Editors' Suggestion

Detailed report on the measurement of the positive muon anomalous magnetic moment to 0.20 ppm

D. P. Aguillard,[33] T. Albahri,[30] D. Allspach,[7] A. Anisenkov,[4,a] K. Badgley,[7] S. Baeßler,[35,b] I. Bailey,[17,c] L. Bailey,[27] V. A. Baranov,[15,*] E. Barlas-Yucel,[28] T. Barrett,[6] E. Barzi,[7] F. Bedeschi,[10] M. Berz,[18] M. Bhattacharya,[7] H. P. Binney,[36] P. Bloom,[19] J. Bono,[7] E. Bottalico,[30,d] T. Bowcock,[30] S. Braun,[36] M. Bressler,[32] G. Cantatore,[12,e] R. M. Carey,[2] B. C. K. Casey,[7] D. Cauz,[26,f] R. Chakraborty,[29] A. Chapelain,[6] S. Chappa,[7] S. Charity,[30] C. Chen,[23,22] M. Cheng,[28] R. Chislett,[27] Z. Chu,[22,g] T. E. Chupp,[33] C. Claessens,[36] M. E. Convery,[7] S. Corrodi,[1] L. Cotrozzi,[10,30,h] J. D. Crnkovic,[7] S. Dabagov,[8,I] P. T. Debevec,[28] S. Di Falco,[10] G. Di Sciascio,[11] S. Donati,[10,h] B. Drendel,[7] A. Driutti,[10,h] V. N. Duginov,[15,*] M. Eads,[20] A. Edmonds,[2,37] J. Esquivel,[7] M. Farooq,[33] R. Fatemi,[29] C. Ferrari,[10,j] M. Fertl,[14] A. T. Fienberg,[36] A. Fioretti,[10,j] D. Flay,[32] S. B. Foster,[2] H. Friedsam,[7] N. S. Froemming,[20] C. Gabbanini,[10,j] I. Gaines,[7] M. D. Galati,[10,h] S. Ganguly,[7] A. Garcia,[36] J. George,[32,k] L. K. Gibbons,[6] A. Gioiosa,[25,l] K. L. Giovanetti,[13] P. Girotti,[10] W. Gohn,[29] L. Goodenough,[7] T. Gorringe,[29] J. Grange,[33] S. Grant,[1,27] F. Gray,[21] S. Haciomeroglu,[5,m] T. Halewood-Leagas,[30] D. Hampai,[8] F. Han,[29] J. Hempstead,[36] D. W. Hertzog,[36] G. Hesketh,[27] E. Hess,[10] A. Hibbert,[30] Z. Hodge,[36] K. W. Hong,[35] R. Hong,[1,29] T. Hu,[23,22] Y. Hu,[22,g] M. Iacovacci,[9,n] M. Incagli,[10] P. Kammel,[36] M. Kargiantoulakis,[7] M. Karuza,[12,o] J. Kaspar,[36] D. Kawall,[32] L. Kelton,[29] A. Keshavarzi,[31] D. S. Kessler,[32] K. S. Khaw,[23,22] Z. Khechadoorian,[6] N. V. Khomutov,[15] B. Kiburg,[7] M. Kiburg,[7,19] O. Kim,[34] N. Kinnaird,[2] E. Kraegeloh,[33] V. A. Krylov,[15] N. A. Kuchinskiy,[15] K. R. Labe,[6] J. LaBounty,[36] M. Lancaster,[31] S. Lee,[5] B. Li,[22,1,p] D. Li,[22,q] L. Li,[22,g] I. Logashenko,[4,a] A. Lorente Campos,[29] Z. Lu,[22,g] A. Lucà,[7] G. Lukicov,[27] A. Lusiani,[10,r] A. L. Lyon,[7] B. MacCoy,[36] R. Madrak,[7] K. Makino,[18] S. Mastroianni,[9] J. P. Miller,[2] S. Miozzi,[11] B. Mitra,[34] J. P. Morgan,[7] W. M. Morse,[3] J. Mott,[7,2] A. Nath,[9,n] J. K. Ng,[23,22] H. Nguyen,[7] Y. Oksuzian,[1] Z. Omarov,[16,5] R. Osofsky,[36] S. Park,[5] G. Pauletta,[26,*,f] G. M. Piacentino,[25,l] R. N. Pilato,[30] K. T. Pitts,[28,s] B. Plaster,[29] D. Počanić,[35] N. Pohlman,[20] C. C. Polly,[7] J. Price,[30] B. Quinn,[34] M. U. H. Qureshi,[14] S. Ramachandran,[1,k] E. Ramberg,[7] R. Reimann,[14] B. L. Roberts,[2] D. L. Rubin,[6] M. Sakurai,[27] L. Santi,[26,f] C. Schlesier,[28,t] A. Schreckenberger,[7] Y. K. Semertzidis,[5,16] D. Shemyakin,[4,a] M. Sorbara,[11,u] J. Stapleton,[7] D. Still,[7] D. Stöckinger,[24] C. Stoughton,[7] D. Stratakis,[7] H. E. Swanson,[36] G. Sweetmore,[31] D. A. Sweigart,[6] M. J. Syphers,[20] D. A. Tarazona,[6,30,18] T. Teubner,[30] A. E. Tewsley-Booth,[29,33] V. Tishchenko,[3] N. H. Tran,[2,v] W. Turner,[30] E. Valetov,[18] D. Vasilkova,[27,30] G. Venanzoni,[30,d] V. P. Volnykh,[15] T. Walton,[7] A. Weisskopf,[18] L. Welty-Rieger,[7] P. Winter,[1] Y. Wu,[1] B. Yu,[34] M. Yucel,[7] Y. Zeng,[23,22] and C. Zhang[30]

(Muon $g-2$ Collaboration)

[1]*Argonne National Laboratory, Lemont, Illinois, USA*
[2]*Boston University, Boston, Massachusetts, USA*
[3]*Brookhaven National Laboratory, Upton, New York, USA*
[4]*Budker Institute of Nuclear Physics, Novosibirsk, Russia*
[5]*Center for Axion and Precision Physics (CAPP) / Institute for Basic Science (IBS), Daejeon, Republic of Korea*
[6]*Cornell University, Ithaca, New York, USA*
[7]*Fermi National Accelerator Laboratory, Batavia, Illinois, USA*
[8]*INFN, Laboratori Nazionali di Frascati, Frascati, Italy*
[9]*INFN, Sezione di Napoli, Naples, Italy*
[10]*INFN, Sezione di Pisa, Pisa, Italy*
[11]*INFN, Sezione di Roma Tor Vergata, Rome, Italy*
[12]*INFN, Sezione di Trieste, Trieste, Italy*
[13]*Department of Physics and Astronomy, James Madison University, Harrisonburg, Virginia, USA*
[14]*Institute of Physics and Cluster of Excellence PRISMA+, Johannes Gutenberg University Mainz, Mainz, Germany*
[15]*Joint Institute for Nuclear Research, Dubna, Russia*
[16]*Department of Physics, Korea Advanced Institute of Science and Technology (KAIST), Daejeon, Republic of Korea*
[17]*Lancaster University, Lancaster, United Kingdom*
[18]*Michigan State University, East Lansing, Michigan, USA*
[19]*North Central College, Naperville, Illinois, USA*

[20]*Northern Illinois University, DeKalb, Illinois, USA*
[21]*Regis University, Denver, Colorado, USA*
[22]*School of Physics and Astronomy, Shanghai Jiao Tong University, Shanghai, China*
[23]*Tsung-Dao Lee Institute, Shanghai Jiao Tong University, Shanghai, China*
[24]*Institut für Kern- und Teilchenphysik, Technische Universität Dresden, Dresden, Germany*
[25]*Università del Molise, Campobasso, Italy*
[26]*Università di Udine, Udine, Italy*
[27]*Department of Physics and Astronomy, University College London, London, United Kingdom*
[28]*University of Illinois at Urbana-Champaign, Urbana, Illinois, USA*
[29]*University of Kentucky, Lexington, Kentucky, USA*
[30]*University of Liverpool, Liverpool, United Kingdom*
[31]*Department of Physics and Astronomy, University of Manchester, Manchester, United Kingdom*
[32]*Department of Physics, University of Massachusetts, Amherst, Massachusetts, USA*
[33]*University of Michigan, Ann Arbor, Michigan, USA*
[34]*University of Mississippi, University, Mississippi, USA*
[35]*University of Virginia, Charlottesville, Virginia, USA*
[36]*University of Washington, Seattle, Washington, USA*
[37]*City University of New York at York College, Jamaica, New York, USA*

(Received 5 March 2024; accepted 21 May 2024; published 8 August 2024)

We present details on a new measurement of the muon magnetic anomaly, $a_\mu = (g_\mu - 2)/2$. The result is based on positive muon data taken at Fermilab's Muon Campus during the 2019 and 2020 accelerator runs. The measurement uses 3.1 GeV/c polarized muons stored in a 7.1-m-radius storage ring with a 1.45 T uniform magnetic field. The value of a_μ is determined from the measured difference between the muon spin precession frequency and its cyclotron frequency. This difference is normalized to the strength of the magnetic field, measured using nuclear magnetic resonance. The ratio is then corrected for small contributions from beam motion, beam dispersion, and transient magnetic fields. We measure $a_\mu = 116592057(25) \times 10^{-11}$ (0.21 ppm). This is the world's most precise measurement of this quantity and represents a factor of 2.2 improvement over our previous result based on the 2018 dataset. In combination, the two datasets yield $a_\mu(\mathrm{FNAL}) = 116592055(24) \times 10^{-11}$ (0.20 ppm). Combining this with the measurements from Brookhaven

[*]Deceased.
[a]Also at Novosibirsk State University.
[b]Also at Oak Ridge National Laboratory.
[c]Also at The Cockcroft Institute of Accelerator Science and Technology, Daresbury, United Kingdom.
[d]Also at INFN, Sezione di Pisa, Pisa, Italy.
[e]Also at Università di Trieste, Trieste, Italy.
[f]Also at INFN Gruppo Collegato di Udine, Sezione di Trieste, Udine, Italy.
[g]Also at Shanghai Key Laboratory for Particle Physics and Cosmology; also at Key Lab for Particle Physics, Astrophysics and Cosmology (MOE).
[h]Also at Università di Pisa, Pisa, Italy.
[i]Also at Lebedev Physical Institute and NRNU MEPhI.
[j]Also at Istituto Nazionale di Ottica—Consiglio Nazionale delle Ricerche, Pisa, Italy.
[k]Now at Alliance University, Bangalore, India.
[l]Also at INFN, Sezione di Roma Tor Vergata, Rome, Italy.
[m]Now at Istinye University, Istanbul, Türkiye.
[n]Also at Università di Napoli, Naples, Italy.
[o]Also at University of Rijeka, Rijeka, Croatia.
[p]Also at Research Center for Graph Computing, Zhejiang Lab, Hangzhou, Zhejiang, China.
[q]Also at Shenzhen Technology University, Shenzhen, Guangdong, China.
[r]Also at Scuola Normale Superiore, Pisa, Italy.
[s]Now at Virginia Tech, Blacksburg, Virginia, USA.
[t]Now at Wellesley College, Wellesley, Massachusetts, USA.
[u]Also at Università di Roma Tor Vergata, Rome, Italy.
[v]Now at Institute for Interdisciplinary Research in Science and Education (ICISE), Quy Nhon, Binh Dinh, Vietnam.

National Laboratory for both positive and negative muons, the new world average is $a_\mu(\text{exp}) = 116592059(22) \times 10^{-11}$ (0.19 ppm).

DOI: 10.1103/PhysRevD.110.032009

I. INTRODUCTION

The anomalous magnetic moment of a charged lepton arises from radiative corrections and interactions with virtual particles. It can be calculated for Standard Model (SM) interactions with high precision. Measurements of the muon magnetic anomaly, expressed as $a_\mu = (g_\mu - 2)/2$, with similar or greater precision thus challenge the SM calculations and probe possible beyond the Standard Model (BSM) physics. Measurement of the electron a_e provides a 0.13-ppt determination of g_e, which is mostly sensitive to electromagnetic interactions [1]. The muon, due to its greater mass, is approximately 43000 times more sensitive to BSM interactions of new heavy particles.

In a series of measurements with both positive and negative muons, the E821 collaboration at Brookhaven National Laboratory (BNL) determined a_μ with a relative precision of 0.54 ppm [2] and found a discrepancy with the SM calculation of about three standard deviations at the time. Improved precision of the SM prediction in subsequent years led to increased significance, and a_μ became one of the largest measured discrepancies with the SM and a possible signal of BSM physics [3,4]. On April 7, 2021, the Muon $g-2$ collaboration released the first result for a_μ based on the Run-1 2018 data campaign at Fermilab [5–8], which was consistent with the BNL results. Meanwhile, newer SM calculations [9] challenge the 2020 $g-2$ Theory Initiative White Paper [10] recommendation. In 2023, the collaboration published the Run-2/3 result [11]. This paper provides the analysis details of that result.

The magnetic anomaly of 3.1 GeV muons is measured in a magnetic storage ring with a uniform vertical magnetic field $\vec{B}$ and weakly focusing quadrupole electric fields $\vec{E}$. For $g_\mu > 2$, the muon spin precession frequency $\vec{\omega}_S$ is greater than the cyclotron frequency $\vec{\omega}_C$, resulting in the anomalous-precession frequency $\vec{\omega}_a = \vec{\omega}_s - \vec{\omega}_c$. For relativistic muons on the ideal orbit with a perfectly uniform magnetic field,

$$\vec{\omega}_a = -a_\mu \frac{q}{m}\vec{B} + \frac{q}{m}\left[\left(a_\mu - \frac{1}{\gamma^2-1}\right)\frac{\vec{\beta}\times\vec{E}}{c} + a_\mu\left(\frac{\gamma}{\gamma+1}\right)(\vec{\beta}\cdot\vec{B})\vec{\beta}\right], \tag{1}$$

where q is the charge, m is the mass, β is the velocity ratio with respect to the speed of light, and γ is the Lorentz factor of the muon. The second term on the right-hand side, proportional to E, vanishes for $\gamma = \sqrt{(1+1/a_\mu)} \approx 29.3$. This corresponds to momentum $p_0 \approx 3.094$ GeV/c, called the "magic momentum." In the absence of vertical betatron motion, the muon velocity is perpendicular to $\vec{B}$, leading to cancellation of the third term.

The magnitude of the measured anomalous-precession frequency, corrected for the momentum spread, betatron motion, and beam-dynamics effects is proportional to $\tilde{B}$, the magnetic field magnitude averaged over the muon distribution in time and space. We express $\tilde{B}$ in terms of the measured nuclear magnetic resonance (NMR) frequency of protons in a spherical water sample at a reference temperature T_r

$$\tilde{\omega}'_p = \gamma'_p(T_r)\tilde{B}, \tag{2}$$

where γ'_p is the gyromagnetic ratio of protons in H_2O known with high precision at T_r. Combining the first term on the right-hand side of Eqs. (1) and (2) allows a_μ to be expressed as a ratio of frequencies,

$$a_\mu \propto \frac{\omega_a}{\tilde{\omega}'_p(T_r)} \equiv \mathcal{R}'_\mu(T_r). \tag{3}$$

Parity violation in the weak decay of the muon allows measurement of the anomalous-precession frequency ω_a. In the muon rest frame, the positron emission direction correlates with the muon spin direction, most strongly for high-energy positrons. In the laboratory frame, this results in a ω_a-dependent modulation of the positron energy spectrum. Fits to the positron time distribution extract the measured frequency ω_a^m. Details are provided in Sec. IV.

Five beam-dynamics-driven corrections are applied to the measured spin precession frequency ω_a^m. The electric-field correction C_e accounts for the electric field contribution due to the muon momentum spread. The pitch correction C_p accounts for the vertical betatron motion of the muons. C_{ml} accounts for the muon losses due to the finite aperture of the storage ring. The phase acceptance correction C_{pa} accounts for the injected muons' phases with respect to the detector acceptance, and finally, the differential decay corrections C_{dd} account for the correlation between spin phase and momentum of the muons. Details are provided in Sec. V.

The muon-averaged magnetic field expressed in the precession frequency of shielded protons $\tilde{\omega}'_p$ is reconstructed from a combination of mapping and tracking the magnetic field in the muon storage region and weighting by the reconstructed muon distribution $M(x, y, \phi, t)$, with x and y the horizontal and vertical transverse coordinates, ϕ the azimuth in the storage ring, and t the time. The magnetic field maps have to be corrected for transient perturbations that are synchronous with the muon injection

due to the eddy currents from the magnetic kick required to move the muons to stored orbit radius (B_K) and due to vibrations induced in the field plates of the pulsed electrostatic quadrupoles (B_Q). Details are provided in Sec. VI.

Including the corrections, we can schematically express the ratio of the measured frequencies as

$$\mathcal{R}'_\mu(T_r) = \frac{\omega_a^m(1 + C_e + C_p + C_{ml} + C_{dd} + C_{pa})}{\langle \omega'_p \times M \rangle (1 + B_K + B_Q)}, \tag{4}$$

where $\langle \omega'_p \times M \rangle$ represents the muon weighting of the magnetic field (Sec. VI).

Following an overview of the experimental setup in Sec. II, we describe the datasets, run conditions, and main differences compared to Run-1 in Sec. III A. The analysis and extraction of ω_a and beam-dynamics corrections are discussed in Secs. IV and V. The determination of $\tilde{\omega}'_p$ is detailed in Sec. VI. Consistency checks over the dataset and the calculation of a_μ are presented in Secs. VII and VIII, and our result is put into the context of the current SM calculation in Sec. IX. Appendices cover details of the analyses and the combination of results.

Throughout this paper, frequencies are expressed as angular frequencies (ω in rad/s) and rotation frequencies ($\omega/2\pi$ or f) as appropriate in the context.

II. THE MUON $g-2$ EXPERIMENTAL SETUP AND SIMULATION PACKAGES

A. Experimental setup

The Fermilab Muon $g-2$ (E989) Experiment uses the same magic-momentum measurement principle developed initially for the CERN III experiment [12]. Furthermore, the Fermilab experiment employs the same storage ring and muon injection principle of E821 at BNL [2] but has improved instrumentation for the magnetic field and muon spin precession frequency measurements.

The superconducting storage ring magnet is made of 12 segments each consisting of a continuous iron yoke [13]. The C shape of the magnet cross section faces the interior of the ring so that positrons from muon decay, which spiral inward, can travel unobstructed by the magnet yoke to detectors placed around the interior of the storage ring. The strong vertical magnetic field is generated by four liquid helium-cooled superconducting coils and shaped by 36 high-purity iron pole pieces on top and the bottom of the opening. To improve the field uniformity, edge shims and iron foils are used to control the transverse gradients and fine tune the magnetic field over the entire azimuthal and transverse storage volume. A set of magnetic coils with individually controlled currents run parallel to the muon beam above and below the vacuum chambers and are trimmed to achieve field uniformity in the storage region to better than one part per million [7] averaged around the ring. The magnet power supply is adjusted continuously by a feedback system that stabilizes the field measured by NMR probes. This compensates for effects such as the thermal expansion of the ring.

Every 1.4 s, a burst of eight bunches or fills every 10 ms, followed by the same pattern approximately 267 ms later, of $\mathcal{O}(10^5) \sim 96\%$ polarized positive muons are delivered to the storage ring [14]. The initial momentum distribution of a fill has a width of 1.6% centered on the magic momentum of $p_0 = 3.094$ GeV/c. Five collimators are positioned inside the storage ring to confine stable muon orbits within a torus of major radius $R \approx R_0$ and minor radius $r \approx 4.5$ cm. Per fill, approximately 5000 muons with a momentum spread around 0.15% rms are stored for up to 700 μs. The central orbit radius is $R_0 = 7.112$ m, with a cyclotron period of $T_c = 149.1$ ns at $B = 1.451$ T.

Before entering the storage ring, the muon beam passes through a scintillator detector and three scintillating fiber detectors. The scintillator detector is a 1-mm-thick plastic scintillator coupled via light guides to two photomultiplier tubes. This detector provides the time reference (called T_0) for each fill, the time profile of the beam, and the integrated beam intensity used for determining the beam storage efficiency and performing quality monitoring. After the T_0 detector, the muons pass through three scintillating fiber detectors that measure the horizontal and vertical beam profile before and after the injection. They comprise the inflector beam monitoring system. The first two are made of a 16×16 grid of 0.5-mm-diameter scintillating fibers read out by 1 mm^2 silicon photomultipliers (SiPMs). The third inflector beam monitoring system detector only has the vertical fibers to measure the horizontal plane profile. It can be deployed to either measure the profile at injection or multiple turns into beam storage. During normal data taking it is in a retracted position to avoid degrading the beam.

Muons tangentially enter the storage ring from a low-field region through a superconducting inflector magnet. This inflector magnet cancels the storage ring magnetic field locally and provides a virtually field-free injection channel. The particles are displaced 77 mm radially outward from the radial center of the storage region and are not on trajectories suitable for storage in the ring. A set of three fast nonferric pulsed magnetic kickers is placed a quarter turn downstream from the injection point. The kickers are composed of three 1.27-m-long aluminum plates. Pulsing the kickers at ~4.3 kA during the first turn after injection reduces the total magnetic field in the kicker region. This brief reduction deflects the muons onto the radially centered trajectory. Ideally, this pulse would last 120 ns, which is a typical length of injected muon bunches. However, significant upgrades to the system were required to reach a FWHM around the cyclotron period to minimize the kick on the second turn. In addition, reflections and eddy currents are induced that have been the subjects of extensive dedicated studies. Detailed characterization of the kicker system and the upgrade effort are described in Ref. [15].

Four electrostatic quadrupoles (ESQs) distributed around the storage ring provide vertical focusing. Each ESQ has a long (spanning 26°) and a short (spanning 13°) section. The ESQ plates are charged before each beam injection, remain powered for about 700 μs after beam injection, and get discharged after the fill. Pulsing is required to ensure a stable operation voltage. Muons can be stored for up to ten times the muon lab-frame lifetime. The pulsing of the ESQ plates results in resonant mechanical vibrations that cause magnetic field perturbations synchronous to the muon injection that have been measured to determine a correction to the muon-averaged magnetic field.

A set of four fiber-detector arrays (harps) positioned around the ring monitors the beam profile and motion directly in the storage region. The fiber harps comprise horizontal and vertical planes of scintillating fibers that destructively measure the stored muons and can be inserted for dedicated systematic runs. Fiber-harp data are used to measure the beam momentum distribution, the cyclotron frequency, and the debunching of the muon beam during a fill.

The magnetic field is determined by mapping within the storage volume and tracking during muon storage and data taking. Mapping is accomplished with a trolley consisting of 17 NMR probes housed in a movable aluminum shell that is pulled through the storage ring on rails. It measures with centimeter-scale spacing in both azimuthal and transverse directions. A high-purity calibrated water NMR probe, mounted on a 3D movable arm [16], calibrated the trolley probes in the storage ring vacuum before Run-2 and after Run-3. The trolley is removed from the storage volume during data taking, and an array of 378 NMR probes, called fixed probes, help track the field. The fixed probes are located in grooves on the outer surfaces of the vacuum chambers above and below the storage volume. While the trolley is mapping the field, fixed probe measurements and trolley measurements are synchronized. The entire chain of NMR measurements is calibrated to provide the precession frequency of shielded protons in a spherical water sample at 34.7 °C.

The positrons from stored positive muon decays are detected in 24 calorimeter stations located equidistantly around the interior arc of the storage ring vacuum chamber. These calorimeters use lead fluoride (PbF_2) crystals as Cherenkov radiators from which signals are read out via SiPMs [17–19]. Each calorimeter consists of a 6×9 ($H \times W$) array of PbF_2 crystals. Each crystal block is 14 cm (15 radiation lengths) long with a 2.5 cm square cross section. In addition to the excellent spatial resolution produced by crystal segmentation, the calorimeters provide sub-ns timing resolution to distinguish individual positron events. A laser-based gain monitoring system [20] is employed to continuously measure the calorimeter response to obtain energy measurements that are stable with respect to the hit rate and the environmental conditions.

An in-vacuum tracking system based on straw trackers [21] is installed at two locations around the storage ring just upstream of a calorimeter to track muon decay electrons headed for the calorimeters. The trackers are used to monitor the beam distribution ($M^T(x, y, t)$) in the storage ring in the proximity of the two tracking stations. These stations are composed of 32 planes of straw-tube detectors assembled into eight modules. The straw tubes are filled with Argon-Ethane gas, and a thin tungsten wire positioned along the central axis of each straw collects the drift electrons arising from the ionization induced by a passing positron. Tracks are reconstructed by registering hits across multiple planes, and the track reconstruction facilitates both a measurement of the positron momentum and extrapolation to the muon decay vertex.

B. Simulation packages

A suite of different simulation packages was developed to validate analysis tools. Simulation results from the three compact packages are cross-checked against each other. Each package's toolkit provides unique properties, which lead to specific advantages or shortcomings depending on the analysis. For example, GM2RINGSIM models with high fidelity the material interactions that determine the properties of the stored beam, whereas symplectic tracking for long-term beam effects is verified with the COSY-INFINITY and BMAD models. Below, we describe the main characteristics of each simulation package. For comparisons of the simulation packages, please refer to Ref. [8].

GM2RINGSIM is a model of the $g-2$ injection line and storage ring that has been implemented in the GEANT4 simulation framework [22–24]. The model consists of a full description of the material structures, as well as the particle detectors that reconstruct the kinematics of the muons and decay positrons [8]. The GM2RINGSIM package includes several particle guns, one that allows for high-fidelity production of decay positrons within the ring and one that allows for muon production, propagation, and decay through the full injection channel. Runge-Kutta integration methods are used to numerically integrate a particle's equation of motion and propagate it through electromagnetic fields and across detector boundaries. The parallel world functionality is used to insert "virtual" tracking planes into the ring, without adding any material. These planes allow for the reconstruction of the motion of the injected particles as they circulate within the ring. The nonsymplectic nature of GEANT4 did not cause any issues for the systematic errors presented.

The COSY-based model [25] is a data-driven computational representation of the storage ring in COSY INFINITY [26]. The magnetic field in the storage volume is an implementation of the azimuthally dependent set of multipole strengths from the experimental data, described as a series of magnetic multipole lattice elements. An optical element superimposed on the magnetic field recreates the ESQ stations. The high-order coefficients of the electrostatic potential's transverse Taylor expansion produce the nonlinear action of the ESQ on the

beam's motion. A recursive iteration of the horizontal midplane coefficients, modeled with conformal mapping methods to satisfy Laplace's equation in curvilinear optical coordinates, provides these coefficients. The boundary element method is utilized in COULOMB's field solver to recreate the ESQ's effective field boundary and fringe fields in the model. The COSY-based model calculates lattice configurations, Twiss parameters, betatron tunes, closed orbits, and dispersion functions of the storage ring.

A third model based on BMAD [27] models the injection line and storage ring, which are arranged as a series of guide field elements referred to as the lattice. The electromagnetic fields of the elements are represented as field maps, or multipole expansions. Particles are tracked by Runge-Kutta or symplectic integration of the equations of motion as required. Muon spin is likewise propagated by numerical integration. Multiple scattering is included at the entrance and exit windows of the inflector and the outer ESQ plate through which particles are injected into the ring. Otherwise, element boundaries are considered apertures, and particles incident on those boundaries are lost. Calorimeters and trackers are represented as simple markers that indicate particle phase space coordinates. BMAD library routines are used to compute beam parameters like beta-functions, chromaticity, dispersion, emittance, etc.

III. DATASETS AND RUN CONDITIONS

A. Datasets

Run-2 and Run-3 data were acquired from March to July 2019 and November 2019 to March 2020, respectively. The data are divided into 9 and 13 data subsets labeled 2A–2I and 3A–3O for Run-2 and Run-3, respectively. Four data subsets (2A, 2I, 3A, and 3H) were excluded from the measurement analysis because systematic studies dominated the periods. The improved stability of the hardware conditions with respect to Run-1 allowed multiple datasets to be combined in the ω_a^m analysis to leverage the higher statistics and minimize the statistical uncertainties of some systematic effects. The smaller data partitions are combined into the following datasets: Run-2 = [2B–2H], Run-3a = [3B–3G, 3I–3M], and Run-3b = [3N–3O]. The three datasets have different beam storage characteristics, ESQ voltage, and kicker strength. The data were hardware-blinded by hiding the true value of the calorimeter digitization clock frequency. This blinding factor was different for Run-2 and Run-3. In Run-2, we performed 25 trolley runs and tracked 17 field periods, and in Run-3, we performed 44 trolley runs and tracked 34 field periods. In each case, only two field periods did not receive a terminal trolley run.

Muon-decay positrons included in the final datasets are selected according to data quality cuts (DQC) based on the quality of fills and magnetic field stability. Selection criteria for good fills include the kick amplitude and timing, beam profiles, and presence of laser synchronization pulses. DQC are based on the average rate of lost muons, the number of positrons detected, and the quality of the magnetic field and monitor data. DQC selection criteria are chosen so that the muon storage conditions are uniform across each of the combined datasets. Overall, roughly 20% of the time periods have been discarded, most of them containing zero or few positron events, which corresponds to ∼2% of the total data. The detector and magnetic field DAQ systems are separate and not synchronized, resulting in short periods between field DAQ runs where the precession data would not have corresponding field data. Elimination of those time periods reduces the precession data by ∼0.3%. Magnetic field quality criteria excluded muon data collected from occasional sudden changes of the magnetic field, probably due to magnet component movement, large field oscillations with a period around two minutes related to variations of the superconducting coils' cryogenics, and rare spikes related to the NMR probes used in the magnetic-field stabilization system. Figure 1 shows the accumulated positrons for Run-2 and Run-3 after DQC. In total, 71×10^9 positrons with an energy above 1 GeV were accumulated.

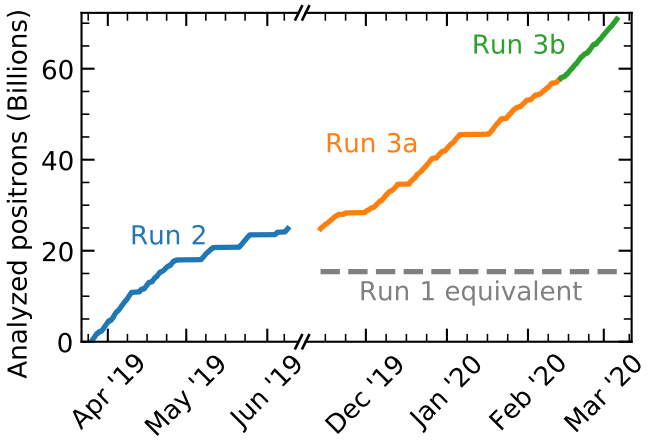

FIG. 1. Muon-decay positrons accumulated in Run-2 and Run-3 after DQC. Positrons with 1 GeV $< E <$ 3 GeV hitting the calorimeters $t > 30$ μs after injection are shown. The Run-1 equivalent (15.4×10^9) is shown for comparison.

B. Run conditions: Run-2/3 vs Run-1

Table I presents the number of fills and reconstructed positrons with energies between 1 and 3 GeV along with the field indices and kicker strengths for the Run-1 and Run-2/3 datasets.

Significant improvements and changes for Run-2/3 with respect to Run-1 [8] include the following:

(1) During Run-1, two resistors electrically connected to the upper and lower plates of the long section of the first ESQ after injection (Q1L) were damaged. Replacing the resistors after Run-1 improved the stability of radial and vertical beam positions. This significantly reduces the phase acceptance correction in Run-2/3.

TABLE I. Dataset statistics and hardware conditions for Run-2/3 compared to Run-1. The number of analyzed positrons (e^+) represents the statistics used in the final ω_a^m fits.

Dataset	Fills ($\times 10^6$)	e^+ ($\times 10^9$)	Field index	Kicker (kV)
Run-1a	1.51	2.0	0.108	130
Run-1b	1.96	2.8	0.120	137
Run-1c	3.33	4.3	0.120	130
Run-1d	7.33	6.3	0.107	125
Run-2	18.60	24.7	0.108	142
Run-3a	33.53	33.1	0.107	142
Run-3b	11.55	11.9	0.108	161

(2) Before Run-3a the operational high-voltage set points for the ESQ system were lowered by 0.1 kV to avoid betatron resonances for beam stability. This shift reduced the muon losses by roughly 20%.
(3) While in Run-1 only two collimators were used, all five collimators were used in Run-2/3, which led to better beam scraping and further reduced the effect of muon losses during storage.
(4) The kicker strengths for Run-1 and Run-2 were limited to 142 kV by the use of A5596 cables [28]. As a result, the beam was not perfectly centered in the storage region. At the end of Run-3a, the cables were upgraded [29] and the kicker voltage was increased to 161 kV in Run-3b to achieve a more optimal kick. This results in a better-centered muon beam, reducing the E-field correction [15].
(5) Between Run-1 and Run-2, the magnet yokes were covered with a thermal insulating blanket to mitigate day-night field oscillations due to temperature drifts. In addition, the experimental hall's air conditioning system was upgraded after Run-2 to further stabilize the temperature of both the magnet yokes and the detector electronics to better than ±0.5 °C. Figure 2 shows the stability improvement for both the magnet and the calorimeter SiPMs since Run-1.
(6) In Run-2/3, the magnetic field hardware operation procedures improved compared to Run-1. The more standardized and automated procedures, especially for trolley runs, made measurements and monitoring of the magnetic field faster and more reliable. In addition, the magnet power supply feedback loop was optimized during Run-2 to suppress oscillations in the magnetic field more efficiently and better decouple from higher-order moment changes.
(7) For Run-2/3, modifications were made to the real-time processing of the digitized waveforms from the calorimeter crystals that are utilized in the positron-based analyses. In Run-1, when an individual crystal exceeded a preset threshold, the digitized waveforms of all 54 crystals of the associated calorimeter were recorded (see Ref. [6] for details). In Run-2/3, when an individual crystal exceeded a preset threshold, only the above-threshold crystals and their neighboring crystals were recorded. This change permitted data collection of positron-based data at higher rates.
(8) For Run-2/3, modifications were also made to the real-time processing of the digitized waveforms from the calorimeter crystals that are utilized in the energy-based analyses. In Run-1, the raw Analog to Digital Converter (ADC) samples from each calorimeter crystal were summed into 75 ns-binned histograms. These per-crystal histograms were then stored for each fill (see Ref. [6] for details). In Run-2/3, the raw ADC samples from each calorimeter crystal were summed into 18.5 ns binned histograms. These per-crystal histograms was then accumulated for 4 fills and stored for every fourth fill. These changes permitted the acquisition of energy-based data with a finer time binning and a greater time range.
(9) During Run-2 (i.e., after dataset 2E), a wedge absorber for muon momentum-spread reduction was installed in the incident muon beamline [30].

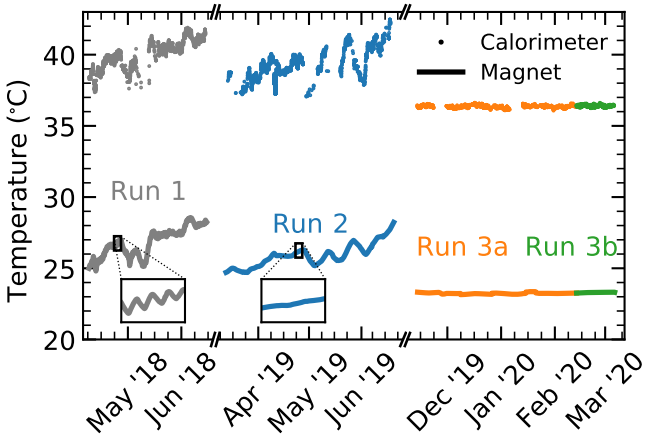

FIG. 2. Temperature of the calorimeter SiPMs (small dots) and the magnet yokes (thicker lines) across Run-1, Run-2, and Run-3. The two inserts show a box of four days with a temperature range of 1 °C. The magnet thermal insulating blanket installed after Run-1 reduced the day-night oscillations of the magnet temperature. The upgraded air conditioning system greatly improved the long-term stability of both the calorimeters and magnet temperature after Run-2.

C. Beam storage conditions

Many of the changes listed in the last chapter define the beam dynamics conditions in the storage ring. The main characteristics, such as typical beam oscillation frequencies, muon losses, and beam distributions, are described in the following subsections.

1. Beam oscillation frequencies

The 120-ns duration of muon injection causes a modulation of positron hits in individual detectors with a

TABLE II. Compilation of frequencies and periods of important beam oscillations for the field index $n = 0.108$ (the anomalous precession frequency fa and cyclotron frequency f_c are given for comparison). Columns 1 and 2 denote the frequency and its symbol. Column 3 gives the relation of the beam frequency to the field index n, cyclotron frequency f_c, and betatron frequencies f_x, f_y, in the continuous ESQ approximation. Columns 4 and 5 list the numerical values of the frequencies and periods for a field index $n = 0.108$ in the continuous ESQ approximation. Note that the measured frequencies differ slightly from the continuous ESQ approximation frequencies.

Term	Symbol	Field index relation	Frequency (MHz) $n = 0.108$	Period (μs) $n = 0.108$
$g-2$	f_a		0.229	4.37
Cyclotron	f_c		6.70	0.149
Horizontal betatron	f_x	$\sqrt{1-n}f_c$	6.33	0.158
Vertical betatron	f_y	$\sqrt{n}f_c$	2.20	0.454
Coherent betatron	f_{CBO}	$f_c - f_x$	0.372	2.69
Vertical waist	f_{VW}	$f_c - 2f_y$	2.30	0.435

cyclotron period T_c. Due to the momentum spread of the stored muons with $p = m_\mu c/\sqrt{a_\mu} \pm 0.15\%$, this initial bunching is gradually debunched [6].

The muons stored in the ring follow both radial and vertical betatron oscillations with frequencies (f_x, f_y) determined by the configuration of the guide fields, characterizing the transverse motion along the azimuth of the ring. In addition, the beam widths (frequencies $2f_x, 2f_y$) and centroids of the stored muons follow the optical lattice (with azimuthal variations smaller than 3%) and closed orbits.

The observed time distribution in a detector is perturbed by these beam oscillations through their coupling to the detector acceptance. In practice, the radial centroid oscillation (f_x) dominates the radial perturbations, and the vertical width oscillation ($2f_y$) dominates the vertical perturbation.

Since muons pass each detector once every cyclotron period, the radial centroid oscillation is observed at an aliased frequency, dubbed coherent betatron oscillation (CBO), $f_{\mathrm{CBO}} = f_c - f_x$. A substantial cancellation of cyclotron period modulation, called fast rotation, is achieved by histogramming data with bin widths as close as achievable to the cyclotron period. Such a binning causes any frequency that exceeds the Nyquist limit $f_c/2$ to also be aliased. The vertical width oscillation appears in the histogram aliased to $f_{\mathrm{VW}} = f_c - 2f_y$. Table II is a summary of these frequencies for the field index (see Ref. [31]) $n = 0.108$.

2. Muon losses

Not all stored muons decay into positrons. Some muons impact material in the storage region, such as aperture-defining collimators, and lose energy to the point where they can no longer be stored. These muons spiral inward, and a subset of them are observed as triple-coincidences of minimum ionizing particles in adjacent calorimeters. The muon loss spectra differ greatly between runs as seen in Fig. 3. The muon loss rate was reduced by an order of magnitude between Run-1 and Run-2 due to the repair of the damaged ESQ resistors. The bump structure (see Sec. IV E 3) observed in Run-2 between 50 and 150 μs was suppressed in Run-3 by better centering the vertical beam.

The presence of lost muons can bias the extraction of ω_a^m in two ways. First, a time-dependent loss of stored muons causes a time-dependent distortion of measured positrons. To avoid biasing the ω_a extraction, the fit must therefore incorporate the effects of muon losses (see Sec. IV E 3). Second, coupling between the muon's momenta and initial spin directions can alter the measured value of ω_a^m, as described with more details in Sec. V C.

3. Beam distributions

The muon beam distribution $M(x, y, \phi)$ is reconstructed by extrapolating beam profiles measured by the two tracker stations. The extrapolation shifts the mean and scales the transverse width of the distributions relative to the tracker

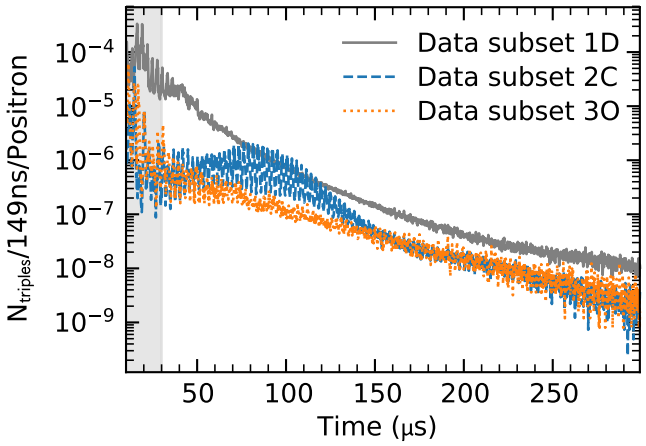

FIG. 3. Muon loss time distribution L(t) for selected Run-1 (gray), Run-2 (blue), and Run-3 (orange) data subsets showing the reduction in losses. The values here are normalized to the number of $e^+ > 1.7$ GeV in each dataset. The large modulation of the muon losses with the frequency f_{CBO} is a reflection of the mechanism of the losses.

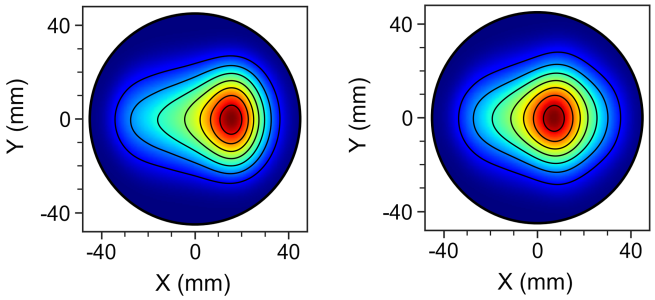

FIG. 4. Azimuthally averaged muon beam distribution summed over $t > 30$ μs ($\langle M(x,y)\rangle_\phi$) from datasets from Run-2 (2B) on the left and Run-3b (3O) on the right. The color represents the intensity, from low intensity in blue (outside) to high intensity in red (inside).

station using characteristic functions obtained from the optical lattice calculated with the COSY INFINITY-based model of the storage ring.

Figure 4 shows azimuthally averaged muon beam distributions based on this beam extrapolation. The increased kick strength in Run-3b moves the beam distribution closer to the center.

IV. MUON ANOMALOUS PRECESSION FREQUENCY MEASUREMENT

This section discusses the analysis of the muon anomalous precession frequency, ω_a^m. It describes the time-distribution reconstructions of positron hits and integrated energy as well as the corrections and the fits that are applied to these distributions. It also discusses the ω_a^m results, systematic uncertainties and consistency checks. We emphasize changes since the Run-1, ω_a^m analysis [6].

The ω_a^m analysis was conducted by seven independent analysis groups using a number of different strategies for the positron hit and integrated-energy reconstruction, handling of cyclotron rotation and positron pileup, and treatment of beam dynamics and muon losses. Herein the analysis groups are denoted by roman numerals I–VII.

A. Analysis methods

The measurement benefits from multiple complementary analysis techniques that can be divided broadly into two categories. The first category is event-based and focuses on reconstructing the energies and times of the individual decay positrons in the calorimeters. The second category is energy-based and focuses on reconstructing the energy versus time in the calorimeters without the positron identification. For each technique, we construct a time distribution that is modulated by the anomalous precession frequency ω_a^m.

In the event-based methods, we applied two data-weighting schemes. In the threshold analysis (denoted the T method), equal weight is given to all positrons above a fixed energy threshold. In the asymmetry-weighted analysis (denoted the A method), each positron is weighted according to the decay-asymmetry corresponding to the positron's energy (see Fig. 5). The asymmetry-weighted analysis achieves the greatest possible statistical power to measure the precession frequency. The integrated-energy approach (denoted the Q method), is logically equivalent to weighting positrons with their energies even though it does not resolve individual positrons.

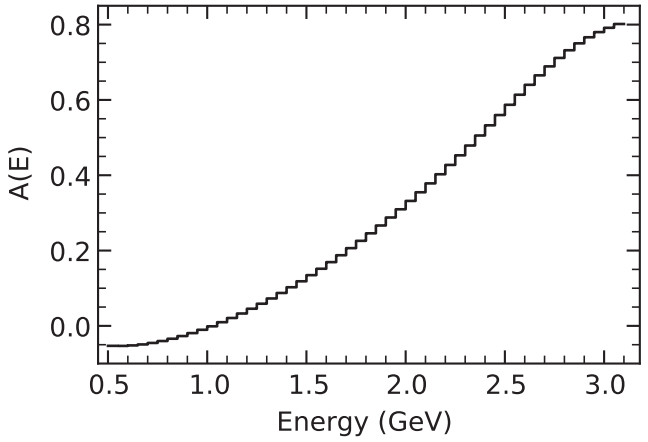

FIG. 5. Representative example of the measured asymmetry $A(E)$ of the anomalous precession signal versus the positron energy E in the region 0.5–3.1 GeV (for the calorimeter summed data and a selected analysis group). In the A method, each positron is weighted by $A(E)$ to achieve the greatest possible statistical power in the anomalous precession frequency measurement. Note the measured asymmetry $A(E)$ incorporates detector acceptance effects.

In a ratio method, the data are split into four subsets, two time-shifted and two unshifted, from which a ratio histogram is constructed. By using time shifts of one-half the anomalous precession period, the ω_a^m modulation is preserved while slow-time variations are mitigated. See Ref. [6] for the details of the construction of the ratio histogram.

B. Reconstruction approaches

For the event-based analyses, we used two distinct reconstruction schemes: a local-fitting approach and a global-fitting approach. The local-fitting approach was used by the groups I through IV and the global-fitting approach was used by groups V and VI. An important difference between these two approaches was the inclusion or exclusion of spatial separation of positron hits in the fitting procedure (see Ref. [6] for details).

The local-fitting approach involves individually fitting the waveform from each crystal. Each crystal waveform is first fit to an empirically determined pulse template to determine its time and energy. The crystal hits occurring in a given time window are then clustered into positron candidates. The cluster time was defined as the time of the crystal hit with the largest energy, and the cluster energy was defined as the sum of the clustered crystal energies.

The global-fitting approach involves simultaneously fitting the waveforms from 3×3 crystal arrays that are centered on the highest-energy crystal. The 3×3 waveforms are simultaneously fit to empirically determined pulse templates to determine a single shared fitted time and individual crystal energies (see Ref. [6] for the details of the construction of the templates). The cluster time was defined as the single shared fitted time and the cluster energy as the sum of the contributing crystal energies.

The group VII, energy-based reconstruction involves the construction of a time distribution of the deposited energy in each calorimeter. The approach utilizes a rolling pedestal with a low-energy threshold in order to extract the integrated energy and mitigate any pedestal variations (see Ref. [6] for details). It negates the need for fitting and clustering of crystal pulses and decision making in positron identification. Although statistically less powerful, its value lies in utilizing different raw data, applying different reconstruction procedures, and inheriting different systematic uncertainties.

C. Data corrections

The analysis methods (Sec. IV A) and reconstruction approaches (Sec. IV B) are used to build time distributions of positrons hits or integrated energy. Before fitting the time distributions to extract ω_a^m we apply several corrections.

One correction applied to the raw data, accounts for any gain changes in the calorimeter electronics. Another correction applied to the time histograms removes the distortions arising from positron pileup. A final correction treats the imprint on the data of the cyclotron rotation of the stored beam. These corrections are described below.

1. Gain corrections

The calorimeter SiPMs and readout electronics suffer from gain fluctuations on multiple timescales from various physical effects. At the longest timescales, temperature variations in the experimental hall lead to gain changes over days or longer (long-term gain correction). Within a muon fill, the initial beam flash causes an immediate gain sag with gradual gain recovery that impacts all calorimeters but especially those near the inflector (in-fill gain correction). At the shortest timescales, the SiPM pixel deadtime causes a short-term gain sag if a second positron is recorded just after an earlier positron (short-term gain correction).

These effects are corrected using dedicated studies with a laser calibration system Ref. [32]. One improvement since Run-1 is the treatment of the temperature dependence of the short-term gain corrections.

Note that the significant improvement in the temperature stability of the experimental hall from Run-2 to Run-3 (see Fig. 2), reduced the size of long-term gain corrections and limited the need for temperature-dependent, short-term gain corrections in Run-3.

2. Pileup corrections

For event-based analyses, it is generally not possible to resolve positron hits in the same calorimeter crystal within a 1.25-ns time interval (we note that the spatial resolution of the global-fitting approach can sometimes identify such pileup events). Consequently, such close-in-time positrons are summed and treated as a single positron with the summed energy of the true positrons. Since the likelihood of positron pileup will decrease during the muon fill, this potentially biases the ω_a extraction.

To account for pileup, the raw time distribution is corrected through a data-driven, statistical reconstruction of a pileup time distribution. Three methods were used in building the pileup distribution: the so-called empirical, semiempirical, and shadow window methods. All three methods model the effects of pileup by computing the difference between the reconstructed energy-time distributions of unresolved positrons and resolved positrons. This pileup time distribution is then subtracted from the raw time distribution.

The pileup modelling is achieved by superimposing data from the same calorimeter with a one cyclotron period delay from the reconstructed positron. This separation randomly samples the calorimeter data with a similar rate. The initial reconstruction provides the individual positrons before the data superposition.

In practice, this superposition of data can be performed at the level of the digitized waveforms, crystal hits, and reconstructed positrons. These levels correspond to the aforementioned empirical, semiempirical, and shadow window methods, respectively [33]. An improvement on Run-1 was the handling of triple pileup in most Run-2/3 analyses.

All three methods show an excellent ability to reproduce the observed pileup energy spectrum in the energy region greater than the 3.1 GeV beam energy. An example using the empirical method is shown in Fig. 6.

The energy-based analyses utilize a nonzero energy threshold and therefore are not completely immune to a pileup distortion. We therefore developed a signal processing algorithm for calculating pedestals and applying thresholds that minimizes pileup effects. The algorithm is described in [6].

3. Fast-rotation handling

Although the fast-rotation modulation (Sec. III C) is greatly reduced by the 30 μs start time of the ω_a fit region, its effect is nonzero. A substantial cancellation of fast rotation is achieved by histogramming data with bin widths as close as possible to the cyclotron period (149.2 ns for the event-based analyses and 150 ns for the energy-based analyses). A further cancellation is achieved by summing the data from the 24 calorimeters (due to the 2π advance of the fast-rotation modulation

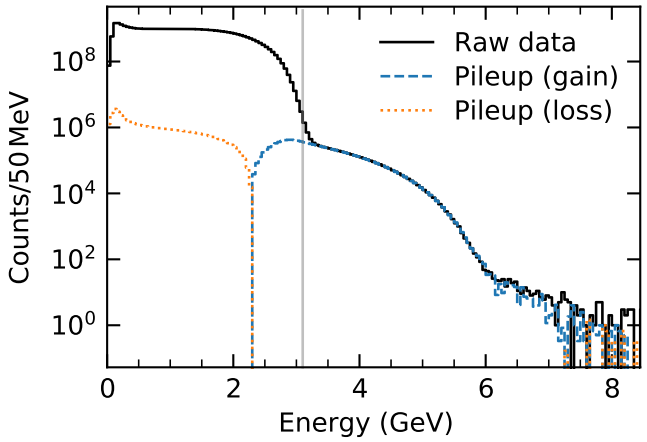

FIG. 6. Illustration of the reconstructed pileup correction for the empirical method. The black curve is the raw energy distribution before the pileup correction. The dashed blue (dotted orange) curves show the reconstructed gain (loss) of positron events due to positron pileup. The agreement between the black curve and the blue curve in the energy region greater than the 3.1 GeV beam energy (vertical gray line) is an indication of the quality of the pileup correction.

around the ring circumference). These procedures were used in all the analyses.

The remaining distortion is handled by either randomizing the histogram entries by one cyclotron period in event-based analyses or uniformly distributing the energy entries over one cyclotron period in energy-based analyses.

D. ω_a^m software blinding procedure

During their analysis processes, each of the seven analysis groups were software-blinded with respect to each other (i.e. in addition to the common hardware blinding).

The procedure parametrized the measured frequency ω_a^m as a fractional shift R from a nominal reference frequency $\omega_{\text{ref}} = 2\pi \times 0.2291$ MHz, where

$$\omega_a^m = \omega_{\text{ref}} \cdot (1 + [R - \Delta R] \times 10^{-6}), \tag{5}$$

and ΔR is that group-dependent, software-blinding offset, which is generated within a ± 24 range. The values of ΔR were derived from group-chosen text phrases whose hash seeded a random number generator.

The relative unblinding of the seven groups to a common software-blinded stage facilitated unbiased comparisons between the analyses and followed internal reviews conducted by the analysis teams. The remaining software and hardware blindings were not removed until the collaboration's decision to publish the result for a_μ.

E. ω_a^m fitting procedure

The measured anomalous precession frequency ω_a^m was extracted by fitting the reconstructed positron or integrated-energy time histograms after correcting for cyclotron rotation and positron pileup. These "ω_a^m-wiggle" fits were performed using either the MINUIT numerical minimization package [34], the PYTHON SciPy.OPTIMIZE package [35], or the PYTHON LMFIT package [36]. They minimized the quantity

$$\chi^2 = \sum_{ij} (y_i - f_i) V_{ij}^{-1} (y_j - f_j), \tag{6}$$

where y_i are the measured data points, f_i are the corresponding fit function values, and V_{ij} is the covariance matrix. The diagonal elements of V_{ij} are the variances σ_i^2 of the data points y_i. The off-diagonal elements of V_{ij} are the covariances σ_{ij}^2 between the data points y_i, y_j. Nonzero covariances were used in some analyses to handle correlations between data points arising from the handling of cyclotron rotation, correction for positron pileup, and construction of ratio histograms. The minimization of χ^2 determines the optimal values of the model parameters of the fit function.

The nominal fit time ranges were 30.1 to 660.0 μs for the event-based analyses and 30.1 to 330.0 μs for the energy-based analyses. The bin widths were 149.2 ns for the event-based analyses and 150.0 ns for the energy-based analyses. The 30.1 μs start time is (i) after the stabilization of beam scraping and (ii) as close as possible to an ω_a anomalous precession node in order to minimize any pull from miscalibration of the calorimeters (see Sec. IV C 1).

1. ω_a^m fit model

The fit function used for extracting ω_a^m from both the event-based and energy-based time distributions has the general form

$$f(t) = N_0 \cdot N_x(t) \cdot N_y(t) \cdot N_{xy}(t) \cdot \Lambda(t) \cdot e^{-t/\gamma\tau_\mu} \\ (1 + A_0 \cdot A_x(t) \cos(\omega_a^m t - (\phi_0 + \phi_x(t)))). \tag{7}$$

The function incorporates the effects of muon decay and anomalous precession through the time-dilated lifetime $\gamma\tau_\mu$, muon decay asymmetry A_0, anomalous precession frequency ω_a^m, and anomalous precession phase ϕ_0. N_0 is an overall normalization. Note that the time-dependent terms N_x, N_y, N_{xy}, A_x, ϕ_x, and Λ are used to handle distortions from beam dynamics and muon losses [37]. These distortions are explained in detail in Secs. IV E 2 and IV E 3, respectively.

In addition, we discuss in Sec. IV E 4 an electronics ringing term that was used in the energy-based analyses and in Sec. IV E 5 a residual slow term that was studied in the event-based analyses.

If N_x, N_y, N_{xy}, and A_x are set to unity and ϕ_x is set to zero in Eq. (7), one obtains a five-parameter function involving N_0, $\gamma\tau_\mu$, A_0, ω_a^m, and ϕ_0. In subsequent sections, we utilize the five-parameter fit residuals and their discrete Fourier transforms to illustrate the effects of beam dynamics.

2. Beam dynamics distortions

In principle, the beam oscillations, in combination with detector acceptances introduced in Sec. III C, perturb the overall normalization (N_0), decay asymmetry (A_0), and precession phase (ϕ_0), in the ω_a^m fit function. In practice, we find the large radial perturbations require accounting for beam distortions to N_0, A_0, and ϕ_0 while the smaller vertical perturbations only require accounting for distortions to N_0.

The time-dependent distortions from beam dynamics were generally modelled by a sinusoidal oscillation with an empirical decoherence envelope. For example, leading effects of CBO perturbations on the normalization N_0 could be modelled by a term

$$N_x(t) = 1 + A_{\mathrm{CBO}} e^{-t/\tau_{\mathrm{CBO}}} \cos(\omega_{\mathrm{CBO}} t + \phi_{\mathrm{CBO}}), \quad (8)$$

where the associated parameters are the CBO amplitude, A_{CBO}, CBO frequency, ω_{CBO}, CBO phase, ϕ_{CBO}, and CBO decoherence time constant τ_{CBO}. Similar functional forms were used for the beam dynamics corrections N_y, N_{xy}, A_x, and ϕ_x. Note that the term $N_{xy}(t)$, with a frequency $\omega_{\mathrm{VW}} - \omega_{\mathrm{CBO}}$, arises from a coupling between the dominant horizontal and vertical oscillations.

In practice, a number of monotonically decreasing functions, which involved combinations of exponential and reciprocal functions, were used for modeling the decoherence envelope. The envelope shape and time constant were found to differ across the three datasets and the event-based and energy-based analyses. The ω_a^m sensitivity to the decoherence envelope is discussed in Sec. IV J 1.

In addition, an effective time variation of the CBO frequency was identified in the time distributions of the individual calorimeters. This effect was modelled through an exponentially decreasing time variation with a 10–20 μs time constant and a fitted amplitude parameter. The ω_a^m sensitivity to the frequency change is discussed in Sec. IV J 1.

3. Muon loss distortions

Muon losses, as described in Sec. III C and shown in Fig. 3, reduce the number of stored muons and, consequently, the number of detected positrons.

As shown, such losses can be measured as a function of time $L(t)$ by muons traversing multiple calorimeters. However, such measurements do not determine the absolute rate of muon losses. An absolute measurement of the muon loss rate would require modeling the calorimeter acceptance of aberrant trajectories to high precision. A data-driven approach was therefore employed.

Note that muon-loss effects on positron rates at time t are determined by the integrated losses up to time t. All ω_a^m fits therefore incorporate a muon loss term

$$\Lambda(t) = 1 - k_{\mathrm{loss}} \int_0^t L(t) e^{t'/\gamma\tau_\mu} dt', \quad (9)$$

where $L(t)$ is the measured muon-loss time distribution and k_{loss} is a fitted normalization parameter.

Figure 3 in Sec. III A compares the measured time distributions $L(t)$ for the different datasets. The changes made to the quadrupole and kicker settings between the three datasets led to related changes in the loss rates and the time distributions. In Run-2 the loss rates were significantly larger as the field index was closer to beam resonances.

Another notable difference between the datasets was the appearance of a bump in the Run-2 time distribution. The bump amplitude and bump time both varied around the storage ring and changed during Run-2 operations. Although the bump's cause is not fully understood, it was found to be correlated with the magnet temperature and the vertical beam position.

Due to the Run-2/3 differences in muon-loss time distributions, the procedures for fitting the losses differed between Run-2 and Run-3. These details are summarized in Table III.

4. Electronics ringing distortions

In the energy-based approach, the time distributions are incremented with above-threshold, pedestal-subtracted energies. The pedestal is calculated from the rolling average of the ADC samples in a window surrounding each above-threshold, ADC sample. Consequently, both drifts and oscillations of the baseline during the fill can bias this calculation.

The largest bias arose from electronics ringing with a period of about 600 ns that resulted from the injection flash in the calorimeters. To determine the effect on calculating the pedestal, we computed the distribution of differences between

(1) ADC samples without above-threshold signals, and
(2) corresponding pedestal estimates from the surrounding pedestal samples.

This data-driven bias was then incorporated in the fit function for the energy-based analyses in a similar manner to the muon loss term.

5. Residual slow effect

Residual slow effects, a change in positron counts or integrated energy over the duration of the fill, have different sources.

One contribution arose in the local-fitting analysis from the handling of the single chopped islands with more than

TABLE III. Summary of the fitting strategies of the seven analysis groups I–VII. Columns 1, 2, and 3 denote the groups, reconstruction and histogramming methods. Column 4 lists the total number of parameters varied in the fits to the datasets. Column 5 lists the strategy for handling the time-dilated muon lifetime. Columns 6 and 7 summarize the strategies for handling the muon-loss term in runs 2 and 3, respectively. The +, − denotes the sign of the muon-loss term in the wiggle fit (see Sec. IV J 3). Columns 8–10 summarize the strategies for handling the various beam dynamics effects where the heading $f_{\mathrm{CBO}}(\mathrm{t})$ denotes a time-dependent CBO frequency, the heading $e^{-t/\tau_{\mathrm{CBO}}}+C$ denotes a CBO envelope with both an exponential and constant term, and the heading VW − CBO denotes the 1.9 MHz oscillation term. An unlabeled check mark indicates the associated fit term was included in all datasets. A check mark with label "r3" or "r3b" indicates the associated fit term was included in the Run-3 or Run-3b datasets only. Note in column 7, "fixed τ_d" indicates the time constant of the CBO frequency change was not varied in the fit. See text for details.

Group	Recon	Method	Number of free parameters 2, 3a/3b	τ_μ handling	Run-2 k_{loss}	Run-3 k_{loss}	$f_{\mathrm{CBO}}(\mathrm{t})$ term	CBO env. $e^{-t/\tau}+C$	VW − CBO term
I	Local	A, T	28/28	Free	Free, +	Free, −		r3b ✓	✓
II	Local	A, T	25/26	Free	Free, +	Fix, 0	✓	r3b ✓	✓
III	Local	A, T	28/28	Free	Free, +	Free, −	✓, fixed τ_d	✓	✓
III	Local	AR, TR	14/14	Fix	Free, +	Free, −	✓, fixed τ_d	✓	
IV	Local	A, T	18/18	Free	Free, +	Fix, 0	✓, fixed τ_d	✓	
IV	Local	AR, TR	15/15	Fix	Fix, +	Fix, 0	✓, fixed τ_d	✓	
V	Global	A, T	30/30	Free	Free, +	Free, −	✓	✓	✓
V	Global	TR	19/19	Fix	Fix, +	Fix, −	✓	✓	
VI	Global	A, T	27/28	Penalize	Free, +	Free −	✓, fixed τ_d	r3b ✓	✓
VII	Energy	Q	34/38	Free	Free, +	Free, −	✓	✓	r3 ✓
VII	Energy	QR	26/24	Fix	Fix, +	Fix, −	✓	✓	r3 ✓

one positron cluster. Such islands—that are more probable at early times in the fill—produced a time-dependent, energy-scale shift.

Another contribution stems from a remaining residual slow term that is common to both local and global fits. Possible sources of this effect include changes in gain, acceptance, or reconstruction over the duration of the fill. The introduction of either an *ad hoc*, time-dependent correction term or an *ad hoc*, time-dependent fit term is utilized to mitigate this residual effect. We noted that this term's magnitude is highly correlated with analysis strategies that are applied to the fitting of other slow terms like the muon lifetime and the muon losses. We chose not to apply the *ad hoc*, time-dependent fit term in the extraction of the frequency ω_a^m.

F. Differences with respect to Run-1

The major differences between the Run-2/3 analysis and the Run-1 analysis are listed below.

(1) In Run-2/3 we introduced a so-called kernel method for building ratio histograms. This method uses four identical copies of the time distributions for the ratio construction. It has the advantage of avoiding the statistical noise originating from the Run-1 randomization approach. It has the disadvantage of introducing bin-to-bin correlations in the ratio histograms.

(2) In Run-2/3 the ratio construction was additionally applied to the asymmetry-weighted positron time distributions and the integrated-energy time distributions. Below we denote the original T-method ratio histograms by RT, the new A-method ratio histograms by RA, and the new Q-method ratio histograms by QR.

(3) In Run-2/3 we introduced several improvements in the local-fitting positron reconstruction. One improvement used the measured energy dependence of the SiPM time resolution [18]. It improved the separation of close-in-time clusters and reduced the positron pileup. Another improvement by group I involved prioritizing the crystal hits with higher energies during clustering. It improved the positron time resolution.

(4) In Run-2/3 we improved the gain correction procedure by incorporating a temperature-dependent, short-term gain correction.

(5) In Run-2/3 a new frequency corresponding to $\omega_{\mathrm{VW}}-\omega_{\mathrm{CBO}}$ was identified in the time distributions and incorporated in the ω_a fits.

G. Multiparameter fits

Table III summarizes the analysis strategies and fitting choices that were made by the seven groups in their multiparameter ω_a^m fits. The discrete Fourier transform of the fit residuals for a representative multiparameter fit to the Run-3b dataset is shown in Fig. 7.

As discussed in detail in Sec. IV B, the analyses span three distinct reconstructions: the event-based, global-fitting reconstruction, the event-based, local-fitting reconstruction, and the energy-based reconstruction. Positron pileup was corrected by three distinct, data-driven approaches involving superimposing ADC waveforms,

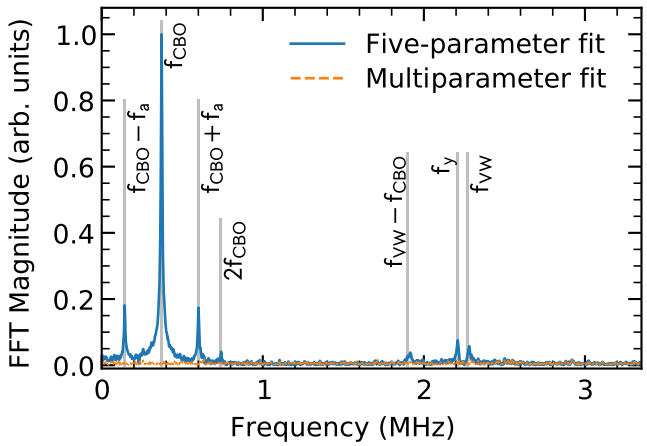

FIG. 7. Representative example of the discrete Fourier transform (FFT) of the fit residuals for a five-parameter fit (solid blue) and a multiparameter fit (dotted orange) to the Run-3b dataset. The five-parameter Fourier transform indicates the presence of perturbations due to beam dynamics, muon losses, *etc*. The five-parameter fit shows peaks corresponding to radial beam oscillations (f_{CBO}, $2f_{\mathrm{CBO}}$), vertical beam oscillations (f_{VW}, f_y), couplings between precession and radial frequencies ($f_{\mathrm{CBO}} \pm f_a$), and radial and vertical frequencies ($f_{\mathrm{VW}} - f_{\mathrm{CBO}}$). Also evident at low frequencies are the effects of muon losses and other slow effects.

crystal hits, or positron hits (see Secs. IV C 2 and IV C 3 for details). The handling of cyclotron rotation involved either randomizing the histogram entries by times $\pm T_c/2$ in event-based analyses or uniformly distributing the histogram entries over times $\pm T_c/2$ in energy-based analyses. The time distributions themselves were constructed with equally weighted positron entries (T method), asymmetry-weighted positron entries (A method), and energy-weighted entries (Q method). Ratio histograms for each weighting were also constructed (TR, AR, and QR methods).

In performing the fits, independent analysis groups used different strategies for handling perturbations from beam dynamics, muon losses, and residual slow effects. Choices included the use of free, penalized, and fixed values for the time-dilated muon lifetime $\gamma\tau_\mu$ [38]; the use of free, fixed, or zero values for the muon loss parameter k_{loss}; and different handlings of the CBO envelope shape and the CBO frequency time-dependence. The total number of free parameters varied with analysis choices and histogramming methods and ranged from 14 parameters (in one AR method fit) to 38 parameters (in the Q method fit).

Note that two analysis groups (III and IV) used a randomization procedure similar to fast rotation randomization to handle the VW beam oscillation. This avoided the need for an associated fit term and reduced the number of fit parameters.

The typical effects the aforementioned corrections have on the extraction of ω_a^m are $\mathcal{O}(1000$ ppb) for the beam dynamics, $\mathcal{O}(10$ ppb) for the muon losses, $\mathcal{O}(100$ ppb) for the positron pileup, and $\mathcal{O}(1$ ppb) for the cyclotron rotation.

TABLE IV. R values in units of ppm for the 19 distinct analyses of the three datasets. Note the muon-weighted magnetic field (Sec. VI F) and beam dynamics corrections (Sec. V) are different for the three datasets. Column 1 denotes the analysis group and column 2 denotes the histogramming method. The remaining columns give the commonly blinded R values and their statistical uncertainties for the Run-2, Run-3a, and Run-3b datasets, respectively. See text for the discussion of the allowed statistical differences between the different analyses.

		Run-2		Run-3a		Run-3b	
Group	Method	R	σ_R	R	σ_R	R	σ_R
I	T	−99.112	0.377	−98.682	0.320	−97.298	0.520
II	T	−99.171	0.376	−98.700	0.323	−97.274	0.519
III	T	−99.198	0.377	−98.690	0.323	−97.267	0.520
IV	T	−99.147	0.382	−98.726	0.329	−97.304	0.528
V	T	−99.029	0.378	−98.603	0.325	−97.191	0.513
VI	T	−99.047	0.378	−98.581	0.325	−97.145	0.522
I	A	−99.197	0.339	−98.355	0.290	−97.453	0.468
II	A	−99.232	0.338	−98.408	0.290	−97.407	0.467
III	A	−99.253	0.337	−98.416	0.291	−97.422	0.468
IV	A	−99.199	0.344	−98.430	0.295	−97.438	0.476
V	A	−99.134	0.340	−98.416	0.291	−97.337	0.466
VI	A	−99.157	0.340	−98.397	0.293	−97.316	0.470
III	RT	−99.189	0.383	−98.693	0.334	−97.279	0.533
IV	RT	−99.160	0.383	−98.710	0.329	−97.244	0.529
V	RT	−99.006	0.384	−98.549	0.325	−97.158	0.513
III	RA	−99.222	0.345	−98.458	0.301	−97.402	0.480
IV	RA	−99.180	0.345	−98.432	0.297	−97.372	0.477
VII	Q	−99.191	0.543	−98.555	0.414	−96.875	0.663
VII	RQ	−99.300	0.491	−98.638	0.386	−97.239	0.616

H. Commonly blinded ω_a^m results

Table IV and Fig. 8 list the commonly blinded ω_a^m values and their statistical uncertainties for 19 distinct analyses covering the Run-2, Run-3a, and Run-3b datasets (the 19 distinct analyses arise from the multiple histogramming techniques applied by the seven analysis groups). The results are expressed in terms of R[ppm] as defined by Eq. (5) and described in Sec. IV D. Across the datasets, the R values may differ due to dataset differences in the muon-averaged magnetic field (Sec. VI F) and ω_a^m beam dynamics corrections (Sec. V).

Within a given dataset the R values from different analyses are highly correlated. The R values should agree within allowed statistical and systematic variations that account for the analysis-to-analysis correlations.

Various sources contribute to the allowed statistical variations between the different analysis approaches. These sources of statistical variations include the following:

(1) differences between event-based and energy-based reconstructions arise from different energy thresholds on crystal pulses and positron candidates,

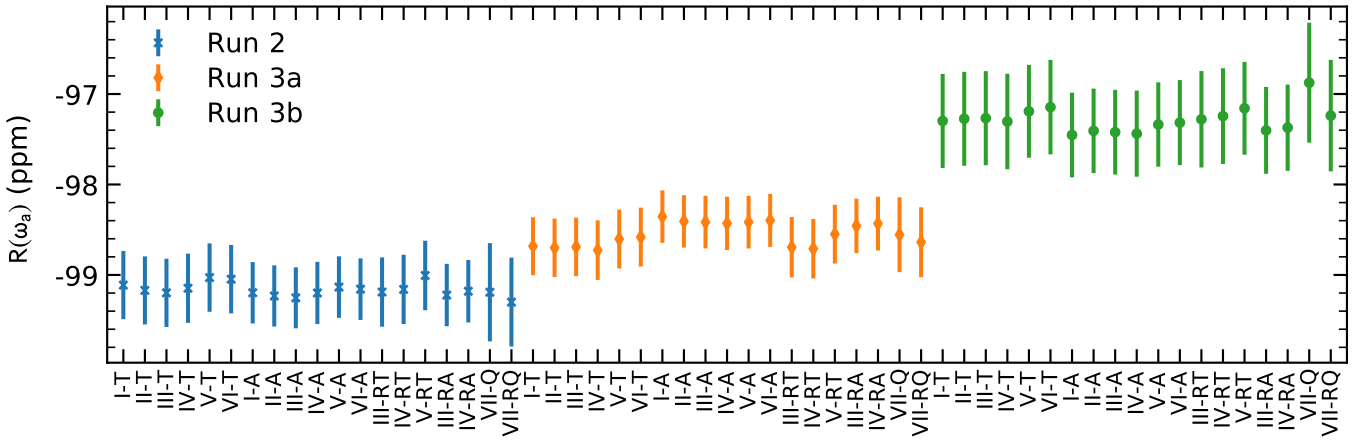

FIG. 8. Plot of the results for the 19 analyses of the three different datasets. Note the muon-weighted magnetic field (Sec. VI F) and beam dynamics corrections (Sec. V) are different for the three datasets. The plotted uncertainties are the statistical uncertainties from the multiparameter fits to the associated time distributions. The allowed statistical and systematic differences between the results for a given dataset are discussed in Sec. IV H.

(2) differences between local-fitting and global-fitting reconstructions arise from different clustering of crystal hits into positron candidates,
(3) differences between T-method and A-method histogramming arise from different thresholds and different weightings of positron candidates,
(4) differences between ratio and nonratio histogramming arise from the ratio-method time shifts and thereby differing data at the beginning and the end of the fit region.

Differing strategies for correcting for positron pileup, handling of beam dynamics, and compensating for muon losses, also introduce allowed differences in the systematic uncertainties for the different analyses. Analysis groups also use different strategies in handling slow effects.

One approach to estimating the analysis-to-analysis correlations uses a Monte Carlo to generate positron candidates and build time distributions. The statistical correlation coefficients between various approaches are then determined by running many Monte Carlo trials, generating many time distributions, and extracting ω_a^m variances between different pairs of analysis approaches.

Another approach to estimating the analysis-to-analysis correlations involves resampling of Run-2/3 data into multiple subsets. These subsets are then separately analyzed using the different analysis approaches. The statistical correlation coefficients between pairs of analyses approaches are then extracted from the measured variances of the ω_a^m differences for the resampled subsets.

In Table XXVII in the Appendix, we list the estimated correlations between all 19 analyses. The largest allowed differences are between event-based analyses and energy-based analyses. The analyses that employ either a common reconstruction approach or a common histogramming approach (the group of six A-method analyses or the group of six T-method analyses) only allow much smaller

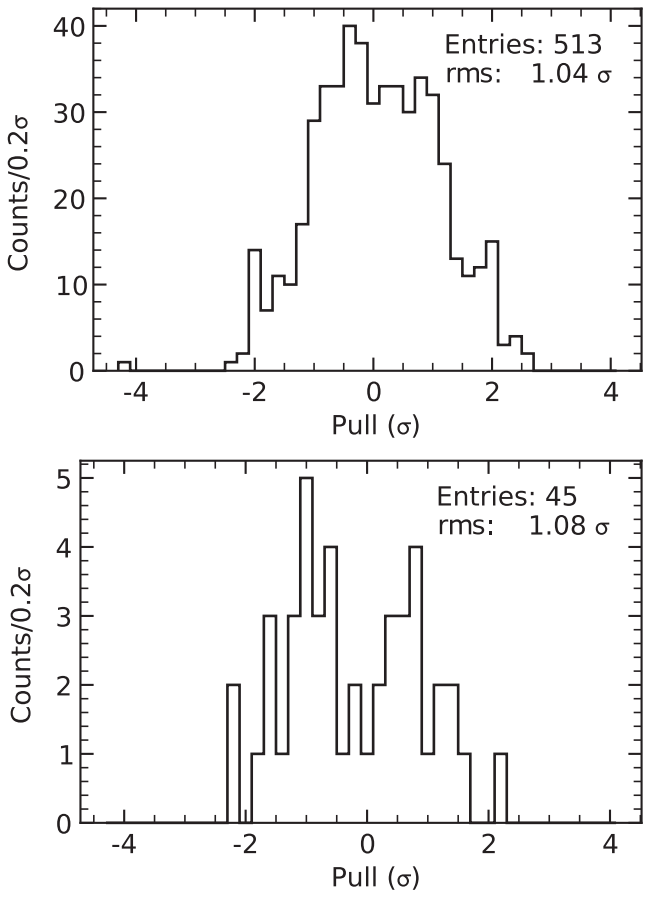

FIG. 9. Pulls between the 513 pairs of all ω_a measurements (top panel) and 45 pairs of A- and RA-method measurements that are used in the ω_a averaging (bottom panel). The pulls are defined as $(y_i - y_j)/\sigma_{ij}$ where y_i, y_j are the two measurements and σ_{ij} is the estimated uncertainty on their difference. The values of σ_{ij} are computed using the statistical and systematic uncertainties and their estimated correlations.

differences. Note in Table IV, the apparent systematic differences between the A-method analyses and the T-method analyses are consistent with the allowed differences between these methods.

We define the pulls between pairs of ω_a^m determinations as $(y_i - y_j)/\sigma_{ij}$, where y_i, y_j is the measurement pair and σ_{ij} is the corresponding allowed statistical and systematic differences. For each set of 19 ω_a^m determinations, there are 171 analysis pairs and therefore a total 513 comparisons across the three datasets.

In Fig. 9, we plot the 513 pulls for all ω_a^m measurements and the 45 pulls from the eight A-method and RA-method measurements that are most relevant to the ω_a^m averaging. Their standard deviations are 1.04 and 1.08, respectively.

I. Consistency checks

Beyond the fit χ^2, fit residuals, and the discrete Fourier transform of the fit residuals, a number of checks were made on the robustness of the results for the frequency ω_a^m and other parameters.

All analyses fit their time distributions with incrementally increasing start times to probe the stability of the fit parameters. A representative start time scan, for an A-method analysis of the Run-3a dataset, is shown in Fig. 10. The start time scan dependence of ω_a^m is sensitive to effects that vary from early to late in fill such as cyclotron rotation, positron pileup, and gain changes. All analyses demonstrated the start time scan stability of fitted ω_a^m values within the allowed statistical deviations.

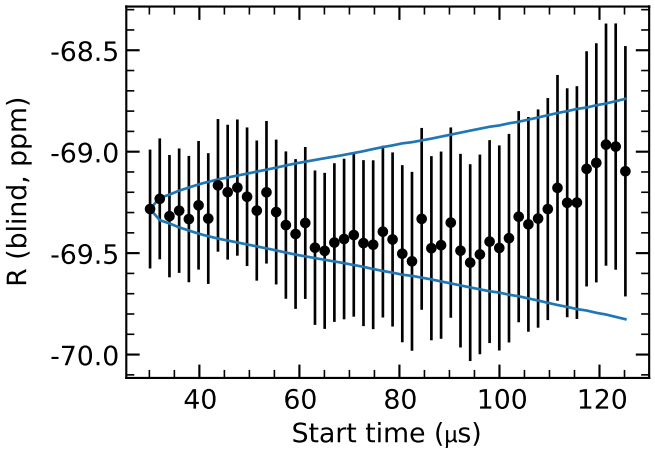

FIG. 10. A representative scan of the blinded R value versus the fit start time for the Run-3a dataset and the asymmetry-weighted histogramming method. The black data points are the R-value fit results. The point-to-point values are highly correlated and the smooth blue curve is the 1 allowed standard deviation band of any fit result from the canonical 30.1 μs fit start time. The allowed deviation band accounts for the statistical correlations between the 30.1 μs and > 30.1 μs fit results. Note the vertical axis includes an analysis-dependent software blinding and cannot be compared to Fig. 8 and Table IV.

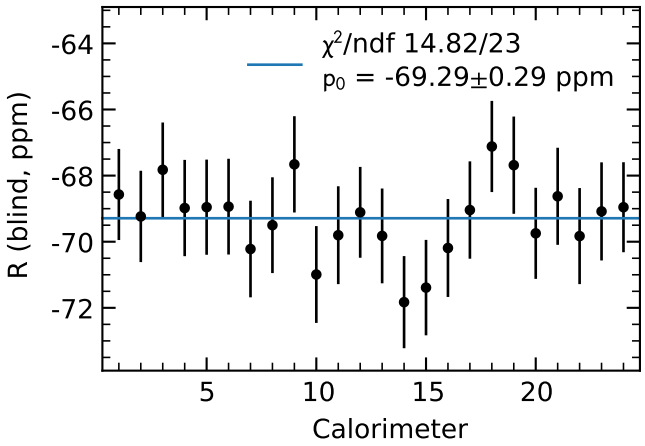

FIG. 11. A representative scan of the blinded R-value versus the calorimeter index for the Run-3a dataset and the asymmetry-weighted histogramming method. The black data points are the R-value fit results, and the solid blue line is a straight-line fit to the 24 individual calorimeter R values. Note the vertical axis includes an analysis-dependent software blinding and cannot be compared to Fig. 8 and Table IV.

All analyses fit the 24 time distributions of the individual calorimeters to perform calorimeter scans. A representative calorimeter scan, for an A-method analysis of the 3a dataset, is shown in Fig. 11. The calorimeter scan dependence of ω_a^m is sensitive to effects from cyclotron rotation and CBO modulation that are larger in the individual calorimeters than the calorimeter sum (as a result of the 2π phase advance of the cyclotron rotation and the CBO modulation around the ring circumference). All analyses demonstrated the calorimeter scan stability of fitted ω_a^m values within the allowed statistical deviations.

Fits as a function of the positron energy were also performed for the event-based analyses. Such energy scans are sensitive to effects of positron pileup and gain changes that vary with energy. No evidence was found for ω_a^m variation with positron energy.

All analyses also reported the correlation coefficients between the fit parameters in their ω_a^m fits. A large, known correlation exists between the frequency ω_a^m and its phase ϕ. A smaller, known correlation exists between the frequency ω_a^m and the frequency and phase parameters of the leading-order CBO term.

J. Systematic uncertainties

The systematic uncertainties reflect the inevitable shortcomings in modeling the true behavior of beam dynamics and other effects. Each analysis made reasonable choices for the required modeling of the various effects in the data, and each analysis made independent estimates of systematic errors. The reported errors are averaged across the analysis groups with the same weightings as the ω_a^m averages.

TABLE V. Summary of the major systematic uncertainties for the ω_a^m analysis of the three datasets. The major systematic uncertainties arose from the handling of CBO effects, the corrections for gain changes and positron pileup, and the presence of a residual slow effect. "Other systematics" refers to the sum of all other systematic uncertainties.

Systematic uncertainty	Run-2 (ppb)	Run-3a (ppb)	Run-3b (ppb)	Run-2/3 (ppb)
CBO handling	22	18	28	21
Pileup corrections	9	6	7	7
Gain corrections	5	4	5	5
Residual slow effect	5	14	10	10
Other systematics	2	5	3	4
Total	25	24	31	25

The major sources of ω_a^m systematic uncertainties are summarized in Table V. The treatment of the CBO distortions of the time distributions provides the largest source of systematic uncertainty. The pileup and gain corrections (see Sec. IV C) and presence of residual slow effects (see Sec. IV E 5) also yield significant systematic uncertainties. The total systematic uncertainty for the three datasets varies from 24 to 31 ppb.

Each of the above systematic categories contains multiple contributions. In general, we assume that the contributions to a specific category may be correlated and are summed linearly [39]. Conversely, we assume that systematics from different categories are not correlated and are summed quadratically.

The total systematic uncertainty for the ω_a^m analysis is about two times smaller than Run-1 (56 ppb). First, in Run-2/3, the CBO systematic was reduced through studies that determined that the contributions from the CBO decoherence envelope and the CBO frequency change are uncorrelated and add in quadrature. Second, in Run-2/3, the pileup systematic was reduced through a combination of improved reconstruction algorithms, which yielded less pileup, and improved correction in more analyses. A pileup phase uncertainty was also shown to be overestimated in the Run-1 analysis. Third, in Run-2/3, the source of the residual slow effect became partially understood, thus reducing this systematic.

The following subsections discuss our procedures for estimating the CBO, pileup, slow term, gain, and other systematics.

1. CBO systematic

Three significant uncertainties from beam dynamics were identified: uncertainty in the shape of the CBO decoherence envelope, uncertainty in the drift of the CBO frequency, and uncertainty in the lifetime of the CBO effects on the precession asymmetry and its phase.

Note that the CBO envelope changed from Run-3a to Run-3b as a result of the increased kicker voltage. For datasets Run-2 and Run-3a, a simple exponential envelope was sufficient to model the CBO decoherence. For Run-3b, an additional constant term was needed to model the CBO decoherence.

To estimate the systematic associated with envelope shapes, the analyses studied a variety of envelope functions. The shapes incorporated constant, exponential, and reciprocal terms and their combinations. The systematic was estimated from the changes of the ω_a^m results for all functions with an acceptable χ^2 value. The average contribution of the CBO decoherence systematic across the datasets and analyses in Table V was about 16 ppb.

The Run-2/3 CBO frequency drift was roughly ten times smaller than the Run-1 drift due to the repair of the ESQ resistors [6]. The Run-2/3 drifts, attributed to the effects of quadrupole scraping and calorimeter acceptance, were modeled as an exponential relaxation of the CBO frequency. The associated systematic uncertainty originates from the poorly known relaxation lifetime. The average contribution of the frequency-drift systematic across the datasets and analyses in Table V was about 10 ppb.

Last, as discussed in Sec. IV E 2, the CBO also modulates the precession asymmetry A_0 and precession phase ϕ_0. These effects are similarly modeled by a sinusoidal oscillation with a decoherence envelope. The effects on A_0 and ϕ_0 are small and their impacts on determining ω_a are negligible compared to the CBO decoherence systematic and the CBO frequency-shift systematic.

2. Pileup systematic

The procedures for correcting the time distribution for pileup distortions are discussed in Sec. IV C 2. The corrections involve superimposing either digitized waveforms, crystal hits, or positron candidates. This pileup modeling is subject to inaccuracies in our knowledge of the detector response and the analysis reconstruction. Further systematics include errors in the pileup rate, errors in the pileup time distribution, and the truncation of the pileup correction at a finite order. Errors arising from unseen pileup—pileup below the threshold for the reconstruction—were also evaluated.

The two largest contributors to the pileup uncertainty are the accuracy of the pileup model, roughly 2 ppb, and the error from the unseen pileup, also roughly 2 ppb. The various other sources of pileup systematic uncertainties were $\mathcal{O}(1\ \text{ppb})$.

We note that the uncertainty in the overall normalization of the pileup correction is about 1%. This is determined by comparing the raw energy and reconstructed-pileup energy distributions in the region above 3.1 GeV (see Fig. 6). This has a negligible contribution to the systematic uncertainty.

3. Residual slow term systematic

As already discussed, both Run-1 data and Run-2/3 data indicated a residual slow effect in the event-based time distributions. Its handling is described in Sec. IV E 5.

In the local-fitting, event-based analyses, we identified an energy-scale shift as a contribution to the residual slow effect. The local-fitting analyses either explicitly corrected their analyses for the energy-scale shift or treated the effect as a systematic as in Run-1.

The remaining effect—about one-third of the size of the energy-scale shift—has unknown origin(s). To evaluate the associated systematic, we applied a "gainlike" correction to accommodate the effect and evaluate its impact on ω_a^m. Two approaches for applying this correction were developed. One method utilized the χ^2 of the fit, and another method equalized the muon-loss normalization across energy bins. Both methods were consistent, and the impact on ω_a^m was 5 to 10 ppb.

Also included within this systematic category—because it is highly correlated with the residual slow term—is the uncertainty assigned to the fit preference for a nonphysical, negative, $k_{\rm loss}$ parameter in Run-3a and 3b [40]. This systematic is estimated from the ω_a^m shift required to return to $k_{\rm loss} \geq 0$. The total systematic for this category was estimated at 5 to 14 ppb.

4. Gain systematic

The procedures for correcting the time distributions for gain changes are discussed in Sec. IV C 1. The long-term gain correction has a negligible effect on extracting ω_a^m, since this correction is a time-independent factor for each muon fill. The two other gain corrections, in-fill and short-term, do change with time in fill.

Both the in-fill gain change and short-term gain change were modeled as exponential relaxations of gain sags. The in-fill gain correction is larger and dominates the gain systematic.

The sensitivity to the in-fill gain parameters is determined by scaling the correction and observing the change in ω_a^m. This sensitivity is then combined with the uncertainty on the parameters obtained from the laser calibration system. Uncertainties are conservatively assumed to be fully correlated across all calorimeter crystals. The resulting in-fill gain systematic is roughly 4 ppb. The same procedure is applied in estimating the smaller short-term gain systematic.

5. Other systematics

The remaining categories of systematic uncertainties considered are the timing calibration of the individual calorimeter channels, the time randomization for the fast rotation handling, the shape of the reconstructed muon loss time distribution, and the requirement of a fixed muon lifetime and precession period in the ratio histogram construction. The largest was the muon loss systematic, which contributed an uncertainty of 1 to 5 ppb.

K. Combination of ω_a^m measurements

To define a single measured value of ω_a^m for each of datasets Run-2, Run-3a, and Run-3b, we performed an equal-weighted average of the six measurements I-A, II-A, III-RA, IV-RA, V-A, and VI-A where I-A, *etc.*, denote the analysis group and histogram method. This strategy combines two local-fitting A-method analyses, two global-fitting A-method analyses, and two ratio histogramming A-method analyses. We did not include measurements using the T, RT, Q, or RQ methods because their statistical uncertainties are significantly larger, their systematic uncertainties are similar or larger, and their estimated correlations imply no appreciable reduction of the uncertainty of the average.

For each dataset, we conservatively assume that the statistical uncertainty and each systematic category uncertainty are fully correlated between the six averaged measurements. In such circumstances, both the statistical uncertainty and the individual systematic uncertainties of the dataset average, are the plain average of the six measurements. Each systematic category uncertainty is also conservatively assumed to be fully correlated across the three datasets.

As mentioned in Sec. IV H, we estimated the statistical correlations between the ω_a^m measurements within the same dataset (see Table XXVII). The statistical correlations between the six averaged analyses range from 0.993 to 1.000. The optimal linear combination of the six measurements in a χ^2 fit using these correlations has an uncertainty that is only 1.5% smaller than the plain average. Consequently, considering that the estimated correlations have significant uncertainties, we use the aforementioned plain average in computing ω_a^m.

V. BEAM DYNAMICS CORRECTIONS

This section reviews the analysis and evaluation of the five beam dynamics corrections to ω_m^a, introduced in Sec. I.

A. Electric-field correction

The radial electric-field contribution from the ESQ to ω_a in Eq. (1) cancels only for magic-momentum muons. The electric-field correction C_e accounts for the spin precession in ω_a^m induced by the momentum spread of the stored muon beam.

Expanding the second term in Eq. (1) to the first order in the muon momentum offset from the magic momentum p_0, the shift relative to the ideal frequency is

$$\frac{\Delta\omega_a}{\omega_a} = -2\frac{\beta_0}{cB_0}\delta E_x, \tag{10}$$

where $\delta = (p - p_0)/p_0$, β_0 is the magic-momentum velocity, B_0 the vertical magnetic field, and E_x the radial component of the ESQ electric field. For small radial displacements, x, from the center of the ESQ, the electric field is approximately linear

$$E_x \approx n\frac{\beta_0 cB_0}{R_0}x, \tag{11}$$

where $n \approx 0.108$ is the effective focusing field index (accounting for the finite lengths of the quadrupole sections) and R_0 is the magic-momentum bending radius. The muon-momentum offset can also be expressed in terms of the radial displacement from R_0, x_e, and the field index via the dispersion relation

$$\delta \approx (1-n)\frac{x_e}{R_0}. \tag{12}$$

The electric-field correction averaged over all momenta is

$$C_e = -\left\langle \frac{\Delta\omega_a}{\omega_a} \right\rangle \approx 2n(1-n)\beta_0^2 \frac{\langle x_e^2 \rangle}{R_0^2}. \tag{13}$$

The following sections describe the two analyses used to evaluate the electric-field correction and the results.

1. Fast-rotation analysis

Because the tangential speed, β_0, is constant to the ppm level for the stored muons, the measured cyclotron angular frequency, ω_c, determines the radial displacement x_e through

$$\beta_0 \approx R\omega_c = (R_0 + x_e)\omega_c. \tag{14}$$

The cyclotron frequency spread of the muons modulates the decay positron intensity detected by the calorimeters and is referred to as the fast-rotation signal. In the fast-rotation analysis, we use this signal to reconstruct the momentum distribution of the stored muons for the determination of C_e. At the start of a fill, the stored muons are tightly bunched. As the fill progresses, the muons spread out azimuthally over time due to the spread in their momenta. This effect leads to decoherence of the fast-rotation signal shown in Fig. 12.

The fast-rotation component of the positron intensity signal is isolated in two ways:

(1) Smearing method: The pulses of the decay positron time spectrum are randomly split into two halves: a numerator and a denominator. Each detection time in the denominator is randomized by an amount uniformly distributed between $\pm T_c/2$, where T_c is the revolution period. This randomization smears out the fast rotation in the denominator while slower features remain intact. Slowly changing features common to the numerator and denominator are eliminated in the ratio, leaving only the fast-rotation signal from the numerator.

(2) Fit method: The decay positron signal is binned at intervals of the expected revolution period, which approximately removes the fast rotation. The resulting histogram is then fit using a simplified version of the ω_a^m analysis fit model, which accounts for the most important features. The finely binned decay positron time spectrum is then divided by the fit function. As in the smearing method, the only prominent oscillation in the resulting ratio histogram is the fast rotation. Figure 12 shows an example of a fast-rotation signal from Run-2 isolated by the fit method.

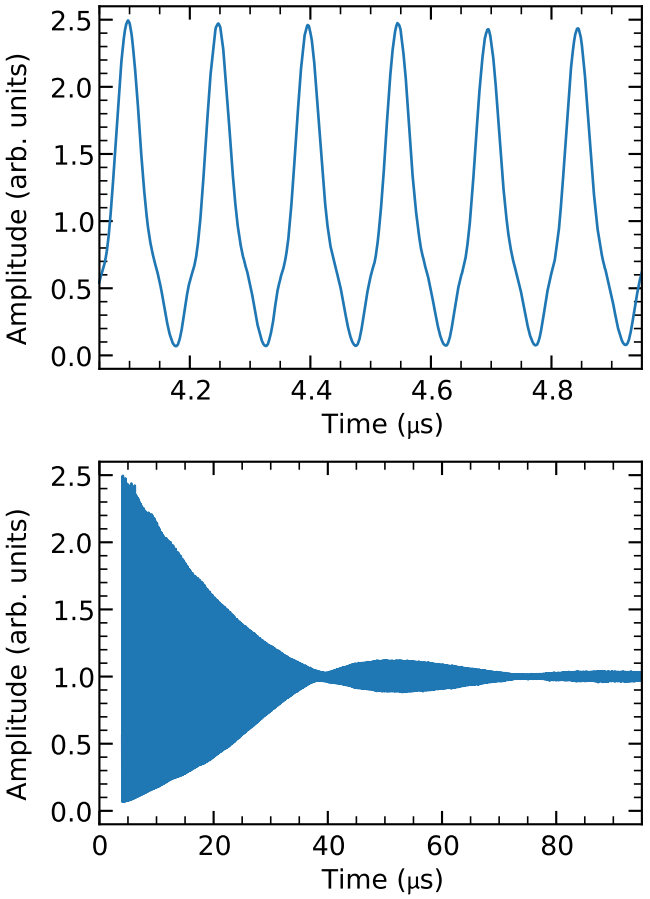

FIG. 12. Fast-rotation signal from Run-2 data, showing individual turns around the storage ring over short timescales (top) and broader decoherence envelope over long time scales (bottom).

The fast-rotation signal $S(t)$ can be modeled as a weighted combination of periodic impulse trains with frequencies ω and time offsets τ, representing periodic detection of the circulating muon bunch, yielding

$$S(t) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\sum_m \delta\left[t - \left(\frac{2\pi m}{\omega} + \tau\right)\right]\rho(\omega,\tau)d\omega d\tau, \tag{15}$$

where m is the turn index around the storage ring and $\rho(\omega,\tau)$ the joint distribution of revolution frequencies and injection times for stored muons. Analysis approaches, based on Fourier analysis or a fit to the time-domain signal, are used to estimate the frequency distribution based on this model.

The Fourier analysis depends on the important assumption that $\rho(\omega,\tau)$ is separable. However, this is generally not true since the kicker pulse is not flat over the width of the injected pulse and preferentially stores different momenta in different time slices of the injected bunch. This "momentum-time correlation" causes a systematic distortion to the Fourier analysis, which depends on the kicker pulse shape. To rectify this feature, an alternative analysis, named the "fast-rotation χ^2 method" and based on a method invented for the CERN storage ring experiments, accounts for the momentum time correlation. The results from this analysis can be used to correct the Fourier method. In the CERN method, the fast-rotation signal $S(t)$ is fit with a simple debunching model. Integrating Eq. (15) over narrow bins for ω and τ, where the weight $\rho(\omega,\tau)$ is approximately constant for each bin, yields the contribution of each (ω,τ) bin to the signal at time t. Denoting this component as $(\beta_{ij})_k$, where i and j label the (ω,τ) bin, and k labels the time bin of the fast-rotation signal, the overall signal S_k may be expressed as a linear combination of these component signals, yielding

$$S_k = \sum_{i,j} (\beta_{ij})_k \rho_{ij}, \tag{16}$$

where ρ_{ij} are the unknown weights of the discretized $\rho(\omega,\tau)$ distribution, treated here as fit parameters determined from the fits.

This prescription typically allows too many free parameters to obtain physically reliable fit results. To impose constraints, the frequency distribution in each injection time slice is assumed to have the same fundamental shape as in the central time slice, but with features of the three lowest moments (mean, standard deviation, and skew) varying smoothly as quartic polynomials over the injection time using the sinh-arcsinh transformation [41]. This modeling reduces the number of parameters to 62: one frequency distribution (25 bins), one overall injection time distribution (25 bins), and 12 polynomial coefficients, which describe the momentum-time correlation. Our χ^2 minimization passes employed both the Davidon-Fletcher-Powell algorithm [42] and refinements with simulated annealing. Each spectrum was fit multiple times from different starting parameters. Because of systematic shape variations in the beam pulses, fits were performed separately on time spectra for each of the bunches delivered by the Fermilab accelerator complex, as well as for the summed spectrum; see Fig. 13 for a momentum distribution and Fig. 14 for a joint distribution obtained in this manner for data subsets from runs 3a and 3b.

We assessed the following systematic errors associated with the fast-rotation analysis methods: late start time, failure to remove stray frequencies from the signal, changes to the distribution created during scraping, and insufficient shape parameters.

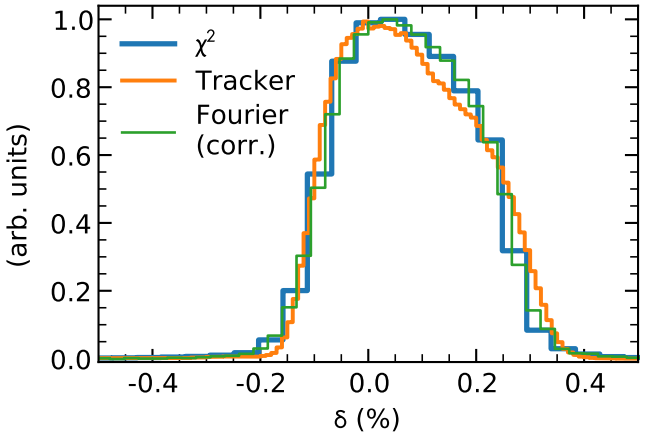

FIG. 13. Fractional momentum distributions from the fast-rotation χ^2 method, the tracking analysis method (data from the straw tracking detector at 180°), and the corrected Fourier analysis for the data subset 3F.

With a quantitative description of the systematic distortions contributed by the correlation between ω and τ, the Fourier analysis may then be corrected by evaluating the correlation-dependent parts using the correlation from the χ^2 method as an external input (see Fig. 13 for an example of the reconstructed momentum distribution obtained in this way). Thus, the corrected Fourier analysis is no longer completely independent from the fitting method, but it does enable a check for consistency between the two methods.

2. *Positron tracking analysis*

The stored beam exhibits a periodic pattern in which the initial narrow width imposed by passage through the inflector grows as the beam circulates due to the momentum dependence of the radial closed orbits. We developed a method for Run-2 and Run-3 datasets to reconstruct the

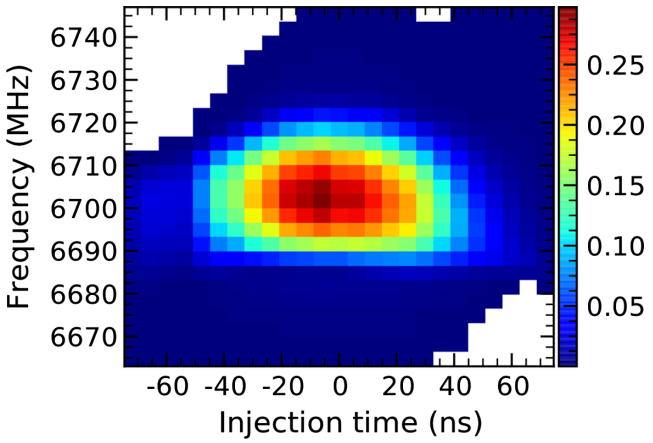

FIG. 14. Joint distribution from the fast-rotation χ^2 method of revolution frequency and injection time determined by the direct fit method for the data subset 3N, first bunch in the beam pulse sequence.

muon momentum distribution based on this behavior of the muons in the radial direction, x, which is directly observed by the positron tracking detectors until the betatron oscillations decohere. Figure 13 includes a sample of a momentum distribution derived from this analysis.

The minimum and maximum radial spreads are apart by half of a betatron period, which appears in data from a detector located at a specific azimuth as the aliased coherent period (see Table II). The momentum-dependent magnetic rigidity $B_0R = p_0(1+\delta)/e$ governs the amount of the spread. The linear matrix of an inhomogeneous magnet with field index n [31] well describes this spectrometric relation between the momentum and radial coordinates, which takes on a simple form for two states, i and f, separated by a phase advance of $\pi/\sqrt{1-n}$ (or, equivalently, separated in time by $\sim T_{\mathrm{CBO}}/2$ at a fixed detector):

$$\begin{pmatrix} x \\ x' \\ \delta \end{pmatrix}_f = \begin{pmatrix} -1 & 0 & 2\frac{R_0}{1-n} \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ x' \\ \delta \end{pmatrix}_i . \tag{17}$$

In Eq. (17), the variables x and x' represent the spatial and angular offsets in radial phase space. From the radial coordinate x_f expressed in terms of the state-i coordinates, the spectrometric relation is

$$\delta = \frac{1-n}{2R_0}(x_i + x_f). \tag{18}$$

From Eq. (18), the radial distribution at state f would equal the momentum distribution, shifted by x_i and scaled by $(1-n)/2R_0$, if all the stored muons were to share the same coordinate x_i. For Run-2 and Run-3, the tracking detectors measured a radial beam that resembled this idealized scenario. Therefore, by defining x_i as the radial mean of the stored beam when the radial width is minimal, we implemented Eq. (18) to reconstruct the momentum spread from which $\langle\delta^2\rangle$ is taken to calculate the electric-field correction via Eq. (13).

The method is validated with realistic beam-tracking simulations using the GM2RINGSIM package [8]. The associated uncertainty is only significant for Run-3b, as shown in Table VI. In this dataset, the beam simulation shows a discrepancy between the truth and reconstructed momentum distributions using the tracking analysis. The discrepancy grows over time while the truth values stay stable, and the reconstructed value falls with time, which is not present in the Run-2 or Run-3a simulations. We see the same behavior in the data analysis of Run-3b, where the reconstructed value of C_e steadily decreases over time, so we consider this behavior a real effect also present in the data. Hence, we apply a 28 ppb correction to the results obtained for Run-3b, which comes directly from comparing truth and reconstruction in the simulation. Given the reliance on simulation, we apply a 100% uncertainty 28 ppb on this correction for the Run-3b dataset.

TABLE VI. Uncertainties of the electric-field correction from the tracking analysis.

	Uncertainty (ppb)		
Description	Run-2	Run-3a	Run-3b
Statistical			
Station 12	0.7	0.3	0.4
Station 18	0.8	0.4	0.5
Systematic			
Method			
Beam simulation	5.4	5.0	27.8
Detector effects			
Tracker resolution	5.0	5.0	5.0
Tracker acceptance	21.8	21.5	18.3
Tracker alignment	21.0	20.3	11.1
Calorimeter acceptance	2.0	2.0	2.0
Other effects			
Tracker station differences	4.0	4.8	1.7
Total	31	31	35

The uncertainties from the tracking analysis are dominated by acceptance correction, alignment, and simulation uncertainties. The acceptance correction uncertainties are approximately 20 ppb for all three datasets. This value comes from conservatively varying the shape of the known correction by ±50%.

The uncertainty in the analysis associated with tracker alignment emerges from the ±0.6 mm uncertainty of the detector radial locations, assumed as uncorrelated between the two tracker stations (its effect is thus reduced by a factor of $1/\sqrt{2}$). This uncertainty is smaller in Run-3b because the systematic bias resulting from an error in tracker alignment scales with the mean value of the muon momentum distribution. In Run-3b, the mean momentum relative to p_0, $\langle\delta\rangle$, is smaller than the width, σ_δ, due to increased kick strength, and thus when we add the sum of squares to get

$$C_e = \frac{2n\beta_0^2}{1-n}(\langle\delta\rangle^2 + \sigma_\delta^2), \tag{19}$$

it is less significant.

The resolution uncertainty in this analysis assumes a detector resolution of ∼3.5 mm on the tracker reconstruction of the transverse muon coordinates. Resolution studies at early times after injection indicate a 25% uncertainty on this value, and we assess the associated systematic uncertainly by scaling the correction by ±25%. The sensitivity of the reconstructions to such resolution uncertainties has an upper limit of 5 ppb, which we assign as a systematic uncertainty. The effect of mismatching

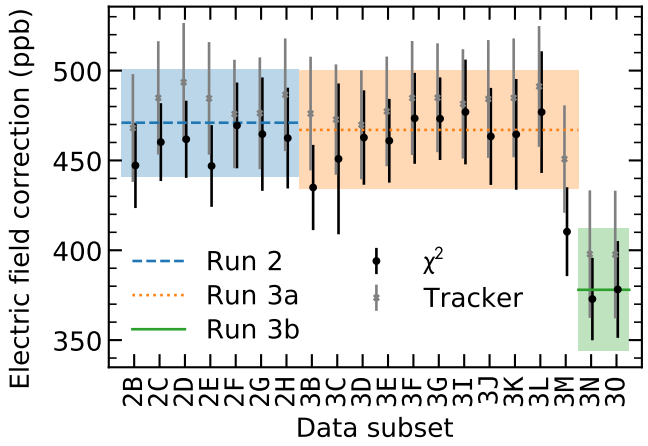

FIG. 15. Electric-field corrections C_e by data subset obtained from the tracking analysis method and the fast-rotation χ^2 method. The final values for runs 2, 3a, and 3b are shown in color, which come from the combination of the calorimeter and tracker-based analyses.

calorimeter-vs-tracker acceptances is small, as shown in Table VI.

The last systematic error in this analysis arises from differences between C_e reconstructions from the two tracker stations. Such difference potentially emerges from additional closed orbit distortions due to ESQ plate misalignments.

3. Results

Figure 15 shows the electric-field correction from the fast-rotation fitting analysis, the positron tracking analysis, and the weighted average of the analyses.

The tracking analysis is insensitive to the momentum-time correlation, whereas the fast-rotation fitting method was designed to incorporate momentum-time correlation, and the fast-rotation Fourier method is subject to significant distortions caused by momentum-time correlation.

Results from the tracking analysis at the data-subset level are generally larger than the fast rotation by 16–31 ppb. The difference in the results from these independent methods is taken into account to estimate the systematic uncertainty of the electric-field correction.

The final results for C_e are presented in Table VII. The combined result is the weighted average, assuming the uncertainties for each are completely uncorrelated. The electric-field correction is significantly smaller for Run-3b due to the better-centered momentum distribution of the stored beam.

A separate class of uncertainty in the final values of the combined result was evaluated, namely, the alignment and voltage errors of the ESQ stations, which correspond to an uncertainty of 6 ppb. This error applies equally to the tracking- and fast-rotation-based analyses and is added in quadrature to the uncertainty of the combined result. We intend to conduct more extensive research to better understand the uncertainties associated with the recently developed techniques for determining the electric-field correction. For this reason, we increase the calculated uncertainties by a factor of 1.5. The final uncertainty values are at the level of 30–33 ppb, as shown in Table VII.

B. Pitch correction

The electric field that keeps the beam confined in the vertical direction drives a radial component of the spin angular frequency [43], which biases ω_a. The pitch correction

$$C_p = \frac{1}{2}\langle \psi^2 \rangle, \tag{20}$$

where $\psi = \frac{dy}{dz}$ is the pitch angle, corrects this bias. This angle is calculated in accordance with sinusoidal vertical betatron motion:

$$y = A\sin(kz + \phi) + \bar{y}, \tag{21}$$

where z and $\bar{y}$ are the longitudinal coordinate and vertical mean position of muons in the storage ring, respectively. This expression allows Eq. (20) to be rewritten as

$$C_p = \frac{n}{4R_0^2}\langle A^2 \rangle. \tag{22}$$

Here, A is the amplitude of the beam's vertical oscillations, n is the field index, and R_0 is the magic momentum radius.

Two independent analyses, "method-1" and "method-2," determine C_p. Both start with the vertical decay distributions measured by the two straw tracking detectors located at 180° and 270°, following equal selection criteria, but apply different corrections for tracker resolution and

TABLE VII. Table of corrections and uncertainties for C_e (ppb) from the fast-rotation and tracking methods. Only the combined values are used for the full Run-2/3 dataset.

	Fast rotation		Tracking		Combined	
Dataset	Correction	Uncertainty	Correction	Uncertainty	Correction	Uncertainty
Run-2	459	24	485	31	469	30
Run-3a	459	28	475	31	466	32
Run-3b	367	27	398	35	378	33

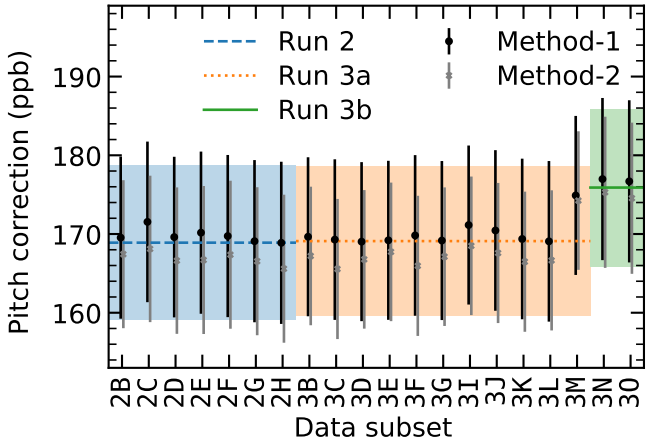

FIG. 16. Comparison between method 1 and method 2 of the pitch correction, C_p, results for all data subsets available in runs 2 and 3. The errors in the two methods are dominated by the tracking uncertainty.

acceptance. The resulting tracker data is transformed into amplitude space, and C_p is calculated using Eq. (22). Both methods then correct for the calorimeter acceptance. In this way, the calculated C_p reflects the bias on ω_a^m for the muon population contributing to the calorimeter measurement. The two methods calculate an average C_p for each dataset, as seen in Fig. 16. To make the switch to the amplitude space, method 1 derives a functional form, whereas method 2 uses a data-driven approach to estimate the amplitude distributions. In the end, results are within ∼2.5 ppb of each other, consistent with the statistical and systematical errors. Central values are calculated for each dataset, and we adopt the average of the final values from the two methods as the final C_p result presented in Table VIII. The ∼8 ppb uncertainty from the tracking hardware and vertical coordinates reconstruction dominate the systematic uncertainties shown in Table VIII, compared to other systematic errors from the amplitude fits, tracker acceptance and resolution correction, calorimeter acceptance, ESQ calibration, and tracker station differences.

C. Muon-loss correction

Muon losses, defined in Sec. III C 2, can bias the extraction of ω_a^m due mainly to the correlation between the $g-2$ phase, ϕ_0, and average momentum, p, of the lost muons distribution. The $g-2$ phase is a single term in the parameter function to extract the anomalous precession frequency (see Sec. IV E 1), and it represents the ensemble-averaged spin phase referenced at the nominal injection time. Since the momentum of the stored beam could change over the data taking as muons are lost, we introduce the muon-loss correction, C_{ml}, to cancel out the resulting biasing on ω_a, where

$$C_{ml} = -\frac{\Delta\omega_a}{\omega_a} = \frac{1}{\omega_a}\frac{d\phi_0}{dt} = \frac{1}{\omega_a}\frac{d\phi_0}{dp}\left(\frac{dp}{dt}\right)_{ml}. \quad (23)$$

The time dependence of the lost muons' momentum distribution, $(dp/dt)_{ml}$, is directly proportional to both the momentum dependence of the loss probability and the overall rate of muon losses [8]. The mechanism in which the phase is correlated with momentum is described in Sec. V D 1.

For Run-1, C_{ml} introduces a $\mathcal{O}$(5–20 ppb) correction [8]. Post Run-1, systematic studies show a momentum dependence of the muon losses for Run-2/3 running conditions similar to Run-1 results; meanwhile, the phase-momentum correlation $d\phi_0/dp$ at injection (which is denoted $t_0 = 0$) is increased in magnitude from -10 ± 1.6 to -13.5 ± 1.4 mrad/(%δ). This increase is attributed to the addition of a momentum cooling wedge in the upstream beamline during Run-2 [30]. The uncertainties of the measurements come from data fitting, magnetic field uncertainties, dataset differences, and gain changes.

Despite these differences, the dominant factor in the determination of the muon loss correction is the order of magnitude reduction in losses from Run-2 onward. Owing to this upgrade, the gradient $(dp/dt)_{ml}$ and therefore C_{ml} is reduced by an order of magnitude, reaching the sub-ppb level. C_{ml} is calculated with a conservative uncertainty attached as 3 ppb:

$$C_{ml} = 0 \pm 3 \text{ ppb}. \quad (24)$$

D. Differential decay correction

The differential decay correction, C_{dd}, accounts for the time dependence of the $g-2$ phase ϕ_0 (defined in Sec. V C) due to the spread of muon lifetimes in the beam. We refer to this spread of decay rate as a function of beam particle momentum as "differential decay." The correction is thus expressed as

$$C_{dd} = -\frac{\Delta\omega_a}{\omega_a} = \frac{1}{\omega_a}\frac{d\phi_0}{dt} = \frac{1}{\omega_a}\frac{d\phi_0}{dp}\left(\frac{dp}{dt}\right)_{dd}, \quad (25)$$

where $(dp/dt)_{dd}$ is the temporal variation of the beam-averaged momentum as muons decay in proportion to their time-dilated lifetimes, $\gamma(p)\tau_\mu$. The evolution of the

TABLE VIII. Pitch correction values, C_p, and associated statistical/systematic uncertainties (ppb) for Run-2, Run-3a, and Run-3b.

Dataset	Correction	Statistical uncertainty	Systematic uncertainty
Run-2	168.9	0.02	9.8
Run-3a	169.1	0.01	9.5
Run-3b	175.9	0.02	10.0

TABLE IX. Orbital variables $\mathbf{r} = \{x, x', y, y', t_0\}$. All the coordinates are relative to the reference axis at injection.

r_i	Definition
x, x'	Spatial and angular offsets in radial phase space
y, y'	Spatial and angular offsets in vertical phase space
t_0	Time relative to the nominal injection time.

momentum distribution can be approximated by

$$\left(\frac{dp}{dt}\right)_{dd} \approx \frac{p_0}{\gamma_0 \tau_\mu}\sigma_\delta^2, \tag{26}$$

where σ_δ^2 is the variance of the fractional-momentum distribution.

In addition to the initial $d\phi_0/dp$ from the upstream beamline (described in Sec. V D 1), there is an additional correlation that develops from the nonsymmetric kicker and longitudinal bunch structure during the injection process. Because of differential decay, the ensemble average phase slightly evolves throughout a fill, interpreted as a slight shift in the value of ω_a^m from the precession data fits. On the basis of the orbital coordinates $\mathbf{r} = \{x, x', y, y', t_0\}$ (see Table IX), the linear momentum dependence of $\phi_0(x, x', y, y', t_0; p)$ is expanded as

$$\frac{d\phi_0}{dp} = \frac{\partial\phi_0}{\partial x}\frac{dx}{dp} + \frac{\partial\phi_0}{\partial x'}\frac{dx'}{dp} + \frac{\partial\phi_0}{\partial y}\frac{dy}{dp} + \frac{\partial\phi_0}{\partial y'}\frac{dy'}{dp} + \frac{\partial\phi_0}{\partial t_0}\frac{dt_0}{dp} + \frac{\partial\phi_0}{\partial p}. \tag{27}$$

Beam tracking studies of the stored muons at injection from GM2RINGSIM simulations confirm the validity of this equality. From Eqs. (25) and (27), we divide the C_{dd} correction into three independent contributions based on their physical origins, namely the following: the *beamline*, *$p - x$ correlation*, and *$p - t_0$ correlation* effects.

1. Beamline effect

The direct correlation between the $g-2$ phase and momentum drives the beamline effect:

$$C_{dd}^{bl} = \frac{1}{\omega_a}\frac{\partial\phi_0}{\partial p}\frac{dp}{dt} \approx \frac{\sigma_\delta^2}{\omega_a \gamma_0 \tau_\mu}\frac{\partial\phi_0}{\partial\delta}. \tag{28}$$

After four revolutions of the muon beam around the delivery ring at Fermilab [44], the magnetic field of the bending dipole magnets contribute to a momentum-dependent angle advance between the muon spin and momentum by $\Delta\phi \approx 8\pi a_\mu \gamma$, which leads to $|\Delta\phi/\Delta\delta| = 8.6\ \text{mrad}/(\%\delta)$ [8]. For Run-1, beam tracking simulations and direct measurements of the correlation determined $|\partial\phi_0/\partial\delta|$ at beam injection to be $10 \pm 1.6\ \text{mrad}/(\%\delta)$; a result in agreement with the delivery-ring-only contribution $|\Delta\phi/\Delta\delta|$.

The first step to calculate C_{dd}^{bl} is to recreate the joint distribution for $\phi_0 - \delta$ of the stored muons at $t = 0$ for each data subset from a bivariate normal distribution. The correlation is defined from the $\partial\phi_0/\partial\delta$ measurements and the momentum projection is scaled with the corresponding momentum distributions, determined in the electric-field correction analysis. Then, a Monte Carlo signal with a simplified five-parameter version of Eq. (7) is prepared out of the $\phi_0 - \delta$ distribution, where the differential decay $e^{-\frac{t}{\gamma(t)\tau}}$ transforms the distribution over time. Finally, we fit the Monte Carlo signal to extract the shift in ω_a^m due to differential decay.

The difference between the results from the steps described above and Eq. (28) is negligible. The main purpose of the step-by-step procedure is to test the sensitivity of C_{dd}^{bl} to two possible systematic effects: correlations of γ and ϕ_0 with the muon-momentum dependence of (1) the asymmetry, A, and (2) emitted positrons, N, based on the leading-order Michel spectrum. Because these effects produce systematic uncertainties below 2 ppb, we assign a conservative upper limit of 3 ppb to the differential-decay beamline correction. Table X summarizes the evaluation of C_{dd}^{bl} for all the datasets based on the weighted results of the procedure for each data subset. The larger $\phi_0 - \delta$ correlation induced by the cooling wedge increases the beamline effect in Run-3a and Run-3b.

2. $p - x$ effect

At the exit of the inflector, the Muon Campus delivers a muon beam where the only sizable momentum-phase correlation is the one that is measured for the differential-decay beamline effect (i.e., $\partial\phi_0/\partial\delta$). This specific feature of the injected beam, which tracking simulations corroborate, is perturbed due to momentum-orbit correlations that develop during beam injection, where the radial and vertical phase-space coordinates x, x', y, and y' are the "orbit" coordinates in this context (see Table IX).

The beam injection is optimized to accommodate the radial beam within the storage ring admittance. The process introduces correlations between the radial phase coordinates and momentum, $dx'/d\delta$ and $dx/d\delta$, of the stored muons at injection time ($t = 0$). The resulting

TABLE X. Differential decay corrections (ppb) for Run-2, Run-3a, and Run-3b. The corresponding uncertainties (ppb) are enclosed in parentheses.

C_{dd}	Run-2	Run-3a	Run-3b	Section
Beamline	−12(3)	−17(3)	−20(3)	V D 1
p-x	−5(6)	−5(6)	−5(6)	V D 2
p-t_0	6(15)	0(16)	23(17)	V D 3
Total	−11(16)	−22(17)	−2(18)	V D

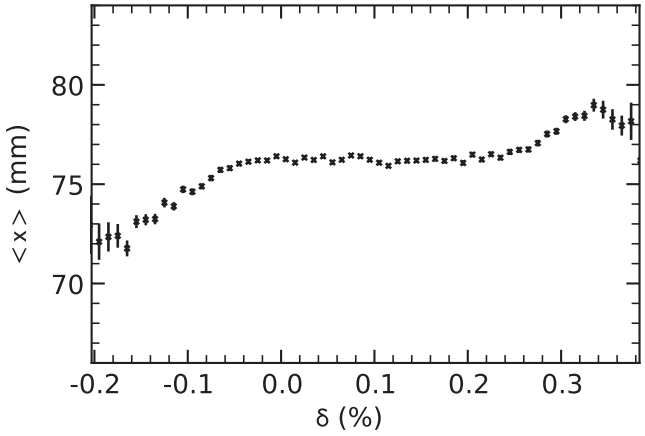

FIG. 17. Average radial coordinate $\langle x \rangle$ of the beam distribution per momentum offset at injection, from a GM2RINGSIM tracking simulation of stored muons. In this example, a nominal configuration of the injection parameters is implemented in the simulation. The $dx/d\delta$ correlations to quantify C_{dd}^{p-x} are obtained from these tracking simulation results.

differential-decay contribution from injection is hence expressed as

$$C_{dd}^{p-x} = \frac{\sigma_\delta^2}{\omega_a \gamma_0 \tau_\mu} \left(\frac{\partial \phi_0}{\partial x} \frac{dx}{d\delta} + \frac{\partial \phi_0}{\partial x'} \frac{dx'}{d\delta} \right). \tag{29}$$

While the pion beam decays into muons as it is transported down the muon-production beamline, the angle ϕ between each muon's momentum in the lab frame and its spin direction depends on the parental pion momentum, p_π, as

$$\sin(\phi) \approx \frac{2m_\mu}{m_\pi^2 - m_\mu^2} \frac{p_\pi}{c} \sin\theta, \tag{30}$$

where θ is the angle between the muon momentum and the pion direction in the lab frame. In our case, as muons are emitted in the lab frame in a forward cone of semiangle $\theta_{\max} \approx 12.7$ mrad, Eq. (30) is further simplified to

$$\sin(\phi) \approx 78.8 x_0', \tag{31}$$

where x_0' is the phase-space coordinate of the muon's trajectory at birth. Therefore, a nonzero correlation $\partial\phi_0/\partial x_0'$ exists, which yields nonzero $\phi_0 - x$ and $\phi_0 - x'$ correlations in Eq. (29) as muons subsequently execute betatron oscillations and cross bending magnets along the muon-production beamline. As shown in Eq. (29), these spin-orbit correlations couple with $dx/d\delta$ and $dx'/d\delta$ to alter the original phase-momentum relationship before injection.

With beam tracking simulations using the BMAD and GM2RINGSIM injection models [8], we calculate the beam correlations necessary to determine the differential-decay $p - x$ effect. Figure 17 shows the radial coordinate versus fractional momentum of the stored muons at injection, which is the dominant momentum-orbit correlation in C_{dd}^{p-x}. With Eq. (29) and the simulation results, the $p - x$-effect contribution to the differential-decay correction for Run-2/3 is

$$C_{dd}^{p-x} = -5 \pm 6 \text{ ppb}. \tag{32}$$

The uncertainty accounts for several simulation configurations in view of injection parameter configurations within operational ranges (i.e., inflector current, beam distributions at the inflector exit, and injection kicker strengths, pulse shapes, and relative timings).

3. $p - t_0$ effect

A muon's spin starts to precess as soon as it enters the storage ring. Typical muon bunches are 120 ns long; the spin of muons at the head of the bunch accumulates an additional precession $\Delta\phi \approx (120 \text{ ns})\omega_a$ relative to muons at the tail while they enter the ring. This longitudinal phase variation across the bunch, together with the t_0-dependent momentum acceptance induced by the time dependence of the injection kicker, produce the momentum-time effect:

$$C_{dd}^{p-t_0} = \frac{1}{\omega_a} \frac{\partial \phi_0}{\partial t_0} \frac{dt_0}{dp} \frac{dp}{dt} \approx \frac{\sigma_\delta^2}{\gamma_0 \tau_\mu} \frac{dt_0}{d\delta}. \tag{33}$$

The method to evaluate $C_{dd}^{p-t_0}$ is similar to the procedure used for the differential-decay beamline effect explained in Sec. V D 1, except for the first step where the muon distributions are prepared from the momentum-time distributions of the electric-field correction analysis; the time coordinates are transformed to relative spin phase advance via $\Delta\phi_0 = \omega_a t_0$ (Fig. 18 shows one example). The $C_{dd}^{p-t_0}$ is evaluated at the bunch level because each of the bunches in a sequence has characteristically different longitudinal

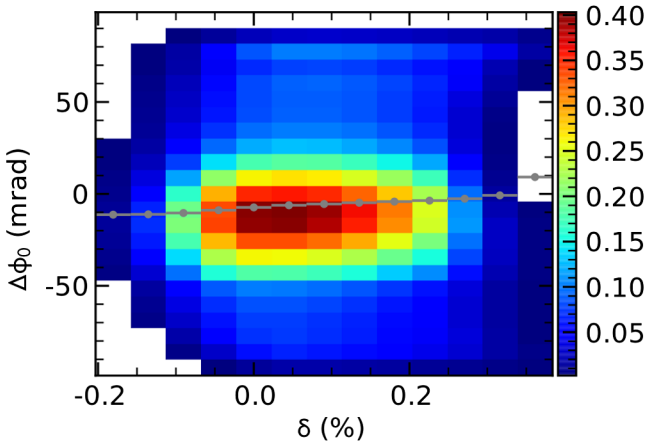

FIG. 18. Momentum-phase distribution from the momentum-time distribution for one bunch in data subset 2C. The gray markers are the averaged relative spin phases per fractional momentum, exhibiting the correlation that drives $C_{dd}^{p-t_0}$.

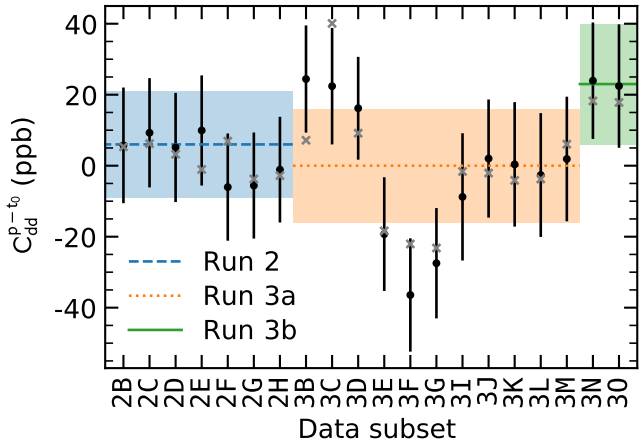

FIG. 19. Momentum-time differential decay correction $C_{dd}^{p-t_0}$ per data subset (black). In gray crosses, correction predictions where the ratio between $p-t_0$ correlations and kicker timing offsets relative to beam injection, based on GM2RINGSIM beam tracking simulations, is scaled in proportion to the per-data-subset kicker timing offsets.

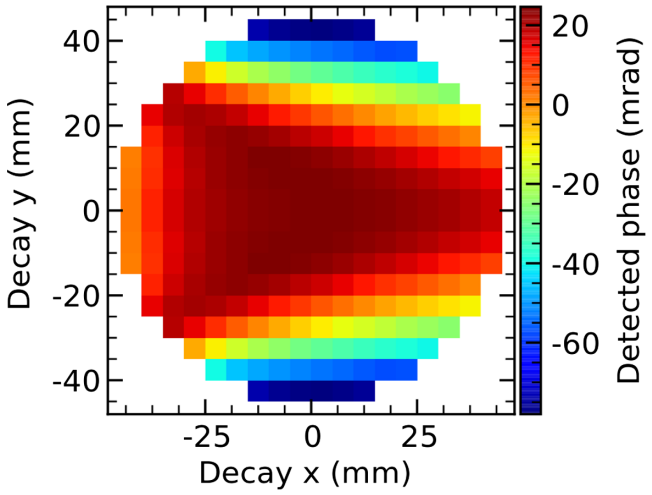

FIG. 20. Simulated azimuthally averaged phase maps for the asymmetry-weighted analysis. The coupling between the overall quadraticlike detected phase acceptance in the vertical direction and the in-fill reduction in vertical beam width is the most significant effect on C_{pa}.

intensity profiles. The results are then combined to obtain the corrections per data subset, as shown in Fig. 19. The final momentum-time corrections per run are summarized in Table X. The effect in Run-2 and Run-3a is consistent with zero, whereas a more constant timing offset between the kicker pulse and injection time leads to the nonzero correction for Run-3b.

To assess the uncertainties in this correction, we prepare 100 momentum-time distributions, each seeded by different initial conditions in the fitting method for the electric-field correction. The $C_{dd}^{p-t_0}$ correction is thereafter calculated for each seed, where the standard deviation for each set of bunches is treated as the uncertainty. The uncertainties per data subset are the correlated combination of the uncertainty from each bunch. An additional uncertainty, added in quadrature with the previously explained errors, is assigned from the rms of all the mean-subtracted data subsets to account for the intrinsic ambiguity in the momentum-time distributions used to calculate the $p-t_0$ effect.

4. Total effect

The total differential decay correction is the combination of the beamline, $p-x$, and $p-t_0$ effects:

$$C_{dd} = C_{dd}^{bl} + C_{dd}^{p-x} + C_{dd}^{p-t_0}, \tag{34}$$

summarized in Table X. To first order, these are uncorrelated; their physical origin is independent of each other. Therefore, the errors of each individual differential-decay effect are added in quadrature.

E. Phase acceptance correction

The detected $g-2$ phase, as measured by the calorimeter detectors, varies over time as a function of the transverse beam coordinates of the muons (x, y). The beam transverse distribution changes with time and creates in-fill variations of the detected phase that could affect the fit model for ω_a^m, where the phase is expected to be time independent. For this detector-acceptance effect, we introduce the phase acceptance correction, C_{pa}.

The time-dependent phase $\phi_{\mathrm{pa}}(t)$ is computed by averaging the measured phase as a function of transverse coordinates (x,y) that are obtained from GM2RINGSIM. The time dependence of the transverse beam coordinates is extracted from tracker beam profiles $M^T(x, y, t)$, which generates a time-dependent phase by virtue of the correlation between the phase and the beam transverse distribution. Figure 20 is a transverse map of $\phi_{\mathrm{pa}}(x, y)$ averaged over the azimuth, obtained by fitting the asymmetry-weighted histogram used to extract ω_a^m (see Sec. IV A).

The tracker stations measure the $M^T(x, y, t)$ distribution at two locations around the ring, but the extraction of the measured ω_a^m is performed by calorimeters at 24 azimuthal locations. Therefore, we extrapolate the $M^T(x, y, t)$ profiles around the ring using GM2RINGSIM and COSY INFINITY beam dynamics simulations. Vertical $(y(\varphi, t))$ and radial $(x(\varphi, t))$ muon coordinates at any given azimuthal position φ are calculated by scaling the transverse coordinates from tracker measurements with the mean and width values from simulated beam distributions as

$$y(\varphi, t) = y_{\mathrm{trk}}(t)\frac{y^{\mathrm{rms}}(\varphi, t)}{y_{\mathrm{trk}}^{\mathrm{rms}}(t)}, \tag{35}$$

for the vertical width, and

$$x(\varphi, t) = \frac{x^{\mathrm{rms}}(\varphi, t)}{x_{\mathrm{trk}}^{\mathrm{rms}}(t)} \cdot [x_{trk}(t) - \bar{x}_{\mathrm{trk}}(t)] + \bar{x}(\varphi, t), \tag{36}$$

for the radial motion of the beam, where $(x^{\mathrm{rms}}, y^{\mathrm{rms}})$ are the root mean squares of the transverse beam distributions and $\bar{x}$ is the radial distribution average. The quantities from simulated distributions on the right-hand side in Eqs. (36) and (35) do not have subscripts, whereas tracker-based values are denoted with the subscript "trk." By modifying the $M^T(x,y,t)$ distribution using Eqs. (36) and (35), we obtain the spatial and time distribution of the muons $M^c(x,y,t)$ at each calorimeter location. Combining the simulated maps with the muon distributions, a time-dependent phase $\phi^c_{pa}(t)$ can be computed for each calorimeter using the following weighted sum:

$$\phi^c_{\mathrm{pa}}(t) = \arctan\left[\frac{\sum_{ij} M^c(x_i,y_j,t)\cdot\varepsilon^c(x_i,y_j)}{\sum_{ij} M^c(x_i,y_j,t)\cdot\varepsilon^c(x_i,y_j)} \times \frac{\cdot A^c(x_i,y_j)\cdot\sin[\phi^c_{\mathrm{pa}}(x_i,y_j)]}{\cdot A^c(x_i,y_j)\cdot\cos[\phi^c_{\mathrm{pa}}(x_i,y_j)]}\right], \tag{37}$$

where acceptance, asymmetry and phase maps for a calorimeter "c" are represented by ε^c, A^c, and ϕ_{pa}, respectively.

The calculation of the phase acceptance correction is done by comparing ω_a^m to the fit of the simulated data. A histogram is generated for each calorimeter and for each parameter of the ω_a^m fit, including the modified $g-2$ phase obtained by fitting $\phi^c_{\mathrm{pa}}(t)$. Simulated data (produced using values extracted from histograms) are fitted with a constant phase. The difference between ω_a^m and the fit result determines C_{pa} for a given calorimeter.

Figure 21 shows the ϕ_{pa} time evolution for a Run-2 data subset, superimposed with one from Run-1d for comparison. After replacing the damaged resistors of the ESQ system from Run-1, the variation of the phase is highly reduced during Run-2/3, and the C_{pa} is hence smaller.

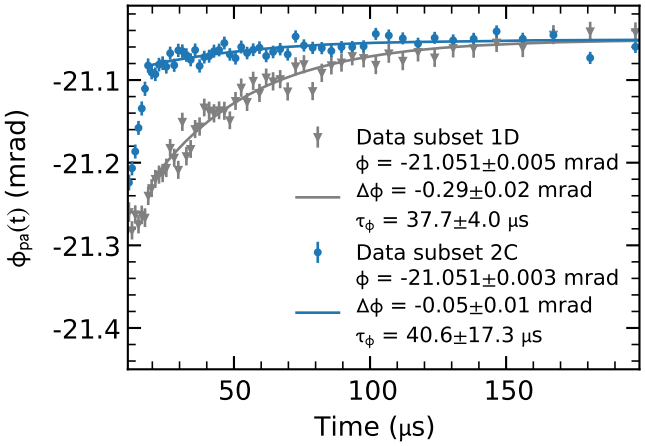

FIG. 21. Calculation of ϕ_{pa} for calorimeter 13 in data subset 1D (gray) and data subset 2C (blue) using data from the tracker station at 180°. The shown fit function is of the form $\phi + \Delta\phi\cdot e^{(-t/\tau_\phi)}$.

TABLE XI. Values of the phase-acceptance correction C_{pa} (ppb) and their statistical, systematic, and total uncertainties (ppb) for each of the Run-2/3 datasets.

Quantity	Run-2	Run-3a	Run-3b
Correction	−50	−16	−13
Statistical uncertainty	9	2	3
Systematic uncertainty			
Tracker and CBO	13	8	7
Phase maps	13	3	3
Beam dynamics	5	3	2
Total uncertainty	21	9	8

The central values of the correction are calculated by taking the average of the results from all calorimeters. The central values are shown in Table XI, where further improvement on the effect is observable in Run-3 with respect to Run-2. This outcome is due to the improved stability of the beam motion thanks to more optimized kicker settings and a better temperature stability of the main magnet. The evaluations of the statistical and systematic uncertainties are also reported in Table XI. The statistical uncertainty, which ranges from 2.0 to 7.8 ppb, originates from the limited number of tracks from the $M^T(x,y,t)$ collected by tracker stations. The sources of systematic uncertainty can be divided into three main groups. The first one stems from imperfect knowledge of the straw trackers' alignment, resolution, and acceptance, which directly affects the measured distribution $M^T(x,y,t)$. Next are the uncertainties associated with the estimation of the phase, asymmetry, and acceptance maps in Eq. (37) estimated using GM2RINGSIM. Lastly, the calculation utilizes beam dynamics functions obtained by simulation to extract the calorimeter $M^c(x,y,t)$ distribution from the tracker-based $M^T(x,y,t)$. Uncertainties are estimated by calculating C_{pa} while varying the beta functions and magnetic field within expected deviations based on the measurements.

F. Summary

The beam dynamics corrections and their uncertainties for Run-2/3 are listed in Table XII.

TABLE XII. Values and uncertainties of the beam dynamics corrections (ppb) for Run-2/3.

Quantity	Correction	Uncertainty
C_e	451	32
C_p	170	10
C_{ml}	0	3
C_{dd}	−15	17
C_{pa}	−27	13
Total	580	40

Each individual correction is highly correlated for different datasets, and therefore, the per-dataset combination of the uncertainties is fully correlated. To obtain the total beam dynamics correction uncertainty, we add the uncertainties of all the individual corrections in quadrature because they are uncorrelated.

A combination of improvements in the experimental setup (listed in Sec. III B) and analysis reduced both the beam dynamics correction magnitudes and uncertainties in Run-2/3 compared to Run-1. The replacement of the ESQ high-voltage resistors damaged in Run-1 leads to a smaller and more precise determination of C_{pa}. The muon loss correction is negligible thanks to the significantly reduced mechanical muon loss rates. With the stronger injection kickers in Run-3b, the more symmetric momentum distribution requires a lower electric-field correction, whereas the determination of the momentum-time beam correlations at injection, as well as an independent reconstruction of the momentum distribution based on the tracker detector data, reduce the uncertainty of C_e. While the differential decay correction was not included in Run-1, the momentum-time correlations analysis for the electric-field correction allowed us to fully quantify this correction in Run-2/3.

VI. MAGNETIC FIELD MEASUREMENT

In Eq. (2), $\tilde{B}$, the magnetic field averaged over space and time by the muons, is expressed as the precession frequency of protons in a spherical water sample at a reference temperature: $\tilde{\omega}'_p(T_r)$. In this notation, the tilde indicates the muon weighting, and the prime indicates that the proton magnetic moment is shielded in H_2O. The reference temperature is $T_r = 34.7\,°\text{C}$, the temperature at which the shielded proton magnetic moment was measured relative to the bound-state electron in hydrogen [45]. This section describes the measurements and analyses leading to $\tilde{\omega}'_p$, which follows from the general approach of Run-1 [7].

A. Magnetic field measurement principle

The muon-weighted magnetic field is derived from time-dependent maps of the magnetic field in the muon storage region $\omega'_p(x, y, \phi, t)$. The maps are derived from measurements by a set of NMR probes in a trolley that is pulled through the storage ring every two to three days and maps the full circumference in about 70 minutes. The field is mapped at the 17 NMR-probe positions $(x,\ y)$ ($x = 0$ at $r = R_0$) and about 9000 azimuthal positions ϕ. Corrections for differences of the physical ring configuration and from magnetic field transients from the kickers and ESQs, which are not operating during the trolley measurements, are discussed in Sec. VI G.

The trolley's NMR probes, described in [7], contain samples of proton-rich petroleum jelly (petrolatum). The trolley probes are calibrated to account for the sample and the different magnetic environment due to magnetic perturbations from the aluminum shell, the wheels of the trolley, the other probes, and other trolley components, including the electronics, cables, etc. A dedicated calibration magnetometer was used to correct each probe to the frequency that would be measured with a spherical water sample at temperature T_r. The details of this calibration procedure are described in Secs. VI B and VI C.

The time-dependent trolley maps are parametrized as

$$\omega'_p(x, y, \phi, t) = \sum_{i=1}^{N_{\max}} m_i(\phi, t) f_i(r, \theta), \tag{38}$$

where

$$f_i(r,\theta) = \begin{cases} 1 & \text{for } i = 1 \\ \left(\frac{r}{r_0}\right)^{\frac{i}{2}} \cos\left(\frac{i}{2}\theta\right) & \text{for even } i > 1 \\ \left(\frac{r}{r_0}\right)^{\frac{i-1}{2}} \sin\left(\frac{i-1}{2}\theta\right) & \text{for odd } i > 1 \end{cases}. \tag{39}$$

Here $r_0 = 4.5$ cm is a reference radius, $x = r\cos(\theta)$, $y = r\sin(\theta)$. The $\cos(\theta)$ and $\sin(\theta)$ terms are referred to as normal and skew moments, and t is the time of the measurement. The moments $m_i(\phi, t)$ are determined from fits of the 17 trolley-probe frequencies at the time t when the trolley is at the position ϕ. The parametrization in Eq. (38) is motivated by solutions to a 2D Laplace equation and is analogous to a 2D Taylor expansion around $(x, y) = (0, 0)$ with constraints. The 2D Laplace-equation solution is strictly valid only if B has no azimuthal dependence; the impact and validation of this parametrization and the effect of truncating the parametrization at $N_{\max}$ are discussed in Sec. VI D.

The time dependence of the moments $m_n(\phi, t)$ between trolley runs is estimated by interpolation making use of a set of 378 NMR magnetometers (fixed probes) mounted on the outside of the vacuum chambers at 72 azimuthal positions, called stations. Each fixed probe is read out with a rate of ~0.5 Hz. Each station has either four or six NMR probes, half above and half below the storage region, and can interpolate the magnetic field moments up to $i = 4$ or $i = 5$, respectively. As a trolley run proceeds, the moments calculated from the fixed probes at the stations near the trolley are set equal to the corresponding moments calculated from the trolley probes at that time, which we call "tying." Moments up to $n = 4, 5$ are tracked with the fixed probes by interpolating in time between two trolley runs, and higher-order moments are interpolated assuming linear time dependence. The limitation of this interpolation results in "tracking errors" that are estimated from the difference between the moments predicted by the fixed probes and the moments actually measured by the subsequent trolley run. Studies with different intervals between trolley runs and at different times after the magnet was ramped to the nominal operating field were used to reduce the tracking errors and uncertainties.

The muon-weighted field is

$$\tilde{\omega}'_p = \frac{\int \omega'_p(x,y,\phi,t) M(x,y,\phi,t) \mathrm{d}x\mathrm{d}y\mathrm{d}\phi\mathrm{d}t}{\int M(x,y,\phi,t)\mathrm{d}x\mathrm{d}y\mathrm{d}\phi\mathrm{d}t}, \quad (40)$$

with the muon distribution $M(x,y,\phi,t)$ determined by a combination of measurements with the trackers and modeling of beam dynamics (Sec. VI F 1). Expanding $M(x,y,\phi,t)$ in the basis introduced in Eq. (38), the muon weighted azimuth- and time-dependent magnetic field is

$$\tilde{\omega}'_p(\phi,t) = \sum_i m_i(\phi,t) k_i(\phi,t), \quad (41)$$

where

$$k_i(\phi,t) = \frac{\int M(x,y,\phi,t) f_i(x,y) dxdy}{\int M(x,y,\phi,t) dxdy}. \quad (42)$$

The time-dependent azimuthally averaged field is

$$\tilde{\omega}'_p(t) = \frac{1}{2\pi}\int_0^{2\pi} \tilde{\omega}'_p(\phi,t)\mathrm{d}\phi, \quad (43)$$

which is weighted by the number of detected muon decays and time averaged over few day intervals.

B. Absolute calibration with a high-purity water probe

Each trolley probe reading is corrected for the field perturbations caused by the trolley components to the NMR frequency expected from a bare spherical water sample at 34.7 °C. This is done using an H_2O absolute calibration probe installed in the $g-2$ storage ring. The calibration probe for Run-2 and Run-3 was similar to that described in detail in [7,16].

Corrections must be applied to the measured calibration probe NMR frequencies to those expected from a bare spherical water sample at T_r. Corrections to the measured calibration frequency are listed in Table XIII and described below. These corrections were cross-checked with respect to a ^{3}He magnetometer in a dedicated high uniform 1.45 T solenoid and with simulations. All corrections are expressed as fractions of the measured NMR frequency, i.e., $\omega^{\mathrm{corr}} = \omega^{\mathrm{meas}}(1+\delta)$, where ω^{corr} is the frequency corrected for the effect δ. For corrections $\ll 1$ (the largest is 1.5 ppm), the combination of two corrections is $(1+\delta_a)(1+\delta_b) \approx (1+\delta_a+\delta_b+\mathcal{O}(\delta^2))$; only the first-order corrections are applied.

TABLE XIII. Calibration probe intrinsic corrections and uncertainties. Shape corrections are temperature dependent and hence different for each trolley probe. Thus, the range of all probes is given.

Description		Correction (ppb)	Uncertainty (ppb)
Shape, susceptibility	δ^b	−1508.7 to −1507.4	6.0
Material effects	δ^s	10.3	5.0
Radiation damping	δ^{RD}	0	3.0
Proton dipolar field	δ^d	0	2.5
Sample purity	δ^P	0	2.0
Subtotal			8.9

Sample-shape correction $\delta^{\mathbf{b}}$. The calibration probe consists of a cylindrical sample filled with high-purity water. The temperature-dependent correction to a spherical sample is

$$\delta^b(T_n) = \chi(T_n)(\epsilon - 1/3), \quad (44)$$

where $\chi(T_n)$ is the susceptibility at the temperature of the calibration probe for calibration of probe n, and $\epsilon = 0.4999(0,-0.0003)$ for the finite cylindrical sample, which was calculated in closed form from [46] and confirmed by numerical simulation ($\epsilon = 1/2$ for an infinite cylinder).

The temperature-dependent volume susceptibility is

$$\chi_V(T) = \chi_V(22\,^\circ\mathrm{C}) \times \left[\frac{\chi_m(T)}{\chi_m(22\,^\circ\mathrm{C})}\right] \times \left[\frac{\rho(T)}{\rho(22\,^\circ\mathrm{C})}\right], \quad (45)$$

where $\chi_V(22\,^\circ\mathrm{C}) = -9.056\times10^{-6}$ is the value recommended by CODATA [47] with 3×10^{-8} uncertainty due to additional measurements at unspecified temperatures [48]. We use the ratio of mass susceptibilities from [49]:

$$\begin{aligned}\frac{\chi_m(T)}{\chi_m(22\,^\circ\mathrm{C})} &= \frac{\chi_m(T)}{\chi_m(20\,^\circ\mathrm{C})}\frac{\chi_m(20\,^\circ\mathrm{C})}{\chi_m(22\,^\circ\mathrm{C})} \\ &\approx 1 + 1.3881\times(T-22\,^\circ\mathrm{C})\frac{10^{-4}}{^\circ\mathrm{C}} \\ &\quad + \mathcal{O}\left(\left((T-20\,^\circ\mathrm{C})\frac{10^{-4}}{^\circ\mathrm{C}}\right)^2\right). \end{aligned} \quad (46)$$

The temperature-dependent density $\rho(T)$ from [50] was used, because that is what was used in the analysis by [49].

Material effects $\delta^{\mathbf{s}}$. The calibration probe consists of the sample contained in a glass cylinder NMR sample tube, a concentric glass cylinder holding the NMR coil wires, a concentric aluminum cylinder shell, end caps, the temperature sensor, tuning capacitors, connectors, and mounting fixtures.

Due to their finite magnetic susceptibility, each of these components becomes magnetized by the external 1.45 T field, and the resulting magnetization contributes to the field measured by the probe. The contribution depends on the orientation (roll and pitch) of the probe with respect to the vertical magnetic field. The approximate cylindrical symmetry of the probe construction mitigates these effects, and a combination of direct measurements of intrinsic-probe effects δ^s, and simulations specific to the configuration in the $g-2$ storage ring are used to determine the remaining material corrections. Additionally, the high-permeability pole pieces of the storage-ring magnet act as magnetic mirrors that create images of the magnetized

calibration-probe components, leading to a correction $\delta^{s,\mathrm{img}}$ that depends on the probe position.

Sample (im)purity $\delta^{\mathbf{P}}$. Potential impurities, in particular, dissolved paramagnetic O_2 and salts, in the water sample could lead to a shift of the NMR frequency. Degassed ultrapure (ASTM type-1) water from several vendors was used, with no observed variation within an uncertainty of 2 ppb. A variety of additional tests were performed in which the glass water sample tube was rotated, and different sample tubes were used. No systematic shifts were observed.

Magnetization dependent effects $\delta^{\mathbf{RD}}$ *and* $\delta^{\mathbf{d}}$. The sample magnetization $\vec{M} = \chi_{\mathrm{H_2O}}\vec{B}$ can lead to two shifts. Radiation damping is the result of the oscillating current in the NMR coil that rotates the magnetization toward the external magnetic field. This leads to a time-dependent precession frequency shift δ^{RD} that depends on the magnetization along the magnetic field, the detuning of the NMR coil, and the coupling between the coil and the precessing spins (filling factor) [51]. A second, shape-dependent frequency shift is caused by the dipolar field from the precessing protons, δ^{d}. Both effects are estimated as in Run-1 [7].

Calibration probe temperature dependence $\delta^{\mathbf{T}}$. The gyromagnetic ratio of protons diamagnetically shielded in a spherical sample of water was measured at 34.7 °C [45]. This diamagnetic shielding is temperature dependent [52]. The correction from T_n^{cp}, the calibration-probe temperature for calibration of trolley probe n, to T_r, is $\delta_n^T = (-10.36 \pm 0.30) \times \frac{10^{-9}}{°\mathrm{C}}(T_r - T_n^{\mathrm{cp}})$. The calibration probe temperature was measured with a platinum resistive temperature device (PT1000 RTD) with an accuracy of 0.5 °C, and a different correction per probe was applied to account for the calibration-probe and trolley temperature during the calibration of each probe as discussed in the next section.

Corrections dependent on the calibration-probe environment. As noted in the discussion of material effects, the magnetized components of the calibration probe contribute to the measured magnitude of the magnetic field that depends on the orientation with respect to $\vec{B}$ and due to magnetic images. Additional corrections for the calibration configuration vary with the individual trolley probe being calibrated and are discussed in Sec. VI C.

1. Calibration-probe cross-checks

Work is underway to cross-check the intrinsic corrections applied to the calibration probe, i.e., corrections not dependent on the environment (δ^b, δ^s, δ^{RD}, δ^d, and δ^P), using ^{3}He magnetometry and a separate H_2O probe based on continuous wave (CW) NMR. The Mark-I ^{3}He absolute magnetometer provided an indirect 42 ppb cross-check on the calibration probe [7,16,53]. A Mark-II ^{3}He probe was designed and constructed with much smaller intrinsic corrections, and a campaign is underway to directly calibrate the muon $g-2$ calibration probes for Run-1 and runs 3–6. Preliminary analysis confirms agreement with uncertainties less than 20 ppb. The calibration probes were also compared to the CW H_2O NMR probe under development for JPARC's MuSEUM and $g-2$/EDM (E34) experiments [54]. Cross-checks with earlier CW prototypes at 1.4 and 1.7T showed a tension on the ∼50 ppb level with a precision around 15 ppb. The same cross-check, with newer probe versions, performed at 3 T, is in good agreement with an uncertainty of 10 ppb. The discrepancy with the earlier version is not yet understood; additional work is ongoing.

C. Trolley-probe calibration

Trolley-probe calibration provides a set of corrections to the frequencies ω_n^{tr} measured by each trolley probe

$$\omega_n' = \omega_n^{\mathrm{tr}}(1 + \delta_n^{\mathrm{calib}}), \tag{47}$$

where ω_n' is the field that would be measured by a spherical water sample at $T_r = 34.7$ °C at the position of probe n. Corrections for the temperature dependence of the vaseline-filled trolley probes are discussed in Sec. VI D.

Calibration campaigns before the start of Run-2 and after Run-3 took place in vacuum in a dedicated region of the storage ring magnet using the calibration probe described in Sec. VI B. Magnetic field gradients applied in all three directions were used to place the effective volumes of the calibration probe and each trolley probe within 0.5 mm of the same position, and the magnetic field in the calibration region was carefully mapped and shimmed.

The calibration correction was determined from a sequence of measurements swapping the trolley and calibration probe into the calibration position. During this swapping, the magnetic field was tracked with fixed probes to mitigate the effect of drifts. Additionally, the Run-2/3 calibration campaigns and the Run-1 calibration campaign provided data on the stability of the trolley-probe calibrations over a three-year period.

Uncertainties from the calibration procedure are listed in Table XIV. These include uncertainties due to misalignment of the calibration probe and trolley probe,

TABLE XIV. Uncertainties from the calibration procedure on the muon-weighted field. The uncertainties for the individual probes are shown in Table XXVIII. The probe individual corrections due to temperature dependence of the diamagnetic shielding range from −126.3 to −59.1 ppb.

Description		Uncertainty (ppb)
Swapping and misalignment	δ^{tr}	1.6
Temperature of diamag. shielding	δ^T	5.2
Variance	δ^{var}	11.0
Active volume	δ^{av}	1.7
Footprint trolley	δ^{fp}	8.0
Footprint CP	δ^{cp}	4.0
Frequency extraction CP	$\delta^{\mathrm{freq}(cp)}$	1.0
Material and mag. image	δ^{img}	9.0
Subtotal		17.8

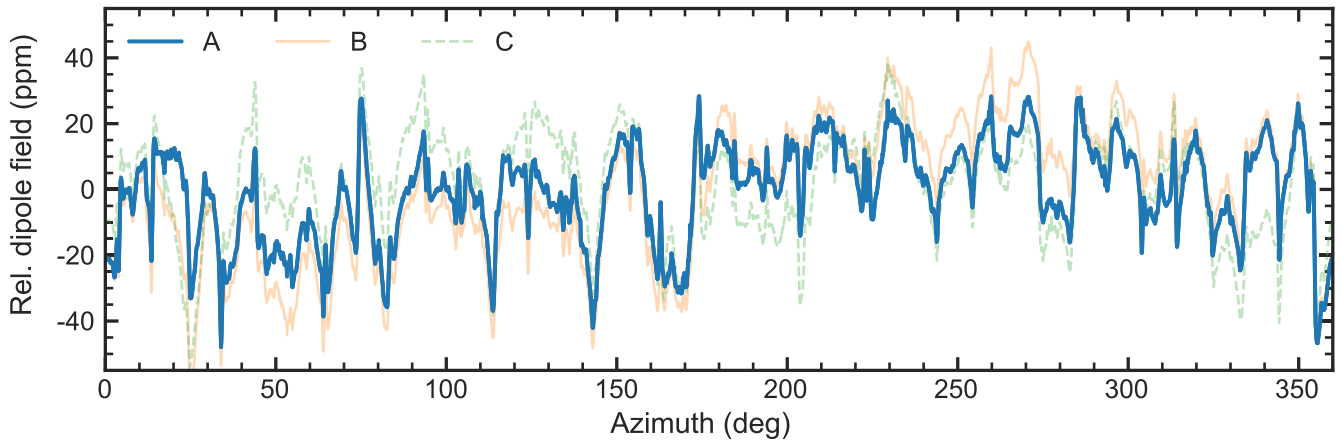

FIG. 22. The relative (Rel.) dipole m_1 coefficient as a function of azimuth for three field maps with respect to its azimuthal average. (A) is from April 8, 2019, the beginning of Run-2, (B) is from June 20, 2019, the end of Run-2, and (C) from March 11, 2020, the end of Run-3. The peak-to-peak amplitudes are 76, 108, and 93 ppm, respectively, with rms of 14.6, 20.5, and 15.8 ppm.

temperature corrections of the diamagnetic shielding δ^T, the variance between the calibration constants of different measurement campaigns and analyzer δ^{var}, the difference between the active volume of the calibration probe and trolley probe δ^{av}, the influence of the trolley and calibration probe's materials on the magnetic environment of the other, called magnetic footprint δ^{fp} and δ^{cp}, the frequency extraction δ^f and the material effects including the magnetic image in the pole pieces δ^{img}. The per-probe calibration constants with a graphical representation is given in Table XXVIII and Fig. 29, in the Appendix.

D. Magnetic field maps

In this section, we describe the detailed extraction of the field maps $\omega_p'^{\text{tr}}(x, y, \phi, t)$ [Eq. (38)]. The transverse positions are fixed by the probe locations, while the trolley position is radially constrained by the trolley rails. The trolley azimuthal position is determined by reading the barcodes etched into the bottom of the vacuum chambers. Encoders that measure the length of the trolley cables are a backup, however, the encoder precision is inferior compared to the barcode due to tension variations in the cables. The 17 trolley NMR probes are triggered in sequence every ~30 ms, resulting in a ~2 Hz sampling rate for each probe. The corrected frequencies are interpolated to a grid of azimuthal positions $\phi_k(t)$. Different interpolation schemes were tested and agreed within 1 ppb.

The multipole coefficients $m_i(\phi_k(t_{\text{tr}}))$ are determined for each ϕ_k by fitting the corrected frequencies to Eq. (39), where t_{tr} is the time when the trolley is at ϕ_k. A lower bound on N_{max} is derived from azimuthal averaged fit residuals, which show a transverse dependence if N_{max} is chosen too small. An upper bound comes from degeneracies of the multipoles with our trolley probe configuration. The truncation at $N_{\text{max}} = 12$ of the parametrization in Eq. (41) is used. The difference between using different minimization algorithms to extract the multipole coefficients is negligible. Representative field maps $m_1(\phi)$ for three different trolley runs are shown in Fig. 22.

Corrections and uncertainties to the trolley multipole coefficients are presented in Table XV and summarized here.

Trolley motion effects (δ^{motion}). The trolley motion in a nonuniform magnetic field generates eddy currents in the conducting components, most significantly the aluminum shell. We use the Run-1 correction for $\delta^{\text{motion}} = (-15 \pm 18)$ ppb from Run-1 analysis [7] estimated from the comparison of standard continuous motion trolley runs with stop-and-go runs and from the comparison for clockwise and counterclockwise trolley runs.

Difference in configuration (δ^{config}). During the trolley runs, the collimators that radially constrain the stored-muon distribution are retracted, and the trolley rails are in a different position than when the muons are stored.

TABLE XV. Corrections and uncertainties from the spatial field maps. A single value per line indicates the same value for all datasets.

			Uncertainty (ppb)		
Description		Corrections (ppb)	Run-2	Run-3a	Run-3b
Motion effects	δ^{motion}	−15.0		18.0	
Configuration	δ^{config}	−7.0		22.0	
Frequency extraction	$\delta^{\text{freq(tr)}}$	⋯	19	18	16
Temperature	δ^{temp}	⋯	9.2	13.8	15.2
Transverse position	δ^{xy}	⋯	10.0	9.9	9.0
Azimuthal position	δ^{azi}	⋯		4.0	
Parametrization	δ^{param}	⋯	3.4	6.3	7.6
Azimuthal averaging	δ^{avg}	⋯	0.8	1.4	1.7
Subtotal			37.2	38.5	38.1

The effect of these two configuration changes is estimated from calculations of the magnetic field produced by the diamagnetic copper and paramagnetic aluminum in the respective configurations. The uncertainty of the Run-1 correction of $\delta^{\text{config}} = (-7 \pm 22)$ ppb [55] is dominated by a discrepancy in the calculation and what a local fixed probe measures. The same value is used for Run-2/3. The effect from the collimators on the azimuthally averaged field is smaller than 1 ppb.

Trolley frequency extraction ($\delta^{\mathbf{freq}}$). Trolley NMR-probe FID analysis is described in [56]. Briefly, the phase function (phase vs time) for the free-induction-decay (FID) signals is extracted from in-phase and Hilbert-transform quadrature signals. The phase functions are fit to polynomials of varying order from two to six and for a varying time ranging from 0.20 to 0.75 of T_2^* (the FIDs are not exponential, so in this case, we refer to the time for the FID amplitude to reach $1/e$ of the maximum). The frequency-extraction correction $\delta^{\mathbf{freq}}$ on m_1 is below 12 ppb. Potential effects from incorrect $t_{\text{FID}} = 0$ on the 100 μs level are shown to be negligible. Temperature changes affect the phase function of FIDs. This effect on the extracted precession frequencies is included in the correction below. The uncertainty due to correcting from the ≈25 °C trolley temperature during field mapping to around ≈33 °C temperature during calibration is 5 ppb.

The total uncertainty from the frequency extraction, taking the Run-2/3 beam shapes and correlations between the multipoles into account, is shown in Table XV. In Run-1, this correction had a different meaning because every trolley NMR position was treated as an independent point with frequency extraction uncertainty of 10 ppb. In fact the NMR sample active volume is ∼1.8 cm, while the measurements are separated by ∼0.5 cm leading to oversampling.

Trolley temperature dependence ($\delta^{\mathbf{temp}}$). A dedicated study in the Argonne National Laboratory magnet facility with two temperature-controlled probes to track magnet drifts revealed a temperature dependence of the vaseline frequency of (-0.8 ± 0.2) ppb/C. However, a conservative uncertainty of 2 ppb/°C is used, since the uncertainty is dominated by the frequency extraction uncertainties discussed above.

The trolley-probe NMR frequencies are not actively temperature corrected, rather, we apply a correction and uncertainty $\delta^{\text{tr,temp}}$. The temperature difference of the trolley probes with respect to the mean temperature during the calibration (33.1 °C) range from −8.0 °C to −1.9 °C. The temperature-dependent frequency correction is calculated using the temperature dependence of (-0.8 ± 20) ppb/C. The muon weighted corrections for the three datasets are −3.6, −5.5, and −6.0 ppb, respectively. In addition, the temperature spread during one field map is (1.8 ± 0.3) °C and an uncertainty of 1 °C on the temperature sensor is used. The resulting uncertainties for Run-2, Run-3a, and Run-3b are listed in Table XV.

Trolley transverse and azimuthal position ($\delta^{\mathbf{xy}}$, $\delta^{\mathbf{azi}}$). The trolley position is constrained in the transverse plane by the rails. A laser tracker was used to estimate rail distortions before the vacuum chambers were installed. The effect in the transverse plane δ^{xy} is evaluated by taking the Run-2 and Run-3 beam shapes into account by running one of the analysis chains with and without incorporating rail distortions. The observed difference of 11.8 ppb (Run-2), 4.1 ppb (Run-3a), and 1.8 ppb (Run-3b) are used to correct the other analysis. The corresponding uncertainties are listed in Table XV. For Run-2/3 the corrections are smaller than for Run-1 due to the smaller higher-order multipole moments.

The azimuthal trolley position is determined using the barcode except for small gaps between adjacent vacuum chambers and for barcode errors, where cable-length encoders are used. A conservative estimate of the azimuthal position resolution of 2 mm leads to a systematic uncertainty of $\delta^{\text{azi}} = 4$ ppb on the average dipole field.

Parametrization ($\delta^{\mathbf{param}}$) *and azimuthal averaging* ($\delta^{\mathbf{avg}}$). The finite number of measurements and the parametrization of Eq. (38) lead to additional uncertainty with three contributions: 1. an uncertainty due to the truncation N_{max} in Eq. (38), 2. uncertainty due to interpolation between the finite number of azimuthal slices, and 3. the use of 2D multipole expansion, which is only valid if there is no azimuthal magnetic field dependence. The uncertainty due to the choice of N_{max} is estimated from the residuals of the fits to Eq. (38) weighted by the azimuthally averaged beam distribution within $\Delta l = 1$–10 mm around each probe.

The uncertainty due to the interpolation between these finite azimuthal slices was determined by interpolating with linear, quadratic, and cubic splines. To estimate the effect of 2D multipole expansion, the averaged magnetic fields following the above analysis approach were compared to an analytic azimuthal average using simulated magnetic fields based on a toroidal 3D multipole-based field description. The observed differences from such comparisons are <1 ppb.

E. Magnetic field tracking

The fixed probes track the magnetic field between trolley runs (see Sec. VI D) for moments up to $i = 5$. For higher-order moments, we use linear interpolation in time. Fixed-probe tracking entails the following steps: (1) extracting fixed-probe moments defined in Eq. (38); (2) tying the fixed-probe moments to the trolley-map moments; (3) parametrizing the moments as a function of azimuth and time.

1. Fixed probe moments

Linear combinations of measurements from the four or six fixed probes at each station provide fixed-probe moments $m_i^{\text{fp}}(\phi_s, t)$ following the procedure described in [7].

To reduce the effect of probe noise, the $m_5^{\mathrm{fp}}(\phi_s, t)$ moment is first tied to the measured m_5 from the trolley run pair (see Sec. VI E 2) before the change of moment basis.

Fixed probes in three stations close to the inflector experience large gradients resulting in very short FIDs and increased frequency uncertainty (noise). Two additional probes with a PEEK housing are installed inside the vacuum chamber at the position of one of the stations. These additional measurements verified that linear interpolation of the moments from neighboring stations gives a better estimate than the determination from the noisy fixed probe frequencies. Therefore, the multipole moments for these three stations are linear interpolations from their neighboring stations.

The relative fixed probe frequency extraction is very robust and the uncertainty from the fixed probe frequency extraction $\delta^{\mathrm{freq(fp)}}$ is ∼1 ppb, consistent with Run-1 [7]. Nonlinear temperature changes of the yoke and thus the fixed probes are on the 0.06 °C level, and thus the uncertainty due to fixed probe temperature is negligible. Linear components are canceled by tracking between two subsequent field maps.

Fixed probe data are subject to general data quality cuts (Sec. III A). Additionally, events with FID amplitudes or FID power more than seven standard deviations from the probe's mean amplitude and power are removed.

2. Tying fixed probe to trolley-map moments

The change of the magnetic field at a fixed-probe station before or after t_s^{tr}, the time the trolley passes the station at ϕ_s during a trolley run, is

$$\Delta m_i^{\mathrm{fp}}(\phi_s, t) = m_i^{\mathrm{fp}}(\phi_s, t) - m_i^{\mathrm{fp}}(\phi_s, t_s^{\mathrm{tr}}), \tag{48}$$

where $m_i^{\mathrm{fp}}(\phi_s, t_s^{\mathrm{tr}})$ is the moment measured using the fixed probes within station s averaged around the time the trolley passes by that station.

To determine t_s^{tr}, we make use of the fact that the material effects of the trolley and its onboard electronics produce a characteristic field perturbation (footprint) that is measured by the fixed probes when the trolley passes. The time of the largest field perturbation sets t_s^{tr} and the trolley's azimuthal location sets ϕ_s. Varying the station positions ϕ_s by ∼0.25 deg has an effect less than 1 ppb.

The field perturbation due to the trolley when passing a fixed probe station is removed from the fixed-probe data and replaced with a linear interpolation of $m_i(\phi_s, t)^{\mathrm{fp}}$ based on the 30 s before and after t_s^{tr}. The effect of the trolley footprint replacement is tested on data in regions without footprint by comparing the field estimated by the replacement algorithm and the actual measured data. The uncertainty is listed in Table XVI and is similar to Run-1, as described in [7].

TABLE XVI. Corrections and uncertainties (in parenthesis) from magnetic field tracking. A single value per line indicates the same value for all datasets. All values are given in units of ppb.

	Correction (Uncertainty)		
Description	Run-2	Run-3a	Run-3b
Tying			
Trolley footprint		(7.0)	
Fixed probe resolution		(1.0)	
Tracking			
Brownian bridge	(15.4)	(10.7)	(16.0)
Magnet ramp effect	−3.0 (3.0)	−10.0 (10.0)	−3.0 (3.0)
Fixed probe temperature	(0)	(0)	(0)
Analysis choices	(1.8)	(2.5)	(1.5)
Subtotal	(17.3)	(16.5)	(17.8)

3. Fixed-probe tracking

For azimuth ϕ and time t for one or more trolley runs at t_k the fixed-probe tracked moments are

$$m_i(\phi, t) = \sum_k W_k(t) \Bigg(m_i^{\mathrm{tr}}(\phi, t_k) + \sum_s W_s(\phi) \sum_j J_{ij}(\phi_s) \Delta m_j^{\mathrm{fp}}(\phi_s, t) \Bigg), \tag{49}$$

where k labels the trolley runs, and $W_k(t)$ is the weighting of each trolley run at time t. The azimuthal weighting factor $W_s(\phi)$ interpolates between stations on either side of ϕ, $J_{ij}(\phi_s) = \frac{\partial m_i^{\mathrm{tr}}(\phi_s)}{\partial m_j^{\mathrm{fp}}(\phi_s)}$ is the Jacobian that relates small changes of the fixed probe moments to changes of the trolley moments for station s, and $\Delta m_i^{\mathrm{fp}}(\phi_s, t)$ is defined in Eq. (48).

Ideally, magnetic field tracking uses two consecutive trolley runs, *e.g.* $k = 1, 2; 2, 3$ *etc*. Occasional unplanned magnet incidents, such as the loss of magnet power allow tracking only from the trolley run before the incident, in which case $W_k(t) = 1$.

Field changes not tracked by the fixed probes lead to errors of the $m_i(\phi, t)$ that is a maximum at the midpoint between the two paired trolley runs. To quantify this, tracking from a single trolley run is used to predict the field moments at the later trolley run. The difference between the predicted and measured field moments for the second trolley run is called the *tracking offset*. The tracking offset can be modeled as a random walk process caused by changes in the magnet shape. For tracking using a pair of consecutive trolley runs, the random walk becomes a Brownian bridge that uses a linear interpolation between the first and second trolley run (see Ref. [7] for details). A single parameter M parametrizes the rate of the process.

The distribution of the azimuthally averaged tracking offsets can be used to account for potential correlations between different stations. In order to reduce the statistical error, the random-walk parameters are determined from the azimuthally averaged tracking offsets for all of Run-2/3. We determine $M = 0.018\ \mathrm{Hz}/\sqrt{\mathrm{s}}$ for the m_1 coefficient. Similar rate of change parameters are determined for each multipole moment. The resulting uncertainties, taking the muon-weighted corrections for the different datasets and the correlations between the different multipole moments into account, are summarized in Table XVI. Note that this uncertainty is statistically independent and hence reduces if multiple datasets are combined.

We observe that the tracking offset depends on the time after the magnet was ramped up and shows a characteristic azimuthal dependence that is largest at magnet yoke boundaries as shown in see Fig. 23. A dedicated measurement was performed, repeatedly measuring the field with the trolley for 60 h after the magnet was ramped. We use the azimuthally averaged tracking offset to estimate the bias.

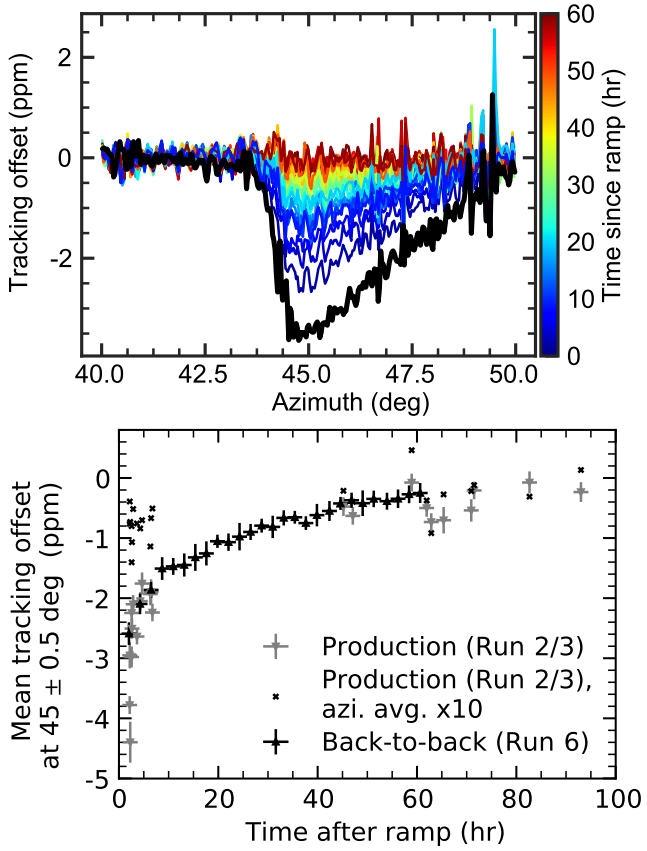

FIG. 23. Top: tracking offset (inability to track field) as a function of azimuth (azi.) around a yoke boundary. Different colors indicate different times after the magnet ramp. Bottom: amplitude of effect at 45° as a function of time after magnet ramp. The "x"s show the azimuthally averaged values scaled up by a factor of x10. A dedicated campaign of back-to-back trolley runs was performed in Run-6 to study this effect.

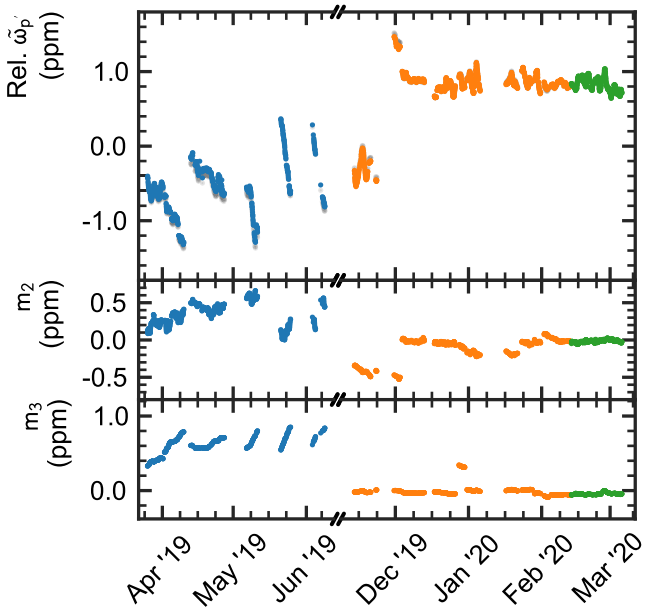

FIG. 24. The relative muon-weighted magnetic field ($\tilde{\omega}_{p'}$) as a function of time for the Run-2 (left side) and Run-3a and Run-3b (right side). The dipole m_1 contribution alone is shown in gray below. On this scale, they barely differ. The lower two plots show the tracked m_2 and m_3 moments.

We model the effect by an exponential function with amplitude and time constants as parameters. The amplitude and time constant may depend on the history of the magnet before the ramp. Therefore, we determine a correction and uncertainty conservatively; the result is an initial amplitude of (100 ± 100) ppb and a relaxation time constant of 12 h. The correction and uncertainty depend on the time periods relative to the magnet ramp time in which muon data have been taken. The resulting correction and uncertainties are listed in Table XVI.

TABLE XVII. Field multipole moments in ppb [see Eq. (38)] averaged over azimuth and time (including DQC) per dataset. The Run-3 experiment hall temperature was more stable than Run-2 due to a climate-control upgrade.

Multipole	Run-2	Run-3a	Run-3b
m_2/m_1	331	−113	−14
m_3/m_1	611	−6	−43
m_4/m_1	−310	23	17
m_5/m_1	383	40	35
m_6/m_1	94	−9	−20
m_7/m_1	217	127	127
m_8/m_1	−24	−22	−21
m_9/m_1	23	15	12
m_{10}/m_1	−697	−725	−727
m_{11}/m_1	−167	−203	−215
m_{12}/m_1	−1068	−1056	−1057

A detailed comparison between interpolation analyses from two groups was performed to identify inconsistencies and bugs in the analysis, while the individual groups had individual software blinds. Comparisons performed on the azimuthal averaged field and on a station-by-station basis agree within a few ppb after relative unblinding. The difference in analysis results due to different analysis choices is added as additional uncertainty and listed in Table XVI.

The multipole moments averaged over azimuth and weighted by the detected muons (including DQC) $\langle m_i \rangle_{\phi,t}$ are listed in Table XVII for all three datasets. The lowest order $\langle m_i \rangle_{\phi,t}$, the normal and skew quadrupoles, is shown as a function of time over the full dataset in Fig. 24.

F. Muon weighted magnetic field

1. Muon beam distribution

The muon beam distribution $M(x, y, \phi)$ is reconstructed from measured positron tracker profiles combined with beam-dynamics calculations of the azimuthal dependence of the muon distribution around the ring. Two trackers provide well-localized muon beam distributions with an azimuthal sensitivity with an rms of 4.9° and 4.8°, respectively. Following Eq. (40), the mapped magnetic field is weighted by the muon distribution to determine the magnetic field seen by the muons.

Tracker profiles $M_i^T(x, y)$ for the muon-weighted magnetic field are accumulated in time intervals of $T_{\text{interval}} =$ 2 h to 3 h and corrected for detector resolution and acceptance. Only positrons with decay times between the analysis start time $t_{\text{start}} = 30.2876$ μs and end time $t_{\text{end}} =$ 650.0644 μs enter the tracker profiles. The time intervals T_{interval} are chosen to contain more than 6×10^5 total tracks, avoid gaps > 6 h, stay within a trolley-run pair and contain entire ω_a DAQ runs.

The measured beam profiles at azimuthal locations where the tracker detectors do not provide beam diagnostics are reconstructed from tracker profiles by shifting the mean and scaling the transverse widths of the distribution relative to the tracker station using

$$\langle x \rangle(\phi) = x_{\text{COD}}(\phi) + D_x(\phi)\langle \delta \rangle, \tag{50}$$

$$\langle y \rangle(\phi) = 0, \tag{51}$$

$$x_{\text{rms}}(\phi) = \left[\frac{\beta_x(\phi)}{\beta_x(\phi_{\text{tkr}})} (x_{\text{rms}}^2(\phi_{\text{tkr}}) - D_x^2(\phi_{\text{tkr}})\delta_{\text{rms}}^2) + D_x^2(\phi)\delta_{\text{rms}}^2 \right]^{1/2}, \tag{52}$$

$$y_{\text{rms}}(\phi) = \left[\frac{\beta_y(\phi)}{\beta_y(\phi_{\text{tkr}})} y_{\text{rms}}^2(\phi_{\text{tkr}}) \right]^{1/2}. \tag{53}$$

The beam widths x_{rms} and y_{rms} at the azimuth of the tracker stations ϕ_{trk} are extracted from the tracker profiles $M_i^T(x, y)$. The beta functions $\beta_x(\phi)$, $\beta_y(\phi)$, and radial dispersion function $D_x(\phi)$ are determined from the optical lattice calculated with the COSY INFINITY-based model of the storage ring. The mean and rms fractional momentum $\langle \delta \rangle$ and δ_{rms} are extracted from the fast-rotation analysis discussed in Sec. V A. The average fractional momentum is ∼0.07% except for Run-3b, which is lower (∼0.01%) owing to stronger injection kickers, whereas the rms of the distribution is ∼0.1%. The field indices are listed in Table I.

Closed orbit distortions (COD) shift the ideally circular closed orbit away from the equilibrium position. Azimuthal variation in the vertical dipole component of the magnetic field causes a radial COD

$$x_{\text{COD}}(\phi) \approx \frac{R_0}{n} \frac{b_1(m_1)}{B_0} \cos(\phi - \phi_1(m_1)), \tag{54}$$

where R_0 is the nominal radius, B_0 is the nominal field, n is the effective field index given in Table I, and $b_1(m_1)$ and $\phi_1(m_1)$ are the $N = 1$ Fourier amplitude and phase of $m_1(\phi)$. The Fourier components are extracted with a discrete Fourier transform from field maps in each T_{interval}, and x_{COD} is calculated for each individual T_{interval}. The amplitudes of the radial COD range from 0.6–1.5 mm and 0.2–0.4 mm for Run-2 and Run-3, respectively.

An azimuthally varying radial magnetic field would cause a vertical COD. Because the radial field dependence on azimuth is not measured during the experiment, y_{COD} is set to zero and considered separately as a systematic. Misalignments of the electric quadrupole plates also cause radial and vertical CODs by steering the beam. These are considered separately as a systematic.

Each tracker station is extrapolated separately, and the reconstructed distributions from both stations are averaged to get the nominal beam distribution.

Figure 4 in Sec. III C 3 illustrates azimuthally averaged muon beam distributions based on the beam extrapolation around the ring of tracker measurements.

2. Muon weighting

Following Eq. (40), the reconstructed muon beam distribution $M(x, y, \phi, t)$ (see Sec. VI F 1) is projected onto the moments used to describe the magnetic field for time intervals T_{interval} and evaluated every 5° because the azimuthal variation of the beam moments is small. Since the tracker profiles and thus the beam moments are only determined every 2–3 h, the field moments $m_i(t, \phi)$ are averaged in time, weighted by the number of muons in the storage region $N_\mu(t)$. Equation (41) is used to calculate the muon-weighted field per T_{interval} and azimuthal bin ϕ_i. Additional averaging over all azimuthal bins and thus implementing Eq. (43) yields the muon-weighed field per time interval T_{interval}. Averaging all time intervals within a dataset, weighting by $N_\mu(t)$ and accounting for DQC cuts, yields the muon weighted magnetic field $\tilde{\omega}_p'$ per

TABLE XVIII. Average beam multipole projections in each dataset, including DQC. Projections are normalized to beam profile intensity and are unitless.

Beam projection	Run-2	Run-3a	Run-3b
k_1	1.000	1.000	1.000
k_2	0.139	0.136	0.073
k_3	−0.001	−0.006	−0.005
k_4	0.001	−0.001	0.000
k_5	0.081	0.076	0.046
k_6	0.000	−0.001	0.000
k_7	−0.001	−0.001	−0.006
k_8	−0.002	−0.001	0.003
k_9	0.001	0.001	0.000
k_{10}	−0.004	−0.003	0.001
k_{11}	0.000	0.000	0.000
k_{12}	−0.001	−0.001	0.001

dataset defined in Eq. (40), listed in Table XXVI for each dataset.

The improvement in the kick for dataset Run-3b reduces the k_2 and k_5 parameters [see Eq. (42)] since the muon distribution is more centered. This has the effect that weighted moments $m_i, i > 1$ are reduced, and thus systematic uncertainties that only couple through moments with $m_i, i > 1$ are reduced as well. The beam multipole projections averaged over azimuth over the times when muons are stored to extract ω_a ($\langle k_i \rangle_{\phi,t}$) are listed in Table XVIII for all three datasets. Figure 24 provides an overview of the muon-weighted field as a function of time.

3. Systematics

Tracker-specific systematics cause uncertainties in the beam distribution, which lead to uncertainties in $\tilde{\omega}'_p$. The relevant uncertainties for muon weighting are tracker resolution $\delta^{\text{reso,tkr}}$, acceptance $\delta^{\text{accept,tkr}}$, and alignment $\delta^{\text{align,x,tkr}}$, $\delta^{\text{align,y,tkr}}$. These systematics are evaluated by varying each parameter by 1σ, producing corresponding beam distributions in the usual time intervals T_{interval} and evaluating the effect on $\tilde{\omega}'_p$ averaged over each dataset. The resulting uncertainties are listed in Table XIX.

The tracker acceptance uncertainty is ≤ 2 ppb from changing the acceptance function by $\pm 20\%$, and the resolution uncertainty is <1 ppb by changing the radial and vertical resolution by ± 0.5 mm. Changing the tracker alignment in x and y by ± 0.6 mm yields uncertainty on the size of 1 ppb. The uncertainty due to tracker profile statistics are insignificant.

The muon-weighted field should be calculated for muons that enter the ω_a determination and thus are seen by the calorimeters. Because the spatial acceptance from tracker and calorimeters is different, the muon distribution from the tracker would have to be corrected for calorimeter acceptance. However, the effect is small and thus is only treated as an uncertainty.

TABLE XIX. Corrections and uncertainties (in parenthesis) due to spatial muon weighting of the magnetic field.

	Correction (Uncertainty) (ppb)		
Description	Run-2	Run-3a	Run-3b
Detector effects			
Tracker acceptance	(2.1)	(1.1)	(0.1)
Tracker resolution	(0.1)	(0.1)	(0.1)
Tracker y alignment	(10.7)	(0.6)	(0.4)
Tracker x alignment	(4.5)	(1.3)	(0.3)
Calorimeter acceptance	(1.0)	(0.2)	(0.2)
Closed Orbit Distortion and azimuthal effects			
yCOD (radial B)	(1.8)	(3.7)	(2.9)
xCOD (quad misalig.)	+1.3 (5.9)	+2.7 (6.7)	+2.5 (6.3)
yCOD (quad misalig.)	−0.9 (0.1)	−0.5 (0.2)	−0.3 (0.2)
Mean momentum offset	(0.2)	(0)	(0)
Subtotal	(13.4)	(7.9)	(6.9)

As discussed above, an azimuthal radial magnetic field variation can contribute to y_{COD}. Since the radial magnetic field was only measured in pre-Run-1 while no vacuum chambers were installed, the effect is estimated by assuming an amplitude of 0.5 mm, which is a factor of 2 larger than the pre-Run-1 measured value, for the $N = 1$ COD and the worst case phase.

Misalignments of the electric quadrupole plates cause an x_{COD} or y_{COD} by steering the beam. The expected COD calculations use the central displacements of the electric quadrupole plates measured in a survey. Survey uncertainties cause uncertainties in the CODs. These effects were evaluated using the same method from Run-1 [7], resulting in a correction and uncertainty listed in Table XIX.

The momentum deviation δ used in the beam reconstruction procedure in Eqs. (50) and (52) slightly differ from different analyzing teams in Sec. IV. The related systematic uncertainty is determined by varying $\langle \delta \rangle$ and δ_{rms} by ± 0.0001.

A changing muon distribution over time in a fill can be caused by magnetic field transient effects from the electric quadrupoles and kicker eddy currents. Tracker profiles are reconstructed for different times in a fill. Studies show that the related uncertainties are negligible in Run-2/3.

G. Transient magnetic fields

The fixed probe system measures the magnetic field at intervals of 1.2–1.4 s asynchronous to beam injection. Thus, any time-dependent, μs-timescale magnetic field transient that is synchronized with beam injection is not accounted for in $\tilde{\omega}'_p$. In addition, the skin-depth effect in the aluminum of the vacuum chambers reduces the effects on high-frequency magnetic field transients. Transient magnetic fields synchronized with beam injection are caused by eddy currents in the kicker and time-varying fields caused by

the pulsing of ESQs. Both effects lead to corrections on the muon-weighted magnetic field and are improved compared to Run-1 by additional measurements. Additional transient effects related to magnetic fields in the booster are <7 ppb as determined for the Run-1 analysis [7].

1. Transient magnetic fields from kickers

The magnetic field kick of 22 mT to store muons on the stable orbit is a fast transient field (∼150 ns) that introduces eddy currents in the region of the kicker magnets that lasts longer than the initial kick. NMR magnetometers are too slow to measure the effect on the magnetic field. The transient magnetic field has been measured with two magnetometers based on Faraday rotation using terbium gallium garnet (TGG) crystals [7]. For Run-2/3, additional measurements with improved setups have been performed using the same magnetometers.

One of the magnetometers utilizes fibers to guide the light from the laser source, which is housed in the center of the storage ring magnet, to the 3D printed magnetometer where the laser light is polarized and sent through two 14.5-mm-long TGG crystals. A polarization-sensitive splitter divides the laser beam into two returning fibers. The two beam intensities are measured by PIN diodes; the polarization is reconstructed from the difference. This differential readout scheme reduced the sensitivity on laser instabilities. The magnetometer base consists of a glass block with small Sorbothane legs, lowering the magnetometer's center of mass and reducing mechanical vibrations.

The measurements in Run-1 [7] were limited by noise picked up from mechanical vibrations of the kicker cage through the magnetometer and the fibers themselves. To reduce the noise in the measurements, a PEEK bridge was machined with Sorbothane legs that allow the magnetometer to be anchored to the vacuum chamber instead of the cage that holds the kicker plates. In addition, the returning fibers are routed on top of silicon bands that dampen out potential vibrations.

Two measurement campaigns in summer 2021 and summer 2022 have been performed. To calibrate the magnetometer, the magnetic field of the main magnet was ramped up and down at a constant rate to 1.4513 T. The calibration constants change from ramp to ramp due to temperature changes affecting the Verdet constant of the TGG crystal and small tilt angles changing the effective length of the crystal.

Since the laser was operated in constant current mode, the calibration factor changed over time, which was tracked by measuring the 12 μT magnetic field transient from charging the kicker plates prior to the kick.

The measured transient field is shown in Fig. 25 for two measurement campaigns one year apart. The average of the two campaigns is used to estimate the effect of the measured field perturbations. The effect on ω_a is estimated by integrating the effect of the transient over the muon lifetime. A five-parameter fit is used to estimate the overall correction. The corrections are estimated based on measurements in the first of the three kickers with upgraded kicker cables and operated at nominal kicker setting of 53.1 kV as present during Run-3b. The results from this measurement are scaled to the other kickers, which operate at slightly different operation voltages (53.1, 53.0, and 55.0 kV), and to the conditions in Run-2 and Run-3a, during which the kickers were operated at lower voltages (47.7, 47.1, and 47.1 kV). Azimuthally, the kicker transient is treated as uniform within the regions occupied by the kicker plates. The steep fall-off at the edges was modeled and confirmed by measurements outside the kicker plates, resulting in a suppression for the azimuthal average of 0.085. Overall, this results in corrections to $\tilde{\omega}'_p$ of −21.1 and −22.5 ppb for Run-2/3a and Run-3b, respectively. The associated uncertainties are summarized in Table XX and described briefly below.

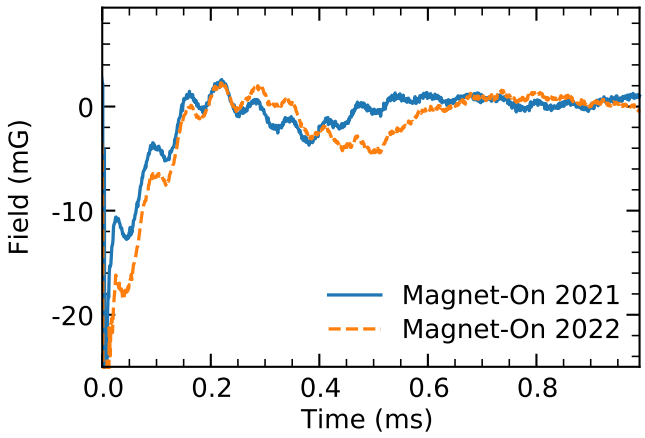

FIG. 25. Magnetic field transient induced by kicker magnets measured by the optical fiber magnetometer in summer 2021 and summer 2022.

TABLE XX. Uncertainties to $\tilde{\omega}'_p$ due to transient magnetic fields from eddy currents in the kicker system. The uncertainties from the two campaigns in 2021 and 2022 are combined for the Run-3b dataset. The values are scaled for the Run-2 and Run-3a datasets accounting for the different run conditions.

	Uncertainty (ppb)			
Description	2021	2022	Run-3b	Run-2/3a
Vibration ambiguity	8.3	12.8	10.5	9.9
Transient variance			4.2	3.9
Azimuthal	3.1	4.7	3.9	3.7
Transverse	4.4	6.8	5.6	5.3
Calibration	0.3	0.2	0.3	0.3
Baseline	2.5	0.2	1.3	1.2
Scaling				1.7
Pulse shape difference				4.2
Subtotal			13.3	13.3

The effect of residual vibrations in the measured signal is estimated by comparing results with the main magnet powered and not powered. The origin of the perturbations with a timescale of about 1 ms and amplitude of a few 0.1 μT remains ambiguous. The measurement data cannot distinguish between an actual change in the total magnetic field and mechanical vibrations of the fibers or the crystal. This ambiguity contributes to the leading systematic uncertainty on the transient measurement. The observed differences between the two campaigns is not fully understood and might indicate local variations of the effect. This ambiguity is accounted for by assigning the observed difference as a "transient variance" uncertainty. Further contributions to the uncertainty come from the azimuthal and transverse modeling, as well as from the above-mentioned calibration procedure and baseline determination. Like the total effect, the uncertainties are scaled to the different run conditions in the Run-2 and Run-3a datasets. The scaling and potential differences in pulse shapes due to using different cables lead to additional uncertainties for these datasets.

2. *Transient magnetic fields from ESQs*

The beam-synchronous pulsing of the ESQ plates causes time-dependent magnetic field changes on the μs timescale. These fast synchronous changes are not captured by the field maps nor tracked by the fixed probe system. Besides the asynchronous operations of the fixed probes with respect to beam injection times, skin depth effects in the aluminum walls of the vacuum chambers suppress field transients on that time scale. *In situ* measurements are required. While the exact mechanism creating this magnetic field transition is not fully understood, the effect is associated with the ESQ plates' and support structure's mechanical vibrations. The injection of muons and associated pulsing of the ESQ plates every 10 ms for 8 bunches drives an oscillation around 100 Hz, close to the system's intrinsic frequencies around 50 Hz. The bottom plot in Fig. 26 shows an example of this effect as a function of time at one fixed location. A second train of eight bunches is injected after 266.7 ms, a gap long enough for the vibration to mostly ring down. This pattern repeats every 1.4 or 1.2 s. Since this field changes during the time muons are stored and are not reflected in the direct measurement of $\tilde{\omega}_p'$, this transient results in a correction term B_Q.

In Run-1, the transient fields from ESQs were measured in a dedicated measurement campaign with a set of trolley NMR probes sealed inside plastic tubes for vacuum compatibility, held in place in the center of the storage volume on static legs sitting on the trolley rails.

The ESQs span 43.3% of the ring and are grouped into four stations, each consisting of a short and a long section. The azimuthal dependence was mapped coarsely for one such section. Significant differences in the oscillation pattern were observed as a function of azimuth. The long

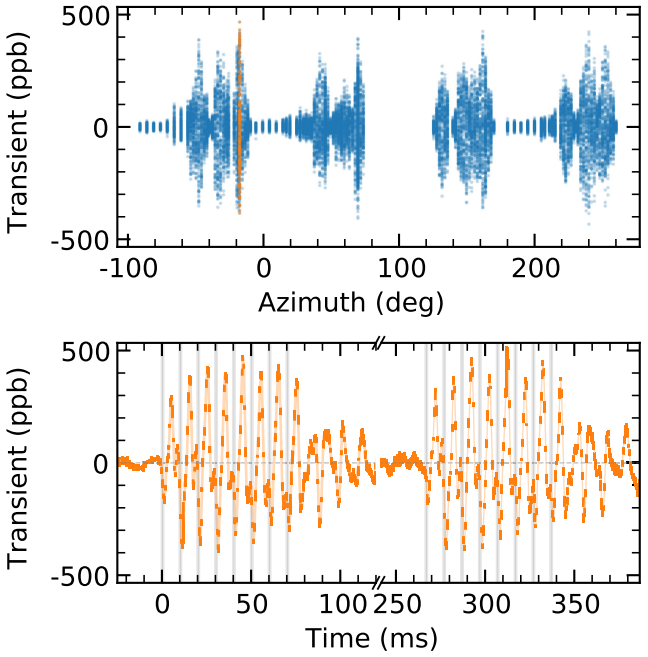

FIG. 26. Top: the transient magnetic field from the vibration caused by the ESQ pulsing for all times as a function of azimuth in the storage ring. Bottom: the transient magnetic field as a function of time at one specific location (−17 deg). The times during which muons are stored are highlighted by gray bands. The shown field transients are scaled up to the ESQ operation voltage.

sections were approximated with two short ones. Due to the static nature of the used probes, only one measurement per section was feasible for most sections. The total shift of the magnetic field during the times the muons are stored averaged around the ring was determined from these spatially sparse measurements, leading to the dominant systematic uncertainty of the Run-1 result [5].

In dedicated measurement campaigns, the identical sealed NMR probes were mounted on a frame that can be moved around the ring using the trolley infrastructure. The NMR probes were pulsed and read out in the same scheme used in Run-1 through a dedicated multiplexer of the fixed probe systems, now through the ∼50 m long trolley cable. This scheme allows mapping of the effect with finer resolution, significantly improving the precision. In the summer of 2020, a quarter of the ring was mapped, and in summer 2021, the full ring was mapped. The top plot in Fig. 26 shows the transients for all times as a function of azimuth around the ring. The measurements were performed at a reduced ESQ voltage of 14 kV. The confirmed voltage-squared dependence was used to scale the measurement to the nominal ESQ operations voltage of 18.2 kV.

The effect of the magnetic field perturbations on ω_a in a particular fill at a particular azimuthal position is estimated by a linear fit of the magnetic field transient over the muon storage time of around 700 μs of this fill. The effect

TABLE XXI. Correction and associated uncertainties to $\tilde{\omega}'_p$ due to transient magnetic fields caused by the pulsing of the ESQ system.

Description	Correction (ppb)	Uncertainty (ppb)
Frequency extraction		5
Skin depth		2
Stability over time		8
Azimuthal averaging		11
Transverse dependence		5.3
Measurement apparatus		10.5
Fill-by-fill variations		2
Second bunch train		5
Subtotal	−21.0	19.5

accumulates over the muon lifetime in the storage ring [7]. The azimuthally resolved effect from the different measurement positions is averaged around the ring, accounting for the different azimuthal spacings between the measurements. Segments outside vacuum chambers containing ESQs and where no time-dependent field perturbations are observed do not contribute. Table XXI shows the total correction $B_Q = -21.0 \pm 19.5$ ppb due to transient magnetic fields from the ESQ and lists the corresponding uncertainties, which are discussed in more detail below.

The frequency extraction from NMR FID signals requires a minimal length of more than ∼0.5 ms for the required resolution. The timescale of the observed transient changes the field within an FID. Hence, magnetic field perturbations from outside the fit window of the transient effect leak into the frequency. Alternatively, the phase function from multiple FIDs with different delays with respect to the muon injection time can be combined and fitted directly in the relevant time window. The NMR probes have a 0.5-mm-thick aluminum shell, and the corresponding skin depth suppresses higher-frequency components. This effect was evaluated in a dedicated measurement. The transient caused by the ESQ was mapped partially one year after Run-3 and around the full ring the year afterward. In addition, starting mid-Run-3, periodic measurements at static positions were taken. The different measurements over time are in good agreement. In addition, the fixed probe system is used to monitor the effect of the transient from outside of the vacuum chambers parasitically during data taking.

All the measurements are point estimates, and the values in between the measurement points are unknown, resulting in uncertainty in the azimuthal averaging. In addition, the mapping was performed in the center of the storage volume. The radial dependence of the transient was measured on the diagonal along the ESQ 0 V-line at one location. A flat dependence was found up to 2 cm, where most of the muon beam is located, and variations up to 25% were observed at a radius of 4 cm, at the edge of the storage volume. As mentioned above, the ESQ can only be operated consistently at 14 kV with the mapper device present. Perturbations of the electric field from the mapping device itself might modify the local forces on the ESQ plates and change the mechanical oscillation of the system. Other sources for uncertainties are fill-by-fill intensity variations not accounted for the averaging between the 16 fills and small changes in the time structures in the second eight bunches between running conditions and the measurements.

H. Summary and differences with respect to Run-1

The dataset averaged $\tilde{\omega}'_p$ are listed in Table XXVI. All non-negligible uncertainties are summarized in Table XXII. For uncertainties that have been determined on a probe-by-probe basis, the uncertainties are translated to multipole moments and further to $\tilde{\omega}'_p$ taking the correlation between moments and the spatial and temporal muon distribution into account. Uncertainties are highly correlated and thus treated as fully correlated, except the Brownian bridge-based tracking uncertainty, which is random in nature and reduced by combining datasets. Calibration constants and corrections are taken into account in the final $\tilde{\omega}'_p$ and are not listed individually. The total uncertainty on the muon-weighted magnetic field, including corrections from magnetic field transients, is ≤ 52 ppb, a factor of ∼2 improvement compared to the Run-1 analysis [7]. The main reason is the improved understanding of the electrostatic quadrupole transient due to additional measurements. Overall, the current uncertainty budget is well below the systematic uncertainty goal from the technical design report of <70 ppb.

The major differences in the Run-2/3 analysis of $\tilde{\omega}'_p$ with respect to the Run-1 analysis are listed below:

TABLE XXII. Summary of uncertainties on $\tilde{\omega}'_p$ for each step in the analysis. A detailed breakdown of each contribution is given in the corresponding section. A single value per line indicates the same value for all datasets. All contributions are assumed to be fully correlated, except the Brownian bridge uncertainty in the Tracking section, which is treated as statistical uncertainty.

	Uncertainty (ppb)			
Description	Run-2	Run-3a	Run-3b	Section
Calibration probe		8.9		VI B
Trolley calibration		17.8		VI C
Spatial field maps	37.2	38.5	38.1	VI D
Tracking	17.3	16.5	17.8	VI E
Muon weighting	13.4	7.9	6.9	VI F
Transient booster		7		VI G
Transient kicker	13.3			VI G 1
Transient ESQ	19.5			VI G 2
Subtotal uncorrelated	15.4	10.7	16.0	
Subtotal correlated	51.3	52.0	50.6	

(1) In Run-1, the transverse multipole expansion was truncated at $N_{\max} = 9$, for Run-2/3, $N_{\max} = 12$ was used.
(2) In the frequency extraction of the trolley FIDs, in Run-2/3, slightly earlier times in the phase function fits were used compared to Run-1.
(3) While in Run-1 only one of the barcode readers was used to determine the azimuthal position, in Run-2 and Run-3 the second barcode reader is used as a cross-check, increasing reliability. This has the advantage that measurements in the small gaps between adjacent vacuum chamber positions can still be reconstructed even though one of the barcode readers fails. In addition, better timing alignment of the barcode and encoder systems is possible due to additional timing information in the raw data of both systems. These two developments led to improved reliability of the position determination.
(4) For Run-2/3, the trolley calibration procedures were improved with respect to Run-1. The improvements include the following: (1) moving the trolley further from the calibration position during measurements with the calibration probe; (2) revised corrections to the calibration-probe mounting configuration; (3) inclusion of improved magnetic image measurements described in Sec. VI B; (4) corrections for second-order gradients near the calibration position due to the different effective sample volumes of the trolley probe and calibration probe.
(5) A ground loop issue that was present in Run-1 was removed between Run-1 and Run-2.
(6) Higher-order multipole moments are smaller in Run-2/3 than in Run-1. They were shimmed out better after Run-1 due to the availability of trolley calibration constants. This reduces the uncertainty from the rail misalignments, as well as from muon weighting.
(7) The temperature dependence of the trolley NMR probes was measured more precisely for Run-2/3. It was evaluated as (-0.8 ± 2.0) ppb/°C. In Run-1, a temperature dependence of (0 ± 5) ppb/°C was used.
(8) The rate of change parameter M used for the uncertainty evaluation of the field tracking with a random walk or Brownian bridge model was evaluated in Run-1 station-by-station, manually including observed correlations. This approach was chosen due to the statistics of field periods. In Run-2/3, M is evaluated directly from azimuthal averages, which intrinsically includes correlations.
(9) Additional measurements with a dedicated magnetometer with significantly reduced vibrations lowered the uncertainty on the measurements of transient magnetic fields from the kickers.
(10) An extensive azimuthal mapping of the transient magnetic field from the ESQ system reduced the corresponding uncertainty significantly.

VII. OVERALL $\omega_a/\tilde{\omega}'_p$ CONSISTENCY CHECKS

The R'_μ ratio values have been investigated for any inconsistencies and unexpected correlations to external parameters. These external parameters are representative of the conditions that the experiment Run-2/3 data had been collected in. Eight external parameters had been identified for these checks, namely, average temperature of the muon storage ring, average vacuum pressure of the muon storage ring, magnet current, inflector current, time of data collection since last magnet ramp up, time of data collection (day or night), amplitude of CBO, and k_{loss}.

A. Methodology

In order to perform these checks the data were split into five slices based on the external parameter values, for each of the three run sets. The ω_a and ω_p values with their respective uncertainties are subsequently extracted from each of the 15 data slices. These in turn are used to calculate the R'_μ ratio and its uncertainty for each of the data slices. It should be noted that for this study the beam dynamics and magnetic field transient corrections are assumed to be constant within the Run-2, Run-3a, and Run-3b datasets. These checks were performed on relatively unblinded but overall still blinded data, and repeated eventually on unblinded data.

For the purposes of these tests, we perform a χ^2 minimization on the calculated R'_μ ratios and their uncertainties in order to evaluate the overall optimal error weighted R'_μ ratio value for each external variable studied. Thereafter, the p value for the sliced R'_μ ratios against the optimal R'_μ ratio is extracted.

Furthermore, the sliced R'_μ ratio values are plotted against the external parameter values for each of the slices and fitted against a constant. The pull histograms for these plots are then evaluated for any skewness in order to identify dependencies on the external parameters at hand.

B. Results

The p values for all the different external parameter cross-checks performed using the methodology described

TABLE XXIII. R'_μ ratio vs external parameter value with optimal R'_μ ratio fit p values, for combined runs 2, 3a, and 3b slicings.

External variable	p value
Average ring temperature	0.43
Inflector current	0.75
Magnet current	0.13
Time since magnet ramp up	0.91
Day/night split	0.70
Average vacuum pressure	0.75
Amplitude of CBO	0.77
k_{loss}	0.93

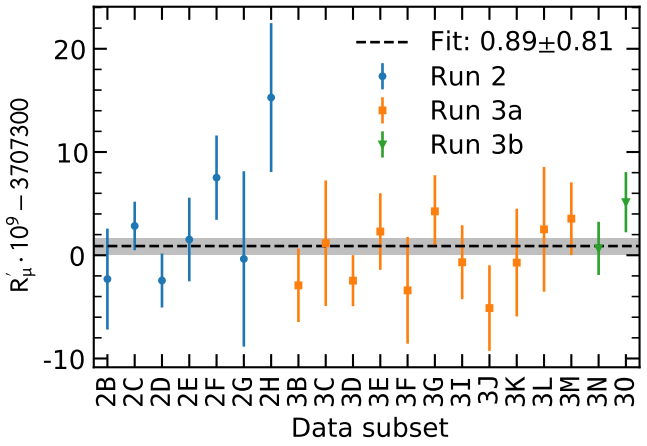

FIG. 27. $R'_\mu(T_r)$ versus data subset. The fit line has a $\chi^2/\mathrm{ndf} = 19.31/19$ with a p value of 44%.

above are summarized in Table XXIII. In the Run-2, Run-3a, and Run-3b overall consistency study, none of the sliced R'_μ ratio values show any direct dependency on the eight investigated external parameters, with p values within nominal ranges. Moreover, the pull histograms for each of the external parameter slicing fits show a Gaussian distribution of the data centered around 0.0 ± 0.2.

The magnet current slicing has a relatively small p value due to a pull from the slices containing data from runs 2F, 2H, and 3N. Detailed analysis cross-checks have been made for datasets 2F, 2H, and 3N, by ω_a^m and $\tilde{\omega}'_p$ analyzers. In these cross-checks no extraordinary anomaly was discovered by the analyzers, consequently, the datasets remain valid datasets with statistical fluctuation.

Additionally, a slicing over different datasets was also performed in order to examine the consistency of the extracted R'_μ ratio values over different datasets and time. The results for this data splitting can be visualized in Fig. 27. There are no observed inconsistencies for the R'_μ ratio values extracted for different datasets.

VIII. CALCULATION OF a_μ

Following Eq. (4), for each dataset, the measured ω_a^m is corrected by adding the beam dynamics corrections, and the ratio $R'_\mu(T_r) = \omega_a/\tilde{\omega}'_p(T_r)$ is computed. Table XXIV provides an overview of all contributions. All uncertainty contributions to ω_a^m, to the beam dynamics corrections and to $\tilde{\omega}'_p(T_r)$, are propagated to $R'_\mu(T_r)$.

Uncertainty contributions that are assumed to be fully correlated between different Run-2/3 datasets and also between different measurements by the Fermilab Muon $g-2$ (E989) collaboration are tracked separately from the statistical uncertainties and the other uncertainty contributions that can be considered uncorrelated: the magnetic field uncorrelated uncertainty. The correlation matrix between the ratios is reported in Table XXV. The three $R'_\mu(T_r)$ values are found to be statistically consistent and are fit to obtain the measured $R'_\mu(T_r)$ for the Run-2/3 sample. The fit χ^2 probability is about 20%. The results are summarized in Table XXVI.

Over the course of this analysis, three small errors in the Run-1 analysis [5] were identified. The total shift in the previous result due to these errors is 28 ppb, resulting in $R'_\mu(T_r)_{\mathrm{run\,1}} = 0.0037073004(16)(6)$. The measured $R'_\mu(T_r)_{\mathrm{run\,2/3}} = 0.00370730088(75)(26)$ is combined with the Run-1 result [5], assuming that the systematic uncertainties are fully correlated, to obtain the

TABLE XXIV. Values and uncertainties of the $\mathcal{R}'_\mu$ terms in Eq. (4) and uncertainties due to the external parameters in Eq. (56) for a_μ. Positive C_i increase a_μ; positive B_i decrease a_μ. The ω_a^m uncertainties are decomposed into statistical and systematic contributions.

Quantity	Correction (ppb)	Uncertainty (ppb)	Section
ω_a^m statistical	⋯	201	IV K
ω_a^m systematic	⋯	25	IV J
C_e	451	32	V A
C_p	170	10	V B
C_{ml}	0	3	V C
C_{dd}	−15	17	V D
C_{pa}	−27	13	V E
$\langle\omega'_p \times M\rangle$	⋯	46	VI H
B_K	−21	13	VI G 1
B_Q	−21	20	VI G 2
$\mu'_p(34.7\,°\mathrm{C})/\mu_e$	⋯	11	[45,57]
m_μ/m_e	⋯	22	[58]
$g_e/2$	⋯	0	[59]
Total systematic	⋯	70	
Total external parameters	⋯	25	
Totals	622	215	

TABLE XXV. Correlation matrix of the Run-2/3 datasets measurements of $R'_\mu(T_r)$.

$R'_\mu(T_r)$	Run-2	Run-3a	Run-3b
Run-2	1.00	0.05	0.03
Run-3a	0.05	1.00	0.03
Run-3b	0.03	0.03	1.00

TABLE XXVI. Run-2/3 datasets measurements of ω_a, $\tilde{\omega}'_p(T_r)$, and their ratios $R'_\mu(T_r)$ multiplied by 1000.

Dataset	$\omega_a/2\pi$ (Hz)	$\tilde{\omega}'_p(T_r)/2\pi$ (Hz)	$R'_\mu(T_r) \times 1000$
Run-2	229077.408(79)	61790875.0(3.3)	3.7073016(13)
Run-3a	229077.591(68)	61790957.5(3.3)	3.7072996(11)
Run-3b	229077.81(11)	61790962.3(3.3)	3.7073029(18)
Run-2/3			3.70730088(79)

Fermilab experimental measurement, $R'_\mu(T_r)_{\text{run 1/2/3}} = 0.00370730082(68)(31)$. This value is combined with the BNL measurement of R_μ for free protons in vacuum [2], $R_\mu = 0.0037072063(20)$, after converting it using the measured diamagnetic shielding correction $\sigma_{p'}(T_r)$ [45]:

$$R'_\mu(T_r) = \frac{R_\mu}{1-\sigma_{p'}(T_r)} = 0.0037073019(20). \quad (55)$$

We compared the systematic uncertainties for the BNL and FNAL measurements and, due to the significant changes in the beam characteristics and detectors between the experiments, concluded that those uncertainties were largely uncorrelated between the two experiments. The resulting experimental average is $R'_\mu(T_r)_{\text{Exp}} = 0.00370730095(70)$.

The muon magnetic anomaly is computed from

$$a_\mu = R'_\mu(T_r)\frac{\mu'_p(T_r)}{\mu_e(H)}\frac{\mu_e(H)}{\mu_e}\frac{m_\mu}{m_e}\frac{g_e}{2}. \quad (56)$$

Here $\mu'_p(T_r)/\mu_e(H)$ is the ratio of the magnetic moment of the proton in a spherical water sample at 34.7 °C and the magnetic moment of the electron in a hydrogen atom [45] (10.5 ppb). $\mu_e(H)/\mu_e$ is the ratio of the magnetic moment of the electron in a hydrogen atom and the magnetic moment of the free electron in vacuum, obtained with a theory QED calculation [57], whose precision is limited to 100 ppt by the number of reported digits. m_μ/m_e is the ratio of the muon and electron masses (22 ppb), taken from the CODATA 2018 fit [58], primarily driven by the LAMPF 1999 measurements of muonium hyperfine splitting [60]. g_e is the electron gyromagnetic factor, computed from the electron anomaly $a_e = (g-2)/2$ world average [59] (100 ppt), dominated by [1].

The measured muon magnetic anomaly for this measurement, this measurement combined with our Run-1 result, and the combined BNL and FNAL results are

$$\begin{aligned} a_\mu^{\text{FNAL run 2/3}} &= 116\,592\,057(25)\times 10^{-11}(0.21\text{ ppm}),\\ a_\mu^{\text{FNAL run 1/2/3}} &= 116\,592\,055(24)\times 10^{-11}(0.20\text{ ppm}),\\ a_\mu^{\text{Exp}} &= 116\,592\,059(22)\times 10^{-11}(0.19\text{ ppm}). \end{aligned}$$

These are displayed in Fig. 28. Values of $R'_\mu(T_r)$ and a_μ with extra digits to facilitate further calculations without loss of precision due to rounding are provided in the Supplemental Material [61].

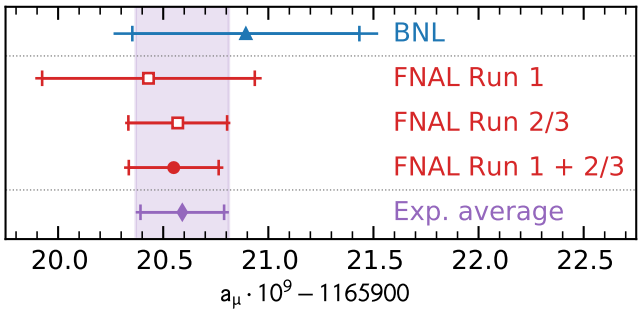

FIG. 28. From top to bottom: experimental values of a_μ from BNL E821, the FNAL 2021 measurement (FNAL Run-1), this measurement (FNAL Run-2/3), the FNAL combined measurement (FNAL Run-1 + 2/3), and the combined experimental average (Exp. average). The inner tick marks indicate the statistical contribution to the total uncertainties.

IX. COMPARISON TO THEORY

In recent years, all aspects of the SM theory prediction a_μ^{SM} have been scrutinized and refined with continued theoretical and computational efforts. These were summarized by the $g-2$ theory initiative [10], using results from Refs. [62–81]. While the QED and electroweak contributions are widely considered noncontroversial, the SM prediction of the muon $g-2$ is limited by our knowledge of the vacuum fluctuations involving strongly interacting particles, comprising effects called hadronic vacuum polarization and hadronic light-by-light scattering. The latter is currently known at a level of precision comparable to a_μ^{Exp}, and it is the leading hadronic vacuum polarization contribution to the muon magnetic anomaly, denoted by a_μ^{HLO}, that gives the dominant uncertainty to the SM prediction. These effects cannot be computed at low-energy scales due to the nonperturbative nature of QCD at large distances. It is possible to overcome this problem by means of a dispersion relation technique involving experimental data on the cross-section of electron-positron annihilation into hadrons, $e^+e^- \to$ hadrons. In the last 20 years, the worldwide efforts of experiments working on $e^+e^- \to$ hadrons data in the energy range below a few GeV have achieved the remarkable uncertainty of 0.6% on a_μ^{HLO} [10,82]. In addition, in the last few years, there has been significant progress on the first-principles calculation of a_μ^{HLO} using lattice QCD which, however, was not yet as precise as the data-driven dispersive approach compiled in [10]. In 2021, the BMW collaboration published the first lattice calculation of a_μ^{HLO} with subpercent precision [9]. This result would move a_μ^{SM} towards a_μ^{Exp} and is compatible with the "no new physics" scenario but discrepant with the dispersive approach. While the evaluation of the whole a_μ^{HLO} from the other lattice groups is in progress, excellent agreement between the different lattice groups is found for the so-called intermediate window observable [83–87]. The evaluation of this intermediate window observable shows a 4 standard deviation discrepancy between the lattice and the data-driven computation. On the $e^+e^- \to$ hadrons side, in addition to the known discrepancy between KLOE [88–91] and *BABAR* [92,93], the recent CMD-3 [94,95] result has shown a discrepancy with all previous measurements used in [10]. The origin of this discrepancy is currently unknown and efforts are in

progress to clarify the situation [96]. In view of this situation, a firm comparison with the theory cannot be established at the moment.

X. CONCLUSION

We have reported a measurement of the muon magnetic anomaly to 0.20 ppm precision, based on the first three years of data. This measurement represents the most precise determination of this quantity. The statistical and systematic errors have been reduced by a factor of 2 with respect to our first measurement [5], due to greater than four times more data and improved running conditions, analysis procedures, dedicated measurements, and systematic studies. This measurement is still statistically limited and the analysis of the remaining data from three additional years of data is expected to result in an improved statistical precision by another factor of approximately 2.

ACKNOWLEDGMENTS

We thank the Fermilab management and staff for their strong support of this experiment, as well as the tremendous support from our university and national laboratory engineers, technicians, and workshops. Greg Bock and Joe Lykken set the blinding clock and diligently monitored its stability. The Muon $g-2$ Experiment was performed at the Fermi National Accelerator Laboratory, a U.S. Department of Energy, Office of Science, HEP User Facility. Fermilab is managed by Fermi Research Alliance, LLC (FRA), acting under Contract No. DE-AC02-07CH11359. Additional support for the experiment was provided by the Department of Energy offices of HEP, NP, and ASCR (USA); the National Science Foundation (USA); the Istituto Nazionale di Fisica Nucleare (Italy); the Science and Technology Facilities Council (UK); the Royal Society (UK); the National Natural Science Foundation of China (Grant No. 12211540001, 12075151); MSIP, NRF, and IBS-R017-D1 (Republic of Korea); the German Research Foundation (DFG) through the Cluster of Excellence PRISMA+ (EXC 2118/1, Project ID 39083149); the European Union Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie Grant Agreements No. 101006726, No. 734303; and European Union STRONG 2020 project under Grant Agreement No. 824093 and the Leverhulme Trust, LIP-2021-01.

APPENDIX A: CORRELATIONS BETWEEN ω_a^m ANALYSES

Table XXVII lists the correlations coefficients between the 19 different ω_a analyses. The largest allowed statistical differences are between the event-based analyses and the energy-based analyses. Smaller allowed statistical differences are between analyses that employ either a common construction approach or a common histogramming method. The correlation coefficients do not account for additional allowed systematic differences between analysis methods.

TABLE XXVII. Table of correlation coefficients based on the allowed statistical differences between the 19 different ω_a analysis approaches. They include the different reconstruction procedures and different histogramming methods. They assume a 100% correlation of systematic uncertainties between analysis approaches.

	C_T	E_T	I_T	S_T	W_T	B_A	C_A	E_A	I_A	S_A	W_A	B_RT	E_RT	I_RT	B_RA	E_RA	K_Q	KR_RQ
B_T	0.967	0.999	0.967	0.999	1.000	0.900	0.871	0.884	0.867	0.884	0.884	0.993	0.995	0.963	0.895	0.904	0.765	0.824
C_T		0.967	1.000	0.965	0.967	0.891	0.900	0.875	0.896	0.874	0.875	0.961	0.963	0.996	0.887	0.895	0.756	0.815
E_T			0.967	0.999	0.999	0.913	0.885	0.898	0.880	0.898	0.897	0.993	0.996	0.963	0.909	0.918	0.753	0.811
I_T				0.965	0.967	0.897	0.906	0.881	0.902	0.880	0.880	0.961	0.963	0.996	0.892	0.901	0.751	0.809
S_T					0.999	0.915	0.886	0.900	0.882	0.902	0.899	0.992	0.995	0.961	0.911	0.920	0.751	0.809
W_T						0.915	0.887	0.899	0.882	0.899	0.899	0.993	0.995	0.963	0.911	0.919	0.752	0.810
B_A							0.994	1.000	0.994	0.999	1.000	0.890	0.886	0.887	0.991	0.994	0.688	0.740
C_A								0.994	1.000	0.993	0.994	0.862	0.857	0.896	0.986	0.988	0.681	0.732
E_A									0.994	0.999	1.000	0.875	0.871	0.871	0.991	0.994	0.676	0.727
I_A										0.993	0.994	0.858	0.853	0.892	0.986	0.988	0.678	0.729
S_A											0.999	0.875	0.871	0.870	0.990	0.993	0.677	0.728
W_A												0.875	0.870	0.871	0.991	0.994	0.676	0.727
B_RT													0.994	0.962	0.902	0.907	0.758	0.825
E_RT														0.967	0.895	0.901	0.767	0.837
I_RT															0.895	0.901	0.750	0.819
B_RA																0.994	0.682	0.743
E_RA																	0.689	0.754
K_Q																		0.994

TABLE XXVIII. Overview of trolley probe calibration constants δ^{calib} and individual contributions for Run-2/3. All values are given in ppb.

	$\delta^{\text{fp,tr}}$		$\delta^{\text{fp,cp}}$		δ^{av}		$\delta^{s,\text{img}}$		δ_n^{calib}
Probe	Value	Uncertainty	Value	Uncertainty	Value	Uncertainty	Value	Uncertainty	Value
1			4.0	4.0	4.9	1.9	−17.2	8.9	1469.0
2			4.0	4.0	−0.2	2.4	−17.8	8.9	1336.9
3			3.7	3.7	1.8	2.9	−17.2	8.9	1523.6
4			4.0	4.0	2.8	4.0	−17.8	8.9	1358.3
5			4.9	4.9	−1.0	3.1	−17.2	8.9	1514.4
6			3.6	3.6	9.4	4.4	−20.2	9.1	1734.5
7			3.2	3.2	−9.5	4.7	−19.4	8.9	1903.0
8			3.2	3.2	−2.8	2.9	−17.8	8.9	1195.8
9	14.3	8	3.1	3.1	7.9	4.0	−17.2	8.9	1367.2
10			3.1	3.1	8.6	3.4	−17.8	8.9	421.1
11			3.1	3.1	19.7	9.1	−19.4	8.9	2878.3
12			3.7	3.7	40.9	8.1	−20.2	9.1	1787.1
13			4.4	4.4	−4.4	4.4	−19.4	8.9	1993.8
14			5.7	5.7	1.5	6.1	−17.8	8.9	1263.9
15			6.5	6.5	−15.2	6.5	−17.2	8.9	1193.0
16			5.5	5.5	−1.0	4.4	−17.8	8.9	337.2
17			4.2	4.2	4.9	8.3	−19.4	8.9	2738.5

APPENDIX B: TROLLEY CALIBRATION CONSTANTS

The trolley calibration constants, including their contributions, are listed in Table XXVIII. A graphic comparison is shown in Fig. 29. In addition to the Run-2/3 average, the values and the differences from the dedicated Run-2 and Run-3 calibration campaigns are shown, in combination with predictions from COMSOL simulations based on a simplified trolley geometry that only takes into account the trolley shell but not the interior details.

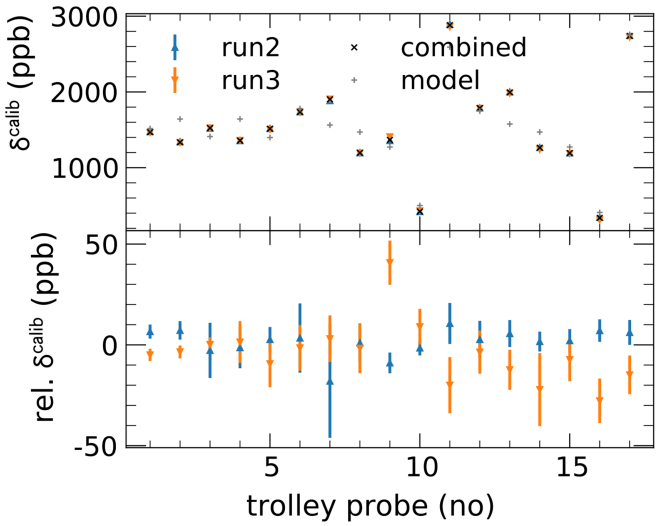

FIG. 29. Top: trolley calibration constants per trolley probe for Run-2 (blue) and Run-3 (orange) and the combination (black). Predictions from COMSOL simulations (gray) with simplified geometry, which only considers the trolley shell, show qualitative consistency. Bottom: the difference of Run-2 and Run-3 calibration constants with respect to the combined value that are used for this analysis.

[1] X. Fan, T. G. Myers, B. A. D. Sukra, and G. Gabrielse, Measurement of the electron magnetic moment, Phys. Rev. Lett. **130**, 071801 (2023).

[2] G. W. Bennett *et al.*, Final report of the E821 muon anomalous magnetic moment measurement at BNL, Phys. Rev. D **73**, 072003 (2006).

[3] D. Stockinger, The muon magnetic moment and supersymmetry, J. Phys. G **34**, R45 (2007).

[4] A. Czarnecki and W. J. Marciano, The muon anomalous magnetic moment: A Harbinger for 'new physics', Phys. Rev. D **64**, 013014 (2001).

[5] B. Abi *et al.* (Muon $g-2$ Collaboration), Measurement of the positive muon anomalous magnetic moment to 0.46 ppm, Phys. Rev. Lett. **126**, 141801 (2021).

[6] T. Albahri *et al.* (Muon $g-2$ Collaboration), Measurement of the anomalous precession frequency of the muon in the Fermilab muon $g-2$ experiment, Phys. Rev. D **103**, 072002 (2021).

[7] T. Albahri *et al.* (Muon $g-2$ Collaboration), Magnetic-field measurement and analysis for the muon $g-2$ experiment at Fermilab, Phys. Rev. A **103**, 042208 (2021).

[8] T. Albahri *et al.* (Muon $g-2$ Collaboration), Beam dynamics corrections to the Run-1 measurement of the muon anomalous magnetic moment at Fermilab, Phys. Rev. Accel. Beams **24**, 044002 (2021).

[9] S. Borsanyi *et al.*, Leading hadronic contribution to the muon magnetic moment from lattice QCD, Nature (London) **593**, 51 (2021).

[10] T. Aoyama *et al.*, The anomalous magnetic moment of the muon in the Standard Model, Phys. Rep. **887**, 1 (2020).

[11] D. P. Aguillard *et al.* (Muon $g-2$ Collaboration), Measurement of the positive muon anomalous magnetic moment to 0.20 ppm, Phys. Rev. Lett. **131**, 161802 (2023).

[12] J. Bailey *et al.*, Final report on the CERN muon storage ring including the anomalous magnetic moment and the electric dipole moment of the muon, and a direct test of relativistic time dilation, Nucl. Phys. **B150**, 1 (1979).

[13] G. T. Danby *et al.*, The Brookhaven muon storage ring magnet, Nucl. Instrum. Methods Phys. Res., Sect. A **457**, 151 (2001).

[14] D. Stratakis, M. E. Convery, C. Johnstone, J. Johnstone, J. P. Morgan, D. Still, J. D. Crnkovic, V. Tishchenko, W. M. Morse, and M. J. Syphers, Accelerator performance analysis of the Fermilab muon campus, Phys. Rev. Accel. Beams **20**, 111003 (2017).

[15] A. P. Schreckenberger *et al.*, The fast non-ferric kicker system for the Muon $g-2$ experiment at Fermilab, Nucl. Instrum. Methods Phys. Res., Sect. A **1011**, 165597 (2021).

[16] D. Flay, D. Kawall, T. Chupp, S. Corrodi, M. Farooq, M. Fertl, J. George, J. Grange, R. Hong, R. Osofsky, S. Ramachandran, E. Swanson, and P. Winter, High-accuracy absolute magnetometry with application to the Fermilab muon $g-2$ experiment, J. Instrum. **16**, P12041 (2021).

[17] J. Kaspar *et al.*, Design and performance of SiPM-based readout of PbF2 crystals for high-rate, precision timing applications, J. Instrum. **12**, P01009 (2017).

[18] K. Khaw *et al.*, Performance of the Muon $g-2$ calorimeter and readout systems measured with test beam data, Nucl. Instrum. Methods Phys. Res., Sect. A **945**, 162558 (2019).

[19] A. Anastasi *et al.*, Electron beam test of key elements of the laser-based calibration system for the muon $g-2$ experiment, Nucl. Instrum. Methods Phys. Res., Sect. A **842**, 86 (2017).

[20] A. Anastasi *et al.*, The laser-based gain monitoring system of the calorimeters in the muon $g-2$ experiment at Fermilab, J. Instrum. **14**, P11025 (2019).

[21] B. King *et al.*, The straw tracking detector for the Fermilab muon $g-2$ experiment, J. Instrum. **17**, P02035 (2022).

[22] J. Allison *et al.*, Recent developments in GEANT4, Nucl. Instrum. Methods Phys. Res., Sect. A **835**, 186 (2016).

[23] J. Allison *et al.*, GEANT4 developments and applications, IEEE Trans. Nucl. Sci. **53**, 270 (2006).

[24] S. Agostinelli *et al.*, Geant4–a simulation toolkit, Nucl. Instrum. Methods Phys. Res., Sect. A **506**, 250 (2003).

[25] D. Tarazona, M. Berz, and K. Makino, Muon loss rates from betatron resonances at the muon $g-2$ storage ring at Fermilab, Int. J. Mod. Phys. A **34**, 1942008 (2019).

[26] K. Makino and M. Berz, COSY INFINITY Version 9, Nucl. Instrum. Methods Phys. Res., Sect. A **558**, 346 (2006).

[27] D. Sagan, Bmad: A relativistic charged particle simulation library, Nucl. Instrum. Methods Phys. Res., Sect. A **558**, 356 (2006).

[28] https://timesmicrowave.com/.

[29] With Dielectric Sciences DS2264; https://www.dielectricsciences.com/.

[30] D. Stratakis, Application of passive wedge absorbers for improving the performance of precision-science experiments, Phys. Rev. Accel. Beams **22**, 053501 (2019).

[31] M. Berz, K. Makino, and W. Wan, *An Introduction to Beam Physics*, Series in High Energy Physics, Cosmology, and Gravitation (Taylor & Francis, London, 2015).

[32] A. Anastasi *et al.*, The laser-based gain monitoring system of the calorimeters in the muon $g-2$ experiment at Fermilab, J. Instrum. **14**, P11025 (2019).

[33] In superimposing waveforms, the waveforms are first superimposed and then the full positron reconstruction is rerun. In superimposing the crystal hits, the hits are first superimposed and then the positron clustering stage is rerun.

[34] F. James and M. Winkler, MINUIT User's Guide (2004).

[35] P. Virtanen, R. Gommers, T. Oliphant *et al.*, SCIPY: Open source scientific tools for PYTHON, Nat. Methods **17**, 261 (2020).

[36] https://lmfit.github.io/lmfit-py/fitting.html.

[37] In Refs. [6,11], we used a simplified version of Eq. (7) with positive phase term ϕ_0.

[38] In fitting the muon lifetime, some analyses added a χ^2 penalty term to constrain the time-dilated lifetime to results from cyclotron rotation studies.

[39] An exception to the policy of adding systematics linearly within a systematics category is the CBO frequency drift and CBO decoherence envelope systematics. A dedicated study showed that the two systematic uncertainties are independent and therefore add in quadrature.

[40] A negative muon loss parameter would imply a gain of stored muons and therefore is considered nonphysical.

[41] M. C. Jones and A. Pewsey, Sinh-arcsinh distributions, Biometrika **96**, 761 (2009).

[42] W. C. Davidon, Variable metric method for minimization, SIAM J. Opt. **1**, 1 (1991).

[43] O. Kim and Y. K. Semertzidis, Analytical estimations of the chromaticity and corrections to the spin precession frequency in weak focusing magnetic storage rings, Phys. Rev. Accel. Beams **25**, 024001 (2022).

[44] D. Stratakis, B. Drendel, J. P. Morgan, M. J. Syphers, and N. S. Froemming, Commissioning and first results of the Fermilab Muon Campus, Phys. Rev. Accel. Beams **22**, 011001 (2019).

[45] W. D. Phillips, W. E. Cooke, and D. Kleppner, Magnetic moment of the proton in H_2O in Bohr magnetons, Metrologia **13**, 179 (1977).

[46] R. Hoffman, Measurement of magnetic susceptibility and calculation of shape factor of NMR samples, J. Magn. Reson. **178**, 237 (2006).

[47] P. J. Mohr and B. N. Taylor, CODATA recommended values of the fundamental physical constants: 1998, Rev. Mod. Phys. **72**, 351 (2000).

[48] B. H. Blott and G. J. Daniell, The determination of magnetic moments of extended samples in a SQUID magnetometer, Meas. Sci. Technol. **4**, 462 (1993).

[49] J. S. Philo and W. M. Fairbank, Temperature dependence of the diamagnetism of water, J. Chem. Phys. **72**, 4429 (1980).

[50] G. S. Kell, Precise representation of volume properties of water at one atmosphere, J. Chem. Eng. Data **12**, 66 (1967).

[51] A. Vlassenbroek, J. Jeener, and P. Broekaert, Radiation damping in high resolution liquid NMR: A simulation study, J. Chem. Phys. **103**, 5886 (1995).

[52] B. W. Petley and R. W. Donaldson, The temperature dependence of the diamagnetic shielding correction for proton NMR in water, Metrologia **20**, 81 (1984).

[53] M. Farooq, T. Chupp, J. Grange, A. Tewsley-Booth, D. Flay, D. Kawall, N. Sachdeva, and P. Winter, Absolute magnetometry with ^{3}He, Phys. Rev. Lett. **124**, 223001 (2020).

[54] H. Yamaguchi, K. Sasaki, T. Tanaka, M. Abe, D. Flay, D. Kawall, T. Mibe, S. Seo, K. Shimomura, Y. Ueno, and P. Winter, Development of a CW-NMR probe for precise measurement of absolute magnetic field, IEEE Trans. Appl. Supercond. **29**, 1 (2019).

[55] Note that in Ref. [7], −5 ppb is used for the central value of the effect from the garage alone.

[56] R. Hong *et al.*, Systematic and statistical uncertainties of the Hilbert-transform based high-precision FID frequency extraction method, J. Magn. Reson. **329**, 107020 (2021).

[57] S. G. Karshenboim and V. G. Ivanov, The g factor of proton, Phys. Lett. B **566**, 27 (2003).

[58] E. Tiesinga, P. J. Mohr, D. B. Newell, and B. N. Taylor, CODATA recommended values of the fundamental physical constants: 2018, Rev. Mod. Phys. **93**, 025010 (2021).

[59] R. L. Workman *et al.* (Particle Data Group), Review of particle physics, Prog. Theor. Exp. Phys. **2022**, 083C01 (2022), and 2023 update.

[60] W. Liu *et al.*, High precision measurements of the ground state hyperfine structure interval of muonium and of the muon magnetic moment, Phys. Rev. Lett. **82**, 711 (1999).

[61] See the Supplemental Material at http://link.aps.org/supplemental/10.1103/PhysRevD.110.032009 for variables and correlation matrices with extra digits that are needed to derive the combined FNAL a_μ result and the experimental average with BNL.

[62] T. Aoyama, M. Hayakawa, T. Kinoshita, and M. Nio, Complete tenth-order QED contribution to the muon $g-2$, Phys. Rev. Lett. **109**, 111808 (2012).

[63] T. Aoyama, T. Kinoshita, and M. Nio, Theory of the anomalous magnetic moment of the electron, Atoms **7**, 28 (2019).

[64] A. Czarnecki, W. J. Marciano, and A. Vainshtein, Refinements in electroweak contributions to the muon anomalous magnetic moment, Phys. Rev. D **67**, 073006 (2003); **73**, 119901(E) (2006).

[65] C. Gnendiger, D. Stöckinger, and H. Stöckinger-Kim, The electroweak contributions to $(g-2)_\mu$ after the Higgs boson mass measurement, Phys. Rev. D **88**, 053005 (2013).

[66] M. Davier, A. Hoecker, B. Malaescu, and Z. Zhang, Reevaluation of the hadronic vacuum polarisation contributions to the standard model predictions of the muon $g-2$ and $\alpha(m_Z^2)$ using newest hadronic cross-section data, Eur. Phys. J. C **77**, 827 (2017).

[67] A. Keshavarzi, D. Nomura, and T. Teubner, Muon $g-2$ and $\alpha(M_Z^2)$: A new data-based analysis, Phys. Rev. D **97**, 114025 (2018).

[68] G. Colangelo, M. Hoferichter, and P. Stoffer, Two-pion contribution to hadronic vacuum polarization, J. High Energy Phys. 02 (2019) 006.

[69] M. Hoferichter, B.-L. Hoid, and B. Kubis, Three-pion contribution to hadronic vacuum polarization, J. High Energy Phys. 08 (2019) 137.

[70] M. Davier, A. Hoecker, B. Malaescu, and Z. Zhang, A new evaluation of the hadronic vacuum polarisation contributions to the muon anomalous magnetic moment and to $\alpha(m_Z^2)$, Eur. Phys. J. C **80**, 241 (2020); **80**, 410(E) (2020).

[71] A. Keshavarzi, D. Nomura, and T. Teubner, The $g-2$ of charged leptons, $\alpha(M_Z^2)$ and the hyperfine splitting of muonium, Phys. Rev. D **101**, 014029 (2020).

[72] A. Kurz, T. Liu, P. Marquard, and M. Steinhauser, Hadronic contribution to the muon anomalous magnetic moment to next-to-next-to-leading order, Phys. Lett. B **734**, 144 (2014).

[73] K. Melnikov and A. Vainshtein, Hadronic light-by-light scattering contribution to the muon anomalous magnetic moment revisited, Phys. Rev. D **70**, 113006 (2004).

[74] P. Masjuan and P. Sánchez-Puertas, Pseudoscalar-pole contribution to the $(g_\mu-2)$: A rational approach, Phys. Rev. D **95**, 054026 (2017).

[75] G. Colangelo, M. Hoferichter, M. Procura, and P. Stoffer, Dispersion relation for hadronic light-by-light scattering: Two-pion contributions, J. High Energy Phys. 04 (2017) 161.

[76] M. Hoferichter, B.-L. Hoid, B. Kubis, S. Leupold, and S. P. Schneider, Dispersion relation for hadronic light-by-light scattering: Pion pole, J. High Energy Phys. 10 (2018) 141.

[77] A. Gérardin, H. B. Meyer, and A. Nyffeler, Lattice calculation of the pion transition form factor with $N_f = 2+1$ Wilson quarks, Phys. Rev. D **100**, 034520 (2019).

[78] J. Bijnens, N. Hermansson-Truedsson, and A. Rodríguez-Sánchez, Short-distance constraints for the HLbL contribution to the muon anomalous magnetic moment, Phys. Lett. B **798**, 134994 (2019).

[79] G. Colangelo, F. Hagelstein, M. Hoferichter, L. Laub, and P. Stoffer, Longitudinal short-distance constraints for

the hadronic light-by-light contribution to $(g-2)_\mu$ with large-N_c Regge models, J. High Energy Phys. 03 (2020) 101.

[80] T. Blum, N. Christ, M. Hayakawa, T. Izubuchi, L. Jin, C. Jung, and C. Lehner, Hadronic light-by-light scattering contribution to the muon anomalous magnetic moment from lattice QCD, Phys. Rev. Lett. **124**, 132002 (2020).

[81] G. Colangelo, M. Hoferichter, A. Nyffeler, M. Passera, and P. Stoffer, Remarks on higher-order hadronic corrections to the muon $g-2$, Phys. Lett. B **735**, 90 (2014).

[82] F. Jegerlehner, *The Anomalous Magnetic Moment of the Muon* (Springer Cham, 2017), 10.1007/978-3-319-63577-4.

[83] G. Colangelo, A. X. El-Khadra, M. Hoferichter, A. Keshavarzi, C. Lehner, P. Stoffer, and T. Teubner, Data-driven evaluations of Euclidean windows to scrutinize hadronic vacuum polarization, Phys. Lett. B **833**, 137313 (2022).

[84] M. Cè *et al.*, Window observable for the hadronic vacuum polarization contribution to the muon $g-2$ from lattice QCD, Phys. Rev. D **106**, 114502 (2022).

[85] C. Alexandrou *et al.* (Extended Twisted Mass Collaboration), Lattice calculation of the short and intermediate time-distance hadronic vacuum polarization contributions to the muon magnetic moment using twisted-mass fermions, Phys. Rev. D **107**, 074506 (2023).

[86] A. Bazavov *et al.* (Fermilab Lattice, HPQCD, MILC Collaborations), Light-quark connected intermediate-window contributions to the muon $g-2$ hadronic vacuum polarization from lattice QCD, Phys. Rev. D **107**, 114514 (2023).

[87] T. Blum *et al.* (RBC, UKQCD Collaborations), Update of Euclidean windows of the hadronic vacuum polarization, Phys. Rev. D **108**, 054507 (2023).

[88] A. Anastasi *et al.* (KLOE-2 Collaboration), Combination of KLOE $\sigma(e^+e^- \to \pi^+\pi^-\gamma(\gamma))$ measurements and determination of $a_\mu^{\pi^+\pi^-}$ in the energy range $0.10 < s < 0.95$ GeV^2, J. High Energy Phys. 03 (2018) 173.

[89] F. Ambrosino *et al.* (KLOE Collaboration), Measurement of $\sigma(e^+e^- \to \pi^+\pi^-\gamma(\gamma)$ and the dipion contribution to the muon anomaly with the KLOE detector, Phys. Lett. B **670**, 285 (2009).

[90] F. Ambrosino *et al.* (KLOE Collaboration), Measurement of $\sigma(e^+e^- \to \pi^+\pi^-)$ from threshold to 0.85 GeV^2 using initial state radiation with the KLOE detector, Phys. Lett. B **700**, 102 (2011).

[91] D. Babusci *et al.* (KLOE Collaboration), Precision measurement of $\sigma(e^+e^- \to \pi^+\pi^-\gamma)/\sigma(e^+e^- \to \mu^+\mu^-\gamma)$ and determination of the $\pi^+\pi^-$ contribution to the muon anomaly with the KLOE detector, Phys. Lett. B **720**, 336 (2013).

[92] J. P. Lees *et al.* (*BABAR* Collaboration), Precise measurement of the $e^+e^- \to \pi^+\pi^-(\gamma)$ cross section with the initial-state radiation method at *BABAR*, Phys. Rev. D **86**, 032013 (2012).

[93] B. Aubert *et al.* (*BABAR* Collaboration), Precise measurement of the $e^+e^- \to \pi^+\pi^-(\gamma)$ cross section with the Initial State Radiation method at *BABAR*, Phys. Rev. Lett. **103**, 231801 (2009).

[94] F. V. Ignatov *et al.* (CMD-3 Collaboration), Measurement of the $e^+e^- \to \pi^+\pi^-$ cross section from threshold to 1.2 GeV with the CMD-3 detector, Phys. Rev. D **109**, 112002 (2024).

[95] F. V. Ignatov *et al.* (CMD-3 Collaboration), Measurement of the pion form factor with CMD-3 detector and its implication to the hadronic contribution to muon $(g-2)$, Phys. Rev. Lett. **132**, 231903 (2009).

[96] G. Colangelo *et al.*, Prospects for precise predictions of a_μ in the standard model, arXiv:2203.15810.