

Analysis Note

Measurement of charged-particle production in single diffractive proton-proton collisions at $\sqrt{s} = 200$ GeV with the STAR detector at RHIC

Leszek Adamczyk¹, Łukasz Fulek¹, Mariusz Przybycień¹, and Rafał Sikora¹

¹*AGH University of Science and Technology, FPACS, Kraków, Poland*

30th November 2020

² We report on the measurement of inclusive and identified (pion, kaon, proton and their anti-particles) charged-particle production in Single Diffractive process in proton-proton collisions at $\sqrt{s} = 200$ GeV with the STAR detector at RHIC. The forward-scattered proton is measured in the Roman Pot system, while the charged-particles are identified based on the specific ionization energy loss of tracks reconstructed in the central detector. The proton-antiproton production asymmetry is measured as a function of transverse momentum and mass of diffractive system, and used to study the baryon number transfer over a large rapidity interval in Single Diffractive process.

In this note we present the analysis of the Single Diffractive Dissociation process with the STAR Roman Pot (RP) detectors at RHIC. The measurement is focused on the charged particle multiplicity, its dependence on the transverse momentum and pseudorapidity in three regions of ξ : $0.02 < \xi < 0.05$, $0.05 < \xi < 0.1$ and $0.1 < \xi < 0.2$. The identified particle to antiparticle (pion, kaon, proton and their antiparticle) multiplicity ratios and K/π ratio as a function of transverse momentum in above three ξ regions are also measured. The data come from proton-proton collisions collected in 2015. The forward proton was tagged in the STAR Roman Pot system while the charged particle tracks were reconstructed in the STAR Time Projection Chamber (TPC). We describe all stages of the analysis involving comparison of the data with MC simulations and systematic uncertainty studies. More technical parts of the analysis are described in a supplementary analysis note [1].

³ **List of contributions**

Leszek Adamczyk	Analysis coordination/supervision, production of picoDST, production of embedded MC samples
Lukasz Fulek*	Main analyzer, write-up author
Mariusz Przybycień	Analysis supervision
Rafal Sikora	Analysis support

⁷ * - contact editor

⁹ **Change log**

30th	ver. 1.0	Initial revision
November		

¹³

Contents

14	List of contributions	2
16	Change log	2
17	1 Introduction	1
18	2 Monte Carlo Samples	2
19	3 Data Sample and Event Selection	4
20	3.1 Track Selection	5
21	3.2 Fiducial Region of the Measurement	8
22	4 Background Contribution	11
23	4.1 Background from Non-Primary Tracks	13
24	4.2 Control Plots	20
25	5 Selection Efficiencies	24
26	5.1 Vertex Reconstruction	24
27	5.2 Correction to BBC-Small	29
28	6 Migrations into and out of the Fiducial Region	33
29	6.1 Migrations of Tracks into and out of the Fiducial Region	33
30	6.2 Migrations in ξ	34
31	7 Event Corrections and Unfolding Procedure	37
32	7.1 Correction to dN/dn_{sel}	37
33	7.2 Correction to Transverse Momentum and Pseudorapidity Distributions	40
34	7.3 Closure Tests	41
35	7.4 EAST-WEST asymmetry	41
36	7.5 Particle Identification	42
37	7.6 Antiparticle-to-Particle Ratios	48
38	8 Systematic Uncertainties	49
39	9 Results	55
40	9.1 Comparison of Charged-Particle Densities at Central Rapidities	64
41	10 Summary and Conclusions	66
42	Appendices	69
43	A Proton and Antiproton DCA Distributions	70
44	B Distributions of $n\sigma_{dE/dx}^i$ in SD	83

⁴⁵ **Acronyms**

⁴⁶	CD	Central Diffraction
⁴⁷	DD	Double Diffraction
⁴⁸	MBR	Minimum Bias Rockefeller
⁴⁹	MC	Monte Carlo
⁵⁰	ND	Non-Diffractive
⁵¹	QCD	Quantum Chromodynamics
⁵²	RP	Roman Pot
⁵³	SaS	Schuler and Sjöstrand
⁵⁴	SD	Single Diffraction
⁵⁵	TPC	Time Projection Chamber

1. Introduction

Inclusive measurements of charged-particle distributions in proton–proton (pp) collisions probe the strong interaction in the low-momentum transfer, non-perturbative regime of Quantum Chromodynamics (QCD). In this kinematic region interactions are usually described by phenomenological models implemented in Monte Carlo (MC) event generators. Measurements can be used to constrain the free parameters of these models. An accurate description of low-energy strong interaction processes is essential for understanding and precise simulation of different types of pp processes and the effects of multiple pp collisions in the same bunch crossing at high instantaneous luminosity at hadron colliders. Measurements with tagging of the forward-scattered proton are of special interest. They give direct access to specific but still significant part of pp processes called diffraction. In addition precise modelling of forward particle production is essential for better understanding of the longitudinal development of air showers observed in experiments studying cosmic radiation.

We present a measurement of charged particle production in events with single forward proton tagging (dominated by Single Diffraction (SD): $p + p \rightarrow p + X$). The following observables are studied:

$$\frac{1}{N} \frac{dN}{dn_{ch}}, \quad \frac{1}{N} \frac{1}{2\pi p_T} \frac{d^2 N_{ch}}{d\bar{\eta} dp_T}, \quad \frac{1}{N} \frac{dN_{ch}}{d\bar{\eta}} \quad (1.1)$$

where n_{ch} is the number of primary charged particles within kinematic range given by $p_T > 200$ MeV and $|\eta| < 0.7$, N is the total number of events with $2 \leq n_{ch} \leq 8$, N_{ch} is the total number of charged particles within the above kinematic acceptance and $\bar{\eta}$ is the pseudorapidity of the charged particle with longitudinal momentum taken with respect to direction of the forward scattered proton. To suppress non-SD events the trigger system required no signal in BBC-small in the direction of forward scattered proton and signal in BBC-small in opposite direction. The measurements are performed in a fiducial phase space of the forward scattered protons of $0.04 < -t < 0.16$ GeV $^2/c^2$ and $0.02 < \xi < 0.2$, where ξ is the fractional energy loss of the scattered proton and t is the squared four momentum transfer. In case of SD process $\xi = M_X^2/s$, where M_X is the mass of the state X into which one of the incoming proton dissociates and s is the center of mass energy squared of the pp system. The Mandelstam variable t is defined by $t = (p_1 - p_3)^2$, where p_1 is the four-momentum of the incoming proton, p_3 is the four-momentum of the outgoing proton. The above mentioned observables are presented in three ξ regions: $0.02 < \xi < 0.05$, $0.05 < \xi < 0.1$ and $0.1 < \xi < 0.2$. In addition their average values are presented as a function of ξ .

We have also studied an identified particle to antiparticle (pion, kaon, proton and their anti-particle) multiplicity ratios as a function of p_T also in the above mentioned three regions of ξ . The system X into which proton diffractively dissociates has net charge and baryon number +1. It is believed that initial charge and barion number should appear in the very forward direction leading to the equal amount of particles and antiparticles in the central region created by fragmentation and hadronization processes. However other scenarios are also possible where extra baryon is uniformly distributed over rapidity [2] or even appear close to the gap edge [3]. It is natural to expect that possible charge and baryon number transfer to central region will be better visible at small ξ where the number of antiparticle-particle pairs creation is smaller due to a generally smaller mass of the dissociated system X or due to the fact that the gap edge appears in the mid-rapidity region inside of the STAR fiducial region of pseudorapidity $|\eta| < 0.7$.

98 2. Monte Carlo Samples

99 MC samples used to correct data for detector effects were obtained by the embedding MC technique
100 [4], in which simulated particles are mixed with the real Zerobias events at the raw data
101 level. Zerobias data events used in the embedding were sampled over the entire data-taking period
102 in order to properly describe the data set used in the analysis. Two samples of embedding MC
103 were produced:

- 104 1. Single particle MC, in which particles are generated from flat distributions in η and p_T , in
105 order to have similar statistics in all bins.
- 106 2. The Schuler and Sjöstrand (SaS) model implemented in PYTHIA 8 with 4C tune.

107 Generated particles were passed through the full simulation of the STAR TPC and RP system
108 detectors using GEANT3 and GEANT4, respectively, and then embedded into real data sample.
109 These embedding events were next processed through the full event reconstruction chain.

110 It is preferred to get the detector corrections from a MC, which is dedicated to simulate
111 the studied physics process. However, for this purpose, the statistics in the MC should be several
112 times greater than in the analysed data sample. Since this is not possible with low efficiency of
113 TPC and TOF, the basic method of corrections used in the analysis for p_T and $\bar{\eta}$ distributions
114 is a method of factorization of global efficiency into the product of single-particle efficiencies. In
115 this way, statistically precise multidimensional corrections on TPC and TOF were obtained from
116 the single particle MC. The energy loss correction was also determined from the same MC sample.
117 The charged-particle multiplicity distributions were unfolded from the measured multiplicities of
118 TPC tracks based on the response matrix, which takes into account all detector effects. In this
119 procedure single particle MC samples were not used.

120 All other detector corrections were obtained from PYTHIA 8 4C (SaS). In order to keep
121 statistical precision coming from the corrections high, samples filtered on true-level values of ξ
122 and t (not necessarily with reconstructed proton track in RP) are used.

123 Several additional MC samples were generated, in which simulated particles were propagated
124 through full simulation and reconstruction chain but were not embedded into Zerobias events.
125 Systematic uncertainty related to hadronization of the diffractive system was determined by using
126 alternative hadronization models as implemented in HERWIG and EPOS. Results are compared
127 to model predictions from PYTHIA 8 4C (SaS), HERWIG, EPOS and alternative PYTHIA 8
128 model Minimum Bias Rockefeller (MBR) with A2 tune. EPOS predicts very large contribution
129 of forward protons, which originate from Non-Diffractive (ND) events and are well separated in
130 rapidity from other final state particles. This is the result of low mass excitation of the proton
131 remnant (< 1 GeV) leading to hadronization of the beam remnant back to the proton. Therefore
132 EPOS predictions were separated in two classes: diffractive (EPOS SD) modelled by Pomeron
133 exchange and ND modelled by low mass excitation of the proton remnant (EPOS SD'). Such
134 remnant treatment is very unique in EPOS compared to other string models, especially, to that
135 used in PYTHIA 8, where ND forward protons are rare and arise from string fragmentation and
136 hadronization. In all PYTHIA 8 models, diffractive cross-sections are scaled by the factors, which
137 were introduced in order to describe the full phase space [5, 6]. In the SaS model, the scaling
138 factors for SD and DD, F_{SD} and F_{DD} , are defined as a function of diffractive masses:

$$F_{SD} = \left(1 - \frac{M^2}{s}\right) \left(1 + \frac{c_{res} M_{res}^2}{M_{res}^2 + M^2}\right) \quad (2.1)$$

$$F_{DD} = \left(1 - \frac{M_a^2 + M_b^2}{s}\right) \left(\frac{sm_p^2}{sm_p^2 + M_a^2 M_b^2}\right) \times \left(1 + \frac{c_{res} M_{res}^2}{M_{res}^2 + M_a^2}\right) \left(1 + \frac{c_{res} M_{res}^2}{M_{res}^2 + M_b^2}\right) \quad (2.2)$$

139 where M and M_a, M_b are the invariant masses of the systems X and X_a, X_b for SD and DD,
140 respectively, $c_{\text{res}} = 2$ and $M_{\text{res}} = 2 \text{ GeV}/c^2$ were obtained from a fit to $pp/\bar{p}p$ data [5]. On
141 the other hand, in the MBR model the scaling factor is given as a function of the rapidity gap [6]:
142

$$S = \frac{1}{2} \left[1 + \text{erf} \left(\frac{\Delta y - \Delta y_S}{\sigma_S} \right) \right] \quad (2.3)$$

143 where Δy is the rapidity gap, $\Delta y_S = 2$ and $\sigma_S = 0.5$. As a result, diffractive cross sections are
144 artificially suppressed at relatively large values of $\xi (> 0.05)$. This artificial suppression significantly
145 changes predicted distribution of ξ and fractions of different processes in our fiducial phase space.
146 Therefore data is also compared with expectations obtained without suppression of the diffractive
147 cross sections (MBR-tuned).
148

149 Figure 2.1 (left) shows the distribution of ξ generated with EPOS (SD and SD+SD') and
150 PYTHIA 8 SD (SaS, MBR and MBR-tuned). PYTHIA 8 (MBR) predicts a strong dependence of
151 the cross section on ξ , which is much weaker in PYTHIA 8 (SaS and MBR-tuned) and the weakest
152 in EPOS. This difference between PYTHIA 8 SaS and MBR models is expected since they are
153 based on different Pomeron trajectories ($\epsilon_{\text{SaS}} = 0, \epsilon_{\text{MBR}} = 0.104$). Only 30% of events in EPOS
154 are SD, while the rest are SD'. Since all MC samples were generated with forward proton filter
155 (a cut on the proton position in front of the RPs), the shapes of $|t|$ distributions for these samples
156 are biased. In order to compare them with each other, only their ratio to PYTHIA 8 (MBR)
157 predictions is presented as a function of $|t|$. EPOS SD is only relevant for very small $|t|$ (below
158 0.04 GeV^2/c^2) and is suppressed in the STAR acceptance region, $0.04 < |t| < 0.16$, where EPOS
159 SD' contribution dominates. The t -slope is very different for EPOS SD and EPOS SD', while it
160 is similar for EPOS SD+SD' and PYTHIA 8 (SaS and MBR-tuned), EPOS SD and PYTHIA 8
(MBR). This is related to the smaller average value of ξ for EPOS SD and PYTHIA 8 (MBR)
compared to EPOS SD+SD' and PYTHIA 8 (SaS and MBR-tuned).

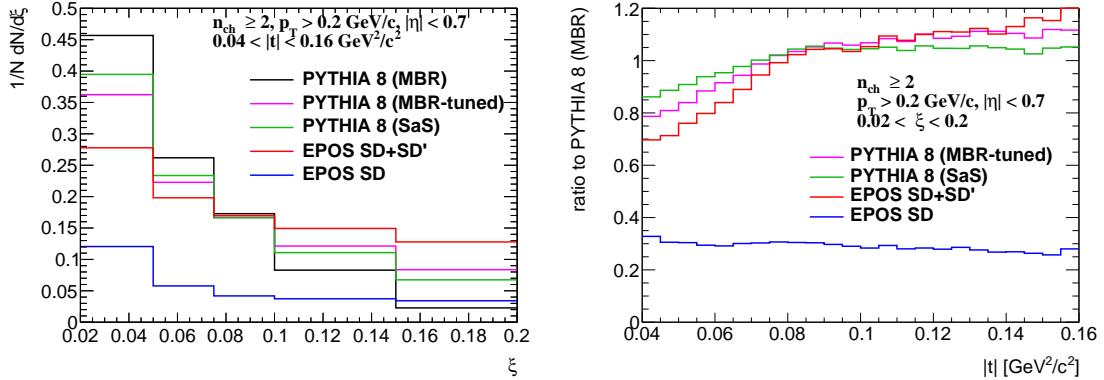


Figure 2.1: (left) ξ distribution for various MC generators and (right) ratios of different MCs to PYTHIA 8 (MBR) predictions as a function of $|t|$ at $\sqrt{s} = 200 \text{ GeV}$.

161

3. Data Sample and Event Selection

The data sample used in this analysis was collected in proton-proton collisions at centre-of-mass energy of $\sqrt{s} = 200$ GeV during RHIC Run 15.

All of the studies in this analysis use data from only the SDT trigger condition, which was the main trigger designed for SD studies in Run 15. The logic of the trigger was formed by the following conditions combined with the logical AND:

1. RP_EOR || RP_WOR - signal in at least one RP on any side of the STAR central detector,
2. veto on any signal in small BBC tiles or ZDC on the triggered RP side of the STAR central detector,
3. at least two TOF hits.

The above requirements were imposed in accordance with the diffractive event topology. Veto on any signal in small BBC tiles and ZDC allowed to accept only events with rapidity gap and reject diffractive events with simultaneous pile-up event. The requirement of at least two TOF hits was applied to ensure activity in the mid-rapidity.

Integrated luminosity delivered by the RHIC to the STAR experiment in pp collisions during Run 15 amounts to 185.1 pb^{-1} [9], whereas about 34.4M SDT events were gathered by the STAR detector, shown in Fig. 3.1, which corresponds to 16 nb^{-1} of integrated luminosity.

Event Selection

Events were selected from those passing the SDT trigger condition. In order to remove events of poor quality and to suppress background the following conditions were required:

1. trigger signals in exactly two stations of one arm of RP system (this requirement divides the sample into four sub-samples, which were later analysed independently, e.g. for background studies),
2. any trigger signal in small BBC tiles on the opposite side of the STAR central detector to the triggered RP station,

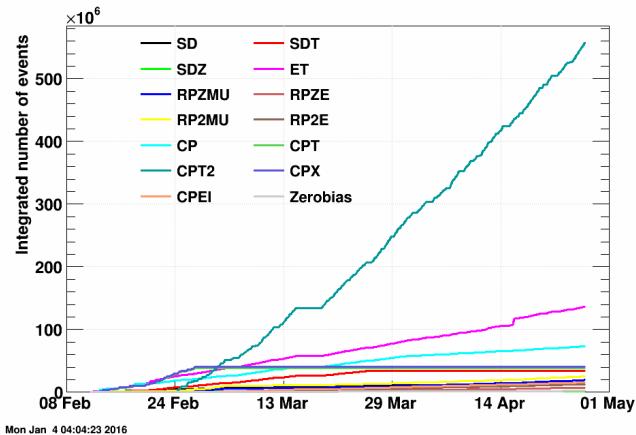


Figure 3.1: Cumulative number of events collected for each trigger in the RP data stream during Run 15 [7, 8].

- 188 3. exactly one proton track in the above RP stations with $0.02 < \xi < 0.2$ and $0.04 < -t <$
 189 $0.16 \text{ GeV}^2/c^2$.
- 190 4. exactly one vertex reconstructed from TPC tracks matched with hits in TOF (later in the text
 191 such vertex is referred as a TOF vertex),
- 192 5. TOF vertex within $|V_z| < 80 \text{ cm}$ - events with vertices away from the nominal IP have low
 193 acceptance for the central and forward tracks,
- 194 6. at least two but no more than eight primary TPC tracks, $2 \leq n_{\text{sel}} \leq 8$, matched with hits
 195 in TOF and satisfying the selection criteria described in Sec. 3.1,
- 196 7. if there are exactly two primary tracks satisfying the above criteria and exactly two global
 197 tracks used in vertex reconstruction (Sec. 5.1), the longitudinal distance between these global
 198 tracks should be smaller than 2 cm, $|\Delta z_0| < 2 \text{ cm}$.

199 Figure 3.2 shows the multiplicity of TOF vertices n_{vrt} (left) and the z -position of reconstructed
 200 vertices in single TOF vertex events (right). Data are compared to embedded PYTHIA 8 SD
 201 sample. These distributions are not significantly process-dependent, therefore, contributions from
 202 other processes are not included in these plots. Most events with $n_{\text{vrt}} > 1$ originate from in-time
 203 pile-up and are excluded from the analysis.

204 **ZDC Veto**

205 The SDT trigger conditions imposed a veto on any signal in the same-side ZDC. However, all MC
 206 samples do not contain ZDC simulation. To check the impact of this veto on the measurement,
 207 the total energy of neutral particles, such as n , γ , π^0 , produced within ZDC acceptance ($|\eta| > 6$)
 208 was measured using true-level PYTHIA 8 (SaS). In most of the events, the energy measured on
 209 the proton side of the IP is smaller than trigger thresholds (as shown in Fig. 3.3). Therefore,
 210 the ZDC veto has a negligible effect on the analysis and ZDC simulation is not needed.

211 **3.1 Track Selection**

212 The following quality cuts had to be passed by the selected primary tracks:

- 213 1. the tracks must be matched with hits reconstructed in TOF,
 214 2. the number of the TPC hits used in the helix fit $N_{\text{hits}}^{\text{fit}}$ must be greater than 24,

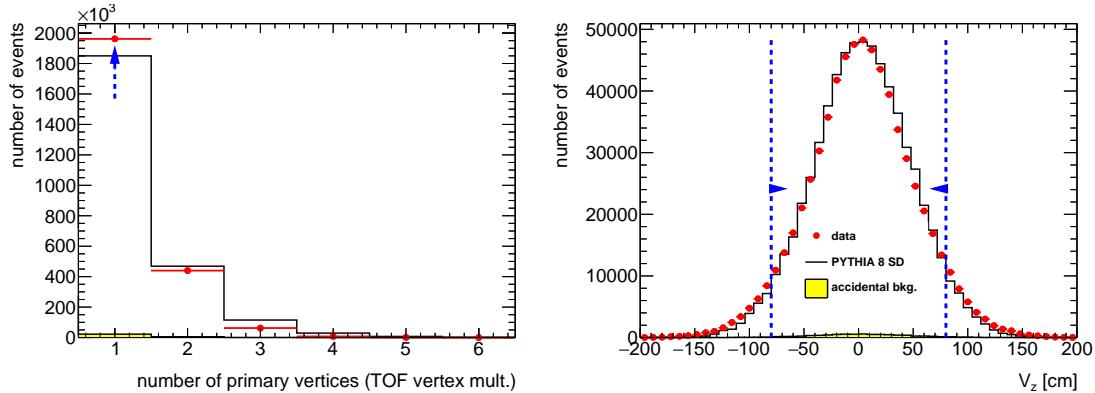


Figure 3.2: (left) Vertex multiplicity and (right) the z -position of reconstructed vertices in single TOF vertex events before applying the cut on the quantity shown. Blue lines indicate regions accepted in the analysis.

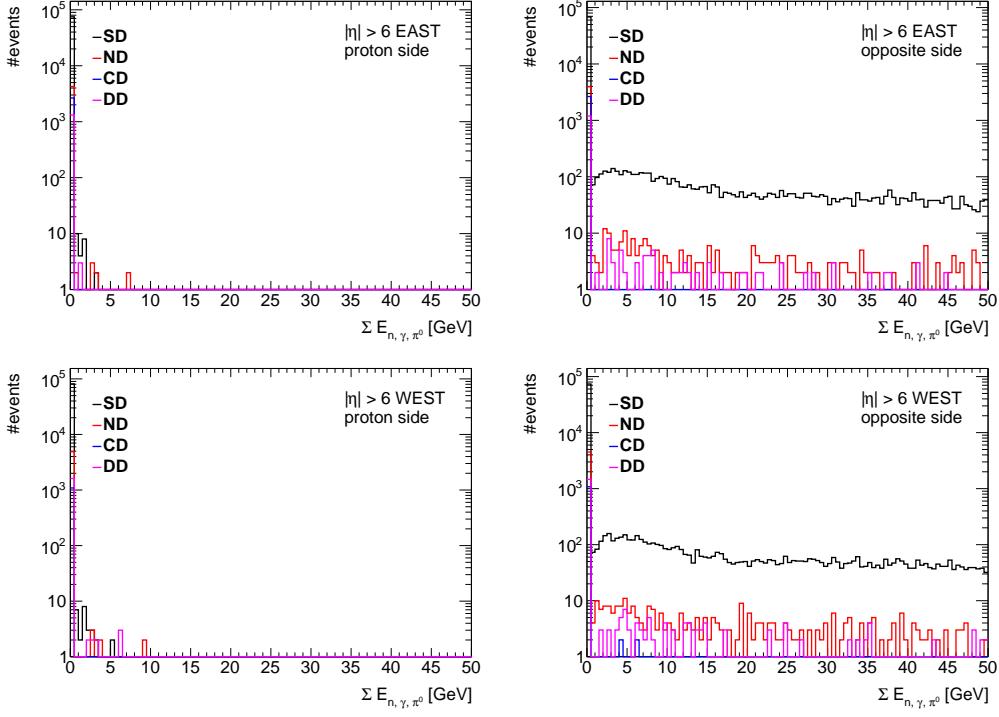


Figure 3.3: Total energy of neutral particles (n, γ, π^0) produced within ZDC acceptance ($|\eta| > 6$) for events in which forward-scattered proton is on (top) west and (bottom) east side of the IP. Distributions are presented separately for neutral particles produced on (left) the proton and (right) opposite side of the IP. PYTHIA 8 predictions for different processes are shown as colour histograms.

- 215 3. the number of the TPC hits used to determine the dE/dx information $N_{\text{hits}}^{\text{dE/dx}}$ must be
216 greater than 14,
- 217 4. the transverse impact parameter with respect to the beamline d_0 must be less than 1.5 cm,
- 218 5. the radial component of the distance of the closest approach between the global helix and
219 the vertex DCA_{xy} must be less than 1.5 cm,
- 220 6. the absolute magnitude of longitudinal component of the distance of the closest approach
221 between the global helix and the vertex $|\text{DCA}_z|$ must be less than 1 cm,
- 222 7. the track's transverse momentum p_T must be greater than 0.2 GeV/c,
- 223 8. the track's absolute value of pseudorapidity $|\eta|$ must be smaller than 0.7.

224 The $N_{\text{hits}}^{\text{fit}}$ and $N_{\text{hits}}^{\text{fit}}/N_{\text{hits}}^{\text{possible}}$ cuts are used to reject low quality TPC tracks and avoid track
225 splitting effects. The d_0 and global DCA_{xy} , $|\text{DCA}_z|$ cuts are used to select tracks that originate
226 from the primary interaction vertex. The cut on $N_{\text{hits}}^{\text{dE/dx}}$ is used to ensure that selected tracks
227 have sufficient energy loss information for particle identification purposes. In this analysis tracks
228 without identification are required to have $p_T > 0.2$ GeV/c and $|\eta| < 0.7$ due to high track
229 reconstruction and TOF matching efficiencies in this region. For the identified particle-antiparticle
230 ratio analysis, where charged pions, charged kaons and (anti)protons are measured, the p_T cut
231 was increased for kaons and (anti)protons to 0.3 and 0.4 GeV/c, respectively. The distributions
232 of the DCA_{xy} , $|\text{DCA}_z|$, d_0 , $N_{\text{hits}}^{\text{fit}}$ and $N_{\text{hits}}^{\text{dE/dx}}$ quantities together with applied cuts are shown in

233 Fig. 3.4, while the p_T , η and the azimuthal angle, ϕ , of the reconstructed tracks are shown in
 234 Fig. 3.5. Data are compared to embedded PYTHIA 8 SD sample.

235 The azimuthal angle of the reconstructed tracks for runs ≤ 16073050 is not described by PY-
 236 THIA 8. The inner sector #19 in the TPC was dead for this running period and some effects
 237 related to it were presumably not taken into account in the TPC detector simulation. There-
 238 fore, additional data-driven corrections to track efficiencies are used [1]. The larger accidental
 239 background is observed for runs > 16073050 , probably due to the higher bunch intensities in this
 240 running period [10].

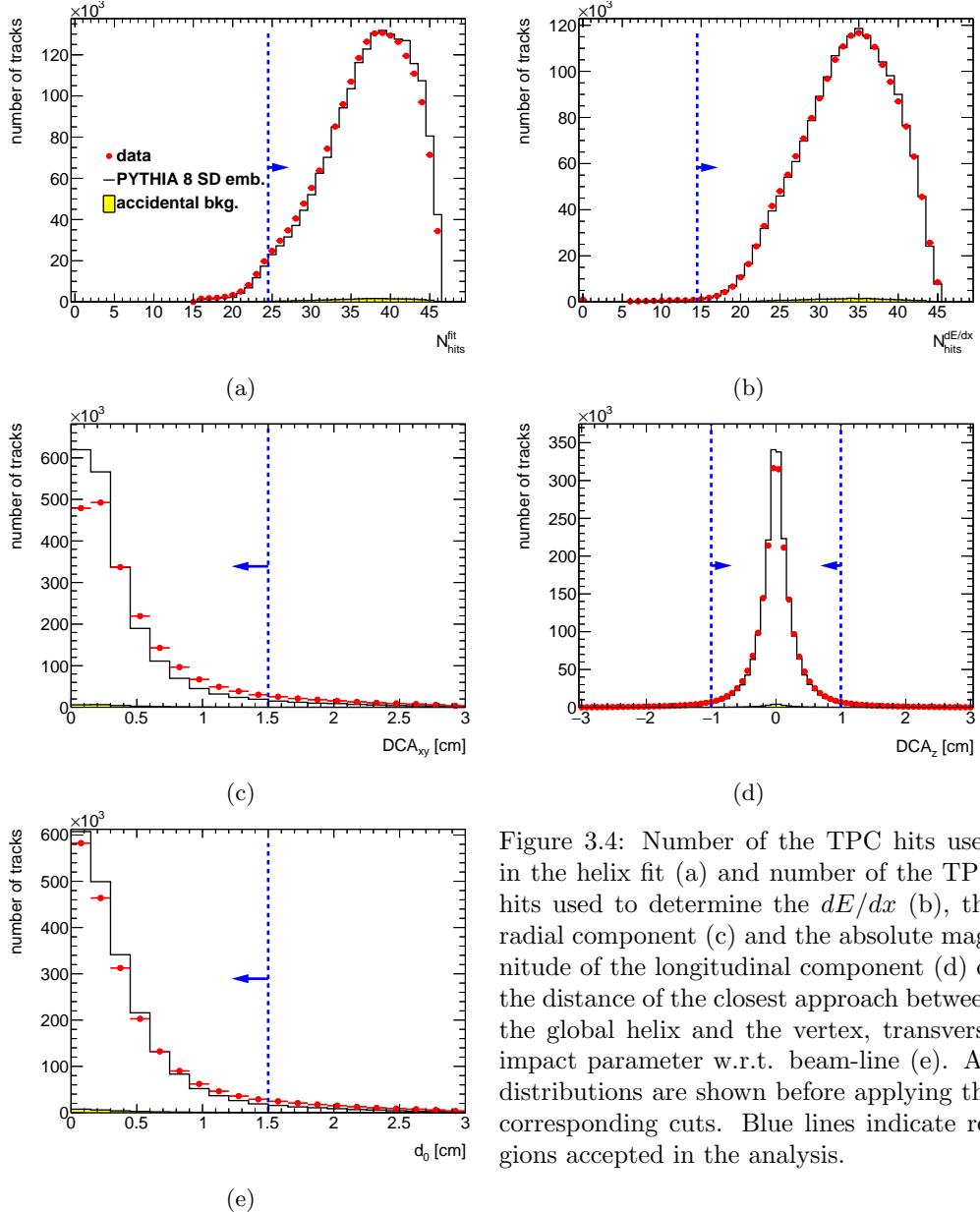


Figure 3.4: Number of the TPC hits used in the helix fit (a) and number of the TPC hits used to determine the dE/dx (b), the radial component (c) and the absolute magnitude of the longitudinal component (d) of the distance of the closest approach between the global helix and the vertex, transverse impact parameter w.r.t. beam-line (e). All distributions are shown before applying the corresponding cuts. Blue lines indicate regions accepted in the analysis.

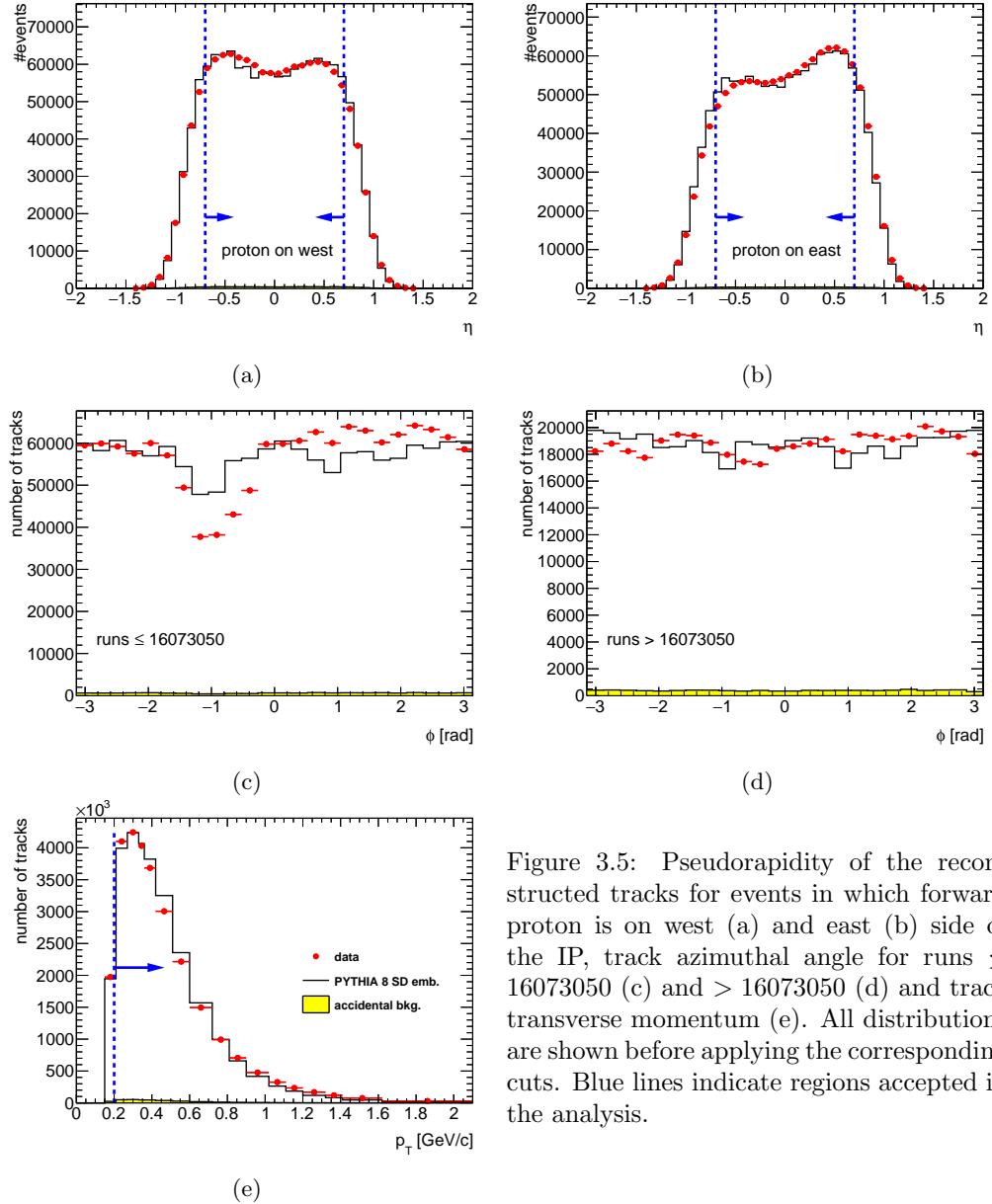


Figure 3.5: Pseudorapidity of the reconstructed tracks for events in which forward proton is on west (a) and east (b) side of the IP, track azimuthal angle for runs ≤ 16073050 (c) and > 16073050 (d) and track transverse momentum (e). All distributions are shown before applying the corresponding cuts. Blue lines indicate regions accepted in the analysis.

241 3.2 Fiducial Region of the Measurement

242 A fiducial phase space of measurement is defined by the following criteria. Primary charged
 243 particles are defined as charged particles with a mean lifetime $\tau > 300$ ps, either directly produced
 244 in pp interaction or from subsequent decays of directly produced particles with $\tau < 30$ ps. Primary
 245 charged particles had to be contained within the kinematic range of $p_T > 0.2$ GeV/c and $|\eta| < 0.7$.
 246 The results are corrected to the region of the total number of primary charged particles (without
 247 identification), $2 \leq n_{ch} \leq 8$. In identified charged antiparticle to particle ratio measurement, the
 248 lower transverse momentum limit was set for the analysed particles as follows: 0.2 GeV/c (pions),
 249 0.3 GeV/c (kaons), 0.4 GeV/c (protons and antiprotons).

250 The measurements were performed in a fiducial phase space of the forward-scattered protons
 251 of $0.04 < -t < 0.16$ GeV $^2/c^2$ and $0.02 < \xi < 0.2$. Figure 3.6 shows that the fraction of events
 252 containing at least two primary charged particles, $\epsilon_{n_{ch}\geq 2}(\log_{10} \xi)$, is reduced by half for $\xi < 0.02$
 253 compared to the region of larger ξ . In addition, the accidental background contribution at $\xi < 0.02$
 254 is significant and approximately equal to 10% (Sec. 4). For these reasons the lower ξ cut was
 255 introduced. The upper ξ cut was required since the region of larger ξ is dominated by Double

256 Diffraction (DD) and ND (Sec. 4.2). The joint RP acceptance and track reconstruction efficiency
 257 was defined as the probability that true-level proton was reconstructed as a track passing the
 258 selection criteria. This efficiency was calculated as a function of $-t$ for three ranges of ξ separately
 259 and is shown in Fig. 3.7. Events were accepted only if the reconstructed values of $-t$ for protons
 260 fall within $> 5\%$ acceptance regions, which were required to be the same for each ξ region and
 261 similar to those defined in the elastic analysis [11]. Therefore, cuts on $0.04 < -t < 0.16 \text{ GeV}^2/c^2$
 262 were introduced. All measured observables are presented in three ξ regions: $0.02 < \xi < 0.05$,
 263 $0.05 < \xi < 0.1$ and $0.1 < \xi < 0.2$.

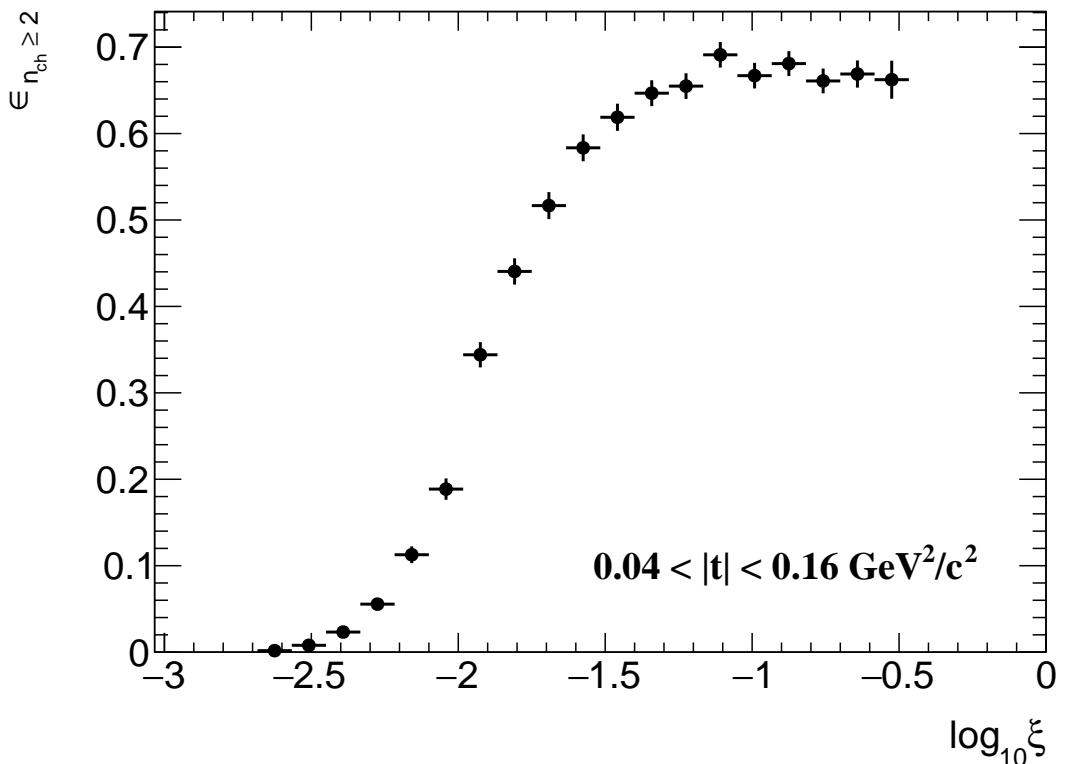


Figure 3.6: $\epsilon_{n_{ch} \geq 2}$ as a function of $\log_{10} \xi$ calculated from PYTHIA 8 (MBR).

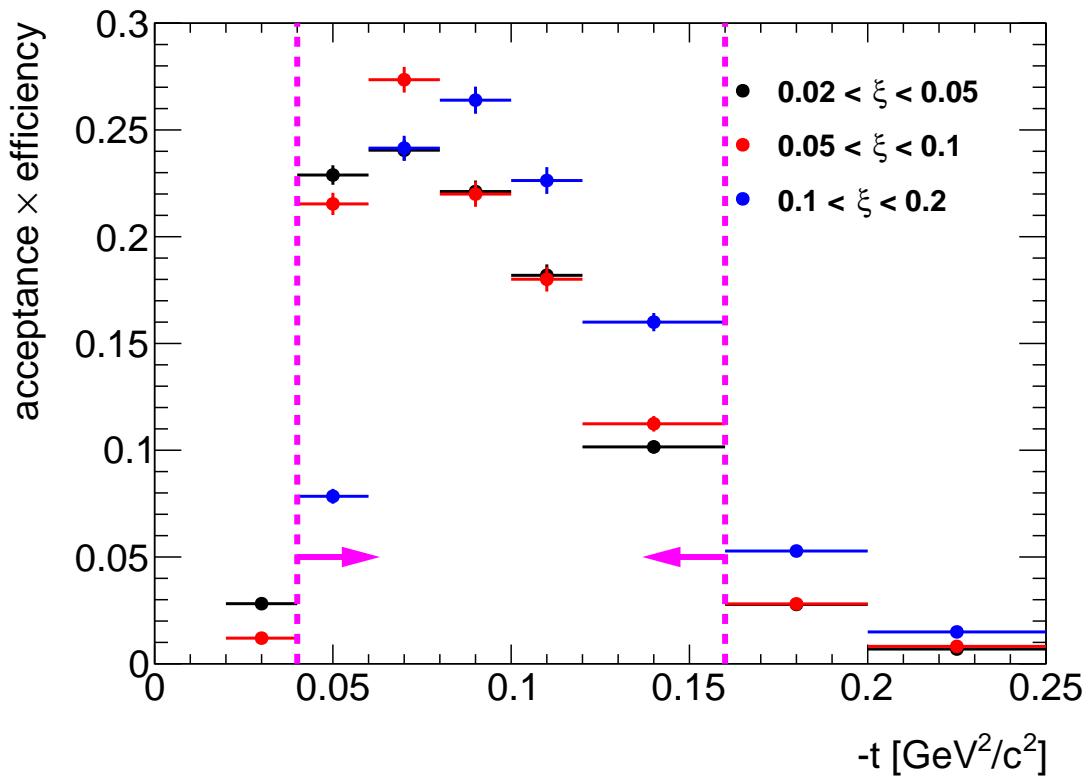


Figure 3.7: RP acceptance and track reconstruction efficiency as a function $-t$ in three ranges of ξ , calculated using PYTHIA 8 4C (SaS). Magenta lines indicate region accepted in the analysis.

4. Background Contribution

The background contributions to the charged-particle distributions can be divided into event-level and track-level backgrounds, and are described in detail below:

- Accidental background refers to events which do not originate from a single collision of two protons.
- Track backgrounds from non-primary tracks consist of secondary tracks and fake tracks; the first come mostly from decays, the short-lived particles with mean life $30 < \tau < 300$ ps, or secondary interactions with the detector dead material, while the second comes from the track reconstruction algorithms and out-of-time pile-up with no corresponding true particles.

Accidental Background

The accidental backgrounds (same bunch pile-up background) are quantified using data-driven method and defined as a process where in single bunch crossing there is coincidence of two interactions, where any single-side proton signal is collected in coincidence with an independent signal in the TPC+TOF+BBC detector. This type of background may come from the overlap of a signal in RP (proton from beam-halo, low mass SD process without activity in TOF, elastic or low mass Central Diffraction (CD) processes with undetected proton on the other side) with a signal in TPC+TOF+BBC (mainly ND events without forward-scattered proton).

The accidental background contribution was calculated from Zerobias data (colliding bunches), where two signatures of such background were investigated: the reconstructed proton in RP and the reconstructed vertex from TPC tracks matched with TOF. The analysis was done for each RP arm separately and thus the Zerobias data was firstly required to pass the following criteria:

1. no trigger in any RP or trigger in exactly one arm (two RPs) with exactly one reconstructed proton track in that arm,
2. veto on any signal in small BBC tiles or ZDC on the same side of the IP as the RP arm under consideration,
3. no or exactly one reconstructed vertex with at least two TOF-matched tracks passing the quality criteria. The latter includes also signal in BBC small tiles on the opposite side of the IP to the RP arm under study.

The sample of selected Zerobias data with total number of events N was divided into four classes:

$$N = N_{PS} + N_{RS} + N_{PT} + N_{RT} \quad (4.1)$$

where: N_{PS} is the number of events with reconstructed proton in exactly one RP and reconstructed TOF vertex, N_{RS} is the number of events with no trigger in any RP and reconstructed TOF vertex, N_{PT} is the number of events with reconstructed proton in exactly one RP and no reconstructed TOF vertex, N_{RT} is the number of events with no trigger in any RP and no reconstructed TOF vertex. Since the signature of the signal is a reconstructed proton in exactly one RP and a reconstructed TOF vertex, the number of such events can be expressed as:

$$N_{PS} = N(p_3 + p_1 p_2) \quad (4.2)$$

where: p_1 is the probability that there is a reconstructed proton in RP and there is no reconstructed TOF vertex, p_2 is the probability that there is no reconstructed proton in RP and there is a reconstructed TOF vertex, p_3 is the probability that there is a reconstructed proton in RP and there is a reconstructed TOF vertex (not accidental).

304 The other classes of events given in Eq. (4.1) can be expressed in terms of the above probabilities
 305 as:

$$\begin{aligned} N_{RS} &= N(1 - p_1)p_2(1 - p_3) \\ N_{PT} &= N(1 - p_2)p_1(1 - p_3) \\ N_{RT} &= N(1 - p_1)(1 - p_2)(1 - p_3) \end{aligned} \quad (4.3)$$

306 Finally, the accidental background contribution $A_{\text{bkg}}^{\text{accidental}}$ is given by:

$$A_{\text{bkg}}^{\text{accidental}} = \frac{p_1 p_2}{p_3 + p_1 p_2} = \frac{N_{RS} N_{PT} N}{N_R N_T N_{PS}} \quad (4.4)$$

307 where: $N_R = N_{RS} + N_{RT}$ and $N_T = N_{PT} + N_{RT}$.

308 The shapes of the accidental background related to TPC distributions come from the above
 309 Zerobias data events which pass all the analysis selection except having no trigger in any RP.
 310 The templates corresponding to RP distributions are from protons in the above data sets but
 311 with no reconstructed TOF vertex. The normalization is given by Eq. (4.4). Figure 4.1 shows
 312 distributions of the reconstructed ξ with the accidental background contribution for events with
 313 proton reconstructed in EU, ED, WU and WD arms. Accidental background in the range of
 314 $0.02 < \xi < 0.2$ is below 1% and increases to 10% at $\xi < 0.02$. Unphysical negative values of
 315 reconstructed ξ are due to the detector resolution.

316 The selection of Zerobias events, which is not unique, may provide some bias to the normalization
 317 of the accidental background. As a systematic check, two criteria for Zerobias selection were
 318 changed to:

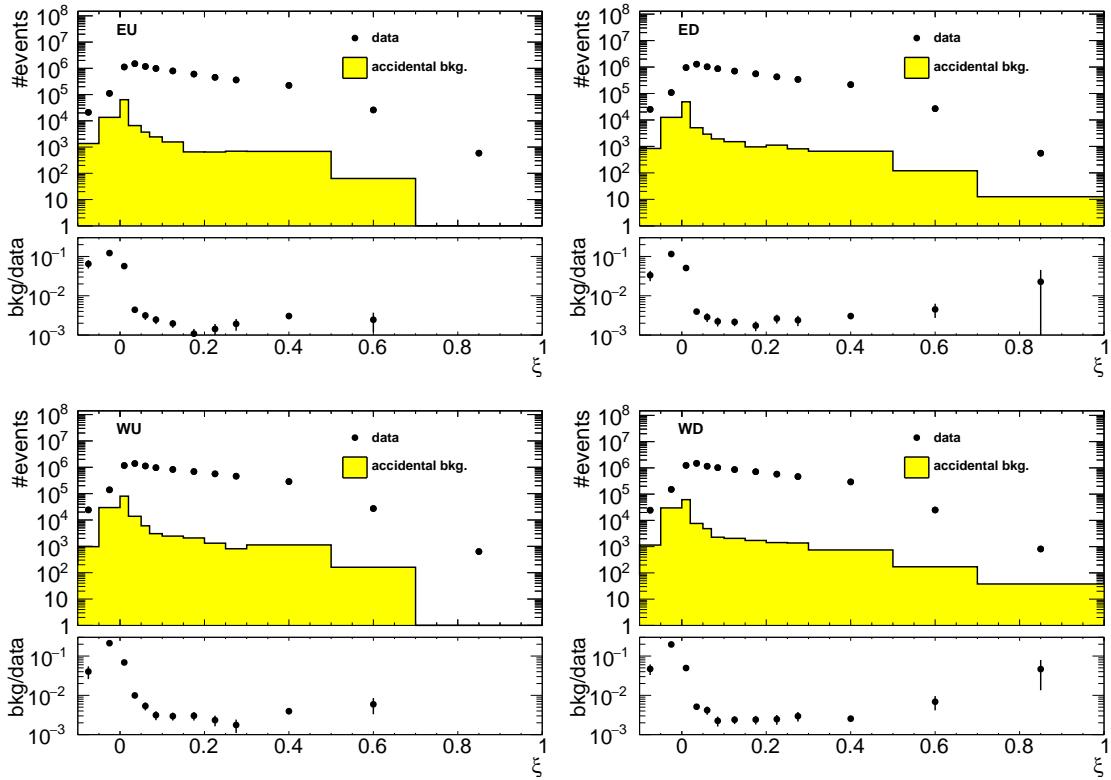


Figure 4.1: Uncorrected distributions of the reconstructed ξ for events with proton reconstructed in (top left) EU, (top right) ED, (bottom left) WU and (bottom right) WD arms. Data is shown as black markers, whereas the accidental background contribution is shown as yellow histogram. The ratio of accidental background and data is shown in the bottom panels.

- 319 1. no trigger in any RP or trigger in exactly one arm (two RPs) with *no more* than one
 320 reconstructed proton track in that arm, i.e. events with trigger signals in exactly one arm
 321 and without reconstructed proton track in that arm were also used,
 322 2. no or exactly one reconstructed TOF vertex (*without any additional requirements*), i.e.
 323 events with a reconstructed TOF vertex that does not have at least two primary tracks
 324 satisfying the selection criteria (Sec. 3.1), or with a reconstructed TOF vertex that is out of
 325 the range of $|V_z| < 80$ cm, were also accepted. The requirement of signal in BBC small tiles
 326 remains unchanged.

327 As a result of this change, the accidental background normalization increases of about 50% with
 328 respect to the nominal value. A symmetric systematic uncertainty of 50% of the normalization of
 329 accidental background is applied to the measurement.

330 4.1 Background from Non-Primary Tracks

331 Reconstructed tracks matched to a non-primary particle, so-called background tracks, originate
 332 mainly from the following sources:

- 333 • decays of short-lived primary particles with strange quark content (mostly K^0, Λ^0),
 334 • photons from π^0 and η decays which are converting to e^+e^- ,
 335 • hadronic interactions of particles with the beam-pipe or detector dead material.

336 Figure 4.2 (left) shows the background from non-primary tracks, $f_{\text{bkg}}(p_T, \eta)$, as a function
 337 of tracks' p_T and η , predicted by PYTHIA 8 SD model. There were no differences observed in
 338 the background contribution in different ξ ranges, hence, all three ξ ranges were merged for this
 339 study. The highest background fraction, which varies between 5 – 10%, was found to be at low
 340 p_T .

341 Figure 4.2 (right) shows the background track contribution to reconstructed tracks as a function
 342 of p_T and η calculated from EPOS SD+SD'. The differences between PYTHIA 8 and EPOS,
 343 which are up to 50% for $p_T > 0.5$ GeV/c (as shown in Fig. 4.3), were symmetrized and taken as
 344 a systematic uncertainty.

345 There is also a small (< 0.5%) contribution from fake tracks, $f_{\text{fake}}(p_T, \eta)$, i.e. tracks not associated
 346 with true-level particles, coming from out-of-time pile-up or formed by a random combination
 347 of TPC hits. The change by $\pm 100\%$ in this contribution was taken as a systematic uncertainty.

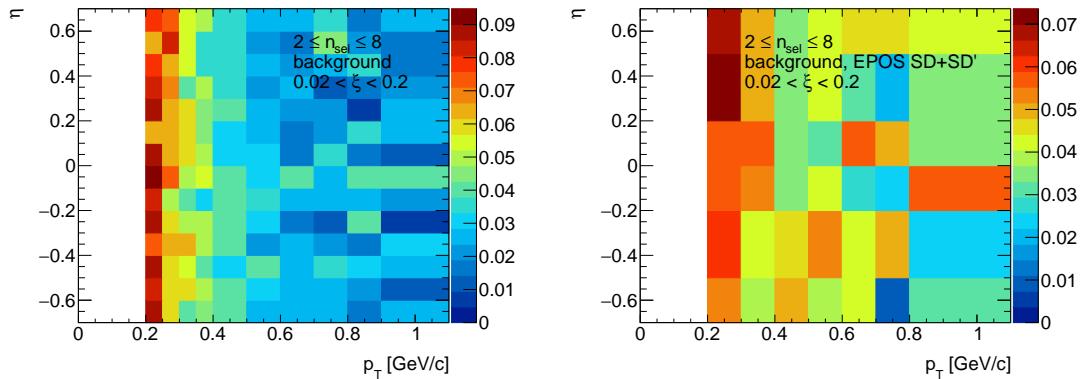


Figure 4.2: Distribution of fraction of selected tracks associated with non-primary particles in the range $0.02 < \xi < 0.2$ as predicted by (left) PYTHIA 8 4C (SaS) embedding and (right) EPOS SD+SD'.

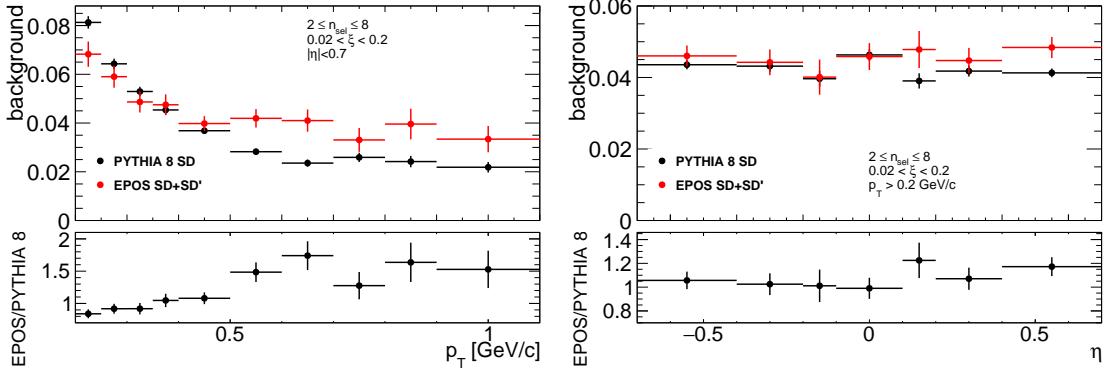


Figure 4.3: PYTHIA 8 SD and EPOS SD+SD' predictions of fraction of selected tracks associated with non-primary particles as a function of (left) p_T and (right) η . The ratio of EPOS and PYTHIA 8 predictions is shown in the bottom panels.

348 Proton Background

349 Secondary particles can be created due to the interaction of particles with detector dead-material.
 350 The proton sample contains background from such protons knocked out from the detector materials [12].
 351 Most of these protons have large DCA to the primary vertex and are not associated with
 352 it. However, the protons with small DCA are included in the primary track sample. Antiprotons
 353 do not have knockout background, hence the DCA tail is almost absent in their DCA distributions.

354 The fraction of knock-out background protons depends on a number of factors, including
 355 the amount of detector material, analysis cuts and the ξ of diffractive proton. While it is natural
 356 to calculate the fractions of primary and background protons in the MC sample, the MC models
 357 do not necessarily predict the fraction of knock-out background protons without any bias. Hence,
 358 data-driven methods should be used to calculate this type of background.

359 In order to correct for the knock-out background protons, sample enriched in proton back-
 360 ground was used for background normalization, where DCA_{xy} , DCA_z and d_0 cuts were aban-
 361 doned. Additionally, at least one, instead of exactly one, reconstructed vertex was allowed in this
 362 sample. Figures 4.4 and 4.5 show the DCA distributions of protons and antiprotons, respectively,
 363 for nominal (bottom) and background enriched (top) samples. The distributions for other p_T
 364 and ξ regions are shown in Appendix A. The protons and antiprotons are selected by a dE/dx
 365 cut of $-1 < n\sigma_{p,\bar{p}} < 3$ where $n\sigma_{p,\bar{p}}$ is given by Eq. (7.10). In some p_T regions, the dE/dx of
 366 (anti)protons and pions starts to overlap, hence, the asymmetric $n\sigma_{p,\bar{p}}$ cut was introduced in or-
 367 der to select as clean (anti)proton sample as possible. The fraction of knock-out protons within
 368 the selected sample is determined via MC template fits. The templates of reconstructed tracks
 369 with dE/dx corresponding to the proton and antiproton are obtained from PYTHIA 8 embedding
 370 MC separately for:

- 371 • primary (anti)protons,
- 372 • knock-out background protons (labeled as dead-material),
- 373 • fake tracks,
- 374 • secondary particles with dE/dx of (anti)proton (labeled as wrong PID - sec.),
- 375 • tracks associated with primary (anti)protons, but with the reconstructed vertex not matched
 376 to true-level primary vertex (labeled as wrong vtx),
- 377 • reconstructed track is partially matched to true-level particle (labeled as wrong match, track
 378 to true-level particle matching is described in [1]), i.e. track and true-level particle have
 379 appropriate number of common hit points but the distance between true-level particle and

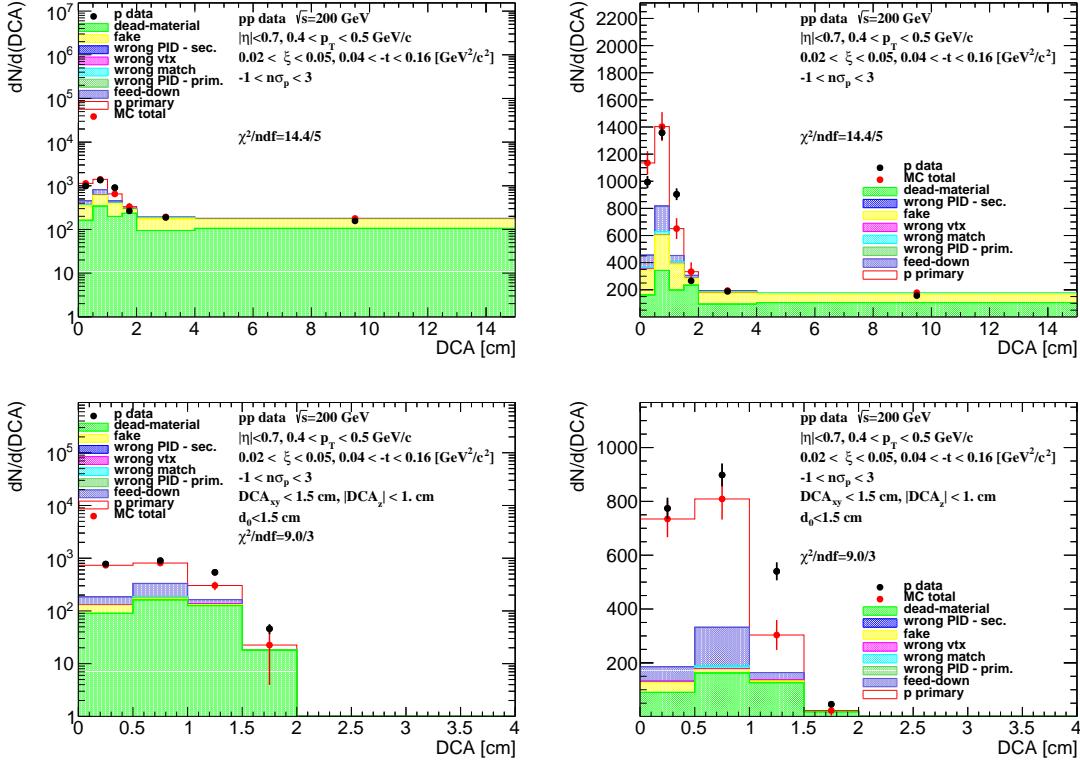


Figure 4.4: The DCA distributions of protons for $0.4 < p_T < 0.5 \text{ GeV}/c$ shown for single range of $0.02 < \xi < 0.05$ (shown in log and linear scale in left and right column, respectively). The MC contributions are shown after scaling the dead-material template to the tail of large DCA values, $2 < \text{DCA} < 15 \text{ cm}$. (top) Background enriched samples were used in the normalization procedure, whereas (bottom) the proton background was estimated from the nominal sample.

380 track is too large, $\delta^2(\eta, \phi) > (0.15)^2$, thus, track is not considered as primary particle
381 according to discussion in [1],

- 382 • primary particles with dE/dx of (anti)proton (labeled as wrong PID - prim.),
383 • (anti)proton as a product of short-lived decays, mainly Λ^0 (labeled as feed-down).

384 First, the background enriched sample was analyzed (Fig. 4.4, top), where the template of
385 knock-out background protons was normalized to the number of events in the fake-subtracted tail
386 of the DCA distribution, $2 < \text{DCA} < 15 \text{ cm}$. Next the knock-out proton and fake background
387 was subtracted from the DCA distribution and the sum of other templates was normalized to
388 the number of events in the signal region, $\text{DCA} < 1.5 \text{ cm}$.

389 The fraction of the knock-out proton background in the signal region, $\text{DCA} < 1.5$, was es-
390 timated from the nominal sample (Fig. 4.4, bottom), where DCA_{xy} , DCA_z and d_0 track cuts
391 were applied and exactly one reconstructed vertex was required. The normalization of each MC
392 contribution was kept the same as that estimated for the background enriched sample. Figure 4.6
393 shows the knock-out proton background as a function of p_T in three ranges of ξ . The following
394 functional form was found to describe the background protons:

$$f_{\text{bkg}}^p(p_T) = p_0 \exp(p_1 p_T) \quad (4.5)$$

395 where p_0 and p_1 are free parameters obtained from a fit.

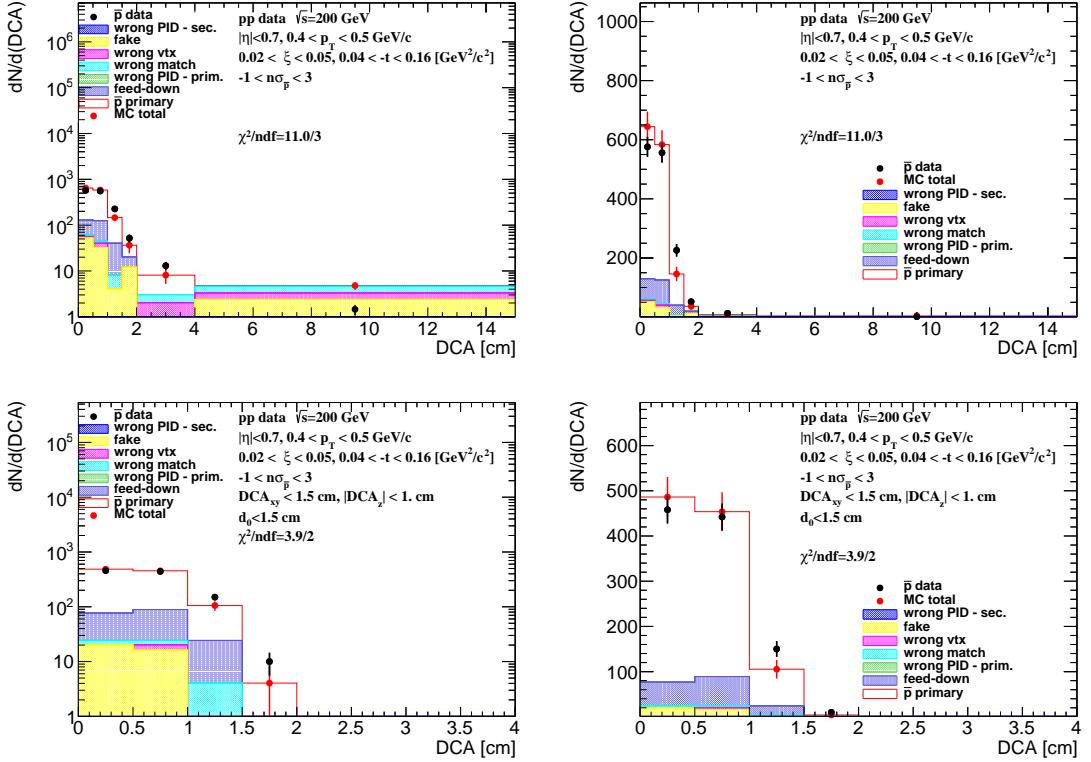


Figure 4.5: The DCA distributions of antiprotons for $0.4 < p_T < 0.5 \text{ GeV}/c$ shown for one range of $0.02 < \xi < 0.05$ (log and linear scale in left and right column, respectively). The MC contributions are shown as colour histograms. (top) Background enriched and (bottom) nominal samples were used.

396 The obtained fraction of knock-out background protons is approximately 20% at $p_T = 0.45$
397 GeV/c and less than 10% at $p_T = 1.0 \text{ GeV}/c$. In PYTHIA 8 SD predictions (also shown in Fig. 4.6),
398 such fraction is much smaller and equals to approximately 7% at $p_T = 0.45 \text{ GeV}/c$ and about 5%
399 at $p_T = 1.0 \text{ GeV}/c$. This may suggest that there are differences in the amount of dead material in
400 front of TPC between data and simulation, which is consistent with the studies presented in [1].

401 Figure 4.5 shows the corresponding DCA distributions with MC templates for antiprotons,
402 where the background from knock-out particles is not present. Therefore, there was no need for
403 any fit to be performed in this comparison. The MC templates fairly well describe the DCA
404 distribution for both, protons, after tuning the fraction of knock-out background to data, and
405 antiprotons.

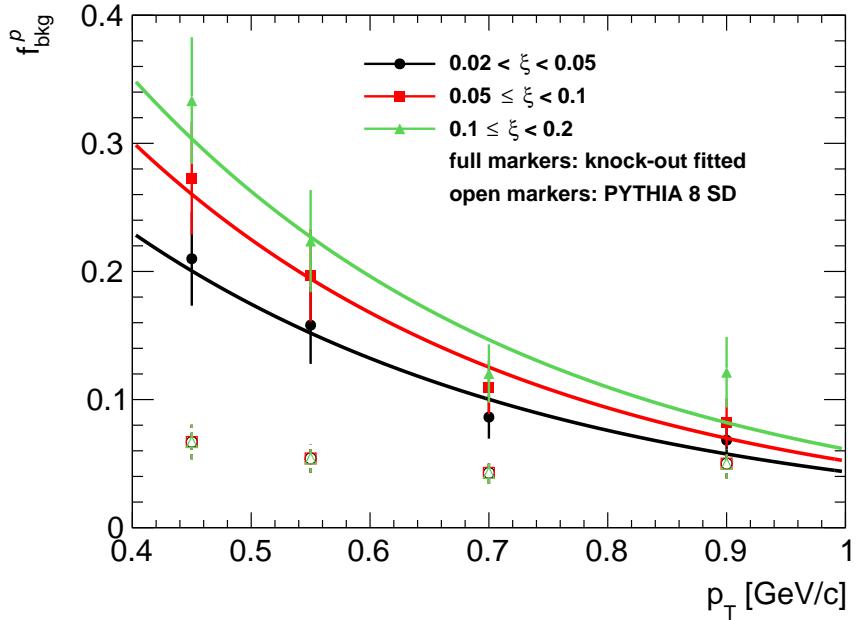


Figure 4.6: The fraction of knock-out proton background as a function of p_T in three ranges of ξ with fitted parametrizations. Full markers represent fitted knock-out background and open markers represent PYTHIA 8 SD predictions.

406 Systematic Uncertainty Related to Proton Background

407 The knock-out proton background estimation introduces systematic uncertainties. First, the nor-
 408 malization interval of the knock-out proton background template in the background enriched
 409 sample was changed to $4 < \text{DCA} < 15$ cm. This introduced a relative systematic uncertainty of
 410 up to 30% for $p_T \approx 0.9$ GeV/c.

411 The knock-out proton background contribution was parameterized as it is shown in Eq. (4.5).
 412 The systematic uncertainty related to the parameterization procedure was estimated by varying the
 413 parameters, p_0 and p_1 , by their statistical uncertainties ($\pm 1\sigma$). As a result, a relative systematic
 414 uncertainties of about 10% were obtained.

415 Differences in the shape of the DCA distribution between data and MC can affect the knock-
 416 out proton background estimation procedure. Figure 4.7 (top left) shows the data to MC ratio of
 417 the number of events in the background dominated region, $2 < \text{DCA} < 15$ cm. Since this region
 418 is used to estimate background normalization, and the shape of the DCA distribution in the data
 419 differs from that observed in the simulation, the predicted background in the $\text{DCA} < 1.5$ cm region
 420 can change. Thus, the following functional form was used to estimate the slope between data and
 421 MC:

$$\frac{\text{data}}{\text{MC}}(\text{DCA}) = A(\text{DCA} - 8.5) + B \quad (4.6)$$

422 where A (slope) and B are fit free parameters. Differences in slope between data and MC were used
 423 to estimate how many more background tracks would fit into the signal region and a systematic
 424 uncertainty, which varies up to 5% for $0.02 < \xi < 0.05$, was introduced.

425 All above components of the systematic uncertainty related to the knock-out proton back-
 426 ground, shown in Fig. 4.7, are added in quadrature. Those related to the fit range and the shape
 427 of the proton background are symmetrized. Figure 4.8 shows the fraction of knock-out proton
 428 background in three ranges of ξ and the total systematic uncertainty related to it.

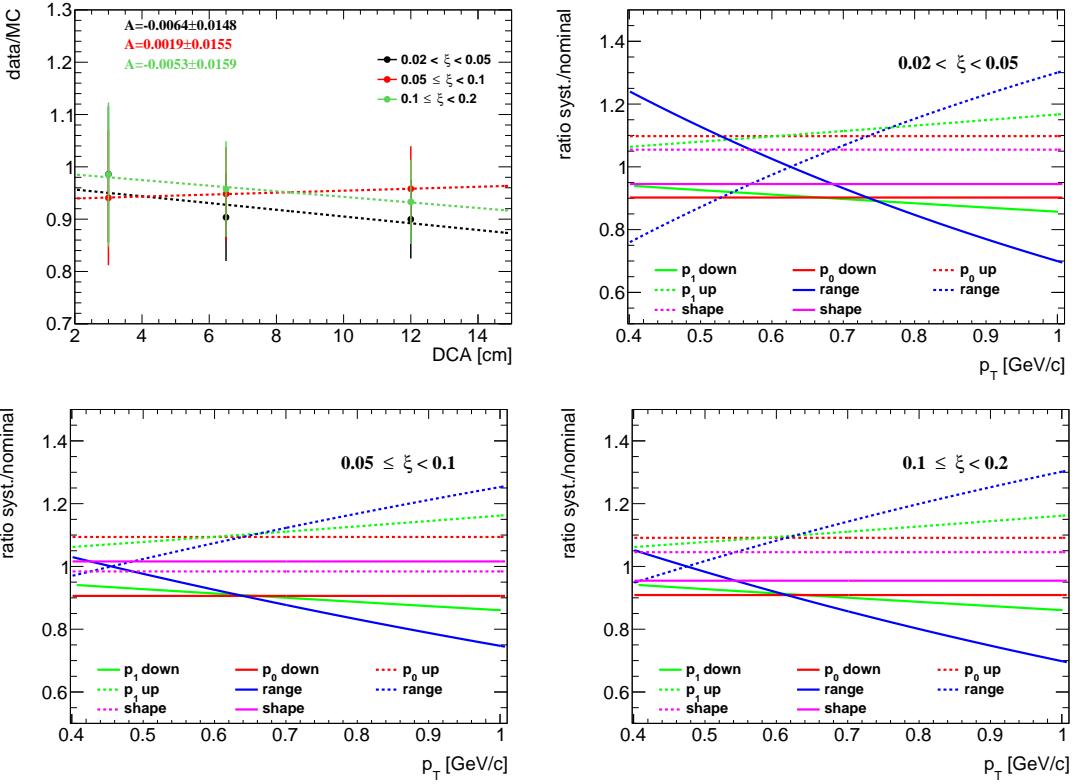


Figure 4.7: (top left) Data to MC ratio of the number of events in the background dominated region in three ranges of ξ with fitted functional form given by Eq. (4.6). (top right and bottom) Components of the systematic uncertainty related to the knock-out background protons contribution in three ξ ranges.

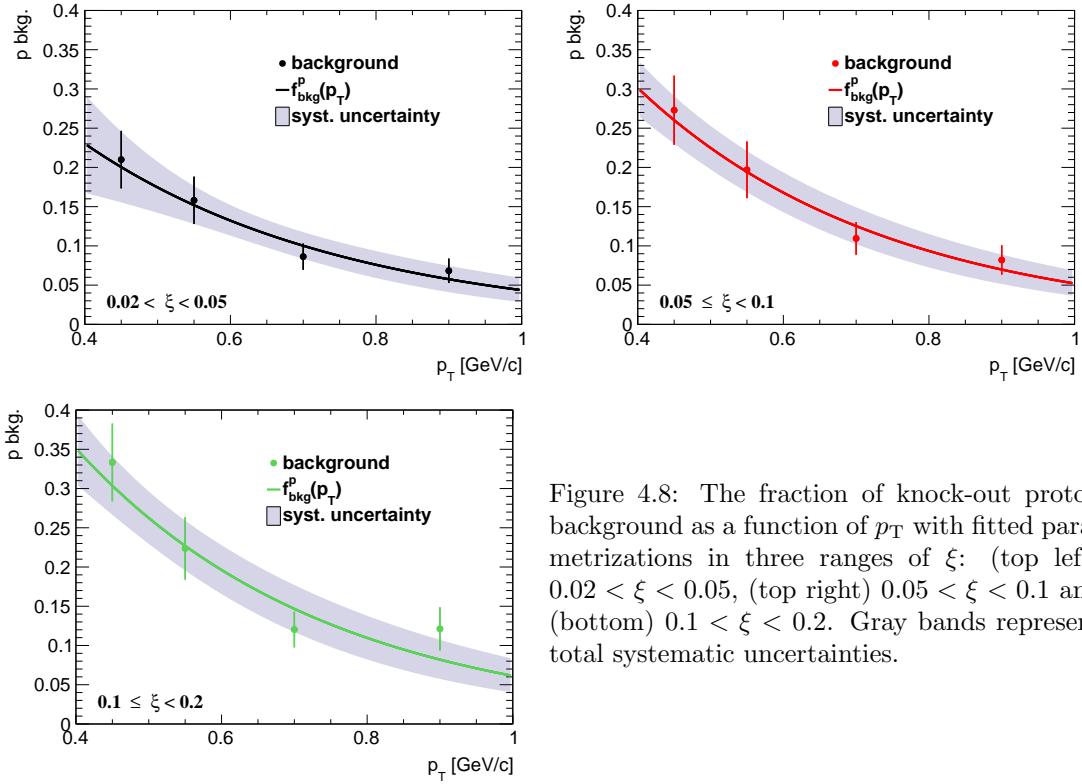


Figure 4.8: The fraction of knock-out proton background as a function of p_T with fitted parametrizations in three ranges of ξ : (top left) $0.02 < \xi < 0.05$, (top right) $0.05 \leq \xi < 0.1$ and (bottom) $0.1 \leq \xi < 0.2$. Gray bands represent total systematic uncertainties.

429 **Pion Background**

430 The pion spectra are corrected for weak decays (mainly of K_S^0 and Λ^0), muon contribution and
 431 background from the detector dead-material interactions. The pion decay muons can be identified
 432 as pions due to the similar masses. These background contributions are obtained from PYTHIA 8
 433 SD. Figure 4.9 shows the background contribution to the pion spectra as a function of p_T in
 434 three ranges of ξ , separately for π^- and π^+ . Since there were negligible differences observed
 435 between these three ranges of ξ , the background contribution was averaged over ξ . The following
 436 parametrization was found to describe it:

$$f_{\text{bkg}}^\pi(p_T) = a_0 \exp(a_1 p_T) + a_2 p_T^2 + a_3 p_T \quad (4.7)$$

437 where a_i , $i = 0, \dots, 3$ are free parameters of the fitted function.

438 The pion background contribution varies between 5% at low- p_T ($p_T = 0.25$ GeV/c) and about
 439 1% at $p_T = 1.0$ GeV/c for both negatively and positively charged pions. In addition, the back-
 440 ground was calculated from EPOS SD+SD' for the full range of ξ . The differences between
 441 PYTHIA 8 and EPOS are up to 1% for π^- .

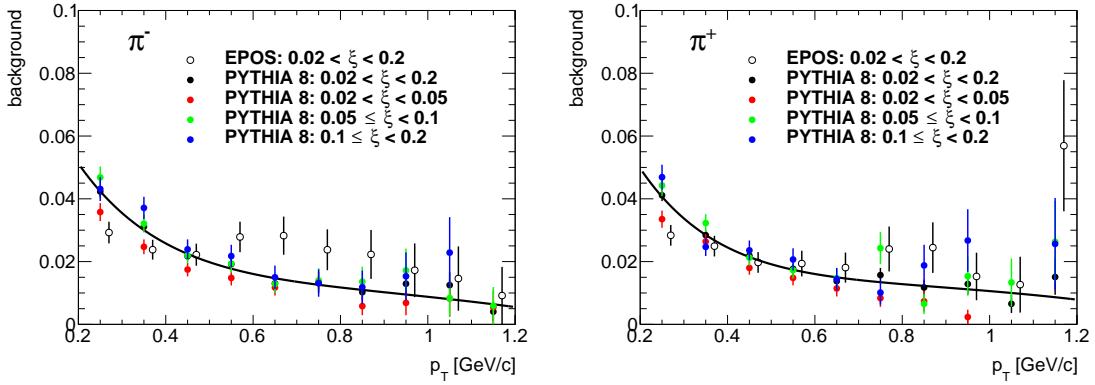


Figure 4.9: Pion background fraction as a function of p_T shown separately for (left) negatively and (right) positively charged pions in three ranges of ξ : (red) $0.02 < \xi < 0.05$, (green) $0.05 < \xi < 0.1$, (blue) $0.1 < \xi < 0.2$. (full black points) The pion background averaged over three ranges of ξ with fitted parametrization is also shown. Open black points represent EPOS predictions for the full ξ range.

4.2 Control Plots

Events, in which forward-scattered proton and reconstructed TOF vertex are the result of the same pp interaction, may originate from ND, DD, SD, and CD processes. It is preferred to estimate the background contribution from data, using dedicated control regions. Since such regions were not found, the relative contributions from the above processes were estimated from MC models and are therefore model dependent. Tracks reconstructed in RPs may also be:

- forward-scattered protons produced in the SD, CD or DD diffractive systems or from ND events,
- secondary particles from showering initiated by interaction of forward-scattered protons with beam-line elements. This contribution is negligible.

Figure 4.10 shows the uncorrected ξ and t distributions in data compared to various MC models: PYTHIA 8 A2 (MBR), PYTHIA 8 A2 (MBR-tuned), PYTHIA 8 4C (SaS) and EPOS. The MC distributions are split into SD, ND, DD and CD components. For EPOS, SD' is separated from the ND events. Additionally, the accidental background is also shown. PYTHIA 8 A2 (MBR) predictions, shown in Fig. 4.10 (a-b), do not agree with the data, especially there is small number of events in the region of large values of ξ . This effect may be due to the scaling factors, which are introduced in PYTHIA 8 to artificially suppress diffractive cross sections in the high mass region, or due to too large Pomeron intercept ($\epsilon = 0.104$). Therefore, additional two samples of PYTHIA 8 were generated: without this artificial suppression (MBR-tuned) and with $\epsilon = 0$ (SaS). Their predictions, shown in Fig. 4.10 (c-f), agree much better with the data than PYTHIA 8 A2 (MBR) and result also in a suppression of non-SD events. Amongst PYTHIA 8 models, PYTHIA 8 A2 (MBR-tuned) shows the best agreement with the data. EPOS predictions, shown in Fig. 4.10 (g-h), describes data better than PYTHIA 8 but shows a dominant contribution of SD' events. The CD contribution in EPOS is several times greater than in PYTHIA 8 (MBR), but it was never tuned to describe any data, as opposed to PYTHIA 8 (MBR) in which the CD cross sections are constrained by CDF measurements [13]. The CD component in the SaS model is based on simple scaling assumption, therefore, it is not usually used by the experimental communities. All MCs predict significant DD and ND background at large ξ , thereby the analysis was limited to $\xi < 0.2$.

Figures 4.11 to 4.13 show the uncorrected distributions of variables used in the later analysis: n_{sel} , p_T and $\bar{\eta}$. The contributions from non-SD (except EPOS SD') interactions differ a bit between each other, i.e. EPOS predicts significantly larger CD contribution, whereas DD and ND are suppressed in PYTHIA 8 A2 (MBR-tuned) and PYTHIA 8 4C (SaS). PYTHIA 8 A2 (MBR) is used as the default model of non-SD contribution subtracted from the data with systematic uncertainty $\pm 50\%$, which covers all differences between the models except EPOS SD'. In this analysis EPOS SD' is considered as an alternative to PYTHIA 8 SD model of events with forward-scattered proton in the final state, where one of the proton remnants hadronizes back to a single proton (non-diffractive process), while in PYTHIA 8 the initial proton stays intact (diffractive process).

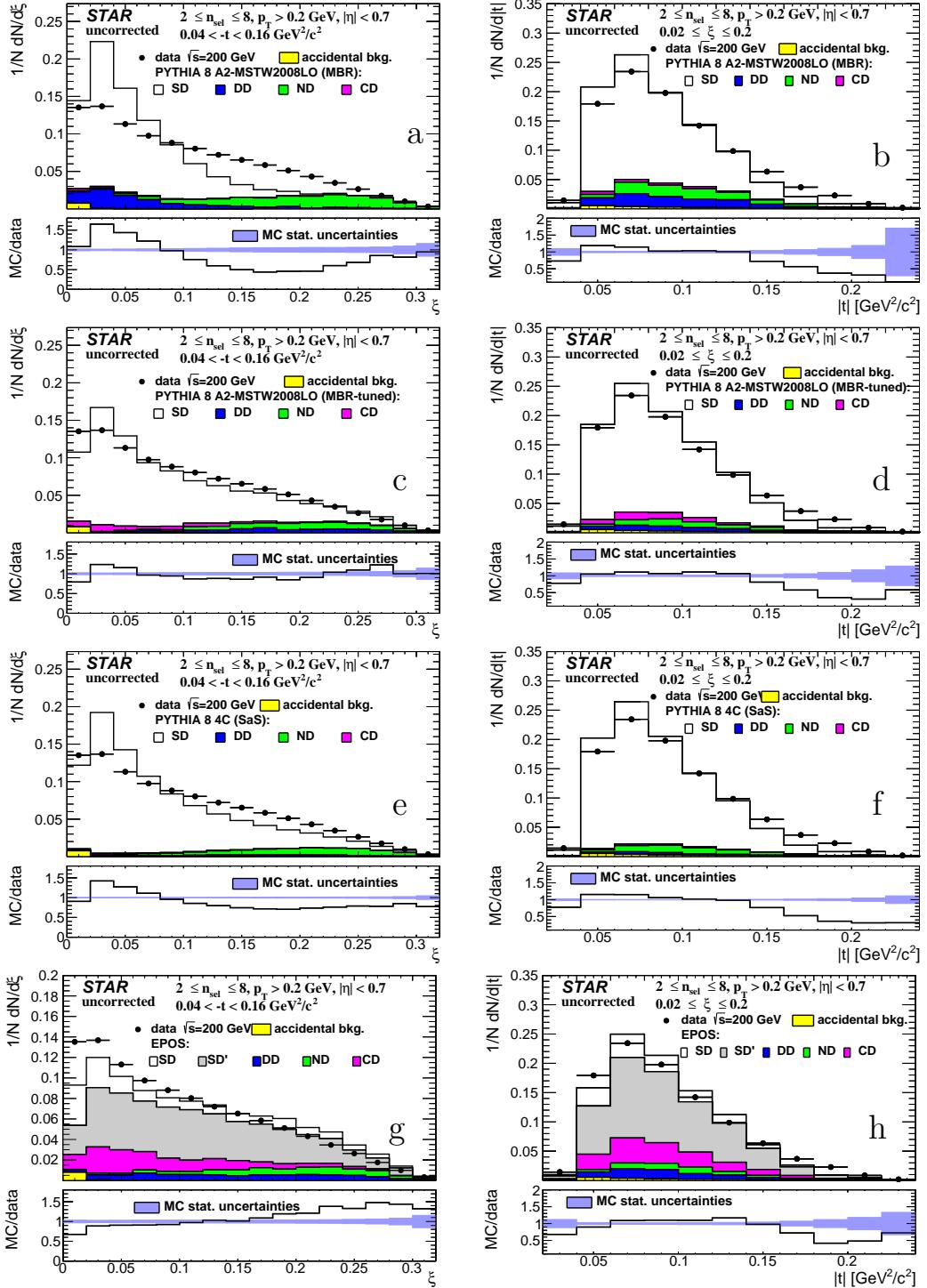


Figure 4.10: Uncorrected distributions of data compared to various MC models: (a-b) PYTHIA 8 A2 (MBR), (c-d) PYTHIA 8 A2 (MBR-tuned), (e-f) PYTHIA 8 4C (SaS) and (g-h) EPOS, as a function of (left column) ξ and (right column) $|t|$. The ratio of MC predictions and data is shown in the bottom panels.

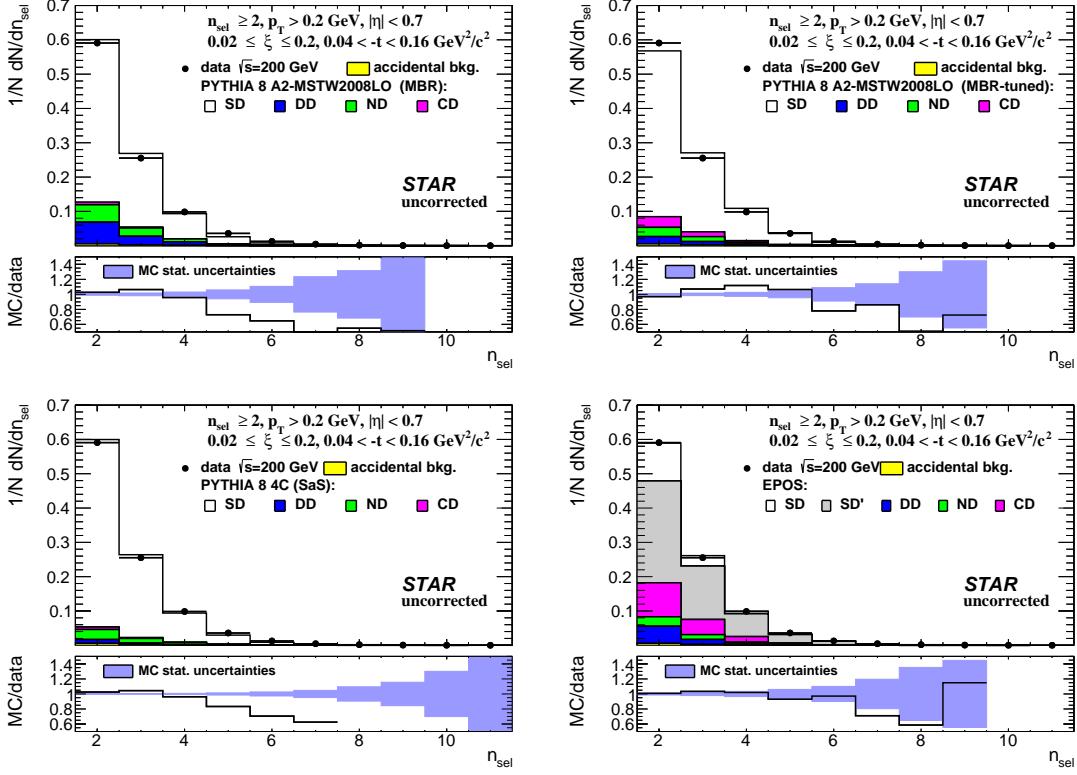


Figure 4.11: Uncorrected distributions of data compared to various MC models: (top left) PYTHIA 8 A2 (MBR), (top right) PYTHIA 8 A2 (MBR-tuned), (bottom left) PYTHIA 8 4C (SaS) and (bottom right) EPOS, as a function of n_{sel} . The ratio of MC predictions and data is shown in the bottom panels.

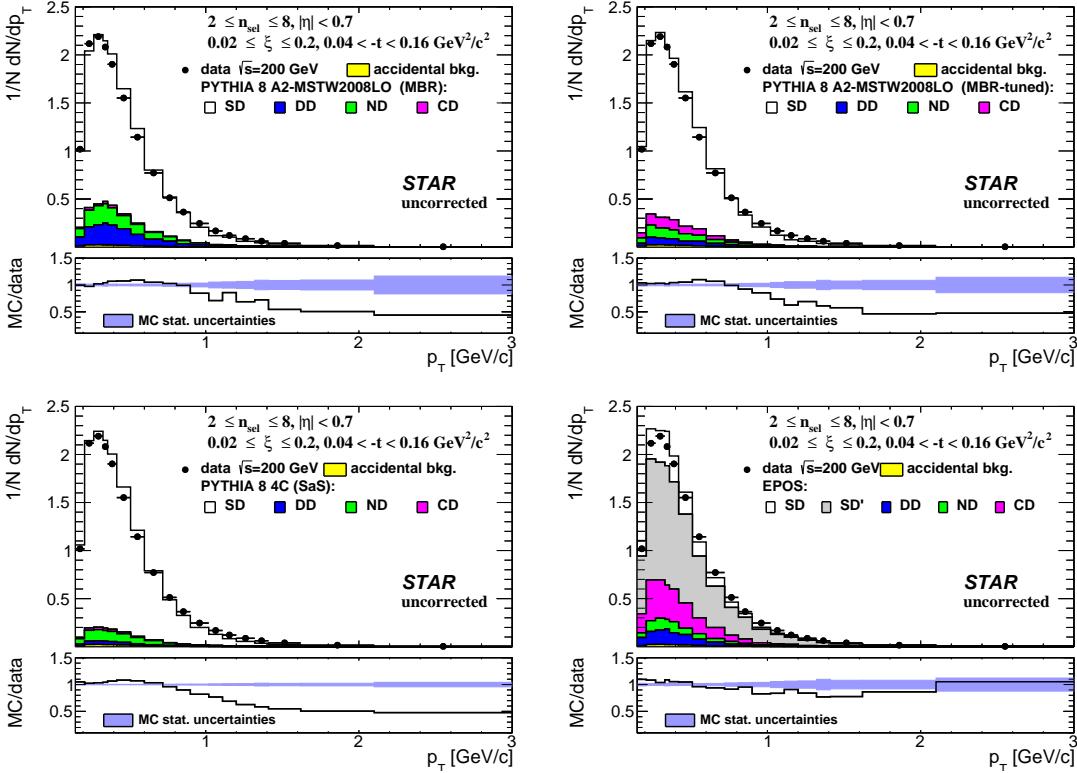


Figure 4.12: Uncorrected distributions of data compared to various MC models: (top left) PYTHIA 8 A2 (MBR), (top right) PYTHIA 8 A2 (MBR-tuned), (bottom left) PYTHIA 8 4C (SaS) and (bottom right) EPOS, as a function of p_T . The ratio of MC predictions and data is shown in the bottom panels.

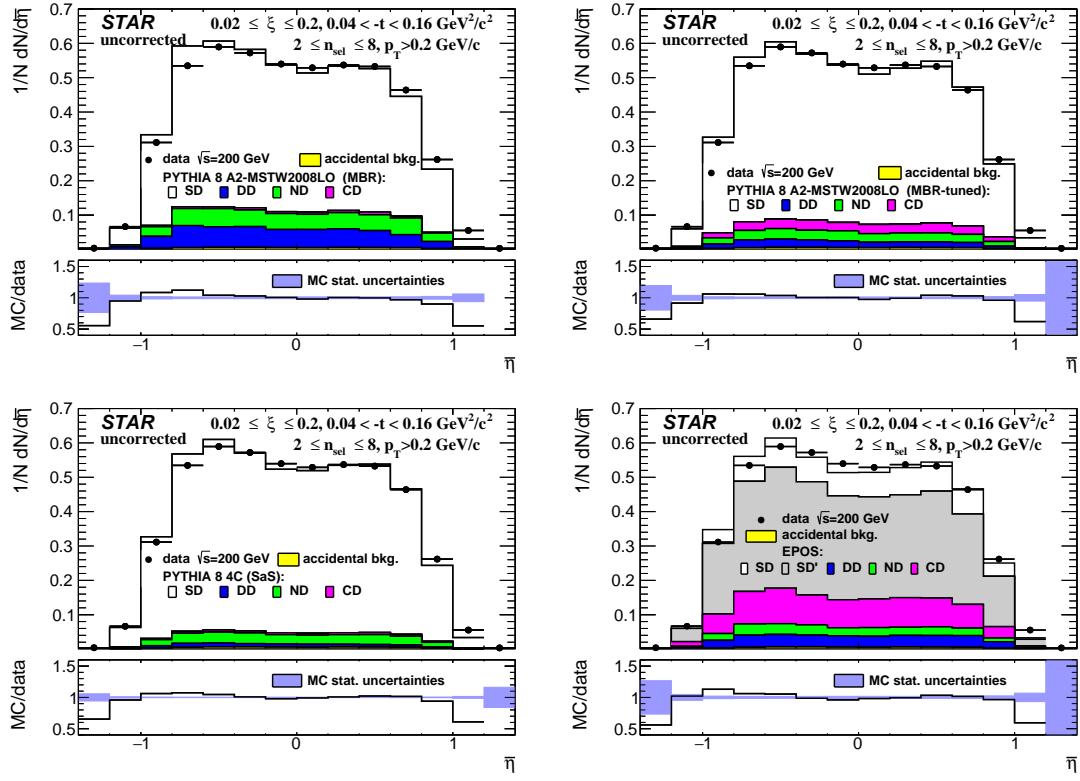


Figure 4.13: Uncorrected distributions of data compared to various MC models: (top left) PYTHIA 8 A2 (MBR), (top right) PYTHIA 8 A2 (MBR-tuned), (bottom left) PYTHIA 8 4C (SaS) and (bottom right) EPOS, as a function of $\bar{\eta}$. The ratio of MC predictions and data is shown in the bottom panels.

5. Selection Efficiencies

5.1 Vertex Reconstruction

When the charged-particle multiplicity is low, the vertex-finding algorithm sometimes fails to find the primary vertex. In addition, at high luminosity, vertex finder can fail due to the contribution of pile-up interactions, providing a wrong reconstructed vertex. In the study of vertex reconstruction efficiency we required at least two reconstructed global tracks $n_{\text{sel}}^{\text{global}} \geq 2$ passing all the quality cuts listed in Sec 3.1, except vertex-related cuts on DCA_{xy} and DCA_z , and associated to true-level primary particles. Additionally, MC events were accepted if the z -coordinate of the true-level primary vertex was between -80 and 80 cm and $n_{\text{ch}} \geq 2$. All corrections, described in this section, were calculated in three ranges of ξ separately using PYTHIA 8 SD embedding MC.

The global tracks (not necessarily associated to a true-level primary particles), which are used by the vertex-finder algorithm, had to pass the following quality cuts:

1. tracks must be matched with hits reconstructed in TOF,
2. the number of the TPC hits used in the helix fit $N_{\text{hits}}^{\text{fit}}$ must be greater than 20,
3. the ratio of the number of TPC hits used in the helix fit to the number of possible TPC hits $N_{\text{hits}}^{\text{fit}}/N_{\text{hits}}^{\text{poss}}$ must be greater than 0.52,
4. the transverse impact parameter with respect to the beamline d_0 must be less than 2 cm,
5. the track's transverse momentum p_T must be greater than 0.2 GeV/c.

The above track selection criteria are different than those used in the nominal analysis. Primary vertex reconstruction efficiency and fake vertex rate were calculated as a function of the number of global tracks used in vertexing $n_{\text{virt}}^{\text{global}}$ instead of $n_{\text{sel}}^{\text{global}}$ ($n_{\text{virt}}^{\text{global}} \geq n_{\text{sel}}^{\text{global}}$).

In the nominal analysis exactly one vertex with $n_{\text{sel}} \geq 2$ is required. However, in the study of vertex reconstruction, events with additional vertices were studied. Therefore, we define the best vertex as the reconstructed vertex with the highest number of TOF-matched tracks. This vertex does not have to be associated to true-level primary vertex (fake or secondary vertex). The algorithm, which matches reconstructed vertices to true-level vertices, checks for reconstructed tracks originating from them. If at least one reconstructed track is assigned to a true-level particle, then the reconstructed vertex is assigned to the true-level vertex from which the true-level particle originates. Since the fake vertices (not matched to the true-level primary vertex) are allowed in the analysis, the overall vertex-finding efficiency, $\epsilon_{\text{virt}}(n_{\text{virt}}^{\text{global}})$, is expressed as:

$$\epsilon_{\text{virt}}(n_{\text{virt}}^{\text{global}}) = \epsilon_{\text{virt}}^{\text{best}}(n_{\text{virt}}^{\text{global}}) + \delta_{\text{virt}}^{\text{fake}}(n_{\text{virt}}^{\text{global}}) \quad (5.1)$$

where:

$\epsilon_{\text{virt}}^{\text{best}}(n_{\text{virt}}^{\text{global}})$ is the primary vertex reconstruction efficiency, determined as the ratio of the number of good reconstructed events (best primary vertex with $n_{\text{sel}} \geq 2$ matched to the true-level primary vertex) to the number of input MC events,

$\delta_{\text{virt}}^{\text{fake}}(n_{\text{virt}}^{\text{global}})$ is the fake vertex rate, determined as the ratio of the number of good reconstructed events (best primary vertex with $n_{\text{sel}} \geq 2$ not matched to the true-level primary vertex) to the number of input MC events. Due to the contribution of pile-up, it is possible that the best vertex originates from fake tracks instead of true-level particles.

518 The vertex-finding efficiency as a function of $n_{\text{virt}}^{\text{global}}$, shown in Fig. 5.1 (left), is larger than 75% for
 519 all $n_{\text{virt}}^{\text{global}}$. However, for $n_{\text{virt}}^{\text{global}} > 8$, there are more fake than true-level primary vertices. When
 520 there are exactly two global tracks used in the vertex reconstruction, $n_{\text{virt}}^{\text{global}} = 2$, the vertex-
 521 finding efficiency depends on the longitudinal distance between these tracks $|\Delta z_0|$. Therefore,
 522 the vertex-finding efficiency for such events $\epsilon_{\text{virt}}(|\Delta z_0|)$ is given by:

$$\epsilon_{\text{virt}}(|\Delta z_0|) = \epsilon_{\text{virt}}^{\text{best}}(|\Delta z_0|) + \delta_{\text{virt}}^{\text{fake}}(|\Delta z_0|) \quad (5.2)$$

523 where: $\epsilon_{\text{virt}}^{\text{best}}(|\Delta z_0|)$ is the primary vertex reconstruction efficiency, $\delta_{\text{virt}}^{\text{fake}}(|\Delta z_0|)$ is the fake vertex
 524 rate.

525 Figure 5.1 (right) shows the vertex-finding efficiency for events with $n_{\text{virt}}^{\text{global}} = 2$. This efficiency
 526 is smaller than 20% for tracks with $|\Delta z_0| > 2$ cm, hence the analysis was limited to events with
 527 $|\Delta z_0| < 2$ cm, when $n_{\text{virt}}^{\text{global}} = 2$. The rate of fake vertices is negligibly low (open points overlap
 528 with full points).

529 Events are rejected if more vertices are reconstructed in addition to the best one. Rejected
 530 events can be classified as:

- 531 a) two or more additional vertices,
- 532 b) additional secondary vertex from interactions with the detector dead-material,
- 533 c) additional fake vertex,
- 534 d) additional primary vertex (vertex splitting or background vertex reconstructed as best ver-
 535 tex),
- 536 e) additional secondary vertex from the decay.

537 The fraction of such events, $f_{\text{virt}}^{\text{veto}}(n_{\text{virt}}^{\text{global}})$, is given by:

$$f_{\text{virt}}^{\text{veto}}(n_{\text{virt}}^{\text{global}}) = \frac{\text{number of events with more than one reconstructed TOF vertex}}{\text{number of events with at least one reconstructed TOF vertex}} \quad (5.3)$$

$$= f_a + f_b + f_c + f_d + f_e$$

538 where f_a to f_e are the fractions of events with additional vertices, with labels corresponding to
 539 the items in the listing above.

540 As before, the fraction was calculated as a function of $|\Delta z_0|$ for events with $n_{\text{virt}}^{\text{global}} = 2$.
 541 Figure 5.2 shows the fraction of multi-vertex events with respect to the $n_{\text{virt}}^{\text{global}}$. There is a large
 542 fraction of events ($> 90\%$) with additional background vertices for $n_{\text{virt}}^{\text{global}} \geq 9$, what would result
 543 in large correction factor. Hence, the analysis was limited to events with $n_{\text{sel}}^{\text{global}} \leq 8$ ($n_{\text{sel}}^{\text{global}} \leq$

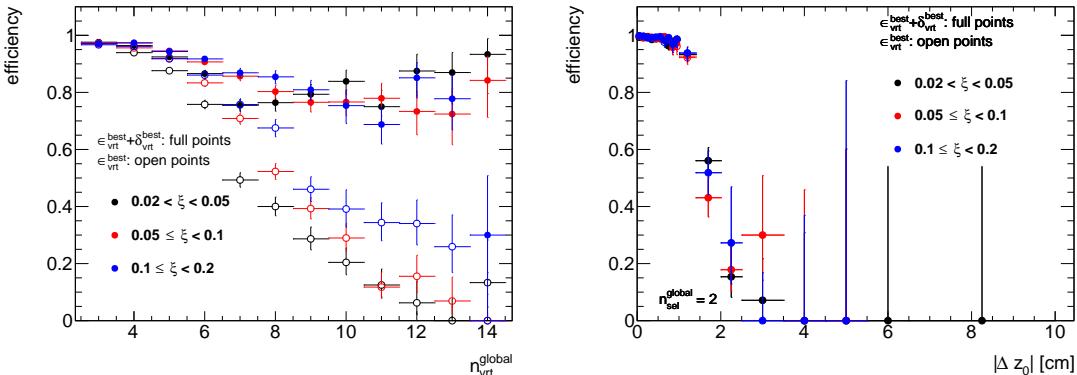


Figure 5.1: Vertex-finding efficiency in three ranges of ξ as a function of (left) $n_{\text{virt}}^{\text{global}}$ and (right)
 with respect to the $|\Delta z_0|$ between reconstructed tracks in events with $n_{\text{virt}}^{\text{global}} = 2$.

544 $n_{\text{vrt}}^{\text{global}}$). The total fraction of multi-vertex events, $f_a + f_b + f_c + f_d + f_e$, as a function of $n_{\text{vrt}}^{\text{global}}$
 545 and $|\Delta z_0|$, shown in Fig. 5.3, demonstrates that $f_{\text{vrt}}^{\text{veto}}(|\Delta z_0|)$ is very small (< 2%) for events with
 546 $n_{\text{vrt}}^{\text{global}} = 2$.

547 Although, the analysis was limited to $n_{\text{sel}}^{\text{global}} \leq 8$ ($n_{\text{sel}}^{\text{global}} \leq n_{\text{vrt}}^{\text{global}}$), a fraction of events with
 548 additional background vertices was still relatively large. Since most of these additional vertices
 549 are fake (and as accidental not correlated with true-level primary distributions), it was checked
 550 whether the charged-particle multiplicity distributions are different for events with and without
 551 reconstructed fake vertices. These distributions, as shown in Fig 5.4, are in good agreement, thus,
 552 above studies of vertex reconstruction were repeated using MC events that do not contain recon-
 553 structed fake vertices. It means that events with additional fake vertex were rejected (similarly to
 554 the analysis of real data) and no correction is needed for such losses since it only affects overall
 555 normalization (not the shapes of distributions under study). The vertex-finding efficiency, which
 556 was calculated from such events, is shown in Fig. 5.5. It is greater than 95% for events with
 557 $2 \leq n_{\text{vrt}}^{\text{global}} \leq 8$. In addition, the corresponding fraction of multi-vertex events, shown in Figs. 5.6

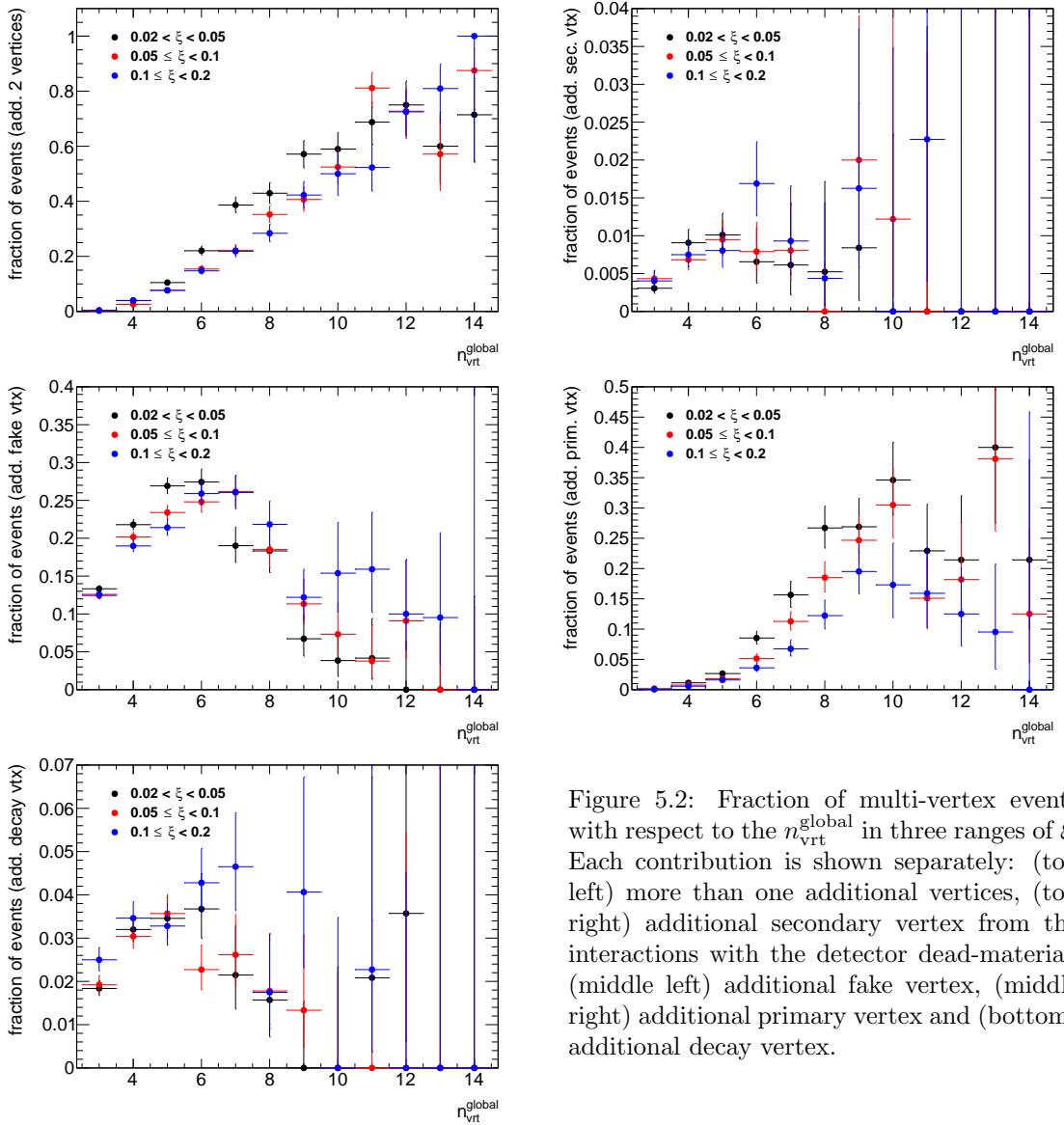


Figure 5.2: Fraction of multi-vertex events with respect to the $n_{\text{vrt}}^{\text{global}}$ in three ranges of ξ . Each contribution is shown separately: (top left) more than one additional vertices, (top right) additional secondary vertex from the interactions with the detector dead-material, (middle left) additional fake vertex, (middle right) additional primary vertex and (bottom) additional decay vertex.

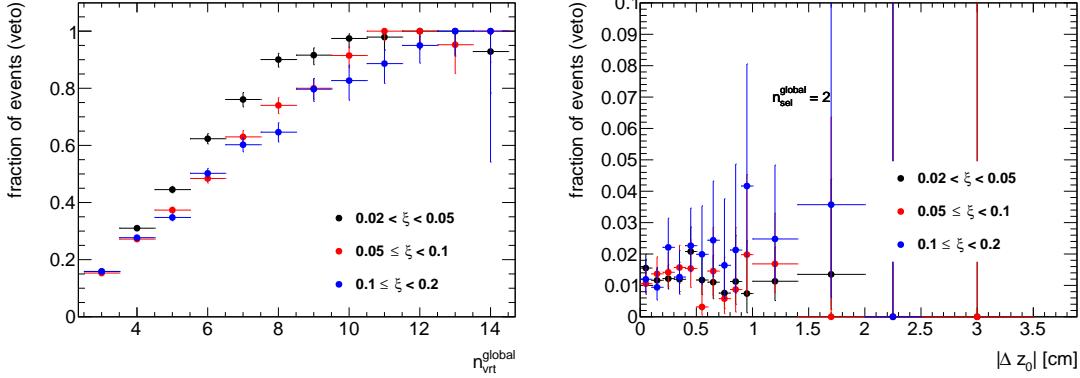


Figure 5.3: Total fraction of multi-vertex events as a function of (left) $n_{\text{vrt}}^{\text{global}}$ for events with $n_{\text{vrt}}^{\text{global}} > 2$ and (right) $|\Delta z_0|$ for events with $n_{\text{vrt}}^{\text{global}} = 2$ in three ranges of ξ .

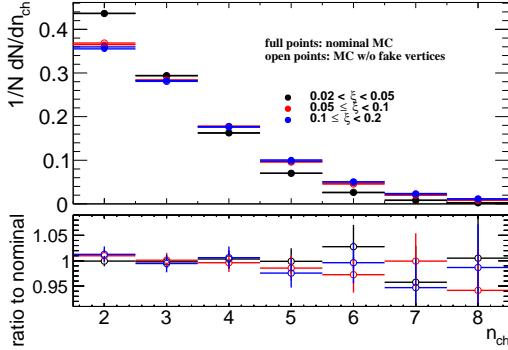


Figure 5.4: Normalized charged-particle multiplicity distributions in three ranges of ξ calculated from PYTHIA 8 SD embedding MC for (full points) all generated events and (open points) events without reconstructed fake vertices.

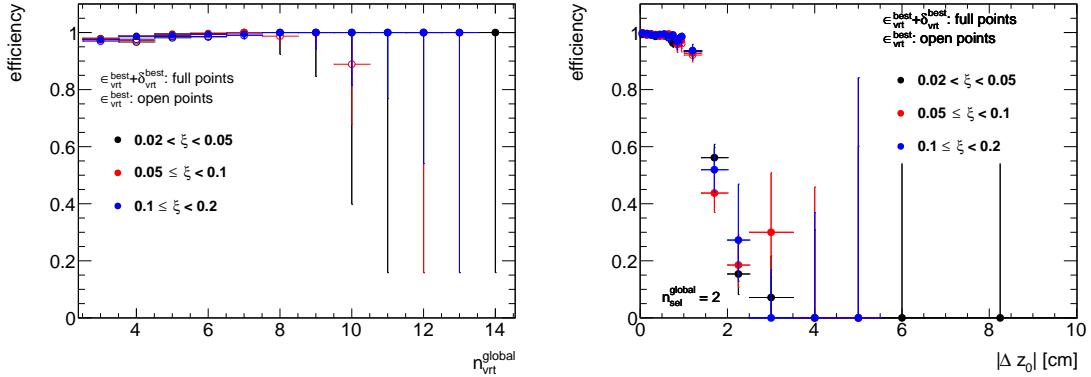


Figure 5.5: Vertex-finding efficiency in three ranges of ξ as a function of (left) $n_{\text{vrt}}^{\text{global}}$ and (right) with respect to the $|\Delta z_0|$ between reconstructed tracks in events with $n_{\text{vrt}}^{\text{global}} = 2$. Only events that do not contain additional fake vertices were used.

and 5.7, is smaller than 20%. Since fake vertices were rejected from this study, the f_c term from Eq. (5.3) is equal to 0. The correction factors calculated from MC events that do not contain reconstructed fake vertices were used in the analysis instead of the one obtained from the full MC sample.

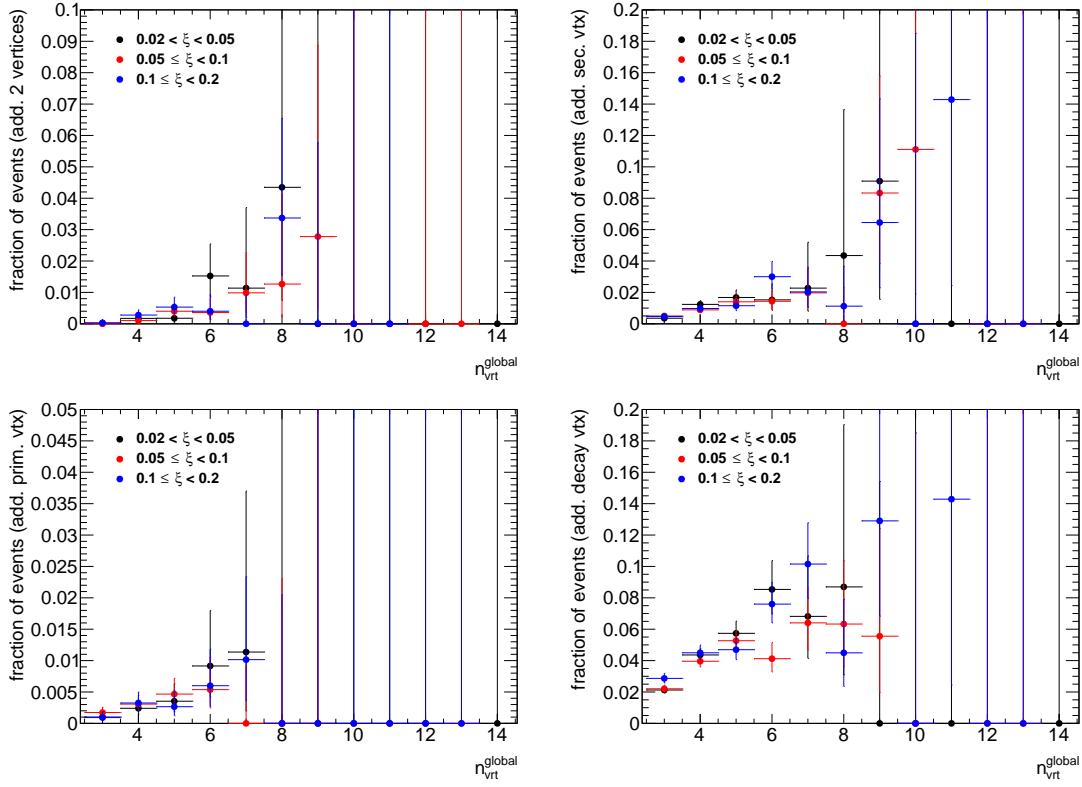


Figure 5.6: Fraction of multi-vertex events with respect to the $n_{\text{vrt}}^{\text{global}}$ in three ranges of ξ . Each contribution is shown separately: (top left) more than one additional vertices, (top right) additional secondary vertex from the interactions with the detector dead-material, (bottom left) additional primary vertex and (bottom right) additional decay vertex. Only events that do not contain additional fake vertices were used.

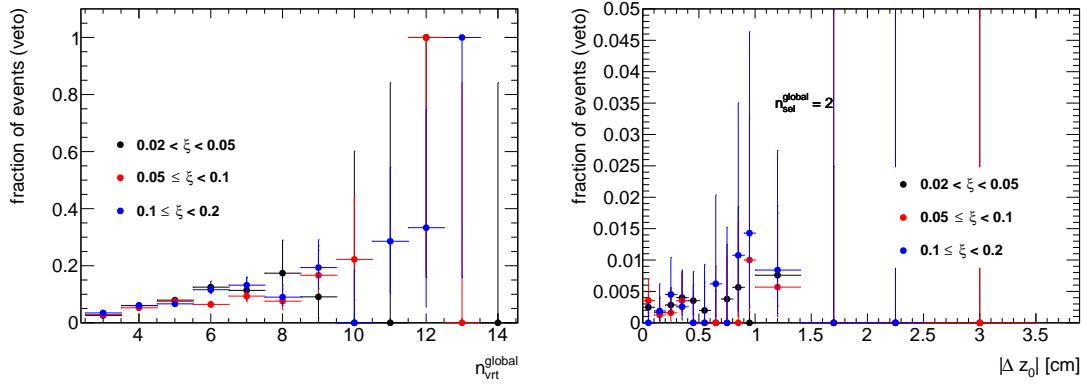


Figure 5.7: Total fraction of multi-vertex events as a function of (left) $n_{\text{vrt}}^{\text{global}}$ for events with $n_{\text{vrt}}^{\text{global}} > 2$ and (right) $|\Delta z_0|$ for events with $n_{\text{vrt}}^{\text{global}} = 2$ in three ranges of ξ . Only events that do not contain additional fake vertices were used.

5.2 Correction to BBC-Small

The SDT trigger conditions imposed a signal in RPs and a veto on any signal in the same-side small BBC tiles, whereas a signal in the opposite-side BBC-small was required by the offline event selection. These requirements were imposed in order to accept only events with rapidity gap and reduce DD, ND and accidental backgrounds. A joined BBC-small efficiency, ϵ_{BBC} , was obtained as a function of each measured quantity using PYTHIA 8 4C (SaS) SD embedded into Zerobias data, EPOS SD+SD' and HERWIG SD MC. The efficiency was calculated for events within fiducial region as follows:

$$\epsilon_{BBC} = \frac{\text{number of MC events satisfying the BBC-small selection criteria}}{\text{number of MC events}} \quad (5.4)$$

Figures 5.8 to 5.10 show the fraction of generated true-level MC events, within the fiducial region of the measurement, in which the selection criteria on BBC-small signal and veto are fulfilled. The efficiency weakly depends on the measured variables (n_{ch} , p_T and $\bar{\eta}$). In addition, veto, signal and joined BBC-small efficiencies are presented separately as a function of ξ in Fig. 5.11. The ϵ_{BBC} strongly depends on ξ and varies from about 90% for events with ξ within 0.02 – 0.05 to about 60% for events with $0.1 < \xi < 0.2$. However, measurements of corrected ξ distributions are out of the scope of this analysis.

Data is corrected for BBC-small efficiency using PYTHIA 8 4C (SaS). The uncertainty related to this correction is estimated by using HERWIG and EPOS SD+SD' samples, where the hadronization models are different from that used in PYTHIA 8. Figure 5.12 shows the PYTHIA 8 prediction on BBC efficiency divided by the HERWIG prediction in three ranges of ξ . The deviations between these two models are of the order of 4% at $0.02 < \xi < 0.05$, 2% at $0.05 < \xi < 0.1$ and about 10% at $0.1 < \xi < 0.2$. The differences between PYTHIA 8 and EPOS SD+SD' predictions are shown in Fig. 5.13. Most of them are of the order of 3%, except $n_{ch} \leq 3$ for which the difference varies up to 6%. The maximum difference between PYTHIA 8 and HERWIG/EPOS hadronization models is used as systematic uncertainty.

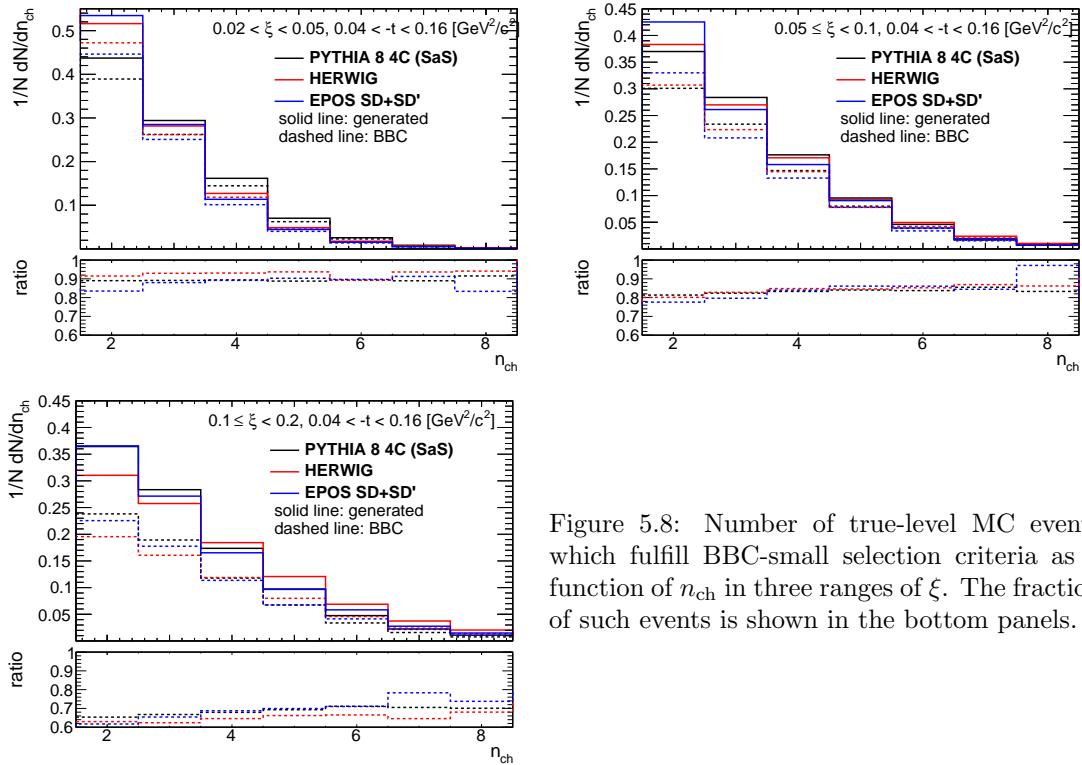


Figure 5.8: Number of true-level MC events which fulfill BBC-small selection criteria as a function of n_{ch} in three ranges of ξ . The fraction of such events is shown in the bottom panels.

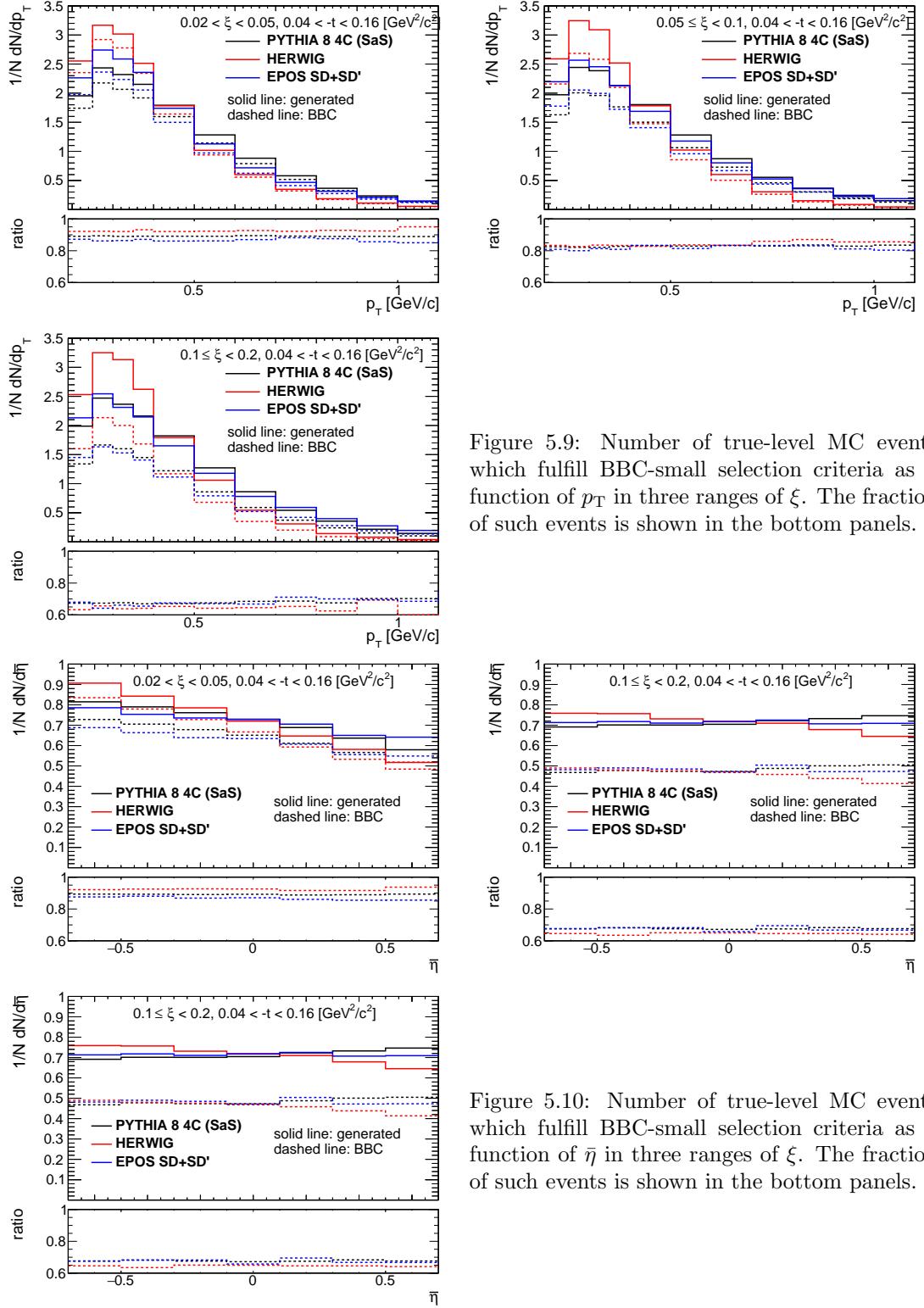


Figure 5.9: Number of true-level MC events which fulfill BBC-small selection criteria as a function of p_T in three ranges of ξ . The fraction of such events is shown in the bottom panels.

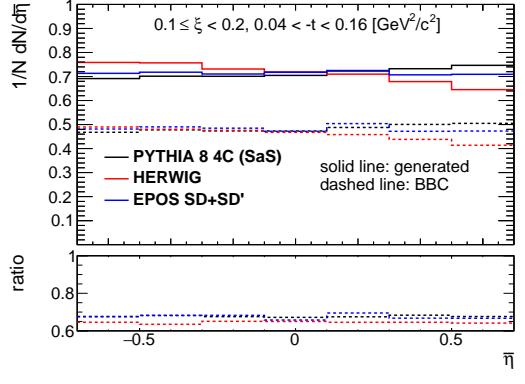
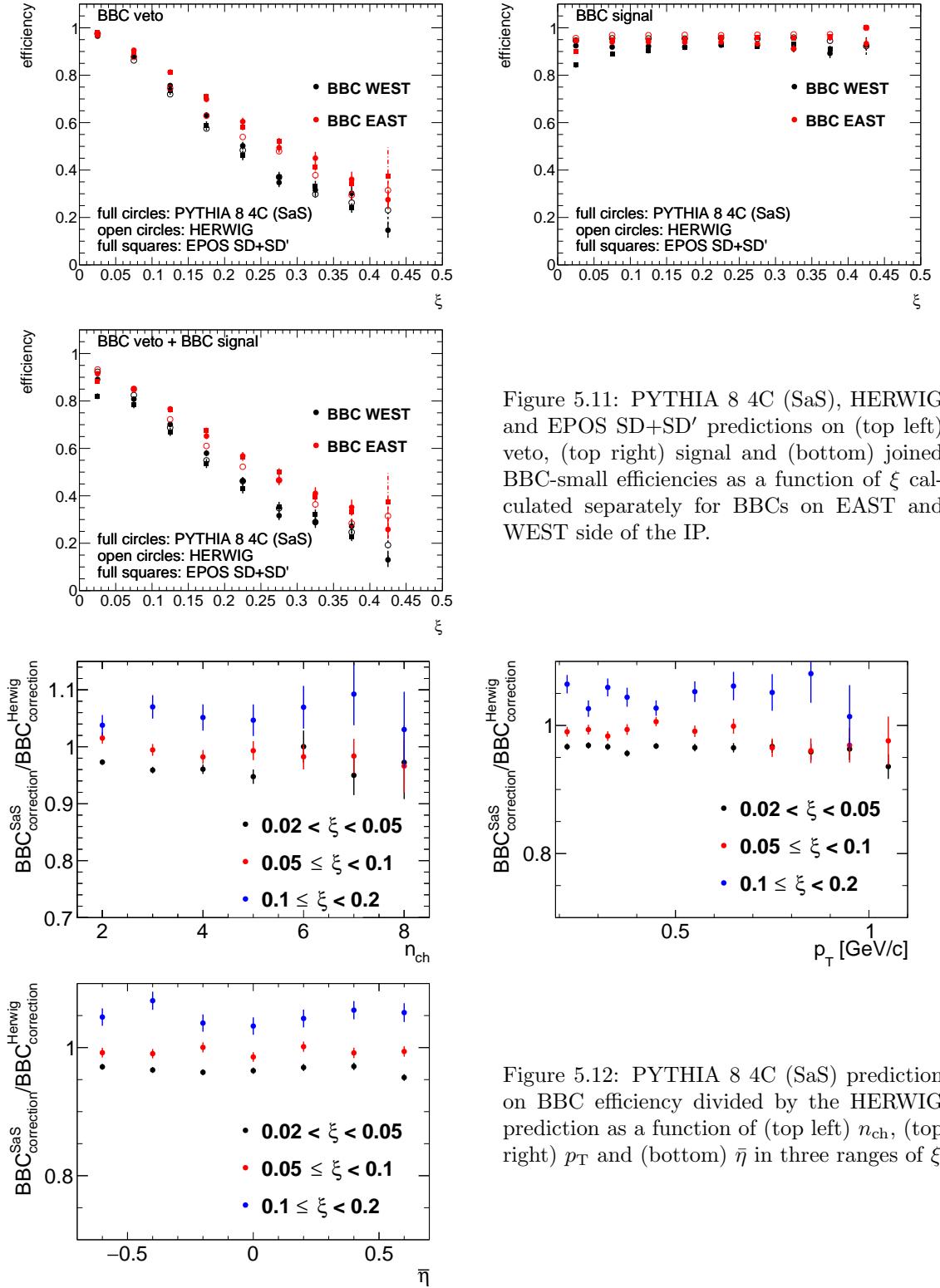


Figure 5.10: Number of true-level MC events which fulfill BBC-small selection criteria as a function of $\bar{\eta}$ in three ranges of ξ . The fraction of such events is shown in the bottom panels.



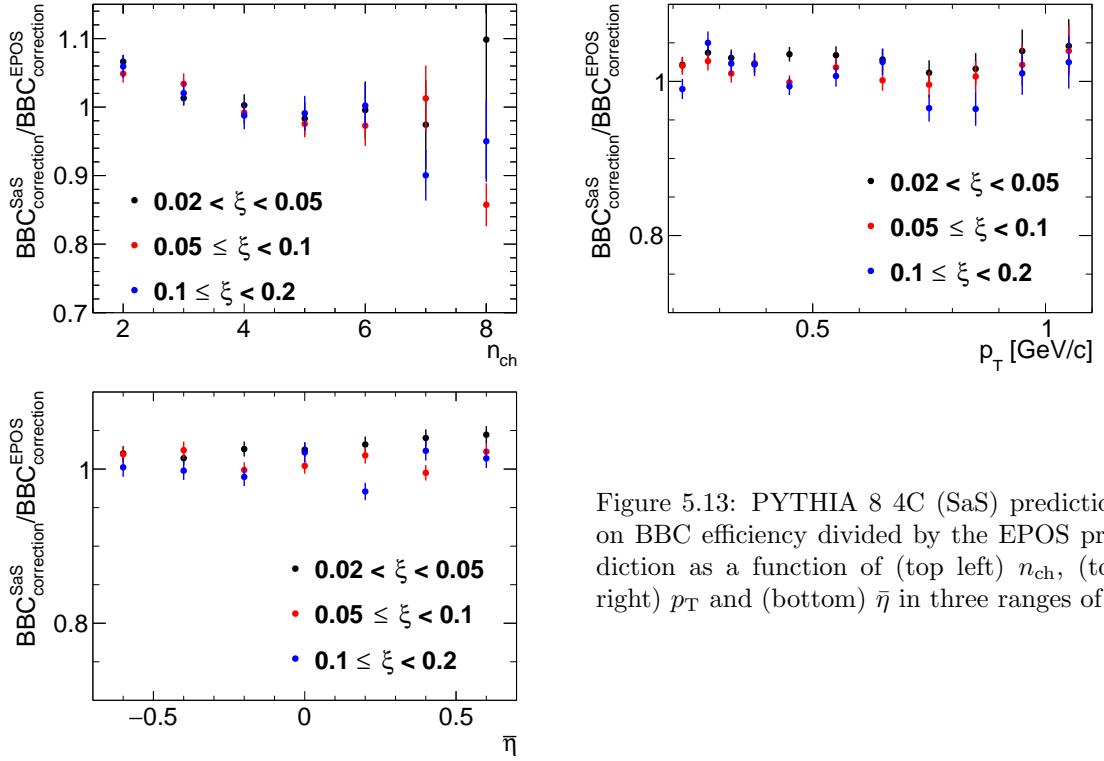


Figure 5.13: PYTHIA 8 4C (SaS) prediction on BBC efficiency divided by the EPOS prediction as a function of (top left) n_{ch} , (top right) p_{T} and (bottom) $\bar{\eta}$ in three ranges of ξ

6. Migrations into and out of the Fiducial Region

In this section the corrections due to the migrations of tracks and forward-scattered protons into and out of the fiducial region are described.

6.1 Migrations of Tracks into and out of the Fiducial Region

The procedure, described in this section, accounts for migrations of tracks into and out of the fiducial region, which originate from TPC resolution effects. The correction factor for such tracks, $f_{\text{okr}}(p_T, \eta)$ is defined as follows:

$$f_{\text{okr}}(p_T, \eta) = \frac{1 - f_{\text{okr}}^-(p_T, \eta)}{1 - f_{\text{okr}}^+(p_T, \eta)} \quad (6.1)$$

where $f_{\text{okr}}^-(p_T, \eta)$ is the fraction of reconstructed tracks for which the corresponding primary particle is outside of the kinematic range of the measurement and $f_{\text{okr}}^+(p_T, \eta)$ is the fraction of primary particles for which the corresponding reconstructed track is outside of the kinematic range of the measurement.

The resulting residual migrations, shown in Fig. 6.1, were estimated using PYTHIA 8 SD embedding MC. The main effect was observed at $|\eta| \sim 0.7$, where about 2 – 6% reconstructed tracks were associated to primary particle outside the fiducial region. However, above contributions to the correction factor, $f_{\text{okr}}(p_T, \eta)$, cancel each other and the resulting factor is about 2% at $|\eta| \sim 0.7$.

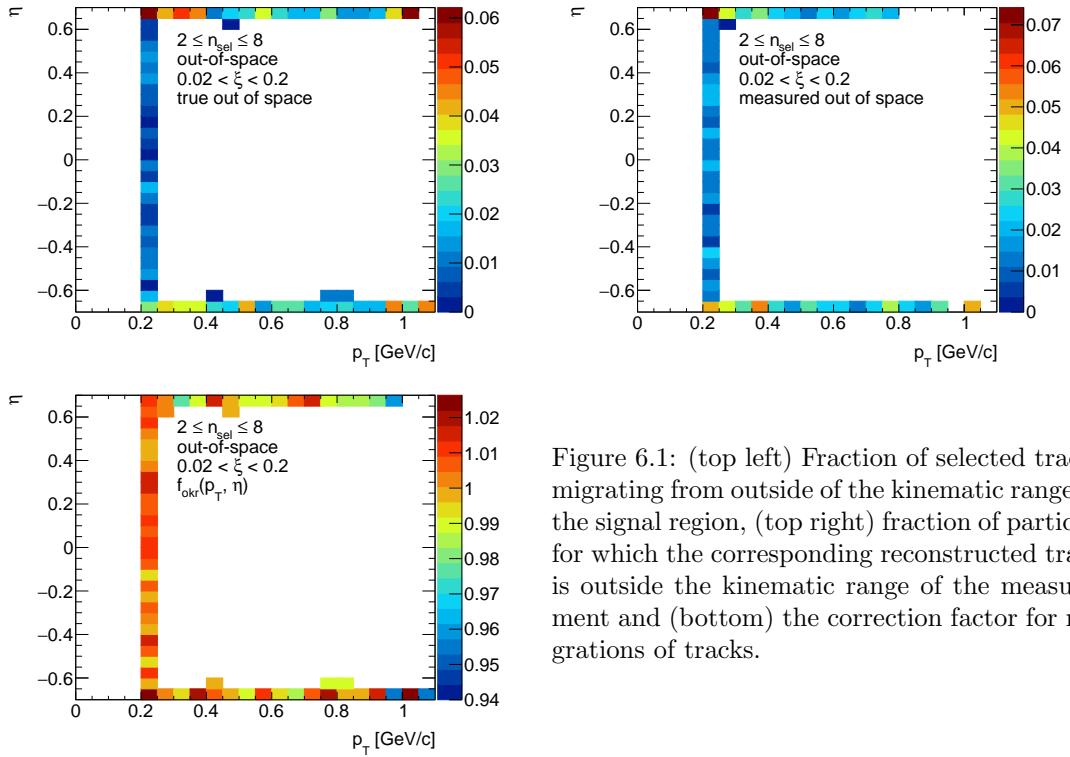


Figure 6.1: (top left) Fraction of selected tracks migrating from outside of the kinematic range to the signal region, (top right) fraction of particles for which the corresponding reconstructed track is outside the kinematic range of the measurement and (bottom) the correction factor for migrations of tracks.

6.2 Migrations in ξ

The analysis was performed in three ranges of ξ . Thus, there are migrations into and out of these ξ regions. They mainly originate from the resolution of ξ reconstructed from RP tracks. Figure 6.2 shows the resolution of ξ as a function of the true-level ξ (denoted as ξ_{true}) with fitted zeroth order polynomial. The resolution of ξ is fairly constant and equals to about 0.3%.

The corrections due to migrations into and out of ξ regions was defined as:

$$f_\xi = \frac{1 - f_\xi^-}{1 - f_\xi^+} \quad (6.2)$$

where f_ξ^- is the fraction of events for which the corresponding true-level, ξ_{true} , is outside of the ξ region and f_ξ^+ is the fraction of events for which the corresponding reconstructed, ξ_{reco} , is outside of the ξ region.

The f_ξ was calculated for each measured variable separately. Figures 6.3 to 6.5 show the fraction of events f_ξ^- and f_ξ^+ as a function of n_{ch} , p_{T} and $\bar{\eta}$. The lower panel in each figure shows the corresponding correction factor f_ξ . The largest differences between migrations into and out of the ξ regions were observed at $0.02 < \xi < 0.05$, where they are of the order of 2 – 4%. In the other ξ regions, the difference between f_ξ^- and f_ξ^+ is smaller than 1%.

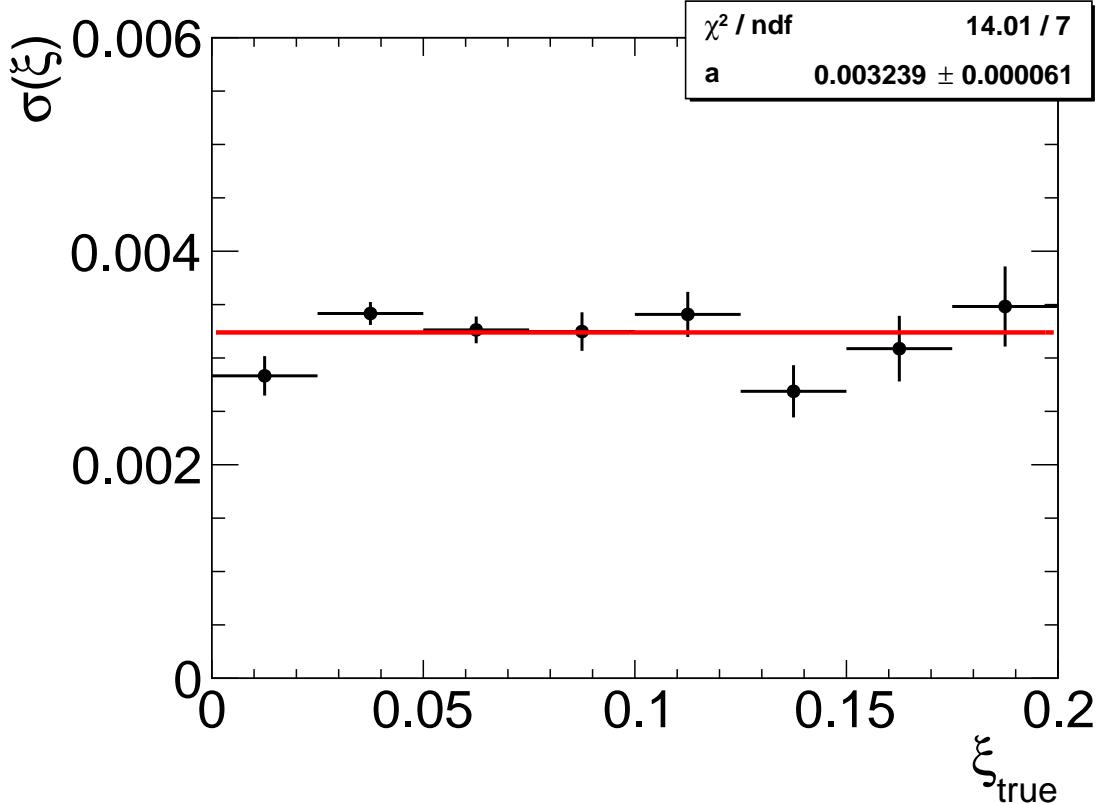


Figure 6.2: The resolution of ξ as a function of ξ_{true} . The zeroth order polynomial, shown as red line, was fitted.

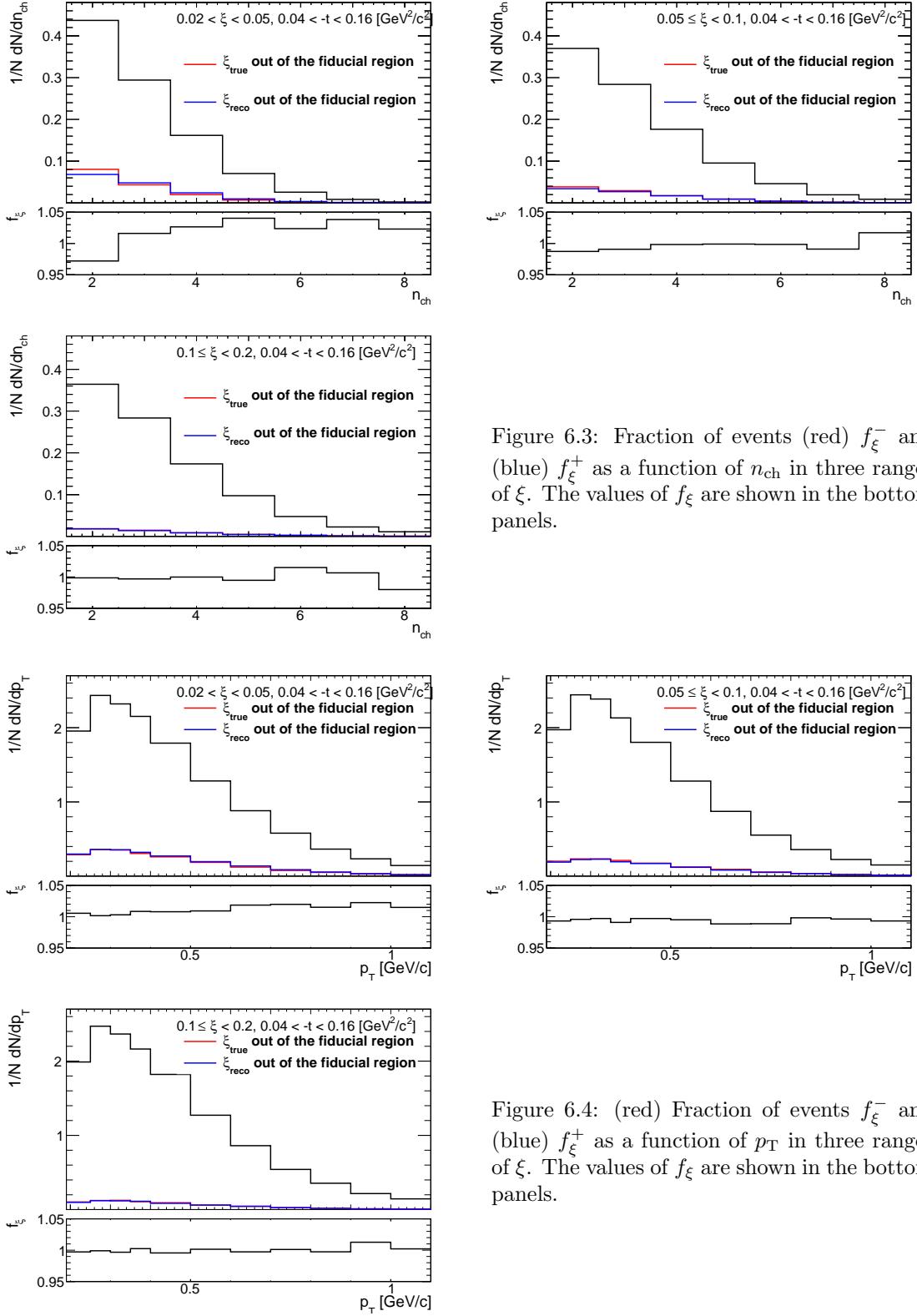


Figure 6.3: Fraction of events (red) f_ξ^- and (blue) f_ξ^+ as a function of n_{ch} in three ranges of ξ . The values of f_ξ are shown in the bottom panels.

Figure 6.4: (red) Fraction of events f_ξ^- and (blue) f_ξ^+ as a function of p_T in three ranges of ξ . The values of f_ξ are shown in the bottom panels.

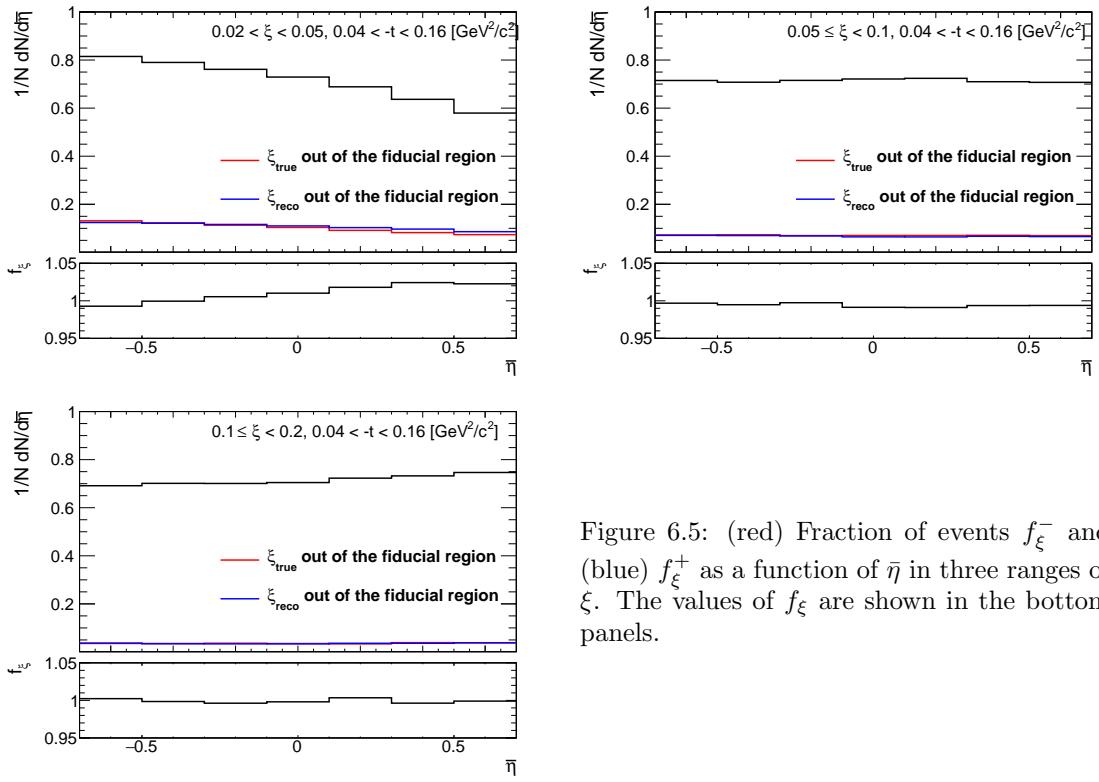


Figure 6.5: (red) Fraction of events f_ξ^- and (blue) f_ξ^+ as a function of $\bar{\eta}$ in three ranges of ξ . The values of f_ξ are shown in the bottom panels.

617 7. Event Corrections and 618 Unfolding Procedure

619 After subtraction of accidental, DD, CD and ND backgrounds (as described in Sec. 4 and 4.2),
620 the data was corrected for detector inefficiencies to obtain the distributions of charged particles
621 and particle to antiparticle (pion, kaon, proton and their antiparticle) multiplicity ratios. These
622 corrections include:

- 623 • event-by-event weights due to vertex reconstruction efficiency:

$$w_{\text{ev}}^{\text{vrt}}(n_{\text{vrt}}^{\text{global}}, |\Delta z_0|) = \frac{1}{\epsilon_{\text{vrt}}(n_{\text{vrt}}^{\text{global}}, |\Delta z_0|)} \cdot \frac{1}{1 - f_{\text{veto}}(n_{\text{vrt}}^{\text{global}}, |\Delta z_0|)} \quad (7.1)$$

624 where the $|\Delta z_0|$ dependence is only applicable for events with $n_{\text{vrt}}^{\text{global}} = 2$ as described in
625 Sec. 5.1.

- 626 • track-by-track weights due to track reconstruction efficiency, track backgrounds, migrations
627 of tracks into and out of the fiducial region:

$$w_{\text{trk}}(p_T, \eta, V_z) = \frac{1 - f_{\text{bkg}}(p_T, \eta) - f_{\text{fake}}(p_T, \eta)}{\epsilon_{\text{TPC}}(p_T, \eta, V_z) \epsilon_{\text{TOF}}(p_T, \eta, V_z)} f_{\text{okr}}(p_T, \eta) \quad (7.2)$$

628 where: $\epsilon_{\text{TPC}}(p_T, \eta, V_z)$ is TPC track reconstruction efficiency [1], $\epsilon_{\text{TOF}}(p_T, \eta, V_z)$ is TOF
629 matching efficiency [1], $f_{\text{okr}}(p_T, \eta)$ is a factor accounting for migrations of tracks into and
630 out of the fiducial region, $f_{\text{bkg}}(p_T, \eta)$ is a fraction of background tracks, and $f_{\text{fake}}(p_T, \eta)$
631 is a fraction of fake tracks. These corrections were not applied for n_{ch} measurements since
632 they were taken into account in the unfolding procedure.

- 633 • event-by-event (for n_{ch} distribution) or track-by-track (for $p_T, \bar{\eta}$ distributions) weights, f_ξ ,
634 due to migrations of events between three ξ regions.

635 Additionally, the obtained distributions were corrected for BBC-small efficiency, ϵ_{BBC} , using
636 the following weight, which was calculated for each true-level quantity $(n_{\text{ch}}, p_T, \bar{\eta})$ in three ranges
637 of ξ separately:

$$w_{\text{BBC}} = \frac{1}{\epsilon_{\text{BBC}}} \quad (7.3)$$

638 In the following sections, the correction procedure for each of the measured distributions is
639 presented separately.

640 7.1 Correction to dN/dn_{sel}

641 In order to express the multiplicity distribution in terms of the number of charged particles, n_{ch} ,
642 instead of the number of selected tracks, n_{sel} , the following procedure based on the Bayesian unfolding
643 [14, 15] was used. First, the n_{sel} distribution was corrected for vertex reconstruction effects
644 by applying event-by-event weights, $w_{\text{ev}}^{\text{vrt}}(n_{\text{vrt}}^{\text{global}}, |\Delta z_0|)$. The number of events in which n_{ch} are
645 produced, $N_{\text{ev}}(n_{\text{ch}})$, can be associated with the number of events in which n_{sel} are reconstructed,
646 $N_{\text{ev}}(n_{\text{sel}})$. Since there are several possible n_{sel} observed in n_{ch} event, $N_{\text{ev}}(n_{\text{ch}})$ is given by:

$$\begin{aligned} N_{\text{ev}}(n_{\text{ch}}) &= \sum_{n_{\text{sel}}=0}^8 P(n_{\text{ch}}|n_{\text{sel}}) \cdot N_{\text{ev}}(n_{\text{sel}}) \\ &= \frac{1}{\epsilon_m(n_{\text{ch}})\epsilon_r(n_{\text{ch}})} \sum_{n_{\text{sel}}=2}^8 P(n_{\text{ch}}|n_{\text{sel}}) \cdot N_{\text{ev}}(n_{\text{sel}}) \end{aligned} \quad (7.4)$$

647 where:

648 $P(n_{\text{ch}}|n_{\text{sel}})$ is the conditional probability of having n_{ch} charged particles in an event in which
 649 n_{sel} tracks were found,

650 $\epsilon_m(n_{\text{ch}})$ is a factor, which recovers events that are lost due to TPC track reconstruction and TOF
 651 matching inefficiencies, i.e. those with $n_{\text{ch}} \geq 2$ but $n_{\text{sel}} < 2$,

652 $\epsilon_r(n_{\text{ch}})$ is a factor, which recovers events which are lost due to fake tracks, i.e. those with $n_{\text{ch}} \leq 8$
 653 but $n_{\text{sel}} > 8$. It was checked that this effect is negligible (smaller than 1%) and can be
 654 omitted.

655 Figure 7.1 shows $\epsilon_m(n_{\text{ch}})$ in three ranges of ξ . It was derived from PYTHIA 8 embedding MC
 656 and varies from about 25% for $n_{\text{ch}} = 2$ to 95% for $n_{\text{ch}} = 8$. Since there are additional data-driven
 657 corrections to TPC and TOF efficiencies, MC simulations were modified by randomly removing
 658 or adding tracks. This was done in accordance with differences in the efficiencies between data
 659 and MC. Figure 7.2 shows $\epsilon_m(n_{\text{ch}})$ calculated in three ranges of ξ using no-pile-up PYTHIA 8 and
 660 EPOS SD+SD'. The differences between these two models, which are up to 12% for $n_{\text{ch}} = 2$ and
 661 $0.02 < \xi < 0.05$, were symmetrized and taken as a systematic uncertainty.

662 The probability $P(n_{\text{ch}}|n_{\text{sel}})$ can be derived using Bayes' theorem, which can be stated mathematically
 663 in terms of charged particle and charged track multiplicities as:

$$P(n_{\text{sel}}|n_{\text{ch}}) \cdot P(n_{\text{ch}}) = P(n_{\text{ch}}|n_{\text{sel}}) \cdot P(n_{\text{sel}}) \quad (7.5)$$

664 where: $P(n_{\text{sel}})$ and $P(n_{\text{ch}})$ are probabilities of observing n_{sel} and n_{ch} respectively, $P(n_{\text{ch}}|n_{\text{sel}})$ and
 665 $P(n_{\text{sel}}|n_{\text{ch}})$ are conditional probabilities.

666 In order to improve the estimate of $P(n_{\text{ch}}|n_{\text{sel}})$, the unfolding is done iteratively:

- 667 • In the first iteration, it is assumed that:

$$P(n_{\text{ch}}|n_{\text{sel}}) = P = P^{\text{MC}}(n_{\text{sel}}|n_{\text{ch}}) \frac{P^{\text{MC}}(n_{\text{ch}})}{P^{\text{MC}}(n_{\text{sel}})} \quad (7.6)$$

$$N_{\text{ev}}(n_{\text{ch}}) = \frac{1}{\epsilon_m(n_{\text{ch}})} \sum_{n_{\text{sel}}=2}^8 N_{\text{ev}}(n_{\text{sel}}) \cdot P \quad (7.7)$$

670 where $P^{\text{MC}}(n_{\text{sel}}|n_{\text{ch}})$, $P^{\text{MC}}(n_{\text{ch}})$ and $P^{\text{MC}}(n_{\text{sel}})$ are obtained from MC. $P^{\text{MC}}(n_{\text{sel}}|n_{\text{ch}})$ is
 671 the same for each iteration.

- 672 • In the $(i + 1)$ th iteration we have:

$$P^{i+1} = P^{\text{MC}}(n_{\text{sel}}|n_{\text{ch}}) \frac{N_{\text{ev}}^i(n_{\text{ch}})}{N_{\text{ev}}(n_{\text{sel}})} \quad (7.8)$$

$$N_{\text{ev}}^{i+1}(n_{\text{ch}}) = \frac{1}{\epsilon_m(n_{\text{ch}})} \sum_{n_{\text{sel}}=2}^8 N_{\text{ev}}(n_{\text{sel}}) \cdot P^{i+1} \quad (7.9)$$

674 where $N_{\text{ev}}^i(n_{\text{ch}})$ is calculated in the previous iteration, and $N_{\text{ev}}(n_{\text{sel}})$ is taken from data.

675 The unfolding matrices $P(n_{\text{ch}}|n_{\text{sel}})$ for each ξ region, shown in Fig. 7.3, were obtained from
 676 PYTHIA 8 embedding MC and used in all iterations of the above procedure. Similarly to $\epsilon_m(n_{\text{ch}})$,
 677 the matrices were modified by randomly removing or adding tracks in order to take into account
 678 additional data-driven corrections to TPC and TOF efficiencies. In order to increase statistical
 679 precision of the unfolding matrices, all simulated events were used, i.e. also those with additional
 680 fake vertices (with n_{sel} defined as a number of primary tracks associated with the best vertex).
 681 The systematic uncertainty related to limited statistics in PYTHIA 8 was estimated by performing
 682 50 pseudo-experiments, in which the unfolding matrices were smeared according to their statistical

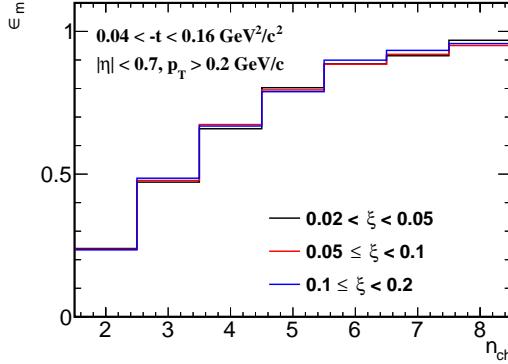


Figure 7.1: $\epsilon_m(n_{ch})$ calculated separately in three ranges of ξ using PYTHIA 8 embedding MC.

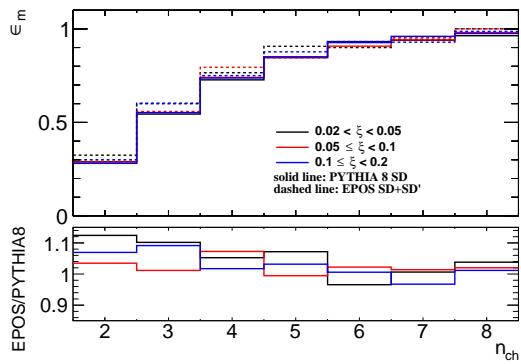


Figure 7.2: Comparison of $\epsilon_m(n_{ch})$ calculated separately in three ranges of ξ using PYTHIA 8 SD and EPOS SD+SD' no-pile-up MCs. The ratios of EPOS to PYTHIA 8 predictions are shown in the bottom panel.

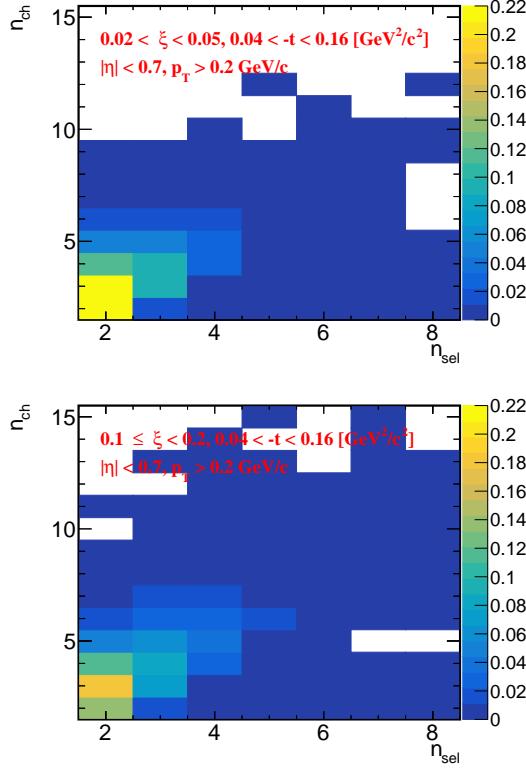


Figure 7.3: The unfolding matrices calculated from PYTHIA 8 embedding MC for three ranges of ξ separately.

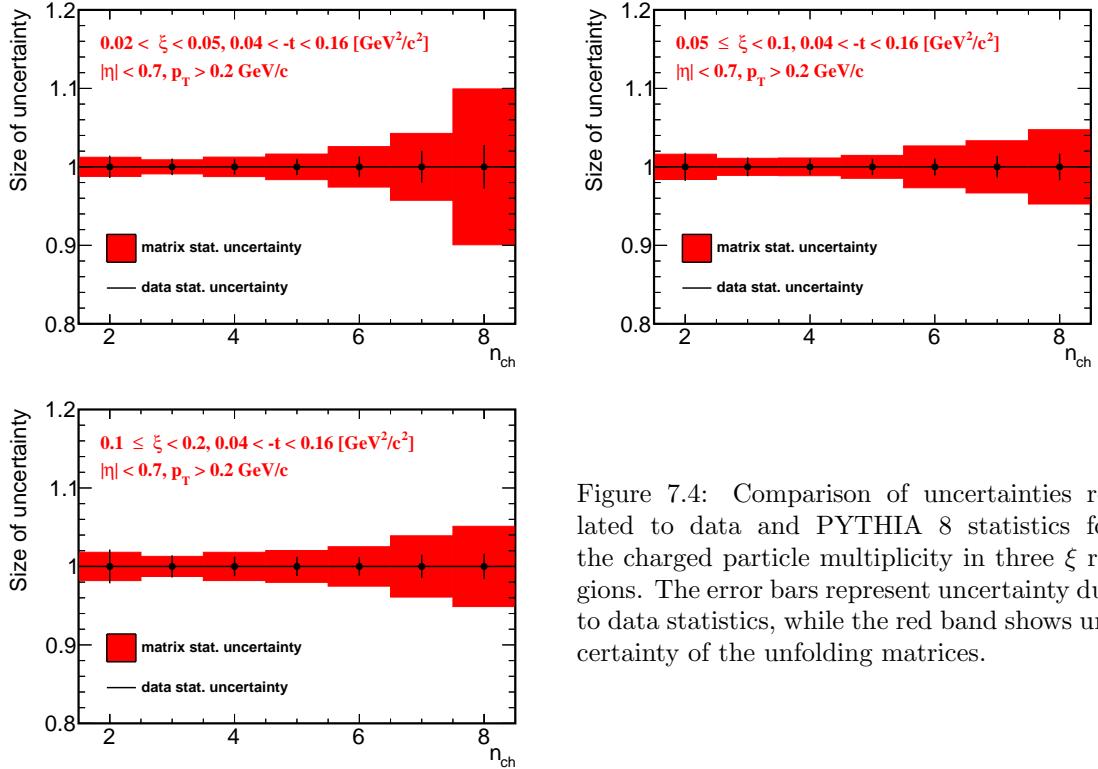


Figure 7.4: Comparison of uncertainties related to data and PYTHIA 8 statistics for the charged particle multiplicity in three ξ regions. The error bars represent uncertainty due to data statistics, while the red band shows uncertainty of the unfolding matrices.

uncertainties. It affects mainly large charged-particle multiplicities, where it is about 8 – 10% (as shown in Fig. 7.4), and is smaller or at the same level as other components contributing to the total systematic uncertainty.

The distribution dN/dn_{ch} obtained after the unfolding procedure was corrected for BBC-small efficiency, through $w_{\text{BBC}}(n_{\text{ch}})$ weights, and migrations of events between ξ ranges, through $f_{\xi}(n_{\text{ch}})$ weights. Since the unfolding matrices contain track reconstruction efficiencies, non-primary track backgrounds, migrations of tracks into and out of the fiducial region, the weight $w_{\text{trk}}(p_{\text{T}}, \eta, V_z)$ was not used.

Finally, the dN/dn_{ch} distribution was normalized to the total number of events, $N_{\text{ev}} = N$, which was calculated as the integral of the unfolded distribution.

7.2 Correction to Transverse Momentum and Pseudorapidity Distributions

First the accidental and non-SD backgrounds were subtracted from the p_{T} and $\bar{\eta}$ distributions. Next, each event was corrected for vertex reconstruction efficiency by applying $w_{\text{ev}}^{\text{vrt}}(n_{\text{vrt}}^{\text{global}}, |\Delta z_0|)$ weights. Then, the tracks were corrected for the track reconstruction efficiency, non-primary track background contribution, track and ξ migrations, BBC-small efficiency (the product of $w_{\text{trk}}(p_{\text{T}}, \eta, V_z)$, f_{ξ} and w_{BBC} weights was applied, f_{ξ} and w_{BBC} were calculated as a function of true-level p_{T} and $\bar{\eta}$ separately).

In order to obtain charged-particle densities, the p_{T} and $\bar{\eta}$ distributions were normalized to unity and scaled by the average charged particle multiplicity in an event $\langle n_{\text{ch}} \rangle$. The latter was calculated from the corrected charged particle multiplicity distribution dN/dn_{ch} (Sec. 7.1). The above procedure was done to correct the data also for events that are lost due to $n_{\text{sel}} < 2$ but $n_{\text{ch}} \geq 2$ since such correction was not included in any event-by-event and track-by-track weights. There was an assumption that p_{T} and η distributions are the same for lost and measured events, but it

707 was validated by the closure tests (Sec. 7.3). The mean p_T and $\bar{\eta}$ in an event, $\langle p_T \rangle$ and $\langle \bar{\eta} \rangle$, were
708 obtained from the measured distributions.

709 7.3 Closure Tests

710 In order to validate the correction procedures, closure tests were performed, i.e. full correction
711 procedure was applied to the MC detector-level distributions and the results were directly com-
712 pared to the true-level distributions. Figure 7.5 shows closure tests of multiplicity, transverse
713 momentum and pseudorapidity distributions for three ranges of ξ , separately. PYTHIA 8 SD
714 embedding MC was used as an input. In order to compare corrected and true-level distributions,
715 the statistical uncertainties of the true-level distributions were assumed to be 0. The difference
716 between true-level and corrected distributions was taken as a systematic uncertainties.

717 7.4 EAST-WEST asymmetry

718 Another kind of consistency check can be performed by comparing the results obtained by tag-
719 ging forward-scattered protons in different detectors. Therefore, each distribution was measured
720 separately for events in which forward-scattered proton is on one and the other side of the IP
721 (east-west). Figure 7.6 shows the tests of multiplicity, transverse momentum and pseudorapidity
722 distributions for three ranges of ξ , separately. Both statistical uncertainty components, due to
723 input data and due to unfolding matrix, are added in quadrature for n_{ch} distributions. The largest
724 difference is observed for charged-particle multiplicity distributions, where it varies up to 20% for
725 $n_{ch} = 8$ and $0.02 < \xi < 0.1$. For the rest multiplicities and ξ ranges, the differences are smaller
726 (< 10%). In case of p_T and $\bar{\eta}$ distributions, a level of these disagreements is below 5%.

727 The deviations between the distributions for events with forward-scattered proton on east and
728 west side of the IP were fitted with a constant (Fig. 7.6). The quality of the fit shows that the
729 disagreements are compatible with statistical fluctuations (χ^2/ndf close to 1) for multiplicity and
730 transverse momentum distributions. For pseudorapidity, the χ^2/ndf is significantly larger than 1.

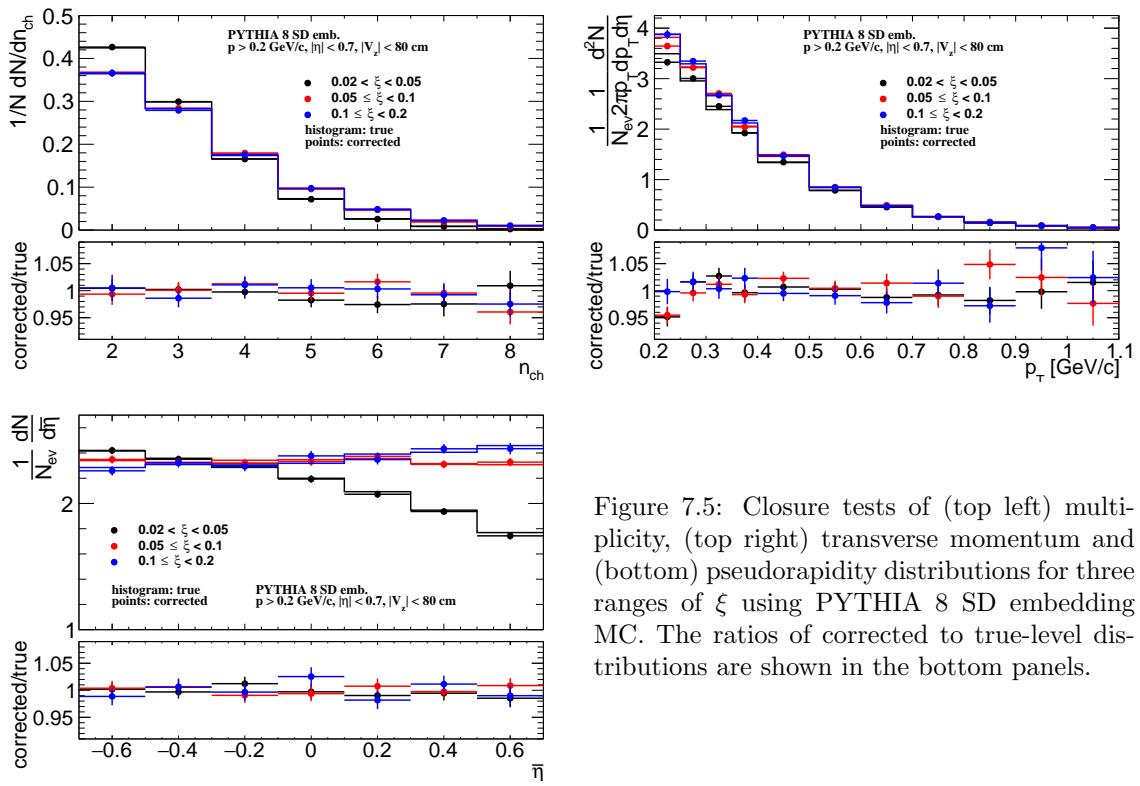


Figure 7.5: Closure tests of (top left) multiplicity, (top right) transverse momentum and (bottom) pseudorapidity distributions for three ranges of ξ using PYTHIA 8 SD embedding MC. The ratios of corrected to true-level distributions are shown in the bottom panels.

731 Therefore, half of the differences between east and west distributions were used to be systematic
 732 uncertainty for $\bar{\eta}$ distributions.

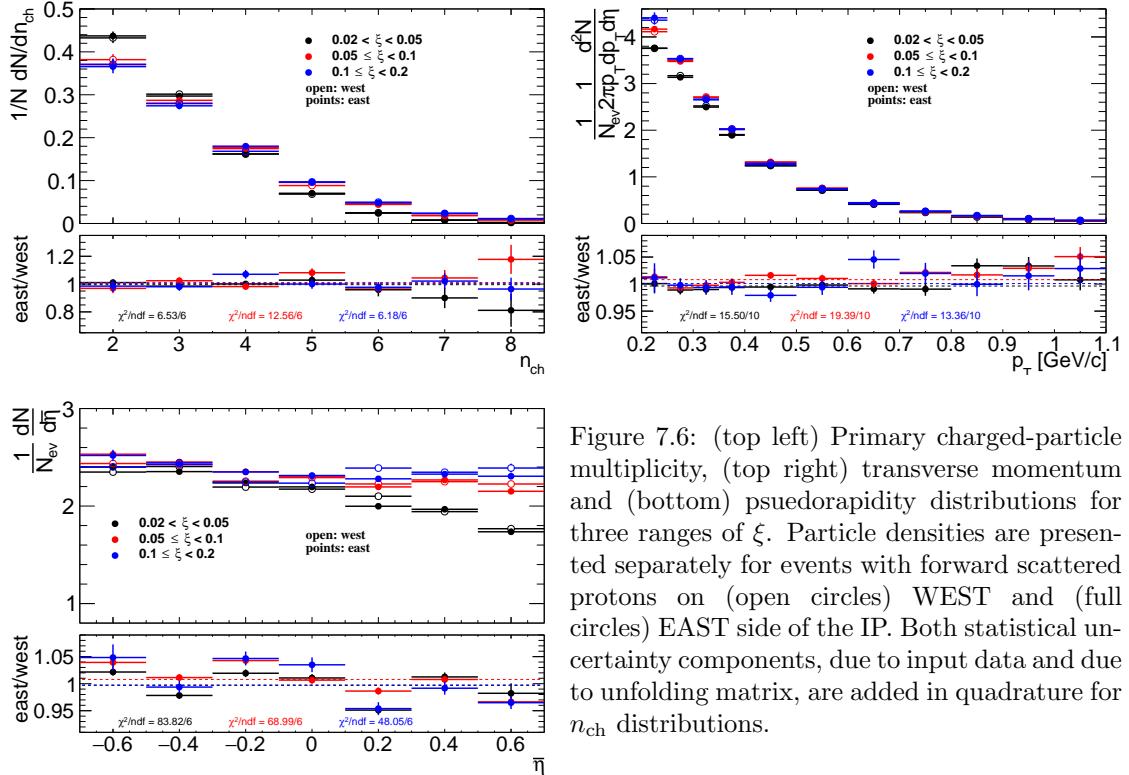


Figure 7.6: (top left) Primary charged-particle multiplicity, (top right) transverse momentum and (bottom) psuedorapidity distributions for three ranges of ξ . Particle densities are presented separately for events with forward scattered protons on (open circles) WEST and (full circles) EAST side of the IP. Both statistical uncertainty components, due to input data and due to unfolding matrix, are added in quadrature for n_{ch} distributions.

7.5 Particle Identification

733 Specific ionization energy loss, the dE/dx , is a function of the magnitude of a particle momentum.
 734 In this section the particle identification with help of dE/dx is described. Due to a low particle
 735 multiplicity and lack of signal in VPDs on the outgoing proton side (presence of the rapidity
 736 gap) in SD events, the time of collision is not defined precisely enough, therefore, the particle
 737 identification by the TOF is not possible and the analysis was limited to identification only by
 738 dE/dx .

739 The ionization energy loss of charged particles in material is given by the Bethe-Bloch formula
 740 and for the STAR TPC by the more precise Bichsel formula [16]. The particle type can be
 741 determined by comparison of particle's dE/dx with the Bethe-Bloch (Bichsel) expectations. Figure
 742 7.7 shows the dE/dx versus rigidity $q \times p$ for particles in $|\eta| < 0.7$. Particles are well separated at
 743 low $|q \times p|$, whereas at higher $|q \times p|$ the dE/dx of different particle species starts to overlap: e^\pm
 744 and K^\pm merge at ~ 0.4 GeV/c, K^\pm and π^\pm merge at ~ 0.65 GeV/c, and $p(\bar{p})$ and π^\pm merge at
 745 ~ 1 GeV/c. Since the dE/dx distribution for a given particle type is not Gaussian, the following
 746 variable for each particle type was defined:

$$n\sigma_{dE/dx}^i = \ln \left(\frac{dE/dx}{(dE/dx)_i^{BB}} \right) / \sigma \quad (7.10)$$

747 where $(dE/dx)_i^{BB}$ is the Bethe-Bloch (Bichsel) expectation of dE/dx for the given particle type
 748 i ($i = \pi, K, p$), σ - the relative dE/dx resolution. The expected value of $n\sigma_{dE/dx}^i$ for the particle
 749 under consideration is 0 and the width equals to 1. The sample $n\sigma_{dE/dx}^i$ distribution for π^\pm , K^\pm
 750 and $p(\bar{p})$ in one ξ range, $0.02 < \xi < 0.05$, is shown in Fig. 7.8.

752 Figure 7.9 shows the $n\sigma_{dE/dx}^{\pi^\pm}$, $n\sigma_{dE/dx}^{K^\pm}$ and $n\sigma_{dE/dx}^{p(\bar{p})}$ distributions for $0.6 < p_T < 0.65$ GeV/c in
 753 the ξ range, $0.02 < \xi < 0.05$, each corrected for the energy loss (mass of i -particle was assumed) [1]
 754 and vertexing (other p_T bins are shown in Appendix B). To extract the particle yield for a given
 755 particle type, a multi-Gaussian fit is applied to the $n\sigma_{dE/dx}^i$ distribution in each p_T bin and ξ range.
 756 The parameters of the multi-Gaussian fit are the centroids μ_{i^-/i^+} , widths σ_{i^-/i^+} , sums and ratios of
 757 yields C_{i^-/i^+} , r_{i^-/i^+} for negative i^- and positive i^+ particles (π^\pm , e^\pm , K^\pm , p and \bar{p}). The positive
 758 and negative particle $n\sigma_{dE/dx}^i$ -distributions are fitted simultaneously, where the centroids and
 759 widths are kept the same for particle and antiparticle. In some p_T regions, the fit does not
 760 converge, because different particle species are not well separated there. Therefore, multiple steps
 761 of the fitting procedure are performed to reduce the number of free parameters in the final fit
 762 and ensure its stability. Almost all centroids and widths are constrained by a function with free
 763 parameters p_k , where $k \in \mathbb{N}$. The function is chosen to describe the data as best as possible. Since
 764 dE/dx is a function of the particle's momentum and its shape should be independent of the process
 765 under study, the values of p_k are obtained only for events with $0.02 < \xi < 0.05$ and kept the same
 766 for other ξ ranges. The electron contributions are fitted only at $p_T < 0.4$ GeV/c, separately for
 767 each particle species and ξ range. For higher p_T ranges, they are estimated from PYTHIA 8
 768 embedding MC, and scaled according to the ratio of PYTHIA 8 predictions and contributions
 769 fitted in the $0.35 < p_T < 0.4$ GeV/c bin. The procedure slightly differs for different particle types.
 770 In each step, the multi-Gaussian fit is performed first, then the widths and centroids are fitted
 771 in p_T ranges in which the fit applied to $n\sigma_{dE/dx}^i$ converges. Later, the widths and centroids are
 772 extrapolated to other p_T ranges, in which particle species are not well separated:

773 1. π^\pm :

774 • Step 1 (Fig. 7.10):

- 775 – Analyze data with $0.2 < p_T < 0.65$ GeV/c
- 776 – Fit μ_{π^-/π^+} and σ_{π^-/π^+} as a function of p_T with a polynomial $p_0 p_T^3 + p_1 p_T^2 + p_2 p_T + p_3$
- 777 – Fit μ_{K^-/K^+} as a function of p_T with $p_0 \exp(p_1 p_T) + p_2$
- 778 – Fit μ_{e^-/e^+} as a function of p_T with $p_0 \exp[-(p_1 p_T)^{p_2}]$
- 779 – Fit σ_{K^-/K^+} as a function of p_T , for $0.3 < p_T < 0.5$ GeV/c, with constant p_0
- 780 – Fit $\mu_{\bar{p}/p}$ and $\sigma_{\bar{p}/p}$ as a function of p_T with $p_0 \exp(p_1 p_T)$

781 • Step 2:

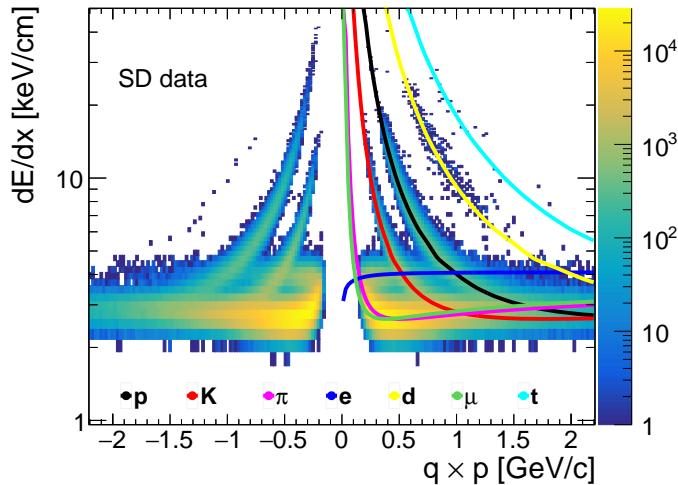


Figure 7.7: Specific ionization energy loss dE/dx as a function of rigidity $q \times p$ for particles in $|\eta| < 0.7$. The Bichsel predictions for each particle species are also shown.

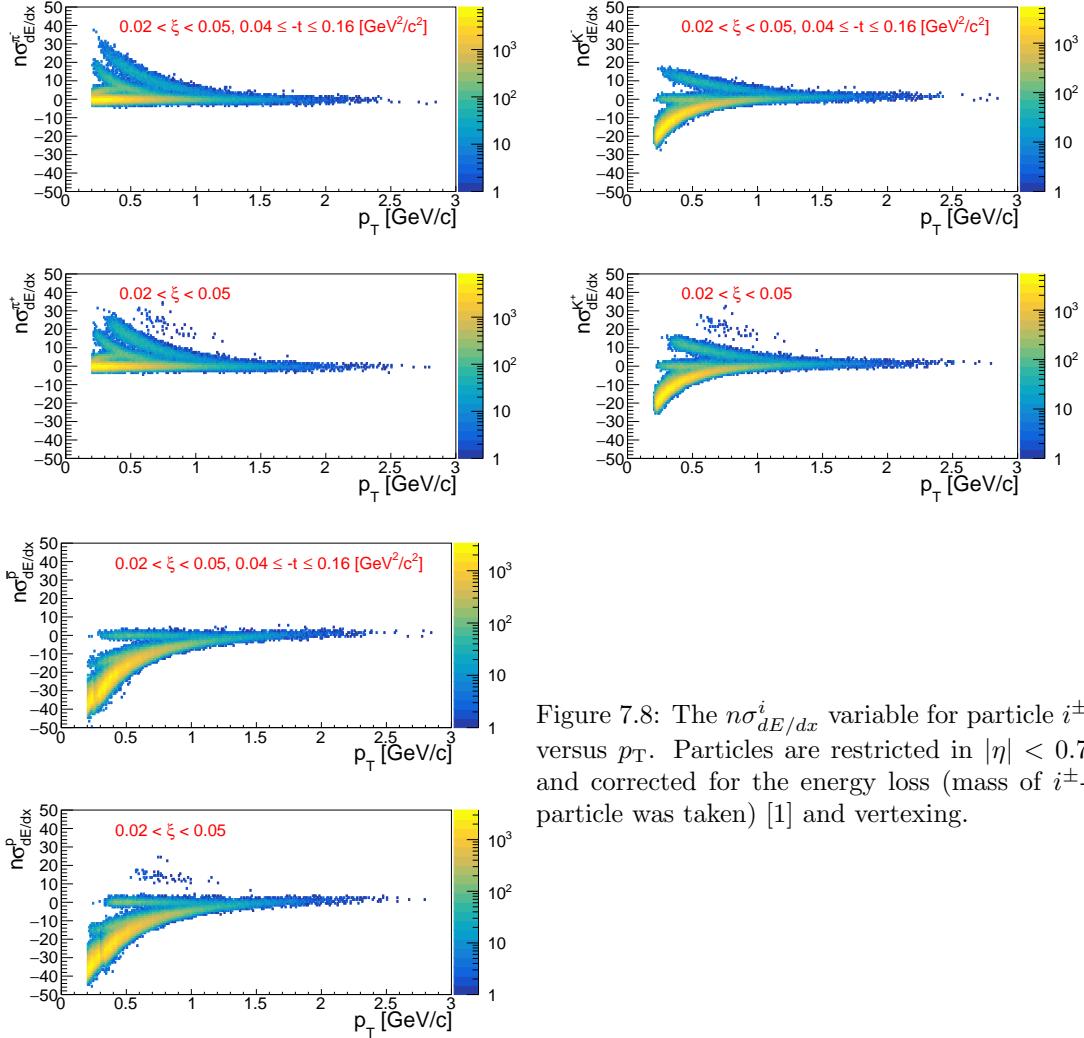


Figure 7.8: The $n\sigma_{dE/dx}^i$ variable for particle i^\pm versus p_T . Particles are restricted in $|\eta| < 0.7$ and corrected for the energy loss (mass of i^\pm -particle was taken) [1] and vertexing.

- 782 – σ_{e^-/e^+} fixed to 1.2 and 0.8 for $0.2 < p_T < 0.4$ and $0.4 < p_T < 0.7$, respectively
- 783 – Fit σ_{K^-/K^+} as a function of p_T , for $0.3 < p_T < 0.7$ GeV/c, with constant p_0 and fix it to the value of p_0
- 784 – The rest parameters from Step 1 are fixed to the values calculated from functions obtained in Step 1: μ_{π^-/π^+} , σ_{π^-/π^+} , μ_{e^-/e^+} , μ_{K^-/K^+} , $\mu_{\bar{p}/p}$, $\sigma_{\bar{p}/p}$

787 2. K^\pm :

- 788 • Step 1 (Fig. 7.11):
 - 789 – Analyze data with $0.2 < p_T < 0.6$ GeV/c
 - 790 – Fit μ_{π^-/π^+} as a function of p_T with $-\exp(p_0 + p_1 p_T)$
 - 791 – Fit σ_{π^-/π^+} , σ_{e^-/e^+} , σ_{K^-/K^+} as a function of p_T with $\exp(p_0 + p_1 p_T)$
 - 792 – Fit μ_{e^-/e^+} as a function of p_T with a polynomial $p_0 p_T^3 + p_1 p_T^2 + p_2 p_T + p_3$
 - 793 – Fit μ_{K^-/K^+} as a function of p_T with a polynomial $p_0 + p_1 p_T^2$
- 794 • Step 2:
 - 795 – All parameters from Step 1 except σ_{e^-/e^+} are fixed to the values calculated from functions obtained in Step 1

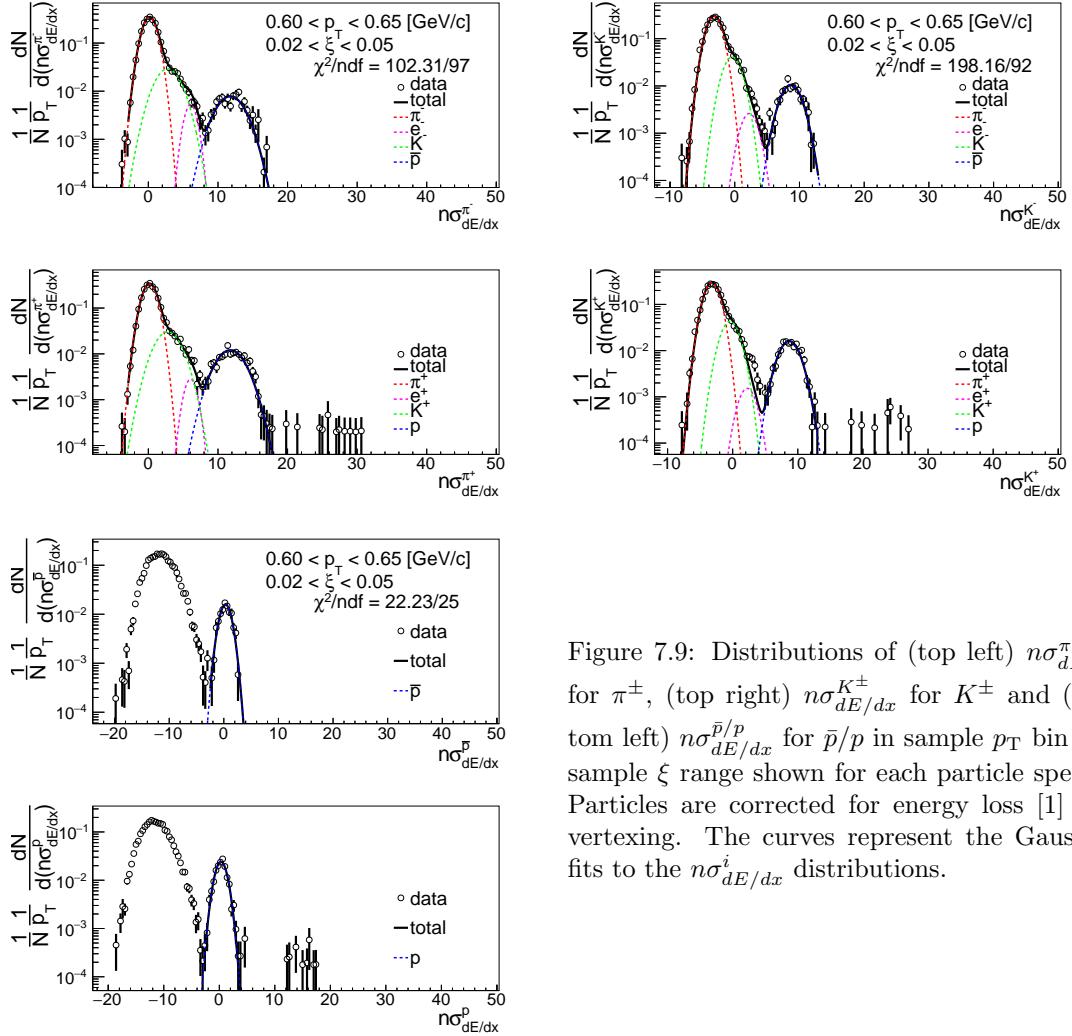


Figure 7.9: Distributions of (top left) $n\sigma_{dE/dx}^{\pi^\pm}$ for π^\pm , (top right) $n\sigma_{dE/dx}^{K^\pm}$ for K^\pm and (bottom left) $n\sigma_{dE/dx}^{\bar{p}/p}$ for \bar{p}/p in sample p_T bin and sample ξ range shown for each particle species. Particles are corrected for energy loss [1] and vertexing. The curves represent the Gaussian fits to the $n\sigma_{dE/dx}^i$ distributions.

- 797 – Fit σ_{e^-/e^+} as a function of p_T , for $0.45 < p_T < 0.65$ GeV/c, with constant p_0
 798 • Step 3:
 799 – σ_{e^-/e^+} fixed to the values calculated from functions obtained in Steps 1 and 2 for
 800 $0.3 < p_T < 0.45$ and $0.45 < p_T < 0.65$, respectively.
 801 – The rest parameters from Step 1 are fixed to the values calculated from functions
 802 obtained in Step 1: μ_{π^-/π^+} , σ_{π^-/π^+} , μ_{e^-/e^+} , μ_{K^-/K^+} , σ_{K^-/K^+}

803 3. \bar{p}, p :

- 804 • Step 1 (Fig. 7.12):
 805 – Analyze data with $0.4 < p_T < 0.9$ GeV/c
 806 – Fit μ_{π^-/π^+} , μ_{K^-/K^+} as a function of p_T with a polynomial $p_0 p_T + p_1$
 807 – Fit σ_{π^-/π^+} as a function of p_T with a polynomial $p_0 p_T^2 + p_1 p_T + p_2$
 808 – Fit σ_{K^-/K^+} as a function of p_T with $\exp(p_0 + p_1 p_T)$
 809 • Step 2:
 810 – μ_{K^-/K^+} fixed to the values calculated from a function obtained in Step 1

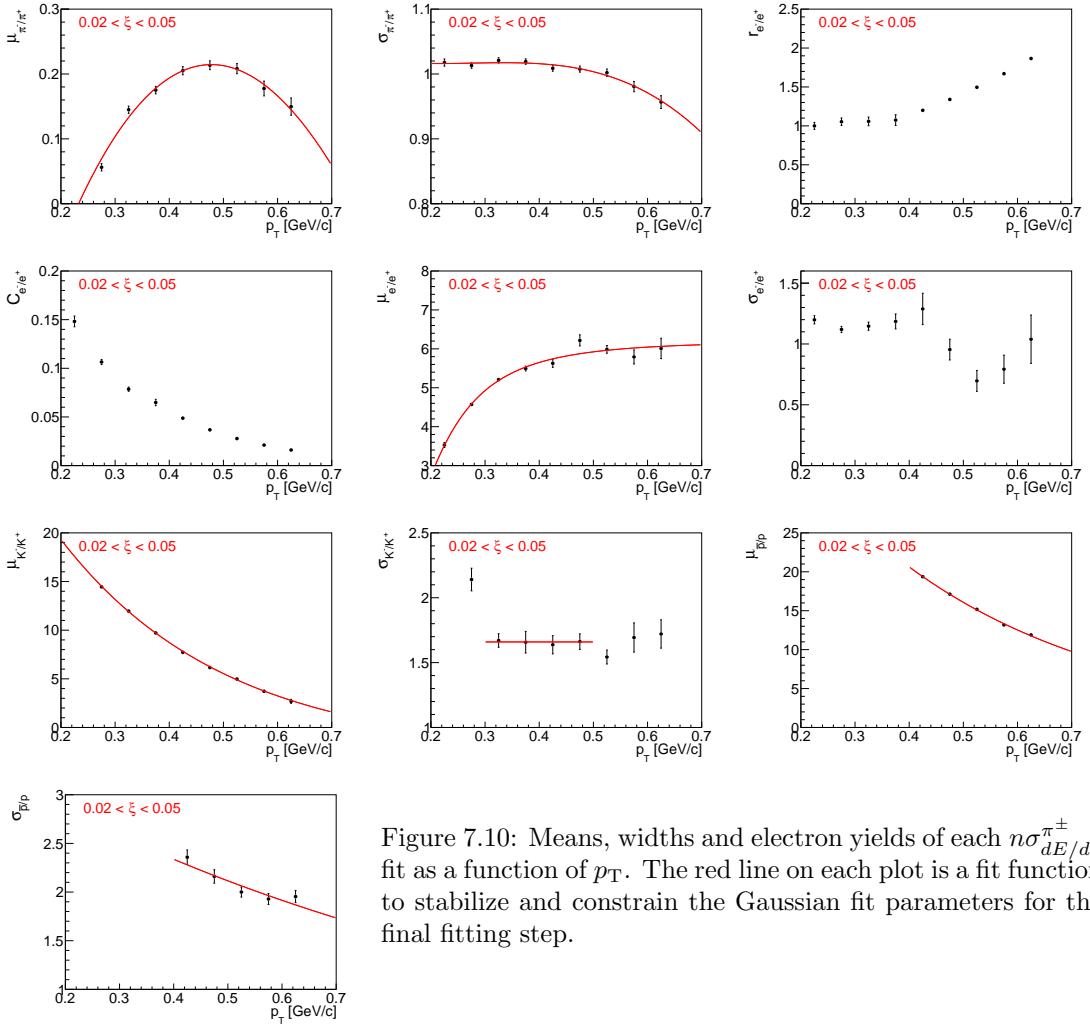


Figure 7.10: Means, widths and electron yields of each $n\sigma_{dE/dx}^{\pi^\pm}$ fit as a function of p_T . The red line on each plot is a fit function to stabilize and constrain the Gaussian fit parameters for the final fitting step.

- 811 – All the rest parameters from Step 1 are limited to the values calculated from functions obtained in Step 1
- 812 – Fit μ_{π^-/π^+} , σ_{π^-/π^+} , σ_{K^-/K^+} as a function of p_T with a polynomial $p_0 p_T^2 + p_1 p_T + p_2$
- 813 – Fit $\mu_{\bar{p}/p}$ as a function of p_T , for $0.7 < p_T < 1.0$ GeV/c, with constant p_0
- 814
- 815 • Step 3:
 - 816 – μ_{K^-/K^+} fixed to the values calculated from a function obtained in Step 1
 - 817 – $\mu_{\bar{p}/p}$ fixed to the values calculated from a function obtained in Step 2 for $0.7 < p_T < 1.0$
 - 818 – The rest parameters from Step 2 are fixed to the values calculated from functions obtained in Step 2: μ_{π^-/π^+} , σ_{π^-/π^+} , σ_{K^-/K^+}

821 The particle yield is extracted from the fit to the corresponding $n\sigma_{dE/dx}^i$ distribution (corrected
 822 only for the energy loss and vertexing). As shown in Fig. 7.8, the dE/dx of each particle type merge
 823 at large p_T . Hence, the particle identification is limited. Pions can be identified in the momentum
 824 range of 0.2 – 0.7 GeV/c, kaons in 0.3 – 0.65 GeV/c and (anti)protons in 0.4 – 1.0 GeV/c.

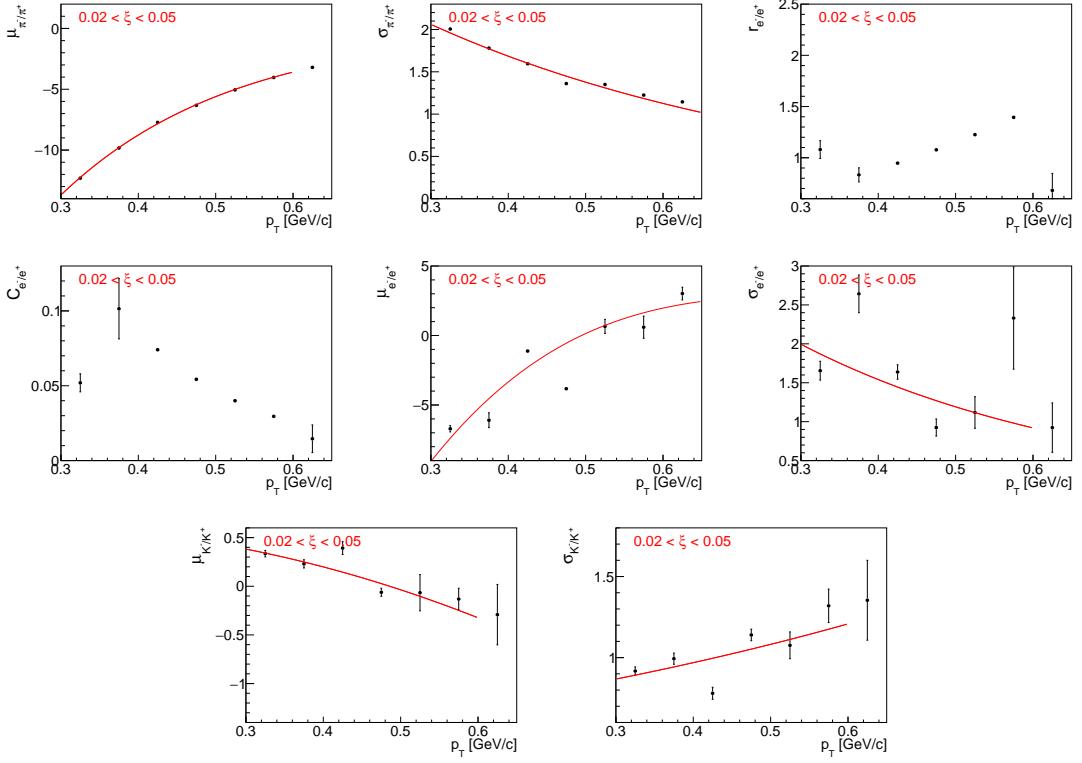


Figure 7.11: Means, widths and electron yields of each $n\sigma_{dE/dx}^{K^\pm}$ fit as a function of p_T . The red line on each plot is a fit function to stabilize and constrain the Gaussian fit parameters for the final fitting step.

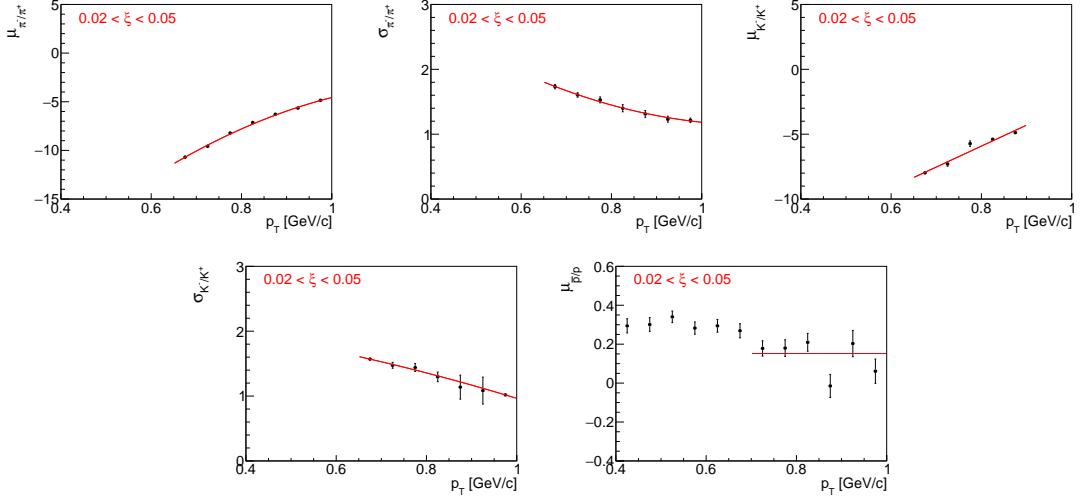


Figure 7.12: Means and widths of each $n\sigma_{dE/dx}^{\bar{p}/p}$ fit as a function of p_T . The red line on each plot is a fit function to stabilize and constrain the Gaussian fit parameters for the final fitting step.

825 7.6 Antiparticle-to-Particle Ratios

826 The following steps were taken to correct the identified antiparticle to particle (pion, kaon, proton
 827 and their antiparticle) multiplicity ratios as a function of p_T in three ranges of ξ :

- 828 • The raw identified particle yields were obtained through multi-Gaussian fits to the $n\sigma_{dE/dx}^i$
 829 distributions (Sec. 7.5), where the vertex reconstruction and energy loss corrections [1] were
 830 applied. The latter depends on the particle type.
- 831 • The non-SD background (Sec. 4.2) is the same for particles and antiparticles, thus, it was
 832 not subtracted. The accidental background contribution (Sec. 4) is very small, hence, any
 833 particle-antiparticle differences have a negligible effect on the result. Therefore, it was as-
 834 sumed that the accidental background does not depend on the particle type and for this
 835 reason it was not subtracted.
- 836 • The particle yields were corrected for track reconstruction efficiencies [1], which depend on
 837 the particle type and charge. These corrections are averaged over η and V_z . The ratio of
 838 particle to antiparticle TPC-TOF efficiencies is shown in Fig. 7.13. It weakly depends on ξ
 839 range, therefore, only sample results for single range of $0.02 < \xi < 0.05$ are presented.
- 840 • The background from non-primary tracks was subtracted (Sec. 4.1):
 - 841 – π^\pm : weak decays pions, muon contribution and background from detector dead-material
 interactions,
 - 842 – p : background from detector dead-material interactions,
 - 843 – p, \bar{p} : reconstructed tracks which have the appropriate number of common hit points
 with true-level particle, but the distance between them is too large (this background is
 negligibly small for other particle types),
 - 844 – fake track contribution was assumed to be the same for each particle type, hence, it
 was not subtracted.
- 845 • Since track and ξ migrations, and BBC-small efficiency, do not depend on the particle type
 846 and charge, these corrections are not applied.
- 847 • Finally, each antiparticle p_T distribution was divided by the corresponding particle p_T dis-
 848 tribution to obtain fully corrected identified antiparticle to particle multiplicity ratios.
- 849 • Additionally, the average antiparticle to particle ratios over fiducial region of p_T in each ξ
 850 region were calculated.

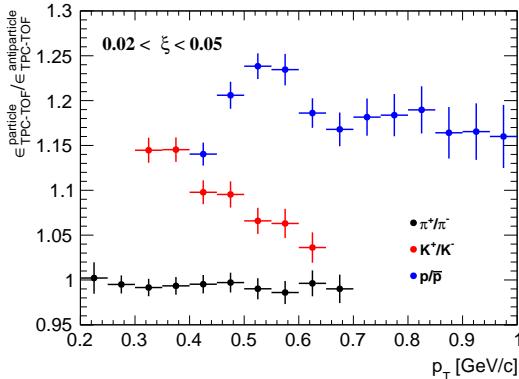


Figure 7.13: Ratio of particle to antiparticle TPC-TOF efficiencies for $0.02 < \xi < 0.05$.

8. Systematic Uncertainties

Apart from the statistical uncertainties there are also systematic uncertainties originating from inefficiencies and limitations of the measurement devices and techniques. The following sources of systematic uncertainties were considered:

- the effect of off-time pile-up on TPC track reconstruction efficiency [1],
- the uncertainty of TPC track reconstruction efficiency related to the description of dead-material in simulation [1],
- representation of data sample in embedding MC [1],
- variation in the track quality cuts [1],
- non-primary track background contribution (Sec. 4.1),
- fake track background contribution (Sec. 4.1),
- TOF system simulation accuracy [1],
- accidental background contribution (Sec. 4),
- the effect of alternative model of hadronization on BBC-small efficiency (Sec. 5.2),
- non-SD background contribution (Sec. 4.2),
- the effect of alternative model on ϵ_m correction (Sec. 7.1),
- non-closure (Sec 7.3),
- non-closure of N_{ev} , applied only to p_T and $\bar{\eta}$ distributions,
- difference in the distributions calculated separately for events in which forward proton is on one and the other side of the IP (east-west, Sec 7.4),
- the effect of $(K^+ + K^-) / (\pi^+ + \pi^-)$ ratio on TPC and TOF efficiencies.

Some of the systematic uncertainties on $1/N dN/dn_{\text{ch}}$ (related to TPC and TOF reconstruction efficiencies, fake track background contribution) are propagated by randomly removing and adding tracks in the n_{sel} distribution before unfolding procedure. For each track, a random number is generated. If this number is smaller than the absolute value of systematic uncertainty, then n_{sel} is increased or decreased, depending on the sign of systematic uncertainty.

Figures 8.1 to 8.3 show the components contributing to the total systematic uncertainty for charged particle distributions without the identification. The dominant systematic uncertainty for p_T and n_{ch} distributions is related to TOF system simulation accuracy. It affects mainly low- p_T particles, where it is about 2 – 3%, and large charged particle multiplicities, where it varies up to 20% for $n_{\text{ch}} = 8$ and $0.02 < \xi < 0.05$. In case of $\bar{\eta}$ distribution, the systematic uncertainty on TOF mainly refers to charged particles produced at the edge of the fiducial region, for which it is about 2%. The largest (up to 30%) systematic uncertainty for $\langle \bar{\eta} \rangle$, is related to the observed difference in the distributions calculated separately with respect to the forward-scattered proton direction. The rest of the components have smaller contributions to the total systematic uncertainty. The systematic uncertainty on non-closure is on average at the level of 2% which proves the accuracy of the correction procedure.

TPC and TOF efficiencies for a charged particle without the identification are calculated by mixing charged pions, kaons and (anti)protons. Their relative contributions are based on PYTHIA 8 model. However, the ratios of production yields of $(K^+ + K^-) / (\pi^+ + \pi^-)$ are measured to be different than predicted PYTHIA 8 (Fig. 9.8). Therefore, the TPC and TOF efficiencies

896 calculated with the modified K^\pm/π^\pm ratios are used to introduce an additional systematic uncertainty.
 897 It is about 2 – 3% for particles with $p_T > 0.6 \text{ GeV}/c$ and up to 5% for $n_{\text{ch}} = 8$.

898 Figures 8.4 to 8.7 show breakdown of all different systematics for the antiparticle-to-particle
 899 multiplicity ratio distributions. An additional systematic contribution for \bar{p}/p multiplicity ratio
 900 due to proton background estimation was introduced. Since most of the corrections are the same
 901 for particle and its antiparticle, nearly all systematic uncertainties cancel out in the antiparticle-to-
 902 particle ratios. The largest sources of systematics, which do not, are related to proton background
 903 estimation and dead-material effect on TPC track reconstruction efficiency. The former was found
 904 to be up to 6%, whereas the latter varies up to 2% for low- p_T \bar{p}/p multiplicity ratio.

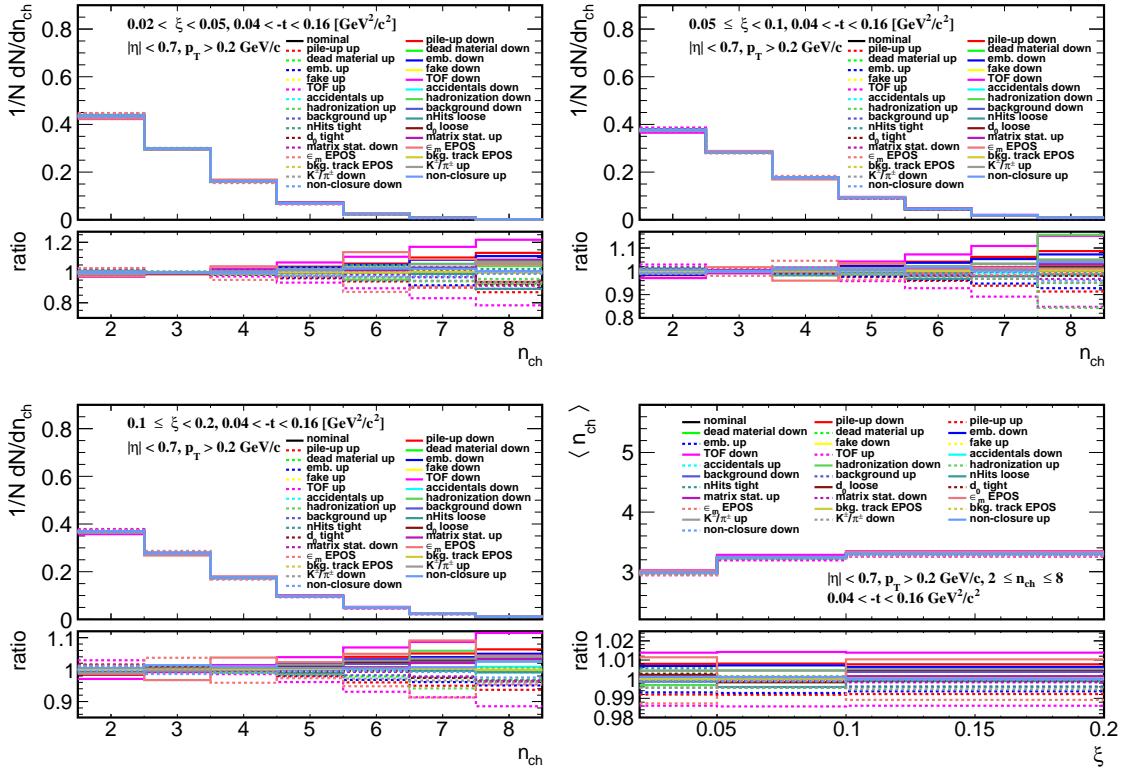


Figure 8.1: Components of the systematic uncertainties for the charged particle multiplicity in three ξ regions and for the average charged particle multiplicity.

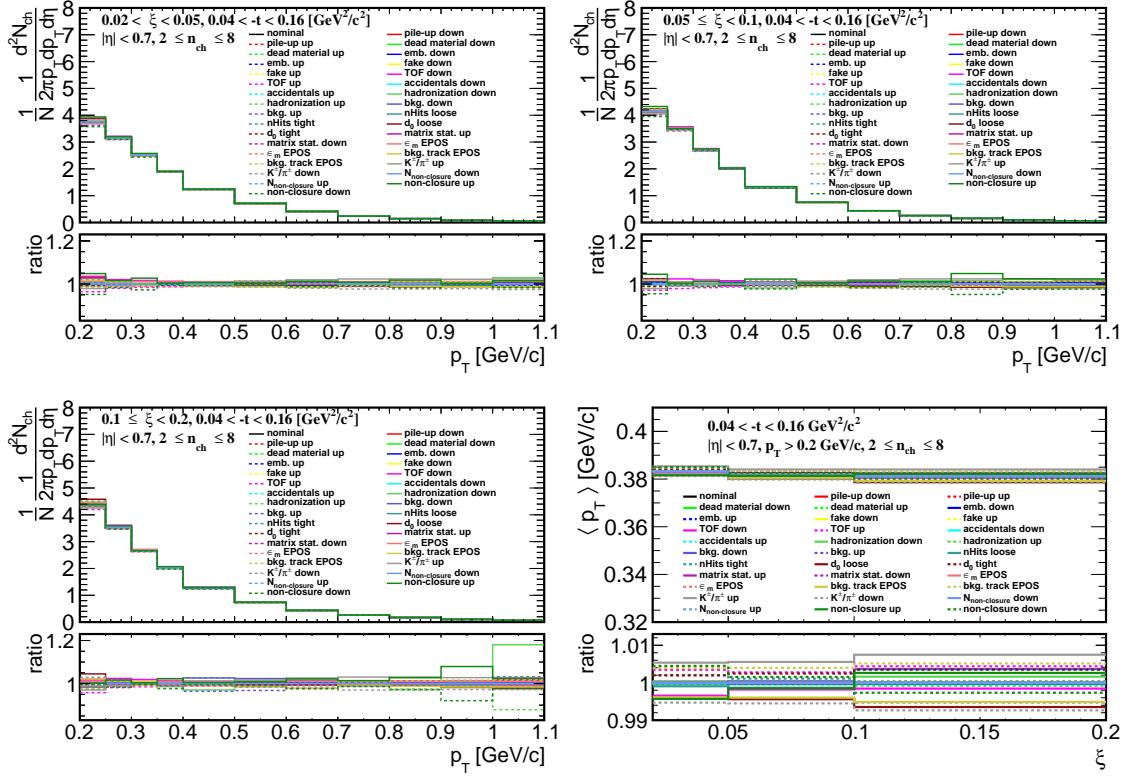


Figure 8.2: Components of the systematic uncertainties for p_T distributions in three ξ regions and for an average p_T distribution.

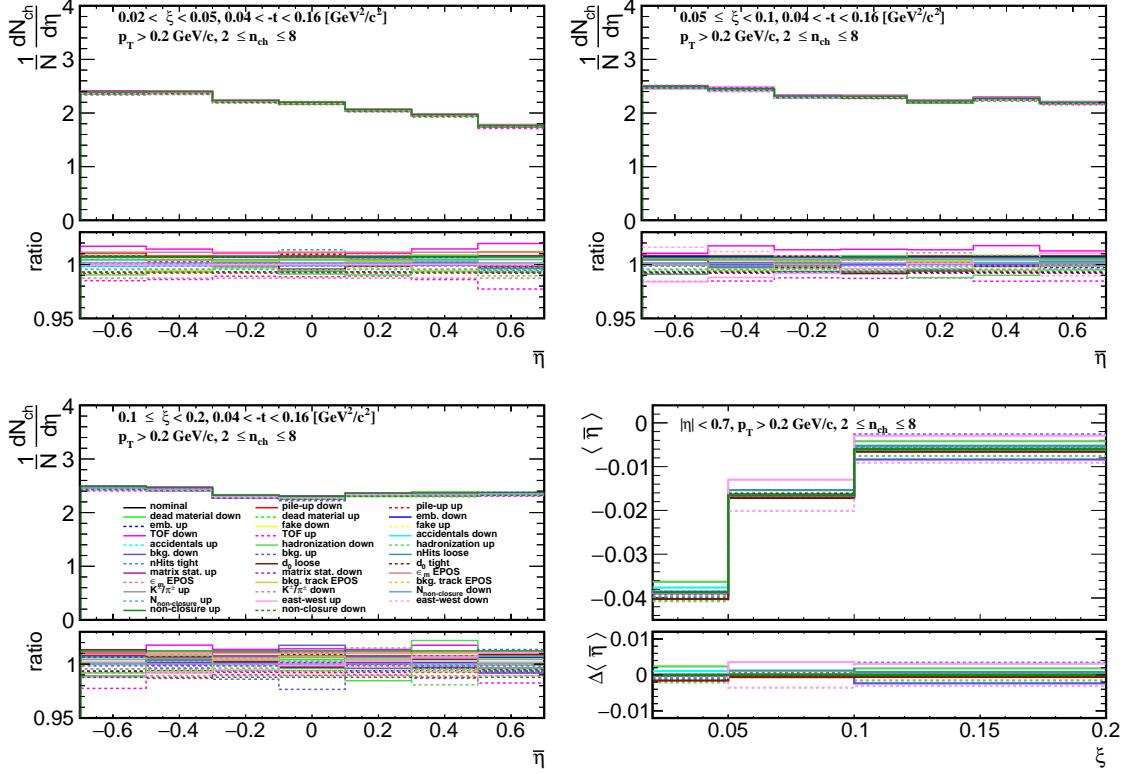


Figure 8.3: Components of the systematic uncertainties for $\bar{\eta}$ distributions in three ξ regions and for an average $\bar{\eta}$ distribution.

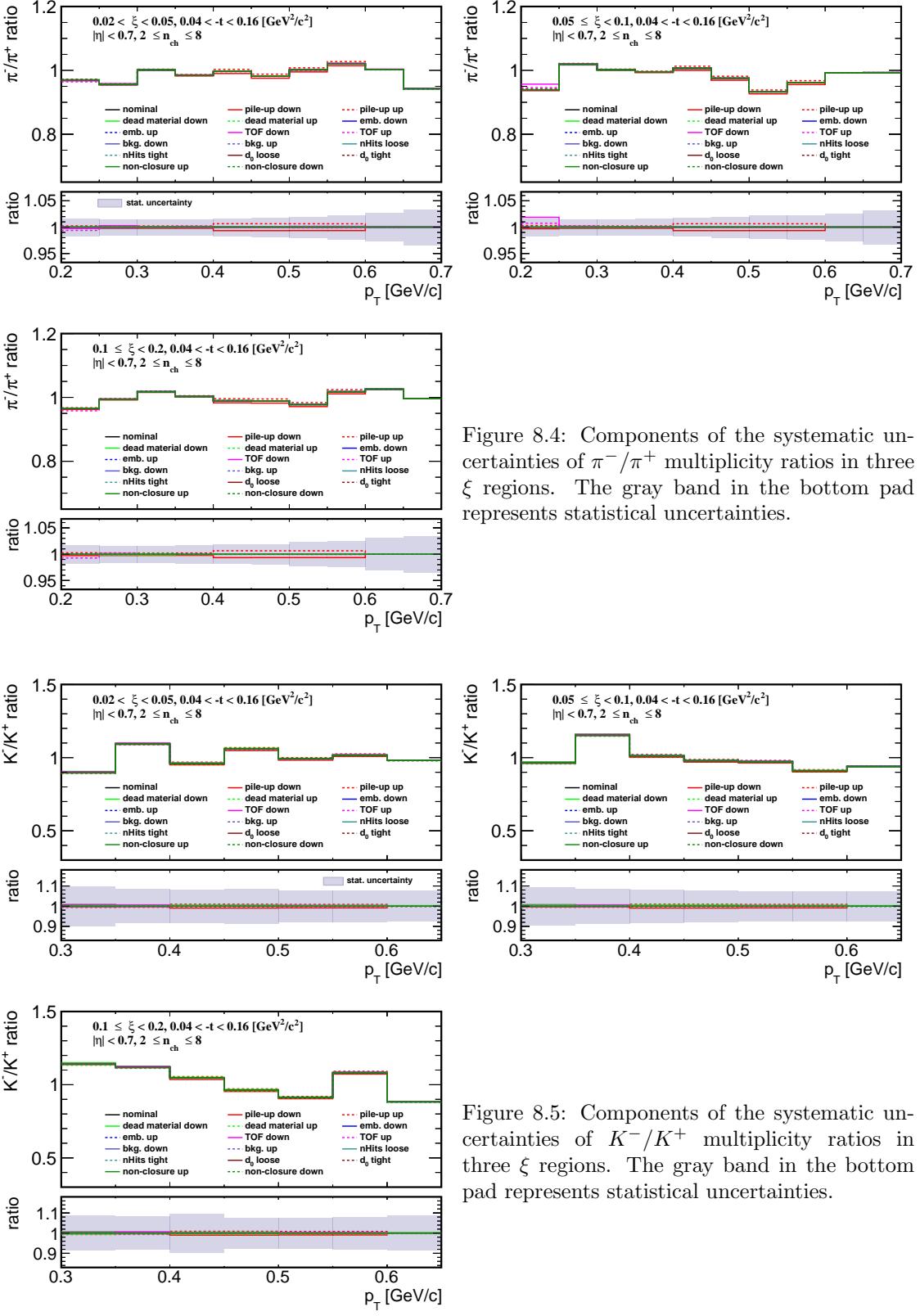


Figure 8.4: Components of the systematic uncertainties of π^-/π^+ multiplicity ratios in three ξ regions. The gray band in the bottom pad represents statistical uncertainties.

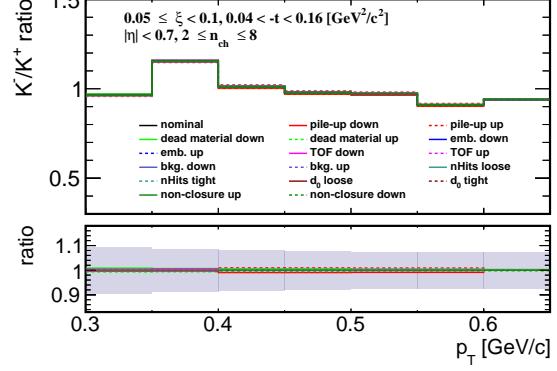


Figure 8.5: Components of the systematic uncertainties of K^-/K^+ multiplicity ratios in three ξ regions. The gray band in the bottom pad represents statistical uncertainties.

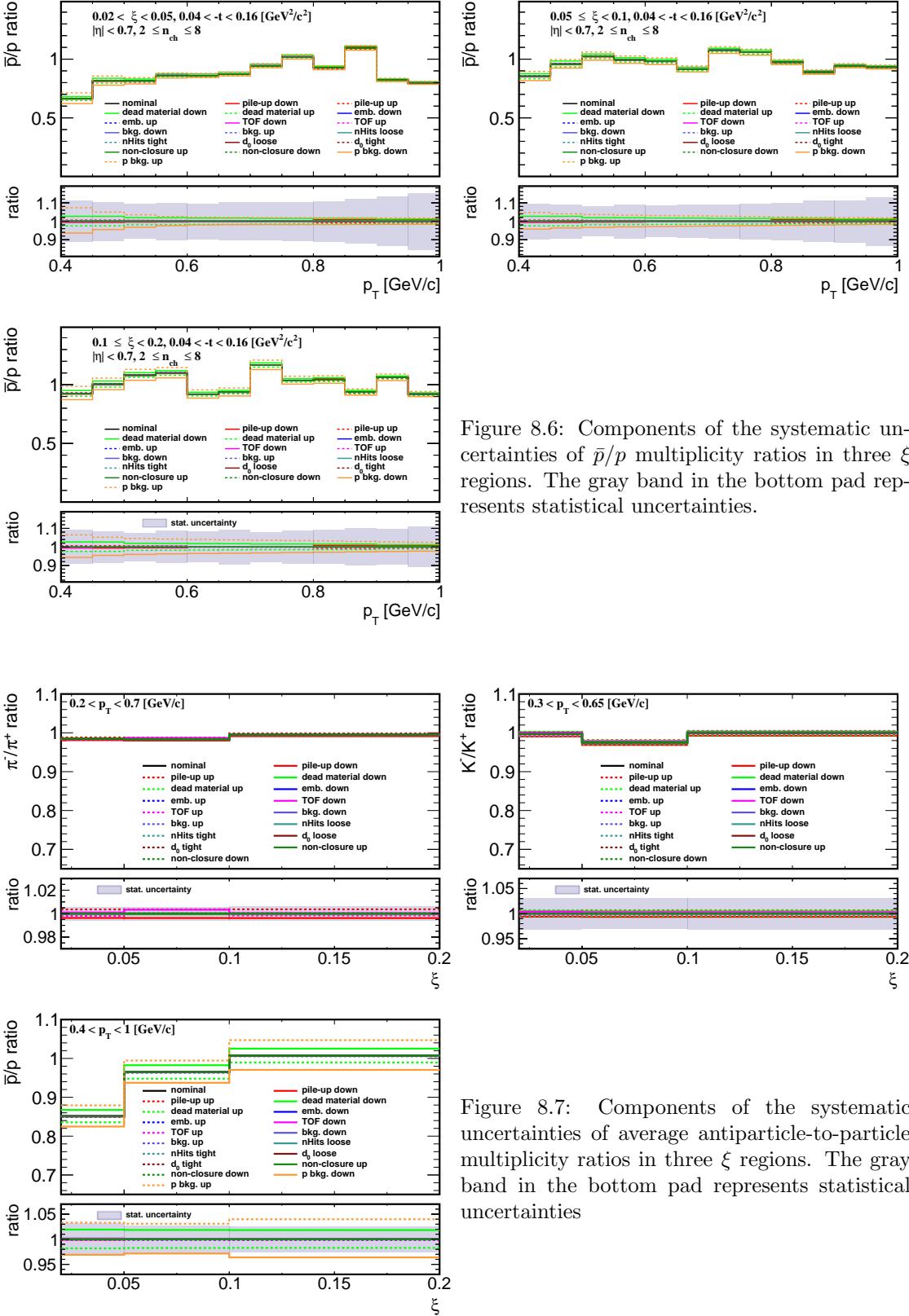


Figure 8.6: Components of the systematic uncertainties of \bar{p}/p multiplicity ratios in three ξ regions. The gray band in the bottom pad represents statistical uncertainties.

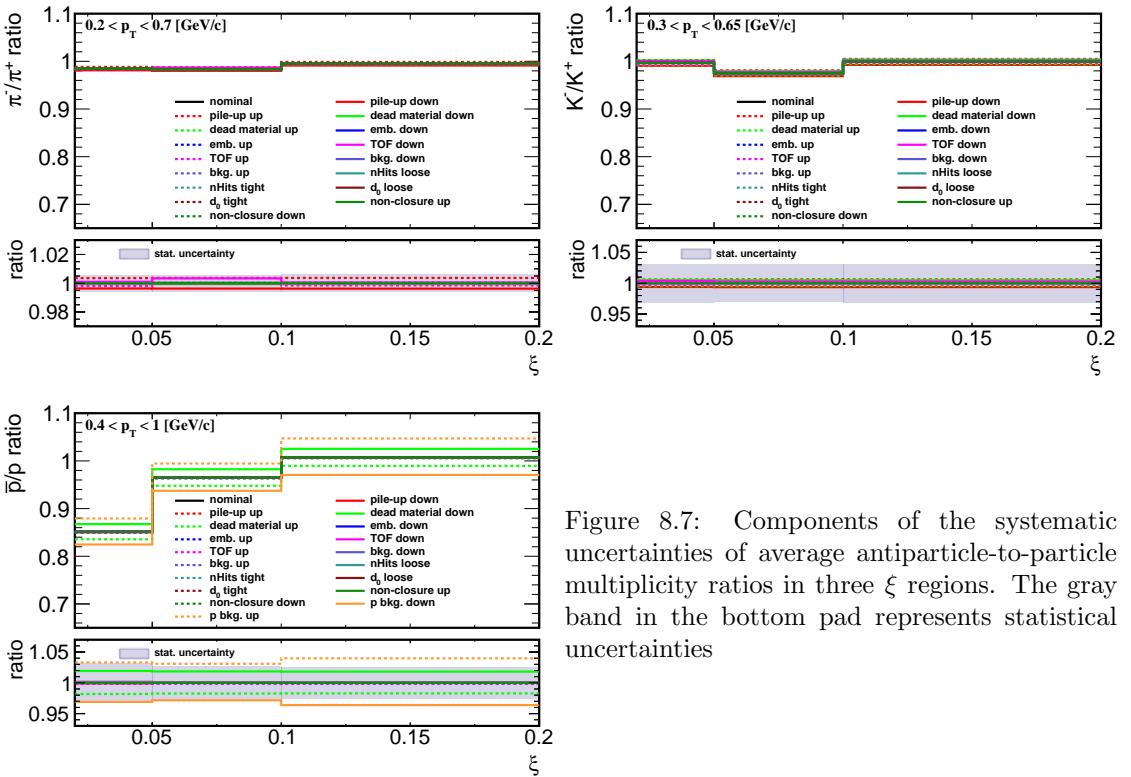


Figure 8.7: Components of the systematic uncertainties of average antiparticle-to-particle multiplicity ratios in three ξ regions. The gray band in the bottom pad represents statistical uncertainties

905 Figures 8.8 and 8.9 show breakdown of all different systematics for the $(K^- + K^+) / (\pi^- + \pi^+)$
 906 multiplicity ratio distributions. The largest sources of systematics are related to TOF reconstruc-
 907 tion efficiencies and dead-material effect on TPC track reconstruction efficiency. The former was
 found to be up to 4%, whereas the latter varies up to 5% for low- p_T multiplicity ratio.

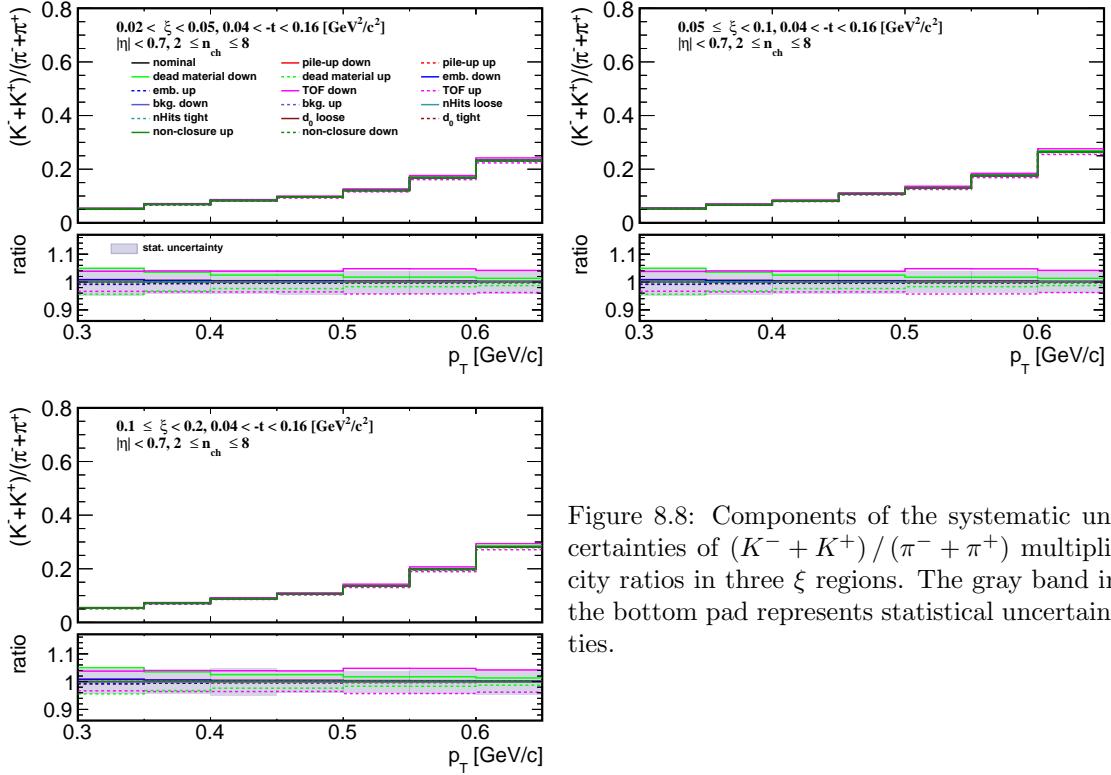


Figure 8.8: Components of the systematic uncertainties of $(K^- + K^+) / (\pi^- + \pi^+)$ multiplicity ratios in three ξ regions. The gray band in the bottom pad represents statistical uncertainties.

908

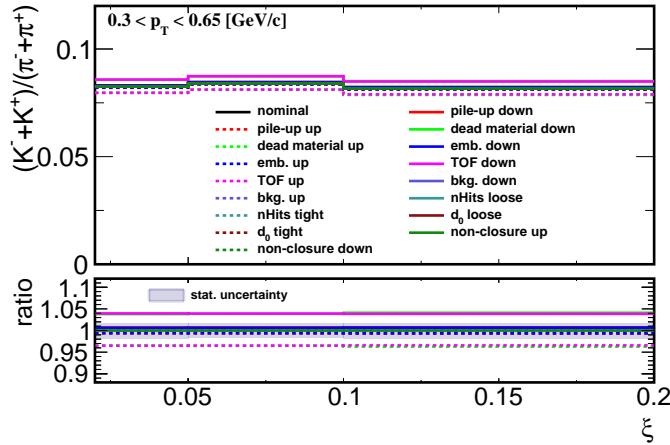


Figure 8.9: Components of the systematic uncertainties of average $(K^- + K^+) / (\pi^- + \pi^+)$ multiplicity ratios in three ξ regions. The gray band in the bottom pad represents statistical uncertainties

9. Results

In the following section, the final-state charged particle distributions are compared with various SD MC predictions, i.e.

- PYTHIA 8 4C (SaS),
- PYTHIA 8 A2 (MBR),
- PYTHIA 8 A2 (MBR-tuned),
- HERWIG 7,
- EPOS LHC with combined two classes of processes: diffractive (EPOS SD) and non-diffractive (EPOS SD'),
- EPOS LHC SD'.

In all figures, data are shown as solid points with error bars representing the statistical uncertainties. Gray boxes represent statistical and systematic uncertainties added in quadrature. Predictions from MC models are shown as colour histograms and markers. The lower panel in each figure shows the ratio of data to the models' predictions. All results are presented separately for three ranges of ξ : $0.02 < \xi < 0.05$, $0.05 < \xi < 0.1$, $0.1 < \xi < 0.2$.

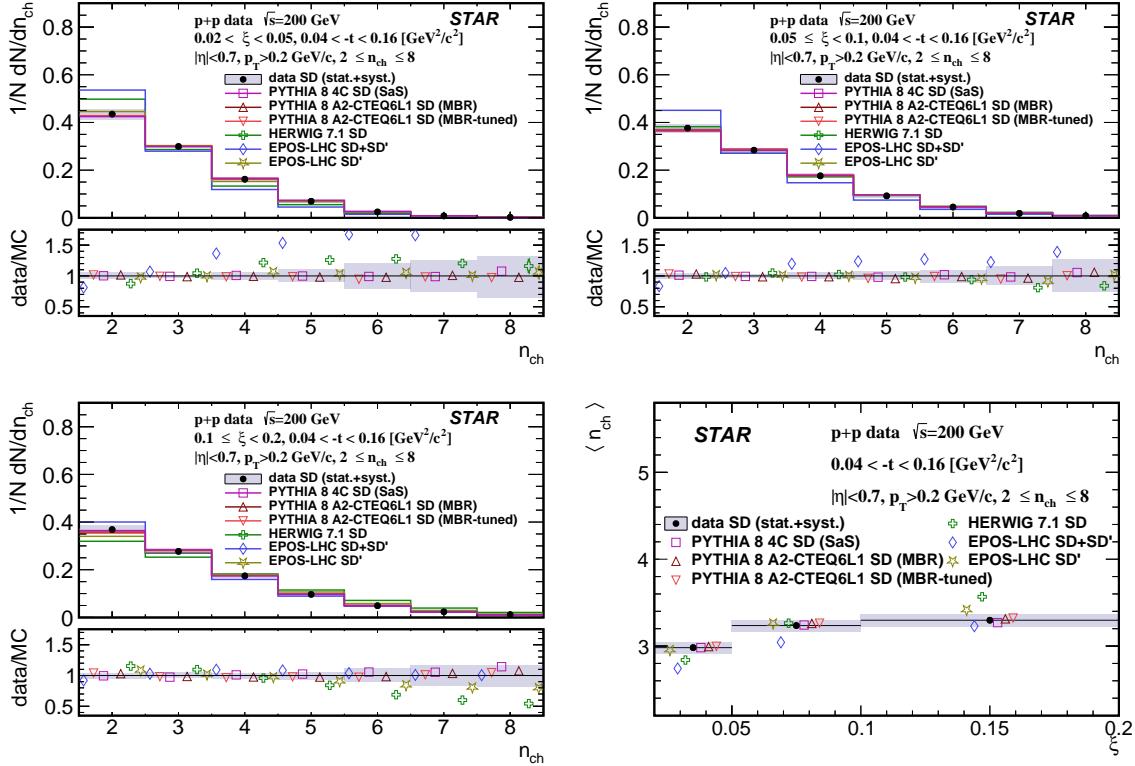


Figure 9.1: Primary charged-particle multiplicity shown separately for the three ranges of ξ : (top left) $0.02 < \xi < 0.05$, (top right) $0.05 < \xi < 0.1$, (bottom left) $0.1 < \xi < 0.2$ and (bottom right) the mean multiplicity $\langle n_{ch} \rangle$ as a function of ξ .

924 In Fig. 9.1 we show multiplicity distributions of charged particles in different intervals of ξ as
 925 well as the average values of n_{ch} in those ξ ranges. Data exhibit an expected increase of the $\langle n_{ch} \rangle$
 926 with ξ due to the larger diffractive masses probed at increasing ξ in SD process. The shapes of the
 927 measured distributions are reproduced reasonably well by all models except the EPOS SD+SD',
 928 which predicts much smaller $\langle n_{ch} \rangle$ at $\xi < 0.1$ and the HERWIG SD, which for $0.1 < \xi < 0.2$
 929 predicts too large $\langle n_{ch} \rangle$. It should be noted, that EPOS SD' describes data much better compared
 930 to EPOS SD+SD'.
 931

Figure 9.2 shows densities of charged-particles as a function of transverse momentum p_T in
 different intervals of ξ and for average values of p_T in those ξ ranges. Data show that $\langle p_T \rangle$ depends
 very weakly on ξ . MC models describe data fairly well predicting $\langle p_T \rangle$ only 0.01 GeV/c higher
 than in data except for the HERWIG SD, which predicts much steeper dependence of particle
 density on p_T in all three ξ ranges.
 932

Figure 9.3 shows densities of charged-particles as a function of $\bar{\eta}$ in different intervals of ξ and
 average values of $\bar{\eta}$ in those ξ ranges. Data show expected flattening of the $\bar{\eta}$ distribution with
 increasing ξ , which reflects SD event-asymmetry and the fact that the position of midrapidity η_m
 at large ξ is closer to the fiducial $|\eta| < 0.7$ region, leading to more flat distribution of particle
 density as a function of $\bar{\eta}$. Models describe data fairly well except EPOS SD+SD', which predicts
 too flat $\bar{\eta}$ distributions in all three ξ ranges and the HERWIG SD, which predicts too steep $\bar{\eta}$
 distributions. Similarly to n_{ch} distributions EPOS SD' describes data better compared to EPOS
 SD+SD'.
 933

In Fig. 9.4 we present the ratios of production yields of π^-/π^+ in three intervals of ξ as
 944 a function of p_T . Data in all three ξ ranges are consistent with equal amounts of π^+ and π^-
 945 with no p_T dependence. MC models agree with data (except HERWIG SD) predicting on average
 946

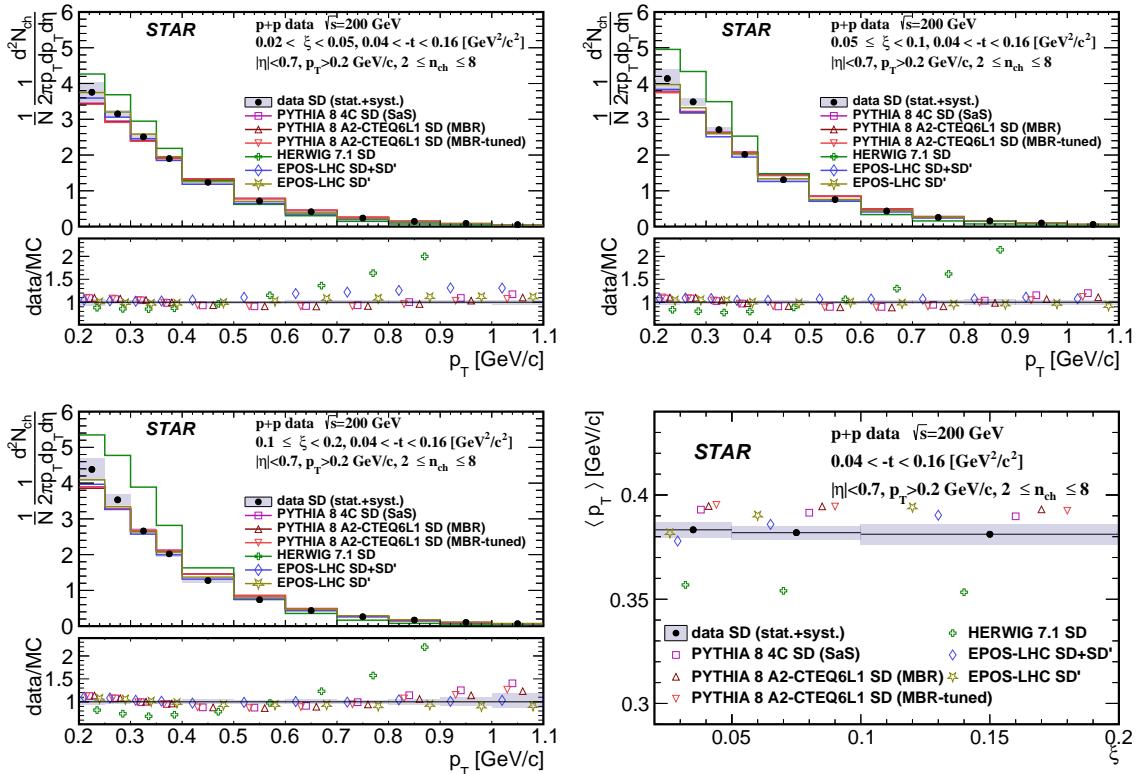


Figure 9.2: Primary charged-particle multiplicities as a function of p_T shown separately for the three ranges of ξ : (top left) $0.02 < \xi < 0.05$, (top right) $0.05 < \xi < 0.1$, (bottom left) $0.1 < \xi < 0.2$ and (bottom right) the mean transverse momentum $\langle p_T \rangle$ as a function of ξ .

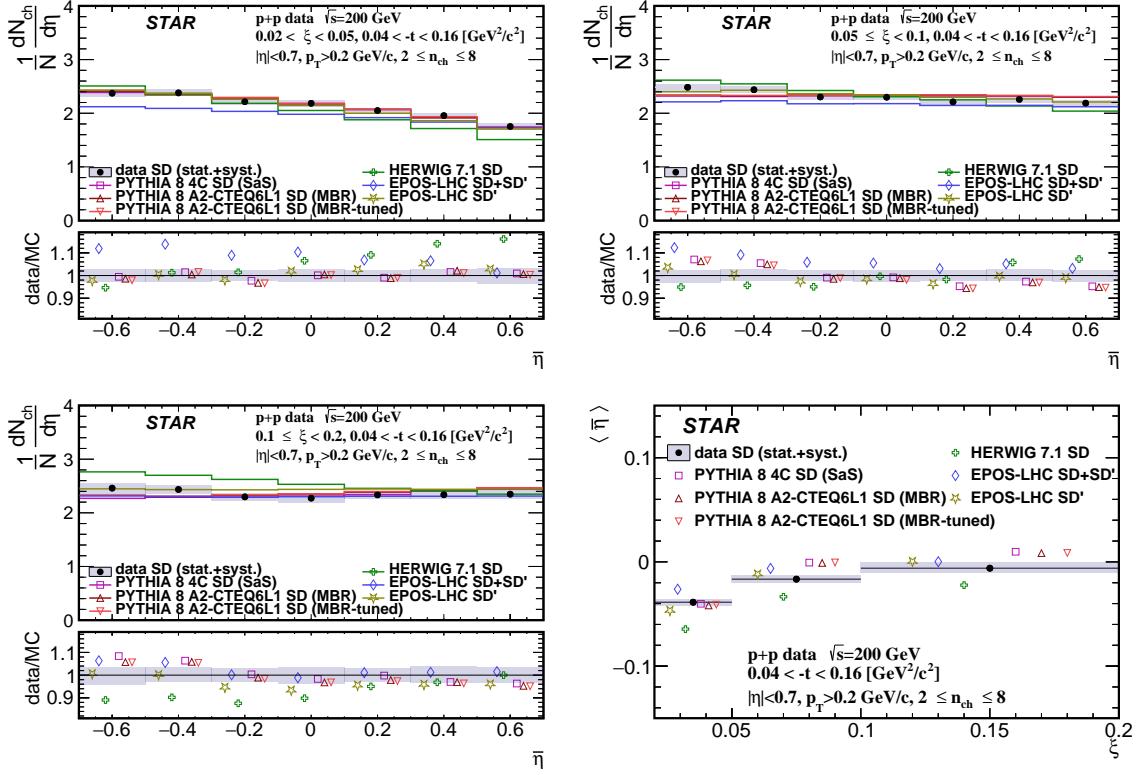


Figure 9.3: Primary charged-particle multiplicity as a function of $\bar{\eta}$ shown separately for the three ranges of ξ : (top left) $0.02 < \xi < 0.05$, (top right) $0.05 < \xi < 0.1$, (bottom left) $0.1 < \xi < 0.2$ and (bottom right) the mean pseudorapidity $\langle \bar{\eta} \rangle$ as a function of ξ .

947 a small deviation from unity by $\sim 2\%$, which is smaller than the data uncertainties. The model
 948 implemented in HERWIG SD, in the first two ξ ranges predicts too large asymmetry between π^+
 949 and π^- .

950 Figure 9.5 shows the ratios of production yields of K^-/K^+ in three intervals of ξ as a function
 951 of p_T . Data in all three ξ ranges are consistent with equal amounts of K^+ and K^- with no p_T
 952 dependence. Models agree with the data except for HERWIG SD in first ξ range, which predicts
 953 too large ratio of K^- to K^+ .

954 Figure 9.6 shows the ratios of production yields of \bar{p}/p in three intervals of ξ as a function
 955 of p_T . In the last two ξ ranges, data are consistent with equal amounts of p and \bar{p} with no p_T
 956 dependence. However, in the first ξ range at $p_T < 0.7$ GeV, data show a significant deviation from
 957 unity indicating a large transfer of the baryon number from the forward to the central region.
 958 PYTHIA 8, EPOS SD' and EPOS SD+SD' agree with data in the last two ξ ranges. In the first
 959 ξ range, they predict small deviation from unity by $\sim 5\%$, which is smaller than $\sim 15\%$ observed
 960 in data. HERWIG SD predicts much larger baryon number transfer compared to data in all three
 961 ξ ranges. This observation is caused by the fact, that HERWIG SD is effectively an extreme
 962 realization of the model in which there is almost always a baryon present at gap edge. Significant
 963 increase of \bar{p}/p ratio with increasing ξ is related to the fact that the gap edge is further away from
 964 the fiducial η region at larger ξ . EPOS SD+SD' predicts slightly higher baryon number transfer
 965 compared to data in first ξ range while EPOS SD' slightly smaller. This suggests, that \bar{p}/p ratio
 966 might be a good observable for tuning relative fractions of SD and SD' processes in EPOS.

967 Figure 9.7 shows the ratios of production yields of π^-/π^+ , K^-/K^+ and \bar{p}/p integrated over
 968 fiducial region in p_T as a function of ξ . The results with increased statistical precision confirm
 969 conclusions from ratios calculated differentially in p_T .

970 Figure 9.8 shows the ratios of production yields of $(K^- + K^+)/(\pi^- + \pi^+)$ in three intervals
 971 of ξ as a function of p_T . The ratio increases from 0.05 at $p_T = 0.3$ GeV to 0.22 – 0.28 at
 972 $p_T = 0.65$ GeV. The slope of the p_T dependence significantly increases at $p_T = 0.5$ GeV in
 973 all three ξ intervals. The change of the p_T slope increases with ξ . All models predict very
 974 similar $(K^+ + K^-)/(\pi^+ + \pi^-)$ ratio except HERWIG, which predicts almost twice larger value
 975 independently from p_T . PYTHIA 8 and EPOS agree very well with data at $0.3 < p_T < 0.5$ GeV
 976 but do not expect a change of the slope of p_T dependence at $p_T > 0.5$ GeV predicting rather
 977 almost twice smaller ratio at the highest p_T value.

978 Figure 9.9 shows the ratio of the production yields of $(K^- + K^+)/(\pi^- + \pi^+)$ as a function of
 979 ξ integrated over $0.3 < p_T < 0.65$ GeV. We do not observe any dependence on ξ . PYTHIA 8 and
 980 EPOS predictions agree with the data. HERWIG predicts much higher $(K^- + K^+)/(\pi^- + \pi^+)$
 981 ratio.

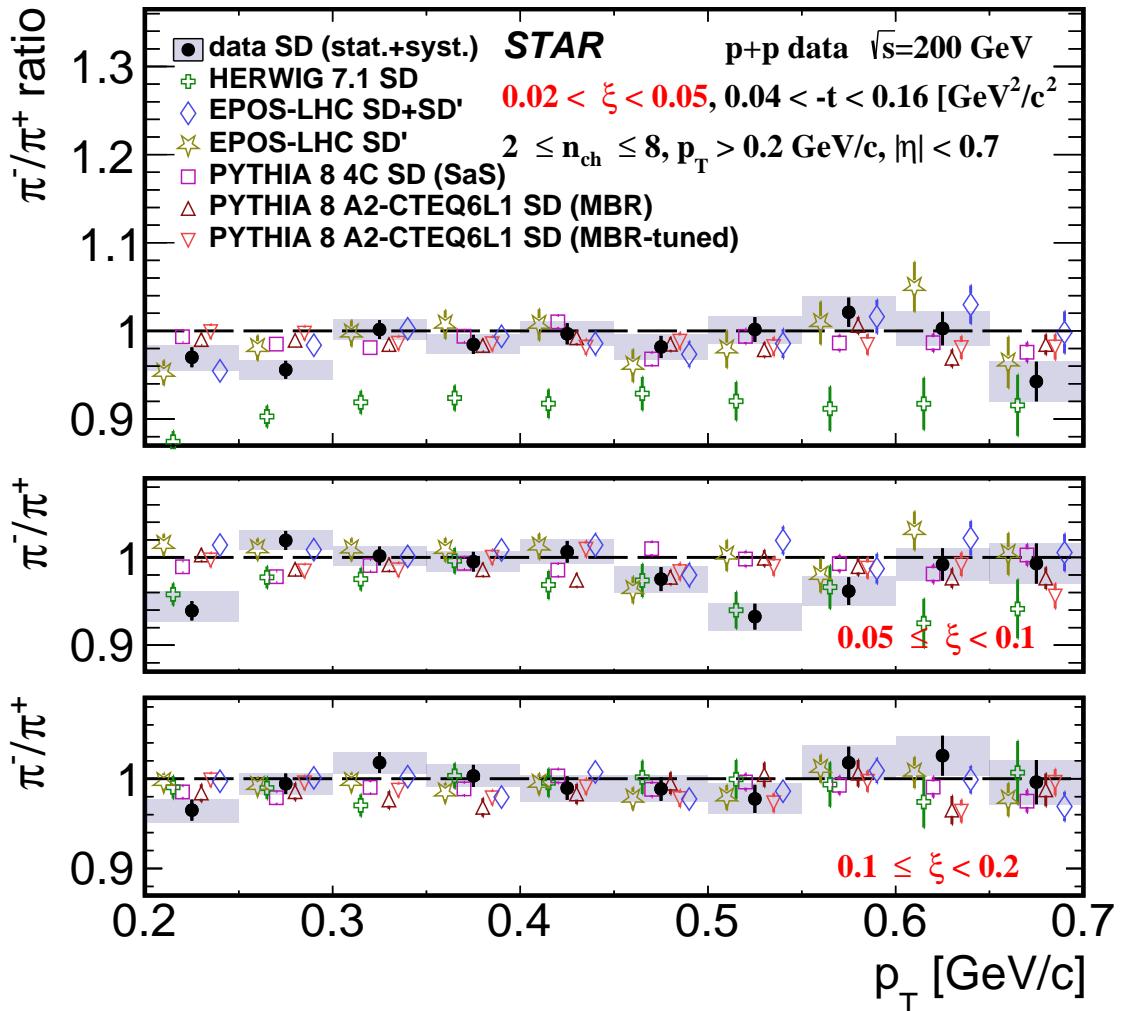


Figure 9.4: Ratio of production yields of π^-/π^+ as a function of p_T shown separately for the three ranges of ξ : (top) $0.02 < \xi < 0.05$, (middle) $0.05 < \xi < 0.1$, (bottom) $0.1 < \xi < 0.2$.

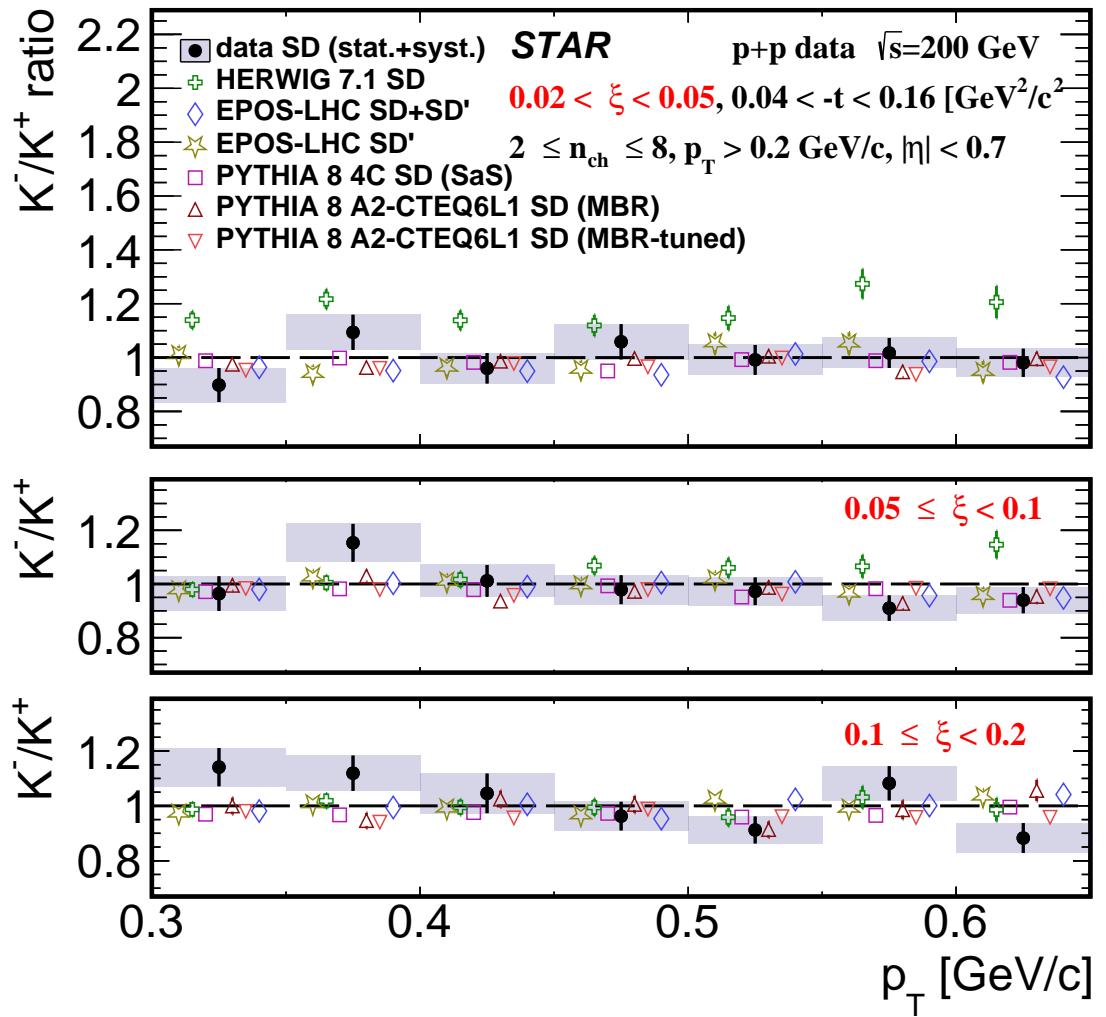


Figure 9.5: Ratio of production yields of K^-/K^+ as a function of p_{T} shown separately for the three ranges of ξ : (top) $0.02 < \xi < 0.05$, (middle) $0.05 < \xi < 0.1$, (bottom) $0.1 < \xi < 0.2$.

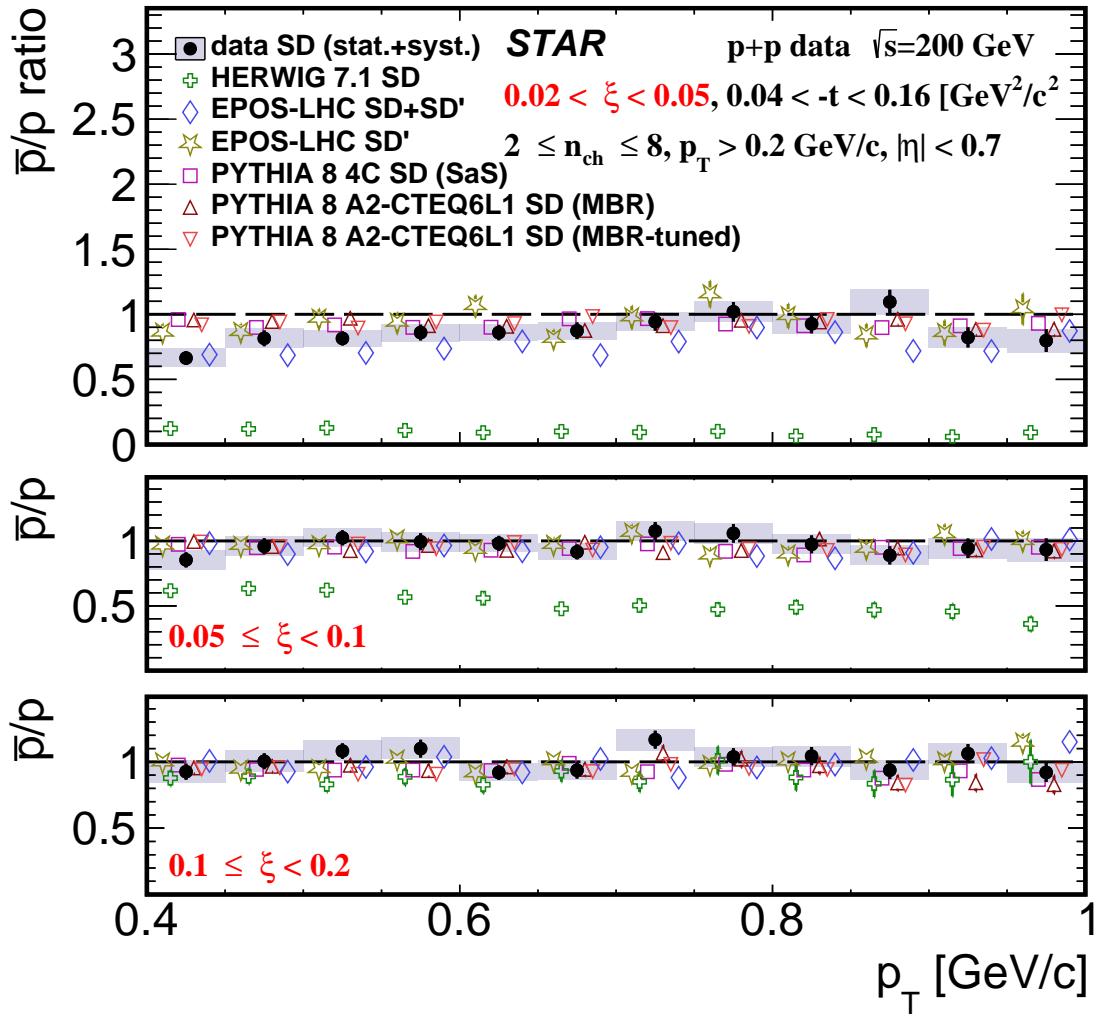


Figure 9.6: Ratio of production yields of \bar{p}/p as a function of p_{T} shown separately for the three ranges of ξ : (top) $0.02 < \xi < 0.05$, (middle) $0.05 < \xi < 0.1$, (bottom) $0.1 < \xi < 0.2$.

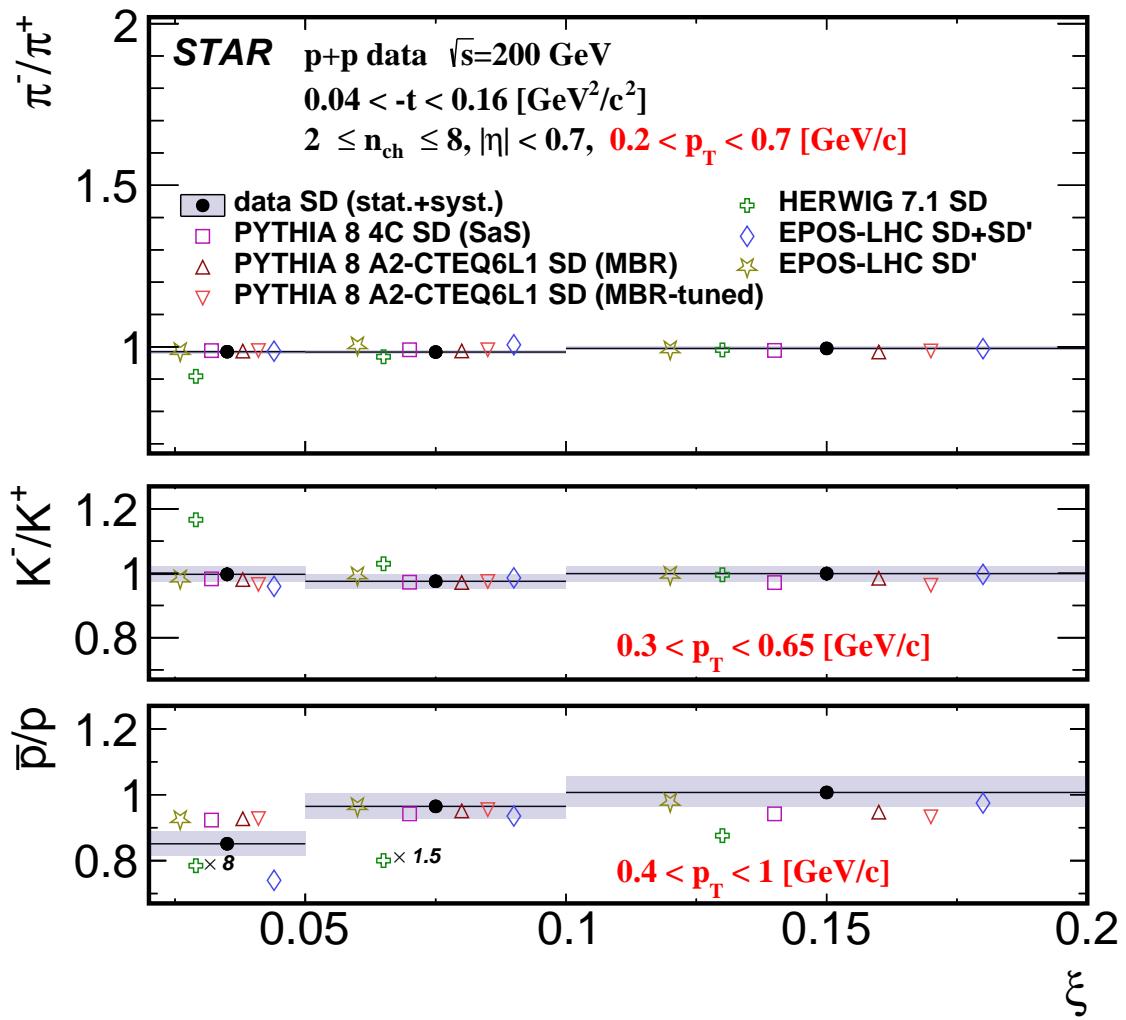


Figure 9.7: Ratio of production yields of π^-/π^+ , K^-/K^+ and \bar{p}/p as a function of ξ .

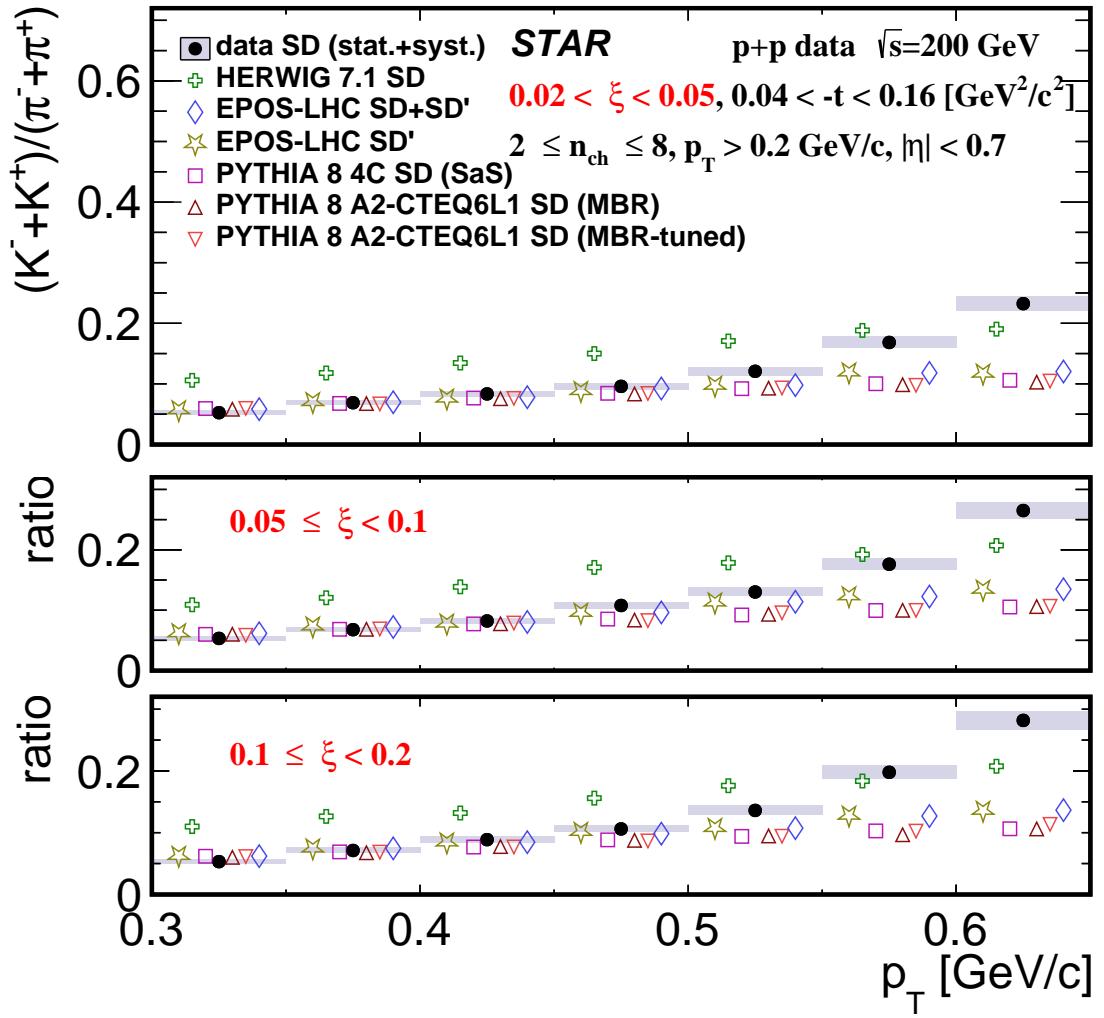


Figure 9.8: Ratio of production yields of $(K^- + K^+)/(\pi^- + \pi^+)$ as a function of p_{T} shown separately for the three ranges of ξ : (top) $0.02 < \xi < 0.05$, (middle) $0.05 \leq \xi < 0.1$, (bottom) $0.1 \leq \xi < 0.2$.

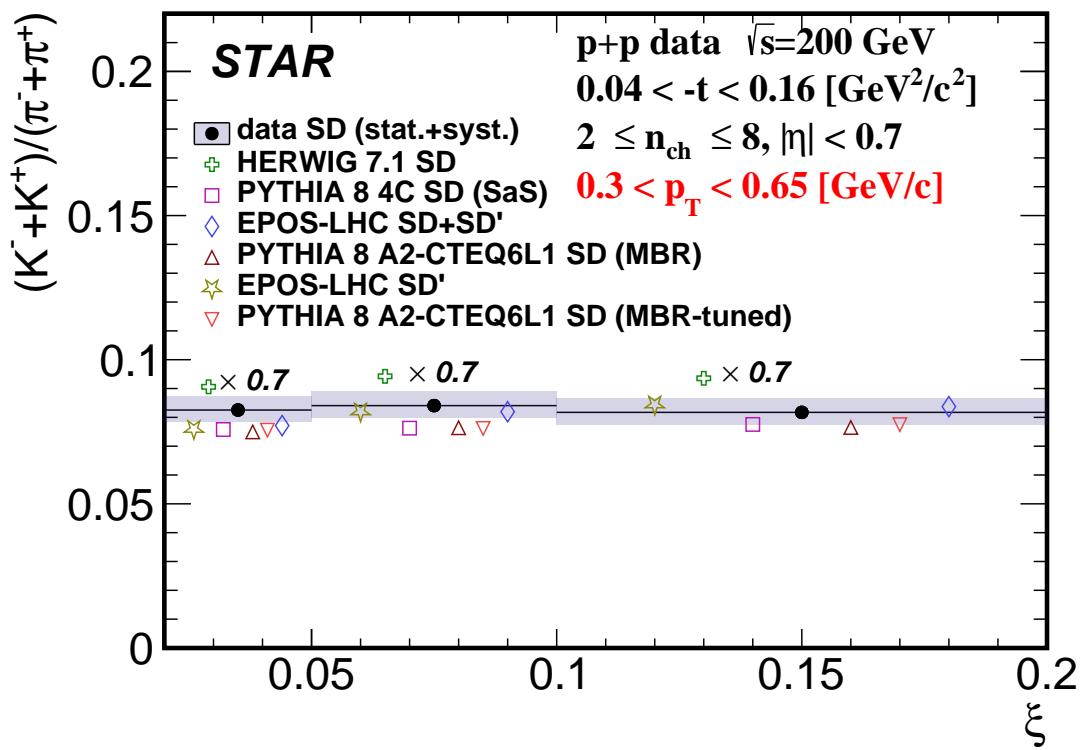


Figure 9.9: Ratio of production yields of $(K^- + K^+) / (\pi^- + \pi^+)$ as a function of ξ .

9.1 Comparison of Charged-Particle Densities at Central Rapidities

The measured charged-particle densities in pseudorapidity near $\eta \approx 0$ are compared to other experimental results from pp and $p\bar{p}$ collisions (shown in Fig. 9.10). Various event selections, based on the topology of the final state, allow to split the data samples into enhanced in non-SD (NSD) and SD events, whose sum forms the inelastic sample. For SD events, the midrapidity region is located at $\bar{\eta} = \eta_m = -\ln(\sqrt{s}/M_X)$, instead of $\eta \approx 0$, and the proper energy scale is given by M_X instead of \sqrt{s} . The values of η_m and $\langle M_X \rangle$, calculated for STAR data, are presented in Tab. 9.1. For all three ranges of ξ , the value of η_m is outside the fiducial region of the measurement. In the case of other experiments the pseudorapidity densities were obtained in the region of the total number of primary charged particles $n_{ch} \geq 1$ (instead of $n_{ch} \geq 2$ as in this analysis). Therefore, the results from STAR analysis were extrapolated to the above fiducial region using PYTHIA 8 A2 (MBR) SD predictions. The uncertainties due to the corrections are not estimated.

ξ range	$\langle M_X \rangle$	η_m	η_{edge}
$0.02 < \xi < 0.05$	37.53 GeV	-1.67	2.02
$0.05 < \xi < 0.1$	53.52 GeV	-1.31	2.73
$0.1 < \xi < 0.2$	72.71 GeV	-1.01	3.34

Table 9.1: Values of $\langle M_X \rangle$ and $\eta_m = \ln(\sqrt{s}/M_X)$ for three ranges of ξ and position of gap edge η_{edge} .

The extrapolation procedure was as follows:

- the ratio of particle density at $\bar{\eta} = \eta_m$ and $n_{ch} \geq 1$ to that at $\bar{\eta} = 0$ and $n_{ch} \geq 2$ was calculated using PYTHIA 8 predictions,
- differences in the slope of the pseudorapidity distribution in the region of $n_{ch} \geq 2$ were observed between data and MC. Therefore, data and MC distributions were normalized to have the same particle density at $\bar{\eta} \approx 0$ and their ratio was fitted with a linear function,
- the correction from step #1, multiplied by the value of the above function at $\bar{\eta} = \eta_m$, was used to scale the measured particle density at $\bar{\eta} \approx 0$.

Figure 9.10 presents the charged-particle densities near $\eta \approx 0$ as a function of \sqrt{s} in inelastic pp and $p\bar{p}$ collisions. The SD results, calculated near $\bar{\eta} = \eta_m$ at \sqrt{s} (M_X), are also shown.

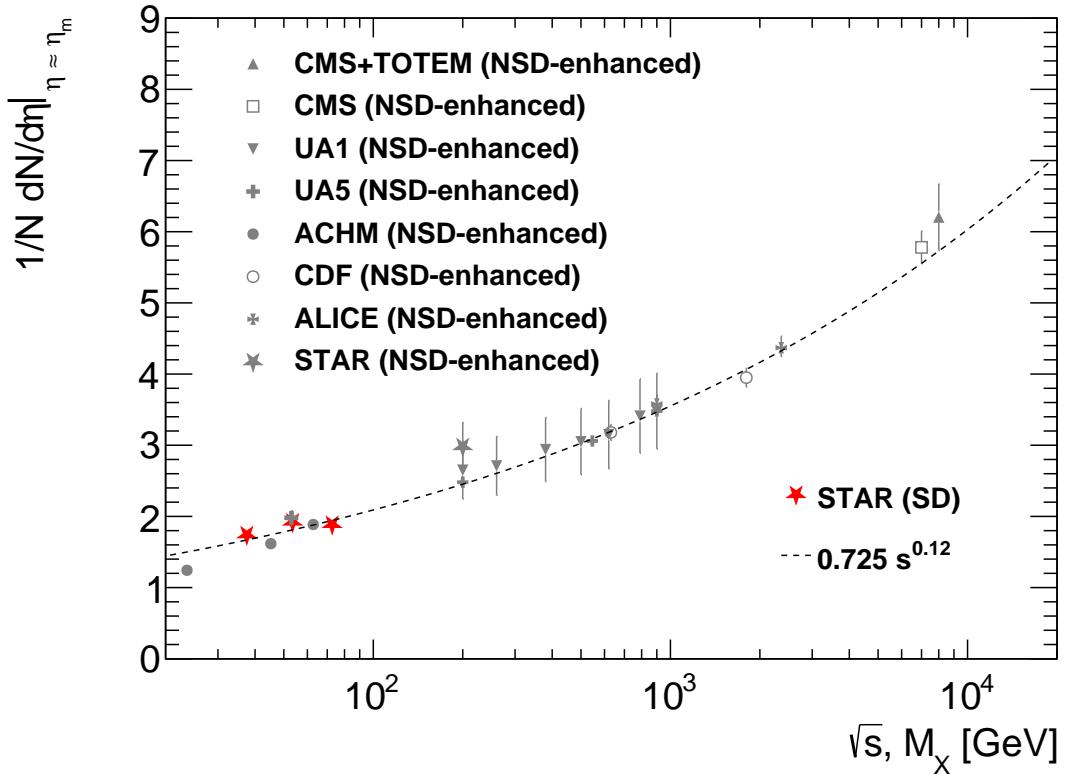


Figure 9.10: Primary charged particle densities at mid-rapidity as a function of \sqrt{s} and M_X for inelastic (non single diffractive (NSD) enhanced) [12, 17, 18, 19, 20, 21, 22, 23] and SD measurements respectively. The dashed line represents power-law fit to the NSD-enhanced measurements [17] excluding ACHM measurements [20] at ISR. The results from this analysis are shown in red.

1005 10. Summary and Conclusions

1006 Inclusive and identified (pion, kaon, proton and their antiparticles) charged-particle production in
1007 Single Diffractive process has been measured in proton-proton collisions at $\sqrt{s} = 200$ GeV with
1008 the STAR detector at RHIC using data corresponding to an integrated luminosity of 15 nb^{-1} .

1009 Significant differences are observed between the measured distributions of ξ and Monte Carlo
1010 model predictions. Among the models considered, EPOS and PYTHIA 8 (MBR) without sup-
1011 pression of diffractive cross sections at large ξ provide the best description of the data. Primary-
1012 charged-particle multiplicities and their densities as functions of pseudorapidity and transverse
1013 momentum are well described by PYTHIA 8 and EPOS SD' models. EPOS SD and HERWIG
1014 do not describe the data. Similarity between the dissociation of a diffractively produced system
1015 of mass M_X and the hadronization of the system resulting from non-diffractive pp collisions at
1016 $\sqrt{s} \approx M_X$ reported for the first time by UA4 Collaboration [24, 25, 26] was confirmed with much
1017 better precision.

1018 π^-/π^+ and K^-/K^+ production ratios are close to unity and consistent with most of model
1019 predictions except for EPOS SD and HERWIG. \bar{p}/p production ratio shows a significant deviation
1020 from unity in the $0.02 < \xi < 0.05$ range indicating a non-negligible transfer of the baryon number
1021 from the forward to the central region. Equal amounts of protons and antiprotons are observed in
1022 the $\xi > 0.05$ range. PYTHIA 8 and EPOS SD' agree with data for $\xi > 0.05$. For $0.02 < \xi < 0.05$
1023 they predict small deviations from unity (0.93), however even larger effect is observed in the data
1024 (0.86 ± 0.02). This observation is consistent with increase of the baryon number transfer to the
1025 central rapidity region with decreasing ξ expected from [3] where an extra baryon can appear close
1026 to the rapidity gap edge (so called backward peak).

1027 At $p_T > 0.5$ GeV measured $(K^- + K^+)/(\pi^- + \pi^+)$ ratio is significantly larger compared to
1028 inclusive inelastic measurements in pp or $\bar{p}p$ collisions. This excess is not predicted by any model.
1029 This enhancement can be due to the high production rate of $gg \rightarrow \bar{s}s$ in the diffraction since
1030 Pomeron is expected to be dominated by gluonic content.

1031 Bibliography

- 1032 [1] L. Adamczyk, L. Fulek, and R. Sikora, *Supplementary note on diffractive analyses of 2015*
1033 *proton-proton data*, <https://drupal.star.bnl.gov/STAR/starnotes/private/psn0732>,
1034 July, 2019.
- 1035 [2] B. Z. Kopeliovich and B. G. Zakharov, *Novel Mechanisms of Baryon Number Flow Over*
1036 *Large Rapidity Gap*, *Z. Phys. C* **43** (1989) 241.
- 1037 [3] F. W. Bopp, *Baryon transport in dual models and the possibility of a backward peak in*
1038 *diffraction*, in *New trends in high-energy physics: Experiment, phenomenology, theory.*
1039 *Proceedings, International School-Conference, Crimea 2000, Yalta, Ukraine, May 27-June 4,*
1040 *2000*. arXiv:hep-ph/0007229 [hep-ph].
- 1041 [4] M. Anderson et al., *The STAR Time Projection Chamber: A Unique tool for studying high*
1042 *multiplicity events at RHIC*, *Nucl. Instrum. Meth. A* **499** (2003) 659–678,
1043 arXiv:nucl-ex/0301015 [nucl-ex].
- 1044 [5] T. Sjöstrand, S. Mrenna, and P. Z. Skands, *PYTHIA 6.4 Physics and Manual*, *JHEP* **05**
1045 (2006) 026.
- 1046 [6] R. Ciesielski and K. Goulianos, *MBR Monte Carlo Simulation in PYTHIA8*, *PoS*
1047 *ICHEP2012* (2013) 301, arXiv:1205.1446 [hep-ph].
- 1048 [7] *Roman Pot Phase II* - online DAQ monitoring*,
1049 <https://online.star.bnl.gov/rp/pp200/>.
- 1050 [8] *STAR RunLog Browser*, <https://online.star.bnl.gov/RunLogRun15/>.
- 1051 [9] W. Fischer, *Run Overview of the Relativistic Heavy Ion Collider*,
1052 <http://www.agsrhichome.bnl.gov/RHIC/Runs/>.
- 1053 [10] *STAR Luminosity Tracker*, <https://www.star.bnl.gov/protected/common/common2015/trigger2015/runningStatepp200GeV/plots.html>.
- 1055 [11] STAR Collaboration, J. Adam et al., *Results on Total and Elastic Cross Sections in*
1056 *Proton-Proton Collisions at $\sqrt{s} = 200$ GeV*, arXiv:2003.12136 [hep-ex].
- 1057 [12] STAR Collaboration, B. I. Abelev et al., *Systematic Measurements of Identified Particle*
1058 *Spectra in pp, d⁺ Au and Au+Au Collisions from STAR*, *Phys. Rev. C* **79** (2009) 034909,
1059 arXiv:0808.2041 [nucl-ex].
- 1060 [13] CDF Collaboration, D. Acosta et al., *Inclusive double pomeron exchange at the Fermilab*
1061 *Tevatron $\bar{p}p$ collider*, *Phys. Rev. Lett.* **93** (2004) 141601, arXiv:hep-ex/0311023 [hep-ex].
- 1062 [14] ATLAS Collaboration, G. Aad et al., *Charged-particle distributions in $\sqrt{s} = 13$ TeV pp*
1063 *interactions measured with the ATLAS detector at the LHC*, *Phys. Lett. B* **758** (2016)
1064 67–88, arXiv:1602.01633 [hep-ex].

- 1065 [15] G. D'Agostini, *A Multidimensional unfolding method based on Bayes' theorem*, Nucl.
 1066 Instrum. Meth. **A362** (1995) 487–498.
- 1067 [16] H. Bichsel, *A method to improve tracking and particle identification in TPCs and silicon
 1068 detectors*, Nucl. Instrum. Meth. **A562** (2006) 154–197.
- 1069 [17] CMS, TOTEM Collaboration, S. Chatrchyan et al., *Measurement of pseudorapidity
 1070 distributions of charged particles in proton-proton collisions at $\sqrt{s} = 8$ TeV by the CMS and
 1071 TOTEM experiments*, Eur. Phys. J. **C74** (2014) 3053.
- 1072 [18] UA1 Collaboration, C. Albajar et al., *A Study of the General Characteristics of $p\bar{p}$
 1073 Collisions at $\sqrt{s} = 0.2$ -TeV to 0.9-TeV*, Nucl. Phys. **B335** (1990) 261–287.
- 1074 [19] UA5 Collaboration, G. J. Alner et al., *Scaling of Pseudorapidity Distributions at c.m.
 1075 Energies Up to 0.9-TeV*, Z. Phys. **C33** (1986) 1–6.
- 1076 [20] ACHM Collaboration, W. Thome et al., *Charged Particle Multiplicity Distributions in $p\ p$
 1077 Collisions at ISR Energies*, Nucl. Phys. **B129** (1977) 365.
- 1078 [21] CDF Collaboration, F. Abe et al., *Pseudorapidity distributions of charged particles produced
 1079 in $\bar{p}p$ interactions at $\sqrt{s} = 630$ GeV and 1800 GeV*, Phys. Rev. **D41** (1990) 2330.
- 1080 [22] ALICE Collaboration, K. Aamodt et al., *Charged-particle multiplicity measurement in
 1081 proton-proton collisions at $\sqrt{s} = 0.9$ and 2.36 TeV with ALICE at LHC*, Eur. Phys. J. **C68**
 1082 (2010) 89–108.
- 1083 [23] CMS Collaboration, V. Khachatryan et al., *Transverse-momentum and pseudorapidity
 1084 distributions of charged hadrons in pp collisions at $\sqrt{s} = 7$ TeV*, Phys. Rev. Lett. **105**
 1085 (2010) 022002, [arXiv:1005.3299 \[hep-ex\]](https://arxiv.org/abs/1005.3299).
- 1086 [24] UA4 Collaboration, M. Bozzo et al., *Single Diffraction Dissociation at the {CERN} {SPS} Collider*, Phys. Lett. B **136** (1984) 2170.
- 1088 [25] UA4 Collaboration, D. Bernard et al., *Pseudorapidity Distribution of Charged Particles in
 1089 Diffraction Dissociation Events at the CERN SPS Collider*, Phys. Lett. **B166** (1986) 459.
- 1090 [26] UA4 Collaboration, D. Bernard et al., *The Cross-section of Diffraction Dissociation at the
 1091 CERN SPS Collider*, Phys. Lett. B **186** (1987) 227.

Appendices

¹⁰⁹³ **A. Proton and Antiproton DCA
Distributions**

¹⁰⁹⁴

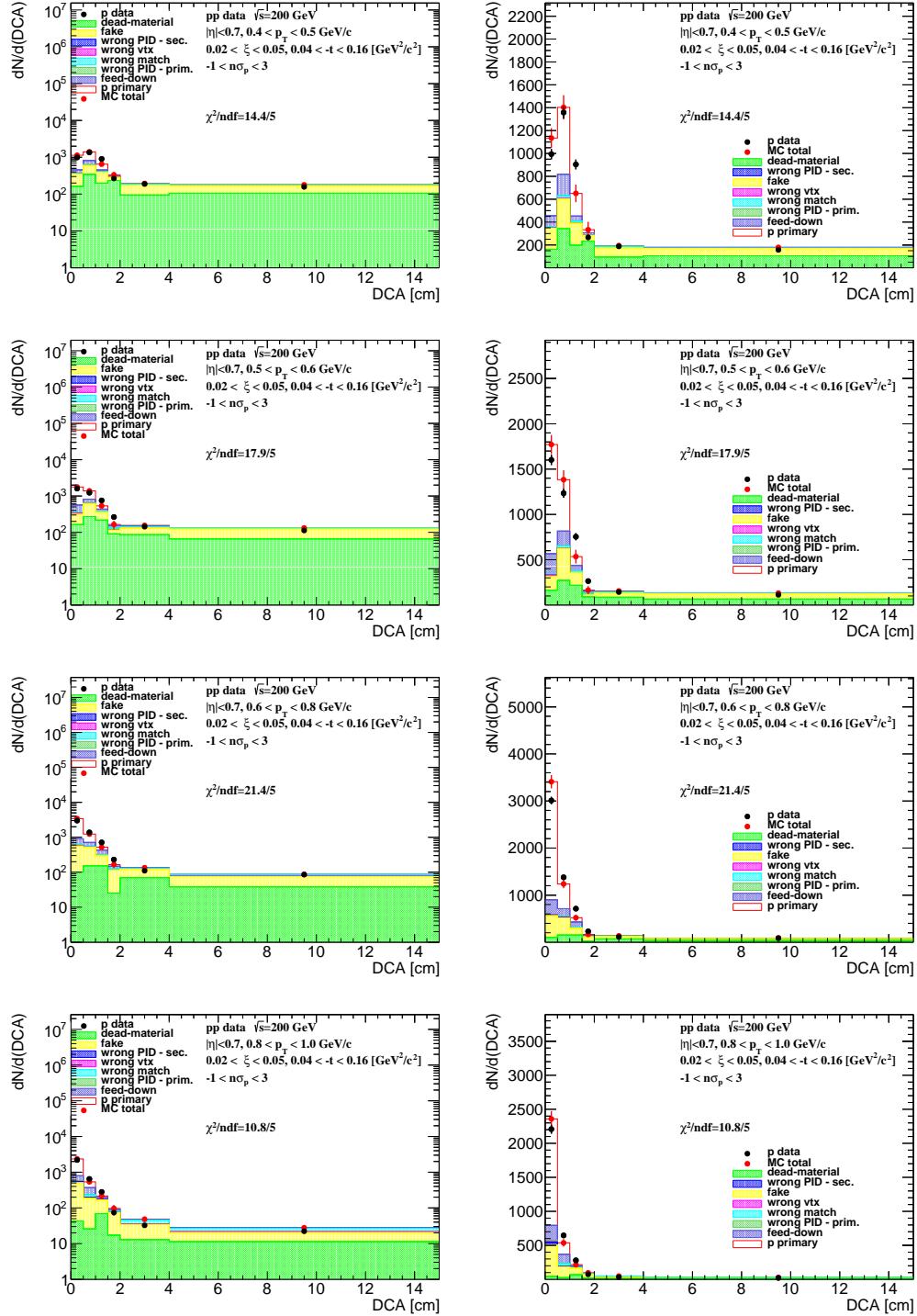


Figure A.1: Distributions of DCA for protons in SD interactions with $0.02 < \xi < 0.05$ and loose selection.

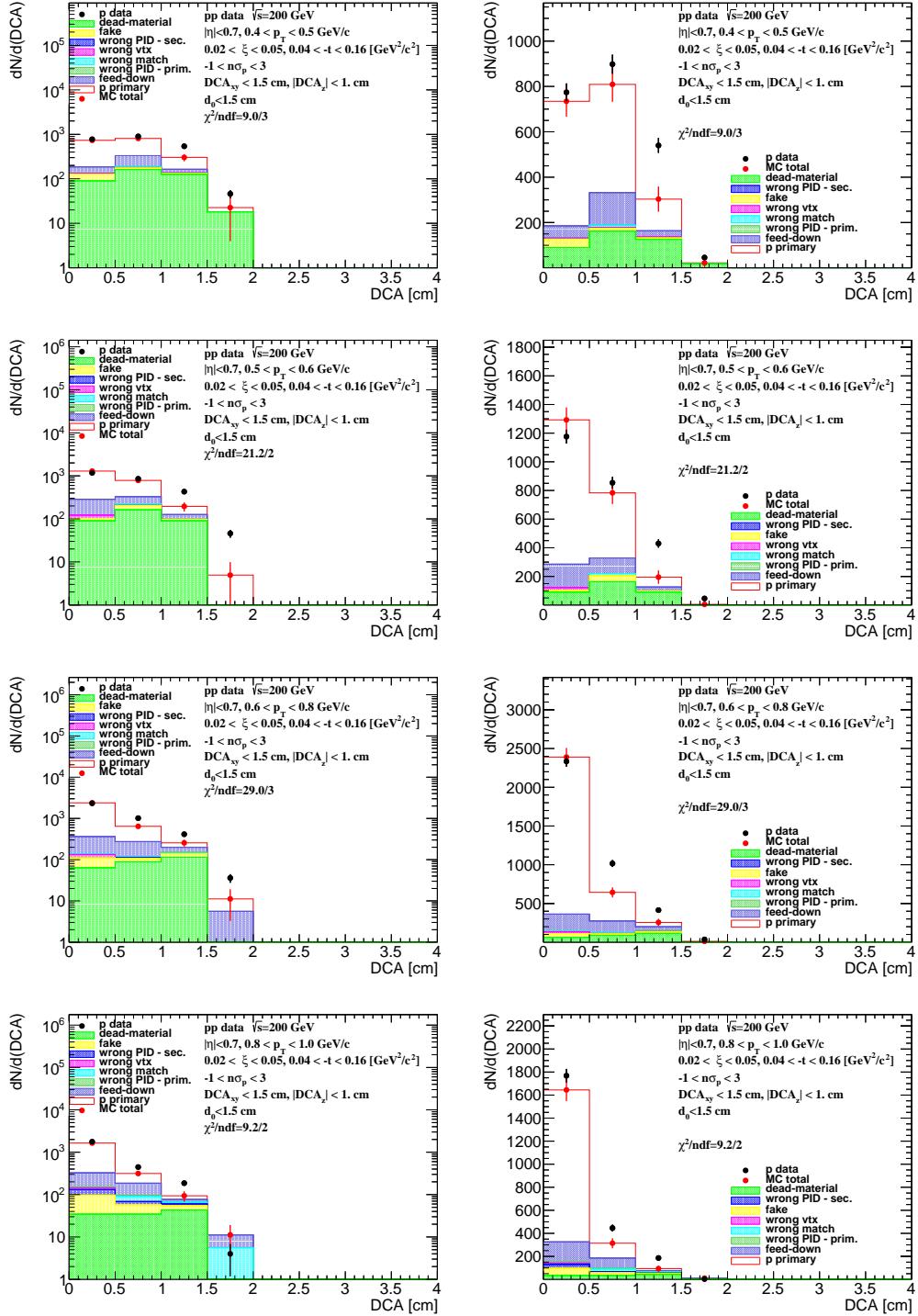


Figure A.2: Distributions of DCA for protons in SD interactions with $0.02 < \xi < 0.05$ and normal selection.

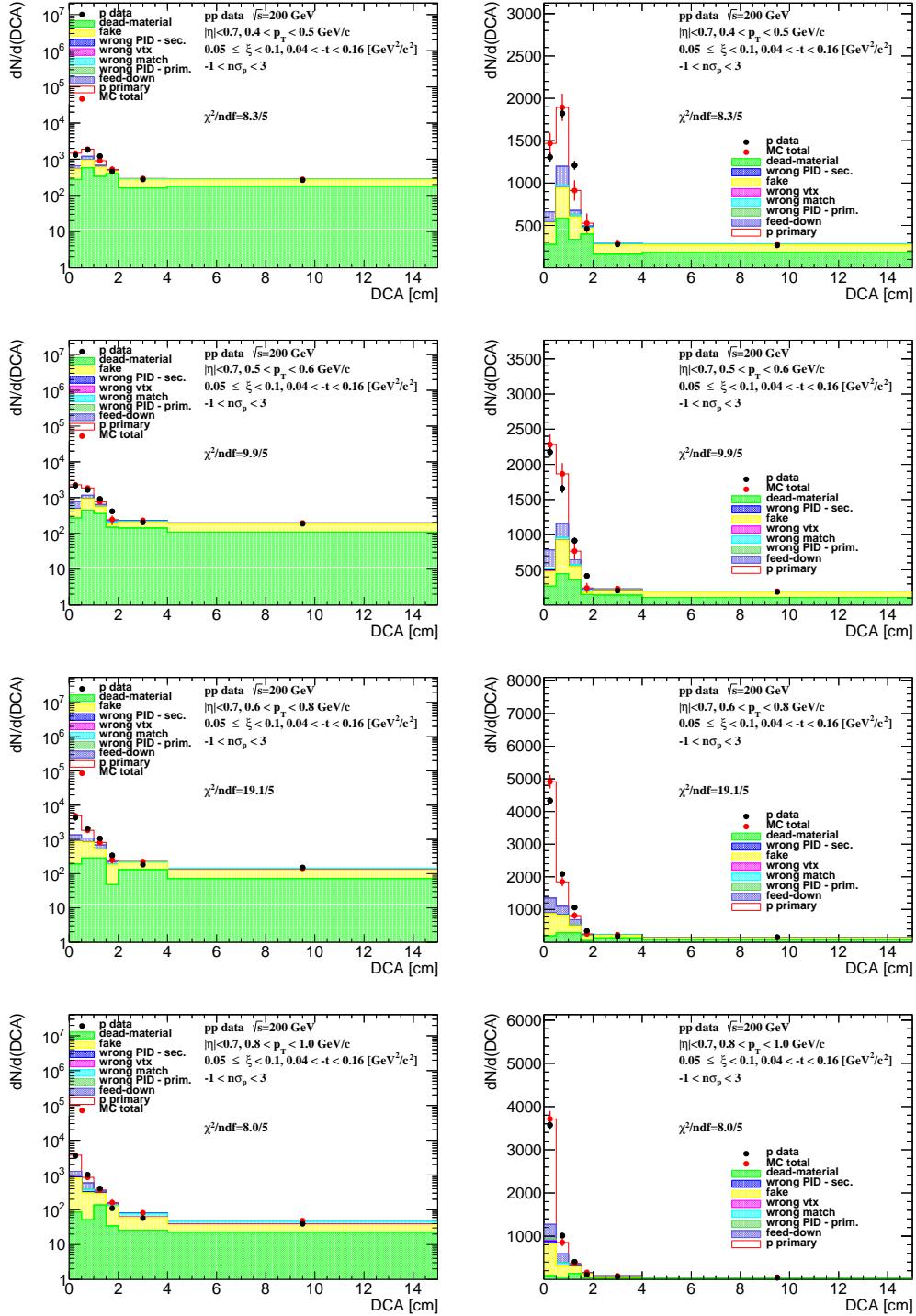


Figure A.3: Distributions of DCA for protons in SD interactions with $0.05 < \xi < 0.1$ and loose selection.

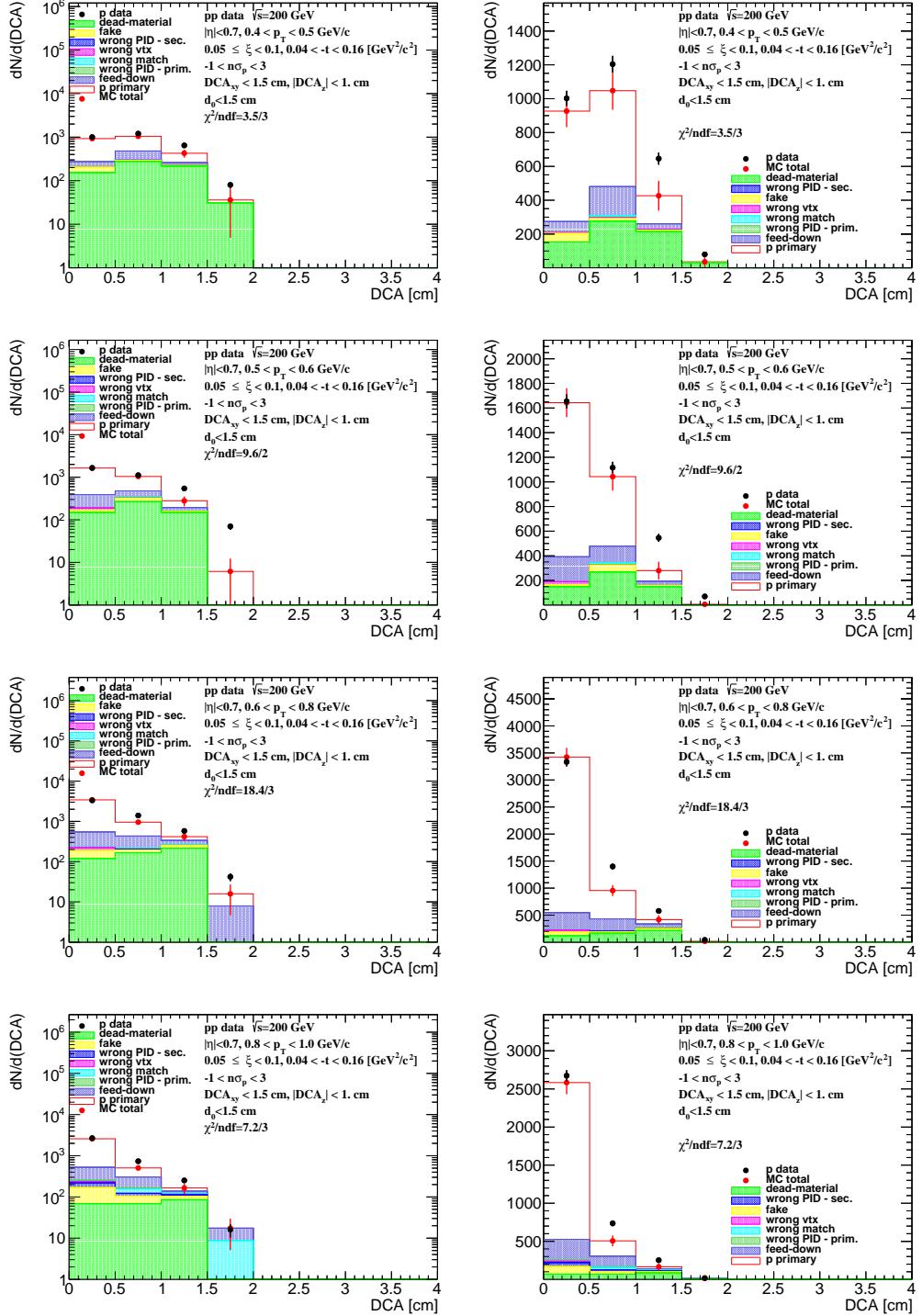


Figure A.4: Distributions of DCA for protons in SD interactions with $0.05 < \xi < 0.1$ and normal selection.

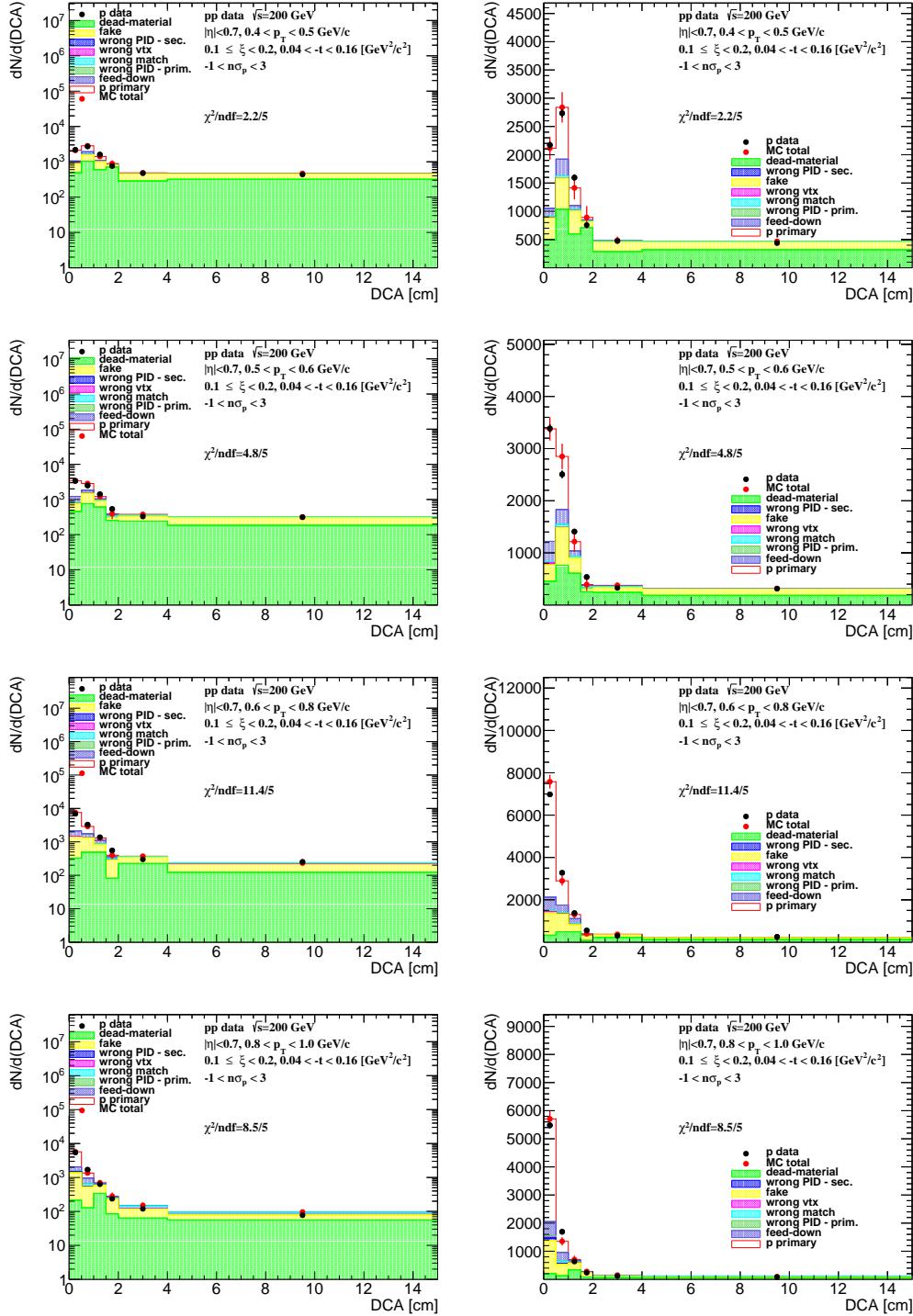


Figure A.5: Distributions of DCA for protons in SD interactions with $0.1 < \xi < 0.2$ and loose selection.

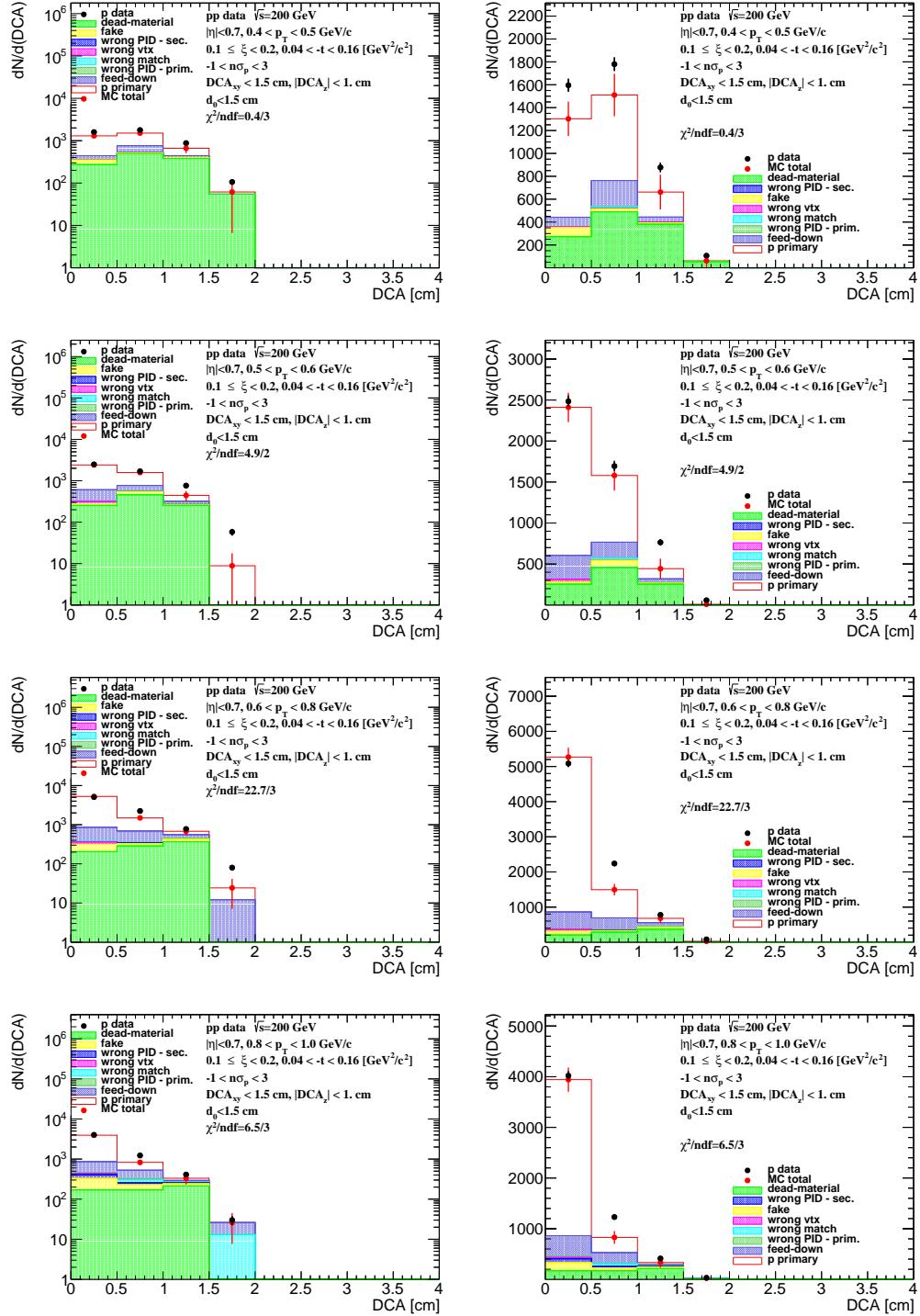


Figure A.6: Distributions of DCA for protons in SD interactions with $0.1 < \xi < 0.2$ and normal selection.

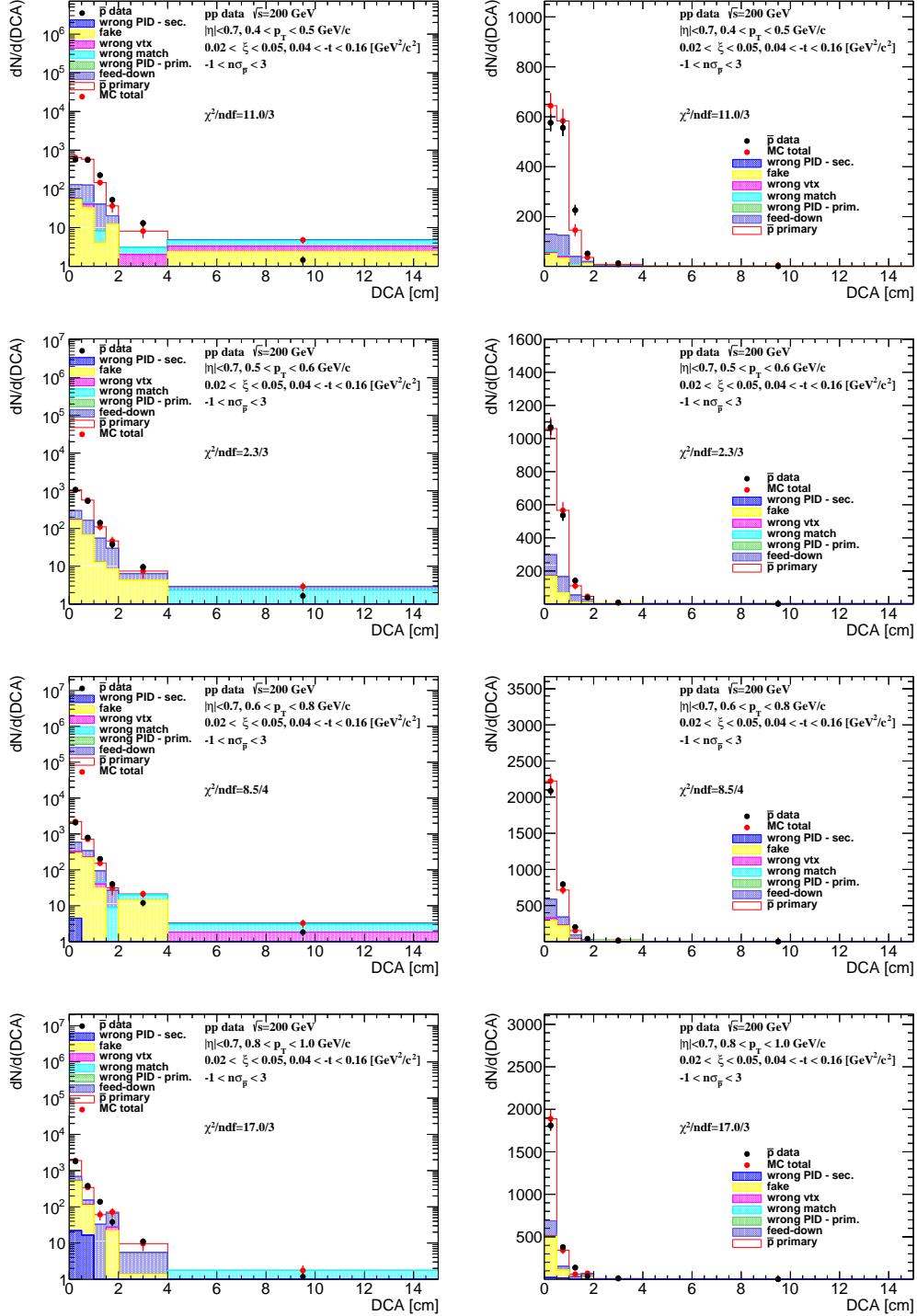


Figure A.7: Distributions of DCA for antiprotons in SD interactions with $0.02 < \xi < 0.05$ and loose selection.

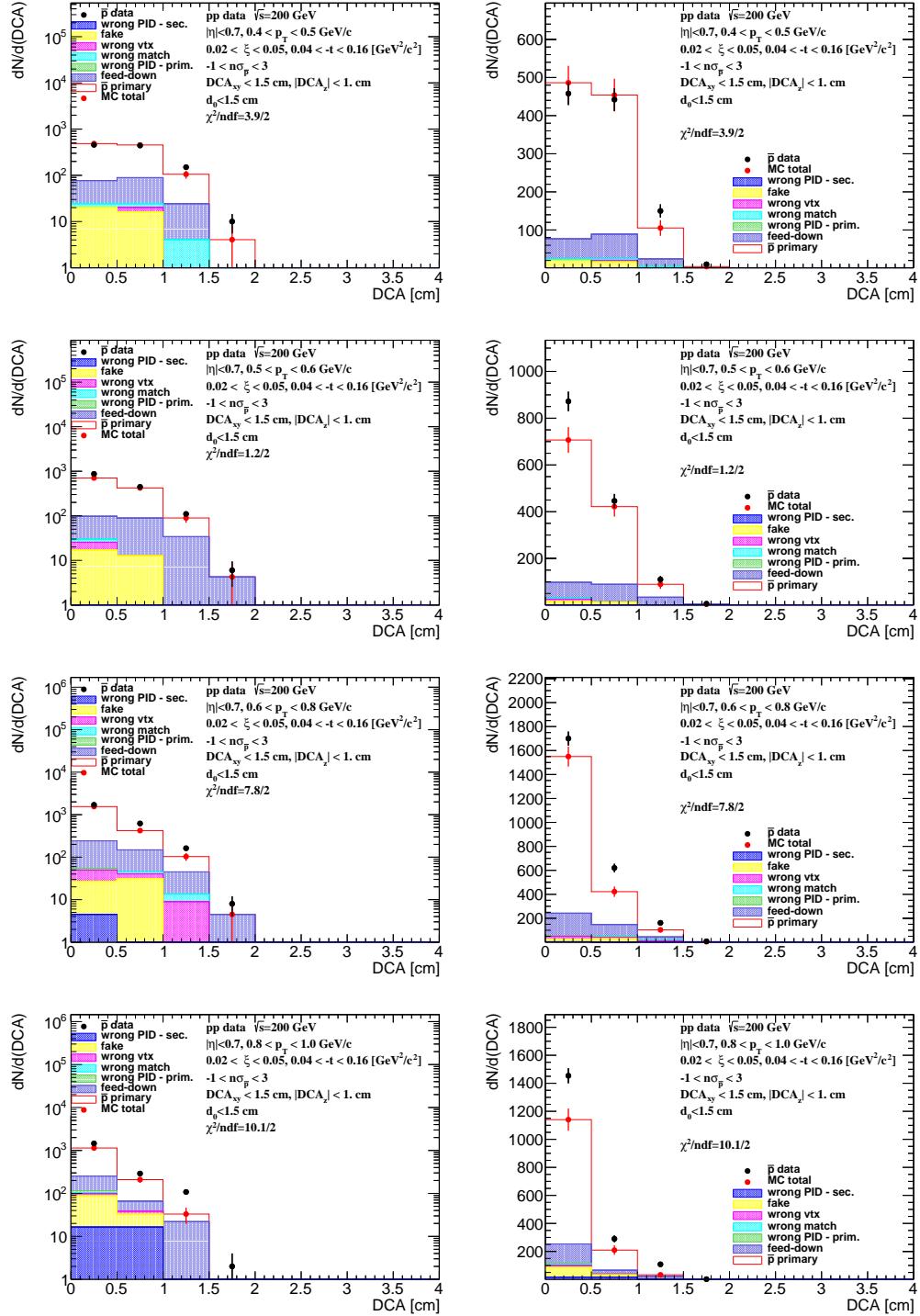


Figure A.8: Distributions of DCA for antiprotons in SD interactions with $0.02 < \xi < 0.05$ and normal selection.

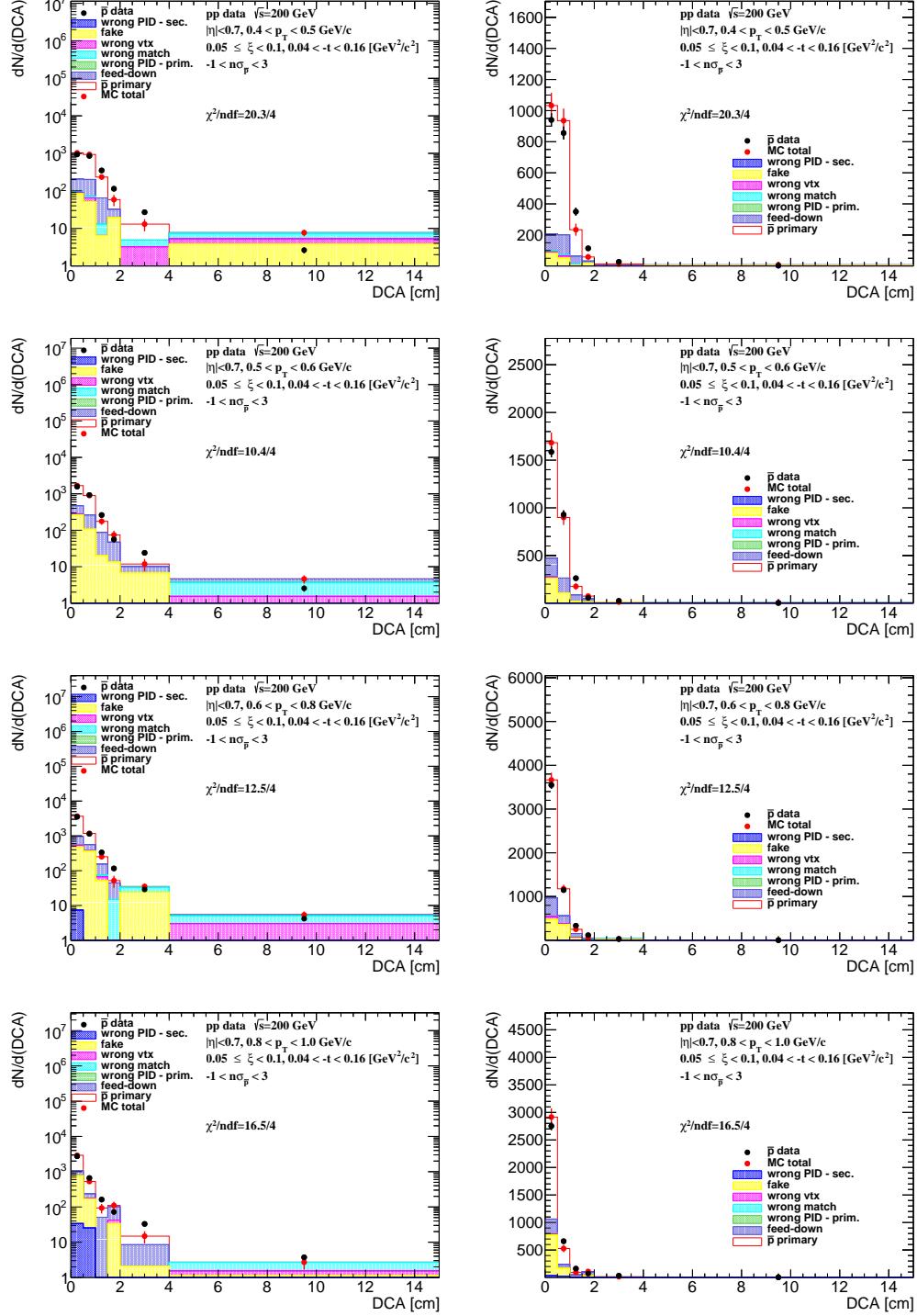


Figure A.9: Distributions of DCA for antiprotons in SD interactions with $0.05 < \xi < 0.1$ and loose selection.

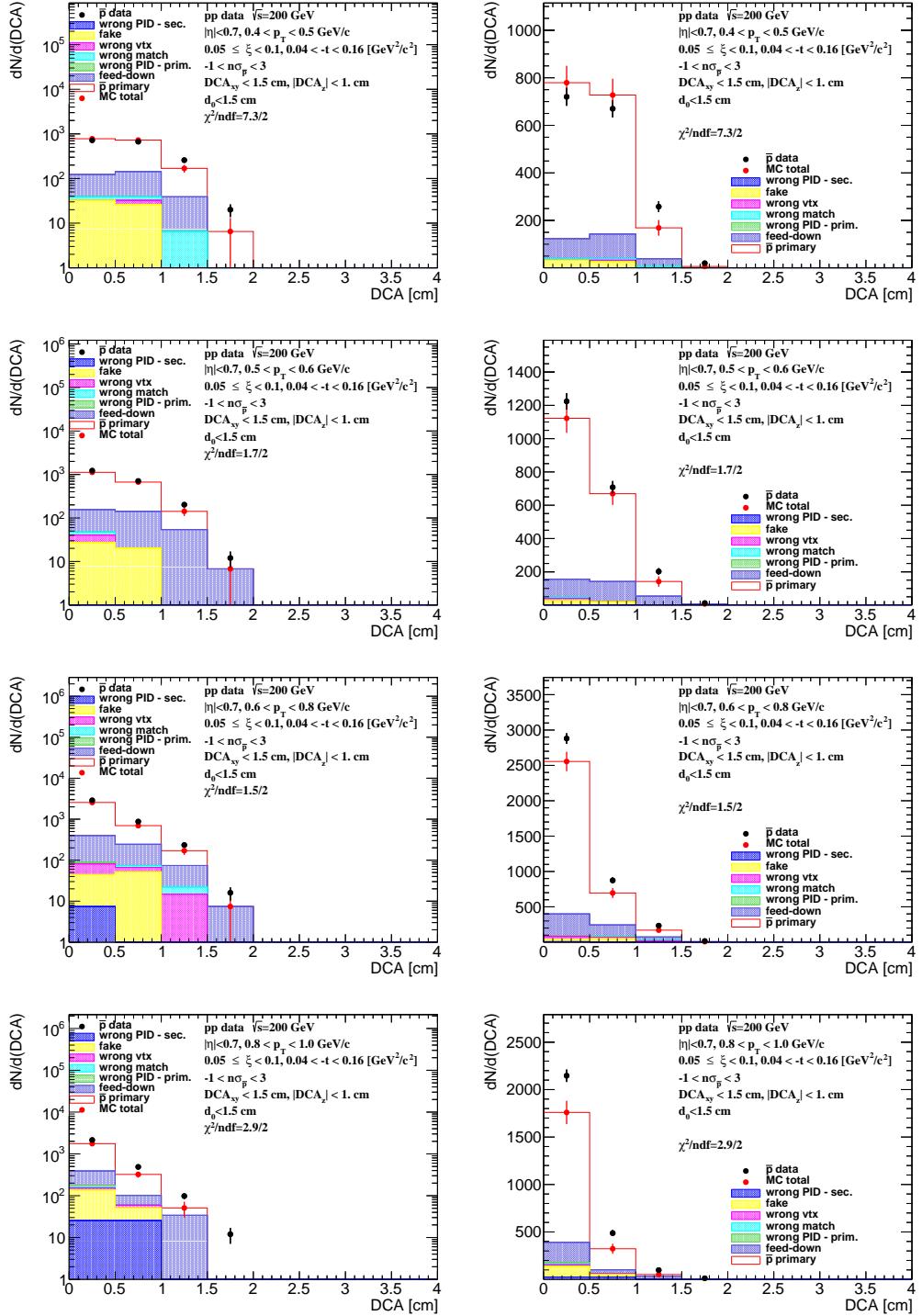


Figure A.10: Distributions of DCA for antiprotons in SD interactions with $0.05 < \xi < 0.1$ and normal selection.

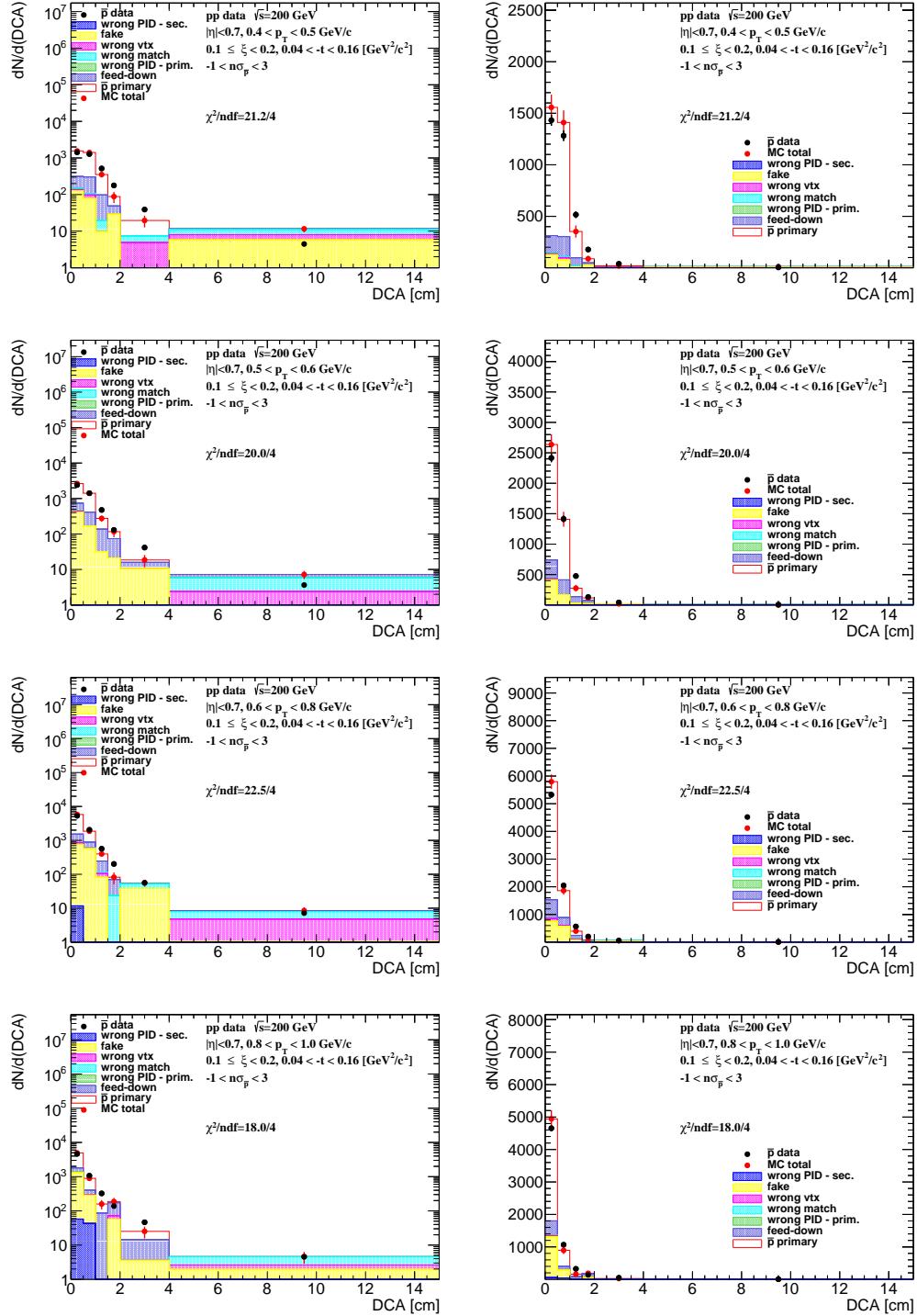


Figure A.11: Distributions of DCA for antiprotons in SD interactions with $0.1 < \xi < 0.2$ and loose selection.

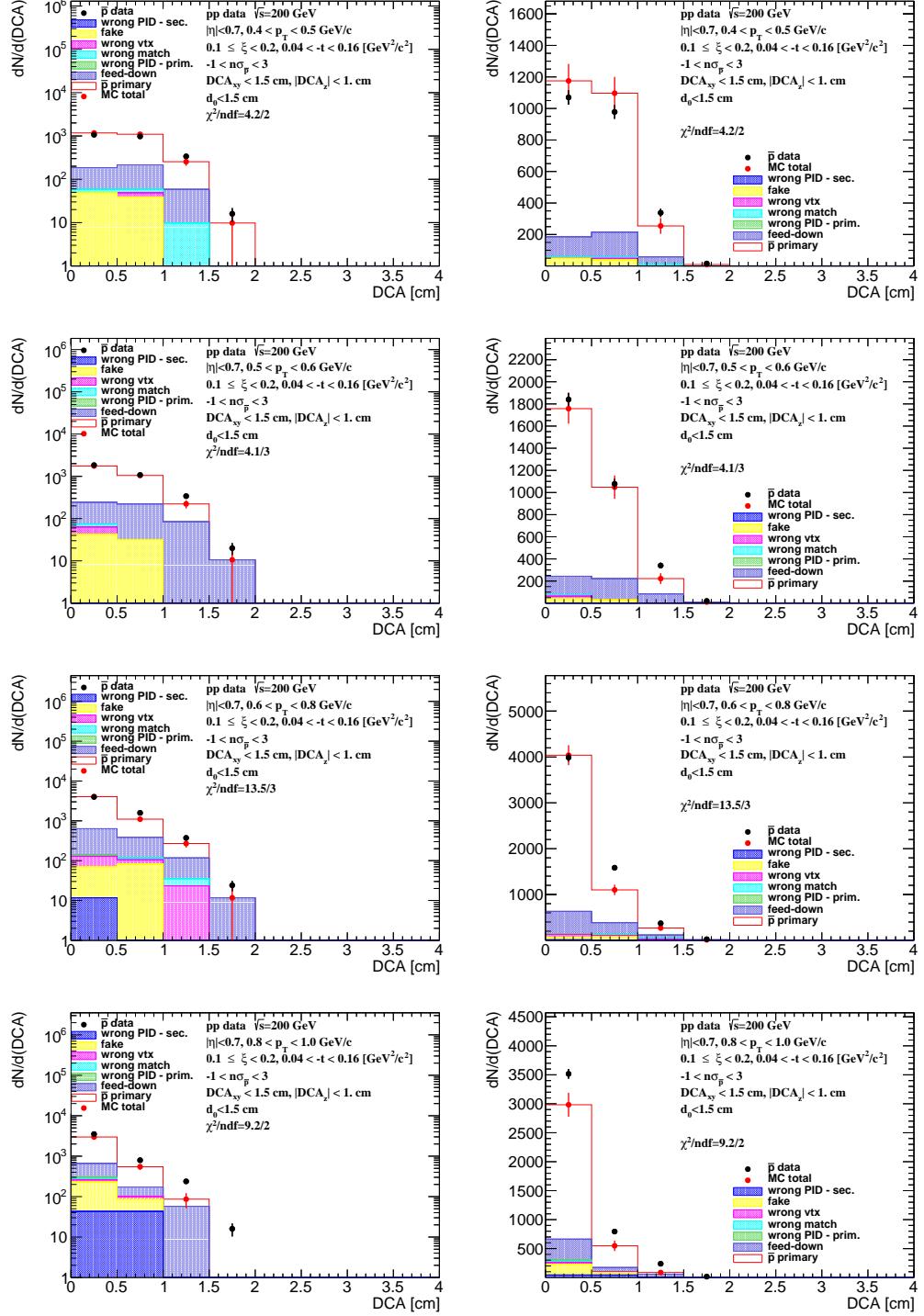


Figure A.12: Distributions of DCA for antiprotons in SD interactions with $0.1 < \xi < 0.2$ and normal selection.

B. Distributions of $n\sigma_{dE/dx}^i$ in SD

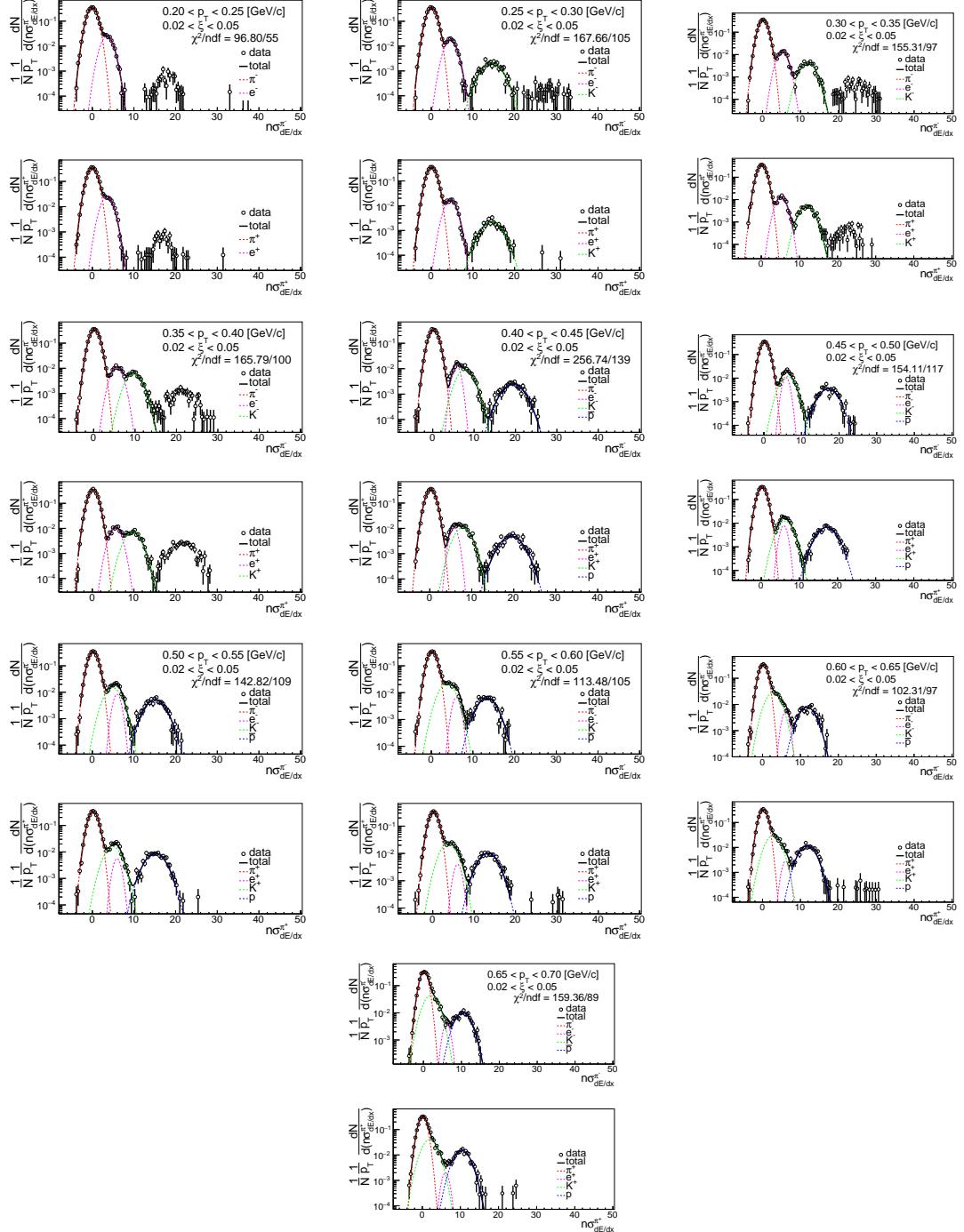


Figure B.1: Distributions of $n\sigma_{dE/dx}^{\pi^\pm}$ for π^\pm in SD interactions with $0.02 < \xi < 0.05$.

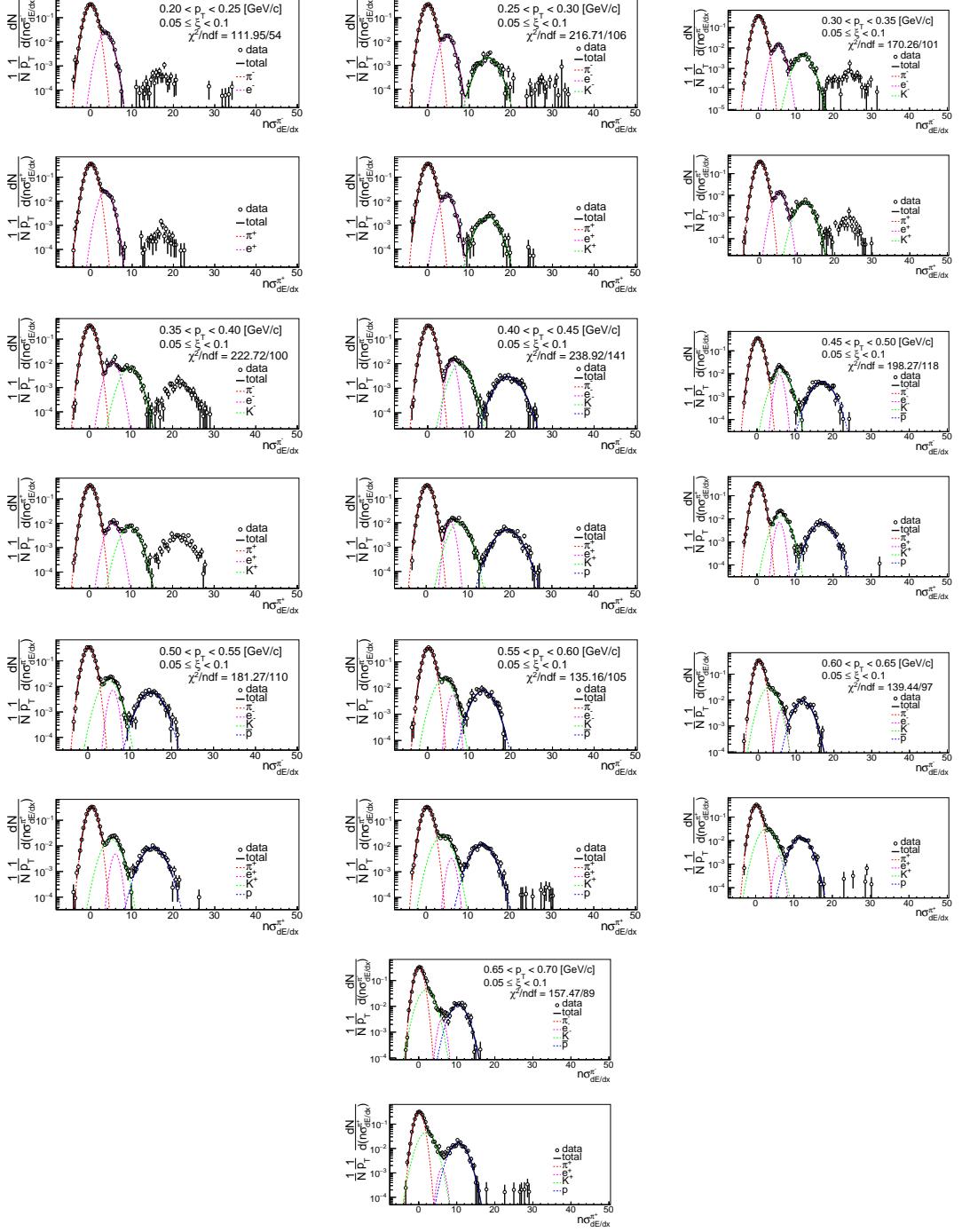


Figure B.2: Distributions of $n\sigma_{dE/dx}^{\pi^\pm}$ for π^\pm in SD interactions with $0.05 < \xi < 0.1$.

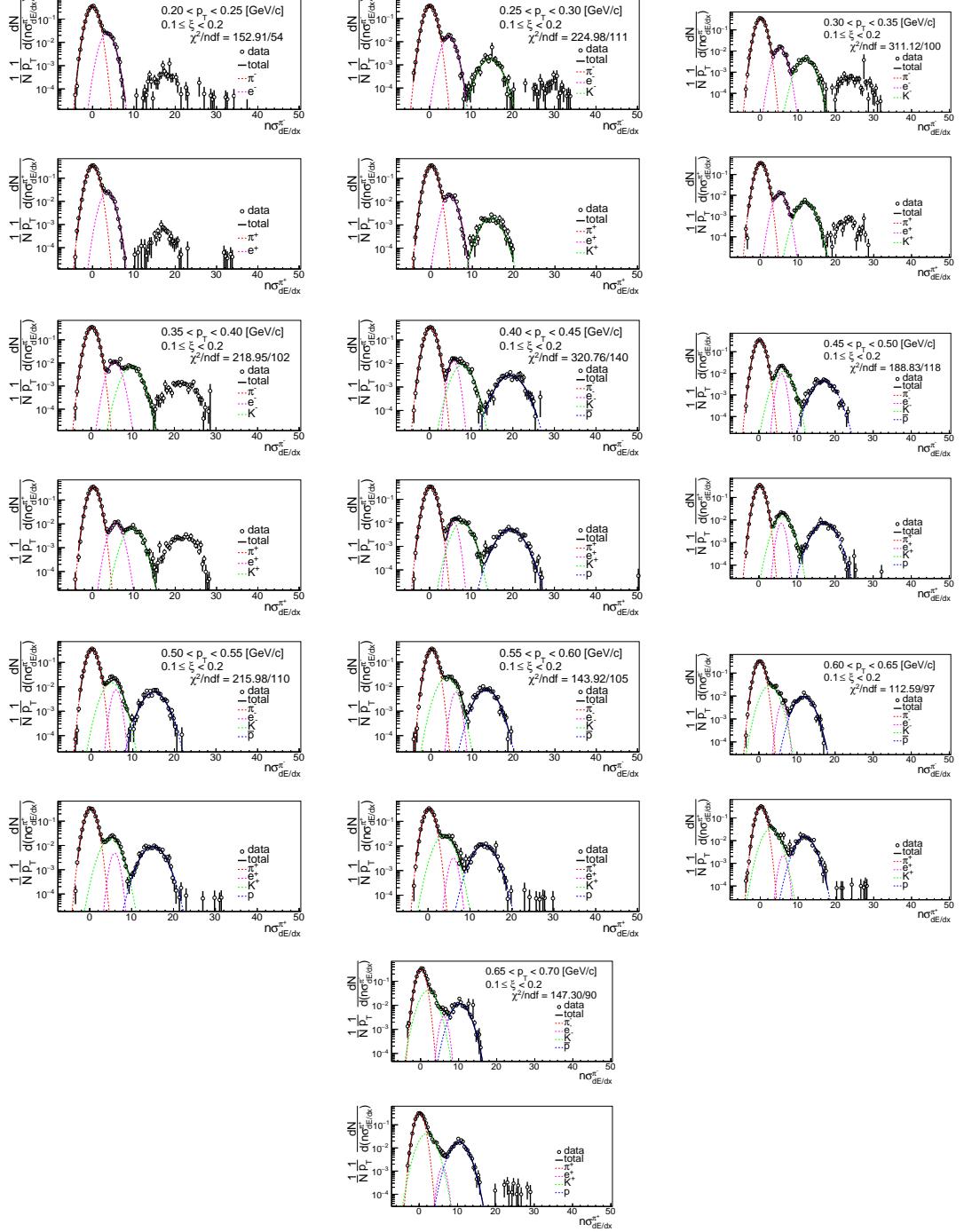


Figure B.3: Distributions of $n\sigma_{dE/dx}^{\pi^\pm}$ for π^\pm in SD interactions with $0.1 < \xi < 0.2$.

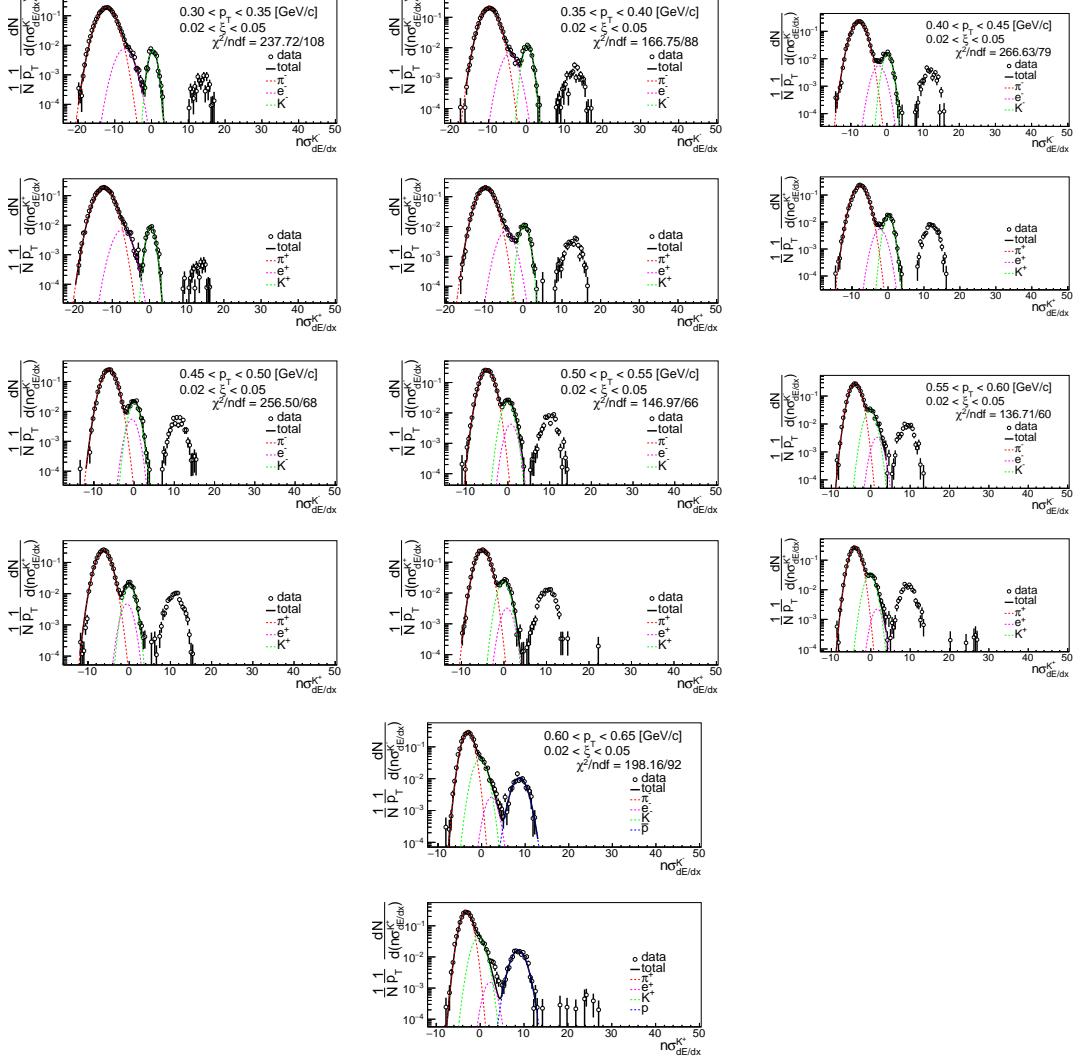


Figure B.4: Distributions of $n\sigma_{dE/dx}^{K^\pm}$ for K^\pm in SD interactions with $0.02 < \xi < 0.05$.

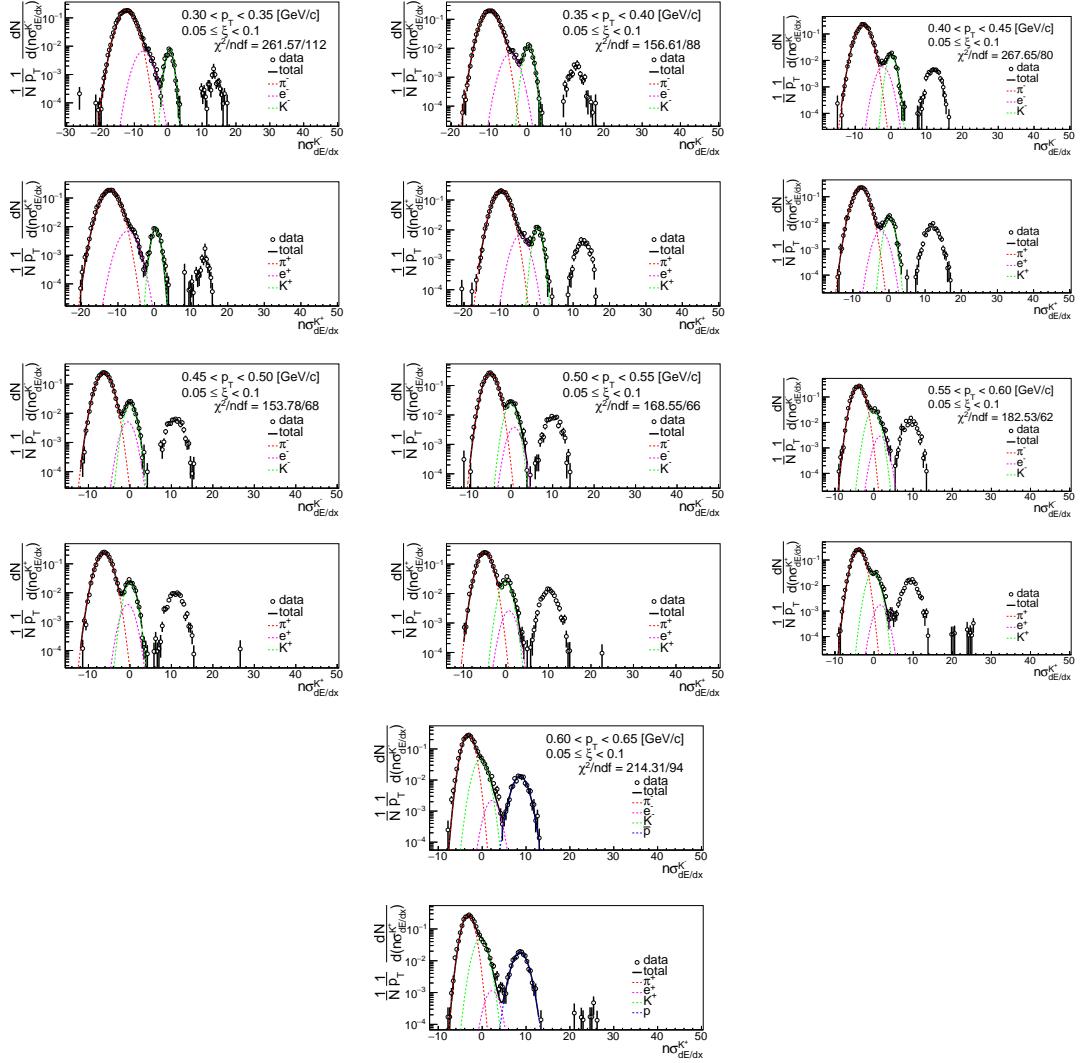


Figure B.5: Distributions of $n\sigma_{dE/dx}^{K^\pm}$ for K^\pm in SD interactions with $0.05 < \xi < 0.1$.

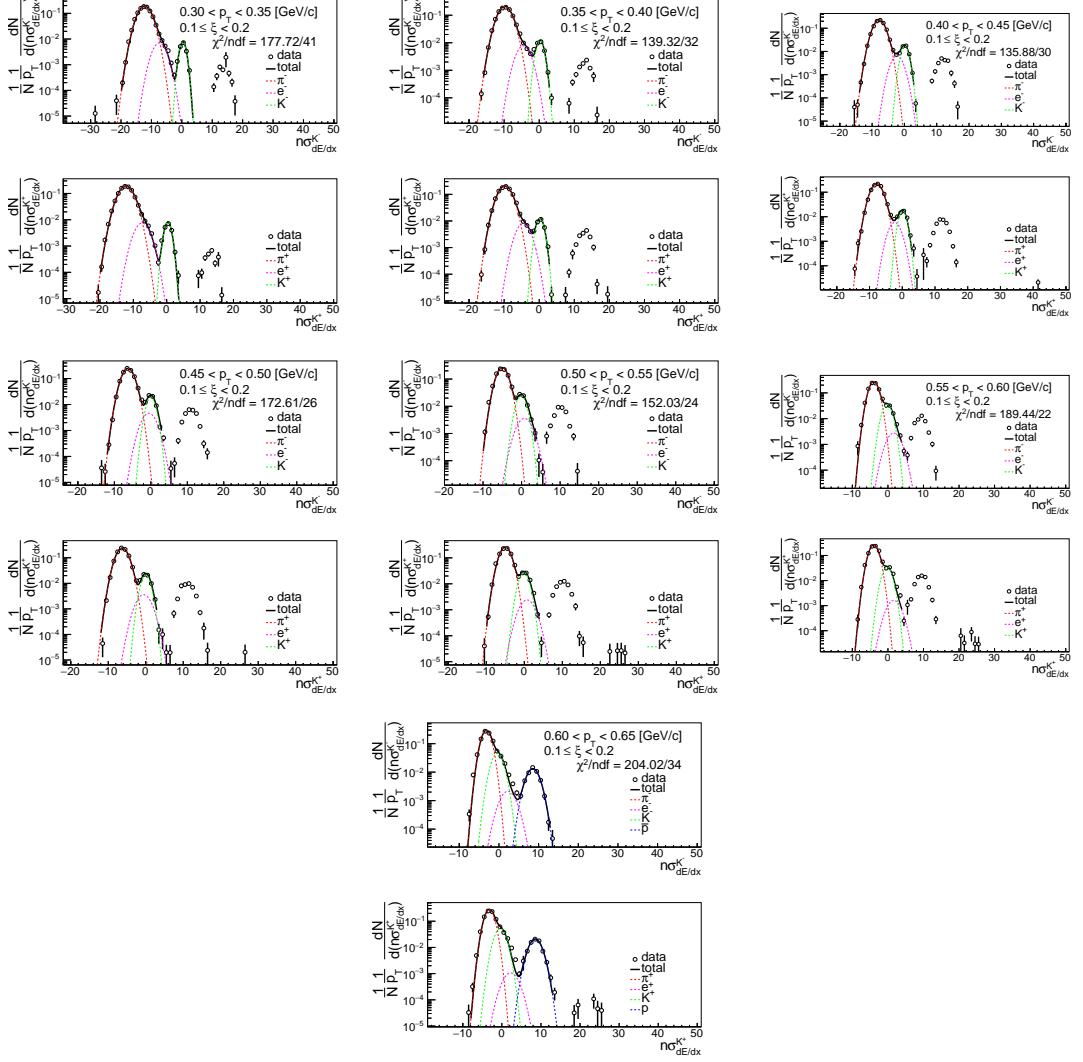


Figure B.6: Distributions of $n\sigma_{dE/dx}^{K^\pm}$ for K^\pm in SD interactions with $0.1 < \xi < 0.2$.

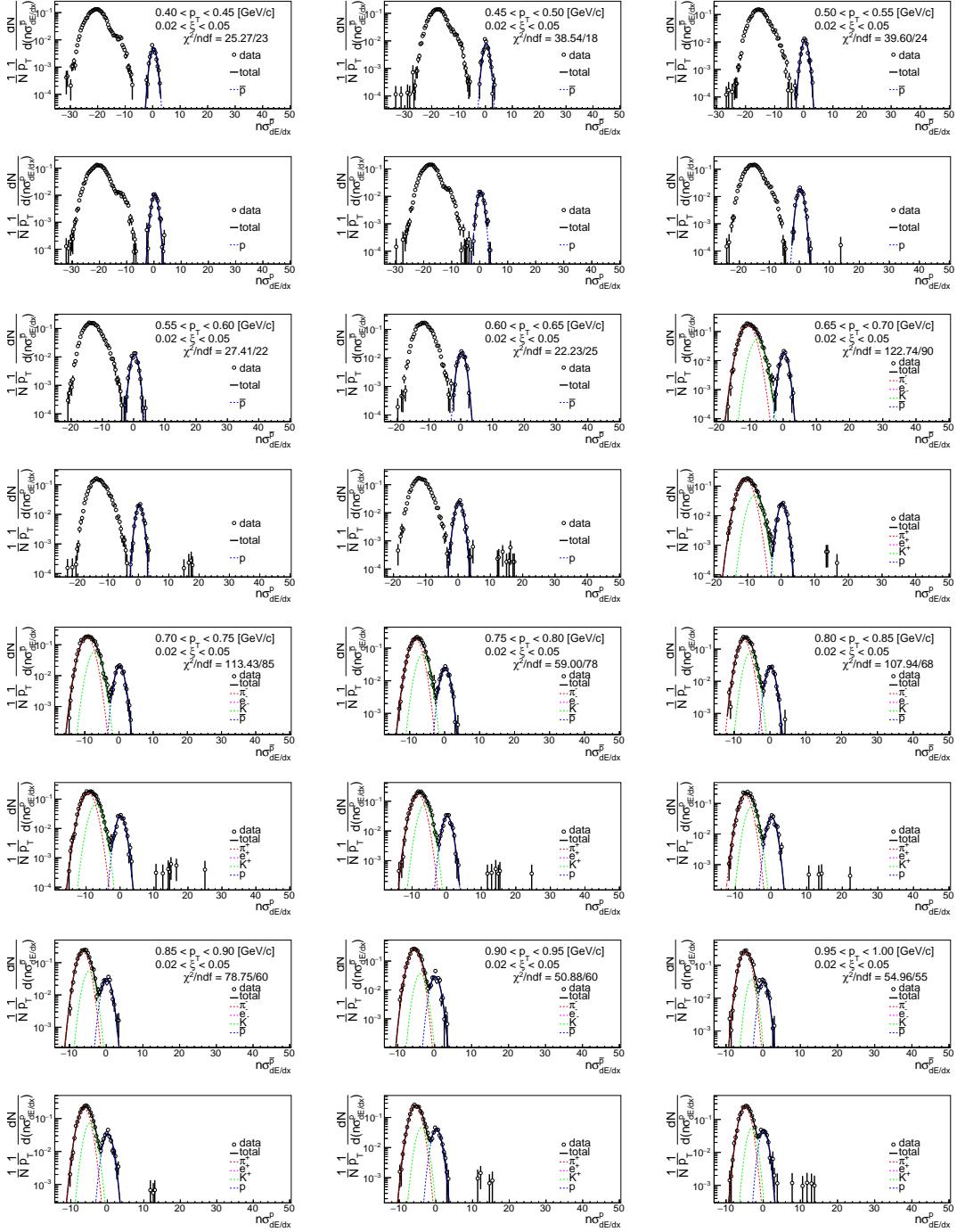


Figure B.7: Distributions of $n\sigma_{dE/dx}^{\bar{p},p}$ for \bar{p}, p in SD interactions with $0.02 < \xi < 0.05$.

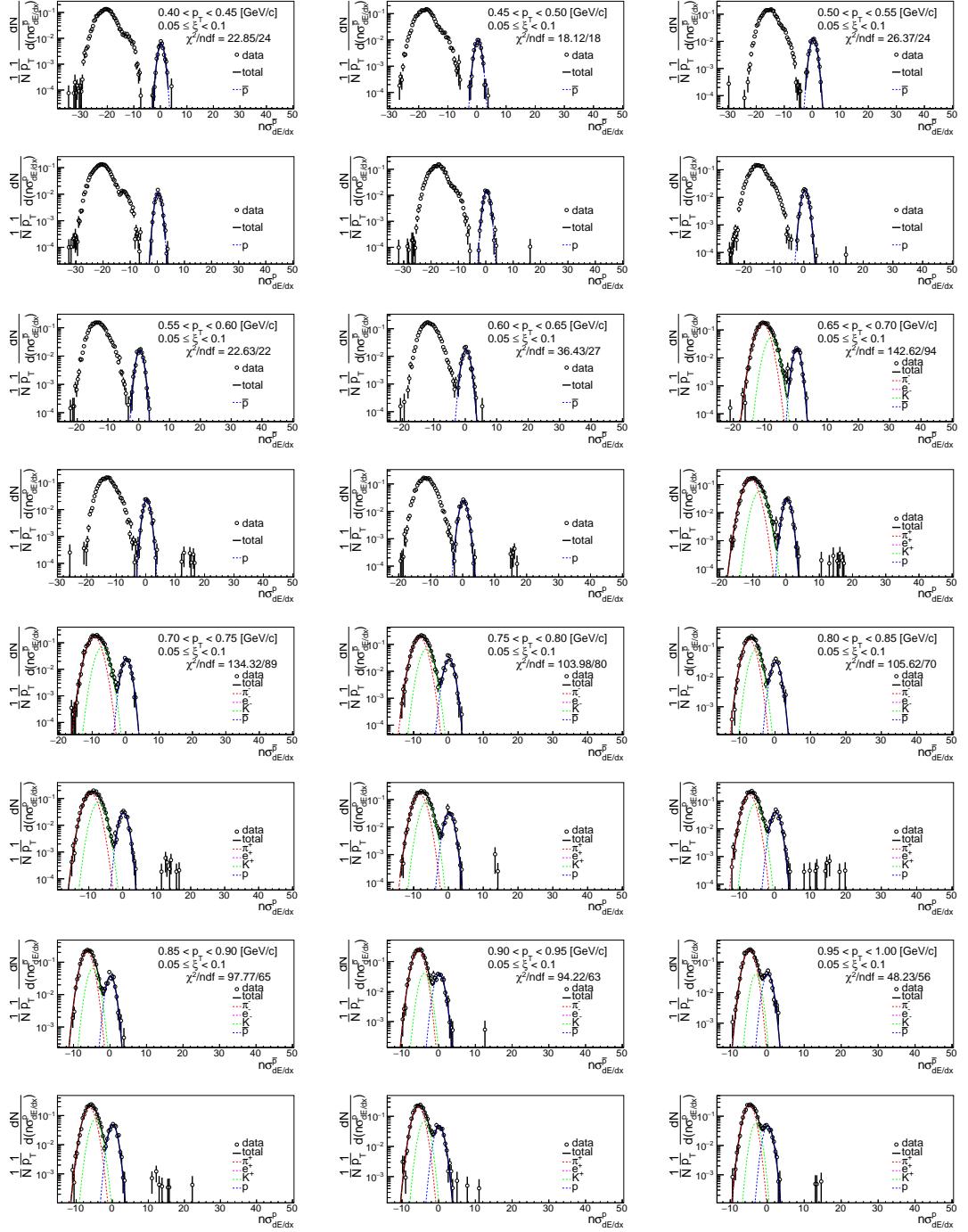


Figure B.8: Distributions of $n\sigma_{dE/dx}^{\bar{p},p}$ for \bar{p}, p in SD interactions with $0.05 < \xi < 0.1$.

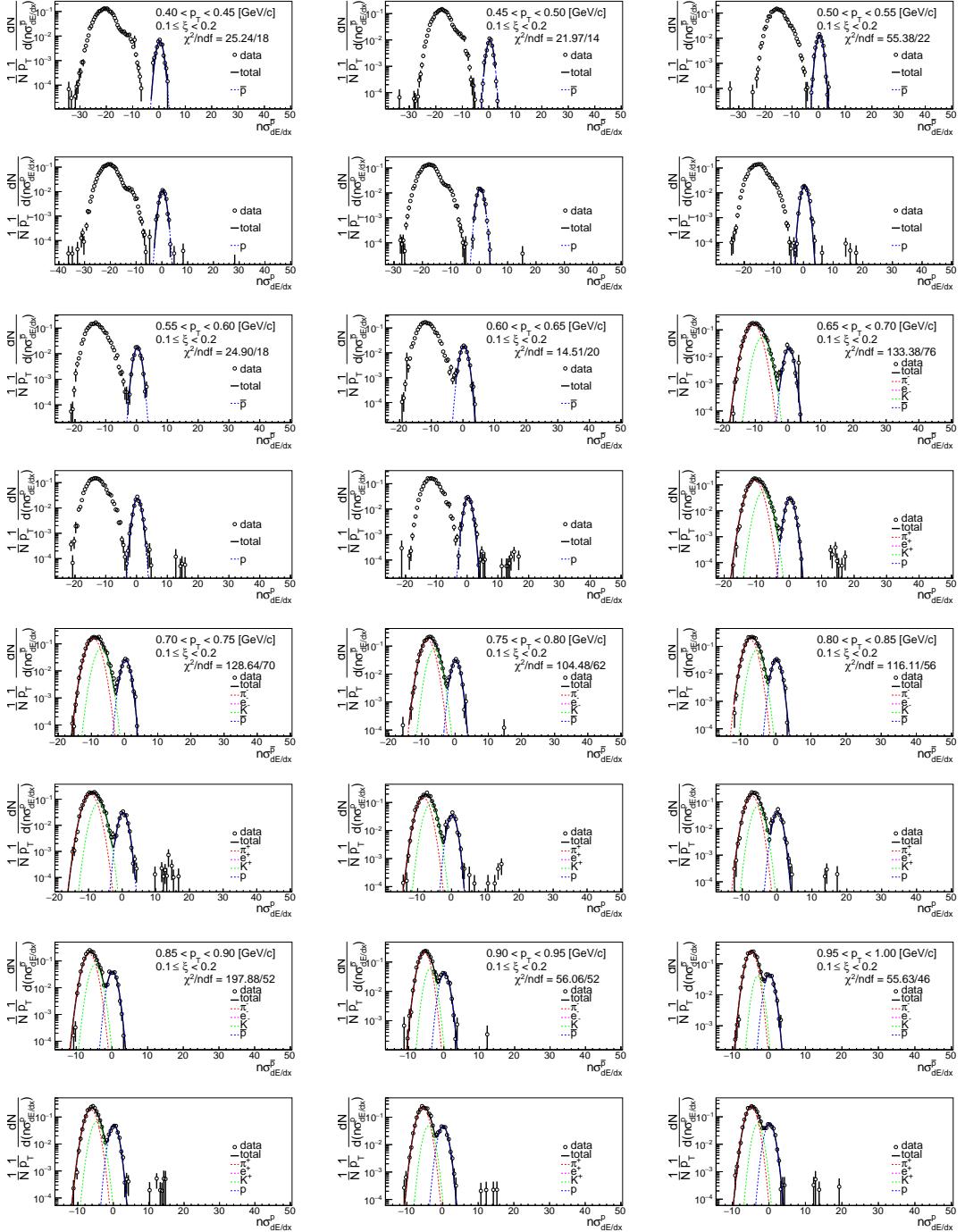


Figure B.9: Distributions of $n\sigma_{dE/dx}^{\bar{p},p}$ for \bar{p}, p in SD interactions with $0.1 < \xi < 0.2$.