THE EUROPEAN
PHYSICAL JOURNAL C

Light quark masses in $N_f = 2 + 1$ lattice QCD with Wilson fermions

ALPHA Collaboration

M. Bruno[1] , **I. Campos**[2] , **P. Fritzsch**[1] , **J. Koponen**[3,a] , **C. Pena**[4,5] , **D. Preti**[6] , **A. Ramos**[7] , **A. Vladikas**[8]

[1] Theoretical Physics Department, CERN, 1211 Geneva 23, Switzerland
[2] Instituto de Física de Cantabria IFCA-CSIC, Avda. de los Castros s/n, 39005 Santander, Spain
[3] High Energy Accelerator Research Organization (KEK), Tsukuba, Ibaraki 305-0801, Japan
[4] Departamento de Física Teórica, Universidad Autónoma de Madrid, Cantoblanco, 28049 Madrid, Spain
[5] Instituto de Física Teórica UAM-CSIC, Universidad Autónoma de Madrid, C/ Nicolás Cabrera 13-15, Cantoblanco, 28049 Madrid, Spain
[6] INFN, Sezione di Torino, Via Pietro Giuria 1, 10125 Turin, Italy
[7] School of Mathematics, Hamilton Mathematics Institute, Trinity College Dublin, Dublin 2, Ireland
[8] INFN, Sezione di Tor Vergata, c/o Dipartimento di Fisica, Università di Roma "Tor Vergata", Via della Ricerca Scientifica 1, 00133 Rome, Italy

Abstract We present a lattice QCD determination of light quark masses with three sea-quark flavours ($N_f = 2 + 1$). Bare quark masses are known from PCAC relations in the framework of CLS lattice computations with a non-perturbatively improved Wilson-Clover action and a tree-level Symanzik improved gauge action. They are fully non-perturbatively improved, including the recently computed Symanzik counter-term $b_A - b_P$. The mass renormalisation at hadronic scales and the renormalisation group running over a wide range of scales are known non-perturbatively in the Schrödinger functional scheme. In the present paper we perform detailed extrapolations to the physical point, obtaining (for the four-flavour theory) $m_{u/d}(2 \text{ GeV}) = 3.54(12)(9)$ MeV and $m_s(2 \text{ GeV}) = 95.7(2.5)(2.4)$ MeV in the $\overline{\text{MS}}$ scheme. For the mass ratio we have $m_s/m_{u/d} = 27.0(1.0)(0.4)$. The RGI values in the three-flavour theory are $M_{u/d} = 4.70(15)(12)$ MeV and $M_s = 127.0(3.1)(3.2)$ MeV.

1 Introduction

The lattice regularisation of QCD provides a well-defined procedure for the determination of the fundamental parameters of the theory (i.e. the gauge coupling and the quark masses) from first principles. The aim of the present work is the determination of the three lightest quark masses (i.e. those of the up, down and strange flavours), in a framework in which the up and down quarks are degenerate and all heavier flavours (i.e. charm and above), if present in the theory, would be quenched (valence) degrees of freedom. This is known as $N_f = 2 + 1$ lattice QCD. Moreover, QED effects are ignored.

The three-flavour theory adopted in this paper is presumably sufficient for determining light quark masses due to the decoupling of heavier quarks [1–4]. Indeed, lattice world averages of light quark masses $m_{u/d}, m_s$ do not show a significant dependence on the number of flavours at low energies for $N_f \geq 2$ within present-day errors [5]. This also holds for the more accurately known renormalisation group independent ratio $m_{u/d}/m_s$. Recently, heavy-flavour decoupling has been substantiated also non-perturbatively [6].

This paper is based on large-scale $N_f = 2 + 1$ flavour ensembles produced by the Coordinated Lattice Simulation (CLS) effort [7,8]. The simulations employ a tree-level Symanzik-improved gauge action and a non-perturbatively improved Wilson fermion action; see references [9–12]. The sea quark content is made of a doublet of light degenerate quarks $m_{q,1} = m_{q,2}$, plus a heavier one $m_{q,3}$. At the physical point $m_{q,1}, m_{q,2} = m_{u/d} \equiv \frac{1}{2}(m_u + m_d)$ and $m_{q,3} = m_s$.

The bare quark masses produced by CLS [7,8] need to be combined with renormalisation and improvement coefficients in order to obtain renormalised quantities with $O(a^2)$ discretisation effects. We use ALPHA collaboration results for the quark mass renormalisation and Renormalisation Group (RG) running [13] in the Schrödinger functional (SF) scheme. Symanzik improvement is implemented for the removal of discretisation effects from correlation functions, leaving us with $O(a^2)$ uncertainties in the bulk and

Springer

$O(g_0^4 a)$ ones at the time boundaries. We find that correlation functions extrapolations to the continuum limit are compatible with an $O(a^2)$ overall behaviour. The counter-terms required for the improvement of the axial current are known from refs. [14–16]. The present work has combined all these elements, obtaining estimates of the up/down and strange quark masses, as well as their ratio. These are expressed as renormalisation scheme-independent and scale-independent quantities, known as Renormalisation Group Invariant (RGI) quark masses. Of course we also give the same results in the $\overline{\text{MS}}$ scheme at scale $\mu = 2$ GeV.

The bare dimensionless parameters of the lattice theory are the strong coupling $g_0^2 \equiv 6/\beta$ and the quark masses expressed in lattice units $am_{q,1} = am_{q,2}$ and $am_{q,3}$, with a the lattice spacing. They can be varied freely in simulations. Having chosen a specific discretisation of the QCD action, its parameters must be calibrated so that three "input" hadronic quantities (one for each bare mass parameter and one for the lattice spacing) attain their physical values. Other physical quantities can subsequently be predicted. Such input quantities are typically very well-known from experiment, but they also need to be precisely computed on the lattice; examples are ground state hadron masses and decay constants (m_π, f_π, m_K, f_K, ...). Since the majority of numerical large-scale simulations do not yet include the small strong-isospin breaking and electromagnetic effects, the physical input quantities have to be corrected accordingly. Following Ref. [8], we use the values of Ref. [17]

$$m_\pi^{\text{phys}} = 134.8(3)\,\text{MeV}, \quad m_K^{\text{phys}} = 494.2(3)\,\text{MeV},$$
$$f_\pi^{\text{phys}} = 130.4(2)\,\text{MeV}, \quad f_K^{\text{phys}} = 156.2(7)\,\text{MeV}. \quad (1.1)$$

The calibration of the lattice spacing, referred to as scale setting, usually singles out a dimensionful quantity as reference scale f_{ref} [MeV]. Its dimensionless counterpart af_{ref} is computed on the lattice for fixed values of the bare coupling at the point where the physical spectrum, such as $[am_\pi/(af_{\text{ref}})]_{g_0^2} \equiv m_\pi/f_{\text{ref}}$, is reproduced in the bare parameter space $(g_0^2, am_{q,i})$ of the lattice theory. In this way, the lattice spacings $a(g_0^2) = [af_{\text{ref}}]_{g_0^2}/f_{\text{ref}}$, and consequently all computed observables, are obtained in physical units. When simulations approach the point of physical mass parameters while the lattice spacing is lowered, computational demands rapidly increase. In the present work results are obtained at non-zero lattice spacings and at quark masses which correspond to unphysical meson and decay constant values. Thus our data need to be extrapolated to the continuum limit and extra/interpolated to the physical quark mass values. This is achieved with a joint chiral and continuum extrapolation. The present work pays particular attention to these extrapolations and interpolations and the ensuing sources of systematic error.

So far we did not specify the reference scale f_{ref}. In Ref. [8] the three-flavor symmetric combination $f_{\pi K}^{\text{phys}} = \frac{2}{3}(f_K^{\text{phys}} + \frac{1}{2}f_\pi^{\text{phys}}) = 147.6(5)$ MeV, (obtained from the physical input of Eqs. (1.1)) was used for calibration, and for the determination of the hadronic gradient flow scale t_0 [18]. [1] An artificial (theoretical) hadronic scale with mass dimension -2, t_0 is precisely computable with small systematic effects [19,20], and thus well-suited as intermediate scale on the lattice. Its physical value determined from CLS ensembles [8] reads

$$\sqrt{8t_0^{\text{phys}}} = 0.415(4)(2)\,\text{fm}, \quad (1.2)$$

at fixed

$$\phi_4 \equiv \left[8t_0(m_K^2 + \tfrac{1}{2}m_\pi^2)\right]^{\text{phys}} = 1.119(21), \quad (1.3)$$

where the first error of $\sqrt{8t_0^{\text{phys}}}$ is statistical and the second systematic.

The theoretical framework of our work is explained in Sect. 2. The definitions of bare current quark masses, their renormalisation parameters and the $O(a)$-improvement counter-terms are provided in standard ALPHA-collaboration fashion. There is also a fairly detailed exposition of how the so-called "chiral trajectory" (a line of constant physics – LCP) is traced by $N_f = 2 + 1$ CLS simulations. In Sect. 3 we outline the computations leading to renormalised current quark masses as functions of the pion squared mass. These are computed in the SF renormalisation scheme at a hadronic (low energy) scale. In Sect. 4 we perform the combined chiral and continuum limit extrapolations in order to obtain estimates of the physical up/down and strange quark masses. Details of the ansätze we have used are provided in Appendix A and in Appendix B. Our final results are gathered in Sect. 5. Preliminary results have been presented in [21].

2 Theoretical framework

We review our strategy for computing light quark masses with improved Wilson fermions. In what follows equations are often written for a general number of flavours N_f. In practice $N_f = 2 + 1$. Flavours 1 and 2 indicate the lighter fermion fields, which are degenerate; at the physical point their mass is the average up/down quark mass. Flavour 3 stands for the heavier fermion, corresponding to the strange quark at the physical point.

---

[1] The gradient flow scale t_0 is defined by the implicit relation $\{t^2\langle E(t)\rangle\}_{t=t_0} = 0.3$, for the finite Yang–Mills energy density $E(t)$ at flow time t; see Section 6 of Ref. [7].

2.1 Quark masses, renormalisation, and improvement

The starting point is the definition of bare correlation functions on a lattice with spacing is a and physical extension $L^3 \times T$:

$$f_{\mathrm{P}}^{ij}(x_0, y_0) \equiv -\frac{a^6}{L^3} \sum_{\mathbf{x}, \mathbf{y}} \langle P^{ij}(x_0, \mathbf{x}) P^{ji}(y_0, \mathbf{y}) \rangle,$$

$$f_{\mathrm{A}}^{ij}(x_0, y_0) \equiv -\frac{a^6}{L^3} \sum_{\mathbf{x}, \mathbf{y}} \langle A_0^{ij}(x_0, \mathbf{x}) P^{ji}(y_0, \mathbf{y}) \rangle, \qquad (2.1)$$

where the pseudoscalar density and axial current are

$$P^{ij}(x) \equiv \bar{\psi}^i(x) \gamma_5 \psi^j(x), \qquad (2.2)$$

$$A_0^{ij}(x) \equiv \bar{\psi}^i(x) \gamma_0 \gamma_5 \psi^j(x). \qquad (2.3)$$

The indices $i, j = 1, 2, 3$ label quark flavours, which are always distinct ($i \neq j$).

The bare current (or PCAC) quark mass is defined via the axial Ward identity at zero momentum and a plateau average between suitable initial and final time-slices $t_{\mathrm{i}} < t_{\mathrm{f}}$,

$$\begin{aligned} m_{ij} \equiv{} & \frac{a}{t_{\mathrm{f}} - t_{\mathrm{i}} + a} \\ & \times \sum_{x_0 = t_{\mathrm{i}}}^{t_{\mathrm{f}}} \frac{\left[\frac{1}{2}(\partial_0 + \partial_0^*) f_{\mathrm{A}_0}^{ij} + c_{\mathrm{A}} a \partial_0 \partial_0^* f_{\mathrm{P}}^{ij} \right](x_0, y_0)}{2 f_{\mathrm{P}}^{ij}(x_0, y_0)}, \end{aligned}$$
$$(2.4)$$

with the source P^{ji} positioned either at $y_0 = a$ or $y_0 = T - a$.[2] The mass-independent improvement coefficient c_{A} is determined non-perturbatively [14]. The average of two renormalised quark masses is then expressed in terms of the PCAC mass m_{ij} as follows:

$$\begin{aligned} \frac{m_{i\mathrm{R}} + m_{j\mathrm{R}}}{2} \equiv{} & m_{ij\mathrm{R}} = \frac{Z_{\mathrm{A}}(g_0^2)}{Z_{\mathrm{P}}(g_0^2, a\mu)} m_{ij} \\ & \times \left[1 + (b_{\mathrm{A}} - b_{\mathrm{P}}) a m_{\mathrm{q}, ij} + (\bar{b}_{\mathrm{A}} - \bar{b}_{\mathrm{P}}) a \mathrm{Tr}[M_{\mathrm{q}}] \right] \\ & + \mathrm{O}(a^2), \end{aligned}$$
$$(2.5)$$

where $M_{\mathrm{q}} \equiv \mathrm{diag}(m_{\mathrm{q},1}, m_{\mathrm{q},2}, \cdots, m_{\mathrm{q}, N_{\mathrm{f}}})$ is the matrix of the sea quark subtracted masses, characteristic of Wilson fermions. Given the bare mass parameter $m_{0,i} \equiv (1/\kappa_i - 8)/(2a)$, with κ_i the Wilson hopping parameter, these are defined as

$$m_{\mathrm{q},i} = 1/(2a\kappa_i) - 1/(2a\kappa_{\mathrm{cr}}) \equiv m_{0,i} - m_{\mathrm{cr}} \qquad (2.6)$$

---

[2] In our simulations we average correlation functions with the source at $y_0 = a$ and (time-reversed) correlation functions with the source at $y_0 = T - a$. Since bare quantities are computed on lattices with open boundary conditions in time, we do not use translation invariance for the source position.

where $m_{\mathrm{cr}} \sim 1/a$ is an additive mass renormalisation arising from the loss of chiral symmetry by the regularisation and κ_{cr} is the critical (chiral) point. The average of two subtracted masses is then denoted by $m_{\mathrm{q}, ij} \equiv \frac{1}{2}(m_{\mathrm{q},i} + m_{\mathrm{q},j})$ in Eq. (2.5).

The axial current normalisation $Z_{\mathrm{A}}(g_0^2)$ is scale-independent, whereas the current quark mass renormalisation parameter $1/Z_{\mathrm{P}}(g_0^2, a\mu)$ depends on the renormalisation scale μ. The renormalisation condition imposed on the pseudoscalar density operator P^{ji} defines the renormalisation scheme for the quark masses. The schemes used in the present work (SF and $\overline{\mathrm{MS}}$) are mass-independent. Pertinent details will be discussed in latter sections.

The improvement coefficients $b_{\mathrm{A}} - b_{\mathrm{P}}$ and $\bar{b}_{\mathrm{A}} - \bar{b}_{\mathrm{P}}$ of Eq. (2.5) cancel O(a) mass-dependent cutoff effects; they are functions of the bare gauge coupling g_0^2. The corresponding counter-terms of Eq. (2.5) contain the subtracted masses $a m_{\mathrm{q}, ij}$ and $\mathrm{Tr}[a M_{\mathrm{q}}]$, which require knowledge on the critical mass m_{cr}. This can be avoided by substituting these masses with current quark masses and their sum. Their relationship is [22],

$$m_{ij} = Z \left[m_{\mathrm{q}, ij} + (r_{\mathrm{m}} - 1) \frac{\mathrm{Tr}[M_{\mathrm{q}}]}{N_{\mathrm{f}}} \right] + \mathrm{O}(a), \qquad (2.7)$$

where $Z(g_0^2) \equiv Z_{\mathrm{P}}/(Z_{\mathrm{S}} Z_{\mathrm{A}})$ and $r_{\mathrm{m}}(g_0^2)$ are finite normalisations. Z_{S} is the renormalisation parameter of the non-singlet scalar density $S^{ij} \equiv \bar{\psi}^i \psi^j$ and $r_{\mathrm{m}}/Z_{\mathrm{S}}$ is the renormalisation parameter of the singlet scalar density, which indirectly defines r_{m}; cf. Ref. [22]. In the above we neglect O(a) terms, as they only contribute to O(a^2) in the b-counter-terms of Eq. (2.5). Substituting $a m_{\mathrm{q}, ij} \rightarrow a m_{ij}$ in the latter expression, we obtain

$$\begin{aligned} m_{ij\mathrm{R}}(\mu_{\mathrm{had}}) ={} & \frac{Z_{\mathrm{A}}(g_0^2)}{Z_{\mathrm{P}}(g_0^2, a\mu)} m_{ij} \left[1 + (\tilde{b}_{\mathrm{A}} - \tilde{b}_{\mathrm{P}}) a m_{ij} \right. \\ & + \left\{ (\tilde{b}_{\mathrm{A}} - \tilde{b}_{\mathrm{P}}) \frac{1 - r_{\mathrm{m}}}{r_{\mathrm{m}}} + (\bar{b}_{\mathrm{A}} - \bar{b}_{\mathrm{P}}) \frac{N_{\mathrm{f}}}{Z r_{\mathrm{m}}} \right\} \frac{a M_{\mathrm{sum}}}{N_{\mathrm{f}}} \right] \\ & + \mathrm{O}(a^2), \end{aligned}$$
$$(2.8)$$

where we define

$$\begin{aligned} & \tilde{b}_{\mathrm{A}} - \tilde{b}_{\mathrm{P}} \equiv \frac{b_{\mathrm{A}} - b_{\mathrm{P}}}{Z}, \\ & M_{\mathrm{sum}} \equiv m_{12} + m_{23} + \cdots + m_{(N_{\mathrm{f}}-1)N_{\mathrm{f}}} + m_{N_{\mathrm{f}}1} \\ & \qquad = Z r_{\mathrm{m}} \mathrm{Tr}[M_{\mathrm{q}}] + \mathrm{O}(a). \end{aligned}$$
$$(2.9)$$

To leading order in perturbation theory the difference $b_{\mathrm{A}} - b_{\mathrm{P}}$ is O(g_0^2) and equals $\tilde{b}_{\mathrm{A}} - \tilde{b}_{\mathrm{P}}$. However, non-perturbative estimates are likely to differ significantly, especially in the range of couplings g_0 considered here ($1.56 \lesssim g_0^2 \lesssim 1.76$). We will employ non-perturbative estimates of $b_{\mathrm{A}} - b_{\mathrm{P}}$ and Z; cf. Ref. [23]. The term multiplying M_{sum} contains $(1 - r_{\mathrm{m}})/r_{\mathrm{m}}$ and $(\bar{b}_{\mathrm{A}} - \bar{b}_{\mathrm{P}})$. In perturbation theory $r_{\mathrm{m}} =$

$1 + 0.001158\, C_F\, N_f\, g_0^4$ [24,25], $(1 - r_m)/r_m \sim O(g_0^4)$ and $(\bar{b}_A - \bar{b}_P) \sim O(g_0^4)$ [22]. A first non-perturbative study of the coefficients $\bar{b}_A$ and $\bar{b}_P$ produced noisy results with 100% errors [26]. Given the lack of robust non-perturbative results and the fact that the term in curly brackets is $O(g_0^4)$ in perturbation theory, it will be dropped in what follows.

Once the quark mass averages m_{12R} and m_{13R} are computed say, in the SF scheme at a scale μ_{had}, the three renormalised quark masses can be determined. Since we are working in the isospin limit ($m_{q,1} = m_{q,2}$), the lighter quark mass is given by m_{12R}. Then one can isolate m_{13R} from the ratio m_{13R}/m_{12R} in which, as seen from Eq. (2.8), the M_{sum} counter-term cancels out.

The ALPHA Collaboration is devoting considerable resources to the determination of the non-perturbative evolution of the renormalised QCD parameters (strong coupling and quark masses) between a hadronic and a perturbative energy scale ($\mu_{had} \leq \mu \leq \mu_{pt}$). Quark masses are renormalised at $\mu_{had} \sim O(\Lambda_{QCD})$ and evolved to $\mu_{pt} \sim O(M_W)$ [13,27–36] in the SF scheme [37,38]. Both renormalisation and RG-running are done non-perturbatively. At μ_{pt} perturbation theory is believed to be reliably controlled and we may safely switch to the conventionally preferred, albeit inherently perturbative $\overline{MS}$ scheme.

We will be quoting results also for the scheme- and scale-independent renormalisation group invariant (RGI) quark masses M_{12} and M_{13} (corresponding to the current masses m_{12} and m_{13}) as well as the physical RGI quark masses $M_{u/d}$ and M_s derived from them. They are conventionally defined in massless schemes [39] by

$$
M_i \equiv m_{iR}(\mu) \left[2b_0 g_R^2(\mu) \right]^{-\frac{d_0}{2b_0}}
$$
$$
\times \exp\left\{ -\int_0^{g_R(\mu)} dg \left[\frac{\tau(g)}{\beta(g)} - \frac{d_0}{b_0 g} \right] \right\}, \qquad (2.10)
$$

for each quark flavour i. In our opinion, M_i is better suited for comparisons either to experimental results or other theoretical determinations. Equation (2.10) is formally exact and independent of perturbation theory as long as the renormalised parameters (g_R, m_{iR}) and the continuum renormalisation group functions (i.e. the Callan–Symanzik β-function and the mass anomalous dimension τ) are known non-perturbatively with satisfactory accuracy [13,33–36]. Their computation in the SF scheme with $N_f = 3$ massless quarks has been carried out in Ref. [13].

Our determination of the renormalised quark masses is based on the bare current mass averages m_{ijR}; cf. Eqs. (2.5) and (2.8). The analogue of these expressions for the RGI mass averages is given by

$$
M_{ij} \equiv \frac{1}{2}(M_i + M_j) = \frac{M}{m_R(\mu_{had})}\, m_{ijR}(\mu_{had}). \qquad (2.11)
$$

Note that the ratio $M/m_R(\mu_{had})$ is flavour-independent; cf. Eq. (2.10). In Ref. [13] it has been computed in the SF scheme for the $N_f = 3$ massless flavours at $\mu_{had} = 233(8)$ MeV.

2.2 The chiral trajectory and scale setting

Our aim is to stay on a line of constant Physics within systematic uncertainties of $O(a^2)$, as we vary the bare parameters of the theory (i.e. the gauge coupling g_0 and the $N_f = 2 + 1$ quark masses). In particular, if the improved bare gauge coupling

$$
\tilde{g}_0^2 \equiv g_0^2 \left(1 + \frac{1}{N_f} b_g(g_0^2)\, a\, \mathrm{Tr}[M_q] \right) \qquad (2.12)
$$

is kept fixed in the simulations, so does the lattice spacing, with any fluctuations being attributed to $O(a^2)$-effects. The problem is that $b_g(g_0^2)$ is only known to one-loop order in perturbation theory [40,41]; $b_g^{PT} = 0.012 N_f g_0^2$. Thus, following refs. [42,43], we vary the quark masses at fixed g_0^2, ensuring that the trace of the quark mass matrix remains constant:

$$
\mathrm{Tr}[M_q] = 2m_{q,1} + m_{q,3} = \mathrm{const.} \qquad (2.13)
$$

In this way the improved bare gauge coupling $\tilde{g}_0^2$ is kept constant at fixed β for any b_g.[3]

This requirement leads to an unusual but unambiguous approach to the physical point, shown in the (M_{12}, M_{13})-plane in the left panel of Fig. 1. Initially, one starts at the symmetric point ($am_{q,1} = am_{q,2} = am_{q,3} = am_q^{sym}$) for some fixed coupling $\beta = 6/g_0^2$, and tunes the mass parameter of the simulation in such a way that $\mathrm{Tr}[M_q] = \mathrm{Tr}[M_q]_{phys}$ to a good approximation.[4] This is achieved by varying am_q^{sym} until $(m_K^2 + \frac{1}{2}m_\pi^2)/f_{ref}$ takes its physical value. Since it is proportional to $\mathrm{Tr}[M_q]$ at leading order in chiral perturbation theory (χPT), it suffices as tuning observable. In subsequent simulations, one successively lifts the mass-degeneracy towards the physical point by decreasing the light quark masses while maintaining the constant-trace condition. By doing so the physical strange quark mass is approached from below as in Fig. 1 (left panel). We call this procedure "the determination of the chiral trajectory".

Note that the improved renormalised quark mass matrix M_R is given by [22]

$$
\mathrm{Tr}[M_R] = Z_m r_m
$$
$$
\times \left[(1 + a\bar{d}_m \mathrm{Tr}[M_q])\mathrm{Tr}[M_q] + a d_m \mathrm{Tr}[M_q^2] \right] + O(a^2). \qquad (2.14)
$$

---

[3] In Ref. [44] (cf. Sect. 5.3.2) it was estimated that when, in some ensembles, $\mathrm{Tr}[M_q]$ is not constant, the resulting effect on the shift of the lattice spacing is about 6 per mille. This estimate was based on the 1-loop value of b_g, $b_A - b_P$ and Z.

[4] In practise one only tunes the bare quark mass am_0, since am_{cr} is unknown a priori, but constant at fixed β.
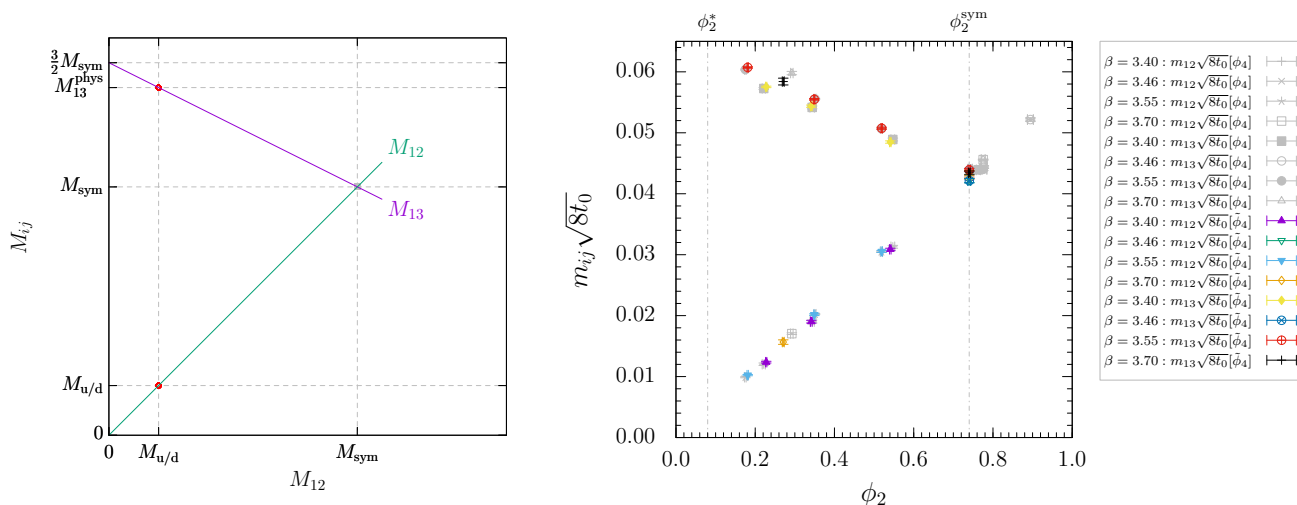
Fig. 1 The left panel shows an idealisation of the chiral trajectory for renormalised RGI current quark masses M_{12} and M_{13} in the continuum. The symmetric point (gray box) is defined by the trace of the renormalised RGI quark mass matrix, $M_{12} = M_{13} = M_{\text{sym}} = \frac{1}{3}\text{Tr}[M]$, and the physical point is indicated by red circles, where $M_{12}^{\text{phys}} = M_{\text{u/d}}$ and

$M_{13}^{\text{phys}} = \frac{1}{2}(M_{\text{u/d}} + M_{\text{s}})$. The right panel shows our data $\phi_{ij} \equiv \sqrt{8t_0}m_{ij}$ versus $\phi_2 \equiv 8t_0 m_\pi^2 \propto m_{12}$. Coloured (gray) points correspond to mass-shifted (-unshifted) points in parameter space, cf. the discussion in the text

Since the d_m-counter-term is proportional to squared bare masses, a constant $\text{Tr}[M_q]$ does not correspond to a constant $\text{Tr}[M_R]$; the latter requirement is violated by $O(a)$ effects. This is an undesirable feature, as it implies that the chiral trajectory is not a line of constant-physics. In practice these violations have been monitored in Ref. [8] (Fig. 4, lowest lhs panel), where $\text{Tr}[M_R]$ has been computed, at constant $\text{Tr}[M_q]$, from the current quark masses with 1-loop perturbative Symanzik b-coefficients. The violations appear to be bigger than what one would expect from $O(a)$ effects.

These considerations have led the authors of Ref. [8] to redefine the chiral trajectory in terms of $\phi_4 = \text{const.}$, where

$$\phi_4 \equiv 8t_0\left(m_K^2 + \frac{1}{2}m_\pi^2\right), \tag{2.15}$$

and t_0 is the gluonic quantity of the Wilson flow [18]; it has mass dimension -2. Here m_π and m_K are the lightest and strange pseudoscalar mesons respectively; at the physical point these are the pion m_π^{phys} and kaon m_K^{phys}. Keeping ϕ_4 constant is a Symanzik-improved constant physics condition. But ϕ_4 is proportional to the sum of the three quark masses only in leading-order (LO) chiral perturbation theory (χPT). Thus, the improved bare coupling $\tilde{g}_0^2$ now suffers from $O(am_q\text{Tr}[M_q])$ discretisation effects due to higher-order χPT contributions. In practice, these turn out to be small, as can be seen from Ref. [8] (Fig. 4, lowest rhs panel), where $\text{Tr}[M_R]$ has been computed, at constant ϕ_4. The violations appear to be at most 1% and thus the variation of the $O(a)$ b_g-term in $\tilde{g}_0^2$ can be ignored.

Obviously, one must also ensure, through careful tuning, that the chosen $\phi_4 = \text{const.}$ trajectory passes through the point corresponding to physical up/down and strange renormalised masses (i.e. quark masses that correspond to the physical pseudoscalar mesons m_π^{phys} and m_K^{phys}). This is done by driving ϕ_4 to its physical value $\phi_4^{\text{phys}} = 8t_0[(m_K^{\text{phys}})^2 + (m_\pi^{\text{phys}})^2/2]$ through mass shifts [8]. The aim is to express the computed quantities of interest (in our case the quark masses) as functions of

$$\phi_2 \equiv 8t_0 m_\pi^2, \tag{2.16}$$

with ϕ_4 held fixed at ϕ_4^{phys}, and eventually extrapolate them to $\phi_2^{\text{phys}} = 8t_0(m_\pi^{\text{phys}})^2$.

The determination of the redefined chiral trajectory is not straightforward. One needs to know the value of ϕ_4^{phys}. The latter is obtained from t_0 and the pseudoscalar masses (corrected for isospin-breaking effects) quoted in Eq. (1.1). But since the value of t_0 is only approximately known, one starts with an initial guess $\tilde{t}_0$, which provides an initial guess $\tilde{\phi}_4$. At each β, the symmetric point with degenerate masses ($\kappa_1 = \kappa_3$) is tuned so that the computed t_0/a^2, am_π and am_K combine as in Eq. (2.15) to give a value close to $\tilde{\phi}_4$. The other ensembles at the same β have been obtained by decreasing the degenerate (lightest) quark mass $m_{q,1} = m_{q,2}$, while increasing the heavier mass $m_{q,3}$ so as to keep $\text{Tr}[M_q]$ constant. Thus they do not correspond exactly to the same $\tilde{\phi}_4$. Small corrections of the subtracted bare quark masses (or hopping parameters) are introduced, using a Taylor expansion discussed in Sect. IV of Ref. [8], in order to shift ϕ_4 to the reference value $\tilde{\phi}_4$ and correct analogously the measured PCAC quark masses

and other quantities of interest such as the decay constants. The procedure is repeated for each β and the same value $\tilde{\phi}_4$ at the starting symmetric point.

All shifted quantities are now known at $\tilde{\phi}_4$ as functions of ϕ_2. Defining the combination of decay constants

$$f_{\pi K} \equiv \frac{2}{3}\left(f_K + \frac{f_\pi}{2}\right), \tag{2.17}$$

the dimensionless $\sqrt{8\tilde{t}_0}\, f_{\pi K}$ is computed for all ϕ_2 and extrapolated to $\tilde{\phi}_2 = 8\tilde{t}_0(m_\pi^{\text{phys}})^2$. The extrapolated $\sqrt{\tilde{t}_0}\, f_{\pi K}$, combined with the experimentally known $f_{\pi K}^{\text{phys}}$, gives a better estimate of $\tilde{t}_0$, and thus of $\tilde{\phi}_4$. As described in Sect. V of Ref. [8], this procedure can be recursively repeated and eventually the physical value of t_0 is fixed through $f_{\pi K}^{\text{phys}}$; the value in Eq. (1.2) from Ref. [8] leads to

$$\phi_4^{\text{phys}} = 1.119(21), \tag{2.18}$$
$$\phi_2^{\text{phys}} = 0.0804(8). \tag{2.19}$$

The main message is that once PCAC quark masses are shifted onto the chiral trajectory defined by the constant ϕ_4^{phys}, they only depend on a single variable, namely ϕ_2.

In analogy to the definitions (2.15) and (2.16), we also define rescaled dimensionless bare current quark masses and their renormalised counterparts at scale μ_{had}

$$\phi_{ij} \equiv \sqrt{8\,t_0}\, m_{ij}, \quad \phi_{ijR}(\mu) \equiv \sqrt{8\,t_0}\, m_{ijR}(\mu). \tag{2.20}$$

The redefined chiral trajectory is shown in Fig. 1 (right panel), where the light-light and heavy-light dimensionless mass averages (ϕ_{12R} and ϕ_{13R} respectively) are plotted as functions of ϕ_2. Extrapolating in ϕ_2 to ϕ_2^{phys} amounts to the simultaneous approach of the light and heavy quark masses to the corresponding physical up/down and strange values. All other physical quantities are then also at the physical point. Section 4 is dedicated to these extrapolations.

3 Quark mass computations

We base our determination of quark masses on the CLS ensembles for $N_f = 2 + 1$ QCD, listed in Table 1. The bare gauge action is the Lüscher–Weisz one, with tree-level coefficients [11]. The bare quark action is the Wilson, Symanzik-improved [10] one. The Clover term coefficient c_{sw} has been tuned non-perturbatively in Ref. [12]. Boundary conditions are periodic in space and open in time, as detailed in Ref. [45].

For details on the generation of these ensembles see Ref. [7]. As seen in Table 1, results have been obtained at four lattice spacings in the range $0.05 \lesssim a/\text{fm} \lesssim 0.086$. For each lattice coupling $\beta = 6/g_0^2$, gauge field ensembles have been generated for a few[5] values of the Wilson hopping param-

---

[5] We note in passing that for the ensemble with $\beta = 3.46$ we only have results for degenerate quark masses.

eters $\kappa_1 = \kappa_2$ and κ_3. The light pseudoscalar meson (pion) varies between 200 and 420 MeV. The heaviest value corresponds to the symmetric point where the three quark masses and the pseudoscalar mesons are degenerate. The strange meson (kaon) varies between 420 and 470 MeV. Given that our lightest pseudoscalars are relatively heavy (200 MeV), the chiral limit ought to be taken with care.

The bare correlation functions f_P^{ij}, f_A^{ij} of Eqs. (2.1) are estimated with stochastic sources located on time slice y_0, with either $y_0 = a$ or $y_0 = T - a$. From them the current quark masses m_{12}, m_{13} are computed as in Eq. (2.4), with the $O(a)$-improvement coefficient c_A determined non-perturbatively in Ref. [14]. The exact procedure to select the plateaux range in the presence of open boundary conditions has been explained in Refs. [7,44,46].

Having obtained the bare current quark masses m_{12}, m_{13} at four values of the coupling g_0^2, we construct the renormalised dimensionless quantities $m_{12R}(\mu_{\text{had}})$ and $m_{13R}(\mu_{\text{had}})$; cf. Eq. (2.8). For this we need the ratio $Z_A(g_0^2)/Z_P(g_0^2, \mu_{\text{had}})$ and the Symanzik b-counter-terms. Results for the axial current normalisation $Z_A(g_0^2)$ are available in Ref. [47], from a separate computation based on the chirally rotated Schrödinger Functional setup of Refs. [48–50]. The computation of $Z_P(g_0^2, \mu_{\text{had}})$ in the SF scheme, for $\mu_{\text{had}} = 233(8)$ MeV, was carried out in Ref. [13] for a theory with $N_f = 3$ massless quarks and the lattice action of the present work. The Z_P results, shown in Eqs. (5.2) and (5.3) of Ref. [13], are in a range of inverse gauge couplings which covers the $\beta \in [3.40, 3.85]$ interval of the large volume simulations of Ref. [8], from which our bare dimensionless PCAC masses are extracted.

Besides the ratio Z_A/Z_P, we also need the improvement coefficient $(\tilde{b}_A - \tilde{b}_P)$, which multiplies the $O(a)$ counter-term proportional to am_{ij} in Eq. (2.8). To leading order in perturbation theory $\tilde{b}_A - \tilde{b}_P = -0.0012g_0^2$. Non-perturbative estimates based on a coordinate-space renormalisation scheme have been provided for $N_f = 2 + 1$ lattice QCD in Ref. [26]. More accurate non-perturbative results have been subsequently obtained by the ALPHA Collaboration, using suitable combinations of valence current quark masses, measured on ensembles with $N_f = 3$ nearly-chiral sea quark masses in small physical volumes [16,23]. Also these simulations have been carried out in an inverse coupling range that spans the interval $\beta \in [3.40, 3.85]$ of the large volume CLS results of Ref. [8]. They are expressed in the form of ratios R_{AP} and R_Z, from which $(b_A - b_P)$ and Z are estimated; thus $(\tilde{b}_A - \tilde{b}_P) = R_{\text{AP}}/R_Z$. In Ref. [23], results are quoted for two values of constant Physics, dubbed LCP-0 and LPC-1. In LCP-0, R_{AP} and R_Z are obtained with all masses in the chiral limit. In LCP-1, one valence flavour is in the chiral limit (so it is equal to the sea quark mass), while a second one is held fixed to a non-zero value. The physical volumes are always kept fixed. In Ref. [23], Eqs. (5.1), (5.2) and (5.3) refer to

Table 1 Details of CLS configuration ensembles, generated as described in Ref. [7]. In the last column, ensembles are labelled by a letter, denoting the lattice geometry, a first digit for the coupling and a further two digits for the quark mass combination

β	$\dfrac{a}{\text{fm}}$	L/a	T/L	κ_1	κ_3	$\dfrac{m_\pi}{\text{MeV}}$	$\dfrac{m_K}{\text{MeV}}$	$m_\pi L$	Label
3.40	0.086	32	3	0.13675962	κ_1	420	420	5.8	H101
		32	3	0.136865	0.136549339	350	440	4.9	H102
		32	3	0.136970	0.136340790	280	460	3.9	H105
		48	2	0.137030	0.136222041	220	470	4.7	C101
3.46	0.076	32	3	0.13688848	κ_1	420	420	5.2	H400
3.55	0.064	32	4	0.137000	κ_1	420	420	4.3	H200
		48	8/3	0.137000	κ_1	420	420	6.5	N202
		48	8/3	0.137080	0.136840284	340	440	5.4	N203
		48	8/3	0.137140	0.136720860	280	460	4.4	N200
		64	2	0.137200	0.136601748	200	480	4.2	D200
3.70	0.050	48	8/3	0.137000	κ_1	420	420	5.1	N300
		64	3	0.137123	0.1367546608	260	470	4.1	J303

LCP-0 results, while those in Eqs. (5.1), (5.4) and (5.5) refer to LCP-1; differences are due to O(a) discretisation effects.

We have opted to use the LCP-0 values of $\tilde{b}_A - \tilde{b}_P$ in the present work. The covariance matrices of the fit parameters of R_{AP} as well as those of R_Z are provided in Ref. [23]. We *assume* that the covariance matrix between fit parameters of R_{AP} and R_Z is nil. This is justified a posteriori, by repeating the analysis with LCP-1 values, as a means to estimate the magnitude of systematic errors arising from our choice. Moreover, we have also compared our LCP-0 results to those obtained from different fit functions, used in the preliminary analysis of Ref. [16], as well as from the perturbative estimate $\tilde{b}_A - \tilde{b}_P$. We find that the contribution arising from such variations is below $\sim 1\%$ of the total error on renormalised quark masses at the physical point.

As discussed in Sect. 2, the complicated Symanzik counter-term in curly brackets, multiplying aM_{sum} in Eq. (2.8), is O($g_0^4 a$) in perturbation theory. As there are no robust non-perturbative estimates of its magnitude at present, we will drop this term, *assuming* that the O($g_0^4 a$) effects it would remove are subdominant compared to O(a^2) uncertainties.

As already explained in Sect. 2.2, our analysis is based on the rescaled dimensionless quantities defined in Eqs. (2.15), (2.16), and (2.20). At each β value and for each gauge field configuration, we have results for t_0/a^2, am_{12} and am_{13} from Refs. [7,8], from which ϕ_{12} and ϕ_{13} are obtained. The error analysis is carried out using the Gamma method approach [51–54] and automatic differentiation for error propagation, using the library described in Ref. [55]. This takes into account all the existing errors and correlations in the data and ancillary quantities (renormalisation constants, improvement coefficients, etc.), and estimates autocorrelation functions (including exponential tails) to rescale the uncertainties correspondingly. Following [8], the estimate of the exponential autocorrelation times τ_{exp} used in the analysis is the one quoted in [7], viz.,

$$\tau_{\text{exp}} = 14(3)\,\frac{t_0}{a^2}. \tag{3.1}$$

We have checked that without attaching exponential tails statistical errors are 40–70% smaller in our final results. The full analysis has been crosschecked by an independent code based on (appropriately) binned jackknife error estimation. Note that one of the strengths of data analysis based on the Gamma-method is that each Monte Carlo ensemble is treated independently, and the final statistical uncertainty is determined as a sum in quadratures of the statistical fluctuations for each ensemble. This allows to trace back which fraction of the statistical variance comes from each ensemble or ancillary quantities, such as renormalisation constants (see References 5–7 in [55]). This feature will be exploited in the error budgets provided below.

The starting values for ϕ_{12R} and ϕ_{13R} on which the analysis is based are shown in Table 2, where renormalised quark masses are in the SF scheme at a scale $\mu_{\text{had}} = 233(8)$ MeV. By suitably fitting these quantities as functions of ϕ_2, and extrapolating to ϕ_2^{phys}, we obtain the results for physical up/down and strange quarks at scale μ_{had}, as detailed in Sect. 4. Only then do we convert them to the RGI masses, by multiplying them with the RG-running factor [13]

$$\frac{M}{m_R(\mu_{\text{had}})} = 0.9148(88), \tag{3.2}$$

with the error added in quadrature; cf. Eq. (2.11).

Before presenting our chiral fits in Sect. 4, we conclude this section with a comment on finite-volume effects. Current quark masses are not expected to be affected by finite-volume corrections, since their values are fixed by Ward identities. On the other hand, meson masses, decay constants, and the ratio t_0/a^2 are expected to suffer from such effects. This can be directly checked in the ensembles H200 and N202, obtained at $\beta = 3.55$ with degenerate masses and corresponding to volumes of about 2 fm and 3 fm respectively. A glance at the

Table 2 Rescaled dimensionless current quark masses ϕ_{12} and ϕ_{13}, renormalised in the SF scheme at μ_{had}, for each CLS ensemble used in our analysis. Note that for simulation points H102, H105, C101 more than one independent ensembles exist, which have been run with different algorithmic setups; we keep those separate before fits. All points have been shifted to the target chiral trajectory as described in the text, and the quoted errors contain both statistical uncertainties and the contribution from renormalisation and the mass shift

β	Ensemble	t_0/a^2	ϕ_2	ϕ_{12}	ϕ_{13}	ϕ_{12}/ϕ_{13}
3.40	H101	2.857(13)	0.747(18)	0.0917(26)	0.0917(26)	1
	H102r001	2.877(19)	0.547(20)	0.0673(27)	0.1047(28)	0.643(10)
	H102r002	2.883(18)	0.549(19)	0.0667(28)	0.1039(29)	0.642(10)
	H105	2.886(11)	0.346(20)	0.0416(25)	0.1167(27)	0.357(15)
	H105r005	2.896(38)	0.355(20)	0.0420(33)	0.1160(34)	0.362(19)
	C101	2.900(19)	0.238(24)	0.0279(31)	0.1236(27)	0.226(21)
	C101r014	2.899(14)	0.233(20)	0.0273(26)	0.1223(30)	0.222(17)
3.46	H400	3.656(20)	0.747(18)	0.0923(28)	0.0923(28)	1
3.55	N202	5.161(23)	0.747(18)	0.0978(26)	0.0978(26)	1
	N203	5.138(16)	0.526(19)	0.0684(27)	0.1128(26)	0.606(10)
	N200	5.155(16)	0.356(18)	0.0455(25)	0.1232(25)	0.369(13)
	D200	5.171(16)	0.189(20)	0.0237(27)	0.1232(29)	0.176(17)
3.70	N300r002	8.592(41)	0.747(18)	0.0988(29)	0.0988(30)	1
	J303	8.628(40)	0.278(20)	0.0364(31)	0.1326(39)	0.274(19)

relevant entries of Table II of Ref. [8] shows that quark masses do not change as the volume is varied, while meson masses and decay constants vary by about 2.5%, which corresponds to differences of about $2 - 3.5\sigma$.[6] Standard SU(3) χPT NLO formulae are available for masses and decay constants [56]; t_0/a^2 does not suffer from finite-volume effects up to NNLO corrections [20]. In particular, the χPT-predicted effects for meson masses are below the percent level, since the lattice spatial size in units of the inverse lightest pseudoscalar meson mass is in the range [3.9, 5.8]. On the other hand, by directly comparing the values in Table 2 obtained at the same lattice spacing and sea quark masses but different volumes (cf. Ref. [44]), it is seen that the finite-volume effects on t_0/a^2 and m_π^2 are comparable and come with opposite signs. As a result, they largely cancel in ϕ_2, the variable in which chiral fits are performed. Decay constants, which generally suffer from larger finite-volume effects than meson masses, enter our computation indirectly only – firstly through NLO terms in chiral fits, where the finite-volume correction is subleading, and secondly through the physical value of $\sqrt{8t_0}$ determined in [8], where these corrections have already been taken into account. We therefore expect that the quantities most affected by finite-volume effects are the rescaled current quark masses ϕ_{12}, ϕ_{13}, due to the presence of $\sqrt{8t_0}/a$ in their definition. As mentioned above, these are much smaller than our statistical uncertainty, cf. Table 2. In the rest of our analysis we will therefore neglect this source of uncertainty.

4 Extrapolations to physical quark masses

Having obtained the dimensionless renormalised current mass combinations $\phi_{12\mathrm{R}}$ and $\phi_{13\mathrm{R}}$ at each β as functions of ϕ_2, we now proceed with the determination of the physical values ϕ_{ud} and ϕ_{s}. This is done in fairly standard fashion through fits and extrapolations. To begin with, we note that the two lighter degenerate quark masses are simply given by ϕ_{12}, whereas the heavier strange one is obtained from the difference[7]

$$\phi_{\mathrm{h}} = 2\phi_{13} - \phi_{12}. \qquad (4.1)$$

It is then possible to perform simultaneous fits of ϕ_{12} and ϕ_{h} as functions of ϕ_2 and the lattice spacing, subsequently extrapolating the results to ϕ_2^{phys} of Eq. (2.19) and the continuum limit, so as to obtain ϕ_{ud} and ϕ_{s}. Variants of this method consist in simultaneous fits and extrapolations of either ϕ_{13} or ϕ_{12} on one hand and their ratio ϕ_{12}/ϕ_{13} on the other. These turn out to be advantageous, as does a certain combination of ratios involving ϕ_{12}, ϕ_{13}, ϕ_2, and ϕ_4, for reasons discussed below. We recall in passing that in the ratio ϕ_{12}/ϕ_{13} all renormalisation factors cancel.

We use fits based in chiral perturbation theory (χPT fits) which are expected to model the data well close to the chiral limit $\phi_2 = 0$. Recall that we have performed $N_{\mathrm{f}} = 2 + 1$ simulations on a chiral trajectory; starting from a symmetric point where all quark masses are degenerate, we increase the mass of the heavy quark while decreasing that of the light one, until the physical point is reached. Since both masses are varying, it is natural to use SU(3)$_{\mathrm{L}} \otimes$ SU(3)$_{\mathrm{R}}$ chiral perturba-

---

[6] The ensemble H200 is only used in this context in the present work. Since at $\beta = 3.55$ we have results at two larger volumes (N202/203/200 and D200), we do not use H200 results in our analysis.

[7] Henceforth all quark masses will be renormalised. In order to simplify the notation, we shall drop the subscript R from $\phi_{12\mathrm{R}}$, $\phi_{13\mathrm{R}}$ in this section, in Appendix A and in Appendix B.

tion theory, which bears explicit dependence on both masses. This works when *all* three quark masses in the simulations are light enough for say, NLO χPT with three flavours to provide reliable fits. In Ref. [57] it is stated that this is the case for their data, obtained with domain wall fermions, as long as the average quark mass satisfies $am_{\mathrm{avg}} < 0.01$. As seen in Table 2 of Ref. [8], our PCAC dimensionless quark masses am_{12} and am_{13} also satisfy this empirical constraint. The real test comes about a posteriori, when the $SU(3)_L \otimes SU(3)_R$ NLO ansätze are seen to fit our results well.

In Appendix A and Appendix B, ansätze for NLO χPT and discretisation effects are adapted to our specific parametrisation in terms of ϕ_2 and ϕ_4. For the current quark masses these are

$$\phi_{12} = \phi_2 \left[p_1 + p_2 \phi_2 + p_3 K \left(\mathcal{L}_2 - \frac{1}{3}\mathcal{L}_\eta \right) \right]$$
$$+ \frac{a^2}{8t_0} [C_0 + C_1 \phi_2], \tag{4.2}$$

$$\phi_{13} = \phi_K \left[p_1 + p_2 \phi_K + \frac{2}{3} p_3 K \mathcal{L}_\eta \right]$$
$$+ \frac{a^2}{8t_0} \left[\widetilde{C}_0 + \widetilde{C}_1 \phi_2 \right], \tag{4.3}$$

where $\phi_K = (2\phi_4 - \phi_2)/2$. The constants p_1, p_2, p_3 and K are related to standard χPT parameters in Eqs. (A.8)-(A.11), whereas the chiral logarithms $\mathcal{L}_2$ and $\mathcal{L}_\eta$ are defined in Eq. (A.12). For justification of the ansatz used for the discretisation effects, see comments after Eqs. (B.8) and (B.9). We stress again that ϕ_{12} and ϕ_{13} are functions of ϕ_2 only, ϕ_4 being held constant. They have common fit parameters p_1, p_2 and p_3.

Using the above expressions and consistently neglecting higher orders in the continuum χPT terms, we obtain the ratio of PCAC masses (cf. Eqs. (A.13) and (B.11))

$$\frac{\phi_{12}}{\phi_{13}} = \frac{2\phi_2}{2\phi_4 - \phi_2} \left[1 + \frac{p_2}{p_1} \left(\frac{3}{2}\phi_2 - \phi_4 \right) - \tilde{K} \left(\mathcal{L}_2 - \mathcal{L}_\eta \right) \right]$$
$$+ \frac{a^2}{8t_0} (2\phi_4 - 3\phi_2) \left[D_0 + D_1 \phi_2 \right]. \tag{4.4}$$

As discussed in Appendix B, the form of the cutoff effects respects the constraint $\phi_{12}/\phi_{13} = 1$ at the symmetric point $m_{\mathrm{q},1} = m_{\mathrm{q},3}$, which is exact at all lattice spacings by construction.

For the combination defined in Eq. (A.14), we have

$$\frac{4\phi_{13}}{2\phi_4 - \phi_2} + \frac{\phi_{12}}{\phi_2} = 3p_1 + 2p_2\phi_4 + p_3 K \left(\mathcal{L}_2 + \mathcal{L}_\eta \right)$$
$$+ \frac{a^2}{8t_0} \left[G_0 + G_1 \phi_2 \right]. \tag{4.5}$$

An alternative to NLO χPT fits is the use of power series, based simply on Taylor expansions around the symmetric point $m_{\mathrm{q},1} = m_{\mathrm{q},2} = m_{\mathrm{q},3}$, for which $\phi_2^{\mathrm{sym}} = 2\phi_4^{\mathrm{phys}}/3$:

$$\phi_{12} = s_0 + s_1(\phi_2 - \phi_2^{\mathrm{sym}}) + s_2(\phi_2 - \phi_2^{\mathrm{sym}})^2$$
$$+ \frac{a^2}{t_0} \left[S_0 + S_1(\phi_2 - \phi_2^{\mathrm{sym}}) \right], \tag{4.6}$$

$$\phi_{13} = s_0 + \tilde{s}_1(\phi_2 - \phi_2^{\mathrm{sym}}) + \tilde{s}_2(\phi_2 - \phi_2^{\mathrm{sym}})^2$$
$$+ \frac{a^2}{t_0} \left[S_0 + \tilde{S}_1(\phi_2 - \phi_2^{\mathrm{sym}}) \right]. \tag{4.7}$$

Note that imposing the constraint $\phi_{12} = \phi_{13}$ at the symmetric point implies that s_0 and S_0 are common fit parameters. These expansions are expected to give reliable results in the higher end of the ϕ_2 range, underperforming close to the chiral limit. They are thus complementary to the chiral fits, which are better suited for the small-mass regime. In this sense the two approaches may provide a handle to estimate the systematic uncertainties due to these fits and extrapolations.

We explore various fit variants, in order to unravel the presence of potentially significant systematic effects. They are encoded as follows:

- **Fitted quantities and ansätze**:
  [chi12] Fit of ϕ_{12} data only, using the χPT ansatz.
  [chi13] Fit of ϕ_{13} data only, using the χPT ansatz.
  [tay12] Fit of ϕ_{12} data only, using the Taylor expansion ansatz.
  [tay13] Fit of ϕ_{13} data only, using the Taylor expansion ansatz.
  [chipc] Combined fit to ϕ_{12} and ϕ_{13}, using χPT.
  [chirc] Combined fit to ϕ_{13} and ϕ_{12}/ϕ_{13}, using χPT.
  [chirr] Combined fit to the ratio ϕ_{12}/ϕ_{13} and the combination $2\phi_{13}/\phi_K + \phi_{12}/\phi_2$ using χPT.
  [tchir] Combined fit to ϕ_{13} and the ratio ϕ_{12}/ϕ_{13}, using the Taylor expansion for ϕ_{13} and χPT for ϕ_{12}/ϕ_{13}.
- **Discretisation effects**:
  [a1] Fits with terms $\propto a^2/t_0$ only.
  [a2] Fits with terms $\propto a^2/t_0$ and $\propto \phi_2 a^2/t_0$.
- **Cuts on pseudoscalar meson masses**:
  [420] Fit all available data, including the symmetric point; i.e. data satisfies $m_\pi \lesssim 420$ MeV.
  [360] Fit excluding the symmetric point; i.e. data satisfies $m_\pi \lesssim 360$ MeV.
  [300] Fit only points for which $m_\pi \leq 300$ MeV.

Any given fit will thus be labelled as [xxxxx] [yy] [zzz], using the above tags.

The results obtained with the various fit methods at the physical point (cf. Eq. (2.19)) are expressed in physical units by dividing them out by $\sqrt{8t_0^{\mathrm{phys}}}$. Multiplication by the factor of Eq. (3.2) subsequently gives the RGI mass estimates

Fig. 2 Results for the RGI light ($M_{u/d}$) and averaged ($M_{13}^{\text{phys}} = (M_{u/d} + M_s)/2$) quark masses from independent fits to either M_{12} or M_{13}. Results are converted to MeV by dividing out with $\sqrt{8t_0^{\text{phys}}}$. Dotted lines indicate the central value of the latest FLAG average [5] for reference
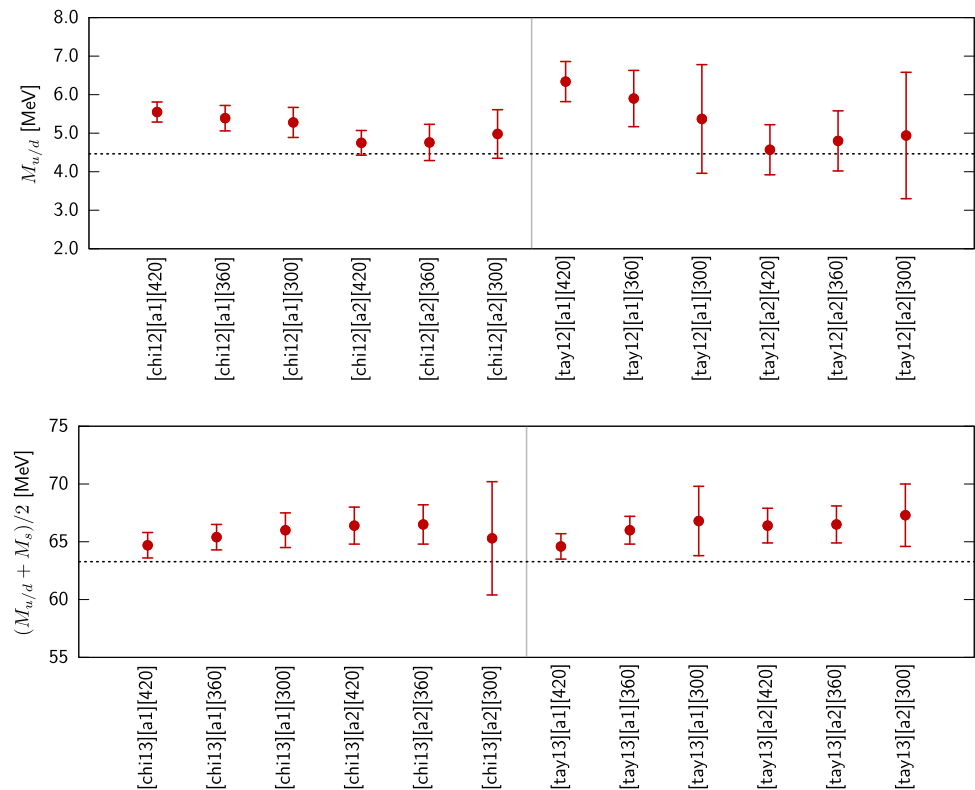
shown in Figs. 2 and 3. We comment on the various fit ansätze:

Independent fits of ϕ_{12} and ϕ_{13}: comparing light quark masses $M_{u/d}$ (upper panel of Fig. 2) from [chi12][a1] and [chi12][a2] we find that they are sensitive to the presence of a discretisation term $\propto a^2/t_0$, albeit within $\sim 1 - 2\sigma$. This difference is attenuated when the more stringent mass cutoff [chi12][300] is enforced, mainly because the error increases as less points are fitted. The same qualitative conclusions are true for the Taylor expansion fits [tay12] of the light quark mass. On the other hand, the lower panel of Fig. 2 shows that the average quark mass M_{13}^{phys} is not sensitive to the details of the fit ansätze. This is not surprising, given that our simulations have been performed in a region of rather heavy pions 220 MeV $\leq m_\pi \leq$ 420 MeV, with data covering the physical point M_{13}^{phys}, while $M_{u/d}$ requires long extrapolations. The conclusion is that independent fits are reliable for ϕ_{13} but less so for ϕ_{12}, and so we discard their results.

Combined fits to ϕ_{12} and ϕ_{13}: Fig. 3 shows that the fits [chipc][a1] and [chipc][a2] give results which are sensitive to the ansatz employed for the cutoff effects. This is more pronounced for $M_{u/d}$ and the ratio $M_s/M_{u/d}$, but persists also for M_s. Moreover, fits [chipc][a1][420] and [chipc][a1][360] display visible differences when compared to fits of the [chipc][a2] variety; the latter agree with results obtained from different fit ansätze. For

these reason we have also discarded results from this analysis.

Combined fits to ϕ_{13} and ϕ-ratios: As previously explained, we have explored three ansätze, namely [chirc], [chirr], and [tchir]. In all cases Fig. 3 shows that there is no significant dependence of the results from the details of these fits, except for a very slight fluctuation of the [tchir][a1][420] results for M_s. Preferring to err on the side of caution, we also discard [tchir] fits.

A few general points concerning the fit analysis deserve to be highlighted:

- In all our fits the $\chi^2/$dof is well below 1. This is partly because our data are correlated – both from the fact that there are common renormalisation factors and improvement coefficients, and because we are including the contribution to the χ^2 from the fluctuations of the meson masses (horizontal errors). Therefore, while the goodness-of-fit is in general satisfactory, we will refrain from quoting the corresponding p-values, since they are not really meaningful.

- Unsurprisingly, the inclusion of a second discretisation term $\propto \phi_2(a^2/t_0)$ in the fits contributes to an increase of the error. This term is often compatible with zero, and almost always so within $\sim 2\sigma$, suggesting that fits [a1] are safe. As stated previously, exceptions are fits [chi12] and [chipc], where inclusion of this term has a strong effect.

Fig. 3 Results for the RGI light ($M_{\mathrm{u/d}}$) and strange (M_{s}) quark masses, and their ratio, from the simultaneous fits [chipc], [chirc], [chirr], and [tchir]. Results are converted to MeV by dividing out with $\sqrt{8t_0^{\mathrm{phys.}}}$. Dotted lines indicate the central value of the latest FLAG average [5] for reference
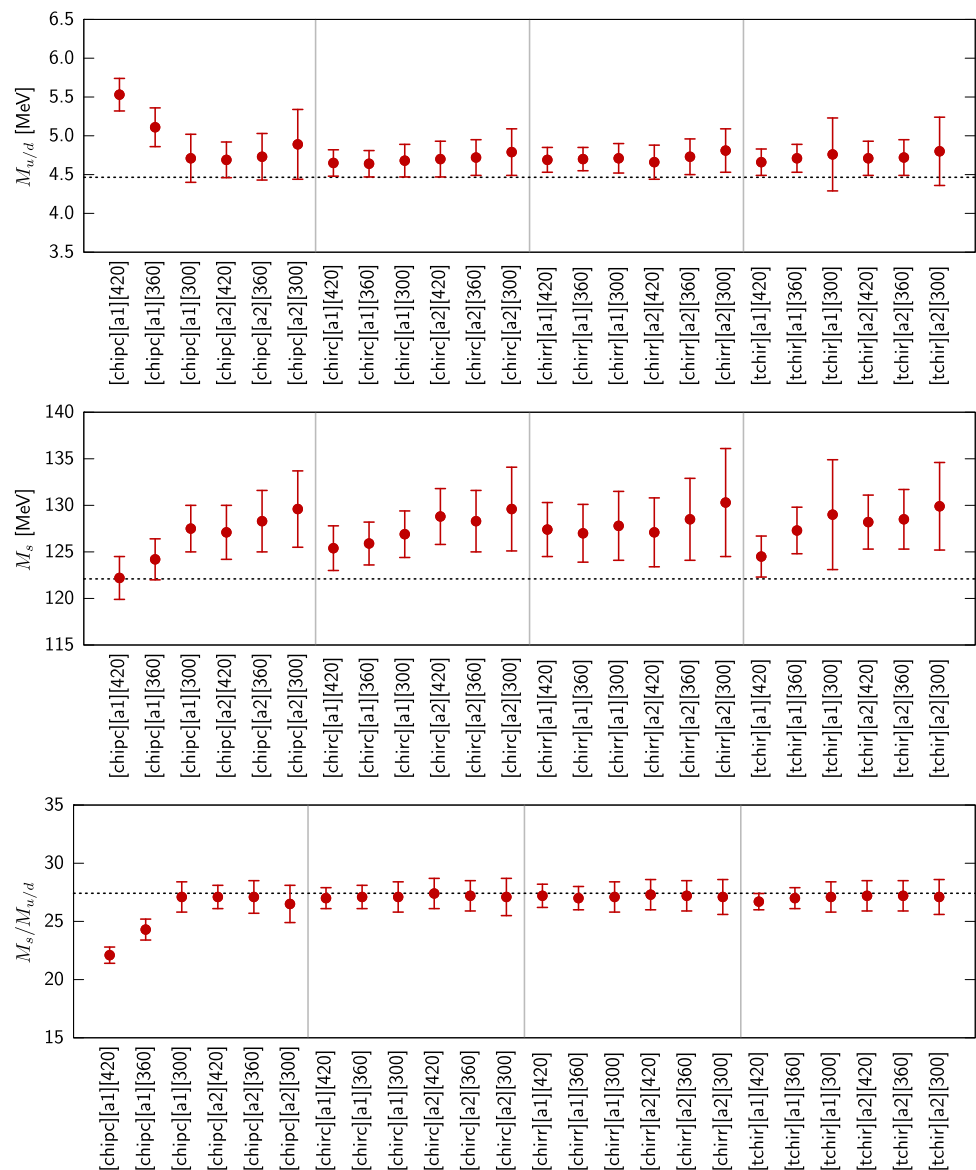
- Within large uncertainties, the coefficients of the leading cutoff effects (i.e. those $\propto a^2/t_0$) depend on the fitted observable, and are larger for ϕ_{13} than for ϕ_{12}.
- The power-series fits [tay12] and [tay13] behave remarkably well. Results from [tay12] vanish within errors in the chiral limit, except for fits going up to the symmetric point, which are sometimes incompatible with naught by 2–3 σ. This is evidence that our data are not precise enough to capture the impact of chiral logs. Fits [tay13] to ϕ_{lh} are very stable, and impressively better than those obtained with the χPT ansatz. Indeed, if one considers fits [texp1] and [texp2], which are safest from the point of view of error estimation, all the fits considered provide compatible results for M_{13} within one sigma. Notice, furthermore, that the constant terms of [tay12] and [tay13] are generally in good agree-

ment, signalling the consistency of the approach. It is also interesting to note that the coefficient of the quadratic term is very small and always compatible with zero within 1σ (save for two cases where it vanishes within 2σ).
- Fits [chirc] and [chirr] appear to be the stablest.
- NLO χPT appears to be suffering around and above 400 MeV.

5 Final results and discussion

Following the analysis of Sect. 4, we quote as final results those obtained from the following procedure:

- The central values are those of a combined fit to the ratio ϕ_{12}/ϕ_{13} and the quantity $2\phi_{13}/\phi_K + \phi_{12}/\phi_2$, using
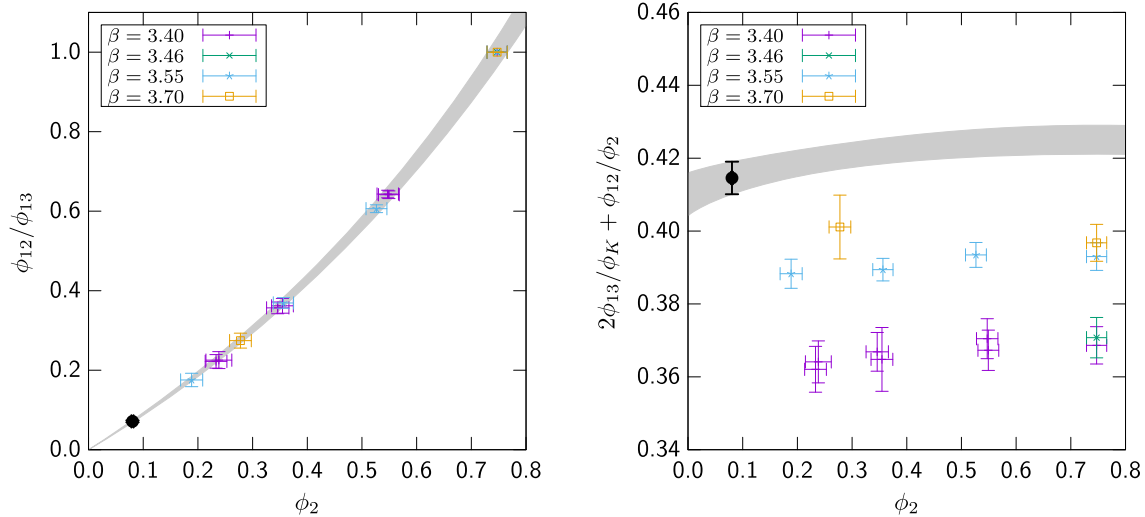
Fig. 4 Illustration of the chiral+continuum fit from which our central values are obtained. The grey band is the continuum limit of our fit, and the full black point corresponds to our extrapolation to the physical point

NLO χPT, with pseudoscalar meson masses less than 360 MeV and a discretisation term proportional to a^2/t_0 (i.e., fit [chirr][a1][360]). The error from this fit will appear as the first uncertainty in the results below. The fit is illustrated in Fig. 4.

- We estimate systematic errors from the spread of central values of all other [chirr] and [chirc] fits, for all pion mass cutoffs, and for both [a1] and [a2]. The spread is intended to be the difference between the central value, obtained as described in the previous item, and the most distant central value of all other [chirr] and [chirc] fits. This is the second error of the results below. Recall that [chirc] are combined fits to ϕ_{13} and the ratio ϕ_{12}/ϕ_{13}, using NLO χPT.
- Discard other fits, including [chipc], considered too unstable.
- All results have been obtained using the Symanzik $\tilde{b}$-parameters computed in the LCP-0 case (see discussion in Sect. 3). Using LPC-1 results instead, has very marginal effects on the error.
- In Sect. 3 we have also argued that for the quantities under consideration finite volume effects are negligible.

The resulting RGI masses are

$$M_s = 127.0(3.1)(3.2) \text{ MeV},$$
$$M_{u/d} = 4.70(15)(12) \text{ MeV}. \tag{5.1}$$

The quark mass ratio is obtained from

$$\frac{M_s}{M_{u/d}} = \frac{2}{\phi_{ll}/\phi_{lh}} - 1. \tag{5.2}$$

Table 3 Contributions to the squared errors of our final quantities from different sources

	$M_{u/d}$	M_s	$M_s/M_{u/d}$
Stat+chiral+cont	56%	40%	86%
Fit systematics	39%	52%	14%
Renormalisation	< 1%	< 1%	n/a
Running	5%	8%	n/a
O(a) impr	Negligible	Negligible	Negligible
Finite volume	Negligible	Negligible	Negligible

Dependence on renormalisation is only implicit, from the joint fit with ϕ_{13}. The same procedures as above yield

$$\frac{M_s}{M_{u/d}} = 27.0(1.0)(0.4). \tag{5.3}$$

The above results for RGI masses refer to the $N_f = 2+1$ theory.

It is customary in phenomenological studies to report light quark masses measured in the $N_f = 2+1$ lattice theory in the $\overline{\text{MS}}$ scheme at 2 GeV, referred to the more physical QCD with four flavours. This entails using $N_f = 3$ perturbative RG-running from 2 GeV down to the charm threshold, followed by $N_f = 4$ perturbative RG-running back to 2 GeV; see for example Ref. [5]. We use 4-loop perturbative RG-running and the value for the $\Lambda_{\text{QCD}}^{\overline{\text{MS}}}$ parameter computed by the ALPHA Collaboration in Ref. [34] to obtain[8]

$$m_{sR}(2 \text{ GeV}) = 95.7(2.5)(2.4) \text{ MeV},$$

---

[8] In converting our results to $\overline{\text{MS}}$ we have taken into account the uncertainty in the matching factor coming from the error on $\Lambda_{\text{QCD}}^{\overline{\text{MS}}}$, as well as the covariance of the latter with our determination of M_i.
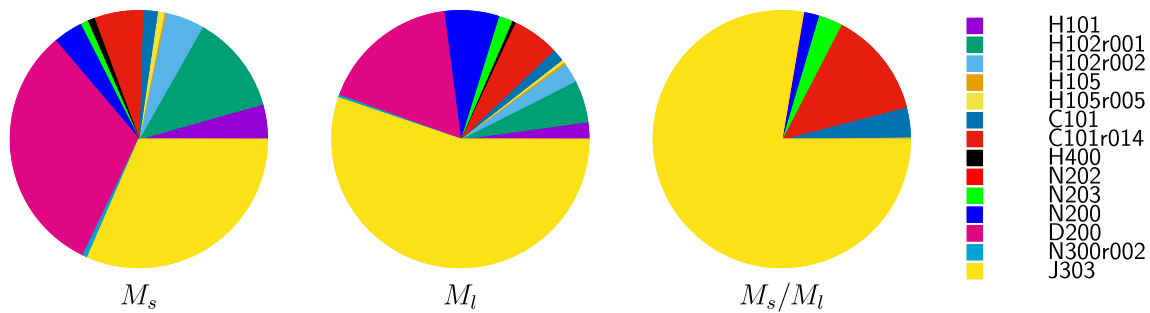
Fig. 5 Contributions to the statistical+chiral extrapolation+continuum limit uncertainties from each ensemble included in our analysis, for our preferred fit `[chirr][a1][360]`

$$m_{\mathrm{u/dR}}(2\,\mathrm{GeV}) = 3.54(12)(9)\,\mathrm{MeV}. \tag{5.4}$$

The mass ratio is obviously the same as in Eq. (5.3). We note in passing that switching to the four-flavour theory has a very small effect on $\overline{\mathrm{MS}}$ results, since at 2 GeV the matching factor is $m_{\mathrm{R}}(N_{\mathrm{f}}=4)/m_{\mathrm{R}}(N_{\mathrm{f}}=3) = 1.002$.

The error budget for our computation is summarised in Table 3 and Fig. 5. Uncertainties are completely dominated by our chiral fits. We have separated these errors into two contributions; see first two lines of Table 3. The first error is that of our best fit `[chirr][a1][360]`, and includes the statistical errors as well as the error from combined fits in ϕ_2 and a. The second uncertainty is the one arising upon varying the fit ansätze and their ϕ_2 range. All other errors are clearly seen to be subdominant. It is worth noting that, as expected, the largest contribution to the uncertainty comes from the ensembles with the lightest sea pion masses, especially the one with the finest lattice spacing. It is then clear that decreasing our errors would require more chiral ensembles, and more extensive simulations at light masses.

The current FLAG 2019 [5] world averages from $N_{\mathrm{f}} = 2 + 1$ simulations, in the $\overline{\mathrm{MS}}$ scheme, reportedly quoted for the $N_{\mathrm{f}} = 4$ theory as explained above, are:

$$m_{\mathrm{sR}}(2\,\mathrm{GeV}) = 92.03(88)\,\mathrm{MeV},$$
$$m_{\mathrm{u/dR}}(2\,\mathrm{GeV}) = 3.364(41)\,\mathrm{MeV}. \tag{5.5}$$

The strange mass estimate is based on the results of Refs. [58–63], while the up/down one is based on Refs. [58–61,64]. For the quark mass ratio, based on Refs. [58–60,63], FLAG quotes

$$\frac{m_{\mathrm{sR}}}{m_{\mathrm{u/dR}}} = 27.42(12)\,\mathrm{MeV}. \tag{5.6}$$

Our results for the strange and light quark masses agree with those of FLAG within 1.7σ and 1.2σ respectively and thus exhibit good compatibility albeit with bigger errors.

Appendix A Chiral perturbation theory expansions

We adapt standard χPT expressions to our specific parametrisation of the data, stemming from our choice of chiral trajectory. We start, for example, from Eqs. (B5) and (B6) of Ref. [57], which are NLO chiral expansions of the light and strange pseudoscalar mesons m_π and m_{K} in terms of light

and strange quark masses $m_1 = m_2$ and m_3. These series are inverted, so that quark masses are functions of meson masses. The PCAC quark mass combinations $m_{12} = (m_1 + m_2)/2$ and $m_{13} = (m_1 + m_3)/2$ are then formed and everything is re-expressed in terms of the dimensionless quark masses ϕ_{12}, ϕ_{13} and the dimensionless quantities ϕ_2 and ϕ_4, so that we arrive at

$$
\phi_{12} = \frac{\phi_2}{2B_0\sqrt{8t_0}} \cdot \left\{ 1 - \frac{16}{8t_0 f_0^2}(2L_8 - L_5)\phi_2 \right.
$$
$$
- \frac{32}{8t_0 f_0^2}(2L_6 - L_4)\phi_4 - \frac{1}{24\pi^2 8t_0 f_0^2}
$$
$$
\left. \times \left[\frac{3}{2}\phi_2 \ln\left(\frac{\phi_2}{8t_0\Lambda_\chi^2}\right) - \frac{1}{2}\phi_\eta \ln\left(\frac{\phi_\eta}{8t_0\Lambda_\chi^2}\right) \right] \right\},
$$
(A.1)

and

$$
\phi_{13} = \frac{2\phi_4 - \phi_2}{4B_0\sqrt{8t_0}} \cdot \left\{ 1 - \frac{8}{8t_0 f_0^2}(-2L_8 + L_5)\phi_2 \right.
$$
$$
- \frac{16}{8t_0 f_0^2}(4L_6 + 2L_8 - 2L_4 - L_5)\phi_4
$$
$$
\left. - \frac{1}{24\pi^2 8t_0 f_0^2}\phi_\eta \ln\left(\frac{\phi_\eta}{8t_0\Lambda_\chi^2}\right) \right\},
$$
(A.2)

where B_0, f_0, $L_k (k = 4, 5, 6, 8)$ are standard χPT parameters and

$$
\phi_\eta \equiv 8t_0 \frac{4m_K^2 - m_\pi^2}{3} = \frac{4\phi_4 - 3\phi_2}{3}.
$$
(A.3)

The NLO LECs L_i and B_0 are implicitly renormalised at scale Λ_χ. It is also useful to consider the ratio

$$
\frac{\phi_{12}}{\phi_{13}} = \frac{2\phi_2}{2\phi_4 - \phi_2}\left\{ 1 - \frac{24}{8t_0 f_0^2}(2L_8 - L_5)\left[\phi_2 - \frac{2}{3}\phi_4\right] \right.
$$
$$
- \frac{1}{16\pi^2(8t_0 f_0^2)}\left[\phi_2 \ln\left(\frac{\phi_2}{8t_0\Lambda_\chi^2}\right) \right.
$$
$$
\left. \left. - \phi_\eta \ln\left(\frac{\phi_\eta}{8t_0\Lambda_\chi^2}\right) \right] \right\},
$$
(A.4)

which does not require renormalisation. Note that at the symmetric point ($\phi_2 = 2\phi_4/3$) the current quark masses of Eqs. (A.1) and (A.2) respect the constraint $\phi_{12} = \phi_{13}$, while the ratio (A.4) is exactly 1. Note that the sum of ratios

$$
\frac{4\phi_{13}}{2\phi_4 - \phi_2} + \frac{\phi_{12}}{\phi_2} = \frac{3}{2B_0\sqrt{8t_0}}
$$
$$
\times \left\{ 1 - \frac{16}{8t_0 f_0^2}\left(\frac{4}{3}L_8 - \frac{2}{3}L_5 + 4L_6 - 2L_4\right)\phi_4 \right.
$$
$$
- \frac{1}{48\pi^2(8t_0 f_0^2)}\left[\phi_2 \ln\left(\frac{\phi_2}{8t_0\Lambda_\chi^2}\right) \right.
$$

$$
\left. + \phi_\eta \ln\left(\frac{\phi_\eta}{8t_0\Lambda_\chi^2}\right) \right] \right\},
$$
(A.5)

has the remarkable advantages of depending on just one combination of NLO LECs, and of being free of polynomial dependence on ϕ_2.

We next rewrite Eqs. (A.1) and (A.2) in forms which are suitable for combined fits, with common coefficients for ϕ_{12} and ϕ_{13}, obtaining

$$
\phi_{12} = \phi_2\left[p_1 + p_2\phi_2 + p_3 K\left(\mathcal{L}_2 - \frac{1}{3}\mathcal{L}_\eta\right)\right],
$$
(A.6)
$$
\phi_{13} = \frac{2\phi_4 - \phi_2}{2}\left[p_1 + p_2\left(\phi_4 - \frac{\phi_2}{2}\right) + \frac{2}{3}p_3 K\mathcal{L}_\eta\right],
$$
(A.7)

where the coefficients p_1, p_2, and p_3 relate to LECs as follows:

$$
p_1 = \frac{1}{2B_0\sqrt{8t_0}}\left[1 - \frac{32}{8t_0 f_0^2}(2L_6 - L_4)\phi_4\right]
$$
$$
\approx \frac{1}{2B_0\sqrt{8t_0}}\left[1 - \frac{32}{8t_0 f_{\pi K}^2}(2L_6 - L_4)\phi_4\right],
$$
(A.8)
$$
p_2 = -\frac{1}{2B_0\sqrt{8t_0}}\frac{16}{8t_0 f_0^2}(2L_8 - L_5)
$$
$$
\approx -\frac{1}{2B_0\sqrt{8t_0}}\frac{16}{8t_0 f_{\pi K}^2}(2L_8 - L_5),
$$
(A.9)
$$
p_3 = -\frac{1}{2B_0\sqrt{8t_0}}.
$$
(A.10)

We also define

$$
K \equiv (8t_0 16\pi^2 f_0^2)^{-1} \approx (8t_0 16\pi^2 f_{\pi K}^2)^{-1},
$$
(A.11)

with $f_{\pi K}$ given by Eq. (2.17). The chiral logarithms are

$$
\mathcal{L}_2 \equiv \phi_2 \ln\phi_2, \quad \mathcal{L}_\eta \equiv \phi_\eta \ln\phi_\eta.
$$
(A.12)

The following points should be kept in mind:

- We are using only configurations along the $\phi_4 = $ constant chiral trajectory. Terms proportional to ϕ_4 are thus reabsorbed into constant fit terms.
- Our expressions are linear in fit parameters, rather than non-linear factors in which LECs appear explicitly. Determination of LECs is beyond the scope of the present work.
- By replacing f_0^2 by $f_{\pi K}^2$ in the above definition of K, the coefficients of chiral logarithms are completely fixed relative to the LO value; cf. Eqs. (A.1) and (A.2). This eliminates one fit parameter, pushing its effect to NNLO

LECs. In practice, the fact that terms with ϕ_4 are reabsorbed into the LO terms nullifies the effect in some fits, e.g., those for ϕ_{12} and ϕ_{13}. A second advantage of this choice is that the resulting ansätze are fully linear in the fit parameters. See also Eq. (2.5) in [8] and comments therein on the reasons that $f_0 \approx f_{\pi K}$ and for preferring $f_{\pi K}$ to f_0.

- We conveniently set the renormalisation scale to $\Lambda_\chi = 1/\sqrt{8t_0} \simeq 476$ MeV, simplifying the chiral logs. There is no need to reabsorb $\ln(8t_0 \Lambda_\chi^2)$ terms in fit parameters. This is an unconventional choice, as common practice consists in providing results for LECs at $\Lambda_\chi = m_\rho$ or $\Lambda_\chi = 4\pi f_0$. Consequently, NLO LECs eventually obtained with our methodology may only be compared to results in the literature after some extra work.

Using the above expressions and consistently neglecting higher mass orders, we obtain for the ratio (A.4) of PCAC masses

$$\frac{\phi_{12}}{\phi_{13}} = \frac{2\phi_2}{2\phi_4 - \phi_2}\left[1 + \frac{p_2}{p_1}\left(\frac{3}{2}\phi_2 - \phi_4\right) - \tilde{K}\left(\mathcal{L}_2 - \mathcal{L}_\eta\right)\right].$$
(A.13)

For the combination (A.5) we have

$$\frac{4\phi_{13}}{2\phi_4 - \phi_2} + \frac{\phi_{12}}{\phi_2} = 3p_1 + 2p_2\phi_4 + p_3 K\left(\mathcal{L}_2 + \mathcal{L}_\eta\right).$$
(A.14)

With ϕ_4 held constant, the quantities of Eqs. (A.6), (A.7), (A.13), and (A.14) are functions of ϕ_2 only. We use these expressions to fit our data, after adding $O(a^2)$ terms which model leading discretisation effects that have been neglected throughout this "Appendix".

Appendix B Discretisation effects

In order to parametrise the discretisation effects of the quantities we fit, we first examine ϕ_{ij}; cf. Eqs. (2.5) and (2.20). It can be written in the very general form

$$\phi_{ij} = \phi_{ij}^{\text{cont}} + f\left(a, \frac{m_i + m_j}{2}, \frac{m_i - m_j}{2}, \text{Tr}[M_q]\right),$$
(B.1)

where ϕ_{ij}^{cont} is the continuum quantity and the function f contains the discretisation effects which in general depend on the lattice spacing a, the quark masses m_i, m_j, and the trace of the mass matrix $\text{Tr}[M_q]$. As we have discussed in Sect. 2, we will ignore $O(g_0^4 \text{Tr}[M_{\text{sum}}])$ discretisation effects and only consider the influence of $O(a^2)$ uncertainties. Also ϕ_{ij} has to be symmetric with respect to the exchange of quarks, $i \leftrightarrow j$. We can thus parametrise f as follows:

$$f\left(a, \frac{m_i + m_j}{2}, \frac{m_i - m_j}{2}, \text{Tr}[M_q]\right)$$

$$= c_0\frac{a^2}{t_0} + c_1\frac{a^2}{t_0}\sqrt{8t_0}\left(\frac{m_i + m_j}{2}\right) + c_2\frac{a^2}{t_0}\sqrt{8t_0}\text{Tr}[M_q]$$

$$+ c_3\frac{a^2}{t_0}8t_0\left(\frac{m_i + m_j}{2}\right)^2 + c_4\frac{a^2}{t_0}8t_0\left(\frac{m_i - m_j}{2}\right)^2$$

$$+ c_5\frac{a^2}{t_0}8t_0\text{Tr}[M_q^2] + c_6\frac{a^2}{t_0}8t_0(\text{Tr}[M_q])^2$$

$$+ c_7\frac{a^2}{t_0}\sqrt{8t_0}\left(\frac{m_i + m_j}{2}\right)\sqrt{8t_0}\text{Tr}[M_q] + O(a^3). \quad \text{(B.2)}$$

A further simplification is brought about by neglecting the dependence of $c_0, \ldots, c_7$ on the bare coupling g_0^2.

Next we write the function f in terms of ϕ_2, recalling that a constant ϕ_4 constrains the relation between the heavier (strange) and light quark masses. This is done by first expressing the current quark masses on the rhs of the above equation in terms of ϕ_{12} and ϕ_{13}, followed by using their LO χPT relations to ϕ_2 and ϕ_4. In particular, with $\beta_0 \equiv 1/(2B_0\sqrt{8t_0})$, we see from Eqs. (A.6), (A.7) that to LO:

$$\phi_{12} \overset{\text{LO}}{=} \beta_0\phi_2,$$
(B.3)

$$\phi_{13} \overset{\text{LO}}{=} \beta_0\frac{1}{2}(2\phi_4 - \phi_2),$$
(B.4)

$$\sqrt{8t_0}\left(\frac{m_1 - m_3}{2}\right)\phi_{12} - \phi_{13} \overset{\text{LO}}{=} \beta_0\left(\frac{3}{2}\phi_2 - \phi_4\right),$$
(B.5)

$$\sqrt{8t_0}\text{Tr}[M_q] = \sqrt{8t_0}[2m_1 + m_3] \overset{\text{LO}}{=} 2\beta_0\phi_4,$$
(B.6)

$$8t_0\text{Tr}[M_q^2] = 8t_0[2m_1^2 + m_3^2] = 2\phi_{12}^2 + \phi_h^2$$
$$\overset{\text{LO}}{=} 2\beta_0^2[3\phi_2^2 + 2\phi_4^2 - 4\phi_4\phi_2].$$
(B.7)

Inserting the above LO expressions in Eq. (B.2) we obtain, after some straightforward algebra, that for two light quarks the discretisation function has the form

$$f_{12}(a, \phi_2) \equiv f(a, m_1, 0, \text{Tr}[M_q])$$
$$= \frac{a^2}{8t_0}[C_0 + C_1\phi_2 + C_2\phi_2^2] + O(a^3), \quad \text{(B.8)}$$

where C_k (with $k = 0, 1, \ldots$) depend on the constants β_0, ϕ_4, and the coefficients c_l, suitably rescaled by factors of $8t_0$ (with $l = 0, 1, \ldots$). Similarly, for the heavier and a light quark we obtain

$$f_{13}(a, \phi_2) \equiv f(a, m_1, m_3, \text{Tr}[M_q])$$
$$= \frac{a^2}{8t_0}[\tilde{C}_0 + \tilde{C}_1\phi_2 + \tilde{C}_2\phi_2^2] + O(a^3). \quad \text{(B.9)}$$

Note that, although in general coefficients C_n and $\tilde{C}_n$ are not the same, in the case of $m_3 = m_1$ (symmetric point) $f_{13} = f_{12}$ trivially.

The very fact that we have used LO χPT to obtain the last two expressions (cf. Eqs. (B.3)–(B.7)) allows us to drop

$O(a^2\phi_2^2)$ contributions of f_{12} and f_{13}. Moreover, standard power-counting schemes in Wilson χPT [65,66] suggest that terms of $O(a^2)$ enter at the same order as $O(m_\pi^2)$, which would imply that the terms of $O(a^2\phi_2)$ should also be dropped. We will nevertheless keep this term and explore its impact.

For the ratio of ϕ_{12} and ϕ_{13} we have that

$$\frac{\phi_{12}}{\phi_{13}} = \frac{\phi_{12}^{\text{cont}} + f_{12}(a, \phi_2)}{\phi_{13}^{\text{cont}} + f_{13}(a, \phi_2)}$$
$$= \frac{\phi_{12}^{\text{cont}}}{\phi_{13}^{\text{cont}}} + \frac{f_{12}}{\phi_{13}^{\text{cont}}} - \frac{f_{13}\phi_{12}^{\text{cont}}}{(\phi_{13}^{\text{cont}})^2} + \cdots . \quad (B.10)$$

We write the discretisation functions f_{12} and f_{13} as in Eqs. (B.8) and (B.9) and then express coefficients C_0, C_1, $\tilde{C}_0$, $\tilde{C}_1, \ldots$ in terms of the original coefficients c_i of Eq. (B.2). After some algebra we end up with

$$\frac{\phi_{12}}{\phi_{13}} = \frac{\phi_{12}^{\text{cont}}}{\phi_{13}^{\text{cont}}} + \frac{a^2}{8t_0}\frac{2\phi_4 - 3\phi_2}{(2\phi_4 - \phi_2)^2}\Big[D_0 + D_1\phi_2 + D_2\phi_2^2\Big]$$
$$+ O(a^3). \quad (B.11)$$

The coefficients $D_0, D_1, D_2, \ldots$ depend on the c_i's. The factor $2\phi_4 - \phi_2$ in the discretisation term vanishes at the symmetric point $\phi_2 = 2\phi_4/3$. This confirms that at the symmetric point the ratio ϕ_{12}/ϕ_{13} is 1 by construction, for any lattice spacing. In analogy to the arguments exposed above for f_{12} and f_{13}, we drop the $D_2\phi_2^2$ term in our fits. Moreover, the variation of the denominator $(2\phi_4 - \phi_2)^2$ is relatively mild, ranging between ~ 2 and ~ 4.6 as ϕ_2 varies between ~ 0.1 and ~ 0.8 in our simulations. To simplify matters, we reabsorb this $O(1)$ term in re-definitions of D_0 and D_1.

Finally, for the combination of Eq. (A.14), we straightforwardly parametrise the discretisation errors in a way analogous to f_{12} and f_{13}; see Eq. (4.5).

References

1. K. Symanzik, Infrared singularities and small distance behavior analysis. Commun. Math. Phys. **34**, 7 (1973). https://doi.org/10.1007/BF01646540
2. T. Appelquist, J. Carazzone, Infrared singularities and massive fields. Phys. Rev. D **11**, 2856 (1975). https://doi.org/10.1103/PhysRevD.11.2856
3. B.A. Ovrut, H.J. Schnitzer, Gauge theories with minimal subtraction and the decoupling theorem. Nucl. Phys. B **179**, 381 (1981). https://doi.org/10.1016/0550-3213(81)90011-0
4. W. Bernreuther, W. Wetzel, Decoupling of heavy quarks in the minimal subtraction scheme. Nucl. Phys. B **197**, 228 (1982). https://doi.org/10.1016/0550-3213(82)90288-7, https://doi.org/10.1016/S0550-3213(97)00811-0
5. S. Aoki et al. (Flavour Lattice Averaging Group), FLAG review (2019). arXiv:1902.08191
6. ALPHA collaboration, F. Knechtli, T. Korzec, B. Leder, G. Moir, Power corrections from decoupling of the charm quark. Phys. Lett. B **774**, 649 (2017). https://doi.org/10.1016/j.physletb.2017.10.025. arXiv:1706.04982
7. M. Bruno et al., Simulation of QCD with $N_f = 2 + 1$ flavors of nonperturbatively improved Wilson fermions. JHEP **02**, 043 (2015). https://doi.org/10.1007/JHEP02(2015)043. arXiv:1411.3982
8. M. Bruno, T. Korzec, S. Schaefer, Setting the scale for the CLS $2 + 1$ flavor ensembles. Phys. Rev. D **95**, 074504 (2017). https://doi.org/10.1103/PhysRevD.95.074504. arXiv:1608.08900
9. K.G. Wilson, Confinement of quarks. Phys. Rev. D **10**, 2445 (1974)
10. B. Sheikholeslami, R. Wohlert, Improved continuum limit lattice action for QCD with Wilson fermions. Nucl. Phys. B **259**, 572 (1985). https://doi.org/10.1016/0550-3213(85)90002-1
11. M. Lüscher, P. Weisz, On-shell improved lattice gauge theories. Commun. Math. Phys. **97**, 59 (1985). https://doi.org/10.1007/BF01206178
12. J. Bulava, S. Schaefer, Improvement of $N_f = 3$ lattice QCD with Wilson fermions and tree-level improved gauge action. Nucl. Phys. B **874**, 188 (2013). https://doi.org/10.1016/j.nuclphysb.2013.05.019. arXiv:1304.7093
13. ALPHA collaboration, I. Campos, P. Fritzsch, C. Pena, D. Preti, A. Ramos, A. Vladikas, Non-perturbative quark mass renormalisation and running in $N_f = 3$ QCD. Eur. Phys. J. C **78**, 387 (2018). https://doi.org/10.1140/epjc/s10052-018-5870-5. arXiv:1802.05243
14. ALPHA collaboration, J. Bulava, M. Della Morte, J. Heitger, C. Wittemeier, Non-perturbative improvement of the axial current in $N_f = 3$ lattice QCD with Wilson fermions and tree-level improved gauge action. Nucl. Phys. B **896**, 555 (2015). https://doi.org/10.1016/j.nuclphysb.2015.05.003. arXiv:1502.04999
15. J. Bulava, M. Della Morte, J. Heitger, C. Wittemeier, Nonperturbative renormalization of the axial current in $N_f = 3$ lattice QCD with Wilson fermions and a tree-level improved gauge action. Phys. Rev. D **93**, 114513 (2016). https://doi.org/10.1103/PhysRevD.93.114513. arXiv:1604.05827
16. ALPHA collaboration, G.M. de Divitiis, M. Firrotta, J. Heitger, C.C. Köster, A. Vladikas, Non-perturbative determination of improvement b-coefficients in $N_f = 3$. EPJ Web Conf. **175**, 10008 (2018). https://doi.org/10.1051/epjconf/201817510008. arXiv:1710.07020
17. S. Aoki et al., Review of lattice results concerning low-energy particle physics. Eur. Phys. J. C **77**, 112 (2017). https://doi.org/10.1140/epjc/s10052-016-4509-7. arXiv:1607.00299
18. M. Lüscher, Properties and uses of the Wilson flow in lattice QCD. JHEP **08**, 071 (2010). https://doi.org/10.1007/JHEP08(2010)071, https://doi.org/10.1007/JHEP03(2014)092. arXiv:1006.4518
19. R. Sommer, Scale setting in lattice QCD. PoS **LATTICE2013**, 015 (2014). https://doi.org/10.22323/1.187.0015. arXiv:1401.3270
20. O. Bär, M. Golterman, Chiral perturbation theory for gradient flow observables. Phys. Rev. D **89**, 034505 (2014). https://doi.org/10.1103/PhysRevD.89.099905,https://doi.org/10.1103/PhysRevD.89.034505. arXiv:1312.4999
21. ALPHA collaboration, M. Bruno, I. Campos, J. Koponen, C. Pena, D. Preti, A. Ramos et al., Light and strange quark masses from $N_f = 2 + 1$ simulations with Wilson fermions, PoS. LATTICE2018, 220 (2019). https://doi.org/10.22323/1.334.0220. arXiv:1903.04094
22. T. Bhattacharya, R. Gupta, W. Lee, S.R. Sharpe, J.M. Wu, Improved bilinears in lattice QCD with non-degenerate quarks. Phys. Rev. D **73**, 034504 (2006). https://doi.org/10.1103/PhysRevD.73.034504. arXiv:hep-lat/0511014
23. ALPHA collaboration, G. M. Divitiis, P. Fritzsch, J. Heitger, C. C. Köster, S. Kuberski, A. Vladikas, Non-perturbative determination of improvement coefficients b_m and $b_A - b_P$ and normalisation factor $Z_m Z_P/Z_A$ with $N_f = 3$ Wilson fermions. Eur. Phys. J. C **79**, 797 (2019). https://doi.org/10.1140/epjc/s10052-019-7287-1. arXiv:1906.03445

24. M. Constantinou, M. Hadjiantonis, H. Panagopoulos, Renormalization of flavor singlet and nonsinglet fermion bilinear operators. PoS **LATTICE2014**, 298 (2014). https://doi.org/10.22323/1.214.0298. arXiv:1411.6990

25. RQCD collaboration, G. S. Bali, E. E. Scholz, J. Simeth, W. Södner, Lattice simulations with $N_f = 2+1$ improved Wilson fermions at a fixed strange quark mass. Phys. Rev. D **94**, 074501 (2016). https://doi.org/10.1103/PhysRevD.94.074501. arXiv:1606.09039

26. P. Korcyl, G.S. Bali, Non-perturbative determination of improvement coefficients using coordinate space correlators in $N_f = 2+1$ lattice QCD. Phys. Rev. D **95**, 014505 (2017). https://doi.org/10.1103/PhysRevD.95.014505. arXiv:1607.07090

27. M. Lüscher, R. Sommer, U. Wolff, P. Weisz, Computation of the running coupling in the SU(2) Yang–Mills theory. Nucl. Phys. B **389**, 247 (1993). https://doi.org/10.1016/0550-3213(93)90292-W. arXiv:hep-lat/9207010

28. M. Lüscher, R. Sommer, P. Weisz, U. Wolff, A precise determination of the running coupling in the SU(3) Yang–Mills theory. Nucl. Phys. B **413**, 481 (1994). https://doi.org/10.1016/0550-3213(94)90629-7. arXiv:hep-lat/9309005

29. ALPHA collaboration, M. Della Morte, R. Frezzotti, J. Heitger, J. Rolf, R. Sommer, U. Wolff, Computation of the strong coupling in QCD with two dynamical flavors. Nucl. Phys. B **713**, 378 (2005). https://doi.org/10.1016/j.nuclphysb.2005.02.013. arXiv:hep-lat/0411025

30. S. Capitani, M. Lüscher, R. Sommer, H. Wittig, Non-perturbative quark mass renormalization in quenched lattice QCD. Nucl. Phys. B **544**, 669 (1999). https://doi.org/10.1016/S0550-3213(00)00163-2,https://doi.org/10.1016/S0550-3213(98)00857-8. arXiv:hep-lat/9810063

31. ALPHA collaboration, M. Della Morte, R. Hoffmann, F. Knechtli, J. Rolf, R. Sommer, I. Wetzorke et al., Non-perturbative quark mass renormalization in two-flavor QCD. Nucl. Phys. B **729**, 117 (2005). https://doi.org/10.1016/j.nuclphysb.2005.09.028. arXiv:hep-lat/0507035

32. P. Fritzsch, A. Ramos, The gradient flow coupling in the Schrödinger functional. JHEP **10**, 008 (2013). https://doi.org/10.1007/JHEP10(2013)008. arXiv:1301.4388

33. ALPHA collaboration, M. Dalla Brida, P. Fritzsch, T. Korzec, A. Ramos, S. Sint, R. Sommer, Slow running of the gradient flow coupling from 200 MeV to 4 GeV in $N_f = 3$ QCD. Phys. Rev. D **95**, 014507 (2017). https://doi.org/10.1103/PhysRevD.95.014507. arXiv:1607.06423

34. ALPHA collaboration, M. Bruno, M. Dalla Brida, P. Fritzsch, T. Korzec, A. Ramos, S. Schaefer et al., QCD coupling from a nonperturbative determination of the three-flavor Λ parameter. Phys. Rev. Lett. **119**, 102001 (2017). https://doi.org/10.1103/PhysRevLett.119.102001. arXiv:1706.03821

35. ALPHA collaboration, M. Dalla Brida, P. Fritzsch, T. Korzec, A. Ramos, S. Sint, R. Sommer, Determination of the QCD Λ-parameter and the accuracy of perturbation theory at high energies. Phys. Rev. Lett. **117**, 182001 (2016). https://doi.org/10.1103/PhysRevLett.117.182001. arXiv:1604.06193

36. ALPHA collaboration, M. Dalla Brida, P. Fritzsch, T. Korzec, A. Ramos, S. Sint, R. Sommer, A non-perturbative exploration of the high energy regime in $N_f = 3$ QCD. Eur. Phys. J. C **78**, 372 (2018). https://doi.org/10.1140/epjc/s10052-018-5838-5. arXiv:1803.10230

37. M. Lüscher, R. Narayanan, P. Weisz, U. Wolff, The Schrödinger functional: a renormalizable probe for nonabelian gauge theories. Nucl. Phys. B **384**, 168 (1992). https://doi.org/10.1016/0550-3213(92)90466-O. arXiv:hep-lat/9207009

38. S. Sint, On the Schrödinger functional in QCD. Nucl. Phys. B **421**, 135 (1994). https://doi.org/10.1016/0550-3213(94)90228-3. arXiv:hep-lat/9312079

39. J. Gasser, H. Leutwyler, Quark masses. Phys. Rep. **87**, 77 (1982). https://doi.org/10.1016/0370-1573(82)90035-7

40. S. Sint, R. Sommer, The running coupling from the QCD Schrödinger functional: a one loop analysis. Nucl. Phys. B **465**, 71 (1996). https://doi.org/10.1016/0550-3213(96)00020-X. arXiv:hep-lat/9508012

41. S. Sint, P. Weisz, Further results on $O(a)$ improved lattice QCD to one loop order of perturbation theory. Nucl. Phys. B **502**, 251 (1997). https://doi.org/10.1016/S0550-3213(97)00372-6. arXiv:hep-lat/9704001

42. W. Bietenholz et al., Tuning the strange quark mass in lattice simulations. Phys. Lett. B **690**, 436 (2010). https://doi.org/10.1016/j.physletb.2010.05.067. arXiv:1003.1114

43. W. Bietenholz et al., Flavour blindness and patterns of flavour symmetry breaking in lattice simulations of up, down and strange quarks. Phys. Rev. D **84**, 054509 (2011). https://doi.org/10.1103/PhysRevD.84.054509. arXiv:1102.5300

44. M. Bruno. The energy scale of the 3-flavour Λ parameter. Ph.D. thesis, Humboldt University, Berlin (2015). https://doi.org/10.18452/17516

45. M. Lüscher, S. Schaefer, Lattice QCD with open boundary conditions and twisted-mass reweighting. Comput. Phys. Commun. **184**, 519 (2013). https://doi.org/10.1016/j.cpc.2012.10.003. arXiv:1206.2809

46. M. Bruno, P. Korcyl, T. Korzec, S. Lottini, S. Schaefer, On the extraction of spectral quantities with open boundary conditions. PoS **LATTICE2014**, 089 (2014). https://doi.org/10.22323/1.214.0089. arXiv:1411.5207

47. M. Dalla Brida, T. Korzec, S. Sint, P. Vilaseca, High precision renormalization of the flavour non-singlet Noether currents in lattice QCD with Wilson quarks. Eur. Phys. J. **C79**, 23 (2019). https://doi.org/10.1140/epjc/s10052-018-6514-5. arXiv:1808.09236

48. S. Sint, The chirally rotated Schrödinger functional with Wilson fermions and automatic $O(a)$ improvement. Nucl. Phys. B **847**, 491 (2011). https://doi.org/10.1016/j.nuclphysb.2011.02.002. arXiv:1008.4857

49. S. Sint, B. Leder, Testing universality and automatic $O(a)$ improvement in massless lattice QCD with Wilson quarks. PoS **LATTICE2010**, 265 (2010). arXiv:1012.2500

50. M. Dalla Brida, S. Sint, A dynamical study of the chirally rotated Schrödinger functional in QCD. PoS **LATTICE2014**, 280 (2014). arXiv:1412.8022

51. ALPHA collaboration, U. Wolff, Monte Carlo errors with less errors. Comput. Phys. Commun. **156**, 143 (2004). https://doi.org/10.1016/S0010-4655(03)00467-3,https://doi.org/10.1016/j.cpc.2006.12.001. arXiv:hep-lat/0306017

52. ALPHA collaboration, S. Schaefer, R. Sommer, F. Virotta, Critical slowing down and error analysis in lattice QCD simulations. Nucl. Phys. B **845**, 93 (2011). https://doi.org/10.1016/j.nuclphysb.2010.11.020. arXiv:1009.5228

53. F. Virotta, Critical slowing down and error analysis of lattice QCD simulations. Ph.D. thesis, Humboldt-Universität zu Berlin, Mathematisch-Naturwissenschaftliche Fakultät I (2012)

54. ALPHA collaboration, H. Simma, R. Sommer, F. Virotta, General error computation in lattice gauge theory. ALPHA Collaboration internal notes (2012–2014)

55. A. Ramos, Automatic differentiation for error analysis of Monte Carlo data. Comput. Phys. Commun. **238**, 19 (2019). https://doi.org/10.1016/j.cpc.2018.12.020. arXiv:1809.01289

56. G. Colangelo, S. Dürr, C. Haefeli, Finite volume effects for meson masses and decay constants. Nucl. Phys. B **721**, 136 (2005). https://doi.org/10.1016/j.nuclphysb.2005.05.015. arXiv:hep-lat/0503014

57. RBC-UKQCD collaboration, C. Allton et al., Physical results from 2+1 flavor domain wall QCD and SU(2) chiral perturbation theory. Phys. Rev. D **78**, 114509 (2008). https://doi.org/10.1103/PhysRevD.78.114509. arXiv:0804.0473

58. RBC, UKQCD collaboration, T. Blum et al., Domain wall QCD with physical quark masses. Phys. Rev. D **93**, 074505 (2016). https://doi.org/10.1103/PhysRevD.93.074505. arXiv:1411.7017

59. S. Dürr, Z. Fodor, C. Hoelbling, S.D. Katz, S. Krieg, T. Kurth et al., Lattice QCD at the physical point: light quark masses. Phys. Lett. B **701**, 265 (2011). https://doi.org/10.1016/j.physletb.2011.05.053. arXiv:1011.2403

60. S. Dürr, Z. Fodor, C. Hoelbling, S. Katz, S. Krieg et al., Lattice QCD at the physical point: simulation and analysis details. JHEP **1108**, 148 (2011). https://doi.org/10.1007/JHEP08(2011)148. arXiv:1011.2711

61. C. McNeile, C.T.H. Davies, E. Follana, K. Hornbostel, G.P. Lepage, High-precision c and b masses, and QCD coupling from current–current correlators in lattice and continuum QCD. Phys. Rev. D **82**, 034512 (2010). https://doi.org/10.1103/PhysRevD.82.034512. arXiv:1004.4285

62. Y. Maezawa, P. Petreczky, Quark masses and strong coupling constant in $2 + 1$ flavor QCD. Phys. Rev. D **94**, 034507 (2016). https://doi.org/10.1103/PhysRevD.94.034507. arXiv:1606.08798

63. MILC collaboration, A. Bazavov et al., MILC results for light pseudoscalars. PoS CD09, 007 (2009). https://doi.org/10.22323/1.086.0007. arXiv:0910.2966

64. A. Bazavov et al., Staggered chiral perturbation theory in the two-flavor case and SU(2) analysis of the MILC data. PoS **LATTICE2010**, 083 (2010)

65. G. Rupak, N. Shoresh, Chiral perturbation theory for the Wilson lattice action. Phys. Rev. D **66**, 054503 (2002). https://doi.org/10.1103/PhysRevD.66.054503. arXiv:hep-lat/0201019

66. O. Bär, G. Rupak, N. Shoresh, Chiral perturbation theory at $O(a^2)$ for lattice QCD. Phys. Rev. D **70**, 034508 (2004). https://doi.org/10.1103/PhysRevD.70.034508. arXiv:hep-lat/0306021