

August 3, 2018

## **D-hadron correlations in pp collisions at $\sqrt{s_{NN}} = 5.02$ TeV**

F. Colamaria, S. Kumar, M. Mazzilli

### **Abstract**

In this note, we present the analysis of azimuthal correlations of D mesons and primary charged  $\pi, K, p, e, \mu$  performed in the ALICE central barrel in pp collisions at  $\sqrt{s_{NN}} = 5.02$  TeV, from 2017 data taking. The analysis is performed in an extended  $p_T$  range and with additional observables with respect to pp 2013 data analysis. After a description of the analysis strategy, corrections and systematic uncertainties, the results obtained for prompt  $D^0$ ,  $D^{*+}$  and  $D^+$  mesons in different ranges of transverse momentum of the D meson and of the associated particles are presented. The results are then compared to perturbative QCD inspired Monte Carlo models and also with pp at 7 and 13 TeV analysis results for the common  $p_T$  ranges as well as with the 2016 p-Pb results.



1 **Contents**

2	<b>1 Introduction and Motivation</b>	<b>2</b>
3	<b>2 Data/Monte Carlo samples and event selection</b>	<b>3</b>
4	<b>3 Analysis strategy</b>	<b>5</b>
5	3.1 Mass plots and cut optimization . . . . .	8
6	3.1.1 Extension to vey-low $p_T$ of D mesons . . . . .	11
7	3.2 Code used for the analysis . . . . .	12
8	3.3 Further details on corrections . . . . .	12
9	3.3.1 Event Mixing . . . . .	12
10	3.3.2 Tracking and D-meson trigger efficiency . . . . .	18
11	3.3.3 Correction for bias on B to D decay topologies . . . . .	24
12	3.3.4 Secondary track contamination . . . . .	27
13	3.3.5 Beauty feed-down . . . . .	31
14	<b>4 Results</b>	<b>32</b>
15	4.1 Comparing the three D meson correlation distributions . . . . .	32
16	4.2 Average of $D^0$ , $D^+$ and $D^{*+}$ results . . . . .	38
17	4.3 Fit observable $p_T$ trends and uncertainties . . . . .	40
18	4.3.1 Results for near-side yield and width, away-side yield and width, and baseline . .	43
19	4.3.2 Comparisons pp and p-Pb . . . . .	43
20	4.3.3 Comparisons pp at 5, 7 and 13 TeV . . . . .	43
21	4.3.4 Comparisons pp@5TeV and Pythia . . . . .	43
22	<b>5 Bibliography</b>	<b>59</b>

## 1 Introduction and Motivation

The study of the azimuthal correlations of heavy-flavour particles and charged particles at the LHC energies provides a way to characterize charm production and fragmentation processes in pp collisions. The measurement also provide a way to probe our understanding of QCD in the perturbative regime, accessible in a large kinematic range given the large mass of heavy quarks. Flavour conservation in QCD implies that charm quarks are always produced as pairs of quarks and anti-quarks. The azimuthal correlations obtained using a meson carrying a heavy quark as trigger particle with the other charged particles in the same event give the possibility to study the underlying charm production mechanism in detail. In particular, prompt charm quark-antiquark pair production is back to back in azimuth at first order in leading-order perturbative-QCD (pQCD). If a hadron from the quark hadronization is taken as trigger particle, a near-side (at  $\Delta\phi = 0$ ) and an away-side (at  $\Delta\phi = \pi$ ) peaks would appear in the azimuthal correlation distributions, coming from the fragmentation of the quark pair. Heavy quarks produced from the splitting of a massless gluon can be rather collimated and may generate sprays of hadrons at small  $\Delta\phi$ . Finally, for hard-scattering topologies classified as “flavour-excitation”, a charm quark undergoes a hard interaction from an initial splitting ( $g \rightarrow c\bar{c}$ ), leading to a big separation in rapidity of the hadrons originating from the antiquark (quark) with respect to the trigger D meson and contribute to a rather flat term to the  $\Delta\phi$ -correlation distribution.

In the following note, we first describe the analysis strategy for the pp 2017 data sample in all its steps, followed by the list of analysis corrections and the estimation of systematic uncertainties. Finally the results of  $\Delta\phi$  correlations, and quantitative observable extracted to fits to those distributions, obtained for prompt  $D^0$ ,  $D^+$  and  $D^{*+}$  in different ranges of transverse momentum for the D-meson (trigger particle) and the associated particles are presented.

The extension of the momentum ranges (both for D mesons and associated particles) with respect to the 2016 pp dataset, as well as the improved precision in the common ranges allow a more thorough investigation of the charm quark fragmentation properties (multiplicity of tracks as a function of momentum, geometrical profile of charm jets,  $p_T$  distribution of the tracks inside the jet). This can also allow us to put better constraints of charm fragmentation and charm jet properties provided by models. Furthermore, 2017 pp data sample allows us a direct and more reasonable comparison with 2016 p-Pb data, since it has the same center-of-mass energy and, thanks to the higher precision and statistics, it was possible to exploit the azimuthal correlations in the same (extended and more differential)  $p_T$  ranges of both the trigger and the associated particle of the 2016 p-Pb data sample. This new pp reference data, together with new p-Pb 2016 data will help to study cold nuclear matter effects affecting the charm fragmentation in p-Pb with better precision. In addition, this new pp data can also be used as solid and precise references in view of an analysis on a Pb-Pb sample at the same energy.

## 57 2 Data/Monte Carlo samples and event selection

58 The data used for the analyses were the AOD samples of the following four datasets: LHC17p\_pass1\_FAST,  
 59 LHC17p\_pass1\_CENT\_woSDD, LHC17q\_pass1\_FAST and LHC17q\_pass1\_CENT\_woSDD . The reason  
 60 for choosing these data samples (in particular, those without the drifts for the CENT cluster) is explained  
 61 later on, in this section. It was verified, by looking at D-meson and associated charged track  $\eta$  and  $\phi$   
 62 distributions, and at the mixed-event correlation distributions for each sub-samples, that no visible differ-  
 63 ences is present for the four periods, hence it was possible to perform the analysis directly on the merged  
 64 samples without any bias.

65 The Monte Carlo productions adopted for this study were:

- 66 1. LHC18a4a2\_fast, a HF production (HIJING with GEANT3) anchored to LHC17p/q is with enrich-  
 67 ment of heavy quarks (charm and beauty) and their decay products in each of the event, performed  
 68 by PYTHIA6 with Perugia2011 tune, and with forced hadronic decays of the charmed hadrons.  
 69 This production was used for D-meson efficiency evaluation, purity estimation and Monte Carlo  
 70 closure test.
- 71 2. LHC17l3b\_fast, minimum-bias samples produced with DPMJET generator, are used for the eval-  
 72 uation of the tracking efficiencies.

73 Table 1 shows the list of runs used for the analysis, for each of the data taking periods, and of the Monte  
 74 Carlo productions used to evaluate the corrections:

The trigger mask request for the event selection is kINT7. Only events with a reconstructed primary vertex within 10 cm from the centre of the detector along the beam line are considered. This choice maximises the detector coverage of the selected events, considering the longitudinal size of the interaction region, and the detector pseudorapidity acceptances. Beam-gas events are removed by offline selections based on the timing information provided by the V0 and the Zero Degree Calorimeters, and the correlation between the number of hits and track segments in the SPD detector. This is automatically performed in the Physics Selection, a positive outcome of which is required during our event selection. The pile-up cuts for out-of-bunch pile-up protection are also invoked when calling the Physics Selection task. The minimum-bias trigger efficiency is 100% for events with D mesons with  $p_T > 1$  GeV/ $c$ . For the analyzed data samples, the probability of pile-up from collisions in the same bunch crossing is below 2% per triggered event (in most of the runs, well below 1%). Events in which more than one primary interaction vertex is reconstructed with the SPD detector (with minimum of 5 contributors, and a  $z$  distance greater than 0.8 cm) are rejected, which effectively removes the impact of in-bunch pile-up events on the analysis. Out-of-bunch tracks are effectively rejected by the Physics Selection pile-up cuts, and also by the request of at least one point in the SPD, which has a very limited time acquisition window (300 ns). Indeed, though the default associated track selection requires a minimum of 2 points in the ITS, as it will be shown later on full compatibility of the corrected results with 2 and 3 minimum ITS clusters are obtained. For FAST and CENT\_woSDD samples, the latter case indirectly forces the presence of a point in the SPD.

Since data collected during pp 2017 data taking are distinguished into two categories - one including SDD detector (CENT\_wSDD sample) and the second one without the SDD in the reconstruction, or in the acquisition (CENT\_woSDD and FAST samples, respectively), a study of performance of the D-hadron correlation analysis with respect to the data samples employed has been carried out on p-Pb 2016 data for  $D^{*+}$  and  $D^+$  mesons (more sensitive to the presence of the SDD w.r.t. the  $D^0$ , due to their reconstruction from three decay tracks) refer to p-Pb 2016 analysis note. By using the same analogy, the pp data analysis also done on the similar data samples.

### 3 Analysis strategy

The analysis follows the same strategy one used in 2016 p-Pb and pp data samples (see published paper [2] and analysis notes [8], [7]). Correlation pairs are formed by trigger particles (D mesons) reconstructed and selected in the following  $p_T^{\text{trig}}$  ranges:  $3 < p_T^{\text{trig}} < 5$  GeV/c,  $5 < p_T^{\text{trig}} < 8$  GeV/c,  $8 < p_T^{\text{trig}} < 16$  GeV/c,  $16 < p_T^{\text{trig}} < 24$  GeV/c (the possibility of extending the analysis in  $2 < p_T^{\text{trig}} < 3$  was also exploited. Further details are furnished in the next paragraph). Associated particles (charged tracks) have been reconstructed in the following  $p_T^{\text{assoc}}$  regions:  $p_T^{\text{assoc}} > 0.3$  GeV/c,  $0.3 < p_T^{\text{assoc}} < 1$  GeV/c,  $1 < p_T^{\text{assoc}} < 2$  GeV/c,  $2 < p_T^{\text{assoc}} < 3$  GeV/c,  $p_T^{\text{assoc}} > 1$  GeV/c. In this analysis, the particle identification defines the trigger particle rather than a momentum cut and therefore the momentum range of the associated particles is not constrained by that of the trigger particle. Our definition of associated particle includes primary particles of the following species: pion, kaon, proton, electron, muon. The primary particle definition comprises particle coming from the primary vertex of interaction, including those coming from strong and electromagnetic decay of unstable particles, and particles deriving from the decay of hadrons with charm or beauty. We therefore include any charged  $\pi, K, p, e, \mu$  except those coming from weak decays of strange particles and particles produced in the interaction with the detector material. This definition corresponds to that used in the method `AliAODMCParticle::IsPyphysicalPrimary()`. All associated particles surviving the selection cuts and not matching the adopted criterion are considered as a contamination whose contribution has to be corrected for.

The analysis is performed through the following steps:

1. **D meson selection and signal extraction.** For each single event, “trigger” particles are defined as the selected D meson candidates ( $D^0$ ,  $D^+$  and  $D^{*+}$ ) within a given  $p_T^{\text{trig}}$  range. The detection strategy for D mesons at central rapidity is the same performed for the analyses of the D-meson production at central rapidity [1], and also applied for the D-h analysis on 2016 pp and 2016 p-Pb samples ([8], [7]). It is based on the reconstruction of decay vertices displayed from the primary vertex by a few hundred  $\mu\text{m}$  and on the identification of the decay-particle species. The identification of the charged kaon and pion in the TPC and TOF detectors is also used, to further reduce the background at low  $p_T$ . An invariant-mass analysis is then used to extract the raw signal yield, using the same fit functions described in [2]. The D mesons are selected in the rapidity range varying from  $|y| < 0.5$  at low  $p_T$  to  $|y| < 0.8$  for  $p_T > 5$  GeV/c.
2. **Correlation of D candidates with associated tracks.** Particle pairs are formed by correlating each trigger particle with the charged primary particles passing the track selection (excluding those coming from the decay of the D-meson candidate) in a specified  $p_T^{\text{assoc}}$  interval (which can overlap with the  $p_T^{\text{trig}}$  range) and in the pseudo-rapidity range  $|\eta| < 0.8$ . For the  $D^0$  meson, also the low-momentum pion tracks from feed-down of  $D^{*+}$  mesons are removed via  $3\sigma$  invariant mass cut on the  $M(K\pi\pi) - M(K\pi)$  difference. This because these soft pion are not related to the charm quark fragmentation chain. For D meson candidates in the invariant mass signal region, defined by a  $\pm 2\sigma$  interval around the D meson mass peak, the azimuthal angle difference  $\varphi^{\text{assoc}} - \varphi^{\text{trig}} \equiv \Delta\varphi$  and the pseudorapidity difference  $\eta^{\text{assoc}} - \eta^{\text{trig}} \equiv \Delta\eta$  are evaluated and stored to build two-dimensional correlation distribution.
3. **Correction for limited acceptance and detector inhomogeneities with Event Mixing** The angular correlation distribution may be affected, even for uncorrelated pair of particles, by structures not due to physical effects, but originating from the limited detector acceptance, as well as from angular inhomogeneities in the trigger and track reconstruction efficiencies as a function of  $\Delta\varphi$  and  $\Delta\eta$ . Effects of this kind are removed using the Event Mixing technique. In this technique, the analysis is executed on the same data sample of the standard one (called “same event” analysis,

SE), but the trigger particles found in each event are correlated to charged particles reconstructed in different events (“Mixed Events” analysis, ME) with similar characteristic, in particular concerning the event multiplicity and z position of the primary vertex (see Section 3.3.1).

The differential yield of associated particles per trigger particle is obtained by

$$\frac{1}{N_{\text{trig}}} \frac{d^2 N^{\text{pair}}}{d\Delta\eta d\Delta\phi} = B_{ME}(0,0) \times \frac{S(\Delta\eta, \Delta\phi)}{B_{ME}(\Delta\eta, \Delta\phi)}, \quad (1)$$

where  $N^{\text{pair}}$  is the total number of correlated D-hadron pairs. The functions  $S(\Delta\eta, \Delta\phi)$  and  $B_{ME}(\Delta\eta, \Delta\phi)$  are the signal and the mixed event background distributions, respectively. The later is normalized to its value in  $(\Delta\eta, \Delta\phi) = (0,0)$ , i.e. ( $B(0,0)$ ). Further details on the mixed-event correction are provided in the next section.

**4. Subtraction of background correlation from signal distribution.** The invariant mass signal region also includes background D-meson candidates. Their contribution to the raw correlation distribution is subtracted as follows. For each  $p_T$  bin, the mean and the sigma of the invariant mass spectrum are extracted. For  $D^0$  and  $D^+$ , a “background” region is defined in the sidebands of the mass distribution as the interval  $4 \text{ GeV}/c^2 < |m - m^{\text{pdg}}| < 8 \text{ GeV}/c^2$  (for the  $D^{*+}$  meson, only the right sideband is used). The angular correlation distribution for background candidates in this region is extracted and normalized with respect to the background in the signal region estimated from the mass fit. This normalized background correlation distribution is then subtracted from the raw signal one to obtain the signal correlation distribution. The normalization factor is the ratio of the number of background candidates under the signal peak (obtained by integrating the background of the fit function within the signal region) over the number of background candidates in the sidebands (obtained via bin-counting in the sideband region). An example of the signal region, sideband and sideband-subtracted 1D correlation distributions (along  $\Delta\phi$ ) is shown in figure 1, together with the comparison of the three distributions after the normalization to the number of triggers.

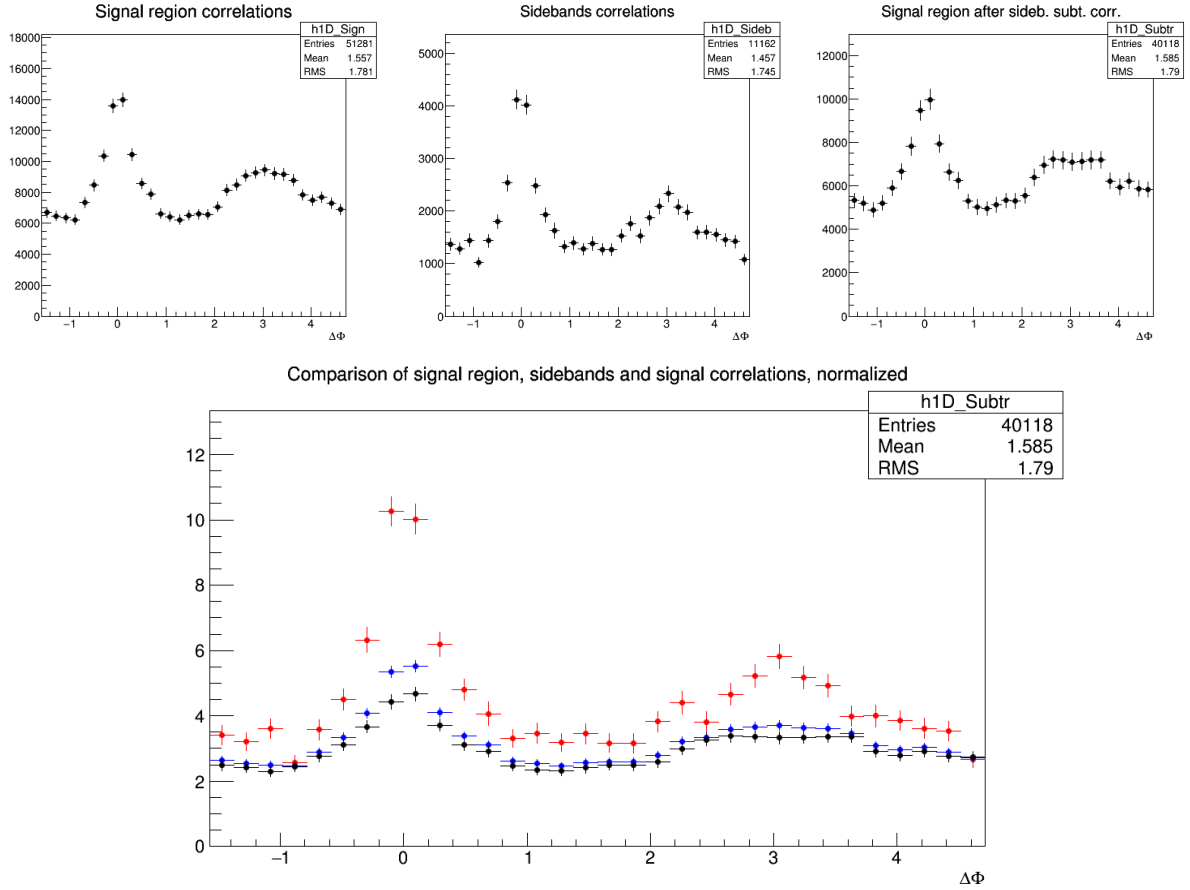
**5. Correction for D meson efficiency and associated track efficiency.** After filling the signal and background correlation distributions, it is necessary to take into account also for the correlations with tracks, those are not reconstructed, or not passing the quality selection due to poor reconstruction. In the same way, the loss of D-mesons which are not reconstructed, or do not pass the selection, impacts the correlation distribution shape. Hence, each pair is weighted by the inverse of the product of the associated track and D meson reconstruction efficiency,  $\epsilon_{\text{trk}}$  and  $\epsilon_{\text{trig}}$ . Further details are provided later on in this section.

**6. Projection in  $\Delta\phi$ .** The limited statistics available does not allow to study the two dimensional  $(\Delta\eta, \Delta\phi)$  distribution, which is therefore projected to the  $\Delta\phi$  axis by integrating on  $|\Delta\eta| < 1$ . Despite, in principle, our maximum  $\Delta\eta$  acceptance is of  $|\Delta\eta| < 1.6$ , removing the large  $|\Delta\eta|$  regions allow us to reject angular regions with very low statistics, where fluctuations would be amplified by a large mixed-event correction, and avoid the so-called wings effect.

As the difference in the azimuthal angle is periodic ( $\Delta\phi = 0 = 2\pi$ ), the  $\Delta\phi$ -range is limited to the essential range of  $2\pi$ . The  $\Delta\phi$ -limits are chosen to be  $[-\pi/2, 3\pi/2]$  in order to provide a good visibility of the correlation pattern, which peaks around 0 and  $\pi$ .

**7. Correction for the contamination of secondary particles** The DCA to primary vertex cut, applied during the associated track selection, has the role of removing the secondary particles from the associated track sample. Secondary particles are indeed produced either from long-lived strange hadrons or from interaction of particles with the detector material. A residual contamination from secondary tracks is hence expected in the correlation distributions. This contamination





**Figure 1:** Top: Example of  $D^{*+}$ -h signal region (left), sideband (middle), and signal minus sideband (right) correlation distributions. Bottom: signal region per-trigger normalized correlation distribution (blue), sideband region per-trigger normalized correlation distribution (red), background-subtracted per-trigger normalized correlation distribution (black).

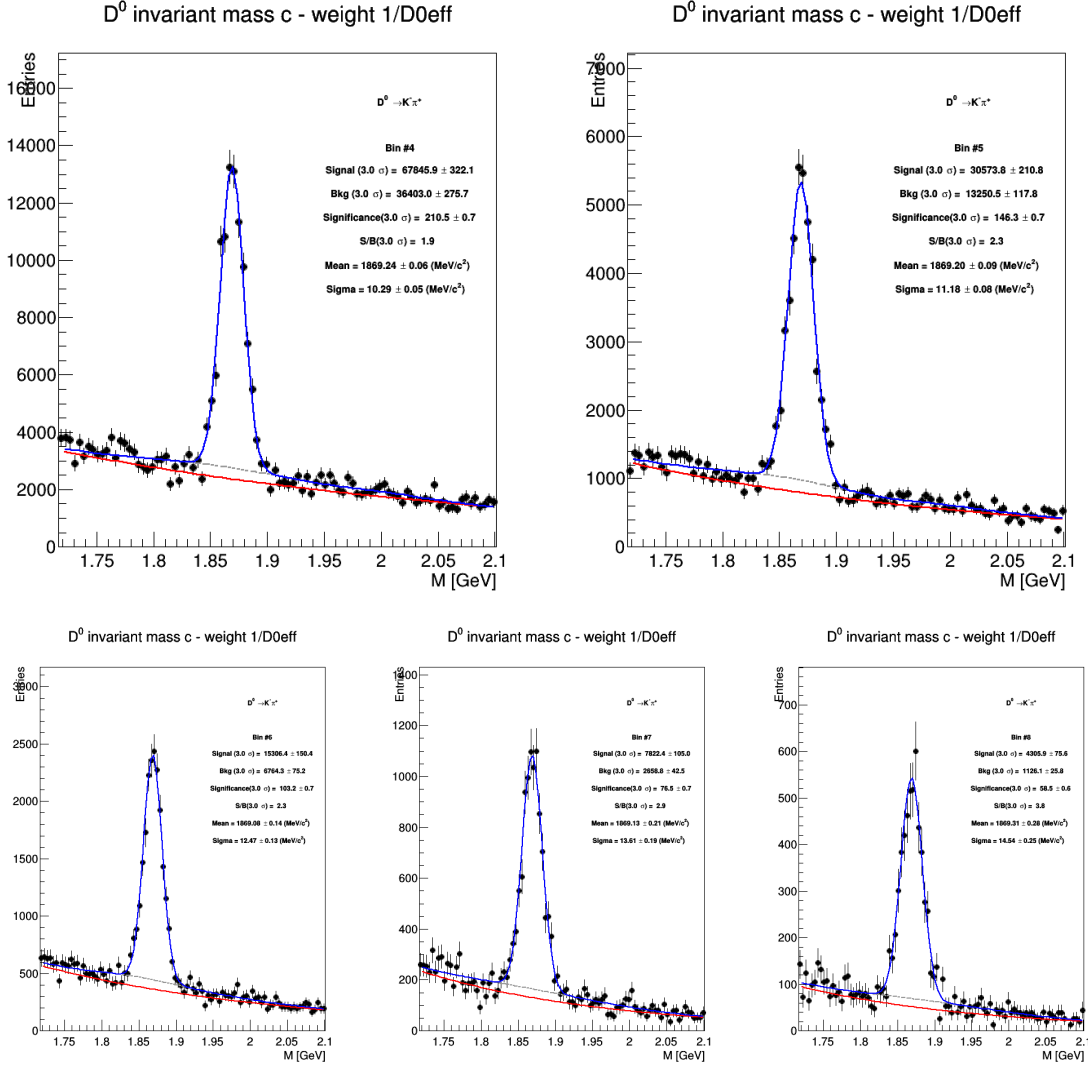
is estimated from Monte Carlo simulation based on Pythia as described more in detail in the next section. The background-subtracted event-mixing corrected correlations are multiplied by a purity factor to encounter this contribution.

8. **Correction for bias on B to D decay topologies** The presence of the topological cuts for the D-meson selection indirectly induce a bias on the topology of the B to D decay topologies, favouring cases with a small opening angle between the D-meson and the other tracks from the B decay. This affects the feed-down component of the data correlation distributions. This effect is corrected for with a procedure described in the subsection 3.3.3.
9. **Correction for feed-down of D meson from b-hadron decay** The selection strategy employed for the D meson candidates selection enhances the fraction of reconstructed D mesons coming from the decay of a b-hadron. Typical values, with the cuts used for the D-meson selection, are of the order of 10% or less. The correlation distribution of these secondary D mesons will be sensitive to the properties of beauty jets and beauty hadron decay, which in general differ from those relative to charm jets and hadrons. The procedure used to subtract this contribution is described in the next paragraphs of this section.
10. **Study of correlation properties.** The properties of the azimuthal correlation distribution are quantified by fitting the distribution with a function composed of two Gaussian functions, modelling the near and the away side peaks, and a constant term describing the baseline. The mean of the

Gaussian are fixed at  $\Delta\phi = 0$  and  $\Delta\phi = \pi$ . To accomplish the  $2\pi$  periodicity of the  $\Delta\phi$  variable, the Gaussian functions are “duplicated” with mean at  $\Delta\phi = 2\pi$  and  $\Delta\phi = -\pi$ . The fitting procedure is described in details in Section 4.

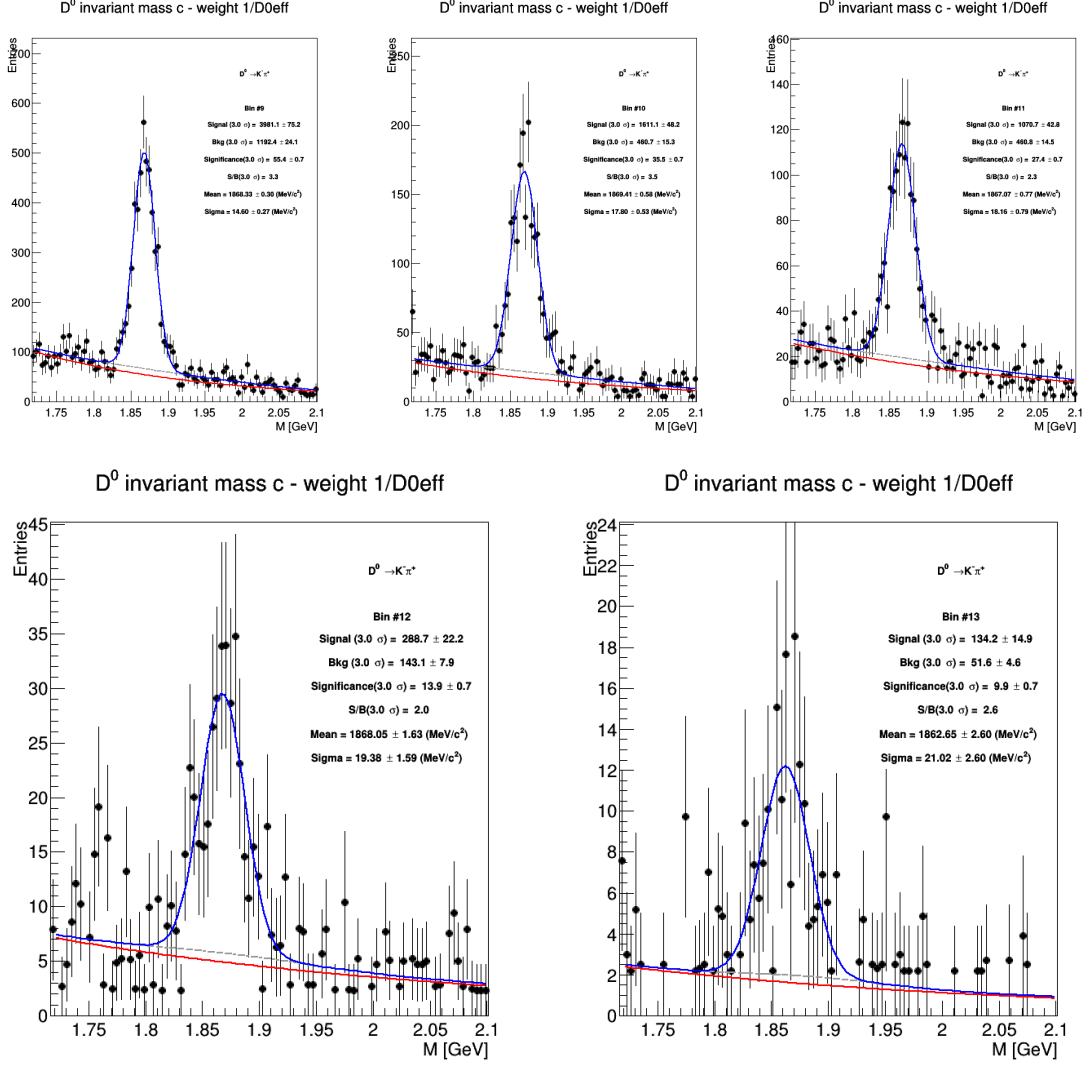
### 3.1 Mass plots and cut optimization

The invariant mass distributions of  $D^0$ ,  $D^{*+}$  and  $D^+$  in the various  $p_T$  ranges are shown in Figure 2, 3 and 4 respectively. Note that the distributions are weighted by the D-meson selection and reconstruction efficiency, to allow a correct normalization of the correlation distributions, which have also these weights.

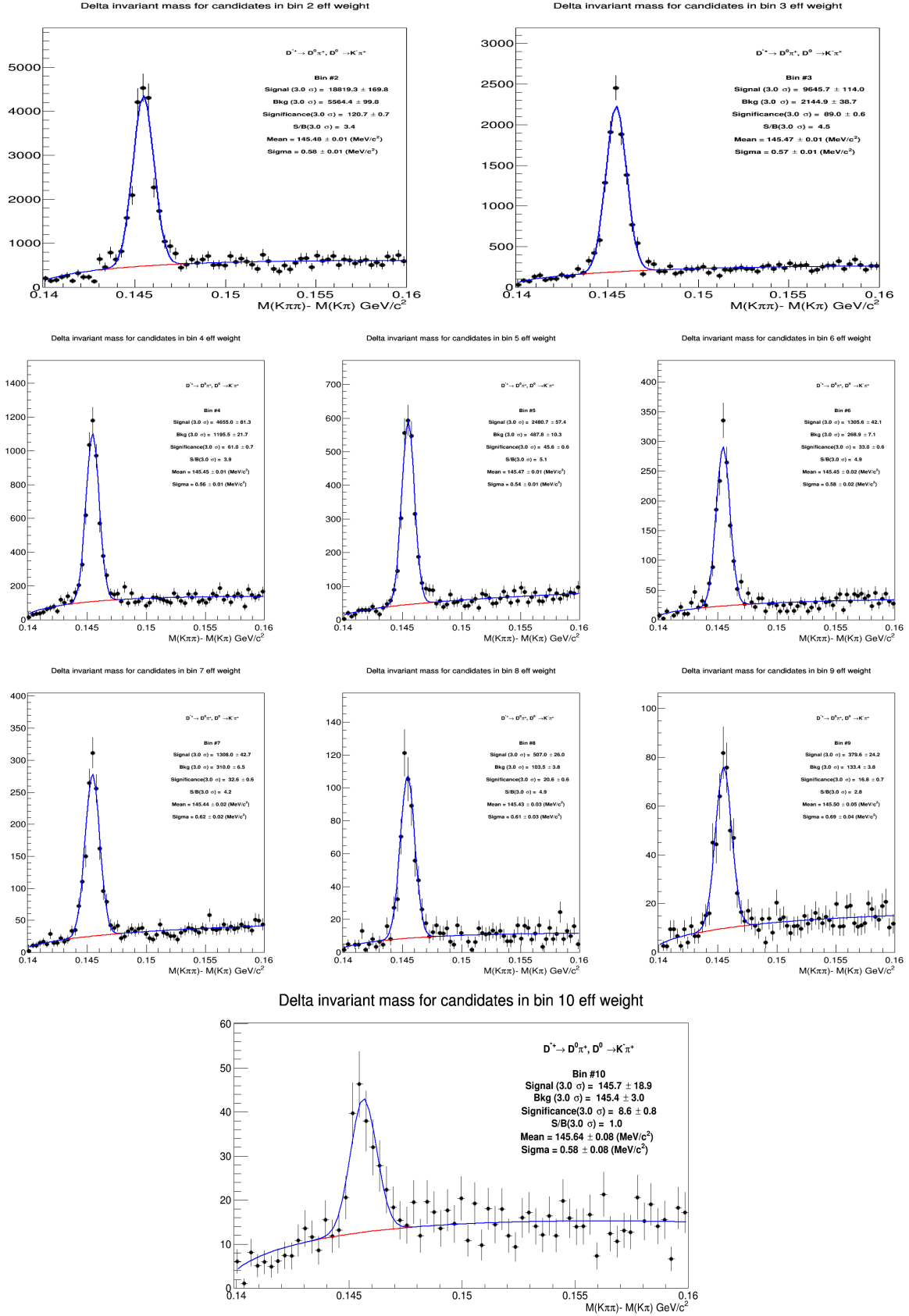


For the  $D^{*+}$ , 3 sets of cuts were compared (the 2016 p-Pb, 2016 pp and the standard D2H 2017 pp cuts). The best performance was obtained with 2016 p-Pb cuts in most of the  $p_T$  bin analysed. Indeed, despite the looser cuts applied from D2H allow to have higher signal, the p-Pb set of cuts assured a better S/B factor without losing too much signal. This allows us to reduce fluctuations induced by the sideband subtraction, that is the limiting factor for the analysis performance. The same holds for the  $D^+$ , but with the addition of cuts on the normalized decay length in  $xy$  plane and of the normalized difference between measured and expected daughter track impact parameters (topomatic cut).

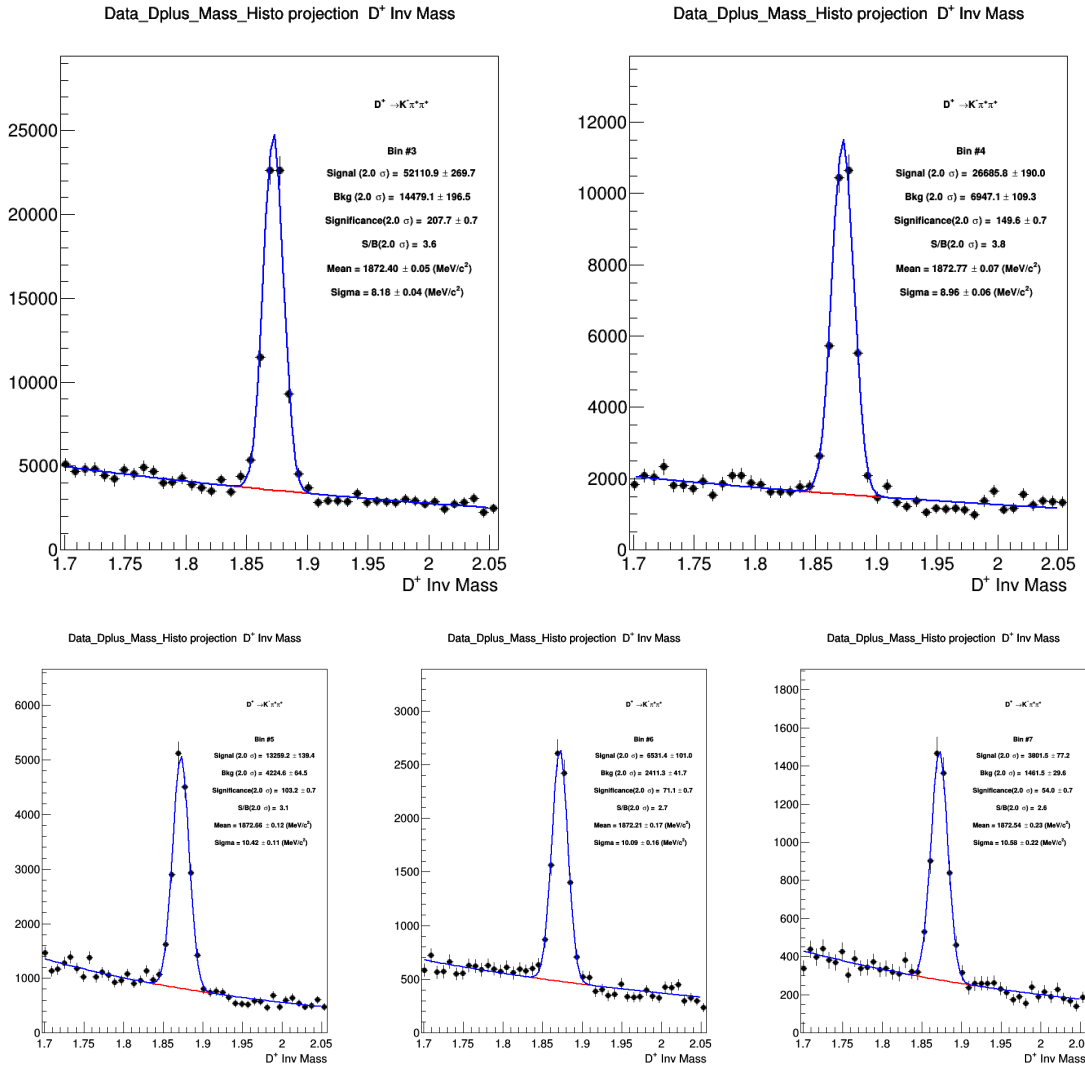
A particular cut optimization was instead performed for the  $D^0$  meson. Twelve cut sets were tried, with the goal of increasing the S/B factor, in order to reduce fluctuations induced by the sideband subtraction as explained. In Figure 3.1 the  $D^0$ -h correlation distributions are shown for the different cut sets, in



**Figure 2:** Invariant mass distributions of  $D^0$  corrected with efficiency in different  $p_T$  regions. Top:  $3 < p_T^D < 4$  GeV/c (left),  $4 < p_T^D < 5$  GeV/c (right), Mid 1:  $5 < p_T^D < 6$  GeV/c (left),  $6 < p_T^D < 7$  GeV/c (middle),  $7 < p_T^D < 8$  GeV/c (right); Mid2:  $8 < p_T^D < 10$  GeV/c,  $10 < p_T^D < 12$  GeV/c (middle),  $12 < p_T^D < 16$  GeV/c (right) and Bottom:  $16 < p_T^D < 24$  GeV/c.



**Figure 3:** Invariant mass distributions of  $D^{*+}$  corrected with efficiency in different  $p_T$  regions. Top:  $3 < p_T^D < 4$  GeV/c (left),  $4 < p_T^D < 5$  GeV/c (right), Mid 1:  $5 < p_T^D < 6$  GeV/c (left),  $6 < p_T^D < 7$  GeV/c (middle),  $7 < p_T^D < 8$  GeV/c (right); Mid2:  $8 < p_T^D < 10$  GeV/c,  $10 < p_T^D < 12$  GeV/c (middle),  $12 < p_T^D < 16$  GeV/c (right) and Bottom:  $16 < p_T^D < 24$  GeV/c.



exemplary kinematic regions (left column), together with the bin-by-bin relative statistical uncertainty on the data points (right column). The best cut set (option G) was defined from the standard cuts used for the p-Pb 2013 cross section analysis, with a tightened selection on the cosine of the pointing angle, and with the addition of a cut on the normalized decay length in  $xy$  plane and of a selection on the normalized difference between measured and expected daughter track impact parameters (topomatic cut).

### 3.1.1 Extension to very-low $p_T$ of $D$ mesons

Thanks to the higher statistic of the 2017 pp data sample, we tried to enlarge our correlation studies to very-low  $p_T$  of the trigger particle. Indeed, the Figure 5 shows the good performance on the invariant mass extraction for all the 3  $D$ -mesons with  $2 < p_T^D < 3$  GeV/ $c$ .

The extension of the analysis to this  $p_T$  interval is of high interest for the jet structure characterization. Indeed, this study allow us to investigate a kinematic region in which the trigger particle has compatible or even lower momentum with respect to the associated particles. This give us the possibility to better understand the production processes. In fact, for example, at the Leading Order (LO) for  $p_T(trig) \sim p_T(ass)$  we don't expect any peak on the near-side region. A peak could arise only from Next-to-Leading-Order production process.

Despite the good statistics, it wasn't enough to perform the correlation analysis. Indeed, the correlation distribution suffered of fluctuation expecially on the away-side region where most of the fit failed.

## 3.2 Code used for the analysis

The code used for D meson-hadron correlation analysis is fully committed in AliPhysics. The analysis classes can be found in \$ALICE\_ROOT/PWGHF/correlationHF/. The D meson specific classes where the aforementioned steps are carried out are AliAnalysisTaskDStarCorrelations, AliAnalysisTaskSED0Correlations and AliAnalysisTaskDplusCorrelations. The classes which are common to the D meson specific analysis which includes the associated particle cuts and the correlation observables are AliHFAssociatedTrackCuts, AliHFCorrelator, AliHFOfflineCorrelator, AliReducedParticle and AliDhCorrelationExtraction. Several additional classes and macros in the same folder deal with the correction steps.

## 3.3 Further details on corrections

### 3.3.1 Event Mixing

The event-mixing technique is used for correcting the raw correlation distribution for effects arising from the detector limited acceptance in rapidity and detector spatial inhomogeneities. The calculation of the Event Mixing correlation distribution is performed online. An event pool is created, where events preceding the one containing a D candidate are stored based on their properties (position of the vertex along the z axis and multiplicity). Each time a D meson candidate is found in an event, only the events contained in the same pool as the event under analysis is used to evaluate the correlations for the event mixing correction.

The multiplicity and z vertex position bins for the pools used in the p-Pb analysis (for both approaches) are the following:

- Multiplicity bins:  $(0, 20); (20, 35); (35, +\infty)$
- Vertex z (cm) =  $(-10, -2.5); (-2.5, 2.5); (2.5, 10)$

In an ideal case, the mixed event distribution is expected to have a constant flat distribution as function of  $\Delta\phi$  and a triangular shaped distribution in  $\Delta\eta$  deriving from the limited  $\eta$  acceptance of the detector. In case, instead of detector inefficient regions, or holes, in the same angular position for D meson and associated tracks, these structures produce an excess of correlations at  $\Delta\phi = 0$  in the  $\Delta\phi$  distribution. The obtained distribution is used as a weight in each correlation bin, i.e, the corrected correlation distribution is calculated as follows:

$$\frac{dN^{corr}(\Delta\phi\Delta\eta)}{d\Delta\phi d\Delta\eta} = \frac{\frac{dN^{SE}(\Delta\phi\Delta\eta)}{d\Delta\phi d\Delta\eta}}{\frac{dN^{ME}(\Delta\phi\Delta\eta)}{d\Delta\phi d\Delta\eta}} \frac{dN^{ME}(0,0)}{d\Delta\phi d\Delta\eta} \quad (2)$$

In Eq.2, the last term stands for the average of the bins in the region  $-0.2 < \Delta\eta < 0.2$ ,  $-0.2 < \Delta\phi < 0.2$  (multiple bins are used to minimize the effect of statistical fluctuations on the normalization of the mixed-event plots). This kind of normalization, adopted in the analysis of hadron-hadron correlations, relies on the fact that at  $(\Delta\eta, \Delta\phi) = (0,0)$  the trigger and associated particle experience the same detector effects. In the D meson case this is true only on average and not at very low  $p_T$ , since D mesons are reconstructed from particles that can go in different detector region. However,  $(\Delta\eta, \Delta\phi) = (0,0)$  is in any case the region with maximum efficiency for the pairs (both correlated and uncorrelated). Thus the same convention was adopted.

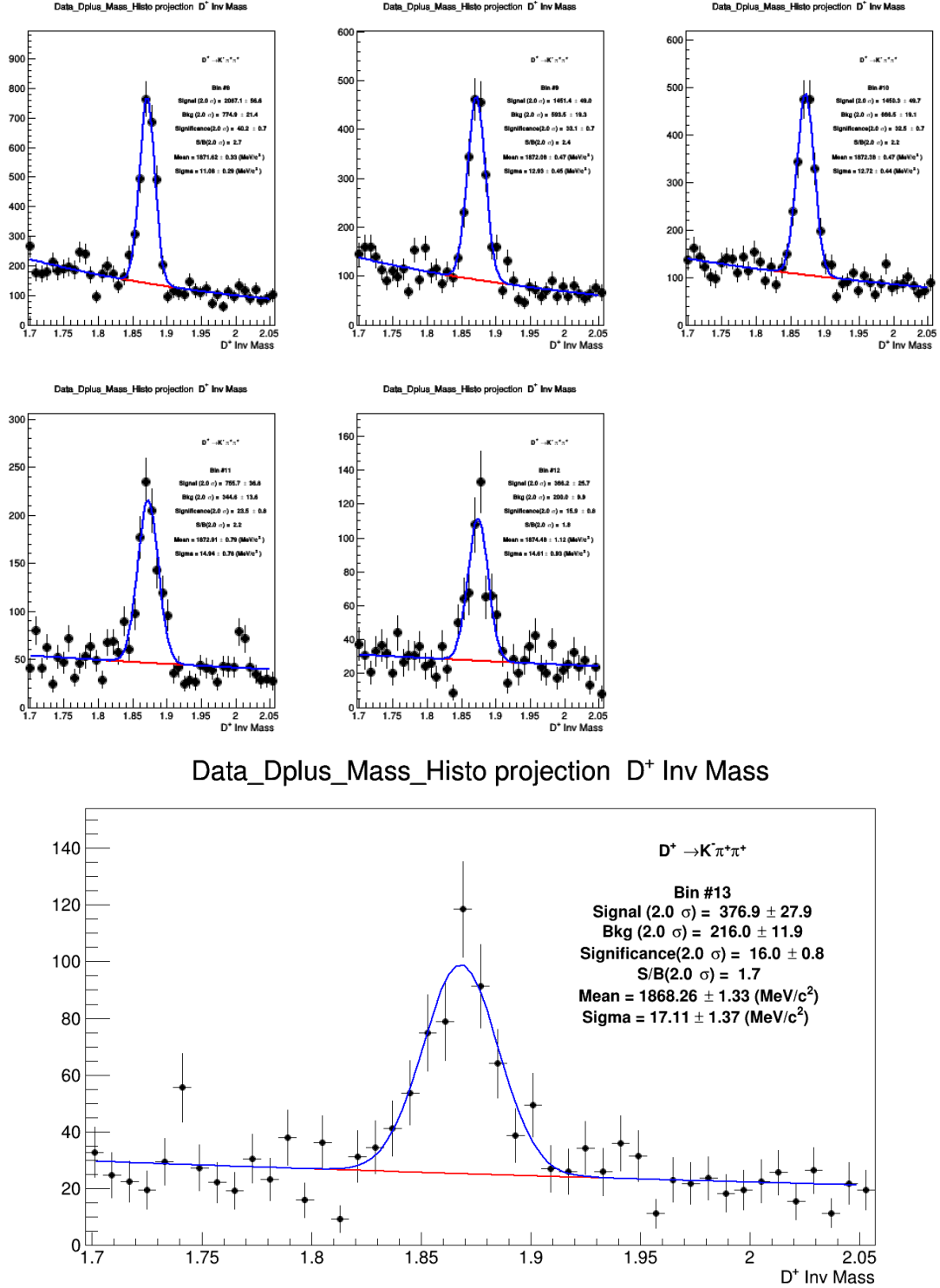
The mixed-event correlation distributions are built in both D meson signal and sideband regions. Both are corrected with the relative distributions. An example of the mixed-event distributions, and of the outcome of the mixed-event correction, is provided in Figures 6 and 3.3.1. The expected triangular shape in  $\Delta\eta$ , for the mixed-event distributions, addresses the effect of the limited detector pseudo-rapidity acceptance. Note that the mixed-event distribution is limited to the interval  $|\Delta\eta| < 1$ : the decision to

284 limit the mixed-event correction, and thus the whole analysis, to this range was taken in order to avoid  
285 the so-called “wing effect”, i.e. the wing-like structures arising in the correlation distribution at large  $\Delta\eta$   
286 due to the limited filling of the correlation bins in that region.

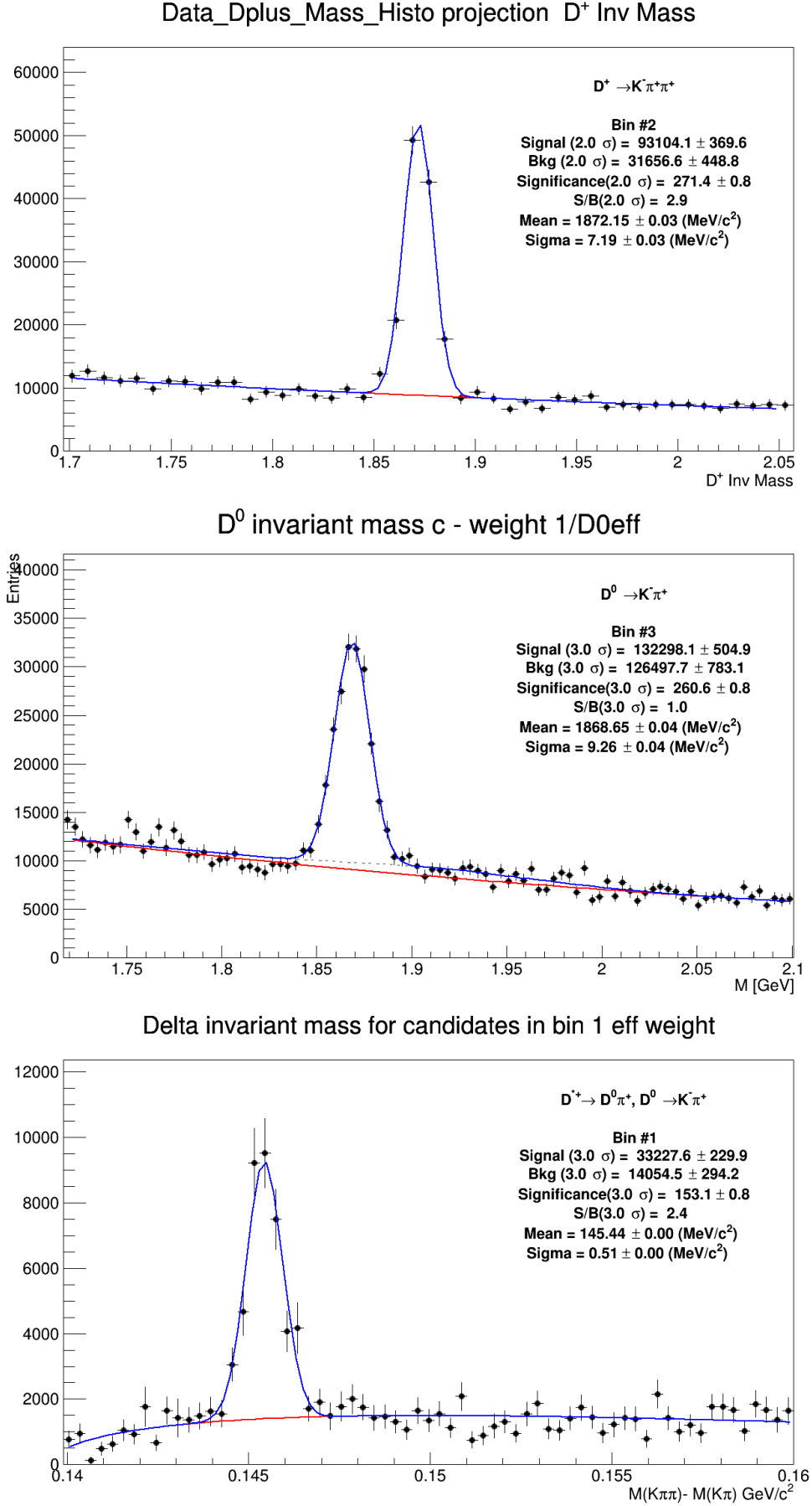
Type	Production	Run list	nEvents
Monte-Carlo	LHC18a4a2_fast (c/b enriched) [GEANT3]	282343, 282342, 282341, 282340, 282314, 282313, 282312, 282309, 282307, 282306, 282305, 282304, 282303, 282302, 282247, 282230, 282229, 282227, 282224, 282206, 282189, 282147, 282146, 282127, 282126, 282125, 282123, 282122, 282120, 282119, 282118, 282099, 282098, 282078, 282051, 282050, 282031, 282025, 282021, 282016, 282008, 282367, 282366, 282365 = <b>[44 runs]</b>	23M
	LHC17l3b_fast (Minimum Bias sample) [GEANT3]	282008, 282016, 282021, 282025, 282031, 282050, 282051, 282078, 282098, 282099, 282118, 282119, 282120, 282122, 282123, 282125, 282126, 282127, 282146, 282147, 282189, 282206, 282224, 282227, 282229, 282230, 282247, 282302, 282303, 282304, 282305, 282306, 282307, 282309, 282312, 282313, 282314, 282340, 282341, 282342, 282343, 282365, 282366, 282367 = <b>[44 runs]</b>	23M
Data	LHC17p_pass1_FAST	282343, 282342, 282341, 282340, 282314, 282313, 282312, 282309, 282307, 282306, 282305, 282304, 282303, 282302, 282247, 282230, 282229, 282227, 282224, 282206, 282189, 282147, 282146, 282127, 282126, 282125, 282123, 282122, 282120, 282119, 282118, 282099, 282098, 282078, 282051, 282050, 282031, 282025, 282021, 282016, 282008 = <b>[41 runs]</b>	985M total
	LHC17p_pass1_CENT_woSDD	282343, 282342, 282341, 282340, 282314, 282313, 282312, 282309, 282307, 282306, 282305, 282304, 282303, 282302, 282247, 282230, 282229, 282227, 282224, 282206, 282189, 282147, 282146, 282127, 282126, 282125, 282123, 282122, 282120, 282119, 282118, 282099, 282098, 282078, 282051, 282050, 282031, 282030, 282025, 282021, 282016, 282008 = <b>[42 runs]</b>	
	LHC17q_pass1_FAST	282367, 282366, 282365 = <b>[3 runs]</b>	
	LHC17q_pass1_CENT_woSDD	282367, 282366, 282365 = <b>[3 runs]</b>	

Table 1: Data Set and Run list

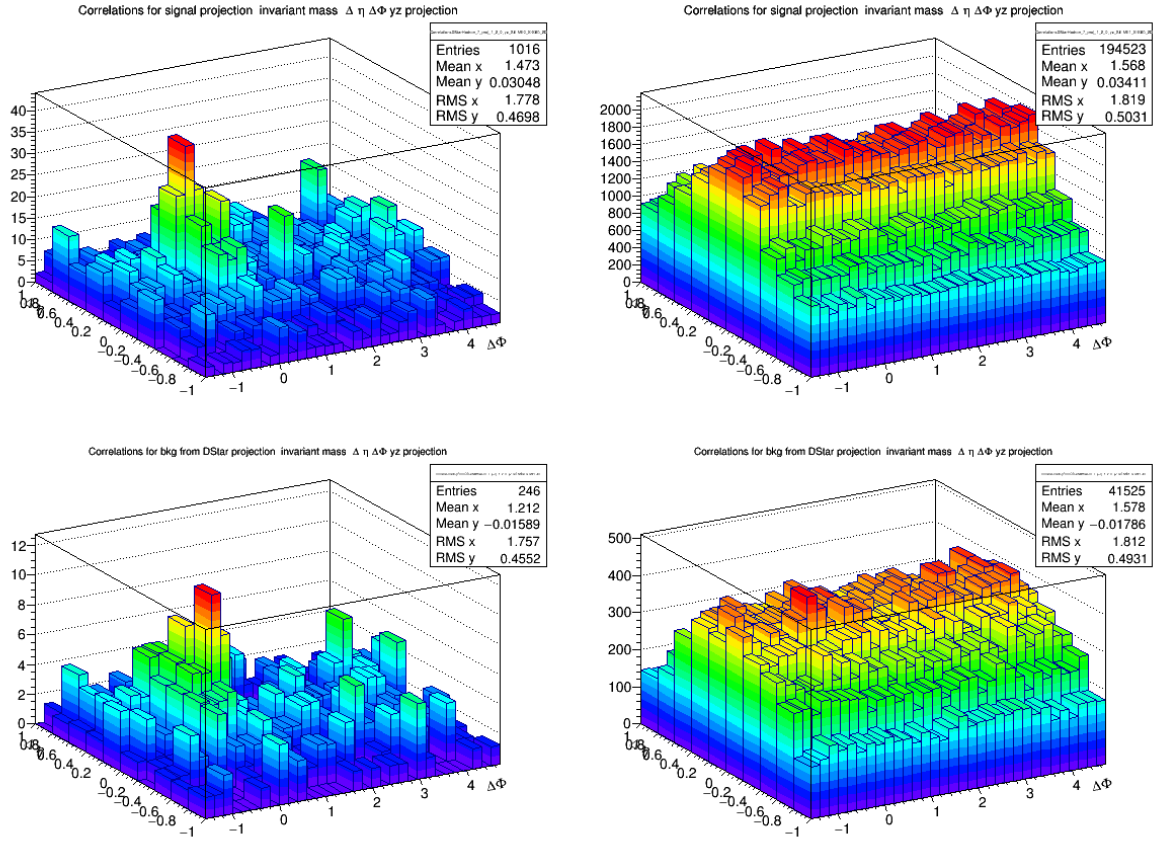




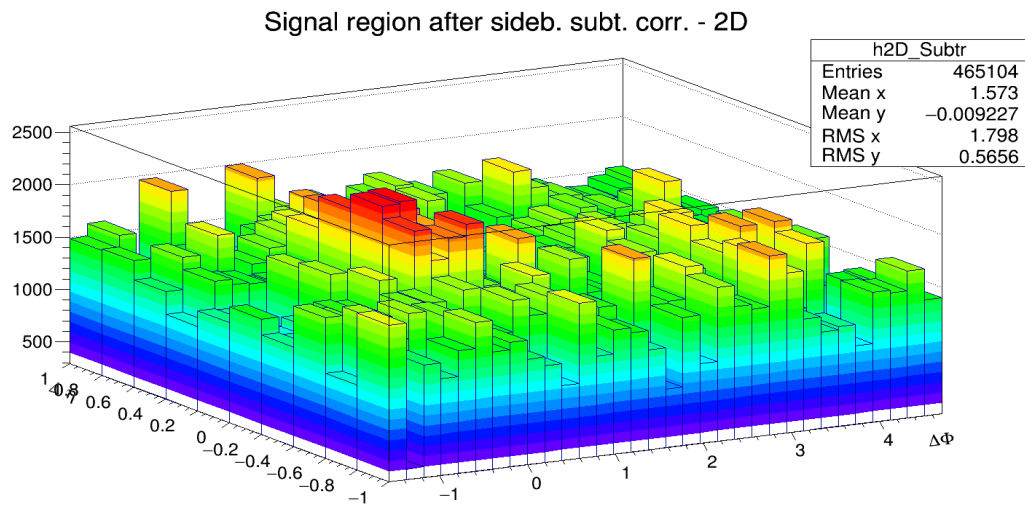
**Figure 4:** Invariant mass distribution of  $D^+$  corrected with efficiency in different  $p_T$  regions. Top:  $3 < p_T^D < 4$  GeV/c (left),  $4 < p_T^D < 5$  GeV/c (right), Mid 1:  $5 < p_T^D < 6$  GeV/c (left),  $6 < p_T^D < 7$  GeV/c (middle),  $7 < p_T^D < 8$  GeV/c (right); Mid2:  $8 < p_T^D < 10$  GeV/c,  $10 < p_T^D < 12$  GeV/c (middle),  $12 < p_T^D < 16$  GeV/c (right) and Bottom:  $16 < p_T^D < 24$  GeV/c.



**Figure 5:** Invariant mass distribution of D<sup>+</sup> (top), D<sup>0</sup> (mid) and D<sup>+</sup> (bottom) corrected with efficiency for  $2 < p_T^D < 3$ .



**Figure 6:**  $D^{*+}$  meson ( $\Delta\phi$ ,  $\Delta\eta$ ) correlation for in the signal region (top row) and sidebands (bottom row) from Single Event (left) and Mixed Event analysis (center) for high  $p_T$ :  $8 < p_T < 16$  GeV/c with associated  $p_T > 0.3$  GeV/c. The right column shows the SE/ME corrected distributions.



Mid:

$(\Delta\phi, \Delta\eta)$  correlation distribution of  $D^{*+}$ -h with  $3 < p_T < 5$  GeV/c and associated track  $p_T$  Threshold:  $p_T > 0.3$  GeV/c

### 3.3.2 Tracking and D-meson trigger efficiency

(i) **Tracking efficiency** - The tracking efficiency was calculated by obtaining the ratio between the yield at the reconstructed level and generated level, for a defined “type” of particles (in our case non-identified particles) and it is estimated differentially in  $p_T$ ,  $\eta$ , and  $z_{vtx}$  of the charged particles.

Tracking efficiency maps were produced as TH3D histograms ( $p_T$ ,  $\eta$ ,  $z_{vtx}$ ) obtained from MC analysis on the minimum-bias samples LHC1713b\_fast and LHC1713b\_cent\_woSDD anchored to LHC17p,q data samples, considering only primary pions, kaons, protons, electrons and muons, and applying at reconstructed level the track selections (summarized in Table. 2). These efficiency maps were used in the analysis tasks to extract single track efficiencies; each correlation pairs found in the data analysis was inserted in correlation plots with a weight of **1/efficiency value**. The 1D ( $p_T$  dependence) tracking efficiency for all the five species as well as the tracking efficiency specie by specie are shown in Fig. 7.

299 Details of cuts at event level and particle/track selection at different steps are listed in Table 2 .

MC Generated	
Stages	Cuts
1.MC Part with Generated Cuts	<b>After Event Selection</b> Charge PDG Code Physical Primary <b>Kinematics Cuts</b> $-0.8<\eta<0.8$ $p_T > 0.3$ (GeV/c)
2. MC Part with Kine Cuts	
MC Reconstructed	
4. Reco tracks	<b>After Event Selection</b> Physical Primary <b>Kinematics Cuts</b> $-0.8<\eta<0.8$ $p_T > 0.3$ (GeV/c) <b>Quality Cuts</b> SetRequireSigmaToVertex(kFALSE) SetDCAToVertex2D(kFALSE) SetMinNCrossedRowsTPC(70) SetMinRatioCrossedRowsOverFindableClustersTPC(0.8) SetMinNClustersITS(2) SetMaxChi2PerClusterTPC(4) SetMaxDCAToVertexZ(1) SetMaxDCAToVertexXY(1) SetRequireTPCRefit(TRUE) SetRequireITSRefit(FALSE)
5. Reco tracks with Kine Cuts	
6. MC true with Quality Cuts	
7. Reco tracks with Quality Cuts	<b>Same as step 6</b>

**Table 2:** The list of event and particle/track selection cuts used in the estimation of single track efficiency

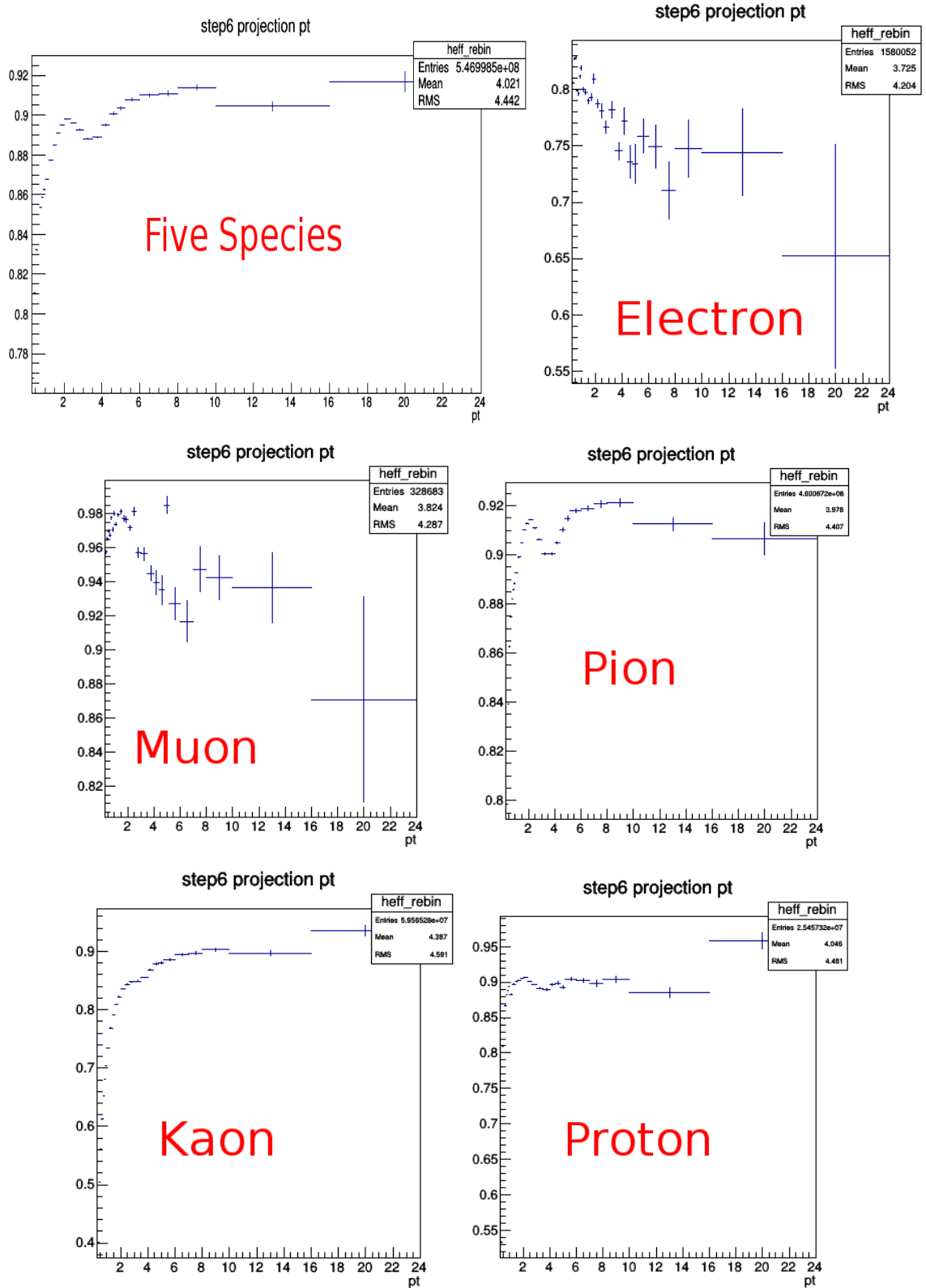
300

301 **(ii) D meson efficiency** - Due to limited statistics, the correlation analysis is performed in quite wide  $p_T$   
 302 bins and in each of them the reconstruction and selection efficiency of D mesons is not flat, in particular  
 303 in the lower  $p_T$  region. We correct for the  $p_T$  dependence of the trigger efficiency within each  $p_T$ -bin.

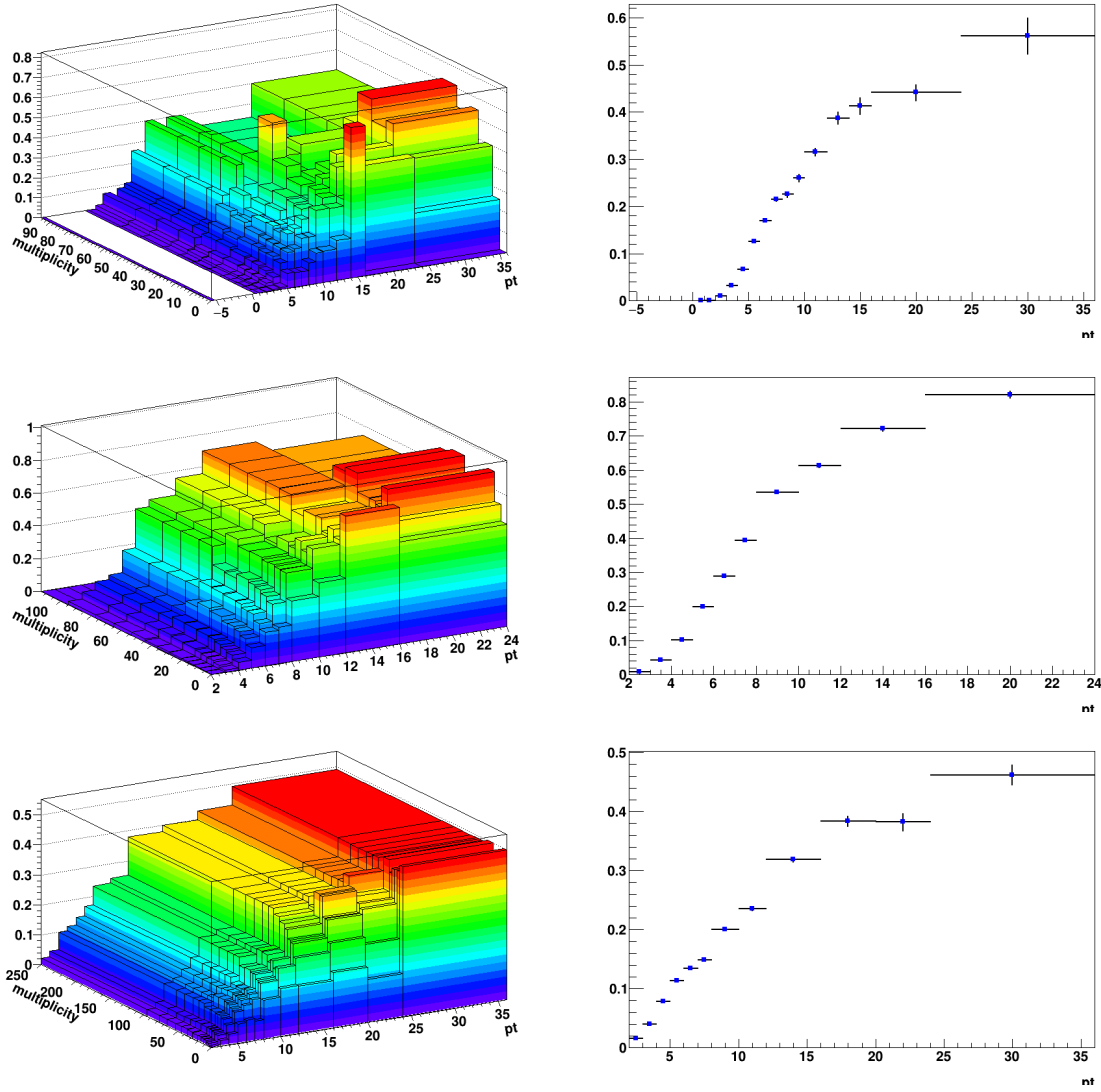
304 This correction is applied online, by using a map of D meson efficiency as a function of  $p_T$  and event  
 305 multiplicity (in terms of SPD tracklets in  $|\eta| < 1$ ) extracted from the enriched Monte Carlo sample  
 306 LHC18a4a2\_fast. The  $\eta$  dependence was neglected due to the statistics of the available Monte Carlo  
 307 sample, which rule out the possibility of performing a 3D study.

308 To properly count the number of trigger particles used to normalize the correlation distributions,  $N_{\text{trig}}$ ,  
 309 each D meson is weighted with the inverse of its efficiency in the invariant mass distribution. The main  
 310 role of the correction for the D meson efficiency is to account for the  $p_T$  dependence of the correlation  
 311 distribution within a given D meson  $p_T$  interval. Indeed, only the  $p_T$  shape of the D meson efficiency  
 312 within the correlation  $p_T^{\text{trig}}$  ranges is relevant while the average value in the  $p_T$  range is simplified due to  
 313 the normalization of the correlation distribution to the number of trigger particles.

<sup>314</sup> Efficiency plots for  $D^0$ ,  $D^+$  and  $D^{*+}$  mesons are shown in Figs. 8 and 9.

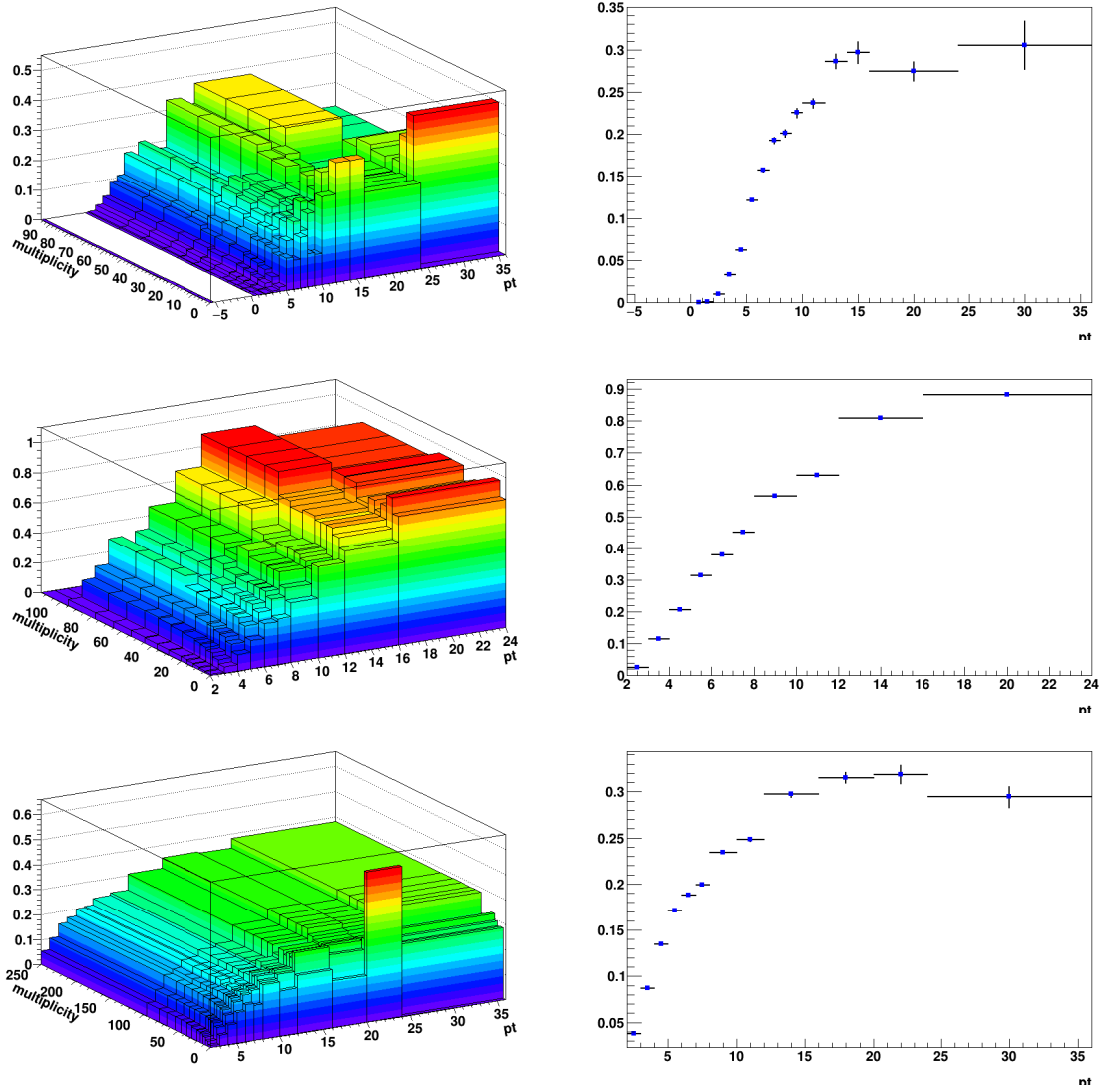


**Figure 7:** 1D (vs  $p_T$ ) tracking efficiency map for standard track selection, evaluated for five species (electron, muon, pion, kaon and proton) and also different species using data sample LHC1713b\_fast.



**Figure 8:** Top panel: ( $p_T$ , multiplicity) dependence (left) and  $p_T$  dependence (right) of prompt  $D^+$  meson efficiency. Mid panel: ( $p_T$ , multiplicity) dependence (left) and  $p_T$  dependence (right) of prompt  $D^{*+}$  meson efficiency. Bottom panel: ( $p_T$ , multiplicity) dependence (left) and  $p_T$  dependence (right) of prompt  $D^0$  meson efficiency.





**Figure 9:** Top panel: ( $p_T$ , multiplicity) dependence (left) and  $p_T$  dependence (right) of feed-down  $D^+$  meson efficiency. Mid panel: ( $p_T$ , multiplicity) dependence (left) and  $p_T$  dependence (right) of feed-down  $D^{*+}$  meson efficiency. Bottom panel: ( $p_T$ , multiplicity) dependence (left) and  $p_T$  dependence (right) of feed-down  $D^0$  meson efficiency.

### 3.3.3 Correction for bias on B to D decay topologies

To verify the consistency of the analysis chain and of the corrections applied to the correlation distributions extracted from data, a Monte Carlo closure test was setup and tried on the  $D^0$ -h analysis.

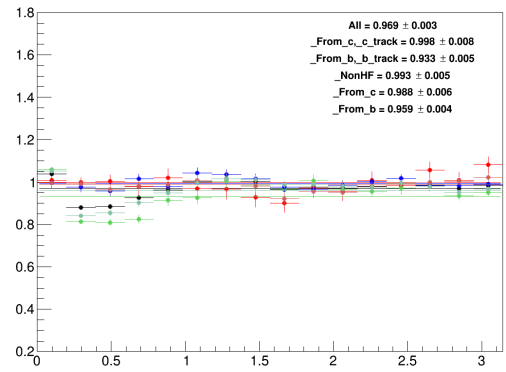
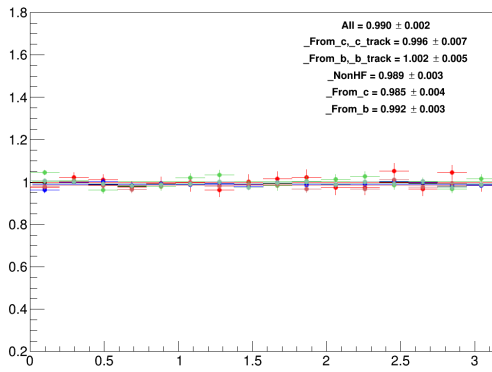
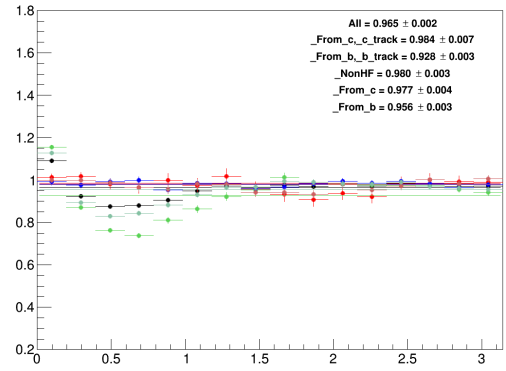
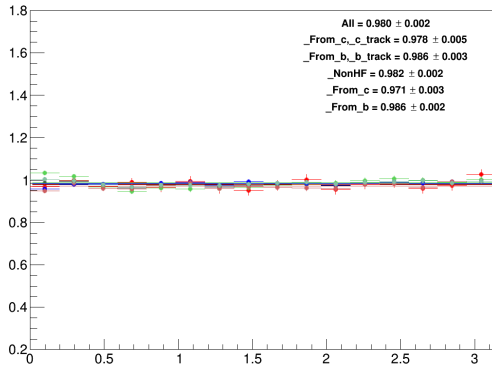
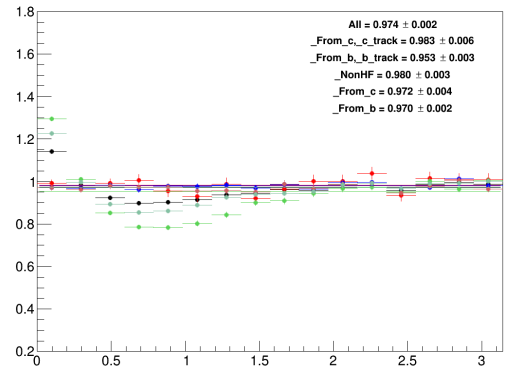
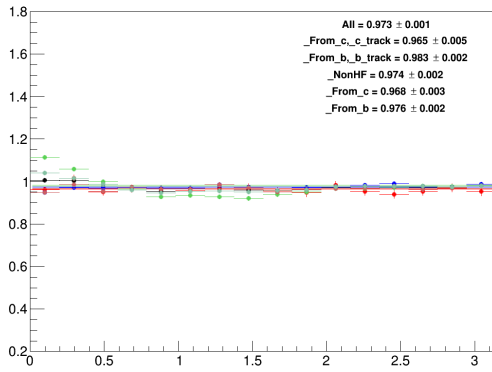
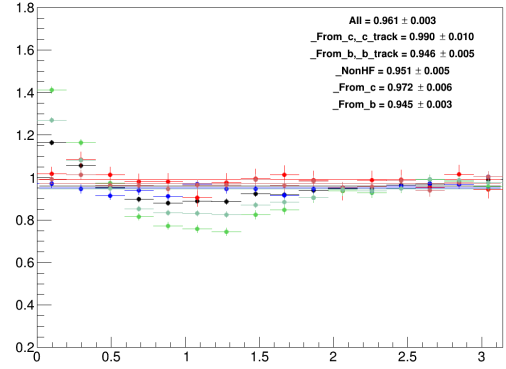
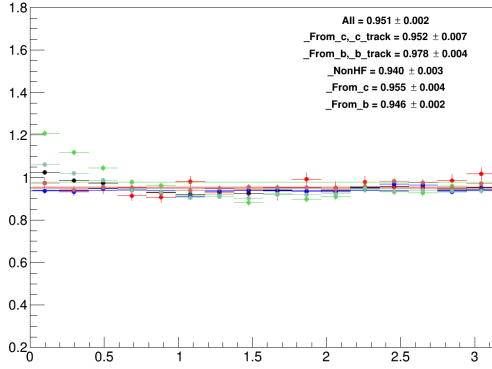
On the Monte Carlo enriched with charm and beauty quarks (LHC18a4a2\_fast, with GEANT3), the correlation analysis was performed both at kinematic level and at reconstructed level. At kinematic level, only acceptance cuts were applied on the D mesons and the associated particles, using the Monte Carlo information for the identification of the D mesons and the hadrons in the event and rejecting the non-primary particles. At reconstructed level, the analysis was performed as if it were executed on data, applying the event selection, the acceptance cuts for D mesons and the associated particles, selecting the D meson candidates with filtering cuts on their daughters, topological cuts and PID selection, and then keeping only the true D mesons by matching with the Monte Carlo truth; non-primary particles were rejected by means of the DCA selection. Event mixing correction was applied both at reconstructed and at kinematic level, where it takes into account just the effects of the acceptance cuts. In addition, at reconstructed level, the efficiency corrections for D mesons and associated tracks were also applied.

The consistency check was performed to verify whether, after having applied all the corrections to the azimuthal correlation plots at reconstructed level, the results were compatible with the ones at kinematic level. Hence, the ratios of fully corrected reconstructed plots over kinematic plots were evaluated in all the  $D^0$   $p_T$  bins and for the various  $p_T$  thresholds for the associated tracks, separating the contributions for the different origins of particles and triggers. The ratios, shown in Figure 10, denote a good compatibility with 1, within the uncertainties, with the only exception being due to some structures in the near side region for the beauty origin case. It was verified that these structures are induced by our topological selection for the D mesons. Indeed, in cases in which the D meson triggers come from B hadrons, applying the topological cuts (especially the cosine of the pointing angle) tends to favour cases with a small angular opening between the products of the B hadron decay (i.e. the D meson trigger itself and other particles), with respect to cases where the B decay particles are less collinear.

In the Monte Carlo closure test, this situation is reflected in the correlation distributions at reconstructed level, where the topological selection is applied, while it does not occur at kinematic level. Hence, in the reconstructed/kinematic ratio, the distribution would show an excess for  $\Delta\phi = 0$  (due to the favoured decays with small opening angle), which is then compensated by a depletion for larger values of  $\Delta\phi = 0$  (corresponding to B decays with larger angles, which are disfavoured). These structures are prominent at low  $D^0$   $p_T$ , where the topological cuts are tighter, and tend to disappear at higher  $p_T$ , where the selections are released. They are also larger in the higher associated track  $p_T$  ranges, where the fraction of B-hadron decay tracks dominate the overall correlation distributions.

The data correlation distribution need to be corrected for this bias, and in particular for the enhancement of b-origin correlation pairs at the centre of the near side region, which would influence the near-side peak features. In order to do this, the amount of the b-origin excess is evaluated from the Reco/Kine ratio, by considering the b- $D^0$ -all tracks case (dark green points). The excess at Reco level (affecting data) is quantified as a  $\Delta\phi$  modulation **modul** for the five points on each side of the  $\Delta\phi = 0$  value (or, equivalently, on the first five points of the reflected distributions, which start from  $\Delta\phi = 0$ ). This is done separately in each  $p_T$  range. Then, the correction is done by applying this modulation to the data correlation distributions, but taking into account that only the correlation entries from  $B \rightarrow D$  are affected, while the  $c \rightarrow D$  correlations need to be left unaltered. In particular, it has to be considered that:

- On data, the  $B \rightarrow D$  correlation pairs are only a fraction  $(1 - f_{\text{prompt}})$  of the total.
- The amplitude of  $B \rightarrow D|_{\text{amplit}}$  correlation pattern is different (greater) than the amplitude of the  $c \rightarrow D|_{\text{amplit}}$  correlation pattern:



**Figure 10:** Ratios of fully corrected azimuthal correlation plots at reconstructed level over azimuthal correlation plots at kinematic level, in the two  $D^0$   $p_T$  bins, for the different associated  $p_T$  ranges. Black points: All  $D^0$ -all hadrons, normalized by all  $D^0$  triggers; light red points:  $D^0$  from c-hadrons from c, normalized by c- $D^0$  triggers; dark red points:  $D^0$  from c-all hadrons, normalized by c- $D^0$  triggers; light green points:  $D^0$  from b-hadrons from b, normalized by b- $D^0$  triggers; dark green points:  $D^0$  from b-all hadrons, normalized by b- $D^0$  triggers; blue points: All  $D^0$ -hadrons from light quarks, normalized by all  $D^0$  triggers. The panels show the ranges:  $2 < p_T(D) < 3 \text{ GeV}/c$ ,  $0.3 < p_T(\text{assoc}) < 1 \text{ GeV}/c$  (1st row-left);  $2 < p_T(D) < 3 \text{ GeV}/c$ ,  $p_T(\text{assoc}) > 1 \text{ GeV}/c$  (1st row-right);  $3 < p_T(D) < 5 \text{ GeV}/c$ ,  $0.3 < p_T(\text{assoc}) < 1 \text{ GeV}/c$  (2nd row-left);  $3 < p_T(D) < 5 \text{ GeV}/c$ ,  $p_T(\text{assoc}) > 1 \text{ GeV}/c$  (2nd row-right);  $5 < p_T(D) < 8 \text{ GeV}/c$ ,  $0.3 < p_T(\text{assoc}) < 1 \text{ GeV}/c$  (3rd row-left);  $5 < p_T(D) < 8 \text{ GeV}/c$ ,  $p_T(\text{assoc}) > 1 \text{ GeV}/c$  (3rd row-right);  $8 < p_T(D) < 16 \text{ GeV}/c$ ,  $0.3 < p_T(\text{assoc}) < 1 \text{ GeV}/c$  (4th row-left);  $8 < p_T(D) < 16 \text{ GeV}/c$ ,  $p_T(\text{assoc}) > 1 \text{ GeV}/c$  (4th row-right).

Thus, the following equation is applied to get the corrected  $C(\Delta\phi)_{\text{corr}}$  data points starting from the raw ones,  $C(\Delta\phi)_{\text{raw}}$ :

$$C(\Delta\phi)_{\text{corr}} = C(\Delta\phi)_{\text{raw}} \cdot \left[ \frac{c \rightarrow D|_{\text{amplit}}}{(B + c) \rightarrow D|_{\text{amplit}}} \cdot f_{\text{prompt}} + \frac{B \rightarrow D|_{\text{amplit}}}{(B + c) \rightarrow D|_{\text{amplit}}} \cdot (1 - f_{\text{prompt}}) \cdot \frac{1}{\text{modul}} \right] \quad (3)$$

where  $(B + c) \rightarrow D|_{\text{amplit}} = c \rightarrow D|_{\text{amplit}} \cdot f_{\text{prompt}} + B \rightarrow D|_{\text{amplit}} \cdot (1 - f_{\text{prompt}})$ , and where the two amplitudes are evaluated from the Monte Carlo distributions of Figure ?? at reconstructed level (so, including the bias), and  $f_{\text{prompt}}$  with the procedure described in ??. Applying the **modul** factor to the beauty part of the data correlation distributions brings its value back to the generated level case, effectively removing the bias. The effect of the correction is a shift of the data points in the near-side region (in general, downward in the first and second points, the upward in the others). The maximum value of the shift is of about 5%, at the centre of the near-side peak, for the lowest D-meson  $p_T$  range ( $3 < p_T < 5 \text{ GeV}/c$ ) and the highest associated track  $p_T$  range ( $p_T > 3 \text{ GeV}/c$ ). The typical values are instead of a couple of percentage points. The correction is zero in the highest D-meson  $p_T$  range. To take into account for possible inaccuracies in the definition of the modulations, or in their rescaling, a systematic uncertainty is applied on the corrected data points, with value  $|C(\Delta\phi)_{\text{corr}} - C(\Delta\phi)_{\text{raw}}|/\sqrt{(12)}$ , on each side of the data points affected by the bias (symmetric uncertainty).

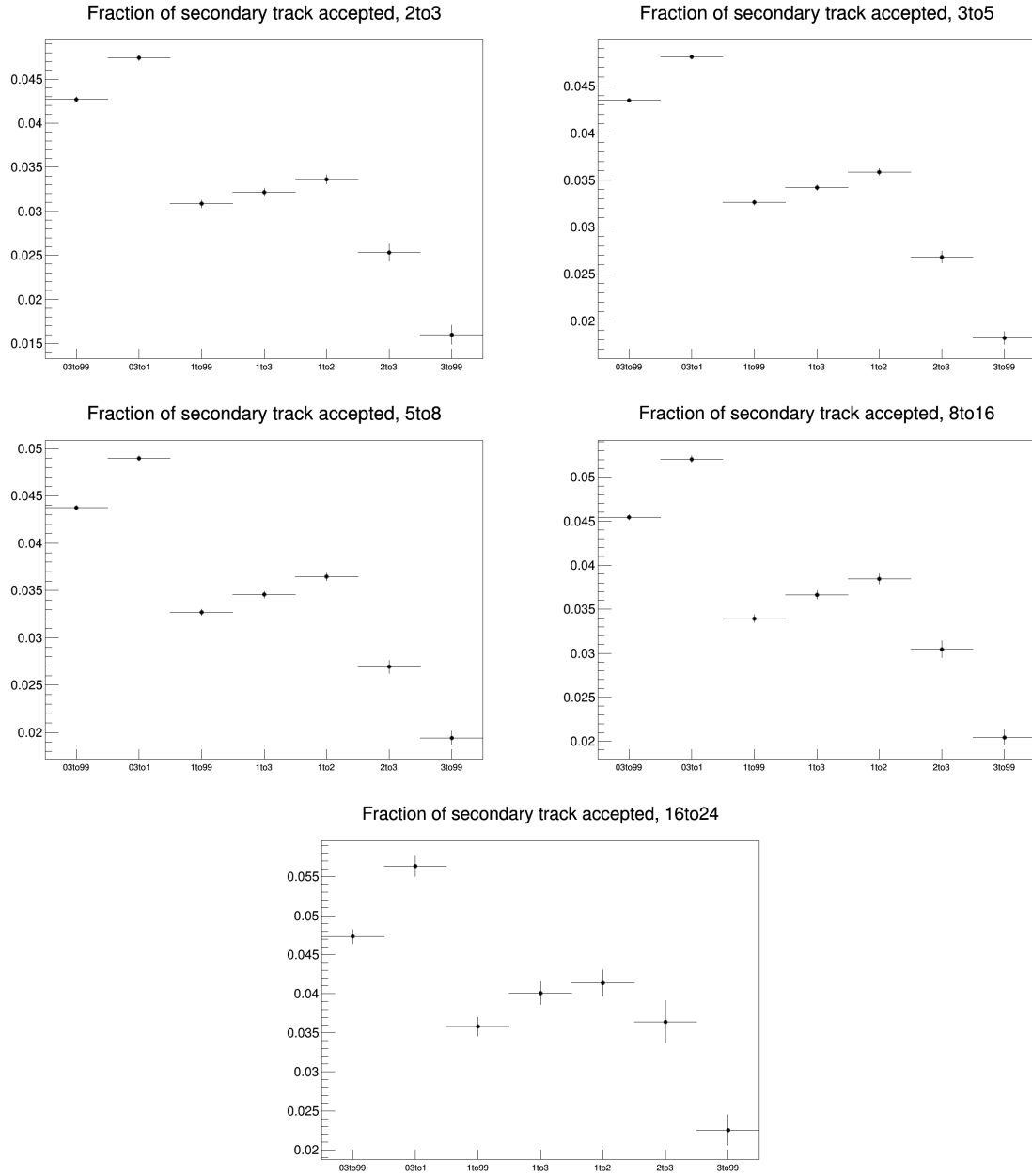
### 3.3.4 Secondary track contamination

The secondary tracks inside the associated track sample, due to interaction of primary track with the detector material or to decays of strange hadrons, are mostly removed by the DCA cuts applied during the cut selection phase ( $DCA(xy) < 1$  cm,  $DCA(z) < 1$  cm). Anyway, a small fraction of secondary tracks survives this cut, and the data correlation distributions have to be corrected for this residual contamination. The fraction of surviving secondary tracks is evaluated via a study on the LHC18a4a2\_fast sample, by counting the number of tracks accepted by the selection whose corresponding generated-level track doesn't satisfy the `IsPhysicalPrimary()` call, and dividing this number by the total number of accepted tracks. The outcome of the check is reported in Figure 11. As it's visible, no more than 5% secondary tracks pass the selection. Moreover, the fraction of residual secondary tracks is flat along the  $\Delta\phi$  axis, as shown, for exemplary  $p_T$  regions, in Figure 12, where the inhomogeneities are always below 1%. For this reason, it is possible to directly scale the data correlation distributions by their purity fraction (i.e. 1 - secondary contamination). This is done with an associated  $p_T$  dependence, due to the increase of the purity with the track  $p_T$ , while the purity fraction is taken flat versus the D-meson  $p_T$ . The purity values that were chosen are the following (NOT FINAL):

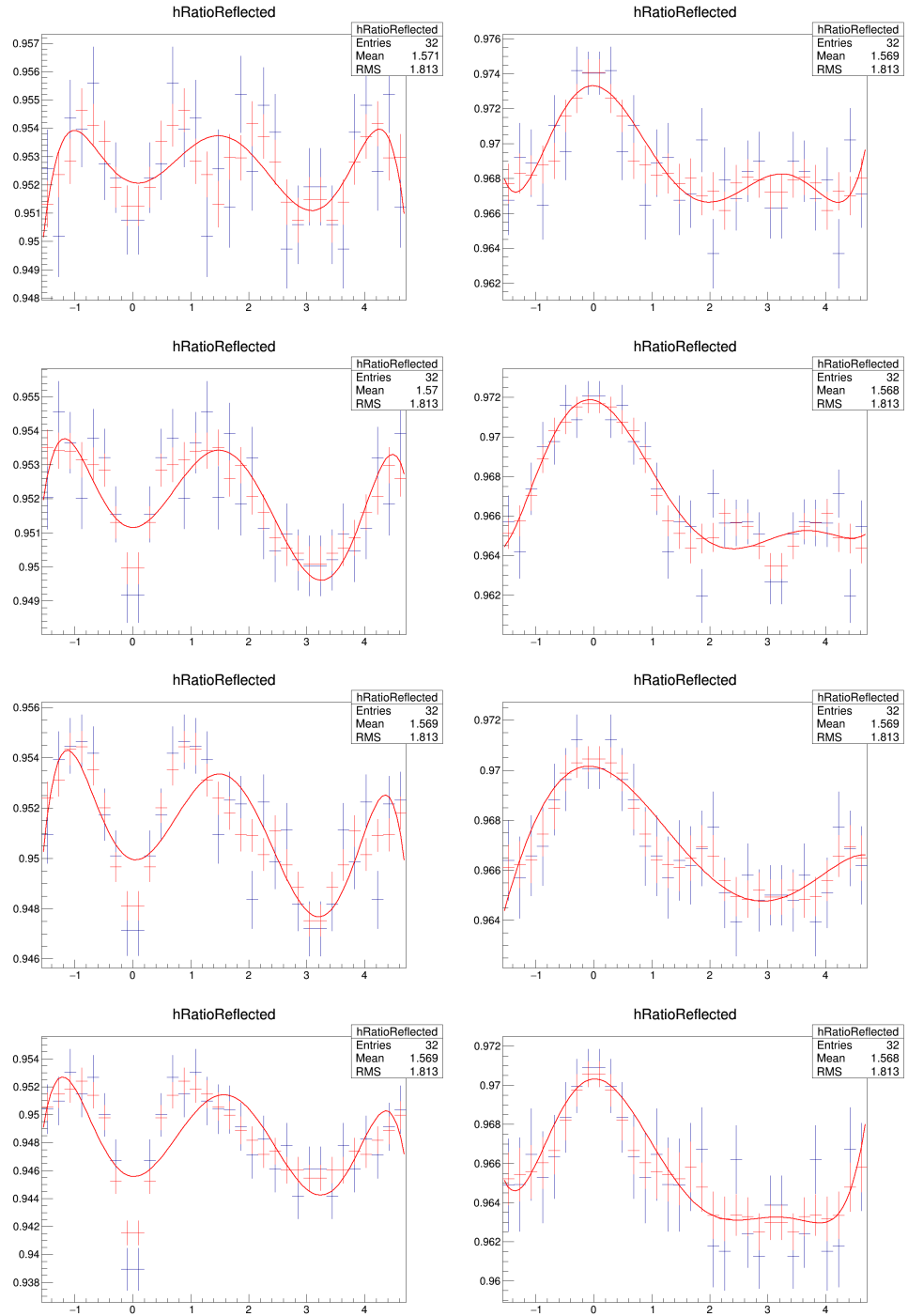
- $p_T(\text{assoc}) > 0.3 \text{ GeV}/c$  : 0.958
- $p_T(\text{assoc}) > 1 \text{ GeV}/c$  : 0.973
- $0.3 < p_T(\text{assoc}) < 1 \text{ GeV}/c$  : 0.953
- $1 < p_T(\text{assoc}) < 2 \text{ GeV}/c$  : 0.969
- $2 < p_T(\text{assoc}) < 3 \text{ GeV}/c$  : 0.982

It was also verified with the same Monte Carlo study that applying the DCA selection rejects less than 0.2% primary tracks (tagged as false positives) from the associated track sample, again with a flat azimuthal distribution, inducing hence a fully negligible bias on the data correlation distributions. This is shown in Figure 13. This was also verified for specific charm-origin and beauty-origin tracks, due to their larger DCA with respect to primary tracks from light quarks. In this case, the fraction of rejected charm and beauty tracks stays below 1% in all the kinematic ranges apart from the associated track  $p_T$  regions 0.3-1 and  $> 0.3 \text{ GeV}/c$ , where the rejection can be as high as 2%. In these kinematic ranges, though, the data correlation distributions are dominated by non-heavy-flavour tracks, as it was verified from the simulations, hence the overall bias is still contained below 1%, thus negligible.

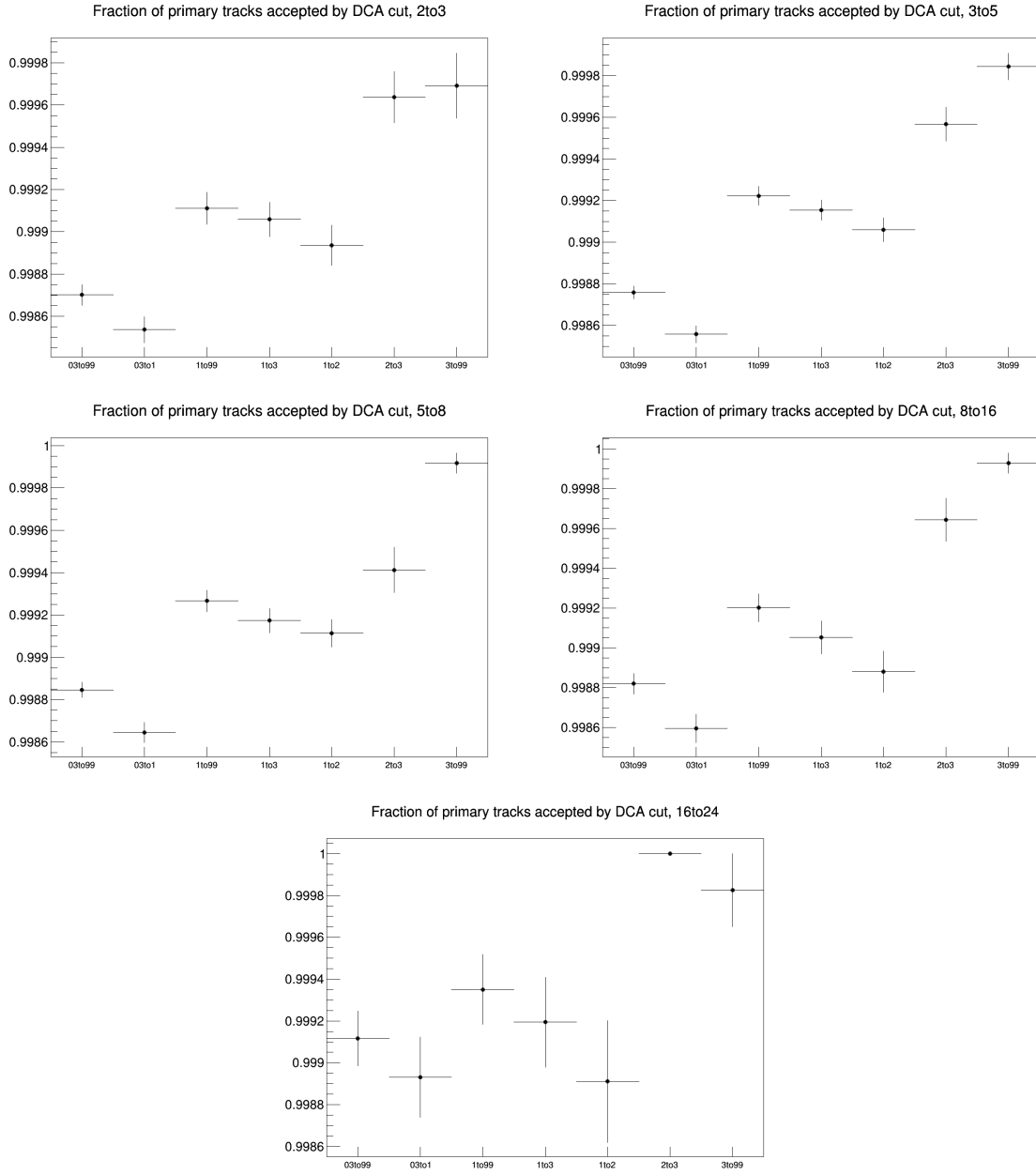
These studies were performed on an enriched Monte Carlo sample, which could not fully reproduce the relative abundancies of the species. Anyway, for events with a reconstructed D-meson, this bias is expected to be minor, and only these events are used in the data analysis. In any case, the percentages obtained from the study were found to be consistent within 1%.



**Figure 11:** Fraction of secondary tracks over total amount of tracks which pass the DCA selection. The four panel show the fractions for the D-meson  $p_T$  ranges: 2-3, 3-5, 5-8, 8-16, 16-24, respectively. Inside each panel, the associated track  $p_T$  ranges are shown on the x-axis.



**Figure 12:** Fraction of primary track in the reconstructed associated track sample (blue histogram). The polynomial fit function (red curve) and the 3-point moving average (red histogram) are also superimposed. The  $p_T(D)$  ranges are 2-3, 3-5, 5-8, 8-16 GeV/c, respectively for each row, and  $0.3 < p_T(assoc) < 1$ ,  $p_T(assoc) > 1$  GeV/c inside each row.



**Figure 13:** Fraction of primary tracks rejected by the DCA selection. The four panel show the fractions for the D-meson  $p_T$  ranges: 2-3, 3-5, 5-8, 8-16, 16-24, respectively. Inside each panel, the associated track  $p_T$  ranges are shown on the x-axis.



### 3.3.5 Beauty feed-down

The contribution of correlations of D meson from b-hadron decay is subtracted from the data correlation distributions as:

$$\tilde{C}_{\text{prompt D}}(\Delta\phi) = \frac{1}{f_{\text{prompt}}} \left( \tilde{C}_{\text{inclusive}}(\Delta\phi) - (1 - f_{\text{prompt}}) \tilde{C}_{\text{feed-down}}^{\text{MC templ}}(\Delta\phi) \right). \quad (4)$$

In the above equation,  $\tilde{C}_{\text{inclusive}}(\Delta\phi)$  and  $\tilde{C}_{\text{prompt D}}(\Delta\phi)$  are per-trigger azimuthal correlation distributions before and after feed-down contribution subtraction,  $f_{\text{prompt}}$  is the fraction of prompt D meson and  $\tilde{C}_{\text{feed-down}}^{\text{MC templ}}$  is a template of the azimuthal correlation distribution for the feed-down component obtained from home-made Monte Carlo simulation at generated level, using PYTHIA6 with Perugia2011 tune. In order to avoid biases related to the different event multiplicity in real and simulated events, the correlation distribution was shifted to have its minimum coinciding with the baseline of the data azimuthal-correlation distribution before feed-down subtraction.

The value of  $f_{\text{prompt}}$  (Figure ??), which depends on D-meson species and varies as a function of the  $p_T$ , is estimated on the basis of FONLL predictions for the production of feed-down D mesons at central rapidity, in pp collisions at  $\sqrt{s} = 5$  TeV, and using the reconstruction efficiency of prompt and feed-down D mesons, following the so-called  $N_b$  approach defined in [1]. Typical values are about 8-10% for the  $D^0$ , about 4-7% for the  $D^+$  and about 5-8% for the  $D^{*+}$ .

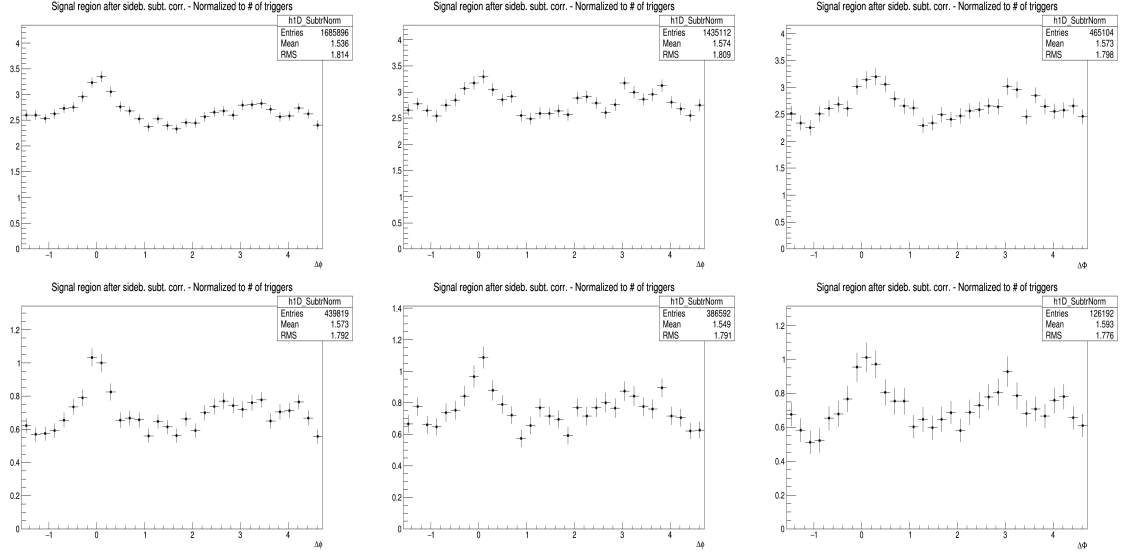
#### 422 **4 Systematic uncertainties on $\Delta\phi$ correlation distributions**

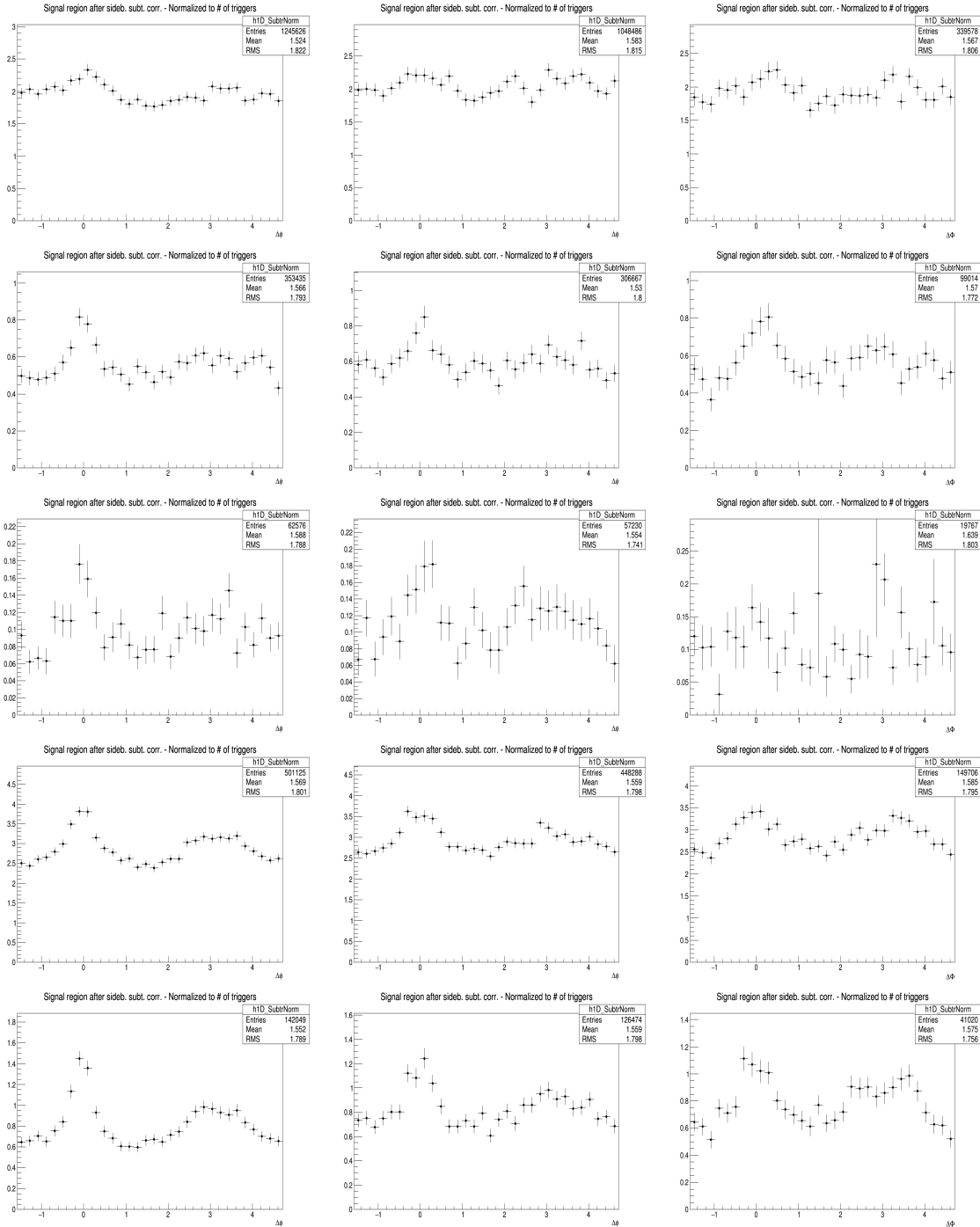
423 TO BE UPDATED

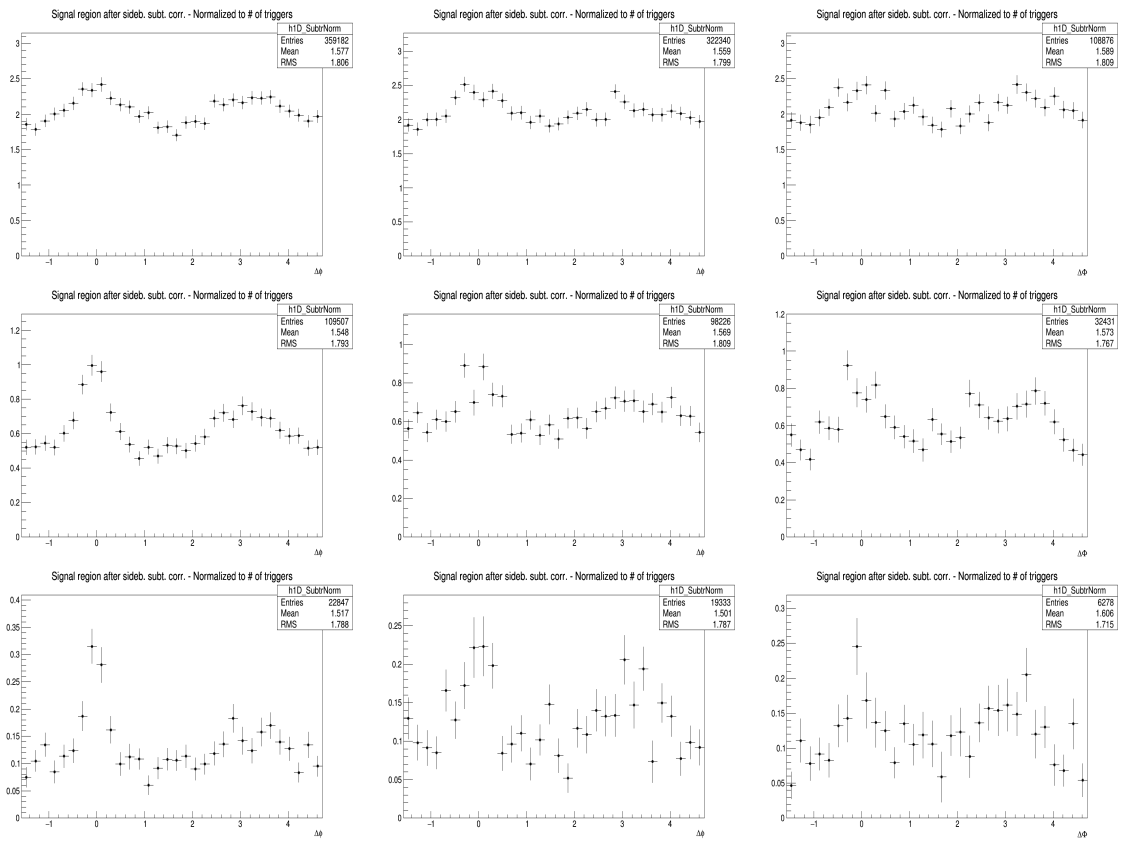
## 5 Results

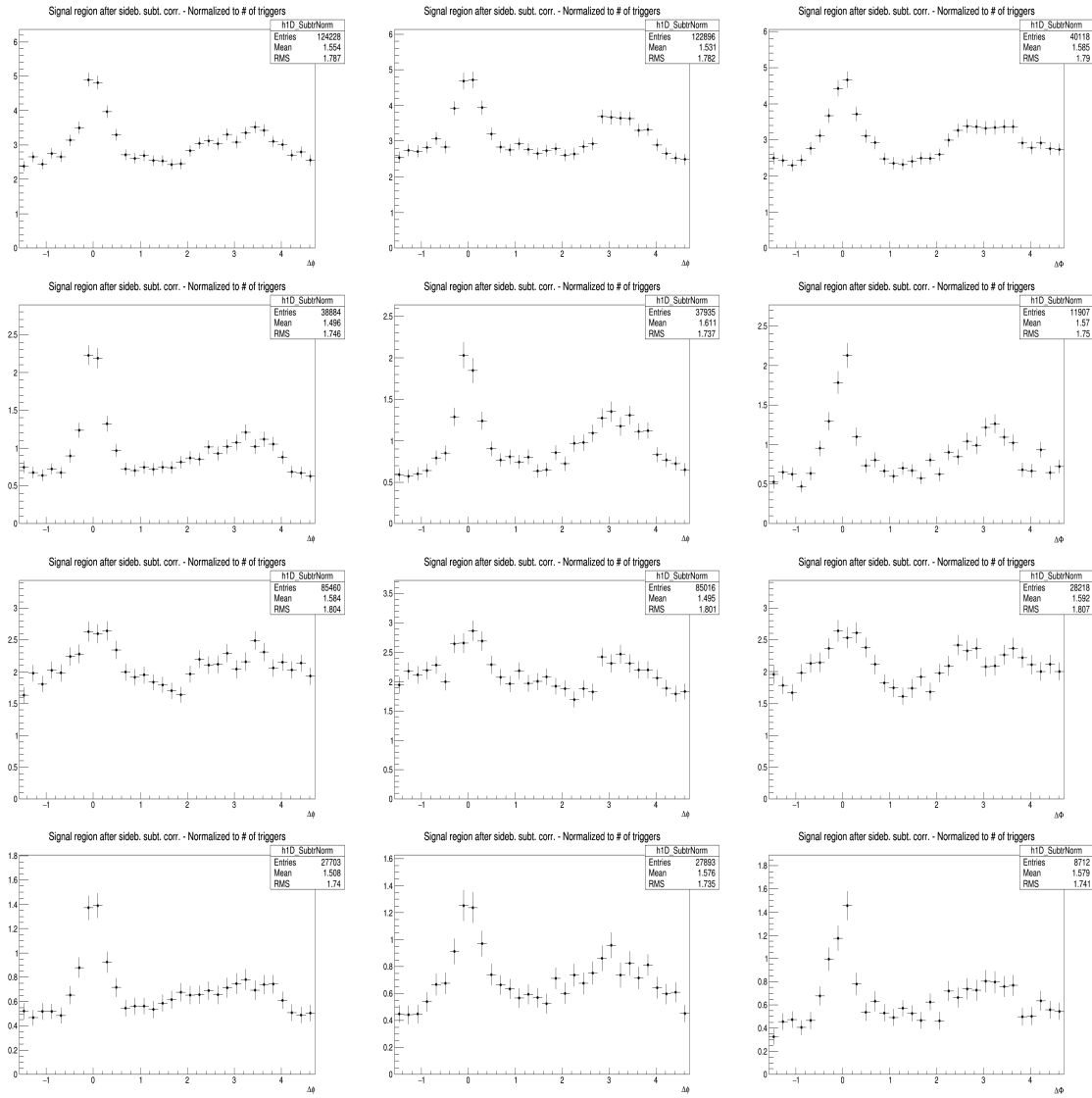
### 5.1 Comparing the three D meson correlation distributions

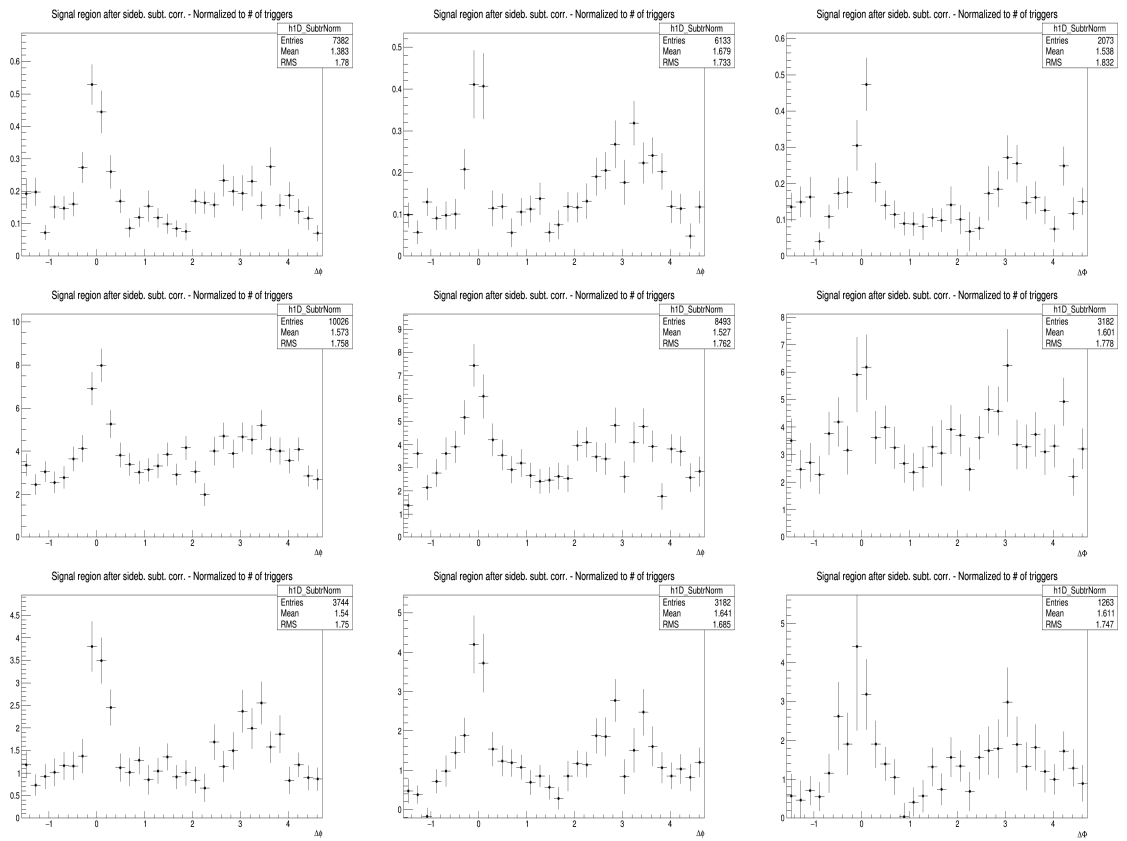
To check the compatibility of three D meson analyses, Figure 14 shows the corrected azimuthal correlation distributions (except for the feed-down subtraction and the secondary contamination removal) for  $D^0$ -h,  $D^{*+}$ -h and  $D^+$ -h, in each column, on the data sample used in the analysis. Results are shown for  $2 < D p_T < 3$  GeV/c,  $3 < D p_T < 5$  GeV/c,  $5 < D p_T < 8$  GeV/c,  $8 < D p_T < 16$  GeV/c and  $16 < D p_T < 24$  GeV/c with associated tracks  $p_T > 0.3$ ,  $p_T > 1$ ,  $0.3 < p_T < 1$  GeV/c,  $1 < p_T < 2$  GeV/c,  $2 < p_T < 3$  GeV/c and  $p_T > 3$  GeV/c.

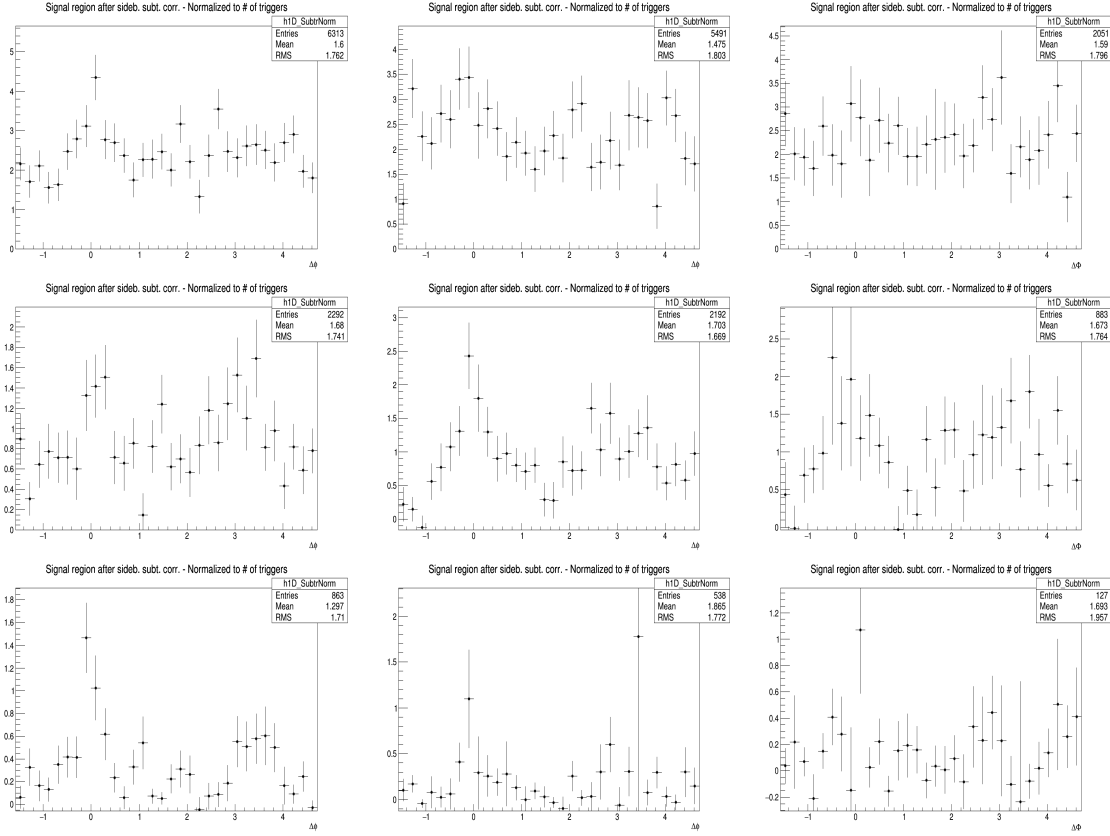












**Figure 14:** Corrected distribution of D-hadrons azimuthal correlations for the three species (apart from feed-down and purity), from analysis on the data sample, for the analyzed D-meson (**Column-Left:**  $D^0$ , **Column-Middle:**  $D^+$  and **Column-Right:**  $D^{*+}$ ) and different associated tracks  $p_T$  ranges (**Row 1-5:**  $3 < Dp_T < 5$  GeV/c,  $p_T$  (Assoc)  $> 0.3, > 1.0, 0.3-1.0, 1.0-2.0$  and  $2.0-3.0$  GeV/c respectively), (**Row 6-10:**  $5 < Dp_T < 8$  GeV/c,  $p_T$  (Assoc)  $> 0.3, > 1.0, 0.3-1.0, 1.0-2.0$  and  $2.0-3.0$  GeV/c respectively), (**Row 11-25:**  $8 < Dp_T < 16$  GeV/c,  $p_T$  (Assoc)  $> 0.3, > 1.0, 0.3-1.0, 1.0-2.0$  and  $2.0-3.0$  GeV/c respectively) and (**Row 16-20:**  $16 < Dp_T < 24$  GeV/c,  $p_T$  (Assoc)  $> 0.3, > 1.0, 0.3-1.0, 1.0-2.0$  and  $2.0-3.0$  GeV/c respectively)



432 An agreement of the distributions from the three mesons within the uncertainties is found in all the  
433 kinematic ranges.

434 Despite being evaluated in the full  $2\pi$  range, the range of final results was then reduced to  $[0, \pi]$  radians,  
435 reflecting the points outside that range over the value of 0. This allowed to reduce the impact of statistical  
436 fluctuations on the data points (supposing equal statistics for a pair of symmetric bins, after the reflection  
437 the relative statistical uncertainty for the resulting bin is reduced by a factor  $1/\sqrt{2}$ ).

## 438 5.2 Average of $D^0$ , $D^+$ and $D^{*+}$ results

439 Given the compatibility within the uncertainties among the  $D^0$ ,  $D^+$  and  $D^{*+}$  azimuthal correlations, and  
440 since no large differences are visible in the correlation distributions observed in Monte Carlo simulations  
441 based on Pythia with Perugia0, 2010 and 2011 tunes<sup>1</sup>, it was possible to perform a weighted average  
442 (eq. 5) of the azimuthal correlation distributions of  $D^0$ ,  $D^+$  and  $D^{*+}$ , in order to reduce the overall  
443 uncertainties. Although some correlation between the mesons could be present (about the 30% of the  
444  $D^0$ , and also part of the  $D^+$ , come from  $D^{*+}$  decays), the three selected D-meson samples can be treated  
445 as uncorrelated. The sum of the statistical uncertainties; the systematics uncertainty on S and B extraction  
446 and on background shape, are added in quadrature and the inverse of this sum was used as weight,  $w_i$ .

$$\left\langle \frac{1}{N_D} \frac{dN^{\text{assoc}}}{dp_T} \right\rangle_{D\text{mesons}} = \frac{\sum_{i=\text{meson}} w_i \frac{1}{N_D} \frac{dN_i^{\text{assoc}}}{d\Delta\phi}}{\sum_{i=\text{meson}} w_i}, w_i = \frac{1}{\sigma_{i,\text{stat}}^2 + \sigma_{i,\text{uncorr.syst.}}^2} \quad (5)$$

447 The statistical uncertainty and the uncertainties on S and B extraction and on background shape (those  
448 used for the weights) on the average were then recalculated using the following formula:

$$\sigma^2 = \frac{1}{n_D} \frac{\sum_{i=\text{meson}} w_i \sigma_i^2}{\sum_{i=\text{meson}} w_i} \quad (6)$$

449 where  $n_D$  is the number of mesons considered in the average. It can be observed that for  $\sigma_i^2 = 1/w_i$  the  
450 formula coincides with the standard one giving the uncertainty on a weighted average. The contribution  
451 to the average systematic uncertainty for those uncertainty sources not included in the weight definition,  
452 was evaluated via error propagation on the formula of the weighted average (5), resulting in equation  
453 (7) and (8) for sources considered uncorrelated and correlated among the mesons. In particular, the  
454 uncertainties on the associated track reconstruction efficiency, on the contamination from secondary, on  
455 the feed-down subtraction, and that resulting from the Monte Carlo closure test were considered fully  
456 correlated among the mesons, while those deriving from the yield extraction (included in the weight  
457 definition) and on the D meson reconstruction and selection efficiency were treated as uncorrelated.

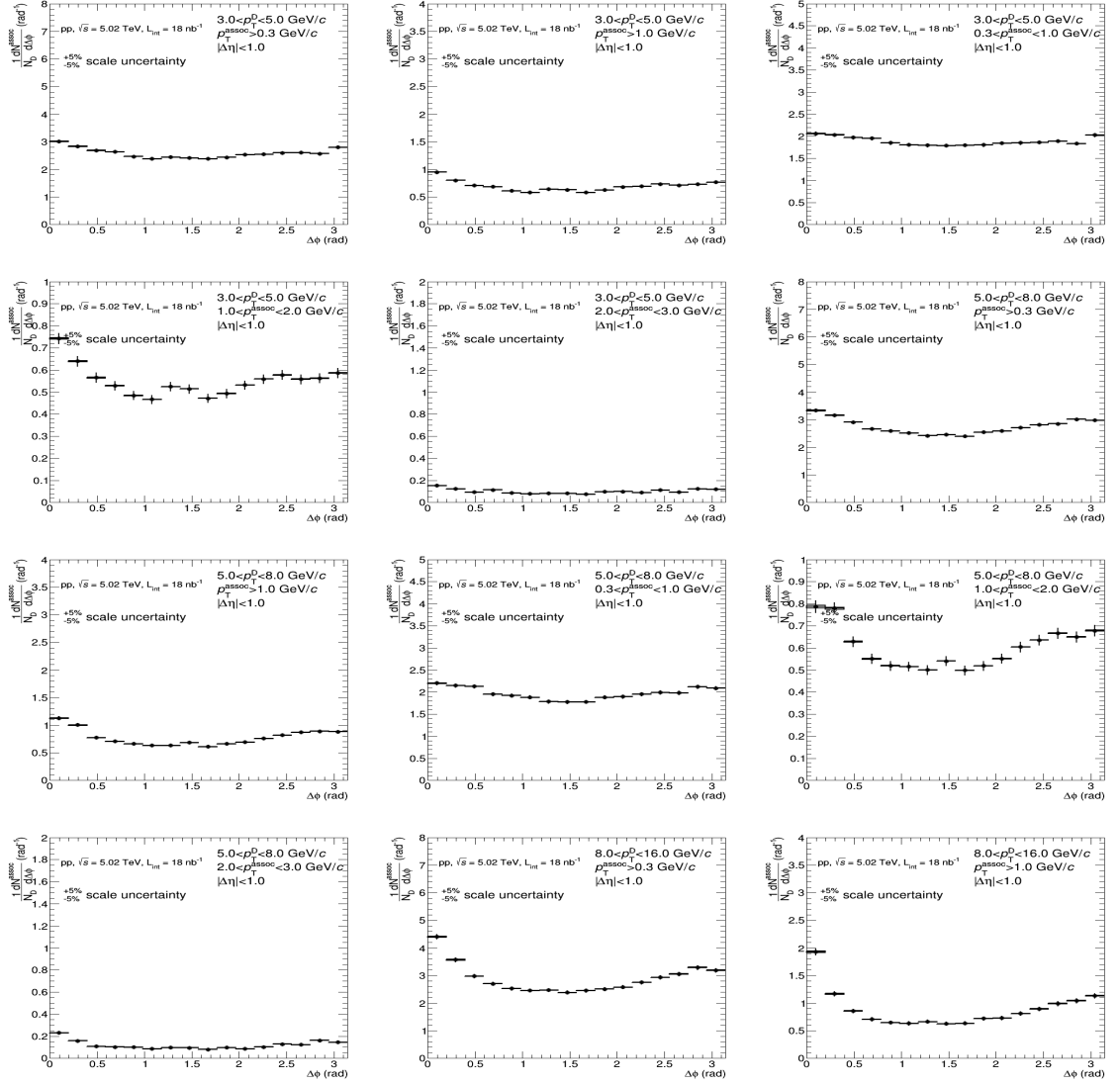
$$\sigma^2 = \frac{\sum_{i=\text{meson}} w_i^2 \sigma_i^2}{(\sum_{i=\text{meson}} w_i)^2} \quad (7)$$

$$\sigma = \frac{\sum_{i=\text{meson}} w_i \sigma_i}{\sum_{i=\text{meson}} w_i} \quad (8)$$

458 Figure 15 shows the averages of the azimuthal correlation distributions of  $D^0$ ,  $D^+$  and  $D^{*+}$  and charged  
459 particles with  $p_T > 0.3$  GeV/c,  $0.3 < p_T < 1$  GeV/c,  $p_T > 1$  GeV/c,  $1 < p_T < 2$  GeV/c,  $2 < p_T <$   
460  $3$  GeV/c in the D meson  $p_T$  ranges  $3 < p_T < 5$  GeV/c,  $5 < p_T < 8$  GeV/c,  $8 < p_T < 16$  GeV/c and  
461  $16 < p_T < 24$  GeV/c. As expected, a rising trend of the height of the near-side peak with increasing D-  
462 meson  $p_T$  is observed, together with a decrease of the baseline level with increasing  $p_T$  of the associated

<sup>1</sup> A slight near side hierarchy is present among the three meson results, with  $D^{*+}$  meson having a lower peak amplitude than  $D^0$  and  $D^+$ . It was verified that this is induced by the presence of  $D^0$  and  $D^+$  mesons coming from  $D^{*+}$ , the latter having on average a larger  $p_T$  and coming, hence, on average, from a larger  $p_T$  quark parton, which fragments in slightly more tracks in the near-side.

463 tracks. To further increase the statistical precision on the averaged correlation distributions, given the  
464 symmetry around 0 on the azimuthal axis, the distributions were reflected and shown in the range  $[0, \pi]$ .  
465 This reduces the statistical uncertainty on the points by, approximately, a factor of  $1/\sqrt{2}$ .



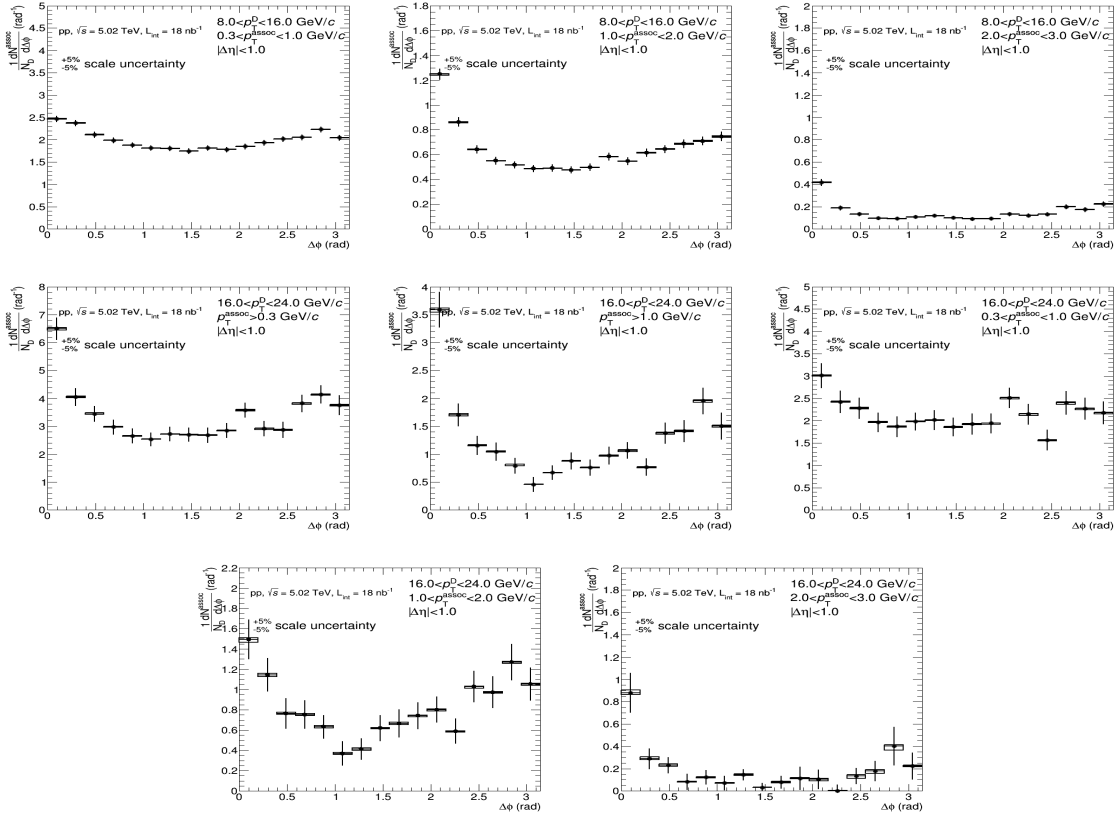
466 The usage of weighted average requires, as an underlying assumption, identical results expected for  
 467 different species (or, at least, compatible within the uncertainties). Anyway, it was also verified that the  
 468 usage of the arithmetic average instead of the weighted average increases the uncertainties on the points,  
 469 but produces a negligible shift of their central values.

### 470 5.3 Fit observable $p_T$ trends and uncertainties

471 In order to extract quantitative and physical information from the data correlation patterns, the averaged  
 472 D-h correlation distributions are fitted with two Gaussian functions (with means fixed at  $\Delta\phi=0$  and  $\Delta\phi=\pi$   
 473 values), plus a constant term (baseline). A periodicity condition is also applied to the fit function to obtain  
 474 the same value at the bounds of  $2\pi$  range. The expression of the fit function is reported below (equation  
 475 9):

$$f(\Delta\phi) = c + \frac{Y_{NS}}{\sqrt{2\pi}\sigma_{NS}} e^{-\frac{(\Delta\phi-\mu_{NS})^2}{2\sigma_{NS}^2}} + \frac{Y_{AS}}{\sqrt{2\pi}\sigma_{AS}} e^{-\frac{(\Delta\phi-\mu_{AS})^2}{2\sigma_{AS}^2}} \quad (9)$$

476 where baseline is calculated as the weighted average of the points lying in the so-called "transverse  
 477 region", i.e. the interval  $\frac{\pi}{4} < |\Delta\phi| < \frac{\pi}{2}$ .



**Figure 15:** Average of  $D^0$ ,  $D^+$  and  $D^{*+}$  azimuthal correlation distributions, in the D meson  $p_T$  ranges  $3 < p_T < 5$  GeV/c,  $5 < p_T < 8$  GeV/c,  $8 < p_T < 16$  GeV/c and  $16 < p_T < 24$  GeV/c, with associated tracks with  $p_T > 0.3$  GeV/c,  $p_T > 1$  GeV/c,  $0.3 < p_T < 1$  GeV/c,  $1 < p_T < 2$  GeV/c and  $2 < p_T < 3$  GeV/c

478 An example of the results from the fit is shown in Figure 16

479 From the fit outcome, it is possible to retrieve the near-side and away-side yield and widths (integral  
480 and sigma of the Gaussian functions, respectively), as well as the baseline height of the correlation  
481 distribution. The near-side observables give information on the multiplicity and angular spread of the  
482 tracks from the fragmentation of the charm jet which gave birth to the D-meson trigger. At first order,  
483 instead, the away-side observables are related to the hadronization of the charm parton produced in the  
484 opposite direction (though the presence of NLO processes for charm production breaks the full validity  
485 of this assumption). The baseline value is a rough indicator of the underlying event multiplicity, though  
486 below the baseline level also charm and beauty-related pairs are contained (especially in cases of NLO  
487 production for the heavy quarks).

488 The evaluation of the systematic uncertainties on the observables obtained from the fits is performed as  
489 follows **(TO BE UPDATED)**:

- 490 – **The fits are repeated by changing the range of the transverse region in which the baseline**  
491 **is evaluated. Alternate definitions of  $\frac{\pi}{4} < |\Delta\phi| < \frac{3\pi}{8}$ ,  $\frac{3\pi}{8} < |\Delta\phi| < \frac{\pi}{2}$  and  $\frac{\pi}{4} < |\Delta\phi| < \frac{5\pi}{8}$  are**  
492 **considered.**
- 493 – **In addition,  $\Delta\phi$  correlation points are shifted to the upper and lower bounds of their uncor-**  
494 **related systematic boxes, and refitted.**
- 495 – **The fits are also repeated by moving the baseline value from its default value (i.e. with the**  
496 **default transverse region) on top and on bottom of its statistic uncertainty. This helps to**  
497 **account, though in a systematic uncertainty, for the statistical uncertainty on the baseline**  
498 **position (since in the fit the baseline is constrained, and its error is not propagated to the**  
499 **other observables).**
- 500 – **The envelope between (i) the RMS of the relative variations of the parameters between the**  
501 **fit outcomes defined in the first two points, and (ii) the relative variations of the parameters**  
502 **from the fit outcomes defined in the third point, is considered as systematic uncertainty for**  
503 **the near-side and away-side widths.**
- 504 – **For the estimation of the baseline and of the near-side and away-side yields, instead, the**  
505 **previous value is added in quadrature with the  $\Delta\phi$ -correlated systematics in the correlation**  
506 **distributions, since these values are affected by a change in the global normalization of the**  
507 **distributions.**
- 508 – **In addition, for all the fit observables, an additional fit variation is performed assuming,**  
509 **instead of a flat baseline, a  $v_{2\Delta}$ -like modulation, with the following  $v_2$  values for the associated**  
510 **tracks (assuming  $v_{2\Delta} = v_2(h) \cdot v_2(D)$ ): 0.04 (0.3-1 GeV/c), 0.06 (>0.3 GeV/c), 0.08 (1-2 GeV/c),**  
511 **0.09 (>1 GeV/c, 2-3 GeV/c), 0.1 (>3 GeV/c), on the basis of ATLAS preliminary results for**  
512 **heavy-flavour muons at 8 TeV; for the D-meson triggers the following  $v_2$  values were instead**  
513 **assumed: 0.05 (3-5 GeV/c), 0.03 (5-8 GeV/c), 0.02 (8-24 GeV/c), on the basis of previous**  
514 **ALICE measurements in p-Pb collisions at 5 TeV [3]. The difference of the fit observables**  
515 **with respect to the standard fits is taken as uncertainty. Due to its peculiarity, this systematic**  
516 **uncertainty is summed in quadrature with the others to obtain the total uncertainty, but is**  
517 **also shown separately in the figures.**

$$\sigma^{syst} = \sqrt{(Max(\Delta par^{ped.mode}, \Delta par^{\Delta\phi point}))^2 + (\sigma_{Syst}^{corr})^2} \quad (10)$$

### 5.3.1 Results for near-side yield and width, away-side yield and width, and baseline

Figures ??, ??, ??, ?? and ?? show the near-side associated yield, width (the sigma of the Gaussian part of the fit functions), away-side associated yield, width and the height of the baseline, for the average correlation distributions, in the kinematic ranges studied in the analysis, together with their statistical and systematic uncertainties. For each kinematic range, the correspondent plot showing the systematic uncertainty of the considered observable from the variation of the fit procedure is reported as well (which is the full systematic uncertainty for the widths).

### 5.3.2 Comparisons $pp$ and $p\text{-Pb}$

Figure 17 shows the average of  $D^0$ ,  $D^+$  and  $D^{*+}$  azimuthal correlations for 2017  $pp$  and 2016  $p\text{Pb}$  for all the kinematic ranges of trigger and associated particles  $p_T$ . Compatibility within uncertainties between the two collision systems is found for all the common kinematic ranges analyzed. In Figure 19, the comparison of the observables extracted from the fits (near-side yield and width) is also presented.

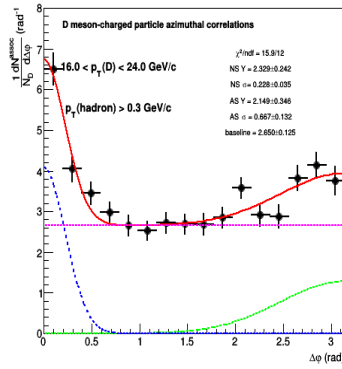
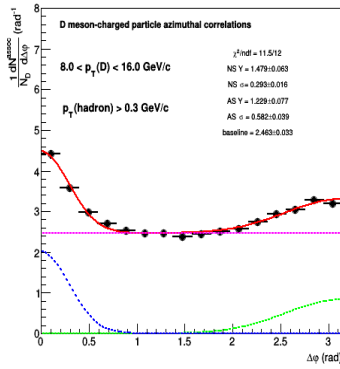
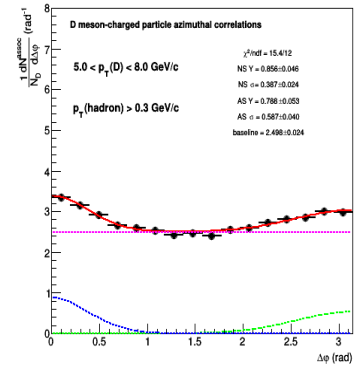
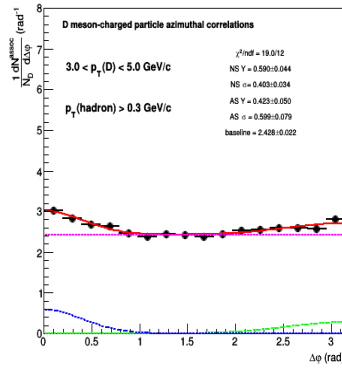
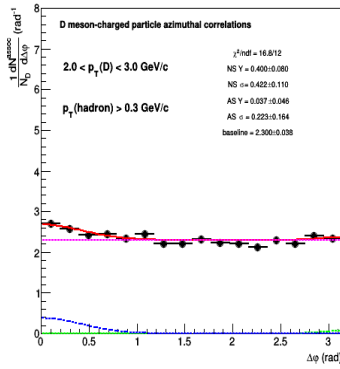
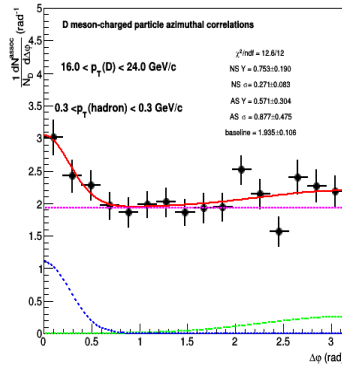
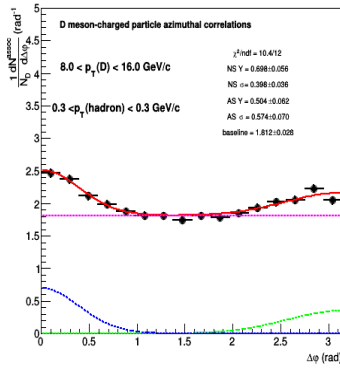
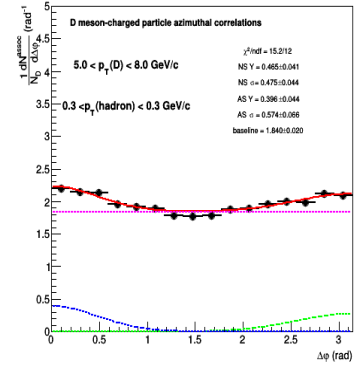
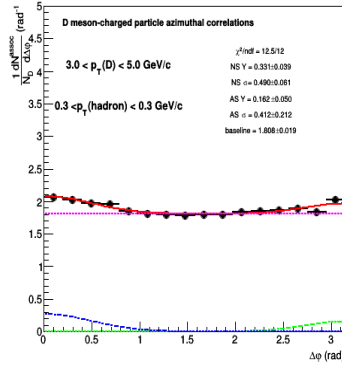
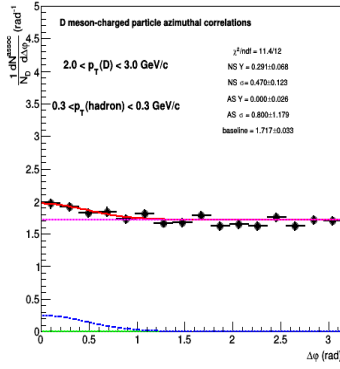
### 5.3.3 Comparisons $pp$ at 5, 7 and 13 TeV

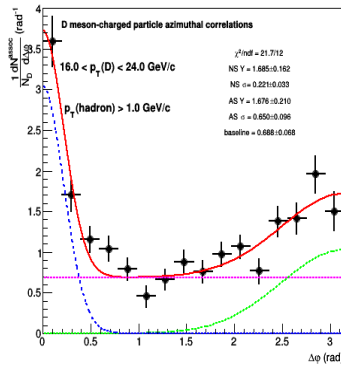
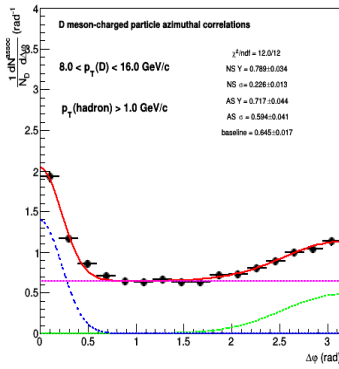
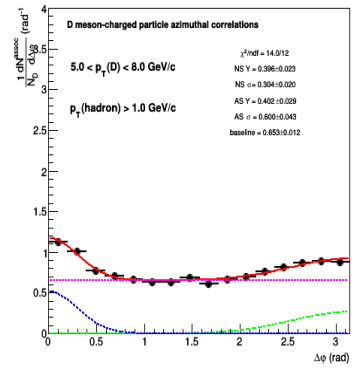
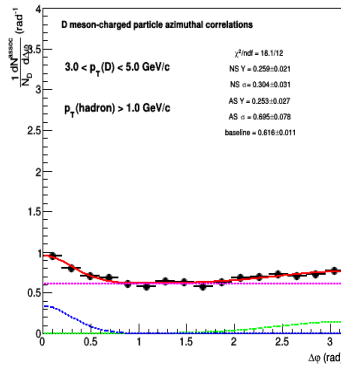
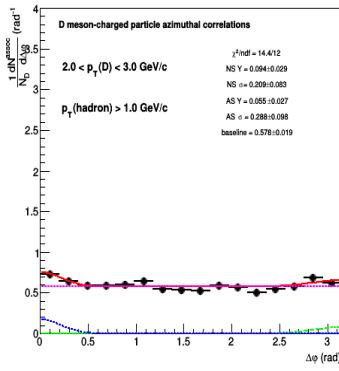
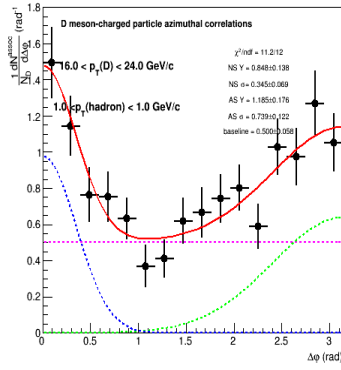
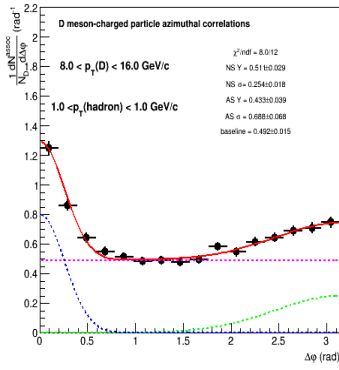
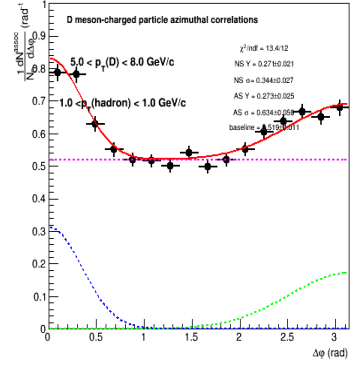
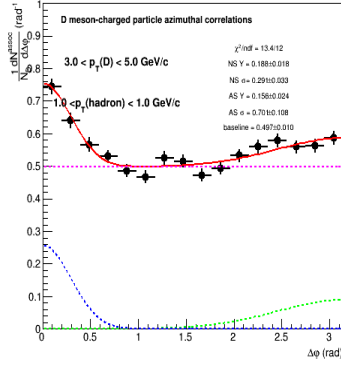
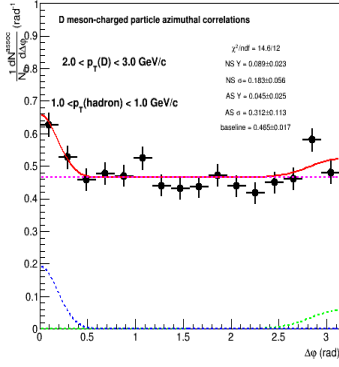
Figure 20 shows the average of  $D^0$ ,  $D^+$  and  $D^{*+}$  azimuthal correlations for  $pp$  at 5 TeV compared with  $pp$  at 7 TeV and 13 TeV for all the common kinematic ranges of trigger and associated particles  $p_T$  analysed. The data distribution of  $pp$  at 5 TeV have much better uncertainties than 7 TeV and also quite better than 13 TeV. Compatibility within uncertainties between the three energy systems is found for all the common kinematic ranges analyzed. In Figure 21 and 22, the comparison of the observables extracted from the fits (near-side yield and width) is also presented.

### 5.3.4 Comparisons $pp@5\text{TeV}$ and *Pythia*

Figure 17 shows the average of  $D^0$ ,  $D^+$  and  $D^{*+}$  azimuthal correlations for  $pp$  for several ranges of trigger and associated  $p_T$ , compared to different *Pythia6* tunes (Perugia 0, 2010, 2011), *Pythia 8* (tune 4C) and POWHEG+PYTHIA at the same collision energy. A substantial agreement in the overall momentum evolution of the correlation pattern is observed within uncertainties. In Figure ?? and ?? the comparison of the extracted physical observables (near-side and away-side yield, width and baseline height).

Figure ?? shows the comparison of the observables extracted from the fit to the correlation distributions, i.e. baseline, near-side yield and near-side width, from data and simulations. Compatibility within the uncertainties is obtained for the three fit parameters.







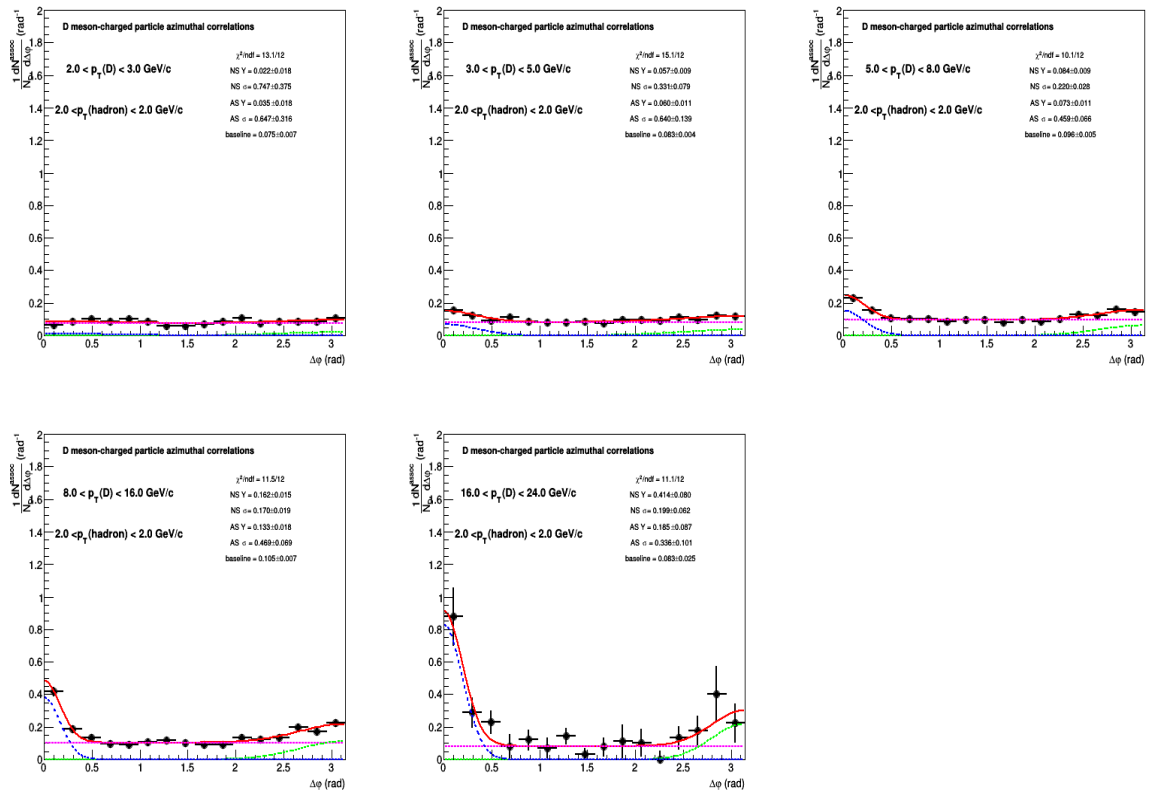
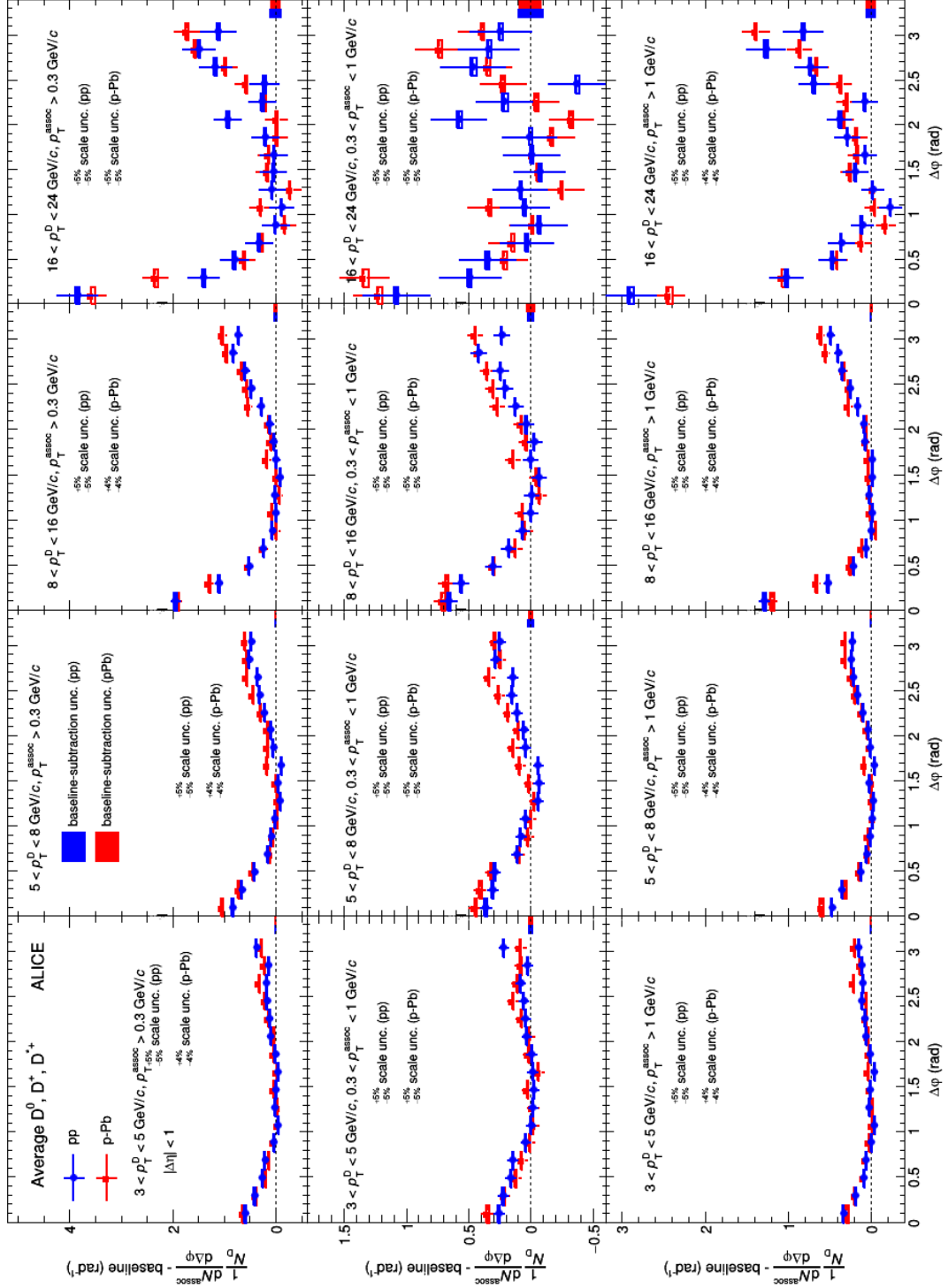
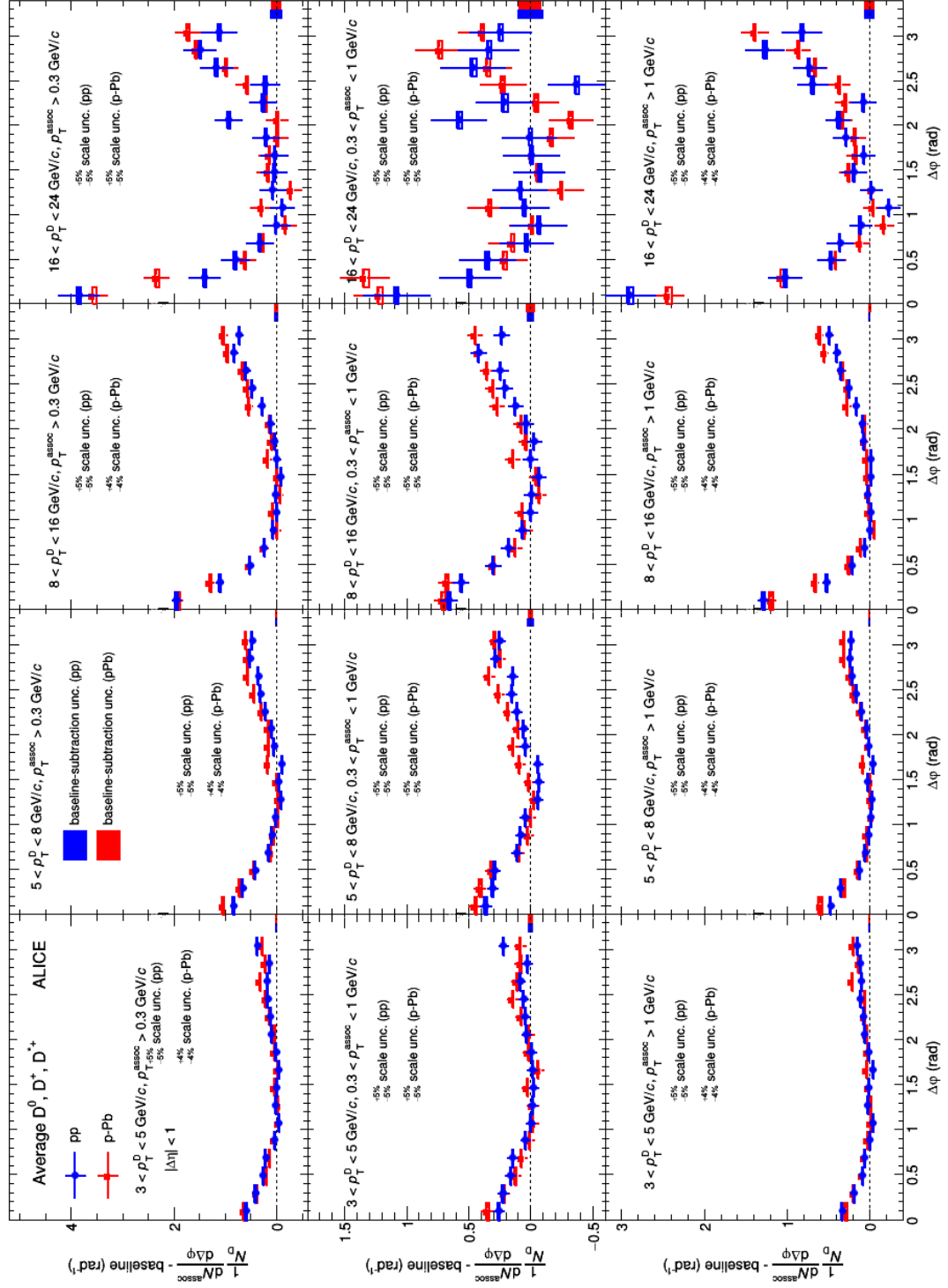
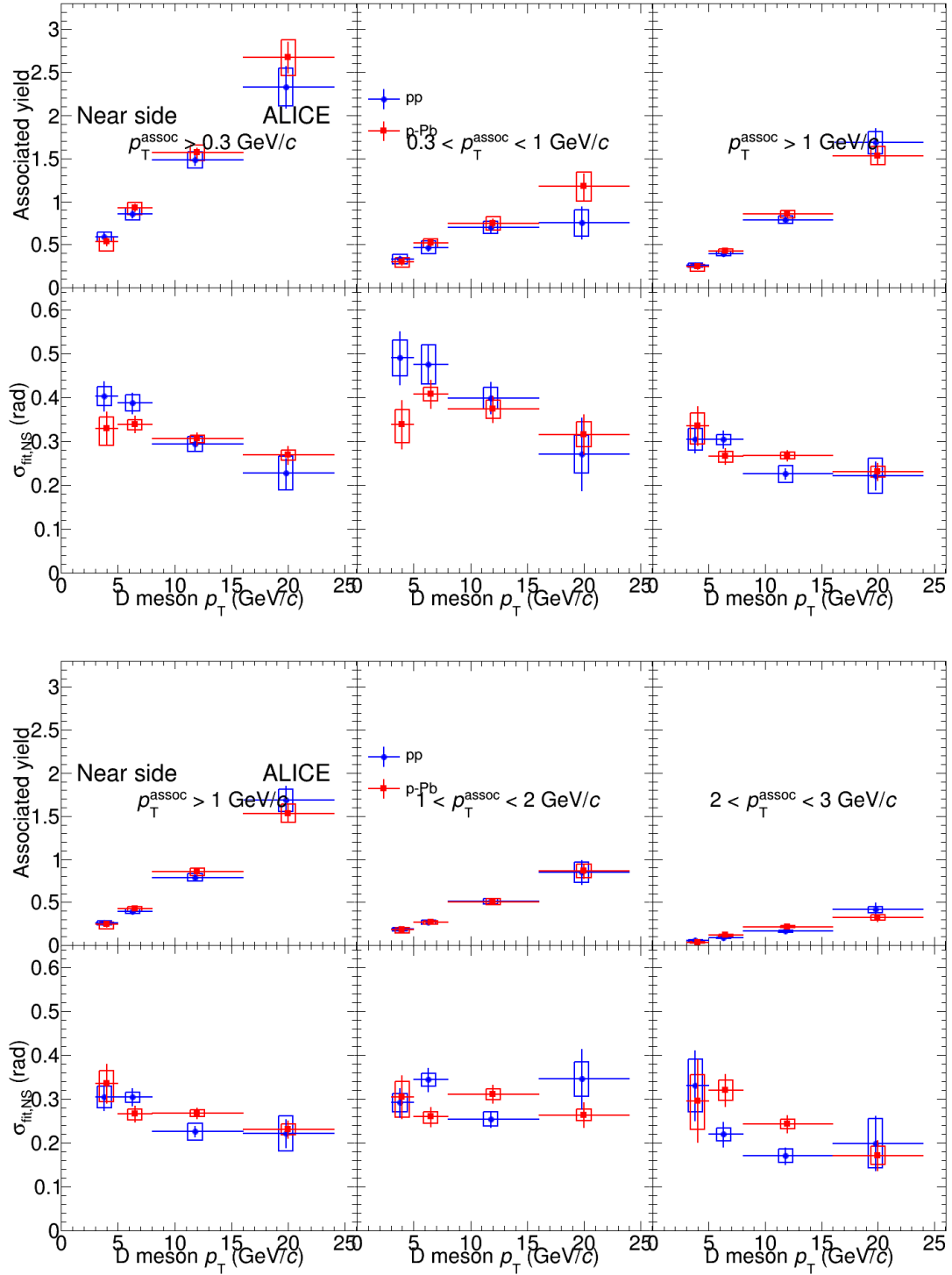


Figure 16: Example of fit to azimuthal correlation distributions and baseline estimation.

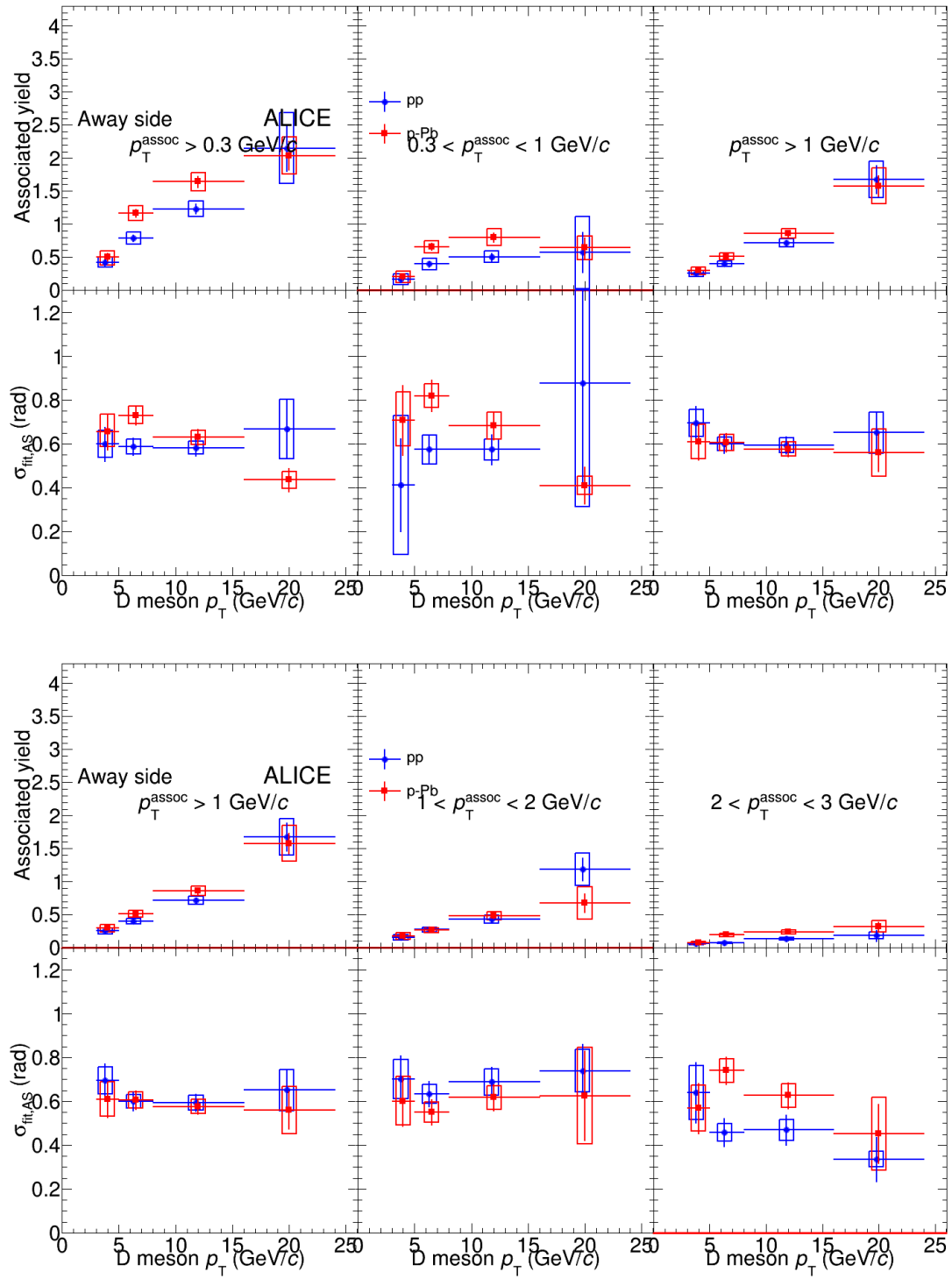




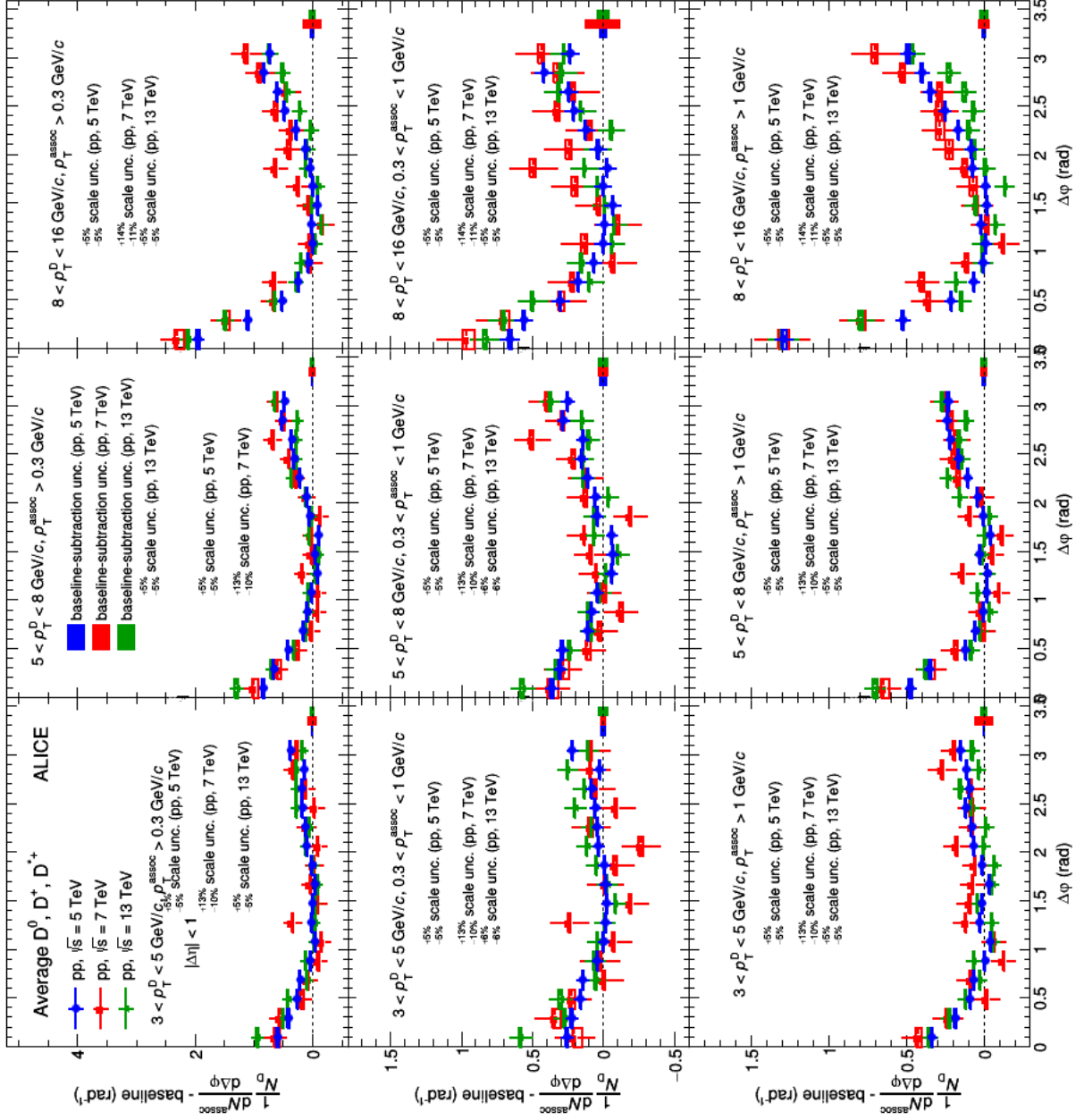
**Figure 17:** Average of  $D^0$ ,  $D^+$  and  $D^{*+}$  azimuthal correlations in pp (blue) and p-Pb (red) in all the kinematic ranges of trigger and associated particles.



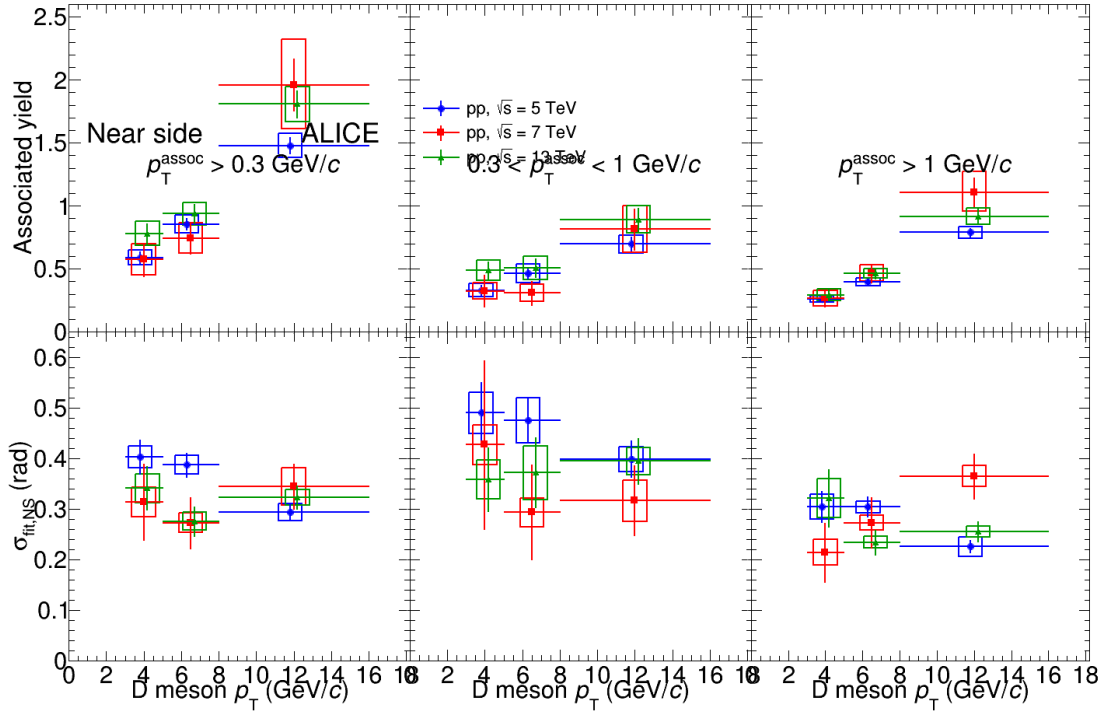
**Figure 18:** Near-side yield and sigmas for the average of  $D^0$ ,  $D^+$  and  $D^{*+}$  azimuthal correlations in pp (red) and p-Pb (black) in all the kinematic regions of trigger and associated track.



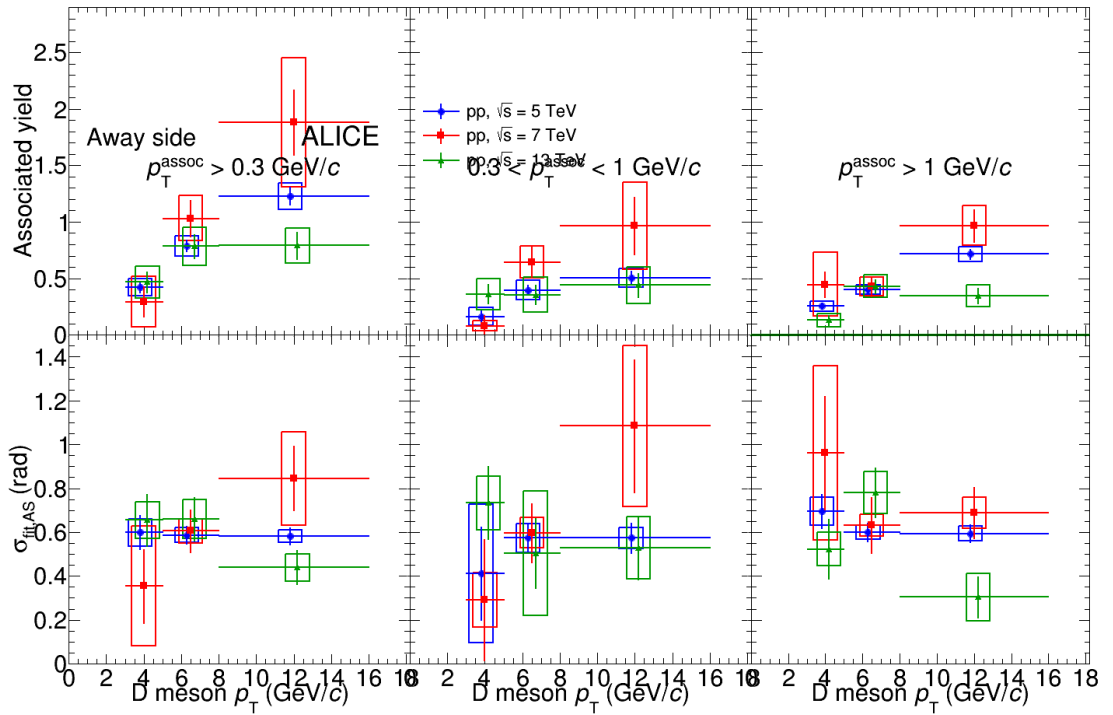
**Figure 19:** Away-side yield and sigmas for the average of  $D^0$ ,  $D^+$  and  $D^{*+}$  azimuthal correlations in pp (red) and p-Pb (black) in all the kinematic regions of trigger and associated track.



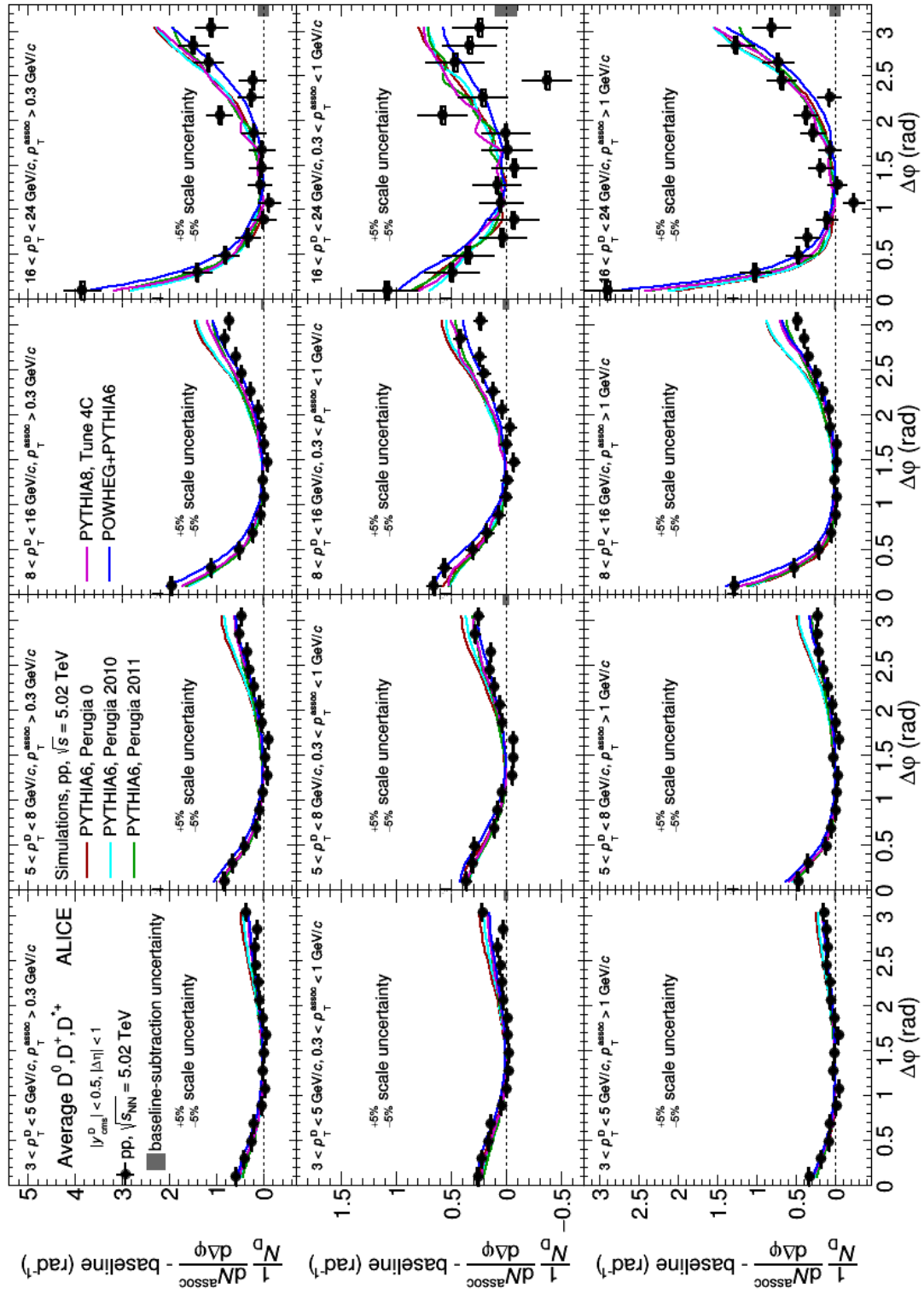
**Figure 20:** Average of  $D^0$ ,  $D^+$  and  $D^{*+}$  azimuthal correlations in pp at 5 (blue), 7 (red) and 13 (green) TeV in all the common kinematic ranges of trigger and associated particles.



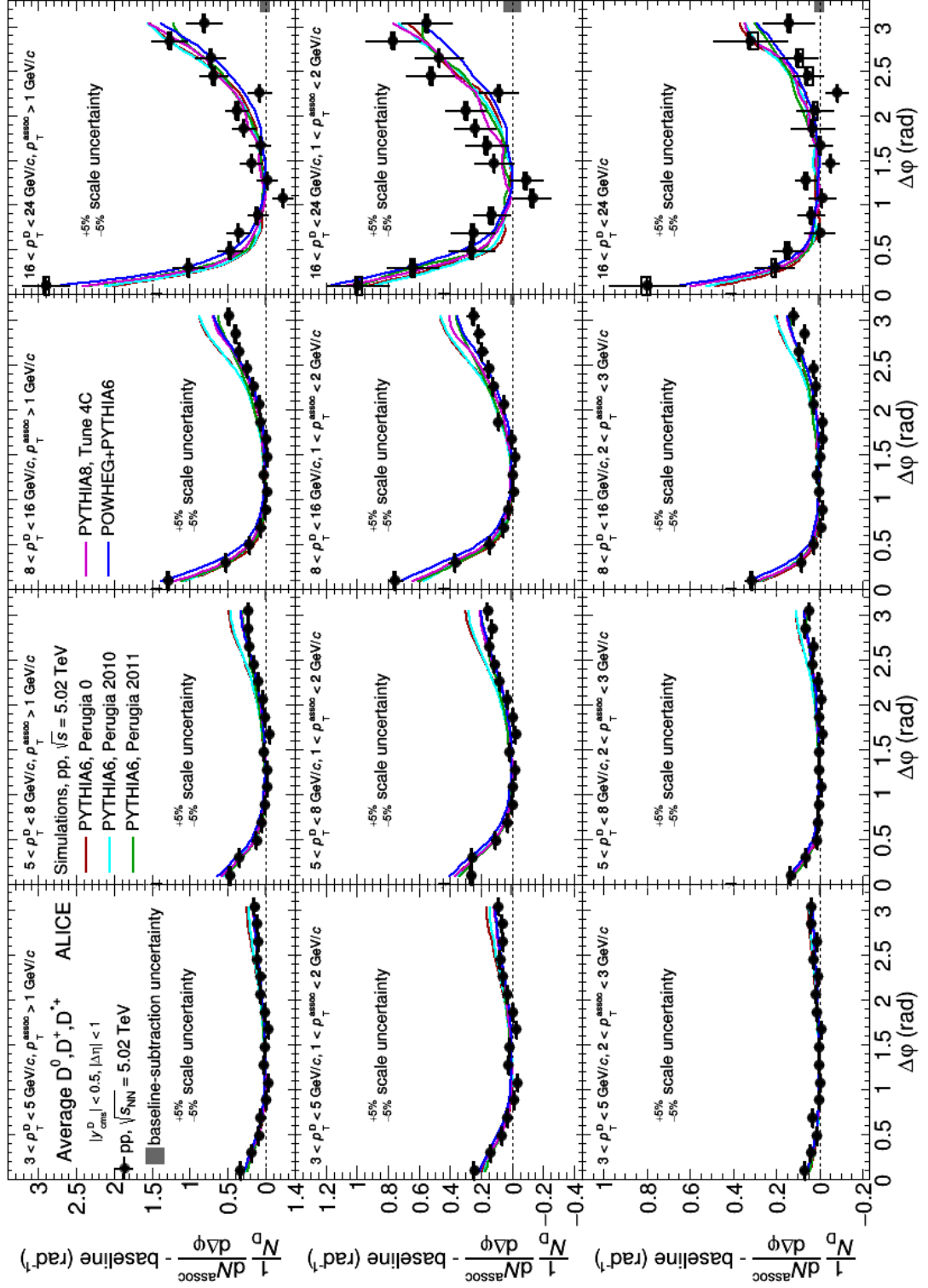
**Figure 21:** Near-side yield and width for the average of  $D^0$ ,  $D^+$  and  $D^{*+}$  azimuthal correlations in pp at 5 (blue), 7 (red) and 13 (green) TeV in all the common kinematic regions of trigger and associated track.



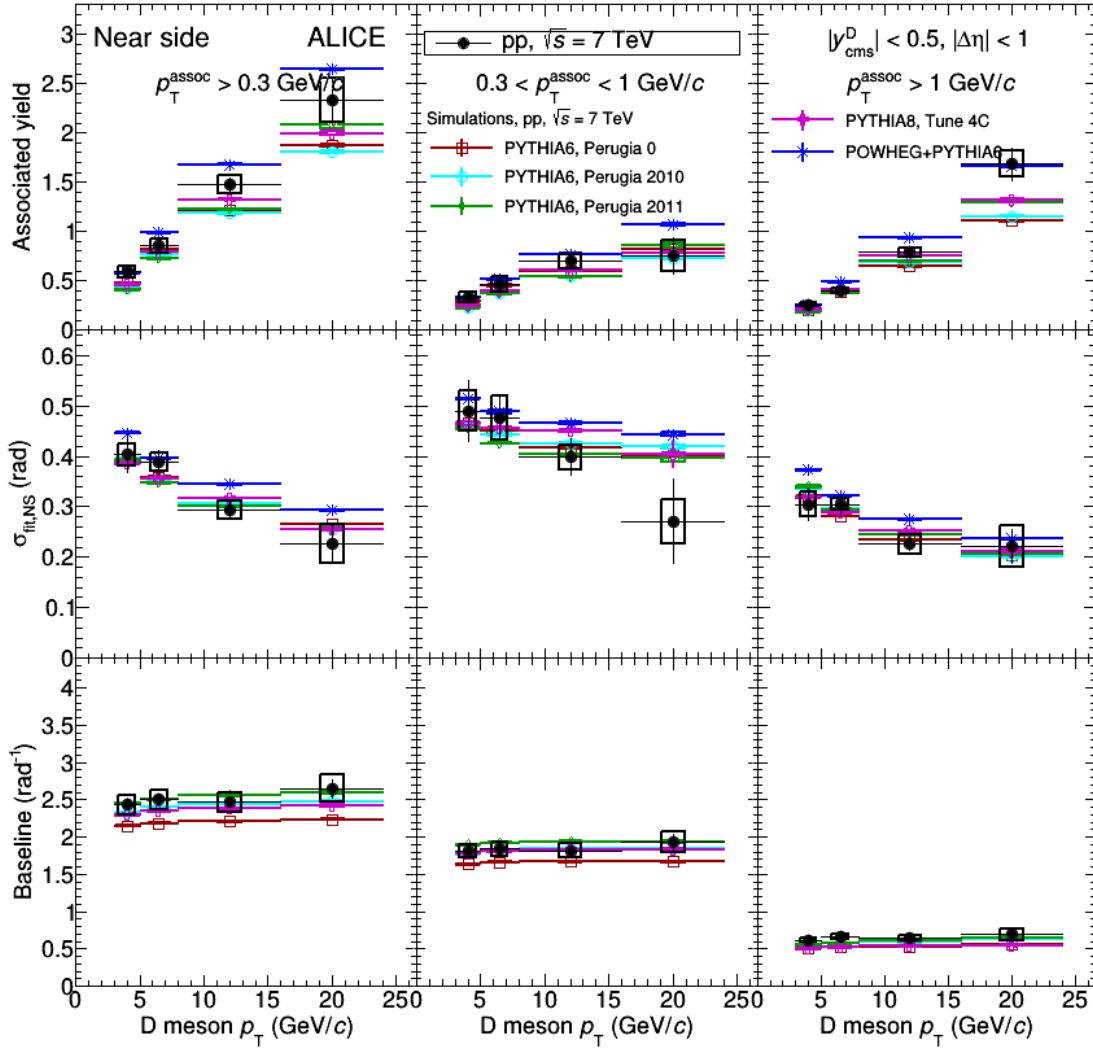
**Figure 22:** Away-side yield and width for the average of  $D^0$ ,  $D^+$  and  $D^{*+}$  azimuthal correlations in pp at 5 (blue), 7 (red) and 13 (green) TeV in all the common kinematic regions of trigger and associated track.

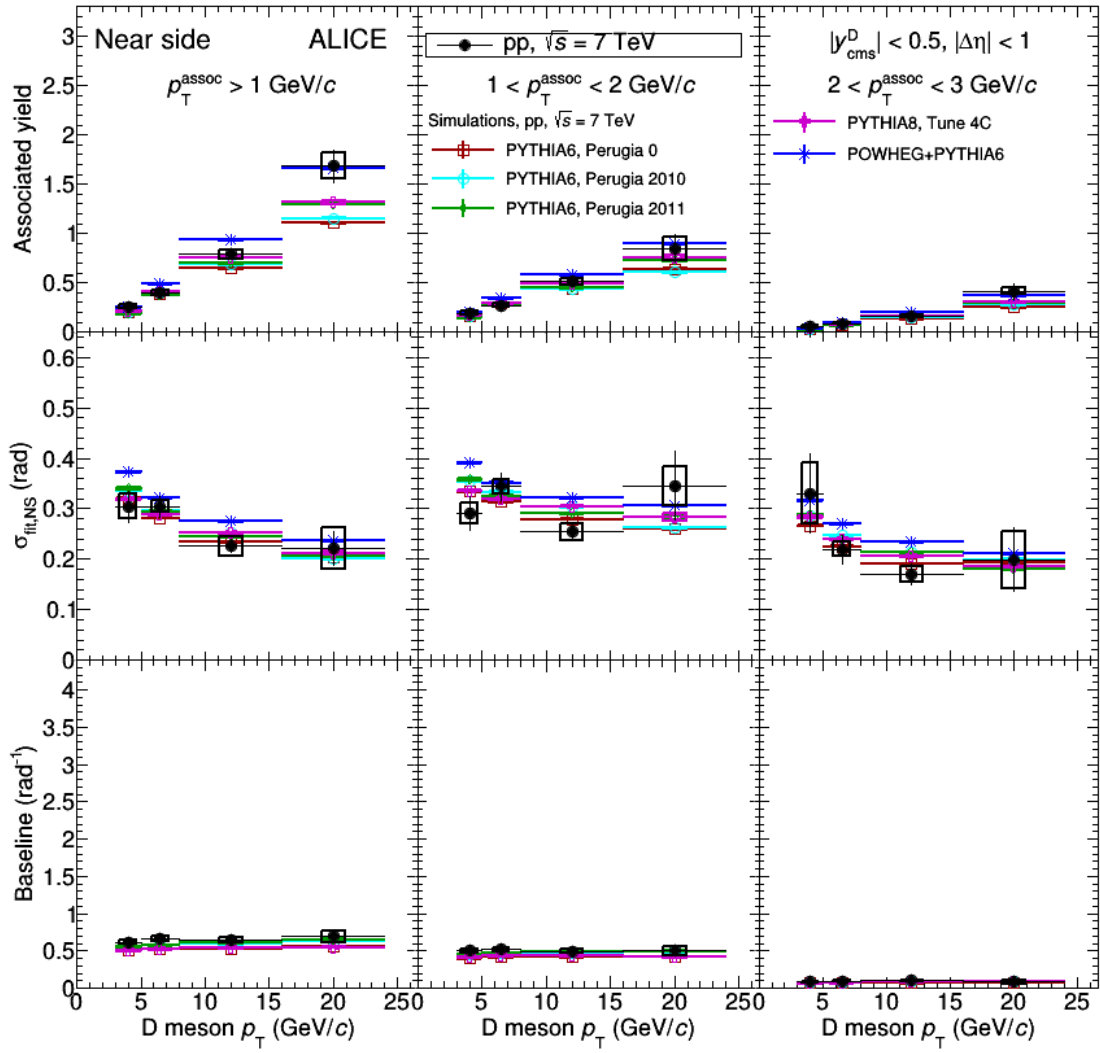




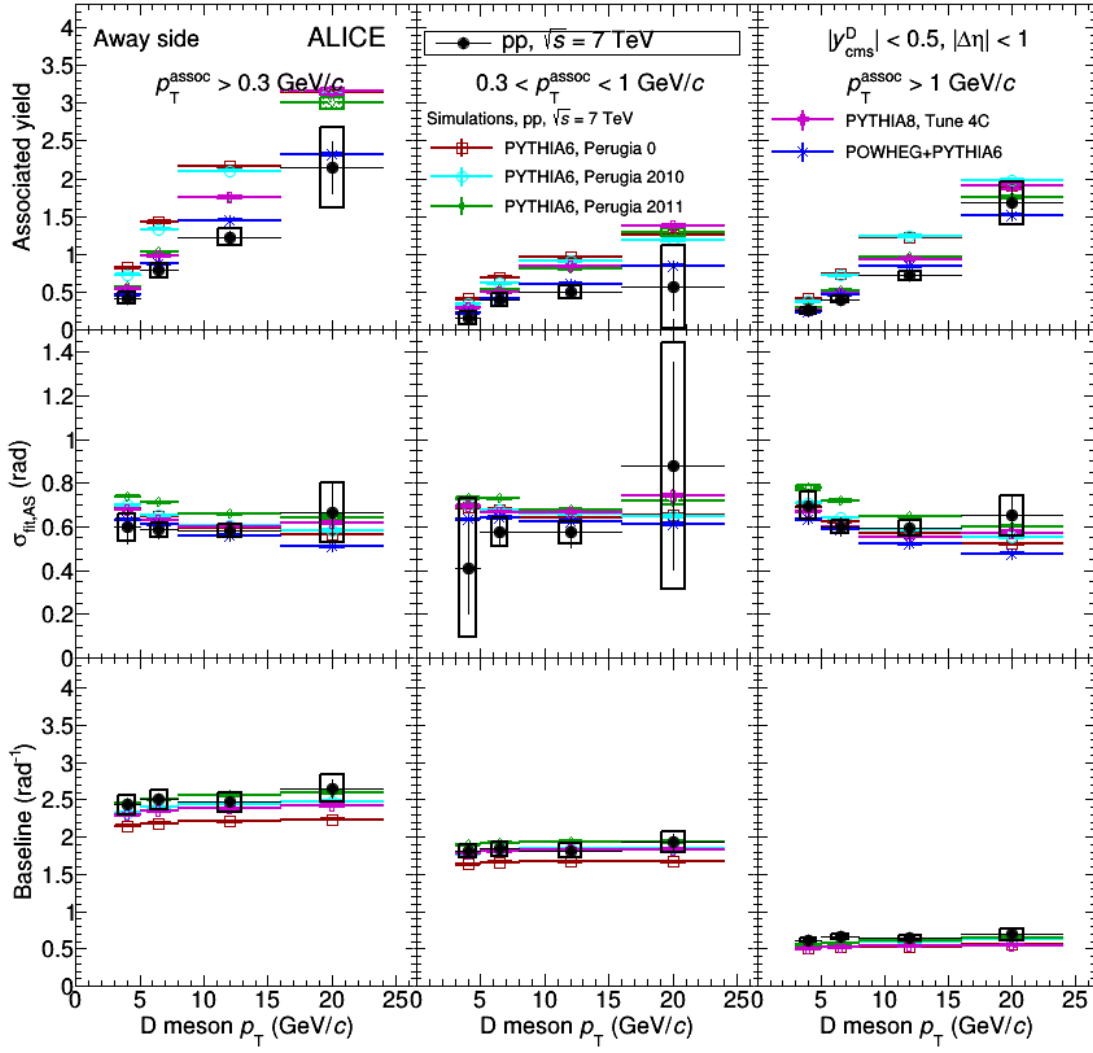


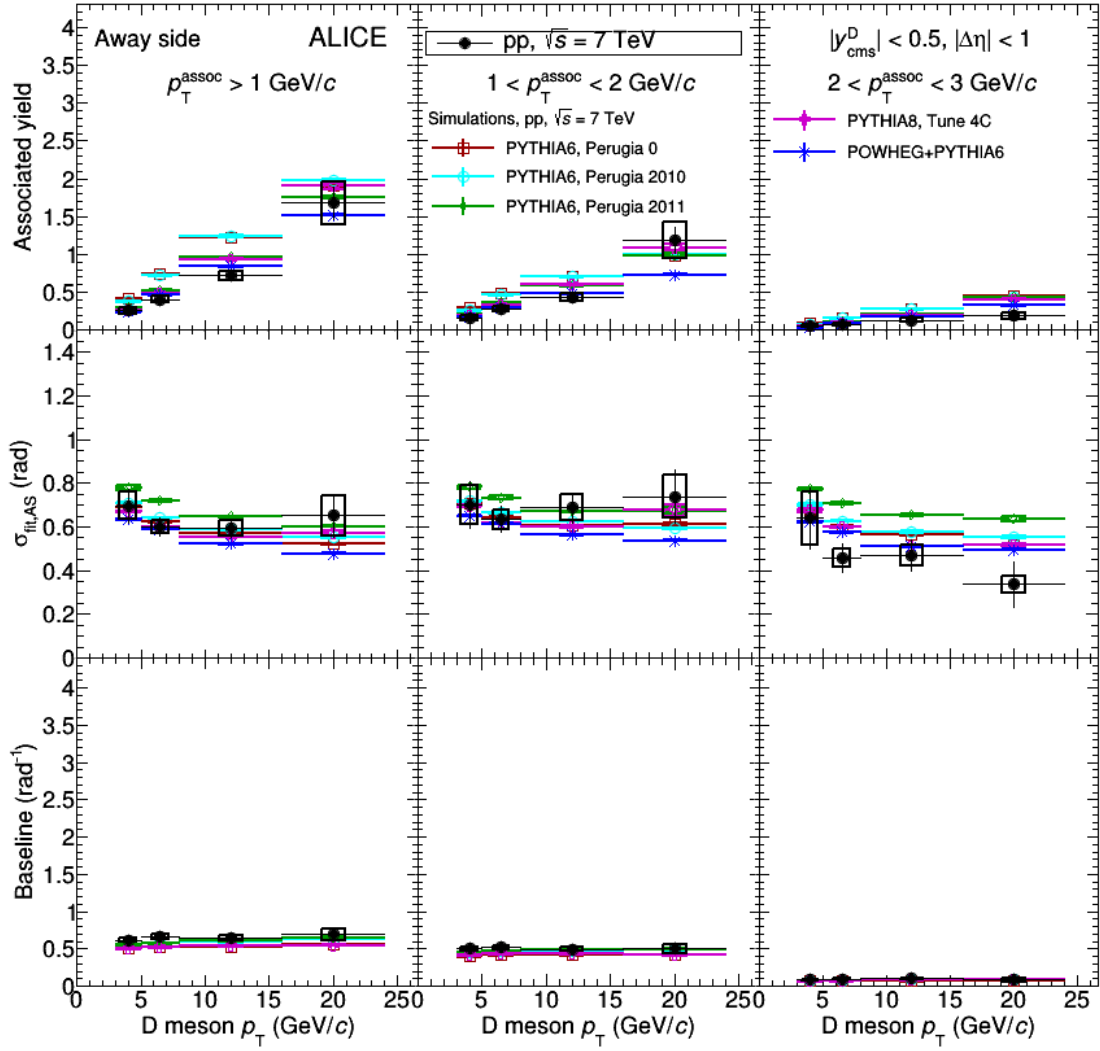
**Figure 23:** Comparison of  $\Delta\phi$  azimuthal distribution for D-meson averages, obtained from data and simulations different event generators (PHYTIA, with three tunes, and POWHEG+PYTHIA), in the different kinematic ranges analyzed.





**Figure 24:** Near-side fit parameters obtained for D-meson averages, extracted from data and simulations different event generators (PYTHIA,6 with three tunes, PYTHIA8 and POWHEG+PYTHIA), in the different kinematic ranges analyzed.





**Figure 25:** Near-side fit parameters obtained for D-meson averages, extracted from data and simulations different event generators (PYTHIA6, with three tunes, PYTHIA8 and POWHEG+PYTHIA), in the different kinematic ranges analyzed.

## 6 Bibliography

### References

- [1] B. Abelev et al. [ALICE Collaboration], JHEP 01 (2012) 128
- [2] B. Abelev et al. [ALICE Collaboration], Eur. Phys. J. C (2017) 77:245
- [3] B. Abelev et al. [ALICE Collaboration], Phys.Lett. B719 (2013) 29-41
- [4] B. Abelev et al. [ALICE Collaboration], Phys.Lett. B726 (2013) 164-177
- [5] <https://aliceinfo.cern.ch/Figure/node/12130> (ALI-PREL-138003)
- [6] ALICE Collaboration, arXiv:1609.06643 [nucl-ex]
- [7] <https://aliceinfo.cern.ch/Notes/node/300>
- [8] <https://aliceinfo.cern.ch/Notes/node/238>
- [9] <https://aliceinfo.cern.ch/Notes/node/201>