

<sup>1</sup>

## Analysis Note

# Measurement of charged particle production in diffractive proton-proton collisions at $\sqrt{s} = 200$ GeV with tagging of the forward scattered proton

Leszek Adamczyk<sup>1</sup>, Łukasz Fulek<sup>1</sup>, Mariusz Przybycień<sup>1</sup>, and Rafał Sikora<sup>1</sup>

<sup>1</sup>*AGH University of Science and Technology, FPACS, Kraków, Poland*

<sup>2</sup>

27th March 2020

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  $\xi$ :  $0.02 < \xi < 0.05$ ,  $0.05 < \xi < 0.1$  and  $0.1 < \xi < 0.2$ . The identified particle to antiparticle (pion, kaon, proton and their antiparticle) multiplicity ratios as a function of transverse momentum in above three  $\xi$  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].

<sup>3</sup> **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

---

<sup>7</sup> \* - contact editor

<sup>9</sup> **Change log**

---

<sup>10</sup> 27th March      ver. 1.0      Initial revision  
<sup>11</sup> 2020

---

# Contents

14	<b>List of contributions</b>	<b>2</b>
16	<b>Change log</b>	<b>2</b>
17	<b>1 Introduction</b>	<b>1</b>
18	<b>2 Monte Carlo Samples</b>	<b>2</b>
19	<b>3 Data Sample and Event Selection</b>	<b>4</b>
20	3.1 Event Selection . . . . .	4
21	3.2 Track Selection . . . . .	5
22	3.3 Fiducial Region of the Measurement . . . . .	8
23	<b>4 Background Contribution</b>	<b>9</b>
24	4.1 Accidental Background . . . . .	9
25	4.2 Background from Non-Primary Tracks . . . . .	11
26	<b>5 Non-SD Contributions</b>	<b>18</b>
27	<b>6 Selection Efficiencies</b>	<b>22</b>
28	6.1 Vertex Reconstruction . . . . .	22
29	6.2 Correction to BBC-Small . . . . .	26
30	<b>7 Migrations into and out of the Fiducial Region</b>	<b>31</b>
31	7.1 Migrations of Tracks into and out of the Fiducial Region . . . . .	31
32	7.2 Migrations of $\xi$ . . . . .	32
33	<b>8 Corrections and Unfolding Procedure</b>	<b>35</b>
34	8.1 Correction to $dN/dn_{sel}$ . . . . .	35
35	8.2 Correction to Transverse Momentum and Pseudorapidity Distributions . . . . .	37
36	8.3 Closure Tests . . . . .	38
37	8.4 EAST-WEST asymmetry . . . . .	38
38	8.5 Particle Identification . . . . .	39
39	8.6 Antiparticle-to-Particle Ratios . . . . .	44
40	<b>9 Systematic Uncertainties</b>	<b>47</b>
41	<b>10 Results</b>	<b>52</b>
42	<b>11 Summary and Conclusions</b>	<b>59</b>
43	<b>Appendices</b>	<b>61</b>
44	<b>A Acronyms</b>	<b>62</b>

<sup>45</sup>	<b>B Proton and Antiproton DCA Distributions</b>	<b>63</b>
<sup>46</sup>	<b>C Distributions of <math>n\sigma_{dE/dx}^i</math> in SD</b>	<b>76</b>

# 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:

$$\frac{1}{N_{\text{ev}}} \frac{dN_{\text{ev}}}{dn_{\text{ch}}}, \quad \frac{1}{N_{\text{ev}}} \frac{1}{2\pi p_{\text{T}}} \frac{d^2 N}{d\bar{\eta} dp_{\text{T}}}, \quad \frac{1}{N_{\text{ev}}} \frac{dN}{d\bar{\eta}} \quad (1.1)$$

where  $n_{\text{ch}}$  is the number of primary charged particles within kinematic range given by  $p_{\text{T}} > 200$  MeV and  $|\eta| < 0.7$ ,  $N_{\text{ev}}$  is the total number of events with  $2 \leq n_{\text{ch}} \leq 8$ ,  $N$  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  $\xi$  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  $\xi$  regions:  $0.02 < \xi < 0.05$ ,  $0.05 < \xi < 0.1$  and  $0.1 < \xi < 0.2$ . In addition their average values ~~in an event~~ are presented as a function of  $\xi$ .

We have also studied an identified particle to antiparticle (pion, kaon, proton and their anti-particle) multiplicity ratios as a function of  $p_{\text{T}}$  also in the above mentioned three regions of  $\xi$ . The system  $X$  into which proton diffractively dissociates has net charge and baryon number +1. It is believed that initial charge and baryon 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  $\xi$  where amount of particle-antiparticle creation is smaller due to the generally smaller particle multiplicity or due to the fact that gap edge is inside our fiducial region of  $|\eta| < 0.7$ .

## 2. Monte Carlo Samples

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

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

The particles were propagated through the full simulation of the STAR TPC and RP system detectors using GEANT3 [5] and GEANT4 [6], respectively. Obtained information for the simulated particles was embedded into the existing information of the real data. These events were next processed through the full reconstruction chain.

An iterative unfolding procedure was used to express the track multiplicity in terms of the number of charged particles. The unfolding matrices were obtained from PYTHIA 8 4C (SaS). In order to keep statistical precision coming from the matrices high, samples filtered on true-level values of  $\xi$  and  $t$  (not necessarily with reconstructed proton track in RP) are used.

It is preferred to get the detector corrections from a MC, which is dedicated to simulate the studied physics process. However, for this purpose, the statistics in the MC should be several times greater than we have in the data for analysis. Since this is not possible with low efficiency of TPC and TOF, the basic method of corrections used in the analysis for  $p_T$  and  $\bar{\eta}$  distributions is a method of factorization of global efficiency into the product of single-particle efficiencies. In this way, statistically precise multidimensional corrections on TPC and TOF were obtained from the single particle MC. ~~These corrections were not applied to charged-particle multiplicity distributions since the unfolding matrices already contain TPC and TOF efficiencies.~~

~~Additionally, several pure MC samples were generated. The simulated particles were propagated through full simulation and reconstruction chain but were not embedded into Zerobias events. Systematic effect related to hadronization of the diffractive system was determined by using an alternative hadronization model implemented in HERWIG. Results are compared to model predictions from PYTHIA 8 4C (SaS), HERWIG, EPOS and alternative PYTHIA 8 model Minimum Bias Rockefeller (MBR) with A2 tune. EPOS predicts very large contribution of forward protons, which originate from non-diffractive events and are well separated in rapidity from other final state particles. This is the result of low mass excitation of the proton remnant ( $< 1$  GeV) leading to hadronization of the beam remnant back to the proton. Therefore EPOS predictions were separated in two classes: diffractive (EPOS SD) modelled by Pomeron exchange and non-diffractive modelled by low mass excitation of the proton remnant (EPOS SD'). Such remnant treatment is very unique in EPOS compared to other string models, especially, to that used in PYTHIA 8, where Non-Diffractive (ND) forward protons are rare and arise from string fragmentation and hadronization. In all PYTHIA 8 models, diffractive cross-sections are scaled by the fudge factors, which were introduced in order to describe the full phase space [7, 8]. As a result, diffractive cross sections are artificially suppressed at relatively large values of  $\xi$  ( $> 0.05$ ). This artificial suppression significantly changes predicted distribution of  $\xi$  and fractions of different processes in our fiducial phase space. Therefore data is also compared with expectations obtained without suppression of the diffractive cross sections (MBR-tuned).~~

Figure 2.1 shows  $\xi$  and  $|t|$  distributions generated with EPOS (SD and SD+SD') and PYTHIA 8 SD (MBR and MBR-tuned). There are differences among models in both the low and high  $\xi$  regions. EPOS SD is only relevant for very small  $|t|$  and is suppressed in the STAR acceptance region,  $0.04 < |t| < 0.16$ , where EPOS SD' contribution dominates. In addition, the  $t$ -slope is very different for EPOS SD and SD', while it is similar for EPOS SD' and PYTHIA 8 predictions.



<sup>137</sup> The PYTHIA 8 (MBR-tuned) expectations, as opposed to the MBR model, lead to the larger  
<sup>138</sup> cross-sections in the high-mass regions.

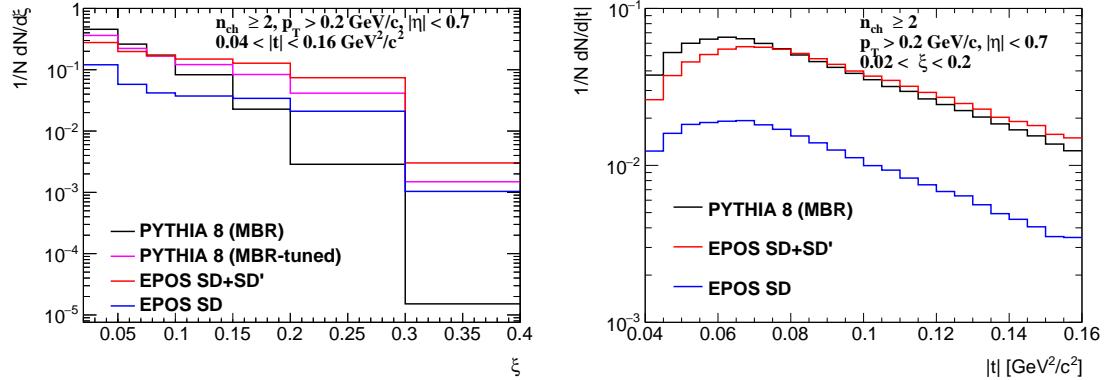


Figure 2.1: (left)  $\xi$  and (right)  $|t|$  distributions for various MC generators at  $\sqrt{s} = 200$  GeV.

# 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, i.e. in year 2015.

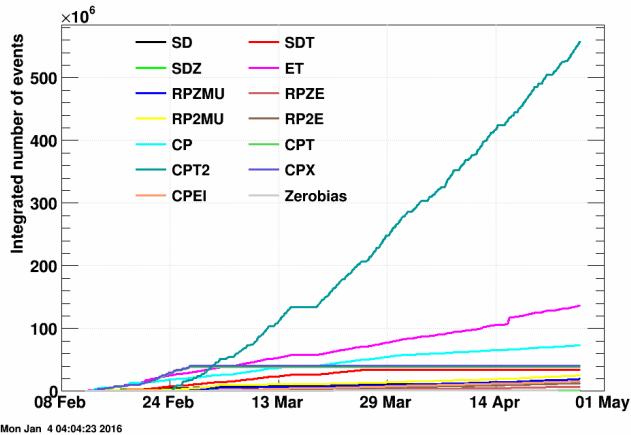


Figure 3.1: Cumulative number of events collected for each trigger in the RP data stream during Run 15 [9, 10].

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

1. RP\_EOR || RP\_WOR - signal in at least one RP on one 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.

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

Integrated luminosity delivered by the RHIC to the STAR detector in  $pp$  collisions during Run 15 amounts to  $185.1 \text{ pb}^{-1}$  [11], whereas about 34.4M SDT events were gathered by the STAR detector, shown in Fig. 3.1, which corresponds to  $16 \text{ nb}^{-1}$  of integrated luminosity.

## 3.1 Event Selection

Events were selected from those passing the SDT trigger condition. In order to remove events having 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 analyzed independently, e.g. for background studies),

- 163     2. Any trigger signal in small BBC tiles on the opposite side of the STAR central detector to  
 164     the triggered RP station,
- 165     3. Exactly one proton track in the above RP stations with  $0.02 < \xi < 0.2$  and  $0.04 < -t <$   
 166      $0.16 \text{ GeV}^2/\text{c}^2$ .
- 167     4. Exactly one vertex with TPC tracks matched with hits in TOF (later in the text such vertex  
 168     is referred as a TOF vertex),
- 169     5. Vertex within  $|V_z| < 80 \text{ cm}$  - events with vertices away from the nominal IP have low  
 170     acceptance for the central and forward tracks [1],
- 171     6. At least two but no more than eight primary TPC tracks,  $2 \leq n_{\text{sel}} \leq 8$ , matched with hits  
 172     in TOF and satisfying the selection criteria described in Sec. 3.2,
- 173     7. If there are exactly two primary tracks satisfying above criteria and exactly two global tracks  
 174     used in vertex reconstruction (Sec. 6.1), the longitudinal distance between these global tracks  
 175     should be smaller than 2 cm,  $|\Delta z_0| < 2 \text{ cm}$ , due to small (< 20%) vertex reconstruction  
 176     efficiency for tracks with  $|\Delta z_0| > 2 \text{ cm}$  (as described in Sec. 6.1).

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

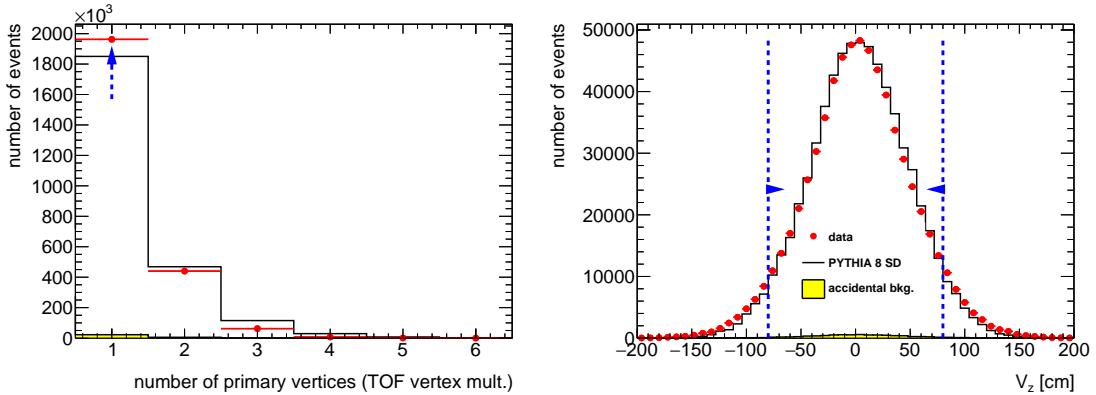


Figure 3.2: (left) TOF vertex multiplicity and (right) the  $z$ -position of reconstructed vertices in single TOF vertex events before applying the cut on the quantity shown. Blue lines indicate regions accepted in the analysis.

## 182     3.2 Track Selection

183     The following quality cuts had to be passed by the selected primary tracks in this analysis (TPC  
 184     and TOF efficiencies are described in [1]):

- 185     1. The tracks must be matched with hits reconstructed in TOF,
- 186     2. The number of the TPC hits used in the helix fit  $N_{\text{hits}}^{\text{fit}}$  must be greater than 24,
- 187     3. The ratio of  $N_{\text{hits}}^{\text{fit}}$  to the number of all possible TPC hits,  $N_{\text{hits}}^{\text{fit}}/N_{\text{hits}}^{\text{possible}}$ , must be greater  
 188     than 0.52,

- 189     4. The number of the TPC hits used to determine the  $dE/dx$  information  $N_{\text{hits}}^{\text{d}E/\text{d}x}$  must be  
 190     greater than 14,  
 191     5. The transverse impact parameter with respect to the beamline  $d_0$  must be less than 1.5 cm,  
 192     6. The radial component of the distance of the closest approach between the global helix and  
 193     the vertex  $\text{DCA}_{xy}$  must be less than 1.5 cm (consistent with the  $d_0$  limit),  
 194     7. The absolute magnitude of longitudinal component of the distance of the closest approach  
 195     between the global helix and the vertex  $|\text{DCA}_z|$  must be less than 1 cm,  
 196     8. The track's transverse momentum  $p_T$  must be greater than 0.2 GeV/c,  
 197     9. The track's absolute value of pseudorapidity  $|\eta|$  must be smaller than 0.7.

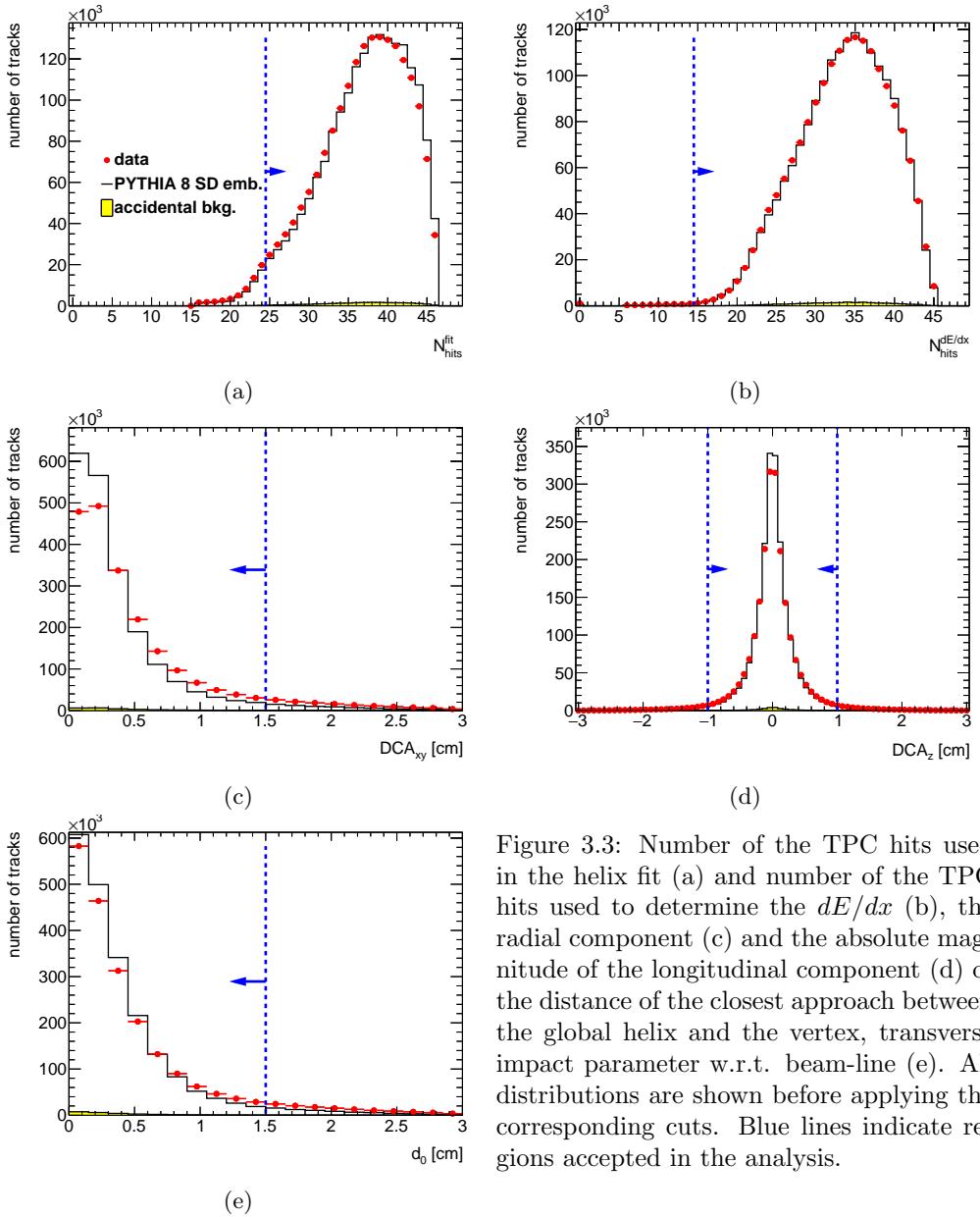


Figure 3.3: Number of the TPC hits used in the helix fit (a) and number of the TPC hits used to determine the  $dE/dx$  (b), the radial component (c) and the absolute magnitude of the longitudinal component (d) of the distance of the closest approach between the global helix and the vertex, transverse 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.

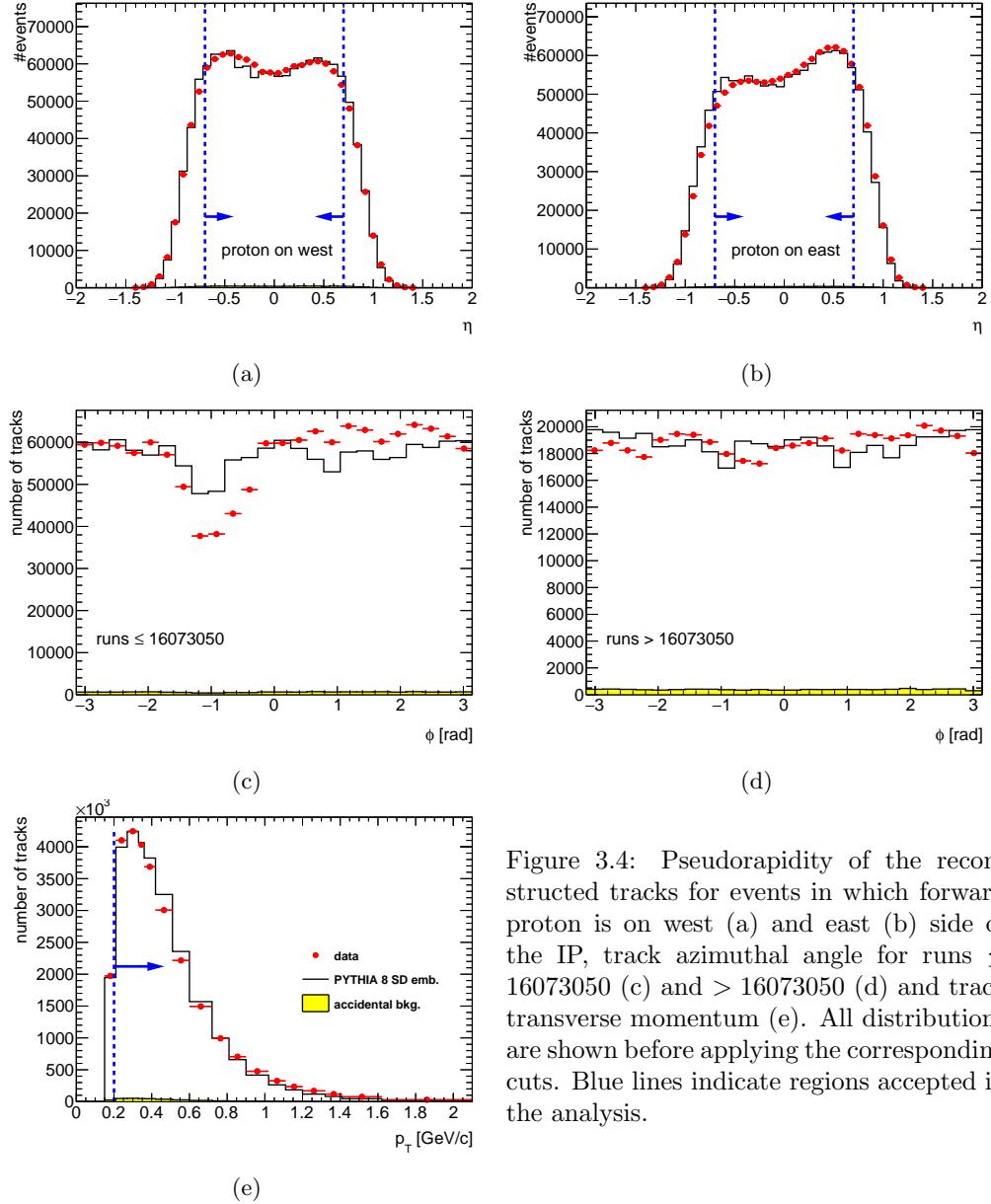


Figure 3.4: Pseudorapidity of the reconstructed tracks for events in which forward proton is on west (a) and east (b) side of the IP, track azimuthal angle for runs  $\leq 16073050$  (c) and  $> 16073050$  (d) and track transverse momentum (e). All distributions are shown before applying the corresponding cuts. Blue lines indicate regions accepted in the analysis.

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 splitting effects. The  $d_0$  and global  $\text{DCA}_{xy}$ ,  $|\text{DCA}_z|$  cuts are used to select tracks that originate from the primary interaction vertex. The cut on  $N_{\text{hits}}^{\text{dE/dx}}$  is used to ensure that selected tracks have sufficient energy loss information for particle identification purposes. In this analysis tracks without identification are required to have  $p_T > 0.2$  GeV/c and  $|\eta| < 0.7$  due to high track reconstruction and TOF matching efficiencies in this region. For the identified particle-antiparticle ratio analysis, where charged pions, charged kaons and (anti)protons are measured, the  $p_T$  cut was increased for kaons and (anti)protons to 0.3 and 0.4 GeV/c, respectively. The distributions of the  $\text{DCA}_{xy}$ ,  $|\text{DCA}_z|$ ,  $d_0$ ,  $N_{\text{hits}}^{\text{fit}}$  and  $N_{\text{hits}}^{\text{dE/dx}}$  qualities together with applied cuts are shown in Fig. 3.3, while the  $p_T$ ,  $\eta$  and the azimuthal angle,  $\phi$ , of the reconstructed tracks are shown in Fig. 3.4. Data are compared to embedded PYTHIA 8 sample. The azimuthal angle of the reconstructed tracks for runs  $\leq 16073050$  is not described by PYTHIA 8, therefore, additional data-driven corrections to track efficiencies are used [1].



### 3.3 Fiducial Region of the Measurement

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

The measurements were performed in a fiducial phase space of the forward scattered protons of  $0.04 < -t < 0.16$  GeV $^2/c^2$  and  $0.02 < \xi < 0.2$ . Figure 3.5 shows that the efficiency to events containing at least two primary charged particles,  $\epsilon_{n_{ch} \geq 2}(\log_{10} \xi)$  is reduced by half for  $\xi < 0.02$  compared to the region of larger  $\xi$ . In addition, the accidental background contribution in this region of  $\xi$  is significant and approximately equal to 20% (Sec. 4.1). For these reasons the lower  $\xi$  cut was introduced. On the other hand, the upper  $\xi$  cut was required since the region of larger  $\xi$  is dominated by Double Diffraction (DD) and ND (Sec. 5). All measured observables are presented in three  $\xi$  regions:  $0.02 < \xi < 0.05$ ,  $0.05 < \xi < 0.1$  and  $0.1 < \xi < 0.2$ .

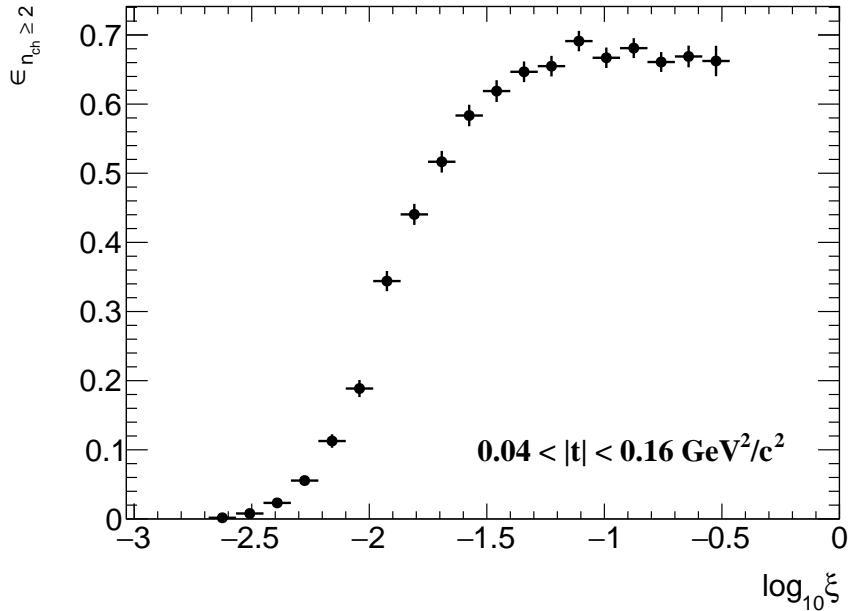


Figure 3.5:  $\epsilon_{n_{ch} \geq 2}$  as a function of  $\log_{10} \xi$  calculated from PYTHIA 8 (MBR).

# 4. Background Contribution

The background contribution to the charged-particle distributions can be broken down into event-level and track-level backgrounds, which are described in detail in the following sections:

- 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.

## 4.1 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 a signal in the TPC-TOF detector. This type of background may come from the overlap of a signal in RP (proton from beamhalo, 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 (mainly ND events without forward proton).

The accidental background contribution was calculated from Zerobias data, where two signatures of such background were investigated: the reconstructed proton in RP and the reconstruction of 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 RP 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 under study.

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

$$N = N(P, S) + N(R, S) + N(P, T) + N(R, T) \quad (4.1)$$

where:  $N(P, S)$  is the number of events with reconstructed proton in exactly one RP and reconstructed TOF vertex,  $N(R, S)$  is the number of events with no trigger in any RP and reconstructed TOF vertex,  $N(P, T)$  is the number of events with reconstructed proton in exactly one RP and no reconstructed TOF vertex,  $N(R, T)$  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:

$$N(P, S) = N(p_3 + p_1 p_2) \quad (4.2)$$

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

268 a reconstructed TOF vertex,  $p_3$  is the probability that there is a reconstructed proton in RP and  
 269 there is a reconstructed TOF vertex (not accidental).

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

$$\begin{aligned} N(R, S) &= N(1 - p_1)p_2(1 - p_3) \\ N(P, T) &= N(1 - p_2)p_1(1 - p_3) \\ N(R, T) &= N(1 - p_1)(1 - p_2)(1 - p_3) \end{aligned} \quad (4.3)$$

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

$$A_{\text{bkg}}^{\text{accidental}} = \frac{p_1 p_2}{p_3 + p_1 p_2} = \frac{N(R, S)N(P, T)N}{N(R)N(T)N(P, S)} \quad (4.4)$$

273 where:  $N(R) = N(R, S) + N(R, T) = N(P, T) + N(R, T)$ .

274 The shapes of the accidental background related to TPC distributions come from the above  
 275 Zerobias data events which pass all the analysis selection except having no trigger in any RP.  
 276 The templates corresponding to RP distributions are from protons in the above data sets but  
 277 with no reconstructed TOF vertex. The normalization is given by Eq. (4.4). Figure 4.1 shows  
 278 distributions of the reconstructed  $\xi$  with the accidental background contribution for events with  
 279 proton reconstructed in EU, ED, WU and WD arms.

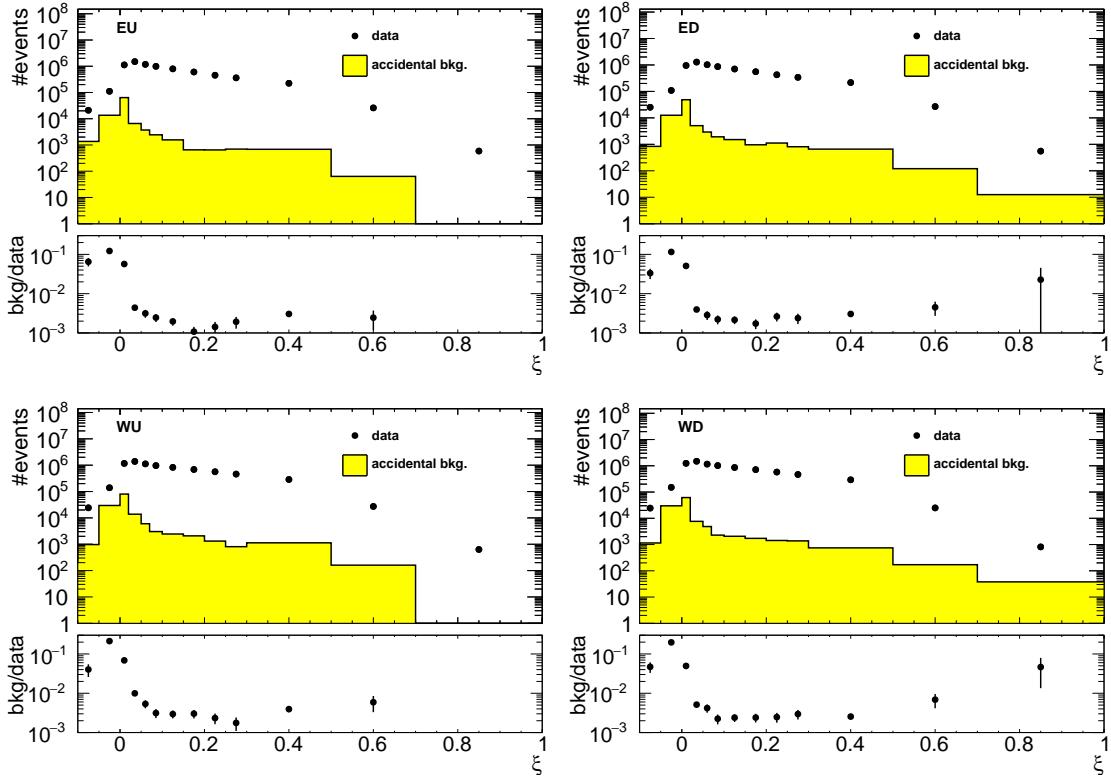


Figure 4.1: Uncorrected distributions of the reconstructed  $\xi$  for events with proton reconstructed in (top left) EU, (top right) ED, (bottom left) WU and (bottom right) WD arms. Data is shown as black markers, whereas the accidental background contribution is shown as yellow histogram. The ratio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

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

- 283     1. no trigger in any RP or trigger in exactly one arm (two RPs) with *no more* than one  
 284       reconstructed proton track in that arm   
 285     2. no or exactly one reconstructed TOF vertex (*without any additional requirement* ). The  
 286       requirement of signal in BBC small tiles remains unchanged.

287 As a result of this change in the procedure, as shown in Fig. 4.2, the accidental background  
 288 normalization increases of about 50% with respect to the nominal value. Therefore, the background  
 289 changes by  $\pm 50\%$  was taken as a systematic uncertainty related to the accidentals.

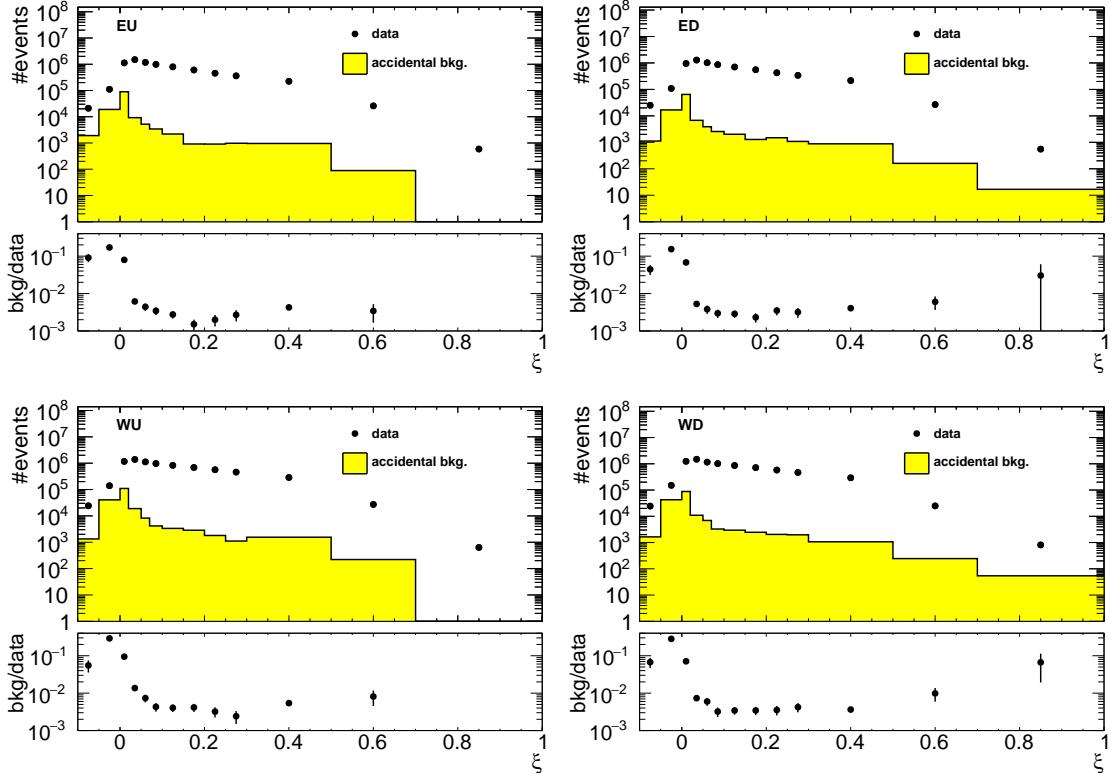


Figure 4.2: Uncorrected distributions of the reconstructed  $\xi$  for events with proton reconstructed in (top left) EU, (top right) ED, (bottom left) WU and (bottom right) WD arms. The accidental background contribution calculated with changed Zerobias selection criteria is also shown.

## 290     4.2 Background from Non-Primary Tracks

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

- 293       • decays of short-lived primary particles with strange quark content (mostly  $K^0, \Lambda^0$ ),  
 294       • photons from  $\pi^0$  and  $\eta$  decays which are converting to  $e^+ e^-$ ,  
 295       • hadronic interactions of particles with the beam-pipe or detector dead material.

296     There is also a contribution from fake tracks, i.e. tracks not associated with true-level particles,  
 297       coming from out-of-time pile-up or formed by a random combination of TPC hits. Figure 4.3  
 298       shows the background  $f_{\text{bkg}}(p_T, \eta)$  and fake track  $f_{\text{fake}}(p_T, \eta)$  contribution to reconstructed tracks  
 299       as a function of  $p_T$  and  $\eta$ . There were no differences observed in the background contribution in

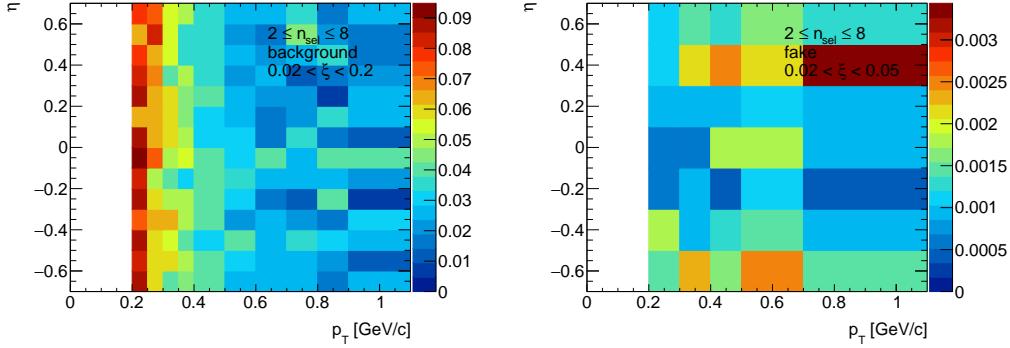


Figure 4.3: (left) Distribution of fraction of selected tracks associated with non-primary particles in the range  $0.02 < \xi < 0.2$  and (right) distribution of fraction of tracks which are not associated with true-level particles in the range  $0.02 < \xi < 0.05$  as predicted by PYTHIA 8 4C (SaS) embedding.

different  $\xi$  ranges, hence, all three  $\xi$  ranges were merged for this study. The highest background fraction, which varies between 5 – 10%, was found to be at low  $p_T$ . Due to too low statistics in PYTHIA 8 embedding MC, the shape of the fake track contribution was assumed to be the same in all three  $\xi$  ranges. However, its normalization was calculated for each  $\xi$  range separately with a ratio between the ranges of 1 : 0.74 : 1.11. The change by  $\pm 100\%$  in fake track contribution was taken as a systematic uncertainty.

### Proton Background

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

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

In order to correct for the knock-out background protons, sample enriched in proton background was used for background normalization, where  $DCA_{xy}$ ,  $DCA_z$  and  $d_0$  cuts were abandoned. Additionally, at least one, instead of exactly one, reconstructed vertex was allowed in this sample. Figures 4.4 and 4.5 show the DCA distributions of protons and antiprotons, respectively, for nominal (bottom) and background enriched (top) samples. The distributions for other  $p_T$  and  $\xi$  regions are shown in Appendix B. The protons and antiprotons are selected by a  $dE/dx$  cut of  $-1 < n\sigma_{p,\bar{p}} < 3$  where  $n\sigma_{p,\bar{p}}$  is given by Eq. (8.10). The fraction of knock-out protons within the selected sample is determined via MC template fits. The templates of reconstructed tracks with  $dE/dx$  corresponding to the proton and antiproton are obtained from PYTHIA 8 embedding MC separately for:

- primary (anti)protons,
- knock-out background protons (labeled as dead-material),
- fake tracks,
- secondary particles with  $dE/dx$  of (anti)proton (labeled as wrong PID - sec.),

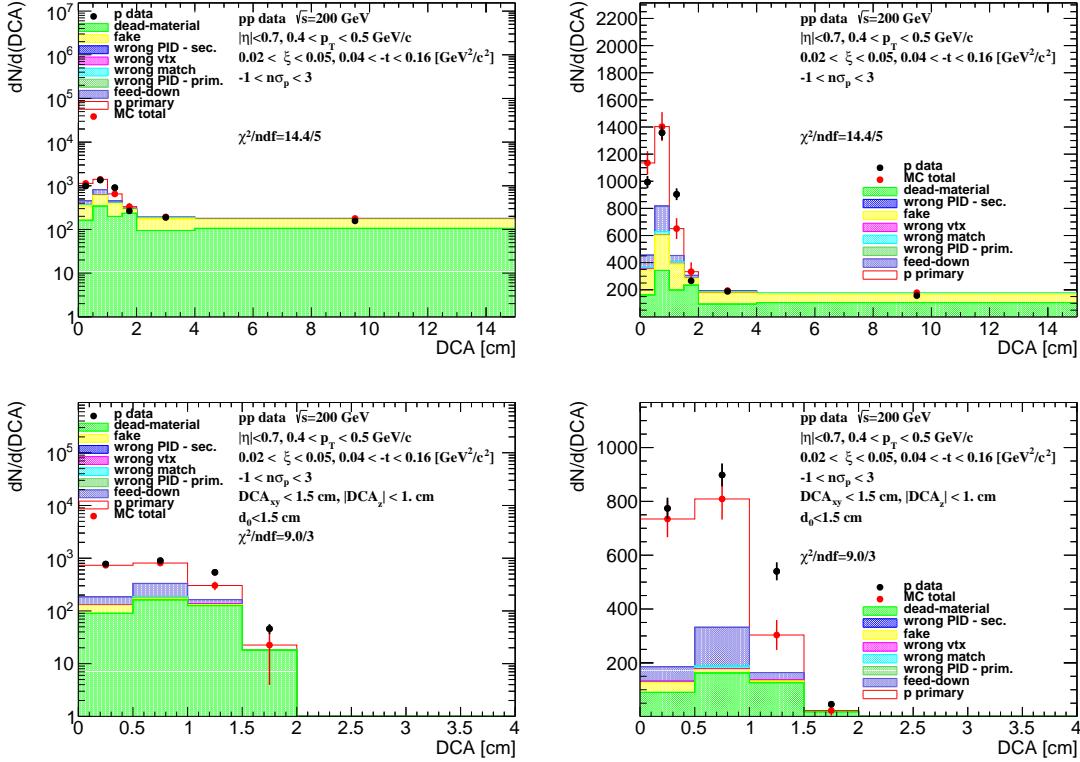


Figure 4.4: The DCA distributions of protons for  $0.4 < p_T < 0.5 \text{ GeV}/c$  shown for single range of  $0.02 < \xi < 0.05$  (shown in log and linear scale in left and right column, respectively). The MC contributions are shown after scaling the dead-material template to the tail of large DCA values,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

- 331     • tracks associated with primary (anti)protons, but with the reconstructed vertex not matched  
332        to true-level primary vertex (labeled as wrong vtx),
- 333     • reconstructed track is partially matched to true-level particle (labeled as wrong match, track  
334        to true-level particle matching is described in [1]), i.e. track and true-level particle have  
335        appropriate number of common hit points but the distance between true-level particle and  
336        track is too large,  $\delta^2(\eta, \phi) > (0.15)^2$ , thus, track is not considered as primary particle  
337        according to discussion in [1],
- 338     • primary particles with  $dE/dx$  of (anti)proton (labeled as wrong PID - prim.),
- 339     • (anti)proton as a product of short-lived decays, mainly  $\Lambda^0$  (labeled as feed-down).

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

345     The fraction of the knock-out proton background in the signal region,  $\text{DCA} < 1.5$ , was es-  
346     timated from the nominal sample (Fig. 4.4, bottom), where  $\text{DCA}_{xy}$ ,  $\text{DCA}_z$  and  $d_0$  track cuts  
347     were applied and exactly one reconstructed vertex was required. The normalization of each MC  
348     contribution was kept the same as that estimated for the background enriched sample. Figure 4.6

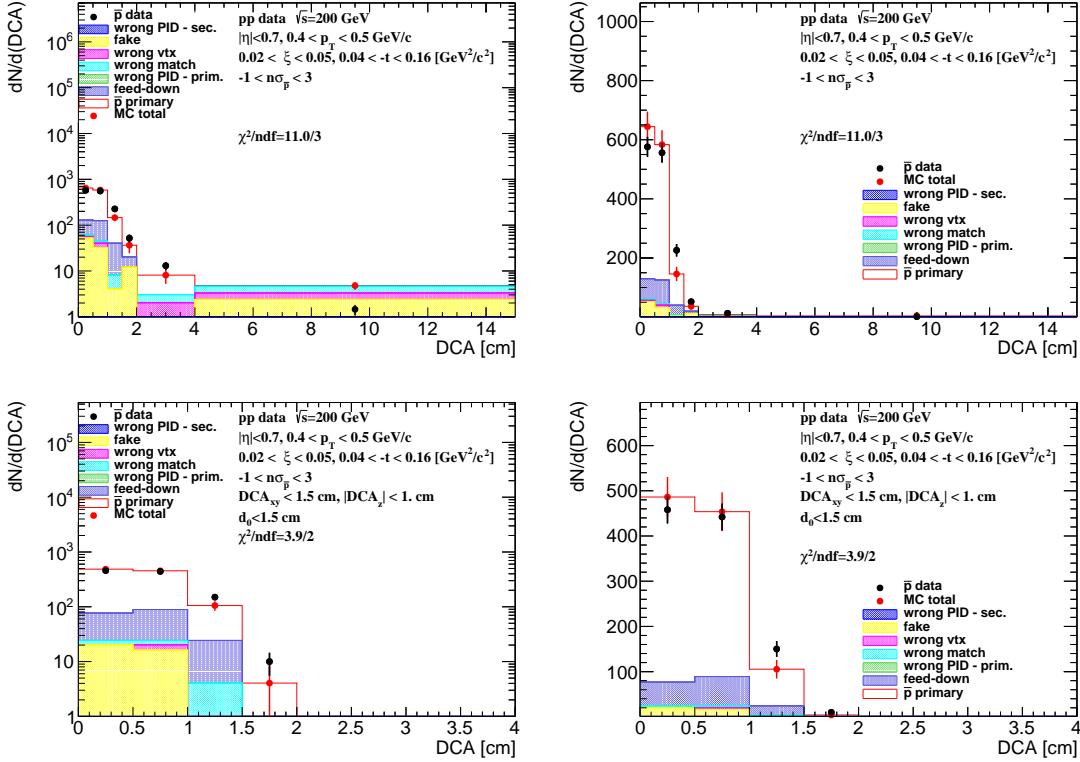


Figure 4.5: The DCA distributions of antiprotons for  $0.4 < p_T < 0.5 \text{ GeV}/c$  shown for one range of  $0.02 < \xi < 0.05$  (log and linear scale in left and right column, respectively). The MC contributions are shown as colour histograms. (top) Background enriched and (bottom) nominal samples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

349 shows the knock-out proton background as a function of  $p_T$  in three ranges of  $\xi$ . The following  
 350 functional form was found to describe the background protons:

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

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

352 The obtained fraction of knock-out background protons is approximately 20% at  $p_T = 0.45$   
 353  $\text{GeV}/c$  and less than 10% at  $p_T = 1.0 \text{ GeV}/c$ . In PYTHIA 8 SD predictions (also shown in Fig. 4.4),  
 354 such fraction is much smaller and equals to approximately 7% at  $p_T = 0.45 \text{ GeV}/c$  and about 5%  
 355 at  $p_T = 1.0 \text{ GeV}/c$ . This may suggest that there are differences in the amount of dead material in  
 356 front of TPC between data and simulation, which is consistent with the studies presented in [1].

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

### 362 Systematic Uncertainty Related to Proton Background

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

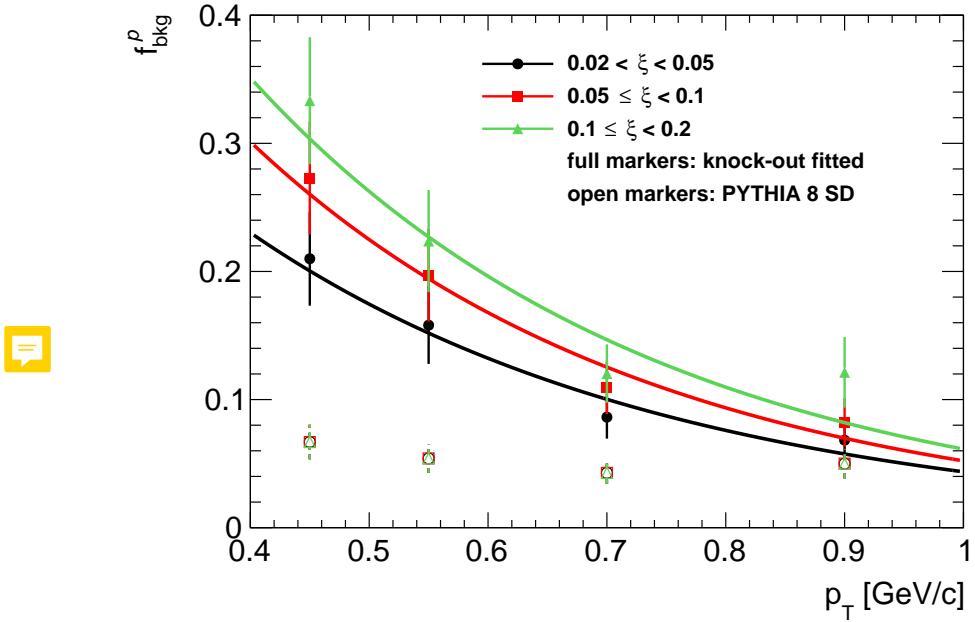


Figure 4.6: The fraction of knock-out proton background as a function of  $p_T$  in three ranges of  $\xi$  with fitted parametrizations. Full markers represent fitted knock-out background and open markers represent PYTHIA 8 SD predictions.

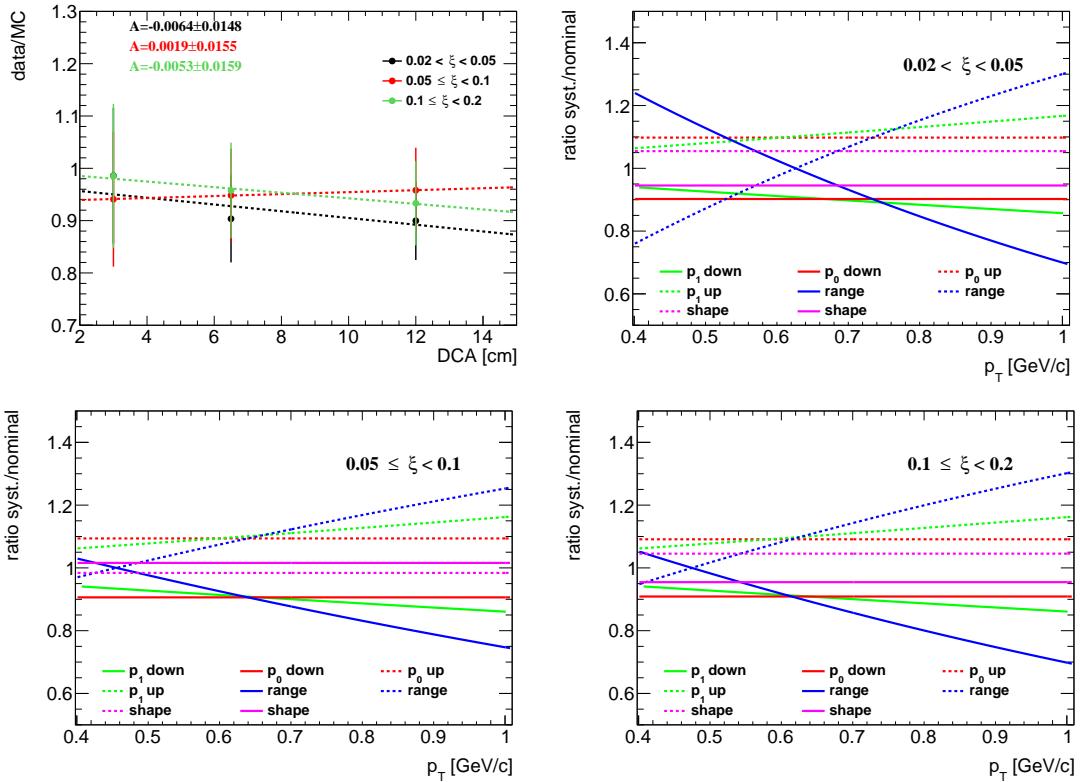


Figure 4.7: (top left) Data to MC ratio of the number of events in the background dominated region in three ranges of  $\xi$  with fitted functional form given by Eq. (4.6). (top right and bottom) Components of the systematic uncertainty related to the knock-out background protons contribution in three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

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

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

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

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

381     All above components of the systematic uncertainty related to the knock-out proton back-  
 382     ground, shown in Fig. 4.7, are added in quadrature. Those related to the fit range and the shape  
 383     of the proton background are symmetrized. Figure 4.8 shows the fraction of knock-out proton  
 384     background in three ranges of  $\xi$  and the total systematic uncertainty related to it.

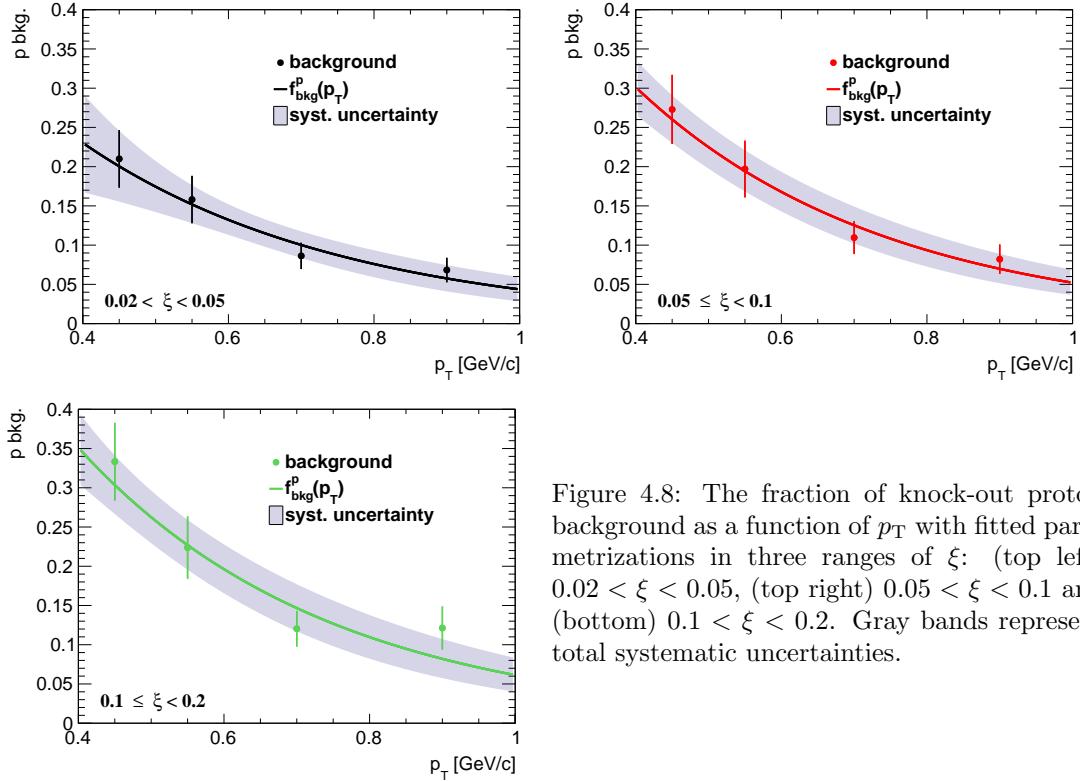


Figure 4.8: The fraction of knock-out proton background as a function of  $p_T$  with fitted parametrizations in three ranges of  $\xi$ : (top left)  $0.02 < \xi < 0.05$ , (top right)  $0.05 < \xi < 0.1$  and (bottom)  $0.1 < \xi < 0.2$ . Gray bands represent total systematic uncertainties.

### 385     Pion Background

386     The pion spectra are corrected for weak decays (mainly of  $K_S^0$  and  $\Lambda^0$ ), muon contribution and  
 387     background from the detector dead-material interactions. The pion decay muons can be identified  
 388     as pions due to the similar masses. These contributions are obtained from PYTHIA 8 SD. Fig-  
 389     ure 4.9 shows the background contribution to the pion spectra as a function of  $p_T$  in three ranges of

<sup>390</sup>  $\xi$ , separately for  $\pi^-$  and  $\pi^+$ . Since there were negligible differences observed between these three  
<sup>391</sup> ranges of  $\xi$ , the background contribution was averaged over  $\xi$ . The following parametrization was  
<sup>392</sup> found to describe it:

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

<sup>393</sup> where  $a_i$ ,  $i = 0, \dots, 3$  are free parameters of the fitted function.

<sup>394</sup> The pion background contribution varies between 5% at low- $p_T$  ( $p_T = 0.25$  GeV/c) and about  
<sup>395</sup> 1% at  $p_T = 1.0$  GeV/c for both negatively and positively charged pions. In addition, the back-  
<sup>396</sup> ground was calculated from EPOS SD+SD' for the full range of  $\xi$ . The differences between  
 PYTHIA 8 and EPOS are up to 1% for  $\pi^-$ .

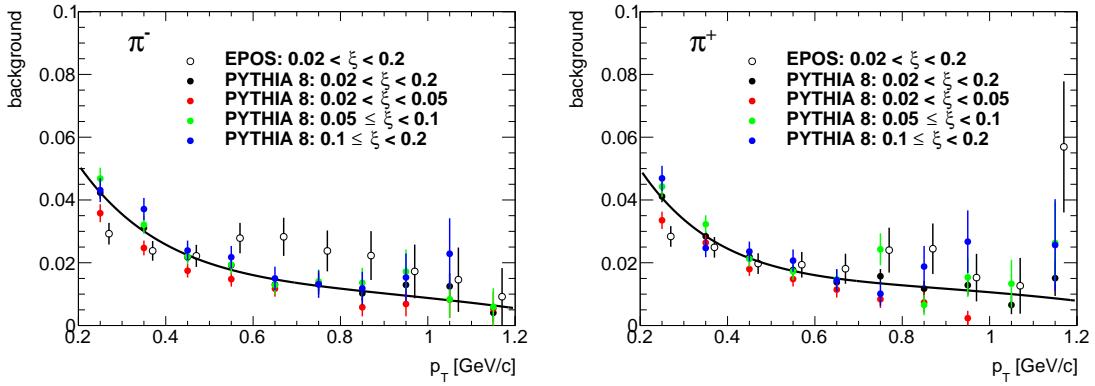


Figure 4.9: Pion background fraction as a function of  $p_T$  calculated from PYTHIA 8 and shown separately for (left) negatively and (right) positively charged pions in three ranges of  $\xi$ : (red)  $0.02 < \xi < 0.05$ , (green)  $0.05 < \xi < 0.1$ , (blue)  $0.1 < \xi < 0.2$ . (full black points) The pion background averaged over three ranges of  $\xi$  with fitted parametrization is also shown. Open black points represent EPOS predictions for the full  $\xi$  range.

<sup>397</sup>

## 5. Non-SD Contributions

Events, in which forward proton and reconstructed TOF vertex are the result of the same  $pp$  interaction, may originate from ND, DD, SD, and CD processes. Their relative contributions must be estimated from MC models and are therefore model dependent. Tracks reconstructed in RPs may also be:

- forward protons produced in the SD, CD or DD diffractive systems or from ND events,
- secondary particles from showering initiated by interaction of forward protons with beam-line elements. This contribution is equal to approximately 2%.

Figure 5.1 shows the uncorrected  $\xi$  and  $t$  distributions in data compared to various MC models: PYTHIA 8 A2 (MBR), PYTHIA 8 A2 (MBR-tuned) 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-tuned) predictions agree much better with the data than PYTHIA 8 A2 (MBR) and result also in a suppression of non-SD events. EPOS describes data better than PYTHIA 8 but shows a dominant contribution of SD' events. All MCs predict significant DD and ND background at large  $\xi$ , thereby the analysis was limited to  $\xi < 0.2$ .

Figures 5.2 to 5.4 show the uncorrected distributions of variables used in the later analysis:  $n_{\text{sel}}$ ,  $p_T$  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). 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 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). As a consequence, results are compared with the sum of SD and SD' processes for EPOS model.

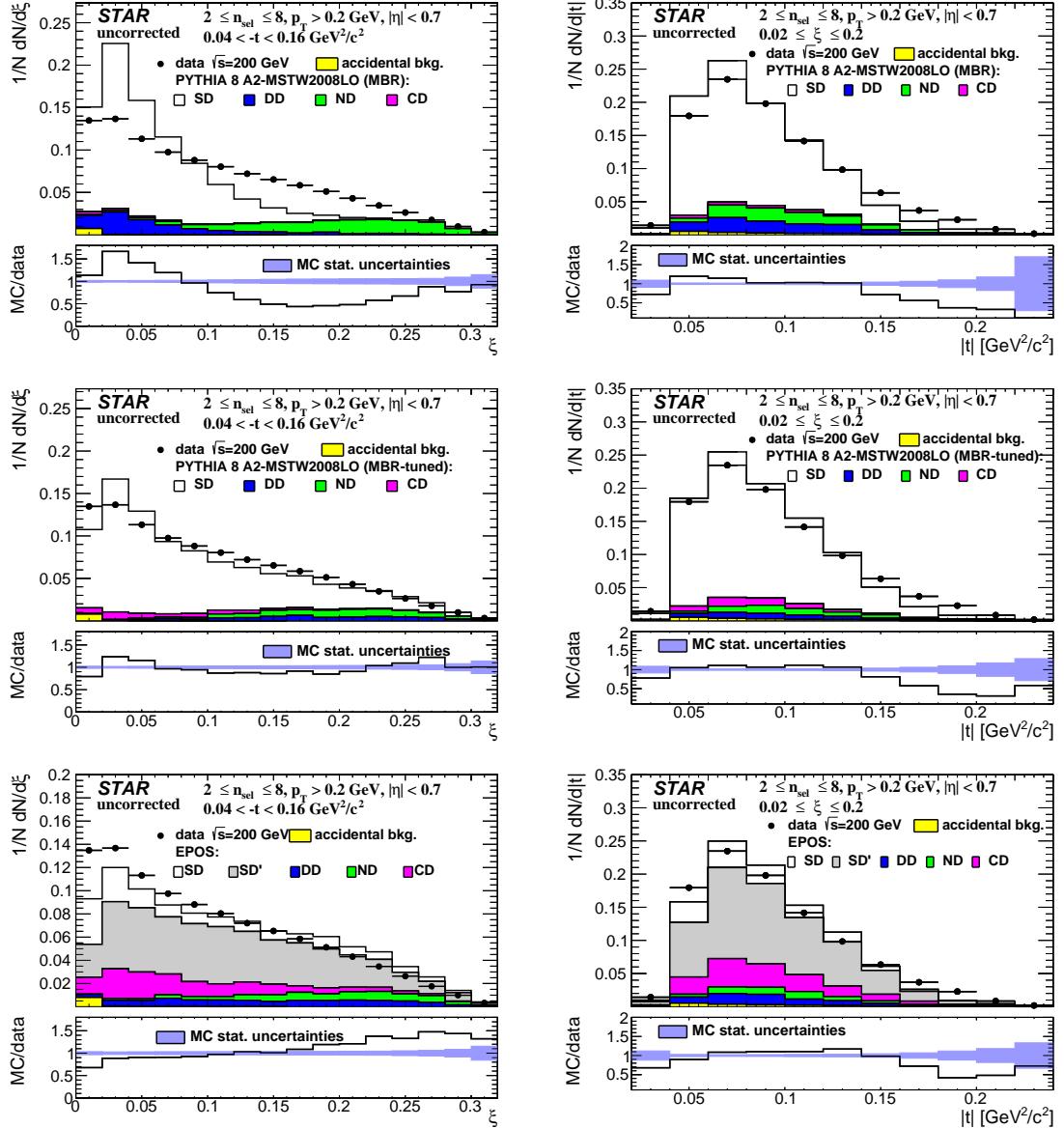


Figure 5.1: Uncorrected distributions of data compared to various MC models: (top) PYTHIA 8 A2 (MBR), (middle) PYTHIA 8 A2 (MBR-tuned) and (bottom) EPOS, as a function of (left column)  $\xi$  and (right column)  $|t|$ .

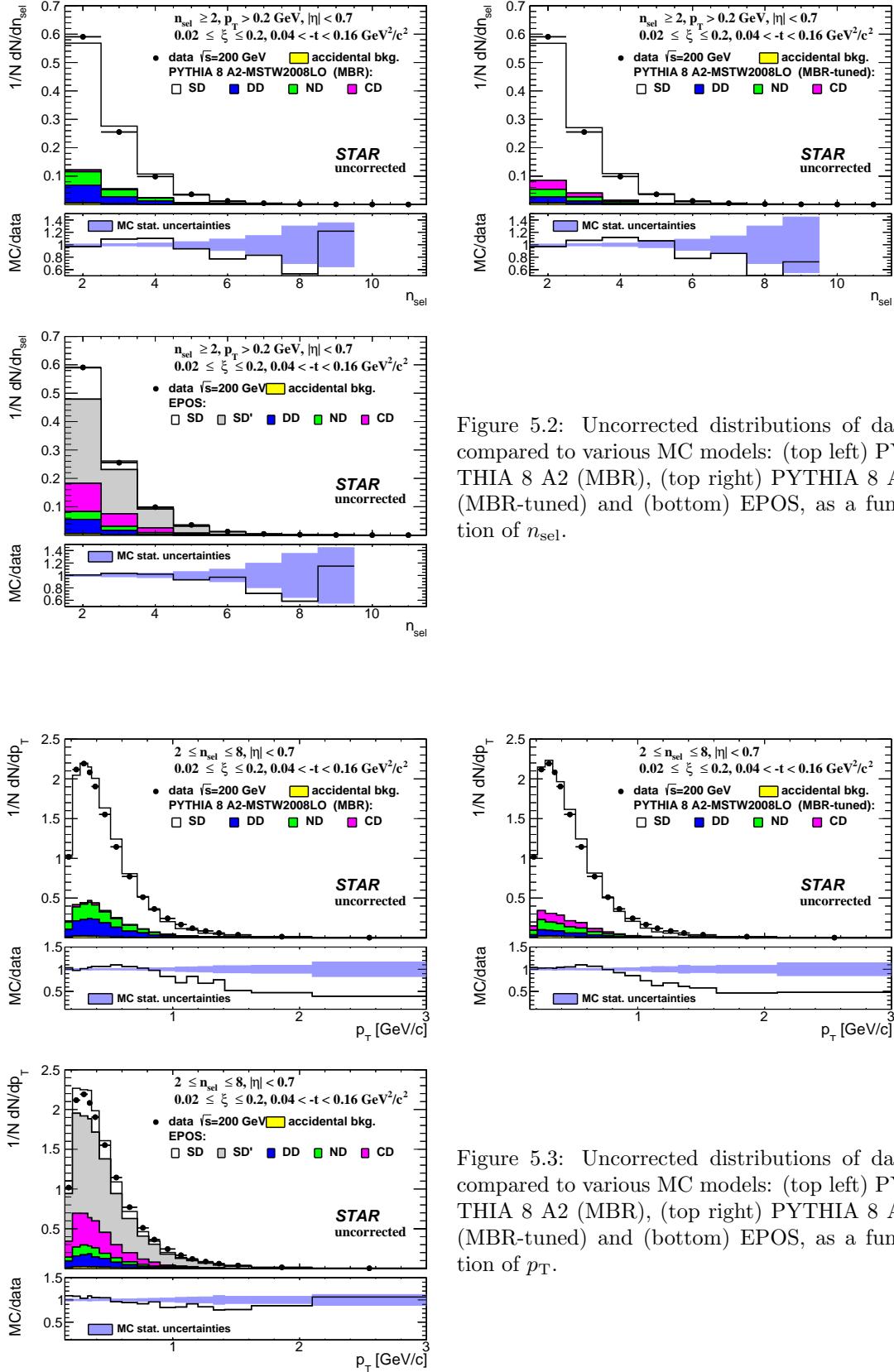


Figure 5.2: Uncorrected distributions of data compared to various MC models: (top left) PYTHIA 8 A2 (MBR), (top right) PYTHIA 8 A2 (MBR-tuned) and (bottom) EPOS, as a function of  $n_{\text{sel}}$ .

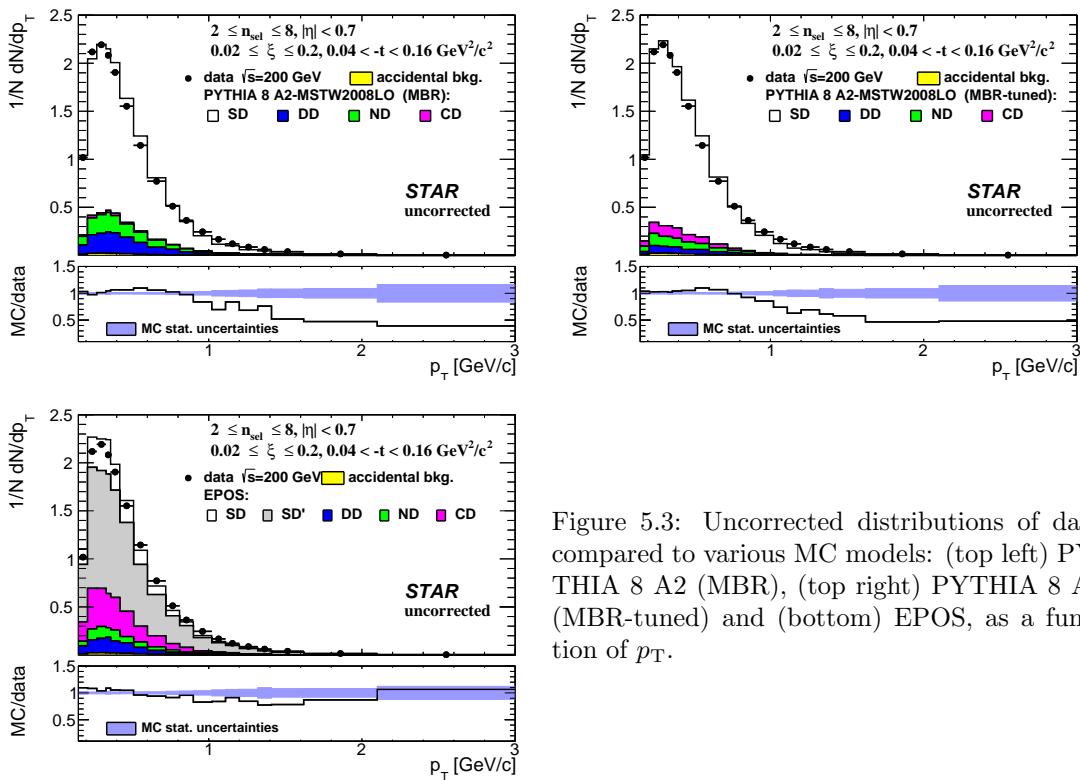


Figure 5.3: Uncorrected distributions of data compared to various MC models: (top left) PYTHIA 8 A2 (MBR), (top right) PYTHIA 8 A2 (MBR-tuned) and (bottom) EPOS, as a function of  $p_T$ .

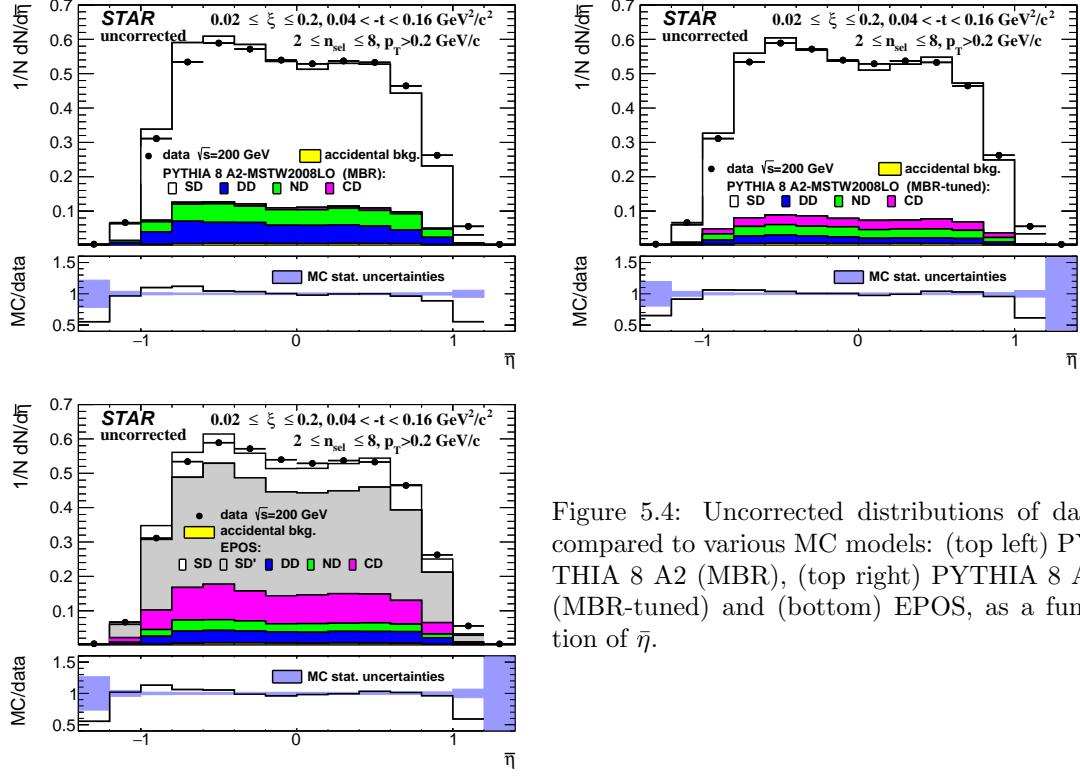


Figure 5.4: Uncorrected distributions of data compared to various MC models: (top left) PYTHIA 8 A2 (MBR), (top right) PYTHIA 8 A2 (MBR-tuned) and (bottom) EPOS, as a function of  $\bar{\eta}$ .

# 6. Selection Efficiencies

## 6.1 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 and 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 3.2, except vertex-related cuts on  $\text{DCA}_{xy}$  and  $\text{DCA}_z$ , 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  $\xi$  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  $p_{\text{T}}$  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:

$$\epsilon_{\text{virt}}(n_{\text{virt}}^{\text{global}}) = \epsilon_{\text{virt}}^{\text{best}}(n_{\text{virt}}^{\text{global}}) + \delta_{\text{virt}}^{\text{fake}}(n_{\text{virt}}^{\text{global}}) \quad (6.1)$$

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.

463 The vertex-finding efficiency as a function of  $n_{\text{vrt}}^{\text{global}}$  is shown in Fig. 6.1 (left) and it is larger  
 464 than 75% for all  $n_{\text{vrt}}^{\text{global}}$ . However, for  $n_{\text{vrt}}^{\text{global}} > 8$ , there are more fake than true-level primary  
 465 vertices. When there are exactly two global tracks used in the vertex reconstruction,  $n_{\text{vrt}}^{\text{global}} = 2$ ,  
 466 the vertex-finding efficiency depends on the longitudinal distance between these tracks  $|\Delta z_0|$ .  
 467 Therefore, the vertex finding efficiency for such events  $\epsilon_{\text{vrt}}(|\Delta z_0|)$  is given by:

$$\epsilon_{\text{vrt}}(|\Delta z_0|) = \epsilon_{\text{vrt}}^{\text{best}}(|\Delta z_0|) + \delta_{\text{vrt}}^{\text{fake}}(|\Delta z_0|) \quad (6.2)$$

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

470 Figure 6.1 (right) shows the vertex finding efficiency for events with  $n_{\text{vrt}}^{\text{global}} = 2$ . This efficiency  
 471 is smaller than 20% for tracks with  $|\Delta z_0| > 2$  cm, hence the analysis was limited to events with  
 472  $|\Delta z_0| < 2$  cm, when  $n_{\text{vrt}}^{\text{global}} = 2$ . The rate of fake vertices is negligibly low (open points overlap  
 with full points).

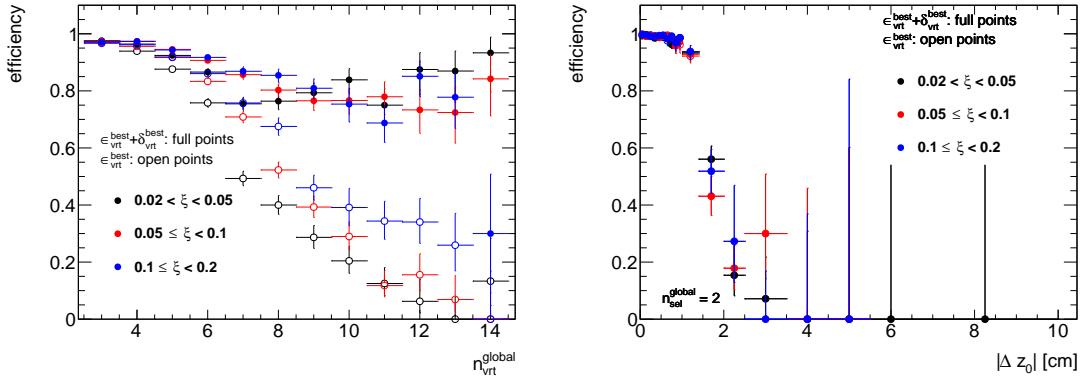


Figure 6.1: Vertex-finding efficiency in three ranges of  $\xi$  as a function of (left)  $n_{\text{vrt}}^{\text{global}}$  and (right)  
 with respect to the  $|\Delta z_0|$  between reconstructed tracks in events with  $n_{\text{vrt}}^{\text{global}} = 2$ .

473 Events are rejected if more vertices are reconstructed in addition to the best one. Rejected  
 474 events can be classified as:  
 475

- 476 a) two or more additional vertices,
- 477 b) additional secondary vertex from interactions with the detector dead-material,
- 478 c) additional fake vertex,
- 479 d) additional primary vertex (vertex splitting or background vertex reconstructed as best ver-  
 480 tex),
- 481 e) additional secondary vertex from the decay.

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

$$f_{\text{vrt}}^{\text{veto}}(n_{\text{vrt}}^{\text{global}}) = \frac{\text{number of events with more than one reconstructed TOF vertex}}{\text{number of events with at least one reconstructed TOF vertex}} \quad (6.3)$$

$$= f_a - f_b - f_c - f_d - f_e$$

483 where  $f_a$  to  $f_e$  are the fractions of events with additional vertices, whose labels are listed above.

484 As before, the fraction was calculated as a function of  $|\Delta z_0|$  for events with  $n_{\text{vrt}}^{\text{global}} = 2$ .  
 485 Figure 6.2 shows the fraction of multi-vertex events with respect to the  $n_{\text{vrt}}^{\text{global}}$ . There is a large  
 486 fraction of events (> 90%) with additional background vertices for  $n_{\text{vrt}}^{\text{global}} \geq 9$ , what would result

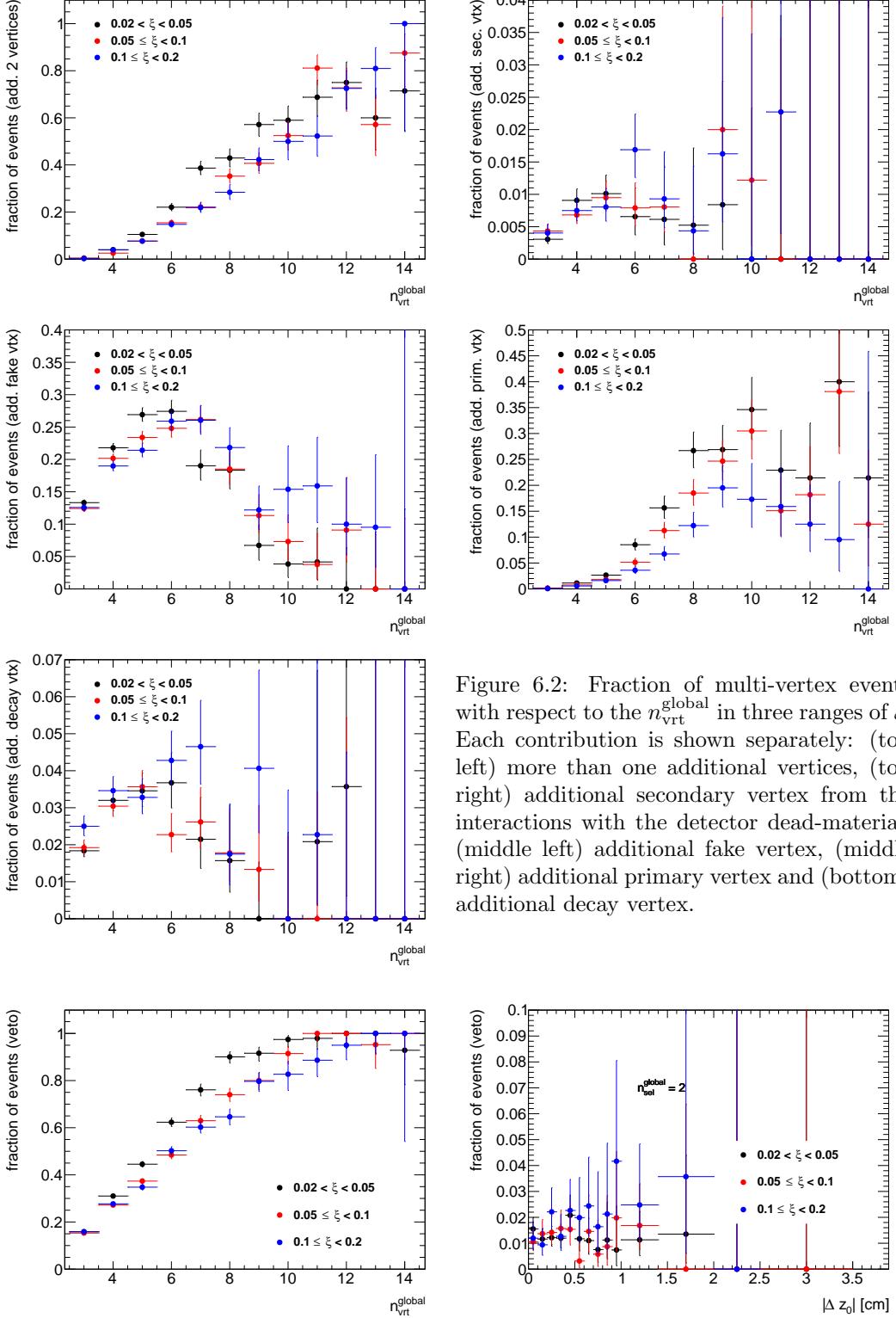


Figure 6.2: Fraction of multi-vertex events with respect to the  $n_{\text{vrt}}^{\text{global}}$  in three ranges of  $\xi$ . Each contribution is shown separately: (top left) more than one additional vertices, (top right) additional secondary vertex from the interactions with the detector dead-material, (middle left) additional fake vertex, (middle right) additional primary vertex and (bottom) additional decay vertex.

Figure 6.3: 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  $\xi$ .

487 in large correction factor. Hence, the analysis was limited to events with  $n_{\text{sel}} \leq 8$ . The total  
 488 fraction of multi-vertex events,  $f_a + f_b + f_c + f_d + f_e$ , as a function of  $n_{\text{vrt}}^{\text{global}}$  and  $|\Delta z_0|$ , shown in  
 489 Fig. 6.3, demonstrates that  $f_{\text{vrt}}^{\text{veto}}(|\Delta z_0|)$  is very small (< 2%) for events with  $n_{\text{vrt}}^{\text{global}} = 2$ .

490 Although, the analysis was limited to  $n_{\text{sel}} \leq 8$ , a fraction of events with additional background  
 491 vertices was still relatively large. Since most of these additional vertices are fake (and as accidental  
 492 not correlated with true-level primary distributions), it was checked whether the charged-particle  
 493 multiplicity distributions are different for events with and without reconstructed fake vertices.  
 494 These distributions, as shown in Fig 6.4, are in good agreement, thus, above studies of vertex  
 495 reconstruction were repeated using MC events that do not contain reconstructed fake vertices.  
 496 It means that events with additional fake vertex were rejected and no correction was applied  
 497 for such losses since it only affects overall normalization (not the shapes of distributions under study).  
 498 The vertex finding efficiency, which was calculated from such events, is shown in Fig. 6.5. It is  
 499 greater than 95% for events with  $2 \leq n_{\text{vrt}}^{\text{global}} \leq 8$ . In addition, the corresponding fraction of multi-  
 500 vertex events, shown in Figs. 6.6 and 6.7, is smaller than 20%. Since fake vertices were rejected  
 501 from this study, the  $f_c$  term from Eq. (6.3) is equal to 0. The correction factors calculated from  
 502 MC events that do not contain reconstructed fake vertices were used in the analysis instead of  
 503 the one obtained from the full MC sample.

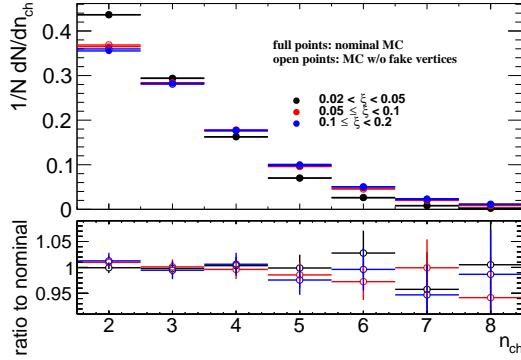


Figure 6.4: Normalized charged-particle multiplicity distributions in three ranges of  $\xi$  calculated from PYTHIA 8 SD embedding MC for (full points) all generated events and (open points) events without reconstructed fake vertices.

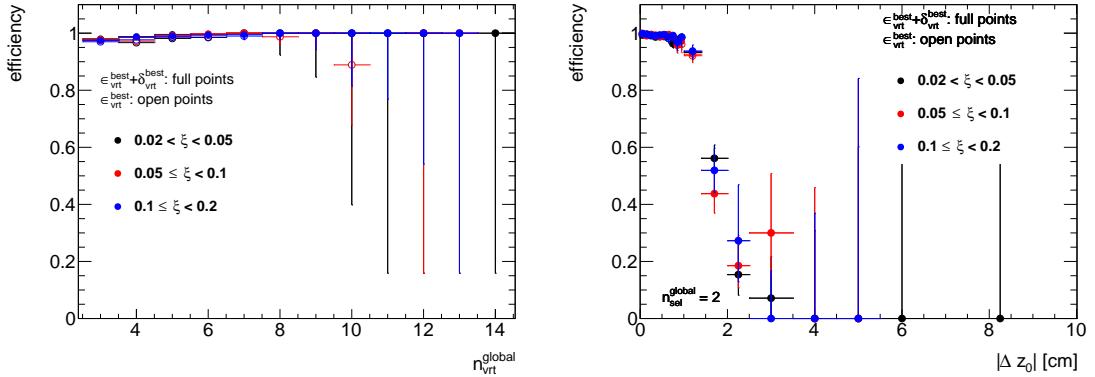


Figure 6.5: Vertex-finding efficiency in three ranges of  $\xi$  as a function of (left)  $n_{\text{vrt}}^{\text{global}}$  and (right)  
 with respect to the  $|\Delta z_0|$  between reconstructed tracks in events with  $n_{\text{vrt}}^{\text{global}} = 2$ . Only events  
 that do not contain additional fake vertices were used.

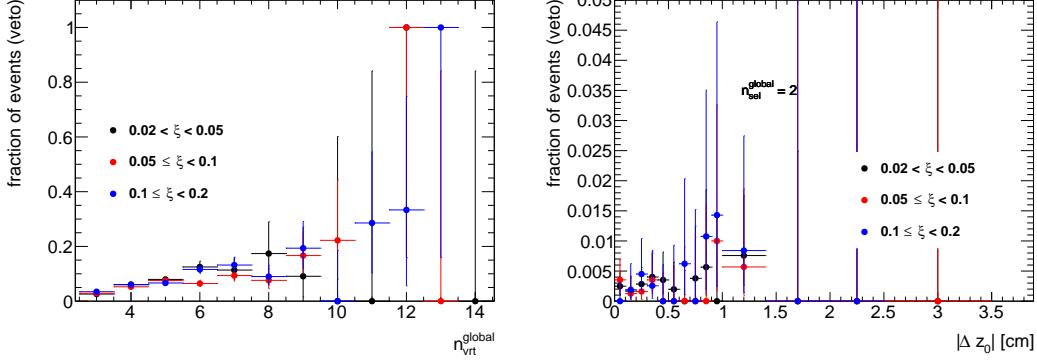


Figure 6.6: Total fraction of multi-vertex events as a function of (left)  $n_{\text{vrt}}^{\text{global}}$  for events with  $n_{\text{vrt}}^{\text{global}} > 2$  and (right)  $|\Delta z_0|$  for events with  $n_{\text{vrt}}^{\text{global}} = 2$  in three ranges of  $\xi$ . Only events that do not contain additional fake vertices were used.

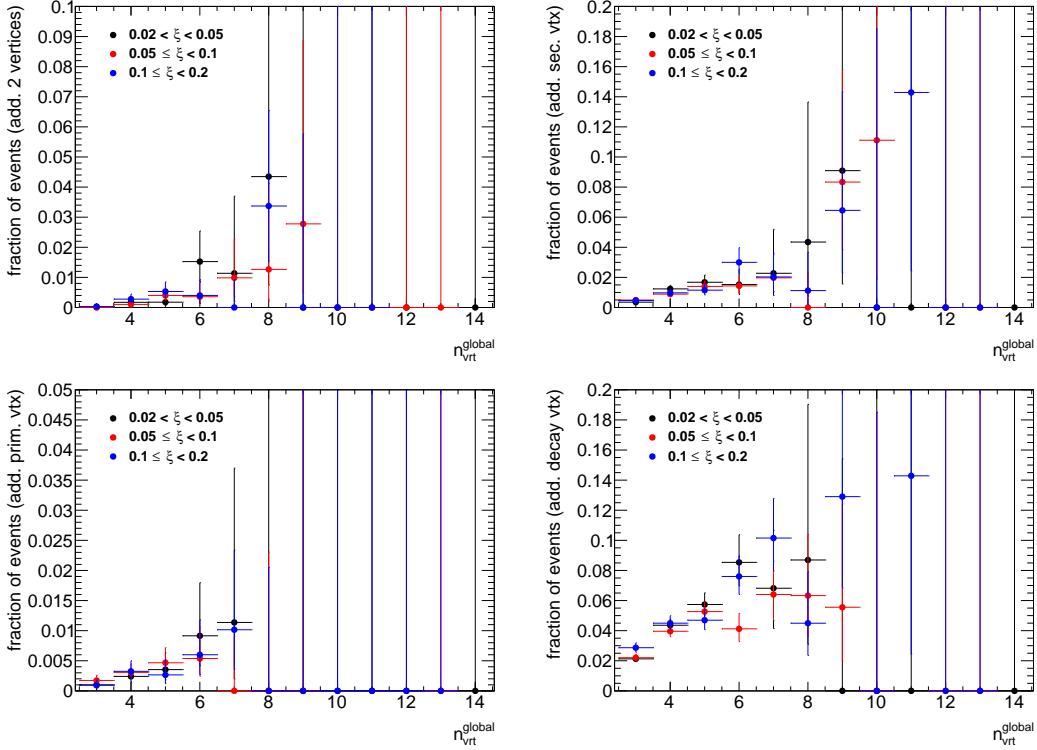


Figure 6.7: Fraction of multi-vertex events with respect to the  $n_{\text{vrt}}^{\text{global}}$  in three ranges of  $\xi$ . Each contribution is shown separately: (top left) more than one additional vertices, (top right) additional secondary vertex from the interactions with the detector dead-material, (bottom left) additional primary vertex and (bottom right) additional decay vertex. Only events that do not contain additional fake vertices were used.

## 504 6.2 Correction to BBC-Small

505 The SDT trigger conditions imposed signal in RPs and veto on any signal in the same-side small  
 506 BBC tiles, whereas signal in the opposite-side BBC-small was required by the offline event se-  
 507 lection. 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,  $\epsilon_{BBC}$ , was obtained as a function of each measured quantity using PYTHIA 8 4C (SaS) embedded into Zerobias data, EPOS SD+SD' and HERWIG MC. The efficiency was calculated for events within fiducial region as follows:

$$\epsilon_{BBC} = \frac{\text{number of MC events satysfying the BBC-small selection criteria}}{\text{number of MC events}} \quad (6.4)$$

Figures 6.8 to 6.10 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 dependents 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  $\xi$  in Fig. 6.11. The  $\epsilon_{BBC}$  strongly depends on  $\xi$  and varies from about 90% for events with  $\xi$  within 0.02 – 0.05 to about 60% for events with  $0.1 < \xi < 0.2$ . However, measurements of corrected  $\xi$  distributions are out of the scope of this analysis.

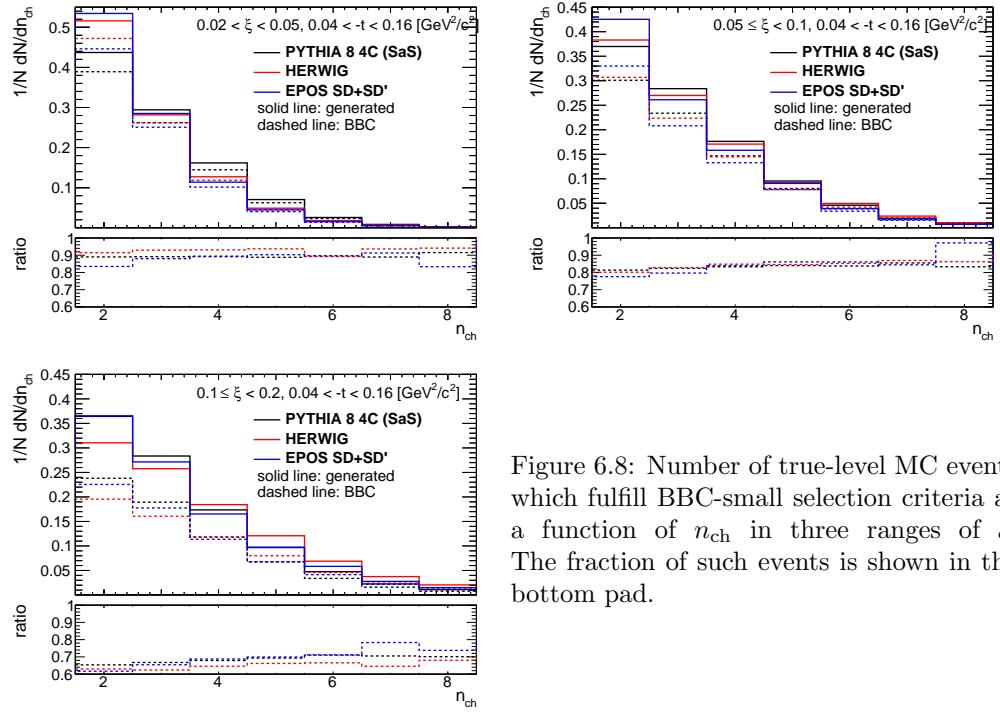


Figure 6.8: Number of true-level MC events which fulfill BBC-small selection criteria as a function of  $n_{ch}$  in three ranges of  $\xi$ . The fraction of such events is shown in the bottom pad.

Data is corrected for BBC-small efficiency using PYTHIA 8 4C (SaS). The uncertainty related to this correction is estimated by using HERWIG MC sample, where the hadronisation model is different from that used in PYTHIA 8. Figure 6.12 shows the PYTHIA 8 prediction on BBC efficiency divided by the HERWIG prediction in three ranges of  $\xi$ . 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 difference between these two hadronisation models is used as systematic uncertainty. For completeness, the differences between PYTHIA 8 and EPOS SD+SD' predictions are shown in Fig. 6.13. Most of them are of the order of 3%, except  $n_{ch} \leq 3$  for which the difference varies up to 6%.

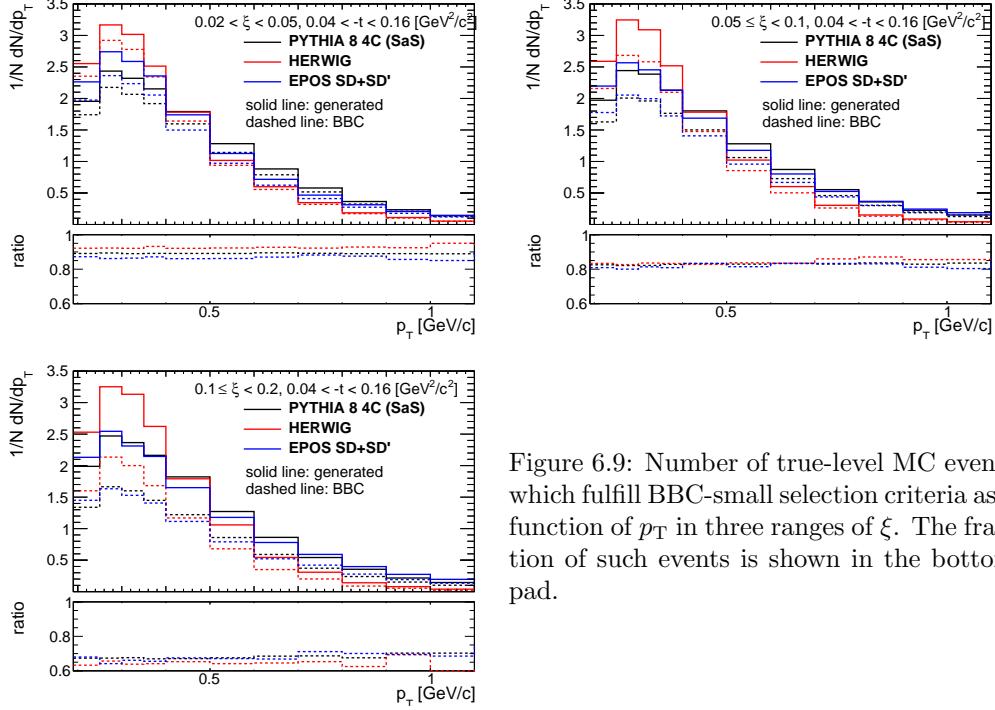


Figure 6.9: Number of true-level MC events which fulfill BBC-small selection criteria as a function of  $p_T$  in three ranges of  $\xi$ . The fraction of such events is shown in the bottom pad.

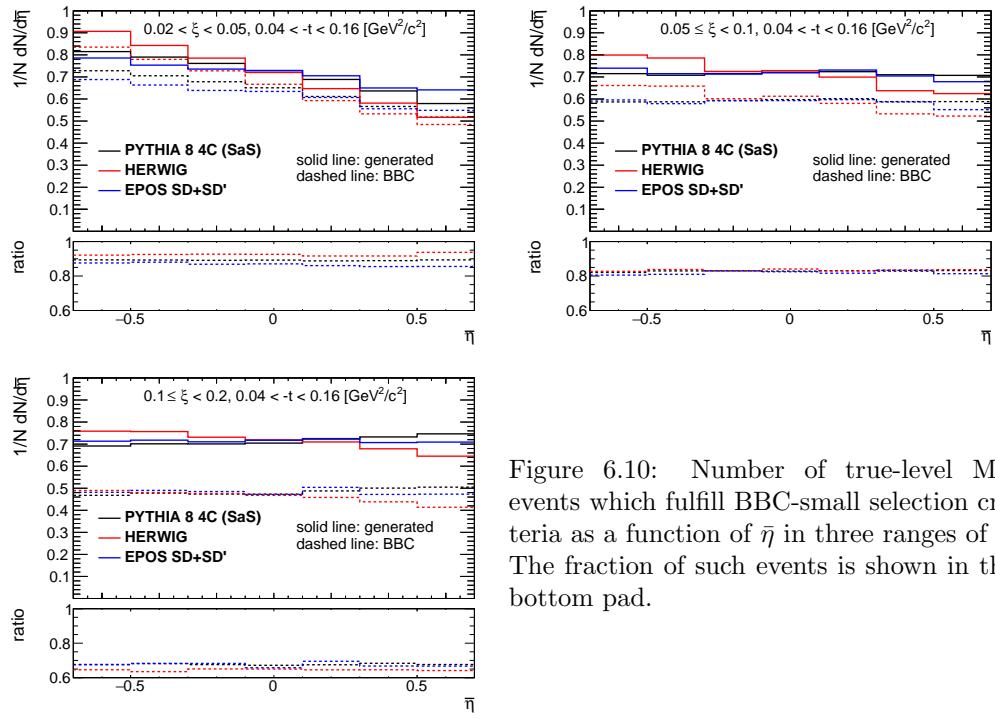


Figure 6.10: Number of true-level MC events which fulfill BBC-small selection criteria as a function of  $\bar{\eta}$  in three ranges of  $\xi$ . The fraction of such events is shown in the bottom pad.

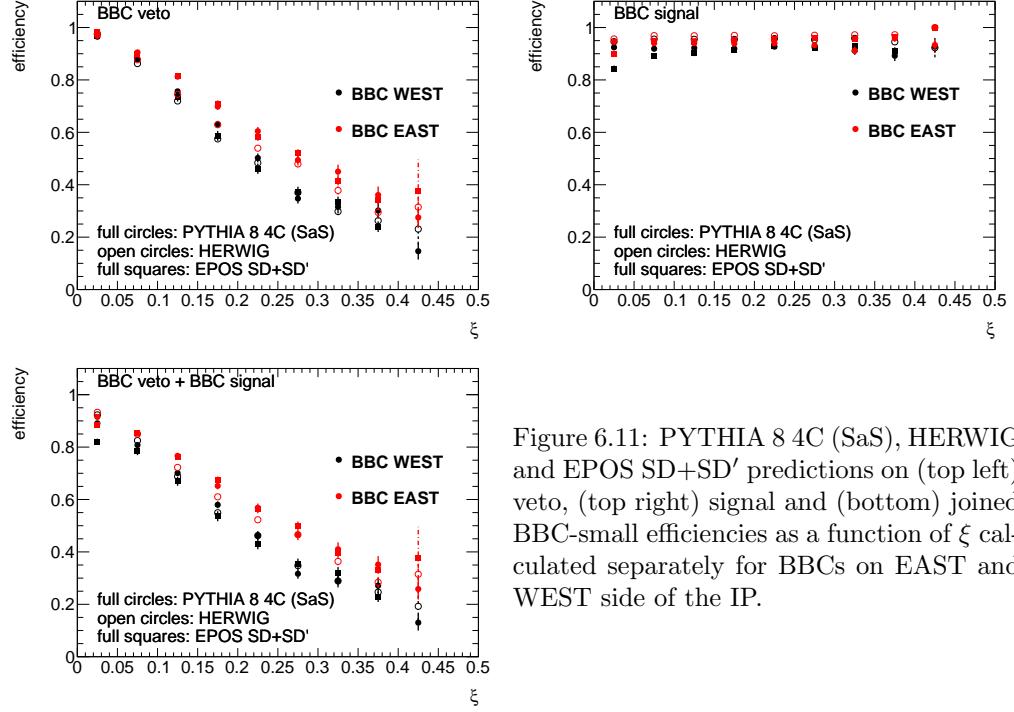


Figure 6.11: PYTHIA 8 4C (SaS), HERWIG and EPOS SD+SD' predictions on (top left) veto, (top right) signal and (bottom) joined BBC-small efficiencies as a function of  $\xi$  calculated separately for BBCs on EAST and WEST side of the IP.

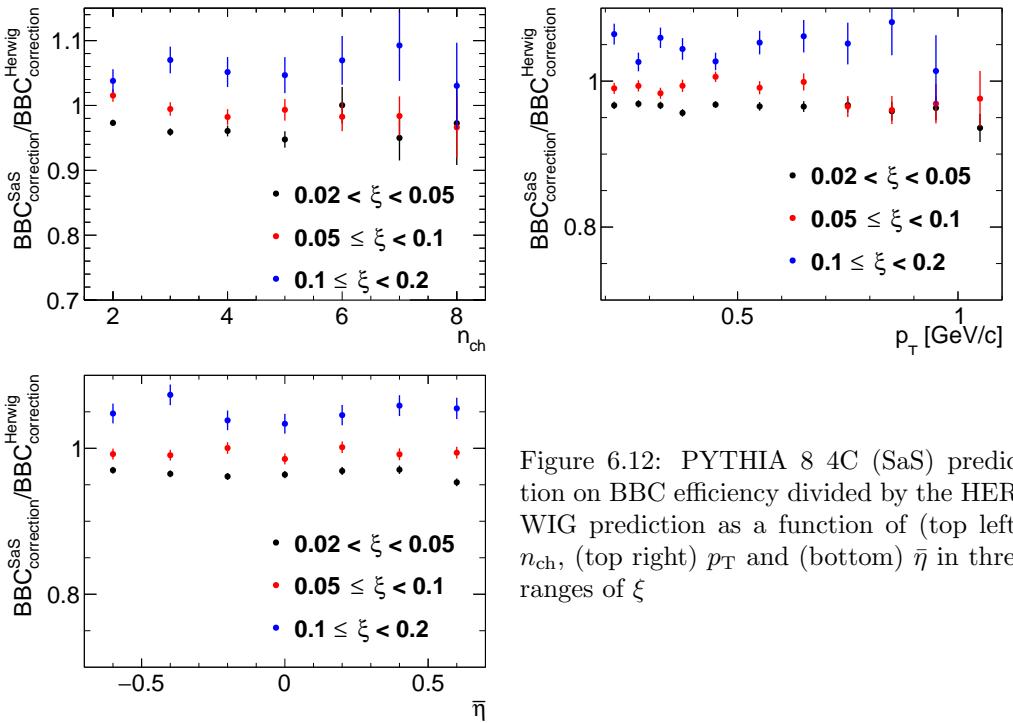


Figure 6.12: PYTHIA 8 4C (SaS) prediction on BBC efficiency divided by the HERWIG prediction as a function of (top left)  $n_{ch}$ , (top right)  $p_T$  and (bottom)  $\bar{\eta}$  in three ranges of  $\xi$

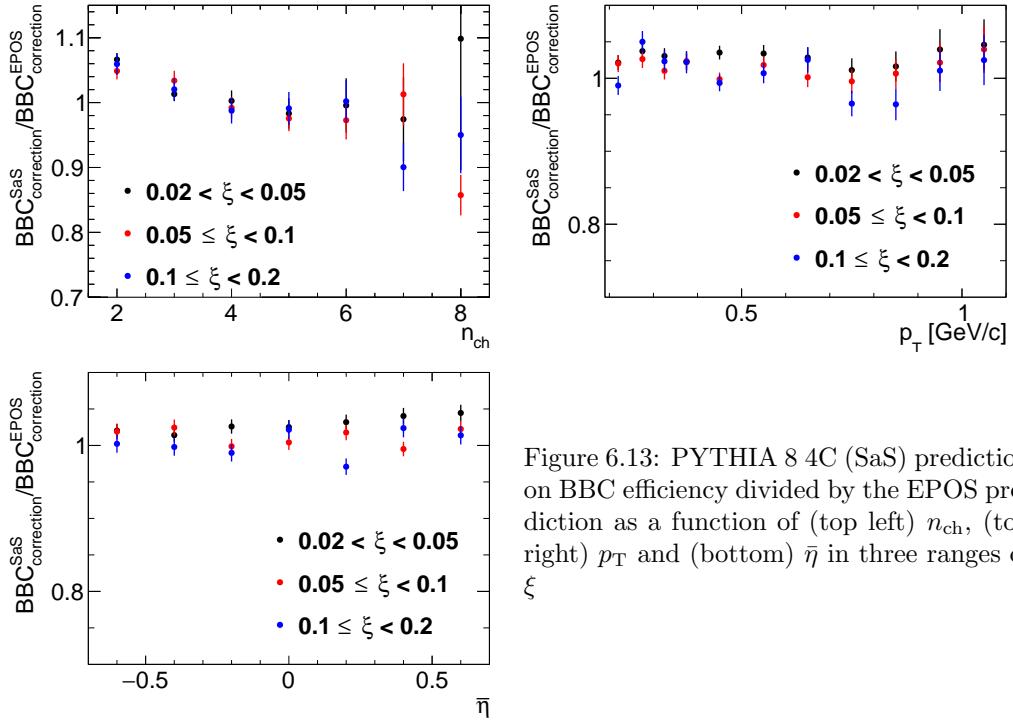


Figure 6.13: PYTHIA 8 4C (SaS) prediction on BBC efficiency divided by the EPOS prediction as a function of (top left)  $n_{\text{ch}}$ , (top right)  $p_{\text{T}}$  and (bottom)  $\bar{\eta}$  in three ranges of  $\xi$

# 528 7. Migrations into and out of 529 the Fiducial Region

530 In this section the corrections due to the migrations of tracks and forward proton into and out of  
531 the fiducial region are described.

## 532 7.1 Migrations of Tracks into and out of the Fiducial Region

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

$$f_{\text{okr}}(p_T, \eta) = \frac{1 - f_{\text{okr}}^-(p_T, \eta)}{1 - f_{\text{okr}}^+(p_T, \eta)} \quad (7.1)$$

536 where:

537  $f_{\text{okr}}^-(p_T, \eta)$  is the fraction of reconstructed tracks for which the corresponding primary particle is  
538 outside of the kinematic range of the measurement,

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

541 The resulting residual migrations, shown in Fig. 7.1, were estimated using PYTHIA 8 SD  
542 embedding MC. The main effect was observed at  $|\eta| \sim 0.7$ , where about 2–6% reconstructed tracks  
543 were associated to primary particle outside the fiducial region. However, above contributions to  
544 the correction factor,  $f_{\text{okr}}(p_T, \eta)$ , cancel each other and a resultant factor is about 2% at  $|\eta| \sim 0.7$ .

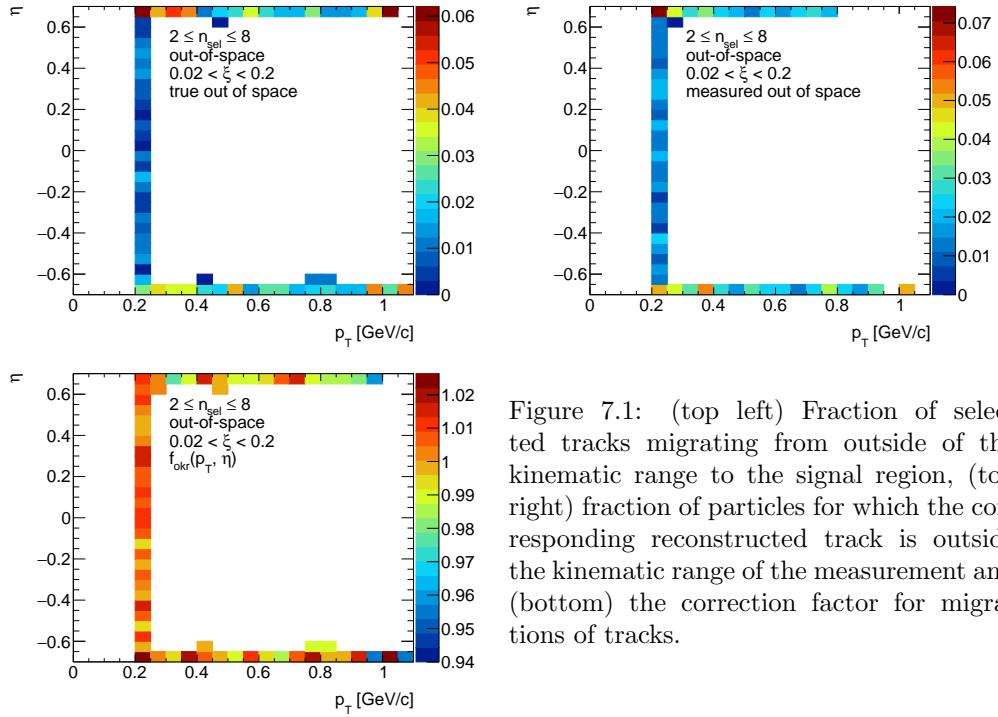


Figure 7.1: (top left) Fraction of selected tracks migrating from outside of the kinematic range to the signal region, (top right) fraction of particles for which the corresponding reconstructed track is outside the kinematic range of the measurement and (bottom) the correction factor for migrations of tracks.

545    **7.2 Migrations in  $\xi$**

546    The analysis was performed in three ranges of  $\xi$ . Thus, there are migrations into and out of these  $\xi$   
 547    regions. They mainly originate from the resolution of  $\xi$  reconstructed from RP tracks. Figure 7.2  
 548    shows the resolution of  $\xi$  as a function of the true-level  $\xi$  (denoted as  $\xi_{\text{true}}$ ) with fitted zeroth order  
 549    polynomial. The resolution of  $\xi$  is fairly constant and equals to about 0.3%.

550    The corrections due to migrations into and out of  $\xi$  regions was defined as:

$$f_\xi = \frac{1 - f_\xi^-}{1 - f_\xi^+} \quad (7.2)$$

551    where:

552     $f_\xi^-$  is the fraction of events for which the corresponding true-level  $\xi_{\text{true}}$ , is outside of the  $\xi$  region,  
 553     $f_\xi^+$  is the fraction of events for which the corresponding reconstructed  $\xi_{\text{reco}}$  is outside of the  $\xi$   
 554    region.

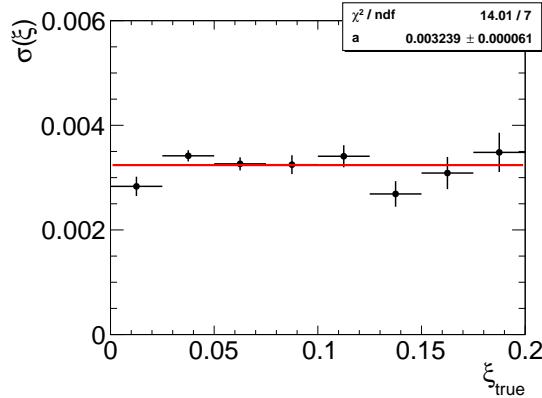


Figure 7.2: The resolution of  $\xi$  as a function of  $\xi_{\text{true}}$ . The zeroth order polynomial, shown as red line, was fitted.

555    The  $f_\xi$  was calculated for each measured variable separately. Figures 7.3 to 7.5 show the  
 556    fraction of events  $f_\xi^-$  and  $f_\xi^+$  as a function of  $n_{\text{ch}}$ ,  $p_T$  and  $\bar{\eta}$ . The lower panel in each figure shows  
 557    the corresponding correction factor  $f_\xi$ . The largest differences between migrations into and out  
 558    of the  $\xi$  regions were observed at  $0.02 < \xi < 0.05$ , where they are of the order of 2 – 4%. In the  
 559    other  $\xi$  regions, the difference between  $f_\xi^-$  and  $f_\xi^+$  is smaller than 1%.

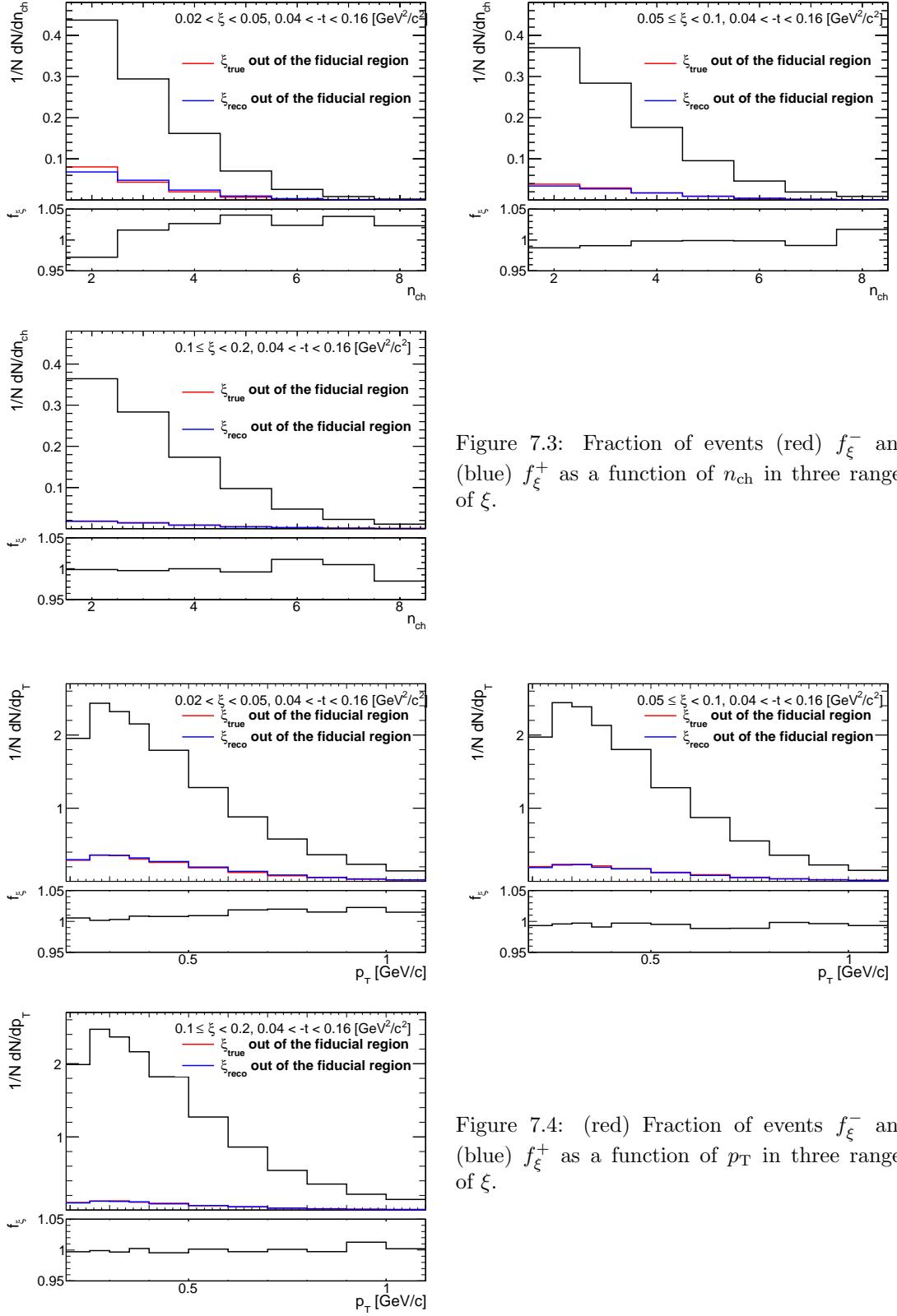


Figure 7.3: Fraction of events (red)  $f_\xi^-$  and (blue)  $f_\xi^+$  as a function of  $n_{\text{ch}}$  in three ranges of  $\xi$ .

Figure 7.4: (red) Fraction of events  $f_\xi^-$  and (blue)  $f_\xi^+$  as a function of  $p_T$  in three ranges of  $\xi$ .

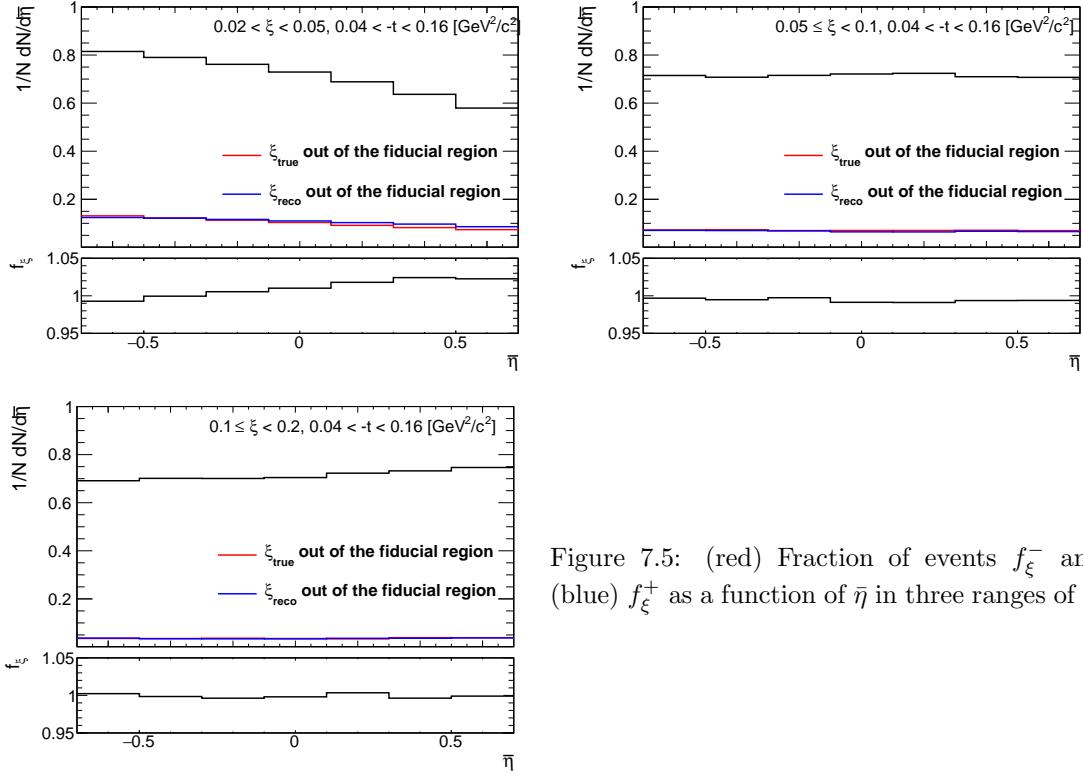


Figure 7.5: (red) Fraction of events  $f_{\xi}^-$  and (blue)  $f_{\xi}^+$  as a function of  $\bar{\eta}$  in three ranges of  $\xi$ .

# 8. Corrections and Unfolding Procedure

After subtraction of accidental, DD, CD and ND backgrounds (as described in Sec. 4.1 and 5), the data was corrected for detector inefficiencies to obtain the distributions of charged particles and particle to antiparticle (pion, kaon, proton and their antiparticle) multiplicity ratios. These corrections include:

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

$$w_{\text{ev}}^{\text{vrt}}(n_{\text{vrt}}^{\text{global}}, |\Delta z_0|) = \frac{1}{\epsilon_{\text{vrt}}(n_{\text{vrt}}^{\text{global}}, |\Delta z_0|)} \cdot \frac{1}{1 - f_{\text{veto}}(n_{\text{vrt}}^{\text{global}}, |\Delta z_0|)} \quad (8.1)$$

where the  $|\Delta z_0|$  dependence is only applicable for events with  $n_{\text{vrt}}^{\text{global}} = 2$  as described in Sec. 6.1.

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

$$w_{\text{trk}}(p_T, \eta, V_z) = \frac{1 - f_{\text{bkg}}(p_T, \eta) - f_{\text{fake}}(p_T, \eta)}{\epsilon_{\text{TPC}}(p_T, \eta, V_z) \epsilon_{\text{TOF}}(p_T, \eta, V_z)} f_{\text{okr}}(p_T, \eta) \quad (8.2)$$

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 matching efficiency [1],  $f_{\text{okr}}(p_T, \eta)$  is a factor accounting for migrations of tracks into and out of the fiducial region,  $f_{\text{bkg}}(p_T, \eta)$  is a fraction of background tracks, and  $f_{\text{fake}}(p_T, \eta)$  is a fraction of fake tracks. These corrections were not applied for  $n_{\text{ch}}$  measurements since they were taken into account in the unfolding procedure.

- event-by-event (for  $n_{\text{ch}}$  distribution ) or track-by-track (for  $p_T, \bar{\eta}$  distributions) weights,  $f_\xi$ , due to migrations of events between three  $\xi$  regions.

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

$$w_{\text{BBC}} = \frac{1}{\epsilon_{\text{BBC}}} \quad (8.3)$$

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

## 8.1 Correction to $dN/dn_{\text{sel}}$

In order to express the multiplicity distribution in terms of the number of charged particles,  $n_{\text{ch}}$ , instead of the number of selected tracks,  $n_{\text{sel}}$ . The following procedure based on the Bayesian unfolding [13, 14] was used. First, the  $n_{\text{sel}}$  distribution was corrected for vertex reconstruction effects 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  $n_{\text{ch}}$  are produced,  $N_{\text{ev}}(n_{\text{ch}})$ , can be associated with the number of events in which  $n_{\text{sel}}$  are reconstructed,  $N_{\text{ev}}(n_{\text{sel}})$ . Since there are several possible  $n_{\text{sel}}$  observed in  $n_{\text{ch}}$  event,  $N_{\text{ev}}(n_{\text{ch}})$  is given by:

$$\begin{aligned} N_{\text{ev}}(n_{\text{ch}}) &= \sum_{n_{\text{sel}}=0}^8 P(n_{\text{ch}}|n_{\text{sel}}) \cdot N_{\text{ev}}(n_{\text{sel}}) \\ &= \frac{1}{\epsilon_m(n_{\text{ch}}) \epsilon_r(n_{\text{ch}})} \sum_{n_{\text{sel}}=2}^8 P(n_{\text{ch}}|n_{\text{sel}}) \cdot N_{\text{ev}}(n_{\text{sel}}) \end{aligned} \quad (8.4)$$

where:

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

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

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

598     Figure 8.1 shows  $\epsilon_m(n_{\text{ch}})$  in three ranges of  $\xi$ . It was derived from MC and varies from about 25%  
599     for  $n_{\text{ch}} = 2$  to 95% for  $n_{\text{ch}} = 8$ . Since there are additional data-driven corrections to TPC and  
600     TOF efficiencies, MC simulations were modified by randomly removing or adding tracks. This was  
601     done in accordance with differences in the efficiencies between data and MC.

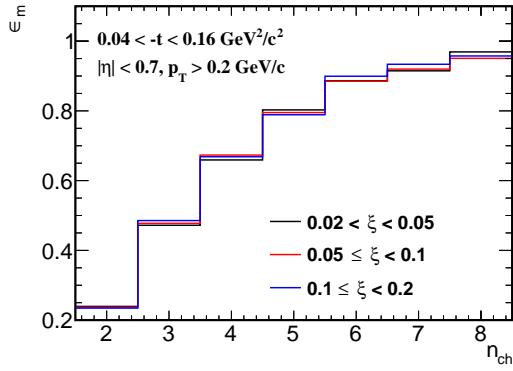


Figure 8.1:  $\epsilon_m(n_{\text{ch}})$  calculated separately in three ranges of  $\xi$  using PYTHIA 8 embedding MC.

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

$$P(n_{\text{ch}}) \cdot P(n_{\text{sel}}|n_{\text{ch}}) = P(n_{\text{ch}}|n_{\text{sel}}) \cdot P(n_{\text{sel}}) \quad (8.5)$$

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

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

607         • In the first iteration, it is assumed that:

$$P(n_{\text{ch}}|n_{\text{sel}}) = P = P^{\text{MC}}(n_{\text{sel}}|n_{\text{ch}}) \frac{P^{\text{MC}}(n_{\text{ch}})}{P^{\text{MC}}(n_{\text{sel}})} \quad (8.6)$$

$$N_{\text{ev}}(n_{\text{ch}}) = \frac{1}{\epsilon_m(n_{\text{ch}})} \sum_{n_{\text{sel}}=2}^8 N_{\text{ev}}(n_{\text{sel}}) \cdot P \quad (8.7)$$

610     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  
611     the same for each iteration.

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

$$P^{i+1} = P^{\text{MC}}(n_{\text{sel}}|n_{\text{ch}}) \frac{N_{\text{ev}}^i(n_{\text{ch}})}{N_{\text{ev}}(n_{\text{sel}})} \quad (8.8)$$

$$N_{\text{ev}}^{i+1}(n_{\text{ch}}) = \frac{1}{\epsilon_m(n_{\text{ch}})} \sum_{n_{\text{sel}}=2}^8 N_{\text{ev}}(n_{\text{sel}}) \cdot P^{i+1} \quad (8.9)$$

613     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.

615 The unfolding matrices  $P(n_{\text{ch}}|n_{\text{sel}})$  for each  $\xi$  region, shown in Fig. 8.2, were obtained from  
 616 PYTHIA 8 embedding MC and used in all iterations of the above procedure. Similarly to  $\epsilon_m(n_{\text{ch}})$ ,  
 617 the matrices were modified in order to take into account differences between data and PYTHIA 8.  
 618 In order to obtain statistically precise unfolding matrices, all simulated events were used, i.e. also  
 619 those with additional fake vertices (with  $n_{\text{sel}}$  defined as a number of primary tracks associated  
 620 with the best vertex). The systematic uncertainty related to limited statistics in PYTHIA 8 was  
 621 estimated by performing 50 pseudo-experiments, in which the unfolding matrices were smeared  
 622 according to their statistical uncertainties. It affects mainly large charged-particle multiplicities,  
 623 where it is about 8 – 10%, as shown in Fig. 8.3.

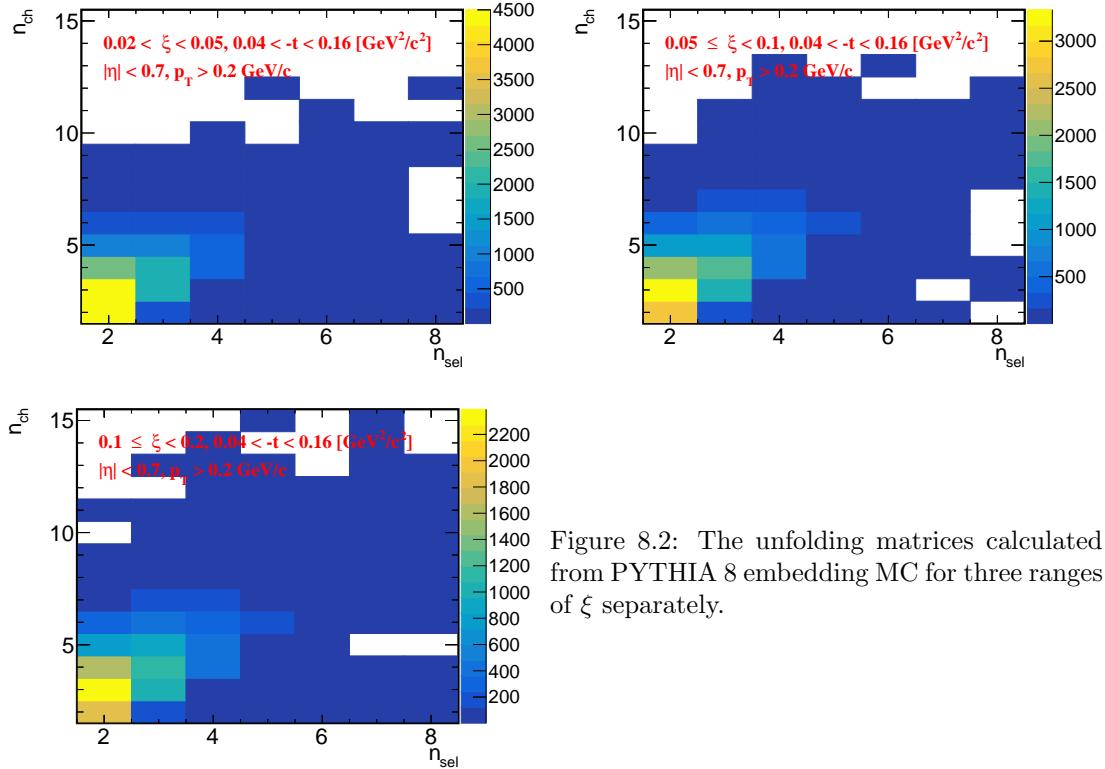


Figure 8.2: The unfolding matrices calculated from PYTHIA 8 embedding MC for three ranges of  $\xi$  separately.

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

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

## 631 8.2 Correction to Transverse Momentum and Pseudorapidity Distributions

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

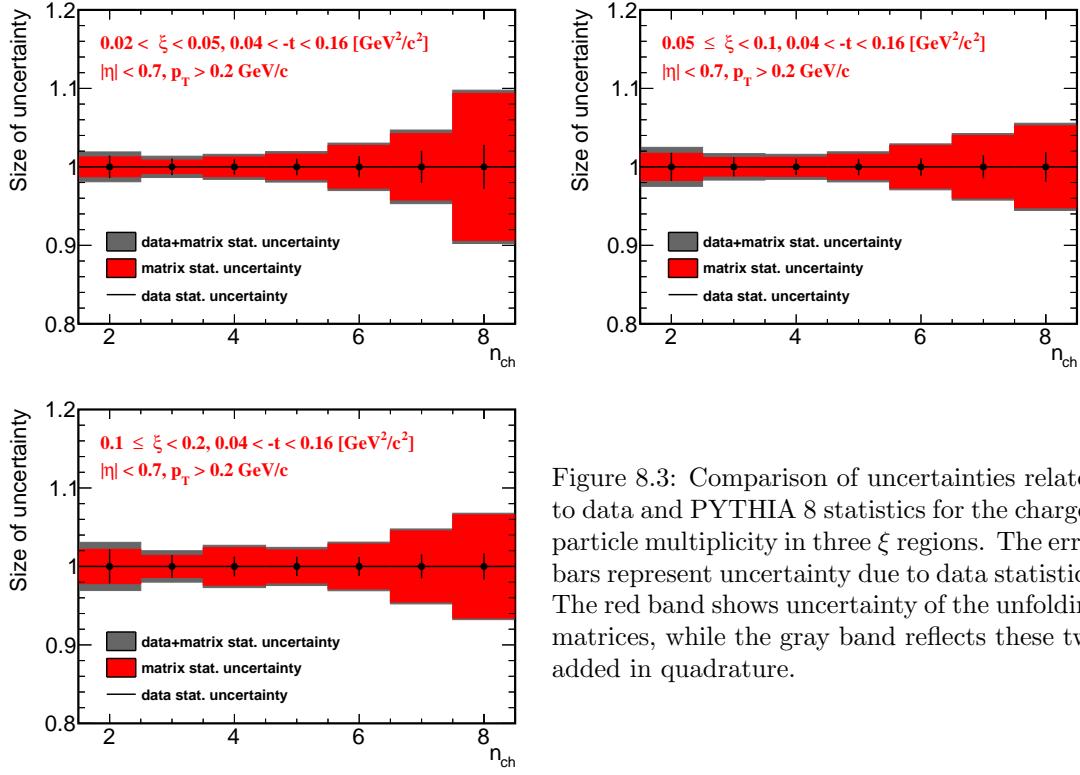


Figure 8.3: Comparison of uncertainties related to data and PYTHIA 8 statistics for the charged particle multiplicity in three  $\xi$  regions. The error bars represent uncertainty due to data statistics. The red band shows uncertainty of the unfolding matrices, while the gray band reflects these two added in quadrature.

639 In order to obtain charged-particle densities, the  $p_T$  and  $\bar{\eta}$  distributions were normalized to  
 640 unity and scaled by the average charged particle multiplicity in an event  $\langle n_{\text{ch}} \rangle$ . The latter was  
 641 calculated from the corrected charged particle multiplicity distribution  $dN/dn_{\text{ch}}$  (Sec. 8.1).

642 The mean particle densities in an event,  $\langle p_T \rangle$  and  $\langle \bar{\eta} \rangle$ , were obtained from the measured  
 643 distributions.

### 644 8.3 Closure Tests

645 In order to validate the correction procedures, closure tests were performed , i.e. full correction  
 646 procedure was applied to the MC detector-level distributions and the results were directly com-  
 647 pared to the true-level distributions. Figure 8.4 shows closure tests of multiplicity, transverse  
 648 momentum and pseudorapidity distributions for three ranges of  $\xi$ , separately. PYTHIA 8 SD  
 649 embedding MC was used as an input. In order to compare corrected and true-level distributions,  
 650 the statistical uncertainties of the true-level distributions were assumed to be 0. The difference  
 651 between true-level and corrected distributions was taken as a systematic uncertainties.

### 652 8.4 EAST-WEST asymmetry

653 Another kind of consistency checks can be performed by comparing the results obtained by tagging  
 654 forward protons in different detectors. Therefore, each distribution was measured separately for  
 655 events in which forward proton is on one and the other side of the IP (east-west). Figure 8.5 shows  
 656 the tests of multiplicity, transverse momentum and pseudorapidity distributions for three ranges  
 657 of  $\xi$ , separately. The largest difference is observed for charged-particle multiplicity distributions,  
 658 where it varies up to 20% for  $n_{\text{ch}} \geq 7$  and  $0.02 < \xi < 0.05$ . For the rest multiplicities and  $\xi$   
 659 ranges, the differences are smaller (< 10%). In case of  $p_T$  and  $\bar{\eta}$  distributions, a level of these

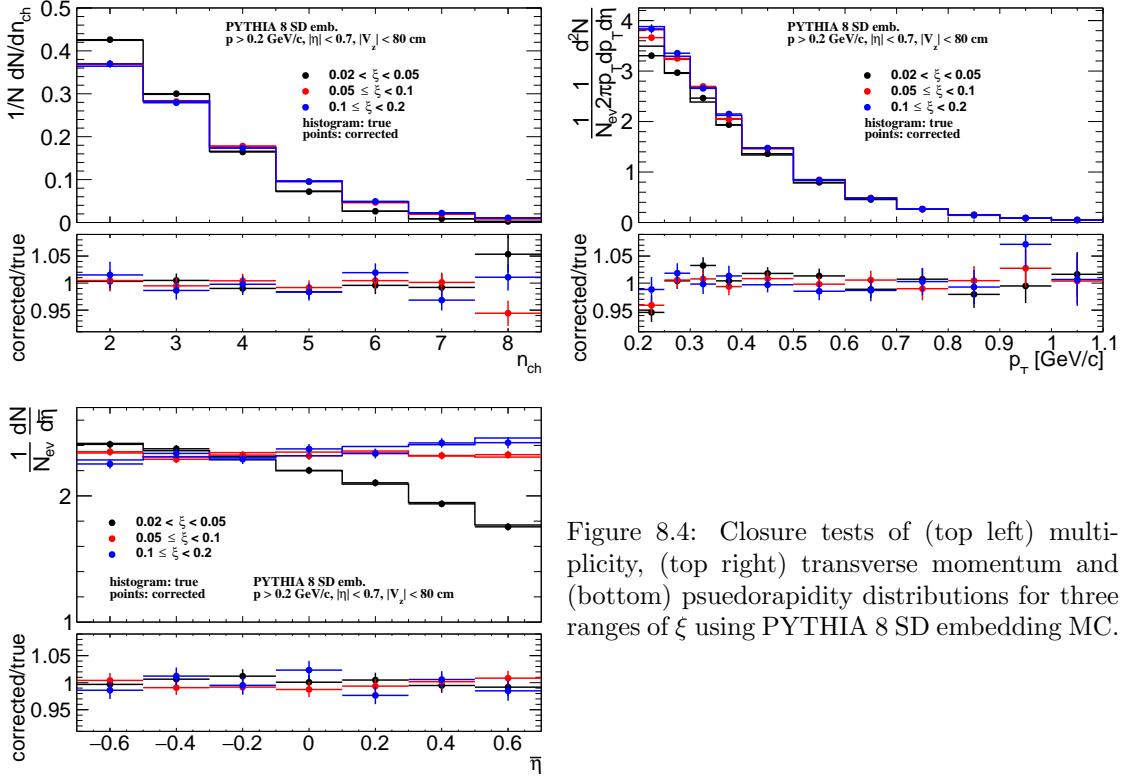


Figure 8.4: Closure tests of (top left) multiplicity, (top right) transverse momentum and (bottom) pseudorapidity distributions for three ranges of  $\xi$  using PYTHIA 8 SD embedding MC.

disagreements is below 5%. As a result, half of the differences between east and west distributions were used to be systematic uncertainty.

## 8.5 Particle Identification

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

The ionization energy loss of charged particles in material is given by the Bethe-Bloch formula and for the STAR TPC by the more precise Bichsel formula [15]. The particle type can be determined by comparison of particle's  $dE/dx$  with the Bethe-Bloch (Bichsel) expectations. Figure 8.6 shows the  $dE/dx$  versus rigidity  $q \times p$  for particles in  $|\eta| < 0.7$ . Particles are well separated at low  $|q \times p|$ , whereas at higher  $|q \times p|$  the  $dE/dx$  of different particle species starts to overlap:  $e^\pm$  and  $K^\pm$  merge at  $\sim 0.4$  GeV/c,  $K^\pm$  and  $\pi^\pm$  merge at  $\sim 0.65$  GeV/c, and  $p(\bar{p})$  and  $\pi^\pm$  merge at  $\sim 1$  GeV/c. Since the  $dE/dx$  distribution for a given particle type is not Gaussian, the following variable for each particle type was defined:

$$n\sigma_{dE/dx}^i = \ln \left( \frac{dE/dx}{(dE/dx)_i^{\text{BB}}} \right) / \sigma \quad (8.10)$$

where  $(dE/dx)_i^{\text{BB}}$  is the Bethe-Bloch (Bichsel) expectation of  $dE/dx$  for the given particle type  $i$  ( $i = \pi, K, p$ ),  $\sigma$  - the  $dE/dx$  resolution. The expected value of  $n\sigma_{dE/dx}^i$  for the particle under consideration is 0 and the width equals to 1. The sample  $n\sigma_{dE/dx}^i$  distribution for  $\pi^\pm$ ,  $K^\pm$  and  $p(\bar{p})$  in one  $\xi$  range,  $0.02 < \xi < 0.05$ , is shown in Fig. 8.7.

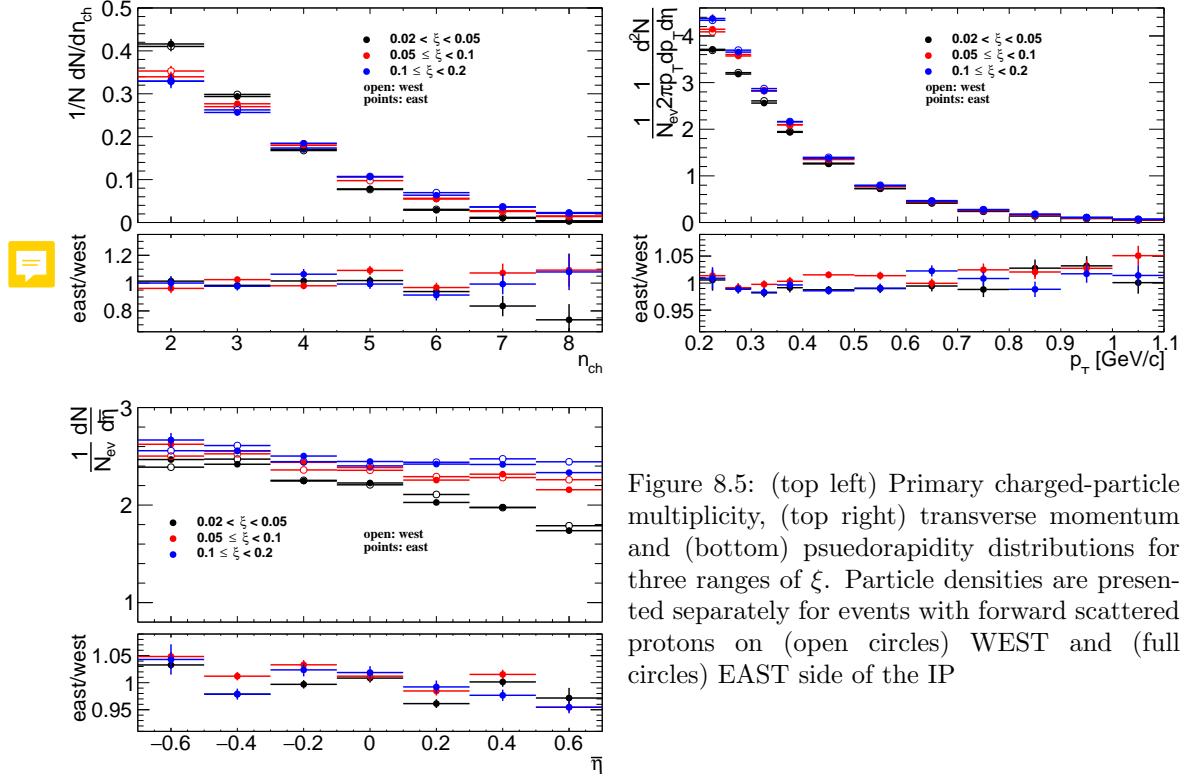


Figure 8.5: (top left) Primary charged-particle multiplicity, (top right) transverse momentum and (bottom) psuedorapidity distributions for three ranges of  $\xi$ . Particle densities are presented separately for events with forward scattered protons on (open circles) WEST and (full circles) EAST side of the IP

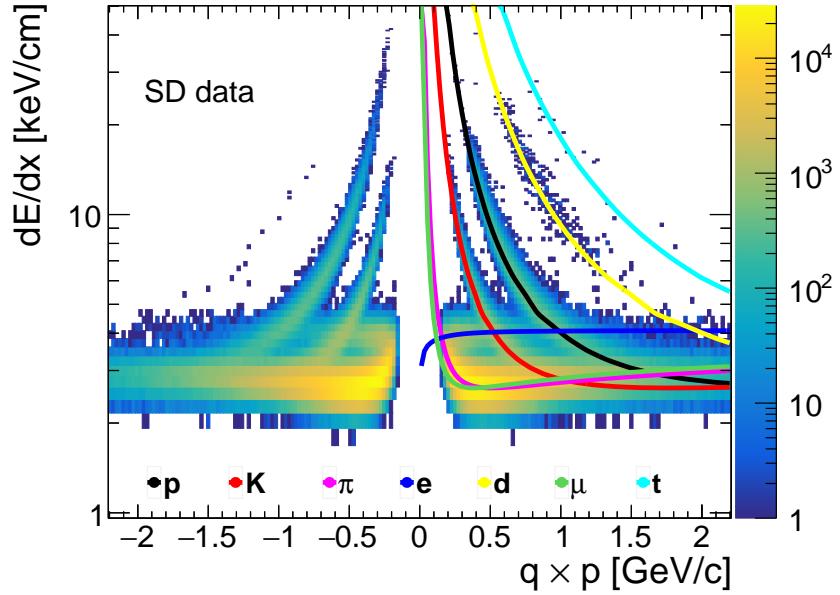


Figure 8.6: Specific ionization energy loss  $dE/dx$  as a function of rigidity  $q \times p$  for particles in  $|\eta| < 0.7$ . The Bichsel predictions for each particle species are also shown.

Figure 8.8 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 the  $\xi$  range,  $0.02 < \xi < 0.05$ , each corrected for the energy loss (mass of  $i$ -particle was assumed) [1] and vertexing (other  $p_T$  bins are shown in Appendix C). To extract the particle yield for a given

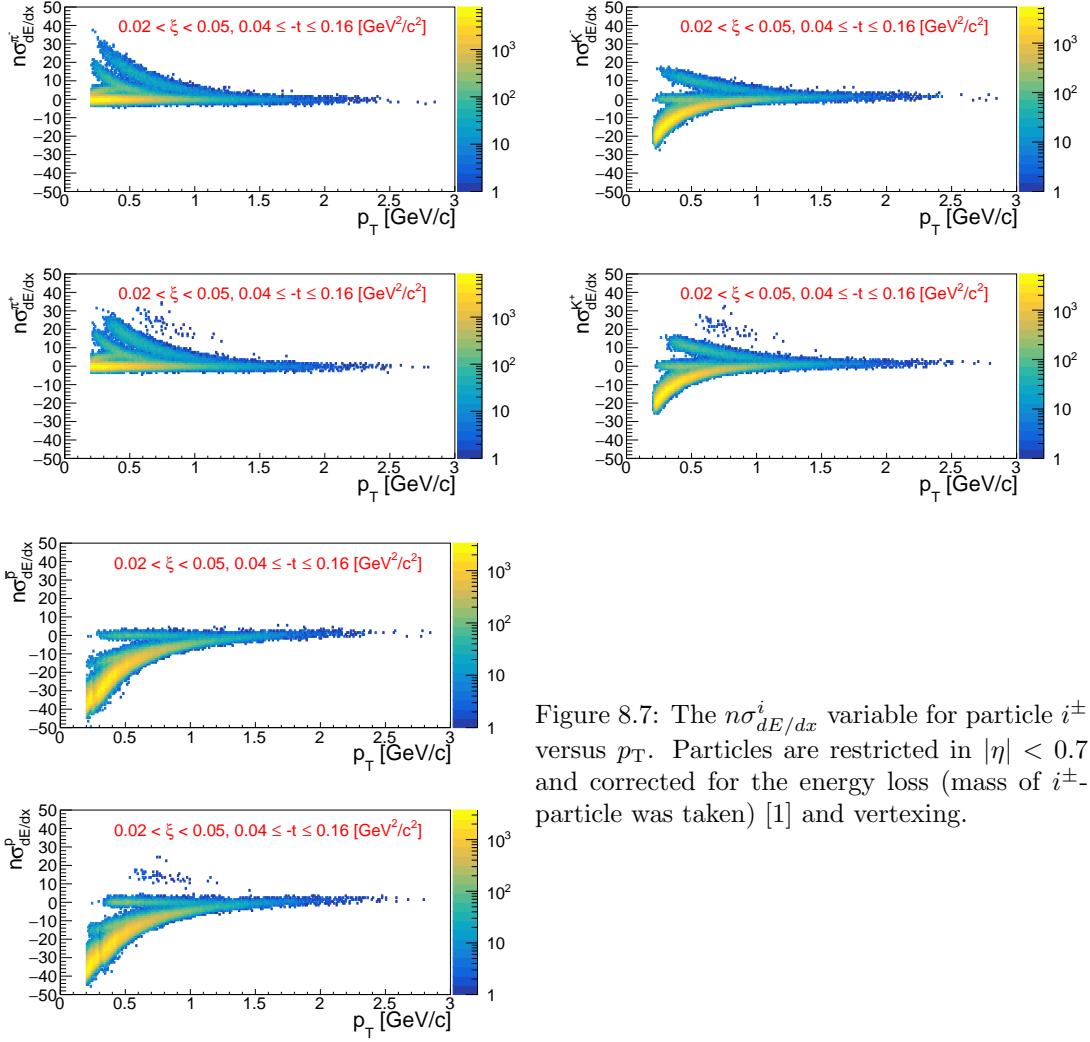


Figure 8.7: The  $n\sigma_{dE/dx}^i$  variable for particle  $i^\pm$  versus  $p_T$ . Particles are restricted in  $|\eta| < 0.7$  and corrected for the energy loss (mass of  $i^\pm$ -particle was taken) [1] and vertexing.

684 particle type, a multi-Gaussian fit is applied to the  $n\sigma_{dE/dx}^i$  distribution in each  $p_T$  bin and  $\xi$  range.  
 685 The parameters of the multi-Gaussian fit are the centroids  $\mu_{i-/i+}$ , widths  $\sigma_{i-/i+}$ , sum and ratios of  
 686 yields  $C_{i-/i+}$ ,  $r_{i-/i+}$  for negative  $i^-$  and positive  $i^+$  particles ( $\pi^\pm$ ,  $e^\pm$ ,  $K^\pm$ ,  $p$  and  $\bar{p}$ ). The positive  
 687 and negative particle  $n\sigma_{dE/dx}^i$ -distributions are fitted simultaneously, where the centroids and  
 688 widths are kept the same for particle and antiparticle. In some  $p_T$  regions, the fit does not  
 689 converge, because different particle species are not well separated there. Therefore, multiple steps  
 690 of fitting are performed to reduce the number of free parameters in the final fit and ensure its  
 691 stability. Almost all centroids and widths are constrained by a function with free parameters  $p_k$ ,  
 692 where  $k \in \mathbb{N}$ . The function is chosen to describe the data as best as possible. Since  $dE/dx$  is  
 693 a function of the particle momentum and its shape should be independent of the process under  
 694 study, the values of  $p_k$  are obtained only for events with  $0.02 < \xi < 0.05$  and kept the same for other  
 695  $\xi$  ranges. The electron contributions are fixed, but separately for each  $\xi$  range. The procedure  
 696 slightly differs for different particle types. In each step, the multi-Gaussian fit is performed first,  
 697 then the widths and centroids are fitted in  $p_T$  ranges in which the fit applied to  $n\sigma_{dE/dx}^i$  converges.  
 698 Later, the widths and centroids are extrapolated to other  $p_T$  ranges, in which particle species are  
 699 not well separated:

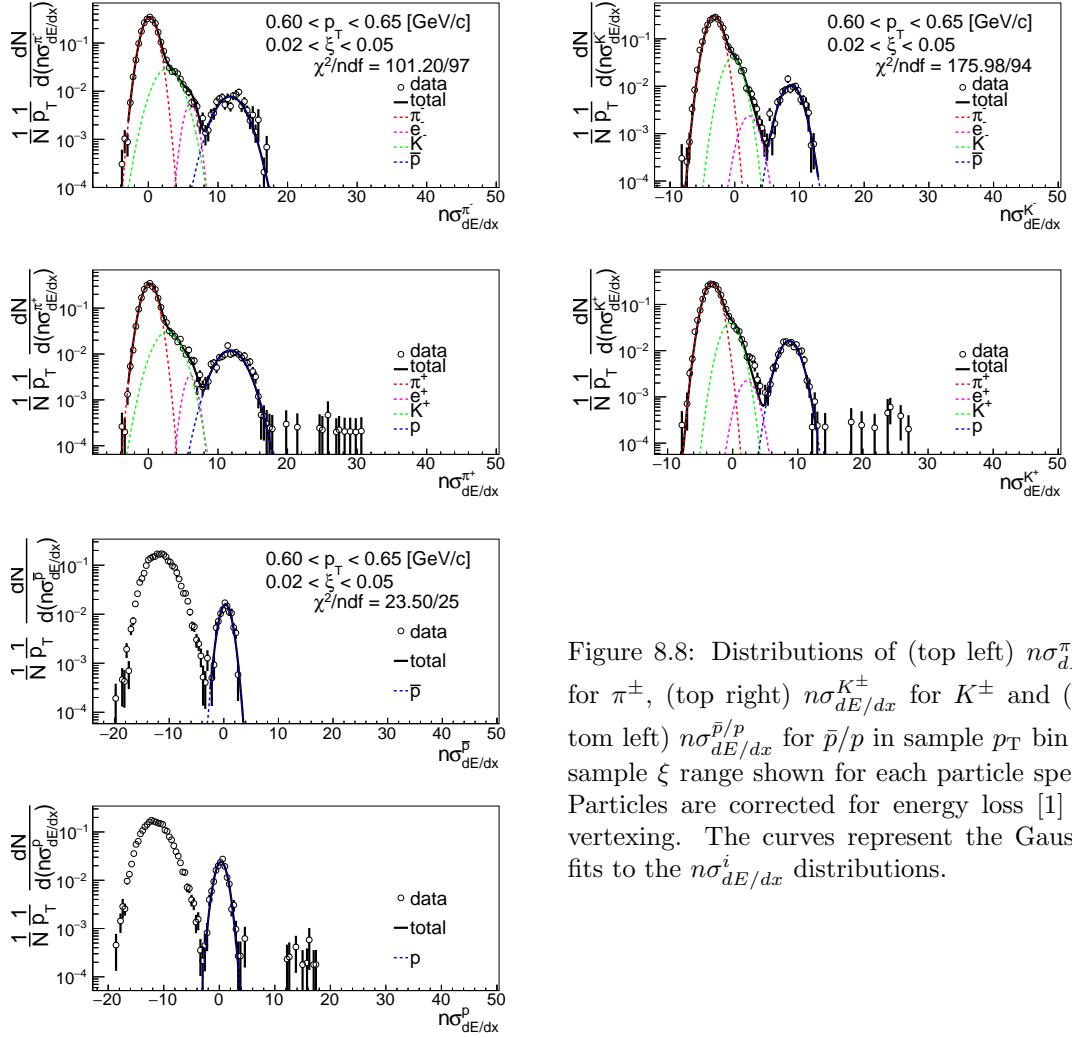


Figure 8.8: Distributions of (top left)  $n\sigma_{dE/dx}^{\pi^\pm}$  for  $\pi^\pm$ , (top right)  $n\sigma_{dE/dx}^{K^\pm}$  for  $K^\pm$  and (bottom left)  $n\sigma_{dE/dx}^{\bar{p}/p}$  for  $\bar{p}/p$  in sample  $p_T$  bin and sample  $\xi$  range shown for each particle species. Particles are corrected for energy loss [1] and vertexing. The curves represent the Gaussian fits to the  $n\sigma_{dE/dx}^i$  distributions.

700      1.  $\pi^\pm$ :

- 701      • Step 1 (Fig. 8.9):
  - 702        – Analyze data with  $0.2 < p_T < 0.65 \text{ GeV/c}$
  - 703        – Fit  $\mu_{\pi^-/\pi^+}$  and  $\sigma_{\pi^-/\pi^+}$  as a function of  $p_T$  with a polynomial  $p_0 p_T^3 + p_1 p_T^2 + p_2 p_T + p_3$
  - 704        – Fit  $r_{e^-/e^+}$  as a function of  $p_T$  with a polynomial  $p_0 p_T^2 + p_1 p_T + p_2$
  - 705        – Fit  $C_{e^-/e^+}$ ,  $\mu_{K^-/K^+}$  as a functions of  $p_T$  with  $p_0 \exp(p_1 p_T) + p_2$
  - 706        – Fit  $\mu_{e^-/e^+}$  as a function of  $p_T$  with  $p_0 \exp[-(p_1 p_T)^{p_2}]$
  - 707        – Fit  $\sigma_{K^-/K^+}$  as a function of  $p_T$ , for  $0.3 < p_T < 0.5 \text{ GeV/c}$ , with constant  $p_0$
  - 708        – Fit  $\mu_{\bar{p}/p}$  and  $\sigma_{\bar{p}/p}$  as a function of  $p_T$  with  $p_0 \exp(p_1 p_T)$
- 709      • Step 2:
  - 710        –  $\sigma_{e^-/e^+}$  fixed to 1.2 and 0.8 for  $0.2 < p_T < 0.4$  and  $0.4 < p_T < 0.7$ , respectively
  - 711        – Fit  $\sigma_{K^-/K^+}$  as a function of  $p_T$ , for  $0.3 < p_T < 0.7 \text{ GeV/c}$ , with constant  $p_0$  and fix it to the value of  $p_0$
  - 712        – The rest parameters from Step 1 are fixed to the values calculated from functions obtained in Step 1:  $\mu_{\pi^-/\pi^+}$ ,  $\sigma_{\pi^-/\pi^+}$ ,  $r_{e^-/e^+}$ ,  $C_{e^-/e^+}$ ,  $\mu_{e^-/e^+}$ ,  $\mu_{K^-/K^+}$ ,  $\mu_{\bar{p}/p}$ ,  $\sigma_{\bar{p}/p}$

716        2.  $K^\pm$ :

- 717     • Step 1 (Fig. 8.10):
  - 718        – Analyze data with  $0.2 < p_T < 0.6 \text{ GeV}/c$
  - 719        – Fit  $\mu_{\pi^-/\pi^+}$  as a function of  $p_T$  with  $-\exp(p_0 + p_1 p_T)$
  - 720        – Fit  $\sigma_{\pi^-/\pi^+}$ ,  $C_{e^-/e^+}$ ,  $\sigma_{e^-/e^+}$ ,  $\sigma_{K^-/K^+}$  as a function of  $p_T$  with  $\exp(p_0 + p_1 p_T)$
  - 721        – Fit  $r_{e^-/e^+}$  as a function of  $p_T$  with constant  $p_0$
  - 722        – Fit  $\mu_{e^-/e^+}$  as a function of  $p_T$  with a polynomial  $p_0 p_T^3 + p_1 p_T^2 + p_2 p_T + p_3$
  - 723        – Fit  $\mu_{K^-/K^+}$  as a function of  $p_T$  with a polynomial  $p_0 + p_1 p_T^2$
- 724     • Step 2:
  - 725        – All parameters from Step 1 except  $\sigma_{e^-/e^+}$  are fixed to the values calculated from functions obtained in Step 1
  - 726        – Fit  $\sigma_{e^-/e^+}$  as a function of  $p_T$ , for  $0.45 < p_T < 0.65 \text{ GeV}/c$ , with constant  $p_0$
- 728     • Step 3:
  - 729        –  $\sigma_{e^-/e^+}$  fixed to the values calculated from functions obtained in Steps 1 and 2 for  $0.3 < p_T < 0.45$  and  $0.45 < p_T < 0.65$ , respectively.
  - 730        – The rest parameters from Step 1 are fixed to the values calculated from functions obtained in Step 1:  $\mu_{\pi^-/\pi^+}$ ,  $\sigma_{\pi^-/\pi^+}$ ,  $r_{e^-/e^+}$ ,  $C_{e^-/e^+}$ ,  $\mu_{e^-/e^+}$ ,  $\mu_{K^-/K^+}$ ,  $\sigma_{K^-/K^+}$

733        3.  $\bar{p}, p$ :

- 734     • Step 1 (Fig. 8.11):
  - 735        – Analyze data with  $0.4 < p_T < 0.9 \text{ GeV}/c$
  - 736        – Fit  $\mu_{\pi^-/\pi^+}$ ,  $\mu_{K^-/K^+}$  as a function of  $p_T$  with a polynomial  $p_0 p_T + p_1$
  - 737        – Fit  $\sigma_{\pi^-/\pi^+}$  as a function of  $p_T$  with a polynomial  $p_0 p_T^2 + p_1 p_T + p_2$
  - 738        – Fit  $\sigma_{K^-/K^+}$  as a function of  $p_T$  with  $\exp(p_0 + p_1 p_T)$
- 739     • Step 2:
  - 740        –  $\mu_{K^-/K^+}$  fixed to the values calculated from a function obtained in Step 1
  - 741        – All the rest parameters from Step 1 are limited to the values calculated from functions obtained in Step 1
  - 742        – Fit  $\mu_{\pi^-/\pi^+}$ ,  $\sigma_{\pi^-/\pi^+}$ ,  $\sigma_{K^-/K^+}$  as a function of  $p_T$  with a polynomial  $p_0 p_T^2 + p_1 p_T + p_2$
  - 743        – Fit  $\mu_{\bar{p}/p}$  as a function of  $p_T$ , for  $0.7 < p_T < 1.0 \text{ GeV}/c$ , with constant  $p_0$
- 745     • Step 3:
  - 746        –  $\mu_{K^-/K^+}$  fixed to the values calculated from a function obtained in Step 1
  - 747        –  $\mu_{\bar{p}/p}$  fixed to the values calculated from a function obtained in Step 2 for  $0.7 < p_T < 1.0$
  - 748        – The rest parameters from Step 2 are fixed to the values calculated from functions obtained in Step 2:  $\mu_{\pi^-/\pi^+}$ ,  $\sigma_{\pi^-/\pi^+}$ ,  $\sigma_{K^-/K^+}$

751        The particle yield is extracted from the fit to the corresponding  $n\sigma_{dE/dx}^i$  distribution (corrected  
 752        only for the energy loss [1] and vertexing). As shown in Fig. 8.7, the  $dE/dx$  of each particle  
 753        type merge at large  $p_T$ . Hence, the particle identification is limited. Pions can be identified  
 754        in the momentum range of  $0.2 - 0.7 \text{ GeV}/c$ , kaons in  $0.3 - 0.65 \text{ GeV}/c$  and (anti)protons in  
 755         $0.4 - 1.0 \text{ GeV}/c$ .

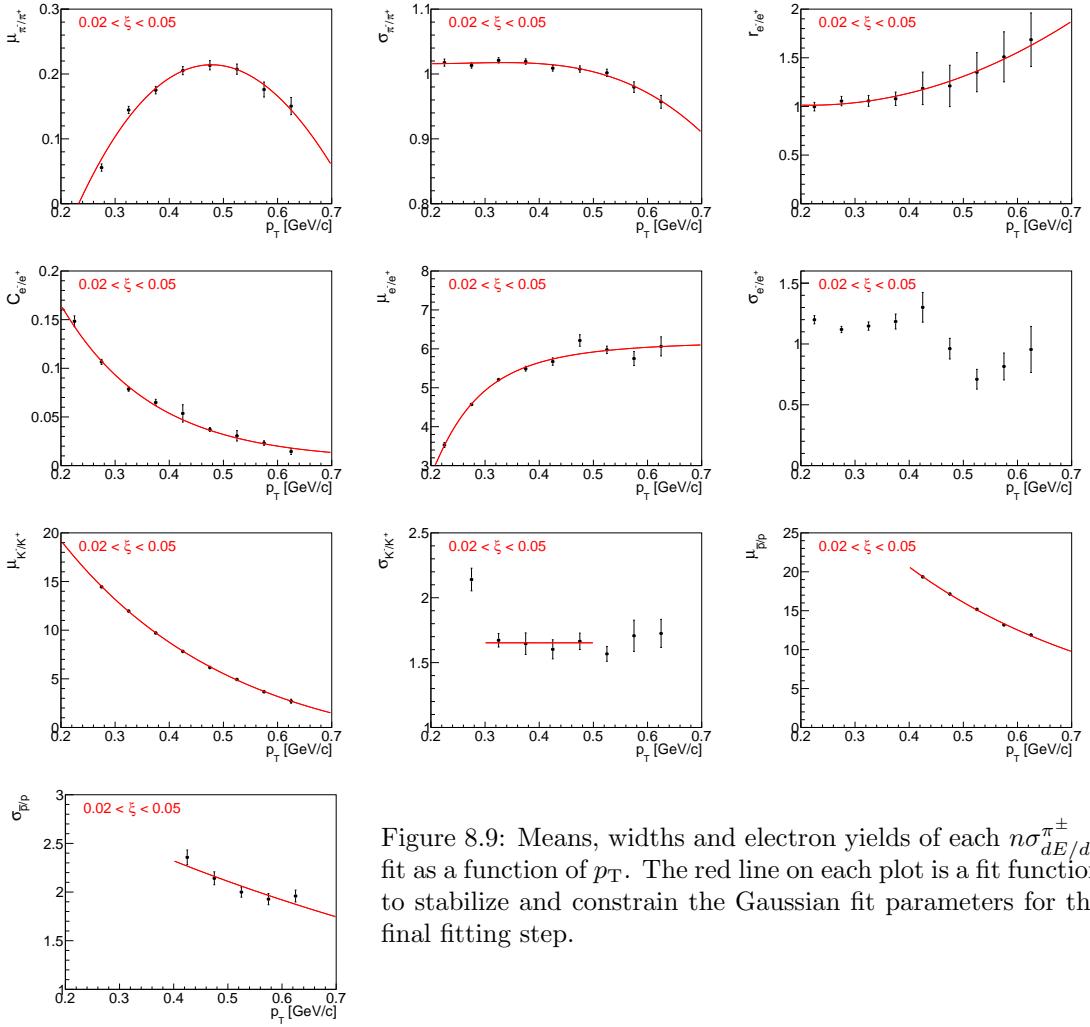


Figure 8.9: Means, widths and electron yields of each  $n\sigma_{dE/dx}^{\pi^\pm}$  fit as a function of  $p_T$ . The red line on each plot is a fit function to stabilize and constrain the Gaussian fit parameters for the final fitting step.

## 756 8.6 Antiparticle-to-Particle Ratios

757 The following steps were taken to correct an identified antiparticle to particle (pion, kaon, proton  
758 and their antiparticle) multiplicity ratios as a function of  $p_T$  in three ranges of  $\xi$ :

- 759 • The raw identified particle yields were obtained through multi-Gaussian fits to the  $n\sigma_{dE/dx}^i$   
760 distributions (Sec. 8.5), where the vertex reconstruction and energy loss corrections [1] were  
761 applied. The latter depends on the particle type.
- 762 • The non-SD background was subtracted. It was assumed that the accidental background  
763 does not depend on the particle type and for this reason it was not subtracted.
- 764 • The particle yields were corrected for track reconstruction efficiencies [1], which depend on  
765 the particle type and charge. These corrections were averaged over  $\eta$  and  $V_z$ .
- 766 • The background from non-primary tracks was subtracted (Sec. 4.2):
  - 767 –  $\pi^\pm$ : weak decays pions, muon contribution and background from detector dead-material  
768 interactions,
  - 769 –  $p$ : background from detector dead-material interactions,

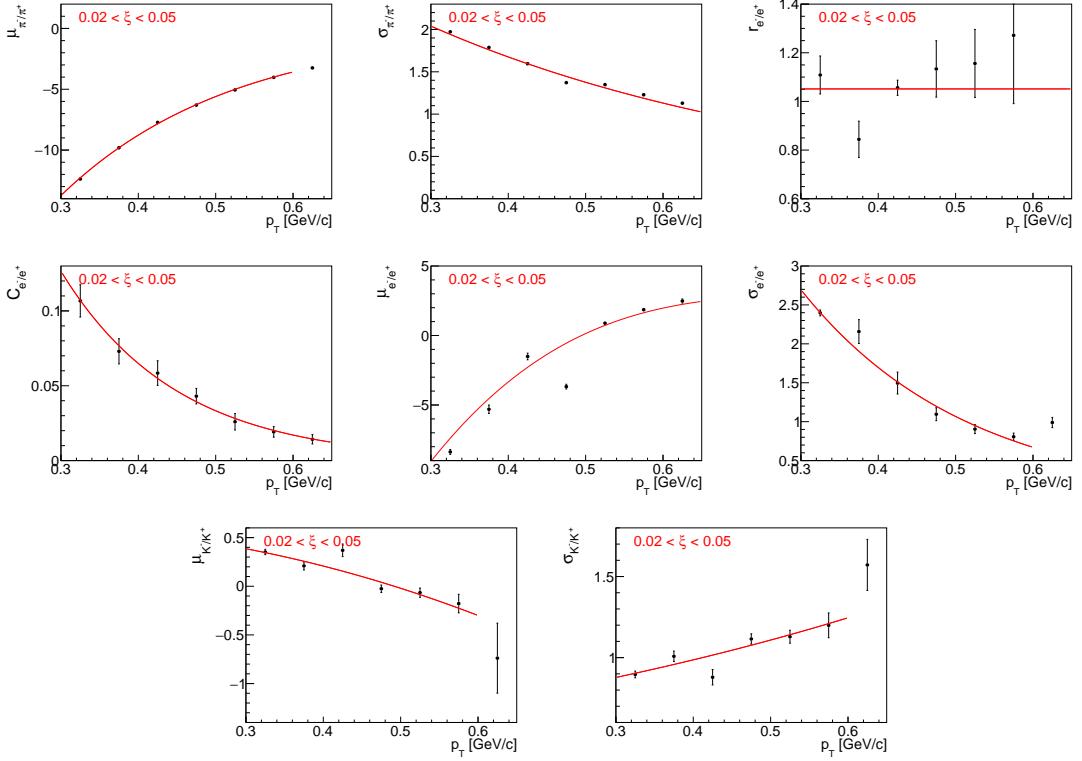


Figure 8.10: Means, widths and electron yields of each  $n\sigma_{dE/dx}^{K^\pm}$  fit as a function of  $p_T$ . The red line on each plot is a fit function to stabilize and constrain the Gaussian fit parameters for the final fitting step.

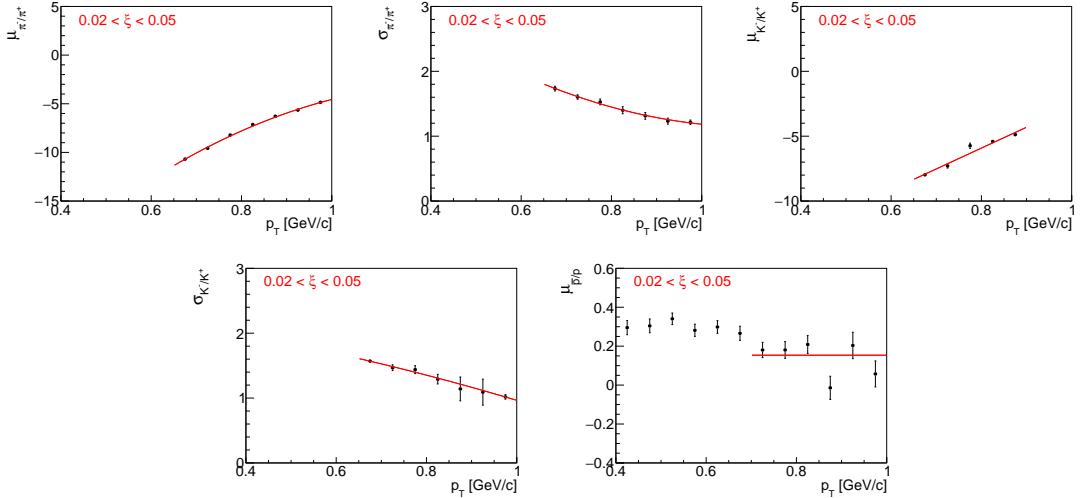


Figure 8.11: Means and widths of each  $n\sigma_{dE/dx}^{\bar{p}/p}$  fit as a function of  $p_T$ . The red line on each plot is a fit function to stabilize and constrain the Gaussian fit parameters for the final fitting step.

- 770 –  $p, \bar{p}$ : reconstructed tracks which have the appropriate number of common hit points  
771 with true-level particle, but the distance between them is too large (this background is  
772 negligibly small for other particle types),

- 773           – fake track contribution was assumed to be the same for each particle type, hence, it  
774            was not subtracted.
- 775       • Since track and  $\xi$  migrations, and BBC-small efficiency, do not depend on the particle type  
776            and charge, these corrections are not applied.
- 777       • Finally, each antiparticle  $p_T$  distribution was divided by the corresponding particle  $p_T$  dis-  
778            tribution to obtain fully corrected identified antiparticle to particle multiplicity ratios.
- 779       • Additionally, the average antiparticle to particle ratios over fiducial region of  $p_T$  in each  $\xi$   
780            region were calculated.

# 9. 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],
- fake track background contribution (Sec. 4.2),
- TOF system simulation accuracy [1],
- accidental background contribution (Sec. 4.1),
- the effect of alternative model of hadronisation on BBC-small efficiency (Sec. 6.2),
- non-SD background contribution (Sec. 5),
- non-closure (Sec 8.3),
- non-closure of  $N_{\text{ev}}$ , applied only to  $p_{\text{T}}$  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 8.4).

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  $n_{\text{sel}}$  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  $n_{\text{sel}}$  is increased or decreased, depending on the sign of systematic uncertainty.

Figures 9.1 to 9.3 show the components contributing to the total systematic uncertainty for charged particle distributions without the identification. The dominant systematic uncertainty for  $p_{\text{T}}$  and  $n_{\text{ch}}$  distributions is related to TOF system simulation accuracy. It affects mainly low- $p_{\text{T}}$  particles, where it is about 6%, and large charged particle multiplicities, where it varies up to 50% 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 4%. The largest (up to 35%) systematic uncertainty for  $\langle \bar{\eta} \rangle$ , is related to the observed difference in the distributions calculated separately with respect to the forward proton direction. The rest of the components have smaller contributions to the total systematic uncertainty. The systematic uncertainty on non-closure is at the level of 2% which proves the accuracy of the correction procedure.

Figures 9.4 to 9.7 show breakdown of all different systematics for the antiparticle-to-particle multiplicity ratio distributions. An additional systematic contribution for  $\bar{p}/p$  multiplicity ratio due to proton background estimation was introduced. Since most of the corrections are the same for particle and its antiparticle, nearly all systematic uncertainties cancel out in the antiparticle-to-particle ratios. The largest sources of systematics, which do not, are related to proton background estimation and dead-material effect on TPC track reconstruction efficiency. The former was found to be up to 5%, whereas the latter varies up to 2% for low- $p_{\text{T}}$   $\bar{p}/p$  multiplicity ratio.

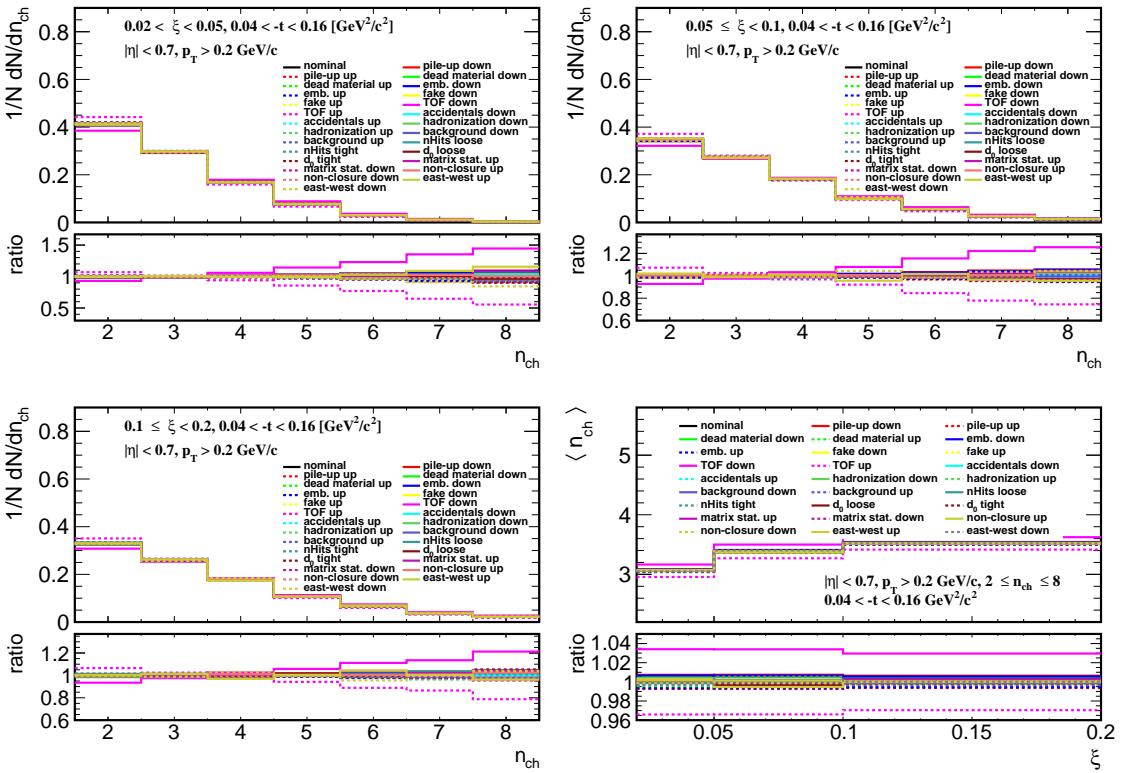


Figure 9.1: Components of the systematic uncertainties for the charged particle multiplicity in three  $\xi$  regions and for the average charged particle multiplicity.

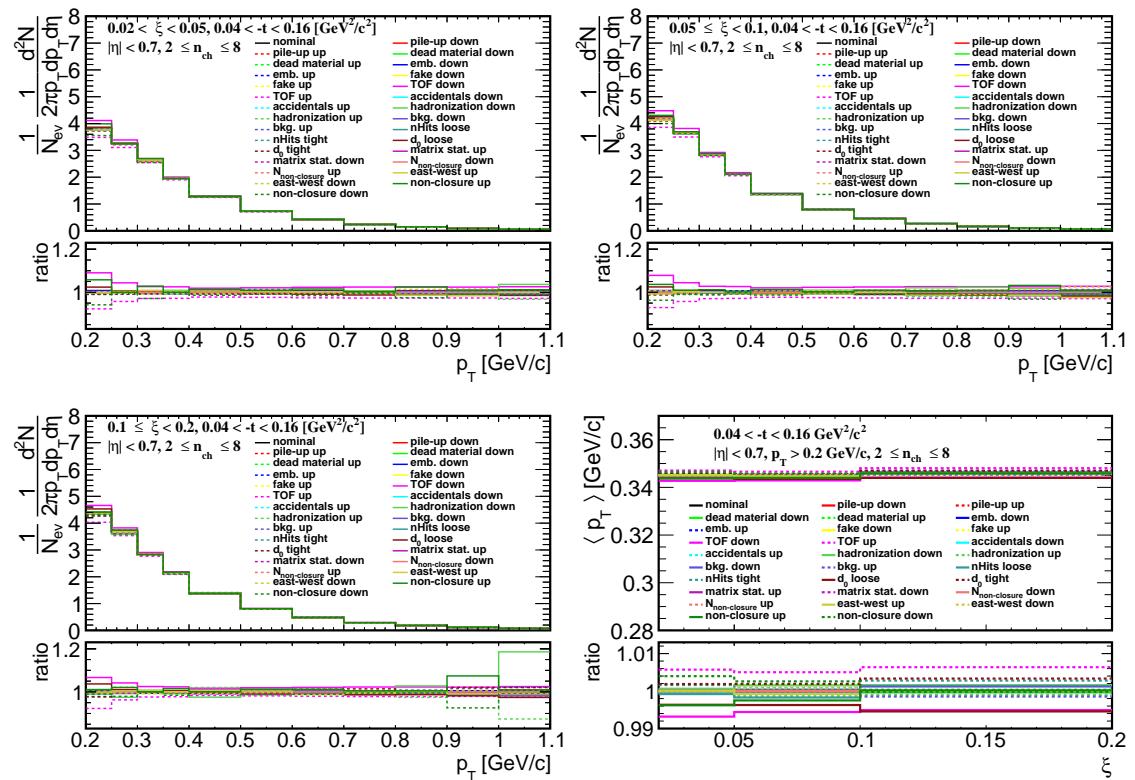


Figure 9.2: Components of the systematic uncertainties for  $p_T$  distributions in three  $\xi$  regions and for an average  $p_T$  distribution.

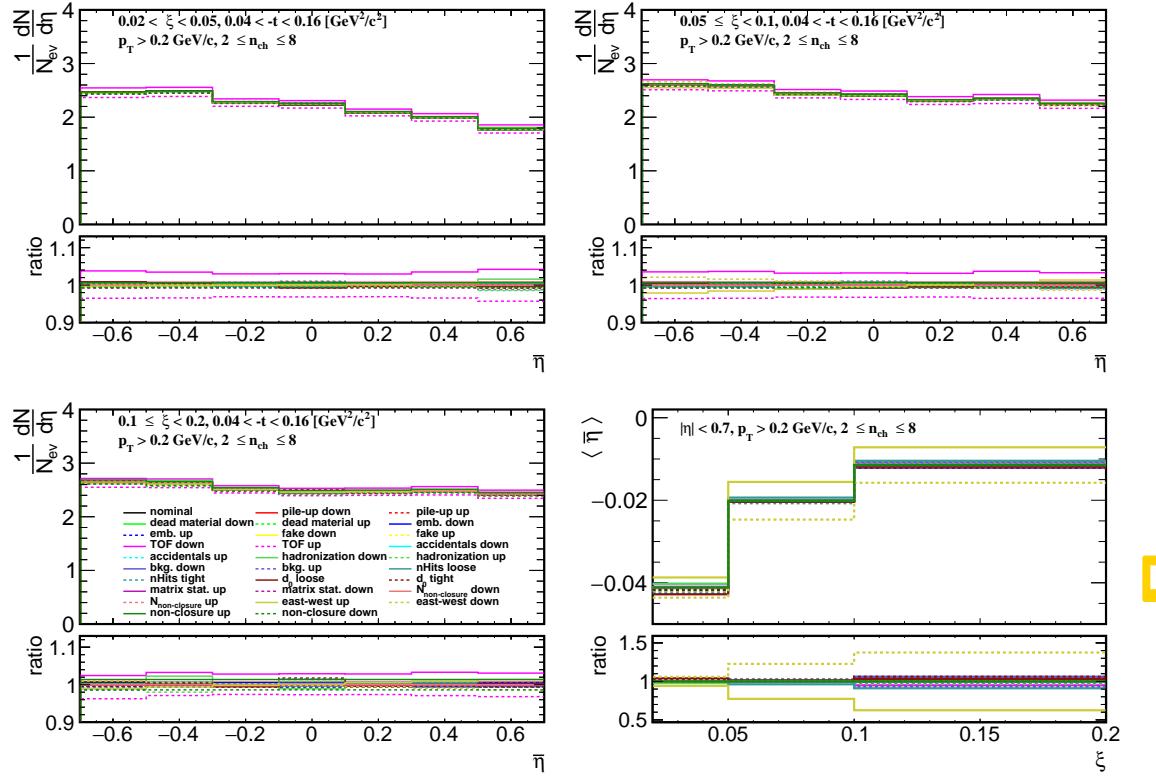


Figure 9.3: Components of the systematic uncertainties for  $\bar{\eta}$  distributions in three  $\xi$  regions and for an average  $\bar{\eta}$  distribution.

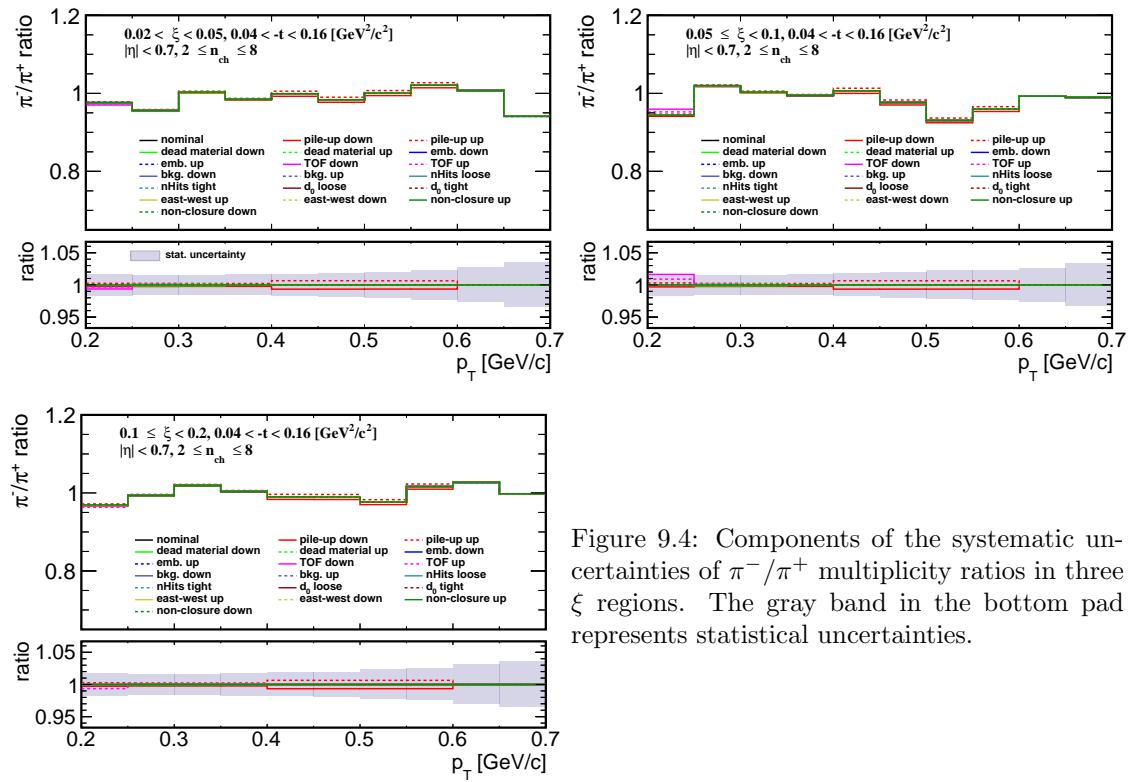


Figure 9.4: Components of the systematic uncertainties of  $\pi^-/\pi^+$  multiplicity ratios in three  $\xi$  regions. The gray band in the bottom pad represents statistical uncertainties.

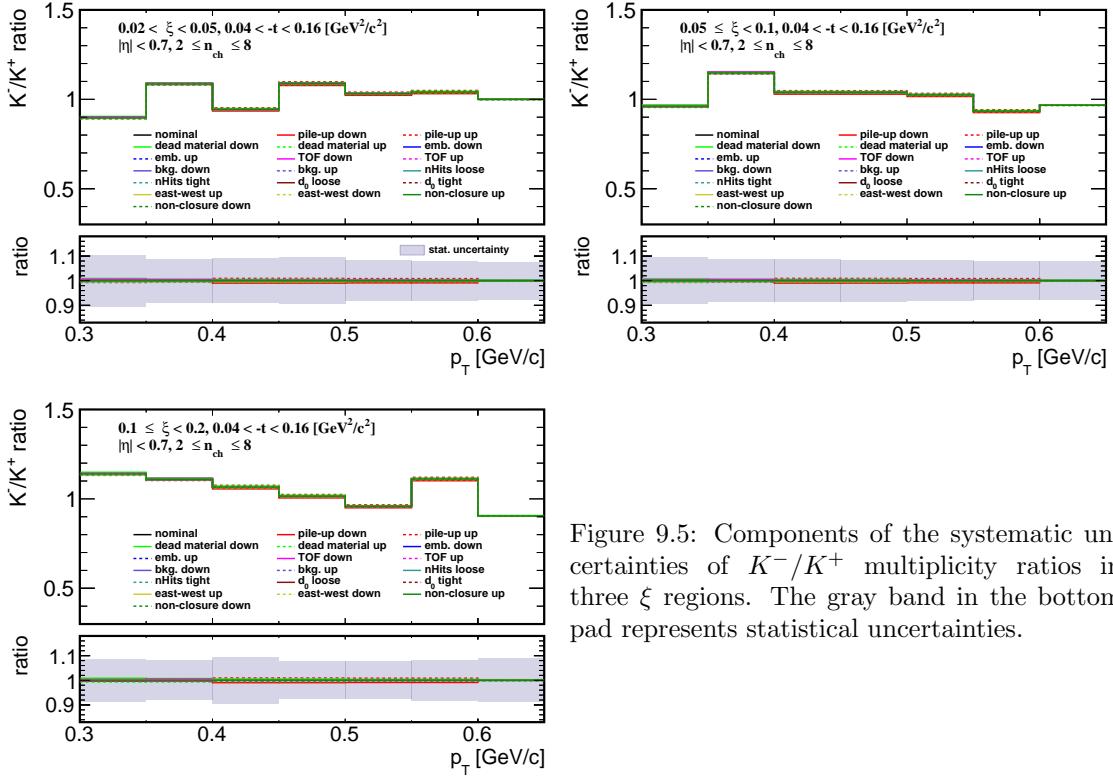


Figure 9.5: Components of the systematic uncertainties of  $K^-/K^+$  multiplicity ratios in three  $\xi$  regions. The gray band in the bottom pad represents statistical uncertainties.

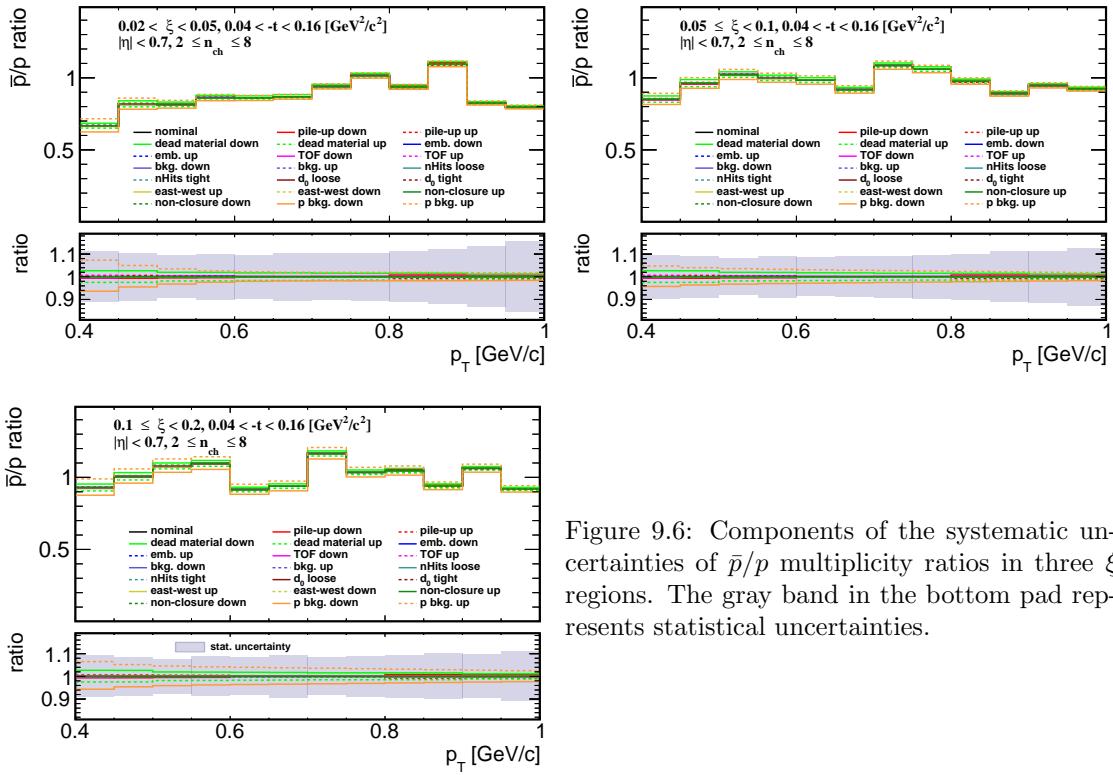


Figure 9.6: Components of the systematic uncertainties of  $\bar{p}/p$  multiplicity ratios in three  $\xi$  regions. The gray band in the bottom pad represents statistical uncertainties.

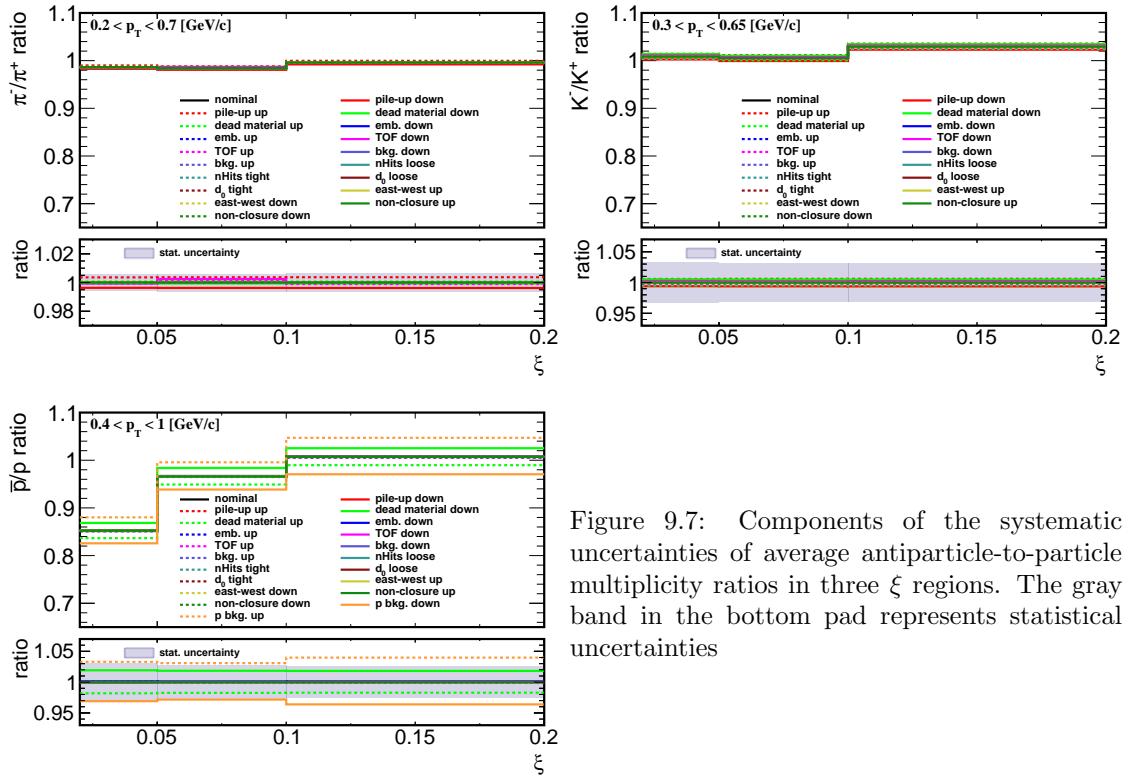


Figure 9.7: Components of the systematic uncertainties of average antiparticle-to-particle multiplicity ratios in three  $\xi$  regions. The gray band in the bottom pad represents statistical uncertainties

# 10. 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) modelled by Pomeron exchange and non-diffractive modelled by low mass excitation of the proton remnant (EPOS-SD'),
- EPOS LHC SD'.

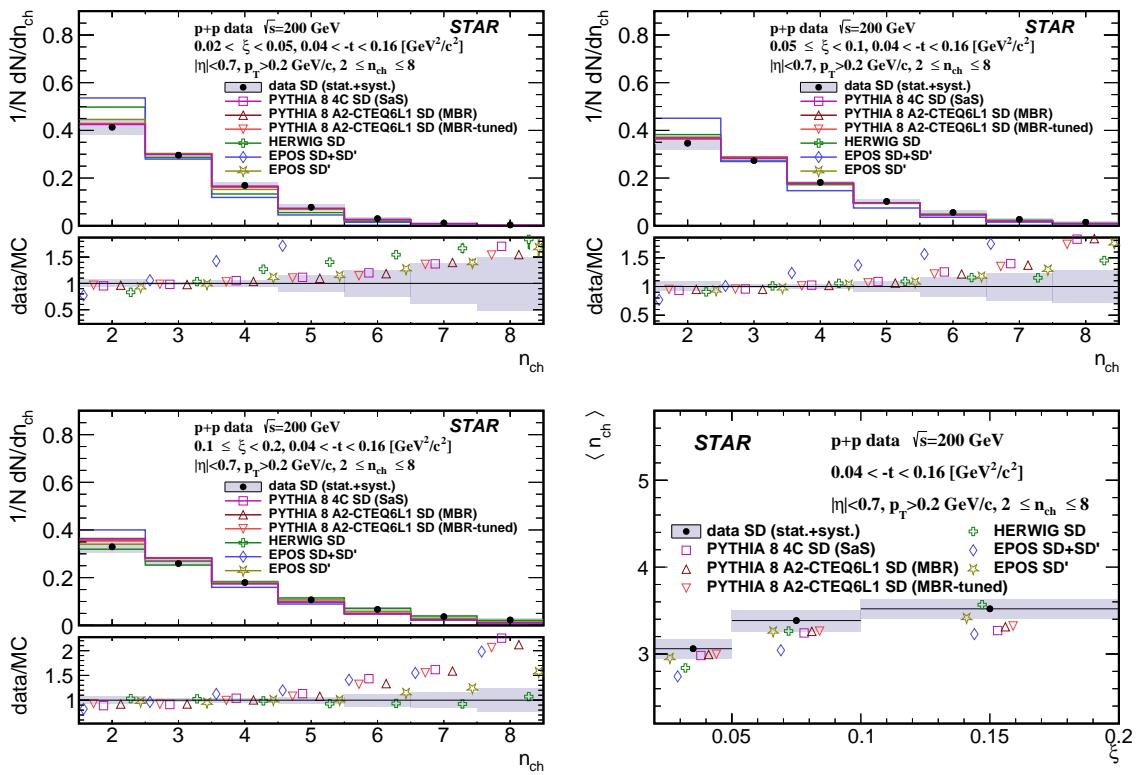


Figure 10.1: Primary charged-particle multiplicity shown separately for the three ranges of  $\xi$ : (top left)  $0.02 < \xi < 0.05$ , (top right)  $0.05 < \xi < 0.1$ , (bottom left)  $0.1 < \xi < 0.2$  and (bottom right) the mean multiplicity  $\langle n_{ch} \rangle$  as a function of  $\xi$ .

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  $\xi$ :  $0.02 < \xi < 0.05$ ,  $0.05 < \xi < 0.1$ ,  $0.1 < \xi < 0.2$ .

838     Figure 10.1 shows primary charged-particle multiplicity separately for the three ranges of  $\xi$   
 839     and the mean multiplicity  $\langle n_{\text{ch}} \rangle$  as a function of  $\xi$ . Data follow the expected increase of  $\langle n_{\text{ch}} \rangle$  with  
 840      $\xi$  due to the larger diffractive masses probed by increasing  $\xi$  in SD process. The shapes of the  
 841     measured distributions are reproduced reasonably well by all models except EPOS SD+SD' and  
 842     HERWIG-SD which predicts smaller  $\langle n_{\text{ch}} \rangle$  for  $0.02 < \xi < 0.1$  and  $0.02 < \xi < 0.05$ , respectively.  
 843     All models, but EPOS SD' and HERWIG-SD, predict too small  $\langle n_{\text{ch}} \rangle$  for  $0.1 < \xi < 0.2$ .

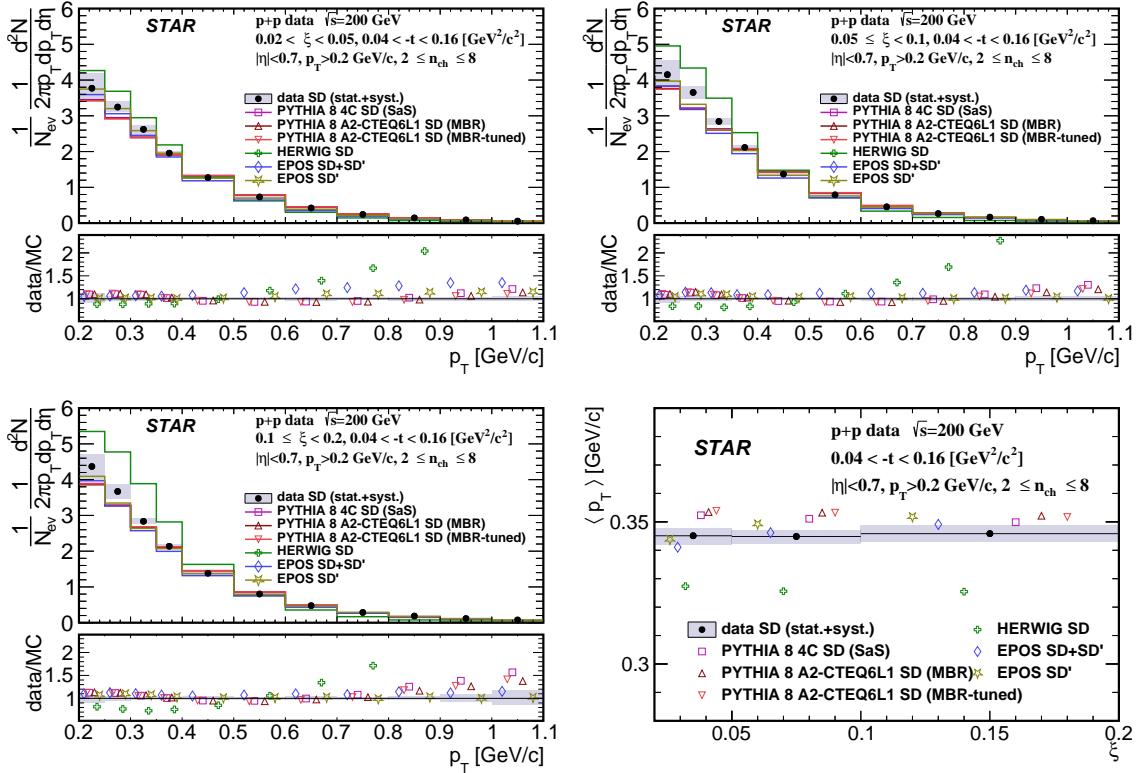


Figure 10.2: Primary charged-particle multiplicities as a function of  $p_T$  shown separately for the three ranges of  $\xi$ : (top left)  $0.02 < \xi < 0.05$ , (top right)  $0.05 < \xi < 0.1$ , (bottom left)  $0.1 < \xi < 0.2$  and (bottom right) the mean transverse momentum  $\langle p_T \rangle$  as a function of  $\xi$ .

844     Figure 10.2 shows primary charged-particle multiplicities as a function of  $p_T$  separately for  
 845     the three ranges of  $\xi$  and the mean transverse momentum  $\langle p_T \rangle$  as a function of  $\xi$ . Data show  
 846     that  $\langle p_T \rangle$  depends very weakly on  $\xi$ . Models describe data fairly well except HERWIG-SD which  
 847     predicts much steeper dependence of particle density with  $p_T$  in all three  $\xi$  ranges.

848     Figure 10.3 shows primary charged-particle multiplicity as a function of  $\bar{\eta}$  separately for  
 849     the three ranges of  $\xi$  and the mean pseudorapidity  $\langle \bar{\eta} \rangle$  as a function of  $\xi$ . Data show expect-  
 850     ed flattening of the  $\bar{\eta}$  distribution with increasing  $\xi$  which reflects SD event-asymmetry and fact  
 851     that the gap-edge at large  $\xi$  is outside  $|\bar{\eta}| < 0.7$  region leading to more flat distribution of particle  
 852     density as a function of  $\bar{\eta}$ . Models describe data fairly well for  $0.02 < \xi < 0.05$  except EPOS  
 853     SD which predicts smaller charged-particle density. All MCs, but EPOS SD' and HERWIG-SD,  
 854     predict less steep dependence of particle density with  $\bar{\eta}$  in the range of  $0.05 < \xi < 0.2$ .

855     Figure 10.4 shows the ratio of production yields of  $\pi^-/\pi^+$  as a function of  $p_T$  separately for  
 856     the three ranges of  $\xi$ . Data in all three  $\xi$  ranges are consistent with equal amounts of  $\pi^+$  and  
 857      $\pi^-$  with no significant  $p_T$  dependence. Models agree with data (except HERWIG) predicting on  
 858     average small deviation from unity by  $\sim 2\%$  what is smaller than data uncertainties. HERWIG  
 859     in first two  $\xi$  ranges predicts too large asymmetry between  $\pi^+$  and  $\pi^-$ .

860     Figure 10.5 shows the ratio of production yields of  $K^-/K^+$  as a function of  $p_T$  separately for

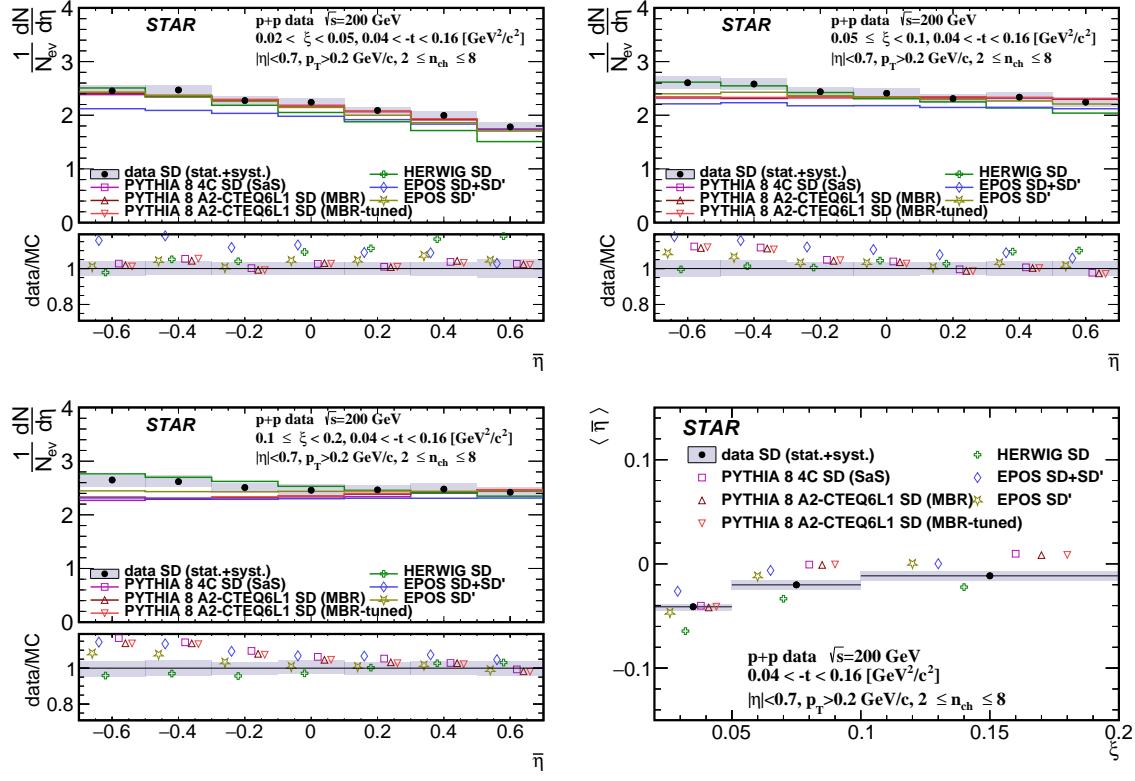


Figure 10.3: Primary charged-particle multiplicity as a function of  $\bar{\eta}$  shown separately for the three ranges of  $\xi$ : (top left)  $0.02 < \xi < 0.05$ , (top right)  $0.05 < \xi < 0.1$ , (bottom left)  $0.1 < \xi < 0.2$  and (bottom right) the mean pseudorapidity  $\langle \bar{\eta} \rangle$  as a function of  $\xi$ .

the three ranges of  $\xi$ . Data in all three  $\xi$  ranges are consistent with equal amounts of  $K^+$  and  $K^-$  with no  $p_T$  dependence. Models agree with data except HERWIG in the first  $\xi$  range predicting too large ratio of  $K^-$  to  $K^+$ .

Figure 10.6 shows the ratio of production yields of  $\bar{p}/p$  as a function of  $p_T$  separately for the three ranges of  $\xi$ . Data in the last two  $\xi$  ranges are consistent with equal amounts of  $p$  and  $\bar{p}$  with no  $p_T$  dependence. However, in the first  $\xi$  range at  $p_T < 0.7$  GeV/c data shows significant deviation from unity indicating a significant transfer of the baryon number from the forward to the central region. PYTHIA8, EPOS SD' and EPOS SD+SD' agree with data in the last two  $\xi$  ranges. In first  $\xi$  range PYTHIA8 and EPOS SD' predict small deviation from unity by  $\sim 5\%$  which is smaller than observed in data, whereas EPOS SD+SD' predicts an asymmetry between  $\bar{p}$  and  $p$  of  $\sim 30\%$  which is larger than observed in data except  $p_T < 0.5$  GeV/c. HERWIG predicts much larger baryon number transfer compared to data in first two  $\xi$  ranges and shows consistency with data in last  $\xi$  range.

Figure 10.7 shows mean ratio of production yields of  $\pi^-/\pi^+$ ,  $K^-/K^+$  and  $\bar{p}/p$  as a function of  $\xi$ .

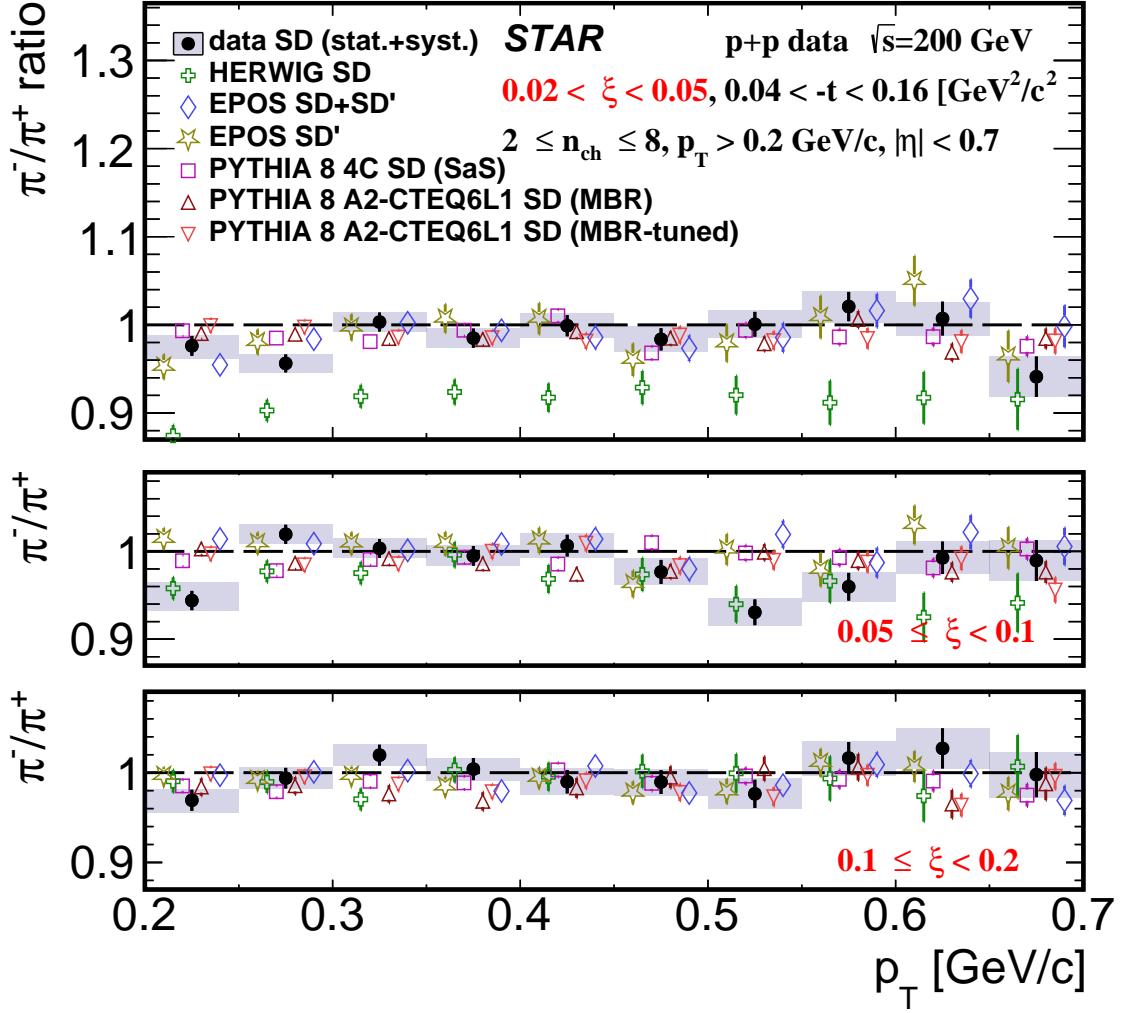


Figure 10.4: Ratio of production yields of  $\pi^-/\pi^+$  as a function of  $p_T$  shown separately for the three ranges of  $\xi$ : (top)  $0.02 < \xi < 0.05$ , (middle)  $0.05 < \xi < 0.1$ , (bottom)  $0.1 < \xi < 0.2$ .

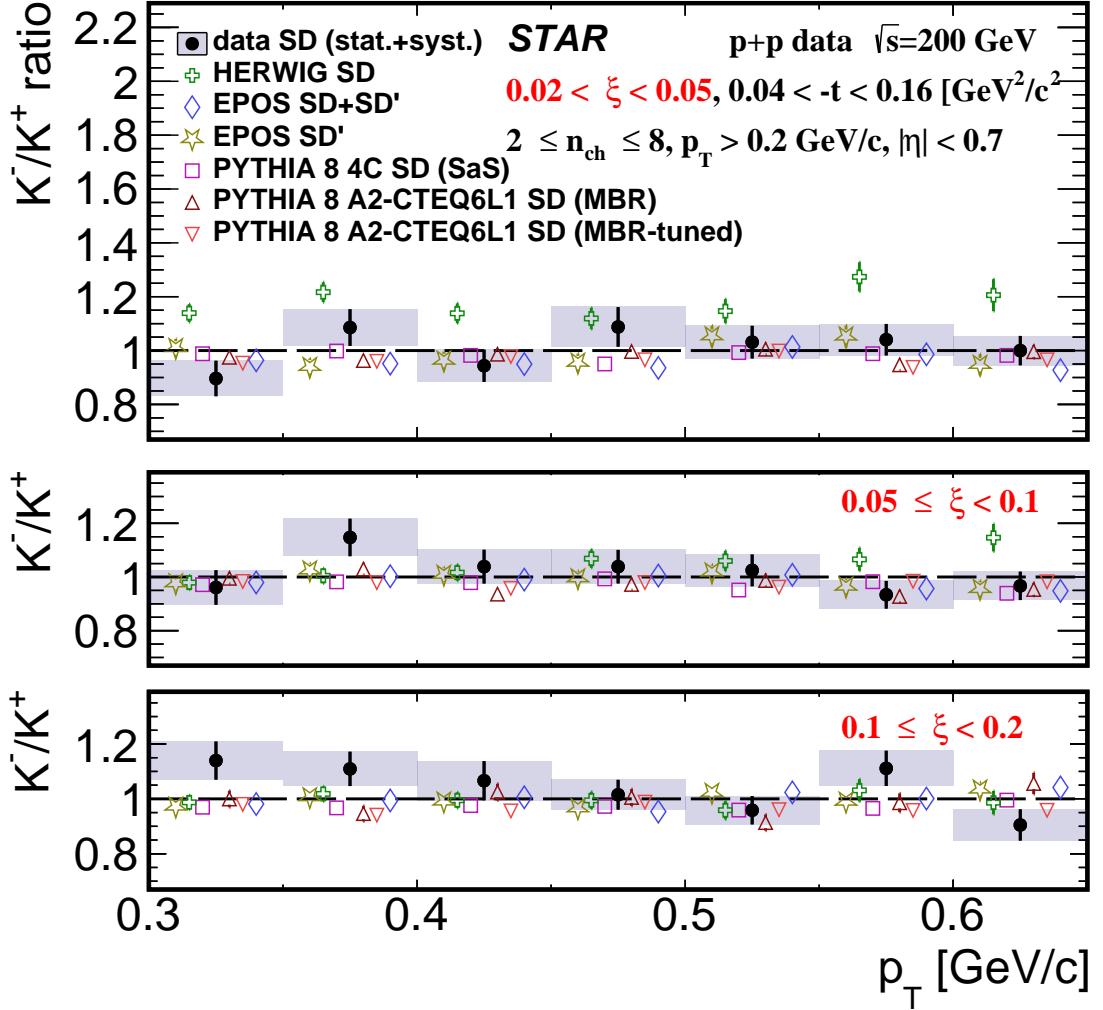


Figure 10.5: Ratio of production yields of  $K^-/K^+$  as a function of  $p_T$  shown separately for the three ranges of  $\xi$ : (top)  $0.02 < \xi < 0.05$ , (middle)  $0.05 < \xi < 0.1$ , (bottom)  $0.1 < \xi < 0.2$ .

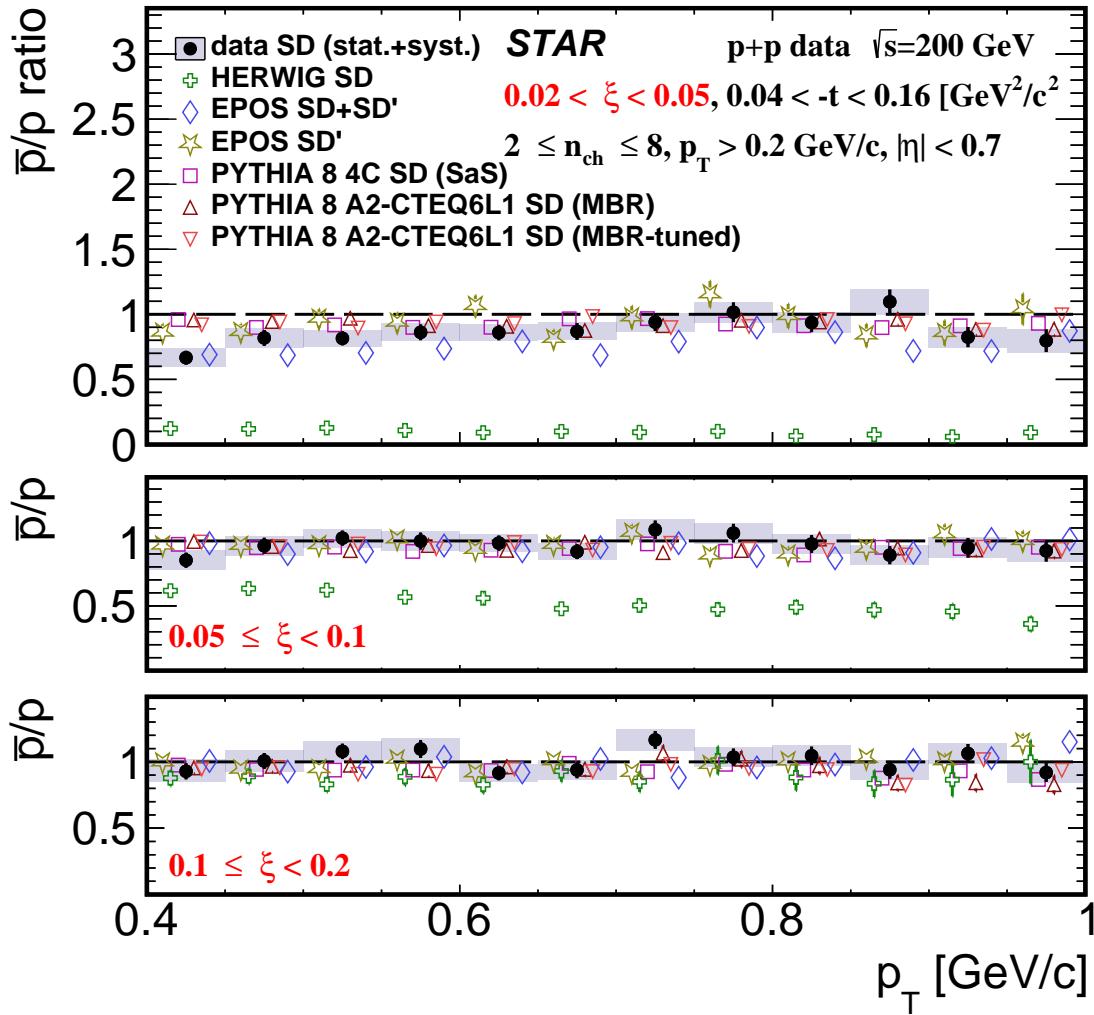


Figure 10.6: Ratio of production yields of  $\bar{p}/p$  as a function of  $p_T$  shown separately for the three ranges of  $\xi$ : (top)  $0.02 < \xi < 0.05$ , (middle)  $0.05 < \xi < 0.1$ , (bottom)  $0.1 < \xi < 0.2$ .

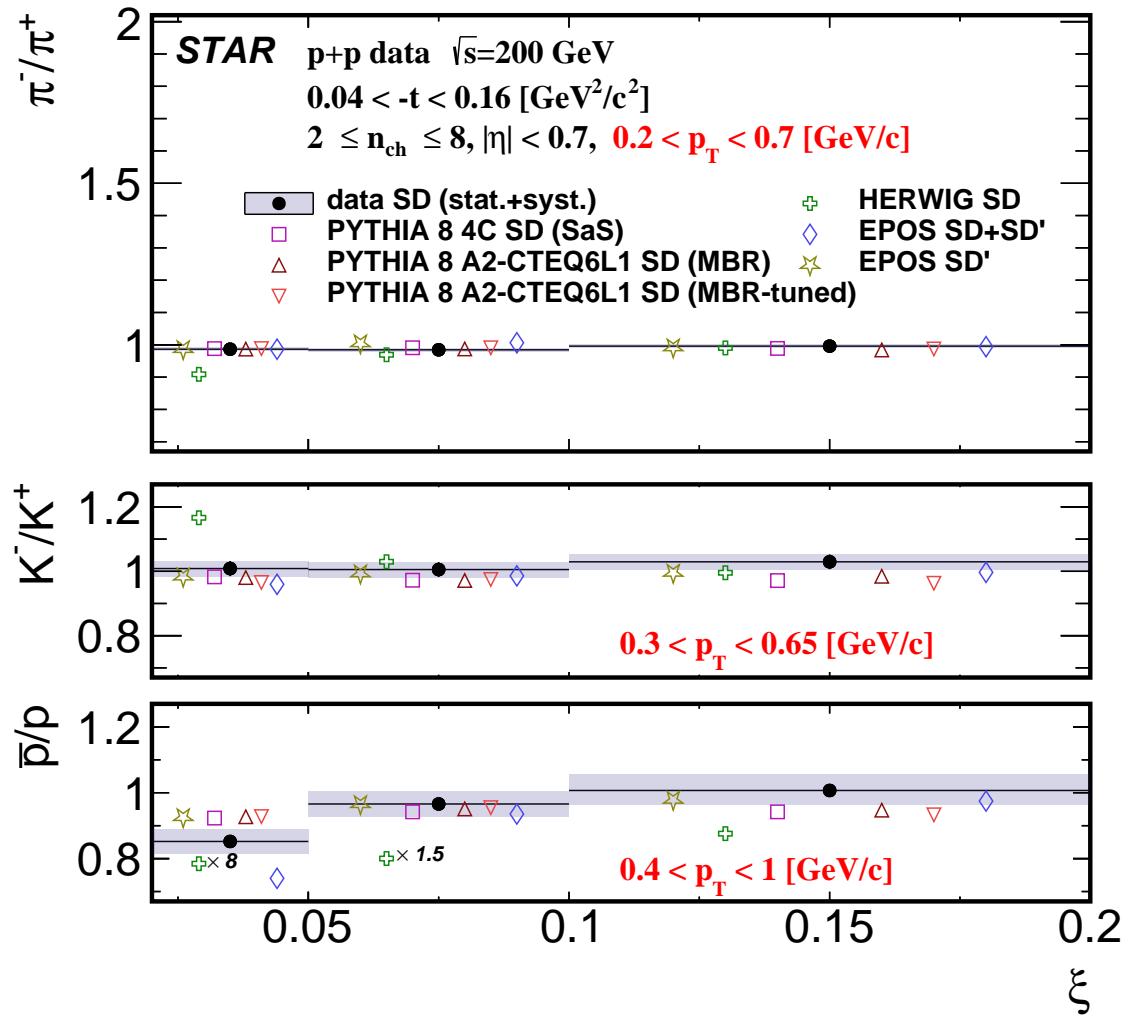


Figure 10.7: Ratio of production yields of  $\pi^-/\pi^+$ ,  $K^-/K^+$  and  $\bar{p}/p$  as a function of  $\xi$ .

# 876 11. Summary and Conclusions

877 Inclusive and identified (pion, kaon, proton and their antiparticles) charged particle production in  
878 Single Diffractive Dissociation process has been measured in proton-proton collisions at  $\sqrt{s} = 200$   
879 GeV with the STAR detector at RHIC using data corresponding to an integrated luminosity of  
880  $15 \text{ nb}^{-1}$ .

881 Significant differences are observed between the measured distributions of  $\xi$  and Monte Carlo  
882 model predictions. Amongst the models considered EPOS and PYTHIA8 (MBR) without sup-  
883 pression of diffractive cross sections at large  $\xi$  provide the best description of the data.

884 Primary-charged-particle multiplicity and its density as a function of pseudorapidity and trans-  
885 verse momentum are well described by PYTHIA8 and EPOS-SD' models. EPOS-SD and HERWIG  
886 do not describe the data.

887  $\pi^-/\pi^+$  and  $K^-/K^+$  production ratios are close to unity and consistent with most of model  
888 predictions except for EPOS-SD and HERWIG.

889  $\bar{p}/p$  production ratio shows a significant deviation from unity in the  $0.02 < \xi < 0.05$  range  
890 indicating a non-negligible transfer of the baryon number from the forward to the central region.  
891 Equal amount of protons and antiprotons are observed in the  $\xi > 0.05$  range. PYTHIA8 and  
892 EPOS-SD' agree with data for  $\xi > 0.05$ . For  $0.02 < \xi < 0.05$  they predict small deviations from  
893 unity (0.93) which is however higher than observed in data ( $0.86 \pm 0.02$ ). HERWIG and EPOS-SD  
894 predict much larger baryon number transfers compared to data for  $\xi < 0.1$  and show consistency  
895 with data for  $\xi > 0.1$ .

896    

# Bibliography

- 897    [1] L. Adamczyk, L. Fulek, and R. Sikora, *Supplementary note on diffractive analyses of 2015*  
898    *proton-proton data*, <https://drupal.star.bnl.gov/STAR/starnotes/private/psn0732>,  
899    July, 2019.
- 900    [2] B. Z. Kopeliovich and B. G. Zakharov, *Novel Mechanisms of Baryon Number Flow Over*  
901    *Large Rapidity Gap*, *Z. Phys. C***43** (1989) 241.
- 902    [3] F. W. Bopp, *Baryon transport in dual models and the possibility of a backward peak in*  
903    *diffraction*, in *New trends in high-energy physics: Experiment, phenomenology, theory.*  
904    *Proceedings, International School-Conference, Crimea 2000, Yalta, Ukraine, May 27-June 4,*  
905    *2000*. [arXiv:hep-ph/0007229](https://arxiv.org/abs/hep-ph/0007229) [hep-ph].
- 906    [4] M. Anderson et al., *The STAR Time Projection Chamber: A Unique tool for studying high*  
907    *multiplicity events at RHIC*, *Nucl. Instrum. Meth. A***499** (2003) 659–678,  
908    [arXiv:nucl-ex/0301015](https://arxiv.org/abs/nucl-ex/0301015) [nucl-ex].
- 909    [5] R. Brun, F. Bruyant, M. Maire, A. C. McPherson, and P. Zanarini, *GEANT3*,  
910    *CERN-DD-EE-84-1* (1987).
- 911    [6] GEANT4 Collaboration, S. Agostinelli et al., *GEANT4: A Simulation toolkit*, *Nucl.*  
912    *Instrum. Meth. A***506** (2003) 250–303.
- 913    [7] T. Sjöstrand, S. Mrenna, and P. Z. Skands, *PYTHIA 6.4 Physics and Manual*, *JHEP* **05**  
914    (2006) 026.
- 915    [8] R. Ciesielski and K. Goulianos, *MBR Monte Carlo Simulation in PYTHIA8*, *PoS*  
916    *ICHEP2012* (2013) 301, [arXiv:1205.1446](https://arxiv.org/abs/1205.1446) [hep-ph].
- 917    [9] *Roman Pot Phase II\* - online DAQ monitoring*,  
918    <https://online.star.bnl.gov/rp/pp200/>.
- 919    [10] *STAR RunLog Browser*, <https://online.star.bnl.gov/RunLogRun15/>.
- 920    [11] W. Fischer, *Run Overview of the Relativistic Heavy Ion Collider*,  
921    <http://www.agsrhichome.bnl.gov/RHIC/Runs/>.
- 922    [12] STAR Collaboration, B. I. Abelev et al., *Systematic Measurements of Identified Particle*  
923    *Spectra in pp, d<sup>+</sup> Au and Au+Au Collisions from STAR*, *Phys. Rev. C***79** (2009) 034909,  
924    [arXiv:0808.2041](https://arxiv.org/abs/0808.2041) [nucl-ex].
- 925    [13] ATLAS Collaboration, G. Aad et al., *Charged-particle distributions in  $\sqrt{s} = 13$  TeV pp*  
926    *interactions measured with the ATLAS detector at the LHC*, *Phys. Lett. B***758** (2016)  
927    67–88, [arXiv:1602.01633](https://arxiv.org/abs/1602.01633) [hep-ex].
- 928    [14] G. D’Agostini, *A Multidimensional unfolding method based on Bayes’ theorem*, *Nucl.*  
929    *Instrum. Meth. A***362** (1995) 487–498.
- 930    [15] H. Bichsel, *A method to improve tracking and particle identification in TPCs and silicon*  
931    *detectors*, *Nucl. Instrum. Meth. A***562** (2006) 154–197.

# Appendices

932

## <sup>933</sup> A. Acronyms

<sup>934</sup>	<b>CD</b>	Central Diffraction
<sup>935</sup>	<b>DD</b>	Double Diffraction
<sup>936</sup>	<b>MBR</b>	Minimum Bias Rockefeller
<sup>937</sup>	<b>MC</b>	Monte Carlo
<sup>938</sup>	<b>ND</b>	Non-Diffractive
<sup>939</sup>	<b>QCD</b>	Quantum Chromodynamics
<sup>940</sup>	<b>RP</b>	Roman Pot
<sup>941</sup>	<b>SaS</b>	Schuler and Sjöstrand
<sup>942</sup>	<b>SD</b>	Single Diffraction
<sup>943</sup>	<b>TPC</b>	Time Projection Chamber

944 **B. Proton and Antiproton DCA  
Distributions**

945

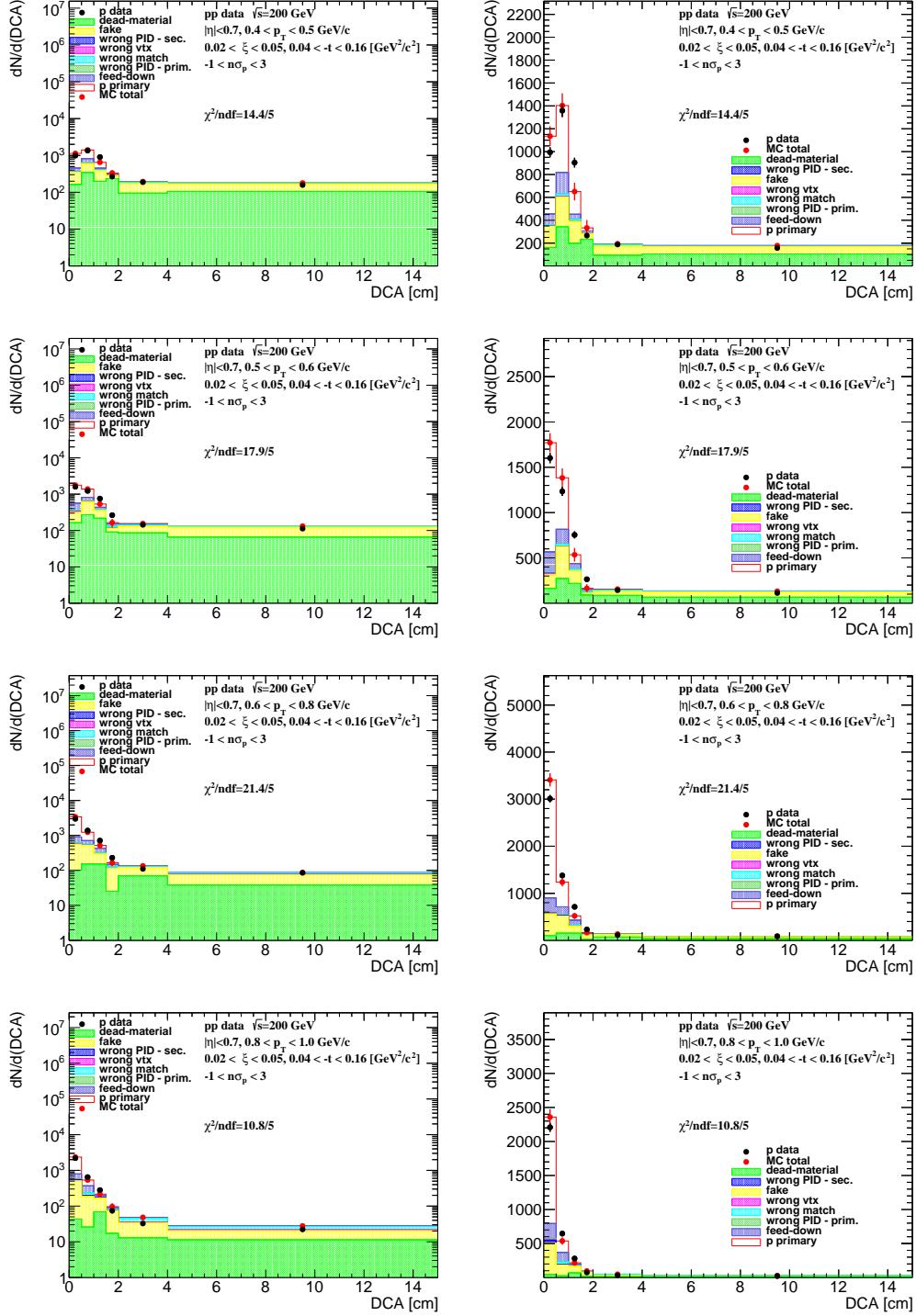


Figure B.1: Distributions of DCA for protons in SD interactions with  $0.02 < \xi < 0.05$  and loose selection.

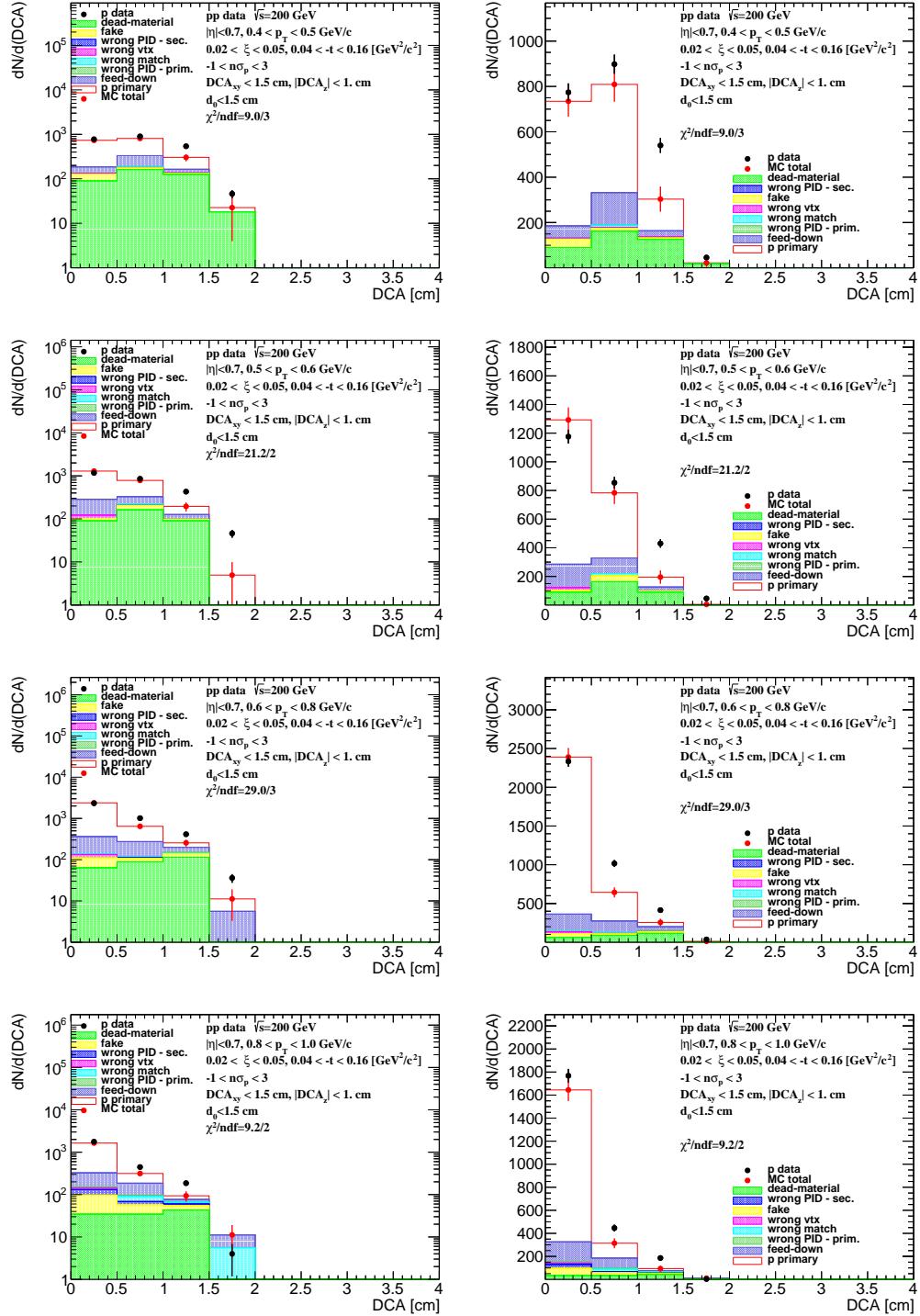


Figure B.2: Distributions of DCA for protons in SD interactions with  $0.02 < \xi < 0.05$  and normal selection.

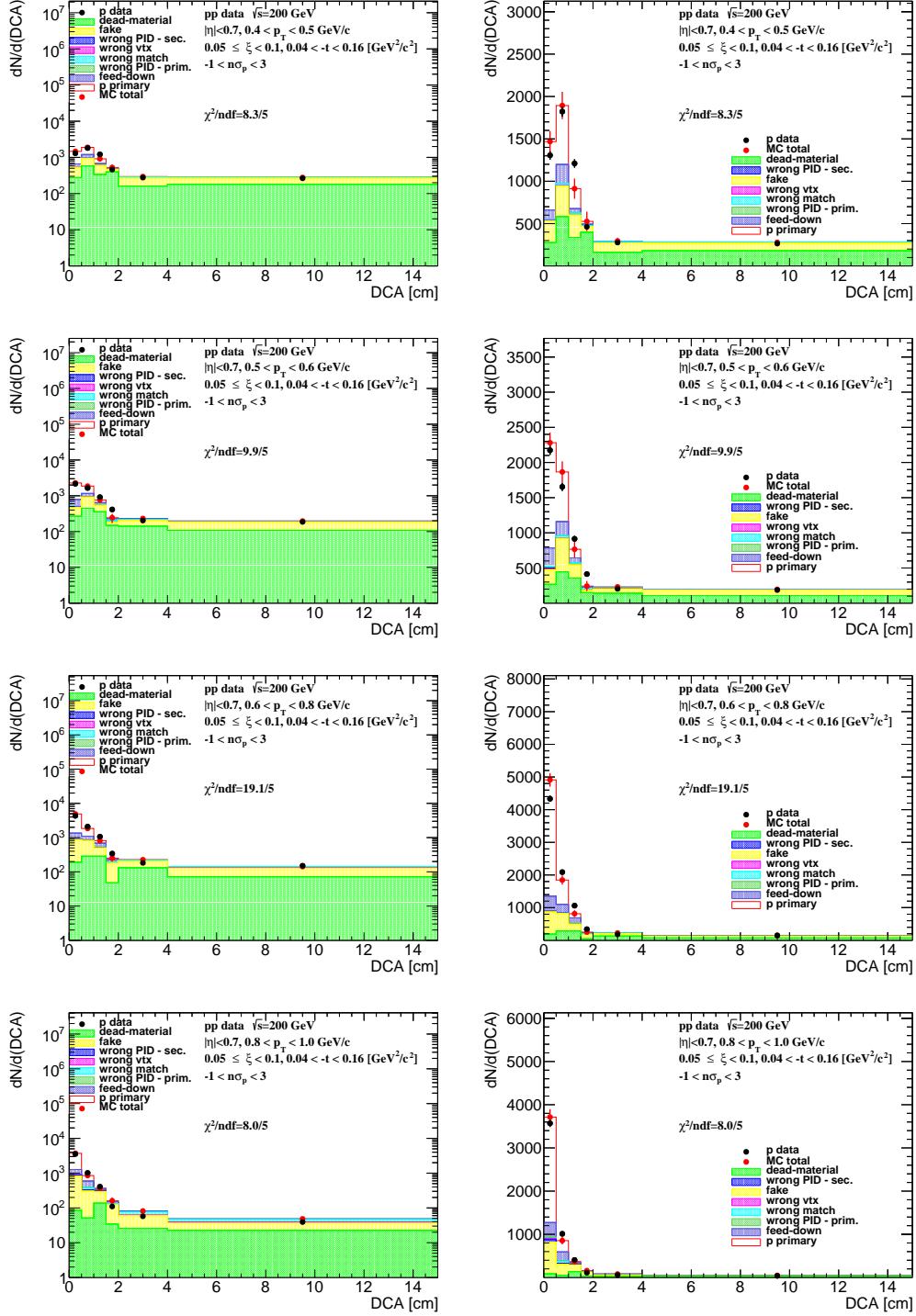


Figure B.3: Distributions of DCA for protons in SD interactions with  $0.05 < \xi < 0.1$  and loose selection.

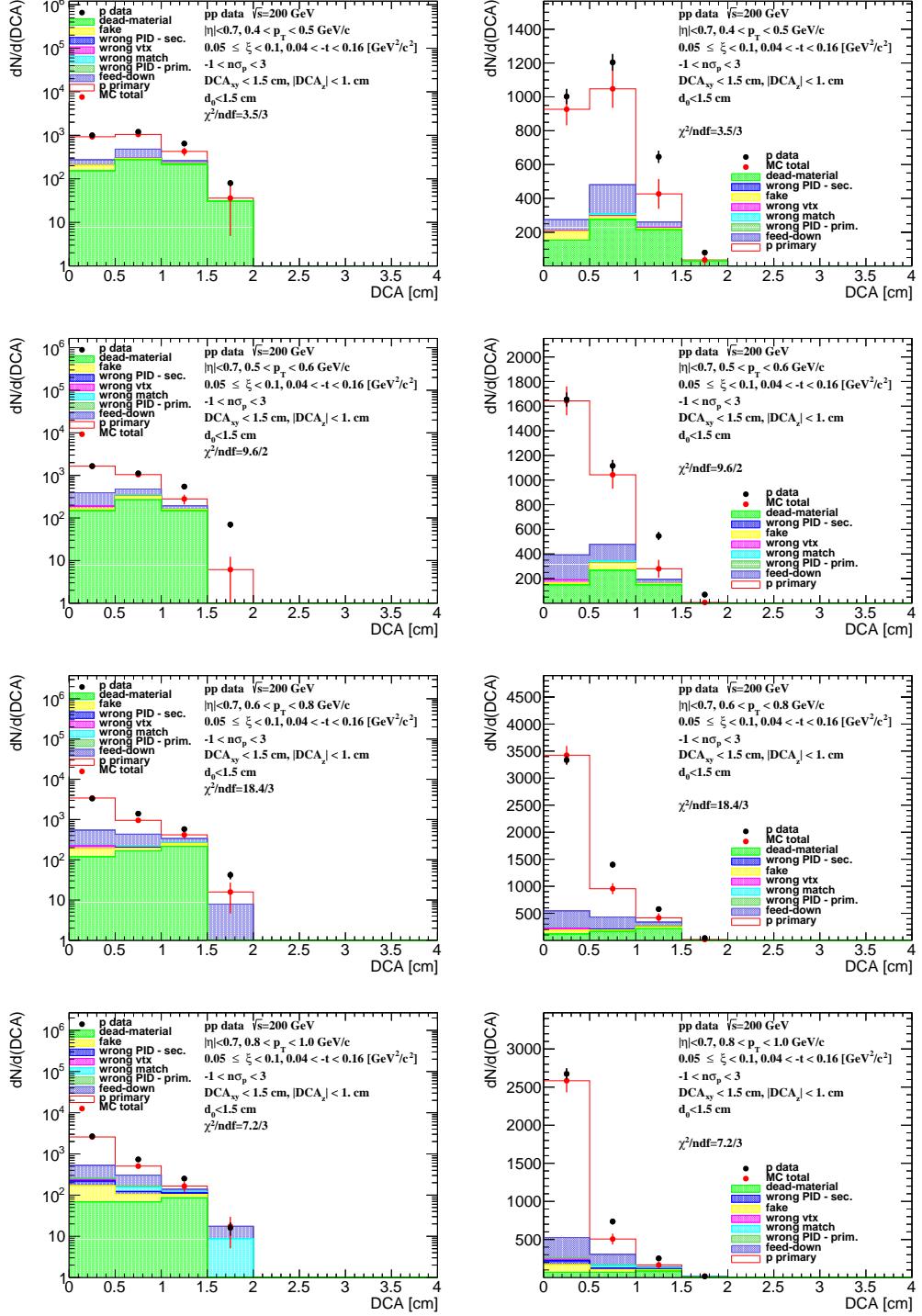


Figure B.4: Distributions of DCA for protons in SD interactions with  $0.05 < \xi < 0.1$  and normal selection.

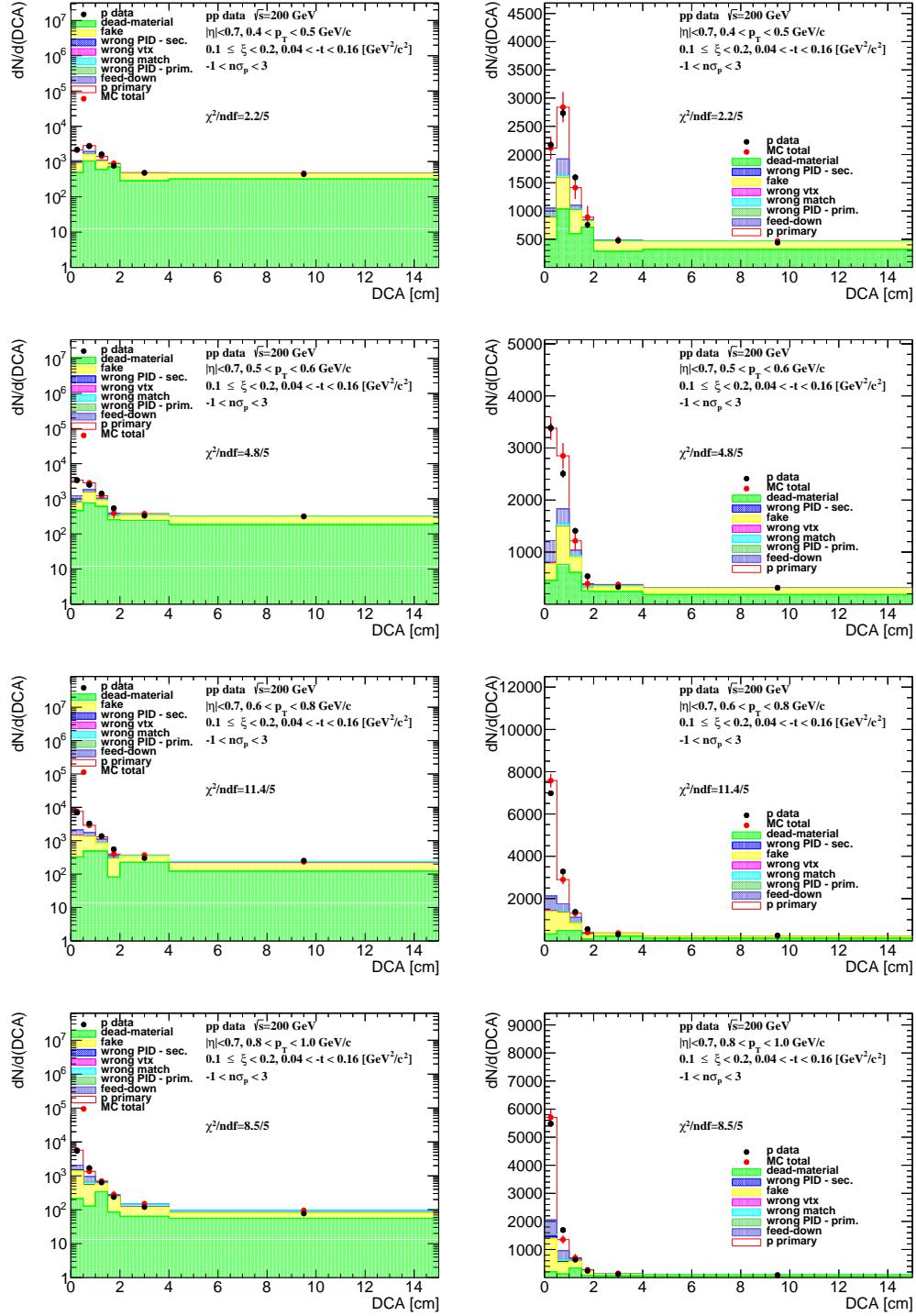


Figure B.5: Distributions of DCA for protons in SD interactions with  $0.1 < \xi < 0.2$  and loose selection.

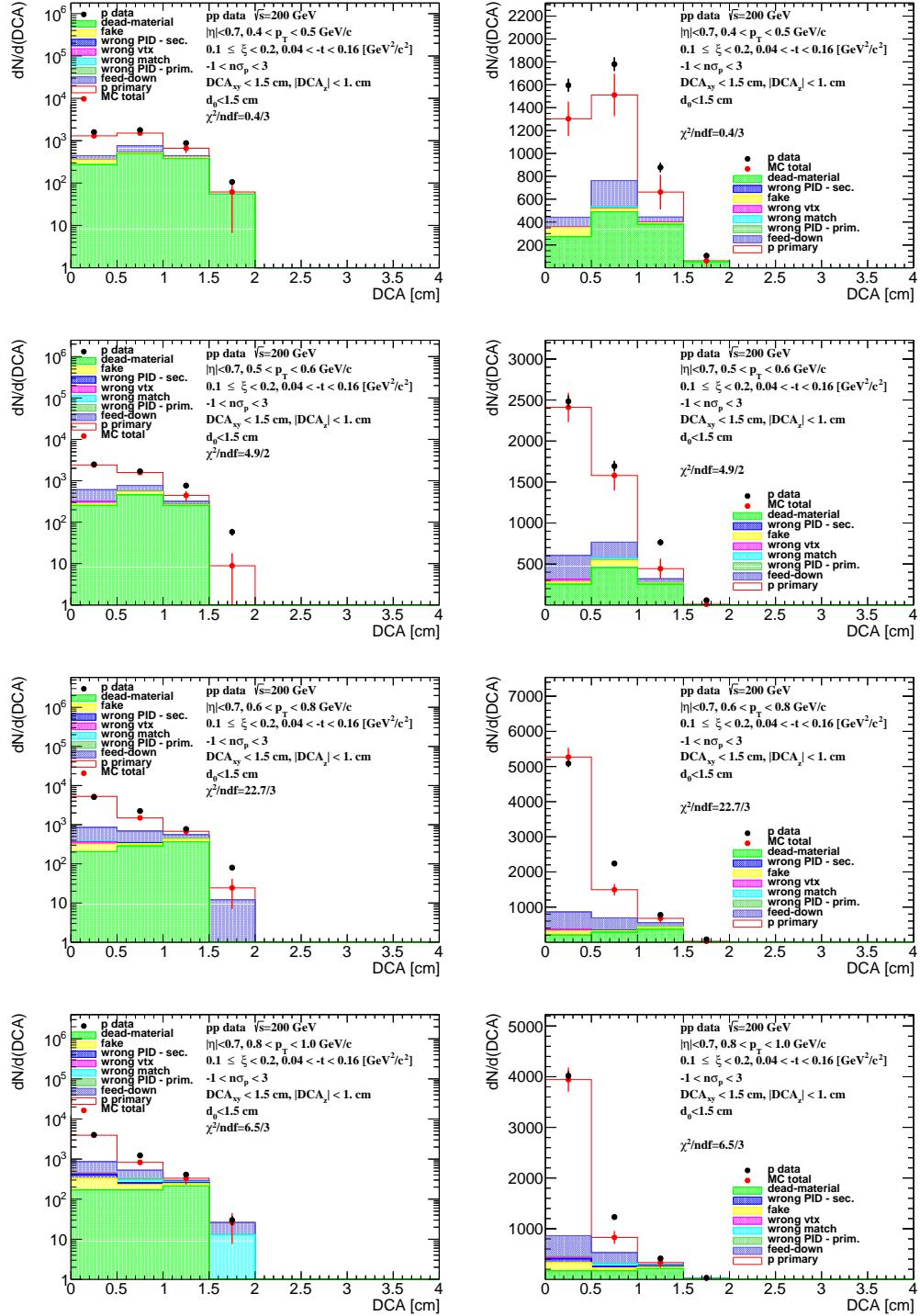


Figure B.6: Distributions of DCA for protons in SD interactions with  $0.1 < \xi < 0.2$  and normal selection.

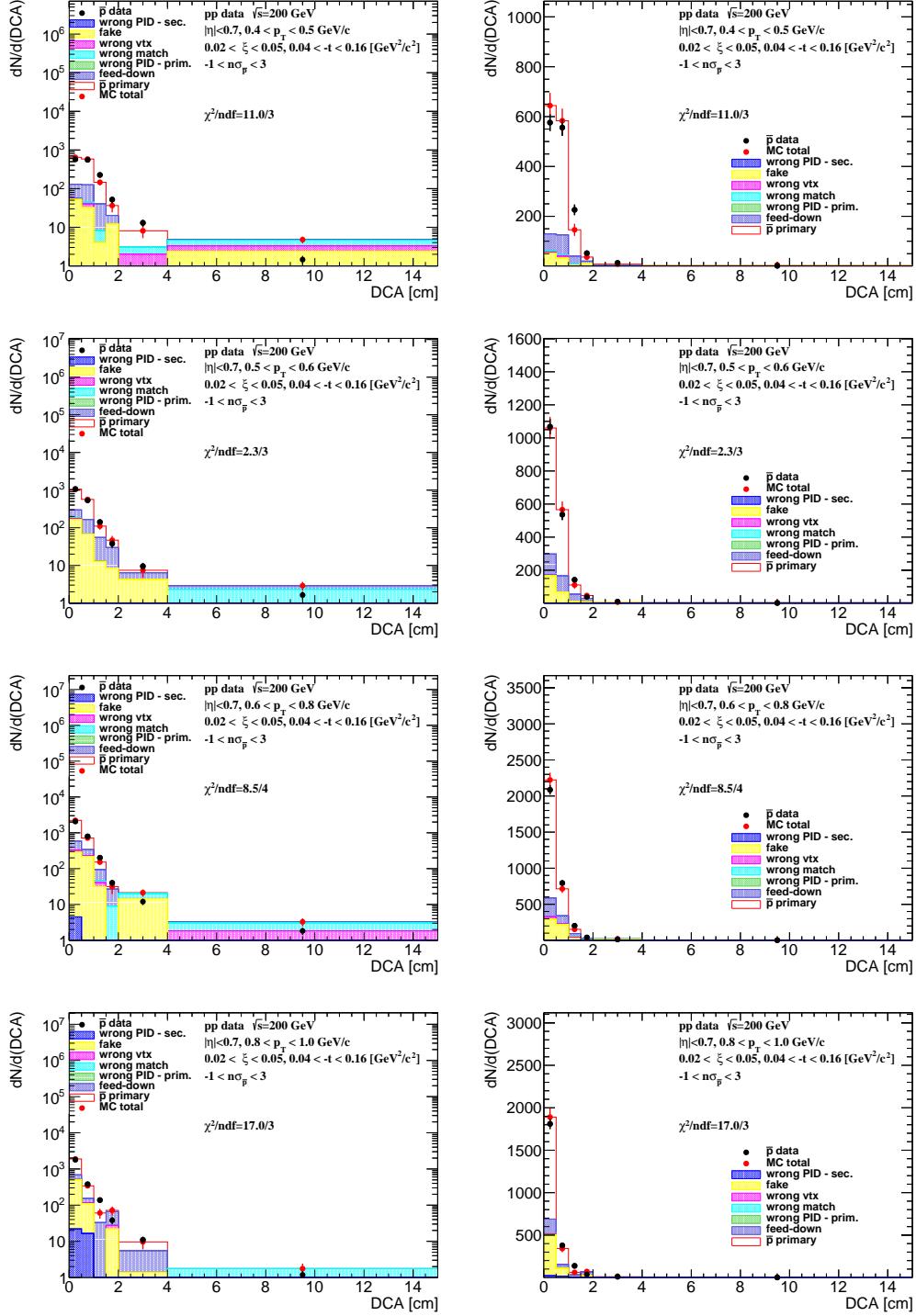


Figure B.7: Distributions of DCA for antiprotons in SD interactions with  $0.02 < \xi < 0.05$  and loose selection.

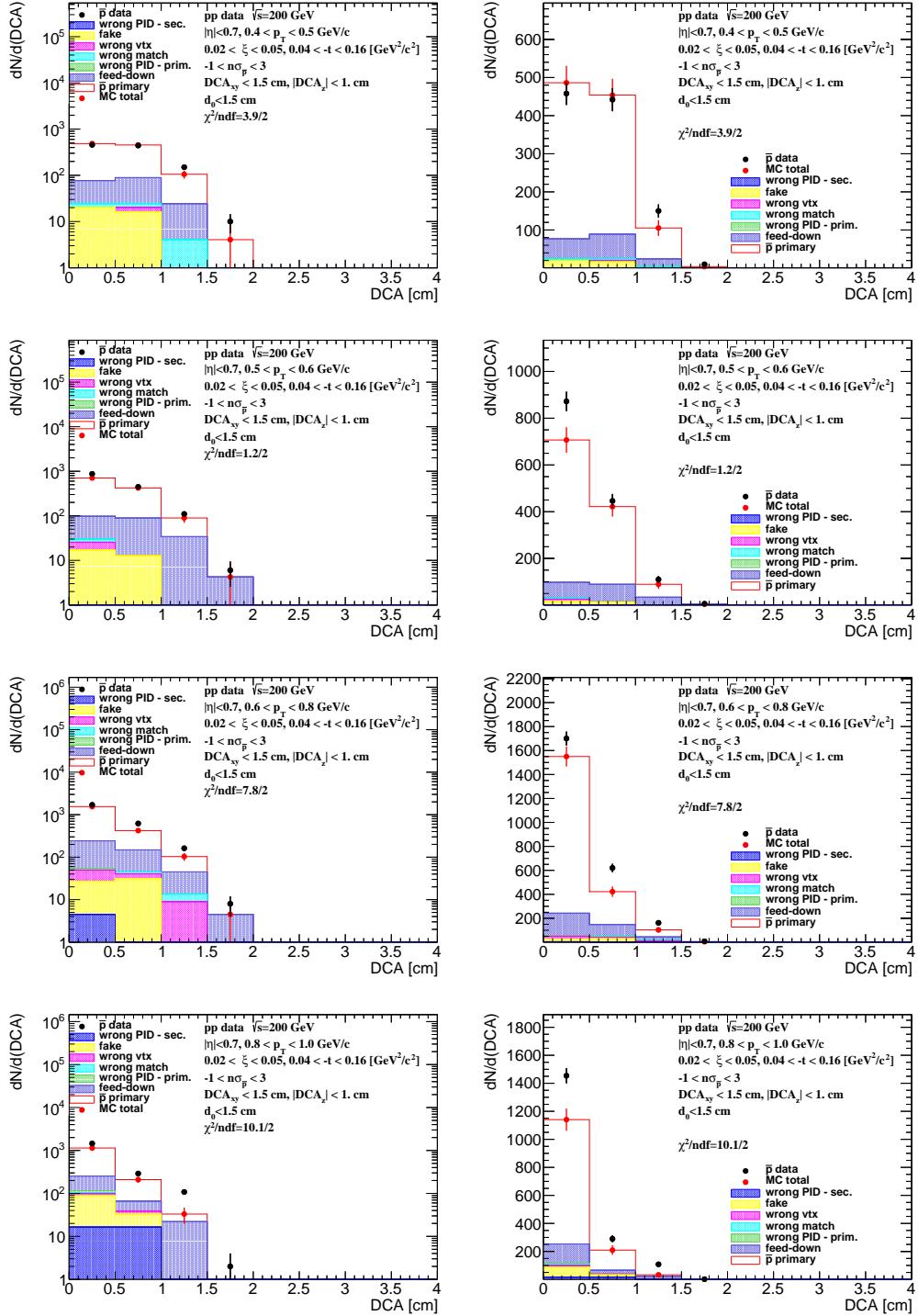


Figure B.8: Distributions of DCA for antiprotons in SD interactions with  $0.02 < \xi < 0.05$  and normal selection.

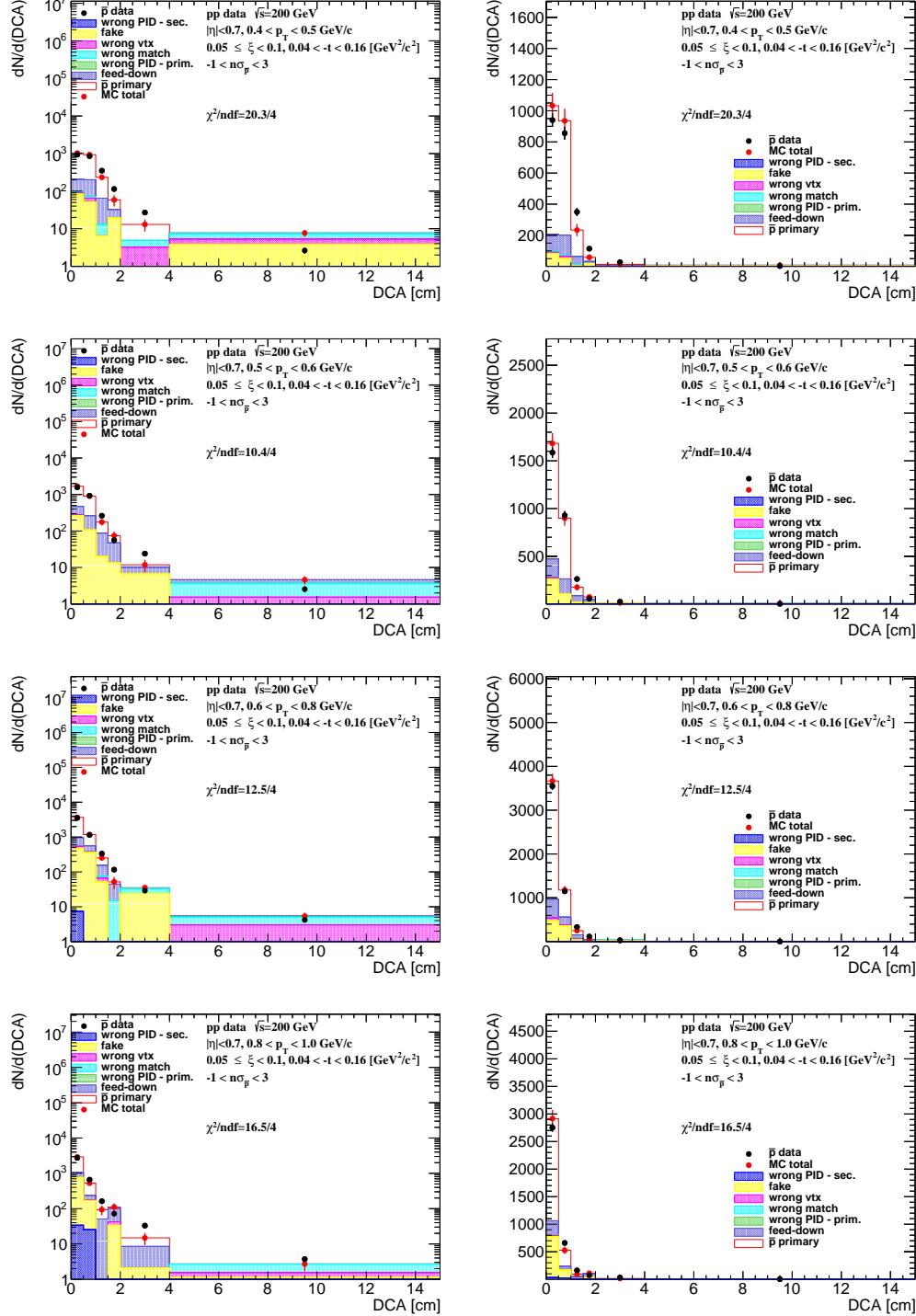


Figure B.9: Distributions of DCA for antiprotons in SD interactions with  $0.05 < \xi < 0.1$  and loose selection.

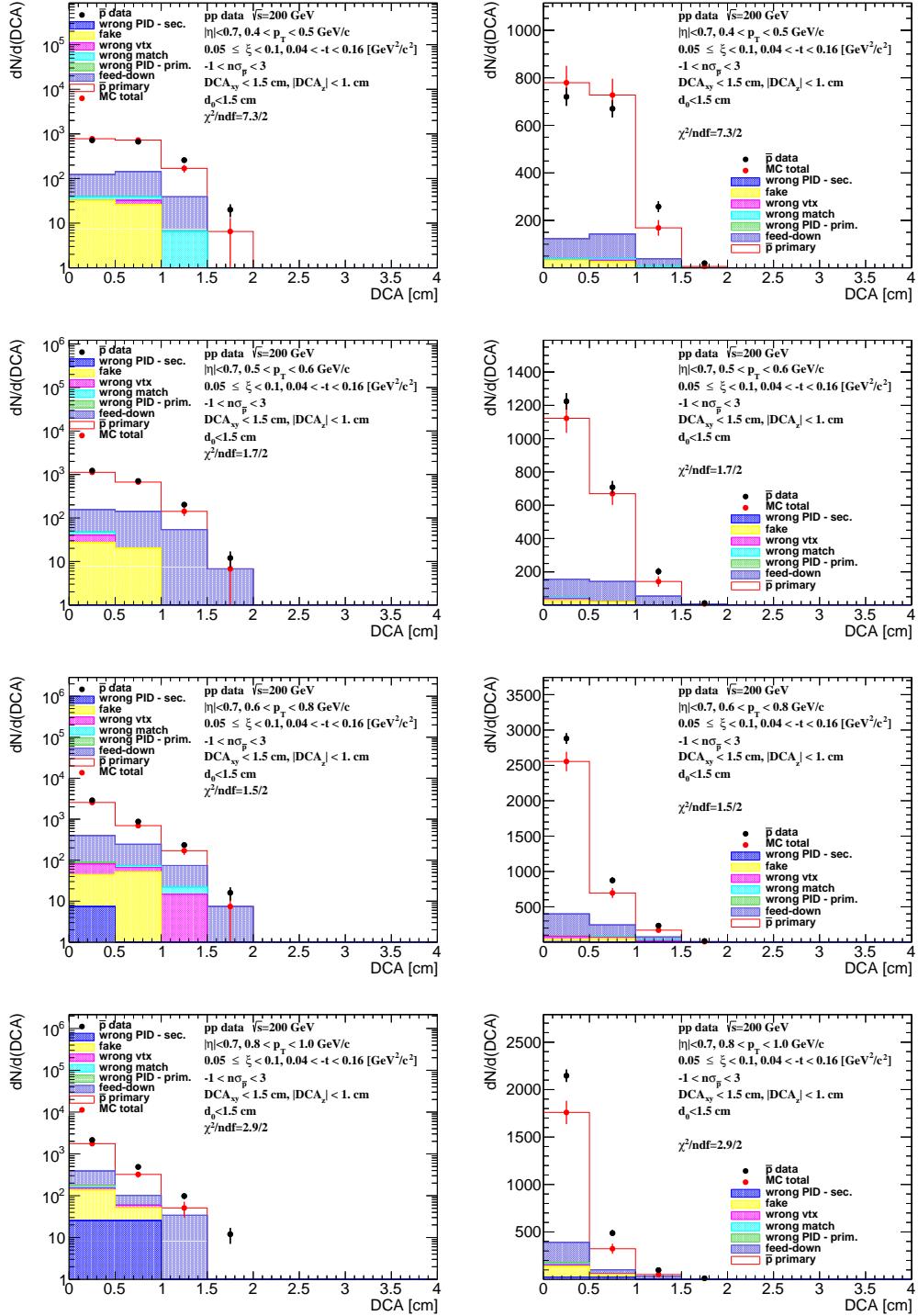


Figure B.10: Distributions of DCA for antiprotons in SD interactions with  $0.05 < \xi < 0.1$  and normal selection.

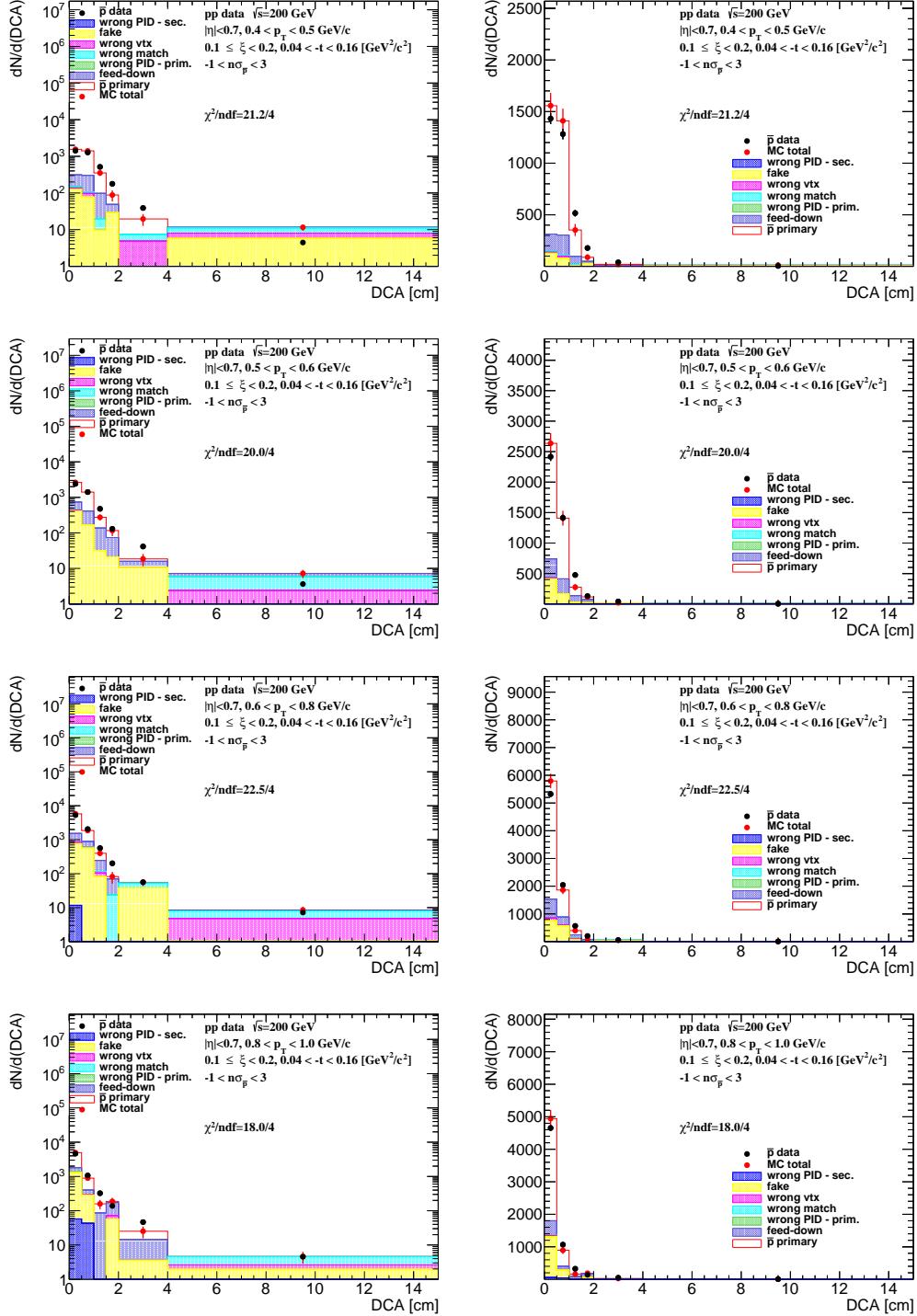


Figure B.11: Distributions of DCA for antiprotons in SD interactions with  $0.1 < \xi < 0.2$  and loose selection.

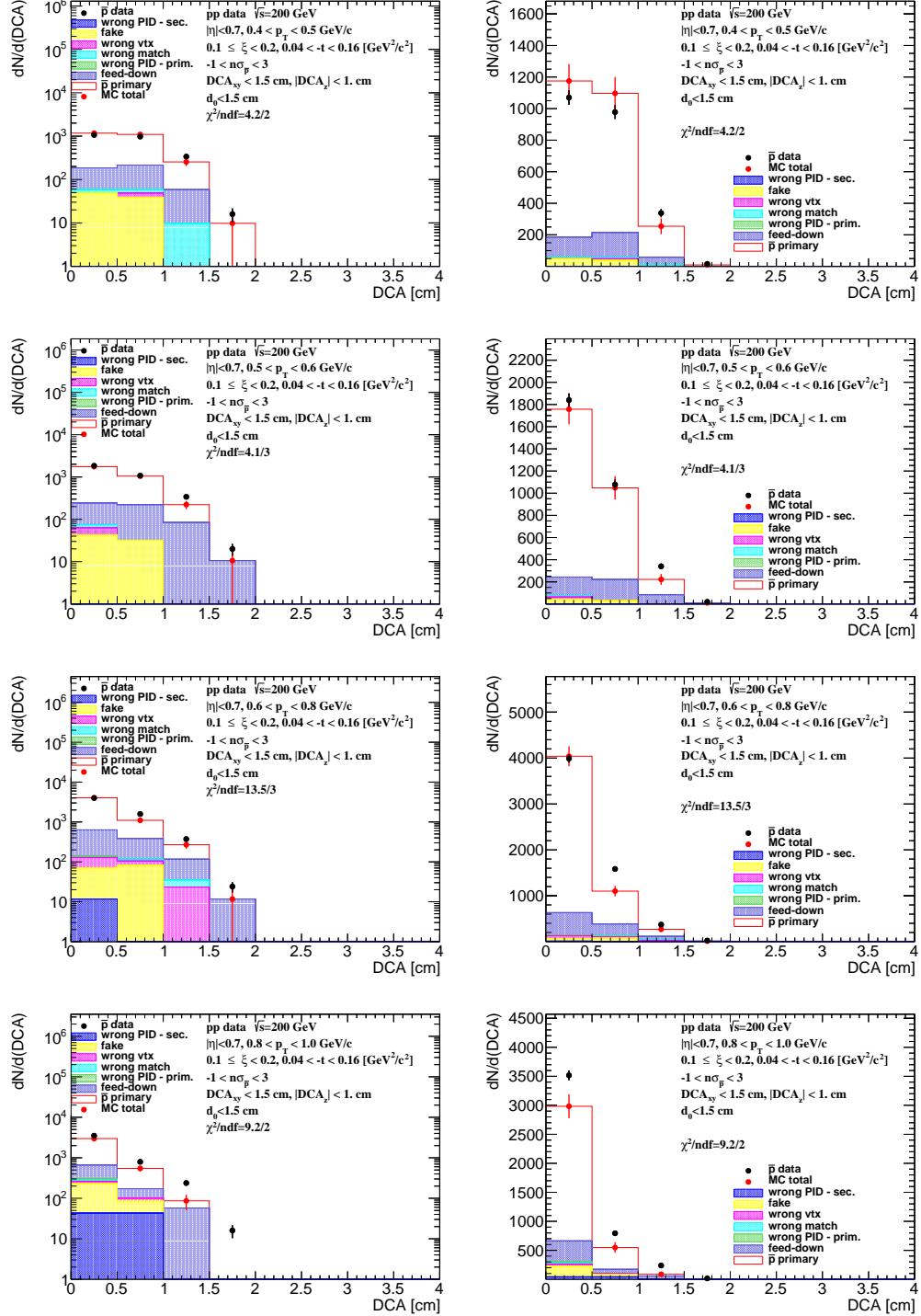


Figure B.12: Distributions of DCA for antiprotons in SD interactions with  $0.1 < \xi < 0.2$  and normal selection.

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

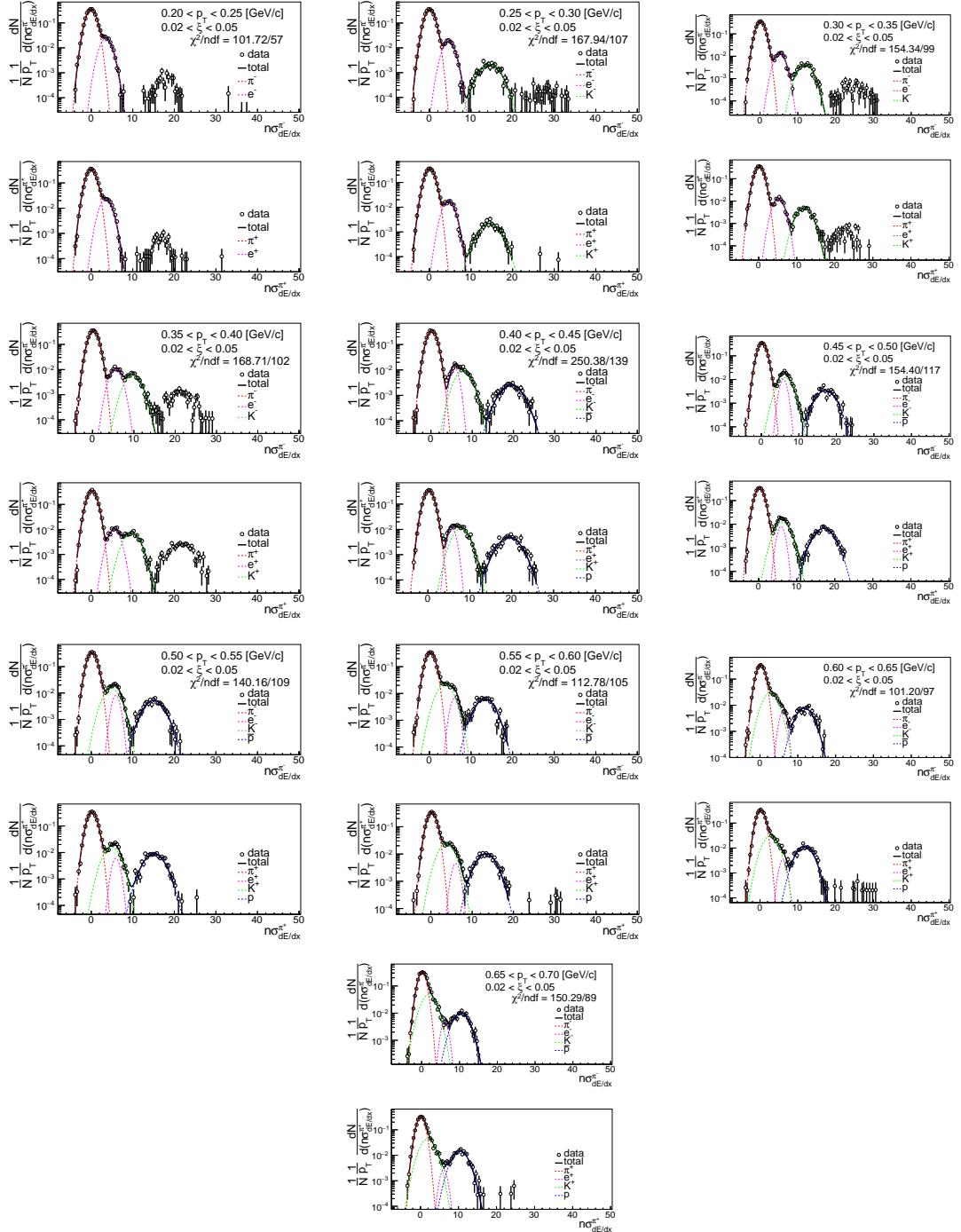


Figure C.1: Distributions of  $n\sigma_{dE/dx}^{\pi^\pm}$  for  $\pi^\pm$  in SD interactions with  $0.02 < \xi < 0.05$ .

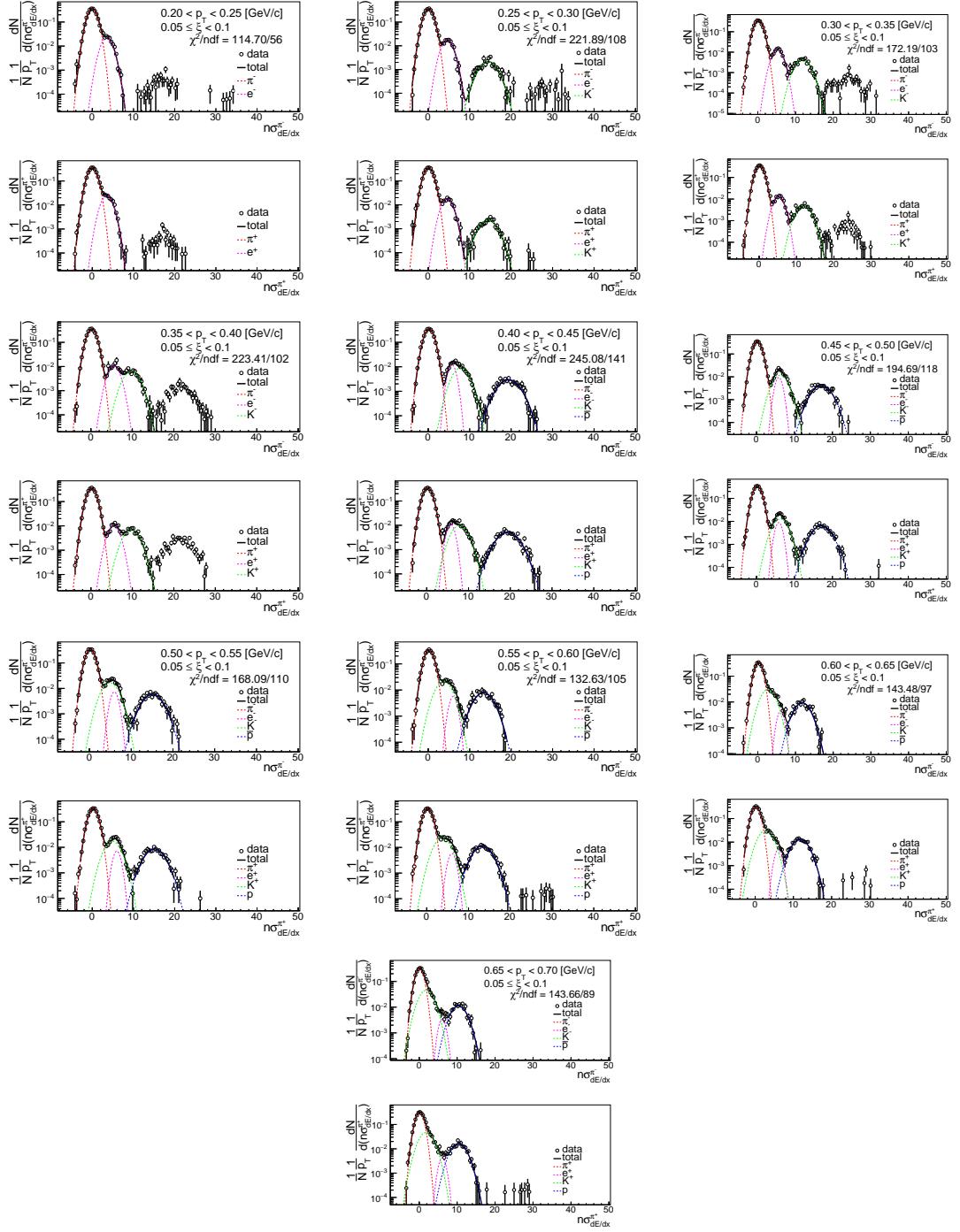


Figure C.2: Distributions of  $n\sigma_{dE/dx}^{\pi^\pm}$  for  $\pi^\pm$  in SD interactions with  $0.05 < \xi < 0.1$ .

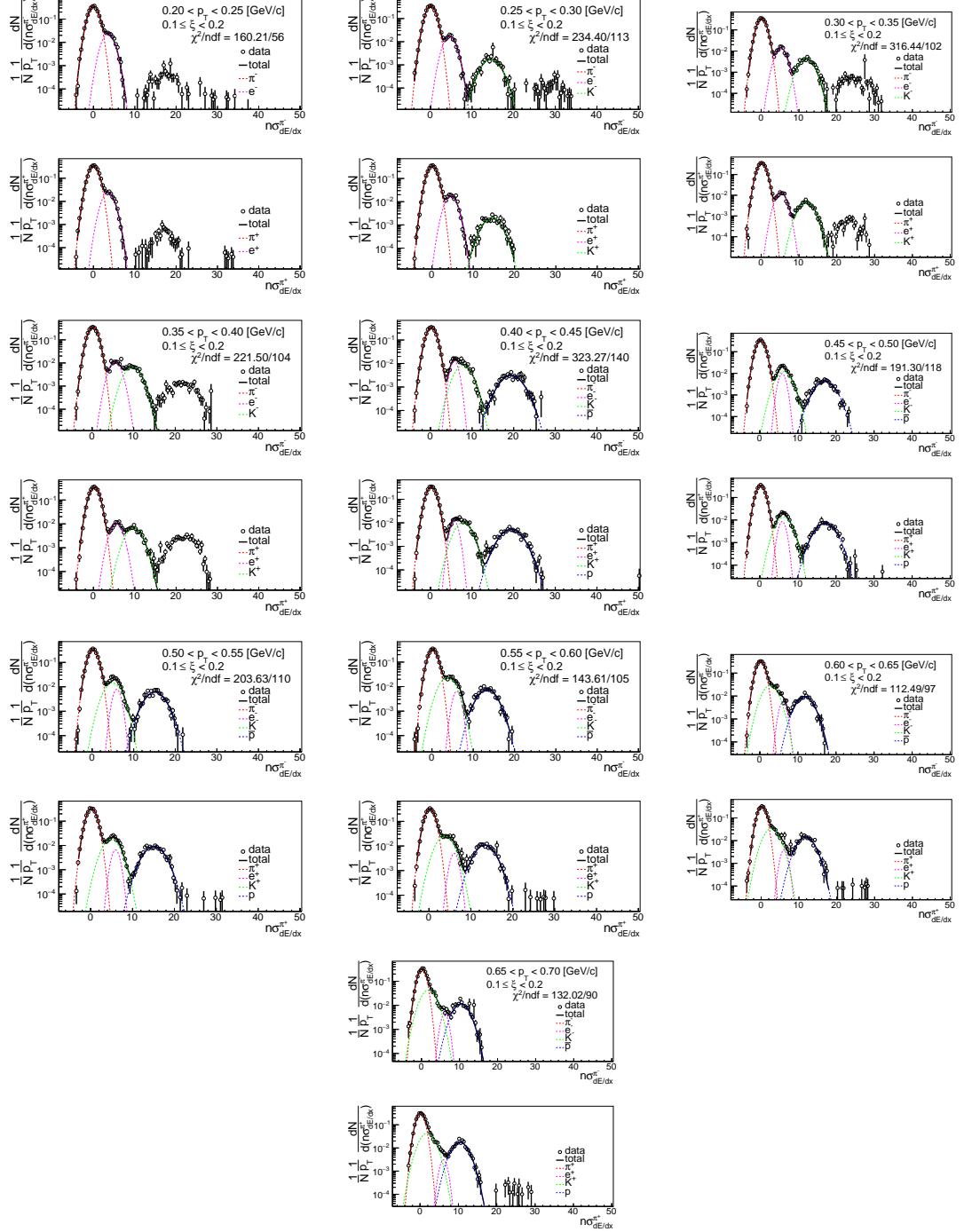


Figure C.3: Distributions of  $n\sigma_{dE/dx}^{\pi^\pm}$  for  $\pi^\pm$  in SD interactions with  $0.1 < \xi < 0.2$ .

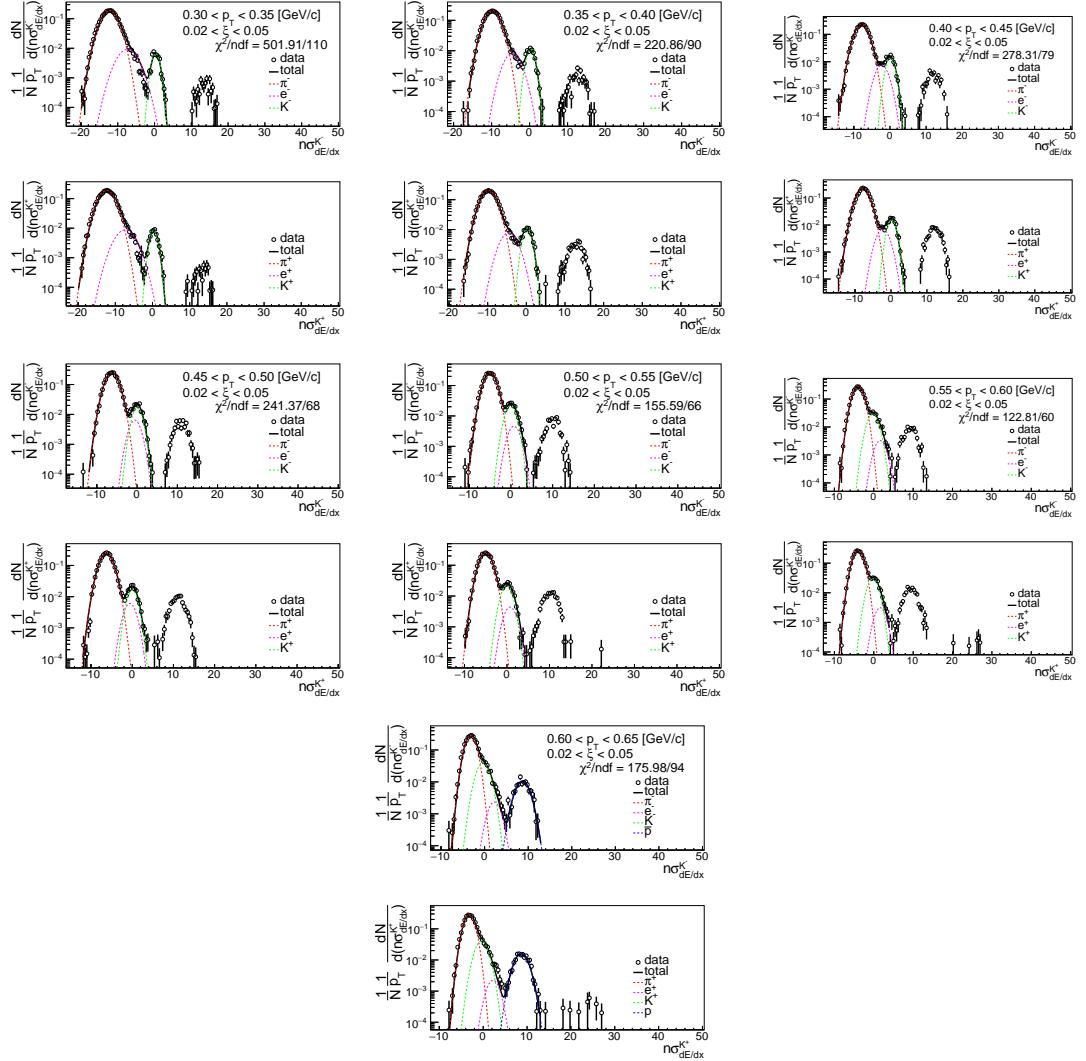


Figure C.4: Distributions of  $n\sigma_{dE/dx}^{K^\pm}$  for  $K^\pm$  in SD interactions with  $0.02 < \xi < 0.05$ .

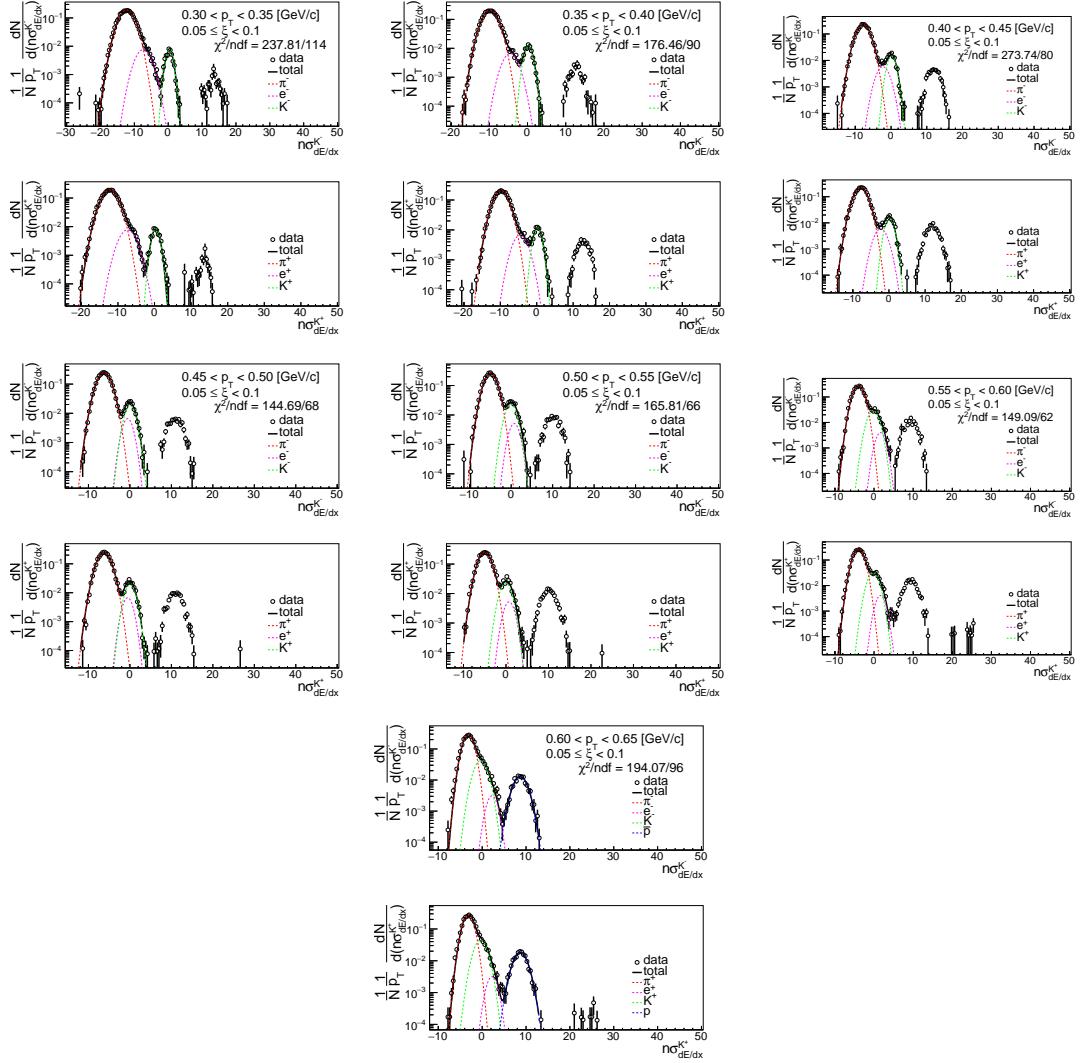


Figure C.5: Distributions of  $n\sigma_{dE/dx}^{K^\pm}$  for  $K^\pm$  in SD interactions with  $0.05 < \xi < 0.1$ .

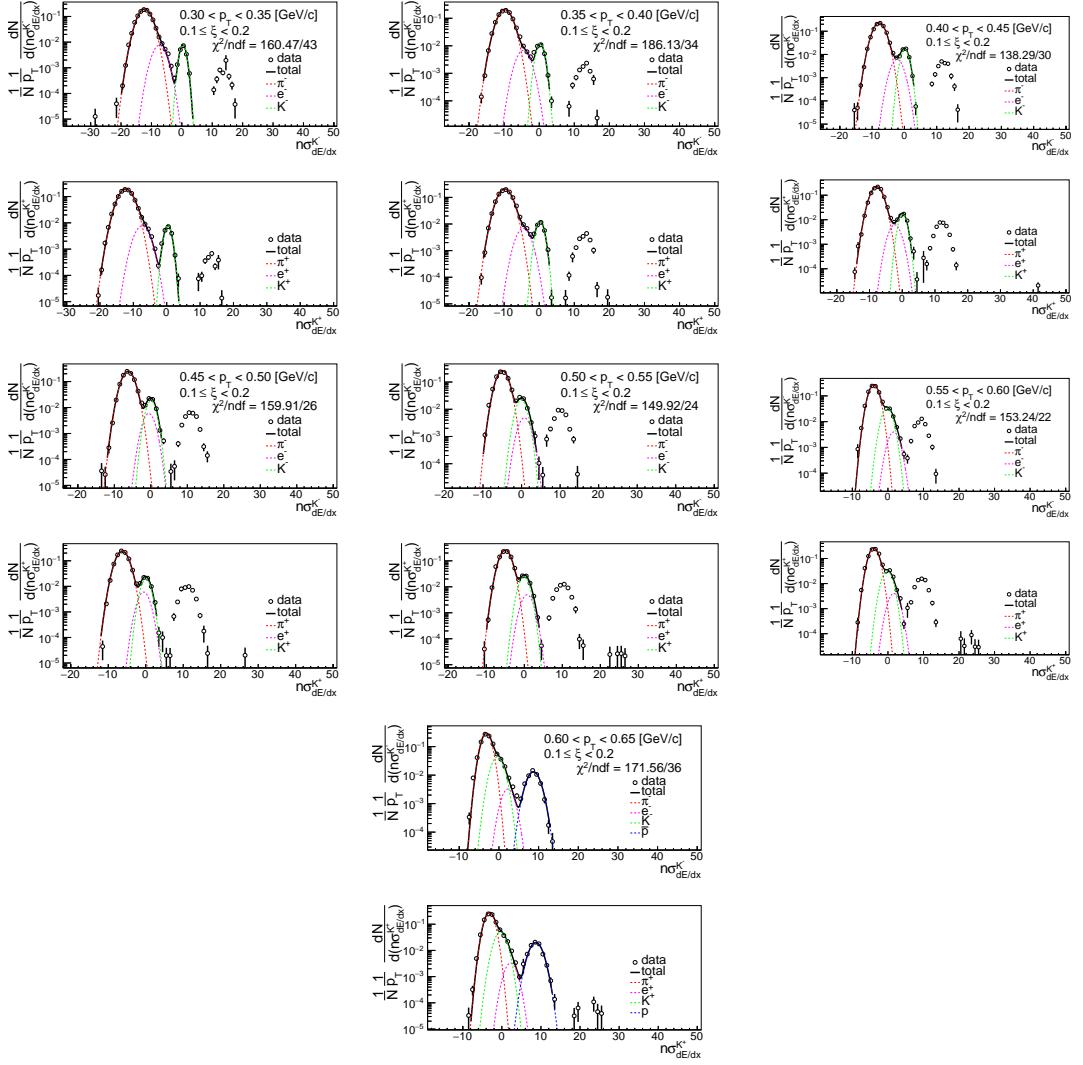


Figure C.6: Distributions of  $n\sigma_{dE/dx}^{K^\pm}$  for  $K^\pm$  in SD interactions with  $0.1 < \xi < 0.2$ .

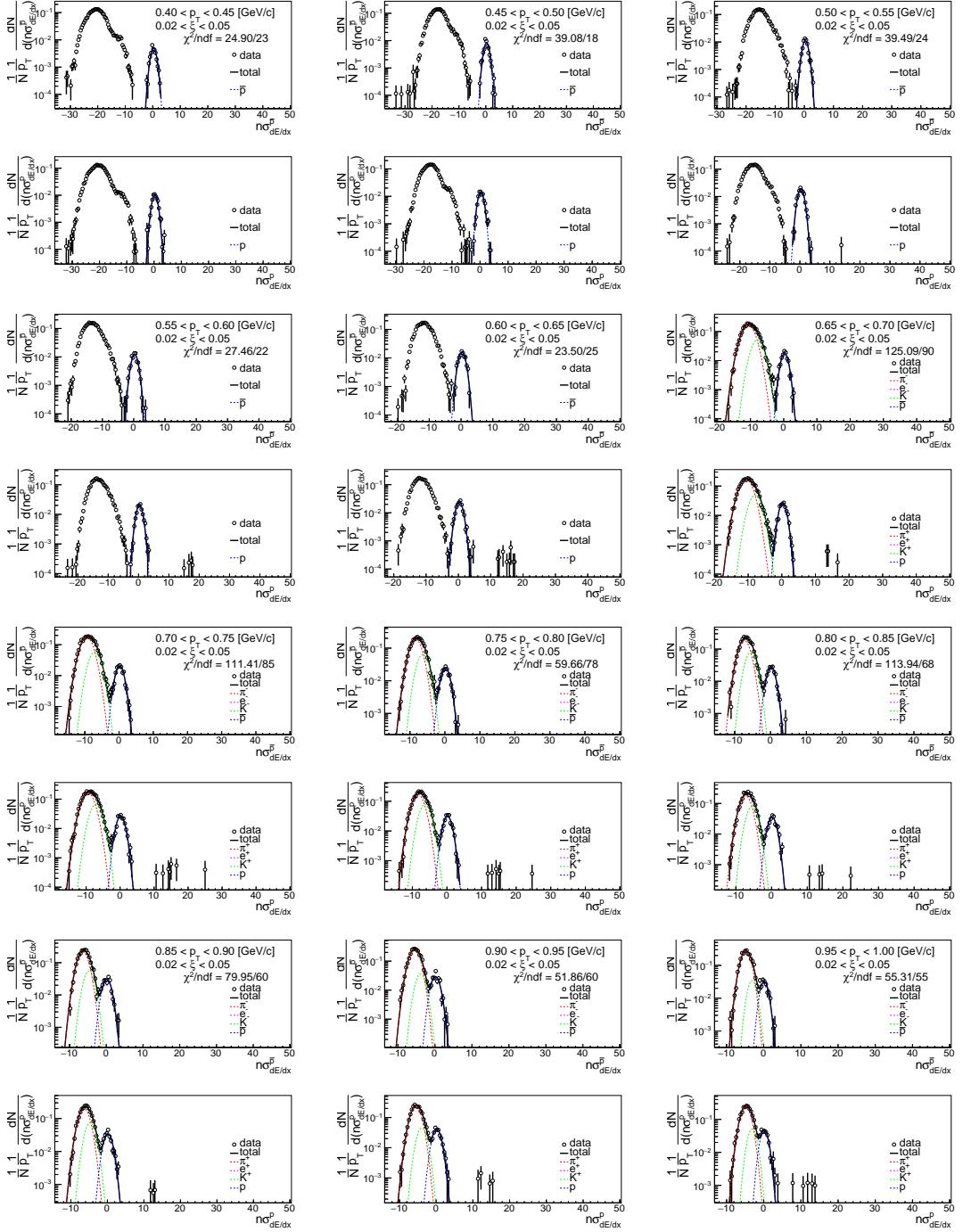


Figure C.7: Distributions of  $n\sigma_{dE/dx}^{\bar{p}, p}$  for  $\bar{p}, p$  in SD interactions with  $0.02 < \xi < 0.05$ .

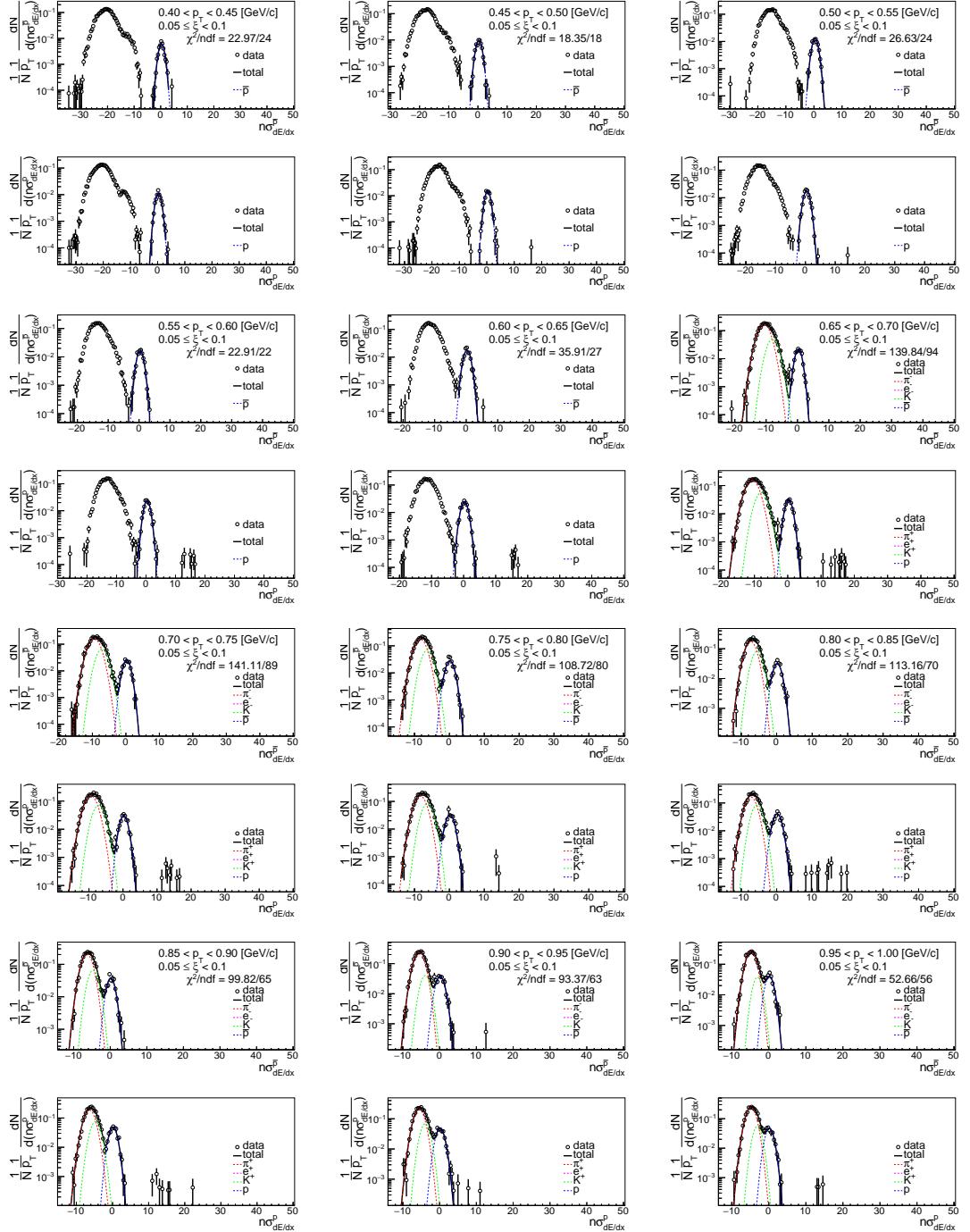


Figure C.8: Distributions of  $n\sigma_{dE/dx}^{\bar{p},p}$  for  $\bar{p}, p$  in SD interactions with  $0.05 < \xi < 0.1$ .

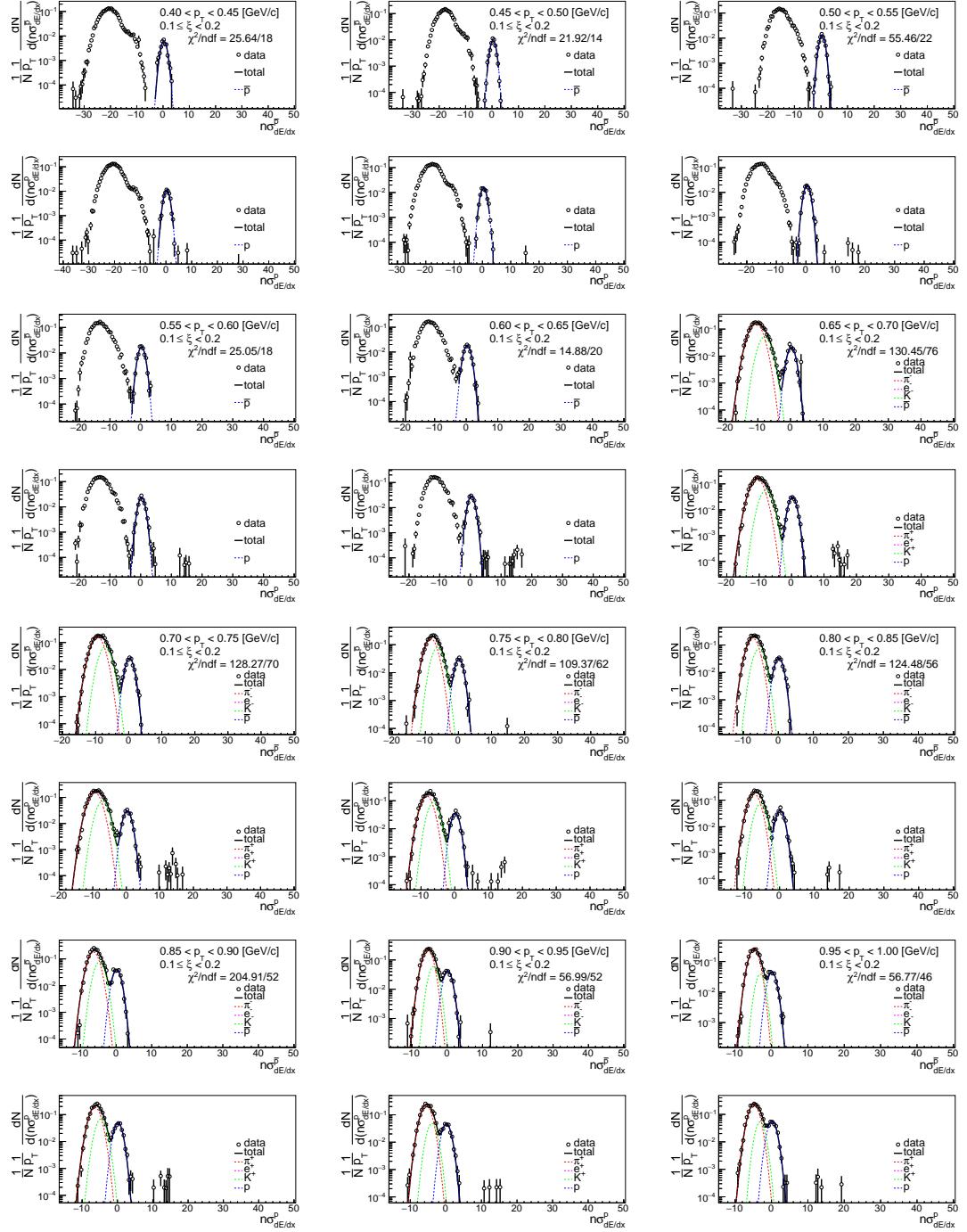


Figure C.9: Distributions of  $n\sigma_{dE/dx}^{\bar{p},p}$  for  $\bar{p}, p$  in SD interactions with  $0.1 < \xi < 0.2$ .