Measurement of the azimuthal anisotropy of charged-particle production in Xe + Xe collisions at $\sqrt{s_{NN}} = 5.44$ TeV with the ATLAS detector

ATLAS Collaboration; Aad, G.

DOI
10.1103/PhysRevC.101.024906

Publication date
2020

Document Version
Final published version

Published in
Physical Review C

License
CC BY

Citation for published version (APA):
Measurement of the azimuthal anisotropy of charged-particle production in Xe + Xe collisions at $\sqrt{s_{NN}} = 5.44$ TeV with the ATLAS detector

G. Aad et al.*
(ATLAS Collaboration)

(Received 13 November 2019; accepted 16 December 2019; published 10 February 2020)

This paper describes the measurements of flow harmonics $v_2$--$v_6$ in 3 $\mu$b$^{-1}$ of Xe + Xe collisions at $\sqrt{s_{NN}} = 5.44$ TeV performed using the ATLAS detector at the Large Hadron Collider (LHC). Measurements of the centrality, multiplicity, and $p_T$ dependence of the $v_n$ obtained using two-particle correlations and the scalar product technique are presented. The measurements are also performed using a template-fit procedure, which was developed to remove nonflow correlations in small collision systems. This nonflow removal is shown to have a significant influence on the measured $v_n$ at high $p_T$, especially in peripheral events. Comparisons of the measured $v_n$ with measurements in Pb + Pb collisions and $p +$ Pb collisions at $\sqrt{s_{NN}} = 5.02$ TeV are also presented. The $v_n$ values in Xe + Xe collisions are observed to be larger than those in Pb + Pb collisions for $n = 2, 3, 4$ in the most central events. However, with decreasing centrality or increasing harmonic order $n$, the $v_n$ values in Xe + Xe collisions become smaller than those in Pb + Pb collisions. The $v_n$ in Xe + Xe and Pb + Pb collisions are also compared as a function of the mean number of participating nucleons, $\langle N_{\text{part}} \rangle$, and the measured charged-particle multiplicity in the detector. The $v_2$ values in Xe + Xe and Pb + Pb collisions are observed to be similar at the same $\langle N_{\text{part}} \rangle$ or multiplicity, but the other harmonics are significantly different. The ratios of the measured $v_n$ in Xe + Xe and Pb + Pb collisions, as a function of centrality, are also compared to theoretical calculations.

DOI: 10.1103/PhysRevC.101.024906

I. INTRODUCTION

Heavy-ion collisions, such as those at the Relativistic Heavy Ion Collider (RHIC) [1–4] and at the Large Hadron Collider (LHC) [5–10], produce a state of matter, with de-confined quarks and gluons, commonly called quark-gluon plasma (QGP). The QGP produced in such collisions expands anisotropically due to spatial anisotropies in the initial geometry, which produce asymmetric pressure gradients between the medium and the outside vacuum. The anisotropic expansion leads to large azimuthal modulations in the final distributions of the produced particles [11]. The single-particle azimuthal yields of particles produced in heavy-ion collisions are typically characterized as a Fourier series [12]:

$$\frac{dN}{d\phi} = \frac{N_0}{2\pi} \left[ 1 + 2 \sum_{n=1}^{\infty} v_n \cos[n(\phi - \Phi_n)] \right],$$

where $\phi$ is the azimuthal angle of the particle momentum and $v_n$ and $\Phi_n$ are the magnitude and phase of the $n$th-order anisotropy. The $v_n$ are referred to as flow harmonics, while the $\Phi_n$ are referred to as event-plane angles. The $v_n$ are functions of the transverse momentum ($p_T$), pseudorapidity$^1$ ($\eta$), event multiplicity, and particle species, and they fluctuate from event to event.

Measurements of the $v_n$ and their comparisons with calculations based on relativistic hydrodynamics have shown that the QGP produced in heavy-ion collisions behaves like a nearly perfect fluid, characterized by a very low ratio of shear viscosity to entropy density, $\eta/s$, close to the conjectured lower limit of $\frac{\pi}{43}$ [13]. Significant experimental progress has been made in recent years in precision measurements of the $v_2$ [5–8,14,15], event-by-event fluctuations in the $v_n$ [16–18], and the correlations between the magnitudes [19,20] and phases [21] of different-order anisotropies. These high-precision measurements have led to significant improvement in constraining the value of $\eta/s$ [22,23]. However, due to large theoretical uncertainties in the understanding of the initial stages of heavy-ion collisions, the precise value of $\eta/s$ or its exact temperature dependence still remains unknown. A possible data-driven way of further constraining the value of $\eta/s$, and at the same time improving the understanding of

$^1$ATLAS 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, with $\phi$ being the azimuthal angle around the $z$ axis. The pseudorapidity is defined in terms of the polar angle $\theta$ as $\eta = -\ln \tan(\theta/2)$. 

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.

2469-9985/2020/101(2)/024906(35) 024906-1 ©2020 CERN, for the ATLAS Collaboration
the initial entropy production in heavy-ion collisions, is to compare the $v_n$ measured across collision systems of different sizes [24].

This paper presents $v_n$ measurements in Xe+Xe collisions at a center-of-mass energy per nucleon pair of $\sqrt{s_{NN}} = 5.44$ TeV by ATLAS using an integrated luminosity of $3 \mu$b$^{-1}$. The flow measurements are performed with the two-particle correlation (2PC) and scalar product (SP) methods. The measurements are also performed using a template-fit procedure [25,26] developed by ATLAS to measure correlations in small systems, such as $pp$ and $p + $ Pb. The template-fit procedure removes “nonflow” correlations that arise from back-to-back jet pairs (dijets) and particle decays, which typically bias the $v_n$ measurements in low-multiplicity events, especially at high $p_T$.

The Xe nucleus is small compared to the Pb nucleus—the specific isotopes used at the LHC are $^{129}$Xe and $^{208}$Pb. Thus, Xe+Xe collisions are expected to have larger event-by-event fluctuations in the initial geometry compared to Pb+Pb collisions [24]. On the other hand, a smaller system implies larger viscous effects in the hydrodynamic expansion of the produced QGP fireball [24,27,28]. Thus, the $v_n$ measurements in Xe+Xe collisions and their comparison with those in Pb+Pb collisions allow the interplay of these two effects to be studied.

The outline of the paper is as follows: Section II describes the ATLAS detector subsystems used in this measurement. Section III describes the dataset, the event selection, and requirements on charged-particle tracks used in the analysis. Section IV gives a brief description of the 2PC, template-fit, and SP methods. The systematic uncertainties in the $v_n$ measurements are described in Sec. V. In Section VI, the results of the measurements are presented, which include the $p_T$, centrality, and multiplicity dependence of the $v_n$ as well as its dependence on the mean number of participating nucleons in the collisions ($\langle N_{\text{part}} \rangle$). In Sec. VII, the present Xe+Xe $v_n$ measurements are compared with previous $v_n$ measurements in Pb+Pb and $p + $ Pb collisions. Finally, Sec. VIII, summarizes the results.

II. THE ATLAS DETECTOR

The measurements presented in this paper were performed using the ATLAS detector [29] at the LHC. The principal components used in this analysis are the inner detector (ID), calorimeter, and the trigger and data acquisition systems.

The ID, consisting of a silicon pixel detector, a silicon microstrip tracker, and a transition radiation tracker, is immersed in a 2 T axial magnetic field. The ID provides charged-particle tracking in the range $|\eta| < 2.5$. The high-granularity silicon pixel detector covers the interaction region and typically provides four measurements per track. The first hit is normally in the “insertable B layer” (IBL) [30,31], which was installed before the 2015 data taking period. The pixel detector is followed by the silicon microstrip tracker (SCT), which typically provides measurements of four three-dimensional points per track. These silicon detectors are complemented by the transition radiation tracker, which enables radially extended track reconstruction up to $|\eta| = 2.0$, providing around 30 hits per track. The ATLAS calorimeter system consists of a liquid argon (LAr) electromagnetic calorimeter covering $|\eta| < 3.2$, a steel-scintillator sampling hadronic calorimeter covering $|\eta| < 1.7$, a LAr hadronic calorimeter covering $1.5 < |\eta| < 3.2$, and two LAr electromagnetic and hadronic forward calorimeters (FCal) covering $3.2 < |\eta| < 4.9$. The ATLAS trigger system [32] consists of a first-level (L1) trigger implemented using a combination of dedicated electronics and programmable logic, and a software-based high-level trigger (HLT).

III. DATASET, EVENT, AND TRACK SELECTIONS

The Xe + Xe data used in this paper were collected in October 2017. A set of minimum-bias events was selected by a pair of complementary and mutually exclusive triggers. The first trigger required the total transverse energy deposited in the calorimeters at L1 ($E_T^{L1}$) to be larger than 4 GeV, without any additional requirement at the HLT. The second trigger required that $E_T^{L1}$ be less than 4 GeV with the additional requirement of a reconstructed track with $p_T > 0.2$ GeV at the HLT. Together, these two triggers selected all events with either $E_T^{L1} > 4$ GeV or a reconstructed track at the HLT.

In the offline analysis, the $z$ coordinate of the primary vertex [33] is required to be within 10 cm of the nominal interaction point. Events containing more than one inelastic interaction (pileup events) were removed by exploiting the correlation between the transverse energy measured in the FCal ($\Sigma E_T^{FCal}$) and the number of tracks associated with the primary vertex. In a typical pileup event, the track multiplicity associated with the primary vertex belongs to a single Xe+Xe collision, while the energy deposited in calorimeters contains contributions from multiple collisions. Therefore, events with small values of measured multiplicity and large $\Sigma E_T^{FCal}$, which differ markedly from those of the majority of Xe+Xe collisions, are removed from the analysis [17]. The fraction of pileup events is estimated to be $\approx 0.1\%$. As in previous ATLAS heavy-ion analyses, the events are classified into centrality percentiles based on the total transverse energy deposited in the FCal in the event [5,14]. The Glauber model [34] is used to obtain a correspondence between the $\Sigma E_T^{FCal}$ distribution and the sampling fraction of the total inelastic Xe + Xe cross section, allowing the setting of the centrality percentiles [5,14]. The Glauber model is also used to obtain the mapping from the observed $\Sigma E_T^{FCal}$ to the primary properties, such as the mean number of nucleons participating in the nuclear collision, $\langle N_{\text{part}} \rangle$, for each centrality interval. Figure 1 shows the distribution of $\Sigma E_T^{FCal}$ in data and thresholds for the selection of several centrality intervals. For the centrality dependence study, this analysis is restricted to the 0–80% most central collisions where the triggers are fully efficient.

Charged-particle tracks are reconstructed from the signals in the ID. A reconstruction procedure developed for tracking in dense environments in $pp$ collisions, and optimized for heavy-ion collisions, is used for this purpose [35]. In the analysis, the set of reconstructed tracks is filtered using several selection criteria. For the nominal selection, the tracks are required to have $p_T > 0.5$ GeV, $|\eta| < 2.5$, at least two pixel hits, with the additional requirement of a hit in the IBL when one
is expected, at least eight SCT hits, and no missing hits in the pixel or SCT. A hit is expected if the extrapolated track crosses an active region of a silicon-sensor module (pixel or SCT) that has not been disabled, and a hit is said to be “missing” when it is expected but not found. Further, the transverse ($d_0$) and longitudinal ($z_0 \sin \theta$) impact parameters of the track relative to the primary vertex are required to be less than 1 mm. The track-fit quality parameter $\chi^2/\text{ndof}$ is required to be less than 6. For cross checks and for estimating systematic uncertainties, a looser and a tighter set of requirements are used. For the looser selection, referred to as “loose,” the requirements on the number of pixel and SCT hits are reduced to one and six, respectively, and the requirement on $d_0$ and $z_0 \sin \theta$ are relaxed to 1.5 mm. These looser requirements on the tracks are also used to study the multiplicity dependence of the $v_n$ as they enhance the per-event multiplicity. For the tighter selection, referred to as “tight,” the topological requirements on the reconstructed track are not altered, but the transverse and longitudinal impact parameters of the track are restricted to be less than 0.5 mm.

In order to study the performance of the ATLAS detector, a sample of 1M minimum-bias Xe + Xe Monte Carlo (MC) events was generated using HIJING version 1.38b [36]. Since the HIJING event generator does not have any intrinsic mechanism to generate flow, the latter is added after the generation using an “afterburner” procedure [37], which slightly shifts the $\phi$ positions of generated particles to mimic flow. The generated sample was passed through a full simulation of the ATLAS detector [38] using GEANT4 [39], and the MC events were reconstructed by the same algorithms as the data. The reconstructed particles in the MC events are used to calculate the reconstruction efficiency—the fraction of the generated charged particles that are successfully reconstructed—as a function of $p_T$ and $\eta$, and denoted by $\epsilon(p_T, \eta)$ below. At midrapidity ($|\eta| < 1$) and for events in the 0–5% centrality interval, the reconstruction efficiency is $\approx 60\%$ at low $p_T$ and increases to $\approx 73\%$ at higher $p_T$. For $|\eta| > 1$, the efficiency decreases to about 40–60% depending on the $p_T$ and centrality. The reconstruction efficiency depends weakly on the centrality for low-$p_T$ tracks, for which it is smaller in the most central events by about 3% as compared to peripheral events. For tracks with $p_T > 1$ GeV, the dependence on centrality is less than 1%. The MC simulation is also used to calculate the fraction of fake tracks—the fraction of the reconstructed particles that do not correspond to any generated particle—also as a function of $p_T$ and $\eta$, and denoted by $f(p_T, \eta)$ below. The fake rate is less than 1% across the $p_T$ and centrality range used in this analysis. Additionally, systematic differences seen between the $v_n$ evaluated using the reconstructed particles in the MC events—using the same techniques as in the data analysis—and the $v_n$ implemented for the generated particles, are used as multiplicative correction factors in the data analysis. The size of this correction is discussed in Sec. V.

The Xe+Xe measurements in this paper are compared with similar measurements in Pb+Pb and $p+Pb$ collisions at $\sqrt{s_{NN}} = 5.02$ TeV. Further, for the template fits, data from $pp$ collisions at $\sqrt{s} = 5.02$ TeV are used [26] (see Sec. IV B). The Pb+Pb dataset, triggers, and track selections used for the comparisons are identical to those used in Ref. [6]. The $pp$ and $p+Pb$ datasets, triggers, and track selections are identical to those used in Ref. [26].

IV. METHODOLOGY

In this section, a brief description of the methods used to measure flow harmonics is provided. First, the two-particle correlation analysis, including the construction of correlation functions, is described. Next, the template-fit procedure is explained, which starts from the 2PCs but implements additional nonflow subtraction. Finally, the steps involved in obtaining the SP results are presented.

A. Two-particle correlations

The 2PC method has been used extensively for flow measurements at RHIC and the LHC [6,14,16,19,25,26,40–42]. In the 2PC method, the distribution of particle pairs in relative azimuthal angle $\Delta \phi = \phi^a - \phi^b$ and pseudorapidity separation $\Delta \eta = \eta^a - \eta^b$ is measured. The particles $a$ and $b$ are conventionally referred to as the “reference” and “associated” particles, respectively. In this analysis, the two particles are charged particles measured by the ATLAS tracking system, over the full azimuth and $|\Delta \eta| = 5\%$ units in $\Delta \eta$. In order to account for the detector acceptance effects, such as localized regions with lower efficiency or dead regions, the correlation is constructed from the ratio of the distribution in which the reference and associated particles are taken from the same event to the distribution in which the reference and associated particles are taken from two different events. The correlation function is defined as

$$ C(\Delta \eta, \Delta \phi) = \frac{S(\Delta \eta, \Delta \phi)}{B(\Delta \eta, \Delta \phi)}, $$

where $S$ and $B$ represent the “same-event” and “mixed-event” pair distributions, respectively [14]. The same-event distribution includes both the physical correlations and the
correlations arising from detector acceptance effects. On the other hand, the mixed-event distribution reflects only the effects of detector inefficiencies and nonuniformity, but contains no physical correlations. Detector acceptance effects largely cancel out in the S/B ratio [40]. To ensure that the acceptance effects in the B distribution closely match those in the S distribution, the B distribution is constructed from particles from two different events that have similar centrality (or multiplicity) and z vertex. When constructing S and B, in order to account for the inefficiency in track reconstruction and for misconstructed tracks, pairs are weighted by their fake-track rates and the inverse product of their reconstruction efficiencies (see Sec. III):

\[ [1 - f\left(p_T^a, \eta^a\right)][1 - f\left(p_T^b, \eta^b\right)]/[\epsilon\left(p_T^a, \eta^a\right)\epsilon\left(p_T^b, \eta^b\right)].\]

To investigate the \(\Delta f\) dependence of the long-range (|\(\Delta \eta| > 2\)) correlation in more detail, the S and B distributions are integrated over \(2 < \Delta \eta < 5\) and a one-dimensional correlation function \(C(\Delta \phi)\) is constructed as follows:

\[ C(\Delta \phi) = \frac{\int_2^5 d|\Delta \eta| S(|\Delta \eta|, \Delta \phi)}{\int_2^5 d|\Delta \eta| B(|\Delta \eta|, \Delta \phi)} = \frac{S(\Delta \phi)}{B(\Delta \phi)}.\]

The \(|\Delta \eta| > 2\) requirement is imposed to reject the short-range correlations and focus on the long-range features of the correlation functions. In a similar fashion to the single-particle distribution [Eq. (1)], the \(C(\Delta \phi)\) can be expanded as a Fourier series [14]:

\[ C(\Delta \phi) = C_0[1 + 2\Sigma_{n=1}^\infty v_{n,n}(p_T^a, p_T^b) \cos(n\Delta \phi)]. \]

In the absence of any nonflow correlations, it can be shown that the Fourier coefficients of the \(C(\Delta \phi)\) factorize into the product of single-particle anisotropies as [40]

\[ v_{n,n}(p_T^a, p_T^b) = v_n(p_T^a) v_n(p_T^b). \]

The factorization of \(v_{n,n}\) given by Eq. (3) is expected to break down at high \(p_T\) and in low-multiplicity events, where the \(v_{n,n}\) separation between the particles is small and short-range correlations dominate [14]. However, the \(|\Delta \eta| > 2\) requirement removes most of such short-range correlations. In the phase-space region where Eq. (3) holds, the \(v_n(p_T^a)\) can be evaluated from the measured \(v_{n,n}\) as

\[ v_n(p_T^a) = \frac{v_{n,n}(p_T^a, p_T^b)}{v_n(p_T^a)} = \frac{v_{n,n}(p_T^a, p_T^b)}{\sqrt{v_{n,n}(p_T^a, p_T^b)}}, \]

where the relation \(v_{n,n}(p_T^a, p_T^b) = v_n^2(p_T^a)\) is used in the denominator. For most of the 2PC results in this analysis the \(v_{n,n}(p_T^a, p_T^b)\) are evaluated using Eq. (4) with \(0.5 \text{ GeV} < p_T^a < 5 \text{ GeV}\). The upper limit on \(p_T^a\) is chosen to exclude high-\(p_T\) particles, which originate predominantly from jets.

### B. Template fits

One drawback of the 2PC method is that in peripheral events with low multiplicity or for 2PCs involving particles at high \(p_T\), the measured \(v_n\) can be biased by correlations arising from back-to-back dijets that are not rejected by the \(|\Delta \eta| > 2\) requirement. This issue is much more severe in smaller collision systems such as \(p + \text{Pb}\) and especially in \(pp\) collisions, in which the 2PC even at large \(|\Delta \eta|\) is completely dominated by the back-to-back dijet correlations. To address this issue, a template-fitting procedure was developed to measure the \(v_n\) in \(pp\) and \(p + \text{Pb}\) collisions [25,26,43]. The template-fit method assumes the following:

1. The shape of the dijet contribution to the 2PC does not change from low- to high-multiplicity events; only its relative contribution to the 2PC changes.
2. The 2PC for low-multiplicity events is dominated by the dijet contribution.

With the above assumptions, the correlation \(C(\Delta \phi)\) in higher multiplicity events is then described by a template fit, \(C_{\text{template}}(\Delta \phi)\), consisting of two components: a scale factor, \(F\), times the correlation measured in low-multiplicity events, \(C_{\text{periph}}(\Delta \phi)\), which accounts for the dijet correlation, and a genuine long-range harmonic modulation, \(C_{\text{ridge}}(\Delta \phi)\):

\[ C_{\text{template}}(\Delta \phi) = F \cdot C_{\text{periph}}(\Delta \phi) + C_{\text{ridge}}(\Delta \phi). \]

where the coefficient \(F\) and the \(v_{n,n}\) are fit parameters adjusted to reproduce the \(C(\Delta \phi)\). The coefficient \(G\) is not a free parameter but is fixed by the requirement that the integrals of the \(C_{\text{template}}(\Delta \phi)\) and \(C(\Delta \phi)\) are equal. There are two variations of the template-fit procedure, depending on how the \(C_{\text{periph}}(\Delta \phi)\), used in Eq. (5), is obtained. In the first case, the entire correlation function in the low-multiplicity interval is used as \(C_{\text{periph}}(\Delta \phi)\). In the second case, only the modulated part of the correlation function is used as \(C_{\text{periph}}(\Delta \phi)\) and the unmodulated “pedestal” is removed. The removal of the pedestal is done using the “zero yield at minimum” (ZYAM) procedure [25], which subtracts a pedestal from the \(C_{\text{periph}}(\Delta \phi)\) such that the value of \(C_{\text{periph}}(\Delta \phi)\) is zero at its minimum. The ZYAM-based method includes the additional assumption that only the part of the low-multiplicity 2PC that is modulated in \(\Delta \phi\) arises from “nonflow” correlations. In general, the standard and ZYAM-based template-fit measurements in \(pp\) and \(p + \text{Pb}\) collisions give significantly different \(v_n\) values at very low multiplicities. But with increasing multiplicities, the difference between the \(v_n\) values obtained from the two methods decreases [25,26]. Since the \(p + \text{Pb}\) and \(\text{Xe} + \text{Xe}\) multiplicities are significantly larger than in \(pp\) and \(p + \text{Pb}\), it is expected that the standard and the ZYAM-based template measurements yield similar \(v_n\) values, except at very low multiplicities. The difference between the \(v_n\) values obtained with and without the ZYAM procedure then gives an estimate of the bias that the nonflow subtraction technique induces in the measured \(v_n\).

In this paper, the \(C_{\text{periph}}(\Delta \phi)\) used in the template fits is constructed using \(pp\) events at \(\sqrt{s} = 5.02\ \text{TeV}\) that have fewer than 20 reconstructed tracks passing the selection requirements listed in Sec. III. Prior measurements using the template-fit method in \(pp\) and \(p + \text{Pb}\) collisions used the low-multiplicity events from the same collision system to generate
the $C_{\text{periph}}^{n}(\Delta \phi)$. In the present analysis, the choice of $pp$ reference is based on the following reasoning. The template-fit method works better when the $C_{\text{periph}}^{n}(\Delta \phi)$ is dominated by nonflow correlations. Even in peripheral $A + A$ collisions, there is still significant flow compared to $pp$ collisions; thus, using the $C_{\text{periph}}^{n}(\Delta \phi)$ constructed from $pp$ collisions at similar collision energy, which has smaller flow-like correlations, is a better alternative. Because of “jet quenching” effects present in heavy-ion collisions [44–47], it is possible that the assumption made in the template-fit procedure, that the shape of the dijet correlation in $\Delta \phi$ does not change from low- to high-multiplicity events, may not be valid. The effects of jet quenching may bias the $v_n$ measured using the template-fit method.

C. Scalar product

While the 2PC method relies on correlations between particle pairs using only information from the ID to measure $v_n$, the SP [48,49] measurement relies on correlations between the flow from energy deposits measured in the FCal and from tracks in the ID. Thus, it allows measurements of the $v_n$ with a larger gap in $\eta$ to strongly suppresses short-range correlations. In fact, the larger $\eta$ gap not only suppresses short-range correlations, it also suppresses correlations from back-to-back dijets, as most of the dijets are at midrapidity.

The SP measurement is based on the construction of flow vectors [48,49] from tracks in the ID, and towers—segmentations of the calorimeter of approximate granularity $0.1 \times 0.1$ in $\eta$ and $\phi$—in the FCal. The flow-vectors are defined as follows:

$$k_n = \frac{1}{\sum_j w_j} \sum_j w_j e^{i \phi_j},$$

where $n$ is the harmonic order. For the construction of $k_n$ from ID tracks, which is labeled as $q_n$, in the following text, $\phi_j$ is the azimuthal angle of the track, and the weight $w_j$, which corrects for tracking performance, is equal to $[1 - f(p_T^j, \eta^j)]/\epsilon(p_T^j, \eta^j)$. The sum runs over a set of particles in a single event, usually restricted to a region of the $p_T-\eta$ space. For the estimate of $k_n$ from the FCal, denoted by $Q_n$, the sum runs over the calorimeter towers, with $\phi_j$ being the azimuthal position of the tower and the weight $w_j$ being the measured $E_T$ in the tower.

Using Eq. (1), it can be shown that the flow vectors $q_n$ are given by $v_n(p_T, \eta)e^{i \phi_n}$. However, due to statistical fluctuations arising from the finite number of tracks used in measuring the $q_n$ event by event [16], the measured $q_n$ fluctuate around the true $v_n(p_T, \eta)e^{i \phi_n}$, and can be written as

$$q_n = v_n e^{i \phi_n} + q_{n\text{fluc}},$$

where $q_{n\text{fluc}}$ is a complex number representing event-by-event statistical fluctuations. Similarly, the $Q_n$ can be written as

$$Q_n = V_n e^{i \phi_n} + Q_{n\text{fluc}},$$

where $V_n$ is used to denote the integrated $v_n$ of particles in the calorimeter acceptance together with the $E_T$ response of the calorimeter folded in, and $Q_{n\text{fluc}}$ represents statistical noise.

Because of the random orientation of the collision geometry from event to event, the flow vector averaged over many events should be equal to zero. Additionally, the distributions for the real and imaginary parts of the flow vector should have identical widths. However, due to nonuniform detector response in $\phi$, these conditions may not be satisfied. To correct for the nonuniform detector response, a procedure described in Ref. [6] is applied, which ensures that the distributions of the real and imaginary parts of the flow vectors are centered at zero and have the same widths.

In the estimation of $v_n$ from the SP method, four flow vectors are involved: $Q_n^+$ measured in the FCal at positive $\eta$, $Q_n^-$ measured in the FCal at negative $\eta$, and the corresponding flow vectors for charged particles measured in the ID, denoted by $q_n^+$ for positive $\eta$ and $q_n^-$ for negative $\eta$. The $v_n$ for $\eta < 0$ from the SP method is then defined as

$$v_n^{\text{SP}}(\eta < 0) = \frac{\text{Re}(q_n^+ Q_n^-)}{\sqrt{|Q_n^+ Q_n^-|^2}},$$

while for $\eta > 0$ the numerator is replaced by $(q_n^- Q_n^+)$. The “$^+$” denotes complex conjugate and the angular brackets indicate an average over all events. Equation (8) can be understood by substituting expressions from Eqs. (6) and (7) for the flow vectors and noting that all terms involving $q_{n\text{fluc}}^+$ and $Q_{n\text{fluc}}^-$ drop out when averaged over many events. This gives

$$v_n^{\text{SP}}(\eta < 0) = \frac{(v_n^+ V_n^-)}{\sqrt{|V_n^+ V_n^-|^2}},$$

where the superscripts $N$ and $P$ indicate whether the quantities involved correspond to $\eta < 0$ or $\eta > 0$, respectively. Equation (9) is similar to Eq. (4) for the 2PC measurements. The final $v_n$ from the SP method is obtained by averaging the results obtained for $\eta > 0$ and $\eta < 0$ in Eq. (8). While Eq. (8) explicitly uses the real part of the flow-vector product to obtain the $v_n$ with the imaginary part of the flow-vector product should be statistically consistent with zero [cf. Eq. (9)]. Any statistically significant nonzero value for the imaginary part of the flow-vector product in Eq. (8) is indicative of detector response effects that are not corrected for in the measurement and is typically used as a systematic uncertainty of the measured $v_n$.

The 2PC and SP methods are very closely related; both methods nominally measure the $\langle \sqrt{v_n^2} \rangle$, where the average is taken over all particles and events in the chosen $p_T$, $\eta$, and centrality range. The 2PC method uses correlation between pairs of tracks while the SP method utilizes correlations between tracks and energy deposition in the calorimeters. One advantage of the SP method in ATLAS is that the larger rapidity gap between the FCal and the ID, which is greater
TABLE I. The relative contributions (in percent) to the systematic uncertainty of $v_n$ in two selected bins of centrality. The contributions are expressed as a percentage of the measured $v_n$ and are rounded up to two significant digits. Items 1–4 are common to all methods. Item 5 is specific to the 2PC and template-fit methods, item 6 is specific to the template-fit method, and items 7 and 8 are specific to the SP method only. The 2PC and template-fit methods are used to measure harmonics $v_2$–$v_5$, while the SP method is used to measure harmonics $v_2$–$v_6$.

<table>
<thead>
<tr>
<th>Uncertainty sources</th>
<th>Harmonic order</th>
<th>5–10% centrality</th>
<th>40–50% centrality</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>0.8–1 GeV</td>
<td>6–8 GeV</td>
</tr>
<tr>
<td>1. Track selection</td>
<td>$v_2$–$v_6$</td>
<td>1.5</td>
<td>1</td>
</tr>
<tr>
<td>2. Tracking efficiency</td>
<td>$v_2$–$v_6$</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>3. Centrality determination</td>
<td>$v_2$–$v_6$</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>4. MC closure</td>
<td>$v_2$–$v_6$</td>
<td>&lt;0.1</td>
<td>&lt;0.1</td>
</tr>
<tr>
<td></td>
<td>$v_2$</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>$v_3$</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>$v_4$</td>
<td>1</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td>$v_5$</td>
<td>4</td>
<td>10</td>
</tr>
<tr>
<td></td>
<td>$v_2$</td>
<td>&lt;0.5</td>
<td>1.5</td>
</tr>
<tr>
<td></td>
<td>$v_3$</td>
<td>0.7</td>
<td>3.5</td>
</tr>
<tr>
<td>5. Event mixing</td>
<td>$v_4$</td>
<td>1.0</td>
<td>14</td>
</tr>
<tr>
<td></td>
<td>$v_5$</td>
<td>5</td>
<td>30</td>
</tr>
<tr>
<td></td>
<td>$v_2$</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>6. Peripheral reference</td>
<td>$v_3$</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>7. $\eta$ asymmetry</td>
<td>$v_4$</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>$v_5$–$v_6$</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td>$v_2$–$v_3$</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>8. Residual sine term</td>
<td>$v_4$</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>$v_5$–$v_6$</td>
<td>10</td>
<td>10</td>
</tr>
</tbody>
</table>

than 3.2 units, leads to larger suppression of nonflow correlations as compared to the 2PC method, where the minimum gap between the reference and associated tracks is 2 units. However, the SP method (in ATLAS) has some disadvantages: For example, the larger $\eta$ gap biases the $v_n$ measurements if longitudinal flow decorrelations are present [50–52]. The flow decorrelations increase with decreasing system-size [50–52] and would affect flow measurements for Xe+Xe more than for Pb+Pb. Additionally, for very low multiplicities, it is very difficult to obtain reliable flow vectors using the calorimeters. This is the reason why for small systems ($pp$ and $p + Pb$) the flow measurements are typically performed using the 2PC or template-fit methods [25,26,43,53,54].

V. SYSTEMATIC UNCERTAINTIES

The systematic uncertainties of the measured $v_n$ are evaluated by varying several aspects of the analysis. Most of the uncertainties are common to the SP and 2PC methods and are discussed together. Since the template-fit measurements start with the 2PCs, all systematic uncertainties related to the 2PC method also affect the measurements based on the template fit. The template-fit measurements have an additional uncertainty related to the choice of the peripheral reference, which is discussed below. The SP method is used to measure harmonics $v_2$–$v_6$, while the 2PC and template-fit methods are used to measure harmonics $v_2$–$v_5$. This difference is due to the fact that the 2PC and template-fit measurements have large systematic uncertainties attributed to the pair acceptance and peripheral reference choice for $v_6$, which make a significant measurement not feasible. The uncertainties for two representative centrality and $p_T$ ranges are summarized in Table I. The following sources of uncertainties are considered:

1. Track selection: The tracking selection cuts control the relative contribution of genuine charged particles and fake tracks entering the analysis. The stability of the results with respect to the track selection is evaluated by varying the requirements imposed on the reconstructed tracks, and including the resulting variation in the $v_n$ as a systematic uncertainty. The two sets of variations termed loose and tight in Sec. III are used for this purpose. At low $p_T$ ($0.5$–$0.8$ GeV), the variation in the $v_n$ obtained from this procedure is most significant in the most central events, typically 5%, as the fake-track rate is largest in this region of phase space. For higher $p_T$ and less central events, changing the set of tracks used in the analysis has less influence on the measurement.

2. Tracking efficiency: As mentioned in Sec. IV, the tracks are weighted by $(1 - f)/\epsilon$ when calculating the $v_n$ to account for the impact of the tracking efficiency. Uncertainties in the efficiency, resulting from, e.g., uncertainty of the amount of material in the detector, are propagated into the measured $v_n$. This uncertainty is evaluated by varying the efficiency up and down within its uncertainties ($\approx \pm 3\%$) in a $p_T$-dependent manner and re-evaluating the $v_n$. This contribution to the overall uncertainty is very small and amounts to less than 1% on average for the $p_T$-integrated $v_n$, and is negligible for the $p_T$-differential $v_n$. This is because the change of efficiency largely cancels out in the
differential $v_n(p_T)$ measurement, and for $v_n$ integrated over $p_T$, the low-$p_T$ particles dominate the measurement. The uncertainty does not change significantly with centrality or with the harmonic order.

Centrality determination: The centrality definitions used to classify the events into centrality percentiles have an ≈1% uncertainty associated with them. This uncertainty arises from uncertainties in the fraction of the inelastic Xe+Xe cross section accepted by the triggers used in this analysis. The impact of the uncertainty from the centrality definition on the $v_n$ is evaluated by varying the centrality interval definitions by 1%, re-evaluating the $v_n$, and including the variation in the $v_n$ as a systematic uncertainty. The impact on all harmonics over the 0–50% centrality range is found to be within 1%. For more peripheral events, this number varies between 1% and 5% depending on the $p_T$, centrality, and harmonic order $n$.

MC closure: The MC closure test consists of comparing the $v_n^{\text{MC}}$ obtained directly from the MC generated particles, and the $v_n^{\text{MC}}$ obtained by applying the same analysis procedures to the MC sample as to the data. The analysis of MC events is done to evaluate the contributions of effects not corrected for in the data analysis. Systematic differences seen between the $v_n$ of the generated particles and reconstructed particles are used to correct the $v_n$ measured in the data and, conservatively, also included as a systematic uncertainty. This uncertainty is at the level of a few percent over the 0.5–0.8 GeV $p_T$ range and the 0–20% centrality range, and reaches 5% for $p_T \sim 0.5$ GeV in the 0–5% centrality interval. It is negligible elsewhere.

Event mixing: As explained in Sec. IV A, the 2PC analysis uses the event-mixing technique to estimate and correct for the detector acceptance effects. Potential systematic uncertainties in the $v_n$ due to the residual pair-acceptance effects, which are not corrected for by the mixed events, are evaluated by varying the multiplicity and $z$-vertex matching criteria used to make the mixed-event distributions, following procedures from Ref. [14]. The resulting uncertainty in the $v_2$–$v_5$ is 1–3% for most of the centrality and $p_T$ ranges considered in this paper. However, at high $p_T$ and for the harmonic orders $n \geq 4$, this uncertainty can become at large as 10%. This uncertainty only contributes to the $v_n$ values measured by the 2PC and template-fit methods.

Choice of peripheral reference: The template-fit procedure uses $pp$ events at $\sqrt{s} = 5.02$ TeV with less than 20 reconstructed tracks to build $C_{\text{template}}(\Delta \phi)$. The choice of 20 tracks is partially motivated by the fact that the mean multiplicity of minimum-bias $pp$ events is close to 20 tracks. To test the stability of the $v_n$ with respect to our choice of the peripheral reference, the analysis is repeated with an alternative $C_{\text{template}}(\Delta \phi)$ constructed from $pp$ events with 0–20, 10–20, and 10–30 reconstructed tracks and the change in the template $v_n$ values is included as a systematic uncertainty. This uncertainty is within ≈4% over the 0–50% centrality range and for $p_T < 4$ GeV, but increases considerably and can become as large as 30% for more peripheral events or at higher $p_T$. This uncertainty only contributes to the $v_n$ values measured by the template-fit method.

$\eta$ asymmetry: Because of the symmetry of the Xe + Xe collision system, the event-averaged $\langle v_\eta(\eta) \rangle$ and $\langle v_\eta(-\eta) \rangle$ are expected to be equal. Any difference between the event-averaged $v_\eta$ at ±$\eta$ arises from residual detector nonuniformity. The difference between the $v_\eta$ values measured in opposite hemispheres is treated as a systematic uncertainty quantifying imperfect detector performance. This uncertainty only contributes to the $v_\eta$ values measured with the SP method. The value of this uncertainty depends on the harmonic order. It is less than 1% for $v_2$ and increases to ≈5% for $v_6$. For the 2PC method, the residual nonuniformity is estimated by variations in the event-mixing procedure.

Residual sine term: The imaginary part of the flow vector product in Eq. (8) should be consistent with zero. Any systematic deviation of the imaginary part from zero indicates the presence of residual detector response effects. The ratio of the imaginary part of
FIG. 3. Two-particle correlations in $\Delta\phi$ for $2 < |\Delta| < 5$ and $2 \text{ GeV} < p_{T}^{a,b} < 3 \text{ GeV}$. The correlations are shown for the 0–5% (left), 30–40% (center), and 60–70% (right) centrality intervals. The lines represent Fourier fits to the correlation functions [Eq. (2)] that include harmonics up to $n = 5$.

the flow-vector product to its real part [Eq. (8)] is therefore included as a contribution to the systematic uncertainty. The contribution from this source is $\leq 2\%$ in most of the phase space, while for the higher harmonics ($n = 5$) it can reach 10%. This uncertainty is only relevant for the $v_n$ values measured by the SP method.

VI. RESULTS

A. Two-particle correlations and template fits

Figure 2 shows the two-dimensional 2PCs in $\Delta\eta$–$\Delta\phi$ for several centrality intervals for $2 \text{ GeV} < p_{T}^{a,b} < 3 \text{ GeV}$, where $p_T^a$ and $p_T^b$ label the $p_T$ ranges of the reference and associated particles used in the correlation. In all cases, a peak is seen in the correlation at $\sim (0, 0)$. This peak arises from short-range correlations such as jet fragmentation, resonance decays, or Hanbury Brown and Twiss (HBT) correlations [55]. The long-range (large $\Delta\eta$) correlations are the result of the global anisotropy of the event and are the focus of the study in this analysis. Thus, all further 2PC results are discussed with the $\Delta\eta > 2$ requirement.

Figure 3 shows corresponding one-dimensional 2PCs along $\Delta\phi$ for $2 \text{ GeV} < p_{T}^{a,b} < 3 \text{ GeV}$ and for several centrality intervals. The line is a Fourier fit to the correlation [Eq. (2)] that includes harmonics up to $n = 5$. The $y$-axis ranges for the different panels are kept identical so that the modulation in the correlation across the different centralities can be easily compared. It is seen that the modulation in the correlation is smallest in the most central events (left panel), increases toward midcentral events, and then decreases slightly. This roughly follows the centrality dependence of most $v_{n,n}$, especially the $v_{2,2}$.

Figure 4 shows the same correlation functions but with a reduced $y$-axis range to make it easier to observe the features of the correlation. The Fourier fit is indicated by the thick black line and the contributions of the individual $v_{n,n}$ are also shown. In the most central 0–5% collisions (left panel of Fig. 4), the $v_{2,2} - v_{4,4}$ are of comparable magnitude, but in other centralities, where the average collision geometry is elliptic,
FIG. 5. Template fits to the 2PCs in Xe + Xe collisions. The fits are shown for the 0–5% (left), 30–40% (center), and 60–70% (right) centrality intervals. The plots are for 0.5 GeV < $p_T^a, b$ < 5 GeV. The solid points indicate the measured 2PC, and the open points and curves show different components of the template (see legend) that are shifted along the y axis by $G$ or by $FC^{\text{norm}}(0)$, where necessary, for presentation. The plots correspond to the template-fit method without using ZYAM subtraction on the $pp$ reference. The template-fit includes harmonics $n$ up to 5.

FIG. 6. The demonstration of factorization as a function of $p_T$ for $v_2$ in the 40–50% centrality interval (top row) and 60–80% centrality interval (bottom row). The left panels show the results for the 2PC method and the right panels show the results for the template-fit method without ZYAM. The error bars indicate statistical uncertainties only.
the $v_{2,2}$ is significantly larger than the other $v_{n,n}$ ($n \geq 3$). In central events, the away-side peak (at $\Delta \phi = \pi$) is also much broader due to the comparable magnitudes of the higher order harmonics and $v_{2,2}$, while in midcentral events (30–40% centrality in Fig. 4), the near ($\Delta \phi = 0$) and away-side peaks are more symmetric as the $v_{2,2}$ dominates. In central and midcentral events, the near-side peak is larger than the away-side peak. However, beyond 60% centrality (right panel of Fig. 4), the away-side peak becomes larger, indicating the presence of a large negative $v_{1,1}$ component. This negative $v_{1,1}$ component in the peripheral 2PCs largely arises from dijets and its contribution to the correlation function increases in
FIG. 8. The $p_T$ dependence of the $v_n$ obtained with the 2PC and template-fit measurements. Each panel represents a different centrality interval. The solid markers show the results obtained using the 2PC analysis. The small and large open markers show the results obtained from the template fit with and without ZYAM, respectively. The vertical error bars and shaded bands indicate statistical and systematic uncertainties, respectively, and are often too small to be visible.

The presence of the jets also results in the breakdown of the factorization relation [Eq. (4)]. This is discussed in more detail in Sec. VI B.

Figure 5 shows examples of template fits to the measured 2PCs [Eq. (5)] that includes harmonics $n$ up to 5. The figure shows the template fits without a ZYAM subtraction.
FIG. 9. The centrality dependence of the $v_n$ obtained with the 2PC and template-fit methods. The solid points represent the 2PC $v_n$. The large open points represent the template fit $v_n$ without ZYAM. The small open points represent the template fit $v_n$ with ZYAM. Each panel shows the results for a different $p_T$ interval. The error bars and shaded bands indicate statistical and systematic uncertainties, respectively.

performed on the $pp$ reference. However, in the plots presented in Fig. 5, the template fits look identical with and without ZYAM as explained in the following. The ZYAM method sets the minimum of the $C_{\text{periph}}(\Delta \phi)$ term to zero by removing a pedestal equal to $C_{\text{periph}}(0)$, since the minimum of $C_{\text{periph}}(\Delta \phi)$ occurs at $\Delta \phi = 0$. The template fits then transfer this pedestal to the $G$ term of the $C_{\text{ridge}}(\Delta \phi)$. However, the components of the template fit are shifted vertically by the pedestal of the other term, i.e., the $C_{\text{ridge}}(\Delta \phi)$ is plotted as $"C_{\text{ridge}}(\Delta \phi) + FC_{\text{periph}}(0)"$ and the $C_{\text{periph}}(\Delta \phi)$ is plotted as $"FC_{\text{periph}}(\Delta \phi) + G"$, where $G$ and $FC_{\text{periph}}(0)$ are the pedestals of the two components. This shift in the plotting undoes the shift in the pedestals of the two components. In Fig. 5, the fits are shown for a central (0–5%), a midcentral (30–40%), and a peripheral (60–70%) centrality interval. The relative difference between $C_{\text{temp}}(\Delta \phi)$ and $C_{\text{ridge}}(\Delta \phi)$ is largest for the peripheral centrality interval and decreases for midcentral events. This is indicative of larger contributions from nonflow correlations in peripheral collisions than in midcentral collisions.

B. Factorization of $v_n$ for 2PC and template fits

Figure 6 shows how well the factorization of the $v_n$ [Eq. (4)] works for the 2PC and template-fit $v_n$ measurements. The plots show the $v_2(p_T)$ obtained using Eq. (4) in the 40–50% and 60–80% centrality ranges, for several $p_T$ intervals: 0.5–1, 1–2, 2–3, 3–4, and 0.5–5 GeV. If factorization worked, then the obtained values of $v_2(p_T)$ would be independent of the choice of $p_T^z$. For the 2PC measurements, factorization breaks down at large values of $p_T^z$ or $p_T^H$. The exact values of $p_T$ at which the breakdown occurs depend on the centrality. The breakdown in factorization occurs mostly because of nonflow effects such as jets. In midcentral events the $v_2$ is largest, and thus the bias from nonflow effects is relatively weak and the factorization holds until much higher $p_T$ values.
than in other centrality ranges. In peripheral events, with the decreased multiplicity, the bias from nonflow effects becomes larger; this, coupled with the decreased \( v_2 \), results in larger or earlier onset (in \( p_T \)) of factorization breakdown. In general, the factorization holds better for \( v_3 \), \( v_4 \), and \( v_5 \) than for \( v_2 \). This is because the factorization breakdown occurs mostly because of the back-to-back dijets, not rejected by the \( \Delta \eta \) cut, which have a larger impact on the \( v_2 \) than the \( n \geq 3 \) harmonics \([14,26]\). These observations are similar to those made in previous \( v_n \) measurements in Pb + Pb collisions at \( \sqrt{s_{NN}} = 5.02 \) TeV \([6]\). However, quantitatively, the breakdown in the Xe + Xe case is somewhat larger than in Pb + Pb collisions at the same centrality, due to smaller multiplicities.

On the other hand, for the template-fit measurements, the
factorization works much better for higher $p_T$ and for more peripheral events, indicating a significant reduction of the bias from dijet correlations.

For all remaining plots involving the 2PC or template-fit measurements, the default $p_T$ interval used for the $v_n$ measurement is $0.5 \text{ GeV} < p_T < 5 \text{ GeV}$. Over this $p_T$ interval, the factorization holds quite well for the 2PC method for most of the centrality ranges.

C. $p_T$ dependence of $v_n$

Figure 7 shows the $p_T$ dependence of the $v_n$ obtained using the 2PC and SP methods. Each panel shows the measurements for a different centrality interval. In general, the 2PC and SP results are quite comparable, although small differences can be seen in the most central events. However, in peripheral events (50–60% and 60–70% centrality classes) and for $p_T > 4 \text{ GeV}$ the $v_2$ from the 2PC method gives systematically higher

---

**FIG. 12.** Comparison of the $p_T$ scale factors for the Xe + Xe and Pb + Pb data as a function of centrality. The left panel shows the comparisons for $v_2$ and the right panel shows the comparisons for $v_3$. The error bars on the scale factors indicate fit uncertainties.

**FIG. 13.** The $p_T$ dependence of scaled flow harmonics, $v_n/v_2^{n/2}$, as a function of centrality in Xe + Xe collisions. The error bars denote statistical errors and the shaded bands represent the systematic uncertainty. The horizontal dashed lines represent fits to a constant, over the 0.6–5 GeV $p_T$ range. The $v_n$ values are measured with the SP method.
values. This increase arises due to bias from dijets and is further investigated below.

Figure 8 compares the $p_T$ dependence of the $v_n$ obtained using the 2PC method, and the template-fit measurement with or without the ZYAM subtraction. The measurements are shown in finer $p_T$ intervals, thereby allowing examination of the $p_T$ dependence and the nonflow bias in detail. For the harmonic $v_5$ the results are truncated for centralities more peripheral than the 40–50% interval, beyond which the statistical errors are typically too large to study the $p_T$ dependence. For all the methods of measurement, the following trends are observed: The $v_n$ values increase at low $p_T$ and reach a maximum between 2 GeV and 4 GeV, and then decrease. For nearly all centralities, the $v_n$ follow the trend $v_2 \gg v_3 > v_4 > v_5$. This hierarchy is violated in the most central 0–5% collisions, where the $v_3$ and $v_4$ values at high enough $p_T$ become larger than the $v_2$ values. These trends are the same as those observed in prior $v_n$ measurements in Pb+Pb collisions [6,14]. For the 2PC measurements, it is observed that at high $p_T$ (the last two $p_T$ intervals shown), the $v_2$ again increases with $p_T$. This feature can be explained as a bias from dijets, which dominate the 2PC at high $p_T$ especially in peripheral events. The back-to-back nature of the dijet correlations enhances (suppresses) the measured values of the

FIG. 14. The $p_T$ dependence of scaled flow harmonics, $v_n/v_3^{n/3}$, as a function of centrality in Xe + Xe collisions. The error bars denote statistical errors and the shaded bands represent the systematic uncertainty. The horizontal dashed lines represent fits to a constant, over the 0.6–5 GeV $p_T$ range. The $v_n$ values are measured with the SP method.

FIG. 15. The centrality dependence of mean values of scaled harmonics (left) $v_n/v_2^{n/2}$ and (right) $v_n/v_3^{n/3}$ calculated in the 0.6–5 GeV $p_T$ range. The errors represent the uncertainty of the fit and do not account for systematic uncertainties.
FIG. 16. Comparisons of the $v_n$ measured in Xe + Xe collisions with those measured in Pb + Pb collisions at $\sqrt{s_{NN}} = 5.02$ TeV. The results are plotted as a function of centrality and obtained using the 2PC method. From top to bottom, each row corresponds to a different $n$. The right panels show the ratio of the Xe + Xe to the Pb + Pb $v_n$ values. The plots are for $0.5 < p_T < 5$ GeV. The error bars and bands represent statistical and systematic uncertainties, respectively. For the ratio plots, the correlated systematic uncertainties between the Xe + Xe and Pb + Pb results are taken into account.

$v_n$ for even-order (odd-order) harmonics. This trend is to some extent also seen in the SP measurements shown previously in Fig. 7. This bias is considerably reduced in the measurements based on the template-fit method. It is interesting to note that the template-fit measurements with or without ZYAM give similar results for the $v_n$. As mentioned earlier, the differences between the $v_n$ values obtained with or without the ZYAM procedure can give an estimate of possible biases in the nonflow subtraction. Thus, similar values obtained from the two template-fit methods, which make different assumptions about the $C_{\text{periph}}(\Delta \phi)$ [Eq. (5)], give more confidence in the robustness of the nonflow subtraction.

D. Centrality dependence of $v_n$

Figure 9 shows the centrality dependence of the $v_n$ obtained from the 2PC and template-fit methods. The different panels represent different $p_T$ intervals. Results from the template-fit method with or without a ZYAM procedure on the reference and from the 2PC method are shown. For the 2PC
measurements and for \( p_T < 3 \text{ GeV} \)—shown in the first three panels—the \( v_n \) values increase from central to midcentral events, reach a maximum between 40% and 60% centrality, and then decrease. The variation with centrality is largest for the \( v_2 \) owing to the large change in the second-order eccentricity \[56\] from central to midcentral collisions. These features are consistent with those observed in other heavy-ion collisions \[1–8,14\]. The effects of the nonflow bias seen in Fig. 8 can also be observed more clearly in Fig. 9, where for the 4–6 GeV \( p_T \) interval the 2PC \( v_2 \) values increase over the 60–80% centrality range. The template-fit \( v_2 \), on the other hand, does not show this increase. In general, for centrality intervals more peripheral than 60% and \( p_T > 3 \text{ GeV} \), the template-fit \( v_2 \) have smaller values for the even-order harmonics and larger values for the odd-order harmonics, compared to the 2PC results. This is consistent with the removal of bias from dijets, which enhances (suppresses) even-order (odd-order) harmonics.

### E. Scaling behavior of \( v_n \)

#### 1. Scaling of \( v_n(p_T) \) across different centralities

In a recent ATLAS paper \[6\] on \( v_n \) measurements in Pb+Pb collisions, it was observed that the \( v_n(p_T) \) at different centralities in Pb+Pb collisions had a very similar \( p_T \) dependence. In fact, after scaling the \( p_T \) and \( v_n \) axes by multiplicative factors that depend only on centrality, the \( v_n(p_T) \) across different centralities were consistent within \( \approx 5\% \) over the 0.5–5 GeV \( p_T \) range. In this section \( p_T \), a similar study is presented for the \( v_n \) measured in Xe+Xe collisions.

The left panels of Fig. 10 show the \( v_n(p_T) \) for several centrality intervals and for \( n = 2 \) and 3 obtained via the 2PC method. The \( v_n(p_T) \) are then scaled along the \( x \) and \( y \) axes to match their shapes across the different centrality intervals. For this matching, the 0–60% centrality interval is chosen as the reference and the \( v_n(p_T) \) for the individual centralities are scaled to best match the \( v_n(p_T) \) in the 0–60% centrality interval over the 0.5–5 GeV \( p_T \) range. The fitting is done over the 0.5–5 GeV \( p_T \) range, with the \( \chi^2 \) minimization taking into account the combined statistical and systematic uncertainties and treating the scale factors along the \( p_T \) and \( v_n \) axes as the fit parameters. The right panels of Fig. 10 show the corresponding scaled \( v_n \). The scaled \( v_n \) match within \( \pm 4\% \) across most of the shown \( p_T \) range.

Figure 11 shows the \( p_T \) and \( v_n \) scale factors obtained for \( v_2 \) and \( v_3 \) as a function of centrality. The \( p_T \) scale factors increase from central to peripheral events over the 10–60% centrality range. Over this interval, the values of the scale factors for the two harmonics match within \( \pm 8\% \). However, in the most central events (0–10%), some significant deviation is observed between the two scale factors, where the scale factor for the \( v_2 \) increases for more central events while that for the \( v_3 \) decreases. This deviation may be due to larger jet-bias effects in the \( v_2 \) as compared to \( v_3 \), but requires additional investigation to properly understand its origin. The \( v_n \) scale factors are quite different between the two harmonics and show a much larger variation for \( v_2 \) compared to \( v_3 \). This scale factor is mostly affected by the changing collision geometry; the ellipticity of the collision geometry changes considerably with the collision centrality, leading to significantly larger variations in the scale factor for \( v_2 \) compared to that for \( v_3 \).

Finally, Fig. 12 compares the \( p_T \) scale factors obtained in Xe+Xe collisions with those obtained in Pb+Pb collisions from Ref. \[6\], as a function of centrality. The scale factors follow a very similar trend as a function of centrality. In fact, for \( v_3 \) the values of the \( p_T \) scale factor are typically consistent within \( \approx 2\% \). The comparison of this scaling and the corresponding scale factors for Xe+Xe and Pb+Pb collisions should be able to provide additional constraints on theoretical models.

#### 2. Scaling of \( v_n(p_T) \) with the harmonic order \( n \)

Another scaling previously observed in the \( v_n \) measurements in Au+Au and Pb+Pb collisions is that the \( v_n \) as a function of \( p_T \) qualitatively follow the power-law relation \( v_n/(v_2)^{n/m} = \text{constant} \) \[14, 57–59\]. This scaling is demonstrated in Fig. 13 for the \( v_n \) in Xe+Xe collisions, for \( n = 3–6 \) and \( m = 2 \), using the \( v_n \) values obtained with the SP method. It is observed that at least in the 0.6–5 GeV \( p_T \) range the ratios do not depend on the transverse momentum. The scaling often holds to considerably higher \( p_T \) values, depending on the harmonic order and the centrality interval. The mean values of \( v_n/(v_2)^{n/2} \) are large in the most central events and...
FIG. 18. Comparisons of the $v_n$ as a function of $\langle N_{\text{part}} \rangle$ measured in Xe + Xe collisions with those measured in Pb + Pb collisions at $\sqrt{s_{NN}} = 5.02$ TeV. Each panel corresponds to a different harmonic order $n$. The plots are for the 2PC method and for 0.5 GeV < $p_T$ < 5 GeV. The error bars and bands represent statistical and systematic uncertainties, respectively.

The $v_n$s become much smaller, even less than one, in the peripheral events. There is also a clear ordering of the ratios over the 0–30% centrality range, with $v_6/v_2^{1/2}$ the largest and $v_3/v_2^{1/2}$ the smallest. The lack of $p_T$ dependence of harmonics scaled by $v_2$ suggests testing for scaling with $v_3$. In Fig. 14, the flow harmonics scaled by triangular flow, $v_n/v_3^{1/3}$, are shown. In this case, not only are the scaled values constant as a function of $p_T$, but they also have similar values for different $n$. In addition, the mean values of $v_n/v_3^{1/3}$ are very similar for all centralities. These observations are summarized in Fig. 15, where the mean values of $v_n/v_2^{1/2}$ and $v_n/v_3^{1/3}$ are shown as a function of centrality. The harmonics scaled by elliptic flow have a considerable variation when going from central to peripheral events. This is because the elliptic flow has considerable centrality dependence, while the other harmonics vary slowly with centrality. Thus, the other harmonics scaled by the elliptic flow also demonstrate considerable centrality dependence. In contrast, scaling by $v_3$ gives a $v_n/v_3^{1/3}$ value of about 1.4 for all centralities and harmonics of all orders larger than $n = 3$. On the basis of the latter scaling, after measuring $v_3$ it is possible to also predict $v_n$ for $n = 4$ up to at least $n = 6$. As mentioned in Ref. [59], the origin of

FIG. 19. Comparison of the $v_n$ obtained with the SP method, as a function of $N_{\text{part}}$, for Xe + Xe collisions at $\sqrt{s_{NN}} = 5.44$ TeV (open markers) and Pb + Pb collisions at $\sqrt{s_{NN}} = 5.02$ TeV (solid markers). The left panel shows the comparison for $v_2$–$v_3$ and the right panel shows the comparison for $v_4$–$v_6$. The error bars and bands represent statistical and systematic uncertainties, respectively.
FIG. 20. Comparisons of the $v_n$ as a function of $p_T$ between Pb + Pb collisions at $\sqrt{s_{NN}} = 5.02$ TeV and Xe + Xe at $\sqrt{s_{NN}} = 5.44$ TeV. The results are for the 0–5% centrality interval and are obtained with the 2PC method. From top to bottom, each row corresponds to a different harmonic order $n$. The right panels show the ratio of the Xe + Xe to the Pb + Pb $v_n$ values. The error bars and bands represent statistical and systematic uncertainties, respectively. For the ratio plots, the correlated systematic uncertainties between the Xe + Xe and Pb + Pb results are taken into account.

this scaling dependence is not well understood. Further insight from theoretical calculations into its origin would be useful.

VII. COMPARISON WITH Pb+Pb MEASUREMENTS AT $\sqrt{s_{NN}} = 5.02$ TeV

Figure 16 shows comparisons of the $v_n$ measured in Xe+Xe collisions at $\sqrt{s_{NN}} = 5.44$ TeV with the corresponding measurements in Pb+Pb collisions at $\sqrt{s_{NN}} = 5.02$ TeV, from Ref. [6], as a function of centrality. The comparisons are for measurements performed using the 2PC method. For $v_2$–$v_4$, the figures also show comparisons with theoretical calculations from Ref. [24]. Figure 17 shows similar comparisons for the SP method. In the 0–5% central events, the Xe+Xe values for $v_2$ are $\approx 35\%$ larger than the Pb+Pb values. This is seen most clearly from the ratio plots shown in Fig. 16. For less central collisions, the ratio decreases and becomes smaller than one by the 10–15% centrality interval.
FIG. 21. Comparisons of the $v_n$ measured in different collision systems as a function of multiplicity. The left panel compares the $v_n$ obtained with the template-fit method for the Xe + Xe measurements with those obtained in Pb + Pb collisions at $\sqrt{s_{NN}} = 5.02$ TeV. The right panel includes comparisons with measurements in $p + Pb$ at $\sqrt{s_{NN}} = 5.02$ TeV, shown over a limited multiplicity range. The plots are for $0.5 \text{GeV} < p_T^b < 5 \text{GeV}$. The error bars and bands represent statistical and systematic uncertainties, respectively.

For more peripheral events, the ratio keeps decreasing, but at a lower rate, and becomes roughly 0.9 by the 50–60% centrality interval. For $v_3$, the Xe+Xe values are larger than the Pb+Pb values over the 0–30% centrality interval, become comparable to the Pb+Pb values over the 30–40% centrality interval, and become smaller than the Pb+Pb values for more peripheral events. In this case, the ratio in the 0–5% most central events is smaller than for $v_2$, and decreases almost linearly over the 0–70% centrality range. For $v_4$, the Xe+Xe values are only marginally larger in the top 0–5% central events. The ratio for $v_4$ becomes comparable to or less than one by the 5–10% centrality and continues to decrease for more peripheral events. For $v_5$, the Xe+Xe values are smaller throughout. For $v_2$–$v_4$, the measured ratios are quite consistent with the theoretical predictions from Ref. [24] and consistent with measurements from the ALICE and CMS Collaborations [9,10]. The observed trends can be explained as follows. Since Xe+Xe is a smaller system than Pb+Pb, the impact of fluctuations is considerably more important. The fluctuations increase the initial eccentricities of the collision geometry and this effect contributes to enhancing the $v_n$ in Xe+Xe compared to Pb+Pb at the same centrality. Additionally, the Xe nucleus may have a slight prolate deformation [60], which would lead to larger $v_2$ in most central Xe+Xe events, compared to Pb+Pb. However, because Xe+Xe is a smaller collision system, the viscous effects, which suppress the $v_n$, are larger and play a bigger role with increasing harmonic order and for less central events. In the most central events, the effect of the increased fluctuations wins for $v_2$. But with increasing harmonic order and/or for less central collisions, eventually, the viscous effects lower the $v_n$ values in Xe+Xe compared to the $v_n$ values in Pb+Pb. Over the centrality and $p_T$ range shown in Fig. 16, the template-fit measurements give results that are similar to those of the 2PC method and are not shown separately, for brevity.

Figures 18 and 19 show the same results but as a function of $N_{\text{part}}$. In general, at the same $N_{\text{part}}$, the Xe+Xe and Pb+Pb $v_n$ are rather different. One exception is the $v_3$, whose values
are comparable between Xe+Xe and Pb+Pb for $N_{\text{part}} < 200$. The difference for the $v_2$ arises because of the very different shapes of the average collision geometry between Xe+Xe and Pb+Pb at the same $N_{\text{part}}$. At the same $N_{\text{part}}$, the Pb+Pb events correspond to collisions with significantly larger elliptic deformation. On the other hand, the $v_3$ is driven purely by event-by-event fluctuations in the initial geometry, and these fluctuations are expected to depend mostly on $N_{\text{part}}$, leading to similar $v_3$ values at the same $N_{\text{part}}$. The higher order harmonics for $n > 3$ are driven by event-by-event fluctuations, as well as nonlinear response to lower order eccentricities [19,21]. Since the second-order eccentricities for the Xe+Xe and Pb+Pb collisions are considerably different at the same $N_{\text{part}}$, the nonlinear responses to the second-order eccentricities, which contribute to the $v_4$ and $v_5$, are different. The $v_4$ and $v_5$ are thus not expected to be similar at the same $N_{\text{part}}$. The difference in the $v_3$ values at the highest Xe+Xe $N_{\text{part}}$ values is intriguing and may be related to centrality fluctuations discussed in Refs. [61,62]. Further input from theoretical calculations to understand this feature would be useful.

Figure 20 shows a similar comparison but as a function of $p_T$ for the 0–5% centrality interval. The plots show that the ratio remains roughly constant over the 0.5–2 GeV $p_T$ range and then decreases for all harmonics. This is indicative of increasing viscous effects and/or breakdown of the hydrodynamic description beyond $\approx 2$ GeV, which lead to the lowering of the ratio.

The left panel of Fig. 21 shows the comparison of the $v_n$ in Xe+Xe and Pb+Pb as a function of the event multiplicity, $N_{\text{ch}}$, defined as the number of reconstructed tracks passing the loose track selection requirements listed in Sec. III. Results are shown for the template-fit method. The right panel of Fig. 21 shows similar results including measurements in $p + \text{Pb}$ but over a limited range. For the Pb+Pb and $p + \text{Pb}$ template-fit measurements, the $C_{n_{\text{eff}}}^\text{rec}(\Delta p)$ [Eq. (5)] is constructed using $p p$ events within the $N_{\text{ch}} < 20$ multiplicity range, similar to the Xe+Xe case. The trends follow what was observed in the $N_{\text{part}}$ dependence: The $v_2$ values are quite similar between the two systems at the same multiplicity over a large multiplicity range (up to $\approx 800$), but the values for the other harmonics are different. The $p + \text{Pb}$ $v_n$ values are smaller than those for Xe+Xe throughout the overlapping $N_{\text{ch}}$ range. The $N_{n_{\text{ch}}}^{\text{rec}}$ dependence of the $v_n$ in $p + \text{Pb}$ collisions is weaker than in Xe+Xe and Pb+Pb collisions. It is interesting to note that the differences between the values (as a fraction) are smallest for the $v_2$, possibly indicating that at similar multiplicity densities the geometry fluctuations in $p + \text{Pb}$ collisions are similar to those in heavy-ion collisions.

\section*{VIII. SUMMARY}

This paper presents ATLAS measurements of azimuthal anisotropy of charged particles in Xe+Xe collisions at $\sqrt{s_{\text{NN}}} = 5.44$ TeV produced by the LHC. The measurements are performed using an integrated luminosity of 3 $\mu b^{-1}$. The azimuthal anisotropies, quantified by the flow harmonics $v_n$, are measured using the SP, 2PC, and template-fit methods for $n = 2–6$. The measurements are performed over wide transverse momentum and centrality ranges. All harmonics show a similar $p_T$ dependence, first increasing with $p_T$ up to a maximum around 3–4 GeV and then decreasing for higher $p_T$. Significant values of the second-order harmonic $v_2$ persist up to 20 GeV. The elliptic flow signal ($v_2$) is strongly dependent on the event centrality and it is largest in midcentral events (30–50%). The higher order harmonics show a weaker centrality dependence, which is consistent with an anisotropy associated with fluctuations in the initial geometry.

Measurements of $v_n$ in peripheral heavy-ion collisions via the 2PC method are known to be biased by nonflow correlations. In particular, the even-order (odd-order) $v_n$ show a rapid increase (decrease) in value in going from midcentral to peripheral events. This bias increases with increasing $p_T$. For peripheral collisions (60–80% centrality), the $v_{n,n}$ obtained by the template-fit method are shown to be less biased by dijets and obey factorization better than the 2PC $v_{n,n}$. The template-fit $v_2$ are found to be significantly different from the 2PC $v_2$ at high $p_T$ and in peripheral events. However, for events over the 0–60% centrality range and $p_T \lesssim 5$ GeV, the template fits and the 2PC measurements typically yield consistent results for the $v_n$.

Two types of scaling behavior observed in prior $v_n$ measurements in Pb+Pb collisions are observed to hold in the Xe+Xe collisions as well. The first scaling is that the $v_n$ for fixed $n$ have the same shape as a function of $p_T$ across different centralities, up to an overall normalization and scaling along the $p_T$ axis. The second scaling is that the ratio $v_n/v_n^{2\text{nd}}$ for two harmonics of order $m$ and $n$ is found to be independent of $p_T$ in a given centrality interval. Neither scaling is understood quantitatively.

The measurements are also compared with prior $v_n$ measurements in Pb+Pb and $p + \text{Pb}$ collisions. Compared to Pb+Pb collisions the Xe+Xe $v_2$ is larger in the most central collisions. This can be attributed to larger fluctuations in Xe+Xe and a possible deformation in the Xe nucleus, which leads to larger eccentricity in Xe+Xe than in Pb+Pb. For midcentral and peripheral events, the $v_2$ in Xe+Xe becomes smaller than the $v_2$ in Pb+Pb, indicating the increased role of viscous effects. This trend with centrality is also observed for the higher order harmonics. The ratio of the Xe+Xe $v_n$ to the Pb+Pb $v_n$, as a function of centrality, is found to be qualitatively consistent with theoretical predictions. As a function of $N_{\text{part}}$ (or multiplicity), the $v_2$, $v_4$, and $v_5$ in Xe+Xe and Pb+Pb are observed to be very different except at small $N_{\text{part}}$ (multiplicity). On the other hand, the $v_3$ values in the two systems are consistent as a function of $N_{\text{part}}$ except for $N_{\text{part}} \gg 200$. The Xe+Xe $v_3$ measurements together with the cross-system comparisons offer a unique opportunity to study the interplay of fluctuations in the collision geometry and the role of viscous effects in heavy-ion collisions and should help to significantly improve the current understanding of the dynamics of heavy-ion collisions.

\section*{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,
Arnas, ANAS, Azerbaijan; SSTC, Belarus; CNPq and FAPESP, Brazil; NSERC, NRC, and CFI, Canada; CERN; CONICYT, Chile; CAS, MOST, and NSFC, China; COLCIENCIAS, Colombia; MSCT CR, MPO CR, and VSC CR, Czech Republic; DNRf and DNNSRC, Denmark; IN2P3-CNRS and CEA-DRF/IRFU, France; SRNSFG, Georgia; BMBF, HGF, and MPG, Germany; GSRT, Greece; RGC, Hong Kong SAR, China; ISF and Benoziyo Center, Israel; INFN, Italy; MEXT and JSPS, Japan; CNRST, Morocco; NWO, Netherlands; RCN, Norway; MNI$\text{\textsc{sw}}$ and NCN, Poland; FCT, Portugal; MNE/IFA, Romania; MES of Russia and NRC KI, Russian Federation; JINR and MESTD, Serbia; MSSR, Slovakia; AARRS and MIZŠ, Slovenia; DST/NRF, South Africa; MINECO, Spain; SRC and Wallenberg Foundation, Sweden; SERI, SNSF, and Cantons of Bern and Geneva, Switzerland; MOST, Taiwan; TAEK, Turkey; STFC, United Kingdom; and DOE and NSF, United States of America. In addition, individual groups and members have received support from BCKDF, CANARIE, CRC, and Compute Canada, Canada; COST, ERC, ERDF, Horizon 2020, and Marie Skłodowska-Curie Actions, European Union; Investissements d’ Avenir Labex and Idex, ANR, France; DFG and AvH Foundation, Germany; Herakleitos, Thales, and Aristeia programmes cofinanced by EU-ESF and the Greek NSRF, Greece; BSF-NSF and GIF, Israel; CERCA Programme Generalitat de Catalunya, Spain; and the Royal Society and Leverhulme Trust, United Kingdom. The crucial computing support from all WLCG partners is acknowledged gratefully, in particular from CERN, the ATLAS Tier-1 facilities at TRIUMF (Canada), NDGF (Denmark, Norway, Sweden), CC-IN2P3 (France), KIT/GridKA (Germany), INFN-CNAF (Italy), NL-T1 (Netherlands), PIC (Spain), ASGC (Taiwan), RAL (UK), and BNL (USA), the Tier-2 facilities worldwide and large non-WLCG resource providers. Major contributors of computing resources are listed in Ref. [63].


