Pion transition form factor from twisted-mass lattice QCD and the hadronic light-by-light π^0-pole contribution to the muon $g-2$

C. Alexandrou,[1,2] S. Bacchio,[2] G. Bergner,[3] S. Burri,[4] J. Finkenrath,[5,2] A. Gasbarro,[4] K. Hadjiyiannakou,[1,2] K. Jansen,[6] G. Kanwar,[4] B. Kostrzewa,[7] G. Koutsou,[2] K. Ottnad,[8] M. Petschlies,[9,10] F. Pittler,[2] F. Steffens,[9,10] C. Urbach,[9,10] and U. Wenger[4]

(Extended Twisted Mass Collaboration)

[1]*Department of Physics, University of Cyprus, 20537 Nicosia, Cyprus*
[2]*Computation-based Science and Technology Research Center, The Cyprus Institute, 20 Konstantinou Kavafi Street, 2121 Nicosia, Cyprus*
[3]*University of Jena, Institute for Theoretical Physics, Max-Wien-Platz 1, D-07743 Jena, Germany*
[4]*Albert Einstein Center for Fundamental Physics, Institute for Theoretical Physics, University of Bern, Sidlerstrasse 5, CH–3012 Bern, Switzerland*
[5]*Bergische Universität Wuppertal, Gaußstraße 20, 42119 Wuppertal, Germany*
[6]*Deutsches Elektronen-Synchrotron DESY, Platanenallee 6, 15738 Zeuthen, Germany*
[7]*High Performance Computing and Analytics Lab, Rheinische Friedrich-Wilhelms-Universität Bonn, Friedrich-Hirzebruch-Allee 8, D-53115 Bonn, Germany*
[8]*PRISMA⁺ Cluster of Excellence and Institut für Kernphysik, Johannes Gutenberg-Universität Mainz, Johann-Joachim-Becher-Weg 45, D-55128 Mainz, Germany*
[9]*HISKP (Theory), Rheinische Friedrich-Wilhelms-Universität Bonn, Nussallee 14-16, D-53115 Bonn, Germany*
[10]*Bethe Center for Theoretical Physics, Rheinische Friedrich-Wilhelms-Universität Bonn, Wegelerstraße 10, D-53115 Bonn, Germany*

The neutral pion generates the leading pole contribution to the hadronic light-by-light tensor, which is given in terms of the nonperturbative transition form factor $\mathcal{F}_{\pi^0\gamma\gamma}(q_1^2, q_2^2)$. Here we present an *ab-initio* lattice calculation of this quantity in the continuum and at the physical point using twisted-mass lattice QCD. We report our results for the transition form factor parametrized using a model-independent conformal expansion valid for arbitrary spacelike kinematics and compare it with experimental measurements of the single-virtual form factor, the two-photon decay width, and the slope parameter. We then use the transition form factors to compute the pion-pole contribution to the hadronic light-by-light scattering in the muon $g-2$, finding $a_\mu^{\pi^0-\text{pole}} = 56.7(3.2) \times 10^{-11}$.

I. INTRODUCTION

The anomalous magnetic moment of the muon (the muon $g-2$) provides a stringent test of the Standard Model at high precision and the possibility to glimpse subtle effects of potential new physics beyond the Standard Model. Recent results from the Fermilab experiment [1,2] in combination with the Brookhaven result [3] have yielded the most precise experimental determination to date at the level of 19 ppm: $a_\mu(\text{exp}) = 116592059(22) \times 10^{-11}$, where $a_\mu \equiv (g-2)_\mu/2$. A comparable precision is striven for on the theoretical side with the recent state of affairs summarized in the white paper [4].

The leading hadronic contributions to the muon anomalous magnetic moment come from diagrams involving the hadronic vacuum polarization (HVP) and the hadronic light-by-light (HLbL) scattering tensors. Both make significant contributions at the level of precision achieved by the experimental results. It is important to improve the theoretical determinations of both contributions to match future targets of experimental precision; this work addresses the HLbL contribution.

Two complementary approaches to the theoretical determination of the HLbL contribution are becoming recognizable for reliable estimates. On one hand, several direct lattice calculations have analyzed the complete HLbL tensor [5–11]. On the other hand, the HLbL tensor can be decomposed into contributions from exchanges of various intermediate on-shell states in a data-driven dispersive approach, where contributions from heavier intermediate states are suppressed [12–15]. The latter approach has provided the most precise theoretical estimates of the HLbL contribution to the muon $g - 2$ to date [4], while the former has the benefit of allowing for a fully self-contained lattice QCD calculation from first principles without resorting to chiral effective field theory and experimental data. However, there has been an ongoing effort to provide also *ab-initio* data based on lattice QCD calculations for the leading contributions within the dispersive decomposition of the HLbL.

The leading single-pole contributions to the HLbL component of the muon $g - 2$ arise from the exchange of neutral pseudoscalar mesons $P \in \{\pi^0, \eta, \eta'\}$ as shown in Fig. 1. The key nonperturbative input to these diagrams are the transition form factors (TFFs) $\mathcal{F}_{P\gamma\gamma}(q_1^2, q_2^2)$, defined by [16]

$$\epsilon_{\mu\nu\alpha\beta} q^\alpha p^\beta \mathcal{F}_{P\gamma\gamma}(q_1^2, q_2^2) \equiv M_{\mu\nu}(p, q_1)$$
$$= i \int d^4x e^{iq_1 x} \langle 0|T\{j_\mu(x)j_\nu(0)\}$$
$$\times |P(p)\rangle, \qquad (1)$$

where

$$j_\mu(x) = \bar{\psi}(x)\gamma_\mu Q\psi(x) \qquad (2)$$

is the electromagnetic current written in terms of the charge matrix Q. The momentum of the pseudoscalar meson satisfies the on-shell condition $p^2 = M_P^2$, while the photon momenta q_1 and $q_2 = p - q_1$ may be either on-shell or virtual. The pole contribution to the HLbL tensor from the pseudoscalar meson P is given by
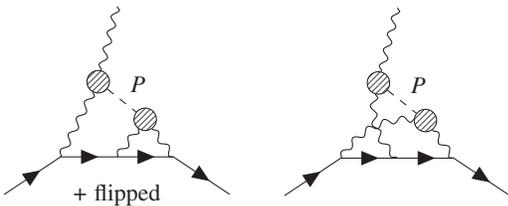
FIG. 1. Pseudoscalar-pole diagrams contributing to the leading order hadronic light-by-light (HLbL)scattering. Each striped circle represents a nonperturbative $P \to \gamma^*\gamma^*$ transition form factor required to evaluate these contributions, where "P" indicates the exchanged pseudoscalar meson. The "flipped" contribution signifies the diagram with the pseudoscalar and photon portions of the diagram horizontally flipped.

$$\Pi_{\mu\nu\lambda\rho}^{(P)}(q_1, q_2, q_3) = i \frac{\mathcal{F}_{P\gamma\gamma}(q_1^2, q_2^2)\mathcal{F}_{P\gamma\gamma}(q_3^2, (q_1 + q_2 + q_3)^2)}{(q_1 + q_2)^2 - M_P^2}$$
$$\times \epsilon_{\mu\nu\alpha\beta} q_1^\alpha q_2^\beta \epsilon_{\lambda\rho\sigma\tau} q_3^\sigma (q_1 + q_2)^\tau$$
$$+ (q_1 \leftrightarrow q_3) + (q_2 \leftrightarrow q_3). \qquad (3)$$

Integrating over spacelike $q_1, q_2,$ and $q_3 = -q_1 - q_2$ with the appropriate kinematical factors from the diagrams in Fig. 1 yields contributions to the total muon $g - 2$ from each pseudoscalar meson exchange.

The π^0-pole contribution to the muon $g - 2$ is larger than those of the η and η' by roughly a factor of four [17,18]. This feature, together with the relative simplicity of accessing the pion TFF in both lattice and data-driven approaches, means that the study of this transition form factor is farthest advanced at this point. Direct experimental data for $\mathcal{F}_{\pi^0\gamma\gamma}(q_1^2, q_2^2)$ is primarily available in the singly virtual (either $q_1^2 = 0$ or $q_2^2 = 0$) regime with large photon virtuality [19–23]. Fits to this data using Canterbury approximants have yielded a data-based estimate across all virtualities [24]. Meanwhile, the pion transition form factor (TFF) has also been determined from indirect experimental data using a dispersive approach [25,26]. On the other hand, an *ab-initio* determination is desirable to remove the dependence on experimental inputs. As such, the neutral pseudoscalar meson TFF has also been computed on the lattice. The calculation of the neutral-pion TFF on the lattice has been pioneered by the Mainz group using Wilson quarks [27,28]. The BMW Collaboration recently added their results using staggered quarks [29]. Our preliminary results with Wilson twisted-mass quarks appeared in Refs. [30,31]. The major sources of uncertainty in these approaches is the extrapolation to the continuum and, in case the quark masses are not tuned to their physical values as in Refs. [27,28], the extrapolation to the physical point. A first calculation of the η-meson TFF and the corresponding pole contribution at one lattice spacing has been reported by our Collaboration in Ref. [32] and results for both the η- and η'-meson TFF and pole contributions addressing all systematic errors have been presented by the BMW Collaboration in Ref. [29].

In the present work, we complement these existing efforts by evaluating the pion transition form factor in the continuum using the twisted-mass lattice QCD discretization, with ensembles generated directly at the physical point. This calculation serves to strengthen the lattice consensus of the continuum limit and minimizes systematic effects from extrapolation of the pion mass. In this work, we present a parametrization of the pion transition form factor and use it to calculate the π^0-pole contribution to the muon $g - 2$. As our main result, we find $a_\mu^{\pi^0-\text{pole}} = 56.7(3.2) \times 10^{-11}$. We further use the transition form factor to estimate the two-photon decay width as $\Gamma(\pi \to \gamma\gamma) = 7.50(0.50)$ eV, and the slope parameter as

$b_\pi = 2.16(0.20)$ GeV^{-2}, which provides input for determining the electromagnetic interaction radius. Quoted uncertainties include statistical errors and systematic uncertainty from analysis choices and the continuum extrapolation.

The structure of this paper is as follows. In Sec. II, we review the theoretical background for the extraction of the required amplitude from lattice QCD simulations. In Sec. III, we describe the calculation of the transition form factors and their extrapolation to photon virtualities required for the observables that we calculate. In Sec. IV, we present the calculation of $a_\mu^{\pi^0-\text{pole}}$ from the transition form factors and the determination of the decay width $\Gamma(\pi \to \gamma\gamma)$ and the slope parameter b_π. Next, in Sec. V, we discuss the estimation of statistical and systematic errors based on model averaging, before presenting our results in Sec. VI, comparing them to experimental and theoretical predictions. Then, in Sec. VIII, we make some concluding remarks and give an outlook on possible future directions for improving the calculation presented here.

II. EUCLIDEAN AMPLITUDE

By analytic continuation, the Minkowski neutral pion transition form factor in Eq. (1) is related to the Euclidean amplitude

$$\tilde{A}_{\mu\nu}(\tau) = \langle 0|T\{j_\mu(\vec{q}_1, \tau)j_\nu(\vec{p} - \vec{q}_1, 0)\}|\pi^0(p)\rangle \quad (4)$$

by [33–35]

$$M_{\mu\nu}(p, q_1) = i^{n_0} M_{\mu\nu}^E(p, q_1), \quad M_{\mu\nu}^E = \int_{-\infty}^{\infty} d\tau\, e^{\omega_1 \tau} \tilde{A}_{\mu\nu}(\tau),$$
$$(5)$$

where $n_0 = \delta_{\mu,0} + \delta_{\nu,0}$, $q_1 = (\omega_1, \vec{q}_1)$, and τ indicates a Euclidean time coordinate. Note that $\tilde{A}_{\mu\nu}$ depends on the momenta $\vec{q}_1$ and $\vec{p}$, but this dependence is suppressed for conciseness of notation. The analytic continuation is valid for virtualities q_1^2 and q_2^2 below the two-pion threshold at $4m_\pi^2$ [33]. In particular, this includes the entire spacelike quadrant $q_{1,2}^2 \le 0$ required for evaluation of the muon $g - 2$.

A. Three-point function

We access the amplitude $\tilde{A}_{\mu\nu}(\tau)$ through the following lattice three-point correlation function

$$C_{\mu\nu}(\tau, t_\pi) = a^6 \sum_{\vec{x}, \vec{z}} \langle j_\mu(\vec{x}, \tau) j_\nu(\vec{0}, 0) P^{0\dagger}(\vec{z}, -t_\pi)\rangle e^{-i\vec{q}_1 \cdot \vec{x}} e^{i\vec{p} \cdot \vec{z}}$$

$$= \langle j_\mu(\vec{q}_1, \tau) j_\nu(\vec{0}, 0) P^{0\dagger}(\vec{p}, -t_\pi)\rangle, \quad (6)$$

where $P^{0\dagger}$ is an interpolating operator for the neutral pion and $\langle \ldots \rangle$ indicates the Euclidean path-integral expectation

value. Translation invariance, or equivalently conservation of momentum, ensures that the current j_ν must carry momentum $\vec{q}_2 = \vec{p} - \vec{q}_1$. We denote the minimal separation between the pion creation operator and the nearest electromagnetic current by $t_{\text{sep}} = \min(\tau, 0) + t_\pi$. When $t_{\text{sep}} \gg 0$, or equivalently $t_\pi \gg 0, -\tau$, the pion creation operator $P^{0\dagger}(\vec{p}, -t_\pi)$ is in the distant Euclidean past, and the Euclidean matrix element in Eq. (4) can be recovered from the three-point function as

$$C_{\mu\nu}(\tau, t_\pi) \stackrel{t_{\text{sep}} \gg 0}{=} e^{-E_\pi t_\pi} \frac{Z_\pi}{2E_\pi} \langle 0|T\{j_\mu(\tau)j_\nu(0)\}|\pi^0(\vec{p})\rangle$$

$$+ O(e^{-\Delta E t_{\text{sep}}})$$

$$= e^{-E_\pi t_\pi} \frac{Z_\pi}{2E_\pi} \tilde{A}_{\mu\nu}(\tau) + O(e^{-\Delta E t_{\text{sep}}}), \quad (7)$$

where

$$Z_\pi \equiv \langle 0|P^0(\vec{0}, 0)|\pi^0(\vec{p})\rangle \quad (8)$$

and ΔE indicates the energy gap to the first excited state produced by $P^{0\dagger}(\vec{p}, -t_\pi)$. Given this limiting behavior, we define

$$\tilde{A}_{\mu\nu}(\tau; t_\pi) \equiv \frac{2E_\pi}{Z_\pi} e^{E_\pi t_\pi} C_{\mu\nu}(\tau, t_\pi)$$

$$= \tilde{A}_{\mu\nu}(\tau) + O(e^{-\Delta E t_{\text{sep}}}). \quad (9)$$

In practice, we compute this quantity for finite values of t_π taken sufficiently large to ensure the excited state contamination is suppressed below our uncertainties. When it is clear from context, we elide the t_π argument in the following.

The parameters E_π and Z_π are extracted from a separate analysis of the τ-dependence of the two-point function

$$C^{(2)}(\tau) \equiv a^3 \sum_{\vec{x}} e^{-i\vec{p}\cdot\vec{x}} \langle 0|P^0(\vec{x}, \tau)P^{0\dagger}(\vec{0}, 0)|0\rangle$$

$$= \frac{|Z_\pi|^2}{2E_\pi} e^{-E_\pi \tau} + O(e^{-\Delta E \tau}). \quad (10)$$

We note that for maximally twisted fermions the value of Z_π is strictly positive and can be unambiguously extracted from Eq. (10), because it is related to the pion decay constant f_π through $Z_\pi = f_\pi m_\pi \sinh(m_\pi)/2\mu_l$ where μ_l is the twisted light-quark mass parameter [36,37].

B. Twisted-mass lattice QCD

We perform our calculation using twisted-basis fermion fields, denoted χ. For the light quarks, these are related to the physical-basis doublet $\psi = (u, d)$ through the maximal twist transformation

$$\psi = e^{i\pi\gamma_5\tau^3/4}\chi, \qquad \bar{\psi} = \bar{\chi}e^{i\pi\gamma_5\tau^3/4}, \qquad (11)$$

where τ^3 is the generator of the third component of isospin, represented by the third Pauli matrix. In the continuum limit, the twist transformation is a trivial change of basis. At finite lattice spacing, however, the Wilson term breaks chiral invariance meaning the twisted basis provides an alternative path towards the continuum limit, with certain desirable properties such as automatic $O(a)$ improvement [36,38].

To relate calculations performed in the twisted basis to the physical amplitude and the TFF, we choose twisted basis operators which correspond to the physical-basis electromagnetic currents and pion operators. Specifically, we choose the pion annihilation operators to be

$$P^0(x) = i\bar{\psi}(x)\gamma_5\tau^3\psi(x) = -\bar{\chi}(x)\chi(x),$$
$$P^\pm(x) = i\bar{\psi}(x)\gamma_5\tau^\pm\psi(x) = i\bar{\chi}(x)\gamma_5\tau^\pm\chi(x), \qquad (12)$$

where $\tau^\pm$ are the raising/lowering operators of isospin. We also choose local electromagnetic currents, given in the physical basis in the light-quark sector by

$$j_\mu^{(l)}(x) = \bar{\psi}(x)\gamma_\mu Q^{(l)}\psi(x), \qquad (13)$$

where $Q^{(l)} = \mathrm{diag}(2/3, -1/3)$ is the light-quark charge matrix. This current can be decomposed into terms of total isospin $I = 1$ and $I = 0$ (always with fixed z-component $I_z = 0$), denoted respectively by $j^{1,0}$ and $j^{0,0}$. The result is then easily written in the twisted basis, giving

$$j_\mu^{(l)}(x) = \frac{1}{6}j_\mu^{0,0}(x) + \frac{1}{2}j_\mu^{1,0}(x)$$
$$= \frac{1}{6}\bar{\chi}(x)\gamma_\mu\chi(x) + \frac{1}{2}\bar{\chi}(x)\gamma_\mu\tau^3\chi(x). \qquad (14)$$

It will also be useful for the following section to define the currents with isospin $I = 1$ and $I_z = \pm 1$,

$$j_\mu^{1,\pm}(x) = \bar{\psi}(x)\gamma_\mu\tau^\pm\psi(x) = \mp i\bar{\chi}(x)\gamma_\mu\gamma_5\tau^\pm\chi(x). \qquad (15)$$

For the strange and charm quark we use a mixed-action approach. In the sea-quark action the strange and charm quark form a twisted-mass quark doublet, including tuned twisted-mass and quark-mass splitting terms. Details about tuning of the twisted-mass parameters for the sea-quark action, in particular to maximal twist, are given in Ref. [39]. In the valence action we use two Osterwalder-Seiler heavy-quark doublets [38] by introducing two doublets $\chi^{(h)} = (\chi_+^{(h)}, \chi_-^{(h)})$, for $h = s, c$, which avoids twisted-mass-related mixing of strange and charm flavor. Both doublets work analogously to the up- and down-quark doublet, they have the same critical hopping parameter, but use individual tuning of the twisted quark-mass parameters.
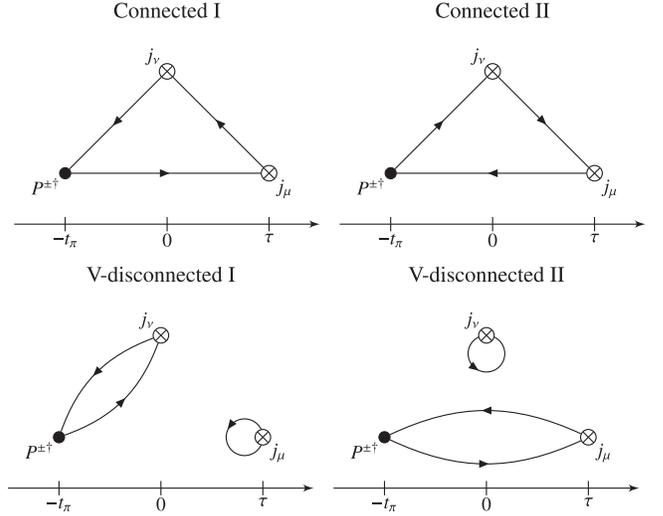
FIG. 2. Wick contractions contributing to $C_{\mu\nu}^\pm(\tau, t_P)$. These are the connected (top row), and V-disconnected (bottom row) diagrams. Note that the connected diagrams only involve light quarks, while the disconnected vector-current loop in the V-disconnected diagrams involves light, strange, and charm quarks for the $2 + 1 + 1$ ensembles used in this work.

The strange-quark mass is tuned such that the Ω baryon takes its physical value, while the charm-quark mass is tuned to reproduce a near-physical Λ_c mass.

The strange- and charm-flavor electromagnetic currents then use the local operators

$$j^{(s)} = -\frac{1}{3}\bar{\psi}^{(s)}(x)\gamma_\mu\psi^{(s)}(x) = -\frac{1}{3}\bar{\chi}^{(s)}(x)\gamma_\mu\chi^{(s)}(x),$$
$$j^{(c)} = \frac{2}{3}\bar{\psi}^{(c)}(x)\gamma_\mu\psi^{(c)}(x) = \frac{2}{3}\bar{\chi}^{(c)}(x)\gamma_\mu\chi^{(c)}(x). \qquad (16)$$

The isoscalar strange- and charm-flavor currents enter correlation functions in Fig. 2 only as vector-current disconnected diagrams "V-disconnected I" and "V-disconnected II" (details on the Wick contractions are provided in the following section). To maintain automatic $\mathcal{O}(a)$ improvement we average the vector-current disconnected contributions from strange and charm propagators generated from both χ_+^h and χ_-^h with $h = s, c$.

We use gauge ensembles produced by the Extended Twisted Mass Collaboration (ETMC) from simulations of isospin-symmetric QCD (isoQCD) with $N_f = 2 + 1 + 1$ flavors of Wilson clover twisted-mass quarks. The properties of the used ensembles most relevant for this work can be found in Table I; for more details see Refs. [39–43].

The local vector currents detailed above renormalize multiplicatively. Because of the twisted-mass formulation, isoscalar and neutral isovector vector currents are renormalized with renormalization factor Z_V, while the charged isovector vector currents renormalize with Z_A:

TABLE I. Parameters of the ETMC ensembles for the analysis presented in this work, adapted from [43]. Further parameters can be found in [39–43].

Ensemble	V/a^4	N_{conf}	β	a [fm]	$a\mu_l$	m_π [MeV]	L [fm]	$m_\pi L$
cB211.072.64	$64^3 \cdot 128$	748	1.778	0.07961(13)	0.00072	140.2(2)	5.09	3.62
cC211.060.80	$80^3 \cdot 160$	397	1.836	0.06821(12)	0.00060	136.7(2)	5.46	3.78
cD211.054.96	$96^3 \cdot 192$	495	1.900	0.05692(10)	0.00054	140.8(2)	5.46	3.90

$$j_{\mu,R}^{0,0} = Z_V j_\mu^{0,0}, \qquad j_{\mu,R}^{1,0} = Z_V j_\mu^{1,0}, \qquad (17)$$

$$j_{\mu,R}^{1,\pm} = Z_A j_\mu^{1,\pm}, \qquad (18)$$

$$j_{\mu,R}^{(s)} = Z_V j_\mu^{(s)}, \qquad j_{\mu,R}^{(c)} = Z_V j_\mu^{(c)}. \qquad (19)$$

These renormalization constants (RCs) were calculated to high precision for the ensembles used in this work in Ref. [43], as shown in Table II. The uncertainties on the RCs are negligible compared to the errors on the bare TFFs, such that we only use the central values of the RCs throughout this work.

Two remarks are in order due to the fact that we are using nonconserved local vector currents for the construction of the three-point correlation function in Eq. (6). First, one can show that possible short-distance singularities are absent, i.e., the local currents in Eq. (6) admit a well-defined continuum limit. The arguments are based on the operator product expansion and are given in Appendix D of Ref. [27] for Wilson fermions. Due to universality, they also apply to the Wilson twisted-mass lattice QCD formulation. Second, we note that the nonconserved local currents do not spoil the automatic $O(a)$ improvement. This can be seen from the fact that all involved lattice quantities are constructed such that they have the correct symmetry property under the combined twisted-mass transformation $\mathcal{P} \times \mathcal{D}_d \times (\mu \to -\mu)$ where $\mathcal{P}$ is the standard parity symmetry transformation, $\mathcal{D}_d$ the operator-dimensionality transformation, and μ the twisted-quark mass parameter. This symmetry forbids the appearance of $O(a)$ terms in physical matrix elements for twisted-mass lattice QCD at maximal twist [36,38], and hence guarantees automatic $O(a)$-improvement of the three-point function in Eq. (6).

TABLE II. Values of Z_V and Z_A for the ETMC ensembles which were used in this work, as determined in Ref. [43] using a hadronic method.

Ensemble	Z_V	Z_A
cB211.072.64	0.706379(24)	0.74294(24)
cC211.060.80	0.725404(19)	0.75830(16)
cD211.054.96	0.744108(12)	0.77395(12)

C. Wick contractions

A naive evaluation of the three-point amplitude Eq. (6) contains disconnected contributions in which the quarks within the neutral pion contract with themselves. These include a singly disconnected contribution which correlates the neutral-pion loop with the connected vector-vector two-point function, as well as a doubly disconnected contribution with three loops. In the limit of exact isospin symmetry, these contributions vanish identically. In the twisted-mass formulation at finite lattice spacing, however, isospin symmetry is broken by $O(a^2)$ lattice artifacts [37] (independently of the twist angle) and the disconnected contributions do not cancel at finite lattice spacing.

Since typically disconnected diagrams are costly to evaluate, we minimize the number of such diagrams by instead considering the isospin-rotated amplitude

$$
\begin{aligned}
C'_{\mu\nu} = &-\frac{1}{6}\langle j_\mu^{0,0} j_\nu^{1,-} P^+ \rangle + \frac{1}{6}\langle j_\mu^{0,0} j_\nu^{1,+} P^- \rangle \\
&+ \frac{1}{3}\langle j_\mu^{(s)} j_\nu^{1,-} P^+ \rangle - \frac{1}{3}\langle j_\mu^{(s)} j_\nu^{1,+} P^- \rangle \\
&- \frac{2}{3}\langle j_\mu^{(c)} j_\nu^{1,-} P^+ \rangle + \frac{2}{3}\langle j_\mu^{(c)} j_\nu^{1,+} P^- \rangle,
\end{aligned} \qquad (20)
$$

where spacetime coordinates have been suppressed for clarity. The amplitude in Eq. (20) does not contain pion-loop disconnected contributions, leaving only the connected and vector-current disconnected (V-disconnected) contributions depicted in Fig. 2. This rotation is valid in the continuum due to the iso-symmetric continuum limit of twisted-mass lattice QCD. Consequently, the rotation only modifies the lattice artifacts at $O(a^2)$ in the twisted-mass formulation, which is taken into account in the continuum extrapolation.

We evaluate the connected diagrams using inversions based on point sources for the current $j_\nu(\vec{0}, 0)$ with sequential inversion through the pseudoscalar operator $P^{\pm\dagger}$. This means that each choice of momentum $\vec{p}$ for the pseudoscalar requires a new inversion, while momentum $\vec{q}_1$ inserted at the sink current $j_\mu(\vec{x}, \tau)$ can be easily varied. For the evaluation of the connected contribution to the three-point amplitude, 16 point sources per configuration are used for cB211.072.64 and 8 point sources per configuration on the other two ensembles cC211.06.80 and cD211.054.96 in Table I.

For the vector-current disconnected diagrams "V-disconnected I/II" we use quark loops generated with hierarchical probing [44] in general, and low-mode deflation [45] in particular on ensemble cB211.072.64. Details on the dilution scheme for the Hadamard vectors and the number of low modes are given in Ref. [46] and in Table 2 of Ref. [47]. The associated two-point correlators built from the charged pion and iso-vector vector current (cf. Fig. 2) are evaluated with stochastic timeslice propagators based on $\mathbb{Z}_2 \times i\mathbb{Z}_2$ noise sources. The number of stochastic samples is chosen equal to the time extent in lattice units, T/a.

D. Kinematics

The pion momentum $\vec{p}$ is set through the pseudoscalar interpolating operator while the momentum $\vec{q}_1$ is set through the current j_μ, as shown in Eq. (6), with both momenta selected from the available finite-volume momenta $\vec{q}_1, \vec{p} = 2\pi/L \cdot \vec{n}$, where $\vec{n} \in \mathbb{Z}^3$. Momentum conservation sets the momentum $\vec{q}_2 = \vec{p} - \vec{q}_1$. The energy E_π of the pion is then imposed by the on-shell condition, while the temporal component ω_1 of q_1 can be continuously varied as indicated in Eq. (5), with the temporal component ω_2 of q_2 fixed by energy conservation to $\omega_2 = E_\pi - \omega_1$. The kinematical range accessible on the lattice can be parametrized by the accessible virtualities of the electromagnetic currents,

$$q_1^2 = \omega_1^2 - \vec{q}_1^2,$$
$$q_2^2 = (E_\pi - \omega_1)^2 - (\vec{p} - \vec{q}_1)^2. \qquad (21)$$

For the Wick rotation to the Euclidean metric to be valid, the virtualities must fall below the two-pion threshold, i.e., $q_i^2 < 4m_\pi^2$.

For a pion at rest, $\vec{p} = 0$, it holds that $E_\pi = m_\pi$ and $\vec{q}_2 = -\vec{q}_1$. In this case, the photon virtualities simplify to

$$q_1^2 = \omega_1^2 - \vec{q}_1^2,$$
$$q_2^2 = (m_\pi - \omega_1)^2 - \vec{q}_1^2. \qquad (22)$$

The analyticity constraint becomes a constraint on ω_1,

$$\sqrt{4m_\pi^2 + \vec{q}_1^2} + m_\pi < \omega_1 < \sqrt{4m_\pi^2 + \vec{q}_1^2}. \qquad (23)$$

Figure 3 depicts the kinematical range for 26 different choices of $|\vec{q}_1|^2$ satisfying $(2\pi/L)^2 \leq |\vec{q}_1|^2 \leq 32(2\pi/L)^2$ for each of the finite volumes accessible on the three distinct ensembles given in Table I. For all three ensembles, the maximum momentum used in the evaluation of the three-point correlation function gives access to virtualities up to $|q_{1,2}^2| \approx 1.7~\text{GeV}^2$.

Using Eqs. (1), (5) and (22) it is straightforward to show that in the rest frame of the pion $\tilde{A}_{\mu\nu}(\tau)$ vanishes when one or both Lorentz indices are temporal, while the spatial components can be written as

$$\tilde{A}_{ij}(\tau) = -im_\pi \epsilon_{ijk} q_1^k \tilde{A}(\tau). \qquad (24)$$

Inverting the relation gives

$$\tilde{A}(\tau) = i\epsilon_{ijk} \frac{q_1^k}{m_\pi \vec{q}_1^2} \tilde{A}_{ij}(\tau), \qquad (25)$$

where $\tilde{A}(\tau)$ is a scalar under the spatial rotation group. Combining Eqs. (1), (5) and (25) shows that the form factor can be extracted from a Laplace transform of this scalar quantity as

$$\mathcal{F}_{\pi^0\gamma\gamma}(q_1^2, q_2^2) = \int_{-\infty}^\infty d\tau \tilde{A}(\tau) e^{\omega_1\tau}. \qquad (26)$$

Since the full TFF can be extracted from the scalar $\tilde{A}(\tau)$ in the rest frame of the pion, we focus on the evaluation of this scalar function for the remainder of this work.
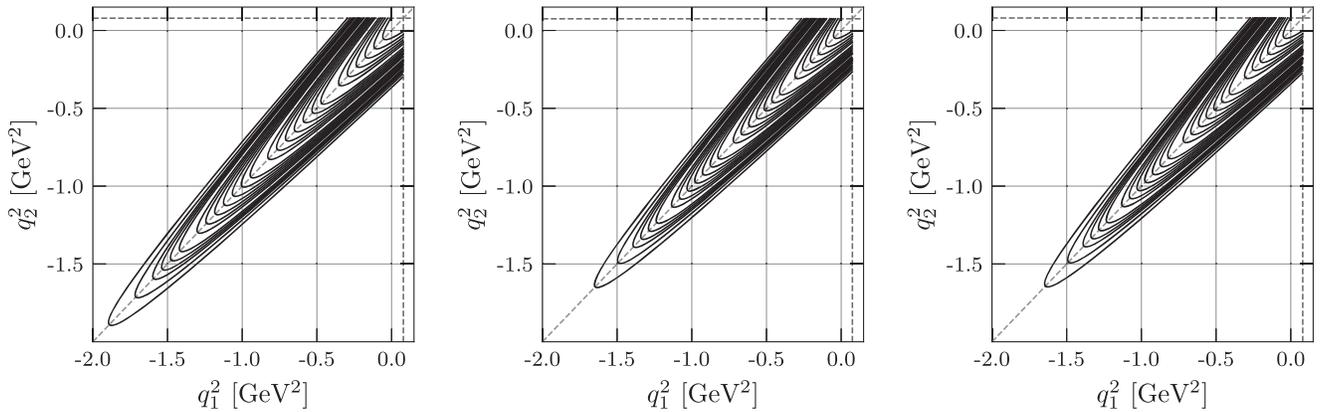
FIG. 3. Range of photon virtualities spanned in our calculation on the ensembles cB211.072.64 (left), cC211.060.80 (middle), and cD211.054.96 (right). Dotted lines at $q_{1,2}^2 = 4m_\pi^2$ indicate the nonanalytic threshold beyond which the rotation to Euclidean spacetime is no longer valid.

The scalar amplitude $\tilde{A}(\tau)$ is implicitly parametrized by the momentum $\vec{q}_1$. All choices of momenta with a fixed value of $|\vec{q}_1|^2$ can be used to evaluate the form factor as in Eq. (26) at identical kinematics. We therefore choose to average the scalar amplitude over such equivalent momenta to increase statistics.

E. Finite-time extent corrections

All ensembles in this work use periodic (anti-periodic) temporal boundaries for the fundamental bosons (fermions). As a result, the pion experiences a periodic boundary condition in time, and the three-point function $C_{\mu\nu}$ receives corrections from states propagating backwards around the finite-time extent of our lattices. Denoting the lattice time extent T, including the leading correction gives

$$
\begin{aligned}
C_{\mu\nu}(\tau, t_\pi) &\approx e^{-E_\pi t_\pi} \frac{Z_\pi}{2E_\pi} \langle 0|T\{j_\mu(\tau)j_\nu(0)\}|\pi^0(\vec{p})\rangle \\
&+ e^{-E_\pi(T-t_\pi)} \frac{Z_\pi}{2E_\pi} \langle \pi^0(\vec{p})|T\{j_\mu(\tau)j_\nu(0)\}|0\rangle \\
&+ O(e^{-\Delta E t_{\text{sep}}} + e^{-\Delta E(T-|\tau|-t_{\text{sep}})}).
\end{aligned}
\tag{27}
$$

In the rest frame of the pion, we can apply Bose symmetry and PT symmetry to define a corrected scalar amplitude $\tilde{A}(\tau)$ [cf. Eq. (25)] to account for this leading-order effect,

$$
\begin{aligned}
\tilde{A}(\tau; t_\pi) &\equiv i\epsilon_{ijk} \frac{q_1^k}{m_\pi \vec{q}_1^2} [1 - e^{E_\pi(T-|\tau|-t_\pi)}]^{-1} \tilde{A}_{ij}(\tau; t_\pi) \\
&= \tilde{A}(\tau) + O(e^{-\Delta E t_{\text{sep}}} + e^{-\Delta E(T-|\tau|-t_{\text{sep}})}).
\end{aligned}
\tag{28}
$$

In all further analysis, this corrected form of the lattice amplitude is used as input.

F. Dependence on t_π

An example of the scalar amplitude $\tilde{A}(\tau)$ with the decomposition into contributions from individual Wick contractions is given in Fig. 4. These results, along with those used in the following analysis are based on the choice of interpolating operator time $t_\pi = 2.23$ fm for ensemble cB211.072.64, $t_\pi = 2.18$ fm for ensemble cC211.060.80, and $t_\pi = 2.28$ fm for ensemble cD211.054.96, with the choices made such that t_π in physical units is similar on all three ensembles. Note that smaller values of t_π lead to a reduction of the statistical errors, however, this requires a careful analysis of excited state effects. To confirm that our choices do not suffer from significant excited state effects, two larger choices of t_π were measured on cB211.072.64 and cC211.060.80 and one larger choice on cD211.054.96. We depict an example of the comparison of $\tilde{A}(\tau; t_\pi)$ across different choices of t_π in Fig. 5. In all cases, $\tilde{A}(\tau; t_\pi)$ does not show any systematic dependence on t_π across the choices.

G. Tail fits

The extraction of the transition form factor using Eq. (5) in principle requires access to $\tilde{A}(\tau)$ for arbitrary $\tau \in (-\infty, \infty)$. To estimate the amplitude over the whole temporal range from our lattice data with finite temporal extent, we use two phenomenological models, the vector-meson-dominance model (VMD) and lowest-meson-dominance model (LMD) [17,48,49]. The VMD and LMD form factors are given, respectively, by

$$
\mathcal{F}_{\pi^0\gamma\gamma}^{\text{VMD}}(q_1^2, q_2^2) = \frac{\alpha M_V^4}{(M_V^2 - q_1^2)(M_V^2 - q_2^2)}
\tag{29}
$$

and
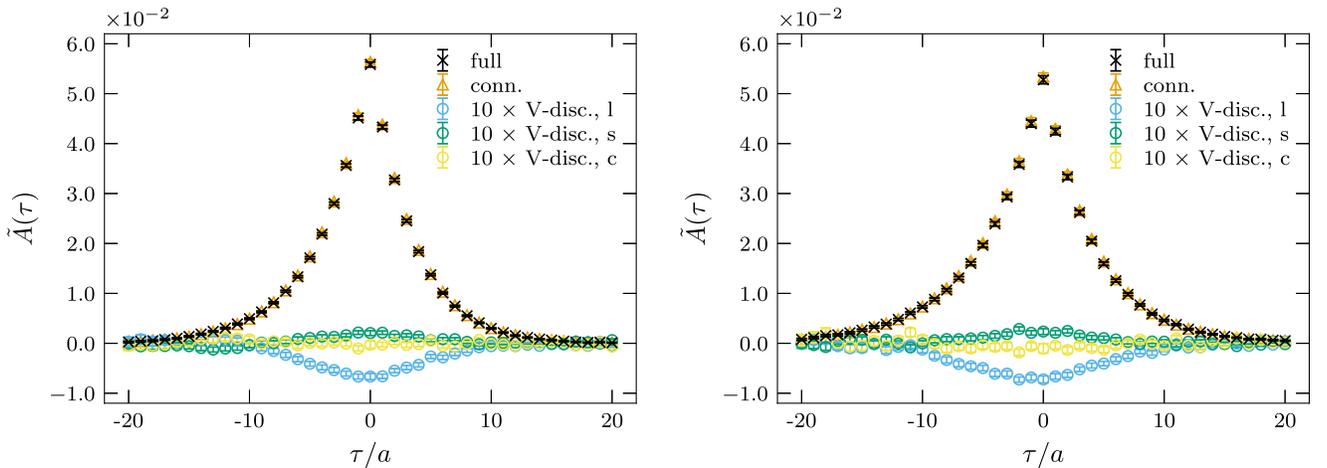
FIG. 4. $\tilde{A}(\tau)$ with $|\vec{q}_1|^2 = 10(2\pi/L)^2$ including finite-time extent corrections on the ensembles cC211.060.80 (left) and cD211.054.96 (right). We show the full $\tilde{A}(\tau)$ as black crosses, the connected contribution as orange triangles, and the V-disconnected contributions (multiplied by 10 for easier comparison) with a light, strange and charm vector-current loop as blue, green and yellow circles, respectively.
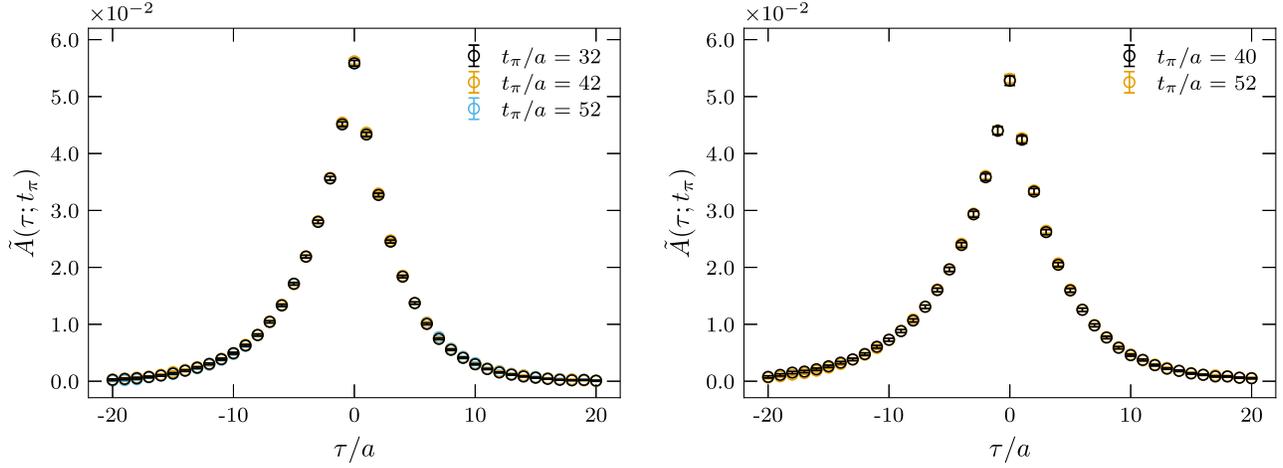
FIG. 5. $\tilde{A}(\tau; t_\pi)$ with $|\vec{q}_1|^2 = 10(2\pi/L)^2$ including finite-time extent corrections on the ensembles cC211.060.80 (left) and cD211.054.96 (right) for the choices $t_\pi \in \{2.18, 2.86, 3.55\}$ fm (left) and $t_\pi \in \{2.28, 3.02\}$ fm (right).

$$\mathcal{F}^{\text{VMD}}_{\pi^0\gamma\gamma}(q_1^2, q_2^2) = \frac{\alpha M_V^4 + \beta(q_1^2 + q_2^2)}{(M_V^2 - q_1^2)(M_V^2 - q_2^2)}. \qquad (30)$$

The phenomenological origin of the models suggests particular parameter choices. In the chiral limit and at low energy, the fermionic triangle diagram gives rise to the Adler-Bell-Jackiw anomaly [50,51], which constrains the form factor in these limits to

$$\mathcal{F}_{\pi^0\gamma\gamma}(0, 0) = \frac{1}{4\pi^2 F_\pi}, \qquad (31)$$

with $F_\pi = 92.3(1)$ MeV the pion decay constant [52], suggesting $\alpha = 1/(4\pi^2 F_\pi) = 0.274$ GeV^{-1}. Meanwhile, the mass $M_V = 775$ MeV is the ρ-meson mass in the VMD and LMD models, and $\beta = -F_\pi/3 = -0.0308$ GeV reproduces the leading operator product expansion (OPE) prediction [27,48,49]. However, we treat α, M_V and β as free model parameters when fitting the models to our data, using the models as inspiration for the fit form rather than a prediction of the precise asymptotic behavior.

Using the relation between the TFF and amplitude in Eq. (1) and inverting the relation in Eq. (5), the VMD and LMD form factors can be converted into fit functions for the Euclidean-time amplitude $\tilde{A}$ as follows. For the LMD fit function one finds

$$\tilde{A}^{\text{LMD}}(\tau) = e^{m_\pi|\tau|\Theta(-\tau)}[C_+ e^{-E_V|\tau|} - C_- e^{-(m_\pi + E_V)|\tau|}], \qquad (32)$$

in terms of

$$C_\pm \equiv \frac{\alpha M_V^4 + \beta(2M_V^2 + m_\pi^2 \mp 2m_\pi E_V)}{m_\pi E_V(2E_V \mp m_\pi)},$$

$$E_V \equiv \sqrt{M_V^2 + |\vec{q}_1|^2}. \qquad (33)$$

The VMD fit function is obtained by setting $\beta = 0$.

The VMD and LMD models should not be expected to fully capture the large-Q^2 behavior of the form factor, or conversely the small-$|\tau|$ behavior of $\tilde{A}(\tau)$. In particular, the Brodsky-Lepage behavior constrains the single-virtual form factor at large Euclidean (spacelike) momentum [53–55], giving

$$\mathcal{F}_{\pi^0\gamma\gamma}(-Q^2, 0) \overset{Q^2 \to \infty}{\to} \frac{2F_\pi}{Q^2} \qquad (34)$$

at leading order in α_s. This behavior is reproduced by the VMD model, but the LMD model tends to a constant at large Euclidean momenta in the single-virtual form factor. Meanwhile, the OPE at short distances [56,57] restricts the double-virtual form factor where both momenta become large at the same time, giving in the chiral limit

$$\mathcal{F}_{\pi^0\gamma\gamma}(-Q^2, -Q^2) \overset{Q^2 \to \infty}{\to} \frac{2F_\pi}{3}\left[\frac{1}{Q^2} + \mathcal{O}\left(\frac{1}{Q^4}\right)\right]. \qquad (35)$$

This behavior is reproduced by the LMD model (with the choice of β given above), but the VMD model in this case falls off as $1/Q^4$ [27]. Despite these shortcomings at large Q^2, both models accurately capture the long-distance behavior. As such, we use the lattice data directly for the short-distance region while replacing only the tails with these fits. Details of this procedure are discussed in the following section.

III. TRANSITION FORM FACTOR

To obtain the transition form factors, we need to evaluate the Laplace transform in Eq. (26), reproduced here for convenience,

$$\mathcal{F}_{\pi^0\gamma\gamma}(q_1^2, q_2^2) = \int_{-\infty}^{\infty} d\tau\, \tilde{A}(\tau) e^{\omega_1 \tau}, \qquad (36)$$

where for a given momentum orbit $|\vec{q}_1|^2$ the squared four-momenta q_1^2 and q_2^2 are given by Eq. (22). The TFF is only directly available on specific sets of orbits $q_{1,2}^2$ determined by the kinematics, and it must therefore be interpolated/extrapolated to other points in the (q_1^2, q_2^2) plane. In the following we describe our strategy for the numerical integration in Eq. (36) and for the extension of the TFF lattice data to the whole momentum plane using the modified z-expansion.

A. Numerical integration

The tails of $\tilde{A}(\tau)$ associated with large $|\tau|$ are expected to decrease exponentially quickly, with a mass scale predicted by the VMD/LMD models to be $E_V = \sqrt{M_V^2 + |\vec{q}_1|^2}$ in one tail and $E_V + m_\pi$ in the other [see Eq. (32)]. However, the exponential factor $e^{\omega_1 \tau}$ enhances one tail of $\tilde{A}(\tau)$ significantly whenever $|\omega_1|$ is sufficiently large compared to this energy scale. This makes the numerical integration difficult: Since the signal-to-noise ratio decreases when going to larger $|\tau|$, fluctuations in one of the tails due to the small signal-to-noise ratio are exponentially enhanced. We observe that this effect is small for virtualities near the diagonal $q_1^2 = q_2^2$, for which $\omega_1 = m_\pi / 2 \ll E_V$, but can be severe in the single-virtual cases corresponding to either $q_1^2 = 0$ or $q_2^2 = 0$, for which $\omega_1 = |\vec{q}_1| \approx E_V$, especially for the larger momentum orbits. Figure 6 shows a representative comparison of the integrand of the Laplace transform for both the diagonal and single-virtual cases. For choices of virtualities between these two cases, the exponential enhancement gets more and more pronounced when going from the diagonal to the single-virtual case.

These issues are addressed by the introduction of cutoff times τ_{cut}^R and τ_{cut}^L, with lattice data replaced by a global fit to $\tilde{A}(\tau)$ for $\tau \geq \tau_{\text{cut}}^R$ and $\tau \leq \tau_{\text{cut}}^L$. In our analysis, we perform a fit to Eq. (32) across data at all available $\vec{q}_1$ and for values of τ selected from symmetrical fit ranges on both sides of the peak, i.e., $\tau \in [-\tau_{\max}, -\tau_{\min}] \cup [\tau_{\min}, \tau_{\max}]$. Variation of the fit ranges allows us to determine a corresponding systematic error. The left and right cutoff times are selected depending on the sign of ω_1 to include as much data as possible on the exponentially suppressed tail of the integrand, in particular by fixing $\tau_{\text{cut}}^L = -T/2$ for $\omega_1 \geq 0$ or $\tau_{\text{cut}}^R = T/2$ for $\omega_1 < 0$. The choice of the cutoff time τ_{cut} is varied to assess the systematic error associated with this choice.

Finally, we filter the choices of cutoff times and kinematics to demand that for a given momentum orbit and value of ω_1 a minimum percentage of $\mathcal{F}_{\pi^0 \gamma\gamma}(q_1^2, q_2^2)$ must come from the lattice data. This is done to minimize the introduction of a model dependence in the final result. To do so, we define the *data content* for a given momentum and value of ω_1 to be

$$\Delta_{\text{latt}} = \frac{\int_{\tau_{\text{cut}}^L}^{\tau_{\text{cut}}^R} d\tau \, \tilde{A}^{(\text{latt})}(\tau) e^{\omega_1 \tau}}{\mathcal{F}_{\pi^0 \gamma\gamma}(q_1^2, q_2^2)}. \tag{37}$$

In this work, we always restrict kinematics to only include TFFs for which $\Delta_{\text{latt}} \geq 95\%$.

B. Sampling in the momentum plane

As illustrated in Fig. 3, the transition form factor obtained from the lattice is a continuous function of ω_1

FIG. 6. Integrands $\tilde{A}(\tau) e^{\omega_1 \tau}$ with $|\vec{q}_1|^2 = 6(2\pi/L)^2$ on the ensemble cB211.072.64, comparing diagonal kinematics $q_1^2 = q_2^2$ (left) and single-virtual kinematics $q_2^2 = 0$ (right). We show the integrand from lattice data as black circles and from a correlated LMD fit to $\tilde{A}(\tau)$ according to Eq. (32) as an orange line. The fit was done globally including all momentum orbits $(2\pi/L)^2 \leq |\vec{q}_1|^2 \leq 32(2\pi/L)^2$ with fit range $[-10, -8] \cup [8, 10]$, indicated by the gray bands, yielding $\chi^2/\text{d.o.f.} = 1.01$ with correlations taken into account across all orbits and fit range. The red data point indicates the time slice where the pseudoscalar operator is inserted, while the grayed-out points are those where the time ordering does not satisfy the one from Eq. (4).

for each spatial momentum orbit $|\vec{q}_1|^2$. In order to interpolate/extrapolate to the rest of the (q_1^2, q_2^2) plane, we first sample choices of ω_1 on which to evaluate the TFF as the input points for the fit. In this work, we do so by selecting ω_1 corresponding to fixed choices of the ratio q_2^2/q_1^2, i.e., by finding the intersection between the available orbits and several diagonal "cuts" through the (q_1^2, q_2^2) plane. Because the underlying lattice data are the same for all choices of ω_1 within each orbit, data at nearby values of ω_1 are strongly correlated. A relatively sparse sampling is therefore possible without sacrificing useful inputs to the z-expansion fits. Both $q_2^2/q_1^2 = 0$ and $q_2^2/q_1^2 = 1$ are useful choices of ratios to include, the former since the single-virtual transition form factor plays an important role in the evaluation of $a_\mu^{\pi^0-\text{pole}}$, and the latter since we get the best signal-to-noise ratio in the transition form factor there.

We use the following procedure for the sampling in the momentum plane. We determine five values of ω_1 on the highest-momentum orbit $|\vec{q}_1|^2 = 32(2\pi/L)^2$, including those corresponding to the diagonal and single-virtual kinematics, such that the arc length of the curve parametrized by ω_1 between neighboring samples is constant. This fixes five ratios q_2^2/q_1^2 which are then used to determine the ω_1 on all other spatial momentum orbits. To better cover the region close to single-virtual kinematics, we additionally include a cut $q_2^2/q_1^2 = 0.1$, as depicted in Fig. 7. We use several subsets of these cuts in the further analysis in order to determine systematic uncertainties associated with the sampling: all combinations of three
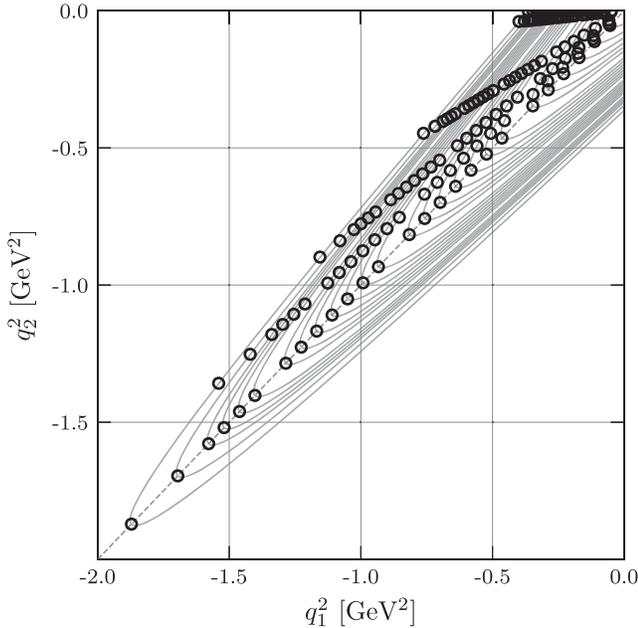
FIG. 7. Cuts $(q_2^2/q_1^2) \in \{1.0, 0.88, 0.78, 0.59, 0.1, 0.0\}$ on ensemble cB211.072.64 illustrating our sampling of $\mathcal{F}_{\pi^0\gamma\gamma}(q_1^2, q_2^2)$ in the momentum plane used as input for the fit of the modified z-expansion.

cuts including both $q_2^2/q_1^2 \in \{1, 0\}$ and one of $q_2^2/q_1^2 \in \{0.88, 0.78, 0.59, 0.1\}$, and in addition a subset containing the four cuts $q_2^2/q_1^2 \in \{1.0, 0.88, 0.1, 0.0\}$. Including more than four cuts usually leads to an ill-conditioned covariance matrix yielding nonconverging or unstable fits.

For each cut, we impose a threshold for a minimal data content $\Delta_{\text{latt}} \geq 95\%$ (see Sec. III A) in the TFF evaluated at each sampled kinematical point, dropping those from further analysis that do not meet the threshold.

C. z-expansion

We use the modified z-expansion proposed in [28] to interpolate/extrapolate the data from the sampling as described above to the whole kinematical range. This is a model-independent way of extending the transition form factor to arbitrary photon momenta which is preconditioned to more easily reproduce the form factor structure. Following Ref. [28], the model-independent fit form is constructed by first defining the conformal variables z_1 and z_2 as [58]

$$z_k = \frac{\sqrt{t_c + Q_k^2} - \sqrt{t_c - t_0}}{\sqrt{t_c + Q_k^2} + \sqrt{t_c - t_0}}, \qquad k \in \{1, 2\}, \quad (38)$$

where $Q_k^2 = -q_k^2$ and $t_c = 4m_\pi^2$ indicates the position of the branch cut due to the two-pion threshold. The parameter t_0 can be freely tuned to optimize the rate of convergence. For a given choice of Q_{max}^2, to best map the region $0 \leq Q_{1,2}^2 \leq Q_{\text{max}}^2$ to conformal variables near the origin, the optimal choice of t_0 is

$$t_0 = t_c \left(1 - \sqrt{1 + Q_{\text{max}}^2/t_c}\right), \quad (39)$$

which minimizes the maximum value of $|z_k|$ in the range $[0, Q_{\text{max}}^2]$. We use $Q_{\text{max}}^2 = 4 \text{ GeV}^2$ in the present study, as this is the region yielding the dominant contribution to the integral for a_μ (see Sec. IV A).

In terms of these conformal variables, one can expand

$$\mathcal{F}_{\pi^0\gamma\gamma}(-Q_1^2, -Q_2^2) = \sum_{n,m=0}^{\infty} c_{nm} z_1^n z_2^m, \quad (40)$$

which is a convergent expansion within the analytic domain $|z_k| < 1$. The coefficients $c_{nm} = c_{mn}$ are symmetric due to the Bose symmetry. By restricting the sum to $m, n \leq N$, one can approximate the form factor with accuracy determined by the choice of maximum order N. In addition, the transition form factor can be multiplied by an arbitrary analytic function $P(Q_1^2, Q_2^2)$ before expanding the resulting product in powers of z_k as above. This preconditioning allows the expansion to more easily reproduce the form factor structure. As shown in Ref. [28], the choice

$$P(Q_1^2, Q_2^2) = 1 + \frac{Q_1^2 + Q_2^2}{M_V^2}, \qquad (41)$$

where $M_V \approx 775$ MeV is the ρ-meson mass, leads to a parametrization of the transition form factor which decreases asymptotically as $1/Q^2$ in all directions in the momentum plane even at finite values of N, thus satisfying the momentum dependence of Eqs. (34) and (35). As shown in Ref. [59], to enforce the appropriate scaling at the two-pion threshold, the expansion should be further modified to fix the derivatives of the transition form factors with respect to z_k at $z_k = -1$ to zero. Including the cutoff N, preconditioning, and threshold constraints gives the modified expansion [28]

$$
\begin{aligned}
&P(Q_1^2, Q_2^2) \mathcal{F}_{\pi^0 \gamma\gamma}(-Q_1^2, -Q_2^2) \\
&\approx \sum_{m,n=0}^{N} c_{nm} \left(z_1^n - (-1)^{N+n+1} \frac{n}{N+1} z_1^{N+1} \right) \\
&\quad \times \left(z_2^m - (-1)^{N+m+1} \frac{m}{N+1} z_2^{N+1} \right).
\end{aligned}
\qquad (42)
$$

In the following, we either fit the coefficients c_{nm} independently to the data for each ensemble or perform a combined fit to all three ensembles with correction terms for $\mathcal{O}(a^2)$ lattice artifacts as

$$c_{nm}(a) = c_{nm}(0) + \delta_{nm} \left(\frac{a}{a_{\mathrm{ref}}} \right)^2, \qquad (43)$$

where we use the cC211.060.80 lattice spacing for a_{ref}, i.e., $a_{\mathrm{ref}} = 0.06821(12)$ fm. The continuum-extrapolation strategies based on these two options are considered below in Sec. VI.

In contrast to Refs. [28,29] we perform fully correlated fits with the modified z-expansion. We use $N = 1, 2$ and all possible combinations of the $N = 1$ coefficients with one of the $N = 2$ coefficients. We find that the resulting fits describe the lattice data well, even along the cuts not included in the fit, as seen in the example in Figs. 8 and 9.

Note that the transition form factors obtained from Eq. (42) multiplied by Q^2 asymptote to a constant by construction, with predictions for the single-virtual and diagonal case described in Eqs. (34) and (35), respectively. We exclude analyses in which the single-virtual transition form factor multiplied by Q^2 turn out to be asymptotically negative, since this clearly violates those predictions.

We find that correlated z-expansion fits with cutoffs set to either $N = 1$ or $N = 2$ already lead to a good quality of fit when including input data from three or four cuts in the momentum plane as described in Sec. III B. For $N = 2$, at least two of the coefficients c_{nm} are strongly correlated, indicating a degeneracy in the parametrization. Even when including only one of the second-order coefficients in
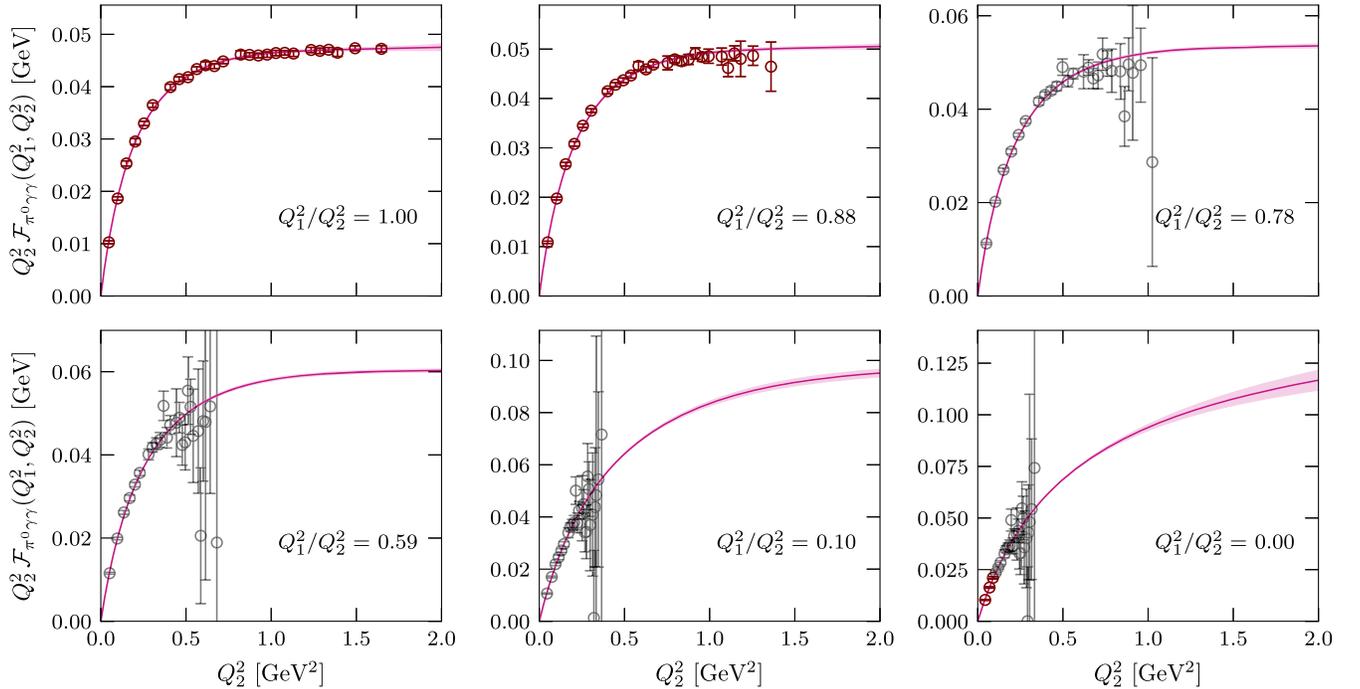
FIG. 8. Correlated z-expansion fit to the TFFs evaluated on ensemble cC211.060.80 using c_{00}, c_{10}, c_{11} and c_{22}. Shown is $Q_2^2 \mathcal{F}_{\pi^0 \gamma\gamma}(Q_1^2, Q_2^2)$, plots showing $\mathcal{F}_{\pi^0 \gamma\gamma}(Q_1^2, Q_2^2)$ can be found in Fig. 9. Only TFF data points with at least 95% data content are included (colored in red for the used cuts $(q_2^2/q_1^2) \in \{1.0, 0.88, 0.0\}$); data in gray with less than 95% data content are shown for illustration only. For $\tilde{A}(\tau)$ a global LMD fit with fit range [20, 21] and $\tau_{\mathrm{cut}} = 23$ in lattice units is used. The reduced χ^2 are $\chi_{\tilde{A}}^2/\mathrm{d.o.f.} = 0.86$ and $\chi_{z-\mathrm{exp}}^2/\mathrm{d.o.f.} = 1.00$.
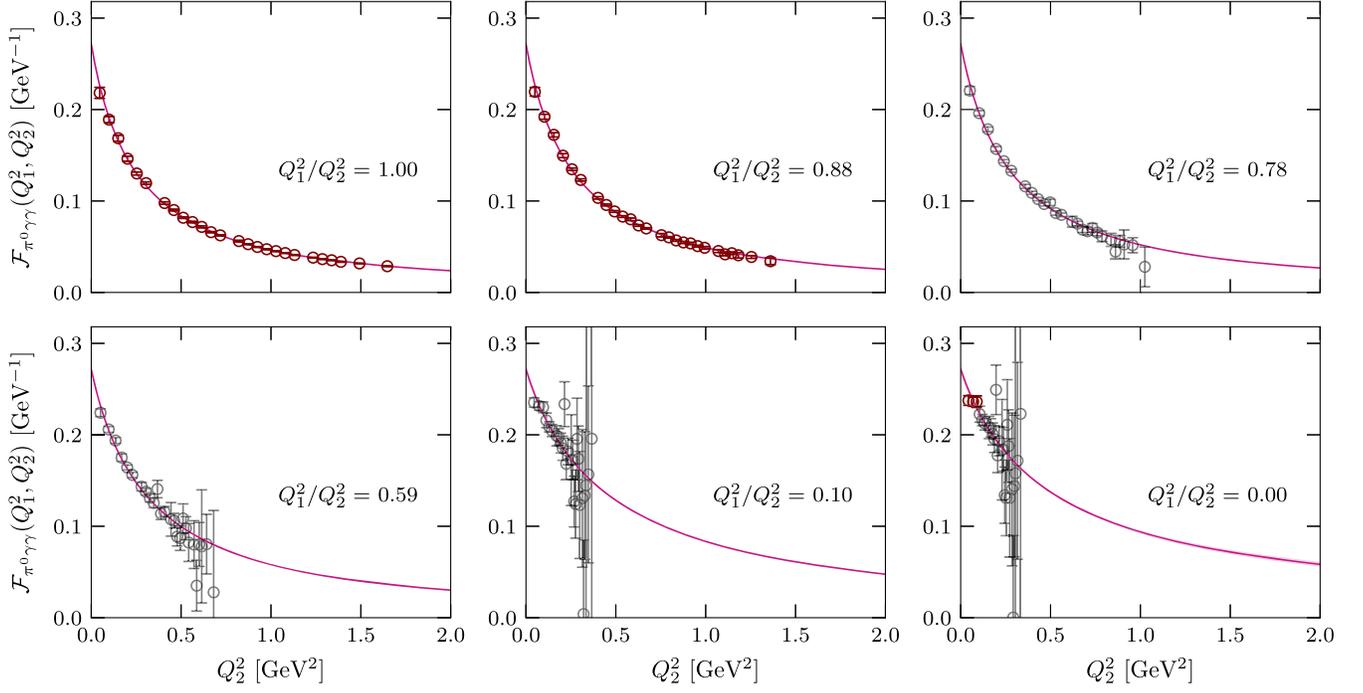
FIG. 9. Same as Fig. 8, but showing $\mathcal{F}_{\pi^0\gamma\gamma}(Q_1^2, Q_2^2)$ instead of $Q_2^2\mathcal{F}_{\pi^0\gamma\gamma}(Q_1^2, Q_2^2)$ in order to highlight the quality of fit at low values of Q_2^2.

conjunction with all three first-order coefficients, we find slight correlations between some of them. See Fig. 10 for a corner plot depicting correlations in a representative case when including coefficients $\{c_{00}, c_{10}, c_{11}, c_{22}\}$.

The corresponding correlation matrix is given by

$$\text{cor}(c_{nm}) = \begin{pmatrix} +1.00 & -0.30 & -0.39 & +0.31 \\ -0.30 & +1.00 & -0.05 & -0.14 \\ -0.39 & -0.05 & +1.00 & -0.88 \\ +0.31 & -0.14 & -0.88 & +1.00 \end{pmatrix}, \quad (44)$$

clearly showing the strong (anti)correlation between c_{11} and c_{22}, and lesser (anti)correlations between all other coefficients except for c_{10} and c_{11} which are almost completely uncorrelated. When including all $N = 2$ coefficients, it turns out that c_{20} and c_{21} are strongly correlated with c_{11} for almost all choices of parameters. To get reliable estimates of the individual coefficients, it is desirable to remove such near degeneracies in the parametrization by only using a subset of the coefficients which are not strongly correlated. In the final analysis, we therefore determine the preferred values of all quantities studied based on a z-expansion parametrization with minimal degeneracy, but we use variations between these choices of parametrizations to determine an additional systematic error.

IV. OBSERVABLES

The main observable studied in this work is $a_\mu^{\pi^0-\text{pole}}$, i.e., the pion-pole contribution to the muon $(g-2)/2$. We further study the two-photon decay width $\Gamma(\pi \to \gamma\gamma)$ and the slope parameter b_π, both of which are related to $\mathcal{F}_{\pi^0\gamma\gamma}$.

A. Pion-pole contribution to a_μ

Following Refs. [17] and [18], we use the three-dimensional integral representation for the pion-pole contribution

$$a_\mu^{\pi^0-\text{pole}} = \left(\frac{\alpha}{\pi}\right)^3 \left[a_\mu^{\pi^0-\text{pole}(1)} + a_\mu^{\pi^0-\text{pole}(2)}\right], \quad (45)$$

where α is the fine-structure constant,

$$a_\mu^{\pi^0-\text{pole}(1)} = \int_0^\infty dQ_1 \int_0^\infty dQ_2 \int_{-1}^1 d\tau \\ \times w_1(Q_1, Q_2, \tau)\mathcal{F}_{\pi^0\gamma\gamma}(-Q_1^2, -Q_3^2)\mathcal{F}_{\pi^0\gamma\gamma}(-Q_2^2, 0), \quad (46)$$

and

$$a_\mu^{\pi^0-\text{pole}(2)} = \int_0^\infty dQ_1 \int_0^\infty dQ_2 \int_{-1}^1 d\tau \\ \times w_2(Q_1, Q_2, \tau)\mathcal{F}_{\pi^0\gamma\gamma}(-Q_1^2, -Q_2^2)\mathcal{F}_{\pi^0\gamma\gamma}(-Q_3^2, 0). \quad (47)$$
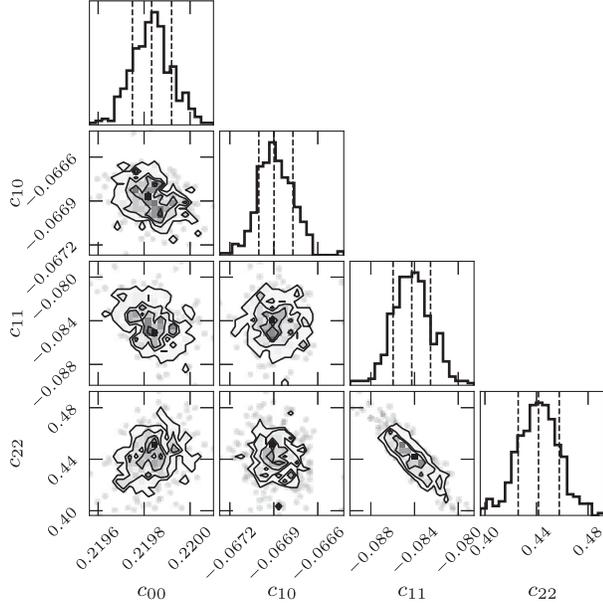
FIG. 10. Corner plot for the coefficients c_{00}, c_{10}, c_{11} and c_{22} from the correlated z-expansion fit to the TFFs on cC211.060.80 as shown in Figs. 8 and 9 from a jackknife resampling. The corresponding correlation matrix is given in Eq. (44).

The integration is performed over the magnitudes Q_1, Q_2 of two of the four-momenta and $\tau = \cos\theta$ describing the angle θ between them, with the magnitude of the third four-momentum fixed by $Q_3^2 = Q_1^2 + Q_2^2 + 2Q_1Q_2\tau$. The weight functions w_1 and w_2 are given in Appendix A. Note that $w_{1,2}$ are both dimensionless and $w_{1,2}(Q_1, Q_2, \tau) \to 0$ for $Q_{1,2} \to 0$ and for $\tau \to \pm 1$. Further, w_2 is symmetric under the exchange of Q_1 and Q_2. In Ref. [17], the weight functions for the pion are studied and discussed in detail. One finds that the momentum region $Q_{1,2} \leq 0.5$ GeV is the

most important in Eqs. (46) and (47) for $a_\mu^{\pi^0-\text{pole}}$, which is where we have the strongest data support from our lattice calculation. For two examples illustrating the concentration of the weight functions at low momenta, see Fig. 11. In particular, note that $w_2(Q_1, Q_2, \tau)$ is roughly an order of magnitude smaller than $w_1(Q_1, Q_2, \tau)$.

By introducing a momentum cutoff Λ, i.e., by replacing

$$\int_0^\infty dQ_{1,2} \to \int_0^\Lambda dQ_{1,2} \tag{48}$$

in Eqs. (46) and (47), one can estimate the importance of various momentum regions in the evaluation of Eqs. (46) and (47). Though the cutoffs are only directly imposed on $Q_{1,2}$, they imply a cutoff of $Q_3 \leq 2\Lambda$ on the third momentum as well. We find that the integrals saturate rapidly, such that we get around 85% of the total result with $\Lambda = 1$ GeV and around 90% with $\Lambda = 1.5$ GeV, i.e., from momentum regions with strong data support from our lattice calculation. An example of this exercise is shown in Fig. 12.

B. Decay width and slope parameter

To leading order in the fine structure constant α, the transition form factors determine the partial decay width through

$$\Gamma(\pi \to \gamma\gamma) = \frac{\pi\alpha^2 m_\pi^3}{4}|\mathcal{F}_{\pi^0\gamma\gamma}(0,0)|^2 + O(\alpha^3). \tag{49}$$

The neutral-pion decay width has been measured in the PrimEx and PrimEx-II experiments [60,61] with a combined result of $\Gamma(\pi \to \gamma\gamma) = 7.802(52)_{\text{stat}}(105)_{\text{sys}}(117)_{\text{tot}}$ eV. At our level of statistical precision,
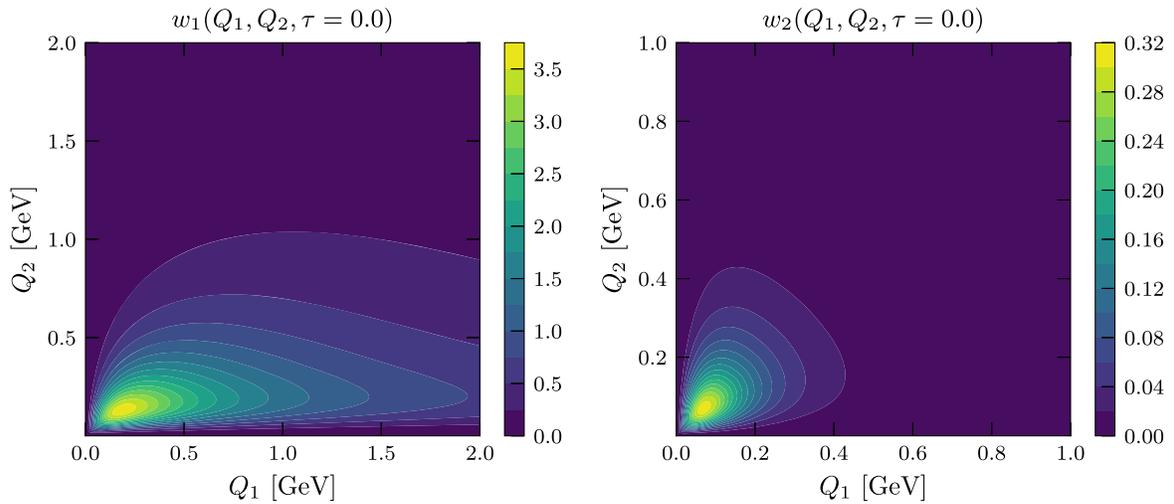
FIG. 11. Weight function $w_1(Q_1, Q_2, \tau)$ (left) and $w_2(Q_1, Q_2, \tau)$ (right) for the pion for momenta Q_1 and Q_2 and $\tau = \cos\theta$ where $\theta = 90°$, showing the concentration at low momenta. Note the different momentum range for and magnitude of w_2 as well as its symmetry with respect to $Q_1 \leftrightarrow Q_2$.
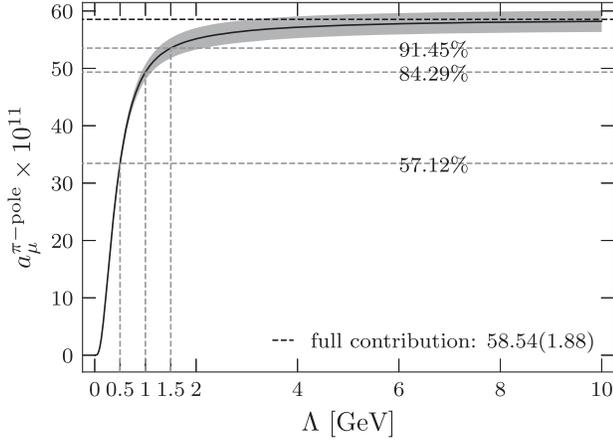
FIG. 12. Contribution to $a_\mu^{\pi^0-\text{pole}}$ for increasing cutoff Λ for ensemble cC211.060.80, using a $N = 2$ correlated z-expansion fit to $(q_2^2/q_1^2) \in \{1.0, 0.88, 0.1, 0.0\}$. Only TFF data points where at least 95% of the transition form factors come from lattice data are included, for $\tilde{A}(\tau)$ a global LMD fit with fit range [7, 8] and $\tau_{\text{cut}} = 26$ in lattice units is used. The reduced χ^2 are $\chi_{\tilde{A}}^2/\text{d.o.f.} = 0.85$ and $\chi_{z-\text{exp}}^2/\text{d.o.f.} = 1.02$. Indicated is the saturation at $\Lambda \in \{0.5, 1.0, 1.5\}$ GeV as well as the value of the full contribution for this particular choice of parameters.

uncertainties from radiative corrections to the decay width are negligible and we therefore quote the leading-order result as our estimate of this quantity.

Further, the transition form factors can be used to extract the slope parameter

$$b_\pi = \frac{1}{\mathcal{F}_{\pi^0\gamma\gamma}(0,0)} \frac{d\mathcal{F}_{\pi^0\gamma\gamma}(q^2,0)}{dq^2}\bigg|_{q^2=0}, \qquad (50)$$

thus providing input for determining the electromagnetic interaction radius of the pion. The averaged experimental result for the slope parameter is $b_\pi = 1.84(17)$ GeV^{-2} [52].

These quantities are easily extracted from the form of the z-expansion fit at $(q_1^2, q_2^2) = (0,0)$. The value of $\mathcal{F}_{\pi^0\gamma\gamma}(0,0)$ comes directly from the z-expansion fit, while the derivative can be acquired by differentiating the form of the z-expansion in Eq. (42) with respect to $-Q_1^2$, yielding the slope parameter

$$b_\pi = -\frac{d}{dQ_1^2} \ln \mathcal{F}_{\pi^0\gamma\gamma}(-Q_1^2, 0)\big|_{Q_1^2=0}$$

$$\approx \frac{1}{M_V^2} - \frac{1}{z_1} \frac{dz_1}{dQ_1^2} \left[\sum_{\substack{m=0 \\ n=1}}^{N} \right.$$

$$\times nc_{nm}(z_1^n - (-1)^{N+n+1}z_1^{N+1})$$

$$\times \left(z_2^m - (-1)^{N+m+1} \frac{m}{N+1} z_2^{N+1} \right) \bigg]_{Q_{1,2}^2=0}. \quad (51)$$

The term in square brackets can be directly evaluated, and the derivative of the conformal parameter is given by

$$\frac{1}{z_1} \frac{dz_1}{dQ_1^2}\bigg|_{Q_1^2=0} = \frac{1}{t_0} \sqrt{1 - t_0/t_c}. \qquad (52)$$

V. MODEL AVERAGING AND ERROR ESTIMATION

We follow different analysis chains from the initial lattice correlation function data to estimates of the quantities $a_\mu^{\pi^0-\text{pole}}$, $\Gamma(\pi \to \gamma\gamma)$ and b_π. Each such chain is determined by a choice of parameters, namely the choice of fit range and fit model in the fit to $\tilde{A}(\tau)$, τ_{cut} when constructing the transition form factors and finally the included cuts in the momentum plane in the fit to the modified z-expansion. This yields $\mathcal{O}(10^3-10^4)$ estimates. For the remainder of this section such a choice of parameters is called an "analysis". *A priori*, only fits with $\chi_{\tilde{A}}^2/\text{d.o.f.}$ close to 1 are included in an analysis chain for the single ensemble analyses, with $0.84 \lesssim \chi_{\tilde{A}}^2/\text{d.o.f.} \lesssim 2.05$. For the combined z-expansion fits the input $\tilde{A}(\tau)$ fit ranges and τ_{cut} were chosen to be similar in physical units across all ensembles, using $t_{\text{min}} \in [0.48, 1.27]$ fm, $t_{\text{max}} \in [0.64, 1.43]$ fm and $\tau_{\text{cut}} \in [1.35, 2.07]$ fm on cB211.072.64 as a reference and remaining within 10% of these physical-units values for the cC211.060.80 and cD211.054.96 ensembles. The fits across all ensembles are of good quality, with χ^2 values in the range $1 \lesssim \chi_{\tilde{A}}^2/\text{d.o.f.} \lesssim 2.4$ with an average of $\bar{\chi}_{\tilde{A}}^2/\text{d.o.f.} \approx 1.5$.

We calculate the statistical errors for each analysis using the jackknife resampling procedure, finding that there is virtually no autocorrelation between the input data $\tilde{A}(\tau)$ amongst the available configurations listed in Table I. Jackknife resampling is applied over the entire analysis pipeline for each set of analysis choices, ensuring that statistical errors are fully propagated.

To take a weighted average of estimates under various analysis choices in an analysis chain, we use a modified version of the Akaike information criterion (AIC) [62,63]. Closely following the method introduced in [64] and [65], the model averaging proceeds as follows. For a target observable y, here $a_\mu^{\pi^0-\text{pole}}$, $\Gamma(\pi \to \gamma\gamma)$ or b_π, we build a histogram from the different analyses, assigning to each analysis a weight given by the AIC. This criterion is derived from the Kullback-Leibler divergence, which measures the distance of a fit function from the true distribution of the points (for a derivation see Ref. [64]). We use the modified AIC introduced in [65],

$$\text{AIC} \sim \exp\left[-\frac{1}{2}(\chi^2 + 2n_{\text{par}} - n_{\text{data}}) \right], \qquad (53)$$

where the χ^2, the number of fit parameters n_{par} and the number of data points n_{data} describe the fit of interest. The first two terms in the exponent correspond to the standard

AIC, and the last term is needed to weigh fits with different lengths in the fit ranges when fitting to $\tilde{A}(\tau)$ or to weight z-expansion fits with a differing number of cuts in the momentum plane when sampling the transition form factors.

In general, two fitting steps are involved in the determination of each of $a_\mu^{\pi^0-\text{pole}}$, $\Gamma(\pi^0 \to \gamma\gamma)$, and b_π on a given ensemble; the global fit to $\tilde{A}(\tau)$ and the z-expansion fit to the transition form factors. A combined weight can be assigned to each combination of fitting choices across both of these steps by multiplying the respective AIC weights $\tilde{w}_i$ given in Eq. (53). In the alternative simultaneous continuum and z-expansion fit procedure described in Sec. III C, the model averaging step is instead performed globally across all three ensembles, and therefore a combined AIC weight is assigned by multiplying the weights of the fits to $\tilde{A}(\tau)$ across all three ensembles as well as the weight of the final global z-expansion fit with lattice artifacts included. In either case, the result is one unnormalized weight $\tilde{w}_i$ per analysis.

Given the weights $\tilde{w}_i$, central values m_i, and statistical error σ_i from each analysis, the global cumulative distribution function (CDF) is expected to be well-described by a weighted combination of normal distribution CDFs. We take the median of the CDF as the central value across all analyses and the half variation between the 16% and 84% percentiles as the total error. To separate statistical and systematic errors, the same procedure can be performed with errors rescaled as $\sigma_i \to \sqrt{\lambda}\sigma_i$. The variation between $\lambda = 2$ and $\lambda = 1$ yields a separate statistical and systematic error [65]. For an illustration, see Fig. 13.

For a better understanding on the composition of the systematic error, we follow the error budgeting procedure suggested in [65]: For each of the choices made during the analysis chain, e.g., the choice between the VMD or LMD fit to $\tilde{A}(\tau)$ or the different fit ranges, we first determine the total error for each possible option, varying all other components of the analysis. We then construct a second CDF as above with m_i the average of the 16% and 84% percentiles, σ_i the total error and w_i the sum of the weights coming from this choice. Using this CDF, we derive the systematic error as described above for the original CDF; this is our result for the systematic error corresponding to the choice. Note that the estimated systematic errors associated with each of the steps of the analysis by this procedure are correlated, thus they do not sum up quadratically to the full systematic error.

VI. LATTICE RESULTS AND CONTINUUM EXTRAPOLATION

Here, our results using the model averaging described in Sec. V are summarized. Comparing to our earlier publication [31] we have refined the analysis by systematically studying different z-expansion fits and by excluding analyses leading to unphysical TFFs as described in Sec. III C.
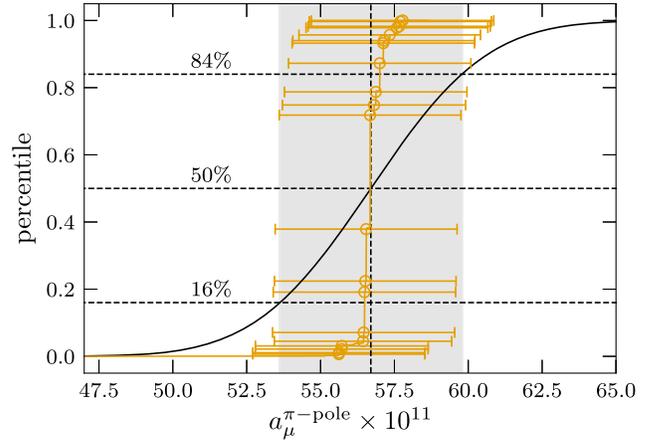
FIG. 13. CDF of $a_\mu^{\pi^0-\text{pole}}$ for the combined fit using $c_{nm} = \{c_{00}, c_{10}, c_{11}, c_{22}\}$. The orange curve shows the CDF of $\approx 15{,}000$ different analyses obtained from their AIC weights, the errors on the orange point show the statistical errors for some analyses with a significant weight in the model averaging. This curve corresponds to statistical errors rescaled by $\lambda = 0$ while the black curve corresponds to rescaling by $\lambda = 1$. The latter gives the central value by the 50% percentile, and the total error as the half variation between the 16% and 84% percentiles shown in gray. To separate the statistical and systematic part of the error, we further calculate the error of the distribution with statistical errors rescaled by $\lambda = 2$ (not shown in the figure), as described in the main text.

For the single ensemble analyses, we perform an AIC averaging over different fit models and fit ranges for $\tilde{A}(\tau)$, different choices of $\tau_{\text{cut}} \approx [1.3, 2.0]$ fm on all three ensembles, different samplings in the momentum plane and different choices of z-expansion fits. The results including error budgeting are summarized in Table III, and are shown in Fig. 16 as diamonds (cD211.054.96), triangles (cC211.060.80), and circles (cB211.072.64), respectively. We note that the error budgets from fit model and fit range are very small, indicating that by restricting the transition form factors in the z-expansion to have at least a 95% contribution from lattice data effectively removes the dependence on the model used to fit the tails of $\tilde{A}(\tau)$ almost completely.

For the z-expansion fits considered in the combined fitting, the set $c_{nm} = \{c_{00}, c_{10}, c_{11}, c_{22}\}$ with the lattice-artifact correction coefficients $\delta_{nm} = \{\delta_{00}, \delta_{10}\}$ leads to the least correlation amongst the coefficients and gives an AIC-averaged fully correlated $\chi^2/\text{d.o.f.} = 1.36$. The AIC-averaged z-expansion coefficients [cf. Eq. (42)] in the continuum limit based on the combined fit across all ensembles are given by

c_{00}	c_{10}	c_{11}	c_{22}
0.2220(48)	$-0.0596(59)$	$-0.050(18)$	0.27(14)

TABLE III. Values of measured observables on each ensemble, with systematic error budgeting indicated below each result.

cB211.072.64	$a_\mu^{\pi^0-\text{pole}} \times 10^{11}$	$\Gamma(\pi \to \gamma\gamma)$ [eV]	b_π [GeV^{-2}]
Value	57.5(1.8)	6.89(0.33)	1.63(0.18)
σ_{stat}	1.39	0.28	0.11
σ_{sys}	1.16	0.18	0.15
Fit model	0.09	0.01	0.01
Fit range	0.00	0.00	0.00
τ_{cut}	0.17	0.06	0.03
Sampling	0.20	0.06	0.04
z-exp.	1.15	0.17	0.14

cC211.060.80	$a_\mu^{\pi^0-\text{pole}} \times 10^{11}$	$\Gamma(\pi \to \gamma\gamma)$ [eV]	b_π [GeV^{-2}]
Value	56.7(2.0)	6.72(0.29)	1.54(0.14)
σ_{stat}	1.33	0.28	0.13
σ_{sys}	1.54	0.08	0.04
Fit model	0.01	0.00	0.00
Fit range	0.07	0.01	0.00
τ_{cut}	0.36	0.03	0.02
Sampling	0.50	0.03	0.03
z-exp.	1.51	0.04	0.02

cD211.054.96	$a_\mu^{\pi^0-\text{pole}} \times 10^{11}$	$\Gamma(\pi \to \gamma\gamma)$ [eV]	b_π [GeV^{-2}]
Value	58.1(2.1)	7.51(0.39)	2.03(0.09)
σ_{stat}	2.08	0.36	0.06
σ_{sys}	0.37	0.13	0.06
Fit model	0.00	0.00	0.00
Fit range	0.01	0.00	0.00
τ_{cut}	0.19	0.08	0.03
Sampling	0.10	0.05	0.03
z-exp.	0.29	0.10	0.06

with correlation matrix

$$\text{cor}(c_{nm}) = \begin{pmatrix} +1.00 & -0.46 & -0.07 & +0.07 \\ -0.46 & +1.00 & +0.03 & -0.08 \\ -0.07 & +0.03 & +1.00 & -0.83 \\ +0.07 & -0.08 & -0.83 & +1.00 \end{pmatrix} \quad (54)$$

and the corresponding corner plot given in Fig. 14. This is our preferred continuum result. We note that the coefficients describing the lattice artifacts are very small and compatible with zero, $\delta_{00} = 0.0005(41)$ and $\delta_{10} = -0.0085(49)$. The resulting continuum TFFs for diagonal and single-virtual kinematics are shown in Fig. 15. The figure further indicates the kinematic region directly supported by lattice data, where the fit is expected to be most reliable. For the single-virtual kinematics we find good agreement with the CELLO [19] bins below $Q^2 \lesssim 2$ GeV^2, while at larger momenta $Q^2 \gtrsim 2$ GeV^2 our
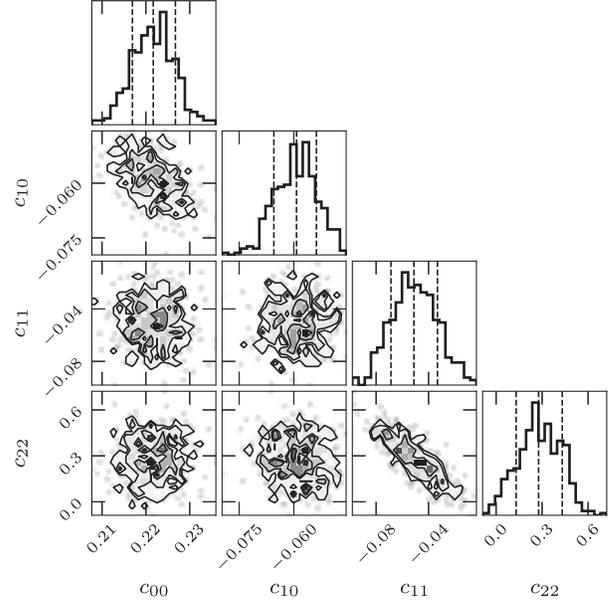
FIG. 14. Corner plot of the AIC averaged coefficients of the combined correlated z-expansion fit, utilizing bootstrap resampling on each ensemble, where only the coefficients c_{00}, c_{10}, c_{11} and c_{22} were used. Notice the anticorrelation between c_{11} and c_{22} and the less severe anticorrelation between c_{00} and c_{10}. All other coefficients are virtually uncorrelated. The corresponding correlation matrix is given in Eq. (B1).

analysis results tend to be systematically lower than the experimental ones.

For the quantities $a_\mu^{\pi^0-\text{pole}}$, $\Gamma(\pi \to \gamma\gamma)$ and b_π, in order to account for different choices of z-expansion parameters in a more conservative way than by using the AIC, we use a procedure inspired by the method described in Ref. [66]. It consists of taking the absolute difference of the central value of the combined fit using the coefficient set $c_{nm} = \{c_{00}, c_{10}, c_{11}, c_{22}\}$ and the weighted average of the results from the other considered fits each weighted with $1/\sigma_{\text{tot}}^2$ as an additional systematic error $\sigma_{\text{sys},z-\text{exp}}$ added in quadrature. We find

$$a_\mu^{\pi^0-\text{pole}} = 56.7(3.1)_{\text{stat}}(1.0)_{\text{sys}}[3.2]_{\text{tot}} \times 10^{-11},$$

$$\Gamma(\pi \to \gamma\gamma) = 7.50(0.48)_{\text{stat}}(0.16)_{\text{sys}}[0.50]_{\text{tot}} \text{ eV},$$

$$b_\pi = 2.16(0.07)_{\text{stat}}(0.19)_{\text{sys}}[0.20]_{\text{tot}} \text{ GeV}^{-2}. \quad (55)$$

These values including a detailed error budget are summarized in Table IV, and are shown in Fig. 16 as crosses. Instead of using the continuum values of the z-expansion coefficients from the combined fit to calculate the continuum values of $a_\mu^{\pi^0-\text{pole}}$, $\Gamma(\pi \to \gamma\gamma)$ and b_π, we could also extrapolate their per-ensemble values in a^2 directly. As depicted in Fig. 16, we find that fits linear in a^2 agree well with the combined fit continuum results, albeit with significantly larger errors at $a^2 = 0$. We note that at the
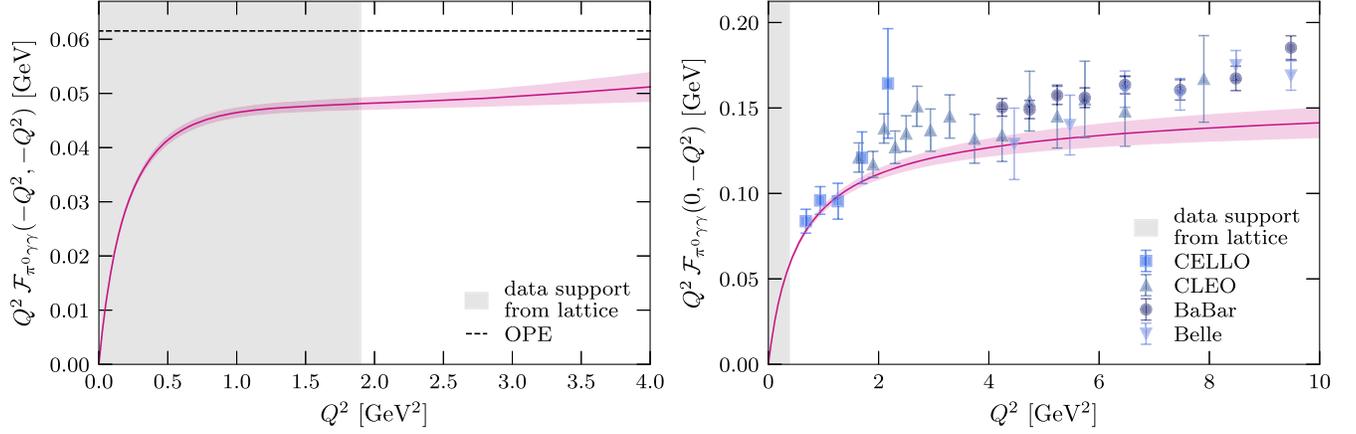
FIG. 15.   Transition form factor from the combined fit for diagonal (left) and single-virtual (right) kinematics. For the diagonal kinematics, the OPE prediction for large Q^2 is indicated by the horizontal dashed line while for the single-virtual kinematics, experimental values from CELLO [19], CLEO [20], *BABAR* [21,22] and Belle [23] are shown. The region with direct support from lattice data is shaded in gray.

current level of accuracy the lattice artifacts in $a_\mu^{\pi^0-\text{pole}}$ are compatible with zero. Hence, a continuum extrapolation with a constant is also possible leading to a better $\chi^2/\text{d.o.f.}$ and a significantly smaller error. However, in order to be on the conservative side, we do not consider this value in our analysis.

Our result for $a_\mu^{\pi^0-\text{pole}}$ is compatible with our earlier analysis [31], where we used a continuum extrapolation in the individual ensemble estimates of each observable instead of one coming from a combined fit to the TFF across ensembles. It is also compatible with the recent lattice result $a_\mu^{\pi^0-\text{pole}} = 57.8 \pm 1.8_{\text{stat}} \pm 0.9_{\text{sys}} \times 10^{-11}$ from the BMW Collaboration in Ref. [29], and $a_\mu^{\pi^0-\text{pole}} = 59.7 \pm 3.6 \times 10^{-11}$ from the Mainz group in Ref. [28]. Comparing to the dispersive result $a_\mu^{\pi^0-\text{pole}} = 63.0^{+2.7}_{-2.1} \times 10^{-11}$ from Refs. [4,25,26], we find that our result is compatible at the level of 1.6σ.

For $\Gamma(\pi \to \gamma\gamma)$, we again find agreement with $\Gamma(\pi \to \gamma\gamma) = 7.11 \pm 0.44_{\text{stat}} \pm 0.21_{\text{sys}}$ GeV^{-2} from Ref. [29]. Our result is

also compatible with the experimental value $\Gamma(\pi \to \gamma\gamma) = 7.802(52)_{\text{stat}}(105)_{\text{sys}}$ eV from Ref. [61].

Finally, for b_π, our result is compatible with the experimental result $b_\pi = 1.84(17)$ GeV^{-2} from Ref. [52] at the level of 1.2σ. We also agree at the level of 1.6σ with $b_\pi = 1.78(12)$ GeV^{-2} from the extraction based on Padé approximants [67] and find a slight tension of 2.1σ with the dispersive result $b_\pi = 1.73(5)$ GeV^{-2} from Refs. [25,26]. In all cases our central value is higher than these results by $O(10\%)$. Further investigation of this quantity from the lattice may therefore be of interest.

VII. COMBINED ANALYSIS OF LATTICE AND EXPERIMENTAL DATA

With our determination of the TFF at the physical point and in the continuum limit we can attempt a combined analysis of our lattice data together with the available experimental data from CELLO [19], CLEO [20], *BABAR* [21,22] and Belle [23]. Such an exercise is interesting for several reasons.

Firstly, we can test whether or not the systematic difference, as evident from Fig. 15, between the experimental data for the single-virtual TFF in the momentum region $Q^2 \gtrsim 2$ GeV^2 and our prediction based on the extrapolation of the lattice data using the modified z-expansion is indeed significant. To this end we perform a combined fit of the same z-expansion as used for our global fit in Sec. VI, i.e., using the coefficient set $c_{nm} = \{c_{00}, c_{10}, c_{11}, c_{22}\}$ with the corresponding lattice-artifact corrections $\delta_{nm} = \{\delta_{00}, \delta_{10}\}$ to the lattice data, while including all the available experimental data up to $Q^2 \simeq 35$ GeV^2. In the $\chi^2/\text{d.o.f.}$ we use equal weights for each set of experimental data and each set of per-ensemble lattice data. As before, we perform a model average over all analysis chains and obtain the AIC-averaged z-expansion coefficients

TABLE IV.   Preferred continuum result from the combined fit using $c_{nm} = \{c_{00}, c_{10}, c_{11}, c_{22}\}$ and $\delta_{nm} = \{\delta_{00}, \delta_{10}\}$. $\sigma_{\text{sys},z-\text{exp}}$. denotes the more conservative systematic error from different z-expansion parameter choices as described in the text.

Combined fit	$a_\mu^{\pi^0-\text{pole}} \times 10^{11}$	$\Gamma(\pi \to \gamma\gamma)$ [eV]	b_π [GeV^{-2}]
Value	56.7(3.2)	7.50(0.50)	2.16(0.20)
σ_{stat}	3.06	0.48	0.07
$\sigma_{\text{sys},z-\text{exp}}.$	0.91	0.14	0.19
σ_{sys}	0.35	0.07	0.01
Fit model	0.00	0.00	0.00
Fit range	0.00	0.00	0.00
τ_{cut}	0.33	0.07	0.00
Sampling	0.19	0.05	0.01

FIG. 16. Combined fit and single ensemble results from Tables III and IV: $a_\mu^{\pi^0-\text{pole}}$ (left), $\Gamma(\pi \to \gamma\gamma)$ (middle) and b_π (right). Indicated are the statistical and total errors. For comparison, we also show linear fits in a^2 on the per-ensemble data points. In each plot, the points with caps on the errorbars correspond to the combined fit (cross), the cD211.054.96 (diamond), the cC211.060.80 (triangle) and the cB211.072.64 (circle) result. In the $\Gamma(\pi \to \gamma\gamma)$ and b_π plots, the square symbols show the experimental values and their errors, namely $\Gamma(\pi \to \gamma\gamma) = 7.802(117)$ eV [61] and $b_\pi = 1.84(17)$ GeV^{-2} [52].

c_{00}	c_{10}	c_{11}	c_{22}
0.2277(31)	−0.0621(52)	−0.1087(85)	0.776(39)

with the correlation matrix given in Appendix B. The resulting TFFs for diagonal and single-virtual kinematics are shown in Fig. 17. We notice the stability and slightly better determination of the z-expansion coefficients c_{00}, c_{10} and the significant shifts in c_{11}, c_{22} with substantially reduced errors. This leads to a much more precise determination of the single-virtual TFF at large values of $Q^2 \gtrsim 1.5$ GeV^2 which is now fully compatible with the experimental data. The AIC-averaged value of $\chi^2/\text{d.o.f.} = 1.35$ demonstrates that the single-virtual data from the lattice and from experiment are indeed completely consistent with each other and can be described by the same TFF function.

For the double-virtual TFF we note the significant difference for $Q^2 \gtrsim 3.0$ GeV^2 which is not surprising given the fact that there is no data support in that momentum range. Additional lattice data is needed there in order to further restrict the TFF.

Secondly, we can test the stability of the lattice-driven determination of $a_\mu^{\pi^0-\text{pole}}$, $\Gamma(\pi \to \gamma\gamma)$ and b_π. For these quantities we obtain from the combined fit to the lattice and experimental data

$$a_\mu^{\pi^0-\text{pole}} = 61.7(2.0)_{\text{stat}}(0.5)_{\text{sys}}[2.0]_{\text{tot}} \times 10^{-11},$$
$$\Gamma(\pi \to \gamma\gamma) = 7.97(0.35)_{\text{stat}}(0.11)_{\text{sys}}[0.37]_{\text{tot}} \text{ eV},$$
$$b_\pi = 2.336(0.049)_{\text{stat}}(0.018)_{\text{sys}}[0.052]_{\text{tot}} \text{ GeV}^{-2}.$$
$$(56)$$
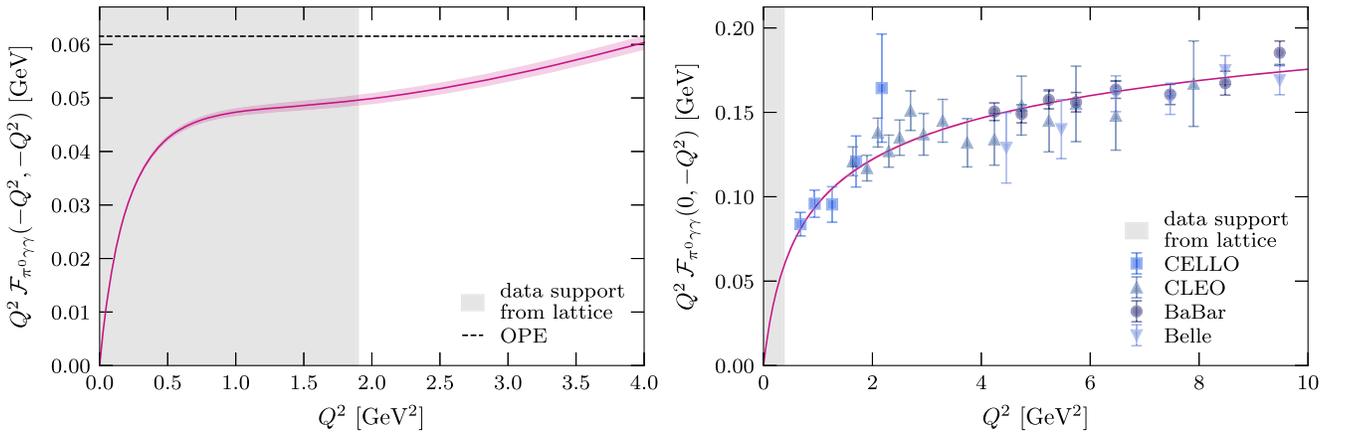
FIG. 17. Transition form factor from the combined fit to lattice and experimental data for diagonal (left) and single-virtual (right) kinematics. For the diagonal kinematics, the OPE prediction for large Q^2 is indicated by the horizontal dashed line while for the single-virtual kinematics, experimental values from CELLO [19], CLEO [20], *BABAR* [21,22] and Belle [23] are shown. The region with direct support from lattice data is shaded in gray.

While these quantities are now much better determined with errors reduced by factors of 1.4, 1.6 and 3.8, the values are still fully compatible with the lattice-only determinations in Eq. (55) in Sec. VI, with the largest shift observed in $a_\mu^{\pi^0-\text{pole}}$ upwards by about 1.5σ. This is easy to understand from the fact that the experimental data shifts the TFF to slightly larger values for $Q^2 \gtrsim 2.5$ GeV^2 outside the lattice support, and the fact that this momentum region may still contribute up to 3% to the total value of $a_\mu^{\pi^0-\text{pole}}$, cf., for example, Fig. 12. Nevertheless, this exercise shows that the low-energy quantities $a_\mu^{\pi^0-\text{pole}}$, $\Gamma(\pi \to \gamma\gamma)$ and b_π are not very sensitive to the large-Q^2 behavior of the TFF and can safely be determined from the lattice alone.

Thirdly, for phenomenological analyses where the large-Q^2 behavior of the TFF is important, the lattice- and data-driven parametrization given here is most useful. However, we would like to add the cautionary remark that our analysis does not contain a full systematic analysis of the dependence on higher-orders in the z-expansion fits. We likely expect that the corresponding systematic error leads to an increased total error for the TFFs, in particular for those close to diagonal virtuality where there is no support from lattice or experimental data.

VIII. CONCLUSIONS AND OUTLOOK

We have presented an *ab-initio* computation of the pion transition form factor at the physical point in the continuum limit using twisted-mass lattice QCD, covering the kinematic range relevant for the extraction of the pion-pole contribution to the HLbL. We are able to include all disconnected Wick contractions contributing to the amplitudes and hence relevant for the calculation of the form factors.

This allows us to provide a parametrization of the transition form factor through the modified z-expansion including the determination of correlations between all coefficients. Our results in the double-virtual and low-Q^2, single-virtual regimes are complementary to the currently available single-virtual experimental data. Our data in the latter regime is of particular interest as new experimental data will become available [68] which can be compared with our prediction from first principles.

In the single-virtual regime we find agreement between our calculated pion form factor and the experimental data up to $Q_2^2 \lesssim 2$ GeV^2, while the extrapolation to $Q_2^2 \gtrsim 2$ GeV^2, based purely on our lattice data, systematically undershoots the experimental data from CELLO [19], CLEO [20], *BABAR* [21,22] and Belle [23], cf. Fig. 15. In contrast, our results for the partial decay width $\Gamma(\pi \to \gamma\gamma)$ and the slope parameter b_π are compatible with the experimental values, cf. Fig. 16. We note, however, that we do not observe

any significant tension between the lattice data and the experimental one, as discussed in Sec. VII.

The main result of this paper, namely the pion-pole contribution to HLbL,

$$a_\mu^{\pi^0-\text{pole}} = 56.7(3.1)_{\text{stat}}(1.0)_{\text{sys}}[3.2]_{\text{tot}} \times 10^{-11}, \quad (57)$$

is compatible with recent lattice and data-driven results, with a relative total error at the sub-6% level. The main systematic effect not fully accounted for in this work is the effect of working at a fixed finite spatial volume. On general grounds finite-size effects can be estimated in QCD to be of the order $\mathcal{O}(\exp(-m_\pi L))$, i.e., they are exponentially suppressed. With our values of $m_\pi L = 3.6$–3.9 we expect them to be at most a few percent. In practice, at the physical lattice sizes of 5.1–5.5 fm employed in this work, finite-size effects appear to be negligible compared to the current precision for $a_\mu^{\pi^0-\text{pole}}$ and $\Gamma(\pi \to \gamma\gamma)$, as demonstrated in Ref. [29] for similar physical lattice sizes. Since there are no lattice results available from previous work for the slope parameter and its sensitivity to finite-size effects, we can not exclude it to be affected more. We leave the study of finite-size effects as future work, potentially using ETMC ensembles at physical pion mass and the same range of lattice spacings with larger lattice sizes up to ≈ 7.5 fm.

To achieve further precision in future work, several options are available: One can use momentum $\vec{p} \neq 0$ for the pseudoscalar creation operator (i.e., the moving frame) to give a better coverage of the single-virtual axis. In addition, a fourth physical point ensemble with approximately the same volume as the three used here, but with an even smaller lattice spacing, is currently in production and could then also be included in this calculation.

ACKNOWLEDGMENTS

APPENDIX A: 3D INTEGRAL REPRESENTATION WEIGHTS

The weight functions $w_{1,2}$ appearing in Eqs. (46) and (47) presented here are taken from [17] and have been derived in [18] using the method of Gegenbauer polynomials [83–87].

The weight functions read

$$w_1(Q_1, Q_2, \tau) = \left(-\frac{2\pi}{3}\right)\sqrt{1-\tau^2}\,\frac{Q_1^3 Q_2^3}{Q_2^2 + m_\pi^2}\, I_1(Q_1, Q_2, \tau),$$
(A1)

$$w_2(Q_1, Q_2, \tau) = \left(-\frac{2\pi}{3}\right)\sqrt{1-\tau^2}\,\frac{Q_1^3 Q_2^3}{(Q_1 + Q_2)^2 + m_\pi^2}$$
$$\times I_2(Q_1, Q_2, \tau),$$
(A2)

with

$$I_1(Q_1, Q_2, \tau) = X(Q_1, Q_2, \tau)[8P_1 P_2 Q_1 Q_2 \tau - 2P_1 P_3(Q_2^4/m_\mu^2 - 2Q_2^2) + 4P_2 P_3 Q_1^2 - 4P_2 - 2P_1(2 - Q_2^2/m_\mu^2 + 2Q_1 Q_2 \tau/m_\mu^2)$$
$$- 2P_3(4 + Q_1^2/m_\mu^2 - 2Q_2^2/m_\mu^2) + 2/m_\mu^2] - 2P_1 P_2(1 + (1 - R_{m1})Q_1 Q_2 \tau/m_\mu^2)$$
$$+ P_1 P_3(2 - (1 - R_{m1})Q_2^2/m_\mu^2) + P_2 P_3(2 + (1 - R_{m1})^2 Q_1 Q_2 \tau/m_\mu^2)$$
$$+ P_1(1 - R_{m1})/m_\mu^2 + 3P_3(1 - R_{m1})/m_\mu^2$$
(A3)

and

$$I_2(Q_1, Q_2, \tau) = X(Q_1, Q_2, \tau)[4P_1 P_2 Q_1 Q_2 \tau + 2P_1 P_3 Q_2^2 - 2P_1 + 2P_2 P_3 Q_1^2 - 2P_2 - 4P_3 - 4/m_\mu^2]$$
$$- 2P_1 P_2 - 3P_1(1 - R_{m2})/(2m_\mu^2) - 3P_2(1 - R_{m1})/(2m_\mu^2) - P_3(2 - R_{m1} - R_{m2})/(2m_\mu^2)$$
$$+ P_1 P_3[2 + 3(1 - R_{m2})Q_2^2/(2m_\mu^2) + (1 - R_{m2})^2 Q_1 Q_2 \tau/(2m_\mu^2)]$$
$$+ P_2 P_3[2 + 3(1 - R_{m1})Q_1^2/(2m_\mu^2) + (1 - R_{m1})^2 Q_1 Q_2 \tau/(2m_\mu^2)],$$
(A4)

where

$$Q_3^2 = (Q_1 + Q_2)^2 = Q_1^2 + 2Q_1 Q_2 \tau + Q_2^2,$$
(A5)

$$\tau = \cos\theta,$$
(A6)

$$P_i = 1/Q_i^2, \quad i \in \{1, 2, 3\}.$$
(A7)

Further,

$$X(Q_1 Q_2, \tau) = \frac{1}{Q_1 Q_2 x}\arctan\left(\frac{zx}{1 - z\tau}\right),$$
(A8)

$$x = \sqrt{1 - \tau^2},$$
(A9)

$$z = \frac{Q_1 Q_2}{4m_\mu^2}(1 - R_{m1})(1 - R_{m2}),$$
(A10)

$$R_{mi} = \sqrt{1 + \frac{4m_\mu^2}{Q_i^2}}, \quad i \in \{1, 2\}.$$
(A11)

For a detailed discussion of the behavior of $w_{1,2}$ in different limits see [17].

APPENDIX B: CORRELATION MATRIX FOR THE COMBINED FIT TO LATTICE AND EXPERIMENTAL DATA

Here we provide the correlation matrix for the combined fit of the modified z-expansion as described in Sec. VII using the coefficient set $c_{nm} = \{c_{00}, c_{10}, c_{11}, c_{22}\}$ with the corresponding lattice-artifact corrections $\delta_{nm} = \{\delta_{00}, \delta_{10}\}$ to the lattice and experimental data,

$$\text{cor}(c_{nm}) = \begin{pmatrix} +1.00 & -0.14 & +0.54 & -0.78 \\ -0.14 & +1.00 & -0.12 & -0.06 \\ +0.54 & -0.12 & +1.00 & +0.07 \\ -0.78 & -0.06 & +0.07 & +1.00 \end{pmatrix}. \quad \text{(B1)}$$

We refrain from providing the corresponding corner plot, since it serves the purpose of illustration only.

[1] B. Abi, T. Albahri, S. Al-Kilani, D. Allspach, L. P. Alonzi, A. Anastasi *et al.* (Muon g-2 Collaboration), Measurement of the positive muon anomalous magnetic moment to 0.46 ppm, Phys. Rev. Lett. **126,** 141801 (2021).

[2] D. P. Aguillard *et al.* (Muon g-2 Collaboration), Measurement of the positive muon anomalous magnetic moment to 0.20 ppm, arXiv:2308.06230.

[3] G. W. Bennett and Others, Final report of the Muon E821 anomalous magnetic moment measurement at BNL, Phys. Rev. D **73,** 072003 (2006).

[4] T. Aoyama, N. Asmussen, M. Benayoun, J. Bijnens, T. Blum, M. Bruno *et al.*, The anomalous magnetic moment of the muon in the standard model, Phys. Rep. **887,** 1 (2020).

[5] T. Blum, S. Chowdhury, M. Hayakawa, and T. Izubuchi, Hadronic light-by-light scattering contribution to the muon anomalous magnetic moment from lattice QCD, Phys. Rev. Lett. **114,** 012001 (2015).

[6] T. Blum, N. Christ, M. Hayakawa, T. Izubuchi, L. Jin, and C. Lehner, Lattice calculation of hadronic light-by-light contribution to the muon anomalous magnetic moment, Phys. Rev. D **93,** 014503 (2016).

[7] T. Blum, N. Christ, M. Hayakawa, T. Izubuchi, L. Jin, C. Jung, and C. Lehner, Connected and leading disconnected hadronic light-by-light contribution to the muon anomalous magnetic moment with a physical pion mass, Phys. Rev. Lett. **118,** 022005 (2017).

[8] T. Blum, N. Christ, M. Hayakawa, T. Izubuchi, L. Jin, C. Jung, and C. Lehner, Using infinite volume, continuum QED and lattice QCD for the hadronic light-by-light contribution to the muon anomalous magnetic moment, Phys. Rev. D **96,** 034515 (2017).

[9] E. H. Chao, R. J. Hudspith, A. Gérardin, J. R. Green, H. B. Meyer, and K. Ottnad, Hadronic light-by-light contribution to $(g-2)_\mu$ from lattice QCD: A complete calculation, Eur. Phys. J. C **81,** 5 (2021).

[10] E.-H. Chao, R. J. Hudspith, A. Gérardin, J. R. Green, and H. B. Meyer, The charm-quark contribution to light-by-light scattering in the muon $g-2$ from lattice QCD, Eur. Phys. J. C **82,** 664 (2022).

[11] N. Asmussen, E.-H. Chao, A. Gérardin, J. R. Green, R. J. Hudspith, H. B. Meyer, and A. Nyffeler, Hadronic light-by-light scattering contribution to the muon $g-2$ from lattice QCD: Semi-analytical calculation of the QED kernel, J. High Energy Phys. 04 (2023) 040.

[12] G. Colangelo, M. Hoferichter, M. Procura, and P. Stoffer, Dispersive approach to hadronic light-by-light scattering, J. High Energy Phys. 09 (2014) 091.

[13] G. Colangelo, M. Hoferichter, B. Kubis, M. Procura, and P. Stoffer, Towards a data-driven analysis of hadronic light-by-light scattering, Phys. Lett. B **738,** 6 (2014).

[14] G. Colangelo, M. Hoferichter, M. Procura, and P. Stoffer, Dispersion relation for hadronic light-by-light scattering: theoretical foundations, J. High Energy Phys. 09 (2015) 074.

[15] V. Pauk and M. Vanderhaeghen, Anomalous magnetic moment of the muon in a dispersive approach, Phys. Rev. D **90,** 113012 (2014).

[16] M. Knecht and A. Nyffeler, Hadronic light by light corrections to the muon $g-2$: The Pion pole contribution, Phys. Rev. D **65,** 073034 (2002).

[17] A. Nyffeler, Precision of a data-driven estimate of hadronic light-by-light scattering in the muon $g-2$: Pseudoscalar-pole contribution, Phys. Rev. D **94,** 053006 (2016).

[18] F. Jegerlehner and A. Nyffeler, The muon $g-2$, Phys. Rep. **477,** 1 (2009).

[19] H. J. Behrend *et al.* (CELLO Collaboration), A Measurement of the π^0, η and η-prime electromagnetic form-factors, Z. Phys. C **49,** 401 (1991).

[20] J. Gronberg *et al.* (CLEO Collaboration), Measurements of the meson—photon transition form-factors of light pseudoscalar mesons at large momentum transfer, Phys. Rev. D **57,** 33 (1998).

[21] B. Aubert *et al.* (BABAR Collaboration), Measurement of the $\gamma\gamma^* \to \pi^0$ transition form factor, Phys. Rev. D **80,** 052002 (2009).

[22] P. del Amo Sanchez *et al.* (BABAR Collaboration), Measurement of the $\gamma\gamma^* \to \eta$ and $\gamma\gamma^* \to \eta'$ transition form factors, Phys. Rev. D **84,** 052001 (2011).

[23] S. Uehara, Y. Watanabe, H. Nakazawa, I. Adachi, H. Aihara, D. M. Asner *et al.* (Belle Collaboration), Measurement of $\gamma\gamma^* \to \pi^0$ transition form factor at Belle, Phys. Rev. D **86,** 092007 (2012).

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

[25] M. Hoferichter, B.-L. Hoid, B. Kubis, S. Leupold, and S. P. Schneider, Pion-pole contribution to hadronic light-by-light scattering in the anomalous magnetic moment of the muon, Phys. Rev. Lett. **121**, 112002 (2018).

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

[27] A. Gérardin, H. B. Meyer, and A. Nyffeler, Lattice calculation of the pion transition form factor $\pi^0 \to \gamma^*\gamma^*$, Phys. Rev. D **94**, 074507 (2016).

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

[29] A. Gérardin, W. E. A. Verplanke, G. Wang, Z. Fodor, J. N. Guenther, L. Lellouch *et al.*, Lattice calculation of the π^0, η and η' transition form factors and the hadronic light-by-light contribution to the muon $g - 2$, arXiv:2305.04570.

[30] S. Burri, C. Alexandrou, S. Bacchio, G. Bergner, J. Finkenrath, A. Gasbarro *et al.*, Pion-pole contribution to HLbL from twisted mass lattice QCD at the physical point, Proc. Sci. LATTICE2021 (**2022**) 519.

[31] S. Burri, G. Kanwar, C. Alexandrou, S. Bacchio, G. Bergner, J. Finkenrath *et al.*, Pseudoscalar-pole contributions to the muon $g - 2$ at the physical point, Proc. Sci. LATTICE2022 (**2023**) 306.

[32] C. Alexandrou, S. Bacchio, S. Burri, J. Finkenrath, A. Gasbarro, K. Hadjiyiannakou *et al.*, The $\eta \to \gamma^*\gamma^*$ transition form factor and the hadronic light-by-light η-pole contribution to the muon $g - 2$ from lattice QCD, Phys. Rev. D **108**, 054509 (2023).

[33] X.-d. X. Ji and C.-w. C. Jung, Studying hadronic structure of the photon in lattice QCD, Phys. Rev. Lett. **86**, 208 (2001).

[34] X.-d. Ji and C.-w. Jung, Photon structure functions from quenched lattice QCD, Phys. Rev. D **64**, 034506 (2001).

[35] J. J. Dudek and R. G. Edwards, Two photon decays of charmonia from lattice QCD, Phys. Rev. Lett. **97**, 172001 (2006).

[36] R. Frezzotti and G. C. Rossi, Chirally improving Wilson fermions. 1. O(a) improvement, J. High Energy Phys. 08 (2004) 007.

[37] R. Frezzotti, P. A. Grassi, S. Sint, and P. Weisz (Alpha Collaboration), Lattice QCD with a chirally twisted mass term, J. High Energy Phys. 08 (2001) 058.

[38] R. Frezzotti and G. C. Rossi, Chirally improving Wilson fermions. II. Four-quark operators, J. High Energy Phys. 10 (2004) 070.

[39] C. Alexandrou *et al.*, Simulating twisted mass fermions at physical light, strange and charm quark masses, Phys. Rev. D **98**, 054518 (2018).

[40] G. Bergner, P. Dimopoulos, J. Finkenrath, E. Fiorenza, R. Frezzotti, M. Garofalo *et al.*, Quark masses and decay constants in $N_f = 2 + 1 + 1$ isoQCD with Wilson clover twisted mass fermions, Proc. Sci. LATTICE2019 (**2020**) 181.

[41] C. Alexandrou, S. Bacchio, G. Bergner, P. Dimopoulos, J. Finkenrath, R. Frezzotti *et al.* (Extended Twisted Mass Collaboration), Ratio of kaon and pion leptonic decay constants with $N_f = 2 + 1 + 1$ Wilson-clover twisted-mass fermions, Phys. Rev. D **104**, 074520 (2021).

[42] J. Finkenrath, C. Alexandrou, S. Bacchio, M. Constantinou, P. Dimopoulos, R. Frezzotti *et al.*, Twisted mass gauge ensembles at physical values of the light, strange and charm quark masses, Proc. Sci. LATTICE2021 (**2022**) 284.

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

[44] A. Stathopoulos, J. Laeuchli, and K. Orginos, Hierarchical probing for estimating the trace of the matrix inverse on toroidal lattices, SIAM J. Sci. Comput. **35**, S299 (2013).

[45] A. S. Gambhir, A. Stathopoulos, and K. Orginos, Deflation as a method of variance reduction for estimating the trace of a matrix inverse, SIAM J. Sci. Comput. **39**, A532 (2017).

[46] C. Alexandrou, S. Bacchio, M. Constantinou, J. Finkenrath, K. Hadjiyiannakou, K. Jansen *et al.*, Complete flavor decomposition of the spin and momentum fraction of the proton using lattice QCD simulations at physical pion mass, Phys. Rev. D **101**, 094513 (2020).

[47] C. Alexandrou *et al.*, Disconnected contribution to the LO HVP term of muon g-2 from ETMC, Proc. Sci., LATTICE2022 (**2023**) 303.

[48] B. Moussallam, Chiral sum rules for $\mathcal{L}^{WZ}_{(6)}$ parameters and its application to π^0, η, η' decays, Phys. Rev. D **51**, 4939 (1995).

[49] M. Knecht, S. Peris, M. Perrottet, and E. de Rafael, Decay of pseudoscalars into lepton pairs and large-N_c QCD, Phys. Rev. Lett. **83**, 5230 (1999).

[50] S. L. Adler, Axial-vector vertex in spinor electrodynamics, Phys. Rev. **177**, 2426 (1969).

[51] J. S. Bell and R. Jackiw, A PCAC puzzle: $\pi^0 \to \gamma\gamma$ in the σ-model, Nuovo Cimento A **60**, 47 (1969).

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

[53] G. P. Lepage and S. J. Brodsky, Exclusive processes in quantum chromodynamics: Evolution equations for hadronic wave functions and the form-factors of mesons, Phys. Lett. **87B**, 359 (1979).

[54] G. P. Lepage and S. J. Brodsky, Exclusive processes in perturbative quantum chromodynamics, Phys. Rev. D **22**, 2157 (1980).

[55] S. J. Brodsky and G. P. Lepage, Large angle two photon exclusive channels in quantum chromodynamics, Phys. Rev. D **24**, 1808 (1981).

[56] V. A. Nesterenko and A. V. Radyushkin, Comparison of the QCD sum rule approach and perturbative QCD analysis for $\gamma^*\gamma^* \to \pi^0$ process, Sov. J. Nucl. Phys. **38**, 284 (1983), https://inspirehep.net/literature/182367.

[57] V. A. Novikov, M. A. Shifman, A. I. Vainshtein, M. B. Voloshin, and V. I. Zakharov, Use and misuse of QCD sum rules, factorization and related topics, Nucl. Phys. **B237**, 525 (1984).

[58] C. G. Boyd, B. Grinstein, and R. F. Lebed, Model independent determinations of $\bar{B} \to Dl\bar{\nu}$, $D^*l\bar{\nu}$ form-factors, Nucl. Phys. **B461**, 493 (1996).

[59] C. Bourrely, I. Caprini, and L. Lellouch, Model-independent description of $B \to \pi l\nu$ decays and a determination of $|V_{ub}|$, Phys. Rev. D **79**, 013008 (2009).

[60] I. Larin *et al.* (PrimEx Collaboration), A new measurement of the π^0 radiative decay width, Phys. Rev. Lett. **106,** 162303 (2011).

[61] I. Larin, Y. Zhang, A. Gasparian, L. Gan, R. Miskimen, M. Khandaker *et al.*, Precision measurement of the neutral pion lifetime, Science **368,** 506 (2020).

[62] H. Akaike, A Bayesian analysis of the minimum AIC procedure, Ann. Inst. Stat. Math. **30,** 9 (1978).

[63] W. I. Jay and E. T. Neil, Bayesian model averaging for analysis of lattice field theory results, Phys. Rev. D **103,** 114502 (2021).

[64] S. Borsanyi, S. Durr, Z. Fodor, C. Hoelbling, S. D. Katz, S. Krieg *et al.*, *Ab initio* calculation of the neutron-proton mass difference, Science **347,** 1452 (2015).

[65] S. Borsanyi, Z. Fodor, J. N. Guenther, C. Hoelbling, S. D. Katz, L. Lellouch *et al.*, Leading hadronic contribution to the muon magnetic moment from lattice QCD, Nature (London) **593,** 51 (2021).

[66] C. Alexandrou, S. Bacchio, G. Bergner, M. Constantinou, M. D. Carlo, P. Dimopoulos *et al.* (Extended Twisted Mass Collaboration), Quark masses using twisted-mass fermion gauge ensembles, Phys. Rev. D **104,** 074515 (2021).

[67] P. Masjuan, $\gamma^*\gamma \to \pi^0$ transition form factor at low-energies from a model-independent approach, Phys. Rev. D **86,** 094021 (2012).

[68] M. Ablikim, M. N. Achasov, P. Adlarson, S. Ahmed, M. Albrecht, M. Alekseev *et al.*, Future physics programme of BESIII, Chin. Phys. C **44,** 040001 (2020).

[69] www.gauss-centre.eu.

[70] D. Krause, JUWELS: Modular tier-0/1 supercomputer at the Jülich supercomputing centre, J. Large-Scale Res. Facil. **5,** A135 (2019).

[71] www.lrz.de.

[72] K. Jansen and C. Urbach, tmLQCD: A program suite to simulate Wilson twisted mass lattice QCD, Comput. Phys. Commun. **180,** 2717 (2009).

[73] A. Deuzeman, K. Jansen, B. Kostrzewa, and C. Urbach, Experiences with OpenMP in tmLQCD, Proc. Sci. LATTICE2013 (**2014**) 416.

[74] A. Abdel-Rehim, F. Burger, A. Deuzeman, K. Jansen, B. Kostrzewa, L. Scorzato *et al.*, Recent developments in the tmLQCD software suite, Proc. Sci. LATTICE2013 (**2014**) 414.

[75] B. Kostrzewa, S. Bacchio, J. Finkenrath, M. Garofalo, F. Pittler, S. Romiti *et al.*, Twisted mass ensemble generation on GPU machines, Proc. Sci. LATTICE2022 (**2023**) 340.

[76] C. Alexandrou, S. Bacchio, J. Finkenrath, A. Frommer, K. Kahl, and M. Rottmann, Adaptive aggregation-based domain decomposition multigrid for twisted mass fermions, Phys. Rev. D **94,** 114509 (2016).

[77] C. Alexandrou, S. Bacchio, and J. Finkenrath, Multigrid approach in shifted linear systems for the non-degenerated twisted mass operator, Comput. Phys. Commun. **236,** 51 (2019).

[78] M. A. Clark, R. Babich, K. Barros, R. C. Brower, and C. Rebbi, Solving Lattice QCD systems of equations using mixed precision solvers on GPUs, Comput. Phys. Commun. **181,** 1517 (2010).

[79] R. Babich, M. A. Clark, B. Joo, G. Shi, R. C. Brower, and S. Gottlieb, Scaling lattice QCD beyond 100 GPUs, in *Proceedings of the SC11 International Conference for High Performance Computing, Networking, Storage and Analysis* (2011), arXiv:1109.2935.

[80] M. A. Clark, B. Joó, A. Strelchenko, M. Cheng, A. Gambhir, and R. Brower, Accelerating lattice QCD multigrid on GPUs using fine-grained parallelization, arXiv:1612.07873.

[81] J. D. Hunter, Matplotlib: A 2d graphics environment, Comput. Sci. Eng. **9,** 90 (2007).

[82] D. Foreman-Mackey, CORNER.PY: Scatterplot matrices in python, J. Open Source Softwaare **1,** 24 (2016).

[83] K. Johnson, M. Baker, and R. Willey, Self-energy of the electron, Phys. Rev. **136,** B1111 (1964).

[84] K. Johnson, R. Willey, and M. Baker, Vacuum polarization in quantum electrodynamics, Phys. Rev. **163,** 1699 (1967).

[85] J. L. Rosner, Higher-order contributions to the divergent part of Z_3 in a model quantum electrodynamics, Ann. Phys. (N.Y.) **44,** 11 (1967).

[86] M. J. Levine and R. Roskies, Hyperspherical approach to quantum electrodynamics: Sixth-order magnetic moment, Phys. Rev. D **9,** 421 (1974).

[87] M. J. Levine, E. Remiddi, and R. Roskies, Analytic contributions to the g factor of the electron in sixth order, Phys. Rev. D **20,** 2068 (1979).