BERGISCHE UNIVERSITÄT WUPPERTAL

Doctoral Thesis

Scale setting in lattice QCD with light and strange quarks using potential scales

Author:
Tom Matty Bo Asmussen

Supervisor:
Prof. Dr. Francesco Knechtli

A thesis submitted in fulfillment of the requirements for the degree of Doctor of Natural Science

June 2026

BERGISCHE
UNIVERSITÄT
WUPPERTAL

Abstract

This thesis develops a precision strategy for potential-based scale setting in $N_f = 2 + 1$ lattice QCD and a complementary study of Teper-link smearing for glueball spectroscopy. Using optimised Wilson-loop measurements on CLS ensembles, it determines the potential scales with controlled statistical and systematic uncertainties, obtaining $r_0 = 0.4729(57)(48)\,\mathrm{fm}$, $r_1 = 0.3127(24)(32)\,\mathrm{fm}$, and $r_0/r_1 = 1.532(12)$. The results are consistent with current lattice averages and indicate that r_0 is the more robust reference scale compared to r_1. The Teper-link analysis yields only qualitative evidence for improved operator overlap.

The primary goal of this work is a first-principles, high-precision determination of the Sommer parameter from the static quark–antiquark potential, together with a transparent account of how operator construction, excited-state suppression, force interpolation, and continuum extrapolation enter the final calculation. The calculation is based on the CLS $N_f = 2 + 1$ ensembles, generated with HMC/RHMC using the tree-level improved Lüscher–Weisz gauge action and non-perturbatively $O(a)$-improved Wilson fermions, mostly with open temporal boundaries, at five lattice spacings between 0.039 and 0.085 fm, large volumes satisfying $m_\pi L \gtrsim 4$, and pion masses ranging from about 420 MeV to the physical point. Static energies are extracted from HYP-smeared Wilson loops by a variational GEVP analysis, fit-based effective-mass determination, tree-level improved distances, and a controlled treatment of open-boundary effects. Statistical uncertainties are propagated with the Γ-method including autocorrelations, while systematic effects are assessed through alternative chiral and continuum fit ansätze, lattice-spacing and pion-mass cuts, and AIC-weighted model averaging. Working with the static force removes the additive self-energy renormalisation of $V(r)$ and permits controlled determinations of r_0 and r_1. Beyond scale setting, the analysis yields $r_0\Lambda^{(3)}_{\overline{\mathrm{MS}}} = 0.820(28)$ and shows through the curvature observable $c(r)$ that the potential interpolates smoothly between perturbative short-distance behaviour and confinement-like intermediate-distance behaviour. The secondary part of the thesis implements Teper blocking for glueball operators and finds a signal mainly for the A_1^{++} ground state; although Teper-blocked operators can modestly improve early-time behaviour in combined bases, the evidence remains exploratory, and a statistically larger study is required for definitive spectral conclusions.

Acknowledgements

My first thanks go to my supervisor, Francesco Knechtli, for his guidance, encouragement, and mentorship throughout my doctoral studies. His insight, patience, and high standards have had a lasting impact on both this thesis and my development as a researcher. His thoughtful comments and suggestions on an earlier version of this manuscript were invaluable. I am also grateful for the many opportunities he provided to present my research internationally, which broadened my perspective and contributed significantly to my professional growth. Despite the demands of his role as Dean, he consistently tried to make time whenever I needed advice or support, and I am deeply grateful for the effort and dedication he invested in supervising this project.

Tomasz Korzec was, in practice, a second supervisor. There was no question I could not bring to him, no problem too half-formed to discuss, and his door, or his inbox, was always open. I also greatly benefited from his careful reading of an earlier version of the thesis and his many constructive suggestions. I am glad he was there.

Roman Höllwieser shared an office with me through the early and middle years of the PhD, and that proximity mattered. Many problems that felt stuck got unstuck through an impromptu conversation, and his experience saved me real time on more than one occasion.

Juan Andrés Urrea Niño's knowledge of glueball spectroscopy was essential to Chapter 7. The discussions we had in the final stages helped me understand the subject in a way I could not have managed alone, and his comments on a draft of that part of the thesis sharpened it significantly.

Andreas Risch helped with several coding problems at moments when I needed it, and I am grateful for his patience with my questions.

I am grateful to all current and former members of the research group for creating an enjoyable and inspiring research environment. Their collegiality, support, and willingness to share ideas made the challenges of research both more manageable and more rewarding.

My family and friends kept me grounded throughout all of this, and I am grateful for their support in ways that are hard to put into words. Growing up between Flensburg and Denmark means many of the people who matter a lot to me are far from Wuppertal, and the distance made staying in touch harder than I would have liked, but those friendships have held, and I hope they keep doing so.

Most of all, thank you to Luisa, my wife. We met just before this PhD began, which means she has known me through the whole of it. I came out of these years a better researcher and a better person, and that is largely because of her. Her love, patience, strength, and constant support carried me through even the most difficult times.

I would also like to remember our daughter, Ida. Though her time with us was far too short, her memory remains an important part of our lives and of the journey that led to this thesis.

Contents

Chapter 1

Introduction

Quantum Chromodynamics (QCD) is the theory of the strong interaction that describes hadrons as bound states of quarks and gluons. Its ultraviolet behaviour is governed by asymptotic freedom [1, 2], which makes perturbation theory quantitatively successful at short distances, while its infrared regime is characterised by confinement (the absence of isolated quarks), mass generation, and the emergence of hadrons as the physically observable states. These long-distance phenomena are intrinsically nonperturbative and therefore require methods that go beyond weak-coupling expansions. In particular, any first-principles determination of hadron masses, matrix elements, or reference scales must ultimately confront the strongly coupled regime in which the QCD running coupling is no longer small. This is the point at which lattice QCD enters. Wilson's lattice formulation [3] replaces continuous Euclidean space-time by a discrete grid. Gauge fields are represented by link variables living on the lattice links, while quark degrees of freedom are discretised as lattice fermion fields. The construction preserves exact local gauge invariance at non-zero lattice spacing and turns the QCD path integral into a problem amenable to numerical evaluation. The pioneering Monte Carlo studies [4] of lattice gauge theory showed that this formulation is not only conceptually clean, but also practically powerful: it provides a first-principles framework in which non-perturbative observables can be computed without model assumptions.

A lattice calculation, however, does not begin by producing answers in MeV or fm. It first produces dimensionless quantities, such as am, where a is the lattice spacing and m a physical mass. Determining a in physical units is therefore not an optional extra: it is the step that turns simulation data into phenomenology. It is also what allows one to compare ensembles generated at different bare couplings, to perform continuum extrapolations, and to state final results in a form that can be confronted with experiment or with other lattice calculations. For that reason, scale setting is one of the central technical and conceptual tasks in lattice QCD. There is no unique way to set the scale. One may use hadron masses or decay constants, whose values are known experimentally and therefore tie the lattice spacing directly to physical observables. Alternatively, flow-based quantities such as t_0 [5] and w_0 [6] are widely used because they can be computed with high precision. In this thesis, however, the main focus is on scales derived from the static quark–antiquark potential [7], the energy of a fixed heavy quark pair separated by a distance r. This is a particularly appealing choice because the static potential is itself one of the most fundamental observables in non-perturbative gauge theory. Defined at all distance scales, it smoothly connects the short-distance regime governed by asymptotic freedom with the long-distance confining regime. A scale derived from this quantity is therefore not merely a conversion factor, it remains closely tied to the underlying physics.

1.1 Potential scales from the static force

The static potential $V(r)$ is extracted on the lattice from Wilson loops [3], i.e. gauge-invariant closed paths of link variables in space-time. Historically, it played an important role in establishing confinement on the lattice, while also providing a non-perturbative observable that can be compared with perturbative expectations at short distances, where the coupling becomes weak. At the same time, the potential comes with an additive self-energy contribution on the lattice, which is ultraviolet divergent and depends on the regularisation. This makes it natural to look not at $V(r)$ itself, but at its derivatives. The static force, $F(r) = \frac{dV(r)}{dr}$, is free of that additive ambiguity and is therefore a renormalised quantity for defining reference scales. In full QCD this is especially attractive, because the string tension, i.e. the coefficient governing the asymptotic linear rise of the static potential, is no longer an ideal scale-setting observable once string breaking, where a quark–antiquark pair is created from the vacuum, enters at large distances.

A seminal step was taken by Rainer Sommer, who proposed an intermediate-distance reference scale r_0 [7] through the condition $r_0^2 F(r_0) = 1.65$. By defining the scale through the force rather than through the potential itself, one avoids the additive self-energy problem; by choosing an intermediate distance, one stays away from both the very short-distance region, where lattice artefacts are more delicate, and the very long-distance region, where the signal deteriorates and string-breaking physics complicates the interpretation in full QCD. In phenomenological potential models, the condition corresponds roughly to $r_0 \approx 0.5\,\mathrm{fm}$, which makes the scale intuitively hadronic as well as numerically convenient. Later work demonstrated the practical usefulness of this construction. Precision quenched studies [8] showed that r_0/a can be determined with high precision and that it behaves smoothly enough to support controlled continuum extrapolations. A shorter-distance companion scale, $r_1^2 F(r_1) = 1.0$ [9], was later introduced and has been used in a number of dynamical-fermion studies. Because r_1 is defined at a somewhat shorter distance than r_0, it is often statistically easier to determine and can be advantageous on finer lattices or in analyses where the extraction of the static potential from large Wilson loops is computationally demanding due to the exponentially degrading signal-to-noise ratio. Together, r_0 and r_1 provide a pair of closely related reference scales that are both rooted in the same physical observable, but probe slightly different distance regimes.

What makes these potential scales especially interesting is that they do more than set the lattice spacing. The same framework also gives access to the shape of the static force, to the interpolation between perturbative and the long-distance behaviour of the static potential, and to potential-based couplings defined from derivatives of $V(r)$. Even though these potential scales are not directly experimentally accessible, their physical values were historically inferred phenomenologically through heavy-quark potential models constrained by the charmonium and bottomonium spectra.

1.2 Glueballs and Teper blocking

A separate line of investigation pursued in this thesis is glueball spectroscopy, which is largely independent of the potential-scale analysis. Glueballs are hypothetical hadronic states composed predominantly of gluons, the carriers of the strong force, without valence quarks [10–12]. Their existence is a natural consequence of QCD because, unlike

photons in QED, gluons themselves carry colour charge and therefore interact with one another. QCD thus allows not only quark-based hadrons, but also bound states built mainly from gluonic degrees of freedom. This possibility was already recognised in the 1970s [10], and lattice calculations in pure gauge theory [13] later mapped out the low-lying glueball spectrum with increasing precision. In particular, the scalar channel emerged early as the most prominent and one of the most phenomenologically relevant candidates. Glueballs are theoretically fascinating precisely because they are such a clean manifestation of non-Abelian gauge dynamics, but they are also particularly challenging. Experimentally, there is still no universally accepted, model-independent glueball identification. In full QCD the problem becomes even harder, because purely gluonic states are expected to mix with flavour-singlet mesons and multi-hadron states, while lattice calculations suffer from a severe signal-to-noise problem [13–16]. For that reason, the quenched spectrum is better established than the unquenched one, and unquenched glueball spectroscopy remains more exploratory.

It is in this context that Teper blocking [17] becomes interesting. Conceptually, this is closely related to gauge-link smearing techniques such as APE smearing [18], which improve the overlap of gauge-invariant operators with low-lying states by suppressing ultraviolet fluctuations. Teper's idea was to recursively construct spatially extended gauge-covariant operators, so that the resulting operators retain the correct symmetry properties while probing the infrared length scales relevant for the low-lying glueball states. In practical terms, this means replacing thin links by iteratively blocked combinations of straight paths and staples, thereby constructing operators with improved overlap onto the states one wants to see. Historically, this was a major step because it made reliable glueball calculations significantly more efficient. Later high-precision glueball studies used increasingly rich bases of extended operators and variational techniques, but the underlying logic remained the same: improved operators are crucial for obtaining cleaner spectral signals. In this thesis, this idea is revisited within the framework developed previously by our research group [19–23], by incorporating Teper blocking into the existing smearing procedure and assessing its impact on the resulting glueball signals.

1.3 Structure of the thesis

The thesis is organised as follows. Chapter 2 reviews the theoretical foundations of lattice gauge theory relevant for this work, including the formulation of static sources and the definition of the static quark–antiquark potential. Chapter 3 describes the computational setup and the ensemble properties used in the analysis. The determination of the static potential from Wilson loops is developed in Chapter 4, while Chapter 5 presents the scale-setting analysis and the continuum extrapolation leading to the final results for the potential scales. Selected physical implications are discussed in Chapter 6. Chapter 7 contains a separate, exploratory study of Teper blocking in glueball spectroscopy. Finally, Chapter 8 summarises the results and provides an outlook.

Part ב

The Fellowship of the Theoretical Foundations

Chapter 2

Theory

Most of the theoretical ingredients assembled in this part are standard and widely accepted within the lattice QCD community; to avoid unnecessary clutter, explicit citations for basic definitions and well-known identities are generally omitted, and we instead take Refs. [24–26] as the primary reference sources for conventions and foundational results.

2.1 Lattice gauge fields: from continuum QCD to the Wilson action

Quantum Chromodynamics (QCD) is the established theory of the strong interaction and has enjoyed striking quantitative success in regimes where a weak-coupling expansion is justified. In particular, asymptotic freedom [27] implies that at sufficiently high momentum transfer the running coupling becomes small, and perturbation theory provides a controlled description of short-distance processes. At the same time, many of the defining phenomena of the strong interaction such as confinement, the emergence of a hadron spectrum, and the generation of nontrivial long-distance scales, are intrinsically non-perturbative. In these regimes, a first-principles treatment requires a formulation beyond ordinary perturbative expansions. The lattice approach provides such a formulation by introducing a non-perturbative ultraviolet regulator that preserves locality and (crucially) exact local gauge invariance at finite cutoff [3]. The key conceptual step is to represent gauge fields by group-valued parallel transporters on the links of a hypercubic spacetime lattice. Gauge-invariant observables are then naturally built from products of link variables around closed loops, and the simplest local gauge action is constructed from elementary plaquettes, i.e. the smallest square Wilson loops on the lattice. This section summarises these starting principles and fixes notation for the gauge sector.

2.1.1 Continuum gauge theory in Euclidean space

Before introducing the lattice formulation, we briefly summarise the continuum formulation of QCD in Euclidean space and fix notation. This provides the reference framework from which the lattice discretisation is constructed and clarifies how continuum quantities are represented on the lattice. We consider QCD with gauge group $SU(N_c)$ and ultimately set $N_c = 3$. Let T^a $(a = 1, \ldots, N_c^2 - 1)$ denote generators of the Lie algebra in the fundamental representation, normalised by

$$\mathrm{Tr}\Big(T^a T^b\Big) = \frac{1}{2}\,\delta^{ab}, \qquad [T^a, T^b] = i f^{abc} T^c, \tag{2.1}$$

where f^{abc} are the structure constants. The gluon field is written as

$$A_\mu(x) = A_\mu^a(x)\, T^a, \tag{2.2}$$

using Einstein's summation notation, and the covariant derivative acting on a quark field in the fundamental representation is

$$D_\mu = \partial_\mu + ig\, A_\mu(x), \tag{2.3}$$

with gauge coupling g. The corresponding field-strength tensor is defined by

$$F_{\mu\nu}(x) \equiv -\frac{i}{g}\,[D_\mu, D_\nu] = \partial_\mu A_\nu - \partial_\nu A_\mu + ig\,[A_\mu, A_\nu], \tag{2.4}$$

so that in components,

$$F_{\mu\nu}^a = \partial_\mu A_\nu^a - \partial_\nu A_\mu^a - g f^{abc} A_\mu^b A_\nu^c. \tag{2.5}$$

We work in four-dimensional Euclidean space with metric $\delta_{\mu\nu}$, and repeated indices are summed using $\delta_{\mu\nu}$. The Euclidean Yang–Mills action is

$$S_g[A] = \frac{1}{4}\int d^4x\; F_{\mu\nu}^a(x)\, F_{\mu\nu}^a(x). \tag{2.6}$$

Under a local gauge transformation $\Omega(x) \in SU(N_c)$, the gauge field transforms as

$$A_\mu(x) \to \Omega(x)\, A_\mu(x)\, \Omega^\dagger(x) - \frac{i}{g}\,(\partial_\mu \Omega(x))\, \Omega^\dagger(x), \tag{2.7}$$

and $F_{\mu\nu}(x) \to \Omega(x) F_{\mu\nu}(x) \Omega^\dagger(x)$, so that $S_g[A]$ is gauge invariant.

2.1.2 Spacetime discretisation

We now proceed to discretise spacetime and introduce the lattice formulation of the gauge field. We introduce a hypercubic lattice with spacing a. Lattice sites are labelled by x, and unit vectors in the coordinate directions are denoted by $\hat{\mu}$ with $\mu \in \{0, 1, 2, 3\}$. A nearest-neighbour site in direction μ is written as $x + a\hat{\mu}$. In finite-volume simulations one typically considers a lattice of extent $L^3 \times T$ (in physical units) corresponding to $(L/a)^3 \times (T/a)$ lattice points. Boundary conditions and details of the simulation setup will be specified in chapter 3. The defining step of lattice gauge theory is to replace the continuum gauge field by group-valued link variables. A link variable $U_\mu(x) \in SU(N_c)$ is associated with the oriented link from x to $x + a\hat{\mu}$ and is interpreted as the parallel transporter along that link. In the formal continuum limit, this relation is captured by

$$U_\mu(x) \;=\; \exp\Big(iag\, A_\mu(x)\Big), \tag{2.8}$$

which serves as motivation rather than a definition at finite a.

Local gauge transformations are represented by group elements $\Omega(x) \in SU(N_c)$ living on lattice sites. The link variables transform as

$$U_\mu(x) \to \Omega(x)\, U_\mu(x)\, \Omega^\dagger(x + a\hat{\mu}). \tag{2.9}$$

This transformation law is the lattice analogue of continuum gauge covariance, and it has the crucial feature that gauge invariance can be implemented exactly at any finite lattice spacing through suitable contractions of link variables into closed loops. The smallest nontrivial closed loop on the hypercubic lattice is the plaquette. The plaquette in the $\mu\nu$ plane at site x is defined as

$$U_{\mu\nu}(x) = U_\mu(x)\, U_\nu(x + a\hat{\mu})\, U_\mu^\dagger(x + a\hat{\nu})\, U_\nu^\dagger(x). \tag{2.10}$$

Under gauge transformations, $U_{\mu\nu}(x) \to \Omega(x) U_{\mu\nu}(x) \Omega^\dagger(x)$, and therefore $\operatorname{Tr} U_{\mu\nu}(x)$ is gauge invariant. More generally, given a closed contour C on the lattice, the associated Wilson loop is defined as the ordered product of link variables along C,

$$U(C) = \prod_{\ell \in C} U_\ell, \tag{2.11}$$

where the product follows the orientation of the contour and U_ℓ denotes the link matrix on each oriented segment. Gauge-invariant loop observables are obtained by taking the trace, e.g.

$$W(C) = \frac{1}{N_c} \operatorname{Re} \operatorname{Tr} U(C). \tag{2.12}$$

Rectangular Wilson loops will play a central role later when defining and extracting the static quark–antiquark potential. A key consistency requirement is that the lattice formulation reproduces the continuum gauge action as $a \to 0$. In the continuum limit, the plaquette is the lattice object associated with the gauge-field strength. For sufficiently smooth gauge fields, the ordered product of link variables around an elementary square measures how much a colour vector is rotated when parallel transported around that loop, and this reproduces the local field-strength tensor in the continuum theory. Although larger Wilson loops can also be used to construct gauge actions with the correct continuum limit, the plaquette is the most local gauge-invariant loop observable and therefore provides the simplest basis for the Wilson gauge action.

2.1.3 The Wilson gauge action

The simplest local, gauge-invariant action for the link variables is the Wilson plaquette action,

$$S_W[U] = \beta \sum_x \sum_{\mu<\nu} \left(1 - \frac{1}{N_c} \operatorname{Re} \operatorname{Tr} U_{\mu\nu}(x)\right), \qquad \beta \equiv \frac{2N_c}{g_0^2}, \tag{2.13}$$

where g_0 is the bare lattice gauge coupling. β is chosen as an input parameter and thereby fixes the lattice spacing a, which is determined a posteriori from a scale-setting observable. Gauge invariance of S_W follows immediately from the gauge-covariant transformation of the plaquette and the cyclicity of the trace. The action is local because each term involves only links around a single plaquette. In the continuum limit $a \to 0$, the Wilson action approaches the Euclidean Yang–Mills action (2.6) (up to an additive constant and corrections that vanish with powers of a). The precise relation is obtained by expanding the plaquette for smooth fields, and it fixes the identification between β and the bare coupling as written in Eq. (2.13). With the Wilson action, the Euclidean partition function of the pure gauge theory is defined as

a finite-dimensional group integral,

$$Z = \int \prod_{x,\mu} dU_\mu(x)\; e^{-S_W[U]}, \tag{2.14}$$

where $dU_\mu(x)$ is the Haar measure on $SU(N_c)$ for each link. Expectation values of gauge-field observables $\mathcal{O}[U]$ are

$$\langle \mathcal{O} \rangle = \frac{1}{Z} \int \prod_{x,\mu} dU_\mu(x)\; \mathcal{O}[U]\; e^{-S_W[U]}. \tag{2.15}$$

In full QCD, the fermionic degrees of freedom are discretised separately and lead, after Grassmann integration, to an additional factor given by the determinant of the lattice Dirac operator.

2.2 Lattice fermions and the fermion determinant

To define lattice QCD, the gauge-field measure introduced in the previous section must be supplemented by a discretisation of the quark fields. In the Euclidean functional integral, quarks are represented by Grassmann-valued fields. Upon discretisation, these fermionic degrees of freedom are placed on the lattice sites, while gauge invariance is maintained by coupling them to the link variables that implement parallel transport between neighbouring sites. This section introduces the basic lattice fermion formulation used throughout this thesis, emphasising the origin of the fermion doubling problem and the Wilson discretisation that resolves it.

2.2.1 Fermion fields on the lattice and gauge-covariant differences

For each quark flavour $f = 1, \ldots, N_f$, we introduce Grassmann fields $\psi_f(x)$ and $\bar{\psi}_f(x)$ defined on lattice sites x. The fields carry colour indices in the fundamental representation of $SU(N_c)$ and Dirac spinor indices. Gauge covariance at finite lattice spacing is ensured by transporting fermion fields between neighbouring sites with the link variables $U_\mu(x)$. A convenient starting point is the gauge-covariant forward and backward difference operators,

$$\nabla_\mu \psi(x) \equiv \frac{1}{a}\Big(U_\mu(x)\, \psi(x + a\hat{\mu}) - \psi(x) \Big), \tag{2.16}$$

$$\nabla^*_\mu \psi(x) \equiv \frac{1}{a}\Big(\psi(x) - U^\dagger_\mu(x - a\hat{\mu})\, \psi(x - a\hat{\mu}) \Big), \tag{2.17}$$

which transform covariantly under local gauge transformations.

In the continuum, the Euclidean Dirac operator is given by

$$D = \gamma_\mu D_\mu + m, \tag{2.18}$$

where D_μ denotes the gauge-covariant derivative. Replacing the continuum derivatives by symmetric covariant finite differences leads to the naive lattice Dirac operator

$$D_{\text{naive}} = \frac{1}{2} \sum_{\mu=0}^{3} \gamma_\mu \left(\nabla_\mu + \nabla^*_\mu \right) + m_0, \tag{2.19}$$

where m_0 is a bare mass parameter. Although D_{naive} reproduces the continuum Dirac operator in the formal limit $a \to 0$, it suffers from the fermion doubling problem: in the free-field case, the lattice dispersion relation contains additional poles corresponding to extra light fermionic modes. In four dimensions this yields 2^4 species in the continuum limit. This is a generic feature of local, translationally invariant discretisations with exact chiral symmetry at finite lattice spacing.

2.2.2 Wilson fermions

A standard resolution of the doubling problem is provided by Wilson fermions, which add an $O(a)$ term (the Wilson term) that lifts the doublers' masses to the cutoff scale while leaving the desired continuum limit unchanged. The Wilson–Dirac operator can be written in the compact form

$$D_W \equiv \frac{1}{2}\sum_{\mu=0}^{3}\left[\gamma_\mu\left(\nabla_\mu + \nabla_\mu^*\right) - ar\,\nabla_\mu^*\nabla_\mu\right] + m_0, \tag{2.20}$$

where r is the Wilson parameter (most commonly $r = 1$). The corresponding lattice fermion action for a single flavour is

$$S_F[\bar\psi, \psi, U] = a^4 \sum_x \bar\psi(x)\, D_W\, \psi(x). \tag{2.21}$$

In an explicit nearest-neighbour form, Eq. (2.21) may be written as

$$\begin{aligned} S_F = a^4 \sum_x \Bigg[& \bar\psi(x)\left(m_0 + \frac{4r}{a}\right)\psi(x) \\ & - \frac{1}{2a}\sum_{\mu=0}^{3}\Big(\bar\psi(x)(r-\gamma_\mu)U_\mu(x)\psi(x+a\hat\mu) + \bar\psi(x)(r+\gamma_\mu)U_\mu^\dagger(x-a\hat\mu)\psi(x-a\hat\mu)\Big)\Bigg]. \end{aligned} \tag{2.22}$$

It is often convenient to trade the bare mass parameter for the hopping parameter κ via

$$\kappa \equiv \frac{1}{2(am_0 + 4r)}, \qquad am_0 = \frac{1}{2\kappa} - 4r, \tag{2.23}$$

in terms of which the action takes a standard hopping form. The Wilson term removes the doublers but explicitly breaks chiral symmetry at finite lattice spacing. One important consequence is an additive mass renormalisation: the critical value of the bare mass (or κ) at which the renormalised quark mass vanishes is shifted from its naive value and must be determined non-perturbatively. In addition, Wilson fermions generically exhibit discretisation effects linear in the lattice spacing a, motivating the improvement programme discussed in Section 2.3.

2.2.3 Integrating out fermions: determinant and propagators

The lattice QCD partition function with N_f flavours can be written formally as

$$Z = \int \Big[\prod_{x,\mu} dU_\mu(x)\Big] \Big[\prod_{f=1}^{N_f} \mathcal{D}\bar{\psi}_f \, \mathcal{D}\psi_f\Big] \, \exp\left\{-S_g[U] - \sum_{f=1}^{N_f} a^4 \sum_x \bar{\psi}_f(x) \, D_f[U] \, \psi_f(x)\right\}, \tag{2.24}$$

where $S_g[U]$ denotes the chosen gauge action and $D_f[U]$ the chosen lattice Dirac operator for flavour f (e.g. D_W with flavour-dependent mass parameter). Since the fermion fields enter quadratically, the Grassmann integrals can be carried out exactly, yielding

$$Z = \int \Big[\prod_{x,\mu} dU_\mu(x)\Big] \, \exp\{-S_g[U]\} \prod_{f=1}^{N_f} \det D_f[U]. \tag{2.25}$$

Fermionic correlation functions are obtained by inserting fermionic operators before integrating out the Grassmann variables; Wick contraction then expresses them in terms of the inverse Dirac operator $D_f[U]^{-1}$ (the quark propagator) evaluated in the background gauge field. For Wilson fermions one has the γ_5-hermiticity property

$$D_W^\dagger = \gamma_5 D_W \gamma_5, \tag{2.26}$$

which implies that eigenvalues occur in complex-conjugate pairs and is useful when discussing positivity properties of fermion determinants for certain flavour combinations. The precise conditions under which the fermion determinant is real and positive depend on the number of flavours and the quark mass parameters. Since these issues are primarily relevant for simulation algorithms and do not affect the formal structure of the path integral used here, we do not discuss them further.

2.3 Discretisation effects and Symanzik improvement

A lattice formulation defines QCD with an explicit ultraviolet cutoff set by the lattice spacing a. Physical predictions are recovered by taking the continuum limit $a \to 0$ at fixed renormalised parameters. At finite a, expectation values of suitably defined observables differ from their continuum values by discretisation effects. A convenient organising principle for these effects is provided by the Symanzik effective theory viewpoint [28, 29]. For local lattice actions with the correct symmetries and a well-defined continuum limit, lattice artefacts can be described by an effective continuum action containing the target continuum action plus a series of higher-dimensional operators suppressed by powers of a,

$$S_{\mathrm{lat}} \;\sim\; S_4 + a\, S_5 + a^2\, S_6 + \cdots . \tag{2.27}$$

Here S_4 represents the dimension-four continuum QCD action, including all operators compatible with the continuum symmetries. More generally, S_d denotes a linear combination of gauge-invariant operators of canonical dimension d allowed by the lattice symmetries. In this language, improvement consists in adjusting the lattice action (and, where required, composite operators) such that the coefficients of the leading correction terms vanish, thereby shifting discretisation errors to higher order in a. The Symanzik framework therefore provides the theoretical basis for continuum extrapolations and for the construction of improved lattice actions.

2.3.1 O(a)-improved Wilson fermions

For Wilson fermions, leading discretisation effects in on-shell quantities are generically of $O(a)$, reflecting the presence of dimension-five operators in the Symanzik effective theory due to explicit chiral symmetry breaking. The standard $O(a)$ improvement of the Wilson fermion action is obtained by adding a single dimension-five operator (the "clover" term) to the lattice action [30],

$$S_F^{\mathrm{imp}} = S_F^{\mathrm{Wilson}} + a^5\, c_{\mathrm{SW}} \sum_x \bar{\psi}(x)\, \frac{i}{4} \sigma_{\mu\nu}\, F_{\mu\nu}^{\mathrm{lat}}(x)\, \psi(x), \tag{2.28}$$

where $\sigma_{\mu\nu} \equiv \frac{i}{2}[\gamma_\mu, \gamma_\nu]$, $F_{\mu\nu}^{\mathrm{lat}}(x)$ is a lattice discretisation of the field strength built from an average of plaquettes in the $\mu\nu$ plane (the "clover" construction), and c_{SW} is an improvement coefficient. When c_{SW} is tuned appropriately and composite operators are improved consistently, the leading cutoff effects of on-shell observables are removed, so that discretisation errors start at $O(a^2)$ (up to possible logarithmic factors). The precise tuning strategy and the values used in the numerical simulations are specified in Chapter 3.

2.3.2 Improved gauge actions: the Lüscher–Weisz class

The Wilson plaquette action in Equation (2.13) provides the simplest gauge-invariant discretisation of the Yang–Mills action, but it exhibits $O(a^2)$ discretisation effects already at the classical level. Gauge-action improvement aims to reduce these effects by adding further local loop terms such that the leading contributions in the Symanzik expansion cancel. A widely used class of improved gauge actions is obtained by taking linear combinations of the plaquette and larger Wilson loops (most commonly planar 1×2 rectangles, and depending on the chosen scheme possibly additional six-link loops) [31],

$$S_g^{\mathrm{imp}}[U] = \beta \sum_x \sum_{\mu<\nu} \left[c_0 \left(1 - \frac{1}{N_c} \mathrm{Re}\,\mathrm{Tr}\, U_{\mu\nu}^{1\times 1}(x) \right) + c_1 \left(1 - \frac{1}{N_c} \mathrm{Re}\,\mathrm{Tr}\, U_{\mu\nu}^{1\times 2}(x) \right) + \cdots \right], \tag{2.29}$$

with coefficients constrained such that the correct continuum limit is recovered (e.g. $c_0 + 8c_1 + \cdots = 1$ for common normalisations). The tree-level Symanzik-improved (Lüscher–Weisz-type) gauge action corresponds to a particular choice of coefficients that cancels the leading $O(a^2)$ terms in the classical continuum expansion. The precise definition (including coefficient conventions and any boundary counterterms required by the boundary conditions) is given later, where the simulation setup is specified.

2.4 Euclidean correlators and spectral representation

A central advantage of the Euclidean formulation is that energies can be extracted from the large-time behaviour of correlation functions. On the lattice, this statement can be formulated cleanly using the transfer matrix: under standard assumptions (locality and reflection positivity) there exists a positive transfer matrix $\mathcal{T}$ whose eigenvalues are related to the energy spectrum of the theory. In practice, one exploits the resulting spectral representation to isolate ground-state energies from exponential falloff at large Euclidean time separations.

2.4.1 Two-point functions and ground-state projection

Let $\mathcal{O}(x)$ be a gauge-invariant (or gauge-covariant with appropriate contractions) interpolating operator with the quantum numbers of interest. A standard object is the zero-momentum projected two-point function

$$C_{\mathcal{O}}(t) \equiv a^3 \sum_{\mathbf{x}} \Big\langle \mathcal{O}(t,\mathbf{x})\, \mathcal{O}^\dagger(0,\mathbf{0}) \Big\rangle, \qquad t = na, \quad n \in \mathbb{Z}_{\geq 0}. \tag{2.30}$$

Here $\langle \cdots \rangle$ denotes the path-integral expectation value with the lattice action and measure defined previously. In a transfer-matrix formulation, one may write $C_{\mathcal{O}}(t)$ as a matrix element of $\mathcal{T}^n$ between appropriate states. This leads to the spectral decomposition

$$C_{\mathcal{O}}(t) = \sum_{k \geq 0} |\langle 0|\mathcal{O}|k\rangle|^2\, e^{-E_k t}, \qquad t > 0, \tag{2.31}$$

where $|0\rangle$ is the vacuum and E_k are energy eigenvalues in the channel of $\mathcal{O}$. For large Euclidean times, excited-state contributions are exponentially suppressed and

$$C_{\mathcal{O}}(t) \xrightarrow[t\to\infty]{} |\langle 0|\mathcal{O}|0^*\rangle|^2\, e^{-E_0 t}, \tag{2.32}$$

where E_0 is the lowest energy in that channel and $|0^*\rangle$ denotes the corresponding lowest-lying state (not the vacuum unless $\mathcal{O}$ has vacuum quantum numbers). On a lattice with temporal extent T and (anti-)periodic boundary conditions, backward propagation around the temporal boundary can contribute, and the asymptotic form,

$$C_{\mathcal{O}}(t) \sim A \left(e^{-E_0 t} \pm e^{-E_0 (T-t)} \right), \tag{2.33}$$

with the sign depending on the operator content and boundary conditions, which can be written in terms of a hyperbolic cosine for the case of periodic boundary conditions. In regimes where $t \ll T$ the second term is negligible, and the simple exponential form is recovered. A useful diagnostic is the effective mass/energy, defined by

$$a\, m_{\mathrm{eff}}(t) \equiv \ln\!\left(\frac{C_{\mathcal{O}}(t)}{C_{\mathcal{O}}(t+a)} \right), \tag{2.34}$$

which approaches aE_0 as $t \to \infty$ if a single exponential dominates. The same logic extends to correlation matrices of multiple operators, where a variational treatment improves ground-state isolation; such methods are part of the numerical analysis and will be described where they are used.

2.5 Static sources and the static quark–antiquark potential

The potential between a static quark and antiquark separated by a distance r is among the most fundamental gauge-invariant quantities that can be computed non-perturbatively in QCD. It provides a single observable that probes both the short-distance regime governed by asymptotic freedom and the long-distance regime associated with confinement physics. Historically, Wilson-loop computations established a concrete, non-perturbative characterisation of confinement in terms of the formation of a flux tube between static colour sources and the resulting rise of the static energy with distance. From a practical standpoint, the static potential and, in particular,

its derivatives are also useful because they lead to clean reference scales such as the potential scales. These are defined directly from the force and do not depend on fixing an additive normalisation of $V(r)$. This is advantageous because $V(r)$ contains an ultraviolet self-energy contribution associated with the static sources, which is regulator dependent and divergent as the cutoff is removed. In the following, we summarise how static sources are represented on the lattice and how rectangular Wilson loops define the static energy levels and the potential. We then discuss the additive renormalisation affecting $V(r)$, which motivates the introduction of force- and curvature-based observables and potential-derived couplings and phenomenological descriptions of the static interaction. The observable-specific procedure of tree-level improvement via improved distances is introduced at the end of this section, where it naturally accompanies the definitions of lattice derivative estimators of $V(r)$.

2.5.1 Static sources and temporal Wilson lines

The static limit corresponds to infinitely heavy colour sources whose spatial positions are fixed and whose propagation is purely along Euclidean time. On the lattice, the propagation of a static source at fixed spatial position $\mathbf{x}$ from time 0 to time t is represented by a straight temporal gauge transporter (a temporal Wilson line),

$$U_0(\mathbf{x}; 0 \to t) \equiv \prod_{\tau=0}^{t-a} U_0(\mathbf{x}, \tau), \qquad t = na, \quad n \in \mathbb{Z}_{\geq 0}, \tag{2.35}$$

where the product is ordered in increasing Euclidean time. Gauge-invariant operators describing a static $Q\bar{Q}$ pair separated by a spatial displacement $\mathbf{r}$ are constructed by combining temporal Wilson lines with spatial transporters that connect the two sources at fixed time slices. Different choices of spatial transporters affect overlap factors onto energy eigenstates but not the energy spectrum itself.

2.5.2 Rectangular Wilson loops and the definition of $V(r)$

We start with the closed-loop observables that were defined earlier for a general contour C by Equation (2.12) and specialise to the rectangular contour $C(r,t)$ of spatial extent r and Euclidean time extent t, and introduce the standard shorthand

$$W(r,t) \equiv W\Big(C(r,t)\Big), \qquad r = |\mathbf{r}|, \quad t = na. \tag{2.36}$$

Geometrically, $W(r,t)$ consists of two temporal Wilson lines at separation r connected by spatial transporters at the initial and final times.

In a transfer-matrix formulation, $W(r,t)$ admits a spectral decomposition [25] analogous to Equation (2.31) in the sector containing a static quark–antiquark pair at separation r,

$$W(r,t) = \sum_{n \geq 0} |c_n(r)|^2 \, e^{-E_n(r)\, t}, \tag{2.37}$$

where the energies $E_n(r)$ are the static energy levels and the coefficients $c_n(r)$ depend on the chosen operator definition (e.g. choice of spatial transporter and smearing) that quantify how strongly the chosen operators couple to the corresponding energy

eigenstates. The static potential is defined as the ground-state energy,

$$V(r) \equiv E_0(r), \qquad V(r) = -\lim_{t\to\infty} \frac{1}{t} \ln W(r,t). \tag{2.38}$$

2.5.3 Short-distance regime and potential-derived couplings

At sufficiently short distances, the static potential is well described by perturbation theory and calculations to two and three loop order have been completed [32–34] [35, 36] both in pure gauge theory and in full QCD. One can define renormalised couplings directly from $V(r)$ and its derivatives. A coupling defined from the potential itself [37] (often referred to as a V-scheme coupling) is

$$\alpha_V(\mu) \equiv -\frac{1}{C_F}\, r\, V(r)\Big|_{\mu=r^{-1}}, \tag{2.39}$$

where $C_F = (N_c^2-1)/(2N_c)$ is the quadratic Casimir invariant associated with quarks in the fundamental representation of SU(N_c). Couplings can also be defined from the force $F(r) \equiv V'(r)$ [7] and from the second derivative $V''(r)$ [38],

$$\alpha_{qq}(\mu) \equiv \frac{1}{C_F}\, r^2\, V'(r)\Big|_{\mu=r^{-1}}, \tag{2.40}$$

$$\alpha_c(\mu) \equiv -\frac{1}{2C_F}\, r^3\, V''(r)\Big|_{\mu=r^{-1}}. \tag{2.41}$$

The shape of $V(r)$ (or equivalently the behaviour of these couplings) at short distances is governed by the corresponding β-functions; in practice, this provides a bridge between non-perturbative lattice determinations at small r and the perturbative running described by continuum renormalisation group methods.

2.5.4 Long-distance behaviour, string picture, and phenomenological potentials

In the absence of dynamical fermions, confinement manifests itself in the asymptotic linear rise of the static potential,

$$V(r) \sim \sigma r + \cdots, \tag{2.42}$$

where σ is the string tension. A widely used parametrisation of the long-distance potential in this setting is motivated by effective string descriptions, which predict a universal $1/r$ correction [39, 40],

$$V(r) = \sigma r + \mu + \frac{\gamma}{r} + O(1/r^2), \qquad \gamma = -\frac{\pi}{24}(d-2), \tag{2.43}$$

in d spacetime dimensions, where μ is a regularisation-dependent constant. In full QCD, the string picture is only approximate because the flux tube can break due to light-quark pair production once the separation is large enough. At such separations the static energy smoothly transitions from a string-like potential to a regime dominated by static–light hadronic states, and a precise definition of the string tension becomes

ambiguous. This is one reason why force-based reference scales (defined at intermediate distances) are used in full QCD.

Phenomenological potential models assume a functional form for $V(r)$ to describe heavy quarkonium spectra. Two frequently discussed examples are the Cornell potential [41],

$$V_{\text{Cornell}}(r) = -\frac{0.52}{r} + \frac{r}{(2.34\ \text{GeV}^{-1})^2}, \tag{2.44}$$

and the Richardson potential [42],

$$V_{\text{Richardson}}(r) = \frac{8\pi}{33 - 2N_f}\,\Lambda\left(\Lambda r - \frac{f(\Lambda r)}{\Lambda r}\right), \tag{2.45}$$

$$f(t) = \frac{4}{\pi}\int_0^\infty dq\,\frac{\sin(qt)}{q}\left[\frac{1}{\ln(1+q^2)} - \frac{1}{q^2}\right]. \tag{2.46}$$

In lattice QCD, the static potential V(r) can be computed non-perturbatively from first principles, allowing direct comparison with phenomenological potential models, which will be discussed further in Section 6.2.

2.5.5 Additive renormalisation, force, and potential scales

When computed on a lattice with spacing a, the static potential contains an additive, linearly divergent self-energy contribution associated with the static sources. This may be expressed schematically as [43, 44]

$$V(r;a) = V_{\text{sub}}(r) + E_{\text{self}}(a), \qquad E_{\text{self}}(a) \sim \frac{1}{a}, \tag{2.47}$$

where $V_{\text{sub}}(r)$ denotes a subtracted potential (defined up to a finite constant by a renormalisation condition) and the precise form of $E_{\text{self}}(a)$ depends on the chosen discretisation of the static action and short-distance regularisation details.

As a consequence, derivatives of the potential are often preferable for defining physical quantities. The static force is

$$F(r) \equiv \frac{dV(r)}{dr}, \tag{2.48}$$

and one defines potential reference scales (also known as Sommer scales) r_i through the dimensionless condition [7, 9]

$$r_i^2\,F(r_i) = c_i, \qquad c_0 = 1.65, \quad c_1 = 1. \tag{2.49}$$

These reference distances are defined at intermediate separations where the lattice determination is precise and where the observable is insensitive to the additive ambiguity in $V(r)$. This makes them particularly suitable for precise determinations from the static potential. In addition, these scales are straightforward to extract numerically and have been widely used in the literature, which allows for direct comparison between different lattice calculations.

In addition to the force, it is useful to consider the curvature of the potential, defined as

$$c(r) = \frac{1}{2}r^3\frac{dF(r)}{dr}. \tag{2.50}$$

This quantity provides additional information on the shape of the static potential

beyond the force itself. In particular, it is sensitive to the variation of the force with distance and therefore probes deviations from simple parametrisations of the potential. At large distances, the behaviour of $c(r)$ can be related to expectations from effective string descriptions of confinement, while at shorter distances it reflects the transition to the perturbative regime. As such, $c(r)$ serves as a useful diagnostic for comparing lattice results with continuum expectations and for assessing the behaviour of the potential across different distance scales.

2.5.6 Tree-level improvement and improved distances for derivatives of V(r)

Derivatives of the static potential are central to the calculation of the potential scales and curvature-based observables. In numerical work these derivatives are obtained from finite differences of $V(r)$ at discrete separations. At finite lattice spacing, such finite-difference estimators inherit sizeable discretisation effects from the lattice geometry, in particular from the fact that the lattice Coulomb potential differs from the continuum Coulomb form at short distances. A way to reduce these geometric artefacts is *tree-level improvement*: one defines an *improved distance* [7, 8, 45] such that, at leading order in perturbation theory, the lattice estimator reproduces the corresponding continuum expression exactly (or to a higher order) when expressed in terms of this improved distance. At leading order in the gauge coupling, the static potential in the continuum is Coulombic,

$$V_{\text{cont}}(r) = -\,C_F \frac{g_0^2}{4\pi}\,\frac{1}{r}, \tag{2.51}$$

with g_0 the bare coupling and $C_F = (N_c^2 - 1)/(2N_c)$ the fundamental Casimir, giving $C_F = 4/3$ for gauge group $SU(3)$. On a spatial lattice, the corresponding tree-level potential is expressed in terms of the lattice Coulomb Green's function. In infinite volume one may write

$$V(\mathbf{r}) = -\,C_F\, g_0^2\, G(\mathbf{r}), \tag{2.52}$$

where $G(\mathbf{r})$ is the Green's function of the (spatial) lattice Laplacian. A convenient momentum-space representation is

$$G(\mathbf{r}) = \int_{-\pi}^{\pi} \frac{d^3k}{(2\pi)^3}\,\frac{e^{i\mathbf{k}\cdot(\mathbf{r}/a)}}{\hat{k}^2}, \qquad \hat{k}^2 \equiv 4\sum_{j=1}^{3} \sin^2\!\left(\frac{k_j}{2}\right). \tag{2.53}$$

Here $k_j = ap_j$ denotes the dimensionless lattice momentum, so that the integration region is the Brillouin zone

$$k_j \in [-\pi, \pi].$$

A commonly used estimator for the static force evaluates a finite difference at the midpoint between two separations,

$$F\!\left(r + \frac{a}{2}\right) \equiv \frac{V(r+a) - V(r)}{a}, \tag{2.54}$$

where r denotes the magnitude of the separation at which the potential is measured (e.g. $r = na$ for on-axis separations). The definition (2.54) is insensitive to additive constants in V. In the continuum Coulomb case (2.51), the corresponding leading-order

force is

$$F_{\mathrm{cont}}(r) = C_F \frac{g_0^2}{4\pi} \frac{1}{r^2}. \tag{2.55}$$

Tree-level improvement amounts to defining, for each lattice midpoint $r + a/2$, an *improved distance* r_I such that the tree-level lattice estimator (2.54), evaluated with the tree-level lattice potential (2.52), matches the continuum form (2.55) when expressed at r_I [46]:

$$F_{\mathrm{tree}}(r_I) = \frac{V_{\mathrm{tree}}(r+a) - V_{\mathrm{tree}}(r)}{a} = C_F \frac{g_0^2}{4\pi} \frac{1}{r_I^2}. \tag{2.56}$$

Using Eqs. (2.52) and (2.56), this can be written equivalently as

$$\frac{G(r) - G(r+a)}{a} = \frac{1}{4\pi\, r_I^2}, \tag{2.57}$$

where $G(r)$ denotes $G(\mathbf{r})$ for the chosen displacement vector of length r (e.g. on-axis). Equation (2.57) shows that the improved distance r_I is independent of the bare coupling g_0, so that once expressed in terms of the lattice Green's function G, its determination is purely geometric (up to the choice of gauge/static discretisation) and can therefore be tabulated once and used for all ensembles with the same action and smearing.

A similar improved distance can be found for the curvature combination Equation (2.50), which is invariant under an additive shift of $V(r)$ and is dimensionless in natural units. For the continuum Coulomb potential (2.51) one finds

$$c_{\mathrm{cont}}(r) = -\, C_F \frac{g_0^2}{4\pi}. \tag{2.58}$$

On the lattice, $c(r)$ may be approximated by a symmetric second difference, where an improved lattice definition is given by [40]

$$c(\tilde{r}) = \frac{1}{2}\tilde{r}^3 \frac{V(r+a) + V(r-a) - 2V(r)}{a^2}. \tag{2.59}$$

Using Equations (2.58) and (2.59) together with Equation (2.52), this becomes a purely geometric condition,

$$\frac{1}{\tilde{r}^3} = 2\pi\, \frac{G(r+a) + G(r-a) - 2G(r)}{a^2}, \tag{2.60}$$

again with $G(r)$ evaluated for the chosen displacement vector(s). With $\tilde{r}(r)$ determined once and for all (typically by numerical evaluation of the integral (2.53)), one then represents the curvature estimator as a function of $\tilde{r}$,

$$c(\tilde{r}) \equiv \frac{1}{2}\, \tilde{r}^3\, V''(r), \tag{2.61}$$

which removes the leading tree-level discretisation effects of the derivative operator in a manner consistent with the continuum Coulomb limit. This is the form relevant for defining potential-based short-distance couplings from curvature observables in later sections.

The expressions above correspond to the tree-level lattice Coulomb potential associated with the standard Wilson gauge action and unsmeared nearest-neighbour lattice

Laplacian. In the case of our improved action in combination with smeared links one can then write the explicit force at tree level [38, 46, 47] using

$$G_{\mathrm{sm}}(\vec{r}) = \frac{1}{a}\int_{-\pi}^{\pi}\frac{d^3k}{(2\pi)^3}\frac{\prod_{j=1}^{3}\cos(x_j k_j/a)\times f_{\mathrm{sm}}(k)}{4\left(\sum_{j=1}^{3}\sin^2(k_j/2) - 4c_1\sum_{j=1}^{3}\sin^4(k_j/2)\right)} \tag{2.62}$$

in combination with Equations (2.56) and (2.57) to get

$$F_{\mathrm{tree}}(r_I) = -\frac{4\,g_0^2}{3\,a}\int_{-\pi}^{\pi}\frac{d^3k}{(2\pi)^3}\frac{[\cos(rk_1/a)-\cos((r-a)k_1/a)]\times f_{\mathrm{sm}}(k)}{4\left(\sum_{j=1}^{3}\sin^2(k_j/2) - 4c_1\sum_{j=1}^{3}\sin^4(k_j/2)\right)}\,, \tag{2.63}$$

with $C_F = 4/3$ and c_1 in the case of the Symanzik tree-level improved action corresponding to $c_1 = -1/12$, so that one finds [46]

$$r_I^2 F_{\mathrm{tree}}(r_I) = C_F\frac{g_0^2}{4\pi}. \tag{2.64}$$

The smearing factor, coming from the HYP smearing to be introduced in Section 2.7.2, is

$$f_{\mathrm{sm}}(k) = \left[1 - \frac{\alpha_1}{6}\sum_{j=1}^{3}\hat{k}_j^2\,\Omega_{j0}(k)\right]^2, \tag{2.65}$$

with $\hat{k}_j \equiv 2\sin(k_j/2)$ and

$$\Omega_{j0}(k) = 1+\alpha_2(1+\alpha_3) - \frac{\alpha_2}{4}(1+2\alpha_3)\left(\hat{k}_1^2+\hat{k}_2^2+\hat{k}_3^2-\hat{k}_j^2\right) + \frac{\alpha_2\alpha_3}{4}\prod_{\tau\neq 0,j}\hat{k}_\tau^2\,, \tag{2.66}$$

where α corresponds to the smearing parameters discussed further in Section 4.3. Similarly, Equation (2.62) is used in Equation (2.60) for $\tilde{r}$.

The values of r_I and $\tilde{r}$, which now have been defined through tree-level matching conditions involving the lattice Coulomb Green's function, are found through computing the required momentum integral numerically by replacing it with discrete momentum sums on a finite spatial torus of extent L, and then extrapolating to the infinite-volume limit $L/a \to \infty$. Denoting the finite-volume approximation by $G(L/a)$, we evaluate $G(L/a)$ for $L/a = \{128, 256, 512\}$ and perform an extrapolation using

$$G(L/a) = G(\infty) + c_2(a/L)^3 + c_3(a/L)^5,$$

as motivated in Appendix A of [38]. To estimate the systematic uncertainty from the extrapolation, we repeat the analysis with extended ansätze that include higher-order corrections, e.g. $(a/L)^7$ and $(a/L)^9$. The infinite-volume value $G(\infty)$ is then inserted into the defining relation for the improved distances. The resulting values up to $r/a = 20$ can be found in Table 2.1.

r/a	r_I/a	$\tilde{r}/a$
3	2.702	3.247
4	3.568	4.053
5	4.541	5.030
6	5.523	6.015
7	6.510	7.001
8	7.504	7.992
9	8.499	8.987
10	9.497	9.984
11	10.496	10.983
12	11.495	11.982
13	12.495	12.982
14	13.494	13.983
15	14.494	14.983
16	15.494	15.984
17	16.495	16.984
18	17.495	17.985
19	18.495	18.985
20	19.495	19.986

TABLE 2.1: Tabulated tree-level improved distances for HYP2 using the Lüscher-Weiß gauge action: r_I/a for the force estimator $F(r)$ and $\tilde{r}/a$ for the curvature estimator $c(r)$. The uncertainties are below the level of the displayed precision.

2.6 Gradient flow and reference scales

A recurring task in lattice QCD is the conversion of dimensionful quantities measured in lattice units into physical units, as well as the construction of dimensionless ratios that allow for controlled comparisons across ensembles and lattice spacings. Besides hadronic observables and quantities derived from the static potential, an important and widely used class of reference scales is provided by the Wilson/gradient flow [5, 48]. The gradient flow defines a continuous smoothing of the gauge field over a fictitious flow time t, which introduces a physical smearing radius $\sqrt{8t}$. Observables constructed from the flowed gauge field at positive flow time are ultraviolet finite after the usual renormalisation of QCD, and typically exhibit very favourable statistical properties. As a result, flow-based reference scales can be determined with high precision and are commonly used to compare lattice ensembles and to form dimensionless combinations of observables. In the continuum, the gradient flow [5, 48] is defined through the evolution of a gauge field $B_\mu(t,x)$ with flow time $t \geq 0$, with initial condition

$$B_\mu(0,x) = A_\mu(x), \tag{2.67}$$

and governed by the flow equation

$$\partial_t B_\mu(t,x) = D_\nu G_{\nu\mu}(t,x), \tag{2.68}$$

where D_ν and the flowed field-strength tensor $G_{\mu\nu}$ are constructed from $B_\mu(t,x)$. At flow time $t = 0$, one recovers the ordinary gauge field $A_\mu(x)$, so that $G_{\mu\nu}(0,x)$ reduces to the continuum field-strength tensor $F_{\mu\nu}$ introduced in Equation (2.4). For $t > 0$, the flow suppresses ultraviolet fluctuations and smooths the gauge field over a radius

$\sqrt{8t}$, as follows from the Gaussian diffusion kernel of the flow equation at leading order. A standard observable is the action density at flow time t,

$$E(t) = \frac{1}{4} G^a_{\mu\nu}(t,x) G^a_{\mu\nu}(t,x), \tag{2.69}$$

whose expectation value defines the reference scale t_0 via the implicit condition [5]

$$t^2 \langle E(t) \rangle \Big|_{t=t_0} = c, \tag{2.70}$$

with a conventional choice $c = 0.3$. The scale t_0 corresponds to an intermediate smoothing range, where discretisation effects are moderate and the observable can be determined with high precision. In the present work, gradient-flow scales are not determined independently on the ensembles under consideration. Instead, values of t_0 are taken from existing determinations associated with the ensemble set. Their role is therefore auxiliary. Specifically, t_0 is used only as a reference scale to form dimensionless ratios, such as $r_0/\sqrt{t_0}$, and to provide a consistency check of the scale-setting procedure based on the static quark–antiquark potential. The primary scale-setting strategy of this thesis is entirely based on the determination of the potential scale r_0, and flow-based quantities enter the analysis as external inputs.

2.6.1 The Λ-parameter

The strong coupling is a fundamental parameter of QCD, and its value at a given energy scale can be obtained from the renormalisation-group equation once the dimensionful Λ-parameter of the chosen scheme is known. The Λ-parameter sets the intrinsic scale of QCD and determines the running of the coupling across energy scales. In lattice QCD, a direct determination of Λ requires connecting low-energy observables to the high-energy regime where perturbation theory is applicable. This is typically achieved through nonperturbative studies of the running coupling in finite-volume renormalisation schemes, such as those developed by the ALPHA Collaboration [49, 50]. These methods allow for a controlled determination of the dimensionless product $L_{\rm had}\Lambda$, where $L_{\rm had}$ is a reference scale defined at hadronic energies. Since lattice simulations are performed in lattice units, results are often expressed in terms of dimensionless combinations involving reference scales. Common choices include $r_0\Lambda$ or $\sqrt{t_0}\Lambda$, which allow for a direct comparison between different determinations. The conversion to physical units can then be performed once a reference observable is fixed.

In the present work, the Λ-parameter is not determined independently. Instead, external results for $L_{\rm had}\Lambda$ are combined with the potential-based determination of the potential scale r_0. This allows one to express Λ in units of r_0, providing a direct connection between the scale-setting analysis of the static potential and the running of the strong coupling. This construction will be used in Chapter 6.

2.7 Smearing

Many lattice operators relevant to hadron spectroscopy and static-source observables are constructed from gauge links and are therefore sensitive to ultraviolet fluctuations of the gauge field. These short-range fluctuations can increase statistical noise and reduce overlap onto the low-lying states whose long-distance behaviour is of primary interest.

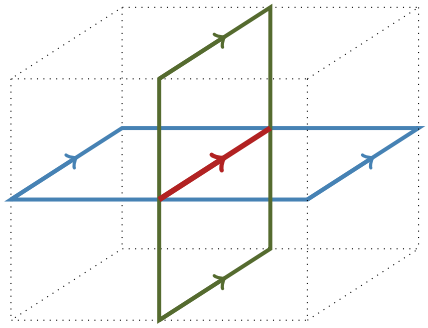

FIGURE 2.1: Schematic illustration of APE smearing for a spatial link $U_i(x)$ (red). The smeared link is constructed as a weighted average of the original link and the staple contributions in the transverse directions $j \neq i$. The blue and green paths indicate the staples associated with the two independent transverse directions, each including forward and backward contributions.

Link smearing addresses this by replacing each link with a gauge-covariant local average built from the link and its neighbouring staples, as illustrated schematically in Figure 2.1, followed by a projection back to $SU(3)$ [18]. Beyond noise reduction, varying the smearing prescription and its parameters naturally generates multiple operators within the same symmetry channel, which is advantageous in variational analyses such as the GEVP. These transformations are applied at the level of operators (and, for static lines, the static discretisation) and do not modify the underlying gauge-field ensemble. In the following we collect the smearing transformations used in this thesis, including APE-type spatial smearing [18], HYP smearing for static links [51], and Teper-style fuzzing/blocking [17]. A broad class of smearing prescriptions can be written in a common form: an *unsmeared* link $U_\mu(x) \in SU(3)$ is replaced by a locally averaged matrix constructed from $U_\mu(x)$ and a set of neighbouring transporters ("staples") that share the same endpoints as $U_\mu(x)$. Gauge covariance is ensured by building the constructed matrix from products of links along paths from x to $x + a\hat{\mu}$, and by subsequently projecting the result back to $SU(3)$ (which we will denote by $\mathrm{Proj}_{SU(3)}\{\cdot\}$).

2.7.1 3D APE smearing

APE smearing is a gauge-covariant smoothing procedure for spatial gauge links, designed to reduce ultraviolet fluctuations and improve the overlap of operators with physical states. Let $S_{\mu\nu}(\vec{x}, t)$ denote a gauge-covariant sum of staples contributing to the link in direction μ. For example, for nearest-neighbour smearing based on plaquette staples one defines

$$\begin{aligned} S_{\mu\nu}(\vec{x}, t) = {} & U_\nu(\vec{x}, t)\, U_\mu(\vec{x} + \hat{\nu}, t)\, U_\nu^\dagger(\vec{x} + \hat{\mu}, t) \\ & + U_\nu^\dagger(\vec{x} - \hat{\nu}, t)\, U_\mu(\vec{x} - \hat{\nu}, t)\, U_\nu(\vec{x} - \hat{\nu} + \hat{\mu}, t). \end{aligned} \tag{2.71}$$

In this work only 3D APE smearing will be used, so in practice the smearing is applied only to spatial directions $\mu = i \in \{1, 2, 3\}$. A generic one-step smeared link is then formed as a linear combination of an unsmeared link and the staples. Denoting by $\tilde{U}_\mu^{(n)}(\vec{x}, t)$ the link after n smearing steps, one APE update is defined by

$$\tilde{U}_\mu^{(n+1)}(\vec{x},t) = \mathrm{Proj}_{SU(3)}\left[U_\mu^{(n)}(\vec{x},t) + \alpha_{\mathrm{APE}} \sum_{\substack{\nu=1 \\ \nu\neq\mu}}^{3} S_{\mu\nu}^{(n)}(\vec{x},t)\right]. \tag{2.72}$$

A schematic illustration of the construction is shown in Figure 2.1. Since the weighted average is not in general unitary, the smeared link is obtained by projection back into $SU(3)$. α_{APE} is a real smearing parameter, typically chosen in the range $0 \leq \alpha_{\mathrm{APE}} \leq 2/3$ for the reasons discussed in [52]. By varying $(\alpha_{\mathrm{APE}}, n)$ one obtains families of gauge-covariant smeared links that suppress ultraviolet fluctuations and improve overlaps of extended operators. This projection and APE smearing is performed using qcdlib library, written by Dr. Tomasz Korzec, that employs the projection corresponding to the prescription described in [16].

2.7.2 HYP smearing (static links)

Hypercubic (HYP) smearing (also referred to as hypercubic blocking) is a gauge-covariant link-smearing transformation in which a thin link $U_\mu(x) \in \mathrm{SU}(3)$ is replaced by a locally averaged "fat" link $V_\mu(x) \in \mathrm{SU}(3)$ constructed from staples confined to the hypercubes adjacent to the original link as illustrated in Figure 2.2. The elementary HYP smearing transformation is restricted to links within the adjacent hypercube, making each iteration ultralocal, while efficiently suppressing short-distance fluctuations [51]. Repeated HYP steps can then be applied to achieve stronger smoothing while retaining good locality properties. The construction proceeds in three nested APE-like steps with intermediate projections back to SU(3). To simplify the notation, sums of the form $\sum_{\pm\nu\neq\mu}$ run over all directions orthogonal to μ, including both orientations. Following the standard definition, one introduces "decorated" links $\widetilde{V}_{\mu;\nu}(x)$ and $\bar{V}_{\mu;\nu\rho}(x)$, where the semicolon indicates directions that are *excluded* from the staple construction at that level. The innermost level reads

$$\bar{V}_{\mu;\nu\rho}(x) = \mathrm{Proj}_{SU(3)}\left[(1-\alpha_3)\,U_\mu(x) + \frac{\alpha_3}{2} \sum_{\pm\eta\neq\mu,\nu,\rho} U_\eta(x)\,U_\mu(x+a\hat{\eta})\,U_\eta^\dagger(x+a\hat{\mu})\right]. \tag{2.73}$$

Using these links, the second level is

$$\widetilde{V}_{\mu;\nu}(x) = \mathrm{Proj}_{SU(3)}\left[(1-\alpha_2)\,U_\mu(x) + \frac{\alpha_2}{4} \sum_{\pm\rho\neq\mu,\nu} \bar{V}_{\rho;\nu\mu}(x)\,\bar{V}_{\mu;\rho\nu}(x+a\hat{\rho})\,\bar{V}_{\rho;\nu\mu}^\dagger(x+a\hat{\mu})\right]. \tag{2.74}$$

Finally, the HYP-smeared link is obtained as

$$V_\mu(x) = \mathrm{Proj}_{SU(3)}\left[(1-\alpha_1)\,U_\mu(x) + \frac{\alpha_1}{6} \sum_{\pm\nu\neq\mu} \widetilde{V}_{\nu;\mu}(x)\,\widetilde{V}_{\mu;\nu}(x+a\hat{\nu})\,\widetilde{V}_{\nu;\mu}^\dagger(x+a\hat{\mu})\right]. \tag{2.75}$$

Equations (2.73)–(2.75) define a single HYP smearing step $U \mapsto V$; the transformation may be iterated to generate operator families with different overlap properties. The explicit projection prescription is not unique. The implementation used in this thesis follows Ref. [43]. In the static limit, HYP smearing is most commonly applied to

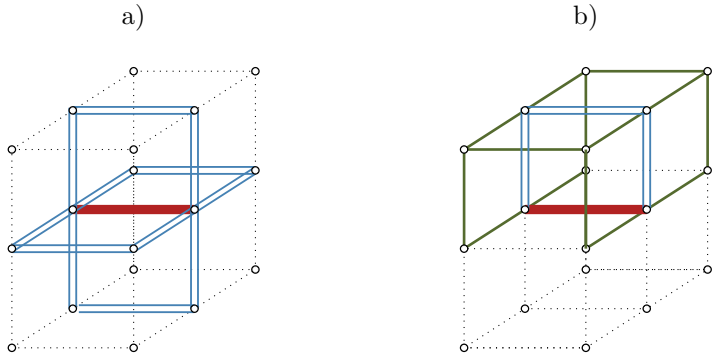

FIGURE 2.2: Schematic illustration of the nested structure of HYP smearing. a) The fat link (red) is constructed from staple contributions (blue) within the attached hypercubes. b) The links entering these staples are themselves built from restricted staples (green) confined to lower-dimensional hypercubes. Redrawn from [51].

the temporal links entering the static Wilson line: one replaces $U_0(x)$ by $V_0(x)$ in the definition of the straight temporal transporter. This yields an alternative (still local) lattice discretisation of the static action, while preserving gauge invariance of the resulting Wilson-loop observables.

2.7.3 Teper fuzzing/blocking (spatial links)

Teper-style fuzzing/blocking introduced in [17] constructs spatially extended gauge-covariant transporters by iterating a local "straight + staples" rule on each fixed timeslice. Similarly as in APE/HYP smearing, the resulting blocked variables need not be elements of $SU(3)$; gauge covariance is sufficient because only closed traces enter the operators of interest [17]. It has, however, been shown in [53] that it is advantageous to project a smeared link back into $SU(3)$ to reduce noise and avoid overflow. The key idea is to iteratively build *composite spatial links* (or paths) that incorporate contributions from an exponentially large number of elementary paths, thereby generating operators with increasing spatial extent while maintaining gauge covariance. Schematic illustration of a single Link is seen in Figure 2.3. Starting from the original spatial link variables one defines blocked links $U_i^{(N)}(x)$ recursively. At each blocking level N, the new link in direction i is constructed as a sum of the straight path and spatial staples built from links at level $N-1$ as

$$\begin{aligned} U_i^{(N)}(x) =& U_i^{(N-1)}(x)\, U_i^{(N-1)}(x + l_{N-1}\hat{i}) \\ &+ \sum_{j\neq i} U_j^{(N-1)}(x)\, U_i^{(N-1)}(x + l_{N-1}\hat{j})\, U_j^{(N-1)\,\dagger}(x + l_{N-1}\hat{i}), \end{aligned} \tag{2.76}$$

where the sum runs over spatial directions only and is illustrated in Figure 2.3. Here l_N denotes the effective length of the blocked link, which grows exponentially with the blocking level,

$$l_N = 2^{N-1}a. \tag{2.77}$$

Thus, each iteration doubles the physical extent of the link, allowing one to match the operator size to the physical size of the state under investigation. An important feature of this construction is that the number of underlying elementary paths contributing to

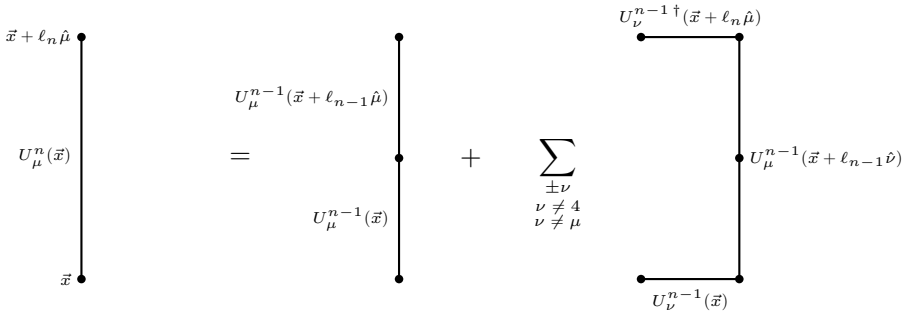

FIGURE 2.3: Schematic illustration of one blocking step in the Teper construction. A blocked transporter of level n is built from the straight product of two level-$(n-1)$ transporters in direction μ and from staple-shaped contributions involving transverse directions $\nu \neq \mu$.

$U_i^{(N)}(x)$ grows extremely rapidly with N, leading to a dense sampling of paths within a given spatial volume. As emphasized in [17], this effectively generates wave functionals with the appropriate physical size, claiming an improvement of the projection onto the lowest-lying states. Gauge-invariant operators are then constructed from these blocked links in the usual way, e.g. through closed Wilson loops built from $U_i^{(N)}(x)$. Importantly, only spatial links are blocked, while temporal links remain unchanged. This preserves the transfer matrix interpretation discussed in Sec. 2.4 and ensures that the spectral decomposition of correlation functions remains valid. We then apply a projection step back onto SU(3) after each iteration to control the growth of the link variables and maintain numerical stability. Early applications of extended fuzzy-link operators in glueball spectroscopy demonstrated their usefulness for improving overlap with low-lying glueball states and reducing statistical noise [16].

Chapter 3

Computational methodology and ensemble properties

The determination of a reference scale in lattice QCD requires gauge ensembles with controlled discretisation effects, large physical volumes, and quark masses close to the physical point. In this thesis, the collaborative effort of the Coordinated Lattice Simulations (CLS) $N_f = 2 + 1$ gauge ensembles [54, 55] form the numerical basis for the non-perturbative determination of the potential scale r_0. The CLS ensembles are designed to enable reliable continuum and chiral extrapolations, which is essential for the scale-setting strategy adopted in this work, based on long-distance quantities derived from the static quark–antiquark potential. The CLS effort aims at the generation of high-quality gauge ensembles, characterised by fine lattice spacings, large physical volumes, and pion masses approaching the physical point. At the same time, a uniform simulation strategy is maintained, using a consistent choice of lattice action and parameters across ensembles, which facilitates controlled continuum and chiral extrapolations. In the present generation, CLS employs $N_f = 2 + 1$ flavours of dynamical quarks, with degenerate light quarks and a separately tuned strange quark. A central design feature is the use of open boundary conditions in the temporal direction, which allows simulations at fine lattice spacings by avoiding topological freezing.

3.1 Lattice action and improvement

This section summarises the lattice discretisation used for the generation of the CLS ensembles. The choice of gauge and fermion actions is guided by the need to control cutoff effects in long-distance observables, which are central to the scale-setting analysis presented in this thesis.

3.1.1 Gauge action

The gluonic degrees of freedom are described by the tree-level improved Lüscher–Weisz gauge action [31] introduced in Equation (2.29),

$$S_g[U] = \frac{\beta}{6}\left(c_0 \sum_p \mathrm{Tr}\{1 - U(p)\} + c_1 \sum_r \mathrm{Tr}\{1 - U(r)\}\right), \tag{3.1}$$

where $U(p)$ and $U(r)$ denote the plaquette and rectangular Wilson loops, respectively and $\beta = 6/g_0^2$, containing the bare gauge coupling g_0. The coefficients are constrained

by the normalisation condition [46, 56]

$$c_0 + 8c_1 = 1, \tag{3.2}$$

and for the tree-level Symanzik action used are $c_0 = 5/3$ and $c_1 = -1/12$. This action reduces $\mathcal{O}(a^2)$ cutoff effects compared to the Wilson plaquette action, which is advantageous for scale setting based on long-distance observables.

3.1.2 Fermion action and $\mathcal{O}(a)$ improvement

The fermionic sector is formulated using $\mathcal{O}(a)$-improved Wilson fermions. For each flavour f, the Dirac operator reads

$$D_{\mathrm{W}}(m_{0,f}) = D_{\mathrm{W}}^{(0)}(m_{0,f}) + c_{\mathrm{SW}}\frac{ia}{4}\sigma_{\mu\nu}F_{\mu\nu}, \tag{3.3}$$

where $D_{\mathrm{W}}^{(0)}$ is the unimproved Wilson–Dirac operator [3] and c_{SW} denotes the Sheikholeslami–Wohlert coefficient [30], determined non-perturbatively, $\sigma_{\mu\nu} = \frac{i}{2}[\gamma_\mu, \gamma_\nu]$ is the antisymmetric spin tensor, and $F_{\mu\nu}$ is a lattice discretisation of the continuum gauge field-strength tensor introduced in Section 2.1.1. The fermion action for $N_f = 2 + 1$ flavours is given by

$$S_f = a^4 \sum_x \sum_{f=u,d,s} \bar{\psi}_f(x) D_{\mathrm{W}}(m_{0,f}) \psi_f(x), \tag{3.4}$$

with degenerate up and down quark masses,

$$m_{0,u} = m_{0,d} \equiv m_{0,\ell}. \tag{3.5}$$

The strange quark mass $m_{0,s}$ is tuned as a function of the light quark mass, corresponding to a trajectory of constant physics. Together, these discretisation choices define the lattice theory underlying the CLS ensembles and provide the basis for the tuning strategy

3.2 Simulation parameters and tuning strategy

The corresponding bare parameters are tuned within a hadronic renormalisation scheme designed to enable controlled matching across lattice spacings. The tuning strategy is designed to minimise systematic uncertainties associated with renormalisation and to enable a controlled matching of ensembles at different lattice spacings. Since CLS has generated data using $N_f = 2 + 1$ flavours, electromagnetic and isospin breaking effects as well as contributions from heavier quarks are not taken into account in the calculations

3.2.1 Hadronic renormalisation scheme

Matching dimensionless combinations of low-energy hadronic observables allows the bare parameters of the theory to be fixed within a hadronic renormalisation scheme. This approach avoids the use of intermediate renormalisation constants and enables tuning quantities to be determined directly from lattice measurements. Residual chiral extrapolations may still be required to reach the physical point. In practice,

pseudoscalar meson masses, the isosymmetric pion and kaon, are combined with the Wilson flow scale t_0, introduced in Section 2.6, to define dimensionless mass parameters,

$$\phi_2 = 8t_0 m_\pi^2, \qquad \phi_4 = 8t_0 \left(m_K^2 + \tfrac{1}{2} m_\pi^2\right), \tag{3.6}$$

which, at leading order in chiral perturbation theory, are proportional to the sum of the light-quark masses and to the trace of the quark mass matrix, respectively [57, 58]. The Ref. [54], which this chapter mainly uses to discuss the CLS setup, used $\sqrt{8t_0} = 0.4144(59)(37)$ by the BMW collaboration using $N_f = 2+1$ flavours [6] together with the QCD values of the pion and kaon masses in the isosymmetric limit without electromagnetic contributions taken from Ref. [59] being $m_\pi = 134.8(3)$MeV and $m_K = 494.2(4)$MeV to estimate ϕ_2 and ϕ_4. Since then Chiral extrapolation has been performed with an increased set of ensembles by Ref. [60], which found the flow scale to be

$$\sqrt{t_0^{\text{phys}}} = 0.1443(7)(13)\,\text{fm}, \tag{3.7}$$

and values of ϕ_2 and ϕ_4 to be

$$\phi_2^{\text{phys}} = 0.0779(7), \qquad \phi_4^{\text{phys}} = 1.098(10). \tag{3.8}$$

3.2.2 Chiral trajectories and tuning

In simulations with $\mathcal{O}(a)$-improved Wilson fermions, the relation between the bare gauge coupling and the lattice spacing is affected by mass-dependent cutoff effects. In particular, within the Symanzik improvement framework, the improved bare coupling receives corrections proportional to the sum of the sea quark masses. As a consequence, variations of the quark masses at fixed bare coupling generally induce changes in the lattice spacing at $\mathcal{O}(am_q)$ [54, 61]. To control these effects CLS adopts the choice of chiral trajectories in the space of bare quark masses, that keep the sum over the bare quark masses fixed,

$$\text{Tr}\, M_q = 2m_{0,\ell} + m_{0,s} = \text{const}, \tag{3.9}$$

where $m_{0,\ell}$ denotes the degenerate up and down quark mass and $m_{0,s}$ the strange quark mass. Up to cutoff effects of $\mathcal{O}(a^2)$, this condition implies that the lattice spacing remains constant along a given trajectory. This property is particularly advantageous for scale-setting studies, since it helps disentangle quark-mass dependence from lattice-spacing effects when comparing dimensionless observables along a trajectory. In practice, ensembles at different values of the bare coupling are matched at a point close to the flavour-symmetric line,

$$m_{0,\ell} = m_{0,s}, \tag{3.10}$$

where the tuning of the quark masses is numerically stable, in the sense that simulations remain well behaved and statistical fluctuations are under control. Starting from this matching point, trajectories towards the physical quark mass point are defined by varying the light and strange quark masses while maintaining a constant trace condition. Nonetheless, small mistunings of the quark masses are unavoidable in practice. As a result, the ensembles at the flavour-symmetric point do not all correspond to exactly identical values of ϕ_2, and the chiral trajectories do not maintain strictly values of ϕ_4. In earlier studies, such mistunings were corrected by explicitly computing quark-mass derivatives of the observables and using these to shift the results accordingly [62].

In the present work, however, the focus lies on purely gluonic observables and on dimensionless ratios constructed from them. For these quantities, the dependence on the quark masses is sufficiently weak that the effects of the residual mistunings can be neglected at the level of statistical precision achieved. This behaviour has already been observed in Ref. [62], where the required corrections to the gluonic reference scale t_0 were found to be negligible. Consistent with this, we observe that potential-based scales exhibit a similarly mild quark-mass dependence on the ensembles for which quark-mass derivatives have been computed. The resulting shifts are small compared to the corresponding statistical uncertainties. For this reason, we do not apply explicit corrections based on quark-mass derivatives, which tend to be noisy in practice. Instead, residual mistunings are accounted for by including terms proportional to deviations from the target value of ϕ_4 from Eq. (3.8) in selected fit ansätze.

3.3 Algorithmic treatment of the quark determinant

The ensembles employed in this work are generated using a Hybrid Monte Carlo [63] (HMC) algorithm with different numerical treatments for the light and strange quark sectors. These choices are motivated by numerical efficiency and stability considerations discussed in detail in [54]. The two degenerate light quarks are simulated using an even–odd preconditioned [64] HMC algorithm combined with a factorisation of the fermion determinant [65–67]. In addition, a small twisted-mass regulator is introduced in the light-quark sector during the simulation to suppress near-zero modes of the Dirac operator [68]. The strange quark is treated separately using a Rational Hybrid Monte Carlo (RHMC) algorithm [69]. The simulated gauge field ensemble differs slightly from the target distribution. Expectation values of observables are therefore corrected by the inclusion of appropriate reweighting factors. These account for the effects of twisted-mass reweighting [68, 70], residual errors of the rational approximation, and the occurrence of a negative strange-quark determinant [71].

3.4 Boundary conditions and their impact on scale observables

In lattice QCD, periodic boundary conditions are used in essentially all older works and are still most commonly used, as they preserve Euclidean translational invariance and minimise boundary artefacts, allowing finite-volume simulations to more closely approximate infinite-volume QCD. But open boundary conditions are now standard in specific regimes. In studies [72, 73] it has been shown that periodic boundary conditions with $a \to 0$ lead to severe autocorrelations and topological freezing as the action barrier between topological sectors grows and HMC transitions between sectors become exponentially suppressed. The ergodicity of the Markov-chain Monte Carlo (MCMC) simulation, which means that for infinite Monte Carlo time any configuration with nonzero Boltzmann weight can be reached, is no longer given, since the Markov chain effectively cannot transition between topological sectors if topological freezing is occurring [67]. Thereby long-range correlations could be invisible in the Monte Carlo history and the generated gauge configurations may not constitute a statistically representative sample of the target QCD ensemble. To mitigate topological freezing at small lattice spacings, open boundary conditions are imposed in the temporal

direction [72]. Gauge fields satisfy

$$F_{0k}(x)\Big|_{x_0=0,T} = 0, \tag{3.11}$$

where $F_{\mu\nu}(x)$ denotes the gauge-field tensor, defined in Equation (2.4), while in the fermion fields Dirichlet-type boundary conditions are imposed with

$$P_+\,\psi(x)\Big|_{x_0=0} = 0, \qquad P_-\,\psi(x)\Big|_{x_0=T} = 0, \tag{3.12}$$

with the projectors defined as

$$P_\pm = \tfrac{1}{2}\left(1 \pm \gamma_0\right). \tag{3.13}$$

The continuum topological charge is defined as

$$Q \equiv \int d^4x\, q(x), \qquad q(x) = \frac{1}{32\pi^2}\,\epsilon_{\mu\nu\rho\sigma}\,\mathrm{tr}[F_{\mu\nu}(x)\,F_{\rho\sigma}(x)]\,. \tag{3.14}$$

With periodic boundary conditions, gauge configurations are separated into topological sectors labelled by an integer-valued topological charge Q. For open temporal boundary conditions, however, the topological charge is no longer quantized, since topology can continuously flow through the temporal boundaries. This removes the topology barriers responsible for topological freezing at fine lattice spacings. Open boundaries break exact translational invariance in the temporal direction, which has implications for observables constructed from Wilson loops. In the determination of the static potential, temporal separations are therefore chosen sufficiently far from the boundaries such that boundary effects are exponentially suppressed. Most CLS ensembles employed in this work, given in Table 3.1, were generated with open temporal boundary conditions to mitigate topological freezing at fine lattice spacings, while only a smaller subset uses periodic boundaries. The treatment of open-boundary effects on the ensembles will be discussed further in Section 4.5.

3.5 Ensemble characteristics

The CLS ensembles used in this work consist of $N_f = 2+1$ gauge configurations generated with $\mathcal{O}(a)$-improved Wilson fermions and primarily open boundary conditions in the temporal direction. They span five lattice spacings in the range $a \approx 0.039$–0.085 fm, allowing for controlled continuum extrapolations. In lattice QCD, both ultraviolet (UV) and infrared (IR) cutoff effects must be controlled. The lattice spacing a introduces a UV cutoff of order $p_{\max} \sim \pi/a$, limiting the resolution of short-distance physics and leading to discretisation effects. Conversely, the finite spatial extent L acts as an IR cutoff, with the smallest non-zero momentum given by $p_{\min} \sim 2\pi/L$, which governs finite-volume effects. Ideally, one would like to realise the hierarchy

$$\frac{2\pi}{L} \ll \Lambda_{\mathrm{QCD}} \ll \frac{\pi}{a},$$

corresponding to large volumes and fine lattice spacings. In practice, this is constrained by computational cost, which grows rapidly with the number of lattice points, $(L/a)^4$. As a result, cutoff effects cannot be eliminated completely, but are instead controlled and quantified. A commonly used criterion [74] to suppress finite-volume effects in

hadronic observables is that the spatial volume satisfies

$$m_\pi L \gtrsim 4. \tag{3.15}$$

The pion masses range from approximately 420 MeV down to the physical pion mass, following the chiral trajectories described in Section 3.2.2. This coverage in lattice spacing, volume, and quark mass allows for controlled interpolations to the physical point and provides a solid basis for the scale setting analysis presented in the following chapters. An overview of the CLS ensembles employed in this work is given in Table 3.1 and visualized in Figure 3.1 in respect to the pion mass, lattice spacing and $m_\pi L$ relationship

3.6 Autocorrelations and error analysis

Monte Carlo simulations of lattice QCD generate ensembles of gauge field configurations through a Markov process, implying that successive measurements can have a certain amount of correlation, which is known as autocorrelation. A reliable determination of statistical uncertainties therefore requires a careful treatment of autocorrelations, in particular for simulations at fine lattice spacings, where critical slowing down is expected to become increasingly severe [75]. Critical slowing down refers to the rapid growth of autocorrelation times as the lattice spacing is reduced, making successive configurations increasingly correlated. In this work we use the Γ-method [76] for the error analysis of the data of auto-correlated Monte-Carlo time series, to take autocorrelations into account. It allows for a controlled determination of statistical errors for both primary observables and derived quantities. Compared to binning/blocking techniques, it provides more robust error estimates and a clearer separation of statistical and systematic effects associated with finite Monte Carlo statistics

3.6.1 Autocorrelation functions and integrated autocorrelation times

One cannot assume that generated data are uncorrelated and to calculate the autocorrelation is the best way to check that. Given a sequence of measurements $\{O_i\}$ of an observable O obtained along a Markov chain, the autocorrelation function is defined as

$$\Gamma_O(t) = \langle (O_i - \langle O \rangle)(O_{i+t} - \langle O \rangle) \rangle, \tag{3.16}$$

where t denotes the separation in Monte Carlo time. Here $\langle \cdot \rangle$ denotes the path integral averages. For most lattice observables, the autocorrelation function exhibits an initial rapid decay at small t, dominated by fast modes of the Markov process, followed by a slower exponential tail at large t governed by the slowest modes. The associated integrated autocorrelation time,

$$\tau_{\mathrm{int}}(O) = \frac{1}{2} + \sum_{t=1}^{\infty} \frac{\Gamma_O(t)}{\Gamma_O(0)}, \tag{3.17}$$

quantifies the effective number of Monte Carlo updates over which measurements remain correlated. It is commonly stated, that independent measurements are separated by $\sim 2\tau_{\mathrm{int}}$. Autocorrelations increase the variance of the estimator for the expectation

value of O according to

$$\mathrm{var}(\bar{O}) = \frac{2\tau_{\mathrm{int}}(O)}{N}\Gamma_O(0), \tag{3.18}$$

where N is the total number of measurements. The quantity $2\tau_{\mathrm{int}}(O)$ thus measures the reduction in the effective number of statistically independent measurements due to autocorrelations. For the topological charge with periodic boundary conditions at small lattice spacing, as mentioned in 3.4, the autocorrelation might become extremely large or unmeasurable, as the sum defining the autocorrelation might not converge reliably. This would break down the error estimation and the ensembles may no longer be representative.

3.6.2 The Γ-method for error estimation

In practice, the autocorrelation function must be estimated from a finite Monte Carlo history and summed only up to a finite window size W [76, 77]. In the Γ-method, the integrated autocorrelation time is therefore estimated as

$$\tau_{\mathrm{int}}(O, W) = \frac{1}{2} + \sum_{t=1}^{W} \frac{\Gamma_O(t)}{\Gamma_O(0)}. \tag{3.19}$$

An automatic windowing procedure is employed by Ref. [76] such that the combined statistical uncertainty and the systematic truncation error of the sum are minimised. To take slowly decaying modes of the auto-correlation functions into accounts, a tail is added onto the autocorrelation time, as proposed in Ref. [75]. The Γ-method generalises straightforwardly to derived observables, such as fit parameters or ratios, by projecting the autocorrelation functions of the primary observables onto the derived quantities. It is described further in [78–80] how the derived observables are treated by having mean values and the corresponding projected fluctuations stored in suitable data structures to allow for a straightforward linear propagation of uncertainties while fully accounting for correlations and autocorrelations. The only practical limitation observed with this approach arises in rare cases from its memory requirements. In order to retain sufficient information for a reliable treatment of autocorrelations in the error propagation, the method stores a large amount of data associated with the derived observables. For data sets comprising on the order of several tens of thousands of configurations, and in particular for analyses involving large correlation matrices, this can lead to substantial memory consumption, which may become a limiting factor on standard desktop machines. All statistical errors presented in this thesis are obtained using the Γ-method, and autocorrelations are taken into account consistently throughout the analysis.

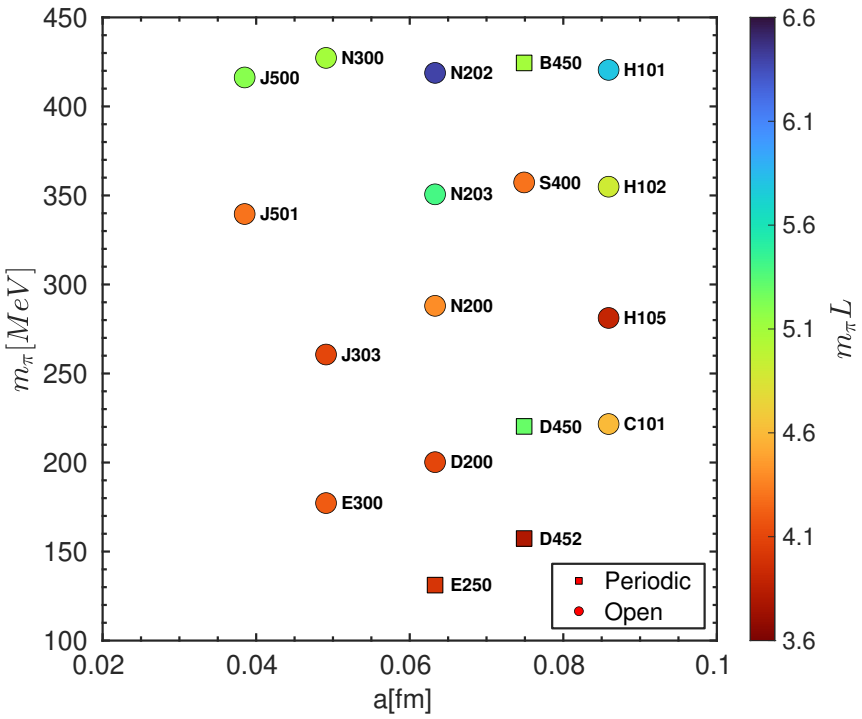

FIGURE 3.1: The CLS $N_f = 2 + 1$ ensembles from Table 3.1 visualized on a lattice spacing and pion mass plane with the colours corresponding to the spatial volume with open or closed boundaries corresponding to the form of the marker

β	a [fm]	id	$\frac{T}{a} \times \frac{L^3}{a^3}$	b.c.	$\kappa_{u,d}$	κ_s	m_π[MeV]	$m_\pi L$	stat. [MDU]	τ_{exp} [MDU]
3.40	0.0849(5)(8)	H101	96×32^3	o	0.13675962	0.13675962	420	5.8	8064	26.6
		H102-0	96×32^3	o	0.136865	0.136549339	353	4.9	3988	26.6
		H102-1	96×32^3	o			356	4.9	4032	26.6
		H105	96×32^3	o	0.13697	0.13634079	281	3.9	7880	26.6
		C101	96×48^3	o	0.137055	0.136172656	221	4.6	8000	26.6
3.46	0.0749(4)(7)	B450	64×32^3	p	0.13689	0.13689	424	5.1	6448	32.9
		S400	128×32^3	o	0.136984	0.136702387	357	4.3	11492	32.9
		D450	128×64^3	p	0.137126	0.136420428639937	220	5.3	2000	32.9
		D452	128×64^3	p	0.137163675	0.136345904546	157	3.8	3996	32.9
3.55	0.0633(4)(6)	N202	128×48^3	o	0.137	0.137	418	6.4	7608	47.3
		N203	128×48^3	o	0.13708	0.136840284	350	5.4	6172	47.3
		N200	128×48^3	o	0.13714	0.13672086	288	4.4	6848	47.3
		D200	128×64^3	o	0.1372	0.136601748	200	4.1	8004	47.3
		E250	192×96^3	p	0.137232867	0.136536633	131	4.0	3800	47.3
3.70	0.0491(3)(4)	N300-0	128×48^3	o	0.137	0.137	427	5.1	2028	77.9
		N300-1	128×48^3	o			427	5.1	5916	77.9
		J303	192×64^3	o	0.137123	0.1367546608	260	4.1	8584	77.9
		E300	192×96^3	o	0.137163	0.1366751636177327	177	4.2	4548	77.9
3.85	0.0385(2)(3)	J500	192×64^3	o	0.136852	0.136852	416	5.2	11552	126.5
		J501	192×64^3	o	0.1369032	0.136749715	339	4.3	14236	126.5

TABLE 3.1: CLS ensembles used in this work. The ensembles are labelled by an id. For the H102 and N300 ensembles, runs with the same physical but different simulation parameters exist. The table furthermore lists the inverse bare coupling β, the lattice sizes, the boundary conditions in time direction (b.c.), where "o" stands for open and "p" for (anti-)periodic, the bare quark masses parametrized by the hopping parameters for up/down and strange quarks, the pion mass in MeV and the statistics in molecular dynamic units. The pion mass and lattice spacings are taken from [60]. More recently the lattice spacings have been calculated using the CLS ensembles by the Wilson flow scale from the octet cascade (Ξ Baryon) [81] The final entry is a rough estimate of the exponential autocorrelation time, which is used during error analysis to attach "tails" to auto-correlation functions.

Part ϖ

The Two Potential Scales

Chapter 4

The static quark–antiquark potential

The static quark–antiquark potential is a fundamental non-perturbative observable of Quantum Chromodynamics. It characterises the interaction energy of a static colour source-anti-source pair as a function of their spatial separation and plays a central role in the study of confinement. Beyond its physical significance, this potential provides a practical foundation for defining reference scales, such as the potential scale r_0 [7], which is a central focus of this thesis. In this chapter, we describe how the static potential is computed on the CLS $N_f = 2 + 1$ gauge ensembles introduced in Chapter 3. The discussion proceeds from the definition of the potential and its lattice representation to the practical aspects of its measurement and extraction. Particular attention is paid to the challenges associated with Wilson-loop measurements in full QCD and to the strategies employed to overcome them. These considerations motivate the use of operator improvement techniques, most notably hypercubic (HYP) smearing [51], which are essential for obtaining a reliable signal for the static potential on dynamical ensembles. The chapter is organised as follows. First, the static potential and its relation to Wilson loops are introduced. The specific difficulties encountered in measuring Wilson loops in the presence of dynamical fermions are then discussed, providing the motivation for the smearing and variational techniques described subsequently. Finally, the static potential is extracted from the Wilson-loop data and used to determine the potential scales.

4.1 Static potential from Wilson loops

In lattice QCD, the static quark–antiquark potential, introduced in Section 2.5, is extracted from rectangular Wilson loops. For a spatial separation r and Euclidean time extent T, the corresponding observable is defined as

$$W(r,T) = \left\langle \operatorname{Tr} \mathcal{P} \prod_{(x,\mu)\in C(r,T)} U_\mu(x) \right\rangle, \tag{4.1}$$

where $C(r,T)$ denotes a rectangular contour of spatial extent r and temporal extent T. In the transfer-matrix formalism, the Wilson loop admits a spectral decomposition analogous to Equation (2.37),

$$W(r,T) = \sum_n |c_n(r)|^2 \, e^{-V_n(r)T}, \tag{4.2}$$

where $V_n(r)$ are the energy levels of the static quark–antiquark system and $c_n(r)$ quantify the overlap of the chosen operators with the corresponding eigenstates. In the limit of large Euclidean time, and provided thermal contributions from the finite temporal extent of the lattice are negligible, excited-state contributions are exponentially suppressed, and the ground-state potential $V(r) = V_0(r)$ can be extracted. The extraction of the static potential from Wilson loops thus requires both sufficient suppression of excited-state contributions and adequate statistical precision at large temporal separations.

4.2 Challenges of Wilson-loop measurements in full QCD

On the lattice, the static quark–antiquark potential is typically extracted from the large-time behaviour of rectangular Wilson loops, although alternative definitions, such as those based on Polyakov loop correlators, can also be considered. In practice, the direct measurement of Wilson loops poses significant challenges, especially in simulations with dynamical fermions. A generic difficulty arises from the exponential degradation of the signal-to-noise ratio with increasing temporal extent and spatial separation of the loop [82]. While the expectation value of a Wilson loop decays exponentially with its area, the variance is dominated by ultraviolet fluctuations, leading to rapidly increasing statistical uncertainties at large distances. This problem is already present in pure gauge theory and becomes more difficult to handle in full QCD. In pure gauge simulations, this issue can be mitigated using the multi-hit (or one-link) method [83] and, more efficiently, multilevel noise-reduction techniques [84], which exploit the locality of the gauge action to factorise contributions from different regions of the lattice. However, once dynamical fermions are included, the fermion determinant induces non-local couplings between gauge fields across time slices, preventing a straightforward application of such methods [38]. As a consequence, alternative strategies are required for Wilson-loop measurements in full QCD. A second, closely related challenge is the contamination from excited states of the static quark–antiquark system. At finite temporal extent, Wilson loops receive contributions from a tower of energy eigenstates, and isolating the ground-state potential requires either very large Euclidean times or operators with sufficiently good overlap onto the ground state. In practice, the former is limited by the rapidly deteriorating signal-to-noise ratio, making operator optimisation indispensable.

4.3 HYP smearing

The challenges outlined in the previous section motivate the use of operator improvement techniques that reduce statistical noise and suppress excited-state contamination in Wilson-loop measurements. In this work, this is achieved through the use of hypercubic (HYP) smearing of gauge links, defined in Section 2.7.2 which follows the approach introduced in Ref. [51] and developed further for observables with static quarks in dynamical lattice QCD in Refs. [43, 85]. HYP smearing has been shown to significantly reduce the noise of Wilson loops both in pure gauge theory and in simulations with dynamical fermions. In pure gauge theory, it was demonstrated in Ref. [86] that Wilson loops constructed from HYP-smeared links achieve a precision comparable to that obtained with the multi-hit method. For dynamical fermions, Ref. [38] showed that the resulting static potential agrees with the continuum potential up to discretisation effects of $\mathcal{O}(a^2)$ after renormalisation.

HYP smearing is a gauge-covariant procedure that replaces the original link variables by locally averaged links constructed from staples contained within a hypercube attached to the link. The smearing is performed in a sequence of steps with fixed parameters that control the relative weights of the staple contributions, followed by a projection back onto the SU(3) group. By construction, HYP smearing preserves gauge invariance while efficiently reducing short-distance fluctuations of the gauge field. In practice, first all gauge links entering the static quark propagators are HYP2 smeared, using the parameter set $\alpha_2 = 1.0$, and $\alpha_3 = 0.5$ [43]. This choice corresponds to a particular discretisation of the static quark action within lattice heavy quark effective theory and may be viewed as a modification of the original Eichten–Hill formulation [87]. In addition, the spatial transporters entering the Wilson loops are subjected to l iterations of three-dimensional HYP smearing, using parameters $\alpha_1 = 0.75, \alpha_2 = 0.6$, and $\alpha_3 = 0.3$. Varying the number of smearing iterations allows the construction of a basis of creation and annihilation operators with different overlaps onto the energy eigenstates of the static quark–antiquark system, which forms the starting point for a variational analysis. After each smearing step, the resulting gauge links are projected back onto the $SU(3)$ group. The projection used is described in Ref. [43], employing Eq. (2.24) and four iterations of Eq. (2.25) therein.

The smeared Wilson loop can be written as

$$W^{(l,k)}(x_0, r, T) = \frac{a^3}{L^3} \sum_{\mu=1}^{3} \sum_{\vec{x}} \mathrm{ReTr}$$

$$\left(\frac{\langle\, \mathcal{W}^{(l)}(x, x + r\hat{\mu})\, \mathcal{W}^{(0)}(x + r\hat{\mu}, x + r\hat{\mu} + T\hat{0}) \times \mathcal{W}^{(k)\dagger}(x + T\hat{0}, x + r\hat{\mu} + T\hat{0})\, \mathcal{W}^{(0)\dagger}(x, x + T\hat{0})\, w \,\rangle}{\langle w \rangle} \right), \tag{4.3}$$

where $\mathcal{W}^{(l)}(x, y)$ stands for a smeared straight Wilson line from x to y and w is the overall reweighting factor, as described in section 3.3. The Wilson loops are measured according to the procedure outlined in Ref. [38], with the implementation provided by the `wloop` package [88]. The smearing levels were initially chosen following the scaling prescription of Ref. [38],

$$n_l \approx \frac{l}{12} \left(\frac{r_0}{a} \right)^2 . \tag{4.4}$$

The values listed in Table 4.1 follow this prescription approximately for the coarser lattice spacings, corresponding to the choices $l \approx 3, 5$. For the finer ensembles, however, the increase of the smearing levels was moderated and the values were chosen empirically. Tests performed on ensemble `E300`, discussed in Chapter A, indicate that the inclusion of substantially larger smearing levels does not lead to a significant improvement of the operator basis, while increasing the numerical cost of the Wilson-loop analysis.

4.4 Variational analysis and extraction of energy levels

At fixed spatial separation r, Wilson loops constructed with different levels of spatial smearing define a matrix of correlation functions. Assuming that the Euclidean time coordinates x_0 and $x_0 + T$ are sufficiently far from the temporal lattice boundaries, each element of this correlation matrix admits a spectral decomposition similar as

β	id	$T_{\max}/a$	$r_{\max}/a$	smearing	N_{meas}
3.40	H101	24	16	8, 14	2016
	H102-0	24	16	8, 14	997
	H102-1	24	16	8, 14	1008
	H105	24	16	8, 14	1970
	C101	24	16	8, 14	2000
3.46	B450	24	16	11, 18	1612
	S400	24	16	11, 18	2873
	D450	24	16	11, 18	500
	D452	24	16	11, 18	999
3.55	N202	24	16	15, 25	1902
	N203	24	16	15, 25	1543
	N200	24	16	15, 25	1712
	D200	24	16	15, 25	2001
	E250	24	16	15, 25	950
3.70	N300-0	28	20	20, 30	507
	N300-1	28	20	20, 30	1479
	J303	28	20	20, 30	1073
	E300	28	20	20, 30, 45, 60	1137
3.85	J500	32	24	25, 35	1444
	J501	32	24	25, 35	3559

TABLE 4.1: The dimensions of the lattices of the CLS ensembles that were used for the scale setting analysis. The number of spatial smearing steps and measurements are listed here. On E300 a possible undersmearing was analysed (see Chapter A).

Equation (4.2),

$$W^{(l,k)}(x_0, r, T) = \sum_{n=0}^{\infty} c_n^{(l,k)}(r)\, \mathrm{e}^{-V_n(r)\,T}, \tag{4.5}$$

where l and k label the smearing levels of the spatial transporters, $V_n(r)$ denote the energy levels of the static quark–antiquark system, and the overlap coefficients satisfy $c_n^{(l,k)} = \left(c_n^{(k,l)}\right)^*$. At short to intermediate separations, the lowest-lying energy level $V_0(r) \equiv V(r)$ corresponds to the static potential. The nature of higher energy levels depends on the distance r. The first excited level $V_1(r)$ may correspond to an excitation of the static potential, to a static potential accompanied by hadronic states with vacuum quantum numbers (such as two-pion states), or to a pair of static–light mesons. As a result, the nature of the excited spectrum can change as a function of the quark–antiquark separation.

In the limit of infinite statistics, the correlation matrix $W^{(l,k)}(x_0, r, T)$ is Hermitian. In practical Monte Carlo simulations, finite-statistics fluctuations lead to small violations of this Hermiticity, and the matrix is therefore explicitly symmetrised before being analysed. The energy levels are then extracted using the generalised eigenvalue problem (GEVP) applied to the correlation matrix [8, 89–92]. This variational procedure allows the construction of optimised operators with enhanced overlap onto individual energy eigenstates, in particular the ground state corresponding to the static potential.

In practice, Wilson loops are averaged over the temporal source position x_0, excluding

regions close to the temporal boundaries as will be discussed in Section 4.5. The resulting symmetrized Wilson loops at different smearing levels form the correlation matrix entering the generalised eigenvalue problem,

$$C(r,T)\vec{v}_n = \lambda_n(T,T_0)C(r,T_0)\vec{v}_n, \qquad \text{with } T > T_0. \tag{4.6}$$

The energies can be found by

$$E_n(T,T_0) = \ln(\lambda_n(T,T_0)/\lambda_n(T+a,T_0)), \tag{4.7}$$

where $n = 0$ corresponds to the ground state and larger n to the excited states.

The use of a variational basis is crucial in view of the severe signal-to-noise problem associated with Wilson loops. Since standard multilevel variance-reduction techniques are not directly applicable in full QCD (see Section 4.2), operator optimisation and HYP smearing play a central role in isolating the ground-state potential. Combined with HYP smearing, the GEVP analysis provides a robust framework for determining the static force and potential scales on the CLS ensembles.

4.5 Treatment of open temporal boundaries

A subset of the CLS ensembles employed in this work uses open boundary conditions in the temporal direction. These boundary conditions are introduced to prevent topological freezing at fine lattice spacings, but they explicitly break time-translation invariance near the temporal boundaries. As a result, correlation functions measured close to the boundaries can receive additional contributions that are absent in the bulk of the lattice and must be excluded from the analysis. For Wilson loops, which are characterised by a temporal extent T and a starting Euclidean time x_0, this implies a potential dependence on x_0 in the presence of open boundaries. To illustrate this effect, Fig. 4.1 shows representative Wilson loops $W(x_0, r, T)$ as a function of x_0 for two ensembles, `J501` and `H101`, corresponding to coarse and fine lattice spacings, respectively. The spatial separation r is chosen to lie in the region relevant for the determination of the potential scale, and the temporal extent T is taken from the plateau region of the effective potential.

As shown in Fig. 4.1, Wilson loops exhibit pronounced distortions near the temporal boundaries, while a broad plateau is observed in the central bulk region of the lattice. The vertical dashed lines indicate the cut parameter N_{cut}, which defines the minimal distance from the temporal boundaries required for a Wilson loop to be included in the analysis. Only Wilson loops satisfying

$$x_0 \geq N_{\text{cut}} \quad \text{and} \quad x_0 + T \leq T_{\text{lat}} - N_{\text{cut}} \tag{4.8}$$

are retained. Within this window, the residual dependence on x_0 is consistent with statistical fluctuations. The choice of N_{cut} represents a compromise between suppressing boundary effects and maintaining sufficient statistical precision. To assess the sensitivity of the final results to this choice, the potential scale r_0/a was determined for a range of N_{cut} values. The resulting dependence is shown in Fig. 4.2 for the same two ensembles. While individual Wilson loops show sizable boundary distortions, the extracted values of r_0/a are remarkably stable once N_{cut} exceeds a moderate threshold. This stability indicates that residual boundary effects largely cancel in the

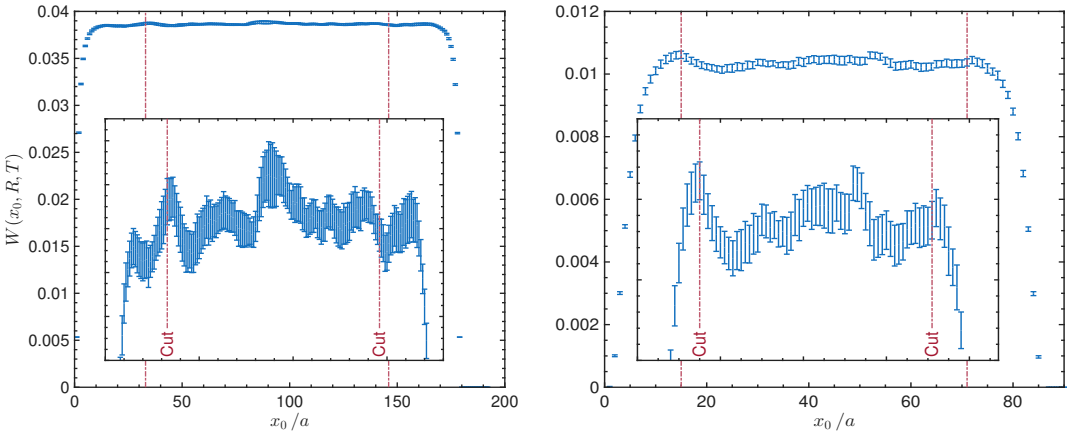

FIGURE 4.1: Wilson loop $W(x_0, R, T)$ with R around the r_0 range and T in the plateau area for Ensembles `J501`(left) and `H101`(right), being in the two end spectrums of our lattice spacings, as a function of x_0. The cut of the timeslices for the left plot being $N_{\text{cut}} = 33$ and $N_{\text{cut}} = 15$ for the right zoomed to the area of interest.

variational analysis and subsequent extraction of the static potential.

Based on this study, a conservative value of N_{cut} is chosen for each lattice spacing scaled with $\frac{1}{a}$, corresponding to a fixed physical distance from the temporal boundaries. The selected values are indicated by the red data points in Fig. 4.2. With this choice, boundary effects are safely suppressed while retaining a sufficiently large number of measurements. The remaining dependence on N_{cut} is well below the statistical uncertainty and is therefore neglected in the final error budget. This procedure ensures a consistent and controlled treatment of open temporal boundary effects across all ensembles used in this work.

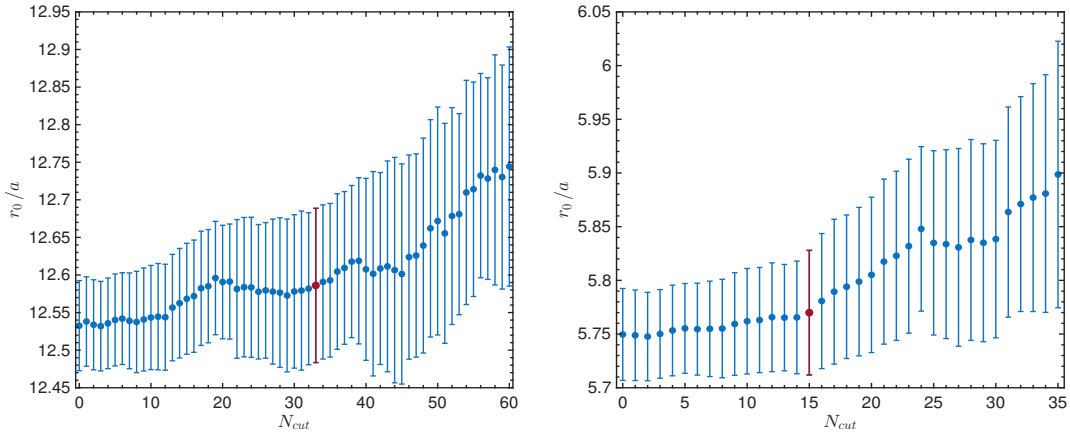

FIGURE 4.2: The dependence of r_0 on N_{cut} for Ensembles `J501`(left) and `H101`(right), with the red datapoints showing the final N_{cut} used of the lattice spacings.

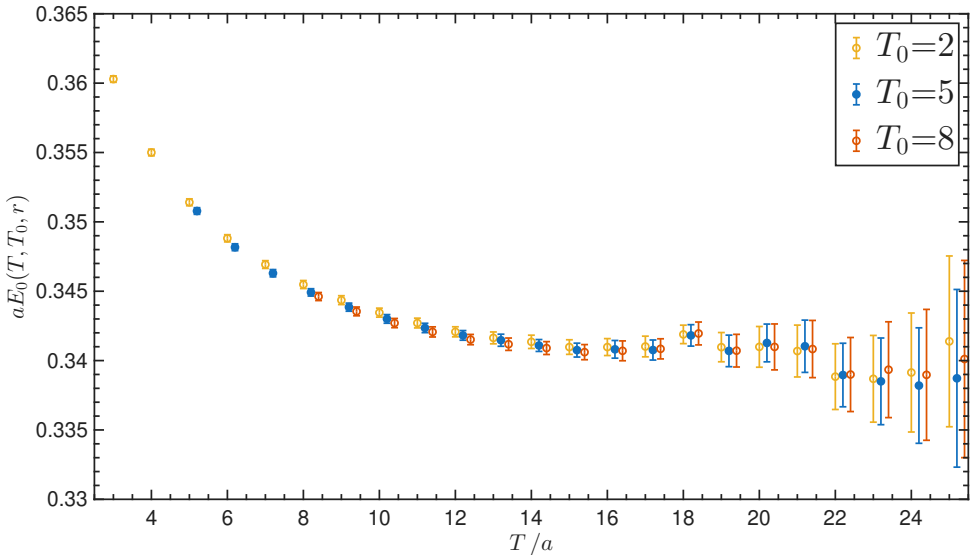

FIGURE 4.3: Effective mass for ensemble `J501` with $r = 13a$ showing the difference of T_0

4.6 Extraction of effective masses

Having defined the variational procedure and the treatment of temporal boundary effects, we now turn to the extraction of effective masses from the generalised eigenvalues obtained from the GEVP. A central challenge in this step is the identification of a reliable plateau region in Euclidean time, where contributions from excited states are sufficiently suppressed while statistical uncertainties remain under control.

Figure 4.3 shows an effective mass, with different T_0 used in the GEVP of Equation (4.6). Larger values of T_0 lead to a faster suppression of excited-state contributions, at the cost of reducing the available temporal fitting range. We chose to keep the value $T_0 = 5a$ constant for all ensembles and distances r. One could argue that this value should be set for one ensemble and scaled inversely proportional to the lattice spacing. As it is a discrete value, some ensembles would anyway share the same value and the effect of increasing by a single step is negligible. As will be discussed in Section 4.6.2 the earlier points, still affected by the excited states, will also be used for an exponential fit, so that not too many points at early times should be cut to keep the fit stable.

4.6.1 Initial plateau-based determination via weighted averages

As an initial approach, effective masses were analysed using weighted average over a manually selected plateau region. As a choice for the plateau range, choosing the "right" end point is typically not crucial, as the signal-to-noise problem arises not too far into the plateau. The high uncertainty in these datapoints make no difference to the end result choosing one over the other. On the other hand, the choice of the starting point T_{min} is crucial, as the earliest points have the smallest statistical uncertainties but are most affected by excited-state contamination. Including or excluding these points can therefore lead to a noticeable shift in the extracted value and its associated uncertainty. For each distance r, the effective mass was inspected visually and a lower bound T_{min} was chosen such that the data appeared compatible with a constant within uncertainties. While this is widely used and allows for a detailed case-by-case

assessment, it relies on subjective judgement in the choice of T_{min} and does not easily scale to a large number of observables.

To reduce subjectivity in the choice of T_{min}, its dependence of the extracted effective mass was studied by constructing T_{min} and T_{max} plots. Figure 4.4 shows such a figure for `J501` at $r = 13a$, where the left shows a T_{min} plot and the right a T_{max}. As a reference one can inspect Figure 4.3 how the effective mass looks like. It should come to no surprise to see that given a fixed T_{min} the right bound makes close to no difference compared to choosing the left bound. These plots provide a useful diagnostic tool and allow the identification of regions where the result is stable under variations of T_{min}. However, it was found that no simple universal criterion, for example requiring stability within a fixed fraction of the statistical uncertainty, proved sufficiently reliable across all ensembles and distances.

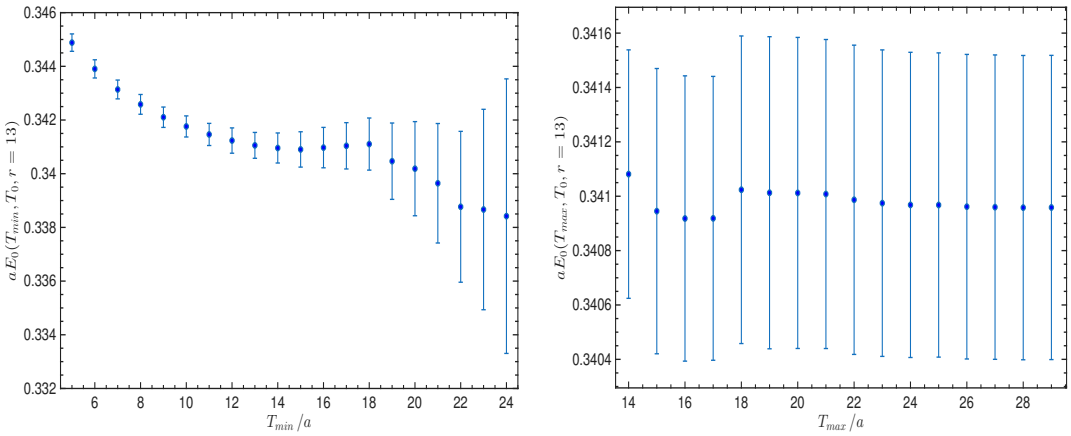

FIGURE 4.4: T_{min}-plot with maximum upper limit on the left and T_{max}-plot with $Tmin = 14$ on the right.

While this approach yields sensible results on a case-by-case basis, it requires continued manual inspection and does not scale efficiently to the analysis of a large number of effective masses. One might observe apparent stability in the T_{min} plot, only to find out that the point clearly shows excited state contamination. For the present study, involving multiple distances on approximately twenty ensembles, a more systematic procedure was required.

One possibility that was discussed[1] was, that since the starting point of the plateau was comparatively similar between ensembles across the same lattice spacing, T_{min} could be determined for effective masses with a reasonable clear plateau range and applied to the other ensembles. Even though this would work in most cases, since it would deliver unsatisfactory results in other cases, this approach was discarded, as one possibly could neglect some unknown lattice artifacts.

4.6.2 Fit-based extraction of effective masses

Motivated by the limitations of explicit plateau identification, a fit-based approach was adopted as the default method for extracting effective masses. This procedure follows the strategy we have deployed in Sec. 4.3 of Ref. [93] and allows for a more uniform and reproducible treatment across ensembles and distances.

[1] Scale setting: Precision lattice QCD for particle and nuclear physics workshop, 03.-07. March 2025, Trento

Instead of selecting a plateau region explicitly, effective masses are analysed by performing fits over temporal ranges in the form of

$$E_0(T, T_0, r) = E_0 + Be^{-\Delta \cdot T} \,. \tag{4.9}$$

In this framework, residual excited-state contributions are accounted for implicitly through the choice of fit range, and the extracted ground-state energy is less sensitive to individual time slices. In cases where a clear plateau could be identified, the fit-based approach yields results that are fully compatible with those obtained from weighted plateau averages. Its main advantage lies in the ability to incorporate data at earlier Euclidean times, in practice including all time separations starting from the reference time T_0. This allows for a precise determination of the ground-state energy even in situations where the onset of a plateau occurs only at very large temporal separations or is only weakly resolved within the available statistical precision. Theoretically, this approach is particularly well suited for large-scale analyses, as it reduces the need for manual intervention while retaining full control over systematic effects. Stability checks with respect to the fit range are used to validate the extracted energies, and the resulting uncertainties are found to be mainly statistical.

While the fit ansatz in Equation (4.9) provides a flexible and physically motivated description of the effective masses, it was found that a direct three-parameter fit could not be stabilised reliably for all distances and ensembles. In particular, for some effective masses the available statistical precision and temporal extent were insufficient to constrain the excited-state contribution parameter Δ simultaneously with the ground-state energy E_0. As a result, unconstrained fits occasionally led to unstable behaviour or failed to describe the data consistently across all ensembles. To address this issue, a two-step fitting strategy was adopted. In a first step, those ensembles and distances for which the three-parameter fit was well behaved were used to determine the excited-state gap Δ as a function of the spatial separation r, the light-quark mass, and the lattice spacing. For each lattice spacing, a global fit was performed to establish a smooth model for the Δ values. This approach effectively reduces the dimensionality of the subsequent fits while retaining sensitivity to the relevant physical dependencies. The global fit ansatz is motivated by the expectation that the energy of the first effective excited state, $E_0 + \Delta$, varies only mildly with the quark–antiquark separation in the region relevant for scale setting. Under this assumption, the gap Δ is expected to follow the same qualitative distance dependence as the static potential itself. Accordingly, the fit model for Δ is chosen in the form

$$\sqrt{t_0}\,\Delta = \hat{\Delta}_0 - \hat{\delta}_0 \frac{r}{\sqrt{t_0}} - \hat{\gamma}_0 \frac{\sqrt{t_0}}{r} + \hat{b}_0\, t_0 \Big(m_\pi^2 - (m_\pi^{\mathrm{phys}})^2 \Big) \,, \tag{4.10}$$

where the parameters $\hat{\Delta}_0$, $\hat{\delta}_0$, $\hat{\gamma}_0$, and $\hat{b}_0$ are dimensionless parameters determined separately for each lattice spacing. The values for $\sqrt{t_0}$ for each ensemble have been calculated previously by Ref. [60]. This functional form captures the dominant distance dependence while allowing for a mild quark mass dependence through the pion mass. It should be emphasised that the gap Δ should be interpreted as an effective parameter that absorbs contributions from the lowest excited states relevant at a given separation r. Depending on the distance, these contributions may arise from different types of excited states, including string excitations or, in the presence of dynamical quarks, multi-hadron states with the appropriate quantum numbers. The adopted parametrisation is therefore not intended to resolve the detailed structure of the excited spectrum, but to provide a stable and economical description of its net

effect on the ground-state extraction.

Once the parameters of this model are fixed, the corresponding value of $\Delta(r)$ can be read off for all ensembles and distances, including those for which the original three-parameter fit was unstable. In a second step, the effective masses are then refitted using Equation (4.9) with Δ fixed to the model value, resulting in a well-conditioned two-parameter fit for E_0 and the amplitude B. This procedure significantly improves the numerical stability of the fits while preserving consistency with the unconstrained results in cases where both approaches can be applied.

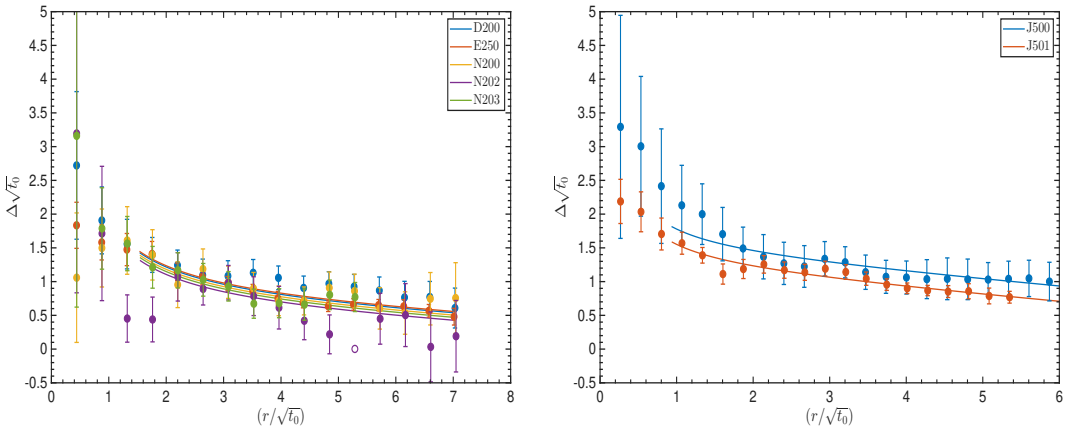

FIGURE 4.5: Excited-state gap Δ as a function of $r/\sqrt{t_0}$ for selected CLS lattice spacings. Solid lines show the global fit of Equation (4.10) performed separately for each lattice spacing. Short-distance points affected by discretisation effects are excluded from the fit but shown for reference.

The quality of the global description of the Δ values is illustrated in Figure 4.5, where the fitted curve is compared to the individual data points. The first 3 points at very short distances, which are affected by large discretisation effects, are excluded from the fit but are shown for illustration. In the range relevant for the determination of the potential scale, $r_1 \leq r \leq r_0$, the model provides a good description of the data. The quality of the global fits for the excited-state gap Δ is summarised in Table 4.2. At the coarser lattice spacings, a larger fraction of data points had to be excluded to ensure fit stability, leading to fewer degrees of freedom despite the larger number of ensembles. The resulting systematic uncertainty associated with fixing Δ is found to be subdominant compared to the statistical errors and is therefore neglected in the final error budget.

β	3.4	3.46	3.55	3.7	3.85
$\frac{\chi^2}{\text{d.o.f.}}$	$\frac{10.2}{20}$	$\frac{19.7}{23}$	$\frac{30.7}{56}$	$\frac{56.1}{50}$	$\frac{8.2}{34}$

TABLE 4.2: Values of χ^2 and corresponding numbers of degrees of freedom obtained from the global fits of the excited-state gap Δ at each lattice spacing.

While the method mentioned above is the tool used for the final determination of the potential of the ensembles, other methods have been analysed. For one it has been considered performing the Fit model Equation (4.10) on the whole data-set instead of for each lattice spacing. In Figure 4.6 the result of it can be seen, where the fits correspond to the whole data, but only the data points of the symmetric points are

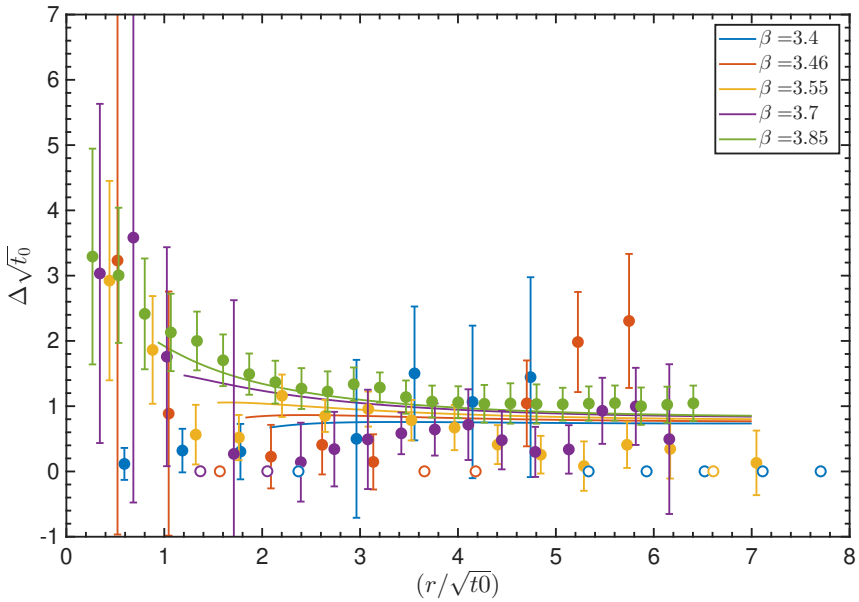

FIGURE 4.6: Representation of the Δ model if performing Equation (4.10) on the whole data-set instead of on each lattice spacing. Only ensembles at the symmetric masses are shown as not to clutter the figure. Empty dots represent data with unstable results not taken into account in the analysis.

shown, instead of all 20 ensembles. The statistical distribution is $\frac{\chi^2}{\text{d.o.f.}} = \frac{215}{196}$. This method could have been used just as well, but we chose to have the determination be contained for each lattice spacing separately. Another approach was to extend the singular exponential fit Equation (4.9) with a second exponential either as having 2 new free parameters or by constraining the second exponential to depend on the pion mass. In this method we saw that there were cases where the two exponentials were highly correlated and no new information was gathered and that this method proved to be significantly more unstable compared to the singular exponential fit, so that it required extensive parameter tuning or the exclusion of a large fraction of data points in the Δ model analysis.

Representative examples of the resulting effective masses together with the corresponding fit functions are shown in Figure 4.7. The figure illustrates the stability of the extracted ground-state energies across different lattice spacings, quark masses, and quark–antiquark separations, as well as the ability of the fit-based approach to reliably describe the data even in cases where a clear plateau is not apparent at early Euclidean times.

Before going to the static force the GEVP and corresponding equation to find the energy values in Equations (4.6) and (4.7) can be used to find an estimate of excited states by evaluating the equation for $n > 0$. Given a small basis of operators, since the generalised eigenvalue problem is solved for a 2×2 correlator matrix constructed from two Wilson-loop operators with different smearing levels, the amount of states is limited. The second eigenvalue provides an estimate of the lowest excited state that has non-negligible overlap with the chosen operator basis. This state is not guaranteed to coincide with the first physical excitation of the static quark–antiquark system, as additional states within the same quantum number channel with smaller overlap may not be resolved by the chosen operator basis. In the static quark–antiquark system this state is commonly interpreted as the lowest flux-tube-like excitation accessible

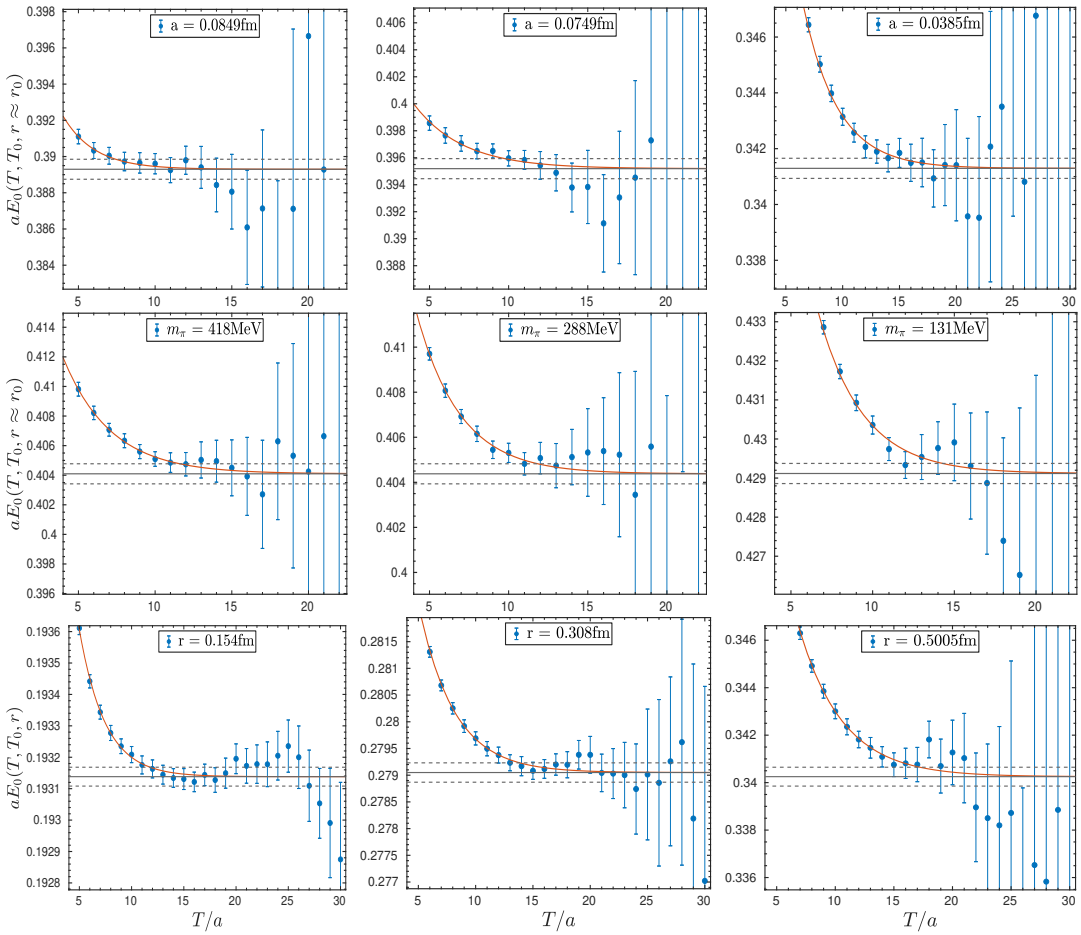

FIGURE 4.7: Representative effective mass plots together with the fit results obtained using the fit-based extraction procedure described in Section 4.6.2. The first row shows results at approximately $r \simeq r_0$ for the symmetric-point ensembles `H101`, `B450`, and `J500` with decreasing lattice spacing. The second row illustrates the quark-mass dependence at fixed lattice spacing, showing ensembles `N202`, `N200`, and `E250` with pion masses decreasing from the symmetric point towards the physical value. The last row displays results for the finest lattice `J501` at $m_\pi \simeq 340\,\mathrm{MeV}$ for several quark–antiquark separations ranging from $r_1/2$ up to r_0. Solid curves represent the fit ansatz of Eq. (4.10), while the horizontal bands indicate the extracted ground-state energies and their statistical uncertainties.

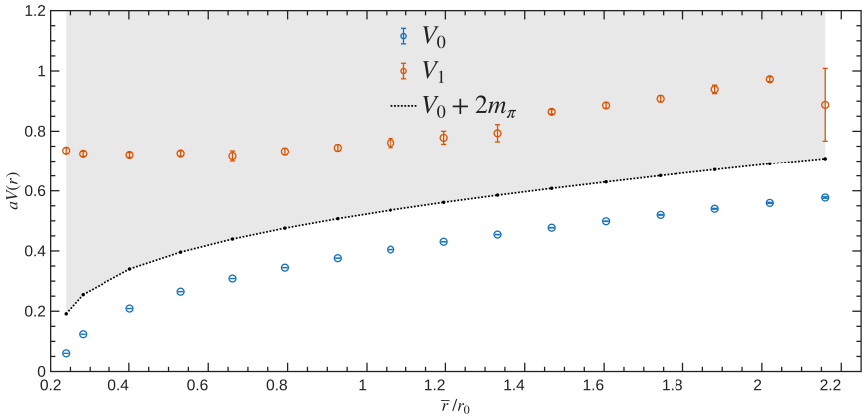

FIGURE 4.8: Static quark–antiquark potential $aV(r)$ as a function of the separation $\bar{r}/r_0$. The ground-state potential (blue) and the next state (orange) are extracted from the two lowest eigenvalues of a 2×2 GEVP constructed from Wilson-loop operators with different smearing levels. The second eigenvalue provides an estimate of the lowest excited state accessible within the chosen operator basis. The dashed line indicates the approximate non-interacting two-meson threshold $V_0 + 2m_\pi$, which is relevant for the onset of string breaking in full QCD.

within the chosen Wilson-loop operator basis [94]. Extracting the eigenvalues for $n = 1$ and estimating the plateaus, which were shorter and in general harder to establish than for the ground state, the values can be seen in Figure 4.8 for ensemble `D200`. The plotted region remains below the expected string-breaking region, which according to string breaking distance of the static light lies at $r_l^\star = 1.21\text{fm}$ [95], which corresponds to $\frac{r_l^\star}{r_0} = 2.56$. Therefore, no clear level crossing associated with string breaking is expected in this range. The shaded area corresponds to energies above $V_0 + 2m_\pi$, where m_π from `D200` is taken from [54]. This is included only as a qualitative reference scale for the onset of light multi-hadron excitations and should not be interpreted as the asymptotic string-breaking energy itself.

4.7 Interpolation of the static force

The potential scales r_0 and r_1 are defined through the static force, introduced in Equation (2.48), as

$$r^2 F(r)\Big|_{r=r_0} = 1.65, \qquad r^2 F(r)\Big|_{r=r_1} = 1.0, \tag{4.11}$$

where $F(r)$ is obtained from the discrete static potential using improved distances, defined in Section 2.5.6. Since the force is only known at a finite set of separations, a local interpolation is required to determine the values of r_0/a and r_1/a, where both two-point and three-point interpolation strategies are considered. The two-point interpolation uses the two neighbouring force values that enclose the target point, while the three-point interpolations include an additional point either at smaller or larger distances. These different choices allow one to assess the sensitivity of the result to the interpolation procedure.

The simplest approach uses the two force values that enclose the target point r_0 and corresponds to a linear interpolation of $r^2F(r)$,

$$F(r_I)\, r_I^2 = \hat{f}_1\, r_I^2 + f_0, \qquad \text{(2-point interpolation)}. \tag{4.12}$$

The two variants of a three-point interpolation are defined as,

$$F(r_I)\, r_I^2 = \hat{f}_1\, r_I^2 + f_0 + \hat{f}_2\, r_I^{-2}, \qquad \text{(3-point interpolation)}. \tag{4.13}$$

Here, f_0, $\hat{f}_1$, and $\hat{f}_2$ denote the interpolation coefficients. The inclusion of the additional term allows for a mild curvature in the force and provides a useful consistency check of the interpolation procedure.

An example of the interpolation procedure on a fine lattice is shown in Figure 4.9, where the ensemble `J500` is displayed. In this case, the different interpolation strategies lead to nearly identical results within uncertainties, and the determination of both r_0 and r_1 is stable. This behaviour reflects the fact that, on fine lattices, the relevant distance region is sufficiently far from both short-distance lattice artefacts and large-distance statistical fluctuations, indicating that systematic effects associated with the choice of interpolation ansatz are subdominant. The purple interpolation corresponds to the three-point ansatz of Equation (4.13), using the two points that enclose the target value together with an additional point at larger distance.

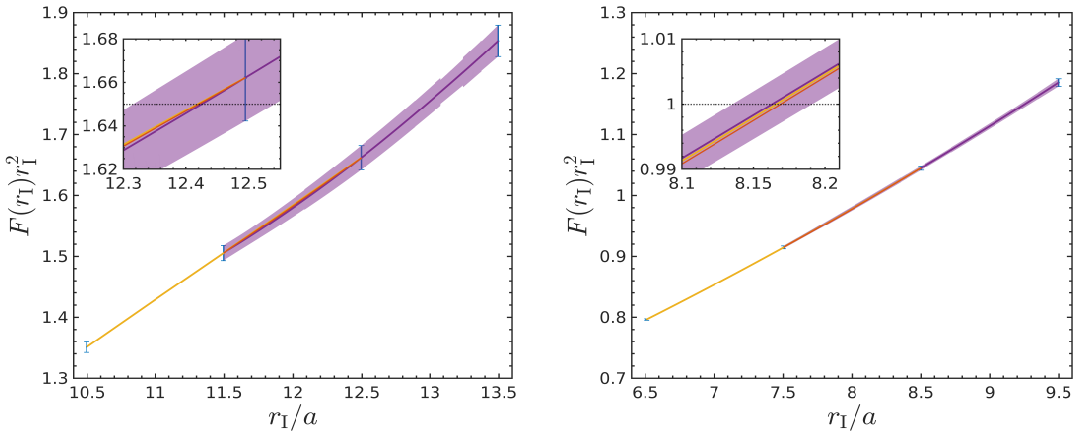

FIGURE 4.9: Local interpolations of the static force $r_I^2F(r_I)$ around the potential scales. The red line shows the two-point interpolation, while the yellow and purple lines correspond to three-point interpolations using an additional point at smaller or larger distance, respectively. The purple interpolation, including its uncertainty band, is used to determine r_0/a. Insets show a zoom of the region around the extraction point.

To further illustrate the behaviour of the force across ensembles, the quantity $r_I^2F(r_I)$ is shown in Figure 4.10 for both the finest ensemble `J500` and the coarsest ensemble `C101`. While on the fine lattice the scales r_0 and r_1 are located in a smooth region of the force, the situation changes on the coarse lattice. In particular, r_1 lies closer to the short-distance region where discretisation effects become significant, whereas r_0 remains in a region where the force exhibits a smoother behaviour. At the largest distances for the coarsest lattice, a deviation from the smooth behaviour of the force is visible. This can be attributed to the increasing statistical uncertainties at large separations and may also receive contributions from finite-volume effects, although no dedicated study is performed here to disentangle these contributions.

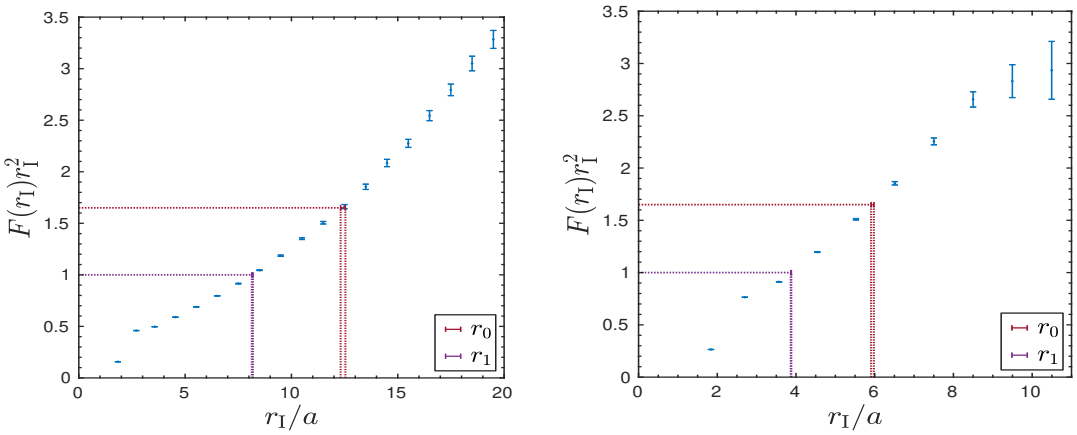

FIGURE 4.10: The static force $r_I^2 F(r_I)$ as a function of the improved distance r_I/a for two representative ensembles. The left panel shows the finest lattice J500, while the right panel corresponds to the coarsest lattice C101. The potential scales r_0 and r_1 are indicated by the intersections with the horizontal reference lines. While both scales are well separated from the short-distance region on the fine lattice, one observes that r_1 approaches the region affected by lattice artefacts on the coarse lattice, indicating an increased sensitivity to discretisation effects at small distances.

This difference between the finest and coarsest lattice is reflected in the interpolation procedure, as shown explicitly in Figure 4.11 for the coarse ensemble C101. While the determination of r_0 remains stable across the different interpolation strategies, a noticeable spread is observed in the determination of r_1. In particular, interpolation methods that include points at the smallest distances are affected by lattice artefacts. Restricting the interpolation to points at larger distances results in a more stable determination. These observations indicate that the extraction of r_1 is more sensitive to discretisation effects on coarse lattices, whereas r_0 remains comparatively robust. This behaviour is consistent with the definition of the two scales, since r_1 probes shorter distances where lattice artefacts are more pronounced. As a consequence, particular care must be taken when including coarse lattice ensembles in analyses based on r_1, and in the subsequent continuum extrapolation the coarsest lattice will be excluded, which will be brought up again in that chapter.

The final values of r_0/a and r_1/a are obtained using the three-point interpolation that avoids the smallest-distance points and thus minimises the impact of discretisation effects. The differences between the various interpolation strategies are found to be small compared to the statistical uncertainties for r_0, and moderate but controlled for r_1, and are therefore not included as a separate systematic uncertainty. The moderate differences between the interpolation strategies decrease significantly towards finer lattice spacings, consistent with the expected reduction of discretisation effects.

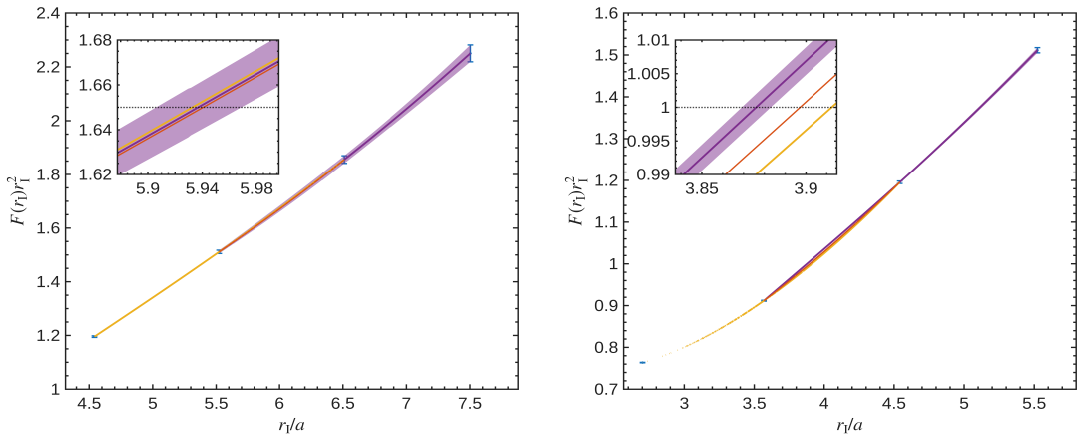

FIGURE 4.11: Comparison of different local interpolation strategies for $r_I^2 F(r_I)$ around the determination of r_0 (left) and r_1 (right) for the coarse ensemble C101. The red line corresponds to the two-point interpolation, while the yellow and purple lines represent three-point interpolations using an additional point at smaller or larger distance, respectively. A noticeable spread between the interpolation methods is observed in the determination of r_1, reflecting the increased sensitivity to short-distance lattice artefacts. In particular, interpolations that include the smallest-distance point are affected by discretisation effects, whereas using points at larger distances leads to a more stable determination.

4.8 Summary of scale determinations

Applying the procedures described in this chapter, the potential scale r_0/a and the auxiliary scale r_1/a have been determined for all CLS ensembles considered in this work. The resulting values, together with the ratio r_0/r_1, are summarised in Table 4.3. These quantities are obtained independently for each ensemble and constitute the final lattice-scale outputs of the static-potential analysis. At this stage, the results are not yet extrapolated to the continuum limit or matched to physical units; rather, they provide a consistent and well-controlled set of inputs for the subsequent scale-setting analysis across lattice spacings and quark masses.

β	id	r_0/a	r_1/a	r_0/r_1
3.40	H101	5.81(7)	3.842(13)	1.511(19)
	H102	5.87(8)	3.886(14)	1.511(20)
	H105	5.89(8)	3.860(11)	1.526(21)
	C101	5.94(4)	3.874(8)	1.534(10)
3.46	B450	6.57(8)	4.299(19)	1.527(20)
	S400	6.71(7)	4.331(17)	1.549(15)
	D450	6.62(3)	4.327(9)	1.530(9)
	D452	6.69(3)	4.340(8)	1.542(8)
3.55	N202	7.72(6)	5.066(19)	1.525(14)
	N203	7.77(5)	5.036(14)	1.543(10)
	N200	7.78(6)	5.025(13)	1.549(12)
	D200	7.69(3)	5.041(9)	1.526(6)
	E250	7.79(2)	5.054(5)	1.541(5)
3.70	N300	9.77(11)	6.413(31)	1.525(20)
	J303	9.84(5)	6.432(13)	1.530(8)
	E300	9.84(11)	6.529(24)	1.506(18)
3.85	J500	12.43(9)	8.163(28)	1.523(12)
	J501	12.61(15)	8.226(32)	1.533(19)

TABLE 4.3: Values of the potential scales r_0/a and r_1/a, and their ratio r_0/r_1, determined on all CLS ensembles used in this work. The quoted uncertainties are statistical only and originate from the extraction of the static potential and the interpolation of the static force. These results form the basis for the scale-setting and continuum extrapolation discussed in the following chapter. For ensemble `E300`, the pencil-of-function method described in Chapter A is used.

Chapter 5

Scale setting and continuum extrapolation

The determination of a reference scale in physical units is a central prerequisite for connecting lattice QCD calculations to phenomenology. In the previous chapter, the static quark–antiquark potential was determined on a set of CLS $N_f = 2 + 1$ gauge ensembles, and the potential scales r_i/a were extracted for each ensemble with controlled statistical uncertainties and subdominant systematic effects. The primary objective of this chapter is to determine the potential scales r_0 and r_1 in physical units, starting from the ensemble-level determinations of r_i/a and combining them with an external input of the Wilson-flow scale $\sqrt{t_0}$. The analysis proceeds through continuum and chiral extrapolations of suitable dimensionless combinations, an assessment of cutoff effects and residual mass dependence, and finally a matching to the physical point. A substantial part of the analysis presented in this chapter is based on the methodology we have developed in [93], of which this thesis work forms an integral part. Here, the discussion is adapted to the structure of the thesis and expanded where necessary to clarify methodological choices and their theoretical motivation.

5.1 Strategy of the scale-setting analysis

The starting point for the scale-setting analysis is the set of ensemble-specific determinations of the potential scales r_i/a obtained from the static force on the CLS ensembles in Chapter 4. These bare scales are defined in terms of the tree-level improved static force at the improved distances introduced earlier and have been extracted with a careful control of excited-state contamination, open-boundary effects and interpolation systematics. Taken by themselves, the ratios r_i/a determine the lattice spacing in lattice units, but not the overall physical scale. To fix the absolute scale, in line with the CLS scale-setting programme based on the gradient-flow scale [62], we express the potential scales in units of $\sqrt{t_0}$ and analyse the dimensionless ratios

$$\frac{r_0/a}{\sqrt{t_0/a^2}} = \frac{r_0}{\sqrt{t_0}}, \qquad \frac{r_1/a}{\sqrt{t_0/a^2}} = \frac{r_1}{\sqrt{t_0}}, \qquad \frac{r_0}{r_1}, \tag{5.1}$$

evaluated on each ensemble using the corresponding determinations of r_i/a and t_0/a^2. The latter are not computed in the present work but are taken from the CLS scale-setting programme, in which the gradient-flow scale t_0 is measured on the fly on the gauge ensembles and analysed as described in [60]. In these combinations, the explicit dependence on the lattice spacing cancels, and the remaining dependence on the bare parameters can be parametrised in terms of dimensionless quark-mass variables,

such as the combinations ϕ_2 and ϕ_4 defined in Equation (3.6), together with cut-off effects that vanish in the continuum limit. Since the flow scale is known with high relative precision, the dominant uncertainties in these ratios are inherited from the potential-scale determination rather than from t_0. The ratios are constructed at the level of ensemble averages, and their statistical uncertainties are obtained using the Γ-method introduced in Section 3.6.2. This framework treats all primary and derived observables from a given ensemble as a correlated set, such that autocorrelations in Monte Carlo time and cross-correlations between r_i/a and t_0/a^2 are propagated consistently to the final errors of $r_i/\sqrt{t_0}$ and r_0/r_1. To disentangle lattice-spacing artefacts from physical mass dependence, it is useful to identify, for each β, a reference point at (or close to) the SU(3)-symmetric line, where the light and strange quark masses are degenerate and ϕ_4 is near its target value. At this point, which will be denoted by the subscript "sym", one can define symmetric-point reference scales r_0^{sym}, r_1^{sym}, and t_0^{sym} that serve as natural normalisations for the leading cut-off terms in the global fits, in the sense that they allow the discretisation effects to be expressed in terms of dimensionless ratios with a common reference scale at each value of β.

The analysis proceeds in three steps:

1. **Ensemble-level observables:** For each ensemble, the quantities r_0/a, r_1/a, t_0/a^2 and their ratios are collected, with statistical correlations propagated using the error-analysis framework already employed for the static potential. This includes the correlations induced by common Monte Carlo histories and reweighting factors.

2. **Continuum and chiral fits:** The ensemble-level ratios $r_0/\sqrt{t_0}$, $r_1/\sqrt{t_0}$ and r_0/r_1 are fitted as functions of the pion mass (or of ϕ_2 and ϕ_4) and the lattice spacing, using ansätze motivated by Symanzik effective theory and chiral expansion. The leading cutoff effects are assumed to be of order a^2, consistent with the $O(a)$-improved fermion action and improved gauge action, while the quark-mass dependence is parametrised by simple low-order polynomials. Several fit variants are considered to estimate systematic effects.

3. **Matching to the physical point:** Once continuum-limit values at the physical mass point have been obtained for the above ratios, the absolute scale is set using an external determination of t_0^{phys} in femtometres. In this way, the physical values of r_0, r_1 and their ratio are obtained without introducing additional dependence on phenomenological potential models.

A key aspect throughout is the separation of different sources of uncertainty. Statistical errors and the effects of autocorrelations are treated consistently using the same methodology as in the preceding chapters. Systematic uncertainties associated with the continuum and chiral extrapolations are estimated by varying the fit ansätze, applying cuts in lattice spacing and pion mass, and, where appropriate, averaging over fit variants using an information-theoretic criterion. The uncertainty associated with the external input for $\sqrt{t_0^{\mathrm{phys}}}$ is propagated separately, so that the final error budget on r_0 and r_1 in physical units can be decomposed into a "simulation part" and a "scale-input part".

5.2 Chiral behaviour at fixed lattice spacing

The starting point of the analysis is the set of values r_0/a and r_1/a obtained for each CLS ensemble in Chapter 4 and tabulated in Table 4.3. For each fixed bare coupling, corresponding to a fixed lattice spacing, these data can be plotted as a function of the pion mass (or, equivalently, of the dimensionless parameter $\phi_2 \propto t_0 m_\pi^2$ introduced in Equation (3.6)).

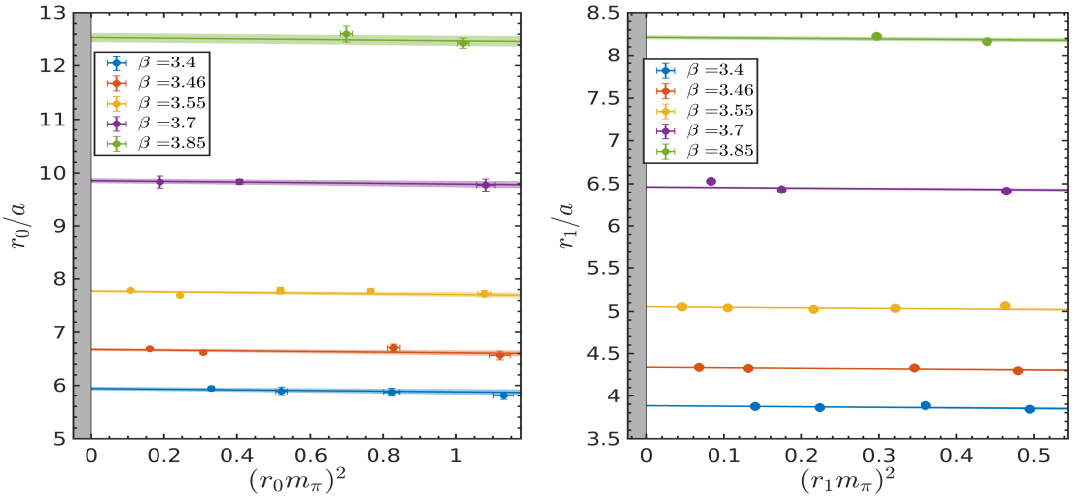

FIGURE 5.1: r_0/a and r_1/a for all ensembles as a function of $(r_0 m_\pi)^2$ left and $(r_1 m_\pi)^2$ right. A global fit eq. (5.2) is performed on the data.

Empirically, both r_0/a and r_1/a exhibit a very mild quark-mass dependence along the CLS chiral trajectories at fixed lattice spacing, as illustrated in Figure 5.1. For each β, the data are well described by a linear ansatz in $(r_0 m_\pi)^2$,

$$\frac{r_i}{a}(\beta, m_\pi^2) = A_i|_\beta + B_i\,(r_0 m_\pi)^2, \qquad i \in \{0, 1\}, \tag{5.2}$$

with slopes B_i that are small compared to the statistical precision of the intercepts $A_i(\beta)$. The chiral-limit values $r_i^\chi/a = A_i(\beta)$ and physical-pion-mass values $r_i(m_\pi^{\text{phys}})/a$ are collected in Table 5.1.

β	r_0^χ/a	$r_0(m_\pi^{\text{phys}})/a$	r_1^χ/a	$r_1(m_\pi^{\text{phys}})/a$
3.4	5.941(35)	5.934(35)	3.883(8)	3.880(8)
3.46	6.676(26)	6.670(26)	4.341(6)	4.338(6)
3.55	7.773(21)	7.766(21)	5.055(5)	5.052(5)
3.7	9.860(46)	9.854(46)	6.462(12)	6.460(12)
3.85	12.540(88)	12.533(88)	8.214(23)	8.211(23)

TABLE 5.1: Chiral-limit r_i^χ/a and physical-point $r_i(m_\pi^{\text{phys}})/a$ values of the potential scales at fixed lattice spacing.

As discussed in Section 3.2.2, explicit mistuning corrections based on quark-mass derivatives are not applied in the present analysis, since their effect on the potential scales is negligible compared to the statistical uncertainties. As a representative example, on ensemble D450 a correction of ϕ_4 by 0.51% leads to a relative shift in r_0 of only 0.46‰, which is negligible compared to the statistical error. Residual

mistuning effects are instead absorbed through the alternative fit ansätze discussed in the following section.

5.3 Global fits and continuum extrapolation in dimensionless ratios

The central observables for continuum-chiral analysis are the ratios of Equation (5.1), which eliminate the explicit a dependence.

The combined continuum and chiral extrapolation is performed by fitting all available ensembles simultaneously using simple ansätze inspired by Symanzik effective theory at leading order and the observed weak quark-mass dependence. The central fit form ,labelled "Fit 1" with the superscript (1) in the following, for $r_0/\sqrt{t_0}$ reads

$$\frac{r_0}{\sqrt{t_0}}(\beta, m_\pi^2) = c_1^{(1)} + c_2^{(1)} \left(\frac{a}{r_0^{\mathrm{sym}}}\right)^2 + c_3^{(1)} \, (r_0 \, m_\pi)^2, \tag{5.3}$$

with fit parameters c_1, c_2, c_3 determined from a correlated least-squares analysis over all ensembles. Here the term proportional to $(a/r_0^{\mathrm{sym}})^2$ represents the leading $O(a^2)$ cutoff effects expected for an $O(a)$-improved Wilson action with tree-level improved gauge sector, while the term proportional to m_π^2 parameterises the residual physical mass dependence. An analogous ansatz is used for $r_1/\sqrt{t_0}$, and a linear form in m_π^2 is also adopted for the ratio r_0/r_1, which is particularly insensitive to cutoff effects.

The qualitative features of the combined continuum and chiral analysis are illustrated in Figures 5.2 and 5.3 using the baseline fit ansatz (Fit 1). In Figure 5.2, the ratios $r_0/\sqrt{t_0}$ and $r_1/\sqrt{t_0}$ are shown as functions of the light-quark mass proxy $(r_i m_\pi)^2$ for all lattice spacings. At fixed inverse bare coupling β, the data are well described by the corresponding fit curves and exhibit only a weak dependence on the quark mass, consistent with the expected behaviour of long-distance gluonic observables, for which the dominant scale is set by Λ_{QCD}, so that variations in the light-quark masses affect them only weakly via sea-quark contributions. At fixed quark mass, the dominant variation of the data are governed by the lattice spacing. This behaviour is made explicit in Figure 5.3, where the same fit functions are evaluated at the symmetric mass point and displayed as functions of a^2. Within uncertainties, the data follow an approximately linear behaviour in a^2, as expected for leading $\mathcal{O}(a^2)$ cutoff effects of the $\mathcal{O}(a)$-improved Wilson action with a tree-level improved gauge sector. The behaviour of the ratio r_0/r_1 is illustrated in Fig. 5.4. As expected for a dimensionless ratio of two gluonic reference scales, the data exhibit only a very weak dependence on the quark mass and no visible sensitivity to the lattice spacing within the current statistical precision. The ensemble data are well described by a simple linear ansatz in the mass variable $(r_0 m_\pi)^2$, and the resulting continuum-limit value is consistent across all lattice spacings. This insensitivity to both quark-mass and cutoff effects motivates the use of a minimal fit form for r_0/r_1 in the combined continuum and chiral analysis. Fit 1, which defines the baseline global fit, is shown here to visualize the mass and lattice-spacing dependence of the data. In all cases, fits including the coarsest lattice spacing, shown by the blue data points in the figures, are used only for diagnostic purposes and are excluded from the final averages as discretisation effects are expected to be largest there and cannot be reliably disentangled from higher-order cutoff contributions within the present precision, as mentioned in Section 4.7, where the determination of r_1 by the force interpolation was rougher for coarse ensembles.

The final results are obtained from a systematic analysis based on multiple fit ansätze and data cuts, as discussed in a following section.

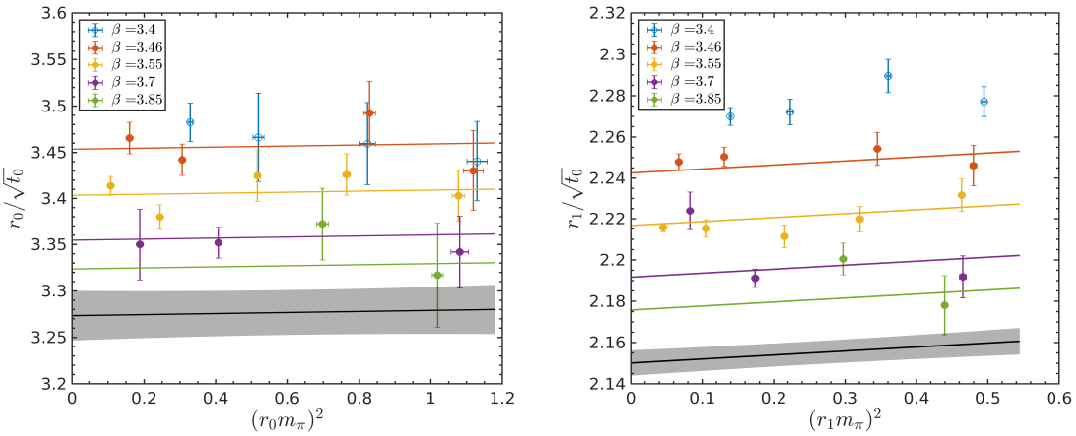

FIGURE 5.2: Ratios $r_0/\sqrt{t_0}$ (left) and $r_1/\sqrt{t_0}$ (right) shown as functions of $(r_i m_\pi)^2$ for all lattice spacings. Data at fixed inverse bare coupling β are shown with common colours, while the corresponding coloured lines represent the global fit functions evaluated at the respective lattice spacings. The grey band indicates the continuum limit obtained from Fit 1, cf.Equation (5.3), after setting the lattice-spacing dependent terms to zero. The horizontal axis $(r_i m_\pi)^2$ serves as a proxy for the light-quark mass and is equivalent, up to an overall rescaling, to a linear dependence on m_π^2 along the CLS chiral trajectories.

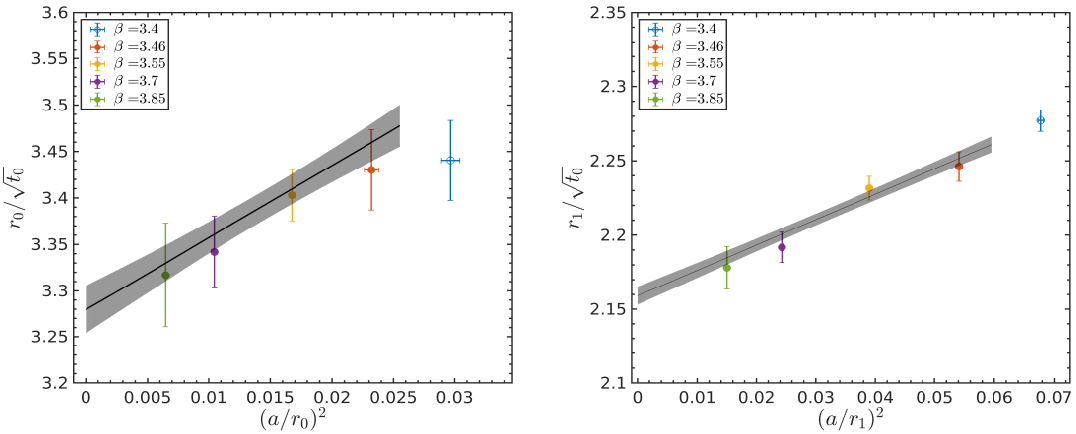

FIGURE 5.3: Same fits results as in Figure 5.2, evaluated at the symmetric mass point, shown as functions of a^2 to illustrate the continuum extrapolation. The data points correspond to the values of $r_0/\sqrt{t_0}$ (left) and $r_1/\sqrt{t_0}$ (right) at the symmetric point for each lattice spacing, while the bands represent the fit functions of Fit 1, cf. Equation (5.3). The extrapolation to the continuum-limit $a^2 \to 0$ at the physical mass is shown by the black line.

To assess systematic effects associated with the fit form, several variants are considered. A first extension ("Fit 2") supplements the leading a^2 term by an additional mass-dependent artefact of the form $m_\pi^2 a^2$, allowing for mild violations of a simple separation between quark-mass dependence and cutoff effects

$$\frac{r_0}{\sqrt{t_0}} = c_1^{(2)} + c_2^{(2)} \left(\frac{a}{r_0^{\text{sym}}} \right)^2 + c_3^{(2)} (r_0\, m_\pi)^2 + c_4^{(2)} m_\pi^2 a^2. \tag{5.4}$$

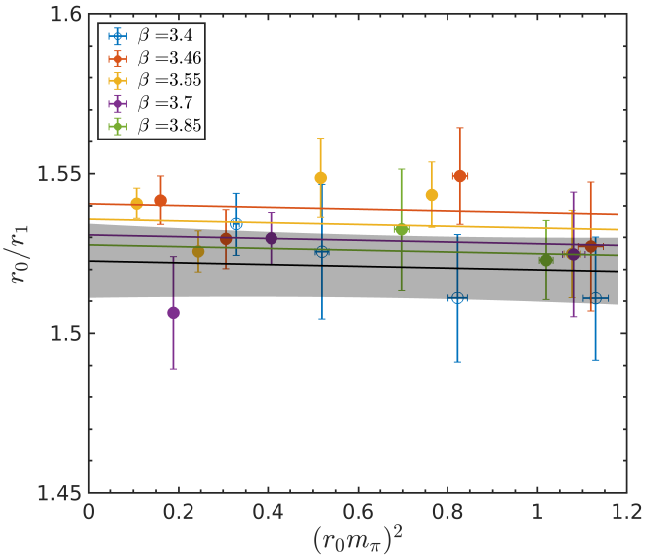

FIGURE 5.4: Ratio r_0/r_1 as a function of the light-quark mass proxy $(r_0 m_\pi)^2$ for all ensembles, together with a global fit of the form Equation (5.3). Data at fixed inverse bare coupling β are shown with common colours. The grey band indicates the continuum-limit result. The weak mass dependence and absence of visible cutoff effects support the use of a simple linear ansatz.

A second class of fits ("Fit 3" and "Fit 4") introduces explicit dependence on the mistuning parameter $\phi_4 - \phi_4^{\text{phys}}$, thereby absorbing residual quark-mass mismatches into the fit rather than neglecting them. In these variants, the leading cutoff term is parameterised either in units of r_0^{sym} or of $\sqrt{t_0^{\text{sym}}}$, which provides an additional cross-check on the robustness of the continuum limit. Concretely, Fit 3 and Fit 4 are defined as

$$\frac{r_0}{\sqrt{t_0}}(\beta, m_\pi^2, \phi_4) = c_1^{(3)} + c_2^{(3)} \left(\frac{a}{r_0^{\text{sym}}}\right)^2 + c_3^{(3)}\,\phi_2 + c_4^{(3)}\left(\phi_4 - \phi_4^{\text{phys}}\right), \tag{5.5}$$

$$\frac{r_0}{\sqrt{t_0}}(\beta, m_\pi^2, \phi_4) = c_1^{(4)} + c_2^{(4)} \left(\frac{a}{\sqrt{t_0^{\text{sym}}}}\right)^2 + c_3^{(4)}\,\phi_2 + c_4^{(4)}\left(\phi_4 - \phi_4^{\text{phys}}\right). \tag{5.6}$$

The role of the different fit ansätze and data cuts, and their impact on the final continuum-limit results, is quantified systematically in the following section.

5.4 Systematic variations and model selection

The continuum and chiral extrapolations described in the previous section rely on a number of assumptions concerning the functional form of cutoff effects, the parameterisation of the quark-mass dependence, and the range of lattice spacings included in the fits. While the baseline ansatz Fit 1 in Equation (5.3) provides a compact and well-motivated description of the data and is sufficient to illustrate their qualitative behaviour, it is not used in isolation to define the final results. In this section, systematic uncertainties associated with the extrapolation procedure are quantified by considering

a set of alternative fit ansätze in Equations (5.4) to (5.6) and data selections. These variations are designed to probe the sensitivity of the continuum-limit results to assumptions about leading scaling behaviour, possible subleading cutoff effects, and the treatment of residual quark-mass mistunings. A central element is the use of explicit cuts in the lattice spacing and pion mass. In particular, the coarsest lattice spacing is excluded from the final determination of the scale, as discretisation effects are expected to be largest there and cannot be reliably disentangled from higher-order cutoff contributions within the present precision. Fits including this lattice spacing are nevertheless performed for diagnostic purposes, in order to assess the stability of the extrapolation and to visualise trends in the data. The different fit strategies considered in the following are grouped into three broad classes. First, extensions of the baseline ansatz are introduced to allow for corrections to leading $\mathcal{O}(a^2)$ scaling violations. Second, alternative parameterisations of the quark-mass dependence are employed to account explicitly for residual mistunings in the light and strange quark masses. Finally, variations in the parametrisation of the cutoff terms are explored as a cross-check of the robustness of the continuum limit. The impact of these systematic variations on the extrapolated continuum values is assessed by comparing the resulting fit outcomes and their associated uncertainties. The procedure used to combine the results of different fits into a final estimate, including the assignment of systematic errors, is described at the end of this section.

5.4.1 Corrections to leading scaling violations

The continuum extrapolations discussed so far are based on the assumption that lattice artefacts are dominated by pure $O(a^2)$ effects. For the action used in this work, Symanzik effective theory [96] predicts additional subleading $O(a^3)$ artifacts, which we neglect. More importantly, the leading scaling violations themselves are not expected to be of a strictly polynomial $O(a^2)$ form, but to receive logarithmic modifications of the type

$$O\Big(a^2 \ln(a\Lambda)^{-\hat{\Gamma}}\Big), \tag{5.7}$$

where the exponents $\hat{\Gamma}$ are related to anomalous dimensions of the dimension–six operators entering the Symanzik action [97–99]. For our particular action, perturbation theory yields representative values such as $\hat{\Gamma} \in \{-0.11111,\, 0.24731,\, 0.51852, \dots\}$, and further logarithmic corrections depending on the specific observable are expected.

A global fit that explicitly includes all such terms is impractical. To assess the potential impact of logarithmic corrections in a controlled way, we therefore revisit Fit 1 from Equation (5.3) and replace its leading a^2 term by a modified ansatz of the form

$$\left(\frac{a}{r_0^{\mathrm{sym}}}\right)^2 \longrightarrow \left(\frac{a}{r_0^{\mathrm{sym}}}\right)^2 \left[\ln\left(\frac{r_0^{\mathrm{sym}}}{a}\right)\right]^{\hat{\Gamma}}, \tag{5.8}$$

where $\hat{\Gamma}$ is varied within a reasonable range motivated by the perturbative values. For each choice of $\hat{\Gamma}$ we repeat the global fits for $r_0/\sqrt{t_0}$, $r_1/\sqrt{t_0}$ and r_0/r_1, and we do so for the same sets of lattice-spacing cuts as in the standard analysis. The resulting continuum values are summarised in Figure 5.5. For r_0 and r_1, a subset of the corresponding fits is also shown explicitly in Figure 5.6, where the data have been shifted to the chiral limit by subtracting the mass dependence from the fit function.

As can be seen from Figure 5.5, the spread of continuum-limit values induced by

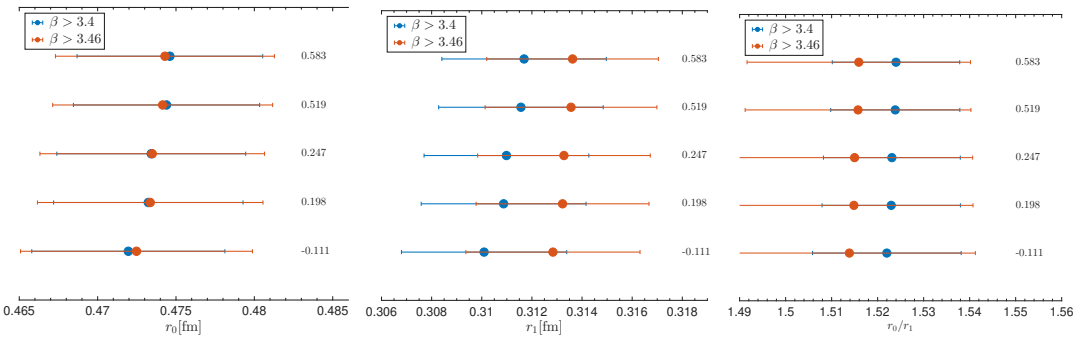

FIGURE 5.5: Impact of logarithmic corrections on the continuum values of r_0 (left), r_1 (centre) and r_0/r_1 (right). Each horizontal band shows the result of a global fit in which the leading lattice artefact term is modified as $(a/r_0^{\mathrm{sym}})^2 \to (a/r_0^{\mathrm{sym}})^2 \left[\ln(r_0^{\mathrm{sym}}/a)\right]^{\hat{\Gamma}}$ for a given exponent $\hat{\Gamma}$. Different colours correspond to different cuts in the coarsest lattice spacings. The value of $\hat{\Gamma}$ used in each fit is printed next to the corresponding error bar.

varying $\hat{\Gamma}$ is comparable to, or smaller than, the spread obtained by omitting the coarsest one or two lattice spacings at fixed $\hat{\Gamma} = 0$. The detailed fits displayed in Figure 5.6 confirm that the different logarithmic variants provide equally acceptable descriptions of the data. We therefore conclude that, at the level of precision targeted in this work, the dominant uncertainty associated with lattice artefacts is already captured by the variation of lattice-spacing cuts in the pure a^2 fits. For our final extrapolations we thus retain the simpler $O(a^2)$ description (corresponding to $\hat{\Gamma} = 0$) as our default ansatz.

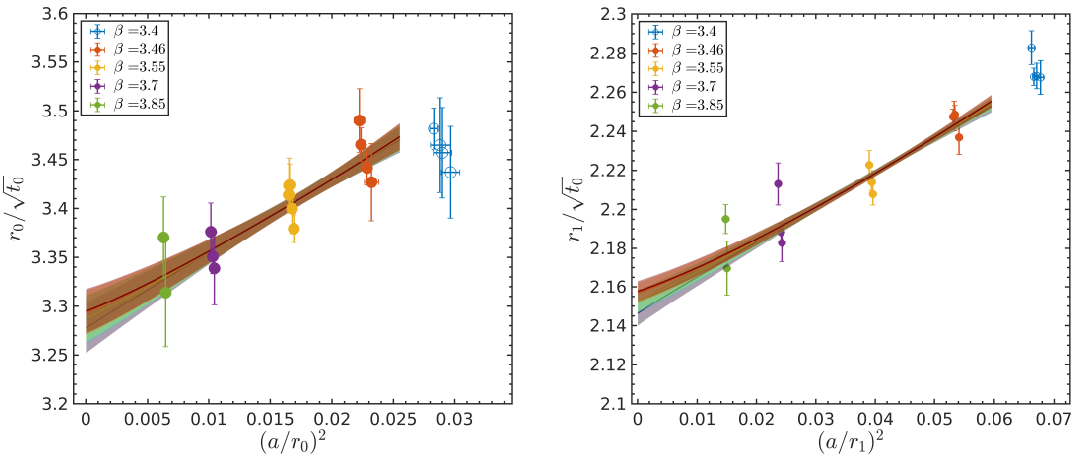

FIGURE 5.6: Examples of continuum extrapolations of $r_0/\sqrt{t_0}$ (left) and $r_1/\sqrt{t_0}$ (right) at the symmetric mass point for several choices of the logarithmic exponent $\hat{\Gamma}$ in the leading artefact term $(a/r_0^{\mathrm{sym}})^2\,[\ln(r_0^{\mathrm{sym}}/a)]^{\hat{\Gamma}}$. To combine all data in each panel, the ensembles have been extrapolated to the physical mass point. The coloured bands correspond to the same values of $\hat{\Gamma}$ as in Fig. 5.5, with the lowest band representing $\hat{\Gamma} = -0.111$ and the highest $\hat{\Gamma} = 0.583$.

5.4.2 Combination of fit results and uncertainty estimate

To further assess systematic uncertainties associated with the continuum and chiral extrapolation, several cuts on the data are applied in addition to the variations of the fit ansätze discussed in the previous sections. In particular, one or two of the coarsest lattice spacings are excluded, and ensembles with pion masses far from the physical point are removed. The latter cut eliminates all data at the symmetric mass point and therefore probes the sensitivity of the extrapolation to the mass range included in the fits. For each fit ansatz and data selection, the extrapolated values of r_0, r_1, and the ratio r_0/r_1 at the physical point are determined, together with their statistical uncertainties and the corresponding χ^2/d.o.f. values, which provide a measure of the fit quality and will later be used in Equation (5.10) to calculate the weights of a given fit/cut-combination. The results of all fit methods and cuts are collected in Table 5.2. Overall, the extrapolated values obtained from the different fits and cuts are found to be mutually consistent within uncertainties. The primary effect of the data cuts is a modest increase in the statistical errors, reflecting the reduced number of data points included in the fits. The largest spread among the fit results is observed for r_1 when discarding the two coarsest lattice spacings. In this case, the slightly outlying E300 receives a comparatively larger statistical weight, leading to a small shift of the central value. This behaviour is consistent with the relative precision of the individual data points and does not indicate a systematic instability of the continuum extrapolation.

In an initial stage of the analysis, the possibility of correcting residual mistunings by computing quark-mass derivatives of the observables with respect to ϕ_4 was also considered, as mentioned in the end of Section 5.2. However, since Fits 3 and 4, which explicitly parameterise such mistunings, yield results consistent with those obtained without explicit corrections, and since the corresponding derivatives are rather noisy for purely gluonic observables, no derivative-based corrections are applied in the final analysis.

The final values of the observables are obtained by combining the results of the

Fit #	Cut Method	r_0[fm]	χ^2 /d.o.f.	r_1[fm]	χ^2 /d.o.f.	r_0/r_1	χ^2 /d.o.f.
Fit 1	No cuts	0.4758(57)	12.6/15	0.3100(32)	30.4/15	1.5344(95)	13.6/15
	$m_\pi < 400$MeV	0.4750(57)	10.5/10	0.3100(32)	27.5/10	1.5325(94)	12.0/10
	β>3.4	0.4724(61)	9.7/11	0.3104(33)	27.0/11	1.5224(127)	10.0/11
	β>3.46	0.4728(73)	7.2/7	0.3130(35)	18.4/7	1.5142(199)	7.6/7
	$m_\pi < 400$MeV & β>3.4	0.4716(62)	8.1/7	0.3104(33)	25.0/7	1.5198(123)	8.5/7
Fit 2	No cuts	0.4763(57)	11.1/14	0.3099(32)	30.1/14	1.5341(93)	13.1/14
	$m_\pi < 400$MeV	0.4778(59)	4.2/9	0.3102(33)	23.6/9	1.5396(95)	8.1/9
	β>3.4	0.4730(62)	8.6/10	0.3104(33)	26.3/10	1.5222(125)	9.8/10
	β>3.46	0.4731(73)	6.5/6	0.3127(35)	18.2/6	1.5124(186)	7.2/6
	$m_\pi < 400$MeV & β>3.4	0.4745(64)	2.3/6	0.3107(33)	21.0/6	1.5272(126)	5.5/6
Fit 3	No cuts	0.4762(63)	12.4/14	0.3099(33)	30.7/14	1.5350(105)	13.4/14
	$m_\pi < 400$MeV	0.4755(68)	10.4/9	0.3099(33)	27.7/9	1.5351(149)	11.5/9
	β>3.4	0.4724(78)	9.8/10	0.3104(35)	27.2/10	1.5225(152)	10.0/10
	β>3.46	0.4724(175)	7.2/6	0.3144(49)	16.4/6	1.5114(295)	7.5/6
	$m_\pi < 400$MeV & β>3.4	0.4716(86)	8.1/6	0.3103(36)	25.0/6	1.5213(230)	8.5/6
Fit 4	No cuts	0.4769(61)	12.3/14	0.3106(33)	30.0/14	1.5351(109)	13.0/14
	$m_\pi < 400$MeV	0.4762(66)	10.3/9	0.3106(33)	27.1/9	1.5350(141)	11.4/14
	β>3.4	0.4733(73)	9.66/10	0.3108(35)	26.8/10	1.5231(160)	10.0/14
	β>3.46	0.4729(166)	7.1/6	0.3145(48)	16.4/6	1.5121(334)	8.1/14
	$m_\pi < 400$MeV & β>3.4	0.4726(82)	8.0/6	0.3107(36)	24.6/6	1.5220(216)	8.3/14

TABLE 5.2: Results of the continuum and physical-point extrapolations for the potential scales r_0, r_1, and their ratio r_0/r_1 obtained from the different fit ansätze and data cuts considered in the systematic analysis. For each fit and cut, the table lists the extrapolated value and the corresponding χ^2/d.o.f.. These results form the basis of the AIC-weighted averages shown in Figures 5.7 and 5.8.

accepted fits using an information-theoretic weighting procedure based on the Akaike Information Criterion (AIC) [100–103]. The AIC-weighted average of an observable $\langle\mathcal{O}\rangle$ is defined as

$$\langle\mathcal{O}\rangle = \sum_{n=1}^{N_{\text{fits}}} w_n^{\text{AIC}} \langle\mathcal{O}\rangle_n, \tag{5.9}$$

where the sum runs over all fits included in the final analysis. The AIC weight w_n^{AIC} for a given fit and data cut labelled by n is given by

$$w_n^{\text{AIC}} = \mathcal{N} \exp\left[-\frac{1}{2}\left(\chi_n^2 + 2N_n^{\text{par}} + 2N_n^{\text{cut}}\right)\right], \tag{5.10}$$

where N_n^{par} denotes the number of fit parameters, N_n^{cut} the number of excluded data points, and $\mathcal{N}$ is a normalisation factor chosen such that $\sum_n w_n^{\text{AIC}} = 1$. The AIC-weighted averages for r_0 and r_1 at the physical point are illustrated in Figure 5.7. The corresponding analysis is repeated for the ratio r_0/r_1, for which the effect of the cuts is found to be more pronounced than for the individual scales, reflecting the higher sensitivity of this ratio to the details of the extrapolation. The outcome of this procedure for the ratio r_0/r_1 is shown in Figure 5.8. Each point represents the continuum and physical-point value obtained from a single fit ansatz and data cut, while the corresponding AIC weights quantify their relative contribution to the final average. Fits including the coarsest lattice spacing are displayed for comparison but are not included in the AIC average. Although the spread of individual fit results is somewhat larger for r_0/r_1 than for the individual scales $r_0/\sqrt{t_0}$ and $r_1/\sqrt{t_0}$, all

accepted fits are mutually consistent, and the final AIC-weighted result is stable within the quoted uncertainties.

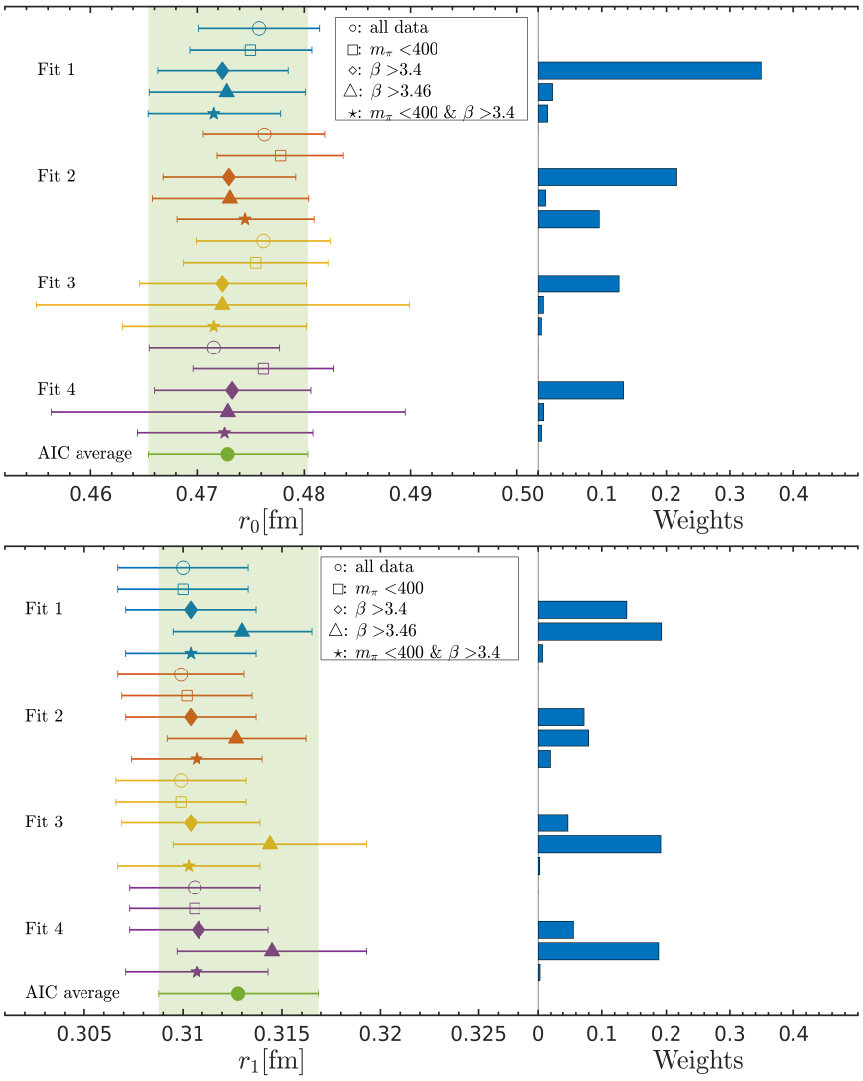

FIGURE 5.7: Summary of the physical-point determinations of r_0 (top) and r_1 (bottom) obtained from the different fit ansätze and data cuts listed in Table 5.2. Each marker corresponds to the continuum-extrapolated value from a single fit, while the horizontal bars indicate the associated statistical uncertainties. The bars on the right show the corresponding AIC weights defined in Equation (5.10). Fits including the coarsest lattice spacing are shown as open symbols and are excluded from the final AIC average. The green marker indicates the final AIC-weighted result.

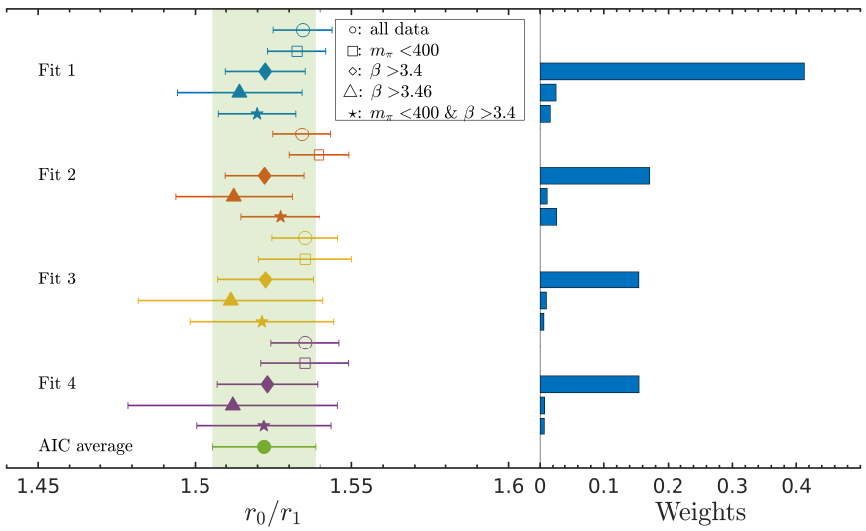

FIGURE 5.8: Continuum and physical-point values of r_0/r_1 obtained from the different fit ansätze and data cuts listed in Table 5.2. The bars on the right show the corresponding AIC weights, cf. Equation (5.10). Open symbols denote fits including the coarsest lattice spacing, which are excluded from the final AIC average. The green marker indicates the final result.

5.5 Final potential scale results and comparison with previous determinations

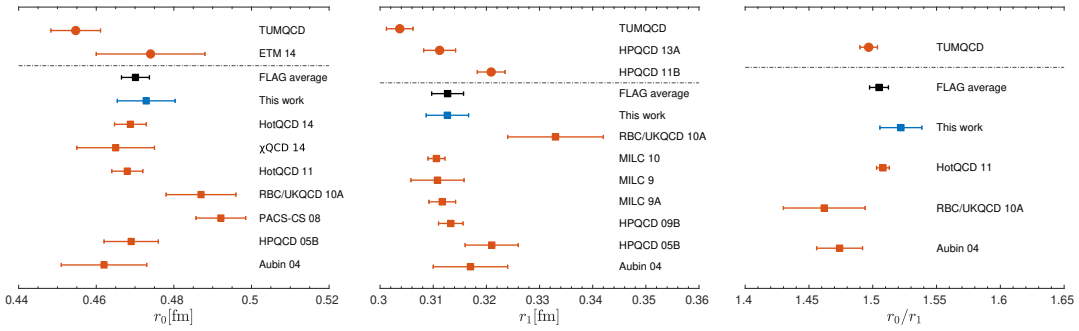

FIGURE 5.9: Comparison of the present results for r_0 (left), r_1 (middle), and the ratio r_0/r_1 (right) with previous determinations. Results obtained in $N_f = 2+1$ QCD are shown together with the FLAG averages [74], while circular markers indicate determinations in $N_f = 2+1+1$ QCD. The results from this work are consistent with earlier determinations within uncertainties. [104–118]

The final outcome of the analysis presented in this chapter is the determination of the potential scales r_0 and r_1 from the static quark–antiquark potential, as well as their ratio, at the physical mass point of $N_f = 2+1$ QCD. The values quoted below are obtained from the AIC-weighted averages described in Section 5.4.2 and include both statistical and systematic uncertainties associated with the continuum and chiral extrapolations.

Using the reference value of $\sqrt{t_0} = 0.1443(7)(13)\text{fm}$ [60], the resulting scales are

$$r_0 = 0.4729(57)(48)\,\text{fm}, \tag{5.11}$$

$$r_1 = 0.3127(24)(32)\,\text{fm}, \tag{5.12}$$

and for their ratio

$$\frac{r_0}{r_1} = 1.532(12). \tag{5.13}$$

For r_0 and r_1, the first uncertainty originates from the determination of the corresponding dimensionless ratios $r_i/\sqrt{t_0}$ in the continuum limit, while the second uncertainty reflects the error of the input scale $\sqrt{t_0}$ in physical units. These two contributions should be combined in quadrature. The autocorrelation tail from the error analysis introduced in Section 3.6 increased the final error by $\tilde{1}5\%$. For completeness, the corresponding continuum-limit values of the dimensionless ratios are

$$\frac{r_0}{\sqrt{t_0}} = 3.277(39), \tag{5.14}$$

$$\frac{r_1}{\sqrt{t_0}} = 2.167(16). \tag{5.15}$$

The results for r_0, r_1, and r_0/r_1 are compared with previous determinations in Figure 5.9. Shown are results obtained in $N_f = 2 + 1$ QCD, together with the FLAG averages, as well as selected determinations in $N_f = 2 + 1 + 1$ QCD [104–119]. The latter are indicated by circular markers. Previous studies [120, 121] have shown that the effect of a dynamical charm quark on low-energy scales is typically at the per-mille level, and agreement between $N_f = 2 + 1$ and $N_f = 2 + 1 + 1$ results is therefore expected at the present level of precision. Overall, the values obtained in this work are consistent with earlier determinations within uncertainties. At the same time, the spread among the published central values, particularly for r_0 and r_1, is somewhat larger than expected from the quoted errors alone, indicating that residual systematic effects may not be fully captured uniformly across different analyses. The comparatively conservative uncertainty estimates obtained in the present work reflect the explicit treatment of continuum-extrapolation, fit-model, and autocorrelation systematics. Beyond the agreement with previous determinations, it is worth emphasizing that these results are obtained from a large set of modern CLS $N_f = 2 + 1$ ensembles covering several lattice spacings, including very fine lattices with open temporal boundary conditions and quark masses close to the physical point. This broad parameter coverage enables controlled simultaneous continuum and chiral extrapolations. Combined with the systematic treatment of excited-state effects, autocorrelations, and fit model uncertainties through the AIC averaging procedure, the present analysis provides a precise and robust determination of the potential scales r_0 and r_1. These results therefore constitute an important benchmark calculation of the potential scales.

Chapter 6

Phenomenological applications and implications

6.1 The Λ-parameter in units of r_0 for $N_f = 2 + 1$ QCD

The strong coupling constant α_s represents a fundamental parameter in quantum chromodynamics (QCD), and its precise determination is essential for accurate predictions across a wide range of phenomena. The value of α_s at any renormalisation scale can be obtained by solving its renormalisation-group equation, provided the dimensionful Λ-parameter, introduced in Section 2.6.1, in the $\overline{\rm MS}$ scheme is known. Lattice QCD has achieved remarkable success in computing the most precise values of $\Lambda_{\overline{\rm MS}}$ by incorporating inputs from low-energy experiments, such as pseudoscalar meson masses and decay rates [122, 123]. To circumvent complications associated with absolute scale setting, lattice collaborations frequently report dimensionless ratios, such as $r_0\Lambda_{\overline{\rm MS}}$, $\sqrt{t_0}\Lambda_{\overline{\rm MS}}$, or $w_0\Lambda_{\overline{\rm MS}}$, enabling direct comparisons across studies. In recent developments, gradient-flow-based scales have largely supplanted r_0, as exemplified by the ALPHA collaboration's computation [124]. The analysis of [124] was performed using two distinct definitions of $L_{\rm had}$, both specified implicitly through fixed reference values of the gradient-flow coupling $\bar{g}^2_{\rm GF}(L^{-1}_{\rm had})$. The corresponding conditions read

$$\bar{g}^2_{\rm GF}(L^{-1}_{\rm had,1}) = 11.31 \quad \Rightarrow \quad L_{\rm had,1}\Lambda^{(3)}_{\overline{\rm MS}} = 1.729(57), \tag{6.1}$$

$$\bar{g}^2_{\rm GF}(L^{-1}_{\rm had,2}) = 10.20 \quad \Rightarrow \quad L_{\rm had,2}\Lambda^{(3)}_{\overline{\rm MS}} = 1.593(53), \tag{6.2}$$

which will be used in our calculation for $r_0\Lambda_{\overline{\rm MS}}$. The corresponding finite-lattice determinations of $L_{{\rm had},i}/a$ are provided in Table V of the supplementary material of [124].

While one could derive $r_0\Lambda^{(3)}_{\overline{\rm MS}}$ by multiplying the original $\sqrt{t_0}\Lambda^{(3)}_{\overline{\rm MS}}$ with our $r_0/\sqrt{t_0}$ ratio, entailing two separate continuum extrapolations, our use of the identical lattice action and improved bare couplings at matching ensembles permits direct extrapolation of $(r_0/a)/(L_{\rm had}/a)$ using our r_0/a results on Table 4.3. This approach eliminates reliance on the flow scale t_0 entirely. The resulting continuum extrapolations, displayed in fig. 6.1, exhibit excellent behaviour, yielding

$$\frac{r_0}{L_{\rm had,1}} = 0.4755(39) \quad \Rightarrow \quad r_0\Lambda^{(3)}_{\overline{\rm MS}} = 0.822(28)\,, \tag{6.3}$$

$$\frac{r_0}{L_{\rm had,2}} = 0.5149(41) \quad \Rightarrow \quad r_0\Lambda^{(3)}_{\overline{\rm MS}} = 0.820(28)\,, \tag{6.4}$$

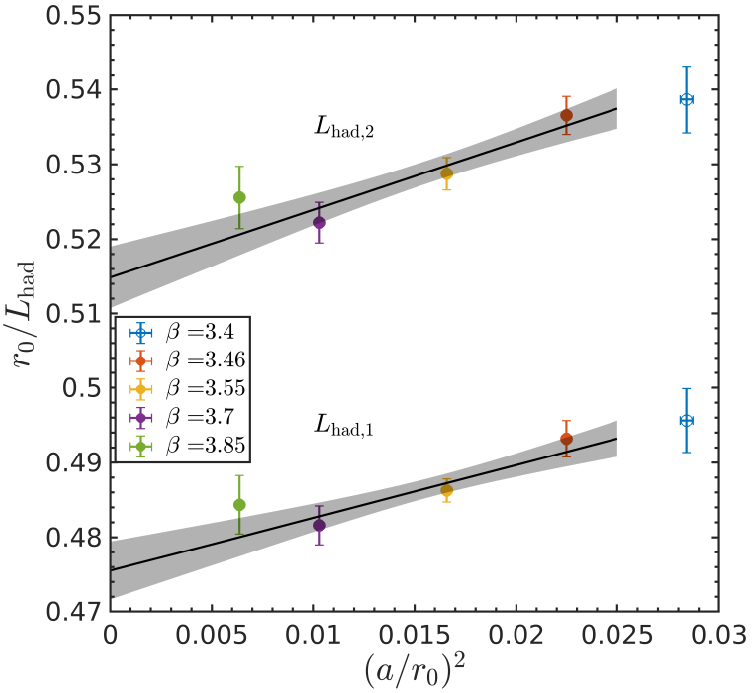

FIGURE 6.1: Continuum extrapolation of the ratio r_0/a to $L_{\rm had}/a$ as a function of $(a/r_0)^2$ for the two definitions of the hadronic scale $L_{\rm had,1}$ and $L_{\rm had,2}$; data points correspond to different bare couplings β, and the grey bands indicate the continuum extrapolation.

These results align closely with the FLAG average [74] $r_0\Lambda^{(3)}_{\overline{\rm MS}} = 0.808(29)$, incorporating computations from Refs. [124–130]. We adopt Equation (6.4) as our final result, given its consistency and self-contained error assessment.

6.2 Shape analysis of the static potential

Following the discussion in Sections 2.5.3 to 2.5.6, the curvature observable is now evaluated using the nonperturbative lattice determinations of the static potential. Beyond the determination of reference scales such as the potential scales, the static potential itself provides important information about the non-perturbative dynamics of QCD. In particular, the functional form of $V(r)$ reflects the interplay between short-distance perturbative behaviour and long-distance confinement physics. Although the primary focus of this thesis is the determination of reference scales, the same analysis also yields direct information on the underlying static quark–antiquark interaction. As discussed in Section 2.5, the static potential is defined from the large Euclidean time behaviour of rectangular Wilson loops,

$$V(r) = -\lim_{t\to\infty} \frac{1}{t} \ln W(r,t), \tag{6.5}$$

and contains an additive ultraviolet self-energy contribution associated with the static sources. While this additive constant does not affect physical observables derived from derivatives of the potential, it must be taken into account when comparing the potential itself with phenomenological models. In practice it is therefore convenient to study quantities that depend on the shape of the potential rather than its overall normalisation. One such observable is the dimensionless curvature combination

introduced in Equation (2.50),

$$c(r) = \frac{1}{2} r^3 V''(r), \tag{6.6}$$

which is particularly sensitive to how the force changes with distance. Unlike the potential scales r_0 and r_1, which probe the potential at isolated reference points, $c(r)$ provides information over a continuous range of separations. Since it is invariant under additive shifts of $V(r)$, it directly probes the functional form of the potential. As already introduced in Equation (2.58), at short distance perturbation theory predicts Coulombic behaviour

$$c(r) \to -\frac{C_F g^2}{4\pi}, \tag{6.7}$$

where g denotes the strong coupling at a scale of order $1/r$. At larger separations the departure from this Coulombic limit reflects the diminishing influence of the short-distance $1/r$ term and the emergence of the approximately linear confining component of the potential. The observable $c(r)$ therefore provides a useful diagnostic of the crossover between asymptotic freedom and confinement. To reduce discretisation effects arising from finite differences of the potential, we evaluate $c(r)$ using the tree-level improved definition introduced in Section 2.5.6,

$$c(\tilde{r}) = \frac{1}{2} \tilde{r}^3 \frac{V(r+a) + V(r-a) - 2V(r)}{a^2}, \tag{6.8}$$

where the improved distance $\tilde{r}$ is defined through the lattice Coulomb Green's function such that the tree level in perturbation theory Equation (6.7) holds. This procedure removes the leading tree-level lattice artefacts associated with the discrete derivative operator and significantly improves the scaling behaviour at finite lattice spacing. The resulting curvature observable provides a direct measure of the shape of the static potential and can be compared to expectations from phenomenological potential models. In particular, the Cornell potential

$$V_{\text{Cornell}}(r) = -\frac{\kappa}{r} + \sigma r \tag{6.9}$$

captures the combination of Coulombic short-distance behaviour and linear confinement at large separations in the Cornell model $c = -\kappa$. A more sophisticated description is given by the Richardson potential, which incorporates the perturbative running of the coupling at short distances while retaining confining behaviour at larger r.

The resulting values of $c(\tilde{r})$ for all CLS ensembles are collected in Table 6.1. Figure 6.2 shows the curvature observable $c(\tilde{r})$ for the ensembles E250 and J500 as a function of the dimensionless separation $\tilde{r}/r_0$. The data points are obtained from the lattice potential using the improved curvature definition described above. For comparison, the curves correspond to the predictions of the Cornell and Richardson potentials, as well as the perturbative expectation at short distances. At the smallest accessible separations the lattice results approach the perturbative prediction for the finest lattice, hinting that the short-distance behaviour of the static potential is well described by perturbation theory. As the separation increases, clear deviations from the perturbative curve emerge and the data enter the intermediate-distance regime where non-perturbative confinement effects become relevant. In this region the curvature observable shows a non-trivial dependence on r, rather than the constant behaviour implied by the Cornell form $c(r) = -\kappa$. The observed trend is therefore more naturally described by the Richardson potential, whose scale-dependent curvature reflects the

running of the coupling while retaining confining behaviour at larger distances. In particular, the smooth interpolation between the perturbative short-distance regime and the non-perturbative confining regime demonstrates that the lattice calculation reproduces the expected physical shape of the potential across the relevant range of scales.

Figure 6.3 on the left compares the curvature observable obtained from the CLS ensemble at $N_f = 2+1$ with earlier lattice determinations for $N_f = 0$ [40] and $N_f = 2$ [38]. The comparison shown in the left panel indicates a visible dependence of the curvature observable on the number of dynamical quark flavours. In particular, the quenched ($N_f = 0$) results differ systematically from the dynamical simulations. By contrast, the right panel shows that for ensembles with similar lattice spacing but different pion masses, the dependence on the light-quark mass is comparatively mild within the present uncertainties.

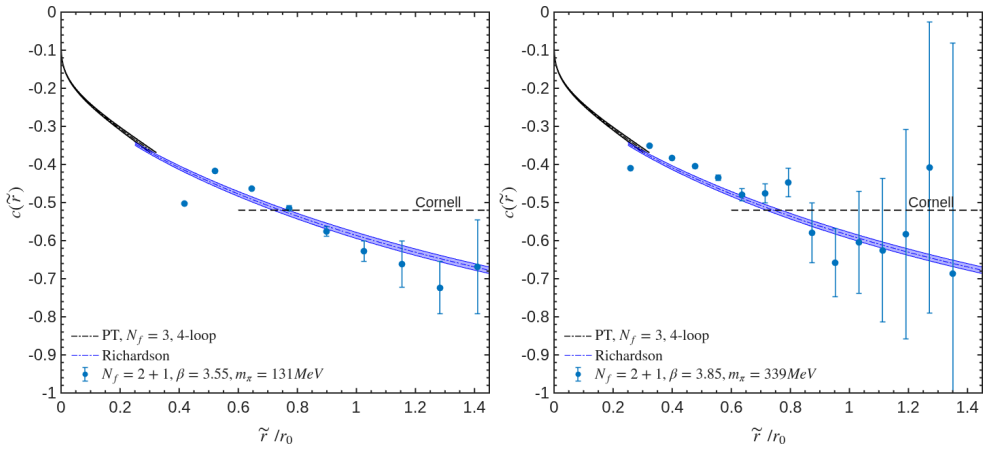

FIGURE 6.2: Shape of the static quark–antiquark potential expressed through the curvature observable $c(\tilde{r})$ defined in Eq. (6.6). The lattice data are shown for the ensembles E250 (left) and J501 (right) as a function of the improved separation $\tilde{r}/r_0$. The dashed line indicates the perturbative prediction, while the solid curves correspond to the Cornell and Richardson phenomenological potentials. The comparison illustrates the transition from perturbative Coulomb behaviour at short distances to the non-perturbative confining regime at larger separations.

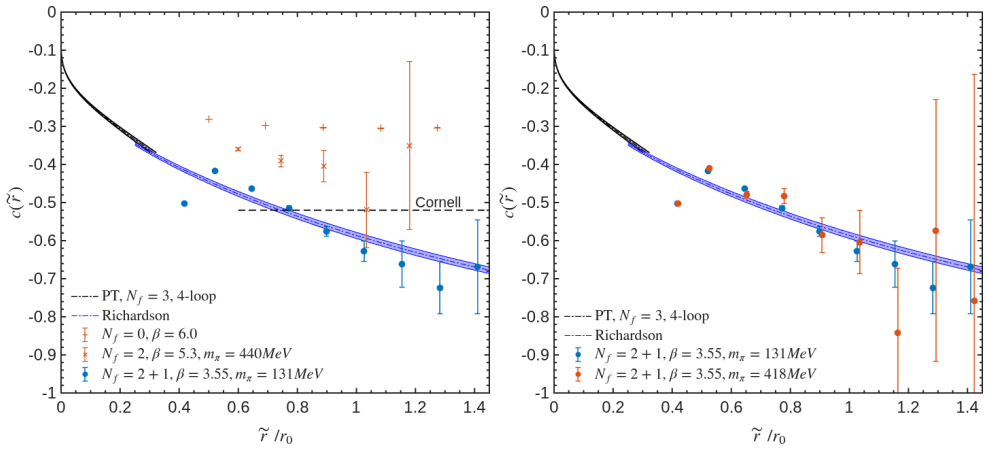

FIGURE 6.3: Comparison of the curvature observable $c(\tilde{r})$ as a function of the improved separation $\tilde{r}/r_0$ with earlier lattice determinations (left) and between different m_π (right). The blue points correspond to the CLS ensembles E250 and the red points on the right N202 obtained in the present work. The additional data sets shown on the left panel for $N_f = 0$ and $N_f = 2$ are taken from Refs. [40] and [38], respectively.

$\tilde{r}/a$	H101	H102-0	H102-1	H105	C101	B450	S400	D450	D452
3.247	-0.5643(27)	-0.5714(39)	-0.564(4)	-0.564(2)	-0.565(2)	-0.538(2)	-0.538(2)	-0.539(1)	-0.540(1)
4.053	-0.4429(90)	-0.4655(124)	-0.443(10)	-0.451(8)	-0.453(4)	-0.434(8)	-0.445(6)	-0.441(3)	-0.439(3)
5.030	-0.4997(284)	-0.4979(380)	-0.454(34)	-0.523(25)	-0.541(14)	-0.494(24)	-0.511(19)	-0.503(11)	-0.501(9)
6.015	-0.6343(799)	-0.7748(1171)	-0.647(91)	-0.69(75)	-0.627(34)	-0.590(62)	-0.617(47)	-0.537(23)	-0.597(24)
7.001	-0.6577(214)	-0.9674(2634)	-0.457(241)	-0.798(179)	-0.634(92)	-0.564(123)	-0.481(103)	-0.599(64)	-0.622(53)
7.992			-0.637(581)	-1.21(41)	-0.836(209)	-0.122(283)	-0.438(222)	-0.545(130)	-0.761(97)
8.987						-0.394(504)		-0.642(238)	-0.631(195)

$\tilde{r}/a$	N202	N203	N200	D200	E250	N300-0	N300-1	J303	E300	J500	J501
3.247	-0.503(1)	-0.502(1)	-0.502(1)	-0.5018(4)	-0.5020(3)	-0.450(1)	-0.450(1)	-0.4515(3)	-0.4514(2)	-0.4090(2)	-0.4092(2)
4.053	-0.410(3)	-0.414(2)	-0.416(2)	-0.414(1)	-0.417(1)	-0.378(4)	-0.378(2)	-0.379(1)	-0.381(1)	-0.349(1)	-0.3501(4)
5.030	-0.479(9)	-0.463(6)	-0.456(6)	-0.464(3)	-0.463(2)	-0.424(10)	-0.422(7)	-0.416(2)	-0.419(2)	-0.381(1)	-0.383(2)
6.015	-0.483(19)	-0.523(17)	-0.531(15)	-0.506(11)	-0.514(6)	-0.434(22)	-0.447(14)	-0.444(6)	-0.463(4)	-0.412(3)	-0.405(4)
7.001	-0.585(46)	-0.597(33)	-0.563(31)	-0.546(24)	-0.576(12)	-0.453(41)	-0.491(29)	-0.490(12)	-0.476(8)	-0.436(6)	-0.435(6)
7.992	-0.604(83)	-0.569(71)	-0.681(66)	-0.616(44)	-0.628(26)	-0.462(82)	-0.542(52)	-0.518(23)	-0.506(18)	-0.455(11)	-0.479(16)
8.987	-0.841(169)	-0.689(122)	-0.520(132)	-0.766(79)	-0.662(61)	-0.531(153)	-0.480(82)	-0.576(33)	-0.547(26)	-0.480(20)	-0.476(25)
9.984	-0.573(344)	-0.564(233)	-0.255(228)	-0.513(142)	-0.724(68)	-0.537(250)	-0.575(144)	-0.544(61)	-0.660(49)	-0.438(40)	-0.447(37)
10.983	-0.758(595)	-0.743(385)	-0.764(396)	-0.865(206)	-0.669(123)	-0.617(451)	-0.666(208)	-0.717(100)	-0.760(84)	-0.576(62)	-0.579(78)
11.982	-0.845(861)			-0.254(390)	-0.904(208)		-0.402(343)	-0.640(166)	-0.610(97)	-0.642(95)	-0.658(89)
12.982							-0.540(566)	-0.799(246)	-0.914(243)	-0.511(142)	-0.604(134)
13.983							-0.957(827)			-0.350(190)	-0.626(188)
14.983										-0.754(283)	-0.583(275)
15.984										-0.253(403)	-0.409(382)
16.984										-0.546(488)	-0.687(606)
17.985										-0.618(606)	
18.985										-0.95(777)	

TABLE 6.1: Values of the curvature observable $c(\tilde{r})$ evaluated at the improved lattice separations $\tilde{r}/a$ for the CLS ensembles considered in this work.

Part ๓

The Return of Teper-link Smearing

Chapter 7

Teper-link smearing: theory and implementation

We now return to smearing techniques in the context of glueball spectroscopy, where the construction of gauge-invariant operators with good overlap onto physical gluonic states is essential for obtaining reliable signals. In previous work of our research group (see [19]), such operators were constructed from spatial loop shapes built from APE-smeared gauge links. The purpose of the present chapter is to investigate whether the signal construction can be improved by combining APE-smearing with Teper-links. The blocking procedure introduced by Teper [17], and explained in Section 2.7.3, generates extended composite transporters by iteratively combining neighbouring gauge links and staple-shaped paths. The resulting operators represent superpositions of many gauge paths and therefore possess a larger spatial extent than those obtained from local link-smearing alone. Such extended structures may provide a better overlap with the wavefunctions of low-lying gluonic states. In this study we therefore extend the existing APE-based operator [19, 21–23] construction by incorporating Teper blocking and examine its impact on the resulting glueball correlation functions. In particular, we compare operators constructed from APE-smeared links with operators obtained from Teper-blocked links as well as combinations of both procedures.

7.1 Glueball operators and lattice symmetry representations

In lattice QCD, the extraction of hadronic states such as glueballs relies on the construction of operators with well-defined transformation properties under the symmetries of the theory. In the continuum, states are classified according to the irreducible representations of the rotation group $SO(3)$, labelled by the total angular momentum J. However, the discretisation of space-time onto a hypercubic lattice explicitly breaks the continuous rotational symmetry down to the finite symmetry group of the cube, the octahedral group O (or O_h when parity is included). As a result, continuum quantum numbers such as J are no longer good labels at finite lattice spacing. To systematically construct operators that overlap with physical states, one instead classifies them according to the irreducible representations (irreps) of the cubic group. This ensures that correlation functions project onto definite symmetry channels, allowing for a clean separation of states in numerical simulations. In particular, glueball operators, typically built from spatial loop shapes, are organized into these irreps to probe the gluonic spectrum[26, 131].

The cubic group O has five irreducible representations [132–134], conventionally denoted

$$A_1, \quad A_2, \quad E, \quad T_1, \quad T_2, \tag{7.1}$$

with dimensions

$$\dim(A_1, A_2) = 1, \quad \dim(E) = 2, \quad \dim(T_1, T_2) = 3. \tag{7.2}$$

These dimensions correspond to the number of independent components transforming among themselves under the symmetry operations of the cube. Each representation can be further labeled by its transformation properties under parity (P) and, for gluonic operators, charge conjugation (C), leading to the full classification

$$\Gamma = R^{PC} \in \{A_1^{\pm}, A_2^{\pm}, E^{\pm}, T_1^{\pm}, T_2^{\pm}\}^{C=\pm}. \tag{7.3}$$

At finite lattice spacing, the continuous rotational symmetry is reduced to the cubic group, such that each lattice irrep R^{PC} receives contributions from several continuum spin states J^{PC}, which can only be disentangled in the continuum limit. That is, the mapping from continuum angular momentum J to lattice irreps is not one-to-one. For example, the A_1 irrep contains states with $J = 0, 4, \ldots$, while T_1 contains $J = 1, 3, \ldots$, and both E and T_2 receive contributions from $J = 2$ states. Therefore, a state observed in a given lattice irrep cannot be uniquely assigned a continuum spin at finite lattice spacing. Instead, continuum quantum numbers are recovered by studying degeneracy patterns across irreps as the lattice spacing is reduced and rotational symmetry is restored. In practice, operators transforming in a given irrep Γ are constructed by projecting gauge-invariant building blocks onto the desired representation. This is achieved by averaging over all symmetry-related orientations with appropriate group-theoretical weights. The resulting operators transform irreducibly under the cubic group O and provide a basis for variational analyses of correlation functions, enabling the extraction of the glueball spectrum. In this work we focus on the irreps A_1^{++}, A_1^{-+}, E^{++}, and T_2^{++}, which contain the lowest-lying scalar, pseudoscalar, and tensor glueball channels. These sectors are expected to yield the most robust signals and allow a direct study of the lightest states of phenomenological interest.

7.1.1 Projection onto irreducible representations

Having established that lattice operators must transform according to irreducible representations of the cubic group, the next step is to explicitly construct such operators from gauge-invariant building blocks. A first essential step is the projection to zero spatial momentum in Section 2.4.1 by summing over all spatial positions, which ensures that the operator creates states with definite (vanishing) momentum. Once projected to zero momentum, the cubic group can act on the loop shapes by spatial rotations. A loop of fixed shape s generates a set of up to 24 orientations, obtained by applying all proper cubic rotations (the elements of the octahedral group O) to the spatial path,

$$W_s^l(\vec{x}, t), \qquad l = 0, \ldots, 23, \tag{7.4}$$

corresponding to the elements of the octahedral group. These loops form a basis for a reducible representation of the cubic group. A list of the different rotations can be found in Section B.1. In the same appendix in Section B.2 a list over all loop shapes can be found, where the first 22 shapes of length up to 8 are found from [131] and the remaining with lengths up to 12 from [22]. Operators are then obtained by projecting

onto a given Irrep R. Following the construction in [22], one defines

$$W^{(R,k,s)}(t) = \frac{a^3}{L^3} \sum_{\vec{x}} \sum_{l=0}^{23} c_l^{R;k} \operatorname{Tr}\Big[W_s^{(l)}(\vec{x},t)\Big], \tag{7.5}$$

where $c_l^{R;k}$ are projection coefficients determined by the representation R, and $k = 0, \ldots, \dim(R) - 1$ labels the rows of the irrep. Further on, we will consider a single copy of each irrep and suppress the index k. The coefficients c_l^R are obtained from the standard group-theoretical projector onto irrep R of the cubic group, constructed by summing over all rotations with weights given by the representation matrices [131, 135, 136]. Acting with this projector on the basis of rotated loop orientations yields the linear combinations in Equation (7.5). Operators with definite parity and charge conjugation are constructed by combining each loop with its parity-transformed counterpart and taking the real or imaginary part of the trace, yielding operators labeled by $\Gamma = R^{PC}$. This construction ensures that the resulting operators transform irreducibly under the lattice symmetry group and provide a suitable basis for extracting glueball states in a given symmetry channel.

7.1.2 Construction of glueball operators

For each shape, one obtains operators transforming according to a given lattice irrep R by combining the rotated loops with the appropriate projection coefficients. Denoting by $W_s^{(l)}(\vec{x},t)$ a loop and by $W_{s,P}^{(l)}(\vec{x},t)$ its parity image, one constructs operators with definite parity $P = \pm$ as symmetric or antisymmetric combinations. Charge conjugation is imposed by taking the real ($C = +$) or imaginary ($C = -$) part of the trace. This follows because charge conjugation maps each link variable to its complex conjugate, so that the Wilson-loop trace is conjugated and its real (imaginary) part is even (odd) under C. This leads to operators of the form [22, 131]

$$O_s^{R^{\pm+}}(t) = \frac{a^3}{L^3} \sum_{\vec{x}} \sum_{l=0}^{23} c_l^R \operatorname{Re}\operatorname{Tr}\Big[W_s^{(l)}(\vec{x},t) \pm W_{s,P}^{(l)}(\vec{x},t)\Big], \tag{7.6}$$

$$O_s^{R^{\pm+}}(t) = \frac{a^3}{L^3} \sum_{\vec{x}} \sum_{l=0}^{23} c_l^R \operatorname{Im}\operatorname{Tr}\Big[W_s^{(l)}(\vec{x},t) \pm W_{s,P}^{(l)}(\vec{x},t)\Big]. \tag{7.7}$$

It should be emphasised that a given loop shape does not contribute to all irreducible representations R^{PC}. For a fixed shape, the set of loops generated by the action of the cubic group defines a representation of the cubic group, whose dimension is determined by the geometry of the underlying path. The irreducible content of this representation can be obtained through a character decomposition, as discussed in [131] (see eq. (3.13)). In particular, the multiplicity of a given irrep in this decomposition may vanish, implying that the corresponding projected operator is identically zero. This reflects the fact that the symmetry properties of a loop shape restrict the set of quantum numbers it can carry. This is particularly relevant for channels such as A_1^{-+}, where only few loop shapes of length 8 contribute. For example, the plaquette does not contain a T_2 or A_1^{-+} component and therefore cannot be used to construct operators transforming in that channel. A complete operator basis for a given channel requires combining multiple loop shapes. To avoid constructing identically vanishing operators, one must first determine which irreducible representations are contained in the decomposition associated with each loop shape, and restrict the operator basis

accordingly. In practice, this information can be tabulated for a given set of loop geometries, providing a systematic guide for selecting operators with non-zero overlap in the desired symmetry channels. One can review table 3.1 and 3.2 in [131] to see which of the first 22 shapes of loops up to length 8 can access which channels. The supplementary 13 shapes from [22] that we're using in Section B.2 can access all the 4 channels we're interested in.

In order to improve the overlap of these operators with the low-lying glueball states, the gauge links entering the loop shapes are typically smeared. Different levels of smearing and different loop shapes are then combined into a basis of operators,

$$O^{\Gamma}_{\alpha}(t), \qquad \alpha = 1, \ldots, N_{\Gamma}, \tag{7.8}$$

where the index α labels both the loop shape and the smearing level. The effectiveness of these operators can depend strongly on the smoothing procedures applied to the gauge links. From this operator basis, we construct the corresponding correlation matrix

$$C^{\Gamma}_{\alpha\beta}(t) = \langle O^{\Gamma}_{\alpha}(t)\, O^{\Gamma}_{\beta}(0)\rangle, \tag{7.9}$$

which generalises the two-point functions introduced in Section 2.4. The energies E_n associated with the quantum numbers Γ are extracted as described in Section 4.4 using the GEVP method. Similar to the static potential, the signal-to-noise ratio of purely gluonic operators deteriorates rapidly with Euclidean time, making the optimisation of the operator basis and the use of smearing techniques essential for a reliable extraction of the low-lying states [13].

Pruning of the operator basis

While the generalised eigenvalue problem provides a powerful framework for extracting energy levels from a basis of interpolating operators, its practical performance depends sensitively on the choice and size of the operator basis. In practice, the construction of operators from many loop geometries and multiple smearing levels leads to very large operator bases. While a large basis in principle increases the chance of overlapping with states of different spatial structures, strong correlations and near linear dependencies between operators often arise [137]. In the presence of strong correlations within the operator basis, the resulting correlation matrix can become nearly rank-deficient, where small statistical fluctuations may lead to large variations in the eigenvectors, rendering the GEVP ill-conditioned. This typically manifests itself in unstable eigenvalues and unreliable effective masses. The solution of the GEVP requires the reference correlation matrix to be well conditioned and positive definite. Operators that are not enough independent can produce very small eigenvalues, so that statistical fluctuations may strongly amplify numerical noise and, in extreme cases, spoil the stability of the problem. Another practical issue that can arise for closely spaced energy levels is *flipping*, i.e. an unstable assignment of eigenvalues at large Euclidean times [92, 138], which is particularly relevant in scenarios with avoided level crossings [139, 140].

To perform the variational analysis in a more stable subspace, a *pruning* (basis reduction) procedure is therefore applied. One selects a time slice t_s at which the states of interest are assumed to dominate the correlation matrix, and determines the N_s singular vectors u_i of $C^{\Gamma}(t_s)$ corresponding to the largest singular values [141, 142]. Using these vectors, a projected (pruned) correlation matrix is defined as

$$\widetilde{C}^{\Gamma}_{ij}(t) \;=\; u_i^{\dagger}\, C^{\Gamma}(t)\, u_j\,, \qquad i, j = 1, \ldots, N_s, \tag{7.10}$$

where typically $N_s < N_\Gamma$ is chosen in order to remove noise-dominated directions. The projected correlation matrix can equivalently be written as the correlation matrix of a new, reduced operator basis[19, 143],

$$\widetilde{C}^{\Gamma}_{ij}(t) \;=\; \left\langle\, Q^{\Gamma}_i(t)\, \overline{Q}^{\Gamma}_j(0) \,\right\rangle , \tag{7.11}$$

where the pruned operators are defined as linear combinations

$$Q^{\Gamma}_i(t) \;=\; \sum_{\alpha=1}^{N_\Gamma} \left(u^{(\alpha)}_i\right)^* \, O^{\Gamma}_\alpha(t)\,. \tag{7.12}$$

Due to the orthogonality of the singular vectors, the coefficients of these operators are orthogonal, and the removal of modes associated with small singular values improves the conditioning of the matrix $\widetilde{C}^{\Gamma}(t)$, thereby stabilizing the subsequent GEVP analysis. The pruned GEVP is then solved analogously to the unpruned formulation; in practical implementations, the reduction to an ordinary eigenvalue problem is often achieved via a Cholesky decomposition [144], for which $\widetilde{C}^{\Gamma}(t_0)$ should be Hermitian and positive definite up to statistical fluctuations. In rare cases different time-reversal symmetries may appear in off-diagonal matrix elements, which can change the usual cosh behaviour of the generalised eigenvalues [145].

7.2 Ensembles and simulation details

Following the analysis in [19] this work is primarily performed on the ensemble `Em1`, which is an $N_f = 2$ lattice generated with clover-improved Wilson fermions and the Wilson plaquette gauge action. This ensemble serves as a controlled setting for spectroscopy studies, with a charm quark mass tuned to approximately half its physical value, allowing for a detailed investigation of operator construction and optimisation without the additional complications of a fully physical setup. It consists of 5 replica comprising of nearly 30,000 configurations with periodic boundary conditions in the spatial directions and anti-periodic boundary conditions for the fermion fields in time. The bare parameters of the ensemble are given by $\beta = 6/g_0^2 = 5.3$ and hopping parameter $\kappa = 0.13270$, and a lattice size of 48×24^3. These parameters yield a lattice spacing of $a \approx 0.0658\,\text{fm}$[146, 147], placing the ensemble at a moderately coarse lattice resolution. The absence of light quarks simplifies the spectrum by preventing decays into states with light hadrons and restricting possible mixing with gluonic operators to the iso-scalar charmonium sector. Furthermore, the relatively large number of configurations allows for high-statistics measurements of gluonic observables, which is essential in view of the significant signal-to-noise challenges associated with glueball correlators.

For comparison, we also consider the $N_f = 3 + 1$ ensemble `A1` generated with a non-perturbatively $\mathcal{O}(a)$-improved Wilson fermion action, Lüscher-Weisz gauge action, and open boundary conditions in the temporal direction [148]. The ensemble has a lattice size of 96×32^3 at a bare coupling $\beta = 3.24$, with hopping parameters $\kappa_l = 0.13440733$ for the degenerate light quarks and $\kappa_c = 0.12784$ for the charm quark with the pion mass at $m_\pi = 420\text{MeV}$. This corresponds to a lattice spacing of $a \approx 0.0536\,\text{fm}$ and a spatial extent of $L \approx 3.65\,\text{fm}$ [149]. In contrast to the Em1 ensemble, A1 includes dynamical light quarks and a charm quark close to its physical value, providing a more realistic setup for QCD. The larger physical volume and finer lattice spacing

allow for a qualitative assessment of finite-volume and discretisation effects in the glueball spectrum.

Ensemble	N_f	$T \times L^3$	β	a [fm]	L [fm]	N_{conf}
Em1	2	48×24^3	5.3	0.0658	1.58	29765
A1	3+1	96×32^3	3.24	0.0536	3.65	2503

TABLE 7.1: Parameters of the ensembles used in this work. For the A1 ensemble, every second configuration of a single replica (5006 configurations) is analysed, resulting in 2503 measurements.

7.3 Operator-basis and smearing optimisation

Before proceeding to the variational analysis, it is necessary to assess the impact of smearing and operator choice on the quality of the glueball correlators. In particular, Teper blocking constructs extended links of increasing spatial extent, such that successive blocking steps rapidly lead to very large operators. This can result in oversmearing, where different loop shapes become effectively degenerate and their overlap with specific symmetry channels is reduced. Given the large number of possible combinations arising from multiple levels of APE smearing and Teper blocking, together with the full set of 35 loop shapes, the resulting operator basis can quickly become prohibitively large. Since the computation of all corresponding correlators is computationally demanding, it is therefore advantageous to perform an exploratory study on a single replica of the ensemble. This allows us to identify the most effective smearing parameters and operator geometries, and to restrict the operator basis to a manageable subset for the subsequent analysis.

7.3.1 Smearing setup and analysis method

In order to study the overlap of the constructed operators with the low-lying glueball states, we investigate different combinations of APE smearing and Teper blocking. For APE smearing, introduced in Section 2.7.1, three levels are considered $N_{APE} = 5, 10, 20$, while for Teper blocking, introduced in Section 2.7.3, three iterations are applied. For both methods an alpha value of 0.5 was applied. Combining these procedures results in a total of 15 distinct smearing setups. For each smearing combination, the full set of 35 loop shapes is evaluated and projected onto the irreducible representations of A_1, E, and T_2. This allows for a systematic study of the dependence of the correlators on both the operator geometry and the smearing procedure. At this stage, we restrict ourselves to the analysis of individual correlation functions without applying the variational method. Effective masses are constructed from the corresponding two-point functions in order to assess the quality of the signal and the overlap with the ground state. This approach allows for a direct comparison of different loop shapes and smearing combinations before constructing an optimised operator basis for the subsequent variational analysis. This is performed on a single replica of 7115 gauge configurations.

7.3.2 Effective mass comparison across smearing levels and loop shapes

We first consider the A_1^{++} channel, which exhibits the cleanest signal, and study the dependence of the effective masses on both the smearing procedure and the choice of loop shape. Figure 7.1 shows the effective masses obtained from the plaquette operator for different smearing prescriptions. Points with the same time separation T/a have been displaced to ensure readability. This is the case for all following effective mass plots. Already at this level, indications of oversmearing become visible. In particular, the data corresponding to three iterations of Teper blocking (T3), shown in purple, exhibit increased fluctuations and reduced plateau stability when combined with APE smearing. This suggests that the operator becomes too extended to maintain a strong overlap with the low-lying states. To further investigate this behaviour, Figure 7.2 shows the effective masses for all loop shapes at fixed smearing levels. For moderate smearing (T2, top panel), a large number of shapes yield consistent and stable signals. In contrast, for stronger Teper blocking (T3, middle panel), the quality of the signal deteriorates significantly: many operators exhibit large statistical uncertainties or fail to produce a meaningful plateau altogether, while only a small subset retains a usable signal. A different effect is observed for APE smearing (APE20, bottom panel), where the effective masses from different shapes become largely degenerate within uncertainties. This indicates that the smearing procedure washes out the geometric differences between loop shapes, reducing the diversity of the operator basis. This is also the case, but less so, for fewer APE iterations.

Taken together, these observations point towards the onset of oversmearing at the level of three Teper iterations, particularly when combined with additional APE smearing. This can be understood geometrically: each Teper iteration effectively doubles the spatial extent of the links, such that after three iterations the operators probe distances of order eight lattice spacings. For loop shapes larger than the plaquette, this implies that the operators extend over a substantial fraction of the lattice volume, which can lead to a loss of sensitivity to the underlying structure of the states. Based on these results, it is therefore advantageous to restrict the smearing levels used in the subsequent analysis to just 2 iterations of Teper blocking in order to retain a well-conditioned and sufficiently diverse operator basis.

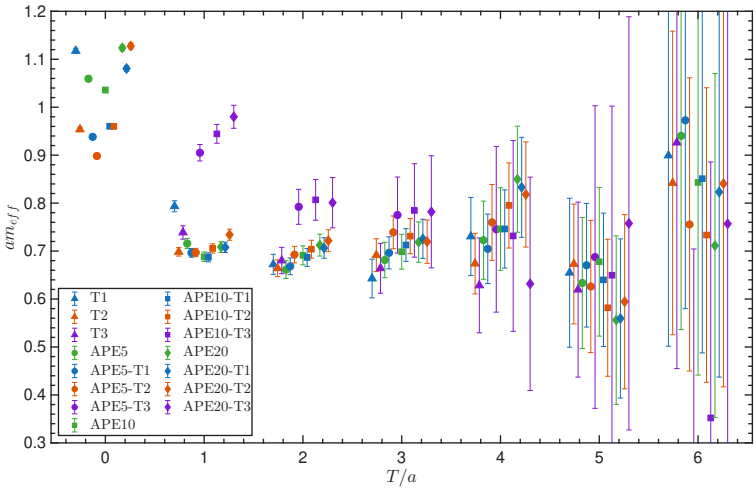

FIGURE 7.1: Effective masses in the A_1^{++} channel for the plaquette operator using different smearing combinations. Results are shown for pure Teper blocking (T1–T3), pure APE smearing (APE5–APE20), and combinations thereof. While moderate smearing levels lead to consistent plateaus, the strongest smearing (in particular T3 combined with APE smearing) exhibits increased fluctuations and indications of degraded signal quality, suggestive of oversmearing effects.

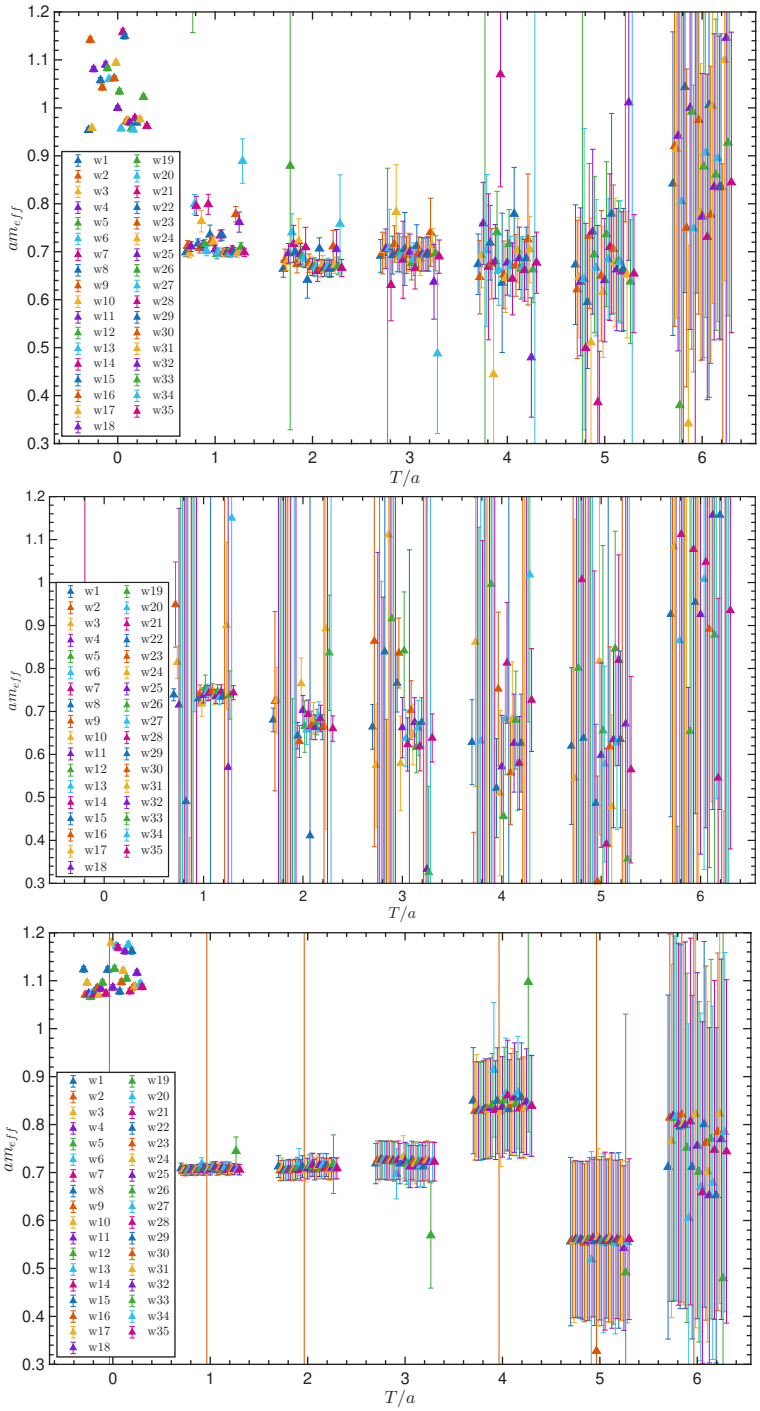

FIGURE 7.2: Effective masses in the A_1^{++} channel for all 35 loop shapes under different smearing prescriptions: top: two iterations of Teper blocking (T2); middle: three iterations of Teper blocking (T3); bottom: APE smearing with 20 iterations (APE20). While moderate smearing (T2) yields consistent signals across many shapes, stronger smearing (T3) leads to a loss of signal for a large fraction of operators. In contrast, APE20 results in a high degree of degeneracy among shapes, indicating reduced sensitivity to operator geometry.

Following the investigation of different smearing prescriptions, we proceed to further restrict the basis by selecting suitable loop shapes. If one were to consider only shapes that have overlap with all irreducible representations of interest, a significant fraction of the operator basis could already be discarded. However, for the ensemble `Em1` the A_1^{++} channel is the only one exhibiting a clearly identifiable signal [19], and all loop shapes can in principle access this channel. It is therefore natural to retain the plaquette as the simplest operator in the basis. To further guide the selection of shapes, Figure 7.3 shows the effective masses in the T_2^{++} channel for all loop geometries using two representative smearing prescriptions. The behaviour observed in this channel is consistent with the findings of [131], where it was shown that not all loop shapes have non-vanishing overlap with a given irreducible representation.

By analysing Figure 7.3 together with the corresponding results in the other irreps, a subset of loop shapes is selected for the subsequent analysis. The final selection of loop shapes is guided by a balance between maximising the overlap with the low-lying states and maintaining a well-conditioned operator basis. In particular, the chosen shapes are those that consistently exhibit a clear signal across the irreps of interest, show early onset of plateaus in the effective masses indicating good ground-state overlap, and possess moderate statistical uncertainties. At the same time, care is taken to avoid redundancies within the operator basis. Many loop geometries become effectively equivalent after smearing, leading to highly correlated signals and little additional information content. The selected subset therefore excludes shapes that are strongly degenerate with others, ensuring that each operator contributes complementary information to the variational basis. Geometrically, the retained shapes probe a variety of spatial structures and operator extents, ensuring sensitivity to different components of the glueball states. In practice, the selection is not constrained to short-distance operators, and several extended loop geometries are retained where they provide a stable signal and non-redundant information. This is particularly relevant in view of the oversmearing effects observed previously, where excessively extended operators were found to lose discriminating power between different shapes. Overall, the resulting set of eight loop shapes provides a compact yet sufficiently diverse operator basis, which is well-suited for a stable and efficient variational analysis of the glueball spectrum. A visualization of the selected shapes is shown in Figure B.1.

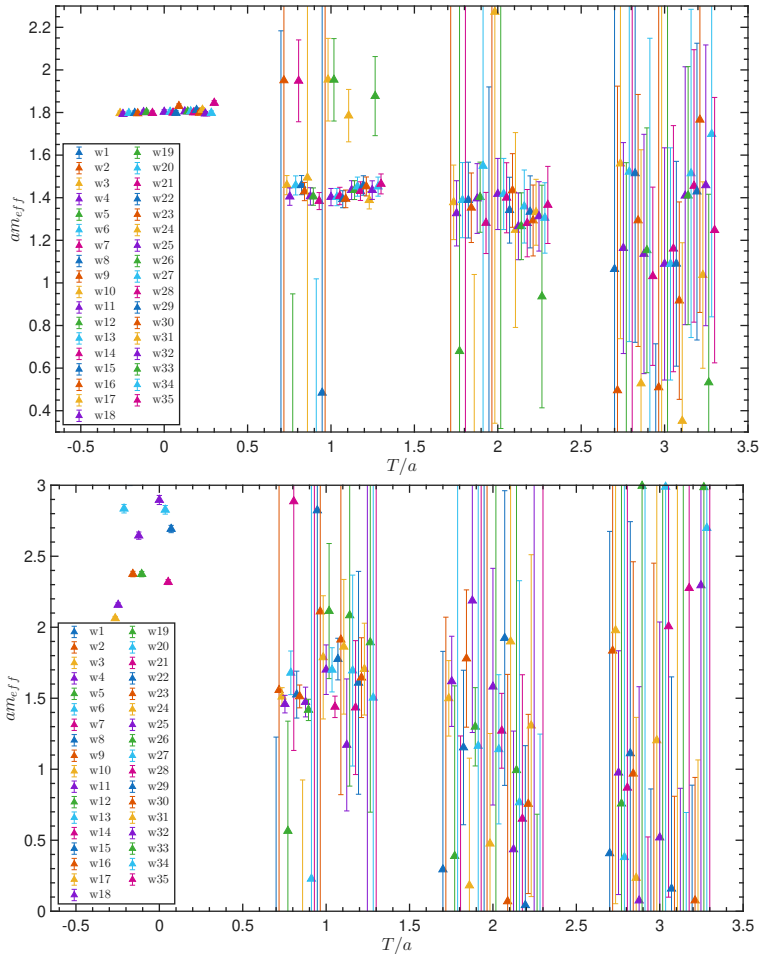

FIGURE 7.3: Effective masses in the T_2^{++} channel for all 35 loop shapes under different smearing prescriptions. Top: APE smearing with 20 iterations (APE20). Bottom: two iterations of Teper blocking (T2). Each point corresponds to a different loop geometry projected onto the T_2^{++} irrep.

7.4 Glueball spectrum from Teper-smeared operators on `Em1`

In this section, we present the results for the glueball spectrum obtained from the variational analysis using Teper-smeared operators. Building on the optimisation study of the previous section, the smearing procedures have been applied to the remaining replicas using the reduced operator basis. The analysis is performed using correlation matrices constructed from the selected set of 8 loop shapes (depicted on Figure B.1) and smearing levels (where we have thrown out the third Teper iteration due to oversmearing). Following the pruning procedure introduced in Equation (7.12), the operator basis is typically reduced to a 3×3 block, ensuring a well-conditioned correlation matrix while retaining the relevant physical information. An examination of several values of pruning was performed where it showed no large effect between the different values. This will become a more prominent when inspecting `A1`. The energy levels are extracted from the generalised eigenvalue problem, and we focus on the corresponding effective masses, as discussed in Section 4.4. After exploratory studies, the reference time is chosen as $t_0 = 0$, similar to [21]. Increasing t_0 was

found not to provide a significant improvement, while using $t_0 = 0$ allows access to earlier Euclidean times. This is particularly useful for investigating the suppression of excited-state contributions and for comparing the behaviour of different operators at small temporal separations.

For each irreducible representation, the analysis is performed separately for different smearing prescriptions. In a first step, pruning and the subsequent GEVP are applied to operator bases constructed from a single smearing level, including all loop shapes accessible to the given irrep. This allows for a direct comparison of the impact of the different smearing procedures on the extracted effective masses. In the upcoming sections, the notation $T1$ and $T2$ denotes one and two iterations of Teper blocking, respectively, while $APE\#$ refers to the number of APE smearing iterations applied. In addition to the single-smearing analysis, we also extend the operator bases combining multiple smearing levels. In the figures, these are labelled as $APEs$, $T1s$, and $T2s$, corresponding to operator bases that include all shapes at fixed Teper levels (0, 1, and 2) combined with the 3 different APE smearings. Selected combinations of smearing procedures are investigated in order to construct compact operator bases with complementary properties. The combination labelled "Combination" consists of the operators obtained from the smearing levels $T1$, $APE5$-$T1$, $APE10$, and $APE20$-$T2$. For the A_1^{++} channel, where the plaquette is included in the operator basis, the enlarged basis involving all shapes and smearing combinations can become particularly large. Since each smearing combination contains 8 shapes, the largest GEVP basis are a combination of 32 operators. Here there is an even higher importance of the pruning procedure in order to obtain a manageable and well-conditioned correlation matrix. For this higher operator basis we investigated several quantities of pruning to ensure that we retain enough information to have a good overlap, while keeping the problem size low to reduce noise. This approach enables a systematic assessment of how the smearing procedures influence both the overlap with the ground state and the stability of the variational analysis.

7.4.1 A_1^{++} channel

We begin by examining the effective masses obtained from the GEVP for the different smearing prescriptions, shown in Figure 7.4. At first glance, the results exhibit a high degree of consistency across the various operator bases. In particular, no significant improvement in the statistical uncertainties is observed when comparing different smearing setups. Furthermore, for the single-smearing cases, no noticeably earlier onset of a plateau can be identified. A closer inspection of the early-time region, however, reveals more subtle differences. In particular, the four combined smearing setups, i.e. operator bases constructed from multiple smearing prescriptions, show a tendency to reach the plateau slightly earlier than the single-smearing cases. At the same time, stronger smearing levels appear to yield a more efficient suppression of excited-state contributions, in line with expectations, as smearing improves the overlap of the operators with the ground state [90]. In particular, at $t/a = 0$ the strongest suppression of excited-state contributions is observed for setups including Teper blocking. To quantify these observations, we consider the $T_{\min}$ dependence of the extracted masses, shown in Figure 7.5. The combined smearing setups exhibit a consistent behaviour, with plateaus that can be identified already at $t/a \approx 2$, compared to $t/a \approx 3$ for the single-smearing cases. No significant difference is observed between combinations including or excluding Teper blocking at this stage, but it is worth noting that the pure Teper smearing cases (T1 and T2), as well as the T2 setup at $t/a = 2$,

show slightly stronger excited-state suppression compared to the other configurations. Overall, however, the differences between the various smearing prescriptions remain modest.

The extracted plateau values are summarized in Table 7.2. For the combined smearing setups, plateau values are taken starting from $t/a = 2$, while for the remaining cases the fit range begins at $t/a = 3$.

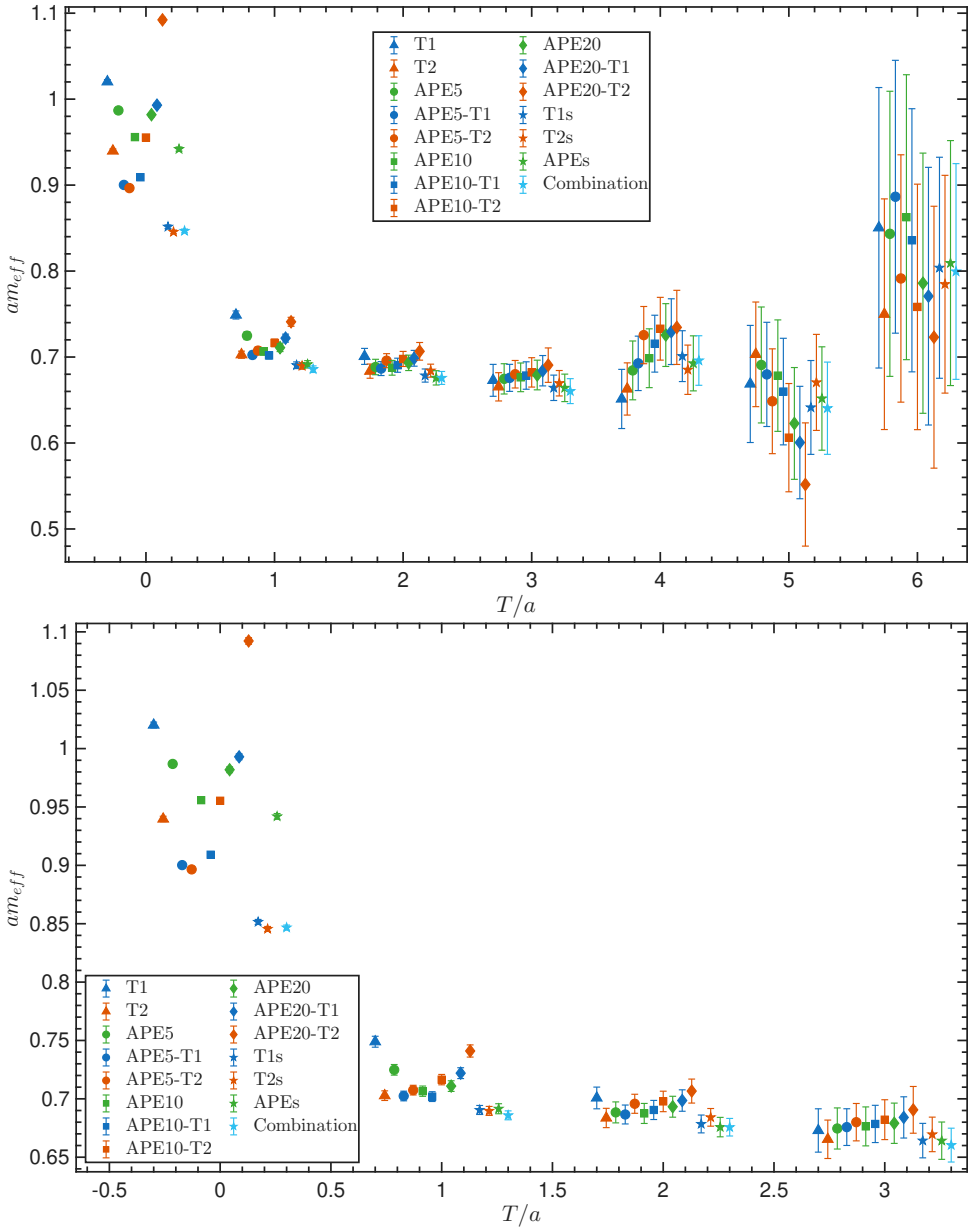

FIGURE 7.4: Effective masses in the A_1^{++} channel obtained from the GEVP using the selected set of eight loop shapes for different smearing prescriptions at $t_g = 0$. Bottom panel shows a zoom of the upper figure focusing on the early times T/a.

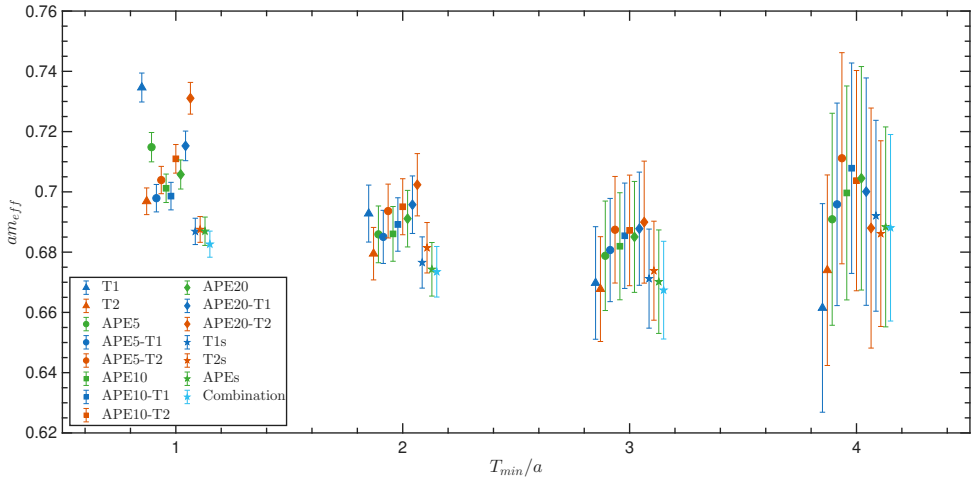

FIGURE 7.5: $T_{\min}$ dependence of the extracted masses for all 15 GEVP setups shown in Figure 7.4.

	T1	T2	APE5	APE5-T1	APE5-T2	APE10	APE10-T1	APE10-T2
am_{eff}	0.670 ± 19	0.668 ± 17	0.679 ± 18	0.681 ± 17	0.687 ± 18	0.682 ± 18	0.685 ± 17	0.687 ± 18
	APE20	APE20-T1	APE20-T2	T1s	T2s	APEs	Comb.	
am_{eff}	0.685 ± 18	0.688 ± 19	0.690 ± 20	0.677 ± 9	0.682 ± 8	0.674 ± 9	0.674 ± 8	

TABLE 7.2: Plateau values of the effective mass for different smearing combinations. Errors are given in units of 10^{-3}. First 12 are in the range of T/a=3-6 and the last 4 in the range of T/a=2-6

First excited state in the A_1^{++} channel

We now turn to the first excited state in the A_1^{++} channel. The corresponding effective masses obtained from the GEVP are shown in Figure 7.6. In contrast to the ground-state analysis, more pronounced differences between the smearing prescriptions become visible. In particular, the single-smearing setups without Teper blocking (i.e. the pure APE smearings) exhibit a noticeably poorer performance. These cases show larger statistical uncertainties and indications of stronger excited-state contamination at early Euclidean times compared to the other smearing prescriptions. A possible explanation for this behaviour is an increased degeneracy within the operator basis when only a single smearing level is used. In this case, the operators probe very similar structures, reducing the effectiveness of the variational method in isolating excited states. This interpretation is supported by the observation that combining different APE smearings improves the signal, bringing it closer to the performance of the other setups, although still with somewhat larger uncertainties than the combined smearing case. While there are indications that some of the smearing combinations may reach a plateau at earlier times, the available data does not provide a sufficient number of points to reliably identify or fit a plateau region for the excited state. This is further illustrated in Figure 7.7, where the $T_{\min}$ dependence of the extracted masses is shown for the selected smearing combinations. The results exhibit sizeable fluctuations and do not allow for a stable determination of the excited-state mass within the current statistics. Overall, the extraction of the first excited state in the A_1^{++} channel remains challenging, and no definitive plateau can be established from the present data.

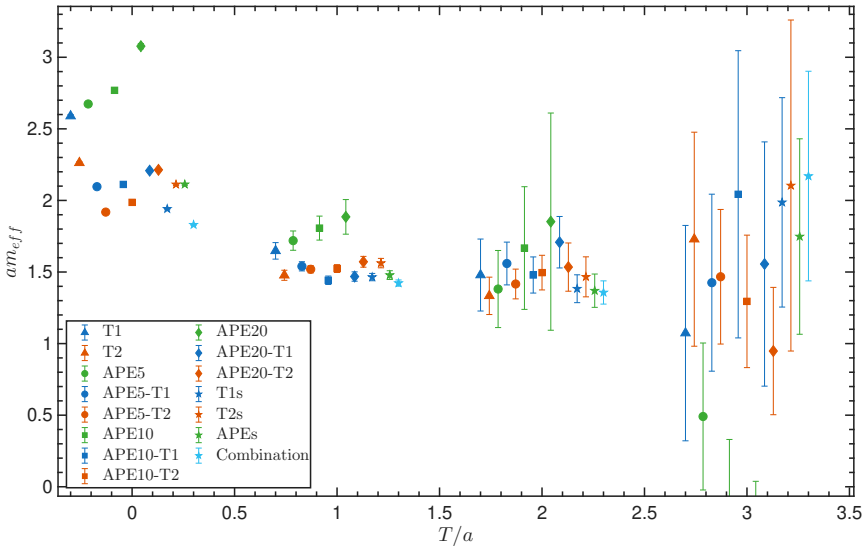

FIGURE 7.6: Effective masses in the A_1^{++} channel of the first excited state obtained from the GEVP

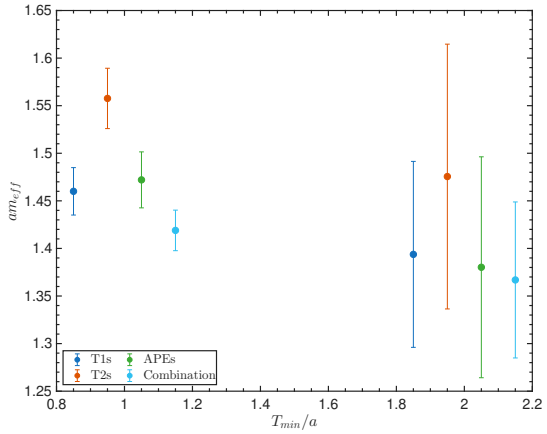

FIGURE 7.7: T_{min} dependence for the first excited state of the combined smearing setups

7.4.2 A_1^{-+}, E^{++} and T_2^{++} channels

We now turn to the remaining channels A_1^{-+}, E^{++} and T_2^{++}, for which the effective masses are shown in Figures 7.8 to 7.10. In contrast to the A_1^{++} channel, the signal quality in these channels is significantly reduced. While a consistent behaviour at early Euclidean times can still be observed, the signal rapidly deteriorates as t/a increases, and statistical uncertainties become dominant. As a consequence, only a limited number of time slices provide usable data for the analysis. Within this restricted range, no clear and stable plateaus can be identified for any of the smearing prescriptions. Although the different operator bases yield broadly consistent results at early times, the available statistics are insufficient to reliably extract mass values or to distinguish between the performance of the various smearing setups. Overall, the extraction of the glueball spectrum in these channels remains inconclusive with the present dataset. This highlights the significantly larger signal-to-noise problem in other channels compared to the A_1^{++} ground state.

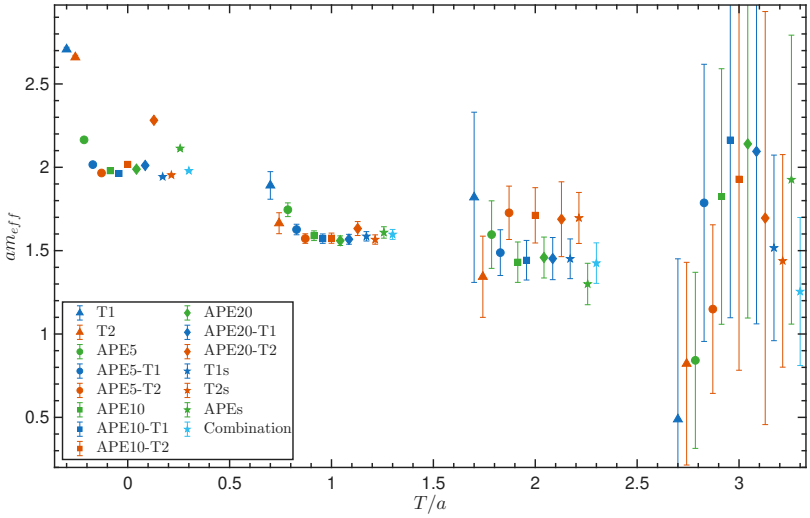

FIGURE 7.8: Effective masses in the A_1^{-+} channel

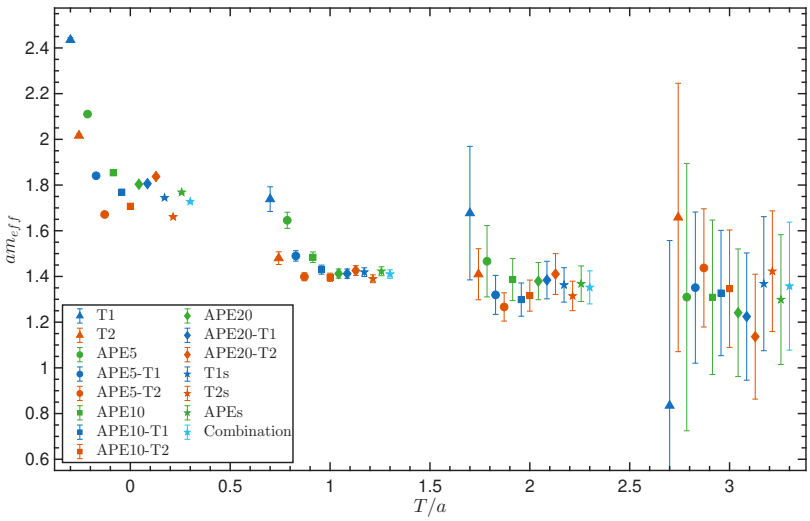

FIGURE 7.9: Effective masses in the E^{++} channel

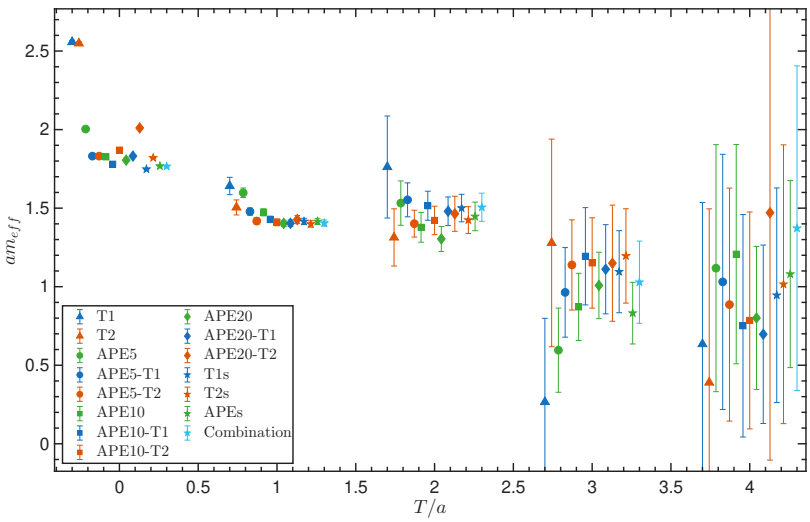

FIGURE 7.10: Effective masses in the T_2^{++} channel

7.5 Results for ensemble A1

We now turn to the analysis of ensemble A1. In contrast to the previous study on Em1, the present analysis is restricted to a single replica. Although additional replicas are available, the computational cost associated with the smearing procedures motivates an initial study based on a representative subset of the data. Consequently, the statistical precision is reduced, and the following discussion is kept primarily qualitative. We begin with the pruning procedure. For the single-smearing bases consisting of the selected eight loop shapes, a reduction to three operators is found to be stable. The main challenge arises for the enlarged combined bases containing 24 or 32 operators, which therefore form the focus of the following discussion. Since the Euclidean correlation matrix is Hermitian and positive definite (up to statistical fluctuations), its singular-value decomposition provides a useful tool for identifying the dominant operator directions. This motivates pruning based on the leading singular vectors. Table 7.3 lists the first 16 singular values of the 24×24 operator basis corresponding to the APE-only combination (APEs). While a pronounced hierarchy of singular values would favour an aggressive truncation, the observed decrease is comparatively gradual. This behaviour may indicate a stronger mixing of operator directions and is consistent with the denser low-lying spectrum expected for the $N_f = 3 + 1$ ensemble A1 [23, 149].

index	Singular value
1	19.986(89)
2	3.235(13)
3	0.3911(14)
4	0.3377(13)
5	0.03621(14)
6	0.005804(22)
7	0.003043(12)
8	0.0023803(91)
9	0.0012076(46)
10	0.0005318(20)
11	0.0004201(16)
12	0.0002775(12)
13	0.0002698(10)
14	0.00013049(63)
15	0.00004691(32)
16	0.00003235(24)

TABLE 7.3: Singular values for the 'APEs' 24by24 operator basis.

The choice of pruning level is therefore particularly relevant. If too few modes are retained, physically relevant information may be lost; if too many are kept, the operator basis remains highly correlated and the resulting GEVP can become noisy or numerically unstable. Figure 7.11 compares different pruning levels between $N_s = 5$ and $N_s = 12$ for the four combined smearing setups. In addition, we tested an alternative two-step procedure in which each single-smearing basis is first reduced from eight to three operators before combining the resulting pruned operators into the enlarged basis. This setup is labelled pr3x in the figure. At first glance, pruning to $N_s = 5$ appears too aggressive for some combinations, in particular for T1s, where

the resulting effective masses are shifted upward relative to the corresponding single-smearing inputs shown in Figure 7.12. Likewise, the two-step pr3x procedure exhibits signs of instability, with the APEs combination already deteriorating at $T/a = 2$. Based on these observations, the subsequent analysis is performed using $N_s = 7$ for the APEs basis and $N_s = 12$ for the remaining three combined smearing setups.

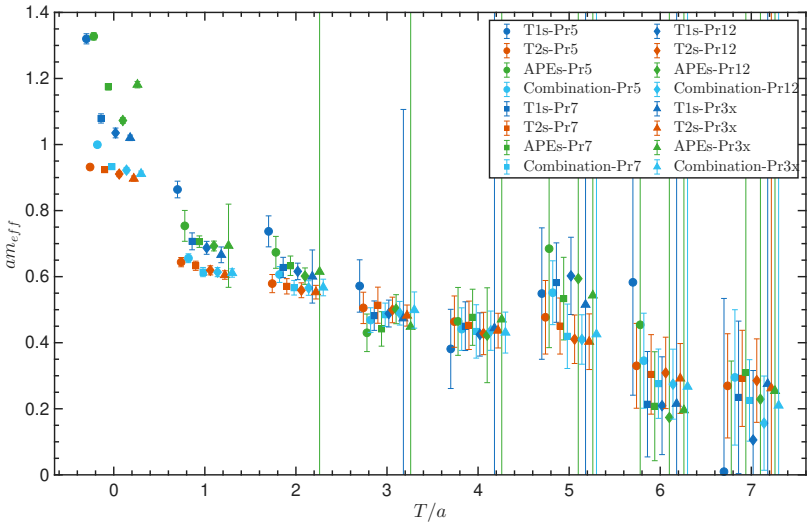

FIGURE 7.11: Effective masses in the A_1^{++} channel for the 4 cases of combined smearings with different amounts of pruning.

The effective masses for the different irreducible representations are shown in Figures 7.12 to 7.15. Overall, the behaviour is similar across channels: while a consistent signal can be observed at early Euclidean times, the signal-to-noise ratio deteriorates rapidly, limiting the range in which meaningful comparisons can be made.

For the ground state in the A_1^{++} channel (Figure 7.12), the combined smearing basis exhibits slightly smaller statistical uncertainties compared to the purely APE-smeared operators. While Teper blocking, in particular the T2 setup, shows indications of stronger suppression of excited-state contributions at early times, this does not translate into a clearly earlier onset of a plateau. A similar trend can be observed for the first excited state in the A_1^{++} channel (Figure 7.13), where the T2 smearing appears to yield a more stable behaviour compared to the other setups. However, the statistical fluctuations remain large, and no reliable plateau region can be identified. It is shown in [21] that you need charmonium-like operators to resolve the state properly. For the E^{++} and T_2^{++} channels (Figures 7.14 and 7.15), the differences between the smearing prescriptions are again modest. The combined smearing basis tends to show slightly reduced uncertainties, and there are weak indications of a faster approach to a plateau compared to the single-smearing cases. Nevertheless, the limited statistics prevent a conclusive assessment of these effects.

In summary, while some qualitative trends are consistent with the observations made for the Em1 ensemble, Teper blocking seems to produce a slighty improved behaviour for combined smearing bases and indications of stronger excited-state suppression at A1, but the reduced statistical precision of the dataset does not allow for a definitive comparison. Further analysis with increased statistics would be required to confirm these tendencies.

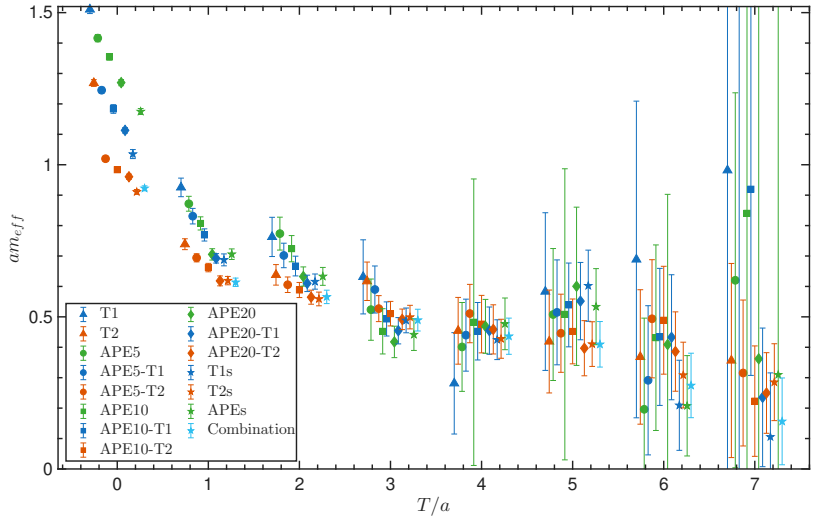

FIGURE 7.12: Effective masses in the A_1^{++} channel

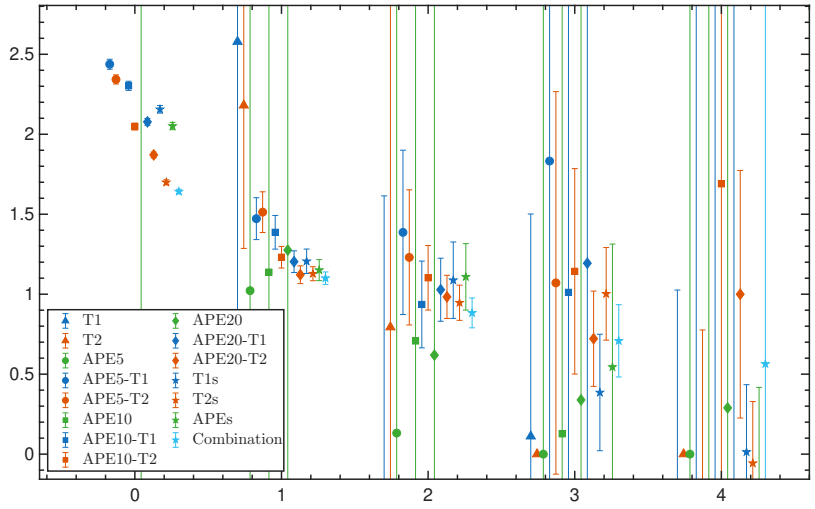

FIGURE 7.13: Effective masses in the A_1^{++} channel 1st excited

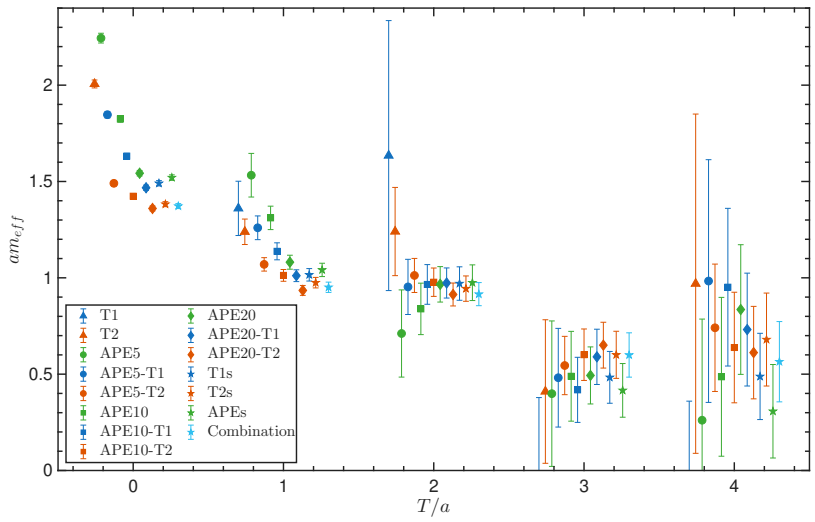

FIGURE 7.14: Effective masses in the E^{++} channel

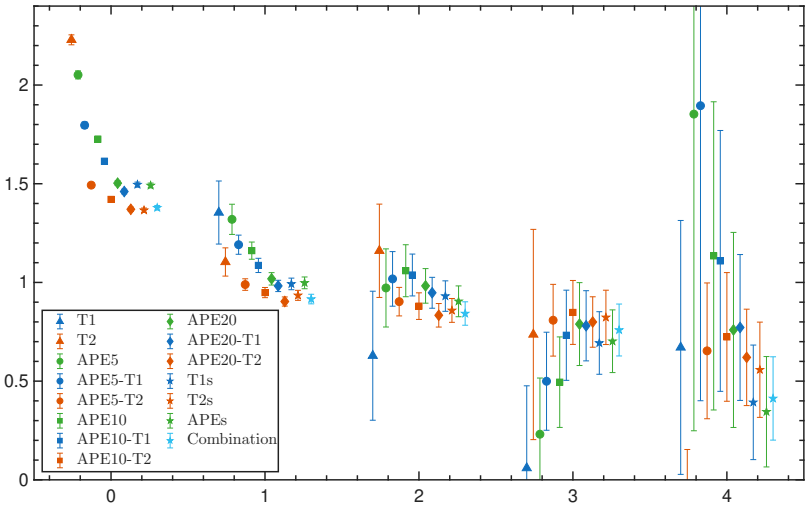

FIGURE 7.15: Effective masses in the T_2^{++} channel

7.6 Final glueball spectrum results

In this section, we summarise the results of the variational analysis using the optimised operator basis and the combined smearing setup identified in the previous sections. The final results for the different irreducible representations are shown in Figures 7.16 and 7.17 for the ensembles `Em1` and `A1`, respectively.

For the ensemble `Em1` (Figure 7.16), a stable signal is observed in the A_1^{++} channel, where a plateau can be identified over a limited range in Euclidean time. The corresponding ground-state result is consistent with the behaviour observed throughout the optimisation analysis and with the most recent glueball study using the same ensemble [21]. Moreover, the extracted A_1^{++} ground-state signal lies below the estimated two-pion threshold shown in the figure. This suggests that the state is unlikely to be dominated by the lowest two-hadron contribution and is consistent with an interpretation as a predominantly glueball-like scalar state. For the first excited state in the same channel, as well as for the E^{++} and T_2^{++} channels, the statistical uncertainties increase rapidly, and the accessible plateau regions are significantly reduced. As a result, no precise determination of these states can be achieved within the present setup.

Turning to the ensemble `A1` (Figure 7.17), the analysis is performed on a restricted subset of configurations, and the results are therefore interpreted with appropriate caution. In contrast to the `Em1` ensemble, the signal extends to slightly larger Euclidean times before the onset of noise, allowing for a marginally broader window in which effective masses can be inspected. While the statistical uncertainties remain sizeable, qualitative features similar to those observed in `Em1` are present. In particular, there are mild indications that the inclusion of Teper-blocked operators leads to an improved suppression of excited-state contributions at early times. This behaviour is most visible when comparing the different smearing setups within the combined basis. From a physical perspective, the interpretation of the low-lying states on `A1` is considerably more involved than for `Em1`. Owing to the lighter dynamical quark content and the denser low-energy spectrum, scalar mesons, multi-hadron states, and glueball-like excitations are expected to mix strongly [23, 149]. Consequently, states extracted from the present purely gluonic operator basis cannot be unambiguously identified, but rather indicate the overlap of the Teper-improved operators with the low-lying scalar sector.

However, given the limited statistical precision, these observations remain tentative. A more definitive assessment of the impact of Teper blocking would require an analysis based on a larger set of configurations and an enlarged variational basis including additional interpolating operators. An extended analysis of this kind would constitute a natural continuation of the present work.

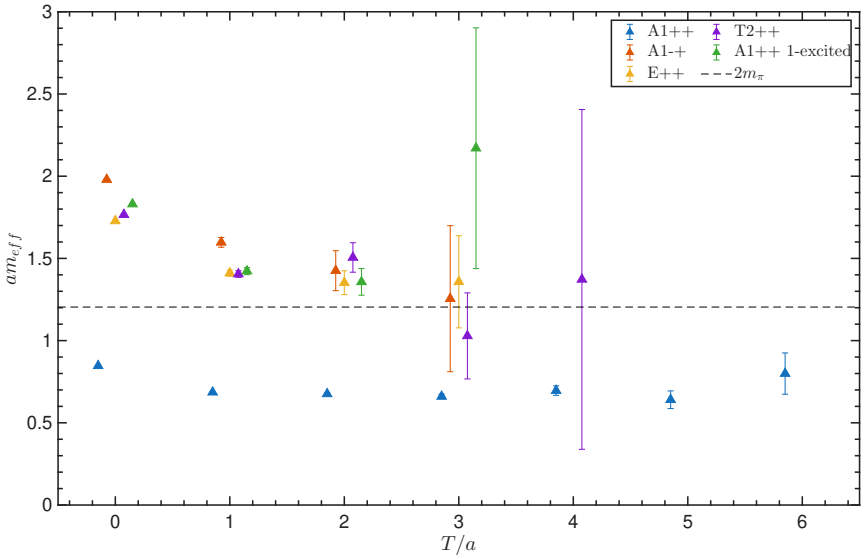

FIGURE 7.16: Effective masses on ensemble `Em1` for the different irreducible representations obtained from the GEVP using the combined smearing basis consisting of $T1$, $APE5$-$T1$, $APE10$, and $APE20$-$T2$. Shown are the ground-state results for the $A_1^{\pm+}$, T_2^{++} and E^{++} channels, as well as the first excited state in the A_1^{++} channel.

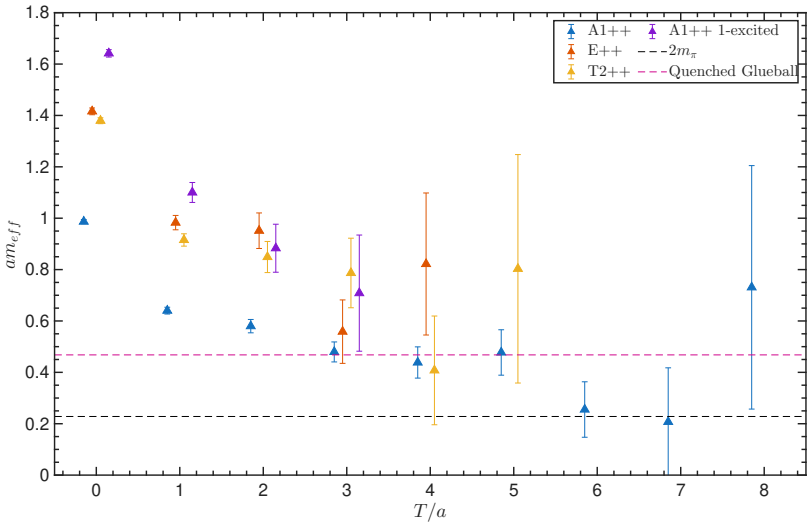

FIGURE 7.17: Effective masses on ensemble `A1` for the different irreducible representations obtained from the GEVP using the combined smearing basis consisting of $T1$, $APE5$-$T1$, $APE10$, and $APE20$-$T2$. Shown are the ground-state results for the $A_1^{\pm+}$, T_2^{++} and E^{++} channels, as well as the first excited state in the A_1^{++} channel.

Part 6

Conclusions and the Road Ahead

Chapter 8

Conclusions and outlook

This thesis pursued two largely independent lines of investigation. Chapters 5 and 6 focused on precision scale setting from the static quark–antiquark potential in $N_\mathrm{f} = 2 + 1$ lattice QCD, culminating in a controlled determination of the potential scales r_i. Chapter 7, in contrast, investigated Teper-style spatial link blocking as a technical ingredient for glueball spectroscopy. In the following, these two parts are summarised and assessed separately, before outlining concrete directions for future work.

8.1 Potential-based scale setting in $N_\mathrm{f} = 2 + 1$ QCD

A central outcome of Chapters 5 and 6 is a high-precision determination of potential-derived reference scales. Starting from Wilson-loop measurements and a variational extraction of the ground-state static energy, the static force was computed using tree-level improved distance definitions, and the resulting ensemble-level values were combined in global continuum and chiral analyses. At the physical point, the final values for the potential scales are

$$r_0 = 0.4729(57)(48)\mathrm{fm}, \tag{8.1}$$

$$r_1 = 0.3127(24)(32)\mathrm{fm}, \tag{8.2}$$

$$\frac{r_0}{r_1} = 1.532(12). \tag{8.3}$$

The first uncertainty in r_0 and r_1 combines the statistical and systematic uncertainties of the lattice determination, while the second reflects the uncertainty of the reference scale $\sqrt{t_0}$ used to convert the results to physical units. The quoted uncertainties reflect a careful separation of statistical and systematic contributions, and the results are compatible with the corresponding $N_\mathrm{f} = 2 + 1$ averages reported by FLAG [74] and with other recent determinations within uncertainties. At the same time, the spread among published central values for r_0 and r_1 appears somewhat larger than would be expected from the quoted uncertainties alone. A simple consistency test based on the available determinations yields p-values of approximately 4×10^{-3} for r_0 and 1×10^{-4} for r_1, indicating that the observed spread is unlikely to arise solely from statistical fluctuations. This may point to residual systematic effects that are treated differently across analyses. By contrast, the determinations of the ratio r_0/r_1 exhibit substantially better consistency, with a corresponding p-value of approximately 0.11. In the present work, particular emphasis is placed on a conservative treatment of continuum-extrapolation, fit-model, and autocorrelation systematics.

A lesson emerging from the analysis is that the shorter-distance scale r_1 is inherently more delicate. In particular, plateau identification and stability were observed to

become more challenging for potentials around $r \simeq r_1$ on a small subset of ensembles, with the coarse lattices and `E300` case providing a concrete example of a more problematic extraction. While the final continuum averages remain stable under the systematic variations considered, these observations motivate a conservative recommendation for future scale-setting work that uses potential-derived reference scales as primary inputs: when choosing between r_0 and r_1 as a reference scale for cross-ensemble comparisons or as an intermediate scale, one should preferentially use r_0, unless a dedicated, statistics- and systematics-driven study demonstrates that r_1 can be controlled to the same robustness level in the specific setup at hand.

8.2 Additional physical results and their interpretation

Chapter 6 presented two applications which, while not the primary goal of this thesis, illustrate the physics reach enabled by a precise potential determination.

First, the $N_\mathrm{f} = 2 + 1$ Λ-parameter in the $\overline{\mathrm{MS}}$ scheme was expressed in units of r_0. Because the calculations involved the same lattice action and matched bare couplings, the analysis could be formulated in terms of a direct continuum extrapolation of r_0/L_had, eliminating the need to propagate an additional flow-scale determination through a second extrapolation step. The final result,

$$r_0\Lambda^{(3)}_{\overline{\mathrm{MS}}} = 0.820(28), \tag{8.4}$$

agrees well with the corresponding FLAG [74] average. The dimensionless combination $r_0\Lambda_{\overline{\mathrm{MS}}}$ has historically played an important role in relating low-energy hadronic scales to the perturbative QCD scale and remains a useful benchmark quantity. In the context of this thesis, the result should primarily be viewed as a non-trivial cross-check: it confirms that the extracted potential scale r_0 is consistent with an independently determined infrared scale and its perturbative matching to the $\overline{\mathrm{MS}}$ scheme. The agreement with dedicated determinations therefore provides additional support for the reliability of the scale-setting analysis presented in this work. Determinations of $r_0\Lambda_{\overline{\mathrm{MS}}}$ are available for several values of N_f, making the quantity a useful point of comparison across theories with different flavour content.

Second, the shape of the static potential was probed through the curvature observable $c(r) = \frac{1}{2}r^3V''(r)$, which is insensitive to the additive self-energy of the static sources and thus directly tests the functional form of $V(r)$. Within the accessible distance range (below the onset of string breaking in full QCD), the observed behaviour interpolates smoothly between a perturbative-like short-distance regime and a confinement-dominated intermediate regime. The measured curvature is clearly not constant over the accessible range of separations, indicating departures from a Cornell form and favouring descriptions based on a running effective coupling.

8.3 Teper-link smearing and glueball spectroscopy: status and limitations

Chapter 7 developed and tested a Teper-link blocking implementation for spatial links and studied its use in glueball operator construction as an extension of previous studies [19, 21–23, 149]. The analysis builds a basis of operators transforming in the irreducible representations of the cubic group, applying smearing/blocking to

improve overlaps onto low-lying states, and extracts energies through a variational analysis of a correlator matrix. The numerical results obtained in this chapter should be interpreted as exploratory. For the ensemble `Em1`, where the scalar glueball is the ground state, a stable signal is observed most clearly for the ground state in the A_1^{++} channel, while statistical uncertainties increase rapidly for higher lying channels. For ensemble `A1`, where the glueball can decay into light hadrons, the analysis was intentionally restricted to a reduced dataset (a single replica), which substantially limits statistical precision. Within this limitation, there are qualitative indications that Teper-blocked operators can contribute to a modest suppression of excited-state contamination at early Euclidean times inside combined operator bases. However, the effect is not established at a level that would justify claiming a clear overall advantage over more conventional choices, and the present results should therefore be seen as motivation for a deeper study rather than as a definitive conclusion.

8.4 Outlook

A natural extension of the present scale-setting analysis could be to repeat it on modern $N_\mathrm{f} = 2 + 1 + 1$ ensembles. While such a study would provide an important validation of the decoupling of the charm quark for low-energy reference scales, existing theoretical expectations and previous investigations suggest that charm-sea effects on r_0 and r_1 are very small. Consequently, significant shifts of the continuum-limit values are not expected at the current level of precision. A complementary direction would be a combined chiral and continuum extrapolation of the static potential itself. While the present work focused on the extraction of the potential scales, the underlying potential contains substantially more information on the dynamics of confinement. Extending the analysis of the force and curvature observables to the continuum limit would allow a quantitative study of the potential shape over a wide range of distances and provide more stringent tests of phenomenological descriptions of the static interaction.

For Chapter 7, the most immediate outlook is statistical: a substantially enlarged dataset for ensemble A1 (including additional replicas) is the key prerequisite for turning the qualitative trends into quantitative statements. With more statistics, one can pursue a systematic optimisation programme for Teper blocking parameters and combined bases, and compare against established glueball spectroscopy results obtained with more conventional operator constructions. At that stage it would also become meaningful to assess, channel by channel, whether Teper-style blocking offers practical benefits (e.g. earlier plateaus, improved overlap, reduced operator-basis size after pruning) relative to more standard operator constructions, or whether its role is best understood as complementary. Beyond improvements in statistics and operator construction, a natural next step would be to incorporate light-hadron operators into the variational basis (see [20, 23]). This would enable the study of glueball–meson mixing and ultimately allow the Teper-blocked operators to be assessed within a more complete spectroscopic analysis.

Appendix A

Behaviour of the E300 ensemble

Among the CLS ensembles analysed in this work some ensemble occasionally exhibits difficulties in the extraction of reliable plateaus for the static potential and in these cases most commonly around the physical distance $r \sim r_1$. Especially the ensemble E300 was troublesome, as it was close to the physical mass point at the second finest lattice one could consider it an important ensemble. Similar behaviour has been noted our earlier studies in [93] and is therefore documented here for completeness. The discussion below summarises the observations relevant for the analysis presented in this thesis. The primary difficulty appears when determining effective energies from Wilson loops at intermediate distances. For very small temporal separations T and distances r, the determination of the potential from the effective energies becomes unstable even before the signal-to-noise problem takes over. In particular, when r approaches r_1 on ensembles with fine lattice spacing, a noticeable dip in the effective mass can appear before the plateau region is reached. This behaviour could suggest that excited-state contamination is enhanced in this regime.

A common strategy to suppress excited-state contributions is to increase the amount of link smearing used to construct the operator basis entering the generalised eigenvalue problem (GEVP). In the standard setup used throughout this work, a basis of HYP-smeared operators with smearing levels between roughly 20 and 30 is employed, see section 4.3. These values were originally chosen as a compromise between suppressing ultraviolet fluctuations and maintaining a stable operator basis. To investigate whether the observed behaviour on the E300 ensemble could be related to an insufficiently rich operator basis, we performed a dedicated study in which the smearing basis was extended significantly. In addition to the standard smearing levels, operators with substantially larger HYP smearing were included, reaching smearing levels of 45 and 60. The motivation behind this extension was to enhance the overlap with the ground state and further suppress excited-state contamination in the Wilson-loop correlation matrix. The resulting analysis is illustrated in Fig. A.1. The left panel shows the effective masses obtained when the extended smearing basis is used in the GEVP analysis. As expected, increasing the smearing leads to slightly smoother correlators at early temporal separations. However, the overall structure of the effective mass remains essentially unchanged. More importantly, the right panel demonstrates that the inclusion of the additional smearing levels has no visible impact on the final GEVP solution. Within statistical precision, the extracted effective masses are consistent with those obtained from the original smearing basis.

Taken together, these results suggest that the behaviour of the E300 ensemble is not primarily driven by the choice of operator basis or smearing parameters. The persistence of the effect under substantial variations of the smearing procedure points instead to an ensemble-specific feature.

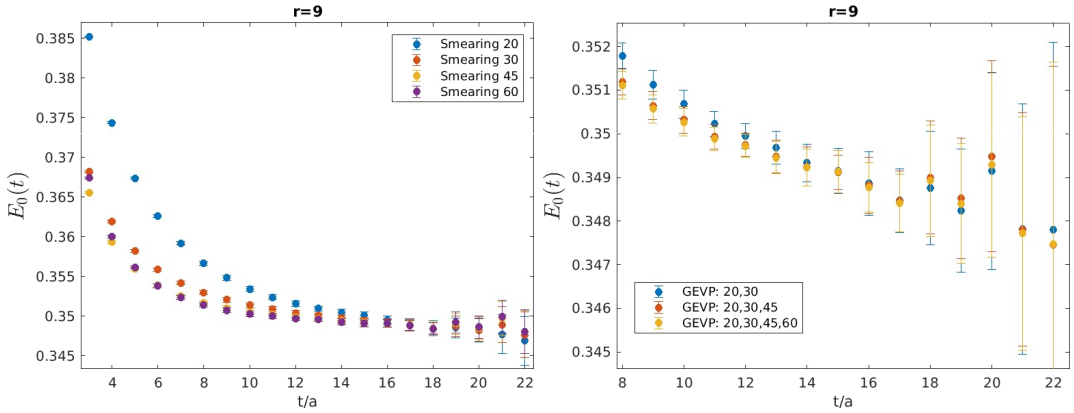

FIGURE A.1: Investigation of the effect of extended HYP smearing on the E300 ensemble. The smearing basis used in the GEVP analysis was enlarged from the standard range (20–30 smearing steps) to include significantly larger smearing levels up to 60. While the increased smearing slightly improves the behaviour of the correlators at small temporal separations (left panel), the resulting GEVP solution remains unchanged within statistical uncertainties (right panel). This indicates that the observed irregularities in the effective masses are not caused by an insufficient smearing basis.

Since enlarging the smearing basis does not resolve the irregular behaviour observed for the E300 ensemble, we also investigated alternative analysis methods for the extraction of the effective energies. In particular, we applied the *pencil-of-function* method [92, 150–153], which provides an alternative approach for extracting exponential components from correlation functions. The basic idea of the method is to enlarge the correlation matrix used in the analysis by incorporating shifted versions of the original data. In practice, starting from the original correlation matrix

$$W^{(k,l)}(r,T),$$

one constructs an extended matrix in which each element is replaced by a block containing correlations at neighbouring temporal separations,

$$W^{(k,l)}(r,T) \;\rightarrow\; \begin{pmatrix} W^{(k,l)}(r,T) & W^{(k,l)}(r,T+\tau) \\ W^{(k,l)}(r,T+\tau) & W^{(k,l)}(r,T+2\tau) \end{pmatrix}.$$

In this way, an original $N \times N$ correlation matrix is promoted to a $2N \times 2N$ matrix, thereby increasing the available information used to separate different exponential contributions. For the present analysis, the original 4×4 correlation matrix was extended to an 8×8 matrix following this procedure. The shift parameter τ was chosen to correspond to two lattice units in Euclidean time, $\tau = 2a$. After constructing the enlarged matrix, the analysis proceeds in analogy to the standard GEVP procedure, and effective energies can be extracted from the resulting eigenvalues. To improve the stability of the analysis, the extended matrix was subsequently *pruned* [142], as described in Equation (7.12). In practice this means that only a subset of the temporal range was retained in order to avoid regions dominated by noise or boundary effects. In the case of the E300 ensemble, the matrix was pruned down to a 6×6 matrix, which was found to provide a stable solution.

The resulting effective masses obtained from the pencil-of-function method are shown in Fig. A.2. For comparison, the standard GEVP results are also displayed. One

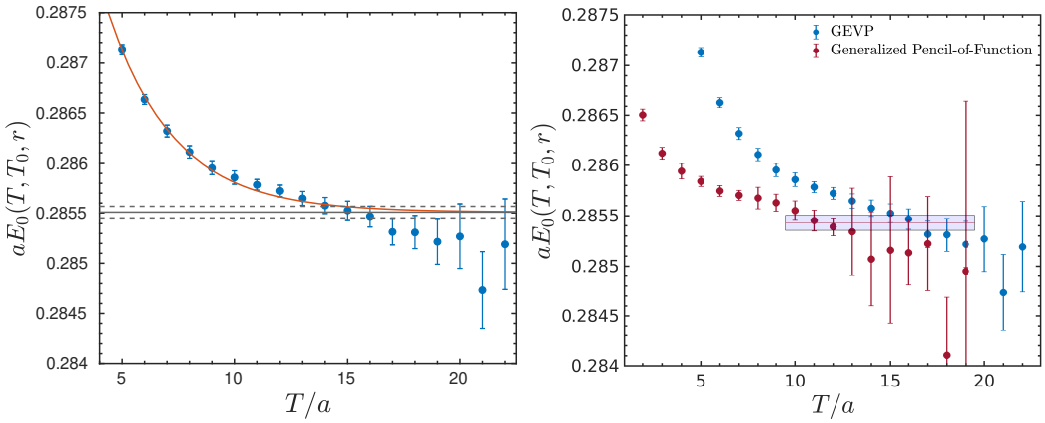

FIGURE A.2: Comparison of effective masses obtained from the standard GEVP analysis and from the pencil-of-function method on the E300 ensemble. The enlarged correlation matrix used in the pencil-of-function approach reduces the dip observed in the effective mass around $r \sim r_1$ and leads to a more stable plateau behaviour.

observes that the alternative analysis leads to a noticeable modification of the effective masses in the region where the original dip was observed.

In particular, the pencil-of-function analysis produces a smoother behaviour of the effective mass as a function of T/a, and the pronounced dip seen in the standard analysis is significantly reduced. This leads to a more identifiable plateau region, especially when guided by the expected onset of plateaus observed in other ensembles at similar lattice spacing. However, the quantitative results obtained from the two methods are not fully consistent. The values of r_0 extracted from the standard analysis exhibit smaller statistical uncertainties but lead to central values that differ from those obtained using the pencil-of-function method. In contrast, the results from the pencil-of-function analysis have larger uncertainties, which encompass the corresponding values from the standard method. This asymmetry is even more pronounced for r_1, where the discrepancy between the central values becomes more significant. These observations indicate that the difference between the two methods is not purely statistical, but reflects a genuine systematic effect related to the stability of the plateau extraction. To quantify the potential systematic effect, the Fit+sys procedure was introduced. Here, the difference between the Fit and Pencil results is added in quadrature as an additional uncertainty to the Fit result. This leads to a substantial increase in the total error, amounting to an increase of approximately 240% for r_0 and 245% for r_1. While this approach yields results that are consistent with both methods, it effectively inflates the uncertainty to a level that significantly reduces the constraining power of the observable.

After extracting the effective masses using both the standard fit method and the pencil-of-function method, we propagated the corresponding results through the full continuum and chiral extrapolation procedure. In addition to these two baseline approaches, the Fit+sys variations and removing E300 altogether was considered in order to assess the systematic impact of the E300 ensemble:

- **Fit**: Standard analysis using the original GEVP results.
- **Pencil**: Analysis based on the pencil-of-function method.

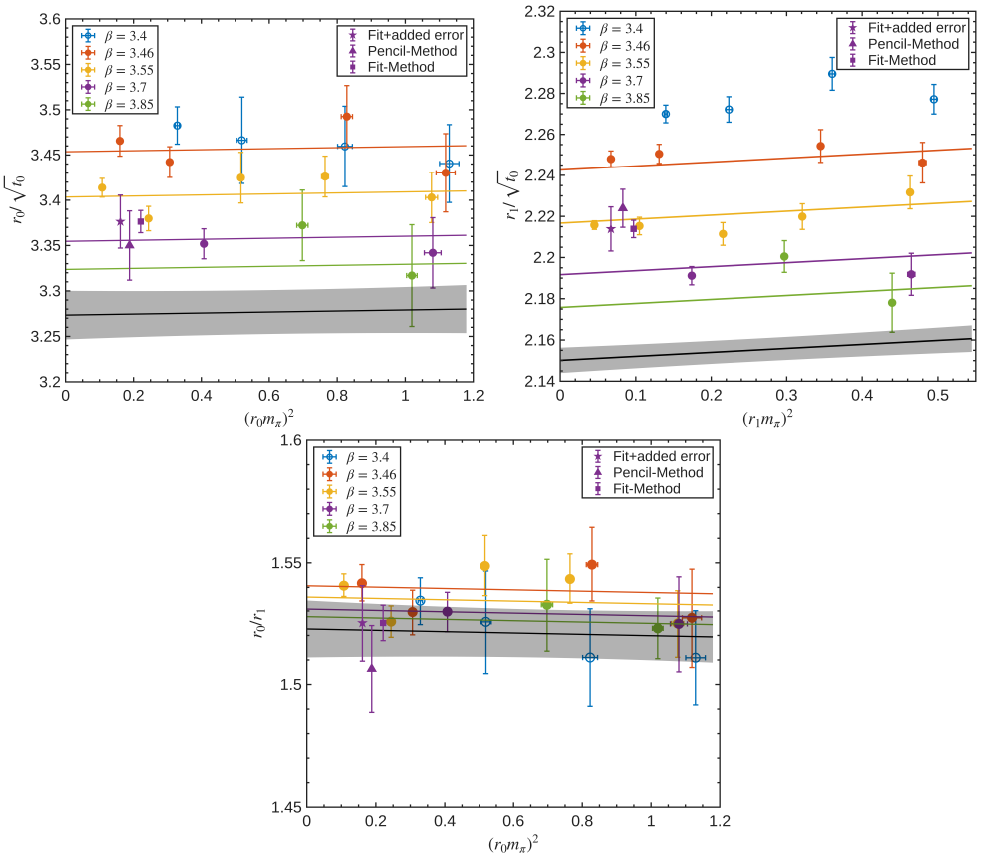

FIGURE A.3: Continuum and chiral extrapolations of r_0, r_1, and r_0/r_1 using different treatments of the E300 ensemble. The three methods (Fit, Pencil, and Fit+Err) are shown simultaneously, with a small horizontal offset applied for visual clarity. While the overall behaviour is consistent across methods, the spread in the extrapolated values reflects the systematic uncertainty associated with the extraction of the effective masses on the E300 ensemble.

- **Fit+sys**: Standard fit result augmented by an additional systematic uncertainty, defined by the difference between the Fit and Pencil central values.
- **No E300**: Analysis excluding the E300 ensemble entirely.

The impact of these different treatments on the continuum and chiral extrapolation is illustrated in Fig. A.3. The three panels show the extrapolations for r_0, r_1, and the ratio r_0/r_1, respectively. For clarity, the data points corresponding to the different methods are slightly shifted along the horizontal axis, although they correspond to the same underlying values of the dimensionless mass variable.

The resulting continuum and chiral extrapolated values for r_0, r_1, and their ratio using the Fit1 ansatz defined in eq. (5.3) are summarised in Table A.1. One could argue that the Fit+Err approach provides the most conservative estimate, as it explicitly accounts for the discrepancy between the two methods. In particular, for r_1 it leads to a significantly improved χ^2/d.o.f compared to the standard Fit result, and even slightly better than the Pencil method. However, this procedure effectively introduces an artificial uncertainty that is only associated with a single ensemble. Instead of inflating the error in an ad hoc manner, we choose to address the underlying issue directly. As discussed in the previous sections, the pencil-of-function method provides

Method	r_0	$\chi^2/d.o.f$	r_1	$\chi^2/d.o.f$	r_0/r_1	$\chi^2/d.o.f$
Fit	0.4754(57)	11/11	0.3116(33)	37/11	1.5216(126)	8.7/11
Fit+sys	0.4733(61)	10/11	0.3100(33)	19/11	1.5252(155)	8.2/11
Pencil	0.4724(61)	9.7/11	0.3104(33)	27/11	1.5224(127)	10/11
No E300	0.4726(62)	9.7/10	0.3096(33)	15/10	1.5268(166)	8/10

TABLE A.1: Comparison of continuum and chiral extrapolation results obtained using different treatments of the E300 ensemble. All results are based on the Fit1 ansatz defined in eq. (5.3). The Fit method corresponds to the standard analysis, while Pencil uses the pencil-of-function method. Fit+Err includes an additional systematic uncertainty based on the difference between the two methods. The "No E300" row shows the result when the ensemble is excluded entirely. The differences between Fit and Pencil are most pronounced for r_1, indicating a systematic effect associated with the plateau extraction.

a more stable plateau behaviour for the E300 ensemble, consistent with expectations from other ensembles at similar lattice spacing. For this reason, the Pencil method is adopted as the preferred treatment for the E300 ensemble in the final analysis. While this choice leads to slightly larger statistical uncertainties, it is based on a more reliable extraction of the ground-state signal and avoids introducing additional systematic errors by hand.

Appendix B

Cubic lattice symmetries and loop shapes

B.1 The cubic group

The symmetry group of the cube can be described through its action on the three orthonormal basis vectors x, y, and z. Each element of the group maps these vectors into permutations of themselves, possibly accompanied by sign changes. In this way, every group element can be specified by its action on the ordered triple (x, y, z).

The 24 elements of the cubic group are listed below in terms of these transformations. The ordering of the elements is not unique; the convention adopted here follows that used in Ref. [19]. A code written by J. A. Urrea-Niño and modified by Andreas Risch was used to apply the transformations to the shapes in C++.

$$
\begin{aligned}
g_0 &: (x,y,z) \to (x,y,z) \\
g_1 &: (x,y,z) \to (-x,-z,-y) \\
g_2 &: (x,y,z) \to (z,x,y) \\
g_3 &: (x,y,z) \to (-z,-y,-x) \\
g_4 &: (x,y,z) \to (y,z,x) \\
g_5 &: (x,y,z) \to (-y,-x,-z) \\
g_6 &: (x,y,z) \to (x,-y,-z) \\
g_7 &: (x,y,z) \to (-x,z,y) \\
g_8 &: (x,y,z) \to (-z,x,-y) \\
g_9 &: (x,y,z) \to (z,y,-x) \\
g_{10} &: (x,y,z) \to (-y,-z,x) \\
g_{11} &: (x,y,z) \to (y,-x,z)
\end{aligned}
\qquad
\begin{aligned}
g_{12} &: (x,y,z) \to (-x,y,-z) \\
g_{13} &: (x,y,z) \to (-z,-x,-y) \\
g_{14} &: (x,y,z) \to (-z,-x,y) \\
g_{15} &: (x,y,z) \to (z,-y,x) \\
g_{16} &: (x,y,z) \to (y,-z,-x) \\
g_{17} &: (x,y,z) \to (-y,x,z) \\
g_{18} &: (x,y,z) \to (-x,-y,z) \\
g_{19} &: (x,y,z) \to (x,-z,y) \\
g_{20} &: (x,y,z) \to (z,-x,-y) \\
g_{21} &: (x,y,z) \to (-z,y,x) \\
g_{22} &: (x,y,z) \to (-y,z,-x) \\
g_{23} &: (x,y,z) \to (y,x,-z)
\end{aligned}
$$

B.2 Loop shapes

The loop paths used in this work are specified as ordered sequences of steps along the spatial lattice directions. Each step corresponds to a displacement along one of the Cartesian axes x, y, or z, while a minus sign indicates the opposite orientation along the corresponding axis. In this way, a loop shape can be represented by an ordered tuple of directed steps on the cubic lattice [131].

Starting from a set of representative loop shapes, the complete set of symmetry-related loops can be generated by applying the transformations of the octahedral group O. Acting with the group elements g_i on the spatial directions maps a given loop into all of its symmetry equivalents under the cubic lattice rotations. The loop operators used in this work are thus constructed from the original loop shapes together with their images under the octahedral transformations. The set of loop shapes employed is listed below, where the first 22 shapes of lengths 4, 6, and 8, and some of their names[1] have been taken from [131, 137] and the remaining w_{22-34} of lengths 10 and 12 from [22]. Figure B.1 shows the 8 shapes that have been chosen from the bigger sample to represent our set of operators used in the glueball study of Teper links.

[1]Most names stem from internal conversations with J. A. Urrea-Niño

w_1 : plaquette
$(x, y, -x, -y)$

w_2 : double plaquette
$(x, x, y, -x, -x, -y)$

w_3 : bent plaquette
$(x, y, -x, z, -y, -z)$

w_4 : twisted plaquette
$(x, y, z, -x, -y, -z)$

w_5 : 2×2 plaquette
$(z, z, x, x, -z, -z, -x, -x)$

w_6 : sofa
$(y, y, -x, z, -y, -y, -z, x)$

w_7 : l
$(z, z, y, -z, y, -z, -y, -y)$

w_8 : hand
$(-y, -y, z, -x, y, x, y, -z)$

w_9 : one-legged chair
$(y, -z, -x, z, z, -y, -z, x)$

w_{10} : triple plaquette
$(y, y, y, z, -y, -y, -y, -z)$

w_{11} : once-bent triple plaquette
$(y, y, -x, z, x, -y, -y, -z)$

w_{12} : double-bent triple plaquette
$(y, -x, z, x, -y, x, -z, -x)$

w_{13} : claw
$(z, -x, y, x, -z, -x, -y, x)$

w_{14} : 2×2 twisted plaquette
$(y, y, z, -x, -y, -y, -z, x)$

w_{15} : double-winded plaquette
$(y, z, -y, -z, y, z, -y, -z)$

w_{16} : twisted 2×1 plaquette
$(y, z, y, -z, -y, z, -y, -z)$

w_{17} : first planar fish
$(y, y, -z, -y, z, z, -y, -z)$

w_{18} : first solid fish
$(y, z, z, -x, -z, x, -y, -z)$

w_{19} : second planar fish
$(y, -z, y, z, -y, z, -y, -z)$

w_{20} : second solid fish
$(y, z, -x, z, x, -z, -y, -z)$

w_{21} : bent twisted plaquette
$(y, z, -x, -z, x, z, -y, -z)$

w_{22} : twisted plaquette with a flap
$(y, y, z, -y, -x, -y, -z, x)$

w_{23} : twisted dumbbell
$(y, -z, -y, z, z, -x, z, x, -z, -z)$

w_{24} : stairs
$(x, y, -x, -y, z, y, -x, -y, x, -z)$

w_{25} : first flag
$(x, y, -x, -y, z, z, -x, -z, x, -z)$

w_{26} : first flag with reversed upper part
$(x, y, -x, -y, z, -x, z, x, -z, -z)$

w_{27} : second flag
$(x, y, -x, -y, z, y, z, -y, -z, -z)$

w_{28} : third flag
$(x, y, -x, -y, z, x, z, -x, -z, -z)$

w_{29} : fourth flag
$(x, y, -x, -y, z, -y, z, y, -z, -z)$

w_{30} : hole
$(x, z, z, y, -x, -z, -x, -y, x, -z)$

w_{31} : flag with a wall
$(x, y, z, x, z, -x, -z, -z, -x, z, -y, -z)$

w_{32} : bus-stop
$(x, y, y, z, -y, -x, y, -z, -y, -x, -y, x)$

w_{33} : swan
$(x, y, -x, -y, -x, y, x, -y, -z, -y, z, y)$

w_{34} : twice-folded cross
$(x, z, z, y, -z, y, -x, -y, -x, -y, x, -z)$

w_{35} : displaced flag
$(z, z, -y, -z, y, -z, y, y, x, -y, -x, -y)$

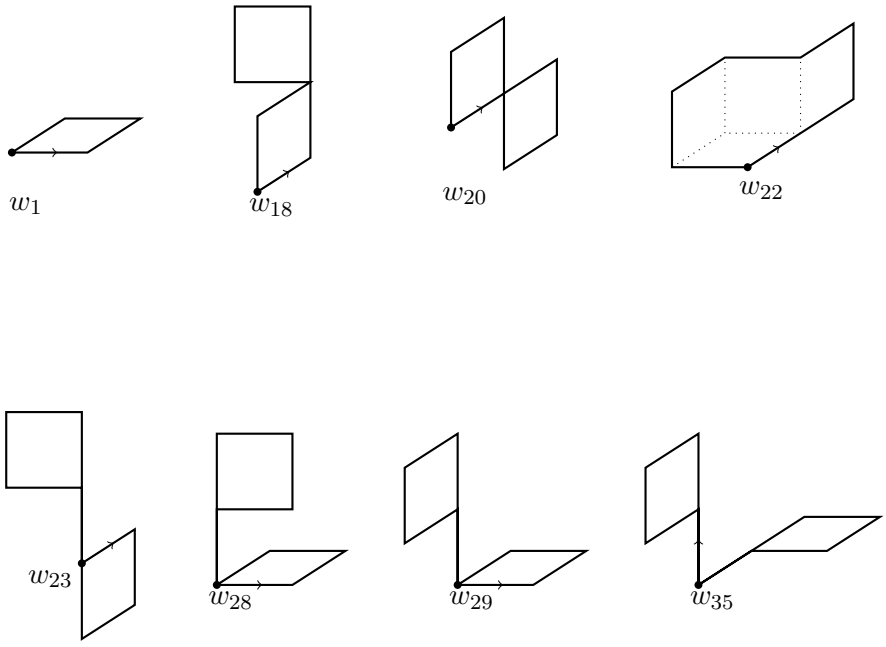

FIGURE B.1: Selected Wilson-loop shapes used in the operator basis.

Bibliography

[1] D. J. Gross and F. Wilczek, "Asymptotically Free Gauge Theories - I", Phys. Rev. D **8**, 3633 (1973).

[2] H. David Politzer, "Asymptotic freedom: an approach to strong interactions", Physics Reports **14**, 129 (1974).

[3] K. G. Wilson, "Confinement of Quarks", Phys. Rev. D **10**, 2445 (1974).

[4] M. Creutz, "Monte carlo study of quantized su(2) gauge theory", Phys. Rev. D **21**, 2308 (1980).

[5] M. Lüscher, "Properties and uses of the Wilson flow in lattice QCD", JHEP **08**, [Erratum: JHEP 03, 092 (2014)], 071 (2010).

[6] S. Borsányi, S. Dürr, Z. Fodor, C. Hoelbling, S. D. Katz, S. Krieg, T. Kurth, L. Lellouch, T. Lippert, and C. McNeile (BMW), "High-precision scale setting in lattice QCD", JHEP **09**, 010 (2012).

[7] R. Sommer, "A New way to set the energy scale in lattice gauge theories and its applications to the static force and alpha-s in SU(2) Yang-Mills theory", Nucl. Phys. B **411**, 839 (1994).

[8] M. Guagnelli, R. Sommer, and H. Wittig (ALPHA), "Precision computation of a low-energy reference scale in quenched lattice QCD", Nucl. Phys. B **535**, 389 (1998).

[9] C. W. Bernard, T. Burch, K. Orginos, D. Toussaint, T. A. DeGrand, C. E. DeTar, S. A. Gottlieb, U. M. Heller, J. E. Hetrick, and B. Sugar, "The Static quark potential in three flavor QCD", Phys. Rev. D **62**, 034503 (2000).

[10] R. L. Jaffe and K. Johnson, "Unconventional states of confined quarks and gluons", Physics Letters B **60**, 201 (1976).

[11] V. Mathieu, N. Kochelev, and V. Vento, "The physics of glueballs", International Journal of Modern Physics E **18**, 1 (2009).

[12] W. Ochs, "The status of glueballs", Journal of Physics G: Nuclear and Particle Physics **40**, 043001 (2013).

[13] C. J. Morningstar and M. J. Peardon, "The Glueball spectrum from an anisotropic lattice study", Phys. Rev. D **60**, 034509 (1999).

[14] Y. Chen et al., "Glueball spectrum and matrix elements on anisotropic lattices", Phys. Rev. D **73**, 014516 (2006).

[15] B. Lucini, M. Teper, and U. Wenger, "Glueballs and k-strings in su(n) gauge theories: calculations with improved operators", Journal of High Energy Physics **2004**, 012 (2004).

[16] Y. Liang, K. Liu, B. Li, S. Dong, and K. Ishikawa, "Lattice calculation of glueball matrix elements", Physics Letters B **307**, 375 (1993).

[17] M. Teper, "An Improved Method for Lattice Glueball Calculations", Phys. Lett. B **183**, 345 (1987).

[18] M. Albanese et al. (APE), "Glueball Masses and String Tension in Lattice QCD", Phys. Lett. B **192**, 163 (1987).

[19] J. A. Urrea-Niño, "Optimal creation operators for charmonium spectroscopy on the lattice", PhD thesis (Bergische Universität Wuppertal, Wuppertal, 2023).

[20] J. A. Urrea-Niño, F. Knechtli, J. Finkenrath, T. Korzec, M. J. Peardon, and R. Höllwieser, "Flavor mixing in charmonium and light mesons with optimal distillation profiles", PoS **LATTICE2024**, 080 (2025).

[21] J. A. Urrea-Niño, F. Knechtli, T. Korzec, and M. Peardon, "Charmonium-Glueball spectroscopy with improved hadron creation operators", (2026).

[22] L. Barca, S. Schaefer, F. Knechtli, J. A. Urrea-Niño, S. Martins, and M. Peardon, "Exponential error reduction for glueball calculations using a two-level algorithm in pure gauge theory", Phys. Rev. D **110**, 054515 (2024).

[23] F. Knechtli, J. Finkenrath, R. Höllwieser, T. Korzec, M. J. Peardon, and J. A. Urrea-Niño, "Mixing of heavy and light quarks in charmonium and light mesons", PoS **QCHSC24**, 121 (2025).

[24] I. Montvay and G. Münster, *Quantum fields on a lattice*, Cambridge Monographs on Mathematical Physics (Cambridge University Press, 1994).

[25] C. Gattringer and C. B. Lang, *Quantum chromodynamics on the lattice*, Vol. 788 (Springer, Berlin, 2010).

[26] F. Knechtli, M. Günther, and M. Peardon, *Lattice Quantum Chromodynamics: Practical Essentials*, SpringerBriefs in Physics (Springer, 2017).

[27] D. J. Gross, "Nobel lecture: the discovery of asymptotic freedom and the emergence of qcd", Rev. Mod. Phys. **77**, 837 (2005).

[28] K. Symanzik, "Continuum limit and improved action in lattice theories: (i). principles and 4 theory", Nuclear Physics B **226**, 187 (1983).

[29] K. Symanzik, "Continuum limit and improved action in lattice theories: (ii). o(n) non-linear sigma model in perturbation theory", Nuclear Physics B **226**, 205 (1983).

[30] B. Sheikholeslami and R. Wohlert, "Improved Continuum Limit Lattice Action for QCD with Wilson Fermions", Nucl. Phys. B **259**, 572 (1985).

[31] M. Lüscher and P. Weisz, "On-shell improved lattice gauge theories", Commun. Math. Phys. **98**, [Erratum: Commun.Math.Phys. 98, 433 (1985)], 433 (1985).

[32] M. Peter, "The Static potential in QCD: A Full two loop calculation", Nucl. Phys. B **501**, 471 (1997).

[33] Y. Schröder, "The Static potential in QCD to two loops", Phys. Lett. B **447**, 321 (1999).

[34] M. Melles, "The Static QCD potential in coordinate space with quark masses through two loops", Phys. Rev. D **62**, 074019 (2000).

[35] C. Anzai, Y. Kiyo, and Y. Sumino, "Static QCD potential at three-loop order", Phys. Rev. Lett. **104**, 112003 (2010).

[36] N. Brambilla, A. Vairo, X. Garcia i Tormo, and J. Soto, "The QCD static energy at NNNLL", Phys. Rev. D **80**, 034016 (2009).

[37] S. J. Brodsky, M. Melles, and J. Rathsman, "The Two loop scale dependence of the static QCD potential including quark masses", Phys. Rev. D **60**, 096006 (1999).

[38] M. Donnellan, F. Knechtli, B. Leder, and R. Sommer, "Determination of the Static Potential with Dynamical Fermions", Nucl. Phys. B **849**, 45 (2011).

[39] Y. Nambu, "QCD and the String Model", Phys. Lett. B **80**, 372 (1979).

[40] M. Lüscher and P. Weisz, "Quark confinement and the bosonic string", JHEP **07**, 049 (2002).

[41] E. Eichten, K. Gottfried, T. Kinoshita, K. D. Lane, and T.-M. Yan, "Charmonium: Comparison with Experiment", Phys. Rev. D **21**, 203 (1980).

[42] J. L. Richardson, "The Heavy Quark Potential and the Upsilon, J/psi Systems", Phys. Lett. B **82**, 272 (1979).

[43] M. Della Morte, A. Shindler, and R. Sommer, "On lattice actions for static quarks", JHEP **08**, 051 (2005).

[44] A. Grimbach, D. Guazzini, F. Knechtli, and F. Palombi, "O(a) improvement of the HYP static axial and vector currents at one-loop order of perturbation theory", JHEP **03**, 039 (2008).

[45] S. Necco and R. Sommer, "The N(f) = 0 heavy quark potential from short to intermediate distances", Nucl. Phys. B **622**, 328 (2002).

[46] P. Weisz and R. Wohlert, "Continuum limit improved lattice action for pure yang-mills theory (ii)", Nuclear Physics B **236**, 397 (1984).

[47] S. Necco, "The static quark potential and scaling behavior of su(3) l attice yang-mills theory", PhD thesis (2003).

[48] M. Lüscher, "Trivializing maps, the wilson flow and the hmc algorithm", Communications in Mathematical Physics **293**, 899 (2009).

[49] M. Lüscher, R. Narayanan, P. Weisz, and U. Wolff, "The schrödinger functional — a renormalizable probe for non-abelian gauge theories", Nuclear Physics B **384**, 168 (1992).

[50] M. Lüscher, R. Sommer, P. Weisz, and U. Wolff, "A precise determination of the running coupling in the su(3) yang-mills theory", Nuclear Physics B **413**, 481 (1994).

[51] A. Hasenfratz and F. Knechtli, "Flavor symmetry and the static potential with hypercubic blocking", Phys. Rev. D **64**, 034504 (2001).

[52] S. Capitani, S. Dürr, and C. Hoelbling, "Rationale for uv-filtered clover fermions", Journal of High Energy Physics **2006**, 028 (2006).

[53] F. Knechtli and A. Hasenfratz, "Dynamical fermions with fat links", Phys. Rev. D **63**, 114502 (2001).

[54] M. Bruno et al., "Simulation of QCD with $N_f = 2 + 1$ flavors of non-perturbatively improved Wilson fermions", JHEP **02**, 043 (2015).

[55] D. Mohler, S. Schaefer, and J. Simeth, "CLS 2+1 flavor simulations at physical light- and strange-quark masses", EPJ Web Conf. **175**, edited by M. Della Morte, P. Fritzsch, E. Gámiz Sánchez, and C. Pena Ruano, 02010 (2018).

[56] P. Weisz, "Continuum Limit Improved Lattice Action for Pure Yang-Mills Theory. 1.", Nucl. Phys. B **212**, 1 (1983).

[57] M. Gell-Mann, R. J. Oakes, and B. Renner, "Behavior of current divergences under SU(3) x SU(3)", Phys. Rev. **175**, 2195 (1968).

[58] O. Bar and M. Golterman, "Chiral perturbation theory for gradient flow observables", Phys. Rev. D **89**, [Erratum: Phys.Rev.D 89, 099905 (2014)], 034505 (2014).

[59] S. Aoki et al., "Review of Lattice Results Concerning Low-Energy Particle Physics", Eur. Phys. J. C **74**, 2890 (2014).

[60] B. Straßberger et al., "Scale setting for CLS 2+1 simulations", PoS **LATTICE2021**, 135 (2022).

[61] M. Luscher, S. Sint, R. Sommer, and P. Weisz, "Chiral symmetry and O(a) improvement in lattice QCD", Nucl. Phys. B **478**, 365 (1996).

[62] M. Bruno, T. Korzec, and S. Schaefer, "Setting the scale for the CLS $2 + 1$ flavor ensembles", Phys. Rev. D **95**, 074504 (2017).

[63] S. Duane, A. D. Kennedy, B. J. Pendleton, and D. Roweth, "Hybrid Monte Carlo", Phys. Lett. B **195**, 216 (1987).

[64] T. A. DeGrand, "A Conditioning Technique for Matrix Inversion for Wilson Fermions", Comput. Phys. Commun. **52**, 161 (1988).

[65] M. Hasenbusch, "Speeding up the hybrid Monte Carlo algorithm for dynamical fermions", Phys. Lett. B **519**, 177 (2001).

[66] M. Hasenbusch and K. Jansen, "Speeding up lattice QCD simulations with clover improved Wilson fermions", Nucl. Phys. B **659**, 299 (2003).

[67] S. Schaefer, "Status and challenges of simulations with dynamical fermions", PoS **LATTICE2012**, edited by D. Leinweber, W. Kamleh, S. Mahbub, H. Matevosyan, A. Thomas, A. G. Williams, R. Young, and J. Zanotti, 001 (2012).

[68] M. Lüscher and F. Palombi, "Fluctuations and reweighting of the quark determinant on large lattices", PoS **LATTICE2008**, edited by C. Aubin, S. Cohen, C. Dawson, J. Dudek, R. Edwards, B. Joo, H.-W. Lin, K. Orginos, D. Richards, and H. Thacker, 049 (2008).

[69] A. D. Kennedy, I. Horvath, and S. Sint, "A New exact method for dynamical fermion computations with nonlocal actions", Nucl. Phys. B Proc. Suppl. **73**, edited by T. A. DeGrand, C. E. DeTar, R. Sugar, and D. Toussaint, 834 (1999).

[70] S. Kuberski, "Low-mode deflation for twisted-mass and RHMC reweighting in lattice QCD", Comput. Phys. Commun. **300**, 109173 (2024).

[71] D. Mohler and S. Schaefer, "Remarks on strange-quark simulations with Wilson fermions", Phys. Rev. D **102**, 074506 (2020).

[72] M. Lüscher and S. Schaefer, "Lattice QCD without topology barriers", JHEP **07**, 036 (2011).

[73] M. Lüscher and S. Schaefer, "Lattice QCD with open boundary conditions and twisted-mass reweighting", Comput. Phys. Commun. **184**, 519 (2013).

[74] Y. Aoki et al. (Flavour Lattice Averaging Group (FLAG)), "FLAG review 2024", Phys. Rev. D **113**, 014508 (2026).

[75] S. Schaefer, R. Sommer, and F. Virotta (ALPHA), "Critical slowing down and error analysis in lattice QCD simulations", Nucl. Phys. B **845**, 93 (2011).

[76] U. Wolff (ALPHA), "Monte Carlo errors with less errors", Comput. Phys. Commun. **156**, [Erratum: Comput.Phys.Commun. 176, 383 (2007)], 143 (2004).

[77] N. Madras and A. D. Sokal, "The Pivot algorithm: a highly efficient Monte Carlo method for selfavoiding walk", J. Statist. Phys. **50**, 109 (1988).

[78] H. Simma, R. Sommer, and F. Virotta, "General error computation in lattice gauge theory", Internal notes of the ALPHA collaboration (2012-2014).

[79] A. Ramos, "Automatic differentiation for error analysis of Monte Carlo data", Comput. Phys. Commun. **238**, 19 (2019).

[80] F. Joswig, S. Kuberski, J. T. Kuhlmann, and J. Neuendorf, "pyerrors: A python framework for error analysis of Monte Carlo data", Comput. Phys. Commun. **288**, 108750 (2023).

[81] G. S. Bali, S. Collins, P. Georg, D. Jenkins, P. Korcyl, A. Schäfer, E. E. Scholz, J. Simeth, W. Söldner, and S. Weishäupl (RQCD), "Scale setting and the light baryon spectrum in $N_f = 2 + 1$ QCD with Wilson fermions", JHEP **05**, 035 (2023).

[82] G. P. Lepage, "The Analysis of Algorithms for Lattice Field Theory", in Theoretical Advanced Study Institute in Elementary Particle Physics (June 1989).

[83] G. Parisi, R. Petronzio, and F. Rapuano, "A Measurement of the String Tension Near the Continuum Limit", Phys. Lett. B **128**, 418 (1983).

[84] M. Lüscher and P. Weisz, "Locality and exponential error reduction in numerical lattice gauge theory", JHEP **09**, 010 (2001).

[85] A. Hasenfratz, R. Hoffmann, and F. Knechtli, "The Static potential with hypercubic blocking", Nucl. Phys. B Proc. Suppl. **106**, edited by M. Muller-Preussker, W. Bietenholz, K. Jansen, F. Jegerlehner, I. Montvay, G. Schierholz, R. Sommer, and U. Wolff, 418 (2002).

[86] C. Alexandrou, P. De Forcrand, and A. Tsapalis, "The Static three quark SU(3) and four quark SU(4) potentials", Phys. Rev. D **65**, 054503 (2002).

[87] E. Eichten and B. R. Hill, "An Effective Field Theory for the Calculation of Matrix Elements Involving Heavy Quarks", Phys. Lett. B **234**, 511 (1990).

[88] B. Leder, *Wilson loop computation*, (2015) `https://github.com/bjoern-leder/wloop`.

[89] N. A. Campbell, A. Huntley, and C. Michael, "Heavy Quark Potentials and Hybrid Mesons From SU(3) Lattice Gauge Theory", Nucl. Phys. B **306**, 51 (1988).

[90] M. Lüscher and U. Wolff, "How to Calculate the Elastic Scattering Matrix in Two-dimensional Quantum Field Theories by Numerical Simulation", Nucl. Phys. B **339**, 222 (1990).

[91] B. Blossier, M. Della Morte, G. von Hippel, T. Mendes, and R. Sommer, "On the generalized eigenvalue method for energies and matrix elements in lattice field theory", JHEP **04**, 094 (2009).

[92] M. Fischer, B. Kostrzewa, J. Ostmeyer, K. Ottnad, M. Ueding, and C. Urbach, "On the generalised eigenvalue method and its relation to Prony and generalised pencil of function methods", Eur. Phys. J. A **56**, 206 (2020).

[93] T. M. B. Asmussen, R. Höllwieser, F. Knechtli, and T. Korzec, "The determination of potential scales in 2+1 flavor QCD", Eur. Phys. J. C **85**, 673 (2025).

[94] K. J. Juge, J. Kuti, and C. J. Morningstar, "Gluon excitations of the static quark potential and the hybrid quarkonium spectrum", Nucl. Phys. B Proc. Suppl. **63**, edited by C. T. H. Davies, I. M. Barbour, K. C. Bowler, R. D. Kenway, B. J. Pendleton, and D. G. Richards, 326 (1998).

[95] J. Bulava, F. Knechtli, V. Koch, C. Morningstar, and M. Peardon, "The quark-mass dependence of the potential energy between static colour sources in the QCD vacuum with light and strange quarks", Phys. Lett. B **854**, 138754 (2024).

[96] K. Symanzik, "Cutoff dependence in lattice ϕ_4^4 theory", NATO Sci. Ser. B **59**, edited by G. 't Hooft, C. Itzykson, A. Jaffe, H. Lehmann, P. K. Mitter, I. M. Singer, and R. Stora, 313 (1980).

[97] J. Balog, F. Niedermayer, and P. Weisz, "Logarithmic corrections to O(a**2) lattice artifacts", Phys. Lett. B **676**, 188 (2009).

[98] N. Husung, P. Marquard, and R. Sommer, "Asymptotic behavior of cutoff effects in Yang–Mills theory and in Wilson's lattice QCD", Eur. Phys. J. C **80**, 200 (2020).

[99] N. Husung, P. Marquard, and R. Sommer, "The asymptotic approach to the continuum of lattice QCD spectral observables", Phys. Lett. B **829**, 137069 (2022).

[100] H. Akaike, "A new look at the statistical model identification", IEEE Transactions on Automatic Control **19**, 716 (1974).

[101] W. I. Jay and E. T. Neil, "Bayesian model averaging for analysis of lattice field theory results", Phys. Rev. D **103**, 114502 (2021).

[102] E. T. Neil and J. W. Sitison, "Improved information criteria for Bayesian model averaging in lattice field theory", Phys. Rev. D **109**, 014510 (2024).

[103] I. Campos Plasencia, M. Dalla Brida, G. M. de Divitiis, A. Lytle, M. Papinutto, L. Pirelli, and A. Vladikas (ALPHA), "Nonperturbative running of the tensor operator for Nf=3 QCD from the chirally rotated Schrödinger functional", Phys. Rev. D **109**, 054511 (2024).

[104] N. Carrasco et al. (European Twisted Mass), "Up, down, strange and charm quark masses with $N_f = 2+1+1$ twisted mass lattice QCD", Nucl. Phys. B **887**, 19 (2014).

[105] A. Bazavov et al. (MILC), "MILC results for light pseudoscalars", PoS **CD09**, 007 (2009).

[106] Y.-B. Yang et al., "Charm and strange quark masses and f_{D_s} from overlap fermions", Phys. Rev. D **92**, 034517 (2015).

[107] R. J. Dowdall, C. T. H. Davies, G. P. Lepage, and C. McNeile, "Vus from pi and K decay constants in full lattice QCD with physical u, d, s and c quarks", Phys. Rev. D **88**, 074504 (2013).

[108] A. Bazavov et al. (MILC), "Results for light pseudoscalar mesons", PoS **LATTICE2010**, edited by G. Rossi, 074 (2010).

[109] A. Bazavov et al. (HotQCD), "Equation of state in (2+1)-flavor QCD", Phys. Rev. D **90**, 094503 (2014).

[110] Y. Aoki et al. (RBC, UKQCD), "Continuum Limit Physics from 2+1 Flavor Domain Wall QCD", Phys. Rev. D **83**, 074508 (2011).

[111] A. Gray, I. Allison, C. T. H. Davies, E. Dalgic, G. P. Lepage, J. Shigemitsu, and M. Wingate, "The Upsilon spectrum and m(b) from full lattice QCD", Phys. Rev. D **72**, 094507 (2005).

[112] C. Aubin, C. Bernard, C. DeTar, J. Osborn, S. Gottlieb, E. B. Gregory, D. Toussaint, U. M. Heller, J. E. Hetrick, and R. Sugar, "Light hadrons with improved staggered quarks: Approaching the continuum limit", Phys. Rev. D **70**, 094505 (2004).

[113] C. T. H. Davies, E. Follana, I. D. Kendall, G. P. Lepage, and C. McNeile (HPQCD), "Precise determination of the lattice spacing in full lattice QCD", Phys. Rev. D **81**, 034506 (2010).

[114] A. Bazavov et al. (MILC), "Nonperturbative QCD Simulations with 2+1 Flavors of Improved Staggered Quarks", Rev. Mod. Phys. **82**, 1349 (2010).

[115] S. Aoki et al. (PACS-CS), "2+1 Flavor Lattice QCD toward the Physical Point", Phys. Rev. D **79**, 034503 (2009).

[116] R. J. Dowdall et al. (HPQCD), "The Upsilon spectrum and the determination of the lattice spacing from lattice QCD including charm quarks in the sea", Phys. Rev. D **85**, 054509 (2012).

[117] A. Bazavov et al., "The chiral and deconfinement aspects of the QCD transition", Phys. Rev. D **85**, 054503 (2012).

[118] N. Brambilla, R. L. Delgado, A. S. Kronfeld, V. Leino, P. Petreczky, S. Steinbeißer, A. Vairo, and J. H. Weber (TUMQCD), "Static energy in $(2+1+1)$-flavor lattice QCD: Scale setting and charm effects", Phys. Rev. D **107**, 074503 (2023).

[119] Y. Aoki et al. (Flavour Lattice Averaging Group (FLAG)), "FLAG Review 2021", Eur. Phys. J. C **82**, 869 (2022).

[120] M. Bruno, J. Finkenrath, F. Knechtli, B. Leder, and R. Sommer (ALPHA), "Effects of Heavy Sea Quarks at Low Energies", Phys. Rev. Lett. **114**, 102001 (2015).

[121] F. Knechtli, T. Korzec, B. Leder, and G. Moir (ALPHA), "Power corrections from decoupling of the charm quark", Phys. Lett. B **774**, 649 (2017).

[122] M. Dalla Brida, "Past, present, and future of precision determinations of the QCD parameters from lattice QCD", Eur. Phys. J. A **57**, 66 (2021).

[123] M. Dalla Brida, R. Höllwieser, F. Knechtli, T. Korzec, A. Ramos, S. Sint, and R. Sommer, "High-precision calculation of the quark–gluon coupling from lattice QCD", Nature **652**, 328 (2026).

[124] M. Bruno, M. Dalla Brida, P. Fritzsch, T. Korzec, A. Ramos, S. Schaefer, H. Simma, S. Sint, and R. Sommer (ALPHA), "QCD Coupling from a Nonperturbative Determination of the Three-Flavor Λ Parameter", Phys. Rev. Lett. **119**, 102001 (2017).

[125] A. Bazavov, N. Brambilla, X. Garcia i Tormo, P. Petreczky, J. Soto, A. Vairo, and J. H. Weber (TUMQCD), "Determination of the QCD coupling from the static energy and the free energy", Phys. Rev. D **100**, 114511 (2019).

[126] S. Cali, K. Cichy, P. Korcyl, and J. Simeth, "Running coupling constant from position-space current-current correlation functions in three-flavor lattice QCD", Phys. Rev. Lett. **125**, 242002 (2020).

[127] C. Ayala, X. Lobregat, and A. Pineda, "Determination of $\alpha(M_z)$ from an hyperasymptotic approximation to the energy of a static quark-antiquark pair", JHEP **09**, 016 (2020).

[128] S. Aoki et al. (PACS-CS), "Precise determination of the strong coupling constant in $N_f = 2{+}1$ lattice QCD with the Schrodinger functional scheme", JHEP **10**, 053 (2009).

[129] K. Maltman, D. Leinweber, P. Moran, and A. Sternbeck, "The Realistic Lattice Determination of alpha(s)(M(Z)) Revisited", Phys. Rev. D **78**, 114504 (2008).

[130] C. McNeile, C. T. H. Davies, E. Follana, K. Hornbostel, and G. P. Lepage, "High-Precision c and b Masses, and QCD Coupling from Current-Current Correlators in Lattice and Continuum QCD", Phys. Rev. D **82**, 034512 (2010).

[131] B. Berg and A. Billoire, "Glueball spectroscopy in 4d su(3) lattice gauge theory (i)", Nuclear Physics B **221**, 109 (1983).

[132] S. L. Altmann and P. M. Herzig, *Point-group theory tables* (Clarendon Press, Oxford, 1994).

[133] J. E. Mandula, G. Zweig, and J. Govaerts, "Representations of the rotation reflection symmetry group of the four-dimensional cubic lattice", Nuclear Physics B **228**, 91 (1983).

[134] R. Johnson, "Angular momentum on a lattice", Physics Letters B **114**, 147 (1982).

[135] H. Georgi, *Lie Algebras In Particle Physics : from Isospin To Unified Theories* (Taylor & Francis, Boca Raton, 2000).

[136] P. R. Bunker and P. Jensen, *Molecular symmetry and spectroscopy*, 2nd (NRC Research Press, Ottawa, 1998).

[137] K. Sakai and S. Sasaki, "Glueball spectroscopy in lattice QCD using gradient flow", Phys. Rev. D **107**, 034510 (2023).

[138] J. J. Dudek, R. G. Edwards, M. J. Peardon, D. G. Richards, and C. E. Thomas (for the Hadron Spectrum Collaboration), "Toward the excited meson spectrum of dynamical qcd", Phys. Rev. D **82**, 034508 (2010).

[139] J. Bulava, B. Hörz, F. Knechtli, V. Koch, G. Moir, C. Morningstar, and M. Peardon, "String breaking by light and strange quarks in QCD", Phys. Lett. B **793**, 493 (2019).

[140] V. Koch, "String breaking from lattice qcd with $N_f = 2 + 1$ dynamical fermions", PhD thesis (Trinity College Dublin, 2019).

[141] J. Balog, M. Niedermaier, F. Niedermayer, A. Patrascioiu, E. Seiler, and P. Weisz, "Comparison of the O(3) bootstrap sigma model with the lattice regularization at low-energies", Phys. Rev. D **60**, 094508 (1999).

[142] F. Niedermayer, P. Rufenacht, and U. Wenger, "Fixed point gauge actions with fat links: Scaling and glueballs", Nucl. Phys. B **597**, 413 (2001).

[143] J. A. U. Niño, F. Knechtli, T. Korzec, and M. Peardon, "Optimizing distillation for charmonium and glueballs", PoS **LATTICE2021**, 314 (2022).

[144] J. J. Dudek, R. G. Edwards, N. Mathur, and D. G. Richards, "Charmonium excited state spectrum in lattice qcd", Phys. Rev. D **77**, 034501 (2008).

[145] G. Bailas, B. Blossier, and V. Morénas, "Some hadronic parameters of charmonia in $N_f = 2$ lattice qcd", The European Physical Journal C **78**, 965 (2018).

[146] P. Fritzsch, F. Knechtli, B. Leder, M. Marinkovic, S. Schaefer, R. Sommer, and F. Virotta, "The strange quark mass and Lambda parameter of two flavor QCD", Nucl. Phys. B **865**, 397 (2012).

[147] S. Cali, F. Knechtli, and T. Korzec, "How much do charm sea quarks affect the charmonium spectrum?", Eur. Phys. J. C **79**, 607 (2019).

[148] P. Fritzsch, R. Sommer, F. Stollenwerk, and U. Wolff (ALPHA), "Symanzik improvement with dynamical charm: a 3+1 scheme for Wilson quarks", JHEP **06**, [Erratum: JHEP 10, 165 (2020)], 025 (2018).

[149] R. Höllwieser, F. Knechtli, and T. Korzec (ALPHA), "Scale setting for $N_f = 3 + 1$ QCD", Eur. Phys. J. C **80**, 349 (2020).

[150] C. Aubin, K. Orginos, D. Armstrong, V. Burkert, J.-P. Chen, W. Detmold, J. Dudek, W. Melnitchouk, and D. Richards, "A new approach for delta form factors", in Aip conference proceedings (2011), pp. 621–624.

[151] C. Aubin and K. Orginos, "An improved method for extracting matrix elements from lattice three-point functions", PoS **Lattice 2011**, 148 (2012).

[152] R. W. Schiel, "Expanding the interpolator basis in the variational method to explicitly account for backward running states", Phys. Rev. D **92**, 034512 (2015).

[153] K. Ottnad, T. Harris, H. Meyer, G. von Hippel, J. Wilhelm, and H. Wittig, "Nucleon average quark momentum fraction with $N_\mathrm{f} = 2 + 1$ Wilson fermions", EPJ Web Conf. **175**, edited by M. Della Morte, P. Fritzsch, E. Gámiz Sánchez, and C. Pena Ruano, 06026 (2018).