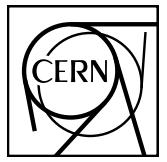


EUROPEAN ORGANIZATION FOR NUCLEAR RESEARCH



August 2, 2018

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

F. Colamaria, S. Kumar, M. Mazzilli

Abstract

In this note, we present the analysis of azimuthal correlations of D mesons and primary charged π, K, p, e, μ performed in the ALICE central barrel in pp collisions at $\sqrt{s_{\text{NN}}} = 5.02 \text{ TeV}$, 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 published 2013 pp analysis results for the common p_T ranges.

1 **Contents**

2	1	Introduction and Motivation	3
3	2	Data/Monte Carlo samples and event selection	4
4	3	Analysis strategy	8
5	3.1	Mass plots and cut optimization	11
6	3.2	Code used for the analysis	11
7	3.3	Further details on corrections	11
8	3.3.1	Event Mixing	11
9	3.3.2	Tracking and D-meson trigger efficiency	21
10	3.3.3	Correction for bias on B to D decay topologies	27
11	3.3.4	Secondary track contamination	32
12	3.3.5	Beauty feed-down	36
13	4	Systematic uncertainties on $\Delta\phi$ correlation distributions	40
14	4.1	Uncertainty on S and B extraction	40
15	4.2	Uncertainty on background correlation shape	40
16	4.3	Uncertainty on D-meson cut stability	44
17	4.4	Uncertainty on tracking efficiency evaluation	44
18	4.5	Uncertainty on secondary particle contamination	48
19	4.6	Uncertainty on feed-down subtraction	49
20	4.7	Uncertainty on correction for the bias on B to D decay topologies	53
21	4.8	Summary table	54
22	5	Results	55
23	5.1	Comparing the three D meson correlation distributions	55
24	5.2	Average of D^0 , D^+ and D^{*+} results	68
25	5.3	Fit observable p_T trends and uncertainties	71
26	5.3.1	Results for near-side yield and width, away-side yield and width, and baseline .	72
27	5.4	Comparison of 2016 p-Pb and 2013 p-Pb results	88
28	5.5	Comparison of 2016 p-Pb and 2010 pp results	88
29	5.6	Comparison of 2016 p-Pb and model expectations	88
30	5.7	Planned results for SQM approvals	88

31	6	Bibliography	99
----	----------	---------------------	-----------

32 1 Introduction and Motivation

33 The study of the azimuthal correlations of heavy-flavour particles and charged particles at the LHC
 34 energies provides a way to characterize charm production and fragmentation processes in pp collisions.
 35 The measurement also provide a way to probe our understanding of QCD in the perturbative regime,
 36 accessible in a large kinematic range given the large mass of heavy quarks. Flavour conservation in
 37 QCD implies that charm quarks are always produced as pairs of quarks and anti-quarks. The azimuthal
 38 correlations obtained using a meson carrying a heavy quark as trigger particle with the other charged
 39 particles in the same event give the possibility to study the underlying charm production mechanism in
 40 detail. In particular, prompt charm quark-antiquark pair production is back to back in azimuth at first
 41 order in leading-order perturbative-QCD (pQCD). If a hadron from the quark hadronization is taken as
 42 trigger particle, a near-side (at $\Delta\varphi = 0$) and an away-side (at $\Delta\varphi = \pi$) peaks would appear in the azimuthal
 43 correlation distributions, coming from the fragmentation of the quark pair. Heavy quarks produced from
 44 the splitting of a massless gluon can be rather collimated and may generate sprays of hadrons at small
 45 $\Delta\varphi$. Finally, for hard-scattering topologies classified as “flavour-excitation”, a charm quark undergoes a
 46 hard interaction from an initial splitting ($g \rightarrow c\bar{c}$), leading to a big separation in rapidity of the hadrons
 47 originating from the antiquark (quark) with respect to the trigger D meson and contribute to a rather flat
 48 term to the $\Delta\varphi$ -correlation distribution.

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

54 The extension of the momentum ranges (both for D mesons and associated particles) with respect to the
 55 2016 pp dataset, as well as the improved precision in the common ranges allow a more thorough investi-
 56 gation of the charm quark fragmentation properties (multiplicity of tracks as a function of momentum,
 57 geometrical profile of charm jets, p_T distribution of the tracks inside the jet). This can also allow us to
 58 put better constraints of charm fragmentation and charm jet properties provided by models. Furthermore,
 59 2017 pp data sample allows us a direct and more reasonable comparison with 2016 p-Pb data, since it
 60 has the same center-of-mass energy and, thanks to the higher precision and statistics, it was possible to
 61 exploit the azimuthal correlations in the same (extended and more differential) p_T ranges of both the trig-
 62 ger and the associated particle of the 2016 p-Pb data sample. This new pp reference data, together with
 63 new p-Pb 2016 data will help to study cold nuclear matter effects affecting the charm fragmentation in
 64 p-Pb with better precision. In addition, this new pp data can also be used as solid and precise references
 65 in view of an analysis on a Pb-Pb sample at the same energy.

66 2 Data/Monte Carlo samples and event selection

67 The data used for the analyses were the AOD samples of the following four datasets: LHC17p_pass1_FAST,
 68 LHC17p_pass1_CENT_woSDD, LHC17q_pass1_FAST and LHC17q_pass1_CENT_woSDD . The reason
 69 for choosing these data samples (in particular, those without the drifts for the CENT cluster) is explained
 70 later on, in this section. It was verified, by looking at D-meson and associated charged track η and φ
 71 distributions, and at the mixed-event correlation distributions for each sub-samples, that no visible differ-
 72 ences is present for the four periods, hence it was possible to perform the analysis directly on the merged
 73 samples without any bias.

74 The Monte Carlo productions adopted for this study were:

- 75 1. LHC18a4a2_fast, a HF production (HIJING with GEANT3) anchored to LHC17p/q is with enrich-
 76 ment of heavy quarks (charm and beauty) and their decay products in each of the event, performed
 77 by PYTHIA6 with Perugia2011 tune, and with forced hadronic decays of the charmed hadrons.
 78 This production was used for D-meson efficiency evaluation, purity estimation and Monte Carlo
 79 closure test.
- 80 2. LHC17l3b_fast, minimum-bias samples produced with DPMJET generator, are used for the eval-
 81 uation of the tracking efficiencies.

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

84 The trigger mask request for the event selection is kINT7. Only events with a reconstructed primary
 85 vertex within 10 cm from the centre of the detector along the beam line are considered. This choice
 86 maximises the detector coverage of the selected events, considering the longitudinal size of the inter-
 87 action region, and the detector pseudorapidity acceptances. Beam-gas events are removed by offline
 88 selections based on the timing information provided by the V0 and the Zero Degree Calorimeters, and
 89 the correlation between the number of hits and track segments in the SPD detector. This is automatically
 90 performed in the Physic Selection, a positive outcome of which is required during our event selection.
 91 The pile-up cuts for out-of-bunch pile-up protection are also invoked when calling the Physics Selection
 92 task. The minimum-bias trigger efficiency is 100% for events with D mesons with $p_T > 1 \text{ GeV}/c$. For
 93 the analyzed data samples, the probability of pile-up from collisions in the same bunch crossing is below
 94 2% per triggered event (in most of the runs, well below 1%). Events in which more than one primary in-
 95 teraction vertex is reconstructed with the SPD detector (with minimum of 5 contributors, and a z distance
 96 greater than 0.8 cm) are rejected, which effectively removes the impact of in-bunch pile-up events on
 97 the analysis. Out-of-bunch tracks are effectively rejected by the Physics Selection pile-up cuts, and also
 98 by the request of at least one point in the SPD, which has a very limited time acquisition window (300
 99 ns). Indeed, though the default associated track selection requires a minimum of 2 points in the ITS, as it
 100 will be shown later on full compatibility of the corrected results with 2 and 3 minimum ITS clusters are
 101 obtained. For FAST and CENT_woSDD samples, the latter case indirectly forces the presence of a point
 102 in the SPD.

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

110 **3 Analysis strategy**

111 The analysis follows the same strategy one used in 2016 p-Pb and pp data samples (see published paper
 112 [2] and analysis notes [8], [7]). Correlation pairs are formed by trigger particles (D mesons) reconstructed
 113 and selected in the following p_T^{trig} ranges: $3 < p_T^{\text{trig}} < 5 \text{ GeV}/c$, $5 < p_T^{\text{trig}} < 8 \text{ GeV}/c$, $8 < p_T^{\text{trig}} < 16$
 114 GeV/c , $16 < p_T^{\text{trig}} < 24 \text{ GeV}/c$ (the possibility of extending the analysis in $2 < p_T^{\text{trig}} < 3$ was also
 115 exploited. Further details are furnished in the next paragraph). Associated particles (charged tracks)
 116 have been reconstructed in the following p_T^{assoc} regions: $p_T^{\text{assoc}} > 0.3 \text{ GeV}/c$, $0.3 < p_T^{\text{assoc}} < 1 \text{ GeV}/c$,
 117 $1 < p_T^{\text{assoc}} < 2 \text{ GeV}/c$, $2 < p_T^{\text{assoc}} < 3 \text{ GeV}/c$, $p_T^{\text{assoc}} > 1 \text{ GeV}/c$. In this analysis, the particle identifica-
 118 tion defines the trigger particle rather than a momentum cut and therefore the momentum range of the
 119 associated particles is not constrained by that of the trigger particle. Our definition of associated particle
 120 includes primary particles of the following species: pion, kaon, proton, electron, muon. The primary par-
 121 ticle definition comprises particle coming from the primary vertex of interaction, including those coming
 122 from strong and electromagnetic decay of unstable particles, and particles deriving from the decay of
 123 hadrons with charm or beauty. We therefore include any charged π, K, p, e, μ except those coming from
 124 weak decays of strange particles and particles produced in the interaction with the detector material.
 125 This definition corresponds to that used in the method `AliAODMCParticle::IsPyphysicalPrimary()`. All
 126 associated particles surviving the selection cuts and not matching the adopted criterion are considered as
 127 a contamination whose contribution has to be corrected for.

128

129 The analysis is performed through the following steps:

- 130 1. **D meson selection and signal extraction.** For each single event, “trigger” particles are defined
 131 as the selected D meson candidates (D^0 , D^+ and D^{*+}) within a given p_T^{trig} range. The detection
 132 strategy for D mesons at central rapidity is the same performed for the analyses of the D-meson
 133 production at central rapidity [1], and also applied for the D-h analysis on 2016 pp and 2016
 134 p-Pb samples ([8], [7]). It is based on the reconstruction of decay vertices displayed from the
 135 primary vertex by a few hundred μm and on the identification of the decay-particle species. The
 136 identification of the charged kaon and pion in the TPC and TOF detectors is also used, to further
 137 reduce the background at low p_T . An invariant-mass analysis is then used to extract the raw signal
 138 yield, using the same fit functions described in [2]. The D mesons are selected in the rapidity range
 139 varying from $|y| < 0.5$ at low p_T to $|y| < 0.8$ for $p_T > 5 \text{ GeV}/c$.
- 140 2. **Correlation of D candidates with associated tracks.** Particle pairs are formed by correlating each
 141 trigger particle with the charged primary particles passing the track selection (excluding those
 142 coming from the decay of the D-meson candidate) in a specified p_T^{assoc} interval (which can overlap
 143 with the p_T^{trig} range) and in the pseudo-rapidity range $|\eta| < 0.8$. For the D^0 meson, also the low-
 144 momentum pion tracks from feed-down of D^{*+} mesons are removed via 3σ invariant mass cut on
 145 the $M(K\pi\pi) - M(K\pi)$ difference. This because these soft pion are not related to the charm quark
 146 fragmentation chain. For D meson candidates in the invariant mass signal region, defined by a \pm
 147 2σ interval around the D meson mass peak, the azimuthal angle difference $\phi^{\text{assoc}} - \phi^{\text{trigg}} \equiv \Delta\phi$ and
 148 the pseudorapidity difference $\eta^{\text{assoc}} - \eta^{\text{trig}} \equiv \Delta\eta$ are evaluated and stored to build two-dimensional
 149 correlation distribution.
- 150 3. **Correction for limited acceptance and detector inhomogeneities with Event Mixing** The angular
 151 correlation distribution may be affected, even for uncorrelated pair of particles, by structures
 152 not due to physical effects, but originating from the limited detector acceptance, as well as from
 153 angular inhomogeneities in the trigger and track reconstruction efficiencies as a function of $\Delta\phi$
 154 and $\Delta\eta$. Effects of this kind are removed using the Event Mixing technique. In this technique, the
 155 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

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

where N^{pair} is the total number of correlated D-hadron pairs. The functions $S(\Delta\eta,\Delta\varphi)$ and $B_{ME}(\Delta\eta,\Delta\varphi)$ are the signal and the mixed event background distributions, respectively. The later is normalized to its value in $(\Delta\eta,\Delta\varphi) = (0,0)$, 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 $\Delta\varphi$) is shown in figure 3, 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, ε_{trk} and $\varepsilon_{\text{trig}}$. Further details are provided later on in this section.
- 6. Projection in $\Delta\varphi$.** The limited statistics available does not allow to study the two dimensional $(\Delta\eta,\Delta\varphi)$ distribution, which is therefore projected to the $\Delta\varphi$ 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 ($\Delta\varphi = 0 = 2\pi$), the $\Delta\varphi$ -range is limited to the essential range of 2π . The $\Delta\varphi$ -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 π .
- 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

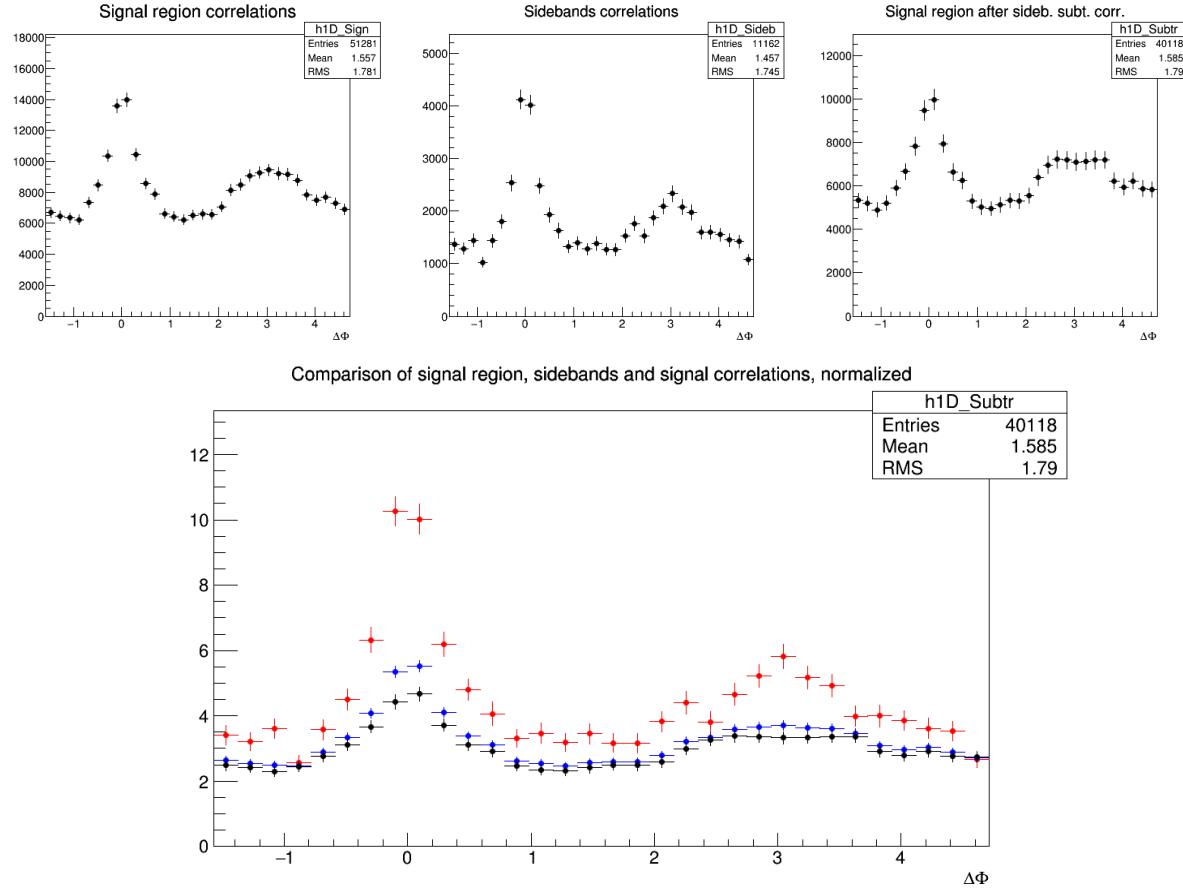


Figure 1: Top: Example of D^0 - h signal region (left), sideband (middle), and signal minus sideband (right) correlation distributions. Bottom: signal region per-trigger normalized correlation distribution (blue), sideband region per-trigger normalized correlation distribution (red), background-subtracted per-trigger normalized correlation distribution (black).

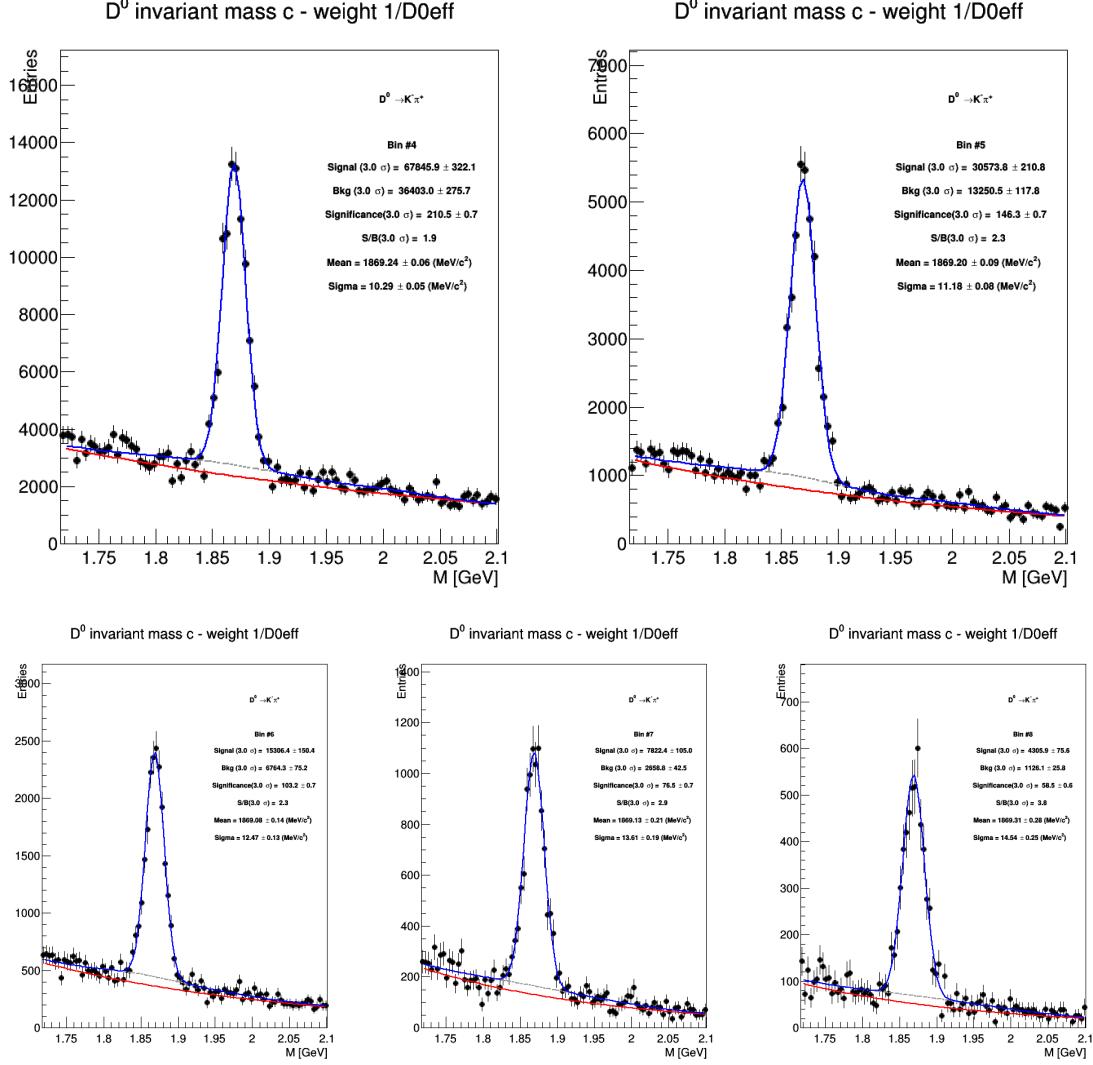
is estimated from Monte Carlo simulation based on Pythia as described more in detail in the next section. The background-subtracted event-mixing corrected correlations are multiplied by a purity factor to encounter this contribution.

- 200 is estimated from Monte Carlo simulation based on Pythia as described more in detail in the next
201 section. The background-subtracted event-mixing corrected correlations are multiplied by a purity
202 factor to encounter this contribution.
- 203 8. **Correction for bias on B to D decay topologies** The presence of the topological cuts for the D-
204 meson selection indirectly induce a bias on the topology of the B to D decay topologies, favouring
205 cases with a small opening angle between the D-meson and the other tracks from the B decay. This
206 affects the feed-down component of the data correlation distributions. This effect is corrected for
207 with a procedure described in the subsection 3.3.3.
- 208 9. **Correction for feed-down of D meson from b-hadron decay** The selection strategy employed
209 for the D meson candidates selection enhances the fraction of reconstructed D mesons coming
210 from the decay of a b-hadron. Typical values, with the cuts used for the D-meson selection, are of
211 the order of 10% or less. The correlation distribution of these secondary D mesons will be sensi-
212 tive to the properties of beauty jets and beauty hadron decay, which in general differ from those
213 relative to charm jets and hadrons. The procedure used to subtract this contribution is described in
214 the next paragraphs of this section.
- 215 10. **Study of correlation properties.** The properties of the azimuthal correlation distribution are quan-
216 tified by fitting the distribution with a function composed of two Gaussian functions, modelling
217 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 2π 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 5.

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 4, 5 and 6 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 were 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 loosing 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 7 the D^0 -h correlation distributions are shown for the different cut sets, in exemplary kinematic regions (left

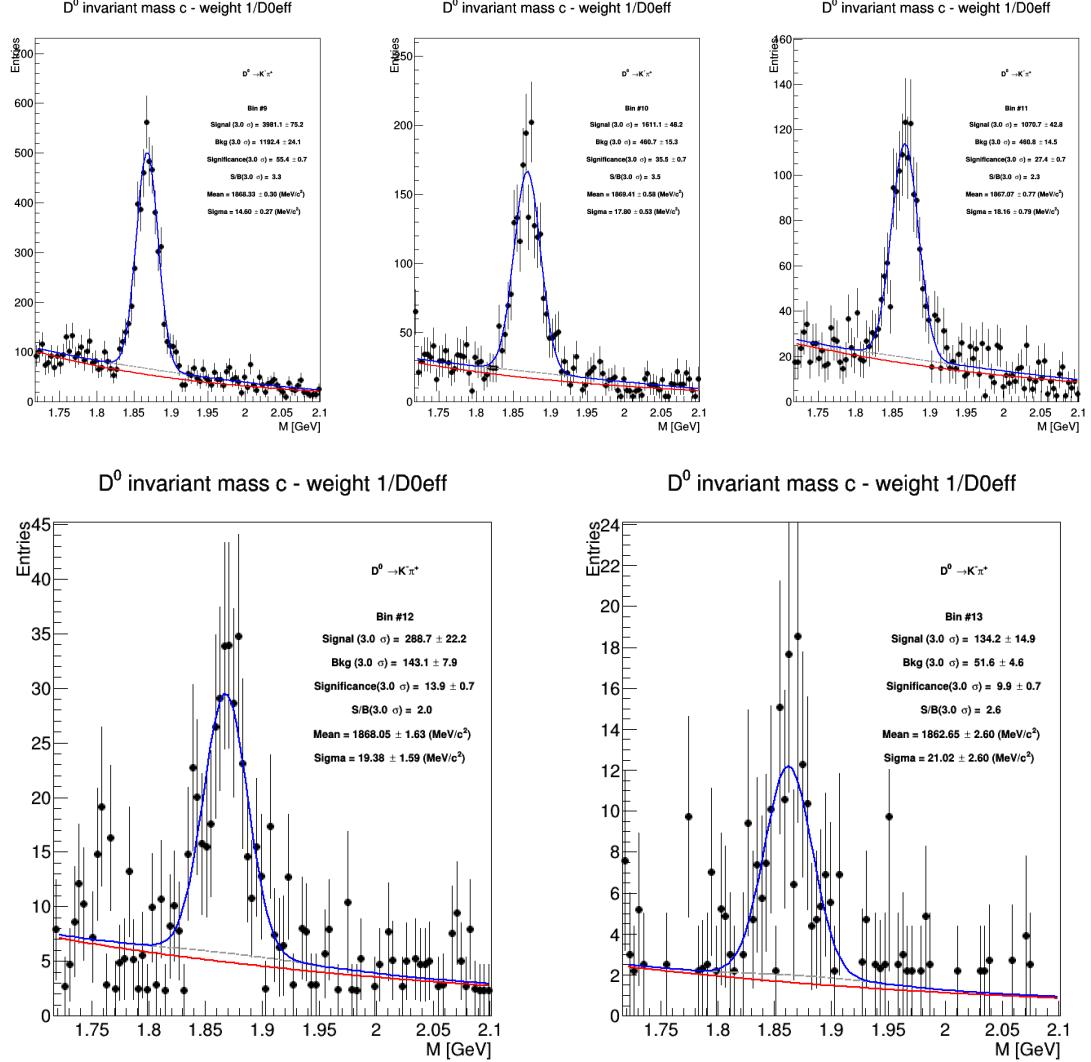


Figure 2: Invariant mass distributions of D^0 corrected with efficiency in different p_T regions. Top: $3 < p_T^D < 4$ GeV/c (left), $4 < p_T^D < 5 \text{ GeV}/c$ right), Mid 1: $5 < p_T^D < 6 \text{ GeV}/c$ (left), $6 < p_T^D < 7 \text{ GeV}/c$ (middle), $7 < p_T^D < 8 \text{ GeV}/c$ (right); Mid2: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

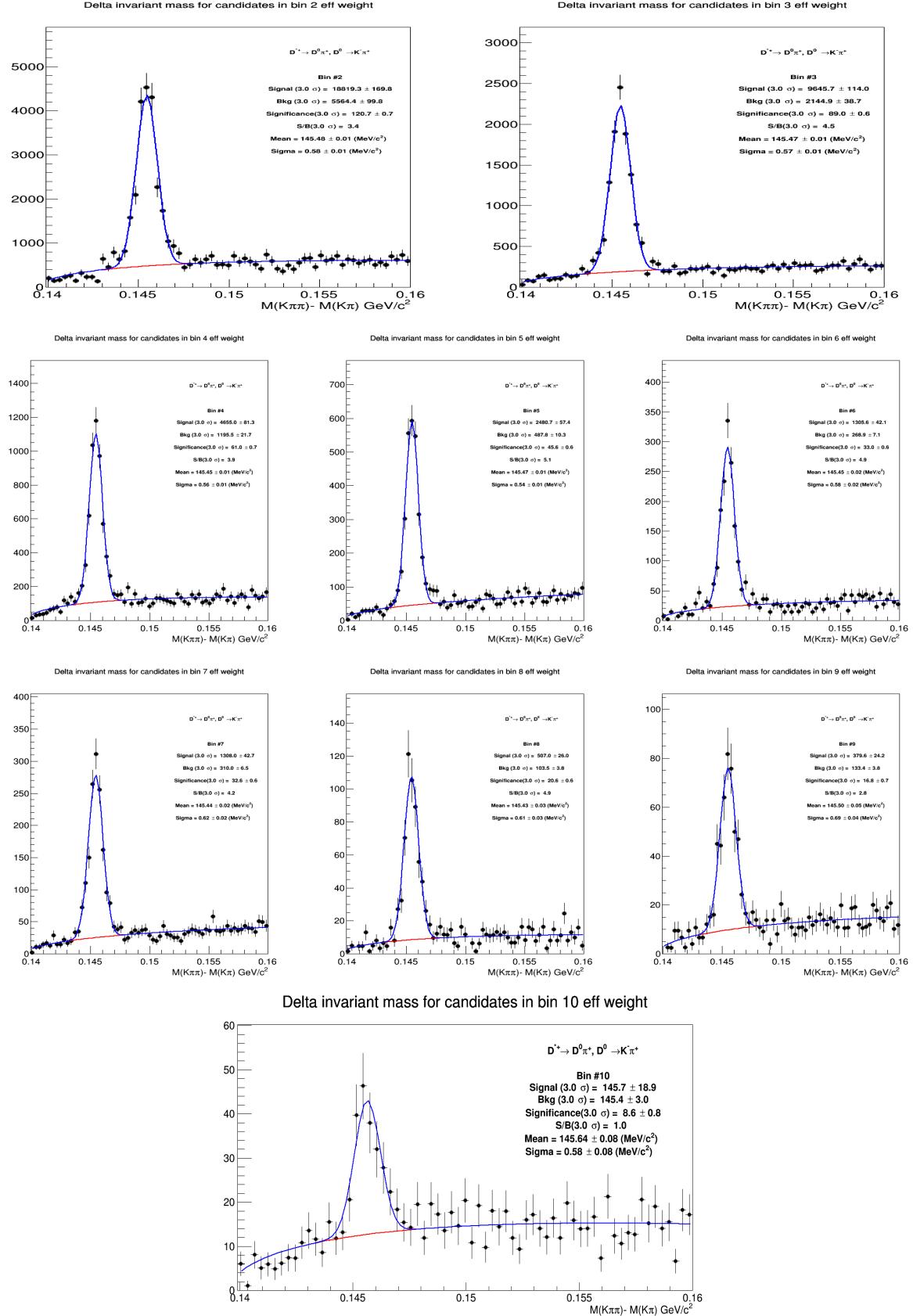
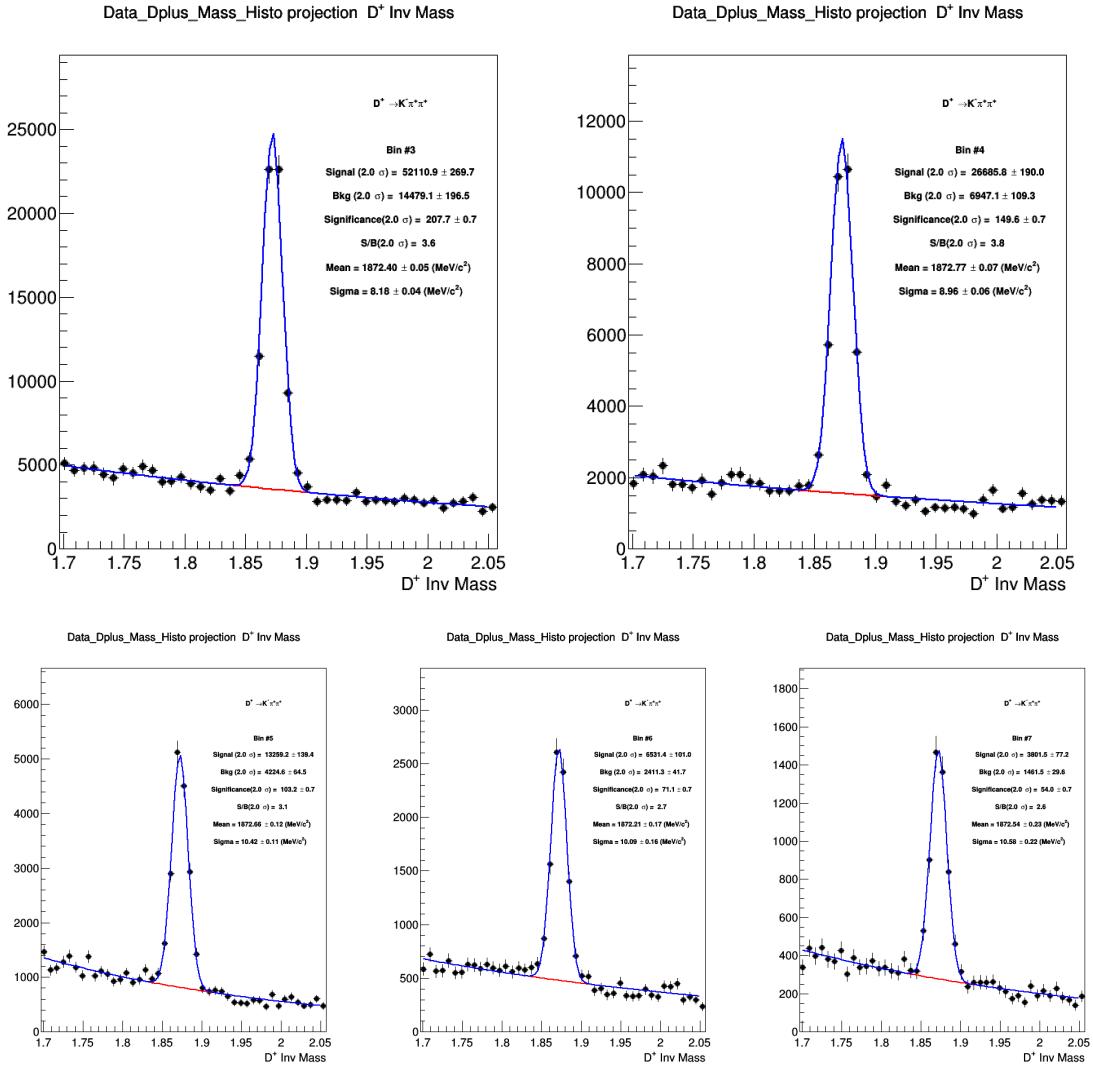


Figure 3: Invariant mass distributions of D^{*+} corrected with efficiency in different p_T^D regions. Top: $3 < p_T^D < 4$ GeV/ c (left), $4 < p_T^D < 5$ GeV/ c (right), Mid 1: $5 < p_T^D < 6$ GeV/ c (left), $6 < p_T^D < 7$ GeV/ c (middle), $7 < p_T^D < 8$ GeV/ c (right); Mid 2: $8 < p_T^D < 10$ GeV/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



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 ?? 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 in 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 $p_T(\text{trig}) \sim p_T(\text{ass})$ 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 especially on the away-side region where most of the fit failed.

252 3.2 Code used for the analysis

253 The code used for D meson-hadron correlation analysis is fully committed in AliPhysics. The anal-
 254 ysis classes can be found in `$ALICE_ROOT/PWGHF/correlationHF/`. The D meson specific classes
 255 where the aforementioned steps are carried out are `AliAnalysisTaskDStarCorrelations`, `AliAnalysis-`
 256 `TaskSED0Correlations` and `AliAnalysisTaskDplusCorrelations`. The classes which are common to the D
 257 meson specific analysis which includes the associated particle cuts and the correlation observables are `Ali-`
 258 `iHFAssociatedTrackCuts`, `AliHFCorrelator`, `AliHFOfflineCorrelator`, `AliReducedParticle` and `AliDhCor-`
 259 `relationExtraction`. Several additional classes and macros in the same folder deal with the correction
 260 steps.

261 3.3 Further details on corrections

262 3.3.1 Event Mixing

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

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

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

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

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

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

288 The mixed-event correlation distributions are built in both D meson signal and sideband regions. Both
 289 are corrected with the relative distributions. An example of the mixed-event distributions, and of the
 290 outcome of the mixed-event correction, is provided in Figures ?? and 12. The expected triangular shape
 291 in $\Delta\eta$, for the mixed-event distributions, addresses the effect of the limited detector pseudo-rapidity
 292 acceptance. Note that the mixed-event distribution is limited to the interval $|\Delta\eta| < 1$: the decision to

293 limit the mixed-event correction, and thus the whole analysis, to this range was taken in order to avoid
294 the so-called “wing effect”, i.e. the wing-like structures arising in the correlation distribution at large $\Delta\eta$
295 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 = [44 runs]	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, 282365, 282366, 282367 = [44 runs]	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 = [41 runs]	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 = [42 runs]	
	LHC17q_pass1_FAST LHC17q_pass1_CENT_woSDD	282367, 282366, 282365 = [3 runs] 282367, 282366, 282365 = [3 runs]	

Table 1: Data Set and Run list

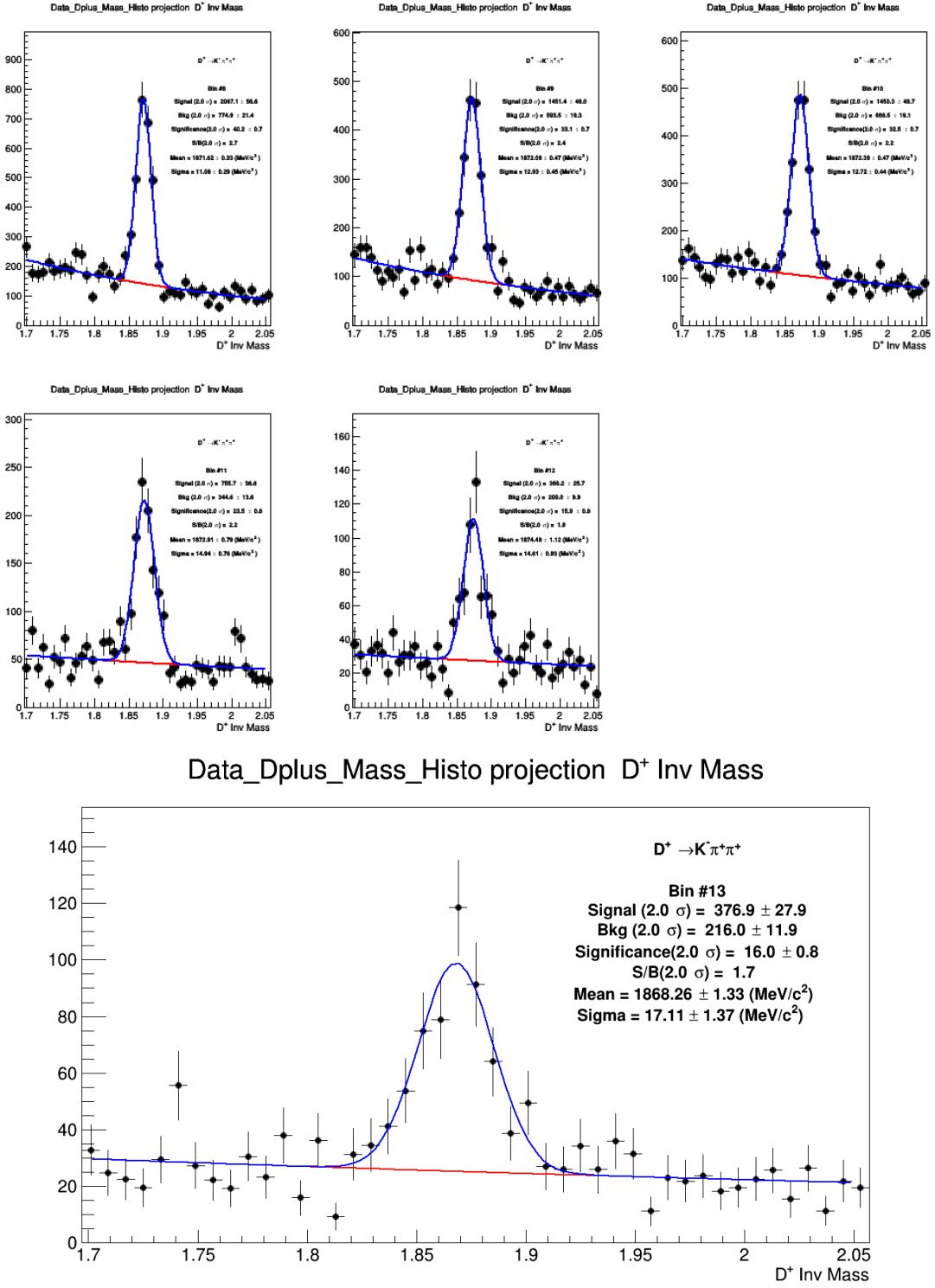


Figure 4: Invariant mass distribution of D⁺ corrected with efficiency in different p_T regions. Top: 3 < p_T^D < 4 GeV/c (left), 4 < p_T^D < 5 GeV/c (right), Mid 1: 5 < p_T^D < 6 GeV/c (left), 6 < p_T^D < 7 GeV/c (middle), 7 < p_T^D < 8 GeV/c (right); Mid2: 8 < p_T^D < 10 GeV/c, 10 < p_T^D < 12 GeV/c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

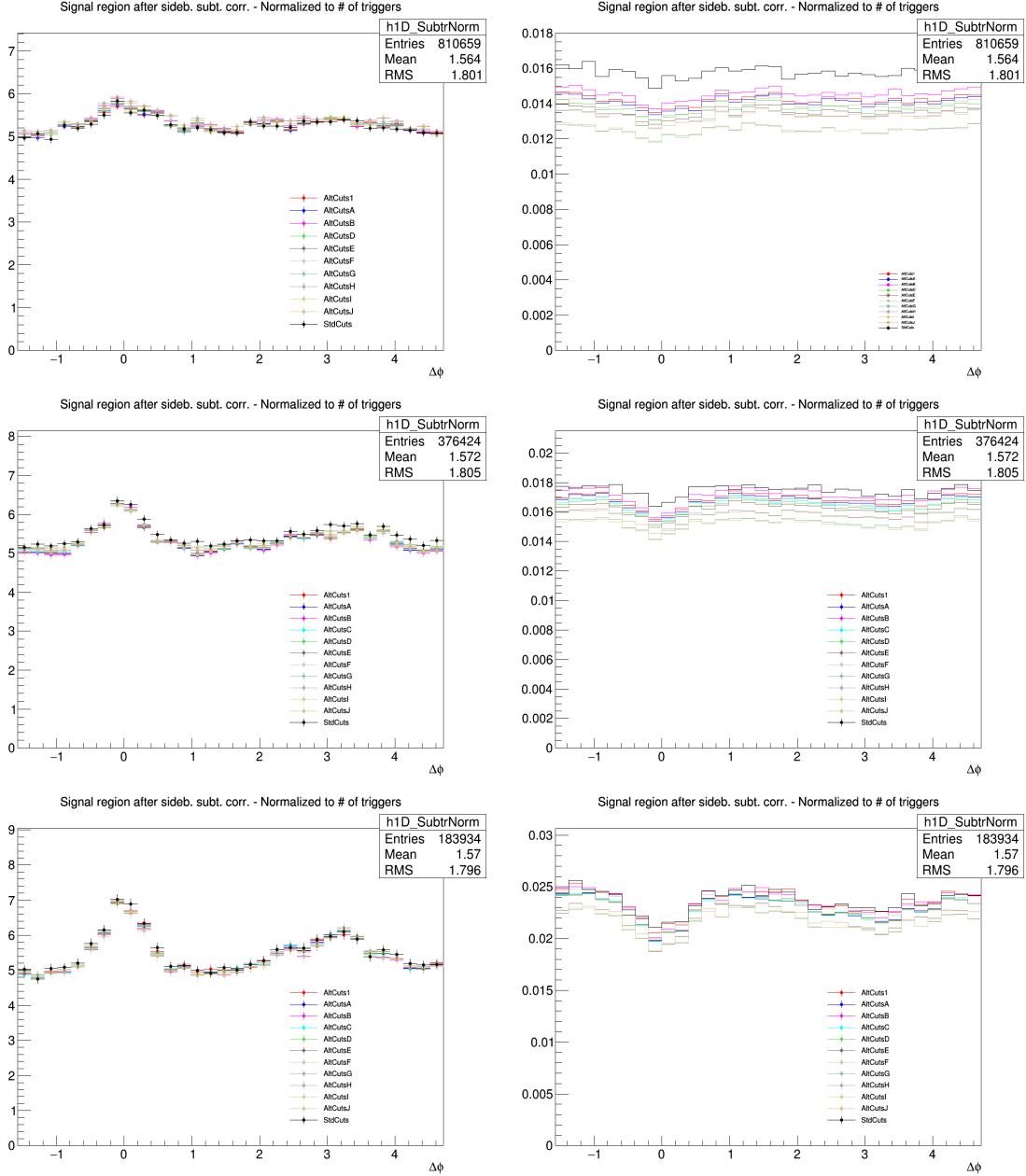
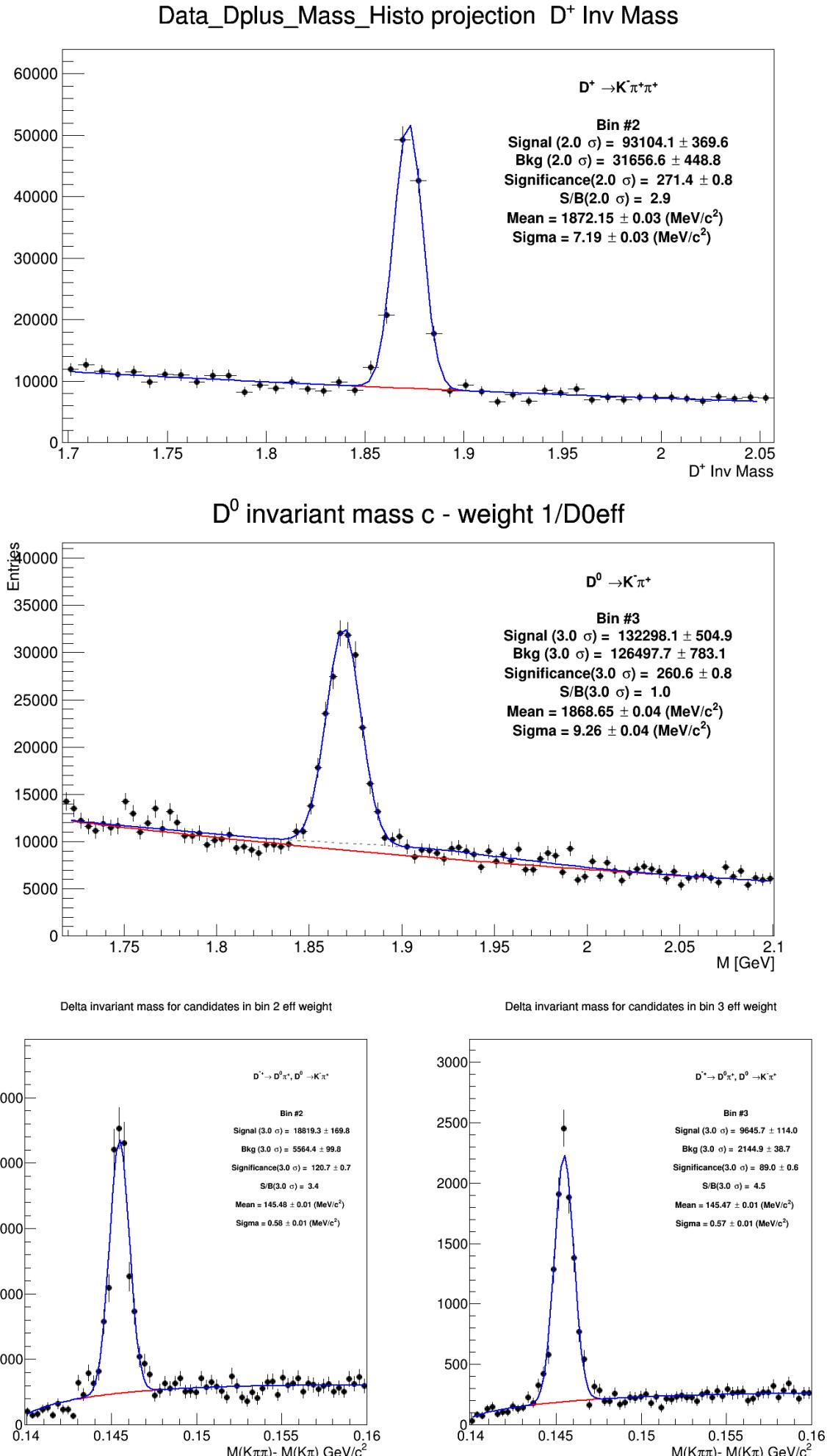


Figure 5: D^0 - h correlation distributions with different cut options (left) and point-by-point relative statistical uncertainty (right) for $3 < p_T^D < 5 \text{ GeV}/c$ (top), $5 < p_T^D < 8 \text{ GeV}/c$ (middle), $8 < p_T^D < 16 \text{ GeV}/c$ (bottom), in all cases with associated track $p_T > 0.3 \text{ GeV}/c$.



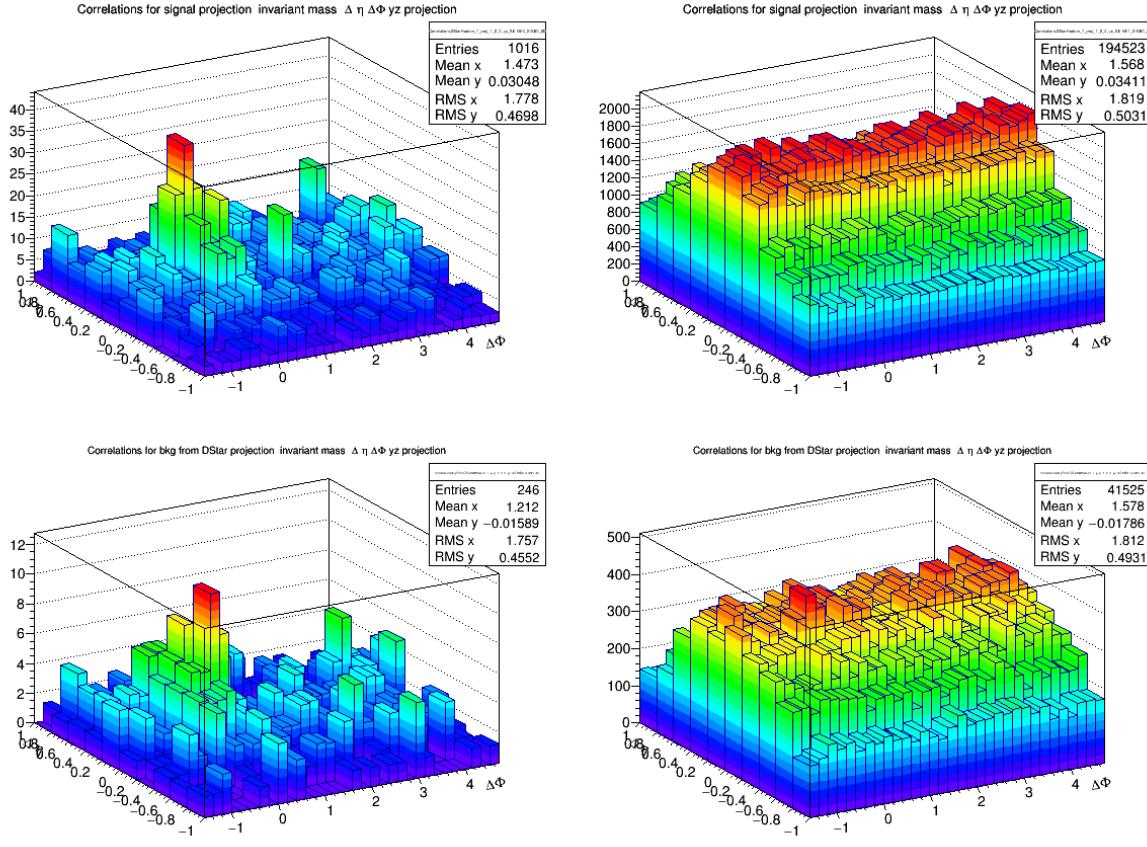
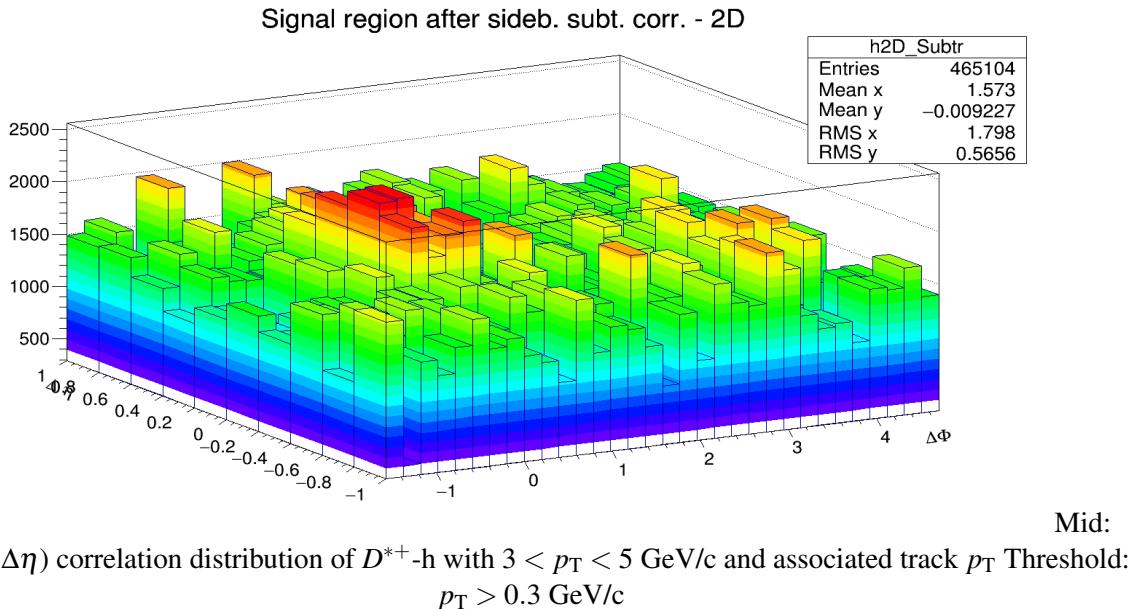


Figure 7: D^{*+} meson ($\Delta\phi$, $\Delta\eta$) correlation for in the signal region (top row) and sidebands (bottom row) from Single Event (left) and Mixed Event analysis (center) for high p_T : $8 < p_T < 16$ GeV/c with associated $p_T > 0.3$ GeV/c. The right column shows the SE/ME corrected distributions.



296 **3.3.2 Tracking and D-meson trigger efficiency**

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

300

301 Tracking efficiency maps were produced as TH3D histograms (p_T , η , z_{vtx}) obtained from MC analy-
302 sis on the minimum-bias samples LHC17l3b_fast and LHC17l3b_cent_woSDD anchored to LHC17p,q
303 data samples, considering only primary pions, kaons, protons, electrons and muons, and applying at
304 reconstructed level the track selections (summarized in Table. 2). These efficiency maps were used in
305 the analysis tasks to extract single track efficiencies; each correlation pairs found in the data analysis
306 was inserted in correlation plots with a weight of **1/efficiency value**. The 1D (p_T dependence) tracking
307 efficiency for all the five species as well as the tracking efficiency specie by specie are shown in Fig. 13.

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

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

309
 310 **(ii) D meson efficiency** - Due to limited statistics, the correlation analysis is performed in quite wide p_T
 311 bins and in each of them the reconstruction and selection efficiency of D mesons is not flat, in particular
 312 in the lower p_T region. We correct for the p_T dependence of the trigger efficiency within each p_T -bin.

313 This correction is applied online, by using a map of D meson efficiency as a function of p_T and event
 314 multiplicity (in terms of SPD tracklets in $|\eta| < 1$) extracted from the enriched Monte Carlo sample
 315 LHC18a4a2_fast. The η dependence was neglected due to the statistics of the available Monte Carlo
 316 sample, which rule out the possibility of performing a 3D study.

317 To properly count the number of trigger particles used to normalize the correlation distributions, N_{trig} ,
 318 each D meson is weighted with the inverse of its efficiency in the invariant mass distribution. The main
 319 role of the correction for the D meson efficiency is to account for the p_T dependence of the correlation
 320 distribution within a given D meson p_T interval. Indeed, only the p_T shape of the D meson efficiency
 321 within the correlation p_T^{trig} ranges is relevant while the average value in the p_T range is simplified due to
 322 the normalization of the correlation distribution to the number of trigger particles.

³²³ Efficiency plots for D^0 , D^+ and D^{*+} mesons are shown in Figs. 14 and 15.

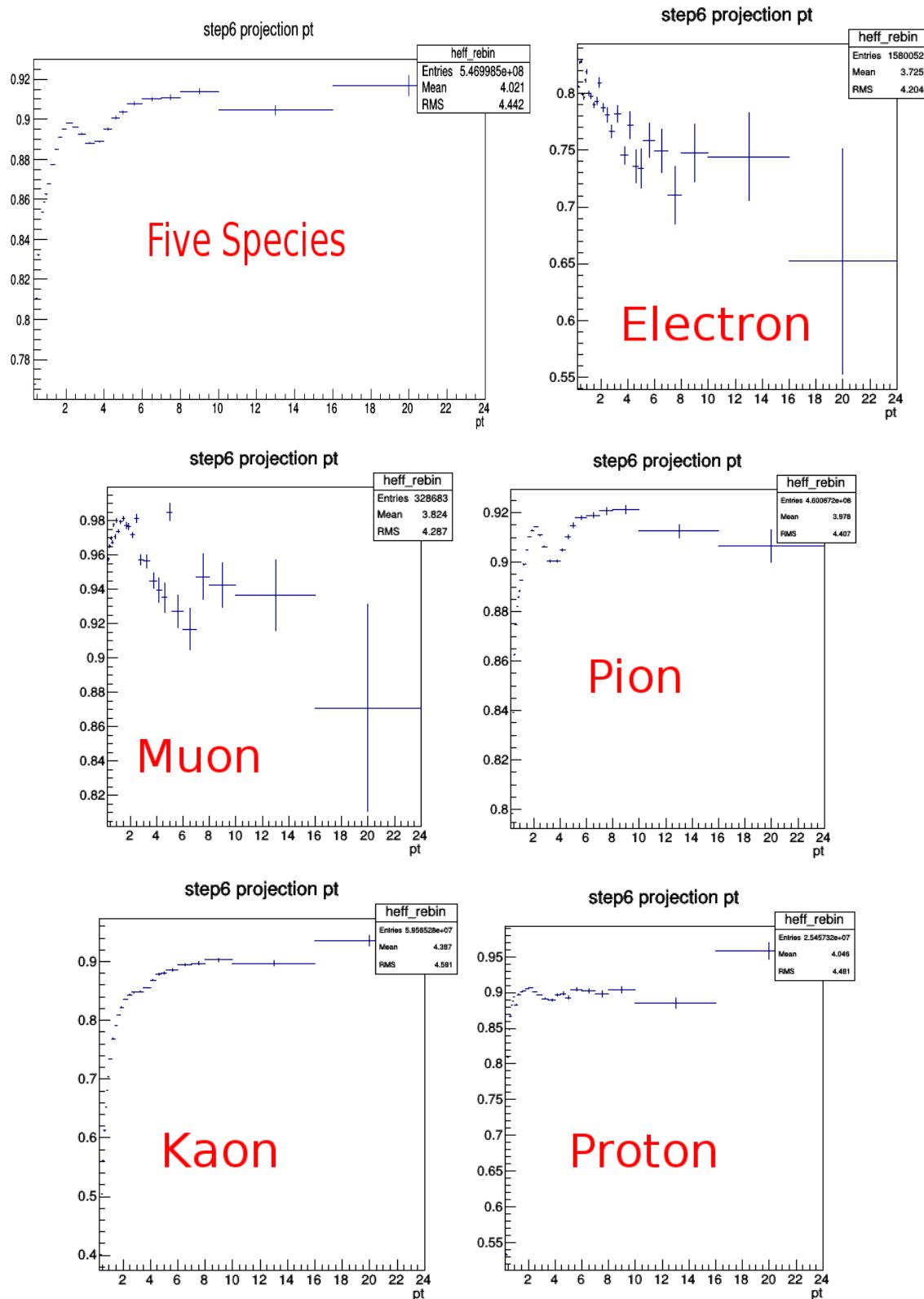


Figure 8: 1D (vs p_T) tracking efficiency map for standard track selection, evaluated for five species (electron, muon, pion, kaon and proton) and also different species using data sample LHC17l3b_fast.

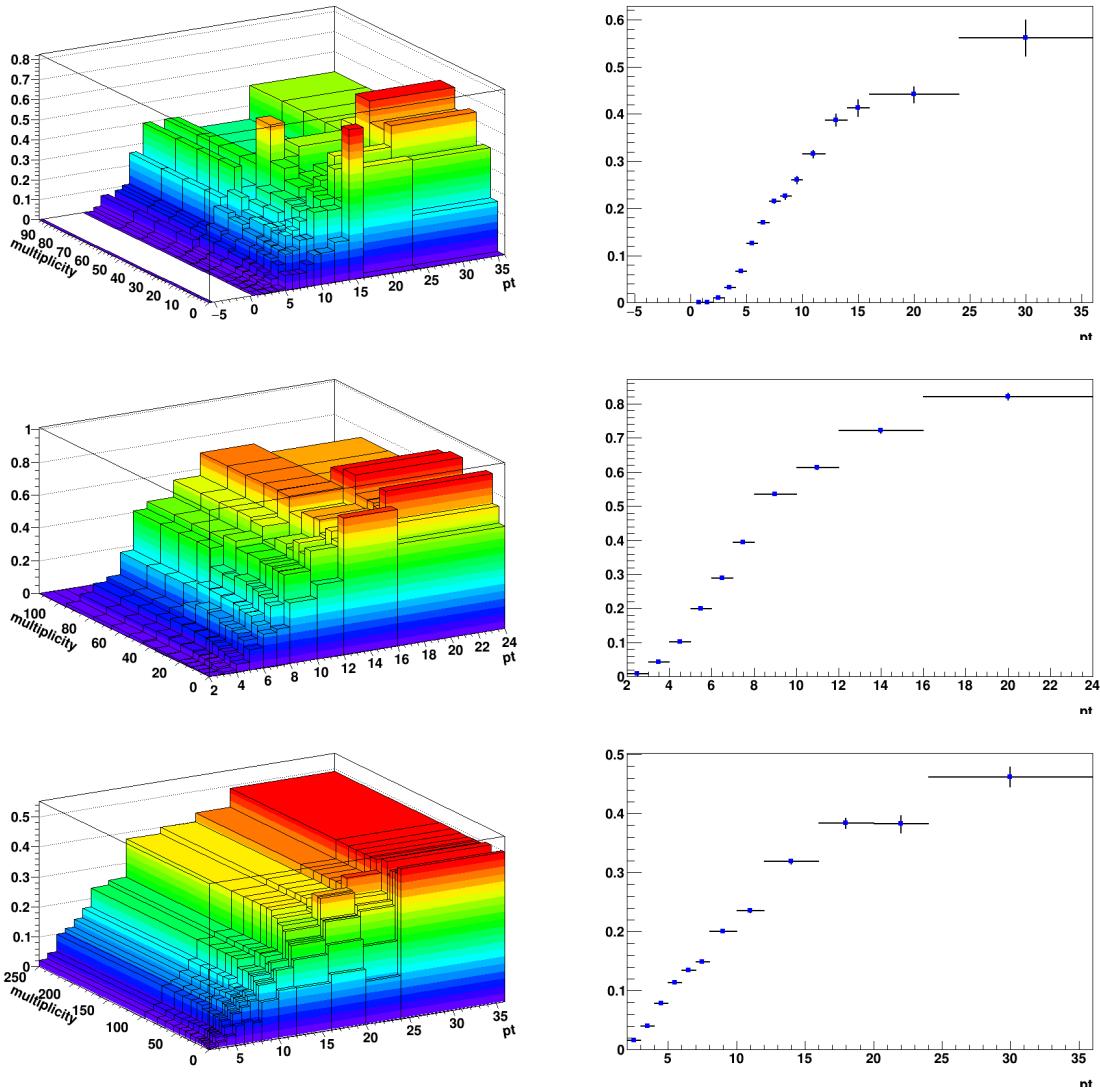


Figure 9: Top panel: $(p_T, \text{multiplicity})$ dependence (left) and p_T dependence (right) of prompt D^+ meson efficiency. Mid panel: $(p_T, \text{multiplicity})$ dependence (left) and p_T dependence (right) of prompt D^{*+} meson efficiency. Bottom panel: $(p_T, \text{multiplicity})$ dependence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

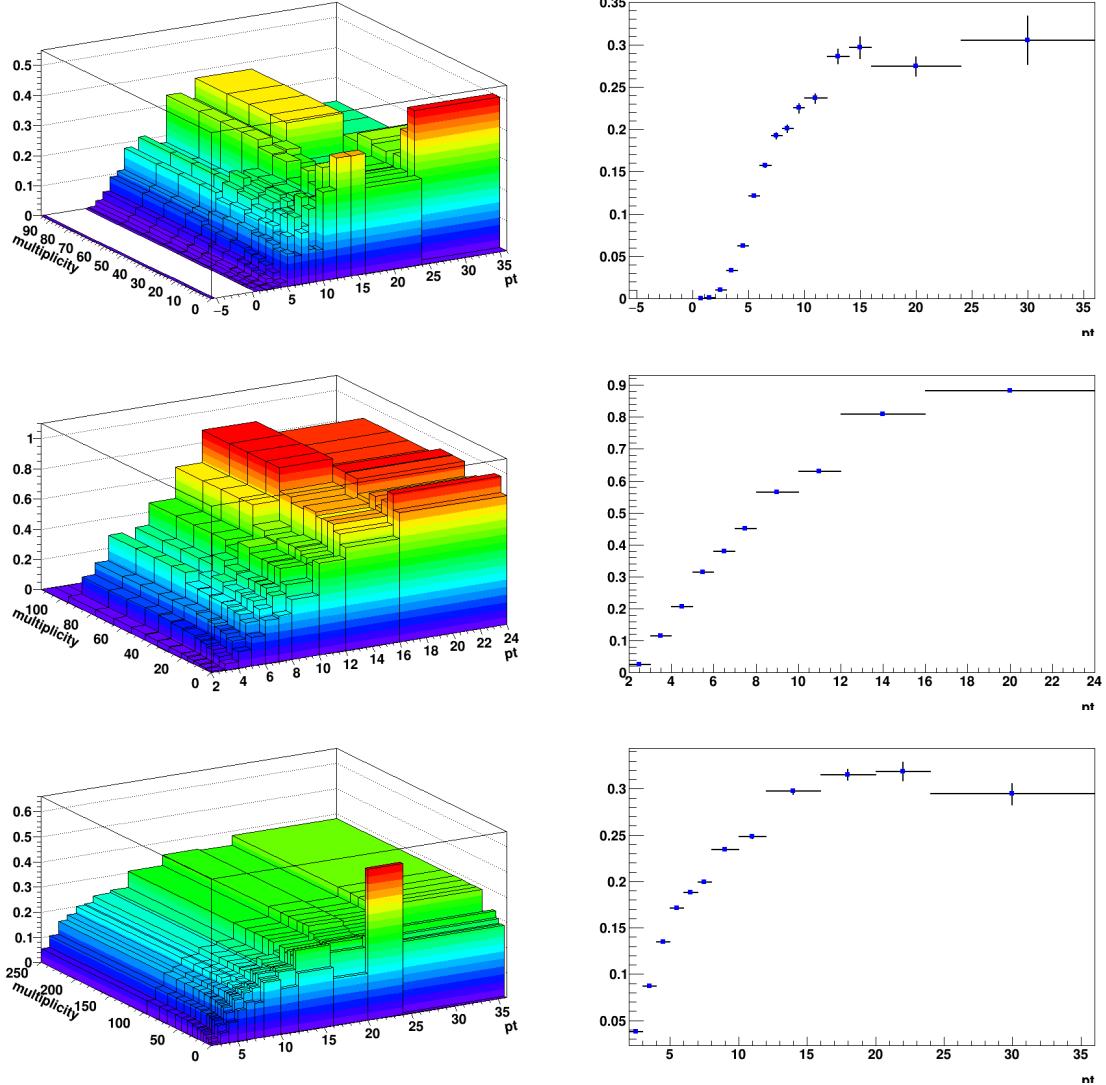


Figure 10: Top panel: (p_T , multiplicity) dependence (left) and p_T dependence (right) of feed-down D^+ meson efficiency. Mid panel: (p_T , multiplicity) dependence (left) and p_T dependence (right) of feed-down D^{*+} meson efficiency. Bottom panel: (p_T , multiplicity) dependence (left) and p_T dependence (right) of feed-down D^0 meson efficiency.

324 **3.3.3 Correction for bias on B to D decay topologies**

325 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.

327 On the Monte Carlo enriched with charm and beauty quarks (LHC18a4a2_fast, with GEANT3), the
 328 correlation analysis was performed both at kinematic level and at reconstructed level. At kinematic
 329 level, only acceptance cuts were applied on the D mesons and the associated particles, using the Monte
 330 Carlo information for the identification of the D mesons and the hadrons in the event and rejecting the
 331 non-primary particles. At reconstructed level, the analysis was performed as if it were executed on data,
 332 applying the event selection, the acceptance cuts for D mesons and the associated particles, selecting the
 333 D meson candidates with filtering cuts on their daughters, topological cuts and PID selection, and then
 334 keeping only the true D mesons by matching with the Monte Carlo truth; non-primary particles were
 335 rejected by means of the DCA selection. Event mixing correction was applied both at reconstructed and
 336 at kinematic level, where it takes into account just the effects of the acceptance cuts. In addition, at
 337 reconstructed level, the efficiency corrections for D mesons and associated tracks were also applied.

338 The consistency check was performed to verify whether, after having applied all the corrections to the
 339 azimuthal correlation plots at reconstructed level, the results were compatible with the ones at kinematic
 340 level. Hence, the ratios of fully corrected reconstructed plots over kinematic plots were evaluated in all
 341 the D^0 p_T bins and for the various p_T thresholds for the associated tracks, separating the contributions for
 342 the different origins of particles and triggers. The ratios, shown in Figure 19, denote a good compatibility
 343 with 1, within the uncertainties, with the only exception being due to some structures in the near side
 344 region for the beauty origin case. It was verified that these structures are induced by our topological
 345 selection for the D mesons. Indeed, in cases in which the D meson triggers come from B hadrons,
 346 applying the topological cuts (especially the cosine of the pointing angle) tends to favour cases with a
 347 small angular opening between the products of the B hadron decay (i.e. the D meson trigger itself and
 348 other particles), with respect to cases where the B decay particles are less collinear.

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

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

- 366 – On data, the $B \rightarrow D$ correlation pairs are only a fraction ($1-f_{\text{prompt}}$) of the total.
- 367 – The amplitude of $B \rightarrow D|_{\text{amplit}}$ correlation pattern is different (greater) than the amplitude of the
 368 $c \rightarrow D|_{\text{amplit}}$ correlation pattern:

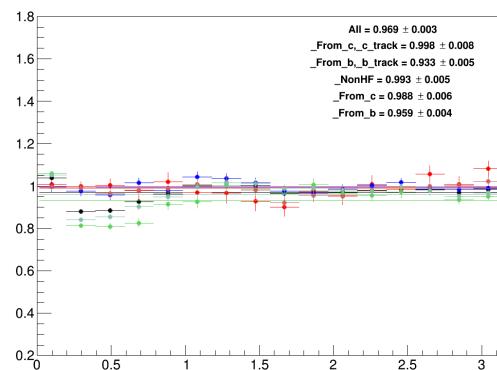
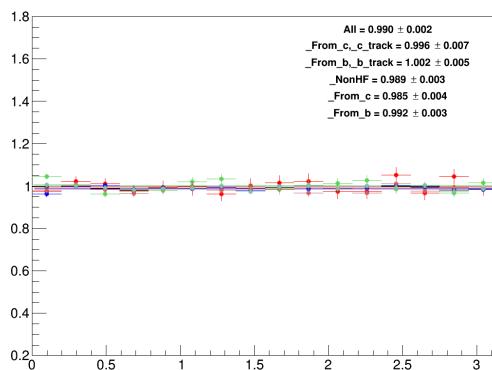
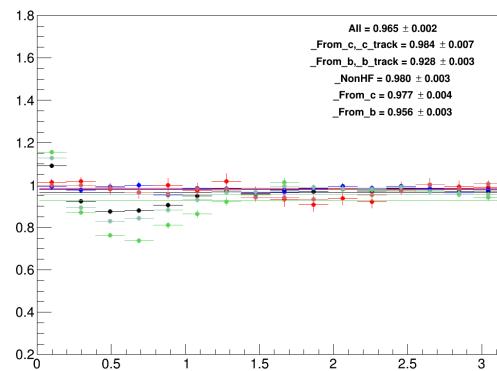
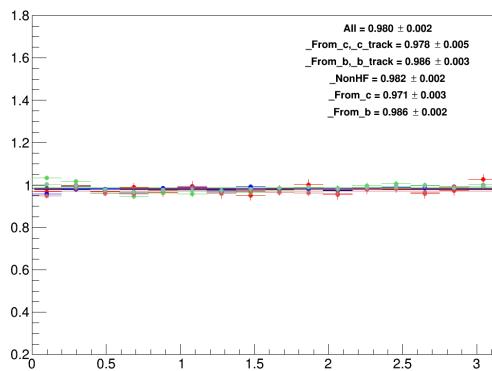
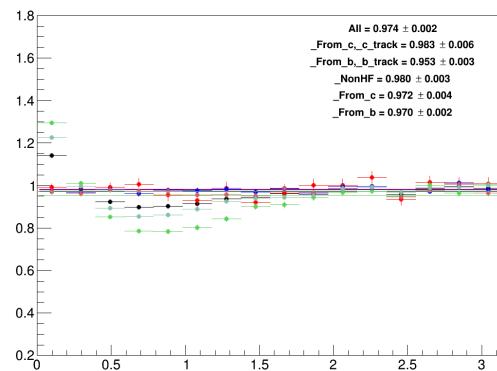
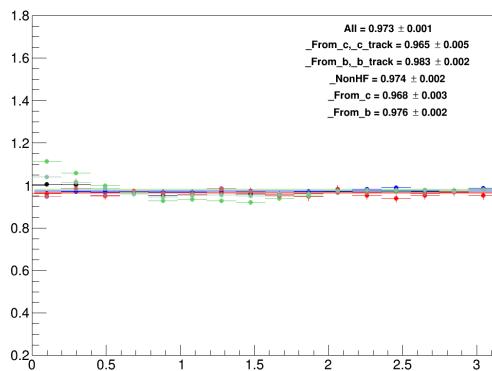
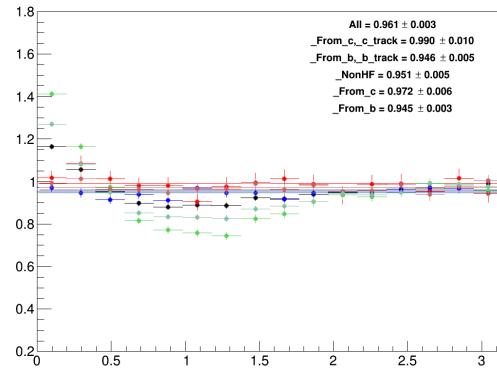
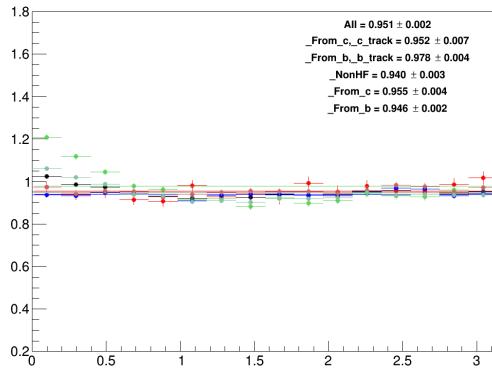


Figure 11: Ratios of fully corrected azimuthal correlation plots at reconstructed level over azimuthal correlation plots at kinematic level, in the two D^0 p_T bins, for the different associated p_T ranges. Black points: All D^0 -all hadrons, normalized by all D^0 triggers; light red points: D^0 from 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).

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

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

371 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
 372 are evaluated from the Monte Carlo distributions of Figure 18 at reconstructed level (so, including
 373 the bias), and f_{prompt} with the procedure described in 3.3.5. Applying the **modul** factor to the beauty part
 374 of the data correlation distributions brings its value back to the generated level case, effectively remov-
 375 ing the bias. The effect of the correction is a shift of the data points in the near-side region (in general,
 376 downward in the first and second points, the upward in the others). The maximum value of the shift is
 377 of about 5%, at the centre of the near-side peak, for the lowest D -meson p_T range ($3 < p_T < 5 \text{ GeV}/c$)
 378 and the highest associated track p_T range ($p_T > 3 \text{ GeV}/c$). The typical values are instead of a couple
 379 of percentage points. The correction is zero in the highest D -meson p_T range. To take into account for
 380 possible inaccuracies in the definition of the modulations, or in their rescaling, a systematic uncertainty
 381 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
 382 data points affected by the bias (symmetric uncertainty).

383 **3.3.4 Secondary track contamination**

384 The secondary tracks inside the associated track sample, due to interaction of primary track with the de-
 385 tector material or to decays of strange hadrons, are mostly removed by the DCA cuts applied during the
 386 cut selection phase ($\text{DCA}(xy) < 1 \text{ cm}$, $\text{DCA}(z) < 1 \text{ cm}$). Anyway, a small fraction of secondary tracks
 387 survives this cut, and the data correlation distributions have to be corrected for this residual contamina-
 388 tion. The fraction of surviving secondary tracks is evaluated via a study on the LHC18a4a2_fast sample,
 389 by counting the number of tracks accepted by the selection whose corresponding generated-level track
 390 doesn't satisfy the `IsPhysicalPrimary()` call, and dividing this number by the total number of accepted
 391 tracks. The outcome of the check is reported in Figure 20. As it's visible, no more than 5% secondary
 392 tracks pass the selection. Moreover, the fraction of residual secondary tracks is flat along the $\Delta\phi$ axis,
 393 as shown, for exemplary p_T regions, in Figure 21, where the inhomogeneities are always below 1%. For
 394 this reason, it is possible to directly scale the data correlation distributions by their purity fraction (i.e. 1
 395 - secondary contamination). This is done with an associated p_T dependence, due to the increase of the
 396 purity with the track p_T , while the purity fraction is taken flat versus the D-meson p_T . The purity values
 397 that were choosen are the following (NOT FINAL):

- 398 – $p_{T(\text{assoc})} > 0.3 \text{ GeV}/c : 0.958$
- 399 – $p_{T(\text{assoc})} > 1 \text{ GeV}/c : 0.973$
- 400 – $0.3 < p_{T(\text{assoc})} < 1 \text{ GeV}/c : 0.953$
- 401 – $1 < p_{T(\text{assoc})} < 2 \text{ GeV}/c : 0.969$
- 402 – $2 < p_{T(\text{assoc})} < 3 \text{ GeV}/c : 0.982$
- 403 – $p_{T(\text{assoc})} > 3 \text{ GeV}/c : 0.990$

404 It was also verified with the same Monte Carlo study that applying the DCA selection rejects less than
 405 0.2% primary tracks (tagged as false positives) from the associated track sample, again with a flat az-
 406 imuthal distribution, inducing hence a fully negligible bias on the data correlation distributions. This is
 407 shown in Figure 22. This was also verified for specific charm-origin and beauty-origin tracks, due to
 408 their larger DCA with respect to primary tracks from light quarks. In this case, the fraction of rejected
 409 charm and beauty tracks stays below 1% in all the kinematic ranges apart from the associated track p_T
 410 regions 0.3-1 and $> 0.3 \text{ GeV}/c$, where the rejection can be as high as 2%. In these kinematic ranges,
 411 though, the data correlation distributions are dominated by non-heavy-flavour tracks, as it was verified
 412 from the simulations, hence the overall bias is still contained below 1%, thus negligible.

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

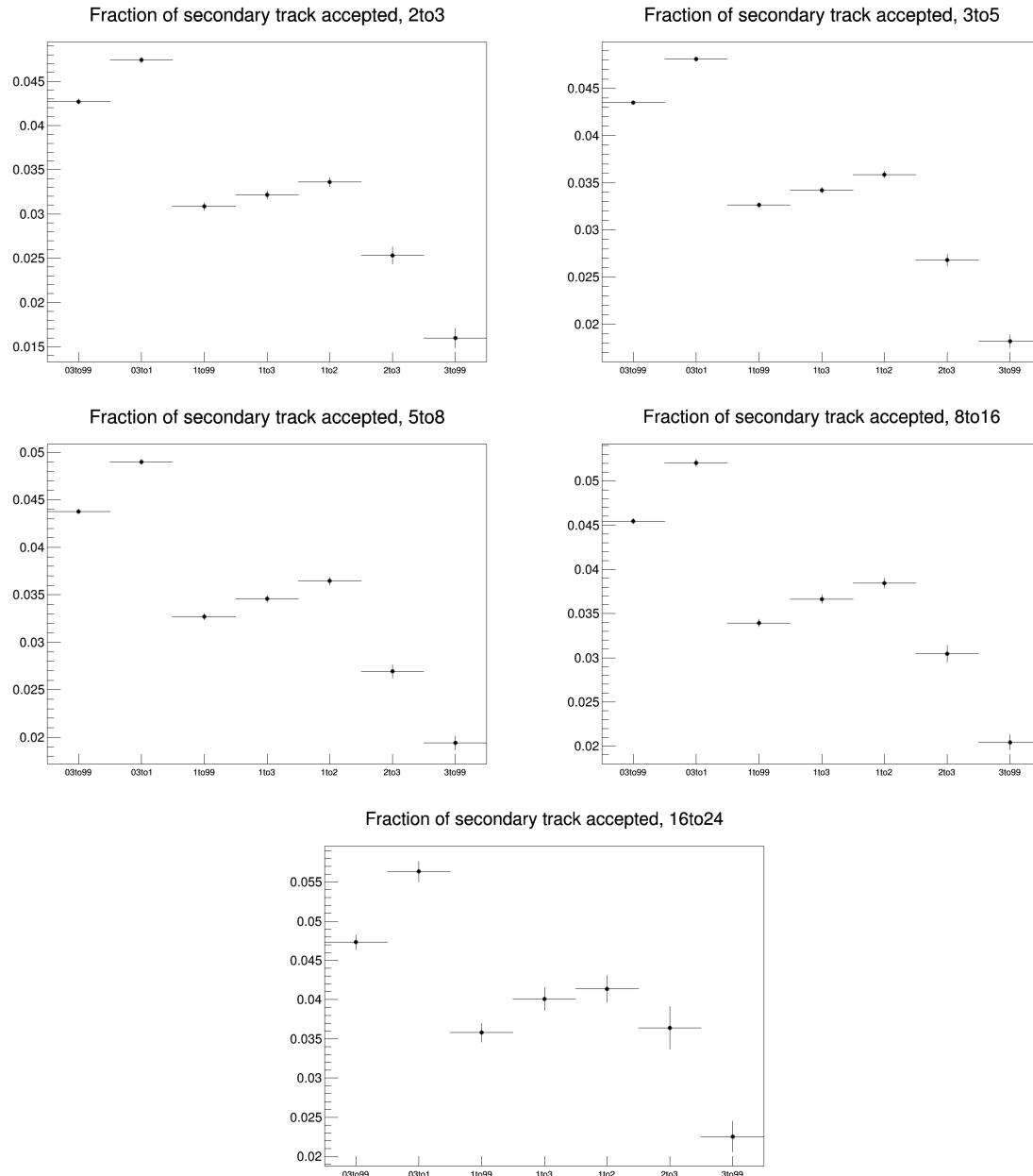


Figure 12: Fraction of secondary tracks over total amount of tracks which pass the DCA selection. The four panel show the fractions for the D-meson p_T ranges: 2-3, 3-5, 5-8, 8-16, 16-24, respectively. Inside each panel, the associated track p_T ranges are shown on the x -axis.

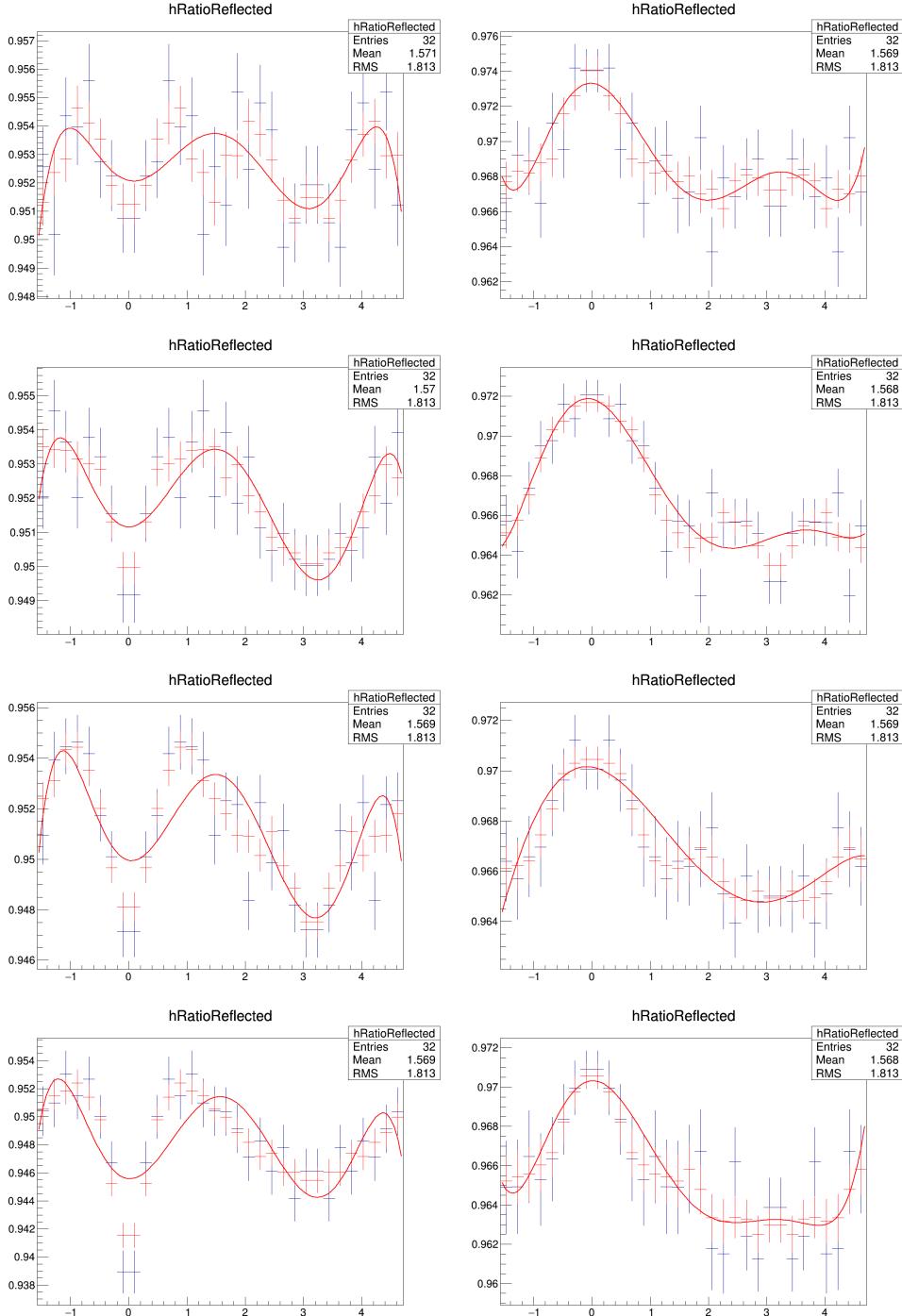


Figure 13: Fraction of primary track in the reconstructed associated track sample (blue histogram). The polynomial fit function (red curve) and the 3-point moving average (red histogram) are also superimposed. The $p_T(D)$ ranges are 2-3, 3-5, 5-8, 8-16 GeV/c, respectively for each row, and $0.3 < p_T(\text{assoc}) < 1$, $p_T(\text{assoc}) > 1$ GeV/c inside each row.

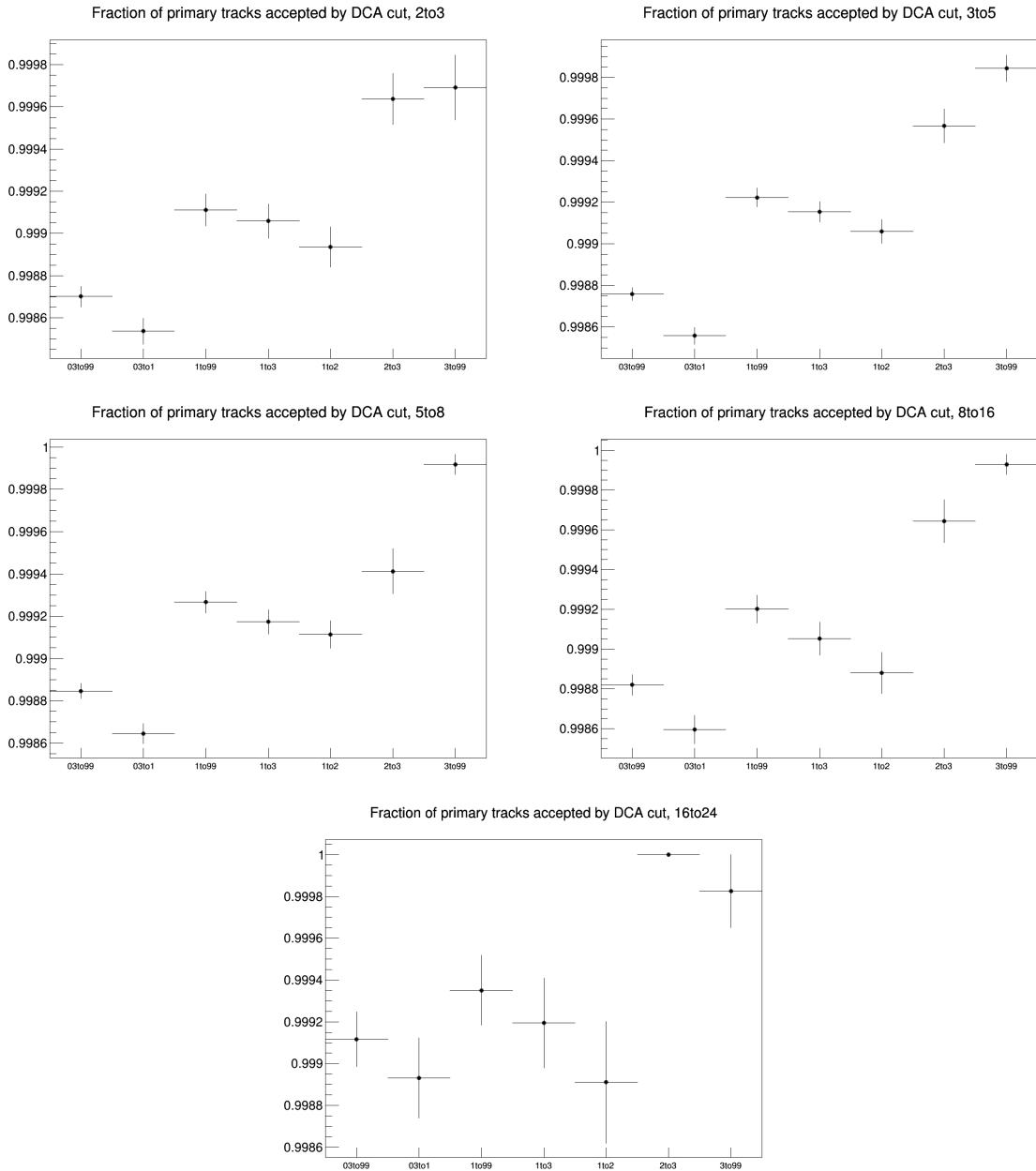


Figure 14: Fraction of primary tracks rejected by the DCA selection. The four panel show the fractions for the D-meson p_T ranges: 2-3, 3-5, 5-8, 8-16, 16-24, respectively. Inside each panel, the associated track p_T ranges are shown on the x -axis.

⁴¹⁷ **3.3.5 Beauty feed-down**

⁴¹⁸ The contribution of correlations of D meson from b-hadron decay is subtracted from the data correlation
⁴¹⁹ distributions as:

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

⁴²⁰ In the above equation, $\tilde{C}_{\text{inclusive}}(\Delta\varphi)$ and $\tilde{C}_{\text{prompt D}}(\Delta\varphi)$ are per-trigger azimuthal correlation distribu-
⁴²¹ tions before and after feed-down contribution subtraction, f_{prompt} is the fraction of prompt D meson
⁴²² and $\tilde{C}_{\text{feed-down}}^{\text{MC templ}}$ is a template of the azimuthal correlation distribution for the feed-down component ob-
⁴²³ tained from home-made Monte Carlo simulation at generated level, using PYTHIA6 with Perugia2011
⁴²⁴ tune. In order to avoid biases related to the different event multiplicity in real and simulated events,
⁴²⁵ the correlation distribution was shifted to have its minimum coinciding with the baseline of the data
⁴²⁶ azimuthal-correlation distribution before feed-down subtraction.

⁴²⁷ The value of f_{prompt} (Figure 23), which depends on D-meson species and varies as a function of the p_T ,
⁴²⁸ is estimated on the basis of FONLL predictions for the production of feed-down D mesons at central
⁴²⁹ rapidity, in pp collisions at $\sqrt{(s)} = 5$ TeV, and using the reconstruction efficiency of prompt and feed-
⁴³⁰ down D mesons, following the so-called N_b approach defined in [1]. Typical values are about 8-10% for
⁴³¹ the D^0 , about 4-7% for the D^+ and about 5-8% for the D^{*+} .

432 **4 Systematic uncertainties on $\Delta\phi$ correlation distributions**

433 **4.1 Uncertainty on S and B extraction**

434 The systematic uncertainty for the D meson yield extraction was determined separately for the three
 435 mesons. It was obtained by evaluating the value of the signal candidate from the invariant mass spectra
 436 with the following differences with respect to the standard approach:

- 437 – Changing the background fit function, for D^0 and D^+ (tried with polynomials of 1st and 2nd order)
 438 and for D^{*+} (tried with polynomials of 2nd order and a power function);
- 439 – Changing the range in which the signal is extracted from the Gaussian fit;
- 440 – Reducing the range of invariant mass axis in which the signal region is defined (and S and B are
 441 extracted);
- 442 – Rebinning the invariant mass distributions before the fit for D^0 and D^+
- 443 – Extracting S and B via integral of the fit functions or B via bin counting and S via integral of the
 444 Gaussian function.

445 Both the value of the yield and the sidebands correlations normalization factor are affected by changing
 446 the yield extraction approach, while the rest of the procedure to extract the azimuthal correlation dis-
 447 tribution is the same as in the standard analysis. The fully corrected azimuthal correlation plots were
 448 evaluated, for each of these approaches, in all D meson p_T bins and for each value of associated tracks
 449 p_T threshold. The ratios of the correlation distributions obtained with the standard yield extraction pro-
 450 cedure and by differentiating the approach were evaluated. From the average of these ratios, which
 451 are found to be flat versus $\Delta\phi$, a systematic uncertainty can be extracted, which was taken of 1% for
 452 $3 < p_T(D) < 16 \text{ GeV}/c$ and of 2% in $16 < p_T(D) < 24 \text{ GeV}/c$. No dependence versus the associated
 453 track p_T was assumed, since from a physics point of view we don't expect a modification of the signal
 454 and sideband values to have a dependence of this kind. Figures 26, show the ratios obtained by the
 455 above mentioned procedure for exemplary p_T ranges, which anyway span over the full kinematic ranges
 456 analyzed, for D^0 -h correlations. Figures 27 and 28 show the same ratios for D^{*+} -h, D^+ -h as well.

457 **4.2 Uncertainty on background correlation shape**

458 The systematic uncertainty for the subtraction of the background correlations includes the effects due to
 459 a potentially biased description of the background correlation shape, which is evaluated from of the side-
 460 bands correlations. In particular, the background correlation shape could present some hidden invariant
 461 mass dependence. To estimate this uncertainty, the invariant mass range of the sidebands definitions was
 462 varied with respect to the default values. For the D^0 meson, the usual range of the sidebands is 4 to 8 σ
 463 from the centre of the peak of the Gaussian fit and it was modified, for both sidebands to:

- 464 – inner half (4 to 6 σ from the centre of the peak);
- 465 – outer half (6 to 8 σ from the centre of the peak)
- 466 – extended to 4 to 10 σ (in case this is possible without exceeding the fitting range of the mass plots)

467 Slightly different variations, but with the same reasoning, were considered for the D^+ meson.

468 For the D^{*+} meson, the usual range of sideband in invariant mass spectra is 5 to 10 σ (only on the right
 469 side) from the centre of the peak of the Gaussian fit of the invariant mass spectra, and it was modified to:

- inner half (5 to 8 σ from the centre of the peak);
- outer half (8 to 13 σ from the centre of the peak);
- extended to 5 to 13 σ from the centre of the peak;
- extended to 6 to 16 σ from the centre of the peak.

The rest of the procedure for the azimuthal correlations distribution was unchanged, and the ratios of the fully corrected azimuthal correlation plots obtained with the standard sidebands range and the correlation plots extracted with different sidebands definitions, were evaluated for each D-meson p_T bin and associated tracks p_T threshold. Results of this check are shown in Figures 29, 30 and 31 for D^0 , D^{*+} , D^+ respectively, for exemplary p_T ranges, spanning over the full kinematic regions analysis. From the values of the ratios extracted from the checks, which do not show any azimuthal dependence a systematic uncertainty for the background subtraction can be evaluated. Also no dependence versus the associated track p_T was assumed also in this case. The uncertainty was hence taken of 1% for all the mesons in $3 < p_T(D) < 16 \text{ GeV}/c$ and of 2%, 2.5% and 4% (for D^+ , D^0 , D^{*+} respectively) in $16 < p_T(D) < 24 \text{ GeV}/c$.

4.3 Uncertainty on D-meson cut stability

To study the systematics due to the topological selections on the D meson, the cut variation approach was used. For each D-meson, alternate sets of released and tightened selection cuts were applied to extract the correlation distribution, varying in particular the cosine of the pointing angle, the maximum DCA among the daughter tracks and the product of the daughter track impact parameters. For each set of cuts new 2D (p_T vs multiplicity), D meson efficiency map was computed. In Figures 32, 33, 34 (for D^0 , D^{*+} and D^+ , respectively) the ratio of the different 1D efficiencies with the alternate cuts with respect to the default cut selection is chosen, to highlight how the different selections effectively varied the efficiency values, especially at low p_T , where cuts are more effective.

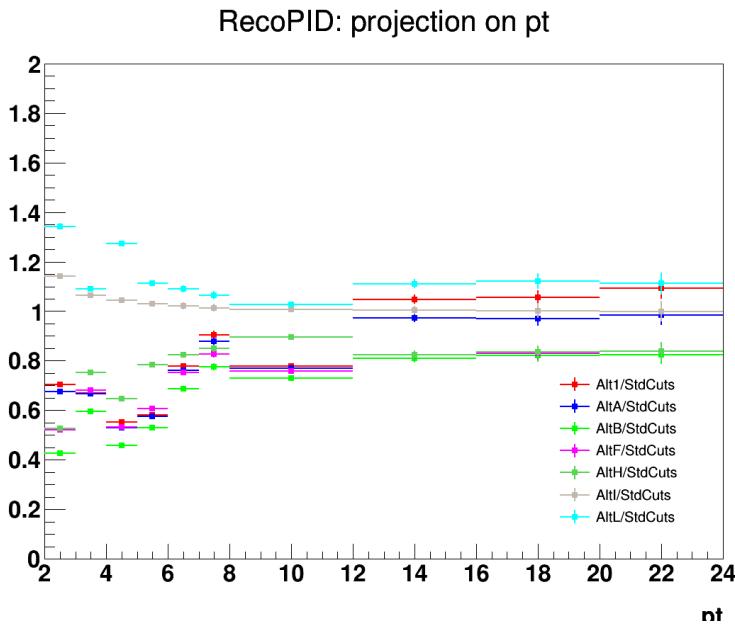


Figure 15: Ratio of D^0 efficiencies with alternate cut variations w.r.t. the standard cut used for the analysis.

Figure 35, 36, 37 show the ratio of the correlation distributions with alternate cut sets over those with the standard approach, for exemplary p_T ranges covering the full kinematic region of interest for the analyses.

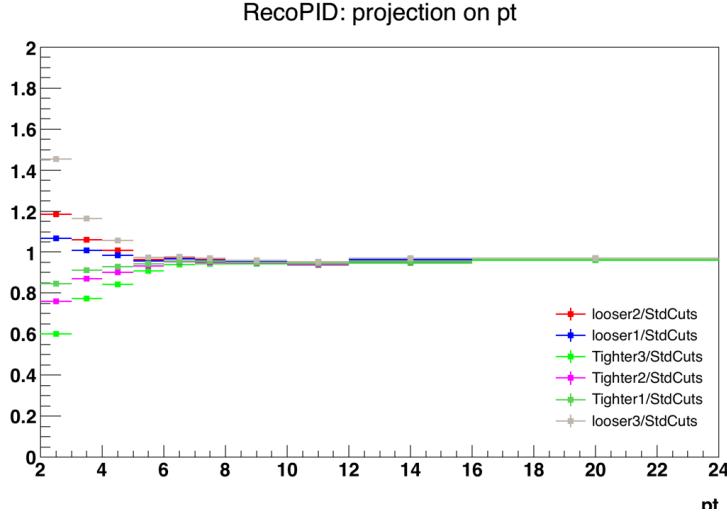


Figure 16: Ratio of D^{*+} efficiencies with alternate cut variations w.r.t. the standard cut used for the analysis.

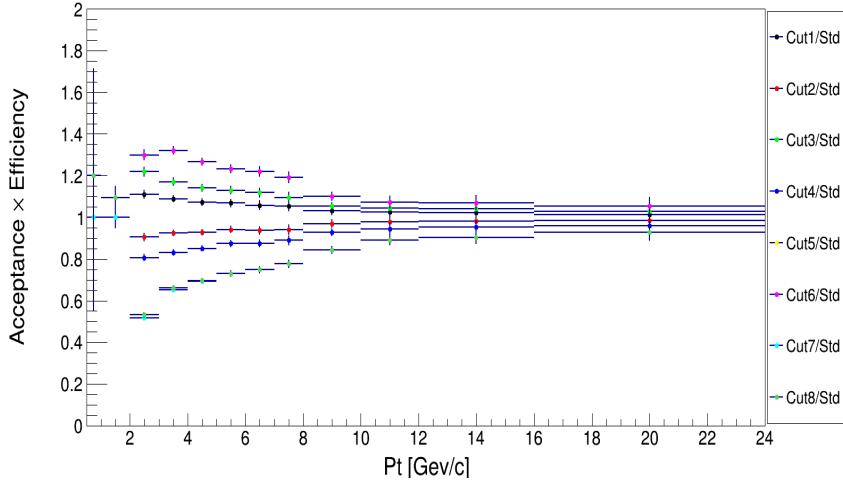


Figure 17: Ratio of D^+ efficiencies with alternate cut variations w.r.t. the standard cut used for the analysis.

495 The ratios are reasonably flat in $\Delta\phi$, hence a flat systematic was evaluated as systematic uncertainty from
 496 D-meson the cut variations. For the D^0 , the uncertainty was considered of 2% for all the p_T ranges of
 497 trigger and tracks analyzed. For the D^{*+} , the uncertainty was considered of 1.5% for $3 < p_T(D) < 8$
 498 GeV/ c and of 1% for $8 < p_T(D) < 24$ GeV/ c . For the D^+ , the uncertainty was considered of 1% for
 499 $3 < p_T(D) < 16$ GeV/ c and of 2% for $16 < p_T(D) < 24$ GeV/ c .

500 4.4 Uncertainty on tracking efficiency evaluation

501 The systematic uncertainty for the tracking efficiency includes the effects related to the set of filtering cuts
 502 defined for the associated tracks selection (mainly requests on the quality of reconstructed tracks for the
 503 TPC and ITS detectors). This uncertainty was determined by repeating the full analysis using different
 504 selections for the cuts on the associated tracks with respect to the usual selection (TPC only tracks with at
 505 least 2 points in the ITS). The alternative selections were: pure TPConly selection, meaning TPC tracks
 506 with no requests on the number of hits in the ITS, and TPC+ITS selection, which requires filterbit 4
 507 with, in addition, at least 3 points in the ITS, ITS refit and a hit in at least an SPD layer. The ratios of the
 508 azimuthal correlation distributions with different sets of tracks selection over distributions with standard
 509 selection were evaluated, and are shown in Figures 38 and 39 for D^0 -h correlations. Their values were

510 used to determine a systematic uncertainty, which as the previous ones could be assigned flat in $\Delta\phi$, and
 511 which was estimated of 3% in all the ranges of p_T analyzed.

512 4.5 Uncertainty on secondary particle contamination

513 Secondary particles, i.e. particles coming from strange hadrons decays or particles produced in inter-
 514 actions with the material, are expected to be tagged and removed by means of a distance of closest
 515 approach (DCA) from primary vertex cut. The uncertainty arising from the residual contamination of
 516 secondary tracks can be estimated from a Monte Carlo study, at reconstructed level. The number of
 517 primary/secondary tracks which are accepted/rejected from the DCA cut was determined for different
 518 values of the DCA selection, and the correlation distributions for the various cases were evaluated. The
 519 variations were done in the xy direction, where the DCA resolution is better, and the following cases
 520 were tried (in addition to the default 1 cm cut): 0.1 cm, 0.25 cm, 0.5 cm, filtering DCA cut (i.e. 2.4 cm).

521 Figure 40 shows the amount of secondary tracks which are accepted by the DCA cut, over the total
 522 number of tracks (primary and secondary) accepted by the selection, for the various DCA selections that
 523 were tried. This is shown for the exemplary case of $5 < p_T < 8 \text{ GeV}/c$ (there's no $p_T(D)$ dependence) and
 524 as a function of the associated track p_T ranges. Hence, this quantity represents the residual contamination
 525 of secondary tracks in our reconstructed track sample. From these values, the corresponding primary
 526 track purities (1-contamination) were extracted, in each of the momentum ranges. It was also verified
 527 that, for all the cut selections, the $\Delta\phi$ distributions of the residual contaminations were flat within 1%.

528 As a second step of the procedure to verify the DCA cut stability, the D^0 -h data analysis was performed
 529 with all the different DCA selection (each time with the proper tracking efficiency map). After having
 530 extracted the correlation distributions, these were rescaled for the corresponding purities and compared
 531 with the purity-corrected correlation distributions obtained with the standard DCA selection. The ratios
 532 of the alternate selections over the standard selection, after the purity correction of both, are shown in
 533 Figures 41 and 42.

534 The ratios show a flat trend along the $\Delta\phi$ axis and, in general, a discrepancy from the value of 1 of no
 535 more than 3% (the worst case being the $0.3\text{-}1 \text{ GeV}/c$ range for the associated track). Hence, a flat and
 536 symmetric 3% systematical uncertainty on the evaluation of the secondary contamination was assigned
 537 on the base of this check in $0.3\text{-}1 \text{ GeV}/c$, reduced to 2.5% in $> 0.3 \text{ GeV}/c$ and to 1.5% for the other
 538 ranges. This amount also covers possible biases in the estimation of the purity (the $\Delta\phi$ distribution of
 539 the residual contamination is always contained inside 1%, as previously said).

540 4.6 Uncertainty on feed-down subtraction

541 As described in the 3.3.5 section, the feed-down subtraction from the data distributions is performed by
 542 means of simulation templates of $B \rightarrow D$ -h correlation distributions from PYTHIA6 generator, with Perugia2011 tune,
 543 and considering the central value of f_{prompt} to extract the feed-down D-meson contribution.
 544 In order to evaluate a systematic uncertainty on this procedure, the feed-down subtraction procedure was
 545 repeated considering, together with PYTHIA6+Perugia2011 templates, also PYTHIA6+Perugia2010
 546 and PYTHIA8 simulations. In each case, not only the central value of the measured f_{prompt} was con-
 547 sidered to rescale the distributions, but also the maximum and minimum values of its total uncertainty.

548 Then, the envelope of nine the different cases obtained by varying the templates and the f_{prompt} assumption
 549 was considered, and a value of the systematics defined as the envelope spread divided by $\sqrt{3}$ was
 550 taken as systematic uncertainty. This uncertainty was assumed uncorrelated among the different $\Delta\phi$
 551 points.

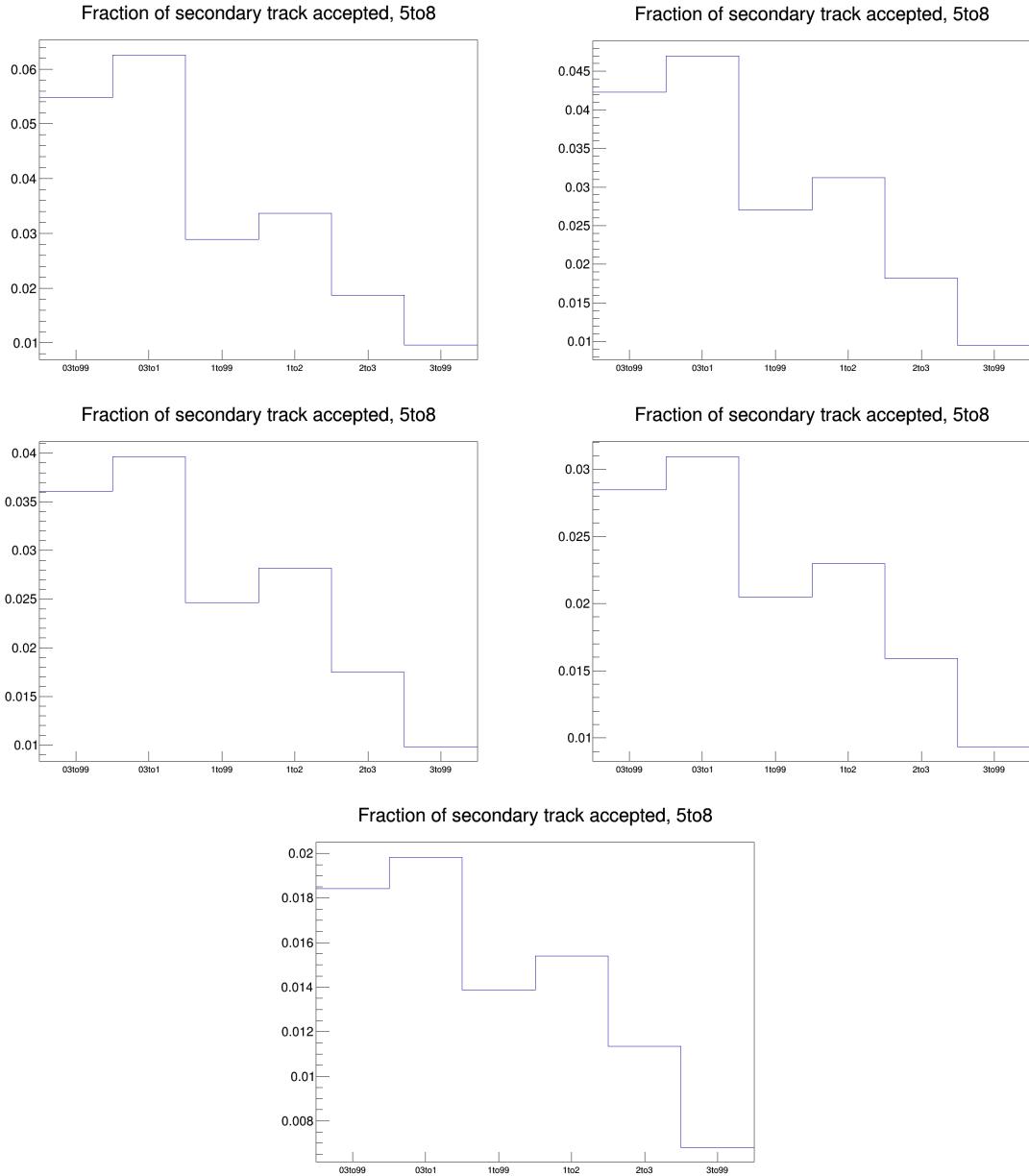


Figure 18: Secondary track contamination as a function of the associated track p_T , for the various DCA selections tried. The plots are ordered from the loosest to the tightest selection, i.e.: $\text{DCA}(xy) < 2.4 \text{ cm}$, $< 1 \text{ cm}$, $< 0.5 \text{ cm}$, $< 0.25 \text{ cm}$, $< 0.1 \text{ cm}$.

552 4.7 Uncertainty on correction for the bias on B to D decay topologies

553 The evaluation of this systematic uncertainty was already explained in Section 3.3.3. For each of the
 554 five data points close to the center of the near-side peak, which are affected by the bias, a bilateral and
 555 symmetric uncertainty of amplitude $|C(\Delta\varphi)_{\text{corr}} - C(\Delta\varphi)_{\text{raw}}|/\sqrt{12}$ was assigned.

556 This because the uncorrected data points are expected to be the extreme (with the current D-meson
 557 selection, the bias is always upwards at the centre of the peak, and always upwards on its sides). We
 558 then assume that, if the correction is properly evaluated, the corrected data points are at the centre of the
 559 possible spread of the true unbiased results. In this case, the span of the possible true results (in case
 560 of underestimation/overestimation of the bias) goes from the uncorrected data points to its symmetric
 561 value, with respect to the corrected data point, on the other direction. If this distribution is uniform,

562 and constrained by these two values, the 1σ confidence region for the position of the is in a bilateral
 563 $|C(\Delta\varphi)_{\text{corr}} - C(\Delta\varphi)_{\text{raw}}|/\sqrt{12}$ window, centered on the $C(\Delta\varphi)_{\text{corr}}$ points.

564 This source of uncertainty was assumed uncorrelated among the $\Delta\varphi$ points.

565 **4.8 Summary table**

566 A summary of the $\Delta\varphi$ -correlated uncertainties affecting the correlation distributions is show in Figure
 567 43. They are the S and B extraction uncertainty, the background shape uncertainty, the cut variation
 568 uncertainty, the tracking efficiency uncertainty and the secondary particle contamination uncertainty.

569 The overall amount of $\Delta\varphi$ -correlated uncertainties is about 5-6% (depending on the p_T bin) for the single
 570 D-meson cases; when evaluating the averages of the distributions (see next section), this uncertainty
 571 shrinks to 4-5%. This uncertainty is a global scale factor of the distributions, and is quoted as a label in
 572 the plots.

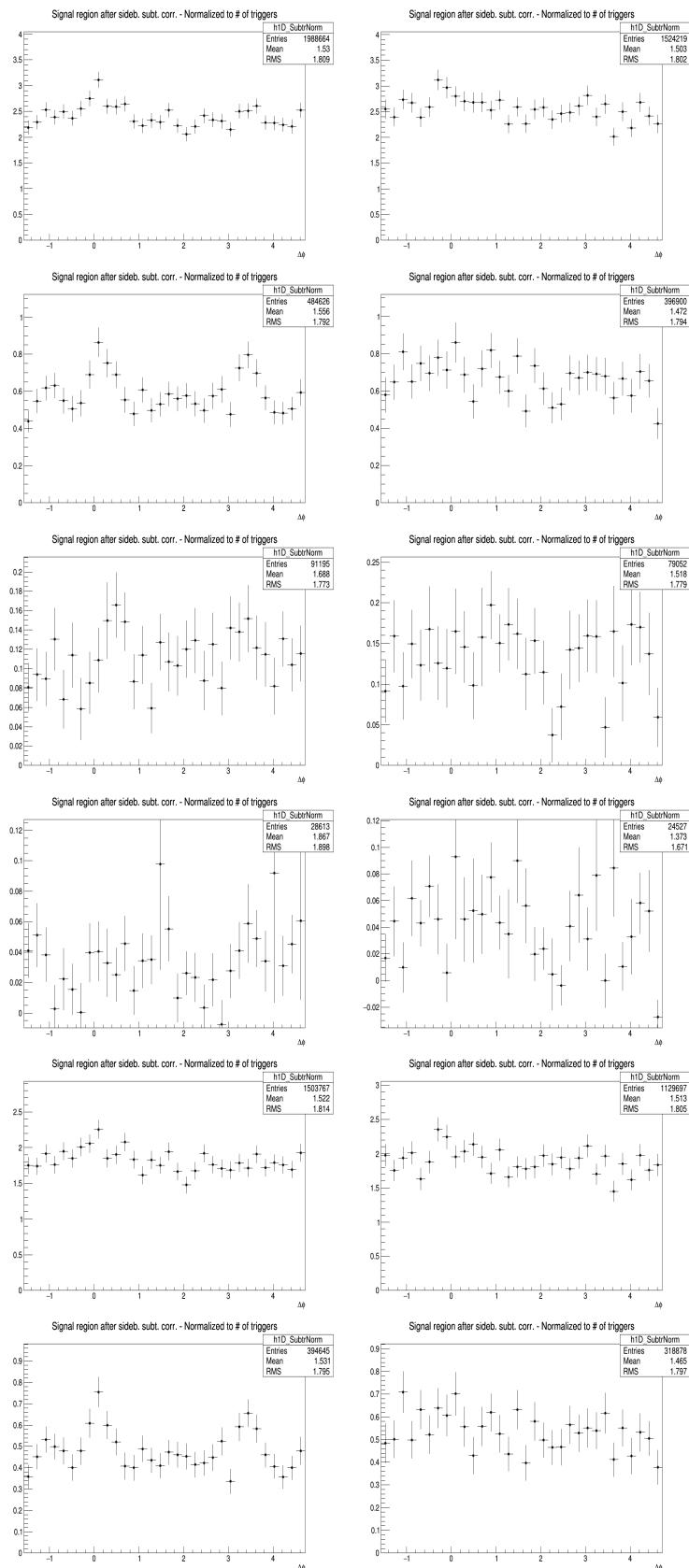
573 The systematics uncertainties from feed-down subtraction and $B \rightarrow D$ decay topology bias, instead are
 574 $\Delta\varphi$ dependent, and are hence reported as uncorrelated boxes in the plots. They do not amount to more
 575 than 4%, in every bin of all the kinematic ranges studied.

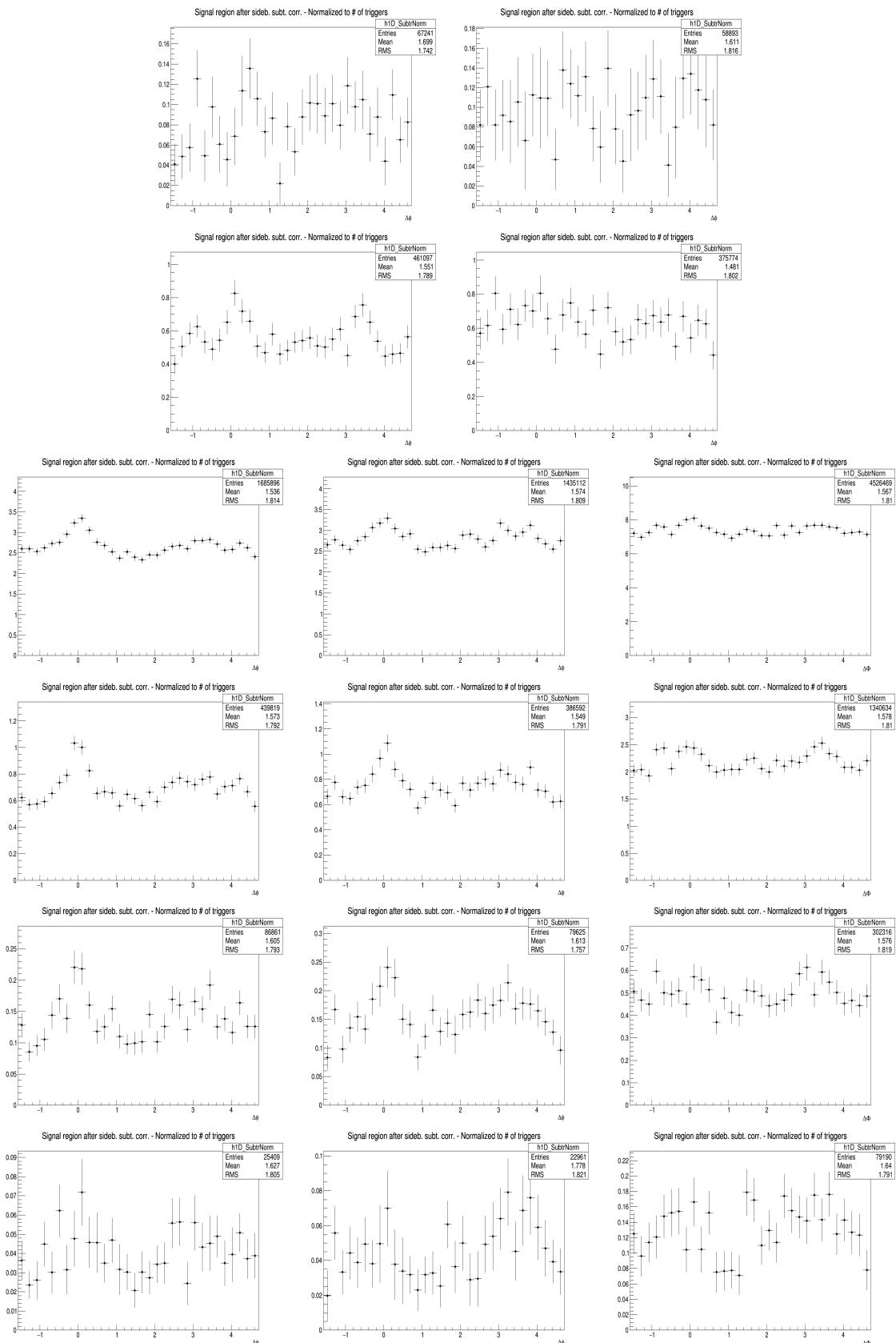
576 **5 Results**

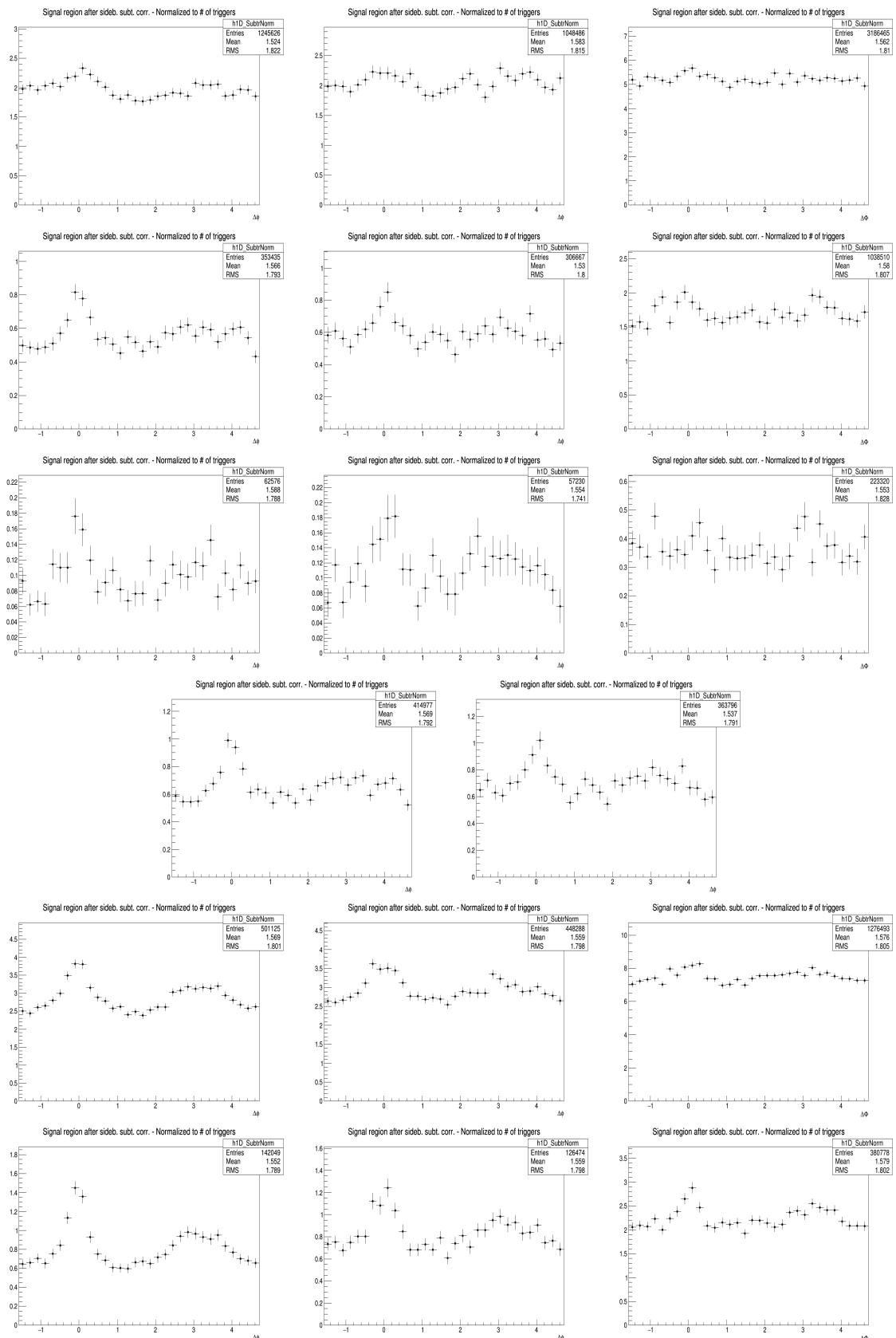
577 **5.1 Comparing the three D meson correlation distributions**

578 To check the compatibility of three D meson analyses, Figure 44 shows the corrected azimuthal cor-
 579 relation distributions (except for the feed-down subtraction and the secondary contamination removal)
 580 for D^0 -h, D^{*+} -h and D^+ -h, in each column, on the data sample used in the analysis. Results are shown
 581 for $2 < D p_T < 3 \text{ GeV}/c$, $3 < D p_T < 5 \text{ GeV}/c$, $5 < D p_T < 8 \text{ GeV}/c$, $8 < D p_T < 16 \text{ GeV}/c$ and
 582 $16 < D p_T < 24 \text{ GeV}/c$ with associated tracks $p_T > 0.3$, $p_T > 1$, $0.3 < p_T < 1 \text{ GeV}/c$, $1 < p_T < 2 \text{ GeV}/c$,
 583 $2 < p_T < 3 \text{ GeV}/c$ and $p_T > 3 \text{ GeV}/c$.

584 Figures 45, 46, 47, 48 show the superimposed correlation distributions from the single-meson analyses
 585 (same plots as previous figure) for better visualize the agreement among the different D-meson species
 586 results.

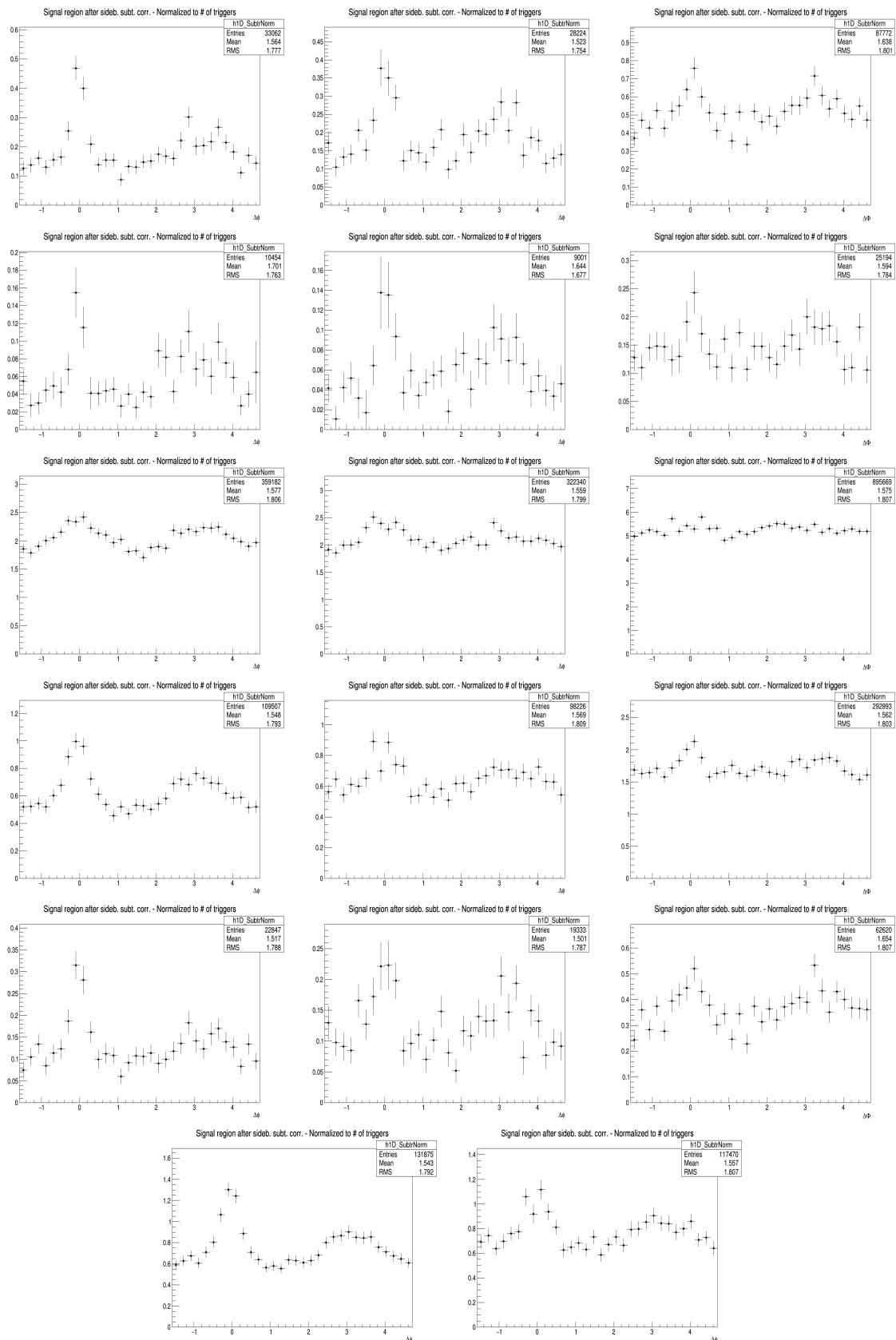


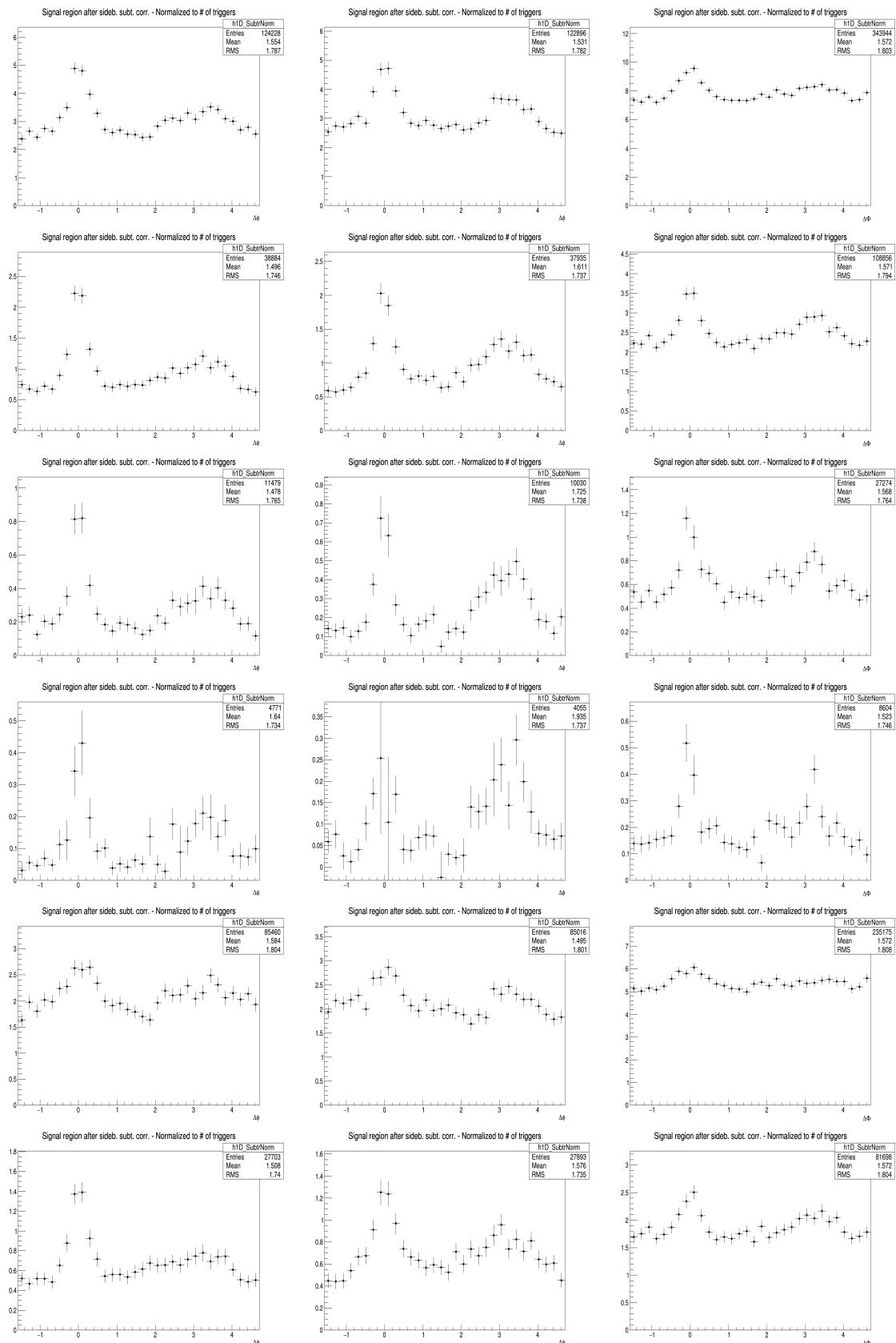


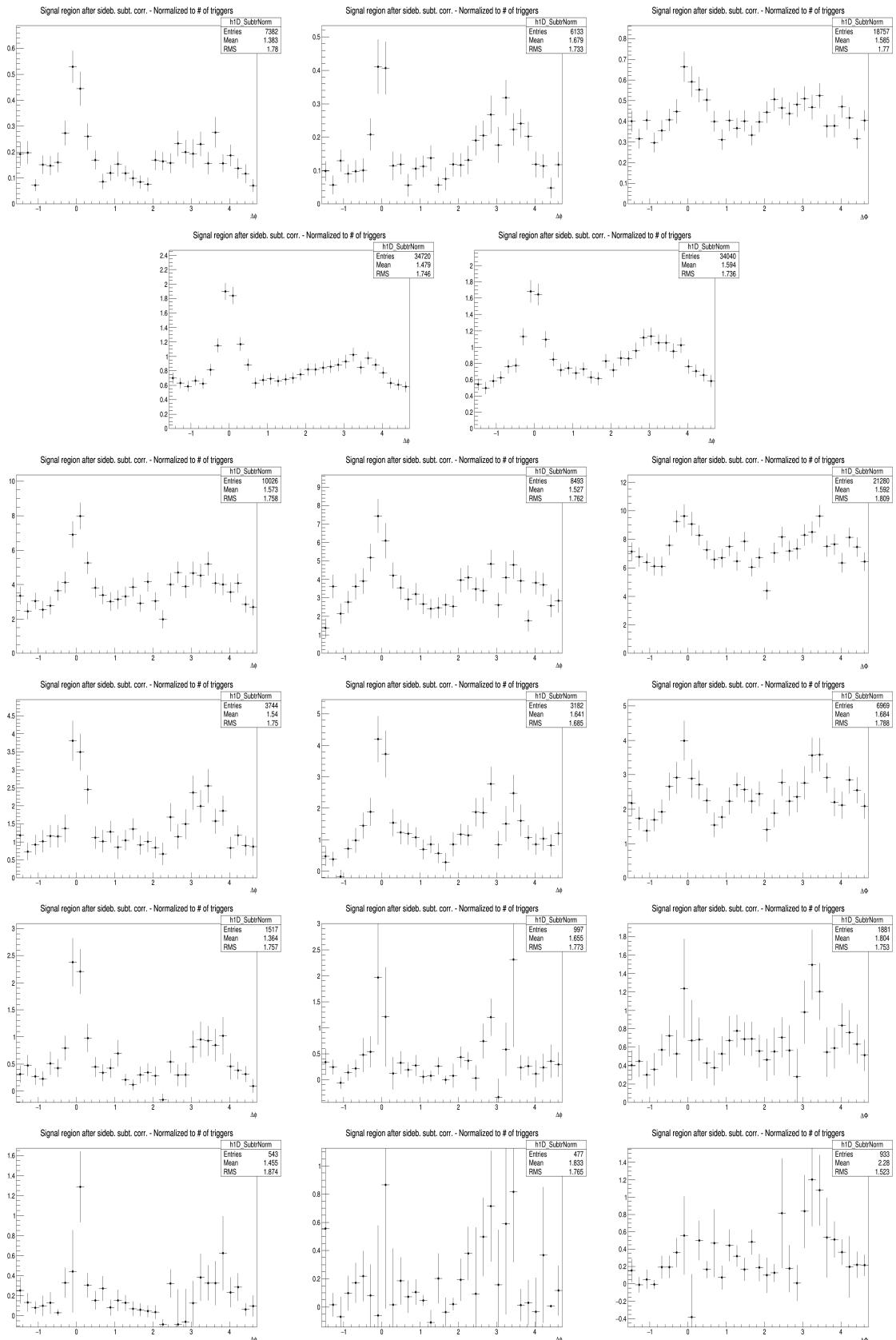


5.1 Comparing the three D meson correlation distributions

43







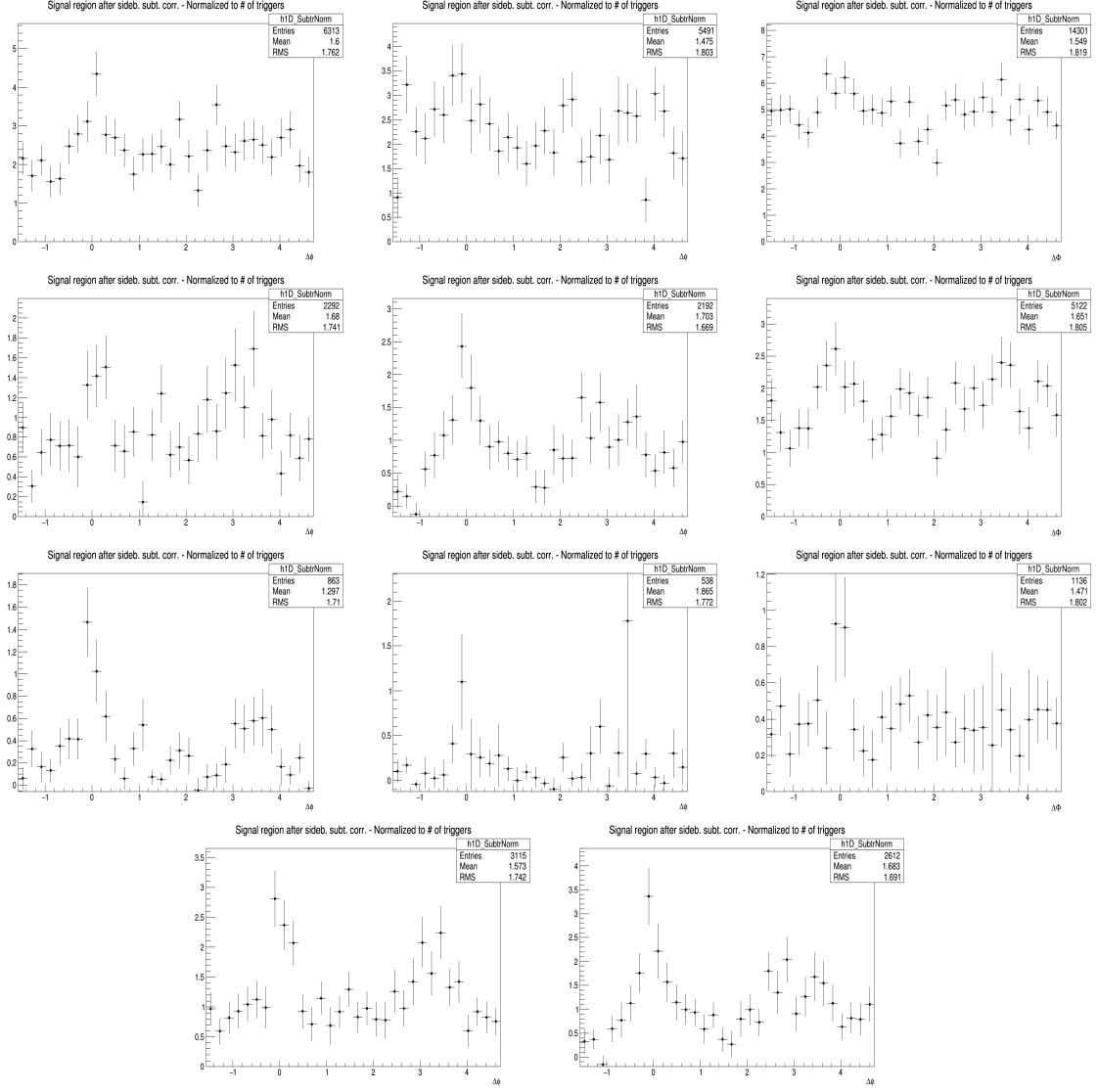


Figure 19: Corrected distribution of D-hadrons azimuthal correlations for the three species (apart from feed-down and purity), from analysis on the data sample, for the analyzed D-meson (**Column-Left:** D^0 , **Column-Middle:** D^+ and **Column-Right:** D^{*+}) and different associated tracks p_T ranges (**Row 1-7:** $3 < Dp_T < 5 \text{ GeV}/c$, $p_T(\text{Assoc}) > 0.3, >1.0, >2.0, >3.0, 0.3-1.0, 1.0-2.0 \text{ and } 2.0-3.0 \text{ GeV}/c$ respectively), (**Row 8-14:** $5 < Dp_T < 8 \text{ GeV}/c$, $p_T(\text{Assoc}) > 0.3, >1.0, >2.0, >3.0, 0.3-1.0, 1.0-2.0 \text{ and } 2.0-3.0 \text{ GeV}/c$ respectively), (**Row 15-21:** $8 < Dp_T < 16 \text{ GeV}/c$, $p_T(\text{Assoc}) > 0.3, >1.0, >2.0, >3.0, 0.3-1.0, 1.0-2.0 \text{ and } 2.0-3.0 \text{ GeV}/c$ respectively) and (**Row 22-28:** $16 < Dp_T < 24 \text{ GeV}/c$, $p_T(\text{Assoc}) > 0.3, >1.0, >2.0, >3.0, 0.3-1.0, 1.0-2.0 \text{ and } 2.0-3.0 \text{ GeV}/c$ respectively)

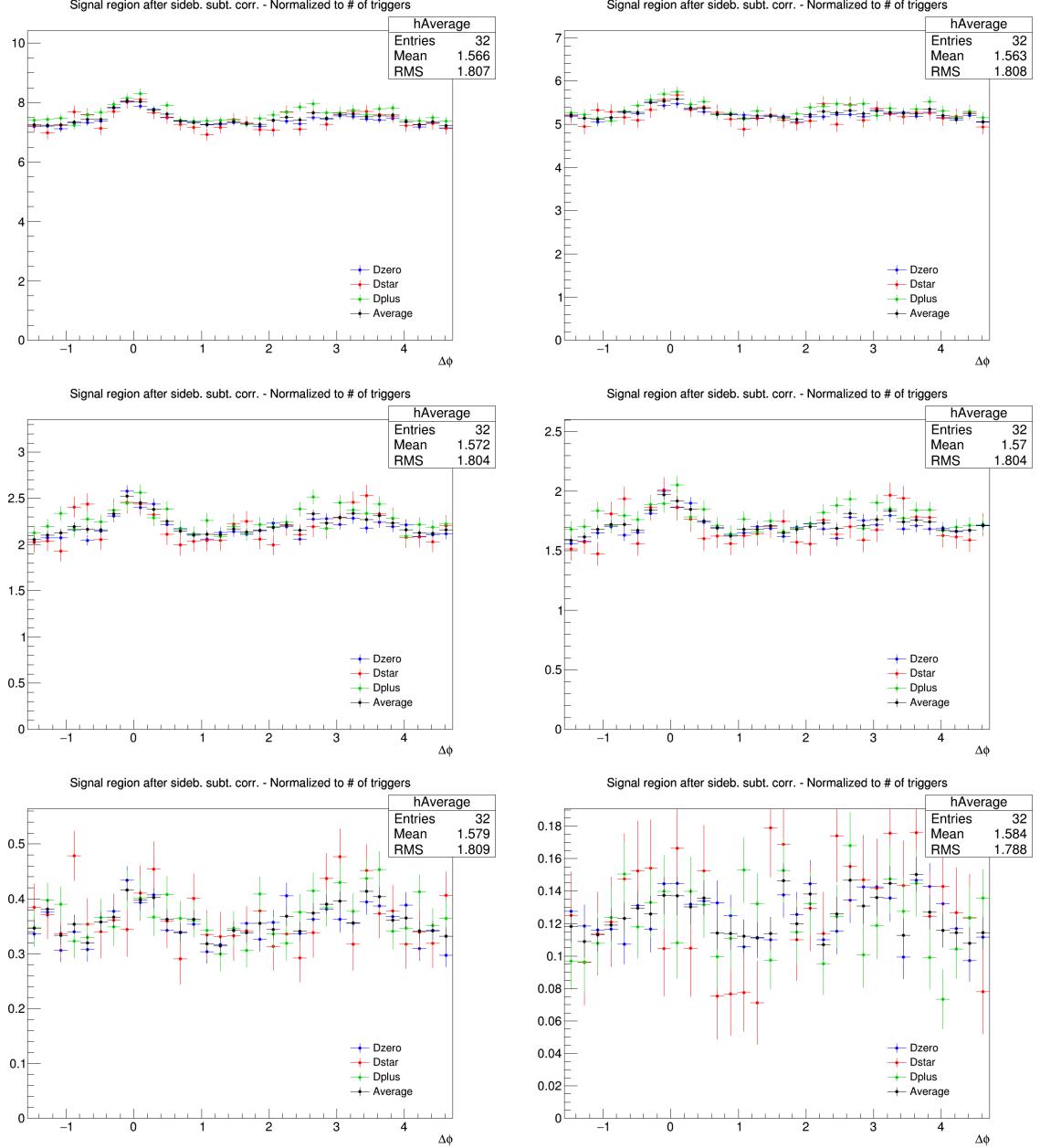


Figure 20: Superimposition of the corrected distribution of D-hadrons azimuthal correlations for the three species (apart from feed-down and purity), from analysis on the data sample, for the analyzed D-meson and different associated track p_T ranges, and D-meson p_T ranges (3-5 GeV/c on this page). **Panels from 1 to 6 of each page:** $p_T(\text{Assoc}) > 0.3, 0.3-1.0, >1.0, 1.0-2.0, 2.0-3.0 and $>3.0 \text{ GeV}/c$$

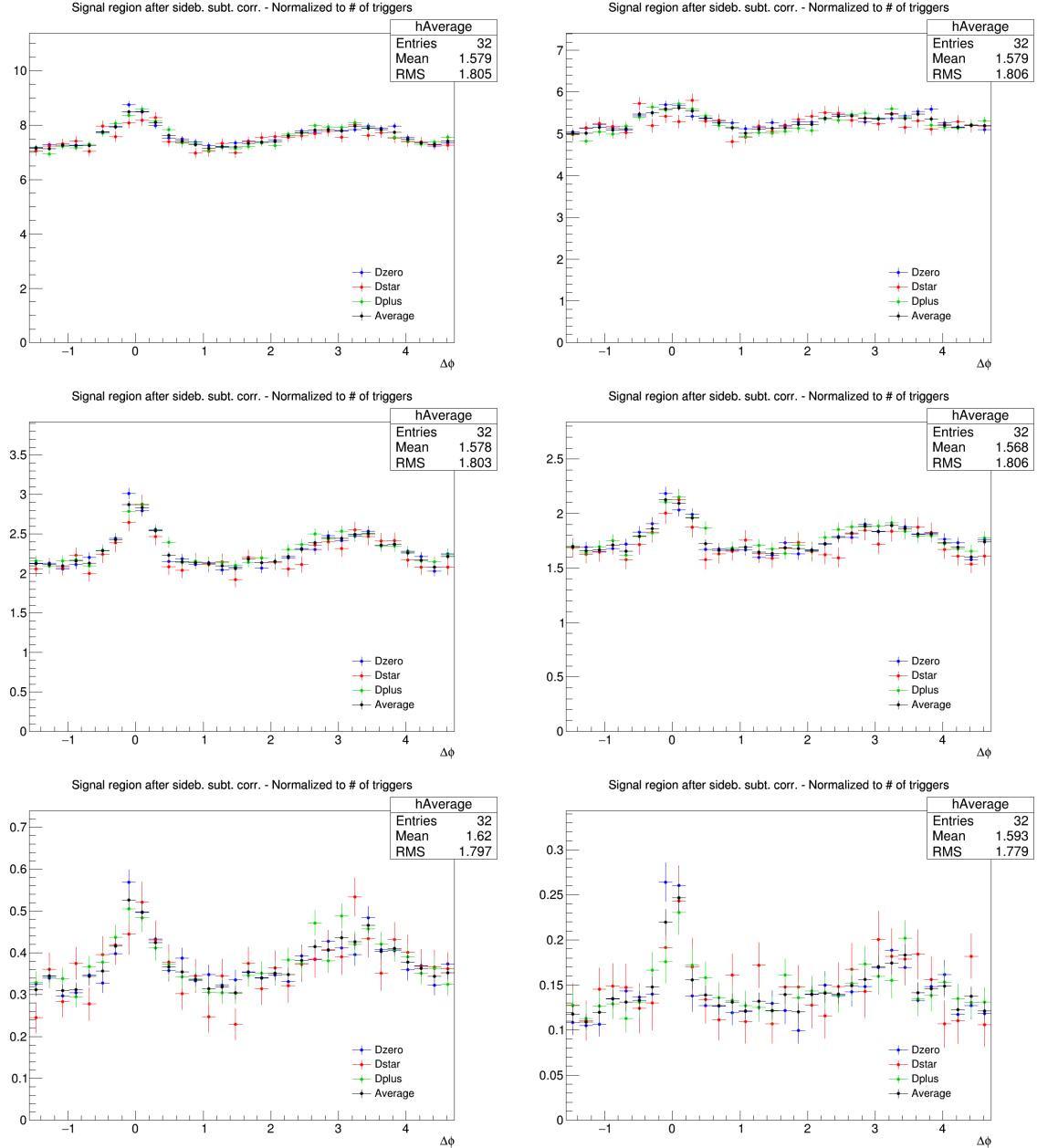


Figure 21: Superimposition of the corrected distribution of D-hadrons azimuthal correlations for the three species (apart from feed-down and purity), from analysis on the data sample, for the analyzed D-meson and different associated track p_T ranges, and D-meson p_T ranges (5-8 GeV/c on this page). **Panels from 1 to 6 of each page:** p_T (Assoc) > 0.3 , $0.3-1.0$, > 1.0 , $1.0-2.0$, $2.0-3.0$ and > 3.0 GeV/c

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

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

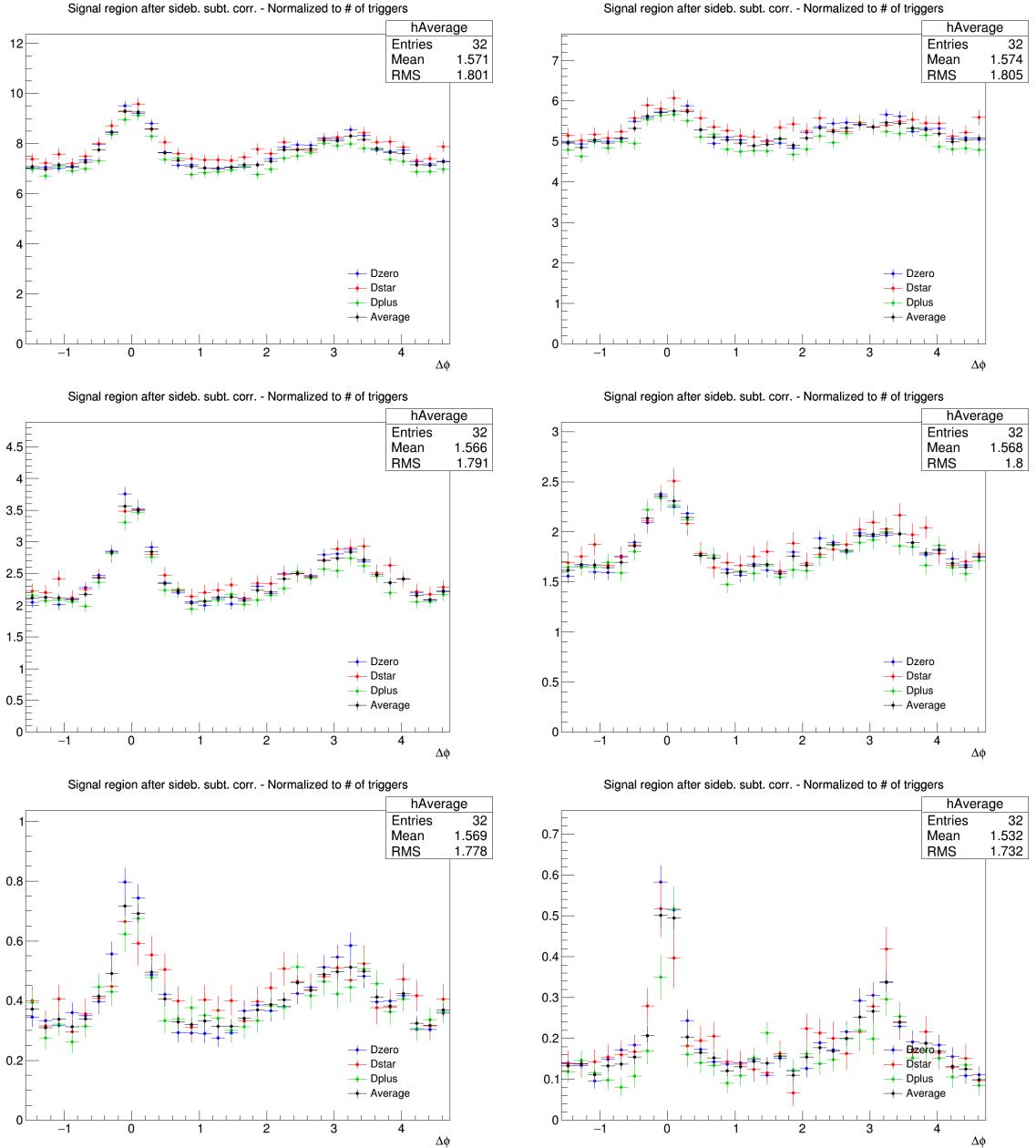


Figure 22: Superimposition of the corrected distribution of D-hadrons azimuthal correlations for the three species (apart from feed-down and purity), from analysis on the data sample, for the analyzed D-meson and different associated track p_T ranges, and D-meson p_T ranges (8-16 GeV/c on this page). **Panels from 1 to 6 of each page:** $p_T(\text{Assoc}) > 0.3, 0.3-1.0, >1.0, 1.0-2.0, 2.0-3.0$ and $>3.0 \text{ GeV}/c$

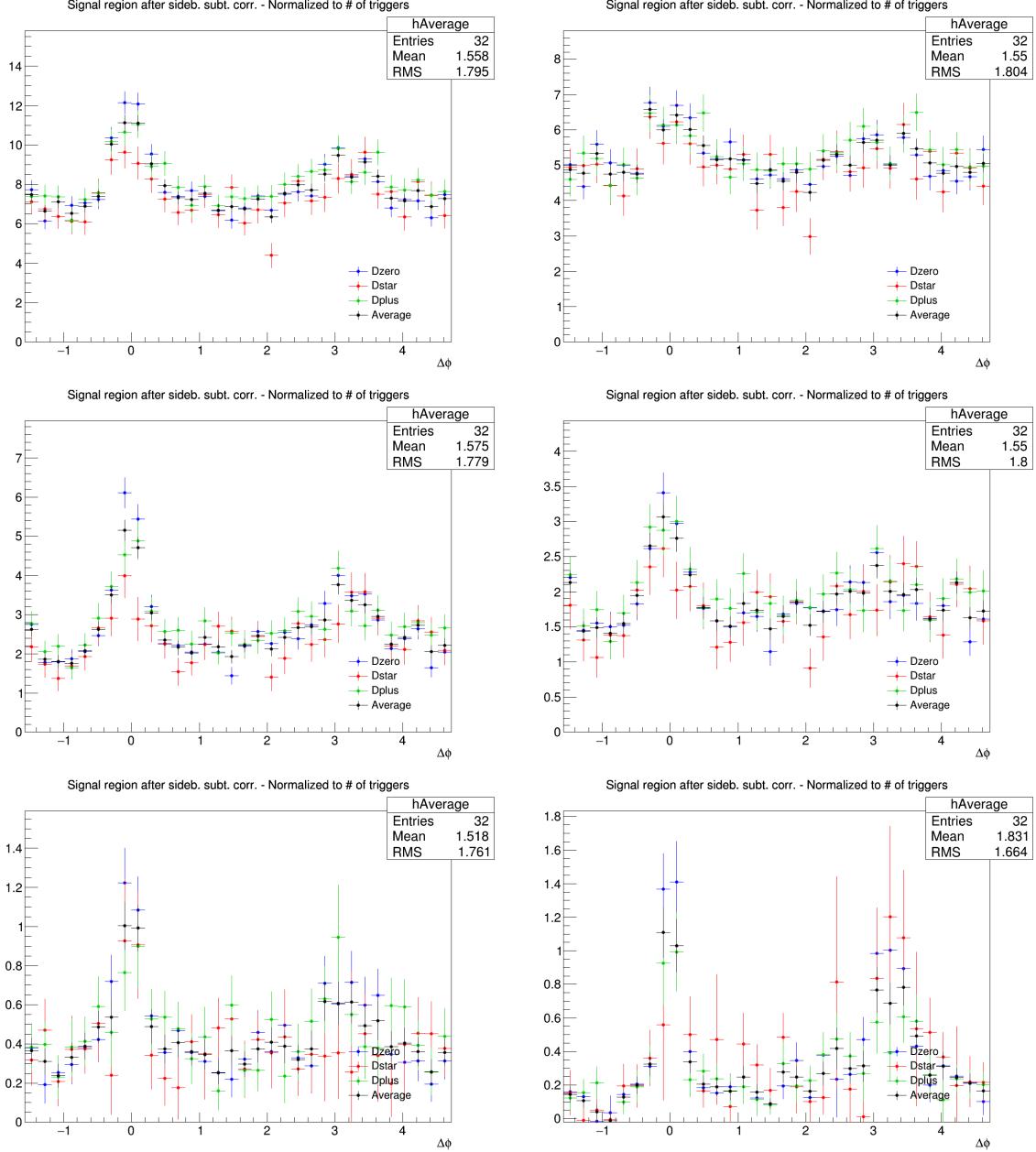


Figure 23: Superimposition of the corrected distribution of D-hadrons azimuthal correlations for the three species (apart from feed-down and purity), from analysis on the data sample, for the analyzed D-meson and different associated track p_T ranges, and D-meson p_T ranges (16-24 GeV/c on this page). **Panels from 1 to 6 of each page:** $p_T(\text{Assoc}) > 0.3, 0.3-1.0, >1.0, 1.0-2.0, 2.0-3.0$ and $>3.0 \text{ GeV}/c$

593 **5.2 Average of D⁰, D⁺ and D*⁺ results**

594 Given the compatibility within the uncertainties among the D⁰, D⁺ and D*⁺ azimuthal correlations, and
 595 since no large differences are visible in the correlation distributions observed in Monte Carlo simulations
 596 based on Pythia with Perugia0, 2010 and 2011 tunes¹, it was possible to perform a weighted average
 597 (eq. 5) of the azimuthal correlation distributions of D⁰, D⁺ and D*⁺, in order to reduce the overall
 598 uncertainties. Although some correlation between the mesons could be present (about the 30% of the
 599 D⁰, and also part of the D⁺, come from D*⁺ decays), the three selected D-meson samples can be treated
 600 as uncorrelated. The sum of the statistical uncertainties; the systematics uncertainty on S and B extraction
 601 and on background shape, are added in quadrature and the inverse of this sum was used as weight, w_i .

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

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

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

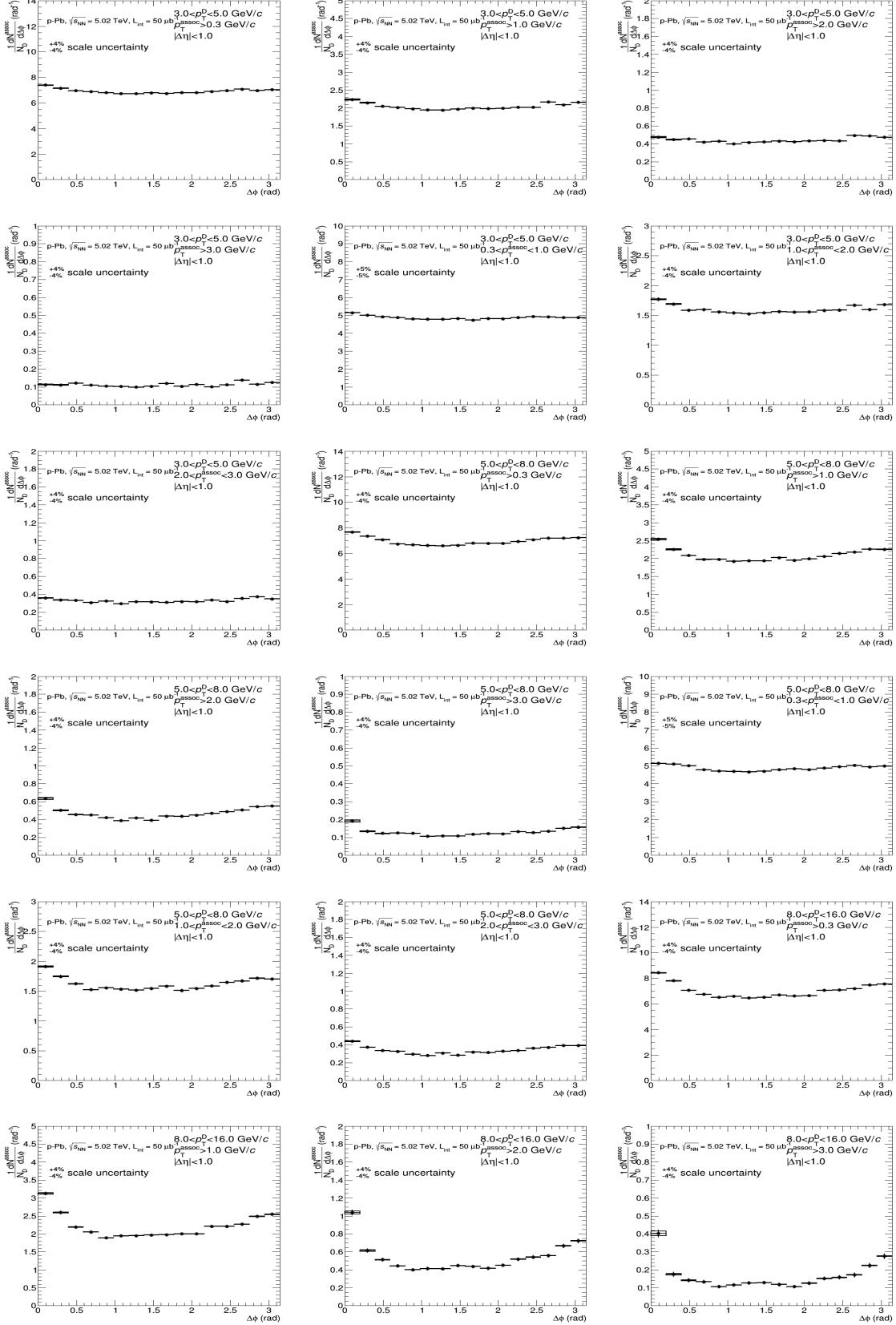
604 where n_D is the number of mesons considered in the average. It can be observed that for $\sigma_i^2 = 1/w_i$ the
 605 formula coincides with the standard one giving the uncertainty on a weighted average. The contribution
 606 to the average systematic uncertainty for those uncertainty sources not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

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

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

613 Figure 49 shows the averages of the azimuthal correlation distributions of D⁰, D⁺ and D*⁺ and charged
 614 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, $2 < p_T <$
 615 3 GeV/c, $p_T < 3$ GeV/c in the D meson p_T ranges $3 < p_T < 5$ GeV/c, $5 < p_T < 8$ GeV/c, $8 < p_T <$
 616 16 GeV/c and $16 < p_T < 24$ GeV/c. As expected, a rising trend of the height of the near-side peak
 617 with increasing D-meson p_T is observed, together with a decrease of the baseline level with increasing
 618 p_T of the associated tracks. To further increase the statistical precision on the averaged correlation
 619 distributions, given the symmetry around 0 on the azimuthal axis, the distributions were reflected and
 620 shown in the range $[0, \pi]$. This reduces the statistical uncertainty on the points by, approximately, a factor
 621 of $1/\sqrt{2}$.

¹A slight near side hierarchy is present among the three meson results, with D*⁺ meson having a lower peak amplitude than D⁰ and D⁺. It was verified that this is induced by the presence of D⁰ 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.



622 The usage of weighted average requires, as an underlying assumption, identical results expected for
 623 different species (or, at least, compatible within the uncertainties). Anyway, it was also verified that the
 624 usage of the arithmetic average instead of the weighted average increases the uncertainties on the points,

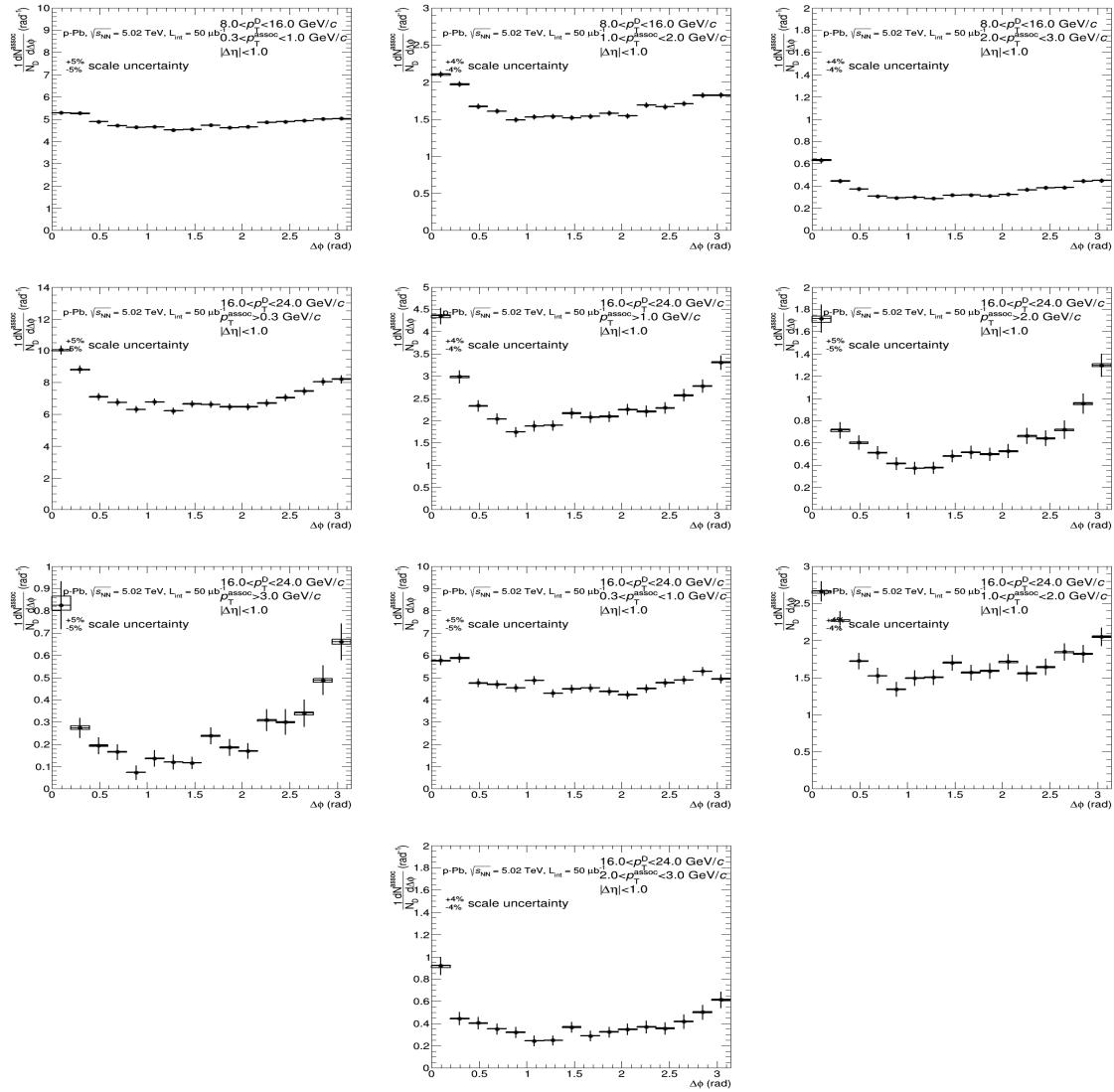


Figure 24: Average of D^0 , D^+ and D^{*+} azimuthal correlation distributions, in the D meson p_T ranges $3 < p_T < 5$ GeV/ c , $5 < p_T < 8$ GeV/ c , $8 < p_T < 16$ GeV/ c and $16 < p_T < 24$ GeV/ c , with associated tracks with $p_T > 0.3$ GeV/ c , $p_T > 1$ GeV/ c and $0.3 < p_T < 1$ GeV/ c .

but produces a negligible shift of their central values.

5.3 Fit observable p_T trends and uncertainties

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

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

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

An example of the results from the fit is shown in Figure 50

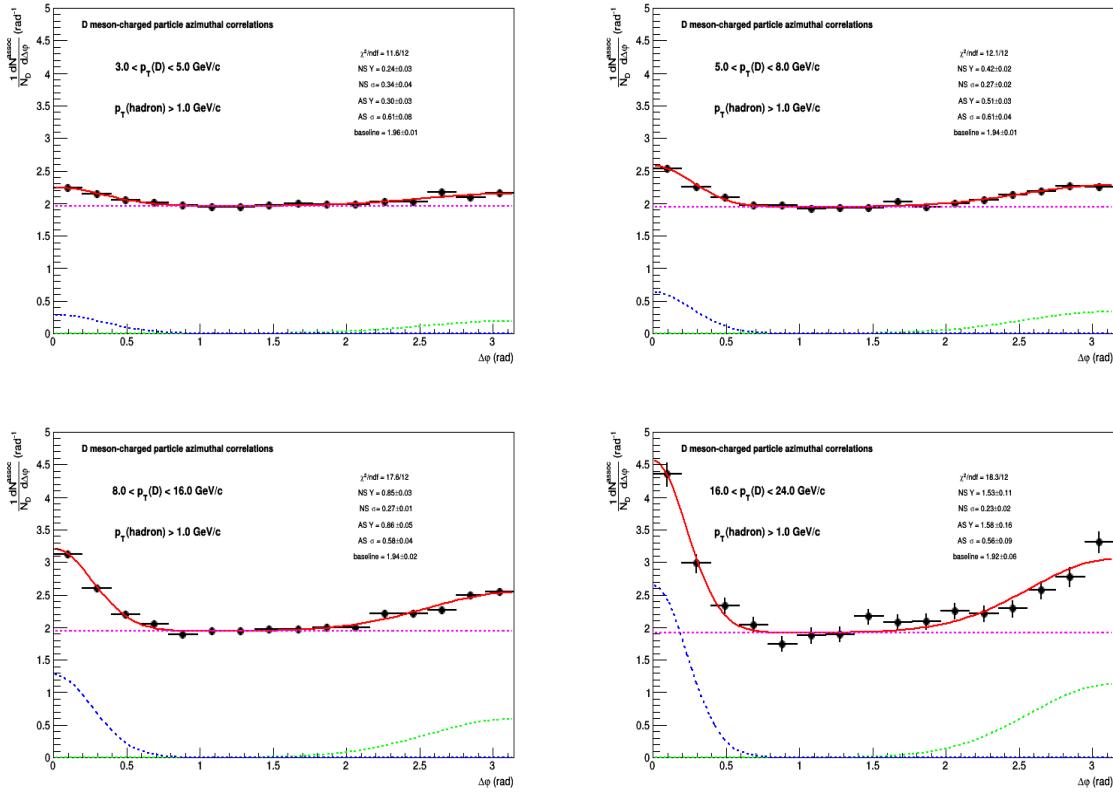


Figure 25: Example of fit to azimuthal correlation distributions and baseline estimation.

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

below the baseline level also charm and beauty-related pairs are contained (especially in cases of NLO production for the heavy quarks).

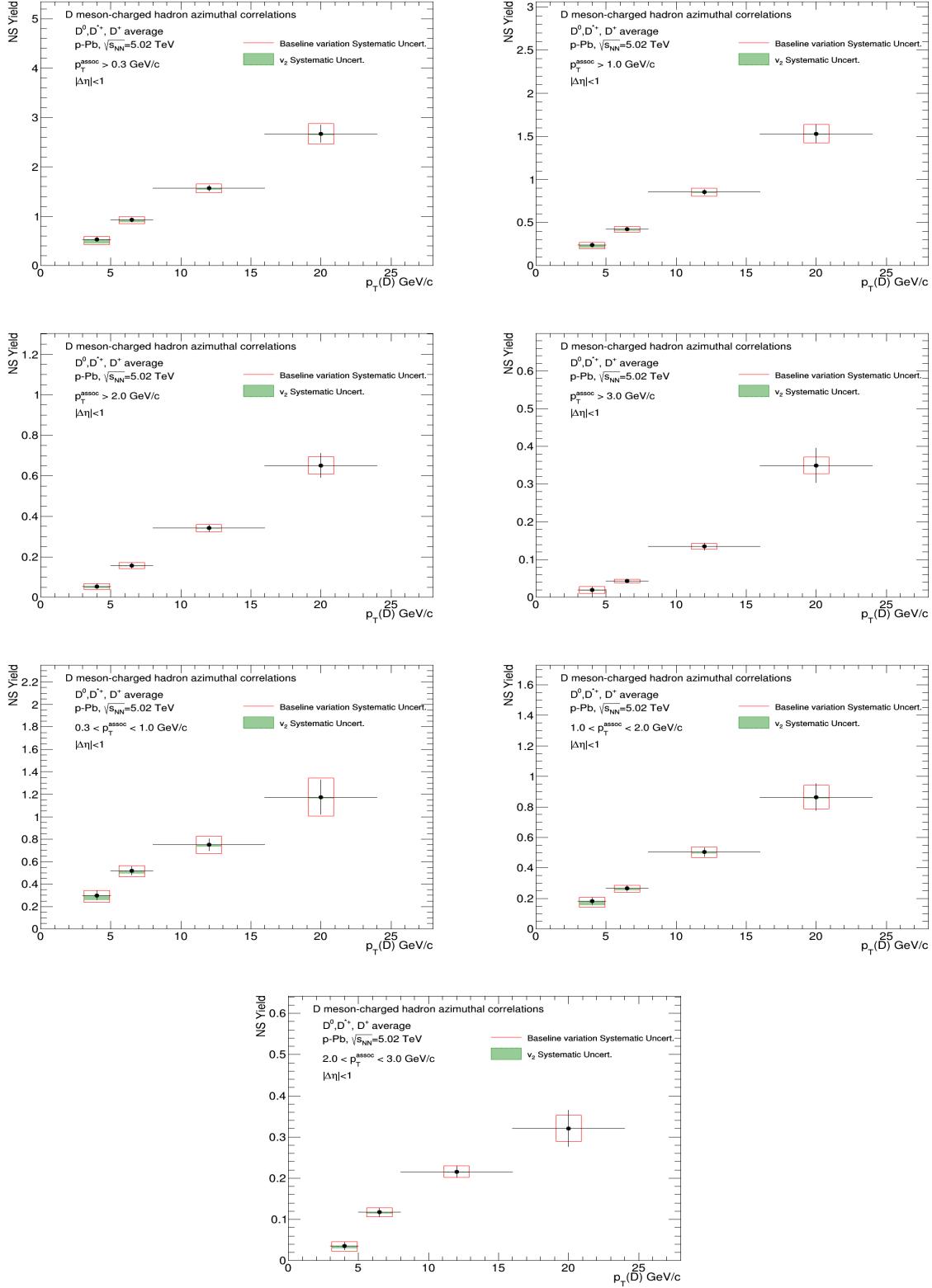
The evaluation of the systematic uncertainties on the observables obtained from the fits is performed as follows:

- The fits are repeated by changing the range of the transverse region in which the baseline 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 considered.
- In addition, $\Delta\phi$ correlation points are shifted to the upper and lower bounds of their uncorrelated systematic boxes, and refitted.
- The fits are also repeated by moving the baseline value from its default value (i.e. with the default transverse region) on top and on bottom of its statistic uncertainty. This helps to account, though in a systematic uncertainty, for the statistical uncertainty on the baseline position (since in the fit the baseline is constrained, and its error is not propagated to the other observables).
- The envelope between (i) the RMS of the relative variations of the parameters between the fit outcomes defined in the first two points, and (ii) the relative variations of the parameters from the fit outcomes defined in the third point, is considered as systematic uncertainty for the near-side and away-side widths.
- For the estimation of the baseline and of the near-side and away-side yields, instead, the previous value is added in quadrature with the $\Delta\phi$ -correlated systematics in the correlation distributions, since these values are affected by a change in the global normalization of the distributions.
- In addition, for all the fit observables, an additional fit variation is performed assuming, instead of a flat baseline, a $v_{2\Delta}$ -like modulation, with the following v_2 values for the associated 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), 0.09 (>1 GeV/c, 2-3 GeV/c), 0.1 (>3 GeV/c), on the basis of ATLAS preliminary results for heavy-flavour muons at 8 TeV; for the D-meson triggers the following v_2 values were instead assumed: 0.05 (3-5 GeV/c), 0.03 (5-8 GeV/c), 0.02 (8-24 GeV/c), on the basis of previous ALICE measurements in p-Pb collisions at 5 TeV [3]. The difference of the fit observables with respect to the standard fits is taken as uncertainty. Due to its peculiarity, this systematic uncertainty is summed in quadrature with the others to obtain the total uncertainty, but is 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 51, 52, 53, 54 and 55 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). Figures 56, 57, 58, 59 and 60 show the full systematic uncertainties for near side yield and width, away side yield and width, and baseline, with the breakdown of fit variation, $v_{2\Delta}$ and $\Delta\phi$ correlated systematic uncertainties.



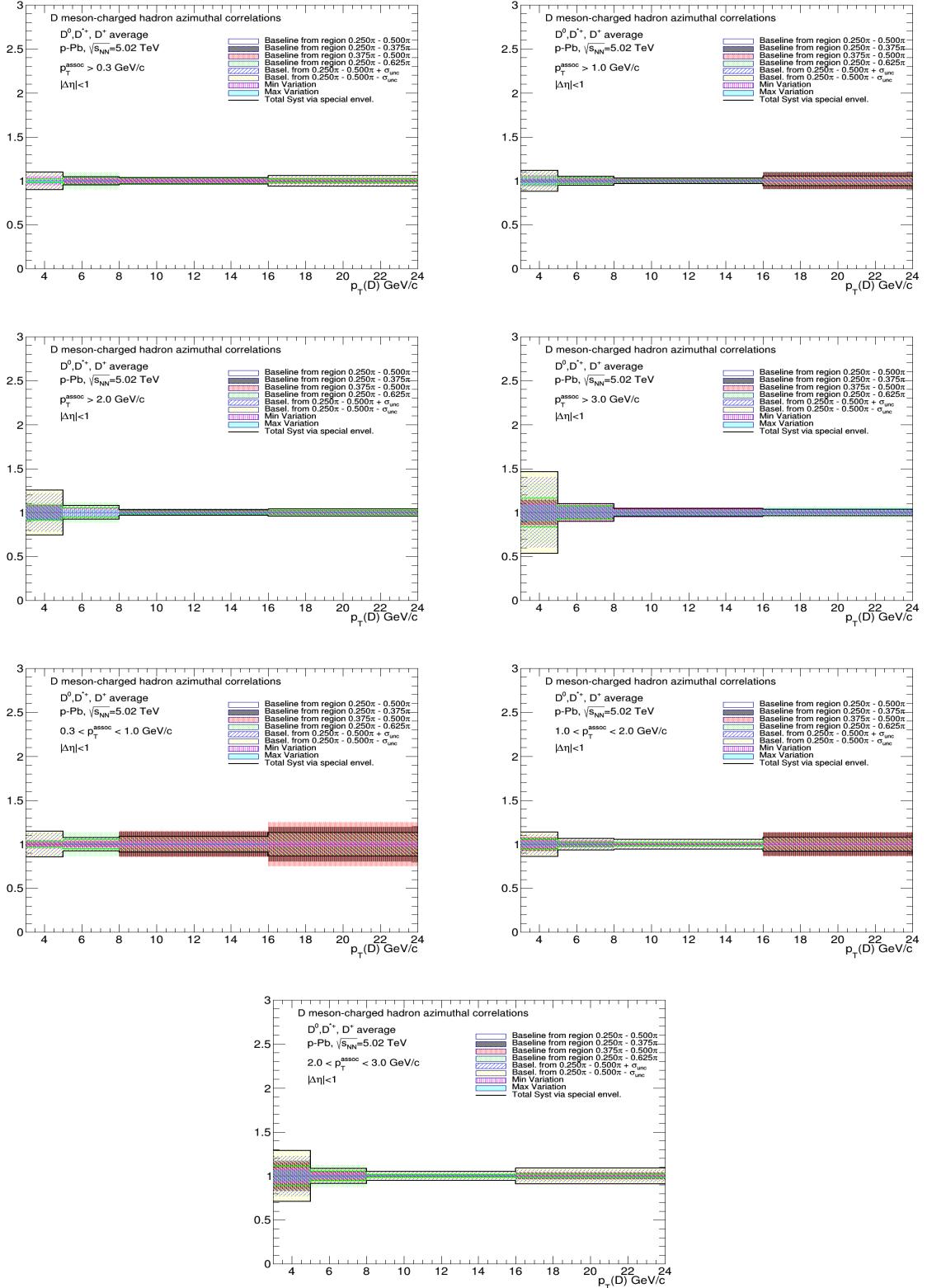
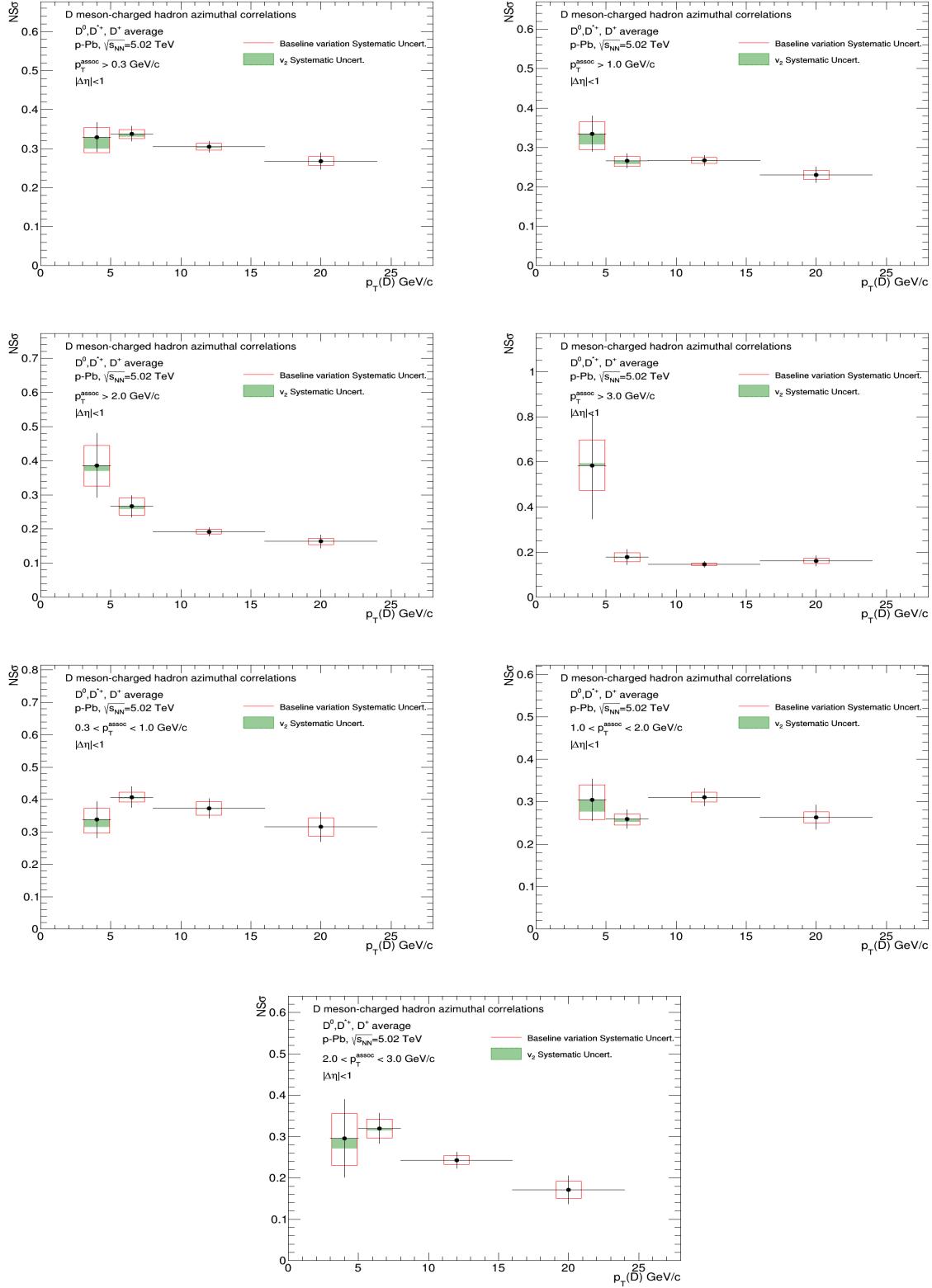


Figure 26: Top panels: near side yield $p_T(D)$ trend for the D-meson average, extracted from fit to the azimuthal correlation distributions, for all the analyzed kinematic ranges of associated track p_T . Bottom panels: for each kinematic region the systematic uncertainties coming from the variation of the fit procedure are shown.



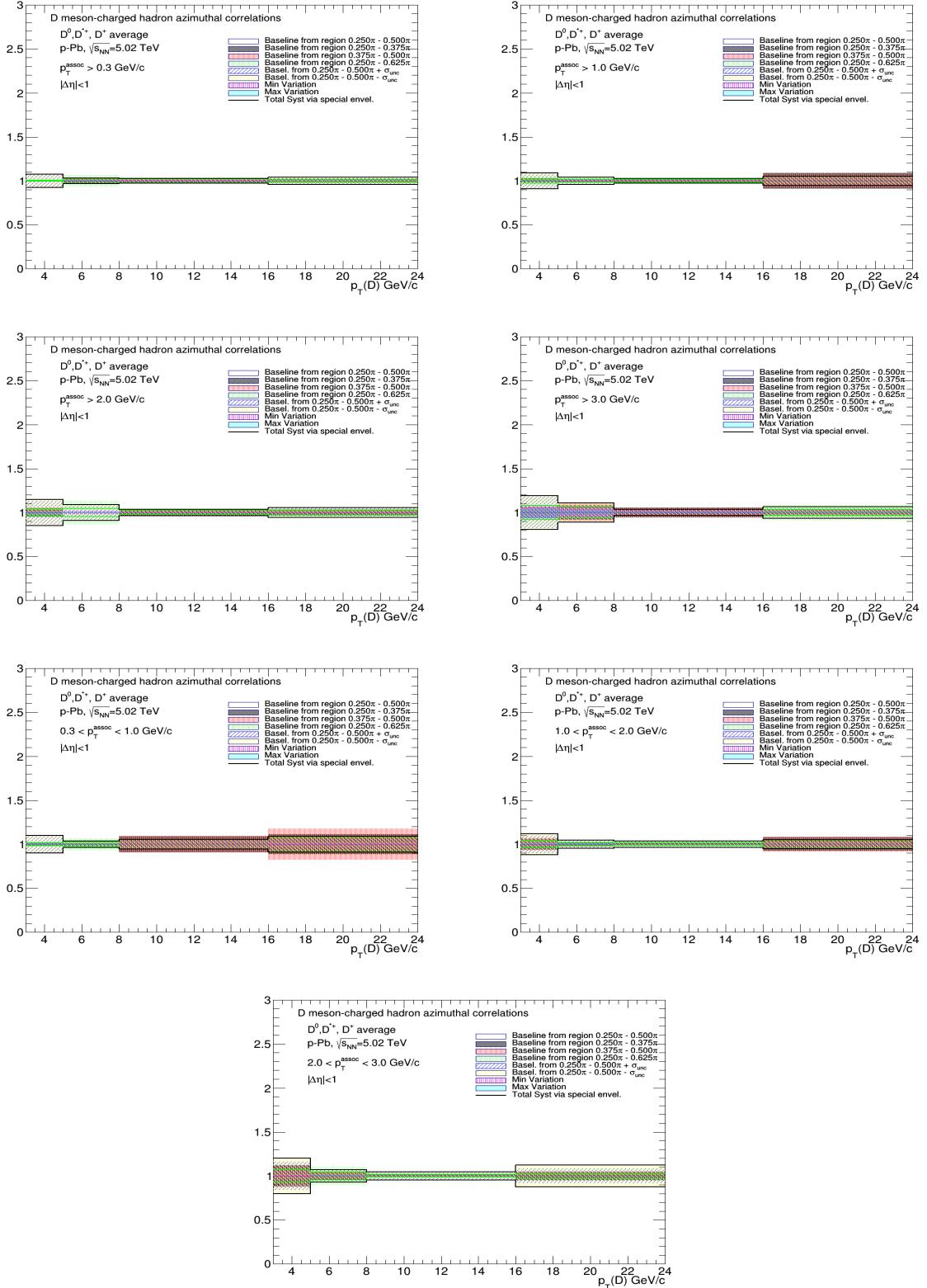
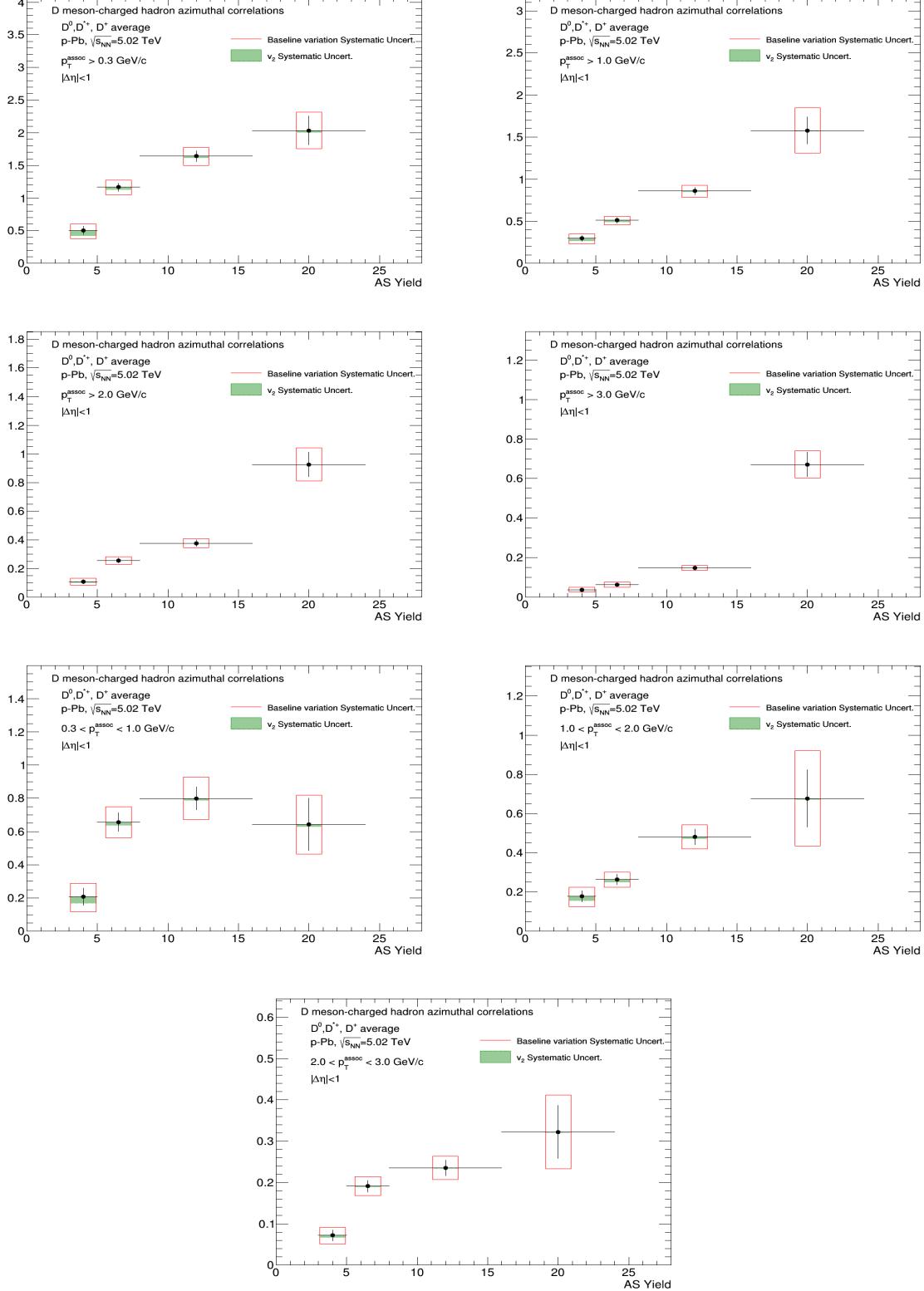


Figure 27: Top panels: near side width $p_T(D)$ trend for the D-meson average, extracted from fit to the azimuthal correlation distributions, for all the analyzed kinematic ranges of associated track p_T . Bottom panels: for each kinematic region the systematic uncertainties coming from the variation of the fit procedure are shown.



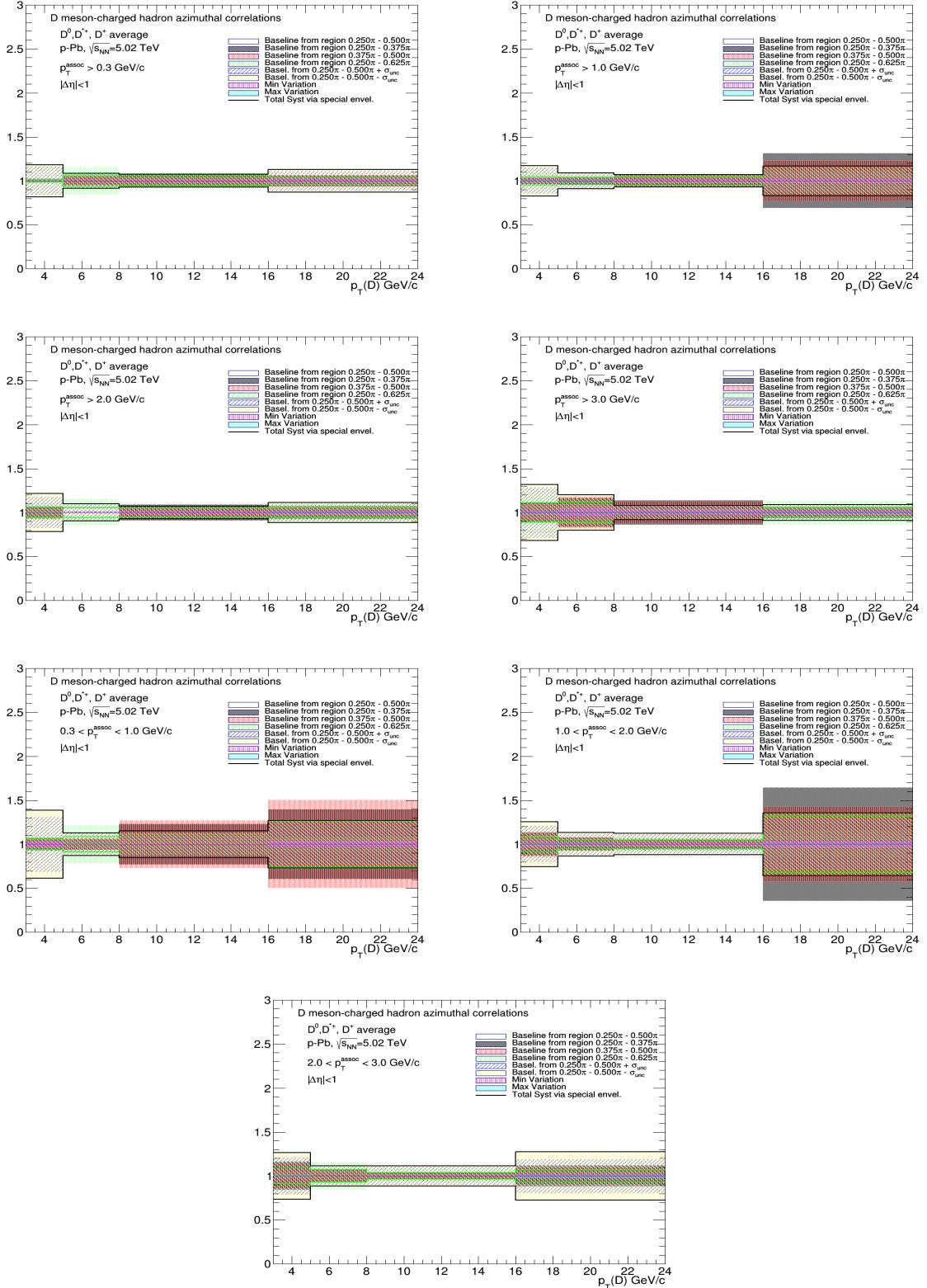
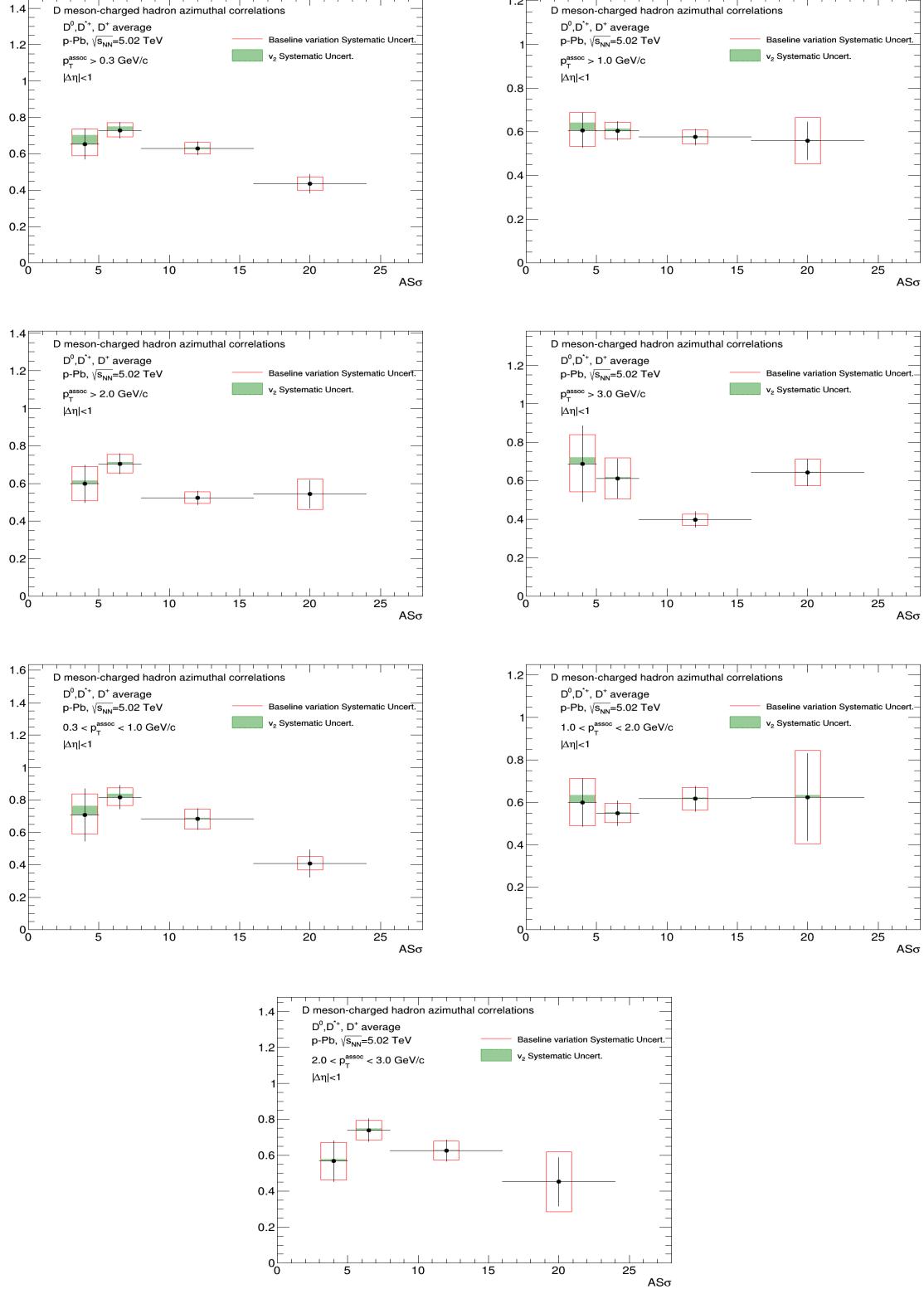


Figure 28: Top panels: away side yield $p_T(D)$ trend for the D-meson average, extracted from fit to the azimuthal correlation distributions, for all the analyzed kinematic ranges of associated track p_T . Bottom panels: for each kinematic region the systematic uncertainties coming from the variation of the fit procedure are shown.



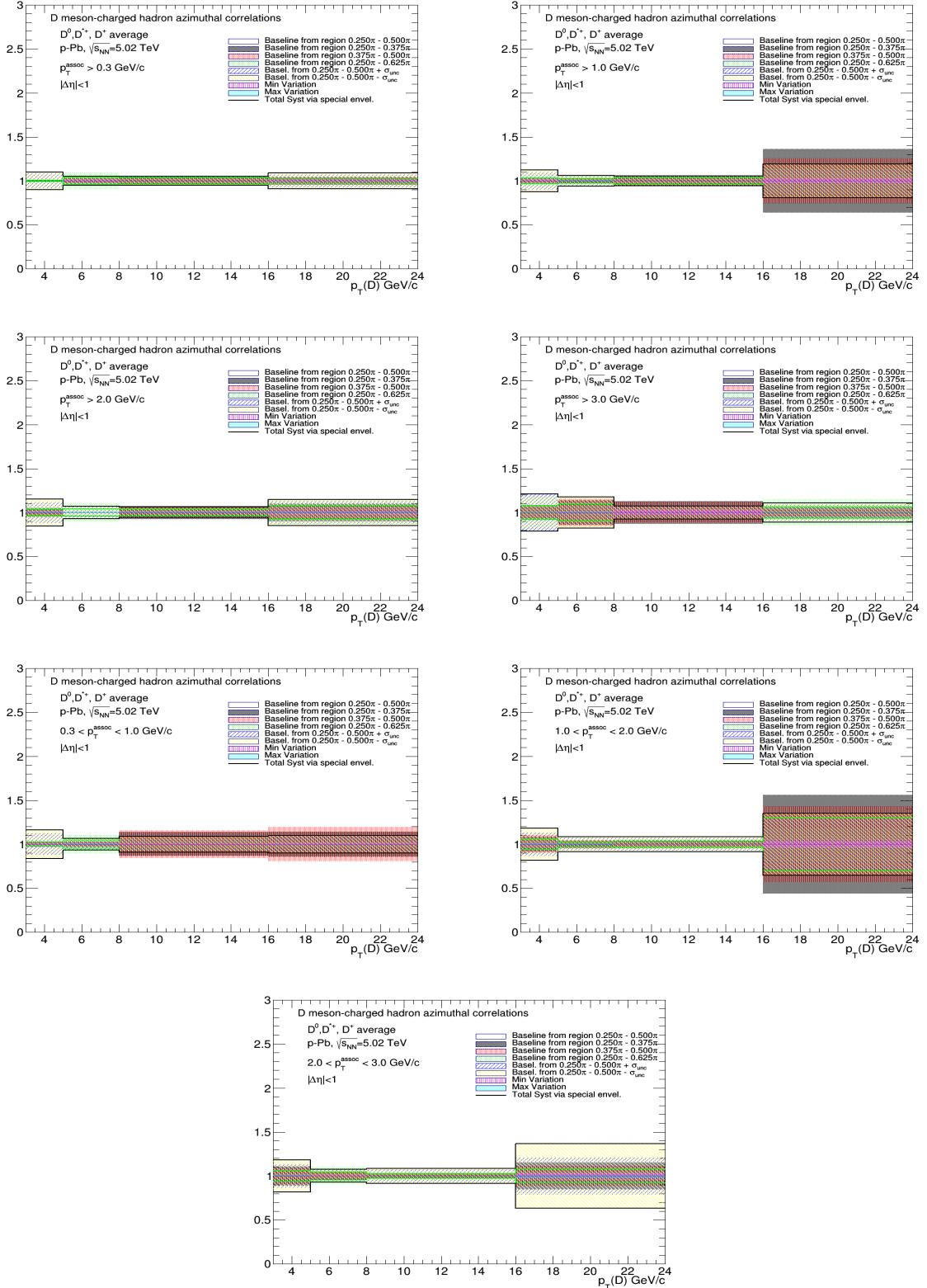
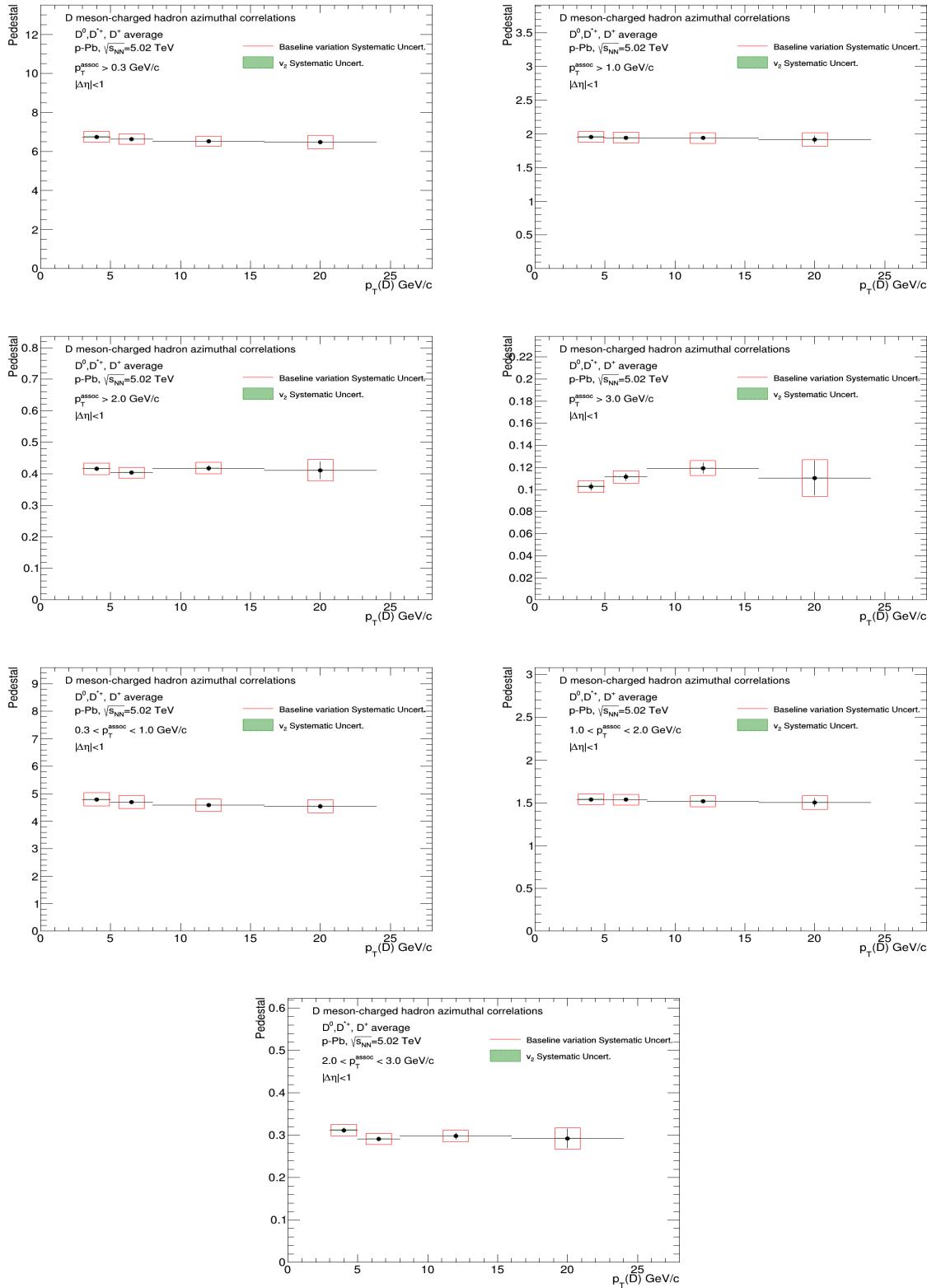


Figure 29: Top panels: away side width $p_T(D)$ trend for the D-meson average, extracted from fit to the azimuthal correlation distributions, for all the analyzed kinematic ranges of associated track p_T . Bottom panels: for each kinematic region the systematic uncertainties coming from the variation of the fit procedure are shown.



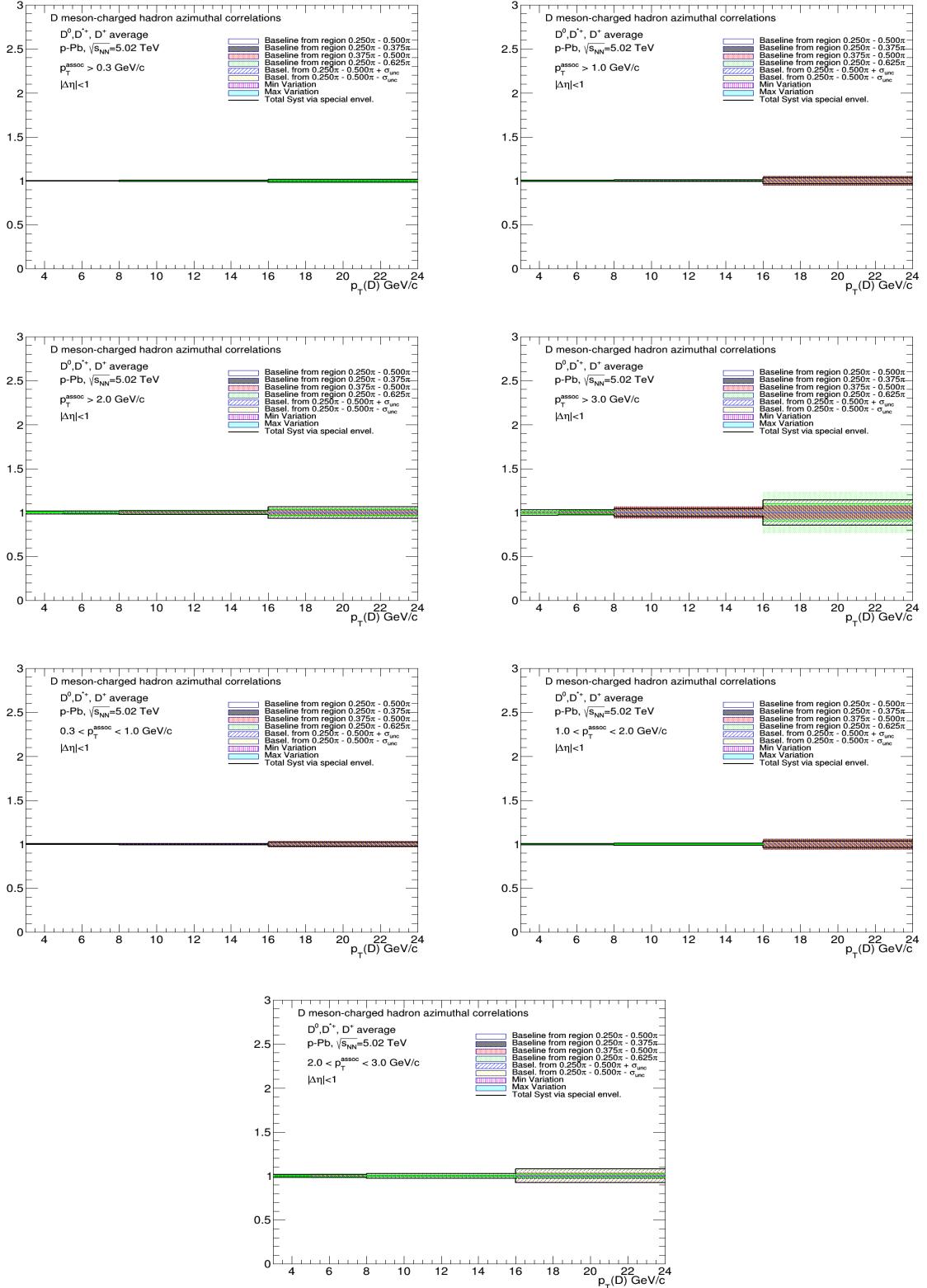


Figure 30: Top panels: baseline height trend for the D-meson average, extracted from fit to the azimuthal correlation distributions, for all the analyzed kinematic ranges of associated track p_T . Bottom panels: for each kinematic region the systematic uncertainties coming from the variation of the fit procedure are shown.

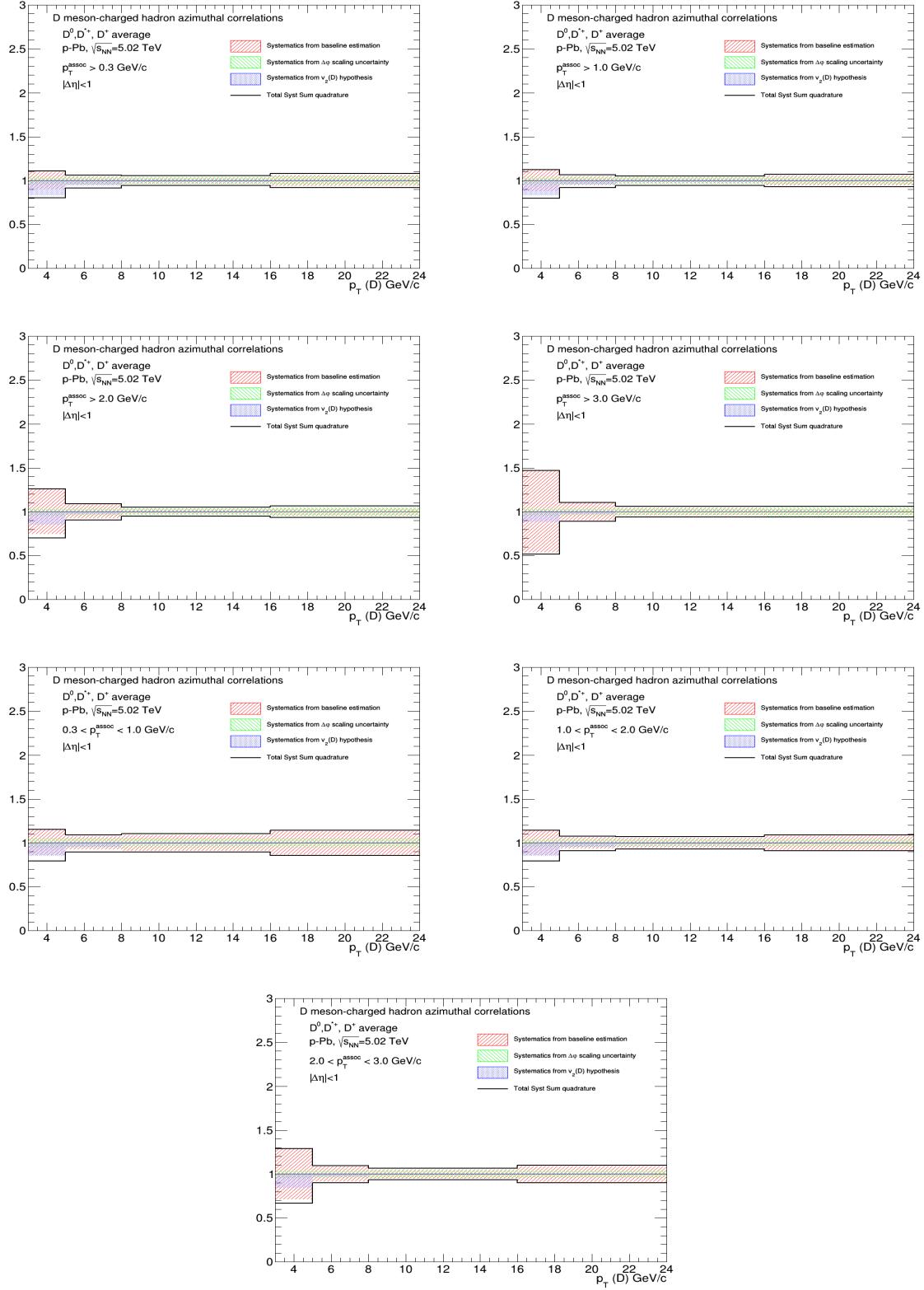


Figure 31: Total systematic uncertainty, and its components, for near-side yields in the different kinematic ranges analyzed

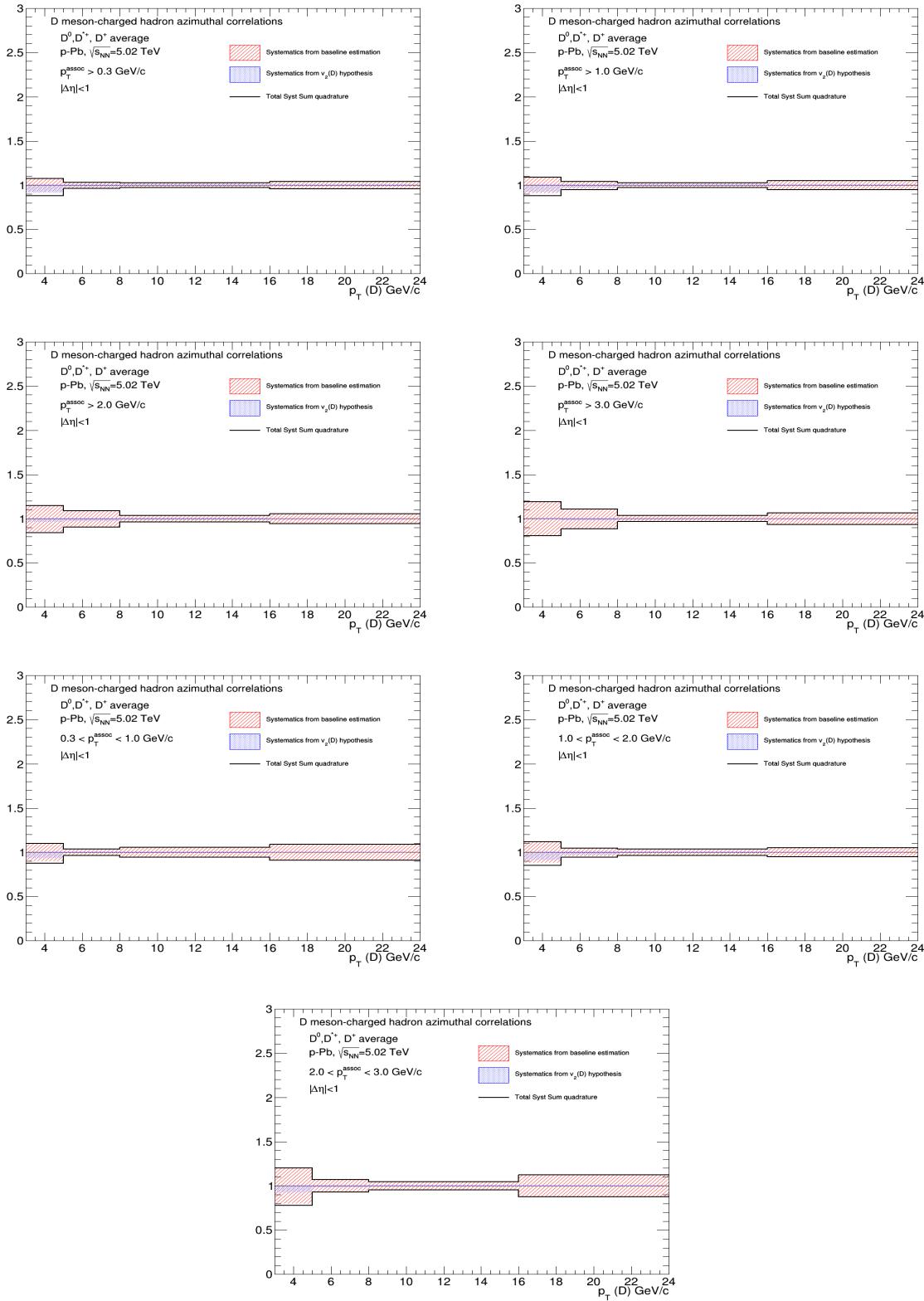


Figure 32: Total systematic uncertainty, and its components, for near-side sigma in the different kinematic ranges analyzed

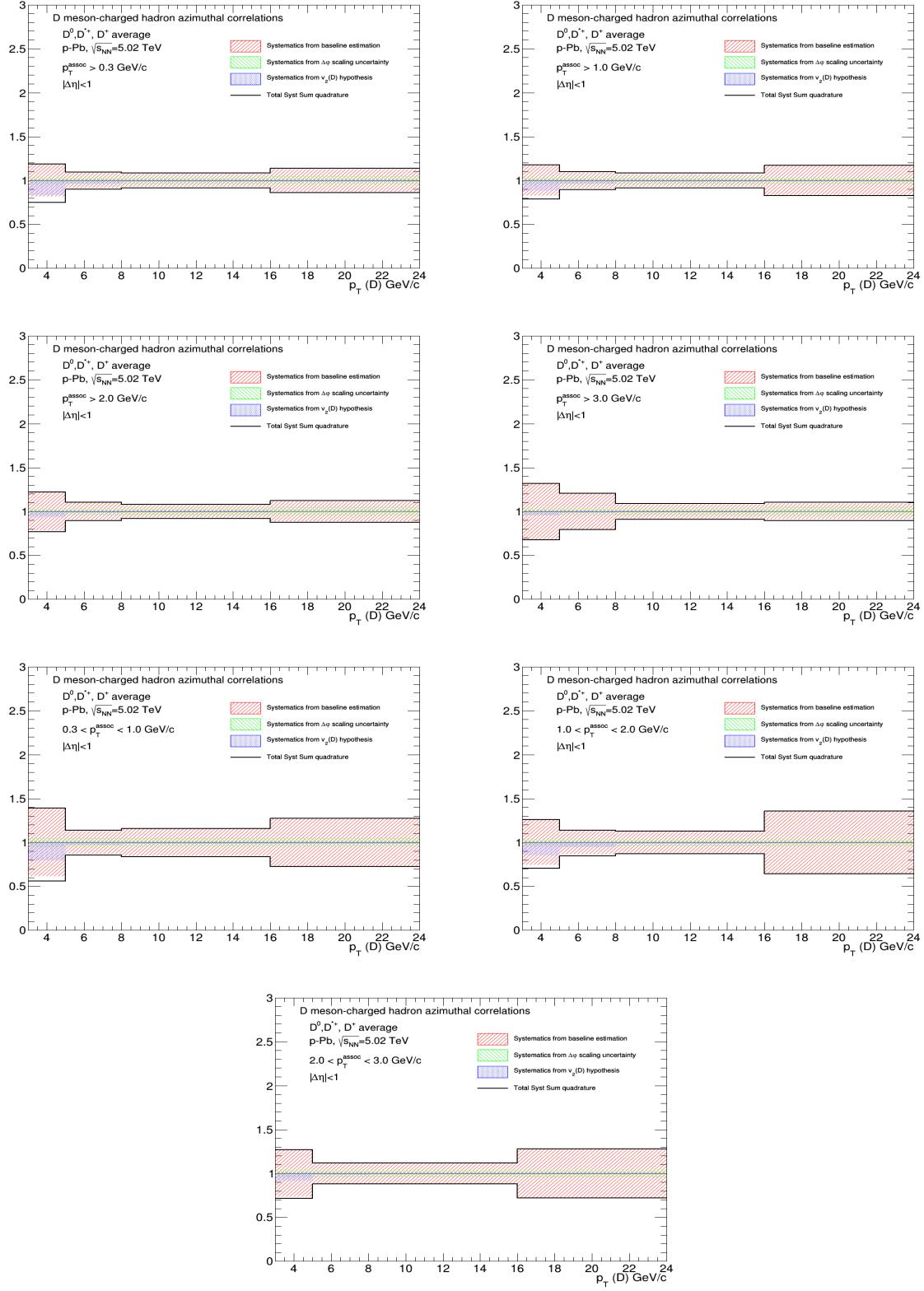


Figure 33: Total systematic uncertainty, and its components, for away-side yields in the different kinematic ranges analyzed

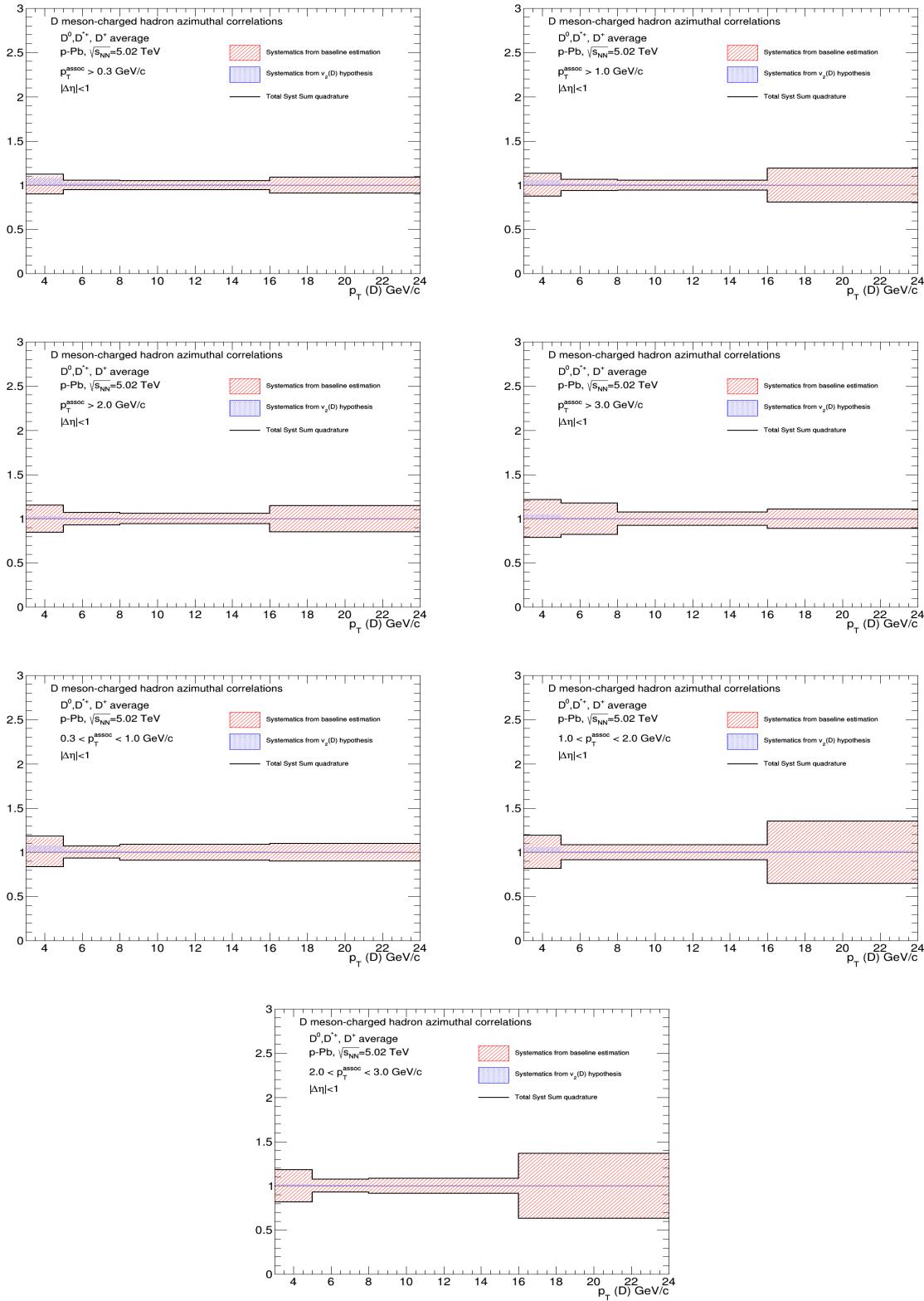


Figure 34: Total systematic uncertainty, and its components, for away-side sigma in the different kinematic ranges analyzed

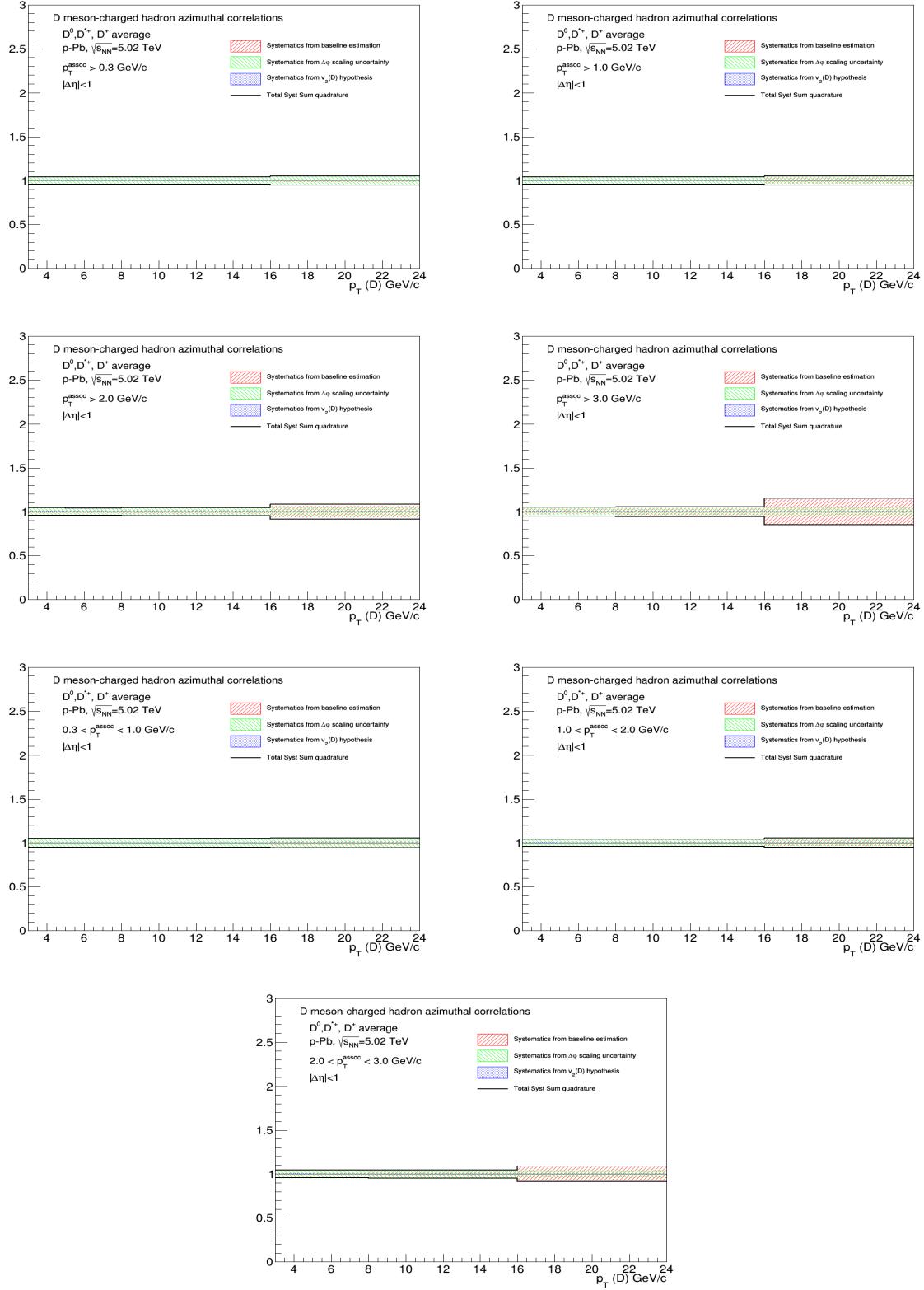


Figure 35: Total systematic uncertainty, and its components, for baseline heights in the different kinematic ranges analyzed.

679 **5.4 Comparison of 2016 p-Pb and 2013 p-Pb results**

680 In Figure 61, the average correlation distributions from the published analysis in p-Pb 2013 sample
 681 (black points) and the new p-Pb 2016 sample (red points), both at 5 TeV, are compared. As it's evident,
 682 the statistical and systematic uncertainties are largely reduced in the new data sample. The feature of
 683 the correlation distributions are the same in both systems, and an overall compatibility of the shapes is
 684 observed.

685 Figure 62 shows the same comparison for the fit observables. Also in this case the uncertainties are
 686 largely reduced for the 2016 analysis. The near-side widths are on top of each other; for the near-side
 687 yields a slight decrease of the 2016 results is observed in some p_T ranges with respect to 2013 results,
 688 though within the uncertainty, which were very large for the 2013 sample. The pedestal values are also
 689 fully compatible within the two samples. The sensitivity to away-side observables was very poor for
 690 2013 results, hence a comparison with 2016 data is difficult, anyway, within the large uncertainties, also
 691 the away-side observables are compatible between the two datasets.

692 **5.5 Comparison of 2016 p-Pb and 2010 pp results**

693 Figure 63 shows the comparison of the average D-h correlation distributions in pp 2010 data sample at
 694 $\sqrt{s} = 7$ TeV (published in [2]) and in the new p-Pb 2016 sample at $\sqrt{s_{NN}} = 5.02$ TeV. The results are
 695 shown after the subtraction of the baseline. The precision of the new p-Pb results is much better than that
 696 of pp results; the correlation distributions show very similar features in the two collision systems.

697 In Figure 64 the comparison is performed for the near-side peak observables, again in the common
 698 kinematic ranges, where the same consideration about the uncertainties holds. The similarity of the
 699 correlation distributions is reflected also in the near-side yield and width values, which do not seem to
 700 differ within the uncertainties, pointing to the absence of strong effects from cold-nuclear matter effects
 701 on the correlation distributions.

702 It has to be said that, on the base of a study performed with Pythia6-Perugia2011 simulations, a scaling
 703 factor of about 0.93 is expected when passing from a center-of-mass energy of $\sqrt{s} = 7$ TeV to $\sqrt{s} = 5$
 704 TeV, difficult to be appreciated with the current uncertainties, especially the pp ones.

705 **5.6 Comparison of 2016 p-Pb and model expectations**

706 A comparison of the average D-h correlation distributions on the new p-Pb data samples with expec-
 707 tations from Monte Carlo simulations (currently Pythia6-Perugia2011, Pythia6-Perugia2010, Pythia6-
 708 Perugia0, PYTHIA8; POWHEG+PYTHIA and EPOS 3 will be added if they come in time) is shown in
 709 Figure 65, after the baseline subtraction (which differs strongly between data and simulations, due to he
 710 very different underlying event). The simulations, though being for pp, include the boost of the center-
 711 of-mass along the beam axis present in p-Pb collisions and nuclear PDF. The shape of the correlation
 712 distributions is well reproduced by all the models, together with their p_T trend and with the evolution of
 713 the correlation peaks.

714 Figures 66 and 67 show the same comparison for the fit observables (peak yields and widths for near-side
 715 and away-side, respectively), for all the addressed p_T ranges.

716 **5.7 Planned results for SQM approvals**

717 We are planning to approve the following results, all shown in the previous figures (the final graphical
 718 style of the plots is still to be finalized):

- 719 – Average D-h correlation distributions, in exemplary pT range
 720 – Fit of D-h correlation distributions, in exemplary pT range

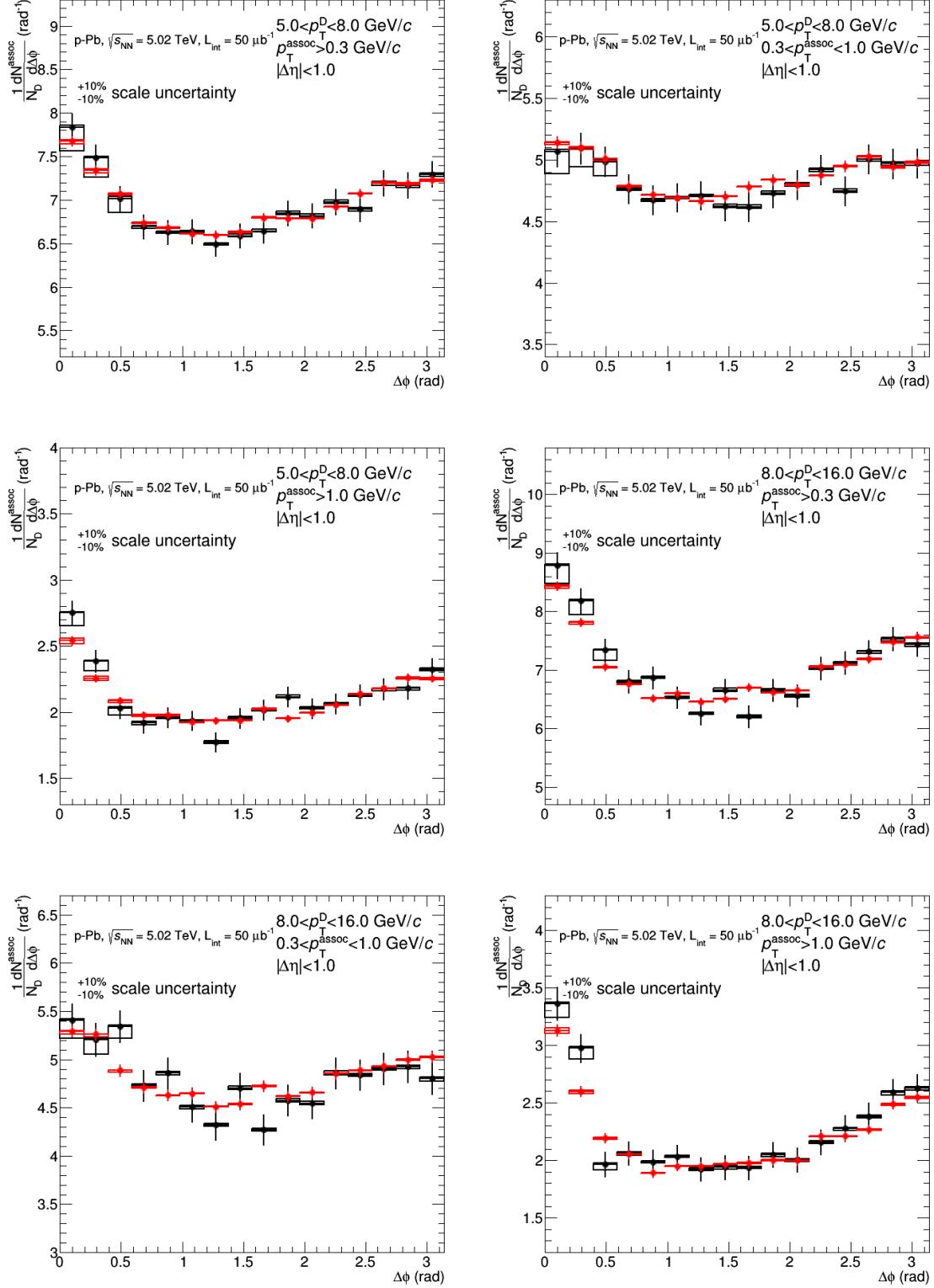


Figure 36: Comparison of 2016 (red) and 2013 (black) results for azimuthal correlation distributions, for the common p_T ranges.

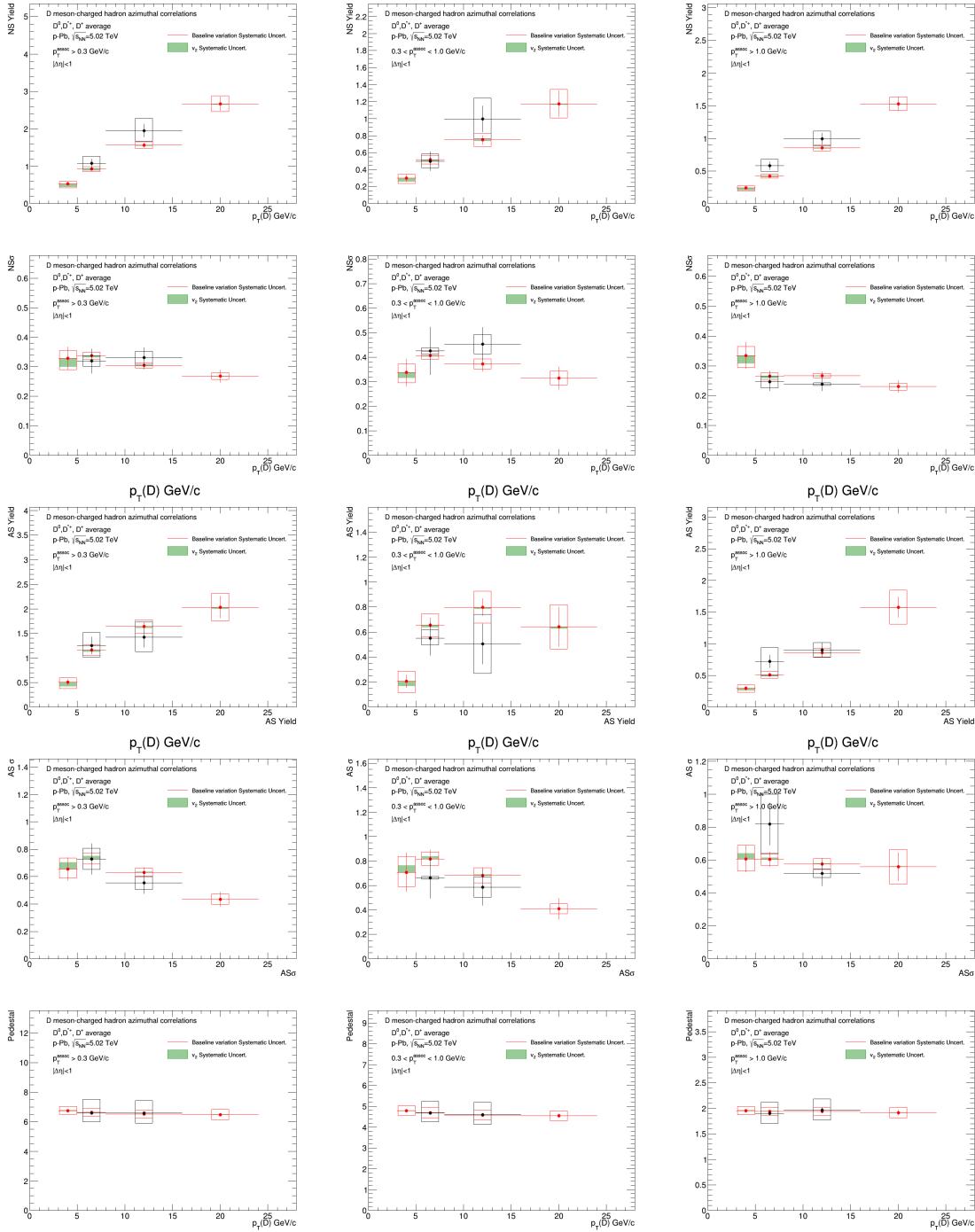


Figure 37: Comparison of the average D-h azimuthal correlation properties between 2016 p-Pb (red) and 2013 p-Pb (black) data analysis, for the common p_T ranges of D meson and associated particles.

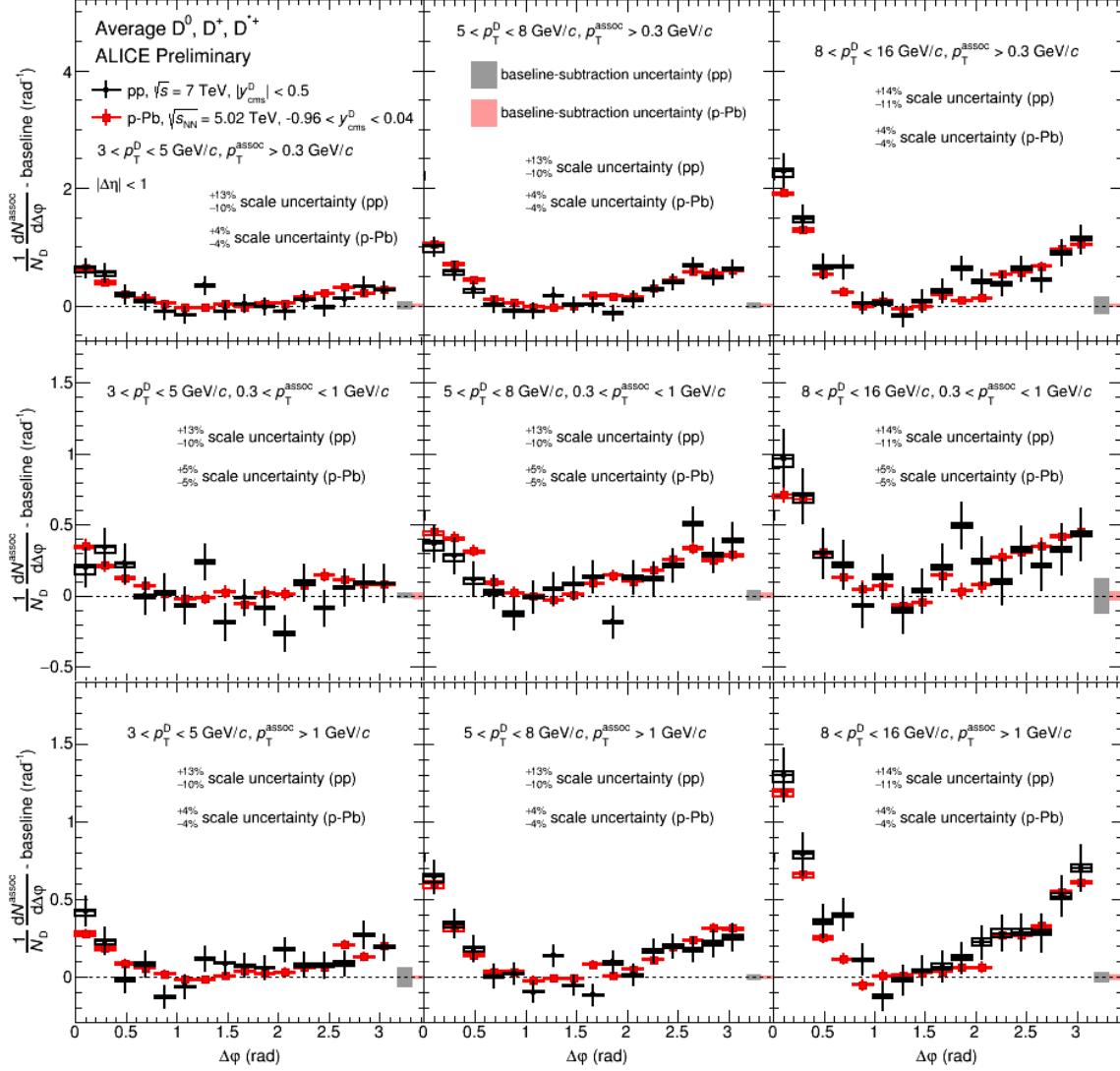


Figure 38: Comparison of pp 2010 (black) and p-Pb 2016 (red) average D-h azimuthal correlation distributions, for the common p_T ranges.

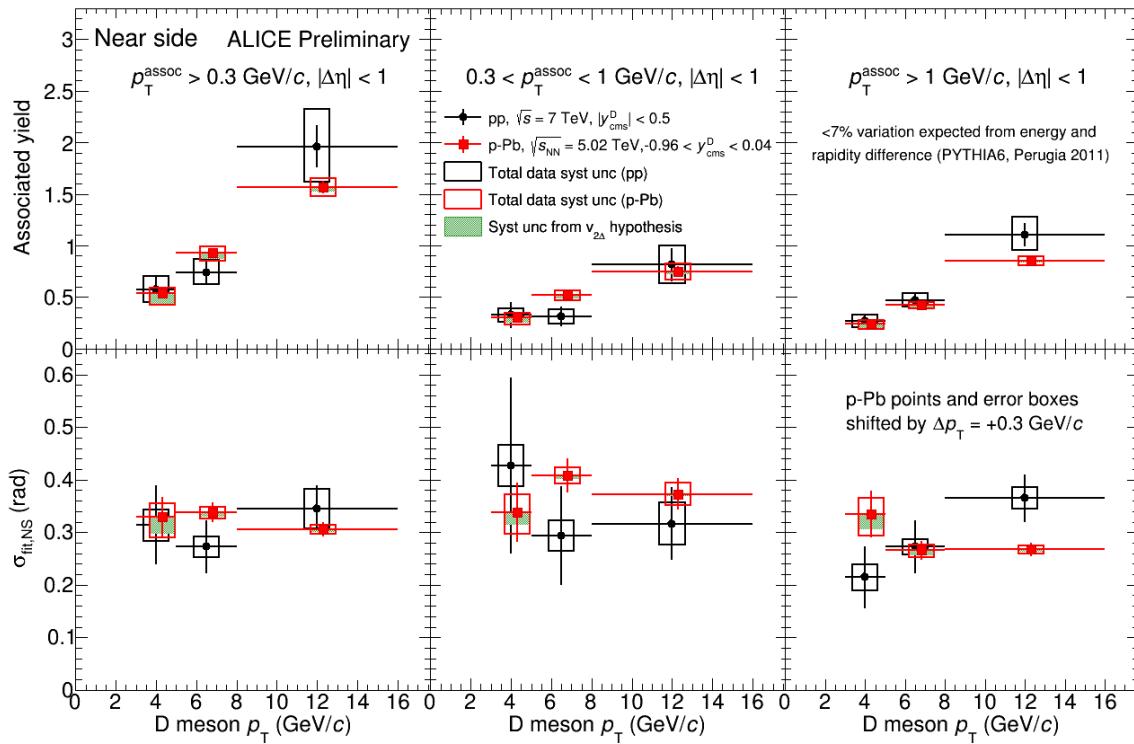
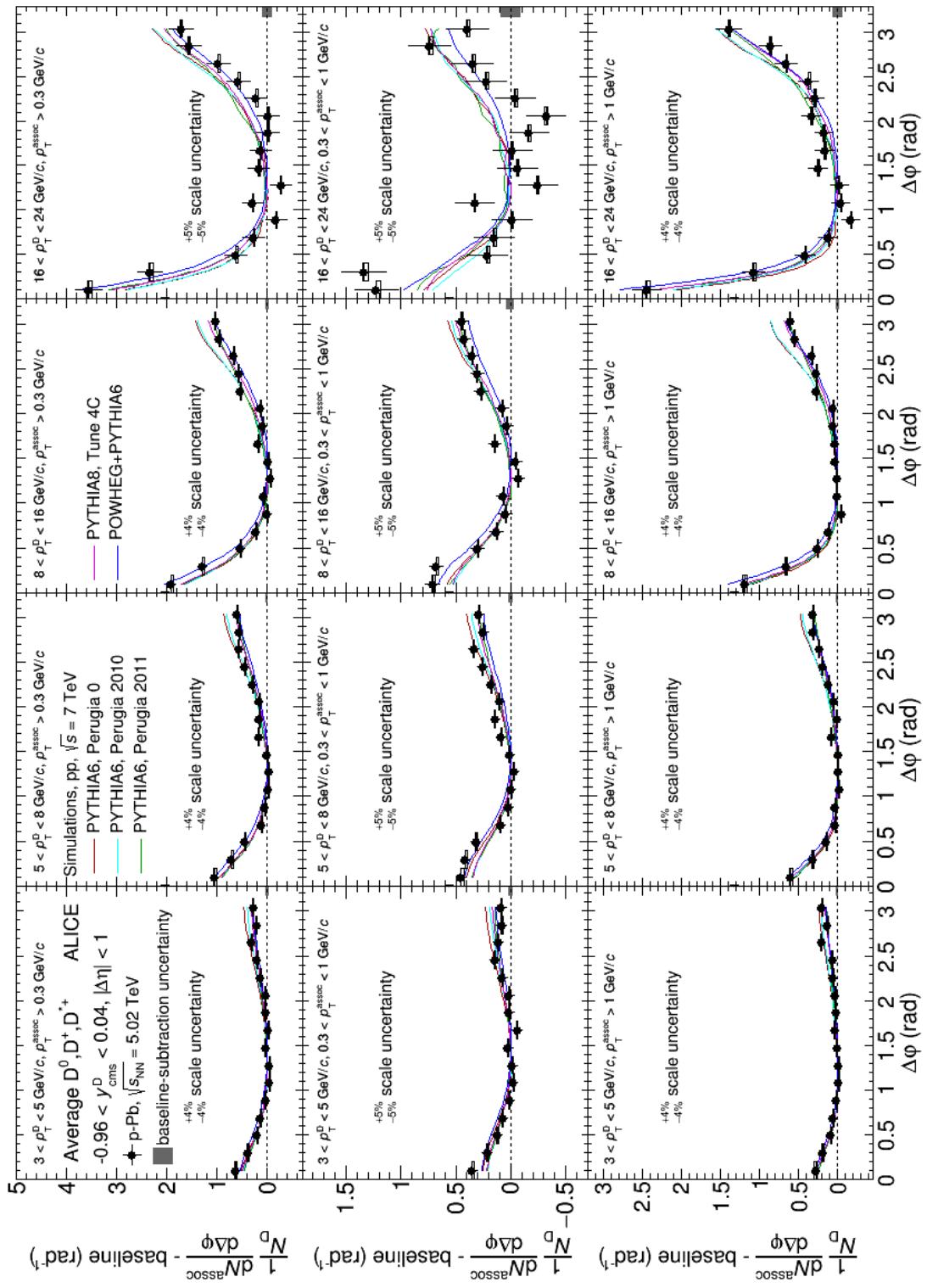


Figure 39: Comparison of pp 2010 (black) and p-Pb 2016 (red) near-side peak yields and widths, for the common p_T ranges.

- 721 – $p_T(D)$, $p_T(\text{assoc})$ trend of NS yield, NS width, AS yield, AS sigma
- 722 – Comparison of correlation distributions with expectations from models (PYTHIA6, PYTHIA8, if
723 in time POWHEG, EPOS)
- 724 – Comparison of fit observables with expectations from models (PYTHIA6, PYTHIA8, if in time
725 POWHEG, EPOS)
- 726 – Comparison of correlation distributions with pp 2010 results
- 727 – Comparison of fit observables with pp 2010 results



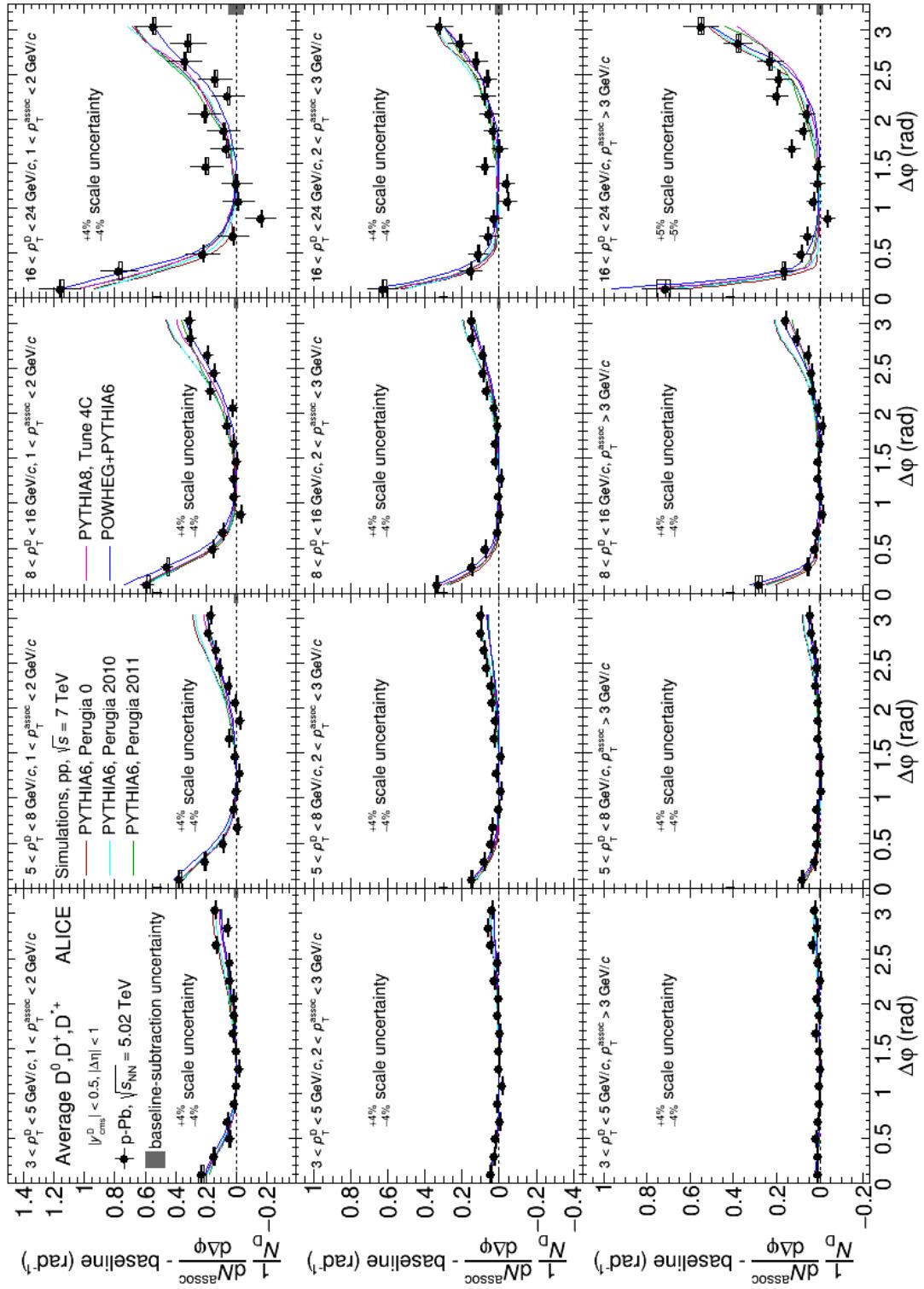
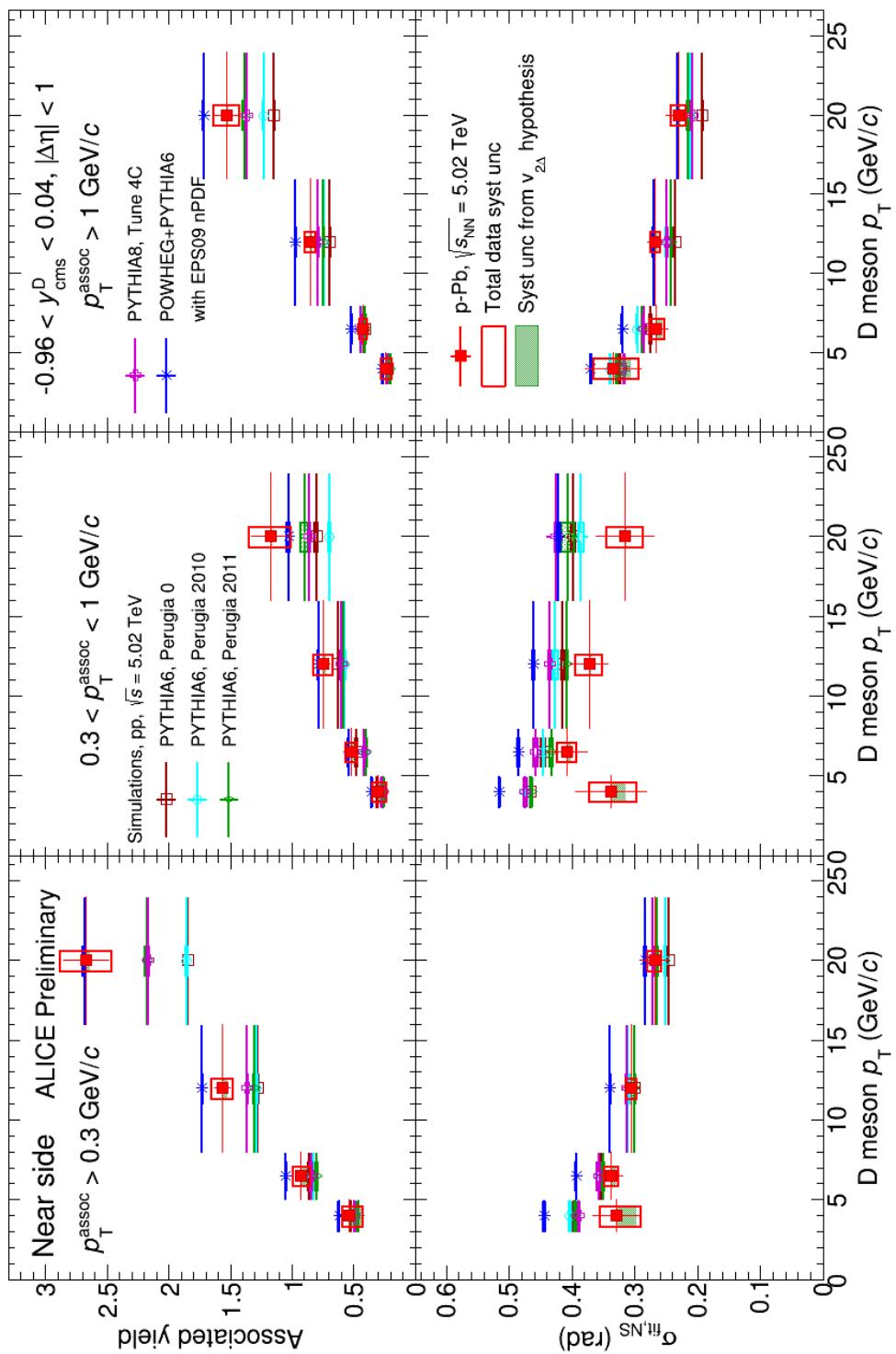


Figure 40: Comparison of p-Pb 2016 average D-h correlation distributions and model expectations, for all the studied kinematic ranges.



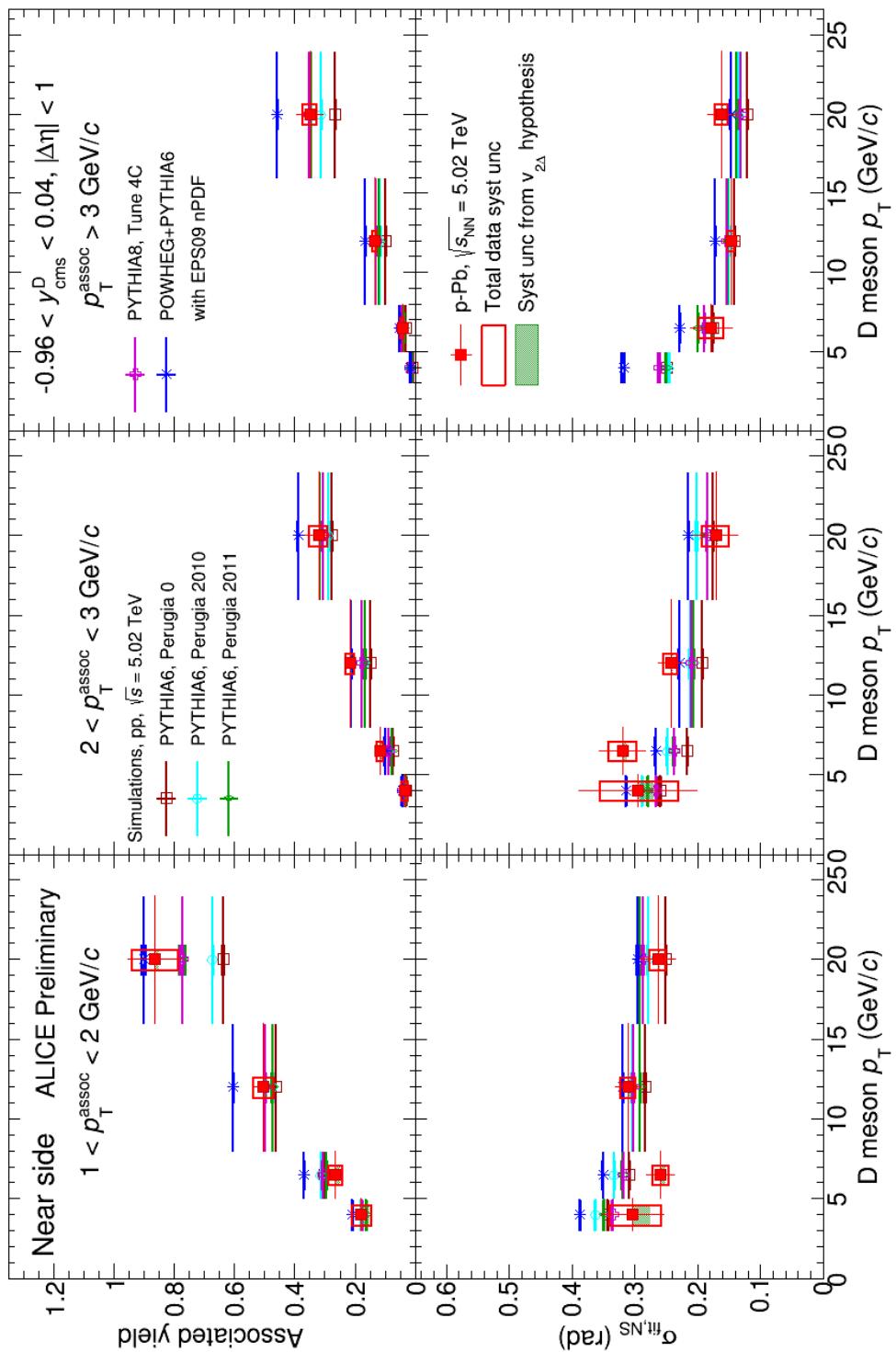
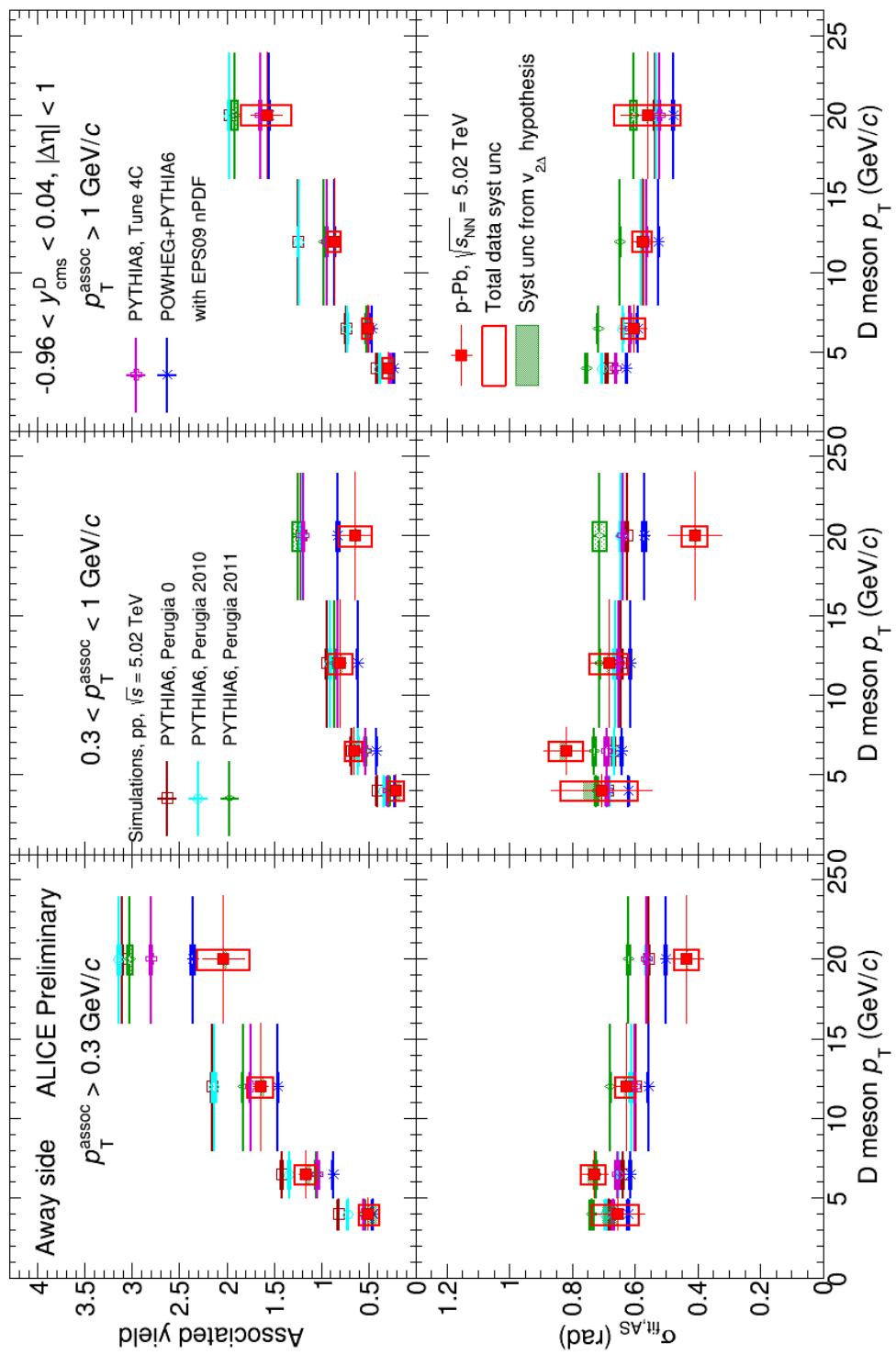


Figure 41: Comparison of near-side peak yields and widths from p-Pb 2016 results and model expectations, for all the studied kinematic ranges.



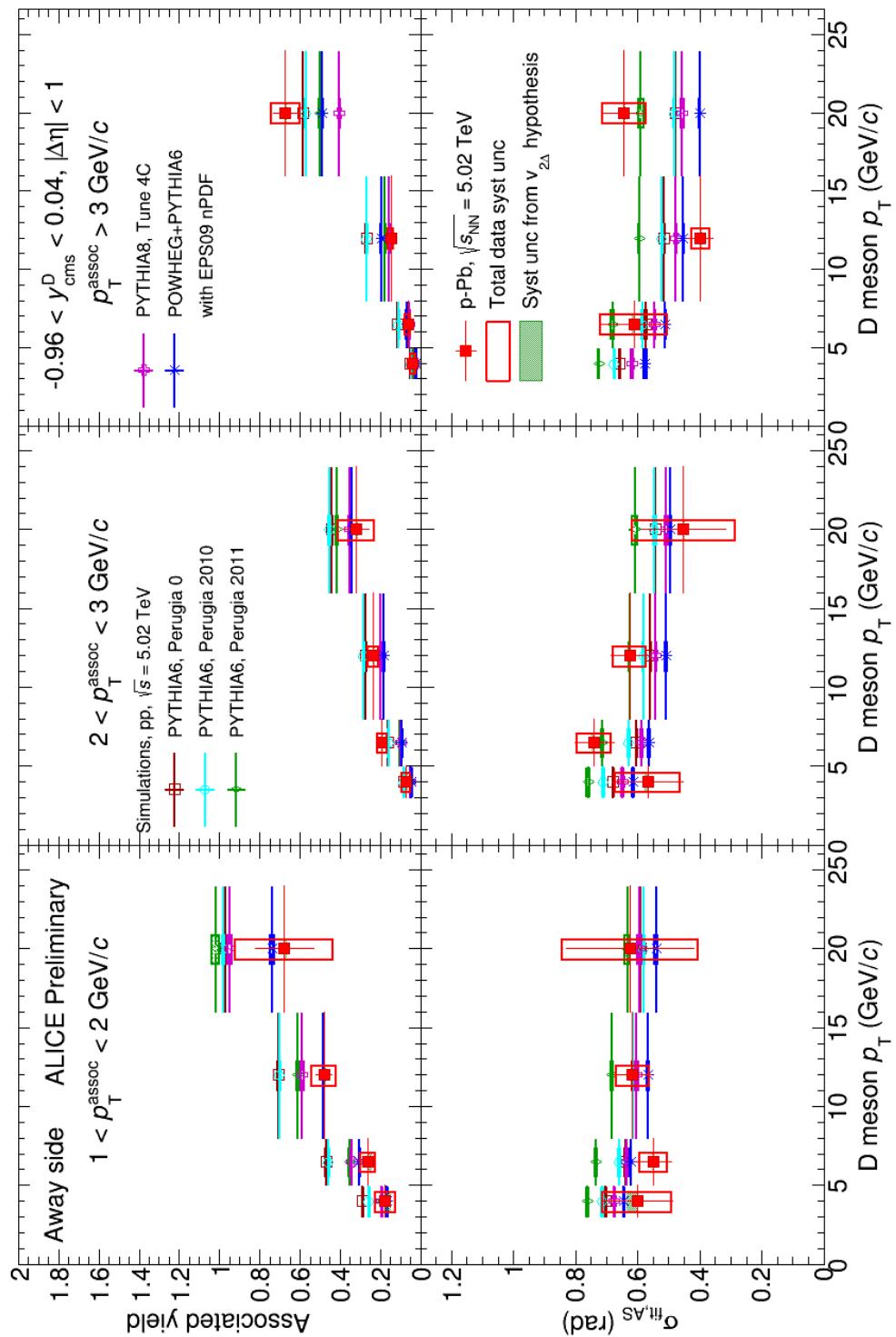


Figure 42: Comparison of away-side peak yields and widths from p-Pb 2016 results and model expectations, for all the studied kinematic ranges.

728 **6 Bibliography**729 **References**

- 730 [1] B. Abelev et al. [ALICE Collaboration], JHEP **01** (2012) 128
731 [2] B. Abelev et al. [ALICE Collaboration], Eur. Phys. J. C (2017) 77:245
732 [3] B. Abelev et al. [ALICE Collaboration], Phys.Lett. B719 (2013) 29-41
733 [4] B. Abelev et al. [ALICE Collaboration], Phys.Lett. B726 (2013) 164-177
734 [5] <https://aliceinfo.cern.ch/Figure/node/12130> (ALI-PREL-138003)
735 [6] ALICE Collaboration, arXiv:1609.06643 [nucl-ex]
736 [7] <https://aliceinfo.cern.ch/Notes/node/300>
737 [8] <https://aliceinfo.cern.ch/Notes/node/238>
738 [9] <https://aliceinfo.cern.ch/Notes/node/201>