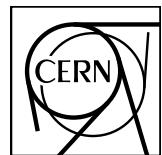


# EUROPEAN ORGANIZATION FOR NUCLEAR RESEARCH

1



v2.0

October 25, 2019

2

3

## Measurement of isolated photon–hadron correlations in 5 TeV pp and p–Pb data

4

5 Alwina Liu, Barbara Jacak, Dhruv Dixit, Fernando Torales - Acosta, Peter Jacobs, Miguel Arratia, Yue Shi Lai.

6

7 University of California, Berkeley, Lawrence Berkeley National Laboratory

8

Email: marratia@cern.ch

9

### Abstract

10

11 We present an analysis of isolated photon–hadron correlations using data collected in 2013 during the  
12  $\sqrt{s_{\text{NN}}} = 5 \text{ TeV}$  p–Pb run and in 2017 during the 5 TeV pp run. We use a combination of isolation and  
13 shower-shape variables to reduce the background from neutral-meson decays. We measure the purity  
14 of our isolated-photon selection by using a template fit technique with a data-driven background  
15 estimate. We perform a measurement of per-trigger associated hadron yields for photons with  $|\eta| <$   
16  $0.67$  and  $12 < p_T < 40 \text{ GeV}/c$  and associated charged particles with  $|\eta| < 0.80$  and  $0.5 < p_T <$   
17  $10 \text{ GeV}/c$ . We do not observe a significant difference between pp and p–Pb data. We also found that  
PYTHIA8.2 describes the data within uncertainties.



**18    Contents**

19	<b>1 Motivation</b>	2
20	<b>2 Analysis summary</b>	3
21	<b>3 Experimental Setup</b>	4
22	<b>4 Datasets</b>	5
23	<b>5 Monte Carlo simulations</b>	6
24	<b>6 Event selection</b>	7
25	<b>7 Calorimeter cluster reconstruction</b>	7
26	7.1 Definition . . . . .	7
27	7.2 Corrections . . . . .	7
28	7.3 Selection . . . . .	8
29	<b>8 Photon Identification</b>	10
30	8.1 Shower-shape variable . . . . .	10
31	<b>9 Isolation</b>	10
32	9.1 Underlying Event estimation . . . . .	11
33	9.2 UE correction to isolation variable . . . . .	12
34	<b>10 Purity Measurement</b>	13
35	10.1 Template fit method . . . . .	14
36	10.1.1 Signal template and background templates . . . . .	14
37	10.1.2 Fit results . . . . .	15
38	10.2 Systematic uncertainties of the purity measurement . . . . .	19
39	10.2.1 Signal template . . . . .	19
40	10.2.2 Sideband variation in the background template . . . . .	21
41	10.2.3 Background template correction . . . . .	21
42	10.3 Summary of systematic uncertainties of purity measurement . . . . .	25
43	<b>11 Tracking</b>	25
44	11.1 Tracking performance with and without TPC . . . . .	26
45	11.1.1 Efficiency and fake rates . . . . .	27

---

46	11.1.2 Resolution, response matrix and bin migration . . . . .	28
47	11.2 Comparison with published data . . . . .	30
48	11.2.1 Angular dependence of tracking efficiency . . . . .	33
49	11.2.2 Validation of $\varphi$ -dependence of efficiency . . . . .	33
50	11.3 Summary of the ITS-only tracking performance studies . . . . .	34
51	<b>12 Isolated-photon hadron correlations</b>	<b>35</b>
52	12.1 $\gamma^{\text{decay}}$ background Subtraction . . . . .	35
53	12.2 Pair-acceptance correction with event mixing . . . . .	36
54	12.3 Corrections applied to data . . . . .	39
55	12.3.1 Weighting for photon purity . . . . .	39
56	12.3.2 Weighting for track efficiency, fake rate, and bin migration . . . . .	39
57	12.4 Corrected correlations . . . . .	40
58	12.5 Underlying event subtraction . . . . .	40
59	12.5.1 Study of neutral energy in isolation variable using PYTHIA8 . . . . .	40
60	12.5.2 Estimate of impact of acceptance difference between pp and p–Pb due to boost .	44
61	<b>13 Systematic uncertainties</b>	<b>48</b>
62	<b>14 Results</b>	<b>50</b>
63	14.1 Azimuthal correlations in pp and p–Pb data . . . . .	50
64	14.2 Fragmentation functions in pp & p–Pb data . . . . .	50
65	<b>15 Conclusions</b>	<b>53</b>
66	<b>Appendices</b>	<b>53</b>

## 67 **1 Motivation**

- 68 The photon-tagged correlation of jets and jet fragments is a promising channel for the study of partonic  
 69 energy loss in heavy-ion collisions. Energetic photons are free from the uncertainties that are associated  
 70 with the fragmentation of partons into hadrons. However, existing measurements using high  $E_T$  photons  
 71 focus on the study of energy loss beyond the region where the largest modification of particle spectra has  
 72 been observed.
- 73 In this note, we present an analysis using pp and p–Pb data with the aim of bench-marking similar studies  
 74 in Pb–Pb collisions. The comparison between p–Pb and Pb–Pb data disentangles effects due to the quark–  
 75 gluon plasma (e.g. parton energy loss) and “cold–nuclear matter” effects such as modification of parton  
 76 distribution functions in nuclei, and elastic, inelastic and coherent multiple parton scattering processes  
 77 inside a large nucleus. This is because final-state effects associated with the quark-gluon plasma are  
 78 expected to be absent or suppressed in p–Pb collisions.

79 **2 Analysis summary**

80 We use data collected during the  $\sqrt{s_{\text{NN}}} = 5$  TeV p–Pb run in 2013 and during the  $\sqrt{s} = 5$  TeV pp run  
 81 in 2017. We use the EMCal trigger to select events with a high-momentum calorimeter cluster. For this  
 82 analysis, we target photons with  $p_T$  in the 12–40 GeV/ $c$  range.

83 In this analysis, our signal are “prompt” photons, which include “direct photons” and “fragmentation  
 84 photons”. At leading order in perturbative QCD, the direct photons are produced in hard scattering  
 85 processes such as quark-gluon Compton scattering ( $qg \rightarrow q\gamma$ ) or quark-antiquark annihilation ( $q\bar{q} \rightarrow g\gamma$ ),  
 86 whereas the fragmentation photons are the product of the collinear fragmentation of a parton ( $q\bar{q}(gg) \rightarrow$   
 87  $\gamma + X$ ). At LHC energies, Compton scattering and gluon fusion ( $gg \rightarrow q\bar{q}\gamma$ ) dominate due to the high-  
 88 gluon density in the proton at small values of Bjorken- $x$ .

89 Beyond the simplistic leading order picture, the direct and fragmentation components have no physical  
 90 meaning and cannot be factorized; the sum of their cross sections is the physical observable. For example,  
 91 the separation between the NLO direct photons and LO fragmentation is arbitrary. However, it is still  
 92 possible to simplify comparisons with theoretical calculations by applying an isolation criteria. We use  
 93 an isolation variable that is the sum of the transverse momentum of the charged particles that are inside  
 94 an angular cone of radius  $R = \sqrt{(\Delta\phi)^2 + (\Delta\eta)^2} = 0.4$  around the photon direction.

95 The main background for our analysis are photons from meson decays, which we will call “decay pho-  
 96 tons” or  $\gamma^{\text{decay}}$ . The challenge that we face in this measurement arises mainly from the small cross-section  
 97 of the signal compared to that of the decay photon background (about 1% at 10 GeV/ $c$  increasing to about  
 98 4% at 30 GeV/ $c$ , according to next-to-leading order calculations [5]).

99 This measurement exploits the difference between the electromagnetic shower profiles of prompt photons  
 100 and of photon pairs from neutral-meson decays. We call the clusters that pass our isolation and shower  
 101 shape selections isolated  $\gamma$  candidates or “ $\gamma^{\text{iso}}$  candidates”.

102 The main background in the  $\gamma^{\text{iso}}$  candidate sample arises from multi-jet events where one jet typically  
 103 contains a  $\pi^0$  or  $\eta$  that carries most of the jet energy and is misidentified as a photon because it decays  
 104 into a photon pair that is collinear with respect to the EMCal cell granularity ( $\Delta\eta \times \Delta\phi \approx 14.3 \times 14.3$   
 105 mrad<sup>2</sup>), that is, the two photons are close enough to deposit most of their energy in the same cell.

106 We measure the signal purity of our  $\gamma^{\text{iso}}$  selection by using the “template-fit method”, in which the  
 107 measured shower-shape distribution is fit with the sum of signal and background templates with the  
 108 relative normalization as the single free parameter<sup>1</sup>. The background template is mostly data-driven,  
 109 calculated with an anti-isolated sideband requirement, but we apply a MC-based correction to account  
 110 for estimated biases. The signal template is obtained from photon-jet simulation. The purity of our  $\gamma^{\text{iso}}$   
 111 selection is measured to be around 20% at 12 GeV/ $c$  and increases to about 55% at 20 GeV/ $c$  and above.

112 We measure the angular correlation of our  $\gamma^{\text{iso}}$  candidates with charged particles. We correct for geo-  
 113 metrical acceptance effects by using the mixed-event technique and then subtract the uncorrelated back-  
 114 ground, estimated by the zero-yield-at-minimum (ZYAM) method and by using a control region at large  
 115  $|\eta_{\text{hadron}} - \eta_{\gamma}|$ . We measure the  $\gamma^{\text{decay}}$ –hadron correlation function by inverting the shower-shape cut to  
 116 select merged-clusters from meson decays. We normalize this correlation with the measured purity and  
 117 subtract the normalized  $\gamma^{\text{decay}}$ –hadron correlation background from the main  $\gamma^{\text{iso}}$  candidate correlations.  
 118 Finally, we integrate the away-side of the resulting correlation function to determine the number of cor-  
 119 related hadrons per  $\gamma^{\text{iso}}$ , i.e. to measure the conditional yield of hadrons. We perform this analysis with  
 120 photons with  $12 < p_T < 40$  GeV/ $c$ , and in intervals of charged particle  $p_T$  and  $z_T \equiv p_{T^h}/p_T^{\gamma^{\text{iso}}}$ .

121 One of the novel aspects of this analysis is the use of ITS standalone tracking. We developed this

---

<sup>1</sup>Note that this is a standard way to estimate QCD background since at least the Tevatron days. The same exact method is used in the CMS  $\gamma^{\text{iso}}$  and  $\gamma^{\text{iso}}$ –jet measurements in pp and PbPb data, for example in Refs. [16, 14].

122 approach to bypass the serious space-charge distortions that compromised the TPC during the high-  
 123 luminosity p–Pb data taking in 2013<sup>2</sup>. Furthermore, the ITS-only tracking allowed the 2017 pp run to  
 124 operate in the CALO mode that yielded a much larger sample than would have been possible otherwise.

125 We have validated the performance of the ITS-standalone tracking (fake rate, efficiency and momentum  
 126 smearing) by measuring the charged particle spectrum and comparing it with published ALICE mea-  
 127 surements at the same center-of-mass energy. Our studies, included below, show agreement between the  
 128 ITS-standalone measurement and the published data to within  $\approx \pm 5\%$  of the corresponding published  
 129 data for the range  $0.5 < p_T < 10 \text{ GeV}/c$ , which is the relevant range in this analysis.

130 One of the main considerations of our analysis strategy was to minimize the use of Monte Carlo simula-  
 131 tions. By using an isolation variable constructed using only charged particles, we reduce the correlations  
 132 between isolation and shower-shape variables due to the opening angle of neutral-meson decays, at the  
 133 expense of a slightly lower purity. In our template fit analysis, we perform checks that are independent  
 134 of any input from simulations, suggesting that we are not sensitive to the detailed simulation of the shape  
 135 of the shower-shape distributions. Moreover, our analysis measures per-trigger quantities such that we  
 136 do not need to correct for efficiency of the  $\gamma^{\text{iso}}$  selection.

137 While we made the effort to collect and use the largest data samples available by pioneering high-rate  
 138 data taking with ITS+EMCal, our measurement turns out to be dominated by statistical uncertainties.  
 139 Faced with this reality, we only make efforts to reduce the systematic uncertainties of the measurement  
 140 such that they are smaller than the statistical uncertainty.

### 141 3 Experimental Setup

142 A comprehensive description of the ALICE experiment and its performance is provided in Ref [4, 9].  
 143 The detector elements most relevant for this study are the electromagnetic calorimeter system, which is  
 144 used to measure and trigger on high  $p_T$  photons, and the inner tracking system, which is used for tracking  
 145 and vertexing. Both are located inside a large solenoid magnet with a field strength of 0.5 T. These are  
 146 briefly described here:

147 The Electromagnetic Calorimeter (EMCal) is a sampling calorimeter composed of 77 alternating layers  
 148 of 1.4 mm lead and 1.7 mm polystyrene scintillators. It has a cellular structure with square cells with a  
 149 transverse size of  $6 \times 6 \text{ cm}^2$  called towers. The towers are arranged in a quasi-projective geometry. The  
 150 tower transverse size is roughly equal to the Molière radius, so that most of the energy of the particle  
 151 is deposited in one cell, about 90% for photons. It is located at 428 cm from the interaction point  
 152 and its cell granularity is  $\Delta\eta \times \Delta\phi = 14.3 \times 14.3 \text{ mrad}^2$ . It has an energy resolution is parametrized as  
 153  $\sigma_E/E = 4.8\%/E \otimes 11.3\%/\sqrt{E} \otimes 1.7\%$  where the energy  $E$  is given in units of GeV [11]. The linearity  
 154 of the response of the detector and electronics has been measured with electron test beams to a precision  
 155 better than 3% for the momentum range probed in this analysis. The non-linearity is negligible for  
 156 cluster energy between 3 and 50  $\text{GeV}/c$ , which is the relevant range for this analysis. The geometrical  
 157 acceptance of the EMCal is  $|\eta| < 0.70$  and  $80^\circ < \phi < 187^\circ$ .

158 The Di-jet Calorimeter (DCal) is back-to-back in azimuth with respect to the EMCal. The DCal uses  
 159 the same technology and material as the EMCal, thus having identical granularity and intrinsic energy  
 160 resolution. It covers  $0.22 < \eta < 0.7$ ,  $260^\circ < \phi < 320^\circ$  and  $|\eta| < 0.7$ ,  $320^\circ < \phi < 327^\circ$ . It was installed  
 161 and commissioned during the LHC long shutdown in 2015, and thus was operational during the 2017 pp  
 162 run but not during the 2013 p–Pb run.

163 The inner tracking system (ITS) consists of six layers of silicon detectors and is located directly around  
 164 the interaction point. The two innermost layers consist of silicon pixel detectors positioned at radial

---

<sup>2</sup>For details, see <https://alice.its.cern.ch/jira/browse/ATO-351>, <https://alice.its.cern.ch/jira/browse/PWGPP-349>, <https://alice.its.cern.ch/jira/browse/PWGPP-314>

distances of 3.9 cm and 7.6 cm, followed by two layers of silicon drift detectors at 15.0 cm and 23.9 cm, and two layers of silicon strip detectors at 38.0 cm and 43.0 cm. The ITS covers  $|\eta| < 0.9$  and has full azimuthal coverage.

We also use the forward scintillators to provide the minimum-bias trigger and to estimate the particle multiplicity in each event. The V0 system consists of two scintillator arrays located on opposite sides of the interaction point at  $z = -340$  cm and  $z = +90$  cm covering  $2.8 < \eta < 5.1$  and  $-3.7 < \eta < -1.7$  respectively. It measures the total charge of the particles produced and the time of their arrival in each of the 64 channels.

## 4 Datasets

The datasets used in this analysis are shown in Table 1. We use the high-luminosity runs of the 2013 p–Pb run (13d,e,f) and the 2017 pp run (17q) that were collected with EMCal triggers, which are listed in Table 2.

**Table 1:** Datasets used in this analysis. The runs listed in the table corresponds to those that are in the good run list appropriate for analysis using the EMCal and ITS detectors.

Name	Config.	Run Number list	Pass	Integrated Luminosity
13b	p–Pb	195344, 195351, 195389,	pass4	$\sim 1.0 \text{ nb}^{-1}$
13d	p–Pb	195872, 195871, 195867, 195831, 195829, 195787, 195783, 195767, 195760, 195724.	pass4	$\sim 1.0 \text{ nb}^{-1}$
13e	p–Pb	196310, 196309, 196308, 196214, 196208, 196201, 196200, 196199, 196197, 196194, 196187, 196185, 196107, 196091, 196090, 196089, 196085, 195958, 195955, 195935.	pass4	$\sim 1.3 \text{ nb}^{-1}$
13f	Pb–p	197342, 197341, 197302, 197300, 197299, 197298, 197297, 197296, 197260, 197258, 197256, 197255, 197254, 197248, 197247, 197189, 197153, 197152, 197138, 197092, 197091, 197027, 197015, 197012, 197011, 197003, 196974, 196973, 196972, 196967, 196965, 196721, 196720, 196714, 196706, 196703, 196702, 196701, 196648, 196646, 196608, 196535, 196528.	pass4	$\sim 2.3 \text{ nb}^{-1}$
13f_new	Pb–p	196433, 196474, 196475, 196477, 196722. 196772, 196773, 196774, 196869, 196870, 196874, 196876, 197139, 197142, 197143, 197144, 197145, 197147, 197148, 197149, 197150, 197348, 197349, 197351, 197386 197387, 197388.	pass4	$\sim 2.3 \text{ nb}^{-1}$
17q	pp	282441, 282440, 282439, 282437, 282415, 282411, 282402, 282399, 282398, 282393, 282392, 282391, 282367, 282366, 282365	pass1_wSDD	$\sim 300 \text{ nb}^{-1}$

**Table 2:** EMCal triggers used in this analysis.

Dataset	Trigger Strings
p–Pb	CEMC7EG1-B-NOPF-CENTNOTRD, CEMC7EG2-B-NOPF-CENTNOTRD,
pp	CEMC7EG2-B-NOPF-CALO, CDMC7DG2-B-NOPF-CALO, CEMC7EG2-B-NOPF-CENT, CDMC7DG2-B-NOPF-CENT

177 The EMCal gamma triggers (EG1, EG2, DG1, DG2) are based on the summed energy in  $2 \times 2$  adjacent  
 178 tiles (a tile is composed of an EMCal module,  $2 \times 2$  adjacent cells). The trigger thresholds were 7 and 11  
 179 GeV/ $c$  during the 2013 p–Pb run and 5 GeV/ $c$  during the 2017 pp run.

180 Due to the 2-in-1 magnet design of the LHC, which requires the same magnetic rigidity for both colliding  
 181 beams, the beams had different energies during the p–Pb run ( $E_p = 4$  TeV,  $E_{Pb} = 4$  TeV  $\times Z$ , where  $Z = 82$   
 182 is the atomic number of lead). In the lead nucleus, the energy per nucleon was therefore 1.56 TeV  
 183  $= (Z/A) \times 4$  TeV, where  $A = 208$  is the nuclear mass number of the lead isotope used. This energy  
 184 asymmetry results in an average nucleon–nucleon center of mass collision energy of  $\sqrt{s_{NN}} = 5$  TeV and  
 185 a rapidity boost of this frame by  $\pm 0.465$  units relative to the ALICE rest frame in the direction of proton  
 186 beam. Around halfway through the 2013 p–Pb run, the beam directions were flipped, yielding similar  
 187 integrated luminosities in both beam configurations.

188 During the 2013 p–Pb run period, the TPC suffered from space-charge distortions<sup>3</sup> that affect tracking,  
 189 leading to a very drastic drop in efficiency for tracks with  $p_T > 4$  GeV/ $c$ . We bypass this issue by using  
 190 ITS-only tracking as detailed in Section 11. For the 2017 pp data, the TPC was also inactive due to the  
 191 high luminosity of the runs considered in this analysis. We use the 17q period, during which all six layers  
 192 of the ITS were active.

193 The average number of inelastic collisions per bunch crossing,  $\mu$ , is 0.020–0.060 for the 2013 p–Pb data  
 194 set and in the range 0.015–0.045 for the 2017 pp dataset<sup>4</sup>.

## 195 5 Monte Carlo simulations

196 We use Monte Carlo (MC) simulations to obtain the signal shower-shape distributions for the template  
 197 fits (section 10) and to study tracking performance (section 11).

198 The simulations of hard processes are based on the PYTHIA event generator. In PYTHIA, the signal  
 199 events are included via  $2 \rightarrow 2$  matrix elements with  $gq \rightarrow \gamma q$  and  $q\bar{q} \rightarrow \gamma g$  hard scatterings, defined  
 200 at the leading order, followed by the leading-logarithm approximation of the partonic shower. The soft  
 201 underlying events in pp collisions as well as fragmentation are included with the default PYTHIA models.

202 For the simulation of p–Pb events, the pp samples are embedded into p–Pb inelastic events generated  
 203 with DPMJET. The boost of  $\Delta y = +0.465$  in the direction of the proton beam is reproduced.

204 Table 3 shows the MC simulations used in this analysis. Each sample is simulated with the detector  
 205 configuration appropriate for the runs used in this analysis.

**Table 3:** Monte Carlo simulations used in this analysis.

Name	Configuration	JIRA ticket link
17g6a1	p–Pb, 5 TeV, PYTHIA8 Gamma-Jet +DPMJET anchored to 13d,e,f	ALIROOT-7271
17g6a3	p–Pb, 5 TeV, PYTHIA8 Jet-Jet +DPMJET anchored to 13def	ALIROOT-7271
13b2	p–Pb, 5 TeV, DPMJET anchored to LHC13b,c	39374
18b10a(b)_calo	pp 5 TeV, PYTHIA8 Gamma-Jet anchored to 17p/q	ALIROOT-7692
18l2a(b)	pp 5 TeV, PYTHIA8 Jet-Jet anchored to 17p/q	ALIROOT-8144

<sup>3</sup>For more information on the problems with space-charge distortions due to high-luminosity in p–Pb run, see: <https://alice.its.cern.ch/jira/browse/PWGPP-314>.

<sup>4</sup>This information can be found in [http://aliquevs.web.cern.ch/aliquevs/data/2013/LHC13d/pass4/global\\_properties.pdf](http://aliquevs.web.cern.ch/aliquevs/data/2013/LHC13d/pass4/global_properties.pdf) and [http://aliquevs.web.cern.ch/aliquevs/data/2017/LHC17q/cpass1\\_pass1/global\\_properties.pdf](http://aliquevs.web.cern.ch/aliquevs/data/2017/LHC17q/cpass1_pass1/global_properties.pdf)

206 **6 Event selection**

207 We use the following event selection criteria to ensure good event quality and uniform acceptance:

- 208 – Run passes QA for EMCal and ITS (the selected runs are listed in Table 1).
- 209 – At least one EMCal cluster with  $p_T > 12 \text{ GeV}/c$ .
- 210 – Selected at least one of the EMCal triggers (logical OR of the trigger strings listed in Table 2).
- 211 – Valid vertex ( $|z| \neq 0.0$ ) and  $|z| < 10 \text{ cm}$

212 The number of events that pass our selection in each sample is shown in Table 4. We report the events  
 213 selected for each trigger separately, as well as the logical OR combination. In p–Pb events, the number  
 214 of events is dominated by the EG1 trigger (11  $\text{GeV}/c$  threshold), and by the EG2 trigger (5  $\text{GeV}/c$ ) in  
 215 pp collisions.

**Table 4:** Number of events that passed our full event selection for each of data taking period used in this analysis. The numbers are also shown separately for EG1 (DG1) and EG2 (DG2) triggers.

Dataset	$N^{\text{EG1}  \text{EG2}}$	$N^{\text{EG1}}$	$N^{\text{EG2}}$
13d	134024	133326	12528
13e	198108	196745	22409
13f	340607	338198	38353
13f_new	241870	240074	30310
Dataset	$N^{\text{EG2}  \text{DG2}}$	$N^{\text{EG2}}$	$N^{\text{DG2}}$
17q	406934	301086	119498

216 **7 Calorimeter cluster reconstruction**

217 **7.1 Definition**

218 EMCal clusters are formed by a clustering algorithm that combines signals from adjacent towers. We  
 219 use calorimeter clusters defined with the “V1” algorithm. This algorithm starts from a “seed” cell, found  
 220 from a local-maximum scan, and adds “neighbor” cells to the cluster if they are above a given threshold.  
 221 The cluster definition is exclusive, i.e. once a cell is assigned to a cluster, it is not considered for other  
 222 clusters. The minimum energy for the seed and neighbor were set to 500 and 100 MeV respectively;  
 223 these values are several times larger than the standard deviation of the electronic noise<sup>5</sup>.

224 **7.2 Corrections**

225 We apply several corrections at the cell level, implemented within the “EMCal Correction Framework,”<sup>6</sup>  
 226 before the clustering algorithm is run over the data and simulations. The following corrections are  
 227 applied:

- 228 – “CELLENERGY”  
 229     This performs an energy calibration of cells, with coefficients obtained with  $\pi^0 \rightarrow \gamma\gamma$  mass measurements.
- 231 – “CELLBADCHANNEL”  
 232     This removes cells that declared hot or dead for a given run period.

<sup>5</sup>Some photon analysis use a 50 MeV threshold, but 100 MeV has been found to improve cell time measurements. The 100 MeV threshold has been used for example in Ref [10].

<sup>6</sup>[http://alidoc.cern.ch/AliPhysics/master/\\_r\\_e\\_a\\_d\\_m\\_eemc\\_corrections.html](http://alidoc.cern.ch/AliPhysics/master/_r_e_a_d_m_eemc_corrections.html)

233 – “CELLTIMECALIB”

234 This correction applies constant offsets, which are arbitrary, to the cell time measurements to  
235 minimize the spread among cells.

236 – “CELLEMULATECROSSTALK”.

237 This correction, described in detail in Ref [6], modifies the simulated cell energies to emulate the  
238 cell cross-talk that has been observed in data. This is applied to all the simulations described in  
239 Table 3.

240 **7.3 Selection**

241 The following selection is applied on the resulting clusters<sup>7</sup>:

242 – Cluster  $p_T$  cut:  $12 < p_T < 40 \text{ GeV}/c$ .

243 – Cluster pseudorapidity:  $|\eta| < 0.67$

244 The cluster pseudorapidity is corrected for the position of the primary interaction vertex.

245 – Number of cells cut:  $N_{\text{cell}} \geq 2$

246 This requirement removes clusters that are likely dominated by noise.

247 – Exotoxicity cut:  $E_{\text{cross}}/E_{\text{cluster}} > 5\%$

248 We remove “exotic” or “spiky” clusters likely coming from slow neutrons or highly-ionizing parti-  
249 cles hitting the avalanche photo-diode of a cell by a requirement on the ratio of the summed energy  
250 around the leading cell to the total cluster energy.

251 – Cluster time cut:  $|t| < 20 \text{ [ns]}$

252 We require a cluster time measurement of  $|t| < 20 \text{ ns}$  to remove out-of-bunch pileup.

253 – Number of local maxima cut:  $N_{LM} < 3$

254 This cuts suppresses background and improves the MC simulation description of the background [2].

255 – Distance seed-cell to bad-channel  $\geq 1$  cells.

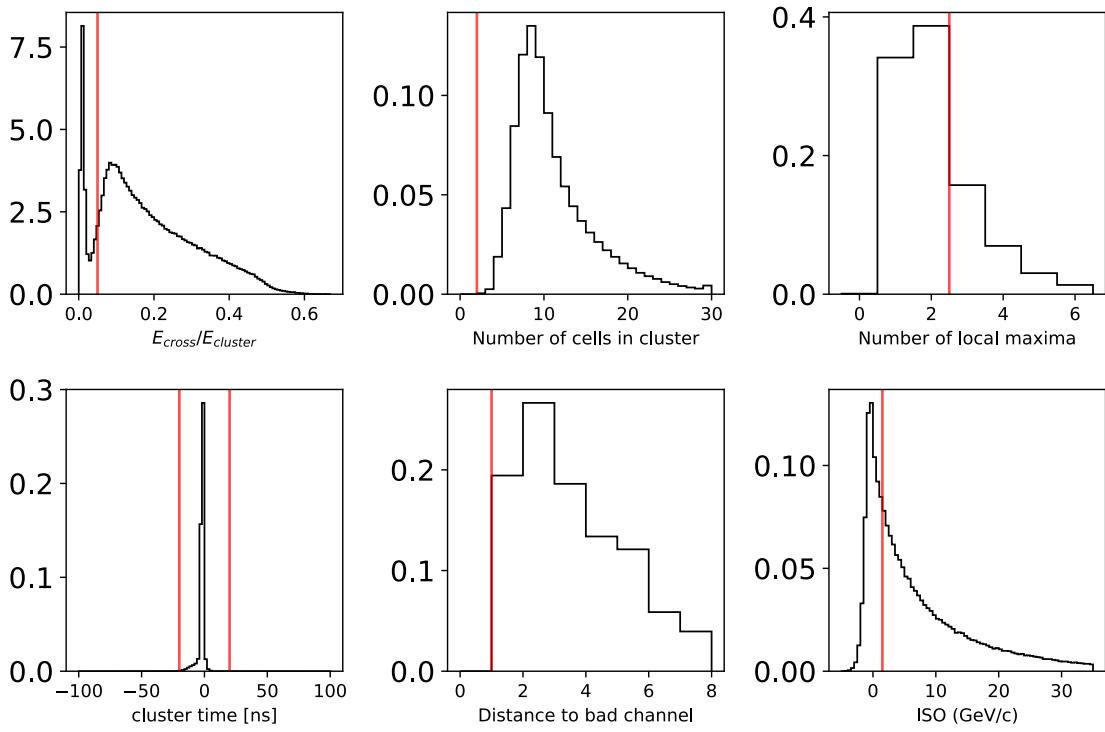
256 Figures 1 and 2 show the distribution of the variables used in the cluster selection and the effect of  
257 sequential selection (“cut flow”) for the p–Pb and pp data respectively. Table 5 shows a summary of the  
258 effect of sequential selection on the number of selected clusters in both pp and p–Pb data.

**Table 5:** Number of clusters, with  $12 < p_T < 40 \text{ GeV}/c$ , that pass our selection in 2013 p–Pb and 2017 pp data.

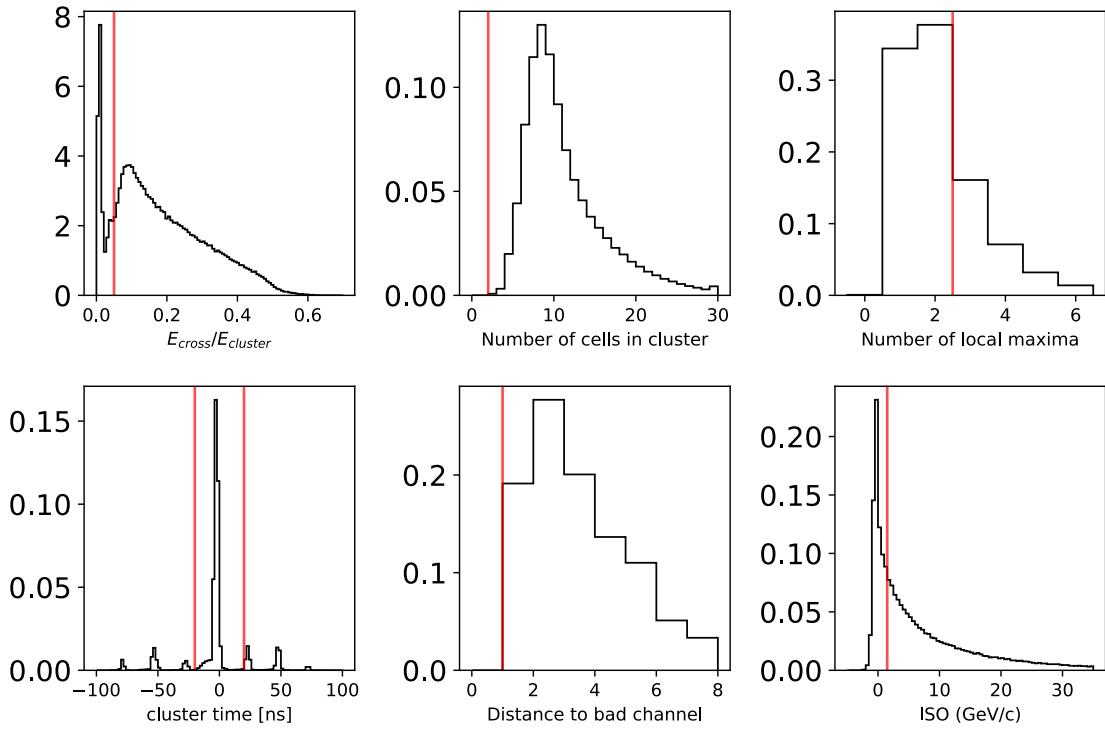
Selection	p–Pb data	pp data
$ \eta  < 0.67$	714834	385220
$E_{\text{cross}}/E_{\text{cluster}} > 5\%$	613560	323750
$N_{\text{cell}} \geq 2$	613560	323750
$N_{LM} < 3$	443102	231490
$ t  < 20 \text{ [ns]}$	441639	171470
Distance-to-bad channel $\geq 1$	441639	171470
$\text{ISO} < 1.5 \text{ GeV}/c$	137895	58638
$0.1 < \sigma_{\text{long}}^2 < 0.3$	40027	16628

---

<sup>7</sup>This event selection also closely follows previous and concurrent isolated-photon spectra analyses in pp and p–Pb data.



**Fig. 1:** Distribution of variables used in the cluster selection of p–Pb data. The red vertical lines represent the cuts used. The cluster cuts get applied sequentially, i.e. the clusters cut with a given variable do not appear in the next.



**Fig. 2:** Distribution of variables used in the cluster selection in pp data. The red vertical lines represent the cuts used. The cluster cuts get applied sequentially, i.e. the clusters cut with a given variable do not appear in the next.

## 259 8 Photon Identification

260 Photons from  $\pi^0$  decays begin to merge into a single cluster in the EMCal above approximately 6 GeV/ $c$ .  
 261 To identify clusters produced by single photons and reject clusters produced by two photons from a  
 262 meson decay, we use variables that encode the shape of the calorimeter shower.

### 263 8.1 Shower-shape variable

264 The  $\sigma_{\text{long}}^2$  variable is the weighted root-mean-square of the shower energy along the major ellipse axis,  
 265 defined according to Ref. [9] as:

$$\sigma_{\text{long}}^2 = \frac{s_{\eta\eta} + s_{\phi\phi}}{2} + \sqrt{\frac{(s_{\eta\eta} - s_{\phi\phi})^2}{4} + s_{\eta\phi}^2}, \quad (1)$$

266 where  $s_{ij} = \langle ij \rangle - \langle i \rangle \langle j \rangle$  are the covariance matrix elements; the  $i, j$  are cell indices in  $\eta$  and  $\phi$  axes;  $\langle ij \rangle$   
 267 and  $\langle i \rangle, \langle j \rangle$  are the second and the first moments of the cluster position cell weighted as follows:

$$\text{weight} = \max(\log(E_{\text{cell}}/E_{\text{cluster}}), w_0). \quad (2)$$

268 Following previous work [13], we chose the cutoff in the log-weighting as  $w_0 = -4.5$ , which means  
 269 that cells that contain less than  $e^{-4.5} = 1.1\%$  of the total cluster energy are not considered in the  $\sigma_{\text{long}}^2$   
 270 calculation.

271 Since  $\sigma_{\text{long}}^2$  represents the extent of the cluster, it discriminates between clusters belonging to single  
 272 photons, for which the  $\sigma_{\text{long}}^2$  distribution is narrow and symmetric, and merged photons from neutral-  
 273 meson decays, for which the distribution is dominated by a long tail towards higher values.

## 274 9 Isolation

275 In leading-order perturbative calculations, prompt photons are produced surrounded by small hadronic  
 276 activity and fragmentation photons are only found within a jet. Beyond leading order, the direct and  
 277 fragmentation components have no physical meaning and cannot be factorized; the sum of their cross  
 278 sections is the physical observable. However, theoretical calculations can be simplified through the use  
 279 of an isolation requirement, which also suppresses the background from decays of neutral mesons.

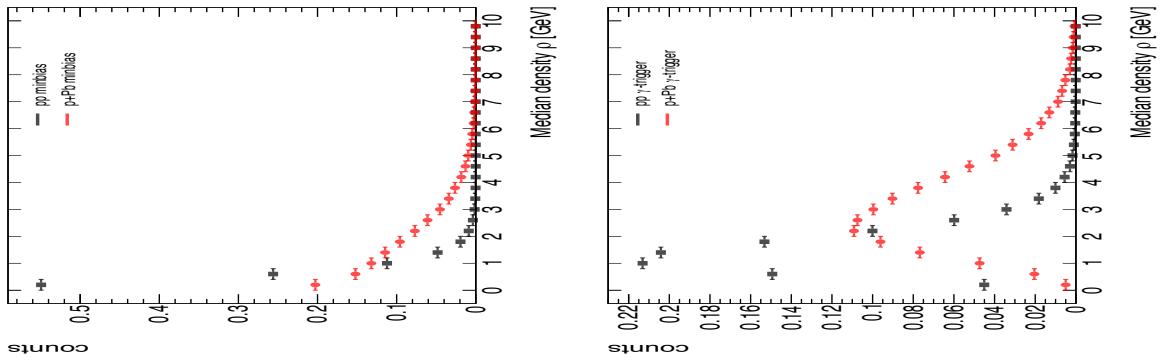
280 We construct an isolation variable using only tracking information and avoid using calorimeter clusters.  
 281 This choice prevents biases due to the correlation between isolation criteria and  $\pi^0$  decay opening angle  
 282 and allow us to use the full acceptance of EMCal. This is at the expense of a lower purity.

283 The isolation variable for this analysis is defined as the scalar sum of the transverse momentum of  
 284 charged particles within an angular radius around the cluster direction,  $R = \sqrt{(\Delta\phi)^2 + (\Delta\eta)^2} = 0.4$   
 285 (with  $\Delta\phi$  measured in radians), thus:

$$\text{ISO}^{\text{raw}} = \sum_{\text{track } \in \Delta R < 0.4} p_{\text{T}}^{\text{track}} \quad (3)$$

286 The charged particles used as input for the isolation calculation have  $0.15 < p_{\text{T}} < 10$  GeV/ $c$ ,  $|\eta| < 0.8$   
 287 and pass the selection described in Section 11.

288 The isolation variable defined in Equation 3 is susceptible to background from the charged particles from  
 289 the underlying event, e.g. a truly isolated photon might appear non isolated due to overlap with particles  
 290 not associated with the hard scattering. In order to correct for this effect, we apply an event-by-event  
 291 underlying event subtraction, which is described in the following section.



**Fig. 3:** Distribution of the median charged-particle transverse momentum density,  $\rho$ , in pp and p–Pb data, for a minimum-bias selection (left panel) and in photon-triggered events (right panel).

### 9.1 Underlying Event estimation

In this section, we describe how we estimate and subtract the “underlying event” (UE). The UE is defined as the particles not associated with the hard-scattering of the collision<sup>8</sup>. We subtract the UE from the measured transverse momentum of jets and in the photon isolation (described in Section 9). Technically, the discrimination between the soft component from the hard component of an event is performed using the FASTJET jet area/median method [7], which uses the median of the distribution of transverse momentum densities of all jets in an event. We use one of the standard jet areas definition implemented in FASTJET called Voronoi area<sup>9</sup>.

The estimation of the UE density uses jets  $J'$  reconstructed by the  $k_T$ -algorithm<sup>10</sup> with distance parameter  $R = 0.3$ . The estimated UE density is defined as:

$$\rho = \text{med} \left\{ \frac{\sum_{i \in J'_k} p_{T,i}}{\sum_{i \in J'_k} A_i} \right\} \quad (4)$$

where  $p_{T,i}$  is the transverse momentum, and  $A_i$  the Voronoi area of the particle  $i$  within the jet reconstructed for UE estimation purpose  $J'_k$ . Following standard practice, the two leading jets are not considered in this observable, to limit the contribution from the hard component of the interaction.

The choice of the median is motivated by its robustness against outliers, which includes jets originated by hard interactions. The observable thus isolates UE by assuming that most of the event is either empty or dominated by soft contributions and that the hard component of the interaction is well contained within the leading jets [7].

Figure 3 shows the median charged-particle density,  $\rho$ , distribution obtained in pp and p–Pb data in minimum bias events and in events that pass the selection in Section 6 and thus have a high- $p_T$  cluster. The distribution in minimum-bias events decreases approximately exponentially. The distribution in photon-triggered events is different and follows sort of an asymmetric Gaussian distribution that peaks at approximately 1.0 GeV/c and 2.5 GeV/c for pp and p–Pb collisions, respectively.

The mean and standard deviation for each distribution is shown in Table 6. The difference in UE-density

<sup>8</sup>In Ref. [8], the UE is defined as “the sum of all the processes that build up the final hadronic state in a collisions excluding the hardest leading order partonic interaction. This includes fragmentation of beam remnants, multi-parton interactions and initial and final-state radiation associated with each interaction.”

<sup>9</sup>The method used is the following `fastjet::VoronoiAreaSpec` [http://www.fastjet.fr/repo/doxygen-2.4.5/classfastjet\\_1\\_1VoronoiAreaSpec.html](http://www.fastjet.fr/repo/doxygen-2.4.5/classfastjet_1_1VoronoiAreaSpec.html)

<sup>10</sup>In contrast to the anti- $k_T$  algorithm, the  $k_T$  algorithm clusters the softest particles first, and thus is more sensitive to the details of the distribution of softer objects and better suited for an investigation of the underlying event.

315 in p–Pb is expected due to the increased number of nucleon-nucleon collisions. The UE-densities shown here are still about a factor of 50 lower than in central Pb–Pb collisions.

**Table 6:** Median transverse momentum density mean and standard deviation in minimum-bias and photon-triggered events in pp and p–Pb data. The statistical uncertainty in these numbers is negligible.

	pp minbias	pp $\gamma$ -trigger	p–Pb minbias	p–Pb $\gamma$ -trigger
$\langle \rho \rangle$	0.49 GeV/c	1.51 GeV/c	1.56 GeV/c	3.19 GeV/c
$\sigma_\rho$	0.47 GeV/c	0.85 GeV/c	1.32 GeV/c	1.60 GeV/c

316  
317 The average  $\rho$  for photon-triggered events reported in Table 6 is consistent with an independent estimate,  
318 based on the “ $\eta$ -band” method, that uses the same dataset and cluster selection [19].

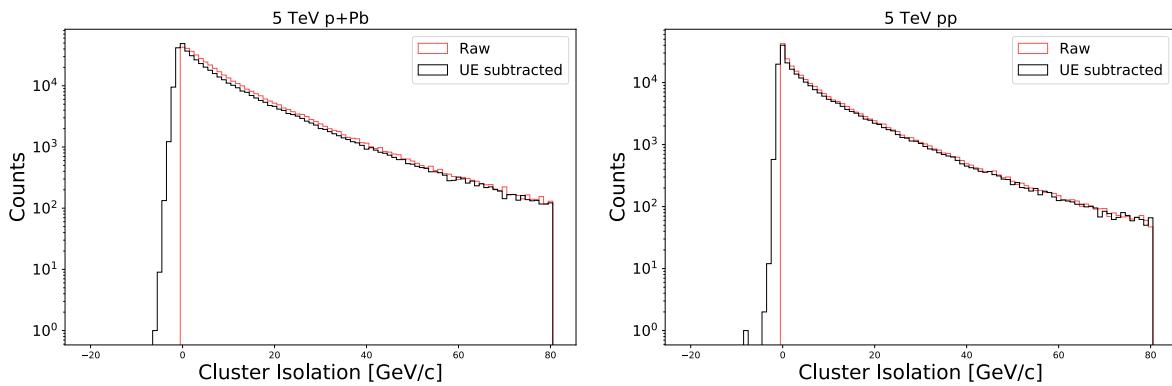
## 319 9.2 UE correction to isolation variable

320 For each event and cluster, we subtract the underlying event using the measured charged-particle density  
321  $\rho$  that is calculated event-by-event as described in Section 9.1:

$$\text{ISO} = \text{ISO}^{\text{raw}} - \rho \times \pi(0.4)^2. \quad (5)$$

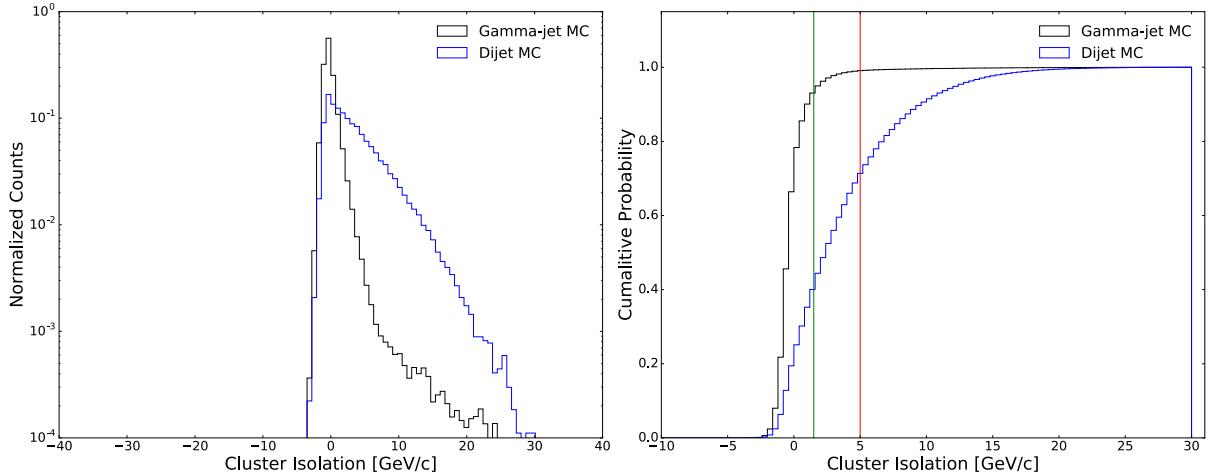
322 Thus, the average subtraction for the isolation cone of  $R = 0.4$  is about 1.6 GeV/c and 0.8 GeV/c for  
323 p–Pb and pp collisions, with a standard deviation of 0.9 GeV/c and 0.4 GeV/c, respectively.

324 Figure 4 shows the isolation distribution before and after underlying event subtraction for p–Pb and pp  
325 collisions. The distributions have a positive tail that decreases exponentially; this is expected as this ob-  
326 servable effectively measures multi-jet production. The difference between the p–Pb and pp distribution  
327 at low ISO values can be attributed to the effect of enhanced soft-particle production in p–Pb collisions.  
328 The underlying event subtraction modifies the isolation distribution only slightly. After subtraction, the  
329 distributions show a negative tail, which arises from a over-subtraction of the underlying event due to  
330 region-to-region fluctuations. In both cases, this tail falls by more than three orders of magnitude by  
331  $\text{ISO} = -3$  GeV/c, indicating that over-subtraction is a small effect.



**Fig. 4:** Cluster isolation before and after underlying event subtraction in p–Pb (left panel) and pp (right panel) collisions.

332 Figure 5 shows the distribution of cluster isolation after UE subtraction for photon-jet and dijet sim-  
333 ulations of p–Pb data (see Table 3). As expected, the distributions are rather different: whereas the  
334 dijet simulation shows a prominent exponential tail at large ISO values, the photon-jet simulation shows  
335 a Gaussian-like shape that is mostly symmetrical except for a very small fraction of events that have  
336 large ISO values. In both cases, the negative tail falls rather sharply, which is expected as it arises from  
337 region-to-region fluctuations of the UE that are independent of the hard-process involved.



**Fig. 5:** Isolation distribution of clusters that pass our selection in p–Pb photon-jet and dijet simulations, and corresponding cumulative distribution. Two vertical lines at ISO = 1.5 GeV/c (green) and ISO = 5.0 GeV/c are shown in the right panel for reference.

338 For the purposes of template fitting, we also need to define a sideband that is dominated by background.  
 339 For this we note that only about 1% of prompt photons of the photon-jet simulation have ISO > 5 GeV/c.  
 340 Given that the cross-section for prompt photons is about two orders of magnitude smaller than the back-  
 341 ground, this region is overwhelmingly dominated by background. The cumulative distributions (Figure 5,  
 342 right panel) show that a ISO < 1.5 GeV/c selection keeps about 90% of the signal and rejects about 60%  
 343 of the background. We use this relatively loose photon isolation criteria to reduce the dependence of  
 344 the results on the details of the simulation of the detector noise, tracking resolution, and the underlying  
 345 event.

346 This isolation cut of ISO < 1.5 GeV/c is used in conjunction with the shower-shape cut to complete  
 347 our isolated-photon selection or “ $\gamma^{\text{iso}}$  selection”. We call “ $\gamma^{\text{iso}}$ -candidates” the clusters that pass our  
 348 isolated photon selection because it still leaves a significant fraction of background (about 40% of the  
 349 cross section, as just shown).

350 The main background present in our  $\gamma^{\text{iso}}$  selection is from multi-jet events where one jet typically contains  
 351 a  $\pi^0$  or  $\eta$ , which carries most of the jet energy, and is misidentified as a photon because it decays into  
 352 a photon pair that is collinear with respect to the EMCAL cell granularity. Other sources of background  
 353 arise from charged-to-neutral fluctuations of jet fragmentation that leads to low observable ISO (that  
 354 considers only charged-particles).

355 This creates the need to measure the purity of our  $\gamma^{\text{iso}}$ -candidate selection, which is described in Sec-  
 356 tion 10.

## 357 10 Purity Measurement

358 While the isolation requirements described in Section 9 remove the bulk of the neutral-meson back-  
 359 ground, a substantial contribution remains; in particular, multi-jet events that produce a  $\pi^0$  or  $\eta$  that  
 360 carries most of the jet energy can pass the selection. Given that the cross-section to produce a  $\pi^0$  is much  
 361 larger than the cross-section for prompt photons [5], the background is still large for the  $p_T$  range that  
 362 we are interested in. In this section, we show measurements of the purity of our  $\gamma^{\text{iso}}$ -candidate sample  
 363 done with a template-fit method.

### 364 10.1 Template fit method

365 The purity of the isolated photon sample is determined with a two-component template fit, as was done  
 366 for example by the CMS collaboration in Ref. [15]. The distribution of the shower shape variable for  
 367 the isolated cluster sample is fit to a linear combination of the signal distribution and the background  
 368 distribution. The shape of the signal distribution is determined by a photon-jet simulation (see Table 3)  
 369 and the shape of the background distribution is determined from data using an anti-isolated sideband<sup>11</sup>  
 370 with an additional correction computed from a dijet simulation. This is described in more detail in the  
 371 following sections.

#### 372 10.1.1 Signal template and background templates

373 We estimate the shape of the background distribution of the shower-shape for isolated clusters with a  
 374 sideband technique. That is, we estimate the shower shape distribution of clusters from isolated decay  
 375 photons with clusters that are anti-isolated but pass all other selection criteria. This method assumes  
 376 that the correlation between the isolation variable and shower shape variable can be corrected for; the  
 377 procedure for doing so is described below. The signal and sideband regions defined using the isolation  
 378 variable are illustrated in Figure 6.

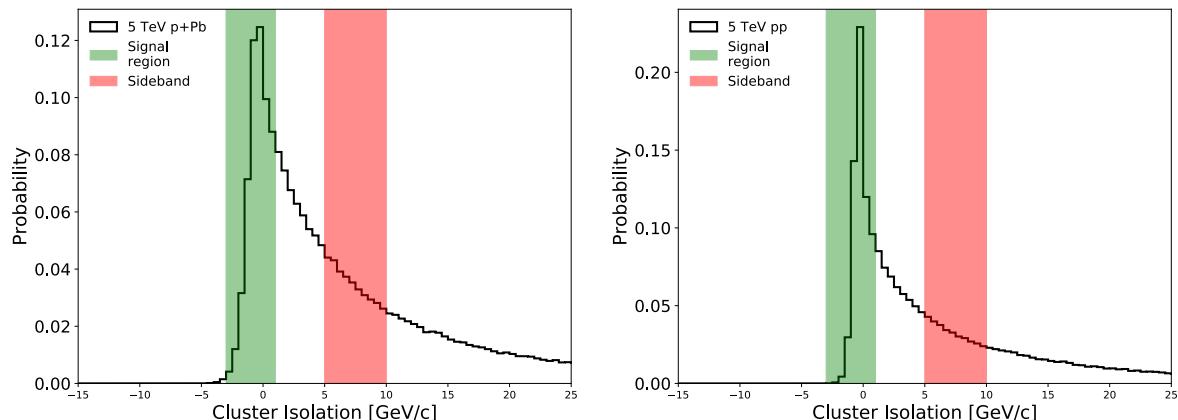


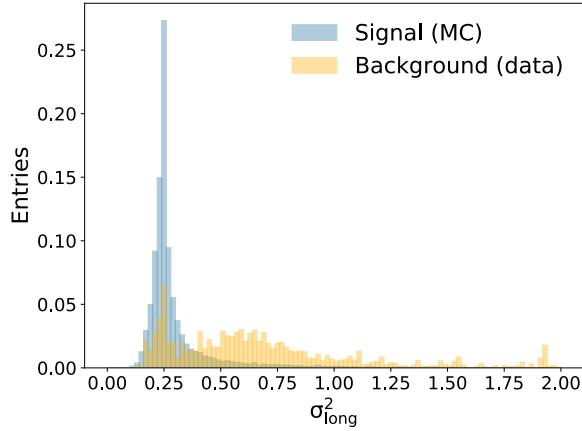
Fig. 6: Isolation variable distribution of clusters with  $p_T$  between 12 and 16  $\text{GeV}/c$  in p–Pb data (left panel) and pp data (right panel). The green shaded area represents the signal region ( $\text{ISO} < 1.5 \text{ GeV}/c$ ); the red represent the sideband ( $5 < \text{ISO} < 10 \text{ GeV}/c$ ) used to estimate the background template.

379 For simplicity, the same definitions are used for pp and p–Pb data. The lower bound of the sideband  
 380 region is defined as  $\text{ISO} = 5 \text{ GeV}/c$ ; according to photon-jet simulations, less than 1% of prompt photons  
 381 are beyond this range. The upper bound is chosen such that the sideband is as narrow as possible, to  
 382 minimize a possible bias to the shower-shape distribution due to a positive correlation with ISO, while  
 383 still containing a number of clusters comparable to the signal region. A more rigorous study on the  
 384 sensitivity of our purity estimate on the choice of sideband region is shown in Section 10.2.2.

385 Figure 7 summarizes the signal and background templates used in the template fit. The distributions  
 386 are rather different, which is key for the stability of the template fit. The background shape in the  $\sigma_{\text{long}}^2$   
 387 variable shows a peak in the single-shower region but a “bump” that reflects a  $\pi^0$  peak. In both cases,  
 388 the peaks in the single-shower region that are observed in the background templates come mostly from  
 389 collinear  $\pi^0 \rightarrow \gamma\gamma$  decays.

390 The background template is corrected for the correlation between the shower shape distribution and  
 391 the isolation energy. A dijet simulation is used to construct the ratio of the shower shape distribution

<sup>11</sup>The inversion of an isolation cut to estimate QCD background is a standard technique in several measurements and searches at the LHC and previous hadron colliders.



**Fig. 7:** Normalized signal (blue) and background (yellow) distributions used as input for the template fit. These distributions correspond to clusters with  $p_T$  in the 14–16 GeV/ $c$  range.

for isolated clusters to the shower shape distribution for anti-isolated clusters. This ratio, as a function of shower shape variable, is then applied as a weight to the anti-isolated clusters in the data, giving a corrected background template that is then used in the template fit (see Equation 6). From this equation, it is clear that, if the MC exactly replicates the data, the Weights function exactly corrects the anti-isolated decay photon  $\sigma_{\text{long}}^2$  distribution back to the isolated decay photon  $\sigma_{\text{long}}^2$  distribution, which is the true background.

$$\begin{aligned} \text{Weights}(\sigma_{\text{long}}^2) &= \frac{\text{Iso}_{\text{MC}}(\sigma_{\text{long}}^2)}{\text{Anti-iso}_{\text{MC}}(\sigma_{\text{long}}^2)} \\ \text{Bkg}^{\text{corrected}}(\sigma_{\text{long}}^2) &= \text{Non-iso}_{\text{data}}(\sigma_{\text{long}}^2) \times \text{Weights}(\sigma_{\text{long}}^2) \end{aligned} \quad (6)$$

The purity as computed with the corrected background template is 8–13% lower in absolute value as compared to the purity as computed with the uncorrected background template; an example of a fit with and without the correction is shown in Figure 8. The correction greatly improves the goodness-of-fit. The Weights( $\sigma_{\text{long}}^2$ ) function for different  $p_T$  ranges is shown in Appendix F and the evaluation of the systematic uncertainty associated with this correction is described in Section 10.2.

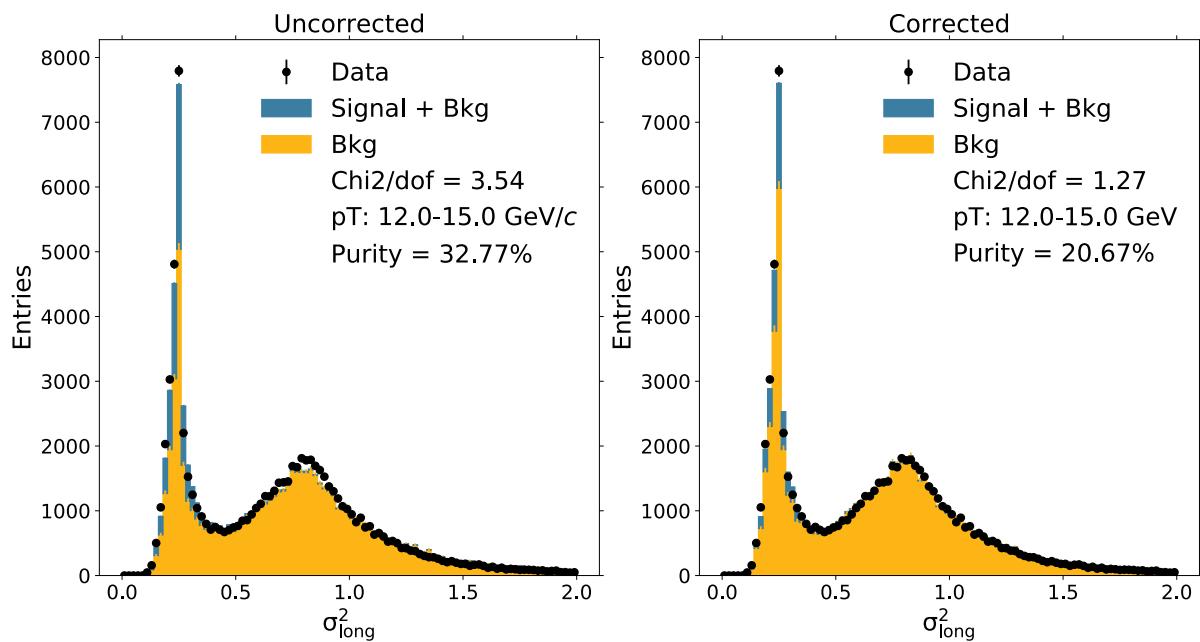
#### 10.1.2 Fit results

The distribution of isolated clusters is fit with a linear combination of the signal and background templates. We use the MINUIT [18] package for  $\chi^2$  minimization and the MIGRAD package for error estimation. The only free parameter in the fit is the number of signal clusters,  $N_{\text{sig}}$ , because the overall normalization,  $N$ , is fixed to the total number of isolated clusters:

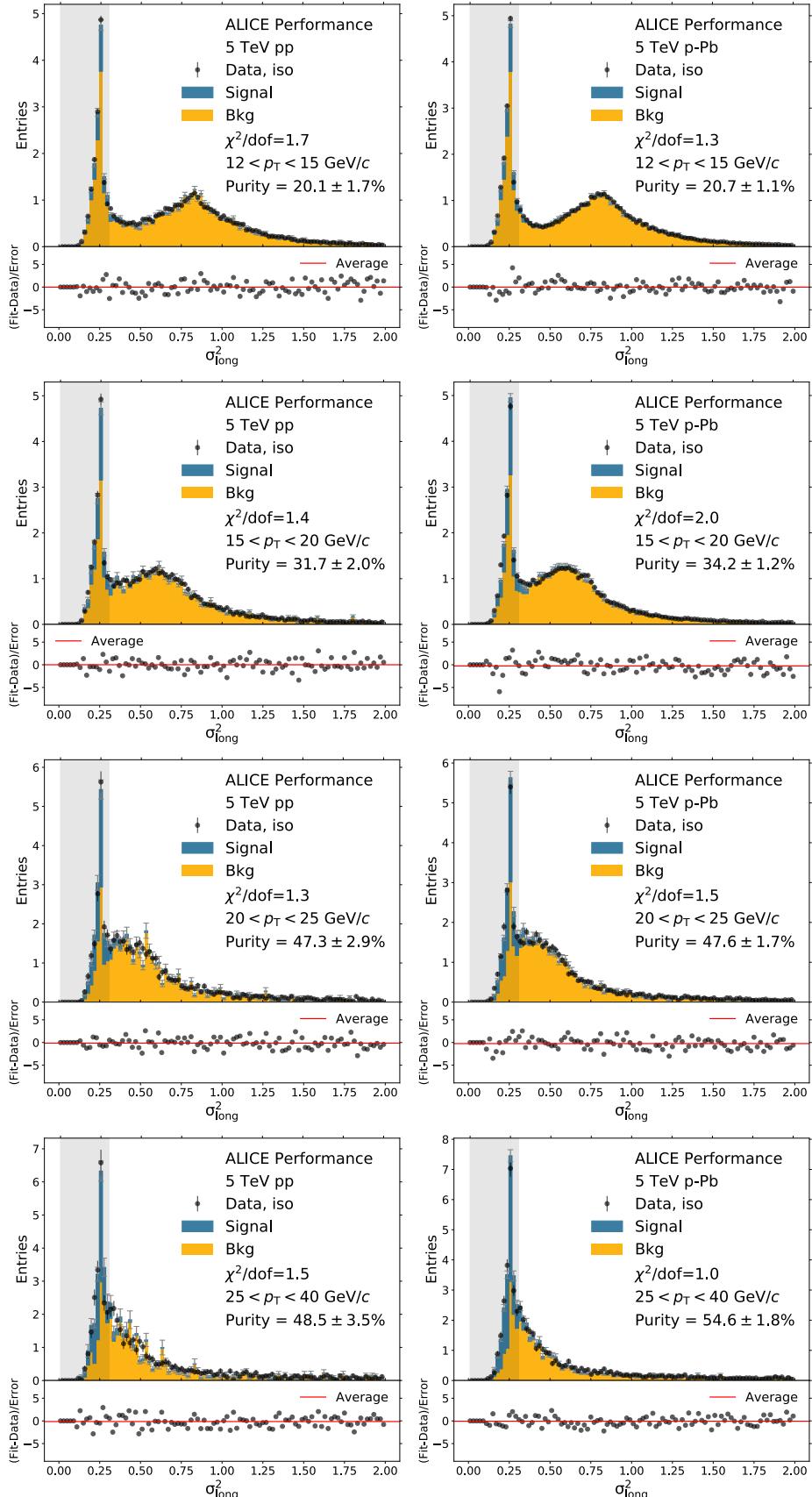
$$N^{\text{observed}} = N_{\text{sig}} \times S + (N - N_{\text{sig}}) \times B, \quad (7)$$

where  $S$  and  $B$  are the normalized signal template and background template.

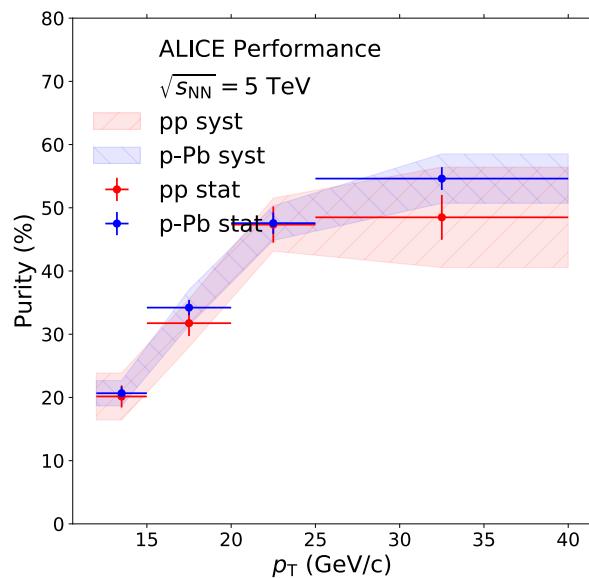
Figures 9 show template fit results for p–Pb and pp data. In all cases a good fit with no systematic pattern in the residuals is achieved over most of the distribution, and the reduced  $\chi^2$  ranges within acceptable values. The purity measurements are presented in graphical form in Figure 10.



**Fig. 8:** An example of the template fit with and without the background template correction in p–Pb for clusters with  $12 < p_{\text{T}} < 15$  GeV/c. The goodness of fit is better after the correction and the purity is significantly lower.



**Fig. 9:** Template fit results in pp and p–Pb data. The stacked histograms (yellow for background, blue for signal) are the predicted counts given the best-fit value of the number of signal photons,  $N_{\text{sig}}$ . The hatched gray area represents the interval considered for the purity estimate. The bottom panels show the normalized residuals of the fit, considering the statistical uncertainty on the isolated cluster data and the background template added in quadrature.



**Fig. 10:** Purity of isolated-photon selection as a function of cluster  $p_T$ . The error bar represents statistical uncertainty only. The error band represents the systematic uncertainty only.

412 **10.2 Systematic uncertainties of the purity measurement**

413 There are two assumptions underlying the template fit procedure. The first is that the signal template from  
 414 simulations is correct. The second is that the shape of the background estimated from the anti-isolated  
 415 sideband, with the correction coming from simulations, reflects the shape of the background in the signal  
 416 region. That is, the assumption is that the correlation between the shower shape and isolation variables  
 417 can be corrected for via an appropriate simulation. The dominant sources of systematic uncertainty on  
 418 this measurement are described in this section and can be summarized as follows: the signal template,  
 419 the sideband region selection, and the background template correction. We also investigated the effect of  
 420 varying our cluster selection but found that the variations on the purity measurements are much smaller  
 421 than the other sources of systematic uncertainties investigated here, so we neglect that. This is shown in  
 422 Appendix E.

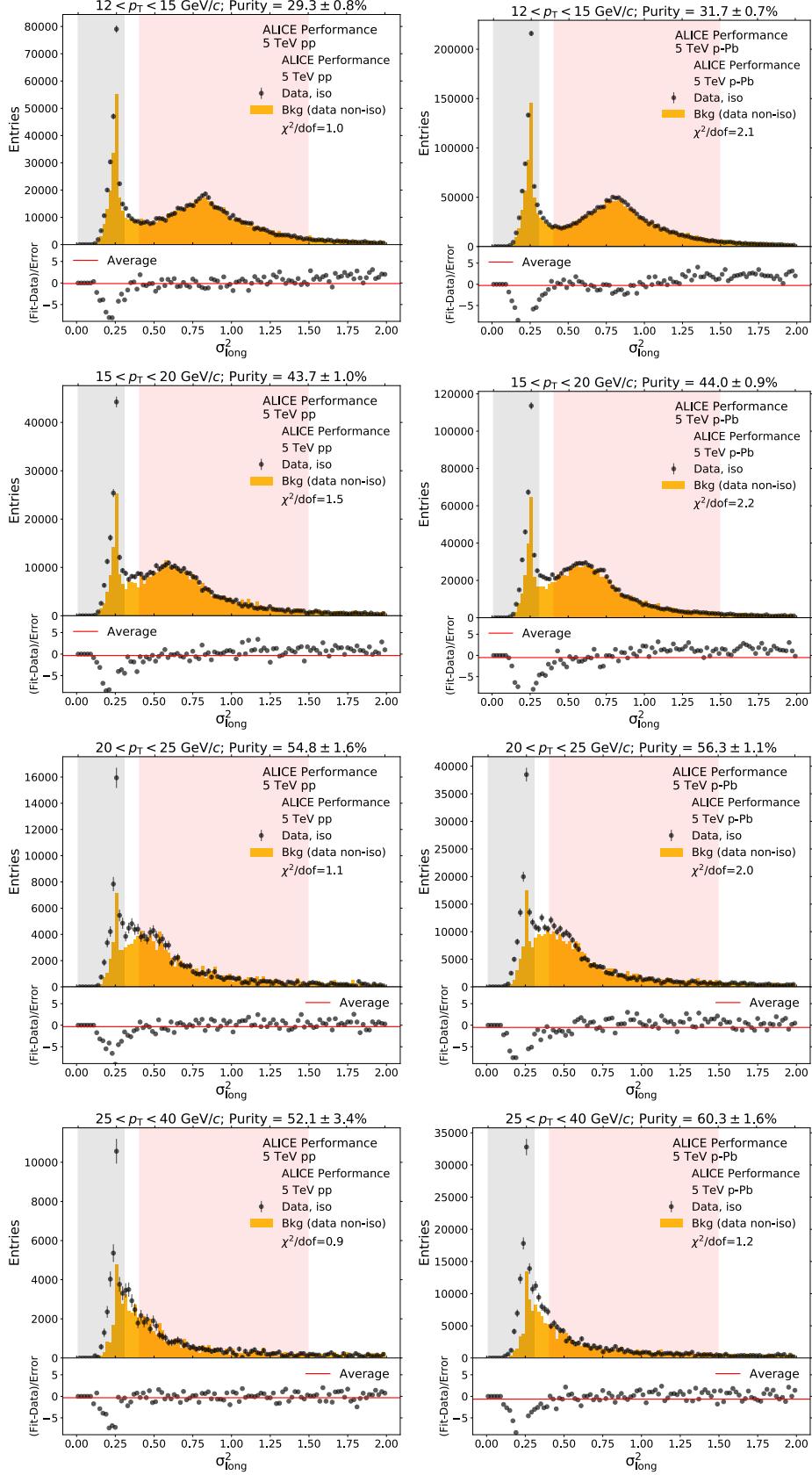
423 **10.2.1 Signal template**

424 We estimate the systematic uncertainty on the purity estimate due to imperfections in the signal template  
 425 by using a data-driven template fit. In this, we restrict the range of the  $\chi^2$  fit to the background-dominated  
 426 region of the shower-shape distribution (0.4–1.5 for  $\sigma_{\text{long}}^2$ ) and we use only the background template to  
 427 fit the isolated data with the normalization as the only free parameter. To factorize the effect of the  
 428 MC-correction to the background template, we do not apply it for this study. Once the background  
 429 normalization is fitted, we consider the signal to be the integral of the isolated data minus the integral of  
 430 the background, both in the signal region of the shower-shape variable.

431 Figure 11 show the results obtained with this method in pp and p–Pb data. We observe some systematic  
 432 pattern in the residuals, which can be attributed to the lack of MC-correction on the background template.

433 This method makes the additional assumption that the fraction of signal misclassified as background is  
 434 small. In a strict sense, this method yields a lower limit on the extracted purity. The results agree with  
 435 our nominal results within a few percent, indicating that we are not particularly sensitive to the details of  
 436 the modeling of the shower shape. As a conservative estimate, we take the full difference between our  
 437 nominal results as a systematic uncertainty in the signal template.

438 As an additional check, we smear the signal template by multiplying the  $\sigma_{\text{long}}^2$  of each cluster by a random  
 439 number selected from a Gaussian with a fixed width before then calculating the purity with this smeared  
 440 distribution. This was done for a variety of widths up to 10%, which is much larger than the expected  
 441 MC simulation mismodelling, and was found to yield a smaller uncertainty than the background-only  
 442 fit (See Appendix G for more details). Thus, for a final estimate of the systematic uncertainty from  
 443 the signal template, we take the uncertainty estimated by the background-only fits as described in the  
 444 previous paragraph.



**Fig. 11:** Template fit results of background-only template method for pp and p-Pb data. The yellow histograms are the predicted counts given the best-fit value of the total number of clusters in the background dominated region. The hatched gray area represents the interval considered for the purity estimate. The bottom panels show the normalized residuals of the fit, considering the statistical uncertainty on the isolated data and the background template added in quadrature.

445 ***10.2.2 Sideband variation in the background template***

446 To estimate the shower-shape distribution for the  $\gamma^{\text{decay}}$  background in the template fit, a sideband in the  
 447 cluster isolation variable is used. Only the shape of this distribution is relevant, as the overall background  
 448 normalization in the signal region (i.e. the purity) is measured with the template fit. As in any analysis  
 449 using a sideband technique, we nominally use a sideband as close as possible to the signal region and  
 450 as narrow as possible. Here, we describe how the sideband region is chosen and address the systematic  
 451 uncertainty that arises from this arbitrary choice.

452 The cluster isolation distribution is divided into narrow ( $2 \text{ GeV}/c$ ) overlapping regions, each of which  
 453 is used to estimate the background shower-shape distribution. A template fit is performed with each  
 454 distribution and the  $\chi^2/\text{dof}$  and purity are calculated for each fit and plotted as a function of the anti-  
 455 isolation region used to create the background template in Figure 12. Then, the  $\chi^2/\text{dof}$  distribution is  
 456 examined to determine which regions of anti-isolation result in good fits in the template fit procedure:  
 457  $5\text{--}10 \text{ GeV}/c$  is chosen to be the sideband definition in the final purity calculation.

458 To calculate the systematic uncertainty on the purity due to this selection, the full range of purities  
 459 reached by the narrow bands of anti-isolation that fall within  $5\text{--}10 \text{ GeV}/c$  is considered. Converting  
 460 the full extent to a systematic uncertainty is a matter of dividing by  $\sqrt{12}$  (i.e., the  $1\sigma$  for a uniform  
 461 distribution). This results in an absolute uncertainty on the purity of 0.7–5.8%, depending on the collision  
 462 system and cluster  $p_{\text{T}}$  range.

463 ***10.2.3 Background template correction***

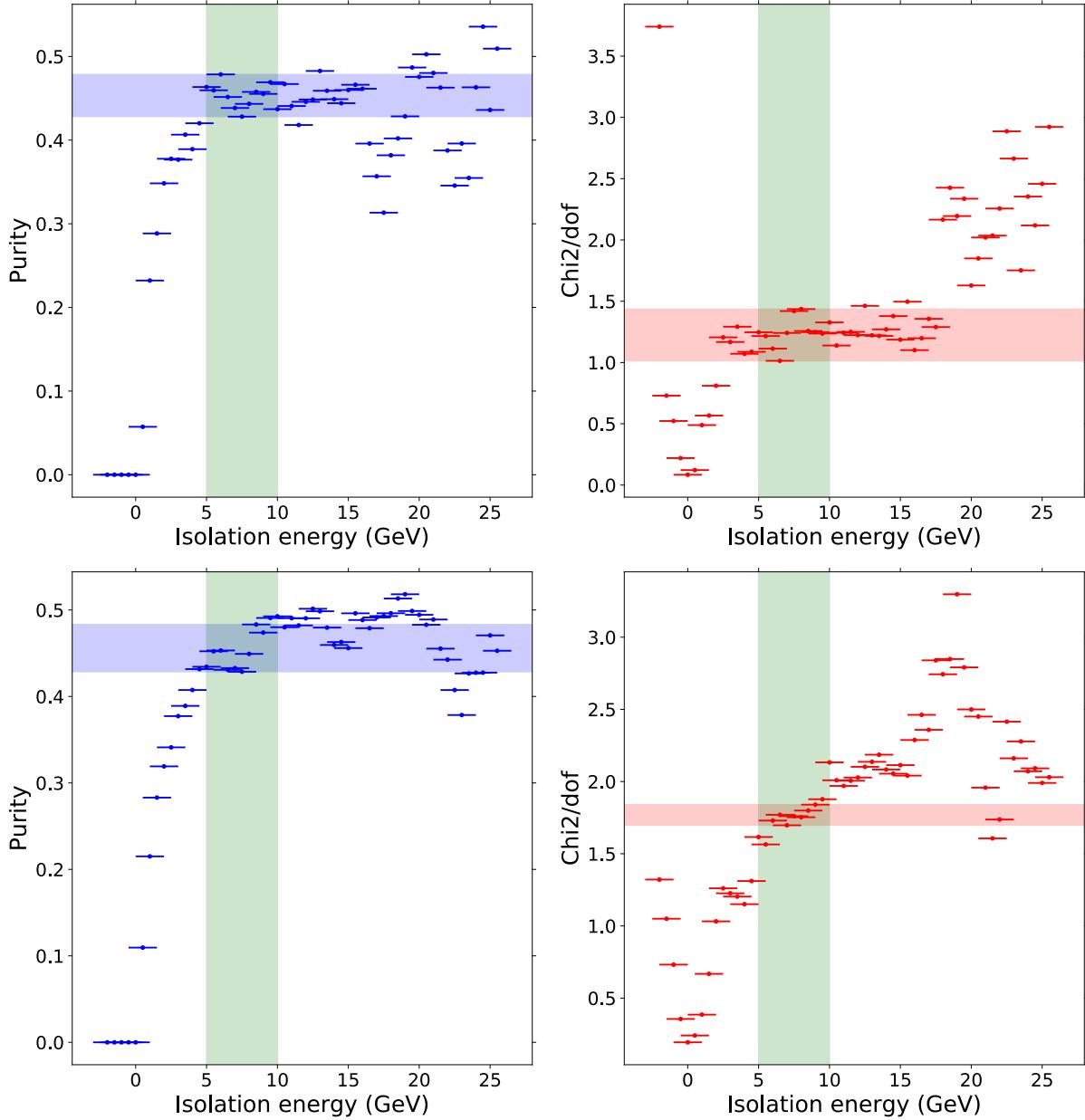
464 Due to the correlation between the isolation and shower shape, the template extracted from the anti-  
 465 isolated sideband does not exactly reflect the shape of the background in the signal region. Clusters in  
 466 the isolation sideband have more associated activity than those in the true isolated background and thus  
 467 emphasize the non-signal region of the shower-shape distribution. Consequently, using the isolation side-  
 468 band instead of the true isolated background yields systematically higher purities. We note that a similar  
 469 observation was made for example by the CMS collaboration in their template-fit purity measurements  
 470 (e.g. Ref. [15])

471 We correct this bias using a dijet MC simulation, as described in Equation 6. However, this correction is  
 472 only valid to the extent that the dijet MC reproduces the data. To estimate the systematic uncertainty on  
 473 this correction, we use a technique based on a method used in the ABCD calculation [19]. In particular,  
 474 we use a double ratio to check to which extend does the dijet MC describe the background-dominated  
 475 region in data:

$$\text{Double ratio} = \frac{\text{Iso}_{\text{data}}/\text{Anti-iso}_{\text{data}}}{\text{Iso}_{\text{MC}}/\text{Anti-iso}_{\text{MC}}} \quad (8)$$

476 In the signal region of the shower shape distribution ( $0.0\text{--}0.3$  for  $\sigma_{\text{long}}^2$ ), this double ratio will be far from  
 477 unity, as the data have prompt photons and the dijet MC do not. However, away from that region, where  
 478 background dominates, the double ratio should be flat (i.e. have no slope) if the dijet MC reproduces  
 479 the background shower-shape of the data. We note that for this analysis only the shape is important  
 480 and overall normalization is irrelevant. At a minimum, we expect the variation in the double ratio to  
 481 be smooth. Thus we fit the double ratio to smooth functions (linear and exponential) in a shower shape  
 482 range away from the signal region and extrapolate the fit back into the signal region. A similar procedure  
 483 was used isolated-photon purity measurements with the ABCD method [2, 19].

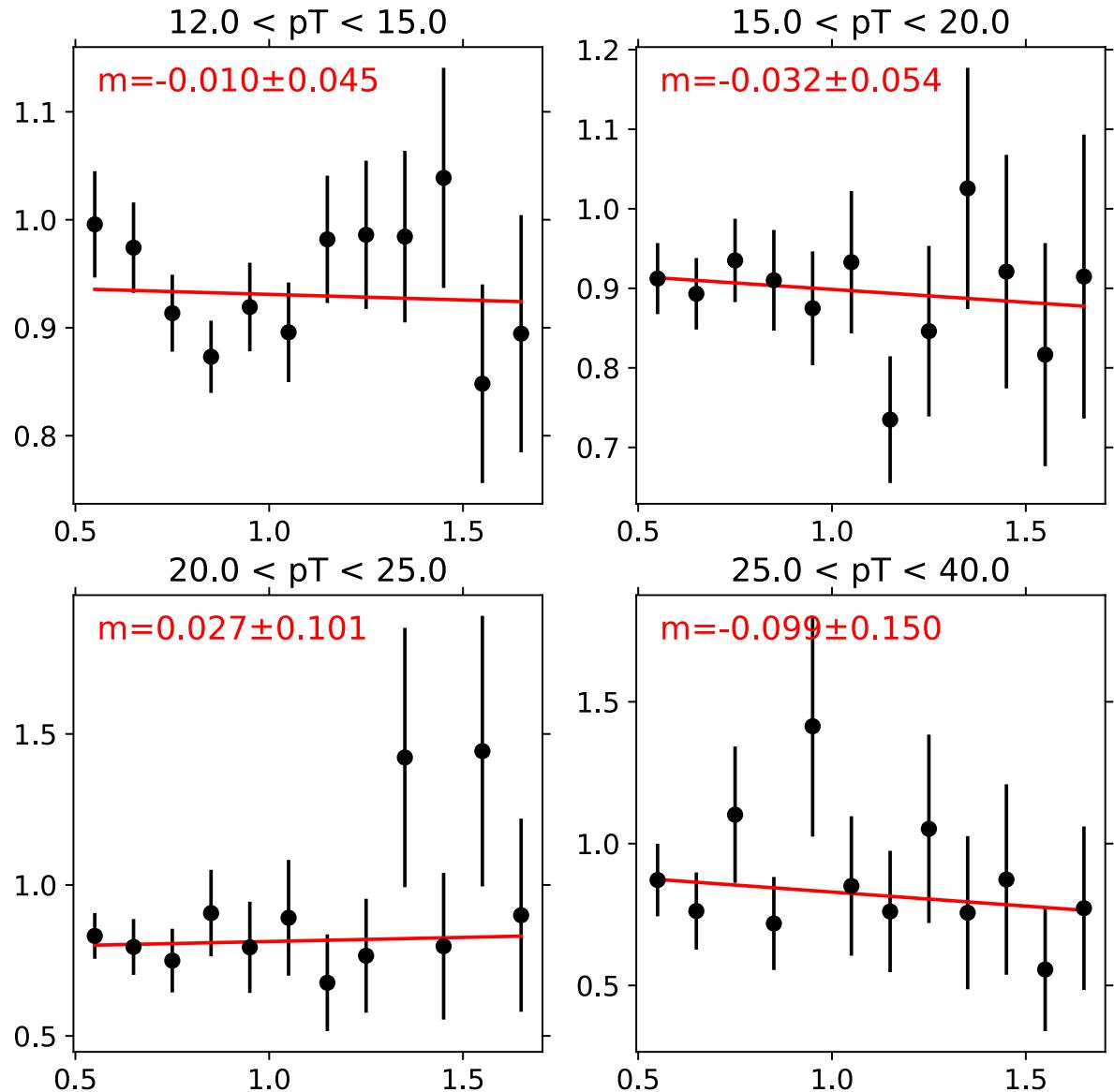
484 Fits to the double ratio are shown in Figures 13 and 14, for pp and p–Pb data respectively. We found that  
 485 the linear and exponential fits gave nearly identical results. In particular, the slope was sufficiently small  
 486 that the higher-order terms in the exponential were negligible. Thus for the purposes of estimating the



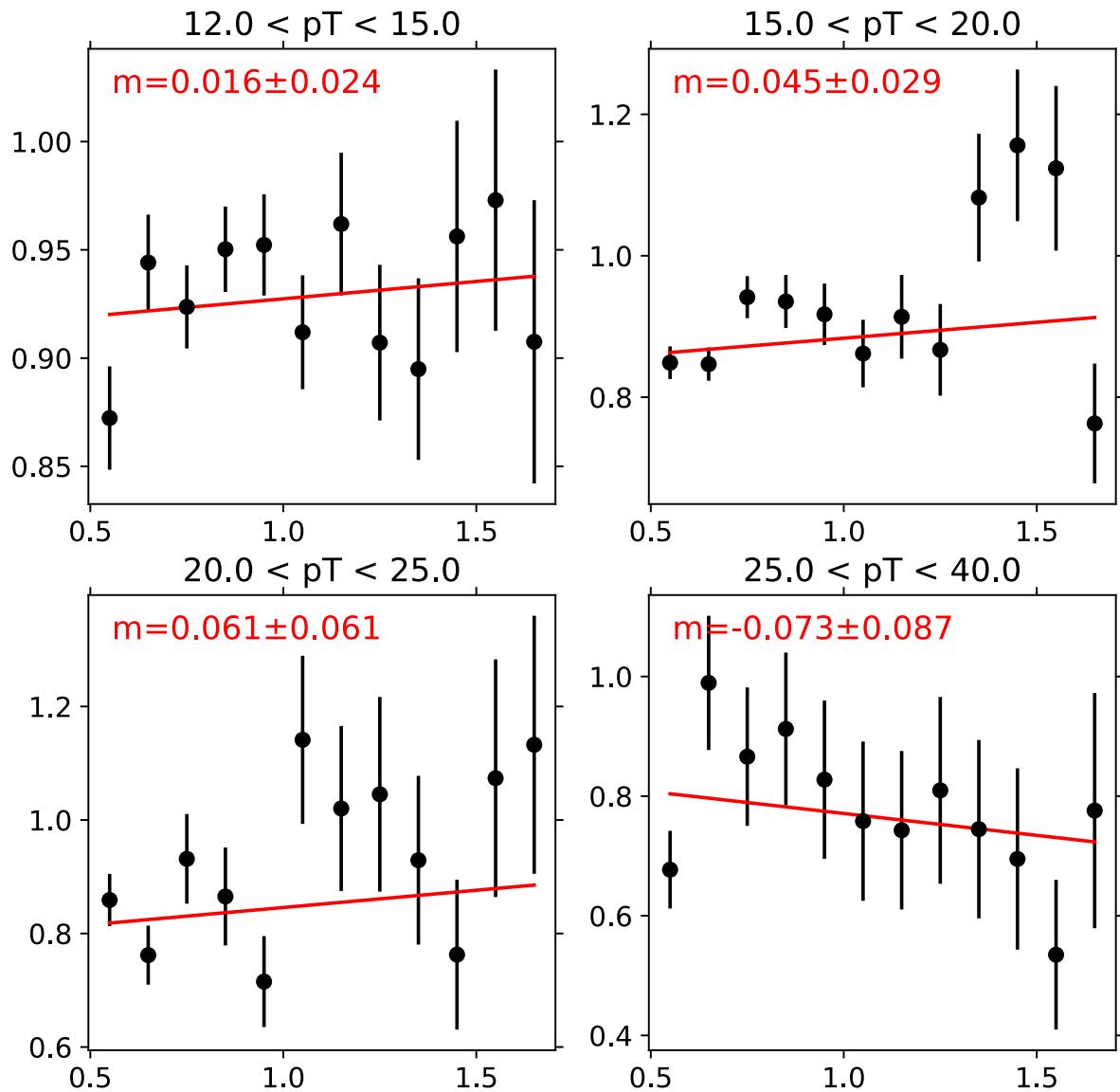
**Fig. 12:** Template fit results (purity and  $\chi^2/\text{dof}$ ) as a function of anti-isolation region for clusters with  $15 < p_T < 20$  GeV/ $c$  in pp (top) and p–Pb (bottom). The green band shows the selected sideband region. The blue and red bands show the full extent of the purity within the selected sideband region.

systematic uncertainty due to the background template correction, we chose to do only linear fits to the double ratio. In order to remove covariance effects between the slope and intercept, the fits were forced to go through the weighted average of the double ratio value within the fit range at the center of the fit range, making it a single-parameter linear fit with only the slope as a free parameter. This allowed us to propagate the fit uncertainty on the slope to an uncertainty on the purity.

These linear fits to the double ratio were done in two fit ranges: 0.5–1.5 and 0.5–1.75 for  $\sigma_{\text{long}}^2$ . In all cases, we found that the slopes were consistent with 0 within the fit uncertainties and thus concluded that the dijet MC was consistent with the data. Therefore, we did not apply an additional double-ratio correction to the Weights function in Equation 6. We also found that the double ratio fits with the different fit ranges gave purities consistent with each other. So in order to minimize the amount of



**Fig. 13:** Linear fits for the double ratio (as described in Equation 8) for the  $\sigma_{\text{long}}^2$  variable in pp data. Included are the value and uncertainty of the fitted slope (in red).



**Fig. 14:** Linear fits for the double ratio (as described in Equation 8) for the  $\sigma_{\text{long}}^2$  variable in p–Pb data. Included are the value and uncertainty of the fitted slope (in red).

497 extrapolation, we did the fit in the largest reasonable fit ranges for each of the variables (the larger of  
 498 each of the ranges described at the beginning of this paragraph).

499 We then took the uncertainty on that double ratio fit and propagated it to a purity uncertainty. This purity  
 500 uncertainty was then taken to be the systematic uncertainty on the background correction. It varies  
 501 between 1.2–5.1% (absolute) depending on cluster  $p_T$  and collision system.

### 502 10.3 Summary of systematic uncertainties of purity measurement

503 Tables 7 and 8 give the full estimates of the systematic uncertainties in both collision systems. No single  
 504 source of systematic uncertainty dominates across  $p_T$  ranges or collision systems.

**Table 7:** Summary of the systematic uncertainties on the purity as measured with  $\sigma_{\text{long}}^2$  in p–Pb collisions. All values are in absolute percentage. “Stat.” refers to the statistical uncertainty; “Signal” refers to the signal template uncertainty; “Anti-iso” refers to the uncertainty due to the sideband selection; “Bkg” refers to the uncertainty due to the background template correction; “Total” is the sum of the previous three columns in quadrature.

$p_T$ (GeV/ $c$ )	Purity	Stat.	Signal	Anti-iso	Bkg	Total syst
12.0–15.0	20.7	1.1	1.1	0.8	1.5	2.0
15.0–20.0	34.2	1.2	2.0	1.6	1.2	2.8
20.0–25.0	47.6	1.7	1.9	1.1	1.7	2.7
25.0–40.0	54.6	1.8	2.3	2.4	2.1	3.9

**Table 8:** Summary of the systematic uncertainties on the purity as measured with  $\sigma_{\text{long}}^2$  in pp collisions. All values are in absolute percentage. “Stat.” refers to the statistical uncertainty; “Signal” refers to the signal template uncertainty; “Anti-iso” refers to the uncertainty due to the sideband selection; “Bkg” refers to the uncertainty due to the background template correction; “Total” is the sum of the previous three columns in quadrature.

$p_T$ (GeV/ $c$ )	Purity	Stat.	Signal	Anti-iso	Bkg	Total syst
12.0–15.0	20.1	1.7	2.0	1.2	2.9	3.7
15.0–20.0	31.7	2.0	2.5	1.5	2.4	3.8
20.0–25.0	47.3	2.9	0.8	3.0	2.8	4.2
25.0–40.0	48.5	3.5	5.9	4.0	3.4	7.9

## 505 11 Tracking

506 This section describes the performance of ITS-only tracking. During the 13def and 17q periods, the TPC  
 507 was either not read out or compromised due to space-charge distortions so we rely on only the ITS for  
 508 track reconstruction. Thus the purpose of this section is to demonstrate that we can use ITS-only tracking  
 509 to measure tracks up to relatively large track  $p_T$ .

510 Our strategy is to measure the charged-particle  $p_T$ -spectrum using ITS-only tracking and to compare it  
 511 with the normal TPC+ITS tracking in the data-taking period when the TPC was active and free of space-  
 512 charge distortions (i.e. the low-luminosity 13b data-taking period of 5 TeV p–Pb minimum-bias data)  
 513 and with published ALICE measurements [12] that use the same dataset. By comparing to published  
 514 measurements in the same system and center-of-mass energy, we avoid any ambiguity related to the  
 515 reference spectrum.

516 Our aim is to validate the combined effect of tracking efficiency, fake rate, and track momentum smearing  
 517 corrections that are based on MC simulations. We also rely on this study to estimate systematic  
 518 uncertainties due to mis-modeling of the tracking performance. So, in effect, we are performing a cross-  
 519 calibration or bootstrapping with the standard ALICE tracking.

520 For this study, we use minimum bias p–Pb events for both data and Monte Carlo. The Monte Carlo  
 521 simulations used for this section are LHC13b2\_efix\_p1, which a DPMJET anchored to LHC13b,c and

- 522 LHC17l3b, which is a pp PYTHIA8, anchored to LHC17p. The data used runs from LHC13b and 17p.  
 523 The p–Pb data sets match the ones used in the published ALICE measurements mentioned previously.
- 524 For this tracking performance analysis, we only use events with the minimum bias trigger (CINT7). We  
 525 also apply the vertex and pileup selections described in Section 6. The tracks reconstructed from the ITS  
 526 (“ITS-only tracks”), are compared with tracks reconstructed from information obtained from both the  
 527 TPC and ITS (“TPC+ITS tracks”). Here, ITS-only tracks are really reconstructed in a standalone way  
 528 and are not simply the ITS-segment of a ITS+TPC track.<sup>12</sup>
- 529 In order to select good-quality tracks emerging from the primary vertex while maintaining a high ef-  
 530 ficiency, each track is required to satisfy the cuts summarized in Table 9. A set of standard PWG-JE  
 531 cuts are applied to all tracks, and additional track cuts are applied depending on whether the track is a  
 TPC+ITS track or an ITS-only track.

**Table 9:** Summary of the cuts used in Track Selection.

<b>Common Cuts</b>	
track $\eta$	$ \eta  < 0.8$
track $p_T$	$p_T \geq 0.150 \text{ GeV}/c$
SetMaxDCAToVertexXY	2.4 cm
SetMaxDCAToVertexZ	3.2 cm
SetDCAToVertex2D	TRUE
<b>TPC+ITS Cuts</b>	
SetMinNClustersTPCPtDep	70.+30./20.*x, 20.0
SetMinNClustersTPC	70
SetMaxChi2PerClusterTPC	4
SetMaxChi2PerClusterITS	36
SetMaxFractionSharedTPCClusters	0.4
SetMaxChi2TPCConstrainedGlobal	36
SetRequireTPCStandAlone	TRUE
SetRequireTPCRefit	TRUE
SetRequireITSRefit	TRUE
SetRequireSigmaToVertex	FALSE
SetAcceptKinkDaughters	FALSE
<b>ITS-Only Cuts</b>	
SetRequireITSPureStandAlone	TRUE
SetMinNClustersITS	4
SetMaxChi2PerClusterITS	3

- 532
- 533 **11.1 Tracking performance with and without TPC**
- 534 Our strategy to validate the use of ITS-only tracking relies on cross-calibration with the normal TPC+ITS  
 535 tracking and comparisons with published data. To this end, we first measure a charged-particle  $p_T$  spec-  
 536 trum using TPC+ITS reconstruction to reproduce the published data and thus demonstrate the validity of  
 537 our method. We then repeat the measurement with ITS-only reconstruction and compare the results.
- 538 The method we use contains the following steps:

- 539 1. Using the MC simulations, obtain the efficiency, the fake rate, and a track- $p_T$  response matrix.

---

<sup>12</sup>The standalone ITS-only tracks can be retrieved from ESD data by demanding a track status of “ALI`ESD-TRACK::KITSPURESA`”. The exact line in the code is as follows: `_track_cut.back().SetRequireITSPureStandAlone(kTRUE);`.

- 540     2. Obtain a  $p_T$  spectrum from the minimum-bias data with either ITS-only or TPC+ITS tracking.  
 541     3. Apply a  $p_T$ -dependent fake rate correction to the spectra of point 2.  
 542     4. Multiply the published spectrum by the track efficiency.  
 543     5. Finally, use the response matrix to smear the result of point 4.  
 544     6. Compare the result of point 5 with the result of point 3.

545 If the efficiency, fake rate, and response matrix, which are all obtained from MC simulations, faithfully  
 546 represent the tracking performance of the data, then it follows that the result of point 3 should be com-  
 547 patible with the result of point 5. We use this fact to constrain the degree to which mis-modeling of the  
 548 detector response could affect single-track  $p_T$  spectra.

549 Note that we could have attempted to directly measure a charged-particle spectrum, applying an effi-  
 550 ciency correction to the result of point 3 and then unfolding that result. That would have been directly  
 551 comparable to the published data. However, the systematic uncertainties intrinsic to the unfolding pro-  
 552 cedure would result in relatively large systematic uncertainties, but these are unrelated to the issue we  
 553 discuss here.

554 We note that using the response matrix to smear the published data is a mathematically well-defined  
 555 procedure (as opposed to unfolding, which requires an inversion of the response matrix subject to the  
 556 statistical uncertainties of the data). Given that our aim is to test the corrections and response matrix in a  
 557 robust and unambiguous way, we follow the above steps to avoid unfolding systematic uncertainties.

### 558 **11.1.1 Efficiency and fake rates**

559 The tracking efficiency is defined as the ratio of the number of reconstructed primary particles<sup>13</sup>,  $N_{\text{prim,rec}}(p_T)$ ,  
 560 to the number of generated primary particles,  $N_{\text{prim,gen}}$ . The truth-to-reconstructed matching is done fol-  
 561 lowing the standard ALICE method<sup>14</sup>. Thus, the efficiency is defined as:

$$\epsilon(p_T^{\text{true}}) = \frac{N_{\text{prim,rec}}(p_T^{\text{true}})}{N_{\text{prim,gen}}(p_T^{\text{true}})}. \quad (9)$$

562 As the equation suggests, we use the simulated or “truth” transverse momentum,  $p_T^{\text{true}}$ , for both the  
 563 numerator and denominator in our efficiency. This way, we factorize out effects of efficiency from bin-  
 564 migration due to momentum smearing.

565 The numerator of Equation 9 is restricted for charged particles with generated pseudorapidity in the range  
 566  $|\eta^{\text{true}}| < 0.8$  and azimuth  $0 < \phi^{\text{true}} < 2\pi$ . Therefore, the correction factor accounts for both geometrical  
 567 acceptance, detector inefficiencies, and dead channels.

568 Fake tracks are defined as reconstructed tracks that do not match to a truth, primary particle. The fake  
 569 rate is determined by taking a ratio of the number of fake tracks to the total reconstructed tracks and is  
 570 parametrized as a function of the reconstructed transverse momentum of the track,  $p_T^{\text{reco}}$ :

$$\text{fakerate}(p_T^{\text{reco}}) = \frac{N_{\text{unmatched}}(p_T^{\text{reco}})}{N_{\text{allreco}}(p_T^{\text{reco}})}. \quad (10)$$

571 Figure 15 shows the efficiency and the fake rates for the TPC+ITS and ITS only tracks. In both cases  
 572 the efficiency grows with  $p_T$  up to about 1 GeV/ $c$  where it dips and it reaches a plateau value with no

---

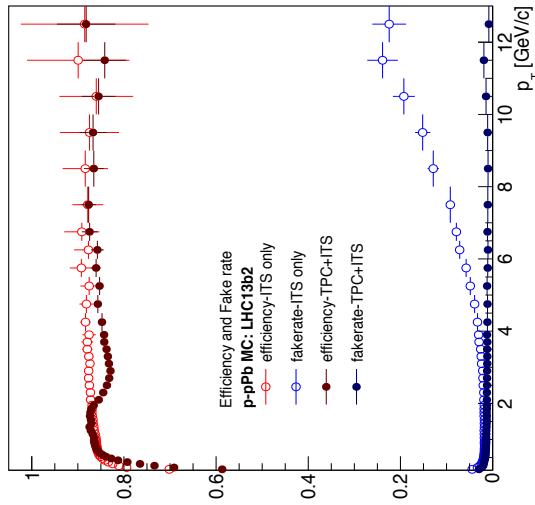
<sup>13</sup>No special tuning of the particle type composition is performed. This typically only matters at low  $p_T$  and enforces a small (percent level) correction to the out-of-the-box results.

<sup>14</sup>The information is retrieved with the method “track→GetLabel()”.

573 significant  $p_T$  dependence. The efficiency starts at about 57% for the TPC+ITS tracks and at 70% for  
 574 ITS-only tracks at 150 MeV and plateaus at 84% and 88% respectively.

575 The lower efficiency for the TPC+ITS tracks compared to ITS-only tracks is expected since the former  
 576 requires a matching between ITS and TPC track segments, which has some inefficiency. This study  
 577 shows that the matching efficiency is high at large  $p_T$  but leads to substantial differences at low  $p_T$ .

578 The fake rate for the TPC+ITS tracks is less than a percent over the entire range shown. In comparison,  
 579 the fake rate is larger in the ITS-only tracks. It is below 2% below 4 GeV/ $c$ , but it grows roughly linearly  
 580 and reaches 20% at 10 GeV/ $c$ . The much higher fake rate is due to the much lower number of clusters  
 associated with ITS-only tracks (maximum of 6) than to TPC+ITS tracks (minimum of 70).



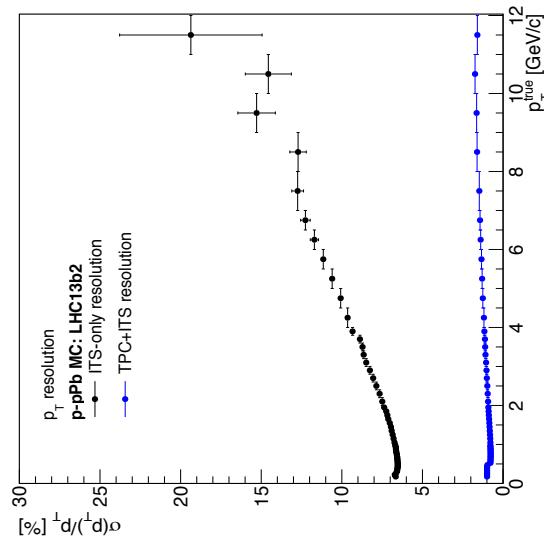
581  
**Fig. 15:** Efficiency and fake rate for combined TPC+ITS tracking (filled circles) and ITS-only tracking (open circles) obtained with p-Pb simulation. The error bars represent statistical uncertainty only.

582    **11.1.2 Resolution, response matrix and bin migration**  
 583 The main difference between TPC+ITS and ITS-only tracking lies in the much worse momentum reso-  
 584 lution of the latter. This is driven by geometry and  $\int B dl$  as the TPC covers up to  $z = 258$  cm but the ITS  
 585 only to  $z = 48$  cm.

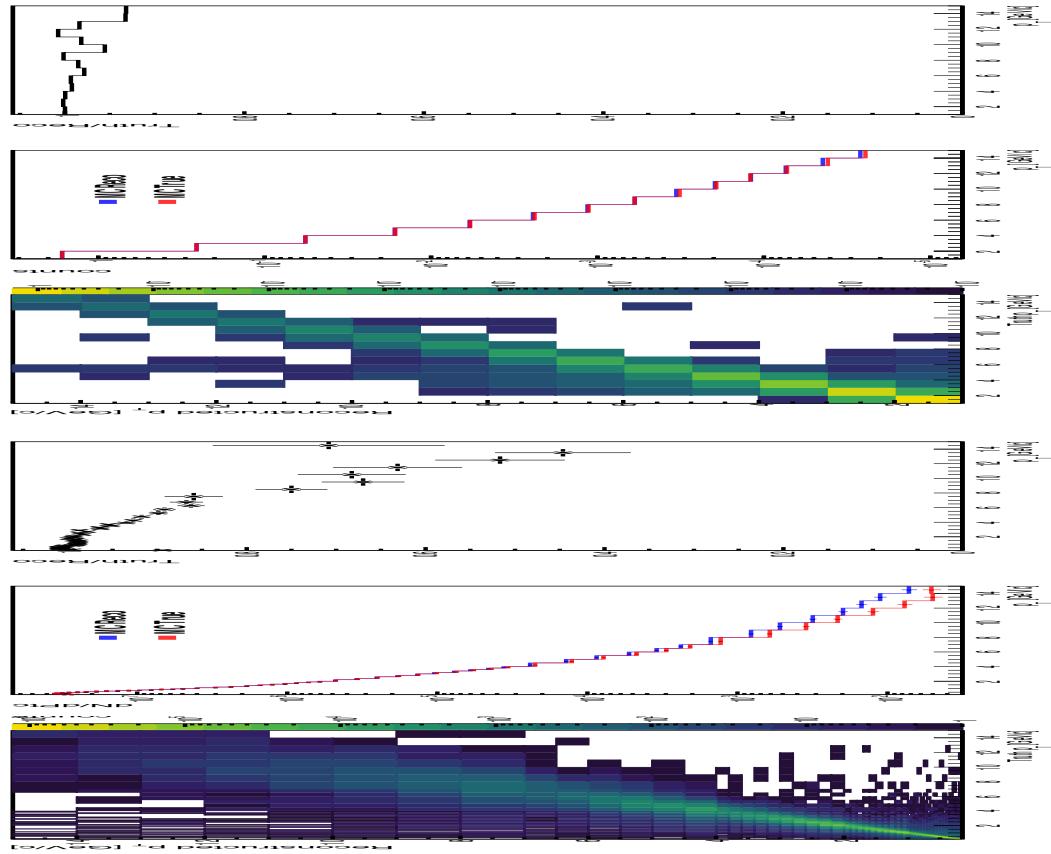
586 Figure 16 shows the momentum resolution as a function of  $p_T^{\text{true}}$  for TPC+ITS and ITS-only tracks. The  
 587 momentum resolution of both increases with  $p_T$ ; however, the resolution for TPC+ITS never exceeds a  
 588 relative 2% below 20 GeV/ $c$ , while the ITS-only tracking resolution is about a factor of 7 worse and  
 589 reaches  $\sim 15\%$  by 10 GeV/ $c$ . In both cases the resolution curves have the expected shape: the growth at  
 590 low momentum is due to multiple-scattering and the linear growth at higher  $p_T$  arises from the number  
 591 and depth of position measurements, and the track bend at the measurement planes.

592 We define the tracking response matrix as the correlation between the reconstructed and the generated  
 593 transverse momentum. This matrix is filled only for reconstructed tracks with a true match; fake tracks  
 594 are explicitly excluded. Figure 17 shows the response matrix, its one-dimensional projections, and the  
 595 ratio of true to reconstructed spectra. The ratio of the true to reconstructed spectra is used to correct of  
 596 the bin migration effects:

$$\text{bin migration}(p_T^{\text{reco}}) = \frac{N_{\text{prim,reco}}(p_T^{\text{true}})}{N_{\text{prim,reco}}(p_T^{\text{reco}})}. \quad (11)$$



**Fig. 16:** The relative  $p_T$  resolution for TPC+ITS tracking and ITS only tracking. The error bar represents statistical uncertainty only.



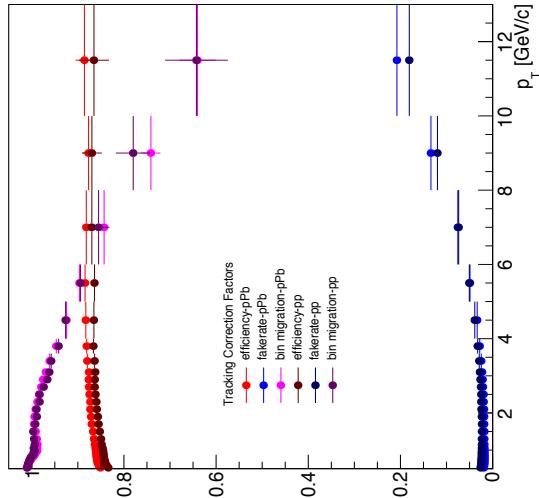
**Fig. 17:** Left panel: correlation matrix between true  $p_T$  and reconstructed  $p_T$  of tracks reconstructed with TPC+ITS (upper row) and ITS-only (bottom row). Middle panel: projections of the response matrix into the true and reconstructed  $p_T$ . Right panel: ratio of true to reconstructed spectra. This ratio used as part of the bin-by-bin correction factors.

597 The effect of the track momentum smearing in the ITS is clearly visible in the projection plots, where

598 the reconstructed spectrum is significantly harder at high  $p_T$ . The ratio of truth to reconstructed  $p_T$  is  
 599 very close to unity in the TPC+ITS case, as expected. On the other hand, the ratio deviates significantly  
 600 from unity in the ITS-only case; it reaches 0.9 at 6 GeV/c and drops quickly, reaching 0.5 at about 13  
 601 GeV/c. The quick drop at high  $p_T$  comes mainly from the linear degradation of the relative momentum  
 602 resolution combined with the fast drop of the true spectrum to produce large effects in the tails.

603 The bin migration (b), along with the tracking efficiency ( $\epsilon$ ) and the fake rate (f) are used as the correction  
 604 factor equation 12 for the charged hadron tracks; all of them are shown together in Figure 18. Based on  
 605 overlapping plot of pp and p–Pb, we can see that correction factors are similar for tracks from a pp  
 606 collision or a p–Pb collision. We conclude that the multiplicity in p–Pb is low enough such that it does  
 607 not affect tracking performance.

$$w = \frac{1}{\epsilon}(1 - f)b \quad (12)$$



**Fig. 18:** The efficiency, fake rate, and momentum smearing correction factors for pp and p–Pb data.

## 608 11.2 Comparison with published data

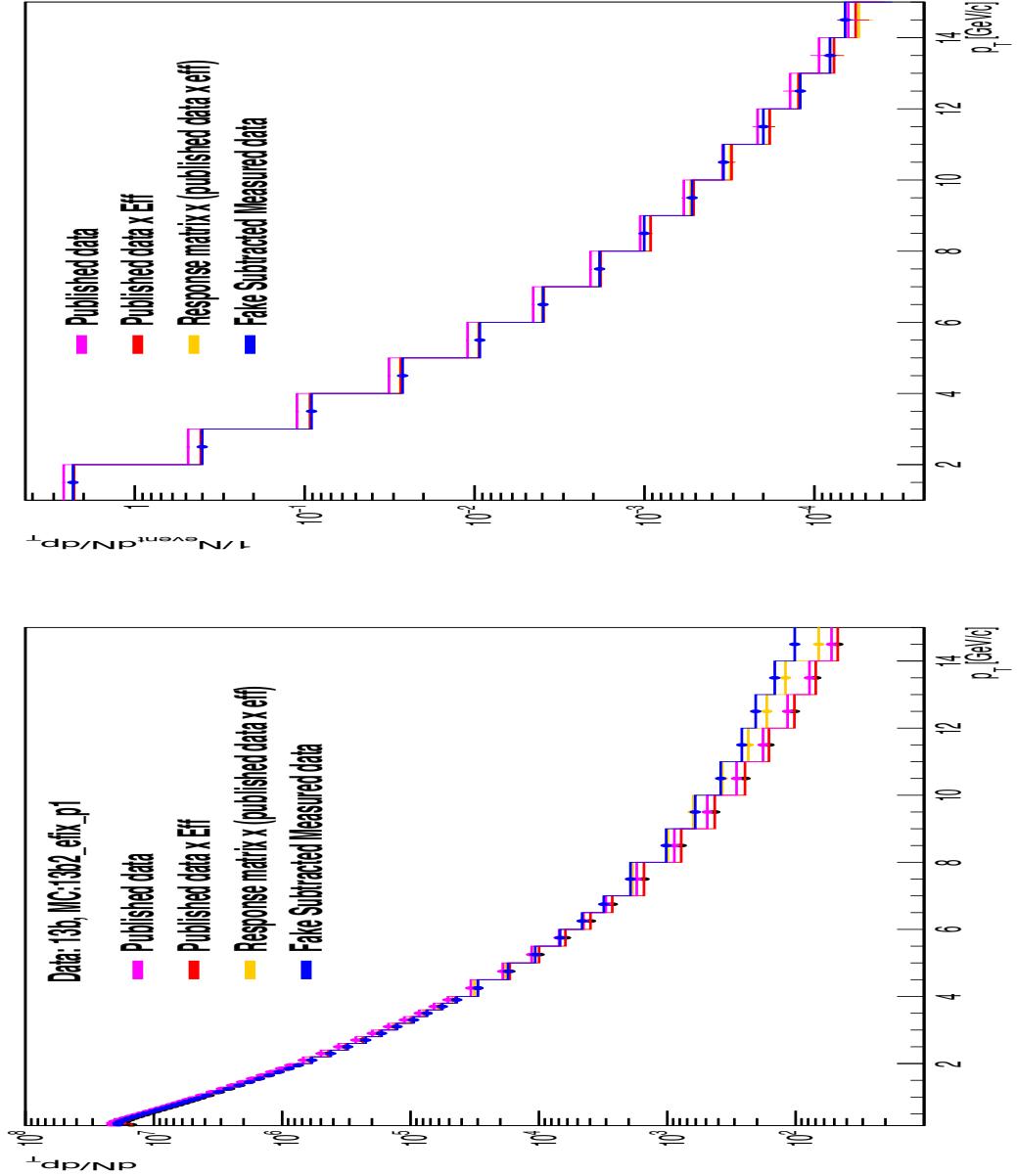
609 In this section we show the result of the closure test, which is described in Section 11.1, that aims to  
 610 validate the MC simulation description of track efficiency, fake rate and momentum smearing corrections  
 611 using the published data.

612 Figure 19 shows the correction procedure on published data at various stages: the pink line is the pub-  
 613 lished charged-particle spectrum in p–Pb collisions at  $\sqrt{s_{NN}} = 5$  TeV from Ref. [12]; the red line is the  
 614 published data spectra after multiplying by the efficiency; the orange line is the red spectrum multiplied  
 615 by the response matrix according to Equation 13:

$$R_j = \sum M_{ij} \cdot P_i \quad (13)$$

616 where  $M$  is the response matrix,  $P$  is the published data times the efficiency,  $R$  is smeared  $p_T$  spectrum.  
 617 The smeared  $p_T$  spectrum should be consistent with the measured spectrum after removal of the fake-  
 618 track contribution.

We perform a “folding” of the published spectra instead of an “unfolding” of the measured spectra because the former is a unique transformation and avoids the need for systematic studies on the stability of the unfolding procedure. This allows us to focus on unambiguously testing the response matrix.



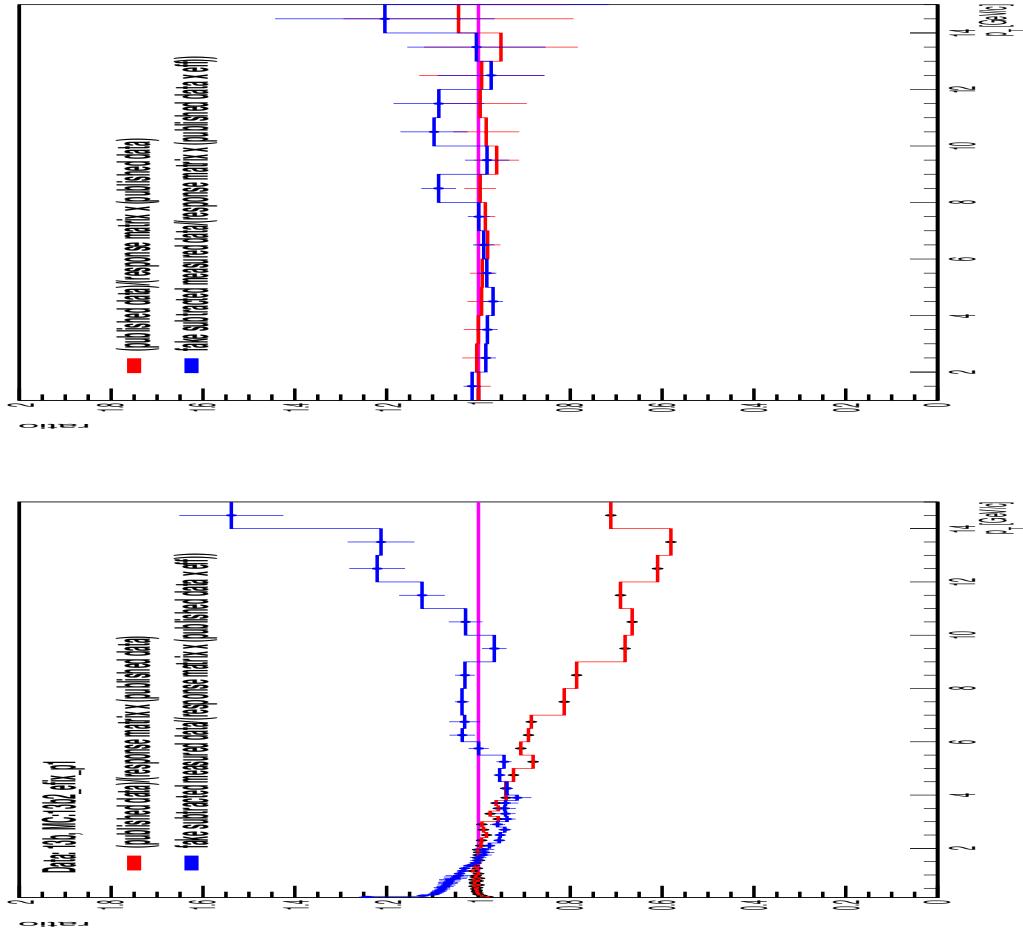
**Fig. 19:** Smearing of published data at various stages, using TPC+ITS (top) and ITS only (bottom) response matrices. The pink is the published data  $p_T$  spectrum. The red is after the pink has been multiplied by the efficiency. The orange is the spectra after applying the response matrix in order to induce the smearing on the red. The blue is fake-rate-subtracted measured data.

- The published data has a total uncertainty (quadrature sum of statistical and systematic uncertainties) that ranges from 1.8% at 1 GeV/ $c$ , reaches 4.8% by 10 GeV/ $c$  and grows quickly to about 20% at 15 GeV/ $c$ , where it is dominated by the statistical uncertainty.
- Figure 20 shows the ratio of the measured fake-subtracted spectrum and the smeared published spectra. The ratio of the published data to the smeared published data is shown to illustrate the impact of the momentum smearing, which is less than a 2% effect for TPC+ITS tracks but it reaches up to a factor

of two in the ITS-only case. The closure-test ratios from TPC+ITS tracking are consistent with unity within uncertainties, which is expected. The more interesting result is that the ratios due to ITS-only tracking are within  $\pm 5\%$  of unity in the range between  $0.5$ – $10$  GeV/ $c$ , which is the range we use for our  $\gamma^{\text{iso}}$ -hadron analysis.

The statistically-significant deviation from the published data with ITS-only tracking at high  $p_{\text{T}}$  (blue curve in lower panel of Figure 20), could be due to several reasons including improper modeling of a rapid deterioration of the momentum resolution and underestimation of fake rate. Further work would be needed to understand and correct these and other effects at high  $p_{\text{T}}$ , but that lies beyond the scope of this work.

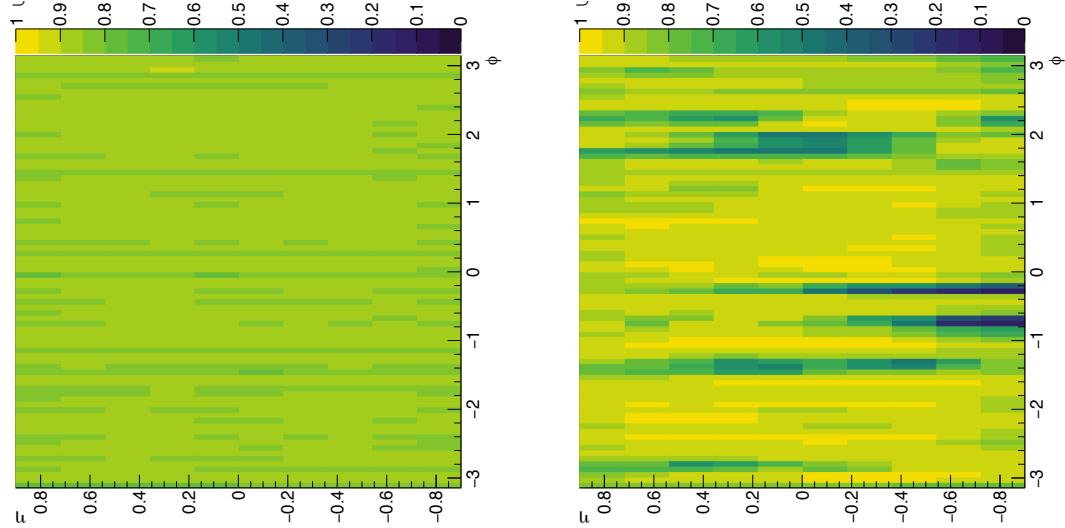
We note that these spectra are normalized per-minimum bias event and not by integral, so the fact that the ratios are close to unity reflect the fact that the efficiency calculations shown in Figure 15 are an accurate description of the detector response. Based on the results shown in this section, we restrict  $p_{\text{T}}$  range to no more than  $10$  GeV/ $c$ , which is beyond the limit of the statistical power of our  $\gamma^{\text{iso}}$ -hadron analysis.



**Fig. 20:** Result of closure test comparing measured data and published data, for TPC+ITS tracking (top) and ITS-only tracking (bottom). The red curves show the ratio of the reference spectra to the smeared reference spectra. The blue curves show the ratio of the fake-subtracted measured data and the smeared reference spectra. Ideally the blue curve would be flat at unity. The error bar represents statistical uncertainty only for the blue curve, and the quadrature sum of statistical and systematic uncertainties for the red curve.

641 **11.2.1 Angular dependence of tracking efficiency**

642 Here, we demonstrate that the  $\varphi$ -dependence of holes in the ITS-only tracking is well reproduced in the  
 643 MC simulation. The 2D  $\varphi$ - $\eta$  efficiency is calculated in a similar way to the  $p_T$  efficiency described in  
 644 Equation 9, but instead of being functions of  $p_T^{\text{true}}$ , the efficiency is a function of the true track azimuthal  
 645 angle and pseudorapidity,  $\varphi^{\text{true}}$  and  $\eta^{\text{true}}$  respectively. Only tracks with  $p_T^{\text{true}} > 1 \text{ GeV}/c$  are considered to  
 646 avoid strong effects of bending due to the magnetic field that would obscure the impact of dead regions.



**Fig. 21:** Tracking efficiency as a function of  $\varphi^{\text{true}}$  and  $\eta^{\text{true}}$  for TPC+ITS (left) and ITS-only (right) tracks.

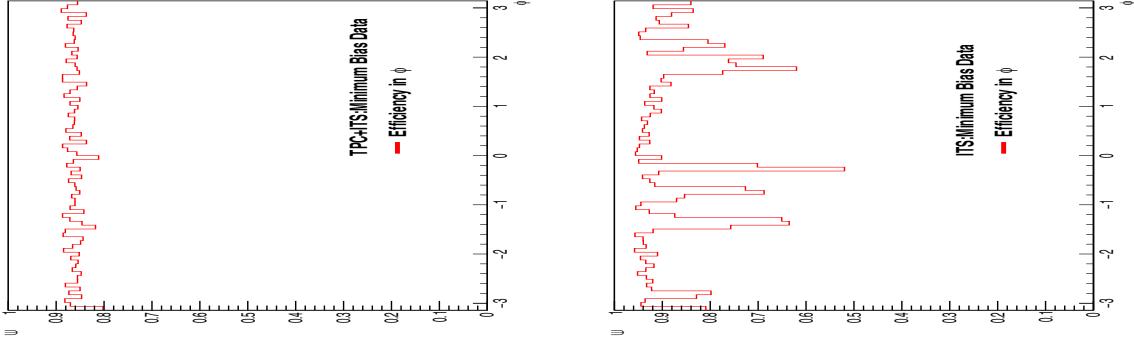
647 Figure 21 shows the resulting efficiency for TPC+ITS and ITS-only tracks. While the TPC+ITS 2D  $\varphi$ - $\eta$   
 648 distribution looks uniform, this is not the case for the ITS-only distribution, which has visible dips in the  
 649 efficiency at various  $\varphi$ . The efficiency is close to unity for most of the phase space covered. There are  
 650 no big  $\eta$  variations, but there are large  $\varphi$  variations. The efficiency holes at  $\varphi = -0.8$  and  $-0.2$  are very  
 651 visible and reach values close to zero. These are attributed to ITS-staves that are completely dead.

652 **11.2.2 Validation of  $\varphi$ -dependence of efficiency**

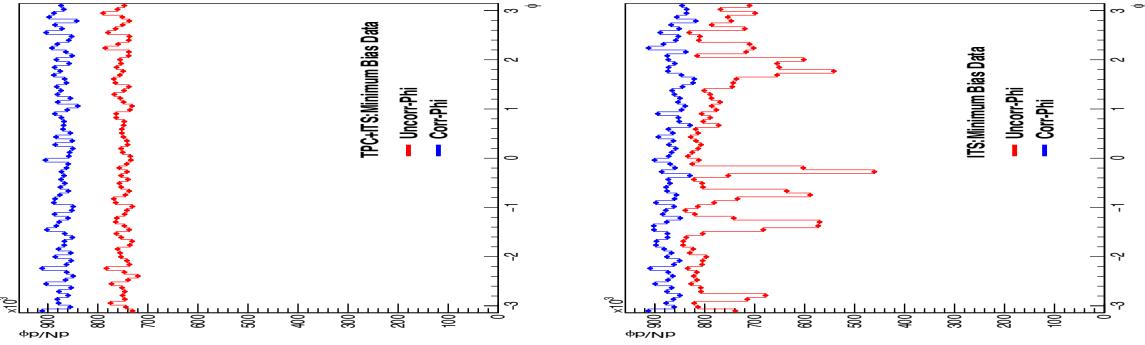
653 In order to validate the description of the ITS-only tracking holes, we do a test with minimum-bias data.  
 654 For this, we rely on the fact that apart from the effect of  $\varphi$ -dependent holes, the track azimuthal angle  
 655 distribution is expected to be uniform in minimum-bias data. So we measure the track  $\varphi$  spectrum, then  
 656 correct for the efficiency and check whether the distribution is flat. The level of flatness gives us a sense  
 657 of the systematic uncertainties associated with mis-modeling of the  $\varphi$ -dependent efficiency.

658 Figure 22 shows the  $\varphi$ -dependence of the tracking efficiency for TPC+ITS and ITS-only tracks. The  
 659 efficiency is calculated using Equation 9, but as a function of  $\varphi^{\text{true}}$  instead of  $p_T^{\text{true}}$ . The TPC+ITS  
 660 tracking efficiency is flat in  $\varphi$  as expected, but there are dips in the efficiency for the ITS-only tracking  
 661 due to dead staves in the ITS. These have little impact on the TPC+ITS tracks because the selection does  
 662 not have the strict requirement on the number of ITS hits,  $N_{\text{ITS}} \geq 4$ , that is applied for ITS-only tracks.

663 Figure 23 shows the  $\varphi$  distribution of TPC+ITS and ITS-only tracks in minimum-bias p–Pb data and  
 664 the effect of applying the efficiency correction to the  $\varphi$  distribution. Before applying the efficiency  
 665 correction, there are visible holes at  $\varphi = -1.04, -0.8, -0.2$  rad in ITS-only tracks which are not present  
 666 in the TPC+ITS tracking. After applying the  $\varphi$  efficiency, the holes are corrected, and we obtain a  
 667 distribution which is flat within  $\pm 2.5\%$ . The TPC+ITS remains flat after the efficiency correction, as  
 668 expected. This shows that the description of dead channels in the ITS is well-described in the simulation.



**Fig. 22:** Tracking efficiency as a function of  $\phi^{\text{true}}$  for TPC+ITS tracks (left) and ITS-only tracks (right). In both cases, the efficiency is calculated for tracks with  $p_T^{\text{true}} > 1 \text{ GeV}/c$  using the LHC13b2\_efix\_p1 Monte Carlo simulation.



**Fig. 23:** Left panel: track  $\phi$  distribution measured in data for tracks with  $|\eta| < 0.8$  before (red) and after (blue) applying the efficiency correction for TPC+ITS tracks. Right panel: track  $\phi$  distribution measured in data for tracks with  $|\eta| < 0.8$  before (red) and after (blue) applying the efficiency correction for ITS-only tracks.

### 669 11.3 Summary of the ITS-only tracking performance studies

670 In this section we summarize the findings of our studies on ITS-only tracking performance:

671 1. Tracking Efficiency:

672 The ITS-only tracking efficiency is 75% at 150 MeV/c and grows to 85% at 1 GeV/c and above.

673 2. Fake rate:

674 The fake rate of ITS-only tracking is about 10 times worse than for TPC+ITS tracks, but still less  
675 than 20% below 10 GeV/c, which is the relevant range for the analyses presented in this note.

676 3. Momentum resolution:

677 The momentum smearing effects are significant for ITS-only tracking. The bin-to-bin correction  
678 factor due to smearing effects for ITS-only tracking is 0.7 at 10 GeV/c and 0.5 at 15 GeV/c. The  
679 smearing effects for ITS+TPC tracking are negligible.

680 4. Description of  $\phi$  holes

681 The efficiency as a function of  $\phi$  shows inhomogeneity not present in the TPC+ITS tracking that  
682 are attributed to dead staves in the ITS. These are concentrated in specific  $\eta$  and  $\phi$  regions. These  
683 are well described in the simulation.

We have validated the MC corrections for tracking efficiency, fake rate and momentum smearing by comparing with published data. From that study, we estimate a combined systematic uncertainty on tracking performance to be a relative  $\pm 5\%$  for ITS-only tracks with  $0.5 < p_T < 10 \text{ GeV}/c$ .

## 12 Isolated-photon hadron correlations

In this section we present the method we use to extract  $\gamma^{\text{iso}}$ -hadron azimuthal correlations. We select tracks with  $|\eta| < 0.8$  and  $0.5 < p_T^{\text{track}} < 10 \text{ GeV}/c$ , following the studies shown in Section 11. We only consider cluster-track pairs within  $|\Delta\eta| < 1.2$ . Our cluster  $p_T$  range selection is  $12 < p_T^{\text{cluster}} < 40 \text{ GeV}/c$ , with isolation criteria of  $\text{ISO} < 1.5 \text{ GeV}/c$ . We apply the same track selection criteria described in Section 11, Table 9. We present our results as a function of  $z_T$ , which is defined as:

$$z_T = \frac{p_T^{\text{track}}}{p_T^{\text{cluster}}} \quad (14)$$

Section 12.1 shows the way we correct for the impurity of our  $\gamma^{\text{iso}}$  selection; Section 12.2 describes the pair-acceptance correction obtained with the event-mixing method.

### 12.1 $\gamma^{\text{decay}}$ background Subtraction

We split the clusters that pass our  $\text{ISO} < 1.5 \text{ GeV}$  selection in two regions of interest, our signal region  $SR$  with the selection  $\sigma_{\text{long}}^2 < 0.3$ , and our background region,  $BR$  as  $\sigma_{\text{long}}^2 > 0.4$

We denote the number of clusters in the background region as  $T_{BR}$  and the number of clusters in the signal region as  $T_{SR}$ . We can write  $T_{SR}$  as:

$$T_{SR} = T_S + T_B. \quad (15)$$

Here we define  $T_S$  as our signal (i.e. isolated prompt photons) and  $T_B$  as the background (decay photons that pass our  $\gamma^{\text{iso}}$  selection).

The number of cluster-track pairs per trigger,  $P$  is:

$$C_{SR} = \frac{1}{T_{SR}} P_{SR}. \quad (16)$$

$$C_{BR} = \frac{1}{T_{BR}} P_{BR}. \quad (17)$$

These quantities are measured directly. We want to separate the correlation function of signal from the correlation function of background as, which are defined as:

$$C_S = \frac{1}{T_S} P_S \quad (18)$$

$$C_B = \frac{1}{T_B} P_B \quad (19)$$

To measure the true signal correlation,  $C_S$ , we decompose the correlation measured in the signal region:

$$C_{SR} = \frac{1}{T_{SR}} P_{SR} = \frac{1}{T_{SR}} (P_S + P_B) \quad (20)$$

$$C_{SR} = \frac{1}{T_{SR}} (T_S C_S + T_B C_B) \quad (21)$$

$$C_{SR} = p C_S + (1 - p) C_B \quad (22)$$

706 where the purity of our  $\gamma^{\text{iso}}$  definition, called  $p$ , is equal to  $T_S/T_{\text{SR}}$  by definition and is measured with the  
 707 template fit method as described in Section 10.

708 Solving for  $C_S$  in Equation 22 yields:

$$C_S = \frac{C_{\text{SR}} - (1-p)C_{\text{BR}}}{p}. \quad (23)$$

709 ,

710 where we have made the approximation that  $C_B$  (per-trigger pairs for decay photons that pass our  $\gamma^{\text{iso}}$   
 711 selection) can be approximated by the measured  $C_{\text{BR}}$  (per-trigger pairs for decay photons that pass our  
 712 isolation criteria but not our shower-shape selection). This is justified because the underlying physics  
 713 process that dictates the number of correlated pairs is independent from the opening angle of the neutral-  
 714 meson decay, which is what drives the shower-shape.

## 715 12.2 Pair-acceptance correction with event mixing

716 Event mixing is a data driven approach to correcting for detector acceptance effects and estimating com-  
 717 binatorial background. By constructing observables with particles from different events, we remove true  
 718 physics correlations from the correlation functions, isolating detector effects from limited acceptance in  
 719  $\eta$  and detector inhomogeneity in  $\eta$  and  $\varphi$ .

720 We mix events which are as similar as possible. To this end, events are often classified by multiplicity  
 721 (V0 amplitude, sum of V0A and V0C) and primary vertex  $z$ -position bins, and then mixed within these  
 722 bins. The mixing in this analysis was carried out using a Gale Shapely matching algorithm [17]. The use  
 723 of this algorithm avoids the need for binning in multiplicity and primary vertex.

724 Instead of bins, events are split into blocks of 2000. For each event in the block, the multiplicity and  
 725  $z$ -vertex position are compared between all other events in the block. The algorithm then creates a  
 726 preference list containing all other events in the block for each event based on how close events are  
 727 in multiplicity and  $z$ -vertex. In order for the mixed event distribution to reflect background instead of  
 728 emulating signal, as well as to fully cover the detector acceptance, we paired each  $\gamma$ -triggered event with  
 729 minimum bias events. Subsequent batches of 200 minimum bias events are used, as necessary, to reach  
 730 the desired number of mixed events per true event. For this analysis, depending on the  $z_T$  bin, each  
 731  $\gamma$ -triggered event was mixed with up to 300 minimum-bias events.

732 After each event has an associated preference list, the algorithm loops through all events in the block,  
 733 and then loops through each event's associated preference list to pair it to the first unpaired event on that  
 734 list. As the loop iterates, if an event towards the end of the main loop has an already-paired event high on  
 735 its preference list, the algorithm loops through the already-paired event's preference list and decides if  
 736 the paired event should stay paired to its current match, or switch to the new event. If the latter is chosen,  
 737 the previously matched event is unpaired and added back into the loop. A stable state is met when all  
 738 paired events have a match that is higher on their preference list than any remaining unpaired events in  
 739 the loop. Such a stable state is guaranteed to eventually be met according to Ref. [17].

740 The pseudo code below follows this description, using  $\gamma$  to denote a  $\gamma$ -triggered event, and  $M_B$  to denote  
 741 a minimum-bias event. The *unrequested* state refers to a  $M_B$  event on a  $\gamma$ -event's preference list that has  
 742 not yet been requested for pairing.

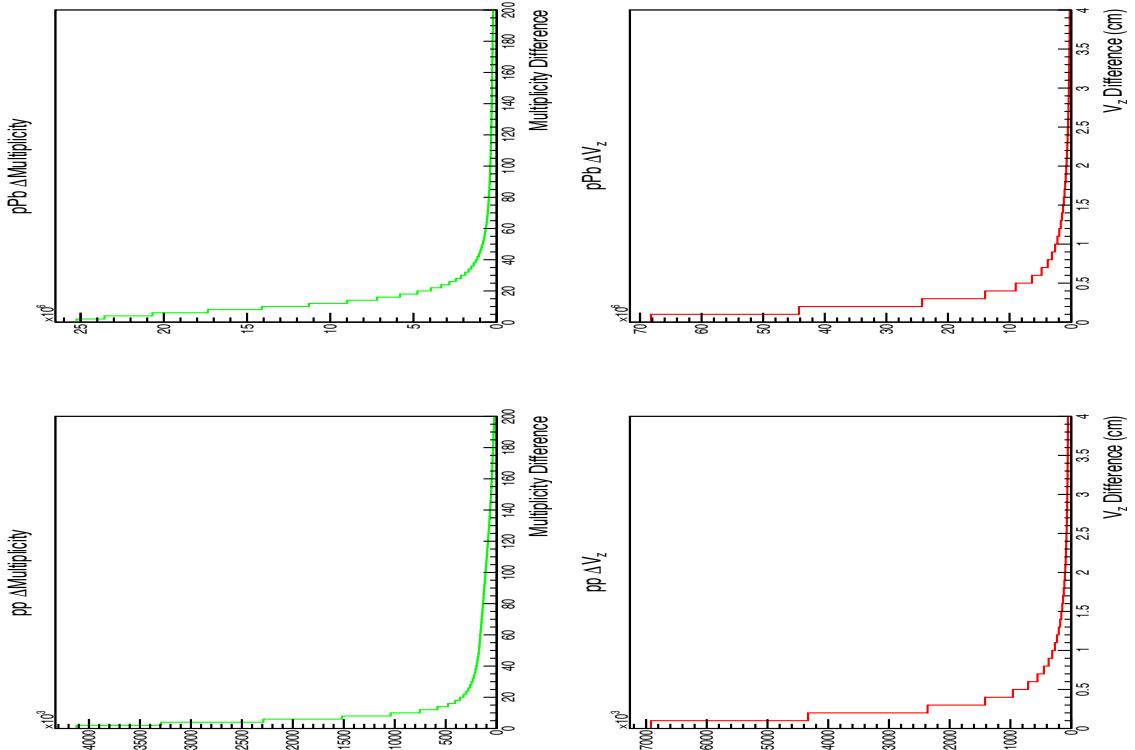
743

```
744 procedure GALESHAPLEYPAIRING
 745   while  $\exists \text{free } \gamma$  with an unrequested  $M_B$  on  $\gamma$ 's list do
 746      $M_B$  = first unrequested MinBias Event on  $\gamma$ 's list.
 747     if  $M_B$  is free then
 748       ( $\gamma, M_B$ ) become paired
```

```

749     else some pair ( $\gamma'$ , $M_B$ ) exists
750         if  $M_B$  prefers  $\gamma$  to  $\gamma'$  then
751              $\gamma'$  becomes free
752             ( $\gamma$ , $M_B$ ) become paired
753         else
754             ( $\gamma'$ , $M_B$ ) remain paired
755         end if
756     end if
757 end while
end procedure

```

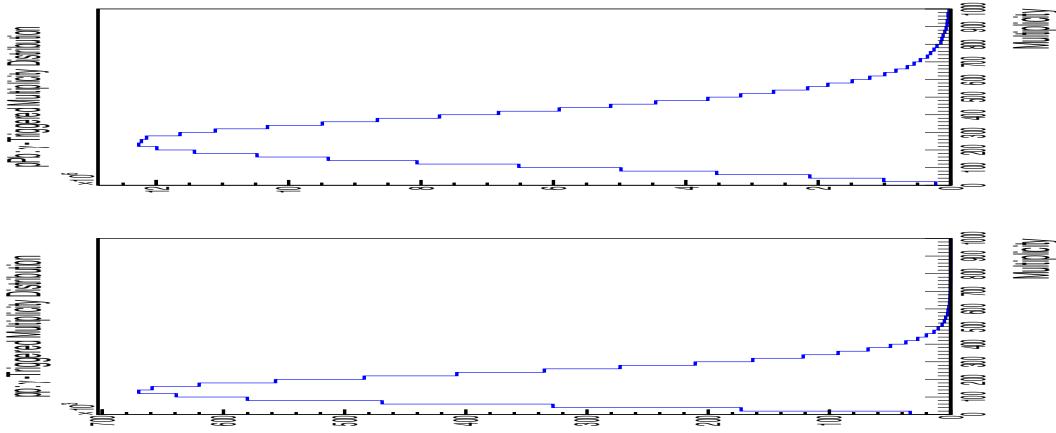


**Fig. 24:** Difference in V0 multiplicity (upper row) and longitudinal vertex position (bottom row) between paired events in pp (left column) and p–Pb right. The pairing algorithm results in sharp peak near zero for these difference distributions, particularly in the longitudinal vertex difference. As described in the text, in the correlation analysis we apply a further selection to cut the large tails observed in these distributions.

The difference distributions for Z-vertex and multiplicity between a  $\gamma$ -triggered and minimum bias events in p–Pb data are shown in Figure 24. The resulting distributions show a sharp peak that is below  $\Delta z < 0.5$  cm and also a long tail. Less than 6 % of the distribution lies beyond  $\Delta z > 2$  cm. The multiplicity difference, however, does not have as sharp a peak near  $\Delta \text{Multiplicity} = 0$ . About 20% of pairs have a multiplicity difference above 40, and cuts at  $\Delta V_z > 2\text{cm}$  and  $\Delta \text{Multiplicity} > 40$  were applied to pairs before calculating correlation functions.

Figure 25 shows the V0 multiplicity distributions for pp and p–Pb data in  $\gamma$ -triggered events. This shows that a multiplicity matching requirement of  $\Delta \text{Multiplicity} < 40$  is indeed very tight.

Ideally, the mixed event distribution should be flat in  $\Delta\phi$  and have a trapezoidal shape in  $\Delta\eta$ , because the limited acceptance in  $\eta$  increases the likelihood to reconstruct pairs with a small  $\Delta\eta$  (i.e, due to the convolution of two uniformly distributed functions). However, the use of ITS-only tracks and holes in



**Fig. 25:** V0 multiplicity distribution, i.e. the sum of V0A and V0C amplitudes , in pp (left) and p–Pb (right) gamma-triggered data.

770 the ITS acceptance result in deviations from a flat distribution in  $\Delta\phi$ .

771 The correlation function corrected by pair-acceptance effects is then given by:

$$C(\Delta\phi, \Delta\eta) = \frac{S(\Delta\phi, \Delta\eta)}{M(\Delta\phi, \Delta\eta)} \quad (24)$$

772 ,

773 where  $S(\Delta\phi, \Delta\eta)$  is the same-event correlation, and  $M(\Delta\phi, \Delta\eta)$  is the mixed-event correlation.  $S(\Delta\phi, \Delta\eta)$   
774 is given by:

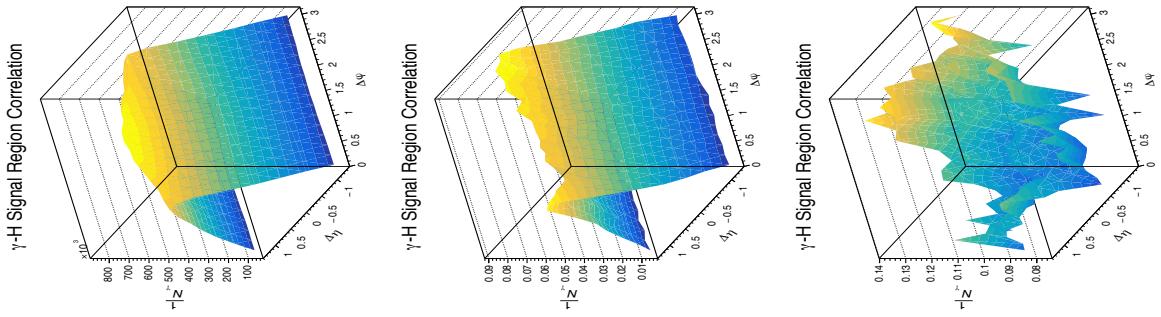
$$S(\Delta\phi, \Delta\eta) = \frac{1}{N_{\text{trig}}} \frac{d^2 N_{\text{same}}}{d\Delta\phi d\Delta\eta} \quad (25)$$

775 With  $N_{\text{trig}}$  as the number of trigger particles and  $N_{\text{same}}$  as the number of same event cluster-track pairs and  
776  $d^2 N_{\text{same}}/d\Delta\phi d\Delta\eta$  is found by pairing trigger particles with tracks from the same event. The mixed-event  
777 distribution,  $M(\Delta\phi, \Delta\eta)$ , is given by

$$M(\Delta\phi, \Delta\eta) = \alpha \frac{d^2 N_{\text{mixed}}}{d\Delta\phi d\Delta\eta}. \quad (26)$$

778 Where  $\alpha$  is the normalization constant that sets the maximum value of the mixed event correlation to  
779 1, and  $N_{\text{mixed}}$  is the number of mixed event cluster-track pairs. The term  $d^2 N_{\text{mixed}}/d\Delta\phi d\Delta\eta$  is obtained  
780 by pairing trigger particles from  $\gamma$ -triggered events with tracks from minimum bias events matched in  
781 z-vertex and multiplicity.

782 Same event correlation functions were divided by the mixed event correlation function within the same  
783  $z_T$  bins, shown for a single  $z_T$  bin in Figure 26. This procedure is carried out identically for clusters in  
784 the signal and background shower-shape regions.



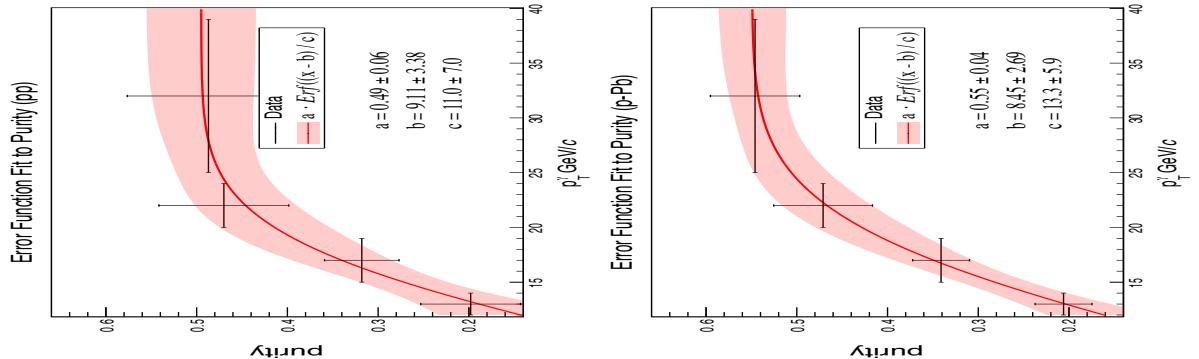
**Fig. 26:** **Left** Mixed Event correlation for a single  $z_T$  bin for gamma-triggered, signal region clusters and hadrons from minimum bias events. **Middle** 2D Correlation for signal region clusters and hadrons from the same events. **Right** Signal region correlation function corrected for detector acceptance effects.

### 785 12.3 Corrections applied to data

786 We apply a series of corrections to the measured correlation functions. We implement these by means of  
787 weights that are applied for each cluster-track pair. These are described in the following sections.

#### 788 12.3.1 Weighting for photon purity

789 Following Equation 23, we weight the background-region correlation  $C_{\text{BR}}$  by  $(1 - p)/p$  and the signal-  
790 region correlation  $C_{\text{SR}}$  by  $1/p$ . We apply this weight cluster-by-cluster. To avoid binning effects, we fit  
791 the purity with a 3-parameter error function, as shown in Figure 27.



**Fig. 27:** A 3-parameter error function is fit to the purity values measured in pp (left) and p–Pb (right) data. The width of the band represents the uncertainty on the fit.

#### 792 12.3.2 Weighting for track efficiency, fake rate, and bin migration

793 In order to correct for the tracking efficiency, fake rate, and bin-migration we apply a track-by-track  
794 weighting according to:

$$w_{\text{tracking}}(p_T^{\text{track}}) = \frac{1}{\epsilon} \times (1 - f) \times b. \quad (27)$$

795 Here  $\epsilon$  is the track efficiency,  $f$  is the fake rate, and  $b$  is the bin-to-bin migration correction. These are  
796 described in Section 11. The corrections are estimated independently for pp and p–Pb data although the  
797 performance is very similar.

---

**798 12.4 Corrected correlations**

799 The fully-corrected  $C_{\text{SR}}$  and  $C_{\text{BR}}$  correlations are shown in Figures 28 and 29. Our  $\gamma^{\text{iso}}$ -hadron cor-  
800 relations are the difference between the scaled- $C_{\text{SR}}$  and the scaled- $C_{\text{BR}}$ , which are shown in blue and  
801 red respectively. While the statistical precision of both  $C_{\text{SR}}$  and  $C_{\text{BR}}$  is high in all  $z_T$  bins and datasets,  
802 this gets diluted in the subtraction. That is, the low-purity leads to the subtraction of two comparable  
803 numbers, which results in a large statistical uncertainty.

**804 12.5 Underlying event subtraction**

805 This section describes the subtraction of underlying-event from the isolated photon-hadron correlations.  
806 We use a modified ZYAM procedure to estimate the uncorrelated background, where we take advantage  
807 of a feature of isolated photons-hadron correlations. The use of an isolated photon removes the near-side  
808 jet peak in the correlation function. We therefore us the  $\Delta\phi$  range of  $0.4 < \Delta\phi < 1.6$  to estimate the  
809 underlying event pedestal, and refer to this method as ZYAM moving forward.

810 As a check on the ZYAM procedure, we use the fact that the UE-estimation is independent of  $\Delta\eta$  and  
811 that that genuine correlations due to hard-scatterings decrease as  $\Delta\eta$  increases. To this end, we pick a  
812 region that is dominated by UE and extrapolate back to the region that would normally contains both  
813 UE and hard-scattering contribution. The UE is estimated by projecting the large  $\Delta\eta$  region defined as  
814  $0.8 < |\Delta\eta| < 1.4$  onto the  $\Delta\phi$  axis. To minimize bias from the isolation cut as well as the away side jet  
815 peak, the uncorrelated background is estimated from the projection in the region  $0.4 < \Delta\phi < 1.4$ . This  
816  $\Delta\phi\Delta\eta$  region is illustrated in Figure 30.

817 The statistical uncertainty in the UE estimate method is taken as a systematic uncertainty for  $\Delta\phi$  corre-  
818 lations as it is completely correlated bin-to-bin in  $\Delta\phi$ .

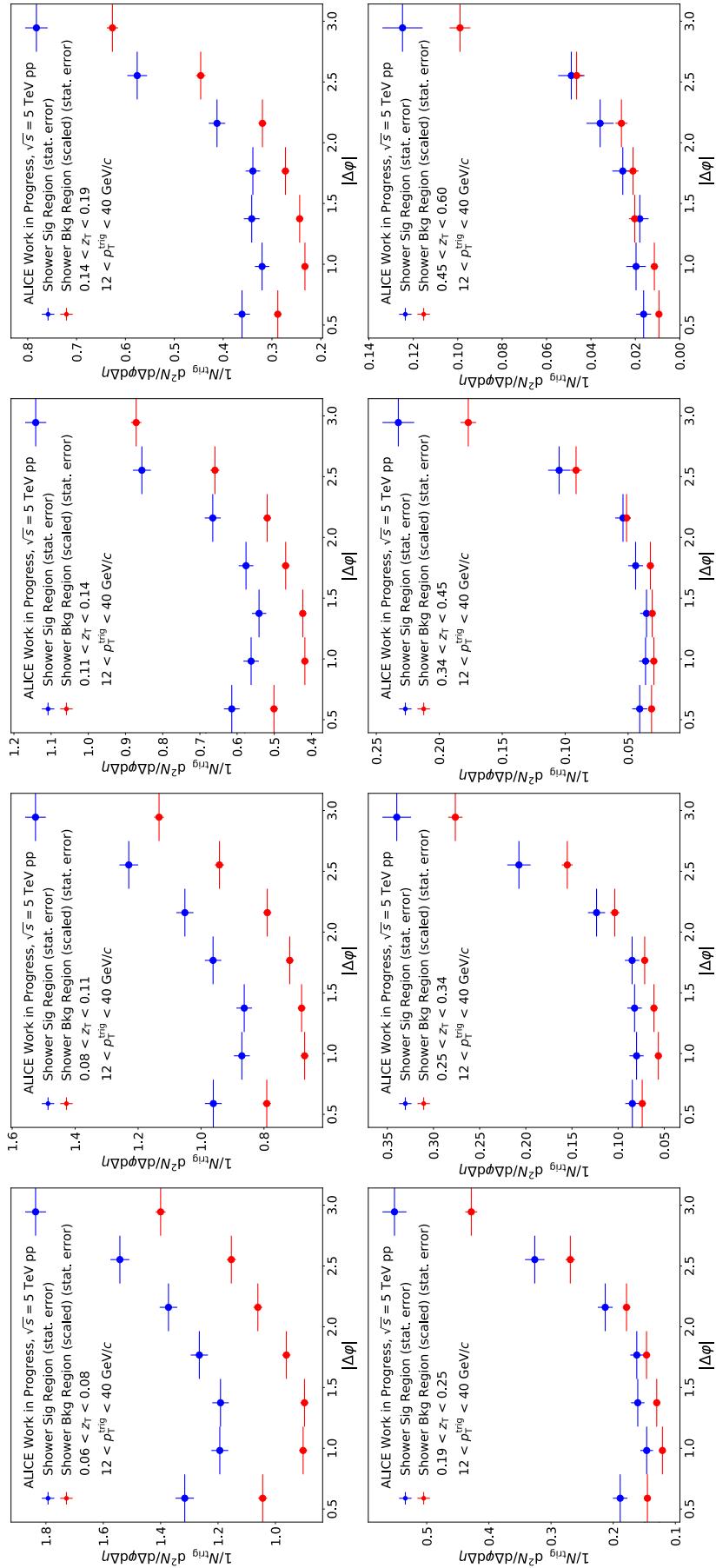
819 Figure 31 shows the two UE estimates compared with the isolated photon-hadron  $\Delta\phi$  correlations for  
820 only 2  $z_T$  bins in order to show detail. The full detail of our two UE estimates is shown in Tables 10 for  
821 pp and p–Pb data. The two estimates are consistent within uncertainties for almost all  $z_T$  bins in both pp  
822 and p–Pb data. For the only case where a significant disagreement is observed, which is for the lowest  
823  $z_T$  bin in p–Pb data, we add in quadrature the difference as an additional systematic uncertainty.

824 In order to show the effect of pedestal subtraction on the correlation functions in pp and p–Pb data, we  
825 overlay the correlation functions in both systems in Figure 32 . By construction, the points at small  $\Delta\phi$   
826 are consistent with zero as demonstrated by the dark grey bands. Additionally, the figures demonstrate  
827 the larger underlying event in p–Pb data, as well as the agreement in away side yields in the two systems  
828 after pedestal subtraction. This also shows visually the fraction of signal to background, particularly at  
829 low  $z_T$  in p–Pb collisions.

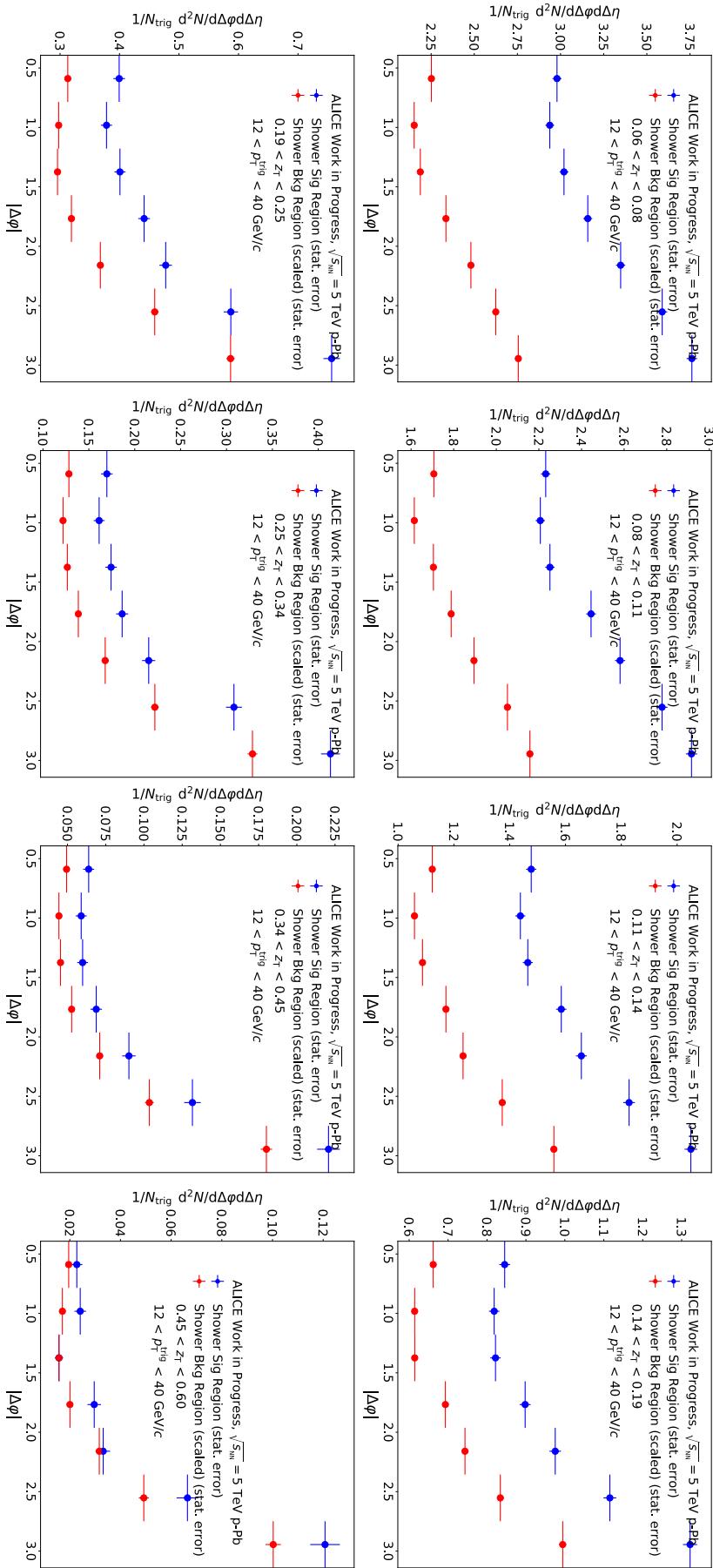
**830 12.5.1 Study of neutral energy in isolation variable using PYTHIA8**

831 In our analysis, we chose to construct the isolation variable using only charged-particles. In principle, we  
832 could have added neutral particles as well in the isolation definition. However, that would have limited  
833 the acceptance our measurement. For example, the recent ALICE isolated photon paper [2] restricted  
834 the pseudorapidity of the  $\gamma^{\text{iso}}$  to  $|\eta| < 0.27$  to ensure a good containment of the isolation cone that has  
835 a radius of  $R = 0.4$  (the EMCAL acceptance is  $|\eta| < 0.67$ ). An acceptance limitation would have a large  
836 impact of this analysis so we chose to use only “charged-only” isolation. This is not different than several  
837 ALICE jet analyses that report “charged-only” jets and not “full-jets”.

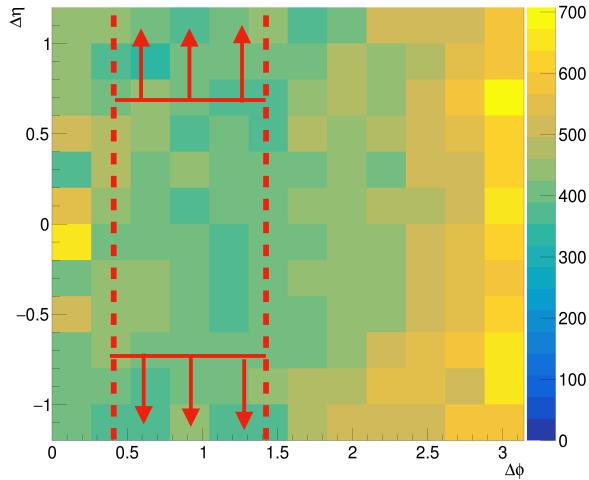
838 We use PYTHIA8 events to estimate what difference would it make in our measurement to include  
839 neutral-particles in our isolation variable. Figure 33 shows the comparison of the prompt-photon hadron  
840 correlations according to PYTHIA8 when using no isolation requirement; with an isolation variable based  
841 on charged particles (this is what we use in our analysis); and with an isolation variable based on both



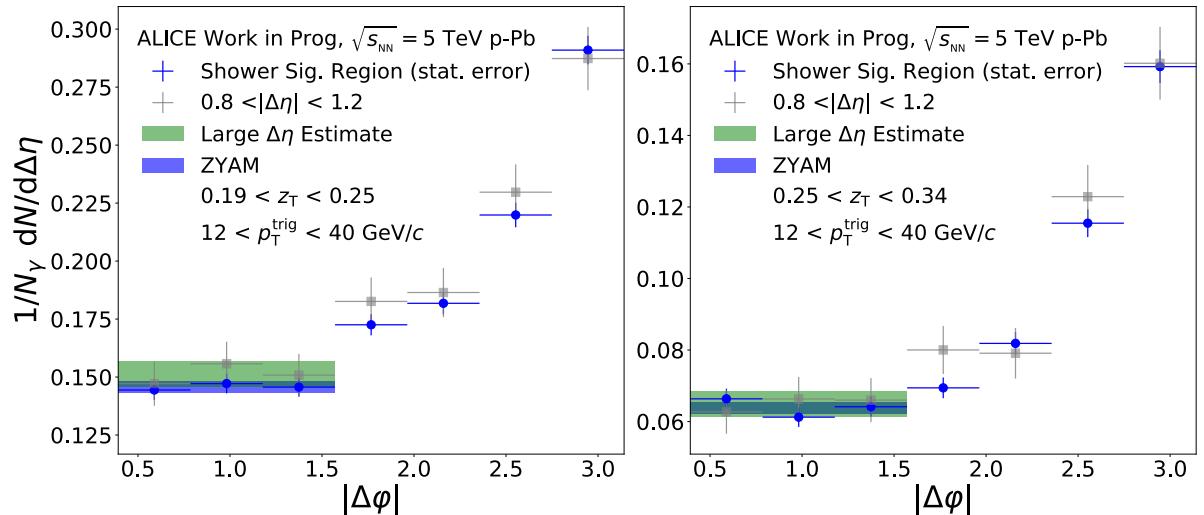
**Fig. 28:** Signal-region correlation  $C_{SR}$  (blue) and background-region correlation  $C_{BR}$  (red) in pp collisions for various  $z_T$  intervals. The error bars represent statistical uncertainties only.



**Fig. 29:** Signal-region correlation  $C_{\text{SR}}$  (blue) and background-region correlation  $C_{\text{BR}}$  (red) in p-Pb collisions for various  $z_T$  intervals. The error bars represent statistical uncertainties only.



**Fig. 30:** The 2D region used to calculate the uncorrelated background. The  $\Delta\phi$  region is chosen to avoid the away side jet peak, as well as our isolation region of  $R=0.4$ . The  $\Delta\eta$  region is chosen assuming that genuine correlations from hard-scatterings decrease as  $\Delta\eta$  increases. The large  $\Delta\eta$  is projected onto the  $\Delta\phi$  axis, and then averaged within region of  $0.4 < \Delta\phi < 1.4$ . ZYAM is estimated in the region  $0.4 < \Delta\phi < 1.4$ , but for our full  $\Delta\eta$  range ( $-1.2 < \Delta\eta < 1.2$ ).



**Fig. 31:** Projections of the  $\gamma^{\text{iso}}$ –hadron correlations in p–Pb collisions in 2  $z_T$  bins after correlated subtraction with UE estimates plotted. The grey points represent the large  $\Delta\eta$  region ( $0.8 < |\Delta\eta| < 1.2$ ) projected onto the  $\Delta\phi$  axis. The blue band represents the region used to calculate ZYAM and the green band represents the region of large  $\Delta\eta$  points used to calculate the Large  $\Delta\eta$  estimate.

842 charged-particles and neutral particles. In all cases, the charged-particles and neutral particles are final-  
 843 state particles and have  $p_T > 150$  MeV/ $c$  and  $|\eta| < 0.8$ . We observe that there is no significant dif-  
 844 ference between the selection of 1.5 GeV/ $c$  based on charged particles and the selection based on 2.0  
 845 GeV/ $c$  based on charged and neutral particles. We conclude therefore than our ISO < 1.5 GeV/ $c$  se-  
 846 lection is enough to suppress the near-side peak in the correlation functions coming from fragmentation  
 847 photons, and that using neutral-particles in our isolation variable would not yield any significant im-  
 848 provement.

**Table 10:** Summary of UE-pedestal estimates with ZYAM and the  $\Delta\eta$  method for various  $z_T$  bins, as well as the difference between LE and ZYAM estimates. The background estimate is shown in units of pairs per trigger. The uncertainty quoted is statistical only.

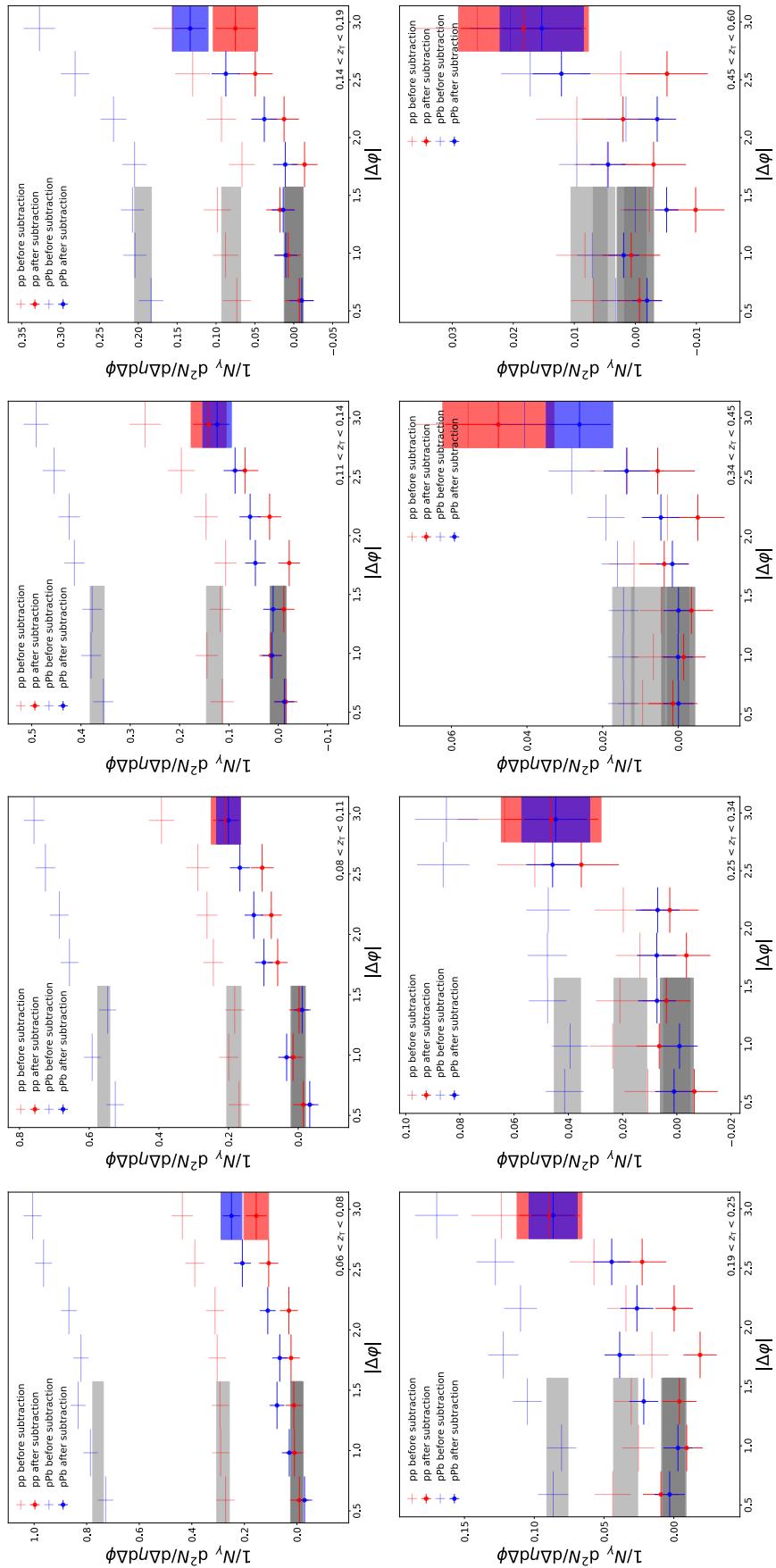
$z_T$ interval	ZYAM	Large $\Delta\eta$	Difference
<b>pp</b>			
0.06 - 0.08	$0.484 \pm 0.007$	$0.467 \pm 0.015$	$0.017 \pm 0.016$
0.08 - 0.11	$0.353 \pm 0.006$	$0.353 \pm 0.013$	$0.000 \pm 0.014$
0.11 - 0.14	$0.225 \pm 0.005$	$0.215 \pm 0.010$	$0.010 \pm 0.011$
0.14 - 0.19	$0.134 \pm 0.004$	$0.130 \pm 0.008$	$0.004 \pm 0.009$
0.19 - 0.25	$0.065 \pm 0.002$	$0.058 \pm 0.005$	$0.006 \pm 0.006$
0.25 - 0.34	$0.032 \pm 0.002$	$0.027 \pm 0.004$	$0.005 \pm 0.004$
0.34 - 0.45	$0.015 \pm 0.001$	$0.016 \pm 0.003$	$0.002 \pm 0.003$
0.45 - 0.60	$0.007 \pm 0.001$	$0.006 \pm 0.002$	$0.001 \pm 0.002$
<b>p–Pb</b>			
0.06 - 0.08	$1.170 \pm 0.006$	$1.212 \pm 0.013$	$0.042 \pm 0.014$
0.08 - 0.11	$0.876 \pm 0.005$	$0.886 \pm 0.011$	$0.010 \pm 0.012$
0.11 - 0.14	$0.573 \pm 0.004$	$0.582 \pm 0.009$	$0.008 \pm 0.010$
0.14 - 0.19	$0.325 \pm 0.003$	$0.324 \pm 0.007$	$0.001 \pm 0.007$
0.19 - 0.25	$0.154 \pm 0.002$	$0.163 \pm 0.005$	$0.008 \pm 0.005$
0.25 - 0.34	$0.066 \pm 0.001$	$0.066 \pm 0.003$	$0.000 \pm 0.003$
0.34 - 0.45	$0.024 \pm 0.001$	$0.023 \pm 0.002$	$0.001 \pm 0.002$
0.45 - 0.60	$0.008 \pm 0.000$	$0.007 \pm 0.001$	$0.001 \pm 0.001$

#### 849    12.5.2 Estimate of impact of acceptance difference between pp and p–Pb due to boost

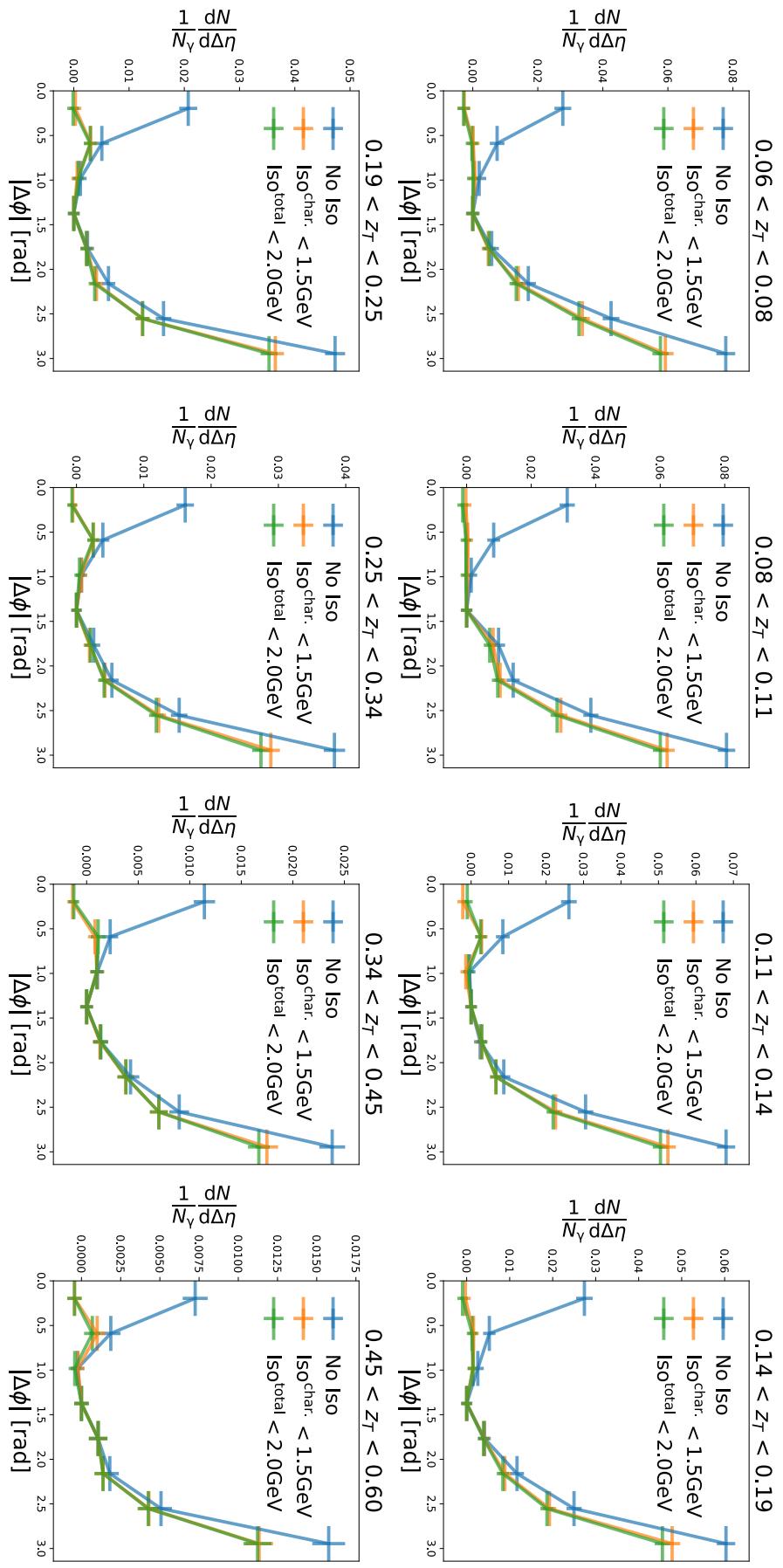
850    In this section, we estimate the impact of the acceptance difference between pp and p–Pb data that arises  
 851    due to the boost in p–Pb data. The boost in p–Pb data arises due to the energy difference between  
 852    the proton and lead beam, and it amounts to a rapidity difference of  $\Delta y = 0.47$  in the proton-going  
 853    direction. That means that in p–Pb collisions, our lab acceptance for photons that is  $-0.67 < \eta < 0.67$   
 854    corresponds to  $-0.2 < \eta < 1.14$  in the center-of-mass frame, whereas our charged-particle acceptance  
 855    of  $-0.8 < \eta < 0.8$  corresponds to  $-0.33 < \eta < 1.27$  in the center-of-mass frame.

856    We use PYTHIA8 events to estimate what is the difference between  $\gamma^{\text{iso}}$ –hadron correlations with the  
 857    acceptance of  $\gamma^{\text{iso}}$  and charged particles  $-0.20 < \eta < 1.14$  and  $-0.33 < \eta < 1.27$  instead of the nominal  
 858    ranges of  $-0.67 < \eta < 0.67$  and  $-0.8 < \eta < 0.8$ . This is shown in Figure 34. We observe that with  
 859    boosted acceptance, the  $\gamma^{\text{iso}}$ –hadron correlation of about 5% lower than with the nominal acceptance,  
 860    irrespective of  $z_T$  range. For illustration purposes, we also show the impact of a boost of  $\Delta y = 1.0$ , which  
 861    shows a decrease of about 15% with respect to the nominal acceptance, irrespective of  $z_T$  range.

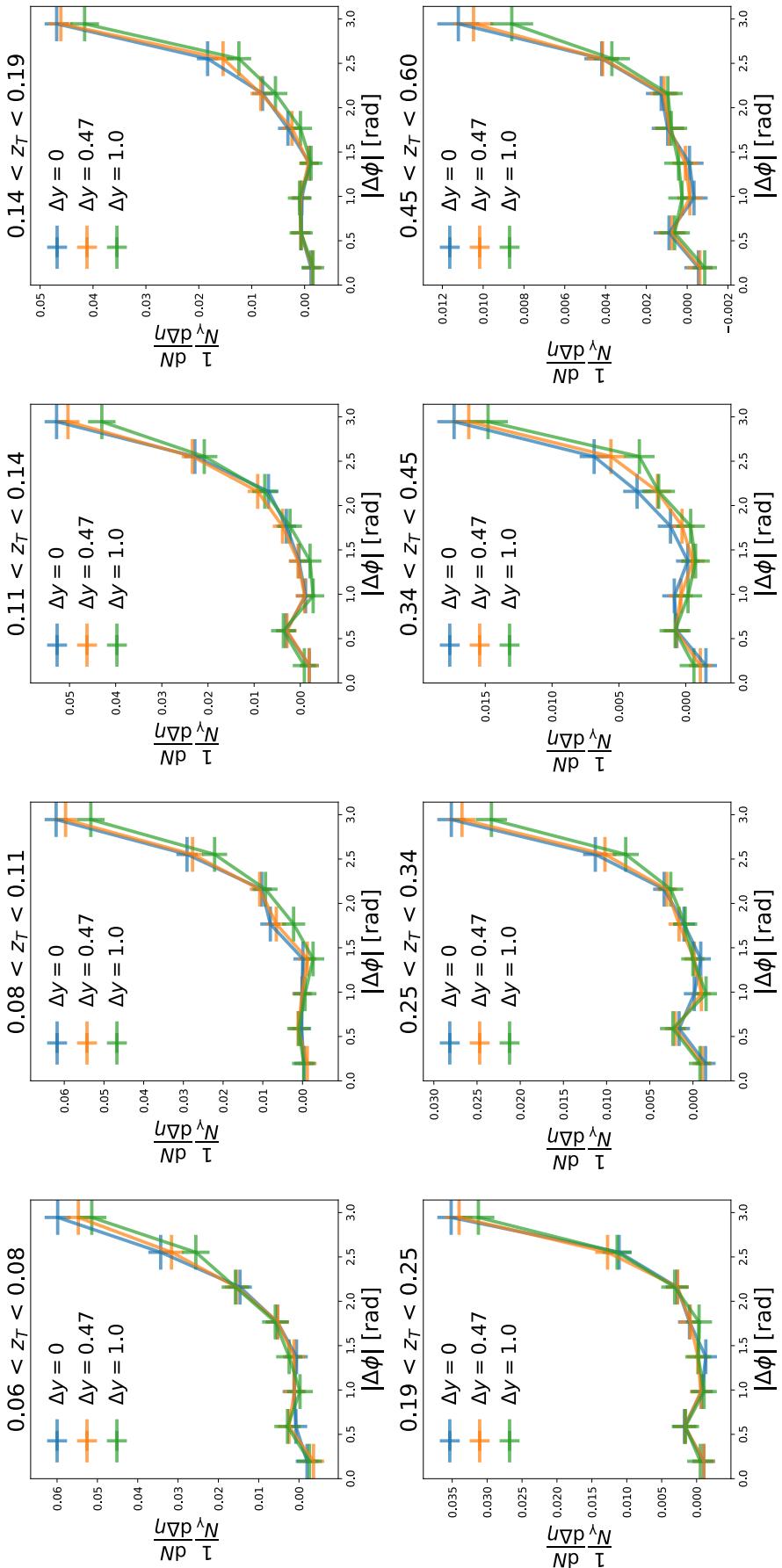
862    We thus conclude that the effect of acceptance mismatch in our analysis is limited due to the relatively  
 863    small boost of  $\Delta y = 0.47$  and the limited acceptance of EMCAL and ITS, which even with boost is within  
 864    mid-rapidity region where the cross-sections do not change drastically.



**Fig. 32:** Correlation functions in pp (red) and p–Pb (blue) before and after pedestal subtraction. The light grey bands represent the ZYAM estimates, while the dark gray bands represent the near side average after pedestal subtraction



**Fig. 33:** Correlation function between prompt photons and hadrons from PYTHIA8 for various  $z_T$  bins. Three selections on the prompt photons based on isolation are presented: no isolation (blue); ISO  $< 1.5\text{GeV}/c$  that considers only charged-particles (orange); and ISO  $< 2.0\text{ GeV}/c$  that considers both charged and neutral particles (green). In all cases, the uncorrelated background has been subtracted using the ZYAM method. The error bar represent statistical uncertainty only.



**Fig. 34:** Correlation function between isolated prompt photons and hadrons from PYTHIA8 for various  $\zeta_T$  bins. The nominal result ( $\Delta y = 0$ , in blue) is compared with results obtained with a kinematic selection that mimics the boost of p–Pb data (orange). For illustration purposes, we show the impact of a boost that is larger than the one of p–Pb data ( $\Delta y = 1.0$  in green).

## 865 13 Systematic uncertainties

866 The source of systematic uncertainties of our  $\gamma^{\text{iso}}$ -hadron measurement are the following:

867 – Purity

868 The uncertainty of the purity measurement, which is described in Section 10.2, is propagated to  
 869 the correlation function measurement following Equation 23. The resulting uncertainty on the  
 870 correlation function is a relative  $\pm 18\%$  for pp data and  $\pm 12\%$  for p–Pb data . As described  
 871 in Section 10.2, a large fraction of the purity total uncertainty is either statistical uncertainty or  
 872 systematic uncertainties that arise due to limited data sample. Therefore, the purity uncertainty in  
 873 pp and p–Pb data are largely uncorrelated. As a conservative approach, we take them to be totally  
 874 uncorrelated.

875 – Underlying Event:

876 As described in Section 9.1, the uncertainty of the UE subtraction originates from statistical fluctuations  
 877 in the ZYAM estimate. It propagates directly to our per-trigger yields. It ranges from 7% to  
 878 15% depending on the  $z_T$  bin and data. This uncertainty is fully correlated in  $\phi$  for a given  $z_T$  bin,  
 879 but totally uncorrelated among  $z_T$  bins, and totally uncorrelated between pp and p–Pb datasets.

880 – Tracking performance :

881 To estimate the systematic uncertainty of our charged-particle  $p_T$  measurement with ITS-only  
 882 track reconstruction, we perform MC simulation studies and make a comparison with published  
 883  $p_T$  spectra that used the ALICE standard tracking (i.e. including TPC) in pp and p–Pb collisions at  
 884 5 TeV [12]. As described in Section 11, the combined uncertainty due to track efficiency, fake rate,  
 885 and bin-to-bin migration corrections amounts to  $\pm 5\%$  added in quadrature with the total systematic  
 886 uncertainty of our reference  $p_T$  spectra, which ranges from a relative 1.6 (2.1%) to 1.9% (2.5%) in  
 887 the range  $0.5 < p_T < 10 \text{ GeV}/c$  for pp (p–Pb) collisions [12].

888 Systematic uncertainties due to secondary-particle contamination and from modelling of the particle-  
 889 type composition in MC simulations are small ( $< 2\%$ ) for the range  $0.5 < p_T < 10 \text{ GeV}/c$ . These  
 890 were already estimated in Ref.[12] for pp and p–Pb data sets and already included in the systematic  
 891 uncertainty estimate described above.

892 The tracking performance between pp and p–Pb datasets is very similar, but as a conservative  
 893 approach we take the systematic uncertainties to be completely uncorrelated.

894 – Acceptance mismatch due to boost:

895 As described in Section 12.5.2, our PYTHIA8 study of  $\gamma^{\text{iso}}$ -hadron correlations show that the  
 896 impact of an acceptance mismatch between pp and p–Pb data that arises from the boost of  $\Delta y =$   
 897 0.47 amounts to about 5% effect irrespective of  $z_T$ . This estimate is subject to PDF uncertainties,  
 898 which are the ones that dictate the shape of the differential cross-section of photons and associated  
 899 hadrons in pseudorapidity. We chose to not apply any correction for this effect, and assign a  $\pm 5\%$   
 900 systematic uncertainty on the per-trigger hadron yields. This systematic uncertainty is taken to be  
 901 completely correlated with  $z_T$ . We assign this systematic uncertainty to our p–Pb measurements  
 902 only.

903 – Luminosity, trigger, photon, and vertex reconstruction:

904 Our observable is normalized per measured photon. Therefore the uncertainties related to overall  
 905 normalization of the  $\gamma^{\text{iso}}$   $p_T$  spectra (such as luminosity scale, vertexing efficiency, trigger effi-  
 906 ciency and photon reconstruction efficiency) cancel completely. Consequently, we do not assign  
 907 any systematic uncertainty associated with these sources in our measurement.

908 – Photon energy scale, resolution and material budget:

909 While we are by construction totally insensitive to overall normalization, we are in principle sen-

**Table 11:** Summary of uncertainties in  $\gamma^{\text{iso}}$ -hadron correlations, which are reported as per-trigger yields of correlated hadrons. The uncertainties quoted are relative. The ranges shown encompass the relative uncertainties for hadron  $z_T$  in 2 ranges: Low (0.06–0.18) and High (0.18–0.6)  $z_T$ .

	Low $z_T$ pp data	High $z_T$ pp data	Low $z_T$ p–Pb data	High $z_T$ p–Pb data
Statistical Uncertainty	19%-40%	30%-49%	16%-23%	29%-39%
Photon Purity	18%	18%	12%	12%
Underlying Event	8%-15%	7%-12%	7%-10%	8%-9%
Tracking performance	5.6%	5.6%	5.6%	5.6%
Acceptance mismatch	—	—	5%	5%
Photon Energy Scale	<1%	<1%	<1%	<1%
Photon Energy Resolution	<1%	<1%	<1%	<1%
Material budget	<1%	<1%	<1%	<1%
Luminosity scale	—	—	—	—
Vertex efficiency	—	—	—	—
Trigger corrections	—	—	—	—
Photon reconstruction efficiency	—	—	—	—
Total Systematic Uncertainty	21%-24%	20%-22%	15%-16%	15%-16%
Total Uncertainty	28%-47%	36%-54%	22%-28%	33%-43%

910 sitive to bin-migration or scale uncertainties that affect the shape of the photon  $p_T$  spectra. This  
 911 potential systematic uncertainty is reduced because we integrate over large photon  $p_T$  range (12–  
 912 40 GeV/ $c$ ). Moreover, the EMCAL performance is such that these effects are small; for a 12 GeV  
 913 cluster the resolution  $\sigma_E/E = 4.8\% / E \otimes 11.3\% / \sqrt{E} \otimes 1.7\%$  yields  $\sigma_E/E = 3.6\%$ , and at 40 GeV  
 914 this yields  $\sigma_E/E = 2.4\%$ . The EMCAL energy scale has been studied with beam-test data [3] and  
 915 comparison of  $\pi^0 \rightarrow \gamma\gamma$  events in data and simulation [1], and has an associated uncertainty if  
 916 0.8%.

917 The uncertainties due to photon energy scale, resolution, and material budget have been estimated  
 918 for the isolated photon cross-section measurement with 7 TeV pp and 5 TeV p–Pb data and are less  
 919 than 3% in the  $p_T$  range covered in this analysis [19, 2]. The effects on the per-trigger correlation  
 920 functions would be even smaller. Given that this level of uncertainty are much smaller than other  
 921 sources of systematic uncertainties for our measurement, we neglect them.

922 Table 11 presents as summary of all uncertainty estimates for our  $\gamma^{\text{iso}}$ –hadron correlation measurement.  
 923 Table 12 shows the uncertainty in the p–Pb to pp ratio.

**Table 12:** Summary of leading relative uncertainties on the fragmentation ratio between proton-lead and proton-proton collisions.

$z_T$ interval	Statistical	UE Estimate	$\gamma^{\text{iso}}$ Purity	Tracking corrections	Acceptance mismatch
0.06–0.08	34%	15%	22%	8%	5%
0.08–0.11	25%	11%	22%	8%	5%
0.11–0.14	33%	13%	22%	8%	5%
0.14–0.19	43%	16%	22%	8%	5%
0.19–0.25	33%	11%	22%	8%	5%
0.25–0.34	49%	15%	22%	8%	5%
0.34–0.45	44%	11%	22%	8%	5%
0.45–0.60	63%	14%	22%	8%	5%

**Table 13:** Summary of uncertainties on integrated away side yields in proton-lead and proton-proton collisions. The uncertainties quoted are absolute.

$z_T$ Range	pp $\pm$ Stat. $\pm$ Sys	p–Pb $\pm$ Stat. $\pm$ Sys
0.06–0.08	$7.55 \pm 2.23 \pm 1.43$	$11.18 \pm 1.89 \pm 1.46$
0.08–0.11	$7.73 \pm 1.48 \pm 1.47$	$7.56 \pm 1.21 \pm 0.99$
0.11–0.14	$4.14 \pm 0.97 \pm 0.79$	$3.44 \pm 0.79 \pm 0.45$
0.14–0.19	$1.44 \pm 0.58 \pm 0.27$	$2.68 \pm 0.46 \pm 0.35$
0.19–0.25	$1.43 \pm 0.36 \pm 0.27$	$1.26 \pm 0.26 \pm 0.16$
0.25–0.34	$0.54 \pm 0.21 \pm 0.10$	$0.50 \pm 0.15 \pm 0.07$
0.34–0.45	$0.43 \pm 0.13 \pm 0.08$	$0.23 \pm 0.08 \pm 0.03$
0.45–0.60	$0.14 \pm 0.07 \pm 0.03$	$0.11 \pm 0.04 \pm 0.01$

## 924 14 Results

### 925 14.1 Azimuthal correlations in pp and p–Pb data

926 Figure J.1 shows the fully-corrected  $\gamma^{\text{iso}}$ –hadron correlations in pp and p–Pb. The band at zero represents  
927 the uncertainty from the uncorrelated background estimate. The vertical error bars represent the statistical  
928 error bars only. We observe an agreement within uncertainties between pp, p–Pb, and PYTHIA 8.2  
929 Monash tune in the presented  $z_T$  range. We quantify this by means of a  $\chi^2$  test between pp and p–  
930 Pb data that considers all uncertainties and their correlation among the different  $\Delta\phi$  bins. No significant  
931 difference is observed ( $p$  values always larger than 0.05).

### 932 14.2 Fragmentation functions in pp & p–Pb data

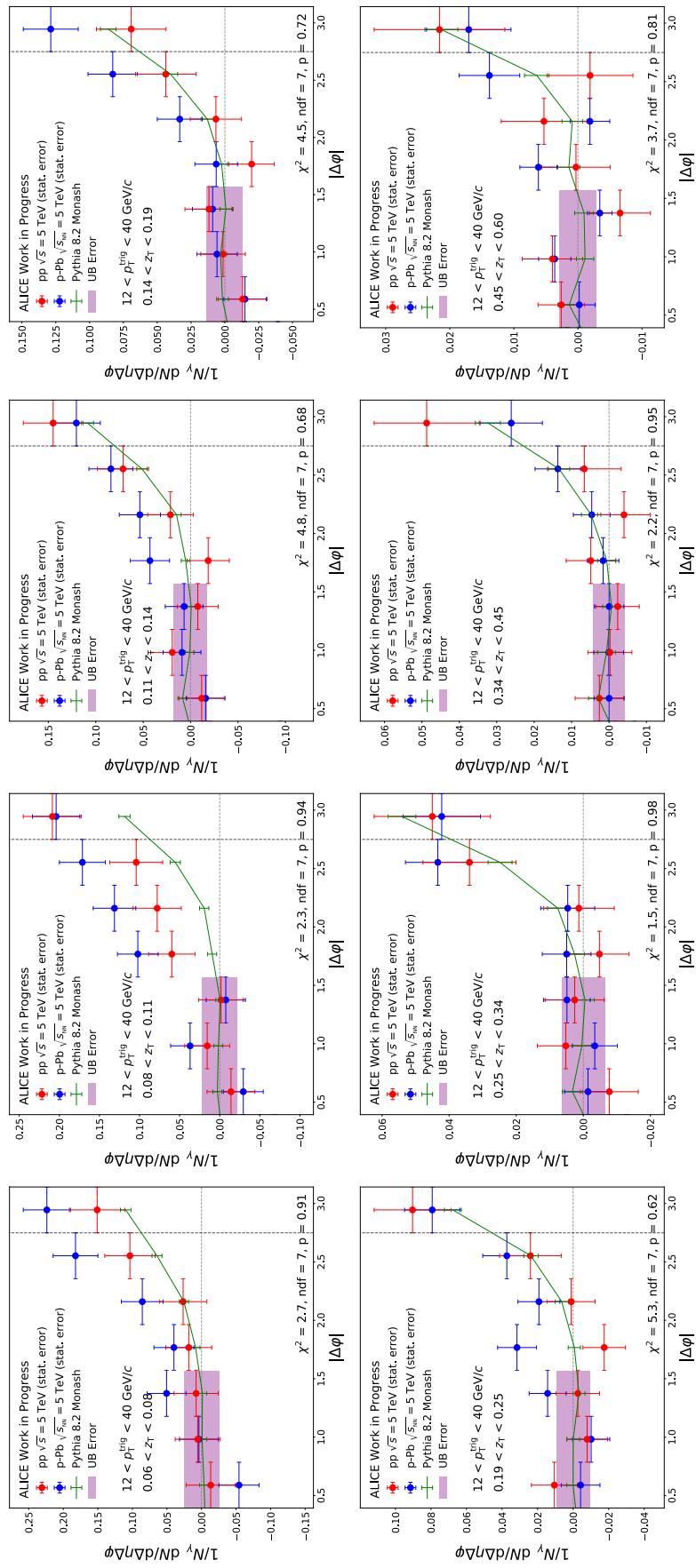
933 Finally, we integrate the  $\gamma^{\text{iso}}$ –hadron correlation functions in the range  $|\Delta\phi| > 7\pi/8$  to roughly corre-  
934 spond to a cone with radius  $R = 0.4$  used in various jet studies. The integrals are presented as a function  
935 of  $z_T$  in Figure 36 and Table 14.

**Table 14:** Number of  $\gamma^{\text{iso}}$ –hadron pairs per  $\gamma^{\text{iso}}$  integrated in  $\Delta\phi > 7\pi/8$ , for different  $z_T$  intervals. The uncertainty  
quoted is statistical only.

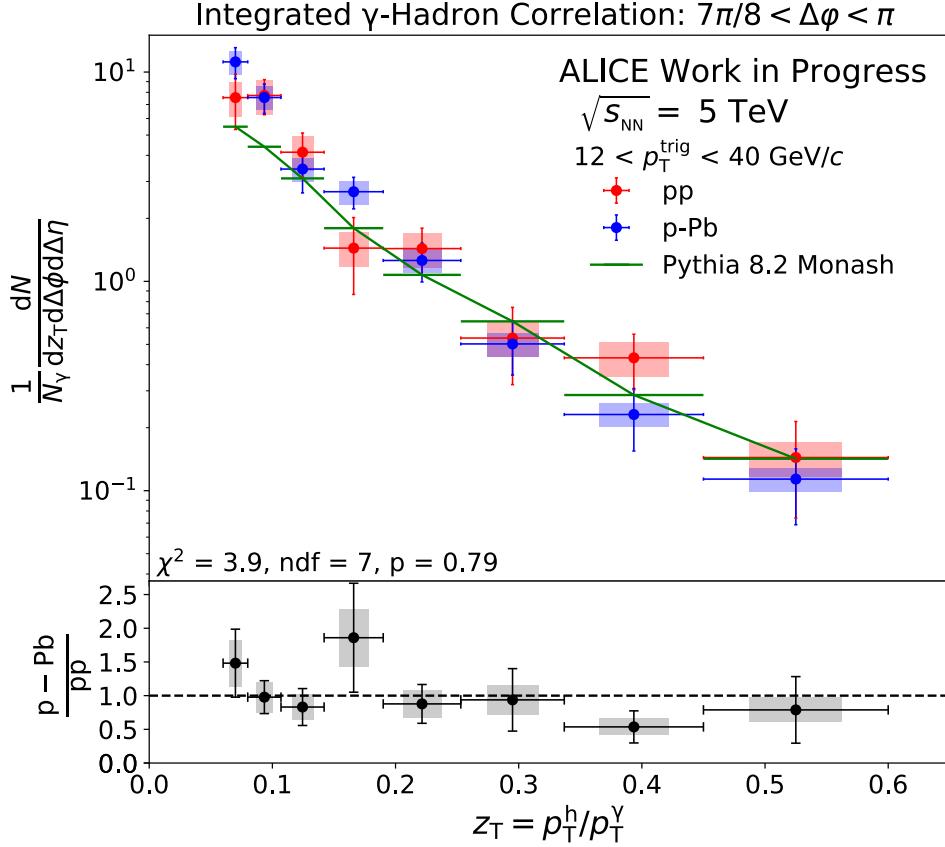
$z_T$ range	pp	p–Pb	p–Pb/pp
0.06 - 0.08	$7.548 \pm 2.230$	$11.185 \pm 1.890$	$1.482 \pm 0.504$
0.08 - 0.11	$7.728 \pm 1.475$	$7.555 \pm 1.215$	$0.978 \pm 0.244$
0.11 - 0.14	$4.140 \pm 0.974$	$3.440 \pm 0.793$	$0.831 \pm 0.274$
0.14 - 0.19	$1.442 \pm 0.576$	$2.681 \pm 0.460$	$1.859 \pm 0.808$
0.19 - 0.25	$1.434 \pm 0.361$	$1.259 \pm 0.265$	$0.878 \pm 0.288$
0.25 - 0.34	$0.536 \pm 0.215$	$0.503 \pm 0.146$	$0.937 \pm 0.463$
0.34 - 0.45	$0.431 \pm 0.128$	$0.231 \pm 0.076$	$0.536 \pm 0.238$
0.45 - 0.60	$0.144 \pm 0.070$	$0.114 \pm 0.045$	$0.789 \pm 0.494$

936 Figure 37 shows the ratio of p–Pb to pp data. The systematic uncertainties in the ratio are described  
937 in Section 13. The uncertainty due to UE-subtraction is fully uncorrelated with  $z_T$  and is combined in  
938 quadrature with the statistical uncertainty and shown as bars. All other uncertainties are correlated with  
939  $z_T$  and shown in boxes.

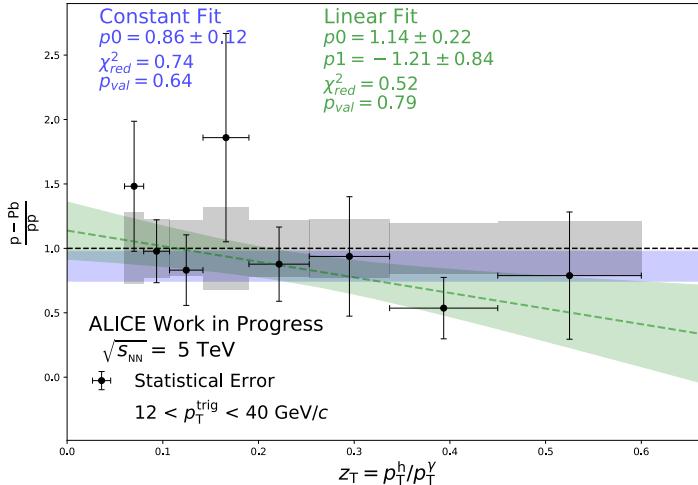
940 The ratio is consistent with unity within uncertainties. We fit the ratio with a constant (using only the  
941 uncertainty that is uncorrelated with  $z_T$ ) and obtain  $0.85 \pm 0.12$  with a reduced  $\chi^2$  of 0.72. Thus, we find  
942 no significant difference between pp and p–Pb  $\gamma^{\text{iso}}$ -tagged fragmentation pattern



**Fig. 35:** Fully-corrected  $\gamma^{\text{iso}}$ –hadron correlation function pp (red) and p–Pb (blue) data. The purple band represents the uncertainty from the underlying event estimate in pp and p–Pb. The error bars represent statistical uncertainty only. The green line is the  $\gamma^{\text{iso}}$ –hadron correlation function obtained with PYTHIA 8.2.



**Fig. 36:**  $\gamma^{\text{iso}}$ -tagged fragmentation function for pp (red) and p-Pb data (blue). The shaded boxes represent the systematic uncertainties. The green line again represents the PYTHIA 8.2 Monash. The  $\chi^2$  test for the comparison of pp and p-Pb data incorporates correlations among different  $z_T$  intervals.



**Fig. 37:** Ratio of the  $\gamma^{\text{iso}}$ -tagged fragmentation function of p-Pb to pp data. The vertical bars represent statistical uncertainties, and the grey boxes centered at unity represent systematic uncertainties. The horizontal blue band along the the plot represents a constant fit to the ratio. The green line represents a linear fit to the ratio. The widths of both fits represent the 68% confidence intervals.

943 **15 Conclusions**

944 We have reported a measurement of  $\gamma^{\text{iso}}$ –hadron correlations in p–Pb and pp data at 5 TeV. Our result  
 945 significantly extends previous results by selecting isolated photons at lower momentum, which is where  
 946 the largest modifications of jet fragmentation in nuclear matter are expected. Our results indicate that  
 947 there are no difference between the two datasets, within uncertainties. We also found that PYTHIA8.2  
 948 describes the data within uncertainties. This constrains the effects of cold-nuclear matter effects. This  
 949 result establishes a benchmark on photon identification for future measurements with higher statistics in  
 950 p–Pb and Pb–Pb collisions.

951 **Appendices**

952 **A Isolated-photon efficiency**

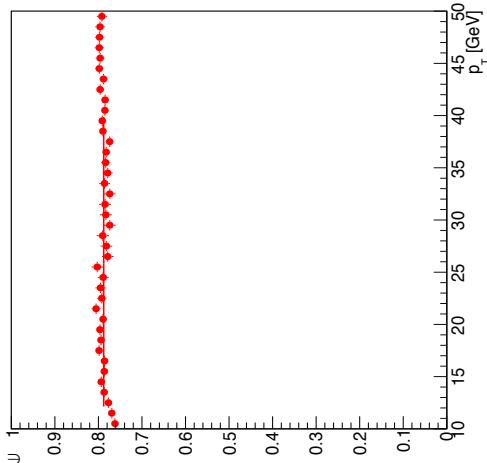
953 We report correlation functions that are normalized per photon trigger so the photon efficiency cancels.  
 954 In principle, we might introduce a bias if the photon efficiency varied rapidly within the photon  $p_{\text{T}}$  range  
 955 that we are using (12–40 GeV/c). Here we show that this is not the case.

956 The efficiency our our isolated-photon selection is shown in Figure A.1. The efficiency is rather inde-  
 957 pendent of  $p_{\text{T}}$  in the range relevant for this analysis. We observe less than 1% variation between the high  
 958 and low of the energy range of our photon triggers (77.7% at 12 GeV/c and 78.5% at 40 GeV/c). This  
 959 level of variation has a negligible impact in our correlation analysis.

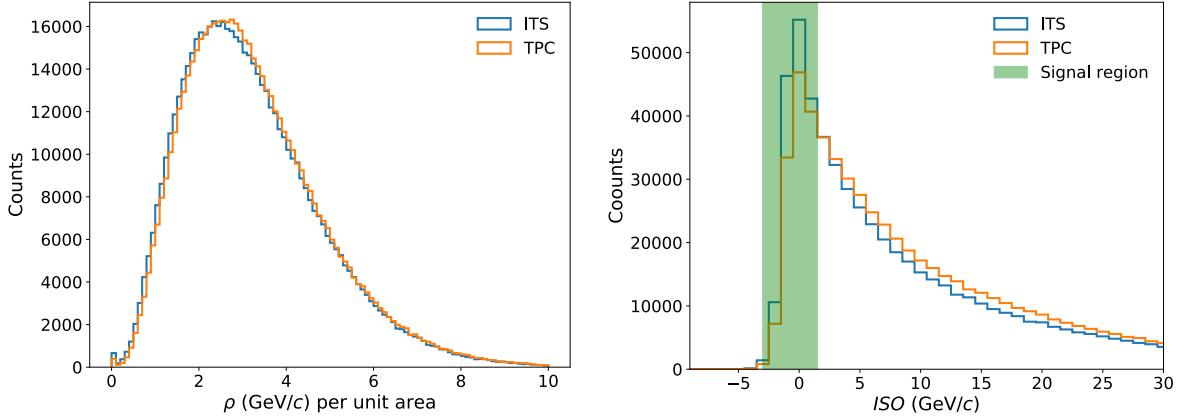
960 **B Checks of UE estimates with standard tracking**

961 As discussed in Section 11, the TPC had space-charge distortions during the 2013 p–Pb run that resulted  
 962 in a drop in efficiency for tracking beyond 4 GeV, which limits our ability to use it for our correlation  
 963 measurements. However, the TPC tracks can still be used for low  $p_{\text{T}}$  tracking, which is the relevant  
 964 region for underlying-event and isolation measurements.

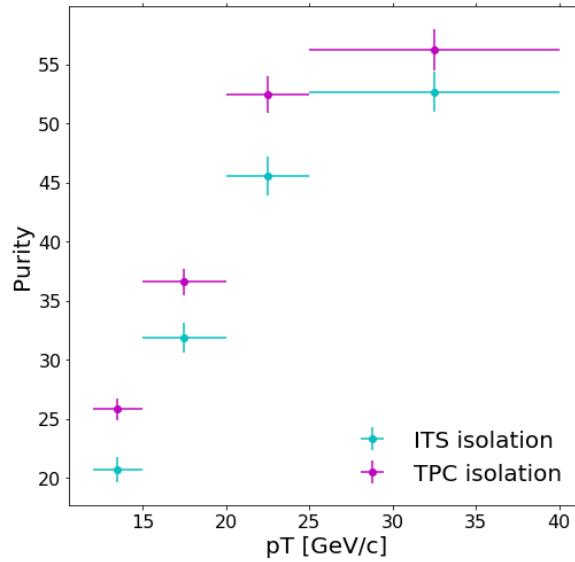
965 Figure B.1 shows the  $\rho$  and isolation distributions measured with ITS-only and TPC+ITS tracks in p–  
 966 Pb data. We found the  $\rho$  distributions to be very similar; the means of  $\rho$  to be 3.129 GeV and 3.202 GeV



**Fig. A.1:** Isolated-photon efficiency obtained with p–Pb simulation.



**Fig. B.1:** Transverse momentum density (left panel) and isolation distributions (right panel) determined with ITS tracks (in blue) and TPC+ITS tracks (in orange) in p–Pb data.

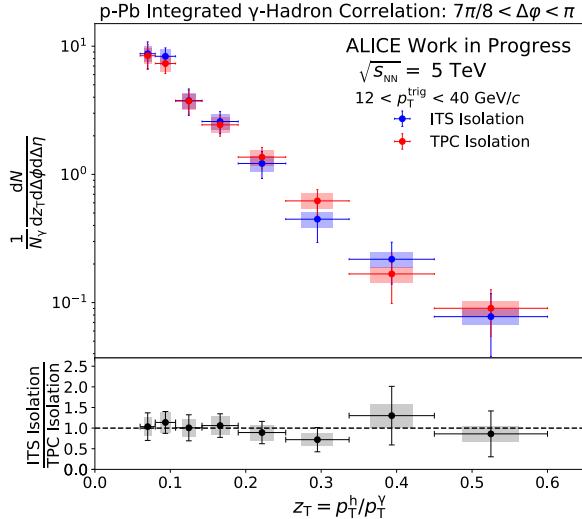


**Fig. B.2:** Comparison of purity obtained in p–Pb collisions with isolation variable obtained with ITS-only tracks (purple) and with ITS+TPC tracks (cyan). The error bars represent statistical uncertainty only.

for ITS and TPC+ITS  $\rho$  values respectively. This is expected because the UE-estimation is dominated by tracks with low momentum, and while the ITS tracking resolution is poorer, the smearing effects are relative small at low momentum. We find some differences in the isolation distribution, which can be attributed to the worse momentum resolution for the ITS-only tracks as the isolation is sensitive to higher  $p_T$  tracks where the momentum resolution worsening is more significant.

For simplicity we use the same threshold of 1.5 GeV/c for our  $\gamma^{\text{iso}}$  candidates for both ITS-only and ITS+TPC tracks. We found that the isolation with ITS+TPC tracks leads to a better rejection of the background, which leads to an increased photon purity. This is shown in Figure B.2.

We checked our main results (correlation function) in p–Pb data by performing our analysis with isolation variable, UE estimate, and corresponding purity values separately for ITS-only tracks and ITS+TPC tracks. As shown in Figure B.3, we obtain consistent results. We obtain a slightly better statistical uncertainty when including TPC (a relative uncertainty of 22% to 41% depending on  $z_T$  vs 24% to 51% for the ITS case), which can be attributed to the corresponding higher purity. However, these slightly better statistical uncertainties do not change the main result of our paper. For consistency with pp results



**Fig. B.3:** Comparison of fragmentation function measurement in p–Pb collisions with isolation variable obtained with ITS-only tracks (blue) and with ITS+TPC tracks (red).

(where we cannot use ITS+TPC tracks because the TPC was not read out), we chose to report our results for ITS-only tracks.

We consider this study comparing ITS-only tracking and ITS+TPC tracks as a check against possible biases due to worse momentum resolution or fake rate of the ITS-only tracking. Given that we have obtained consistent results with the standard ITS+TPC tracking, we do not assign an additional systematic uncertainty for the tracking performance on UE-estimate and isolation variables.

### C Comparison to ABCD method results

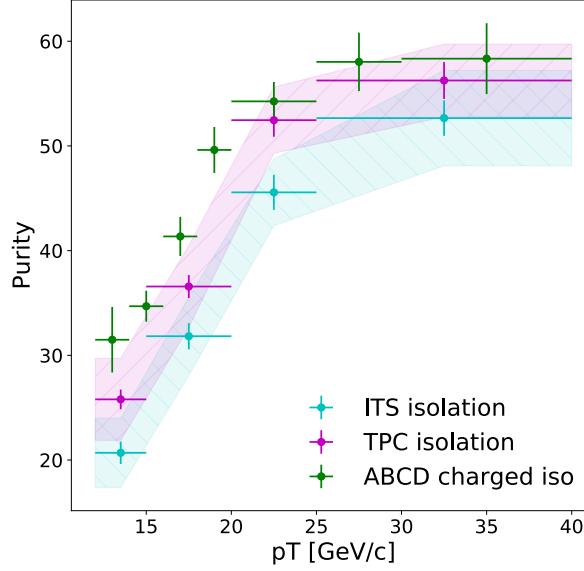
In this section we compare our results obtained with the template fit method to the ABCD method results obtained by Erwann Masson for his isolated-photon analysis in p–Pb data [19]. That study is performed with the same p–Pb data using the same event and cluster selections and a similar isolation criterion that uses only charged-particles (shown in the Appendix of Ref. [19]).

Figure C.1 shows our results compared to the results of the ABCD method. We show both our results using ITS-only tracks and ITS+TPC tracks (which is more directly comparable to the ABCD result). Our results are systematically below the ABCD method, but they are consistent within  $1\sigma$  systematic uncertainty for most of the  $p_T$  range.

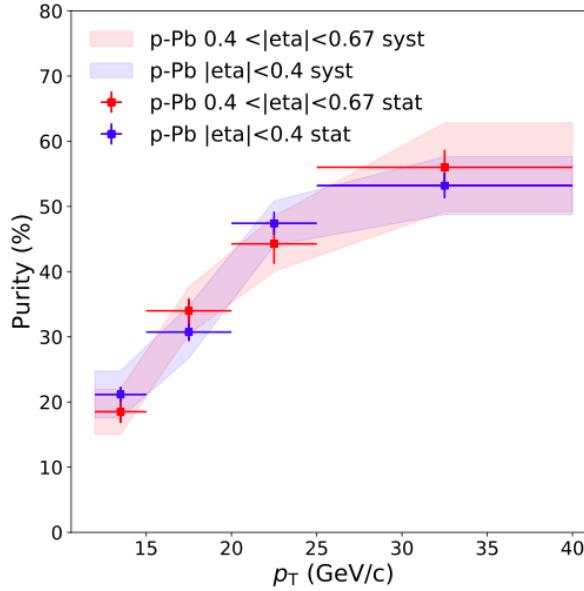
### D Check by splitting clusters in $|\eta| < 0.4$ and $0.4 < |\eta| < 0.67$

In our analysis, we use clusters with  $|\eta| < 0.67$  (Section 7); we consider an isolation variable constructed with tracks with  $|\eta| < 0.8$  (Section 11) and a cone size of  $R = 0.4$  (Section 9). The isolation cone is thus fully contained in the tracking acceptance only for clusters with  $|\eta| < 0.4$ ; the isolation cone for clusters with  $0.4 < |\eta| < 0.67$  is only partially covered in pseudorapidity angle. Note that the tracking acceptance covers the full azimuthal angle so this is not an issue in azimuth.

To check for possible biases that this partial containment of isolation cone in pseudorapidity we did split all the measured clusters into two categories:  $|\eta| < 0.4$  and  $0.4 < |\eta| < 0.67$ . In principle, if the bias introduced by the lack of total coverage of the isolation cone would lead to higher background in the  $0.4 < |\eta| < 0.67$  region (background might appear less-isolated than in reality, and might pass the selection), and thus lead to a lower purity. We perform this study in the p–Pb data, which has

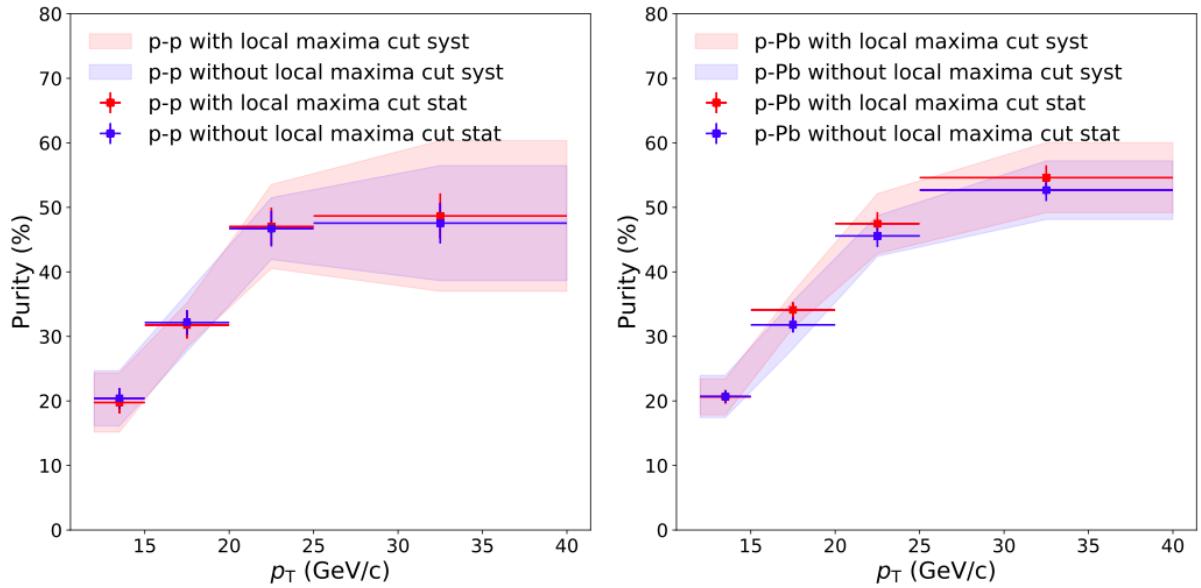


**Fig. C.1:** Comparison between purity measured with the template fit and ABCD method, from Ref. [19], in p–Pb data. The error bars represent statistical uncertainty only and the bands represent the systematic uncertainty only.



**Fig. D.1:** Purity measurement in p–Pb collisions for clusters with  $|\eta| < 0.4$  and  $0.4 < |\eta| < 0.67$ .

better statistical precision than the pp data and would allow us to better constrain small biases. The comparison of purity measurements is shown in Figure D.1. Both set of measurements are compatible within statistical uncertainties. Thus, we conclude that the potential bias due to incomplete coverage of isolation cone is negligible and we do not assign any systematic uncertainty to this source. We explain this observation by noting that the azimuthal angle is fully covered, and that most of the energy in the isolation cone is within small angles of the neutral-mesons that dominate the background (the background is primarily high-z neutral mesons in jets).



**Fig. E.1:** Purity measurement with and without cluster selection based on number-of-local maxima in pp (left) and p–Pb (right) collisions.

## 1014 E Cluster selection variations

1015 In this section, we study the impact on our purity measurement of variations in our cluster selection  
 1016 (Section 7)

1017 First, we study the impact of our number-of-local maximum criteria for clusters. The NLM < 3 selection  
 1018 was used in previous isolated photon analyzes (e.g. Ref [2, 19]), where it was found to help to improve  
 1019 the simulation description of the background shower-shape. Figure E.1 shows the purity measurement  
 1020 with and without this selection in pp and p–Pb data. We observe no significant difference.

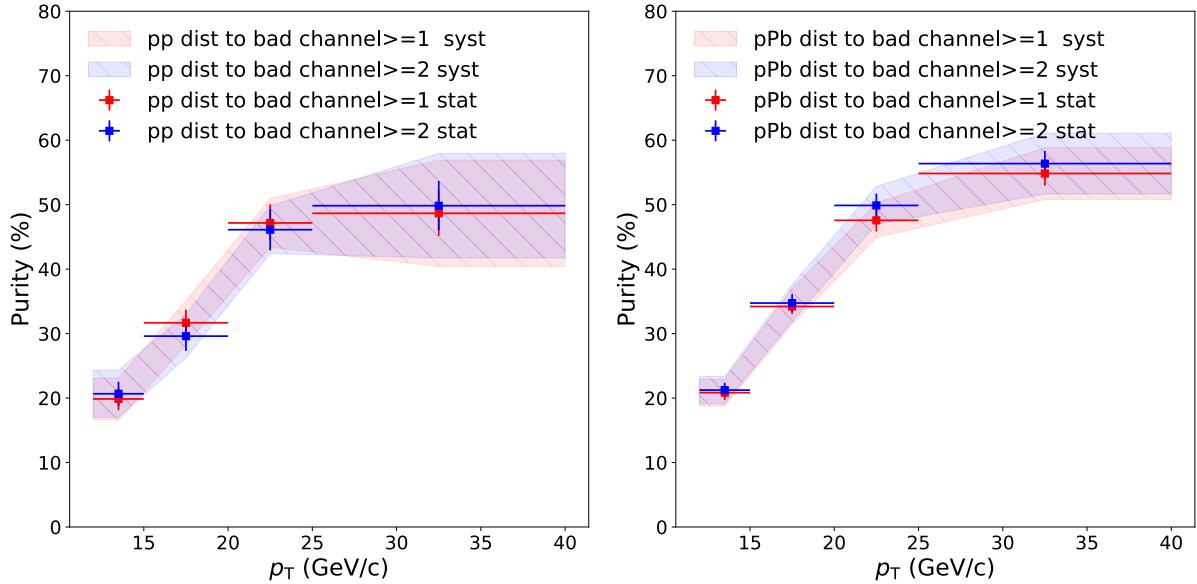
1021 In Figure E.2 we show our purity measurement by varying the distance-to-bad channel cut from our  
 1022 nominal  $\geq 1$  to  $\geq 2$ . This change would remove about 20% of our  $\gamma^{\text{iso}}$  candidates. We observe no  
 1023 significant difference.

1024 In Figure E.4 we show our purity measurement with and without cluster time selection. We observe no  
 1025 significant difference.

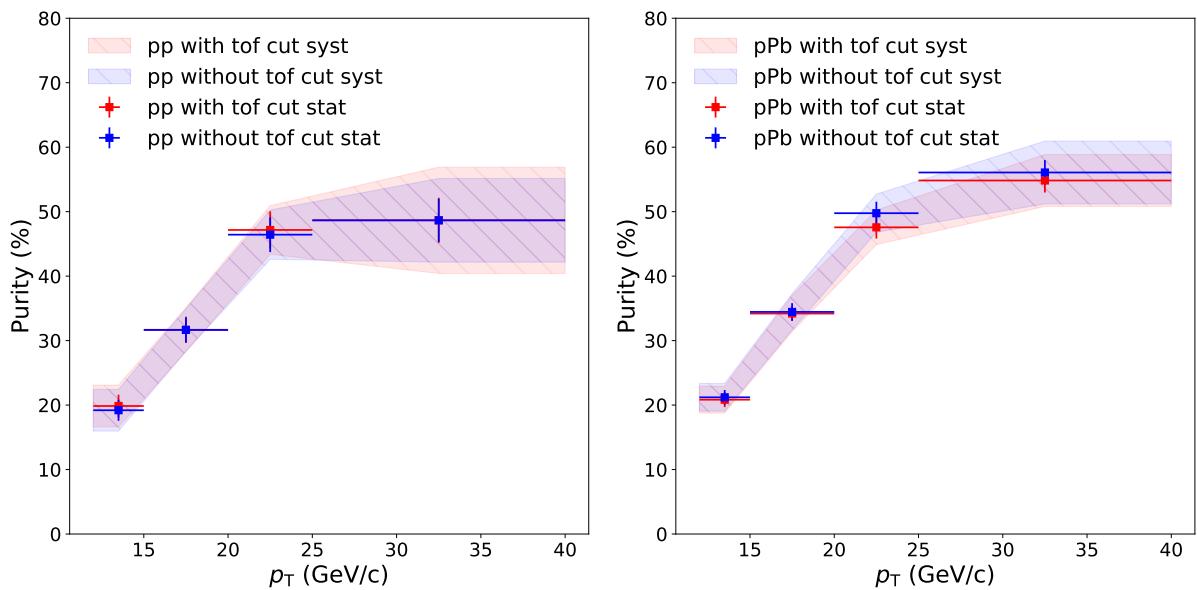
1026 In Figure E.4 we show our purity measurement with an exotoxicity cut with a threshold of 5% (nominal)  
 1027 and a variation of 3%. We observe no significant difference.

1028 In Figure E.5 we show our purity measurement with different isolation thresholds. We vary our nominal  
 1029 threshold of 1.5 GeV/c by  $\pm 0.2$  GeV/c. As expected the higher (lower) isolation threshold results in a  
 1030 lower (higher) purity. However, the difference with respect to our nominal result is small with respect to  
 1031 our statistical and systematic uncertainties. This indicates that the purity measurement has only a weak  
 1032 dependence on the choice of isolation threshold.

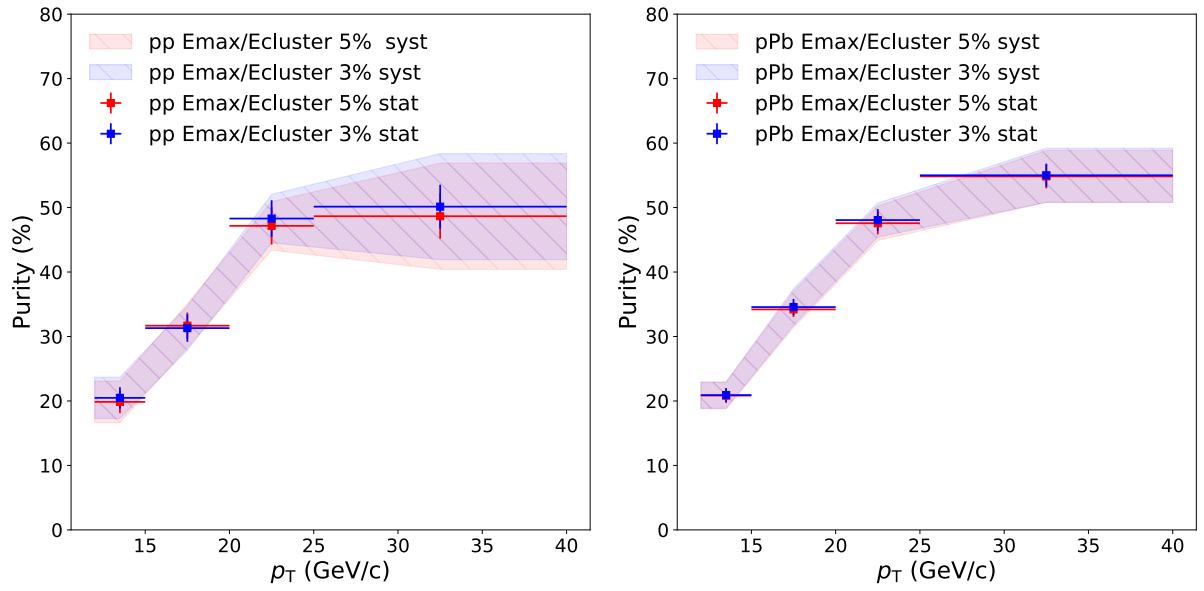
1033 Given that the impact of these cluster variations on our purity measurement is small compared to our  
 1034 estimated systematic uncertainties (described in Section 10.2), we do not assign an additional systematic  
 1035 uncertainty due to cluster selection.



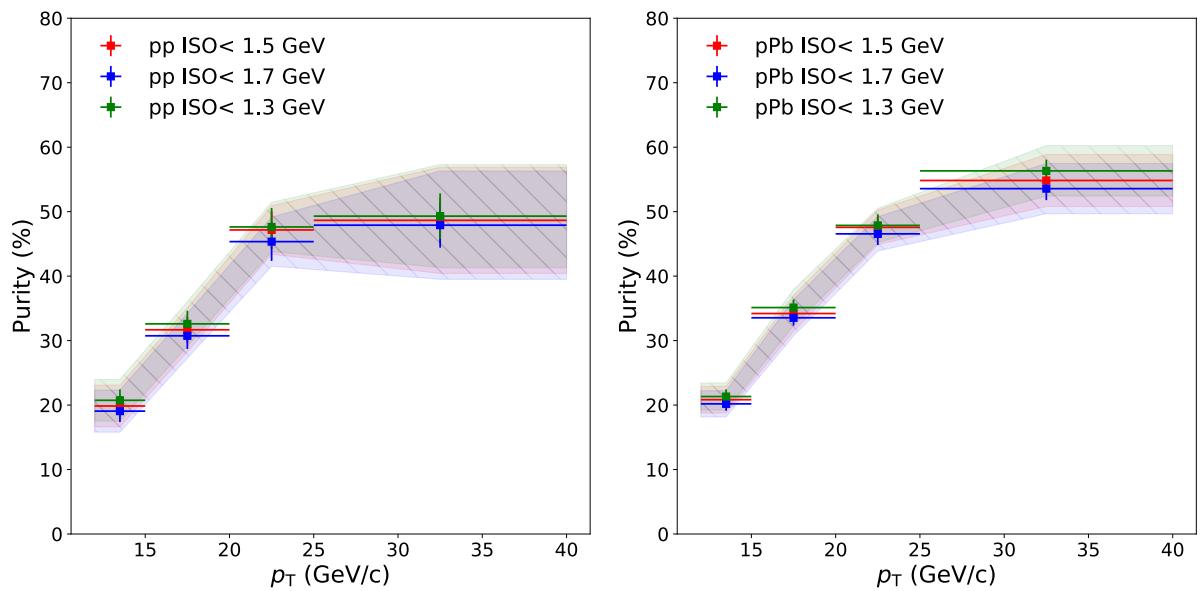
**Fig. E.2:** Purity measurement with distance-to-bad channel  $\geq 1$  (nominal) and  $\geq 2$  in pp (left) and p–Pb (right) data.



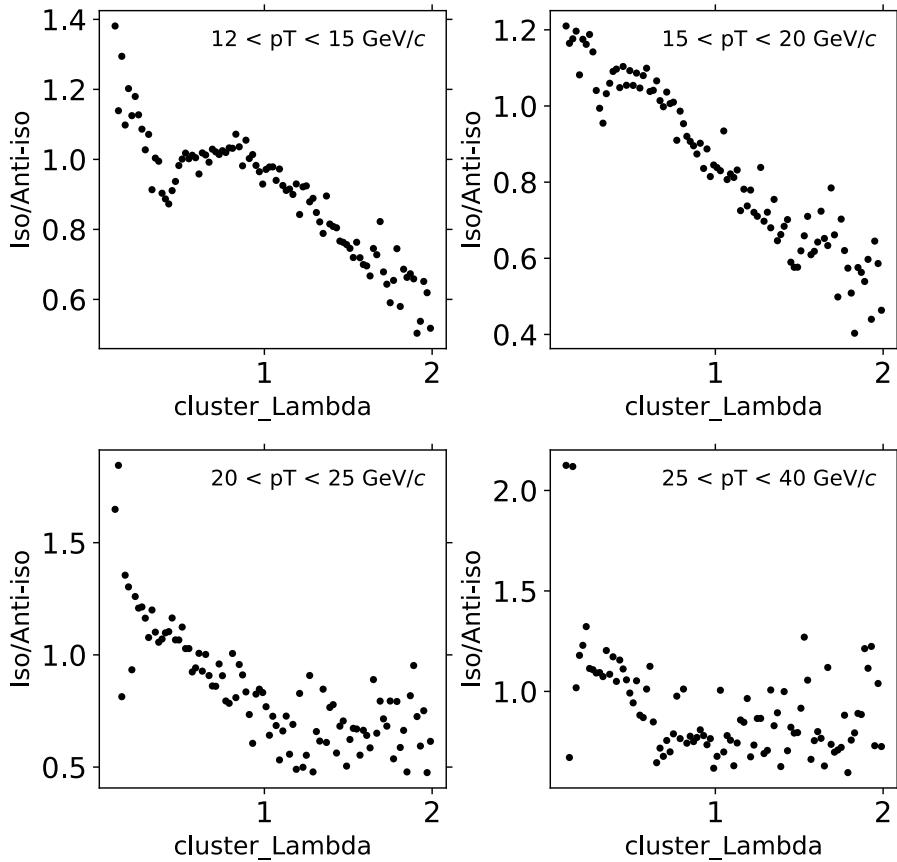
**Fig. E.3:** Purity measurement with and without time cut ( $\Delta t < 20$  ns) in pp (left) and p–Pb (right) data.



**Fig. E.4:** Purity measurement with an exotoxicity cut of 5% (nominal) and 3% in pp (left) and p–Pb (right) data.



**Fig. E.5:** Purity measurement with different isolation threshold requirements in pp (left) and p–Pb (right) data.



**Fig. F.1:** MC-based correction applied to the background shower-shape template for p–Pb collisions in various  $p_T$  ranges.

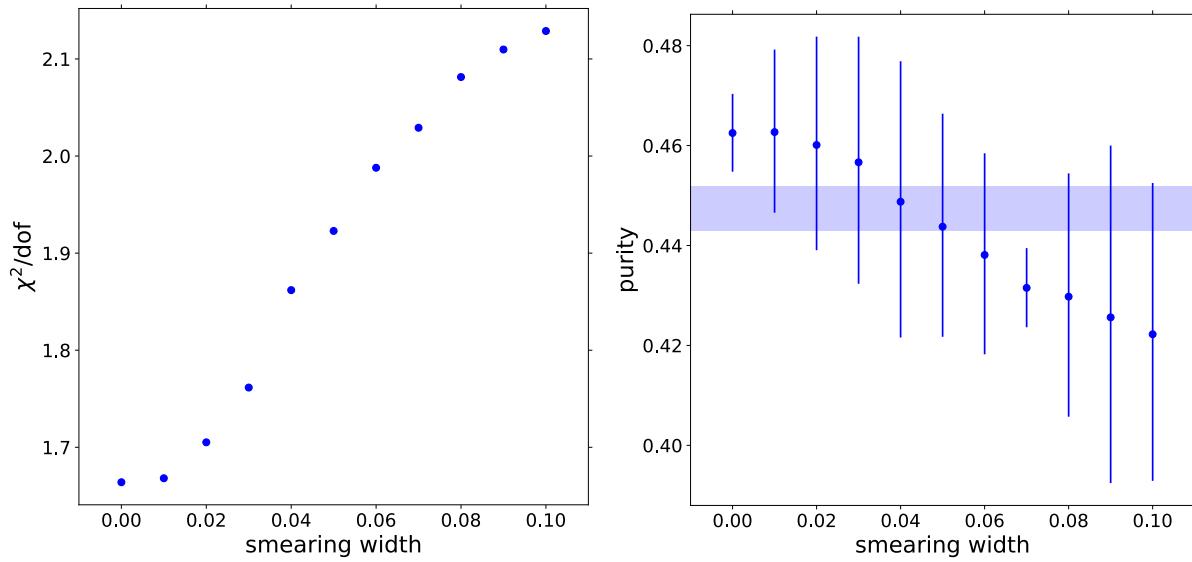
## 1036 F MC-based correction for background template

1037 Figure F.1 shows the weight correction (see Equation 6) applied to the anti-isolation shower-shape dis-  
 1038 tribution, which is obtained from dijet MC simulation. As described in more detail in Section 10, only  
 1039 the shape of this correction is relevant as the normalization is actually fixed in the template fit. The  
 1040 systematic uncertainty associated with this correction is obtained with a double-ratio using data in the  
 1041 background-dominated region ( $\sigma_{\text{long}}^2 > 0.4$ ), as described in more detail in Section 10.2.

## 1042 G Smearing signal template

1043 As part of our investigations to evaluate the template fit systematic uncertainty, we performed studies  
 1044 “smearing” of the signal-shape template. That is, for each cluster we smear the measured  $\sigma_{\text{long}}^2$  by  
 1045 multiplying it by some random number drawn from a Gaussian centered at unity with a given standard  
 1046 deviation “smearing width”.

1047 Figure G.1 shows the purity obtained for  $15 < p_T < 20 \text{ GeV}/c$  as a function of smearing width used  
 1048 and the corresponding  $\chi^2/\text{DOF}$ . We observe rather significant deterioration of the  $\chi^2$  with increasing  
 1049 smearing width, which indicates that smearing beyond  $\approx 3\%$  are strongly disfavoured by the data. To  
 1050 account for the deterioration of the  $\chi^2$  that we observe with the smearing, we multiply the statistical  
 1051 error of the purity by a scaling factor of  $\sqrt{\chi^2/\text{DOF}}$ . We find a trend of decreasing purity with increasing  
 1052 smearing width but this the trend is “not significant”, i.e. it can be accounted by the artificial worsening  
 1053 of the goodness-of-fit.



**Fig. G.1:** Left: Signal template distribution for various smearing widths. Right: purity vs smearing width. The error bar represents the statistical uncertainty, which has been scaled by  $\sqrt{\chi^2/\text{DOF}}$ .

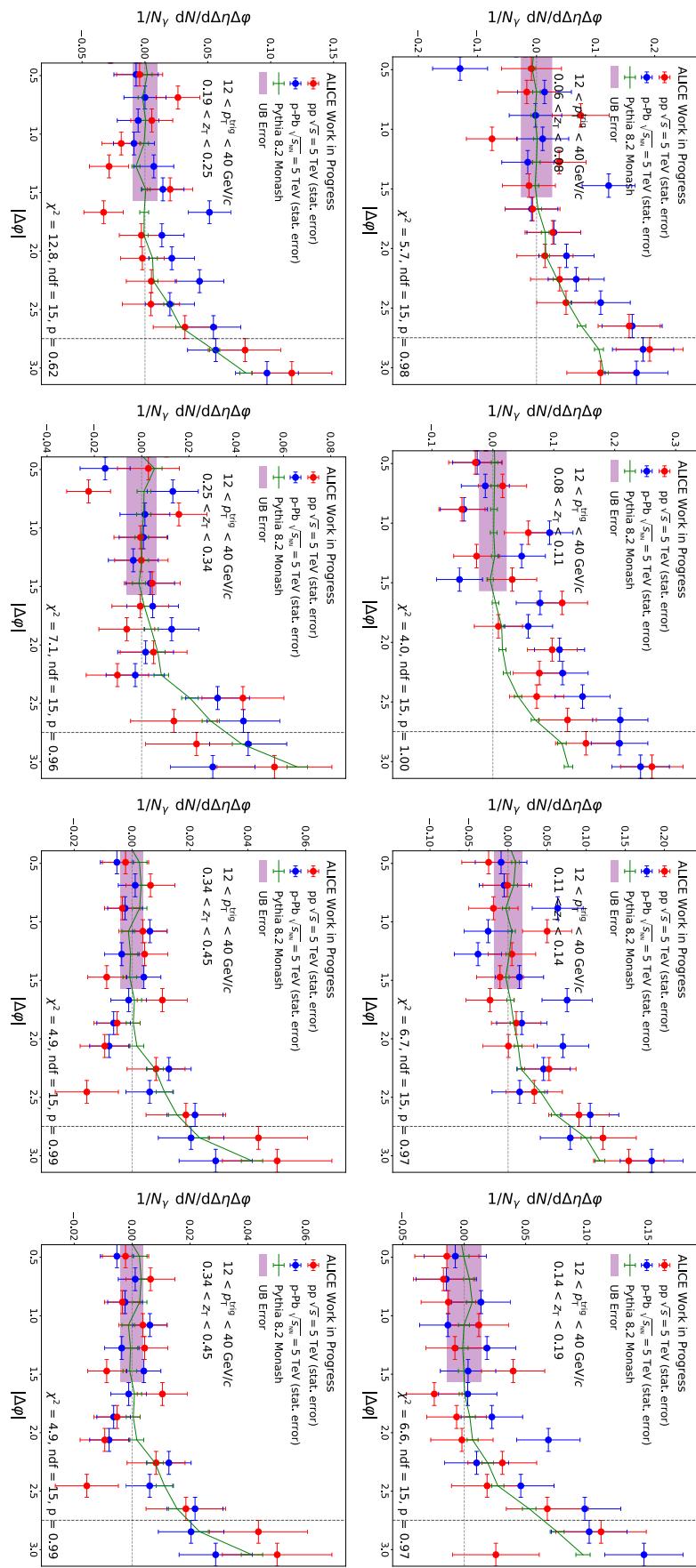
1054 We fit a constant to the purity values vs smearing widths (shown as band in Figure G.1). The error on  
 1055 that constant fit and its deviation of its central value from the nominal result (smearing width equal to  
 1056 zero) is below and absolute 1% difference, which is much smaller than our signal-template uncertainty  
 1057 based on the background-only template fit described in Section 10.2. We reach the same conclusion for  
 1058 all  $p_T$  ranges and for pp and p–Pb data.

## 1059 **H Checking sensitivity to $\Delta\phi$ binning**

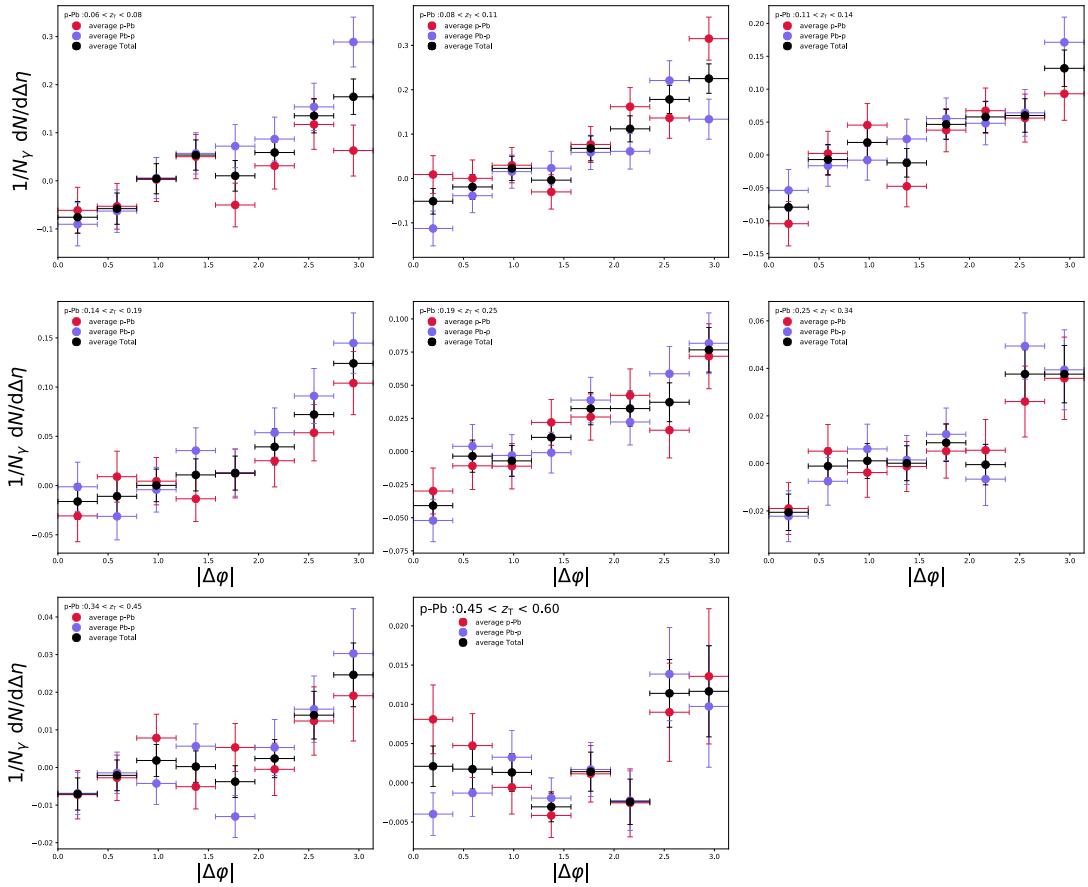
1060 Here we quantify the impact of our choice of binning in the correlation functions. The binning in principle  
 1061 could affect the ZYAM estimate and the shape of the correlation peak. We show results doubling the  
 1062 number of  $\Delta\phi$  bins. We do this for simplicity given that we want to be able to integrate the correlation  
 1063 in exactly the same window as reported in our nominal results ( $\Delta\phi > 7\pi/8$ ). Figure H shows the cor-  
 1064 relation function obtained with double the number of bins as the nominal result. The main conclusion  
 1065 of our study does not change: in every  $z_T$  bin, the pp and p–Pb data are compatible within uncertainties  
 1066 (as quantified in the figure by the  $\chi^2$  and corresponding p-value), and the PYTHIA8 describes the data  
 1067 within uncertainties. The comparison of the ratios is shown in Figure ??.

## 1068 **I Checking Sensitivity to Beam Flip**

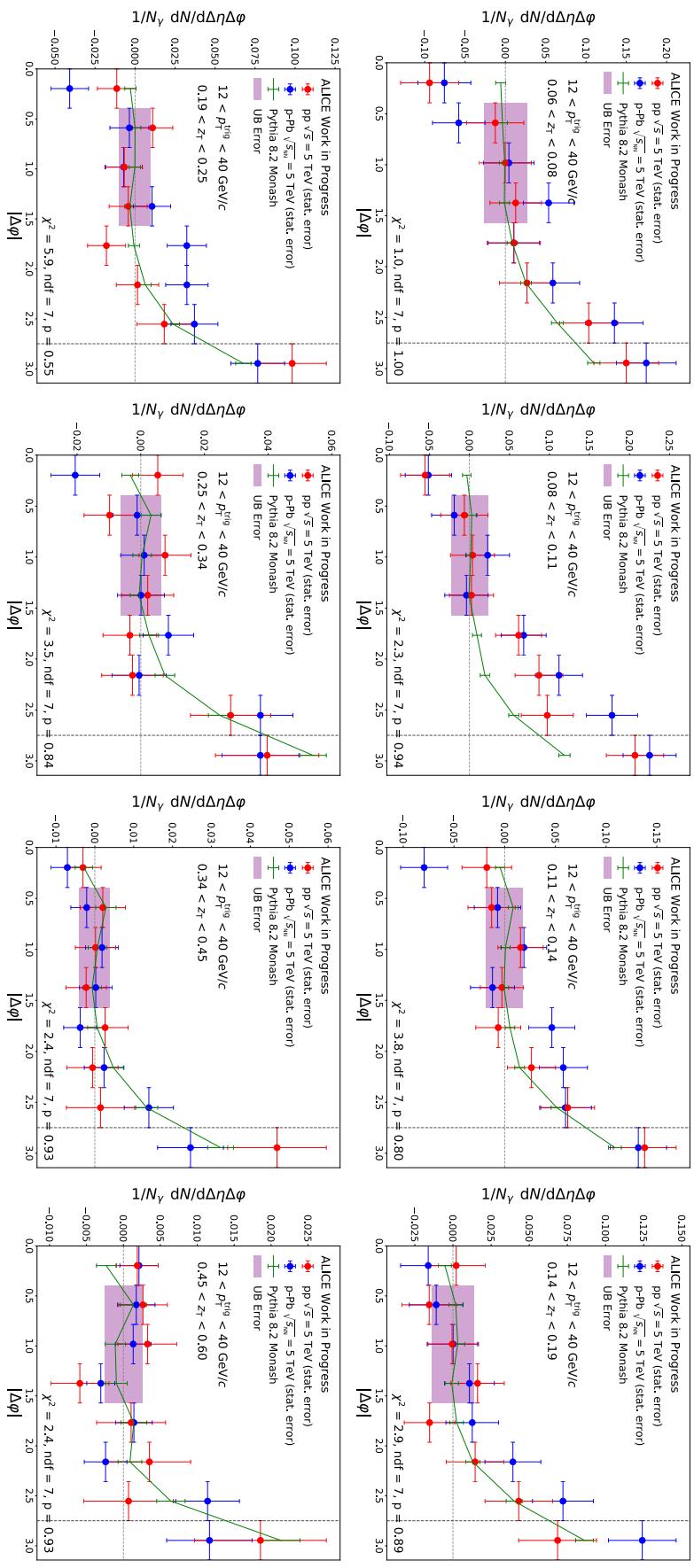
## 1069 **J Correlations including $\Delta\phi = 0$**



**Fig. H.1:** Fully-corrected  $\gamma^{\text{iso}}$ –hadron correlation function pp (red) and p–Pb (blue) data. The purple band represents the uncertainty from the underlying event estimate in pp and p–Pb. The error bars represent statistical uncertainty only. The green line is the  $\gamma^{\text{iso}}$ –hadron correlation function obtained with PYTHIA 8.2. This difference with the nominal results shown in Section 14 is that here we double the number of  $\Delta\phi$  bins.



**Fig. I.1:** Comparison of final correlations functions for p–Pb (red) and Pb–p (blue) collisions.



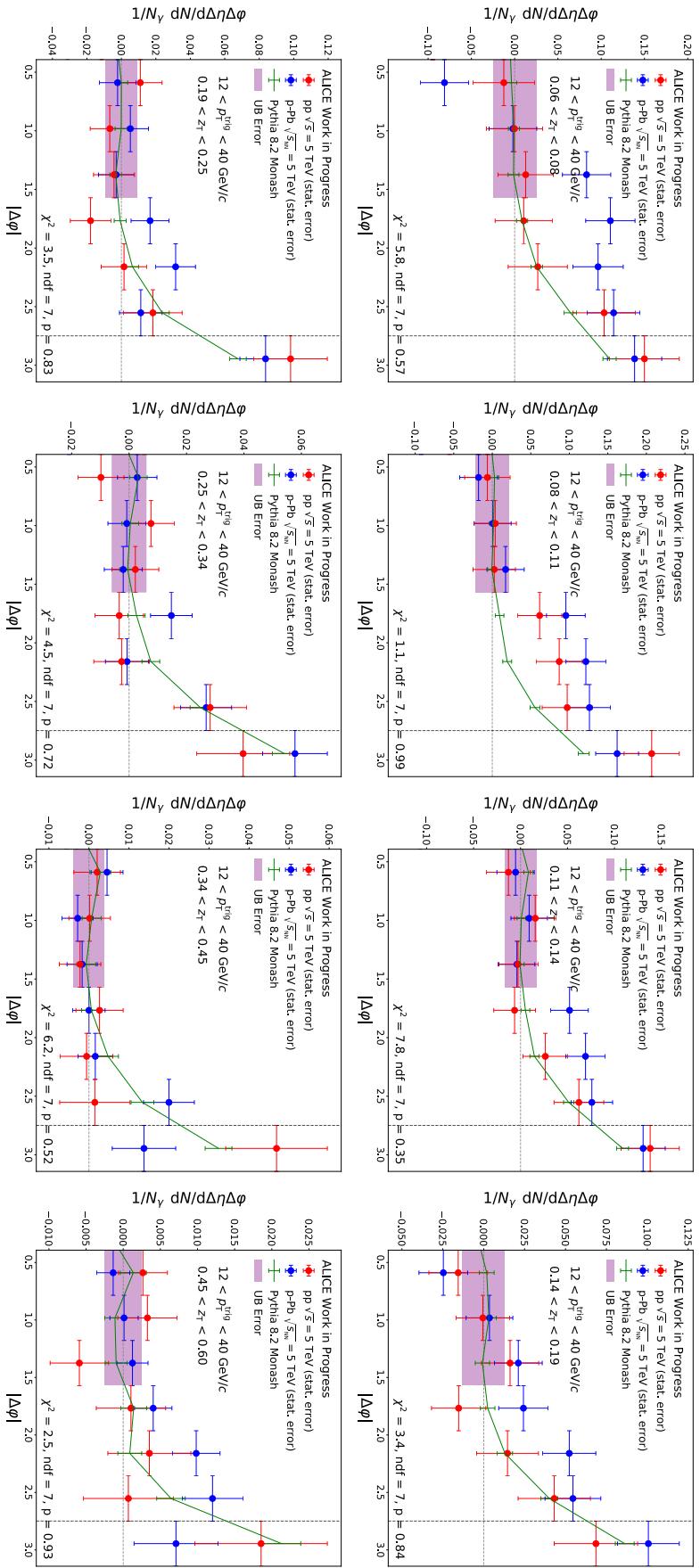
**Fig. J.1:** Fully-corrected  $\gamma^{\text{iso}}$ -hadron correlation function pp (red) and p-Pb (blue) data. Identical to Figure J.1 with the exception that we show the correlations down to  $\Delta\phi$ . After decay-photon background subtraction and isolation a complete depletion of the near-side peak.

**1070 K Comparing ITS only tracking to ITS+TPC Tracking**

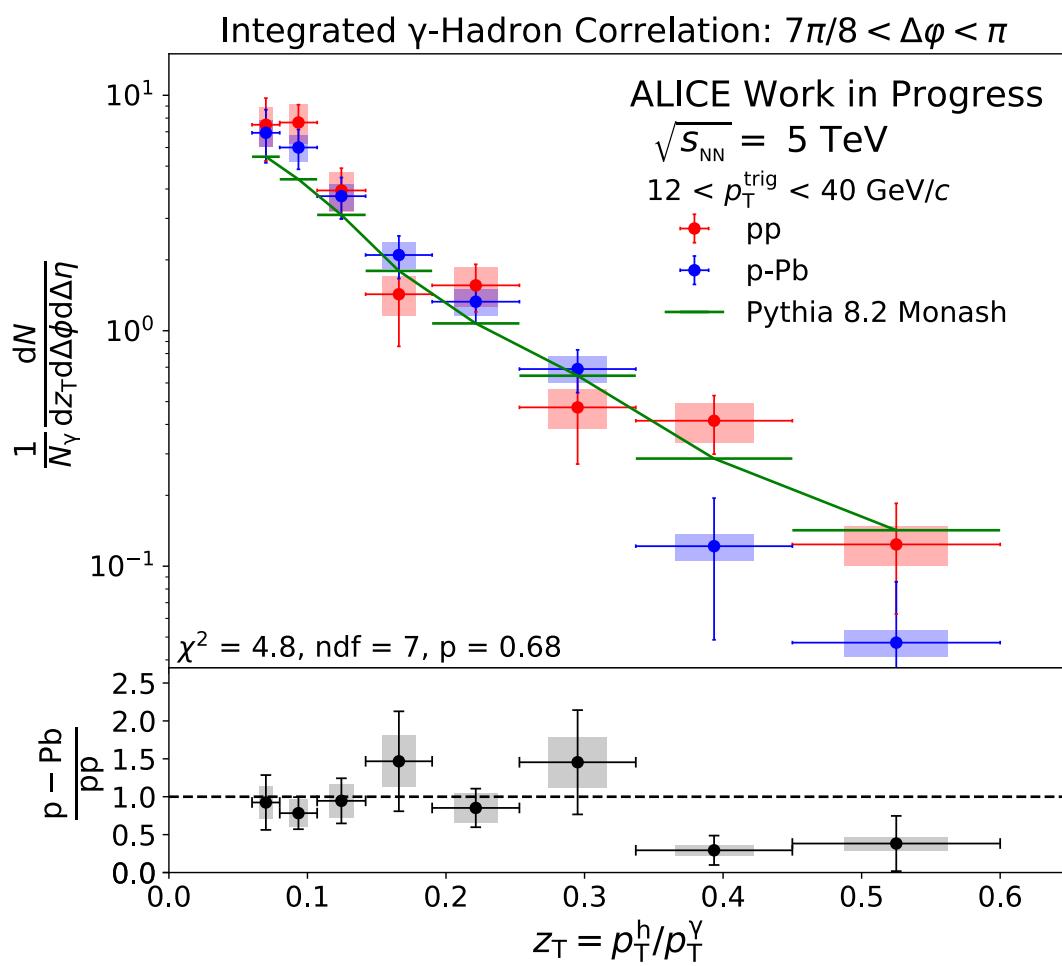
1071 We have performed our analysis with hybrid (ITS + TPC) tracks in the triggered 13def data (p-Pb). Other  
1072 than the change in the appropriate tracking corrections, the analysis chain is identical to the one that uses  
1073 our ITS-only tracking. Figure K.1 shows the correlation functions for p-Pb which uses hybrid tracks and  
1074 pp which only uses ITS tracks (the TPC was not read out during the 17q data taking).

1075 Fig K.2 shows the resulting fragmentation function using TPC+ITS hybrid tracks. It can be seen when  
1076 compared to Figure 36. A more direct comparison is done in Figure K.3.

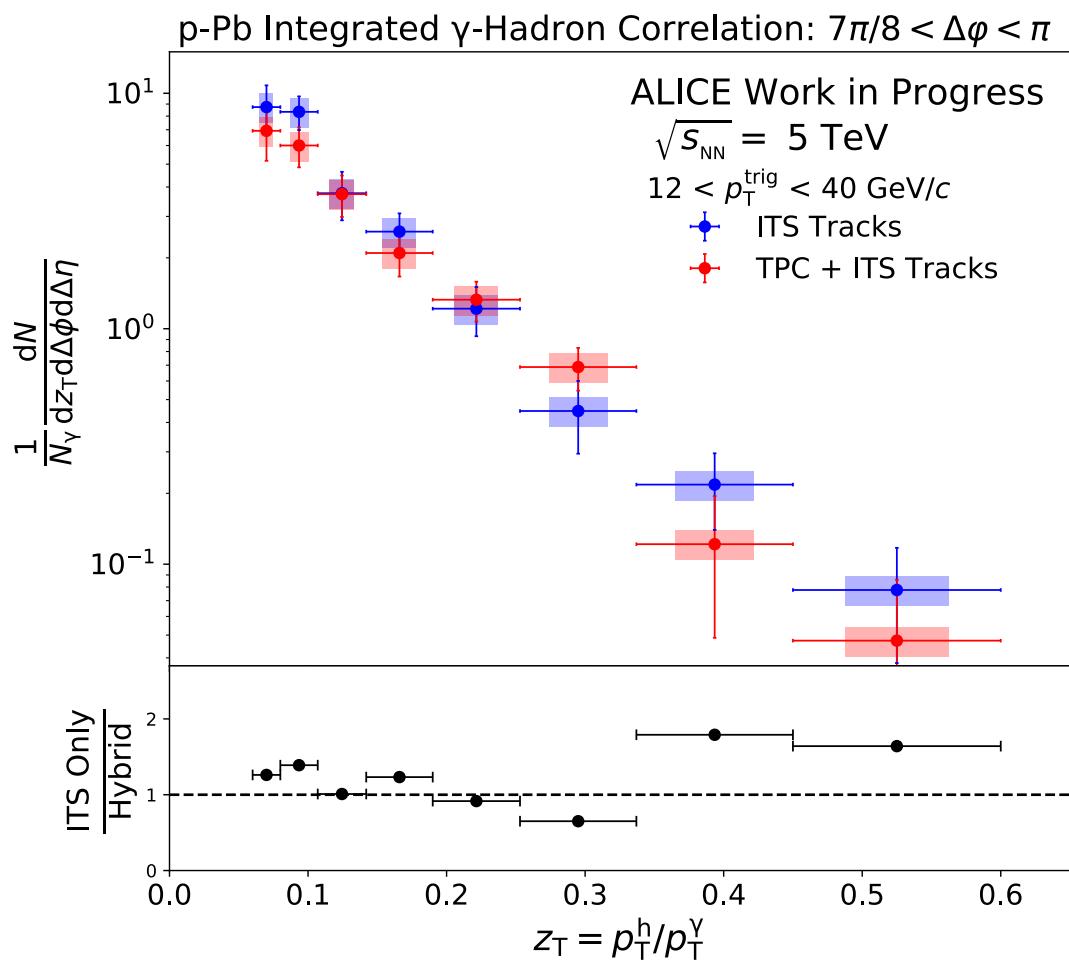
1077 Figure K.3 Shows the fragmentation studies in p–Pbwith ITS only tracks and Hybrid tracks. We can see  
1078 in the ratio of ITS Only/ Hybrid that the hybrid tracks have a significantly lower yield at high  $z_T$ , which  
1079 roughly corresponds to tracks with high  $p_T$ . This is unsurprising as there are issues with TPC charge  
1080 distortions that have been documented at length [<https://alice.its.cern.ch/jira/browse/ATO-351>].



**Fig. K.1:** Fully-corrected  $\gamma_{iso}$ -hadron correlation function in pp using ITS only tracks (red) and p-Pb using hybrid tracks (blue).



**Fig. K.2:** Comparison of final correlations functions for p–Pb Hybrid tracks (blue) and pp ITS only tracks (red).



**Fig. K.3:** Comparison of final correlations functions for p-Pb Hybrid tracks (red) and Pb-p ITS only tracks (blue).

1081 **L Checking sensitivity to  $z_T$  binning**

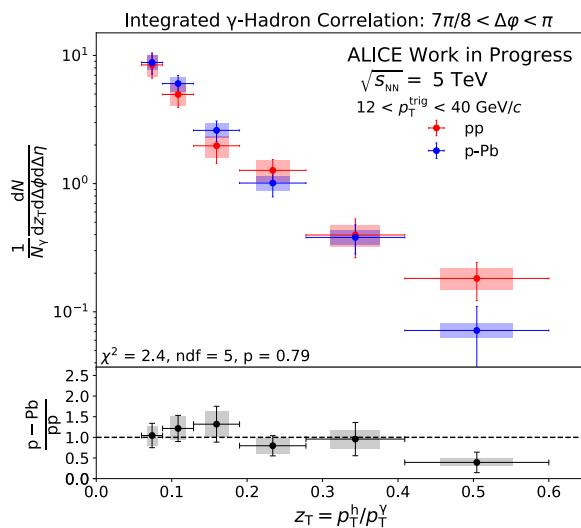
1082 Figure L shows the resulting fragmentation function in 6  $z_T$  bins. The conclusion we draw from the data  
1083 does not affect by this change.

1084 **M Checking Maximum Track  $p_T$  Cut**

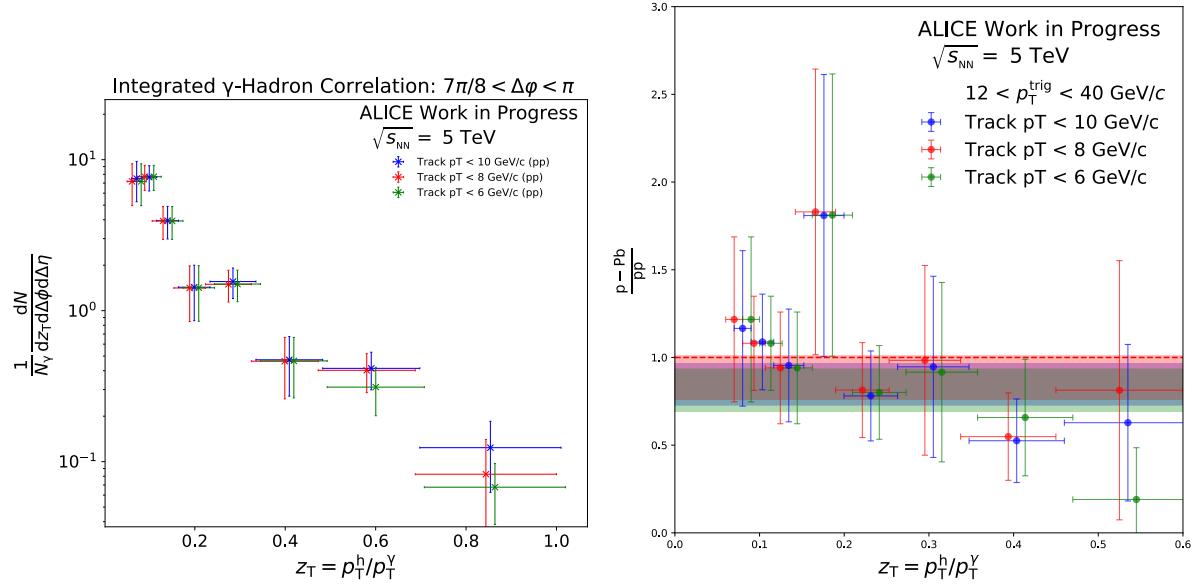
1085 In this section we show the effect of varying the maximum  $p_T$  cut on the final away-side yields in pp and  
1086 p–Pb. We then show the effect this variation has on the final ratio.

1087 **N Sensitivity to  $\chi^2/\text{ITS}_{\text{clus}}$  in p–Pb**

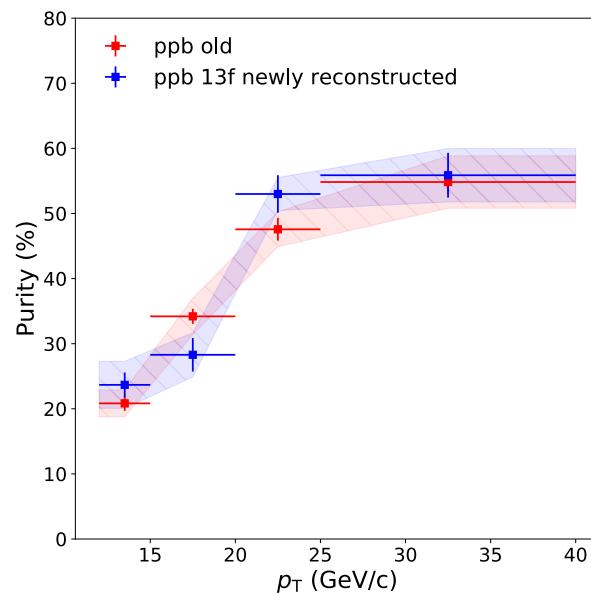
1088 The effect of the  $\chi^2/\text{ITS}_{\text{clus}}$  cut on the tracking efficiency, fake rate, and bin migration is studied to select  
1089 the most effective  $\chi^2/\text{ITS}_{\text{clus}}$  maximum limit. Four values have been tested:  $\chi^2/\text{ITS}_{\text{clus}} = 1, 2, 3$ , and  
1090 36. The following figures N.1, ??, ??, and ?? show the effect of  $\chi^2/\text{ITS}_{\text{clus}}$  cuts on the efficiency, fake  
1091 rate, bin migration, and track correction weight. The weight is calculated from the efficiency, fake rate,  
1092 and bin migration using equation 12. The efficiency seems to be similar for all cut variation except for  
1093  $\chi^2/\text{ITS}_{\text{clus}} \geq 1$ , where it is lower by 5%. However, the fake rate and bin migration effects are improved as  
1094 the cut is tightened. It is to note that regardless of the threshold for the  $\chi^2/\text{ITS}_{\text{clus}}$  cut, as long as identical  
1095 cuts are applied to both the Monte Carlo and the data sets used in the section 11, the fine agreement with  
1096 published data is always present.



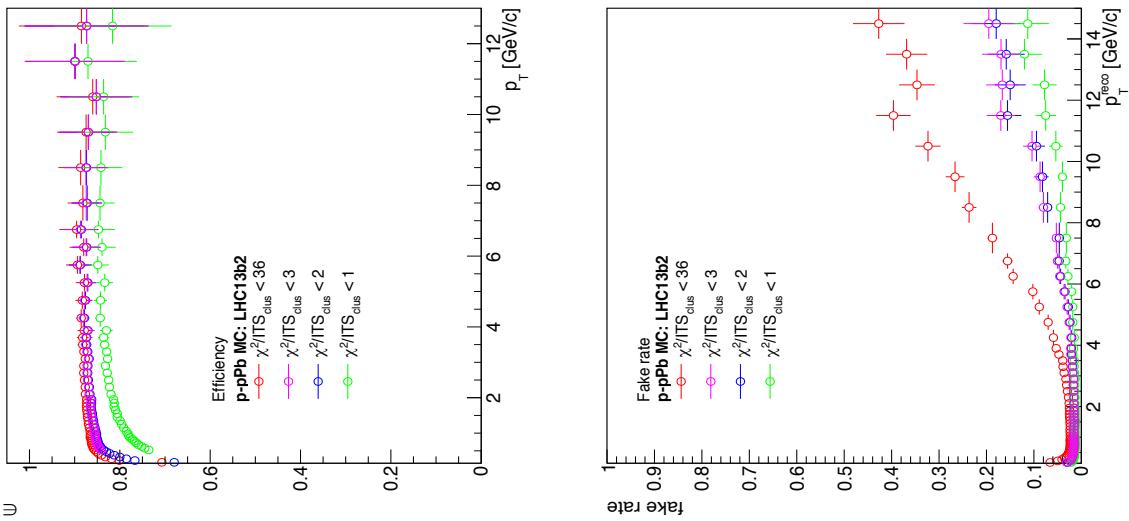
**Fig. L.1:** Results with variation of numbers of  $z_T$  bins.



**Fig. M.1:** Left: The away side yields in pp with a maximum  $p_T^{\text{Track}}$  cut of 10,8, and 6 GeV/c. Right: The Ratio of p-Pb to pp with a maximum  $p_T^{\text{Track}}$  cut of 10,8, and 6 GeV/c



**Fig. M.2:** Nominal results for purity and purity measured with the newly reconstructed 13f runs.



**Fig. N.1:** Left: The efficiency comparison for the various  $\chi^2/\text{ITS}_{\text{clus}}$  cuts. Right: The fake rate comparison for the various  $\chi^2/\text{ITS}_{\text{clus}}$  cuts

## Bibliography

- [1] Jaroslav Adam et al. Measurement of the production of high- $p_T$  electrons from heavy-flavour hadron decays in Pb-Pb collisions at  $\sqrt{s_{\text{NN}}} = 2.76$  TeV. *Phys. Lett.*, B771:467–481, 2017.
- [2] ALICE Collaboration. Measurement of the inclusive isolated photon production cross section in pp collisions at  $\sqrt{s} = 7$  TeV. *Submitted to: Eur. Phys. J.*, 2019.
- [3] J. Allen et al. Performance of prototypes for the ALICE electromagnetic calorimeter. *Nucl. Instrum. Meth.*, A615:6–13, 2010.
- [4] J. Allen et al. ALICE DCal: An Addendum to the EMCal Technical Design Report Di-Jet and Hadron-Jet correlation measurements in ALICE. CERN-LHCC-2010-011, ALICE-TDR-14-add-1, 2010.
- [5] F. Arleo et al. Hard probes in heavy-ion collisions at the LHC: Photon physics in heavy ion collisions at the LHC. *arXiv:hep-ph/0311131*, 2004.
- [6] Gustavo Conesa Balbastre. Study and emulation of cell energy cross-talk in EMCal . ANA-837, 2018.
- [7] Matteo Cacciari, Gavin P. Salam, and Sebastian Sapeta. On the characterisation of the underlying event. *JHEP*, 04:065, 2010.
- [8] ALICE Collaboration. Underlying Event measurements in  $pp$  collisions at  $\sqrt{s} = 0.9$  and 7 TeV with the ALICE experiment at the LHC. *JHEP*, 07:116, 2012.
- [9] ALICE Collaboration. Performance of the ALICE Experiment at the CERN LHC. *Int. J. Mod. Phys.*, A29:1430044, 2014.
- [10] ALICE Collaboration.  $\pi^0$  and  $\eta$  meson production in proton-proton collisions at  $\sqrt{s} = 8$  TeV. *Eur. Phys. J.*, C78(3):263, 2018.
- [11] ALICE Collaboration. ALICE EMCal Physics Performance Report. *arXiv:1008.0413*, 2010.

- 1120 [12] ALICE Collaboration. Transverse momentum spectra and nuclear modification factors of charged  
1121 particles in pp, p-Pb and Pb-Pb collisions at the LHC. arXiv:1802.09145, 2018.
- 1122 [13] ALICE Collaboration. Direct photon production at low transverse momentum in proton-proton  
1123 collisions at  $\sqrt{s} = 2.76$  and 8 TeV. arXiv:1803.09857, 2018.
- 1124 [14] CMS Collaboration. Studies of jet quenching using isolated-photon+jet correlations in PbPb and  
1125  $pp$  collisions at  $\sqrt{s_{NN}} = 2.76$  TeV. *Phys. Lett.*, B718:773–794, 2013.
- 1126 [15] CMS Collaboration. Study of jet quenching with isolated-photon+jet correlations in PbPb and pp  
1127 collisions at  $\sqrt{s_{NN}} = 5.02$  TeV. arXiv:1711.09738, 2017.
- 1128 [16] CMS Collaboration. Measurement of differential cross sections for inclusive isolated-photon and  
1129 photon+jets production in proton-proton collisions at  $\sqrt{s} = 13$  TeV. arXiv:1807.00782, 2018.
- 1130 [17] D. Gale and L. S. Shapley. College admissions and the stability of marriage. *The American Mathematical Monthly*, 69(1):9–15, 1962;.
- 1132 [18] F. James and M. Roos. Minuit: A System for Function Minimization and Analysis of the Parameter  
1133 Errors and Correlations. *Comput. Phys. Commun.*, 10:343–367, 1975.
- 1134 [19] Erwann Masson. Measurement of isolated photons in pPb collisions at 5.02 TeV with the EM-  
1135 Cal detector in ALICE. <https://alice-notes.web.cern.ch/system/files/notes/analysis/741/>. ALICE-  
1136 ANA-2018-4677.