Eur. Phys. J. C (2024) 84:721
https://doi.org/10.1140/epjc/s10052-024-12964-7

THE EUROPEAN PHYSICAL JOURNAL C

Regular Article - Theoretical Physics

Tensions in $e^+e^- \to \pi^+\pi^-(\gamma)$ measurements: the new landscape of data-driven hadronic vacuum polarization predictions for the muon $g-2$

Michel Davier[1,a], **Andreas Hoecker**[2], **Anne-Marie Lutz**[1], **Bogdan Malaescu**[3], **Zhiqing Zhang**[1]

[1] IJCLab, Université Paris-Saclay et CNRS/IN2P3, Orsay, France
[2] CERN, 1211 Geneva 23, Switzerland
[3] LPNHE, Sorbonne Université, Université Paris Cité, CNRS/IN2P3, Paris, France

Received: 19 December 2023 / Accepted: 30 May 2024

Abstract The situation of the experimental data used in the dispersive evaluation of the hadronic vacuum polarization contribution to the anomalous magnetic moment of the muon is assessed in view of two recent measurements: $e^+e^- \to \pi^+\pi^-$ cross sections in the ρ resonance region by CMD-3 and a study of higher-order radiative effects in the initial-state-radiation processes $e^+e^- \to \mu^+\mu^-\gamma$ and $e^+e^- \to \pi^+\pi^-\gamma$ by BABAR. The impact of the latter study on the KLOE and BESIII cross-section measurements is evaluated and found to be indicative of larger systematic effects than uncertainties assigned. The new situation also warrants a reappraisal of the independent information provided by hadronic τ decays, including state-of-the-art isospin-breaking corrections. The findings cast a new light on the longstanding deviation between the muon $g-2$ measurement and the Standard Model prediction using the data-driven dispersive approach, and the comparison with lattice QCD calculations.

1 Introduction

The muon anomalous magnetic moment, characterized by the gyromagnetic anomaly, $a_\mu = (g-2)/2$, is the subject of notable current interest. Recent measurements at Fermilab [1,2] with increased precision exhibit a spectacular excess compared to the Standard Model (SM) prediction [3] whose lowest-order (LO) hadronic vacuum polarization (HVP) contribution is evaluated with dispersion integrals involving $e^+e^- \to$ hadrons cross-section data [4–9]. Other contributions are from quantum electrodynamics [10,11], electroweak interactions [12,13], NLO HVP [9], NNLO HVP [14], hadronic light-by-light [15–29]. While the observed 5σ deviation could be regarded as a serious clue for physics beyond the SM, it must be taken with extreme caution in view of significant tensions among the data sets entering the HVP calculations. Although the decade-long discrepancy between the two most precise results of the $e^+e^- \to \pi^+\pi^-(\gamma)$ cross section[1] from KLOE [30–33] and BABAR [34,35] was already taken into account in the systematic uncertainty assigned to the prediction [3,8], the recent measurement of the same process by CMD-3 [36,37] is in conflict with all previous determinations, thus requiring a close scrutiny of all the e^+e^- input data.

A tension of a different nature arose almost four years ago with the first precise HVP calculation using QCD on the lattice [38] that resulted in a 2.1σ larger lowest-order contribution than the dispersive analysis. The tension is exacerbated to 3.7σ if the comparison is restricted to an intermediate HVP window in Euclidean time [39], which can be calculated more precisely on the lattice. Confirmation of this discrepancy has since then been obtained by several independent lattice groups [40–43]. This situation calls again for specific studies cross-checking both approaches [44,45].

This paper reviews the existing tensions among the $e^+e^- \to \pi^+\pi^-$ cross-section measurements, and discusses in detail systematic uncertainties related to higher-order (HO) effects [46] in the measurements relying on initial state photon radiation. In view of the results obtained we reappraise the use of τ hadronic spectral functions in the dispersive approach, and discuss the discrepancies of the dispersive HVP calculations with lattice QCD and the a_μ experimental result.

[a] e-mail: davier@lal.in2p3.fr (corresponding author)

[1] If not explicitly stated, final state photon radiation is implied throughout this paper for all hadronic final states.

Published online: 22 July 2024

2 Tensions among the $e^+e^- \to \pi^+\pi^-(\gamma)$ data sets

The $e^+e^- \to \pi^+\pi^-$ channel contributes with 73% to the lowest-order HVP contribution to a_μ in the dispersive approach, and 58% to its uncertainty-squared. It also leads to the largest observed discrepancies among some of the most precise data sets. The studies in this paper therefore focus on that process.

The longest known and most critical tensions occur between precise cross section measurements from KLOE and BABAR. Albeit heavily discussed in the framework of the Muon $g-2$ Theory Initiative [47], no understanding of the difference could be achieved and consequently no solution to the problem emerged. The discrepancy was bridged by inflated uncertainties in the corresponding HVP contribution.

The available $e^+e^- \to \pi^+\pi^-$ cross-section measurements, zoomed into the ρ peak region, are shown in Fig. 1. Their combination and 1σ uncertainty, obtained using the DHMZ methodology implemented in the HVPTools software [48,49], are indicated by the green band. The spline-based combination procedure[2] takes into account all known correlations and accounts for measurement tensions. It has been thoroughly validated through closure tests [48]. Compared to our last update [8], we added the more recent SND20 [50] and CMD-3 [36,37] data, while also employing an updated version of the covariance matrix provided by BESIII [51].

Relative comparisons between the most precise individual measurements and the combination are shown for the ρ resonance region in Fig. 2, and for the BABAR and CMD-3 data in a wider window in Fig. 3. A large tension arises between

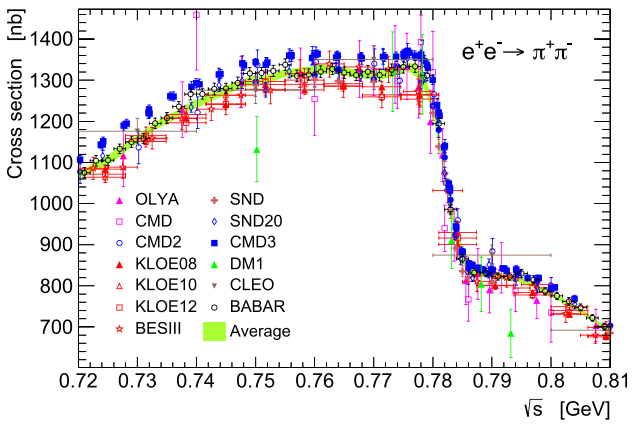

Fig. 1 Bare $e^+e^- \to \pi^+\pi^-$ cross section versus centre-of-mass energy in the ρ peak region. The error bars of the data points include statistical and systematic uncertainties added in quadrature. The green band shows the HVPTools combination within its 1σ uncertainty

[2] Since the main purpose of the combination here is to provide a common reference for comparing the various measurements, we do not employ the analyticity-based constraints used in Ref. [8].

CMD-3 and KLOE, which provide the, respectively, largest and smallest cross-section measurements. Tensions are also observed between BABAR and CMD-3 in the central ρ resonance region, while they agree at low and high energies. The CMD-3 data also exhibit a 2.8σ discrepancy with the older CMD-2 results by the same collaboration [52]. Extensive discussions with CMD-3/2 physicists in the framework of the Muon $g-2$ Theory Initiative [53] did not reveal any obvious problem in the new results. A summary of these discussions is available [54].

Figure 4 (top) shows the local combination weights versus $\sqrt{s}$ for each data set. They take into account the uncertainties of the measurements and their correlations, as well as the corresponding point-spacing and binning [48,49]. While previously the BABAR and KLOE measurements dominated the combination over the entire energy range, the more recent CMD-3 and SND20 data receive important weights, too. The group of experiments labelled "Other exp" corresponds to older data, often with incomplete radiative corrections, which receive small weights throughout.

The bottom panel of Fig. 4 displays the uncertainty scale factor versus $\sqrt{s}$, derived based on the local compatibility among the measurements [48,49].[3] Large scale factors due to tensions indicate the presence of systematic effects that are not included in the measurement uncertainties. They require a conservative uncertainty treatment in the combination [3,8].

Figure 5 shows the pull magnitude (significance) between pairs of the three most precise $e^+e^- \to \pi^+\pi^-$ experiments, computed as the absolute value of the difference of the contributions to a_μ divided by its uncertainty, in various energy intervals. The three KLOE measurements [30–32] have been combined into one data set [33]. The difference between BABAR and CMD-3 rises to a significance of 2–3σ on the ρ peak, while reasonable agreement is seen at lower and higher energies. The differences between BABAR and KLOE are also at the 2–3σ level in the ρ peak region, reaching up to 4σ at higher energy, while good agreement is seen at lower energy. The largest differences are observed between CMD-3 and KLOE, with significance above 5σ around the ρ peak. When probing the broader energy interval 0.6−0.975 GeV, covering the ρ peak, the significance of the difference between BABAR and CMD-3 is 2.2σ, that between BABAR and KLOE is 3.0σ, while CMD-3 and KLOE differ by 5.1σ (Fig. 5, bottom). When extending the comparisons to the maximal regions of overlap between pairs of experiments, the differences are diluted to 2.1σ between BABAR and CMD-3, 1.5σ between BABAR and KLOE, and 3.3σ between CMD-3 and KLOE, respectively, owing

[3] While the uncertainty rescaling is applied to the combined $\pi^+\pi^-$ cross-section uncertainty to account for local inconsistencies among the measurements, a global systematic tension must also be taken into account in the HVP calculation [8].

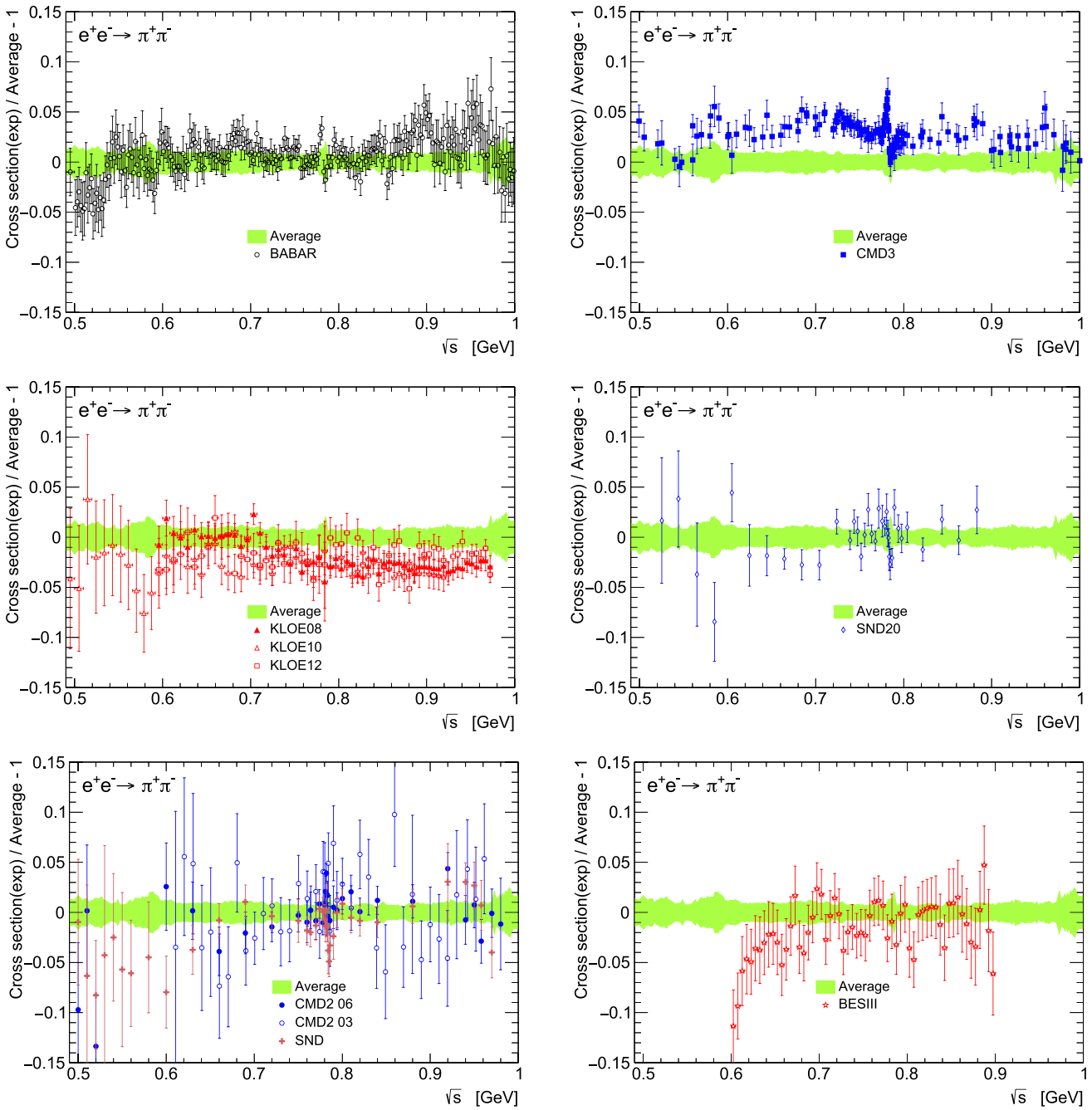

Fig. 2 Comparison between $e^+e^- \to \pi^+\pi^-$ cross-section measurements from BABAR [34,35], KLOE08 [30], KLOE10 [31], KLOE12 [32], BESIII [51], CMD-203 [55], CMD-206 [52], SND [56], SND20 [50], CMD-3 [36,37], and the HVPTools combination. The error bars include statistical and systematic uncertainties added in quadrature

to the better inter-experiment agreement and larger KLOE uncertainties below and above the peak of the resonance.

The interesting possibility to resolve the tensions between different data sets using basic theoretical constraints on the pion form factor from analyticity and unitary has been investigated [57]. However, the theory-constrained fits are loose enough to accommodate even the extreme cases of KLOE and CMD-3 [58].

3 BABAR study of additional photon radiation

The BABAR collaboration performed unique measurements of additional photon radiation in the initial state radiation (ISR) processes $e^+e^- \to \mu^+\mu^-\gamma$ and $e^+e^- \to \pi^+\pi^-\gamma$. Hard NLO radiation with one additional photon was studied in Refs. [34,35]. A new analysis [46] based on the full available data set extended that study and included for the

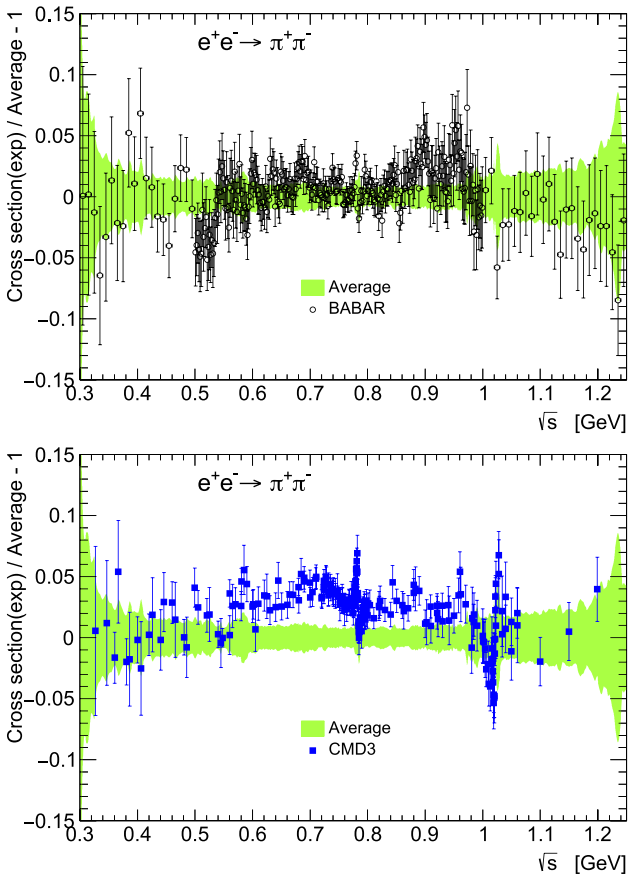

Fig. 3 Comparison between $e^+e^- \to \pi^+\pi^-$ cross-section measurements from BABAR [34,35] (top panel), CMD-3 [36,37] (bottom), and the HVPTools combination of all available data in the $\sqrt{s}$ range covered by CMD-3. The error bars include statistical and systematic uncertainties added in quadrature

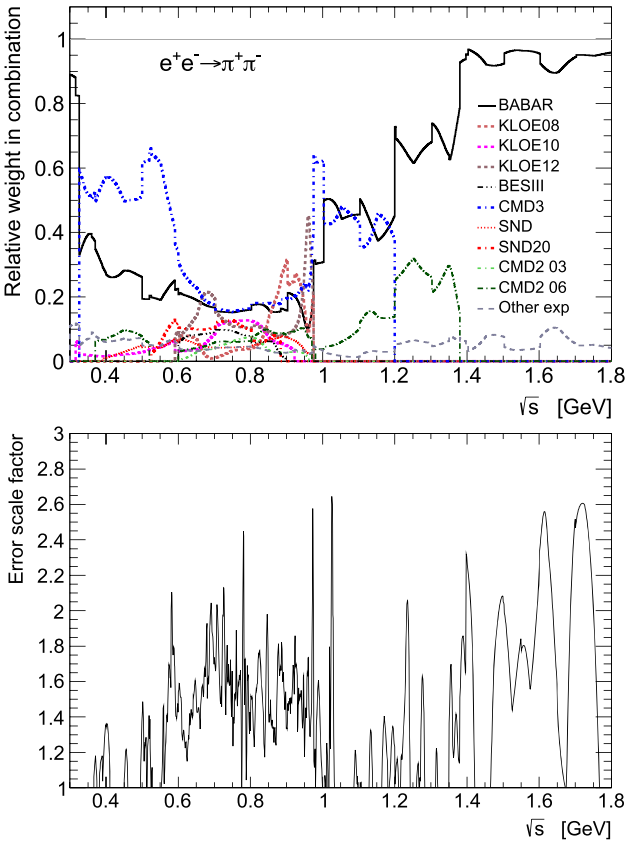

Fig. 4 Top: relative local weight per measurement contributing to the $e^+e^- \to \pi^+\pi^-$ cross-section combination versus centre-of-mass energy. Bottom: local uncertainty scale factor versus centre-of-mass energy applied to the combined $\pi^+\pi^-$ cross-section uncertainty to account for inconsistencies among the measurements

first time the measurement of hard NNLO processes with two additional photons from either initial or final state radiation (FSR). The paper also includes comparisons with predictions from the NLO PHOKHARA and the partial NNLO AFKQED [59–62] Monte Carlo generators.

In the following we use the notation LO, NLO, NNLO to specify the true QED order defined with respect to the lowest-order ISR process, while the same symbols taken within quotes, 'LO', 'NLO', 'NNLO', refer to reconstructed topologies with various photon multiplicities. We summarise here the main findings of the BABAR study [46]:

- 'NNLO' contributions with additional photon energies above 200 MeV (100 MeV) for the most (least) energetic one (representing (1.9–3.8)% of the beam energy in the centre-of-mass frame) are observed in $(3.47 \pm 0.38)\%$ and $(3.36 \pm 0.39)\%$ of the dimuon and dipion events, respectively. These events are dominated by small-angle additional ISR photons.
- The 'NLO' event fractions, with one additional detected or kinematically reconstructed photon above 200 MeV, predicted by the PHOKHARA generator exceed the BABAR data, particularly for additional ISR photons at small angle. Over the full measured phase space, including additional ISR and FSR photons, PHOKHARA overpredicts the hard 'NLO' contribution by a factor of 1.25 ± 0.05.
- The BABAR cross-section measurements [34,35] are found to be insensitive to the missing NNLO contributions in (and hard 'NLO' excess of) the PHOKHARA generator.
- The AFKQED generator approximates real and virtual higher-order (HO) corrections by resumming the leading logarithms.[4] It provides a reasonable description of

[4] AFKQED generates the additional NLO (or NNLO) ISR photons in the collinear approximation using the structure function technique, which is not consistent with the 'NLO' angular distribution with respect to the beam axis observed in the BABAR data. This shortcoming is not present in the KKMC event generator [63], which uses a different technique for multi-photon emission, while maintaining the same advantage

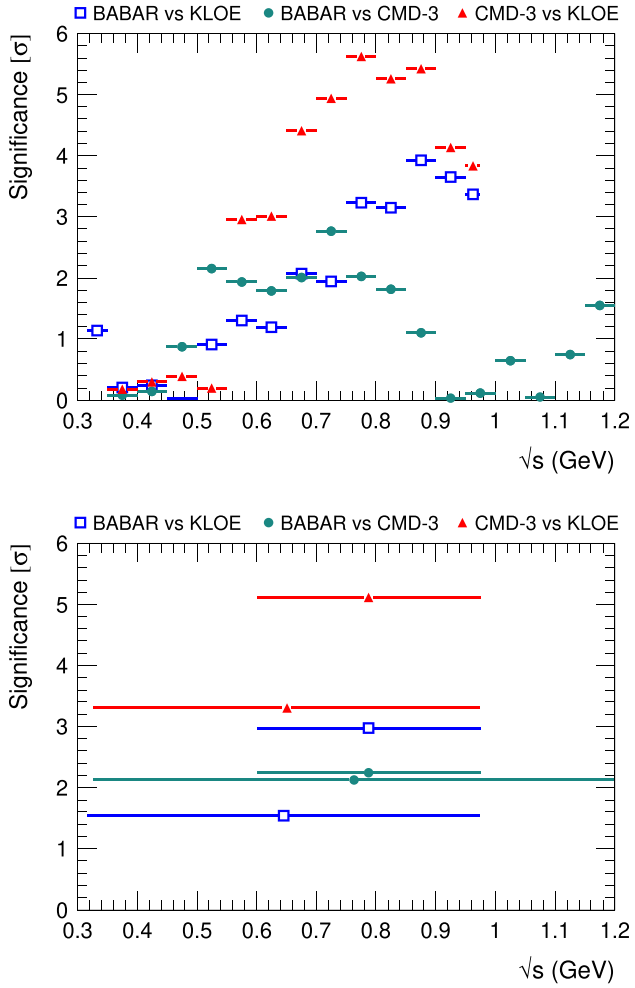

Fig. 5 Significance of the difference between pairs of the three most precise $e^+e^- \to \pi^+\pi^-$ experiments for narrow energy intervals of 50 MeV or less (top) and larger energy intervals (bottom) indicated by the horizontal lines

the rates and energy distributions of the measured 'NLO' and 'NNLO' topologies.

4 Cancellation between soft/virtual and hard photon corrections

4.1 Expected behaviour of NLO events and experimental procedures

To assess the effect of higher-order radiative corrections one needs to evaluate the sensitivity of a given measurement to the presence of additional photon radiation, particularly ISR. An analysis that rejects part of the additional ISR requires a compensating correction from an NLO Monte Carlo generator to be consistent, at that order, with the corresponding ISR luminosity computed with the same generator. The PHOKHARA event generator version 9.1 [64] incorporates all contributions from NLO QED and thus provides a complete prediction of the ISR process $e^+e^- \to \mu^+\mu^-\gamma(\gamma)$. It includes the lowest-order (LO) ISR and FSR processes and NLO contributions from real photon emission by the $e^\pm$ beams and the outgoing muons, as well as soft photon emission and virtual corrections. The sum of the soft and virtual terms is infrared finite and the transition energy between soft and hard emission is chosen within a safe range (5 MeV for BABAR simulations) so that both contributions are under control. From an experimental point of view, both LO and soft plus virtual NLO lead to event configurations that are reconstructed in the 'LO' topology and kinematics, whereas sufficiently hard NLO radiation necessitates a different kinematic treatment. The lowest energy for NLO photon contributions is experiment dependent. In BABAR a value of 50 MeV, the energy threshold for a detected photon included in kinematic fits, is representative, although a higher threshold (200 MeV) is applied to detected or kinematically reconstructed photons to separate the 'NLO' from the 'LO' topologies for the final results.

The effects of HO radiative corrections are evaluated using samples of ISR muon-pair events generated with PHOKHARA in the BABAR conditions: ISR (or FSR) photon at large polar angle (20°–160°) in the e^+e^- centre-of-mass (CM) system; two-charged-particle mass from threshold to 1.4 GeV; $\sqrt{s} = 10.58$ GeV CM energy. Soft and virtual corrections are studied with the use of samples generated at LO with either ISR only or with ISR and FSR, and samples generated at NLO with either ISR only or the full NLO configuration with ISR, FSR, and their interference. The fraction of hard photon radiation turns out to be rather large because NLO ISR is enhanced by a factor $\ln(s/m_e^2)$. It strongly depends on the energy threshold of the additional photon: a fraction of 60% for E_γ^* above 5 MeV in the centre-of-mass decreases to 38% above 50 MeV and to 25% above 200 MeV. All contributions are dominated by NLO ISR at small angle with respect to the beam axis. For example, with 50 MeV photon energy threshold the NLO ISR fraction at small angle outside the BABAR acceptance is 27%, NLO ISR at large angle 8%, and NLO FSR 3%. These values illustrate the importance of a thorough understanding and robust correction of effects from HO radiative corrections. The situation is very similar for the $e^+e^- \to \pi^+\pi^-\gamma(\gamma)$ ISR process.[5]

Footnote 4 continued
of leading-log resummation. We have verified that the 'LO', 'NLO', and 'NNLO' event fractions predicted by AfkQed and KKMC are in agreement with the data.

[5] For pions, additional radiation requires a model, usually assuming scalar QED with point-like pions and improved with dispersive methods [65].

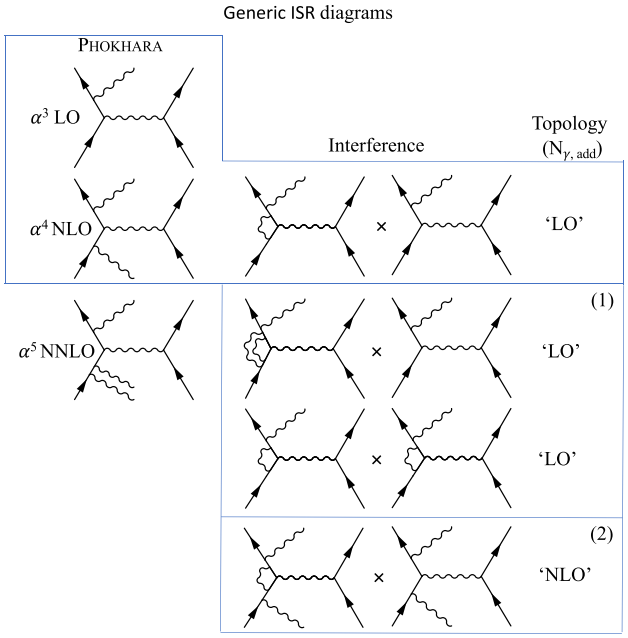

Fig. 6 Generic Feynman diagrams for the ISR $e^+e^- \to \mu\mu\gamma(\gamma)$ process at LO, NLO and NNLO. At each specified QED order indicated on the left, generic diagrams, ignoring specific topologies with different particle permutations, are drawn for the virtual (loop) and real photon emission processes. For each order, the latter processes are given on the left-hand side, while the interference contributions are specified on the right together with their experimental topology labeled with quotes. $N_{\gamma,\mathrm{add}}$ refers to the number of real photons emitted, beyond the main ISR photon: $N_{\gamma,\mathrm{add}} = 0$ for the 'LO' topology, $N_{\gamma,\mathrm{add}} = 1$ for the 'NLO' topology. In the case of NNLO, the two interference contributions labelled (1) and (2) lead to 'LO' and 'NLO' topologies, respectively

It is instructive to compare the PHOKHARA predictions at different orders. For the BABAR conditions the full NLO (LO) cross section for $e^+e^- \to \mu^+\mu^-\gamma(\gamma)$ amounts to 17.16 pb (17.45 pb), a reduction by -1.7% at NLO. Since the NLO cross-section contribution with an additional photon above 50 MeV corresponds to $38\% \times 17.16/17.45 \simeq 37\%$, it is almost compensated by a reduction of 39% due to the soft and virtual contribution. This large cancellation between hard and soft/virtual effects is well-known in QED [66,67]. It requires a careful assessment of the measured and theoretically corrected cross-section fractions.

4.2 Going from NLO to NNLO processes

At present there exists no complete NNLO calculation of the $e^+e^- \to \mu^+\mu^-\gamma(\gamma)(\gamma)$ process. A behaviour similar to NLO is expected, i.e., an overall small effect on the cross section, possibly at the level of a few per mil, and significantly larger contributions from hard radiation, which may affect the fiducial acceptance of the analyses.

The investigation of hard and soft/virtual radiative corrections at NNLO is more intricate than at NLO. The situation is illustrated in Fig. 6, which shows the relevant generic Feynman diagrams. For each order in QED, positive contributions with one to three real photons are separated from contributions from interfering amplitudes involving soft/virtual photons. The first two rows correspond to the diagrams considered in the NLO generator PHOKHARA. They illustrate the large cancellation occurring at this level as the result of the interference term within the 'LO' topology.

At NNLO, the cancellation occurs between the positive three real photon emission contribution and the generic interference contributions leading to an 'LO' topology, for the processes labelled (1), or to an 'NLO' topology for those in part (2). The interpretation of the results from the radiative BABAR study [46] depends on the relative importance of these two components. Two extreme scenarios may be considered:

- Scenario 1: the processes labelled (1) dominate the NNLO interference term. Since they fall into the 'LO' topology, the 'NLO' contribution is unaffected by NNLO and the large excess of events predicted by PHOKHARA compared to the data for the 'NLO' topology would need to be interpreted as a generator issue at NLO.
- Scenario 2: the processes labelled (2) are the dominant NNLO interference contribution. Being negative, it will affect the 'NLO' photon energy distribution in a way uncorrected by the NLO generator. In this situation, the observed deficit in data would arise from NNLO virtual contributions and PHOKHARA is safe.

The true situation is likely in-between these two extreme scenarios. Only complete NNLO calculations, at fixed order or in an event generator, will help resolve this ambiguity and should be a high priority for the field. Since the interference contribution listed in the second row of part (1) is obviously positive, it will tend to reduce the negative contribution originating from the first row, perhaps to an overall level smaller than part (2). Also, the interpretation of the BABAR results appears more natural in the second scenario as the NNLO contributions, real and virtual, would explain all the observed features without having to question the validity of the PHOKHARA generator at NLO.

5 Impact of higher-order radiative effects

The observation that PHOKHARA does not correctly predict the 'NLO' contribution raises potential issues for ISR experiments measuring only part of the cross section because of event selection criteria. The fiducial acceptance of an analysis is evaluated with a Monte Carlo generator interfaced with a simulation of the detector response. KLOE, BESIII and CLEOc [68] rely on PHOKHARA to estimate the unse-

lected 'NLO' part. Hard NNLO contributions are ignored. As explained in Sect. 4, a mis-evaluation of hard NLO and NNLO contributions is not compensated by soft/virtual contributions at the same order since the latter are included in the selection of lower-order-like events. This unbalance will generate a bias in the cross section measurement.

5.1 Procedure

It is not possible to accurately compute the bias without full knowledge of the respective analyses and associated detector performance. The purpose of the following study is limited to estimating the possible extent of the bias by reproducing the kinematic conditions of the published analyses with a simplified generic detector. The study is further restricted to two configurations: 'KLOE08' with small-angle undetected ISR [30] and 'BESIII' with large-angle measured ISR [51], the quotes indicating their generic nature. In both cases, Monte Carlo samples are generated with PHOKHARA in the kinematic conditions of the experiments with a fast simulation of the tracking and calorimeter performance. The 4-vectors from the event generator are converted into pseudo "reconstructed" data using the acceptances and resolutions of the detector, as found in the papers published by the experiment. Analysis steps are then applied to these pseudo data to reproduce the overall acceptance and efficiency taking into account the analysis cuts used by the experiment eg., fiducial acceptance in the $(M_{\rm trk},\ M_{\pi\pi})$ plane for KLOE or $\chi^2_{\rm LO}$ selection for BES-III.

Moreover, two assumptions are made: first, the hard NNLO fraction is taken from the BABAR measurements and assumed to hold independently of the experiment's CM energy. Secondly, real hard NNLO and soft/virtual NNLO radiative corrections are assumed to cancel in the cross section. In absence of a complete NNLO calculation, the effect of the observed hard NNLO contribution on the 'NLO' spectrum is not known. This ambiguity is related to the relative importance of parts (1) and (2) in Fig. 6, which will be approached by considering the extreme scenarios introduced above.

While scenario 2 can be readily transposed to any ISR experiment by estimating the effect of missing NNLO corrections, evaluating the impact of scenario 1 is more delicate without knowing the origin of the issue in the PHOKHARA generator. It is worth mentioning in this context that all tests documented in the PHOKHARA publications to evaluate the impact of NLO versus LO corrections relate to the integrated cross section as a function of the two-pion mass [69]. Albeit PHOKHARA was used by the experiments to evaluate the fiducial acceptance and efficiency of energy and angular selection requirements on additional ISR photons, the modelling accuracy was to our knowledge never tested. Experiments exploiting ISR measure the $M_{\pi\pi}$ spectrum of the selected $\pi^+\pi^-\gamma$ sample and correct it for acceptance and selection efficiencies to determine the Born-level $e^+e^- \to \pi^+\pi^-$ cross section

$$\sigma_{\pi\pi} = \frac{dN_{\pi\pi\gamma}}{dM_{\pi\pi}} \cdot \frac{s}{2M_{\pi\pi}H(M_{\pi\pi})\varepsilon_{\rm acc}\varepsilon_{\rm sel}L_{ee}}\,, \tag{1}$$

where s is the CM energy squared, $H(M_{\pi\pi})$ the ISR radiation function (radiator), and L_{ee} the e^+e^- luminosity. The acceptance $\varepsilon_{\rm acc}$, selection efficiency $\varepsilon_{\rm sel}$, and $H(M_{\pi\pi})$ are evaluated with PHOKHARA. Although we have studied NNLO effects for all three variables, results will only be reported for the selection efficiency which is affected most.

5.2 Generic 'KLOE08' configuration

In the experimental configuration with $\sqrt{s} = 1.02$ GeV, only the two charged particles are detected in a polar angle range between 50° and 130°. Their three-momenta are measured accurately, while the energy and polar angle of the putative ISR photon are calculated assuming LO kinematics for the ISR process. The ISR photon is required to be emitted in a dead cone of 15° around the beams. The common track mass $M_{\rm trk}$ of the two charged particles, computed under the LO assumption, allows to separate dimuon from dipion processes. The selection of $\pi^+\pi^-\gamma$ events is defined in the $(M_{\rm trk}, M_{\pi\pi})$ plane in a region avoiding background from $\phi \to \pi^+\pi^-\pi^0$ and the muon band [30]. Fast simulation follows the KLOE performance for charged particle reconstruction [70]. Since the selection is very sensitive to the NLO radiative tail, only events satisfying the acceptance cuts are considered here.

Despite the LO-like selection, half of the NLO events are automatically kept in the selected sample when the additional ISR photon is emitted in the same hemisphere as the primary ISR photon, reducing thereby the dependence of the selection efficiency on the PHOKHARA generator. This situation occurs since all ISR emissions are sharply peaked in the beam direction, resulting in a small invariant mass of the diphoton system consistent with the zero-mass assumption in the $M_{\rm trk}$ calculation. However, when the additional photon is emitted along the opposite beam direction, the diphoton mass can be relatively large, introducing a long tail in the $M_{\rm trk}$ distribution. In that case, the selection efficiency depends on the validity of the photon distribution predicted by PHOKHARA. The simulated event distributions within the fiducial acceptance in the $(M_{\rm trk}, M_{\pi\pi})$ plane for the same-side and opposite-side samples are shown for all events (top) and the selected events (bottom) in Fig. 7. Figure 8 shows the corresponding $M_{\rm trk}$ distributions integrated over $0.6 < M_{\pi\pi} < 0.95$ GeV, for events with additional photon energies larger than 10 MeV. Here the opposite side

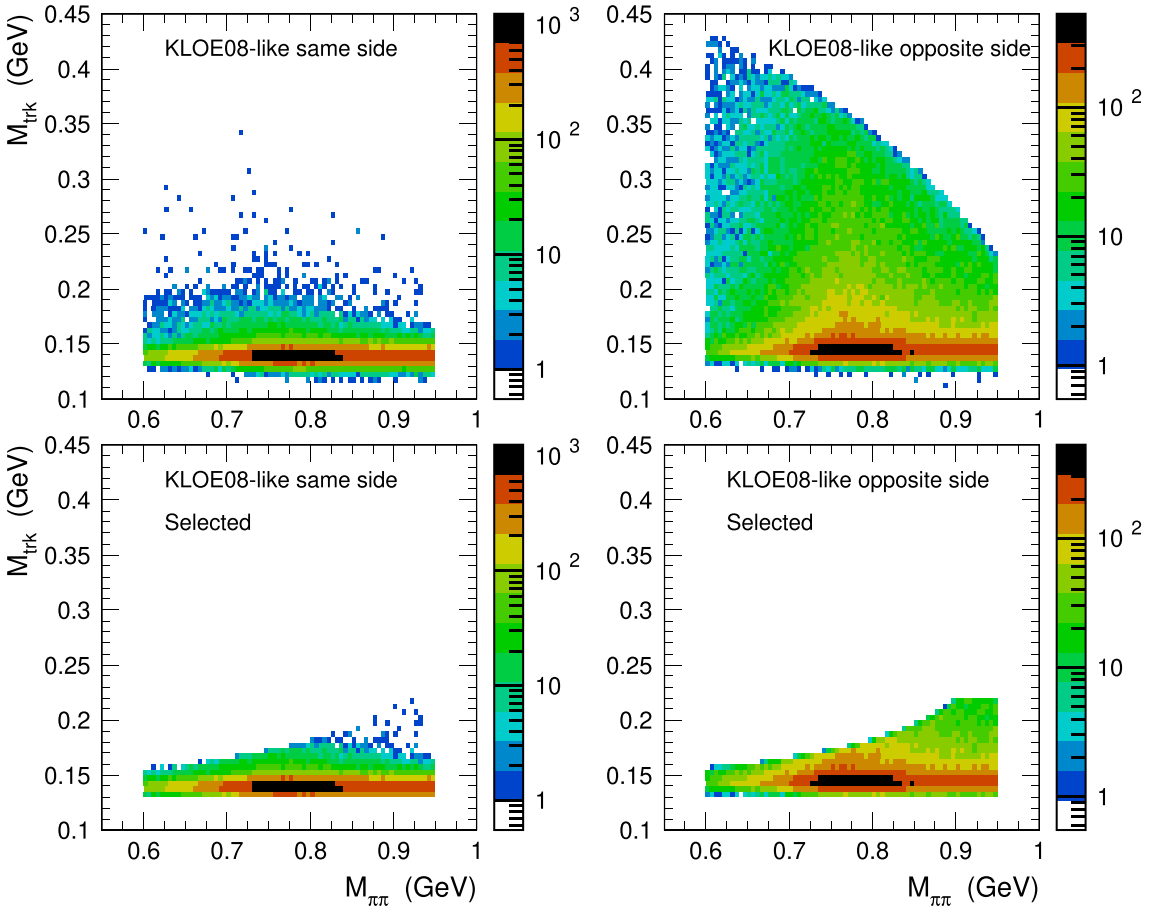

Fig. 7 Distributions of simulated $e^+e^- \to \pi^+\pi^-\gamma(\gamma)$ events within the fiducial acceptance for the 'KLOE08' configuration in the $(M_{\rm trk}, M_{\pi\pi})$ plane. Top: events with an additional NLO photon with energy larger than 5 MeV emitted in the same hemisphere as the ISR photon (left) and in the opposite hemisphere (right). Bottom: the corresponding distributions after applying selection requirements

configuration leads to $M_{\rm trk}$ values above the pion peak, which are selected with an average efficiency $\varepsilon_{\rm sel}^{\rm oppo}$ of only 25%.

To estimate the effect of missing hard NNLO radiation in PHOKHARA, the fraction of $(3.5\pm0.4)\%$ observed by BABAR is assumed, as the relative thresholds for the additional photon, 100 MeV in BABAR versus 10 MeV in KLOE, scaled by the beam energies, are comparable.[6] Taking further the NNLO contribution as a perturbation of the much larger hard NLO component, the selection efficiency is assumed to be unaffected. Following the previous discussion, three out of four configurations feature the emission of at least one of the two additional photons opposite to the ISR photon and thus contribute to the radiative tail of the $M_{\rm trk}$ distribution. Averaging over the KLOE08 mass range (dominated by the larger statistics at high mass), the resulting cross section change from the reduced selection efficiency amounts to roughly $-3.5 \cdot 3/4(1-\varepsilon_{\rm sel}^{\rm oppo})\% = -2.0\%$. The $M_{\pi\pi}$ dependence is small across the ρ mass region, with a value of -2.3% at the peak.

In scenario 1 the NLO excess in PHOKHARA is assumed to be a generator issue. Were the NLO fractional excess at the same level as that observed by BABAR, the resulting effect on the selection efficiency would partially cancel the bias from missing NNLO radiation, with a residual effect of order -1% at the ρ peak.

In scenario 2, the use of PHOKHARA is safe and the only bias originates from missing NNLO corrections. Hard NNLO radiation contributes as in scenario 1, but its effect is reduced by the negative interference contributions with an 'NLO' topology (cf. part (2) in Fig. 6), of which only one half with opposite-side radiation contributes to the $M_{\rm trk}$ radiative tail. The resulting cross section change is estimated to be $-3.5 \cdot (3/4-1/2)(1-\varepsilon_{\rm sel}^{\rm oppo})\% = -0.7\%$ for the average over $M_{\pi\pi}$, and -0.8% at the ρ peak.

Both scenarios lead to cross section changes that exceed the 0.5% uncertainty assigned by KLOE08 [30] to radiative corrections.

[6] This scaling is approximate because of the phase-space reduction due to the non-negligible muon mass at the low centre-of-mass energy for KLOE.

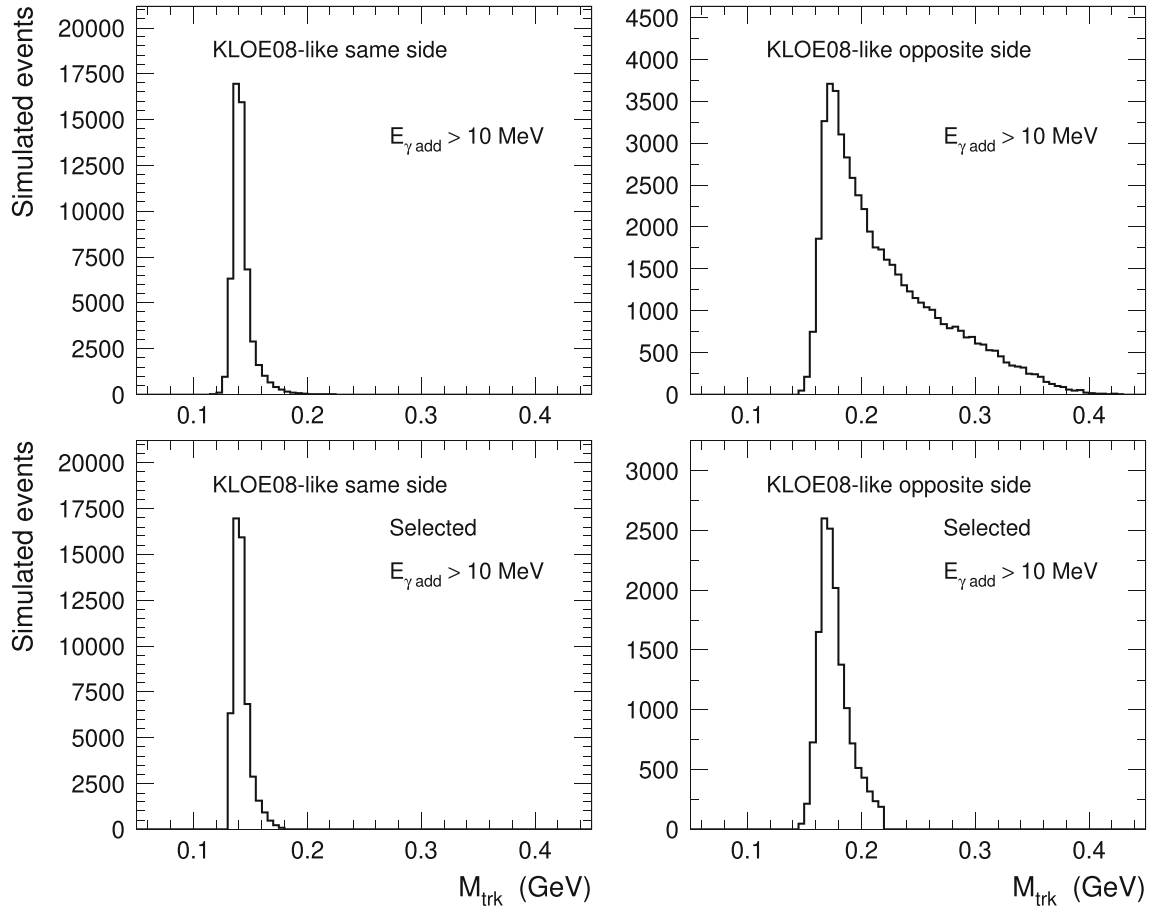

Fig. 8 Distributions of $M_{\rm trk}$ for simulated $e^+e^- \to \pi^+\pi^-\gamma(\gamma)$ events within the fiducial acceptance for the 'KLOE08' configuration. Top: events with an additional NLO photon with energy larger than 10 MeV emitted in the same hemisphere as the ISR photon (left) and in the opposite hemisphere (right). Bottom: the corresponding distributions after applying selection requirements

5.3 Other KLOE measurements

The 'KLOE10' configuration with the ISR photon detected at large angle and the two pions in the same range as 'KLOE08' may be treated in a similar way. Because additional ISR photons predominantly emitted along the beams are well separated from the detected ISR photon, one expects both same and opposite sides to contribute to the $M_{\rm trk}$ radiative tail. The cross-section change in scenario 1 is therefore expected to be larger than for the 'KLOE08' configuration. In scenario 2, however, since the NNLO positive real photon and negative virtual/soft interference contributions approximately cancel, and lead to photon topologies in the rejected radiative tail, there is no bias for KLOE10.

In the KLOE12 measurement [32] the cross section was directly obtained from the ratio of the $\pi^+\pi^-\gamma$ to $\mu^+\mu^-\gamma$ mass spectra, protecting the result against modelling biases. However, in practice, the protection is incomplete as pion and muon selection requirements differ in the $(M_{\rm trk}, M_{\pi\pi/\mu\mu})$ plane. While part of the pion radiative tail is retained, it is almost entirely removed in the selected muon sample by a tight $80 < M_{\rm trk} < 115$ MeV requirement applied to reduce the pion background. This asymmetry in the selection of the pion and muon samples reintroduces a modelling dependence.

KLOE12 features a comparison of the measured muon ISR cross section with the QED NLO prediction by PHOKHARA. The results show agreement within the quoted systematic uncertainty of 1%, which is however insufficient to validate the 0.5% uncertainty assigned to radiative corrections in the two-pion cross-section measurement. A newer dimuon study based on a much larger data set does not improve in precision [71]. We may proceed as in the pion case to estimate the effect of missing NNLO corrections in PHOKHARA. The cross section change under scenario 1 is found to be of order -2.6%, while the increase from a potential PHOKHARA NLO excess cannot be estimated. As for the pions, an NLO excess as the one observed in BABAR would essentially cancel that cross-section change. In scenario 2, the muon cross-section change is found to be reduced, as for the

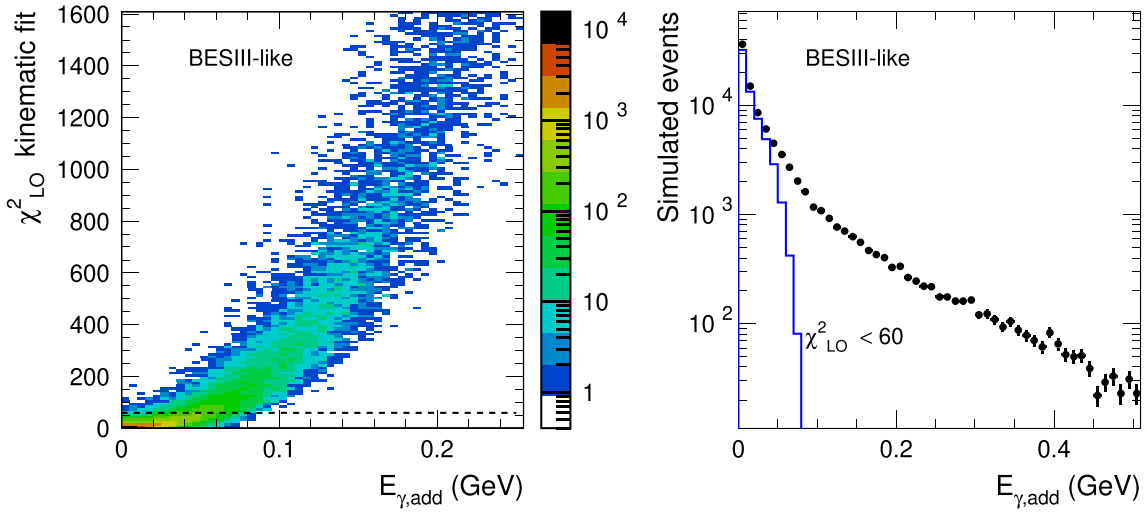

Fig. 9 Distributions of simulated $e^+e^- \to \pi^+\pi^-\gamma(\gamma)$ events in the 'BESIII' configuration. Left: $\chi^2_{\rm LO}$ of the $\pi^+\pi^-\gamma$ kinematic fit as a function of the true energy of the NLO additional photon. The dashed horizontal line indicates the $\chi^2_{\rm LO} < 60$ selection requirement. Right: the energy spectrum of the additional NLO photon with (blue line) and without (dots) applying the compatibility requirement

pions, by a factor 1/3 to −0.9%. Such a bias would amount to twice the quoted radiative correction uncertainty, but would not be detectable given the 1% systematic uncertainty of the test.

5.4 Generic 'BESIII' configuration

BESIII reported ISR based $e^+e^- \to \pi^+\pi^-\gamma$ cross-section results [51] using data taken at $\sqrt{s} = 3.773$ GeV, a factor of three below (above) the BABAR (KLOE) CM energy. The analysis requires detection of the two pions and a large-angle ISR photon, while additional photons are ignored. A kinematic fit using the $\pi^+\pi^-\gamma$ hypothesis selects LO and NLO soft/virtual events with the requirement $\chi^2_{\rm LO} < 60$. A fast simulation of the 'BESIII' configuration and detector performance [72], using the PHOKHARA generator and the same assumptions as in the 'KLOE08' study, allows to investigate the effects of additional photon radiation.

Figure 9 (left) shows the distribution of $\chi^2_{\rm LO}$ as a function of the true NLO additional photon energy $E_{\gamma,\rm add}$, exhibiting a strong correlation. Events with about $E_{\gamma,\rm add} > 50$ MeV are subject to rejection. As this maximum accepted $E_{\gamma,\rm add}$ is consistent with the BABAR threshold of 100–200 MeV for NLO/NNLO photons when normalized to the respective beam energies, one expects very low selection efficiencies of NLO/NNLO events in BESIII. The distributions of $E_{\gamma,\rm add}$ for all radiative events and after the $\chi^2_{\rm LO} < 60$ selection is shown on the right panel of Fig. 9. The fraction of rejected events with $E_{\gamma,\rm add} > 50$ MeV is 92%.

The fractional cross-section change due to missing NNLO in PHOKHARA in scenario 1 amounts to approximately $-3.5 \cdot 0.92\% = -3.2\%$, again significantly exceeding the assigned systematic uncertainty of 0.5%. As in the case of KLOE, this large effect might be partially cancelled by an NLO excess in PHOKHARA under scenario 1 that we are unable to propagate to CM energies lower than BABAR.

Similarly to KLOE10, the tight 'LO' selected topology preserves BESIII from any bias under scenario 2 as the NNLO positive and negative contributions approximately cancel in the rejected radiative tail.

5.5 Additional remarks

The quantitative effects of higher order radiative corrections on the KLOE and BESIII two-pion cross-section results estimated here cannot be taken at face value. Rather, they indicate the potential size of systematic effects encountered from the use of PHOKHARA in view of the findings reported by BABAR [46]. According to our study, effects from neglected NNLO contributions may suggest upward cross-section corrections that exceed the quoted systematic uncertainties, potentially reducing the difference seen with BABAR. The concomitant effect of the hard NLO excess in PHOKHARA [46] is more speculative and may depend on CM energy. Investigations by the generator authors should allow to shed light on this issue [73]. Any definitive assessment needs to be carried out by the KLOE and BESIII collaborations with the full machinery of their analyses.

In this context we also performed a test comparing dimuon samples generated with PHOKHARA and KKMC [63] in the 'KLOE08' configuration. Differences in the energy distributions of additional photons at 'NLO' level lead to different acceptance predictions among the two generators. By construction, KKMC produces higher photon multiplicities, but

it predicts the fraction of three or more photons above 10 MeV at KLOE energies to be 1.3%, which is lower than the corresponding 'NNLO' rate found by BABAR. Of course the two generators operate in different ways. Whereas PHOKHARA is designed for ISR with an NLO matrix element, KKMC works from the Born level up with multiple ISR photon emission approximating higher orders. It is beyond the scope of this paper to conduct a detailed evaluation of these generators, but we note their different predictions. Contrary to our above estimates for PHOKHARA, KKMC would predict a downward shift of the measured cross sections, albeit again larger than the quoted systematic uncertainty assigned to radiative corrections.

In a recent paper [74], Belle II confirms the BABAR finding of an 'NLO' excess by PHOKHARA with respect to their $\pi^+\pi^-\pi^0$ data. To account for this excess and the missing NNLO contributions, they assign a 1.2% systematic error for the generator. It is critical that future (re-)analyses of ISR based cross section measurements perform data-driven tests of the kinematic properties of additional photons as Belle II has done. Such tests allow to investigate the sensitivity to mismodelling and higher order radiative effects, and help design robust selection criteria. The loose selection used by BABAR could be implemented rather straightforwardly in the BESIII analysis since the setup follows the same topology with a large-angle ISR photon, and the detector allows the measurement of large-angle additional photons. The situation is more complicated for KLOE as the selection method, at least in the small-angle ISR topology used in KLOE08 and KLOE12, lacks kinematic constraints, preventing the reconstruction of additional small-angle photons. Such an approach would be possible in the KLOE10 topology, but would require independent charged particle identification. Another difficulty for KLOE arises from the low centre-of-mass energy and the proximity of the ρ resonance leading to low ISR photon energies that are not as well separated from additional photons as in the case of BESIII and BABAR. This also presents an obstacle to the experimental separation of additional large-angle ISR and FSR photons as was done by BABAR and would be possible with BESIII, as seen from Fig. 10.

6 Reappraisal of τ spectral functions

Spectral functions derived from measurements of mass spectra in hadronic τ decays provide a complementary input, under isospin symmetry and accounting for isospin-breaking corrections, to compute HVP integrals [75]. In the late 1990 s, thanks to LEP experiments (particularly ALEPH), τ spectral functions in the two-pion channel were more precise than the available e^+e^- cross sections. In the following decade, both τ and e^+e^- data were therefore used by our group [49,76–81].

The τ spectral function $v_{\pi^-\pi^0}(s)$ in the $\pi^-\pi^0$ channel is defined by

$$v_{\pi^-\pi^0}(s) = \frac{m_\tau^2}{6\,|V_{ud}|^2}\,\frac{B_{\pi^-\pi^0}}{B_e}\,\frac{1}{N_{\pi^-\pi^0}}\,\frac{dN_{\pi^-\pi^0}}{ds} \times \left(1-\frac{s}{m_\tau^2}\right)^{-2}\left(1+\frac{2s}{m_\tau^2}\right)^{-1}\frac{R_{\rm IB}(s)}{S_{\rm EW}}, \tag{2}$$

with

$$R_{\rm IB}(s) = \frac{{\rm FSR}(s)}{G_{\rm EM}(s)}\,\frac{\beta_0^3(s)}{\beta_-^3(s)}\left|\frac{F_0(s)}{F_-(s)}\right|^2, \tag{3}$$

and where $(1/N_{\pi^-\pi^0})dN_{\pi^-\pi^0}/ds$ is the normalised invariant mass-squared (s) spectrum of the $\pi^-\pi^0$ final state obtained from the combination of spectra from several experiments, $B_{\pi^-\pi^0}$ (B_e) are the corresponding τ branching fractions (final state photon radiation implied), and $S_{\rm EW}$ is an electroweak radiative correction. The s-dependent isospin-breaking (IB) corrections are included in $R_{\rm IB}(s)$. In Eq. (3), $\beta_{0,-}$ denote the pion velocities in the two-pion CM system for the $\pi^+\pi^-$ and $\pi^-\pi^0$ final states, respectively. $G_{\rm EM}(s)$ is the radiative function, correcting from the $\pi^-\pi^0(\gamma)$ to the $\pi^+\pi^-$ final states, requiring the addition of the specific FSR contribution to the neutral case. Several model-dependent approaches exist for the small long-distance radiative correction $G_{\rm EM}(s)$. The pioneering work of Cirigliano-Ecker-Neufeld [82,83] used Chiral Perturbation Theory (ChPT), while vector dominance was the basis of further work by Lopez Castro et al. [84,85]. The two methods have been known to be in good agreement. More recently, other studies extended the order in ChPT while satisfying short-distance constraints [86]. Additional free parameters, however, deteriorate the precision of the prediction. In the longer term, lattice QCD based estimates are expected to become available [87] and will provide an important cross check. The form factor ratio F_0/F_- takes into account the different masses and widths of the charged and neutral ρ mesons and the $\rho-\omega$ interference only present in the neutral final state.

The idea of significant $\rho-\gamma$ mixing, motivated by the well-founded $Z-\gamma$ mixing in high-energy e^+e^- collisions, was put forward by Jegerlehner and Szafron [88] and introduced large IB corrections on top of what had been previously estimated. However, a justification for applying the same $Z-\gamma$ formalism to the composite ρ meson was never given. It was exacerbated by the proposal by Jegerlehner to reverse the correction and apply it to e^+e^- rather than τ data [89]. In a consistent dispersive approach of the pion form factor there is no room for $\rho-\gamma$ mixing as differences between charged and neutral ρ line shapes are embedded in their respective resonance parameters (mass, width) [90]. The consideration of $\rho-\gamma$ mixing is therefore dropped.

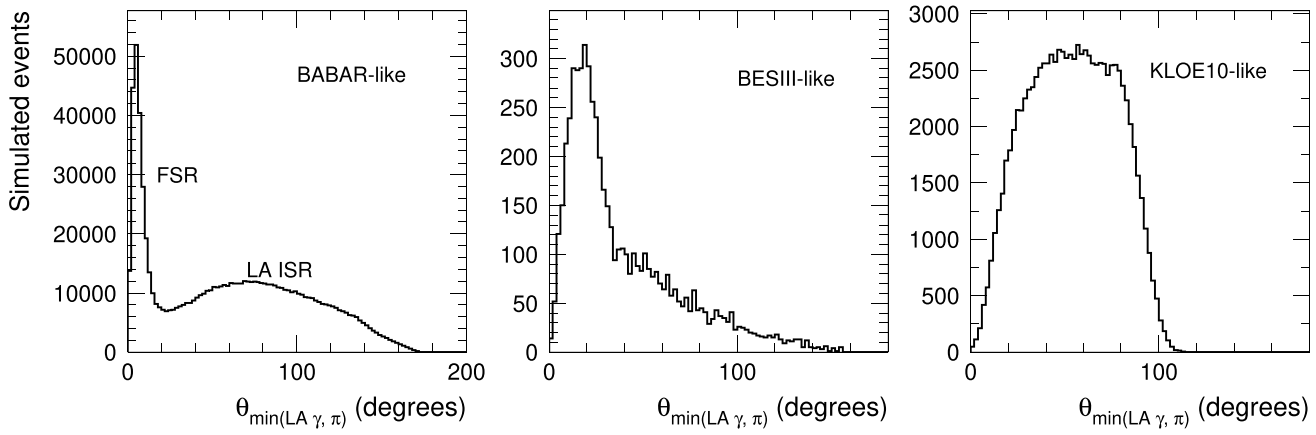

Fig. 10 The minimum angle between the additional large-angle (LA) photon and the two pions within the detector acceptance for simulated $e^+e^- \to \pi^+\pi^-\gamma(\gamma)$ events in the BABAR, BESIII, and KLOE10 conditions (left to right panels). The separation between FSR and LA ISR events is pronounced at high CM energy (BABAR), still visible at intermediate CM energy (BESIII), and vanishes at low CM energy (KLOE)

The use of τ spectral functions was at some point discontinued owing to the improved e^+e^- cross-section data from KLOE and BABAR not requiring IB corrections. Given the discrepancies among the e^+e^- data sets and the progress on the understanding of IB corrections, we reconsider them here and present an update of the 2π HVP contribution to the muon $g-2$ from τ decays. The combined τ mass spectrum, after an update of the ALEPH data, is unchanged from Ref. [81]. A small change is introduced by updated IB corrections, essentially the ρ–ω contribution. The parameters used in Eq. (2) are $m_\tau = (1776.84 \pm 0.17)$ MeV, the CKM matrix element $|V_{ud}| = 0.97418 \pm 0.00019$, and $B_e = (17.818 \pm 0.032)\%$. Short-distance electroweak radiative effects [91–94], relevant for the $\pi\pi$ decay give $S_{\rm EW} = 1.0235 \pm 0.0003$ [78].

Most corrections to the τ-based 2π contribution to a_μ are unchanged from our previous work [49,81]. They amount to (all in 10^{-10} units): -12.21 ± 0.15 from $S_{\rm EW}$, -1.92 ± 0.90 from $G_{\rm EM}$, $+4.67 \pm 0.47$ from FSR, -7.88 from $m_{\pi^-} - m_{\pi^0}$ in the cross section, $+4.09$ from $m_{\pi^-} - m_{\pi^0}$ in Γ_ρ, $+0.20^{+0.27}_{-0.19}$ from $m_{\rho^-} - m_{\rho^0}$, -5.91 ± 0.59 from $\pi\pi\gamma$ and other electromagnetic ρ decays. The last four corrections are affected by a systematic uncertainty from the choice of the analytic model for the ρ lineshape, which we estimate from the difference between the Gounaris-Sakurai and Kühn-Santamaria resonance parameterisations and add linearly.

Due to its fast bipolar dependence on mass the contribution of ρ–ω interference to the dispersion integral is relatively small. It depends on the ω mass, the mixing amplitude $\varepsilon_{\rho\omega}$ and its phase $\phi_{\rho\omega}$, all determined from fits to the pion form factor in e^+e^- data. The value for $\phi_{\rho\omega}$ used in our previous analyses [49,81] was unexpectedly large [95]. Here, we use updated results from a fit to the combined e^+e^- data before CMD-3 [57,96] giving $m_\omega = 782.07 \pm 0.15$ MeV, $\varepsilon_{\rho\omega} = (1.99 \pm 0.03) \times 10^{-3}$, and $\phi_{\rho\omega} = (3.8 \pm 1.8)^\circ$. Including CMD-3 [58,96] gives similar results with the full difference added as systematic uncertainty. The resulting IB correction from ρ–ω mixing is $+(4.0 \pm 0.4) \times 10^{-10}$.

Summing up all the effects, the total IB correction to the τ-based 2π contribution is estimated to be $-(14.9 \pm 1.9) \times 10^{-10}$ to be compared to our previous estimate of $-(16.1 \pm 1.9) \times 10^{-10}$ [49,81]. Finally the contribution to a_μ from the combined τ data reads

$$a_\mu^\tau[2\pi] = (517.3 \pm 1.9 \pm 2.2 \pm 1.9) \times 10^{-10}\,, \qquad (4)$$

where the uncertainties are from the combined mass spectrum, the branching fractions, and the IB corrections, respectively.

The result (4) differs from that obtained in Ref. [86], $(519.6 \pm 2.8[\exp]^{+1.9}_{-2.1}[\rm IB]) \times 10^{-10}$ using $\mathcal{O}(p^4)$ ChPT. Most of the difference is accounted for by their $S_{\rm EW}$ value (1.0201), which does not take into account double counting between $S_{\rm EW}$ and $G_{\rm EM}$ for the subleading non-logarithmic short-distance correction for quarks. This effect is responsible for a shift of 1.7×10^{-10} in $a_\mu^\tau[2\pi]$. The remaining difference[7] (0.6×10^{-10}) originates mostly from the ρ width corrections in the pion form factor.

7 A new perspective on the muon $g-2$ HVP contribution from the dispersive method

Having discussed the tensions among the $e^+e^- \to \pi^+\pi^-$ cross-section measurements and their possible origins, and

[7] Larger differences are seen when comparing results from individual experiments.

reappraised the use of the complementary τ spectral functions, we proceed with a quantitative study of the dominant HVP contributions to a_μ. We consider here only the most precise results. We do not include the CMD-2 measurements [52,55], whose discrepancy with CMD-3 is currently under investigation [97], and the SND results, which are in a state of flux from the older [56] to the new measurements [50] that are still being updated [98].

For the following exercise, we consider the LO HVP contributions from the $\pi^+\pi^-$ channel in the wide mass range from threshold to 1.8 GeV for each experiment. BABAR and the τ spectral functions extend over the entire interval, while the other experiments cover a more restricted range and are completed near threshold and at large mass with the combination discussed in Sect. 2. For KLOE, we use the original combined data from Ref. [33] and consider two cases: the full available range and a restricted range of 0.6–0.975 GeV, where the data are most precise and KLOE's weight in the combination is largest (cf. top panel of Fig. 4). The two-pion contributions are complemented by the remaining LO HVP, NLO and NNLO HVP, hadronic light-by-light, as well as QED and electroweak contributions, all taken from Ref. [3]. The differences in the resulting a_μ predictions therefore reflect the differences in the two-pion contributions from each experiment, whose uncertainties correspond to the original ones, that is without rescaling to accommodate inconsistencies among data sets.

The results are shown in Fig. 11 as differences between the a_μ predictions and experiment [2]. The uncertainties drawn are from the $\pi^+\pi^-$ measurements (inner bars) and the total contributions (outer bars). The quoted uncertainties are separated into the $\pi^+\pi^-$ and remaining non-$\pi^+\pi^-$ contributions.

The BABAR and τ based results are in agreement. Combining both with CMD-3 gives $a_\mu^{\text{had, LO}} = (7057 \pm 33 \pm 22) \times 10^{-11}$, where the first uncertainty is from the $\pi^+\pi^-$ contribution, scaled by a factor 1.5 according to the χ^2 value of 4.5 for 2 degrees of freedom and the second from the non-$\pi^+\pi^-$ contribution. This average results into $\Delta a_\mu = a_\mu^{\text{SM}} - a_\mu^{\text{exp}} = -(123 \pm 33 \pm 29 \pm 22) \times 10^{-11}$, where the first uncertainty is from the $\pi^+\pi^-$ contribution, the second from all the other terms in the a_μ prediction, and the third from the $g-2$ experimental world average [2]. The significance of a non-zero Δa_μ is 2.5σ. As expected from the known tensions, the a_μ value for KLOE in the restricted range lies well below (3.8σ) the above combination.

The BABAR, τ, CMD-3 combination agrees with the only result available so far from lattice QCD for the full a_μ prediction, BMW [38], who find $\Delta a_\mu = -(105 \pm 55 \pm 22) \times 10^{-11}$, shedding a new light on the apparent discrepancy between BMW and the dispersive approach. Combining the values of BABAR, τ, CMD-3 and BMW, the difference with experiment is 2.8σ.

In the light of these results, we extend the study to the intermediate window 0.4–1.0 fm in Euclidean time, which

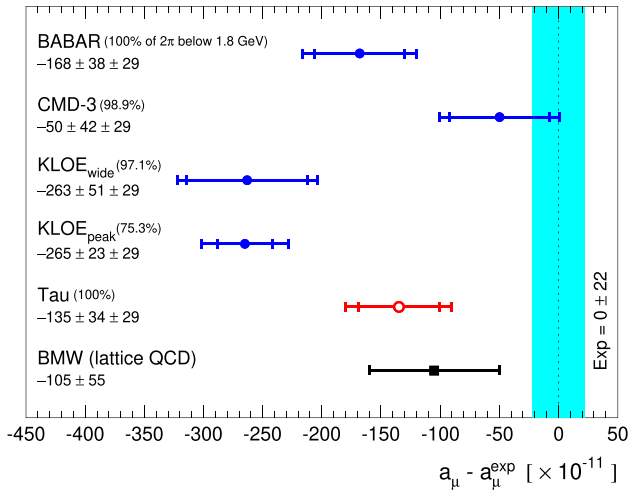

Fig. 11 Compilation of a_μ predictions subtracted by the central value of the experimental world average [2]. The predictions are computed from the individual $\pi^+\pi^-$ contributions between threshold and 1.8 GeV, complemented by common non-$\pi^+\pi^-$ contributions taken from Ref. [3] (circles). The quoted uncertainties correspond to the two contributions and do not include that of the subtracted experimental value shown by the vertical band. The error bars indicate the $\pi^+\pi^-$ and total uncertainties, respectively. The percentage given for each experiment represents the fraction of $a_\mu[\pi^+\pi^-, \text{threshold}-1.8\ \text{GeV}]$ used from a given experiment (see text for details, particularly concerning the two values for KLOE). The lattice result from BMW [38] is shown as filled square

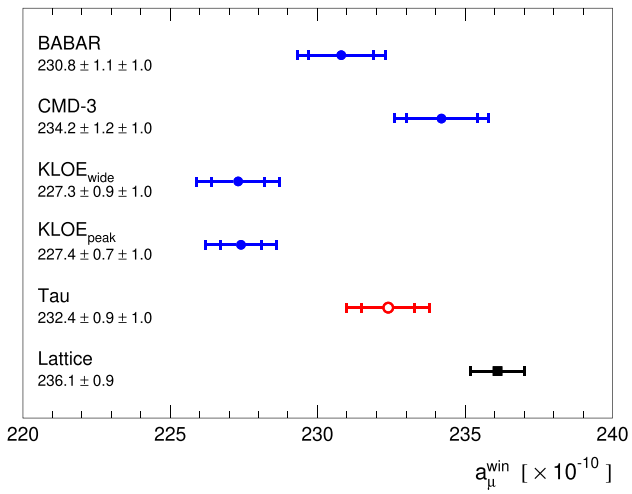

Fig. 12 Compilation of LO HVP a_μ^{win} predictions in the intermediate Euclidean time window (0.4–1.0 fm) [39], computed from the individual $\pi^+\pi^-$ measurements between threshold and 1.8 GeV (when only part of this interval is available it is extended to the full range using Ref. [8]), complemented by non-$\pi^+\pi^-$ combined spectra taken from Ref. [8]. Also shown is the average of the available lattice QCD results [45]

is favourable for lattice QCD. The corresponding $a_\mu^{\rm win}$ values are displayed in Fig. 12, where the quoted uncertainties are again separated into $\pi^+\pi^-$ and non-$\pi^+\pi^-$ contributions, the latter contribution using the combined spectra from Ref. [8].[8] All dispersive predictions are found below that from lattice QCD with significance of 1.1σ for CMD-3, 2.5σ for τ, 3.1σ for BABAR, 5.4σ for full KLOE, and 5.8σ for restricted-range KLOE, exacerbating the pattern seen for a_μ. The weighted average of BABAR, CMD-3 and τ gives 232.0 ± 1.1, to be compared with 236.1 ± 0.9 from lattice QCD. To further understand the discrepancy, additional lattice QCD studies, splitting the range of the lattice window into smaller intervals, possibly around the present optimal window, could be helpful.[9]

8 Conclusions

This paper reviewed existing tensions among the most precise $e^+e^- \to \pi^+\pi^-$ cross-section measurements used in the dispersive evaluation of the hadronic vacuum polarization (HVP) contribution to the anomalous magnetic moment of the muon. Local discrepancies between KLOE on one hand and BABAR and CMD-3 on the other hand exceed significances of 3σ and 5σ, respectively, while that between BABAR and CMD-3 is generally at the 2σ level. CMD-3 data lie systematically above all other data, while KLOE data lie below.

A dedicated analysis of radiative processes in $e^+e^- \to \mu^+\mu^-\gamma$ and $e^+e^- \to \pi^+\pi^-\gamma$ at NLO and NNLO by BABAR [46] prompted a study of related systematic uncertainties in the measurements using initial state photon radiation. In absence of an NNLO Monte Carlo generator the studies relied on approximate assumptions and fast simulation. They indicate potential problems for radiative event acceptances in the KLOE and BESIII measurements, not covered by the quoted systematic uncertainties.

In view of these difficulties with e^+e^- results we reappraised the use of τ hadronic spectral functions in the dispersive approach with an updated treatment of isospin-breaking corrections. The τ-based HVP contribution comes out close to the larger values provided by BABAR and CMD-3.

We reevaluated the compatibility of the dispersive HVP calculations with lattice QCD and with the $g-2$ experiment. Combining BABAR, τ, and CMD-3 measurements for the $e^+e^- \to \pi^+\pi^-$ HVP contribution, and adding all other contributions, the dispersive calculation of a_μ agrees with the lattice QCD result from BMW [38], while a discrepancy in the restricted observable $a_\mu^{\rm win}$ persists.

The discrepancy of the dispersive prediction with the $g-2$ experimental world average reduces from more than 5σ when KLOE measurements are included but neither CMD-3 nor τ data, as in [3], to the new prediction of 2.5σ when CMD-3 and τ measurements are included but not KLOE.

Acknowledgements We wish to thank Vincenzo Cirigliano, Gilberto Colangelo, Henryk Czyż, Laurent Lellouch, Gabriel Lopez Castro, and Pablo Roig for discussions that were very helpful to clarify various theoretical aspects. We thank Zbigniew Wąs for advice on the KKMC generator, Swagato Banerjee for providing an operational version of the code and Xiaowen Su for generating samples. This work benefited from funding by the French National Research Agency under contract ANR-22-CE31-0011.

Data Availability Statement This manuscript has no associated data. [Authors' comment: This is a theoretical study and no experimental data is produced beyond already published data.]

Code Availability Statement The manuscript has no associated code/software. [Author's comment: The code used to combine published experimental data is not public and available from the authors on reasonable request.]

Funded by SCOAP[3].

References

1. B. Abi et al., [Muon g-2], Phys. Rev. Lett. **126**(14), 141801 (2021). https://doi.org/10.1103/PhysRevLett.126.141801. arXiv:2104.03281 [hep-ex]
2. D.P. Aguillard et al., [Muon g-2], Phys. Rev. Lett. **131**(16), 161802 (2023). https://doi.org/10.1103/PhysRevLett.131.161802. arXiv:2308.06230 [hep-ex]
3. T. Aoyama et al., Phys. Rep. **887**, 1–166 (2020). https://doi.org/10.1016/j.physrep.2020.07.006. arXiv:2006.04822 [hep-ph]
4. M. Davier, A. Hoecker, B. Malaescu, Z. Zhang, Eur. Phys. J. C **77**(12), 827 (2017). https://doi.org/10.1140/epjc/s10052-017-5161-6. arXiv:1706.09436 [hep-ph]
5. A. Keshavarzi, D. Nomura, T. Teubner, Phys. Rev. D **97**(11), 114025 (2018). https://doi.org/10.1103/PhysRevD.97.114025. arXiv:1802.02995 [hep-ph]

[8] The τ based $a_\mu^{\rm win}$ result differs strongly from those given in the first versions of Ref. [99], particularly when using a non-$\pi^+\pi^-$ contribution derived from the full spectrum. Because of the different weighting of the mass spectrum, the $\pi^+\pi^-$ fraction of the HVP contribution is different for the window (0.64) and the full range (0.73), thus invalidating their procedure. This inconsistency has been corrected in the published version of Ref. [99] now in fair agreement with our result.

[9] See also the discussion on the limitations of such splitting in Ref. [45].

6. G. Colangelo, M. Hoferichter, P. Stoffer, JHEP **02**, 006 (2019). https://doi.org/10.1007/JHEP02(2019)006. arXiv:1810.00007 [hep-ph]
7. M. Hoferichter, B.L. Hoid, B. Kubis, JHEP **08**, 137 (2019). https://doi.org/10.1007/JHEP08(2019)137. arXiv:1907.01556 [hep-ph]
8. M. Davier, A. Hoecker, B. Malaescu, Z. Zhang, Eur. Phys. J. C **80**(3), 241 (2020) (Erratum: Eur. Phys. J. C 80(5), 410 (2020)). https://doi.org/10.1140/epjc/s10052-020-7792-2. arXiv:1908.00921 [hep-ph]
9. A. Keshavarzi, D. Nomura, T. Teubner, Phys. Rev. D **101**(1), 014029 (2020). https://doi.org/10.1103/PhysRevD.101.014029. arXiv:1911.00367 [hep-ph]
10. T. Aoyama, M. Hayakawa, T. Kinoshita, M. Nio, Phys. Rev. Lett. **109**, 111808 (2012). https://doi.org/10.1103/PhysRevLett.109.111808. arXiv:1205.5370 [hep-ph]
11. T. Aoyama, T. Kinoshita, M. Nio, Atoms **7**, 28 (2019). https://doi.org/10.3390/atoms7010028
12. A. Czarnecki, W.J. Marciano, A. Vainshtein, Phys. Rev. D **67**, 073006 (2003) (Erratum: Phys. Rev. D 73, 119901 (2006)). https://doi.org/10.1103/PhysRevD.67.073006. arXiv:0212229 [hep-ph]
13. C. Gnendiger, D. Stöckinger, H. Stöckinger-Kim, Phys. Rev. D **88**, 053005 (2013). https://doi.org/10.1103/PhysRevD.88.053005. arXiv:1306.5546 [hep-ph]
14. A. Kurz, T. Liu, P. Marquard, M. Steinhauser, Phys. Lett. B **734**, 144–147 (2014). https://doi.org/10.1016/j.physletb.2014.05.043. arXiv:1403.6400 [hep-ph]
15. K. Melnikov, A. Vainshtein, Phys. Rev. D **70**, 113006 (2004). https://doi.org/10.1103/PhysRevD.70.113006. arXiv:hep-ph/0312226
16. P. Masjuan, P. Sánchez-Puertas, Phys. Rev. D **95**, 054026 (2017). https://doi.org/10.1103/PhysRevD.95.054026. arXiv:1701.05829 [hep-ph]
17. G. Colangelo, M. Hoferichter, M. Procura, P. Stoffer, JHEP **04**, 161 (2017). https://doi.org/10.1007/JHEP04(2017)161. arXiv:1702.07347 [hep-ph]
18. M. Hoferichter, B.L. Hoid, B. Kubis, S. Leupold, S. Schneider, JHEP **10**, 141 (2018). https://doi.org/10.1007/JHEP10(2018)141. arXiv:1808.04823 [hep-ph]
19. A. Gérardin, H.B. Meyer, A. Nyffeler, Phys. Rev. D **100**, 034520 (2019). https://doi.org/10.1103/PhysRevD.100.034520. arXiv:1903.09471 [hep-ph]
20. J. Bijnens, N. Hermansson-Truedsson, A. Rodríguez-Sánchez, Phys. Lett. B **798**, 134994 (2019). https://doi.org/10.1016/j.physletb.2019.134994. arXiv:1908.03331 [hep-ph]
21. G. Colangelo, F. Hagelstein, M. Hoferichter, L. Laub, P. Stoffer, JHEP **03**, 101 (2020). https://doi.org/10.1007/JHEP03(2020)101. arXiv:1910.13432 [hep-ph]
22. V. Pauk, M. Vanderhaeghen, Eur. Phys. J. C **74**, 3008 (2014). https://doi.org/10.1140/epjc/s10052-014-3008-y. arXiv:1401.0832 [hep-ph]
23. I. Danilkin, M. Vanderhaeghen, Phys. Rev. D **95**, 014019 (2017). https://doi.org/10.1103/PhysRevD.95.014019. arXiv:1611.04646 [hep-ph]
24. F. Jegerlehner, Springer Tracts Mod. Phys. **274**, 1 (2017). https://doi.org/10.1007/978-3-319-63577-4
25. M. Knecht, S. Narison, A. Rabemananjara, D. Rabetiarivony, Phys. Lett. B **787**, 111 (2018). https://doi.org/10.1016/j.physletb.2018.10.048. arXiv:1808.03848 [hep-ph]
26. G. Eichmann, C. Fischer, R. Williams, Phys. Rev. D **101**, 054015 (2020). https://doi.org/10.1103/PhysRevD.101.054015. arXiv:1910.06795 [hep-ph]
27. P. Roig, P. Sánchez-Puertas, Phys. Rev. D **101**, 074019 (2020). https://doi.org/10.1103/PhysRevD.101.074019. arXiv:1910.02881 [hep-ph]
28. T. Blum, N. Christ, M. Hayakawa, T. Izubuchi, L. Jin, C. Jung, C. Lehner, Phys. Rev. Lett. **124**, 132002 (2020). https://doi.org/10.1103/PhysRevLett.124.132002. arXiv:1911.08123 [hep-ph]
29. G. Colangelo, M. Hoferichter, A. Nyffeler, M. Passera, P. Stoffer, Phys. Lett. B **735**, 90 (2014). https://doi.org/10.1016/j.physletb.2014.06.012. arXiv:1403.7512 [hep-ph]
30. F. Ambrosino et al., [KLOE], Phys. Lett. B **670**, 285–291 (2009). https://doi.org/10.1016/j.physletb.2008.10.060. arXiv:0809.3950 [hep-ex]
31. F. Ambrosino et al., [KLOE], Phys. Lett. B **700**, 102–110 (2011). https://doi.org/10.1016/j.physletb.2011.04.055. arXiv:1006.5313 [hep-ex]
32. D. Babusci et al., [KLOE], Phys. Lett. B **720**, 336–343 (2013). https://doi.org/10.1016/j.physletb.2013.02.029. arXiv:1212.4524 [hep-ex]
33. A. Anastasi et al., [KLOE-2]. JHEP **03**, 173 (2018). https://doi.org/10.1007/JHEP03(2018)173. arXiv:1711.03085 [hep-ex]
34. B. Aubert et al., [BABAR], Phys. Rev. Lett. **103**, 231801 (2009). https://doi.org/10.1103/PhysRevLett.103.231801. arXiv:0908.3589 [hep-ex]
35. J.P. Lees et al., [BABAR], Phys. Rev. D **86**, 032013 (2012). https://doi.org/10.1103/PhysRevD.86.032013. arXiv:1205.2228 [hep-ex]
36. F.V. Ignatov et al., [CMD-3], Phys. Rev. D **109**, 112002 (2024). https://doi.org/10.1103/PhysRevD.109.112002. arXiv:2302.08834 [hep-ex]
37. F.V. Ignatov et al., [CMD-3], Phys. Rev. Lett. **132**, 231903 (2024). https://doi.org/10.1103/PhysRevLett.132.231903. arXiv:2309.12910 [hep-ex]
38. S. Borsanyi et al., Nature **593**(7857), 51–55 (2021). https://doi.org/10.1038/s41586-021-03418-1. arXiv:2002.12347 [hep-lat]
39. T. Blum et al., [RBC and UKQCD], Phys. Rev. Lett. **121**(2), 022003 (2018). https://doi.org/10.1103/PhysRevLett.121.022003. arXiv:1801.07224 [hep-lat]
40. M. Cè et al., Phys. Rev. D **106**(11), 114502 (2022). https://doi.org/10.1103/PhysRevD.106.114502. arXiv:2206.06582 [hep-lat]
41. C. Alexandrou et al., [Extended Twisted Mass], Phys. Rev. D **107**(7), 074506 (2023). https://doi.org/10.1103/PhysRevD.107.074506. arXiv:2206.15084 [hep-lat]
42. T. Blum et al., [RBC and UKQCD], Phys. Rev. D **108**(5), 054507 (2023). https://doi.org/10.1103/PhysRevD.108.054507. arXiv:2301.08696 [hep-lat]
43. A. Bazavov et al., [Fermilab Lattice, HPQCD, and MILC], Phys. Rev. D **107**(11), 114514 (2023). https://doi.org/10.1103/PhysRevD.107.114514. arXiv:2301.08274 [hep-lat]
44. G. Colangelo, A.X. El-Khadra, M. Hoferichter, A. Keshavarzi, C. Lehner, P. Stoffer, T. Teubner, Phys. Lett. B **833**, 137313 (2022). https://doi.org/10.1016/j.physletb.2022.137313. arXiv:2205.12963 [hep-ph]
45. M. Davier et al., Phys. Rev. D **109**, 076019 (2024). https://doi.org/10.1103/PhysRevD.109.076019. arXiv:2308.04221 [hep-ph]
46. J.P. Lees et al., Phys. Rev. D **108**, L111103 (2023). https://doi.org/10.1103/PhysRevD.108.L111103 [BABAR], arXiv:2308.05233 [hep-ex]
47. Muon $g-2$ Theory Initiative: https://muon-gm2-theory.illinois.edu
48. M. Davier, A. Hoecker, B. Malaescu, C.Z. Yuan, Z. Zhang, Eur. Phys. J. C **66**, 1–9 (2010). https://doi.org/10.1140/epjc/s10052-010-1246-1. arXiv:0908.4300 [hep-ph]
49. M. Davier et al., Eur. Phys. J. C **66**, 127–136 (2010). https://doi.org/10.1140/epjc/s10052-009-1219-4. arXiv:0906.5443 [hep-ph]
50. M.N. Achasov et al., SND, JHEP **01**, 113 (2021). https://doi.org/10.1007/JHEP01(2021)113. arXiv:2004.00263 [hep-ex]
51. M. Ablikim et al., [BESIII], Phys. Lett. B **753**, 629–638, (2016) (Erratum: Phys. Lett. B 812 (2021), 135982). https://doi.org/10.1016/j.physletb.2015.11.043. arXiv:1507.08188 [hep-ex]

52. R.R. Akhmetshin et al., [CMD-2], Phys. Lett. B **648**, 28–38 (2007). https://doi.org/10.1016/j.physletb.2007.01.073. arXiv:hep-ex/0610021 [hep-ex]
53. Seminars and discussions organised by the Muon $g-2$ Theory Initiative: March 27, 2023 https://indico.fnal.gov/event/59052/; July 20, 2023. https://indico.ijclab.in2p3.fr/event/9697/
54. M. Davier, Muon $g-2$ Theory Initiative Workshop, Bern, 4-8 September 2023, https://indico.cern.ch/event/1258310/
55. R.R. Akhmetshin et al., [CMD-2], Phys. Lett. B **578**, 285–289 (2004). https://doi.org/10.1016/j.physletb.2003.10.108. arXiv:hep-ex/0308008 [hep-ex]
56. M.N. Achasov et al., [SND], J. Exp. Theor. Phys. **103**, 380–384 (2006). https://doi.org/10.1134/S106377610609007X. arXiv:hep-ex/0605013
57. G. Colangelo, M. Hoferichter, P. Stoffer, Phys. Lett. B **814**, 136073 (2021). https://doi.org/10.1016/j.physletb.2021.136073. arXiv:2010.07943 [hep-ph]
58. G. Colangelo, M. Hoferichter, P. Stoffer, PoS **FPCP2023**, 042. https://doi.org/10.22323/1.445.0042. arXiv:2308.04217 [hep-ph]
59. The AfkQed program is based on the following references: H. Czyż, J.H. Kühn, Eur. Phys. J. C **18**, 497–509 (2001). https://doi.org/10.1007/s100520000553. arXiv:hep-ph/0008262
60. S. Binner, J.H. Kühn, K. Melnikov, Phys. Lett. B **459**, 279–287 (1999). https://doi.org/10.1016/S0370-2693(99)00658-9. arXiv:hep-ph/9902399
61. M. Caffo, H. Czyż, E. Remiddi, Nuovo Cim. A **110**, 515–536 (1997). https://doi.org/10.1007/BF03035898. arXiv:hep-ph/9704443
62. E. Barberio, B. van Eijk, Z. Wąs, Comput. Phys. Commun. **66**, 115–128 (1991). https://doi.org/10.1016/0010-4655(91)90012-A
63. S. Jadach, B.F.L. Ward, Z. Wąs, Comput. Phys. Commun. **130**, 260–325 (2000). https://doi.org/10.1016/S0010-4655(00)00048-5. arXiv:hep-ph/9912214 [hep-ph]
64. F. Campanario et al., JHEP **02**, 114 (2014). https://doi.org/10.1007/JHEP02(2014)114. arXiv:1312.3610 [hep-ph]
65. G. Colangelo, M. Hoferichter, J. Monnard, J.R. de Elvira, JHEP **08**, 295 https://doi.org/10.1007/JHEP08(2022)295. arXiv:2207.03495 [hep-ph]
66. T. Kinoshita, J. Math. Phys. **3**, 650 (1962)
67. T.D. Lee, M. Nauenberg, Phys. Rev. **133**, 1549 (1964)
68. T. Xiao, S. Dobbs, A. Tomaradze, K.K. Seth, G. Bonvicini, Phys. Rev. D **97**(3), 032012 (2018). https://doi.org/10.1103/PhysRevD.97.032012. arXiv:1712.04530 [hep-ex]
69. Private communication from Henryk Czyż
70. S . Müller et al., KLOE note 221 (2008)
71. A. Anastasi et al., [KLOE-2], Phys. Lett. B **767**, 485–492 (2017). https://doi.org/10.1016/j.physletb.2016.12.016. arXiv:1609.06631 [hep-ex]
72. C.Z. Yuan, L.L. Wang, private communication
73. We shared the information about the found problems with Henryk Czyż, one of the PHOKHARA authors
74. I. Adachi et al., [Belle-II], arXiv:2404.04915 [hep-ex]
75. R. Alemany, M. Davier, A. Hoecker, Eur. Phys. J. C **2**, 123–135 (1998). https://doi.org/10.1007/s100520050127. arXiv:hep-ph/9703220
76. M. Davier, A. Hoecker, Phys. Lett. B **419**, 419–431 (1998). https://doi.org/10.1016/S0370-2693(97)01512-8. arXiv:hep-ph/9801361
77. M. Davier, A. Hoecker, Phys. Lett. B **435**, 427–440 (1998). https://doi.org/10.1016/S0370-2693(98)00825-9. arXiv:hep-ph/9805470
78. M. Davier, S. Eidelman, A. Hoecker, Z. Zhang, Eur. Phys. J. C **27**, 497–521 (2003). https://doi.org/10.1140/epjc/s2003-01136-2. arXiv:hep-ph/0208177
79. M. Davier, S. Eidelman, A. Hoecker, Z. Zhang, Eur. Phys. J. C **31**, 503–510 (2003). https://doi.org/10.1140/epjc/s2003-01362-6. arXiv:hep-ph/0308213
80. M. Davier, A. Hoecker, B. Malaescu, Z. Zhang, Eur. Phys. J. C **71**, 1515 (2011).(Erratum: Eur. Phys. J. C 72 (2012), 1874). https://doi.org/10.1140/epjc/s10052-012-1874-8. arXiv:1010.4180 [hep-ph]
81. M. Davier, A. Hoecker, B. Malaescu, C.Z. Yuan, Z. Zhang, Eur. Phys. J. C **74**, 2803 (2014). https://doi.org/10.1140/epjc/s10052-014-2803-9. arXiv:1312.1501 [hep-ex]
82. V. Cirigliano, G. Ecker, H. Neufeld, Phys. Lett. B **513**, 361–370 (2001). https://doi.org/10.1016/S0370-2693(01)00764-X. arXiv:hep-ph/0104267
83. V. Cirigliano, G. Ecker, H. Neufeld, JHEP **08**, 002 (2002). https://doi.org/10.1088/1126-6708/2002/08/002. arXiv:hep-ph/0207310
84. F. Flores-Baez, A. Flores-Tlalpa, G. Lopez Castro, G. Toledo Sanchez, Phys. Rev. D **74**, 071301 (2006). https://doi.org/10.1103/PhysRevD.74.071301. arXiv:hep-ph/0608084
85. A. Flores-Tlalpa, F. Flores-Baez, G. Lopez Castro, G. Toledo Sanchez, Nucl. Phys. B Proc. Suppl. **169**, 250–254 (2007). https://doi.org/10.1016/j.nuclphysbps.2007.03.011. arXiv:hep-ph/0611226
86. J.A. Miranda, P. Roig, Phys. Rev. D **102**, 114017 (2020). https://doi.org/10.1103/PhysRevD.102.114017. arXiv:2007.11019 [hep-ph]
87. M. Bruno, Muon $g-2$ Theory Initiative Workshop, Bern, 4-8 September 2023, https://indico.cern.ch/event/1258310/
88. F. Jegerlehner, R. Szafron, Eur. Phys. J. C **71**, 1632 (2011). https://doi.org/10.1140/epjc/s10052-011-1632-3. arXiv:1101.2872 [hep-ph]
89. F. Jegerlehner, ECFA workshop e^+e^- Higgs/EW/Top Factory working group WG1 (July 14 2022), https://indico.cern.ch/event/1173700/
90. Clarifying discussions with G. Colangelo are gratefully acknowledged
91. A. Sirlin, Rev. Mod. Phys. **50**, 573 (1978) (Erratum: Rev. Mod. Phys. **50** (1978), 905). https://doi.org/10.1103/RevModPhys.50.573
92. W.J. Marciano, A. Sirlin, Phys. Rev. Lett. **61**, 1815–1818 (1988). https://doi.org/10.1103/PhysRevLett.61.1815
93. A. Sirlin, Nucl. Phys. B **196**, 83–92 (1982). https://doi.org/10.1016/0550-3213(82)90303-0
94. E. Braaten, C.S. Li, Phys. Rev. D **42**, 3888–3891 (1990). https://doi.org/10.1103/PhysRevD.42.3888
95. G. Colangelo, M. Hoferichter, B. Kubis, P. Stoffer, JHEP **10**, 032 (2022). https://doi.org/10.1007/JHEP10(2022)032. arXiv:2208.08993 [hep-ph]
96. P. Stoffer, presentation in the discussion session on e^+e^- data, Muon $g-2$ Theory Initiative Workshop, Bern, 4-8 September 2023, https://indico.cern.ch/event/1258310/
97. I. Logashenko, Muon $g-2$ Theory Initiative Workshop, Bern, 4-8 September 2023, https://indico.cern.ch/event/1258310/
98. A. Kupich, Muon $g-2$ Theory Initiative Workshop, Bern, 4-8 September 2023, https://indico.cern.ch/event/1258310/
99. P. Masjuan, A. Miranda, P. Roig, Phys. Lett. B **850**, 138492 (2024). https://doi.org/10.1016/j.physletb.2024.138492. arXiv:2305.20005 [hep-ph]