10^{-2}$
[6-7]	727 ± 28	25496	$(2.85 \pm 0.10) \times 10^{-2}$	724 ± 34	25496	$(2.84 \pm 0.10) \times 10^{-2}$
[7-8]	670 ± 29	16054	$(4.17 \pm 0.16) \times 10^{-2}$	669 ± 26	16054	$(4.17 \pm 0.16) \times 10^{-2}$
[8-12]	1869 ± 44	25705	$(7.27 \pm 0.16) \times 10^{-2}$	1871 ± 46	25705	$(7.28 \pm 0.16) \times 10^{-2}$
[12-16]	805 ± 29	5983	$(1.34 \pm 0.04) \times 10^{-1}$	803 ± 29	5983	$(1.34 \pm 0.04) \times 10^{-1}$
[16-24]	551 ± 24	2680	$(2.06 \pm 0.08) \times 10^{-1}$	546 ± 24	2680	$(2.04 \pm 0.08) \times 10^{-1}$
[24-40]	154 ± 7	629	$(2.45 \pm 0.17) \times 10^{-1}$	160 ± 11	629	$(2.54 \pm 0.17) \times 10^{-1}$
[40-100]	22 ± 2	99	$(2.2 \pm 0.4) \times 10^{-1}$	22 ± 3	99	$(2.2 \pm 0.4) \times 10^{-1}$

Table 30: Efficiency calculation in p_T bins for D mesons in the pseudorapidity range $|\eta| < 2.1$ with different matching conditions: $\Delta R < 0.5$ and $\Delta R < 0.6$.

- 535 • we applied a very tight request on the direction of the charmed mesons (from the
 536 PV to the SV) as explained in Section 4;
 537 • we used the same MC sample for the tracking efficiency measurement [10] (that
 538 involves the D^* reconstruction through the same decay considered in this analysis),

$ \eta $ range	$\Delta R < 0.5$			$\Delta R < 0.6$		
	N_i	TOT_i	ε_i	N_i	TOT_i	ε_i
[0-0.2]	826 ± 29	17964	$(4.60 \pm 0.16) \times 10^{-2}$	830 ± 24	17964	$(4.62 \pm 0.16) \times 10^{-2}$
[0.2-0.4]	810 ± 28	17744	$(4.57 \pm 0.16) \times 10^{-2}$	809 ± 29	17744	$(4.56 \pm 0.16) \times 10^{-2}$
[0.4-0.6]	840 ± 29	17770	$(4.73 \pm 0.16) \times 10^{-2}$	838 ± 23	17770	$(4.72 \pm 0.16) \times 10^{-2}$
[0.6-0.8]	786 ± 29	17890	$(4.39 \pm 0.15) \times 10^{-2}$	786 ± 28	17890	$(4.39 \pm 0.15) \times 10^{-2}$
[0.8-1]	678 ± 26	17585	$(3.86 \pm 0.15) \times 10^{-2}$	676 ± 29	17585	$(3.85 \pm 0.15) \times 10^{-2}$
[1-1.2]	598 ± 25	17692	$(3.38 \pm 0.14) \times 10^{-2}$	599 ± 41	17692	$(3.39 \pm 0.14) \times 10^{-2}$
[1.2-1.4]	445 ± 27	17202	$(2.59 \pm 0.12) \times 10^{-2}$	460 ± 22	17202	$(2.67 \pm 0.12) \times 10^{-2}$
[1.4-1.6]	358 ± 15	16825	$(2.13 \pm 0.11) \times 10^{-2}$	367 ± 17	16825	$(2.18 \pm 0.11) \times 10^{-2}$
[1.6-1.8]	329 ± 8	16432	$(2.00 \pm 0.11) \times 10^{-2}$	334 ± 7	16432	$(2.03 \pm 0.11) \times 10^{-2}$
[1.8-2]	240 ± 7	15687	$(1.53 \pm 0.10) \times 10^{-2}$	250 ± 9	15687	$(1.59 \pm 0.10) \times 10^{-2}$
[2-2.1]	109 ± 12	7698	$(1.41 \pm 0.14) \times 10^{-2}$	116 ± 12	7698	$(1.51 \pm 0.14) \times 10^{-2}$

Table 31: Efficiency calculation in η bins for D mesons with $p_T > 4$ GeV/c with different matching conditions: $\Delta R < 0.5$ and $\Delta R < 0.6$.

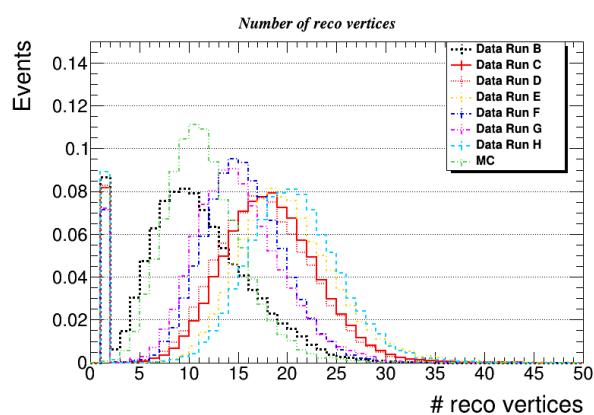


Figure 46: Distributions of reconstructed vertex in data, for the different runs, and MC.

comparing MC to both 2015 and 2016 data with no evident differences in terms of performances. Furthermore we saw that the efficiency calculated with two different MC configuration (2015 and 2017) seems to be not dependent on PU. In particular $\varepsilon_{2015} = 0.013734 \pm 0.000002$ and $\varepsilon_{2017} = 0.01836 \pm 0.00004$; the efficiency seems not decreasing with the considerable increasing of PU but the two values are expected to be different taking into account the additional layer in the 2017 configuration that improves the reconstruction performances.

Several tests have been performed in order to demonstrate such statements.

A first test has been performed to figure out any possible dependence of our reconstruction strategy on the PU. We quoted the signal yield for the D^* reconstruction in each different data taking period (from Run B to Run F) and we normalised all the values to the relative integrated lumisection; the numbers are reported in Table 32.

Even if the values are not all in perfect agreement, the difference doesn't seem to be related to the difference on the PU conditions. In fact a linear correlation can't be trivially figured out, also because the first value, corresponding to the minimum $\langle PU \rangle$, is in agreement within the uncertainty with the last one that corresponds to the maximum $\langle PU \rangle$. To make it more easily visible in Figure 47 all the values reported in the table are plotted normalised to the first

Run	$\langle PU \rangle$	Integrated lumi [μb^{-1}]	Signal Yield	Yield norm to lumi
Run B	10	12275 ± 307	4733 ± 168	0.39 ± 0.02
Run C	17	1267 ± 32	510 ± 39	0.40 ± 0.03
Run D	16	1899 ± 48	594 ± 41	0.31 ± 0.02
Run E	18	3011 ± 75	905 ± 88	0.30 ± 0.03
Run F	14	1720 ± 43	538 ± 66	0.31 ± 0.04
Run G	14	3166 ± 79	1050 ± 73	0.33 ± 0.03
Run H	19	5701 ± 143	1970 ± 188	0.35 ± 0.03

Table 32: D^* signal yield normalised to the respective luminosity for each 2016 data taking run taken at different PU conditions.

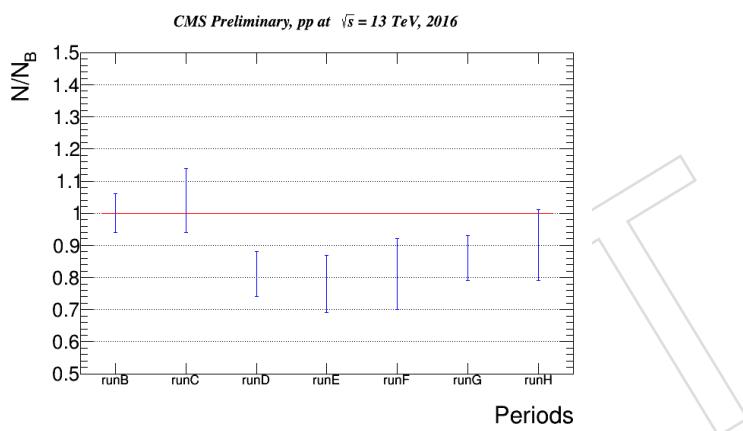


Figure 47: The values of table 32 are reported normalised to the first one.

556 one.

557 The second more accurate test has been performed considering different values of PU (chosen
 558 according to the available statistics) in both data and MC samples and then evaluating the cross
 559 section integrating in p_T and $|\eta|$. The measurements are reported in the Table 33 and in Figure
 560 48, where all the values are normalised to the previous one to make the results more clearly
 561 visible. Even if there is not a perfect agreement between all the values, a trend can't be trivially
 562 observed and we can conclude that there is not an evident dependence on PU.

$\langle PU \rangle$ value	MC effi	Data Signal Yield	xsec [$\mu barn$]
10	0.0341 ± 0.0001	507 ± 54	20 ± 2
15	0.0383 ± 0.0001	752 ± 97	26 ± 3
20	0.0321 ± 0.0004	570 ± 45	23 ± 2
25	0.0165 ± 0.0013	261 ± 68	21 ± 5

Table 33: D^* cross section values as a function of different PU values.

563 11.3 Minimum Bias and enriched MC comparison

564 For this analysis two different MC samples have been used:

- 565 • the enriched one for the efficiency calculation, with a quite high statistics in order to
 566 minimise the statiscal error on that - first row in Table 3
- 567 • the Minimum Bias one for the contamination evaluation in order to have the best

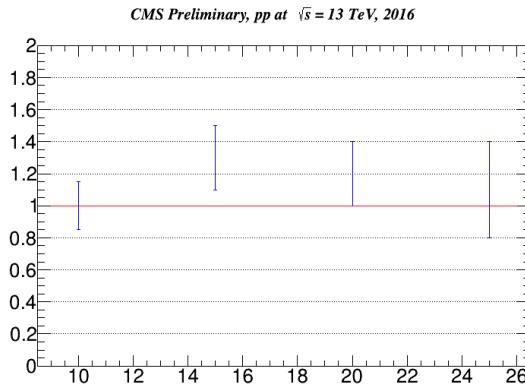


Figure 48: The values of table 33 are reported normalised to the first one.

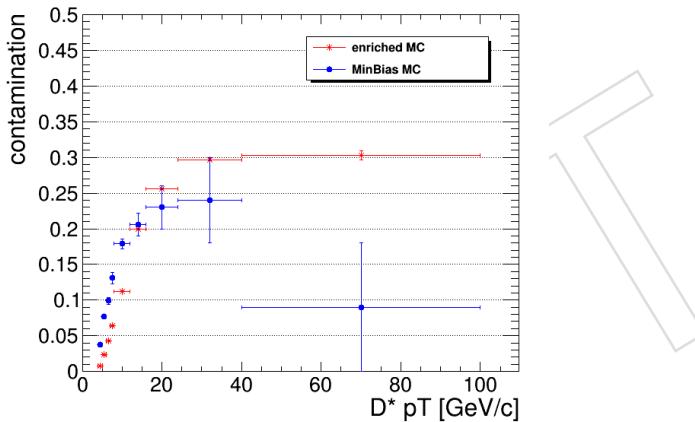


Figure 49: The D^* reconstruction efficiency reported as a function of p_T for the enriched (red) and the Minimum Bias (blue) MC samples.

568 simulation of data, without any filters in generation to avoid any bias in the production - second row in Table 3

570 These two samples, that are used in the analysis for two independent goals, differ mainly for
571 the presence of PU. The enriched MC has a 2015 PU scenario, while the MB one is without
572 PU. In the previous section we already demonstrated that the different PU scenario between
573 the enriched sample and the 2016 data doesn't affect this measurement.

574 Here we want to compare the two MC sample features. A deep study started comparing the
575 reconstruction efficiency in the two MC. In principle we expected compatible values but this
576 was not the case; in Figure 49 the D^* reconstruction efficiency is reported as a function of p_T
577 for the enriched MC and the Minimum Bias MC samples. The difference between the points is
578 not negligible, in any of the p_T bins analysed.

579 This was the starting point of a long investigation with the aim to understand if we really could
580 rely on the samples we were using.

581 As first step we studied the gen-level kinematic distributions, reported in Figure 50 for the
582 three mesons.

583 While the η and ϕ distributions appear in a quite good agreement, the p_T shapes show some
584 inconsistencies, in particular at lower values. This could in principle explain the differences

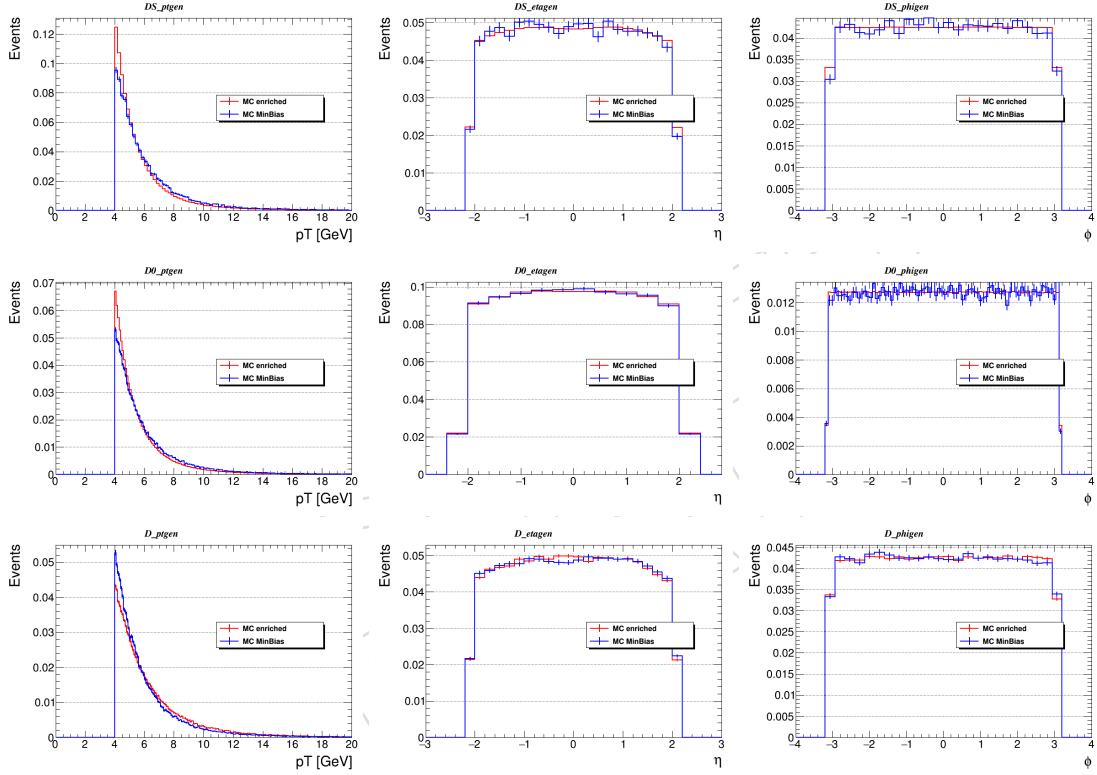


Figure 50: The kinematic variables (p_T , η and ϕ from the left to the right) are shown for the three mesons D^* (top), D^0 (center) and D^+ (bottom) at gen level comparing the enriched MC (red) to the Minimum Bias (blue) MC samples.

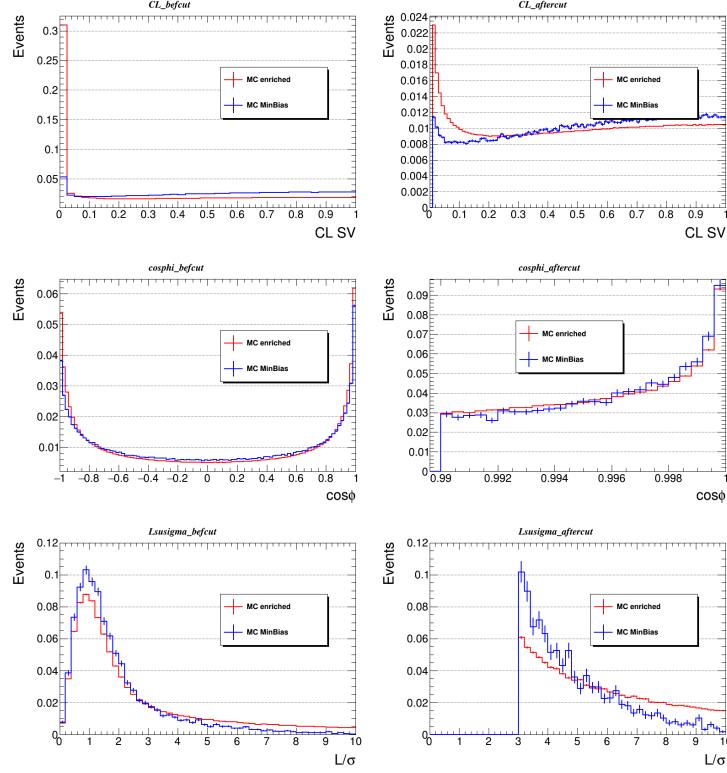


Figure 51: Three key variables for the D^* meson selection are shown: the CL of the SV (top), the $\cos\phi$ (center) and the L/σ (bottom). A comparison between the two MC samples is shown before (left) and just after (right) the application of the cut on the variable. N.B. these are the final D^* candidates.

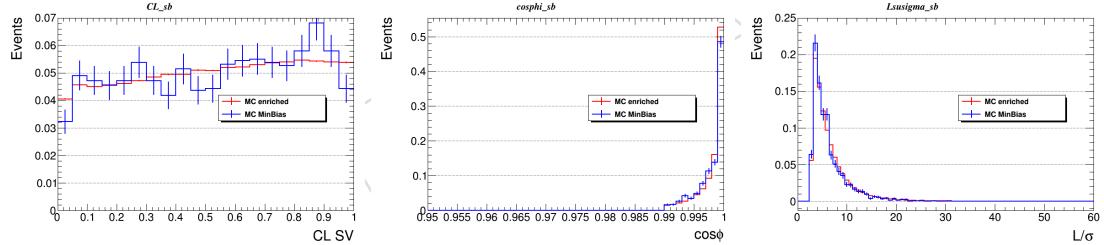


Figure 52: The same key variables for the D^* meson fully reconstructed are shown: the CL of the SV (left), the $\cos\phi$ (center) and the L/σ (right) comparing the two MC samples.

seen in the efficiency calculation, but the reason of the discrepancies is not clear anyway also because the generation tune of the two samples is the same.

Moving to the reconstructed quantities, in Figure 51 some key variables are shown for the D^* meson comparing again the two samples, the distributions on the left report the variables before the cuts while on the right the cuts are applied in order, as discussed in Chapter 4 and reported in Table 7. Many differences in the shapes are visible, both before and after the cuts, even if the statistic of the two samples is not comparable. It's important to underline that the distributions reported are not referring to the final D^* candidates.

The same distributions are then reported for the final reconstructed D^* candidates in Figure 52. After the whole selection strategy is applied, the discrepancies between the two MC samples are much more reduced.

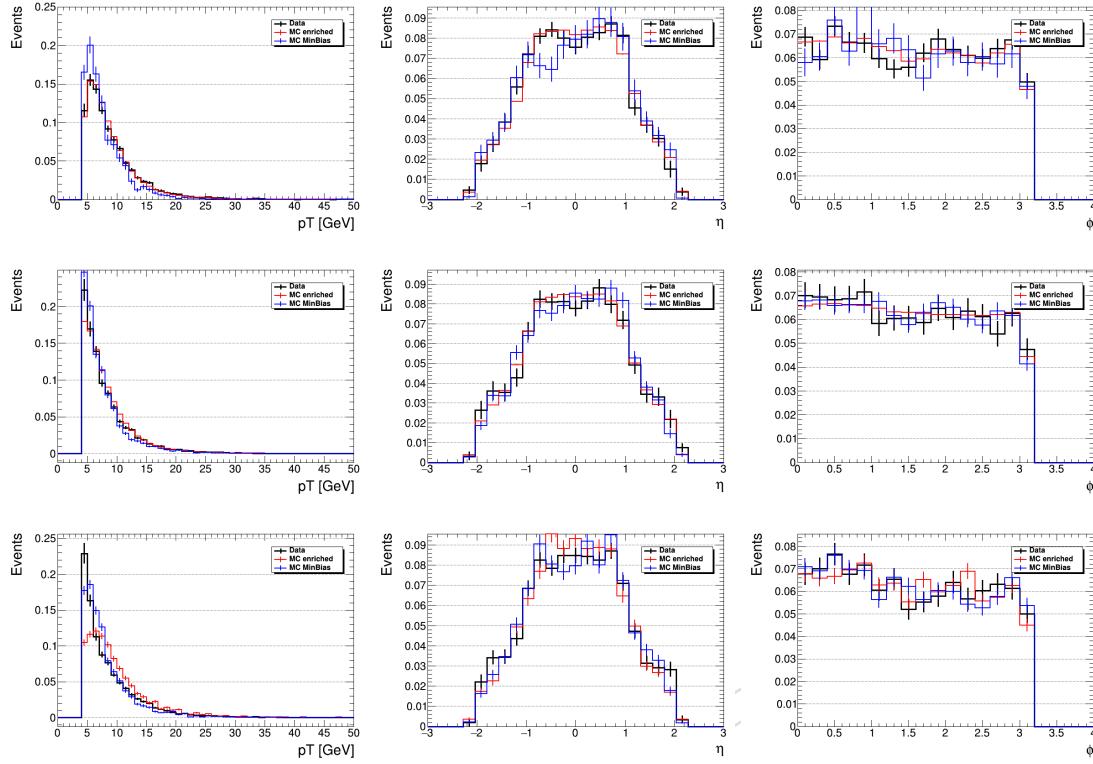


Figure 53: The kinematic variables (p_T , η and ϕ from the left to the right) are shown for the three mesons D^* (top), D^0 (center) and D^+ (bottom), comparing data (black) to the enriched MC (red) and to the Minimum Bias (blue) MC samples.

596 To confirm that the selected physics objects are coherent between the two samples, in Figure
 597 54 the kinematic variables are shown for the three mesons comparing data, the enriched MC
 598 and the Minimum Bias MC: a quite good agreement is shown, in according to what has been
 599 already shown in the previous chapters.

600 In Figure ?? the L/σ and $\cos\phi$ distributions are shown, again for the three mesons, comparing
 601 data to the two MC samples. Even in this case, the same agreement shown previously in the
 602 analysis note is confirmed here.

603 At this point the first considerations have been drawn:

- 604 • The efficiency values obtained in the MinBias MC we used for the contamination
 605 evaluation are different from the ones used in the analysis for the xsec calculation,
 606 obtained from the enriched MC sample. We tend to trust the enriched sample be-
 607 cause it provides a very good description of the charmed mesons kinematic and
 608 event topology. The reliability of such sample has been already proved and stressed
 609 in [10]
- 610 • We checked the key variables of the event selection strategy and we see that before
 611 and after the cuts application they show different shapes between the two samples,
 612 but once we apply the complete strategy for the candidates selection, the shapes are
 613 in a good agreement.
- 614 • We tend to address the differences to the presence/absence of PU that is the main
 615 difference between the two samples.

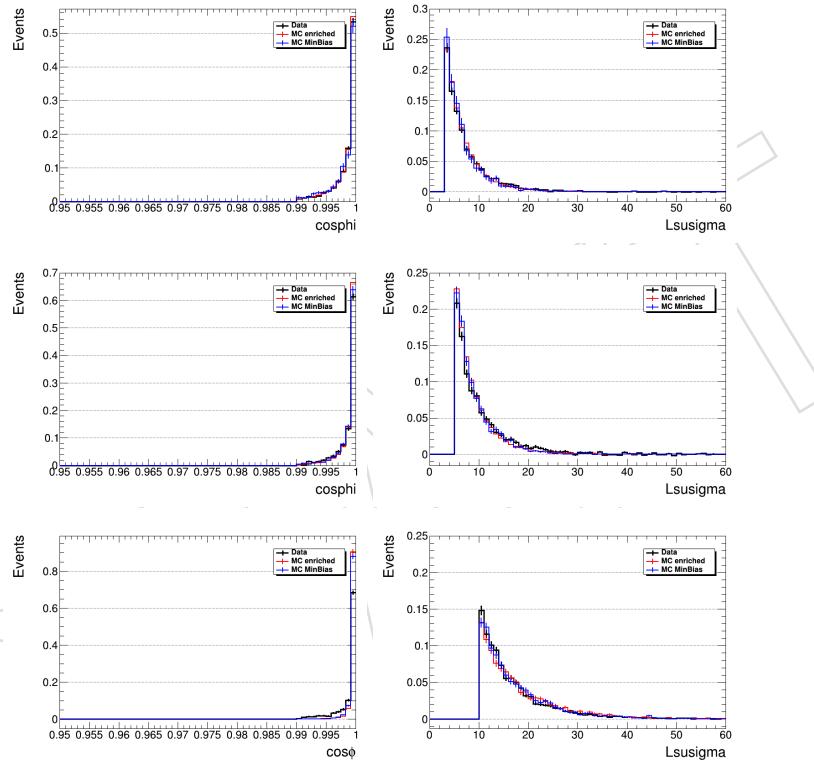


Figure 54: The $\cos\phi$ (left) and L/σ (right) variables are shown for the three mesons D^* (top), D^0 (center) and D^+ (bottom), comparing data (black) to the enriched MC (red) and to the Minimum Bias (blue) MC samples.

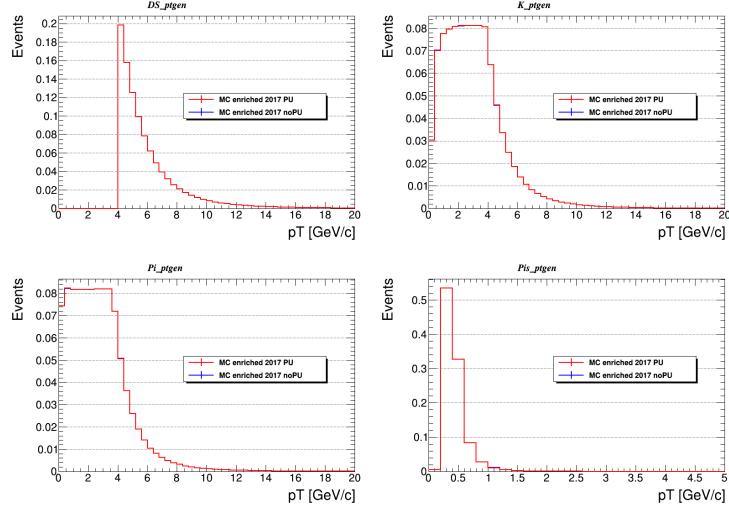


Figure 55: The p_T distributions at generator level are shown for all the objects in the D^* final state, comparing the 2017 MC sample with PU (red) and without PU (blue).

616 The real point is thus whether we could or not trust the contamination values we got from the
617 Minimum Bias sample that is simulated without PU.

618 In order to understand if PU is really the source of such discrepancies, and to quantify the
619 impact on the analysis, we used two MC samples enriched in D^* produced with the 2017 con-
620 figuration:

- 621 • /DStarToD0PiD0KPiDStarFilterTuneCP513TeV-pythia8-evtgen/RunIIFall17DRPremix-
622 94Xmc2017realisticv10-v1/AODSIM
- 623 • /DStarToD0PiD0KPiDStarFilterTuneCP513TeV-pythia8-evtgen/RunIIFall17DRStdmix-
624 NoPU94Xmc2017realisticv10-v1/AODSIM

625 As it is clear from the names, the two samples have identical configuration at gen level and
626 identical reconstruction conditions, except for the PU: the former is with 2017 PU scenario, the
627 second one is without PU.

628 The following studies have been conducted only for the D^* meson, for brevity.

629 Repeating what has been previously shown, in Figure 56 the p_T distributions for the particles
630 in the D^* decay mode are reported at generator level comparing the two MC with the 2017
631 samples: in red the one with PU and in blue the one without PU.

632 As expected, the distribution are perfectly superimposed since the generator process is exactly
633 the same. In this way we are confident that at least the denominator in the efficiency calculation
634 is coherent between these two samples and we minimised the sources of discrepancies basically
635 only to the presence/absence of PU.

636 Moving to the reconstructed quantities, the same distributions of the p_T spectra are reported
637 in Figure ??.

638 The shapes are not in a perfect agreement, with the spectra of the PU sample that seems to be
639 harder than the ones in the noPU sample, similarly to what has been shown in 54.

640 We proceeded then with the efficiency calculation, that is reported in Figure 57.

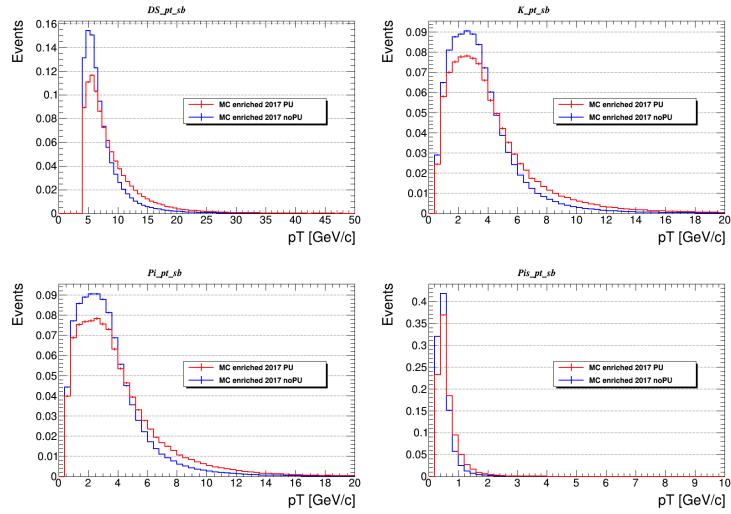


Figure 56: The p_T distributions are shown for all the reconstructed objects in the D^* final state, comparing the 2017 MC sample with PU (red) and without PU (blue).

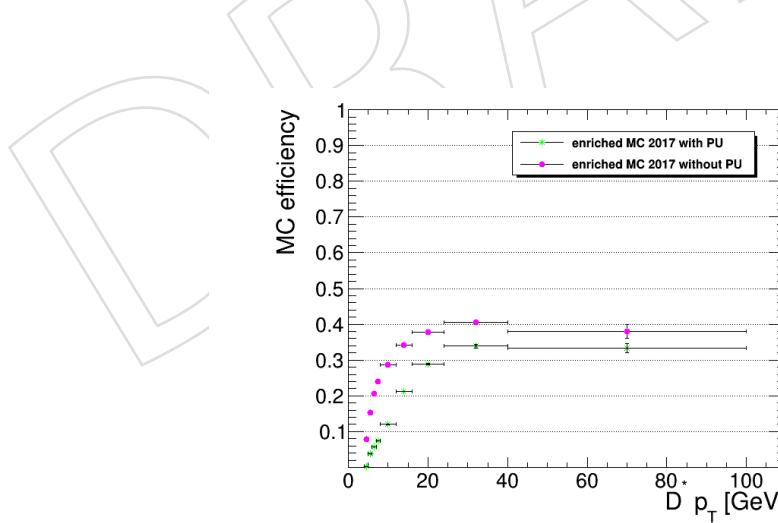


Figure 57: The D^* reconstruction efficiency reported as a function of p_T for the two enriched samples with the 2017 configuration: in green the points of the sample with PU and in magenta the points referring to the noPU sample.

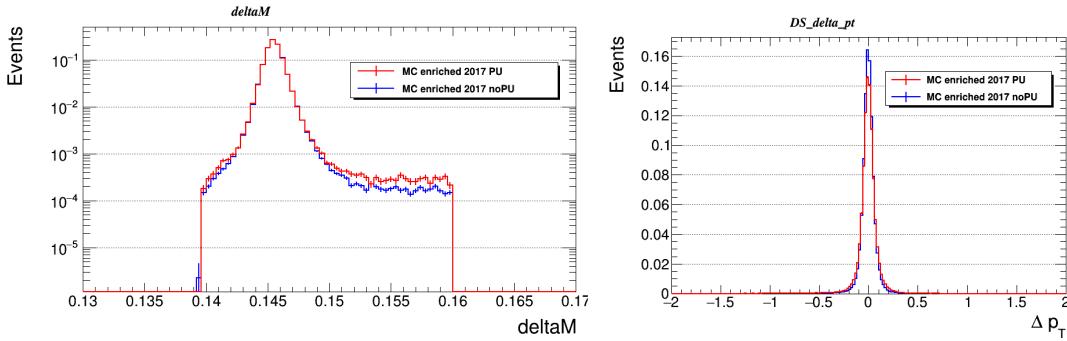


Figure 58: The ΔM (left) and the D^* Δp_T are shown with the comparison between the PU sample (red) and the noPU sample (blue).

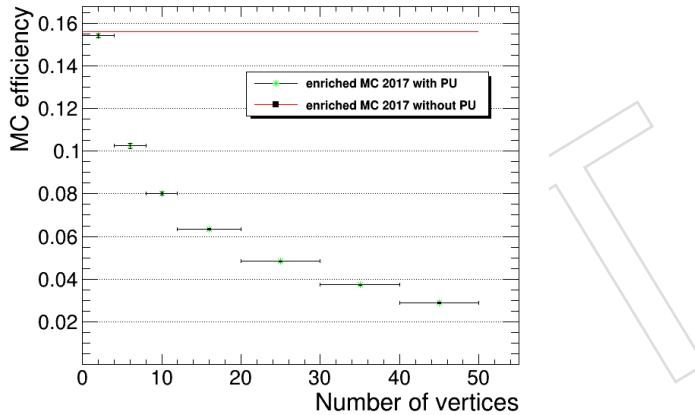


Figure 59: The D^* reconstruction efficiency reported as a function of the number of vertices: the points refer to the PU sample, while the red line report the efficiency for the noPU sample.

641 A very similar behaviour with respect to what has been shown in 49 is also reported here, with
642 the efficiency in the sample without PU greater than the one calculated in the PU sample.

643 Figure 58 shows the ΔM and the Δp_T ($|p_T^{gen_{D^*}} - p_T^{reco_{D^*}}|$) distributions that show, again, that the
644 physics objects are correctly selected.

645 The differences between the two samples are very small and can't justify the big discrepancies
646 in the efficiency calculation.

647 The key point is shown in Figure 59, where the efficiency is reported as a function of the number
648 of vertices.

649 The two efficiency values are in a very good agreement in the first bin, where the PU values
650 are comparable to each other ($0 < PU < 4$), then the efficiency drops down with the increasing
651 of the number of vertices in the event. This confirms the original hypothesis: the source of
652 the efficiency discrepancy lies on the PU presence. In fact, taking only the first PU bin and
653 comparing the efficiency as a function of p_T , as shown in Figure 60, a much better agreement is
654 achieved.

655 So once the reason of the differences in the efficiency calculation has been clearly addressed to
656 the presence/absence of PU in the sample what was still needed to understand is whether this
657 affect or not the contamination evaluation. The easiest way to check this was to select some
658 MinimumBias samples with PU, calculate the contamination in some bins and compare the
659 values to the one obtained with the sample used for the analysis.

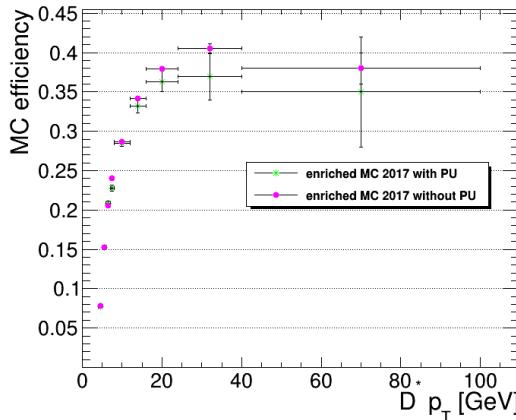


Figure 60: The D^* reconstruction efficiency reported as a function of p_T for the first PU bin, ($0 < PU < 4$), comparing the PU sample (green) to the noPU sample (magenta).

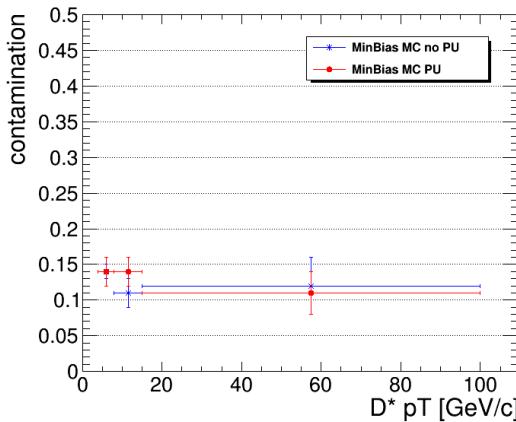


Figure 61: The contamination for the D^* meson is reported as a function of the p_T , with a comparison between the Minimum Bias PU sample (red) and the noPU sample used in the analysis (blue).

660 Since the available samples with PU are quite poor in statistics is has been necessary to merge
661 them:

- 662 • /MinBiasPt-10to35fwdJetbwdJetTuneCUETP8M113TeV-pythia8/RunIIFall15DR76-
- 663 PU25nsData2015v1castor76XmcRun2asymptoticv12-v1/AODSIM
- 664 • /MinBiasPt-35toInffwdJetbwdJetTuneCUETP8M113TeV-pythia8/RunIIFall15DR76-
- 665 PU25nsData2015v1castor76XmcRun2asymptoticv12-v1/AODSIM
- 666 • /MinBiasPt-35toInffwdJetbwdJetTuneCUETP8M113TeV-pythia8/RunIIFall15DR76-
- 667 PU25nsData2015v1castor76XmcRun2asymptoticv12ext1-v1/AODSIM

668 A direct comparison is shown in Figure 61, because of the statistics it was possible to study
669 only three p_T bins.

670 These points seem to be enough to conclude, at the end of all these studies, that the although
671 the reconstruction efficiency is different between the two samples, the contamination calcula-
672 tion is not afflicted in the whole p_T range considered. All the points, in fact, are in a good
673 agreement within the statistical uncertainties. This means that there is no needed to add any
674 systematic uncertainty related to that in the analysis and that the Minimum Bias sample used
675 in the analysis for the contamination evaluation is a good choice.

676 References

- 677 [1] *Measurement of charm production at central rapidity in proton-proton collisions at $\sqrt{s} = 7\text{TeV}$* ,
 678 Alice Collaboration, JHEP01(2012)128
- 679 [2] *Measurements of prompt charm production cross-sections in pp collisions at $\sqrt{s} = 13\text{TeV}$* , LHCb
 680 Collaboration, JHEP03(2016)159
- 681 [3] *Measurement of $D^{*\pm}$, D^\pm and D_s^\pm meson production cross sections in pp collisions at $\sqrt{s} = 7\text{TeV}$
 682 with the ATLAS detector*, ATLAS Collaboration, Nucl. Phys. B 907 (2016) 717
- 683 [4] *Nuclear modification factor of D0 mesons in PbPb collisions at $\sqrt{s_{NN}} = 5.02\text{ TeV}$* , CMS Collabora-
 684 tion, Phys. Lett. B 782 (2017) 474-496. 23 p
- 685 [5] *Open Heavy Flavor Measurements in Heavy Ion Collisions with CMS*, CMS Collaboration, Nu-
 686 clear Physics A 956 (2016) 481-484
- 687 [6] *The Review of Particle Physics*, Chin. Phys. C, 40, 100001 (2016) and 2017 update.
- 688 [7] <http://cms-service-lumi.web.cern.ch/cms-service-lumi;brilwsdoc.html>.
- 689 [8] *TMVA: Toolkit for Multivariate Data Analysis*, A. Hoecker, P. Speckmayer, J. Stelzer, J. Ther-
 690 haag, E. von Toerne, and H. Voss; PoS A CAT 040 (2007) [physics/0703039].
- 691 [9] *CMS luminosity measurements for the 2016 data-taking period* CMS Collaboration, CMS PAS
 692 LUM-17-001
- 693 [10] *Tracking POG results for pion efficiency with the D^* meson using data from 2016 and 2017* CMS
 694 Collaboration, CMS DP-2018/050
- 695 [11] *The p_T spectrum in heavy-flavour hadroproduction*, M.Cacciari, S.Frixione, P.Nason, JHEP
 696 (9805) (1998) 007
- 697 [12] *PYTHIA 6.4 Physics and Manual*, T. Sjostrand, S.Mrenna P.Skands, JHEP 0605:026,2006
- 698 [13] *A Brief Introduction to PYTHIA 8.1*, T. Sjostrand, S.Mrenna P.Skands, arXiv:0710.3820
- 699 [14] *Jet pair production in POWHEG*, S. Alioli, K. Hamilton, P. Nason, C. Oleari, E. Re, JHEP
 700 1104 (2011) 081