FERMILAB-PUB-23-006-T

Light-quark connected intermediate-window contributions to the muon $g-2$ hadronic vacuum polarization from lattice QCD

Alexei Bazavov,[1] Christine Davies,[2] Carleton DeTar,[3] Aida X. El-Khadra,[4,5] Elvira Gámiz,[6] Steven Gottlieb,[7] William I. Jay,[8] Hwancheol Jeong,[7] Andreas S. Kronfeld,[9] Shaun Lahert,[4,5,*] G. Peter Lepage,[10] Michael Lynch,[4,5] Andrew T. Lytle,[4,5] Paul B. Mackenzie,[9] Craig McNeile,[11] Ethan T. Neil,[12] Curtis T. Peterson,[12,†] Gaurav Ray,[11] James N. Simone,[9] Ruth S. Van de Water,[9] and Alejandro Vaquero[3,13]

(Fermilab Lattice, HPQCD, and MILC Collaborations)

[1]*Department of Computational Mathematics, Science and Engineering, and Department of Physics and Astronomy, Michigan State University, East Lansing, Michigan 48824, USA*
[2]*SUPA, School of Physics and Astronomy, University of Glasgow, Glasgow, G12 8QQ, United Kingdom*
[3]*Department of Physics and Astronomy, University of Utah, Salt Lake City, UT 84112, USA*
[4]*Department of Physics, University of Illinois, Urbana, Illinois, 61801, USA*
[5]*Illinois Center for Advanced Studies of the Universe, University of Illinois, Urbana, Illinois, 61801, USA*
[6]*CAFPE and Departamento de Física Teórica y del Cosmos, Universidad de Granada, E-18071 Granada, Spain*
[7]*Department of Physics, Indiana University, Bloomington, Indiana 47405, USA*
[8]*Center for Theoretical Physics, Massachusetts Institute of Technology, Cambridge, MA 02139, USA*
[9]*Theory Division, Fermi National Accelerator Laboratory, Batavia, Illinois, 60510, USA*
[10]*Laboratory for Elementary-Particle Physics, Cornell University, Ithaca, NY 14853, USA*
[11]*Centre for Mathematical Sciences, University of Plymouth, Plymouth PL4 8AA, United Kingdom*
[12]*Department of Physics, University of Colorado, Boulder, Colorado 80309, USA*
[13]*Departmento de Física Teórica, Universidad de Zaragoza, 50009 Zaragoza, Spain*

(Dated: June 30, 2023)

Abstract

We present a lattice-QCD calculation of the light-quark connected contribution to window observables associated with the leading-order hadronic vacuum polarization contribution to the anomalous magnetic moment of the muon, $a_\mu^{\mathrm{HVP,LO}}$. We employ the MILC Collaboration's isospin-symmetric QCD gauge-field ensembles, which contain four flavors of dynamical highly-improved-staggered quarks with four lattice spacings between $a \approx 0.06$–0.15 fm and close-to-physical quark masses. We consider several effective-field-theory-based schemes for finite-volume and other lattice corrections and combine the results via Bayesian model averaging to obtain robust estimates of the associated systematic uncertainties. After unblinding, our final results for the intermediate and "W2" windows are $a_\mu^{ll,\mathrm{W}}(\mathrm{conn.}) = 206.6(1.0) \times 10^{-10}$ and $a_\mu^{ll,\mathrm{W2}}(\mathrm{conn.}) = 100.7(3.2) \times 10^{-10}$, respectively.

* slahert@illinois.edu
† curtis.peterson@colorado.edu

I. INTRODUCTION

In April 2021, the Fermilab muon $g-2$ experiment, E989, released their first result for the muon's anomalous magnetic moment $a_\mu \equiv (g_\mu - 2)/2$ based on Run-1 data collected in 2018 [1]. When combined with the previous measurement from Brookhaven National Lab experiment E821 [2], the new result for the muon's anomalous magnetic moment increases the disagreement with the Standard Model (SM) theory prediction [3][1] from 3.7σ to 4.2σ. Because the anomaly arises from loop effects, it is sensitive to the contributions of yet-undiscovered particles that could give rise to small deviations from the theoretical prediction. Increased precision is now essential to say conclusively if this substantial difference is from physics beyond the SM.

The error on the experimental average of the muon's anomalous magnetic moment is now 0.35 parts per million (ppm), and is limited by statistics. Fermilab E989 continues to collect data and improve the experimental apparatus, and ultimately aims to measure a_μ to a precision at or below 0.14 ppm by the end of its lifetime. Additionally, a new complementary experiment to measure the muon's anomalous magnetic moment and electric dipole moment is planned for later this decade at J-PARC in Japan [24, 25]. The J-PARC E34 experiment will employ a different method to determine a_μ than the "magic momentum" approach of both Fermilab E989 and BNL E821 and aims for a precision of 0.45 ppm from its initial run in 2027 [24, 26].

Corresponding theoretical efforts are underway to reduce the uncertainty on the prediction for a_μ in the SM, which currently stands at 0.37 ppm [3]. At present, over 90% of the SM theory error comes from the leading-order hadronic vacuum polarization (HVP) contribution to the anomaly, $a_\mu^{\mathrm{HVP,LO}}$. The HVP contribution is difficult to determine precisely because the bulk of it comes from the low-energy, nonperturbative regime of quantum chromodynamics (QCD). To date, the most precise theoretical results for $a_\mu^{\mathrm{HVP,LO}}$ are obtained from a data-driven, dispersive approach [27, 28] using experimental measurements of the total cross section for $e^+e^- \to$ hadrons (the so-called R ratio) as input. These data-driven determinations have achieved around 0.5% precision on $a_\mu^{\mathrm{HVP,LO}}$ corresponding to 0.34 ppm uncertainty on a_μ [8–13] and are the basis of the Muon $g-2$ Theory Initiative's SM prediction for $a_\mu^{\mathrm{HVP,LO}}$ [3].

Lattice QCD provides an alternative, *ab initio*, approach for calculating the leading-order HVP contribution that is independent of experimentally measured cross sections.[2] The most precise lattice QCD calculation of $a_\mu^{\mathrm{HVP,LO}}$ to date (and the first with sub-percent precision) comes from the BMW collaboration [29]. Although BMW's result implies a SM value for a_μ that is within 1.5σ of experiment, it differs from the R-ratio based prediction of Ref. [3] by 2.1σ. Independent lattice-QCD calculations with commensurate precision are therefore urgently needed to address this theoretical discrepancy.

The leading-order HVP contribution to the muon's anomalous magnetic moment is computed in lattice QCD as a weighted integral over Euclidean time of the two-point correlation function of the quarks' electromagnetic vector current [30, 31]. By judiciously restricting the integration range (or "window"), one can construct sub-quantities of $a_\mu^{\mathrm{HVP,LO}}$ that avoid problematic statistical and/or systematic effects [32–34]. The same Euclidean-time window observables can also be obtained from R-ratio data by including a suitable weight function

[1] The SM prediction is based on a large body of theoretical work [4–23], and reflects the consensus of the muon $g-2$ theory community.

[2] A small number of experimentally-measured quantities are employed in lattice-QCD calculations to fix the quark masses and lattice spacing in the QCD Lagrangian.

in the dispersive integral for $a_\mu^{\rm HVP,LO}$ [35]. Because Euclidean-time windows allow for more detailed and sensitive comparisons between independent $a_\mu^{\rm HVP,LO}$ calculations, they are a valuable tool both for diagnosing sources of disagreement between lattice-QCD results and for quantifying differences (if any) between data-driven and lattice determinations.

Various Euclidean-time windows with differing positive features and drawbacks have been proposed [32–34]. In 2018, the RBC and UKQCD Collaborations separated the Euclidean-time integral into contributions from "short" ($t \lesssim 0.4$ fm), "intermediate" ($0.4 \lesssim t \lesssim 1.0$ fm), and "long" ($t \gtrsim 1.0$ fm) times [32]. The intermediate window observable $a_\mu^{\rm W}$ can be computed in lattice QCD with high statistical precision. Hence, it has been adopted by the muon $g-2$ theory community as a benchmark quantity. Several independent three- and four-flavor lattice-QCD calculations of $a_\mu^{\rm W}$ are now available [29, 32, 36, 37], but the results are not fully consistent. RBC/UKQCD's initial intermediate-window result [32] is within about 1σ of the determination from R-ratio data [35]. More recent lattice-QCD calculations of $a_\mu^{\rm W}$ by the BMW [29], Mainz/CLS [36], and ETM [37] collaborations, however, are all more than 3σ higher than the data-driven value.[3] Further scrutiny of the intermediate window is therefore needed to clarify the picture.

In this work, we calculate the intermediate-window contribution to $a_\mu^{\rm HVP,LO}$ in four-flavor lattice QCD. Using the same methods, we also calculate the "W2" window observable introduced by Aubin *et al.* [33], which corresponds to the Euclidean-time range $t \in [1.5, 1.9]$ fm. As pointed out in that work, although $a_\mu^{\rm W2}$ is statistically noisier than $a_\mu^{\rm W}$, effective-field theory (EFT) estimates of finite-volume, lattice-discretization, and pion-mass corrections are more reliable at larger times. We focus exclusively on the connected contribution from light (up and down) quarks in the isospin-symmetric limit, which accounts for about 90% of $a_\mu^{\rm HVP,LO}$. Calculations of the heavier quark flavors, isospin-breaking corrections and quark-disconnected contributions are in progress [39–42].

Our calculation of the intermediate and W2 window observables in this work builds upon our 2019 calculation of $a_\mu^{\rm HVP,LO}$ [43]. As before, we employ the MILC collaboration's dynamical-QCD gauge-field configurations [44] with four flavors of highly-improved-staggered quarks (HISQ) [45]. Our numerical simulations are again performed at the physical pion mass and with four lattice spacings ranging from about 0.15 to 0.06 fm. Since our earlier work, however, we have increased statistics significantly at our three finest lattice spacings. The new data give better control of the lattice-dependence of the Euclidean-time window observables and enable stringent tests of the EFT-based corrections, which inform our analysis of the associated systematic errors. We estimate the uncertainties on $a_\mu^{\rm W}$ and $a_\mu^{\rm W2}$ from making different, reasonable analysis choices for finite-volume corrections and treating discretization effects, among others, via Bayesian model averaging [46, 47]. Finally, to avoid confirmation bias, the Euclidean-time window observables were blinded until the analysis and error budgets were finalized. (See Sec. III A for details.)

This paper is organized as follows. First, in Sec. II A we provide analytic expressions for $a_\mu^{\rm W}$ and $a_\mu^{\rm W2}$ in terms of the Euclidean-time vector-current correlation function. Next, in Sec. II B we define the isospin-symmetric QCD limit employed here and describe our numerical correlator computations in Sec. II C. In Sec. III we present a detailed description of our analysis procedures, starting with how blinding was applied and removed in Sec. III A. Briefly, in Sec. III B, we use the lattice correlators to calculate the Euclidean-time windows corresponding to our numerical-simulation parameters. We then correct these "raw" window

[3] The RBC/UKQCD collaboration's update [38], which appeared on arXiv on the same day as our paper, is in good agreement with these recent results.

values on each ensemble for the finite lattice volume, slight mistunings of the simulated pion mass, and (optionally) remove taste-breaking discretization effects in Sec. III C. Next, we extrapolate the corrected window values to zero lattice spacing in Sec. III D. Sections III E and III F describe our procedure for Bayesian model averaging and the resulting systematic error budget. We conclude in Sec. IV by presenting our final results for a_μ^{W} and a_μ^{W2} and comparing them with previous determinations. Appendices A and B provide additional details on obtaining the Euclidean-time windows from staggered correlators and on computing corrections to the windows using the chiral model of pions, photons, and ρ mesons introduced in Ref. [48], respectively. Progress reports on related ongoing work can be found in Refs. [40–42, 49, 50].

II. PRELIMINARIES

A. Definitions of Euclidean-time window observables

The hadronic vacuum polarization function $\Pi(Q^2)$ can be obtained from Euclidean vector-current correlation functions through the equations

$$\Pi^{\mu\nu}(Q^2) = \left(\delta^{\mu\nu}Q^2 - Q^\mu Q^\nu\right)\Pi(Q^2) = \int \mathrm{d}^4x e^{iQx}\left\langle J^\mu(x)J^\nu(0)\right\rangle, \tag{2.1}$$

$$J^\mu(x) = \sum_f q_f \bar{\psi}_f(x)\gamma^\mu\psi_f(x), \tag{2.2}$$

where $J^\mu(x)$ is the electromagnetic current summed over quark flavors $f = \{u, d, s, c, b, t\}$, q_f are the corresponding electric charges in units of e, and $\langle J^\mu(x)J^v(0)\rangle$ includes both quark-line connected and disconnected Wick contractions. The HVP contribution to the muon's anomalous magnetic moment can then be obtained from a weighted integral of $\hat{\Pi}(Q^2) \equiv \Pi(Q^2) - \Pi(0)$ via Eq. (B1).

It is now standard for lattice-QCD HVP calculations, however, to employ the alternative time-momentum representation introduced by Bernecker and Meyer [31]. This formulation is more convenient for an inherently space-time approach such as lattice QCD and allows the construction of Euclidean-time windows. Starting with the spatial-vector-current correlation function $C(t)$, defined as

$$C(t) = \frac{1}{3}\sum_{\mathbf{x},k}\left\langle J^k(\mathbf{x},t)J^k(0)\right\rangle, \qquad k = 1,2,3, \tag{2.3}$$

$a_\mu^{\mathrm{HVP,LO}}$ is obtained via

$$a_\mu^{\mathrm{HVP,LO}} = 4\alpha^2\int_0^\infty \mathrm{d}t\, C(t)\tilde{K}(t), \tag{2.4}$$

$$\tilde{K}(t) = 2\int_0^\infty \frac{\mathrm{d}Q}{Q}\, K_E(Q^2)\left[Q^2t^2 - 4\sin^2\left(\frac{Qt}{2}\right)\right], \tag{2.5}$$

where $K_E(Q^2)$ is given in Eq. (B2).

The window observables are then easily obtained by introducing the window function $\mathcal{W}$, limiting the Euclidean-time region over which $C(t)$ is integrated [32]:

$$a_\mu^{\mathrm{win}(t_0,t_1,\Delta)} = 4\alpha^2\int_0^\infty \mathrm{d}t\, C(t)\tilde{K}(t)\mathcal{W}\left(t,t_0,t_1,\Delta\right), \tag{2.6}$$

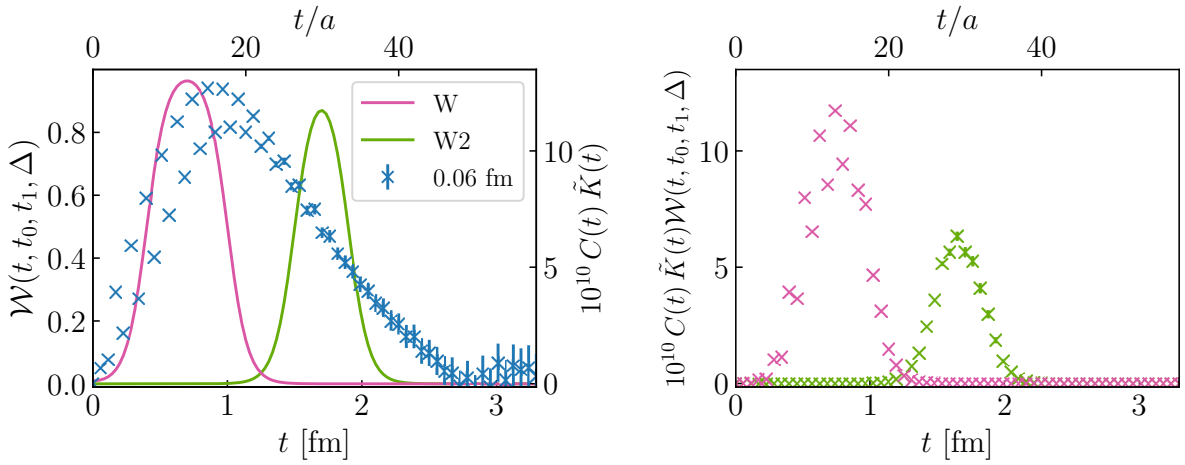

FIG. 1. (Left) The W (magenta) and W2 (green) window functions (corresponding to the parameters in Eqs. (2.8) and (2.9)) overlaid with raw lattice data for the integrand of Eq. (2.4) (blue crosses) from our finest ensemble. (Right) The windowed integrand of Eq. (2.6) for the corresponding window functions using the lattice data in the left panel.

$$\mathcal{W}\left(t, t_0, t_1, \Delta\right) = \frac{1}{2}\left[\tanh\left(\frac{t - t_0}{\Delta}\right) - \tanh\left(\frac{t - t_1}{\Delta}\right)\right] + (t \to -t)\,. \tag{2.7}$$

The parameters t_0 and t_1 of $\mathcal{W}$ control the location of the window's boundaries, while Δ controls the sharpness of its edges. In this work, we consider two such windows, the intermediate window W,

$$a_\mu^{\mathrm{W}} \equiv a_\mu^{\mathrm{win}(0.4,\ 1,\ 0.15)}, \tag{2.8}$$

and W2,

$$a_\mu^{\mathrm{W2}} \equiv a_\mu^{\mathrm{win}(1.5,\ 1.9,\ 0.15)}, \tag{2.9}$$

with the parameters in fm. We plot $\mathcal{W}$ in Eq. (2.7) for these window regions in the left panel of Fig. 1.

It is convenient in lattice-QCD calculations of $a_\mu^{\mathrm{HVP,LO}}$ to separately compute and then sum up the contributions from each quark flavor and from connected and disconnected Wick contractions. Here we focus on the light-quark connected contribution to the Euclidean-time windows, $a_\mu^{ll,\mathrm{W}}(\mathrm{conn.})$ and $a_\mu^{ll,\mathrm{W2}}(\mathrm{conn.})$, in the isospin-symmetric limit. Therefore, our electromagnetic vector current $J^\mu(x)$ includes only the terms for light quarks with $m_l = (m_u + m_d)/2$ and our correlation function $C(t)$ includes only the connected contractions. Figure 1, left, shows the light-quark connected contribution to the integrand [$\tilde{K}(t)C(t)$ in Eq. (2.4)] using the lattice correlation function obtained on our finest ensemble (see Sec. II C). Figure 1, right, shows the corresponding window integrands [$C(t)\tilde{K}(t)\mathcal{W}\left(t, t_0, t_1, \Delta\right)$ in Eq. (2.6)] for the W and W2 windows.

B. Prescription for isospin-symmetric QCD

Both the MILC HISQ gauge-field configurations and light-quark connected correlators employed in this work correspond to the isospin symmetric limit of QCD, *i.e.*, a pure-QCD

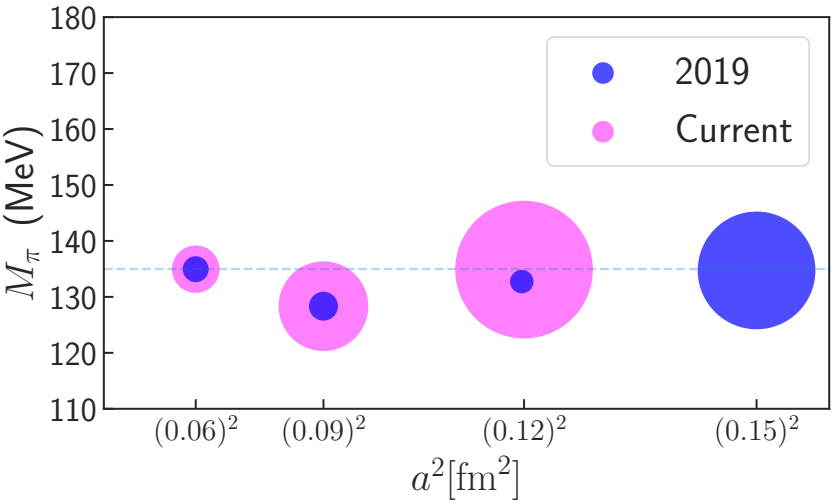

FIG. 2. Visualization of the ensemble parameters and statistics employed in this work (labeled "Current") and in our previous $a_\mu^{\mathrm{HVP,LO}}$ calculation [43] (labeled "2019"). Each disk is centered at the corresponding ensemble's squared lattice spacing and pion mass (a^2 and $M_{\pi 5}$ in Table I), while the disk areas are proportional to the size of each data set ($N_{\mathrm{conf}} \times N_{\mathrm{loose\ sources}}$ in Table I).

world with equal-mass up- and down-quark masses and without electromagnetism. Following the prescription introduced (for three flavors) in Refs. [51, 52] and later extended to include the charm quark in Ref. [53], we set the light-quark masses and lattice scale in physical units using the pion mass and decay constant. We then set the strange- and charm-quark masses using the kaon and D_s-meson masses, respectively.

Prior to the tuning procedure, however, electromagnetic effects must be removed from the experimental inputs. Neither M_{π^0} nor f_{π^+} are affected significantly by the quarks' electric charges, so their pure-QCD values are defined to be $M_\pi \equiv M_{\pi^0}$ and $f_\pi \equiv f_{\pi^+}$. In the Fermilab Lattice and MILC Collaboration's most recent analysis of pseudoscalar-meson masses and decay constants [54], the numerical values for these inputs were taken from the 2016 Particle Data Group: $M_{\pi^0} = 134.977$ MeV and $f_{\pi^+} = 130.50(1)_{\mathrm{exp.}}(3)_{V_{ud}}(13)_{\mathrm{EM}}$ MeV [55]. The remaining pure-QCD meson masses employed in Ref. [54] are $M_{K^0} = 497.567$ MeV, $M_{K^+} = 491.405$ MeV, and $M_{D_s} = 1967.02$ MeV. Details on how these values were obtained can be found in Sec. IV of that work and references therein. For the inputs in isospin-symmetric QCD we use the same values for M_{π^0} and f_{π^+} as above, while the kaon mass is defined as the average of M_{K^0} and M_{K^+}, giving $M_K = 494.486$ MeV.

C. Lattice-QCD ensembles and correlation functions

Our calculation employs the MILC Collaboration's four-flavor lattice-QCD configurations with dynamical up, down, strange, and charm quarks. The ensembles use the HISQ action [45] for the sea quarks, a Symanzik-improved gauge action [56–60] that includes the plaquette, the 1×2 rectangle, and the so-called bent-chair 6-link term for the gluon fields as well as tadpole improvement [61] based on the plaquette. Details of the configuration generation can be found in Ref. [62].

In this work, we employ a subset of the available MILC ensembles, for which the quark

TABLE I. Ensemble parameters used in this work. The first column lists the approximate lattice spacings in fm. The second column gives the spatial length L of the lattices in fm. The third column lists the volumes of the lattices in number of space-time points. The fourth column gives the sea-quark masses in lattice-spacing units. The fifth column lists the ratios of the gradient-flow scale w_0 [63] to the lattice spacing, where we take these values from Ref. [43] except for the newer ensemble with $a \approx 0.12$ fm. To convert simulation results to physical units, we take $w_0 = 0.1715(9)$ fm from Ref. [64]. The sixth column gives the taste-Goldstone pion masses [44]. The seventh column lists the renormalization factors for the local vector current, taken from Ref. [65]. The second-last column lists the number of configurations analyzed. The last column gives the number of loose-residual solves per configuration used in the truncated solver method [66, 67].

$\approx a$/fm	L/fm	$N_s^3 \times N_t$	$am_l^{\rm sea}/am_s^{\rm sea}/am_c^{\rm sea}$	w_0/a	M_{π_5}/MeV	Z_V	$N_{\rm conf}$	$N_{\rm loose}$
0.15	4.85	$32^3 \times 48$	0.002426/0.0673/0.8447	1.13215(35)	134.73(71)	0.9881(10)	9362	48
0.12	5.83	$48^3 \times 64$	0.001907/0.05252/0.6382	1.41110(59)	134.86(71)	0.9922(4)	9637	64
0.09	5.62	$64^3 \times 96$	0.00120/0.0363/0.432	1.95180(70)	128.34(68)	0.9940(5)	5384	48
0.06	5.46	$96^3 \times 128$	0.0008/0.022/0.260	3.0170(23)	134.95(72)	0.9950(6)	2621	24

masses are well tuned to their physical values. We include ensembles at four lattice spacings $a \approx 0.15$, 0.12, 0.09, and 0.06 fm. The high-statistics ensemble at $a \approx 0.15$ fm is unchanged from Ref. [43], where additional details can be found. The ensemble at $a \approx 0.12$ fm was generated specifically for our muon $g-2$ project, and has better tuned sea-quark masses compared with the ensemble with the same bare coupling used previously [43]. It contains about 10,000 configurations. We also extended the ensemble with $a \approx 0.09$ fm [44] to include over 5,000 configurations; this is a factor of roughly 3.5 beyond what was used in Ref. [43]. The pion mass for this ensemble is less accurately tuned than for the other three ensembles used in our study, which were generated more recently using quark masses obtained from a detailed analysis of pseudoscalar mesons and their decay constants [54]. Finally, we increased the number of configurations in our finest ensemble with lattice spacing $a \approx 0.06$ fm [54] by about a factor of two compared with Ref. [43]. We are continuing to extend this ensemble in anticipation of future needs. Our ensemble set is visualized in Fig. 2 and detailed in Table I.

The tuned quark masses listed in Table I are determined from the analysis in Ref. [54] in which pseudoscalar-meson masses and decay constants were computed using 24 gauge ensembles with six lattice spacings ranging from ≈ 0.15 to 0.03 fm. The pion decay constant is used to set the scale and the meson masses used to determine the up, down, strange, and charm masses are given in Sec. II B. Further details may also be found in Ref. [53] which used fewer configurations for a similar study.

The light-quark propagators from which the correlation functions $C(t)$ are constructed are computed using the HISQ action and truncated solver method (TSM) [66, 67]. Using random-wall sources, we compute one fine-residual conjugate gradient solve and the number of loose-residual solves ($N_{\rm loose}$) shown in the last column of Table I. Compared with Ref. [43], we have increased the number of loose sources per configuration by factors of 4, 3, and 1.5 for the ensembles at $a \approx 0.12, 0.09, 0.06$ fm, respectively. Exploiting time-reversal invariance, we further increase statistics by averaging the correlator values at times t and $N_t - t$ on each configuration. For the electromagnetic-current operator, we use the same local taste-vector vector current as in Ref. [43]. To match the local vector current to continuum QCD, we use the nonperturbatively computed renormalization factors obtained by the HPQCD

TABLE II. Additional valence-quark masses used to study the pion-mass dependence of a_μ^{W} and a_μ^{W2}. Simulation parameters not listed are the same as in Table I.

$\approx a/\mathrm{fm}$	am_q^{val}	$M_{\pi_5}^{\mathrm{val}}/\mathrm{MeV}$
0.15	0.001524	107.28(56)
	0.003328	157.16(83)
0.12	0.001190	107.05(56)
	0.002625	157.65(83)

Collaboration in Ref. [65]. Specifically, for the $a \approx 0.15, 0.12$, and 0.09 fm ensembles, we take the "H-H" Z_{V^4} values from Table IV, while for the $a \approx 0.06$ fm ensemble, we use the extrapolated value at this lattice spacing given in Appendix B of that work. To test our corrections for pion-mass mistuning (see Sec. III C 2), we generated additional vector- and pseudoscalar-current correlation functions with unphysical valence-quark masses on our two coarsest ensembles. Table II lists the valence-quark masses used in these partially quenched simulations.

In the course of our current analysis, we discovered two mistakes in Ref. [43] pertaining to the $a \approx 0.06$ fm ensemble. First, a small subset ($\approx 5\%$) of this ensemble's correlation functions were affected by a software bug in the data processing script. Second, the renormalization factor employed for this ensemble was taken from the arXiv version of Ref. [65], and differs from the published result by $\approx 0.1\%$. The latter error was realized after we unblinded our analysis (see Sec. III A). Hence, while keeping the analysis procedure frozen, we now use the published value of Z_{V^4} at $a \approx 0.06$ fm from Ref. [65] in our determinations of the window observables. We are preparing errata to Refs. [34, 43], but do not expect the results for $a_\mu^{ll}(\mathrm{conn.})$, $a_\mu^{\mathrm{HVP,LO}}$, or the one-sided Euclidean-time windows to change significantly.

III. DATA ANALYSIS

Here we present the analysis procedure to obtain the window observables $a_\mu^{ll,\mathrm{W}}(\mathrm{conn.})$ and $a_\mu^{ll,\mathrm{W2}}(\mathrm{conn.})$ at the physical, isospin-symmetric, pion mass and in the continuum and infinite-volume limits. First, as discussed in Sec. III A, we blinded the analyses of all components of $a_\mu^{\mathrm{HVP,LO}}$ to avoid unintentional introduction of bias. Second, because the correlation functions $C(t)$ are obtained at discrete Euclidean times, the integral in Eq. (2.6) must be approximated by a discrete integration rule. As described in Sec. III B, we use both the trapezoidal and Simpson's rules to quantify the associated discretization effects.

The resulting lattice data for $a_\mu^{ll,\mathrm{W}}(\mathrm{conn.})$ and $a_\mu^{ll,\mathrm{W2}}(\mathrm{conn.})$ must then be corrected for finite-volume effects, pion-mass mistunings and (optionally) taste-breaking effects. Our estimates of these lattice corrections to the intermediate and W2 window observables in Sec. III C are based on the use of EFTs and EFT-inspired models that capture the dominant low-energy, two-pion physics contribution to $a_\mu^{\mathrm{HVP,LO}}$. In particular, we consider variations obtained from four different approaches: next-to-leading-order (NLO) and next-to-next-to-leading order (NNLO) chiral perturbation theory (χPT) [29, 33, 68–71]; the Chiral Model (CM) [48] employed in Ref. [43]; the Meyer-Lellouch-Lüscher-Gounaris-Sakurai (MLLGS) approach [72–79]; and the relativistic pion EFT approach of Hansen and Patella (HP) [80]. In the absence of data-driven guidance, the spread of EFT-based corrections provides an

especially important indicator of the underlying uncertainties.

As is well known, staggered actions include additional, unphysical degrees of freedom (so-called "tastes"), yielding a 16-fold enlarged meson spectrum at finite lattice spacing [44, 62, 81, 82]. The splittings between the tastes are a lattice artifact that vanishes in the continuum limit. At finite lattice spacing, taste splittings of the pion masses are a significant discretization effect in $a_\mu^{\mathrm{HVP,LO}}$ observables (so-called taste-breaking effects). As discussed in Sec. III C, the splittings also affect the pion-mass and finite-volume dependencies, resulting in an interplay between them. In the case of pseudoscalar meson masses and weak matrix elements, discretization effects due to the taste-splittings are well-described by staggered chiral perturbation theory [83, 84], providing an additional handle on the continuum extrapolations.

Our continuum extrapolation analysis in Sec. III D includes a comprehensive study of taste breaking and other discretization effects. For each of the two window observables, we perform continuum extrapolations with and without first correcting for taste-breaking effects. In addition, we vary the fit function used for the continuum extrapolations, and we also include continuum-limit fits dropping the data at the coarsest lattice spacing ($a \approx 0.15$ fm). The fit function used in our continuum extrapolation contains the strong coupling constant α_s. In this work, following Ref. [54] we use $\alpha_s = \alpha_V(2/a)$ and take $\alpha_V(n_f = 4, \mu = 5.0\ \mathrm{GeV}) = 0.2530(38)$ from Ref. [85], where we evolve the coupling using the four-loop beta function.

Our lattice-QCD calculations of $a_\mu^{ll,\mathrm{W}}$(conn.) and $a_\mu^{ll,\mathrm{W2}}$(conn.) entail numerous analysis choices. As described in Sec. III E, we incorporate the systematic uncertainties due to these variations using Bayesian model averaging (BMA) [46, 47]. The remaining uncertainties in the corrected data sets include the statistical errors from the Monte-Carlo integration and parametric errors from w_0, w_0/a, and Z_V, which are propagated through the analysis as Gaussian random variables. Our final results for $a_\mu^{ll,\mathrm{W}}$(conn.) and $a_\mu^{ll,\mathrm{W2}}$(conn.) and error budgets from the respective BMA analyses are presented and discussed in Sec. III F.

A. Blinding

To avoid confirmation bias, we blinded this analysis until the systematic error budgets were finalized. The analysis was then frozen and used to generate the unblinded results and figures presented here. We employ a software blinding procedure, in which each observable is multiplied by an unknown random factor, chosen from a uniform distribution between $[0.7, 1.3]$. As the correlation function on the $a \approx 0.15$ fm ensemble is unchanged from Ref. [43], we additionally blinded the results from this ensemble by adding to it an offset equal to its standard deviation times an undisclosed random number between $[-1, 1]$. Each observable receives its own unique blinding factor, which is kept the same for all lattice spacings, except 0.15 fm. This procedure allows us to unblind specific sub-quantities, such as the intermediate window observables discussed here, without unblinding other quantities for which our analyses are ongoing.

B. Extraction of window observables

After the light-quark-connected vector-current correlation functions $C(t)$ are obtained on each ensemble as described in Sec. II C, lattice values for $a_\mu^{ll,\mathrm{W}}$(conn.) and $a_\mu^{ll,\mathrm{W2}}$(conn.) are

computed from these correlators via Eq. (2.6), using a chosen numerical integration scheme.

Because we employ a single-time-slice vector-current operator in our computations, the spectral representation of our staggered correlation functions consists of a sum of positive contributions from states with the desired parity and, additionally, contributions that oscillate in time as $(-1)^{t/a}$ from opposite-parity states (see Eq. (A1) of Appendix A). These oscillations are discretization effects, and should in principle be removed via the continuum extrapolation in Sec. III D. To quantify any residual uncertainty or bias on $a_\mu^{ll,\mathrm{W}}(\mathrm{conn.})$ and $a_\mu^{ll,\mathrm{W2}}(\mathrm{conn.})$ from oscillations in our correlation functions, we also perform our full analysis using the fit-reconstructed correlator without oscillations, $C_{\text{no osc.}}(t)$, which is defined in Eq. (A3). Appendix A also presents a number of alternate schemes for removing the unwanted oscillating terms from $C(t)$.

Numerical integration of the lattice correlators introduces additional discretization errors that depend upon the method used. Here we consider two integration schemes: the trapezoidal rule and Simpson's rule, which is formally higher order in the lattice spacing. Given a Euclidean-time correlator $C(t)$, windows of a_μ are obtained with the trapezoidal rule via

$$a_{\mu,\text{ Trap.}}^{\text{win}(t_0,t_1,\Delta)} = 4\alpha^2 a \sum_{t=1}^{N_t/2-1} C(t)\tilde{K}(t)\mathcal{W}\left(t,t_0,t_1,\Delta\right), \tag{3.1}$$

where the integration kernel $\tilde{K}(t)$ and window function $\mathcal{W}(t,t_0,t_1,\Delta)$ are given in Eqs. (2.5) and (2.7), respectively, and the boundary terms are omitted as $\tilde{K}(t)\mathcal{W}(t,t_0,t_1,\Delta) = 0$ for these cases. Similarly, Euclidean-time windows are obtained with Simpson's rule via

$$a_{\mu,\text{ Simp.}}^{\text{win}(t_0,t_1,\Delta)} = 4\alpha^2 \frac{a}{3}\left[\left(4\sum_{t\in\{t_{\text{odd}}\}}^{N_t/2-1} + 2\sum_{t\in\{t_{\text{even}}\}}^{N_t/2-1}\right) C(t)\tilde{K}(t)\mathcal{W}\left(t,t_0,t_1,\Delta\right)\right]. \tag{3.2}$$

Figure 3 compares lattice data for $a_\mu^{ll,\mathrm{W}}(\mathrm{conn.})$ (left) and $a_\mu^{ll,\mathrm{W2}}(\mathrm{conn.})$ (right) obtained by integrating $C(t)$ using the trapezoidal rule (blue squares), integrating $C_{\text{no osc.}}(t)$ using the trapezoidal rule (orange squares), and integrating $C_{\text{no osc.}}(t)$ using Simpson's rule (red circles). To enable meaningful comparisons between $a_\mu^{ll,\mathrm{W}}(\mathrm{conn.})$ (or $a_\mu^{ll,\mathrm{W2}}(\mathrm{conn.})$) at different lattice spacings, all data in these plots include corrections for the finite spatial volumes and pion-mass mistuning using the CM. (See Sec. III C for details.)

The impact of temporal oscillations in our staggered lattice correlators on $a_\mu^{ll,\mathrm{W}}(\mathrm{conn.})$ and $a_\mu^{ll,\mathrm{W2}}(\mathrm{conn.})$ can be assessed by comparing the data sets obtained from integrating both $C(t)$ and $C_{\text{no osc.}}(t)$ using the trapezoidal rule (blue and orange squares in Fig. 3, respectively). For $a_\mu^{ll,\mathrm{W}}(\mathrm{conn.})$, the trapezoidal-rule data sets are statistically indistinguishable at our two finest lattice spacings (see Table VI in Appendix A, which provides the correlated pairwise differences). Further, the differences between them decrease rapidly with the lattice spacing. For $a_\mu^{ll,\mathrm{W2}}(\mathrm{conn.})$, which corresponds to a later Euclidean time range, oscillations in the correlator from heavier opposite-parity states have largely died out (see Fig. 1). Consequently, the trapezoidal-rule data sets are statistically consistent on all ensembles and there is no clear lattice-spacing dependence in their correlated differences. As, for both $a_\mu^{ll,\mathrm{W}}(\mathrm{conn.})$ and $a_\mu^{ll,\mathrm{W2}}(\mathrm{conn.})$, the continuum extrapolations of the trapezoidal-rule data sets are in excellent agreement, we therefore conclude that temporal oscillations are an insignificant source of discretization error in our calculation.

Similarly, discretization errors stemming from the numerical integration can be estimated comparing the data sets obtained by integrating $C_{\text{no osc.}}(t)$ with either the trapezoidal rule

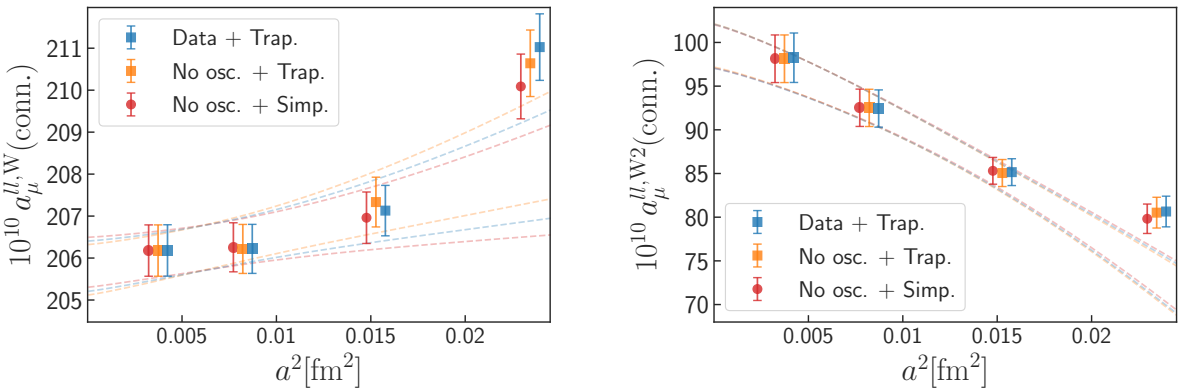

FIG. 3. Comparison of results for $a_\mu^{ll,\mathrm{W}}(\mathrm{conn.})$ (left) and $a_\mu^{ll,\mathrm{W2}}(\mathrm{conn.})$ (right) from integrating the raw lattice correlator $C(t)$ with the trapezoidal rule (blue squares), fit reconstruction with oscillating-state contributions removed $C_{\text{no osc.}}(t)$ with the trapezoidal rule (orange squares), and $C_{\text{no osc.}}(t)$ with Simpson's rule (red circles). Data at the same lattice spacing are offset horizontally for visibility. As described in Secs. III C 1 and III C 2, each point is corrected for finite-volume effects using the CM and pion-mass mistuning effects using the data-driven approach. For each integration scheme, we fit the data for the three finest ensembles to a function linear in $\alpha_s a^2$. The dashed curves show the fits' error bands, with colors matching the corresponding plot symbols.

or Simpson's rule (orange squares and red circles in Fig. 3, respectively). As is displayed in the figure and quantified in Table VII, on our coarse ensembles the differences between integration schemes are statistically significant for both $a_\mu^{ll,\mathrm{W}}(\mathrm{conn.})$ and $a_\mu^{ll,\mathrm{W2}}(\mathrm{conn.})$. These differences, however, decrease with lattice spacing much faster than a^2 and are already at the per-mille level at our finest lattice spacing. We therefore conclude that lattice artifacts from the choice of numerical integration scheme are negligible compared to the leading discretization terms in the Symanzik effective Lagrangian, which are of $\mathcal{O}(\alpha_s a^2)$ (see Sec. III D for details).

Based on these observables, we generate two data sets for each of the two observables ($a_\mu^{ll,\mathrm{W}}(\mathrm{conn.})$ and $a_\mu^{ll,\mathrm{W2}}(\mathrm{conn.})$). The first is obtained from integrating the original correlation function data $C(t)$ with the trapezoidal rule, and the second is obtained from integrating the reconstructed correlation function data $C_{\text{no osc.}}(t)$ with Simpson's rule. The inclusion of both data sets in the subsequent analysis accounts for any residual systematic effects due to both $\mathcal{O}(a^2)$ artifacts induced by the trapezoidal rule as well as the oscillating contributions that remain after the continuum extrapolation. For each observable, the two data sets are taken as inputs in the next step of the analysis, where the corrections are applied to the a_μ data.

C. Lattice corrections

Before taking the continuum limit, we correct our lattice $a_\mu^{ll,\mathrm{W}}(\mathrm{conn.})$ and $a_\mu^{ll,\mathrm{W2}}(\mathrm{conn.})$ data in two (or three) separate steps: first for finite volume (FV), second for pion-mass mistuning (M_π), and (sometimes) third for the effects of taste splittings (TB). The last step is optional, since changing discretization effects will only alter the window observables' lattice-

spacing dependence, not their continuum-limit values. Further, the pion-mass corrections to the intermediate and W2 windows are only numerically significant on the $a \approx 0.09$ fm ensemble for which the simulated pion mass is $\sim 5\%$ below the physical value (see Table II).

Mathematically, our correction scheme is defined via the following equations:

$$a_\mu \left(L_\infty, M_{\pi_{\rm phys}}\right) = a_\mu \left(L_{\rm latt}, M_{\pi_{\rm lat\ ,\xi_1}}, \cdots, M_{\pi_{\rm lat\ ,\xi_{16}}}\right) + \Delta_{\rm FV} + \Delta_{M_\pi} + \Delta_{\rm TB}, \tag{3.3}$$

where

$$\Delta_{\rm FV} = a_\mu \left(L_\infty, M_{\pi_{\rm lat,\xi_1}}, \cdots, M_{\pi_{\rm lat,\xi_{16}}}\right) - a_\mu \left(L_{\rm latt}\ , M_{\pi_{\rm lat,\xi_1}}, \cdots, M_{\pi_{\rm lat,\xi_{16}}}\right), \tag{3.4}$$

$$\Delta_{M_\pi} = a_\mu \left(L_\infty, M_{\pi_{\rm phys,\xi_1}}, \cdots, M_{\pi_{\rm phys,\xi_{16}}}\right) - a_\mu \left(L_\infty, M_{\pi_{\rm lat,\xi_1}}, \cdots, M_{\pi_{\rm lat,\xi_{16}}}\right), \tag{3.5}$$

$$\Delta_{\rm TB} = a_\mu \left(L_\infty, M_{\pi_{\rm phys}}\right) - a_\mu \left(L_\infty, M_{\pi_{\rm phys,\xi_1}}, \cdots, M_{\pi_{\rm phys,\xi_{16}}}\right). \tag{3.6}$$

"a_μ" is shorthand for either $a_\mu^{ll,\rm W}$(conn.) or $a_\mu^{ll,\rm W2}$(conn.). The first term on the right-hand-side of Eq. (3.3) is the window observable on each ensemble obtained in Sec. III B. The three corrections in Eqs. (3.4)–(3.6) are evaluated for each ensemble and added to the lattice values for a_μ. The first correction, $\Delta_{\rm FV}$ in Eq. (3.4), takes $a_\mu^{ll,\rm W}$(conn.) or $a_\mu^{ll,\rm W2}$(conn.) from the simulated spatial volume, indicated by $L_{\rm latt}$, to the infinite-volume limit, denoted by L_∞. The second correction, Δ_{M_π}, takes the simulated taste-Goldstone pion mass to the physical value, while the final correction, $\Delta_{\rm TB}$, removes the effects of the pion taste-splittings, $a^2\Delta_{\xi_i}$. In practice, the lattice and "physical" staggered-pion masses $M_{\pi_{\rm lat,\xi_i}}$ and $M_{\pi_{\rm phys,\xi_i}}$ in Eqs. (3.4)–(3.6) are calculated via the leading-order staggered χPT relationship $M^2_{\pi_{\xi_i}} = M^2_{\pi_{\xi_1}} + a^2\Delta_{\xi_i}$ with $M_{\pi_{\rm lat,\xi_1}}$ and $M_{\pi_{\rm phys,\xi_1}}$ fixed to the taste-Goldstone pion mass (column six of Table I) and the experimentally-measured π^0 mass, respectively.

The order in which the finite-volume, pion-mass, and taste-breaking corrections are applied impacts the form of the corrections. In our case, we apply the corrections in Eq. (3.3) from left to right. Therefore, the finite-volume and pion-mass mistuning corrections in Eqs. (3.4) and (3.5) must preserve the taste splittings. The left-hand-side of Eq. (3.3) is the infinite-volume, physical pion-mass and finite-lattice-spacing a_μ with correct physical parameters, which are inputs to the continuum extrapolations described in Sec. III D.

We employ four different effective-field-theory-based schemes for the finite-volume and taste-breaking corrections:

- Chiral Perturbation Theory (χPT) at next-to-leading order (NLO) and next-to-next-to-leading order (NNLO) [29, 69, 71]. The staggered NNLO χPT expressions of Refs. [29, 33, 71] are derived for a taste-singlet vector current, which couples to taste-diagonal pion pairs. We adapt the expressions of Ref. [33] to the taste-vector vector current employed here, which couples to taste-nondiagonal two-pion states. We replace the pion energies in Eq. (3.3) of Ref. [33] with averages of the energies of the two pions in the two-pion states which contribute to the taste-vector vector current.[4] We test this approximation for the case of NLO χPT (and for the CM and MLLGS approaches discussed below) where we have exact formulas for the taste-vector current. These tests reveal at most sub-percent differences in the corrections computed using the exact approach versus the approximation.

[4] For the term labeled NNLO,4 in the finite-volume correction of [29, 71], we substitute the average masses of the two-pion state into the energies instead of substituting the average energies. This avoids a numerical instability in this term for the case of unequal energies.

- The Chiral Model (CM) is an extension of χPT, where the ρ meson is included explicitly through a massive spin-1 vector field. This model was introduced by Jegerlehner and Szafron to study $\rho - \gamma$ mixing in $e^+e^- \to \pi\pi$ scattering [86]. It was first applied to Euclidean-space lattice-QCD calculations of the muon $g-2$ HVP (with modifications to incorporate the staggered-pion mass spectrum) by the HPQCD Collaboration [48].[5] The equations for the CM corrections computed in this analysis differ from Refs. [43, 48] slightly in that the effects of taste-breaking are included as specified in Eqs. (B8)–(B10).

- The Meyer-Lellouch-Lüscher-Gounaris-Sakurai (MLLGS) approach combines the pion form-factor parameterization of Gounaris-Sakurai with the mapping (due to Meyer-Lellouch-Lüscher [72–79]) between the infinite-volume scattering amplitude and finite-volume energies and amplitudes of the two-pion states. We account for taste-breaking effects in the same fashion as Ref. [29], by including contributions from two-pion states constructed with all 16 tastes of pions. Here, we modify the expressions of Ref. [29] to the case of the taste-vector vector current which couples to taste-nondiagonal two-pion states. As in [29], we fix the number of finite-volume states to $n = 8$.

- The relativistic-pion effective-field-theory approach by Hansen and Patella (HP) for finite-volume effects [80]. We obtain the correction defined in Eq. (3.4) using the same replacement as described in the χPT description above.

We describe the above schemes as "effective-field-theory-based" because, in some parts of our analysis, they may be employed outside the schemes' ranges of validity. Except for the CM, these EFTs and phenomenological models include only the contributions to $a_\mu^{\rm HVP,LO}$ observables from two-pion intermediate states.[6] Because contributions to the Euclidean-time correlation function fall off as $\exp(-Et)$ (see Appendix A), those from low-lying $\pi\pi$ states are most important at large Euclidean times. Consequently, the correction schemes listed above should best describe the volume and pion-mass dependence of $C(t)$ and, hence, $a_\mu^{{\rm win}(t_0,t_1,\Delta)}$, for later time ranges. Indeed, we and other collaborations find that, for $t_0 \gtrsim 1.5$ fm (which includes the 'W2' window), all of the higher-order correction schemes enumerated above (*i.e.*, excluding NLO χPT) yield similar predictions for the finite-volume, pion-mass, and taste-breaking corrections to $a_\mu^{{\rm win}(t_0,t_1,\Delta)}$. Further, the estimates from these schemes for the sum of finite-volume, pion-mass, and taste-breaking corrections reasonably describe the observed differences between lattice data in this region. [29, 33, 79, 87]. Therefore, they can be reliably used to calculate lattice corrections to $a_\mu^{ll,{\rm W2}}$(conn.). In the intermediate-window region, the predicted corrections from the EFT-based schemes display a wider variation. The sizes of the finite-volume and pion-mass corrections to $a_\mu^{ll,{\rm W}}$(conn.), however, are numerically small ($\Delta^{\rm W}_{{\rm FV},M_\pi} \lesssim 0.5\%$), and we incorporate the spread in $a_\mu^{ll,{\rm W}}$(conn.) results obtained with different correction schemes in our systematic error estimate in Sec. III E.

1. *Finite-volume corrections*

Figure 4 shows the finite-volume corrections to $a_\mu^{ll,{\rm W}}$(conn.) and $a_\mu^{ll,{\rm W2}}$(conn.) computed via Eq. (3.4) for each ensemble listed in Table I and the four correction schemes discussed

[5] In Refs. [29] and [33], the staggered Chiral Model is denoted "SRHO."

[6] Although χPT, MLLGS, and HP do not treat the ρ meson as a dynamical degree-of-freedom, they implicitly incorporate some resonance effects through parameters that are tuned to match experiment.

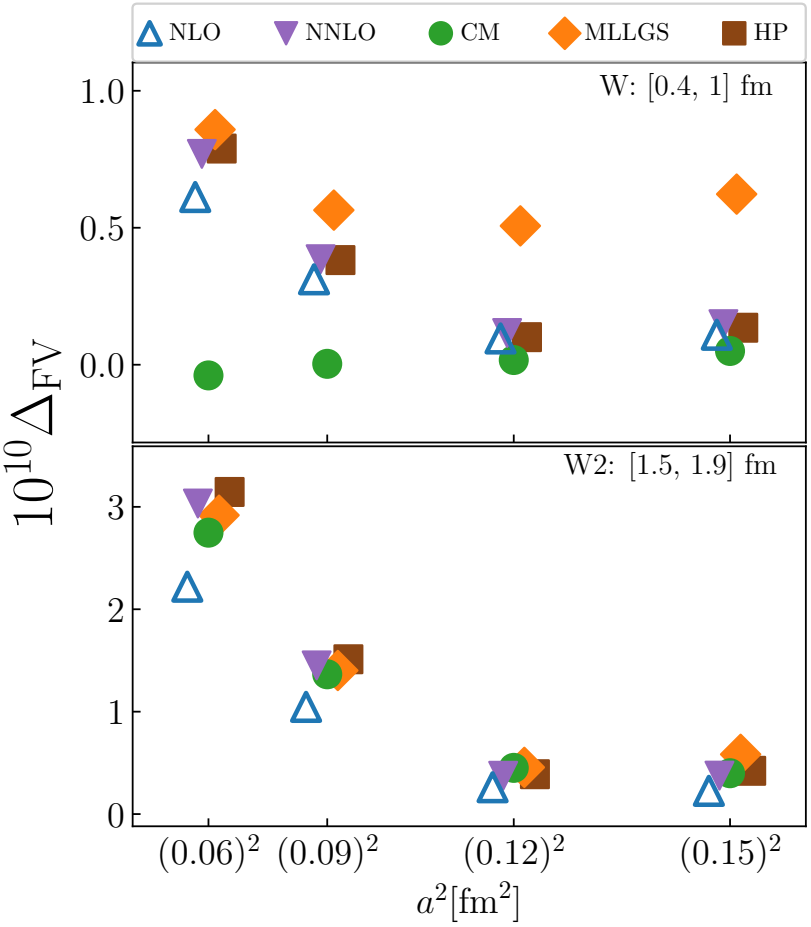

FIG. 4. Finite-volume corrections to $a_\mu^{ll,\mathrm{W}}$(conn.) (top) and $a_\mu^{ll,\mathrm{W2}}$(conn.) (bottom) obtained from NLO χPT (open blue triangles), NNLO χPT (purple downward triangles), CM (green circles), MLLGS (orange diamonds), and HP (brown squares). The data points at each lattice spacing are offset horizontally for visual clarity. Lattice spatial volumes are given in Table I.

above. For $a_\mu^{ll,\mathrm{W}}$(conn.) (top panel), the finite-volume corrections $\Delta_{\mathrm{FV}}^{\mathrm{W}}$ are always less than 0.5%. There is, however, a significant spread between the different schemes. In particular, the finite-volume corrections obtained from the CM (green circles) are close to zero on all ensembles. This is because in the CM, the renormalized vacuum polarization function, Eq. (B3), is comprised of two terms: the first is identical to NLO χPT, while the second accounts for ρ-π-π interactions. For $a_\mu^{ll,\mathrm{W}}$(conn.), the latter contribution produces a correction opposite in sign to the former. In contrast, in χPT, the NLO (open blue triangles) and NNLO contributions to Δ_{FV} have the same sign, making the total NNLO corrections (purple downward triangles) larger. The spread between the finite-volume corrections in the top panel of Fig. 4 reflects the limitations of the correction schemes in the intermediate window region, as discussed earlier.

By design [33], χPT (and the other EFTs) should work better in the W2 region, for which contributions from low-lying $\pi\pi$ states are more important. Hence, we expect better consistency between the finite-volume corrections to $a_\mu^{ll,\mathrm{W2}}$(conn.) from the different approaches. Indeed, these expectations are borne out in the bottom panel of Fig. 4, where the correc-

tions from the higher-order schemes have a much-reduced (relative) spread compared to $a_\mu^{ll,\mathrm{W}}$(conn.). Additionally, the finite-volume corrections to $a_\mu^{ll,\mathrm{W2}}$(conn.) are larger than for $a_\mu^{ll,\mathrm{W}}$(conn.) ($\Delta_{\mathrm{FV}}^{\mathrm{W2}} \sim 3\%$ at the finest lattice spacing) due to the increased sensitivity to long-distance contributions at later Euclidean times.

Below $a \lesssim 0.12$ fm, the size of finite-volume corrections to $a_\mu^{ll,\mathrm{W}}$(conn.) and $a_\mu^{ll,\mathrm{W2}}$(conn.) decrease with increasing lattice spacing. This is because the pion taste splittings are larger on coarser lattices, and finite-volume corrections in systems with heavier masses are smaller. The finite-volume corrections at $a \approx 0.15$ fm are generally larger than at $a \approx 0.12$ fm, however, because the spatial volume of our coarsest ensemble is substantially smaller than the others (see Table VII).

In the absence of guidance from a direct finite-volume study, we take the range of finite-volume corrections for the schemes we consider here as an estimate of the associated systematic uncertainty. Some or all of the EFT-based models considered are of questionable reliability in the intermediate-window region. Motivated by this, we generate a second set of corrections to $a_\mu^{ll,\mathrm{W}}$(conn.) obtained from restricting the window to higher t, namely $[0.7, 1.0]$ fm. The spread of these restricted corrections is $\sim 20\%$ smaller than the full W window case. Therefore, in total we include ten sets of finite-volume-corrected data for each input $a_\mu^{ll,\mathrm{W}}$(conn.) data set in our analysis: two each for NLO χPT, NNLO χPT, CM, MLLGS, and HP. For the W2 region, the EFTs are on more solid theoretical footing and the higher-order schemes (NNLO χPT, CM, MLLGS, and HP) yield consistent results. Therefore, in our analysis we include four sets of finite-volume-corrected data for $a_\mu^{ll,\mathrm{W2}}$(conn.), omitting NLO χPT because NNLO χPT should be more accurate in this region. These finite-volume-corrected data sets are inputs into the next step, and eventually feed into the BMA analysis of Sec. III E.

2. *Pion-mass adjustment*

We next consider the effects of pion-mass mistuning on $a_\mu^{ll,\mathrm{W}}$(conn.) and $a_\mu^{ll,\mathrm{W2}}$(conn.) and estimate the pion-mass adjustments to these quantities, Δ_{M_π} in Eq. (3.5), using a data-driven approach. As stated in Sec. II C, on our ensembles with $a \approx 0.15$ and 0.12 fm, in addition to the unitary correlation functions listed in Table I, we have partially quenched correlators (and hence lattice data for $a_\mu^{ll,\mathrm{W}}$(conn.) and $a_\mu^{ll,\mathrm{W2}}$(conn.)) with valence-quark masses bracketing the physical light quark (see Table II)[7]. Together with our unitary data at $a \approx 0.09$ and 0.06 fm, this allows us to predict the size of pion-mass adjustments to $a_\mu^{ll,\mathrm{W}}$(conn.) and $a_\mu^{ll,\mathrm{W2}}$(conn.) on all of our ensembles as follows.

First, we correct our entire dataset for finite-volume effects as described in Sec. III C 1. We then fit the corrected $a_\mu^{ll,\mathrm{W}}$(conn.) and $a_\mu^{ll,\mathrm{W2}}$(conn.) data to an interpolating function of the form

$$a_\mu^{ll,\mathrm{win}}\big(a, M_\pi\big) = \sum_{i=-1}^{1} c_i(a)(M_\pi/\Lambda)^{2i};$$
$$c_i(a) = \sum_{j=0}^{n_i} c_{ij}(a\Lambda)^{2j}, \tag{3.7}$$

[7] The $a \approx 0.15$ fm correlators were also employed in Ref. [39] to study strong-isospin-breaking effects in a_μ^{ll}(conn.).

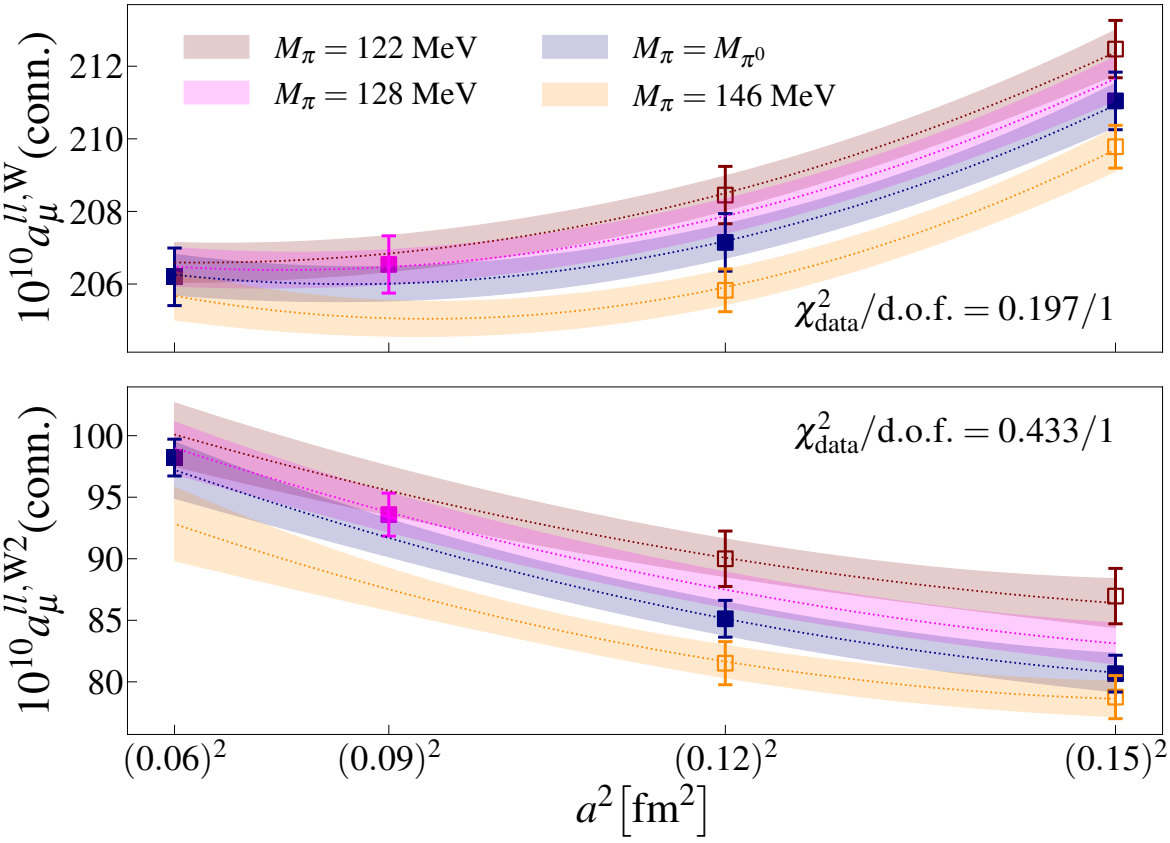

FIG. 5. Data-driven estimate of pion-mass adjustments to $a_\mu^{ll,\mathrm{W}}(\mathrm{conn.})$ (top) and $a_\mu^{ll,\mathrm{W2}}(\mathrm{conn.})$ (bottom) on all ensembles. Lattice data (corrected for finite-volume effects using the CM scheme) are shown as open/filled squares with error bars, with each color denoting a different simulation pion mass: M_{π^0} (blue), the taste-Goldstone pion mass at $a \approx 0.09$ fm (magenta), and the partially-quenched pion masses bracketing M_{π^0} at $a \approx 0.12$ and 0.15 fm (maroon and orange). The results of fitting these data to an interpolating function in M_π^2 and a^2 (specifically, Eq. (3.7) with $n_{\{-1,0,1\}} = \{1, 2, 1\}$) are shown for fixed M_π as dashed curves with error bands, and share the same color coding as the data points.

where $\Lambda = 500$ MeV (following Ref. [43]) and M_π is the taste-Goldstone valence-*sea* pion mass, which is what enters the leading-order pion loops in χPT. The parametric dependence on M_π in Eq. (3.7) is motivated by χPT, with an additional $1/M_\pi^2$ term accounting for the expected infrared-divergent behavior of a_μ^{ll} in the $M_\pi \to 0$ limit [48, 88]. For each value of i in Eq. (3.7), we consider several values for $n_i \geq 0$, requiring only that the $\{n_i\}$ are the same for both $a_\mu^{ll,\mathrm{W}}(\mathrm{conn.})$ and $a_\mu^{ll,\mathrm{W2}}(\mathrm{conn.})$, and that each fit has at least one degree-of-freedom (d.o.f.). (Note that if $n_i = 0$, then $j = 0$ and $c_i(a)$ is independent of $a\Lambda$.) Following Ref. [89], we account for correlations between the independent variables (a and M_π), as well as between the independent and dependent variables ($a_\mu^{ll,\mathrm{W}}(\mathrm{conn.})$ and $a_\mu^{ll,\mathrm{W2}}(\mathrm{conn.})$), using Bayesian priors. We monitor the $\chi^2_{\mathrm{data}}/\mathrm{d.o.f.}$ of each fit variation, preferring fits with $\chi^2_{\mathrm{data}}/\mathrm{d.o.f.}$ closest to 1. (A $\chi^2_{\mathrm{data}}/\mathrm{d.o.f.} \gg 1$ indicates that the fit function does not describe the data, while a $\chi^2_{\mathrm{data}}/\mathrm{d.o.f.} \ll 1$ suggests that we are overfitting.) After trying several combinations of $\{n_i\}$, we select $n_{-1} = 1$, $n_0 = 2$, and $n_1 = 1$ for our central analysis because this functional form gives the best interpolation of our data for both Euclidean-time windows

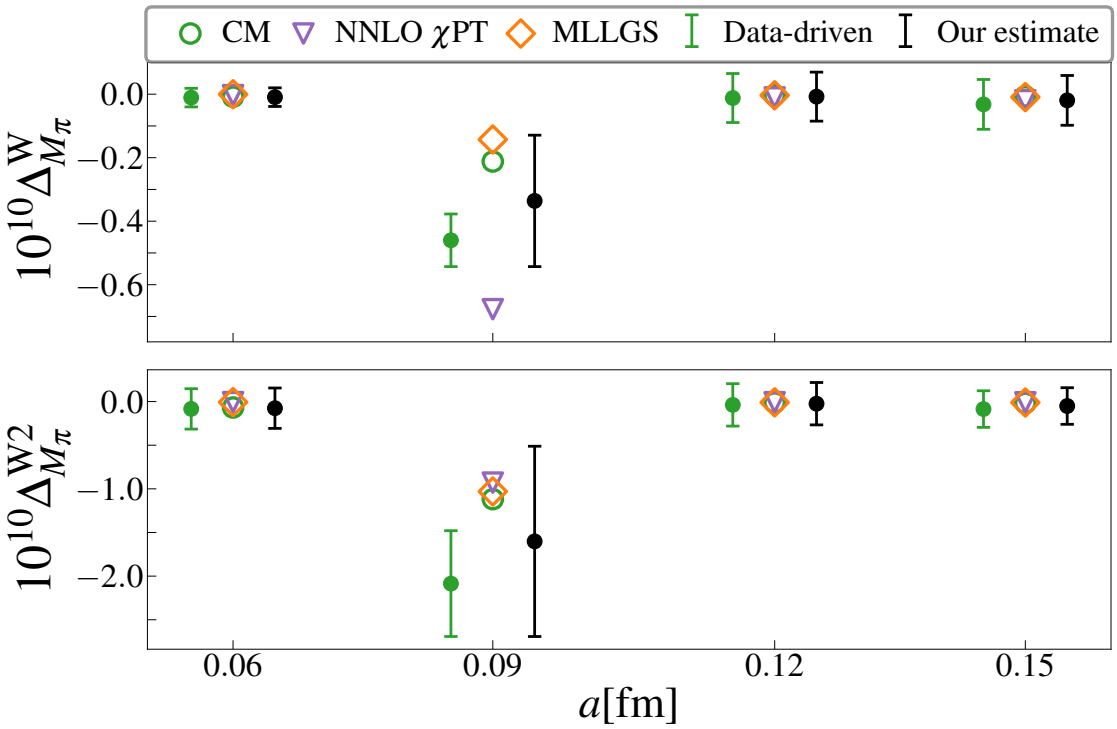

FIG. 6. Comparison of predictions for the pion-mass adjustments with error bars show the predictions of our data-driven analysis, which employs finite-volume corrections from the CM. Open symbols show predictions from the CM (green circles), NNLO χPT (empty purple upside down triangle), MLLGS (orange diamonds) correction schemes. Black points with error bars show our final estimates for the pion-mass adjustments on each ensemble, which account for the spread between predictions as described in the text.

simultaneously.

Once the coefficients c_{ij} are determined for a given set of $\{n_i\}$, we can use Eq. (3.7) to predict $a_\mu^{ll,\mathrm{W}}(\mathrm{conn.})$ and $a_\mu^{ll,\mathrm{W2}}(\mathrm{conn.})$ at the target physical pion mass, $M_{\pi,\mathrm{phys}} = M_{\pi^0}$ (see Sec. III B) for each ensemble. Figure 5 shows our central fits for $a_\mu^{ll,\mathrm{W}}(\mathrm{conn.})$ (upper panel) and $a_\mu^{ll,\mathrm{W2}}(\mathrm{conn.})$ (lower panel). At each lattice spacing, we take the difference in $a_\mu^{ll,\mathrm{win}}$ between the fit prediction at the physical-pion mass (blue dashed curve) and the unitary lattice data (filled squares) as our data-driven estimate of the pion-mass adjustment $\Delta_{M_\pi,\mathrm{DD}}$. As seen in Fig. 5, our data-driven analysis finds that a correction to $a_\mu^{ll,\mathrm{W}}(\mathrm{conn.})$ and $a_\mu^{ll,\mathrm{W2}}(\mathrm{conn.})$ of about 1 sigma is needed on the $a \approx 0.09$ fm ensemble, for which the simulation pion mass is about 5% below the physical value.

The $a_\mu^{ll,\mathrm{W}}(\mathrm{conn.})$ and $a_\mu^{ll,\mathrm{W2}}(\mathrm{conn.})$ data entering our central fits are corrected for finite-volume effects using the CM. Repeating this analysis using other finite-volume correction schemes yields almost identical predictions for the pion-mass adjustments. Replacing the $1/M_\pi^2$ term in Eq. (3.7) with $\log\left(M_\pi^2\right)$ also leads to negligible changes in the predicted values for Δ_{M_π}.

Figure 6 compares the pion-mass adjustments to $a_\mu^{ll,\mathrm{W}}(\mathrm{conn.})$ (upper panel) and $a_\mu^{ll,\mathrm{W2}}(\mathrm{conn.})$ (lower panel) obtained in our data-driven analysis (filled green circles with error bars) and those estimated within three of the EFT-based correction schemes introduced in Sec. III C:

the CM (empty green circles), NNLO χPT (empty purple upside down triangle), MLLGS (empty orange square triangles). On the three ensembles for which the pion mass is well tuned ($a \approx 0.06$, 0.12, and 0.15 fm), the pion-mass adjustment $\Delta_{M_\pi,\mathrm{DD}}$ is negligible in all correction schemes. At $a \approx 0.09$ fm, however, the picture is less clear. For $a_\mu^{ll,\mathrm{W}}$(conn.), the spread in model estimates is significantly larger than the error bar on the data-driven evaluation. For $a_\mu^{ll,\mathrm{W2}}$(conn.), the models agree with each other, but differ from the data-driven prediction by $\approx 1.5\sigma$.

In light of the differences between predicted corrections at $a \approx 0.09$ fm, we adopt the following conservative procedure to obtain our final estimates for Δ_{M_π} (black filled circles with error bars in Fig. 6). For the central value, we use the average of the data-driven and chiral-model predictions, *i.e.*, $\Delta_{M_\pi} \approx (\Delta_{M_\pi,\mathrm{DD}} + \Delta_{M_\pi,\mathrm{CM}})/2$. For the error on $a \approx 0.09$ fm, we add (linearly) to the uncertainty on the data-driven prediction $\sigma_{M_\pi,\mathrm{DD}}$ an additional systematic uncertainty given by half the absolute difference between the data-driven and chiral-model predictions, *i.e.*, $\sigma_{M_\pi,\mathrm{DD}} + |\Delta_{M_\pi,\mathrm{DD}} - \Delta_{M_\pi,\mathrm{CM}}|/2$. On $a \approx 0.06, 0.12$, and 0.15 fm we take the uncertainty to be just the uncertainty on the data-driven prediction. As shown in Fig. 6, our final estimates for the pion-mass adjustment at $a \approx 0.06$, 0.12, and 0.15 fm are essentially those from our data-driven analysis. At $a \approx 0.09$ fm, our final estimate for the pion-mass adjustment covers most (all) of the model spread for $a_\mu^{ll,\mathrm{W}}$(conn.) ($a_\mu^{ll,\mathrm{W2}}$(conn.)).

3. *Taste-breaking corrections*

The final lattice correction, Δ_{TB} in Eq. (3.6), accounts for the mass differences at finite lattice spacing between staggered pions with different taste quantum numbers. For the HISQ action, these taste splittings arise from discretization effects of $O(\alpha_s^2 a^2)$ and higher. It is well known, however, that the HISQ taste splittings do not scale linearly with $\alpha_s^2 a^2$ [44, 62].[8] As shown in Ref. [44], the HISQ pion-taste splittings decrease faster than naive expectations at lattice spacings below around 0.09 fm, while increasing more slowly at very coarse lattice spacings above roughly 0.12 fm. The former observation is likely due to the HISQ smearing [45] suppressing the leading $\alpha_s^2 a^2$ taste-breaking discretization contributions, making higher-order, *e.g.*, $\mathcal{O}(\alpha_s^3 a^2, a^4)$, effects more prominent, while the latter indicates the presence of additional higher-order terms.

Figure 7 shows the taste-breaking corrections to $a_\mu^{ll,\mathrm{W}}$(conn.) (left panel) and $a_\mu^{ll,\mathrm{W2}}$(conn.) (right panel) obtained within the χPT, CM, and MLLGS correction schemes introduced at the beginning of Sec. III C. Qualitatively, they display the same behavior as the taste splittings, with Δ_{TB} decreasing more rapidly at finer lattice spacings. Quantitatively, the estimated corrections span a wide range of values between $0 \lesssim \Delta_{\mathrm{TB}} \lesssim 30 \times 10^{-10}$. This corresponds to corrections to $a_\mu^{ll,\mathrm{W}}$(conn.) and $a_\mu^{ll,\mathrm{W2}}$(conn.) on our coarsest ensemble of up to $\sim 15\%$ and $\sim 30\%$, respectively. Differences between correction schemes and sizes of the corrections vanish at zero lattice spacing by construction.

For $a_\mu^{ll,\mathrm{W}}$(conn.), the NLO χPT, CM, and MLLGS results are in broad agreement, while the NNLO χPT prediction is about 2–4 times larger for $a \gtrapprox 0.09$ fm. In contrast, for $a_\mu^{ll,\mathrm{W2}}$(conn.) the three higher-order schemes (NNLO χPT, CM, and MLLGS) predict similar corrections, while NLO χPT is the outlier. As with the finite-volume corrections (see Sec. III C 1), the spread of predicted taste-breaking corrections for $a_\mu^{ll,\mathrm{W}}$(conn.) is likely due to the correction schemes becoming less reliable at short distances. Indeed, this expectation

[8] See also Ref. [33] for a discussion of the HISQ taste splittings as they pertain to calculations of $a_\mu^{\mathrm{HVP,LO}}$.

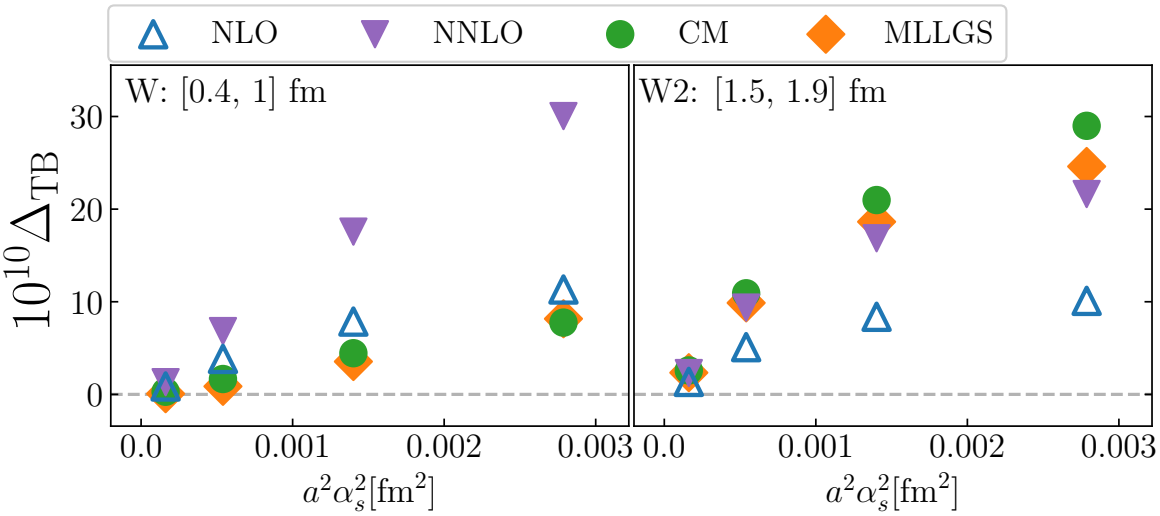

FIG. 7. Taste-breaking corrections to $a_\mu^{ll,\mathrm{W}}$(conn.) (left) and $a_\mu^{ll,\mathrm{W2}}$(conn.) (right) obtained from NLO χPT (open blue triangles), NNLO χPT (purple downward triangles), CM (green circles), and MLLGS (orange diamonds).

is borne out in Figs. 8 and 9, which show the continuum extrapolations of $a_\mu^{ll,\mathrm{W}}$(conn.) and $a_\mu^{ll,\mathrm{W2}}$(conn.), respectively, with and without taste-breaking corrections. For $a_\mu^{ll,\mathrm{W}}$(conn.), applying taste-breaking corrections increases the lattice-spacing dependence. In contrast, for $a_\mu^{ll,\mathrm{W2}}$(conn.) the taste-breaking corrections computed in the three higher-order correction schemes all substantially reduce the lattice-spacing dependence, indicating that they capture the dominant discretization effects in this window.

Although the inclusion of taste-breaking corrections (and choice of scheme) will alter the lattice-spacing dependence of $a_\mu^{ll,\mathrm{W}}$(conn.) and $a_\mu^{ll,\mathrm{W2}}$(conn.), it should not change the continuum-limit values. Consequently, varying the treatment of taste-breaking in the continuum extrapolation provides an additional measure of the continuum-extrapolation error. For the analysis of the intermediate window, we generate taste-breaking-corrected data sets with NLO χPT, NNLO χPT, CM, and MLLGS corrections for each input $a_\mu^{ll,\mathrm{W}}$(conn.) data set from the previous sections. We follow the reasoning of Sec. III C 1 and compute the corrections in two regions, the full intermediate window interval [0.4, 1] fm and the smaller interval of [0.7, 1] fm, resulting in a total of eight taste-breaking-corrected data sets for each input set. For each input $a_\mu^{ll,\mathrm{W2}}$(conn.) data set we generate three sets of taste-breaking-corrected data, one each for NNLO χPT, CM, and MLLGS, dropping NLO χPT, as in Sec. III C 1. In both cases, we also keep the data sets uncorrected for taste-breaking effects. The corrected and uncorrected $a_\mu^{ll,\mathrm{W}}$(conn.) and $a_\mu^{ll,\mathrm{W2}}$(conn.) data sets are then taken as inputs into the continuum limit extrapolations, and feed ultimately into the Bayesian model averaging analysis of Sec. III E.

D. Continuum extrapolation

To perform our continuum extrapolation we consider fit functions of the form:

$$a_\mu^{ll}(a,\{m_f\}) = a_\mu^{ll}\left(1 + F^{\mathrm{disc.}}(a) + F^{m}(\{\delta m_f\})\right), \tag{3.8}$$

where

$$F^{\text{disc.}}(a) = C_{a^2,n}\left[(a\Lambda)^2\alpha_s^n\right] + C_{a^4}(a\Lambda)^4 + C_{a^6}(a\Lambda)^6 \tag{3.9}$$

$$F^m(\{\delta m_f\}) = C_{\text{sea}} \sum_{f=l,l,s} \delta m_f/\Lambda. \tag{3.10}$$

The function $F^{\text{disc.}}(a)$ describes discretization effects and $F^m(\{\delta m_f\})$ accounts for quark mass differences in the sea, where δm_f is the difference between the physical and the simulation quark masses (see Secs. II B and II C). In $F^{\text{disc.}}(a)$ we include variations where the coefficient C_{a^6} is set to zero and where the power of α_s in the a^2 term varies as $n = 1, 2$. For both variations of $n = 1, 2$, we label fit functions with $C_{a^6} = 0$ as "quadratic" and fit functions where all terms listed in $F^{\text{disc.}}(a)$ are included as "cubic". Following Ref. [43] we take $\Lambda = 500$ MeV and impose the Gaussian prior constraint $C_{\text{sea}} = 0.0(3)$. Here, the C_{sea} term accounts for residual light sea-quark mass miss-tuning effects, remaining after performing the correction in Sec. III C 2, and also strange sea-quark miss-tuning effects. As in Ref. [43], the sea-quark masses in the ensembles employed here are so close to their physical values that our fits are insensitive to the $F^m(\{\delta m_f\})$ term and return posteriors for C_{sea} with central values close to zero and uncertainties close to the initial prior width. This also means that higher-order terms involving δm_f can be safely neglected. Additionally, to regulate the degrees of freedom in fits with an a^6 term, we constrain its coefficient with the Gaussian prior

$$C_{a^6} = 0(2). \tag{3.11}$$

This prior width conservatively accommodates instances among the over two thousand continuum fits when the posterior central values are close to or slightly larger than unity. We also include fits to three ensembles, dropping the coarsest. In this case we include an additional prior constraint on the quadratic term $C_{a^4} = 0(2)$ with the same reasoning as above for C_{a^6}. We include continuum-extrapolation fit variations, both with and without including the C_{sea} term in our Bayesian averaging process.

As illustration, in Fig. 8 we show results for quadratic continuum extrapolations of the $a_\mu^{ll,\text{W}}$(conn.) data with $n = 1$ in Eq. (3.8). The four panels show finite-volume-corrected data computed from the CM (top left), NLO χPT (top right), NNLO χPT (bottom left) and MLLGS (bottom right). For each scheme, we compare continuum extrapolations of data with and without taste-breaking corrections, where we include fits to all four ensembles as well as fits to ensembles at only the three finest lattice spacings. For the data sets corrected with NLO χPT or the CM, we find very good agreement between the four continuum extrapolated results, whereas the NNLO χPT- and MLLGS-corrected data sets show larger spreads. Taking into account this variance, we find that the continuum results obtained with all four correction schemes are consistent with each other. The corresponding continuum extrapolations for $a_\mu^{ll,\text{W2}}$(conn.) are shown in Fig. 9. In this case we find good agreement between the continuum extrapolated results, both within each scheme as well as across the different schemes, albeit with larger uncertainties. We observe, however, that the NLO χPT taste-breaking corrections do a poor job at removing lattice spacing dependence compared to the higher-order schemes.

In summary, for each input data set, we perform twelve different continuum extrapolations, the results of which become inputs to the Bayesian model averaging analysis of Sec. III E: in Eq. (3.8), we take $n = 1, 2$ in the linear term, and include or don't include the

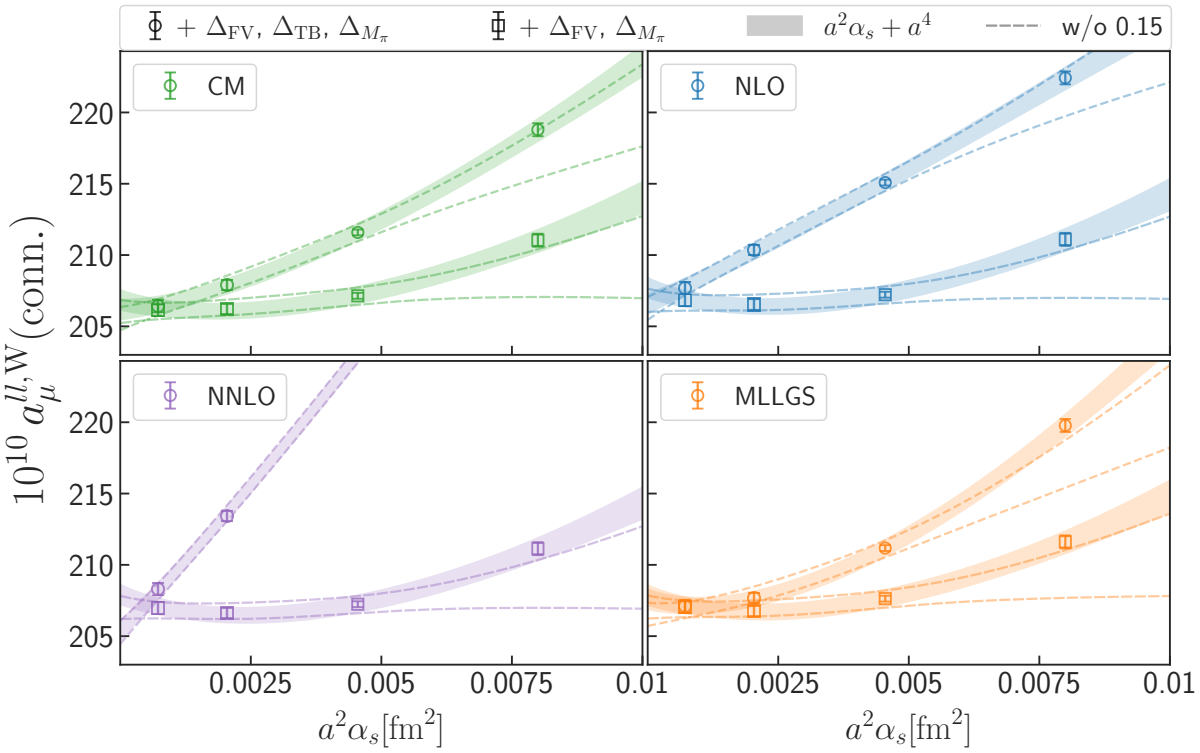

FIG. 8. Continuum extrapolations of $a_\mu^{ll,\rm W}$(conn.) using the CM (top left), NLO χPT (top right), NNLO χPT (bottom left) and MLLGS (bottom right) correction schemes. All data are obtained from integrating the lattice correlator $C(t)$ using the trapezoidal rule, and corrected for finite-volume effects and adjusted for pion-mass mistuning. Data sets that also include taste-breaking corrections are shown as circles, while data without these optional corrections are shown as squares. All corrections come from the full window region. Solid bands (dashed lines) show the fit results of continuum extrapolations with (without) data at our coarsest lattice spacing (right-most point in each panel). All fits employ the same fit function, Eq. (3.8) with terms through $O(a^4)$.

$C_{\rm sea}$ term[9]. With these four variations, we perform fits to the data at four lattice spacings with and without the cubic (a^6) term in Eq. (3.8), as well as with quadratic fits to data at the finest three lattice spacings.

Separately, as an independent analysis cross check of our continuum extrapolations and associated error estimate, we allow for higher-order terms in $a_\mu^{ll,\rm W}$(conn.) and $a_\mu^{ll,\rm W2}$(conn.) using the empirical Bayes (or maximum marginal likelihood) approach described in Sec. 5.2 of Ref. [90]. For this analysis, the discretization term $F^{\rm disc.}(a)$ in Eq. (3.8) takes the form:

$$F_{\rm alt}^{\rm disc.}(a) = \sum_{j=1}^{3} c_{1j} x \alpha_s^j + \sum_{i=2}^{5}\sum_{j=0}^{3} c_{ij} x^i \alpha_s^j + c_{60} x^6, \qquad x \equiv \left(\frac{Q_{\rm eff}}{\pi/a}\right)^2. \tag{3.12}$$

The coefficients in Eq. (3.12) are constrained with Gaussian priors $c_{ij} = 0(1)$, while the scale $Q_{\rm eff}$ is chosen to maximize the Gaussian Bayes Factor, Eq. (28) of Ref. [90], which is proportional to the marginal likelihood (model evidence). The results of this comparison are discussed in Sec. III E.

[9] We find that the fit results are virtually unchanged when the $C_{\rm sea}$ term is included and, in addition, that they are insensitive to the prior width of $C_{\rm sea}$, after increasing it by up to a factor of ten.

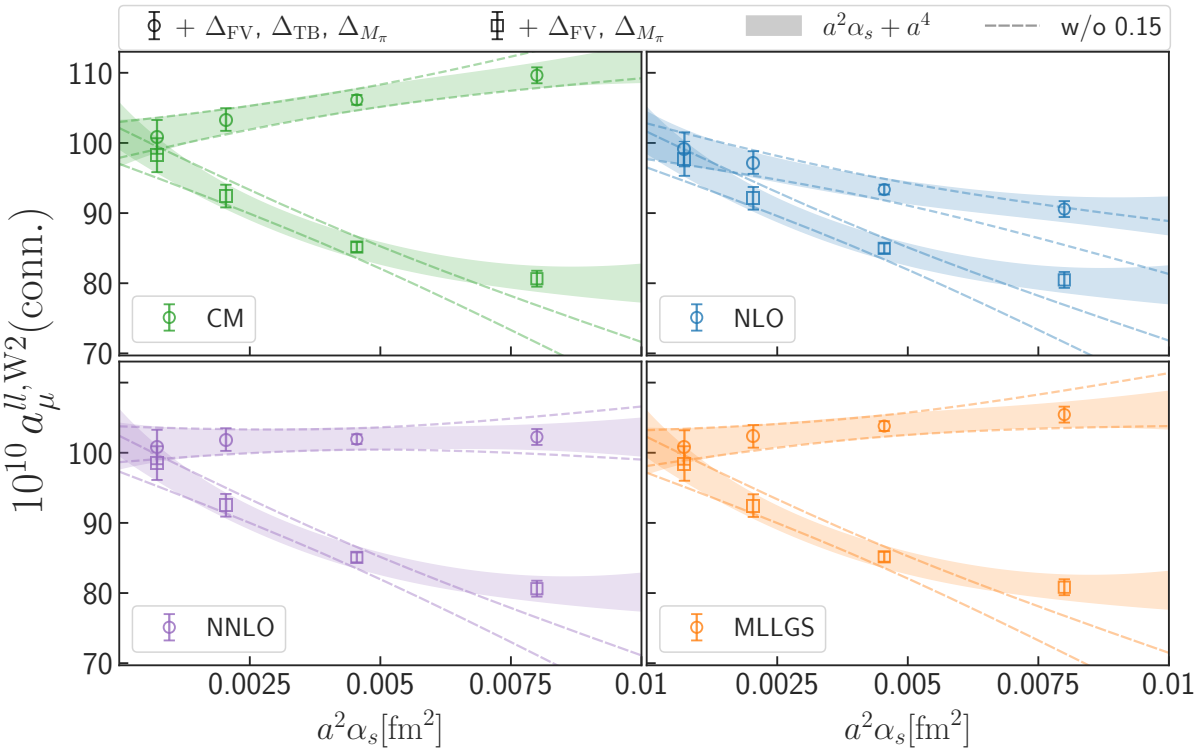

FIG. 9. Continuum extrapolations of $a_\mu^{ll,\mathrm{W2}}(\mathrm{conn.})$. Figure is described in the caption of Fig. 8.

E. Bayesian model averaging

In order to quantify the systematic uncertainty due to the analysis choices described in the previous sections, we employ Bayesian model averaging (BMA) [46, 47]. Summarizing these choices, we include variations of:

- **Observable extraction** - Two methods are used to extract the uncorrected values of $a_\mu^{ll,\mathrm{W}}(\mathrm{conn.})$ and $a_\mu^{ll,\mathrm{W2}}(\mathrm{conn.})$ from the correlation function data, as described in Sec. III B:
  - Raw correlation function data, $C(t)$, integrated with the trapezoidal rule.
  - Fit-reconstructed correlation function data $C_{\mathrm{noosc.}}(t)$ integrated with Simpson's rule.
- **Finite-volume correction** - All correction schemes discussed in Sec. III C 1 above: χPT, CM, MLLGS, and HP. We include the NLO χPT variation for $a_\mu^{ll,\mathrm{W}}(\mathrm{conn.})$ but not $a_\mu^{ll,\mathrm{W2}}(\mathrm{conn.})$.
- **Taste-breaking correction** - We include χPT, CM, and MLLGS as well as a_μ data sets which are not corrected for taste-breaking effects prior to continuum extrapolation.
- **Correction region** - For $a_\mu^{ll,\mathrm{W}}(\mathrm{conn.})$, we include a variation on the corrections where they are computed from the range $[0.7, 1]$ fm instead of over the full W window interval.
- **Continuum fit** - We perform continuum extrapolations using all 12 fit function variations described in Sec. III D including fits to the three finest ensembles.

In the context of BMA, a "model" M is defined as the set of analysis choices that yield a given result for the desired continuum, infinite-volume, physical observable from a single data set D. In our case, M is given by a set of choices from the options listed above, while D consists of the unmodified correlation function data.[10] In order to carry out the averaging, each M is assigned a probability weight given by

$$\mathrm{pr}(M \mid D) \equiv \mathrm{pr}(M) \exp\left[-\frac{1}{2}\left(\chi^2_{\mathrm{data}}\left(\mathbf{a}^\star\right) + 2k + 2N_{\mathrm{cut}}\right)\right]. \tag{3.13}$$

This is the "Bayesian Akaike information criterion" (BAIC) as defined in [47]. Here, χ^2_{data} is the standard chi-squared function, *not* including the contribution of the priors, and $\mathbf{a}^\star$ is the posterior mode (*i.e.*, the best-fit point for the vector of fit parameters $\mathbf{a}$ when optimized against the augmented chi-squared function [90].) N_{cut} is the number of data points cut from a data set—in this case, the number of ensembles omitted from a given extrapolation. The parameter k is the number of independent parameters in a given fit function. The factor $\mathrm{pr}(M)$ is the prior probability of a given M; we adopt a flat prior, so that this factor is a constant over all analysis variations and drops out of the model averaging results. The BMA mean and variance are then obtained from the following formulas:

$$\langle a_\mu \rangle = \sum_{i=1}^{N_M} \langle a_\mu \rangle_i \, \mathrm{pr}\left(M_i \mid D\right), \tag{3.14}$$

$$\sigma^2_{a_\mu} = \sum_{i=1}^{N_M} \sigma^2_{a_\mu,i} \mathrm{pr}\left(M_i \mid D\right) + \sum_{i=1}^{N_M} \langle a_\mu \rangle_i^2 \, \mathrm{pr}\left(M_i \mid D\right) - \left(\sum_{i=1}^{N_M} \langle a_\mu \rangle_i \, \mathrm{pr}\left(M_i \mid D\right)\right)^2. \tag{3.15}$$

The first term on the right-hand side of Eq. (3.15) is a weighted average over the variances of the individual results. The second and third terms reflect the spread in results obtained with different analysis choices (in our case, correction schemes and fit functions). Because they encapsulate the systematic uncertainty due to analysis choices, we refer to their sum as the "model variance."

In Fig. 10, we show the results of the Bayesian model average for $a_\mu^{ll,\mathrm{W}}(\mathrm{conn.})$. The top-right panel illustrates the continuum extrapolations on two data sets, the first corrected with NNLO χPT and the second corrected with the CM, in both cases computed from the full W window interval. The dashed lines indicate the continuum extrapolations for each data set. In total, we include over two thousand separate fit results in the model average. The resulting distribution is shown in the top left panel of Fig. 10, where it is overlaid on the BMA result (red line and error band) obtained using Eqs. (3.14) and (3.15). The middle panel shows the results from the 24 best individual fits for each correction choice, ordered by the BAIC, in comparison to the BMA result, while the bottom panel gives the associated Q values [91] computed from χ^2_{aug}. We find that our best fits, as determined by the Q value, tend to have the smallest BAIC and hence largest model probability. We also note that the continuum results for data sets corrected for taste breaking using NNLO χPT tend to be smaller than those from the other variations and also return some of the largest model probabilities (points in middle panel with lower half purple).

[10] Note that for BMA, the single data set D is held fixed even in variations where ensembles are dropped since this is treated as a model change (see the discussion of data subset selection in [46, 47]). Also note that throughout this work, we also use the more colloquial definition of "data set" outside the context of BMA to refer to any set of a_μ data points before continuum extrapolation.

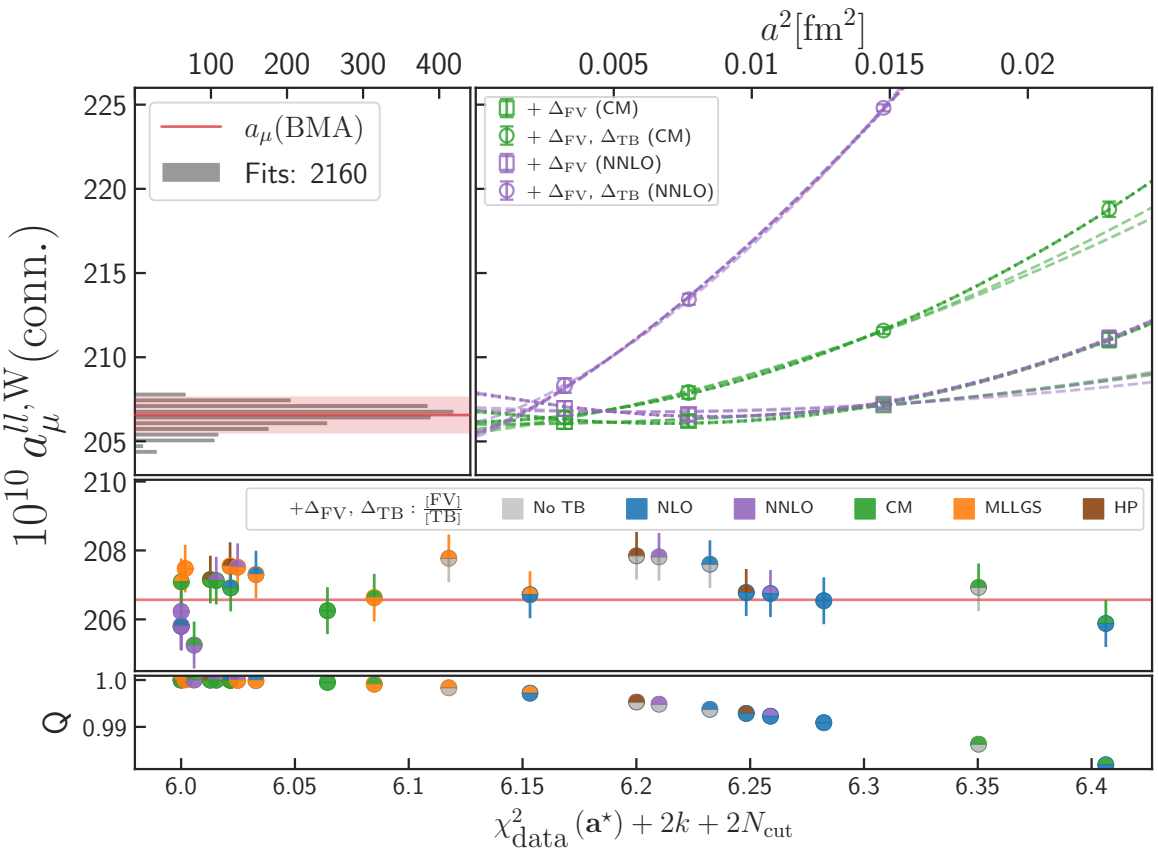

FIG. 10. Results of the Bayesian model averaging (BMA) procedure applied to $a_\mu^{ll,\mathrm{W}}$(conn.). Top left: Histogram of all continuum extrapolations used in the BMA, the light-red band is the BMA result. Top right: The subset of data sets and extrapolations corresponding to correcting the data with the CM and NNLO χPT. Data (without) with taste-breaking corrections are shown as (squares) circles. Different extrapolations correspond to variations of the fit function and ensembles included. Lower panels: The best fits according to model probability, Eq. (3.13). The middle panel shows the fit results, while the bottom one shows the corresponding Q-values [91]. In both panels, the correction schemes employed for Δ_{FV} and Δ_{TB} are indicated by the symbols' top and bottom colors, respectively.

Figure 11 shows the analogous BMA result for $a_\mu^{ll,\mathrm{W2}}$(conn.). Here we include 384 fit results, which is fewer than for $a_\mu^{ll,\mathrm{W}}$(conn.). This stems from the absence of NLO χPT corrections and from employing only a single correction region. The general features of this figure are the same as for Fig. 10. In the top left panel, we note that the BMA uncertainty for $a_\mu^{ll,\mathrm{W2}}$(conn.) is larger than the spread of the histogram. This is because the bulk of the uncertainty in this case comes from the first term in Eq. (3.15) with relatively large statistical and scale-setting uncertainty contributions.

In order to better understand and test the model-averaging results, we also perform Bayesian model averages on specific subsets of the variations. That is, we fix one of the analysis choices but vary the rest as usual. The results of these subset averages for $a_\mu^{ll,\mathrm{W}}$(conn.) are shown in Fig. 12 (left). The top data point is our BMA result from Fig. 10. The two data points below it, show the BMA results for the two observable extraction choices described in Sec. III B. They are in excellent agreement with each other and with the full BMA result,

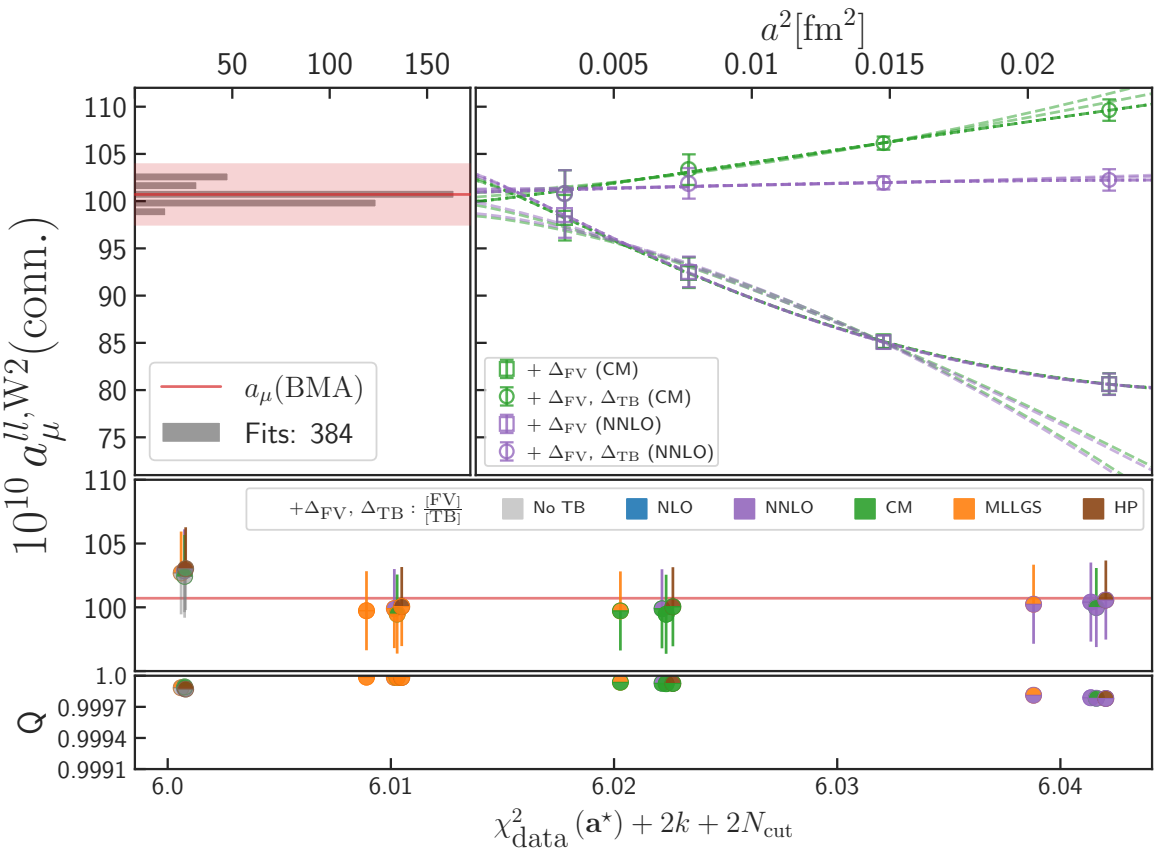

FIG. 11. Results of the Bayesian model averaging (BMA) procedure applied to $a_\mu^{ll,\mathrm{W2}}$(conn.). Figure is described in the caption of Fig. 10.

signifying, as expected, that residual effects of oscillating contributions and of $O(a^2)$ errors of the trapezoidal rule are negligibly small. The next five data points are the BMA results obtained from subsets with specific taste-breaking correction schemes. While these results are statistically consistent with the overall average, the differences in the central values contribute significantly to the systematic uncertainty through the latter two terms of Eq. (3.15) (outer uncertainty of the BMA result). In particular, as shown in Fig. 10, the fit results obtained from NNLO χPT corrected data tend to lie below the average. The following three data points are BMA results obtained from subsets of specific continuum-extrapolation fit functions, which agree well with each other and with the full BMA result. The last block of data points (below the dashed line) are BMA results from subsets that use the same schemes for finite-volume and taste-breaking corrections, where the top data point (BMA w/o mix) averages all four schemes (NLO χPT, NNLO χPT, CM, MLLGS), followed by results from the subsets corresponding to each single scheme, all of which are consistent with the full BMA result with small variations in central values.

The probability weights defined in Eq. (3.13) can be used to assess the relative weight of specific analysis choices in the BMA. Comparison of these weights can identify if one particular choice of observable extraction method, correction scheme or fit-function variation is preferred by the averaging procedure. More specifically, letting S denote a subset of the full space of models $\{M\}$. We can define the "subset probability" of S by the relative

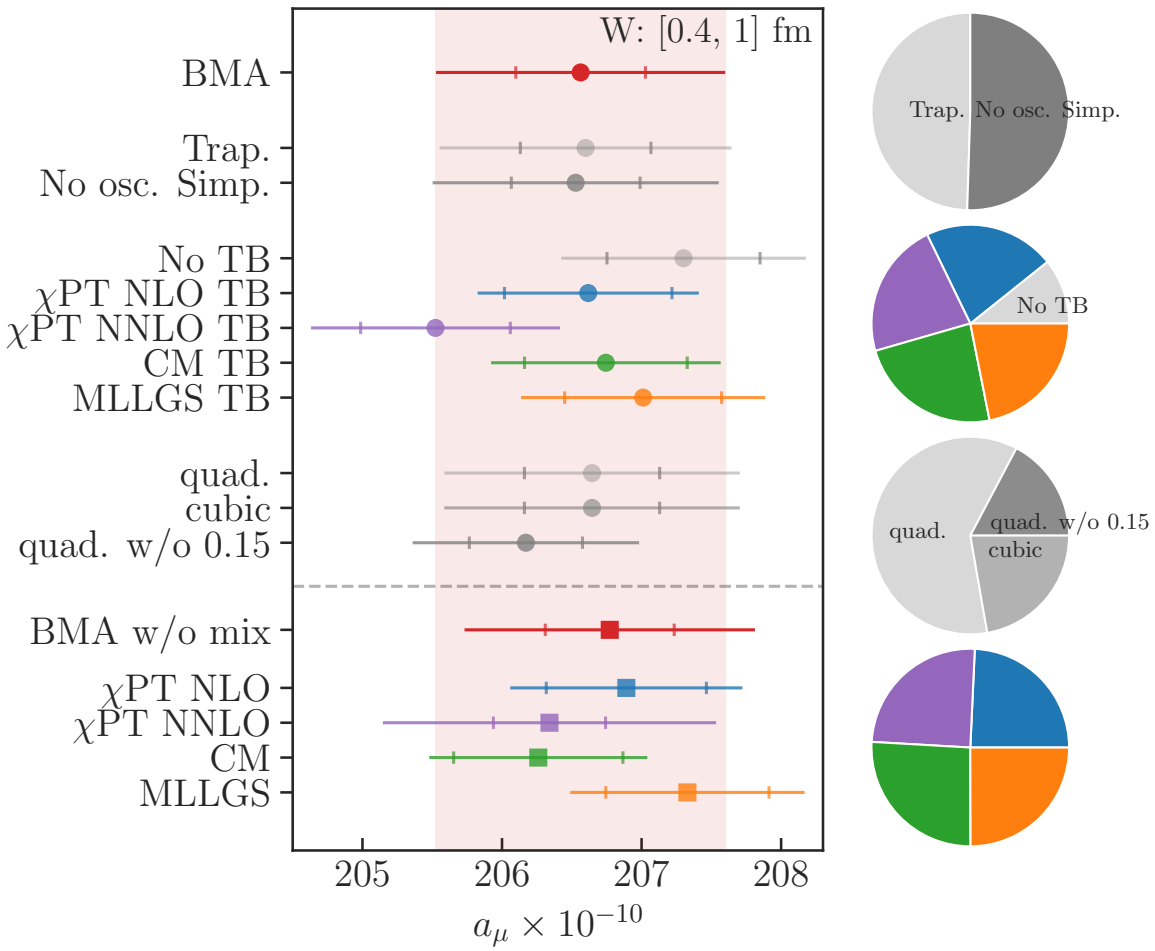

FIG. 12. Breakdown of the results from the Bayesian model averaging applied to $a_\mu^{ll,\mathrm{W}}$(conn.). Left: From top to bottom, the first, main result (BMA) includes all data sets, schemes, and other variations. The next two use data integrated with either the trapezoidal rule (Trap.) or Simpson's rule (No osc. Simp.). The following five results are obtained from subsets with specific taste-breaking corrections. The next three are subsets with specific continuum fit functions: quadratic, cubic, or quadratic without the 0.15 fm ensemble. The last block of results (below the dashed line) uses the same scheme for finite-volume and taste-breaking corrections. The top ("BMA w/o mix") includes all four schemes; the final four are breakdowns using only a single correction scheme in the BMA. The inner error bar on the data points corresponds to the first term in Eq. (3.15), while the outer is the total error. Right: Pie-charts showing the contributions to the BMA corresponding to the breakdowns in the left panel. The percentages are computed by summing over Eq. (3.13) for the particular subsets.

posterior probability of the variations contained in S:

$$\mathrm{pr}(S|D) = \sum_{M_i \in S} \mathrm{pr}(M_i|D). \tag{3.16}$$

The subset probability encapsulates the relative weight of the models in a given subset compared to the whole model space, informed by the data. For example, we can estimate the subset probability of using NNLO χPT for taste-breaking and finite-volume corrections

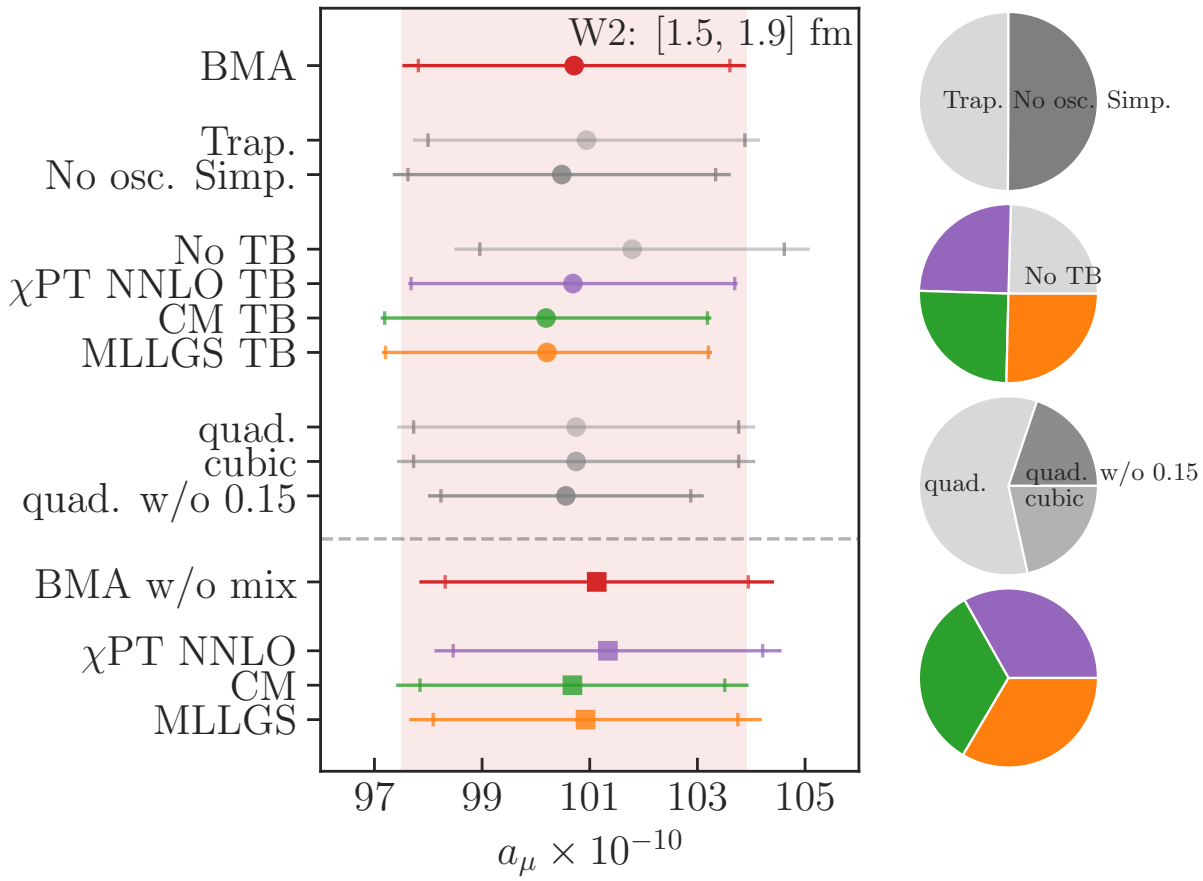

FIG. 13. Breakdown of the results from the Bayesian model averaging applied to $a_\mu^{ll,\mathrm{W2}}$(conn.). Figure is described in the caption of Fig. 12 (with "four results" replacing "five results" because NLO χPT is not employed here).

as

$$\mathrm{pr}(\mathrm{NNLO}|D) = \sum_{M_i \in \mathrm{NNLO}} \mathrm{pr}(M_i|D). \tag{3.17}$$

where "$M_i \in$ NNLO" denotes the subset of models (*i.e.*, analysis choices) in which NNLO χPT is used for both corrections. Using this definition, we show the relative probabilities of the subsets considered above as pie charts in Fig. 12 (right). From the top pie chart, for the two methods of observable extraction, we find roughly equal contributions to the overall BMA result, indicating no preference by the BMA procedure. The second pie-chart from the top shows the subset probabilities for specific taste-breaking corrections. The probability of the subset in which the data are not corrected prior to continuum extrapolation is smaller by slightly more than a factor of two compared with the other subsets. This is because the taste-breaking corrections are computed in two window regions, [0.4, 1] and [0.7, 1] fm in addition to the continuum fits to data without taste-breaking corrections having larger χ^2 values, indicating a preference for data corrected for taste-breaking. The third pie chart shows that quadratic continuum fits to the full set of four ensembles are preferred over cubic fits or fits to just three ensembles. In the case of the fits to three ensembles the smaller subset probability can be traced back to the penalty incurred, N_{cut}, in Eq. (3.13) due to dropping a data point. For subsets in which the same correction scheme is used for finite

TABLE III. Comparison of results for $a_\mu^{ll,\mathrm{W}}$(conn.) and $a_\mu^{ll,\mathrm{W2}}$(conn.) obtained from the BMA analysis with an empirical Bayes approach. Both analyses use data sets corrected in the CM scheme. The third and fifth columns show a_μ results obtained in the empirical Bayes approach from fits to data sets without and with first correcting for taste splittings, respectively. The fourth and sixth columns list the effective scales Q_{eff} obtained by maximizing the Bayes Factor log(GBF).

	BMA: CM	Empirical Bayes			
		$\Delta_{\mathrm{TB}} = 0$	Q_{eff}/GeV	$\Delta_{\mathrm{TB}} \neq 0$	Q_{eff}/GeV
$a_\mu^{ll,\mathrm{W}}$(conn.)	206.28(81)	206.52(69)	1.8	205.83(67)	1.7
$a_\mu^{ll,\mathrm{W2}}$(conn.)	100.9(3.3)	98.7(2.8)	2.9	102.0(2.0)	1.9

volume and taste splittings (bottom pie chart) we find a slight preference for NNLO χPT and slight disinclination for NLO χPT.

Figure 13 shows the BMA subsets for $a_\mu^{ll,\mathrm{W2}}$(conn.), with results similar to those for $a_\mu^{ll,\mathrm{W}}$(conn.). As expected, there is greater consistency among the subset averages from specific correction schemes compared to the $a_\mu^{ll,\mathrm{W}}$(conn.) case, with the largest variation in central value coming from continuum extrapolations to data not corrected for taste-breaking effects. The pie charts in the right panel of Fig. 13 reveal roughly equal subset probabilities in each case, except for the third (from the top) pie chart, which illustrates that here too quadratic continuum fits to all four ensembles are preferred for the same reasons as above.

The AIC criterion used in Ref. [29] differs from Eq. (3.13) in that the weight assigned to cutting data points is given as N_{cut} instead of $2N_{\mathrm{cut}}$. In order to test the robustness of the model-averaging procedure, we repeat the analysis by replacing $2N_{\mathrm{cut}}$ with N_{cut} in Eq. (3.13). We find that this yields central values and uncertainties on the final results are essentially the same as before, with at most minor changes to the weights in the third pie-chart from the top in Figs. 12 and 13. This result is not unexpected, because in our case, $N_{\mathrm{cut}} \leq 1$, and only a small fraction of the total variations in our averages have $N_{\mathrm{cut}} \neq 0$.

In order to cross check our main continuum-limit extrapolations and the subsequent BMA analysis, we use an empirical Bayes approach to perform independent continuum-limit extrapolations (see Eq. (3.12)). In the comparison of the two approaches, we use the a_μ data sets obtained from integrating the correlation functions with Trapezoidal rule and corrected using the CM scheme, with and without first correcting for taste splittings. When performing continuum extrapolations using all the terms in Eq. (3.12), we observe that most of the posterior coefficients are small and consistent with zero — only the linear $\sim c_{11}a^2\alpha_s$ and quadratic $\sim c_{20}a^4$ terms in Eq. (3.12) are needed to describe the data. This observation is consistent with our main continuum-extrapolation analysis, described by Eqs. (3.8) and (3.9). Table III shows the comparison of the BMA analysis (restricted to the same CM-corrected data sets) with the empirical Bayes fits for both $a_\mu^{ll,\mathrm{W}}$(conn.) and $a_\mu^{ll,\mathrm{W2}}$(conn.). We find good agreement in central values and error bars, after considering the spread between the empirical Bayes results from data sets with and without first correcting for taste splittings.

TABLE IV. Approximate error budgets for $a_\mu^{ll,\mathrm{W}}$(conn.) and $a_\mu^{ll,\mathrm{W2}}$(conn.).

Source	$\delta a_\mu^{ll,\mathrm{W}}$(conn.) (%)	$\delta a_\mu^{ll,\mathrm{W2}}$(conn.) (%)
Monte Carlo statistics	0.19	2.44
Continuum extrapolation ($a \to 0$, Δ_{TB})	0.34	1.05
Finite-volume correction (Δ_{FV})	0.16	0.23
Pion-mass adjustment (Δ_{M_π})	0.06	0.96
Scale setting (w_0 (fm), w_0/a)	0.21	1.28
Current renormalization (Z_V)	0.17	0.16
Total	0.50%	3.18%

F. Results and error budgets

Our results for the light-quark-connected contributions to a_μ^{W} and a_μ^{W2} are

$$a_\mu^{ll,\mathrm{W}}(\mathrm{conn.}) = 206.6(1.0) \times 10^{-10} \tag{3.18}$$

and

$$a_\mu^{ll,\mathrm{W2}}(\mathrm{conn.}) = 100.7(3.2) \times 10^{-10}\ , \tag{3.19}$$

where the errors are those obtained from the BMA procedure described in the previous section, and include both statistical and systematic uncertainties.

Although Bayesian model averaging provides a robust estimate of the total uncertainties in our results, the construction of detailed error budgets from the BMA is not straightforward. We start from the expression for the BMA variance in Eq. (3.15). The first term on the right-hand side is linear in the variances, and hence can be trivially separated into individual contributions from Monte Carlo statistics and each of the parametric inputs w_0, Δ_{M_π}, and Z_V. For example, the statistical uncertainty is given by

$$\sigma^2_{a_\mu}(\mathrm{stat.}) = \sum_{i=1}^{N_M} \sigma^2_{a_\mu,i}(\mathrm{stat.})\mathrm{pr}\left(M_i \mid D\right) \tag{3.20}$$

where we average over all analysis variations using the probability weights of Eq. (3.13). Repeating this procedure for all the above-mentioned contributions yields the error estimates in Table IV in the rows marked "Monte Carlo statistics", "Scale setting", "Pion-mass adjustment" and "Current renormalization". The second and third terms in Eq. (3.15) depend solely and non-linearly on the central value of each variation, with the latter term including pairwise differences between all possible model pairs in the full BMA result. This makes it impossible to strictly disentangle the contribution from only a subset of model variations, (*e.g.*, finite-volume corrections or treatment of discretization effects). We can obtain an approximate error budget, however, as follows.

First, to estimate the systematic uncertainty associated with the finite-volume correction, we perform subset model averages separately for each finite-volume correction scheme. These results are shown in Fig. 14. Taking the variance in central values of these results yields the "finite-volume correction" error in Table IV. Next, we subtract (in quadrature) the so-estimated finite-volume error from the total model variance. The remaining uncertainty is associated with variations in the treatment of oscillating states in $C(t)$, the taste-breaking

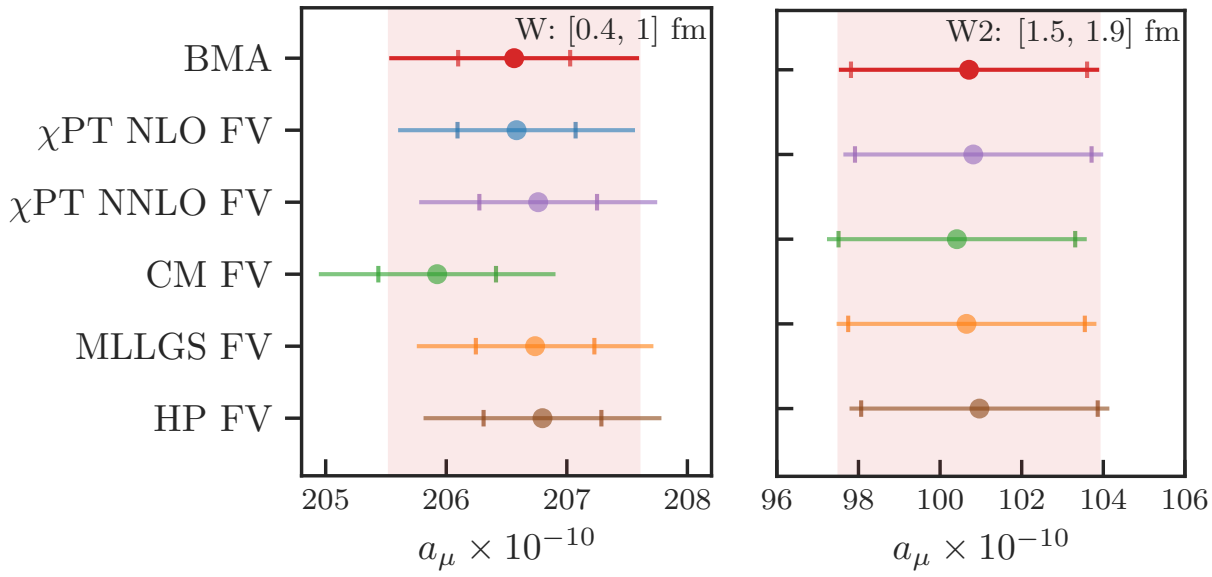

FIG. 14. Breakdown of the BMA result into subsets that contain only one choice of finite-volume correction for $a_\mu^{ll,\mathrm{W}}$(conn.) (left) and $a_\mu^{ll,\mathrm{W2}}$(conn.) (right).

corrections, and the continuum-extrapolation fit function. Combining this uncertainty (in quadrature) with those on the fit-function coefficient posteriors yields the "continuum extrapolation" error in Table IV.

Table IV presents the approximate error budgets for $a_\mu^{ll,\mathrm{W}}$(conn.) and $a_\mu^{ll,\mathrm{W2}}$(conn.) obtained from the above approach. For $a_\mu^{ll,\mathrm{W}}$(conn.), the largest error is from the continuum extrapolation, and is driven by the spread in results using different taste-breaking correction schemes. Here we note that the consistency between quadratic and cubic continuum extrapolations (as illustrated in Figs. 12 and 13) as well as between our main results and those from the empirical Bayes approach (see Table III) indicate that systematic errors due residual higher-order discretization effects are well encompassed by our uncertainties. Next is the parametric uncertainty from the gradient-flow scale, which is about 30% smaller. Errors from Monte-Carlo statistics, finite-volume corrections, and current renormalization are also non-neglible, and are roughly commensurate. For $a_\mu^{ll,\mathrm{W2}}$(conn.), Monte-Carlo statistics are by far the largest source of uncertainty. Following that, the contributions from scale setting, the continuum extrapolation, and the pion-mass adjustment, which are $\sim$ 50–60% smaller. Although finite-volume and current-renormalization errors are negligible compared with these other uncertainties, they will be important for calculations of $a_\mu^{\mathrm{HVP,LO}}$ aiming for $\lesssim 0.5\%$ precision.

IV. SUMMARY AND OUTLOOK

In Fig. 15, we compare our intermediate-window result, Eq. (3.18), with other lattice-QCD calculations of this quantity [29, 32, 33, 36–38, 71, 92, 93], which were obtained using different lattice actions and analysis methods. Of the results to date, ours has the smallest statistical uncertainty, 0.19%. Ours is also the first result for $a_\mu^{ll,\mathrm{W}}$(conn.) obtained from

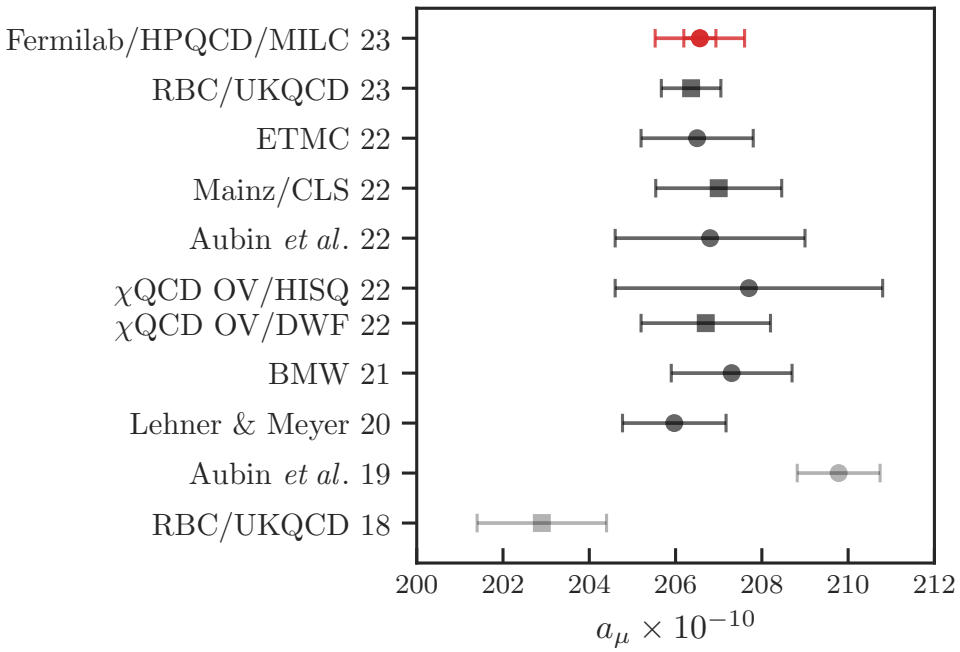

FIG. 15. Comparison of our lattice determination of $a_\mu^{ll,\mathrm{W}}$(conn.) (red circle) labeled "Fermilab/HPQCD/MILC 23" to $n_f = 2{+}1{+}1$ (black circles) and $n_f = 2{+}1$ (black squares) lattice-QCD calculations by RBC/UKQCD 23 [38], ETMC 22 [37], Mainz/CLS 22 [36], Aubin *et al.* 22 [33], χQCD 22 [92], BMW 21 [29] and Lehner & Meyer 20 [93]. Results by Aubin *et al.* 19 [71] and RBC/UKQCD 18 [32], shown in grey, are superseded by Aubin *et al.* 22 and RBC/UKQCD 23, respectively. The inner error bar shown for our result is from Monte Carlo statistics.

a blind analysis. While some form of EFT-inspired correction schemes were employed in every calculation, our analysis is the first to include all of them. Because we incorporate uncertainties due to analysis choices via Bayesian model averaging [46, 47], our systematic error estimate is robust without being overly conservative.

In Fig. 16, we compare our result for the "W2" window observable, Eq. (3.19), with the only other available lattice-QCD result for this quantity [33]. Although the results appear consistent, they are not wholly independent because the analysis in Ref. [33] is based on some of the same ensembles as employed in this work. Statistical and systematic correlations due to the shared configurations must be taken into account to make a quantitative comparison. Other independent lattice-QCD calculations of $a_\mu^{ll,\mathrm{W2}}$(conn.) would provide welcome consistency checks.

Before our results for $a_\mu^{ll,\mathrm{W}}$(conn.) and $a_\mu^{ll,\mathrm{W2}}$(conn.) can be directly compared with data-driven determinations, the contributions from heavier flavors must be added as well as those from quark-line disconnected contractions and isospin-breaking corrections (QED and $m_u \neq m_d$). The s-, c-, and b-quark-connected contributions to $a_\mu^{\mathrm{HVP,LO}}$ have already been computed on the HISQ ensembles with high precision [94–96]; windowing these results will be straightforward. The remaining contributions are being computed in ongoing projects; see Refs. [40–42, 50].

Looking at the big picture, the observed consistency between so many different, largely independent, results for the light-quark connected contribution to the intermediate-window

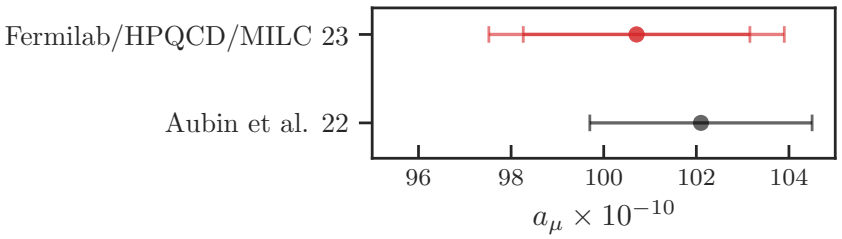

FIG. 16. Comparison of our lattice determination of $a_\mu^{ll,\mathrm{W2}}$(conn.) (red circle) labeled "Fermilab/HPQCD/MILC 23" with the result of Ref. [33] (black circle) labeled Aubin *et al.* 22. The inner error bar shown for our result is from Monte Carlo statistics.

observable (see Fig. 15) indicates that the systematic errors in lattice-QCD calculations of this quantity are under reasonable control. It is therefore unlikely that the differences between the lattice-QCD calculations reported in Refs. [29, 36–38] and the data-driven result of Ref. [35] will be resolved by further improvements in lattice-QCD calculations of $a_\mu^{ll,\mathrm{W}}$(conn.). Lattice-QCD calculations of the quark-connected contributions from heavier flavors are also unlikely causes of the difference, since their uncertainties are smaller by an order of magnitude [29, 95, 96]. The quark-disconnected and isospin-breaking contributions to $a_\mu^{\mathrm{HVP,LO}}$, however, have been computed by only a few collaborations [29, 32, 38, 39, 87, 97].[11] Although these contributions are too small to change a_μ^{W} substantially, additional independent lattice-QCD calculations are needed to solidify the central value and uncertainty in order to better quantify the significance of the difference.

In Ref. [34], we pointed out that other windowed observables can provide more stringent comparisons between lattice-QCD and data-driven results right now. Because intermediate-window observables cut out low-t contributions to $a_\mu^{\mathrm{HVP,LO}}$ where lattice-QCD statistical errors are smallest, "one-sided windows" without a lower bound on the Euclidean time can capture a larger fraction of the total $a_\mu^{\mathrm{HVP,LO}}$ while retaining controlled uncertainties. We are currently repeating the analysis of Ref. [34] using the larger data set employed in this work.

The light-quark connected contribution to the intermediate-window observable represents only around a third of the total leading-order HVP contribution to the muon's anomalous magnetic moment. Thus, the work presented in this paper is only a part of a multi-year project to compute $a_\mu^{\mathrm{HVP,LO}}$ with $\lesssim 0.5\%$ precision. Several of our ongoing efforts aim to reduce the dominant sources of uncertainty in our published result for the total light-quark connected contribution to $a_\mu^{\mathrm{HVP,LO}}$ [43]; these will also improve our determinations of the intermediate-window observables in this work. For example, recently we introduced a "low-mode-improved" method into our analysis that substantially reduces statistical errors at large Euclidean times [98]. The uncertainty on the scale-setting quantity w_0 is an important source of uncertainty not only for all $a_\mu^{\mathrm{HVP,LO}}$ observables, but also for many other analyses based on the MILC HISQ ensembles. We are therefore working to compute precisely the Ω-baryon mass on these ensembles [99], as well as the relative scale w_0/a, and plan to use the results to determine the scale in physical units with reduced uncertainty.

With these ongoing efforts, we expect to obtain $a_\mu^{\mathrm{HVP,LO}}$ with sub-percent-level precision in the near future. In order to further reduce the precision to match that of the Fermi-

[11] Indeed, only the BMW collaboration [29] has presented a complete calculation of all contributions to $a_\mu^{\mathrm{HVP,LO}}$ including the disconnected QED and disconnected strong-isospin-breaking corrections.

lab [1] and JPARC [24, 25] experiments, however, it seems likely that considerable exascale computing resources will be needed. In particular, the inclusion of MILC's physical-mass HISQ ensemble with $a \approx 0.042$ fm would enable more robust continuum extrapolations of all $a_\mu^{\mathrm{HVP,LO}}$ observables and provide better control over this important source of systematic error. A direct finite-volume study is needed to better quantify the finite-volume corrections and reduce the corresponding uncertainty. This would require the generation and analysis of new ensembles with different spatial volumes and all other parameters held fixed. Finally, further control over long-distance effects and statistical noise could be achieved by computing directly the two-pion contributions to the vector current correlation functions [49, 100, 101].

ACKNOWLEDGMENTS

We thank Claude Bernard, Urs Heller, Jack Laiho, Bob Sugar, and Doug Toussaint for their scientific leadership and collaboration. In particular, we are grateful to Bob for his tireless efforts to obtain computational resources, to Claude for guidance on chiral perturbation theory, to Doug for his invaluable expertise in creating so many of our gauge-field ensembles, and to Jack and Urs for essential contributions to previous projects that formed the basis for this work. We thank Anthony Grebe for his contributions to the chiral perturbation theory codes used in this work. We thank Maarten Golterman for useful comments and suggestions. Computations for this work were carried out in part with resources provided by the USQCD Collaboration, the National Energy Research Scientific Computing Center (Cori), the Argonne Leadership Computing Facility (Mira) under the INCITE program, and the Oak Ridge Leadership Computing Facility (Summit) under the Innovative and Novel Computational Impact on Theory and Experiment (INCITE) and the ASCR Leadership Computing Challenge (ALCC) programs, which are funded by the Office of Science of the U.S. Department of Energy. This work used the Extreme Science and Engineering Discovery Environment (XSEDE) supercomputer Stampede 2 at the Texas Advanced Computing Center (TACC) through allocation TG-MCA93S002. The XSEDE program is supported by the National Science Foundation under grant number ACI-1548562. Computations on the Big Red II+ and Big Red 3 supercomputers were supported in part by Lilly Endowment, Inc., through its support for the Indiana University Pervasive Technology Institute. The parallel file system employed by Big Red II+ is supported by the National Science Foundation under Grant No. CNS-0521433. This work utilized the RMACC Summit supercomputer, which is supported by the National Science Foundation (awards ACI-1532235 and ACI-1532236), the University of Colorado Boulder, and Colorado State University. The Summit supercomputer is a joint effort of the University of Colorado Boulder and Colorado State University. Some of the computations were done using the Blue Waters sustained-petascale computer, which was supported by the National Science Foundation (awards OCI-0725070 and ACI-1238993) and the state of Illinois. Blue Waters was a joint effort of the University of Illinois at Urbana-Champaign and its National Center for Supercomputing Applications. Some computations also used the Cambridge Service for Data Driven Discovery (CSD3) operated by University of Cambridge Research Computing on behalf of the STFC DiRAC HPC Facility. The DiRAC component of CSD3 was funded by BEIS and STFC under grants ST/P002307/1, ST/R002452/1 and ST/R00689X/1.

This work was supported in part by the U.S. Department of Energy, Office of Science, under Awards No. DE-SC0010005 (E.T.N.), No. DE-SC0010120 (S.G.), Nos. DE-SC0011090 and DE-SC0021006 (W.J.), No. DE-SC0015655 (A.X.K., S.L., M.L., A.T.L.), and the

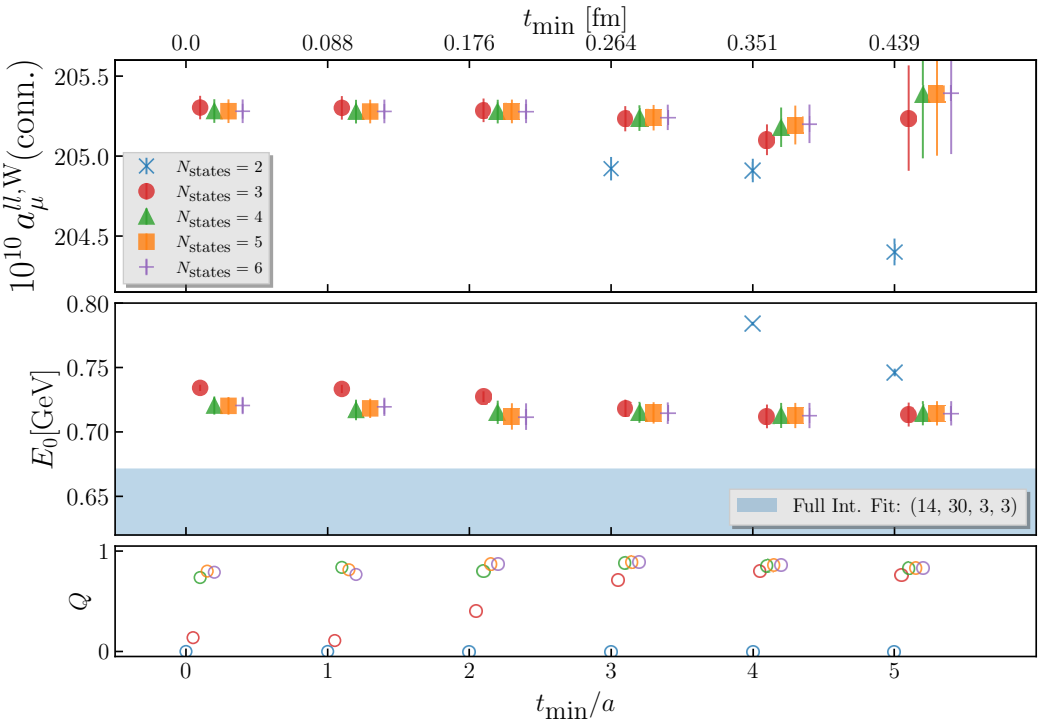

FIG. 17. Fit and reconstruction results on the 0.09 fm ensemble. (Top) $a_\mu^{ll,{\rm W}}$(conn.) from the reconstruction for a range of $t_{\rm min}$ and number of exponentials in the fit function values with $t_{\rm max} = 1.3$ fm fixed. (Middle) Ground state energies determined by the fit and fit result for the late-time correlation function obtained from a fit with three states, $t_{\rm min}/a = 14$ and $t_{\rm max}/a = 30$ (blue band). (Bottom) Fit quality, Q, coming from the augmented chi squared fit.

Funding Opportunity Announcement Scientific Discovery through Advanced Computing: High Energy Physics, LAB 22-2580; by the National Science Foundation under Grants Nos. PHY17-19626 and PHY20-13064 (C.E.D., A.V.) and from their Graduate Research Fellowship under Grant DGE 2040434 (C.T.P); by the Simons Foundation under their Simons Fellows in Theoretical Physics program (A.X.K.); by the Universities Research Association Visiting Scholarship awards 20-S-12 and 21-S-05 (S.L.); by SRA (Spain) under Grant No. PID2019-106087GB-C21 / 10.13039/501100011033 (E.G.); by the Junta de Andalucía (Spain) under Grants No. FQM-101, A-FQM-467-UGR18 (FEDER), and P18-FR-4314 (E.G.); by AEI (Spain) under Grant No. RYC2020-030244-I / AEI / 10.13039/501100011033 (A.V.); and by UK Science and Technology Facilities Council under Grant ST/T000945/1 (C.T.H.D). This document was prepared by the Fermilab Lattice, HPQCD, and MILC Collaborations using the resources of the Fermi National Accelerator Laboratory (Fermilab), 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.

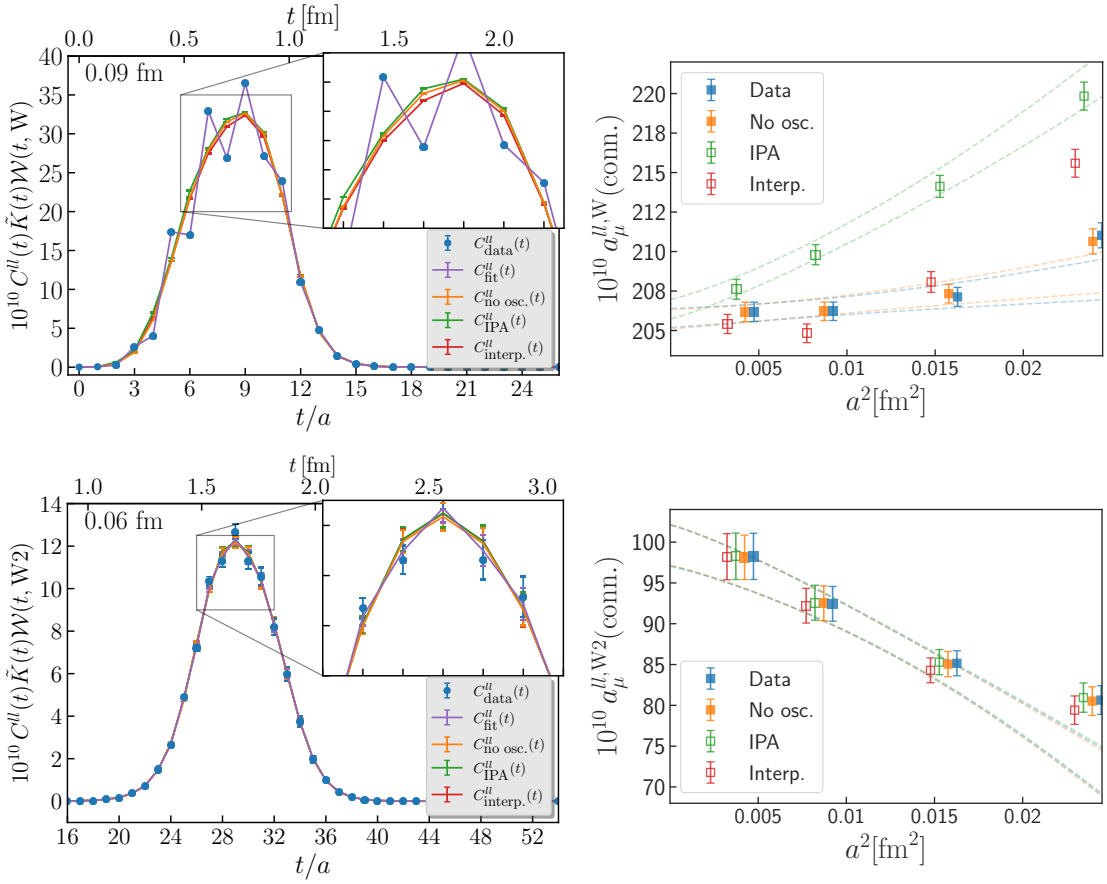

FIG. 18. (Left) Comparison of the methods used to remove the oscillating contribution to the integrand for $a_\mu^{ll,\mathrm{W}}(\mathrm{conn.})$ at 0.09 fm (top left) and $a_\mu^{ll,\mathrm{W2}}(\mathrm{conn.})$ at 0.06 fm (bottom left). Shown are the integrands obtained with raw correlation-function data $C(t)$ (blue circles), the reconstruction from the fit including oscillating states $C_{\mathrm{fit}}(t)$ (purple), without oscillating states $C_{\mathrm{no\ osc.}}(t)$ (orange), improved parity averaged correlator $C_{\mathrm{IPA}}(t)$, (Eq. (A4)) (green), and interpolated correlator $C_{\mathrm{interp}}(t)$ (Eq. (A5)) (red). (Right) Lattice-spacing dependence of $a_\mu^{ll,\mathrm{W}}(\mathrm{conn.})$ (top right) and $a_\mu^{ll,\mathrm{W2}}(\mathrm{conn.})$ (bottom right) data obtained from the correlation functions modified with the oscillation removal techniques discussed. All data sets are corrected for finite-volume effects using the Chiral Model and pion-mass mistuning effects using the data-driven approach, described in Secs. III C 1 and III C 2. The data points are slightly displaced horizontally for clarity. A linear fit function (see Sec. III D) is used to fit the $a_\mu^{ll,\mathrm{W}}(\mathrm{conn.})$ and $a_\mu^{ll,\mathrm{W2}}(\mathrm{conn.})$ data at the three finest lattice spacings.

Appendix A: Cross-checks of window determinations from staggered correlation functions

In this section, we detail the methods used to obtain the windowed $a_\mu^{\mathrm{HVP,LO}}$ from the staggered correlation function, $C(t)$. First, we compare three different approaches for treating the oscillating contribution to the correlator. The first method, which we use in our main analysis, is to fit $C(t)$ over the region of interest and reconstruct it from the fit posteriors excluding the oscillatory contribution. The correlation function has the spectral

representation,

$$C(t) = \sum_n \left[Z_n^2 e^{-E_n t} + (-1)^t Z_{n,\mathrm{osc}}^2 e^{-E_{n,\mathrm{osc}} t} \right], \tag{A1}$$

where the sum is over all possible contributing states. We use this expression to craft a model fit function that separates the oscillating and nonoscillating contributions. For this purpose, we truncate the sum:

$$C_{\mathrm{fit}}(t) = \sum_n^{N_{\mathrm{states}}} \left[Z_n^2 e^{-E_n t} + (-1)^t Z_{n,\mathrm{osc}}^2 e^{-E_{n,\mathrm{osc}} t} \right]. \tag{A2}$$

For simplicity, we keep the same number N_{states} of regular and oscillating states. We restrict the fit range $[t_{\mathrm{min}}, t_{\mathrm{max}}]$ to cover the window region of interest. The use of t_{min} justifies the truncation to a finite N_{states} by suppressing contributions from states with large energy. On the other hand, the use of a fixed t_{max} carries a risk that the lowest-lying energies and amplitudes may not be accurately resolved with finite statistical precision. However, since we are simply using the expression as a useful model for removing the unwanted oscillations, it is not critical that our estimates of all energy levels are asymptotically correct for $t_{\mathrm{max}} \to \infty$. For the energies and amplitudes of the light-quark-connected correlator, we take the Gaussian priors associated with the local (unsmeared) data in Eqs. (A3) and (A4) of Ref. [48]. We then have the corresponding fit reconstruction of the correlation function without the oscillating contribution

$$C_{\text{no osc.}}(t) = \sum_n^{N_{\mathrm{states}}} Z_n^2 e^{-E_n t}. \tag{A3}$$

Results for $a_\mu^{ll,\mathrm{W}}$(conn.) computed from the fit reconstruction on the 0.09 fm ensemble are shown in Fig. 17 (top). Here, for the fit range, we fix $t_{\mathrm{max}} = 1.3$ fm $(t_1 + 2\Delta)$ and vary t_{min}. We also fit up to six states with good stability obtained at four, which we take to be our value for N_{states} on all ensembles. The second panel of Fig. 17 shows the ground state energies obtained from these fits; shown also is the ground state energy obtained from a fit to the full correlation function (blue band). We see a significant difference in these energies, perhaps because the full fit picks up some mixture of the hard-to-determine two-pion states in the large-time region. Nonetheless, we observe in Fig. 18 (as described below) that C_{fit} accurately reconstructs the correlation function data in the window region of interest. For $a_\mu^{ll,\mathrm{W}}$(conn.) and $a_\mu^{ll,\mathrm{W2}}$(conn.) we take t_{min} and t_{max} to be 2Δ beyond the t_0 and t_1 boundaries in the corresponding window definition. For the coarsest two ensembles, this would correspond to a $t_{\mathrm{min}}/a = 0, 1$. To avoid possible staggered-operator complications at small t/a, we take $t_{\mathrm{min}}/a = 2$ for those two ensembles. As a test of the fidelity of this method, we show results for the correlated differences of $a_\mu^{ll,\mathrm{W}}$(conn.) and $a_\mu^{ll,\mathrm{W2}}$(conn.) computed from the fit reconstruction with the oscillating states, $C_{\mathrm{fit}}(t)$, and the original correlation function in Table V. One can see tiny differences on the coarsest ensembles for $a_\mu^{ll,\mathrm{W}}$(conn.), likely due to the restriction of not using the first two time-slices; however, these differences are well within the uncertainties of the results for $a_\mu^{ll,\mathrm{W}}$(conn.).

The second method we examined is improved parity averaging (IPA) as employed in Ref. [93] for computing $a_\mu^{\mathrm{HVP,LO}}$, a modification of the method developed in Ref. [102]. Here,

TABLE V. $a_\mu^{ll,\mathrm{W}}$(conn.) and $a_\mu^{ll,\mathrm{W2}}$(conn.) computed from the raw data (columns two and five), the fit reconstruction with oscillating states (columns three and six) and the correlated difference between them (columns four and seven).

$\approx a$	$a_\mu^{ll,\mathrm{W}}$(conn.)	$a_{\mu,\,\mathrm{fit}}^{ll,\mathrm{W}}$(conn.)	$\Delta a_\mu^{ll,\mathrm{W}}$(conn.)	$a_\mu^{ll,\mathrm{W2}}$(conn.)	$a_{\mu,\,\mathrm{fit}}^{ll,\mathrm{W2}}$(conn.)	$\Delta a_\mu^{ll,\mathrm{W2}}$(conn.)
0.15	211.01(79)	211.15(80)	−0.14(11)	80.3(1.7)	80.1(1.7)	0.20(19)
0.12	207.13(60)	207.16(60)	−0.025(29)	84.7(1.5)	84.6(1.5)	0.09(10)
0.09	206.56(55)	206.58(55)	−0.016(10)	92.7(1.8)	92.7(1.8)	−0.07(22)
0.06	206.22(61)	206.22(61)	0.003(61)	95.6(2.8)	95.5(2.7)	0.12(73)

the correlation function is replaced by the following equation:

$$C_{\mathrm{IPA}}(t) = \frac{e^{-m_\rho t}}{4}\left[\frac{C(t-1)}{e^{-m_\rho(t-1)}} + 2\frac{C(t)}{e^{-m_\rho(t)}} + \frac{C(t+1)}{e^{-m_\rho(t+1)}}\right] \tag{A4}$$

The exponent used is the PDG value of the ρ meson mass [103], to give the best cancellation in the ρ resonance peak which dominates in the regions of W and W2. This approach introduces additional discretization effects; however, one expects a consistent continuum limit as the oscillations become small at finer lattice spacing.

The final approach, originally used in Ref. [104], is performed by interpolating the even- and odd-site correlation functions separately, then averaging the two interpolations to obtain a new correlation function where the oscillating contribution has been removed.

$$C_{\mathrm{interp}}(t) = \frac{1}{2}\left(C^{\text{even. interp}}(t) + C^{\text{odd. interp}}(t)\right) \tag{A5}$$

We use a cubic-spline interpolation with the Steffen algorithm implemented in the GVAR Python package [105] to interpolate the correlation functions.

In Fig. 18 (top left), we compare the $a_\mu^{ll,\mathrm{W}}$(conn.) integrand on the 0.09 fm ensemble obtained from the raw correlator data (blue circles), $C_{\mathrm{fit}}(t)$ (purple line), $C_{\text{no osc.}}(t)$ (orange line), $C_{\mathrm{IPA}}(t)$ (green line), and $C_{\mathrm{interp}}(t)$ (red line). We find that the $C_{\mathrm{fit}}(t)$ integrand is in excellent agreement with the raw data in the region of interest, suggesting that $C_{\text{no osc.}}(t)$ is an accurate representation of the correlation function without the oscillating contribution. However, we see some differences between the $C_{\mathrm{IPA}}(t)$ and $C_{\mathrm{interp}}(t)$ integrands and the $C_{\text{no osc.}}(t)$ integrand, especially at shorter times where a large number of excited states contribute.

Figure 18 (top right) examines the lattice spacing dependence of the $a_\mu^{ll,\mathrm{W}}$(conn.) data obtained with all of the different oscillation removal techniques. In the case of $a_\mu^{ll,\mathrm{W}}$(conn.) data obtained from $C_{\text{no osc.}}(t)$ we see only small deviations compared to the $a_\mu^{ll,\mathrm{W}}$(conn.) from the raw data which are more significant at coarser lattice spacing. As a result, the continuum extrapolations (which use a simple linear fit in $a^2\alpha_s(2/a)$ to the three finest ensembles, leaving out the 0.15 fm data point) of the two data sets are in excellent agreement. While the IPA method does yield a consistent result in the continuum limit, it exibits much larger discretization effects. The interpolation method modifies the lattice spacing dependence so significantly that a linear fit is not enough to describe the observed behavior. This is likely due to the interpolation scheme not capturing the high energy state contributions sufficiently. In Fig. 18 (bottom), we compare these methods applied to $a_\mu^{ll,\mathrm{W2}}$(conn.); here the different

TABLE VI. $a_\mu^{ll,\mathrm{W}}$(conn.) and $a_\mu^{ll,\mathrm{W2}}$(conn.) computed from the raw data (columns two and five), the fit reconstruction without oscillating states (columns three and six) and the correlated difference between them (columns four and seven).

$\approx a$	$a_\mu^{ll,\mathrm{W}}$(conn.)	$a_{\mu,\,\mathrm{No\ osc.}}^{ll,\mathrm{W}}$(conn.)	$\Delta a_\mu^{ll,\mathrm{W}}$(conn.)	$a_\mu^{ll,\mathrm{W2}}$(conn.)	$a_{\mu,\,\mathrm{No\ osc.}}^{ll,\mathrm{W2}}$(conn.)	$\Delta a_\mu^{ll,\mathrm{W2}}$(conn.)
0.15	211.01(79)	210.62(79)	0.39(20)	80.3(1.7)	80.2(1.7)	0.13(19)
0.12	207.13(60)	207.34(59)	−0.204(34)	84.7(1.5)	84.6(1.5)	0.10(11)
0.09	206.56(55)	206.56(55)	0.001(10)	92.7(1.8)	92.7(1.8)	−0.07(22)
0.06	206.22(61)	206.22(61)	0.003(60)	95.6(2.8)	95.5(2.7)	0.12(73)

TABLE VII. $a_\mu^{ll,\mathrm{W}}$(conn.) and $a_\mu^{ll,\mathrm{W2}}$(conn.) computed from the fit reconstruction without oscillating states with the trapezoidal rule (columns two and five), Simpson's rule (columns three and six) and the correlated difference between them (columns four and seven).

$\approx a$	$a_{\mu,\ \mathrm{Simp.}}^{ll,\mathrm{W}}$(conn.)	$a_{\mu,\ \mathrm{Trap.}}^{ll,\mathrm{W}}$(conn.)	$\Delta a_\mu^{ll,\mathrm{W}}$(conn.)	$a_{\mu,\ \mathrm{Simp.}}^{ll,\mathrm{W2}}$(conn.)	$a_{\mu,\ \mathrm{Trap.}}^{ll,\mathrm{W2}}$(conn.)	$\Delta a_\mu^{ll,\mathrm{W2}}$(conn.)
0.15	210.62(79)	210.07(77)	0.55(26)	80.2(1.7)	79.5(1.7)	0.70(16)
0.12	207.34(59)	206.96(61)	0.373(49)	84.6(1.5)	84.9(1.5)	0.254(59)
0.09	206.56(55)	206.60(55)	−0.039(12)	92.7(1.8)	92.7(1.8)	−0.01(13)
0.06	206.22(61)	206.22(61)	0.0002(691)	95.5(2.7)	95.5(2.7)	−0.0003(4285)

methods give nearly identical results because the oscillations are less pronounced and fewer excited states contribute significantly.

In order to quantify the effects of the oscillations in $a_\mu^{ll,\mathrm{W}}$(conn.) and $a_\mu^{ll,\mathrm{W2}}$(conn.), we use the fit approach, our preferred method of removing them, in Table VI, where we compare results and correlated differences obtained using the trapezoidal rule (see Sec. III B) for the raw correlation function data vs. $C_{\mathrm{no\ osc.}}(t)$. For $a_\mu^{ll,\mathrm{W}}$(conn.), we find the differences to be small but statistically significant on the coarsest two ensembles and statistically zero on the finer ones. For $a_\mu^{ll,\mathrm{W2}}$(conn.), we find the differences to be zero on all ensembles, which is expected because the oscillating contributions are from heavier states which contribute significantly less in the long time region.

Finally, we examine the truncation effects associated with the trapezoidal rule by comparing a_μ observables computed from it to results obtained with Simpson's rule. Simpson's rule cannot be applied to the raw correlation function data because of the presence of oscillatory contributions. Hence, the comparisons in Table VII employ the $C_{\mathrm{no\ osc.}}(t)$ correlation functions. Here, the differences are within errors on $a_\mu^{ll,\mathrm{W}}$(conn.) and $a_\mu^{ll,\mathrm{W2}}$(conn.) and decrease much faster than a^2.

In summary, truncation effects from numerical integration and discretization effects due to the oscillatory contributions are clearly well-controlled and small compared to other systematic effects. To make certain that any systematic error due to integration and removal of oscillatory states is included, we include variations on both numerical integration and removal of oscillatory contributions in our main analysis, as described in Sec. III E.

Appendix B: Chiral-model expressions for the Euclidean-space vacuum polarization function

In this appendix, we provide expressions for calculating lattice corrections to $a_\mu^{\rm HVP,LO}$ (and windows thereof) within the chiral model of pions, photons, and ρ mesons denoted "CM" in Sec. III and employed in our 2019 work [43].

We begin with Blum's formulation of the $\mathcal{O}(\alpha^2)$ Standard-Model HVP contribution as an integral over the Euclidean-space momentum transfer Q^2 [30]

$$a_\mu^{\rm HVP,LO} = 4\alpha^2 \int_0^\infty {\rm d}Q^2 K_E\big(Q^2\big)\widehat{\Pi}\big(Q^2\big), \tag{B1}$$

where $\hat{\Pi}\left(Q^2\right) = \Pi\left(Q^2\right) - \Pi(0)$ is the renormalized vacuum polarization function and the integration kernel $K_E(Q^2)$ depends on the muon's mass:

$$K_E\left(Q^2\right) = \frac{m_\mu^2 Q^2 Z^3(1-Q^2Z)}{1+m_\mu^2}, \qquad Z = -\frac{Q^2-(Q^4+4m_\mu^2Q^2)^{1/2}}{2m_\mu^2Q^2}. \tag{B2}$$

In the chiral model [48, 86], the renormalized light-quark hadronic vacuum polarization function is given by

$$\hat{\Pi}\left(Q^2\right) = -\hat{\Sigma}\left(Q^2\right) + \frac{\hat{f}_\rho^2}{2\hat{m}_\rho^2}\frac{q^2\left(1+g_\rho g_{\rho\pi\pi}\hat{\Sigma}\left(Q^2\right)\right)^2}{Q^2\left(1+g_{\rho\pi\pi}^2\hat{\Sigma}\left(Q^2\right)\right)+\hat{m}_\rho^2}, \tag{B3}$$

where $\hat{\Sigma}\left(Q^2\right) \equiv {\rm Re}\,\Sigma\left(Q^2\right) - \Sigma(0)$ is the renormalized photon self energy and $\hat{m}_\rho$ ($\hat{f}_\rho$) are the renormalized ρ-meson mass (decay constant). In the chiral model, the leading contribution to $\Sigma\left(Q^2\right)$ arises from $\pi\pi$ loops, and is given by the integral

$$-\hat{\Sigma}\left(Q^2, m_a, m_b\right) \equiv \frac{4Q^2}{3}\int \frac{{\rm d}^3\mathbf{k}}{(2\pi)^3 2E_aE_b}\frac{\mathbf{k}^2}{(E_a+E_b)^3\left(Q^2+(E_a+E_b)^2\right)}\,, \tag{B4}$$

where m_a, m_b are the masses of the two pions in the loop. The renormalized ρ parameters can be expressed in terms of the bare mass, $\rho\gamma$ coupling, $\rho\pi\pi$ coupling, and $\Sigma(0)$ as

$$\hat{m}_\rho^2 \equiv m_{0\rho}^2\left(1-g_{\rho\pi\pi}^2\Sigma(0)\right) \tag{B5}$$

$$\frac{\hat{f}_\rho}{\hat{m}_\rho} \equiv \frac{\sqrt{2}}{g_\rho}\left(1+g_\rho g_{\rho\pi\pi}\Sigma(0) - \frac{1}{2}g_{\rho\pi\pi}^2\Sigma(0)\right)\,. \tag{B6}$$

We take the values of the bare parameters from Ref. [48]:

$$m_{0\rho} = 0.766{\rm GeV} \quad g_\rho = 5.4 \quad g_{\rho\pi\pi} = 6.0. \tag{B7}$$

In the chiral model, lattice effects are incorporated by modifying the pion self energy in two ways. First, to account for the finite volume, the continuous momentum integrals in Eq. (B4) are replaced by sums over the discrete lattice momenta, *i.e.*,

$$\int \frac{{\rm d}^3\mathbf{k}}{(2\pi)^3} \to \frac{1}{L^3}\sum_{k_x=-\infty}^{\infty}\sum_{k_y=-\infty}^{\infty}\sum_{k_z=-\infty}^{\infty} \tag{B8}$$

where $k_i = \frac{2\pi}{L} n_i$, $n_i = 1, 2, \ldots$. Second, taste-breaking effects are incorporated by replacing the renormalized photon self energy with an average over sea-pion tastes [43, 48]

$$\hat{\Sigma}\left(Q^2, m_\pi, m_\pi\right) \;\to\; \frac{1}{16} \sum_{\xi_a, \xi_b} \hat{\Sigma}\left(Q^2, m_\pi\left(\xi_a\right), m_\pi\left(\xi_b\right)\right). \tag{B9}$$

As stated in Sec. III C, for the analysis in this work we also include taste-breaking contributions to $\Sigma(0)$ via the replacement

$$\Sigma\left(0, m_\pi, m_\pi\right) \to \frac{1}{16} \sum_{\xi_a, \xi_b} \Sigma\left(0, m_\pi\left(\xi_a\right), m_\pi\left(\xi_b\right)\right) \;. \tag{B10}$$

Finally, the windowed HVP can be computed in the chiral model via [29]

$$\hat{\Pi}\left(Q^2\right) \to \hat{\Pi}_{\text{win.}}\left(Q^2\right) = \int_{-\infty}^{\infty} \frac{\mathrm{d}P}{2\pi} \frac{1}{Q^2} \left[\widetilde{\mathcal{W}}(P-Q) - \widetilde{\mathcal{W}}(P) - \frac{Q^2}{2} \frac{\mathrm{d}^2 \widetilde{\mathcal{W}}(P)}{\mathrm{d}P^2}\right] P^2 \hat{\Pi}\left(P^2\right), \tag{B11}$$

where $\hat{\Pi}$ is given by Eq. (B3) and $\widetilde{\mathcal{W}}$ is the Fourier transform of the window function defined in Eq. (2.7).

---

[1] B. Abi *et al.* (Muon $g-2$), Phys. Rev. Lett. **126**, 141801 (2021).
[2] G. W. Bennett *et al.* (Muon $g-2$), Phys. Rev. **D73**, 072003 (2006).
[3] T. Aoyama *et al.*, Phys. Rept. **887**, 1 (2020).
[4] T. Aoyama, M. Hayakawa, T. Kinoshita, and M. Nio, Phys. Rev. Lett. **109**, 111808 (2012).
[5] T. Aoyama, T. Kinoshita, and M. Nio, Atoms **7**, 28 (2019).
[6] A. Czarnecki, W. J. Marciano, and A. Vainshtein, Phys. Rev. **D67**, 073006 (2003), [Erratum: Phys. Rev. **D73**, 119901 (2006)].
[7] C. Gnendiger, D. Stöckinger, and H. Stöckinger-Kim, Phys. Rev. **D88**, 053005 (2013).
[8] M. Davier, A. Hoecker, B. Malaescu, and Z. Zhang, Eur. Phys. J. C **77**, 827 (2017).
[9] A. Keshavarzi, D. Nomura, and T. Teubner, Phys. Rev. D **97**, 114025 (2018).
[10] G. Colangelo, M. Hoferichter, and P. Stoffer, JHEP **02** (2019), 006.
[11] M. Hoferichter, B.-L. Hoid, and B. Kubis, JHEP **08** (2019), 137.
[12] M. Davier, A. Hoecker, B. Malaescu, and Z. Zhang, Eur. Phys. J. C **80**, 241 (2020), [Erratum: Eur. Phys. J. C **80**, 410 (2020)].
[13] A. Keshavarzi, D. Nomura, and T. Teubner, Phys. Rev. D **101**, 014029 (2020).
[14] A. Kurz, T. Liu, P. Marquard, and M. Steinhauser, Phys. Lett. B **734**, 144 (2014).
[15] K. Melnikov and A. Vainshtein, Phys. Rev. **D70**, 113006 (2004).
[16] P. Masjuan and P. Sánchez-Puertas, Phys. Rev. **D95**, 054026 (2017).
[17] G. Colangelo, M. Hoferichter, M. Procura, and P. Stoffer, JHEP **04** (2017), 161.
[18] M. Hoferichter, B.-L. Hoid, B. Kubis, S. Leupold, and S. P. Schneider, JHEP **10** (2018), 141.
[19] A. Gérardin, H. B. Meyer, and A. Nyffeler, Phys. Rev. **D100**, 034520 (2019).
[20] J. Bijnens, N. Hermansson-Truedsson, and A. Rodríguez-Sánchez, Phys. Lett. **B798**, 134994 (2019).
[21] G. Colangelo, F. Hagelstein, M. Hoferichter, L. Laub, and P. Stoffer, JHEP **03** (2020), 101.

[22] T. Blum, N. Christ, M. Hayakawa, T. Izubuchi, L. Jin, C. Jung, and C. Lehner (RBC), Phys. Rev. Lett. **124**, 132002 (2020).
[23] G. Colangelo, M. Hoferichter, A. Nyffeler, M. Passera, and P. Stoffer, Phys. Lett. **B735**, 90 (2014).
[24] M. Abe *et al.*, PTEP **2019**, 053C02 (2019).
[25] J-PARC muon $g-2$/EDM experiment, `https://g-2.kek.jp/`.
[26] A. Keshavarzi, K. S. Khaw, and T. Yoshioka, Nucl. Phys. B **975**, 115675 (2022).
[27] C. Bouchiat and L. Michel, J. Phys. Radium **22**, 121 (1961).
[28] S. J. Brodsky and E. de Rafael, Phys. Rev. **168**, 1620 (1968).
[29] S. Borsanyi *et al.* (BMW), Nature **593**, 51 (2021).
[30] T. Blum, Phys. Rev. Lett. **91**, 052001 (2003).
[31] D. Bernecker and H. B. Meyer, Eur. Phys. J. A **47**, 148 (2011).
[32] T. Blum *et al.* (RBC, UKQCD), Phys. Rev. Lett. **121**, 022003 (2018).
[33] C. Aubin, T. Blum, M. Golterman, and S. Peris, Phys. Rev. D **106**, 054503 (2022).
[34] C. T. H. Davies *et al.* (Fermilab Lattice, HPQCD, MILC), Phys. Rev. D **106**, 074509 (2022).
[35] G. Colangelo, A. X. El-Khadra, M. Hoferichter, A. Keshavarzi, C. Lehner, P. Stoffer, and T. Teubner, Phys. Lett. B **833**, 137313 (2022).
[36] M. Cè *et al.*, Phys. Rev. D **106**, 114502 (2022).
[37] C. Alexandrou *et al.* (ETM), arXiv:2206.15084 [hep-lat] (2022).
[38] T. Blum *et al.*, arXiv:2301.08696 [hep-lat] (2023).
[39] B. Chakraborty *et al.* (Fermilab Lattice, HPQCD, MILC), Phys. Rev. Lett. **120**, 152001 (2018).
[40] S. Yamamoto, C. DeTar, A. X. El-Khadra, C. McNeile, R. S. Van de Water, and A. Vaquero (Fermilab Lattice, HPQCD, MILC), PoS **LATTICE2018**, 322 (2019).
[41] C. E. DeTar *et al.* (Fermilab Lattice, HPQCD, MILC), PoS **LATTICE2019**, 070 (2019).
[42] C. McNeile *et al.* (Fermilab Lattice, HPQCD, MILC), PoS **LATTICE2021**, 039 (2022).
[43] C. Davies *et al.* (Fermilab Lattice, HPQCD, MILC), Phys. Rev. D **101**, 034512 (2020).
[44] A. Bazavov *et al.* (MILC), Phys. Rev. D **87**, 054505 (2013).
[45] E. Follana *et al.* (HPQCD, UKQCD), Phys. Rev. D **75**, 054502 (2007).
[46] W. I. Jay and E. T. Neil, Phys. Rev. D **103**, 114502 (2021).
[47] E. T. Neil and J. W. Sitison, arXiv:2208.14983 [stat.ME] (2022).
[48] B. Chakraborty, C. T. H. Davies, P. G. de Oliviera, J. Koponen, G. P. Lepage, and R. S. Van de Water (HPQCD), Phys. Rev. D **96**, 034516 (2017).
[49] S. Lahert *et al.* (Fermilab Lattice, HPQCD, MILC), PoS **LATTICE2021**, 526 (2022).
[50] G. Ray *et al.* (Fermilab Lattice, HPQCD, MILC), in *39th International Symposium on Lattice Field Theory* (2022) arXiv:2212.12031 [hep-lat].
[51] C. Aubin *et al.*, Phys. Rev. D **70**, 031504 (2004).
[52] C. Aubin, C. Bernard, C. E. DeTar, J. Osborn, S. Gottlieb, E. B. Gregory, D. Toussaint, U. M. Heller, J. E. Hetrick, and R. Sugar (MILC), Phys. Rev. D **70**, 114501 (2004).
[53] A. Bazavov *et al.* (Fermilab Lattice, MILC), Phys. Rev. D **90**, 074509 (2014).
[54] A. Bazavov *et al.* (Fermilab Lattice, MILC), Phys. Rev. D **98**, 074512 (2018).
[55] C. Patrignani *et al.* (Particle Data Group), Chin. Phys. C **40**, 100001 (2016).
[56] K. Symanzik, Nucl. Phys. B **226**, 205 (1983).
[57] M. Lüscher and P. Weisz, Commun. Math. Phys. **97**, 59 (1985), [Erratum: Commun. Math. Phys. 98, 433 (1985)].
[58] M. Lüscher and P. Weisz, Phys. Lett. B **158**, 250 (1985).

[59] M. G. Alford, W. Dimm, G. P. Lepage, G. Hockney, and P. B. Mackenzie, Phys. Lett. B **361**, 87 (1995).
[60] A. Hart, G. M. von Hippel, and R. R. Horgan (HPQCD), Phys. Rev. D **79**, 074008 (2009).
[61] G. P. Lepage and P. B. Mackenzie, Phys. Rev. D **48**, 2250 (1993).
[62] A. Bazavov *et al.* (MILC), Phys. Rev. D **82**, 074501 (2010).
[63] S. Borsanyi *et al.* (BMW), JHEP **09** (2012), 010.
[64] R. J. Dowdall, C. T. H. Davies, G. P. Lepage, and C. McNeile (HPQCD), Phys. Rev. D **88**, 074504 (2013).
[65] B. Chakraborty, C. T. H. Davies, G. C. Donald, J. Koponen, and G. P. Lepage (HPQCD), Phys. Rev. D **96**, 074502 (2017).
[66] J. Gasser, A. Rusetsky, and I. Scimemi, Eur. Phys. J. C **32**, 97–114 (2003).
[67] G. S. Bali, S. Collins, and A. Schäfer, Comput. Phys. Commun. **181**, 1570–1583 (2010).
[68] C. Aubin, T. Blum, P. Chau, M. Golterman, S. Peris, and C. Tu, Phys. Rev. D **93**, 054508 (2016).
[69] J. Bijnens and J. Relefors, JHEP **12** (2017), 114.
[70] C. Aubin, T. Blum, M. Golterman, and S. Peris, Phys. Rev. D **102**, 094511 (2020).
[71] C. Aubin, T. Blum, C. Tu, M. Golterman, C. Jung, and S. Peris, Phys. Rev. D **101**, 014503 (2020).
[72] G. J. Gounaris and J. J. Sakurai, Phys. Rev. Lett. **21**, 244 (1968).
[73] M. Lüscher, Nucl. Phys. B **364**, 237 (1991).
[74] M. Lüscher, Nucl. Phys. B **354**, 531 (1991).
[75] L. Lellouch and M. Lüscher, Commun. Math. Phys. **219**, 31 (2001).
[76] C. J. D. Lin, G. Martinelli, C. T. Sachrajda, and M. Testa, Nucl. Phys. B **619**, 467 (2001).
[77] H. B. Meyer, Phys. Rev. Lett. **107**, 072002 (2011).
[78] A. Francis, B. Jäger, H. B. Meyer, and H. Wittig, Phys. Rev. D **88**, 054502 (2013).
[79] M. Della Morte, A. Francis, V. Gülpers, G. Herdoíza, G. von Hippel, H. Horch, B. Jäger, H. B. Meyer, A. Nyffeler, and H. Wittig, JHEP **10** (2017), 020.
[80] M. T. Hansen and A. Patella, JHEP **10** (2020), 029.
[81] C. Aubin, C. Bernard, C. DeTar, J. Osborn, S. Gottlieb, E. B. Gregory, D. Toussaint, U. M. Heller, J. E. Hetrick, and R. Sugar (MILC), Phys. Rev. D **70**, 094505 (2004).
[82] A. Bazavov *et al.* (MILC), Rev. Mod. Phys. **82**, 1349 (2010).
[83] W.-J. Lee and S. R. Sharpe, Phys. Rev. D **60**, 114503 (1999).
[84] C. Bernard (MILC), Phys. Rev. D **65**, 054031 (2002).
[85] B. Chakraborty *et al.* (HPQCD), Phys. Rev. D **91**, 054508 (2015).
[86] F. Jegerlehner and R. Szafron, Eur. Phys. J. C **71** (2011).
[87] A. Gérardin, M. Cè, G. von Hippel, B. Hörz, H. B. Meyer, D. Mohler, K. Ottnad, J. Wilhelm, and H. Wittig, Phys. Rev. D **100**, 014510 (2019).
[88] M. Golterman, K. Maltman, and S. Peris, EPJ Web Conf. **175**, 06010 (2018).
[89] G. Lepage and C. Gohlke, `gplepage/lsqfit` (2022).
[90] G. P. Lepage *et al.*, Nucl. Phys. B Proc. Suppl. **106**, 12 (2002).
[91] A. Bazavov *et al.* (Fermilab Lattice, MILC), Phys. Rev. D **93**, 113016 (2016).
[92] G. Wang, T. Draper, K.-F. Liu, and Y.-B. Yang (χQCD), arXiv:2204.01280 [hep-lat] (2022).
[93] C. Lehner and A. S. Meyer, Phys. Rev. D **101**, 074515 (2020).
[94] B. Chakraborty, C. T. H. Davies, G. C. Donald, R. J. Dowdall, J. Koponen, G. P. Lepage, and T. Teubner (HPQCD), Phys. Rev. D **89**, 114501 (2014).
[95] D. Hatton, C. T. H. Davies, B. Galloway, J. Koponen, G. P. Lepage, and A. T. Lytle

(HPQCD), Phys. Rev. D **102**, 054511 (2020).
[96] D. Hatton, C. T. H. Davies, J. Koponen, G. P. Lepage, and A. T. Lytle (HPQCD), Phys. Rev. D **103**, 054512 (2021).
[97] D. Giusti, V. Lubicz, G. Martinelli, F. Sanfilippo, and S. Simula (ETM), Phys. Rev. D **99**, 114502 (2019).
[98] M. T. Lynch and C. DeTar, *poster presented at the 39th International Symposium on Lattice Field Theory* (2022).
[99] Y. Lin, A. S. Meyer, C. Hughes, A. S. Kronfeld, J. N. Simone, and A. Strelchenko (Fermilab Lattice), Phys. Rev. D **103**, 034501 (2021).
[100] F. Erben, J. R. Green, D. Mohler, and H. Wittig, Phys. Rev. D **101**, 054504 (2020), arXiv:1910.01083 [hep-lat].
[101] M. Bruno, T. Izubuchi, C. Lehner, and A. S. Meyer, PoS **LATTICE2019**, 239 (2019), arXiv:1910.11745 [hep-lat].
[102] J. A. Bailey *et al.* (Fermilab Lattice, MILC), Phys. Rev. D **79**, 054507 (2009).
[103] R. L. Workman *et al.* (Particle Data Group), PTEP **2022**, 083C01 (2022).
[104] G. Aarts, C. Allton, J. Foley, S. Hands, and S. Kim, Phys. Rev. Lett. **99**, 022002 (2007).
[105] G. Lepage, C. Gohlke, and D. Hackett, `gplepage/gvar v11.10` (2022).