Dispersive determination of the HVP contribution to the muon $g - 2$

Martin Hoferichter

Abstract. The determination of hadronic vacuum polarization (HVP) from $e^+e^-\rightarrow$ hadrons cross-section data, in the energy region relevant for the anomalous magnetic moment of the muon, has recently been challenged by lattice-QCD calculations, especially for the intermediate window in Euclidean time. In these proceedings we review some frequently-asked questions on the comparison between data-driven and lattice-QCD evaluations of the HVP contribution.

1 Introduction

Traditionally, the HVP contribution to $a_\mu$ has been evaluated using cross-section data for $e^+e^-\rightarrow$ hadrons, leading to the current $4.2\sigma$ tension between Standard-Model prediction [1–25] and the experimental world average [26–30]. However, this data-driven evaluation lies 2.1$\sigma$ lower than the lattice-QCD calculation by BMWc [31], which was subsequently confirmed by other collaborations at least for the intermediate window quantity (following the definition with parameters from Ref. [32]). For this partial quantity, the current situation is summarized in Fig. 1, indicating a much stronger tension than observed globally by BMWc.

The consequences of this pattern have been studied in the literature to the extent that this is possible based on the currently available information, including the relation to the global electroweak fit via the hadronic running of the fine-structure constant [55–58] and changes to the $2\pi$ cross section [59]. In particular, the latter analysis shows that while, in principle, modifications to the $2\pi$ cross section consistent with analyticity and unitarity constraints would be able to accommodate the global difference to the BMWc result, a simultaneous explanation of the tension in the intermediate window is not possible. Assuming a rather uniform shift in the low-energy $2\pi$ region as well as no significant negative shifts, the form of the kernel functions suggests that at least 40% of the changes would need to originate from the energy region above 1 GeV. Such a conclusion is line with direct lattice-QCD results for the hadronic running itself [31, 60], given that the moderate increase observed therein implies that the changes cannot occur too high in energy, because then the effect on the running would be too large, but also not concentrated at very low energies, as in this case no effect would be visible at all. 

Extracting more detailed information on the energy dependence is a complicated endeavor, amounting to the inversion of a Laplace transform, but could be possible if finer windows become available [37], with linear combinations to be optimized according to lattice and data-driven covariance matrices, or by considering spectral-
weight sum rules [61]. At the moment, the situation summarized in Fig. 1 presents a puzzle, whose resolution remains far from obvious. It is therefore prudent to reconsider the relation between lattice-QCD calculations and data-driven evaluations in more detail, here, we review some of the frequently-asked questions that arise in this context.

2 Conventions for higher-order corrections

By convention, the cross section that enters the HVP master formula [62, 63]

\[
\sigma_{\mu}^{\text{HVP, LO}}(s) = \left(\frac{e m_{\mu}}{3 \pi}\right)^2 \int_{s_0}^{s_{\text{max}}} d s \frac{\hat{K}(s)}{s^2} R_{\text{had}}(s),
\]

\[
R_{\text{had}}(s) = -\frac{3}{4 \pi \alpha^2} \sigma(e^+ e^- \rightarrow \text{hadrons}(\gamma))
\]

is defined in a photon-inclusive way, including radiative intermediate states and final-state radiation (FSR). Initial-state radiation (ISR) and VP are subtracted to avoid double counting with higher-order HVP insertions. This bookkeeping is illustrated in Fig. 2, where diagram (a) gives photonic corrections and muon VP, diagram (b) VP corrections from \(\ell = e, \tau\) (leading to a dependence on the respective lepton mass ratio), and diagram (c) the double HVP iteration. Numerically, these three classes sum up to [11, 64–66]

\[
\sigma_{\mu}^{\text{HVP, NLO}} = -20.7 + 10.6 + 0.3 \times 10^{-10}
\]

\[
\approx -9.8 \times 10^{-10},
\]

in such a way that the double HVP iteration, which potentially requires some care in the definition of the hadronic two-point function, only produces a very small effect. Defining each VP function as the one-particle-irreducible amplitude, in accordance with the running of the fine-structure constant

\[
\alpha(q^2) = \frac{\alpha(0)}{1 - \Delta \alpha_{\text{ lept}}(q^2) - \Delta \alpha_{\text{ had}}(q^2)},
\]

\[
\Delta \alpha_{\text{ had}}(q^2) = \frac{-\alpha^2}{3 \pi} \int_{s_0}^{s_{\text{max}}} d s \frac{R_{\text{had}}(s)}{s(s - q^2)}
\]

one can thus unfold the experimental cross sections and iteratively determine the \(R\)-ratio \(R_{\text{had}}(s)\) to be inserted in Eq. (1). In this way, it suffices to consider the exclusive hadronic channels that contribute, without the need to ever specify the role of a specific resonance such as \(\rho, \omega, \phi\). The only subtlety concerns very narrow \(\epsilon\) and \(bb\) resonances for which the Dyson series no longer converges [7, 67]. In these cases, one needs to take out the resonance that is being corrected in \(R_{\text{had}}\) in the VP undressing, but such a procedure is not even necessary for the low-energy cross sections that dominate the HVP integral. The same conventions are used also for yet higher-order HVP insertions [13, 68].

To match these conventions in lattice QCD, isospin-breaking (IB) corrections need to be included, both arising from QED and strong IB. The latter is, in principle, straightforward to include by means of insertions of the quark mass difference, while the QED corrections need to be defined in accordance with the conventions for \(\sigma_{\mu}^{\text{HVP, LO}}\) in the data-driven approach [32]. This mainly concerns the quark-level diagram shown in Fig. 3, which only becomes one-particle-irreducible if dressed with additional gluon interactions between the quark lines. Accordingly, the diagram without any additional gluons needs to be subtracted, as can be implemented by subtracting the separate quantum averages of the quark loops. The final sum of isospin-limit calculation and IB corrections can then be directly compared to phenomenology, and it is this comparison that is shown in Fig. 1.\footnote{Note that at present only Refs. [31, 32] provide complete calculations of the IB corrections, while Ref. [34] uses a partial estimate and Ref. [35] the result from Ref. [31]. Both QED and strong IB corrections are found to be small for the intermediate window.}

The size of the IB corrections can also be estimated directly from phenomenology, either by summing up contributions from the exclusive channels [69] or, in the case of strong IB, from chiral perturbation theory [70], with the critical low-energy constant \(\delta C_{10}^{(1)}\) extracted from \(\tau\) decays. In the exclusive approach, sizable effects arise from the radiative channels \(\pi^0 \gamma, \eta \gamma\), FSR, \(\rho\) mixing, and pion-mass effects in \(2\pi, 71, 72\), as well as similar, resonance-enhanced effects in \(KK\) [73]. Separating all effects into QED and strong contributions, and decomposing onto the Euclidean-time windows, there is reasonable agreement with Refs. [31, 32], especially in view of limitations in the phenomenological approach due to unknown IB in resonance couplings and the truncation in the sum over hadronic channels. In fact, Refs. [69, 70] both indicate a larger strong IB correction for the full HVP contribution than Ref. [31] (albeit largely consistent within uncertainties) and Ref. [69] finds an increased QED correction in

\begin{figure}
\centering
\includegraphics[width=0.5\textwidth]{figure3}
\caption{Critical diagram for defining QED corrections in lattice QCD (solid lines refer to quarks, the wiggly line to the photon, and gluons are implied everywhere), diagram (f) \(F\) in the notation of Ref. [32]. The one-particle-reducible piece without any additional gluons needs to be subtracted.}
\end{figure}
the intermediate window. In either case the lattice-QCD result for HVP would increase even further if these phenomenological estimates were adopted, suggesting that issues with IB corrections are unlikely to resolve the observed tension.

3 Radiative corrections and Monte-Carlo generators

In practice, a key question becomes whether radiative corrections as correspond to the conventions in Sec. 2 can be controlled and implemented in Monte-Carlo (MC) generators at the sufficient level of precision. The status of MC generators around the time the BaBar [74, 75] and KLOE [76–79] 2π data appeared is documented in detail in Ref. [80]. From a theory perspective the formalism is based on scalar QED supplemented by the pion vector form factor whenever possible (“FsQED”), based on the observation that this strategy should capture the dominant, infrared-enhanced (IR) effects [81–84]. Corrections beyond IR enhancement were evaluated in Ref. [85] for the πτ channel and found to be small. Another possible concern regarding the ISR data sets arose from the fact that the MC generator PHOJET [86, 87] did not include full NLO corrections, but these previously only estimated diagrams were recently incorporated, not revealing any sizable corrections [88].

Going forward, improved MC generators are key for the success of future experimental campaigns [91]. Recent developments involve the forward–backward asymmetry in e⁺e⁻ → π⁺π⁻, which can be considered a valuable test case, since due to the C-odd nature the asymmetry vanishes at tree level, while radiative corrections generate a non-vanishing result from the interference of ISR and FSR contributions [92–94]. In Ref. [89] it was observed that scalar QED without form-factor modifications was not able to describe preliminary CMD-3 data on the asymmetry, but from the perspective of dispersion relations the required form-factor corrections do correspond to the IR enhanced effects, as expected [90]. Ultimately, this is the reason why the full dispersive calculation largely reproduces the outcome of the model estimate, see Fig. 4, as the dominant pole–dispersive contribution receives both significant IR and form-factor enhancements. The study of the C-even case is in progress [95], and it remains to be seen if the dispersive analysis reveals unexpectedly large effects. The test case of the asymmetry, however, makes it appear unlikely that such effects can resolve the tension with lattice QCD either.

4 τ data

The use of τ data for the evaluation of the HVP contribution relies on the relation between the vector-current-mediated decay τ → Xντ and the cross section e⁺e⁻ → X⁰ involving the corresponding isovector final state X⁰, which becomes exact in the limit of isospin invariance [96]. In practice, this implies that IB corrections need to be controlled very accurately if data for X = 2π [97, 98] are to be included in the HVP evaluation. Phenomenological estimates of these corrections, including Refs. [99–102], are typically separated into FSR, phase-space factors, long-range radiative corrections, and the pion form factors corresponding to the two charge channels of the ρ(770). Here, the need to argue in terms of hadronic states, with ρ resonance parameters to be determined from elsewhere, incurred a significant model dependence, which, together with the superior precision from e⁺e⁻ data, ultimately led to τ data not being considered in most data-driven HVP evaluations [1].

Given the tension between lattice QCD and e⁺e⁻ data, an independent check via τ decays would be highly valuable, also in view of potential experimental improvements at Belle II [103], but would require a convincing solution to the issue of IB corrections. Recent progress in this direction was reported using lattice QCD [104], allowing one to address the calculation of IB effects without the need to specify ρ properties. Challenges in this program include the matching to the experimental observables, e.g., the 2π channel needs to be isolated from the inclusive lattice calculation and long-range QED corrections need to be implemented in a consistent manner. For this last step, also the interplay with dispersive methods developed in the context of radiative corrections in e⁺e⁻ → π⁺π⁻ [95] could prove beneficial.

5 Avenues for future progress

Prospects for improved predictions of α₂ in the Standard Model are described in detail in Ref. [105]. In lattice QCD, this includes full calculations from other collaborations at a level comparable to BMWc, given that a comparison as in Fig. 1 is so far only available for the intermediate window. Moreover, results for other windows are
expected, both as a cross check that their sum reproduces the full HVP integral and as a diagnostic tool to scrutinize first consistency among lattice-QCD calculations and then the comparison to $e^+e^-$ data.

On the data-driven side, new input has already become available after Ref. [1], for $2\pi$ [106], $3\pi$ [107, 108], and inclusive measurements [109], but so far with a small impact on global HVP compilations.\(^3\) Addressing the longstanding tension between KLOE and BaBar will require new data at a similar level of precision, as are expected from CMD-3 [110], BaBar [91], BESIII [111], and Belle II [103] in the coming years. Together with continued scrutiny of radiative corrections and, potentially, new input from $\tau$ decays, this should allow one to corroborate the situation for time-like data-driven evaluations.

In addition, the MuonE experiment [112–114] aims at a completely independent determination of the HVP contribution, by means of a space-like extraction from muon–electron scattering. Controlling the systematics of the cross section at a level of $10^{-5}$ presents unique challenges to both experiment and theory [115], with a resulting measurement that would give a direct determination in the space-like region and thus be fully complementary to the time-like approach.

In conclusion, the present situation in the determination of HVP remains puzzling, especially given the increased tension in the intermediate window shown in Fig. 1, to the extent that even scenarios for beyond-Standard-Model contamination in the $e^+e^-$ data have been considered [116–118]. To resolve this puzzle, a large, concerted effort in lattice QCD, experiment, and phenomenology is currently underway, which should allow for critical new insights on the timescale of the Fermilab experiment.

Acknowledgements

I thank Gilberto Colangelo and Peter Stoffer for longstanding collaboration on many of the topics discussed in these proceedings, as well as for comments on the manuscript. Support by the Swiss National Science Foundation, under Project No. PCEFP2_181117, is gratefully acknowledged.

References


\(^3\)Note, however, that in Ref. [71] tensions with analyticity and unitarity constraints were pointed out for the SND data set [106], as were tensions in the $2\pi$ data base in the context of the $\rho$–$\omega$ interference.