High precision renormalization of the flavour non-singlet Noether currents in lattice QCD with Wilson quarks

Mattia Dalla Brida¹,a, Tomasz Korzec², Stefan Sint³, Pol Vilaseca⁴

¹ Dipartimento di Fisica, Università di Milano-Bicocca, and INFN, Sezione di Milano-Bicocca, Piazza della Scienza 3, 20126 Milan, Italy
² Department of Physics, Bergische Universität Wuppertal, Gaußstr. 20, 42119 Wuppertal, Germany
³ School of Mathematics, Trinity College Dublin, Dublin 2, Ireland
⁴ Dipartimento di Fisica, Università La Sapienza di Roma, and INFN, Sezione di Roma, Piazzale A. Moro 2, 00185 Rome, Italy

Received: 29 September 2018 / Accepted: 12 December 2018 / Published online: 14 January 2019

© The Author(s) 2019

Abstract We determine the non-perturbatively renormalized axial current for O(α) improved lattice QCD with Wilson quarks. Our strategy is based on the chirally rotated Schrödinger functional and can be generalized to other finite (ratios of) renormalization constants which are traditionally obtained by imposing continuum chiral Ward identities as normalization conditions. Compared to the latter we achieve an error reduction by up to one order of magnitude. Our results have already enabled the setting of the scale for the N_f = 2 + 1 CLS ensembles (Bruno et al. in Phys Rev D 95(7):074504. arXiv:1608.08900, 2017) and are thus an essential ingredient for the recent α_s determination by the ALPHA collaboration (Phys Rev Lett 119(10):102001. arXiv:1706.03821, 2017). In this paper we shortly review the strategy and present our results for both N_f = 2 and N_f = 3 lattice QCD, where we match the β-values of the CLS gauge configurations. In addition to the axial current renormalization, we also present precise results for the renormalized local vector current.

1 Introduction

Lattice regularizations with Wilson type fermions are widely used in current lattice QCD simulations. The ultra-locality of the action enables numerical efficiency and thus access to a wide range of lattice spacings and spatial volumes. Furthermore, Wilson fermions maintain the full flavour symmetry of the continuum action, as well as the discrete symmetries such as parity, charge conjugation and time reversal. Unitarity is either realized exactly, or, in the case of Symanzik-improved actions, approximately up to cutoff effects which vanish in the continuum limit.

The price to pay for these advantages consists in the explicit breaking of all chiral symmetries by the Wilson term in the action. Well-known consequences include the additive renormalization of quark masses, the mixing under renormalization of composite operators in different chiral multiplets and discretization effects linear in the lattice spacing. Furthermore, the Noether currents of chiral symmetry are no longer protected against renormalization.

The matrix elements of the axial Noether currents between pion or kaon states and the vacuum, parametrized by the decay constants f_π, f_K, e.g.

\[ \langle 0 | A_{\mu}^{ud} (0) | \pi^-, p \rangle = i p_\mu f_\pi, \]

\[ A_{\mu}^{ud} (x) = -\bar{\psi}_u (x) \gamma_\mu \gamma_5 \psi_d (x), \]

(1.1)

can be related to the measured life times of pions and kaons. The decay constants are finite in the chiral limit, can be precisely measured in numerical simulations and are ideally suited to set the scale in physical units. In order to achieve this with Wilson quarks one needs to determine the correctly renormalized axial currents,

\[ (A_R)^{f_1 f_2}_\mu (x) = Z_A A^{f_1 f_2}_\mu (x), \]

(1.2)

(with flavour indices f_1,2 = u, d, s), which are to be inserted into the matrix elements. Of course it is desirable that the error of the matrix elements is not dominated by the uncertainty of the current normalization constant.

Over the last 30 years many efforts have been made to control the consequences of explicit chiral symmetry breaking with Wilson quarks. The main strategy consists in imposing continuum chiral symmetry relations as normalization conditions at finite lattice spacing. This is usually done using chiral Ward identities, which follow from an infinitesimal chiral change of variables in the QCD path integral. An
example is the PCAC relation which determines the additive quark mass renormalization constant, as the "critical value" of the bare mass parameter, where the axial current is conserved. The fact that chiral symmetry is fully recovered only in the continuum limit implies that the choice of normalization condition matters at the cutoff level; at a fixed value of the lattice spacing the numerical results may occasionally differ substantially between any two such choices. Rather than interpreting this scatter as a systematic error, the modern approach consists in choosing a particular normalization condition and in fixing all dimensionful parameters (such as momenta or distances or background fields) in terms of a physical scale. This defines a so-called “line of constant physics” (LCP), along which the continuum limit is taken. As the lattice spacing \( a \) (or, equivalently, the bare coupling, \( g_0^2 = 6/\beta \), is varied, this defines a function \( Z_A = Z_A(\beta) \). Obviously, another choice for the LCP will result in a different function \( Z'_A(\beta) \). However, their difference will be, within errors, a smooth function of \( \beta \) which vanishes asymptotically \( \propto a \) or \( \propto a^2 \) if \( O(a) \) improvement is implemented. Hence, following a LCP ensures that cutoff effects are smooth functions of \( \beta \) and the choice of LCP becomes irrelevant in the continuum limit. Adopting this viewpoint, the relevant systematic error is therefore determined by the precision to which a chosen LCP can be followed.

In this paper we apply a recently developed method to lattice QCD with \( N_f = 2 \) and \( N_f = 3 \) flavours, matching the lattice actions chosen by the CLS initiative \([8,10]\). Our method is based on the chirally rotated Schrödinger functional (\( \chi \) SF) \([13,14]\). The theoretical foundation of this framework has been explained in \([14]\) and it has passed a number of perturbative and non-perturbative tests \([15-20]\). In contrast to the Ward identity method the axial current renormalization conditions follow from a finite chiral rotation in the massless QCD path integral with Schrödinger functional (SF) boundary conditions. The renormalization constants are then obtained from ratios of simple 2-point functions. For the axial current, this represents a significant advantage over the Ward identity method \([12,21,22]\) which involves 3- and 4-point functions. Hence, we observe a dramatic improvement in the attainable statistical precision for \( Z_A \) and some care is required to ensure that systematic errors are under control at a similar level of precision. We also discuss the normalization procedure for the local vector current. While flavour symmetry remains unbroken on the lattice with (mass-degenerate) Wilson quarks, the corresponding Noether current lives on neighbouring lattice points connected by a gauge link, so that the use of the local vector current is often more practical.

This paper is organized as follows: after a short reminder of the \( \chi \) SF correlation functions in the continuum and the normalization conditions derived from them in Sect. 2, we define in Sect. 3 a couple of different LCPs which we have followed. We then present the \( Z_A \) and \( Z_V \) determinations for lattice QCD with \( N_f = 2 \) and \( N_f = 3 \) quark flavours in Sects. 4 and 5, respectively, together with various tests we have carried out. Section 6 contains a summary of the main results of this work and some concluding remarks. Finally, the paper ends with three technical appendices: Appendix A collects the parameters and results of the simulations, Appendix B provides a detailed discussion on the systematic error estimates for our determinations, and Appendix C gathers our set of chosen fit functions which smoothly interpolate our \( Z_{A,V} \) results in \( \beta \).

The main results for \( N_f = 2 \) are collected in Table 4, while those for \( N_f = 3 \) are given in Tables 6 and 7. These results can be directly applied to data obtained from the CLS 2- and 3-flavour configurations, respectively \([8,10]\). The \( N_f = 3 \) results have, in fact, already been used, and enabled the precise CLS scale setting in Ref. \([1]\) and the accurate quark-mass renormalization of Ref. \([23]\).

## 2 Renormalization conditions from universality relations

### 2.1 The Schrödinger functional and chiral field rotations

We start by considering massless two-flavour continuum QCD. The Euclidean space-time is taken to be a hypercylinder of volume \( L^4 \) with Schrödinger functional boundary conditions \([24,25]\). In particular, in the Euclidean time direction, the quark and anti-quark fields satisfy

\[
P_+ \psi(x) \big|_{x_0=0} = 0 = \overline{\psi}(x) P_- \big|_{x_0=0},
\]

and similarly at time \( x_0 = L \) with the change \( P_\pm \to P_\mp \). The \( \text{SU}(2) \times \text{SU}(2) \) chiral and flavour symmetry leads to conserved isovector Noether currents, given by

\[
A^a_\mu(x) = \overline{\psi}(x) \gamma_\mu g^a s_a \psi(x),
\]

\[
V^a_\mu(x) = \overline{\psi}(x) \gamma_\mu g^a s_a \psi(x),
\]

with Pauli matrices \( s^a \) and isospin index \( a = 1, 2, 3 \). SF correlation functions of these currents with isovector boundary sources \( O^a_S \) and \( O^a_K \) have been defined in \([26,27]\) and are given by

\[
\langle A^a_\mu(x) O^b_S \rangle = -\delta^{ab} f_A(x_0),
\]

\[
\sum_{k=1}^{3} \langle V^a_\mu(x) O^b_K \rangle = -3 \delta^{ab} k_V(x_0).
\]

Passing from the isospin notation to fields with definite flavour assignments,

\[
A^{f_1 f_2}_\mu(x) = \overline{\psi}_{f_1}(x) \gamma_\mu g^a \psi_{f_2}(x),
\]
and similarly for the boundary sources, the correlation functions for isospin indices \( a = 1, 2 \), can be written in terms of the flavour off-diagonal fields,

\[
f_A(x_0) = -\frac{1}{2} \langle A^{au}_0 (x) O^{da}_5 \rangle,
\]

\[
k_V(x_0) = -\frac{1}{6} \sum_{k=1}^{3} \langle V^{au}_k (x) O^{du}_k \rangle.
\]

For the flavour diagonal fields in the isospin \( a = 3 \) components, e.g.

\[
A^3_\mu = \frac{1}{2} \left( A^{uu}_\mu - A^{dd}_\mu \right),
\]

one may use flavour symmetry to write

\[
f_A(x_0) = -\frac{1}{2} \langle A^{uu'}_0 (x) O^{u'u}_5 \rangle,
\]

and analogously for \( k_V \). Note that the additional up-type flavour \( u' \) is merely a notational device to indicate the fermionic contractions taken into account when applying Wick's theorem. Indeed, the sum of all the disconnected contributions for the flavour diagonal \( a = 3 \) components of SF correlation functions cancels exactly due to flavour symmetry.

We now apply a flavour diagonal chiral rotation to the fields,

\[ \psi \rightarrow \exp \left( i \frac{\alpha}{2} \gamma_s r^3 \right) \psi, \quad \bar{\psi} \rightarrow \bar{\psi} \exp \left( i \frac{\alpha}{2} \gamma_s r^3 \right). \]

Choosing the rotation angle \( \alpha = \pi/2 \) then leads to the chirally rotated SF boundary conditions,

\[ \tilde{Q}_+ \psi(x)|_{x_0=0} = 0 = \bar{\psi}(x) \tilde{Q}_+|_{x_0=0}, \]

where \( \tilde{Q}_\pm = \frac{1}{2}(1 \pm i \gamma_5) \gamma^3 i \gamma^{a_5} \). Analogous boundary conditions with reverted projectors are obtained at \( x_0 = L \).

Applying the same chiral field rotation to the axial currents,

\[ A^{uu}_\mu (x) \rightarrow -i V^{uu}_\mu (x), \quad A^{du}_\mu (x) \rightarrow A^{uu}_\mu (x), \]

one obtains either a vector current or remains with an axial current, depending on the flavour assignments. If the chiral rotation of the field variables is performed as a change of variables in the functional integral, one arrives at the formal continuum identities

\[ f_A = g_A^{uu} = -i g_\Sigma^{uu}, \quad k_V = i k_\Sigma = -i k_A^{ad}, \]

where the \( g \) - and \( l \)-functions are defined with \( \chi \) SF boundary conditions, Eq. (2.9), for instance

\[
g_A^{f_1f_2}(x_0) = -\frac{1}{2} \left( A^{f_1f_2}_0 (x) Q_A^{f_1f_2} \right|_{(\bar{q}_+)}.
\]

Here, the boundary operators \( Q_A^{f_1f_2} \) denote the chirally rotated versions of their SF counterparts, \( O_A^{f_1f_2} \). For the complete expressions and further details we refer to Ref. [20].

Regarding the case of QCD with \( N_f = 3 \) quark flavours we note that the very same steps can be taken provided the massless third quark does not take part in the chiral rotation and thus remains with standard SF boundary conditions [14].

Correlation functions are then considered for the doublet fields only, i.e. the third quark never appears as a valence quark.

2.2 Renormalization conditions

In the lattice regularized theory with Wilson type quarks, relations such as (2.11) can only be expected to hold after renormalization and up to cutoff effects. One first has to ensure that massless QCD with \( \chi \) SF boundary conditions has been correctly regularized. This is achieved by tuning the bare mass parameter \( m_0 \) to its critical value, \( m_{cr} \), where the axial current is conserved, and by tuning a boundary counterterm coefficient \( z_f \) such that physical parity is restored (cf. [20] for more details). In terms of the bare \( \chi \) SF correlation functions one may choose the two conditions,

\[ m = \left. \frac{\tilde{\partial}_0 g_u^{ad}(x_0)}{2 g_u^{ad}(x_0)} \right|_{x_0=L/2} = 0, \quad g_u^{ad}(L/2) = 0 \]

(with \( \tilde{\partial}_0 \) the symmetric lattice derivative). The division by the pseudo-scalar correlation function \( g_0^{ad} \) is not really necessary, however it is done for convenience, as it gives rise to the definition of a (bare) PCAC quark mass \( m \). Solutions to these equations define \( m_{cr} \) and \( z_f \) as functions of the bare coupling \( g_0 \), and the lattice size, \( L/a \).

Once the lattice regularization is correctly implemented, we expect e.g.

\[ Z_A Z_\xi Z_A^2 g_u^{ad}(x_0) = -i Z_V Z_\xi Z_\Sigma^2 \bar{\psi}^{ad}(x_0) + O(a^2), \]

where \( Z_\xi \) renormalizes a boundary quark or anti-quark field [14,26,28] and \( Z_A, V \) are the current normalization constants of interest. Requiring such identities to hold exactly at finite lattice spacing thus fixes the relative normalization of axial and vector current. Replacing the latter by the exactly conserved lattice vector current \( \bar{V}_\mu(x) \) (cf. Ref. [20]), for which \( Z_V = 1 \), one may obtain \( Z_A \) from either one of the ratios...
Assuming that the parameters \(x_0\) (here set to \(L/2\)), the boundary angle \(\theta\) [29], the background gauge field [24], and the precise definition for the zero mass and \(\alpha = \pi/2\) point (2.13) are fixed, we define, on an \((L/a)^4\) lattice and for a given bare coupling \(g_0^2 = 6/\beta\),

\[
Z^{u,l}_A(\beta, L/a) = R^{u,l}_A. 
\]

Finally the choice of a line of constant physics (cf. Sect. 3) defines a smooth function \((L/a)(\beta)\) such that the normalization constants become functions of \(\beta\) alone, with the difference between any two definitions vanishing smoothly with a rate \(\propto a^2\).

We also comment on the appearance of a second up-type flavour \(u'\) in (2.15). When applying the chiral rotation (2.8) to the diagonal components of \(f_A\), the disconnected diagrams are mapped to disconnected diagrams on the \(\chi\) SF side which can be shown to add up to a pure cutoff effect. Their omission is thus perfectly legitimate, even if the formulation of the renormalization conditions then has an element of partial quenching to it. The situation is comparable with the Ward identity method in two-flavour QCD [12,21], where a fictitious \(s\)-quark can be introduced to eliminate the disconnected diagrams.

Even though there exists a conserved vector current, in practice the local current is often used and then requires renormalization, too. Its renormalization constant can be obtained from,

\[
R^{u}_V = \left. \frac{g^{u(d)}_V(x_0)}{g^{u(d)}_V(x_0)} \right|_{x_0=L/2} \quad \text{or} \quad R^{l}_V = \left. \frac{\mu^{u(d)}_V(x_0)}{\mu^{u(d)}_V(x_0)} \right|_{x_0=L/2}.
\]

The same remarks as for the axial current normalization apply here, and with definite choices for all parameters we set,

\[
Z^{u,l}_V(\beta, L/a) = R^{u,l}_V. 
\]

As in the case of the axial current normalization conditions, only 2-point functions are required, which connect the boundary quark bilinear sources with the currents in the bulk. This is a major advantage over the Ward identity method [12,21] where 3- and 4-point functions are required. Hence, one expects better statistical precision from the simpler 2-point functions, and this will be confirmed below. Furthermore, as discussed in [20], the cutoff effects in the ratios are \(O(a^2)\), due to the mechanism of automatic \(O(a)\) improvement [30], even if the PCAC mass and the axial current are not \(O(a)\) improved by the counterterm \(\propto c_A [26,\] or if the vector currents are not improved by the corresponding counterterms \(\propto c_V, c_{\pi} [27,31]\).

Finally, we emphasize that similar renormalization conditions can be devised for other finite renormalization constants. An interesting example is the ratio \(Z_P/Z_S\), where \(Z_P\) and \(Z_S\) are the pseudo-scalar and scalar renormalization constant, respectively. We refer the reader to Ref. [20] for more details.

### 3 Lines of constant physics and choice of renormalization conditions

#### 3.1 General considerations

A line of constant physics requires to specify a physical (length) scale \(r\) which is kept fixed as the continuum limit is taken. A typical choice would be the pion decay constant, \(r = 1/f_{\pi}\), either at the physical quark masses or in the chiral limit. Once calculated for a range of lattice spacings, this scale defines a function \((r/a)(\beta)\) of the bare coupling \(\beta = 6/g_0^2\) which fixes the lattice spacing \(a\) in units of the chosen physical scale. Choosing the spatial lattice extent \(L/a\), at a given beta, such that

\[
(L/a)(\beta) = L/r = C_r
\]

(with a numerical constant \(C_r\)) then fixes the spatial size of the finite volume system in units of \(r\). In practice we will choose \(C_r\) such that the physical size of \(L\) will be somewhat larger than half a femto metre. Note that this equation can be read in two ways: first, if one fixes \(C_r\) and then chooses a set of \(\beta\)-values for which \(r/a\) is known, one obtains a corresponding set of values \((L/a)(\beta)\), which will not necessarily be integers. To evaluate the normalization constants at these non-integer lattice sizes then requires some interpolation of results from neighbouring integer \(L/a\)-values at the same \(\beta\). Alternatively, one could choose a set of integer \(L/a\)-values such that a choice for \(C_r\) will imply a set of \(\beta\)-values. In general this means that the data for \(r/a\) may have to be interpolated in \(\beta\). We will here choose the first option, with the set of \(\beta\)-values taken over from the large volume simulations by the CLS project [8,10].

Having set the scale one needs to ensure the correlation functions are calculated in the desired situation of massless QCD and for the chosen chirally rotated boundary conditions at \(\alpha = \pi/2\). This means one needs to tune the bare quark mass \(am_0\) and \(z_f\) as functions of \(\beta\). We will discuss this in more detail below. Finally, the correlation functions depend on kinematic parameters, such as \(x_0\) or background field parameters such as \(\theta\). We have already set \(x_0 = L/2\)
in Eqs. (2.15), (2.17) and we choose \( \theta = 0 \) and work with vanishing SU(3) background field.

With these parameter choices we will have, for a given \( r \) and \( C_r \) in Eq. (3.1), two definitions each for \( Z_A \) and \( Z_V \), namely

\[
Z_{A,V}^{g,l}(\beta, a/L) = R_{A,V}^{g,l}(\beta, a/L) |_{L/r=C_r; \ m=0, \ a=\pi/2}, \tag{3.2}
\]

either based on the \( g \)- or the \( l \)-ratios. We then expect e.g. that

\[
Z_A^g(\beta) = Z_A^l(\beta) + O(a^2), \tag{3.3}
\]

where the \( a^2 \)-effects are now expected to be smooth functions of the bare coupling.

### 3.2 Perturbative subtraction of cutoff effects

A possible refinement consists in using perturbation theory to reduce the cutoff effects perturbatively. This requires to compute the \( R \)-ratios (2.15), (2.17) perturbatively, with the exact same parameter choices as in the numerical simulations. We have performed this calculation to 1-loop order,

\[
R_{A,V}^{g,l}(g_0^2, a/L) = R_{A,V}^{g,l}(0) (a/L) + g_0^2 R_{A,V}^{g,l}(1) (a/L) + O(g_0^4), \tag{3.4}
\]

and for the chosen parameters we always find \( R_{A,V}^{g,l}(0) (a/L) = 1 \), exactly. We may then define a 1-loop correction factor,

\[
r_{A,V}^{g,l}(\beta, L/a) = \frac{1 + g_0^2 R_{A,V}^{g,l}(0)}{1 + g_0^2 R_{A,V}^{g,l}(1) (a/L)}, \tag{3.5}
\]

and results for the coefficients \( R_{A,V}^{g,l} \) are collected in Table 1, for the relevant lattice resolutions \( L/a \) and the two lattice gauge actions used by CLS. Note that the 1-loop results are \( N_t \)-independent and are thus obtained along the lines of Ref. [20], the only difference being the form of the free gluon propagator in the case of the Lüscher–Weiss gauge action [32]. As an aside we note that our results converge to the known 1-loop results \( Z_{A,V}^{l,l}(\beta) \) for an infinitely extended lattice [33–35], i.e. for \( a/L = 0 \). We also observe that the 1-loop cutoff effects for the \( l \)-definitions are generally much smaller than for the \( g \)-definitions.

For given \( L/a \) and \( \beta \), the perturbatively improved current normalization constants are now defined by

\[
Z_{A,V}^{g,l}_{\text{sub}}(\beta, L/a) = r_{A,V}^{g,l}(\beta, L/a) \times Z_{A,V}^{g,l}(\beta, L/a), \tag{3.6}
\]

and, by construction, the \( O(a^2) \) cutoff effects are subtracted to \( O(g_0^2) \), reducing them to \( O(a^2 g_0^2) \). The subtracted data for the \( Z \)-factors are then treated as before: a choice of a line of constant physics implies a set of \( \beta \)- and corresponding \( L/a \)-values to which the data must be interpolated. We will see evidence for the effectiveness of this perturbative subtraction of cutoff effects in Sects. 4 and 5.

### 3.3 Choices of LCP for \( N_t = 2 \) and \( N_t = 3 \)

In order to fix the physical scale \( r \), we choose either the kaon decay constant \( r = 1/f_K / (N_t = 2) \), or the gradient flow scale \( r = \sqrt{\beta_0} / (N_t = 3) [36]. \)

1 In order to fix the respective constants \( C_r \), we proceed as follows. Given the set of values \( \beta_i \) for \( i = 1, 2, \ldots \) (taken from CLS), we choose as a reference value \( \beta_{\text{ref}} \) either the largest or the smallest of the set. Choosing an integer lattice size \( L/a \) at the reference point \( \beta_{\text{ref}} \) now fixes \( C_r \) through

\[
C_r = \frac{(L/a)(\beta_{\text{ref}})}{(r/a)(\beta_{\text{ref}})}. \tag{3.7}
\]

Having set the scale in this way, the \( L/a \)-values at the remaining \( \beta_i \), follow from Eq. (3.1). For all our choices the physical size of our space-time extent will be \( L \approx 0.6 \text{–} 0.7 \text{fm} \). As mentioned before, except at the chosen reference value for \( \beta \) this requires interpolations of simulation results at integer \( L/a \) and our current simulation code, which is based on the openQCD package [37,38], requires that \( L/a \) is also even.

### 3.4 Topology freezing

Numerical simulations of the SF by means of standard Monte Carlo algorithms are known to suffer from the topology freezing problem (see e.g. Ref. [39] for a discussion). A possible solution is to follow the proposal of Ref. [39] and simulate the theory with open-SF boundary conditions. However, if for the given choice of parameters the problem is “mild”, one can circumvent the issue in a straightforward manner by simply imposing the renormalization conditions (2.18) and (2.13) within the trivial topological sector [40,41]. In a continuum notation, the correlation functions entering these definitions are modified as follows,

\[
g_A^{ud}(x_0) \rightarrow g_A^{ud}(x_0) = \frac{-1}{2} \frac{\langle A_0^{ud}(x) Q_{\delta Q,0}(\frac{\partial}{\partial Q}) \rangle_{Q_+}}{\langle \delta Q, 0 \rangle_{Q_+}}, \tag{3.8}
\]

and analogously in all other cases. Here, the Kronecker \( \delta \) in the functional integral selects the gauge field configurations

\footnote{The choice of the scale \( f_K \) seems somewhat circular, as its measurement requires the correctly normalized axial current. We use the results from Ref. [8] which were obtained using \( Z_A \) from a standard SF Ward identity determination.}

\footnote{For ease of notation, in the following the subscript \( Q \) is implicitly understood, and we assume that all relevant correlation functions are restricted to the \( Q = 0 \) sector.}
with topological charge $Q = 0$. Since relations based on chiral flavour symmetries should hold separately in each topological charge sector, this restriction to the trivial sector is a legitimate modification of the current renormalization conditions. It provides a viable solution to the algorithmic problem of topology freezing in cases where this problem becomes marginally relevant; this means when the fraction of topologically non-trivial gauge field configurations in the relevant ensembles is not too large. For our choices of parameters, the percentage of gauge field configurations with $Q \neq 0$ is generally below 10%, and reaches approximately 30% only in a couple of cases (cf. Tables 8, 9).

On the lattice the topological charge is not unambiguously defined. We follow Refs. [40, 41] and define the trivial topological sector as the set of gauge field configurations for which $|Q| < 0.5$, where $Q$ is discretized in terms of the Wilson flow and the clover definition of the field strength tensor [36]. The flow time $t$ is then kept fixed in physical units by requiring $\sqrt{8t} = 0.6 \times L$.

### 3.5 On the tuning of $am_0$ and $z_f$

The current normalization conditions require the $\chi$SF correlation functions at zero quark mass and with a chiral twist angle of $\pi/2$. In practice this is achieved by the simultaneous tuning of $m_0$ and $z_f$, such that Eq. (2.13) are satisfied. In general a 2-parameter tuning can be quite involved. However, here the non-perturbative $O(a)$ improvement of the action implies that the $O(a)$ uncertainty of the zero mass point is very much reduced. Since a change in $z_f$ merely re-defines the matrix element used to define the PCAC mass, a variation of $z_f$ is expected to induce a small variation of $m$ within this $O(a)$ uncertainty. The latter could in principle be reduced to $O(a^2)$ by including the $c_A$-counterterm to the axial current, but this will not be pursued here. Another important observation is that, once $m_0$ and $z_f$ are within $O(a)$ of their target values, the sensitivity of the PCAC mass to a variation of $z_f$ is reduced to order $a^2$ [cf. Appendix B, discussion after Eq. (B.8)]. One is therefore led to conclude that the PCAC mass $m$ is to a good approximation independent of $z_f$, and the tuning of $m_0$ and $z_f$ thus becomes straightforward; given a reasonable guess for $z_f$, one can first tune $m_0$, and then turn to $z_f$.

As an illustration of this situation we discuss the $N_f = 2$ case for $L/a = 8$, $\beta = 5.3$. For the tuning we considered 3 values of $\kappa = 1/(2am_0 + 8)$ and 4 values of $z_f$. We then generated around 2000 gauge field configurations separated by 10 MDUs for each of the 12 ensembles, and measured the relevant correlation functions. Figure 1 collects the results for the PCAC mass as a function of the bare quark mass, for the 4 different values of $z_f$. Within statistical errors, the PCAC mass depends linearly on $m_0$ and is essentially independent of $z_f$. A linear fit of $m$ vs. $m_0$ yields an estimate of $m_0 = m_{\text{cr}}(g_0, L/a)$ for which $m$ vanishes: these are collected in Table 2. The results are perfectly compatible with each other, and we take as our estimate for $m_{\text{cr}}$ the result of a weighted average of these four.

Once the critical bare mass is fixed, a smooth interpolation of $g_{\Lambda}^{\text{sd}}(L/2)$ in $m_0$ gives the results shown in Fig. 2. Over the chosen range, $g_{\Lambda}^{\text{sd}}(L/2)$ so interpolated is perfectly linear in $z_f$, and it is thus straightforward to determine the point $z_f^*$ where $g_{\Lambda}^{\text{sd}}(L/2)$ vanishes i.e. $z_f^* = 1.2877(5)$ in this example.

### Table 1 Finite $L/a$ estimators for the current normalization constants at 1-loop order, and our estimates for their asymptotic values; the latter agree with previous results in the literature [33–35]. All results are given for SU(3)

| $L/a$     | $R_A^{(1)}(a/L)$ | $R_A^{(1)}(a/L)$ | $R_N^{(1)}(a/L)$ | $R_N^{(1)}(a/L)$ |
|-----------|-----------------|-----------------|-----------------|-----------------|
| Wilson gauge action |
| 6         | −0.104309       | −0.116808       | −0.118728       | −0.130549       |
| 8         | −0.109076       | −0.116640       | −0.122586       | −0.129383       |
| 10        | −0.111857       | −0.116595       | −0.125088       | −0.129662       |
| 12        | −0.113308       | −0.116564       | −0.126426       | −0.129588       |
| 16        | −0.114714       | −0.116526       | −0.127747       | −0.129519       |
| $\infty$  | $Z_A^{(1)} = -0.116458(2)$ | $Z_N^{(1)} = -0.129430(2)$ |

| Lüscher–Weisz gauge action |
| 6         | −0.078368       | −0.091011       | −0.089760       | −0.101750       |
| 8         | −0.083286       | −0.090737       | −0.093819       | −0.101006       |
| 10        | −0.085958       | −0.090650       | −0.096256       | −0.100811       |
| 12        | −0.087374       | −0.090604       | −0.097578       | −0.100730       |
| 16        | −0.088756       | −0.090557       | −0.098889       | −0.100657       |
| $\infty$  | $Z_A^{(1)} = -0.090488(5)$ | $Z_N^{(1)} = -0.100567(2)$ |
Table 2 Results for $am_{ct}(g_0, L/a)$ for four different values of $z_f$, for $L/a = 8$ and $\beta = 5.3$. The weighted average of the results is also given in the last row of the table.

| $z_f$  | $am_{ct}$  | $\kappa_{ct}$  |
|--------|------------|----------------|
| 1.280  | −0.32808(13) | 0.136168(47)    |
| 1.283  | −0.32828(14) | 0.136176(51)    |
| 1.288  | −0.32808(12) | 0.136168(45)    |
| 1.293  | −0.32831(14) | 0.136177(51)    |
| Average| −0.328179(65)| 0.136172(24)    |

The estimated values of $am_{ct}$ and $z_f^*$ determined in this way turn out to be quite accurate in practice, cf. Table 8. We remark that results for $m_{ct}$ could also be taken from a different source, for instance from standard SF simulations. In this case only $z_f$ needs to be tuned. The differences to the above procedure would be $O(a)$ both in $m_{ct}$ and in $z_f^*$ which, by the mechanism of automatic $O(a)$ improvement, induce $O(a^2)$ differences in observables such as the current normalization constants [14,20]. One also expects that a precise tuning of $m_0$ is less crucial in the $\chi$SF than in the SF; the quark mass dependence of physical observables around the chiral limit is quadratic rather than linear [42].

3.6 Sources of uncertainties

Besides statistical errors directly affecting the estimators for the current normalization constants, the other source of uncertainty originates from the precision to which a line of constant physics can be followed. In principle also this latter effect is of a statistical nature, however, some elements of modellng or estimates may be involved when propagating these errors to the normalization constants, so that it is partly justified labelling these effects as systematic.

Our procedure consists of the following steps:

1. The LCP together with the set of values $\beta_i$ translates to target values $(L/a)(\beta_i)$. At each $\beta_i$ we choose lattices with even $L/a$ straddling the target values. We have anticipated that with our choices of LCPs the required lattice sizes are in the range $L/a = 8$ to $L/a = 16$. Note that all target values $(L/a)(\beta_i)$ come with statistical errors except for $\beta = \beta_{ct}$, where, by definition, $L/a$ is given as an (even) integer.

2. For given $\beta$ and $L/a$ we determine the solutions $am_0 = am_{ct}$ and $z_f = z_f^*$ of Eq. (2.13). In order to find their statistical errors which follow from the statistical uncertainties on $m$ and $g_{ud}^{ad}(L/2)$, we use estimates for the relevant derivatives,

$$
\frac{\partial mL}{\partial m_0 L}, \quad \frac{\partial mL}{\partial z_f}, \quad \frac{\partial g_{ud}^{ad}}{\partial m_0 L}, \quad \frac{\partial g_{ud}^{ad}}{\partial z_f}. \quad (3.9)
$$

3. We then determine the induced error on the $Z$-factors by estimating their derivatives with respect to the bare parameters,

$$
\frac{\partial Z_{A,N}}{\partial z_f}, \quad \frac{\partial Z_{A,N}}{\partial m_0 L}. \quad (3.10)
$$

It turns out that the derivatives (3.9) and (3.10) scale quite well with lattice size and lattice spacing, so that it is unnecessary to evaluate them for all parameter choices. Some cross checks are sufficient. The errors coming from the uncertainties in $m_0$ and $z_f$ are then combined in quadrature and added, again in quadrature, to the statistical error.
4. Where necessary, the results for \( Z_{A,V} \) at the different \( L/a \)-values and fixed \( \beta_i \) are interpolated to the target \((L/a)(\beta_i)\); and the statistical error on \((L/a)(\beta_i)\) is propagated at this point. In the case where only one value of \( L/a \) has been simulated, an estimate for the derivative

\[
\frac{\partial Z_{A,V}}{\partial (L/a)}
\]

is used to assign a systematic error due to the difference \( \Delta(L/a) \equiv L/a - (L/a)(\beta_i) \), also taking into account the statistical uncertainty on \((L/a)(\beta_i)\). The resulting systematic error is again added in quadrature.

We emphasize that all systematic effects become essentially statistical errors provided enough data is produced to estimate the derivatives required to propagate the errors to the normalization constants. In the following two sections we will present the lattice set-up and results for \( N_f = 2 \) and \( N_f = 3 \) lattice QCD. We will also come back to some of the above points.

4 Numerical results for \( N_f = 2 \) flavours

4.1 Lattice set-up and parameter choices

The CLS large volume simulations of 2-flavour QCD \cite{8} were performed using non-perturbatively \( O(a) \) improved Wilson quarks and the Wilson gauge action. The matching to CLS data via the bare coupling requires that we use the same action in the \( \chi \) SF. As for the details of the action near the time boundaries we refer to Ref. \cite{20}. In particular the counterterm coefficients \( c_1(g_0) \) and \( d_1(g_0) \) were set to their perturbative one-loop values using the results of that reference. In general, the incomplete cancellation of boundary \( O(a) \) artefacts implies some remnant \( O(a) \) effects in observables. However, for the estimators of the current normalization constants, Eqs. (2.15), (2.17), it can be shown that such \( O(a) \) effects only cause \( O(a^2) \) differences \cite{20}.

The CLS simulations were carried out for 3 values of the lattice spacing \cite{8}, corresponding to the \( \beta \)-values 5.2, 5.3 and 5.5. For future applications we have added a finer lattice spacing corresponding to \( \beta = 5.7 \). We choose the smallest CLS-value \( \beta = 5.2 \) as reference value and set

\[
L/a = 8 \quad \text{at} \quad \beta = 5.2,
\]

(4.1)

to define the starting point for the line of constant physics. We then fix the space-time volume of the \( \chi \) SF simulations in terms of the kaon decay constant, \( f_K \), evaluated at physical quark masses. Taking \( a f_K \) from Table 3 at \( \beta = 5.2 \) yields

\[
f_K L = 0.4744(74),
\]

(4.2)

and corresponds to \( L \approx 0.6 \) fm. Imposing this condition at the other \( \beta \)-values then leads to the (non-integer) \( L/a \)-values given in Table 3. The quoted errors are a combination of statistical uncertainties, propagated from Eq. (4.2) and the error on \( a f_K \) at the given \( \beta \)'s.

While the first 3 results for \( a f_K \) in Table 3 have been directly measured \cite{8} we have estimated \( a f_K \) at the fourth value, \( \beta = 5.7 \), as follows: with \( a f_K \) at \( \beta = 5.5 \) taken as starting point we used the three-loop \( \beta \)-function for the bare coupling \cite{43}, in order to determine the ratio of lattice spacings. The error is obtained by summing (in quadrature) the statistical error propagated from the result at \( \beta = 5.5 \), and a systematic error due to the use of perturbation theory. The latter is estimated as the difference between the non-perturbative result for \( a f_K \) at \( \beta = 5.5 \), and the same
perturbative procedure, applied between $\beta = 5.3$ and $\beta = 5.5$. This systematic error is about 2.7 times larger than the statistical one, and thus dominates the error on $L/\alpha$ at $\beta = 5.7$.

Except for $\beta = 5.3$, the target values $(L/\alpha)(\beta_i)$ resulting from condition (4.2), are very close to even integer values of $L/\alpha$, so that interpolations between simulations at different $L/\alpha$ can be avoided. At $\beta = 5.3$ we simulated at the three $L/\alpha$-values given in the last column of Table 3 and interpolated to the target value (see Appendix B.4 for more details). For each choice of $\beta$ and $L/\alpha$, along the lines of the discussion in Sect. 3.5, we have carried out various tuning runs covering a range of $am_0$ and $z_f$, so as to determine the parameters satisfying the conditions (2.13). The values of the tuned parameters and the results for $m$ and $g_A^{ud}(L/2)$ at these parameters are given in Table 8.

4.2 Results and error budget

In Table 4 we collect the results for $Z_A(V)$, both $g$ and $l$ definitions, at the four values of the lattice spacing. The statistics range from 1800 to 12,000 measurements depending on the ensemble, cf. Table 8. The quoted uncertainties combine the statistical and systematic errors. The statistical errors are at the level of $0.1$–$0.4/\sqrt{N}$, depending on the $Z$-factor and ensemble considered. Hence a significant contribution to the error comes from systematic uncertainties.

As discussed in Sect. 3.6, systematic errors result from uncertainties or deviations in following a chosen LCP, which correspond with statistical errors and deviations from zero in $m$ and $g_A^{ud}(L/\alpha)$, as well as uncertainties in the target lattice extent $L/\alpha$ and systematic errors arising from inter- or extrapolations from the simulated lattices sizes, if applicable. Tables 3 and 8 contain the relevant information for the case $N_f = 2$. The propagation of these uncertainties to the $Z$-factors is then performed following the steps outlined in Sect. 3.6. We have carried out some additional simulations to estimate the derivatives in Eqs. (3.9), (3.10), and some perturbative calculation to check the expected scaling of the derivatives with the lattice size. We delegate a detailed discussion to Appendix B. Here we just note that with our statistics and our rather conservative approach, the propagated uncertainties are typically larger than the statistical errors for the $R$-estimators Eqs. (2.15), (2.17) (cf. Tables 10, 11).

4.2.1 Effect of perturbative one-loop improvement

As discussed in Sect. 3.2, we have also computed the relevant $\chi$SF correlation functions in perturbation theory to order $g_0^2 = 6/\beta$. Besides consistency checks and qualitative insight the main application consists in the perturbative subtraction of cutoff effects from the data. Note that this requires to emulate the non-perturbative procedure in all details, in particular the determination of $am_{cr}$ and $z_f^\nu$ according to Eq. (2.13). The lower part of Table 4 contains the results for $Z_A(V)$ after perturbative improvement. Comparing with the unimproved results in the upper part of Table 4, one can see that the $g$-definitions are more affected, and are brought closer to the

| $\beta$ | $Z_A^k$ | $Z_A^l | Z_V^k | Z_V^l |
|---|---|---|---|---|
| 5.2 | 0.78022(55) | 0.76944(94) | 0.74673(47) | 0.73849(97) |
| 5.3 | 0.78411(61) | 0.77576(66) | 0.75220(70) | 0.74607(69) |
| 5.5 | 0.7945(13) | 0.7895(13) | 0.7663(14) | 0.7625(14) |
| 5.7 | 0.80526(97) | 0.80277(93) | 0.7800(11) | 0.7780(19) |

| $\beta$ | $Z_A^k$, sub | $Z_A^l$, sub | $Z_V^k$, sub | $Z_V^l$, sub |
|---|---|---|---|---|
| 5.2 | 0.7726(54) | 0.76963(94) | 0.73986(47) | 0.73890(97) |
| 5.3 | 0.77847(42) | 0.77591(66) | 0.74706(49) | 0.74636(70) |
| 5.5 | 0.79138(89) | 0.78971(13) | 0.7634(11) | 0.7627(14) |
| 5.7 | 0.80358(63) | 0.80283(93) | 0.77836(79) | 0.7781(10) |
corresponding $l$-definitions by the perturbative improvement (cf. also Fig. 5). In any case, the perturbative corrections are at the level of 1 per cent at most.

In conclusion, our final results for $Z_{A,V}$, either with or without perturbative improvement, turn out to be very precise and improve significantly on the standard SF determination based on chiral Ward identities (WIs) [8,21,44]. This is particularly true for the case of $Z_A$, which can be appreciated in Fig. 3 where the determinations of Table 4 are compared with those of Refs. [8,44]. In Fig. 4 we show instead a comparison for the case of $Z_V$, as obtained from the $\chi$SF, cf. Table 4, and from the standard SF (cf. Ref. [21]). We note that a relevant contribution to the error of our results comes from propagating the uncertainties associated with maintaining the condition (4.2) i.e. keeping $L$ constant (cf. Tables 10, 11). We anticipate that due to the much more accurate knowledge of the LCP in terms of $t_0$ (cf. Table 5), and by using interpolations in $L/a$ at all relevant $\beta$ values, this source of error will be essentially eliminated in the case of $N_f = 3$ (cf. Sect. 5).

4.3 Universality and automatic $O(a)$ improvement
The $\chi$SF determinations (2.15) and (2.17) are expected to be automatically $O(a)$ improved once the bare parameters $m_0$ and $z_f$ are properly tuned (cf. Sect. 2.2). This means that neither bulk nor boundary $O(a)$ counterterms are necessary to cancel $O(a)$ discretization errors in these quanti-
ties. This was confirmed to one-loop order in perturbation theory [20] and should hold generally. To this end we now look at the ratios between $Z$-factors coming from the $g$- and $l$-definitions. The expectation that these ratios converge to 1 with $O(a^2)$ corrections is indeed very well borne out by the data, cf. Fig. 5, where we also include fits to this expected behaviour. We emphasize that this is a non-trivial result: even though the bulk action is improved to match the CLS set-up, we did not $O(a)$ improve the currents entering the definitions (2.15), (2.17) and (2.13). This result thus confirms automatic $O(a)$ improvement at the non-perturbative level, and, indirectly, the universality relations between the $\chi$ SF and SF formulations. A direct way to test universality between the $\chi$ SF and SF formulations would be simply to study the continuum scaling of ratios of $Z$-factors as obtained from one and the other formulation. Provided the SF determinations are properly improved, these should also approach 1 in the continuum limit with $O(a^2)$ corrections. The large errors on the SF determinations do not allow us for a precise test of this expectation. However, the results in Figs. 3 and 4 clearly show that our determinations are in fact compatible with the SF ones within errors.

| $\beta$ | $t_0/a^2$ | $(L_1/a)(\beta)$ | $(L_2/a)(\beta)$ | $L/a$ |
|---------|-----------|------------------|------------------|-------|
| 3.40    | 2.8619(55)| 8                | 7.225(16)        | 6, 8, 10, 12 |
| 3.46    | 3.662(13) | 9.049(18)        | 8.172(22)        | 6, 8, 10, 12 |
| 3.55    | 5.166(17) | 10.748(21)       | 9.706(25)        | 8, 10, 12, 16 |
| 3.70    | 8.596(31) | 13.864(29)       | 12.521(34)       | 8, 10, 12, 16 |
| 3.85    | 14.036(57)| –                | 16               | 16    |

Fig. 5 Continuum limit of the ratios between the $l$ and $g$ definitions of $Z_A$ (left panel) and $Z_V$ (right panel) for the case of $N_f = 2$ quark-flavours; the effect of subtracting the lattice artefacts from the $Z$-factors to $O(g_2^0)$ is also shown. The dashed lines correspond to linear fits to the data, constrained to extrapolate to 1 for $a/L = 0$

5 Numerical results for $N_f = 3$ flavours

5.1 Lattice set-up and parameter choices

The CLS simulations with $N_f = 2 + 1$ flavours of non-perturbatively $O(a)$ improved Wilson fermions [45] and Lüscher–Weisz (LW) gauge action, have been carried out for 5 values of the lattice spacing, with $\beta$-values between 3.4 and 3.85 [1, 2, 10]. For completeness we note that CLS has also tried to simulate at a coarser lattice spacing corresponding to $\beta = 3.3$. However, these ensembles have been discarded for the scale determination in [1] due to very large cutoff effects observed e.g. in $t_0$ [10]. For this reason we will not consider this $\beta$-value in our study, however, we mention that it was adopted as starting point for the Ward identity determination of $Z_A$ in Ref. [22]. Given the relatively large set of lattice spacings we here consider two different LCPs, with slightly different physical extent, $L_1$ and $L_2$, which we define through the gradient flow time $t_0$ [36].
L_1/\sqrt{8t_0} = 1.6719(16) \text{ and } L_2/\sqrt{8t_0} = 1.5099(30),
\end{equation}

where the right hand sides were chosen in order to have exactly,
\begin{equation}
L_1/a = 8 \text{ at } \beta = 3.4 \text{ and } L_2/a = 16 \text{ at } \beta = 3.85,
\end{equation}

respectively. Using the result for \( t_0 \) in physical units \([1]\), Eq. (5.1) translates to \( L_1 \approx 0.7 \text{ fm} \) and \( L_2 \approx 0.6 \text{ fm} \).

In Table 5 we collect the relevant \( \beta \) values of the CLS simulations and the corresponding results for \( t_0/a^2 \) \([2]\). The latter are evaluated for equal up-, down-, and strange-quark masses, which are close to the physical average quark mass (see Refs. \([1,2]\)). Table 5 also gives the lattice sizes \( (L_1,2)/a(\beta) \) which satisfy the conditions (5.1). Compared to the \( N_f = 2 \) case (cf. Table 3), it is obvious that these \( N_f = 3 \) LCPs are much more accurately determined. In order to exploit this higher precision, we performed simulations for several \( L/a \)-values at each \( \beta \) (cf. Table 5). This allowed us to accurately interpolate the \( Z \)-factors to the target values (see Appendix B.5 for more details). Table 9 contains a summary of all simulations performed with the corresponding parameters. Due to both technical and historical reasons, we do not use the finest lattice spacing for the LCP defined in terms of \( L_1 \). Following this LCP up to \( \beta = 3.85 \) would have required simulating lattices with \( L/a = 18,20 \), which are particularly inconvenient to parallelize with our current simulation program. Note also that CLS simulations at \( \beta = 3.85 \) are ongoing and currently limited to a single ensemble, so that the LCP with \( L_1 \) may remain useful for a while. More importantly, however, the comparison between both LCPs allows us to perform additional tests on our results (cf. Sect. 5.3).

The lattice action we employ for the finite volume simulations matches the CLS action in the bulk, i.e. the Lüscher-Weisz tree-level improved gauge action and 3 flavours of non-perturbatively improved Wilson quarks \([45]\). Close to the time boundaries of the lattice there is some freedom regarding the implementation of Schrödinger functional boundary conditions. For the gauge fields we choose option B of Ref. \([32]\); we refer the reader to this reference for the details. Regarding the fermions, two quark flavours satisfy \( \chi \text{SF} \) boundary conditions (option \( r = 1 \) of \([14]\)), while the third one obeys the standard \( \chi \text{SF} \) boundary conditions \([25]\). In general, such a mixed set-up increases the number of \( O(a) \) improvement coefficients which need to be tuned in order to eliminate \( O(a) \) discretization errors from the time boundaries. As in the \( N_f = 2 \) case, however, one can show that the corresponding counterterms affect the renormalization constants \( Z_{A,V} \) only at \( O(a^2) \). For definiteness we have used the one-loop estimate \( c_1 = 1 + g_0^2 c_1^{(1)} \), where the one-loop coefficient decomposes as follows,
\begin{equation}
c_1^{(1)} = c_1^{(1,0)} + 2 \times c_1^{(1,1)}(\chi \text{SF}) + 1 \times c_1^{(1,1)}(\text{SF}).
\end{equation}

The pure gauge contribution is taken from Ref. \([46]\), the fermionic \( \chi \text{SF} \) contribution from Ref. \([20]\) and the SF contribution from Ref. \([29]\).\(^4\) Furthermore, we use the tree-level values \( d_t = 1/2 \) \([20]\) and \( c_t = 1 \) \([26]\).

5.2 Results and error budget

In Tables 6 and 7 we collect the results for \( Z_{A,V} \), corresponding to the \( L_1 \)- and \( L_2 \)-LCP, respectively. The latter we accumulated for the different ensembles ranges between 3200 and 31,000 measurements, with exact numbers given in Table 9. The corresponding statistical precision on the \( Z \)-factors is between 0.1 and 0.55\%/\(\chi\text{f}\), depending on the exact quantity and ensemble. The errors quoted in the tables then combine the statistical errors with the systematic errors originating from the uncertainties on the LPCs.

Like in the \( N_f = 2 \) case, the high statistical precision requires a careful assessment of the systematic errors in order to arrive at reliable error estimates. Tables 5 and 9 contain information on the accuracy with which the chosen LCPs are realized for our simulation parameters. Our estimates for the systematic uncertainties due to deviations from the chosen LCP were then obtained analogously to the case of \( N_f = 2 \); we refer the reader to Appendix B for the details. Here it is worth noting that, similarly to this case, the propagated uncertainties are typically larger than the statistical errors for the \( R \)-estimators, Eqs. (2.15), (2.17), cf. Table 12.

5.2.1 Effect of perturbative one-loop improvement

In the lower halves of Tables 6 and 7 we give the results for \( Z_{A,V} \) after perturbatively subtracting the lattice artefacts to one-loop order. The results have been obtained by first improving the \( Z_{A,V} \) determinations for each \( L/a \) and \( g_0 \) value, and then interpolating to the proper \( (L_1,2)/a(\beta) \) (see Appendix B.5).

Comparing the results for \( Z_{A,V} \) before and after perturbative improvement, one sees that the \( g \)-definitions are the most affected, and are brought closer to the corresponding \( l \)-definitions. All in all, the effect of the perturbative improvement is at most at the level of a couple of percent (cf. Fig. 7). Hence, not too surprisingly perhaps, the situation is very much the same as for the \( N_f = 2 \) case.

\(^4\) Even though the fermionic contributions were calculated with the Wilson gauge action, to this order the calculation only depends on the gauge background field, which is not modified when using the LW action with option B of \([32]\).
Table 6  $N_f = 3$ results for $Z_{A,V}$ using the $L_1$-LCP, both for $g$ and $l$ definitions. The lower part of the table contains the results after subtraction of the one-loop cutoff effects, cf. Eq. (3.6)

| $\beta$ | $Z_A^g$ | $Z_A^{l}$ | $Z_V^g$ | $Z_V^{l}$ |
|---------|---------|-----------|---------|---------|
| 3.40    | 0.76847(35) | 0.75446(68) | 0.72923(27) | 0.71940(70) |
| 3.46    | 0.77128(44) | 0.76018(80) | 0.73399(36) | 0.72637(77) |
| 3.55    | 0.77703(30) | 0.76879(42) | 0.74261(23) | 0.73758(44) |
| 3.70    | 0.78831(30) | 0.78327(43) | 0.75833(31) | 0.75521(44) |

| $\beta$ | $Z_{A,sub}^g$ | $Z_{A,sub}^{l}$ | $Z_{V,sub}^g$ | $Z_{V,sub}^{l}$ |
|---------|---------------|-----------------|---------------|-----------------|
| 3.40    | 0.75702(35)  | 0.75485(68)    | 0.71882(27)   | 0.72008(70)     |
| 3.46    | 0.76245(44)  | 0.76048(80)    | 0.72578(36)   | 0.72683(77)     |
| 3.55    | 0.77103(29)  | 0.76900(42)    | 0.73701(23)   | 0.73789(44)     |
| 3.70    | 0.78485(30)  | 0.78340(43)    | 0.75506(31)   | 0.75538(44)     |

Table 7  Same as Table 6 but for the $L_2$-LCP

| $\beta$ | $Z_A^g$ | $Z_A^{l}$ | $Z_V^g$ | $Z_V^{l}$ |
|---------|---------|-----------|---------|---------|
| 3.40    | 0.77129(39) | 0.75592(72) | 0.73368(30) | 0.72164(74) |
| 3.46    | 0.77371(51) | 0.76132(93) | 0.73721(42) | 0.72782(89) |
| 3.55    | 0.77856(31) | 0.76953(43) | 0.74468(25) | 0.73846(46) |
| 3.70    | 0.78925(31) | 0.78362(47) | 0.75936(33) | 0.75552(48) |
| 3.85    | 0.79985(31) | 0.79657(47) | 0.77304(33) | 0.77061(49) |

| $\beta$ | $Z_{A,sub}^g$ | $Z_{A,sub}^{l}$ | $Z_{V,sub}^g$ | $Z_{V,sub}^{l}$ |
|---------|---------------|-----------------|---------------|-----------------|
| 3.40    | 0.75741(38)  | 0.75642(72)    | 0.72120(29)   | 0.72259(74)     |
| 3.46    | 0.76288(51)  | 0.76169(93)    | 0.72732(41)   | 0.72845(89)     |
| 3.55    | 0.77115(31)  | 0.76979(43)    | 0.73780(24)   | 0.73886(46)     |
| 3.70    | 0.78499(31)  | 0.78378(47)    | 0.75534(33)   | 0.75574(48)     |
| 3.85    | 0.79734(31)  | 0.79667(47)    | 0.77065(33)   | 0.77074(49)     |

In conclusion, our final results for $Z_{A,V}$ are very precise for both LCPs. Similarly to the $N_f = 2$ case, the results for $Z_A$ are significantly more accurate than the standard SF determination based on Ward identities [22]. This can be appreciated in Fig. 6, where the results from Table 7 are displayed together with the 2 alternative definitions $Z_{A,0}$ and $Z_{A,0}^{on}$ of Ref. [22].

5.3 Universality and automatic $O(a)$ improvement

Given our estimates for $Z_{A,V}$ we can study the approach to the continuum limit of the ratio between different definitions. We begin with Fig. 7 where the ratios between the $g$- and $l$- definitions are considered for the $L_1$- and $L_2$-LCPs; both the results before and after perturbative improvement are shown. The conclusions are very much the same as for the $N_f = 2$ case. Considering the results before perturbative improvement, along both LCPs, the $g$ and $l$ definitions deviate by at most a couple of per-cent. These differences then perfectly scale with $a^2$ to zero as the continuum limit is approached. If perturbative improvement is implemented, these differences almost vanish even at the coarsest lattice spacings. There is no significant deviation from $a^2$ scaling, however, some small admixture of higher order effects cannot be excluded either.
Fig. 6  Comparison between different $Z_A$ determinations for $N_f = 3$, obtained either from WIs in the standard SF or from universality relations in the $\chi$SF. The $\chi$SF results are taken from Table 7 and the effect of the perturbative one-loop improvement is shown in the right panel. The individual SF points labelled $Z_A^{SF}$ and $Z_{A,\text{con}}^{SF}$ are taken from Ref. [22] and correspond to the definitions $Z_{A,0}$ and $Z_{A,\text{con},0}$, respectively, of that reference. The solid black line is the fit formula to $Z_A^{SF}$ also given in [22] and the dashed lines delimit the 1σ region of the fit. Note that this fit function enforces the perturbative 1-loop behaviour for $g_0^2 \rightarrow 0$.

Fig. 7  Continuum limit of the ratios between the $l$ and $g$ definitions of $Z_A$ (left panels) and $Z_V$ (right panels) for the case of $N_f = 3$ quark-flavours; the effect of subtracting the lattice artefacts from the $Z$-factors to $O(g_0^2)$ is also shown. The upper panels show the $L_1$-LCP results while the lower ones show those of the $L_2$-LCP. In all cases, the dashed lines correspond to linear fits to the data constrained to extrapolate to 1 for $a/L_{1,2} = 0$. Note that the (tiny) effect of the statistical correlation between numerator and denominator has been neglected in these ratios.
It is also interesting to consider the continuum limit of the ratio between the definitions belonging to different LCPs i.e. the $L_1$- and $L_2$-LCP. An example of such a ratio is shown in Fig. 8. Also in this case, the continuum scaling of this ratio is the one expected, and the initial difference is at the 2 per cent level. Apart from providing an important check of universality and automatic $O(a)$ improvement, these results show that considering one definition or the other for the renormalization of matrix elements of the axial and vector currents, will only introduce small $O(a^2)$ differences over the whole range of lattice spacings covered.

Finally, we look at ratios between $\chi$SF and standard SF determinations. Towards the continuum limit these should also scale like $1 + O(a^2)$, if the SF determinations are $O(a)$ improved. In Fig. 9 we show the continuum limit of the ratios between the standard SF determinations of Ref. [22] and the $\chi$SF results of table 7. We here consider both definitions of this reference, and label them as $Z_{A,con}^F = Z_{A,0}^F$ and $Z_{A,sub}^F = Z_{A,0}^{SF}$, respectively (cf. [22] for the exact definitions).

As one can see in Fig. 9, for their preferred definition, $Z_{A,con}^F$, the expected scaling is only setting in around $a^2/t_0 < 0.2$, where the SF and $\chi$SF determinations differ by a couple of per cent. At the coarsest lattice spacing, corresponding to $a^2/t_0 \approx 0.05$, the deviation from the $O(a^2)$ scaling is significant. The results for $Z_{A,con}^F$ show the largest deviation from the SF determination, which is about 6%. Considering the perturbatively improved $\chi$SF results this difference is somewhat reduced to 4–5%, but $O(a^2)$ scaling is not observed either. If we consider instead the alternative definition, $Z_{A,sub}^F$, the
deviation is reduced to about 2 per cent at the coarsest lattice spacing for $Z_A^l$, while, remarkably, the results for $Z_A^x$ and $Z_{A,0}^{\text{con}}$ are compatible within errors. In particular, the difference between this SF and both our $\chi$ SF determinations is perfectly compatible with an $O(a^2)$ effect over the whole range of lattice spacings considered. While discretization effects can only be defined with respect to some reference definition, we conclude that the alternative SF definition $Z_{A,\text{con}}^{\text{SF}}$ is, within errors, perfectly scaling with $a^2$ for $\beta \geq 3.4$ relative to all $\chi$ SF definitions, whereas the preferred definition $Z_A^{\text{SF}}$ of Ref. [22] requires much finer lattices before this expected asymptotic behaviour sets in. With hindsight, $Z_{A,\text{con}}^{\text{SF}}$ seems to be a better choice within the SF framework and also has been the preferred SF definition within the $N_f = 2$ setup of Refs. [8,44].

6 Summary and conclusions

We have used a new method [20] based on the chirally rotated Schrödinger functional [14] to obtain high precision results for the normalization constants of the Noether currents corresponding to non-singlet chiral and flavour symmetries. The matrix elements of these axial and vector currents play a crucial rôle in various contexts of hadronic physics. Our method differs from the traditional Ward identity method [11,12] in that it compares correlation functions which are related by finite chiral or flavour rotations, rather than infinitesimal ones. The major advantage compared to the Ward identity method consists in the avoidance of 3- and 4-point functions in favour of simple 2-point functions. This very significantly improves on the precision achieved in previous determinations [8,21,22,44,47]. In particular, for the case of $Z_A$, we obtain a reduction of the error by up to an order of magnitude (cf. Figs. 3, 6). The relatively poor precision obtained for $Z_A$ with the traditional Ward identity methods [8,21,22,44] (around the percent level at the coarsest lattice spacings of interest), has now become a limiting factor in several applications. For this reason, our results are in high demand and have already been used in several works [1,23,48]. In particular, the precise $N_f = 2 + 1$ scale setting from a linear combination of $f_K$ and $f_{\pi}$ in Ref. [1] crucially relies on our values of $Z_A^l$ in Table 6 and the associated uncertainty is negligible compared to the statistical error of the bare hadronic matrix elements. In turn, the precise scale setting result of [1] is entering almost all studies done with CLS gauge configurations: in particular it has enabled the precise result for the 3-flavour QCD $\Lambda$-parameter and thus $\alpha_s(m_Z)$ by the ALPHA-collaboration [2,41,49,50]. Further applications of our $Z_A$-results include the non-perturbative quark mass renormalization factor in [23] and the related determination of the light and strange quark masses [48]. Regarding the $N_f = 2$ case, the potential improvement of the scale setting in Ref. [8] due to our $Z_A$-results would be very significant, too. Tentative estimates anticipate a gain by a factor 3–6 in precision, when going from the finest to the coarsest lattice spacing [51].

In order to maximize the usefulness of our results we have chosen the same actions and the same $\beta$-values for $N_f = 2$ and $N_f = 3$ lattice QCD as used by the CLS initiative [8,10]. Hence, anyone working with CLS gauge configurations will be able to directly use our results: for $N_f = 2$ we recommend to use $Z_{A,V,\text{sub}}^l$ from Table 4, and for $N_f = 3$ we recommend using $Z_{A,V,\text{sub}}^{x}$ either of Tables 6 or 7. Although the results for $Z_{A,V,\text{sub}}^l$ are slightly less precise than those for $Z_{A,V,\text{sub}}^x$, their $L/a$-interpolations turn out to be more robust. Furthermore, the effect of the perturbative subtraction of cutoff effects is rather small and only marginally significant with current errors. While the precise choice of the $\chi$ SF results for the $Z$-factors is not crucial, it is however very important to be consistent and to not switch definitions when changing $\beta$. Only then cutoff effects are guaranteed to vanish smoothly at a rate $\propto a^2$.

Our determination of $Z_{A,V}(\beta)$ was carried out for each $\beta$-value independently, in order to avoid adding statistical correlation between physics results at different lattice spacings. However, it is straightforward to fit our $Z$-factors to a smooth function of $\beta$ (or $g_3^2$), which interpolates to any intermediate $\beta$-value. We have included a few such fits in Appendix C to our preferred definitions $Z_{A,V,\text{sub}}^l$. We also include fits which incorporate the expected perturbative behaviour to 1-loop order. However, the high precision obtained in the $\beta$-range covered by the data cannot be guaranteed outside this range. If a similar precision is required at higher $\beta$, an extension of our non-perturbative determination will be required. If $I_0/a^2$ was known for higher $\beta$-values one could extend our chosen line of constant physics covering another factor of 2 or so in the lattice spacing. The required simulations of the $\chi$ SF for lattice sizes up to $L/a = 32$ would be feasible with current resources. Going beyond this range it may be advisable to choose a different line of constant physics from a finite volume observable, or at least estimate the errors incurred by deviating from the original choice.
In applications to hadronic physics one would also like to control the $O(a)$ effects cancelled by the counterto-

terms to the currents. Close to the chiral limit, one essentially requires the counterterm coefficients $c_{A,V}$ [26,27]. We

emphasize that our method of determining the $Z$-factors does not rely on any assumptions about these counterto-
terms and can therefore be combined with results for $c_{A,V}$ from other studies, e.g. [47,52]. The same remark applies to
the $b$-coefficients multiplying $O(am)$ counterterms, which have recently been determined for the vector current in
Ref. [53].

Looking beyond direct applications of our results in the CLS context, it is quite obvious that the precision gains of this
method are generic and could be implemented with any other choice of Wilson type fermions. One would need to imple-
ment the $\chi$SF boundary conditions following Ref. [14], as well as the $\chi$SF correlation functions [20]. We also note that
the computer resources required are rather modest: in fact our largest lattice size was $16^4$; indeed, the main work for
the present results went into painstakingly following lines of constant physics and the determination of the corresponding
uncertainties and their propagation to the $Z$-factors. We have reported many technical details in the hope that any further
applications of the method will be able to benefit from our experience. One possible improvement we did not explore
was to measure the derivatives (3.9), (3.10) by computing the corresponding operator insertions into the correlation func-
tions directly on the tuned ensembles; this was done e.g. in Refs. [1,2] for the PCAC mass, $t_0$, and other observables,
and this would certainly allow one to further improve on the precision, as no assumptions on the derivatives need to be
made.

Possible future applications of the $\chi$SF include the deter-
mination of the ratio between pseudo-scalar and scalar renor-
malization constants, $Z_P/Z_S$. Advantages of the $\chi$SF are also
expected for scale-dependent problems, such as the renor-
malization of 4-quark operators, where the contamination by
$O(a)$ effects could be significantly reduced by the mecha-
nism of automatic $O(a)$ improvement [19]. Finally the $\chi$SF
offers new methods for the determination of $O(a)$ improve-
ment coefficients, which we hope to explore in the future.

Acknowledgements We would like to thank Rainer Sommer and Ste-
fano Lottini for useful discussions, and Patrick Fritzsch, Tim Harris, and
Alberto Ramos for helpful comments on the analysis of the data. We are
moreover thankful to the computer centers at ICHEC, LRZ (project id
pr84mi), and DESY-Zeuthen for the allocated computer resources and
support. The code we used for the simulations is based on the openQCD
package developed at CERN [38].

Data Availability Statement This manuscript has no associated data
or the data will not be deposited. [Authors’ comment: The paper con-
tains all necessary data to reproduce the results presented. If additional
material or the raw data files of the simulations are desired, these can
be obtained by contacting directly the authors.]

Open Access This article is distributed under the terms of the Creative
Commons Attribution 4.0 International License (http://creativecommons
ons.org/licenses/by/4.0/), which permits unrestricted use, distribution,
and reproduction in any medium, provided you give appropriate credit
to the original author(s) and the source, provide a link to the Creative
Commons license, and indicate if changes were made.
Funded by SCOAP3.

Appendix A: Simulation parameters and results

See Tables 8, 9, 10, 11 and 12.
### Table 8
Parameters of the $N_f = 2$ ensembles and corresponding results for $mL$ and $g_A^{ud}(L/2)$. The total number of measurements we collected is given by $N_{\text{ms}}$; these are spaced by 10 MDUs. In the table we also give the percentage $P_Q$ of gauge fields with $Q = 0$.

| $L/a$ | $\beta$ | $\kappa$ | $z_f$ | $mL \times 10^3$ | $g_A^{ud}(L/2) \times 10^3$ | $P_Q$ (%) | $N_{\text{ms}}$ |
|-------|---------|----------|------|----------------|---------------------------|----------|-----------|
| 8     | 5.2     | 0.1356450| 1.28300| $-0.8(1.4)$ | $-2.9(2.2)$ | 99.8     | 8002      |
| 8     | 5.3     | 0.1361712| 1.28680| $0.9(1.0)$  | $1.0(1.5)$  | 99.9     | 12,002    |
| 10    | 5.3     | 0.1362811| 1.30900| $-3.1(1.7)$ | $6.6(2.3)$  | 98.6     | 4004      |
| 12    | 5.3     | 0.1363310| 1.32280| $0.1(2.3)$  | $7.5(2.8)$  | 93.8     | 2403      |
| 12    | 5.5     | 0.1367093| 1.31120| $1.8(1.6)$  | $-2.0(1.9)$ | 99.4     | 2403      |
| 16    | 5.7     | 0.1367058| 1.30600| $-0.2(1.4)$ | $1.4(1.6)$  | 100      | 1803      |

### Table 9
Parameters of the $N_f = 3$ ensembles and corresponding results for $mL$ and $g_A^{ud}(L/2)$. The total number of measurements we collected is given by $N_{\text{ms}}$; these are spaced by 10 MDUs. In the table we also give the percentage $P_Q$ of gauge fields with $Q = 0$.

| $L/a$ | $\beta$ | $\kappa$ | $z_f$ | $mL \times 10^3$ | $g_A^{ud}(L/2) \times 10^3$ | $P_Q$ (%) | $N_{\text{ms}}$ |
|-------|---------|----------|------|----------------|---------------------------|----------|-----------|
| 6     | 3.40    | 0.1364794| 1.33331| $-0.19(99)$  | $3.1(1.6)$   | 99.9     | 31,208    |
| 8     | 3.40    | 0.1366405| 1.37100| $-0.3(1.0)$  | $-1.4(1.6)$  | 98.4     | 18,620    |
| 10    | 3.40    | 0.1367529| 1.40741| $-2.0(1.8)$  | $-2.7(2.5)$  | 92.3     | 14,416    |
| 12    | 3.40    | 0.1368158| 1.43650| $0.5(2.0)$   | $-2.0(2.8)$  | 72.9     | 21,685    |
| 6     | 3.46    | 0.1367283| 1.33580| $0.06(84)$   | $-0.9(1.5)$  | 99.95    | 31,208    |
| 8     | 3.46    | 0.1368457| 1.36250| $1.2(1.5)$   | $0.2(2.1)$   | 99.4     | 8002      |
| 10    | 3.46    | 0.1369430| 1.39170| $-3.7(1.5)$  | $3.4(2.1)$   | 96.1     | 12,441    |
| 12    | 3.46    | 0.1369731| 1.40600| $2.9(2.2)$   | $-0.5(2.8)$  | 83.1     | 5109      |
| 8     | 3.55    | 0.1370247| 1.35500| $-0.3(1.3)$  | $-0.4(1.8)$  | 99.85    | 8002      |
| 10    | 3.55    | 0.1370827| 1.37200| $-0.45(95)$  | $0.7(1.3)$   | 99.1     | 7712      |
| 12    | 3.55    | 0.1371100| 1.38320| $-0.6(1.6)$  | $2.0(2.1)$   | 96.4     | 4004      |
| 16    | 3.55    | 0.1371487| 1.40530| $-0.3(1.7)$  | $-3.3(1.9)$  | 71.2     | 6404      |
| 8     | 3.70    | 0.1370673| 1.34250| $-0.55(99)$  | $0.8(1.5)$   | 100      | 8002      |
| 10    | 3.70    | 0.1370938| 1.35164| $0.9(1.6)$   | $-3.9(2.4)$  | 99.4     | 3208      |
| 12    | 3.70    | 0.1371160| 1.35860| $-2.2(1.3)$  | $0.0(2.0)$   | 99.9     | 8000      |
| 16    | 3.70    | 0.1371370| 1.36790| $-1.0(1.2)$  | $0.7(1.3)$   | 91.9     | 4004      |
| 16    | 3.85    | 0.1369595| 1.34540| $0.6(1.0)$   | $1.4(1.1)$   | 99.1     | 4354      |

### Table 10
Renormalization constants $Z_{\Lambda,V}^{\Lambda,V}$ for the different $N_f = 2$ ensembles. The four errors refer to (cf. Appendix B): statistical, systematic coming from the uncertainty on $am_{ct}$ ($\Delta_{\text{sys}} Z_{\Lambda,V}$), systematic coming from the uncertainty on $z_f^\gamma$ ($\Delta z_f^\gamma Z_{\Lambda,V}$), and systematic coming from maintaining the condition (4.2) ($\Delta z_f^\gamma Z_{\Lambda,V}$); the latter is only relevant for the last two ensembles.

| $L/a$ | $\beta$ | $\kappa$ | $z_f$ | $Z_{\Lambda}^{\Lambda}$ | $Z_{\Lambda}^V$ | $Z_{\Lambda}^\gamma$ | $Z_{\Lambda}^{\gamma}$ |
|-------|---------|----------|------|----------------|----------------|-----------------|----------------|
| 8     | 5.2     | 0.1356450| 1.28300| 0.76944(19)(43)82 | 0.746732(28)(75)(34)94 | 0.73849(21)(46)(83) | 0.73849(21)(46)(83) |
| 8     | 5.3     | 0.1361712| 1.28680| 0.78787(16)(34)(15) | 0.77754(13)(33)(55) | 0.75661(93)(34)(62) | 0.74806(14)(36)(55) |
| 10    | 5.3     | 0.1362811| 1.30900| 0.78204(22)(76)(36) | 0.77505(18)(74)(135) | 0.75000(12)(77)(15)  | 0.74539(19)(80)(136) |
| 12    | 5.3     | 0.1363310| 1.32280| 0.77957(30)(55)(33) | 0.77331(22)(54)(124) | 0.74639(15)(56)(14)  | 0.74372(25)(58)(126) |
| 12    | 5.5     | 0.1367093| 1.31120| 0.79450(18)(57)(24)(111) | 0.78955(13)(56)(88)(79) | 0.76634(12)(58)(10)(131) | 0.76253(15)(60)(89)(86) |
| 16    | 5.7     | 0.1367058| 1.30600| 0.80526(12)(35)(16)(88) | 0.80276(86)(345)(587)(624) | 0.77995(88)(358)(617)(1039) | 0.778014(99)(371)(595)(676) |
Appendix B: Error propagation, systematic error estimates and comparison with perturbation theory

In this appendix we describe in some detail the elements required to carry out the steps 2, 3 and 4 sketched in Sect. 3.6, for the propagation of uncertainties. We first report on the numerical estimates of the various derivatives (steps 2, 3) for both $N_f = 2$ and $N_f = 3$. These estimates are obtained on smaller lattices $L/a = 8$ and $L/a = 6, 8$ respectively, and then used for all lattice sizes. We therefore also summarize the expected scaling with $L/a$ of these derivatives with maintaining the condition (4.2), are different. The mean values and corresponding statistical errors can be obtained from those of Table 10 through Eq. (3.6). For determining the systematic errors associated with the uncertainty on $am_{cf}$ and $z_f^s$ we use the same estimates for the relevant derivatives as for the results in Table 10: we thus obtain the same values. The systematic errors associated with maintaining the condition (4.2), instead, involve different derivatives for the data in Table 10 and the perturbatively improved ones (cf. Appendix B)

Table 11 Renormalization constants $Z_{A,V}^{g,l}$ for the different $N_f = 2$ ensembles. Lattice artefacts have been subtracted at O($a^3$) in perturbation theory. The four errors refer to (cf. Appendix B): statistical, systematic coming from the uncertainty on $am_{cf}$ ($\Delta \mu, Z_{A,V}$), systematic coming from the uncertainty on $z_f^s$ ($\Delta z_f^s Z_{A,V}$), and systematic coming from maintaining the condition (4.2) ($\Delta (z_f^s Z_{A,V})$); the latter is only relevant for the last two ensembles. Comparing with the results of Table 10, only the mean values, statistical errors, and errors associated

| $L/a$ | $\beta$ | $\kappa$ | $z_f$ | $Z_{A,sub}^{l}$ | $Z_{A,sub}^{g}$ | $Z_{V,sub}^{l}$ | $Z_{V,sub}^{g}$ |
|-------|--------|--------|-------|----------------|----------------|----------------|----------------|
| 8     | 5.2    | 1.356450 | 1.28300 | 0.77262(24) | 0.76963(19) | 0.739864(127) | 0.73890(21) |
| 8     | 5.3    | 1.361712 | 1.28680 | 0.78016(15) | 0.77772(13) | 0.749804(92) | 0.74846(14) |
| 10    | 5.3    | 1.362811 | 1.30900 | 0.77738(22) | 0.77519(18) | 0.74570(12) | 0.74562(19) |
| 12    | 5.3    | 1.363310 | 1.32280 | 0.77638(30) | 0.77342(22) | 0.74343(15) | 0.74342(25) |
| 12    | 5.5    | 1.367093 | 1.31120 | 0.79138(18) | 0.78965(13) | 0.76343(12) | 0.76268(15) |
| 16    | 5.7    | 1.367058 | 1.30600 | 0.80358(12) | 0.80283(86) | 0.77835(87) | 0.77809(99) |

Table 12 Renormalization constants $Z_{A,V}^{g,l}$ for the different $N_f = 3$ ensembles. The three errors refer to (cf. Appendix B): statistical, systematic coming from the uncertainty on $am_{cf}$ ($\Delta \mu, Z_{A,V}$), and systematic coming from the uncertainty on $z_f^s$ ($\Delta (z_f^s Z_{A,V})$). The corresponding results with the O($a^3$) lattice artefacts subtracted are obtained from those below by applying Eq. (3.6) to the mean values and corresponding statistical errors. The systematic uncertainties are instead left unchanged (cf. Table 11)

| $L/a$ | $\beta$ | $\kappa$ | $z_f$ | $Z_{A}^{l}$ | $Z_{A}^{g}$ | $Z_{V}^{l}$ | $Z_{V}^{g}$ |
|-------|--------|--------|-------|----------|----------|----------|----------|
| 6     | 3.40   | 0.136479 | 1.33331 | 0.77905(18) | 0.75899(15) | 0.74294(11) | 0.72534(18) |
| 8     | 3.40   | 0.136640 | 1.37100 | 0.76847(184) | 0.75446(14) | 0.72920(84) | 0.71940(15) |
| 10    | 3.40   | 0.136752 | 1.40741 | 0.76315(28) | 0.75190(20) | 0.72122(93) | 0.71552(27) |
| 12    | 3.40   | 0.136815 | 1.43650 | 0.76227(41) | 0.75037(22) | 0.71663(76) | 0.71220(33) |
| 6     | 3.46   | 0.136728 | 1.33580 | 0.784318(159) | 0.76501(13) | 0.75016(100) | 0.72359(15) |
| 8     | 3.46   | 0.136845 | 1.36250 | 0.77445(25) | 0.76154(21) | 0.73805(142) | 0.72806(22) |
| 10    | 3.46   | 0.136943 | 1.39170 | 0.76852(21) | 0.75933(18) | 0.73070(84) | 0.71255(21) |
| 12    | 3.46   | 0.136973 | 1.40600 | 0.76711(41) | 0.75760(29) | 0.72743(14) | 0.72293(27) |
| 8     | 3.55   | 0.137024 | 1.35500 | 0.78311(20) | 0.77108(16) | 0.74962(124) | 0.73993(17) |
| 10    | 3.55   | 0.137082 | 1.37200 | 0.77803(13) | 0.76936(10) | 0.74399(68) | 0.73813(11) |
| 12    | 3.55   | 0.137110 | 1.38320 | 0.77595(25) | 0.76786(21) | 0.74086(125) | 0.73692(17) |
| 16    | 3.55   | 0.137148 | 1.40530 | 0.77378(42) | 0.76676(28) | 0.73723(10) | 0.73456(20) |
| 8     | 3.70   | 0.137067 | 1.34250 | 0.79662(144) | 0.78571(12) | 0.76699(103) | 0.75716(12) |
| 10    | 3.70   | 0.137093 | 1.35164 | 0.79223(18) | 0.78454(16) | 0.76247(12) | 0.75626(18) |
| 12    | 3.70   | 0.137116 | 1.35860 | 0.789547(139) | 0.78386(11) | 0.75981(87) | 0.75576(13) |
| 16    | 3.70   | 0.137137 | 1.36790 | 0.787298(140) | 0.78283(15) | 0.75719(76) | 0.75481(99) |
| 16    | 3.85   | 0.136959 | 1.34540 | 0.799848(82) | 0.796571(88) | 0.77304(58) | 0.77060(70) |

B.1: Estimating the derivatives: $N_f = 2$

As described in Sect. 3.6, to estimate the uncertainty in $Z_{A,V}$ originating from those of $am_{cf}$ and $z_f^s$, we require the deriva-
We estimated these through dedicated simulations at $L/a = 8$ and $\beta = 5.2$, measuring the relevant quantities for several different values of $\kappa$ ($z_f$) at fixed $L$ ($\kappa$), straddling the tuned values given in Table 8. The results we obtained for the derivatives (3.9) are:

$$\frac{\partial mL}{\partial m0L} = 1.251(68), \quad \frac{\partial mL}{\partial z_f} = 0.048(73),$$
$$\frac{\partial g_{A}^{ud}}{\partial m0L} = -2.719(96), \quad \frac{\partial g_{A}^{ud} \partial mL}{\partial z_f} = -2.39(10). \quad (B.1)$$

We observe that the $z_f$-derivative of $mL$ vanishes within an uncertainty much smaller than the values of the other derivatives, so that we may safely set it to zero. These results were then used for all other ensembles and lattice sizes listed in Table 8. The uncertainty in the tuning of $mL$ and $g_{A}^{ud}(L/2)$ is thus converted to one in $m_{cr}L$ and $z_{f}^{*}$. Specifically, one has that,

$$\left( \Delta (m_{cr}L) \right) = A^{-1} \left( \Delta (mL) \right) \quad \left( \Delta z_{f}^{*} \right) = \Delta g_{A}^{ud}$$

where $A = \left( \begin{array}{cc} \frac{\partial mL}{\partial m0L} & \frac{\partial mL}{\partial z_f} \\ \frac{\partial g_{A}^{ud}}{\partial m0L} & \frac{\partial g_{A}^{ud} \partial mL}{\partial z_f} \end{array} \right)$. \quad (B.2)

and $\Delta (mL)$, $\Delta g_{A}^{ud}$, $\Delta (m_{cr}L)$ and $\Delta z_{f}^{*}$ are the uncertainties in $mL$, $g_{A}^{ud}(L/2)$, $m_{cr}L$ and $z_{f}^{*}$, respectively. For the uncertainties $\Delta (mL)$ and $\Delta g_{A}^{ud}$ we took the (absolute) values of $mL$ and $g_{A}^{ud}(L/2)$ measured on the given ensemble (cf. Table 8), plus 2 times the corresponding statistical errors. The corresponding systematic errors on the Z-factors can then be estimated as,

$$(\Delta_{m_{o}Z_{X}})^{2} + (\Delta_{z_{f}Z_{X}})^{2} = \left( \frac{\partial Z_{X}}{\partial m_{o}L} (\Delta m_{cr}L) \right)^{2}$$
$$+ \left( \frac{\partial Z_{X}}{\partial z_{f}} \right)^{2} \left( \Delta z_{f}^{*} \right)^{2}, \quad X = A, V. \quad (B.3)$$

Through the very same simulations used to determine the derivatives (B.1) we obtained,

$$\frac{\partial Z_{A}}{\partial m0L} = 0.126(11), \quad \frac{\partial Z_{A}^{l}}{\partial m0L} = -0.1266(88),$$
$$\frac{\partial Z_{V}}{\partial m0L} = 0.1376(60), \quad \frac{\partial Z_{V}^{l}}{\partial m0L} = -0.1356(98), \quad (B.4)$$

and

$$\frac{\partial Z_{A}}{\partial z_{f}} = -0.011(12), \quad \frac{\partial Z_{A}^{l}}{\partial z_{f}} = -0.1092(93),$$
$$\frac{\partial Z_{V}}{\partial z_{f}} = -0.0017(65), \quad \frac{\partial Z_{V}^{l}}{\partial z_{f}} = -0.108(11). \quad (B.5)$$

We then used these values at all other $L/a$- and $\beta$-values of Table 8. More precisely, we propagated the errors according to Eq. (B.3) using the measured absolute mean values to which we added twice the statistical errors. Taking these results at the coarsest available lattice spacing is a conservative choice which should be safe, also given that some favourable expected scaling of the derivative towards larger $L/a$ (s. below) is not taken advantage of. Complementary information obtained during the tuning runs for finding $m_{cr}$ and $z_{f}^{*}$, further corroborates this assumption. Looking at Eqs. (B.4), (B.5) it is clear that the $l$-definitions have in general larger systematic uncertainties due to their larger sensitivity to $z_{f}$. This is then reflected in the Z-factors in Tables 10 and 11 where the uncertainties are listed separately.

B.2: Estimating the derivatives: $N_{f} = 3$

For $N_{f} = 3$ we proceeded in very much the same way, except that we carried out a more complete study of $L/a = 8$ lattices at all $\beta$-values, except $\beta = 3.85$, and also included additional $L/a = 6$ lattices at $\beta = 3.4$ and 3.46. As before for the derivatives (3.9) we used their mean values and set $\partial mL/\partial z_{f} = 0$, while for the derivatives (3.10) we considered their (absolute) mean values plus twice their statistical errors. However, here we did this separately for all $\beta$-values, with $\beta = 3.7$ results also applied at $\beta = 3.85$. Table 12 contains the results for $Z_{A,V}$ for all ensembles of Table 9, including the different estimated uncertainties. As with $N_{f} = 2$, the soundness of our assumption is supported by experience gained during the tuning runs to find $m_{cr}$ and $z_{f}^{*}$.

B.3: Expected scaling with $L/a$ and perturbative calculations

B.3.1: Expected $L/a$-scaling

Using general arguments based on the Symanzik expansion and $P_{5}$ parity [20] one may obtain the expected scaling with the lattice spacing $a$ of the derivatives of the PCAC mass and $g_{A}^{ud}$ with respect to $m_{0}$ and $z_{f}$,

$$\frac{\partial g_{A}^{ud}}{\partial m_{0}L} = O(1), \quad \frac{\partial g_{A}^{ud} \partial mL}{\partial z_{f}} = O(1), \quad (B.6)$$

and

$$\frac{\partial mL}{\partial m0L} = O(1), \quad \frac{\partial mL}{\partial z_{f}} = O(a^{2}). \quad (B.7)$$

To explain how one arrives at these scaling properties we go through the example of the PCAC mass:
where we have introduced the modified PCAC mass $m'$ through the equation

$$\frac{\partial m_L}{\partial z_f} = L \frac{\partial}{\partial z_f} \left( \tilde{\partial}_0 g_{A;2;\tilde{p}}^{ad} \right) = L (m' - m) \times \frac{g_{P;\tilde{f}}^{ad}}{g_{P;\tilde{f}}^{ad}}, \quad (B.8)$$

and the notation $z_f$ indicates differentiation with respect to $z_f$ [20]. The crucial point to note is that this differentiation merely modifies the fields at the time boundaries and thus produces a different matrix element for the PCAC relation and hence the modified PCAC mass $m'$. Since the difference $m' - m$ between two PCAC masses is of $O(a)$ in general, and the $z_f$-derivative of the $P_5$-even correlation function $g_{P}^{ad}$ is $P_3$-odd and thus of $O(a)$, we arrive at $O(a^2)$ for the complete expression. It is reassuring to see that this derivative is indeed found to be small in the simulations.

Similar arguments lead to

$$\frac{\partial Z_{A;V}^{g,d}}{\partial m_0 L} = O(a), \quad \frac{\partial Z_{A;V}^{g,d}}{\partial z_f} = O(a), \quad (B.10)$$

where the derivatives are taken at fixed $\beta, z_f$ and $\beta, m_0$, respectively.

### B.3.2: Comparison with perturbation theory

Perturbation theory confirms all of these expected scaling properties. The derivatives (3.9) have only been considered to tree-level, which gives,

$$\frac{\partial m_L}{\partial m_0 L} = 1 + O(g_0^2), \quad \frac{\partial m_L}{\partial z_f} = O(g_0^2),$$

$$\frac{\partial g_{A;V}^{ad}}{\partial m_0 L} = -3 + O(g_0^2), \quad \frac{\partial g_{A;V}^{ad}}{\partial z_f} = -6 + O(g_0^2). \quad (B.11)$$

For the mass sensitivity of the axial current we have, to one-loop order,

$$\frac{\partial Z_{A}^{g}}{\partial m_0 L} \approx a \frac{L}{L} \times \left\{ \begin{array}{l} 1 - 0.47 \times g_0^2 + \cdots \text{(Wilson action)} \hfill \\ 1 - 0.43 \times g_0^2 + \cdots \text{(LW action)} \end{array} \right. \quad (B.12)$$

$$\frac{\partial Z_{A}^{l}}{\partial m_0 L} \approx a \frac{L}{L} \times \left\{ \begin{array}{l} -0.16 \times g_0^2 + \cdots \text{(LW&Wilson action)} \end{array} \right. \quad (B.13)$$

and, for the vector current,

$$\frac{\partial Z_{V}^{g}}{\partial m_0 L} \approx a \frac{L}{L} \times \left\{ \begin{array}{l} 1 - 0.48 \times g_0^2 + \cdots \text{(Wilson action)} \hfill \\ 1 - 0.44 \times g_0^2 + \cdots \text{(LW action)} \end{array} \right. \quad (B.14)$$

Note that perturbative results without explicit group factors assume gauge group SU(3) and fermions in the fundamental representation.

Here, the one-loop coefficients are the values at $L/a = 12$ and are stable within 3–10% for the range $L/a$ from 8 to 16. Incidentally, this result resolves qualitatively a puzzle posed by the non-perturbative results, Eq. (B.4), where the derivatives of the $g$- and $l$-definitions are almost the same in magnitude and opposite in sign, whereas the tree level results are 1 and 0, respectively, as first noticed in [20].

The $z_f$-sensitivity of the $Z$-factors is easily described: all $z_f$-derivatives vanish at tree level and the one-loop coefficients for the $g$-definitions are very small and vanish with a rate roughly proportional to $a$ for $Z_{A;V}^g$ and both gauge actions. On the other hand, the $l$-definitions behave as expected (s. above): very similar numbers are obtained which are, for both gauge actions, within a few percent given by

$$\frac{\partial Z_{A;V}^{l}}{\partial z_f} = -0.32 \times g_0^2 \frac{a}{L} + O(g_0^4). \quad (B.16)$$

For completeness we report the perturbative results for $am_{cr}$ and $z_f^*$ we obtain from the same computation. For the critical mass to order $g_0^2$, the known values [32,54,55] (with $C_F = 4/3$ for gauge group SU(3)),

$$am_{cr} = g_0^2 C_F \times \left\{ \begin{array}{l} -0.2025565(3), \text{ (Wilson action)} \hfill \\ -0.1509201(1), \text{ (LW action)} \end{array} \right. \quad (B.17)$$

are already reproduced to 4–5 digits on lattices with $L/a$ in the range from 8 to 16. As for $z_f^*$, we have, to order $g_0^2$ and for $a/L \rightarrow 0$,

$$z_f^* = 1 + g_0^2 C_F \times \left\{ \begin{array}{l} 0.16759(1), \text{ (Wilson action)} \hfill \\ 0.12923(5), \text{ (LW action)} \end{array} \right. \quad (B.18)$$

where the Wilson action result is from Ref. [20], whereas the LW action value is the one of $L/a = 16$ with a generous guess for the error. Also in this case the values at finite $L/a$ from 8 and 16 coincide with these numbers to 4–5 digits precision.

We observe that the quantitative comparison of non-perturbative data with bare perturbation theory to order $g_0^2$ works quite well in certain cases. For instance, for $N_f = 2$ at $\beta = 5.2$, we compare the non-perturbative value $am_{cr} = -0.3282$ to $am_{cr} = -0.3116 + O(g_0^4)$, and similarly for $z_f^* = 1.288$ we need to compare to $z_f^* = 1.258 + O(g_0^4)$. Also the non-perturbative current normalization constants themselves are reproduced by one-loop perturbation theory at the 5–10% level (compare Table 1 with Tables 4 and 6, 7). On
the other hand, the majority of the derivatives differ very significantly, for instance,

$$\frac{\partial Z_A^g}{\partial m_0 L} \bigg|_{\beta=5.2, L/a=8} = 0.126(11) \text{ vs. } 0.057 + O(g_0^4). \quad (B.19)$$

is off by a factor 2, and for the $l$-definition the comparison is between $-0.127(9)$ and $-0.023$, which differs by a factor 5. For the $z_f$-derivatives we note that perturbation theory correctly predicts the smallness of the sensitivity in the $g$-definition. Quantitatively, the perturbative $z_f$-derivatives of $Z_{A,V}^{l_i}$ at $\beta = 5.2$ and $L/a = 8$ are $-0.046 + O(g_0^4)$, to be compared with Eq. (B.5), so again we observe a difference by a factor 2.

Finally, for the interpolations in $L/a$ of the data (s. below) we are also interested in the derivatives of the $Z$-factors with respect to $x = (a/L)^2$ at fixed bare coupling. In perturbation theory we can obtain approximate results from Table 1. Expanding

$$R_{A,V}^{x} = 1 + g^2_0 \left( Z_{A,V}^{(1)} + k_{A,V}^{x} \times \frac{a^2}{L^2} + O(a^3) \right) + O(g_0^4), \quad (B.20)$$

the $x$-derivatives to order $g_0^2$ are approximately given by

$$\frac{\partial Z_{A,V}^{x}}{\partial x} \approx g^2_0 \times k_{A,V}^{x}, \quad (B.21)$$

provided higher order cutoff effects are small. We find that this is quite well satisfied, with very similar coefficients for both Wilson and LW actions, given approximately by

$$k_{A}^{x} \approx 0.45, \quad k_{V}^{x} \approx 0.43, \quad (B.22)$$

whereas $k_{A,V}^{l}$ are ca. 20–30 times smaller in magnitude, for axial and vector cases, respectively, and come with the opposite sign. Comparing this with the $\beta = 5.3$ non-perturbative data in Eq. (B.23) we see again that for the $g$-definitions these derivatives are reproduced by perturbation theory up to a factor 2, while for the $l$-definitions, perturbation theory to $O(g_0^2)$ is clearly missing the bulk of the effect. While, as expected, the non-perturbative derivatives are smaller than for the $g$-definitions, it seems that the smallness of the $O(g_0^2)$ term is an accident and higher orders are dominating at these values of $\beta$.

To conclude this comparison, perturbation theory often gives valuable qualitative information and may provide reasonable starting values for the tuning of $am_0$ and $z_f$. However, quantitatively, the agreement with non-perturbative data at lattice spacings of interest for hadronic physics hugely varies for different observables. Hence, the main practical use of perturbation theory consists in the perturbative subtraction of cutoff effects. Here, even a qualitative agreement, which may be quantitatively off by a factor 2, still means a welcome reduction of cutoff effects by 50 percent, and our data analysis does indeed point to such benefits.

Given this situation, we have refrained from using perturbative data in our estimates of the derivatives, and we have decided to ignore the favourable $O(a)$ scaling, Eq. (B.10), when applying the results obtained at $L/a = 8$ at all other $L/a$-values, too. In this respect, the tuning runs for $am_0$ and $z_f$, both for $N_f = 2$ and $N_f = 3$, provided some consistency checks which make us confident that the chosen procedure is indeed sound and rather conservative.

B.4: Interpolation in $L/a$ for $N_f = 2$

Once the uncertainties associated with the conditions (2.13) have been propagated to $Z_{A,V}$ at given $\beta$ and $L/a$, one needs to keep the LCP condition (4.2) and also take into account the corresponding uncertainties. The choice (4.2) is made such that the $L/a = 8$ results at $\beta = 5.2$ satisfy this condition by definition, while at $\beta = 5.3$ we need to interpolate the results for $L/a = 8, 10, 12$ to the target value $(L/a)(5.3) = 9.18(21)$ (cf. Table 3). We have performed a simple linear interpolation in $(a/L)^2$ which describes the data very well, similarly to the case $N_f = 3$ which will be discussed in more detail below. For $\beta = 5.5$ and 5.7, the simulated $L/a$ values are, within errors, compatible with the target values in Table 3. Systematic errors related to the condition (4.2) were then estimated by using the slope of the interpolation in $x = (a/L)^2$ at $\beta = 5.3$ also for the higher $\beta$-values. Note that we interpolate results with bare parameters $am_c$ and $z_f$ tuned at the given $\beta$- and $L/a$-values. Hence the derivative defines the sensitivity to a change of the physical size of the system. This is a pure cutoff effect of $O(a^2)$ on the $Z$-factors. Since, by the choice of the variable $x$, a factor $a^2$ is also divided out, the $x$-derivative is expected to be of $O(1)$ and we expect a smooth dependence of this derivative on $\beta$; this expectation is in fact confirmed by the results for $N_f = 3$ where the slope shows a very mild $\beta$-dependence over the whole range (cf. Table 13). As we are looking at an $O(a^2)$ effect in disguise, it is no surprise that the results depend on whether or not the cutoff effects have been subtracted perturbatively.

Without perturbative subtraction we obtained the results,

$$\frac{\partial Z_A^g}{\partial x} = 0.946(89), \quad \frac{\partial Z_A^l}{\partial x} = 0.48(16),$$

$^6$ We recall that at leading order in PT the $x$-derivatives of the $Z$-factors are of $O(g_0^2)$ (cf. Eq. (B.21)). They are thus expected to diminish, and eventually vanish, as $g_0 \to 0$. 
\[ \frac{\partial Z_V}{\partial x} = 1.177(77), \quad \frac{\partial Z^l_V}{\partial x} = 0.54(17), \quad \text{(B.23)} \]

while for the perturbatively improved ones we obtain,

\[ \frac{\partial Z_A}{\partial x} = 0.447(89), \quad \frac{\partial Z^l_A}{\partial x} = 0.49(16), \quad \text{and} \quad \frac{\partial Z^g_A}{\partial x} = 0.734(77), \quad \frac{\partial Z^l_A}{\partial x} = 0.56(17). \quad \text{(B.24)} \]

The corresponding systematic error is then simply taken to be,

\[ (\Delta x Z_X)^2 = \left( \frac{\partial Z_X}{\partial x} \right)^2 (\Delta x)^2, \quad X = A, V, \quad \text{(B.25)} \]

which is summed in quadrature to (B.3) and the statistical error from the Monte Carlo simulations. The uncertainty \( \Delta x \) was estimated as:

\[ \Delta x = |x - x(\beta)| + 2 \sigma(x(\beta)), \quad \text{(B.26)} \]

where \( x(\beta) = ((a/L)(\beta))^2 \) and \( \sigma(x(\beta)) \) is the associated error. As a further safeguard we took for the derivatives (B.23) and (B.24) the (absolute) mean value plus twice their statistical error. We observe that the \( l \)-definitions have a milder \( L/a \)-dependence than the \( g \)-based ones, unless perturbative improvement is implemented.

**B.5: Interpolation in \( L/a \) for \( N_l = 3 \)**

Once the systematic errors deriving from the tuning of \( a m_0 \) and \( z_l \) have been taken into account, the results for \( Z_{A,V} \) at different \( L/a \) and fixed \( \beta \) must be interpolated to either \( (L_1/a)(\beta) \) or \( (L_2/a)(\beta) \), depending on the LCP; for completeness the values of \( Z_{A,V} \) prior to interpolation are given in Table 12. We have considered three types of interpolation in \( x = (a/L)^2 \); these are: linear using all 4 available values of \( L/a \) (cf. Table 5), linear using only the 3 closest \( L/a \)-values to the target \( (L_{1,2}/a)(\beta) \), and quadratic using all 4 \( L/a \)-values. Given this choice, the interpolations needed for the \( L_1 \)- and \( L_2 \)-LCPs only differ for \( \beta = 3.4 \), where, by definition, \( L_1/a = 8 \) is exact, and in the case of linear interpolations with 3 points at \( \beta = 3.55 \). Recall that for \( \beta = 3.85 \) no interpolation is required, as \( L_2/a = 16 \) is exact and this \( \beta \)-value has been excluded for the \( L_1 \)-LCP.

Starting with the \( L_1 \)-LCP, the different interpolations describe the data quite well in general, particularly so for the results at the two smallest lattice spacings and for definitions based on the \( l \)-correlators. The most relevant exception is given indeed by the linear interpolation of \( Z_0^V \) at \( \beta = 3.46 \) using all 4 values of \( L/a \), for which we find a \( \chi^2/\text{d.o.f.} \approx 2.2 \). It should be noted, however, that the \( \chi^2 \)-criterion does not come with the usual probability interpretation due to the errors being dominated by systematics. In any case, the interpolated values are generally compatible at the 1\( \sigma \) level.

Considering the perturbatively improved data, the quality of the interpolations is generally improved, and all fits have excellent \( \chi^2 \). The beneficial effect of the perturbative improvement can be appreciated by comparing Figs. 10 and 11, where the \( Z_{A,V} \) interpolations at \( \beta = 3.46 \) are shown for the cases before and after perturbative improvement, respectively. This example also illustrates the general feature that, before perturbative improvement, the \( l \)-definitions have a significantly milder \( L/a \)- and hence \( x \)-dependence. In addition, it is interesting to note that the \( x \)-dependence of the \( Z \)-factors does not change significantly over the range of \( \beta \) considered, but seems in general to diminish, as expected, as \( \beta \to \infty \) (cf. Table 13). Based on these observations, we take as our final estimates for the \( Z \)-factors the results of the quadratic fits, which have the largest errors.

Regarding the \( L_2 \)-LCP, the situation is more complicated due to the fact that we need to interpolate the data at the coarsest lattice spacing, \( \beta = 3.4 \). The quality of the interpolations is still good in general, but there are a few significant exceptions. We note that all these cases involve \( g \)-definitions: indeed, we obtain a pretty large \( \chi^2/\text{d.o.f.} \) for the linear fits of \( Z_0^V \) at \( \beta = 3.4 \) and 3.46, around 7.8 and 5 respectively. Also the quadratic fit for \( Z_A^V \) at \( \beta = 3.4 \) has a large \( \chi^2/\text{d.o.f.} \approx 1.8 \). While in this case, however, the results of the interpolation are compatible with those of the linear fits within less than one standard deviation, in the case of \( Z_0^V \) at \( \beta = 3.4 \) the discrepancy between the linear and quadratic interpolations is close to 3 standard deviations. The situation definitely improves when the perturbatively improved data are considered. In this case, with the exception of the linear interpolations with 4 points of \( Z_0^V,A,\text{sub} \) at \( \beta = 3.4 \), all fits have very good \( \chi^2 \), and give compatible results within one standard deviation or so. As in the case of the \( L_1 \)-LCP, we take as our final estimates for \( Z_{A,V} \) the results of the quadratic fits, which have the best \( \chi^2 \)-values and the largest errors.

### Appendix C: Fit formulas for \( Z_{A,V}(g_0^2) \)

In this appendix we collect some useful fit formulas for the \( Z_{A,V} \) results, both for \( N_l = 2 \) and \( N_l = 3 \). We will focus on the data for \( Z_{A,V}^l,A,\text{sub} \), cf. the discussion in Sect. 6.

**C.1: \( N_l = 2 \)**

For \( N_l = 2 \) the final \( Z_{A,V} \) results are given in Table 4. Over the whole range of \( \beta \in [5.2, 5.7] \), the data for \( Z_{A,\text{sub}}^l \) is well described by a simple linear fit function,

\[ Z_{A,\text{sub}}^l = c_1 + c_2 g_0^2, \]
Fig. 10 $L/\alpha$-interpolations for $\beta = 3.46$ and the $L_1$-LCP. The upper two sets of points correspond to the $Z_{g,l}^\Lambda$ results while the lower two sets are the $Z_{g,l}^V$ results. The dashed lines are our preferred, quadratic, fits to the data, and the interpolation points are marked by a black vertical line.

Fig. 11 Same as Fig. 10, for the $Z$-factors with perturbative subtraction of the cutoff effects.

Table 13 Results for $\partial Z_{g,l}^{\Lambda,V}/\partial x$, where $x = (a/L)^2$, as a function of $\beta$ for $N_f = 3$ quark-flavours. The derivatives are estimated along the $L_1$-LCP from the linear fits using the 3 closest $L/\alpha$-values to the target $(L_1/\alpha)(\beta)$.

| $\beta$ | $\partial Z_{g}^\Lambda/\partial x$ | $\partial Z_{l}^\Lambda/\partial x$ | $\partial Z_{g}^V/\partial x$ | $\partial Z_{l}^V/\partial x$ |
|---------|-----------------|-----------------|-----------------|-----------------|
| 3.46    | 0.90(11)        | 0.44(21)        | 1.25(09)        | 0.56(20)        |
| 3.55    | 0.69(11)        | 0.44(15)        | 1.10(08)        | 0.56(16)        |
| 3.70    | 0.81(10)        | 0.29(16)        | 0.87(11)        | 0.24(17)        |

| $\beta$ | $\partial Z_{g,l}^\Lambda,sub/\partial x$ | $\partial Z_{g,l}^V,sub/\partial x$ |
|---------|---------------------------------|---------------------------------|
| 3.46    | 0.16(11)                        | 0.46(21)                        |
| 3.55    | -0.01(11)                       | 0.46(15)                        |
| 3.70    | 0.12(10)                        | 0.31(16)                        |
\[
\begin{align*}
c_{1.2} &= \begin{pmatrix} 1.15183 \\ -0.33176 \end{pmatrix} \\
\text{Cov} &= 10^{-3} \times \begin{pmatrix} 0.17228874 & -0.15443145 \\ -0.15443145 & 0.13858044 \end{pmatrix}, \quad (C.1)
\end{align*}
\]

which has a \( \chi^2/\text{d.o.f.} = 0.759/2 \).

Similarly, for the vector current renormalization, \( Z_{V,\text{sub}} \), a good description of the data is given by,
\[
\begin{align*}
Z_{V,\text{sub}} &= c_1 + c_2 g_0^2, \\
c_{1.2} &= \begin{pmatrix} 1.18984 \\ -0.39138 \end{pmatrix} \\
\text{Cov} &= 10^{-3} \times \begin{pmatrix} 0.19505967 & -0.17469696 \\ -0.17469696 & 0.15663400 \end{pmatrix}, \quad (C.2)
\end{align*}
\]

which has a \( \chi^2/\text{d.o.f.} = 0.866/2 \).

\section*{C.1.1: Matching with perturbation theory}

It is also interesting to consider fit functions with the correct perturbative 1-loop behaviour for \( g_0^2 \to 0 \) (cf. Sect. 3.2). In the case of \( Z_{A,\text{sub}} \) this is possible using a 2-parameter polynomial fit,
\[
\begin{align*}
Z_{A,\text{sub}}^l &= 1 - 0.116458 g_0^2 + c_1 g_0^4 + c_2 g_0^6, \\
c_{1.2} &= \begin{pmatrix} -0.015248 \\ -0.049793 \end{pmatrix} \\
\text{Cov} &= 10^{-3} \times \begin{pmatrix} 0.12545440 & -0.11198363 \\ -0.11198363 & 0.10005857 \end{pmatrix}, \quad (C.3)
\end{align*}
\]

which gives a \( \chi^2/\text{d.o.f.} = 1.519/2 \). We note that the same fit ansatz was used to fit the standard SF results of ref. [8]. Similarly, for the vector current data, \( Z_{V,\text{sub}} \), we have,
\[
\begin{align*}
Z_{V,\text{sub}}^l &= 1 - 0.129430 g_0^2 + c_1 g_0^4 + c_2 g_0^6, \\
c_{1.2} &= \begin{pmatrix} -0.005952 \\ -0.068180 \end{pmatrix} \\
\text{Cov} &= 10^{-3} \times \begin{pmatrix} 0.14221117 & -0.12684203 \\ -0.12684203 & 0.11324469 \end{pmatrix}, \quad (C.4)
\end{align*}
\]

which gives a \( \chi^2/\text{d.o.f.} = 1.866/2 \). We stress that although the latter fit functions encode the expected asymptotic behaviour far outside the \( \beta \)-range covered by the data, it is not recommended to use them for \( \beta \) values much outside this range. For \( \beta \in [5.2, 5.7] \), the two sets of fit functions agree within less than 1\( \sigma \) deviations.

\section*{C.2: \( N_f = 3 \)}

For the case \( N_f = 3 \), our final \( Z_{A,V} \) results are given in Tables 6 and 7. Having one additional \( \beta \)-value, it is natural to prefer an interpolation of the \( L_2 \)-LCP data of Table 7. The higher precision of the data compared to \( N_f = 2 \), and the availability of a fifth data point suggests to use 3-parameter fits in this case. We find that, for the whole range of \( \beta \in [3.4, 3.85] \), \( Z_{A,\text{sub}}^l \), is well described by the quadratic fit:
\[
Z_{A,\text{sub}}^l = c_1 + c_2 g_0^2 + c_3 g_0^4, \quad (C.5)
\]

with coefficients and covariance given by
\[
\begin{align*}
c_{1.2.3} &= \begin{pmatrix} 1.35510 \\ -0.501106 \\ 0.091656 \end{pmatrix} \\
\text{Cov} &= 10^{-1} \times \begin{pmatrix} 0.229571866 & -0.278151898 & 0.084105454 \\ -0.278151898 & 0.337131945 & -0.101975449 \\ 0.084105454 & -0.101975449 & 0.030856380 \end{pmatrix}, \quad (C.5)
\end{align*}
\]

and \( \chi^2/\text{d.o.f.} = 0.622/2 \).

For the vector current data, \( Z_{V,\text{sub}}^l \), we use the same fit function,
\[
Z_{V,\text{sub}}^l = c_1 + c_2 g_0^2 + c_3 g_0^4, \quad (C.6)
\]

and obtain
\[
\begin{align*}
c_{1.2.3} &= \begin{pmatrix} 1.32353 \\ -0.459016 \\ 0.066995 \end{pmatrix} \\
\text{Cov} &= 10^{-1} \times \begin{pmatrix} 0.247244906 & -0.299424391 & 0.0909493399 \\ -0.299424391 & 0.362743281 & -0.109668066 \\ 0.0909493399 & -0.109668066 & 0.033167490 \end{pmatrix}, \quad (C.5)
\end{align*}
\]

which gives a \( \chi^2/\text{d.o.f.} = 1.801/2 \).

\section*{C.2.1: Matching with perturbation theory}

Also in this case we consider fit functions with the correct perturbative 1-loop behaviour for \( g_0^2 \to 0 \) (cf. Sect. 3.2). Applying a 3-parameter polynomial fit of the form
\[
\begin{align*}
Z_{A,\text{sub}}^l &= 1 - 0.090488 g_0^2 + c_1 g_0^4 + c_2 g_0^6 + c_3 g_0^8, \quad (C.7)
\end{align*}
\]

we obtain
\[
\begin{align*}
c_{1.2.3} &= \begin{pmatrix} 0.127163 \\ -0.178785 \\ 0.051814 \end{pmatrix} \\
\text{Cov} &= 10^{-2} \times \begin{pmatrix} 0.29841165 & -0.36050066 & 0.10868891 \\ -0.36050066 & 0.43567202 & -0.13140137 \\ 0.10868891 & -0.13140137 & 0.03964605 \end{pmatrix}, \quad (C.5)
\end{align*}
\]

which gives a \( \chi^2/\text{d.o.f.} = 0.403/2 \). We have also tried various Padé fits e.g. of the type used in [22]. With these fits we experienced some technical problems with the bootstrap technique, when trying to determine the covariance matrix for the fit parameters. We therefore also tried the automatic differentiation procedure of Ref. [56] which completely solved this technical problem. It turns out, however, that the best
fit function of this type develops a singularity at a $\beta$-value slightly above 4, and therefore does not provide a smooth interpolation to the perturbative region. Given the good quality of the linear fits we did not pursue any further non-linear options.

Regarding the vector current data, $Z_{V,\text{sub}}^i$, we have,

$$Z_{V,\text{sub}}^i = 1 - 0.100567 \beta_0^2 + c_1 \beta_0^4 + c_2 \beta_0^6 + c_3 \beta_0^8, \quad (C.8)$$

with

$$c_{1,2,3} = \begin{pmatrix}
0.130134 \\
-0.182926 \\
0.051526
\end{pmatrix}$$

$$\text{Cov} = 10^{-2} \times \begin{pmatrix}
0.32342173 & -0.39055319 & 0.11769835 \\
-0.39055319 & 0.47179235 & -0.14223242 \\
0.11769835 & -0.14223242 & 0.04289459
\end{pmatrix},$$

which gives a $\chi^2/\text{d.o.f.} = 1.443/2$.

To conclude this appendix, we emphasize again that all given fits to the data are very good if used as interpolations in the range of the non-perturbative data. Using them outside this range is at the user’s own risk, even where perturbative information is used as a constraint.

References

1. M. Bruno, T. Korzec, S. Schaefer, Setting the scale for the CLS 2 + 1 flavor ensembles. Phys. Rev. D 95(7), 074504 (2017). arXiv:1608.08900

2. ALPHA Collaboration, M. Bruno, M. Dalla Brida, P. Fritzsch, T. Korzec, A. Ramos, S. Schaefer, H. Simma, S. Sint, R. Sommer, QCD Coupling from a non-perturbative determination of the three-flavor $\Lambda$ parameter. Phys. Rev. Lett. 119(10), 120001 (2017). arXiv:1706.03821

3. K.G. Wilson, Confinement of quarks. Phys. Rev. D 10, 2445–2459 (1974)

4. Hadron Spectrum Collaboration, H.-W. Lin, et al., First results from 2 + 1 dynamical quark flavors on an anisotropic lattice: light-hadron spectroscopy and setting the strange-quark mass. Phys. Rev. D 79 034502, (2009). arXiv:0810.3588

5. PACS-CS Collaboration, S. Aoki, et al., Physical point simulation in 2 + 1 flavor lattice QCD. Phys. Rev. D 81, 074503 (2010). arXiv:0911.2561

6. W. Bietenholz et al., Tuning the strange quark mass in lattice simulations. Phys. Lett. B 690, 436–441 (2010). arXiv:1003.1114

7. R. Baron et al., Light hadrons from lattice QCD with light (u, d), strange and charm dynamical quarks. JHEP 06, 111 (2010). arXiv:1004.5284

8. P. Fritzsch et al., The strange quark mass and Lambda parameter of two flavor QCD. Nucl. Phys. B 865, 397–429 (2012). arXiv:1205.5380

9. S. Borsanyi et al., Ab initio calculation of the neutron-proton mass difference. Science 347, 1452–1455 (2015). arXiv:1406.4088

10. M. Bruno et al., Simulation of QCD with $N_f = 2+1$ flavors of non-perturbatively improved Wilson fermions. JHEP 02, 043 (2015). arXiv:1411.3982

11. M. Bochicchio, L. Maiani, G. Martinelli, G.C. Rossi, M. Testa, Chiral symmetry on the lattice with Wilson fermions. Nucl. Phys. B 262, 331 (1985)

12. M. Lüscher, S. Sint, R. Sommer, H. Wittig, Nonperturbative determination of the axial current normalization constant in $O(a)$ improved lattice QCD. Nucl. Phys. B 491, 344–364 (1997). arXiv:hep-lat/9611015

13. S. Sint, The Schrödinger functional with chirally rotated boundary conditions. PoS LAT 2005, 235 (2006). arXiv:hep-lat/0511034

14. S. Sint, The chirally rotated Schrödinger functional with Wilson fermions and automatic O(a) improvement. Nucl. Phys. B 847, 491–531 (2011). arXiv:1008.4857

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

16. J.G. Lopez, K. Jansen, D. Renner, A. Shindler, A quenched study of the Schroedinger functional with chirally rotated boundary conditions: non-perturbative tuning. Nucl. Phys. B 867, 567–608 (2013). arXiv:1208.4591

17. J.G. Lopez, K. Jansen, D. Renner, A. Shindler, A quenched study of the Schrödinger functional with chirally rotated boundary conditions: applications. Nucl. Phys. B 867, 609–635 (2013). arXiv:1208.4661

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

19. P. Vilaseca, M. Dalla Brida, M. Papinutto, Perturbative renormalization of $\Delta \beta = 2$ four-fermion operators with the chirally rotated Schrödinger functional. PoS LATTICE 2015, 252 (2016)

20. M. Dalla Brida, S. Sint, P. Vilaseca, The chirally rotated Schrödinger functional: theoretical expectations and perturbative tests. JHEP 08, 102 (2016). arXiv:1603.00046

21. M. Della Morte, R. Hoffmann, F. Knechtli, R. Sommer, U. Wolff. Non-perturbative renormalization of the axial current with dynamical Wilson fermions. JHEP 0507, 007 (2005). arXiv:hep-lat/0505026

22. 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(11), 114513 (2016). arXiv:1604.05827

23. 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(5), 387 (2018). arXiv:1802.05243

24. M. Lüscher, R. Narayanan, P. Weisz, U. Wolff, The Schrödinger functional: a renormalizable probe for non-Abelian gauge theories. Nucl. Phys. B 384, 168–228 (1992). arXiv:hep-lat/9207009

25. S. Sint, On the Schrödinger functional in QCD. Nucl. Phys. B 421, 135–158 (1994). arXiv:hep-lat/9312079

26. M. Lüscher, S. Sint, R. Sommer, P. Weisz, Chiral symmetry and O(a) improvement in lattice QCD. Nucl. Phys. B 478, 365–400 (1996). arXiv:hep-lat/9605038

27. S. Sint, P. Weisz, Further results on O(a) improved lattice QCD to one loop order of perturbation theory. Nucl. Phys. B 502, 251–268 (1997). arXiv:hep-lat/9704001

28. S. Sint, One loop renormalization of the QCD Schrödinger functional. Nucl. Phys. B 451, 416–444 (1995). arXiv:hep-lat/9504005

29. S. Sint, R. Sommer, The running coupling from the QCD Schrödinger functional: a one loop analysis. Nucl. Phys. B 465, 71–98 (1996). arXiv:hep-lat/9508012

30. R. Frezzotti, G. Rossi, Chirally improving Wilson fermions. 1. O(a) improvement. JHEP 0408, 007 (2004). arXiv:hep-lat/0306014

31. ALPHA Collaboration, R. Frezzotti, S. Sint, P. Weisz, O(a) improved twisted mass lattice QCD. JHEP 07, 048 (2001). arXiv:hep-lat/0104014
32. S. Aoki, R. Frezzotti, P. Weisz, Computation of the improvement coefficient $c_{sw}$ to one loop with improved gluon actions. Nucl. Phys. B 540, 501–519 (1999). arXiv:hep-lat/9808007
33. E. Gabrielli, G. Martinelli, C. Pittori, G. Heatlie, C.T. Sachrajda, Renormalization of lattice two fermion operators with improved nearest neighbor action. Nucl. Phys. B 362, 475–486 (1991)
34. M. Göckeler, R. Horsley, E.-M. Ilgenfritz, H. Oelrich, H. Perlt, P.E.L. Rakow, G. Schierholz, A. Schiller, P. Stephenson, Perturbative renormalization of bilinear quark and gluon operators. Nucl. Phys. Proc. Suppl. 53, 896–898 (1997). arXiv:hep-lat/9608033
35. S. Aoki, K.-I. Nagai, Y. Taniguchi, A. Ukawa, Perturbative renormalization factors of bilinear quark operators for improved gluon and quark actions in lattice QCD. Phys. Rev. D 58, 074505 (1998). arXiv:hep-lat/9802034
36. M. Lüscher, Properties and uses of the Wilson flow in lattice QCD. JHEP 1008, 071 (2010). arXiv:1006.4518
37. M. Lüscher, S. Schaefer, Lattice QCD with open boundary conditions and twisted-mass reweighting. Comput. Phys. Commun. 184, 519–528 (2013). arXiv:1206.2809
38. openQCD: simulation program for lattice QCD. http://luscher.web.cern.ch/luscher/openQCD/
39. M. Lüscher, Step scaling and the Yang–Mills gradient flow. JHEP 06, 105 (2014). arXiv:1404.5930
40. P. Fritzsch, A. Ramos, F. Stollenwerk, Critical slowing down and the gradient flow coupling in the Schrödinger functional. PoS Lat- tice 2013, 461 (2013). arXiv:1311.7304
41. ALPHA Collaboration, M. Dalla Brida, P. Fritzsch, T. Korzec, A. Ramos, S. Sint, R. Sommer, Determination of the QCD $\Lambda$-parameter and the accuracy of perturbation theory at high energies. Phys. Rev. Lett. 117(18), 182001 (2016). arXiv:1604.06193
42. S. Sint, Lattice QCD with a chiral twist, in Workshop on Perspectives in Lattice QCD Nara, Japan, October 31–November 11, 2005 (2007). arXiv:hep-lat/0702008
43. A. Bode, H. Panagopoulos, The three loop beta function of QCD with the clover action. Nucl. Phys. B 625, 198–210 (2002). arXiv:hep-lat/0110211
44. M. Della Morte, R. Sommer, S. Takeda, On cutoff effects in lattice QCD from short to long distances. Phys. Lett. B 672, 407–412 (2009). arXiv:0807.1120
45. J. Bulava, S. Schaefer, Improvement of $N_f = 3$ lattice QCD with Wilson fermions and tree-level improved gauge action. Nucl. Phys. B 874, 188–197 (2013). arXiv:1304.7093
46. S. Takeda, S. Aoki, K. Ide, A perturbative determination of $O(\alpha)$ boundary improvement coefficients for the Schrödinger functional coupling at one loop with improved gauge actions. Phys. Rev. D 68, 014505 (2003). arXiv:hep-lat/0304013
47. J. Heitger, F. Joswig, A. Vladikas, C. Wittemeier, Non-perturbative determination of $c_v$, $Z_v$ and $Z_s/Z_p$ in $N_f = 3$ lattice QCD. EPJ Web Conf. 175, 10004 (2018). arXiv:1711.03924
48. J. Koponen, et al., Light and strange quark masses for $N_f = 2 + 1$ simulations with Wilson fermions, in Proceedings, 36th International Symposium on Lattice Field Theory (Lattice2018): East Lansing, Michigan, USA, July 22–28, 2018. EPJ Web Conf
49. ALPHA Collaboration, M. Dalla Brida, P. Fritzsch, T. Korzec, A. Ramos, S. Sint, R. Sommer, Determination of the QCD $\Lambda$-parameter and the accuracy of perturbation theory at high energies. Phys. Rev. Lett. 117(18), 182001 (2016). arXiv:1604.06193
50. 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(5), 372 (2018). arXiv:1803.10230
51. S. Lottini, private communication (2014)
52. ALPHA Collaboration, J. Bulava, M. Della Morte, J. Heitger, and 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, 555–568 (2015). arXiv:1502.04999
53. P. Fritzsch, Mass-improvement of the vector current in three-flavor QCD. JHEP 06, 015 (2018). arXiv:1805.07401
54. R. Wohlert, Improved continuum limit lattice action for quarks. DESY-87-069 (unpublished)
55. M. Lüscher, P. Weisz, O($a$) improvement of the axial current in lattice QCD to one loop order of perturbation theory. Nucl. Phys. B 479, 429–458 (1996). arXiv:hep-lat/9606016
56. A. Ramos, Automatic differentiation for error analysis of Monte Carlo data. arXiv:1809.01289