Measurement of the suppression and azimuthal anisotropy of muons from heavy-flavor decays in Pb+Pb collisions at $\sqrt{s_{NN}} = 2.76$ TeV with the ATLAS detector

M. Aaboud et al.* (ATLAS Collaboration)

(Received 15 May 2018; published 22 October 2018)

ATLAS measurements of the production of muons from heavy-flavor decays in $\sqrt{s_{NN}} = 2.76$ TeV Pb+Pb collisions and $\sqrt{s} = 2.76$ TeV pp collisions at the LHC are presented. Integrated luminosities of 0.14 nb$^{-1}$ and 570 nb$^{-1}$ are used for the Pb+Pb and pp measurements, respectively, which are performed over the muon transverse momentum range $4 < p_T < 14$ GeV and for five Pb+Pb centrality intervals. Backgrounds arising from in-flight pion and kaon decays, hadronic showers, and misreconstructed muons are statistically removed using a template-fitting procedure. The heavy-flavor muon differential cross sections and per-event yields are measured in pp and Pb+Pb collisions, respectively. The nuclear modification factor $R_{AA}$ obtained from these is observed to be independent of $p_T$, within uncertainties, and to be less than unity, which indicates suppressed production of heavy-flavor muons in Pb+Pb collisions. For the 10% most central Pb+Pb events, the measured $R_{AA}$ is approximately 0.35. The azimuthal modulation of the heavy-flavor muon yields is also measured and the associated Fourier coefficients $v_n$ for $n = 2, 3,$ and 4 are given as a function of $p_T$ and centrality. They vary slowly with $p_T$ and show a systematic variation with centrality which is characteristic of other anisotropy measurements, such as that observed for inclusive hadrons. The measured $R_{AA}$ and $v_n$ values are also compared with theoretical calculations.

DOI: 10.1103/PhysRevC.98.044905

I. INTRODUCTION

Heavy quarks, especially bottom quarks, provide an important probe of the properties of the quark-gluon plasma created in high-energy nuclear ($A+A$) collisions [1–8]. The masses of the charm and bottom quarks are much larger than the temperatures of 200–500 MeV attained in the plasma (Ref. [9] and references therein). As a result, the heavy quarks are mostly thermalized and lose energy similarly to light quarks but with mass-dependent modification to the pattern of collisional and radiative energy loss [3,10–15]. At lower transverse momenta, $p_T \lesssim m_b$, the quarks are expected to diffuse in the plasma [4,7,16], losing energy and partially thermalizing [1,17]. As a result of their interactions with the collectively expanding medium, the heavy quarks may acquire an azimuthal anisotropy. Previous measurements of heavy-flavor production in $A+A$ collisions [18–21] and direct reconstruction of heavy-flavor mesons [22–26], have shown both substantial suppression in the yield of heavy quarks due to energy loss and significant azimuthal anisotropy. Measurements of the heavy-quark yield and azimuthal anisotropy in Pb+Pb collisions at the LHC can provide valuable constraints on plasma transport parameters, such as the heavy-quark diffusion coefficient, and potentially distinguish between weak- and strong-coupling models for heavy-quark interactions in the plasma [5,27–31].

The yield of particles produced in hard-scattering processes in $A+A$ collisions is often characterized using the nuclear modification factor

$$R_{AA} = \frac{1}{(T_{AA})} \frac{1}{N_{evt}} \left| \frac{d\sigma/dp_T d\eta}{d\sigma/dp_T d\eta}_{cent} \right|,$$

where $\eta$ is the pseudorapidity, the numerator is the differential per-event yield in $A+A$ collisions for a given centrality interval, the denominator is the pp differential cross section for producing the given particles, and $(T_{AA})$ represents the nuclear overlap function averaged over the centrality interval [32]. In the absence of significant modification to the nuclear parton distributions and of final-state interactions of the outgoing partons, $R_{AA}$ should be unity. Measurements of the production of vector bosons [33–37] in Pb+Pb collisions at the LHC have verified this expectation. In contrast, measurements of $R_{AA}$ for jets [38,39] and single hadrons [40–42] have shown a centrality-dependent suppression that is understood to result from the energy loss of the parent quarks and gluons (Refs. [43–45] and references therein). Measurements of $D$-meson production in Pb+Pb collisions at the LHC [24] have shown a centrality- and $p_T$-dependent suppression similar to

*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.
that observed for single hadrons. A measurement of $b$-hadron production, via their inclusive decays to $J/\psi$ mesons, has also shown significant suppression [46]. Separate measurements of the production of forward heavy-flavor-electron mesons [47] and muons [20] that are predominantly produced in semileptonic $B$- and $D$-meson decays give $R_{AA}$ values that are significantly larger than those observed for inclusive hadrons. However, the $b \rightarrow J/\psi X$ and forward muon measurements are statistically limited and insufficient to test theoretical calculations.

The azimuthal anisotropy of particles produced in an $A+A$ collision is often characterized by harmonic coefficients $v_n$ in a Fourier expansion of the particle yield as a function of azimuthal angle $\phi$ [48],

$$\frac{dN}{d\phi} = \left( \frac{dN}{d\phi} \right) \left( 1 + 2 \sum_{n \geq 1} v_n \cos(n(\phi - \Phi_n)) \right),$$

where $\Phi_n$ represents the event-plane angle for the $n$th harmonic. In noncentral collisions, the azimuthal anisotropy is usually dominated by the $n = 2$ term due to the almondlike shape of the collision geometry in the transverse plane resulting from the nonzero impact parameter. Measurements of inclusive [49–53] and identified hadron [54,55] $v_n$ values in $A+A$ collisions at the LHC and at RHIC show the presence of significant azimuthal anisotropies, which are well reproduced by hydrodynamic calculations. These results provide the basis for the interpretation that the medium created in heavy-ion collisions is strongly coupled. The elliptic flow of heavy-flavor hadrons depends both on the coupling of the heavy quark with the medium and on the transfer of the collective motion of the medium to the heavy-flavor hadron in the hadronization process [56]. The measurements of $D$-meson elliptic flow at midrapidity at the LHC [25,26] give $v_2$ values that are similar to those measured for light hadrons, while the forward-rapidity heavy-flavor $v_2$ values measured using semileptonic decays to muons are significantly smaller. However, those measurements are statistically limited and, thus, do not provide stringent constraints on theoretical calculations of the heavy-flavor elliptic flow. This paper presents ATLAS measurements of muons from heavy-flavor semileptonic decays (heavy-flavor muons, hereafter) in $pp$ collisions at $\sqrt{s} = 2.76$ TeV and $\sqrt{s_{NN}} = 2.76$ TeV. The $pp$ and Pb+Pb data were recorded during 2011, and the $pp$ data were recorded during 2013. The measurements are performed using data sets with integrated luminosities of 570 and 0.14 nb$^{-1}$ for $pp$ and Pb+Pb collisions, respectively. They are performed for several intervals of collision centrality, characterized using the total transverse energy measured in the forward calorimeters, and for different muon $p_T$ intervals spanning the range 4–14 GeV. Heavy-flavor muons are statistically separated from background muons resulting from pion and kaon decays and from hadronic interactions using a “momentum-imbalance” variable (Sec. III C) that compares the momenta of the muons measured in the inner detector and muon spectrometer.

Over the $p_T$ range of the measurement, the residual irreducible contamination by non-heavy-flavor muons, including contributions from $J/\psi$ decays [57,58], is less than 1% and is neglected in the following. The heavy-flavor muon differential per-event yields in Pb+Pb collisions and differential cross sections in $pp$ collisions measured over the pseudorapidity interval $|\eta| < 1$ are used to calculate the heavy-flavor muon $R_{AA}$ as a function of $p_T$ in different Pb+Pb centrality intervals. In addition, heavy-flavor muon $v_n$ values are measured for $n = 2–4$ as a function of $p_T$ and collision centrality over $|\eta| < 2$ using both the event-plane and scalar-product [59] methods. The scalar-product method has become the de facto standard procedure for $v_n$ measurements using event-plane reconstruction. However, the method introduces additional complexity to the background subtraction procedure (see Sec. III D), so results obtained using both methods are provided. The results presented in this paper provide significantly improved statistical precision over previous measurements of the suppression and the anisotropic flow of semileptonically decaying heavy-flavor hadrons in Pb+Pb collisions at the LHC.

This paper is structured as follows. Section II describes the components of the ATLAS detector and trigger system used in the measurement, Sec. III describes the data analysis, Sec. IV discusses the systematic uncertainties, and the results are discussed in Sec. V. Section VI provides a summary and outlook.

II. ATLAS DETECTOR

The measurements presented in this paper use the ATLAS muon spectrometer (MS), inner detector (ID), calorimeter, trigger, and data acquisition systems. A detailed description of these detectors and their performance in $pp$ collisions is given in Ref. [60]. Muons are reconstructed by combining independent measurements of the muon trajectories from the ID and the MS. The ID measures charged particles within the pseudorapidity interval $|\eta| < 2.5$ using silicon pixel detectors, silicon microstrip detectors (SCTs), and a straw-tube tracker, all immersed in a 2-T axial magnetic field. A charged particle typically traverses three layers of silicon pixel detectors, four layers of double-sided microstrip sensors, and 36 straws. The ID is surrounded by electromagnetic and hadronic calorimeters that absorb efficiently the copious charged and neutral hadrons produced in Pb+Pb collisions. A muon typically loses 3–5 GeV of energy, depending on the muon pseudorapidity, while crossing the calorimeters. The MS surrounds the calorimeters and provides tracking for muons within $|\eta| < 2.7$ in the magnetic field produced by three air-core toroid magnet systems. Muon momenta are measured in the MS using three stations of precision drift chambers. Fast tracking detectors are used to trigger on muons in the MS.

Two forward calorimeters (FCal) are placed symmetrically with respect to $z = 0$ and cover $3.2 < |\eta| < 4.9$. They are composed of tungsten and copper absorbers with liquid argon.
as the active medium; each calorimeter has a total thickness of about ten interaction lengths.

Minimum-bias Pb+Pb collisions are identified using the zero-degree calorimeters (ZDCs) and the minimum-bias trigger scintillator (MBTS) counters [60]. The ZDCs are located symmetrically at $z = \pm 140$ m and cover $|\eta| > 8.3$. They are used only in Pb+Pb collisions where they primarily measure “spectator” neutrons, which originate from the incident nuclei and do not scatter hadronically during the collision. The MBTS system detects charged particles over 2.1 < $|\eta|$ < 3.9 using two hodoscopes of 16 counters each, placed at $z = \pm 3.6$ m. The MBTS counters provide measurements of both the pulse heights and arrival times of ionization energy depositions in each hodoscope.

The ATLAS trigger system [61] consists of a first-level (L1) trigger implemented using a combination of dedicated electronics with programmable logic, and a software-based high-level trigger (HLT). Data used for this analysis were selected using a combination of minimum-bias triggers, which provided a uniform sampling of the Pb+Pb inelastic cross section, and triggers that selected rare physics signatures such as muons. The measurements presented here are primarily obtained from muon triggers. Events from the minimum-bias triggers are used only for cross checks.

The muon triggers are formed using a combination of a L1 trigger and an HLT muon trigger whose configuration differed between Pb+Pb and $pp$ operation. For the Pb+Pb data, the L1 trigger selected events having a total transverse energy of more than 50 GeV, and the HLT trigger selected events containing a track in the MS whose $p_T$, when corrected for the average muon energy loss in the calorimeter, is greater than 4 GeV. In $pp$ data, the muon trigger required a stand-alone muon track in the MS at L1, and a muon track reconstructed using both the ID and MS with $p_T > 4$ GeV at the HLT. The muon trigger was unprescaled throughout the Pb+Pb run and sampled essentially all of the delivered luminosity. In the $pp$ run, the trigger was prescaled such that it sampled $\sim$14% (570 nb$^{-1}$) of the 4 pb$^{-1}$ delivered luminosity.

III. DATA ANALYSIS

A. Event selection

Charged-particle tracks and vertices are reconstructed from hits in the ID using a track reconstruction algorithm [62] whose configuration changed between the $pp$ and Pb+Pb measurements to account for the high hit density in heavy-ion collisions [50]. To remove noncollision backgrounds, Pb+Pb events are required to have a reconstructed primary vertex, at least one hit in each MBTS counter, and a time difference between the two MBTS time measurements of less than 5 ns; $pp$ events are required to have at least one reconstructed primary vertex.

The centrality of Pb+Pb collisions is characterized by $\Sigma E^\text{FCal}_T$, the total transverse energy measured in the FCal [50]. For the results presented in this paper, the minimum-bias $\Sigma E^\text{FCal}_T$ distribution is divided into centrality intervals according to the following percentiles of the $\Sigma E^\text{FCal}_T$ distribution ordered from the most central to the most peripheral collisions: 0–10%, 10–20%, 20–30%, 30–40%, and 40–60%. A Glauber Monte Carlo analysis [63] is used to estimate $(T_{AA})$ for each of the centrality intervals [38]. The results are provided in Table I.

B. Muon reconstruction

Muons in this analysis are formed by combining tracks reconstructed in the MS [57] with the tracks measured in the ID. The associated ID tracks are required to satisfy criteria for the number of hits in the SCT and pixel detectors which are the same for the $pp$ and Pb+Pb data, but which are optimized for the Pb+Pb analysis [50]. In particular, for both data sets, ID tracks are required to have transverse and longitudinal impact parameters relative to the reconstructed primary vertex of less than 5 mm and to have a momentum $p > 3$ GeV. The requirements on the longitudinal and transverse impact parameters are relaxed to 5 mm, compared to the 1 mm (or 1.5 mm) typically used in heavy-ion analyses [50,52], to allow selection of muons from off-vertex heavy-flavor decays. The ID tracks are also required to have at least one pixel hit, with the additional requirement of a hit in the first pixel layer when one is expected,\(^2\)\footnote{A hit is expected if the extrapolated track crosses an active region of a pixel module that has not been disabled.} at least seven SHT hits, and at most one hit that is expected but not found in the pixel and SCT detectors taken together. The transverse momentum measured in the MS ($p_T^\text{MS}$) is required to be greater than 1.2 GeV for both the $pp$ and Pb+Pb data. In the Pb+Pb analysis, this selection removes muons for which the Pb+Pb trigger efficiency is less than 50%.

The results presented here use muons having $4 < p_T < 14$ GeV and having $|\eta| < 1$ for the heavy-flavor-suppression analysis or $|\eta| < 2$ for the flow measurements. The lower limit of the $p_T$ range is constrained by the $p_T$ dependence of the muon trigger and reconstruction efficiencies, while the upper limit is determined by the number of events available in the Pb+Pb data. For the $R_{AA}$ measurements, a muon $\eta$ interval of $|\eta| < 1$ is chosen, as the muon trigger and reconstruction have optimal performance over this $\eta$ range. The $\eta$ range is extended to $|\eta| < 2$ for the $v_n$ measurements.

<table>
<thead>
<tr>
<th>Centrality interval (%)</th>
<th>$(T_{AA})$ (mb$^{-1}$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0–10</td>
<td>23.45 ± 0.37</td>
</tr>
<tr>
<td>10–20</td>
<td>14.43 ± 0.30</td>
</tr>
<tr>
<td>20–30</td>
<td>8.73 ± 0.26</td>
</tr>
<tr>
<td>30–40</td>
<td>5.04 ± 0.22</td>
</tr>
<tr>
<td>40–60</td>
<td>2.02 ± 0.15</td>
</tr>
</tbody>
</table>

044905-3
as they are not sensitive to the effects of trigger and tracking efficiency. A total of 9.2 million (1.8 million) muons are reconstructed within these kinematic ranges from 8.7 million (1.8 million) events recorded using the Pb+Pb \((pp)\) muon triggers. The performance of the ATLAS detector and offline analysis in measuring muons in \(pp\) collisions is evaluated by a \textsc{geant4} [64] simulation of the ATLAS detector [65] using Monte Carlo (MC) \(\sqrt{s} = 2.76\text{ TeV}\) \(pp\) events produced with the \textsc{pythia} event generator [66] (version 6.423 with parameters chosen according to the AUET2B set of tuned parameters [67]). The reconstruction performance in \(Pb+Pb\) collisions is evaluated by “overlaying” simulated \textsc{pythia} \(pp\) events on minimum-bias \(Pb+Pb\) events. In this overlay procedure, the simulated hits are combined with the data from minimum-bias events to produce the final sample. The minimum-bias \(Pb+Pb\) events used in the overlay procedure were recorded by ATLAS during the same data-taking period as the data used in this analysis. For both the \(pp\) and \(Pb+Pb\) measurements, the muon reconstruction efficiency increases by about 30% from \(p_T = 4\text{ GeV}\) to \(p_T = 6\text{ GeV}\), above which it is approximately constant at 0.80 and 0.77 for the \(pp\) and \(Pb+Pb\) data, respectively. The \(Pb+Pb\) muon reconstruction efficiency is independent of the centrality within uncertainties.

The \(Pb+Pb\) muon trigger efficiency is measured for fully reconstructed muons using the minimum-bias \(Pb+Pb\) data set. The efficiency is evaluated as the fraction of reconstructed muons for which the HLT finds a matching muon with \(p_T > 4\text{ GeV}\). It is observed to be independent of centrality, within statistical uncertainties, and increases from about 0.6 at \(p_T = 4\text{ GeV}\) to about 0.8 at 6 GeV, above which it is approximately constant. The \(pp\) muon trigger efficiency is similarly evaluated using \(pp\) events selected by a set of minimum-bias triggers. The efficiency increases from 0.40 for \(p_T = 4\text{ GeV}\) to 0.75 for \(p_T = 12\text{ GeV}\).

### C. Heavy-flavor-suppression measurement

The muons measured in the \(pp\) and \(Pb+Pb\) data sets contain background from in-flight decays of pions and kaons, muons produced from the decays of particles produced in hadronic showers in the material of the detector, and misassociations of ID and MS tracks. Previous studies have shown that the signal and background contributions to the reconstructed muon sample can be discriminated statistically [57]. This analysis relies solely on the fractional momentum imbalance \(\Delta p/p_{\text{ID}}\), which quantifies the difference between the ID and MS measurements of the muon momentum after accounting for the energy loss of the muon in the calorimeters. It is defined as

\[
\frac{\Delta p}{p_{\text{ID}}} = \frac{p_{\text{ID}} - p_{\text{MS}} - \Delta p_{\text{calo}}(p, \eta, \phi)}{p_{\text{ID}}},
\]

where \(p_{\text{ID}}\) and \(p_{\text{MS}}\) represent the reconstructed muon momenta from the ID and MS, respectively, and \(\Delta p_{\text{calo}}\) represents the momentum- and angle-dependent average momentum loss of muons in the calorimeter obtained from simulations. Muons resulting from background processes typically have \(p_{\text{MS}}\) values smaller than would be expected for a muon produced directly in \(pp\) or \(Pb+Pb\) collisions or via the decays of heavy-flavor hadrons. This is because the background muons from pion/kaon decays or from hadronic interactions in the calorimeter have, on average, smaller \(p_T\) compared to the parent particle. As a result, background muons are expected to have \(\Delta p/p_{\text{ID}} > 0\).

Distributions for \(\Delta p/p_{\text{ID}}\) are obtained from the simulated samples separately for signal muons and for background muons. The signal muons include muons directly produced in electromagnetic decays of hadrons, in decays of \(\tau\) leptons, in decays of W and Z bosons, in decays of top quarks, and in semileptonic decays of heavy-flavor hadrons. The last contribution dominates the signal sample, contributing about 99% of the muons over the \(p_T\) range measured in this analysis (Ref. [57] and references therein). The different contributions to the background—pion decays in flight, kaon decays in flight, muons produced by secondary interactions of prompt particles, and misassociations—are evaluated separately. Figure 1 shows MC distributions of \(\Delta p/p_{\text{ID}}\) for signal and background muons having 5 < \(p_T\) < 6 GeV for \(Pb+Pb\) collisions in the centrality range 0–60% and for \(pp\) collisions. The \(\Delta p/p_{\text{ID}}\) distribution for signal muons is centered at zero while the distribution for background muons is shifted to positive values. The signal distributions show only modest differences between \(pp\) and \(Pb+Pb\) collisions. Similarly, when making separate templates for different \(Pb+Pb\) collision centralities, a weak dependence of the signal template on centrality is observed. The background \(\Delta p/p_{\text{ID}}\) distributions are much broader and are insensitive to the centrality-dependent effects seen in the signal distributions.

A template-fitting procedure is used to estimate statistically the signal fraction for each kinematic and centrality selection used in the analysis. The measured \(\Delta p/p_{\text{ID}}\) distribution is assumed to result from a combination of signal and background

![FIG. 1](image-url)
distributions, 
\[ \frac{1}{N_\mu} \frac{dN_\mu}{d\Delta p/p_{ID}} = f^{sig} \frac{dP^{sig}}{d\Delta p/p_{ID}} + (1 - f^{sig}) \frac{dP^{bkg}}{d\Delta p/p_{ID}}, \]
where \( N_\mu \) is the total number of muons in the sample, \( dP^{sig}/d\Delta p/p_{ID} \) and \( dP^{bkg}/d\Delta p/p_{ID} \) represent the signal and background \( \Delta p/p_{ID} \) probability distributions, respectively, and \( f^{sig} \) represents the signal fraction.

For Pb+Pb data, centrality-dependent templates are used for the signal while centrality-integrated templates are used for the background. The latter is motivated by the observed centrality independence of the background templates and the limited size of the background sample. Template fits are performed using binned \( \chi^2 \) fits that account for the statistical precision of the signal and background templates. The fits are performed using MINUIT \cite{68} with \( f^{sig} \) as the free parameter. The uncertainties from the fits are used as statistical uncertainties of the yields and propagated into the final results. Example template fits are shown for two muon \( p_T \) intervals in Fig. 2 for Pb+Pb events in the 0–10% and 40–60% centrality intervals and for \( pp \) data. As shown in Fig. 2, the measured \( \Delta p/p_{ID} \) distributions are well described by a combination of the signal and background templates, and this holds for all studied kinematic and centrality intervals.

The signal fractions \( f^{sig} \) obtained from the template fits using these intervals are shown in Fig. 3 for the Pb+Pb and \( pp \) data. The signal fractions increase with \( p_T \) for \( p_T > 5 \text{ GeV} \), indicating that at higher \( p_T \) a larger fraction of the reconstructed muons are heavy-flavor (HF) muons. The increase in \( f^{sig} \) at low \( p_T \) results from the trigger, which is less efficient for background muons that have low \( p_T^{MS} \). Such an increase is not observed when repeating this analysis using the minimum-bias Pb+Pb data set. This increase in the \( f^{sig} \) due to the trigger does not affect the measurement, as is demonstrated by studies of variations in the \( p_T^{MS} \) criterion in Sec. IV A.

With the \( f^{sig} \) obtained from the template fits, the \( pp \) differential cross section for producing heavy-flavor muons is calculated according to
\[ \frac{d^2\sigma_{HF\mu}}{dp_T d\eta} = \frac{1}{\mathcal{L}} \frac{\Delta N_\mu f^{sig} \frac{1}{\epsilon_{\text{trig}}\epsilon_{\text{rec}}}}{\Delta p_T \Delta \eta}, \]
where \( \mathcal{L} \) is the integrated luminosity of the \( pp \) measurement, \( \Delta p_T \) is the width of the given \( p_T \) interval, \( \Delta \eta = 2 \) is the size of the pseudorapidity interval, \( \Delta N_\mu \) represents the number of muons in the given \( p_T \) and \( \eta \) intervals, and \( \epsilon_{\text{trig}} \) and \( \epsilon_{\text{rec}} \) represent the trigger and reconstruction efficiencies, respectively. The luminosity is calibrated using a set of beam-separation scans performed in February 2013. It has a relative uncertainty of 3.1% that was derived following a methodology similar to that detailed in Ref. \cite{69}.

The Pb+Pb differential per-event yields for producing heavy-flavor muons are calculated according to
\[ \frac{1}{N_{\text{evt}}} \frac{d^2N_{HF\mu}}{dp_T d\eta}_{\text{cent}} = \frac{\Delta N_{\mu}^{\text{cent}} f^{sig} \frac{1}{\epsilon_{\text{trig}}\epsilon_{\text{rec}}}}{\Delta p_T \Delta \eta N_{\text{evt}}^{\text{cent}}}. \]
where $N_{\text{cent}}$ is the number of Pb+Pb collisions in a given centrality interval, $\Delta N_{\text{cent}}$ represents the number of total muons with $|\eta| < 1$ measured in the given $p_T$ and centrality interval, $f^{\text{sig}}$ represents the corresponding signal fraction obtained from the template fits, and $\varepsilon_{\text{trg}}$ and $\varepsilon_{\text{rec}}$ represent the trigger and reconstruction efficiencies, respectively.

D. Azimuthal anisotropy measurement

The $v_n$ measurements additionally require determination of the event-plane (EP) angles $\Phi_n$ [Eq. (2)]. However, due to detector acceptance effects and finite particle multiplicity in an event, the measured EP angles, denoted $\Psi_n$, fluctuate event by event around the true EP angles [48]. The “observed” $v_n$, $v_n^{\text{obs}}$, is obtained by measuring the distribution of the particle directions relative to the $\Psi_n$ planes:

$$dN\over d\phi = N_0 \left[ 1 + 2 \sum_{n \geq 1} v_n^{\text{obs}} \cos(n(\phi - \Psi_n)) \right].$$

The $v_n^{\text{obs}}$ are smaller in magnitude than the true $v_n$ because they are calculated around the $\Psi_n$ planes rather than the $\Phi_n$ planes. To account for this, the $v_n^{\text{obs}}$ are corrected by the EP resolution factor $\text{Res}(n\Psi_n)$, which accounts for the smearing of $\Psi_n$ relative to $\Phi_n$ [48]:

$$v_n = {v_n^{\text{obs}}} \over \text{Res}(n\Psi_n), \quad \text{Res}(n\Psi_n) = \langle \cos(n(\Psi_n - \Phi_n)) \rangle_{\text{evts}},$$

where the $\langle \ldots \rangle_{\text{evts}}$ indicates averaging over all events in a given centrality class. In this analysis, the $\Psi_n$ angle is determined using the flow vector or “$q$-vector” method [48], in which the $q$ vector is calculated from the $E_T$ deposited in the FCal according to

$$q_{n,x} = \frac{\sum E_{T,i} \cos(n\phi_i) - \langle \sum E_{T,i} \cos(n\phi_i) \rangle_{\text{evts}}}{\sum E_{T,i}},$$

$$q_{n,y} = \frac{\sum E_{T,i} \sin(n\phi_i) - \langle \sum E_{T,i} \sin(n\phi_i) \rangle_{\text{evts}}}{\sum E_{T,i}},$$

where the sum is over all the calorimeter towers in the FCal, $E_{T,i}$ is the transverse energy deposited in the $i$th tower, and $\phi_i$ denotes the azimuthal angle of the position of the center of the tower. The event-averaged terms $\langle \sum E_{T,i} \cos(n\phi_i) \rangle_{\text{evts}}$ and $\langle \sum E_{T,i} \sin(n\phi_i) \rangle_{\text{evts}}$ are subtracted in order to remove detector effects [70]. From the $q_n$ vectors, the EP angles $\Psi_n$, are determined as [71]

$$\tan(n\Psi_n) = \frac{q_{n,y}}{q_{n,x}}.$$  

The parameter $\text{Res}(n\Psi_n)$ is determined by the two-subevents (2SE) method [48]. In the 2SE method, the signal from a detector used to measure the event plane is divided into two “subevents” covering equal pseudorapidity ranges in opposite $\eta$ hemispheres, such that the two subevents nominally have the same resolution. The FCal detectors located at positive and negative $\eta$, FCal$^P$ and FCal$^N$, provide such a division. The resolution of the FCal$^{P(N)}$ is calculated from the correlation between the two subevents

$$\text{Res}(n\Psi_n^{P(N)}) = \sqrt{\langle \cos(n(\Psi_n^P - \Psi_n^N)) \rangle_{\text{evts}}},$$

where $\Psi_n^{P(N)}$ is the event-plane angle determined from the positive (negative) side of the FCal. From the subevent resolution the full FCal resolution can be determined by the procedure

3 Calorimeter towers are localized groups of calorimeter cells that have a $\delta\eta \times \delta\phi$ segmentation of 0.1 $\times$ 0.1.
are significant bin-to-bin correlations between the statistical uncertainties due to the use of the same signal and background templates in all 

described in Ref. [48]. The $\text{Res}(n \Psi_n)$ for the FCal and their associated systematic uncertainties were determined in a previous ATLAS analysis [52]. Those values and uncertainties are directly used in this paper. Depending on the centrality class, the EP resolution factor for the FCal varies between 0.7 and 0.9, 0.3 and 0.65, and 0.2 and 0.4 for $v_2$, $v_3$, and $v_4$, respectively. The uncertainties in the EP resolution factor are less than 3%, 4%, and 6% for $v_2$, $v_3$, and $v_4$, respectively, for all the centrality classes used in this analysis.

The heavy-flavor muon $v_n^{obs}$ values are measured by evaluating the yields differentially relative to the $\Psi_n$ plane. For this, the template-fitting procedure is repeated in intervals of $n|\phi - \Psi_n|$ for each $p_T$ and centrality interval. Utilizing the $n$-fold symmetry of the $\Psi_n$ plane and the fact that $\cos[n(\phi - \Psi_n)]$ is an even function, it is sufficient to bin only over the interval $(0, \pi)$ in $n|\phi - \Psi_n|$. Four intervals of $n|\phi - \Psi_n|$ [(0, $\pi/4$), ($\pi/4$, $\pi/2$), ($\pi/2$, $3\pi/4$), and ($3\pi/4$, $\pi$)] are used. The same signal and background templates are used for the four $n|\phi - \Psi_n|$ intervals in a given $p_T$ and centrality interval. As a result, there is a significant correlation between the statistical uncertainties of the signal fractions measured in the four $\cos[n(\phi - \Psi_2)]$ intervals. This correlation is accounted for in the statistical uncertainties of the final $v_n$ values.

Figure 4 shows examples of the differential yields of heavy-flavor muons obtained from the template fits as a function of $2|\phi - \Psi_2|$ for two centrality and two $p_T$ intervals.

A clear dependence of the yields on $2|\phi - \Psi_2|$ can be observed, with a larger yield in the “in-plane” direction ($2|\phi - \Psi_2| \sim 0$) compared to the “out-of-plane” direction ($2|\phi - \Psi_2| \sim \pi$), implying a significant $v_2$ signal. The differential yields are fitted with a second-order Fourier function of the form in Eq. (5) to obtain the $v_n^{obs}$ values. In the fits, the $\chi^2$ minimization takes into account the correlations between the statistical uncertainties of the yields in the different $2|\phi - \Psi_2|$ bins. These fits are indicated by the continuous lines in Fig. 4. The $v_n^{obs}$ values are then corrected to account for the EP resolution [Eq. (6)] for the final results presented in Sec. V.

One drawback of the EP method is that there is an ambiguity in the interpretation of the $v_n$ values obtained from it (from here on the $v_n$ values obtained from the event-plane method are denoted by $v_n^{EP}$). In the limit of perfect EP resolution, $\text{Res}(n \Psi_n) \rightarrow 1$, $v_n^{EP} \rightarrow \langle v_n \rangle$, while in the limit of poor resolution, $\text{Res}(n \Psi_n) \rightarrow 0$, $v_n^{EP} \rightarrow \sqrt{\langle v_n^2 \rangle}$ where the $\langle \ldots \rangle$ indicates an average over all events [59]. In general, the $v_n$ values measured with the EP method lie somewhere between $\langle v_n \rangle$ and $\sqrt{\langle v_n^2 \rangle}$, depending on the value of the resolution. For this reason, the scalar-product (SP) method is considered to be a superior measurement technique, as it always measures the r.m.s. $v_n$ value, i.e., $\sqrt{\langle v_n^2 \rangle}$ [59]. The ideal SP method entails weighting the contribution of each measured signal muon by the magnitude of the $q$ vector [Eq. (7)] measured in the FCal.
giving
\[ n_{SP} = \frac{(q_n \cos[n(\phi - \Psi_n)])_{\text{evts}}}{\text{Res}^{SP}(n\Psi_n)}, \]
where \( \text{Res}^{SP}(n\Psi_n) \) is the resolution for the SP method, given by
\[ \text{Res}^{SP}(n\Psi_n) = \sqrt{|q_n^{P(N)} N_n| \cos(n(\Psi_n^{P} - \Psi_n^{N}))}, \]
where \( q_n^{P(N)} \) is the magnitude of the \( n \)th-order \( q \) vector measured in the positive \( z \) (negative \( z \)) side of the FCal. Previous ATLAS measurements for inclusive charged particles show that \( v_{n}^{EP} \) values differ by less than 5% from the r.m.s. \( v_n \) values for \( v_2 \), and harmonics of order \( n \geq 3 \) are consistent with the r.m.s. \( v_n \) within systematic uncertainties [72]. However, Eq. (8) cannot be directly used in the present analysis, since a priori
\[ \text{Eq. (8)} \]
cannot be directly used in the present analysis, since a priori
\[ \text{Eq. (8)} \], presented later in Sec. V.

While the SP method has advantages over the EP method, only a modified version of the SP method can be used in the present analysis. Thus, the results obtained from both the SP and EP methods are presented.

E. Jet bias in the \( v_n \) measurement

The heavy-flavor muons measured in this analysis often result from heavy-flavor jets that have an associated back-to-back recoil jet. If the recoil jet is in the FCal, it can bias the orientation of the \( \Psi_n \), to be aligned with the azimuthal angle of the muon, yielding a larger measured \( v_n \). This “jet bias effect” needs to be estimated and corrected for in the measurement. The magnitude of this effect is estimated using the simulated-data overlay events described in Sec. III B, where PYTHIA-generated events are overlaid on minimum-bias Pb+Pb data. The overlay is done independently of the \( \Psi_n \) angles and, thus, should yield a zero \( v_n \) value when the analysis procedure used in the data is applied to the simulated events. Any systematic deviation from \( v_n = 0 \) seen in the simulated data is, then, a result of jet bias. The procedure used to evaluate the jet bias in \( v_n \) values is as follows.

The presence of the recoil jet biases the observed \( q \) vector in the FCal as
\[ q_n^{\text{Biased}} = q_n e^{in\Psi_n} + k e^{in\phi^{\text{jet}}}, \]
where the first term on the right is the unbiased \( q \) vector, which only has the natural statistical smearing. The second term on the right is the bias introduced by the recoil jet, which shifts the event-plane angle to be aligned with the recoil jet direction. The factor \( k \) represents the strength of the bias and may depend on the \( p_T \) of the recoil jet as well as the centrality, and \( \phi^{\text{jet}} \) is the direction of the jet. Since the recoil jet is nominally back to back with the muon, its direction can be written as
\[ \phi^{\text{jet}} = \phi^{\text{mu}} + \pi + \delta, \]
where \( \phi^{\text{mu}} \) is the azimuthal angle of the muon and \( \delta \) represents event-by-event fluctuations in the jet direction. This bias affects the numerator in the SP method [Eq. (8)]; the resolution [denominator in Eq. (8)] is not affected by the bias, as the resolution is calculated using minimum-bias events and not from events that are triggered by muons. The dot product between the muon’s transverse direction and the biased \( q \) vector, averaged over many events, [equation of \( \text{Eq. (8)} \), now becomes
\[ \langle q_n e^{-in\Psi_n} + ke^{-in(\phi^{\text{mu}}+\pi+\delta)} \rangle_{\text{evts}} \]

The first term on the right is the numerator of Eq. (8) for no bias, and the second term is the bias, which conveniently separates out as an additive contribution. The second term on the right of Eq. (9) corrected by \( \text{Res}^{SP}(n\Psi_n) \) is the jet bias in \( v_n^{SP} \).

The bias determined in this manner is independent of \( p_T \) within statistical errors. The magnitude of the bias varies with centrality. It is smallest in the most central events—where the underlying event is quite large, and the additional energy deposited by the jet does not cause a significant perturbation—and increases with decreasing centrality. For \( v_2 \), the \( p_T \)-averaged value of this bias is 0.0025 in the 0–10% centrality interval; it increases to 0.011 in the 40–60% centrality interval. For comparison, the \( v_2 \) at \( p_T = 4 \text{ GeV} \) in the 0–10% and 40–60% centrality intervals is about 0.04 and 0.07, respectively. Because the jet yield is suppressed by as much as a factor of 2 in Pb+Pb collisions [38], only half of this estimated bias is applied as a correction. Half of this estimated bias is also conservatively taken as the systematic uncertainty of the correction. In principle, the jet bias also affects the \( R_{AA} \) measurements since the correlated jet, if it falls within the FCal acceptance, also alters the centrality interval to which the event is assigned. However, this effect, estimated from the simulated-data overlay sample, is negligible compared to the systematic uncertainties in the \( R_{AA} \) measurement (Sec. IV A), and corrections for it are not applicable.

IV. SYSTEMATIC UNCERTAINTIES

A. Yield, cross-section, and \( R_{AA} \) systematic uncertainties

The measurements of the heavy-flavor muon differential cross sections and per-event yields are subject to systematic uncertainties arising from the muon-trigger selection, muon-reconstruction efficiencies, the template-fitting procedure, muon \( p_T \) resolution, and the \( pp \) luminosity. They are

\[ n_{SP} = \frac{(q_n \cos[n(\phi - \Psi_n)])_{\text{evts}}}{\text{Res}^{SP}(n\Psi_n)}, \]
described below. Where appropriate, the uncertainties are smoothed as a function of $p_T$, to reduce the statistical fluctuations in the uncertainty estimates. The systematic uncertainties for $\text{Pb+Pb}$ data do not show any significant variation with collision centrality.

The systematic uncertainty in the $\text{Pb+Pb}$ muon-trigger efficiency is evaluated by varying the selections applied to the offline-reconstructed muons in the minimum-bias reference sample and re-evaluating the trigger efficiency. The resulting changes in the trigger efficiency are less than 0.5% over $4 < p_T < 14$ GeV and are taken as the estimate of the systematic uncertainty in $\varepsilon_{\text{trig}}$. The uncertainty in the $pp$ muon-trigger efficiency is evaluated similarly, and is less than 2.5% for $p_T < 6$ GeV and less than 1.5% for $p_T > 6$ GeV. The systematic uncertainty associated with the muon-reconstruction efficiency is evaluated by varying the muon selections, evaluating the reconstruction efficiency for the new selections, and repeating the analysis with the updated muon selection and reconstruction efficiency. This uncertainty is less than about 4% for the $pp$ data and less than about 2.5% for the $\text{Pb+Pb}$ data. Separately, the minimum $p_{\text{TMS}}^M$ (default value of 1.2 GeV, Sec. III B) is varied from 0.5 to 1.5 GeV, and the entire analysis is repeated. This variation affects the template fitting but also is sensitive to potential systematic uncertainties in the muon reconstruction and trigger efficiencies. The change in the $\text{Pb+Pb}$ muon yields from varying the minimum $p_{\text{TMS}}^M$, taken as a systematic uncertainty in the heavy-flavor muon yields, decreases with $p_T$ from $\sim10\%$ to $\sim1.5\%$ over the measured $p_T$ range. For the $pp$ cross-section measurements, the systematic uncertainty decreases with $p_T$ from $\sim11.5\%$ to $\sim3\%$. The systematic uncertainty associated with the $p_{\text{TMS}}^M$ criterion is somewhat correlated with the systematic uncertainty associated with the trigger efficiency; however, they are conservatively treated as independent uncertainties.

Systematic uncertainties resulting from the construction of the templates, particularly the background template, are evaluated by changing the relative proportions of different background contributions. The pion and kaon decay-in-flight

<table>
<thead>
<tr>
<th>$p_T$ interval</th>
<th>$4 &lt; p_T &lt; 4.5$ GeV</th>
<th>$6 &lt; p_T &lt; 7$ GeV</th>
<th>$10 &lt; p_T &lt; 12$ GeV</th>
</tr>
</thead>
<tbody>
<tr>
<td>Muon selection (%)</td>
<td>2.5</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>$p_{\text{TMS}}^M$ selection (%)</td>
<td>7.5</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>Background template variation (%)</td>
<td>0.5</td>
<td>0.5</td>
<td>0.5</td>
</tr>
<tr>
<td>Template fitting (%)</td>
<td>13</td>
<td>7</td>
<td>5</td>
</tr>
<tr>
<td>Efficiency (%)</td>
<td>2.5</td>
<td>1.5</td>
<td>1.5</td>
</tr>
</tbody>
</table>

**Table II.** Relative systematic uncertainties in the heavy-flavor muon $R_{AA}$, quoted in percent, for selected $p_T$ intervals.

**Table III.** Systematic uncertainties in the heavy-flavor muon $v_n$ for selected $p_T$ and centrality intervals. The values are for the EP method and are quoted either as absolute values or in percent. They are averaged over $p_T$ intervals that are larger than the intervals used for the measurement.
components of the background are separately increased by a factor of 2 and then separately decreased by a factor of 2, as motivated by differences observed in the kaon to pion yields between PYTHIA—which is used to generate the MC templates—and data [74]. For each variation, the template fitting is performed, and a new value for $f^{\text{sig}}$ is obtained. The average of the unsigned differences between the varied and nominal $f^{\text{sig}}$ values is taken as the systematic uncertainty in the template fitting due to the background composition. This is less than 0.5% over the $p_T$ range of the measurement for both the Pb+Pb and $pp$ data.

In order to account for possible inconsistencies between the data and MC templates that may arise from the effect of the trigger, or other factors that may not be properly accounted for in the MC simulation, a separate systematic uncertainty in the template-fitting method is estimated using a “cut-and-correct” procedure applied to the $\Delta p/p_{ID}$ distributions. In this procedure, the fraction of muons having $\Delta p/p_{ID} < \Delta p/p_{ID}|_{\text{cut}}$, $f^c$, is measured in the data in each centrality and $p_T$ interval. This fraction provides an estimate of the signal muon fraction, but it must be corrected for true muons having $\Delta p/p_{ID} > \Delta p/p_{ID}|_{\text{cut}}$ (inefficiency) and background muons having $\Delta p/p_{ID} < \Delta p/p_{ID}|_{\text{cut}}$ (fakes). The corrections are obtained from the MC signal and background $\Delta p/p_{ID}$ distributions and are expressed in terms of the efficiencies, $\varepsilon_{\text{true}}$ and $\varepsilon_{\text{bkg}}$, for true and background muons, respectively, to pass the $p/p_{ID} < \Delta p/p_{ID}|_{\text{cut}}$. In terms of these efficiencies, $f^c$ is given by

$$f^c = f^{\text{sig}} \varepsilon_{\text{true}} + (1 - f^{\text{sig}}) \varepsilon_{\text{bkg}}.$$  

Inverting this equation, the signal fraction estimated using the cut-and-correct procedure is

$$f^{\text{sig}} = \frac{f^c - \varepsilon_{\text{bkg}}}{\varepsilon_{\text{true}} - \varepsilon_{\text{bkg}}}.$$  

If the MC exactly describes the signal and background $\Delta p/p_{ID}$ distributions in the data, then the cut-and-correct $f^{\text{sig}}$ values will be identical to the signal fractions obtained from the template fitting. Differences from the template-fit signal fractions quantify the impact of inaccuracies in the MC templates and are taken as a systematic uncertainty. The cut-and-correct $f^{\text{sig}}$ values were evaluated using
Δ$p_T/p_T^{ID}|_{cut} = 0.1$. The obtained signal fractions were found to be systematically higher than the results from the template fits at both low and high $p_T$ and in both the $pp$ and Pb+Pb data. The relative difference is largest in the lowest $p_T$ interval where it is ~11% and 6% for the $pp$ and Pb+Pb data, respectively. It decreases with increasing $p_T$, and for the highest $p_T$ interval, is ~6% and 3% for the $pp$ and Pb+Pb data, respectively.

The $pp$ cross sections and Pb+Pb per-event yields are not corrected for any bin migrations that result from the muon momentum resolution. An evaluation of MC bin-by-bin correction factors gives values that are typically within 1% (2%) of unity for $pp$ (Pb+Pb) data. These corrections are sufficiently small that they are not applied to the data. However, the deviations from unity are included in the systematic uncertainties of the cross sections and per-event yields.
FIG. 8. Comparison of the Pb+Pb heavy-flavor muon $R_{AA}$ measured in this analysis to similar measurements for muons at forward rapidity ($2.5 < y < 4$) and heavy-flavor electrons at midrapidity ($|y| < 0.6$) from the ALICE Collaboration. The error bars represent systematic and statistical uncertainties added in quadrature. The $\langle T_{AA} \rangle$ errors are identical between the three measurements and are excluded from the comparison.

FIG. 9. Comparison of the Pb+Pb heavy-flavor muon $R_{AA}$ measured in this analysis to the $R_{AA}$ for inclusive charged hadrons from ATLAS and the $R_{AA}$ for identified $D^0$ mesons from the CMS Collaboration. The error bars represent systematic and statistical uncertainties added in quadrature. The $\langle T_{AA} \rangle$ errors are identical between the three measurements and are excluded from the comparison. The inclusive charged hadron $R_{AA}$ values shown in the top left panel are for the 0–5% centrality interval.
FIG. 10. The $p_T$ dependence of the Pb+Pb heavy-flavor muon $v_2$. Results are shown for both the EP and SP methods. Each panel represents a different centrality interval. The error bars and shaded bands represent statistical and total uncertainties, respectively, and are shown only for the EP $v_2$. The horizontal dashed lines indicate $v_2 = 0$.

The measured $pp$ cross section has an additional normalization systematic uncertainty of 3.1% due to uncertainties in the integrated luminosity.

For the $R_{AA}$ measurement, the systematic uncertainties from the $pp$ cross section and Pb+Pb per-event yields are propagated as if they are correlated, i.e., the systematic variations are simultaneously performed in the $pp$ and Pb+Pb data and the change in the $R_{AA}$ value is taken as the systematic uncertainty. Besides the systematic uncertainties from the $pp$ cross section and Pb+Pb per-event yields, additional systematic uncertainties in the $R_{AA}$ measurement come from theoretical uncertainties in $\langle T_{AA} \rangle$, which are listed in Table I. Table II summarizes the final experimental systematic uncertainties in $R_{AA}$. The total uncertainty is obtained by adding the individual uncertainties in quadrature.

B. Systematic uncertainties in $v_n$

The sources of the systematic uncertainties in the $v_n$ measurements are primarily the same as those in the $R_{AA}$ measurements (Sec. IV A). However, several sources of systematic uncertainty that affect $R_{AA}$ do not have a significant effect on the $v_n$ values. The $v_n$ measurements are independent of the trigger and tracking efficiencies. While these efficiencies have an impact on the absolute muon yields, the $v_n$ values, which measure the relative or fractional modulation in yields, are insensitive to them. Therefore, the uncertainties in the efficiencies do not have any effect on the $v_n$ measurements. Varying the muon selection as described in Sec. IV A changes the measured value of $v_2$ by $(1–2) \times 10^{-3}$ below $p_T$ of 6 GeV. The $p_T^{MS}$ criterion variation changes the measured value of $v_2$ by $(0.5–1) \times 10^{-3}$ for $p_T < 6$ GeV. At higher $p_T$ the effect of this criterion on $v_2$ is about $0.2 \times 10^{-3}$. For $v_3$ and $v_4$ the effect of the $p_T^{MS}$ criterion is $(0.5–1) \times 10^{-3}$ across the measured $p_T$ range. The effects of the muon selection and the $p_T^{MS}$ criterion are evaluated not just by applying the selection in the data but also by rebuilding the templates in the MC simulation while applying the variations, and then repeating the entire analysis. The variation in the shape of the background template, when varying the relative contribution of the pion and kaon backgrounds, results in variations in the $v_n$ values that are less than $0.5 \times 10^{-3}$ across most of the centrality and $p_T$ ranges. The systematic uncertainty in $v_n$ due to $p_T$-resolution effects is estimated to be less than 1% (relative) for $p_T < 10$ GeV. This estimate is obtained by first determining the $p_T$ resolution using MC simulation.
FIG. 11. The centrality dependence of the Pb+Pb heavy-flavor muon $v_2$ (the horizontal scale decreases in centrality). Each panel represents a different $p_T$ interval. The error bars and shaded bands represent statistical and total uncertainties, respectively. The dashed lines indicate $v_2 = 0$. The results are for the EP method.

(Sec. III B), and then evaluating the change in the $v_2$ values when smearing the $p_T$ of the reconstructed muons by this resolution. The uncertainty arising from the $p_T$ resolution is treated as a fractional uncertainty; since if $v_n$ changes, then the $p_T$ resolution effects that result in migration of muons from one $p_T$ interval to an adjacent one also increase proportionally. For $p_T > 10$ GeV, the systematic uncertainties from all the above sources are partially correlated with the statistical uncertainties, and are thus somewhat larger.

Additional systematic uncertainties that affect only the $v_n$ but not the $R_{AA}$ measurements are the uncertainty in the EP resolution for $\Psi_1$ and the jet bias correction discussed in Sec. III D. The uncertainty in the EP resolution is a relative uncertainty and depends only on the centrality. It varies between 1% and 5.5% depending on the harmonic and centrality. The systematic uncertainty associated with the jet bias correction is the leading uncertainty in the measurement. The absolute value of this uncertainty depends on the centrality and the harmonic order but is independent of $p_T$. It increases monotonically from central to peripheral events and is much larger for $v_3$ and $v_4$ than for $v_2$. Table III summarizes the systematic uncertainties for the $v_n$ in three different $p_T$ ranges and for two centrality intervals. The uncertainties associated with the $p_T$ resolution and EP resolution are intrinsically fractional uncertainties and are listed as percentages. All other uncertainties are listed as absolute values.
FIG. 12. The $p_T$ dependence of the Pb+Pb heavy-flavor muon $v_3$. Results are shown for both the EP and SP methods. Each panel represents a different centrality interval. The error bars and shaded bands represent statistical and total uncertainties, respectively, and are shown only for the EP $v_3$. The horizontal dashed lines indicate $v_3 = 0$.

V. RESULTS

A. Heavy-flavor muon $R_{AA}$

Figure 5 shows the measured heavy-flavor muon cross sections, calculated via Eq. (3), in the $\sqrt{s} = 2.76$-TeV $pp$ data as a function of the muon $p_T$. The error bars show statistical uncertainties resulting from combining the statistical uncertainties of $\Delta N_\mu$ and $f^{3\ell}$. The measured cross sections are compared with fixed-order plus next-to-leading-logarithm (FONLL) [75–78] calculations using CTEQ 6.6 PDFs [79]. The FONLL calculations are based on three main components: (1) the heavy-quark production cross-section calculated in perturbative QCD by matching the fixed-order next-to-leading-order (NLO) terms with the next-to-leading-logarithms (NLLs) high-$p_T$ resummation, (2) the nonperturbative heavy-flavor fragmentation functions determined from $e^+e^-$ collisions and extracted in the same framework, and (3) the decays of the heavy hadrons to leptons using decay tables and form factors from $B$ factories. The middle panel of Fig. 5 presents the ratios of the measured and FONLL cross sections. The FONLL calculation agrees with the data within systematic uncertainties. The individual contributions of the bottom and charm quarks to the heavy-flavor muon cross section obtained from the FONLL calculations are compared in the lower panel of Fig. 5. It is seen that at 4 GeV the contribution of the bottom quark to the muon cross section is about 40% of that of the charm quark. The relative contribution increases monotonically with the muon $p_T$, and at $p_T = 14$ GeV, the contributions from bottom and charm decays are comparable.

Figure 6 shows the differential per-event heavy-flavor muon yields in Pb+Pb collisions [Eq. (4)] scaled by the corresponding $\langle T_{AA} \rangle$ for the centrality intervals in this analysis. The statistical uncertainties are the combined statistical uncertainties of $\Delta N_\mu$ and $f^{3\ell}$. Figure 6 also compares the $\langle T_{AA} \rangle$ scaled yields to the measured $pp$ cross section. There are significant differences between the scaled Pb+Pb yields and the $pp$ cross section, which monotonically increase with increasing centrality.

The heavy-flavor muon $R_{AA}$ is calculated according to Eq. (1) using the results in Fig. 6 and is shown in Fig. 7. The parameter $R_{AA}$ does not depend on $p_T$ within the uncertainties of the measurement. This is of note because the suppression of bottom and charm quarks in the quark-gluon plasma (QGP) is expected to be different, and the FONLL calculations show that the contribution of bottom and charm quarks changes...
with $p_T$ in the $pp$ case, as shown in Fig. 5. The parameter $R_{AA}$ decreases between peripheral 40–60% collisions, where it is about 0.65, to more central collisions, reaching a value of about 0.35 in the 0–10% centrality interval.

Figure 8 shows a comparison of the $R_{AA}$ measurements in this paper with similar measurements for muons at forward rapidity ($2.5 < y < 4$) [20] and heavy-flavor electrons at midrapidity ($|y| < 0.6$) [47] from the ALICE Collaboration. In general, the results are consistent; however, the present measurements have considerably smaller uncertainties.

Figure 9 compares the $R_{AA}$ measurement presented in this paper with the $R_{AA}$ of inclusive charged hadrons [42] at $\sqrt{s_{NN}} = 2.76$ TeV and identified $D^0$ mesons [80] from the CMS Collaboration at $\sqrt{s_{NN}} = 5.02$ TeV. The $R_{AA}$ from $D^0$ analyses is similar to that of inclusive hadrons for $p_T > 5$ GeV [80], implying that the charm suppression is very similar to that for the light quarks and gluons. On the other hand, the heavy-flavor muon $R_{AA}$, which includes contributions from bottom and charm, is observed to be larger than that of inclusive hadrons. This would imply a significantly smaller suppression for muons from the decays of $b$ hadrons. One caveat is that the $D^0$ $p_T$ and the HF muon $p_T$ are related differently to the $p_T$ of the HF quark that produced them. However, this effect is mitigated by the relatively weak $p_T$ dependence of both the $D^0$ and HF muon $R_{AA}$ over the 4–14-GeV $p_T$ range.

B. Heavy-flavor muon $v_2$

Figure 10 shows the $v_2$ values measured using the EP method as a function of $p_T$ for the five centrality intervals in this analysis, including the statistical and total uncertainties. The evaluation of the total uncertainty includes the correlation between the statistical uncertainties and the systematic uncertainties that are proportional to $v_2$, i.e., the relative uncertainties associated with the EP and $p_T$ resolutions. This correlation arises because as the measured $v_2$ is varied within its statistical uncertainty, the relative uncertainties that are proportional to $v_2$ also vary. The other (absolute) systematic uncertainties are added in quadrature to the correlated uncertainty to get the total uncertainty. Over the 10–40% centrality range, $v_2$ is largest at the lowest measured $p_T$ of 4 GeV and decreases for higher $p_T$. However, in the 0–10% and 40–60% centrality intervals, no clear $p_T$ dependence is visible. For all centralities, a significantly nonzero $v_2$ is observed up to a $p_T$ of 12 GeV. Figure 10 also shows the $v_2^{SP}$ values, which are slightly higher than the EP values. The systematic uncertainties and a significant fraction of the statistical

\begin{figure}
\centering
\includegraphics[width=\textwidth]{fig13.png}
\caption{The $p_T$ dependence of the Pb+Pb heavy-flavor muon $v_4$. Results are shown for both the EP and SP methods. Each panel represents a different centrality interval. The error bars and shaded bands represent statistical and total uncertainties, respectively, and are shown only for the EP $v_4$. The horizontal dashed lines indicate $v_4 = 0$.}
\end{figure}
uncertainties are correlated between the EP and SP $v_2$ values, and for clarity are not shown for the $v_2^{SP}$. These measurements are consistent with previous $v_2$ measurements of heavy-flavor muons [21] and heavy-flavor electrons [81] from the ALICE Collaboration, but have significantly smaller statistical and systematic uncertainties, and are performed over wider centrality and $p_T$ ranges.

Figure 11 shows the $v_2$ obtained from the EP method plotted as a function of centrality for different $p_T$ intervals. For $p_T$ in the range 4–8 GeV, the centrality dependencies of the heavy-flavor muon $v_2$ are qualitatively similar in shape, but considerably smaller in magnitude, to those for charged hadrons of similar $p_T$ [50,52]. In this $p_T$ range, the $v_2$ first increases from central to midcentral events, reaches a maximum between 20% and 40% centrality, and then decreases. Over the $p_T$ range of 8–12 GeV, some deviation from this trend is observed, with the $v_2$ increasing monotonically from central to peripheral events. However, the associated statistical and systematic uncertainties are considerably larger. This monotonically increasing centrality dependence of the $v_2$ at high $p_T$ is also seen in the inclusive charged hadron $v_2$ [50,52]. For the highest $p_T$ interval of $12 < p_T < 14$ GeV, the statistical and systematic errors are too large to identify a clear centrality dependence of $v_2$.

Figure 12 shows the $p_T$ dependence of $v_3$. At a given $p_T$ and centrality, $v_3$ is a factor of 2–3 smaller than the corresponding $v_2$. As with $v_2$, $v_3$ also decreases with increasing $p_T$ over the 4–8-GeV $p_T$ range. At higher $p_T$, the statistical uncertainties are too large to observe clear $p_T$-dependent trends. The parameter $v_3$ shows a much weaker variation with centrality: the $v_3$ values at a given $p_T$ are consistent within uncertainties across the different centrality intervals. These features for the centrality and $p_T$ dependence are consistent with observations of the inclusive charged-hadron $v_1$ [52]. Figure 13 shows the $p_T$ dependence of $v_3$. The statistical uncertainties in $v_3$ do not allow inference of any significant $p_T$- or centrality-dependent trends.

C. Comparison with theoretical models

In this section, the measured $R_{AA}$ and $v_2$ values are compared with calculations from the TAMU transport model [82] and the DABMod model [83]. TAMU is a transport model for heavy flavor within the QGP and subsequent hadronic phase. The initial heavy-quark spectra used in the model are obtained from FONLL calculations, accounting for shadowing effects in Pb+Pb collisions. The space-time evolution of the bulk QGP medium, in which the heavy quarks diffuse, is modeled using ideal relativistic hydrodynamics, tuned to reproduce the charged-hadron $p_T$ spectra and inclusive elliptic flow measured in Pb+Pb collisions at the LHC. The initial conditions for the hydrodynamic modeling are obtained from the Glauber model and do not include initial state fluctuations or initial flow. After this tuning, there are no free parameters in the model. The hadronization of heavy-flavor quarks is done partially via recombination of heavy quarks with light-flavor hadrons in the QGP and partially by fragmentation. Finally, the diffusion of heavy-flavor hadrons in the hadronic phase is continued until kinetic freeze-out. DABMod is an energy-loss model for heavy quarks traversing the QGP. The energy loss is a parametrized analytic function of the velocity of the heavy quark and the local temperature. The initial $p_T$ distribution of
heavy quarks is obtained from FONLL calculations. The underlying QGP is modeled using (2+1)-dimensional relativistic viscous hydrodynamics including event-by-event fluctuations in the initial conditions and subsequent hydrodynamic expansion. All the hydrodynamic parameters are tuned to describe the experimental flow data at low $p_T$. The heavy quarks are evolved on top of the hydrodynamic underlying event until they reach a decoupling temperature below which they are hadronized via fragmentation. Any subsequent hadronic rescattering is neglected. The DABMod model calculations are available for $R_{AA}$ and $v_2$ for all the centrality intervals over which the measurements are performed in this paper. The TAMU calculations for $R_{AA}$ are available for the 0–10%, 20–40%, and 40–60% centrality intervals, and for $v_2$ for the 20–40% and 40–60% centrality intervals only.

Figure 14 compares the measured heavy-flavor muon $R_{AA}$ values with theoretical calculations from the TAMU and DABMod models. Generally, the TAMU model describes many features of the data well, especially the weak $p_T$ dependence of $R_{AA}$, while DABMod only reproduces the measured $R_{AA}$ for $p_T > 12$ GeV. The failure of the DABMod model at low $p_T$ is understood to result from incomplete modeling of heavy-flavor suppression for $p_T \lesssim m_b$. The TAMU model predicts a larger suppression in the 40–60% centrality interval and a lower suppression in the 0–10% centrality interval than what is measured. Thus, the range of the suppression seen in the data is larger than in the TAMU model. As stated above, the TAMU model does not implement event-by-event fluctuations in the initial geometry, which are known to affect the dynamical evolution of bulk medium [63,84]. This may be one of the possible reasons for the smaller dynamical range of $R_{AA}$ predicted by the model.

Figure 15 compares the measured heavy-flavor $v_2$ values with calculations from the TAMU and DABMod models. The DABMod $v_2$ values are systematically larger than the TAMU values and closer to the measured $v_2$. Unlike TAMU, the DABMod calculations include event-by-event fluctuations which are known to increase $v_n$ [63,84]. This could be a possible reason for the systematically larger $v_2$ values obtained in the DABMod model. The DABMod calculations are consistent with the measured values for $p_T > 6$ GeV for all centralities. However, for $4 < p_T < 6$ GeV and for the 10–40% centrality range, the calculated values are significantly smaller than the measured $v_2$ values. The TAMU $v_2$ values are significantly smaller than the measured $v_2$ over the $4 < p_T < 10$ GeV $p_T$ range. Figure 16 compares the measured $v_3$ values
FIG. 16. Comparison of the Pb+Pb heavy-flavor muon $v_3$ with calculations from the DABMod model. Each panel represents a different centrality interval. For the data, the error bars and shaded bands represent statistical and total uncertainties, respectively. For the model calculations, the bands represent theoretical systematic uncertainties.

to calculations from the DABMod model. Features similar to the $v_2$ comparison are observed; the model predictions are smaller than the measured $v_3$ for $4 < p_T < 6$ GeV but become consistent with the data at higher $p_T$. The DABMod calculations are also compared with the $v_4$ measurements. However, the large experimental uncertainties do not allow detailed comparisons with the model predictions.

VI. CONCLUSION

This paper presents ATLAS measurements of heavy-flavor muon production in 0.14 nb$^{-1}$ of $\sqrt{s_{\text{NN}}} = 2.76$ TeV Pb+Pb collisions and 570 nb$^{-1}$ of $\sqrt{s} = 2.76$ TeV pp collisions at the LHC. The measurements are performed over the transverse momentum range of $4 < p_T < 14$ GeV. Backgrounds arising from in-flight pion and kaon decays, hadronic showers, and misreconstructed muons are statistically removed using a template-fitting procedure based on the relative difference between the muon track momenta in the muon spectrometer and inner detector, corrected for energy loss in the calorimeter system. The heavy-flavor muon differential cross sections and per-event yields are measured in $pp$ and Pb+Pb collisions, respectively. The nuclear modification factor $R_{AA}$ calculated from these quantities shows a centrality-dependent suppression that does not depend on $p_T$ within uncertainties. In the 0–10% centrality interval, $R_{AA} \sim 0.35$. In Pb+Pb collisions, measurements of the heavy-flavor muon yields as a function of $\phi - \Psi_n$, the azimuthal angle of the muons relative to the event-plane angles, show a clear sinusoidal modulation of the yield in all centrality intervals. The heavy-flavor muon $v_n$, for $n = 2–4$, is measured in Pb+Pb collisions as a function of $p_T$ for five centrality intervals covering the 0–60% centrality range. Significant $v_2$ values up to about 0.08 are observed at $p_T = 4$ GeV. In the 10–20%, 20–30%, and 30–40% intervals, the $v_2$ decreases with $p_T$ but is still significant at 10 GeV. At fixed $p_T$, the $v_2$ values show a systematic variation with centrality which is typical of elliptic-flow measurements. For most centrality intervals, $v_2$ also decreases with increasing $p_T$ over the 4–8-GeV $p_T$ range. For $p_T > 8$ GeV, the statistical uncertainties in the measured $v_2$ values are too large to discern any $p_T$-dependent trends. At a given $p_T$ and centrality, the $v_3$ values are smaller than the $v_2$ values by a factor of 2–4. Further, $v_3$ shows a much weaker centrality dependence than $v_2$. Conclusions about any $p_T$- or centrality-dependent trends in the $v_4$ are limited by the statistical precision.

The measured $R_{AA}$ and $v_2$ are also compared with theoretical predictions from the TAMU and DABMod models. The $R_{AA}$ values from the TAMU model show a weak $p_T$ dependence over the 4–14-GeV $p_T$ range, qualitatively
similar to the measured $R_{AA}$. However, the predicted $R_{AA}$ values are smaller than the measured values in the 40–60% centrality interval, and larger than the measured values in the 0–10% centrality interval. On the other hand, the DABMod model predicts a strong $p_T$ dependence for $R_{AA}$, which is not observed in the data. The $R_{AA}$ value at $p_T = 4$ GeV predicted by DABMod is significantly smaller than the measured values but increases with increasing $p_T$ and becomes comparable to the measured values at $p_T = 12$ GeV. For $v_2$, the TAMU and DABMod qualitatively reproduce the observed $p_T$ dependence but the DABMod calculations are more consistent with the measured values. Thus both models fail to simultaneously reproduce $v_2$ and $R_{AA}$ over the measured $p_T$ range.

The $R_{AA}$ values measured here for $|y| < 1$ and $v_2$ values for $|y| < 2$ are compatible with, but are substantially more precise than, similar measurements of heavy-flavor muons at forward rapidity (2.5 < $y$ < 4) and heavy-flavor electrons at midrapidity ($|y| < 0.6$) from the ALICE Collaboration. Thus, they should provide improved insight into the propagation of heavy quarks in the quark-gluon plasma created in Pb+Pb collisions.

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; YerPhl, Armenia; ARC, Australia; BMWF and FWF, Austria; ANAS, Azerbaijan; SSTC, Belarus; CNPq and FAPESP, Brazil; NSERC, NRC and CFI, Canada; CERN; CONICYT, Chile; CAS, MOST and NSFC, China; COLCIENCIAS, Colombia; MSMT CR, MPO CR and VSC CR, Czech Republic; DNRF and DNSRC, Denmark; IN2P3-CNRS, CEA-DRF/IRFU, France; SRNSFG, Georgia; BMBF, HGF, and MPG, Germany; GSRT, Greece; RGC, Hong Kong SAR, China; ISF, I-CORE and Benoziyo Center, Israel; INFN, Italy; MEXT and JSPS, Japan; CNRST, Morocco; NWO, Netherlands; RCN, Norway; MNiSW and NCN, Poland; FCT, Portugal; MNE/IFA, Romania; MES of Russia and NRC KI, Russian Federation; JINR; MESTD, Serbia; MSSR, Slovakia; ARRS 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; DOE and NSF, United States of America. In addition, individual groups and members have received support from BCKDF, the Canada Council, CANARIE, CRC, Compute Canada, FORNT, and the Ontario Innovation Trust, Canada; EPLANET, ERC, ERDF, FP7, Horizon 2020 and Marie Sklodowska-Curie Actions, European Union; Investissements d’Avenir Labex and Idex, ANR, Région Auvergne and Fondation Partager le Savoir, France; DFG and AVH Foundation, Germany; Herakleitos, Thales and Aristeia programmes co-financed by EU-ESF and the Greek NSRF; BSF, GIF and Minerva, Israel; BRF, Norway; CERCA Programme Generalitat de Catalunya, Generalitat Valenciana, Spain; 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. [85].


[37] ATLAS Collaboration, Centrality, rapidity and transverse momentum dependence of isolated prompt photon production in lead–lead collisions at $\sqrt{s_{NN}} = 2.76$ TeV measured with the ATLAS detector, Phys. Rev. C 93, 034914 (2016).


044905-22