



Draft version 2



ATLAS NOTE

October 17, 2018

1 Measurement of multi-particle azimuthal correlation in proton-lead and 2 proton-proton collisions with the ATLAS detector

3 Mingliang Zhou^a, Jiangyong Jia^{a,b}, Adam Trzupek^c, Peng Huo^a

4 ^a*Department of Chemistry, Stony Brook University, Stony Brook, NY 11794, USA*

5 ^b*Brookhaven National Laboratory, Physics Department, Bldg. 510A, Upton, NY 11973, United States of America*

6 ^c*Institute of Nuclear Physics Polish Academy of Sciences, Krakow*

7 Abstract

8 This note presents details of the analysis of the multi-particle azimuthal anisotropy of
9 charged particles produced in pp and $p+Pb$ collisions with the new sub-event multi-particle
10 cumulant method. Cumulant has previously been measured in $p+Pb$ and $Pb+Pb$ collision
11 systems, and recently was first measured in pp collision by CMS collaboration. In this note
12 we will show that the traditional cumulant measurement is very sensitive to the multiplicity
13 fluctuation, and existed results might not indicate collectivity even with presence of negative
14 $C_2\{4\}$. For this reason, we proposed a new sub-event method within the cumulant frame-
15 work to further suppress the non-flow in small systems. This new method has been validated
16 in PYTHIA sample, where $C_2\{4\}$ calculated from the new method is much closer to 0 than
17 the traditional method. The 4-particle cumulant are calculated for high multiplicity pp data
18 at $\sqrt{s} = 13$ TeV as well as 5.02 TeV. Results from the new method give negative $C_2\{4\}$
19 down to 40 tracks, and the values of $C_2\{4\}$ are independent of multiplicity fluctuation. The
20 new results indicate potential collective behavior in small collision system.

21 **Contents**

22 1	Introduction	4
23 2	Sub-event multi-particle cumulants	5
24 2.1	Basic idea of subevent cumulant	5
25 2.2	Traditional cumulant	5
26 2.3	2 sub-event method	6
27 2.4	3 sub-event method	6
28 3	Data samples, event and tracking selections	8
29 3.1	<i>pp</i> 13 TeV	8
30 3.1.1	Run period 1	8
31 3.1.2	Run period 2	11
32 3.1.3	Run period 3	13
33 3.1.4	Run period 4	15
34 3.1.5	PYTHIA for <i>pp</i> 13 TeV data	17
35 3.1.6	Tracking validation for run 305359, 309314 and 309346	22
36 3.2	<i>pp</i> 5.02 TeV	27
37 3.2.1	PYTHIA for <i>pp</i> 5.02 TeV data	29
38 3.2.2	Impact of ID misalignment	29
39 3.3	<i>p+Pb</i> 5.02 TeV	30
40 3.3.1	2013 <i>p+Pb</i>	30
41 3.3.2	HIJING for 2013 <i>p+Pb</i> data	35
42 3.3.3	2016 <i>p+Pb</i>	35
43 3.3.4	HIJING for 2016 <i>p+Pb</i> data	38
44 4	Data analysis	41
45 4.1	Outline	41
46 4.2	<i>p_T</i> range of the reference particles N_{ch}^{ref}	41
47 4.3	Calculation of <i>Q</i> -vector	41
48 4.4	Flattening procedure	43
49 4.5	2- and 4-particle correlations with direct cumulant method	47
50 4.6	Definition of event classes	47
51 4.7	X-axis of cumulant results	50
52 4.8	Event weighting	52
53 4.9	2- and 4-particle cumulant	53
54 4.10	Validation of the subevent method using PYTHIA truth particles	53
55 5	Systematic uncertainties and cross-checks	55
56 5.1	<i>pp</i> 13 TeV	55
57 5.1.1	Trigger efficiency	55
58 5.1.2	Tracking efficiency	57
59 5.1.3	Track selection	59
60 5.1.4	Pile-up condition	60
61 5.1.5	Residual detector effects	61
62 5.1.6	Event class bin width	63
63 5.1.7	η gap for sub-event	65
64 5.1.8	Charge dependence	66

65	5.1.9	Deviations from the truth in the reconstruction of Monte Carlo data	67
66	5.1.10	Summary of systematics	69
67	5.2	pp 5.02 TeV	71
68	5.2.1	Summary of systematics	71
69	5.3	$p+Pb$ 5.02 TeV	72
70	5.3.1	Pileup rejection in run 1	72
71	5.3.2	Summary of systematics	73
72	5.4	Summary of systematic table	74
73	6	Results	76
74	6.1	13 TeV pp $C_2\{2\}$	76
75	6.2	13 TeV pp $C_3\{2\}$	78
76	6.3	13 TeV pp $C_2\{4\}$	80
77	6.4	13 TeV pp $C_3\{4\}$	82
78	6.5	5.02 TeV pp $C_2\{2\}$	84
79	6.6	5.02 TeV pp $C_3\{2\}$	86
80	6.7	5.02 TeV pp $C_2\{4\}$	88
81	6.8	5.02 TeV pp $C_3\{4\}$	90
82	6.9	5.02 TeV $p+Pb$ $C_2\{2\}$	92
83	6.10	5.02 TeV $p+Pb$ $C_3\{2\}$	94
84	6.11	5.02 TeV $p+Pb$ $C_2\{4\}$	96
85	6.12	5.02 TeV $p+Pb$ $C_3\{4\}$	98
86	6.13	Comparison among three collision systems	100
87	6.14	Comparison with 2PC results	101
88	6.15	Number of sources N_s in the initial state geometry	102
89	7	Summary	104
90	8	Appendix	105
91	8.1	Statistics of all triggers in pp 13 TeV	105
92	8.2	Statistics of all triggers in pp 5.02 TeV	109
93	8.3	Auxiliary plots 13 TeV PYTHIA	109
94	8.4	Formulas for direct cumulant with particle weight	113
95	8.4.1	Traditional method	113
96	8.4.2	2 sub-event method: 1 st kind	116
97	8.4.3	2 sub-event method: 2 nd kind	117
98	8.4.4	3 sub-event method: 1 st kind	117
99	8.4.5	3 sub-event method: 2 nd kind	118

100 List of contributions

101 • Mingliang Zhou

- 102 – Submit production jobs for 13 TeV pp , 5.02 TeV pp and 5.02 TeV $p+Pb$;
- 103 – Development and optimization of the method for sub-event cumulant method;
- 104 – Evaluation of the systematic uncertainties in all systems;
- 105 – Writing the internal note;

106 • Jiangyong Jia

- 107 – Development and optimization of the method for sub-event cumulant method;
- 108 – Evaluation of the systematic uncertainties in all systems;
- 109 – Producing the final physics plots;
- 110 – Writing the paper draft;

111 • Adam Trzupek

- 112 – Analyzer for the results and systematics from traditional cumulant method;

113 • Peng Huo

- 114 – Submit production jobs for 13 TeV pp , 5.02 TeV pp and 5.02 TeV $p+Pb$;
- 115 – Cross-check of the subevent analysis code;

¹¹⁶ 1 Introduction

¹¹⁷ High energy heavy-ion collisions at RHIC and LHC create a strongly-interacting nuclear matter that
¹¹⁸ exhibit many interesting characteristics. One of which is the collimated emission of particle pairs with
¹¹⁹ small azimuthal-angle separation, $\Delta\phi$, that extends over large range of pseudorapidity differences, $\Delta\eta$.
¹²⁰ This so called "ridge" correlation was first observed in A+A collisions [1, 2, 3, 4, 5, 6], but later was
¹²¹ also observed in proton-nucleus and light-ion-nucleus collisions [7, 8, 9, 10, 11, 12], and more recently
¹²² also observed in high multiplicity proton-proton collisions [13, 14, 15, 16]. In A+A collisions, the ridge
¹²³ is believed to be the consequence of collective emission of particles in the azimuthal direction, and
¹²⁴ the collectivity is generated in the final state after local thermalization, described by relativistic viscous
¹²⁵ hydrodynamic models. For small system such as pp and $p+A$ collisions, the origin of the ridge is less
¹²⁶ clear. There are considerable debates in the theoretical community on whether the ridge in small systems
¹²⁷ is of hydrodynamic origin similar to A+A collisions [17] or it is created in the initial state due to gluon
¹²⁸ saturation effects [18].

¹²⁹ The ridge signal from two-particle correlation (2PC) is characterized by a Fourier decomposition
¹³⁰ $\sim 1 + 2v_n^2 \cos(n\Delta\phi)$, where the v_n denotes the single-particle anisotropy harmonics. The second-order
¹³¹ coefficient v_2 is by far the largest, followed by v_3 . In small collision systems, the extraction of ridge
¹³² signal requires a careful removal of a large contribution from dijets, which is estimated from 2PC in very
¹³³ low multiplicity events and then subtracted from higher multiplicity events [8, 9, 11, 14, 15, 19]. On the
¹³⁴ other hand, since collectivity is intrinsically a multi-particle phenomenon, it can be probed more directly
¹³⁵ using multi-particle correlation (or cumulant) technique [20]. One of the perceived hallmark feature
¹³⁶ of collectivity is the observation of positively defined signal from $2k$ -particle correlation $v_2\{2k\}, k \geq 2$.
¹³⁷ The $v_2\{4\}$ and $v_2\{6\}$ have been measured in high-multiplicity pp and $p+Pb$ collisions [12, 16, 19, 21].
¹³⁸ However, this perception is quite misleading in small collision systems, where the non-flow correlation
¹³⁹ can be as large or bigger than genuine long-range ridge correlation. Recently, an improved cumulant
¹⁴⁰ method based on η -separated sub-events has been proposed to further reduce the non-flow correlations,
¹⁴¹ in particular from jets and dijets [22]. The performance of the method in suppressing non-flow correlations
¹⁴² has been validated using PYTHIA 8 model, which contains no long-range collective effects.

¹⁴³ Multi-particle cumulants suppress jet and short-range correlations, but not completely removing it. In
¹⁴⁴ fact $v_2\{4\}$ is observed to flip sign at smaller N_{ch} , number of charged particle, in pp and $p+Pb$ collisions.
¹⁴⁵ Recently ATLAS observed that the N_{ch} value where sign-flip happens and the magnitude of $v_2\{4\}$ depend
¹⁴⁶ also on how the event class are chosen for the calculation of cumulants. In this paper, we show that the
¹⁴⁷ choice of event class influences the probability of non-flow, and consequently the non-flow distribution
¹⁴⁸ to the $v_2\{4\}$. An improved cumulant method is proposed to further suppress non-flow, and therefore
¹⁴⁹ reduce the sensitivity to non-flow fluctuations. In this method, cumulants are constructed from particles
¹⁵⁰ in several subevents separated in η , which allows measurement of multi-particle correlations to much
¹⁵¹ lower N_{ch} .

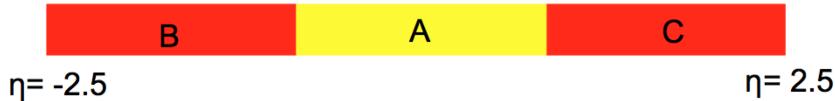
¹⁵² The structure of the internal note is as follows. In section 2 subevent cumulant will be briefly in-
¹⁵³ troduced, with all the formulas listed. In section 3, data sets together with event and track selections
¹⁵⁴ will be described. Detailed procedures of the analysis will be presented in section 4. In section 5, we
¹⁵⁵ will discuss all the primary systematics and cross-checks. In section 6, all the physics plots in different
¹⁵⁶ collision systems will be shown. Section 7 summarizes the whole analysis.

¹⁵⁷ 2 Sub-event multi-particle cumulants

¹⁵⁸ 2.1 Basic idea of subevent cumulant

¹⁵⁹ In this analysis we focus mainly on four-particle subevent cumulant, but the generalization to higher-
¹⁶⁰ order is straightforward. The basic idea for subevent cumulants is shown in Fig. 1. The event is equally
¹⁶¹ divided into three non-overlapping rapidity ranges, labelled as A , B and C . In this analysis, A is defined
¹⁶² with $-2.5/3 < \eta < 2.5/3$, B is defined with $-2.5 < \eta < -2.5/3$ and C is defined with $2.5/3 < \eta < 2.5$.
¹⁶³ Four-particle correlators are then constructed by choosing two particles from subevent A and one particle
¹⁶⁴ each from subevents B and C . All other permutations are also considered to increase the number of pairs
¹⁶⁵ (e.g. two particles from subevent B and one particle each from subevent A and C). Dijets contributions,
¹⁶⁶ the main source for non-flow, are suppressed, since they can only produce particles in each subevent. For
¹⁶⁷ comparison, we also consider cumulants based on two subevents as shown in Fig. 1. In this case, two
¹⁶⁸ particles each are chosen from A and B , which effectively suppress contribution from single jet, but not
¹⁶⁹ away-side jets when it lands in different subevents. For the 2 subevent, in this analysis, A is defined with
¹⁷⁰ $\eta < 0$ and B is defined with $\eta > 0$.

(a) Three subevents



(b) Two subevents



Figure 1: The η ranges for the subevents in three-subevent method (a) and two-subevent method (b).

¹⁷¹ In this section, we will quickly go over all the formulas used for the analysis. Detailed descriptions
¹⁷² will be discussed in the reference method paper [22].

¹⁷³ 2.2 Traditional cumulant

¹⁷⁴ Compared with newly developed subevent cumulant, the existing cumulant method is referred as tradi-
¹⁷⁵ tional cumulant, where the 4-particle cumulant is defined as:

$$C_n\{4\} \equiv \langle 4 \rangle - 2 \langle 2 \rangle^2 \quad (1)$$

¹⁷⁶ Using the direct or Q-cumulant technique, 4-particle correlation $\langle 4 \rangle$ and 2-particle correlation $\langle 2 \rangle$ are
¹⁷⁷ defined as [23, 24]:

$$\begin{aligned} \langle 2 \rangle &\equiv \frac{|Q_n|^2 - M}{M(M-1)} \\ \langle 4 \rangle &\equiv \frac{|Q_n|^4 + |Q_{2n}|^2 - 2\text{Re}(Q_{2n}Q_n^*Q_n^*) - 4(M-2)|Q_n|^2 + 2M(M-3)}{M(M-1)(M-2)(M-3)} \end{aligned} \quad (2)$$

¹⁷⁸ where M is the multiplicity in each event and Q_n is calculated event-by-event:

$$Q_n \equiv \sum_i e^{in\phi_i} \quad (3)$$

¹⁷⁹ and without special mention, in this analysis, only the real part of the inner products (e.g. $Q_{2n}^2 -$
¹⁸⁰ $2Q_{2n}Q_n^*Q_n^*$) is taken.

¹⁸¹ 2.3 2 sub-event method

¹⁸² The cumulant is defined in a way that the non-flow sources associated with lower number of particles will
¹⁸³ be subtracted. Since in the framework of sub-event cumulant, the way non-flow sources are correlated is
¹⁸⁴ changed, the corresponding cumulant definition also needs to be changed:

$$C_n^{a,a|b,b} \{4\} \equiv \langle 4 \rangle_{a,a|b,b} - 2 \langle 2 \rangle_{a|b}^2 \quad (4)$$

¹⁸⁵ where subscript $a, a|b, b$ defines how the four particles are arranged in the two sub-events. In this type,
¹⁸⁶ particles in the same sub-events are always correlated with particles in the other sub-events. In other
¹⁸⁷ words, particles in the same sub-events will not be correlated with each other, so that the short-range
¹⁸⁸ non-flow sources are suppressed. There will be an alternative arrangement for 2 sub-event cumulant:
¹⁸⁹ $a, b|a, b$. In this analysis, we will only discuss the first type $a, a|b, b$ since there are less residual non-flow
¹⁹⁰ contribution compared with the other type.

¹⁹¹ Following similar direct cumulant technique, 4-particle correlation $\langle 4 \rangle_{a,a|b,b}$ and 2-particle correlation
¹⁹² $\langle 2 \rangle_{a|b}$ can be calculated in a single event loop:

$$\begin{aligned} \langle 2 \rangle_{a|b} &\equiv \frac{Q_{n,a}Q_{n,b}}{M_aM_b} \\ \langle 4 \rangle_{a,a|b,b} &\equiv \frac{(Q_{n,a}^2 - Q_{2n,a})(Q_{n,b}^2 - Q_{2n,b})^*}{M_a(M_a-1)M_b(M_b-1)} \end{aligned} \quad (5)$$

¹⁹³ where M_a and M_b are number of particles in each sub-event. $Q_{n,a}$ and $Q_{n,b}$ are the Q-vector in each
¹⁹⁴ sub-event, where the definition is same as the traditional cumulant.

¹⁹⁵ 2.4 3 sub-event method

¹⁹⁶ Similar as the 2 sub-event method, depending on how the particles are correlated between sub-events,
¹⁹⁷ there are two types for the cumulant definition: $C_n^{a,a|b,c} \{4\}$ and $C_n^{a,b|a,c} \{4\}$. For the same reason as 2
¹⁹⁸ sub-event method, we only consider the first type, where there is no correlation within the sub-event.

¹⁹⁹ The definition of the 4-particle 3 sub-event cumulant $C_n^{a,b|a,c}$ is:

$$C_n^{a,a|b,c} \{4\} \equiv \langle 4 \rangle_{a,a|b,c} - 2 \langle 2 \rangle_{a|b} \langle 2 \rangle_{a|c} \quad (6)$$

²⁰⁰ where subscript $a, b|a, c$ defines how the four particles are arranged in the two sub-events: 2 particles
²⁰¹ from sub-event A and the other two from B and C separately. It is worth mentioning that there are two
²⁰² unique combinations by rotating the sub-event A, B and C, which will contribute to the total number of
²⁰³ pairs.

²⁰⁴ Similarly, 4-particle correlation $\langle 4 \rangle_{a,a|b,c}$ and 2-particle correlation $\langle 2 \rangle_{a|b}$ and $\langle 2 \rangle_{a|c}$ can be calculated
²⁰⁵ using the direct cumulant technique:

$$\begin{aligned}\langle 2 \rangle_{a|b} &\equiv \frac{Q_{n,a} Q_{n,b}}{M_a M_b} \\ \langle 2 \rangle_{a|c} &\equiv \frac{Q_{n,a} Q_{n,c}}{M_a M_c} \\ \langle 4 \rangle_{a,a|b,c} &\equiv \frac{(Q_{n,a}^2 - Q_{2n,a}) Q_{n,b}^* Q_{n,c}^*}{M_a (M_a - 1) M_b M_c}\end{aligned}\tag{7}$$

²⁰⁶ where the definitions are the same as 2 sub-event method.

²⁰⁷ All the formulas in this section are without particles weights. The formulas with particle weights
²⁰⁸ (tracking efficiency and detector effects) will be introduced in the following section [25]. In this analysis,
²⁰⁹ we will only focus on the 4-particle cumulant. One reason is that due to the large statistical errors, it is not
²¹⁰ practical to measure 6-particle cumulant in small systems like pp . But the main reason is that traditional
²¹¹ cumulant relies on correlation of more particles to suppress the non-flow sources, since non-flow source
²¹² is usually associated with few particles. However, this is not true for sub-event method. By purposely put
²¹³ the particles into sub-events extended in a wide η range, the non-flow correlation is already suppressed.
²¹⁴ In other word, the advantage of sub-event method is to suppress the non-flow without involving more
²¹⁵ particles. For these reasons, we will only discuss the results of 4-particle cumulants in this note.

216 3 Data samples, event and tracking selections

217 3.1 pp 13 TeV

218 In 2015 and 2016, low- μ and intermediate- μ pp data was collected with ATLAS detectors [26]. An
219 additional pixel layer, the "Insertable B Layer" (IBL) [27, 28] installed between Run 1 and Run 2, is
220 used in the pp measurement. There are four run periods depending on the triggers used and they will be
221 discussed in the following sections.

222 The good events are selected by Good-Run-List and with $|z_{vtx}| < 150\text{mm}$.

223 To reduce the fraction of fake tracks, standard track selection has been applied:

- 224** • Tracks are from primary vertex
- 225** • If IBL hit is expected: at least 1 IBL hit
- 226** • If on IBL hit is expected: a Layer-0 hit if expected
- 227** • At least 1 pixel hit + dead sensors
- 228** • $p_T < 300 \text{ MeV}$: at least 2 SCT hits + dead sensors
- 229** • $p_T < 400 \text{ MeV}$: at least 4 SCT hits + dead sensors
- 230** • $p_T > 400 \text{ MeV}$: at least 6 SCT hits + dead sensors
- 231** • If $p_T > 10 \text{ GeV}$: χ^2 probability < 0.01
- 232** • $|d_0| < 1.5$
- 233** • $|z_0 - v_z| \sin\theta < 1.5$
- 234** • $|\eta| \leq 2.5$
- 235** • $p_T \geq 200 \text{ MeV}$

236 In the section of systematics, the selection will be tightened to check the stability of the results.

237 3.1.1 Run period 1

238 Low- μ data was collected in June and intermediate- μ data was collected in October. All the runs, with
239 Good Run List (GRL) selections, are included in this analysis and they are listed as below:

- 240** • Run 267358, peak $\mu = 0.00255$, 8.6 million events

241 `data15_13TeV.00267358.physics_MinBias.merge.AOD.f597_m1441`

- 242** • Run 267359, peak $\mu = 0.00612$, 12.3 million events

243 `data15_13TeV.00267359.physics_MinBias.merge.AOD.f597_m1441`

- 244** • Run 267360, peak $\mu = 0.374$, 12.6 million events

245 `data15_13TeV.00267360.physics_MinBias.merge.AOD.f597_m1441`

- 246** • Run 267367, peak $\mu = 0.46$, 17.1 million events

247 `data15_13TeV.00267367.physics_MinBias.merge.AOD.f597_m1441`

- 248 • Run 267385, peak $\mu = 0.466$, 71.2 million events
249 data15_13TeV.00267385.physics_MinBias.merge.AOD.f597_m1441
- 250 • Run 267599, peak $\mu = 0.488$, 105.6 million events
251 data15_13TeV.00267599.physics_MinBias.merge.AOD.f597_m1441
- 252 • Run 277025, peak $\mu = 0.613$, 19.9 million events
253 data15_13TeV.00277025.physics_MinBias.merge.AOD.f624_m1486
- 254 • Run 277081, peak $\mu = 0.62$, 44.5 million events
255 data15_13TeV.00277081.physics_MinBias.merge.AOD.f624_m1486

256 where peak μ is the highest μ value during data taking.

257 The triggers [29] that have been used in this analysis have both minimum bias (MinBias) and high
258 multiplicity track (HMT) triggers [15, 30] The HMT triggers are developed to enhance the statistics in
259 high multiplicity region. Since the main uncertainty in cumulant analysis is from statistical errors, HMT
260 triggers are crucial in order to extend the measurement to higher multiplicity region. This is especially
261 more important in sub-event method, since not all the particles pairs are included. The list of all triggers
262 used in this run period are listed as follows:

- 263 • HLT_noalg_mb_L1MBTS_1
- 264 • HLT_mb_sp400_trk40_hmt_L1MBTS_1_1
- 265 • HLT_mb_sp700_trk50_hmt_L1MBTS_1_1
- 266 • HLT_mb_sp900_trk60_hmt_L1MBTS_1_1
- 267 • HLT_mb_sp1400_trk90_hmt_L1TE10

268 where the items of the triggers are explained below:

- 269 • L1MBTS_1: L1 trigger requires at least 1 hit at either side of MBTS detector;
- 270 • L1MBTS_1_1: L1 trigger requires at least 1 hit at both side of MBTS detector;
- 271 • L1TEX: L1 trigger is seeded at L1 total energy from the full η range and the threshold is X GeV;
- 272 • spX: HLT trigger requires X number of space points in the SCT detector;
- 273 • trkX: HLT trigger requires X number of online reconstructed tracks from primary vertex;

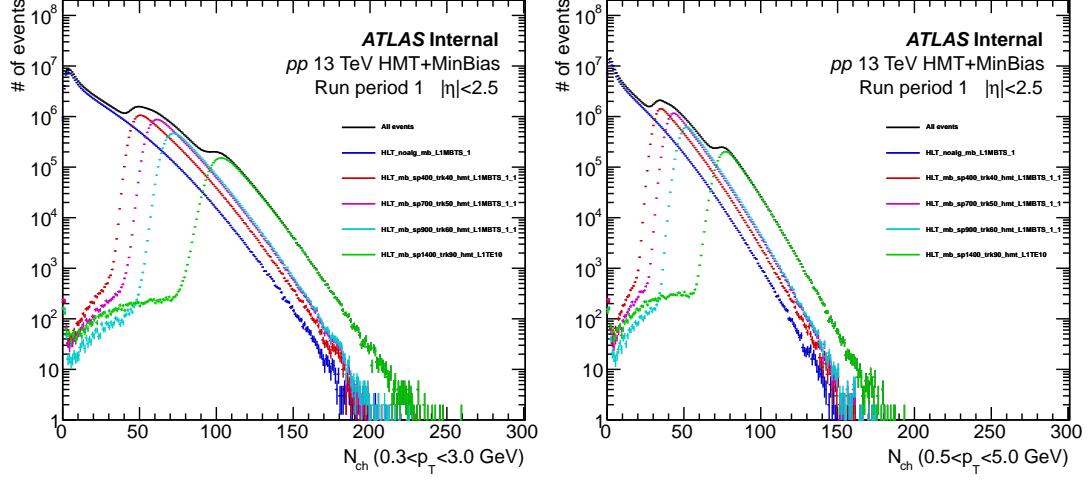
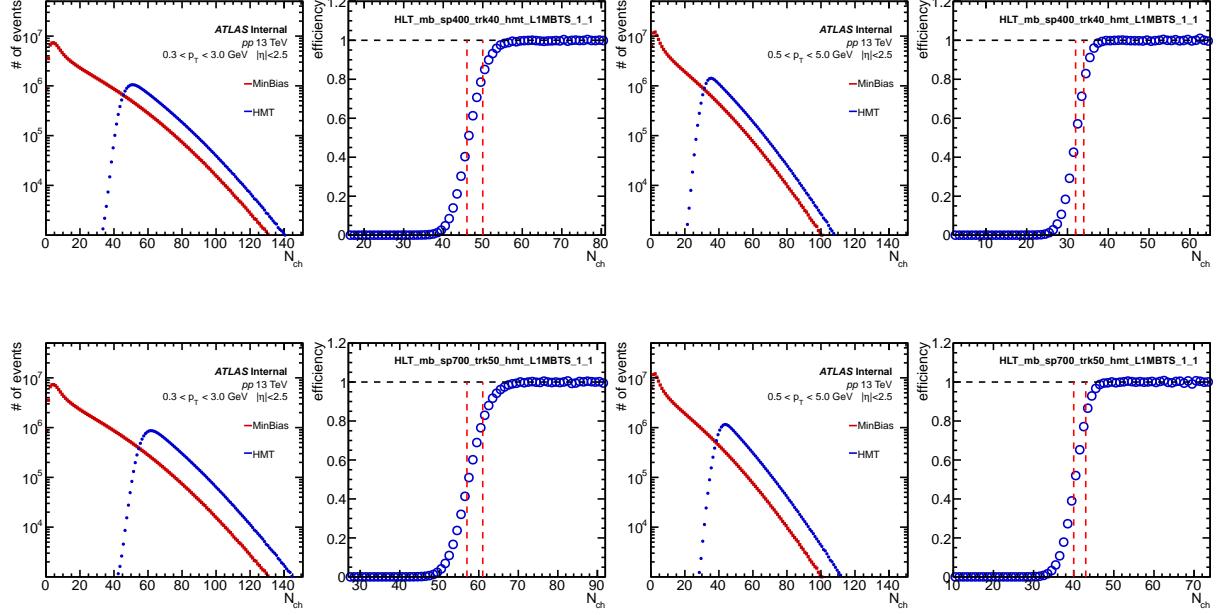


Figure 2: Distribution of number of tracks with two p_T thresholds: $0.3 < p_T < 3.0$ GeV and $0.5 < p_T < 5.0$ GeV, in 13 TeV pp run period 1. The major MinBias and HMT triggers are plotted separately. "All events" means all the MinBias and HMT triggers included, mainly from L1TE and HMT performance triggers.

As a summary of total statistics with all runs combined, Fig.2 shows the distributions of number of tracks with two p_T cuts used in this analysis: $0.3 < p_T < 3.0$ GeV and $0.5 < p_T < 5.0$ GeV. In each case, events collected with different MinBias and HMT triggers are shown separately, since in the end we will apply lower N_{ch} cuts to reduce the selection bias introduced by HMT triggers. Statistics of all the triggers in the MinBias stream (including those triggers not used in this analysis) are listed in the Appendix.



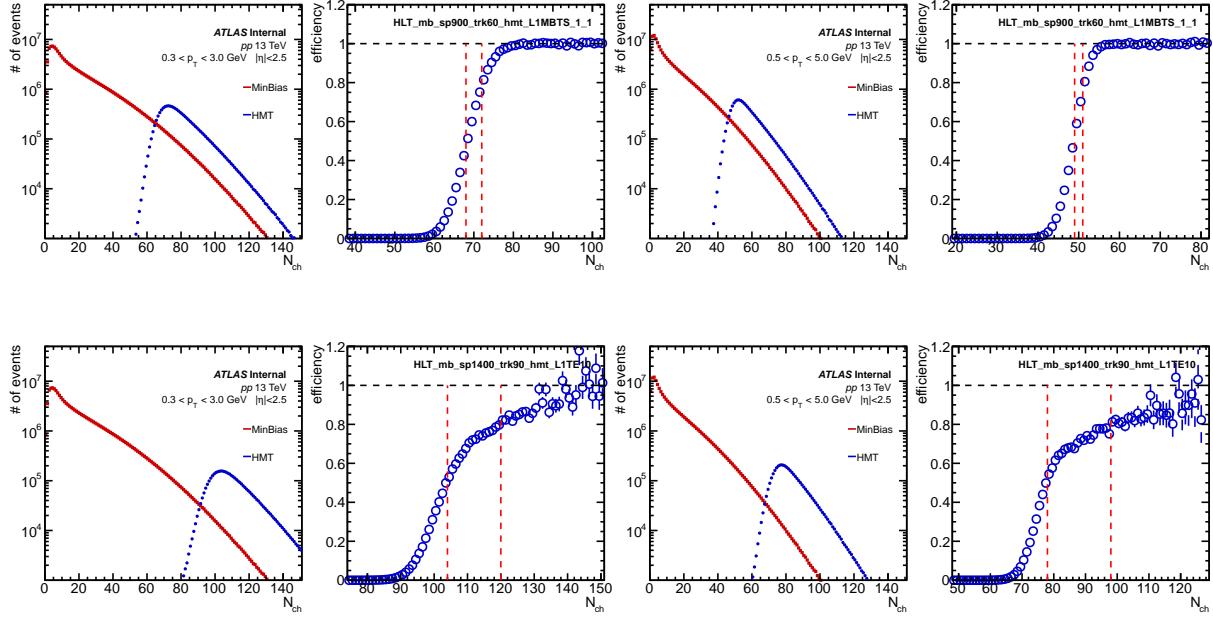


Figure 3: Trigger efficiencies of all major HMT triggers as a function of number of tracks in two p_T ranges: $0.3 < p_T < 3.0$ GeV and $0.5 < p_T < 5.0$ GeV, from 13 TeV pp run period 1. Efficiency is calculated relative to the corresponding MinBias trigger in this run period then scaled to 1.0 in the large N_{ch} region. The two red dash lines indicate 50% and 80% efficiency cuts.

During the data taking, both minbias and HMT triggers are prescaled. The triggers efficiency ϵ is estimated on the statistical level:

$$\epsilon^{\text{HMT}} \equiv \frac{\text{number of events passed HMT trigger} * \text{PS of HMT trigger}}{\text{number of events passed MB trigger} * \text{PS of MB trigger}} \quad (8)$$

Since the number of events with high N_{ch} triggered by minbias trigger is limited (during some runs the prescale of minbias triggers are very high), the estimated HMT trigger efficiency could fluctuate around 1 in some cases. However, because the turn-on of HMT trigger is very sharp, N_{ch} cut with 50% and 80% efficiency are not affected by this estimation. Trigger efficiencies of all the major HMT triggers are summarized in Fig. 3, where efficiencies are shown for two p_T ranges separately: $0.3 < p_T < 3.0$ GeV and $0.5 < p_T < 5.0$ GeV. In order to remove the potential bias introduced by trigger selections, HMT events are only used when their corresponding trigger efficiency is larger than 50%. Another criteria of 80% is applied as a cross check and will be included as one of the systematics.

3.1.2 Run period 2

The 2nd run period data was taken in 2016 and there are 2 runs used for this analysis:

- Run 299390, peak $\mu = 0.052$, 4.1 million events

`data16_13TeV.00299390.physics_MinBias.recon.AOD.r8358/`

- Run 300287, peak $\mu = 0.052$, 8.1 million events

`data16_13TeV.00300287.physics_MinBias.recon.AOD.r8358/`

The major MinBias and HMT triggers applied in this run period are:

- 296 • HLT_noalg_mb_L1MBTS_1_1
- 297 • HLT_mb_sp1800_hmtperf_L1TE5
- 298 • HLT_mb_sp1500_hmtperf_L1TE10
- 299 • HLT_mb_sp900_trk60_hmt_L1MBTS_1_1
- 300 • HLT_mb_sp1000_trk70_hmt_L1TE5
- 301 • HLT_mb_sp1400_trk90_hmt_L1TE10

302 where there is one new trigger item:

- 303 • hmt_perf: performance triggers of HMT, usually without online track cuts;

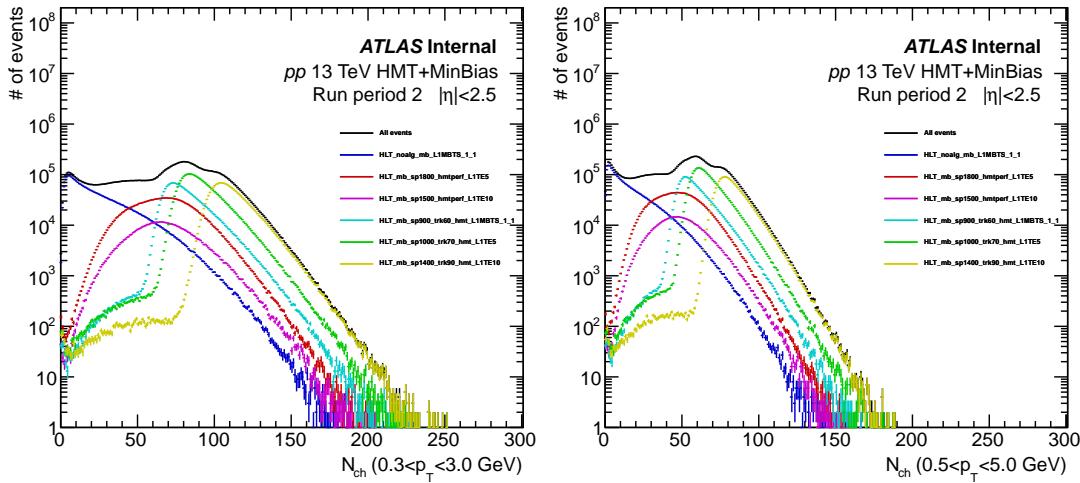
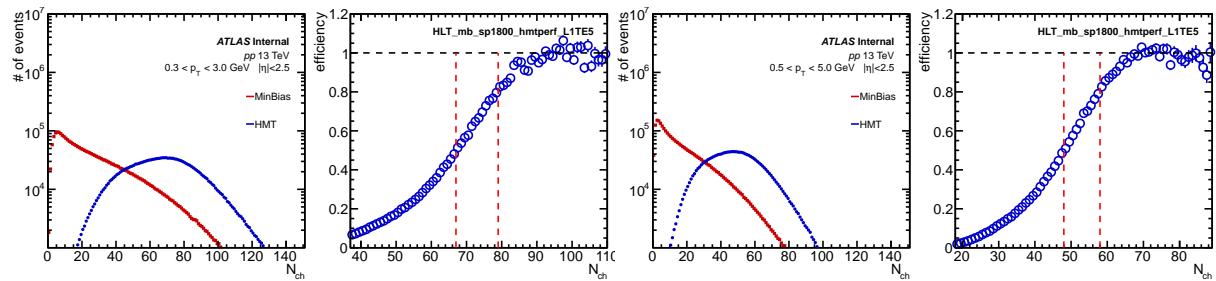


Figure 4: Distribution of number of tracks with two p_T thresholds: $0.3 < p_T < 3.0$ GeV and $0.5 < p_T < 5.0$ GeV, in 13 TeV pp run period 2. The major MinBias and HMT triggers are plotted separately.

304 The summary of statistics with all the major triggers used in this analysis are shown in Fig. 4.



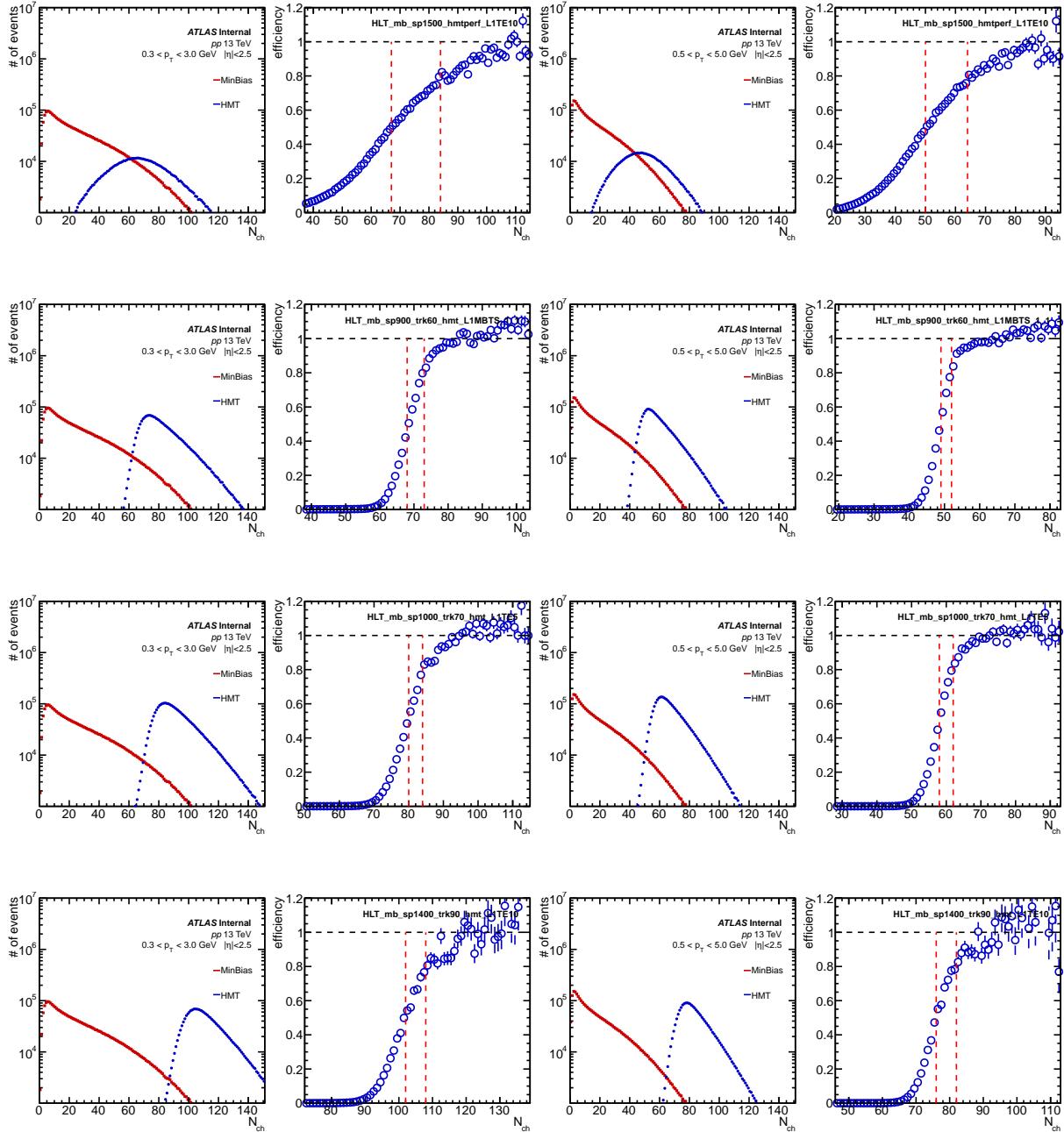


Figure 5: Trigger efficiencies of all major HMT triggers as a function of number of tracks in two p_T ranges: $0.3 < p_T < 3.0$ GeV and $0.5 < p_T < 5.0$ GeV, from 13 TeV pp run period 2. Efficiency is calculated relative to the corresponding MinBias trigger in this run period then scaled to 1.0 in the large N_{ch} region. The two red dash lines indicate 50% and 80% efficiency cuts.

Trigger efficiencies of all the major HMT triggers are summarized in Fig. 5, where efficiencies are shown for two p_T ranges separately: $0.3 < p_T < 3.0$ GeV and $0.5 < p_T < 5.0$ GeV.

3.1.3 Run period 3

The 3rd run period data was taken in 2016 and there is one low- μ run used for this analysis:

- Run 305359, peak $\mu = 0.06$, 16.5 million events

310 **data16_13TeV.00305359.physics_MinBias.recon.AOD.r8489/**

311 The major MinBias and HMT triggers applied in this run period are:

- 312 • HLT_mb_mbts_L1MBTS_1_1
- 313 • HLT_noalg_mb_L1TE30
- 314 • HLT_mb_sp700_hmtperf_L1TE5
- 315 • HLT_mb_sp1500_hmtperf_L1TE10
- 316 • HLT_mb_sp1000_trk70_hmt_L1TE5
- 317 • HLT_mb_sp1400_trk90_hmt_L1TE10

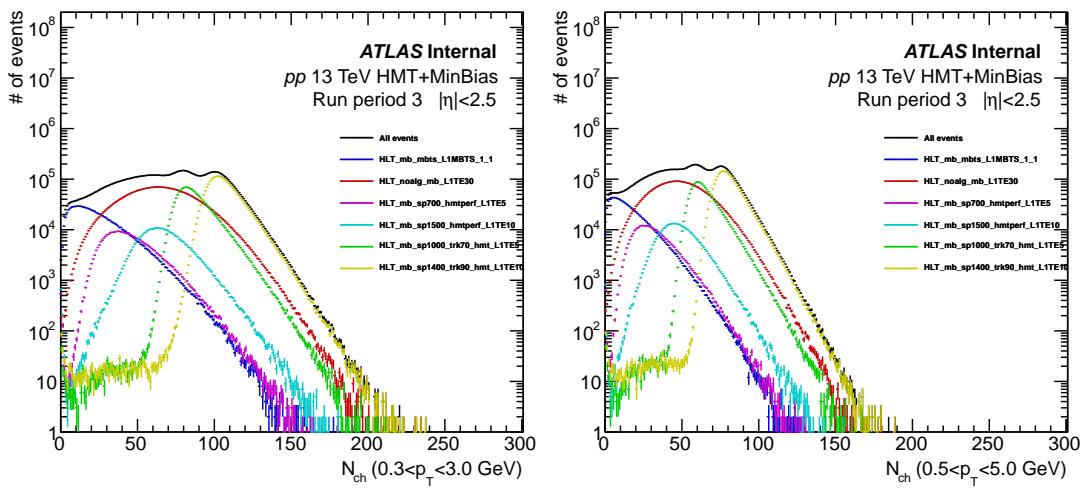
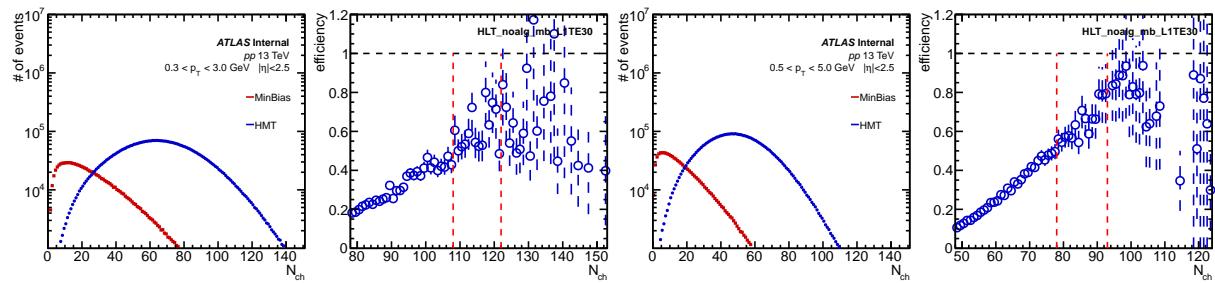


Figure 6: Distribution of number of tracks with two p_T thresholds: $0.3 < p_T < 3.0 \text{ GeV}$ and $0.5 < p_T < 5.0 \text{ GeV}$, in 13 TeV pp run period 3. The major MinBias and HMT triggers are plotted separately.

318 The summary of statistics with all the major triggers used in this analysis are shown in Fig. 6.



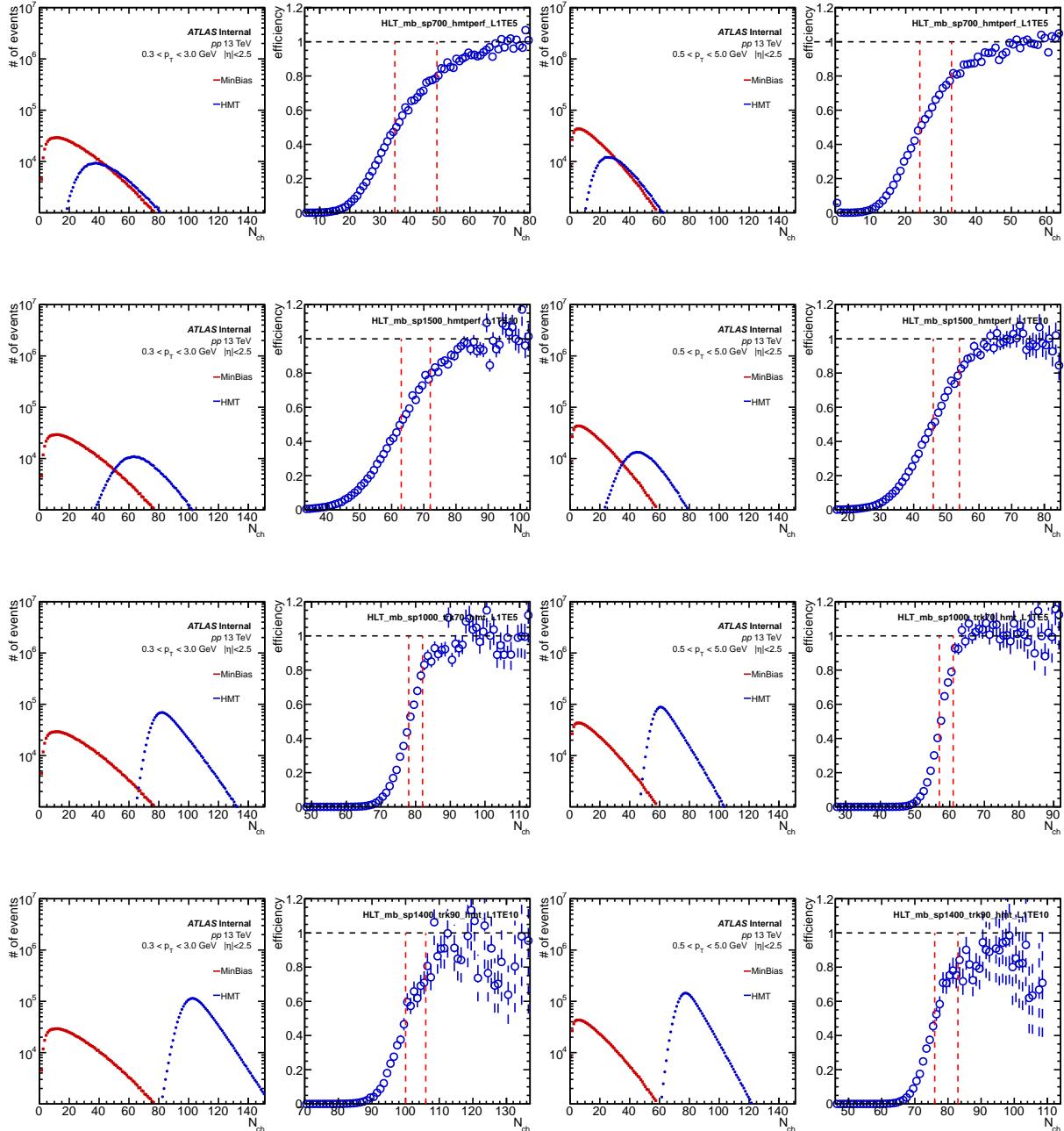


Figure 7: Trigger efficiencies of all major HMT triggers as a function of number of tracks in two p_T ranges: $0.3 < p_T < 3.0$ GeV and $0.5 < p_T < 5.0$ GeV, from 13 TeV pp run period 3. Efficiency is calculated relative to the corresponding MinBias trigger in this run period then scaled to 1.0 in the large N_{ch} region. The two red dash lines indicate 50% and 80% efficiency cuts.

Trigger efficiencies of all the major HMT triggers are summarized in Fig. 7, where efficiencies are shown for two p_T ranges separately: $0.3 < p_T < 3.0$ GeV and $0.5 < p_T < 5.0$ GeV.

3.1.4 Run period 4

The 4th run period data was taken in 2016 and there are 3 intermediate- μ runs used for this analysis:

- Run 309314, peak $\mu = 0.36$, 13.2 million events

324 **data16_13TeV.00309314.physics_MinBias.recon.AOD.r8576/**

- 325 • Run 309346, peak $\mu = 0.33$, 7.5 million events

326 **data16_13TeV.00309346.physics_MinBias.recon.AOD.r8576/**

- 327 • Run 310216, peak $\mu = 0.31$, 79.9 million events

328 **data16_13TeV.00310216.physics_MinBias.recon.AOD.r8600/**

329 The major MinBias and HMT triggers applied in this run period are:

- 330 • **HLT_mb_sptrk**
- 331 • **HLT_mb_sp900_trk50_hmt_L1TE5**
- 332 • **HLT_mb_sp1000_trk70_hmt_L1TE10.0ETA24**
- 333 • **HLT_mb_sp1400_trk90_hmt_L1TE20.0ETA24**

334 where a new trigger item has been introduced:

- 335 • **L1TEX.0ETA24**: L1 trigger is seeded at L1 total energy from the restricted η range: $-2.4 < \eta < 2.4$
336 and the threshold is X GeV;

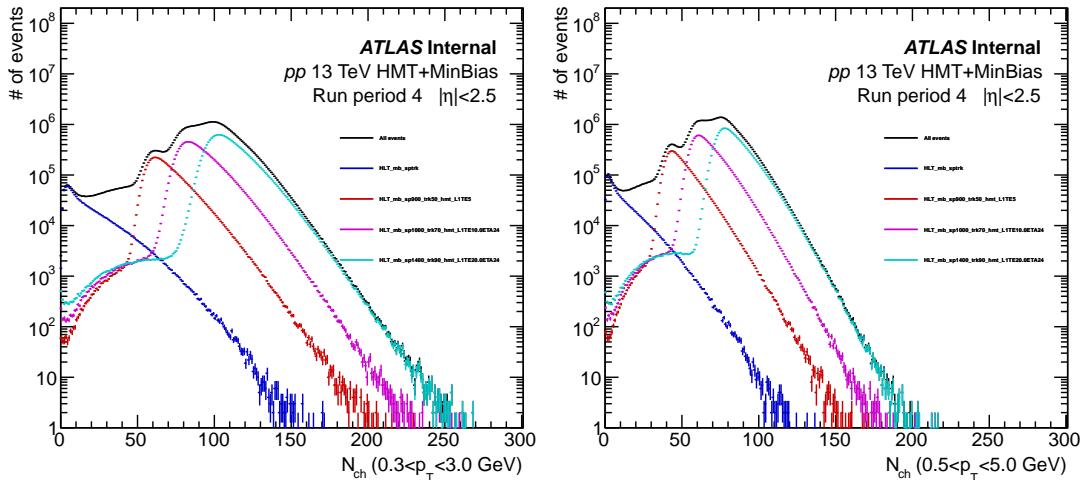
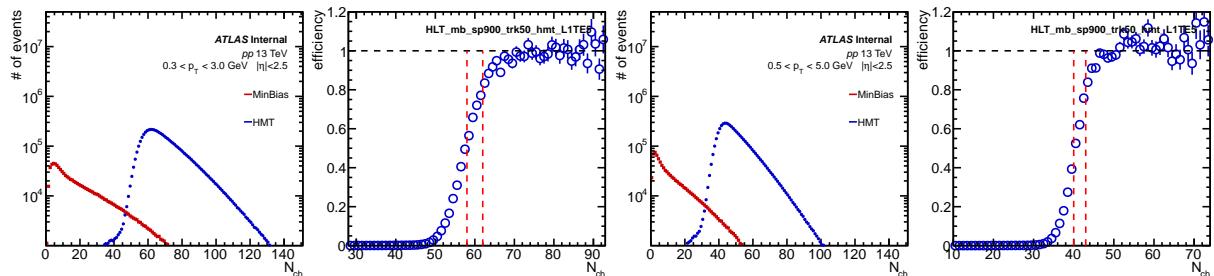


Figure 8: Distribution of number of tracks with two p_T thresholds: $0.3 < p_T < 3.0$ GeV and $0.5 < p_T < 5.0$ GeV, in 13 TeV pp run period 4. The major MinBias and HMT triggers are plotted separately.

337 The summary of statistics with all the major triggers used in this analysis are shown in Fig. 8.



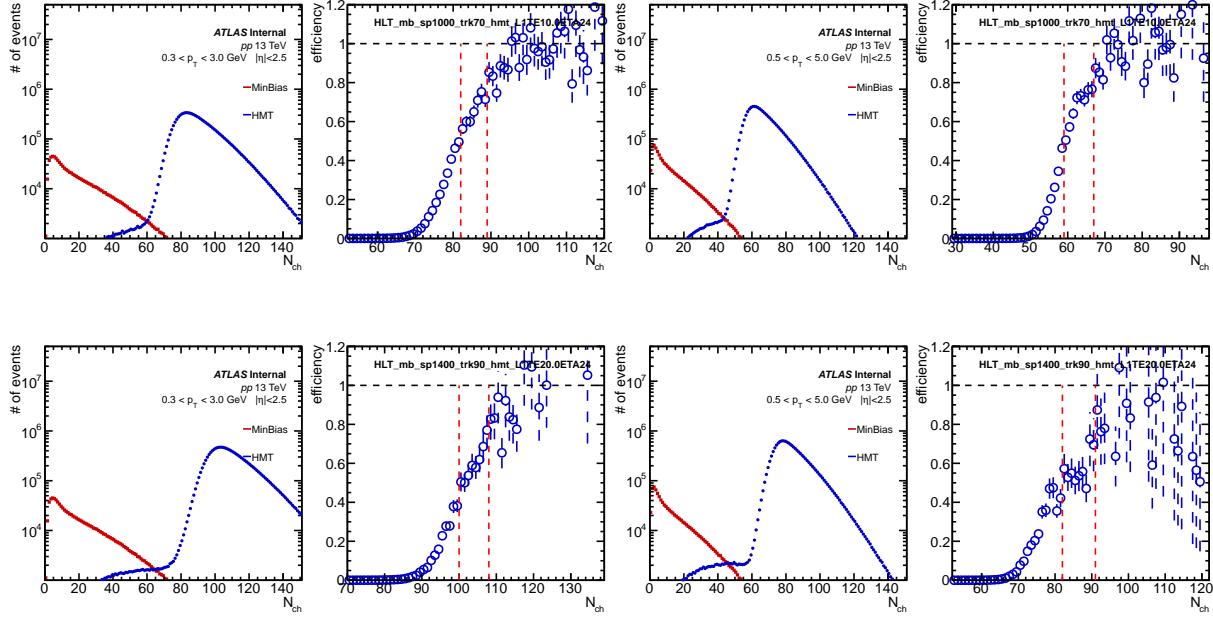


Figure 9: Trigger efficiencies of all major HMT triggers as a function of number of tracks in two p_T ranges: $0.3 < p_T < 3.0$ GeV and $0.5 < p_T < 5.0$ GeV, from 13 TeV pp run period 4. Efficiency is calculated relative to the corresponding MinBias trigger in this run period then scaled to 1.0 in the large N_{ch} region. The two red dash lines indicate 50% and 80% efficiency cuts.

Trigger efficiencies of all the major HMT triggers are summarized in Fig. 9, where efficiencies are shown for two p_T ranges separately: $0.3 < p_T < 3.0$ GeV and $0.5 < p_T < 5.0$ GeV.

3.1.5 PYTHIA for pp 13 TeV data

The tracking efficiency applied in this analysis were identical to the ATLAS previous cumulant measurement ATL-COM-PHYS-2015-740. The following studies are independent to the other analysis team, and the final results of c_{24} are checked to be consistent no matter which tracking efficiency map is applied.

The PYTHIA [31] A2 tune [32] was used to produce pp collisions with the same energy as in the data. The detector response is simulated with GEANT 4 with conditions matching those present during the data-taking. The simulated events are reconstructed with the same algorithms as data, in particular using the same track reconstruction as for the data.

The list of PYTHIA samples is shown in the following:

- 20 million ND sample

mc15_13TeV.361203.Pythia8_A2_MSTW2008LO_minbias.merge.AOD.
e3639_a782_a787_r6264/

- 1 million HMT sample with $N_{ch} > 120$

mc15_13TeV.361214.Pythia8_A2MSTW2008LO_minbias_NDnch120.merge.AOD.
e3908_a782_a787_r6264/

- 1 million HMT sample with $N_{ch} > 160$

mc15_13TeV.361215.Pythia8_A2MSTW2008LO_minbias_NDnch160.merge.AOD.

357 e3908_a782_a787_r6264/

- 358 • 1 million HMT sample with $N_{ch} > 200$

359 mc15_13TeV.361216.Pythia8_A2MSTW2008LO_minbias_NDnch200.merge.AOD.

360 e3908_a782_a787_r6264/

361 To study the tracking performance, the same track selection requirements were applied to PYTHIA
362 as 13 TeV data.

363 For the reconstructed tracks, the primary tracks are defined as:

- 364 • Pass the track selection requirements
- 365 • Truth match probability > 0.5
- 366 • Associated truth particle is a primary particle

367 where primary particle is defined on the truth level:

- 368 • Status = 1, charge! = 0
- 369 • $p_T > 200$ MeV
- 370 • $|\eta| \leq 2.5$
- 371 • $0 < \text{Barcode} < 2e5$
- 372 • strange baryons are excluded

373 The tracking efficiency ϵ is then defined as:

$$\epsilon(p_T, \eta, N_{ch}, z_{vtx}) \equiv \frac{N_{ch}^{\text{primary}}}{N_{ch}^{\text{truth}}} \quad (9)$$

374 where N_{ch}^{primary} denotes the number of primary tracks on reconstructed level and N_{ch}^{truth} denotes the num-
375 ber of primary particles on truth level. In previous pp analysis, the tracking efficiency is measure only
376 as a function of p_T and η . To evaluate the possible variation due to changing N_{ch} (In this analysis, N_{ch} is
377 defined as number of reconstructed tracks with $p_T > 0.4$ GeV.) and z_{vtx} , we also added these two into the
378 efficiency map.

379 The fake track is defined as:

- 380 • Pass the track selection requirements
- 381 • Fulfill one of the following:
 - 382 – Truth match probability < 0.5
 - 383 – Not associated with truth particle
 - 384 – Barcode = 0 of associated truth particle

385 Then the fraction of fake tracks f is defined as:

$$f(p_T, \eta, N_{ch}^{\text{primary}}, z_{vtx}) \equiv \frac{N_{ch}^{\text{fake}}}{N_{ch}^{\text{primary}} + N_{ch}^{\text{fake}}} \quad (10)$$

386 where N_{ch}^{fake} denotes the number of fake tracks.

387 To compensate the contribution from fake tracks, the efficiency ϵ can be corrected by defining ϵ'

$$\epsilon'(p_T, \eta, N_{ch}, z_{vtx}) \equiv \frac{N_{ch}^{primary} + N_{ch}^{fake}}{N_{ch}^{truth}} = \frac{\epsilon}{1-f} \quad (11)$$

388 where in this definition an additional correction of $1-f$ is added to the tracking efficiency ϵ .

389 As the z position of the vertex changes, average multiplicity distribution along η will also change
390 slightly. To evaluate the impact due to the vertex position changes, the efficiency and fake fraction are
391 determined as a function of z_{vtx} . Furthermore, the efficiency could also change slightly for different N_{ch}
392 region, so the efficiency and fake fraction are also measured as a function of N_{ch} .

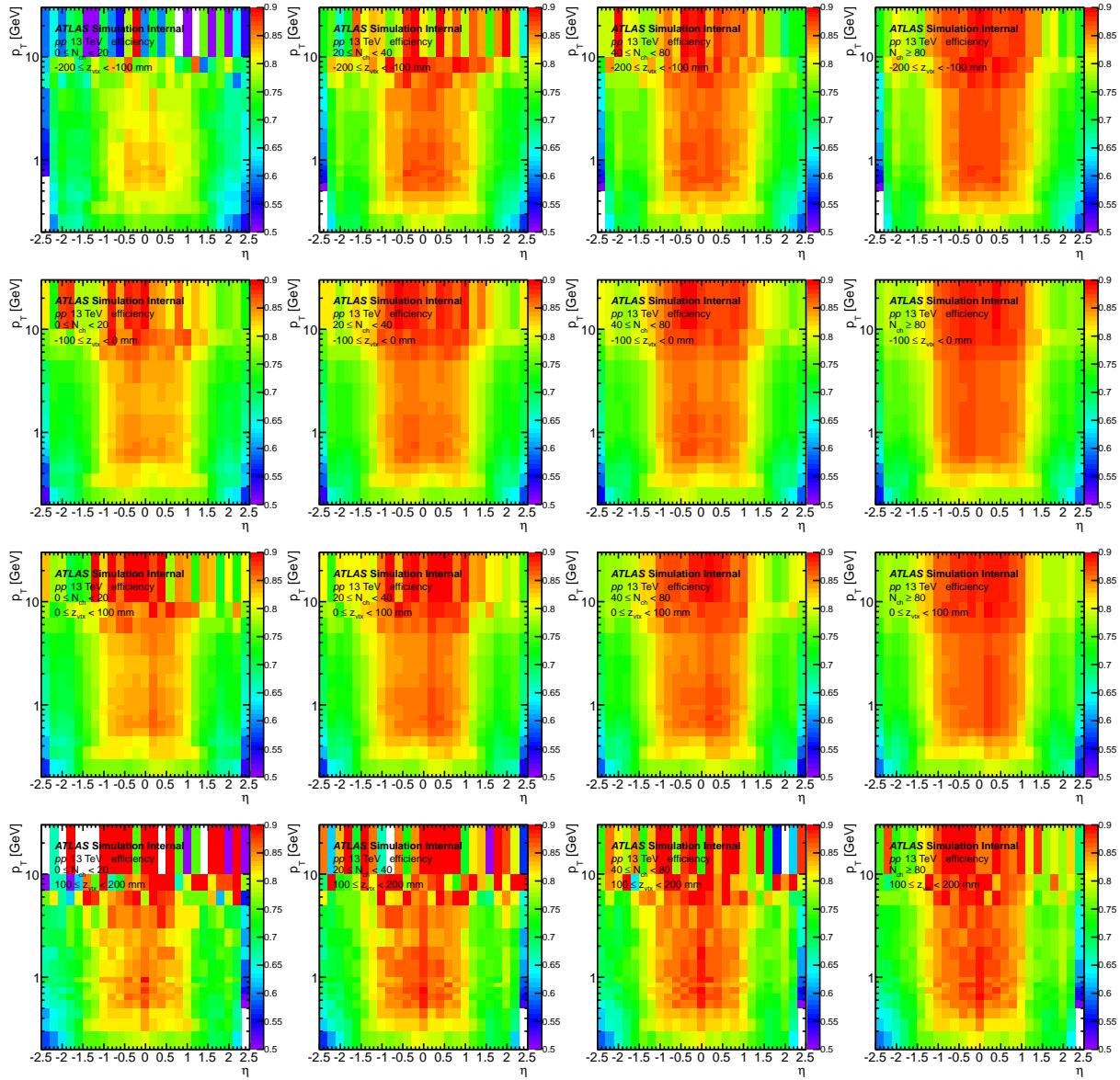


Figure 10: Tracking efficiency $\epsilon(\eta, p_T)$. Each column is for different N_{ch} ranges, and each row is for different z_{vtx} range.

393 The efficiency $\epsilon(p_T, \eta, N_{ch}, z_{vtx})$ is summarized in Fig. 10, where different rows represent different

394 z_{vtx} and different columns for different N_{ch} ranges. The efficiency is not uniform in η : higher in mid- η
 395 region. It is also not uniform along p_T : high- p_T has higher efficiency. In this analysis, the default p_T
 396 range is either $0.3 < p_T < 3.0$ GeV or $0.5 < p_T < 5.0$ GeV, and in the plots the p_T starts from 0.2 GeV.
 397 So the efficiency is in low- p_T is still above 60%. Moving from low- N_{ch} to high- N_{ch} the efficiency
 398 slightly increases: evaluating ϵ in different N_{ch} is thus required. However, when move from $z_{vtx} = 200$
 399 mm to z_{vtx} 200 mm, the efficiency doesn't change much. To gain better statistics, we decide to merge
 400 different z_{vtx} ranges when applying efficiency in the data. Besides, since this measurement is not sensitive
 401 to rapidity, the minimal changes of the efficiency as a function of z_{vtx} only have negligible impact on the
 402 final results.

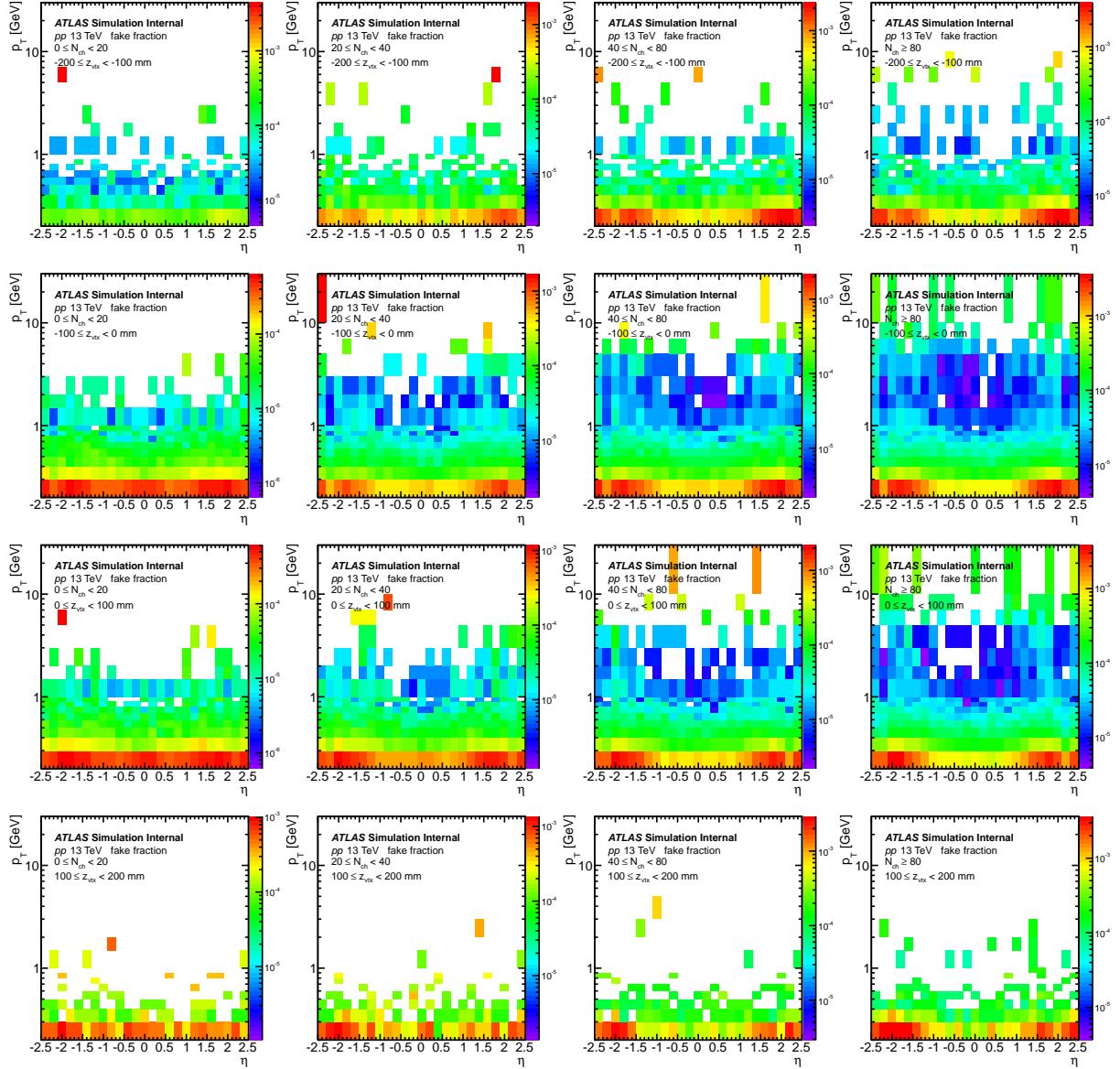


Figure 11: Fake fraction $f(\eta, p_T)$. Each column is for different N_{ch} ranges, and each row is for different z_{vtx} range.

403 Fig. 11 summarizes the fake fraction $f(\eta, p_T)$ for different z_{vtx} and N_{ch} ranges, where the layouts of
 404 the panels are same as previous efficiency plots. The fake fraction is higher in the low- p_T region and is

consistent with 0 in high- p_T . The fake fraction is also η dependent, but opposite to efficiency: lower in mid- η . Overall the fake fraction is on the level of below 0.1%. Like efficiency, the fake fraction is also dependent of N_{ch} , and has very weak dependence of z_{vtx} . So we will merge the fake fraction map among different z_{vtx} ranges in the end.

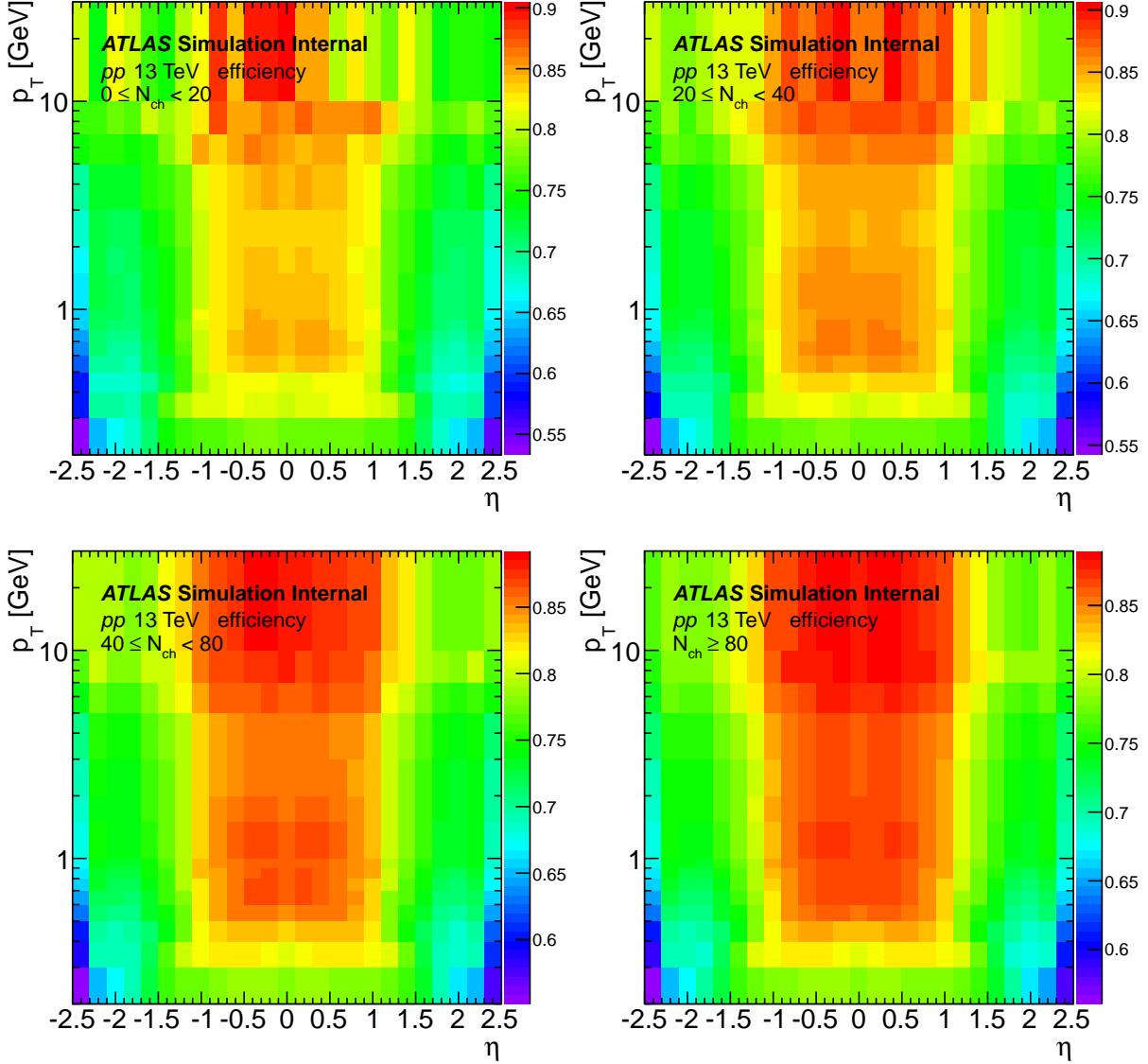


Figure 12: Tracking efficiency $\epsilon(\eta, p_T)$. Different panels are for different N_{ch} ranges, where different z_{vtx} ranges have been merged.

The merged tracking efficiency is shown in Fig. 12, where four panels are for four different N_{ch} ranges. The boundaries of ranges are chosen to make each range have enough statistics for the efficiency measurement. The efficiency is higher in high multiplicity. These efficiency maps are applied in this analysis, where linear interpolation is used to obtain the more precise efficiency when p_T or η are not in the bin center. The c_{24} are compared with using tracking efficiency from previous cumulant measurement and the results are consistent.

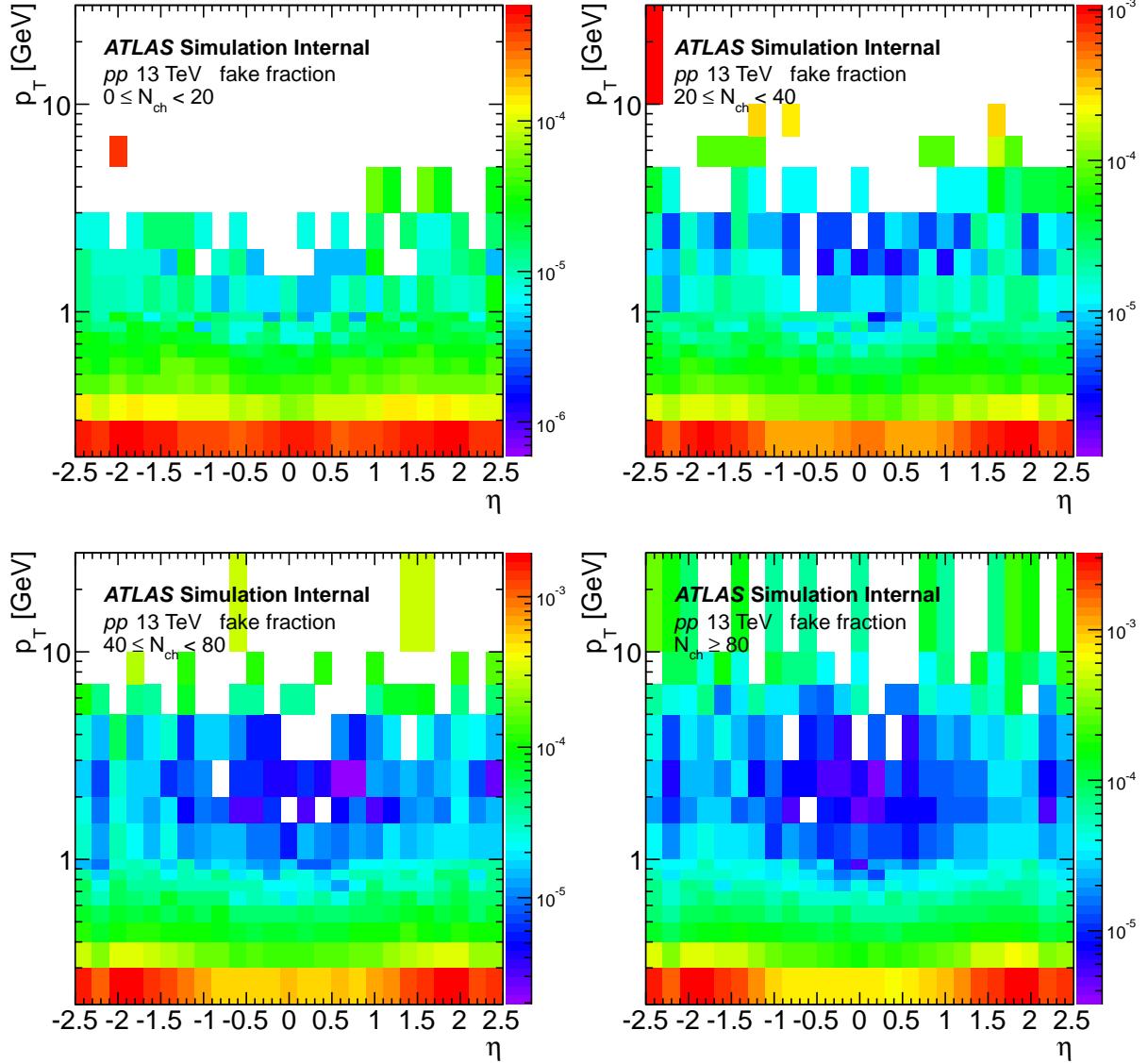


Figure 13: Fake fraction $f(\eta, p_T)$. Different panels are for different N_{ch} ranges, where different z_{vtx} ranges have been merged.

415 The merged fake fraction is shown in Fig. 13, where four panels are for four different N_{ch} ranges.
 416 Since the maximum fake fraction is smaller than the level of 0.1%, even for the lowest p_T , we will not
 417 apply this additional correction in the data analysis.

418 In the systematics section we will test the stability of the results by varying the tracking efficiency.
 419 Appendix 8 summarizes many detailed monitoring plots for this PYTHIA sample.

420 3.1.6 Tracking validation for run 305359, 309314 and 309346

421 When producing the GRL for the low- μ pp runs in 2016, three runs are tagged with following issues:

- 422
- Run 305359: BTAG_BLAYER_SERIOUS_PROBLEM
 - Run 309314: ID_BS_RUNAVERAGE
- 423

424 • Run 309346: ID_BS_RUNAVERAGE

425 where in run 305359, the IBL detector is absent while in the other two runs, there is no constrain on the
 426 beam spot position at the trigger level. (The offline reconstruction DOES impose the beam spot constraint
 427 in all runs.) When beam spot is constraint, fluctuation of rate of HMT triggers has been observed, due to
 428 online reconstruction issues.

429 The purpose of this section is to validate the basic tracking quantities in these three runs, to check
 430 whether IBL and beam spot can have an impact on the results. Run 310216 is used as a reference for
 431 comparison. All the events are required to pass the MB trigger `HLT_mb_sptrk` and the track selection is
 432 identical to the pp analysis.

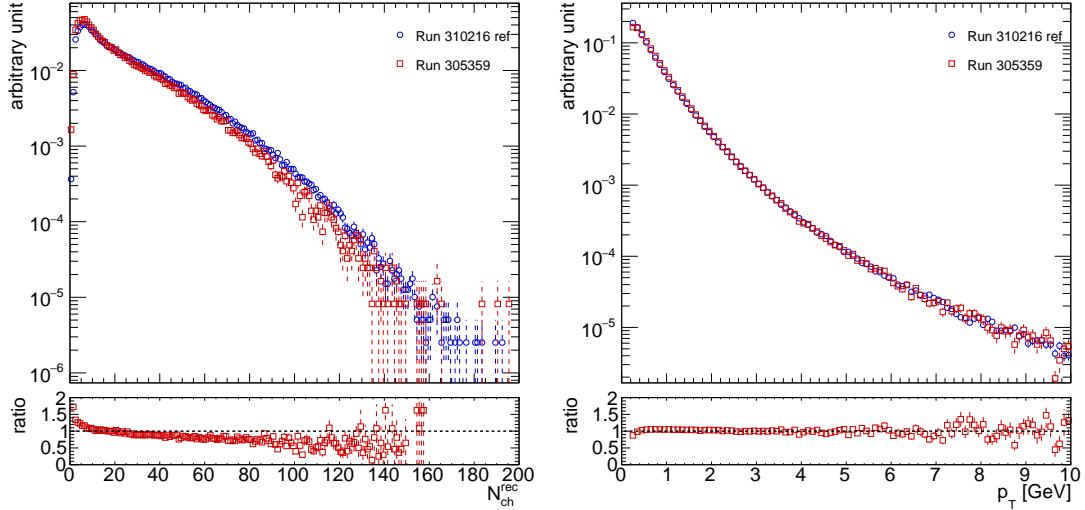


Figure 14: Distribution of number of offline reconstructed tracks and p_T spectrum: a comparison between run 305359 and run 310216.

433 Fig. 14 shows the comparison of number of reconstructed tracks distribution and p_T spectrum, be-
 434 tween run 305359 and 310216. The mean value of number of tracks is slightly smaller in 305359 than
 435 310216. This is probably due to that the IBL is absent in 305359, and same cut on number of Pixel
 436 hits will reject slightly more events in this run. Meanwhile, the p_T spectrums are consistent between
 437 the two runs, except for the lowest p_T region, where again it is due to the absence of IBL detector, thus
 438 the tracking reconstruction has a poorer efficiency in the lowest p_T region. We will come back to this
 439 behavior again once we show the comparison of Run 1 and Run 2 $p+Pb$ tracking efficiency.

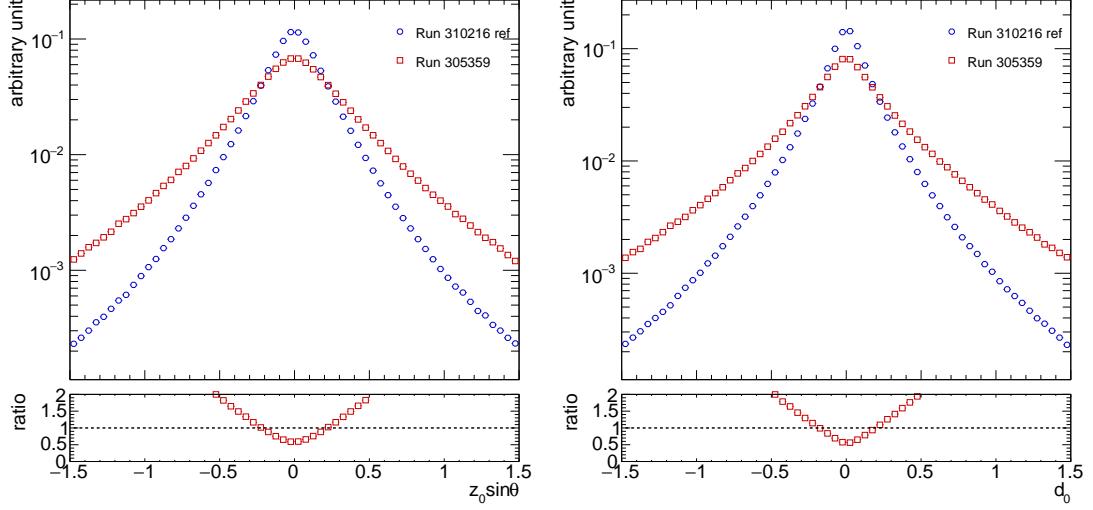


Figure 15: Distribution of $z_0 \sin\theta$ and d_0 of reconstructed tracks: a comparison between run 305359 and run 310216.

Fig. 15 shows the comparison of $z_0 \sin\theta$ and d_0 distribution of reconstructed tracks, between run 305359 and 310216. Both z_0 and d_0 distributions are much wider when IBL is absent. The default tracking pointing cut is at 1.5 mm, which will result in different tracks because of the different z_0 and d_0 distribution.

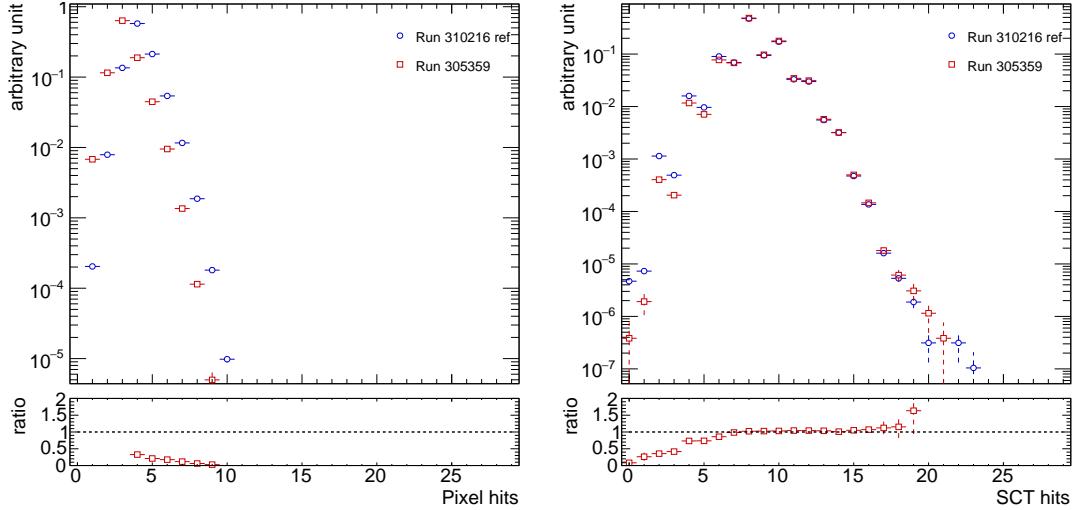


Figure 16: Distribution of number of Pixel and SCT hits of reconstructed tracks: a comparison between run 305359 and run 310216.

Fig. 16 shows the comparison of number of Pixel and SCT hits, between run 305359 and 310216. Due to absence of IBL, the number of Pixel hits in run 305359 is shifted by 1 compared with 310216. The number of SCT hits are consistent between the two runs, above the default track selection cut.

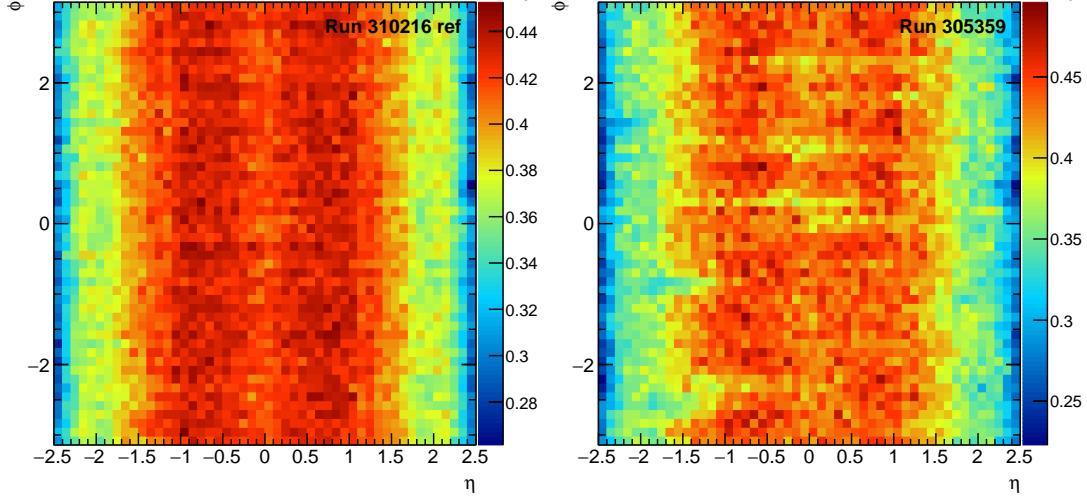


Figure 17: 2D map of raw η - ϕ distribution: a comparison between run 305359 and run 310216.

Finally, fig. 17 shows the 2D map of raw η - ϕ distribution. "raw" means the tracks are not weighted by tracking efficiency. The distribution is much more uniform in run 310216 than 305359.

As a conclusion, because of the absence of IBL, z_0 , d_0 and η - ϕ distributions are quite different between run 310216 and 305359. In principle, most of the differences can be covered by a dedicated Monte-Carlo sample with special detector setup for the run 305359. However, after discussion with the experts who prepared the GRL, we decide not to include run 305359, which only contributes to less than 10% of the total statistics.

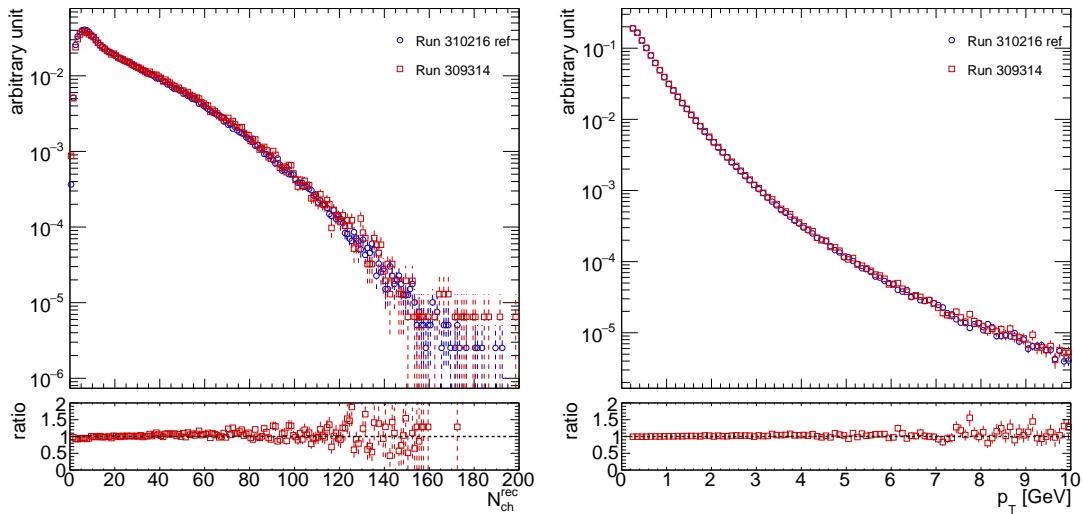


Figure 18: Distribution of number of offline reconstructed tracks and p_T spectrum: a comparison between run 309314 and run 310216.

The rest two runs are tagged with the error ID_BS_RUNAVERAGE, since the beam spot position is not constraint. In the following we will demonstrate that tracking quantities will be consistent without beam spot constraint. Fig. 18 shows the comparison of number of reconstructed tracks distribution and p_T spectrum, between run 309314 and 310216. The ratios show that they are very consistent.

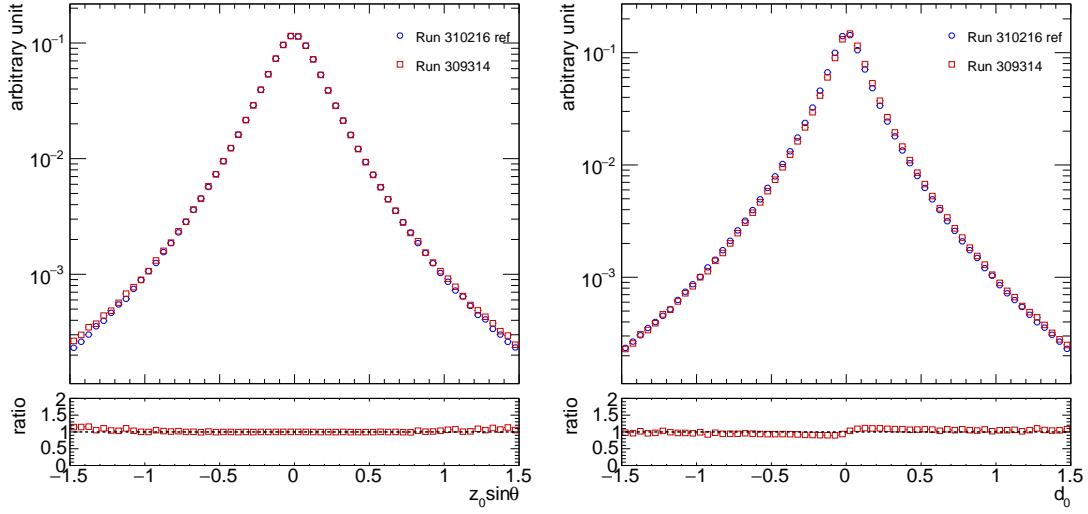


Figure 19: Distribution of $z_0 \sin \theta$ and d_0 of reconstructed tracks: a comparison between run 309314 and run 310216.

458 Fig. 19 shows the comparison of $z_0 \sin \theta$ and d_0 distribution of reconstructed tracks, between run
459 309314 and 310216. They are very consistent.

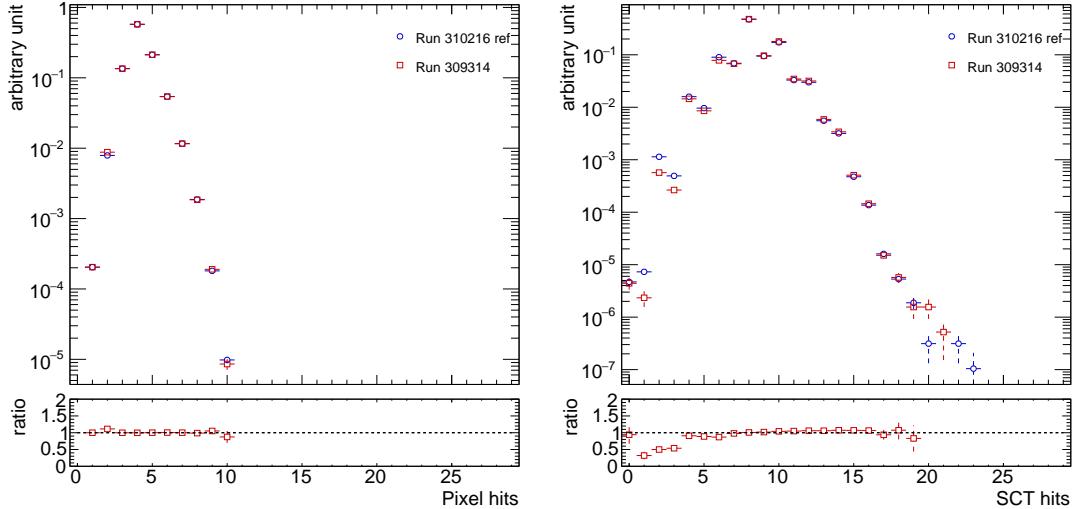


Figure 20: Distribution of number of Pixel and SCT hits of reconstructed tracks: a comparison between run 309314 and run 310216.

460 Fig. 20 shows the comparison of number of Pixel and SCT hits, between run 309314 and 310216.
461 They are very consistent.

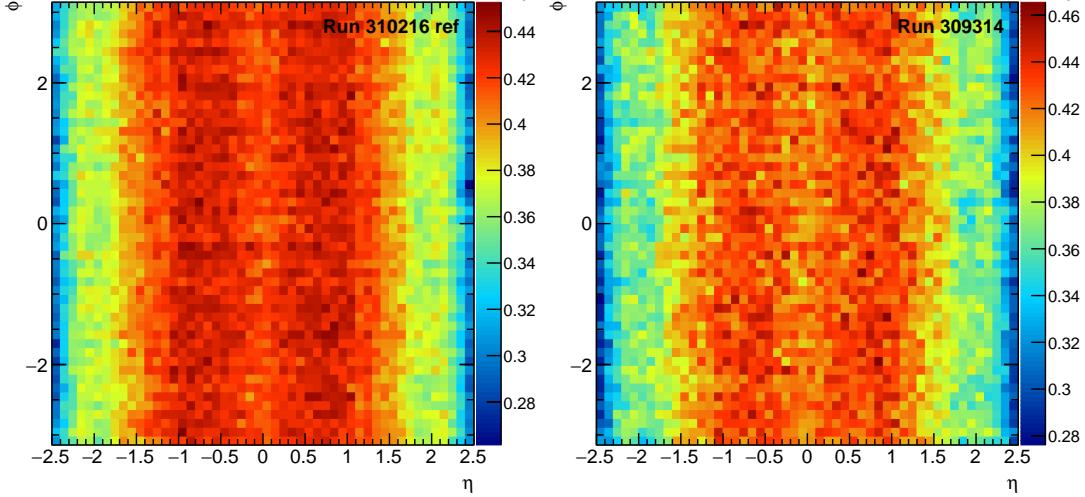


Figure 21: 2D map of raw η - ϕ distribution: a comparison between run 309314 and run 310216.

Finally, fig. 21 shows the 2D map of raw η - ϕ distribution. "raw" means the tracks are not weighted by tracking efficiency. The distributions are very consistent.

As a conclusion, even though beam spot is not constraint in these two runs, various tracking quantities are very consistent with the reference run. Due to these reasons, we will include these two runs into the pp analysis.

3.2 pp 5.02 TeV

In 2015 before the Pb+Pb run, low- μ and intermediate- μ 5.02 TeV pp data was collected with ATLAS detectors. Since this reference run is right after previous pp runs, the track selection is identical to 13 TeV pp . All the runs, with GRL selection, are included in this analysis and they are listed as below:

- Run 286282, peak $\mu = 0.58$, 37.3 million events from MinBias stream
data15_5TeV.00286282.physics_MinBias.recon.AOD.r7744/
- Run 286328, peak $\mu = 1.58$, 5.8 million events from MinBias stream
data15_5TeV.00286328.physics_MinBias.recon.AOD.r7744/
- Run 286361, peak $\mu = 1.34$, 3.5 million events from MinBias stream
data15_5TeV.00286361.physics_MinBias.recon.AOD.r7744/
- Run 286364, peak $\mu = 1.58$, 7.9 million events from MinBias stream
data15_5TeV.00286364.physics_MinBias.recon.AOD.r7744/
- Run 286367, peak $\mu = 0.67$, 1.7 million events from MinBias stream
data15_5TeV.00286367.physics_MinBias.recon.AOD.r7744/
- Run 286411, peak $\mu = 1.33$, 10.8 million events from MinBias stream
data15_5TeV.00286411.physics_MinBias.recon.AOD.r7744/
- Run 286474, peak $\mu = 1.25$, 5.3 million events from MinBias stream

484 **data15_5TeV.00286474.physics_MinBias.recon.AOD.r7744/**

485 Similar as 13 TeV pp , the triggers applied in 5 TeV pp have two components: MinBias and HMT. A
486 list of all the major MinBias and HMT triggers used in this analysis is summarized as follows:

- 487 • HLT_mb_sptrk
488 • HLT_noalg_mb_L1MBTS_1
489 • HLT_mb_sp800_pusup400_trk50_hmt_L1TE5
490 • HLT_mb_sp900_pusup500_trk60_hmt_L1TE5
491 • HLT_mb_sp1200_pusup700_trk70_hmt_L1TE5
492 • HLT_mb_sp1400_pusup550_trk90_hmt_L1TE10

493 where a new trigger item has been introduced:

- 494 • pusup: maximum number of hits from a vertex, used to suppress pile-up events before track
495 reconstruction;

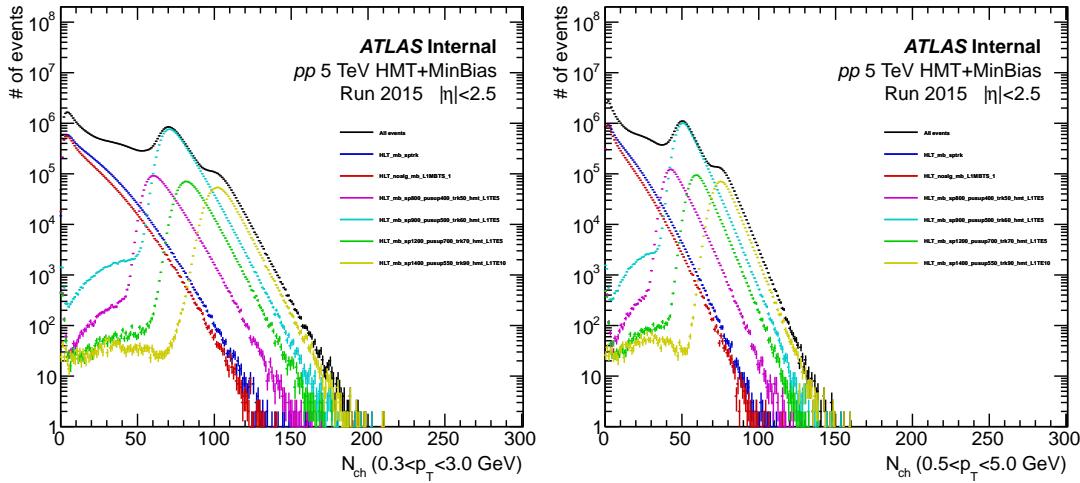
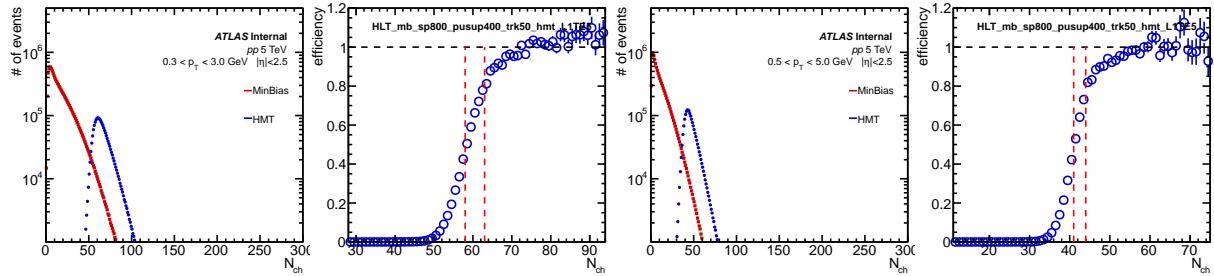


Figure 22: Distribution of number of tracks with two p_T thresholds: $0.3 < p_T < 3.0$ GeV and $0.5 < p_T < 5.0$ GeV, in 13 TeV pp run period 4. The major MinBias and HMT triggers are plotted separately.

496 The summary of statistics with all the major triggers used in this analysis are shown in Fig. 22.



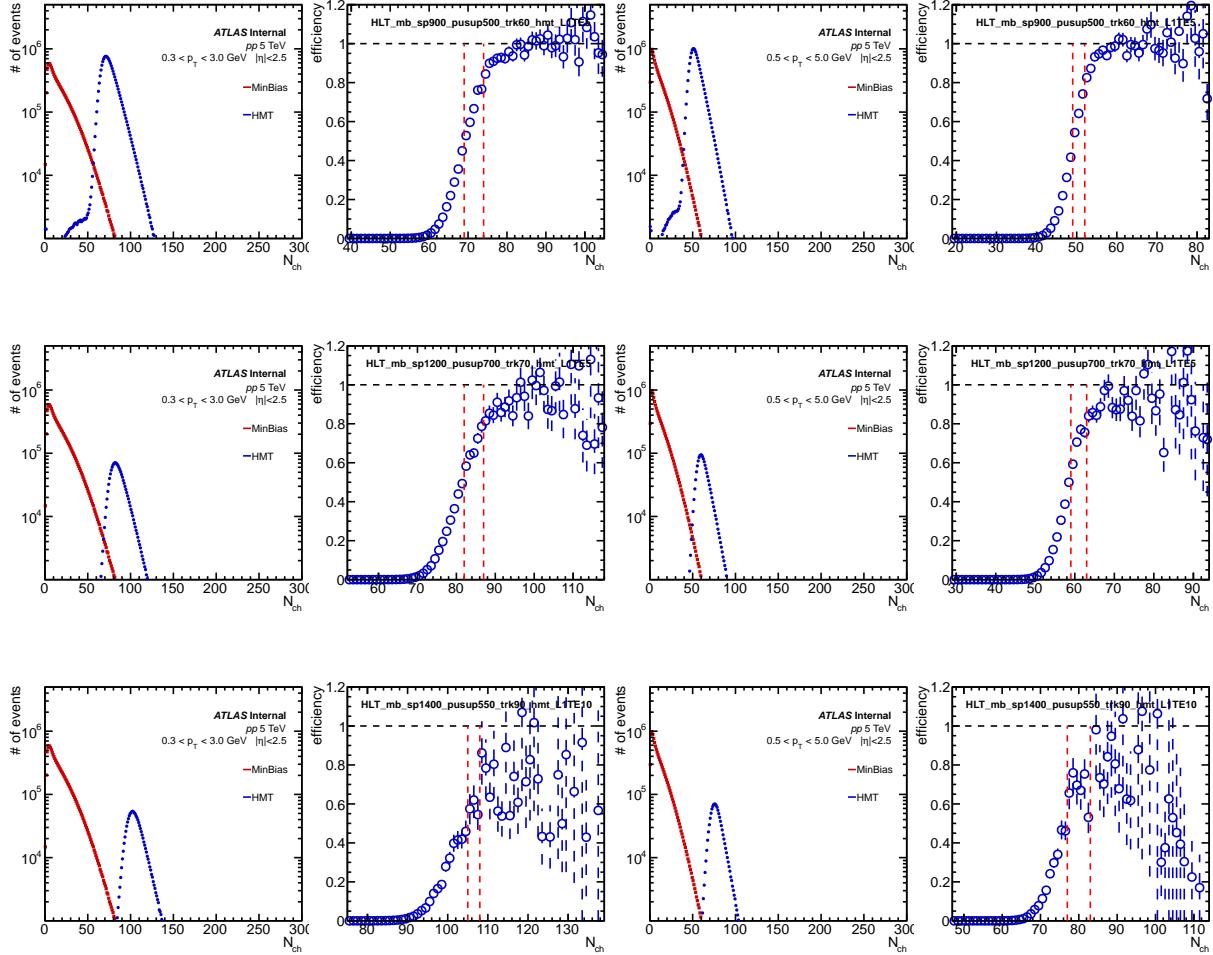


Figure 23: Trigger efficiencies of all major HMT triggers as a function of number of tracks in two p_T ranges: $0.3 < p_T < 3.0 \text{ GeV}$ and $0.5 < p_T < 5.0 \text{ GeV}$, from 5 TeV pp run. Efficiency is calculated relative to the corresponding MinBias trigger in this run period then scaled to 1.0 in the large N_{ch} region. The two red dash lines indicate 50% and 80% efficiency cuts.

Trigger efficiencies of all the major HMT triggers are summarized in Fig. 23, where efficiencies are shown for two p_T ranges separately: $0.3 < p_T < 3.0 \text{ GeV}$ and $0.5 < p_T < 5.0 \text{ GeV}$.

3.2.1 PYTHIA for pp 5.02 TeV data

Since the 5 TeV pp run is close to 13 TeV run in time, the detector effects between the two runs should be similar. For this reason, tracking efficiency and fraction of fake tracks extracted from 13 TeV PYTHIA is applied to 5 TeV pp data.

3.2.2 Impact of ID misalignment

In order to study the effect of ID misalignment, the package `InDetTrackSystema9csTools-00-00-18` was used. The 5 TeV pp cumulant are recalculated with the corrected track momenta and compared to the default measurement.

For $c24$ the effects of ID misalignment for low multiplicity (up to 60) is at 0.1% level. At higher multiplicity the difference is smaller than statistical errors. This concludes that the ID misalignment has minimal impact on this analysis.

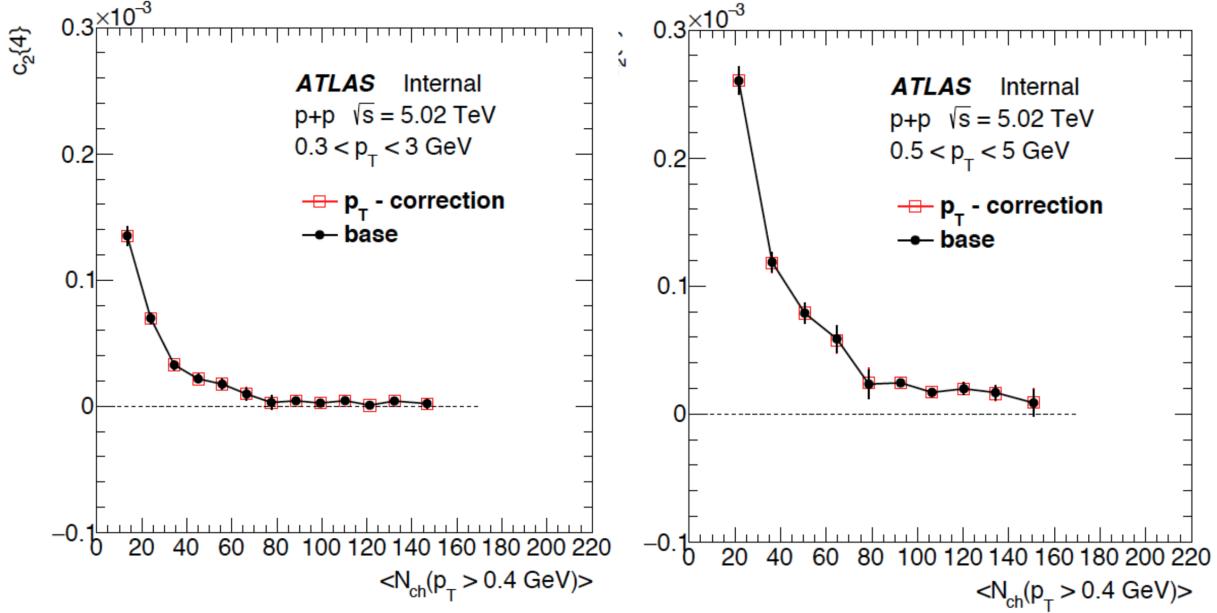


Figure 24: c_{24} measured with default and p_T corrected, for $0.3 < p_T < 3.0$ GeV (left) and $0.5 < p_T < 5.0$ GeV (right).

510 3.3 $p+\text{Pb}$ 5.02 TeV

511 5.02 TeV $p+\text{Pb}$ data have been collected in 2013 [5] and 2016. Since the event and tracking selections
 512 are quite different between Run 1 and Run 2, they will be discussed separately in the following sections.

513 3.3.1 2013 $p+\text{Pb}$

514 In 2013, low- μ 5.02 TeV $p+\text{Pb}$ data was collected with ATLAS detectors. In addition to GRL selection,
 515 each event also needs to pass the MBTS cuts:

- 516 • $|\text{time}_A|$ or $|\text{time}_C|$ must not equal 75 or 0 ns;
- 517 • $|\text{time}_A - \text{time}_C| < 10$ ns;

518 and the pile-up events are suppressed by rejecting events containing more than one good reconstructed
 519 vertex. The remaining pileup events are further suppressed based on the signal in the ZDC on the Pb-
 520 fragmentation side. This signal is calibrated to the number of detected neutrons (N_n) based on the location
 521 of the peak corresponding to a single neutron. The distribution of N_n in events with pileup is broader than
 522 that for the events without pileup. Hence, a simple cut on the high tail end of the ZDC signal distribution
 523 further suppresses the pileup, while retaining more than 98% of the events without pileup. After this
 524 pileup rejection procedure, the residual pileup fraction is estimated to be 1% in the event class with the
 525 highest track multiplicity studied in this analysis. For details of this pile-up rejection, refer to the internal
 526 note of flow measurements in $p+\text{Pb}$ [33].

527 The track selection criteria is $p+\text{Pb}$ is slightly different from pp [14, 30]:

- 528 • Tracks are from primary vertex
- 529 • Present hit in B-Layer if expected
- 530 • At least 1 pixel hit

- 531 • $p_T < 300$ MeV: at least 4 SCT hits + dead sensors
- 532 • $p_T > 300$ MeV: at least 6 SCT hits + dead sensors
- 533 • significance of d_0 is < 3.0
- 534 • significance of $z_0 \sin\theta$ is < 3.0
- 535 • $|\eta| \leq 2.5$
- 536 • $p_T \geq 200$ MeV

537 In addition, an additional tighter cut defined to be the track selection requirements in Pb-Pb analysis is
 538 used to check the stability of results in the systematics section.

539 All the runs included in this analysis are listed as follows. A special reprocessing was done for these
 540 runs so that the lowest p_T goes down to 0.1 GeV.

- 541 • Run 217999, 4.7 million events from MinBias stream
 542 data13_hip.00217999.physics_MinBias.merge.NTUP_HI.r5813_p1729/
- 543 • Run 218006, 2.7 million events from MinBias stream
 544 data13_hip.00218006.physics_MinBias.merge.NTUP_HI.r5813_p1729/
- 545 • Run 218048, 6.6 million events from MinBias stream
 546 data13_hip.00218048.physics_MinBias.merge.NTUP_HI.r5813_p1729/
- 547 • Run 218118, 3.9 million events from MinBias stream
 548 data13_hip.00218118.physics_MinBias.merge.NTUP_HI.r5813_p1729/
- 549 • Run 218168, 4.2 million events from MinBias stream
 550 data13_hip.00218168.physics_MinBias.merge.NTUP_HI.r5813_p1729/
- 551 • Run 218179, 4.5 million events from MinBias stream
 552 data13_hip.00218179.physics_MinBias.merge.NTUP_HI.r5813_p1729/
- 553 • Run 218213, 1.8 million events from MinBias stream
 554 data13_hip.00218213.physics_MinBias.merge.NTUP_HI.r5813_p1729/
- 555 • Run 218222, 0.5 million events from MinBias stream
 556 data13_hip.00218222.physics_MinBias.merge.NTUP_HI.r5813_p1729/
- 557 • Run 218301, 3.8 million events from MinBias stream
 558 data13_hip.00218301.physics_MinBias.merge.NTUP_HI.r5813_p1729/
- 559 • Run 218338, 3.8 million events from MinBias stream
 560 data13_hip.00218338.physics_MinBias.merge.NTUP_HI.r5813_p1729/
- 561 • Run 218391, 8.0 million events from MinBias stream

562 **data13_**hip.**00218391.**physics_MinBias.merge.NTUP_HI.r5813_p1729/
563 • Run 218436, 5.1 million events from MinBias stream
564 **data13_**hip.**00218436.**physics_MinBias.merge.NTUP_HI.r5813_p1729/
565 • Run 218473, 3.9 million events from MinBias stream
566 **data13_**hip.**00218473.**physics_MinBias.merge.NTUP_HI.r5813_p1729/
567 • Run 218589, 3.3 million events from MinBias stream
568 **data13_**hip.**00218589.**physics_MinBias.merge.NTUP_HI.r5813_p1729/
569 • Run 218677, 1.3 million events from MinBias stream
570 **data13_**hip.**00218677.**physics_MinBias.merge.NTUP_HI.r5813_p1729/
571 • Run 218679, 6.9 million events from MinBias stream
572 **data13_**hip.**00218679.**physics_MinBias.merge.NTUP_HI.r5813_p1729/
573 • Run 218716, 8.1 million events from MinBias stream
574 **data13_**hip.**00218716.**physics_MinBias.merge.NTUP_HI.r5813_p1729/
575 • Run 218751, 3.1 million events from MinBias stream
576 **data13_**hip.**00218751.**physics_MinBias.merge.NTUP_HI.r5813_p1729/
577 • Run 218771, 1.7 million events from MinBias stream
578 **data13_**hip.**00218771.**physics_MinBias.merge.NTUP_HI.r5813_p1729/
579 • Run 218783, 2.9 million events from MinBias stream
580 **data13_**hip.**00218783.**physics_MinBias.merge.NTUP_HI.r5813_p1729/
581 • Run 218829, 1.7 million events from MinBias stream
582 **data13_**hip.**00218829.**physics_MinBias.merge.NTUP_HI.r5813_p1729/
583 • Run 218898, 3.1 million events from MinBias stream
584 **data13_**hip.**00218898.**physics_MinBias.merge.NTUP_HI.r5813_p1729/
585 • Run 218940, 4.0 million events from MinBias stream
586 **data13_**hip.**00218940.**physics_MinBias.merge.NTUP_HI.r5813_p1729/
587 • Run 218968, 4.0 million events from MinBias stream
588 **data13_**hip.**00218968.**physics_MinBias.merge.NTUP_HI.r5813_p1729/
589 • Run 219001, 4.0 million events from MinBias stream
590 **data13_**hip.**00219001.**physics_MinBias.merge.NTUP_HI.r5813_p1729/
591 • Run 219028, 1.6 million events from MinBias stream

592 **data13_hip.00219028.physics_MinBias.merge.NTUP_HI.r5813_p1729/**
 593 • Run 219055, 4.1 million events from MinBias stream
 594 **data13_hip.00219055.physics_MinBias.merge.NTUP_HI.r5813_p1729/**
 595 • Run 219089, 3.5 million events from MinBias stream
 596 **data13_hip.00219089.physics_MinBias.merge.NTUP_HI.r5813_p1729/**
 597 • Run 219111, 5.5 million events from MinBias stream
 598 **data13_hip.00219111.physics_MinBias.merge.NTUP_HI.r5813_p1729/**
 599 • Run 219114, 2.0 million events from MinBias stream
 600 **data13_hip.00219114.physics_MinBias.merge.NTUP_HI.r5813_p1729/**

601 Similar as pp , the triggers applied in 5 TeV $p+\text{Pb}$ have two components: MinBias and HMT [11, 34].
 602 A list of all the major MinBias and HMT triggers used in this analysis is summarized as follows:

- 603 • EF_mbMbts_1_1_counter
- 604 • EF_hip_trk100_TE10_counter
- 605 • EF_hip_trk130_TE10_counter
- 606 • EF_hip_trk150_TE50_counter
- 607 • EF_hip_trk185_TE50_counter
- 608 • EF_hip_trk200_TE65_counter
- 609 • EF_hip_trk225_TE65_counter

610 where a new trigger item has been introduced:

- 611 • mbMbts_1_1: HLT trigger requires at least 1 hit on both sides of MBTS;

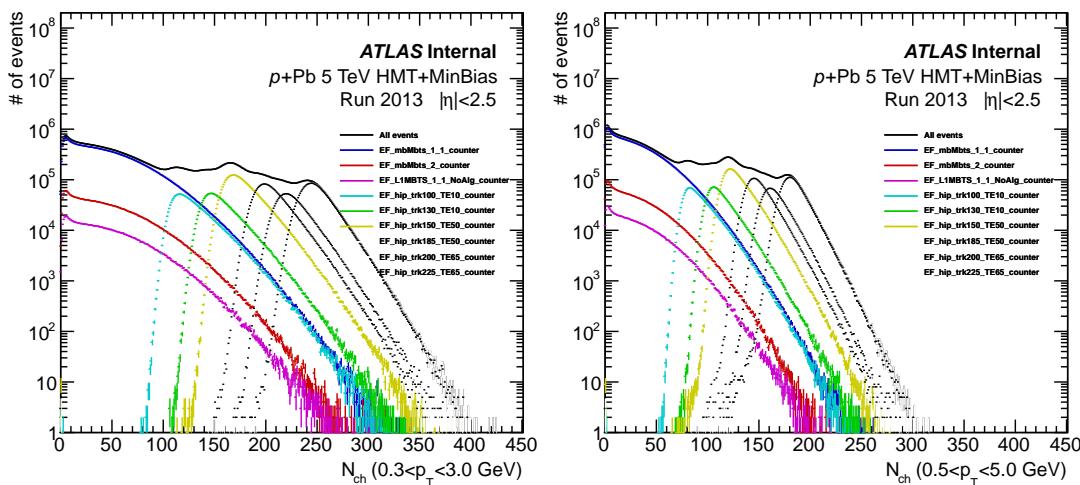
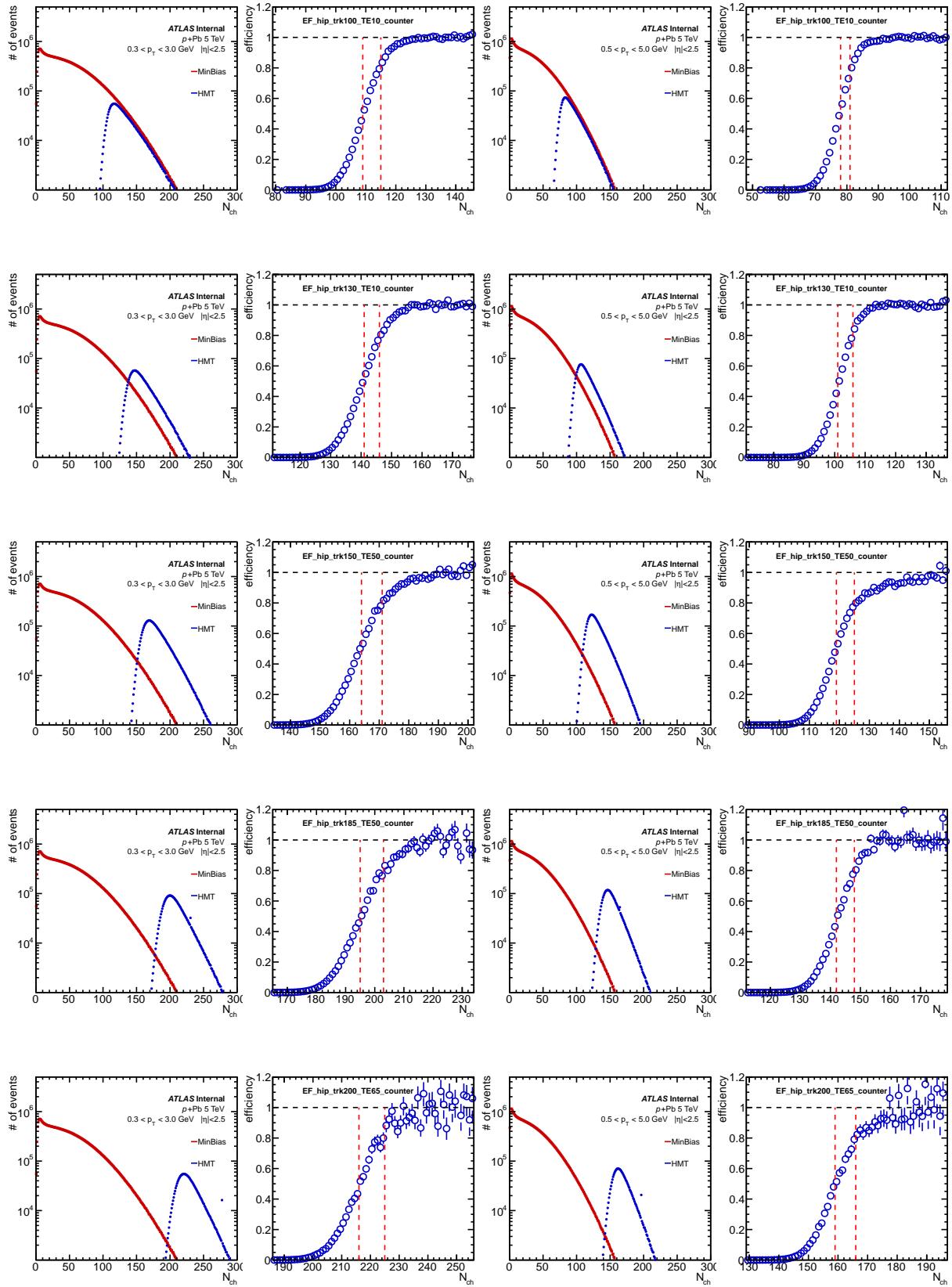


Figure 25: Distribution of number of tracks with two p_T thresholds: $0.3 < p_T < 3.0$ GeV and $0.5 < p_T < 5.0$ GeV, in 5 TeV $p+\text{Pb}$ run 2013. The major MinBias and HMT triggers are plotted separately.

612 The summary of statistics with all the major triggers used in this analysis are shown in Fig. 25.



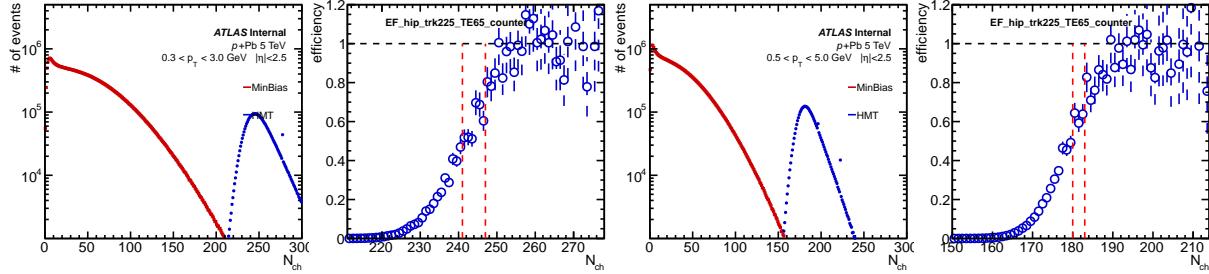


Figure 26: Trigger efficiencies of all major HMT triggers as a function of number of tracks in two p_T ranges: $0.3 < p_T < 3.0$ GeV and $0.5 < p_T < 5.0$ GeV, from 5 TeV $p+Pb$ run. Efficiency is calculated relative to the corresponding MinBias trigger in this run period then scaled to 1.0 in the large N_{ch} region. The two red dash lines indicate 50% and 80% efficiency cuts.

613 Trigger efficiencies of all the major HMT triggers are summarized in Fig. 26, where efficiencies are
614 shown for two p_T ranges separately: $0.3 < p_T < 3.0$ GeV and $0.5 < p_T < 5.0$ GeV.

615 3.3.2 HIJING for 2013 $p+Pb$ data

616 Tracking efficiency and fraction of fake tracks have been extensively studied in previous $p+Pb$ flow
617 analyses [11, 34]. In this analysis we are re-using the efficiency map applied in the forward-backward
618 multiplicity fluctuation paper ATL-COM-PHYS-2015-655, where the tracking efficiency is estimated as a
619 function of η , p_T , N_{ch} and z position of the vertex.

620 3.3.3 2016 $p+Pb$

621 In 2016, 5.02 TeV $p+Pb$ data was collected with ATLAS detectors, with lower μ value than 2013. In
622 additional to GRL selection, each event is required to have at least one vertex. Besides, problematic
623 events are also removed:

- 624 • due to the liquid argon system
- 625 • due to the tile calorimeter system
- 626 • due to the SCT inner detector system
- 627 • due to incomplete events (event information missing after TTC restarts)

628 and since the μ value of this year's run is much lower than 2013, cleaning pile-up events is not crucial.
629 In the later section, with 2016 8.16 TeV $p+Pb$ data, we will show that for events with $\mu < 0.01$, pile-up
630 effects are negligible.

631 Since this year's $p+Pb$ run is right after the pp run, the track selection criteria is identical to pp . An
632 additional tighter d_0 and z_0 pointing cut is used to check the stability of results in the systematics section.
633 All the runs included in this analysis are listed as follows:

- 634 • Run 312649, 8.9 million events from MinBias stream
635 data16_hip5TeV.00312649.physics_Main.recon.AOD.f784_m1741
- 636 • Run 312796, 43 million events from MinBias stream
637 data16_hip5TeV.00312796.physics_Main.recon.AOD.f784_m1741

- 638 • Run 312837, 85 million events from MinBias stream
- 639 data16_hip5TeV.00312837.physics_Main.recon.AOD.f774_m1736
- 640 • Run 312937, 26 million events from MinBias stream
- 641 data16_hip5TeV.00312937.physics_Main.recon.AOD.f774_m1736
- 642 • Run 312945, 29 million events from MinBias stream
- 643 data16_hip5TeV.00312945.physics_Main.recon.AOD.f774_m1736
- 644 • Run 312968, 37 million events from MinBias stream
- 645 data16_hip5TeV.00312968.physics_Main.recon.AOD.f774_m1736
- 646 • Run 314199, 240 million events from MinBias stream
- 647 data16_hip5TeV.00314199.physics_Main.recon.AOD.f781_m1741

648 Due to the reason that the intermediate N_{ch} region is not covered by HMT triggers, several HMT
 649 triggers with intermediate N_{ch} thresholds are specially designed before 2016 data taking. Together with
 650 MinBias triggers, all the major triggers used in this analysis is summarized as follows:

- 651 • HLT_noalg_mb_L1MBTS_1
- 652 • HLT_noalg_mb_L1MBTS_1_1
- 653 • HLT_mb_sp100_trk10_hmt_L1MBTS_1_1
- 654 • HLT_mb_sp100_trk20_hmt_L1MBTS_1_1
- 655 • HLT_mb_sp100_trk30_hmt_L1MBTS_1_1
- 656 • HLT_mb_sp100_trk60_hmt_L1MBTS_1_1
- 657 • HLT_mb_sp100_trk80_hmt_L1MBTS_1_1
- 658 • HLT_mb_sp100_trk100_hmt_L1MBTS_1_1
- 659 • HLT_mb_sp100_trk110_hmt_L1MBTS_1_1

660 where all the HMT triggers are seeded on L1MBTS_1_1, which is different from 2013 since the luminosity
 661 is much lower.

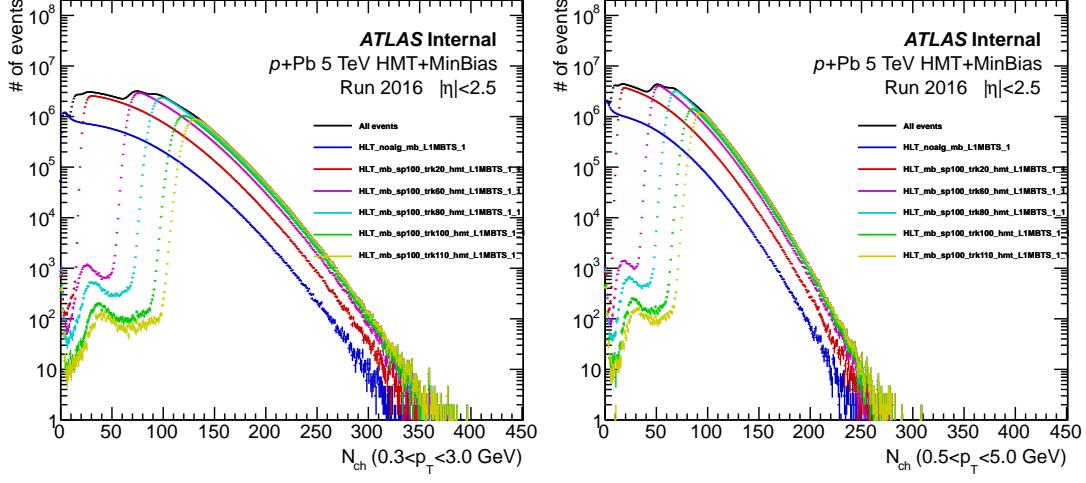
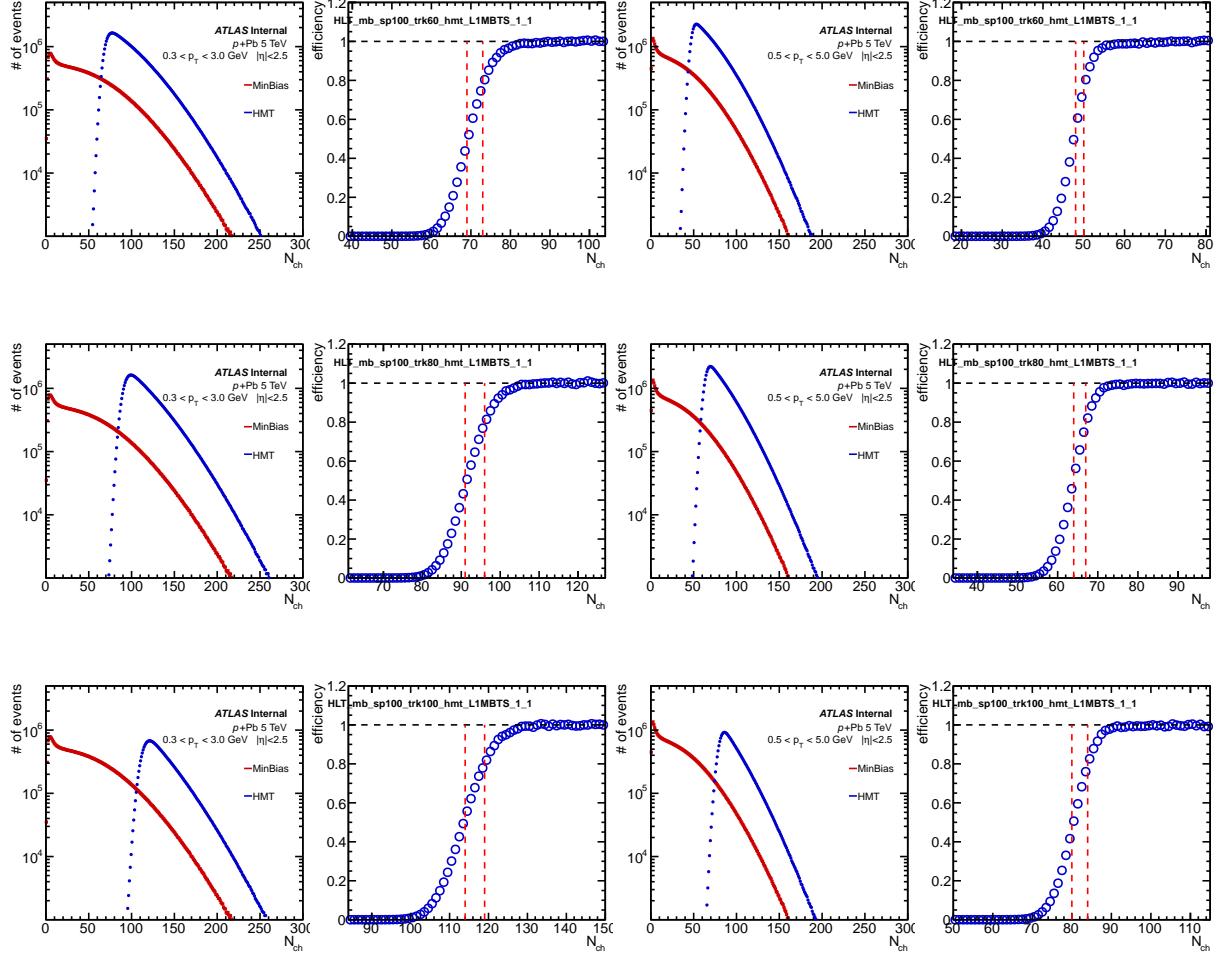


Figure 27: Distribution of number of tracks with two p_T thresholds: $0.3 < p_T < 3.0$ GeV and $0.5 < p_T < 5.0$ GeV, in 5 TeV $p+Pb$ run 2016. The major MinBias and HMT triggers are plotted separately.

662 The summary of statistics with all the major triggers used in this analysis are shown in Fig. 27.



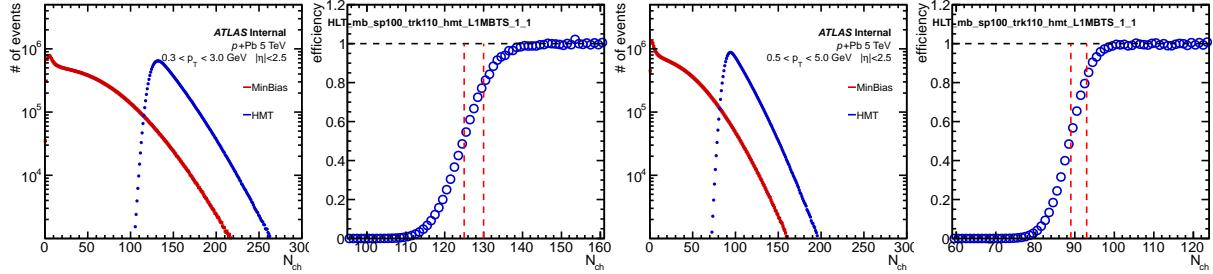


Figure 28: Trigger efficiencies of all major HMT triggers as a function of number of tracks in two p_T ranges: $0.3 < p_T < 3.0$ GeV and $0.5 < p_T < 5.0$ GeV, from 2016 5 TeV $p+Pb$ run. Efficiency is calculated relative to the corresponding MinBias trigger in this run period then scaled to 1.0 in the large N_{ch} region. The two red dash lines indicate 50% and 80% efficiency cuts.

Trigger efficiencies of all the major HMT triggers are summarized in Fig. 28, where efficiencies are shown for two p_T ranges separately: $0.3 < p_T < 3.0$ GeV and $0.5 < p_T < 5.0$ GeV.

3.3.4 HIJING for 2016 $p+Pb$ data

By comparing Fig. 25 and Fig. 27, it is obvious that 2016 collected much more statistics than 2013 in low and intermediate N_{ch} region. Since cumulant measurement is dominated by statistics errors, 2016 5 TeV $p+Pb$ data will be included in this analysis. There existed 1 million HIJING [35] sample before this year's run with Run 2 configuration. In this section, we will show the this HIJING sample is sufficient for the tracking efficiency estimation.

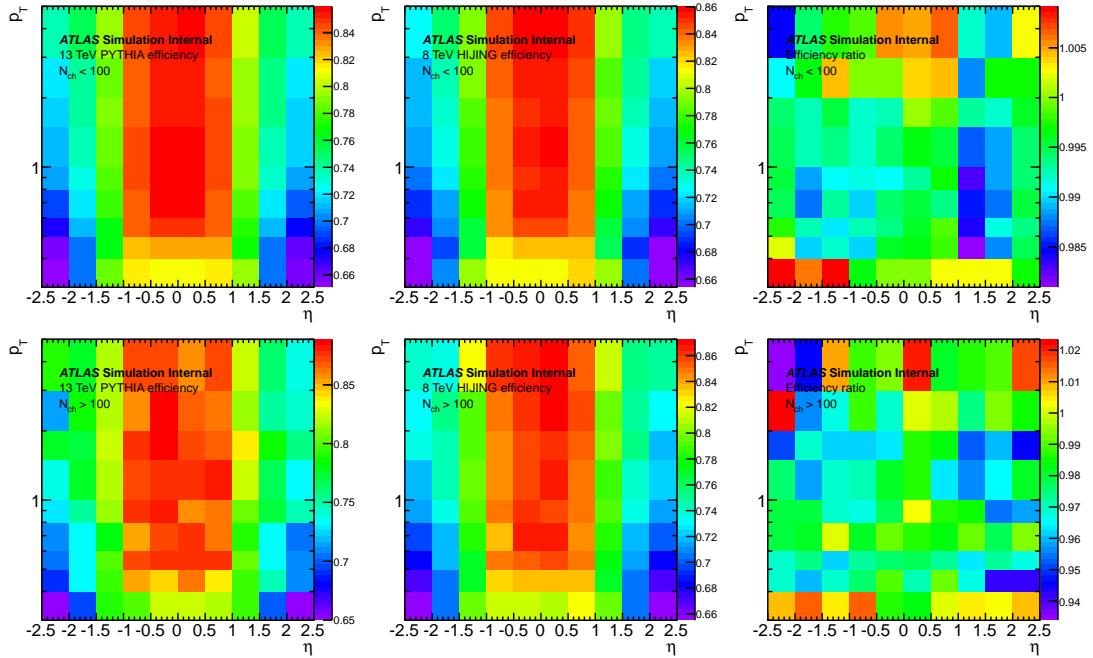


Figure 29: Tracking efficiency as a function of η and p_T . Top three panels are estimated in $N_{ch} < 100$ and bottom panels are for $N_{ch} > 100$. The 1st column is the tracking efficiency estimated from PYTHIA with Run 2 configuration; 2nd column is from the HIJING sample before 2016 5 TeV $p+Pb$ run and 3rd column is the ratio between HIJING and PYTHIA.

Fig. 29 shows the comparisons of tracking efficiencies as a function of η and p_T . Based on the ratio plot, even though tracking efficiencies are estimated from PYTHIA and HIJING, since both the Monte-Carlo samples have Run 2 configurations, the relative differences are within 2% for $N_{ch} < 100$. For $N_{ch} > 100$, the relative difference is larger because of larger statistical errors in larger N_{ch} region. This result means that the tracking efficiency stays stable in Run 2, from Year 2015 to 2016. Thus one should not expect big changes for 2016 p +Pb data.

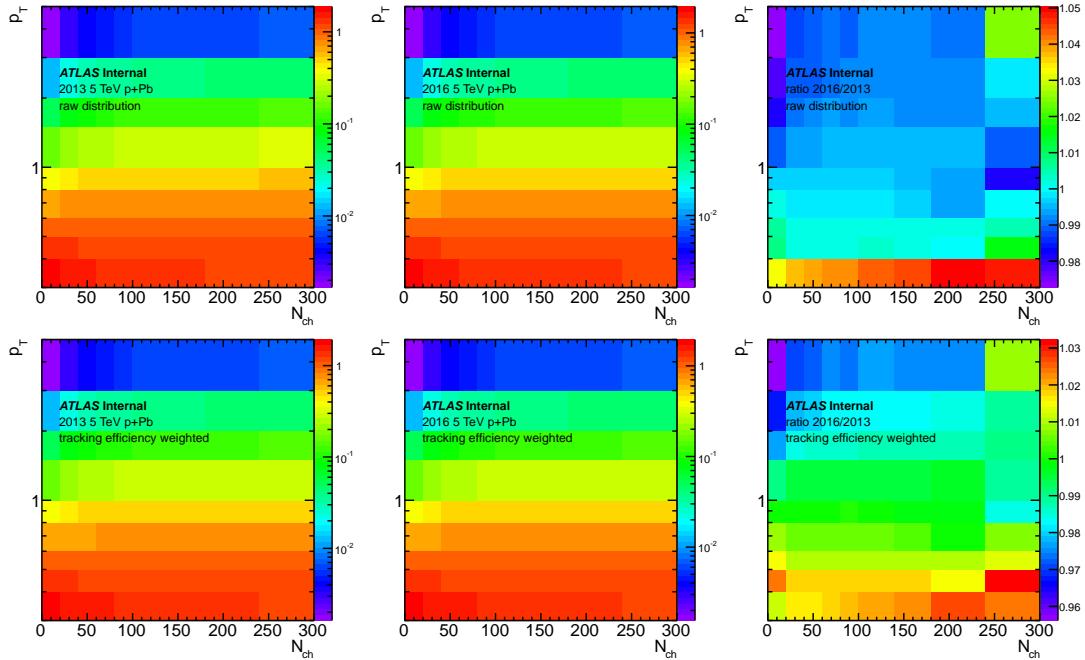


Figure 30: Particle distribution as a function of N_{ch} and p_T . Top three panels are the raw distribution while bottom three panels are weighted by inverse of the tracking efficiency. The 1st column is from 2013 5 TeV p +Pb, 2nd column is from 2016 5 TeV p +Pb and 3rd column is the ratio between 2016 and 2013.

Fig. 30 shows the comparison of particle distributions as a function of N_{ch} and p_T , using 2013 and 2016 5 TeV p +Pb data. From ratio of raw distributions, 2016 data reconstructs 5% more tracks in the lowest p_T region. The reason is that in Run 2, IBL was implemented to the inner detector, which provides better tracking reconstruction in lower p_T region. Meanwhile, once the particles are weighted by the inverse of tracking efficiencies, the difference reduces down to 2%. This means the HIJING sample we used correctly estimated the tracking efficiency in the new data, since one should not expect any difference between the p_T spectrum from 2013 and 2016. The remaining 1% to 2% difference in the lowest p_T region is negligible since in the systematic checks of tracking efficiency, a difference on the level of 10% has been evaluated and included as one of the systematics.

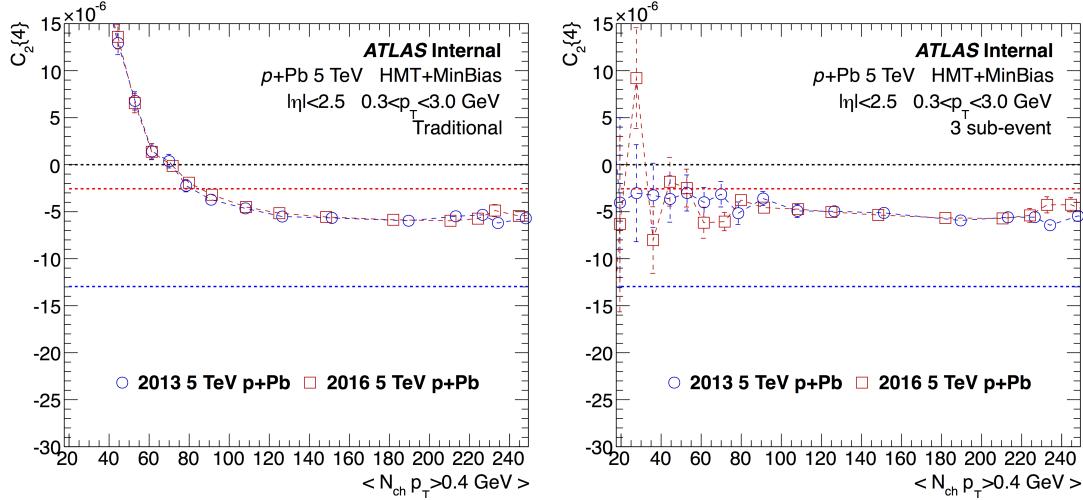


Figure 31: Comparison of $C_2\{4\}$ measured in 2013 and 2016 5 TeV $p+Pb$ data, with traditional method (left) and 3 sub-event method (right). Particles have $0.3 < p_T < 3.0 \text{ GeV}$.

686 Last but not least, it is worthwhile to directly compare the physics observable using 2013 and 2016
 687 data, with tracking efficiency correction. As shown in Fig. 31, $C_2\{4\}$ measured in the 2013 and 2016
 688 are consistent within statistical errors, for both traditional and 3 sub-event methods. With all the three
 689 evidences listed above, we decide to use this mentioned HIJING sample to estimate tracking efficiency
 690 and 2016 5 TeV $p+Pb$ has been included in the final results.

691 4 Data analysis

692 4.1 Outline

693 The analysis is carried out in the following steps:

- 694 • p_T range of the particles;
- 695 • Calculate the Q -vector event-by-event;
- 696 • Calculate the 2- and 4-particle correlation event-by-event;
- 697 • Define event classes;
- 698 • Determine the mean value of N_{ch} with $p_T > 0.4$ GeV;
- 699 • Calculate the mean value of 2- and 4-particle correlation in each event class;
- 700 • Calculate the 2- and 4-particle cumulant in each event class;
- 701 • Calculate the corresponding flow signal from 2- and 4-particle cumulant.

702 Since the cumulant signal in small system is dominated by non-flow contribution and small changes
 703 in the analysis procedure may potentially change the final results. In this section, we will go through the
 704 above procedures step by step in details, using 13 TeV pp collision system as one example. The same
 705 procedures are applied across all the other systems.

706 4.2 p_T range of the reference particles N_{ch}^{ref}

707 From 2PC measurements we know that v_2 signal increases then decreases as a function of p_T , in order
 708 to compare our cumulant results with existing v_2 measurements using other methods, we choose the p_T
 709 range of the particles as following:

- 710 • $0.3 < p_T < 3.0$ GeV: to compare with CMS measurements, where they used peripheral subtraction
 711 and traditional cumulant methods;
- 712 • $0.5 < p_T < 5.0$ GeV: to compare with ATLAS measurement, where they used peripheral subtraction
 713 method with template fitting.

714 Since later we will use particles from other p_T ranges for other purposes, we stress that the p_T ranges
 715 listed above are used to select particles for Q -vector measurement, which is directly related to the final
 716 cumulant results. It is denoted as reference particle N_{ch}^{ref} .

717 4.3 Calculation of Q -vector

718 To reduce the CPU burden of calculating multi-particle correlation, Q-cumulant method (or direct cumu-
 719 lant) is developed to calculate the event-by-event Q -vector, without multiple loops through all the tracks.

720 Since multi-particle correlation only counts distinguishable pairs, the duplicates, or self-correlation, need
 721 to be removed while calculating the 2- and 4-particle correlations.

722 For each event, the Q -vector is calculated by summing all the reference particles:

$$723 Q_{n,k} \equiv \sum_{i=1}^{N_{ch}^{ref}} w_i^k e^{in\phi_i} \quad (12)$$

723 where i loops through all the reference particles. Index n denotes the order of flow harmonics. In this
 724 analysis we will mainly focus on the second harmonic where $n = 2$. With presence of the tracking
 725 efficiency and detector effects, we introduce power k into the Q -vector definition. Each reference particle
 726 is weighted by the track-weight w_i^k , which accounts for the tracking efficiency and detector effects:

$$w_i \equiv \frac{w_\phi(\eta_i, \phi_i)(1 - f(\eta_i, p_{T,i}))}{\epsilon(\eta_i, p_{T,i})} \approx \frac{w_\phi(\eta_i, \phi_i)}{\epsilon(\eta_i, p_{T,i})} \quad (13)$$

727 where w_ϕ is the track-weight calculated from flattening produce (we will discuss it shortly), to further
 728 remove the detector effects. ϵ and f are the tracking efficiency and fake fraction evaluated from the
 729 Monte-Carlo. It has been shown that the maximum fake fraction is on the level of 0.1% so we will not
 730 correct the fake fraction for simplicity.

731 4.4 Flattening procedure

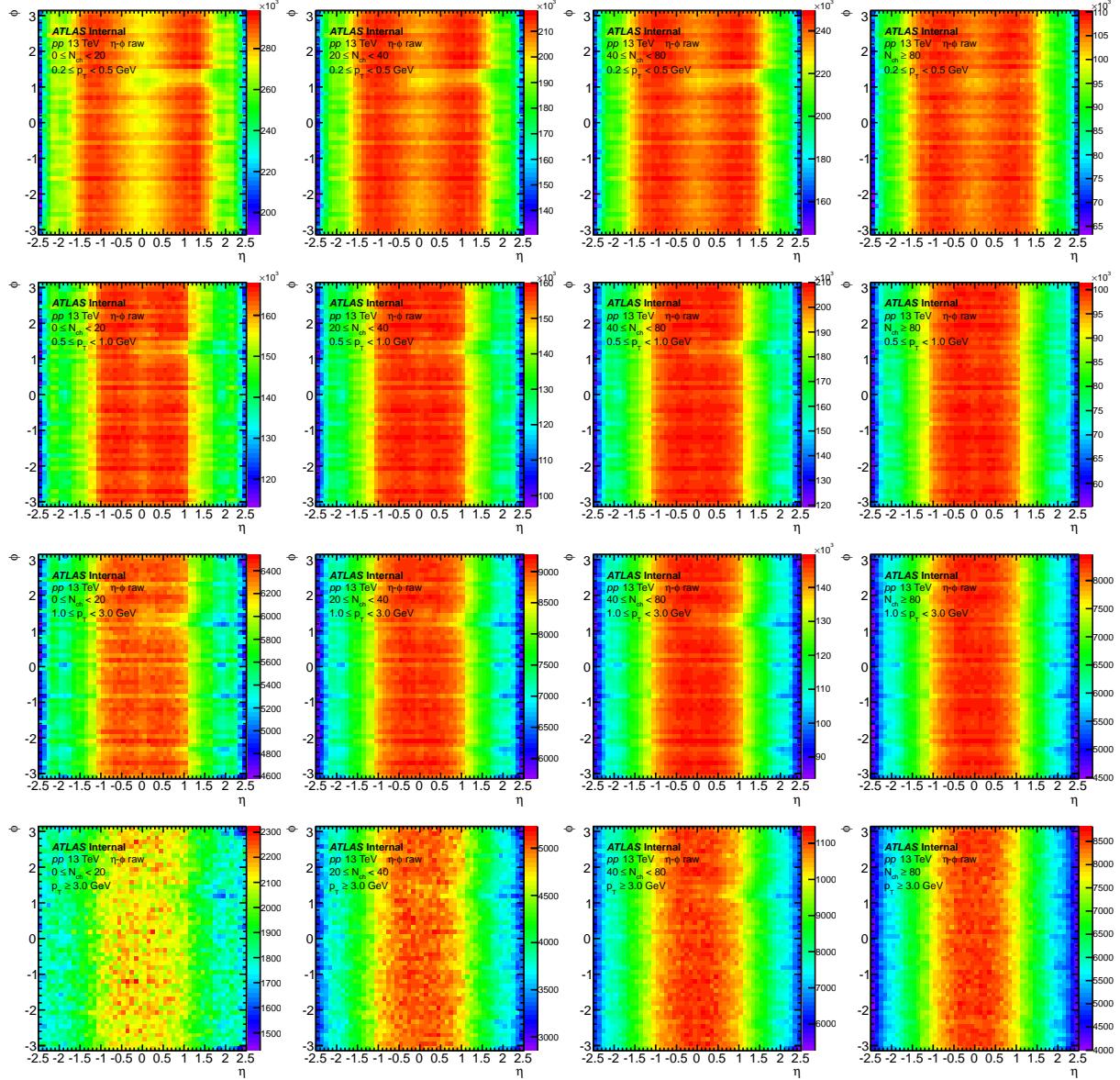


Figure 32: Raw $N(\eta, \phi)$ distribution of reconstructed tracks. Different columns are for different N_{ch} bins and different rows are from different p_T ranges. The uniformness in η is independent of N_{ch} and p_T .

732 The detector imperfections may result in a non-uniform distribution of particle azimuthal angles and such
 733 non-uniform distribution may vary from run to run. Fig. 32 shows the $N(\eta, \phi)$ distribution from four 13
 734 TeV pp runs. The $N(\eta, \phi)$ is determined by counting number of reconstructed tracks, without tracking
 735 efficiency correction, in narrow ($\delta\eta, \delta\phi$) slices. The different panels are for different N_{ch} bins and p_T
 736 ranges, cumulated over all the runs. By comparing the different panels, it is evident that the raw $N(\eta, \phi)$
 737 is not strongly dependent of N_{ch} and p_T . In the following plots, we will merge the all the N_{ch} and p_T and
 738 only compare the run dependence.

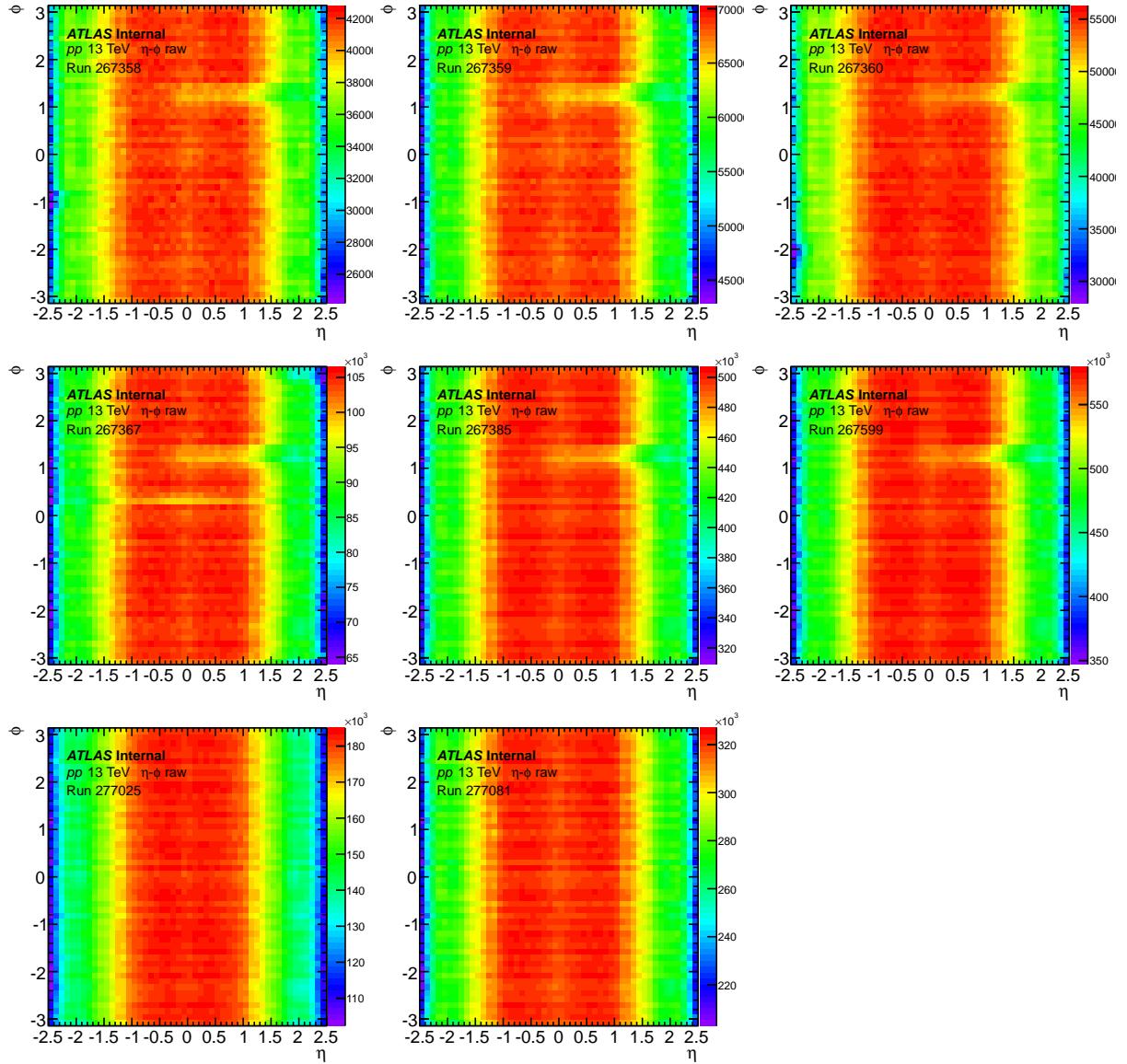


Figure 33: Raw $N(\eta, \phi)$ distribution of reconstructed tracks. Different panels are from different pp runs, where first 6 runs are from the low- μ period and the last 2 runs are from the intermediate- μ runs. Intermediate- μ runs are more uniform in ϕ than low- μ runs.

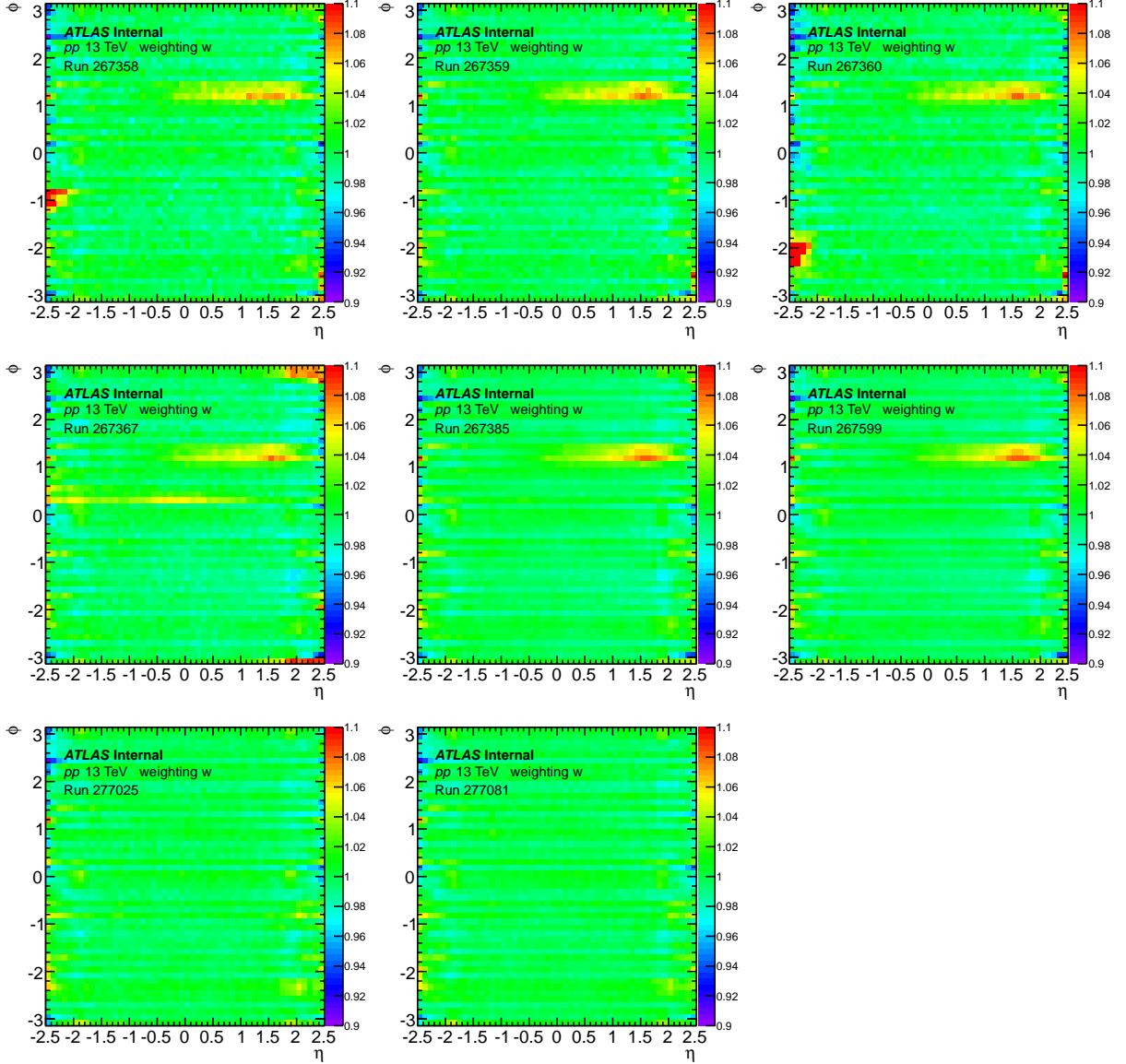


Figure 34: Raw $N(\eta, \phi)$ distribution of reconstructed tracks. Different panels are from different pp runs, where first 6 runs are from the low- μ period and the last 2 runs are from the intermediate- μ runs. Intermediate- μ runs are more uniform in ϕ than low- μ runs.

739 Fig. 33 shows the raw $N(\eta, \phi)$ distribution of reconstructed tracks, for 8 different runs. In the eight
 740 runs, 267358, 267359, 267360, 267367, 267385, 267599 are from the low- μ run period, while 277025
 741 and 277081 are from the intermediate- μ run period. In the low- μ run period, some detector effects can
 742 be clearly seen: at $\phi \approx 1.2$ and $0 < \eta < 1.5$. To better show these detector effects, we plot the weight w_ϕ for
 743 different runs in Fig. 34. ϕ weighting w_ϕ is defined as:

$$w_\phi = \frac{\langle N(\delta\eta) \rangle}{N(\delta\eta, \delta\phi)} \quad (14)$$

744 where $\langle N(\delta\eta, \delta\phi) \rangle$ is the mean value of number of reconstructed tracks in small $\delta\eta$ slice, averaged over
 745 ϕ . In the low- μ run period, some clear hot spots have been seen which is absent in intermediate- μ period.
 746 Based on these, we will evaluate and apply the w_ϕ weighting run-by-run.

To show how the flattening works, Fig. 35 shows the $N(\eta, \phi)$ distribution after w_ϕ weighting, for 8 different runs. It is clear that after the correction all the runs are very uniform in ϕ , while the small residual structures are due to the interpolation while applying the weights. In the systematics section we will evaluate the changes of results due to this flattening procedure and include them into the systematics.

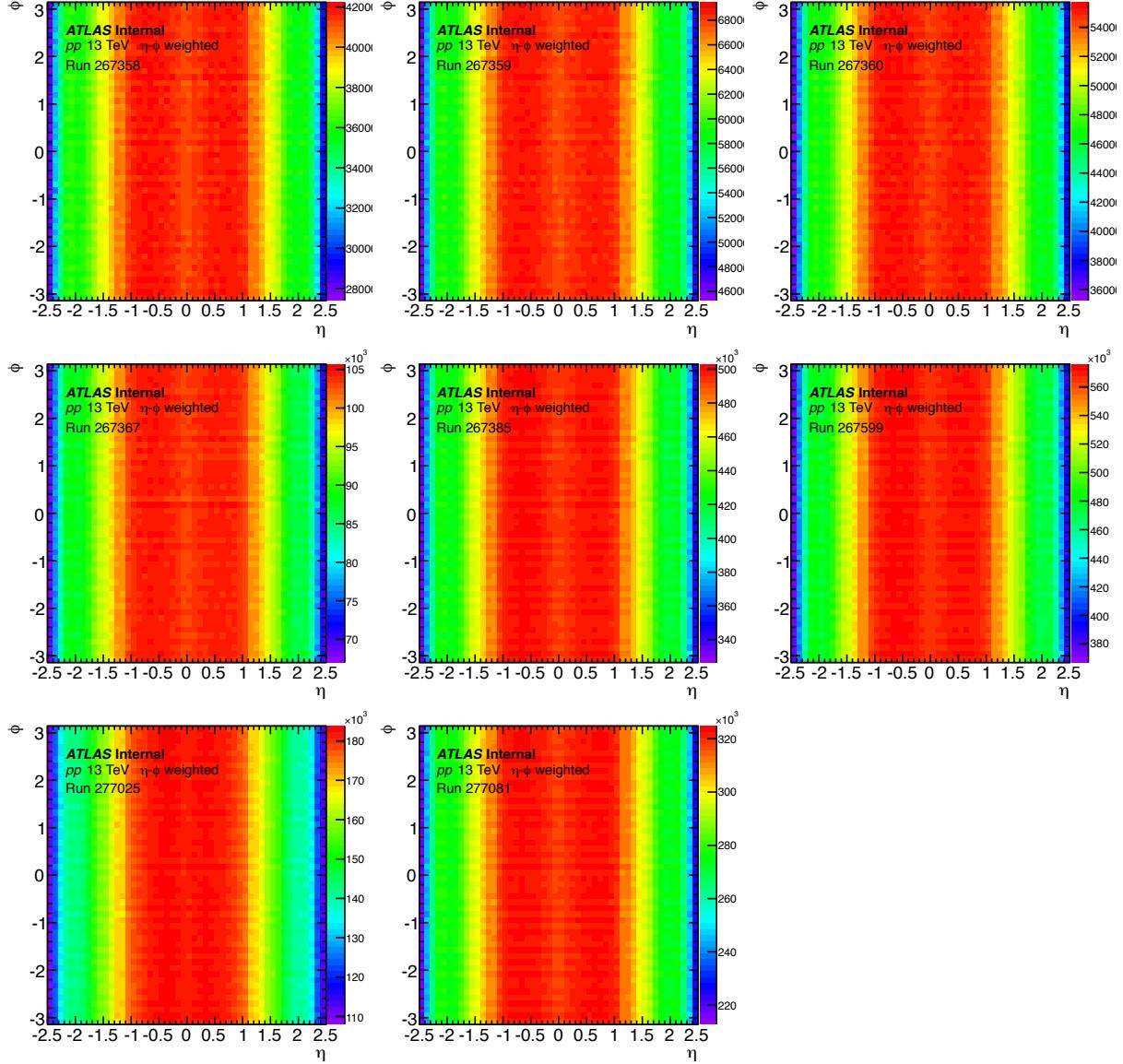


Figure 35: $N(\eta, \phi)$ distribution of reconstructed tracks, with w_ϕ weighting. Different panels are from different pp runs, where first 6 runs are from the low- μ period and the last 2 runs are from the intermediate- μ runs. After applying the w_ϕ weights run-by-run, all runs are now equally uniform.

751 4.5 2- and 4-particle correlations with direct cumulant method

752 In this section, we will extend all the formulas introduced in the methodology section to the cases with
 753 track weights. The 2- and 4-particle correlations with track weight are written as:

$$\begin{aligned} corr_n\{2\}_{ev} &\equiv \frac{1}{W_{(2)}} \sum'_{i,j=1} w_i w_j e^{in(\phi_i - \phi_j)} \\ corr_n\{4\}_{ev} &\equiv \frac{1}{W_{(4)}} \sum'_{i,j,k,l=1} w_i w_j w_k w_l e^{in(\phi_i + \phi_j - \phi_k - \phi_l)} \end{aligned} \quad (15)$$

754 where w_i is the track weight and $'$ in the summation means $i \neq j$ and $i \neq j \neq k \neq l$. $W_{(2)}$ and $W_{(4)}$ are the
 755 unique number of pairs weighted by track weight:

$$\begin{aligned} W_{(2)} &\equiv \sum'_{i,j=1} w_i w_j \\ W_{(4)} &\equiv \sum'_{i,j,k,l=1} w_i w_j w_k w_l \end{aligned} \quad (16)$$

756 To reduce the CPU burden, we will follow the direct cumulant method, and derive the new formulas
 757 for sub-event methods with particle weight. The case without particle weight has been introduced in
 758 Section 2. All the formulas are discussed in the Appendix 8.

759 4.6 Definition of event classes

760 After 2- and 4-particle correlations are calculated event-by-event, many similar events are merged ac-
 761 cording to event class criteria. The optimal event class definition should be based on the N_{ch}^{ref} particles
 762 that have been used in the cumulant calculation, since the cumulant is very sensitive to the multiplicity
 763 fluctuation, and in this definition the multiplicity fluctuation is minimum. However, in order to show how
 764 the multiplicity fluctuation can affect the final results, we also define the three other event class criteria
 765 N_{ch}^{evtCls} , one of which has been applied in CMS recent paper. All the event class definitions are listed
 766 below, for two N_{ch}^{ref} classes separately:

- 767 • For reference particles N_{ch}^{ref} with $0.3 < p_T < 3.0$ GeV, event classes are defined by:
 - 768 – Particles with $0.3 < p_T < 3.0$ GeV;
 - 769 – Particles with $p_T > 0.2$ GeV;
 - 770 – Particles with $p_T > 0.4$ GeV;
- 771 • For reference particles N_{ch}^{ref} with $0.5 < p_T < 5.0$ GeV, event classes are defined by:
 - 772 – Particles with $0.5 < p_T < 5.0$ GeV;
 - 773 – Particles with $p_T > 0.4$ GeV;
 - 774 – Particles with $p_T > 0.6$ GeV;

775 where for each N_{ch}^{ref} p_T ranges, neighbouring p_T cuts were used to define the event class. Note that
 776 both N_{ch}^{ref} and N_{ch}^{evtCls} are tracking efficiency and flattening weighted: w_ϕ/ϵ . To further minimize the
 777 multiplicity fluctuation, the bin width of N_{ch} for each event class is always 1. The results will be merged
 778 on the final cumulant level, to gain sufficient statistics for small negative signal.

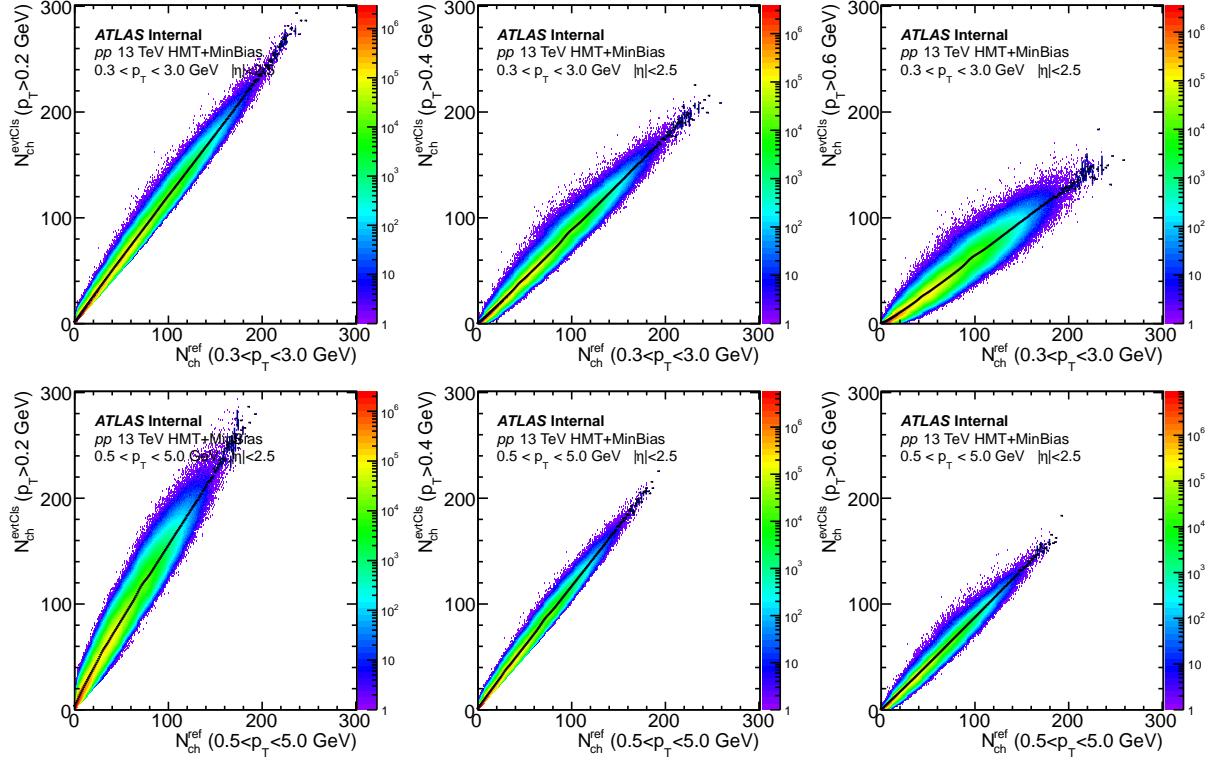


Figure 36: Correlation of N_{ch}^{ref} and $N_{ch}^{evtClis}$, in 2015 13 TeV pp . Different rows are two p_T ranges for N_{ch}^{ref} : $0.3 < p_T < 3.0 \text{ GeV}$ and $0.5 < p_T < 5.0 \text{ GeV}$ and different columns are three p_T ranges for $N_{ch}^{evtClis}$: $p_T > 0.2 \text{ GeV}$, $p_T > 0.4 \text{ GeV}$ and $p_T > 0.6 \text{ GeV}$. The correlation strengths are very different for various combinations.

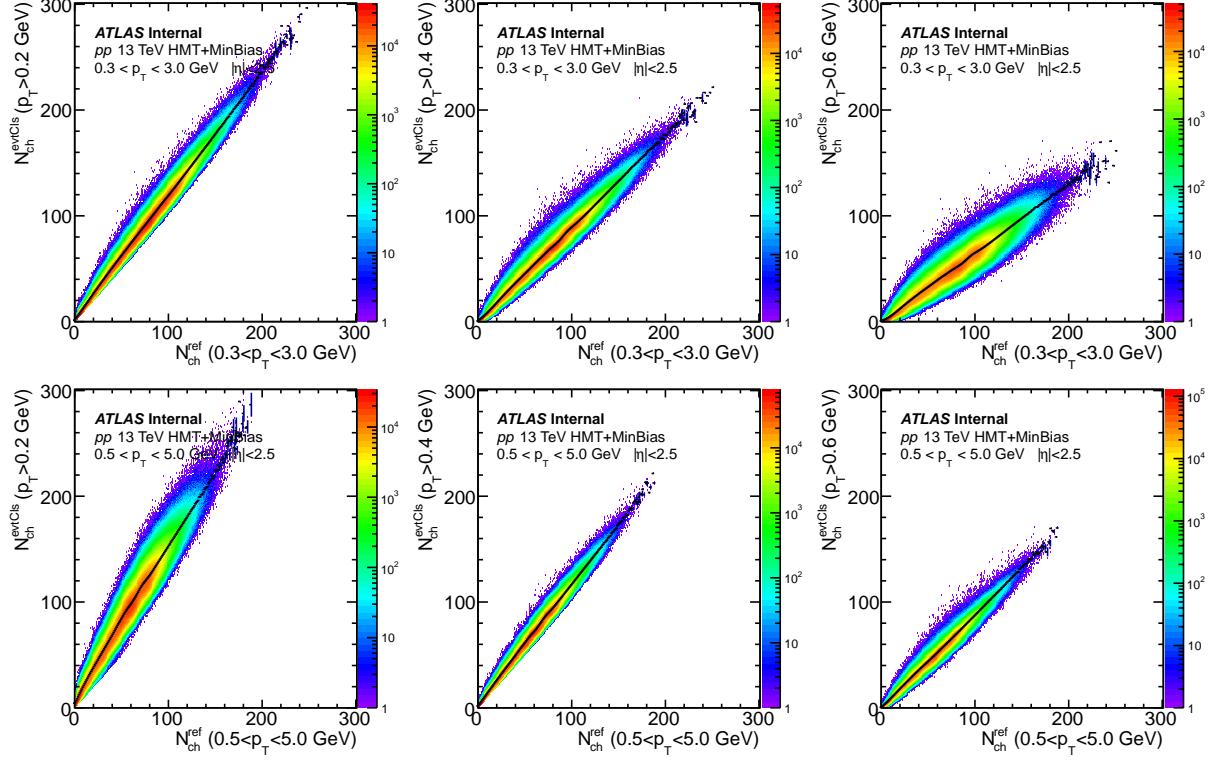


Figure 37: Correlation of N_{ch}^{ref} and N_{ch}^{evtCls} , in 2016 13 TeV pp . Different rows are two p_T ranges for N_{ch}^{ref} : $0.3 < p_T < 3.0 \text{ GeV}$ and $0.5 < p_T < 5.0 \text{ GeV}$ and different columns are three p_T ranges for N_{ch}^{evtCls} : $p_T > 0.2 \text{ GeV}$, $p_T > 0.4 \text{ GeV}$ and $p_T > 0.6 \text{ GeV}$. The correlation strengths are very different for various combinations.

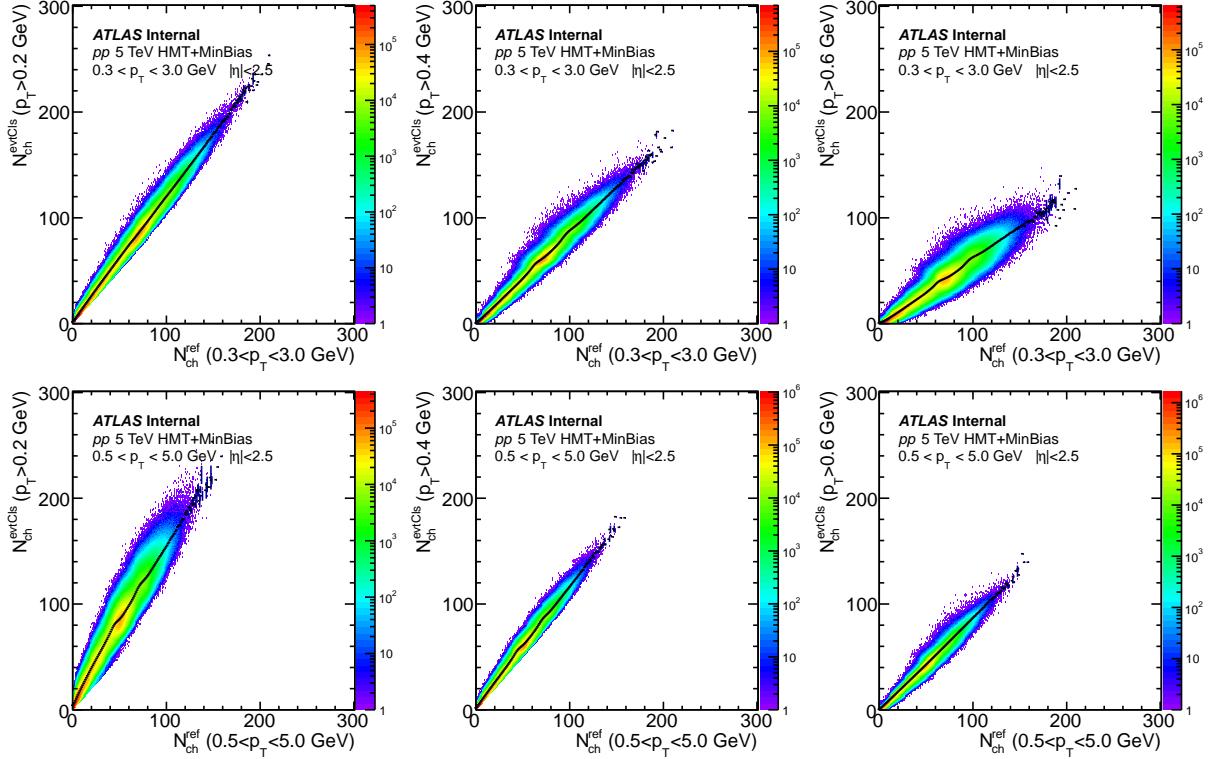


Figure 38: Correlation of N_{ch}^{ref} and N_{ch}^{evtCls} , in 2015 5 TeV pp . Different rows are two p_T ranges for N_{ch}^{ref} : $0.3 < p_T < 3.0 \text{ GeV}$ and $0.5 < p_T < 5.0 \text{ GeV}$ and different columns are three p_T ranges for N_{ch}^{evtCls} : $p_T > 0.2 \text{ GeV}$, $p_T > 0.4 \text{ GeV}$ and $p_T > 0.6 \text{ GeV}$. The correlation strengths are very different for various combinations.

779 The correlations between N_{ch}^{ref} and N_{ch}^{evtCls} , for various combinations, are summarized in Fig. 36. For
 780 N_{ch}^{ref} with $0.3 < p_T < 3.0 \text{ GeV}$, the correlation is weakest for N_{ch}^{evtCls} with $p_T > 0.6 \text{ GeV}$. While for N_{ch}^{ref}
 781 with $0.5 < p_T < 5.0 \text{ GeV}$, the correlation is weakest for N_{ch}^{evtCls} with $p_T > 0.2 \text{ GeV}$. As will be seen in
 782 the results section, the different event binning will give different traditional cumulant results. A simple
 783 explanation for this is that by defining the different event class criteria with different p_T , different events
 784 with different non-flow contributions are mixed. Previous studies have shown that the cumulant mea-
 785 surement is sensitive to flow fluctuation, one could imagine that the cumulant should also be sensitive to
 786 the non-flow (this conclusion can be tested in our method paper on the same topic). If there is remaining
 787 non-flow in the cumulant, by re-arranging different events into the same group, how the non-flow fluc-
 788 tuates is changed, which results in different $C_2\{4\}$ values. In principle, one could verify the changes in
 789 non-flow and flow fluctuations by measuring Q -distributions in different event classes. However, since
 790 the changes due to event class binning are much smaller than the statistical fluctuation, it is impractical
 791 to distinguish the differences with current statistics. We have tested this approach in PYTHIA, where
 792 only non-flow fluctuation exists, and found that the Q -distributions from different event classes are hard
 793 to tell.

794 4.7 X-axis of cumulant results

795 Since there are two p_T ranges for the reference particles and three p_T ranges for the particles for event
 796 class definition. In order to properly compare all the results, the same X-axis need to be defined. Previous
 797 flow measurements all choose N_{ch} with $p_T > 0.4 \text{ GeV}$ as X-axis, we will also follow this convention and

calculate the mean value of N_{ch} with $p_T > 0.4$ GeV in each event class. The number of particles to determine the X-axis is denoted as N_{ch}^X .

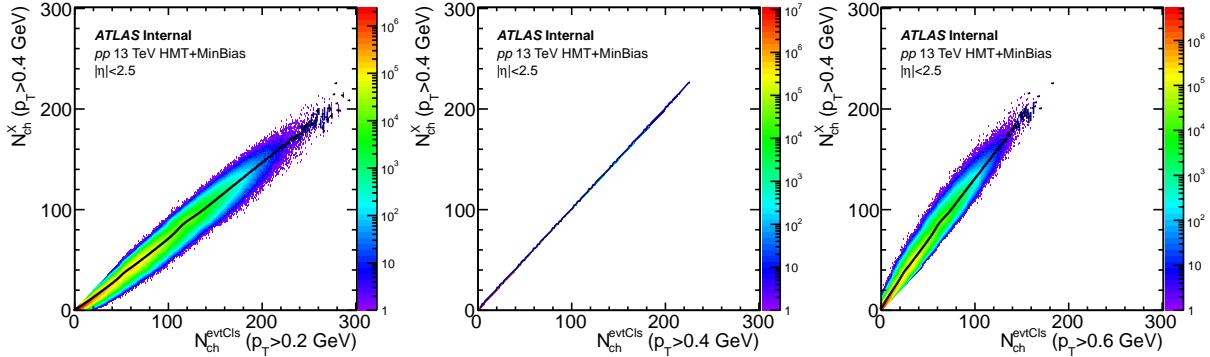


Figure 39: Correlation of N_{ch}^X and N_{ch}^{evtCls} , in 2015 13 TeV pp , and there are three N_{ch}^{evtCls} with different p_T ranges: $p_T > 0.2$ GeV, $p_T > 0.4$ GeV and $p_T > 0.6$ GeV. As a convention, the N_{ch}^X is defined as number of particles with $p_T > 0.4$ GeV, which explains why in the second plot the correlation factor is 1. The black dots on the correlations are profiles, which provides the $\langle N_{ch}(p_T > 0.4\text{GeV}) \rangle$ as the X-axis of cumulant results.

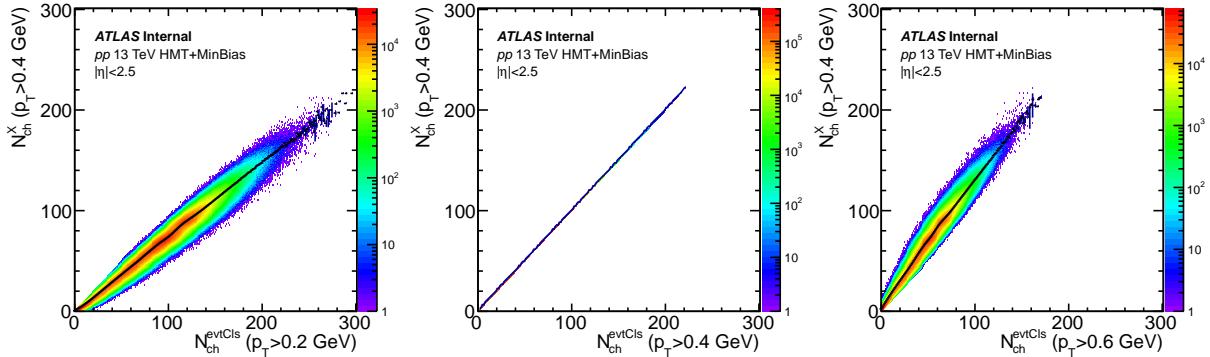


Figure 40: Correlation of N_{ch}^X and N_{ch}^{evtCls} , in 2016 13 TeV pp , and there are three N_{ch}^{evtCls} with different p_T ranges: $p_T > 0.2$ GeV, $p_T > 0.4$ GeV and $p_T > 0.6$ GeV. As a convention, the N_{ch}^X is defined as number of particles with $p_T > 0.4$ GeV, which explains why in the second plot the correlation factor is 1. The black dots on the correlations are profiles, which provides the $\langle N_{ch}(p_T > 0.4\text{GeV}) \rangle$ as the X-axis of cumulant results.

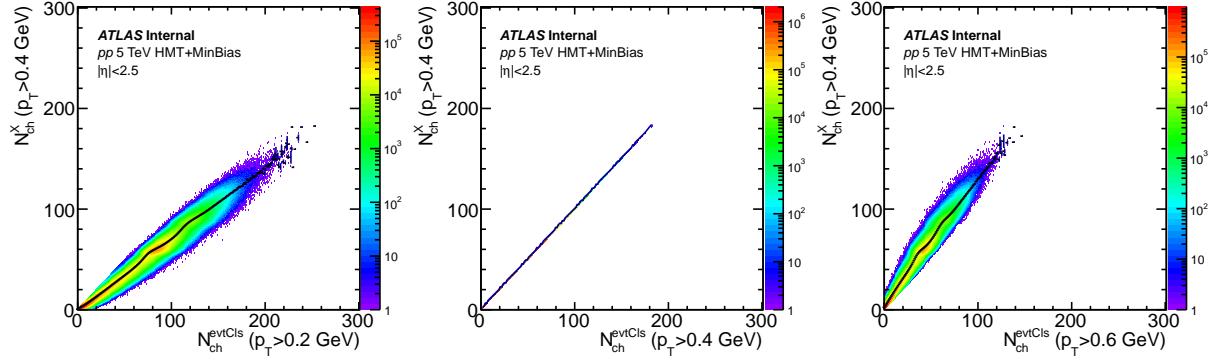


Figure 41: Correlation of N_{ch}^X and $N_{ch}^{evtClus}$, in 2015 5 TeV pp , and there are three $N_{ch}^{evtClus}$ with different p_T ranges: $p_T > 0.2$ GeV, $p_T > 0.4$ GeV and $p_T > 0.6$ GeV. As a convention, the N_{ch}^X is defined as number of particles with $p_T > 0.4$ GeV, which explains why in the second plot the correlation factor is 1. The black dots on the correlations are profiles, which provides the $\langle N_{ch}(p_T > 0.4\text{GeV}) \rangle$ as the X-axis of cumulant results.

Following last section, where we defined three event classes, based on N_{ch} with $p_T > 0.2$ GeV, $p_T > 0.4$ GeV and $p_T > 0.6$ GeV, we will determine the $\langle N_{ch}(p_T > 0.4\text{GeV}) \rangle$ in each event class. Fig. 39 shows the correlation between $N_{ch}(p_T > 0.2\text{GeV})$, $N_{ch}(p_T > 0.4\text{GeV})$ and $N_{ch}(p_T > 0.6\text{GeV})$. The correlation factor of the second plot is 1 by definition. The black dots on the correlations are profiles, which provides the $\langle N_{ch}(p_T > 0.4\text{GeV}) \rangle$ as the X-axis of the cumulant results.

4.8 Event weighting

At current stage the event-by-event 2- and 4-particle correlations have been calculated, and the event class has been determined. Now we will determine the mean value of 2- and 4-particle correlations in each event class. Since the original cumulant definition is based on looping through all the particles in all the events, in order to reflect this feature in the event-by-event calculation, proper event weighting needs to be applied during the averaging.

Since the cumulant is calculated by counting number of pairs, the unique combinations in 2- and 4-particle correlation provide the natural weight for event averaging:

$$\begin{aligned} W_{(2)} &\equiv M(M-1) \\ W_{(4)} &\equiv M(M-1)(M-2)(M-3) \end{aligned} \quad (17)$$

where M is the number of reference particles used in the Q -vector calculation. For the sub-event cumulant method, the weights can also be derived by counting the number of unique pairs:

$$\begin{aligned} W_{(2)a|b} &\equiv M_A M_B \\ W_{(4)a,a|b,b} &\equiv M_A(M_A-1)M_B(M_B-1) \end{aligned} \quad (18)$$

for the 2 sub-event method, and

$$\begin{aligned} W_{(2)a|a} &\equiv M_A(M_A-1) \\ W_{(2)b|c} &\equiv M_B M_C \\ W_{(2)a|b} &\equiv M_A M_B \\ W_{(2)a|c} &\equiv M_A M_C \\ W_{(4)a,a|b,c} &\equiv M_A(M_A-1)M_B M_C \end{aligned} \quad (19)$$

for the two kinds of 3 sub-event methods. The weight for other permutations will be similar.

In ALICE and CMS cumulant papers, instead of weighting events on the multi-particle correlation level, the events are weighted on the cumulant level instead, where 4-particle weights are used. It is obvious that the two different weighting methods are consistent when either the multiplicity is large, or the measured v_n is independent of multiplicity. To keep consistent with previous published results, we choose to weight the events on the cumulant level.

4.9 2- and 4-particle cumulant

The formulas for 2- and 4-particle cumulant are already listed in the method section. All the formulas are the same, independent of the efficiency correction and sub-event method. The flow signal can also be easily calculated once the cumulant is measured. In this analysis we will focus on the 2- and 4-particle $v_2\{2\}$ and $v_2\{4\}$ signals.

4.10 Validation of the subevent method using PYTHIA truth particles

Particle production in pp collisions is typically described by QCD-inspired models implemented in Monte Carlo (MC) event generators such as PYTHIA. PYTHIA model contains significant non-flow correlations from jets, dijets and resonance decays but no genuine long-range ridge correlations. In this analysis, 200 million pp collisions at $\sqrt{s} = 13\text{TeV}$ are generated with PYTHIA 8. Multi-particle cumulants based on standard method as well as subevent methods are calculated to quantify how they are biased by non-flow correlations as a function of charged particle multiplicity. Furthermore, flow signal is added to the generated event using a flow afterburner, and the performance for recovering the input flow signal is studied.

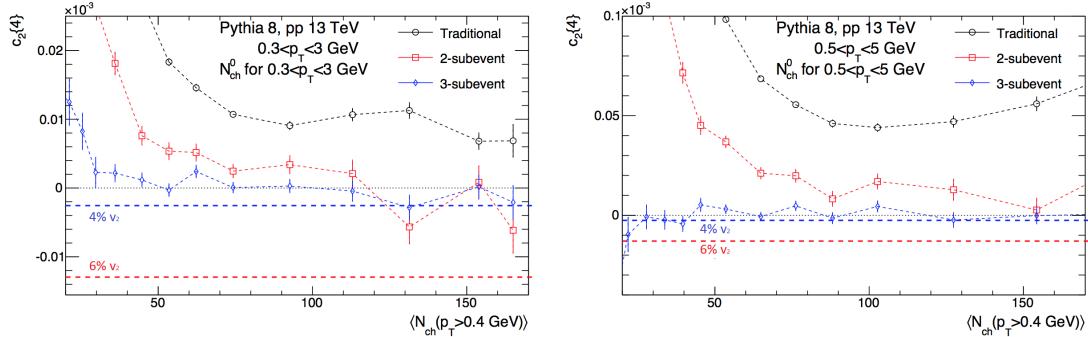


Figure 42: The $c_2\{4\}$ calculated for particles in $0.3 < p_T < 3.0 \text{ GeV}$ (left panel) or $0.5 < p_T < 5.0 \text{ GeV}$ (right panel) compared between the three cumulant methods. The event averaging is performed for N_{ch}^0 calculated for same p_T range, which is then mapped to $\langle N_{ch}(p_T > 0.4\text{GeV}) \rangle$, the average number of charged particles with $p_T > 0.4 \text{ GeV}$.

Fig. 42 shows a direct comparison between standard method and the two- and three-subevent methods for $c_2\{4\}$ in two p_T ranges. The three-subevent method has the best performance in suppressing the non-flow effects.

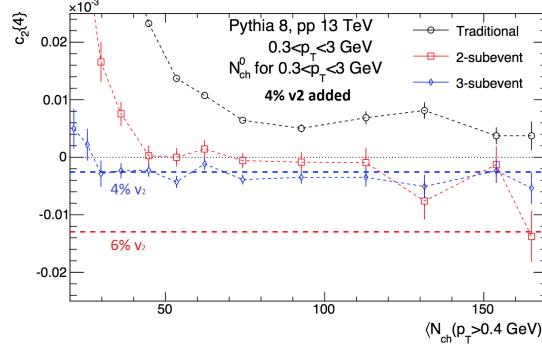


Figure 43: The $c_2\{4\}$ calculated for particles in $0.3 < p_T < 3.0$ GeV compared between the three cumulant methods with 4% v_2 imposed. The event averaging is performed for N_{ch}^0 calculated for same p_T range, which is then mapped to $\langle N_{ch}(p_T > 0.4 \text{ GeV}) \rangle$, the average number of charged particles with $p_T > 0.4$ GeV.

839 To quantify the performance of the three methods for recovering the underlying flow signal, a flow
 840 afterburner is used to add a constant v_2 to the generated PYTHIA events. Fig. 43 shows the calculated
 841 $c_2\{4\}$ with 4% imposed to the generated events. In the case 4% input flow, only the three-subevent
 842 method can recover the input.

843 5 Systematic uncertainties and cross-checks

844 5.1 pp 13 TeV

845 In this section, we will go through all the systematics and cross-checks for 13 TeV pp . Each check has
 846 been done separately in two p_T ranges that used to calculate the cumulants:

- 847 • Reference particles with $0.3 < p_T < 3.0$ GeV;

- 848 • Reference particles with $0.5 < p_T < 5.0$ GeV;

849 Relative errors δ_{sys} are calculated in two situations:

- 850 • If one check is compared with the default: $\delta_{sys} \equiv \frac{C_{check} - C_{default}}{C_{default}}$;

- 851 • If data sample is divided into two sub-samples: $\delta_{sys} \equiv \frac{C_{check1} - C_{check2}}{C_{check1} + C_{check2}}$,

852 where $C_{default}$ is the $C_2\{4\}$ from default setup, while the rest are $C_2\{4\}$ from other checks. In this
 853 analysis, without special mentioning, the default setup is: events with > 50% trigger efficiency, track-
 854 ing efficiency corrected, standard track selection, with pile-up events cleaned and with detector effects
 855 removed.

856 Since $C_2\{4\}$ is a very small quantity, our quoted relative systematic uncertainties can sometimes be
 857 very large. This is especially true when $C_2\{4\}$ goes across 0. Due to this reason, we might quote absolute
 858 errors in the end. Another issue is that since cumulant measurement is usually dominated by statistical
 859 errors instead of systematics, if the statistical errors are much larger than systematics, we will re-bin the
 860 N_{ch} to increase statistical significance.

861 When combining different systematic sources, we assume that different systematic checks are uncor-
 862 related, which is true for most of the systematic uncertainties we included in the end.

863 5.1.1 Trigger efficiency

864 To extent the cumulant measurement to high multiplicity region, HMT triggers are applied in this analy-
 865 sis. The trigger efficiency is evaluated for all the major HMT triggers. When trigger efficiency reaches
 866 1, that means there is no selection bias due to this trigger. However, when trigger efficiency is below
 867 1, for example in the turn-on region, there could possibly exist selection bias due to this trigger: HMT
 868 trigger could select events with higher (or lower) p_T , with higher fraction of non-flow contribution and
 869 so on. In the ideal case, there won't be any bias if one only uses triggered events with efficiency equals 1.
 870 However, in that case, since the N_{ch} distribution is decreasing very fast in pp and $p+Pb$, a large fraction
 871 of events will be rejected. Under this circumstance, we will still include the triggered events even though
 872 the efficiency is below 1 and we will evaluate the impact on the results due to this "loose" selection
 873 criteria.

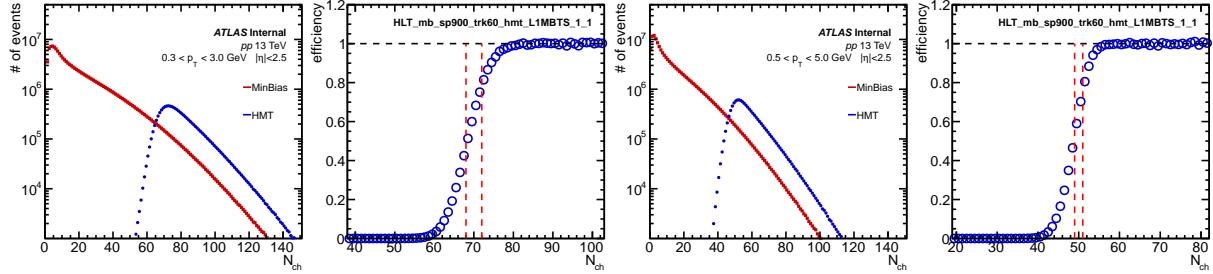


Figure 44: Trigger efficiencies of one major HMT triggers as a function of number of tracks in two p_T ranges: $0.3 < p_T < 3.0$ GeV and $0.5 < p_T < 5.0$ GeV, from 13 TeV pp run period 1. Efficiency is calculated relative to the corresponding MinBias trigger in this run period then scaled to 1.0 in the large N_{ch} region. The two red dash lines indicate 50% and 80% efficiency cuts.

Fig. 44 shows an example of N_{ch} distribution and efficiency of HMT trigger HLT_mb_sp900_trk60, in lower and higher p_T ranges separately. The two red dash lines indicate 50% and 80% efficiency cuts. The default trigger efficiency cut is 50% and 80% cut is used as a cross-check. A summary of all the efficiency cuts, together with collected statistics, is shown in table 45.

pp 13 TeV period 1									
Trigger Name	Total Events / 1000	50% efficiency cut (0.3< p_T <3.0)	Events after previous cut	80% efficiency cut (0.3< p_T <3.0)	Events after previous cut	50% efficiency cut (0.5< p_T <5.0)	Events after previous cut	80% efficiency cut (0.5< p_T <5.0)	Events after previous cut
[0] HLT_noalg_mb_L1MBTS_1	130275	0	130275	0	130275	0	130275	0	130275
[11] HLT_mb_sp400_trk40_hmt_L1MBTS_1_1	25387	46	23165	50	19478	32	22776	34	20361
[12] HLT_mb_sp700_trk50_hmt_L1MBTS_1_1	20498	57	18235	61	15128	40	18247	43	15209
[13] HLT_mb_sp900_trk60_hmt_L1MBTS_1_1	10782	68	9303	72	7600	49	9095	51	7985
[15] HLT_mb_sp1400_trk90_hmt_L1TE10	3950	104	2471	120	704	78	2436	98	300

pp 13 TeV period 2 (299390, 300287)									
Trigger Name	Total Events / 1000	50% efficiency cut (0.3< p_T <3.0)	Events after previous cut	80% efficiency cut (0.3< p_T <3.0)	Events after previous cut	50% efficiency cut (0.5< p_T <5.0)	Events after previous cut	80% efficiency cut (0.5< p_T <5.0)	Events after previous cut
[0] HLT_noalg_mb_L1MBTS_1_1	2598	0	2598	0	2598	0	2598	0	2598
[17] HLT_mb_sp1800_hmtpf_L1TE5	1832	67	922	79	521	48	894	58	482
[18] HLT_mb_sp1500_hmtpf_L1TE10	567	67	280	84	117	50	259	64	99
[19] HLT_mb_sp900_trk60_hmt_L1MBTS_1_1	1847	68	1638	73	1330	49	1597	52	1350
[22] HLT_mb_sp1000_trk70_hmt_L1TE5	2771	80	2292	84	1898	58	2323	62	1803
[24] HLT_mb_sp1400_trk90_hmt_L1TE10	1851	102	1389	108	984	76	1413	82	889

pp 13 TeV period 3 (305359)									
Trigger Name	Total Events / 1000	50% efficiency cut (0.3< p_T <3.0)	Events after previous cut	80% efficiency cut (0.3< p_T <3.0)	Events after previous cut	50% efficiency cut (0.5< p_T <5.0)	Events after previous cut	80% efficiency cut (0.5< p_T <5.0)	Events after previous cut
[0] HLT_mb_mbts_L1MBTS_1_1	988	0	988	0	988	0	988	0	988
[2] HLT_noalg_mb_L1TE30	4139	108	176	122	53	78	268	93	58
[6] HLT_mb_sp700_hmtpf_L1TE5	337	35	240	49	121	24	240	33	137
[7] HLT_mb_sp1500_hmtpf_L1TE10	388	63	225	72	133	46	213	54	118
[9] HLT_mb_sp1000_trk70_hmt_L1TE5	1586	78	1303	82	1041	57	1342	61	1014
[11] HLT_mb_sp1400_trk90_hmt_L1TE10	2630	100	1927	106	1259	76	1856	83	928

pp 13 TeV period 4 (309314, 309346, 310216)									
Trigger Name	Total Events / 1000	50% efficiency cut (0.3< p_T <3.0)	Events after previous cut	80% efficiency cut (0.3< p_T <3.0)	Events after previous cut	50% efficiency cut (0.5< p_T <5.0)	Events after previous cut	80% efficiency cut (0.5< p_T <5.0)	Events after previous cut
[0] HLT_mb_spTRK	877	0	877	0	877	0	877	0	0
[18] HLT_mb_sp900_trk50_hmt_L1TE5	5326	58	4520	62	3696	40	4711	43	3946
[24] HLT_mb_sp1000_trk70_hmt_L1TE10.0ETA24	9174	82	6639	89	4380	59	7256	67	3908
[30] HLT_mb_sp1400_trk90_hmt_L1TE20.0ETA24	13615	100	10422	108	6757	82	6785	91	3054

Figure 45: Table of N_{ch} cuts for 50% and 80% trigger efficiencies. The number of events for each trigger after N_{ch} cuts are also listed.

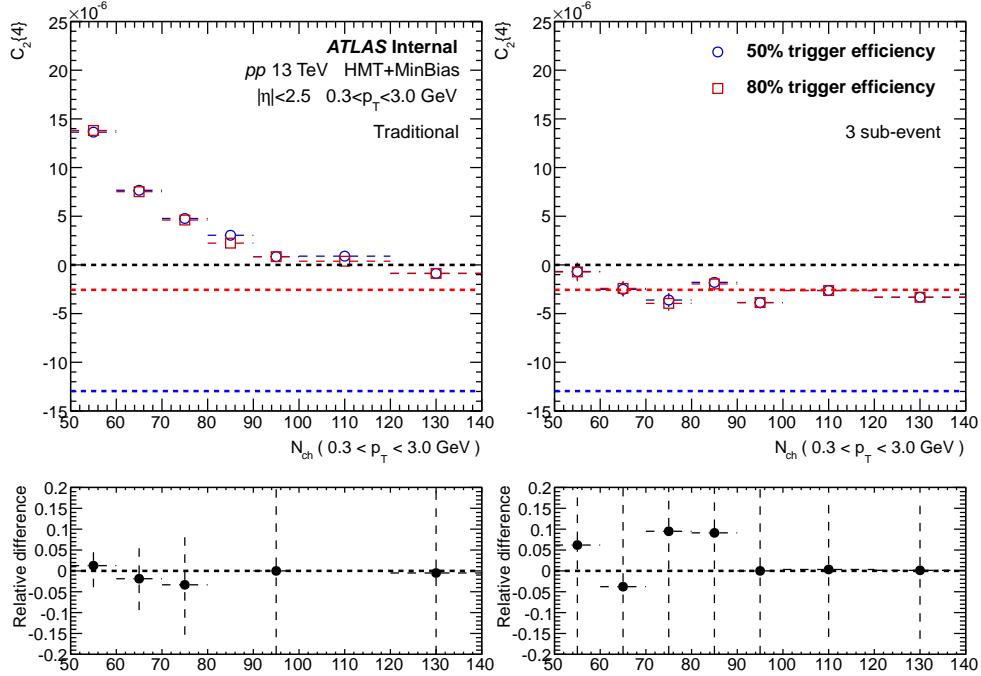


Figure 46: $C_2\{4\}$ measured with 50% and 80% trigger efficiency cuts. The left column has the results from traditional method and the right column from 3 sub-event method.

Fig. 46 shows the $C_2\{4\}$ calculated with 50% and 80% trigger efficiency cuts. For the traditional method, the relative difference is on the level of 5%, and it goes to larger than 20% when $C_2\{4\}$ goes across 0. As a comparison, the relative difference from 3 sub-event method is less than 10% in the region $N_{ch} > 50$, and the difference is probably caused by the statistical fluctuation, since 3 sub-event method always has larger statistical error bars. Even with large statistical errors, the relative differences are still quoted as systematic uncertainties.

5.1.2 Tracking efficiency

During the offline track reconstruction, not all the tracks will be reconstructed. The fraction of successfully reconstructed tracks are estimated through Monte-Carlo generators with a simulation of ATLAS detector, known as tracking efficiency ϵ . In previous section, the tracking efficiency is evaluated as a function of η , p_T and N_{ch} , and while calculating the multi-particle correlations, each particle is weighted by $1/\epsilon$. Since flow is a global quantity and its signal has very weak dependence of η , the η dependence of tracking efficiency should not have a large influence on the results. Furthermore, since the $C_2\{4\}$ is measured as a function of N_{ch} , this means the only factor that could contribute to the results is the p_T weighting in the tracking efficiency. In other words, if tracking efficiency is independent of p_T , with and without tracking efficiency weight should not cause any difference.

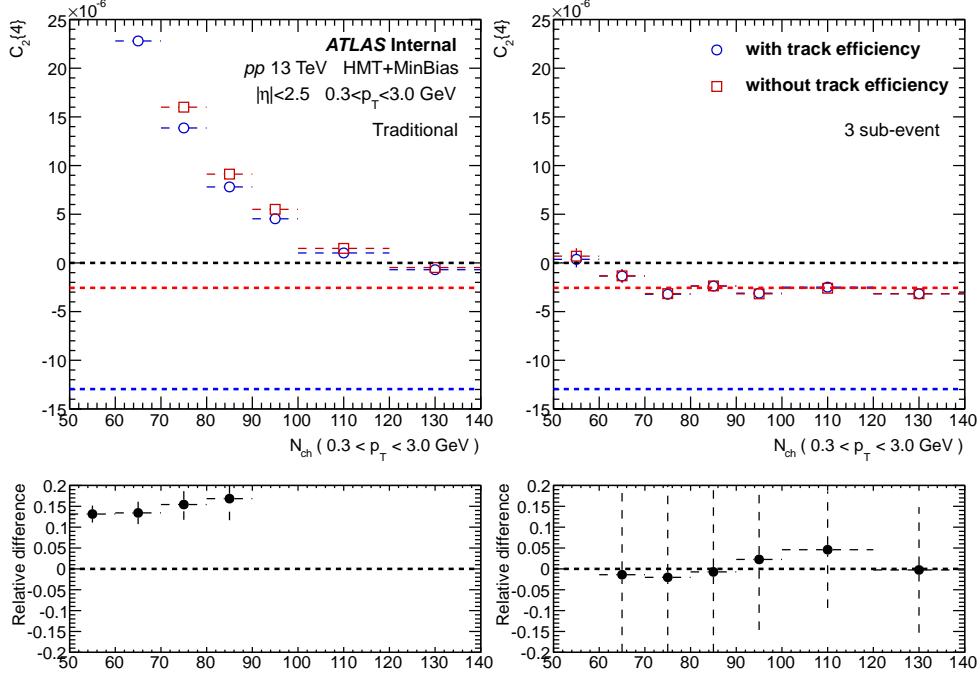


Figure 47: $C_2\{4\}$ measured with and without tracking efficiency correction. The left column has the results from traditional method and the right column from 3 sub-event method.

Fig.47 shows the $C_2\{4\}$ calculated with and without tracking efficiency correction. For the traditional method, the relative difference is on the level of 10% larger without efficiency correction and keeps increasing as N_{ch} increases. Meanwhile, the relative difference from 3 sub-event method is less than 5% in the region $N_{ch} > 50\%$. As discussed before, what tracking efficiency does is to put different weights to particles with different p_T . From many previous flow measurements, it is known that v_2 usually increases with p_T until p_T reaches 2-3 GeV, and then decrease. p_T dependence of v_2 is relatively small in the region where tracking efficiency dramatically changes. This means that different p_T weighting should have minimal impact on the flow measurement, as been shown in previous $p+Pb$ and $Pb + Pb$ results, where with and without tracking efficiency correction does not make a big difference. This explains why with 3 sub-event method, the relative difference is minimal. However, for the traditional method, it is dominated by the non-flow contribution, which might be more sensitive to p_T weighting than flow. This could be the reason why traditional method is dependent of tracking efficiency weighting. But in any case, since tracking efficiency correction will introduce additional p_T weighting to the results, this check will not be quoted as one of the systematics.

In order to correctly estimate impact from uncertainty of tracking efficiency, the systematic uncertainty on the tracking efficiency in different $|\eta|$ ranges are applied, as shown in Fig. 48. These values are obtained from the 13 TeV multiplicity analysis and include material uncertainties. The entire analysis is re-done by varying the tracking efficiency within its systematic uncertainty. Two extreme cases are chosen:

- High tracking efficiency: the tracking efficiency is increased by its uncertainty for all p_T and η ;
- Low tracking efficiency: the tracking efficiency is decreased by its uncertainty for all p_T and η .

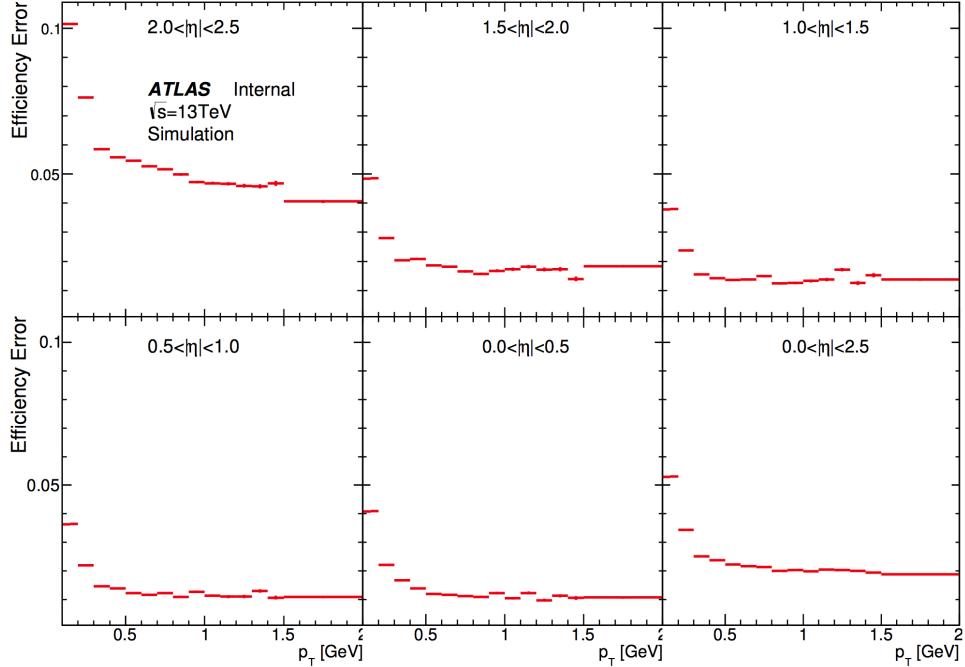


Figure 48: The systematic uncertainties of the tracking efficiency plotted as a function of p_T for several $|\eta|$ slices. These include the material uncertainties. These were obtained from the 13TeV multiplicity analysis.

With above variations, the 4-particle cumulant from all three methods are re-evaluated and the results is shown in the summary of systematics.

5.1.3 Track selection

Another check is concerning the track selection. The default track selection is listed as follows:

- Tracks are from primary vertex
- If IBL hit is expected: at least 1 IBL hit
- If on IBL hit is expected: a Layer-0 hit if expected
- At least 1 pixel hit + dead sensors
- $p_T < 300$ MeV: at least 2 SCT hits + dead sensors
- $p_T < 400$ MeV: at least 4 SCT hits + dead sensors
- $p_T > 400$ MeV: at least 6 SCT hits + dead sensors
- If $p_T > 10$ GeV: χ^2 probability < 0.01
- $|d_0| < 1.5$
- $|z_0 - v_z| \sin\theta < 1.5$
- $|\eta| \leq 2.5$

- 930 • $p_T \geq 200$ MeV

931 and as a check, the track pointing cut is tightened

- 932 • $|d_0| < 1.0$

- 933 • $|z_0 - v_z| \sin\theta < 1.0$

934 while keeping all the rest selection cuts the same. This check will test the stability of results due to the
935 track selection cuts.

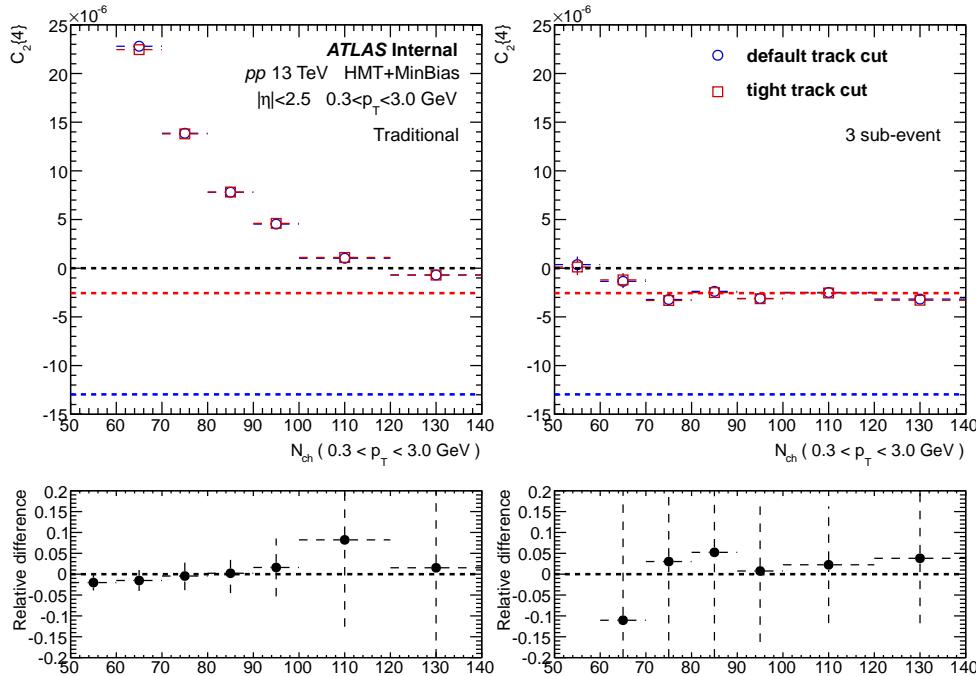


Figure 49: $C_2\{4\}$ measured with default and tighter track selection cuts. The left column has the results from traditional method and the right column from 3 sub-event method.

936 Fig.49 shows the $C_2\{4\}$ calculated with default and tighter track selection cuts. It is obvious that
937 both traditional and 3 event methods have minimal dependence of track selection. This is not hard
938 to interpret since the tighter cuts only reject less than < 1% of the tracks, and such loss is compensated
939 by the tracking efficiency. This check will be quoted as one of the systematics.

940 5.1.4 Pile-up condition

941 In this analysis all the tracks used to calculate cumulants are from the primary vertex. In principle, in
942 pile-up events, tracks from pile-up vertex should not contribute to the measurement. However, during
943 the track and vertex reconstruction, when a pile-up vertex is too close to the primary vertex, two vertices
944 might be merged. Since the particles from two different vertices are totally uncorrelated, including these
945 events with pile-up vertex will reduce the signal of flow signal.

946 13 TeV pp runs have 4 run periods, and the μ values varies from 0.003 to 0.6. In order to estimate
947 how pile-up will affect the results, the whole data sample is divided into 2 sub-sample:

- 948 • Runs with peak $\mu < 0.4$;

- 949 • Runs with peak $\mu > 0.4$;

and the μ value across all the runs are summarized in table. 50. The threshold $\mu = 0.4$ is chosen so that the statistics in two run groups are roughly the same. In the table, runs tagged with green are the runs with relatively low- μ and runs tagged with red are with relatively high- μ .

Run Number	267358	267359	267360	267367	267385	267599	277025	277081	299390	300287	305359	309314	309346	310216
Events	8.6	12.3	12.6	17.1	71.2	105.6	19.9	44.5	4.1	8.1	16.5	13.2	7.5	79.9
Peak mu	0.003	0.006	0.37	0.46	0.47	0.49	0.61	0.62	0.52	0.52	0.06	0.36	0.33	0.31

Figure 50: Table showing the μ value and statistics in all the runs from 13 TeV pp . Runs tagged with green are the runs with relatively low- μ and runs tagged with red are with relatively high- μ .

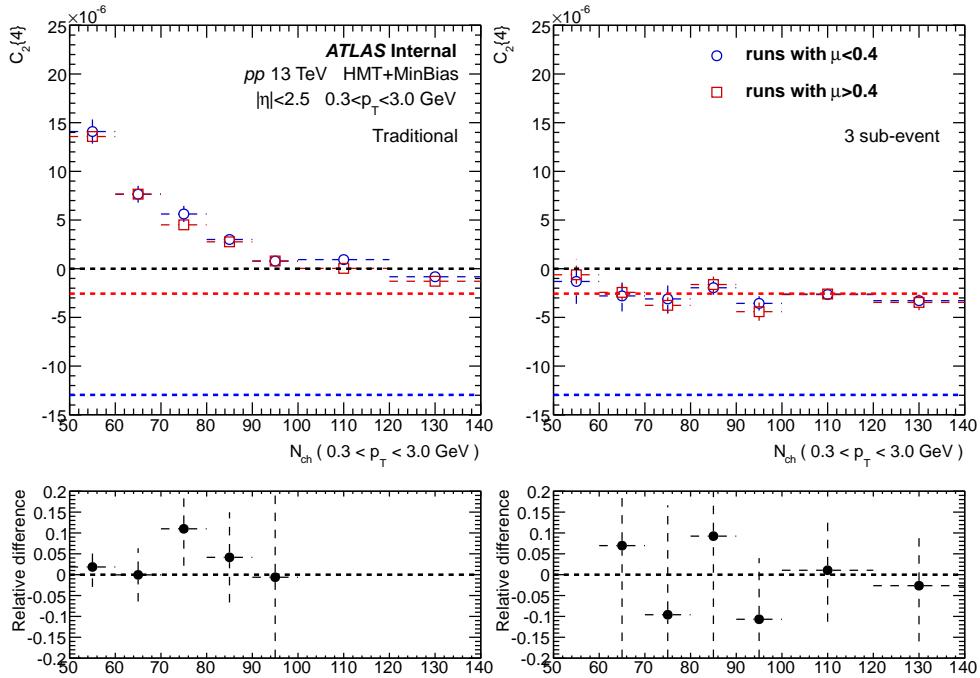


Figure 51: $C_2\{4\}$ measured in two run groups with different μ values. The left column has the results from traditional method and the right column from 3 sub-event method.

Since in this check the whole data set is divided into two run groups, which means the statistical fluctuation is larger than previous check, which is reflected in the results. In both traditional and 3 sub-event cumulant results, relative differences are fluctuating around 0, meaning the differences are dominated by the statistical fluctuation. Under this circumstance, we will re-bin the X-axis to increase the statistical significance. The merged results will be shown in the summary of systematics.

5.1.5 Residual detector effects

Tracking efficiency weighting corrects the possible detector effects as a function of η and p_T , but the residual detector effects could still remain in the ϕ direction. In pp collision, since the "event plane" angle differs from one event to another, the ϕ distribution over many events should be flat and the discrepancy upon that is caused by the residual detector effects.

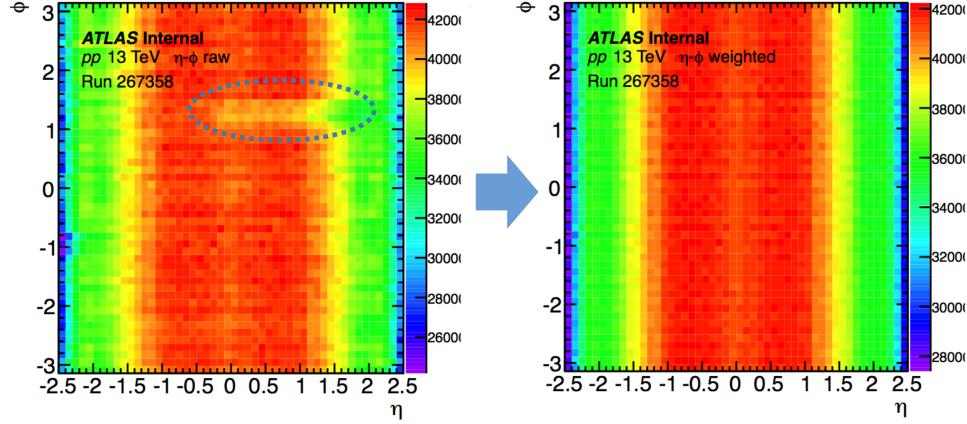


Figure 52: An example of $\eta - \phi$ distribution of all the events in one run. Left plot is the raw distribution, and right plot is after the correction.

In order to show this, Fig. 52 presents one example of $\eta - \phi$ distribution of all the events in one run. Left plot is the raw distribution, where some discrepancy has been observed (circled in blue). In order to correct this detector effect, an additional weighting has been applied to the particles while calculating the multi-particle correlation:

$$w_\phi(\eta, \phi) \equiv \frac{\langle N(\delta\eta) \rangle}{N(\delta\eta, \delta\phi)} \quad (20)$$

and after the correlation, by construction, the $\eta - \phi$ map becomes flat in ϕ . This procedure has been used in many other flow analyses and known as "flattening". Since the detector condition could vary from run to run, w_ϕ is calculated run-by-run and the summary is shown in the data analysis section.

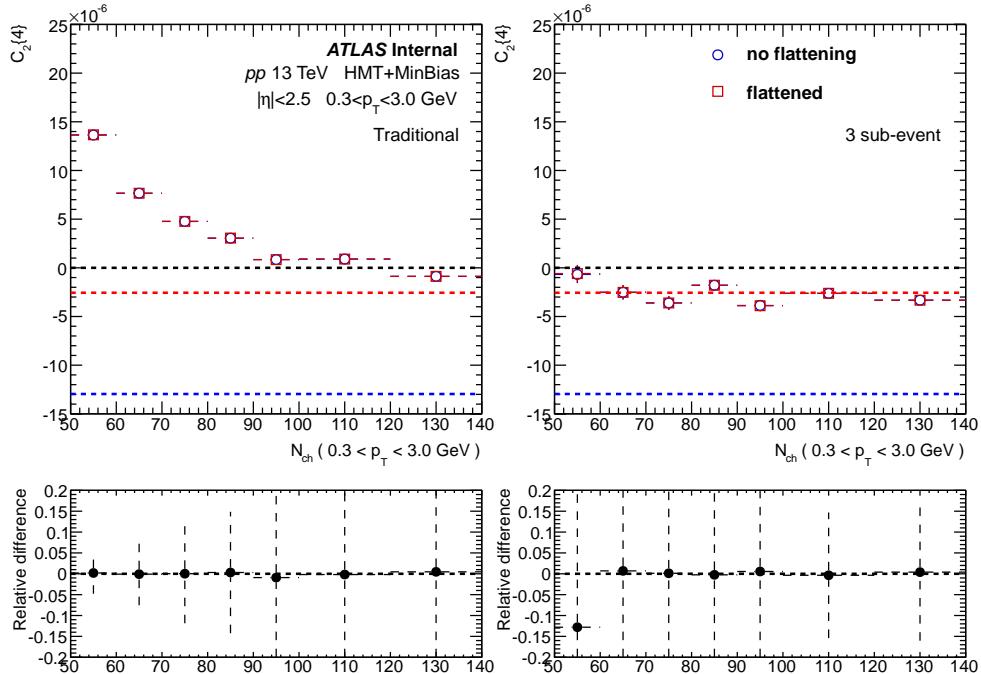


Figure 53: $C_2\{4\}$ measured with and without correcting the residual detector effects. The left column has the results from traditional method and the right column from 3 sub-event method.

Fig. 53 compares the $C_2\{4\}$ with and without flattening procedure. For both methods, the relative differences are within 1%, meaning such residual detector effects have minimal influence on the results. This check will be quoted as one of the systematics.

5.1.6 Event class bin width

As has been discussed before, traditional cumulant is sensitive to the multiplicity fluctuation, which is originated from the observation that non-flow fluctuation is changed when combining different events into groups. For this reason, while calculating the cumulant, the events are always binned with 1 N_{ch} bin width, and the particles used for event class definition are the same as particles used in cumulant calculation. In the results section, we will discuss the influence of using particles with different p_T for event class definition. Here we will discuss another factor: event class bin width.

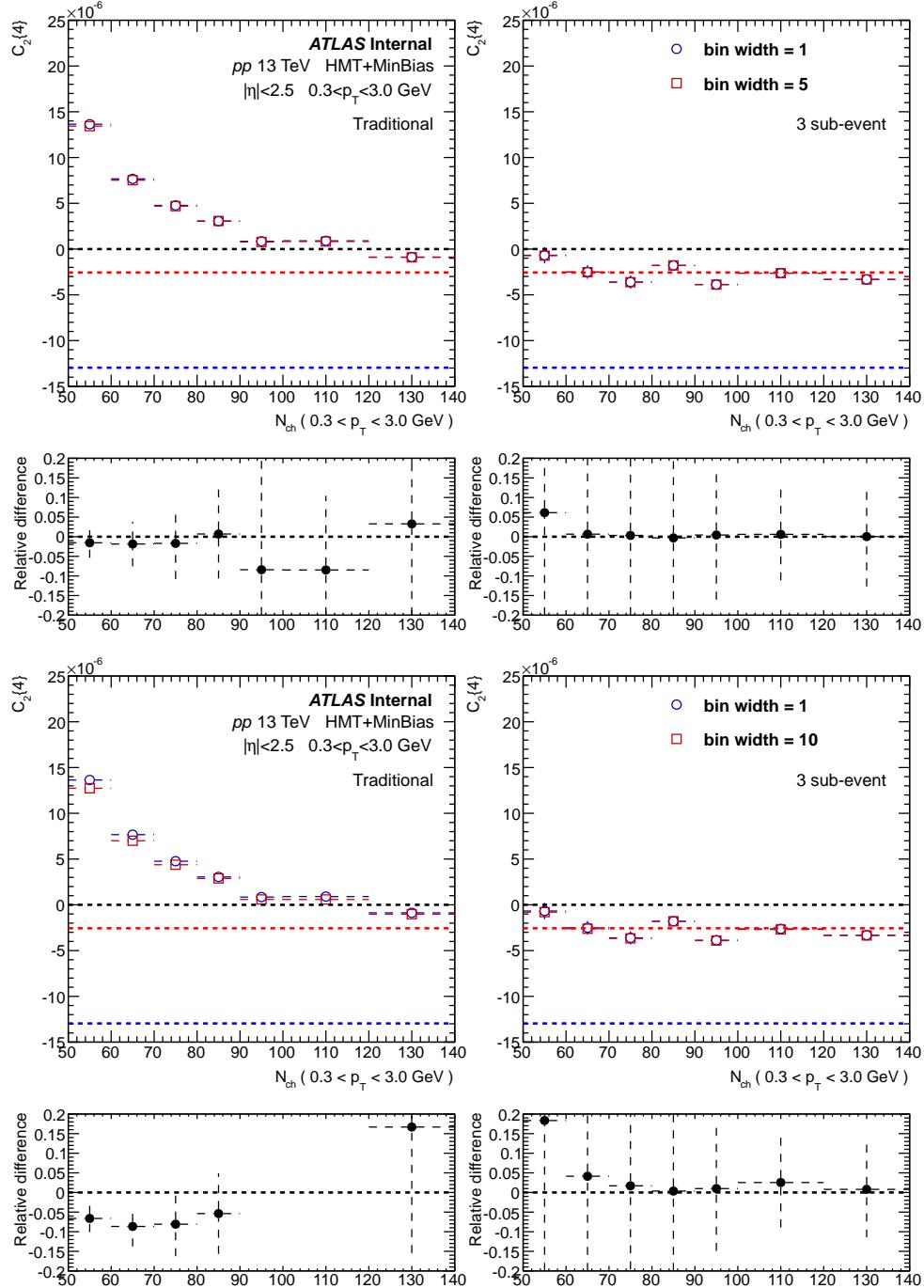


Figure 54: $C_2\{4\}$ measured with different event class bin widths. The left column has the results from traditional method and the right column from 3 sub-event method.

If the bin width is increased from default 1 to 5 or even 10, the non-flow fluctuation in each event class could be different. Fig. 54 compares the $C_2\{4\}$ with different bin widths. When event class bin width is increase to 5 tracks, $C_2\{4\}$ from traditional method slightly changes, within 5%. As a comparison, there is little change for $C_2\{4\}$ from the 3 sub-event method, the relative difference is minimal. This is even the case when the event class bin width is increased to 10 tracks: $C_2\{4\}$ from 3 sub-event method still has a relative difference within 5%, meanwhile the difference in traditional method is beyond 5%. This

observation is exactly consistent with our expectation, since the 3 sub-event method further suppress the non-flow contribution and the additional multiplicity fluctuation due to the bin width changes does not cause any difference. We will come back to this point when we discuss the different event class definition based on p_T of the particles. This check is only a cross-check to show the advantage of sub-event cumulant and will not be quoted as systematics.

5.1.7 η gap for sub-event

Compared with 2 sub-event method, 3 sub-event method is designed mainly to remove the dijet-like correlation, since the chance that dijet still contribute to all the 3 sub-events are dramatically lowered compared with traditional cumulant. The only situation where dijet correlation still contributes to the 3 sub-event cumulant is when two jets from the dijet rest on the two boundaries of the 3 sub-events. In order to show whether the $C_2\{4\}$ is due to such little remaining dijet contributions, we introduced η gaps between the sub-events to further suppress this contribution.

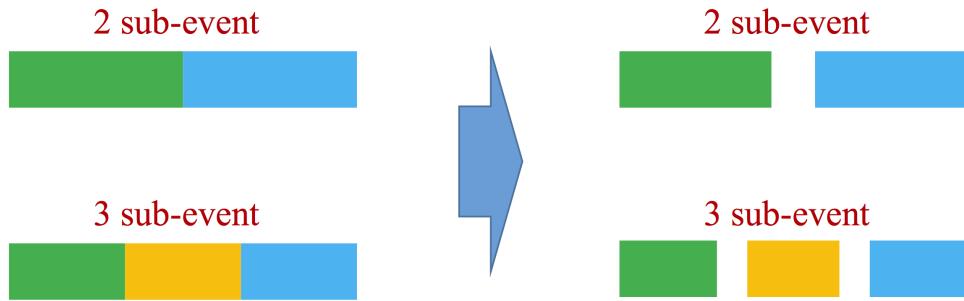


Figure 55: A cartoon showing the how the η gaps are introduced to the 2 and 3 sub-event. The η gap is $\Delta\eta = 0.5$.

Fig. 55 describes the basic idea of the η gap between the sub-events: in the 2 sub-event case, we introduced a η gap with $\Delta\eta = 0.5$ between 2 sub-events and in the 3 sub-event case, two η gaps with $\Delta\eta = 0.5$ have been added among three sub-events. The advantage of add η gaps is removing the residual dijet correlation and the drawback is reducing the number of particle pairs, especially in 4-particle cumulant. Thus the purpose of this check is estimating the contributing of remaining non-flow, the default method is still without η gap, to gain better statistical significance.

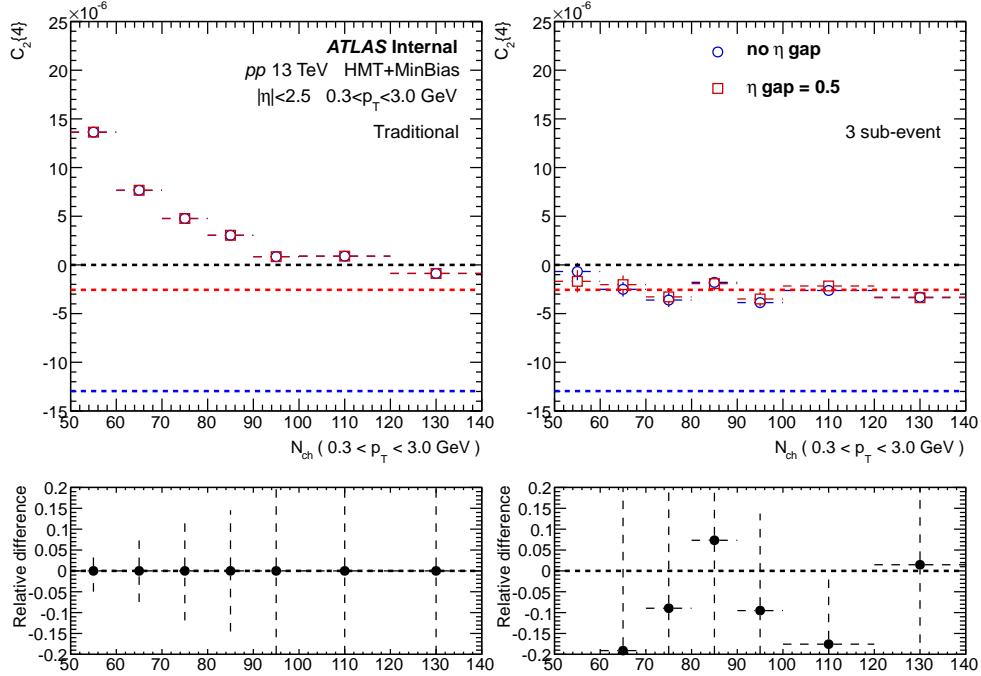


Figure 56: $C_2\{4\}$ measured with and without η gaps between sub-events. The left column has the results from traditional method and the right column from 3 sub-event method.

Fig. 56 shows the $C_2\{4\}$ measured with and without η gaps between sub-events. Since there is no η gap introduced for traditional cumulant, the relative difference is 0 by construction. For 3 sub-event method, with η gap does not change the observation: even though the relative difference is fluctuating around 0 on the level of more than 10%, but that is partially due to the larger statistical errors with presence of η gap. Due to these reasons, the default case is without η gap and this check will not be quoted as systematics.

5.1.8 Charge dependence

Since most short-range non-flow correlations are originated from decay and jets, they are supposed to be stronger in the opposite charge combination than same charge. Subevent method provides a natural way to test this charge dependence.

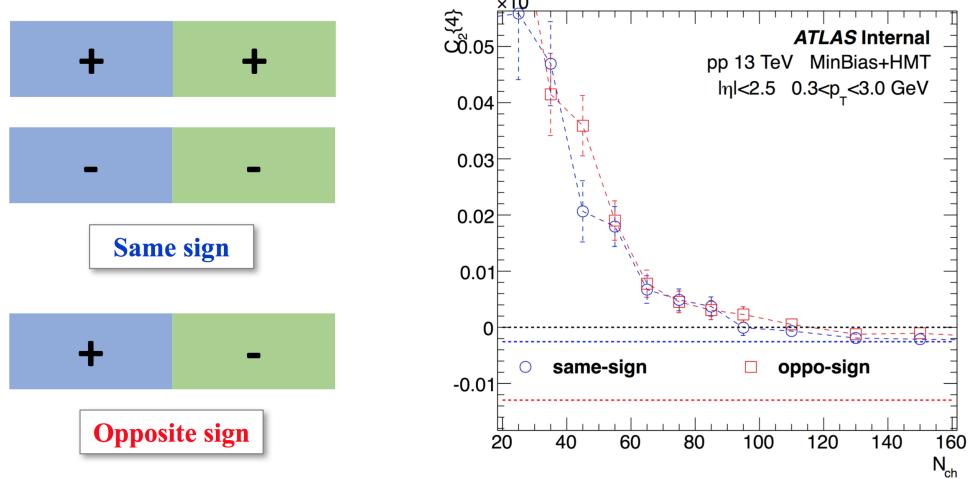


Figure 57: Left plot shows the configuration for same and opposite change in 2 subevent cumulant. Right plot show the comparison of $c_2\{4\}$ measured using only same change and opposite charge. The blue and red dash line indicate the 4% and 6% flow respectively.

1014 The configurations of same-sign and opposite-sign are shown in the left plot in Fig. 57, where same-
 1015 sign combination has both ++ and -- charges in the 2 subevents, while one + and one - charges in
 1016 2 subevents for the opposite-sign. The comparison of $c_2\{4\}$ measured in these two configurations are
 1017 shown in the right plot. $c_2\{4\}$ from opposite sign is slightly larger than the same sign, which is expected
 1018 since the majority of short-range non-flow are suppressed in the 2 subevent cumulant.

1019 5.1.9 Deviations from the truth in the reconstruction of Monte Carlo data

1020 The PYTHIA 8 Monte Carlo simulations were used to evaluate the difference between 4-particle cumu-
 1021 lants in pp data calculated using the generated and reconstructed charged particles obtained using the
 1022 same analysis method. In some analysis it is considered as a crosscheck, since it assesses the quality of
 1023 tracking, which are separately accounted for in previous systematics. The argument for not accounting
 1024 it as a systematic uncertainties also relies on the fact that MC generators do not properly describe the
 1025 investigated particle correlations. Nevertheless, we think that it still be considered as a validation of the
 1026 analysis method, even though PYTHIA does not include any flow effect. This is only a cross-check, and
 1027 will not be included as the systematics.

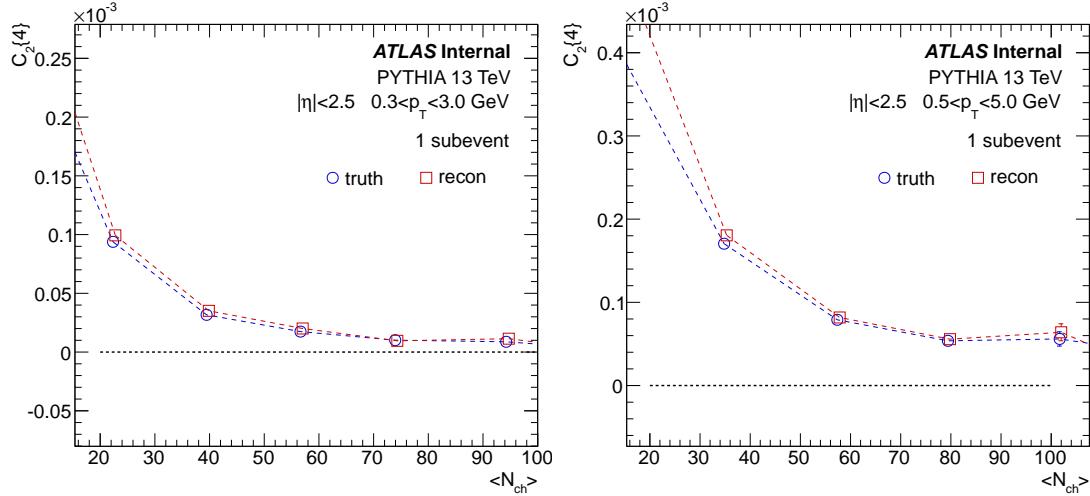


Figure 58: Comparison of the 4-particle cumulants $c_2\{4\}$ for 13 TeV PYTHIA, calculated with the reconstructed and generated charged particle for $0.3 < p_T < 3.0$ GeV (left) and $0.5 < p_T < 5.0$ GeV (right). The cumulant is calculated using the standard method (1 subevent).

Fig. 58 illustrates the evaluation of the MC closure test for pp , for $0.3 < p_T < 3.0$ GeV (left) and $0.5 < p_T < 5.0$ GeV (right). For the standard cumulant, $c_2\{4\}$ from reconstructed tracks is slightly larger than the $c_2\{4\}$ from truth particles. The small difference could be due to several reasons: 1) tracking efficiency is slightly dependent of N_{ch} , which has not been taken into account when applying the efficiency; 2) even though events are binned based on the corrected number of charged particles, in each event class, the events from truth and reconstructed are not identical due to event-by-event fluctuations, which might have an impact on the 4-particle cumulant $c_2\{4\}$; 3) tracking efficiency does not take non-uniformness in ϕ into account, which could also result in small differences.

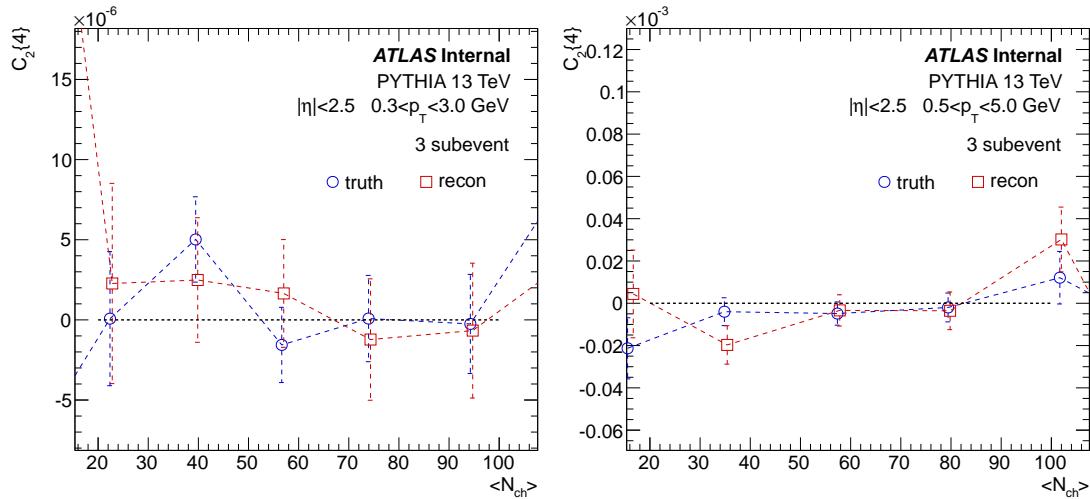


Figure 59: Comparison of the 4-particle cumulants $c_2\{4\}$ for 13 TeV PYTHIA, calculated with the reconstructed and generated charged particle for $0.3 < p_T < 3.0$ GeV (left) and $0.5 < p_T < 5.0$ GeV (right). The cumulant is calculated using the 3 subevent method.

As a comparison, Fig. 59 shows the same check but calculated with 3 subevent method. With current statistics, the difference of $c_2\{4\}$ between truth and reconstructed seems to be smaller than standard

method, and both of them fluctuates around 0. Since PYTHIA does not include any flow effect, this plot supports that 3 subevent method can effectively suppress the non-flow: $c_2\{4\}$ is much closer to 0 compared with standard method. Furthermore, the smaller difference between truth and reconstructed could also indicate that the second reason listed above could be the reason why in standard cumulant $c_2\{4\}$ from reconstructed tracks is slightly higher than the $c_2\{4\}$ from truth particles: additional event-by-event non-flow fluctuation can cause the difference. While with 3 subevent method, since the non-flow is suppressed, the difference is much smaller.

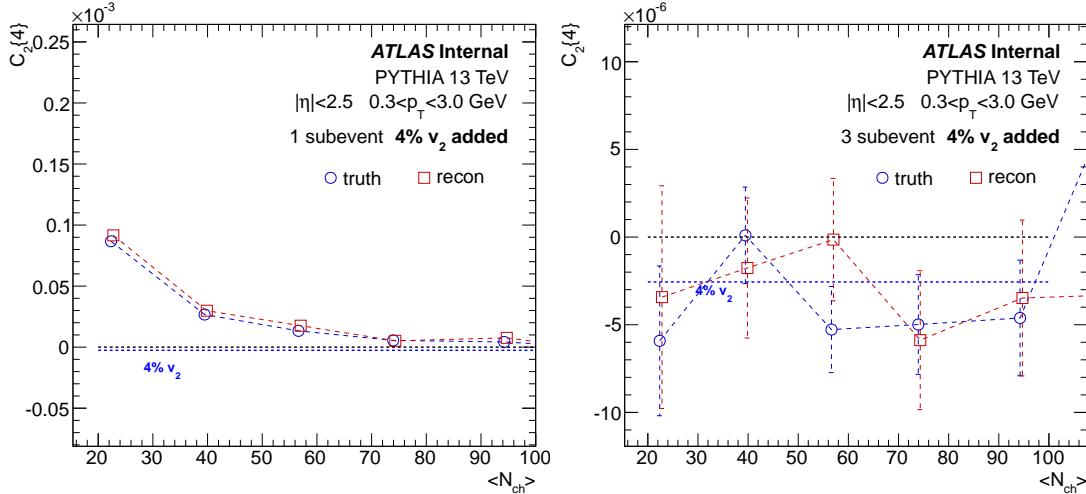


Figure 60: Comparison of the 4-particle cumulants $c_2\{4\}$ for 13 TeV PYTHIA, calculated with the reconstructed and generated charged particle for $0.3 < p_T < 3.0$ GeV, with 4% v_2 imposed to the generated and reconstructed events (indicated by the blue dash line). The left plot is calculated using standard method and the right plot using 3 subevent method.

To quantify the performance of the three methods for recovering the underlying flow signal, a flow afterburner is used to add a constant v_2 to the generated and reconstructed PYTHIA events. Fig. ?? shows the calculated $c_2\{4\}$ with 4% imposed to the generated and reconstructed events. In the case 4% input flow, standard cumulant still give $c_2\{4\}$ with positive sign, while $c_2\{4\}$ from the three-subevent method fluctuates around the 4% imposed signal.

5.1.10 Summary of systematics

In this section, the major systematics sources in 13 TeV pp are summarized. Since the 4-particle cumulant $C_n\{4\}$ is a very small quantity and it usually changes sign in the low multiplicity region, we choose absolute uncertainties instead of relative uncertainty to show the systematics. In this way, the trend of systematics will remain stable and will not go to much larger value when $C_n\{4\}$ goes across 0.

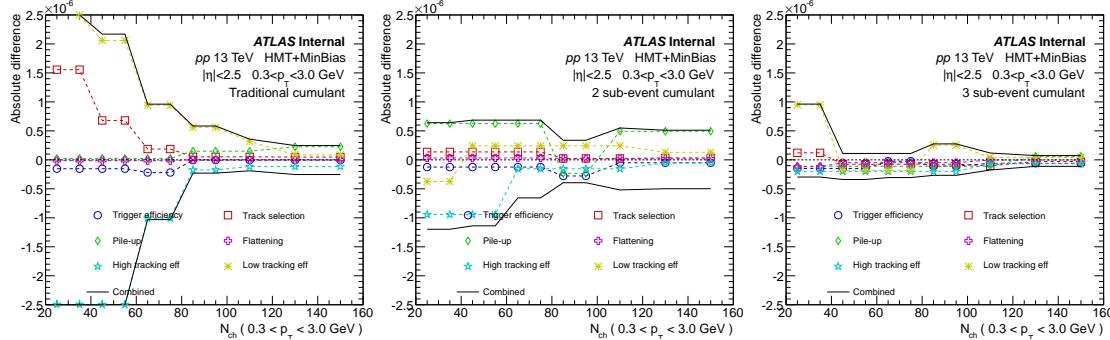


Figure 61: Summary of systematics in 13 TeV pp , for three different methods, with $0.3 < p_T < 3.0$ GeV. The statistical errors are not shown for better readability. The total systematics combined systematics from trigger efficiency, track selection, pile-up effects, flattening procedure and tracking efficiency.

Fig. 61 summarizes five major systematic sources with the traditional cumulant method. The Y-axis shows the absolute difference for each check. The dominating uncertainty comes from when lowering the tracking efficiency by its uncertainty. Since the tracking efficiency adds different weights to particles with different p_T , and the non-flow sources usually contains high p_T particles, lowering the tracking efficiency will equivalently adding more weights to the particles with higher p_T , causing a larger non-flow contribution. Since there is significant residue non-flow with the traditional method, this is the reason why varying tracking efficiency has the largest uncertainty. The next leading uncertainty comes from track selection, and it becomes larger towards the low multiplicity region, where the fraction of non-flow contribution is larger. For other sources, the absolute differences are relatively smaller: within 0.2×10^{-6} .

As a comparison, fig. 61 summarizes five major systematic sources with the 2 sub-event cumulant method. The Y-axis shows the absolute difference for each check. The uncertainty from lower tracking efficiency is much smaller than traditional cumulant method, since 2 sub-event method already largely suppressed the short-range non-flow sources. The uncertainty from the pile-up effects becomes larger and dominates the high multiplicity region, but that is partially due to a larger statistical errors (not shown in the plots for better readability) with the sub-event method. But in any case, the systematic from pile-up effects is still quoted even with large statistical errors. Overall, the combined systematics are within 0.5×10^{-6} .

In the end, the summary of systematics from 3 sub-event method is shown in Fig. 61, which has the smallest systematic uncertainties among all 3 methods. The combined absolute difference is 0.3×10^{-6} .

For completeness, the summary of systematics for 3 methods, but with higher p_T region: $0.5 < p_T < 5.0$ GeV, are shown in Fig. 62. Compared with lower p_T region, the uncertainties from all three methods grows larger. This is partially due to the fact that the statistical errors becomes larger thus the systematic fluctuation also becomes larger. But the main reason is that the fraction of non-flow contribution is larger when moving to the higher p_T range. A comparison between systematics from three methods shows that 3 sub-event method still has the smallest uncertainty, consistent with the lower p_T case.

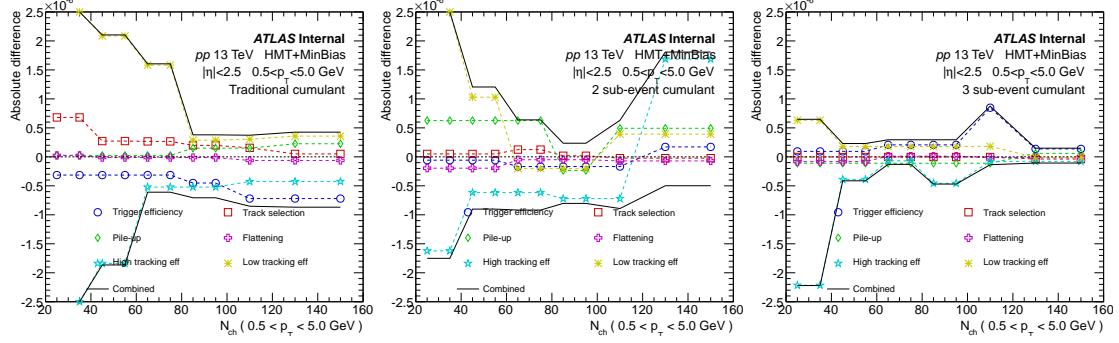


Figure 62: Summary of systematics in 13 TeV pp , for 3 different cumulant methods, with $0.5 < p_T < 5.0$ GeV. The statistical errors are not shown for better readability.

1081 5.2 pp 5.02 TeV

1082 5.2.1 Summary of systematics

1083 The systematic checks for 5.02 TeV pp are identical to 13 TeV pp :

- 1084 • Trigger efficiency: comparison between 50% and 80% HMT efficiency cut;
- 1085 • Track selection: tighten the z_0 and d_0 pointing cut to 1.0;
- 1086 • Pile-up effects: comparison between runs with $\mu < 0.7$ and $\mu > 1.5$;
- 1087 • Flattening procedure: comparison between with and without flattening in ϕ ;
- 1088 • High tracking efficiency: increasing tracking efficiency by its systematic uncertainty;
- 1089 • Low tracking efficiency: decreasing tracking efficiency by its systematic uncertainty;

1090 For simplicity, we will not go through the systematics one by one, only the summary of systematics are
1091 shown in the following plots. The total statistics in 5 TeV pp is much smaller than 13 TeV pp . Since
1092 the systematics will be larger due to larger statistical fluctuations. The results are re-binned to reduce the
1093 statistical fluctuation when necessary.

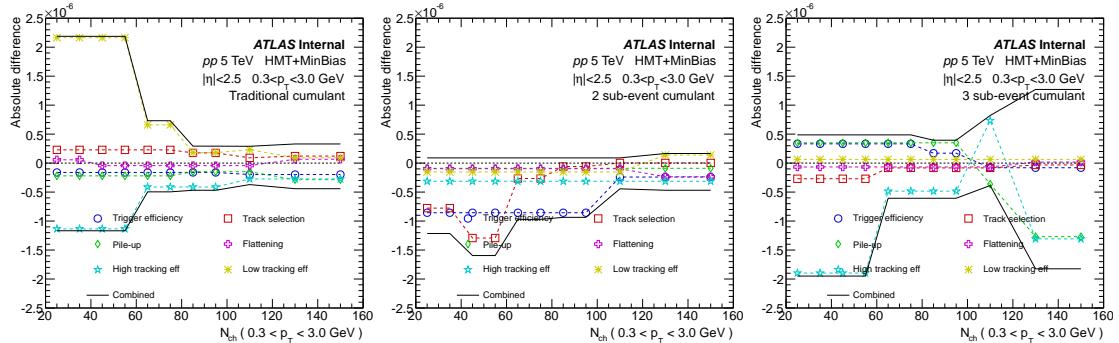


Figure 63: Summary of systematics in 5.02 TeV pp , for three different cumulant methods, with $0.3 < p_T < 3.0$ GeV. The statistical errors are not shown for better readability.

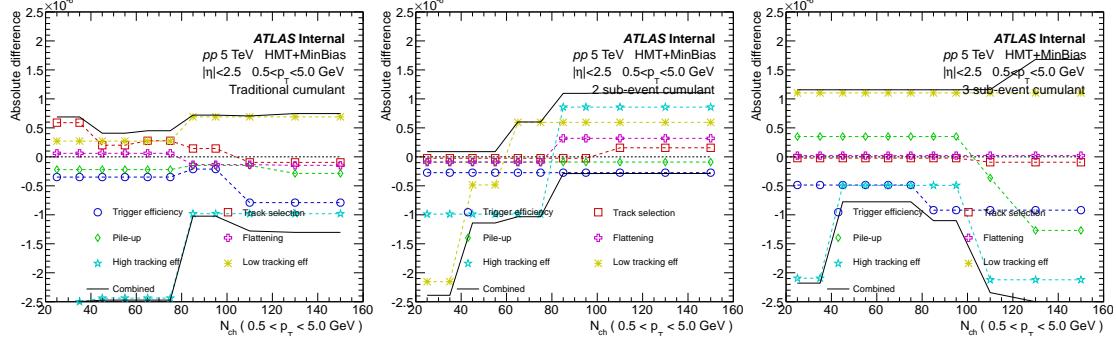


Figure 64: Summary of systematics in 5.02 TeV $p\bar{p}$, for three different cumulant methods, with $0.5 < p_T < 5.0$ GeV. The statistical errors are not shown for better readability.

1094 5.3 $p+\text{Pb}$ 5.02 TeV

1095 5.3.1 Pileup rejection in run 1

1096 In the 2013 $p+\text{Pb}$ run, the luminosity conditions provided by the LHC results in an average probability
 1097 of 3% that an event contains two or more $p+\text{Pb}$ collisions (pileup). The pileup events are suppressed by
 1098 the following two criteria: [33]

- 1099 • Rejecting events containing more than one good reconstructed vertex (sum $p_T > 5$ GeV);
- 1100 • A simple cut on the high tail end of the ZDC signal distribution;

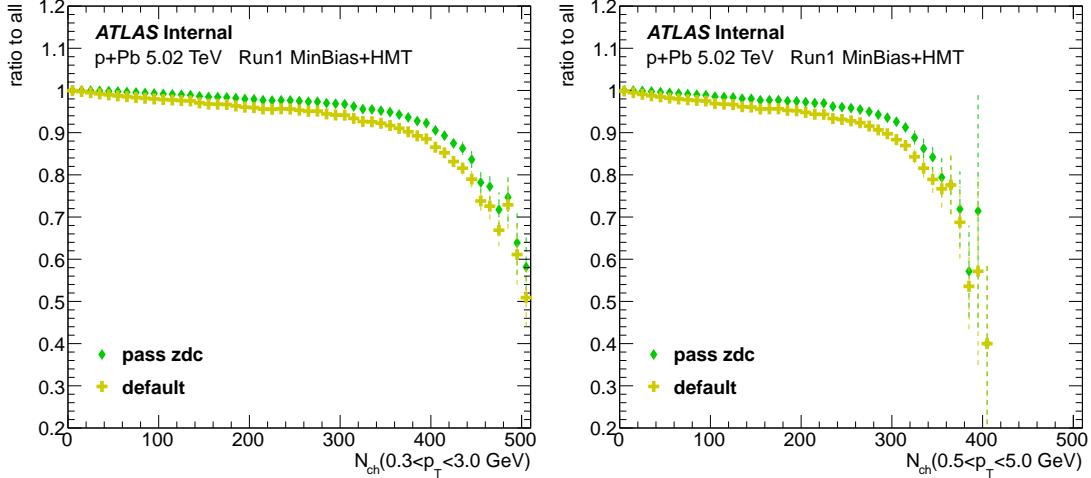


Figure 65: Fractions of remaining events after ZDC cut and additional vertex cut (default), in Run 1 5.02 TeV $p+\text{Pb}$, for two p_T ranges: $0.3 < p_T < 3.0$ GeV (left) and $0.5 < p_T < 5.0$ GeV (right).

1101 Fig. 65 shows the fractions of remaining events after ZDC cut and additional vertex cut (default), for
 1102 two p_T ranges. In this analysis, the default procedure is with both vertex and ZDC cut. As expected, the
 1103 fraction of pileup events increase as N_{ch} increases, reaches 50% in the highest N_{ch} region. Compared
 1104 with ZDC cut, additional vertex cut can further reject pileup events.

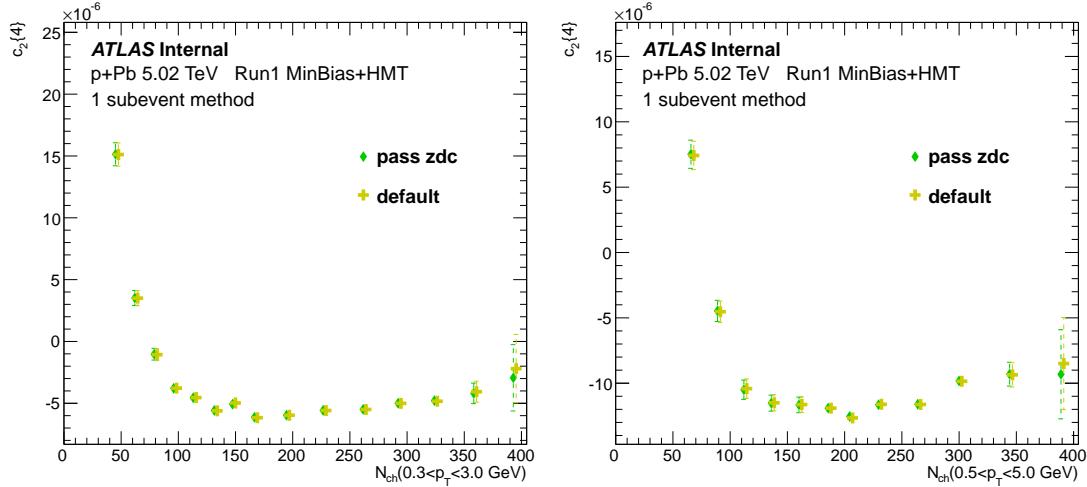


Figure 66: Comparisons of $c_2\{4\}$ calculated in events with different pileup rejections, after ZDC cut and additional vertex cut (default), with standard cumulant method, in Run 1 5.02 TeV $p+Pb$, for two p_T ranges: $0.3 < p_T < 3.0$ GeV (left) and $0.5 < p_T < 5.0$ GeV (right). The x axis are shifted on purpose for better comparison.

To gain an idea how pileup rejection could affect the final results, Fig. 66 shows the comparisons of $c_2\{4\}$ calculated in events with different pileup rejections: after ZDC cut and additional vertex cut (default), with standard cumulant method. The results are very consistent between two cuts: even though additional vertex cut further cleans up pileup events, the results are already very stable even with the residual pileup events. The 2 and 3 subevent methods give similar behaviors but with larger errors.

5.3.2 Summary of systematics

The systematic checks for 5.02 TeV $p+Pb$ are similar to pp , pile-up effects are replaced with run period dependence:

- Trigger efficiency: comparison between 50% and 80% HMT efficiency cut;
- Track selection: tighten the z_0 and d_0 pointing cut;
- Run period: comparison results before and after beam reversal;
- Flattening procedure: comparison between with and without flattening in ϕ ;
- High tracking efficiency: increasing tracking efficiency by its systematic uncertainty;
- Low tracking efficiency: decreasing tracking efficiency by its systematic uncertainty;

For simplicity, we will not go through the systematics one by one, only the summary of systematics are shown in the following plots.

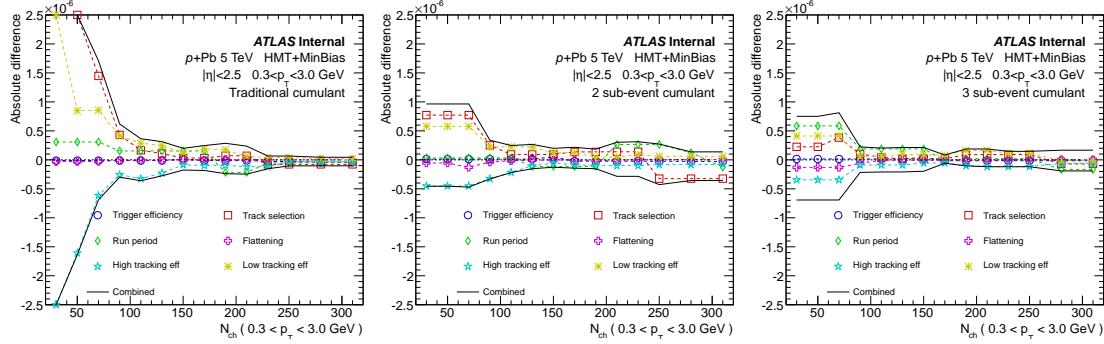


Figure 67: Summary of systematics in 5.02 TeV pPb , for three different cumulant methods, with $0.3 < p_T < 3.0$ GeV. The statistical errors are not shown for better readability.

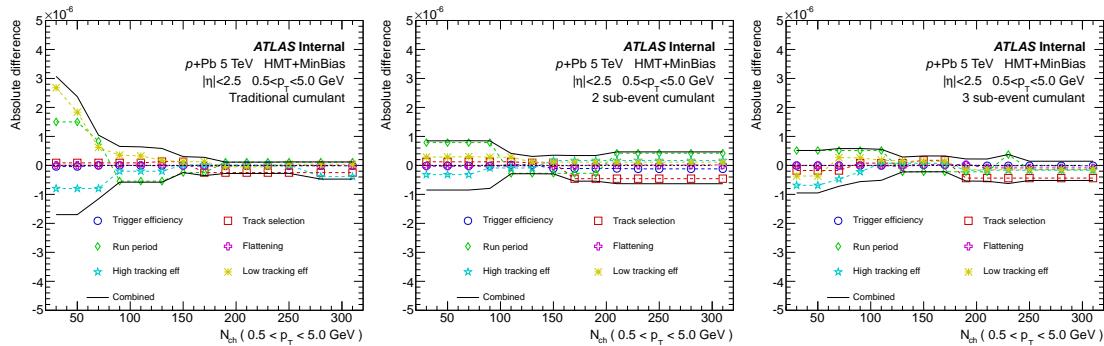


Figure 68: Summary of systematics in 5.02 TeV pPb , for three different cumulant methods, with $0.5 < p_T < 5.0$ GeV. The statistical errors are not shown for better readability.

1121 5.4 Summary of systematic table

1122 The systematics of three collision systems, three cumulant methods and two p_T ranges are summarized
 1123 in Fig. 69, as a function of N_{ch} .

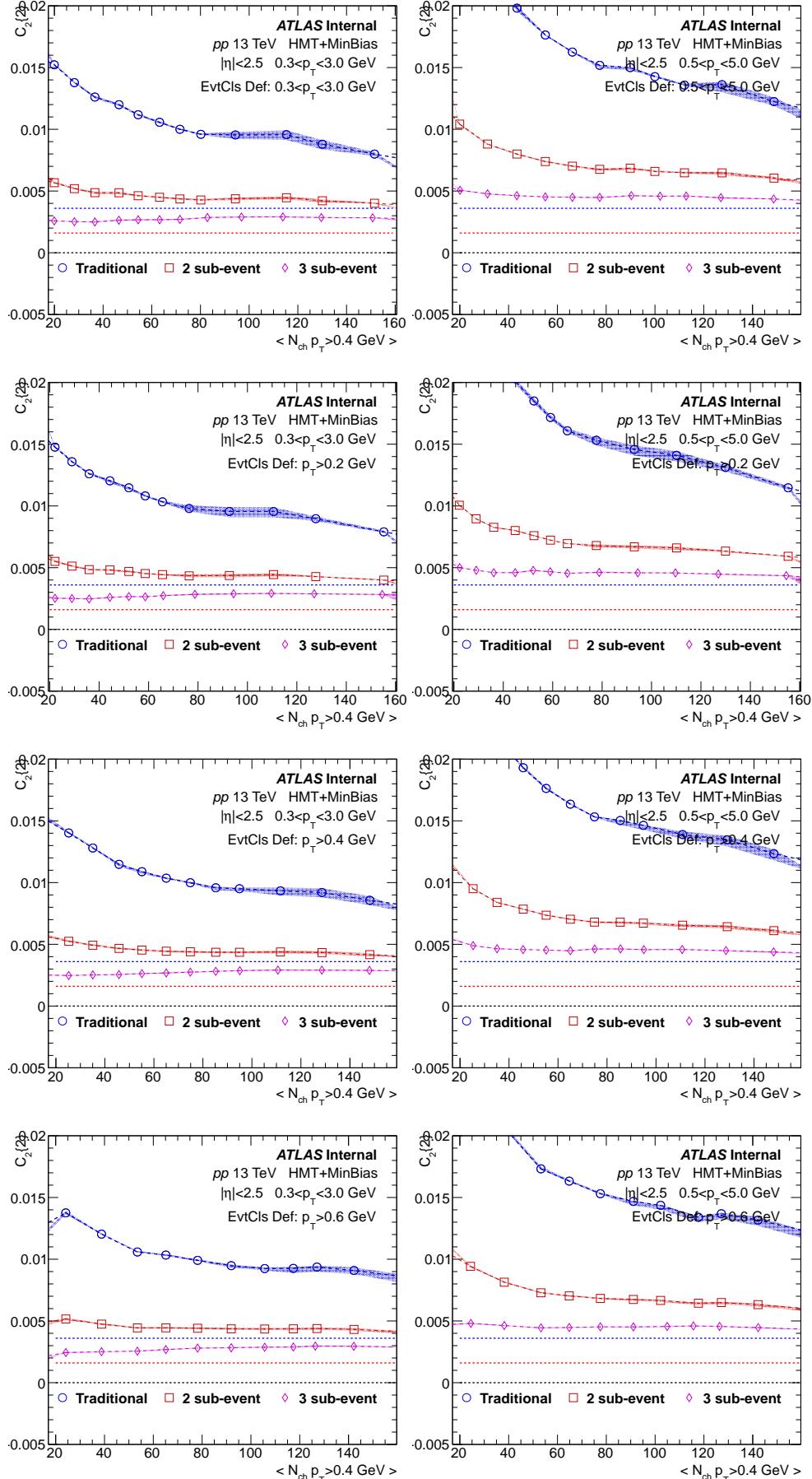
13 TeV $p\bar{p}$ [10E-6]	0.3 < p_T < 3.0 GeV												0.5 < p_T < 5.0 GeV												
	Traditional			2 sub-event			3 sub-event			Traditional			2 sub-event			3 sub-event									
	(0,50)	(50,100)	>100	(0,50)	(50,100)	>100	(0,50)	(50,100)	>100	(0,50)	(50,100)	>100	(0,50)	(50,100)	>100	(0,50)	(50,100)	>100	(0,50)	(50,100)	>100	(0,50)	(50,100)	>100	
Trigger efficiency	-0.15	-0.085	-0.01	-0.12	-0.16	-0.051	-0.15	-0.067	-0.06	-0.31	-0.34	-0.64	-0.057	-0.13	-0.083	0.094	0.16	0.48							
Track selection	2.1	0.12	0.049	0.14	0.068	0.034	0.06	-0.083	-0.035	1.4	0.26	0.14	0.05	0.08	-0.01	-0.01	-0.01	-0.01	-0.01	-0.01	-0.01	-0.01	-0.01	-0.01	
Pile-up effects	0.025	0.1	0.22	0.63	0.29	0.5	-0.11	-0.11	0.045	0.025	0.045	0.17	0.63	0.49	0.28	-0.11	-0.11	-0.068							
Flattening	-0.014	-0.01	0.033	0.032	0.023	-0.01	-0.097	-0.037	-0.01	0.022	-0.012	-0.047	-0.2	-0.1	-0.064	-0.082	-0.023	-0.01							
High track eff	-3.5	-0.53	-0.11	-0.31	-0.18	-0.061	-0.2	-0.16	-0.094	-8.6	-1	-0.45	-3	-0.63	-0.12	-2.1	-0.25	-0.19							
Low track eff	3.1	0.68	0.11	-0.18	0.24	0.16	0.6	0.029	0.042	9.7	1.6	0.31	2.6	0.26	0.22	0.59	0.18	0.13							
Upper band	3.9	0.71	0.27	0.81	0.53	0.53	0.7	0.17	0.077	9.8	1.6	0.39	2.7	0.78	0.81	0.61	0.27	0.52							
Lower band	-3.5	-0.58	-0.25	-1	-0.55	-0.51	-0.31	-0.26	-0.14	-8.6	-1.1	-0.81	-3.1	-0.89	-0.77	-2.1	-0.29	-0.23							
5 TeV $p\bar{p}$ [10E-6]	0.3 < p_T < 3.0 GeV												0.5 < p_T < 5.0 GeV												
	Traditional			2 sub-event			3 sub-event			Traditional			2 sub-event			3 sub-event									
	(0,40)	(40,80)	>80	(0,40)	(40,80)	>80	(0,40)	(40,80)	>80	(0,40)	(40,80)	>80	(0,40)	(40,80)	>80	(0,40)	(40,80)	>80	(0,40)	(40,80)	>80	(0,40)	(40,80)	>80	
Trigger efficiency	-0.16	-0.16	-0.19	-0.86	-0.86	-0.35	0.33	0.29	-0.036	-0.35	-0.35	-0.33	-0.27	-0.27	-0.27	-0.49	-0.49	-0.8							
Track selection	0.23	0.21	0.12	-0.84	-0.52	-0.01	-0.27	-0.14	-0.063	1	0.28	0.15	-0.024	-0.024	0.01	-0.022	-0.022	-0.032							
Pile-up effects	-0.22	-0.2	-0.2	-0.09	-0.09	-0.09	0.35	0.35	-0.59	-0.22	-0.22	-0.17	-0.09	-0.09	-0.09	0.35	0.35	0.25							
Flattening	0.044	-0.043	-0.01	-0.096	-0.096	-0.15	-0.067	-0.067	-0.034	0.061	0.061	-0.079	-0.091	-0.091	0.2	0.022	0.022	0.022							
High track eff	-1.1	-0.63	-0.3	-0.31	-0.31	-0.31	-1.9	-0.91	-0.28	-5.9	-2.5	-1.4	-0.99	-0.99	0.33	-2.1	-0.72	-0.73							
Low track eff	2.2	0.99	0.17	-0.15	-0.15	-0.039	0.064	0.064	0.064	0.27	0.27	0.57	-2.2	-0.41	0.6	-0.98	1.1	1.1							
Upper band	2.2	1.1	0.31	0.09	0.09	0.12	0.49	0.46	0.92	1.1	0.46	0.64	0.09	0.24	0.96	0.7	1.2	1.2							
Lower band	-1.2	-0.69	-0.42	-1.3	-1.2	-0.54	-1.9	-1	-0.99	-6	-2.6	-1.5	-2.4	-1.3	-0.5	-2.8	-0.98	-0.53							
5 TeV $p\bar{p}+Pb$ [10E-6]	0.3 < p_T < 3.0 GeV												0.5 < p_T < 5.0 GeV												
	Traditional			2 sub-event			3 sub-event			Traditional			2 sub-event			3 sub-event									
	0-100	100,200	>200	0-100	100,200	>200	0-100	100,200	>200	0-100	100,200	>200	0-100	100,200	>200	0-100	100,200	>200	0-100	100,200	>200	0-100	100,200	>200	
Trigger efficiency	-0.013	-0.01	-0.04	0.013	-0.01	-0.026	0.015	0.01	-0.015	-0.029	-0.01	-0.014	-0.016	-0.033	-0.11	-0.01	0.01	-0.01							
Track selection	3.4	0.048	-0.078	0.58	0.12	-0.28	0.2	0.052	-0.035	0.087	0.06	-0.25	0.12	0.026	-0.46	-0.16	0.11	-0.38							
Run period	0.26	0.01	-0.042	0.035	0.01	0.047	0.46	0.076	-0.12	1.2	-0.43	0.079	0.79	-0.1	0.19	0.51	0.07	-0.066							
Flattening	-0.022	-0.01	-0.01	-0.059	-0.01	0.01	-0.1	-0.01	0.01	0.01	-0.01	-0.01	-0.026	-0.017	0.012	-0.01	-0.051								
High track eff	-1.6	-0.13	-0.061	-0.4	-0.093	-0.082	-0.26	-0.087	-0.073	-0.76	-0.12	0.016	-0.3	0.01	0.17	-0.61	0.026	-0.14							
Low track eff	2.4	0.15	0.029	0.47	0.11	0.063	0.3	0.12	0.015	2.9	0.22	-0.015	0.29	0.18	0.11	-0.18	0.17	-0.12							
Upper band	4.3	0.24	0.051	0.75	0.24	0.17	0.59	0.17	0.16	3.3	0.5	0.13	0.85	0.43	0.42	0.53	0.41	-0.6	-0.87	-0.35	-0.53				
Lower band	-1.7	-0.22	-0.12	-0.4	-0.19	-0.35	-0.54	-0.14	-0.18	-1.5	-0.47	-0.28	-0.85	-0.41	-0.6	-0.87	-0.35								

Figure 69: Summary of systematic table of three collision systems, three cumulant methods and two p_T ranges, as a function of N_{ch} . The values in the table are absolute differences of each check, in the unit of 10^{-6} .

1124 6 Results

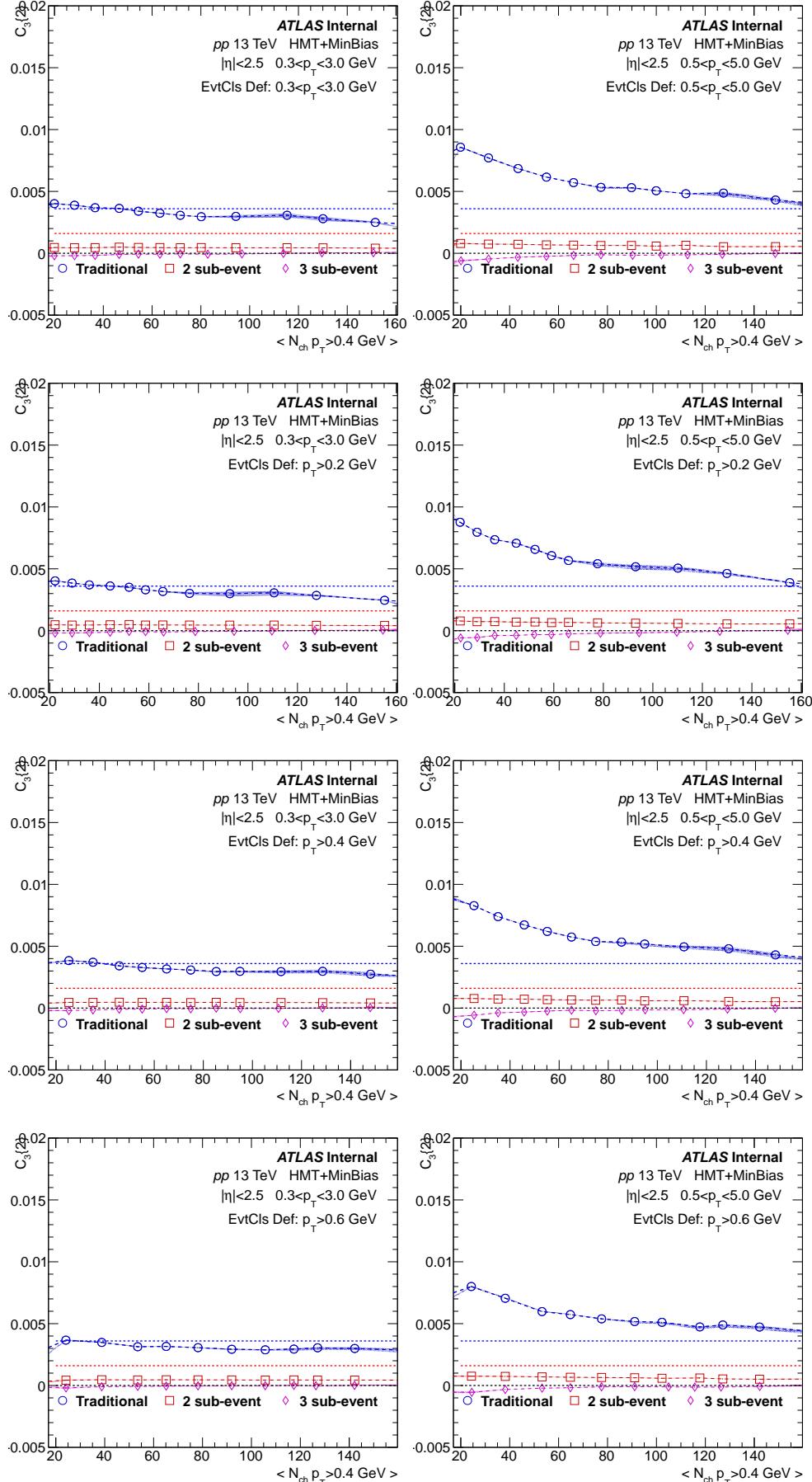
1125 6.1 13 TeV pp $C_2\{2\}$

1126 2-particle cumulant results of v_2 harmonic from 13 TeV pp are summarized in Fig. 70. Four rows have
1127 different event class definitions and two columns are particles with different p_T ranges. In each panel,
1128 $C_2\{2\}$ calculated using three cumulant methods are compared. In particular, for 2 sub-event method,
1129 two particles come from two sub-events, while for 3 sub-event method, two particles are separated by
1130 one sub-event in the mid- η . Red dash line represents 4% v_2 signal while blue dash line represents 6%
1131 v_2 signal. Traditional method has largest $C_2\{2\}$, due to largest residual non-flow contribution. 2 sub-
1132 event cumulant already suppresses non-flow and gives smaller $C_2\{2\}$ values. $C_2\{2\}$ from 3 sub-event
1133 cumulant is the smallest since most short-range non-flow correlations are removed with the η gap. $C_2\{2\}$
1134 from all methods increase moving from lower p_T range to higher p_T range, which is consistent with 2PC
1135 results using template fit method [11, 15]. The 2-particle cumulant results are not sensitive to the event
1136 class definition.

Figure 70: Comparison of $C_2\{2\}$ calculated with 3 cumulant methods, from 13 TeV pp .

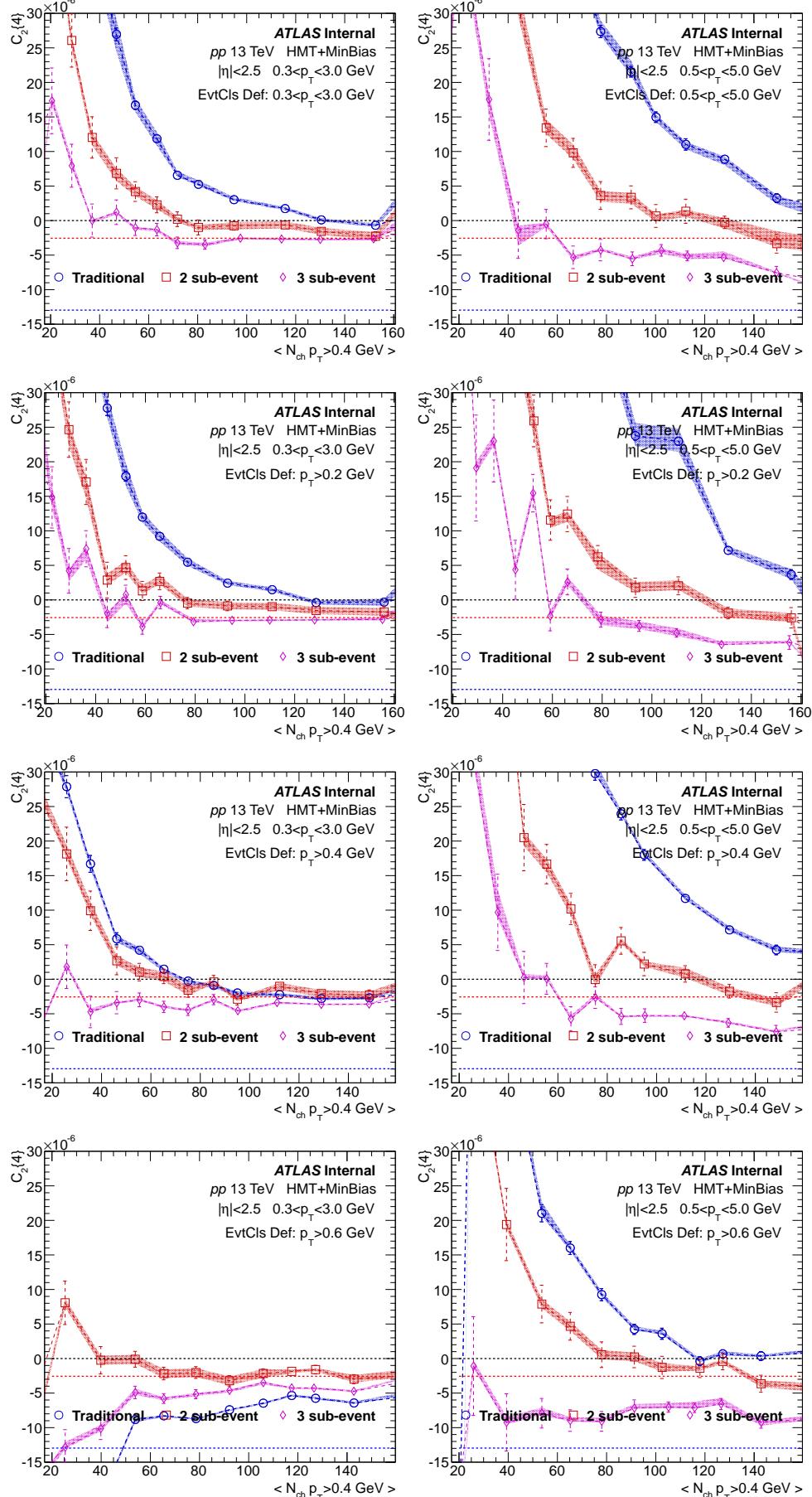
¹¹³⁷ **6.2 13 TeV pp $C_3\{2\}$**

¹¹³⁸ 2-particle cumulant results of v_3 harmonic from 13 TeV pp are summarized in Fig. 71. Four rows have
¹¹³⁹ different event class definitions and two columns are particles with different p_T ranges. In each panel,
¹¹⁴⁰ $C_3\{2\}$ calculated using three cumulant methods are compared. In particular, for 2 sub-event method,
¹¹⁴¹ two particles come from two sub-events, while for 3 sub-event method, two particles are separated by
¹¹⁴² one sub-event in the mid- η . Red dash line represents 4% v_3 signal while blue dash line represents 6% v_3
¹¹⁴³ signal. Traditional cumulant measures positive v_3 signal, and it increases as p_T moves to higher range.
¹¹⁴⁴ Meanwhile, $C_3\{2\}$ from 2 sub-event method is much smaller and $C_3\{2\}$ from 3 sub-event method is
¹¹⁴⁵ consistent with 0 with $0.3 < p_T < 3.0$ GeV, and it even goes to negative (wrong sign) in the low-
¹¹⁴⁶ multiplicity with $0.5 < p_T < 5.0$ GeV. Like the $C_2\{2\}$ results, $C_3\{2\}$ are not sensitive to the event class
¹¹⁴⁷ definition.

Figure 71: Comparison of $C_3\{2\}$ calculated with 3 cumulant methods, from 13 TeV pp .

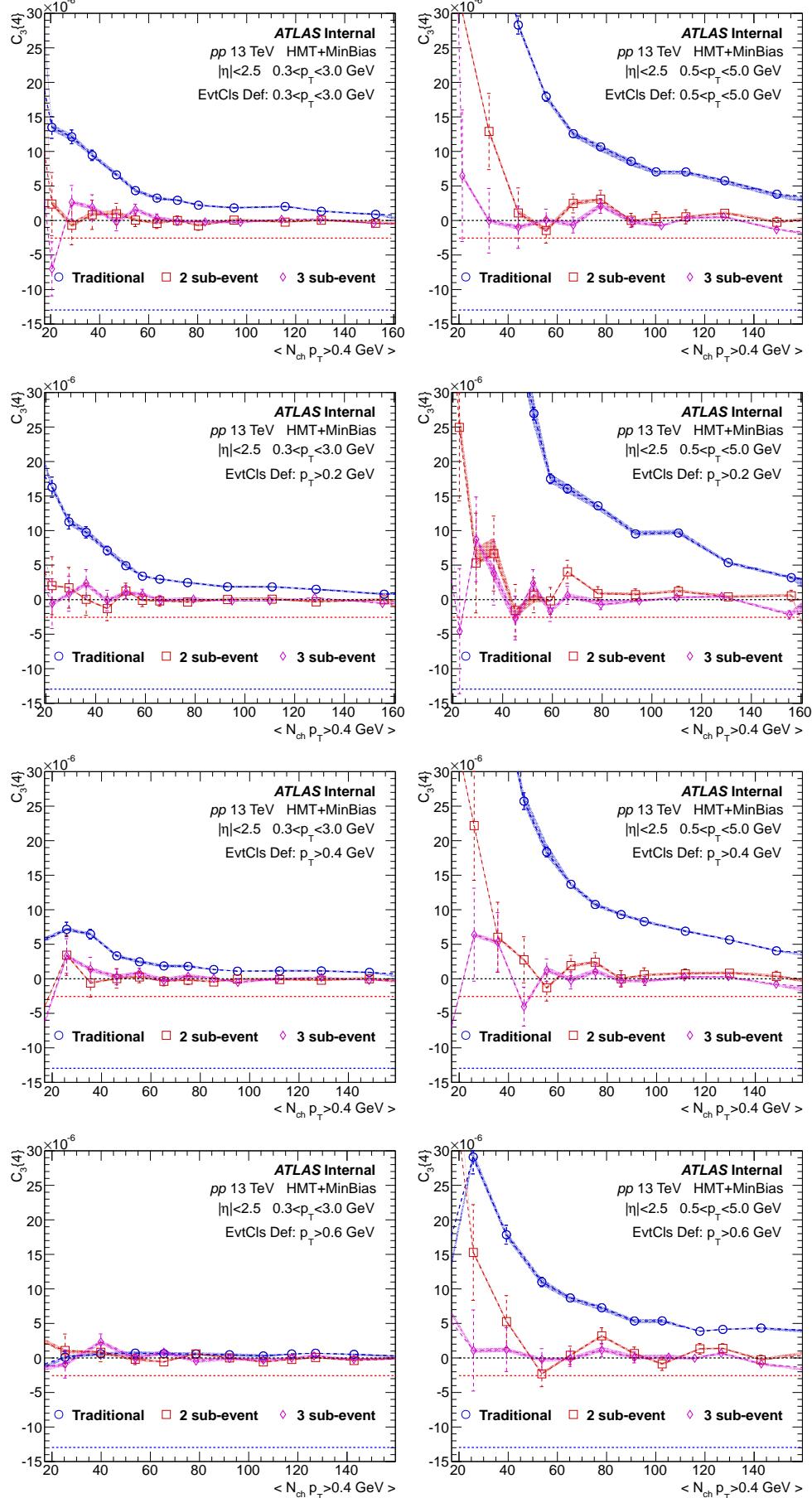
¹¹⁴⁸ 6.3 13 TeV pp $C_2\{4\}$

¹¹⁴⁹ 4-particle cumulant results of v_2 harmonic from 13 TeV pp are summarized in Fig. 72. Four rows have
¹¹⁵⁰ different event class definitions and two columns are particles with different p_T ranges. In each panel,
¹¹⁵¹ $C_2\{4\}$ calculated using three cumulant methods are compared. Red dash line represents 4% v_2 signal
¹¹⁵² while blue dash line represents 6% v_2 signal. Compared with 2-particle cumulant, 4-particle cumulant has
¹¹⁵³ much larger statistical errors thus fluctuation from point to point is expected. For a well-defined $v_2\{4\}$,
¹¹⁵⁴ $C_2\{4\}$ has to be negative. However, traditional cumulant always gives the positive $C_2\{4\}$ with default
¹¹⁵⁵ event class definition. This means the residual non-flow can even change the sign of $C_2\{4\}$. Meanwhile,
¹¹⁵⁶ $C_2\{4\}$ from 2 sub-event method is already much suppressed and $C_2\{4\}$ from 3 sub-event method stays
¹¹⁵⁷ negative in most of N_{ch} ranges, except in the lowest N_{ch} region. As moved from $0.3 < p_T < 3.0$ GeV to
¹¹⁵⁸ $0.5 < p_T < 5.0$ GeV, because of larger fraction of non-flow, both traditional and 2 sub-event cumulant
¹¹⁵⁹ tend to give the wrong sign of $C_2\{4\}$. Only $C_2\{4\}$ from 3 sub-event becomes more negative as p_T goes
¹¹⁶⁰ higher, which is consistent with the existing 2PC results. This provides another evidence that only 3 sub-
¹¹⁶¹ event cumulant can effectively suppress the non-flow and give reasonable v_2 measurement. Last but not
¹¹⁶² least, both traditional and 2 sub-event cumulant results are sensitive to the event class definition, while 3
¹¹⁶³ sub-event method gives rather stable results.

Figure 72: Comparison of $C_2\{4\}$ calculated with 3 cumulant methods, from 13 TeV pp .

¹¹⁶⁴ **6.4 13 TeV pp $C_3\{4\}$**

¹¹⁶⁵ 4-particle cumulant results of v_3 harmonic from 13 TeV pp are summarized in Fig. 73. Four rows have
¹¹⁶⁶ different event class definitions and two columns are particles with different p_T ranges. In each panel,
¹¹⁶⁷ $C_3\{4\}$ calculated using three cumulant methods are compared. Red dash line represents 4% v_2 signal
¹¹⁶⁸ while blue dash line represents 6% v_2 signal. Like the $C_2\{4\}$, non-flow contaminates the results from
¹¹⁶⁹ traditional method, and it becomes even larger as p_T goes higher. Meanwhile, $C_3\{4\}$ from 2 sub-event
¹¹⁷⁰ and 3 sub-event methods are consistent with 0, which is partially due to that the mean value of v_3 is much
¹¹⁷¹ smaller than v_2 , and fluctuation of v_3 makes it very hard to measure in small systems, using cumulant
¹¹⁷² method. It is interesting to repeat the measurement in Pb+Pb to see whether non-zero $v_3\{4\}$ can be
¹¹⁷³ measured.

Figure 73: Comparison of $C_3\{4\}$ calculated with 3 cumulant methods, from 13 TeV pp .

¹¹⁷⁴ **6.5 5.02 TeV pp $C_2\{2\}$**

¹¹⁷⁵ 2-particle cumulant results of v_2 harmonic from 5.02 TeV pp are summarized in Fig. 74. Compared
¹¹⁷⁶ with 13 TeV pp results, because of much less statistics, larger point-to-point fluctuations are expected.
¹¹⁷⁷ Four rows have different event class definitions and two columns are particles with different p_T ranges.
¹¹⁷⁸ In each panel, $C_2\{2\}$ calculated using three cumulant methods are compared. In particular, for 2 sub-
¹¹⁷⁹ event method, two particles come from two sub-events, while for 3 sub-event method, two particles
¹¹⁸⁰ are separated by one sub-event in the mid- η . Red dash line represents 4% v_2 signal while blue dash
¹¹⁸¹ line represents 6% v_2 signal. Traditional method has largest $C_2\{2\}$, due to largest residual non-flow
¹¹⁸² contribution. 2 sub-event cumulant already suppresses non-flow and gives smaller $C_2\{2\}$ values. $C_2\{2\}$
¹¹⁸³ from 3 sub-event cumulant is the smallest since most short-range non-flow correlations are removed with
¹¹⁸⁴ the η gap. $C_2\{2\}$ from all methods increase moving from lower p_T range to higher p_T range, which is
¹¹⁸⁵ consistent with 2PC results using template fit method. The 2-particle cumulant results are not sensitive
¹¹⁸⁶ to the event class definition.

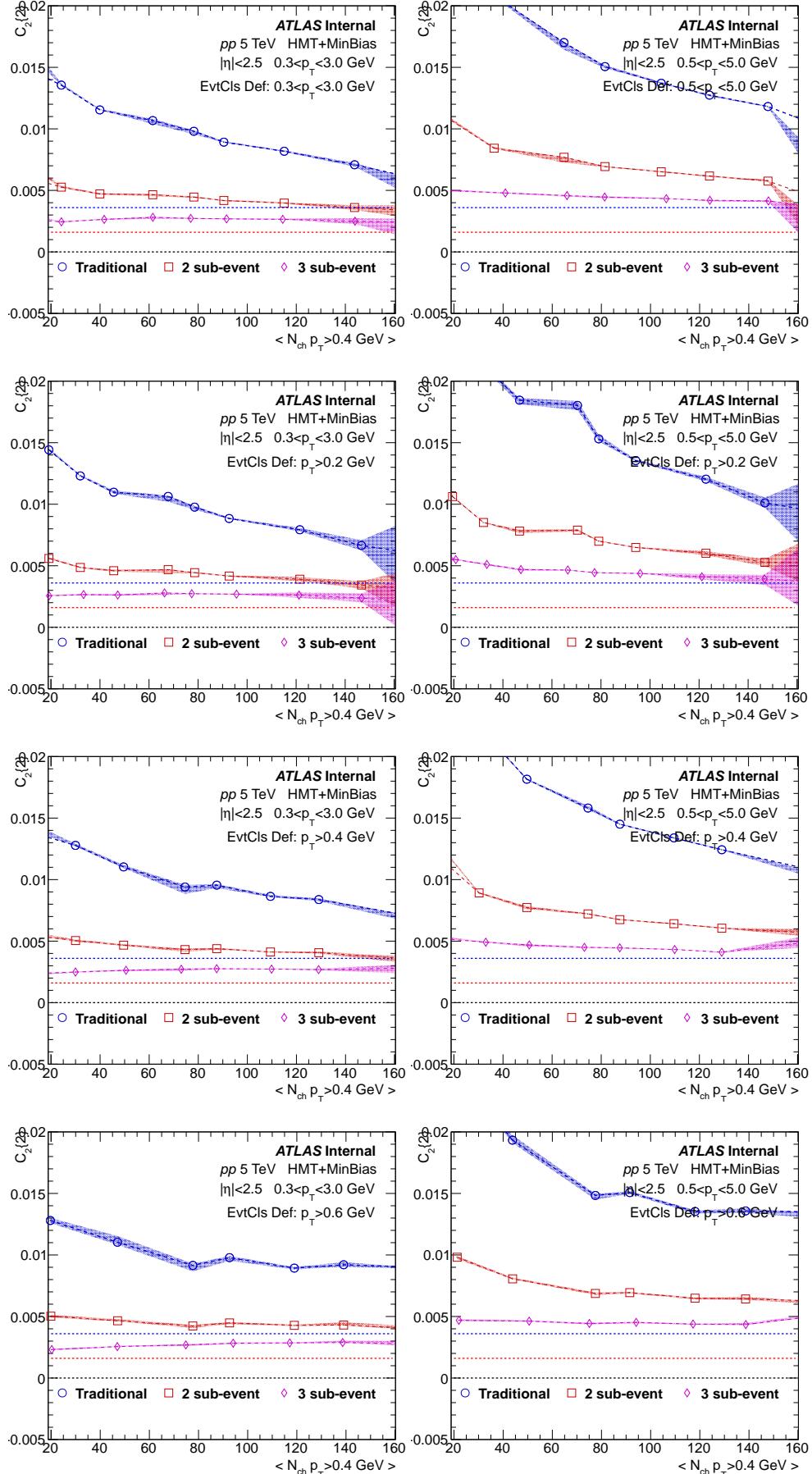


Figure 74: Comparison of $C_2\{2\}$ calculated with 3 cumulant methods, from 5.02 TeV pp .

¹¹⁸⁷ **6.6 5.02 TeV $pp C_3\{2\}$**

¹¹⁸⁸ 2-particle cumulant results of v_3 harmonic from 5.02 TeV pp are summarized in Fig. 75. Four rows have
¹¹⁸⁹ different event class definitions and two columns are particles with different p_T ranges. In each panel,
¹¹⁹⁰ $C_3\{2\}$ calculated using three cumulant methods are compared. In particular, for 2 sub-event method,
¹¹⁹¹ two particles come from two sub-events, while for 3 sub-event method, two particles are separated by
¹¹⁹² one sub-event in the mid- η . Red dash line represents 4% v_3 signal while blue dash line represents 6% v_3
¹¹⁹³ signal. Traditional cumulant measures positive v_3 signal, and it increases as p_T moves to higher range.
¹¹⁹⁴ Meanwhile, $C_3\{2\}$ from 2 sub-event method is much smaller and $C_3\{2\}$ from 3 sub-event method is
¹¹⁹⁵ consistent with 0 with $0.3 < p_T < 3.0$ GeV, and it even goes to negative (wrong sign) in the low-
¹¹⁹⁶ multiplicity with $0.5 < p_T < 5.0$ GeV. Like the $C_2\{2\}$ results, $C_3\{2\}$ are not sensitive to the event class
¹¹⁹⁷ definition.

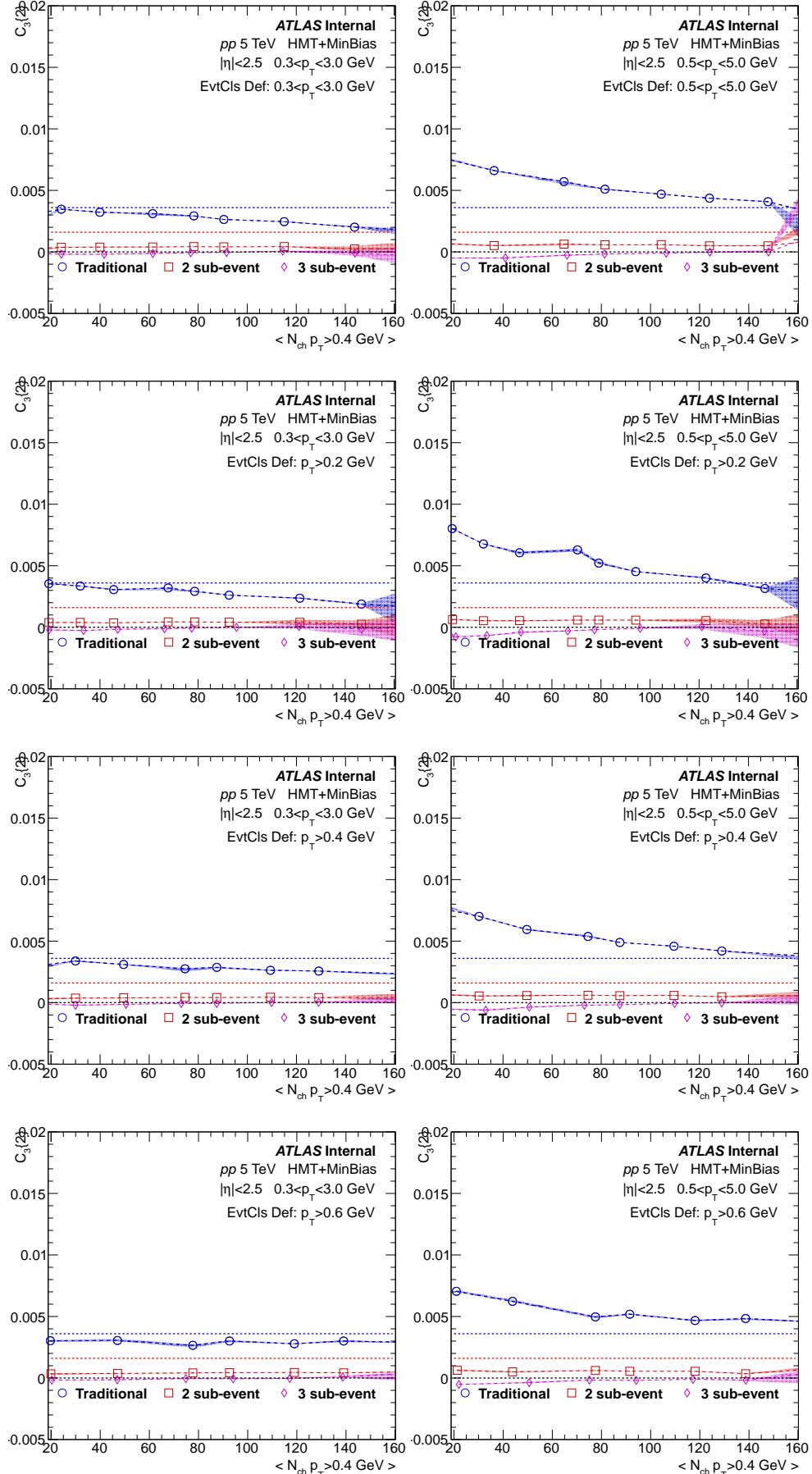
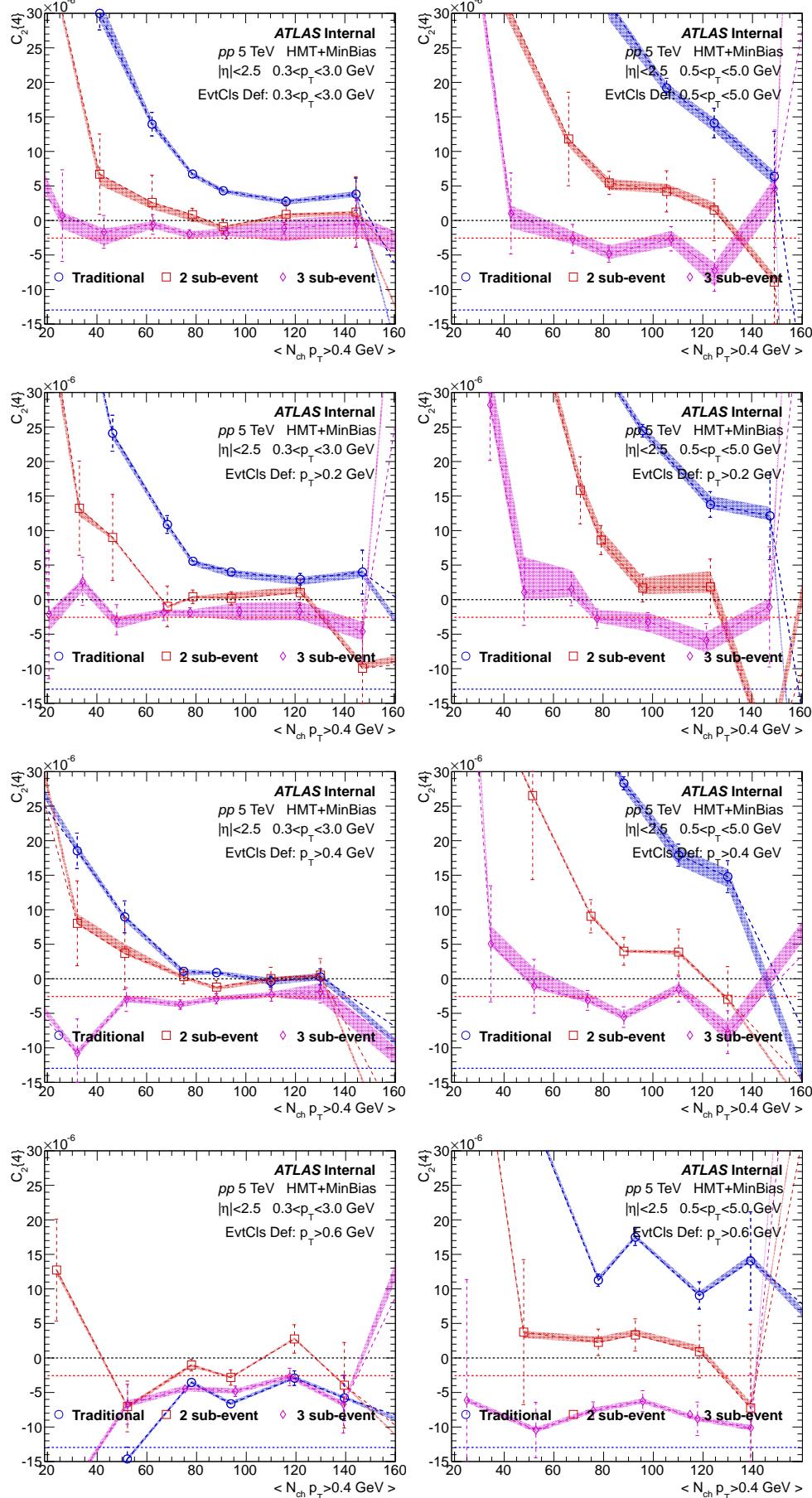


Figure 75: Comparison of $C_3\{2\}$ calculated with 3 cumulant methods, from 5.02 TeV pp .

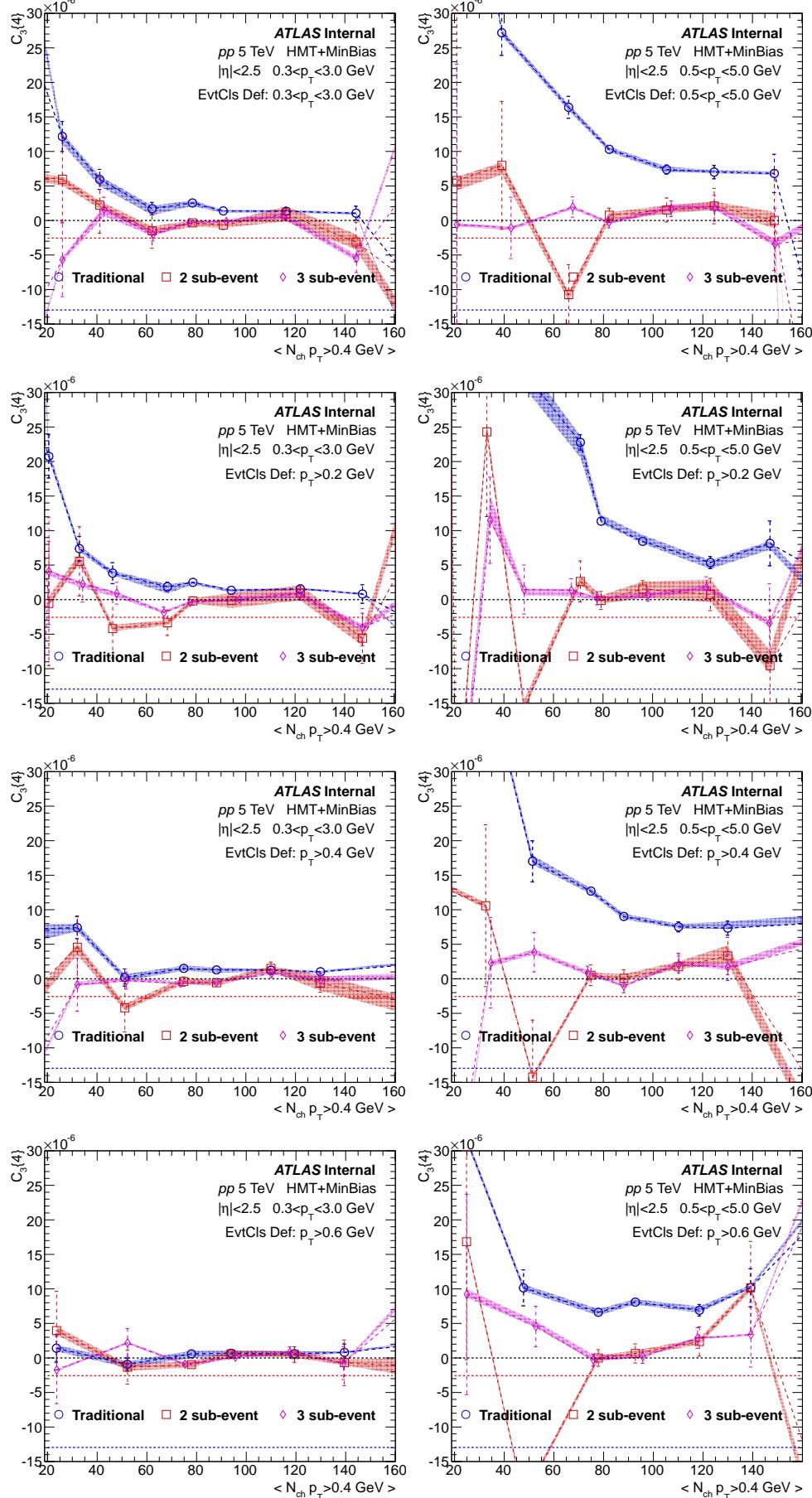
¹¹⁹⁸ **6.7 5.02 TeV pp $C_2\{4\}$**

¹¹⁹⁹ 4-particle cumulant results of v_2 harmonic from 5.02 TeV pp are summarized in Fig. 76. Four rows have
¹²⁰⁰ different event class definitions and two columns are particles with different p_T ranges. In each panel,
¹²⁰¹ $C_2\{4\}$ calculated using three cumulant methods are compared. Red dash line represents 4% v_2 signal
¹²⁰² while blue dash line represents 6% v_2 signal. Compared with 2-particle cumulant, 4-particle cumulant has
¹²⁰³ much larger statistical errors thus fluctuation from point to point is expected. For a well-defined $v_2\{4\}$,
¹²⁰⁴ $C_2\{4\}$ has to be negative. However, traditional cumulant always gives the positive $C_2\{4\}$ with default
¹²⁰⁵ event class definition. This means the residual non-flow can even change the sign of $C_2\{4\}$. Meanwhile,
¹²⁰⁶ $C_2\{4\}$ from 2 sub-event method is already much suppressed and $C_2\{4\}$ from 3 sub-event method stays
¹²⁰⁷ negative in most of N_{ch} ranges, except in the lowest N_{ch} region. As moved from $0.3 < p_T < 3.0$ GeV to
¹²⁰⁸ $0.5 < p_T < 5.0$ GeV, because of larger fraction of non-flow, both traditional and 2 sub-event cumulant
¹²⁰⁹ tend to give the wrong sign of $C_2\{4\}$. Only $C_2\{4\}$ from 3 sub-event becomes more negative as p_T goes
¹²¹⁰ higher, which is consistent with the existing 2PC results. This provides another evidence that only 3 sub-
¹²¹¹ event cumulant can effectively suppress the non-flow and give reasonable v_2 measurement. Last but not
¹²¹² least, both traditional and 2 sub-event cumulant results are sensitive to the event class definition, while 3
¹²¹³ sub-event method gives rather stable results.

Figure 76: Comparison of $C_2\{4\}$ calculated with 3 cumulant methods, from 5.02 TeV pp .

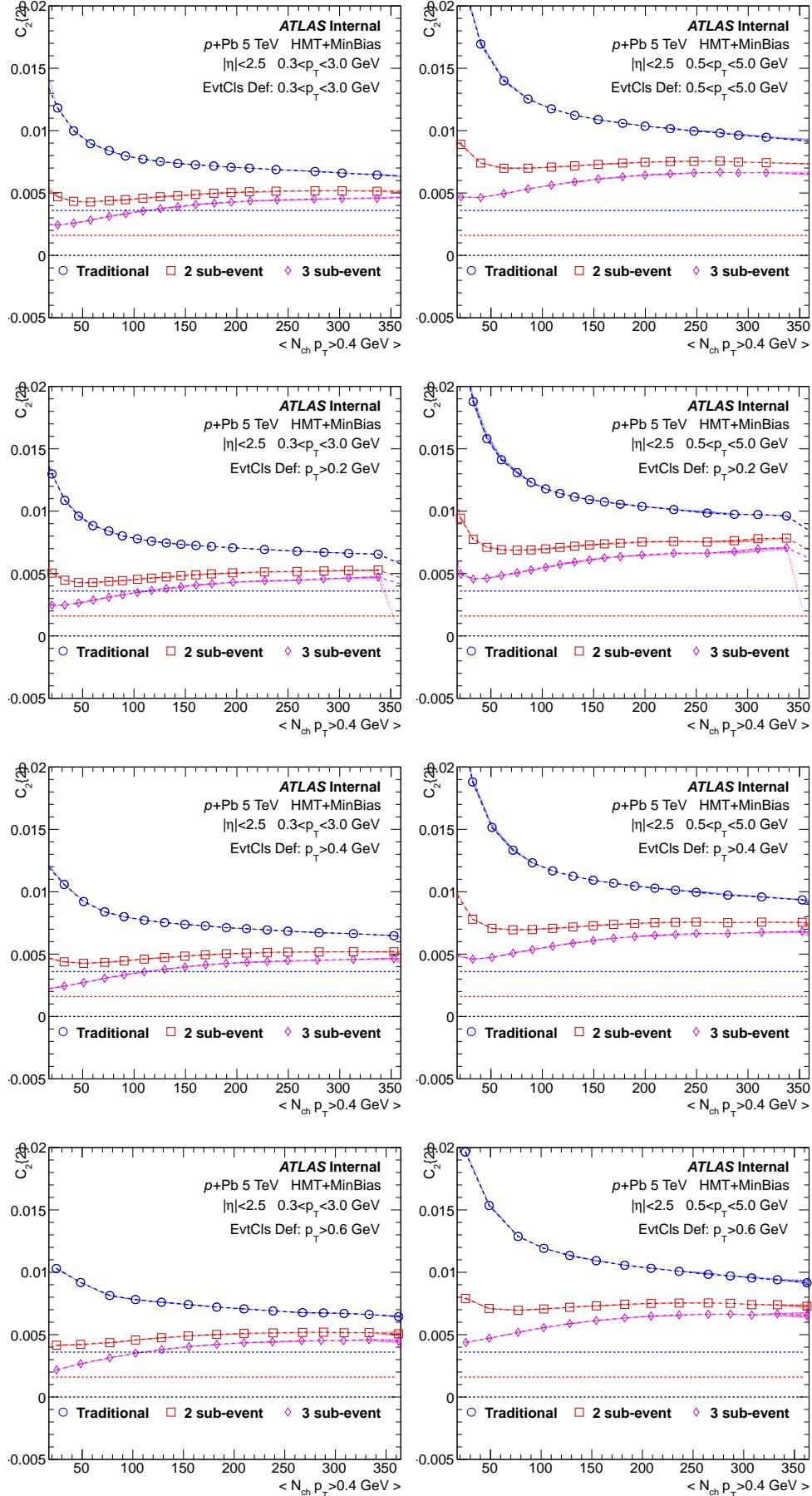
1214 **6.8 5.02 TeV pp $C_3\{4\}$**

1215 4-particle cumulant results of v_3 harmonic from 5.02 TeV pp are summarized in Fig. 77. Four rows have
1216 different event class definitions and two columns are particles with different p_T ranges. In each panel,
1217 $C_3\{4\}$ calculated using three cumulant methods are compared. Red dash line represents 4% v_2 signal
1218 while blue dash line represents 6% v_2 signal. Like the $C_2\{4\}$, non-flow contaminates the results from
1219 traditional method, and it becomes even larger as p_T goes higher. Meanwhile, $C_3\{4\}$ from 2 sub-event
1220 and 3 sub-event methods are consistent with 0, which is partially due to that the mean value of v_3 is much
1221 smaller than v_2 , and fluctuation of v_3 makes it very hard to measure in small systems, using cumulant
1222 method. It is interesting to repeat the measurement in Pb+Pb to see whether non-zero $v_3\{4\}$ can be
1223 measured.

Figure 77: Comparison of $C_3\{4\}$ calculated with 3 cumulant methods, from 5.02 TeV pp .

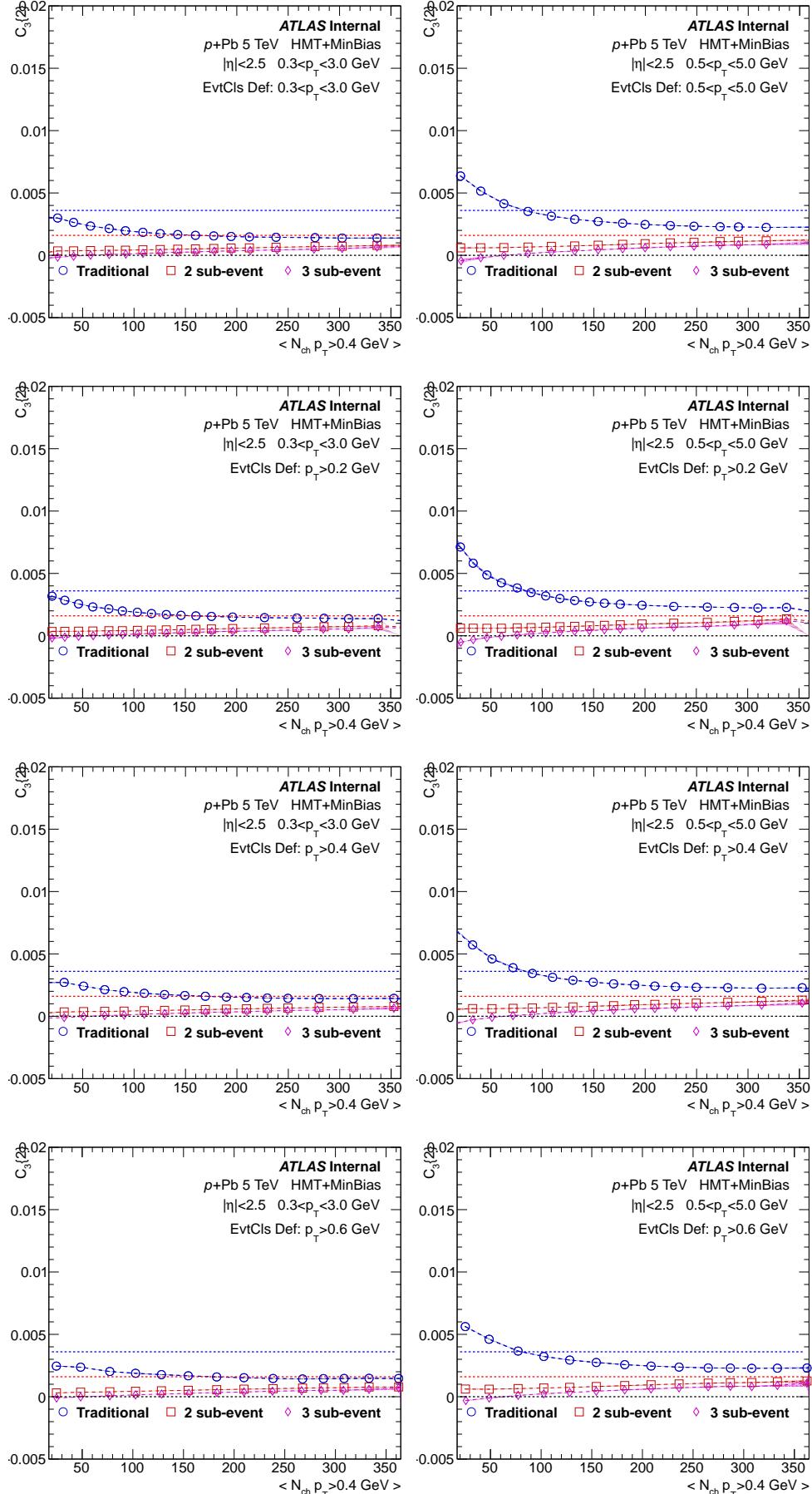
1224 **6.9 5.02 TeV $p+\text{Pb}$ $C_2\{2\}$**

1225 2-particle cumulant results of v_2 harmonic from 5.02 TeV $p+\text{Pb}$ are summarized in Fig. 78. Four rows
1226 have different event class definitions and two columns are particles with different p_T ranges. In each
1227 panel, $C_2\{2\}$ calculated using three cumulant methods are compared. In particular, for 2 sub-event
1228 method, two particles come from two sub-events, while for 3 sub-event method, two particles are sepa-
1229 rated by one sub-event in the mid- η . Red dash line represents 4% v_2 signal while blue dash line represents
1230 6% v_2 signal. Traditional method has largest $C_2\{2\}$, due to largest residual non-flow contribution. 2 sub-
1231 event cumulant already suppresses non-flow and gives smaller $C_2\{2\}$ values. $C_2\{2\}$ from 3 sub-event
1232 cumulant is the smallest since most short-range non-flow correlations are removed with the η gap. $C_2\{2\}$
1233 from all methods increase moving from lower p_T range to higher p_T range, which is consistent with 2PC
1234 results using template fit method. The 2-particle cumulant results are not sensitive to the event class
1235 definition.

Figure 78: Comparison of $C_2\{2\}$ calculated with 3 cumulant methods, from 5.02 TeV $p+Pb$.

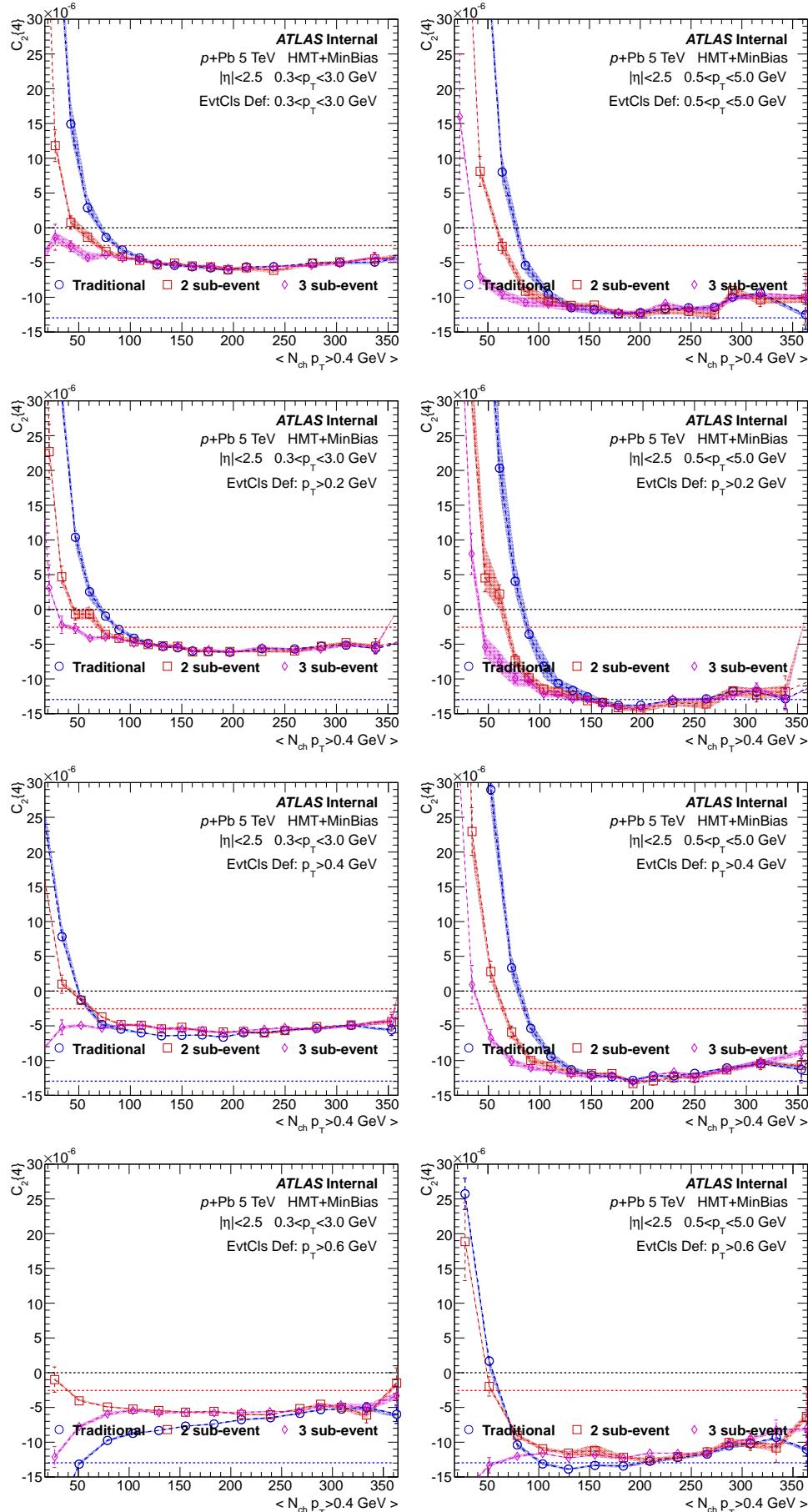
1236 **6.10 5.02 TeV $p+\text{Pb}$ $C_3\{2\}$**

1237 2-particle cumulant results of v_3 harmonic from 5.02 TeV $p+\text{Pb}$ are summarized in Fig. 79. Four rows
1238 have different event class definitions and two columns are particles with different p_{T} ranges. In each
1239 panel, $C_3\{2\}$ calculated using three cumulant methods are compared. In particular, for 2 sub-event
1240 method, two particles come from two sub-events, while for 3 sub-event method, two particles are sep-
1241 arated by one sub-event in the mid- η . Red dash line represents 4% v_3 signal while blue dash line rep-
1242 resents 6% v_3 signal. Traditional cumulant measures positive v_3 signal, and it increases as p_{T} moves to
1243 higher range. Meanwhile, $C_3\{2\}$ from 2 sub-event method is much smaller and $C_3\{2\}$ from 3 sub-event
1244 method is consistent with 0 with $0.3 < p_{\text{T}} < 3.0$ GeV, and it even goes to negative (wrong sign) in the
1245 low-multiplicity with $0.5 < p_{\text{T}} < 5.0$ GeV. Like the $C_2\{2\}$ results, $C_3\{2\}$ are not sensitive to the event
1246 class definition.

Figure 79: Comparison of $C_3\{2\}$ calculated with 3 cumulant methods, from 5.02 TeV $p+Pb$.

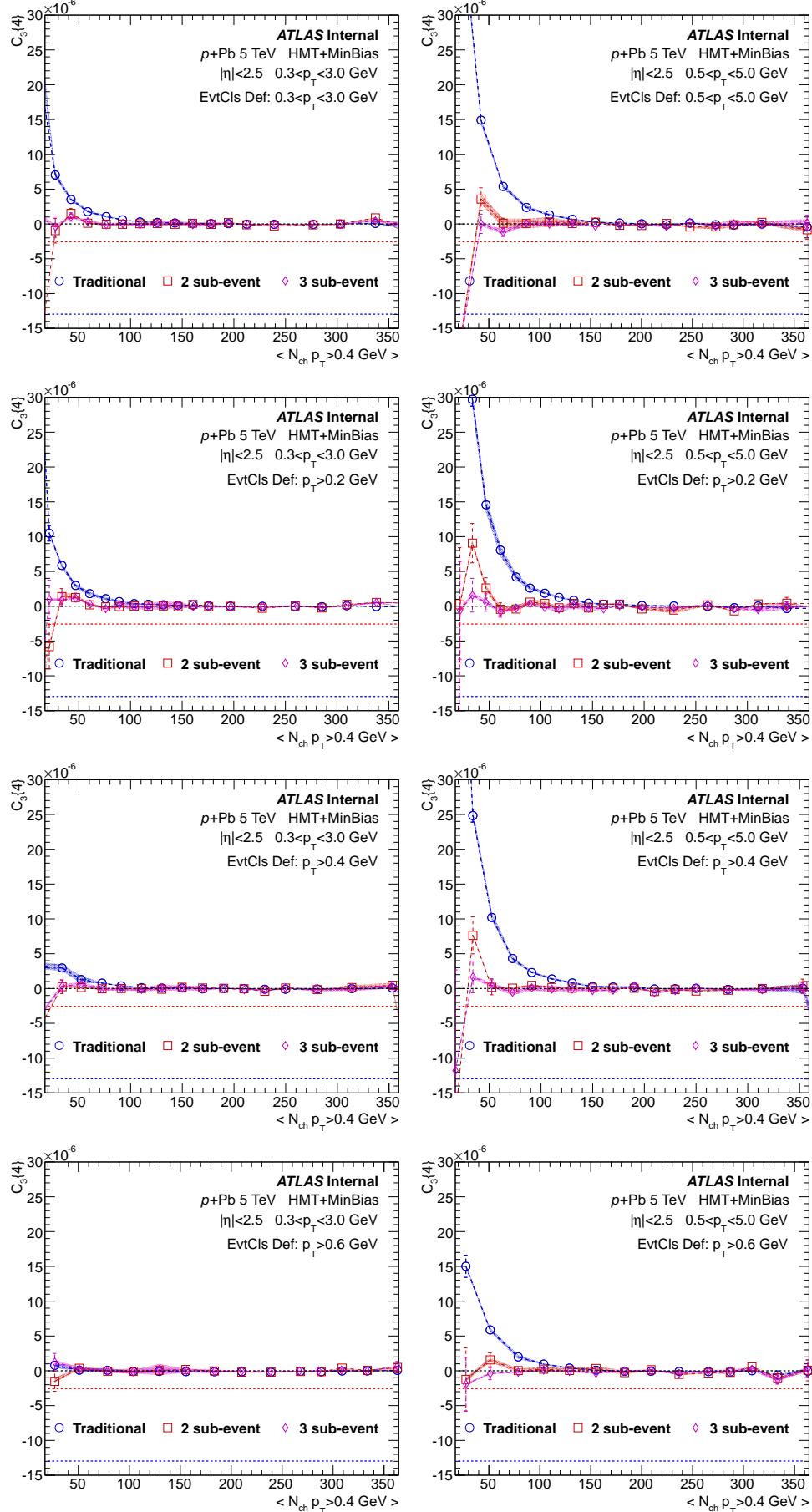
1247 6.11 5.02 TeV $p+\text{Pb}$ $C_2\{4\}$

1248 4-particle cumulant results of v_2 harmonic from 5 TeV $p+\text{Pb}$ are summarized in Fig. 80. Four rows
1249 have different event class definitions and two columns are particles with different p_T ranges. In each
1250 panel, $C_2\{4\}$ calculated using three cumulant methods are compared. Red dash line represents 4% v_2
1251 signal while blue dash line represents 6% v_2 signal. Three cumulant methods only differentiate at low
1252 $N_{ch} < 100$ region and in the following discussion we will only focus on this region. Compared with
1253 2-particle cumulant, 4-particle cumulant has much larger statistical errors thus fluctuation from point to
1254 point is expected. For a well-defined $v_2\{4\}$, $C_2\{4\}$ has to be negative. However, traditional cumulant
1255 always gives the positive $C_2\{4\}$ with default event class definition. This means the residual non-flow can
1256 even change the sign of $C_2\{4\}$. Meanwhile, $C_2\{4\}$ from 2 sub-event method is already much suppressed
1257 and $C_2\{4\}$ from 3 sub-event method stays negative in most of N_{ch} ranges, except in the lowest N_{ch}
1258 region. As moved from $0.3 < p_T < 3.0$ GeV to $0.5 < p_T < 5.0$ GeV, because of larger fraction of non-
1259 flow, both traditional and 2 sub-event cumulant tend to give the wrong sign of $C_2\{4\}$. Only $C_2\{4\}$ from
1260 3 sub-event becomes more negative as p_T goes higher, which is consistent with the existing 2PC results.
1261 This provides another evidence that only 3 sub-event cumulant can effectively suppress the non-flow and
1262 give reasonable v_2 measurement. Last but not least, both traditional and 2 sub-event cumulant results are
1263 sensitive to the event class definition, while 3 sub-event method gives rather stable results.

Figure 80: Comparison of $C_2\{4\}$ calculated with 3 cumulant methods, from 5.02 TeV $p+Pb$.

1264 **6.12 5.02 TeV $p+\text{Pb}$ $C_3\{4\}$**

1265 4-particle cumulant results of v_3 harmonic from 5.02 TeV $p+\text{Pb}$ are summarized in Fig. 81. Four rows
1266 have different event class definitions and two columns are particles with different p_{T} ranges. In each
1267 panel, $C_3\{4\}$ calculated using three cumulant methods are compared. Red dash line represents 4% v_2
1268 signal while blue dash line represents 6% v_2 signal. Like the $C_2\{4\}$, non-flow contaminates the results
1269 from traditional method, and it becomes even larger as p_{T} goes higher. Meanwhile, $C_3\{4\}$ from 2 sub-
1270 event and 3 sub-event methods are consistent with 0, which is partially due to that the mean value of
1271 v_3 is much smaller than v_2 , and fluctuation of v_3 makes it very hard to measure in small systems, using
1272 cumulant method. It is interesting to repeat the measurement in Pb+Pb to see whether non-zero $v_3\{4\}$
1273 can be measured.

Figure 81: Comparison of $C_3\{4\}$ calculated with 3 cumulant methods, from 5.02 TeV $p+Pb$.

¹²⁷⁴ **6.13 Comparison among three collision systems**

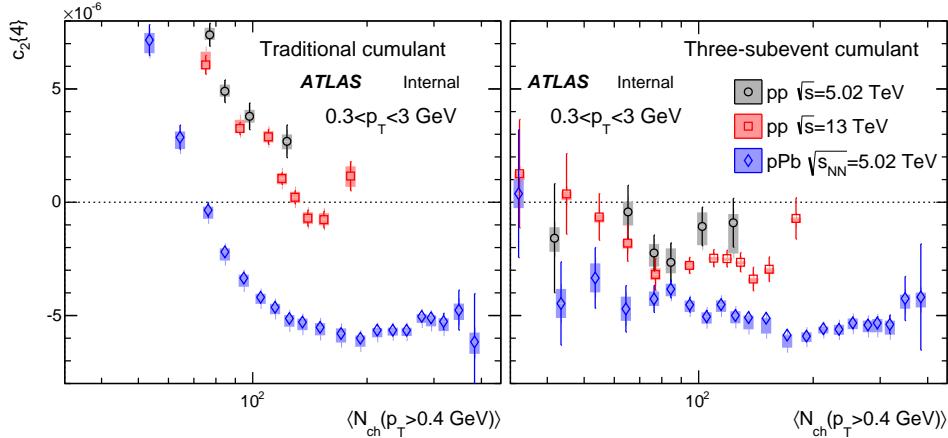


Figure 82: The $C_2\{4\}$ calculated for charged particles in $0.3 < p_T < 3.0$ GeV using the traditional cumulant (left panel) and 3 sub-event method (right panel) compared between 5.02 TeV pp , 13 TeV pp and 5.02 TeV $p+Pb$. The event averaging is performed for N_{ch} calculated for the same p_T range, which is then mapped to $\langle N_{ch} \rangle$, the average number of charged particles with $p_T > 0.4$ GeV.

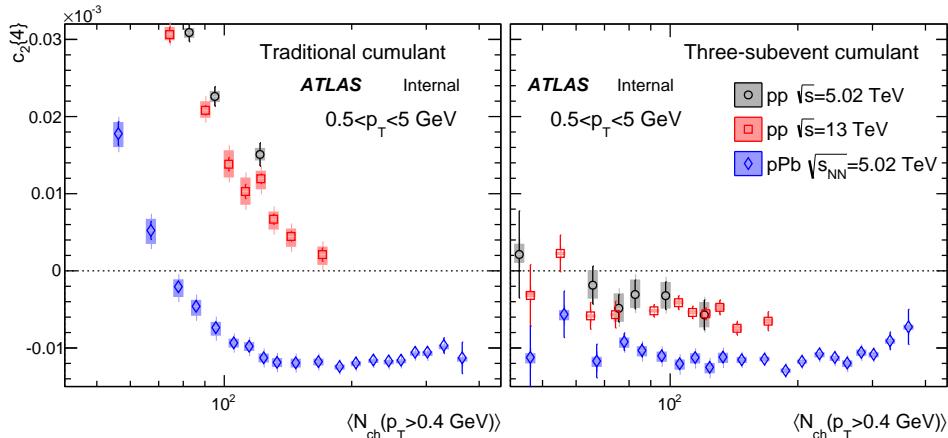


Figure 83: The $C_2\{4\}$ calculated for charged particles in $0.5 < p_T < 5.0$ GeV using the traditional cumulant (left panel) and 3 sub-event method (right panel) compared between 5.02 TeV pp , 13 TeV pp and 5.02 TeV $p+Pb$. The event averaging is performed for N_{ch} calculated for the same p_T range, which is then mapped to $\langle N_{ch} \rangle$, the average number of charged particles with $p_T > 0.4$ GeV.

¹²⁷⁵ Fig. 82 and Fig. 83 compare the $C_2\{4\}$ values among three collision systems obtained with the traditional
¹²⁷⁶ method and 3 sub-event method. The large positive $C_2\{4\}$ values observed in small $\langle N_{ch} \rangle$ region in the
¹²⁷⁷ traditional method is likely due to non-flow correlations, which is absent in the 3 sub-event cumulant
¹²⁷⁸ method. In $p+Pb$ collisions, the magnitude of $C_2\{4\}$ seems to decrease slightly for $\langle N_{ch} \rangle > 200$ region.

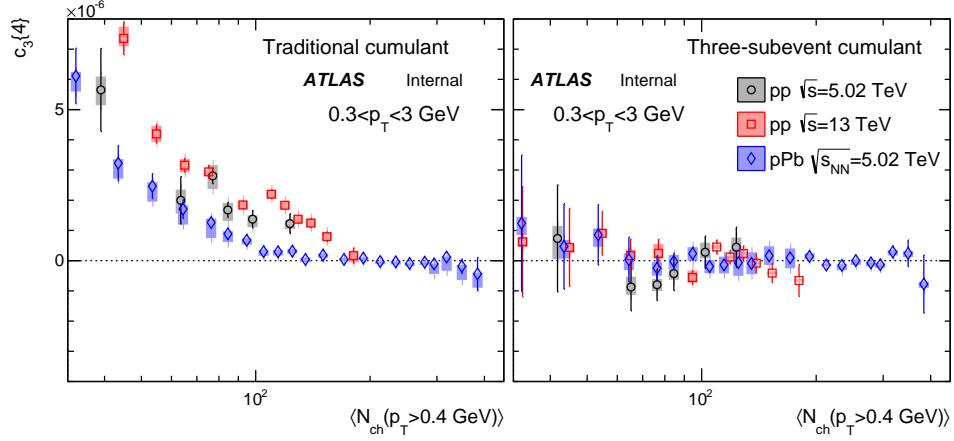


Figure 84: The $C_3\{4\}$ calculated for charged particles in $0.3 < p_T < 3.0$ GeV using the traditional cumulant (left panel) and 3 sub-event method (right panel) compared between 5.02 TeV pp , 13 TeV pp and 5.02 TeV $p+Pb$. The event averaging is performed for N_{ch} calculated for the same p_T range, which is then mapped to $\langle N_{ch} \rangle$, the average number of charged particles with $p_T > 0.4$ GeV.

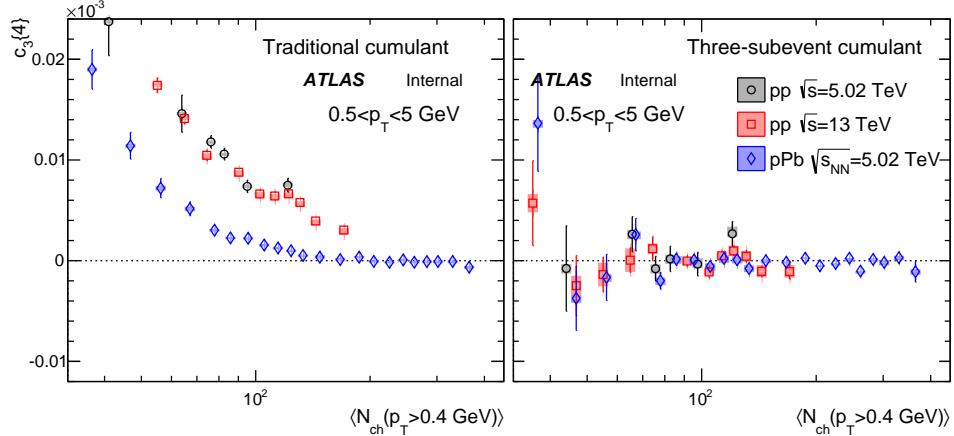


Figure 85: The $C_3\{4\}$ calculated for charged particles in $0.5 < p_T < 5.0$ GeV using the traditional cumulant (left panel) and 3 sub-event method (right panel) compared between 5.02 TeV pp , 13 TeV pp and 5.02 TeV $p+Pb$. The event averaging is performed for N_{ch} calculated for the same p_T range, which is then mapped to $\langle N_{ch} \rangle$, the average number of charged particles with $p_T > 0.4$ GeV.

Fig. 84 and Fig. 85 compare the $C_3\{4\}$ values among three collision systems for the traditional cumulant method and 3 sub-event method. The positive $C_3\{4\}$ values in small $\langle N_{ch} \rangle$ region in the traditional method is indicative of the non-flow correlations, with a smaller magnitude in comparison to $C_2\{4\}$. In contrast, the $C_3\{4\}$ values from the 3 sub-event method are consistent with zero in all three systems.

6.14 Comparison with 2PC results

The $v_2\{4\}$ results are compared to the $v_2\{2\}$ obtained from a two-particle correlation analysis [11, 15] where the non-flow from dijets is estimated using low-multiplicity events ($\langle N_{ch} \rangle < 10$) and then subtracted. The subtraction was done by either including or not including the pedestal in the low multiplicity

events (labeled as "template fit" and "peripheral subtraction" respectively), where the pedestal is determined by a zero-yield at minimum (ZYAM) procedure [1]. Not including the pedestal in low-multiplicity events in the subtraction was shown to significantly reduces the measured v_2 value since it explicitly assumes no long-range v_2 in the peripheral bin and therefore forcing the v_2 to be zero as $\langle N_{ch} \rangle$ approaches that of the low-multiplicity events.

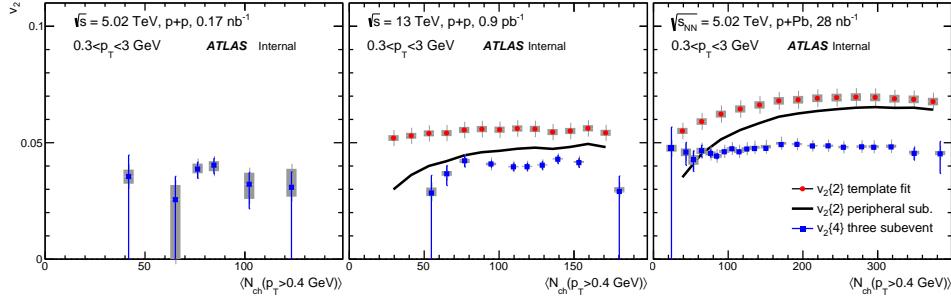


Figure 86: The $v_2\{4\}$ calculated for charged particles in $0.3 < p_T < 3.0$ GeV using the three-subevent method in 5.02 TeV pp (left panel), 13 TeV pp (middle panel) and 5.02 TeV $p+Pb$ collisions (right panel). They are compared to $v_2\{2\}$ obtained from a two-particle correlation analysis where the non-flow effect is removed by template fit procedure without ZYAM assumption (solid circles) and with ZYAM assumption (solid line).

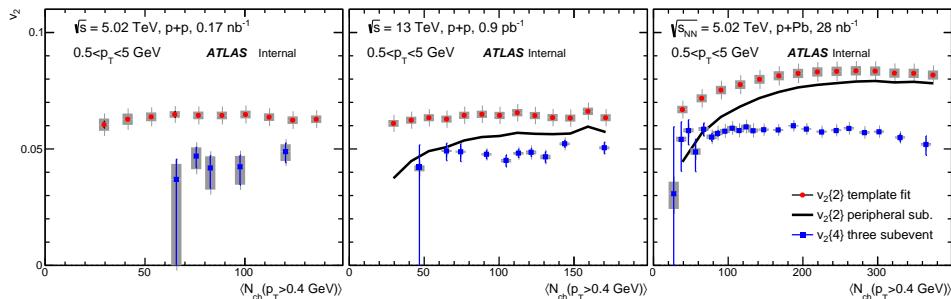


Figure 87: The $v_2\{4\}$ calculated for charged particles in $0.5 < p_T < 5.0$ GeV using the three-subevent method in 5.02 TeV pp (left panel), 13 TeV pp (middle panel) and 5.02 TeV $p+Pb$ collisions (right panel). They are compared to $v_2\{2\}$ obtained from a two-particle correlation analysis where the non-flow effect is removed by template fit procedure without ZYAM assumption (solid circles) and with ZYAM assumption (solid line).

Fig. 86 and Fig. 87 show that the $v_2\{4\}$ are below $v_2\{2\}$ from the template-fit method in both pp and $p+Pb$ collisions. This is not surprising since cumulant method also measures the fluctuation of flow signal, which results in a lower value. In other words, these results mean that the peripheral subtraction method underestimates the v_2 value since it assumes no long-range v_2 in the peripheral bin.

6.15 Number of sources N_s in the initial state geometry

In hydrodynamic models for small collision systems, this difference between $c_2\{4\}$ and $c_2\{2\}$ from 2-particle correlation has been interpreted as influence of event-by-event flow fluctuations associated with fluctuating initial condition, which is closely related to the effective number of sources N_s in the

¹³⁰¹ transverse density distribution of the initial state [36, 37].

$$\frac{v_2\{4\}}{v_2\{2\}} = \left(\frac{4}{3 + N_s}\right)^{1/4}$$

$$N_s = 4 \frac{v_2\{2\}^4}{v_2\{4\}^4} - 3 \quad (21)$$

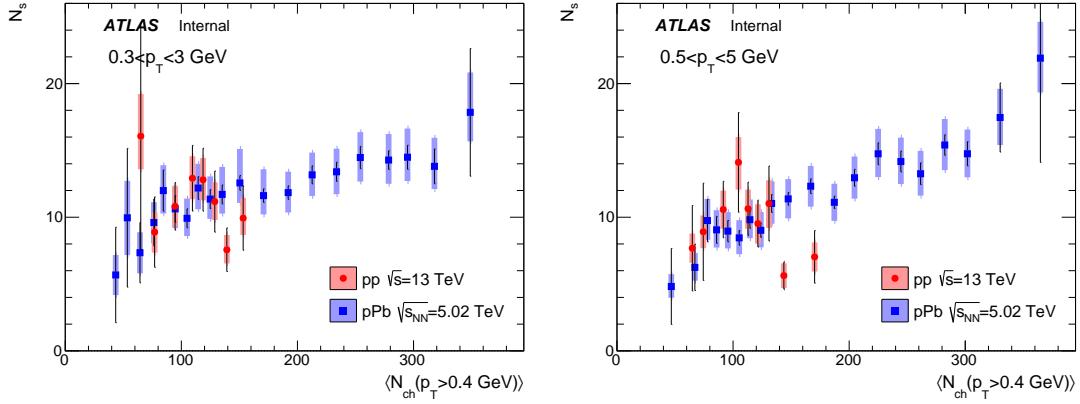


Figure 88: The number of sources inferred from $v_2\{2\}$ and $v_2\{4\}$ via a model-dependent equation in 13 TeV pp and 5.02 TeV $p+Pb$ collisions. The correlation is performed in $0.3 < p_T < 3.0$ GeV (left panel) and $0.5 < p_T < 5.0$ GeV (right panel).

¹³⁰² Fig. 88 shows the N_s values as a function $\langle N_{ch} \rangle$ in 13 TeV pp and 5.02 $p+Pb$ collisions, estimated
¹³⁰³ with changed particles in $0.3 < p_T < 3.0$ GeV and $0.5 < p_T < 5.0$ GeV ranges. The number of sources
¹³⁰⁴ increases with $\langle N_{ch} \rangle$ in $p+Pb$ collisions up to $N_s \sim 20$ in the highest multiplicity class. As N_s becomes
¹³⁰⁵ large, the flow fluctuation is expected to approach Gaussian, and $|C_2\{4\}|$ or $v_2\{4\}$ is expected to decrease.
¹³⁰⁶ This is consistent with the slight decrease of $C_2\{4\}$ shown for $\langle N_{ch} \rangle > 200$. The results for 13 TeV pp
¹³⁰⁷ collisions has large uncertainties and cover a much limited $\langle N_{ch} \rangle$ range, but are approximately consistent
¹³⁰⁸ with $p+Pb$ collisions at comparable $\langle N_{ch} \rangle$ value, which is expected if the initial eccentricity is driven by
¹³⁰⁹ similar underlying physics.

1310 7 Summary

1311 The measurement of four-particle cumulant elliptic flow coefficients $C_2\{4\}$ for charged particles are
 1312 presented in 0.17 pb^{-1} pp data at $\sqrt{s} = 5.02 \text{ TeV}$, 0.9 pb^{-1} pp data at $\sqrt{s} = 13 \text{ TeV}$ and 28 nb^{-1} $p+\text{Pb}$ at
 1313 $\sqrt{s_{NN}} = 5.02 \text{ TeV}$. The $C_2\{4\}$ are calculated using traditional cumulant, as well as the recently proposed
 1314 two-subevent method and three-subevent method, they are presented as a function of $\langle N_{ch} \rangle$, the average
 1315 number of charged particles with $p_T > 0.4 \text{ GeV}$. It is found that the $C_2\{4\}$ from traditional method is
 1316 sensitive to the choice of event classes used for averaging, such sensitivity is greatly reduced in the two-
 1317 subevent method and nearly removed in the three-subevent method, suggesting that the three-subevent
 1318 method is much less affected by the non-flow effects. Negative $C_2\{4\}$ is obtained in all three collision
 1319 systems in a broad range of $\langle N_{ch} \rangle$ using the three-subevent method. The magnitude of the $C_2\{4\}$ is
 1320 nearly independent of $\langle N_{ch} \rangle$ but a slight decrease is observed in $p+\text{Pb}$ collisions in the large multiplicity
 1321 region. The single-particle harmonic coefficient $v_2\{4\} = (-C_2\{4\})^{1/4}$ is calculated and compared with
 1322 v_2 obtained previously using a two-particle correlation method, where the non-flow effects was estimated
 1323 and subtracted. The $v_2\{4\}$ is smaller than $v_2\{2\}$ as expected for long-range collective ridge, and the
 1324 difference between $v_2\{4\}$ and $v_2\{2\}$ are used to infer the number of sources N_s in the initial state collision
 1325 geometry. The N_s is found to increase with $\langle N_{ch} \rangle$ in $p+\text{Pb}$ collisions and reaches around 20 in highest
 1326 multiplicity events.

¹³²⁷ **8 Appendix**

¹³²⁸ **8.1 Statistics of all triggers in pp 13 TeV**

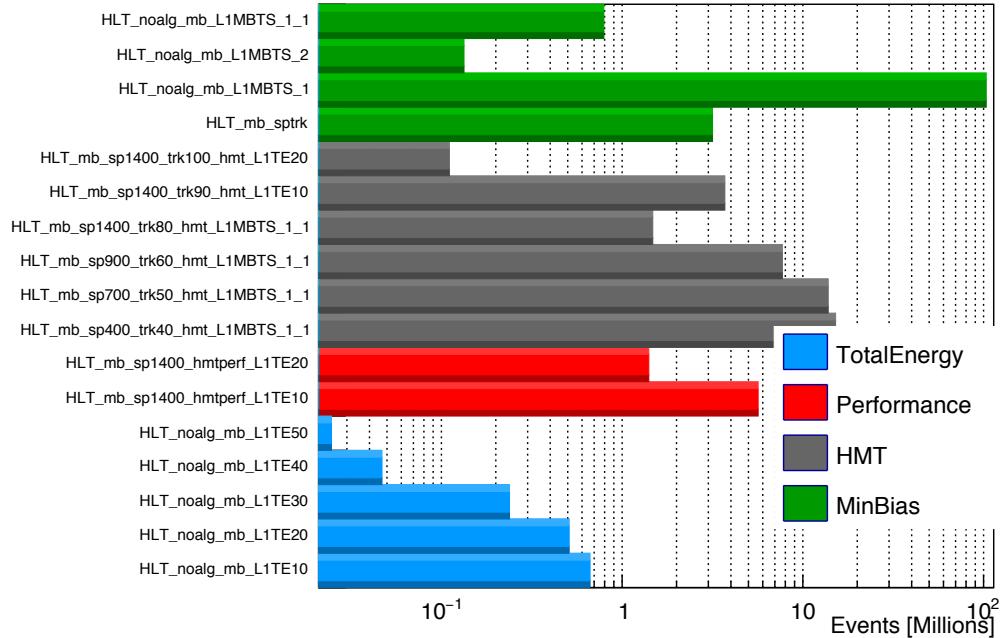


Figure 89: Total statistics of all the MinBias and HMT related triggers, recorded in 13 TeV pp run period 1.

Run : 299390

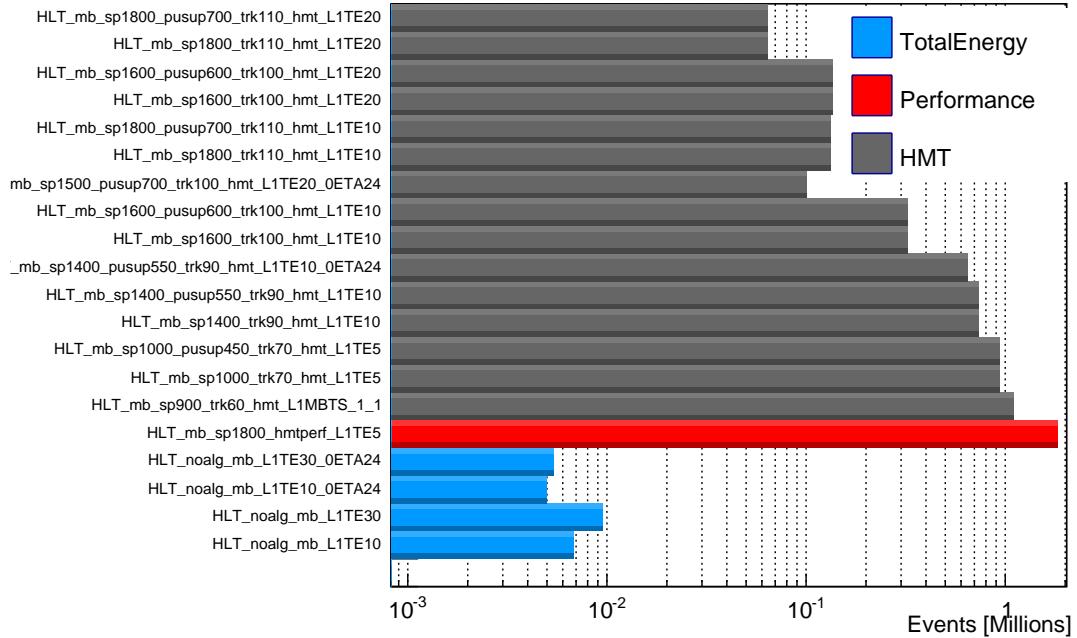


Figure 90: Total statistics of all the MinBias and HMT related triggers, recorded in 13 TeV pp run period 2 (299390).

Run : 300287

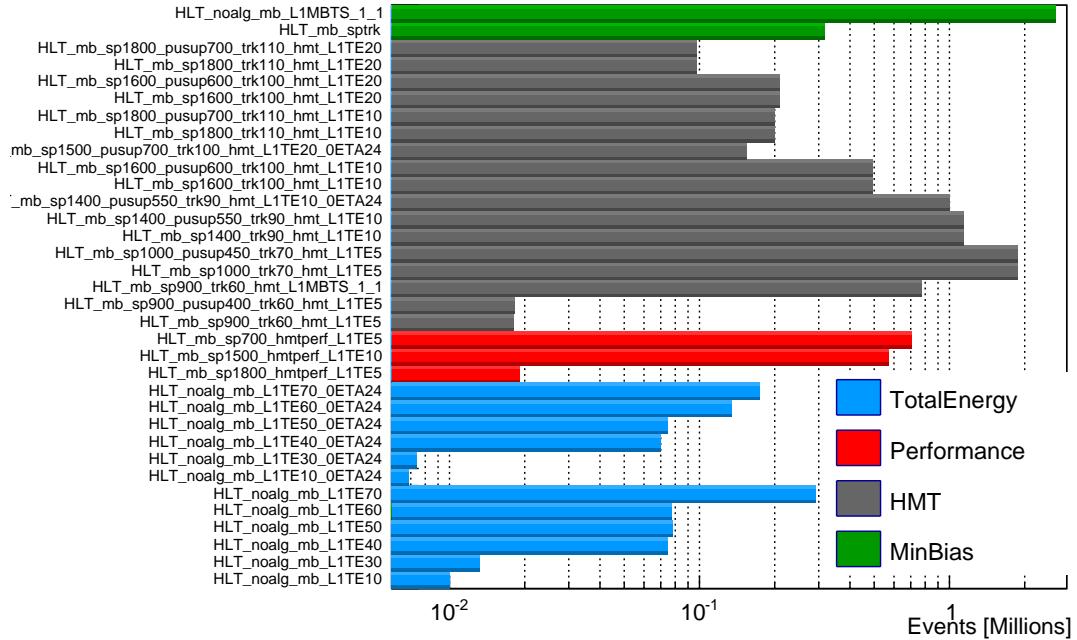


Figure 91: Total statistics of all the MinBias and HMT related triggers, recorded in 13 TeV pp run period 2 (300287).

Run : 305359

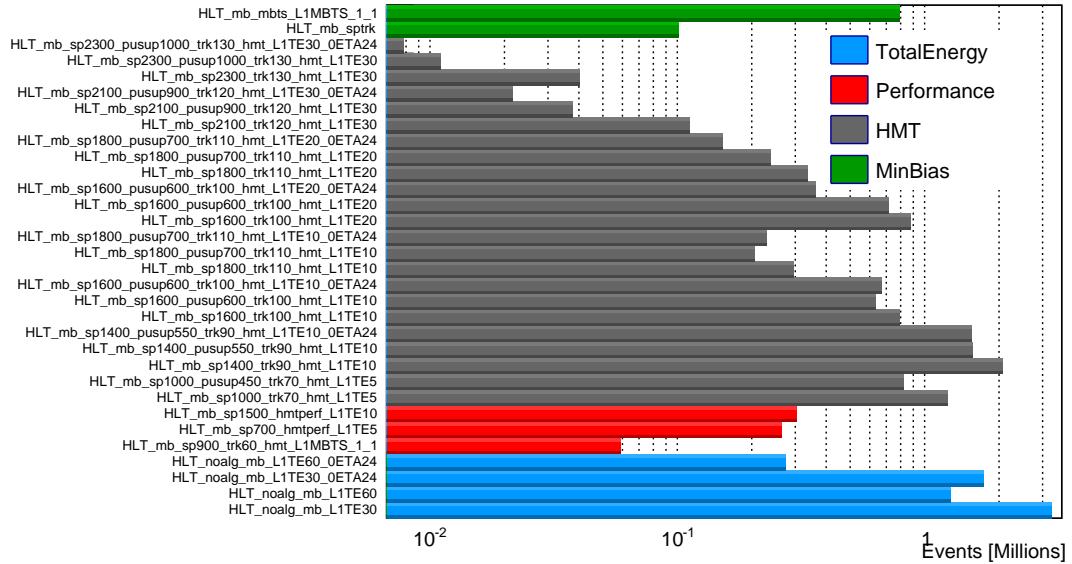


Figure 92: Total statistics of all the MinBias and HMT related triggers, recorded in 13 TeV pp run period 3.

Run : 309314

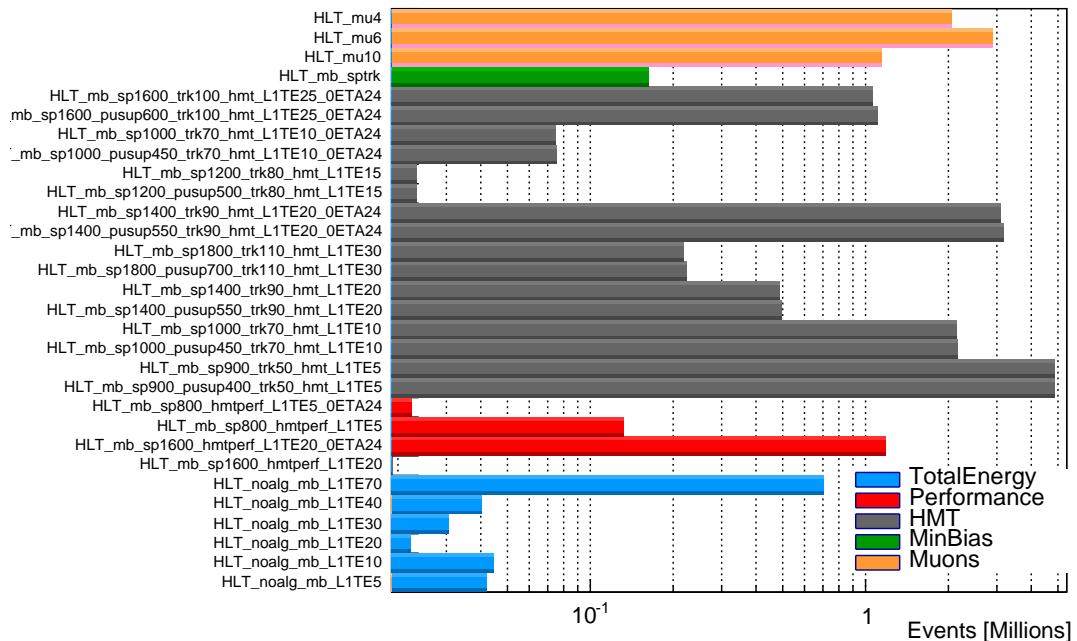


Figure 93: Total statistics of all the MinBias and HMT related triggers, recorded in 13 TeV pp run period 4 (309314).

Run : 309346

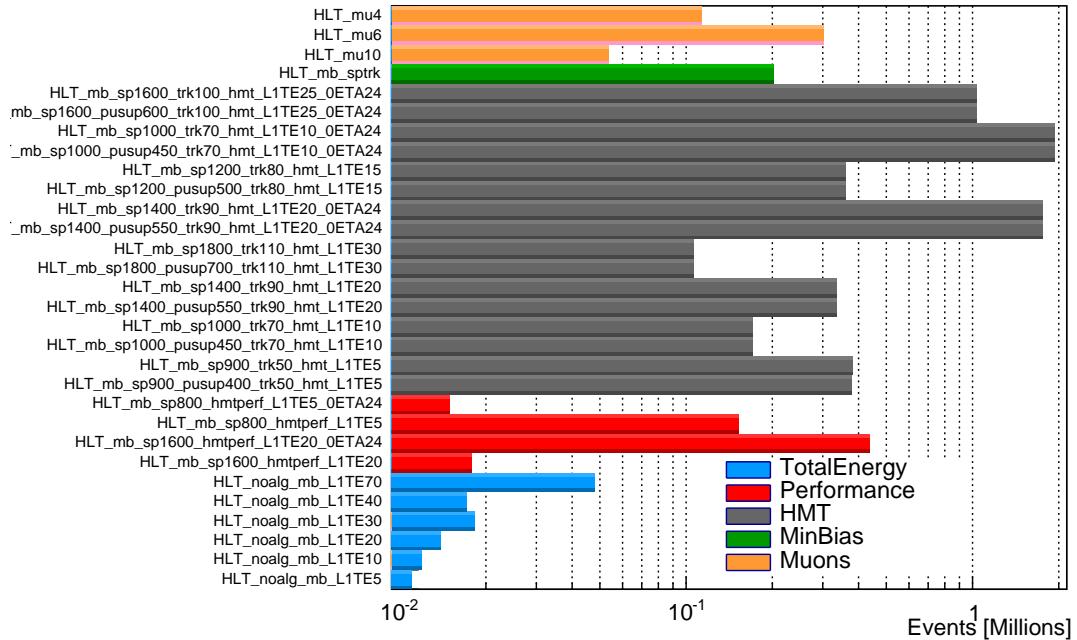


Figure 94: Total statistics of all the MinBias and HMT related triggers, recorded in 13 TeV pp run period 4 (309346).

Run : 310216

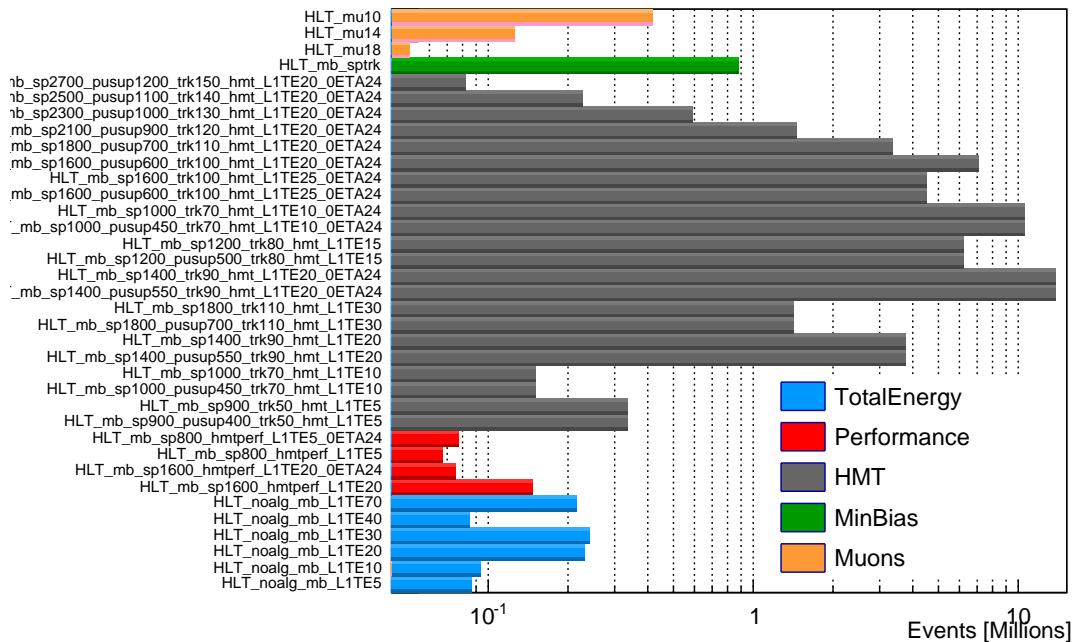


Figure 95: Total statistics of all the MinBias and HMT related triggers, recorded in 13 TeV pp run period 4 (310216).

1329 8.2 Statistics of all triggers in pp 5.02 TeV

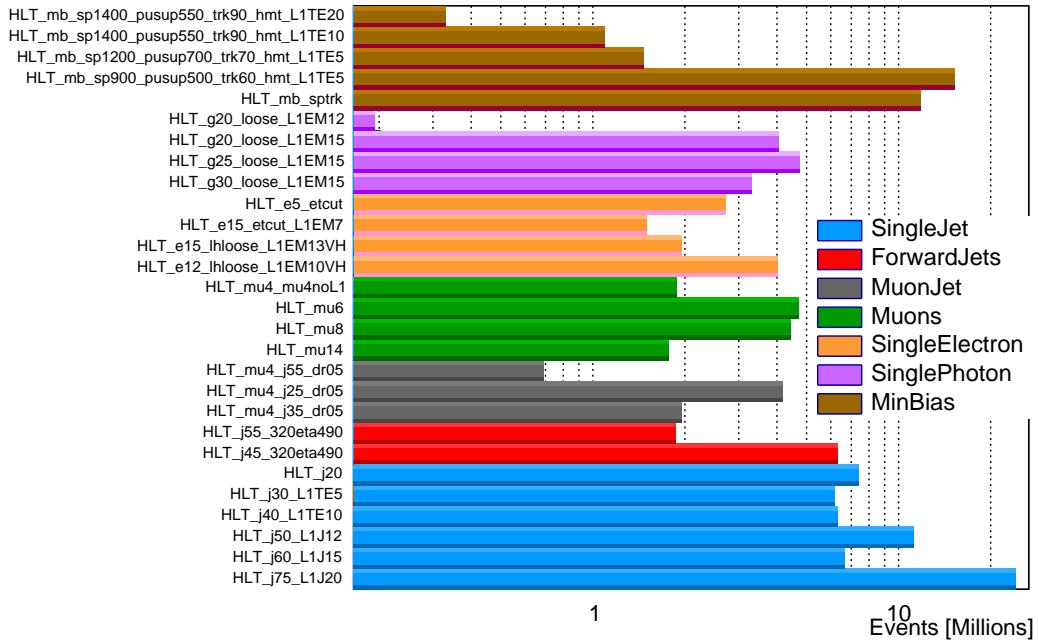
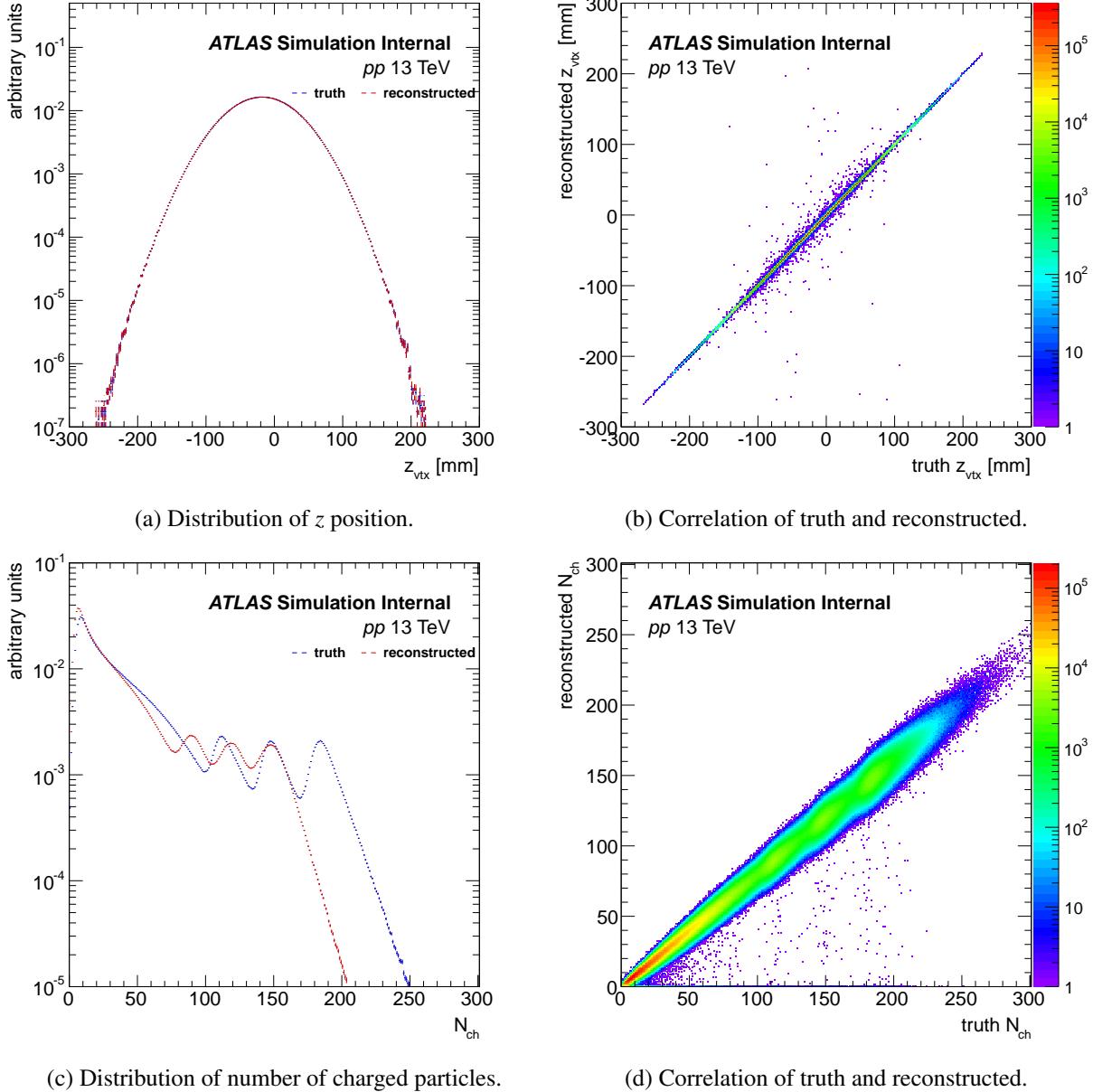


Figure 96: Total statistics of all the MinBias and HMT related triggers, recorded in 5 TeV pp reference run.

1330 8.3 Auxiliary plots 13 TeV PYTHIA

1331 Fig. 97 shows the performance of event-level quantities in this PYTHIA sample, both for truth and
 1332 reconstructed. For the z position, the mean value is shifted by $-10mm$ and the distribution is roughly
 1333 Gaussian. The correlation between z positions of truth and reconstructed vertex are very strong. The
 1334 number of charged particles is determined by counting number of tracking passing the tracking selection
 1335 criteria, which will be introduced later. The 3 bumps shown in the distribution represents the 3 million
 1336 high-multiplicity events. The number of reconstructed tracks is smaller than truth due to the efficiency
 1337 loss. The correlation between the number of truth and reconstructed tracks is also strong, as shown in
 1338 the right plot.

Figure 97: Performance of event-level quantities in PYTHIA 13 TeV pp .

1339 The results of particle-level qualities are summarized in Fig. 99, where we compared the three basic
 1340 quantities of tracking: p_T , η and ϕ . The p_T spectrum is very consistent between truth and reconstructed,
 1341 while the correlation shows some broad structure in the low- p_T region. The distributions of η are very
 1342 different between truth and reconstructed, due to the different efficiency in different η region. In this
 1343 analysis, we will evaluate and apply the tracking efficiency as a function of η . Based on the correlation
 1344 plot, the reconstructed and associated truth tracks have consistent η value. For ϕ of the charged particles,
 1345 the reconstructed has additional small structures over truth. We will not apply this efficiency correction
 1346 in the analysis, instead, we will apply flattening procedure run-by-run. The truth and reconstructed ϕ
 1347 are strongly correlated: the events seen in the two corners are due to the periodic boundary condition.

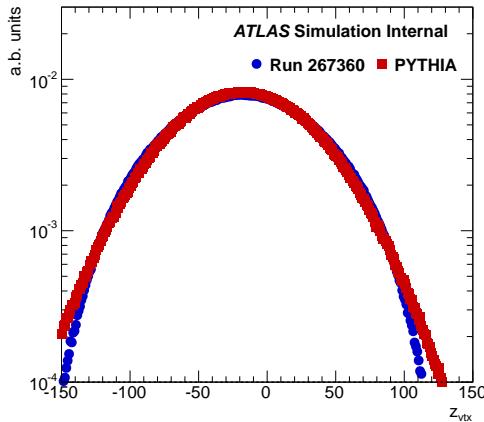
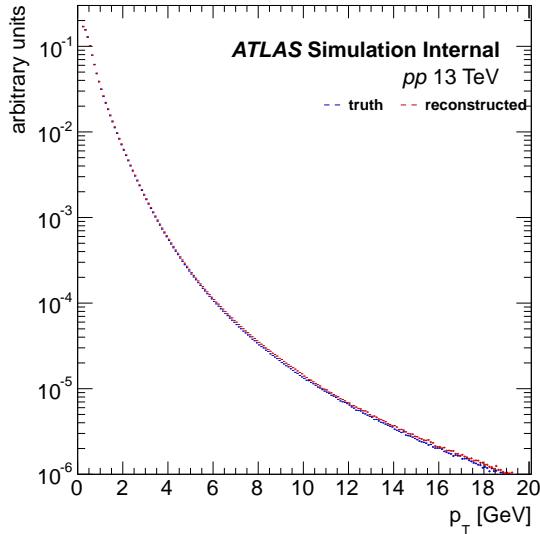
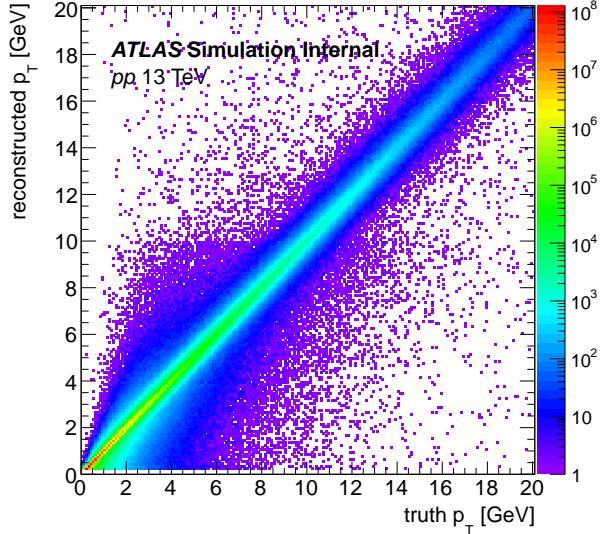
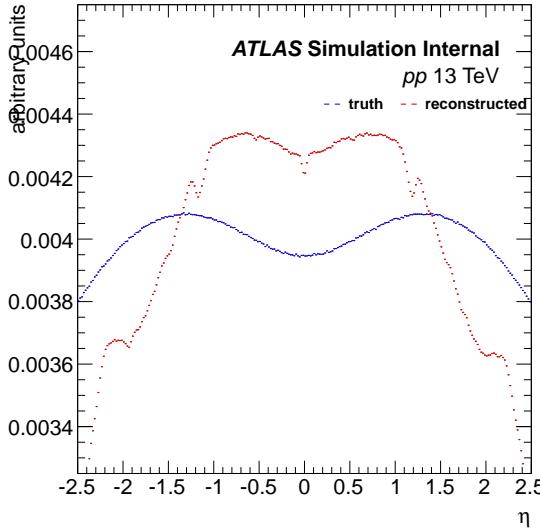
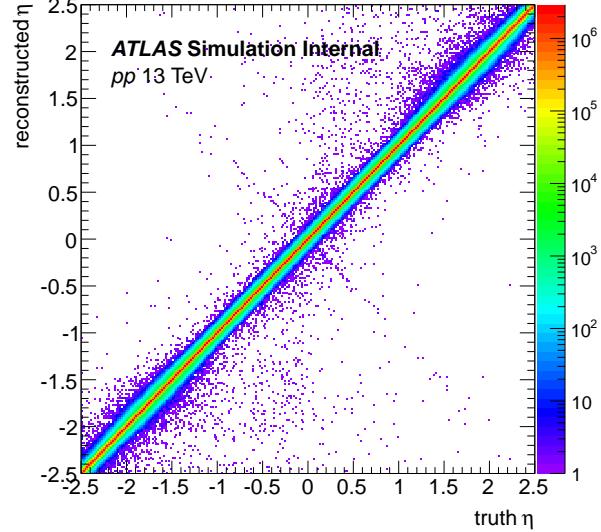


Figure 98: The distributions of z_{vertex} , from Run 267360 and PYTHIA.

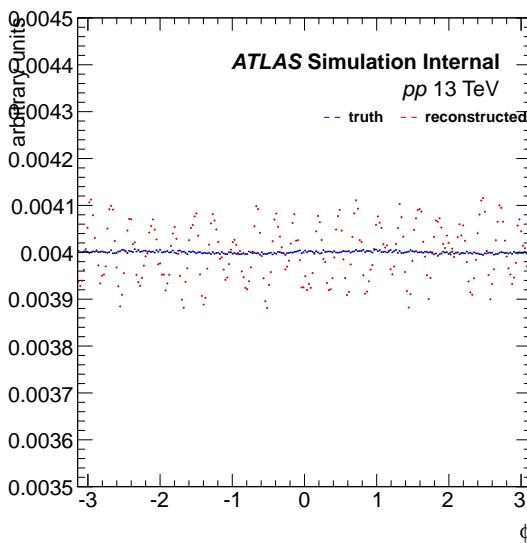
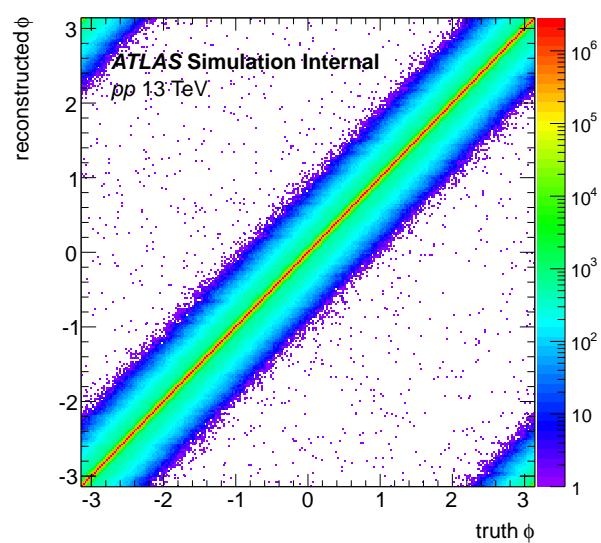
1348 The distribution of z_{vertex} is compared between Run 267360 and PYTHIA, as shown in Fig. 98. The
 1349 width of z_{vertex} distribution in PYTHIA is slightly larger than data.

(a) p_T spectrum.

(b) Correlation of truth and reconstructed.

(c) Distribution of η .

(d) Correlation of truth and reconstructed.

(e) Distribution of ϕ .

(f) Correlation of truth and reconstructed.

Figure 99: Performance of particle-level quantities in PYTHIA 13 TeV pp .

1350 In this analysis we also included the PYTHIA truth results as a comparison with data. In order to
 1351 include more statistics, the min-bias data sets used are on the EVGEN level:

- 1352 • 200 million non-diffractive events

1353 `mc15_13TeV.361203.Pythia8_A2_MSTW2008LO_ND_minbias.evgen.EVNT.e3639/`

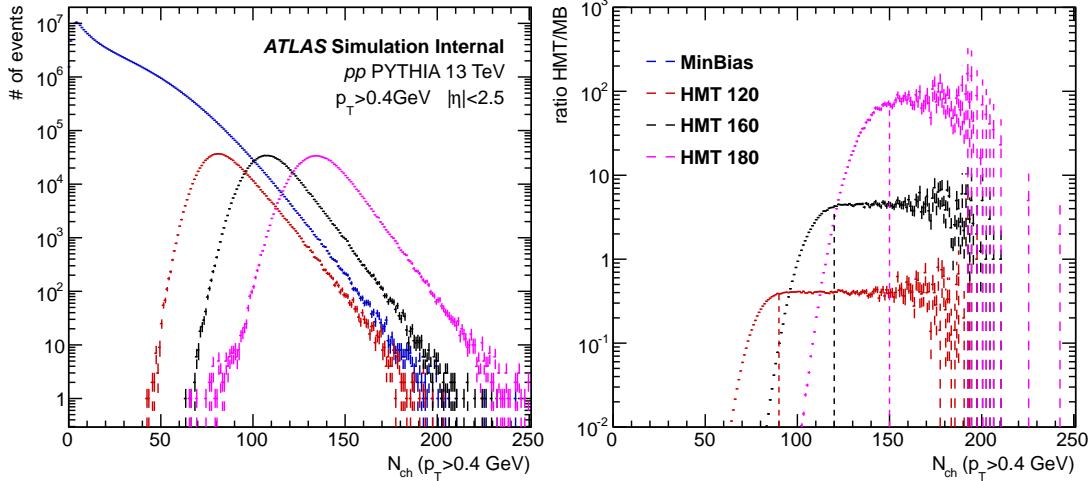


Figure 100: Left plot shows the N_{ch} distribution of MinBias and HMT events separately. Right plot shows the ratio of HMT events and MinBias events. The dash lines in the right plot indicate the minimum N_{ch} cuts while including HMT events, in order to not introduce additional bias.

1354 As to the HMT events, EVGEN provides exactly the same statistics as reconstructed sample. In pp
 1355 data, the high-multiplicity events were selected by using online algorithms, where most of the bias come
 1356 from the difference between online and offline tracking reconstruction. While in PYTHIA, since it is
 1357 not clear how the high-multiplicity events are selected, in order not to introduce additional bias, we only
 1358 select the HMT events where "trigger efficiency" is 100%, as indicated in Fig. 100. For three HMT data
 1359 sets, the N_{ch} cuts are determined as: 90, 120 and 150.

1360 8.4 Formulas for direct cumulant with particle weight

1361 8.4.1 Traditional method

1362 The idea of direct cumulant is to calculate the multi-particle correlation in a single loop. However, since
 1363 the multi-particle correlation only counts the unique pairs, some duplicated terms need to be subtracted
 1364 out.

1365 We start with 2-particle correlation, to calculate $W_{(2)}$:

$$(\sum w_i)^2 = \sum w_i w_j + \sum w_i^2 \quad (22)$$

1366 where in the following equations, all summation symbols mean summing the unique pairs for simplicity.
 1367 We also introduce the notation S_p^k for better book keeping:

$$S_p^k \equiv (\sum w_i^p)^k \quad (23)$$

¹³⁶⁸ So the goal of direct cumulant is to express the 2- and 4-particle correlation as a function of Q_n, k
¹³⁶⁹ and S_p^k , so that they could be calculated in a single loop.

$$\begin{aligned} corr_n\{2\} &\rightarrow f(Q_{n,k}, S_p^k) \\ corr_n\{4\} &\rightarrow g(Q_{n,k}, S_p^k) \end{aligned} \quad (24)$$

¹³⁷⁰ In the same way, in order to calculate $\sum w_i w_j e^{in(\phi_i - \phi_j)}$, we expand:

$$\begin{aligned} (\sum w_i e^{in\phi_i})^2 &= \sum w_i^2 + \sum w_i w_j e^{in(\phi_i - \phi_j)} \\ \Rightarrow |Q_{n,1}|^2 &= S_2^1 + \sum w_i w_j e^{in(\phi_i - \phi_j)} \end{aligned} \quad (25)$$

¹³⁷¹ and plug in all the formulas above, we get the expression for $corr_n\{2\}$

$$\begin{aligned} W_{(2)} &= S_1^2 - S_2^1 \\ corr_n\{2\} &= \frac{|Q_{n,1}|^2 - S_2^1}{S_1^2 - S_2^1} \end{aligned} \quad (26)$$

¹³⁷² The formula for 4-particle correlations will be more complicated. We will start with $W_{(4)}$ by expand-
¹³⁷³ ing S_1^4 :

$$S_1^4 = \sum w_i w_j w_k w_l + 6 \sum w_i^2 w_j w_k + 3 \sum w_i^2 w_j^2 + 4 \sum w_i^3 w_j + S_4^1 \quad (27)$$

¹³⁷⁴ where in order to calculate the remaining non- S terms, proper terms also need to be expanded:

$$\begin{aligned} S_3^1 S_1^1 &= \sum w_i^3 w_j + S_4^1 \\ S_2^2 &= \sum w_i^2 w_j^2 + S_4^1 \\ S_2^1 S_1^2 &= \sum w_i^2 w_j w_k + 2 \sum w_i^3 w_j + \sum w_i^2 w_j^2 + S_4^1 \end{aligned} \quad (28)$$

¹³⁷⁵ By solving the linear equation, we get:

$$\sum w_i w_j w_k w_l = S_1^4 + 8S_3^1 S_1^1 - 6S_2^1 S_1^2 + 3S_2^2 - 6S_4^1 \quad (29)$$

¹³⁷⁶ In order to simply the writing, we define:

$$e^{in\phi_i} \equiv q_i \quad (30)$$

¹³⁷⁷ where we omitted the harmonic n and we will follow a new approach to derive the $corr_n\{4\}$ without
¹³⁷⁸ tracking weight, which proves its benefits in the case of track efficiency correction:

$$\begin{aligned} \sum q_i \sum q_i^* \sum q_i \sum q_i^* &= \sum q_i q_j^* q_k q_l^* \\ &+ 4 \sum q_i q_i^* q_j q_k^* + 2 \sum q_i q_j^* q_i q_k^* \\ &+ 2 \sum q_i q_i^* q_j q_j^* + \sum q_i q_j^* q_i q_j^* \\ &+ 4 \sum q_i q_i^* q_i q_j^* \\ &+ \sum q_i q_i^* q_i q_i^* \end{aligned} \quad (31)$$

¹³⁷⁹ where $q_i \equiv Q$ and $q_i^* \equiv Q^*$. Follow the similar way, many terms in above formulas can be further

¹³⁸⁰ expanded. Like calculating $W_{(4)}$, the goal is to express the final form as a function of Q :

$$\begin{aligned}
 \sum q_i q_i^* q_i \sum q_i^* &= \sum q_i q_i^* q_i q_j^* + \sum q_i q_i^* q_i q_i^* \\
 \Rightarrow \sum q_i q_i^* q_i q_j^* &= |Q_n|^2 - M \\
 \\
 \sum q_i q_i^* \sum q_i q_i^* &= \sum q_i q_i^* q_j q_j^* + \sum q_i q_i^* q_i q_i^* \\
 \Rightarrow \sum q_i q_i^* q_j q_j^* &= M^2 - M \\
 \\
 \sum q_i q_i \sum q_i^* q_i^* &= \sum q_i q_j^* q_i q_j^* + \sum q_i q_i^* q_i q_i^* \\
 \Rightarrow \sum q_i q_j^* q_i q_j^* &= |Q_{2n}|^2 - M \\
 \\
 \sum q_i q_i^* \sum q_i \sum q_i^* &= \sum q_i q_i^* q_j q_k^* + 2 \sum q_i q_i^* q_i q_j^* + \sum q_i q_i^* q_j q_j^* + \sum q_i q_i^* q_i q_i^* \\
 \Rightarrow \sum q_i q_i^* q_j q_k^* &= (M-2)|Q_n|^2 - M^2 + 2M \\
 \\
 \sum q_i q_i \sum q_i^* \sum q_i^* &= \sum q_i q_j^* q_i q_k^* + 2 \sum q_i q_i^* q_i q_j^* + \sum q_i q_j^* q_i q_j^* + \sum q_i q_i^* q_i q_i^* \\
 \Rightarrow \sum q_i q_j^* q_i q_k^* &= |Q_{2n} Q_n^* Q_n| - |Q_{2n}|^2 - 2|Q_n|^2 + 2M
 \end{aligned} \tag{32}$$

¹³⁸¹ By solving the linear equations above, we could get:

$$\sum q_i q_j^* q_k q_l^* = |Q_n|^4 + |Q_{2n}|^2 - 2|Q_{2n} Q_n^* Q_n| - 4(M-2)|Q_n|^2 + 2M(M-3) \tag{33}$$

¹³⁸² where the result is identical to the procedure described in the direct cumulant paper.

¹³⁸³ The advantage of this approach is to deal with track weights. q_i is then modified as:

$$w_i e^{jn\phi_i} \equiv q_i \tag{34}$$

¹³⁸⁴ and then all the expansions can be easily obtained by slightly modifying some of the summation terms:

$$\begin{aligned}
 \sum q_i q_i^* q_i \sum q_i^* &= \sum q_i q_i^* q_i q_j^* + \sum q_i q_i^* q_i q_i^* \\
 \Rightarrow \sum q_i q_i^* q_i q_j^* &= |Q_{n,3} Q_{n,1}^*| - S_4^1 \\
 \\
 \sum q_i q_i^* \sum q_i q_i^* &= \sum q_i q_i^* q_j q_j^* + \sum q_i q_i^* q_i q_i^* \\
 \Rightarrow \sum q_i q_i^* q_j q_j^* &= S_2^2 - S_4^1 \\
 \\
 \sum q_i q_i \sum q_i^* q_i^* &= \sum q_i q_j^* q_i q_j^* + \sum q_i q_i^* q_i q_i^* \\
 \Rightarrow \sum q_i q_j^* q_i q_j^* &= |Q_{2n,2}|^2 - S_4^1
 \end{aligned} \tag{35}$$

$$\begin{aligned}
 \sum q_i q_i^* \sum q_i \sum q_i^* &= \sum q_i q_i^* q_j q_k^* + 2 \sum q_i q_i^* q_i q_j^* + \sum q_i q_i^* q_j q_j^* + \sum q_i q_i^* q_i q_i^* \\
 \Rightarrow \sum q_i q_i^* q_j q_k^* &= S_2^1 |Q_{n,1}|^2 - 2|Q_{n,3} Q_{n,1}^*| - S_2^2 + 2S_4^1
 \end{aligned}$$

$$\begin{aligned}
 \sum q_i q_i \sum q_i^* \sum q_i^* &= \sum q_i q_j^* q_i q_k^* + 2 \sum q_i q_i^* q_i q_j^* + \sum q_i q_j^* q_i q_j^* + \sum q_i q_i^* q_i q_i^* \\
 \Rightarrow \sum q_i q_j^* q_i q_k^* &= |Q_{2n,2} Q_{n,1}^* Q_{n,1}| - |Q_{2n,2}|^2 - 2|Q_{n,3} Q_{n,1}^*| + 2S_4^1
 \end{aligned}$$

1385 By solving the linear equations above, we could get:

$$\sum q_i q_j^* q_k q_l^* = |Q_{n,1}|^4 + |Q_{2n,2}|^2 - 2|Q_{2n,2} Q_{n,1}^* Q_{n,1}| + 8|Q_{n,3} Q_{n,1}| - 4S_2^1 |Q_{n,1}|^2 + 2S_2^2 - 6S_4^1 \quad (36)$$

1386 and the result is consistent with original direct cumulant paper.

1387 8.4.2 2 sub-event method: 1st kind

1388 The formulas for 2 sub-event method will be simpler compared with traditional cumulant, because some
1389 duplicated terms no longer show up. For the purpose of better book keeping, we will define the new
1390 $Q_{n,k}$ and S_p^k for the sub-event case. In the following formulas, subscript i and j denote particles from
1391 sub-event A and subscript k and l denote particles from sub-event B :

$$\begin{aligned} A_{n,k} &\equiv \sum w_i^k e^{in\phi_i} \\ B_{n,k} &\equiv \sum w_l^k e^{in\phi_l} \end{aligned} \quad (37)$$

1392

$$\begin{aligned} X_p^k &\equiv (\sum w_i^p)^k \\ Y_p^k &\equiv (\sum w_l^p)^k \end{aligned} \quad (38)$$

1393 The expression for $W_{\{2_{a|b}\}}$ is much simpler since two particles come from two sub-events and they
1394 can never be the same particle:

$$W_{\{2_{a|b}\}} \equiv \sum w_i w_k = \sum w_i \sum w_j = X_1^1 Y_1^1 \quad (39)$$

1395 In a similar way, $W_{\{4_{a,a|b,b}\}}$ can also be determined:

$$W_{\{4_{a,a|b,b}\}} = X_1^2 Y_1^2 - X_2^1 Y_1^2 - X_1^2 Y_2^1 + X_2^1 Y_2^1 \quad (40)$$

1396 For the 1st kind of 2 sub-event method, two particles from the same sub-event always have the same
1397 sign. We still use the same notation q_i as the traditional cumulant method, only the subscript determines
1398 which sub-event it comes from: i, j from one sub-event and k, l from the other sub-event:

$$\sum q_i \sum q_k^* = \sum q_i q_k^* \quad (41)$$

1399 and for 4-particle cumulant:

$$\begin{aligned} \sum q_i \sum q_j \sum q_k^* \sum q_l^* &= \sum q_i q_j q_k^* q_l^* + \sum q_i q_j q_k^* q_l^* + \sum q_i q_j q_k^* q_k^* + \sum q_i q_j q_k^* q_k^* \\ \sum q_i q_j \sum q_k^* q_k^* &= \sum q_i q_j \sum q_k^* q_k^* \\ \sum q_i q_j \sum q_k^* \sum q_l^* &= \sum q_i q_j q_k^* q_l^* + \sum q_i q_j q_k^* q_k^* \\ \sum q_i \sum q_j \sum q_k^* q_k^* &= \sum q_i q_j q_k^* q_k^* + \sum q_i q_j q_k^* q_k^* \end{aligned} \quad (42)$$

1400 Solve the linear equations and plug in $A_{n,k}$ and $B_{n,k}$:

$$\sum q_i q_j q_k^* q_l^* = |A_{n,1} B_{n,1}^* A_{n,1} B_{n,1}^*| - |A_{2n,2} B_{n,1}^* B_{n,1}^*| - |A_{n,1} A_{n,1} B_{2n,2}^*| + |A_{2n,2} B_{2n,2}^*| \quad (43)$$

1401 where ' \parallel ' means taking the real part of the inner product.

¹⁴⁰² **8.4.3 2 sub-event method: 2nd kind**

¹⁴⁰³ Keeping all the same notations from 1st kind, here we list all the formulas for the 2 sub-event 2nd kind.

¹⁴⁰⁴ For $W_{(2)}$, there are three items:

$$\begin{aligned} W_{(2_{a|a})} &\equiv \sum w_i w_j = \sum w_i \sum w_i - \sum w_i^2 \\ W_{(2_{b|b})} &\equiv \sum w_k w_l = \sum w_k \sum w_k - \sum w_k^2 \\ W_{(2_{a|b})} &\equiv \sum w_i w_k = \sum w_i \sum w_k \end{aligned} \quad (44)$$

¹⁴⁰⁵ and expression for $W_{(4_{a,b|a,b})}$ is same as 1st kind because conjugate will not change the track weight,
¹⁴⁰⁶ which is a real number:

$$W_{(4_{a,b|a,b})} = X_1^2 Y_1^2 - X_2^1 Y_1^2 - X_1^2 Y_2^1 + X_2^1 Y_2^1 \quad (45)$$

¹⁴⁰⁷ Compared with 1st kind, the expression for $\sum q_i q_j^* q_k q_l^*$ is similar:

$$\begin{aligned} \sum q_i \sum q_i^* \sum q_k \sum q_k^* &= \sum q_i q_j^* q_k q_l^* + \sum q_i q_i^* q_k q_l^* + \sum q_i q_j^* q_k q_k^* + \sum q_i q_i^* q_k q_k^* \\ \sum q_i q_i^* \sum q_k q_k^* &= \sum q_i q_i^* \sum q_k q_k^* \\ \sum q_i q_i^* \sum q_k \sum q_k^* &= \sum q_i q_i^* q_k q_l^* + \sum q_i q_i^* q_k q_k^* \\ \sum q_i \sum q_i^* \sum q_k q_k^* &= \sum q_i q_j^* q_k q_k^* + \sum q_i q_i^* q_k q_k^* \end{aligned} \quad (46)$$

¹⁴⁰⁸ and the 4-particle correlation:

$$\sum q_i q_j^* q_k q_l^* = |A_{n,1} A_{n,1}^* B_{n,1} B_{n,1}^*| - X_2^1 |B_{n,1}|^2 - Y_2^1 |A_{n,1}|^2 + X_2^1 Y_2^1 \quad (47)$$

¹⁴⁰⁹ **8.4.4 3 sub-event method: 1st kind**

¹⁴¹⁰ Since there is one more sub-event, we will change our notations as:

$$\begin{aligned} A_{n,k} &\equiv \sum w_i^k e^{in\phi_i} \\ B_{n,k} &\equiv \sum w_k^k e^{in\phi_k} \\ C_{n,k} &\equiv \sum w_l^k e^{in\phi_l} \end{aligned} \quad (48)$$

¹⁴¹¹

$$\begin{aligned} X_p^k &\equiv (\sum w_i^p)^k \\ Y_p^k &\equiv (\sum w_k^p)^k \\ Z_p^k &\equiv (\sum w_l^p)^k \end{aligned} \quad (49)$$

¹⁴¹² where in the following formulas 2 of 4 particles always come from sub-event A , denoted as subscript i and
¹⁴¹³ j . Particle k comes from sub-event B and particle l comes from sub-event C . All the other permutations
¹⁴¹⁴ can be easily obtained by rotating the A, B and C .

¹⁴¹⁵ All the formulas can be directly "guessed" from without weights case and they are summarized
¹⁴¹⁶ below:

$$\begin{aligned} W_{(2_{a|b})} &= X_1^1 Y_1^1 \\ W_{(2_{a|c})} &= X_1^1 Z_1^1 \\ W_{(4_{a,a|b,c})} &= (X_1^2 - X_2^1) Y_1^1 Z_1^1 \\ \sum q_i q_k^* &= |A_{n,1} B_{n,1}^*| \\ \sum q_i q_l^* &= |A_{n,1} C_{n,1}^*| \\ \sum q_i q_j^* q_k q_l^* &= |A_{n,1} B_{n,1}^* A_{n,1} C_{n,1}^*| - |A_{2n,2} B_{n,1}^* C_{n,1}^*| \end{aligned} \quad (50)$$

¹⁴¹⁷ **8.4.5 3 sub-event method: 2nd kind**

¹⁴¹⁸ Use same notation as the 1st kind, all the formulas are listed below:

$$\begin{aligned}
 W_{\langle 2_{a|b} \rangle} &= X_1^1 Y_1^1 \\
 W_{\langle 2_{a|c} \rangle} &= X_1^1 Z_1^1 \\
 W_{\langle 2_{b|c} \rangle} &= Y_1^1 Z_1^1 \\
 W_{\langle 2_{a|a} \rangle} &= X_1^2 - X_2^1 \\
 W_{\langle 4_{a,b|a,c} \rangle} &= (X_1^2 - X_2^1) Y_1^1 Z_1^1 \\
 \sum q_i q_k^* &= |A_{n,1} B_{n,1}^*| \\
 \sum q_i q_l^* &= |A_{n,1} C_{n,1}^*| \\
 \sum q_k q_l^* &= |B_{n,1} C_{n,1}^*| \\
 \sum q_i q_j^* &= |A_{n,1}|^2 - X_2^1 \\
 \sum q_i q_j^* q_k q_l^* &= |A_{n,1} A_{n,1}^* B_{n,1} C_{n,1}^*| - X_2^1 |B_{n,1}^* C_{n,1}^*|
 \end{aligned} \tag{51}$$

1419 References

- 1420 [1] A. Adare *et al.*, “Dihadron azimuthal correlations in Au+Au collisions at $s_{\text{NN}}^{**}(1/2) = 200$ -
1421 GeV,” *Phys. Rev.*, vol. C78, p. 014901, 2008.
- 1422 [2] B. I. Abelev *et al.*, “Long range rapidity correlations and jet production in high energy nuclear
1423 collisions,” *Phys. Rev.*, vol. C80, p. 064912, 2009.
- 1424 [3] B. Alver *et al.*, “High transverse momentum triggered correlations over a large pseudorapidity
1425 acceptance in Au+Au collisions at $s_{\text{NN}}^{**}1/2 = 200$ GeV,” *Phys. Rev. Lett.*, vol. 104, p. 062301,
1426 2010.
- 1427 [4] K. Aamodt *et al.*, “Higher harmonic anisotropic flow measurements of charged particles in Pb-Pb
1428 collisions at $\sqrt{s_{\text{NN}}} = 2.76$ TeV,” *Phys. Rev. Lett.*, vol. 107, p. 032301, 2011.
- 1429 [5] G. Aad *et al.*, “Measurement of the azimuthal anisotropy for charged particle production in $\sqrt{s_{\text{NN}}} =$
1430 2.76 TeV lead-lead collisions with the ATLAS detector,” *Phys. Rev. C*, vol. 86, p. 014907, 2012.
- 1431 [6] S. Chatrchyan *et al.*, “Measurement of higher-order harmonic azimuthal anisotropy in PbPb colli-
1432 sions at $\sqrt{s_{\text{NN}}} = 2.76$ TeV,” *Phys. Rev. C*, vol. 89, p. 044906, 2014.
- 1433 [7] S. Chatrchyan *et al.*, “Observation of long-range near-side angular correlations in proton-lead col-
1434 lisions at the LHC,” *Phys. Lett. B*, vol. 718, p. 795, 2013.
- 1435 [8] B. Abelev *et al.*, “Long-range angular correlations on the near and away side in p -Pb collisions at
1436 $\sqrt{s_{\text{NN}}} = 5.02$ TeV,” *Phys. Lett. B*, vol. 719, p. 29, 2013.
- 1437 [9] G. Aad *et al.*, “Observation of Associated Near-side and Away-side Long-range Correlations in
1438 $\sqrt{s_{\text{NN}}}=5.02$ TeV Proton-lead Collisions with the ATLAS Detector,” *Phys. Rev. Lett.*, vol. 110,
1439 p. 182302, 2013.
- 1440 [10] A. Adare *et al.*, “Quadrupole Anisotropy in Dihadron Azimuthal Correlations in Central d +Au
1441 Collisions at $\sqrt{s_{\text{NN}}}=200$ GeV,” *Phys. Rev. Lett.*, vol. 111, p. 212301, 2013.
- 1442 [11] G. Aad *et al.*, “Measurement of long-range pseudorapidity correlations and azimuthal harmonics
1443 in $\sqrt{s_{\text{NN}}} = 5.02$ TeV proton-lead collisions with the ATLAS detector,” *Phys. Rev. C*, vol. 90,
1444 p. 044906, 2014.
- 1445 [12] V. Khachatryan *et al.*, “Evidence for Collective Multiparticle Correlations in p-Pb Collisions,” *Phys.*
1446 *Rev. Lett.*, vol. 115, no. 1, p. 012301, 2015.
- 1447 [13] V. Khachatryan *et al.*, “Observation of Long-Range Near-Side Angular Correlations in Proton-
1448 Proton Collisions at the LHC,” *JHEP*, vol. 1009, p. 091, 2010.
- 1449 [14] G. Aad *et al.*, “Observation of Long-Range Elliptic Azimuthal Anisotropies in $\sqrt{s} = 13$ and 2.76
1450 TeV pp Collisions with the ATLAS Detector,” *Phys. Rev. Lett.*, vol. 116, no. 17, p. 172301, 2016.
- 1451 [15] M. Aaboud *et al.*, “Measurements of long-range azimuthal anisotropies and associated Fourier
1452 coefficients for pp collisions at $\sqrt{s} = 5.02$ and 13 TeV and p +Pb collisions at $\sqrt{s_{\text{NN}}} = 5.02$ TeV
1453 with the ATLAS detector,” 2016.
- 1454 [16] “Measurement of multi-particle azimuthal correlations in pp collisions at $\text{sqrt}(s) = 5.02$ and 13 TeV
1455 with the ATLAS detector,” *ATLAS-CONF-2016-106*.

- 1456 [17] P. Bozek and W. Broniowski, “Collective dynamics in high-energy proton-nucleus collisions,” *Phys.*
 1457 *Rev.*, vol. C88, no. 1, p. 014903, 2013.
- 1458 [18] K. Dusling and R. Venugopalan, “Comparison of the color glass condensate to dihadron correlations
 1459 in proton-proton and proton-nucleus collisions,” *Phys. Rev.*, vol. D87, no. 9, p. 094034, 2013.
- 1460 [19] V. Khachatryan *et al.*, “Evidence for collectivity in pp collisions at the LHC,” *Submitted to: Physics*
 1461 *Letters B*, 2016.
- 1462 [20] N. Borghini, P. M. Dinh, and J.-Y. Ollitrault, “A New method for measuring azimuthal distributions
 1463 in nucleus-nucleus collisions,” *Phys. Rev. C*, vol. 63, p. 054906, 2001.
- 1464 [21] G. Aad *et al.*, “Measurement with the ATLAS detector of multi-particle azimuthal correlations in
 1465 p+Pb collisions at $\sqrt{s_{NN}} = 5.02$ TeV,” *Phys. Lett.*, vol. B725, pp. 60–78, 2013.
- 1466 [22] J. Jia, M. Zhou, and A. Trzupek, “,” *Note on the subevent cumulant method*.
- 1467 [23] A. Bilandzic, R. Snellings, and S. Voloshin, “Flow analysis with cumulants: Direct calculations,”
 1468 *Phys. Rev. C*, vol. 83, p. 044913, 2011.
- 1469 [24] A. Bilandzic, C. H. Christensen, K. Gulbrandsen, A. Hansen, and Y. Zhou, “Generic framework
 1470 for anisotropic flow analyses with multiparticle azimuthal correlations,” *Phys. Rev.*, vol. C89, no. 6,
 1471 p. 064904, 2014.
- 1472 [25] B. B. Abelev *et al.*, “Multi-particle azimuthal correlations in p-Pb and Pb-Pb collisions at the CERN
 1473 Large Hadron Collider,” *Phys. Rev. C*, vol. 90, p. 054901, 2014.
- 1474 [26] “The ATLAS Experiment at the CERN Large Hadron Collider,” *JINST*, vol. 3, p. S08003, 2008.
- 1475 [27] “,” *ATLAS-LHCC-2010-013. ATLAS-TDR-19*.
- 1476 [28] “,” *ATLAS-LHCC-2012-009. ATLAS-TDR-19-ADD-1*.
- 1477 [29] “Performance of the ATLAS Trigger System in 2010,” *Eur. Phys. J. C*, vol. 72, p. 1849, 2012.
- 1478 [30] G. Aad *et al.*, “Charged-particle distributions in $\sqrt{s}=13$ TeV pp interactions measured with the
 1479 ATLAS detector at the LHC,” *Phys. Lett.*, vol. B758, pp. 67–88, 2016.
- 1480 [31] T. Sjostrand, S. Mrenna, and P. Z. Skands, “A Brief Introduction to PYTHIA 8.1,” *Comput. Phys.*
 1481 *Commun.*, vol. 178, pp. 852–867, 2008.
- 1482 [32] “ATLAS tunes of PYTHIA 6 and Pythia 8 for MC11,” *ATLAS-PHYS-PUB-2011-009*.
- 1483 [33] “Measurement of the long-range pseudo-rapidity correlation and associated Fourier harmonics in
 1484 $\sqrt{s_{NN}} = 5.02$ TeV proton-lead collisions with the ATLAS detector,”
- 1485 [34] “Performance of the ATLAS Minimum Bias and Forward Detector Triggers in pPb collisions,”
 1486 *ATLAS-CONF-2013-104*.
- 1487 [35] M. Gyulassy and X.-N. Wang, “HIJING 1.0: A Monte Carlo program for parton and particle pro-
 1488 duction in high-energy hadronic and nuclear collisions,” *Comput. Phys. Commun.*, vol. 83, p. 307,
 1489 1994.
- 1490 [36] A. Bzdak, P. Bozek, and L. McLerran, “Fluctuation induced equality of multi-particle eccentricities
 1491 for four or more particles,” *Nucl. Phys.*, vol. A927, pp. 15–23, 2014.

- ¹⁴⁹² [37] L. Yan and J.-Y. Ollitrault, “Universal fluctuation-driven eccentricities in proton-proton, proton-
¹⁴⁹³ nucleus and nucleus-nucleus collisions,” *Phys. Rev. Lett.*, vol. 112, p. 082301, 2014.