

# Isospin-breaking corrections to light-meson leptonic decays from lattice simulations at physical quark masses

Peter Boyle,<sup>a,b</sup> Matteo Di Carlo,<sup>b</sup> Felix Erben,<sup>b</sup> Vera Gülpers,<sup>b</sup> Maxwell T. Hansen,<sup>b</sup> Tim Harris,<sup>b</sup> Nils Hermansson-Truedsson,<sup>c,d</sup> Raoul Hodgson,<sup>b</sup> Andreas Jüttner,<sup>e,f</sup> Fionn Ó hÓgáin,<sup>b</sup> Antonin Portelli,<sup>b</sup> James Richings<sup>b,e,g</sup> and Andrew Zhen Ning Yong<sup>b</sup>

<sup>a</sup>Physics Department, Brookhaven National Laboratory,  
Upton NY 11973, U.S.A.

<sup>b</sup>School of Physics and Astronomy, University of Edinburgh,  
Edinburgh EH9 3FD, United Kingdom

<sup>c</sup>Albert Einstein Center for Fundamental Physics, Institute for Theoretical Physics,  
Universität Bern,  
Sidlerstrasse 5, CH-3012 Bern, Switzerland

<sup>d</sup>Department of Astronomy and Theoretical Physics, Lund University,  
Sölvegatan 14A, 223 62 Lund, Sweden

<sup>e</sup>Physics and Astronomy, University of Southampton,  
Southampton SO17 1BJ, United Kingdom

<sup>f</sup>CERN, Theoretical Physics Department,  
CH-1211 Geneva, Switzerland

<sup>g</sup>EPCC, University of Edinburgh,  
EH8 9BT, Edinburgh, United Kingdom

E-mail: [pboyle@bnl.gov](mailto:pboyle@bnl.gov), [matteo.dicarlo@ed.ac.uk](mailto:matteo.dicarlo@ed.ac.uk), [felix.erben@ed.ac.uk](mailto:felix.erben@ed.ac.uk),  
[vera.gulpers@ed.ac.uk](mailto:vera.gulpers@ed.ac.uk), [maxwell.hansen@ed.ac.uk](mailto:maxwell.hansen@ed.ac.uk), [tharris@ed.ac.uk](mailto:tharris@ed.ac.uk),  
[nils.hermansson-truedsson@thep.lu.se](mailto:nils.hermansson-truedsson@thep.lu.se), [raoul.hodgson@ed.ac.uk](mailto:raoul.hodgson@ed.ac.uk),  
[andreas.juttner@cern.ch](mailto:andreas.juttner@cern.ch), [fionn.o.hogain@ed.ac.uk](mailto:fionn.o.hogain@ed.ac.uk),  
[antonin.portelli@ed.ac.uk](mailto:antonin.portelli@ed.ac.uk), [j.richings@epcc.ed.ac.uk](mailto:j.richings@epcc.ed.ac.uk),  
[andrew.yong@ed.ac.uk](mailto:andrew.yong@ed.ac.uk)

**ABSTRACT:** The decreasing uncertainties in theoretical predictions and experimental measurements of several hadronic observables related to weak processes, which in many cases are now smaller than  $O(1\%)$ , require theoretical calculations to include subleading corrections that were neglected so far. Precise determinations of leptonic and semi-leptonic decay rates, including QED and strong isospin-breaking effects, can play a central role in solving the current tensions in the first-row unitarity of the CKM matrix. In this work we present the first RBC/UKQCD lattice calculation of the isospin-breaking corrections to the ratio of leptonic decay rates of kaons and pions into muons and neutrinos. The calculation is performed at fixed lattice spacing ( $a^{-1} \simeq 1.730$  GeV) on a  $48^3 \times 96$  volume with  $N_f = 2 + 1$  dynamical quarks close to the physical point and domain wall fermions in the Möbius formulation are employed. Long-distance QED interactions are included according to the QED<sub>L</sub> prescription and the crucial role of finite-volume electromagnetic corrections in the determination of leptonic decay rates, which produce a large systematic uncertainty, is extensively discussed. Finally, we study the different sources of uncertainty on  $|V_{us}|/|V_{ud}|$  and observe that, if finite-volume systematics can be reduced, the error from isospin-breaking corrections is potentially sub-dominant in the final precision of the ratio of the CKM matrix elements.

**KEYWORDS:** CKM Parameters, Hadronic Matrix Elements and Weak Decays, Lattice QCD, Standard Model Parameters

**ARXIV EPRINT:** [2211.12865](https://arxiv.org/abs/2211.12865)

---

**Contents**

<b>1</b>	<b>Introduction</b>	<b>1</b>
<b>2</b>	<b>Isospin-breaking corrections</b>	<b>3</b>
2.1	Renormalizing the full theory	4
2.2	Defining QCD and its isospin-symmetric limit	5
2.3	Computing isospin-breaking effects on the lattice	7
2.4	Scheme ambiguities	8
<b>3</b>	<b>Matrix elements from Euclidean correlation functions</b>	<b>9</b>
3.1	Virtual corrections to the leptonic decay rate	10
3.2	Extracting matrix elements from Euclidean correlation functions	15
3.3	Subtraction of finite-volume effects	17
3.4	Inclusion of real photon emission	19
<b>4</b>	<b>Lattice methodology</b>	<b>20</b>
4.1	Lattice QCD+QED path integrals	20
4.2	Lattice setup	21
4.3	Implementation of the hadronic correlators	22
<b>5</b>	<b>Numerical analysis</b>	<b>28</b>
5.1	Strategy for correlator fits	29
5.2	Tuning of the bare parameters	33
5.3	Estimation of model uncertainties	35
<b>6</b>	<b>Results</b>	<b>36</b>
<b>7</b>	<b>Conclusions</b>	<b>40</b>
<b>A</b>	<b>Comparison of local and conserved electromagnetic currents</b>	<b>42</b>
A.1	Theory	42
A.2	Numerical check	44
<b>B</b>	<b>zMöbius to Möbius correction</b>	<b>45</b>
<b>C</b>	<b>Free domain wall fermion propagators</b>	<b>46</b>
C.1	Feynman rules for free propagator	48
C.2	Pole mass of the free propagator	49
C.3	Projectors on definite spinor structure	50
<b>D</b>	<b>Correlation functions for the <math>\Omega</math> baryon</b>	<b>53</b>
<b>E</b>	<b>Determining best fits with a genetic algorithm</b>	<b>56</b>

---

## 1 Introduction

Flavour physics offers a unique opportunity in the search for new physics at the precision frontier of the Standard Model (SM). Discrepancies between SM predictions and experimental observations of processes where yet undiscovered particles or fields may play a tiny but measurable role can in fact be signals of new physics beyond the SM. In the hadronic sector, the study of processes mediated by the weak force gives access to the elements of the Cabibbo-Kobayashi-Maskawa (CKM) matrix describing quark-flavour mixing. The accurate determination of the CKM matrix elements  $V_{ud}$  and  $V_{us}$  is of crucial importance to test the first-row unitarity  $|V_{ud}|^2 + |V_{us}|^2 + |V_{ub}|^2 = 1$  imposed by the SM and to probe emerging tensions that are approaching the  $3\sigma$  confidence level [1–3]. A complete understanding of SM processes like the leptonic and semi-leptonic decay modes of pseudoscalar mesons or nuclear beta decays which underpin these constraints is therefore necessary to test CKM unitarity and eventually put bounds on the new physics energy scale and couplings.

In particular, in this work we are concerned with the precision determination of the ratio  $|V_{us}|/|V_{ud}|$  obtained by combining the experimental leptonic decay rates of the pion ( $\pi_{\mu 2}$ ) and kaon ( $K_{\mu 2}$ ) into a muon and a neutrino with hadronic matrix elements which parameterize the SM prediction. Given the non-perturbative dynamics of strong interactions at low energies, these theoretical determinations can be obtained in a reliable and systematically improvable way from first principles lattice field theory computations. Lattice QCD has now entered the precision era and is able to provide many hadronic quantities with percent precision, e.g. the ratio of kaon and pion leptonic decay constants  $f_K/f_\pi$  and the kaon semi-leptonic decay ( $K_{\ell 3}$ ) vector form factor  $f^+(0)$ , which play a central role in the determination of the CKM quantities  $|V_{us}|/|V_{ud}|$  and  $|V_{us}|$ , respectively [2]. To date, most lattice QCD computations in flavour physics neglect isospin-breaking (IB) effects, namely the inclusion of electromagnetism and the difference of the up and down quark masses, which are required to go beyond percent level precision. These contributions have been historically included using effective field theories such as chiral perturbation theory ( $\chi$ PT) [4–6], where, however, it can be difficult to systematically assess uncertainties emerging from effective expansions. The RM123+Southampton (RM123S) collaboration pioneered the first lattice calculations beyond the QCD isospin limit [7, 8], although with an extrapolation of the result from unphysical quark masses. In this work we provide a first determination of the IB effects in  $f_K/f_\pi$  using an ab initio computation of lattice QCD and QED using a regularization with good chiral properties directly at physical quark masses, and examine its impact on the determination of  $|V_{us}|/|V_{ud}|$ . The calculation is performed on a single  $48^3 \times 96$  lattice with spacing  $a^{-1} \simeq 1.730$  GeV.

When electromagnetism is included, leptonic decay amplitudes can no longer be factorised into QCD and non-QCD contributions, as the lepton can interact with the pseudoscalar meson. Additionally, this new interaction generates infrared (IR) divergences which only cancel when summing diagrams containing virtual and real photon corrections [9]. Thus, the decay rate can be properly written including these effects as

$$\Gamma(P^\pm \rightarrow \ell^\pm \nu[\gamma]) = \lim_{\Lambda_{\text{IR}} \rightarrow 0} [\Gamma_0(\Lambda_{\text{IR}}) + \Gamma_1(\Lambda_{\text{IR}})], \quad (1.1)$$

where we indicate the contribution to the decay rate with virtual photon corrections as  $\Gamma_0$  (virtual decay rate), and that with one real photon in the final state as  $\Gamma_1$  (real decay rate), and where  $\Lambda_{\text{IR}}$  is an arbitrary IR energy cutoff (e.g. a photon mass). A practical strategy for non-perturbative computations was put forward by the RM123S group in ref. [10] (and applied in successive calculations [7, 8]), which consists in defining the inclusive rate as the sum of two contributions which are separately IR safe, namely

$$\Gamma(P^\pm \rightarrow \ell^\pm \nu[\gamma]) = \lim_{\Lambda_{\text{IR}} \rightarrow 0} \left[ \Gamma_0(\Lambda_{\text{IR}}) - \Gamma_0^{\text{uni}}(\Lambda_{\text{IR}}) \right] + \lim_{\Lambda'_{\text{IR}} \rightarrow 0} \left[ \Gamma_0^{\text{uni}}(\Lambda'_{\text{IR}}) + \Gamma_1(\Lambda'_{\text{IR}}) \right], \quad (1.2)$$

where we note that different IR regulators can be used in both terms, as will be the case in practice. The quantity  $\Gamma_0^{\text{uni}}(\Lambda_{\text{IR}})$  corresponds to the *universal* (structure-independent) IR-divergent part of the virtual decay rate, which can be computed perturbatively assuming the decaying meson to be a point-like particle. This term also exactly cancels the divergence in  $\Gamma_1(\Lambda'_{\text{IR}})$ , when evaluated with  $\Lambda'_{\text{IR}}$  as an IR regulator.

In principle, both virtual and real decay rates should be computed non-perturbatively since photons with sufficiently high energy can resolve the internal structure of the decaying meson. While there is no choice for virtual corrections, as all photon modes contribute to the rate, one can impose a cut on the real photon energy,  $\omega_\gamma^{\text{cut}}$ , such that its sensitivity to the structure of the meson is suppressed. In this case  $\Gamma_1(\Lambda_{\text{IR}})$  can be computed analytically in perturbation theory in the point-like approximation, namely  $\Gamma_1^{\text{pt}}(\omega_\gamma^{\text{cut}}, \Lambda_{\text{IR}})$ . As predicted by  $\chi$ PT [6] and confirmed by lattice calculations [10, 11], structure-dependent contributions are negligible for the decay channels studied in this work, namely the decay of pions or kaons into muons and neutrinos, and therefore at our level of precision we can reliably consider  $\Gamma_1(\Lambda_{\text{IR}}) \simeq \Gamma_1^{\text{pt}}(\omega_\gamma^{\text{max}}, \Lambda_{\text{IR}})$ , where  $\omega_\gamma^{\text{max}}$  is the maximum photon energy kinematically allowed. We will focus on the non-perturbative calculation of the virtual decay rate  $\Gamma_0$  and will use the finite spatial extent of the lattice,  $L$ , along with a suitable prescription for QED in a finite volume called QED<sub>L</sub>, as an IR regulator [12]. The real decay rate will be evaluated in the well-motivated point-like approximation and directly in infinite volume, regularizing the IR divergence with a photon mass  $\Lambda'_{\text{IR}} = m_\gamma$  in eq. (1.2). Thus, we follow the approach outlined in refs. [8, 10].

Going beyond the isospin-symmetric limit, the leading corrections to the pseudoscalar decay rates from the electromagnetic fine structure constant  $\alpha_{\text{em}} \approx 1/137$  and from the renormalized (e.g. in  $\overline{\text{MS}}$  at 2 GeV) up-down quark mass difference  $(m_d^{\text{R}} - m_u^{\text{R}})/\Lambda_{\text{QCD}} \sim 1\%$  are both of the order of 1%, which we denote universally by  $\mathcal{O}(\varepsilon) = \mathcal{O}[\alpha_{\text{em}}, (m_d^{\text{R}} - m_u^{\text{R}})/\Lambda_{\text{QCD}}]$ . The correction to the ratio of kaon and pion decay rates is then parameterized by  $\delta R_{K\pi}$ , which can be expressed through the relation [1]

$$\frac{\Gamma(K^+ \rightarrow \mu^+ \nu_\mu[\gamma])}{\Gamma(\pi^+ \rightarrow \mu^+ \nu_\mu[\gamma])} = \frac{|V_{\text{us}}|^2}{|V_{\text{ud}}|^2} \frac{m_\pi}{m_K} \frac{(m_K^2 - m_\mu^2)}{(m_\pi^2 - m_\mu^2)} \frac{f_K^2}{f_\pi^2} (1 + \delta R_{K\pi}) + \mathcal{O}(\varepsilon^2), \quad (1.3)$$

where  $\mathcal{O}(\varepsilon^2)$  is understood as a second-order correction in  $(\alpha_{\text{em}}, (m_d^{\text{R}} - m_u^{\text{R}})/\Lambda_{\text{QCD}})$ . The specific choice defining the isospin-symmetric theory, implicit in the definition of  $f_K/f_\pi$ , will be the subject of section 2, where we discuss the consistency with other choices in the literature and the advantage of simulating with close-to-physical quark masses. With that

in mind, we anticipate our final result for the leading IB corrections to  $f_K^2/f_\pi^2$  as

$$\delta R_{K\pi} = -0.0086 (3)_{\text{stat.}} ({}^{+11}_{-4})_{\text{fit}} (5)_{\text{disc.}} (5)_{\text{quench.}} (39)_{\text{vol.}}, \quad (1.4)$$

where the first error is statistical, while the others are systematic uncertainties and will be discussed in section 6. We note that our result is compatible with the only other lattice determination by the RM123S collaboration [8] and with  $\chi$ PT [13]

$$\delta R_{K\pi}^{\text{RM123S}} = -0.0126 (14), \quad (1.5)$$

$$\delta R_{K\pi}^{\chi\text{PT}} = -0.0112 (21). \quad (1.6)$$

All these numbers are dependent on a choice of scheme to define the separation of isospin-breaking effects. However, as we will discuss later in this paper, one can provide quantitative evidence that the prescriptions used in the results above are close enough so that the prescription-dependence lies well below the quoted uncertainties. In addition to our numerical result, one of the main findings of this work is a refined investigation of the large power-law finite-size corrections which are induced by the QED<sub>L</sub> treatment of electromagnetism in a finite volume, and are reflected in a large systematic error in our result. As the leading structure-dependent finite-size effects are now known to be negligible [14], the dominant point-like corrections which are investigated in section 3 exhibit unexpectedly large higher-order corrections. In that section, the connection between the Euclidean correlation functions and the hadronic matrix elements of interest is outlined. The details of the lattice implementation using domain wall fermions and the gauge ensemble generated by the RBC/UKQCD collaboration may be found in section 4, while the analysis of the numerical data including the estimation of the systematic effects is detailed in section 5. The discussion of the result and the implications for the extraction of  $|V_{us}|/|V_{ud}|$  are found in section 6, before the conclusion.

## 2 Isospin-breaking corrections

Leptonic decays of pions and kaons are low-energy processes that can be studied in an effective Fermi theory where the  $W$ -boson is integrated out, and the process is mediated by a local four-fermion interaction. At first order in the Fermi constant  $G_F$ , we can then assume that low-energy observables can be predicted to a high degree of precision within a theory of QCD+QED. We will refer to this as the full or physical theory in the rest of the paper. In the full theory, quantities like the meson decay constant  $f_P$  entering eq. (1.3) are ambiguous as they are defined in the unphysical iso-symmetric limit of QCD (iso-QCD) where  $m_u = m_d$  and  $\alpha_{\text{em}} = 0$ . Such unphysical definitions are related to the fact that QCD and QCD+QED interactions generate different ultraviolet (UV) divergences and hence require different renormalization procedures to fix the bare parameters of the respective actions. In order to give a meaning to (iso-)QCD observables within the full QCD+QED theory, additional renormalization conditions are then required.

The discussion in this section is divided in two: first, we will discuss how to non-perturbatively renormalize the full theory on the lattice and make well-defined physical

predictions in the continuum limit. Then, we will define QCD and its iso-symmetric limit, which is employed in our numerical lattice calculation and fixes the definition of IB corrections. However, the discussion here is general and not restricted to lattice QCD calculations. In fact, the definitions of the QCD+QED, QCD and iso-QCD theory discussed in the rest of the section hold for any other non-perturbative approach (like e.g. effective field theories).

## 2.1 Renormalizing the full theory

In the full QCD+QED theory with  $N_f$  flavours of quarks, once a UV regulator is introduced (in our case the lattice spacing  $a$ ), the action depends only on the bare quark masses in lattice units  $\hat{\mathbf{m}} = (\hat{m}_1, \dots, \hat{m}_{N_f})$  and the bare strong and electromagnetic couplings,  $g$  and  $e$ , respectively. In this theory every physical observable can be predicted once the bare parameters of the action are defined and the regulator is removed. Note that, since we are only working at first order  $O(\varepsilon)$  in the IB effects, we can neglect the running of the electromagnetic coupling and safely fix it to its Thomson limit,  $e^\phi = (4\pi\alpha_{\text{em}}^\phi)^{1/2}$  with  $\alpha_{\text{em}}^\phi = 1/137.035999084$  (21) [15], without the need to impose a specific renormalization condition. Moreover, when working at first order in  $\alpha_{\text{em}}$  and when a lepton is also included in the theory, its mass  $m_\ell$  can be renormalized perturbatively in the usual way by imposing that its on-shell value coincides with the experimental one, i.e.  $m_\ell^\phi = m_\ell^{\text{PDG}} = 105.6583755$  (23) MeV [1]. The superscript  $\phi$  is used to denote quantities evaluated in the physical (QCD+QED) theory.

At a fixed value of the bare strong coupling  $g$ , we define the bare lattice quark masses in the full theory  $\hat{\mathbf{m}}^\phi$  by identifying  $N_f + 1$  dimensionful quantities, which we assume without loss of generality<sup>1</sup> to have mass dimension 1, namely  $M_1^\phi, \dots, M_{N_f}^\phi, \Lambda^\phi$ , and requiring the following ratios to take on the correct values when  $e$  is also at its physical value  $e^\phi$ ,

$$\left[ \frac{\hat{M}_j}{\hat{\Lambda}} \right]^2 (g, e^\phi, \hat{\mathbf{m}}^\phi) = \left( \frac{M_j^\phi}{\Lambda^\phi} \right)^2, \quad \text{for } j = 1, \dots, N_f. \quad (2.1)$$

Here, the  $\hat{M}_j$  and  $\hat{\Lambda}$  denote the same quantities evaluated in lattice units at the physical point  $(g, e^\phi, \hat{\mathbf{m}}^\phi)$ . For later use, we define this point as  $\boldsymbol{\sigma}^\phi = (g, e^\phi, \hat{\mathbf{m}}^\phi)$ . Note that the procedure for fixing  $\hat{\mathbf{m}}^\phi$  must be performed at every value of the coupling  $g$ , so in this sense we can think of the bare quark masses as a function of this coupling,  $\hat{\mathbf{m}}^\phi(g)$ . Moreover, the lattice quantities appearing on the left-hand side of eq. (2.1) are considered to be evaluated in the infinite volume limit. In practice, in the full QCD+QED theory, electromagnetic interactions can generate sizeable power-like finite-volume effects and should be removed, as discussed in section 3.3. Once the bare quark masses  $\hat{\mathbf{m}}^\phi$  are determined, we can predict any other quantity  $\hat{X}^\phi$  in lattice units and in the full theory as a function of  $g$ , namely

$$\hat{X}^\phi(g) = \hat{X}(g, e^\phi, \hat{\mathbf{m}}^\phi(g)). \quad (2.2)$$

---

<sup>1</sup>In practice one only requires  $N_f + 1$  quantities that can be used to form  $N_f$  independent dimensionless ratios.

From this, we can give a physical value to  $a$  as a function of  $g$  using a suitable dimensionful external input. For concreteness, we envision using  $\Lambda^\phi$

$$a(g) = \frac{\hat{\Lambda}^\phi(g)}{\Lambda^\phi}. \tag{2.3}$$

With this definition, one can predict dimensionful quantities as

$$X^\phi(g) = a(g)^{-[X]} \hat{X}^\phi(g), \tag{2.4}$$

where  $[X]$  is the mass dimension of  $X$ . As a consequence of asymptotic freedom, the limit  $g \rightarrow 0$  implies  $a(g) \rightarrow 0$ , and for renormalized quantities the equation above has a  $g \rightarrow 0$  limit which is cutoff independent. At non-zero  $g$  and  $a$  there is a family of choices that have the same continuum limit. For a given discretization of the QCD+QED action, this family is defined by (i) the set of  $N_f$  ratios that we use to define  $\hat{\mathbf{m}}(g)$ , and (ii) the physical quantity  $\Lambda^\phi$  that we use to set the scale.

In this calculation we employ three flavours of quarks, so we require four hadronic observables to fix the bare quark masses  $\mathbf{m} = (m_u, m_d, m_s)$  and the scale, which we choose to be  $\mathbf{M} = (M_1, M_2, M_3) = (m_{\pi^+}, m_{K^+}, m_{K^0})$  and  $\Lambda = m_{\Omega^-}$ . The physical values of such hadronic masses are taken as their experimental measurements, reported in the PDG [1], i.e.  $\mathbf{M}^\phi = (m_{\pi^+}^{\text{PDG}}, m_{K^+}^{\text{PDG}}, m_{K^0}^{\text{PDG}}) = (139.57039(18), 493.677(16), 497.611(13))$  MeV and  $\Lambda^\phi = m_{\Omega^-}^{\text{PDG}} = 1672.45(29)$  MeV.

## 2.2 Defining QCD and its isospin-symmetric limit

The calculation of IB corrections requires the definition of an isospin-symmetric limit of QCD. If we write the parameters of the bare QCD+QED Lagrangian as  $\sigma = (g, e, \hat{\mathbf{m}})$ , then the point  $\sigma^\phi = (g, e^\phi, \hat{\mathbf{m}}^\phi)$  identifies the full theory defined above in section 2.1. Points where  $e = 0$  correspond instead to the bare parameters for a QCD Lagrangian, that we denote as  $\sigma^{\text{QCD}} = (g^{\text{QCD}}, 0, \hat{\mathbf{m}}^{\text{QCD}})$ , where the vector  $\hat{\mathbf{m}}^{\text{QCD}} = (\hat{m}_u^{\text{QCD}}, \hat{m}_d^{\text{QCD}}, \hat{m}_s^{\text{QCD}})$  contains the QCD bare quark masses, with  $\hat{m}_u^{\text{QCD}} \neq \hat{m}_d^{\text{QCD}}$ . Therefore, we see that defining QCD within the full QCD+QED theory consists of choosing *one* point by imposing an additional renormalization condition. The same holds for iso-QCD theories, all belonging to the set identified by  $\sigma^{(0)} = (g^{(0)}, 0, \hat{\mathbf{m}}^{(0)})$ , with  $\hat{\mathbf{m}}^{(0)} = (\hat{m}_{ud}^{(0)}, \hat{m}_{ud}^{(0)}, \hat{m}_s^{(0)})$ , i.e. with  $\delta m = 0$ . Here we denoted the average light quark mass as  $m_{ud} = (m_u + m_d)/2$  and the up-down quark mass difference as  $\delta m = m_u - m_d$ . Again, the definition of a given iso-QCD theory requires imposing specific renormalization conditions. The possibility of choosing different renormalization prescriptions for the definition of QCD and iso-QCD translates into a *scheme dependence* in any observable computed in such theories.

Since there are many possible valid choices such that the IB corrections are small, there is no single generally accepted scheme and lattice collaborations have used in the past different prescriptions to define the iso-QCD theory [2]. Although no significant differences have been observed so far in the use of different schemes [2], when aiming at sub-percent precision calculations the ambiguities related to the different renormalization prescriptions adopted might become relevant when comparing results for (scheme-dependent) iso-QCD observables or IB effects. We therefore advocate that lattice collaborations be as clear and



transparent as possible in defining the scheme used to define iso-QCD in their calculations, and eventually agree on a common reference scheme.

In this work we adopt a scheme very similar to that introduced by the BMW Collaboration in refs. [16–18], which relies on the use of the neutral mesonic observables  $\mathbf{M}^{\text{BMW}} = (M_{\text{ud}}, \Delta M, M_{K\chi})$ , where

$$M_{\text{ud}}^2 = \frac{1}{2} \left( M_{\text{uu}}^2 + M_{\text{dd}}^2 \right), \quad \Delta M^2 = M_{\text{uu}}^2 - M_{\text{dd}}^2, \quad M_{K\chi}^2 = \frac{1}{2} \left( m_{K^+}^2 + m_{K^0}^2 - m_{\pi^+}^2 \right), \quad (2.5)$$

and  $M_{\text{qq}}^2$  denotes the squared mass of the connected  $\bar{q}q$  neutral pseudoscalar mesons. One can show that the leading-order partially-quenched chiral corrections to such quantities are given by [16, 19]

$$M_{\text{ud}}^2 = 2Bm_{\text{ud}}^{\text{QCD,R}} + \dots, \quad \Delta M^2 = 2B\delta m^{\text{QCD,R}} + \dots, \quad M_{K\chi}^2 = 2Bm_s^{\text{QCD,R}} + \dots, \quad (2.6)$$

where  $B$  is the QCD chiral condensate, the superscript R denotes the QCD-renormalized quark masses in a given scheme (e.g.  $\overline{\text{MS}}$  at 2 GeV) and the ellipses are next-to-leading order SU(3) chiral corrections [19]. An important feature of this choice of variables is the systematic absence of  $\mathcal{O}(\alpha_{\text{em}})$  corrections at leading order. Therefore, these specific meson masses are expected to be dominated by the part proportional to the quark masses, allowing comparisons with quark-mass schemes like the Gasser-Rusetsky-Scimemi (GRS) one [20] without the need of using short-distance schemes (e.g. the  $\overline{\text{MS}}$  scheme) to renormalize the quark masses.

In this scheme we assume that the bare strong coupling of any unphysical theory identified by the vector  $\sigma^*$  is kept equal to that of the full theory, i.e.  $\sigma^* = (g, e^*, \mathbf{m}^*)$ . The bare quark masses of QCD and iso-QCD are then obtained by imposing the following renormalization conditions, respectively

$$\left[ \frac{\hat{\mathbf{M}}^{\text{BMW}}}{\hat{m}_\Omega} \right]^2 (g, 0, \hat{\mathbf{m}}^{\text{QCD}}) = \left( \left[ \frac{\hat{M}_{\text{ud}}}{\hat{m}_\Omega} \right]^2 (g, e^\phi, \hat{\mathbf{m}}^\phi), \left[ \frac{\Delta \hat{M}}{\hat{m}_\Omega} \right]^2 (g, e^\phi, \hat{\mathbf{m}}^\phi), \left[ \frac{\hat{M}_{K\chi}}{\hat{m}_\Omega} \right]^2 (g, e^\phi, \hat{\mathbf{m}}^\phi) \right), \quad (2.7)$$

$$\left[ \frac{\hat{\mathbf{M}}^{\text{BMW}}}{\hat{m}_\Omega} \right]^2 (g, 0, \hat{\mathbf{m}}^{(0)}) = \left( \left[ \frac{\hat{M}_{\text{ud}}}{\hat{m}_\Omega} \right]^2 (g, e^\phi, \hat{\mathbf{m}}^\phi), 0, \left[ \frac{\hat{M}_{K\chi}}{\hat{m}_\Omega} \right]^2 (g, e^\phi, \hat{\mathbf{m}}^\phi) \right). \quad (2.8)$$

The above conditions can be extended beyond  $N_f = 3$  flavours by choosing ratios of hadron masses with a dependence on the heavier quarks. The lattice spacings for the two theories can be evaluated by imposing the additional conditions

$$a^{\text{QCD}}(g) = \frac{\hat{m}_\Omega(g, 0, \hat{\mathbf{m}}^{\text{QCD}})}{m_{\Omega^-}^{\text{PDG}}} \quad \text{and} \quad a^{(0)}(g) = \frac{\hat{m}_\Omega(g, 0, \hat{\mathbf{m}}^{(0)})}{m_{\Omega^-}^{\text{PDG}}}. \quad (2.9)$$

In principle, one could also impose the simpler condition  $a^{\text{QCD}}(g) = a^{(0)}(g)$ . Since the UV divergences of the two theories do not depend on quark masses, the difference between the two approaches would result in cut-off effects that eventually vanish in the continuum limit. However, in this work, the quantities we aim to compute are dimensionless, and therefore we will not need, in practice, to make dimensions explicit in the (iso-)QCD theory.

### 2.3 Computing isospin-breaking effects on the lattice

Once the bare parameters of both the full theory and iso-QCD have been determined through the renormalization procedure described above, one can define the expansion of a QCD+QED observable  $\hat{X}(\boldsymbol{\sigma}^\phi)$  around the iso-QCD point  $\boldsymbol{\sigma}^{(0)}$  as

$$\hat{X}(\boldsymbol{\sigma}^\phi) = \hat{X}(\boldsymbol{\sigma}^{(0)}) + \delta\hat{X}(\boldsymbol{\sigma}^{(0)}) + \mathcal{O}(\varepsilon^2), \quad (2.10)$$

where the quantity  $\delta\hat{X}(\boldsymbol{\sigma}^{(0)})$  encodes the leading order IB corrections of  $\mathcal{O}(\varepsilon)$  relative to the specific iso-QCD point  $\boldsymbol{\sigma}^{(0)}$ . Suppose now that the lattice setup used for the calculation has been tuned to some iso-QCD point  $\boldsymbol{\sigma} = (g, 0, \hat{\mathbf{m}})$ , close to the physical point  $\boldsymbol{\sigma}^\phi$ . In practice this point is generally the result of a simulation-parameter tuning procedure and might differ from the iso-QCD point defined in the previous section, i.e.  $\boldsymbol{\sigma} \neq \boldsymbol{\sigma}^{(0)}$ . However, we assume that this simulation point is sufficiently close to  $\boldsymbol{\sigma}^\phi$ ,  $\boldsymbol{\sigma}^{(0)}$ , and  $\boldsymbol{\sigma}^{\text{QCD}}$  that the value of  $\hat{X}$  at any of those points can be described accurately enough by a linear correction to the simulation point. This is a fairly strong assumption that would not be valid in a number of lattice calculations, particularly when working at non-physical quark masses where non-linear corrections are expected to be sizeable. However, as we will demonstrate in section 5.2, this applies to the close-to-physical point simulations used in this work. More explicitly, under this linearity assumption the physical value of  $\hat{X}$  can be written as

$$\hat{X}(\boldsymbol{\sigma}^\phi) = \hat{X}(\boldsymbol{\sigma}) + \alpha_{\text{em}}^\phi \frac{\partial \hat{X}}{\partial \alpha_{\text{em}}}(\boldsymbol{\sigma}) + \sum_{q=1}^{N_f} (\hat{m}_q^\phi - \hat{m}_q) \frac{\partial \hat{X}}{\partial \hat{m}_q}(\boldsymbol{\sigma}) + \mathcal{O}(\bar{\varepsilon}^2), \quad (2.11)$$

where the  $\mathcal{O}(\bar{\varepsilon}^2)$  represent any second order corrections in  $\hat{m}_q^\phi - \hat{m}_q$  and  $\alpha_{\text{em}}$ , which we consider to be of similar size to higher-order isospin-breaking corrections  $\mathcal{O}(\varepsilon^2)$  that are taken to be negligible and are discarded throughout this work. This expanded power-counting is necessary as some isospin-symmetric parameters like  $m_s$  and  $m_{\text{ud}}$  might be slightly mistuned at the simulation point. If we consider  $\hat{X}$  to be the ratio of hadron masses in eq. (2.1), by using lattice data for  $\hat{X}(\boldsymbol{\sigma})$  and its derivatives, and applying the linear equation above, we can solve the resulting system to obtain the lattice bare quark masses for the physical point  $\hat{\mathbf{m}}^\phi$ .

A similar equation is obtained by linearizing the first term on the right-hand side of eq. (2.10),

$$\hat{X}(\boldsymbol{\sigma}^{(0)}) = \hat{X}(\boldsymbol{\sigma}) + \sum_{q=1}^{N_f} (\hat{m}_q^{(0)} - \hat{m}_q) \frac{\partial \hat{X}}{\partial \hat{m}_q}(\boldsymbol{\sigma}) + \mathcal{O}(\bar{\varepsilon}^2), \quad (2.12)$$

and can be solved for  $\hat{\mathbf{m}}^{(0)}$  applying the renormalization conditions in eq. (2.8). Combining eqs. (2.11) and (2.12), the isospin-breaking correction in eq. (2.10) can be identified as

$$\delta\hat{X}(\boldsymbol{\sigma}^{(0)}) = \alpha_{\text{em}}^\phi \frac{\partial \hat{X}}{\partial \alpha_{\text{em}}}(\boldsymbol{\sigma}) + \sum_{q=1}^{N_f} (\hat{m}_q^\phi - \hat{m}_q^{(0)}) \frac{\partial \hat{X}}{\partial \hat{m}_q}(\boldsymbol{\sigma}). \quad (2.13)$$

The QCD masses and  $\hat{X}(\boldsymbol{\sigma}^{\text{QCD}})$  can be determined using analogous linear expansions.

We can then define a separation of  $\delta\hat{X}(\boldsymbol{\sigma}^{(0)})$  into a contribution due to the strong isospin breaking (SIB) and another due to electromagnetic interactions, namely

$$\delta\hat{X}^{\text{SIB}}(\boldsymbol{\sigma}^{(0)}) = \hat{X}(\boldsymbol{\sigma}^{\text{QCD}}) - \hat{X}(\boldsymbol{\sigma}^{(0)}) = \sum_{q=1}^{N_f} (\hat{m}_q^{\text{QCD}} - \hat{m}_q^{(0)}) \frac{\partial\hat{X}}{\partial\hat{m}_q}(\boldsymbol{\sigma}), \quad (2.14)$$

$$\delta\hat{X}^\gamma(\boldsymbol{\sigma}^{(0)}) = \hat{X}(\boldsymbol{\sigma}^\phi) - \hat{X}(\boldsymbol{\sigma}^{\text{QCD}}) = \alpha_{\text{em}}^\phi \frac{\partial\hat{X}}{\partial\alpha_{\text{em}}}(\boldsymbol{\sigma}) + \sum_{q=1}^{N_f} (\hat{m}_q^\phi - \hat{m}_q^{\text{QCD}}) \frac{\partial\hat{X}}{\partial\hat{m}_q}(\boldsymbol{\sigma}), \quad (2.15)$$

such that

$$\hat{X}(\boldsymbol{\sigma}^\phi) = \hat{X}(\boldsymbol{\sigma}^{(0)}) + \delta\hat{X}^{\text{SIB}}(\boldsymbol{\sigma}^{(0)}) + \delta\hat{X}^\gamma(\boldsymbol{\sigma}^{(0)}) + \mathcal{O}(\varepsilon^2). \quad (2.16)$$

The definition of isospin-breaking effects provided in eqs. (2.10) and (2.13) for hadronic observables in lattice units can be rewritten in terms of quantities expressed in physical units making use of eqs. (2.3) and (2.4). For this discussion we make the arbitrary (but simple) choice that the same dimensionful quantity  $\Lambda^\phi$  adopted in eq. (2.3) to define the scale in the full theory is used to define the scale in all other theories (as done in eq. (2.9) above). This gives (see also eq. (B15) of ref. [8])

$$X(\boldsymbol{\sigma}^\phi) = (a^{(0)})^{-[X]} \left[ \hat{X}(\boldsymbol{\sigma}^{(0)}) + \delta\hat{X}(\boldsymbol{\sigma}^{(0)}) - [X] \frac{a^\phi - a^{(0)}}{a^{(0)}} \hat{X}(\boldsymbol{\sigma}^{(0)}) \right] + \mathcal{O}(\varepsilon^2),$$

where the isospin-breaking correction to the lattice spacing amounts to

$$\frac{\delta a}{a^{(0)}} = \frac{a^\phi - a^{(0)}}{a^{(0)}} = \frac{1}{\hat{\Lambda}(\boldsymbol{\sigma})} \left[ \alpha_{\text{em}}^\phi \frac{\partial\hat{\Lambda}}{\partial\alpha_{\text{em}}}(\boldsymbol{\sigma}) + \sum_{q=1}^{N_f} (\hat{m}_q^\phi - \hat{m}_q^{(0)}) \frac{\partial\hat{\Lambda}}{\partial\hat{m}_q}(\boldsymbol{\sigma}) \right].$$

As already stressed at the end of the previous section, in this work we compute dimensionless quantities ( $[X] = 0$ ), and hence the corrections to the lattice spacing will play no role in the calculation.

From the above discussion it is clear that in order to determine IB corrections one needs to evaluate numerically the linear coefficients of the expansion of a given observable in terms of  $\alpha_{\text{em}}$  and quark mass shifts. The QED effects have been computed in the past by different collaborations following two approaches. On the one hand, one can include QED gauge links in the fermion operator to be inverted and produce QCD+QED quark propagators to construct hadronic correlation functions, as introduced first in ref. [21], and used in a wide range of lattice calculations [2]. On the other hand, it is possible to obtain the same corrections by perturbatively expanding the path integral for  $\hat{X}$  with respect to  $\alpha_{\text{em}}$ , with the result of evaluating diagrams with electromagnetic current insertions, as originally proposed by the RM123 collaboration in refs. [22, 23]. In the following, the latter method will be adopted to determine the IB corrections to hadron masses and to the ratio of the leptonic decay rates of kaons and pions into muons. The implementation of the method will be discussed in section 4.

## 2.4 Scheme ambiguities

The isospin decomposition in eq. (2.16), as discussed in section 2.2, depends on the prescriptions in eqs. (2.7) and (2.8), although the physical observable,  $\hat{X}(\boldsymbol{\sigma}^\phi)$ , is not. Varying

these prescriptions will lead to different bare masses  $\hat{\mathbf{m}}_q^{\text{QCD}}$  and  $\hat{\mathbf{m}}_q^{(0)}$ , which generate the scheme-dependence of  $\delta\hat{X}^{\text{SIB}}$  and  $\delta\hat{X}^\gamma$  through eqs. (2.14) and (2.15). In principle this ambiguity is an  $\mathcal{O}(\varepsilon)$  effect on  $\hat{\mathbf{m}}_q^{\text{QCD}}$  and  $\hat{\mathbf{m}}_q^{(0)}$ , and therefore the ambiguity on the isosymmetric component  $\hat{X}(\boldsymbol{\sigma}^{(0)})$  can potentially be as large as the isospin-breaking corrections  $\delta\hat{X}(\boldsymbol{\sigma}^{(0)})$  themselves. It is therefore important to identify classes of schemes which are phenomenologically relevant with a minimal level of ambiguity.

For instance, the GRS scheme [20] assumes that the renormalized quark masses and the renormalized strong coupling constant, in a given scheme and at a chosen renormalization scale, are kept constant for any value of  $\alpha_{\text{em}}$ . This scheme itself depends then on a choice of renormalization procedure for these quantities. The GRS scheme is generally the prescription used in phenomenological calculations, such as chiral perturbation theory predictions for weak decay rates including the ones discussed in this paper [4–6]. It is therefore attractive for lattice calculations to use prescriptions which produce predictions close to the GRS scheme. Defining precisely such a class of schemes is a rich technical topic which is relevant for precision physics, and will be the topic of a future publication based on the lattice data presented in this paper. The BMW variables chosen here are designed to be hadronic quantities providing an isospin decomposition close to the GRS scheme, and in section 5.2 we will explicitly check that by comparing with existing results from ref. [8].

### 3 Matrix elements from Euclidean correlation functions

As discussed in section 1, the inclusion of IB corrections in the calculation of decay rates is complicated by the appearance of IR divergences, generated by  $\mathcal{O}(\alpha_{\text{em}})$  QED corrections to the decay amplitude. In this calculation we adopt the RM123S method of eq. (1.2) to regularize IR divergences. We choose two separate regulators for the two terms. In particular, we compute the virtual decay rate on the lattice using the finite volume with the QED<sub>L</sub> prescription as an IR regulator, while the real decay width is evaluated in perturbation theory using a photon mass, namely

$$\Gamma(P^\pm \rightarrow \ell^\pm \nu[\gamma]) = \lim_{L \rightarrow \infty} [\Gamma_0(L) - \Gamma_0^{\text{uni}}(L)] + \lim_{m_\gamma \rightarrow 0} [\Gamma_0^{\text{uni}}(m_\gamma) + \Gamma_1(\omega_\gamma^{\text{max}}, m_\gamma)] . \quad (3.1)$$

The first bracketed term removes the universal (structure-independent) logarithmic IR divergence and finite volume effects (FVE) up to  $\mathcal{O}(1/L)$  [24, 25]. Recently, the  $\mathcal{O}(1/L^2)$  corrections to  $\Gamma_0(L)$  have been calculated in QED<sub>L</sub>, including structure-dependent contributions, in ref. [14]. Thus, we can extend eq. (3.1) to

$$\Gamma(P^+ \rightarrow \mu^+ \nu_\mu[\gamma]) = \lim_{L \rightarrow \infty} [\Gamma_0(L) - \Gamma_0^{(2)}(L)] + \lim_{m_\gamma \rightarrow 0} [\Gamma_0^{\text{uni}}(m_\gamma) + \Gamma_1(\omega_\gamma^{\text{max}}, m_\gamma)] , \quad (3.2)$$

where  $\Gamma_0^{(2)}(L)$  contains the finite-volume effects up to  $\mathcal{O}(1/L^2)$  and will be discussed in detail in section 3.3. Here, we only note that the residual finite-volume effects in the quantity  $\Gamma_0(L) - \Gamma_0^{(2)}(L)$  now begin at  $\mathcal{O}(1/L^3)$ . The second bracketed term in eq. (3.2) has been instead calculated analytically in ref. [10] and is reported below in section 3.4. For convenience, we distinguish the two contributions, the one computed on the lattice

with a finite volume and the one evaluated in perturbation theory with a massive photon, respectively as

$$\Gamma_P^{\text{latt}}(L) \equiv [\Gamma_0(L) - \Gamma_0^{(2)}(L)] \quad \text{and} \quad \Gamma_P^{\text{pert}}(m_\gamma) \equiv [\Gamma_0^{\text{uni}}(m_\gamma) + \Gamma_1(\omega_\gamma^{\text{max}}, m_\gamma)]. \quad (3.3)$$

We can expand these expressions at leading order in the IB corrections

$$\Gamma_P^{\text{latt}}(L) = \Gamma_P^{\text{tree}} \left[ 1 + \delta R_P^{\text{latt}}(L) - \delta R_P^{(2)}(L) \right] + \mathcal{O}(\varepsilon^2), \quad (3.4)$$

$$\Gamma_P^{\text{pert}}(m_\gamma) = \Gamma_P^{\text{tree}} \delta R_P^{\text{pert}}(\omega_\gamma^{\text{max}}, m_\gamma) + \mathcal{O}(\varepsilon^2), \quad (3.5)$$

having defined

$$\Gamma_0(L) = \Gamma_P^{\text{tree}} \left[ 1 + \delta R_P^{\text{latt}}(L) \right] + \mathcal{O}(\varepsilon^2) \quad \text{and} \quad \Gamma_0^{(2)}(L) = \Gamma_P^{\text{tree}} \left[ 1 + \delta R_P^{(2)}(L) \right] + \mathcal{O}(\varepsilon^2). \quad (3.6)$$

It follows that, writing the leptonic decay rate as

$$\Gamma(P^+ \rightarrow \mu^+ \nu_\mu[\gamma]) = \Gamma_P^{\text{tree}}(1 + \delta R_P) + \mathcal{O}(\varepsilon^2) \quad (3.7)$$

its leading IB correction is given by

$$\delta R_P = \lim_{L \rightarrow \infty} \left[ \delta R_P^{\text{latt}}(L) - \delta R_P^{(2)}(L) \right] + \lim_{m_\gamma \rightarrow 0} \delta R_P^{\text{pert}}(\omega_\gamma^{\text{max}}, m_\gamma). \quad (3.8)$$

The outline of the rest of the section is as follows. In section 3.1 we give our definition of  $\Gamma_P^{\text{tree}}$  and derive the corresponding correction  $\delta R_P^{\text{latt}}(L)$  at finite volume in terms of IB corrections to matrix elements and meson masses. Then, in section 3.2 we describe how to obtain such corrections starting from Euclidean lattice correlation functions. The subtraction of QED<sub>L</sub> finite-volume effects  $\delta R_P^{(2)}(L)$  and the calculation of the real photon emission in perturbation theory are then discussed in sections 3.3 and 3.4, respectively.

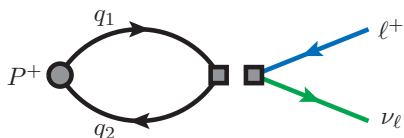
### 3.1 Virtual corrections to the leptonic decay rate

We focus here on the determination of the matrix element associated to the virtual decay of a positive pseudoscalar meson,  $P^+ \rightarrow \ell^+ \nu_\ell$ , without including real photons in the final state. As this is an IR divergent quantity, we assume that an IR regulator is in place throughout the section. For concreteness, we regulate IR divergences on a finite volume of size  $L$  adopting the QED<sub>L</sub> prescription [12] to remove the spatial zero modes of the lattice photon propagator (see section 3.3). At the lowest order in QED and QCD, pseudoscalar mesons decay into a lepton-neutrino pair via the exchange of a  $W$ -boson between the constituent quarks of the meson and the leptons. Since for both pions and kaons the process  $P^+ \rightarrow \ell^+ \nu_\ell$  has a momentum transfer much smaller than the  $W$ -boson mass  $m_W$ , we can study it in an effective theory with a local four-fermion interaction described by the effective Hamiltonian

$$\mathcal{H}_W = \frac{G_F}{\sqrt{2}} V_{q_1 q_2}^* O_W \quad (3.9)$$

where  $G_F$  is the Fermi constant and  $V_{q_1 q_2}^*$  the relevant CKM matrix element. The four-fermion operator mediating the process is

$$O_W \equiv J_H^\rho J_L^\rho = (\bar{q}_2 \gamma^\rho (1 - \gamma_5) q_1) (\bar{\nu}_\ell \gamma^\rho (1 - \gamma_5) \ell), \quad (3.10)$$



**Figure 1.** Feynman diagram of the tree-level contribution to the weak decay of a positive pseudoscalar meson  $P^+ \in \{\pi^+, K^+\}$  into a lepton-neutrino pair. The double-square vertex represents the effective weak Hamiltonian eq. (3.9).

with  $q_1$  being a  $u$ -type quark and  $q_2$  a  $d$ -type quark and  $J_H^\rho$  and  $J_L^\rho$  denoting the weak  $(V - A)$  hadronic and leptonic currents, respectively. The Feynman diagram associated with this tree-level term is represented by figure 1.

When including QED at  $O(\alpha_{\text{em}})$ , the UV corrections to matrix elements of the local operator  $O_W$  differ from those of the Standard Model and a matching between the two theories is therefore needed. This is usually performed in the so-called  $W$ -regularization [26, 27], and we refer to refs. [8, 10] for detailed discussions on the argument. After the inclusion of QCD and QED at  $O(\alpha_{\text{em}})$  and assuming that chiral symmetry is preserved, the effective Hamiltonian reads<sup>2</sup>

$$\mathcal{H}_W = \frac{G_F}{\sqrt{2}} V_{q_1 q_2}^* \left( \mathcal{Z}_0 + \frac{\alpha_{\text{em}}}{4\pi} \delta\mathcal{Z} \right) O_W. \tag{3.11}$$

Here  $\mathcal{Z}_0$  is the non-perturbative QCD renormalization constant of the operator  $O_W$ . The quantity  $\delta\mathcal{Z}$  encodes instead the short-distance matching between the effective theory in the  $W$ -regularization and the Standard Model, as well as the electromagnetic corrections to the matching of the four-fermion operator  $O_W$  renormalized non-perturbatively in a given scheme to the  $W$ -regularization one. If  $O_W$  is a lattice operator and the regularization used for the fermionic action introduces an explicit chiral symmetry breaking, then the operator  $O_W$  undergoes an additive renormalization due to the mixing with other lattice operators with different chirality and the mixing pattern would be more complicated than that in eq. (3.11) (see e.g. refs. [8, 10]). In the lattice calculation presented in this work, however, chiral fermions are employed and therefore in the following we will consider the operator  $O_W$  renormalizing multiplicatively as in eq. (3.11), with  $\mathcal{Z}_0 = Z_V = Z_A$ . Moreover, if a mass-independent scheme is adopted to renormalize the four-fermion operator, then the quantities  $\mathcal{Z}_0$  and  $\delta\mathcal{Z}$  will be the same regardless of the masses of the particles involved in the process. As a consequence, the contribution of the electromagnetic corrections proportional to  $\delta\mathcal{Z}$  will cancel in the calculation of our quantity of interest,  $\delta R_{K\pi} = \delta R_K - \delta R_\pi$ , entering eq. (1.3).

In the full theory the (IR regulated) virtual decay rate can be written as

$$\Gamma_0(L) = \mathcal{K} |\mathcal{M}_P|^2, \tag{3.12}$$

where  $\mathcal{K}$  is a factor containing the electro-weak coupling, the CKM matrix elements and the integration over the two-body phase space, while  $|\mathcal{M}_P|^2 = \sum_{r,s} |\mathcal{M}_P^{rs}|^2$  is the

---

<sup>2</sup>When including electromagnetic corrections at  $O(\alpha_{\text{em}})$ , the Fermi constant  $G_F$  has to be defined accordingly. This is conventionally obtained from the muon lifetime including one-loop electromagnetic corrections and reads  $G_F = 1.16634 \times 10^{-5} \text{ GeV}^{-2}$  [28, 29].

magnitude squared of the QCD+QED virtual amplitude summed over the lepton and neutrino polarisations  $r$  and  $s$ . In the rest frame of the decaying meson  $P$  ( $\mathbf{p}_P = \mathbf{0}$ ) the on-shell lepton and neutrino (Euclidean) momenta,  $p_\ell = (i\omega_\ell, \mathbf{p}_\ell)$  and  $p_\nu = (i\omega_\nu, \mathbf{p}_\nu)$ , are such that  $\mathbf{p}_\ell + \mathbf{p}_\nu = \mathbf{0}$  and the decay rate can be written in terms of

$$\mathcal{K} = \frac{G_F^2}{16\pi} |V_{q_1 q_2}|^2 \frac{1}{2m_P} \left(1 - \frac{m_\ell^2}{m_P^2}\right), \quad (3.13)$$

and the renormalized QCD+QED matrix element

$$\mathcal{M}_P^{rs}(\mathbf{p}_\ell) = \mathcal{Z} \overline{\mathcal{M}}_P^{rs}(\mathbf{p}_\ell) = \mathcal{Z} \langle \ell^+, r, \mathbf{p}_\ell; \nu_\ell, s, \mathbf{p}_\nu | O_W | P^+, \mathbf{0} \rangle^\phi. \quad (3.14)$$

Here  $\overline{\mathcal{M}}_P^{rs}(\mathbf{p}_\ell)$  is the bare matrix element computed in the full QCD+QED theory, as indicated by the superscript  $\phi$ , while the factor  $\mathcal{Z} = \mathcal{Z}_0 + \frac{\alpha}{4\pi} \delta\mathcal{Z}$  denotes the renormalization constant of the weak operator  $O_W$  entering the effective Hamiltonian of eq. (3.11). Note that  $|\mathcal{M}_P(\mathbf{p}_\ell)|^2 = |\mathcal{M}_P(|\mathbf{p}_\ell|)|^2$  is a rotationally symmetric function of the lepton momentum  $\mathbf{p}_\ell$  in the continuum and infinite-volume limit. Energy conservation ( $m_P = \omega_\ell + \omega_\nu$ ) has also been employed to rewrite

$$|\mathbf{p}_\ell| = \omega_\nu = \frac{m_P}{2} \left(1 - \frac{m_\ell^2}{m_P^2}\right) \quad \text{and} \quad \omega_\ell = \frac{m_P}{2} \left(1 + \frac{m_\ell^2}{m_P^2}\right). \quad (3.15)$$

We may expand now the QCD+QED squared matrix element  $|\mathcal{M}_P(\mathbf{p}_\ell)|^2$  in eq. (3.12) around the iso-symmetric QCD point keeping the lepton momentum to its physical value as

$$|\mathcal{M}_P(\mathbf{p}_\ell)|^2 = |\mathcal{M}_P^{(0)}(\mathbf{p}_\ell)|^2 + \delta|\mathcal{M}_P(\mathbf{p}_\ell)|^2 + \mathcal{O}(\varepsilon^2). \quad (3.16)$$

In iso-QCD the matrix element factorizes into a hadronic and leptonic part, namely

$$\begin{aligned} |\mathcal{M}_P^{(0)}(\mathbf{p}_\ell)|^2 &= \mathcal{Z}_0^2 |\langle \ell^+, r, \mathbf{p}_\ell; \nu_\ell, s, \mathbf{p}_\nu | O_W | P, \mathbf{0} \rangle^{(0)}|^2 \\ &= \mathcal{Z}_0^2 |\mathcal{A}_P^{(0)}|^2 |\mathcal{L}(\mathbf{p}_\ell)|^2, \end{aligned} \quad (3.17)$$

where

$$\mathcal{Z}_0 \mathcal{A}_P^{(0)} \equiv -\mathcal{Z}_0 \langle 0 | J_H^0 | P, \mathbf{0} \rangle^{(0)} = \mathcal{Z}_0 \langle 0 | \bar{q}_2 \gamma^0 \gamma_5 q_1 | P, \mathbf{0} \rangle^{(0)} = i m_P^{(0)} f_P \quad (3.18)$$

is the iso-QCD renormalized axial matrix element expressed in terms of the mass  $m_P^{(0)}$  of the meson state  $|P, \mathbf{0}\rangle^{(0)}$  and the decay constant  $f_P$ , while

$$\mathcal{L}^{rs}(\mathbf{p}_\ell) = \langle \ell^+, r, \mathbf{p}_\ell; \nu_\ell, s, \mathbf{p}_\nu | J_L^0 | 0 \rangle^{(0)} = \bar{u}_\nu^r(\mathbf{p}_\nu) \gamma^0 (1 - \gamma_5) v_\ell^s(\mathbf{p}_\ell) \quad (3.19)$$

is the tree-level leptonic tensor with  $v_\ell^r(\mathbf{p}_\ell) = \langle \ell^+, r, \mathbf{p}_\ell | \ell | 0 \rangle^{(0)}$  and  $\bar{u}_\nu^r(\mathbf{p}_\nu) = \langle \nu_\ell, s, \mathbf{p}_\nu | \bar{\nu}_\ell | 0 \rangle^{(0)}$  the free Dirac spinors. We have considered here only the  $\rho = 0$  component of  $O_W$  (eq. (3.10)) since this is the only one contributing to the axial matrix element in the meson rest frame. Using the completeness relations for spinors in Euclidean space

$$\sum_{r,r'} v_\ell^{r'}(\mathbf{p}_\ell) \bar{v}_\ell^r(\mathbf{p}_\ell) = -i \not{p}_\ell - m_\ell, \quad \sum_{s,s'} u_\nu^s(\mathbf{p}_\nu) \bar{u}_\nu^{s'}(\mathbf{p}_\nu) = -i \not{p}_\nu, \quad (3.20)$$

one gets

$$|\mathcal{L}(\mathbf{p}_\ell)|^2 = \sum_{r,s} |\mathcal{L}^{rs}(\mathbf{p}_\ell)|^2 = 8|\mathbf{p}_\ell|(\omega_\ell - |\mathbf{p}_\ell|) = 4m_\ell^2 \left(1 - \frac{m_\ell^2}{m_P^2}\right) \quad (3.21)$$

and hence

$$|\mathcal{M}_P^{(0)}(\mathbf{p}_\ell)|^2 = \mathcal{Z}_0^2 |\mathcal{A}_P^{(0)}|^2 |\mathcal{L}(\mathbf{p}_\ell)|^2 = 4m_\ell^2 \left(1 - \frac{m_\ell^2}{m_P^2}\right)^2 m_P^{(0)2} f_P^2. \quad (3.22)$$

Following the convention of the PDG [1], we define the “tree-level” decay rate as

$$\Gamma_P^{\text{tree}} = \mathcal{K} |\mathcal{M}_P^{\text{tree}}|^2 = \mathcal{K} \left(\frac{m_P}{m_P^{(0)}}\right)^2 |\mathcal{M}_P^{(0)}|^2 = \frac{G_F^2}{8\pi} |V_{q_1 q_2}|^2 m_\ell^2 \left(1 - \frac{m_\ell^2}{m_P^2}\right)^2 m_P f_P^2, \quad (3.23)$$

i.e. with all masses defined in the full theory and only the decay constant  $f_P$  defined in iso-QCD. Combining the above eqs. (3.14), (3.16) and (3.23) with eq. (3.4) we obtain

$$\delta R_P^{\text{latt}} = \frac{\delta |\mathcal{M}_P(\mathbf{p}_\ell)|^2}{|\mathcal{M}_P^{(0)}(\mathbf{p}_\ell)|^2} - 2 \frac{\delta m_P}{m_P^{(0)}} = 2 \left( \frac{\delta \mathcal{A}_P}{\mathcal{A}_P^{(0)}} - \frac{\delta m_P}{m_P^{(0)}} + \frac{\delta \mathcal{Z}}{\mathcal{Z}_0} \right), \quad (3.24)$$

where we have defined the leading IB corrections to the meson mass  $\delta m_P$ ,

$$m_P = m_P^{(0)} + \delta m_P + \mathcal{O}(\varepsilon^2), \quad (3.25)$$

and those to the bare matrix element as

$$\frac{\delta \mathcal{A}_P}{\mathcal{A}_P^{(0)}} \equiv \text{Re} \left\{ - \frac{\sum_{r,s} \delta \bar{\mathcal{M}}_P^{rs}(\mathbf{p}_\ell) [\mathcal{L}^{rs}(\mathbf{p}_\ell)]^\dagger}{\mathcal{A}_P^{(0)} |\mathcal{L}(\mathbf{p}_\ell)|^2} \right\}. \quad (3.26)$$

As discussed above, the quantity  $\delta \mathcal{Z}/\mathcal{Z}_0$  does not depend on the masses of the decaying meson and hence our target quantity  $\delta R_{K\pi}^{\text{latt}}$  is given by

$$\delta R_{K\pi}^{\text{latt}} = 2 \left( \frac{\delta \mathcal{A}_K}{\mathcal{A}_K^{(0)}} - \frac{\delta m_K}{m_K^{(0)}} \right) - 2 \left( \frac{\delta \mathcal{A}_\pi}{\mathcal{A}_\pi^{(0)}} - \frac{\delta m_\pi}{m_\pi^{(0)}} \right). \quad (3.27)$$

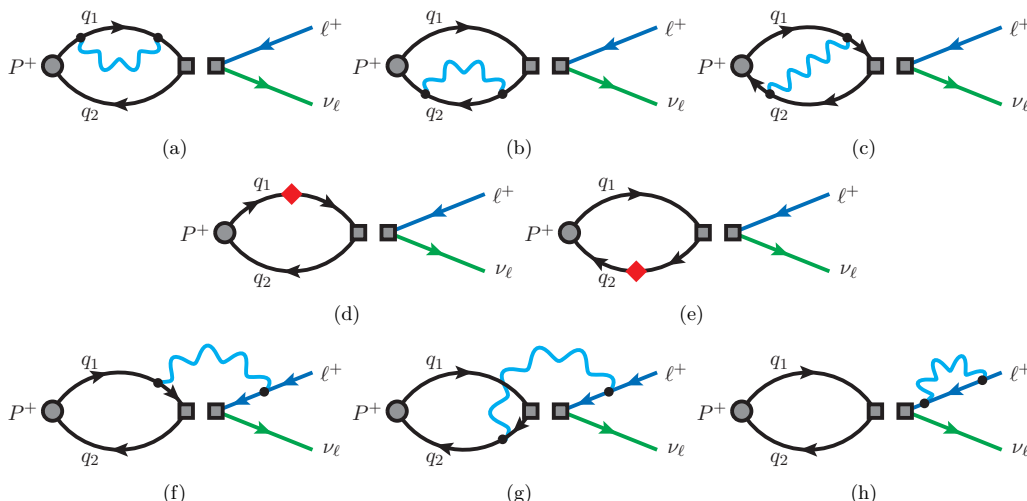
We can distinguish three kinds of corrections to the matrix element, that we denote as

$$\delta \mathcal{A}_P = \delta \mathcal{A}_P^f + \delta \mathcal{A}_P^{\text{nf}} + \delta \mathcal{A}_P^\ell. \quad (3.28)$$

The first term contains corrections involving only the quarks and these are proportional to either the quark fractional charges or to the bare quark mass splittings. These are obtained from the corrections to the bare matrix element

$$\begin{aligned} [\delta \bar{\mathcal{M}}_P^{rs}(\mathbf{p}_\ell)]^f &= \left[ \frac{1}{2} \sum_{q,q'} e_q e_{q'} \frac{\partial^2}{\partial e_q \partial e_{q'}} + \sum_q (\hat{m}_q^\phi - \hat{m}_q^{(0)}) \frac{\partial}{\partial \hat{m}_q} \right] \bar{\mathcal{M}}_P^{rs}(\mathbf{p}_\ell) \Big|_{\sigma^{(0)}} \\ &= -\mathcal{L}^{rs}(\mathbf{p}_\ell) \left[ \frac{1}{2} \sum_{q,q'} e_q e_{q'} \frac{\partial^2}{\partial e_q \partial e_{q'}} + \sum_q (\hat{m}_q^\phi - \hat{m}_q^{(0)}) \frac{\partial}{\partial \hat{m}_q} \right] \mathcal{A}_P^\phi \Big|_{\sigma^{(0)}}, \end{aligned} \quad (3.29)$$





**Figure 2.** Quark-connected Feynman diagrams contributing to the leading IB corrections to the weak decay. The wiggly lines correspond to photons, and the diamond-shaped vertices are scalar insertions.

with  $e_{q_1} = +2/3|e|$  and  $e_{q_2} = -1/3|e|$ , and  $\mathcal{A}_P^\phi$  the axial matrix element evaluated in the full theory.  $\sigma^{(0)}$  indicates that the quantities are evaluated in the target iso-QCD theory,  $\sigma^{(0)} = (g, 0, \hat{\mathbf{m}}^{(0)})$  as discussed in section 2.2. Since in this case the decay amplitude factorizes into a hadronic and a leptonic part, we refer to these contributions as “factorizable”. The relevant diagrams contributing to these corrections are depicted in figures 2(a)–2(e) and 3(a)–3(e). The second term in eq. (3.28) corresponds instead to the “non-factorizable” corrections to the matrix element where a photon is exchanged between a quark and the charged lepton ( $e_\ell = -|e|$ ). These are given by

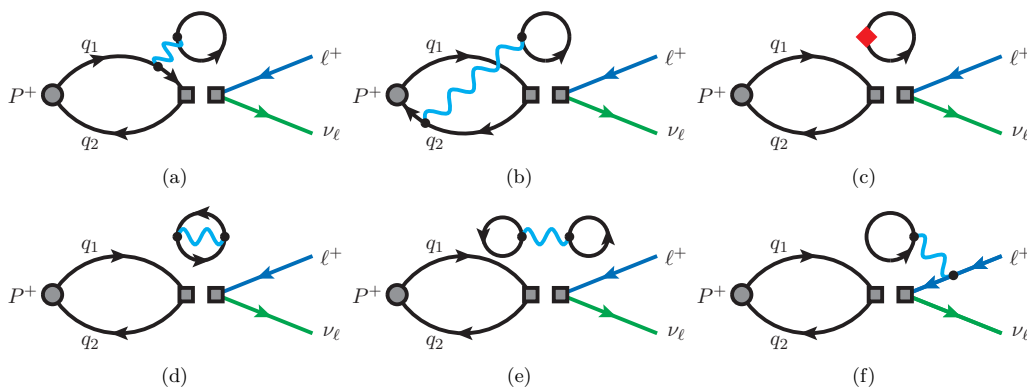
$$[\delta \bar{\mathcal{M}}_P^{rs}(\mathbf{p}\ell)]^{\text{nf}} = \left[ \frac{1}{2} e_\ell \sum_q e_q \frac{\partial^2}{\partial e_q \partial e_\ell} \right] \bar{\mathcal{M}}_P^{rs}(\mathbf{p}\ell) \Big|_{\sigma^{(0)}}, \quad (3.30)$$

and the corresponding diagrams are shown in figures 2(f)–2(g) and 3(f). Finally, the third term in eq. (3.28) consists in the  $\mathcal{O}(e_\ell^2)$  contribution of the lepton self-energy in figure 2(h), which is proportional to  $\mathcal{A}_P^{(0)}$  with a factor that can be computed analytically in perturbation theory,

$$[\delta \bar{\mathcal{M}}_P^{rs}(\mathbf{p}\ell)]^\ell = \left[ \frac{1}{2} e_\ell^2 \frac{\partial^2}{\partial e_\ell^2} \right] \bar{\mathcal{M}}_P^{rs}(\mathbf{p}\ell) \Big|_{\sigma^{(0)}} = -\mathcal{A}_P^{(0)} \left[ \frac{1}{2} e_\ell^2 \frac{\partial^2}{\partial e_\ell^2} \right] \mathcal{L}_\phi^{rs}(\mathbf{p}\ell) \Big|_{\sigma^{(0)}}, \quad (3.31)$$

with  $\mathcal{L}_\phi^{rs}(\mathbf{p}\ell) = \langle \ell^+, r, \mathbf{p}_\ell; \nu_\ell, s, \mathbf{p}_\nu | J_L^0 | 0 \rangle^\phi$ . This perturbative correction, however, cancels in the difference  $[\Gamma_0(L) - \Gamma_0^{(2)}(L)]$  in eq. (3.2) and therefore can be neglected in practice in the calculation. Of course, the lepton self-energy must be included in  $\Gamma_P^{\text{pert}}(m_\gamma)$ .

Due to the numerical difficulty of evaluating the quark disconnected diagrams in figure 3 on the lattice, in this work we employ the electro-quenched approximation. This consists in treating the sea quarks as if they were electrically neutral and hence, in practice, neglecting the diagrams in figure 3. The deviations from this approximation are expected to be small,



**Figure 3.** Quark-disconnected Feynman diagrams contributing to the leading IB corrections to the weak decay.

and we assign an associated systematic uncertainty in our final prediction. We are currently working on overcoming this approximation and the progress of our preliminary study has been reported in ref. [30].

### 3.2 Extracting matrix elements from Euclidean correlation functions

The IB corrections to meson masses,  $\delta m_P$ , and to the decay amplitude,  $\delta \mathcal{A}_P$ , which are needed to compute  $\delta R_P^{\text{latt}}$  in eq. (3.24), can be obtained from the study of the large time behaviour of suitably defined Euclidean correlation functions. Here the correlation functions are studied in the continuum and in a volume with infinite temporal extent. The subtraction of the effects due to the finite spatial extent of the lattice,  $L$ , are discussed later in section 3.3, while finite-time corrections to these quantities will be addressed in section 4.3, together with the details on the lattice implementation of the correlation functions.

**Tree-level correlation function:** we start by defining the tree-level correlation function for the decay  $P^+ \rightarrow \ell^+ \nu_\ell$ , with the aim of extracting the tree-level matrix element  $\mathcal{A}_P^{(0)}$  defined in eq. (3.18). As discussed in section 3.1, in the absence of QED the matrix element for the operator  $O_W$  is factorisable into a hadronic and a leptonic part. As a consequence, we can extract the hadronic matrix element  $\mathcal{A}_P^{(0)}$  from a pure QCD two-point correlation function without the need of including leptons in the calculation. Let  $\phi_P^\dagger(x) = \bar{q}_1(x) \gamma_5 q_2(x)$  be the interpolating operator for the pseudoscalar meson  $P^+$  and define the Euclidean correlation functions

$$C_{\text{PA}}(t) \equiv \int d^3\mathbf{x} \langle 0 | T[A^0(t, \mathbf{x}) \phi_P^\dagger(0)] | 0 \rangle, \quad C_{\text{PP}}(t) \equiv \int d^3\mathbf{x} \langle 0 | T[\phi_P(t, \mathbf{x}) \phi_P^\dagger(0)] | 0 \rangle, \quad (3.32)$$

with  $A^0(x) = \bar{q}_2(x) \gamma^0 \gamma_5 q_1(x)$  the temporal component of the hadronic axial current and the meson being projected on zero spatial momentum. For simplicity, we use translational invariance to create the meson at the origin. In practice, lattice correlators have been computed for several positions  $x_P = (t_P, \mathbf{x}_P)$  and then shifted and averaged over all the volume to improve the statistical precision (see section 4.3). Note that these are generic correlation functions evaluated at a given point  $\sigma$ . Fixing  $t > 0$ , the correlation functions

in eq. (3.32) have the following spectral decomposition

$$C_{\text{PA}}(t) = \frac{\mathcal{A}_P Z_P}{2m_P} e^{-m_P t} + \dots, \quad C_{\text{PP}}(t) = \frac{|Z_P|^2}{2m_P} e^{-m_P t} + \dots, \quad (3.33)$$

where  $Z_P = \langle P, \mathbf{0} | \phi^\dagger(0) | 0 \rangle$  and the ellipses stand for contributions of heavier states that decay exponentially faster than the leading terms. The combined study of the two correlation functions evaluated in iso-QCD allows one to extract the meson mass  $m_P^{(0)}$  and the matrix elements  $Z_P^{(0)}$  and  $\mathcal{A}_P^{(0)}$ .

**Factorizable correlators:** when IB corrections only involve the constituent quarks of the decaying meson, the matrix element is still factorizable into a hadronic and a leptonic part. Also in this case we can make use of the correlation functions in eq. (3.32). Defining the leading factorizable corrections to the correlators as

$$\delta C_{\text{PA}}^f(t) = \left[ \frac{1}{2} \sum_{q,q'} e_q e_{q'} \frac{\partial^2}{\partial e_q \partial e_{q'}} + \sum_q (\hat{m}_q^\phi - \hat{m}_q^{(0)}) \frac{\partial}{\partial \hat{m}_q} \right] C_{\text{PA}}(t) \Big|_{\sigma^{(0)}} \quad (3.34)$$

and analogously for  $\delta C_{\text{PP}}^f(t)$ , one gets the following decomposition for their ratios with the corresponding tree-level correlators

$$\mathcal{R}_{\text{PA}}^f(t) \equiv \frac{\delta C_{\text{PA}}^f(t)}{C_{\text{PA}}^{(0)}(t)} = \frac{\delta \mathcal{A}_P^f}{\mathcal{A}_P^{(0)}} + \frac{\delta Z_P}{Z_P^{(0)}} - \frac{\delta m_P}{m_P^{(0)}} (1 + m_P^{(0)} t) + \dots, \quad (3.35)$$

$$\mathcal{R}_{\text{PP}}^f(t) = \frac{\delta C_{\text{PP}}^f(t)}{C_{\text{PP}}^{(0)}(t)} = 2 \frac{\delta Z_P}{Z_P^{(0)}} - \frac{\delta m_P}{m_P^{(0)}} (1 + m_P^{(0)} t) + \dots, \quad (3.36)$$

from a Taylor expansion of the spectral decomposition of the form eq. (3.33). The slope in  $t$  of the above ratios corresponds to the mass shift  $\delta m_P$ , and by combining the constant coefficients we can obtain the correction  $\delta \mathcal{A}_P^f / \mathcal{A}_P^{(0)}$ .

**Non-factorizable correlators:** in order to obtain the non-factorizable IB corrections to the decay amplitude, we start defining the following QCD+QED correlation function

$$C_{\text{P}\ell}(t, t_\ell) = \int d^3 \mathbf{x} d^3 \mathbf{x}_\ell d^3 \mathbf{x}_\nu e^{-i\mathbf{p}_\nu \cdot \mathbf{x}_\nu - i\mathbf{p}_\ell \cdot \mathbf{x}_\ell} \langle 0 | T[\nu(t_\ell, \mathbf{x}_\nu) \bar{\ell}(t_\ell, \mathbf{x}_\ell) O_W(0) \phi_P^\dagger(-t, -\mathbf{x})] | 0 \rangle, \quad (3.37)$$

where for simplicity we have set the temporal coordinate of the neutrino and the lepton to be equal. Also in this case we have used translational invariance to insert the weak Hamiltonian at the origin. Fixing  $t > 0$  and  $t_\ell > 0$  we have that in iso-QCD the above correlator becomes

$$C_{\text{P}\ell}^{(0)}(t, t_\ell) = - \langle 0 | J_H^\rho(0) \phi_P^\dagger(-t, \mathbf{p}_P = \mathbf{0}) | 0 \rangle \times S_\nu(t_\ell, \mathbf{p}_\nu | 0) \gamma^\rho (1 - \gamma_5) S_\ell(0 | t_\ell, \mathbf{p}_\ell), \quad (3.38)$$

with

$$S_\nu(t_\ell, \mathbf{p}_\nu | 0) = \int d^3 \mathbf{x}_\nu e^{-i\mathbf{p}_\nu \cdot \mathbf{x}_\nu} S_\nu(t_\ell, \mathbf{x}_\nu | 0) = \frac{e^{-\omega_\nu t_\ell}}{2\omega_\nu} \sum_s u_\nu^s(\mathbf{p}_\nu) \bar{u}_\nu(\mathbf{p}_\nu), \quad (3.39)$$

$$S_\ell(0 | t_\ell, \mathbf{p}_\ell) = \int d^3 \mathbf{x}_\ell e^{-i\mathbf{p}_\ell \cdot \mathbf{x}_\ell} S_\ell(0 | t_\ell, \mathbf{x}_\ell) = - \frac{e^{-\omega_\ell t_\ell}}{2\omega_\ell} \sum_r v_\ell^r(\mathbf{p}_\ell) \bar{v}_\ell^r(\mathbf{p}_\ell). \quad (3.40)$$

Using eq. (3.33) we get the following spectral decomposition in iso-QCD

$$C_{P\ell}^{(0)}(t, t_\ell) = - \sum_{r,s} \frac{e^{-m_P^{(0)}t} e^{-\omega_\ell t_\ell} e^{-\omega_\nu t_\ell}}{8m_P^{(0)}\omega_\ell\omega_\nu} Z_P^{(0)} u_\nu^r(\mathbf{p}_\nu) \left\{ \mathcal{A}_P^{(0)} \mathcal{L}^{rs}(\mathbf{p}_\ell) \right\} \bar{v}_\ell^s(\mathbf{p}_\ell) + \dots, \quad (3.41)$$

with  $\mathcal{L}^{rs}(\mathbf{p}_\ell)$  defined in eq. (3.19). Note that  $C_{P\ell}^{(0)}(t, t_\ell)$  is a matrix in Dirac space and that tracing with  $\gamma_L^0 = \gamma^0(1 - \gamma_5)$  gives

$$\text{Tr}[\gamma_L^0 C_{P\ell}^{(0)}(t, t_\ell)] = - \frac{e^{-m_P^{(0)}t} e^{-\omega_\ell t_\ell} e^{-\omega_\nu t_\ell}}{8m_P^{(0)}\omega_\ell\omega_\nu} Z_P^{(0)} \mathcal{A}_P^{(0)} |\mathcal{L}(\mathbf{p}_\ell)|^2 + \dots \quad (3.42)$$

We now define the non-factorizable correlator as

$$\begin{aligned} \delta C_{P\ell}^{\text{nf}}(t, t_\ell) &= \int d^4x d^4y \langle 0 | \text{T} [J_H^\rho(0) V^\mu(x) \phi_P^\dagger(-t, \mathbf{p}_P = \mathbf{0})] | 0 \rangle \times \Delta_{\mu\nu}(x - y) \\ &\quad \times S_\nu(t_\ell, \mathbf{p}_\nu | 0) \gamma^\rho (1 - \gamma_5) S_\ell(0 | y) \gamma^\nu S_\ell(y | t_\ell, \mathbf{p}_\ell), \end{aligned} \quad (3.43)$$

where  $V^\mu(x) = \sum_{q=1}^{N_f} V_q^\mu(x)$  is the Euclidean quark electromagnetic current and  $\Delta_{\mu\nu}(x - y)$  the photon propagator. Here we have used  $V_\ell^\nu(y) = \bar{\ell}(y) \gamma^\nu \ell(y)$  for the leptonic electromagnetic current. The correlator in eq. (3.43) is obtained by applying the derivatives of eq. (3.30) to the QCD+QED correlator  $C_{P\ell}(t, t_\ell)$ . The asymptotic behaviour of the non-factorizable correlator is

$$\delta C_{P\ell}^{\text{nf}}(t, t_\ell) = \sum_{r,s} \frac{e^{-m_P^{(0)}t} e^{-\omega_\ell t_\ell} e^{-\omega_\nu t_\ell}}{8m_P^{(0)}\omega_\ell\omega_\nu} Z_P^{(0)} u_\nu^r(\mathbf{p}_\ell) [\delta \bar{\mathcal{M}}_P^{rs}(\mathbf{p}_\ell)]^{\text{nf}} \bar{v}_\ell^s(\mathbf{p}_\ell) + \dots \quad (3.44)$$

Tracing the correlator with  $\gamma_L^0$  and making use of eq. (3.26) we can obtain the desired non-factorizable correction to the decay amplitude as

$$\mathcal{R}_{P\ell}^{\text{nf}} = \text{Re} \left[ \frac{\text{Tr}[\gamma_L^0 \delta C_{P\ell}^{\text{nf}}(t, t_\ell)]}{\text{Tr}[\gamma_L^0 C_{P\ell}^{(0)}(t, t_\ell)]} \right] = \frac{\delta \mathcal{A}_P^{\text{nf}}}{\mathcal{A}_P^{(0)}} + \dots \quad (3.45)$$

### 3.3 Subtraction of finite-volume effects

In the calculation presented in this work we adopt the QED<sub>L</sub> prescription, first introduced in ref. [12]. As discussed in a number of publications [12, 31–33], charged states are not well-defined in a naive implementation of finite-volume QED (in which periodic boundary conditions are applied to the photon fields). The QED<sub>L</sub> approach solves this by discarding the zero spatial-momentum mode of the photon on each energy slice. The resulting momentum-space photon propagator in Feynman gauge then takes the simple form

$$\Delta_L^{\mu\nu}(k_0, \mathbf{k}) = \delta^{\mu\nu} \frac{1 - \delta_{\mathbf{k}, \mathbf{0}}}{k_0^2 + \mathbf{k}^2}. \quad (3.46)$$

This prescription solves the issue of zero-mode singularities in a periodic volume, at the cost of violating locality in space at finite volume. Nevertheless, this theory has a well-defined and local limit if the infinite-volume extrapolation is performed before the continuum limit. Additionally, QED<sub>L</sub> has been the dominant prescription so far in high-precision

lattice QCD+QED calculations, including radiative corrections to leptonic decays [7, 8] and isospin-breaking corrections to the muon anomalous magnetic moment [18, 34]. Alternative strategies exist which preserve locality, such as introducing a photon mass [35] or using non-periodic boundary conditions [36]. However, these approaches affect other fundamental symmetries (gauge invariance and charge conservation, respectively) and their finite-volume behaviour for processes such as weak decays is currently not as well studied as in the case of QED<sub>L</sub>.

As it is described in detail in ref. [14], the Feynman rules of QED<sub>L</sub> can be used to predict the  $L$  dependence of any lattice quantity by representing the latter in terms of QCD vertex functions at fixed order in QED. In particular, this allows one to analytically predict the power-like volume dependence, order by order in  $1/L$ , for the virtual-photon contribution to the leptonic decay rate targeted in this work. This strategy is already implicit in eq. (3.2), where the subtracted quantity in the first term, denoted by  $\Gamma_0^{(2)}(L)$ , is defined as the analytic QED<sub>L</sub> prediction through  $O(1/L^2)$ .

An extension of  $\Gamma_0^{(2)}(L)$  to  $n$  orders in  $1/L$  can be written as

$$\Gamma_0^{(n)}(L) = \Gamma_0^{\text{tree}} \left[ 1 + \delta R_P^{(n)}(L) \right], \tag{3.47}$$

where

$$\delta R_P^{(n)}(L) = 2 \frac{\alpha_{\text{em}}}{4\pi} \left( \tilde{Y}_P(L) + \sum_{i=0}^n \frac{Y_{P,i}}{L^i} \right), \tag{3.48}$$

isolates the  $O(\alpha_{\text{em}})$  contribution of direct interest to us. The first term in parentheses,  $\tilde{Y}_P(L)$ , combines the infinite-volume universal (point-like) contributions to the decay rate with those that are logarithmic in  $L$ .<sup>3</sup> The functional form is given by [14, 24]

$$\tilde{Y}_P(L) = -\frac{5}{4} + 2 \log \left( \frac{m_\ell}{m_W} \right) + 2 \log \left( \frac{m_\ell L}{2\pi} \right) + \frac{2 \log r_\ell}{|\mathbf{v}_\ell|} \left[ \log \frac{m_P L}{2\pi} + \log \frac{m_\ell L}{2\pi} - 1 \right], \tag{3.49}$$

with  $\mathbf{v}_\ell = \mathbf{p}_\ell/\omega_\ell$  defining the 3-velocity of the lepton and  $r_\ell = m_\ell/m_P$  the lepton-pseudoscalar mass ratio.

Equation (3.49) depends only on the masses of particles and is, in this sense, universal or structure-independent. In fact, one can show that the same is true for  $Y_{P,0}$  and  $Y_{P,1}$ , while for  $Y_{P,n>1}$  structure dependence enters through contributions from, e.g., form factors and their derivatives. For this reason, the point-like approximation can be used to calculate  $Y_{P,0}$  and  $Y_{P,1}$  and the full machinery introduced in ref. [14] is first required for the determination of  $Y_{P,2}$  and higher-order coefficients.

---

<sup>3</sup>In the notation of ref. [14] this quantity can be defined introducing a photon mass  $\lambda$  as

$$\tilde{Y}_P(L) = \lim_{\lambda \rightarrow 0} \left[ Y_{P,1V}^{\text{mi}}(\lambda) + Y_{P,\log} \log \frac{L\lambda}{2\pi} \right].$$

A summary of the knowledge to-date on these coefficients is given by the following:

$$\begin{aligned}
 Y_{P,0} &= \frac{c_3 - 2(c_3(\mathbf{v}_\ell) - B_1(\mathbf{v}_\ell))}{2\pi} + 2(1 - \log 2), \\
 Y_{P,1} &= -\frac{(1 + r_\ell^2)^2 c_2 - 4r_\ell^2 c_2(\mathbf{v}_\ell)}{m_P(1 - r_\ell^4)}, \\
 Y_{P,2} &= -\frac{F_A^P}{f_P} \frac{4\pi [(1 + r_\ell^2)^2 - 4r_\ell^2 c_1(\mathbf{v}_\ell)]}{m_P(1 - r_\ell^4)} + \frac{8\pi [(1 + r_\ell^2)c_1 - 2c_1(\mathbf{v}_\ell)]}{m_P^2(1 - r_\ell^4)}, \\
 Y_{P,3} &= \frac{32\pi^2 c_0 (2 + r_\ell^2)}{m_P^3(1 + r_\ell^2)^3} + Y_{P,3}^{\text{sd}},
 \end{aligned} \tag{3.50}$$

where  $c_j$  and  $c_j(\mathbf{v}_\ell)$  are known finite-volume coefficients, and  $B_1(\mathbf{v}_\ell)$  is a known special function. These quantities are all defined in ref. [14].

In  $Y_{P,2}$ , the structure-dependent ratio  $F_A^P/f_P$  appears where  $F_A^P$  is the on-shell zero-momentum axial form factor describing radiative leptonic decays and  $f_P$  is the iso-QCD pseudoscalar decay constant. A key message is that, while the full result including structure dependence is known for  $Y_{P,2}$ , the same is not true for  $Y_{P,3}$ , for which the structure-dependent piece, denoted  $Y_{P,3}^{\text{sd}}$ , has yet to be determined. As a result, the finite-volume subtractions available currently include  $\delta R_P^{(2)}(L)$  and  $\delta R_P^{(3),\text{pt}}(L)$ , where  $Y_{P,3}^{\text{sd}}$  is set to zero in the latter. In this work we use  $\delta R_P^{(2)}(L)$  to determine our central value and take the absolute difference  $|\delta R_P^{(3),\text{pt}}(L) - \delta R_P^{(2)}(L)|$  to estimate a systematic uncertainty associated with neglected finite-volume effects.

Finally, we also need to consider finite-size effects from QED corrections to the meson mass  $m_P$ . For the finite-volume state with zero spatial momentum, these are given by [14, 25, 31–33]

$$\Delta m_P^2(L) = e^2 m_P^2 \left\{ \frac{c_2}{4\pi^2 m_P L} + \frac{c_1}{2\pi(m_P L)^2} + \frac{m_P^2 \langle r_P^2 \rangle}{3(m_P L)^3} + \frac{\mathcal{C}}{(m_P L)^3} + \mathcal{O}\left[\frac{1}{(m_P L)^4}\right] \right\}, \tag{3.51}$$

where  $\langle r_P^2 \rangle$  is the squared electromagnetic charge radius known from experiments, dispersion theory and lattice simulations [1, 2], and  $\mathcal{C}$  is an unknown contribution, arising from the branch-cut in the Compton amplitude evaluated with zero spatial momentum for both the photon and the pseudoscalar. In this work we use the predicted volume dependence through  $\mathcal{O}(1/L^2)$  to estimate the infinite-volume pseudoscalar mass. As with the decay rate, we take the difference between the  $1/L^2$  and partial  $1/L^3$  results as a systematic uncertainty.

### 3.4 Inclusion of real photon emission

Lastly, we need to include the contributions from a real photon emission, namely the quantity  $\delta R_P^{\text{pert}}(\omega_\gamma^{\text{max}})$  in eq. (3.8). To this end, we adopt the formulation discussed in detail in ref. [10]. Most notably, if the photon energy threshold,  $\omega_\gamma^{\text{cut}}$ , is small enough, one may treat the initial hadron as a point-like particle and compute the inner bremsstrahlung term analytically. However, since structure-dependent contributions are negligible for the decays studied in this work, we can set  $\omega_\gamma^{\text{cut}}$  to the maximum value allowed for the photon energy,

namely  $\omega_\gamma^{\max} = m_P(1 - r_\ell^2)/2$ . We report here the result obtained in ref. [10],

$$\begin{aligned} \delta R_P^{\text{pert}}(\omega_\gamma^{\max}) &= \lim_{m_\gamma \rightarrow 0} \delta R_P^{\text{pert}}(\omega_\gamma^{\max}, m_\gamma) \\ &= \frac{\alpha_{\text{em}}}{4\pi} \left[ 3 \log\left(\frac{m_P^2}{m_W^2}\right) - 8 \log(1 - r_\ell^2) - \frac{3r_\ell^4}{(1 - r_\ell^2)^2} \log r_\ell^2 - 8 \frac{1 + r_\ell^2}{1 - r_\ell^2} \text{Li}_2(1 - r_\ell^2) \right. \\ &\quad \left. + \frac{13 - 19r_\ell^2}{2(1 - r_\ell^2)} + \frac{6 - 14r_\ell^2 - 4(1 + r_\ell^2) \log(1 - r_\ell^2)}{1 - r_\ell^2} \log r_\ell^2 \right]. \end{aligned} \quad (3.52)$$

## 4 Lattice methodology

In this section we discuss the lattice implementation of the correlation functions relevant for the calculation of IB corrections to the leptonic decay rate and give the details of our lattice setup.

### 4.1 Lattice QCD+QED path integrals

As anticipated in section 2, IB corrections are computed in this work using the RM123 perturbative method [22], which consists in expanding the path integral for a given physical observable around the iso-QCD point. In practice, since our lattice setup has been tuned to an iso-symmetric point different from the target one described in section 2, we follow a two-step procedure to get our perturbative corrections. This consists in expanding both the full QCD+QED and the iso-QCD path integral around the simulation point, and get the desired correction as the difference of the two.

Let  $\langle \hat{\mathcal{O}} \rangle_{\sigma^\phi}$  be the expectation value of an observable  $\hat{\mathcal{O}}$  calculated (in lattice units) in terms of the discretized Euclidean path integral in the full QCD+QED theory with bare parameters  $\sigma^\phi = (g, e^\phi, \hat{\mathbf{m}}^\phi)$ ,

$$\langle \hat{\mathcal{O}} \rangle_{\sigma^\phi} = \frac{1}{Z^\phi} \int \mathcal{D}[U] \mathcal{D}[A] \mathcal{D}[\psi, \bar{\psi}] \hat{\mathcal{O}}[\psi, \bar{\psi}, U, A; \sigma^\phi] e^{-S_F[\psi, \bar{\psi}, U, A; \sigma^\phi]} e^{-S_\gamma[A]} e^{-\frac{1}{g^2} S_G[U]}, \quad (4.1)$$

with  $S_F[\psi, \bar{\psi}, U, A; \sigma^\phi]$  being the fermionic action, and  $S_\gamma[A]$  and  $S_G[U]$  the QED<sub>L</sub> and QCD gauge actions, respectively.  $Z^\phi$  denotes instead the QCD+QED partition function. Here we keep the discussion general and allow the observable  $\hat{\mathcal{O}}$  to depend on the electromagnetic coupling  $e^\phi$ . Let  $\langle \hat{\mathcal{O}} \rangle_{\sigma^{(0)}}$  be the corresponding expectation values calculated in the target iso-QCD theory,  $\sigma^{(0)} = (g, 0, \hat{\mathbf{m}}^{(0)})$ ,

$$\langle \hat{\mathcal{O}} \rangle_{\sigma^{(0)}} = \frac{1}{Z^{(0)}} \int \mathcal{D}[U] \mathcal{D}[\psi, \bar{\psi}] \hat{\mathcal{O}}[\psi, \bar{\psi}, U] e^{-S_F[\psi, \bar{\psi}, U; \sigma^{(0)}]} e^{-\frac{1}{g^2} S_G[U]}, \quad (4.2)$$

and  $\langle \hat{\mathcal{O}} \rangle_\sigma$  that evaluated at the simulation point  $\sigma = (g, 0, \hat{\mathbf{m}})$ , which is obtained from the previous equation by substituting  $\sigma^{(0)} \rightarrow \sigma$ .

The expansion of  $\langle \hat{\mathcal{O}} \rangle_{\sigma^\phi}$  around the simulation point  $\sigma$  is then given by

$$\langle \hat{\mathcal{O}} \rangle_{\sigma^\phi} = \langle \hat{\mathcal{O}} \rangle_\sigma + \langle \delta \hat{\mathcal{O}} \rangle_\sigma - \langle \hat{\mathcal{O}} \delta S_F \rangle_\sigma + \frac{1}{2} \langle \hat{\mathcal{O}} (\delta S_F)^2 \rangle_\sigma + \mathcal{O}(\bar{\epsilon}^2). \quad (4.3)$$

The correction  $\langle \delta \hat{\mathcal{O}} \rangle_\sigma$  in eq. (4.3) only appears if the observable  $\mathcal{O}$  itself depends on the electromagnetic coupling  $e^\phi$ , which is not the case for the correlation functions studied in this work. The quantity  $\delta S_F$  is instead the IB correction to the lattice fermionic action, i.e.

$$\delta S_F = \sum_x \sum_f \left[ (\hat{m}_f^\phi - \hat{m}_f) \hat{\mathcal{S}}_f(x) + i e_f \hat{V}_{c,f}^\mu(x) \hat{A}_\mu(x) - \frac{1}{2} e_f^2 \hat{T}_f^\mu(x) \hat{A}_\mu^2(x) \right] + \mathcal{O}(\bar{\varepsilon}^2). \quad (4.4)$$

Here  $\mathcal{S}_f(x) = \bar{\psi}_f(x) \psi_f(x)$  is the scalar density, while  $V_{c,f}^\mu(x)$  and  $T_f^\mu(x)$  are the electromagnetic conserved current and the seagull (or tadpole) current, respectively, which depend on the lattice regularization adopted (see e.g. refs. [22, 37]). The hats denote that all quantities are expressed in lattice units. In this work, however, we employ a definition of the fermion-photon coupling similar to the continuum one, where we use the renormalized local vector current,  $V_f^\mu(x) = Z_V \bar{\psi}_f(x) \gamma^\mu \psi_f(x)$ , instead of the electromagnetic conserved one and do not include the tadpole current.<sup>4</sup> This results in

$$\delta S_F = \sum_x \sum_f \left[ (\hat{m}_f^\phi - \hat{m}_f) \hat{\mathcal{S}}_f(x) + i e_f \hat{V}_f^\mu(x) \hat{A}_\mu(x) \right] + \mathcal{O}(\bar{\varepsilon}^2), \quad (4.5)$$

and the two approaches are expected to differ just by cut-off effects. The comparison of the two approaches has been thoroughly investigated, and we report on that in appendix A. Since the simulation point and the iso-QCD point only differ by the choice of the quark masses, the expansion of the iso-QCD path integral around the simulation point is given by

$$\langle \hat{\mathcal{O}} \rangle_{\sigma^{(0)}} = \langle \hat{\mathcal{O}} \rangle_\sigma + \sum_f (\hat{m}_f^{(0)} - \hat{m}_f) \langle \hat{\mathcal{O}} \hat{\mathcal{S}}_f \rangle_\sigma + \mathcal{O}(\bar{\varepsilon}^2). \quad (4.6)$$

From eqs. (4.3) and (4.6) it is then clear that IB corrections are obtained by computing correlation functions at the simulation point with the insertion of the operators  $\mathcal{S}_f(x)$  and  $V_f^\mu(x)$ . We repeat that throughout this paper we work in the electro-quenched approximation. In practice, the bare parameters of the sea quark are kept fixed to their simulated values, which amounts to neglecting all quark-line disconnected diagrams.

In the perturbative approach adopted in this work the U(1) gauge fields  $A_\mu(x)$  are generated as stochastic fields sampled according to the QED<sub>L</sub> gauge action in Feynman gauge [23, 37]

$$S_\gamma[A] = \frac{1}{V} \sum_{k:k \neq 0} \bar{k}^2 \sum_\mu |\tilde{A}_\mu(k)|^2 \quad \text{with} \quad \bar{k}^\mu = \frac{2}{a} \sin\left(\frac{ak^\mu}{2}\right), \quad (4.7)$$

with  $\tilde{A}_\mu(k)$  being the photon field in momentum space, so that the expectation value  $\langle A_\mu(x) A_\nu(y) \rangle_\gamma$  reproduces the photon propagator  $\Delta_{\mu\nu}(x - y)$ .

## 4.2 Lattice setup

For this calculation, we generate correlators for a  $(L/a)^3 \times (T/a) = 48^3 \times 96$  lattice using Möbius Domain Wall Fermions (DWF) [38] with close-to-physical masses. The Domain wall height and the length of the fifth dimension are  $aM_5 = 1.8$  and  $L_s/a = 24$ , respectively. See

<sup>4</sup>Note that for leptons  $Z_V = 1$  and hence  $V_\ell^\mu(x) = \bar{\ell}(x) \gamma^\mu \ell(x)$ .



ref. [39] for more details. The QCD gauge configurations are generated by the RBC/UKQCD collaboration using the Iwasaki gauge action [40] with bare coupling  $\beta = 2.13$ . The sea quark masses are  $\hat{m}_{\text{ud}}^{\text{sea}} = 0.00078$  for the light quarks and  $\hat{m}_{\text{s}}^{\text{sea}} = 0.0362$  for the strange quark. We work in a unitary setup where we choose the valence light-quark masses to have the same value as the sea,  $\hat{m}_{\text{ud}} = \hat{m}_{\text{ud}}^{\text{sea}}$  and similarly for the valence strange quarks,  $\hat{m}_{\text{s}} = \hat{m}_{\text{s}}^{\text{sea}}$ . In this setup, that we refer to as our *simulation point*  $\sigma$ , the lattice spacing has been determined without QED to be  $1/a = 1.7295(38)$  GeV and the simulated pion mass of this ensemble is  $m_{\pi} = 139.15(36)$  MeV, corresponding to  $m_{\pi}L = 3.863(6)$ .

To reduce the computational cost of inverting the Dirac operator for near-physical light quarks, we employ zMöbius fermions, which are a rational approximation of the Möbius formalism (see ref. [41] and references therein), together with the deflation eigenvectors generated by the RBC/UKQCD collaboration for this  $48^3 \times 96$  ensemble. Light-quark propagators can then be obtained with a smaller value of  $L_s$ , thereby reducing the simulation cost. This rational approximation of the Möbius DWF action must be corrected for, and we defer this discussion to appendix B.

### 4.3 Implementation of the hadronic correlators

We now turn to discuss the lattice implementation of the correlation functions introduced in section 3.2, where the relations with the corresponding matrix elements were obtained in the continuum and infinite-volume limit. As explained in section 4.1, IB corrections to the expectation value of a given observable can be obtained in the iso-QCD simulated theory by inserting additional operators in the correlation function [22]. As we discuss in the following this is obtained, in practice, by iteratively inverting the Dirac operator using suitable sources to get the appropriate sequential propagators. All the correlation functions used in this calculation are generated using a set of 60 statistically independent QCD configurations and are then resampled with the bootstrap method. The QED gauge fields  $A_{\mu}$  are generated using one stochastic source on each QCD gauge configuration. In this way the averages over QED and QCD gauge configurations are simultaneous. The inversions of the Dirac operator and the quark field contractions have been performed using the Grid/Hadrons software framework [42–44].

In this calculation we study the decay of the meson  $P^+$  in its rest frame,  $\mathbf{p}_P = \mathbf{0}$ . To create the meson we use gauge-fixed wall sources. This corresponds to defining a zero-momentum interpolating operator of the form

$$\phi_P^{\dagger}(t) \equiv \phi_P^{\dagger}(t, \mathbf{p}_P = \mathbf{0}) = a^6 \sum_{\mathbf{x}_1, \mathbf{x}_2} \bar{q}_1(t, \mathbf{x}_1) \gamma_5 q_2(t, \mathbf{x}_2), \quad (4.8)$$

and evaluating expectation values of this operator fixed to Coulomb gauge. Any gauge-fixed expectation value involving  $\phi_P^{\dagger}(t)$  can be re-expressed as a gauge invariant correlator with an alternative operator that includes a Wilson line between the quark fields. Crucially, the gauge-invariant equivalent is local in time so that we can perform spectral decompositions using the standard Hilbert space of lattice QCD with pseudoscalar quantum numbers.<sup>5</sup> Note that using such definition of the meson interpolating operator, the dimensions of

---

<sup>5</sup>The same would not be true for gauge fixings that affect temporal gauge links. The idea of equivalence

the correlators are different from those described in section 3.2 because of the additional integration over the spatial coordinates of the quark fields.

**Tree-level correlation function:** the tree-level correlation functions in eq. (3.32) are implemented at the simulation point in terms of quark propagators as

$$C_{\text{PA}}(t) = a^3 \sum_{\mathbf{x}} \langle \text{Tr}[S_{q_2}(t, \mathbf{x}|0)^\dagger \gamma^0 S_{q_1}(t, \mathbf{x}|0)] \rangle, \quad (4.9)$$

$$C_{\text{PP}}(t) = a^6 \sum_{\mathbf{x}_1, \mathbf{x}_2} \langle \text{Tr}[S_{q_2}(t, \mathbf{x}_2|0)^\dagger S_{q_1}(t, \mathbf{x}_1|0)] \rangle, \quad (4.10)$$

where we have used  $\gamma_5$ -hermiticity,  $S_q(t_2, \mathbf{x}_2|t_1, \mathbf{x}_1) = \gamma_5 S_q(t_1, \mathbf{x}_1|t_2, \mathbf{x}_2)^\dagger \gamma_5$ , and defined the quark propagator with one end projected on zero momentum as

$$S_q(t_1, \mathbf{x}_1|t_2) \equiv S_q(t_1, \mathbf{x}_1|t_2, \mathbf{p} = \mathbf{0}) = a^3 \sum_{\mathbf{x}_2} S_q(t_1, \mathbf{x}_1|t_2, \mathbf{x}_2), \quad (4.11)$$

while the symbol  $\langle \cdot \rangle$  denotes the average over the gauge configurations. Note that here we have generated the pseudoscalar meson at  $t_P = 0$  for simplicity. In the lattice calculation we have instead evaluated the correlation functions on each gauge configuration inserting the source at every timeslice  $t_P/a = \{1, \dots, T/a = 96\}$  and then shifted and averaged over the source positions to improve the statistical uncertainty.

By considering the asymptotic form for the pseudoscalar correlator in eq. (3.33) on a torus with period  $T$  in the temporal direction and periodic boundary conditions we obtain

$$C_{\text{PA}}(t) = \frac{\mathcal{A}_P Z_P}{2 m_P} \{e^{-m_P t} - e^{-m_P(T-t)}\}, \quad (4.12)$$

$$L^3 C_{\text{PP}}(t) = \frac{|Z_P|^2}{2 m_P} \{e^{-m_P t} + e^{-m_P(T-t)}\}, \quad (4.13)$$

having neglected exponentially suppressed contributions of excited states.

**Factorizable correlators:** let us define the sequential propagators obtained by inserting the correction to the fermionic action  $\delta S_F$  (see eq. (4.5)) along the quark line as

$$S_q^{(1)}(x|y) = e_q S_q^A(x|y) + (\hat{m}_q^\phi - \hat{m}_q) S_q^S(x|y), \quad (4.14)$$

where

$$S_q^A(x|y) = i a^4 Z_V \sum_z S_q(x|z) \gamma^\mu A_\mu(z) S_q(z|y) \quad \text{and} \quad S_q^S(x|y) = a^3 \sum_z S_q(x|z) S_q(z|y). \quad (4.15)$$

We can analogously define the sequential quark propagator with a double insertion of  $\delta S_F$ , which generates the quark self-energy, as

$$S_q^{(2)}(x|y) = e_q^2 S_q^{\text{self}}(x|y) + \mathcal{O}(\bar{\varepsilon}^2) \quad (4.16)$$

---

between gauge-invariant and gauge-fixed formulations is discussed in the context of QED in a seminal paper by Dirac [45] and more recently in ref. [46].

with

$$S_q^{\text{self}}(x|y) = -a^8 Z_V^2 \sum_{z,w} S_q(x|z) \gamma^\mu A_\mu(z) S_q(z|w) \gamma^\nu A_\nu(w) S_q(w|y). \quad (4.17)$$

Note that the propagators  $S_q^{(1)}(x|y)$  and  $S_q^{(2)}(x|y)$  are both  $\gamma_5$ -hermitian.

The factorizable correlators  $\delta C_{\text{PA}}^f(t)$  and  $\delta C_{\text{PP}}^f(t)$  can then be evaluated in terms of such sequential propagators. We define

$$\begin{aligned} \delta C_{\text{PA}}^f(t) &= 4\pi\alpha_{\text{em}} \delta C_{\text{PA}}^{\text{em}}(t) + \sum_q (\hat{m}_q^\phi - \hat{m}_q^{(0)}) \delta C_{\text{PA}}^{\text{S},q}(t) \\ &= \sum_q e_q^2 \delta C_{\text{PA}}^{\text{self},q}(t) + e_{q_1} e_{q_2} \delta C_{\text{PA}}^{\text{exch}}(t) + \sum_q (\hat{m}_q^\phi - \hat{m}_q^{(0)}) \delta C_{\text{PA}}^{\text{S},q}(t), \end{aligned} \quad (4.18)$$

and analogously  $\delta C_{\text{PP}}^f(t)$ , where

$$\begin{aligned} \delta C_{\text{PA}}^{\text{self},q_1}(t) &= a^3 \sum_{\mathbf{x}} \langle \text{Tr}[S_{q_2}(t, \mathbf{x}|0)^\dagger \gamma^0 S_{q_1}^{\text{self}}(t, \mathbf{x}|0)] \rangle, \\ \delta C_{\text{PA}}^{\text{self},q_2}(t) &= a^3 \sum_{\mathbf{x}} \langle \text{Tr}[S_{q_2}^{\text{self}}(t, \mathbf{x}|0)^\dagger \gamma^0 S_{q_1}(t, \mathbf{x}|0)] \rangle, \\ \delta C_{\text{PA}}^{\text{exch}}(t) &= a^3 \sum_{\mathbf{x}} \langle \text{Tr}[S_{q_2}^A(t, \mathbf{x}|0)^\dagger \gamma^0 S_{q_1}^A(t, \mathbf{x}|0)] \rangle, \\ \delta C_{\text{PA}}^{\text{S},q_1}(t) &= a^3 \sum_{\mathbf{x}} \langle \text{Tr}[S_{q_2}(t, \mathbf{x}|0)^\dagger \gamma^0 S_{q_1}^{\text{S}}(t, \mathbf{x}|0)] \rangle, \\ \delta C_{\text{PA}}^{\text{S},q_2}(t) &= a^3 \sum_{\mathbf{x}} \langle \text{Tr}[S_{q_2}^{\text{S}}(t, \mathbf{x}|0)^\dagger \gamma^0 S_{q_1}(t, \mathbf{x}|0)] \rangle, \end{aligned} \quad (4.19)$$

and

$$\begin{aligned} \delta C_{\text{PP}}^{\text{self},q_1}(t) &= a^6 \sum_{\mathbf{x}_1, \mathbf{x}_2} \langle \text{Tr}[S_{q_2}(t, \mathbf{x}_2|0)^\dagger S_{q_1}^{\text{self}}(t, \mathbf{x}_1|0)] \rangle, \\ \delta C_{\text{PP}}^{\text{self},q_2}(t) &= a^6 \sum_{\mathbf{x}_1, \mathbf{x}_2} \langle \text{Tr}[S_{q_2}^{\text{self}}(t, \mathbf{x}_2|0)^\dagger S_{q_1}(t, \mathbf{x}_1|0)] \rangle, \\ \delta C_{\text{PP}}^{\text{exch}}(t) &= a^6 \sum_{\mathbf{x}_1, \mathbf{x}_2} \langle \text{Tr}[S_{q_2}^A(t, \mathbf{x}_2|0)^\dagger S_{q_1}^A(t, \mathbf{x}_1|0)] \rangle, \\ \delta C_{\text{PP}}^{\text{S},q_1}(t) &= a^6 \sum_{\mathbf{x}_1, \mathbf{x}_2} \langle \text{Tr}[S_{q_2}(t, \mathbf{x}_2|0)^\dagger S_{q_1}^{\text{S}}(t, \mathbf{x}_1|0)] \rangle, \\ \delta C_{\text{PP}}^{\text{S},q_2}(t) &= a^6 \sum_{\mathbf{x}_1, \mathbf{x}_2} \langle \text{Tr}[S_{q_2}^{\text{S}}(t, \mathbf{x}_2|0)^\dagger S_{q_1}(t, \mathbf{x}_1|0)] \rangle, \end{aligned} \quad (4.20)$$

having used again  $\gamma_5$ -hermiticity together with eq. (4.11). Note that the symmetries of the correlators ensure that  $\delta C_{\text{PP}}^{\text{self},u} = \delta C_{\text{PP}}^{\text{self},d}$  and  $\delta C_{\text{PA}}^{\text{self},u} = \delta C_{\text{PA}}^{\text{self},d}$ , as well as  $\delta C_{\text{PP}}^{\text{S},u} = \delta C_{\text{PP}}^{\text{S},d}$  and  $\delta C_{\text{PA}}^{\text{S},u} = \delta C_{\text{PA}}^{\text{S},d}$  when  $m_u = m_d = m_{\text{ud}}$ . The five correlators in eq. (4.19) correspond to the hadronic part of the Feynman diagrams shown in figures 2(a)-(e), respectively.

For the factorizable correlators, correcting the asymptotic behaviour in eqs. (3.35) and (3.36) for finite-time  $T$  effects with (anti-)periodic boundary conditions and neglecting the contribution of excited states results in

$$\mathcal{R}_{\text{PA}}^f(t) = \frac{\delta \mathcal{A}_P^f}{\mathcal{A}_P} + \frac{\delta Z_P}{Z_P} - \frac{\delta m_P}{m_P} f_{\text{PA}}(t, T), \quad (4.21)$$

$$\mathcal{R}_{\text{PP}}^f(t) = 2 \frac{\delta Z_P}{Z_P} - \frac{\delta m_P}{m_P} f_{\text{PP}}(t, T), \quad (4.22)$$

with

$$f_{\text{PA}}(t, T) = 1 + m_P \left\{ \frac{T}{2} - \left( t - \frac{T}{2} \right) \coth \left[ m_P \left( t - \frac{T}{2} \right) \right] \right\}, \quad (4.23)$$

$$f_{\text{PP}}(t, T) = 1 + m_P \left\{ \frac{T}{2} - \left( t - \frac{T}{2} \right) \tanh \left[ m_P \left( t - \frac{T}{2} \right) \right] \right\}, \quad (4.24)$$

and  $f_{\text{PA}}(t, T) = f_{\text{PP}}(t, T) \approx 1 + m_P t$  for  $t \ll T/2$ .

In the following we will make use of the notation  $\mathcal{R}_{\text{PA}}^{\text{x}}(t)$  (and analogously for  $\mathcal{R}_{\text{PP}}^{\text{x}}(t)$ ) with  $\text{x} = \{\text{self}, \text{q}; \text{exch}; \mathcal{S}, \text{q}\}$ . This has to be interpreted as the contributions to  $\mathcal{R}_{\text{PA}}^{\text{f}}(t)$  coming from the corresponding corrections to the correlator  $\delta C_{\text{PA}}^{\text{f}}(t)$  in eq. (4.18). Equivalently, we can decompose the correction to the meson mass up to  $\mathcal{O}(\epsilon^2)$  as follows

$$\begin{aligned} \delta m_P &= 4\pi\alpha_{\text{em}} \delta m_P^{\text{em}} + \sum_{\text{q}} (\hat{m}_{\text{q}}^{\phi} - \hat{m}_{\text{q}}^{(0)}) \delta m_P^{\mathcal{S}, \text{q}} \\ &= \sum_{\text{q}} e_{\text{q}}^2 \delta m_P^{\text{self}, \text{q}} + e_{\text{q}_1} e_{\text{q}_2} \delta m_P^{\text{exch}} + \sum_{\text{q}} (\hat{m}_{\text{q}}^{\phi} - \hat{m}_{\text{q}}^{(0)}) \delta m_P^{\mathcal{S}, \text{q}}. \end{aligned} \quad (4.25)$$

For the mesons studied in this work we have  $\text{q} = \{\text{u}, \text{d}\}$  for  $\delta m_{\pi^+}$ ,  $\text{q} = \{\text{u}, \text{s}\}$  for  $\delta m_{K^+}$ ,  $\text{q} = \{\text{d}, \text{s}\}$  for  $\delta m_{K^0}$ ,  $\text{q} = \{\text{u}, \text{u}\}$  for  $\delta M_{\text{uu}}$  and  $\text{q} = \{\text{d}, \text{d}\}$  for  $\delta M_{\text{dd}}$ .

**Non-factorizable correlators:** the non-factorizable correlator introduced in eq. (3.43) can also be evaluated on the lattice by using the sequential propagators described above. Defining

$$\delta C_{\text{P}\ell}^{\text{nf}}(t, t_{\ell}) = e_{\text{q}_1} e_{\ell} \delta C_{\text{P}\ell}^{\text{nf}, \text{q}_1}(t, t_{\ell}) + e_{\text{q}_2} e_{\ell} \delta C_{\text{P}\ell}^{\text{nf}, \text{q}_2}(t, t_{\ell}), \quad (4.26)$$

and using eq. (4.11) one has

$$\delta C_{\text{P}\ell}^{\text{nf}, \text{q}_1}(t, t_{\ell}) = \langle \text{Tr} [S_{\text{q}_2}(0|-t)^{\dagger} \gamma_L^{\rho} S_{\text{q}_1}^A(0|-t)] \times S_{\nu}(t_{\ell}, \mathbf{p}_{\nu}|0) \gamma_L^{\rho} S_{\ell}^A(0|t_{\ell}, \mathbf{p}_{\ell}) \rangle, \quad (4.27)$$

$$\delta C_{\text{P}\ell}^{\text{nf}, \text{q}_2}(t, t_{\ell}) = \langle \text{Tr} [S_{\text{q}_2}^A(0|-t)^{\dagger} \gamma_L^{\rho} S_{\text{q}_1}(0|-t)] \times S_{\nu}(t_{\ell}, \mathbf{p}_{\nu}|0) \gamma_L^{\rho} S_{\ell}^A(0|t_{\ell}, \mathbf{p}_{\ell}) \rangle, \quad (4.28)$$

which correspond to the Feynman diagrams in figures 2(f) and 2(g), respectively. Here we have defined the (sequential) propagator of an anti-lepton with the insertion of an electromagnetic current and projected on definite external momentum as

$$S_{\ell}^A(0|t_{\ell}, \mathbf{p}_{\ell}) = i a^{\tau} \sum_{z, \mathbf{x}_{\ell}} S_{\ell}(0|z) \gamma^{\mu} A_{\mu}(z) S_{\ell}(z|t_{\ell}, \mathbf{x}_{\ell}) e^{-i\mathbf{p}_{\ell} \cdot \mathbf{x}_{\ell}}. \quad (4.29)$$

The tree-level correlator of eq. (3.38) evaluated at the simulated iso-symmetric point takes the form

$$C_{\text{P}\ell}(t, t_{\ell}) = \langle \text{Tr} [S_{\text{q}_2}(0|-t)^{\dagger} \gamma_L^{\rho} S_{\text{q}_1}(0|-t)] \rangle \times S_{\nu}(t_{\ell}, \mathbf{p}_{\nu}|0) \gamma_L^{\rho} S_{\ell}(0|t_{\ell}, \mathbf{p}_{\ell}). \quad (4.30)$$

Also in this case translational invariance has been used to simplify the notation such that the weak current is inserted in the origin. However, lattice correlators have been computed by inserting the weak current on all possible timeslices  $t_H/a = \{1, \dots, T/a = 96\}$  and at all positions  $\mathbf{x}_H$ , and then averaged over the volume. The lepton propagator has been computed for 8 different lepton source-sink separations  $t_{\ell}/a = \{12, 16, \dots, 40\}$  and its momentum is chosen in such a way that energy and momentum are conserved in the process.

Some comments concerning lattice lepton propagators are in order. First, we note that when evaluated on a torus, the lepton propagator  $S_\ell(0|t_\ell, \mathbf{p}_\ell)$  takes the form (neglecting possible contact terms)

$$S_\ell(0|t_\ell, \mathbf{p}_\ell) = - \sum_r \left[ \frac{e^{-\omega_\ell t_\ell}}{2\Omega_\ell} v_\ell^r(\mathbf{p}_\ell) \bar{v}_\ell^r(\mathbf{p}_\ell) + \frac{e^{-\omega_\ell(T-t_\ell)}}{2\Omega_\ell} u_\ell^r(-\mathbf{p}_\ell) \bar{u}_\ell^r(-\mathbf{p}_\ell) \right] \frac{1}{1 + e^{-\omega_\ell T}}. \quad (4.31)$$

The backward signal has a different Dirac structure compared to the forward one and  $(2\Omega_\ell)^{-1}$  appears in the residue at the pole, with  $\lim_{a \rightarrow 0} \Omega_\ell = \omega_\ell$ . Such a backward term would contribute to the traces in eq. (3.45). However, this contribution is not related to the matrix element  $\mathcal{M}_P^{rs}(\mathbf{p}_\ell)$  of our interest and therefore it has to be subtracted. To this end it is possible to define a projector  $\mathcal{P}_{v_\ell(\mathbf{p}_\ell)}$  only onto the forward-propagating part, namely

$$S_\ell(0|t_\ell, \mathbf{p}_\ell) \cdot \mathcal{P}_{v_\ell(\mathbf{p}_\ell)} = - \frac{e^{-\omega_\ell t_\ell}}{2\Omega_\ell} \sum_r v_\ell^r(\mathbf{p}_\ell) \bar{v}_\ell^r(\mathbf{p}_\ell) \frac{1}{1 + e^{-\omega_\ell T}}. \quad (4.32)$$

The definition and derivation of the projector  $\mathcal{P}_{v_\ell(\mathbf{p}_\ell)}$  is discussed in appendix C.3. Note that the same feature would appear also in the lattice neutrino propagator. However, being electrically neutral, the neutrino does not couple to the photon and, in addition, the term  $e^{-\omega_\nu t_\ell}/(2\omega_\nu)$  in its time-momentum representation (see eq. (3.39)) cancels in the ratio of eq. (3.45). Therefore we can amputate the neutrino propagator and substitute it with the (continuum) completeness relation  $[\sum_s u_\nu^s(\mathbf{p}_\nu) \bar{u}_\nu(\mathbf{p}_\nu)]^{\text{cont}} = -i\not{p}_\nu$ .

The lattice correlators employed in the numerical calculation are then defined as

$$\begin{aligned} \delta\tilde{\mathcal{C}}_{P\ell}^{\text{nf},q_1}(t, t_\ell) &= -i \langle \text{Tr} [S_{q_2}(0|-t)^\dagger \gamma_L^{\rho} S_{q_1}^A(0|-t)] \times \not{p}_\nu \gamma_L^{\rho} S_\ell^A(0|t_\ell, \mathbf{p}_\ell) \cdot \mathcal{P}_{v_\ell(\mathbf{p}_\ell)} \rangle, \\ \delta\tilde{\mathcal{C}}_{P\ell}^{\text{nf},q_2}(t, t_\ell) &= -i \langle \text{Tr} [S_{q_2}^A(0|-t)^\dagger \gamma_L^{\rho} S_{q_1}(0|-t)] \times \not{p}_\nu \gamma_L^{\rho} S_\ell^A(0|t_\ell, \mathbf{p}_\ell) \cdot \mathcal{P}_{v_\ell(\mathbf{p}_\ell)} \rangle, \\ \tilde{\mathcal{C}}_{P\ell}(t, t_\ell) &= -i \langle \text{Tr} [S_{q_2}(0|-t)^\dagger \gamma_L^{\rho} S_{q_1}(0|-t)] \rangle \times \not{p}_\nu \gamma_L^{\rho} S_\ell(0|t_\ell, \mathbf{p}_\ell) \cdot \mathcal{P}_{v_\ell(\mathbf{p}_\ell)}. \end{aligned} \quad (4.33)$$

The spectral decompositions of  $\delta\tilde{\mathcal{C}}_{P\ell}^{\text{nf}}(t, t_\ell)$  and  $\tilde{\mathcal{C}}_{P\ell}(t, t_\ell)$ , taking into account also the backward propagation of the meson on the torus, become<sup>6</sup>

$$\delta\tilde{\mathcal{C}}_{P\ell}^{\text{nf}}(t, t_\ell) = \frac{1}{L^3} \sum_{r,s} \frac{e^{-\omega_\ell t_\ell}}{4m_P \bar{\Omega}_\ell} \{e^{-m_P t} + \kappa_{P\ell} e^{-m_P(T-t)}\} Z_P u_\nu^r(\mathbf{p}_\ell) [\delta\bar{\mathcal{M}}_P^{rs}(\mathbf{p}_\ell)]^{\text{nf}} \bar{v}_\ell^s(\mathbf{p}_\ell), \quad (4.34)$$

$$\tilde{\mathcal{C}}_{P\ell}(t, t_\ell) = \frac{1}{L^3} \sum_{r,s} \frac{e^{-\omega_\ell t_\ell}}{4m_P \bar{\Omega}_\ell} \{e^{-m_P t} - e^{-m_P(T-t)}\} Z_P u_\nu^r(\mathbf{p}_\ell) \{-\mathcal{A}_P \mathcal{L}^{rs}(\mathbf{p}_\ell)\} \bar{v}_\ell^s(\mathbf{p}_\ell), \quad (4.35)$$

where  $\bar{\Omega}_\ell = \Omega_\ell(1 + e^{-\omega_\ell T})$  and  $\kappa_{P\ell}$  (which has a residual dependence on  $t_\ell$ ) parametrizes the correction to the matrix element due to the interaction of the backward propagating meson and the lepton. It follows that eq. (3.45) becomes

$$\mathcal{R}_{P\ell}^{\text{nf}}(t, t_\ell) = \text{Re} \left[ \frac{\text{Tr}[\gamma_L^0 \delta\tilde{\mathcal{C}}_{P\ell}^{\text{nf}}(t, t_\ell)]}{\text{Tr}[\gamma_L^0 \tilde{\mathcal{C}}_{P\ell}(t, t_\ell)]} \right] = \frac{\delta\mathcal{A}_P^{\text{nf}}}{\mathcal{A}_P} f_{P\ell}(t, T), \quad (4.36)$$

<sup>6</sup>Note that the spectral decomposition for  $\delta\tilde{\mathcal{C}}_{P\ell}^{\text{nf}}(t, t_\ell)$  given in eq. (4.34) is valid only for  $t < T - t_\ell$ . In this work we restrict the analysis of non-factorizable correlators in the region  $t < T/2$ , where the condition  $t < T - t_\ell$  is satisfied for all values of  $t_\ell$  used.

where

$$f_{\mathbf{P}\ell}(t, T) = \frac{1}{2} \{ (1 + \kappa_{\mathbf{P}\ell}) - (1 - \kappa_{\mathbf{P}\ell}) \coth [m_P(t - \frac{T}{2})] \} \approx 1 \quad \text{for } t \ll \frac{T}{2}. \quad (4.37)$$

For the lepton propagator we use the free Shamir DWF action [47] with  $aM_5 = 1.0$  and  $L_s/a = 8$ . The Feynman rules for the free DWF propagator have been derived in ref. [48] and we give details of the relevant Feynman rules in the conventions used in the Grid software framework [42, 43] in appendix C.1. We have determined the bare input mass for the lepton such that the pole mass of the free propagator corresponds to the physical muon mass  $m_\mu^\phi = 105.6583755$  MeV [1]. This results in a bare input lepton mass of  $am_\ell^{\text{input}} = 0.06107$  when using a previous determination of the lattice spacing  $a^{-1} = 1.730$  GeV [39]. Details on how to determine the input bare mass for a desired target pole mass of the free Shamir DWF propagator are given in appendix C.2.

We use twisted boundary conditions [49–52] for the lepton propagator in order to fix the momentum of the lepton such that energy and momentum are conserved at the weak Hamiltonian. This is the case when the momentum of the lepton is given by  $|\mathbf{p}_\ell| = \frac{m_P}{2} (1 - (m_\ell/m_P)^2)$  for the pseudoscalar meson at rest. For the determination of  $|\mathbf{p}_\ell|$  we used the physical mass for the muon  $m_\ell \equiv m_\mu^\phi$  and the simulation point masses  $m_P$  for pion and kaon as determined previously in ref. [53] on this gauge ensemble. We find  $a|\mathbf{p}_\ell| = 0.017054$  for the pion and  $a|\mathbf{p}_\ell| = 0.13783$  for the kaon. We distribute the momentum of the lepton equally in all three spatial directions, such that  $\mathbf{p}_\ell = -\frac{|\mathbf{p}_\ell|}{\sqrt{3}} \{1, 1, 1\}$ .

**Omega baryon correlators:** before closing the section we give details about the correlators for the  $\Omega^-$  baryon, which is employed in the renormalization conditions imposed in section 2 to fix the bare parameters of the QCD+QED, QCD and iso-QCD actions. We define the zero momentum two-point function as

$$C_{\Omega\Omega}(t) = \frac{a^3}{2} \sum_i \sum_{\mathbf{x}} \langle 0 | \text{T} [\psi_\Omega^i(t, \mathbf{x}) \bar{\psi}_\Omega^i(0)] | 0 \rangle, \quad (4.38)$$

where the operator  $\bar{\psi}_\Omega^\mu(x) = \psi_\Omega^{\mu\dagger}(x)\gamma^0$  denotes the spin-3/2 interpolating operator for the  $\Omega^-$  and we have summed over the spatial directions  $i$ . One form of baryon interpolator is given by

$$\psi_\Omega^\mu(x) = \epsilon^{abc} P_+ s_a(x) \left[ s_b^T(x) C \gamma^\mu s_c(x) \right] \quad (4.39)$$

where the  $s$  represent the strange quark fields,  $C$  is the charge conjugation matrix  $C = i\gamma_2\gamma_0$ , and Roman indices identify color components of the fields. The projector  $P_+ = (1 + \gamma^0)/2$  ensures that the interpolating operator  $\bar{\psi}_\Omega^\mu$  generates states with positive parity quantum number ( $\mathcal{P} = +1$ ) and annihilates states with negative parity quantum number ( $\mathcal{P} = -1$ ). In order to improve the signal for the correlation function, in this calculation we employ Gaussian smearing for the strange quark fields  $\tilde{s}(t, \mathbf{x}) = a^3 \sum_{\mathbf{y}} \exp[-(\mathbf{x} - \mathbf{y})^2/(2\sigma^2)] s(t, \mathbf{y})$  with a width of  $\sigma/a = 9$ , which requires gauge fixing of the QCD gauge configurations.

One feature of lattice baryon interpolating operators is that, on a torus, they couple to negative parity states propagating backward in time. As a consequence, assuming ground

state dominance, the correlator has the form

$$C_{\Omega\Omega}(t) = \left( |Z_{\Omega}|^2 e^{-m_{\Omega}t} + |\bar{Z}_{\Omega}|^2 e^{-\bar{\omega}_{\Omega}(T-t)} \right) P_+ \quad (4.40)$$

where  $\bar{\omega}_{\Omega}$  is the energy of the state with parity  $\mathcal{P} = -1$ . The operator-state overlaps for a state with spin projection  $s \in \{\pm\frac{3}{2}, \pm\frac{1}{2}\}$  are defined by  $Z_{\Omega} u_s^{\mu} = \langle 0 | \psi_{\Omega}^{\mu}(0) | \Omega, s \rangle$  and  $\bar{Z}_{\Omega} \gamma_5 u_s^{\mu} = \langle 0 | \psi_{\Omega}^{\mu}(0) | \bar{\Omega}, s \rangle$ , where  $u_s^{\mu}$  is the positive energy solution to the spin-3/2 Rarita-Schwinger equation (see e.g. [54] for a recent review), and  $|\Omega, s\rangle$  and  $|\bar{\Omega}, s\rangle$  are states with positive and negative parity respectively. In addition, quarks with anti-periodic boundary conditions in time have been assumed. Since baryon correlators are significantly affected by an exponential signal-to-noise-ratio problem, we restrict our analysis of the correlator to the time region  $t \ll T/2$ . In this interval we can then neglect the backward propagating signal and take for  $t \gg 0$ ,

$$\tilde{C}_{\Omega\Omega}(t) = \frac{1}{2} \text{Tr}[C_{\Omega\Omega}(t)] \approx |Z_{\Omega}|^2 e^{-m_{\Omega}t}. \quad (4.41)$$

In analogy with eq. (4.18), we can define the IB corrections to the correlator as

$$\begin{aligned} \delta\tilde{C}_{\Omega\Omega}(t) &= 4\pi\alpha_{\text{em}} \delta\tilde{C}_{\Omega\Omega}^{\text{em}}(t) + (\hat{m}_s^{\phi} - \hat{m}_s) \delta\tilde{C}_{\Omega\Omega}^{\mathcal{S},s}(t) \\ &= e_s^2 [\delta\tilde{C}_{\Omega\Omega}^{\text{self},s}(t) + \delta\tilde{C}_{\Omega\Omega}^{\text{exch}}(t)] + (\hat{m}_s^{\phi} - \hat{m}_s) \delta\tilde{C}_{\Omega\Omega}^{\mathcal{S},s}(t), \end{aligned} \quad (4.42)$$

where  $\tilde{C}_{\Omega\Omega}^{\text{self},s}(t)$  and  $\tilde{C}_{\Omega\Omega}^{\text{exch}}(t)$  denote the corrections due to the photon exchange between the constituent-strange quarks and  $\tilde{C}_{\Omega\Omega}^{\mathcal{S},s}(t)$  the correction given by the insertion of the quark scalar density on the quark lines. The ratio with the iso-QCD correlator has then the following asymptotic behaviour

$$\mathcal{R}_{\Omega\Omega}(t) = \frac{\delta\tilde{C}_{\Omega\Omega}(t)}{\tilde{C}_{\Omega\Omega}(t)} = 2 \frac{\delta Z_{\Omega}}{Z_{\Omega}} - \delta m_{\Omega} t. \quad (4.43)$$

Also in this case we can decompose the correction to the  $\Omega^-$  mass as

$$\delta m_{\Omega} = 4\pi\alpha_{\text{em}} \delta m_{\Omega}^{\text{em}} + (\hat{m}_s^{\phi} - \hat{m}_s) \delta m_{\Omega}^{\mathcal{S},s}. \quad (4.44)$$

Details on the quark contractions for the  $\Omega^-$  correlator, as well as a discussion on the derivation of its spectral decomposition can be found in appendix D.

## 5 Numerical analysis

The virtual IB corrections to the ratio of inclusive decay rates evaluated on the lattice, as defined in eq. (3.27), is built from the IB corrections to the kaon and pion decay amplitudes and to their masses. As discussed in the previous section, such quantities can be extracted from the large-time behaviour of suitably defined Euclidean lattice correlators. In this section, the strategy for extracting the relevant quantities from lattice correlators using a global-fit analysis is presented. Due to the various classes of correlators involved in this calculation, we adopt a data-driven approach to standardize the fitting criteria, which we explain below.

## 5.1 Strategy for correlator fits

Extracting physical quantities from lattice correlators using a fit procedure requires that optimal fit ranges are identified for each correlator. In our work, when multiple lattice correlators have fit parameters in common, e.g. the meson mass  $m_P$ , these data are fitted simultaneously fully taking into account such a constraint including the statistical correlation between the data. In this way, all parameters can be extracted from 7 independent frequentist fits.

In the case of the analyses of factorizable corrections, there are 12 correlators to study for the kaon, while for the pion the flavour symmetries of the correlation functions reduce the number of independent ones to 8. These correlators are listed below in eq. (5.4). The functional forms of the fit ansätze used for the correlators are based on the spectral decompositions eqs. (4.12), (4.13), (4.21) and (4.22), where only the ground-state contribution is included. For both mesons the tree-level correlators depend on two parameters, while all the factorizable correlators depend on 3 parameters each, namely a constant term containing the relative corrections to the matrix elements  $\mathcal{A}_P$  and  $Z_P$ , the correction to the meson mass and the simulation point mass  $m_P$  entering the tanh/coth functions in eqs. (4.23) and (4.24). The exact relation between the fit parameters and the physical quantities of interest is given in eqs. (5.5)–5.7. Since all the correlators for a given meson depend on the same simulation point mass, we combine the fits as described below. For what concerns the non-factorizable pion and kaon correlators, we decide instead to fit the ratios  $\mathcal{R}_{P\ell}^{\text{nf}}(t, t_\ell)$  using a constant fit ansatz, i.e. setting  $f_{P\ell}(t, t_\ell) = 1$  in eq. (4.36). This approximation corresponds to neglecting the contribution of backward signals and excited states and does not have a significant effect on the  $\chi^2$  for the range considered. In this case there is then only one parameter for each meson. The  $\Omega^-$  correlators, due to the usual rapidly degrading signal-to-noise ratio in baryon correlators, are also fitted in a region of small  $t$ , where we can safely neglect the contribution of the backward propagating baryon and excited states. This simplifies the fit ansätze for the tree level correlator  $\tilde{C}_{\Omega\Omega}$ , and the ratios  $\mathcal{R}_{\Omega\Omega}^{\text{em}}$  and  $\mathcal{R}_{\Omega\Omega}^{S,s}$  to those given in eqs. (4.41) and (4.43), respectively. Both the ansätze have two free parameters.

In order to select the best fit ranges we choose those with the maximum value for the Akaike Information Criterion (AIC) [55, 56] similarly to the strategy followed by refs. [18, 32, 57]

$$w = \exp \left[ -\frac{1}{2} (\chi^2 - 2n_{\text{dof}}) \right], \quad (5.1)$$

where  $n_{\text{dof}} = n_{\text{data}} - n_{\text{par}}$  is the number of degrees of freedom of the fit and the  $\chi^2$  function is defined as

$$\chi^2 = (\mathbf{C} - \mathbf{C}_M(\mathbf{a}))^T \boldsymbol{\Sigma}^{-1} (\mathbf{C} - \mathbf{C}_M(\mathbf{a})). \quad (5.2)$$

Here  $\mathbf{C}$  is a vector containing the data (i.e. the time correlators),  $\mathbf{C}_M(\mathbf{a})$  the corresponding model as a function of the fit parameters  $\mathbf{a}$  and  $\boldsymbol{\Sigma}$  the covariance matrix

$$\boldsymbol{\Sigma} = \frac{1}{n_B - 1} \sum_{i=1}^{n_B} (\mathbf{C}_i - \langle \mathbf{C} \rangle) (\mathbf{C}_i - \langle \mathbf{C} \rangle)^T, \quad (5.3)$$



with  $n_B$  the number of bootstrap samples. The AIC weight function favours fits that have minimal  $\chi^2$  with the largest  $n_{\text{dof}}$  possible, which penalises fits with a low  $\chi^2$  per degree of freedom resulting from over-fitting the data.

The datasets  $\mathbf{C}$  used for the 7 analyses can be summarized as follows

$$\begin{aligned}
 & \mathbf{1.} \quad \mathbf{C}_K^f = (C_{Kx}, \mathcal{R}_{Kx}^{\text{self,u}}, \mathcal{R}_{Kx}^{\text{self,s}}, \mathcal{R}_{Kx}^{\text{exch}}, \mathcal{R}_{Kx}^{S,u}, \mathcal{R}_{Kx}^{S,s}) \quad \text{with } x = (K, A), \\
 & \mathbf{2.} \quad \mathbf{C}_\pi^f = (C_{\pi y}, \mathcal{R}_{\pi y}^{\text{self,u}}, \mathcal{R}_{\pi y}^{\text{exch}}, \mathcal{R}_{\pi y}^{S,u}) \quad \text{with } y = (\pi, A), \quad \mathbf{3.} \quad \mathbf{C}_K^{\text{nf}} = (\mathcal{R}_{K\ell}^{\text{nf}}), \\
 & \mathbf{4.} \quad \mathbf{C}_\pi^{\text{nf}} = (\mathcal{R}_{\pi\ell}^{\text{nf}}), \quad \mathbf{5.} \quad \mathbf{C}_\Omega = (C_{\Omega\Omega}), \quad \mathbf{6.} \quad \mathbf{C}_\Omega^{\text{em}} = (\mathcal{R}_{\Omega\Omega}^{\text{em}}), \quad \mathbf{7.} \quad \mathbf{C}_\Omega^{S,s} = (\mathcal{R}_{\Omega\Omega}^{S,s}).
 \end{aligned} \tag{5.4}$$

The corresponding sets of fit parameters  $\mathbf{a}$  are

$$\begin{aligned}
 & \mathbf{1.} \quad \mathbf{a}_K^f = (\mathbf{a}_K, \mathbf{a}_K^{\text{self,u}}, \mathbf{a}_K^{\text{self,s}}, \mathbf{a}_K^{\text{exch}}, \mathbf{a}_K^{S,u}, \mathbf{a}_K^{S,s}), \\
 & \mathbf{2.} \quad \mathbf{a}_\pi^f = (\mathbf{a}_\pi, \mathbf{a}_\pi^{\text{self,u}}, \mathbf{a}_\pi^{\text{exch}}, \mathbf{a}_\pi^{S,u}), \quad \mathbf{3.} \quad \mathbf{a}_K^{\text{nf}} = (a_{K\ell}^{\text{nf}}), \\
 & \mathbf{4.} \quad \mathbf{a}_\pi^{\text{nf}} = (a_{\pi\ell}^{\text{nf}}), \quad \mathbf{5.} \quad \mathbf{a}_\Omega = (\mathbf{a}_{\Omega\Omega}), \quad \mathbf{6.} \quad \mathbf{a}_\Omega^{\text{em}} = (\mathbf{a}_{\Omega\Omega}^{\text{em}}), \quad \mathbf{7.} \quad \mathbf{a}_\Omega^{S,s} = (\mathbf{a}_{\Omega\Omega}^{S,s}),
 \end{aligned} \tag{5.5}$$

where we have defined

$$\mathbf{a}_P = (m_P, |Z_P|^2, \mathcal{A}_P Z_P), \quad \mathbf{a}_P^x = \left( \delta m_P^x, 2 \frac{\delta Z_P^x}{Z_P}, \frac{\delta \mathcal{A}_P^x}{\mathcal{A}_P} + \frac{\delta Z_P^x}{Z_P} \right), \quad a_{P\ell}^{\text{nf}} = \frac{\delta \mathcal{A}_P^{\text{nf}}}{\mathcal{A}_P}, \tag{5.6}$$

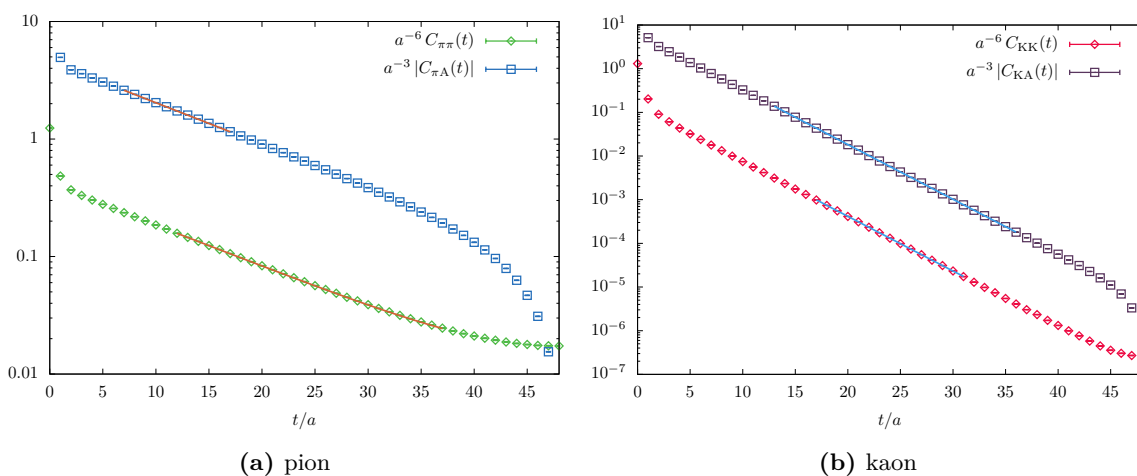
$$\mathbf{a}_{\Omega\Omega} = (m_\Omega, |Z_\Omega|^2), \quad \mathbf{a}_{\Omega\Omega}^x = \left( \delta m_\Omega^x, 2 \frac{\delta Z_\Omega^x}{Z_\Omega} \right). \tag{5.7}$$

In the case of the factorizable correlators, however, the bootstrap covariance is rank-deficient as the number of original samples  $n_{\text{cfg}} = 60$  is smaller than the dimension of the covariance matrix. Some form of regularisation is then required to make the  $\chi^2$ -problem well-conditioned. To this end we choose to neglect the covariance between the rows of  $\mathbf{C}_K^f$  and  $\mathbf{C}_\pi^f$  with and without photon lines. This choice is motivated by the fact that the correlation matrix is approximately block diagonal, and furthermore, we verified that the optimum parameters do not change significantly if correlation is also neglected between the correlation functions with different operator insertions. Finally, to reduce the number of degrees of freedom further, only a subsequence of correlator data separated by the *thinning* parameter  $\Delta t$  are included in the fit, which are reported for each fit in table 1. The regulated  $\chi^2$  thus defined, the best-fit parameters are determined by minimizing the  $\chi^2$  function for a given fit range.

The choice of the fit ranges for each correlator is made using two different approaches depending on the number of possibilities. For non-combined fits, like those on the  $\Omega$  correlators, the maximum number of fit ranges spanning the region  $t \in [0, T/2 - 1]$  (with  $T/a = 96$ ) is 1128. In the case of non-factorizable diagrams, including also all possible ranges in the lepton-time variable  $t_\ell$ , the maximum number of fit ranges is of  $\mathcal{O}(10^5)$ . In this case it is computationally feasible to do fits for all possible fit ranges and to compare the values of  $w$ . However, applying the same strategy to the combined factorizable fits would be computationally unfeasible, as the maximum number of possible fits is  $\mathcal{O}(10^{24})$  or  $\mathcal{O}(10^{36})$  for pion and kaon, respectively. To find good fit range(s) with large AIC we utilize

	$n_{\text{corr}}$	$n_{\text{par}}$	$n_{\text{dof}}$	$\Delta t/a$	$\chi^2$	$p$ -value
<b>1</b>	12	18	80	2	49.98	1.00
<b>2</b>	8	12	95	2	65.00	0.99
<b>3</b>	5	1	24	2	21.42	0.61
<b>4</b>	3	1	32	2	29.41	0.60
<b>5</b>	1	2	7	1	5.14	0.64
<b>6</b>	1	2	9	1	5.32	0.81
<b>7</b>	1	2	6	1	1.73	0.94

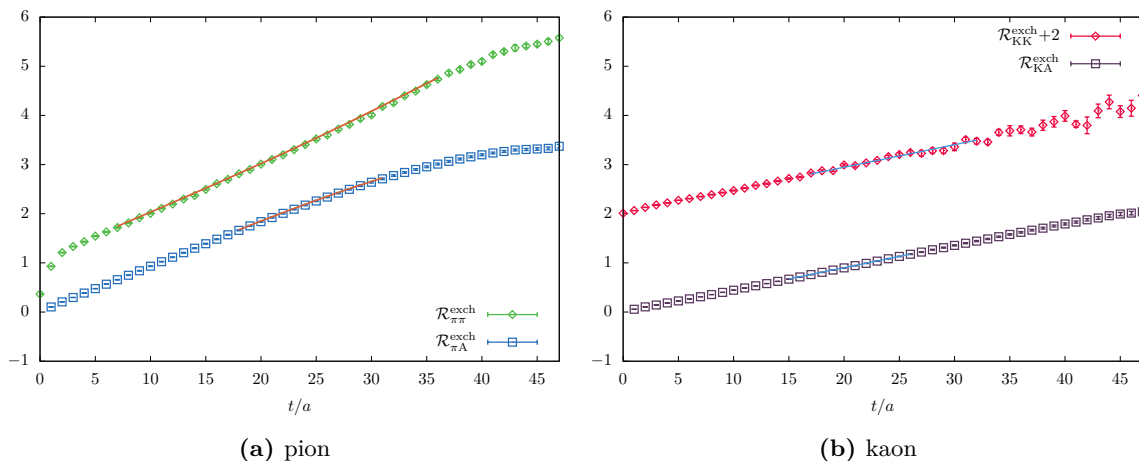
**Table 1.** Details of the fits with highest AIC weight for the 7 analyses performed in this work and presented in figures 4–8, 19 and 20: number of correlators ( $n_{\text{corr}}$ ), number of parameters ( $n_{\text{par}}$ ), number of degrees of freedom ( $n_{\text{dof}}$ ), size of thinning interval ( $\Delta t/a$ ), chi-squared ( $\chi^2$ ) and one-sided  $p$ -value.



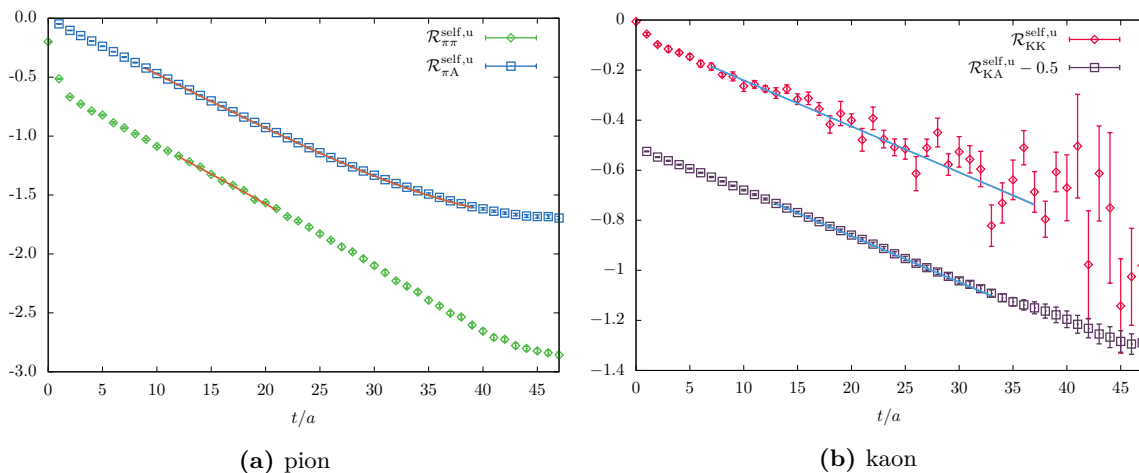
**Figure 4.** Tree-level correlators  $C_{PP}(t)$  and  $|C_{PA}(t)|$  for pion (a) and kaon (b). The solid lines with error band correspond to the best fits of the data.

a genetic algorithm as described in appendix E to perform the optimization. The outcome of this procedure is a set of fit ranges and their associated AIC weights  $w$  from each analysis. There is, however, a large multiplicity of good fit results. In order to capture the variability in the resulting good fits, we consider the 5 fits from each analysis that correspond to the highest AIC. This is an arbitrary and seemingly small number, which however already leads to a large multiplicity of  $n_{\text{fit}} = 5^7 = 78125$  alternative combinations for the fit parameters. The propagation of the variations due to these alternatives to the final results is discussed in section 5.3.

Here we only show the representative best fits of the correlators for each analysis, i.e. those corresponding to the highest AIC weight. In figure 4 the tree-level pion and kaon correlators of eqs. (4.12) and (4.13) are shown on a logarithmic scale, their slope being related to the tree-level meson mass  $m_P$ . The electromagnetic corrections due to the exchange of photons between the two constituent quarks and to the  $u$ -quark self-energy are reported in figures 5 and 6, respectively, normalized by the tree-level diagrams. In this



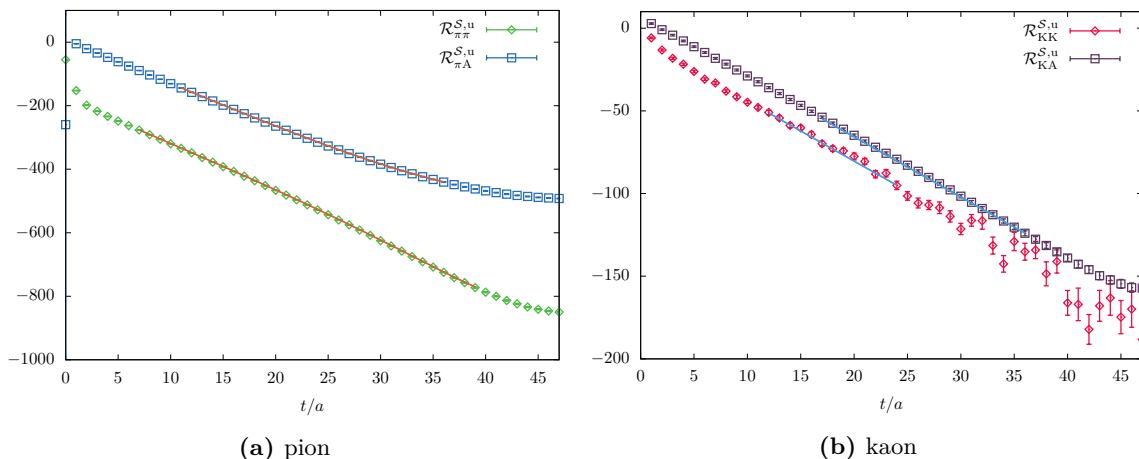
**Figure 5.** Factorizable diagram with a photon exchanged between the two constituent quarks,  $\mathcal{R}_{PP}^{\text{exch}}(t)$  and  $\mathcal{R}_{PA}^{\text{exch}}(t)$ , for pion (a) and kaon (b). The solid lines with error band correspond to the best fits of the data.



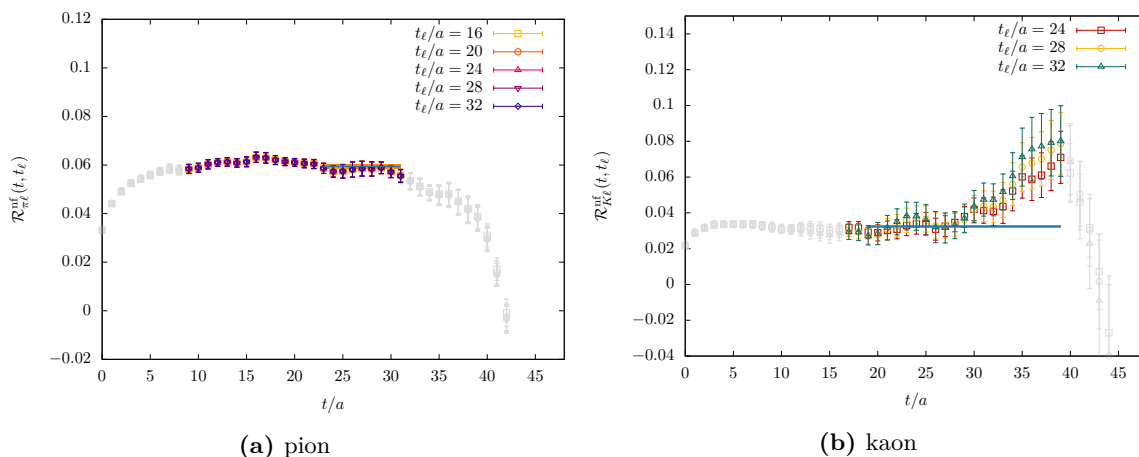
**Figure 6.** Factorizable diagram with  $u$ -quark self-energy correction,  $\mathcal{R}_{PP}^{\text{self},u}(t)$  and  $\mathcal{R}_{PA}^{\text{self},u}(t)$ , for pion (a) and kaon (b). The solid lines correspond to the best fits of the data.

case the slopes of the correlators correspond to the corrections to the meson mass  $\delta m_P^{\text{exch}}$  and  $\delta m_P^{\text{self},u}$  (see eqs. (4.21) and (4.22)). The correction due to the scalar insertion on the  $u$ -quark leg is shown instead in figure 7.

The non-factorizable correlators  $\mathcal{R}_{P\ell}^{\text{nf}}$  defined in eq. (4.36) are reported in figure 8 for both pions (left) and kaons (right). The expected time behaviour  $f_{P\ell}(t, T)$  is visible from the data, with plateaus in the region  $t \ll T/2$ . The dependence on the lepton source-sink time separation  $t_\ell$  is suppressed by the use of the projector on the forward propagating signal (see appendix C.3 for more details). The constant fits to the data corresponding to the highest value of the AIC weight are reported in the figures, while the grey points identify the data which are not included in any of the top 5 best fits selected in our analysis. The details for the best fits are reported in table 1 for the 7 analyses performed in this work.



**Figure 7.** Factorizable diagram with a scalar insertion along the  $u$ -quark line,  $\mathcal{R}_{PP}^{S,u}(t)$  and  $\mathcal{R}_{PA}^{S,u}(t)$ , for pion (a) and kaon (b). The solid lines correspond to the best fits of the data.



**Figure 8.** Non-factorizable diagram  $\mathcal{R}_{P\ell}^{nf}(t, t_\ell)$  for pion (a) and kaon (b), for multiple values of lepton source-sink separation  $t_\ell/a$ . The solid lines correspond to the best fits of the data. The grey points denote the data that are not included in any of the top 5 best fits.

## 5.2 Tuning of the bare parameters

From each of the fits performed in the factorizable analyses (1) and (2) outlined in eqs. (5.4) and (5.5) we obtain an estimate of the masses of the charged pion, the charged and neutral kaon and the neutral BMW mesons at the simulation iso-QCD point, together with their leading IB corrections. Analogously, we obtain the mass of the  $\Omega^-$  baryon and its corrections from the analyses (5), (6) and (7). Imposing the renormalization conditions in section 2, we can then obtain the relevant mass shifts  $(\hat{\mathbf{m}}^\phi - \hat{\mathbf{m}}^{(0)})$ ,  $(\hat{\mathbf{m}}^{\text{QCD}} - \hat{\mathbf{m}}^{(0)})$  and  $(\hat{\mathbf{m}}^\phi - \hat{\mathbf{m}}^{\text{QCD}})$  that allow one to define the IB correction  $\delta\hat{X}(\sigma^{(0)})$  to a given observable  $\hat{X}$ , as well as its decomposition into strong isospin-breaking and electromagnetic effects (see eqs. (2.14) and (2.15)).

The mass shift ( $\hat{\mathbf{m}}^\phi - \hat{\mathbf{m}}$ ) from the physical to the simulation point is obtained by imposing eq. (2.1) and simultaneously solving the following system of equations

$$\frac{M_j^2}{m_\Omega^2} \left[ 1 + 2\alpha_{\text{em}} \left( \frac{\delta M_j^{\text{em}}}{M_j} - \frac{\delta m_\Omega^{\text{em}}}{m_\Omega} \right) + 2 \sum_q \left( \frac{\delta M_j^{\mathcal{S},q}}{M_j} - \frac{\delta m_\Omega^{\mathcal{S},q}}{m_\Omega} \right) (\hat{m}_q^\phi - \hat{m}_q) \right] = \left( \frac{M_j^2}{m_\Omega^2} \right)^{\text{PDG}}, \quad (5.8)$$

where  $j = 1, 2, 3$  and  $\mathbf{M} = \{m_{\pi^+}, m_{K^+}, m_{K^0}\}$ . Finite-volume effects are applied to the meson masses on the right-hand side of eq. (5.8) making use of the formula in eq. (3.51). Once the vector ( $\hat{\mathbf{m}}^\phi - \hat{\mathbf{m}}$ ) is known, the QCD mass shifts ( $\hat{\mathbf{m}}^{\text{QCD}} - \hat{\mathbf{m}}$ ) are obtained from eq. (2.7) using the BMW mesons  $\mathbf{N} = \{M_{\text{ud}}, \Delta M, M_{K\chi}\}$  and solving the system

$$\begin{aligned} \frac{N_j^2}{m_\Omega^2} \left[ 1 + 2 \sum_q \left( \frac{\delta N_j^{\mathcal{S},q}}{N_j} - \frac{\delta m_\Omega^{\mathcal{S},q}}{m_\Omega} \right) (\hat{m}_q^{\text{QCD}} - \hat{m}_q) \right] = \\ \frac{N_j^2}{m_\Omega^2} \left[ 1 + 2\alpha_{\text{em}} \left( \frac{\delta N_j^{\text{em}}}{N_j} - \frac{\delta m_\Omega^{\text{em}}}{m_\Omega} \right) + 2 \sum_q \left( \frac{\delta N_j^{\mathcal{S},q}}{N_j} - \frac{\delta m_\Omega^{\mathcal{S},q}}{m_\Omega} \right) (\hat{m}_q^\phi - \hat{m}_q) \right], \end{aligned} \quad (5.9)$$

with  $j = 1, 2, 3$ . Finally, the iso-QCD point is determined solving the system in eq. (2.8) for ( $\hat{\mathbf{m}}^{(0)} - \hat{\mathbf{m}}$ ), namely for  $j = 1, 2, 3$

$$\begin{aligned} \frac{N_j^2}{m_\Omega^2} \left[ 1 + 2 \sum_q \left( \frac{\delta N_j^{\mathcal{S},q}}{N_j} - \frac{\delta m_\Omega^{\mathcal{S},q}}{m_\Omega} \right) (\hat{m}_q^{(0)} - \hat{m}_q) \right] = \\ \frac{N_j^2}{m_\Omega^2} \left[ 1 + 2\alpha_{\text{em}} \left( \frac{\delta N_j^{\text{em}}}{N_j} - \frac{\delta m_\Omega^{\text{em}}}{m_\Omega} \right) + 2 \sum_q \left( \frac{\delta N_j^{\mathcal{S},q}}{N_j} - \frac{\delta m_\Omega^{\mathcal{S},q}}{m_\Omega} \right) (\hat{m}_q^\phi - \hat{m}_q) \right] (1 - \delta_{j,2}). \end{aligned} \quad (5.10)$$

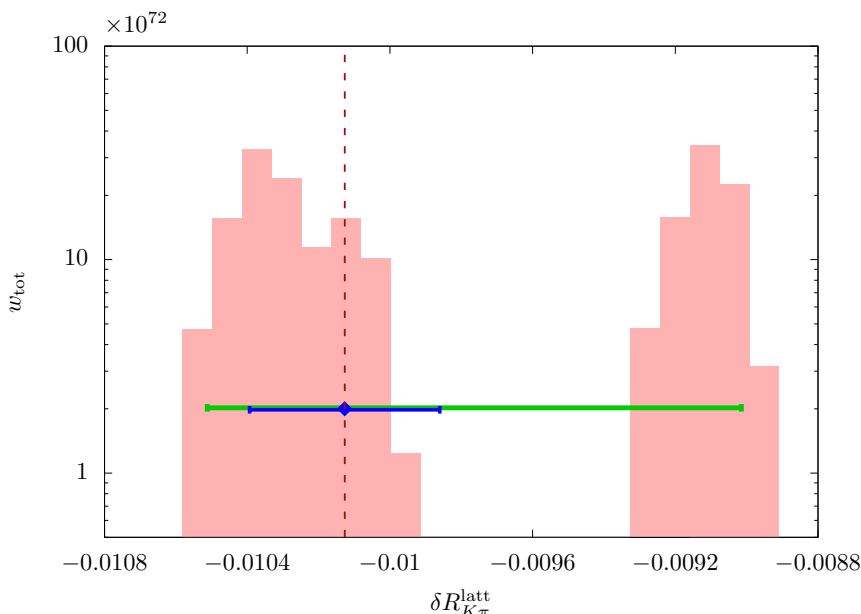
Using only the best fit from each of the analyses (i.e. the one corresponding to the highest AIC weight), we obtain the following bare quark masses in lattice units

$$\begin{pmatrix} \hat{m}_{\text{ud}}^{(0)} \\ \delta \hat{m}^{(0)} \\ \hat{m}_{\text{s}}^{(0)} \end{pmatrix} = \begin{pmatrix} 0.00068(2) \\ 0 \\ 0.0353(4) \end{pmatrix}, \quad \begin{pmatrix} \hat{m}_{\text{ud}}^{\text{QCD}} \\ \delta \hat{m}^{\text{QCD}} \\ \hat{m}_{\text{s}}^{\text{QCD}} \end{pmatrix} = \begin{pmatrix} 0.00068(2) \\ -0.0010(4) \\ 0.0353(4) \end{pmatrix}, \quad \begin{pmatrix} \hat{m}_{\text{ud}}^\phi \\ \delta \hat{m}^\phi \\ \hat{m}_{\text{s}}^\phi \end{pmatrix} = \begin{pmatrix} 0.00068(2) \\ -0.0010(4) \\ 0.0352(4) \end{pmatrix}. \quad (5.11)$$

The difference between the simulation point and the physical point is given by

$$\begin{pmatrix} \hat{m}_{\text{ud}} - \hat{m}_{\text{ud}}^\phi \\ \delta \hat{m} - \delta \hat{m}^\phi \\ \hat{m}_{\text{s}} - \hat{m}_{\text{s}}^\phi \end{pmatrix} = \begin{pmatrix} 0.00010(2) \\ 0.0010(4) \\ 0.0010(4) \end{pmatrix}, \quad (5.12)$$

and an important feature to notice is the similar size between the deviations in  $\hat{m}_{\text{ud}}$ ,  $\hat{m}_{\text{s}}$ , and  $\delta \hat{m}^\phi$ . This justifies the linearity assumption made in section 2.3, where we assumed that the  $\hat{m}_{\text{ud}}$  and  $\hat{m}_{\text{s}}$  corrections to match with the physical point were of the same size as the isospin-breaking effects.



**Figure 9.** Histogram displaying the distribution of  $\delta R_{K\pi}^{\text{latt}}$ . The blue and green error bands are the statistical and fit systematic errors, respectively.

Finally, we also obtain the following ratios

$$\left[ \frac{\mathbf{M}^{(0)}}{m_{\Omega}^{(0)}} \right]^2 = (0.006530 (4), 0.08761 (3), 0.08761 (3)), \quad (5.13)$$

$$\left[ \frac{\mathbf{M}^{\text{QCD}}}{m_{\Omega}^{\text{QCD}}} \right]^2 = (0.006530 (4), 0.08653 (2), 0.08869 (3)), \quad (5.14)$$

$$\left[ \frac{\mathbf{N}^{\phi}}{m_{\Omega}^{\phi}} \right]^2 = (0.006530 (4), -0.00464 (2), 0.08434 (2)). \quad (5.15)$$

Assuming  $m_{\Omega}^{(0)} = m_{\Omega}^{\phi}$ , we can form the ratio in eq. (5.13) using the iso-QCD meson masses in the GRS scheme quoted in ref. [8],

$$\left[ \frac{\mathbf{M}^{(0)}}{m_{\Omega}^{(0)}} \right]_{\text{GRS}}^2 = (0.00652 (2), 0.08746 (4), 0.08746 (4)). \quad (5.16)$$

The pion component agrees between the two schemes, the difference in the kaon part is more significant, but represents only a per-mille relative difference, which as we will see in section 6 is well covered by our systematic errors.

### 5.3 Estimation of model uncertainties

As described in section 5.1, the fit-scan procedure selects a set of fit ranges and their associated AIC weights from each of the seven analyses. We choose to consider the five best fits from each analysis, thus obtaining a total of  $n_{\text{fit}} = 78125$  determinations of the fit parameters for each bootstrap sample. We can then combine the fit parameters, tune

the bare-quark masses and use eq. (3.27) to get  $n_{\text{fit}}$  estimates of  $\delta R_{K\pi}^{\text{latt}}$  for each bootstrap. In order to extract a value from this set, we build a histogram of the  $n_{\text{fit}}$  values of  $\delta R_{K\pi}^{\text{latt}}$  reweighting each entry with the total AIC weight for that choice of analyses, namely

$$w_{\text{tot}} = \prod_{i=1}^7 w_i = \exp \left[ -\frac{1}{2} \sum_{i=1}^7 (\chi_i^2 - 2n_{\text{dof},i}) \right]. \quad (5.17)$$

Here the summation applies because the seven analyses are independent. The relative size of the  $n_{\text{fit}}$  different weights informs us which predictions are preferable over others. The choice of limiting our study to only the fit ranges associated to the top five AIC weights in each analysis is motivated by the fact that, with this reweighting procedure, the exponential suppresses the relatively inferior fit results. We checked, however, that varying the input parameters of the genetic algorithm used for the factorizable analyses and hence varying the best fits selected, yields consistent results (see appendix E). Given the reweighted histogram built from the  $n_{\text{fit}}$  values of  $\delta R_{K\pi}^{\text{latt}}$ , which is shown in figure 9, we determine the central value for this quantity as the median of the histogram. Choosing the median instead of the mean makes the result not subject to drastic variations due to outlier predictions. In figure 9 the median is indicated in blue together with its statistical error, while the green error bar is the fit systematics. The statistical error is estimated from the variance of the bootstrap samples of the medians, while the systematic error is determined from the distribution of  $\delta R_{K\pi}^{\text{latt}}$  as the  $2\sigma$  interval around the central value (i.e. the central 95% band). The distribution of  $\delta R_{K\pi}^{\text{latt}}$  in figure 9 shows two peaks. They suggest that there are two sets of fit intervals with statistically distinct fit results but with comparably good AIC weights. However, we note that both peaks are covered by our systematic error. Alternative strategies were attempted to stress the stability of our result, including different assumptions about correlation and different weight functions,<sup>7</sup> all leading to results within the quoted systematic uncertainty. The value obtained for  $\delta R_{K\pi}^{\text{latt}}$  is then

$$\delta R_{K\pi}^{\text{latt}} = -0.0101 (3)_{\text{stat.}} \binom{+11}{-4}_{\text{fit}}. \quad (5.18)$$

## 6 Results

The finite-volume lattice estimate of  $\delta R_{K\pi}^{\text{latt}}$  obtained in the previous section can be combined with the function  $\delta R_P^{(n)}(L)$  discussed in section 3.3 in order to subtract the logarithmic divergence and power-like electromagnetic finite-volume effects up to  $\mathcal{O}(1/L^2)$ . The prediction of  $\delta R_{K\pi}$  is then obtained according to eq. (3.8) by adding the contribution of the real-photon emission  $\delta R_{K\pi}^{\text{pert}}(\omega_\gamma^{\text{max}})$ , which is computed in perturbation theory [10] and reported in eq. (3.52). To evaluate the finite-volume correction, we compute eqs. (3.48) and (3.50) using the finite-volume coefficients determined in ref. [14] and the simulation point meson masses and decay constants, together with  $F_A^\pi$  and  $F_A^K$  from  $\chi$ Pt at  $\mathcal{O}(p^4)$  and  $\mathcal{O}(p^6)$ , respectively [11, 58, 59]. For our lattice of size  $L_{48} \equiv 48a$  we get

$$\delta R_{K\pi}^{(2)}(L_{48}) = \delta R_K^{(2)}(L_{48}) - \delta R_\pi^{(2)}(L_{48}) = -0.00730. \quad (6.1)$$

<sup>7</sup>We tried the flat distribution, the two-sided  $p$ -value, Bayesian model averaging [32, 57] and ad-hoc functions favouring high number of degrees of freedom with small  $\chi^2$ .

Evaluating eq. (3.52) for the physical values of the meson masses  $m_\pi$  and  $m_K$  [1] we obtain instead

$$\delta R_{K\pi}^{\text{pert}}(\omega_\gamma^{\text{max}}) = \delta R_K^{\text{pert}}(\omega_\gamma^{\text{max}}) - \delta R_\pi^{\text{pert}}(\omega_\gamma^{\text{max}}) = -0.00583. \quad (6.2)$$

Combining the previous results and including all sources of systematic uncertainties, which we are going to discuss in the rest of the section, our result for  $\delta R_{K\pi}$  obtained at  $L = L_{48}$  amounts to

$$\delta R_{K\pi} = -0.0086 (3)_{\text{stat.}} ({}^{+11}_{-4})_{\text{fit}} (5)_{\text{disc.}} (5)_{\text{quench.}} (39)_{\text{vol.}}. \quad (6.3)$$

The first error is statistical, and it is obtained from the variance of the bootstrap distribution of  $\delta R_{K\pi}^{\text{latt}}$ . The second error is the systematic uncertainty associated with our fit strategy and estimated as the  $2\sigma$  interval around the median of the distribution of  $\delta R_{K\pi}^{\text{latt}}$  (see figure 9), as discussed in section 5.3.

The calculation presented in this work has been performed on a single lattice spacing and, as a consequence, we are not able to extrapolate  $\delta R_{K\pi}^{\text{latt}}$  to the continuum limit. Thus, we quote a systematic uncertainty associated with the residual  $\mathcal{O}(a^2)$  discretization effects. This is estimated as  $(a\Lambda)^2$  with  $a^{-1} = 1730$  MeV and  $\Lambda = 400$  MeV [34]. This gives  $(a\Lambda)^2 \sim 5\%$ , which is applied to the central value of  $\delta R_{K\pi}^{\text{latt}}$  before the finite-volume subtraction and results in  $-0.0086 (5)_{\text{disc.}}$ .

Electromagnetic interactions involving sea quarks have been neglected in this work. Such electro-quenching effects are SU(3) and  $1/N_c$  suppressed for  $\mathcal{O}(\alpha_{\text{em}})$  contributions and expected to be of  $\sim \mathcal{O}(10\%)$  [16] of the QED correction to the rate. Separating  $\delta R_{K\pi}^{\text{latt}}$  into its electromagnetic and strong-isospin breaking contributions  $(\delta R_{K\pi}^{\text{latt}})^\gamma$  and  $(\delta R_{K\pi}^{\text{latt}})^{\text{SIB}}$  (according to the separation scheme outlined in section 2) we take the 10% of the e.m. part  $(\delta R_{K\pi}^{\text{latt}})^\gamma$  as our electro-quenching error. Using the median of the  $(\delta R_{K\pi}^{\text{latt}})^\gamma$  distribution,  $(\delta R_{K\pi}^{\text{latt}})^\gamma = -0.0047$ , we get  $-0.0086 (5)_{\text{quench.}}$ .

As discussed above, we use the finite-volume correction including the full  $1/L^2$  scaling (denoted by  $\delta R_P^{(2)}(L)$ ) in order to determine our central value for the infinite-volume observable  $\delta R_P$ . We then estimate the systematic uncertainty, associated with the truncation of the finite-volume expansion, by forming the difference between  $\delta R_P^{(2)}(L)$  and the correction including the point-like  $1/L^3$  contribution (denoted by  $\delta R_P^{(3),\text{pt}}(L)$ ). These quantities are given explicitly by combining eqs. (3.48)–(3.50) from section 3.3.

Since we are only targeting the difference between pion and kaon decay rates, the finite-volume correction we actually require is the difference

$$\delta R_{K\pi}^{(n)}(L) = \delta R_K^{(n)}(L) - \delta R_\pi^{(n)}(L). \quad (6.4)$$

The systematic uncertainty on this is then estimated via

$$\sigma_L \equiv \delta R_{K\pi}^{(3),\text{pt}}(L) - \delta R_{K\pi}^{(2)}(L), \quad (6.5)$$

$$= \frac{\alpha_{\text{em}}}{2\pi} \frac{32\pi^2}{(m_\pi L)^3} \left[ \frac{2 + (m_\ell/m_\pi)^2}{[1 + (m_\ell/m_\pi)^2]^3} - \frac{2 + (m_\ell/m_K)^2}{[1 + (m_\ell/m_K)^2]^3} \frac{m_\pi^3}{m_K^3} \right], \quad (6.6)$$

where we have given the explicit expression as it will play a crucial role in our error budget. We stress that  $\sigma_L$  is positive. As we will see below, both  $\delta R_{K\pi}^{(2)}(L)$  and the final observable



$\delta R_{K\pi}$  are negative. This implies that, if one were to estimate  $\delta R_{K\pi}$  using  $\delta R_{K\pi}^{(3),\text{pt}}(L)$ , the result would be reduced (a negative number with increased magnitude) as compared to the central value we report using  $\delta R_{K\pi}^{(2)}(L)$ .

To give numerical results for  $\delta R_{K\pi}^{(n)}$ , we require values for the meson masses and decay constants, the muon and  $W$ -boson mass, and also values for the form factors  $F_A^K$  and  $F_A^\pi$ . As above, we take  $F_A^K$  and  $F_A^\pi$  from  $\chi$ PT at  $O(p^6)$  and  $O(p^4)$ , respectively [11, 58, 59], and meson masses and decay constants from our simulation. The full set of inputs is then

$$\begin{aligned}
 m_\pi &= 0.1395 \text{ GeV}, & m_K &= 0.4992 \text{ GeV}, \\
 m_\mu &= 0.1057 \text{ GeV}, & m_W &= 80.38 \text{ GeV}, \\
 f_\pi &= 0.1310 \text{ GeV}, & f_K &= 0.1564 \text{ GeV}, \\
 F_A^{\pi, \chi\text{PT}} &= 0.0119, & F_A^{K, \chi\text{PT}} &= 0.0340,
 \end{aligned}
 \tag{6.7}$$

where results for  $F_A$  are reported to three digits, all other numbers to four digits and uncertainties are neglected, since these are completely subdominant in our determination.

Substituting these values into the expressions for  $\delta R_{K\pi}^{(n)}(L)$  and evaluating at the lattice volume used in this calculation,  $L_{48}$ , one finds

$$\delta R_{K\pi}^{(1)}(L_{48}) = -0.00468, \quad \delta R_{K\pi}^{(2)}(L_{48}) = -0.00730, \quad \delta R_{K\pi}^{(3),\text{pt}}(L_{48}) = -0.00337.$$

From these numerical results it is clear that the convergence appears quite poor for the volume used. In particular the ratio

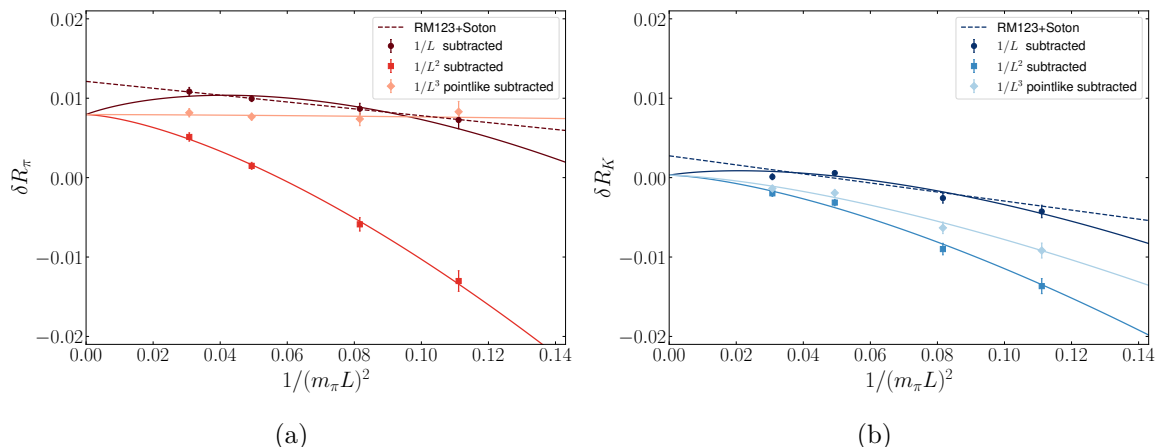
$$\frac{\sigma_{L_{48}}}{\delta R_{K\pi}^{(2)}(L_{48})} \simeq -0.54,
 \tag{6.8}$$

implies that the finite-volume correction is assigned a 54% systematic error in our method. As emphasized above, this is due to the fact that we have only incomplete knowledge of the correction through  $1/L^3$ , since the structure-dependent piece has not been calculated. Propagating this through eq. (3.8), we obtain

$$\delta R_{K\pi} = -0.0086 (39)_{\text{vol}}.
 \tag{6.9}$$

We close this section by presenting additional information on the finite-volume expansion, making use of the analytic results of ref. [14] as well as data from the previously published lattice calculation by the RM123S group [8]. This calculation uses a different lattice discretization and also extrapolates from heavier-than-physical pions. A key advantage relative to this work, however, is that it includes results at multiple volumes. The data are displayed in figure 10, separately for  $\delta R_\pi$  and  $\delta R_K$ . The results are for  $m_\pi \approx 320$  MeV and  $m_K \approx 580$  MeV, and four different volumes.

Our aim is to examine this data in light of a key conclusion of ref. [14], namely that the structure-dependent part of  $Y_{P,2}$  (the  $1/L^2$  coefficient) is numerically negligible. As this was not known at the time, the approach of ref. [8] was to subtract the point-like  $1/L$  prediction and to numerically investigate the residual volume dependence. The circular data points



**Figure 10.** The volume-dependence of (a)  $\delta R_\pi$  and (b)  $\delta R_K$  based on data taken from ref. [8] supplemented with analytic knowledge from ref. [14]. The  $L$  dependence is presented at fixed, heavier-than-physical quark masses corresponding to  $m_\pi \approx 320$  MeV and  $m_K \approx 580$  MeV. As indicated in the legend and described in detail in the text, the various points correspond to different subtractions and the curves to fits of residual  $L$  dependence.

in figure 10, labelled in the legend as “ $1/L$  subtracted”, show the result of this analysis and coincide to figure 9 of ref. [8]. As can be seen from the  $1/L$ -subtracted data and the dashed curves in figure 10, a linear description vs.  $1/L^2$  for the residual volume dependence is realistic. This results in a numerical prediction for the structure dependence that is much larger than the analytic result of ref. [14]. Another way to reach this same conclusion is to examine the residual  $L$  dependence in the pointlike  $1/L^2$  subtracted data. The fact that this shows a clear residual slope was interpreted as the effect of the structure dependence at  $1/L^2$ . This can also be seen in figure 10 in the square data points labelled “ $1/L^2$  subtracted” (strictly, here we subtract the full  $1/L^2$  behaviour and ref. [8] the point-like part, but the distinction is numerically insignificant.)

We argue that the puzzle is resolved by the observation that the data is equally well described by  $1/L^3$  behaviour. To explore this we first subtract the point-like  $1/L^3$  prediction and find that the  $L$  dependence is reduced. This is shown in figure 10 as the diamonds, labelled in the legend as “ $1/L^3$  pointlike subtracted”. We then perform a fit of the form  $a + b/L^3$  to the point-like  $1/L^3$ -subtracted data. We find this describes the data reasonably and can be interpreted as an estimate of the residual  $1/L^3$  behaviour, again arising from structure dependence.<sup>8</sup> The three solid curves in figure 10 show the result of the  $a + b/L^3$  fit for each of the three subtraction scenarios. We stress that the curves are related by analytic terms and that only one fit was performed.

From these considerations, we conclude that the  $L \rightarrow \infty$  limit is challenging for  $\text{QED}_L$  and that analytical knowledge of the  $L$  dependence can be of great importance in controlling the systematic error associated to these extrapolations. We are working on several directions

<sup>8</sup>Given the discussion above, the reader might note that we are mimicking the approach of ref. [8] but one order higher in  $1/L$ . To this point we stress one key difference; the point-like  $1/L^n$  contributions are known to vanish for  $n > 3$ .

to address this issues, including an analytic determination of the structure dependence at  $1/L^3$ , both in the continuum and at finite lattice spacing (where the interplay between cut-off effects and finite volume effects is not trivial due to the non-locality of QED<sub>L</sub> [60]). As is discussed in ref. [14], this will require evaluating a branch-cut contribution (similar to that appearing for the pseudoscalar mass in eq. (3.51)). In addition, when working with physical pion masses at such level of precision, exponentially suppressed finite-volume effects might become relevant and require a careful investigation.

We now turn to the determination of  $|V_{us}|/|V_{ud}|$ . For this purpose, symmetrizing the fit systematic in eq. (6.3) and summing in quadrature all the errors but the “vol.” one, we get  $\delta R_{K\pi} = -0.0086 (13)(39)_{\text{vol.}}$ . Combining this result with the value of the iso-QCD ratio  $f_K/f_\pi$  we can predict  $|V_{us}|/|V_{ud}|$  at leading order in IB corrections as

$$\frac{|V_{us}|}{|V_{ud}|} = \left[ \frac{\Gamma(K^+ \rightarrow \mu^+ \nu_\mu [\gamma])}{\Gamma(\pi^+ \rightarrow \mu^+ \nu_\mu [\gamma])} \frac{m_K}{m_\pi} \frac{(m_\pi^2 - m_\mu^2)}{(m_K^2 - m_\mu^2)} \right]^{1/2} \frac{f_\pi}{f_K} \left( 1 - \frac{1}{2} \delta R_{K\pi} \right). \quad (6.10)$$

Averaging<sup>9</sup> the  $N_f = 2 + 1$  lattice results reviewed in FLAG [2, 39, 61–67], and using the PDG average for the ratio of experimental decay widths [1], we obtain

$$|V_{us}|/|V_{ud}| = 0.23154 (28)_{\text{exp.}} (15)_{\delta R_P} (45)_{\delta R_P, \text{vol.}} (65)_{f_P}, \quad (6.11)$$

where the first error comes from the experimental measurements, the second is our uncertainty on  $\delta R_{K\pi}$  excluding the finite-volume systematics quoted separately, and the last error comes from the average of lattice determinations for  $f_K/f_\pi$ . Interestingly, we find that the error from  $f_K/f_\pi$  dominates the uncertainty on  $|V_{us}|/|V_{ud}|$ . The same conclusion is obtained using the RM123S result. In fact, taking  $\delta R_{K\pi}^{\text{RM123S}} = -0.0126 (14)$  and the  $N_f = 2 + 1 + 1$  average  $f_K/f_\pi = 1.1966 (18)$  from ref. [8], one obtains  $(|V_{us}|/|V_{ud}|)^{\text{RM123S}} = 0.23131 (28)_{\text{exp}} (17)_{\delta R_P} (35)_{f_P}$ . This is a clear motivation for future new computations of  $f_K/f_\pi$  on the lattice, with the aim of reducing the uncertainty by a factor 2 to 3 to bring it below the current experimental uncertainties on the decay width ratio. Finally, the second-largest uncertainty in eq. (6.11) comes from the challenges with finite-volume QED as discussed above. It is foreseeable that this conservative uncertainty will be drastically reduced in the near future, which can be done through the addition of multiple volumes to compute the  $1/L^3$  coefficient or the usage of a different QED formulation with smaller volume corrections. In conclusion, there are identified ways forward to reduce in the short-term future the two main systematic errors on  $|V_{us}|/|V_{ud}|$ , and beyond those the precision reached on  $\delta R_{K\pi}$  is sufficient and below the experimental input uncertainties.

## 7 Conclusions

The study of light-meson leptonic decays is of great relevance for the extraction of the CKM matrix elements  $|V_{us}|$  and  $|V_{ud}|$ , especially in light of current outstanding  $3\sigma$  tensions

---

<sup>9</sup>FLAG does not quote an average for  $f_K/f_\pi$ , but for the isospin-corrected ratio  $f_{K^+}/f_{\pi^+}$ . We produced the value  $f_K/f_\pi = 1.1930 (33)$  following exactly the averaging procedure described in the review. Although iso-QCD has been tuned in slightly different ways in the calculations entering this average, from the corresponding values of  $m_\pi$  and  $m_K$  we expect scheme ambiguities to be below the quoted uncertainty (see discussion in sections 2.4 and 5.2).

in the first-row unitarity [1, 3]. To either confirm or resolve such tensions, a combined effort of both theory and experiment is necessary. New experimental measurements and analyses are possible for some facilities (e.g. at NA62, as suggested by the authors of ref. [3]) and can help to clarify the situation. On the theoretical side, precise and controlled calculations of leptonic and semi-leptonic decay rates, including non-perturbative effects of strong interactions, as well as QED and strong isospin-breaking effects, would allow stringent tests of the SM.

In this paper we have presented the first physical-quark-mass lattice calculation of the leading isospin-breaking effects on the ratio of the rates of kaon and pion decays into muons. This has been performed using chiral domain wall fermions with close-to-physical masses on a single gauge ensemble, i.e. at a fixed value of the lattice spacing and on a finite volume. Finite-volume QED interactions have been regulated according to the QED<sub>L</sub> prescription by removing the spatial zero mode of the photon propagator and the electro-quenched approximation has been employed, thus assigning zero electric charge to the sea quarks. Including all sources of systematic uncertainty, we obtain

$$\delta R_{K\pi} = -0.0086 (3)_{\text{stat.}} (+11)_{\text{fit}} (5)_{\text{disc.}} (5)_{\text{quench.}} (39)_{\text{vol.}} . \quad (7.1)$$

This result is compatible with the lattice result obtained by the RM123S collaboration [8], as well as with the  $\chi$ PT estimate of ref. [13].

Although our statistical error is very competitive with e.g. the RM123S calculation, the final precision of our estimate of  $\delta R_{K\pi}$  is affected by a large systematic uncertainty. This is dominated by the error associated with residual finite-volume effects, which amounts to around 45% of the central value of  $\delta R_{K\pi}$ . The origin of such a large uncertainty, as explained in section 6, is due to the lack of knowledge of structure-dependent effects at  $O(1/L^3)$ , which are specific to the QED<sub>L</sub> prescription. The discussion in section 6 emphasises the crucial role of finite-volume effects in the extraction of  $\delta R_{K\pi}$  and the need for a dedicated study of the  $O(1/L^3)$  contributions. Two ways of reducing the finite-volume systematic error will be explored in future calculations. On the one hand, work is in progress to understand and determine the  $1/L^3$  finite-volume QED<sub>L</sub> contributions. On the other hand, performing the same calculation on multiple volumes can certainly help to extrapolate to the infinite-volume limit. Repeating the calculation on gauge ensembles with different lattice spacings would also allow to reduce the systematic uncertainties associated to discretization effects. For what concerns electro-quenching, a plan is in place to overcome this approximation calculating quark-disconnected electromagnetic corrections. The progress of our preliminary study has been reported in ref. [30].

To conclude, our calculation provides an important step towards future flavour physics precision tests. The anticipated extensions of the calculation presented in this work, resulting in smaller systematic uncertainties, will allow for a new theoretical prediction for the ratio  $|V_{us}|/|V_{ud}|$ . However, as discussed at the end of section 6, a real progress will only be possible if also the precision of the iso-QCD decay constants  $f_K/f_\pi$  is improved. At this point, the uncertainties coming from theoretical predictions will no longer dominate over those from experimental inputs in the extraction of  $|V_{us}|/|V_{ud}|$ . We are currently also investigating the prospects for a non-perturbative determination of the leading isospin-breaking corrections

to semi-leptonic  $K \rightarrow \pi \ell \nu$  decays, which are relevant for an independent determination of  $|V_{us}|$ . Together, these results will provide novel and stringent precision tests of the CKM matrix unitarity.

## Acknowledgments

We warmly thank Luigi Del Debbio, Fabian Joswig and our colleagues in the RBC and UKQCD collaborations for many helpful discussions. We also thank the RM123-Soton collaboration for kindly providing lattice data (from ref. [8]) for our study of finite-volume effects. V.G. thanks M. Tomii and J. Flynn for useful discussions about free domain wall fermion propagators. N. H.-T. wishes to thank the Higgs Centre for Theoretical Physics at the University of Edinburgh for hosting visits where part of this work was completed. This work used the DiRAC Extreme Scaling service at the University of Edinburgh, operated by the Edinburgh Parallel Computing Centre on behalf of the STFC DiRAC HPC Facility ([www.dirac.ac.uk](http://www.dirac.ac.uk)). This equipment was funded by BEIS capital funding via STFC capital grant ST/R00238X/1 and STFC DiRAC Operations grant ST/R001006/1. DiRAC is part of the National e-Infrastructure. P.B. has been supported in part by the U.S. Department of Energy, Office of Science, Office of Nuclear Physics under the Contract No. DE-SC-0012704 (BNL). M.D.C., F.E., T.H., V.G., M.T.H., and A.P. are supported in part by UK STFC grant ST/P000630/1. Additionally M.T.H. is supported by UKRI Future Leader Fellowship MR/T019956/1. F.E., V.G., R.H., F.Óh., A.P. and A.Z.N.Y. received funding from the European Research Council (ERC) under the European Union’s Horizon 2020 research and innovation programme under grant agreement No 757646 and A.P. additionally under grant agreement No 813942. N. H.-T. is funded in part by the Albert Einstein Center for Fundamental Physics at the University of Bern, and in part by the Swedish Research Council, project number 2021-06638. A.J. and J.R. have been supported in part by UK STFC grant ST/P000711/1 and ST/T000775/1. J.R. is also supported in part by UK STFC DiRAC operational grants ST/S003762/1 and ST/W002701/1.

## A Comparison of local and conserved electromagnetic currents

### A.1 Theory

A certain freedom always exists in the detailed choice of how to discretize local composite fields in a lattice calculation. In particular, various equally valid discretizations can be defined that differ in their renormalization and cut-off effects. In this work, we use the (ultra-)local discretization of the electromagnetic current on the lattice, defined as

$$V_{fg}^\mu(x) = Z_V \bar{\psi}_f(x) \gamma^\mu \psi_g(x). \tag{A.1}$$

This is an extension of the current appearing in eq. (4.5), as here we allow the possibility of an off-diagonal flavour current in order to simplify the discussion of particular quark contractions in isolation.

The local current does not exactly satisfy the QED Ward-Takahashi identity. In other words, the coupling to the photon field breaks QED gauge invariance explicitly for non-zero

lattice spacing. In addition to introducing a finite renormalization of the electromagnetic current at order  $\alpha_{\text{em}}$ , the lack of gauge symmetry implies that new singularities may arise when the position of the current coincides with other local fields. Such short-distance effects occur, for example, when the vertex is integrated over the space-time volume and as a result coincides with the axial current as in the first correlation function defined in eq. (4.19). By contrast, when gauge-invariance is preserved using a discretization of the current which is exactly conserved, singularities associated with overlapping operators are highly constrained by the Ward-Takahashi identities.

Nevertheless, by power counting one can show that in our set-up no such extra divergences arise, nor is the automatic  $O(a)$  improvement of the chiral fermion discretization spoiled. To see this, first consider diagram (a) of figure 2. We examine the limit in which both electromagnetic vertices approach the position of the axial current. To identify this diagram in isolation we introduce fictitious valence-quark flavours denoted 1, 2, 3, 4 (discussed in more detail below) and write

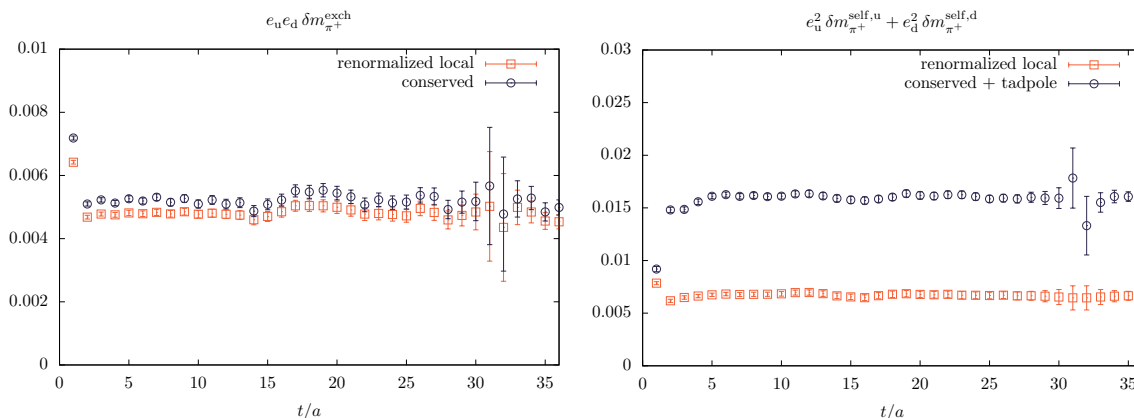
$$a^8 \frac{1}{x^2} \left[ V_{12}^\mu(x) V_{23}^\mu(x) A_{34}^\nu(0) \right] \stackrel{x \rightarrow a}{\equiv} \delta Z_A A_{14}^\nu(0) + O(a^2). \tag{A.2}$$

Here the factor  $x^{-2}$  on the left-hand side arises from the short-distance behaviour of the photon propagator and the  $a^8$  arises from the discretized space-time measure. The three key claims in this equation, all justified in the following paragraphs, are (i) that no power divergences (positive powers of  $1/a$ ) arise, (ii) that the constant order simply defines a contribution to the renormalization of the axial current, and (iii) that the leading corrections that vanish as  $a \rightarrow 0$  are  $O(a^2)$  rather than  $O(a)$ .

We have introduced additional (degenerate) flavours 1, 2, 3, 4 in the paradigm of a partially quenched theory to isolate the contribution from the diagram of interest. This flavour structure ensures that only operators with energy dimension greater than or equal to three can contribute to the right-hand side, since all contributing operators must carry anti-1 and 4 quantum numbers, and must therefore be built from at least two quarks. The difference between the lowest dimension of operators contributing (three) and the dimension on the left-hand side (nine) is therefore six, and this leads to a  $1/a^6$  scaling accompanying the quark bilinear. This is however cancelled by the power of  $a^8/x^2 \rightarrow a^6$  on the left-hand side, implying that no power divergences arise. This demonstrates point (i) above.

Without any additional symmetries, all rotationally covariant quark bilinears could contribute. However, in our set-up we have an approximate chiral symmetry, broken due to an exponentially suppressed contribution from the finite extent of the fifth dimension. Taking this to be negligible, we need only catalogue dimension three operators with the correct chiral rotation properties, and the {1,4} axial-current is the unique choice with dimension three. Thus, the effect of breaking gauge invariance results in an additional renormalization of the axial current at the next-to-leading order in the electromagnetic coupling. This demonstrates point (ii) above.

Finally, the discrete lattice chiral symmetry [68] forbids mixing with dimension-four operators with the appropriate definition [69], which might otherwise introduce linear lattice artefacts in such off-shell correlation functions. This is our third and final point (iii) and



**Figure 11.** A comparison of the QED correction to the pion mass from the exchange (left) and the self energy+tadpole (right) diagrams between the local and conserved currents.

a similar analysis of the remaining diagrams illustrates that the use of the local current poses no particular difficulties with our chosen discretization. We now turn to a numerical demonstration that the discrepancy in  $\delta R_{K\pi}$  between this current and the conserved vector current at fixed lattice spacing has a value consistent with our expectations for an  $O(a^2)$  effect.

### A.2 Numerical check

We perform a numerical test on a smaller  $24^3 \times 64$  lattice using Shamir-Domain-Wall fermions [70], with  $aM_5 = 1.8$  and  $L_s/a = 16$ . We limit the statistics to 10 QCD configurations, with interpolating operator inserted on every other timeslice (32 in total). The pion mass for this ensemble is  $m_\pi \approx 339.789$  MeV.

The difference between the formulation of local and conserved electromagnetic current is the presence of a tadpole diagram in the latter, which arises from the second derivative of the Dirac operator with respect to the electric charge. We may extract the contributions to the QED mass corrections from correlator ratios via

$$\delta m_P^x = -\frac{\mathcal{R}_{PP}^x(t+1) - \mathcal{R}_{PP}^x(t)}{f_{PP}(t+1, T) - f_{PP}(t, T)}, \tag{A.3}$$

where  $x = \{\text{self}, q_1; \text{self}, q_2; \text{exch}\}$  while the ratios  $\mathcal{R}_{PP}^x(t)$  and the function  $f_{PP}(t, T)$  are defined in eqs. (4.22) and (4.24), respectively. The result for pions is shown in figure 11. For the exchange diagram where the electromagnetic current is inserted on both propagators, we notice that the use of the renormalized local current or the conserved current give very similar results. However, in the case where there are two current insertions on the same quark propagator, the presence of the tadpole contributes additionally to the mass correction, as expected. This discrepancy will manifest in the results obtained from the combined fits performed on the tree-level and factorizable correlators.

Since, for this numerical check, we are working on a gauge ensemble away from the physical point we simply define the iso-QCD point to be equal to the simulation one,

	local	conserved
$(\delta\hat{m}_{\pi^+}^2)^{\text{e.m.}}$	0.005476 (38)	0.009160 (30)
$(\delta\hat{M}_{\text{uu}}^2)^{\text{e.m.}}$	0.000434 (39)	0.005782 (29)
$(\delta\hat{m}_{K^+}^2)^{\text{e.m.}}$	0.008148 (47)	0.012260 (54)
$(\delta\hat{m}_{K^0}^2)^{\text{e.m.}}$	0.0005142 (98)	0.002250 (13)
$(\delta\mathcal{A}_{\pi^+}/\mathcal{A}_{\pi})^{\text{f}}$	$5.388 (49) \times 10^{-2}$	$1.1743 (64) \times 10^{-1}$
$(\delta\mathcal{A}_{K^+}/\mathcal{A}_K)^{\text{f}}$	$2.218 (48) \times 10^{-2}$	$5.265 (57) \times 10^{-2}$
$(\delta\mathcal{A}_{\pi^+}/\mathcal{A}_{\pi})^{\text{nf}}$	$5.374 (59) \times 10^{-2}$	$5.287 (43) \times 10^{-2}$
$(\delta\mathcal{A}_{K^+}/\mathcal{A}_K)^{\text{nf}}$	$4.493 (41) \times 10^{-2}$	$4.494 (48) \times 10^{-2}$

**Table 2.** Comparison of photon corrections to meson masses and to the decay amplitude computed using local and conserved current.

i.e.  $\sigma^{(0)} = \sigma$ , such that  $m_P^{(0)} = m_P$ . Furthermore, the (fictitious) physical point  $\sigma^\phi$  is defined imposing the following conditions

$$(m_{\pi^+}^\phi)^2 = m_\pi^2 + (m_{\pi^+}^{\text{PDG}})^2 - (m_{\pi^0}^{\text{PDG}})^2, \quad (\text{A.4})$$

$$(m_{K^+}^\phi)^2 = m_K^2 + \frac{1}{2} \left[ (m_{K^+}^{\text{PDG}})^2 - (m_{K^0}^{\text{PDG}})^2 \right], \quad (\text{A.5})$$

$$(m_{K^0}^\phi)^2 = m_K^2 - \frac{1}{2} \left[ (m_{K^+}^{\text{PDG}})^2 - (m_{K^0}^{\text{PDG}})^2 \right], \quad (\text{A.6})$$

i.e. we keep the pion and kaon mass splittings at their experimentally measured values (taken from PDG [1]). These conditions allow us to obtain the quark mass shifts ( $\hat{\mathbf{m}}^\phi - \hat{\mathbf{m}}$ ) needed to compute IB corrections. The physical value of  $\alpha_{\text{em}}$  is tuned instead to its Thomson limit, as done in section 2.

In table 2 we report the photon corrections to the meson masses, as well as the factorizable and non-factorizable contributions to the decay amplitude, obtained using either the conserved electromagnetic current or the renormalized local vector current. We see that all diagrams except the non-factorizable correction give as expected significantly different results using the two different approaches. However, when combining these corrections with those obtained from the insertion of the scalar density,  $\sum_x (\hat{m}_q^\phi - \hat{m}_q) \hat{\mathcal{S}}_q(x)$  (see eq. (4.5)), the estimates for a physical observable obtained with the two approaches become comparable. In fact, this is the case for  $\delta R_{K\pi}^{\text{latt}}$ ,

$$(\delta R_{K\pi}^{\text{latt}})_{\text{loc}} = -7.04 (20) \times 10^{-3}, \quad (\delta R_{K\pi}^{\text{latt}})_{\text{cons}} = -6.91 (20) \times 10^{-3}. \quad (\text{A.7})$$

We can see that the two results are compatible with each other, the difference  $(\delta R_{K\pi}^{\text{latt}})_{\text{loc}} - (\delta R_{K\pi}^{\text{latt}})_{\text{cons}}$  being consistent with zero within errors. The slightly larger value ( $\sim 2\%$ ) of  $(\delta R_{K\pi}^{\text{latt}})_{\text{loc}}$  can be associated to  $O(a^2)$  cut-off effects, which as explained in the previous subsection are expected to contribute.

## B zMöbius to Möbius correction

The zMöbius DWF action [41] is an approximation of the Möbius DWF action [38] and is used in this work due to faster numerical convergence. The real parameters of the Möbius



DWF action are matched to complex ones in the zMöbius DWF action, using the Remez algorithm, leading to a reduced  $L_s$  dimension. On the ensemble used in this work, an  $L_s/a = 24$  is used for Möbius and  $L_s/a = 10$  for zMöbius.

A further drastic improvement in the iterations needed for a light-quark inversion is achieved via deflation: we compute the lowest  $N_{\text{vec}} = 2000$  eigenvectors of the Dirac operator to obtain a starting guess, reducing computational cost of light-quark inversions substantially.

For the light-quark inversions using the Möbius action, for which we do not have eigenvectors available on disk, we employ the Möbius accelerated DWF (MADWF) algorithm [71]. This algorithm constructs a guess for the final solve by transforming the 5D Domain-Wall Dirac operator  $D_{\text{DW}}^{5\text{D}}$  via Pauli-Villars solves into a 4D approximation of the overlap operator  $D_{\text{ov}}^{4\text{D}}$ . The solution of the  $D_{\text{ov}}^{4\text{D}}$  inversion is then used to reconstruct an approximated solution for  $D_{\text{DW}}^{5\text{D}}$ . Using this solution as a guess for the final solve on  $D_{\text{DW}}^{5\text{D}}$  leads to an overall reduction in computational cost. One key insight used in this work is that the Domain-Wall Dirac operator  $D_{\text{DW}}^{5\text{D}}$  does not have to be the same in the first and last step of this algorithm. For our light-quark solves with the Möbius action, we therefore produced the guess of the MADWF algorithm using a zMöbius Dirac operator, allowing us to benefit from deflation. We found that the zMöbius MADWF guess was able to significantly speed up the final Möbius solve. Compared to an undeflated light-quark Möbius solve, the deflated zMöbius solve has an iteration count reduced by a factor 20 and the MADWF Möbius solve with a deflated zMöbius MADWF guess is faster by a factor 10.

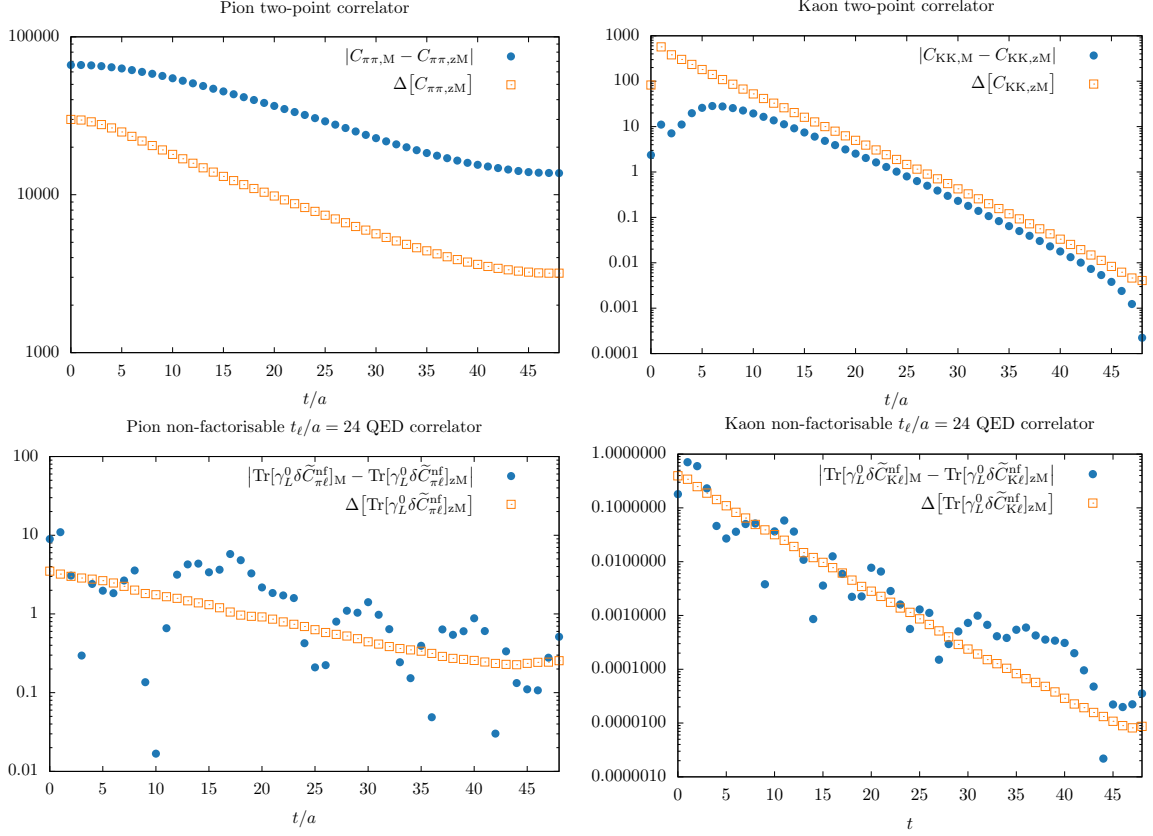
To correct for the bias introduced by the zMöbius approximation, we perform an all-mode averaging (AMA) [72] correction step. Within AMA, for each observable  $O$  we compute the estimator  $\langle \tilde{O} \rangle_{\text{M}}$  using the Möbius action from two source times ( $t_{\text{src}} = 0, T/2$ ). On the same source times, we compute the cheaper estimator  $\langle \tilde{O} \rangle_{\text{zM}}$  using the zMöbius action. Finally, we compute another zMöbius estimator  $\langle O \rangle_{\text{zM}}$  from all 96 source times available on the ensemble used in this work. The final bias-corrected estimator is then given by

$$\langle O \rangle = \langle O \rangle_{\text{zM}} + \langle \tilde{O} \rangle_{\text{M}} - \langle \tilde{O} \rangle_{\text{zM}}. \tag{B.1}$$

A comparison of the magnitude of bias correction  $\langle \tilde{O} \rangle_{\text{M}} - \langle \tilde{O} \rangle_{\text{zM}}$  to the statistical error of the estimator  $\langle O \rangle_{\text{zM}}$  is shown in figure 12. We find that the correction is negligible on most observables, with the exception of the non-factorisable correlation functions and the pion two-point correlation function.

## C Free domain wall fermion propagators

In this appendix we discuss the free domain wall fermion propagators, which have been used for the implementation of the lepton in the non-factorizable correlation functions. Throughout this section all quantities are expressed in lattice units.



**Figure 12.** Comparison of the magnitude of the zMöbius-to-Möbius bias correction (blue circles) to the statistical error of the zMöbius estimators (orange squares) for the  $\pi$  and  $K$  meson. All plots show the correlations functions, in lattice units, computed using point sources and wall sinks. In the pion two-point correlation function (top left) the correction is larger than the noise, while for the kaon two-point function (top right) the correction is smaller than our statistical precision. In the non-factorisable correlation functions (bottom plots) signal and correction are of compatible magnitude.

In this work we use the following convention for the five-dimensional Shamir-Domain-Wall-Fermion-Dirac operator [42, 43]

$$\begin{aligned}
 D_{s,t}(x,y) = & -\delta_{s,t} \frac{1}{2} \sum_{\mu} \left[ (1 - \gamma_{\mu}) U_{\mu} \delta_{y,x+\mu} + (1 + \gamma_{\mu}) U_{\mu}^{\dagger} \delta_{y,x-\mu} \right] - \delta_{s,t} (M_5 - 1 - 4) \delta_{x,y} \\
 & - \delta_{t,s+1} P_- \delta_{x,y} - \delta_{t,s-1} P_+ \delta_{x,y} + m \delta_{x,y} \delta_{s,L_s} \delta_{t,1} P_- + m \delta_{x,y} \delta_{s,1} \delta_{t,L_s} P_+ \quad (C.1)
 \end{aligned}$$

where  $s, t \in \{1, L_s\}$  label the slices in the fifth dimension and  $M_5$  is the Domain Wall height. The Domain Wall Fermion action is given by

$$S[\Psi, \bar{\Psi}, U] = - \sum_{x,y} \sum_{s,t=1}^{L_s} \bar{\Psi}_s(x) D_{s,t}(x,y) \Psi_t(y). \quad (C.2)$$

The physical quark fields are given by

$$q(x) = P_- \Psi_1(x) + P_+ \Psi_{L_s}(x) \quad \text{and} \quad \bar{q}(x) = \bar{\Psi}_1(x) P_+ + \bar{\Psi}_{L_s}(x) P_-, \quad (C.3)$$

with  $P_{\pm} = (1 \pm \gamma_5)/2$ .

### C.1 Feynman rules for free propagator

A derivation of the free Domain-Wall-Fermion propagator in momentum space can be found in ref. [48]. However, the conventions used for the five dimensional Dirac operator in ref. [48] differ from the ones given in eq. (C.1) and thus, in the following, we give results for the convention used in our work. These can be obtained by following the same steps as the derivation in ref. [48].

The free momentum-space action is given by

$$\begin{aligned} \tilde{D}_{st}(p) = & i \sum_{\mu} \gamma_{\mu} \sin p_{\mu} \delta_{st} + (W(p) \delta_{s,t} - \delta_{s-1,t} + m \delta_{s,1} \delta_{t,L_s}) P_+ \\ & + (W(p) \delta_{s,t} - \delta_{s+1,t} + m \delta_{L_s,1} \delta_{t,1}) P_-, \end{aligned} \quad (\text{C.4})$$

with

$$W(p) = 1 - M_5 + 2 \sum_{\mu} \sin^2 \frac{p_{\mu}}{2}. \quad (\text{C.5})$$

The inverse of the operator in eq. (C.4) (i.e. the propagator) can be written as

$$S_{s,t}(p) = - \left( i \sum_{\mu} \gamma_{\mu} \sin p_{\mu} \delta_{su} + (W_m^+)_{su} \right) G_{u,t}^R P_+ - \left( i \sum_{\mu} \gamma_{\mu} \sin p_{\mu} \delta_{su} + (W_m^-)_{su} \right) G_{u,t}^L P_-, \quad (\text{C.6})$$

where we use the notation

$$(W_m^+)_{st} = -W(p) \delta_{st} + \delta_{s+1,t} - m \delta_{s,L_s} \delta_{t,1}, \quad (\text{C.7})$$

$$(W_m^-)_{st} = -W(p) \delta_{st} + \delta_{s-1,t} - m \delta_{s,1} \delta_{t,L_s}, \quad (\text{C.8})$$

and define

$$G_{s,t}^R = \left( \sum_{\mu} \sin^2 p_{\mu} + W_m^- W_m^+ \right)_{s,t}^{-1}, \quad G_{s,t}^L = \left( \sum_{\mu} \sin^2 p_{\mu} + W_m^+ W_m^- \right)_{s,t}^{-1}. \quad (\text{C.9})$$

Following the steps in ref. [48] for the conventions used in this work, the inverses in eq. (C.9) can be calculated and are given by

$$G_{s,t}^R = G(s,t) + A_{++} e^{\alpha(s+t)} + A_{+-} e^{\alpha(s-t)} + A_{-+} e^{\alpha(-s+t)} + A_{--} e^{\alpha(-s-t)}, \quad (\text{C.10})$$

$$G_{s,t}^L = G(s,t) + B_{++} e^{\alpha(s+t)} + B_{+-} e^{\alpha(s-t)} + B_{-+} e^{\alpha(-s+t)} + B_{--} e^{\alpha(-s-t)}, \quad (\text{C.11})$$

where

$$G(s,t) = A \left( e^{\alpha(L_s - |s-t|)} + e^{-\alpha(L_s - |s-t|)} \right), \quad (\text{C.12})$$

with

$$A = \frac{1}{2|W| \sinh \alpha} \cdot \frac{1}{2 \sinh(\alpha L_s)}, \quad (\text{C.13})$$

and  $\alpha$  can be defined via

$$\cosh \alpha = \frac{1 + W^2 + \sum_{\mu} \sin^2 p_{\mu}}{2|W|}. \quad (\text{C.14})$$

The coefficients  $A_{\pm\pm}$  and  $B_{\pm\pm}$  are determined such that the boundary conditions ( $s = 1, L_s$ ) in  $D_{s,t}S_{t,u} = \delta_{s,u}$  are fulfilled

$$\begin{aligned}
A_{++} &= \frac{A}{F} \left( e^{-2\alpha L_s} - 1 \right) e^{-\alpha} (e^{-\alpha} - |W|) (1 - m^2), \\
A_{--} &= \frac{A}{F} \left( 1 - e^{2\alpha L_s} \right) e^{\alpha} (e^{\alpha} - |W|) (1 - m^2), \\
B_{++} &= \frac{A}{F} \left( e^{-2\alpha L_s} - 1 \right) (1 - e^{-\alpha} |W|) (1 - m^2), \\
B_{--} &= \frac{A}{F} \left( 1 - e^{2\alpha L_s} \right) (1 - e^{\alpha} |W|) (1 - m^2), \\
A_{+-} &= A_{-+} = B_{+-} = B_{-+} = \frac{A}{F} 2|W| \sinh(\alpha) \left( 1 + 2m \cosh(\alpha L_s) + m^2 \right),
\end{aligned} \tag{C.15}$$

with

$$\begin{aligned}
F &= e^{\alpha L_s} \left[ 1 - |W|e^{\alpha} + m^2(|W|e^{-\alpha} - 1) \right] + e^{-\alpha L_s} \left[ |W|e^{-\alpha} - 1 + m^2(1 - |W|e^{\alpha}) \right] \\
&\quad - 4|W|m \sinh(\alpha).
\end{aligned} \tag{C.16}$$

**NB:** the five dimensional free propagator in eq. (C.6) can be projected to four dimensions by

$$S^{4D}(p) = P_- S_{1,1}(p) P_+ + P_+ S_{L_s, L_s}(p) P_- + P_- S_{1, L_s}(p) P_- + P_+ S_{L_s, 1}(p) P_+. \tag{C.17}$$

It can be shown that in the infinite  $L_s$  limit, the four dimensional propagator is given by

$$S^{4D} \longrightarrow \frac{-i \sum_{\mu} \gamma_{\mu} \sin p_{\mu} + m(1 - W e^{-\alpha})}{-(1 - |W|e^{\alpha}) - m^2(|W|e^{-\alpha} - 1)} \quad \text{for} \quad L_s \rightarrow \infty \tag{C.18}$$

in agreement with the expression given in ref. [73].

## C.2 Pole mass of the free propagator

For the calculation of the QED correction from the factorisable diagram, we want to fix the free lepton propagator to the physical muon mass as its pole mass. In the following we describe how to determine the correct input-mass parameter  $m$  for the free propagator to reproduce a desired pole mass. The 4D propagator can be written with a common denominator  $A/F$  (see eqs. (C.13) and (C.16))

$$\begin{aligned}
\left( \frac{A}{F} \right)^{-1} &= 2|W| \sinh(\alpha) 2 \sinh(\alpha L_s) \left\{ e^{\alpha L_s} \left[ 1 - |W|e^{\alpha} + m^2(|W|e^{-\alpha} - 1) \right] \right. \\
&\quad \left. - 4|W|m \sinh(\alpha) + e^{-\alpha L_s} \left[ |W|e^{-\alpha} - 1 + m^2(1 - |W|e^{\alpha}) \right] \right\}.
\end{aligned} \tag{C.19}$$

We now have to find  $m_{\text{pole}}$  where  $(A/F)^{-1}|_{p^2 = -m_{\text{pole}}^2} = 0$ .  $(A/F)^{-1}$  has some trivial zeros, where  $\sinh(\alpha) = 0$ , which we are not interested in. We are interested in the case  $F = 0$ , i.e.

$$F|_{p^2 = -m_{\text{pole}}^2} \equiv F_{\text{pole}} = 0. \tag{C.20}$$

In practice, we want to choose a desired pole mass  $m_{\text{pole}}^2$  (e.g. the muon mass) and determine the input mass  $m$  that corresponds to this pole mass, i.e. we have to solve with respect to  $m$ :

$$F_{\text{pole}} = e^{\alpha_{\text{pole}} L_s} \left[ 1 - |W_{\text{pole}}| e^{\alpha_{\text{pole}}} + m^2 (|W_{\text{pole}}| e^{-\alpha_{\text{pole}}} - 1) \right] - 4|W_{\text{pole}}| m \sinh(\alpha_{\text{pole}}) + e^{-\alpha_{\text{pole}} L_s} \left[ |W_{\text{pole}}| e^{-\alpha_{\text{pole}}} - 1 + m^2 (1 - |W_{\text{pole}}| e^{\alpha_{\text{pole}}}) \right] = 0, \quad (\text{C.21})$$

with

$$W_{\text{pole}} = 1 - M_5 + 2 \sum_{\mu} \sin^2 \left( \frac{p_{\text{pole}}^{\mu}}{2} \right), \quad \cosh \alpha_{\text{pole}} = \frac{1 + W_{\text{pole}}^2 + \sum_{\mu} \sin^2 \left( p_{\text{pole}}^{\mu} \right)}{2|W_{\text{pole}}|}. \quad (\text{C.22})$$

This is a simple quadratic equation in  $m$  and the solutions are easily obtained from

$$m = -\frac{p}{2} \pm \sqrt{\left(\frac{p}{2}\right)^2 - q}, \quad (\text{C.23})$$

with

$$p = \frac{-4|W_{\text{pole}}| \sinh(\alpha_{\text{pole}})}{e^{\alpha_{\text{pole}} L_s} (|W_{\text{pole}}| e^{-\alpha_{\text{pole}}} - 1) + e^{-\alpha_{\text{pole}} L_s} (1 - |W_{\text{pole}}| e^{\alpha_{\text{pole}}})}, \quad (\text{C.24})$$

$$q = \frac{e^{\alpha_{\text{pole}} L_s} (1 - |W_{\text{pole}}| e^{\alpha_{\text{pole}}}) + e^{-\alpha_{\text{pole}} L_s} (|W_{\text{pole}}| e^{-\alpha_{\text{pole}}} - 1)}{e^{\alpha_{\text{pole}} L_s} (|W_{\text{pole}}| e^{-\alpha_{\text{pole}}} - 1) + e^{-\alpha_{\text{pole}} L_s} (1 - |W_{\text{pole}}| e^{\alpha_{\text{pole}}})}. \quad (\text{C.25})$$

For large  $L_s$  one finds

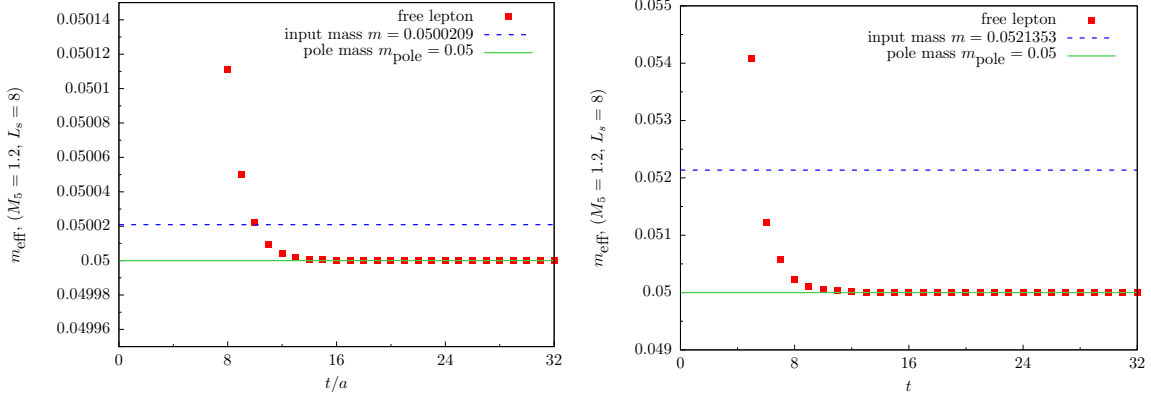
$$m \longrightarrow \pm \sqrt{-\frac{1 - |W_{\text{pole}}| e^{\alpha_{\text{pole}}}}{|W_{\text{pole}}| e^{-\alpha_{\text{pole}}} - 1}} \quad \text{for} \quad L_s \rightarrow \infty. \quad (\text{C.26})$$

In figure 13 we show the effective mass of a free propagator calculated using our implementation in Grid from Feynman rules (see appendix C.1) on a  $24^3 \times 64$  lattice. The plot on the left corresponds to a Domain Wall height of  $aM_5 = 1.0$ , while the plot on the right to  $aM_5 = 1.2$ , both with length  $L_s/a = 8$  in the fifth dimension. Red points show the numerical results for the effective mass, the solid green line shows the target pole mass of  $am_{\text{pole}} = 0.05$ , while the dashed blue line is the required input mass determined according to eqs. (C.23)–(C.25). For large-enough times  $t$  the effective mass of the free propagator plateaus at the desired target value of the pole mass  $am_{\text{pole}} = 0.05$ . The deviation from the plateau at small  $t$  is due to unphysical poles in the free Domain Wall Fermion propagator (see, e.g., the discussion in ref. [74]).

### C.3 Projectors on definite spinor structure

Euclidean free Dirac spinors satisfy the following on-shell Dirac equations

$$D(i\omega, \mathbf{p}) u(\mathbf{p}) = 0, \quad D(-i\omega, -\mathbf{p}) v(\mathbf{p}) = 0, \quad (\text{C.27})$$



**Figure 13.** The effective mass of a free propagator calculated from Feynman rules on a  $24^3 \times 64$  lattice for  $aM_5 = 1.0$  (left) and  $aM_5 = 1.2$  (right).

where  $D(p)$  is the Dirac operator in momentum space and  $\omega$  is the energy satisfying the dispersion relation  $D(i\omega, \mathbf{p})D(-i\omega, -\mathbf{p}) = 0$ . The spinors also respect the following completeness relations

$$\sum_r u^r(\mathbf{p})\bar{u}^r(\mathbf{p}) = D(-i\omega, -\mathbf{p}), \quad \sum_r v^r(\mathbf{p})\bar{v}^r(\mathbf{p}) = -D(i\omega, \mathbf{p}), \quad (\text{C.28})$$

and orthogonality relations

$$\begin{aligned} \bar{v}^r(-\mathbf{p})\Gamma_0 u^s(\mathbf{p}) &= 0, & \bar{u}^r(\mathbf{p})\Gamma_0 u^s(\mathbf{p}) &= 2\mathcal{E}\delta^{rs}, & \bar{u}^r(\mathbf{p})u^s(\mathbf{p}) &= 2\mathcal{M}\delta^{rs}, \\ \bar{u}^r(-\mathbf{p})\Gamma_0 v^s(\mathbf{p}) &= 0, & \bar{v}^r(\mathbf{p})\Gamma_0 v^s(\mathbf{p}) &= 2\mathcal{E}\delta^{rs}, & \bar{v}^r(\mathbf{p})v^s(\mathbf{p}) &= -2\mathcal{M}\delta^{rs}, \end{aligned} \quad (\text{C.29})$$

where  $\mathcal{E}$  and  $\mathcal{M}$  are quantities that in the continuum limit reduce to  $\lim_{a \rightarrow 0} \mathcal{E} = \sqrt{m^2 + |\mathbf{p}|^2}$  and  $\lim_{a \rightarrow 0} \mathcal{M} = m$ , respectively.

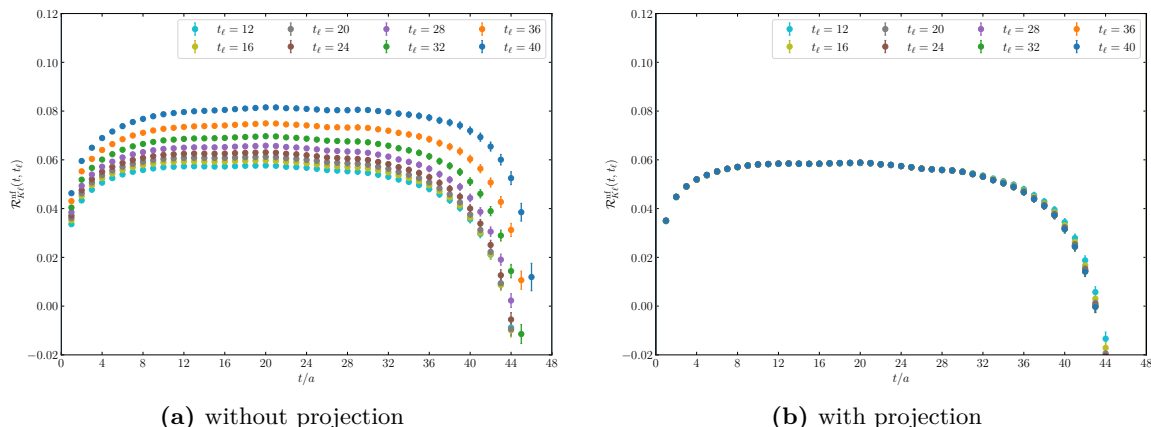
As discussed in section 4.3, the external anti-lepton propagator projected on momentum  $\mathbf{p}_\ell$ , when evaluated on the lattice with finite time  $T$  and anti-periodic boundary conditions takes the following form (neglecting possible contact terms)

$$\begin{aligned} S_\ell(0|t_\ell, \mathbf{p}_\ell) &= \left\{ e^{-\omega_\ell t_\ell} \frac{D(i\omega_\ell, \mathbf{p}_\ell)}{2\Omega_\ell} - e^{-\omega_\ell(T-t_\ell)} \frac{D(-i\omega_\ell, \mathbf{p}_\ell)}{2\Omega_\ell} \right\} \times \frac{1}{1 + e^{-\omega_\ell T}} \\ &= - \sum_r \left\{ e^{-\omega_\ell t_\ell} \frac{v_r(\mathbf{p})\bar{v}_r(\mathbf{p})}{2\Omega_\ell} + e^{-\omega_\ell(T-t_\ell)} \frac{u_r(-\mathbf{p})\bar{u}_r(-\mathbf{p})}{2\Omega_\ell} \right\} \times \frac{1}{1 + e^{-\omega_\ell T}}, \end{aligned} \quad (\text{C.30})$$

where we observe that the backward signal has a different Dirac structure compared to the forward one. Here  $\Omega_\ell$  is a quantity that in the continuum limit gives  $\lim_{a \rightarrow 0} \Omega_\ell = \sqrt{m_\ell^2 + |\mathbf{p}_\ell|^2}$ .

By using the orthogonality relations above and the fact that

$$D(-i\omega, \mathbf{p}) - D(i\omega, \mathbf{p}) = 2\mathcal{E}\Gamma_0, \quad (\text{C.31})$$



**Figure 14.** Comparison of the pion non-factorizable correlator  $\mathcal{R}_{\pi\ell}^{\text{nf}}$  obtained without the use of the projector  $\mathcal{P}_{v(\mathbf{p}_\ell)}$  (a) and with the backward-propagating signal removed (b).

we can define two projectors

$$\begin{aligned}
 \mathcal{P}_{v(\mathbf{p})} &= \{D(-iE, \mathbf{p}) - D(iE, \mathbf{p})\}^{-1} [-D(iE, \mathbf{p})] \\
 &= \{u_t(-\mathbf{p})\bar{u}_t(-\mathbf{p}) + v_s(\mathbf{p})\bar{v}_s(\mathbf{p})\}^{-1} [v_r(\mathbf{p})\bar{v}_r(\mathbf{p})] , \\
 \mathcal{P}_{u(-\mathbf{p})} &= \{D(-iE, \mathbf{p}) - D(iE, \mathbf{p})\}^{-1} D(-iE, \mathbf{p}) \\
 &= \{u_t(-\mathbf{p})\bar{u}_t(-\mathbf{p}) + v_s(\mathbf{p})\bar{v}_s(\mathbf{p})\}^{-1} [u_r(-\mathbf{p})\bar{u}_r(-\mathbf{p})] ,
 \end{aligned}
 \tag{C.32}$$

such that for the lepton propagator we have

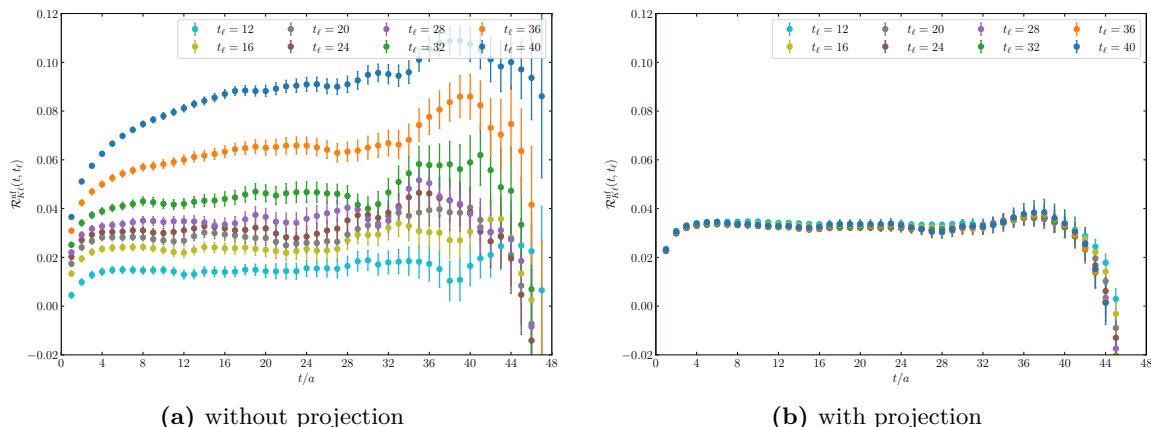
$$\begin{aligned}
 S_\ell(0|t_\ell, \mathbf{p}_\ell) \cdot \mathcal{P}_{v(\mathbf{p}_\ell)} &= - \sum_r \left\{ e^{-\omega_\ell t_\ell} \frac{v_r(\mathbf{p}_\ell)\bar{v}_r(\mathbf{p}_\ell)}{2\Omega_\ell} \right\} \times \frac{1}{1 + e^{-\omega_\ell T}} , \\
 S_\ell(0|t_\ell, \mathbf{p}_\ell) \cdot \mathcal{P}_{u(-\mathbf{p}_\ell)} &= - \sum_r \left\{ e^{-\omega_\ell(T-t_\ell)} \frac{u_r(-\mathbf{p}_\ell)\bar{u}_r(-\mathbf{p}_\ell)}{2\Omega_\ell} \right\} \times \frac{1}{1 + e^{-\omega_\ell T}} .
 \end{aligned}
 \tag{C.33}$$

In order to construct the projectors  $\mathcal{P}_{v(\mathbf{p}_\ell)}$  and  $\mathcal{P}_{u(-\mathbf{p}_\ell)}$  we then compute on the lattice the free domain-wall lepton propagator  $S_\ell(t_\ell, -\mathbf{p}_\ell|0)$ , projected on definite momentum  $-\mathbf{p}_\ell$ , having the following temporal behaviour

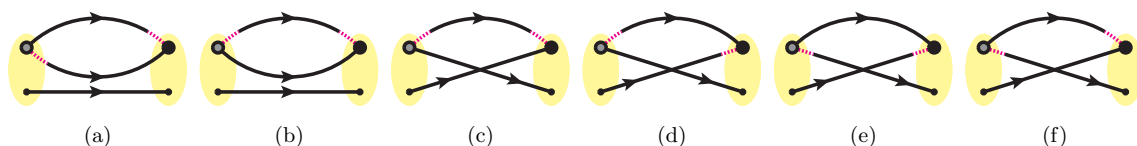
$$S_\ell(t_\ell, -\mathbf{p}_\ell|0) = \sum_r \left\{ e^{-\omega_\ell t_\ell} \frac{u_r(-\mathbf{p})\bar{u}_r(-\mathbf{p})}{2\Omega_\ell} + e^{-\omega_\ell(T-t_\ell)} \frac{v_r(\mathbf{p})\bar{v}_r(\mathbf{p})}{2\Omega_\ell} \right\} \times \frac{1}{1 + e^{-\omega_\ell T}} . \tag{C.34}$$

Since the definitions of the projectors in eq. (C.32) do not depend on the spinor normalization  $2\Omega_\ell(1 + e^{-\omega_\ell T})$ , because it cancels out in the matrix multiplications, they can be easily obtained from the free lattice lepton propagator just by extracting and combining the coefficients of the forward and backward exponentials.

The effect of using the projector  $\mathcal{P}_{v(\mathbf{p}_\ell)}$  on the non-factorizable correlator  $\mathcal{R}_{P_\ell}^{\text{nf}}$  defined in eq. (4.36) is shown in figures 14 and 15 for the pion and kaon decay, respectively (computed with zMöbius fermions). We note that the backward signal is drastically suppressed and the dependence on the lepton source-sink separation  $t_\ell$  is barely visible for  $t \ll T/2$ . The use of these projectors makes then a crucial difference in the extraction of a clear signal from the lattice data.



**Figure 15.** Comparison of the kaon non-factorizable correlator  $\mathcal{R}_{K\ell}^{\text{nf}}$  obtained without the use of the projector  $\mathcal{P}_{v(\mathbf{p}_\ell)}$  (a) and with the backward propagating signal removed (b).



**Figure 16.** All Feynman diagrams corresponding to the tree-level correlation function  $\tilde{C}_{\Omega\Omega}(t)$ . Points connecting two propagators are contractions of a diquark pair, and dashed magenta portions of a propagator indicate contraction with a transposed quark field.

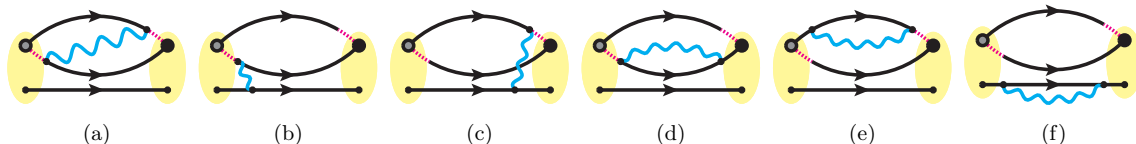
### D Correlation functions for the $\Omega$ baryon

In this appendix we discuss the construction of the  $\Omega^-$  baryon correlation functions used in this work, as well as their spectral representation. We begin by considering the Wick contractions for the tree-level iso-QCD correlator given in eq. (4.38), of which there are 6 contributions. These are shown diagrammatically in figure 16, where the points connecting two propagators are contractions of a diquark pair in eq. (4.39), and dashed magenta portions of a propagator indicate contraction with a transposed quark field. The colour structure of these contractions is not represented in these diagrams.

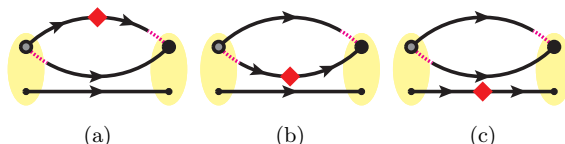
For the QED corrections to this correlator, we require two insertions of the quark-photon interaction  $iZ_V \sum_x \bar{s} \mathcal{A} s$  (see eq. (4.15)) which corresponds to a photon propagator connecting the quark legs, as well as a quark-disconnected contribution that is omitted in this work. Taking for example diagram (a) in figure 16, the corresponding QED corrections are shown in figure 17 where (a), (b) and (c) are the exchange diagrams contributing to  $\delta\tilde{C}_{\Omega\Omega}^{\text{exch}}(t)$  and (d), (e) and (f) are the self energy diagrams contributing to  $\delta\tilde{C}_{\Omega\Omega}^{\text{self},s}(t)$  in eq. (4.42). Similarly, quark-mass corrections are given by the insertion of the scalar density  $\sum_x \bar{s} s$  (see eq. (4.15)). Again taking diagram (a) in figure 16 as an example, the mass corrections are given by the diagrams in figure 18, as well as a disconnected contribution that is also omitted.

Once the  $\Omega^-$  correlators have been constructed, their spectral representation must be evaluated. First note that the interpolator in eq. (4.39) contains a parity projector  $P_+$





**Figure 17.** All (connected) Feynman diagrams contributing to  $\mathcal{R}_{\Omega\Omega}^{e,m.}(t)$  originating from the tree-level contribution shown in figure 16 (a). Similar diagrams exist for the other contributions.



**Figure 18.** All (connected) Feynman diagrams contributing to  $\mathcal{R}_{\Omega\Omega}^{S,s}(t)$  originating from the tree-level contribution shown in figure 16 (a). Similar diagrams exist for the other contributions.

which causes  $\bar{\psi}_{\Omega}^{\mu}$  to create states of positive parity, but also annihilate states of negative parity. Therefore, the ground state spectral representation will have the form

$$C_{\Omega\Omega}(t) = \frac{a^3}{2} \sum_i \sum_s \left[ \frac{\langle 0 | \psi_{\Omega}^i(0) | \Omega, s \rangle \langle \Omega, s | \bar{\psi}_{\Omega}^i(0) | 0 \rangle}{2m_{\Omega}} e^{-m_{\Omega}t} - \frac{\langle \bar{\Omega}, s | \psi_{\Omega}^i(0) | 0 \rangle \langle 0 | \bar{\psi}_{\Omega}^i(0) | \bar{\Omega}, s \rangle}{2\bar{m}_{\Omega}} e^{-\bar{m}_{\Omega}(T-t)} \right], \quad (D.1)$$

where the relative sign change between the forward and backward-propagating components comes from assuming anti-periodic boundary conditions in time on the quarks, and therefore also on the baryon fields. We have additionally distinguished the notation of the rest energy of the negative parity state  $\bar{m}_{\Omega}$  from the positive parity one  $m_{\Omega}$  due to the fact that, at the physical point, the negative parity  $\Omega^{-}$  baryon is not simply a single state in the QCD Fock space, but is instead a resonance in the  $\Xi K$  channel and therefore there is a whole spectrum of finite volume multi-particle states contributing in the backward time direction. However, this does not complicate our analysis since we are restricted to early times where the backward propagating contributions are negligible.

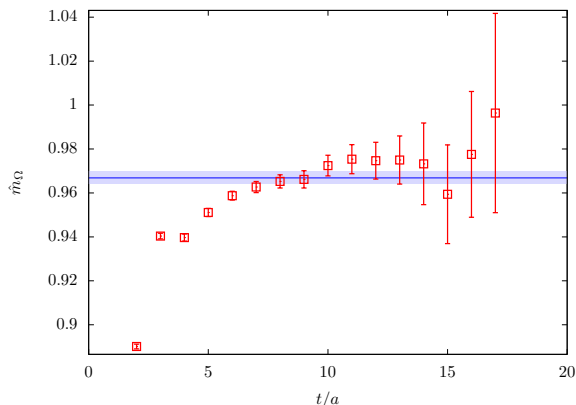
The operator-state overlaps have the form

$$\langle 0 | \psi_{\Omega}^{\mu}(0) | \Omega, s \rangle = Z_{\Omega} u_s^{\mu}, \quad \langle \Omega, s | \bar{\psi}_{\Omega}^{\mu}(0) | 0 \rangle = Z_{\Omega}^* \bar{u}_s^{\mu}, \quad (D.2)$$

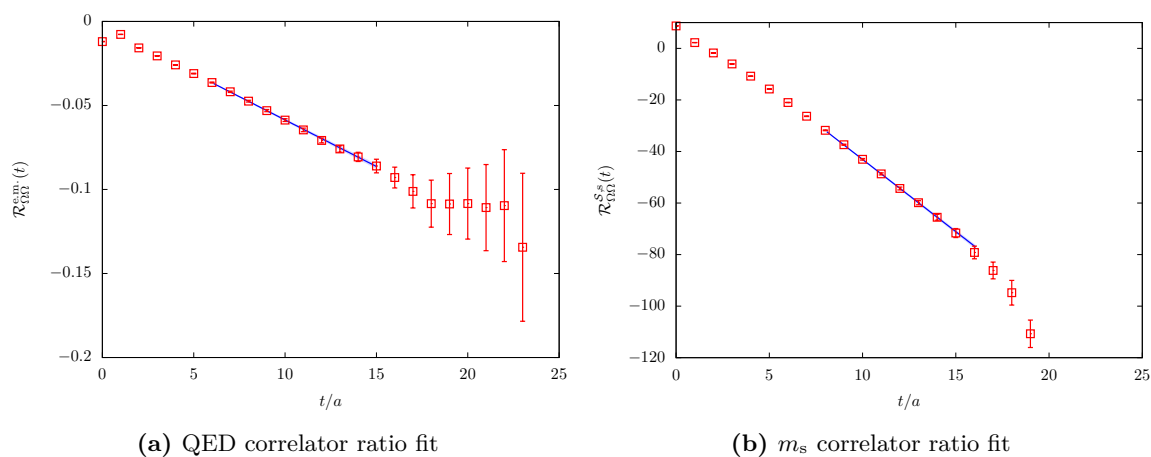
$$\langle \bar{\Omega}, s | \psi_{\Omega}^{\mu}(0) | 0 \rangle = \bar{Z}_{\Omega} \gamma_5 v_s^{\mu}, \quad \langle 0 | \bar{\psi}_{\Omega}^{\mu}(0) | \bar{\Omega}, s \rangle = \bar{Z}_{\Omega}^* \bar{v}_s^{\mu} \gamma_5, \quad (D.3)$$

where  $u_s^{\mu}$  and  $v_s^{\mu}$  are the positive and negative energy solutions to the spin-3/2 Rarita-Schwinger equation respectively (see e.g. [54] for a recent review), and the  $\gamma_5$  is present in the negative parity matrix elements to obtain the correct transformation properties. Using Euclidean conventions, the completeness relations for zero momentum spinors with mass  $m$  are given by

$$\sum_s u_s^i \bar{u}_s^j = 2m P_+ \left( \delta^{ij} - \frac{1}{3} \gamma^i \gamma^j \right), \quad \sum_s v_s^i \bar{v}_s^j = -2m P_- \left( \delta^{ij} - \frac{1}{3} \gamma^i \gamma^j \right), \quad (D.4)$$



**Figure 19.** Log effective mass of the tree-level  $\Omega^-$  baryon (in lattice units) in red and the fit result of the mass parameter in blue.



**Figure 20.** Ratios of  $\Omega^-$  baryon QED corrections (a) and quark mass corrections (b) to the tree-level correlator,  $\mathcal{R}_{\Omega\Omega}^{e,m}(t)$  and  $\mathcal{R}_{\Omega\Omega}^{S,s}(t)$  respectively, in red and the fit to the data in blue.

which result in the form of the correlator given in eq. (4.40). The spectral representation of the QED and quark mass corrections to this tree-level correlator are simply found by expanding the two parameters  $Z_\Omega$  and  $m_\Omega$  to first order in the respective isospin breaking parameter.

Figure 19 shows the log effective mass of the tree-level correlator  $\tilde{C}_{\Omega\Omega}(t)$  along with the fit result of the iso-QCD mass. It should be noted that the fit was performed to the correlator and not directly to the effective mass. The value obtained for the mass of the  $\Omega^-$  baryon,  $\hat{m}_\Omega = 0.967(3)$  for the best fit shown in figure 19, is in agreement with that obtained in ref. [39] using the same gauge ensemble. Figure 20 shows the ratio of the QED and  $m_s$  corrections of the  $\Omega^-$  correlator to the tree-level result,  $\mathcal{R}_{\Omega\Omega}^{e,m}(t)$  and  $\mathcal{R}_{\Omega\Omega}^{S,s}(t)$  respectively. Included is the fit to this ratio using the linear fit model given in eq. (4.43).

## E Determining best fits with a genetic algorithm

In this appendix, we discuss in detail the setup of the genetic algorithm (GA) used in the data analysis described in section 5.1 to select best correlator fits. After summarising the purpose of a GA, we describe the genetic operators used, and we present the GA hyperparameters used to produce the factorisable analysis fit results in section 5.

GAs form a class of global optimizers which stochastically evolve a set of candidate solutions toward ones which maximize/minimize a given objective function. The evolution process is inspired by natural selection in biological systems by proposing new solutions attempting to combine best features from a previous generation of solutions. Because of this analogy, the set of candidate solutions is generally referred as the *population*, and the state of the population at a given iteration of the algorithm is called a *generation*. The step between one generation to the next is done via genetics-inspired operators called *crossover* and *mutation* operators. The crossover operator aims at producing a better solution to the optimisation problem by combining two members of the current generation. The mutation operator make arbitrary random changes to members of the current generation, increasing the space of solutions explored by the algorithm.

In the context of the factorisable analyses, a population member is a vector  $\boldsymbol{\tau} = (\tau^{(1)}, \dots, \tau^{(n_{\text{corr}})})$  where each component is a fit interval  $\tau^{(j)} = [t_{\text{min}}^{(j)}, t_{\text{max}}^{(j)}]$  for the  $j$ -th correlator, and  $n_{\text{corr}}$  is the total number of correlators to fit. The objective function here is the AIC weight introduced in section 5.1, which we aim at maximising. The space of all possible fits is finite, although it contains a very large number of elements. However, this finiteness guarantees that at least one solution to the optimisation problem exists. A high-level description of the algorithm is as follows:

1. in each generation, begin with an initial population of  $\{\boldsymbol{\tau}_k\}$  with  $P_0$  elements;
2. evolve  $\{\boldsymbol{\tau}_k\}$  with genetic operators to produce  $n_{\text{off}}$  new elements (called *offspring*);
3. compute the AIC for all population members through  $\chi^2$  minimization as described in section 5.1;
4. choose among  $\{\boldsymbol{\tau}_k\}$  the best  $P_0$  fit ranges with the largest AIC weights and discard all other elements;
5. repeat steps 2-4 until a termination condition is satisfied.

Let us now introduce the genetic operators. Consider two candidate fit intervals in the initial population,  $\boldsymbol{\tau}_k$  and  $\boldsymbol{\tau}_{k'}$ . The crossover operator,  $X$ , generates a new  $\boldsymbol{\tau}$  with fit intervals from either of the parent members based on random numbers  $0 \leq p_j \leq 1$  for  $1 \leq j \leq n_{\text{corr}}$ . That is,

$$\boldsymbol{\tau}_{k''} \equiv X(\boldsymbol{\tau}_k, \boldsymbol{\tau}_{k'}) = \left\{ X(\tau_k^{(1)}, \tau_{k'}^{(1)}), \dots, X(\tau_k^{(n_{\text{corr}})}, \tau_{k'}^{(n_{\text{corr}})}) \right\}, \quad (\text{E.1})$$

where

$$X(\tau_k^{(j)}, \tau_{k'}^{(j)}) = \begin{cases} \tau_k^{(j)} & \text{if } p_j < 0.5, \\ \tau_{k'}^{(j)} & \text{otherwise.} \end{cases} \quad (\text{E.2})$$

This is repeated until one obtains a population size of  $P > P_0$ . The mutation operator,  $M$ , then mutates the population at a given rate  $m$ . That is, for some randomly drawn value of  $p$  ( $0 \leq p \leq 1$ ), one has

$$M(\tau_k) = \begin{cases} \tau_{P+1}, & \text{if } p < m \\ \tau_k & \text{otherwise,} \end{cases} \quad (\text{E.3})$$

with

$$\tau_{P+1} \equiv \left\{ \tau_{P+1}^{(1)}, \dots, M(\tau_{P+1}^{(j)}), \dots, \tau_{P+1}^{(n_{\text{corr}})} \right\} \quad (\text{E.4})$$

where the index  $j$  ( $1 \leq j \leq n_{\text{corr}}$ ), is also randomly drawn. The mutation  $M(\tau_{P+1}^{(j)})$  is a fit interval where either  $t_{\text{min}}$  or  $t_{\text{max}}$  or both have been modified randomly.

With the operators defined, we can discuss the GA parameters used for this work. To begin, the free parameters in a GA are: the size of the initial population,  $P_0$ ; the crossover rate, which is parametrised in this work by the population size after crossover  $P$ ; the mutation rate  $m$ ; the weight function to optimize  $w$ ; the maximum number of generations  $G_{\text{max}}$  and the termination condition. We studied three GA setups to check the validity of our fit conclusions. These are summarised in table 3. First, let

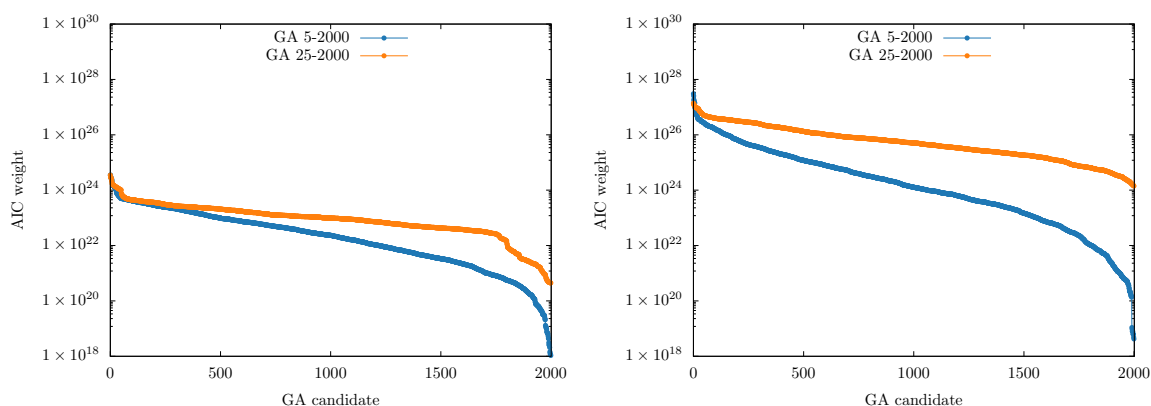
$$\bar{w}_N = \frac{1}{N} \sum_{i=1}^N w_i \quad (\text{E.5})$$

be the average of the top  $N$  weights in each generation. In all three setups, we aim to maximize the AIC in each GA run, which terminates when the average of the  $N = 5$  top fit,  $\bar{w}_5$ , does not improve over 1000 successive generations. If this condition cannot be satisfied, we impose a cut-off when a GA run exceeds  $G_{\text{max}}$  generations. In practice, however, none of the runs hit this cut-off limit. To accelerate the GA in its exploration of the  $\tau$ -space, we cache all fit results during the process.

Let the label ‘GA  $X$ - $Y$ ’ refer to a GA setup with a population size  $P_0 = X$  which has been run multiple times until obtaining a total of  $Y$  candidates. Across the different runs are varied the initial condition of the algorithm and the random number sequence used in the genetic operators. We consider 3 different setups GA 5-2000, GA 25-2000, and GA 25-5000, summarised in table 3. Additionally, figure 21 compare the AICs of the 2000 outcomes, sorted by their weights in descending order. It is worth commenting on two features. First, the best AIC fits in both setups are very similar, demonstrating some level of independence between the optimal solution found and the hyperparameters of the GA. Beyond that, it is clear that the range of AIC weights is narrower in GA 25-2000 than GA 5-2000 for both pion and kaon. This suggests that a population size of  $P_0 = 25$  allows one to discover more optimal solutions than  $P_0 = 5$  for a given target number of candidates. This is expected as the algorithm will try more candidates at each generation. However, this also suggests that the  $P_0 = 5$  set of runs is not saturating its exploration of the best AIC fit space. To address this, the GA 25-5000 setup was designed to check that such saturation was achieved for GA 25-2000. In table 4, we give the average number of distinct fits explored by each setup for each analysis. As we can see, considerably increasing the number of GA runs in the  $P_0 = 25$  case does not lead to a significant volume of new fits

	GA 5-2000	GA 25-2000	GA 25-5000
number of runs	400	80	200
$P_0$	5	25	25
$P$	20	100	100
$m$	0.3	0.3	0.3
$G_{\max}$	25000	10000	10000
termination cond. n. 1	$\bar{w}_5$ unchanged for 1000 generations		
termination cond. n. 2	GA exceeds $G_{\max}$		
total GA candidates	2000	2000	5000

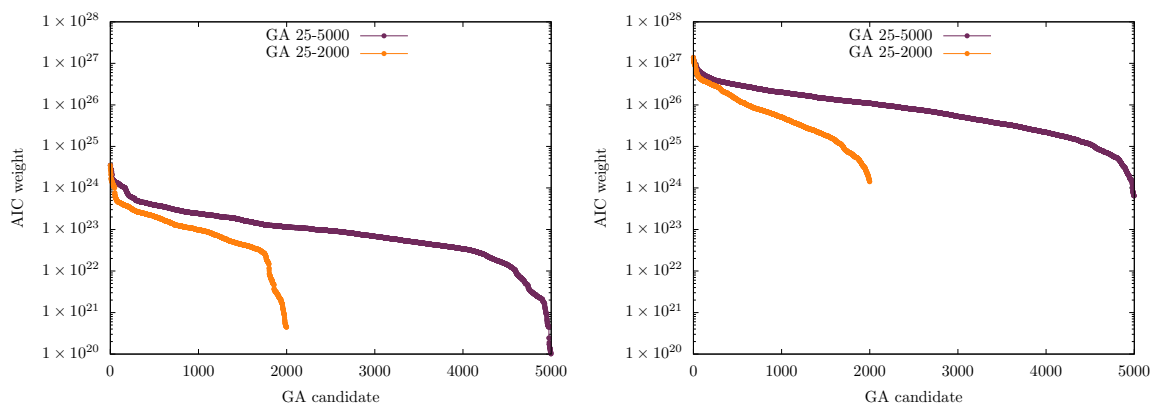
**Table 3.** Table of three different GA setups used in this work. All setups maximize the AIC weight.  $\bar{w}_5$  is defined in the text below eq. (E.5).



**Figure 21.** Comparison of the 2000 GA candidates between two setups GA 5-2000 and GA 25-2000, sorted in descending order of AIC weights, for  $\pi$  (left) and  $K$  (right) meson correlator analysis.

tried, meaning that the additional runs were to a large extent redundant in terms of optimal solutions found.

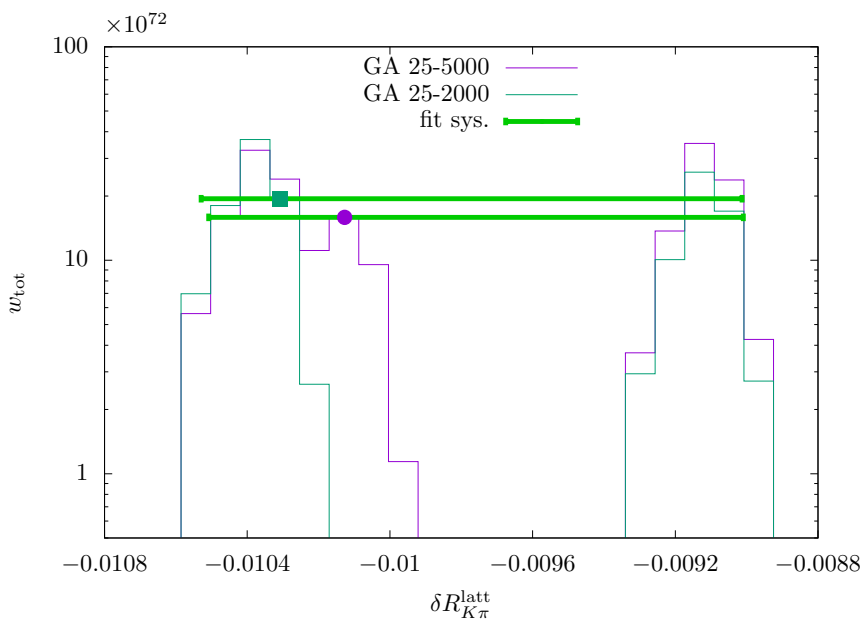
Ultimately, as it is generally the case with GAs in this type of context, it is not possible to demonstrate with absolute certainty that the GA found the best fits without knowing the exact solution to the problem. However, we remind the reader that the main aim here is to establish a representative spread in our final result for  $\delta R_{K\pi}$  in order to assign a systematic error related to the selection of fit ranges. To check the stability of our systematic error under variations of the GA setup, we take the top 5 fits of both GA 25-2000 and GA 25-5000 setups and generate an AIC-weighted histogram of  $\delta R_{K\pi}^{\text{latt}}$  for each setup. These are shown in figure 23, along with their median and the fit systematic errors as defined in section 5.3. We see that they lead to very similar conclusions in terms of median and systematic spread, and we consider that as a compelling evidence that the GA is converging on a set of fit candidates which is representative enough to estimate the fit range selection systematic uncertainty. Finally, the GA 25-5000 setup was used to produce the final result of this work.



**Figure 22.** Comparison of the GA candidates between setups GA 25-5000 and GA 25-2000, sorted in descending order of AIC weights, for the  $\pi$  (left) and  $K$  (right) meson correlator analysis.

Analysis	GA 5-2000	GA 25-2000	GA 25-5000
$\pi$	4085.915	32854.5375	33767.365
$K$	4726.1425	34430.625	33975.085

**Table 4.** The average total number of fits tried per GA setup for each meson correlator analysis.



**Figure 23.** AIC-weighted histogram generated from top 5 fits of each analysis as described in section 5.3. The median of each histogram and the fit systematics are superimposed on the histogram.

**Open Access.** This article is distributed under the terms of the Creative Commons Attribution License ([CC-BY 4.0](https://creativecommons.org/licenses/by/4.0/)), which permits any use, distribution and reproduction in any medium, provided the original author(s) and source are credited. SCOAP<sup>3</sup> supports the goals of the International Year of Basic Sciences for Sustainable Development.

## References

- [1] PARTICLE DATA GROUP collaboration, *Review of Particle Physics*, *PTEP* **2022** (2022) 083C01 [[INSPIRE](#)].
- [2] FLAVOUR LATTICE AVERAGING GROUP (FLAG) collaboration, *FLAG Review 2021*, *Eur. Phys. J. C* **82** (2022) 869 [[arXiv:2111.09849](#)] [[INSPIRE](#)].
- [3] V. Cirigliano, A. Crivellin, M. Hoferichter and M. Moulson, *Scrutinizing CKM unitarity with a new measurement of the  $K\mu 3/K\mu 2$  branching fraction*, *Phys. Lett. B* **838** (2023) 137748 [[arXiv:2208.11707](#)] [[INSPIRE](#)].
- [4] B. Ananthanarayan and B. Moussallam, *Four-point correlator constraints on electromagnetic chiral parameters and resonance effective Lagrangians*, *JHEP* **06** (2004) 047 [[hep-ph/0405206](#)] [[INSPIRE](#)].
- [5] S. Descotes-Genon and B. Moussallam, *Radiative corrections in weak semi-leptonic processes at low energy: A Two-step matching determination*, *Eur. Phys. J. C* **42** (2005) 403 [[hep-ph/0505077](#)] [[INSPIRE](#)].
- [6] V. Cirigliano and I. Rosell,  *$\pi/K \rightarrow e\bar{\nu}_e$  branching ratios to  $O(e^2p^4)$  in Chiral Perturbation Theory*, *JHEP* **10** (2007) 005 [[arXiv:0707.4464](#)] [[INSPIRE](#)].
- [7] D. Giusti et al., *First lattice calculation of the QED corrections to leptonic decay rates*, *Phys. Rev. Lett.* **120** (2018) 072001 [[arXiv:1711.06537](#)] [[INSPIRE](#)].
- [8] M. Di Carlo et al., *Light-meson leptonic decay rates in lattice QCD+QED*, *Phys. Rev. D* **100** (2019) 034514 [[arXiv:1904.08731](#)] [[INSPIRE](#)].
- [9] F. Bloch and A. Nordsieck, *Note on the Radiation Field of the electron*, *Phys. Rev.* **52** (1937) 54 [[INSPIRE](#)].
- [10] N. Carrasco et al., *QED Corrections to Hadronic Processes in Lattice QCD*, *Phys. Rev. D* **91** (2015) 074506 [[arXiv:1502.00257](#)] [[INSPIRE](#)].
- [11] A. Desiderio et al., *First lattice calculation of radiative leptonic decay rates of pseudoscalar mesons*, *Phys. Rev. D* **103** (2021) 014502 [[arXiv:2006.05358](#)] [[INSPIRE](#)].
- [12] M. Hayakawa and S. Uno, *QED in finite volume and finite size scaling effect on electromagnetic properties of hadrons*, *Prog. Theor. Phys.* **120** (2008) 413 [[arXiv:0804.2044](#)] [[INSPIRE](#)].
- [13] V. Cirigliano and H. Neufeld, *A note on isospin violation in  $P_{\ell 2}(\gamma)$  decays*, *Phys. Lett. B* **700** (2011) 7 [[arXiv:1102.0563](#)] [[INSPIRE](#)].
- [14] M. Di Carlo, M.T. Hansen, A. Portelli and N. Hermansson-Truedsson, *Relativistic, model-independent determination of electromagnetic finite-size effects beyond the pointlike approximation*, *Phys. Rev. D* **105** (2022) 074509 [[arXiv:2109.05002](#)] [[INSPIRE](#)].
- [15] E. Tiesinga, P.J. Mohr, D.B. Newell and B.N. Taylor, *CODATA recommended values of the fundamental physical constants: 2018*, *Rev. Mod. Phys.* **93** (2021) 025010 [[INSPIRE](#)].
- [16] BUDAPEST-MARSEILLE-WUPPERTAL collaboration, *Isospin splittings in the light baryon octet from lattice QCD and QED*, *Phys. Rev. Lett.* **111** (2013) 252001 [[arXiv:1306.2287](#)] [[INSPIRE](#)].

- [17] Z. Fodor et al., *Up and down quark masses and corrections to Dashen's theorem from lattice QCD and quenched QED*, *Phys. Rev. Lett.* **117** (2016) 082001 [[arXiv:1604.07112](#)] [[INSPIRE](#)].
- [18] S. Borsanyi et al., *Leading hadronic contribution to the muon magnetic moment from lattice QCD*, *Nature* **593** (2021) 51 [[arXiv:2002.12347](#)] [[INSPIRE](#)].
- [19] J. Bijnens and N. Danielsson, *Electromagnetic Corrections in Partially Quenched Chiral Perturbation Theory*, *Phys. Rev. D* **75** (2007) 014505 [[hep-lat/0610127](#)] [[INSPIRE](#)].
- [20] J. Gasser, A. Rusetsky and I. Scimemi, *Electromagnetic corrections in hadronic processes*, *Eur. Phys. J. C* **32** (2003) 97 [[hep-ph/0305260](#)] [[INSPIRE](#)].
- [21] A. Duncan, E. Eichten and H. Thacker, *Electromagnetic splittings and light quark masses in lattice QCD*, *Phys. Rev. Lett.* **76** (1996) 3894 [[hep-lat/9602005](#)] [[INSPIRE](#)].
- [22] RM123 collaboration, *Leading isospin breaking effects on the lattice*, *Phys. Rev. D* **87** (2013) 114505 [[arXiv:1303.4896](#)] [[INSPIRE](#)].
- [23] D. Giusti et al., *Leading isospin-breaking corrections to pion, kaon and charmed-meson masses with Twisted-Mass fermions*, *Phys. Rev. D* **95** (2017) 114504 [[arXiv:1704.06561](#)] [[INSPIRE](#)].
- [24] V. Lubicz, G. Martinelli, C.T. Sachrajda, F. Sanfilippo, S. Simula and N. Tantalo, *Finite-Volume QED Corrections to Decay Amplitudes in Lattice QCD*, *Phys. Rev. D* **95** (2017) 034504 [[arXiv:1611.08497](#)] [[INSPIRE](#)].
- [25] N. Tantalo, V. Lubicz, G. Martinelli, C.T. Sachrajda, F. Sanfilippo and S. Simula, *Electromagnetic corrections to leptonic decay rates of charged pseudoscalar mesons: finite-volume effects*, [arXiv:1612.00199](#) [[INSPIRE](#)].
- [26] A. Sirlin, *Radiative Corrections in the  $SU(2)_L \times U(1)$  Theory: A Simple Renormalization Framework*, *Phys. Rev. D* **22** (1980) 971 [[INSPIRE](#)].
- [27] A. Sirlin, *Large  $m_W$ ,  $m_Z$  Behavior of the  $O(\alpha)$  Corrections to Semileptonic Processes Mediated by  $W$* , *Nucl. Phys. B* **196** (1982) 83 [[INSPIRE](#)].
- [28] S.M. Berman, *Radiative corrections to muon and neutron decay*, *Phys. Rev.* **112** (1958) 267 [[INSPIRE](#)].
- [29] T. Kinoshita and A. Sirlin, *Radiative corrections to Fermi interactions*, *Phys. Rev.* **113** (1959) 1652 [[INSPIRE](#)].
- [30] T. Harris, V. Gülpers, A. Portelli and J. Richings, *Efficiently unquenching QCD+QED at  $O(\alpha)$* , [arXiv:2301.03995](#).
- [31] Z. Davoudi and M.J. Savage, *Finite-Volume Electromagnetic Corrections to the Masses of Mesons, Baryons and Nuclei*, *Phys. Rev. D* **90** (2014) 054503 [[arXiv:1402.6741](#)] [[INSPIRE](#)].
- [32] S. Borsanyi et al., *Ab initio calculation of the neutron-proton mass difference*, *Science* **347** (2015) 1452 [[arXiv:1406.4088](#)] [[INSPIRE](#)].
- [33] Z. Davoudi, J. Harrison, A. Jüttner, A. Portelli and M.J. Savage, *Theoretical aspects of quantum electrodynamics in a finite volume with periodic boundary conditions*, *Phys. Rev. D* **99** (2019) 034510 [[arXiv:1810.05923](#)] [[INSPIRE](#)].
- [34] RBC, UKQCD collaborations, *Calculation of the hadronic vacuum polarization contribution to the muon anomalous magnetic moment*, *Phys. Rev. Lett.* **121** (2018) 022003 [[arXiv:1801.07224](#)] [[INSPIRE](#)].
- [35] M.G. Endres, A. Shindler, B.C. Tiburzi and A. Walker-Loud, *Massive photons: an infrared regularization scheme for lattice QCD+QED*, *Phys. Rev. Lett.* **117** (2016) 072002 [[arXiv:1507.08916](#)] [[INSPIRE](#)].



- [36] B. Lucini, A. Patella, A. Ramos and N. Tantalo, *Charged hadrons in local finite-volume QED+QCD with  $C^*$  boundary conditions*, *JHEP* **02** (2016) 076 [[arXiv:1509.01636](#)] [[INSPIRE](#)].
- [37] P. Boyle et al., *Isospin breaking corrections to meson masses and the hadronic vacuum polarization: a comparative study*, *JHEP* **09** (2017) 153 [[arXiv:1706.05293](#)] [[INSPIRE](#)].
- [38] R.C. Brower, H. Neff and K. Orginos, *The Möbius domain wall fermion algorithm*, *Comput. Phys. Commun.* **220** (2017) 1 [[arXiv:1206.5214](#)] [[INSPIRE](#)].
- [39] RBC, UKQCD collaborations, *Domain wall QCD with physical quark masses*, *Phys. Rev. D* **93** (2016) 074505 [[arXiv:1411.7017](#)] [[INSPIRE](#)].
- [40] Y. Iwasaki, *Renormalization group analysis of lattice theories and improved lattice action: Two-dimensional non-linear  $O(N)$  sigma model*, *Nucl. Phys. B* **258** (1985) 141 [[INSPIRE](#)].
- [41] G. Mcglynn, *Algorithmic improvements for weak coupling simulations of domain wall fermions*, *PoS LATTICE2015* (2016) 019 [[INSPIRE](#)].
- [42] P.A. Boyle, G. Cossu, A. Yamaguchi and A. Portelli, *Grid: A next generation data parallel C++ QCD library*, *PoS LATTICE2015* (2016) 023 [[INSPIRE](#)].
- [43] A. Yamaguchi, P. Boyle, G. Cossu, G. Filaci, C. Lehner and A. Portelli, *Grid: OneCode and FourAPIs*, *PoS LATTICE2021* (2022) 035 [[arXiv:2203.06777](#)] [[INSPIRE](#)].
- [44] A. Portelli et al., *aportelli/hadrons: Hadrons v1.3*, *Zenodo* (Mar. 2022).
- [45] P.A.M. Dirac, *Gauge invariant formulation of quantum electrodynamics*, *Can. J. Phys.* **33** (1955) 650 [[INSPIRE](#)].
- [46] M. Hansen, B. Lucini, A. Patella and N. Tantalo, *Gauge invariant determination of charged hadron masses*, *JHEP* **05** (2018) 146 [[arXiv:1802.05474](#)] [[INSPIRE](#)].
- [47] V. Furman and Y. Shamir, *Axial symmetries in lattice QCD with Kaplan fermions*, *Nucl. Phys. B* **439** (1995) 54 [[hep-lat/9405004](#)] [[INSPIRE](#)].
- [48] S. Aoki and Y. Taniguchi, *One loop calculation in lattice QCD with domain wall quarks*, *Phys. Rev. D* **59** (1999) 054510 [[hep-lat/9711004](#)] [[INSPIRE](#)].
- [49] P.A. Boyle, *Four momentum boosted fermion fields*, *Nucl. Phys. B Proc. Suppl.* **129** (2004) 358 [[hep-lat/0309100](#)] [[INSPIRE](#)].
- [50] P.F. Bedaque, *Aharonov-Bohm effect and nucleon nucleon phase shifts on the lattice*, *Phys. Lett. B* **593** (2004) 82 [[nucl-th/0402051](#)] [[INSPIRE](#)].
- [51] G.M. de Divitiis, R. Petronzio and N. Tantalo, *On the discretization of physical momenta in lattice QCD*, *Phys. Lett. B* **595** (2004) 408 [[hep-lat/0405002](#)] [[INSPIRE](#)].
- [52] C.T. Sachrajda and G. Villadoro, *Twisted boundary conditions in lattice simulations*, *Phys. Lett. B* **609** (2005) 73 [[hep-lat/0411033](#)] [[INSPIRE](#)].
- [53] RBC/UKQCD collaboration, *The kaon semileptonic form factor in  $N_f = 2 + 1$  domain wall lattice QCD with physical light quark masses*, *JHEP* **06** (2015) 164 [[arXiv:1504.01692](#)] [[INSPIRE](#)].
- [54] S.-Z. Huang, T.-N. Ruan, N. Wu and Z.-P. Zheng, *Solution to the Rarita-Schwinger equations*, *Eur. Phys. J. C* **26** (2003) 609 [[INSPIRE](#)].
- [55] H. Akaike, *A new look at the statistical model identification*, *IEEE Trans. Automatic Control* **19** (1974) 716.
- [56] R.D. Piyadi Gamage, W. Ning and A.K. Gupta, *Adjusted Empirical Likelihood for Time Series Models*, [arXiv:1602.09128](#).

- [57] W.I. Jay and E.T. Neil, *Bayesian model averaging for analysis of lattice field theory results*, *Phys. Rev. D* **103** (2021) 114502 [[arXiv:2008.01069](#)] [[INSPIRE](#)].
- [58] J. Bijnens, G. Ecker and J. Gasser, *Radiative semileptonic kaon decays*, *Nucl. Phys. B* **396** (1993) 81 [[hep-ph/9209261](#)] [[INSPIRE](#)].
- [59] V. Cirigliano, G. Ecker, H. Neufeld, A. Pich and J. Portoles, *Kaon Decays in the Standard Model*, *Rev. Mod. Phys.* **84** (2012) 399 [[arXiv:1107.6001](#)] [[INSPIRE](#)].
- [60] A. Patella, *QED Corrections to Hadronic Observables*, *PoS LATTICE2016* (2017) 020 [[arXiv:1702.03857](#)] [[INSPIRE](#)].
- [61] HPQCD, UKQCD collaborations, *High Precision determination of the  $\pi$ ,  $K$ ,  $D$  and  $D(s)$  decay constants from lattice QCD*, *Phys. Rev. Lett.* **100** (2008) 062002 [[arXiv:0706.1726](#)] [[INSPIRE](#)].
- [62] S. Durr et al., *The ratio  $F_K/F_\pi$  in QCD*, *Phys. Rev. D* **81** (2010) 054507 [[arXiv:1001.4692](#)] [[INSPIRE](#)].
- [63] RBC, UKQCD collaborations, *Continuum Limit Physics from 2+1 Flavor Domain Wall QCD*, *Phys. Rev. D* **83** (2011) 074508 [[arXiv:1011.0892](#)] [[INSPIRE](#)].
- [64] RBC, UKQCD collaborations, *Domain Wall QCD with Near-Physical Pions*, *Phys. Rev. D* **87** (2013) 094514 [[arXiv:1208.4412](#)] [[INSPIRE](#)].
- [65] E.E. Scholz and S. Durr, *Leptonic decay-constant ratio  $f_K/f_\pi$  from clover-improved  $N_f = 2 + 1$  QCD*, *PoS LATTICE2016* (2016) 283 [[arXiv:1610.00932](#)] [[INSPIRE](#)].
- [66] S. Dürer et al., *Leptonic decay-constant ratio  $f_K/f_\pi$  from lattice QCD using 2+1 clover-improved fermion flavors with 2-HEX smearing*, *Phys. Rev. D* **95** (2017) 054513 [[arXiv:1601.05998](#)] [[INSPIRE](#)].
- [67] QCDSF–UKQCD collaborations, *Flavour breaking effects in the pseudoscalar meson decay constants*, *Phys. Lett. B* **767** (2017) 366 [[arXiv:1612.04798](#)] [[INSPIRE](#)].
- [68] F. Niedermayer, *Exact chiral symmetry, topological charge and related topics*, *Nucl. Phys. B Proc. Suppl.* **73** (1999) 105 [[hep-lat/9810026](#)] [[INSPIRE](#)].
- [69] S. Capitani et al., *Renormalization and off-shell improvement in lattice perturbation theory*, *Nucl. Phys. B* **593** (2001) 183 [[hep-lat/0007004](#)] [[INSPIRE](#)].
- [70] Y. Shamir, *Chiral fermions from lattice boundaries*, *Nucl. Phys. B* **406** (1993) 90 [[hep-lat/9303005](#)] [[INSPIRE](#)].
- [71] H. Yin and R.D. Mawhinney, *Improving DWF Simulations: the Force Gradient Integrator and the Möbius Accelerated DWF Solver*, *PoS LATTICE2011* (2011) 051 [[arXiv:1111.5059](#)] [[INSPIRE](#)].
- [72] T. Blum, T. Izubuchi and E. Shintani, *New class of variance-reduction techniques using lattice symmetries*, *Phys. Rev. D* **88** (2013) 094503 [[arXiv:1208.4349](#)] [[INSPIRE](#)].
- [73] S. Capitani, *Lattice perturbation theory*, *Phys. Rept.* **382** (2003) 113 [[hep-lat/0211036](#)] [[INSPIRE](#)].
- [74] M. Tomii, *Unphysical poles and dispersion relations for Möbius domain-wall fermions in free field theory at finite  $L_s$* , *Phys. Rev. D* **96** (2017) 074504 [[arXiv:1706.03099](#)] [[INSPIRE](#)].