Measurement of long-range multiparticle azimuthal correlations with the subevent cumulant method in pp and p+Pb collisions with the ATLAS detector at the CERN Large Hadron Collider

M. Aaboud et al.*
(ATLAS Collaboration)

(Received 14 August 2017; published 12 February 2018)

A detailed study of multiparticle azimuthal correlations is presented using pp data at √s = 5.02 and 13 TeV, and p+Pb data at √sNN = 5.02 TeV, recorded with the ATLAS detector at the CERN Large Hadron Collider. The azimuthal correlations are probed using four-particle cumulants c₄[4] and flow coefficients v₂[4] = (-c₄[4])½ for n = 2 and 3, with the goal of extracting long-range multiparticle azimuthal correlation signals and suppressing the short-range correlations. The values of c₄[4] are obtained as a function of the average number of charged particles per event, ⟨Nch⟩, using the recently proposed two-subevent and three-subevent cumulant methods, and compared with results obtained with the standard cumulant method. The standard method is found to be strongly biased by short-range correlations, which originate mostly from jets with a positive contribution to c₄[4]. The three-subevent method, on the other hand, is found to be least sensitive to short-range correlations. The three-subevent method gives a negative c₄[4], and therefore a well-defined v₂[4], nearly independent of ⟨Nch⟩, which implies that the long-range multiparticle azimuthal correlations persist to events with low multiplicity. Furthermore, v₂[4] is found to be smaller than the v₂[2] measured using the two-particle correlation method, as expected for long-range collective behavior. Finally, the measured values of v₂[4] and v₂[2] are used to estimate the number of sources relevant for the initial eccentricity in the collision geometry. The results based on the subevent cumulant technique provide direct evidence, in small collision systems, for a long-range collectivity involving many particles distributed across a broad rapidity interval.

DOI: 10.1103/PhysRevC.97.024904

1. INTRODUCTION

The study of azimuthal correlations in high-energy nuclear collisions at the Relativistic Heavy Ion Collider (RHIC) and the Large Hadron Collider (LHC) has been important for understanding the multiparton dynamics of QCD in the strongly coupled nonperturbative regime. One striking observation is the long-range ridge [1–5] in two-particle angular correlations (2PC): an apparent collimated emission of particle pairs with small relative azimuthal angle ∆φ and large separation in pseudorapidity ∆η. The ridge signature from 2PC is characterized by a Fourier decomposition of the correlation function C(∆φ) ~ 1 + 2 ∑ₙ vₙ² cos(n ∆φ), where vₙ denotes the single-particle anisotropy harmonic coefficients. The second-order coefficient v₂ is observed to be the largest, followed by v₃ [3,4]. These coefficients carry information about the collective behavior of the produced system. The ridge was first discovered in nucleus-nucleus (A+A) collisions [1–6], but was later observed in small systems such as proton-nucleus (p+A) collisions [7–11], light-ion–nucleus collisions [12], and more recently in proton-proton (pp) collisions [13–16]. The ridge in large systems, such as central or midcentral A+A collisions, is commonly interpreted as the result of collective hydrodynamic expansion of hot and dense nuclear matter created in the overlap region of the colliding nuclei. Since the formation of an extended region of nuclear matter is not expected in small collision systems such as p+A and pp, the origin of the ridge there could be different from that formed in large collision systems. There remains considerable debate in the theoretical community as to whether the ridge in small systems is of hydrodynamic origin, like it is in A+A collisions [17], or stems from other effects such as initial-state gluon saturation [18].

An important question about the ridge is whether it involves all particles in the event (collective flow) or if it arises merely from correlations among a few particles, due to resonance decays, jets, or multijet production (nonflow). In small systems the contributions from nonflow sources, in particular from jets and dijets, are large. The extraction of a ridge signal using the 2PC method requires a large ∆η gap and careful removal of the significant contribution from dijet production [8–10,14,15,19]. Since collective flow is intrinsically a multiparticle phenomenon, it can be probed more directly using cumulants based on multiparticle correlation techniques [20]. Azimuthal correlations involving four, six, and eight particles have been measured in p+Pb, d+Au, and pp collisions, and a significant v₄ signal has been obtained [11,19,21,22]. One weakness of the standard multiparticle cumulant method is that it does not suppress adequately the nonflow correlations.

*Full author list given at the end of the article.

Published by the American Physical Society under the terms of the Creative Commons Attribution 4.0 International license. Further distribution of this work must maintain attribution to the author(s) and the published article’s title, journal citation, and DOI. Funded by SCOAP³.
Furthermore, the magnitude of $c_2(4)$ at smaller values of the charged particle multiplicity, $N_{ch}$ [11,16,19,21]. Furthermore, the magnitude of $c_2(4)$ and the $N_{ch}$ value at which the sign change occurs are found to depend sensitively on the exact definition of $N_{ch}$ used to categorize the events. These observations suggest that the standard cumulant method, on which several previous measurements in small systems are based, is strongly contaminated by nonflow correlations [11,19,21,22], especially in $pp$ collisions and low $N_{ch}$ region.

Recently an improved cumulant method based on the correlation between particles from different subevents separated in $\eta$ has been proposed to further reduce the nonflow correlations [23]. The effectiveness of this method for suppressing nonflow correlations has been validated using the PYTHIA8 event generator [24], which contains only nonflow correlations.

This paper presents measurements of $c_{2(4)}$ and $c_{3(4)}$ in $pp$ collisions at $\sqrt{s} = 5.02$ and $13$ TeV, as well as $p+Pb$ collisions at $\sqrt{s_{NN}} = 5.02$ TeV. They are obtained using two- and three-subevent cumulant methods and are compared with the standard cumulant method. The $c_2(4)$ cumulant is converted to the corresponding $v_2$ coefficient and compared with the results obtained using the two-particle correlation method in Refs. [10,15] to assess the nature of the event-by-event fluctuation of the collective flow in these collisions.

The paper is organized as follows. Section II describes the framework for the standard, two-subevent and three-subevent four-particle cumulant methods used in this analysis. Details of the detector, trigger, data sets, as well as event and track selections are provided in Secs. III–V. The correlation analysis and systematic uncertainties are described in Secs. VI and VII, respectively. The measured cumulants from the three data sets are provided in Sec. VIII. A summary is given in Sec. IX.

II. FOUR-PARTICLE CUMULANTS

The multiparticle cumulant method [20] is used to extract the amplitude of long-range azimuthal correlations of particles produced in high-energy collisions. This method has the advantage of suppressing correlations from jets and dijets, instead of relying on an explicit procedure to correct $v_n$ harmonics for dijet contributions in the 2PC approach, as done in Refs. [10,14]. The framework for the standard cumulant is described in Refs. [25,26], which was recently extended to the case of subevent cumulants in Ref. [23]. This paper presents measurements of four-particle cumulants obtained with the standard, two-subevent, and three-subevent methods. The following discussion first describes the standard cumulant method, then describes the two- and three-subevent methods focusing on the differences from the standard method.

The cumulant methods involve the calculation of 2$k$-particle azimuthal correlations $\langle (2k)_{n}) \rangle$, and 2$k$-particle cumulants, $c_n(2k)$, for the $n$th-order flow harmonics. The two- or four-particle azimuthal correlations in one event are evaluated as [23,25,26]:

$$\langle (2)_{n} \rangle = \langle e^{i(n \phi_1 - \phi_2)} \rangle = \frac{q_n^2 - \tau_1}{1 - \tau_1},$$

$$\langle (4)_{n} \rangle = \langle e^{i(n \phi_1 + \phi_1 - \phi_1)} \rangle = \frac{q_n^4 - 2\tau_1 \langle \text{Re}[q_{2n>2}q_{n>2}^2] + 2q_n^2 \rangle + 8\tau_2 \text{Re}[q_{n>2}q_{n>2}^*] + \tau_1^2 (2 + q_{2n>2}^2 - 6\tau_1)}{1 - 6\tau_1 + 8\tau_2 + 3\tau_1^2 - 6\tau_1},$$

where $\langle \rangle$ denotes a single-event average over all pairs or quadruplets, respectively. The averages from Eqs. (1) and (2) are expanded into per-particle normalized flow vectors $q_{n,l}$ and factors $\tau_l$ with $l = 1, 2, \ldots$:

$$q_{n,l} \equiv \frac{\sum_j w_j^l e^{i\phi_j}}{\sum_j w_j^l}, \quad q_{n,l} \equiv |q_{n,l}|, \quad q_n \equiv q_{n,1}, \quad \tau_l \equiv \frac{\sum_j w_j^{l+1}}{\sum_j w_j^{l+1}}$$

where the sum runs over all $M$ particles in the event and $w_j$ is a weight assigned to the $j$th particle. This weight is constructed to correct for both detector nonuniformity and tracking inefficiency as explained in Sec. VI. For unit weight $w_j = 1$, then $q_{m,m} = q_{m,n}$, and $\tau_1 = 1/M'$. The two- and four-particle cumulants are obtained from the azimuthal correlations as:

$$c_n(2) = \langle (2)_{n} \rangle,$$

$$c_n(4) = \langle (4)_{n} \rangle - 2\langle (2)_{n} \rangle^2,$$

where $\langle \rangle$ represents a weighted average of $\langle (2k)_{n} \rangle$ over an event ensemble. In the absence of nonflow correlations, $c_n(2k)$ reflects the moments of the distribution of the flow coefficient $v_n$:

$$c_n(2)_{\text{flow}} = \langle v_n^2 \rangle, \quad c_n(4)_{\text{flow}} = \langle v_n^4 \rangle - 2\langle v_n^2 \rangle^2.$$

If harmonic coefficients do not fluctuate event by event, Eq. (6) gives $c_n(2)_{\text{flow}} = v_n^2, c_n(4)_{\text{flow}} = v_n^4$, and $c_n(4)_{\text{flow}}$ is expected to be negative. Therefore, the flow coefficients from two- and four-particle cumulants are defined as:

$$v_n(2) = \sqrt{c_n(2)}, \quad v_n(4) = \sqrt{-c_n(4)}.$$

In the standard cumulant method described so far, all $2k$-particle multiplets involved in $\langle (2k)_{n} \rangle$ are selected using the entire detector acceptance. To further suppress the nonflow correlations that typically involve particles emitted within a localized region in $\eta$, the particles can be grouped into several subevents, each covering a nonoverlapping $\eta$ interval [23]. The multiparticle correlations are then constructed by correlating particles between different subevents, further reducing nonflow correlations. This analysis uses the subevent cumulant methods based on two and three subevents as described in the following.

In the two-subevent cumulant method, the entire event is divided into two subevents, labeled as $a$ and $b$, for example, ac-
according to $-\eta_{\text{max}} < \eta < 0$ and $0 < \eta < \eta_{\text{max}}$, where $\eta_{\text{max}} = 2.5$ is the maximum $\eta$ used in the analysis and corresponds to the ATLAS detector acceptance for charged particles. The per-event two- and four-particle azimuthal correlations are then evaluated as:

$$
\langle \{2\rangle_n \rangle_{a,b} = \langle e^{i(\phi_a - \phi_b)} \rangle_{a,b} = \text{Re}\{q_{a,b}g_{n,b}^*\},
$$

(8)

$$
\langle \{4\rangle_n \rangle_{2a,2b} = \langle e^{i(\phi_a + \phi_b - \phi_c - \phi_d)} \rangle_{a,b,c,d} = \frac{\langle q_n^2 - \tau_1 q_{2a})_n - \tau_2 q_{2b})_n \rangle}{(1 - \tau_1)_{ab} (1 - \tau_2)_{bc}},
$$

(9)

where the superscript or subscript $a$ ($b$) indicates particles chosen from the subevent $a$ ($b$). Here the four-particle cumulant is defined as:

$$
c_{n}^{2a(2)} = \langle \{4\rangle_n \rangle_{2a,2b} - 2\langle \{2\rangle_n \rangle_{a,b} \rangle_{a,b}.
$$

(10)

The two-subevent method should suppress correlations within a single jet (intrajet correlations), since each jet usually emits particles into only one subevent.

In the three-subevent cumulant method, the event is divided into three subevents $a$, $b$, and $c$ each covering a unique range, for example $-\eta_{\text{max}} < \eta_a < -|\eta_{\text{max}}|/3$, $|\eta_b| < |\eta_{\text{max}}|/3$, and $\eta_{\text{max}}/3 < \eta_c < \eta_{\text{max}}$. The four-particle azimuthal correlations and cumulants are then evaluated as:

$$
\langle \{4\rangle_n \rangle_{2a,2b,c} = \langle e^{i(\phi_a + \phi_b - \phi_c - \phi_d)} \rangle_{a,b,c,d} = \frac{\langle q_n^2 - \tau_1 q_{2a})_n + \tau_2 q_{2b})_n + \tau_3 q_{2c})_n \rangle}{(1 - \tau_1)_{ab} (1 - \tau_2)_{bc} (1 - \tau_3)_{bc}},
$$

(11)

$$
c_{n}^{2a,b,c} \equiv \langle \{4\rangle_n \rangle_{2a,2b,c} - 2\langle \{2\rangle_n \rangle_{a,b} \rangle_{a,b} \rangle_{a,b}.
$$

(12)

where $\langle \{2\rangle_n \rangle_{a,b}$ and $\langle \{4\rangle_n \rangle_{a,b}$ are two-particle correlators defined as in Eq. (8). Since the two jets in a dijet event usually produce particles in at most two subevents, the three-subevent method further suppresses nonflow contributions from interjet correlations associated with dijets. To enhance the statistical precision, the $\eta$ range for subevent $a$ is also interchanged with that for subevent $b$ or $c$, and the resulting three $c_n^{2a,b,c}[\{4\}]$ values are averaged to obtain the final result.

**III. DETECTOR AND TRIGGER**

The ATLAS detector [27] provides nearly full solid-angle coverage around the collision point with tracking detectors, calorimeters, and muon chambers, and is well suited for measurement of multiparticle correlations over a large pseudorapidity range.  

\footnote{ATLAS typically uses a right-handed coordinate system with its origin at the nominal interaction point (IP) in the center of the detector and the $z$ axis along the beam pipe. The $x$ axis points from the IP to the center of the LHC ring, and the $y$ axis points upward. Cylindrical coordinates ($r,\phi$) are used in the transverse plane, $\phi$ being the azimuthal angle around the beam pipe. By default, the pseudorapidity is defined in terms of the polar angle $\theta$ as $\eta = -\ln \tan(\theta/2)$. However, for asymmetric $p+Pb$ or $Pb+p$ collisions, the $-\eta$ direction is always defined as the direction of the Pb beam.}

The measurements were performed primarily using the inner detector (ID), minimum-bias trigger scintillators (MBTS), and the zero-degree calorimeters (ZDCs). The ID detects charged particles within $|\eta| < 2.5$ using a combination of silicon pixel detector, a silicon microstrip detector (SCT), and a straw-tube transition radiation tracker, all immersed in a 2 T axial magnetic field [28]. An additional pixel layer, the insertable B-layer (IBL) [29] installed between Run 1 (2010–2013) and Run 2 (2015–2018), is available for the Run-2 data sets. The MBTS, rebuilt before Run 2, detects charged particles within $2.1 \leq |\eta| \leq 5.9$ using two hodoscopes of counters positioned at $z = \pm 3.6 \text{ m}$. The ZDCs are positioned at $\pm 140 \text{ m}$ from the collision point, and detect neutral particles, primarily neutrons and photons, with $|\eta| > 8.3$.

The ATLAS trigger system [30] consists of a Level-1 (L1) trigger implemented using a combination of dedicated electronics and programmable logic, and a high-level trigger (HLT) implemented in processors. The HLT reconstructs charged-particle tracks using methods similar to those applied in the offline analysis, allowing high-multiplicity track (HMT) triggers that select events based on the number of tracks with $p_T > 0.4 \text{ GeV}$ associated with the vertex with the largest number of tracks. The different HMT triggers also apply additional requirements on either the transverse energy ($E_T$) in the calorimeters or on the number of hits in the MBTS at L1, and on the number of charged-particle tracks reconstructed by the HLT. The $pp$ and $p+Pb$ data were collected using a combination of the minimum-bias and HMT triggers. More details of the triggers used for the $pp$ and $p+Pb$ data can be found in Refs. [15,31] and Refs. [10,32], respectively.

**IV. DATA SETS AND MONTE CARLO SIMULATIONS**

This analysis uses integrated luminosities of 28 nb$^{-1}$ of $p+Pb$ data recorded at $\sqrt{s_{NN}} = 5.02 \text{ TeV}$, 0.17 pb$^{-1}$ of $pp$ data recorded at $\sqrt{s} = 5.02 \text{ TeV}$, and 0.9 pb$^{-1}$ of $pp$ data recorded at $\sqrt{s} = 13 \text{ TeV}$, all taken by the ATLAS experiment at the LHC. The $p+Pb$ data were mainly collected in 2013, but also include 0.3 nb$^{-1}$ data collected in November 2016, which increases the number of events at moderate multiplicity (see Sec. V). During both $p+Pb$ runs, the LHC was configured with a 4 TeV proton beam and a 1.57 TeV per-nucleon Pb beam that together produced collisions at $\sqrt{s_{NN}} = 5.02 \text{ TeV}$, with a rapidity shift of 0.465 of the nucleon–nucleon center-of-mass frame towards the proton beam direction relative to the ATLAS rest frame. The direction of the Pb beam is always defined to have negative pseudorapidity. The 5.02 TeV $pp$ data were collected in November 2015. The 13 TeV $pp$ data were collected during several special low-luminosity runs of the LHC in 2015 and 2016.

Monte Carlo (MC) simulated event samples are used to determine the track reconstruction efficiency (Sec. V). The 13 TeV and 5.02 TeV $pp$ data were simulated by the PYTHIA8 MC event generator [24] using the A2 set of tuned parameters with MSTW2008LO parton distribution functions [33]. The HIJING event generator [34] was used to produce $p+Pb$ collisions with the same energy and the same boost of the center-of-mass system as in the data. The detector response was simulated using GEANT4 [35,36] with detector conditions matching those during the data taking. The simulated events
and data events are reconstructed with the same algorithms, including those for track reconstruction.

V. EVENT AND TRACK SELECTION

The offline event selection for the p+Pb and pp data requires at least one reconstructed vertex with its longitudinal position satisfying $|z_{\text{rec}}| < 100$ mm. The vertex is required to have at least two associated tracks with $p_T > 0.4$ GeV. The mean collision rate per bunch crossing $\mu$ was approximately 0.03 for the 2013 p+Pb data, 0.001–0.006 for the 2016 p+Pb data, 0.02–1.5 for 5.02 TeV pp data, and 0.002–0.8 for the 13 TeV pp data. In order to suppress additional interactions in the same bunch crossing (referred to as pileup) in pp collisions, events containing additional vertices with at least four associated tracks are rejected. In p+Pb collisions, events with more than one good vertex, defined as any vertex for which the scalar sum of the $p_T$ of the associated tracks is greater than 5 GeV, are rejected. The remaining pileup events are further suppressed by using the signal in the ZDC on the Pb-fragmentation side. This signal is calibrated to the number of detected neutrons ($N_{\text{ch}}$) by using the location of the peak corresponding to a single neutron. The distribution of $N_{\text{ch}}$ in events with pileup is broader than that for the events without pileup. Hence a simple requirement on the ZDC signal distribution is used to further suppress events with pileup, while retaining more than 98% of the events without pileup. The impact of residual pileup, at a level of $\lesssim 10^{-3}$, is studied by comparing the results obtained from data with different $\mu$ values.

Charged-particle tracks and collision vertices are reconstructed using the same algorithms and methods applied in previous minimum-bias pp and p+Pb measurements [10,14,31]. For the 2013 p+Pb analysis, tracks are required to have a $p_T$-dependent minimum number of hits in the SCT. The transverse ($d_0$) and longitudinal ($z_0 \sin \theta$) impact parameters of the track relative to the primary vertex are both required to be less than 1.5 mm. A more detailed description of the track selection for the 2013 p+Pb data can be found in Ref. [10].

For all the data taken since the start of Run 2, the track selection criteria make use of the IBL, as described in Refs. [14,31]. Furthermore, the requirements of $|d_0^{\text{BL}}| < 1.5$ mm and $|z_0 \sin \theta| < 1.5$ mm are applied, where $d_0^{\text{BL}}$ is the transverse impact parameter of the track relative to the beam line (BL).

The cumulants are calculated using tracks passing the above selection requirements, and having $|\eta| < 2.5$ and $0.3 < p_T < 3$ GeV or $0.5 < p_T < 5$ GeV. These two $p_T$ ranges are chosen because they were often used in previous ridge measurements at the LHC [11,14–16,19]. However, to count the number of reconstructed charged particles for event-class definition (denoted by $N_{\text{ch}}^{\text{rec}}$), tracks with $p_T > 0.4$ GeV and $|\eta| < 2.5$ are used for compatibility with the requirements in the HLT selections described above. Due to different trigger requirements, most of the p+Pb events with $N_{\text{ch}}^{\text{rec}} > 150$ are provided by the 2013 data set, while the 2016 data set provides most of the events at lower $N_{\text{ch}}^{\text{rec}}$.

The efficiency of the combined track reconstruction and selection requirements in data is estimated using the MC samples reconstructed with the same tracking algorithms and the same track selection requirements. Efficiencies, $\epsilon(\eta, p_T)$, are evaluated as a function of track $\eta$, $p_T$, and the number of reconstructed charged-particle tracks, but averaged over the full range in azimuth. For all collision systems, the efficiency increases by about 4% as $p_T$ increases from 0.3 GeV to 0.6 GeV. Above 0.6 GeV, the efficiency is independent of $p_T$ and reaches 86% (72%) at $\eta \approx 0$ ($|\eta| > 2$) for pp collisions and 83% (70%) for p+Pb collisions, respectively. The efficiency is independent of the event multiplicity for $N_{\text{ch}}^{\text{rec}} > 40$. For lower-multiplicity events the efficiency is smaller by up to a few percent due to broader $d_0^{\text{BL}}$ and $z_0 \sin \theta$ distributions.

The rate of falsely reconstructed charged-particle tracks is also estimated and found to be negligibly small in all data sets. This rate decreases with increasing $p_T$, and even at the lowest transverse momenta of 0.2 GeV it is below 1% of the total number of tracks. Therefore, there is no correction for the presence of these tracks in the analysis.

In the simulated events, the reconstruction efficiency reduces the measured charged-particle multiplicity relative to the generated multiplicity for primary charged particles. The multiplicity correction factor $b$ is used to correct $N_{\text{ch}}^{\text{rec}}$ to obtain the efficiency-corrected number of charged particles per event, $\langle N_{\text{ch}} \rangle = b(N_{\text{ch}}^{\text{rec}})$. The value of the correction factor is found to be independent of $N_{\text{ch}}^{\text{rec}}$ in the range used in this analysis. Its value and the associated uncertainties are $b = 1.29 \pm 0.05$ for the 2013 p+Pb collisions and $b = 1.18 \pm 0.05$ for Run-2 p+Pb and pp collisions [37]. Both $c_n(4)$ and $v_n(4)$ are then studied as a function of $\langle N_{\text{ch}} \rangle$.

VI. DATA ANALYSIS

The multiparticle cumulants are calculated in three steps using charged particles with $|\eta| < 2.5$. In the first step, the multiparticle correlators $\langle (2k)_n \rangle$ from Eqs. (1), (2), (8), (9), and (11) are calculated for each event from particles in one of two $p_T$ ranges, $0.3 < p_T < 3$ GeV and $0.5 < p_T < 5$ GeV.

In the second step, the correlators $\langle (2k)_n \rangle$ are averaged over events with the same $N_{\text{ch}}^{\text{rec}}$, the number of reconstructed charged particles in a given $p_T$ range, to obtain $\langle (2k)_n \rangle$ and $c_n(2k)$ from Eqs. (4), (10), and (12). In a previous study [16], it was observed that the $c_n(2k)$ values varied with the exact definition of $N_{\text{ch}}^{\text{rec}}$. This is because different definitions of $N_{\text{ch}}^{\text{rec}}$ lead to different multiplicity fluctuations and therefore different nonflow correlations associated with these multiplicity fluctuations. The observed dependence of $c_n(2k)$ on the definition of $N_{\text{ch}}^{\text{rec}}$ has been attributed to the change in the nonflow correlations when $N_{\text{ch}}^{\text{rec}}$ is changed [16].

In order to further test the sensitivity of $c_n(2k)$ to the exact definition of $N_{\text{ch}}^{\text{rec}}$, four different $p_T$ requirements are used to define $N_{\text{ch}}^{\text{rec}}$ as follows: when $\langle (2k)_n \rangle$ is calculated in the range $0.3 < p_T < 3$ GeV, $N_{\text{ch}}^{\text{rec}}$ is evaluated in four different track $p_T$ ranges: $0.3 < p_T < 3$ GeV, $p_T > 0.2$ GeV, $p_T > 0.4$ GeV, and $p_T > 0.6$ GeV. When $\langle (2k)_n \rangle$ is calculated in $0.5 < p_T < 5$ GeV, $N_{\text{ch}}^{\text{rec}}$ is evaluated in four different track $p_T$ ranges: $0.5 < p_T < 5$ GeV, $p_T > 0.2$ GeV, $p_T > 0.4$ GeV, and $p_T > 0.6$ GeV. In each case, the $c_n(2k)$ value is first calculated for events with the same $N_{\text{ch}}^{\text{rec}}$, the $c_n(2k)$ values are
then combined in the broader $N_{\text{ch}}^{\text{sel}}$ range of the event ensemble to obtain statistically significant results.

In the third step, the $c_n(2k)$ and $v_n(2k)$ values obtained for a given $N_{\text{ch}}^{\text{sel}}$ are mapped to a given $\langle N_{\text{ch}}^{\text{rec}} \rangle$, the average number of reconstructed charged particles with $p_T > 0.4$ GeV. The mapping procedure is necessary so that $c_n(2k)$ obtained for different $N_{\text{ch}}^{\text{sel}}$ can be compared using a common $x$ axis defined by $\langle N_{\text{ch}}^{\text{rec}} \rangle$. The $\langle N_{\text{ch}}^{\text{rec}} \rangle$ value is then converted to $\langle N_{\text{ch}} \rangle$, the efficiency-corrected average number of charged particles with $p_T > 0.4$ GeV, as discussed in Sec. V.

In order to account for detector inefficiencies and nonuniformity, particle weights used in Eq. (3) are defined as:

$$ w_i(\phi, \eta, p_T) = d(\phi, \eta) / e(\eta, p_T). \tag{13} $$

The additional weight factor $d(\phi, \eta)$ accounts for nonuniformities in the azimuthal acceptance of the detector as a function of $\eta$. All reconstructed charged particles with $p_T > 0.2$ GeV are entered into a two-dimensional histogram $N(\phi, \eta)$, and the weight factor is then obtained as $d(\phi, \eta) = (N(\eta))/N(\phi, \eta)$, where $N(\eta)$ is the track density averaged over $\phi$ in the given $\eta$ bin. This procedure removes most $\phi$-dependent nonuniformity from track reconstruction for any azimuthal correlation analysis [16].

VII. SYSTEMATIC UNCERTAINTIES

The main sources of systematic uncertainty are related to the detector azimuthal nonuniformity, track selection, track reconstruction efficiency, trigger efficiency, and pileup. Most of the systematic uncertainties enter the analysis through the particle weights, Eq. (13). Since $c_2(4)$ often changes sign in the low $\langle N_{\text{ch}} \rangle$ region, the absolute uncertainties (instead of relative uncertainties) in $c_2(4)$ are determined for each source. The uncertainties are typically of the order of $10^{-6}$, which translates into an absolute uncertainty of $\sqrt{10^{-6}} = 0.032$ for zero flow signal.

The effect of detector azimuthal nonuniformity is accounted for using the weight factor $d(\phi, \eta)$. The impact of the reweighting procedure is studied by fixing the weight to unity and repeating the analysis. The results are mostly consistent with the nominal results within statistical uncertainties. As a cross check, the multiparticle correlations are calculated using a mixed-event procedure, where each particle in a 2k multiplet is selected from a different event with similar $N_{\text{ch}}^{\text{rec}}$ ($|\Delta N_{\text{ch}}^{\text{rec}}| < 10$) and similar $z_{\text{vtx}}$ ($|\Delta z_{\text{vtx}}| < 10$ mm). The particle weights defined in Eq. (13) are applied for each particle forming the mixed event. The $c_2(4)$ signal obtained from the mixed events is less than $0.2 \times 10^{-6}$ in all data sets.

The systematic uncertainty associated with the track selection is estimated by tightening the $|d_0|$ and $|z_0 \sin \theta|$ requirements. For each variation, the tracking efficiency is reevaluated and the analysis is repeated. The maximum differences from the nominal results are observed to be less than $0.3 \times 10^{-6}$, $0.2 \times 10^{-6}$, and $0.1 \times 10^{-6}$ in 5.02 TeV pp, 13 TeV pp, and $p+p$ Pb collisions, respectively.

Previous measurements indicate that the azimuthal correlations (both the flow and nonflow components) have a strong dependence on $p_T$, but a relatively weak dependence on $\eta$ [10,15]. Therefore, $p_T$-dependent systematic effects in the track reconstruction efficiency could affect $c_n(2k)$ and $v_n(2k)$ values. The uncertainty in the track reconstruction efficiency is mainly due to differences in the detector conditions and material description between the simulation and the data. The efficiency uncertainty varies between 1% and 4%, depending on track $\eta$ and $p_T$ [15,16]. Its impact on multiparticle cumulants is evaluated by repeating the analysis with the tracking efficiency varied up and down by its corresponding uncertainty as a function of $p_T$. For the standard cumulant method, which is more sensitive to jets and dijets, the evaluated uncertainty amounts to $(0.1-1.5) \times 10^{-6}$ in pp collisions and less than $0.3 \times 10^{-6}$ in $p+p$ Pb collisions for $\langle N_{\text{ch}} \rangle > 50$. For the two- and three-subevent methods, the evaluated uncertainty is typically less than $0.3 \times 10^{-6}$ for most of the $\langle N_{\text{ch}} \rangle$ ranges.

Most events used in the analysis are collected with the HMT triggers with several $N_{\text{ch}}^{\text{rec}}$ thresholds. In order to estimate the possible bias due to trigger inefficiency as a function of $\langle N_{\text{ch}} \rangle$, the offline $N_{\text{ch}}^{\text{rec}}$ requirements are changed such

![FIG. 1. The $c_2(4)$ values calculated for charged particles with $0.3 < p_T < 3$ GeV (left) and $0.5 < p_T < 5$ GeV (right) with the standard cumulant method from the 13 TeV pp data. The event averaging is performed for $N_{\text{ch}}^{\text{rec}}$ calculated for various $p_T$ selections as indicated in the figure, which is then mapped to $\langle N_{\text{ch}} \rangle$, the average number of charged particles with $p_T > 0.4$ GeV. The error bars and shaded boxes represent the statistical and systematic uncertainties, respectively.](image-url)
that the HMT trigger efficiency is at least 50% or 80%. The results are obtained independently for each variation. These results are found to be consistent with each other for the two- and three-subevent methods, and show a small difference for the standard cumulant method in the low \( \langle N_{ch} \rangle \) region. The nominal analysis is performed using the 50% efficiency selection and the differences between the nominal results and those from the 80% efficiency selection are used as a systematic uncertainty. The change amounts to \((0.1–0.7) \times 10^{-6}\).

In this analysis, a pileup rejection criterion is applied to reject events containing additional vertices. In order to check the impact of residual pileup, the analysis is repeated without the pileup rejection criterion, and no difference is observed. For the 5.02 and 13 TeV pp data sets, which have relatively high pileup, the data is divided into two samples based on the \( \mu \) value: \( \mu > 0.4 \) and \( \mu < 0.4 \), and the results are compared. The average \( \mu \) values differ by a factor of two between the two samples, and the difference in \( c_2[4] \) is found to be less than \( 0.5 \times 10^{-6}\).

To check the impact of dijet events, where both jets have pseudorapidities close to the boundaries of relevant subevent regions, the three-subevent cumulants are calculated by requiring a \( \Delta \eta = 0.5 \) gap between the adjacent regions. The results are found to be consistent with the nominal result. The systematic uncertainties from different sources are added in quadrature to determine the total systematic uncertainty. The uncertainty is \((0.1–1) \times 10^{-6}\) for two- and three-subevent methods in the region \( \langle N_{ch} \rangle > 50 \), where there is a negative \( c_2[4] \) signal. The total systematic uncertainty for the standard method is typically about a factor of two larger.

The systematic uncertainty studies described above are also carried out for \( c_3[4] \), and the absolute uncertainties are found to be smaller than those for \( c_2[4] \), presumably because \( c_3[4] \) is less sensitive to the influence from dijets.

---

**FIG. 2.** The \( c_2[4] \) values calculated for charged particles with \( 0.3 < p_T < 3 \text{ GeV} \) (left) and \( 0.5 < p_T < 5 \text{ GeV} \) (right) with the two-subevent cumulant method from the 13 TeV pp data. The event averaging is performed for \( N_{ch}^{sel} \) calculated for various \( p_T \) selections as indicated in the figure, which is then mapped to \( \langle N_{ch} \rangle \), the average number of charged particles with \( p_T > 0.4 \text{ GeV} \). The error bars and shaded boxes represent the statistical and systematic uncertainties, respectively.

**FIG. 3.** The \( c_2[4] \) values calculated for charged particles with \( 0.3 < p_T < 3 \text{ GeV} \) (left) and \( 0.5 < p_T < 5 \text{ GeV} \) (right) with the three-subevent cumulant method from the 13 TeV pp data. The event averaging is performed for \( N_{ch}^{sel} \) calculated for various \( p_T \) selections as indicated in the figure, which is then mapped to \( \langle N_{ch} \rangle \), the average number of charged particles with \( p_T > 0.4 \text{ GeV} \). The error bars and shaded boxes represent the statistical and systematic uncertainties, respectively.
FIG. 4. The $c_2[4]$ values calculated for charged particles with $0.3 < p_T < 3$ GeV (left) and $0.5 < p_T < 5$ GeV (right) compared for the three cumulant methods from the 13 TeV $pp$ data. The event averaging is performed for $N_{ch}^{sel}$ 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. The dashed line indicates the $c_2[4]$ value corresponding to a 4% $v_2$ signal. The error bars and shaded boxes represent the statistical and systematic uncertainties, respectively.

VIII. RESULTS

A. Dependence on the event-class definition

This section presents the sensitivity of $c_2[4]$ to $N_{ch}^{sel}$, which defines the event class used to calculate $\langle [2]_n \rangle$ and $\langle [4]_n \rangle$ in Eqs. (10)–(12). The discussion is based on results obtained from the 13 TeV $pp$ data, but the observations for the 5.02 TeV $pp$ and $p+Pb$ data are qualitatively similar.

Figure 1 shows the $c_2[4]$ values obtained using the standard method for four event-class definitions based on $N_{ch}^{sel}$. The $c_2[4]$ values change dramatically as the event-class definition is varied, which, as points out in Ref. [23], reflects different amount of nonflow fluctuations associated with different $N_{ch}^{sel}$.

The $c_2[4]$ values for $0.3 < p_T < 3$ GeV become negative when the reference $N_{ch}^{sel}$ is obtained for $p_T > 0.4$ GeV or higher, but the four cases do not converge to the same $c_2[4]$ values. On the other hand, $c_2[4]$ values for $0.5 < p_T < 5$ GeV are always positive, independent of the definition of $N_{ch}^{sel}$. These behaviors suggest that the $c_2[4]$ values from the standard method are strongly influenced by nonflow effects in all $\langle N_{ch} \rangle$ and $p_T$ ranges. Therefore the previously observed negative $c_2[4]$ in $pp$ collisions for $0.3 < p_T < 3$ GeV and $N_{ch}^{sel}$ with $p_T > 0.4$ GeV [19] may be dominated by nonflow correlations instead of long-range collective flow.

Figure 2 shows that the $c_2[4]$ values calculated using the two-subevent method are closer to each other among different event-class definitions. The $c_2[4]$ values decrease gradually with $\langle N_{ch} \rangle$ and become negative for $\langle N_{ch} \rangle > 70$ when $c_2[4]$ is calculated in the range $0.3 < p_T < 3$ GeV range and for $\langle N_{ch} \rangle > 150$ when $c_2[4]$ is calculated in the range $0.5 < p_T < 5$ GeV. Therefore, the $c_2[4]$ values from the two-subevent method are more sensitive to long-range ridge correlations, but nevertheless may still be affected by nonflow effects, especially in the low $\langle N_{ch} \rangle$ region and higher $p_T$.

Figure 3 shows the results from the three-subevent method. For most of the $\langle N_{ch} \rangle$ range, the $c_2[4]$ values are negative, i.e., having the sign expected for long-range ridge correlations.

FIG. 5. The $c_2[4]$ values calculated for charged particles with $0.3 < p_T < 3$ GeV (left) and $0.5 < p_T < 5$ GeV (right) compared for the three cumulant methods from the 5.02 TeV $pp$ data. The event averaging is performed for $N_{ch}^{sel}$ 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. The dashed line indicates the $c_2[4]$ value corresponding to a 4% $v_2$ signal. The error bars and shaded boxes represent the statistical and systematic uncertainties, respectively.
The $c_2(4)$ values show some sensitivity to the definition of the reference $N_{\text{ch}}^{sel}$ but they are close to each other for all definitions in the region $\langle N_{\text{ch}} \rangle > 100$. This suggests that the residual nonflow effects may still be important at small $\langle N_{\text{ch}} \rangle$, but are negligible at $\langle N_{\text{ch}} \rangle > 100$. It is also observed that the $c_2(4)$ values for $0.5 < p_T < 5$ GeV are more negative than those for $0.3 < p_T < 3$ GeV, which is consistent with the observation that the $v_2$ value associated with the long-range collectivity increases with $p_T$ [10,15].

Given the relatively small dependence of $c_2(4)$ on the reference $N_{\text{ch}}^{sel}$ in the three-subevent method, the remaining discussion focuses on cases where the reference $N_{\text{ch}}^{sel}$ is calculated in the same $p_T$ ranges as those used for calculating $c_2(4)$, i.e., $0.3 < p_T < 3$ GeV and $0.5 < p_T < 5$ GeV.

### B. Comparison between different cumulant methods

Figures 4–6 show direct comparisons of the results for the standard, two-subevent, and three-subevent methods for pp collisions at $\sqrt{s} = 13$ TeV, pp at $\sqrt{s} = 5.02$ TeV, and p+Pb collisions at $\sqrt{s_{\text{NN}}} = 5.02$ TeV, respectively. The results from 5.02 TeV pp collisions are qualitatively similar to those from the 13 TeV pp collisions, i.e., the $c_2(4)$ values are smallest for the three-subevent method and largest for the standard method. The same hierarchy between the three methods is also observed in p+Pb collisions, but only for the $\langle N_{\text{ch}} \rangle < 100$ region, suggesting that nonflow effects in p+Pb collisions are much smaller than those in pp collisions at comparable $\langle N_{\text{ch}} \rangle$. In p+Pb collisions, all three methods give consistent results for $\langle N_{\text{ch}} \rangle > 100$. Furthermore, the three-subevent method
gives negative $c_2\{4\}$ values in most of the measured $\langle N_{ch} \rangle$ range.

The comparison of the $c_2\{4\}$ values between the three data sets, for the standard and the three-subevent methods, is shown in Figs. 7 and 8. The large positive $c_2\{4\}$ values observed in the small $\langle N_{ch} \rangle$ region in the standard method are likely due to nonflow correlations, since this trend is absent when using the three-subevent cumulant method. In $p+Pb$ collisions, the absolute value of $c_2\{4\}$ seems to become smaller for $\langle N_{ch} \rangle > 200$.

The same analysis is performed for the third-order harmonics. Figures 9 and 10 compare the $c_3\{4\}$ values between the three data sets for the standard cumulant method and the three-subevent method. The $c_3\{4\}$ values from the three-subevent method are close to zero in all three systems. For the standard method, the positive $c_3\{4\}$ values in the small $\langle N_{ch} \rangle$ region indicate the influence of nonflow correlations, but the influence is not as strong as that for $c_2\{4\}$.

Figure 11 shows the $c_3\{4\}$ values from $p+Pb$ collisions in the two $p_T$ ranges, obtained with the three-subevent method;
they are zoomed-in version of the $p+\text{Pb}$ data shown in Figs. 8 and 9. Within their large statistical and systematic uncertainties, the values of $c_1[4]$ are systematically below zero, especially for $0.5 < p_T < 5$ GeV, where the $c_1[4]$ values are comparable to $-0.16 \times 10^{-6}$, corresponding to a $v_3$ value of 2% as indicated in the figure. The negative $c_1[4]$ values from the three-subevent method support the existence of long-range multiparticle triangular flow in $p+\text{Pb}$ collisions.

C. Three-subevent flow harmonic $v_2[4]$

The harmonic flow coefficients $v_2[4]$ can be obtained from the measured values of $c_2[4]$ according to Eq. (7). Figure 12 shows the $v_2[4]$ values for charged particles with $0.3 < p_T < 3$ GeV calculated using the three-subevent method in the three data sets. Results for the higher $p_T$ range ($0.5 < p_T < 5$ GeV) are presented in Fig. 13. The value of $v_2[4]$ is measured down to $\langle N_{ch}\rangle \approx 50$ in $pp$ collisions and down to $\langle N_{ch}\rangle \approx 20–40$ in $p+\text{Pb}$ collisions. The $v_2[4]$ values are observed to be approximately independent of $\langle N_{ch}\rangle$ in the measured range in the three data sets: $50 < \langle N_{ch}\rangle < 150$ for 5.02 TeV $pp$, $50 < \langle N_{ch}\rangle < 200$ for 13 TeV $pp$, and $20 < \langle N_{ch}\rangle < 380$ for 5.02 TeV $p+\text{Pb}$, respectively. Moreover, the $p+\text{Pb}$ data suggest the value of $v_2[4]$ is lower for $\langle N_{ch}\rangle > 200$, as expected from the similar behavior of $|c_2[4]|$ in Figs. 7 and 8 at large $\langle N_{ch}\rangle$.

FIG. 10. The $c_1[4]$ values calculated for charged particles with $0.5 < p_T < 5$ GeV using the standard cumulants (left) and the three-subevent method (right) compared between 5.02 TeV $pp$, 13 TeV $pp$, and 5.02 TeV $p+\text{Pb}$. The event averaging is performed for $N_{ch}^{sel}$ 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. The error bars and shaded boxes represent the statistical and systematic uncertainties, respectively.

FIG. 11. The $c_1[4]$ values calculated for charged particles with $0.3 < p_T < 3$ GeV (left) or $0.5 < p_T < 5$ GeV (right) with the three-subevent cumulant method for the $p+\text{Pb}$ data. The event averaging is performed for $N_{ch}^{sel}$ calculated for various $p_T$ selections as indicated in the figure, which is then mapped to $\langle N_{ch}\rangle$, the average number of charged particles with $p_T > 0.4$ GeV. The dashed line indicates the $c_1[4]$ value corresponding to a 2% $v_3$ signal. The error bars and shaded boxes represent the statistical and systematic uncertainties, respectively.
The values of $v_2[4]$ presented in Figs. 12 and 13 are also compared to the values of $v_2[2]$ obtained from the 2PC measurements [10,15] where the nonflow effects are estimated using low-multiplicity events ($\langle N_{ch} \rangle < 20$) and then subtracted. The subtraction was performed either by a template fit, which includes the pedestal level from the $\langle N_{ch} \rangle < 20$ events, or by a peripheral subtraction, which sets the pedestal level by a zero-yield at minimum (ZYAM) procedure [6]. The peripheral subtraction explicitly assumes that the most peripheral events do not contain any long-range correlations [15], and so $v_2$ is forced to be zero at the corresponding $\langle N_{ch} \rangle$ value, which biases $v_2$ to a lower value in other multiplicity ranges.

D. Dependence on the number of sources in the initial state

Figures 12 and 13 show that the $v_2[4]$ values are smaller than the $v_2[2]$ values extracted using the template-fit method in both the $pp$ and $p+Pb$ collisions. In various hydrodynamic models for small collision systems [38,39], this difference can be interpreted as the influence of event-by-event flow fluctuations associated with the initial state, which is closely related to the effective number of sources $N_s$ for particle production in the transverse density distribution of the initial state [39]:

$$\frac{v_2[4]}{v_2[2]} = \left( \frac{4}{3 + N_s} \right)^{1/4} \quad \text{or} \quad N_s = \frac{4v_2[2]^4}{v_2[4]^4} - 3. \quad (14)$$

Figure 14 shows the extracted values of $N_s$ as a function of $\langle N_{ch} \rangle$ in 13 TeV $pp$ and 5.02 TeV $p+Pb$ collisions, estimated using charged particles with $0.3 < p_T < 3$ GeV and $0.5 < p_T < 5$ GeV. It is observed that the $N_s$ value increases with $\langle N_{ch} \rangle$ in $p+Pb$ collisions, reaching $N_s \approx 20$ in the highest multiplicity class, and it is consistent between the two $p_T$ ranges.

In the model framework in Refs. [38,39], the values of $|c_2[4]|$ and $v_2[4]$ are expected to decrease for large $N_s$, which is compatible with the presented results. The slight decreases of $|c_2[4]|$ shown in Figs. 7 and 8 for $p+Pb$ collisions are compatible with the model predictions. The results for 13 TeV $pp$ collisions cover a limited $\langle N_{ch} \rangle$ range compared to $p+Pb$, but agree with $p+Pb$ collisions in this range.
The subevent cumulant technique and the new results provide direct evidence for the presence of long-range multiparticle azimuthal correlations in broad \(\langle N_{ch} \rangle\) ranges in \(pp\) and \(p+Pb\) collisions, and these long-range multiparticle correlations persist even in events with rather low multiplicity of \(\langle N_{ch} \rangle \sim 40\). The \(c_2(4)\) values are consistent with zero in \(pp\) collisions, but are systematically below zero in \(p+Pb\) collisions, compatible with the presence of significant long-range multiparticle triangular flow in \(p+Pb\) collisions.

The single-particle harmonic coefficient \(v_2(4) = (-c_2(4))^{1/4}\) is calculated and compared with \(v_2(2)\) obtained previously using the two-particle correlation method, where the nonflow contributions were estimated and subtracted. The magnitude of \(v_2(4)\) is smaller than that for \(v_2(2)\), as expected for a long-range final-state hydrodynamic collective effect. The ratio of \(v_2(4)\) to \(v_2(2)\) is used, in a model-dependent framework, to infer the number of particle-emitting sources in the initial-state geometric configuration. The number of sources extracted within this framework is found to increase with \(\langle N_{ch} \rangle\) in \(p+Pb\) collisions.

The subevent cumulant technique and the new results provide direct evidence that the ridge is indeed a long-range collective phenomenon involving many particles distributed across a broad rapidity interval. The results of \(v_2(4)\) and its dependence on \(p_T\) and \(\langle N_{ch} \rangle\), largely free from nonflow effects, can be used to understand the space-time dynamics and the properties of the medium created in small collision systems.

ACKNOWLEDGMENTS

We thank CERN for the very successful operation of the LHC, as well as the support staff from our institutions without whom ATLAS could not be operated efficiently. We acknowledge the support of ANPCyT, Argentina; YerPhI, Armenia; ARC, Australia; BMWFW and FWF, Austria; ANAS, Azerbaijan; SSTC, Belarus; CNPq and FAPESP, Brazil; NSERC, NRC and CFI, Canada; CERN; CONICYT, Chile; CAS,


aAlso at Department of Physics, King’s College London, London, United Kingdom.
bAlso at Institute of Physics, Azerbaijan Academy of Sciences, Baku, Azerbaijan.
cAlso at Novosibirsk State University, Novosibirsk, Russia.
dAlso at TRIUMF, Vancouver BC, Canada.
eAlso at Department of Physics & Astronomy, University of Louisville, Louisville, KY, USA.
fAlso at Physics Department, An-Najah National University, Nablus, Palestine.
gAlso at Department of Physics, California State University, Fresno CA, USA.
hAlso at Department of Physics, University of Fribourg, Fribourg, Switzerland.
iAlso at II Physikalisches Institut, Georg-August-Universität, Göttingen, Germany.
jAlso at Departamento de Fisica de la Universitat Autonoma de Barcelona, Barcelona, Spain.
kAlso at Departamento de Fisica e Astronomia, Faculdade de Ciencias, Universidade do Porto, Portugal.
lAlso at Tomsk State University, Tomsk, and Moscow Institute of Physics and Technology State University, Dolgoprudny, Russia.
mAlso at The Collaborative Innovation Center of Quantum Matter (CICQM), Beijing, China.
nAlso at the Centro de Calculo de l’Institut National de Physique Nuclrale et de Physique des Particules (IN2P3), Villeurbanne, France.
Academia Sinica Grid Computing, Institute of Physics, Academia Sinica, Taipei, Taiwan.

Also at...