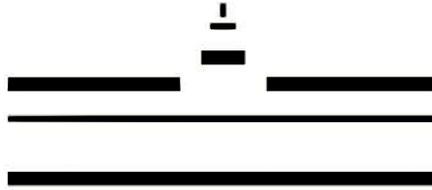

PATRICK FRITZSCH

B-meson properties from non-perturbative matching of

HQET to finite-volume two-flavour QCD

2009



Theoretische Physik

**B-meson properties from non-perturbative
matching of HQET to finite-volume
two-flavour QCD**

Inaugural-Dissertation
zur Erlangung des Doktorgrades
der Mathematisch-Naturwissenschaftlichen Fakultät
der Westfälischen Wilhelms-Universität Münster

vorgelegt von
PATRICK FRITZSCH
aus Stollberg

–2009–

Dekan: Prof. Dr. Johannes P. Wessels
1. Gutachter: Priv.-Doz. Dr. Jochen Heitger
2. Gutachter: Prof. Dr. Gernot Münster
Tag der mündlichen Prüfung: 15.09.2009
Tag der Promotion: 15.09.2009

für meine Familie

Abstract

A precise value for the b-quark mass is of great importance in the present experimental program of precision flavour physics at B-factories. Studying B-meson properties allows to overconstrain the elements of the Cabibbo–Kobayashi–Maskawa (CKM) quark-mixing matrix. Thus it provides sensitive tests of the flavour sector of the Standard Model (SM). The Heavy Quark Effective Theory (HQET) is a widely used phenomenological and theoretical tool to describe QCD bound states if only one sufficiently large quark mass like m_b or m_c is involved. It phenomenologically describes the spectrum of B - and D -mesons quite successfully. But in determinations where only perturbative HQET enters, we still lack a sound theoretical understanding. Often the error introduced by such approximations cannot be estimated accurately enough if even possible.

From the theoretical point of view non-perturbative calculations are desired because they are independent and non-perturbative tests of HQET may provide a deeper insight into the feasibility of the effective theory approach. Through a matching procedure of HQET and QCD in small-volume of about 0.5 fm, we want to perform non-perturbative tests of predictions made by HQET in limit of large quark masses. To this end we extract the heavy quark mass dependence of certain heavy-light meson observables in finite-volume QCD by means of lattice QCD and cover a wide range of heavy quark masses. In terms of the renormalization group invariant heavy quark mass M , this range is given by $M \approx (1.55 - 8.13)$ GeV and covers both the charm-quark and bottom-quark region.

The non-perturbative calculations are performed in the Schrödinger functional (SF) renormalization scheme of lattice QCD. To overcome the quenched approximation we simulate two dynamical massless flavours of quarks and thus take appropriate virtual quark loops into account. Much of our effort is dedicated to the non-perturbatively setup of the on-shell $O(a)$ improved lattice theory. This is a crucial step to obtain a controlled error estimate in the end.

Keywords: lattice QCD, B physics, non-perturbative methods, renormalization, HQET

Zusammenfassung

Eine präzise Bestimmung der b-quark Masse ist von großer Bedeutung für das aktuelle experimentelle Flavourphysikprogramm an sogenannten B-factories. Das Studium der Eigenschaften von B-Meson erlaubt eine Überbestimmung der Elemente in der Cabibbo–Kobayashi–Maskawa Quarkmischungsmatrix. Dies gewährleistet sehr genaue Tests im Flavoursektor des Standardmodells. Die effektive Theorie schwerer Quarks (HQET) ist ein weit verbreitetes phänomenologisches wie auch theoretisches Werkzeug um QCD-Bindungszustände zu beschreiben solange große Quarkmassen wie m_b oder m_c involviert sind. Die HQET beschreibt das Spektrum von B- und D-Mesonen recht erfolgreich doch Bestimmungen in denen störungstheoretische HQET eingeht, bedürfen immer noch eines genaueren physikalischen Verständnisses. In vielen Fällen kann man die durch solche Näherungen eingeführten Abweichungen nur schwer oder überhaupt nicht hinreichend genau bestimmen.

Nichtstörungstheoretische Berechnungen sind wünschenswert weil sie unabhängige und gerade eben auch nichtstörungstheoretische Tests der effektiven Theorie ermöglichen. Desweiteren können tiefere Einsichten über die Verwendbarkeit der effektiven Theorie gewonnen werden. Eine nichtstörungstheoretische Anpassung von HQET und QCD in einem kleinen physikalischen Volumen von ungefähr 0.5 fm Ausdehnung, ermöglicht es nichtstörungstheoretische Tests über die Vorhersagekraft der HQET im Limes großer Quarkmassen durchzuführen. Dazu extrahieren wir die Quarkmassenabhängigkeit ausgewählter Observablen im schwer-leichten Mesonsystem. Die zugehörigen Rechnungen werden in QCD in einem endlichen physikalischen Volumen für einen großen Bereich schwerer Quarkmassen mittels Gitterregularisierung durchgeführt. Der zugehörige Bereich für die renormierungsgruppeninvariante schwere Quarkmasse ist $M \approx (1.55 - 8.13)$ GeV und deckt damit sowohl den Bereich des c-Quarks als auch den des b-Quarks ab.

Für die nichtstörungstheoretischen Berechnungen wird das Schrödingerfunktional Renormierungsschema herangezogen. Desweiteren simulieren wir ein Doublet masseloser dynamischer Quarks um den Effekten virtueller Quarkschleifen Rechnung zu tragen. Um kontrollierte Fehlerabschätzungen zu erhalten ist es notwendig eine auf der Massenschale $O(a)$ verbesserte Gittertheorie zu verwenden. Wir widmen daher dem korrekten Setup dieser Theorie große Aufmerksamkeit.

Schlüsselwörter: Gitter QCD, B-Physik, nichtstörungstheoretische Methoden, Renormierung, HQET

Contents

Introduction	13
1 Non-perturbative renormalization in QCD	19
1.1 Renormalization applied to QCD	19
1.2 The running coupling	20
1.3 The running mass	22
1.4 Running of multiplicatively renormalized operators	23
1.5 Examples for a non-perturbative running	23
2 The lattice regularisation of QCD	25
2.1 Lattice gauge theory	26
2.1.1 The pure lattice gauge action	27
2.1.2 Fermion actions	28
2.2 $O(a)$ -Improvement	31
2.2.1 The Symanzik improvement programme	31
2.2.2 Operator improvement	33
2.3 Remarks	35
3 The Schrödinger functional renormalization scheme	37
3.1 Formal definition	37
3.2 The SF renormalized coupling	42
3.3 Quarks in the SF	43
3.4 The Wilson action revisited	45
3.5 Expectation values of composite operators	46
3.6 Fermion correlation functions	47
3.7 Renormalization of correlation function	52
3.8 Effective masses	53
3.9 Running of matrix elements	55
4 Heavy Quark Effective Theory and its non-perturbative matching to QCD	57
4.1 The physical picture	58
4.2 HQET on the lattice	60
4.3 $1/m_b$ -expansion and renormalization of heavy-light currents	62
4.4 Strategy for a non-perturbative matching of QCD and HQET in finite volume	64
4.5 Example: The matching strategy in the static case	66
4.6 Perturbative conversion factors between QCD and HQET	68
5 Setup of lattice simulations	71
5.1 Fixing the finite-volume heavy quark mass	73
5.2 Simulating a doublet of degenerate dynamical quarks	76

6	Computation of improvement and renormalization factors	83
6.1	The PCAC relation	83
6.2	Strategy to compute $b_A - b_P$, b_m and Z	85
6.2.1	Setup of Improvement conditions at constant physics	88
6.3	Results for estimators R_X	91
6.4	Hopping parameters at fixed RGI heavy quark mass $z = LM$	97
7	Non-perturbative tests of HQET	101
7.1	Observables for non-perturbative tests of HQET	101
7.2	Conversion functions between QCD and HQET	104
7.3	Tree-level improvement of test observables	106
7.4	Results	107
7.5	An outlook	114
8	Summary	117
A	The Renormalization group	119
B	Definitions and conventions	122
B.1	The group $SU(N)$	122
B.2	The Dirac algebra	124
B.3	Definitions on the lattice	125
B.3.1	Lattice derivatives	125
B.4	Finite volume continuum gauge fields	128
B.5	Schrödinger functional correlation functions	129
C	Optimization techniques	132
D	Error estimation techniques	140
D.1	The Jackknife method	140
D.2	The Gamma method	141
E	Tables for matching QCD and HQET	143
F	Further figures & tables	146
	References	149
	Acknowledgments	158
	Selbstständigkeitserklärung	161

List of Figures

1	Running coupling and mass in the SF	24
2	Depiction of the lattice quantities	27
3	Sketch of the Schrödinger functional space-time manifold	37
4	Schrödinger functional correlation functions	48
5	Strategy for a non-perturbative matching of QCD and HQET in a small volume	65
6	Schematic view for the continuum limit at L_1	73
7	Normalised distribution of lowest eigenvalues at $L = L_1$	81
8	Time dependence of PCAC masses $L_0 m_{ii}$ after parameter tuning.	90
9	Typical plateau dependence of R_X	91
10	Results for $b_m, b_A - b_P$ and Z	93
11	Residual cutoff dependence of estimators R_X	94
12	$O(a)$ ambiguity: test of universality	99
13	Mass dependence of QCD-HQET conversion functions in the matching scheme	104
14	Error of $L\Gamma_{av}$ vs. bin-size N	107
15	Continuum extrapolations of the spin-averaged mass $L\Gamma_{av}(z, \theta = 0.5)$	108
16	$(1/z)$ -asymptotics of $L_1\Gamma_{av}, R_{spin}, Y_{PS}$ and Y_V in the matching scheme	111
17	$(1/z)$ -asymptotics of $R_{PS/V}$ and $R_{PS/P}$ in the matching scheme	112
18	Unconstrained $(1/z)$ -asymptotics of $L_1\Gamma_{av}, R_{PS/V}$ and $R_{PS/P}$ in the matching scheme	113
19	apeNEXT processor design	134

List of Tables

1	Phenomenology of quarks	14
2	Perturbative values of improvement coefficients b_X	35
3	Heavy-light meson phenomenology	59
4	Simulation parameters for computations in L_0	72
5	Simulation parameters for computations in L_1	72
6	Recursive finite-size scaling of the mass starting from L_0	75
7	Algorithmic parameters in L_0	80
8	Algorithmic parameters in L_1	80
9	HQET simulations in physical volumes L_1 and L_2	82
10	Tuned hopping parameters to set up improvement conditions in $L = L_0$	89
11	Non-perturbative results of improvement estimators R_X	92
12	Simulation parameters for R_X in the charm region	95
13	Results for estimators R_X at $\beta = 5.2$	95
14	Results for estimators R_X used in charm applications	96
15	Summary of interpolated improvement coefficients and renormalization constants	97
16	Hopping parameters at fixed z and various L/a	98
17	$Z_P(g_0, L/a)$ and normalization factors $c(L_1/a)$ at $L = L_1$	102

18	Statistic for measuring test observables	107
19	Continuum extrapolations of QCD test observables (raw data)	109
20	Fit results for the $(1/z)$ -asymptotics in QCD test observables	115
21	A typical run-time profile of HMC code on apeNEXT	133
22	Abbreviations appearing in the static performance analysis	136
23	apeNEXT address generation and memory access instructions	136
24	Microcode example on apeNEXT	137
25	Run-time of Dirac routines	139
26	Run-time of linear algebra routines	139
27	Some scheme independent perturbative coefficients	144
28	Some scheme dependent perturbative coefficients in the $\overline{\text{MS}}$ scheme	144
29	Perturbative coefficients in m_Q/\overline{m}_*	144
30	QCD-HQET matching coefficients of heavy-light currents	145
31	Non-perturbative results for QCD test observables in volume L_1 at $\theta = 0$	146
32	Non-perturbative results for QCD test observables in volume L_1 at $\theta = 0.5$	147
33	Non-perturbative results for QCD test observables in volume L_1 at $\theta = 1$	148

*Likelihood also entails that
the improbable can happen.*

Aristotle

Introduction

During the last decades the physics community developed a view of nature which unifies three of the four fundamental forces in nature over a wide range of energy scales. On one side, the classical Theory of General Relativity (GR) is a very accurate macroscopic description of gravitational interactions which is known to break down at energy scales that are dominated by quantum interactions. The *Standard Model* (SM) on the other hand is by construction a quantum field theory that describes the nature of electromagnetic, weak and strong interactions in the microscopic world of elementary particles. Both are consistent in its own as long as they are applied to problems that appear at interaction energies which are typical to the corresponding interaction length scale.

The most accurate agreement between a theoretical prediction and experiment made so far was achieved within the SM, namely the electron anomalous magnetic moment. To date almost all predictions of the SM do agree well with experiments which thus still confirm its predictability. However, from the constructive point of view there are many open questions like that where all the parameters in the SM come from. Both, GR and the SM are believed to be effective theories (models) which need to be modified at some point. To explore the frontiers of Standard Model predictions, more and more sophisticated high energy experiments have to be performed. The forthcoming experiments where the largest impact is expected, are measurements at the Large Hadron Collider (LHC) at CERN near Geneva. The hope in the physics community to find evidence for (so-called) new physics is enormous. But no insights can be gained, if we lack of accurate theoretical results with controlled error estimates that can be confronted with experiments.

The mathematical structure of the Standard Model is that of a gauged quantum field theory. This means that each gauge subgroup in the SM gives rise to a local gauge field that mediates the corresponding interaction between elementary particles. The gauge groups in the SM,

$$SU_c(3) \times SU_L(2) \times U_Y(1)$$

naturally separate through *spontaneous symmetry breaking* in a part covered by the so-called *electroweak theory* with gauge group $SU_L(2) \times U_Y(1)$ and that of strong interactions. The latter is described by Quantum Chromodynamics (QCD) with non-Abelian gauge group $SU_c(3)$ for three different *colour charges*. The mediators of the colour interaction are called *gluons*. They are vector bosons which couple to each particle that carries colour charge. These particles are the matter

quark label	electric charge	flavour quantum number						mass (MeV)
		I	I_3	S	C	B	T	
u	2/3	1/2	1/2	0	0	0	0	[1.5, 3.0]
d	-1/3	1/2	-1/2	0	0	0	0	[3.0, 7.0]
s	-1/3	0	0	-1	0	0	0	95 ± 25
c	2/3	0	0	0	1	0	0	~ 1270
b	-1/3	0	0	0	0	-1	0	~ 4200
t	2/3	0	0	0	0	0	1	~ 171200

Table 1: Quark flavour phenomenology with flavour quantum numbers: Isospin I with third component I_3 , Strangeness S , Charm C , Bottom B and Top T . Values are taken from Particle Data Group 2008 [1].

constituents, i.e. fermions fields called *quarks*, and the gluons themselves. This self-interaction is due to the non-Abelian structure of the gauge group and thus responsible for *asymptotic freedom* and *confinement*. These special properties make it impossible to just rely on calculations performed as perturbations in the effective couplings constant. While asymptotic freedom still allows to perform perturbative calculations in the (small) coupling constant at high energies or short distances, this becomes impossible due to confinement in the low energy region. Asymptotic freedom is a synonym for quarks and gluons which are weakly coupled at high energies and become non-interacting in the asymptotic limit of arbitrarily large energies due to a vanishing effective coupling. At low energies, quarks are confined into hadronic bound states and cannot be observed as free particles. While several experiments like deep-inelastic scatterings as well as perturbative calculations confirm asymptotic freedom at high energies, confinement at low energies cannot be proved that easily. This is because the description of physical processes by means of Feynman diagrams, as perturbatively derived from the functional integral, become ill-defined. However, due to lattice gauge theory calculations as proposed by K.G. Wilson in 1974 [2] the presence of confinement in the strong-coupling limit is accepted beyond any doubts. Furthermore, a rich phenomenology of hadrons as bound states of constituent quarks and gluons is well described by the quark model. This comprehends that hadrons usually decompose into two colour neutral classes, the mesons and baryons. The former are quark–anti-quark bound states and the latter quark triplets.

In the Standard Model we have six different *flavours* of quarks: up (u), down (d), strange (s), charm (c), bottom (b) and top (t); here and in table 1 listed in order of increasing mass. While the masses of the first three quarks are referred to as light masses with respect to the energy scale $\Lambda \sim 250$ MeV, the others have a large mass. As heaviest quark, the top has a very short lifetime and thus cannot form any bound states that would live long enough to be measured.

Because the dynamics governed by QCD is expected to be simplified in the limit of large quark masses, it is important to gain insights from studying apparently easy systems consisting of one heavy and one light quark before passing into more sophisticated ones. In 1964 J. Cronin & V. Fitch [3] at BNL¹ found evidence of CP -violation in the SM. While their measurements were performed in the light-light system of neutral K-mesons, later it was realized that heavy-light systems which contain a b-quark are much better suited to study this effect and hence to obtain more

¹Brookhaven National Laboratory, USA

accurate results on the experimental side. CP -violation as it is part of the electroweak Lagrangian of the SM is the prerequisite of a matter–antimatter asymmetry as observed in our universe. A theoretical description was given by M. Kobayashi & T. Maskawa in 1973 [4] who extended the work of N. Cabibbo [5] from two to three generations of quarks. Their analysis showed that starting with three generations of quarks, a complex phase appears which causes CP -violation. They predicted a third generation of quarks by introducing CP -violation to the SM.² The study of B-meson physics and its rich phenomenology has seen a growing interest and activity since then. Especially in recent years accurate data became available through a second generation of B-factories, like the BaBar Detector at SLAC³, the Collider Detector at Fermilab (CDF) as well as the Belle experiment at the KEKB factory. Further experimental progress in flavour physics is expected, mainly due to the new LHCb experiment at CERN and the forthcoming SuperKEKB at KEK⁴ [6].

The Kobayashi–Maskawa mechanism of CP -violation is manifest in the CKM quark-mixing matrix which describes flavour changing currents. They appear because eigenstates of the weak interaction are not eigenstates of the quark mass matrix. Entries of the CKM matrix have to fulfil various unitarity constraints in the Standard Model and account for the so-called *unitarity triangles*. Deviations from this description are expected to yield hints for New Physics, i.e. physics that is not incorporated into the SM yet. As mentioned earlier, to confront our picture of the world with experiment accurate theoretical predictions are extremely important. On the theoretical side this involves among hadronic states also matrix elements of the operators in the weak effective Hamiltonian and therefore requires a non-perturbative approach. The most appropriate tool to perform such non-perturbative calculations from first principles is Lattice QCD. One example where lattice QCD computations can provide valuable input is for instance the *B-meson decay constant* F_B :

$$M_B F_B = \langle 0 | (A_R)_0(0) | B \rangle, \quad A_\mu(x) = \bar{\psi}_1(x) \gamma_\mu \gamma_5 \psi_b(x), \quad (0.1)$$

with a zero momentum B-meson state $|B\rangle$ of mass M_B and a heavy-light axial current operator insertion $A_\mu(x)$. Another example is the bag parameter B_B which appears in the mixing matrix element

$$\frac{8}{3} B_B M_B^2 F_B^2 = \langle B^0 | [\bar{\psi}_d(x) \gamma_\mu (1 - \gamma_5) \psi_b(x)] [\bar{\psi}_d(x) \gamma_\mu (1 - \gamma_5) \psi_b(x)] | \bar{B}^0 \rangle. \quad (0.2)$$

and can be estimated if the RHS and $M_B F_B$ are known.

In lattice QCD, light quarks as widely spread objects ($1/m_\pi \simeq 1/(140 \text{ MeV}) \simeq 1.5 \text{ fm}$) demand for a treatment in large-volume simulations. In contrast, heavy quarks are extremely localized objects ($1/m_b \simeq 1/(4 \text{ GeV}) \simeq 0.04 \text{ fm}$) and thus require very fine lattice resolutions $a \ll 1/m_b$. This two scale problem is the main reason that realistic simulations of heavy-light systems involving a b-quark are still impossible. A theoretically clean solution is given by a recourse to the *Heavy Quark Effective Theory (HQET)* [7, 8, 9]. This effective theory starts from the static approximation describing the asymptotics as $m_b \rightarrow \infty$. Corrections to this limit have to be

²This was honoured together with Y. Nambu by the Nobel Prize in 2008.

³Stanford Linear Accelerator Center, USA

⁴高エネルギー加速器研究機構 (High Energy Accelerator Research Organization), Japan

computed by a $1/m_b$ expansion. But there is no lunch for free. By the renormalization properties of HQET, physical quantities as derived from expectation values calculated in the effective theory are affected by power-law divergences in the lattice spacing a when the theory is formulated on the lattice. The static Lagrangian of HQET with bare coupling g_0 produces a divergence due to the static quark self energy E_{self} ,

$$E_{\text{stat}} \sim E_{\text{self}} + \mathcal{O}(a^0), \quad E_{\text{self}} = \frac{1}{a} \{ e^{(1)} g_0^2 + \mathcal{O}(g_0^4) \}, \quad e^{(1)} = \frac{1}{12\pi^2} \times 19.95.$$

Here E_{stat} is the binding energy of the static-light system. Formally, the divergence has to be canceled by a mass counterterm $\delta m \sim -E_{\text{self}} + \mathcal{O}(a^0)$. But they cannot be subtracted perturbatively in a clean way: the continuum limit does not exist unless the theory is renormalized non-perturbatively [10]. A method to overcome these deficiencies that can be applied to general renormalization problems in HQET was proposed in [11]. The power-law divergences can be removed by a *non-perturbative matching procedure of HQET to QCD in a finite volume*. Only the smallness of the physical volume allows to incorporate the b-quark as a relativistic fermion and thus stands at the core of the matching strategy. In principle this strategy can be applied to higher orders in the $1/m_b$ expansion but beyond $\mathcal{O}(1/m_b)$ this becomes practically impossible due to the various terms and effective parameters which have to be taken into account. In [12] for instance, the mass of the b-quark was computed in the quenched approximation of lattice HQET to subleading order and also a strategy to compute the heavy-light decay constant to that order was given [13]. In each case the knowledge of the heavy quark mass dependence of QCD observables is crucial. It motivates to also investigate QCD observables in a small-volume setup for $N_f = 2$ mass degenerate dynamical fermions to overcome the quenched approximation.

In this work we will focus on the setup of the corresponding small-volume simulations in the framework of an on-shell $\mathcal{O}(a)$ improved lattice QCD. Furthermore, we will compute the non-perturbative heavy quark mass dependence of effective heavy-light meson observables in the continuum limit. They allow for quantitative non-perturbative tests of predictions made by HQET in comparison to the large quark mass limit of small-volume QCD. These independent non-perturbative tests may also provide a deeper insight into the feasibility of the effective theory approach. To explicitly make HQET an effective theory of QCD requires matching calculations to express the parameters in the effective theory Lagrangian by those of QCD. While we will explain this matching strategy and how the work presented here fits into it, the full work is out of our scope. In the quenched approximation, such tests were performed and discussed in [14].

Simulating relativistic heavy quarks like the b-quark by means of lattice QCD is still demanding nowadays. The great challenge for realistic lattice QCD simulations is to deal with a multi-scale problem, mainly introduced by the quark masses, non-perturbatively. Furthermore, all systematic errors that may be introduced by some approximation have to be controlled in a theoretically sound way. They read

- *Finite-size effects*

To extract physical matrix elements by lattice simulations one needs to perform computations in a physically large volume in order to avoid (or at least suppress) effects introduced

by simulating QCD in a finite volume. We only compute effective quantities in small volume QCD. Thus we are not immediately faced with this issue. However, the heavy-light observables considered here are in a consistent way connected to observables in large volume computations as will become clear in section 4.

- *Lattice spacing/cutoff effects*

In general one expects to find discretization errors that are linear in the lattice spacing a . In physical units a has to be small enough compared to *all* other scales that are involved in the computation. The lattice artifacts are inherent in the discretization and need to be estimated. To this end and if even possible, one varies a over a wide range while keeping the physical parameters of the theory fixed. This allows to estimate the cutoff-dependence in lattice quantities and hence the rate at which the continuum limit $a \rightarrow 0^+$ is reached. Such lattice artifacts can become quite large and one is faced with the problem of finding a better discretization of the theory with improved convergence properties. A systematic approach to cancel lattice artifacts non-perturbatively was given by Symanzik and is discussed in more detail in section 2.2.1. It allows to improve the full lattice theory order by order in a non-perturbatively. But this is a very complex task and one usually restricts to remove only the leading (linear) lattice artifacts. This is called $O(a)$ -improvement and done in the first part of this work for some selected quantities.

- *Heavy quark mass effects*

For our simulation of relativistic heavy quarks, mass dependent cutoff effects at finite lattice spacing can become severe and at worst also spoil the $O(a)$ -improvement program mentioned above. This problem is partially lifted by choosing a very fine lattice resolution which is only possible in a small volume simulation. However, with increasing mass this problem arises again and thus has to be monitored.

- *Non-perturbative vs. perturbative matching*

In principle one could employ perturbation theory in the matching step which connects HQET and QCD. Owing to the difficulty of a reliable error estimation it may be hard to disentangle *and* quantify deviations coming from higher orders or non-perturbative effects in particular.

This thesis is structured as follows: First, I summarize features of a non-perturbative renormalization as given by the renormalization group approach to QCD quite generally. Special emphasis is put to the renormalization scale running of physical observables. Especially that of the effective coupling and masses is fundamental to the strategy follow throughout this work. In section 2 I discuss the lattice discretization as our non-perturbative regulator of QCD. After introducing some general terms and definitions I will focus on $O(a)$ -improvement of the lattice theory. After that, in section 3 I introduce the Schrödinger functional renormalization scheme (SF) as our basic tool for solving renormalization problems non-perturbatively on the lattice. Because it is regularisation independent this will mainly be done in the continuum notation. Before defining some relevant fermion correlation functions in this scheme I describe how to set up the SF as an $O(a)$ improved

lattice theory and give the definition for expectation values of composite operators. This section ends with an explanation how such operators get renormalized.

By then the discussion was quite general or restricted to QCD, so I focus on HQET in section 4. After some general properties were discussed, I give a definition of HQET to subleading order in the lattice regularisation and motivate the necessity of a non-perturbative renormalization. The previously defined heavy-light correlation functions are expanded to subleading order in HQET. Additional parameters that appear in the effective theory to this order are described. Then I present the strategy to match HQET non-perturbatively to QCD which is still under investigation in our collaborative effort. This matching strategy relies on various different lattice QCD and HQET simulations. Thus they have to be set up very carefully. The involved parameter tuning and most of the technical details to setup these lattice computations with $N_f = 2$ dynamical massless quarks is discussed in section 5.

In section 6 I explain how some special improvement coefficients and renormalization constants can be computed non-perturbatively. Then I present the results obtained on some of the lattices introduced in the foregoing section. The results are a prerequisite for the computation of relativistic heavy-light meson observables that allow to test predictions made by HQET. In section 7 we finally define appropriate QCD observables to confront with predictions of HQET. These observables are computed at various values of the heavy quark mass and thus allow for an interpolation from the charm to the bottom quark mass region. Thus estimators for $(1/m)$ -corrections to the leading order of HQET are obtained. Furthermore, I present some forthcoming results and conclude in section 8.

1 Non-perturbative renormalization in QCD

We directly start with the renormalization of the parameters of QCD. Special emphasis is put to the renormalization scale dependence of effective parameters of the theory as well as of local composite operators which get multiplicatively renormalized. In later sections we will always come back to the basics introduced here. Beside the standard textbooks [143, 144] also [146, 147] provide a good introduction to the subject of Renormalization and Quantum Field Theory (QFT) in general.

1.1 Renormalization applied to QCD

The general Lagrangian density of QCD in a convenient normalization with n_f flavours of quarks ψ_i of mass m_i reads

$$\mathcal{L} = -\frac{1}{2g_0^2} \text{Tr} \{F^{\mu\nu}(x)F_{\mu\nu}(x)\} + \sum_{i=1}^{n_f} \bar{\psi}_i(x) [\gamma_\mu(\partial^\mu + iA^\mu(x)) + m_{0,i}] \psi_i(x). \quad (1.1)$$

It connects the gauge field A_μ to the fermionic degrees of freedom, $\bar{\psi}_i, \psi_i$ through minimal coupling. The (local) gauge symmetry group is the non-Abelian colour group $SU(3)$, see appendix B.1 for details. The quark and anti-quark fields belong to the fundamental 3 and $\bar{3}$ representation, respectively, whereas the gauge field belongs to the adjoint 8 representation. QCD has $n_f + 1$ free parameters, the bare coupling g_0 and the bare quark masses $\{m_{0,i} | i = 1, \dots, n_f\}$.

A special and very convenient class of renormalization schemes (RS) are the *mass independent* ones in which the fields and parameters of the Lagrangian renormalize multiplicatively and the counterterms⁵ do not depend on the mass. Without further notice we always assume a mass independent scheme in the following, where the bare parameters of the Lagrangian g_0 and m_0 , get renormalized multiplicatively by

$$\bar{g}(\mu) = Z_g(\mu, g_0)g_0, \quad \bar{m}(\mu) = Z_m(\mu, g_0)m_0. \quad (1.2)$$

But physical measurable (renormalized) quantities Q only depend on the renormalized parameters of the theory, \bar{g} and \bar{m} , and not on the arbitrary value of the renormalization point μ . Thus they have to be *renormalization scale invariant*, expressed by

$$\frac{d}{d\mu} Q(\{\bar{g}(\mu)\}, \{\bar{m}(\mu)\}, \mu) \equiv 0. \quad (1.3)$$

From this invariance condition a partial differential equation can be deduced, the so-called *Callan-Symanzik* or *Renormalization Group Equation (RGE)*,

$$\left(\mu \frac{\partial}{\partial \mu} + \beta(\bar{g}) \frac{\partial}{\partial \bar{g}} + \tau(\bar{g}) \bar{m} \frac{\partial}{\partial \bar{m}} - n \gamma_Q(\bar{g}) \right) Q(\{\bar{g}(\mu)\}, \{\bar{m}(\mu)\}, \mu) \equiv 0. \quad (1.4)$$

⁵Counterterms Z are dimensionless and should have a finite limit as $m \rightarrow 0$ in a mass-independent scheme. PT shows that the only way of Z to depend on m would be in powers of $\ln(m/\mu)$ which is divergent as $m \rightarrow 0$ and therefore is ruled out.

In a mass independent RS, the *Gell-Mann–Low function* β , the *mass anomalous dimension* τ and the *field anomalous dimension* γ_Q only depend on the renormalized coupling \bar{g} . With the deduction given in appendix A, their definition reads

$$\beta(\bar{g}) = \mu \frac{\partial \bar{g}}{\partial \mu}, \quad \tau(\bar{g}) = \frac{\mu}{\bar{m}} \frac{\partial \bar{m}}{\partial \mu}, \quad \gamma_Q(\bar{g}) = \frac{1}{2} \frac{\mu}{Z} \frac{\partial Z}{\partial \mu}. \quad (1.5)$$

Their formal definition does not depend on perturbation theory. Obviously, the behaviour of the parameters of the theory, \bar{g} and \bar{m} , under a rescaling is encrypted in the β and τ function, respectively. In order to solve the renormalization group equations at least to some extend, one needs to choose a renormalization scheme. This in turn makes the anomalous dimensions dependent on that scheme. The β function in QCD describes (in a given scheme) the scale dependence of the strong coupling constant, which is the fundamental expansion parameter in perturbative calculations. Hence, it is the most important object and should be known beyond PT.

The most commonly used renormalization scheme in QCD, the modified minimal subtraction ($\overline{\text{MS}}$) scheme, belongs to the class of mass-independent renormalization schemes and relies on *dimensional regularisation* [15]. Almost all physical results are expressed in that scheme, usually at the fixed-reference scale $\mu_{\text{ref}} = m_Z$, the mass of the Z boson. The current world average as given by the Particle Data Group [1] is

$$\alpha_s^{\overline{\text{MS}}}(m_Z) = 0.1176(20), \quad m_Z \equiv \bar{m}_{\overline{\text{MS}}}(m_Z) = 91.1876(21) \text{ GeV}, \quad (1.6)$$

obtained from various experimental results as well as lattice QCD input. Beside its direct relation to a physical particle, this scale is high enough to safely apply a perturbative renormalization scheme.

1.2 The running coupling

As we are interested in expressing physical results in terms of the renormalized parameters of the theory, we have to solve the equations in (1.5). Formally that is an easy task and for the coupling one gets from its definition

$$\beta(\bar{g}) = \mu \frac{\partial \bar{g}}{\partial \mu} = \frac{\partial \bar{g}}{\partial \ln \mu}, \quad (1.7)$$

by a standard separation of variables the solution

$$\int_{\mu_a}^{\mu_b} d(\ln \mu) = \int_{\bar{g}(\mu_a)}^{\bar{g}(\mu_b)} \frac{dg}{\beta(g)} \quad \Leftrightarrow \quad \ln \frac{\mu_b}{\mu_a} = \int_{\bar{g}(\mu_a)}^{\bar{g}(\mu_b)} \frac{dg}{\beta(g)}. \quad (1.8)$$

This is well-defined too as long as the β function has no zero in the integration domain. A zero in the β function corresponds to a fixed point in the renormalized coupling with important physical consequences. Apart from that, one integration constant is required for the initial value problem described by the partial differential equation of first order, eq. (1.7). This is the first time where physical input from the theory under consideration, QCD, is needed.

Nearly 40 years ago, during the search for a model to describe strong interacting processes the calculation of the one-loop β function has led to the discovery of *asymptotic freedom*⁶ and to the establishment of QCD as the theory of strong interactions [16, 17]. Later the two-loop and three-loop β function was derived in [18] and [19] in the $\overline{\text{MS}}$ -scheme, respectively. The latest analytic computation of the β function was done in the $\overline{\text{MS}}$ -scheme up to four-loops, [20].

The main disadvantage of those computations – beside the fact that they become more and more complicated from one order to another – is that they are inherently perturbative. One computes the coefficients in the following expansion of the (non-perturbative) QCD β function,

$$\beta(\bar{g}) = -\bar{g}^3(b_0 + b_1\bar{g}^2 + b_2\bar{g}^4 + b_3\bar{g}^6 + \dots), \quad (1.9)$$

which is obviously defined for not too large renormalized couplings \bar{g} only. In mass-independent renormalization schemes the two leading coefficients in this expansions,

$$b_0 = (11 - \frac{2}{3}N_f)/(4\pi)^2, \quad b_1 = (102 - \frac{38}{3}N_f)/(4\pi)^4, \quad (1.10)$$

are *universal*, i.e. scheme independent, whereas all higher order coefficients depend on the chosen renormalization scheme. Nevertheless, at very high energies where the QCD coupling is small, perturbation theory is a trustworthy tool as it was in QED before the emergence of QCD. In fact, any non-perturbative computation has to coincide at high energies with perturbation theory due to asymptotic freedom. As can be easily seen in eq. (1.9), as the coupling approaches zero in the high energy limit $\mu \rightarrow \infty$ the β function tends to zero, *the ultra-violet fixed point*, and loses their scheme dependence as the leading universal coefficients become more and more dominant.

Although eq. (1.8) seems to diverge in this limit, one can write down a solution which reintroduces an artificial scheme dependence. The (exact) solution is often stated as the integration constant one has to introduce

$$\Lambda = \mu [b_0 \bar{g}^2(\mu)]^{-b_1/(2b_0^2)} e^{-1/[2b_0 \bar{g}^2(\mu)]} \exp \left\{ - \int_0^{\bar{g}(\mu)} dg \left[\frac{1}{\beta(g)} + \frac{1}{b_0 g^3} - \frac{b_1}{b_0^2 g} \right] \right\}. \quad (1.11)$$

This constant, also known as Λ_{QCD} , depends on the number of flavours and the scheme it is computed in. We provide a derivation in appendix A. It is a *Renormalization Group Invariant* (RGI) as it does not depend on the renormalization scale μ . However, different schemes can be related *exactly* by

$$\Lambda_b/\Lambda_a = \exp \left\{ c_1^{(ba)}/2b_0 \right\}, \quad \bar{g}_b^2(\mu) = \bar{g}_a^2(\mu) + c_1^{(ba)}[\bar{g}_a^2(\mu)]^2 + \dots, \quad (1.12)$$

which connects the Lambda parameter of scheme a and scheme b , c.f. [21]. One just has to know the one-loop coefficient $c_1^{(ab)}$ that relates the running couplings at the same scale μ . The connection of the Λ parameters in QCD with two dynamical flavours between the $\overline{\text{MS}}$ and the Schrödinger

⁶... and cause for the Nobel Prize in Physics 2004

functional scheme for instance is known very accurately from [22],

$$\Lambda_{\overline{\text{MS}}} = 2.382035(3)\Lambda_{\text{SF}}. \quad (1.13)$$

Λ depends on the number of active flavours and characterises the scale dependence of the renormalized coupling in a given scheme. A remarkable thing to note about the renormalization procedure is that even in a theory without a dimensionful parameter – as QCD with massless quarks where the remaining parameter is the dimensionless coupling in the bare Lagrangian – a scale appears. This is known by *dimensional transmutation* and reflects the appearance of the scale parameter Λ . Beside the bare coupling we also have to take care of the other parameters in the Lagrangian, the masses.

1.3 The running mass

Again starting from the basic equation for the corresponding anomalous dimension,

$$\tau(\bar{g}) = \frac{\mu}{\bar{m}} \frac{\partial \bar{m}}{\partial \mu} = \frac{\partial \ln \bar{m}}{\partial \ln \mu}, \quad (1.14)$$

one can formally integrate to end up with

$$\ln \frac{\bar{m}(\mu_b)}{\bar{m}(\mu_a)} = \int_{\bar{g}(\mu_a)}^{\bar{g}(\mu_b)} d\bar{g} \frac{\tau(\bar{g})}{\beta(\bar{g})}. \quad (1.15)$$

As in the case of the running coupling the β function appears in the denominator and causes problems in the limit $\bar{g} \rightarrow 0$, because the anomalous dimension of the mass approaches the limit only quadratically in \bar{g} ,

$$\tau(\bar{g}) = -\bar{g}^2(d_0 + d_1\bar{g}^2 + d_2\bar{g}^4 + \dots), \quad d_0 = 8/(4\pi)^2. \quad (1.16)$$

Here only the first coefficient d_0 is scheme independent. The integration constant in the UV limit is now given by the RGI quark mass

$$M = \bar{m}(\mu) [2b_0\bar{g}^2(\mu)]^{-d_0/(2b_0)} \exp \left\{ - \int_0^{\bar{g}(\mu)} d\bar{g} \left[\frac{\tau(\bar{g})}{\beta(\bar{g})} - \frac{d_0}{b_0\bar{g}} \right] \right\}. \quad (1.17)$$

It is a scale and also a scheme independent quantity and exists for each quark flavour separately. Due to the universality of the RGI quark masses and the Λ -parameter, it seems more naturally to take Λ and $\{M_k | k = 1, \dots, n_f\}$ as the fundamental parameters of QCD. Whereas it exists an accepted normalisation convention for the Λ -parameter, there are different conventions in case of the masses M . The one applied here goes back to the conventions used by Gasser and Leutwyler, [23, 24, 25]. Beside the running of renormalized parameters of a theory, some renormalized matrix elements that are not protected against renormalization are also scale dependent. If their renormalization scale dependence is multiplicatively the same strategy to compute the corresponding RGI operator applies.

1.4 Running of multiplicatively renormalized operators

For matrix elements $\Phi_0 = \langle \text{out} | \mathcal{O} | \text{in} \rangle$ that include insertions of local (composite) operators \mathcal{O} an additional inhomogeneity arises in the renormalization group equations. Accordingly a renormalization group invariant matrix element can be introduced by

$$\Phi_{\text{RGI}} = \Phi(\mu) [2b_0 \bar{g}^2(\mu)]^{-\gamma_0/(2b_0)} \exp \left\{ - \int_0^{\bar{g}(\mu)} dg \left[\frac{\gamma(g)}{\beta(g)} - \frac{\gamma_0}{b_0 g} \right] \right\}, \quad (1.18)$$

with

$$\Phi(\mu) = Z_\Phi(\mu, g_0) \Phi_0(g_0), \quad (1.19)$$

$$\gamma(\bar{g}) \equiv \frac{\partial \ln \Phi(\mu)}{\partial \ln \mu}, \quad \gamma(\bar{g}) = -\bar{g}^2 (\gamma_0 + \bar{g}^2 \gamma_1 + \dots). \quad (1.20)$$

No matter what the composite operator is, the first coefficient in the perturbative expansion has to be universal to ensure a scheme independence of the matrix element at the fixed point. A typical example is the static axial current in Heavy Quark Effective Theory (HQET) [26, 155].

1.5 Examples for a non-perturbative running

In perturbation theory the two-loop asymptotics for $\mu \rightarrow \infty$ of the renormalized strong coupling reads

$$\alpha_s(x) \equiv \frac{\bar{g}^2(x)}{4\pi} \simeq \frac{1}{4\pi b_0 \ln x} \left[1 - \frac{b_1}{b_0^2} \frac{\ln \ln x}{\ln x} \right] + \mathcal{O} \left(\frac{\ln^2 \ln x}{\ln^3 x} \right), \quad x = \left(\frac{\mu}{\Lambda} \right)^2. \quad (1.21)$$

This is again universal in the sense that a finite change in the scale – no matter if it arises from a change of the renormalization scheme by $\mu \rightarrow \mu' = \mu \Lambda / \Lambda'$ or not – does not affect the value of $\alpha_s(x)$ in this limit. Furthermore to be within a certain accuracy for decreasing μ one has to add higher order terms to the expansion which then becomes apparently scheme dependent due to the higher order coefficients $\{b_2, \dots\}$. In [27] various perturbative and non-perturbative techniques to determine the running coupling in QCD are discussed. One of them relies on a finite-size scaling method which allows to study the scale dependence of different quantities non-perturbatively in QCD or other quantum field theories. The first step to solve the renormalization group equations is to determine the scale dependence of the coupling. Thus, the β -function is known non-perturbatively and required to determine the running of the mass in a secondary step. In figure 1 we show the running of the strong coupling and mass as carried out in the Schrödinger functional renormalization scheme with $N_f = 2$ dynamical flavours of massless quarks. The non-perturbative values (points) are taken from [28] for the coupling and from [29] for the mass. For comparison we added the two-loop estimate (1.21) as dashed line in the upper plot for the coupling. A corresponding one-loop estimate of the running mass, which follows from eq. (1.17)

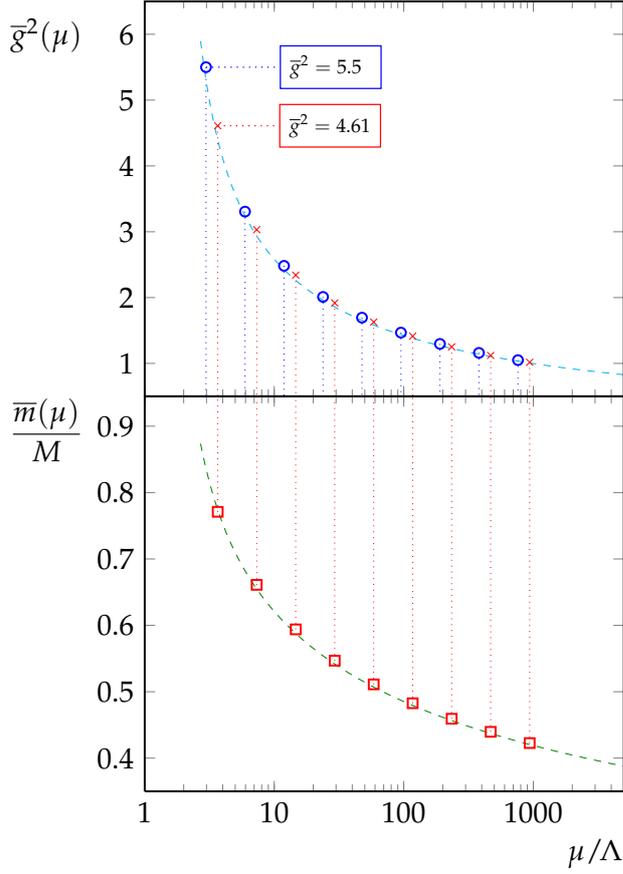


Figure 1: By applying a finite step scaling technique the running of the strong coupling (*top*) and the quark mass (*bottom*) was obtained non-perturbatively in the Schrödinger functional scheme with $N_f = 2$ dynamical quark flavours.

From the running of the coupling the (RGI) Λ parameter was obtained by implicitly fixing the low-energy scale at $\bar{g}^2 = 5.5$. For practical reasons, the RGI mass M and other scale dependent matrix elements are subsequently evaluated by fixing the coupling at a scale corresponding to $\bar{g}^2 = 4.61$.

to

$$\frac{M}{\bar{m}(\mu)} = [2b_0\bar{g}^2(\mu)]^{-d_0/(2b_0)} \quad (1.22)$$

is shown as dashed line in the lower plot together with the non-perturbative estimates. For the perturbative running, eq. (1.21) was used for the coupling that appears in (1.22). A comparison of the NP running of the strong coupling [28] to the PT running using the highest available order in β has shown significant deviations above $\bar{g}^2 = 2.51$. Our work partially relies on the results of these non-perturbative estimations and we will come back to it at the appropriate places.

A general framework to carry out such calculations is provided by the *lattice regularization*, which is yet the only known regularization where computations can be performed from first principles non-perturbatively. This method is deeply connected to path integral formulation in QFT. The regularized functional integral – by its relation to the partition function in statistical mechanics – is the starting point to carry out non-perturbative computations by means of stochastic methods.

2 The lattice regularisation of QCD

Before we present the basic concepts and some obstacles that appear when regularising Quantum Chromodynamics by a specific lattice approximation, I want to emphasize some important points. A general mathematical foundation is the *axiomatic* or *constructive quantum field theory* and goes back to Wightman [30]. Wightman and others introduced a set of axioms that describe vacuum expectation values of products of fields which are sufficient to reconstruct the corresponding quantum field theory in Minkowski space. After the formulation of equivalent axioms on Euclidean Green's functions by Osterwalder & Schrader [31, 32], Euclidean fields became the fundamental tool to study a Minkowski space field theory. Thus it became feasible to link relativistic QFTs in Feynman's functional integral representation of expectation values to the partition function of a corresponding classical system of statistical mechanics. In this approach the Feynman–Kac formula [33] allows to solve certain partial differential equations by simulating random paths of stochastic processes. The functional integral is re-interpreted – after analytic continuation to imaginary time – as average over the configuration space of Euclidean field configurations, weighted by a Boltzmann probability e^{-S_E} of the classical (Euclidean) action S_E . This average gives the desired expectation values and is computed by means of quasi Monte-Carlo simulations that generate ensembles of field configuration with respect to the given probability distribution driven by the field equations, see section 5.2.

To describe colour confinement of quarks, Wilson invented the notion of a lattice gauge field theory in 1974 [2, 34]. However, one of the Osterwalder–Schrader conditions, *reflection positivity*, is crucial for the existence of a positive-energy Hamiltonian (spectrum). To not alter the physical content of the corresponding lattice theory it should hold there as well or at least be recovered when removing the regulator. Using the *transfer matrix formalism* already known in quantum mechanics, Lüscher [35] showed in 1977 that the lattice gauge field theory in Wilson's original setting leads to a Hamiltonian with real eigenvalues only. But it took ten more years till the consistency of his lattice regularisation was proven rigorously by Reisz, [36, 37]. But what is a lattice regularisation?

The four-dimensional Minkowski spacetime continuum gets replaced by a Euclidean hyper-cubic discrete grid of points with finite distance a , the *lattice spacing*. The regularisation is given by a^{-1} which serves as momentum cutoff that – in the usual way – modifies the theory at short distances and renders UV divergences finite. Obviously, the introduction of a *fundamental length scale* breaks rotational symmetry which gets replaced by a hyper-cubic symmetry (c.f. [38]). The original symmetry is recovered in the *continuum limit* $a \rightarrow 0$, i.e. when the cutoff is removed. For practical reasons one usually considers a compact space-time, i.e. one restricts the setup of a lattice regularised field theory to a finite volume $V = T \times L^3$, with extent T in time and L in the spatial directions. Now the Feynman path integral is a product integral over a countable, finite number of degrees of freedom and hence well-defined. In this setup one usually treats the finite volume as a source of error which has to be taken into account [39, 40]. Another way to look at this is to immediately start from an effective finite volume theory and regularise it. This will be our point of view in section 3. However, the regularisation parameter a is believed to smoothly encode the difference of the regularised (effective) quantum field theory from the original one,

continuum QCD. Thus, a is a new degree of freedom and introduces a kind of arbitrariness to the system. To define a theory on the lattice completely, one has to specify lattice approximants for every quantity the theory is made of. But this mapping is not unique, each quantity – a derivative or the action itself for instance – has several lattice approximants. As a fundamental principle it is mandatory that they share the same symmetry properties as their counterpart in the continuum. The converse of that diversity is *universality* which just states that different lattice discretisations finally lead to the same continuum theory if they are a member of the same universality class. Like dimensional regularisation, the lattice regularisation does not break *gauge invariance* if the prescription for discretising the action belonging to the Lagrangian density under consideration is done carefully. For a long time lattice QCD was faced with explicit chiral symmetry breaking at finite lattice spacing. This was a major problem because in the beginning the only way to get a chirally symmetric theory on the lattice resulted in an unphysical continuum spectrum. But with increasing understanding of the mechanism behind it, one was able to formulate theories on the lattice which break chiral symmetry only in a minimal way. In the following I refer to the standard lattice field theory textbooks [150, 151] and the one devoted only to lattice QCD [152].

2.1 Lattice gauge theory

To achieve a lattice discretisation of QCD one has to restrict the quark and anti-quark fields ψ and $\bar{\psi}$ to the lattice sites $x_\mu = an_\mu$, $n_\mu \in \mathbb{Z}^4$, $\mu = 1, \dots, 4$. Unit vectors in μ -direction are denoted by $\hat{\mu}$. Like in the continuum fermion fields carry colour, Dirac and flavour indices. The gauge field is approximated by a gauge link variable $U(x, \mu) \in \text{SU}(3)$, pointing from site x to site $x + a\hat{\mu}$. The corresponding (fundamental) directed link is an ordered pair of points $b = \langle x + a\hat{\mu}, x \rangle \equiv (x, \mu)$. The U act as *parallel transporters* on the gauge group $\text{SU}(3)$ between fermion fields defined at different points on the lattice. They maintain gauge-covariance by the matrix representation

$$U(x, \mu) \equiv U(x + a\hat{\mu}, x) = \mathcal{T} \exp \left\{ a \int_0^1 dt A_\mu(x + (1-t)a\hat{\mu}) \right\} \quad (2.1)$$

$$= \mathbb{1} + aA_\mu(x) + \mathcal{O}(a^2), \quad (2.2)$$

with respect to the local vector field (*connection*) $A_\mu(x)$. However, on the lattice the parallel transporters have to be considered as the fundamental gauge field representative and not the vector field itself. The symbol \mathcal{T} preserves the ordering of the exponential along the path, needed in case of non-Abelian gauge groups, see [145]. A field $\psi(y) = U(y, x)\psi(x)$ is the parallel transported version of $\psi(x)$ at point y . The link variables obey the composition law

$$U(\mathcal{C}) = U(b_n) \cdots U(b_1) \equiv \prod_{b \in \mathcal{C}} U(b), \quad \mathcal{C} = b_n \circ \dots \circ b_1, \quad (2.3)$$

for arbitrary paths \mathcal{C} on the lattice, built from fundamental links. Hence, the point y is not restricted to be a next neighbour of x . The following is notationally convenient,

$$U(x, \mu) \equiv U_{x\mu} \equiv U_{x+a\hat{\mu},x} \equiv U(x + a\hat{\mu}, x), \quad (2.4)$$

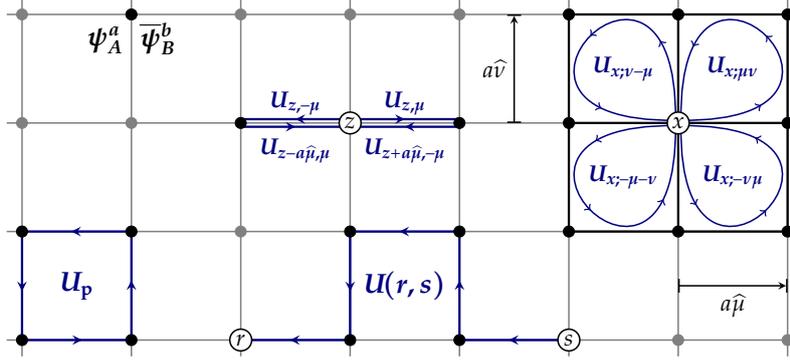


Figure 2: Lattice representatives of different quantities in a region of the (μ, ν) -plane. For each fermion quark flavour the lattice sites carry a quark ψ and anti-quark field $\bar{\psi}$ with colour and Dirac index. Each site connecting link variable $U_{z, \mu}$ represents a parallel transporter and as shown $(U_{z, \mu})^{-1} \equiv U_{z+a\hat{\mu}, -\mu}$ has to hold for uniqueness reasons. A plaquette U_p is made of four link variables and represents the smallest (gauge-invariant) closed loop on the lattice. The clover term on the other hand (*upper right*) is built from four plaquettes, corresponding to eq. (2.35). The blue paths there just show the orientation of the involved plaquettes.

with $U_{x+a\hat{\mu}, x}^{-1} = U_{x, x+a\hat{\mu}}$, i.e. $U^{-1} \equiv U^\dagger$. For a visual depiction see fig. 2. The smallest set of all gauge links forms the lattice gauge field. A local gauge transformation Λ on the lattice is given by

$$\psi(x) \rightarrow \Lambda(x)\psi(x), \quad U(x, \mu) \rightarrow \Lambda(x+a\hat{\mu})U(x, \mu)\Lambda^{-1}(x). \quad (2.5)$$

Derivatives become finite difference operator as defined in appendix B.3.1. On the lattice we distinguish the forward, backward and symmetric partial derivative, denoted by ∂_μ , ∂_μ^* and $\tilde{\partial}_\mu$, respectively. The corresponding gauge-covariant lattice derivatives ∇_μ , ∇_μ^* and $\tilde{\nabla}_\mu$ contain gauge-links and do not commute. All those derivatives only include nearest neighbours, which is also true for the second order derivatives $\partial_\mu\partial_\mu^*$ and $\nabla_\mu\nabla_\mu^*$. Furthermore, d -dimensional integrals $\int d^d x$ reduce to finite sums over smallest hyper-cube volumes containing the lattice point x , i.e. $a^d \sum_x$. With this at hand one is able to formulate a lattice gauge and fermion action for QCD à la Wilson, [2].

2.1.1 The pure lattice gauge action

What is the lattice representative of the gauge field strength tensor $F_{\mu\nu}$? Again, one could start from the continuum expression of the (Euclidean) parallel transporters,

$$U(y, x) = \mathcal{T} \exp \left\{ \int_x^y dz_\mu A_\mu(z) \right\}. \quad (2.6)$$

Stokes theorem tells us that we can measure the *curvature* of the gauge field, i.e. the gauge field strength, by parallel transport along a closed (non-vanishing) loop $\mathcal{C}_{x,x}$ in accordance with

$$U(\mathcal{C}_{x,x}) = \mathcal{T} \exp \left\{ \oint_x dz_\mu A_\mu(z) \right\} = \mathcal{T} \exp \left\{ \int_{A_C} d\sigma_{\mu\nu} F_{\mu\nu} \right\} \sim \mathbb{1} + F_{\mu\nu} d\sigma_{\mu\nu} + \dots \quad (2.7)$$

where A_C is the area enclosed by the loop $\mathcal{C}_{x,x}$. The smallest loop on the lattice is called a *plaquette*, $p = (x; \mu, \nu)$, that is an area element of size a^2 oriented by the direction the boundary

loop is passed through. Thus, a gauge-invariant lattice action that reduces to the continuum Yang–Mills action can be build from products of gauge links, the *plaquette variables*

$$U_p \equiv U_{x,\mu\nu} \equiv U(x, x + a\hat{\nu})U(x + a\hat{\nu}, x + a\hat{\mu} + a\hat{\nu})U(x + a\hat{\mu} + a\hat{\nu}, x + a\hat{\mu})U(x + a\hat{\mu}, x) . \quad (2.8)$$

Then the *Wilson action* is a sum of traces over all plaquette variables with fixed orientation,

$$S_G[U] = \sum_p S_p(U_p) , \quad \sum_p \equiv \sum_x \sum_{1 \leq \mu < \nu \leq 4} = \frac{1}{2} \sum_{x,\mu,\nu} , \quad (2.9a)$$

$$S_p(U_p) = \beta \left\{ 1 - \frac{1}{2 \text{Tr} \mathbb{1}} (\text{Tr} U_p + \text{Tr} U_p^{-1}) \right\} = \beta \left\{ 1 - \frac{1}{N} \text{Re Tr} U_p \right\} , \quad (2.9b)$$

where the second equal sign in (2.9b) holds for gauge group $SU(N)$. The constant β is a convenient normalization constant and needs to be determined in order to recover the right continuum Yang-Mills action. An explicit calculation shows

$$S_G = -\frac{\beta}{4N} a^4 \sum_x \text{Tr} F_{\mu\nu}(x) F^{\mu\nu}(x) + O(a^5) , \quad (2.10)$$

and therefore

$$\beta = 2N/g_0^2 , \quad \text{i.e.} \quad \beta = 6/g_0^2 , \quad (2.11)$$

in case of the $SU(3)$ colour gauge group of QCD.

2.1.2 Fermion actions

Focussing on a one flavour theory for the moment, a *naive* discretization of the fermionic part of the QCD action is given by

$$S_F[U, \bar{\psi}, \psi] = a^4 \sum_x \bar{\psi}(x) (D + m_0) \psi(x) , \quad D \equiv \gamma_\mu \tilde{\nabla}_\mu = \frac{1}{2} \sum_{\mu=0}^3 \gamma_\mu (\nabla_\mu^* + \nabla_\mu) . \quad (2.12)$$

m_0 is the bare quark mass in the original Lagrangian and D the lattice Dirac operator using the symmetric gauge covariant derivative and the Euclidean Dirac matrices as given in appendix B.2. This straight-forward discretisation suffers from the so-called *fermion doubling problem*, which is easy explained but deeply connected to the problem of formulating chirally invariant theories on the lattice. To work with free fermions we switch off the interaction with the gauge field by setting $\tilde{\nabla}_\mu \rightarrow \tilde{\partial}_\mu$. Now it is an easy task to write down the *free fermion propagator in momentum space* for a lattice theory on a 4-torus,

$$S(p) = 1 / \left\{ m_0 + \frac{i}{a} \sum_\mu \gamma_\mu \sin ap_\mu \right\} \quad (2.13)$$

$$= \left\{ m_0 - \frac{i}{a} \sum_\mu \gamma_\mu \sin ap_\mu \right\} / \left\{ m_0^2 + a^{-2} \sum_\mu \sin^2 ap_\mu \right\} . \quad (2.14)$$

The momenta $-\pi \leq ap_\mu \leq \pi$ are now quantized with respect to $2\pi/L$ and restricted to the first Brillouin zone. From the denominator it follows that apart from one physical pole, 15 additional singularities arise when taking the continuum limit. The points in momentum space where those poles are located for $m_0 = 0$, is given by the set

$$\Pi = \{ \pi_{ijkl} = (i, j, k, l)\pi \mid i, j, k, l \in \{0, 1\} \} . \quad (2.15)$$

In d space-time dimensions the degeneration is 2^d . The reason for this degeneration is an exact global symmetry of the naive fermion action as shown in [41]. A deeper understanding was achieved with the no-go theorem of Nielsen and Ninomiya, [42, 43, 44]. It states that on the lattice there is always an equal number of left- and right-handed fermions. Their difference is classically given by a conserved quantum number, the *axial charge* which vanishes in that case. It is not conserved any more in the corresponding quantum field theory. The reason that classical symmetries – like chiral symmetry in QCD – are violated at the quantum level is the emergence of *anomalies* and related to topological properties of the theory, [148].

However, to get rid of the unphysical additional fermion modes (doublers), one has to lift the degeneracy. As in quantum mechanics this can be done by breaking the corresponding symmetry. Again, Wilson was the first who did this by adding an irrelevant (Wilson-) term to the naive fermion action. Irrelevant terms in the action vanish in the continuum limit as they are proportional to some positive power of a . This is the aforementioned ambiguity in the choice of a lattice action. Since the Wilson-term gives the doublers a mass proportional to the cutoff, they disappear from the spectrum when the continuum limit (CL) is taken and chiral symmetry restored.

Wilson fermions

The chiral symmetry breaking term that Wilson added to the naive fermion action is $-\frac{1}{2}ar\nabla_\mu^*\nabla_\mu$ with some parameter $0 < r \leq 1$. This term respects the symmetry of the action and the most natural choice is to set $r \equiv 1$ which for instance respects *reflection positivity* even at finite lattice spacing in contrast to $r \neq 1$. In our strategy to keep things simple we therefore set the parameter to its standard value, $r = 1$, and never mention it again. The Wilson–Dirac operator now reads

$$D_W \equiv \frac{1}{2} \sum_{\mu=0}^3 \{ \gamma_\mu (\nabla_\mu^* + \nabla_\mu) - a \nabla_\mu^* \nabla_\mu \} , \quad (2.16)$$

and leads with a non-vanishing mass term to the following free propagator

$$S(p) = 1 / \left\{ m_0 + a^{-1} \sum_\mu [i\gamma_\mu \sin(ap_\mu) - (\cos(ap_\mu) - 1)] \right\} \quad (2.17)$$

$$= \left\{ m_0 - a^{-1} \sum_\mu [i\gamma_\mu \sin(ap_\mu) + (\cos(ap_\mu) - 1)] \right\} / \left\{ a^{-2} \sum_\mu \sin^2(ap_\mu) + a^{-2} [am_0 - \sum_\mu (\cos(ap_\mu) - 1)]^2 \right\} . \quad (2.18)$$

Now, in the continuum limit only the physical fermion mode at the origin $ap_\mu = \pi_{0000} = (0, 0, 0, 0) \in \Pi$ survives as the doubler contributions scale like a^{-1} and disappear from the spectrum. Another disadvantage that results from explicit chiral symmetry breaking by the Wilson

term in (2.16) is obviously an increase of discretisation errors from $O(a^2)$ to $O(a)$. Therefore the convergence to the continuum limit is reduced but can be systematically restored as will be explained in section 2.2. However, to get back an $O(a^2)$ convergence to the continuum limit at the non-perturbative level, additional work has to be done, see also section 6.

Maybe a more severe problem of chiral symmetry breaking of Wilson fermions is that beside the usual multiplicative mass renormalization, $m_R = Z_m m_0$, an additional *additive mass renormalization* has to be applied. In this case the quark mass renormalization reads

$$m_R = Z_m m_q, \quad \text{with} \quad m_q = m_0 - m_c. \quad (2.19)$$

m_c is called the (bare) *critical quark mass* and m_q the *bare subtracted quark mass*. Hence, the latter measures the deviation from the *critical line*, defined at

$$m_q \equiv 0 \quad \Leftrightarrow \quad m_c = m_0, \quad (2.20)$$

where the renormalized quark mass vanishes. In a theory with vanishing quark mass the remaining parameters are the couplings, therefore the critical line is parametrized by g_0 . Actually, by analysing chiral Ward–Takahashi identities (WTI) on the lattice for Wilson fermions, [45], it has been shown that operators and bare parameters can suitably be redefined such that the continuum renormalized WTI takes the form it assumes when chiral symmetry is preserved. To this end the mass m that generically depends on the bare parameters g_0, m_0 – and that appears in the axial WTI on the lattice – implicitly defines the critical (chiral) line $m_c(g_0)$ by the requirement $m(g_0, m_0) \equiv m_0$ in the space of bare parameters. There is no way to estimate $m_c(g_0)$ in advance. The parameter space always has to be explored by direct simulations, but perturbation theory can give a hint where to start. The fermion action with the chiral symmetry breaking Wilson–Dirac operator of eq. (2.16) is known as *Wilson action* and reads

$$S_W[U, \bar{\psi}, \psi] = a^4 \sum_x \bar{\psi}(x) (\gamma_\mu \tilde{\nabla}_\mu - \frac{a}{2} \nabla_\mu^* \nabla_\mu + m_0) \psi(x) \quad (2.21)$$

$$= a^8 \sum_{x,y} \bar{\psi}_x \left[\frac{1}{a} (am_0 + 4) \delta_{x,y} \right. \quad (2.22)$$

$$\left. - \frac{1}{2a} \sum_\mu \left\{ (1 - \gamma_\mu) U_{x,\mu} \delta_{x+a\hat{\mu},y} + (1 + \gamma_\mu) U_{x-a\hat{\mu},\mu} \delta_{x-a\hat{\mu},y} \right\} \right] \psi_y$$

$$\equiv a^8 \sum_{x,y} \bar{\psi}(x) M_{[U]}(x,y) \psi(y), \quad (2.23)$$

where the representations of appendix B.3.1 were used. In a theory with several flavours, m_0 becomes a mass matrix in flavour space and the fermion fields carry a flavour label which is summed over. The last line defines the *fermion matrix* M that depends on the gauge background $[U]$ and only contains next-neighbour interactions. Conventionally one applies an irrelevant field reparametrisation by introducing the so-called *hopping parameter* κ via

$$\psi \rightarrow \sqrt{2\kappa} \psi, \quad \bar{\psi} \rightarrow \bar{\psi} \sqrt{2\kappa}, \quad \kappa = 1/(2am_0 + 8), \quad (2.24)$$

which is the relevant (mass) parameter when doing explicit calculations on the lattice. Thus, the parameter set (β, κ) is in one-to-one correspondence to (g_0^2, m_0) and can be used alternately. The chiral limit at $m_0 = m_c$ defines the corresponding *critical hopping parameter* κ_c and the general bare subtracted quark mass in (2.19) becomes

$$am_q = \frac{1}{2} \left(\frac{1}{\kappa} - \frac{1}{\kappa_c} \right). \quad (2.25)$$

In actual simulations tuning of κ is required to fix the quark mass – either bare or renormalized – to a specific value. The Wilson fermion matrix with respect to the reparametrised field is

$$M_{xy}[U] = \frac{1}{a} \left[\delta_{xy} - \kappa \sum_{\mu} \left[(1 - \gamma_{\mu}) U_{x,\mu} \delta_{x,y-a\hat{\mu}} + (1 + \gamma_{\mu}) U_{x-a\hat{\mu},\mu}^{-1} \delta_{x,y+a\hat{\mu}} \right] \right]. \quad (2.26)$$

2.2 $O(a)$ -Improvement

Assuming R is a dimensionless variable like a ratio of masses for instance, the difference of the expectation value of R computed on the lattice R^{lat} to that of continuum R^{cl} would be proportional to some power p of the lattice spacing,

$$R^{\text{lat}} = R^{\text{cl}} + O(a^p). \quad (2.27)$$

p depends on the chosen lattice discretization of the action and on R itself. It can be increased by adding additional terms to the lattice action and R which serve as counterterms to the leading order cutoff dependence and therefore improve the rate of convergence to the continuum limit.

This rigorous treatment of the cutoff dependence in the lattice regularisation was invented by Symanzik for the case of the ϕ^4 theory [46] and the $O(N)$ non-linear sigma model [47]. His proposal was that the structural properties he found by applying lattice perturbation theory should also hold at the non-perturbative level.

2.2.1 The Symanzik improvement programme

Symanzik proposed that the lattice theory – at energies below the cutoff – should be equivalent to an effective continuum theory with effective action

$$S_{\text{eff}} = S_0 + \sum_{n=1}^{\infty} a^n S_n, \quad S_n = \int d^4x \mathcal{L}_n(x), \quad (2.28)$$

where S_0 is the original continuum action and the \mathcal{L}_n are linear combinations of local operators of dimension $4 + n$ with the symmetries of the lattice theory. Furthermore, local lattice fields renormalized in the effective theory, $\phi_R = Z_{\phi} \phi_{\text{eff}}$ are represented by local operators (with the same quantum numbers) ϕ_n of the continuum theory,

$$\phi_{\text{eff}} = \phi_0 + \sum_{n=1}^{\infty} a^n \phi_n. \quad (2.29)$$

Then the connected renormalized lattice n -point correlation functions to first order in a are [48]

$$G_n(x_1, \dots, x_n) = (Z_\phi)^n \langle \phi_{\text{eff}}(x_1) \cdots \phi_{\text{eff}}(x_n) \rangle_{\text{con}} \quad (2.30)$$

$$\begin{aligned} &= (Z_\phi)^n \left\{ \langle \phi_0(x_1) \cdots \phi_0(x_n) \rangle_{\text{con}} \right. \\ &\quad - a \int d^4 y \langle \phi_0(x_1) \cdots \phi_0(x_n) \mathcal{L}_1(y) \rangle_{\text{con}} \\ &\quad \left. + a \sum_{k=1}^n \langle \phi_0(x_1) \cdots \phi_1(x_k) \cdots \phi_0(x_n) \rangle_{\text{con}} + O(a^2) \right\}, \quad (2.31) \end{aligned}$$

where all points x_k are fixed and separated from each other to avoid contact terms. Expectation values are supposed to be taken in the continuum. In Wilsons formulation of QCD a basis of dimension five operators to describe at least the leading cutoff dependence is given by

$$\mathcal{O}_1 = \bar{\psi} \sigma_{\mu\nu} F_{\mu\nu} \psi, \quad (2.32a)$$

$$\mathcal{O}_2 = \bar{\psi} D_\mu D_\mu \psi + \bar{\psi} \overleftarrow{D}_\mu \overleftarrow{D}_\mu \psi, \quad (2.32b)$$

$$\mathcal{O}_3 = m \text{Tr} F_{\mu\nu} F_{\mu\nu}, \quad (2.32c)$$

$$\mathcal{O}_4 = m \{ \bar{\psi} \gamma_\mu D_\mu \psi - \bar{\psi} \overleftarrow{D}_\mu \gamma_\mu \psi \}, \quad (2.32d)$$

$$\mathcal{O}_5 = m^2 \bar{\psi} \psi. \quad (2.32e)$$

This set can be further reduced by restriction to on-shell quantities where some operators become redundant: when inserted into a correlation function it follows from the field equations $\mathcal{O}_4 \propto \mathcal{O}_5$, up to contact terms that can be absorbed by a redefinition of ϕ_1 . As \mathcal{O}_4 also \mathcal{O}_2 can be eliminated in favour of \mathcal{O}_5 . Independent of the restriction to on-shell quantities also \mathcal{O}_3 and \mathcal{O}_5 itself can be dropped as they correspond to a redefinition $(g_0^2, m_0) \rightarrow (\tilde{g}_0^2, \tilde{m}_0)$ of the bare coupling and mass, see below. Therefore the only counterterm needed to construct an on-shell $O(a)$ improved lattice action of QCD with Wilson fermions is the one first proposed by Sheikholeslami and Wohlert [49],

$$\delta S_G[U, \bar{\psi}, \psi] = S_{\text{sw}}[U, \bar{\psi}, \psi] = a c_{\text{sw}}(g_0) \int d^4 x \bar{\psi}(x) \frac{1}{4} \sigma_{\mu\nu} F_{\mu\nu} \psi(x). \quad (2.33)$$

Here $\sigma_{\mu\nu}$ is defined in (B.16) and the lattice representation $\hat{F}_{\mu\nu}$ of the field strength tensor $F_{\mu\nu}$ is the standard one,⁷

$$\hat{F}_{\mu\nu}(x) = \frac{1}{8a^2} \{ Q_{\mu\nu}(x) - Q_{\nu\mu}(x) \}, \quad (2.34)$$

$$\begin{aligned} Q_{\mu\nu}(x) &= U(x, \mu) U(x + a\hat{\mu}, \nu) U(x + a\hat{\nu}, \mu)^{-1} U(x, \nu)^{-1} \\ &\quad + U(x, \nu) U(x - a\hat{\mu} + a\hat{\nu}, \mu)^{-1} U(x - a\hat{\mu}, \nu)^{-1} U(x - a\hat{\mu}, \mu) \\ &\quad + U(x - a\hat{\mu}, \mu)^{-1} U(x - a\hat{\mu} - a\hat{\nu}, \nu)^{-1} U(x - a\hat{\mu} - a\hat{\nu}, \mu) U(x - a\hat{\nu}, \nu) \\ &\quad + U(x - a\hat{\nu}, \nu)^{-1} U(x - a\hat{\nu}, \mu) U(x + a\hat{\mu} - a\hat{\nu}, \nu) U(x, \mu)^{-1}. \quad (2.35) \end{aligned}$$

⁷ Here we point out a negligible inconsistency in our notation. In order to maintain consistency with the original nomenclature in the literature, these gauge link variables are inverse to those introduced in sec. 2.1, taken over from [150].

Due to the four plaquette loops appearing in eq. (2.35) and shown in figure 2, this counterterm is also called *clover term*. Each counterterm comes with an *improvement coefficient* that depends on the bare coupling g_0 – here this is c_{SW} . These coefficients intrinsically also depend on the number of active/dynamical flavours and the gauge and fermion action used. For other versions of c_{SW} see for instance [50] and references therein. The improvement coefficients have to be tuned carefully to cancel the $O(a)$ cutoff effects in on-shell quantities. It has been shown first in the quenched approximation of lattice QCD [51] that c_{SW} can and should be determined non-perturbatively. For $N_f = 2$ dynamical quark simulations as used in this work, c_{SW} was determined non-perturbatively in [52]:

$$c_{\text{SW}}(g_0)_{\text{NP}} = \frac{1 - 0.454g_0^2 - 0.175g_0^4 + 0.012g_0^6 + 0.045g_0^8}{1 - 0.720g_0^2}. \quad (2.36)$$

Currently, the computation for $N_f = 4$ Wilson fermions is under investigation [53]. To achieve $O(a)$ -improvement in spectral quantities, it is sufficient to improve the action. But if we define our lattice effective theory of QCD to be $O(a)$ improved as such, i.e. every QCD observable we can think of has to have a leading $O(a^2)$ cutoff dependence, we also need to improve all dimension 4 lattice operators non-perturbatively. In some cases where it is too demanding to get an improvement coefficient one has to restrict to perturbative improvement even if available.

2.2.2 Operator improvement

In QCD with two degenerate quark flavours a full set of bare iso-vector operators bilinear in the quark fields is given by

$$S^a(x) = \bar{\psi}(x) \frac{1}{2} \tau^a \psi(x), \quad (2.37a)$$

$$V_\mu^a(x) = \bar{\psi}(x) \gamma_\mu \frac{1}{2} \tau^a \psi(x), \quad (2.37b)$$

$$T_{\mu\nu}^a(x) = \bar{\psi}(x) i\sigma_{\mu\nu} \frac{1}{2} \tau^a \psi(x), \quad (2.37c)$$

$$A_\mu^a(x) = \bar{\psi}(x) \gamma_\mu \gamma_5 \frac{1}{2} \tau^a \psi(x), \quad (2.37d)$$

$$P^a(x) = \bar{\psi}(x) \gamma_5 \frac{1}{2} \tau^a \psi(x), \quad (2.37e)$$

with Pauli matrices (τ^a) from (B.6). In this order they are referred to as *scalar*, *vector*, *tensor*, *axial-vector* and *pseudo-scalar current*, respectively. Again the symmetry properties dictate the allowed improvement terms. The corresponding unrenormalized $O(a)$ improved fields are

$$(S_I)^a(x) = S^a(x), \quad (2.38a)$$

$$(V_I)_\mu^a(x) = V_\mu^a(x) + c_V \cdot a \tilde{\partial}_\nu T_{\mu\nu}^a(x), \quad (2.38b)$$

$$(T_I)_{\mu\nu}^a(x) = T_{\mu\nu}^a(x) + c_T \cdot a \left[\tilde{\partial}_\mu V_\nu^a(x) - \tilde{\partial}_\nu V_\mu^a(x) \right], \quad (2.38c)$$

$$(A_I)_\mu^a(x) = A_\mu^a(x) + c_A \cdot a \tilde{\partial}_\mu P^a(x), \quad (2.38d)$$

$$(P_I)^a(x) = P^a(x). \quad (2.38e)$$

The g_0 -dependent improvement coefficients c_V , c_T and c_A have to be computed such that the remaining $O(a)$ terms cancel in correlation functions involving the corresponding operators. Their first estimation obtained in perturbation theory,

$$c_V(g_0) = -0.01225(1) \times C_F g_0^2 + O(g_0^4), \quad (2.39a)$$

$$c_T(g_0) = 0.00896(1) \times C_F g_0^2 + O(g_0^4), \quad (2.39b)$$

$$c_A(g_0) = -0.005680(2) \times C_F g_0^2 + O(g_0^4), \quad (2.39c)$$

was given in [54, 55]. All leading order coefficients do not depend on N_f and the perturbative series vanish in the limit $g_0 \rightarrow 0$. An independent estimation obtained later by a different method [56, 57] shows consistency. For $N_f = 2$ dynamical quark flavours only c_A was computed non-perturbatively in [58]:

$$c_A(g_0)_{\text{NP}} = -0.00756g_0^2 \times \frac{1 - 0.4485g_0^2}{1 - 0.8098g_0^2}. \quad (2.40)$$

The range of validity is given by $\beta \geq 5.4$ and the parametrization respects the one-loop estimate shown above. c_V was only computed non-perturbatively in the quenched theory [59]. Thus we have to use the one-loop estimate later on. As we focus on observables composed of P , A , and V , we do not need the improvement coefficient c_T . But nevertheless T is computed in order to improve the vector current.

If we introduce the *pole mass* m_p as the energy of a free quark at zero spatial momentum, we get from the denominator of the quark propagator of free Wilson fermions,⁸

$$m_p = a^{-1} \ln(1 + am_0) = m_q - \frac{1}{2}am_q^2 + O(a^2) = m_q(1 - \frac{1}{2}am_q) + O(a^2), \quad (2.41)$$

which clearly shows a term of order a even at *tree-level* ($g_0^2 \equiv 0$). This leads to uncanceled $O(am)$ effects in various places and spoils the improvement programme as well as the set up of a mass-independent renormalization scheme (RS) at order a . As also the bare coupling suffers from such a contribution one already has to take it into account at the level of bare parameters of the theory by which renormalization factors are expressed. To employ a massless RS at the level of an on-shell $O(a)$ improved lattice theory we have to introduce a *modified bare coupling* and *bare quark mass* through

$$\tilde{g}_0^2 \equiv g_0^2 \left[1 + b_g(g_0)am_q \right], \quad \tilde{m}_q \equiv m_q \left[1 + b_m(g_0)am_q \right], \quad (2.42)$$

where again m_q is the subtracted bare quark mass of eq. (2.20). In such a scheme the renormalization of the bare parameters of the theory is given by

$$g_R^2 \equiv \tilde{g}_0^2 Z_g(\tilde{g}_0^2, a\mu), \quad m_R \equiv \tilde{m}_q Z_m(\tilde{g}_0^2, a\mu). \quad (2.43)$$

As pointed out the scaling of (g_0^2, m_q) necessarily depends on the subtracted bare quark mass

⁸corresponds to $g_0 = 0$ in the lattice action where the quarks decouple and $m_c(0) = 0$ and therefore $m_q = m_0$ holds

Table 2: Known perturbative values of improvement coefficients b for additive mass renormalization as given in [55]. As some of these coefficients will be computed non-perturbatively, only the perturbative expansion of b_V needs to be used in section 7. In the limit $g_0 \rightarrow 0$ the NP estimates have to be consistent with their PT counterparts. To this order only the one-loop coefficient to the coupling depends on the number of quark flavours, $N_f = 2$ here. For SU(3) we have $C_F = 4/3$. The lower part contains estimates we will need in section 6 where also the definition of Z is given, eq. (6.6).

X	$b_X^{(0)}$	$b_X^{(1)}$
g	0	$0.012000(2) \times N_f$
m	$-1/2$	$-0.07217(2) \times C_F$
S	1	$0.14434(5) \times C_F$
V	1	$0.11492(4) \times C_F$
T	1	$0.10434(4) \times C_F$
A	1	$0.11414(4) \times C_F$
P	1	$0.11484(2) \times C_F$
X	$X^{(0)}$	$X^{(1)}$
$b_A - b_P$	0	$-0.00093(8)$
Z	1	$0.090514(2)$

while the scaling of $(\tilde{g}_0^2, \tilde{m}_q)$ does not in an $O(a)$ improved theory. This in turn makes b_g and b_m well-determined and as we are dealing with the bare parameters independent of a particular renormalization scheme chosen.

Beside of this additive parameter renormalization we also need to track such additive mass contributions in the renormalization of multiplicatively renormalizable composite fields ϕ to achieve full $O(a)$ -improvement. As contributions like $1 + b_\phi am_q$ only arise at non-vanishing quark mass and the improvement of local fields itself (2.38) do not depend on the mass, one usually includes the corresponding $O(a)$ counterterm in the definition of the renormalized fields

$$\phi_R(x) = Z_\phi(\tilde{g}_0^2, a\mu) \left[1 + b_\phi(g_0) am_q \right] \phi_I(x). \quad (2.44)$$

Again, all the improvement coefficients b_ϕ are universal, i.e. they are independent of the renormalization conditions that fix Z_ϕ , and depend only on the bare coupling of the action. The usual perturbative expansion looks like

$$b_X(g_0) = \sum_{i=0}^{\infty} b_X^{(i)} (g_0^2)^i, \quad (2.45)$$

for all $X \in \{g, m, \{\phi\}\}$. As consistency constraint to the continuum expressions one can immediately deduce the tree-level values $b_g^{(0)} = 0$, $b_m^{(0)} = -\frac{1}{2}$ and $b_\phi^{(0)} = 1$ for $\phi \in \{S, V, T, A, P\}$. Known higher order coefficients of perturbation theory (PT) are summarized in table 2 and a strategy to non-perturbatively (NP) compute some of the coefficients $b_X(g_0)$ is presented in section 6.2. Without further notice, every perturbative expansion in the bare coupling follows the same counting scheme as given in (2.45).

2.3 Remarks

Besides the fact that we are working in the lattice regularisation of QCD no special mass-independent renormalization scheme was specified yet. It is a highly non-trivial task to define such a

scheme at the non-perturbative level in order to be able to compute the renormalization constants Z_g , Z_m and the Z_ϕ 's through numerical simulations from first principles.

The main obstacle is to fulfill the condition of vanishing quark mass that guaranties a simplification of the renormalization group equations and ease the connection to other massless renormalization schemes like the commonly used $\overline{\text{MS}}$ scheme. This is because it is still very demanding to simulate QCD with very light quarks.⁹ The feasibility of numerical inversions of the lattice Dirac equation strongly depends on the mass parameter. Simulations near the critical mass which controls the deviation from the physical mass up to renormalization, are dominated by the pure lattice Dirac operator and there is a point at which the mass cannot prevent the solution of the Dirac equation any more from approaching zero eigenvalue. But the duration of an algorithm for inverting the Dirac equation – and therefore the computational costs of a simulation – depends on the ratio between the largest λ_{max} and smallest eigenvalue λ_{min} , the *condition number*,

$$k = \lambda_{\text{max}} / \lambda_{\text{min}} . \quad (2.46)$$

Usually the largest eigenvalue is insensitive to the mass parameter and therefore approximately constant, such that the condition number is highly controlled by the minimal eigenvalue. The smaller λ_{min} the larger the condition number and even at moderate (small) masses it becomes rapidly inefficient to do the inversion. As long as there is no other quantity that serves as a sizeable gap in the lattice Dirac operator the situation remains worse for all lattice discretisations.

⁹This is the case not only for dynamical quarks but also for simulations in the quenched approximation due to so-called exceptional configurations.

3 The Schrödinger functional renormalization scheme

Since the early days of quantum field theory also the Schrödinger representation was known but first considered to be non-renormalizable. Then, Symanzik [60] proved in 1981 that the Schrödinger picture exists in renormalizable quantum field theories. His result was that in a quantum field theory with boundaries one needs further counterterms besides the usual ones to render the theory finite. Because these new counterterms appear due to the presence of boundaries, they also have to be imposed at the boundary of the manifold itself. He studied the massless ϕ^4 theory in detail and gave an outline how to deal with models including spin- $\frac{1}{2}$ fermions.

In 1986, Wolff studied scale transformations in asymptotically free theories with boundaries, [61, 62], which leads to the invention of a finite-size scaling method to compute the running coupling in a lattice regularized (gauge) theory. This method was proposed together with Lüscher and Weisz in [63], where the key idea was the

use of a *finite volume renormalization scheme with boundary conditions in time*, the Schrödinger functional. To be more precise, the *Schrödinger functional* (SF) is the propagation kernel that connects some field configuration at time $x_0 = 0$ to some other configuration at time $x_0 = T$, see figure 3. In Euclidean space-time this is just the functional integral over all fields with specified initial and final values. By an analysis of the SF to one-loop order in perturbation theory for gauge group $SU(2)$ [64], it was argued that the SF is renormalizable by means of the lattice regularization for generic non-Abelian gauge theories. Even if this is not proved to any order in PT, there is no doubt that this statement is true as various non-perturbative numerical computations confirmed it over the years. A definition and review of its applicability was also given in [65].

3.1 Formal definition

First, following [64] we repeat the definition of the Schrödinger functional for non-Abelian gauge theories without matter fields in continuum notation. A rigorous treatment in terms of a lattice regularised theory can be given as well and when it is helpful we switch between lattice and continuum notation.

In view of the applicability of the finite-size scaling technique mentioned above, one introduces the SF as finite volume scheme with spatial volume L^3 , made of a box of size $L \times L \times L$ with periodic boundary conditions. The extension in time direction is given by T and separates the initial and final states which are defined by Dirichlet boundary conditions at $x_0 = 0$ and $x_0 = T$, respectively. Thus, the space-time manifold the QFT relies on has the topology of a cylinder $([0, 1] \times S^1 \times S^1 \times S^1)$ with volume $V = T \times L^3$, see figure 3.

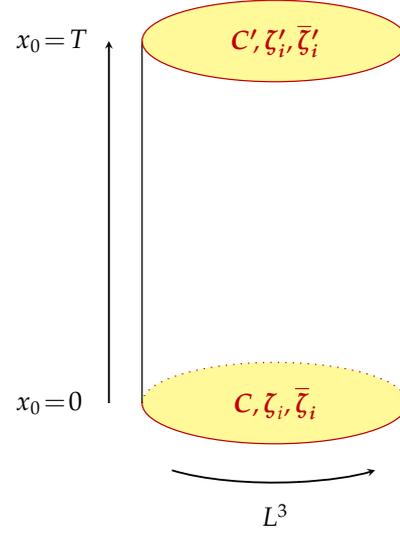


Figure 3: Sketch of the Schrödinger functional (SF) space-time manifold by suppressing two spatial dimensions.

The Dirichlet boundary conditions on the spatial components of the gauge vector field $A_\mu(x)$ with values in the Lie algebra of $SU(N)$ are given by some smooth classical gauge fields C, C' :

$$A_k(x)|_{x_0=0} = C_k^\Lambda(x), \quad A_k(x)|_{x_0=T} = C'_k(x), \quad k = 1, 2, 3. \quad (3.1)$$

Here, Λ denotes a local gauge transformation, defined by

$$A_k^\Lambda(\mathbf{x}) = \Lambda(\mathbf{x})A_k(\mathbf{x})\Lambda(\mathbf{x})^{-1} + \Lambda(\mathbf{x})\partial_k\Lambda(\mathbf{x})^{-1}, \quad \Lambda \in SU(N), \quad (3.2)$$

where only periodic, time-independent gauge functions $\Lambda(\mathbf{x})$ are admitted to preserve periodicity of the gauge fields,

$$A_k(x + L\hat{k}) = A_k(x), \quad \Lambda(\mathbf{x} + L\hat{k}) = \Lambda(\mathbf{x}). \quad (3.3)$$

Now, the pure gauge theory Schrödinger functional representation is formally given as an Euclidean functional integral,

$$\mathcal{Z}[C', C] \equiv \int \mathcal{D}[\Lambda] \int \mathcal{D}[A] e^{-S_g[A]}, \quad (3.4a)$$

$$\mathcal{D}[A] = \prod_{\mathbf{x}, \mu, a} dA_\mu^a(x), \quad \mathcal{D}[\Lambda] = \prod_{\mathbf{x}} d\Lambda(\mathbf{x}), \quad (3.4b)$$

with Haar measure $d\Lambda(\mathbf{x})$ on $SU(N)$ and gauge action

$$S_g[A] = -\frac{1}{2g_0^2} \int_V d^4x \text{Tr} F_{\mu\nu} F_{\mu\nu} = -\frac{1}{2g_0^2} \int_0^T dx_0 \int_0^L d^3\mathbf{x} \text{Tr} F_{\mu\nu} F_{\mu\nu}, \quad (3.5a)$$

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu + [A_\mu, A_\nu]. \quad (3.5b)$$

As usual, g_0 is the bare gauge coupling of the theory. While in the functional integral the time component of the gauge field appears, no attempt was made to restrict it on the boundaries. All definitions made so far are induced by the cylindrical topology of the SF. One is still left with gauge invariance, because eq. (3.2) and (3.4) are invariant under gauge transformations $\Omega(x)$,

$$A_\mu(x) \rightarrow \Omega(x)A_\mu(x)\Omega(x)^{-1} + \Omega(x)\partial_\mu\Omega(x)^{-1}, \quad (3.6a)$$

$$\Lambda(\mathbf{x}) \rightarrow \Omega(x)|_{x_0=0}\Lambda(\mathbf{x}), \quad \Omega(x)|_{x_0=T} = \mathbb{1}. \quad (3.6b)$$

A natural gauge fixing condition for this symmetry would be the temporal gauge $A_0 = 0$, which itself is ghost free as the Faddeev–Popov determinant is field independent and only contributes to an overall normalization of the functional integral.¹⁰ In general no gauge fixing procedure is needed within the Schrödinger functional formalism. However, in case of lattice perturbation theory for example, one is forced to choose a gauge.

¹⁰Actually, as in the case of Yang–Mills theories on compact manifolds without boundaries, gauge invariance reduces the integration over the gauge transformations Λ to a sum over topological classes. For practical considerations we set the vacuum angle θ to vanish here.

Quantum mechanical interpretation

The Schrödinger functional is the Schrödinger representation of quantum mechanics extended to quantum field theory. The states in QFT are wave functionals $\Psi[A]$ of the gauge field A as described above. With (3.4b) a scalar product can be defined by

$$\langle \Psi | \Phi \rangle = \int \mathcal{D}[A] \Psi[A]^* \Phi[A]. \quad (3.7)$$

Physically states are gauge invariant, i.e. they fulfill $\Psi[A^\Lambda] = \Psi[A]$ for all gauge transformations Λ . Using (3.4b), a projector for any given wave functional $\Psi[A]$ onto physical states is given by

$$\mathbb{P} \Psi[A] = \int \mathcal{D}[\Lambda] \Psi[A^\Lambda]. \quad (3.8)$$

The dynamical variables of the quantum field theory, $A_k^a(\mathbf{x})$, act as operator fields on those wave functionals. Their canonically conjugate field is the *chromoelectric field*

$$E_k^a(\mathbf{x}) \equiv F_{0k}^a = \frac{1}{i} \frac{\delta}{\delta A_k^a(\mathbf{x})}, \quad (3.9)$$

i.e. the non-vanishing time-space components of the field strength tensor $F_{\mu\nu}$. Together with the magnetic components of the colour field tensor,

$$F_{kl}^a(\mathbf{x}) = \partial_k A_l^a(\mathbf{x}) - \partial_l A_k^a(\mathbf{x}) + f^{abc} A_k^b(\mathbf{x}) A_l^c(\mathbf{x}), \quad (3.10)$$

the Hamilton operator takes the form

$$\mathbb{H} = \int_0^L d^3\mathbf{x} \left\{ \frac{g_0^2}{2} F_{0k}^a(\mathbf{x}) F_{0k}^a(\mathbf{x}) + \frac{1}{4g_0^2} F_{kl}^a(\mathbf{x}) F_{kl}^a(\mathbf{x}) \right\} \quad (3.11)$$

$$= \int_0^L d^3\mathbf{x} \frac{1}{2} \left\{ g_0^2 E_k^a(\mathbf{x}) E_k^a(\mathbf{x}) + \frac{1}{g_0^2} B_i^a(\mathbf{x}) B_i^a(\mathbf{x}) \right\}, \quad (3.12)$$

where the second equation only holds in four space-time dimension due to $B_i^a = \frac{1}{2} \varepsilon_{ikl} F_{kl}^a$. Each classical gauge field $C_k(\mathbf{x})$ defines a state vector $|C\rangle$ through

$$\langle C | \Psi \rangle = \Psi[C], \quad (3.13)$$

which can be made gauge invariant after applying the projector \mathbb{P} of equation (3.8). Finally, the Euclidean Schrödinger functional $\mathcal{Z}[C', C]$ is defined by

$$\mathcal{Z}[C', C] \equiv \langle C' | e^{-T\mathbb{H}} \mathbb{P} | C \rangle = \sum_{n=0}^{\infty} e^{-E_n T} \Psi_n[C'] \Psi_n[C]^*. \quad (3.14)$$

The last term shows the *spectral representation*, obtained after inserting an orthogonal basis $|\Psi_n\rangle$, $n = 0, 1, 2, \dots$ of gauge invariant energy eigenstates with eigenvalues E_n .

The induced background field

Since proposed in [66], it is well known that classical solutions of the four dimensional Euclidean Yang–Mills equations exist which are topological in nature. These solutions are called instantons now and give a lower bound to the gauge field action $S_g[A]$.

It was not mentioned yet explicitly but it should be clear that by imposing boundary values C and C' to the system, this also has an effect to the (absolute) minimum of the action. That means, if C and C' are small, there should exist a configuration $B_\mu(x)$ of minimal action,

$$S[B] < S[A], \forall A \neq B^\Omega \quad (3.15)$$

which is unique up to gauge transformations and simply connected to the boundary conditions. It was explicitly shown in [64] that this is indeed the case and for obvious reasons, B is called the (boundary condition) *induced background field*.¹¹ Since B is a classical solution to the equation of motion and therefore dominates the functional integral (3.4) in the weak coupling domain, the Schrödinger functional can be computed by a saddle point expansion [67] about B . Thus, one obtains a series expansion of the *effective action*,

$$\Gamma[B] \equiv -\ln \mathcal{Z}[C', C], \quad (3.16)$$

given by

$$\Gamma[B] = \frac{1}{g_0^2} \Gamma_0[B] + \Gamma_1[B] + g_0^2 \Gamma_2[B] + \dots, \quad \Gamma_0[B] \equiv g_0^2 S[B]. \quad (3.17)$$

The Feynman diagrams involved in this expansion still need an ultra-violet regulator to be well-defined. To preserve the gauge invariance of the theory, dimensional or lattice regularization should be used. In case of the latter, a careful choice of B is necessary to obtain lattice corrections to the effective action which are tolerable.

A special (and simple) class of background fields that comply with this are *Abelian background fields*. These background fields are spatially constant and diagonal. Their general construction was given in [64]. Restricted to gauge group $SU(3)$ they read

$$C_k = \frac{i}{L} \text{diag}(\phi_1, \phi_2, \phi_3), \quad C'_k = \frac{i}{L} \text{diag}(\phi'_1, \phi'_2, \phi'_3), \quad k = 1, 2, 3, \quad (3.18)$$

for two sets of real angles both satisfying $\sum_i \phi_i = \text{Tr} C_k = 0$ to ensure that C_k, C'_k are elements of $SU(3)$. In principle one could choose another set of angles for each spatial index k . Abelian background fields obey the following solution to the field equations,

$$B_0(x) = 0, \quad B_k(x) = [x_0 C'_k + (L - x_0) C_k] / L, \quad k = 1, 2, 3. \quad (3.19)$$

¹¹Even if it is usually impossible to get $B[C', C]$ in closed analytical form, it exists a one-to-one correspondence between B and the boundary values C, C' due to the assumptions made. Hence, the dependence on the boundary values is interchangeable with that of B .

The associated non-vanishing components of the field strength tensor are

$$\partial_0 B_k = (C'_k - C_k)/L, \quad k = 1, 2, 3. \quad (3.20)$$

Actually, there are only two sets of boundary field conditions that are of interest here,

- the vanishing background field configuration $\mathcal{BF} = 0$, with

$$(\phi_1, \phi_2, \phi_3) = (\phi'_1, \phi'_2, \phi'_3) = 0 \quad \Leftrightarrow \quad C = 0 = C', \quad (3.21)$$

- and the background field configuration $\mathcal{BF} = A$:

$$(\phi_1, \phi_2, \phi_3) = \frac{1}{3}(-\pi, 0, \pi) + \frac{1}{2}(2, -1, -1)\eta, \quad (3.22a)$$

$$(\phi'_1, \phi'_2, \phi'_3) = -\frac{2}{3}(2, -1, -1) - (\phi_1, \phi_3, \phi_2). \quad (3.22b)$$

The second choice obviously forms a one parameter family of boundary fields in the parameter η which is restricted to $0 < \eta < \pi$. The mapping $\phi_i \mapsto \phi'_i$, i.e. eq. (3.22b), is a discrete symmetry of the SF. For an explicit construction see [68]. It turns out that this choice leads to a smaller statistical error in the numerical computation of the renormalized coupling \bar{g}^2 , see also section 3.2.

Boundary gauge fields

In accordance with (3.1) the spatial gauge link variables connected to the boundaries of the SF cylinder are given by the constant Abelian (continuum) gauge fields C_k and C'_k ,

$$U(x, k)|_{x_0=0} = \exp(aC_k), \quad U(x, k)|_{x_0=T} = \exp(aC'_k). \quad (3.23)$$

In the continuum this setting leads to an up to gauge transformations unique minimal action configuration V which is expressible by the background field,

$$V(x, \mu) = \exp(aB_\mu(x)). \quad (3.24)$$

In this sense it is the lattice representative of the background field B .

Each dimensionless quantity that is finite on removal of the regulator and that depends on exactly one scale can serve as a renormalized coupling. A priori there are a lot of equivalent formulations that fulfill this condition. However, one usually chooses a definition that has some preferable properties for the problem under consideration and which preserve a well-behaved perturbative expansion to connect it to other renormalization schemes. Furthermore, in contrast to eq. (1.21) the definition of a coupling should ideally be non-perturbatively from the beginning. Such a definition is given in the next subsection and recently a new scheme was proposed in [69]. Both are capable to study the scaling properties of the strong coupling over a wide range of energy scales.

3.2 The SF renormalized coupling

Here the idea is to define a non-perturbative renormalized coupling $\bar{g}^2(\mu)$ by identifying the renormalization point μ with the inverse of the spatial box length L of the Schrödinger functional. This definition allows to apply a finite-size scaling technique to study the scale dependence in non-trivial theories like Yang–Mills or the non-linear sigma model. Furthermore, it naturally separates the different scales that appear when studying renormalization problems on the lattice,

$$L \gg \frac{1}{m_\pi} \sim \frac{1}{0.14 \text{ GeV}} \gg \frac{1}{\mu_{\text{pt}}} \sim \frac{1}{10 \text{ GeV}} \gg a. \quad (3.25)$$

The origin for a non-perturbative definition of a strong coupling in the SF is the one parameter family (3.22) in view of the known perturbative expansion of the effective action, eq. (3.17). The former allows a variation of the latter on the gauge background and results in an extraction of a coupling by means of the definition

$$\bar{g}^2(L) \equiv \frac{\Gamma'_0[B]}{\Gamma'[B]}, \quad \Gamma'[B] \equiv \left. \frac{\partial \Gamma[B_\eta]}{\partial \eta} \right|_{\eta=0} = \left\langle \left. \frac{dS}{d\eta} \right|_{\eta=0} \right\rangle. \quad (3.26)$$

In this sense the running coupling is the response of the system to the induced background field. The normalisation with the constant Γ'_0 is such that the usual leading behaviour in the perturbative expansion, g_0^2 , is recovered. Since the background field B scales like an inverse length, it is made dimensionless by a multiplication with L such that the remaining dependence on the parameter η is apparent. Note that L is the only available scale in the problem as we refer to the finite volume SF scheme in the continuum. Hence, \bar{g}^2 scales with L and the definition of the renormalized coupling is not restricted to any particular regularization. However, to compute the scale dependence of the renormalized coupling from first principles non-perturbatively one has to do two things: (a) choose the lattice as regulator and perform Monte-Carlo simulations at different lattice resolutions a/L in order to extract the continuum value of the coupling at fixed scale $\mu = 1/L$, and (b) repeat those steps at different scales which are connected in a clean way to the previous one in order to extract the running.

To keep the discussion clear, we introduced the SF running coupling in the pure gauge theory. But as we study QCD with two dynamical quark flavours we also have to introduce and specify the lattice fermion action which naturally influences the running of the coupling as well as it appears in the total action S in eq. (3.26). That does not change the definition of the coupling which just becomes dependent on the flavour content of the theory via the action in contrast to quenched QCD. To complete the definition of the SF running coupling we have to specify the renormalization condition which is hold fixed in course of the computational steps mentioned above. In the pure gauge theory these are the boundary field \mathcal{BF} and the ratio T/L . The introduction of quarks presented in the next subsection contributes with two more conditions, the value of the (renormalized) mass¹² and an angle θ . Last but not least, we have to fix a (low-energy/hadronic) scale $L^* = 1/\mu^*$ at which we start the recursive computation of the running coupling. This constraint is given by a

¹² To get a mass-independent RS this mass clearly has to vanish such that $am_q = 0$ holds.

(parameter) *window of applicability* that bounds every numerical method in the sense of costs and reliability. Usually, the hadronic scale is implicitly defined by specifying the SF coupling at that point, i.e. $\bar{g}^2(L)|_{L=L^*}$. In ref. [28] the running coupling in $N_f = 2$ massless QCD was computed using the following renormalization conditions,

$$\{\bar{g}^2(L^*), \mathcal{BF}, T/L, am_q, \theta\} = \{5.5, A, 1, 0, \pi/5\}. \quad (3.27)$$

See for instance figure 1. This was accompanied by a two-loop computation in [70] such that the error stemming from the remaining perturbative running at high energies becomes negligible.

3.3 Quarks in the SF

Due to the natural separation of bulk and boundary in the SF, the usual Wilson lattice actions adopt some minor modifications and some additional boundary improvement terms have to be introduced to restore an $O(a)$ improved lattice theory. Following Symanzik, the general form of the correction terms (2.28) that are encountered when the lattice SF approaches the continuum one, reads

$$S_k = \int d^4x \mathcal{L}_k(x) + \lim_{\epsilon \rightarrow 0^+} \int d^3\mathbf{x} \left\{ \mathcal{B}_k(x)|_{x_0=\epsilon} + \mathcal{B}_k(x)|_{x_0=T-\epsilon} \right\}. \quad (3.28)$$

The construction of the SF on top of Wilson's lattice QCD goes back to Sint, [71, 72]. It is based upon the transfer matrix formulation used by Lüscher to prove positivity of Wilson's lattice gauge theory [35] which is the pre-requisite for a quantum mechanical Hamiltonian with only real eigenvalues. The continuum fermion action in the SF scheme was found to be

$$S_F[A, \bar{\psi}, \psi] = \int d^4x \bar{\psi}(x) [\gamma_\mu D_\mu + m] \psi(x) - \int d^3\mathbf{x} [\bar{\psi}(x) P_- \psi(x)]_{x_0=0} - \int d^3\mathbf{x} [\bar{\psi}(x) P_+ \psi(x)]_{x_0=T}, \quad (3.29)$$

where due to the Dirichlet boundary conditions on the fermion fields in Euclidean time direction,

$$P_+ \psi(x)|_{x_0=0} = \rho(\mathbf{x}), \quad P_- \psi(x)|_{x_0=L} = \rho'(\mathbf{x}), \quad (3.30a)$$

$$\bar{\psi}(x) P_-|_{x_0=0} = \bar{\rho}(\mathbf{x}), \quad \bar{\psi}(x) P_+|_{x_0=L} = \bar{\rho}'(\mathbf{x}), \quad (3.30b)$$

additional boundary terms with the corresponding (time) projection operators,

$$P_\pm = \frac{1}{2}(1 \pm \gamma_0), \quad (3.31)$$

appear. The assumptions that are made in this construction are invariance of the action under parity due to P_\pm and the existence of smooth classical solutions $\bar{\psi}_{cl}, \psi_{cl}$ to the equations of motion,

$$\psi(x) = \psi_{cl}(x) + \chi(x) \quad \text{and} \quad (\gamma_\mu D_\mu + m)\psi_{cl}(x) = 0 \quad \forall \mathbf{x}, 0 < x_0 < T, \quad (3.32)$$

with ψ expanded around the saddle point ψ_{cl} . While the projections in (3.30) on ψ vanish for

the quantum components χ , the boundary fields represent the classical solution and the quantum dynamics is restricted to the *bulk* of the SF cylinder, $0 < x_0 < T$. The same holds for the anti-fermion field $\bar{\psi}$. In fact, the appearance of the boundary terms in the action result from uniqueness reasons of the Dirac equation, a partial differential equation of first order. That means that only half of the components of the 4-component Dirac fields can be specified at the boundaries. In addition to the gauge boundary fields C and C' , the Schrödinger functional now also depends on the boundary fermion and anti-fermion field ρ, ρ' and $\bar{\rho}, \bar{\rho}'$ – one for each flavour of quarks.

In contrast to the periodic boundary conditions in space that apply for the gauge fields (3.3), the most general boundary conditions for the fermion fields read

$$\psi(x + L\hat{k}) = e^{i\theta_k} \psi(x) , \quad \bar{\psi}(x + L\hat{k}) = e^{-i\theta_k} \bar{\psi}(x) . \quad (3.33)$$

The phase angles $\theta_k, k = 1, 2, 3$ are real numbers and we restrict ourselves immediately to $\theta_1 = \theta_2 = \theta_3 \equiv \theta$. This angle serves as an additional degree of freedom that has to be specified and $\theta = 0$ refers to standard periodic boundary conditions. The additional degree of freedom can be utilised to improve numerical computations as it was done for the renormalized coupling with $\theta = \pi/5$ as suggested by a one-loop computation in [22]. It has been shown that this choice increases the lowest eigenvalue at fixed T by more than a factor of two, compared to $\theta = 0$. The result is a speed up in the corresponding Monte-Carlo simulation that can be used to increase the statistics and thus to reduce the uncertainty. It is equivalent and practical to implement θ at the level of the gauge covariant derivatives, see appendix B.3.1. In section 7 we will see and discuss some results for different values of θ that apparently parametrise a family of fermionic boundary conditions. Note that imposing conditions like (3.33) is not special to the SF. For a more general discussion see [73].

The finite gap of the Schrödinger functional

A very important feature of the SF is that the fermionic Dirichlet boundary conditions induce a finite gap in the spectrum of the continuum Dirac operator and hence serve as a natural infrared cutoff. This has been emphasized in [71], where it has been shown that also in case of vanishing quark mass the free Dirac operator exhibits a purely discrete spectrum without zero modes and smallest eigenvalue

$$\lambda_0^2 = \left(\frac{\pi}{2T} \right)^2 , \quad \text{at} \quad \theta = 0 , \quad C = C' = 0 . \quad (3.34)$$

Turning on the interaction and/or altering some of these parameters of course influences the value of the minimal eigenvalue. For even larger time extension T , the gap decreases and one is confronted with the usual practical problems of simulating Wilson fermions on the lattice. We will come back to this in section 5.2.

3.4 The Wilson action revisited

Improvement of the gauge action (2.10) is achieved by setting

$$S_G[U] = \frac{1}{g_0^2} \sum_{\mathbf{p}} w(\mathbf{p}) \text{Tr} \{1 - U_{\mathbf{p}}\} , \quad (3.35)$$

with the following weight factors [48] at the oriented plaquette \mathbf{p} :

$$w(\mathbf{p}) = \begin{cases} 1 & \text{for } \mathbf{p} \text{ in the bulk} \\ c_t(g_0) & \text{for time-like } \mathbf{p} \text{ connected to } x_0 = 0, T \\ \frac{1}{2}c_s(g_0) & \text{for space-like } \mathbf{p} \text{ at } x_0 = 0, T \end{cases} . \quad (3.36)$$

To reduce the lattice artefacts, the improvement coefficients $c_t(g_0)$ and $c_s(g_0)$ need to be tuned as well. The clover term (2.33) reduces to a sum in the bulk,

$$S_{\text{sw}}[U, \bar{\psi}, \psi] = a^5 c_{\text{sw}}(g_0) \sum_{x_0=a}^{T-a} \sum_{\mathbf{x}} \bar{\psi}(x) \frac{i}{4} \sigma_{\mu\nu} \hat{F}_{\mu\nu}(x) \psi(x) \equiv a^4 \sum_{x_0=a}^{T-a} \sum_{\mathbf{x}} \bar{\psi}(x) \delta D_{\mathbf{v}} \psi(x) . \quad (3.37)$$

For counterterms due to the quark fields the field equations can again be used to reduce the total set of dimension 5 composite fields that are compatible with the symmetries. Some counterterms can be eliminated by (irrelevant) rescaling of the boundary values ρ and $\bar{\rho}$. The remaining four boundary counterterms that preserve the time reversal symmetry of the theory in the combination

$$\delta S_{F,b}[U, \bar{\psi}, \psi] = a^4 \sum_{\mathbf{x}} \left\{ (\tilde{c}_s - 1) [\mathcal{O}_s(\mathbf{x}) - \mathcal{O}'_s(\mathbf{x})] + (\tilde{c}_t - 1) [\mathcal{O}_t(\mathbf{x}) - \mathcal{O}'_t(\mathbf{x})] \right\} , \quad (3.38a)$$

are given by

$$\mathcal{O}_s(\mathbf{x}) = \frac{1}{2} \bar{\rho}(\mathbf{x}) \gamma_k (\nabla_k^* + \nabla_k) \rho(\mathbf{x}) , \quad (3.38b)$$

$$\mathcal{O}'_s(\mathbf{x}) = \frac{1}{2} \bar{\rho}'(\mathbf{x}) \gamma_k (\nabla_k^* + \nabla_k) \bar{\rho}'(\mathbf{x}) , \quad (3.38c)$$

$$\mathcal{O}_t(\mathbf{x}) = \frac{1}{2} \left[\bar{\psi}(x) (P_- \nabla_0 + P_+ \overleftarrow{\nabla}_0^*) \psi(x) \right]_{x_0=a} , \quad (3.38d)$$

$$\mathcal{O}'_t(\mathbf{x}) = \frac{1}{2} \left[\bar{\psi}(x) (P_+ \nabla_0 + P_- \overleftarrow{\nabla}_0^*) \psi(x) \right]_{x_0=T-a} . \quad (3.38e)$$

Thus, another set of improvement coefficients $\tilde{c}_t(g_0)$, $\tilde{c}_s(g_0)$ has to be introduced. Fortunately, for Abelian or spatial homogeneous boundary fields C , C' as considered here, the space-like improvement coefficients are $c_s = \tilde{c}_s = 1$ and their counterterms give no contribution. The time-like boundary improvement coefficients are only known perturbatively. c_t is known to two-loop [70]

$$c_t(g_0) = 1 + (-0.08900(5) + 0.0191410(1)N_f) g_0^2 + (-0.0294(3) + 0.002(1)N_f + 0.0000(1)N_f^2) g_0^4 , \quad (3.39)$$

while \tilde{c}_t was yet only computed to one-loop [48, 70, 74],

$$\tilde{c}_t(g_0) = 1 - 0.01795(2)g_0^2. \quad (3.40)$$

The latter is known to be independent of N_f and we set c_t to its proper value at $N_f = 2$ in the following.

All in all the action which describes QCD as an on-shell $O(a)$ improved lattice theory in the Schrödinger functional is given by the sum of (3.35), (3.29), (3.37) and (3.38), that is

$$S[U, \bar{\psi}, \psi] = S_G[U] + S_F[U, \bar{\psi}, \psi] + S_{\text{sw}}[U, \bar{\psi}, \psi] + \delta S_{F,b}[U, \bar{\psi}, \psi] + S_{F,s}[U, \bar{\psi}, \psi]. \quad (3.41)$$

The last piece that was introduced here, is the additional source term

$$S_{F,s}[U, \bar{\psi}, \psi] = a^4 \sum_{x_0=a}^{T-a} \sum_{\mathbf{x}} [\bar{\psi}(x)\eta(x) + \bar{\eta}(x)\psi(x)], \quad (3.42)$$

that allows us to compute expectation values of local composite fields by the usual functional integral approach. To shorten notation, I never explicitly write the dependence of the action on the various source fields $C, C', \bar{\rho}, \rho, \bar{\rho}', \rho', \bar{\eta}$ and η .

3.5 Expectation values of composite operators

With the lattice action (3.41) the Schrödinger functional reads

$$\mathcal{Z}[C, C'; \bar{\rho}', \rho'; \bar{\rho}, \rho; \bar{\eta}, \eta] = \int \mathcal{D}[U] \mathcal{D}[\psi] \mathcal{D}[\bar{\psi}] e^{-S[U, \bar{\psi}, \psi]} \quad (3.43)$$

$$= \int \mathcal{D}[U] \mathcal{Z}_F[U; \bar{\rho}', \rho'; \bar{\rho}, \rho; \bar{\eta}, \eta] e^{-S_G[U, \bar{\psi}, \psi]}, \quad (3.44)$$

$$\mathcal{Z}_F[U; \bar{\rho}', \rho'; \bar{\rho}, \rho; \bar{\eta}, \eta] = \int \mathcal{D}[\psi] \mathcal{D}[\bar{\psi}] e^{-S_{F,I}[U, \bar{\psi}, \psi]}. \quad (3.45)$$

This natural decomposition into a gauge functional enclosing a fermionic generating functional is the key when simulating lattice QCD and building fermion expectation values. It follows directly from the continuum expression. The weight in the fermionic generating functional is the improved fermion action, given through eq. (3.41) by $S_{F,I} = S - S_G$. The expectation value of any operator \mathcal{O} is formally given by

$$\langle \mathcal{O} \rangle = \left\{ \mathcal{Z}^{-1} \int \mathcal{D}[U] \mathcal{D}[\psi] \mathcal{D}[\bar{\psi}] \mathcal{O} e^{-S[U, \bar{\psi}, \psi]} \right\}_{\substack{\bar{\rho}' = \bar{\rho} = \bar{\eta} = \\ \rho' = \rho = \eta = 0}}. \quad (3.46)$$

In the standard way we can identify the fermion fields in the SF as functional derivatives with respect to the corresponding sources in the action. These are

$$\psi(x) = \frac{\delta}{\delta\eta(x)}, \quad \zeta(\mathbf{x}) = \frac{\delta}{\delta\bar{\rho}(\mathbf{x})}, \quad \zeta'(\mathbf{x}) = \frac{\delta}{\delta\bar{\rho}'(\mathbf{x})}, \quad (3.47a)$$

$$\bar{\psi}(x) = -\frac{\delta}{\delta\eta(x)}, \quad \bar{\zeta}(\mathbf{x}) = -\frac{\delta}{\delta\rho(\mathbf{x})}, \quad \bar{\zeta}'(\mathbf{x}) = -\frac{\delta}{\delta\rho'(\mathbf{x})}. \quad (3.47b)$$

Now the fields appearing in the operator \mathcal{O} are to be substituted by the corresponding functional derivative and the expectation value takes the form

$$\langle \mathcal{O} \rangle = \left\{ \mathcal{Z}^{-1} \mathcal{O} \mathcal{Z} \right\}_{\substack{\bar{\rho}'=\bar{\rho}=\bar{\eta}= \\ \rho'=\rho=\eta=0}}. \quad (3.48)$$

We call $\zeta, \bar{\zeta}$ *boundary quark* and *boundary anti-quark field* located at $x_0 = 0$, respectively. Hence, $\zeta', \bar{\zeta}'$ are located at the boundary $x_0 = T$. The minus sign in front arises due to the Grassmann nature of fermions and the structure of the fermion bilinears. As is well-known, a computer cannot deal with Grassmann numbers intrinsically. But fortunately due to the bilinear structure in the action, the fermion fields can be analytically integrated out. Then, the expectation value (3.46) is given by

$$\langle \mathcal{O} \rangle \equiv \left\langle [\mathcal{O}]_{\text{F}} \right\rangle_U = \left\{ \mathcal{Z}_{\text{eff}}^{-1} \int \mathcal{D}[U] [\mathcal{O}]_{\text{F}} e^{-S_{\text{eff}}[U]} \right\}, \quad \mathcal{Z}_{\text{eff}} = \int \mathcal{D}[U] e^{-S_{\text{eff}}[U]}, \quad (3.49a)$$

$$S_{\text{eff}}[U] = S_G[U] - \ln [\det (M[U])], \quad (3.49b)$$

where the fermionic contribution $[\mathcal{O}]_{\text{F}}$ has to be evaluated on the given gauge background U ,

$$[\mathcal{O}[\bar{\zeta}', \zeta'; \bar{\zeta}, \zeta; \bar{\psi}, \psi]]_{\text{F}} = \left\{ \mathcal{Z}_{\text{F}}^{-1} \mathcal{O} \left[-\frac{\delta}{\delta\bar{\rho}'}, \dots, \frac{\delta}{\delta\bar{\eta}} \right] \mathcal{Z}_{\text{F}} \right\}_{\substack{\bar{\rho}'=\bar{\rho}=\bar{\eta}= \\ \rho'=\rho=\eta=0}}. \quad (3.49c)$$

The functional derivatives in $[\bullet]_{\text{F}}$ are carried out by applying Wick's theorem in the usual way, i.e. writing $[\mathcal{O}]_{\text{F}}$ as sums of Wick contractions which can be expressed through appropriate propagators. A complete list of non-zero contractions is given in appendix B.5.

In the effective action which defines the Boltzmann weight factor, $M[U]$ is the fermion matrix on a fixed gauge background which in the sense of eq. (2.23) is given by the Dirac operator of $S_{F,I}$. The appearance of $\det(M[U])$ makes the problem of computing expectation values through Monte-Carlo techniques highly sophisticated as the determinant couples every lattice site to each other and thus is highly non-local. By setting this determinant to one, all contributions from quark loops are neglected. This is the so-called *quenched approximation*. Further details about the Monte-Carlo technique and our setup of dynamical fermion simulations are presented in section 5.2.

3.6 Fermion correlation functions

Our focus is on meson observables. Thus, we need operators that create and annihilate the corresponding physical states as well as the propagators of their constituent quarks. The corresponding

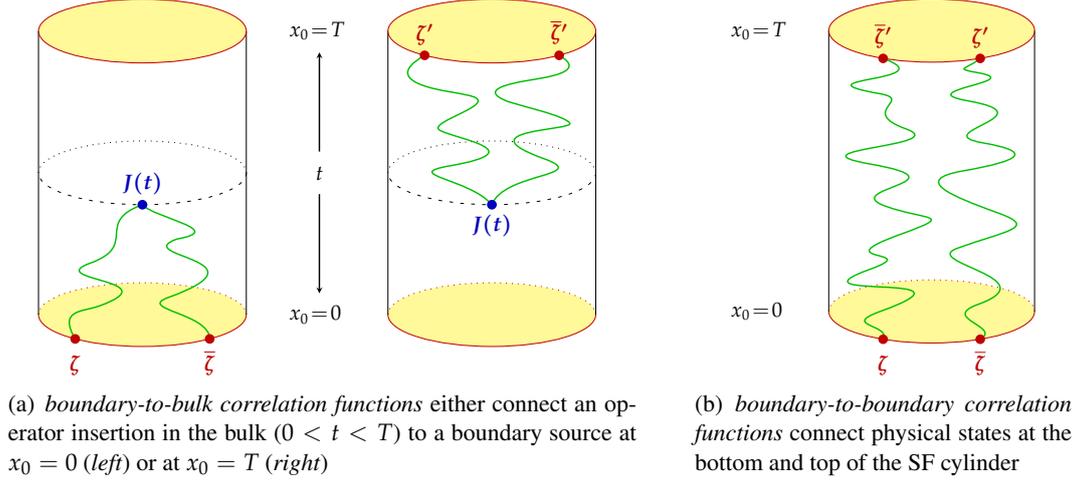


Figure 4: In the SF we have two different classes of correlation functions, boundary-to-bulk (a) and boundary-to-boundary (b). The operator insertions under consideration, which can appear in the bulk or at the boundaries, are P , A_0 , V_k and T_{0k} .

transition amplitude can be build in the SF in three different ways, see figure 4:

1. create a state at $x_0 = 0$ that propagates through the bulk and gets annihilated at $x_0 = T$ – we call this a *boundary-to-boundary* correlation function
2. create a state at $x_0 = 0$ that will propagate forward in time and gets annihilated somewhere in the bulk $0 < x_0 < T$ – this is a *forward boundary-to-bulk* correlation function
3. create a state at $x_0 = T$ that propagates backward in time and gets annihilated in the bulk – a *backward boundary-to-bulk* correlation function

We formerly mentioned that our main interest is in the currents $X \in \{P, A, V, T\}$ as defined in (2.37). From the point of view of quantum numbers, the first two belong to the *pseudo-scalar*, $J^P = 0^-$, and the latter to the *vector channel*, $J^P = 1^-$.¹³ Additionally, the corresponding mesons are characterized by their flavour content, that is the mass of the quarks they are built of. But this is of secondary importance as we can vary their mass by the hopping parameter κ . We denote the isospin operators that create a pseudo-scalar/vector state at $x_0 = 0$ (left) and $x_0 = T$ (right) by

$$\mathcal{O}^a = \frac{a^6}{L^3} \sum_{\mathbf{u}, \mathbf{v}} \bar{\zeta}(\mathbf{u}) \gamma_{5\frac{1}{2}} \tau^a \zeta(\mathbf{v}), \quad \mathcal{O}'^a = \frac{a^6}{L^3} \sum_{\mathbf{u}, \mathbf{v}} \bar{\zeta}'(\mathbf{u}) \gamma_{5\frac{1}{2}} \tau^a \zeta'(\mathbf{v}), \quad (3.50a)$$

$$\mathcal{Q}_k^a = \frac{a^6}{L^3} \sum_{\mathbf{y}, \mathbf{z}} \bar{\zeta}(\mathbf{y}) \gamma_{k\frac{1}{2}} \tau^a \zeta(\mathbf{z}), \quad \mathcal{Q}'_k{}^a = \frac{a^6}{L^3} \sum_{\mathbf{y}, \mathbf{z}} \bar{\zeta}'(\mathbf{y}) \gamma_{k\frac{1}{2}} \tau^a \zeta'(\mathbf{z}). \quad (3.50b)$$

By summing over the whole spatial volume we build zero momentum projected sources which means that the corresponding states at the boundary are at rest. Then the forward correlation

¹³ J : total angular momentum; P : parity

functions for the currents $X \in \{P, A, V, T\}$ read

$$f_A(x_0) = -\frac{a^3}{2} \sum_{\mathbf{x}} \langle A_0^a(x) \mathcal{O}^a \rangle, \quad f_P(x_0) = -\frac{a^3}{2} \sum_{\mathbf{x}} \langle P^a(x) \mathcal{O}^a \rangle, \quad (3.51a)$$

$$k_V(x_0) = -\frac{a^3}{6} \sum_{\mathbf{x}} \langle V_k^a(x) \mathcal{Q}_k^a \rangle, \quad k_T(x_0) = -\frac{a^3}{6} \sum_{\mathbf{x}} \langle T_{k0}^a(x) \mathcal{Q}_k^a \rangle, \quad (3.51b)$$

where summation over isospin index a and spatial index k is assumed. Analogously the backward correlation functions are constructed by

$$g_A(T - x_0) = -\frac{a^3}{2} \sum_{\mathbf{x}} \langle \mathcal{O}'^a A_0^a(x) \rangle, \quad g_P(T - x_0) = -\frac{a^3}{2} \sum_{\mathbf{x}} \langle \mathcal{O}'^a P^a(x) \rangle, \quad (3.52a)$$

$$l_V(T - x_0) = -\frac{a^3}{6} \sum_{\mathbf{x}} \langle \mathcal{Q}'^a V_k^a(x) \rangle, \quad l_T(T - x_0) = -\frac{a^3}{6} \sum_{\mathbf{x}} \langle \mathcal{Q}'^a T_{k0}^a(x) \rangle. \quad (3.52b)$$

The parametrisation chosen here is such that they fall-off with $T - x_0$ as the forward correlation functions do with x_0 . We finally note the definition of boundary-to-boundary correlators:

$$f_1 = -\frac{1}{3L^6} \langle \mathcal{O}'^a \mathcal{O}^a \rangle, \quad k_1 = -\frac{1}{6L^6} \langle \mathcal{Q}'^a \mathcal{Q}_k^a \rangle. \quad (3.53)$$

Under a time reflection transformation of the SF the boundary gauge fields C and C' are interchanged. If the boundary fields are equal, this is also true for the correlation functions f_X and g_X , $X = A_0, P$. In this case the signal of the numerical computation can be smoothed by averaging.

It is often more convenient to work with *off-diagonal bilinears*, especially if the isospin symmetry is explicitly broken or when we investigate currents build from different flavours. For those cases the basis $\{\tau^+, \tau^-, \tau^3\}$ is better suited than $\{\tau^1, \tau^2, \tau^3\}$. For an explicit representation see appendix B.1. For later purposes we count different flavours by $i, j = 1, 2, \dots$ which only differ by their mass or hopping parameter in our simulations. For an arbitrary bispinor $\psi(x)$ we now write $\psi = (\psi^i \psi^j)^T$ and the anti-bispinor accordingly for flavour i and j . In the new basis the time component of the axial current (2.37d) becomes

$$A_0^+ \equiv \bar{\psi} \gamma_0 \gamma_5 \frac{1}{2} \tau^+ \psi = \bar{\psi}^i \gamma_0 \gamma_5 \psi^j \equiv A_0^{ij}, \quad (3.54a)$$

$$A_0^- \equiv \bar{\psi} \gamma_0 \gamma_5 \frac{1}{2} \tau^- \psi = \bar{\psi}^j \gamma_0 \gamma_5 \psi^i \equiv A_0^{ji}. \quad (3.54b)$$

The very same construction applies to all other currents as well as to the boundary sources such that the definitions above become

$$f_A^{ij}(x_0) = -\frac{1}{2} \langle A_0^+(x_0) \mathcal{O}^- \rangle, \quad f_P^{ij}(x_0) = -\frac{1}{2} \langle P^+(x_0) \mathcal{O}^- \rangle, \quad (3.55a)$$

$$k_V^{ij}(x_0) = -\frac{1}{6} \langle V_k^+(x_0) \mathcal{Q}_k^- \rangle, \quad k_T^{ij}(x_0) = -\frac{1}{6} \langle T_{k0}^+(x_0) \mathcal{Q}_k^- \rangle, \quad (3.55b)$$

$$f_1^{ij} = -\frac{1}{3L^6} \langle \mathcal{O}'^+ \mathcal{O}^- \rangle, \quad k_1^{ij} = -\frac{1}{6L^6} \langle \mathcal{Q}'^+ \mathcal{Q}_k^- \rangle. \quad (3.55c)$$

Interchanging (i, j) corresponds to interchanging $(+, -)$. An explicit construction of these correlation functions is given in appendix B.5.

Transfer matrix formalism

Even if we are not interested in spectroscopy as we will work in physically small volume with length of approximately 0.5 fm, it is instructive to have a closer look at the transfer matrix formalism. The following discussion mainly follows that of [75]. We already noticed, eq. (3.14), that the Schrödinger functional can be represented as transition amplitude. For the more general case with fermions it now reads

$$\mathcal{Z}[C', C; \bar{\rho}', \rho'; \bar{\rho}, \rho; \bar{\eta}, \eta] \equiv \langle l' | e^{-T\mathbb{H}\mathbb{P}} | l \rangle, \quad (3.56)$$

for some initial and final state vector $|l\rangle$ and $|l'\rangle$. In case of vanishing fermion and gauge boundary fields both correspond to the vacuum state $|l_0\rangle$. In this quantum mechanical representation, the correlation functions in the pseudo-scalar channel are described by

$$f_X(x_0) = \mathcal{Z}^{-1} L^3 \frac{1}{2} \langle \Omega | e^{-(T-x_0)\mathbb{H}} \mathbb{P} \mathbb{X}_{\text{PS}}(x_0) e^{-x_0\mathbb{H}\mathbb{P}} | l_{\text{PS}} \rangle, \quad 0 < x_0 < T, \quad (3.57a)$$

$$f_1 = \mathcal{Z}^{-1} \frac{1}{2} \langle l'_{\text{PS}} | e^{-T\mathbb{H}\mathbb{P}} | l_{\text{PS}} \rangle, \quad X = A_0, P. \quad (3.57b)$$

Actually, in the Schrödinger picture of quantum mechanics this is quite obvious in view of figure 4. \mathbb{X}_{PS} is the operator in the Schrödinger picture that represents the pseudo-scalars A_0, P . The states $|l_{\text{PS}}\rangle, |l'_{\text{PS}}\rangle$ have quantum numbers of a pseudo-scalar meson without spatial momentum. The next step is to introduce a complete set of energy eigenstates $|n, q\rangle$,

$$\mathbb{H}|n, q\rangle = E_n^{(q)} |n, q\rangle, \quad \langle n', q' | n, q \rangle = \delta_{n', n} \delta_{q', q}, \quad (3.58)$$

where n enumerates the energy levels in the Hilbert space of quantum number q , e.g. $q = 0$ for the vacuum and $q = \text{PS}$ for the pseudo-scalar channel. This construction is possible because \mathbb{H} is known to be hermitean [35].¹⁴ After inserting the completeness relation $\mathbb{1} = \sum_{n, q} |n, q\rangle \langle n, q|$ at appropriate places and applying the eigenvalue equation (3.58), an expansion shows that the dominant contributions to the correlation functions are,

$$f_X(x_0) \approx -L^3 \frac{1}{2} \langle 0, 0 | \mathbb{X}_{\text{PS}} | 0, \text{PS} \rangle e^{-x_0 m_{\text{PS}}} \times \left\{ 1 + \eta_X^{\text{PS}} e^{-x_0 \Delta} + \eta_X^0 e^{-(T-x_0) m_G} \right\}, \quad (3.59a)$$

$$f_1 \approx \frac{1}{2} \rho^2 e^{-T m_{\text{PS}}}, \quad (3.59b)$$

with the following ratios of matrix elements

$$\eta_X^{\text{PS}} = \frac{\langle 0, 0 | \mathbb{X}_{\text{PS}} | 1, \text{PS} \rangle \langle 1, \text{PS} | 0, l_{\text{PS}} \rangle}{\langle 0, 0 | \mathbb{X}_{\text{PS}} | 0, \text{PS} \rangle \langle 0, \text{PS} | 0, l_{\text{PS}} \rangle}, \quad \eta_X^0 = \frac{\langle l_0 | 1, 0 \rangle \langle 1, 0 | \mathbb{X}_{\text{PS}} | 0, \text{PS} \rangle}{\langle l_0 | 0, 0 \rangle \langle 0, 0 | \mathbb{X}_{\text{PS}} | 0, \text{PS} \rangle}, \quad (3.60a)$$

$$\rho = \frac{\langle 0, \text{PS} | l_{\text{PS}} \rangle}{\langle 0, 0 | l_0 \rangle}. \quad (3.60b)$$

¹⁴ at least for the unimproved Wilson action

$m_{\text{PS}} = E_0^{(\text{PS})}$ is the mass of the ground state meson corresponding to the inserted operator \mathbb{X}_{PS} . $m_G = E_1^{(0)} - E_0^{(0)}$ is the mass of the 0^{++} glueball and $\Delta = E_1^{(\text{PS})} - E_0^{(\text{PS})}$ denotes the gap in the pseudo-scalar channel. Contributions which arise from even higher excited states decay even faster as x_0 and $T - x_0$ becomes large and henceforth are neglected. In lattice simulations one can choose the value of the quark masses by the corresponding hopping parameter within certain bounds. Thus, we can simulate a bunch of (unphysical) mesons and obtain the physical ones – usually denoted as $\text{PS} = \pi, B_s, \dots$ for instance – by extra-/interpolations. We will study such a quark mass dependence in small volume QCD in section 7.

Note that under the reasonable assumption $\eta_X^{\text{PS}} \approx \eta_X^0$ the contribution of the higher states in eq. (3.59a) are minimized at $x_0 = T/2$. This is also true for even higher excited state contributions and a result of the very construction. The effect of the boundary (source) fields decays with the distance from the boundary and hence becomes minimal in the middle of the SF cylinder.

Each lattice computation suffers from the presence of higher excited states as straightforward operator insertions do not map onto a special energy eigenstate. This behaviour can be improved by different techniques to increase the sensitivity (overlap) to the state one is looking for. In most cases the presence of the SF boundaries itself serves as a kind of improvement and gives a clear signal. Here, we will not use any improvements for the signal because we focus on effective ground state contributions which only correspond to m_{PS} in large volume simulations. Superfluously to say that the whole discussion applies to the vector channel as well. Now, the representation (3.59) immediately offers us a way to extract an effective mass.

Addendum

To systematically achieve on-shell $\mathcal{O}(a)$ -improvement in our quantities we build the *improved correlation functions* corresponding to (2.38) as

$$f_A^{\text{I}}(x_0) = f_A(x_0) + c_A(g_0) a \tilde{\partial}_0 f_{\text{P}}(x_0), \quad (3.61)$$

$$k_V^{\text{I}}(x_0) = k_V(x_0) + c_V(g_0) a \tilde{\partial}_0 k_{\text{T}}(x_0), \quad (3.62)$$

and analogous for g_A, l_V . The boundary-to-boundary correlation functions f_1 and k_1 as well as f_{P} are already at $\mathcal{O}(a^2)$. This is only true as long as residual $\mathcal{O}(a)$ terms that may result from improperly tuned improvement coefficients are negligible. Here the weak point is evidently the use of c_V, \tilde{c}_t and c_t which are only known in one-loop, respective two-loop order of perturbation theory. To what extend this may become noticeable cannot be estimated in advance and depends on the actual simulation. However, such residual $\mathcal{O}(a)$ terms would mix and could not be disentangled in the end. Furthermore, they could cancel to some extend among themselves accidentally, resulting in an even smaller contribution. So far no simulation is known where those coefficients have had a large impact to the outcome. But an even larger contribution in renormalized quantities is expected from the mass improvement coefficients b_χ at non-vanishing am_q . Thus, we focus on their non-perturbative estimation in section 6 to get reliable estimates for mass-dependent quantities.

3.7 Renormalization of correlation function

Yet, we only considered correlation functions of bare currents. Renormalized correlation functions are constructed straightforwardly as we already discussed more generally in section 2.2.2 on operator improvement. The renormalization for each flavour of SF boundary fields is given by,

$$[\zeta]_{\text{R}} = Z_{\zeta}(1 + b_{\zeta}am_{\text{q}})\zeta, \quad Z_{\zeta'} = Z_{\zeta}, \quad b_{\zeta'} = b_{\zeta}. \quad (3.63)$$

The same relation holds for the anti-fermion boundary fields $\bar{\zeta}$ and $\bar{\zeta}'$ which are identical to those given in (3.63). According to equation (2.44) the renormalized SF correlation functions read

$$[f_X(x_0)]_{\text{R}} = Z_X(1 + b_Xam_{\text{q}})Z_{\zeta}^2(1 + b_{\zeta}am_{\text{q}}) \times f_X(x_0), \quad X = A, P, \quad (3.64a)$$

$$[f_1]_{\text{R}} = Z_{\zeta}^4(1 + b_{\zeta}am_{\text{q}})^2 \times f_1, \quad (3.64b)$$

$$[k_X(x_0)]_{\text{R}} = Z_X(1 + b_Xam_{\text{q}})Z_{\zeta}^2(1 + b_{\zeta}am_{\text{q}}) \times k_X(x_0), \quad X = V, T, \quad (3.64c)$$

$$[k_1]_{\text{R}} = Z_{\zeta}^4(1 + b_{\zeta}am_{\text{q}})^2 \times k_1. \quad (3.64d)$$

An important point to note here is that the contributions from the boundary quark field renormalization and improvement will (by construction) always cancel in our observables such that we do not need to worry about them.

As in the continuum the renormalization constants Z_A, Z_V for dynamical Wilson fermions are not scale dependent. But due to the chiral symmetry breaking of the Wilson action at finite lattice spacing they get a finite renormalization on the lattice. Fortunately, this was already computed non-perturbatively in [77, 78] for $\beta \geq 5.2$,

$$Z_A(g_0^2) = 1 - 0.116g_0^2 + 0.011g_0^4 - 0.072g_0^6, \quad (3.65)$$

$$Z_V(g_0^2) = \frac{1 - 0.6715g_0^2 + 0.0388g_0^4}{1 - 0.5421g_0^2}, \quad (3.66)$$

and will always be used in our numerical analysis. Also the running of Z_P in the continuum was computed in the SF [29]. However, for our purpose we need to compute them directly in course of our numerical simulations. Hence this is a good point to discuss it now. From the representation of the pseudo-scalar current P in the SF

$$[f_P(x_0)]_{\text{R}} = Z_P(g_0^2, a\mu)(1 + b_P(g_0)am_{\text{q}})Z_{\zeta}^2(1 + b_{\zeta}am_{\text{q}})f_P(x_0), \quad (3.67)$$

we immediately conclude, that in order to estimate Z_P we need to simulate at vanishing mass m_{q} . Furthermore, we need to cancel the multiplicative renormalization of the boundary fields. According to (3.64) this is achieved if we divide by $\sqrt{f_1}$ and take $x_0 = T/2$ to reduce contaminations from higher excited states. We are left with the relation

$$\left[\frac{f_P(T/2)}{\sqrt{f_1}} \right]_{\text{R}} = Z_P(g_0^2, a\mu) \frac{f_P(T/2)}{\sqrt{f_1}}, \quad \mu = 1/L. \quad (3.68)$$

By enforcing the *tree-level normalization condition*

$$Z_P(g_0^2, L/a) \frac{f_P(T/2)}{\sqrt{f_1}} = \frac{f_P(T/2)}{\sqrt{f_1}} \Big|_{g_0^2=0} \Leftrightarrow Z_P(0, L/a) \equiv 1, \quad (3.69)$$

for all L/a , the definition for the renormalization factor of the pseudo-scalar density reads

$$Z_P(g_0^2, L/a) = c(L/a) \frac{\sqrt{3f_1}}{f_P(T/2)}, \quad c(L/a) = \frac{f_P(T/2)}{\sqrt{f_1}} \Big|_{g_0^2=0}. \quad (3.70)$$

The tree-level normalization constant c can be evaluated in perturbation theory and is independent of the dynamical flavour content of the theory. The additional (colour) factor $\sqrt{3}$ appears due to a redefinition of the constant c between reference [21] and [79], such that the c in (3.70) approaches one in the limit $L/a \rightarrow \infty$. In table 17 we list some values of $c(L/a)$ that are relevant for this work. Note that $Z_P = 1$ at tree-level is a desired condition as all renormalization constants in the minimal subtraction scheme take the form $Z(g_0^2, a\mu) = 1 + Z^{(1)}g_0^2 + \mathcal{O}(g_0^4)$ and the Schrödinger functional is not special to one particular renormalization scheme.

The PCAC quark mass

For further purposes we define the notion of the PCAC quark mass here because we recently defined renormalization constants which appear in its renormalization. It follows from the PCAC relation as known in the continuum. For an elaborated discussion see for instance [45, 135]. In terms of improved SF correlation functions (3.61) the bare PCAC quark mass is given by

$$m(x_0) = \frac{\tilde{\partial}_0 f_A^I(x_0) + ac_A \partial_0^* \partial_0 f_P(x_0)}{2f_P(x_0)}. \quad (3.71)$$

This is not fully improved at non-vanishing quark mass am_q because the additive mass contributions $b_\chi am_q$ are still absent. Note that for convenience we do not just let $\tilde{\partial}_0$ act on f_A^I , eq. (3.61), because this includes contributions at $x_0 \pm 2a$ and increases the contaminations by higher state contributions. But due to the associated Ward–Takahashi identity its signal is much better behaved than that of effective masses introduced in the next subsection. The *renormalized on-shell $\mathcal{O}(a)$ improved current quark mass* finally reads

$$m_R(x_0) = \frac{Z_A(g_0)(1 + b_A(g_0)am_q)}{Z_P(g_0, a\mu)(1 + b_P(g_0)am_q)} m(x_0) + \mathcal{O}(a^2). \quad (3.72)$$

3.8 Effective masses

Taking the logarithm of eq. (3.59a) enables us to separate the mass term ($-x_0 m_{PS}$) from other contributions. Hence, in leading order we obtain the positive mass m_{PS} by applying the negative time derivative to it. This gives the usual definition of an effective mass in lattice QCD and is not special to the SF as the construction follows directly in the transfer matrix formalism. Again we have to choose a lattice representative for the time derivative which gives the following three

definitions,¹⁵

$$\Gamma(x_0 + \frac{a}{2}) = -\partial_0 \ln [f_X(x_0)] \equiv \frac{1}{a} \ln \left[\frac{f_X(x_0)}{f_X(x_0+a)} \right], \quad (3.73a)$$

$$\Gamma(x_0 - \frac{a}{2}) = -\partial_0^* \ln [f_X(x_0)] \equiv \frac{1}{a} \ln \left[\frac{f_X(x_0-a)}{f_X(x_0)} \right], \quad (3.73b)$$

$$\Gamma(x_0) = -\tilde{\partial}_0 \ln [f_X(x_0)] \equiv \frac{1}{2a} \ln \left[\frac{f_X(x_0-a)}{f_X(x_0+a)} \right] = \frac{1}{2} [\Gamma(x_0 - \frac{a}{2}) + \Gamma(x_0 + \frac{a}{2})], \quad (3.73c)$$

the *forward*, *backward* and *symmetric effective mass*, respectively. We denote them by the same symbol because the argument allows us to distinguish them. These definitions are quite general and if we have to specify a specific channel, we will use PS or V as subscript. Note the shift symmetry between the first two definitions. With $x_0 \mapsto x_0 - a$ in the argument of the forward effective mass, the backward effective mass is obtained and thus they are equal.

Let us have a look at their leading order behaviour by inserting (3.59a). Clearly the overall normalization, i.e. the constant in front, cancel when building ratios of different time slices,

$$\begin{aligned} \Gamma(x_0 + \frac{a}{2}) &= \frac{1}{a} \ln \left[\frac{f_X(x_0)}{f_X(x_0+a)} \right] \\ &\simeq \frac{1}{a} \ln \left[\frac{e^{am_{\text{PS}}} \frac{1 + \eta_X^{\text{PS}} e^{-x_0 \Delta_{\text{PS}}} + \eta_X^0 e^{-(T-x_0)m_G}}{1 + \eta_X^{\text{PS}} e^{-(x_0+a)\Delta_{\text{PS}}} + \eta_X^0 e^{-(T-x_0-a)m_G}}}{1} \right] \\ &= m_{\text{PS}} + \frac{1}{a} \left\{ \eta_X^{\text{PS}} e^{-x_0 \Delta_{\text{PS}}} (1 - e^{-a\Delta_{\text{PS}}}) + \eta_X^0 e^{-(T-x_0)m_G} (1 - e^{am_G}) \right\} + \mathcal{O}(\eta^2) \\ &= m_{\text{PS}} + \frac{2}{a} \sinh(a\Delta_{\text{PS}}/2) e^{-\frac{a\Delta_{\text{PS}}}{2}} \eta_X^{\text{PS}} e^{-x_0 \Delta_{\text{PS}}} \\ &\quad - \frac{2}{a} \sinh(am_G/2) e^{-\frac{am_G}{2}} \eta_X^0 e^{-(T-x_0)m_G} + \mathcal{O}(\eta^2), \end{aligned} \quad (3.74)$$

Here $\mathcal{O}(\eta^2)$ encloses contributions in η_X^{PS} and η_X^0 that arise from approximating the logarithm by $\ln(1 + \eta) = \eta + \mathcal{O}(\eta^2)$. For $x_0 \mapsto x_0 - a$ one gets from eq. (3.74),

$$\begin{aligned} \Gamma(x_0 - \frac{a}{2}) &= m_{\text{PS}} + \frac{2}{a} \sinh(a\Delta_{\text{PS}}/2) e^{\frac{a\Delta_{\text{PS}}}{2}} \eta_X^{\text{PS}} e^{-x_0 \Delta_{\text{PS}}} \\ &\quad - \frac{2}{a} \sinh(am_G/2) e^{-\frac{am_G}{2}} \eta_X^0 e^{-(T-x_0)m_G} + \mathcal{O}(\eta^2), \end{aligned} \quad (3.75)$$

and the symmetrized mass follows to

$$\begin{aligned} \Gamma(x_0) &= m_{\text{PS}} + \frac{1}{a} 2 \sinh(a\Delta_{\text{PS}}/2) \cosh(a\Delta_{\text{PS}}/2) \eta_X^{\text{PS}} e^{-x_0 \Delta_{\text{PS}}} \\ &\quad - \frac{1}{a} 2 \sinh(am_G/2) \cosh(am_G/2) \eta_X^0 e^{-(T-x_0)m_G} + \mathcal{O}(\eta^2) \\ &= m_{\text{PS}} + \frac{1}{a} \sinh(a\Delta_{\text{PS}}) \eta_X^{\text{PS}} e^{-x_0 \Delta_{\text{PS}}} - \frac{1}{a} \sinh(am_G) \eta_X^0 e^{-(T-x_0)m_G} + \mathcal{O}(\eta^2) \end{aligned} \quad (3.76)$$

where $2 \sinh(x) \cosh(x) = \sinh(2x)$ has been used. As expected the leading order term in all three definitions is dominated by the ground state mass. We can go one step further and expand

¹⁵ Just for notational convenience we use the standard cf.s f_X instead of the improved ones, f_X^1 . In actual calculations we will always use the improved ones if not stated otherwise.

the explicit a -dependent terms using $\sinh(a) \simeq a(1 + a^2/6)$ and $\exp(a) \simeq 1 + a$. It follows

$$\Gamma = m_{\text{PS}} + \Delta_{\text{PS}} \eta_X^{\text{PS}} e^{-x_0 \Delta_{\text{PS}}} - m_G \eta_X^0 e^{-(T-x_0)m_G} + \mathcal{O}(a^2) + \mathcal{O}(\eta^2), \quad (3.77)$$

for each definition, i.e. differences become visible at order a^2 to the leading contributions, which in the middle of the SF cylinder is given by

$$\Gamma = m_{\text{PS}} + \Delta_{\text{PS}} \eta_X^{\text{PS}} e^{-\Delta_{\text{PS}} T/2} - m_G \eta_X^0 e^{-m_G T/2}. \quad (3.78)$$

By now it should be clear why we call this an effective mass. There are always contaminations from other states. However, a simple fit to the time dependence of the effective mass would allow to separate the ground state from higher contributions. Depending on the real problem that is investigated one sometimes also needs more sophisticated (variational) techniques like that proposed in [76] for example.

Using an asymptotic expansion of the transfer matrix we have shown that in physically large volumes (~ 2 fm) and small lattice spacings the definitions (3.73) allow to extract spectral quantities like m_{PS} , Δ_{PS} and so on. But our focus is on computations in physically small volume such that a clear separation becomes impossible and indeed is not even needed. However, our main concern is about the effective energies Γ in eq. (3.73) which are renormalized spectral quantities and hence are well-suited for the strategy of a non-perturbative matching of QCD and HQET, see section 4.4.

3.9 Running of matrix elements

An exact matching condition to relate hadronic matrix elements in any two different schemes was first suggested in [129]. It makes use of the universality of the renormalization group invariant matrix element as defined in eq. (1.18). The running of such an operator could be carried out non-perturbatively in the Schrödinger functional scheme. Starting at a low-energy scale μ one can reach a perturbative scale μ_{pt} by a recursive finite-size scaling technique. Hence one has to compute the matrix element on successive scales $\mu_i = 2^i \mu$, $i = 1, \dots, n$, from the hadronic scale $\mu \equiv \mu_0$ to the high energy scale $\mu_n \equiv \mu_{\text{pt}}$. Starting from the identity

$$\begin{aligned} \Phi_{\text{RGI}} &= \Phi_{\text{SF}}(\mu) \times \frac{\Phi_{\text{RGI}}}{\Phi_{\text{SF}}(\mu)} \\ &= \Phi_{\text{SF}}(\mu) \times \frac{\Phi_{\text{SF}}(\mu_1)}{\Phi_{\text{SF}}(\mu)} \frac{\Phi_{\text{SF}}(\mu_2)}{\Phi_{\text{SF}}(\mu_1)} \cdots \frac{\Phi_{\text{SF}}(\mu_{\text{pt}})}{\Phi_{\text{SF}}(\mu_{n-1})} \times \frac{\Phi_{\text{RGI}}}{\Phi_{\text{SF}}(\mu_{\text{pt}})}, \end{aligned} \quad (3.79)$$

one splits the estimation of Φ_{RGI} (from the right to the left) in a perturbative part,

$$\frac{\Phi_{\text{RGI}}}{\Phi_{\text{SF}}(\mu_{\text{pt}})} = [2b_0 \bar{g}_{\text{SF}}^2(\mu_{\text{pt}})]^{\frac{-\gamma_0}{2b_0}} \exp \left\{ - \int_0^{\bar{g}_{\text{SF}}(\mu_{\text{pt}})} dg \left[\frac{\gamma_{\text{SF}}(g)}{\beta_{\text{SF}}(g)} - \frac{\gamma_0}{b_0 g} \right] \right\}, \quad (3.80)$$

and the non-perturbative running. The latter gives estimates for the analytic expression

$$\frac{\Phi_{\text{SF}}(\mu_{i+1})}{\Phi_{\text{SF}}(\mu_i)} = \left[\frac{\bar{g}_{\text{SF}}(\mu_i)}{\bar{g}_{\text{SF}}(\mu_{i+1})} \right]^{-\gamma_0/b_0} \exp \left\{ - \int_{\bar{g}_{\text{SF}}(\mu_{i+1})}^{\bar{g}_{\text{SF}}(\mu_i)} dg \frac{\gamma_{\text{SF}}(g)}{\beta_{\text{SF}}(g)} \right\}, \quad (3.81)$$

by means of non-perturbative lattice computations of the matrix elements $\Phi_{\text{SF}}(\mu_i)$ and $\Phi_{\text{SF}}(\mu_{i+1})$ while keeping g_0^2 fixed. A final computation of $\Phi_{\text{SF}}(\mu)$ at the low energy scale μ in a physically large volume establish the contact to experiment.

In this way one solves the renormalization group running of an unprotected operator in a purely non-perturbative fashion without the necessity of knowing the β -function and the anomalous dimension of the operator to high orders in perturbation theory. Furthermore, the regulator is removed in each ratio of the matrix elements, such that only the scheme and scale dependence is left. By this procedure a reliable error estimate can be given as the uncertainty in each step of the recursion is known and the remaining contribution from the perturbative approximation in the integration is negligible compared to the statistical and systematical errors of the former.

The bounds for the high- and low-energy scale μ_0 and μ_{pt} are dictated by the tools at hand, the Schrödinger functional scheme in our case. Here, the energy scale is given by the inverse box length, $\mu = 1/L$. Therefore, a high-energy scale corresponds to a small box and one has to make sure to reach the perturbative region. The bound for the low-energy scale is given by large volume simulations, where computations become expensive in terms of computer time with the usual problems encountered in lattice QCD simulations. Then the renormalized parameters of QCD are known at any scale $\mu \in [\mu_0, \infty)$. But to connect high- and low-energy physics one usually wants to fix the renormalization about μ_0 as it is the most distant point that can be reached through numerical simulations.

After a non-perturbative calculation of the RGI matrix element was performed it is an easy task to get the corresponding matrix element in the $\overline{\text{MS}}$ scheme by inverting eq. (1.18),

$$\Phi_{\overline{\text{MS}}}(\mu) = \Phi_{\text{RGI}}[2b_0\bar{g}_{\overline{\text{MS}}}^2(\mu)]^{\gamma_0/(2b_0)} \exp \left\{ \int_0^{\bar{g}_{\overline{\text{MS}}}(\mu)} dg \left[\frac{\gamma(g)}{\beta(g)} - \frac{\gamma_0}{b_0 g} \right] \right\}, \quad (3.82)$$

as long as the corresponding anomalous dimension is known to higher orders and to a scale where perturbation theory is still feasible.

4 Heavy Quark Effective Theory and its non-perturbative matching to QCD

The Heavy Quark Effective Theory (HQET) is an expansion of QCD about the inverse of the heavy quark mass $1/m_b$ [80]. It is renormalizable at any finite order n in $1/m_b^n$ by means of power counting, i.e. after fixing the order of the expansion there exists a finite number of parameters that need to be renormalized in order to render the effective theory finite. HQET is a standard phenomenological tool which simplifies the QCD dynamics in the limit of large masses like that of the c- or b-quark while still treating the dynamics of light quarks and gluons exactly.

We use HQET to subleading order ($n = 1$) and hence refer to the part of the Lagrangian density given by,

$$\mathcal{L}_{\text{HQET}}(x) = \bar{\psi}_h(x) \left[D_0 + m - \omega_{\text{kin}} \mathbf{D}^2 - \omega_{\text{spin}} \boldsymbol{\sigma} \mathbf{B} \right] \psi_h(x) + \mathcal{O}(1/m^2) , \quad (4.1)$$

that accommodates the heavy quark degrees of freedom. We pass the derivation to focus on its implementation within our matching strategy. For details see [81, 82, 153, 154] and references therein. To distinguish the heavy quark fields ψ_h and $\bar{\psi}_h$ in HQET from their counterparts in QCD, we conveniently denote a relativistic heavy quark and anti-quark by ψ_b and $\bar{\psi}_b$, respectively. Every quark field that is not heavy will be called light, ψ_l or $\bar{\psi}_l$. Before we treat HQET more systematically on the lattice, some remarks are in order:

- The zeroth-order approximation in (4.1) which describes the dynamics in the limit of an infinitely heavy quark mass ($1/m_b \rightarrow 0 \Rightarrow \omega_{\text{kin}} = \omega_{\text{spin}} \rightarrow 0$) is called *static approximation* and the quark fields fulfill

$$\psi_h = P_+ \psi_b , \quad \bar{\psi}_h = \bar{\psi}_b P_- , \quad P_{\pm} = (1 \pm \gamma_0)/2 . \quad (4.2)$$

Here the projectors map onto the 'large' components of the usual four-component spinor field leaving a 2-component spinor behind. Accordingly, we call ψ_h a *static quark*. Thus the leading Lagrangian only contains the time component of the Dirac-operator D_0 and the bare heavy quark mass m . It immediately follows that the static quark is at rest, only propagating through time while interacting with the light degrees of freedom of the theory. The static Lagrangian has an enhanced symmetry. Apart from the usual ones, first there is the *heavy quark spin symmetry*, the invariance under global $\text{SU}(2)$ rotations,

$$\psi_h(x) \longrightarrow V \psi_h(x) , \quad \bar{\psi}_h(x) \longrightarrow \bar{\psi}_h(x) V^{-1} , \quad V = \exp(-i\phi_i \epsilon_{ijk} \sigma_{jk}) , \quad (4.3)$$

for transformation parameters ϕ_i and $[\sigma_i, \sigma_j] = i\epsilon_{ijk} \sigma_k$. Secondly, we have a *local conservation of the b-quark number* expressed by the invariance under time-independent, local phase transformations $\eta(\mathbf{x})$, with

$$\psi_h(x) \longrightarrow e^{i\eta(\mathbf{x})} \psi_h(x) , \quad \bar{\psi}_h(x) \longrightarrow \bar{\psi}_h(x) e^{-i\eta(\mathbf{x})} . \quad (4.4)$$

- At subleading order two dimension five operators appear, the (spatial) kinetic term \mathbf{D}^2 and the spin splitting term $\boldsymbol{\sigma}\mathbf{B}$ with the chromomagnetic field \mathbf{B} . Their coefficients are $\omega_{\text{kin}} = \omega_{\text{spin}} = 1/(2m)$ in the classical theory and get renormalized at the quantum level. Due to reparametrization invariance [83, 84, 85] the kinetic operator in the effective theory is protected against renormalization.

4.1 The physical picture

Equation (4.1) tells us that in the static approximation, the spin of the heavy quark and the total angular momentum of the light degrees of freedom, j , are separately conserved by the strong interaction. The dynamics does not depend on the spin and mass of the heavy quark, thus hadronic states are classified by the quantum numbers of the light degrees of freedom. From the spin symmetry it follows the existence of a doublet of degenerated states with total spin $J = j \pm \frac{1}{2}, j \neq 0$. We consider ground state mesons containing one heavy quark and the light degrees of freedom carry the quantum numbers of the light anti-quark. The degenerated states are the pseudoscalar ($J = 0$) and the vector ($J = 1$) mesons. Whether they contain a b- or c-quark these are B, B^* or D, D^* mesons, respectively. The static approximation of HQET is flavour blind and we cannot distinguish between these four states. This changes if we go to subleading order. When turning on the interactions in (4.1) by working at finite $1/m \gtrsim 0$, the spin interaction now allows to separate $J = 0$ and $J = 1$ states. Furthermore, the fact that we have to choose either $1/m_b$ or $1/m_c$ naturally distinguishes the B - from the D -channel. In lattice QCD we are free to vary the heavy quark masses within certain limits such that we are able to cover both regions.

The evidence that HQET is a good approximation to mesons containing a b- or c-quark is given by experiment. In table 3 we list some results from the Particle Data Group (PDG) [1] which can phenomenologically be explained by the following expansions for masses in HQET to subleading order [86],

$$m_B = m_b + \bar{\Lambda} - \frac{\lambda_1}{2m_b} - \frac{3\lambda_2(m_b)}{2m_b}, \quad m_{B^*} = m_b + \bar{\Lambda} - \frac{\lambda_1}{2m_b} + \frac{\lambda_2(m_b)}{2m_b}. \quad (4.5)$$

Here λ_1 and λ_2 are properly normalized matrix elements in HQET with operator insertions \mathcal{O}_{kin} and $\mathcal{O}_{\text{spin}}$, respectively. The dependence of the latter on m_b signals that it is renormalized in the effective theory at scale $\mu = m_b$. The parameter $\bar{\Lambda}$ determines the effective mass of meson states in HQET due to the strong interaction of the heavy quark with the light degrees of freedom. Because there is no unique non-perturbative definition of the mass m_b , the binding energy $\bar{\Lambda} = m_B - m_b + \mathcal{O}(1/m_b)$ cannot be obtained as a prediction of HQET. Hence $\bar{\Lambda}$ has an ambiguity of order $\Lambda_{\text{QCD}} \sim (200 - 300) \text{ MeV} \ll m_b, m_c$. For dimensional reasons it also follows that the subleading contributions λ_1, λ_2 are of order Λ_{QCD}^2 . All particles within the same spin-flavour multiplet like the B, B^*, D and D^* states share the same value of $\bar{\Lambda}$. Therefore the corresponding expressions for D and D^* are obtained by substituting $m_b \rightarrow m_c$ in eq. (4.5). It is convenient and part of our main strategy to separately treat the spin splitting term. For this purpose one introduces

Hadron	Mass (MeV)	Quark Content	Hadron	Mass (MeV)	Quark Content	I	J^P
\bar{B}^-	5279.15 ± 0.31	$b\bar{u}$	D^0	1864.84 ± 0.17	$c\bar{u}$	$\frac{1}{2}$	0^-
\bar{B}^0	5279.53 ± 0.33	$b\bar{d}$	D^+	1869.62 ± 0.20	$c\bar{d}$	$\frac{1}{2}$	0^-
\bar{B}_s^0	5366.30 ± 0.60	$b\bar{s}$	D_s^+	1968.49 ± 0.34	$c\bar{s}$	0	0^-
\bar{B}^{*-}	5325.10 ± 0.50	$b\bar{u}$	D^{*0}	2006.97 ± 0.19	$c\bar{u}$	$\frac{1}{2}$	1^-
\bar{B}^{*0}		$b\bar{d}$	D^{*+}	2010.27 ± 0.17	$c\bar{d}$		
\bar{B}_s^*	5412.8 ± 1.3	$b\bar{s}$	D_s^{*+}	2112.30 ± 0.50	$c\bar{s}$	0	1^-

Table 3: Phenomenology and masses of heavy-light mesons [1].

the *spin-averaged mass* m_B^{av} and *spin-splitting* Δm_B by

$$m_B^{\text{av}} \equiv \frac{1}{4}(m_B + 3m_{B^*}) \quad \Delta m_B \equiv m_B - m_{B^*} \quad (4.6a)$$

$$= m_b + \bar{\Lambda} - \frac{\lambda_1}{2m_b}, \quad = \frac{4\lambda_2(m_b)}{2m_b}. \quad (4.6b)$$

As well as their counterparts in (4.5) they are expected to be violated by terms of order $\Lambda_{\text{QCD}}^3/m_b^2$. The experimental values

$$m_{B^*} - m_B = (45.78 \pm 0.35) \text{ MeV}, \quad (4.7)$$

$$m_{D^{*0}} - m_{D^0} = (145.421 \pm 0.010) \text{ MeV}, \quad m_{D^{*+}} - m_{D^+} = (140.64 \pm 0.10) \text{ MeV}, \quad (4.8)$$

are reasonable small compared to the masses in table 3 to expect HQET to describe QCD in the high energy region if a heavy quark is involved. Indeed, by HQET to subleading order a much better prediction is given by

$$m_{B^*}^2 - m_B^2 \equiv (m_{B^*} - m_B)(m_B + m_{B^*}) = 4\lambda_2(m_b) + \mathcal{O}(\Lambda_{\text{QCD}}/m_b), \quad (4.9)$$

where $m_B + m_{B^*} = 2m_b + \mathcal{O}(\Lambda_{\text{QCD}})$ has been used. Here the leading term is mass-independent up to a mild intrinsic dependence on the renormalization scale. Hence one expects $m_{B^*}^2 - m_B^2 \approx m_{D^{*0}}^2 - m_{D^0}^2$, which is very well fulfilled,

$$m_{B^*}^2 - m_B^2 \approx 0.49 \text{ GeV}^2, \quad m_{D^{*0}}^2 - m_{D^0}^2 \approx 0.56 \text{ GeV}^2. \quad (4.10)$$

As λ_1, λ_2 are non-perturbative parameters of HQET they must be determined by first principle computations.

4.2 HQET on the lattice

If we split the total action \mathcal{S} into a relativistic part \mathcal{S}_{rel} and the heavy quark part $\mathcal{S}_{\text{HQET}}$ we can write to some order n of the HQET expansion,

$$\mathcal{S}_{\text{HQET}} = a^4 \sum_x \mathcal{L}_{\text{HQET}}(x) = a^4 \sum_x \left\{ \mathcal{L}_{\text{stat}}(x) + \sum_{i=1}^n \mathcal{L}^{(i)}(x) \right\}, \quad (4.11a)$$

$$\mathcal{L}_{\text{stat}}(x) = \bar{\psi}_h(x) [D_0 + \delta m] \psi_h(x), \quad (4.11b)$$

$$\mathcal{L}^{(i)} = \sum_j \omega_j^{(i)} \mathcal{L}_j^{(i)}. \quad (4.11c)$$

Here, $\mathcal{L}_{\text{stat}}$ is the Lagrangian of the static theory with mass counterterm δm . By power counting one gets expansion coefficients $\omega_j^{(i)}$, i.e. the bare parameters of the effective theory, that multiply the various operators $\mathcal{O}_j^{(i)}$, $j = 1, 2, \dots, j_i$ of mass dimension $4 + i$ which appear in the Lagrangian $\mathcal{L}^{(i)}$ at order $i = 1, \dots, n$. They are functions of the bare coupling and heavy mass. According to this, the Boltzmann factor in the Euclidean path integral can be written as

$$\begin{aligned} e^{-(\mathcal{S}_{\text{rel}} + \mathcal{S}_{\text{HQET}})} &= \exp \left\{ - \left(\mathcal{S}_{\text{rel}} + a^4 \sum_x \mathcal{L}_{\text{stat}}(x) \right) \right\} \times \\ &\quad \left\{ 1 - a^4 \sum_x \mathcal{L}^{(1)}(x) + \frac{1}{2} \left[a^4 \sum_x \mathcal{L}^{(1)}(x) \right]^2 - a^4 \sum_x \mathcal{L}^{(2)}(x) + \dots \right\}. \end{aligned} \quad (4.12)$$

Only the leading non-trivial term contributes at order $1/m_b$. Following Symanzik's idea of an effective lattice theory the path integral can be taken with respect to the static action,

$$\mathcal{S} \equiv \mathcal{S}_{\text{rel}} + \mathcal{S}_{\text{stat}}, \quad \mathcal{S}_{\text{stat}} \equiv a^4 \sum_x \mathcal{L}_{\text{stat}}(x). \quad (4.13a)$$

Then $(1/m_b)$ -corrections appear as insertions of local operators $\mathcal{O}_j^{(1)}(x)$ and insertions $\delta \mathcal{O}_j^{(1)}(x)$ into correlation functions. In shorthand notation with $\phi \in \{U, \psi_l, \bar{\psi}_l, \psi_h, \bar{\psi}_h\}$ this reads

$$\langle \mathcal{O} \rangle_{\text{HQET}} = \mathcal{Z}^{-1} \int \mathcal{D}[\phi] \left\{ \mathcal{O}_{\text{stat}} + \delta \mathcal{O}^{(1)} \right\} \left\{ 1 - a^4 \sum_x \mathcal{L}^{(1)}(x) \right\} e^{-\mathcal{S}_{\text{rel}} - \mathcal{S}_{\text{stat}}}. \quad (4.13b)$$

This is the essential definition of HQET up to and including order $1/m_b$ as it provides renormalizability by the continuum limit and preserves the continuum asymptotic expansion in $1/m_b$. To complete its definition at the classical level we summarize what was already said:

$$\mathcal{L}_{\text{stat}}(x) = \bar{\psi}_h(x) D_0 \psi_h(x), \quad \mathcal{L}^{(1)}(x) = -\omega_{\text{spin}} \mathcal{O}_{\text{spin}}(x) - \omega_{\text{kin}} \mathcal{O}_{\text{kin}}(x), \quad (4.13c)$$

$$\mathcal{O}_{\text{kin}}(x) = \bar{\psi}_h(x) \mathbf{D}^2 \psi_h(x), \quad \omega_{\text{kin}} = 1/(2m_b), \quad (4.13d)$$

$$\mathcal{O}_{\text{spin}}(x) = \bar{\psi}_h(x) \boldsymbol{\sigma} \cdot \mathbf{B} \psi_h(x), \quad \omega_{\text{spin}} = 1/(2m_b). \quad (4.13e)$$

The mass counterterm δm appearing in (4.11b) is usually set to zero and appears later on as an overall energy shift m_{bare} . The expectation value (4.13b) can now be written as

$$\langle \mathcal{O} \rangle = \langle \mathcal{O} \rangle_{\text{stat}} + \omega_{\text{kin}} a^4 \sum_x \langle \mathcal{O} \mathcal{O}_{\text{kin}}(x) \rangle_{\text{stat}} + \omega_{\text{spin}} a^4 \sum_x \langle \mathcal{O} \mathcal{O}_{\text{spin}}(x) \rangle_{\text{stat}} \quad (4.14a)$$

$$\equiv \langle \mathcal{O} \rangle_{\text{stat}} + \omega_{\text{kin}} \langle \mathcal{O} \rangle_{\text{kin}} + \omega_{\text{spin}} \langle \mathcal{O} \rangle_{\text{spin}} , \quad (4.14b)$$

because $\mathcal{O} = \mathcal{O}^{(0)} + \delta \mathcal{O}^{(1)} = \mathcal{O}_{\text{stat}} + \delta \mathcal{O}^{(1)}$. Combining $\delta \mathcal{O}^{(1)}$ and $\mathcal{L}^{(1)}$ produces terms of order $n = 2$ that are to be neglected. Some final comments on this read

- That the outcome is renormalizable and the expansion survives the removal of the regulator is not automatic for an effective theory. Chiral perturbation theory (ChPT) for example shares these properties but non-relativistic QCD (NRQCD) for instance does not because the $\mathcal{O}(1/m)$ term remains in the Boltzmann weight factor to compute the path-integral. Thus new divergences occur in perturbation theory at each order of the loop expansion which renders the theory not renormalizable.
- Opposed to ChPT, interactions are not turned off when reaching the leading order approximation in HQET, i.e. the static theory in the limit $1/m \rightarrow 0$. One still needs the lattice formulation of HQET to evaluate non-perturbatively the static approximation.
- Power counting arguments do not distinguish $\mathcal{O}(a)$ and $\mathcal{O}(1/m)$ expansions because both have a mass dimension of minus one. So both have to be considered as a single expansion in terms of the dimension of the local fields. By starting from a set of operators obtained by a formal $1/m$ expansion in the continuum, they will mix with all operators of the same or lower dimensions which are allowed by the lattice and not only by the continuum symmetries. Hence one starts from the beginning with the complete basis of operators compatible with the lattice symmetries and power counting with $a = \mathcal{O}(1/m)$. This means that to go to order $1/m$ in HQET it is crucial that \mathcal{S}_{rel} is $\mathcal{O}(a)$ improved.
- Terms that have to be taken into account are restricted by the HQET lattice symmetries that correspond to the two symmetries mentioned in the beginning and the 3-dimensional cubic group as remnant of the hyper-cubic group of the lattice. Since our interest is in on-shell quantities, keeping all fields at non-zero distance we can again reduce the set of operators to be considered by the equations of motion derived from $\mathcal{S}_{\text{rel}} + \mathcal{S}_{\text{stat}}$.

Renormalization of HQET to subleading order involves renormalization of certain dimension five operators which for simple dimensional reasons read

$$\mathcal{O}_{\text{R}}^{d=5} = Z_{\mathcal{O}} \left[\mathcal{O}^{d=5} + \sum_k c_k \mathcal{O}_k^{d=4} \right] , \quad \text{with} \quad c_k = \frac{1}{a} \sum_{i=0}^{\infty} c_k^{(i)} (g_0^2)^i . \quad (4.15)$$

In practice such perturbative computations have to be restricted to some finite order loop l in g_0^2 and the error in the truncation is

$$\Delta c_k \sim \frac{g_0^{2(l+1)}}{a} \sim \frac{1}{a[\ln(a\Lambda)]^{l+1}} \xrightarrow{a \rightarrow 0} \infty, \quad (4.16)$$

showing that the continuum limit does not exist and the theory *is not perturbatively renormalizable*. This underscores the necessity for a non-perturbative renormalization procedure for HQET. But before we explain our strategy in more detail in section 4.4 we first need to introduce some HQET observables in the Schrödinger functional framework.

4.3 $1/m_b$ -expansion and renormalization of heavy-light currents

For our concerns the most important correlation functions are given by the time component of the heavy-light axial current and the space components of the heavy-light vector current. In the notation from above we thus have $\mathcal{O} \in \{A_0^{\text{stat}}, V_k^{\text{stat}}\}$ and $\delta\mathcal{O} \in \{\delta A_0^{\text{stat}}, \delta V_k^{\text{stat}}\}$ to represent the local fields in the effective theory. Their renormalization reads

$$A_0^{\text{HQET}}(x) = Z_A^{\text{HQET}} [A_0^{\text{stat}}(x) + c_A^{\text{HQET}} \delta A_0^{\text{stat}}(x)], \quad (4.17a)$$

$$V_k^{\text{HQET}}(x) = Z_V^{\text{HQET}} [V_k^{\text{stat}}(x) + c_V^{\text{HQET}} \delta V_k^{\text{stat}}(x)], \quad (4.17b)$$

with the definitions

$$A_0^{\text{stat}}(x) = \bar{\psi}_1(x) \gamma_0 \gamma_5 \psi_h(x), \quad \delta A_0^{\text{stat}}(x) = \bar{\psi}_1(x) \gamma \mathbf{D} \gamma_5 \psi_h(x), \quad (4.17c)$$

$$V_k^{\text{stat}}(x) = \bar{\psi}_1(x) \gamma_k \psi_h(x), \quad \delta V_k^{\text{stat}}(x) = -\bar{\psi}_1(x) \gamma \mathbf{D} \gamma_k \psi_h(x). \quad (4.17d)$$

On the lattice we take the (contracted) left-acting spatial covariant derivative to be $\gamma \mathbf{D} = \gamma_i (\overleftarrow{\nabla}_i + \overleftarrow{\nabla}_i^*)/2$ and the minus sign in (4.17d) as our convention. By this choice for $V_k^{\text{stat}}, \delta V_k^{\text{stat}}$, they are exactly related to $A_0^{\text{stat}}, \delta A_0^{\text{stat}}$ by a spin rotation. The coefficients $\omega_{\text{kin}}, \omega_{\text{spin}}, Z_A^{\text{HQET}}, c_A^{\text{HQET}}, Z_V^{\text{HQET}}$ and c_V^{HQET} are functions of g_0 and am_b . These are the bare parameters of the effective theory and together with an overall energy shift m_{bare} are sufficient to absorb all divergences in the effective theory to order $1/m_b$. As formerly mentioned we can either take finite m_{bare} with vanishing mass counterterm $\delta m = 0$ or vice versa. By power counting the parameters are decomposed into

$$Z_X^{\text{HQET}} = Z_X^{\text{stat}} + Z_X^{(1)}, \quad c_X^{\text{HQET}} = c_X^{(1)}, \quad \omega_{\text{kin}} = 1/(2m_b), \quad (4.18a)$$

$$m_{\text{bare}} = m_{\text{bare}}^{\text{stat}} + m_{\text{bare}}^{(1)}, \quad \omega_{\text{spin}} = 1/(2m_b), \quad (4.18b)$$

where Z_X^{stat} and $m_{\text{bare}}^{\text{stat}}$ are $\mathcal{O}(1/m_b^0)$, all other $\mathcal{O}(1/m_b^1)$. Terms like $\omega_{\text{kin}} c_X^{\text{HQET}} \propto 1/m_b^2$ are to be dropped in the expansion to first order. We also have an $\mathcal{O}(a)$ -improvement term in the static approximation, which introduces a term proportional to some c_X^{stat} at $\mathcal{O}(1)$. This is part of our definition of improved currents in the static limit. So c_X^{HQET} starts at order $1/m_b$.

With the known structure of the currents at hand (4.17), it is straightforward to implement the corresponding correlation functions in the SF by introducing boundary fields for static quarks [26].

The static propagator is exactly/analytically known and just has to be evaluated on the given gauge background. Its standard implementation on the lattice goes back to Eichten and Hill [8, 9] but suffers from a worse signal-to-noise ratio. This main issue in simulating static quarks on the lattice was overcome by applying smearing techniques that are consistent with the heavy quark symmetries [87].

In order to perform our non-perturbative tests of HQET we are currently only interested in an expansion of those relativistic correlation functions from which we will build our test observables.

Heavy-light SF correlation functions to subleading order

For the sake of clarity and completeness we introduce the abbreviation

$$Z_{\text{hl}} = Z_{\zeta_{\text{h}}} Z_{\zeta_{\text{l}}} (1 + \frac{1}{2} b_X a m_{q,l}) (1 + \frac{1}{2} b_{\zeta_{\text{l}}} a m_{q,l}), \quad \text{with} \quad X = A, P, V, \quad (4.19)$$

at the appropriate places. There is no mass improvement term for the heavy quark in the effective theory. Following the definition of expectation values in HQET to subleading order, eq. (4.14), the SF correlation functions read

$$[f_X]_{\text{R}} = Z_X^{\text{HQET}} Z_{\text{hl}} e^{-m_{\text{bare}} x_0} \left\{ f_X^{\text{stat}} + c_X^{\text{HQET}} f_{\delta X}^{\text{stat}} + \omega_{\text{kin}} f_X^{\text{kin}} + \omega_{\text{spin}} f_X^{\text{spin}} \right\}, \quad (4.20a)$$

$$[f_1]_{\text{R}} = Z_{\text{hl}}^2 e^{-m_{\text{bare}} T} \left\{ f_1^{\text{stat}} + \omega_{\text{kin}} f_1^{\text{kin}} + \omega_{\text{spin}} f_1^{\text{spin}} \right\}, \quad (4.20b)$$

$$[k_1]_{\text{R}} = Z_{\text{hl}}^2 e^{-m_{\text{bare}} T} \left\{ k_1^{\text{stat}} + \omega_{\text{kin}} k_1^{\text{kin}} + \omega_{\text{spin}} k_1^{\text{spin}} \right\}. \quad (4.20c)$$

For f_1, k_1 there is no b_X -term in Z_{hl} . By construction these expansions are smoothly connected to the static limit. From the static limit with its enhanced symmetries there are certain relations among these correlation functions [12, 65] as discussed in the previous subsections. They also hold at finite lattice spacing and we have for example $f_A = -f_P = -k_V$ in the static approximation or $f_A^{\text{kin}} = -k_V^{\text{kin}}$ and $f_A^{\text{spin}} = 3k_V^{\text{spin}}$ as consequence of the spin symmetry of the static action. Using these identities the $O(a)$ improved heavy-light correlation functions to subleading order are given by

$$[f_A]_{\text{R}} = Z_A^{\text{HQET}} Z_{\text{hl}} e^{-m_{\text{bare}} x_0} \left\{ f_A^{\text{stat}} + c_A^{\text{HQET}} f_{\delta A}^{\text{stat}} + \omega_{\text{kin}} f_A^{\text{kin}} + \omega_{\text{spin}} f_A^{\text{spin}} \right\}, \quad (4.21a)$$

$$[f_P]_{\text{R}} = -Z_A^{\text{HQET}} Z_{\text{hl}} e^{-m_{\text{bare}} x_0} \left\{ f_A^{\text{stat}} + c_P^{\text{HQET}} f_{\delta A}^{\text{stat}} + \omega_{\text{kin}} f_A^{\text{kin}} + \omega_{\text{spin}} f_A^{\text{spin}} \right\}, \quad (4.21b)$$

$$[k_V]_{\text{R}} = -Z_A^{\text{HQET}} Z_{\text{hl}} e^{-m_{\text{bare}} x_0} \left\{ f_A^{\text{stat}} + c_V^{\text{HQET}} f_{\delta A}^{\text{stat}} + \omega_{\text{kin}} f_A^{\text{kin}} - \frac{1}{3} \omega_{\text{spin}} f_A^{\text{spin}} \right\}, \quad (4.21c)$$

$$[f_1]_{\text{R}} = Z_{\text{hl}}^2 e^{-m_{\text{bare}} T} \left\{ f_1^{\text{stat}} + \omega_{\text{kin}} f_1^{\text{kin}} + \omega_{\text{spin}} f_1^{\text{spin}} \right\}, \quad (4.21d)$$

$$[k_1]_{\text{R}} = Z_{\text{hl}}^2 e^{-m_{\text{bare}} T} \left\{ f_1^{\text{stat}} + \omega_{\text{kin}} f_1^{\text{kin}} - \frac{1}{3} \omega_{\text{spin}} f_1^{\text{spin}} \right\}. \quad (4.21e)$$

4.4 Strategy for a non-perturbative matching of QCD and HQET in finite volume

At any given order n in HQET we naively expect that properly chosen HQET observables approximate their QCD counterparts by means of

$$\Phi^{\text{QCD}}(M) = \Phi^{\text{HQET}}(M) + \mathcal{O}\left(\left(\Lambda_{\text{QCD}}/M\right)^{n+1}\right), \quad (4.22)$$

when expressed in terms of the RGI heavy quark mass M . We denote the number of bare couplings of HQET to order n by N_n^{HQET} . Its value at fixed order depends on the number of composite operators that need to be evaluated. We already became familiar with some of them ($m_{\text{bare}}, \omega_{\text{kin}}, \omega_{\text{spin}}, c_X^{\text{HQET}}, Z_X^{\text{HQET}}, \dots$). The equivalence between HQET and QCD only becomes true if we fix the bare couplings of HQET.

First of all one has to fix the parameters of (lattice) QCD as usual by requiring a set of observables to agree with experiment. If this is done, lattice QCD gets its predictive power and one can compute e.g. matrix elements or hadron masses and compare them with experiments. Next the determination of the bare couplings of HQET at order n can be done by imposing

$$\Phi_k^{\text{QCD}}(M) \equiv \Phi_k^{\text{HQET}}(M), \quad k = 1, 2, \dots, N_n^{\text{HQET}} \quad (4.23)$$

Then, predictions made by HQET would be correct up to $\mathcal{O}(aM^l/M^{n+1}) = \mathcal{O}(1/m^{n+1})$ terms with $l = 0, \dots, n+1$, i.e. at $n = 1$ one gets $1/M^0$ terms with $\mathcal{O}(a^2)$ errors as well as $1/M^1$ terms with $\mathcal{O}(a)$ errors.

One might think that each $\Phi_k^{\text{HQET}}(M)$ could be determined by a physical observable, available through experiments. But N_n^{HQET} increases rapidly even at fixed order if new composite operators are introduced, accompanied with a (complete) loss of its predictive power. However, from the constructive point of view it is more natural to compare HQET to QCD with a fixed number of parameters, $n_f + 1$. Hence, one may instead use the quantities $\Phi_k^{\text{QCD}}(M)$ computed in the continuum limit of lattice QCD. But this is the main practical problem since it involves computations with heavy quarks treated fully relativistically on the lattice. One knows that finite volume effects in lattice QCD simulations become negligible above a physical extent of approximately 2 fm. For a lattice of medium size like $L/a = 40$ this corresponds to a physical lattice spacing of $a = 0.05$ fm. With m_{B} from table 3 the corresponding mass in lattice units would be $am_{\text{B}} \approx 1.3$, i.e. a subtracted quark mass am_{q} of $\mathcal{O}(1)$. QCD lattice artefacts are thus expected to be very large and a real non-perturbative treatment of HQET seemed to be impossible for some time. This apparent contradiction was solved by Heitger & Sommer [11].

Finite volume matching

The solution to this matching problem is to consider QCD in a (finite) small volume where the lattice spacing can be made very fine, i.e. $am_{\text{B}} \ll 1$, and simultaneously $1/(m_{\text{B}}L) \ll 1$ to get a well-behaved $1/m$ -expansion with small residual $\mathcal{O}(1/m^{n+1})$ corrections.

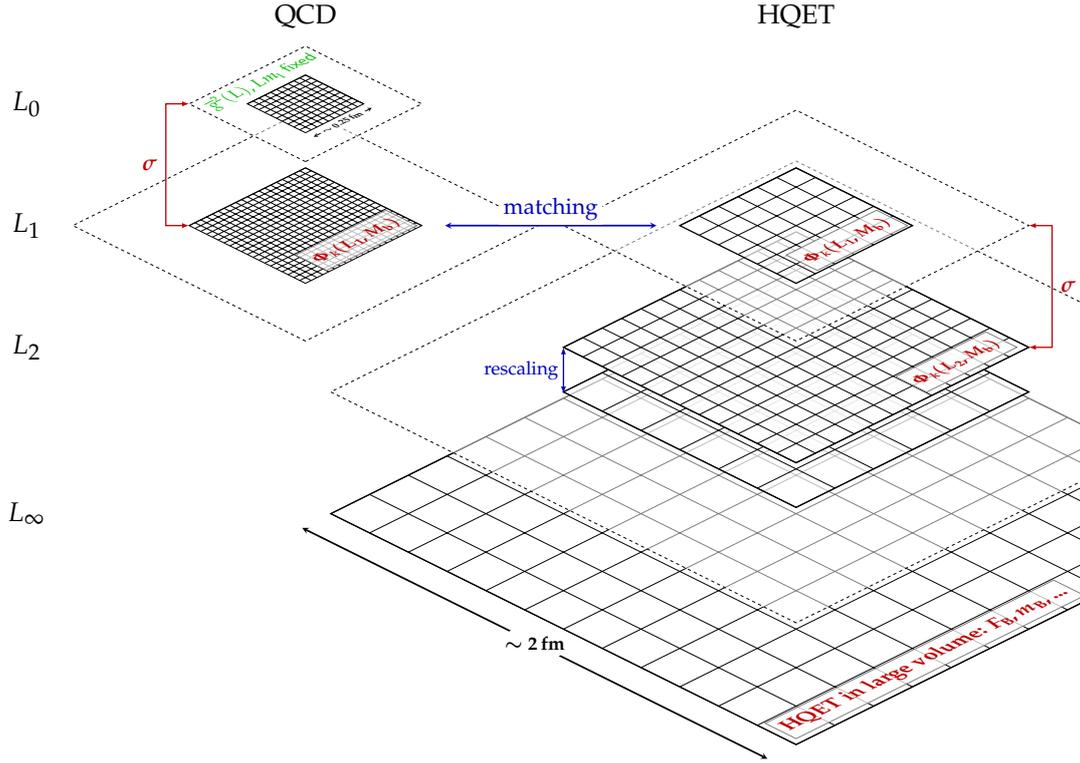


Figure 5: Strategy for a non-perturbative matching of QCD and HQET in a small volume L_1^4 , $L_1 \sim 0.5$ fm. Computations on the left/right are performed in lattice QCD/HQET. The determination of HQET parameters has to be performed by QCD simulations at the same physical parameters. σ denotes a well-defined step-scaling of the physical parameters and each lattice shown here corresponds to a set of lattices such that the continuum limit is taken in each step.

After the computations of the small volume QCD observables – now denoted by $\Phi_k^{\text{QCD}}(L, M)$ – are performed, one computes the corresponding observables $\Phi_k^{\text{HQET}}(L, M)$ in the same physical volume but at much coarser lattice spacing. We are free to do so because the matching itself is performed after taking the continuum limit for both sets of observables at fixed renormalized parameters of QCD. We call the physical volume where this happens the *matching volume* and label it for future purposes by L_1 . To make the whole strategy more visible it is shown in fig. 5.

The main and much tested assumption behind this strategy is that both *QCD* and *HQET* can be well-defined in a finite volume and that the parameters in the Lagrangian do not depend on the volume.

However, to make predictions with HQET that can be compared with experiment like the B-meson decay constant F_{B_q} or the bag parameter B_{B_q} one still needs *large volume computations* given by a physical volume $L_\infty \approx 2$ fm. The connection to the matching volume can be established by a finite-size scaling of the physical volume observables. This is practically achieved by one finite-size scaling step σ from L_1 to $L_2 = 2L_1$ followed by a rescaling of the lattice spacing $a \rightarrow 2a$ at fixed volume L_2 . This rescaling causes no problems but is needed in order to reach the final volume L_∞ at a lattice spacing that can still be reached in large volume simulations.

4.5 Example: The matching strategy in the static case

It was made clear in foregoing discussions that in the static approximation the spin splitting term and the kinetic term disappear and the ground state of the B-meson is degenerated. In this case the quantum ground state is fully characterized by the B-meson mass, or effectively by the mass of the b-quark. As we prefer to express physics in terms of RGIs, this would be the non-perturbatively renormalized RGI mass of the b-quark in the static approximation, M_b . The non-perturbative matching of HQET and QCD can be applied to determine this mass. To do this we have to identify the quantities Φ_k with suitable observables that can be calculated by numerical simulations at fixed renormalized parameters. We already know how this can be done when using the Schrödinger functional as intermediate renormalization scheme. Because HQET does not alter the light sector of QCD it is naturally to choose

$$\Phi_1^{\text{HQET}}(L) = \bar{g}^2(L) \equiv \text{const}, \quad \text{and} \quad \Phi_2^{\text{HQET}}(L) = Lm_1(L) \equiv 0. \quad (4.24)$$

This trivially fulfills the *finite-volume matching conditions*

$$\Phi_k^{\text{QCD}}(L, M) \equiv \Phi_k^{\text{HQET}}(L, M), \quad (4.25)$$

for $k = 1, 2$. Above, we have not indicated a heavy mass dependence because these observables are clearly independent of it. They belong to the light degrees of freedom. By fixing (4.24) we will automatically get parameter triples $(L/a, \beta = 6/g_0^2, \kappa_1)$ that allow us to take the limit $a/L \rightarrow 0$. The computation of the b-quark mass in leading order of HQET clearly requires to define one more condition in order to fix the remaining parameter $a\delta m$ of the static Lagrangian. From our discussion in section 3.8 we already know candidates that do the job. These are the effective energies Γ_X , $X = \text{PS}, V$ which carry the entire quark mass dependence if defined in small volume QCD with a relativistic b-quark. We thus set

$$L\Gamma_X(L, M) \equiv \Phi_3^{\text{QCD}}(L, M) \equiv \Phi_3^{\text{HQET}}(L, M) \equiv L(\Gamma_{\text{stat}}(L, M) + m), \quad (4.26)$$

where Γ_{stat} denotes the analogue of Γ_X in the static effective theory that implicitly contains the linearly divergent mass counterterm $a\delta m$ which causes all the problems. *Due to the degeneration Γ_{stat} is the same for each of the $X = \text{PS}, V$.* Before we proceed, I want to explain on a formal level the strategy [11] that allows for the cancellation of the divergent term $a\delta m$ in the static limit. From the explicit form of the static quark propagator it can be deduced that the energy Γ_{stat} in the static approximation obeys the decomposition

$$\begin{aligned} \Gamma_{\text{stat}}(L, M) &= \Gamma_{\text{stat}}(L, M)|_{\delta m=0} + a^{-1} \ln(1 + a\delta m) \\ &= \Gamma_{\text{stat}}(L, M)|_{\delta m=0} + m_{\text{bare}} - m, \quad m_{\text{bare}} = m + a^{-1} \ln(1 + a\delta m). \end{aligned} \quad (4.27)$$

Now $\Gamma_{\text{stat}}(L, M)|_{\delta m=0}$ is finite and immediately offers us a way to cancel the divergent part which is independent of L . Hence a difference like $\Gamma_{\text{stat}}(2L, M) - \Gamma_{\text{stat}}(L, M)$ is made finite without further knowledge of the divergence. With $E_{\text{stat}} = \lim_{L \rightarrow \infty} \Gamma_{\text{stat}}(L, M_b)$ and $m_B =$

$\lim_{L \rightarrow \infty} \Gamma_\chi(L, M_b)$ we have $m_B = E_{\text{stat}} + m_{\text{bare}}$ as infinite volume relation. By multiplying with L_1 and using (4.26) for finite L_1 we get

$$L_1 m_B = L_1 E_{\text{stat}} + L_1 m_{\text{bare}} \quad (4.28)$$

$$= L_1 E_{\text{stat}} - L_1 \Gamma_{\text{stat}}(L_1) + \Phi_3^{\text{QCD}}(L_1, M_b) \quad (4.29)$$

$$= L_1 [E_{\text{stat}} - \Gamma_{\text{stat}}(2L_1)] + L_1 [\Gamma_{\text{stat}}(2L_1) - \Gamma_{\text{stat}}(L_1)] + \Phi_3^{\text{QCD}}(L_1, M_b) \quad (4.30)$$

Note that by inserting a 'zero' the terms can be rearranged such that the $1/a$ power divergences carried by each term in the square brackets cancel each other, provided that the lattice spacing is equal for both. Thus their continuum limit exists and can be taken. The matching condition now becomes

$$L_1 m_B - L_1 [E_{\text{stat}} - \Gamma_{\text{stat}}(2L_1)] + L_1 [\Gamma_{\text{stat}}(2L_1) - \Gamma_{\text{stat}}(L_1)] = \Phi_3^{\text{QCD}}(L_1, M_b). \quad (4.31)$$

The first term on the LHS is input from experiment, i.e. m_B taken from table 3. In view of figure 5 the second one connects the HQET large-volume simulation to the finite volume $L_2 = 2L_1$. The difference has to be evaluated before taking the continuum limit. This applies to the third term as well which finally through one finite-size scaling step establish the connection to the matching volume L_1 , where Φ_3^{QCD} can be evaluated. The example above again shows the beautiful entanglement between finite-volume theory and renormalization and the need to first produce continuum results before taking the infinite volume limit. Furthermore, it is evident that the quantitative knowledge of $\Phi_3^{\text{QCD}} = L_1 \Gamma_\chi$ as a function of M in the relevant quark mass region allows to non-perturbatively solve this equation in favour of M_b in the static approximation.

The computation of Φ_3^{QCD} over a wide range of masses M is our final goal, examined in section 7. Before we are able to extract the M -dependence of Φ_3^{QCD} , we have to solve another tuning problem. We need to know the hopping parameters for the heavy quark κ_h on each of our lattices in table 5 for various choices of M such that the continuum limit is reached at fixed M .

Remarks

We described the matching procedure for the simple case of a minimal set of matching functions Φ_k in the static approximation. Even in that case there are various ways to reach the goal, starting with a concrete choice for Γ_χ . Which one is best could be guided through symmetry considerations and in fact in [11] evidence was found that the spin-averaged mass Γ_{av} has smaller $\mathcal{O}(1/m)$ -corrections. But a final answer can only be obtained by numerical computations. However, to obtain a better physical picture using HQET one should go the subleading order. In quenched QCD this was addressed in [12] for the simplest but still intricate case of spin-averaged quantities. In order to perform the non-perturbative power-subtractions in a more general setting, *step-scaling functions* F_k are involved that implicitly describe the change of a complete set of observables $\{\Phi_k^{\text{HQET}}\}$ under a finite rescaling $L \rightarrow sL$ by the matrix relation

$$\Phi_k^{\text{HQET}}(sL, M) = F_k \left(\{ \Phi_j^{\text{HQET}}(L, M), j = 1, \dots, N_n^{\text{HQET}} \} \right), \quad k = 1, \dots, N_n^{\text{HQET}}. \quad (4.32)$$

Practicability dictates to set $s = 2$ and to find a set $\{\Phi_k^{\text{HQET}}\}$ with a minimal mixing behaviour under (4.32). But the most important point is still that (4.32) needs to be evaluated at finite lattice spacing and fixed renormalized mass M . A quite simple implementation of such F_k was given by the brackets in (4.31). In [13] for instance the method described here was applied to non-perturbatively determine the parameters in the effective theory and to non-perturbatively compute the decay constant of a heavy-light meson in HQET including $(1/m)$ -corrections.

4.6 Perturbative conversion factors between QCD and HQET

We perform non-perturbative tests of HQET as an additional tool to a full non-perturbative matching procedure which was explained in the previous sections. It allows us to get first impressions of the asymptotic mass behaviour, to locate possible issues which may arise in the matching procedure and to get an estimate of the errors associated to some of the matching functions Φ_k . Thus we will obtain further insights about the convergence to the continuum limit of relativistic heavy-light QCD observables. If there are large cutoff effects for instance which could alter the continuum limit behaviour especially at large quark masses, we would have to make additional effort to really control the cutoff effects in an already $\mathcal{O}(a)$ improved theory.

In contrast to the non-perturbative matching procedure, we will use conversion functions to connect our QCD results to the predictions made by HQET. Here perturbation theory enters in a well-controlled way. In view of dynamical fermion simulations we thus have to extend the work that has been done in quenched QCD [14] to the present case.

The interaction of the light degrees of freedom with heavy quarks in QCD can produce high exchanging momenta which are truncated in the HQET expansion of the partition function. Thus short-distance properties of the heavy quark cannot be resolved in HQET. But in order to get predictive descriptions of heavy-light mesons, short-distance corrections are needed. They entail a logarithmic dependence on the heavy quark mass m through the strong coupling $\bar{g}^2(m)$. For large values of m as anticipated here, the coupling is small and can be safely computed in perturbation theory, allowing for a perturbative matching of QCD and HQET observables.

A matrix element $\langle \mathcal{O}_R(\mu, m) \rangle$ renormalized in QCD at scale μ can be related to its equivalent in the effective theory up to order $\mathcal{O}(1/m)$ through the Wilson coefficient C_{match} by means of the so-called *short-distance expansion* [88, 89]

$$\langle \mathcal{O}_R(\mu, m) \rangle_{\text{QCD}} = C_{\text{match}}^{\mathcal{O}}(\mu, m) \langle \mathcal{O}_R(\mu) \rangle_{\text{stat}} + \mathcal{O}(1/m). \quad (4.33)$$

In this expansion the coefficients $C_{\text{match}}^{\mathcal{O}}(\mu, m)$ account for the short-distance corrections of QCD which is not describable in HQET [90, 91]. They are perturbatively computable and add QCD short-distance corrections to the relevant non-perturbatively renormalized HQET operator. It is convenient to compute these coefficients in the $\overline{\text{MS}}$ scheme with an expansion in the $\overline{\text{MS}}$ coupling constant $\bar{g}_{\overline{\text{MS}}}$ according to

$$C_{\text{match}}^{\mathcal{O}}(\mu, m) = 1 + c_1^{\mathcal{O}}(\mu, m) \bar{g}_{\overline{\text{MS}}}^2(\mu) + \mathcal{O}\left(\bar{g}_{\overline{\text{MS}}}^4(\mu)\right). \quad (4.34)$$

Applying equation (1.18) to (1.20) to the present case of an operator in the static effective theory

renormalized in the $\overline{\text{MS}}$ scheme gives the corresponding RGI matrix element:

$$\langle \mathcal{O}_{\text{RGI}}^{\overline{\text{MS}}} \rangle_{\text{stat}} = \langle \mathcal{O}_{\text{R}}^{\overline{\text{MS}}}(\mu) \rangle_{\text{stat}} [2b_0 \bar{g}_{\overline{\text{MS}}}^2(\mu)]^{\frac{-\gamma_0^{\mathcal{O}}}{2b_0}} \exp \left\{ - \int_0^{\bar{g}_{\overline{\text{MS}}}(\mu)} dg \left[\frac{\gamma_{\overline{\text{MS}}}^{\mathcal{O}}(g)}{\beta_{\overline{\text{MS}}}(g)} - \frac{\gamma_0^{\mathcal{O}}}{b_0 g} \right] \right\}, \quad (4.35)$$

Clearly we can write equation (4.33) on the level of renormalization group invariants, because all quantities are well defined in the low coupling limit. With the corresponding RGI matrix element on the QCD side, $\langle \mathcal{O}_{\text{RGI}}(m) \rangle_{\text{QCD}}$, it reads

$$\langle \mathcal{O}_{\text{RGI}}(m) \rangle_{\text{QCD}} = C^{\mathcal{O}}(m) \langle \mathcal{O}_{\text{RGI}} \rangle_{\text{stat}} + \mathcal{O}(1/m). \quad (4.36)$$

The matching coefficient that appears here is then related to that in eq. (4.33) by

$$C^{\mathcal{O}}(m) = C_{\text{match}}^{\mathcal{O}}(\mu, m) \frac{\langle \mathcal{O}_{\text{RGI}}(m) \rangle_{\text{QCD}}}{\langle \mathcal{O}_{\text{R}}(\mu, m) \rangle_{\text{QCD}}} \Big|_{\mu=m} \frac{\langle \mathcal{O}_{\text{R}}(\mu, m) \rangle_{\text{stat}}}{\langle \mathcal{O}_{\text{RGI}}(m) \rangle_{\text{stat}}} \Big|_{\mu=m} + \mathcal{O}(1/m). \quad (4.37)$$

To not introduce an artificial log-dependence the renormalization scale μ is set to the heavy quark mass m . Note that for operators that do not renormalize, like the axial vector current in QCD, the corresponding ratio is one and does not contribute. Equation (4.37) implicitly defines the anomalous dimension $\gamma_{\text{match}}^{\mathcal{O}}$ of the operator \mathcal{O} in the matching scheme through

$$C^{\mathcal{O}}(m) = [2b_0 \bar{g}_{\overline{\text{MS}}}^2(m)]^{\frac{\gamma_0^{\mathcal{O}}}{2b_0}} \exp \left\{ \int_0^{\bar{g}_{\overline{\text{MS}}}(m)} dg \left[\frac{\gamma_{\text{match}}^{\mathcal{O}}(g)}{\beta_{\overline{\text{MS}}}(g)} - \frac{\gamma_0^{\mathcal{O}}}{b_0 g} \right] \right\}. \quad (4.38a)$$

The leading coefficient in the usual perturbative expansion $\gamma_{\text{match}}^{\mathcal{O}}(\bar{g}_{\overline{\text{MS}}})$ is still the universal one, $\gamma_0^{\mathcal{O}}$. But the higher order, scheme-dependent coefficients $\{\gamma_{i>0}^{\mathcal{O}}\}$ are altered by the matching coefficients:

$$\gamma_{\text{match}}^{\mathcal{O}}(\bar{g}_{\overline{\text{MS}}}) = -\bar{g}_{\overline{\text{MS}}}^2 \left\{ \gamma_0^{\mathcal{O}} + \gamma_1^{\mathcal{O}} \bar{g}_{\overline{\text{MS}}}^2 + \gamma_2^{\mathcal{O}} \bar{g}_{\overline{\text{MS}}}^4 + \mathcal{O}(\bar{g}_{\overline{\text{MS}}}^6) \right\} \quad (4.38b)$$

$$\gamma_1^{\mathcal{O}} = \gamma_1^{\mathcal{O}, \overline{\text{MS}}} + 2b_0 c_1^{\mathcal{O}} \quad (4.38c)$$

$$\gamma_2^{\mathcal{O}} = \gamma_2^{\mathcal{O}, \overline{\text{MS}}} + 4b_0 (c_2^{\mathcal{O}} + k_1 \gamma_0^{\mathcal{O}}) + 2b_1 c_1^{\mathcal{O}} - 2b_0 [c_1^{\mathcal{O}}]^2 \quad (4.38d)$$

...

The coefficients $\{c_i^{\mathcal{O}}\}$ are those of the perturbative expansion of the Wilson coefficients (4.34). Note a subtlety in our conventions: while $\gamma_1^{\mathcal{O}}$ is the two-loop coefficient, $c_1^{\mathcal{O}}$ is the one-loop coefficient. Thus at $(n+1)$ -loop order of the anomalous dimension in the matching scheme the highest loop order of the matching coefficient is n . The additional constant k_1 appears due to a reparametrization of the mass, because in the literature on HQET the matching is usually done at the scale of the heavy quark's pole mass, m_Q . But its definition is ambiguous due to the presence of infra-red renormalon singularities [92, 93]. This issue disappears after the pole mass was eliminated in favour of another (more physical) mass definition like $\bar{m}_{\overline{\text{MS}}}(\mu)$. Both are related

by

$$\bar{m}_{\overline{\text{MS}}}(\bar{m}_*) = \bar{m}_* , \quad \frac{m_Q}{\bar{m}_*} = 1 + k_1 \bar{g}_{\overline{\text{MS}}}^2(\bar{m}_*) + \mathcal{O}(\bar{g}_{\overline{\text{MS}}}^4) . \quad (4.39)$$

where the first equation conventionally defines the $\overline{\text{MS}}$ mass without any logarithmic dependencies. In view of its application in section 7 and because it is NP computable in lattice QCD, we go one step further and finally also express the mass dependence of (4.38a) in terms of the RGI mass M . To this end we need to recall the well-known formula (1.17) in the $\overline{\text{MS}}$ scheme:

$$\frac{M}{\bar{m}_*} = [2b_0 \bar{g}_{\overline{\text{MS}}}^2(\bar{m}_*)]^{-\frac{d_0}{2b_0}} \exp \left\{ - \int_0^{\bar{g}_{\overline{\text{MS}}}(\bar{m}_*)} dg \left[\frac{\tau_{\overline{\text{MS}}}(g)}{\beta_{\overline{\text{MS}}}(g)} - \frac{d_0}{b_0 g} \right] \right\} . \quad (4.40)$$

Then the function that relates the heavy quark's pole mass to the renormalization group invariant one reads

$$C_{\text{mass}}(M/\Lambda_{\overline{\text{MS}}}) \equiv \frac{m_Q}{\bar{m}_*} \frac{\bar{m}_*}{M} . \quad (4.41)$$

The second factor on the right hand side is known very accurately because the mass anomalous dimension $\tau_{\overline{\text{MS}}}$ is known up to four-loop [94, 20]. Unfortunately, the perturbative expansion of the first factor which is known to three-loop [95] is not well behaved since even the highest term is still significant. However, the pole mass only appears in (4.41) and not in the conversion functions (4.38a). They are parametrized by the RGI mass M which is unambiguously defined.

5 Setup of lattice simulations

To study the quark mass dependence of heavy-light observables using lattice simulations it is crucial to carefully set up a *line of constant physics*. To this end the renormalized parameters in the light quark sector have to be fixed. In order to achieve this in the continuum limit, one has to compute these observables for a series of bare parameters $(L/a, \beta, \kappa_1)$ with the hopping parameter of the dynamical light quarks, κ_1 . We define the *constant physics condition* by fixing the renormalized SF coupling and the PCAC current quark mass m_1 of the light flavours to¹⁶

$$\bar{g}^2(L)|_{L=L_0} \equiv 2.989, \quad z_1(L)|_{L=L_0} \equiv L_0 m_1(L_0) = 0, \quad L_0 = L_1/2. \quad (5.1)$$

This *defines* the spatial extend L_1 of our matching volume. From the running of the SF coupling as it is known for $N_f = 2$ [28], we can infer that $L_0 \approx 0.25$ fm holds. Its exact value does not matter at the moment. But what matters is that this volume is exactly related to the matching volume by one finite size scaling step, i.e. $L_1 = 2L_0 \approx 0.5$ fm. In this volume the lattice spacings a can be chosen sufficiently smaller than $1/m_b$ to allow for a relativistic treatment of the heavy quark with controllable discretization errors of order a^2 in the $O(a)$ improved theory.

Supported by former simulations it seems possible, and necessary to reach the continuum limit, to get a maximal resolution of $L_1/a = 40$ within our computational resources which translates to $L_0/a \leq 20$. After tuning the simulation parameters (β, κ_1) by numerical computation on the largest resolution $L_0/a = 20$ with kinematic parameters specified in (3.27) to fulfill (5.1), we made tentative interpolations in (β, κ_1) relying on known data in bare parameter space [28]. The outcome is summarized in table 4. These bare parameters (β, κ_1) are the same as for the matching volume which is just given by doubling the number of lattice points in each direction. The coupling in the matching volume turns out to be

$$\bar{g}^2(L_1) = \sigma(2.989) = 4.484(48), \quad (5.2)$$

where the (continuum) step-scaling function σ of the coupling [28] was used. We aim for a precision of $\Delta\bar{g}^2 \leq 0.04$ in the coupling that corresponds to an uncertainty in the b-quark mass of at most 0.5%. Furthermore, a mismatch $|L_0 m_1(L_0)| < 0.05$ on the condition in (5.1) by a not properly tuned hopping parameter $\kappa_1 \approx \kappa_c$ is still tolerable.

In the matching volume L_1 where the QCD parameters are defined by (5.2) and vanishing light quark mass, we must fix the HQET simulation parameters for resolutions $6 \leq L_1/a \leq 16$ to be able to reach $L_2 = 2L_1$ for all of them. Thus another set of tuning runs had to be carried out to get $(L_1/a, \beta, \kappa)$ for HQET. Small mismatches in the outcome of such tuning simulations with respect to the target values (\bar{g}^2, z_1) can always be corrected by the perturbative β -function and the mass derivative of the coupling. This was done for instance in [28, 157]. Together they provide a function that relates the β parameter of our $N_f = 2$ simulations in the range $\beta \in [5.3, 5.8]$ to the

¹⁶ As might be clear it is not possible in real simulations to set such values to arbitrary precision. This is not even required as long as the impact of the associated error is far below that of the final result.

L_0/a	β	κ	\bar{g}^2	$L_0 m_1$
20	6.6380	0.1351631	2.989(43)	+0.0190(5)
16	6.5113	0.1354410	2.878(39)	-0.0033(7)
12	6.3158	0.1357930	2.893(33)	+0.0054(6)
10	6.1906	0.1360160	2.923(38)	+0.0089(10)

Table 4: Parameters after initial tuning in L_0 . NP values for the coupling even within the bounds forced us to refine the tuning leading to β -values of table 5. The fermionic phase angle was set to $\theta = 0.5$.

L_1/a	β	κ	$L_1 m_1$
40	6.6380	0.135192285	+0.0019(8)
32	6.4574	0.135521000	+0.0047(10)
24	6.2483	0.135910400	-0.0035(16)
20	6.1569	0.136053600	+0.0110(26)

Table 5: Parameters for relativistic heavy quark simulations in L_1 after tuning. Note that only the β value of the largest lattice corresponds to its counterpart in table 4 and the hopping parameter can be tuned independently. The line of constant physics is very well approached within bounds. Again, θ was set to 0.5.

lattice resolution by

$$\ln(L^*/a) = 2.3338 + 1.4025(\beta - 5.5) \quad \Leftrightarrow \quad L^*/a \in [7.8, 16.1] , \quad (5.3)$$

up to a relative error on L^*/a of two percent. The *low-energy reference length scale* L^* was implicitly defined through [28]¹⁷

$$\bar{g}^2(L^*) \equiv 5.5 \quad \Rightarrow \quad L^* \Lambda_{\overline{\text{MS}}} = 0.801(56) . \quad (5.4)$$

This is necessary to calibrate the overall energy scale by relating L^* to a non-perturbative large-volume observable. In [158] the physical value of $\Lambda_{\overline{\text{MS}}}$ could be refined with respect to the original value by relating the length scale by $L^* m_K$ to the physical Kaon mass. For this input from [96] was needed, giving a value of $\Lambda_{\overline{\text{MS}}} = 257(56)$ MeV.

However, for further purposes it is important to know the connection between the scales L_1 and L^* . Both are defined in terms of the coupling and the connection is established by computing $\ln(\Lambda_{\text{SF}} L_1)$ and $\ln(\Lambda_{\text{SF}} L^*)$ as described in [158]. The result is

$$r_1 = L_1/L^* \Big|_{\text{continuum}} = 0.810(12) , \quad \text{and hence} \quad L_1 \Lambda_{\overline{\text{MS}}} = 0.649(57) . \quad (5.5)$$

With this ratio (5.3) translates into $L_1/a \in [6.3, 13.0]$ and could be used to give some trial estimates for the tuning runs. The resolutions that will be simulated in HQET finally are $L_1/a \in \{6, 8, 10, 12, 16\}$ and twice as large at L_2/a . To finally perform non-perturbative subtractions as in [12] to circumvent the problem of power divergences in a perturbative treatment, lattices for HQET computations with $T = L$ and $T = L/2$ for both, L_1 and L_2 , have to be prepared. They all share the same set of parameters (β, κ_1) .

¹⁷ For practical reason all subsequent scale dependent renormalizations are carried out at a low-energy renormalization scale $1/\mu_{\text{ren}} = L_{\text{ren}} > L^*$ which itself is defined by $\bar{g}^2(L_{\text{ren}}) = 4.61$, c.f. fig. 1.

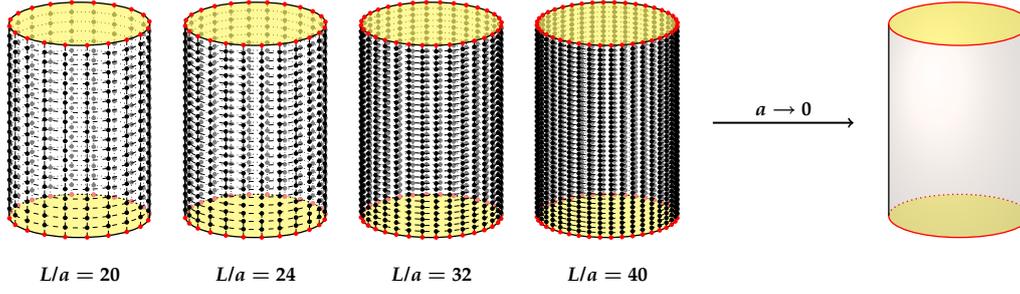


Figure 6: Schematic view how the continuum limit is reached at fixed spatial extend $L = L_1$ of approximately 0.5 fm. The physical volume is kept fixed while varying the lattice spacing by increasing the number of simulation points.

5.1 Fixing the finite-volume heavy quark mass

For the purpose of studying the heavy quark mass dependence of finite-volume QCD observables in the continuum limit of the matching volume L_1 , we need to compute the hopping parameters for each lattice resolution that corresponds to the same value of the dimensionless RGI heavy quark mass $z = LM$. With those at hand the continuum limit as sketched in figure 6 can be taken at fixed z . But let me remind you that we fixed our line of constant physics in the volume $L_0 = L_1/2$.

We denote the masses associated to the heavy quark flavour by a subscript h. The renormalized PCAC current quark mass (3.72) and the multiplicative renormalization of the subtracted quark mass parameter (2.43) for the heavy quark flavour are equal up to $O(a^2)$ in the improved theory and read at renormalization scale $\mu_0 = 1/L_0$:

$$\bar{m}_h(\mu_0) = \frac{Z_A(g_0)(1 + b_A am_{q,h})}{Z_P(g_0, \mu_0)(1 + b_P am_{q,h})} m_h, \quad (5.6)$$

$$\bar{m}_h(\mu_0) = Z_m(g_0, \mu_0)(1 + b_m am_{q,h}) m_{q,h}. \quad (5.7)$$

The connection to the RGI heavy quark mass M_h is provided through the step-scaling procedure of [29], i.e. their connection is decomposed via

$$M_h = \frac{M_h}{\bar{m}_h(\mu_{pt})} \frac{\bar{m}_h(\mu_{pt})}{\bar{m}_h(\mu_0)} \times \bar{m}_h(\mu_0), \quad \mu_0 = 1/L_0, \quad (5.8)$$

into finite ratios with the

- non-perturbative part $\bar{m}_h(\mu_{pt})/\bar{m}_h(\mu_0)$ that consists of a joint recursive finite-size scaling of the coupling and mass from the low-energy scale μ_0 to a high-energy scale μ_{pt} in the

continuum. In the notation of [29] this is characterized by¹⁸

$$u_k = \bar{g}^2(L_k), \quad L_k = L_0/2^{-k}, \quad u_0 = \bar{g}^2(L_0) = 2.989, \quad (5.9a)$$

$$\frac{\bar{m}_h(\mu_{\text{pt}})}{\bar{m}_h(\mu_0)} = \left[\prod_{k=-1}^n \sigma_{\text{P}}(u_k) \right]^{-1}, \quad \mu_k = 2^{-k} \mu_0, \quad \mu_{\text{pt}} \equiv \mu_n = 2^{-n} \mu_0, \quad (5.9b)$$

in terms of the step-scaling function of the mass σ_{P} and coupling σ [28], and

- perturbative part that connects the scale μ_{pt} by the asymptotic RG equation

$$\frac{M_h}{\bar{m}_h(\mu_{\text{pt}})} = [2b_0 \bar{g}^2(\mu_{\text{pt}})]^{-d_0/2b_0} \exp \left\{ - \int_0^{\bar{g}^2(\mu_{\text{pt}})} dg \left[\frac{\tau(g)}{\beta(g)} - \frac{d_0}{b_0 g} \right] \right\} \quad (5.10)$$

with negligible errors to the RGI heavy quark mass M_h .

We will abbreviate the total factor as it was done in the literature [14, 97] for the quenched case by the universal, regularization independent ratio

$$h(L_0) \equiv \frac{M_h}{\bar{m}_h(\mu_0)} = \frac{M_h}{\bar{m}_h(\mu_{\text{pt}})} \frac{\bar{m}_h(\mu_{\text{pt}})}{\bar{m}_h(\mu_0)}, \quad (5.11)$$

which still indicates the dependence on the renormalization scale. To compute different hopping parameters of the heavy quark κ_h that represent the same values for M_h on each lattice, we need to establish a connection between them. This is given by eq. (5.7) together with (2.25) and subsequently amounts to an estimation of $Z_m(g_0, \mu_0)$ and $b_m(g_0)$. However, it turned out to be advantageous to not directly compute Z_m but to make a detour for its estimation. The indirect method we follow here relies on the equivalence of (5.6) and (5.7). This allows to get a really accurate estimate for the joint renormalization factor

$$Z(g_0) = \frac{Z_{\text{P}}(g_0, \mu_0) Z_{\text{M}}(g_0, \mu_0)}{Z_{\text{A}}(g_0)}. \quad (5.12)$$

The direct estimation of Z is part of section 6 and we postpone further details to it. Because of $Z_m \propto Z_{\text{P}}^{-1}$, the scale dependence cancels out and Z only depends on g_0 . With $Z(g_0)$ at hand and estimates for Z_{A} and Z_{P} which are standard, we can finally define the following dimensionless RGI heavy quark mass

$$z \equiv LM_h \equiv LZ_{\text{M}}(g_0, \mu_0) \tilde{m}_{\text{q,h}} \quad (5.13a)$$

$$= Lh(L_0) \frac{Z(g_0) Z_{\text{A}}(g_0)}{Z_{\text{P}}(g_0, \mu_0)} (1 + b_m a m_{\text{q,h}}) m_{\text{q,h}} \quad (5.13b)$$

$$= \left[\frac{L}{a} \right] h(L_0) \frac{Z(g_0) Z_{\text{A}}(g_0)}{Z_{\text{P}}(g_0, \mu_0)} (1 + b_m a m_{\text{q,h}}) a m_{\text{q,h}}, \quad (5.13c)$$

¹⁸In contrast to the cited reference we have some unusual minus signs here. This is just to not clash with our notation used in the matching strategy, where L_1 is equal to $2L_0$ instead of $L_0/2$. Doing it this way, we have the sequence $\mu_1 < \mu_0 < \mu_{-1} < \dots < \mu_{\text{pt}} < \dots < \infty$, starting from the QCD matching volume .

k	u_k	$M/\bar{m}(\mu_0)$
0	2.989	1.521
-1	2.317(13)	1.519(06)
-2	1.903(15)	1.518(09)
-3	1.616(13)	1.518(11)
-4	1.405(12)	1.520(13)
-5	1.244(11)	1.521(14)
-6	1.116(10)	1.523(15)
-7	1.013(09)	1.525(16)
-8	0.927(09)	1.526(17)

Table 6: Recursive finite-size scaling for the mass in the continuum according to (5.9). The 3-loop β -function and 2-loop mass anomalous dimension were used for the perturbative part when starting at a renormalized coupling $\bar{g}^2(\mu_{\text{pt}}) = u_k$.

with the overall renormalization factor

$$Z_M(g_0, \mu_0) \equiv h(L_0) \times Z_m(g_0, \mu_0) . \quad (5.14)$$

We only wrote the g_0 -dependence for the renormalization constants explicitly. As already known, b_m and am_q depend on g_0 . They are all well-defined due to our line of constant physics. Hence, also z becomes a smooth function on g_0 . Using (2.25) for the heavy quark flavour, the quadratic equation $z = z(1/\kappa_h)$ can be inverted in favour of

$$\kappa_h(z) = \left[\frac{1}{\kappa_c} - \frac{1}{b_m} \left(1 - \sqrt{1 + z \cdot \frac{4b_m}{[L/a]Z_M(g_0, \mu_0)}} \right) \right]^{-1} . \quad (5.15)$$

Choosing the solution of the quadratic equation with the negative square root fulfills the requirement $\kappa(z=0) \equiv \kappa_c$. We should emphasize that there is an upper bound on z because still the tree-level value for b_m is negative and remains so for $g_0 > 0$. This bound z_{max} is obviously given by

$$0 \leq z \leq z_{\text{max}} \equiv - \left[\frac{L}{a} \right] \frac{Z_M(g_0, \mu_0)}{4b_m} \quad \Leftrightarrow \quad 0 \leq a\tilde{m}_q \leq - \frac{1}{4b_m} . \quad (5.16)$$

We can already determine the value of $h(L_0)$ since the coupling is fixed and the scale evolution was given in [29]. The result for the finite-size scaling is summarized in table 6. We choose the value given at $n = 5$, namely

$$h(L_0) = 1.521(14) , \quad (5.17)$$

because the corresponding $\bar{g}^2(\mu_{\text{pt}}) = 1.244(11)$ is already well in the weak coupling regime such that perturbation theory may be applied to evaluate (5.10). We complete the estimation of $\kappa_h(z)$ after Z which enters the estimation of Z_M has been computed in section 6.3. The final results, i.e. the values of $\kappa_h(z)$ that are later used to test predictions of HQET are given table 16.

5.2 Simulating a doublet of degenerate dynamical quarks

In section 3.5 the question arises how we compute general expectation values in lattice field theory. We recall

$$\langle \mathcal{O} \rangle \equiv \langle [\mathcal{O}]_{\text{F}} \rangle_U = \left\{ \mathcal{Z}_{\text{eff}}^{-1} \int \mathcal{D}[U] [\mathcal{O}]_{\text{F}}[U] e^{-S_{\text{eff}}[U]} \right\}, \quad \mathcal{Z}_{\text{eff}} = \int \mathcal{D}[U] e^{-S_{\text{eff}}[U]}, \quad (3.49a)$$

$$S_{\text{eff}}[U] = \mathcal{S}_{\text{G}}[U] - \ln [\det (M[U])]. \quad (3.49b)$$

with $\mathcal{D}[U] = \prod_{x,\mu} dU(x,\mu)$. Such a high dimensional integral can only be solved numerically. The best suited technique to do that is the (quasi) *Monte-Carlo* method. The path integral is replaced by a *sample average* over an ensemble of N gauge field configurations $\{U_n | n = 1, \dots, N\}$,

$$\langle \mathcal{O} \rangle \equiv \langle \mathcal{O}[U] \rangle_U \equiv \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{n=1}^N \mathcal{O}(U_n) = \frac{1}{N} \sum_{n=1}^N \mathcal{O}(U_n) + \mathcal{O}(1/\sqrt{N}). \quad (5.18)$$

This is *exact* in the limit $N \rightarrow \infty$, but when truncated to just N so-called measurements, the residue is of order $N^{-1/2}$. This is the convergence rate predicted by the law of large numbers, i.e. for statistically independent measurements. But in a real simulation successive measurements are not independent. Thus it is crucial to get a reliable estimate for the uncertainty of observables in Monte-Carlo simulations and an estimate of the correlation among those. There are different methods to determine this *statistical error*, like Jackknife or Bootstrap. Beside Jackknife we will use the so-called T -method which is based on an explicit evaluation of autocorrelation functions. Sometimes we also make consistency checks by using both methods. They are elucidated in more detail in appendix D.

Before one is able to compute observables, the gauge field configurations have to be generated. The method of choice for dynamical fermion simulations is the *Hybrid Monte-Carlo* (HMC) [98] algorithm or rather one of the various implementations thereof. This is an ergodic algorithm that combines the Markovian Langevin approach of stochastic processes with a microcanonical simulation [149]. I will sketch the general idea in the following:

The pseudo-fermion field

The Grassmannian nature of fermions forbids us to put them in a straightforward way in a computer as we do it with bosons. But the simplicity of the Grassmann algebra allows us to handle them by explicitly integrating out the fermion bilinear in the functional integral,

$$\int \mathcal{D}[\bar{\psi}] \mathcal{D}[\psi] e^{-\bar{\psi} M \psi} = \det M, \quad (5.19)$$

in favour of the fermion determinant which gives the effective action (3.49b). A priori $\det M[U]$ has neither to be real nor positive but *importance sampling*, a synonym for the use of a Boltzmann weight, requires positive definiteness. However, the γ_5 -hermiticity property of the Wilson fermion matrix, $M^\dagger = \gamma_5 M \gamma_5$, at least implies that $\det M$ is real. Positivity is still a problem for an odd number of fermion fields, but we directly simulate a doublet of massless degenerated quark

flavours. In this case the γ_5 -hermitean representation of the Dirac matrix can be exploited again, i.e. we can write

$$\det M^2 = \det [M^\dagger M] = \frac{1}{\det[M^\dagger M]^{-1}} \equiv \int \mathcal{D}[\phi^\dagger] \mathcal{D}[\phi] e^{-\phi^\dagger [M^\dagger M]^{-1} \phi}. \quad (5.20)$$

This bosonisation allows us to treat the fermion doublet by pseudo-fermions ϕ that inherit the spin and colour structure.¹⁹ Thus the Boltzmann weight respective probability measure becomes

$$p[U, \phi^\dagger, \phi] = \mathcal{Z}_{\text{eff}}^{-1} e^{-S_{\text{eff}}} = \mathcal{Z}_{\text{eff}}^{-1} \exp \left\{ -\mathcal{S}_G[U] - \phi^\dagger \left(M^\dagger[U] M[U] \right)^{-1} \phi \right\}. \quad (5.21)$$

Molecular dynamics

Take the (classical) Hamiltonian of your system to be like

$$\mathcal{H}[\pi, \phi] = \frac{1}{2} \sum_{x, \mu} \text{Tr} [\pi^2(x, \mu)] + S_{\text{eff}}[U, \phi^\dagger, \phi], \quad (5.22)$$

$$S_{\text{eff}}[U, \phi^\dagger, \phi] = \mathcal{S}_G[U] + \mathcal{S}_{\text{det}}[U] + \mathcal{S}_{\text{PF}}[U, \phi^\dagger, \phi], \quad (5.23)$$

where $\pi(x, \mu)$ will serve as the momentum conjugate of the bosonic field $\phi(x)$ [99, 100]. At this point a π is introduced as an independent parameter for every degree of freedom ϕ and can be taken as traceless hermitean matrices in the gauge group $\text{SU}(3)$. \mathcal{S}_{PF} is the pseudo-fermion part of the action. Thereby the probability measure that serves as the statistical (Boltzmann) weight in the path integral gets changed

$$p[U, \phi^\dagger, \phi] \rightarrow p[U, \phi^\dagger, \phi, \pi] = \mathcal{Z}_{\text{eff}}^{-1} e^{-\mathcal{H}[U, \phi^\dagger, \phi, \pi]}. \quad (5.24)$$

The momentum fields π do not contribute to expectation values because the exponential separates and the path integral is not performed over these fields such that they drop out. This approach allows to establish the equations of motion

$$\dot{\phi} = \pi, \quad \dot{\pi} = -\frac{\partial S_{\text{eff}}}{\partial \phi}, \quad (5.25)$$

in terms of a fictitious Monte-Carlo or *Molecular dynamics time*. That means, the evolution of the system and hence the sequence of generated gauge configuration depends on this virtual time. Due to Liouville's theorem they define an area-preserving reversible mapping on the phase space $\{\phi, \pi\}$ which can be passed through by a suitable iteration procedure, called *update*. To evolve the system it is necessary to solve these equations which is analytically impossible in practice. Thus it has to be done numerically by introducing a smallest time interval $\delta\tau$ and a piecewise integration of the equations of motion for \mathcal{H} over some time interval, the *trajectory length* τ . *Integrators* that avoid finite-step-size errors for the equations of motion are the heart of the HMC method [98]. A

¹⁹Actually, this also enables to simulate odd numbers of flavours by representing one of them as $|\det M| = \det \sqrt{M^\dagger M}$. In each case numerically efficient ways of representing $\sqrt{M^\dagger M}$ in the pseudo-fermion action are needed. This leads to the variants called PHMC [101] or RHMC [102].

class of numerical integrators that almost conserve \mathcal{H} and retain the most important properties of *reversibility* and *area-preserving* are the *symplectic integrators*. These are integrators that preserve the symplectic structure of the classical phase space. The leap-frog, the Omelyan [103] and the Sexton–Weingarten [104] integrators belong to this class. A general HMC algorithm now reads as follows:

- draw new conjugate momenta $\pi(x, \mu)$ from the Gaussian distribution

$$P[\pi] \propto \exp \left\{ -\frac{1}{2} \text{Tr} [\pi^2(x, \mu)] \right\} \quad (5.26)$$

- generate each flavour of pseudo-fermion fields ϕ_i by a *heatbath algorithm* according to the probability distribution

$$P[\phi_i] \propto \exp \left\{ -\phi_i^\dagger [M^\dagger M]^{-1} \phi_i \right\} \propto \exp \left\{ -R^\dagger R \right\}, \quad \phi_i = MR, \quad (5.27)$$

this is usually done by generating a complex Gaussian random vector R

- integrate the equations of motion over a time interval τ using a reversible symplectic integrator with step-size $\delta\tau$, such that $\{\pi, \phi\} \rightarrow \{\pi', \phi'\}$
- *Metropolis accept/reject step* [105]: accept the proposed configuration $\{\pi', U'\}$ as entry in the Markov chain with probability

$$P_{\text{acc}} = \min \left[1, e^{-\Delta\mathcal{H}} \right], \quad \Delta\mathcal{H} = \mathcal{H}[\pi', U', \phi_i] - \mathcal{H}[\pi, U, \phi_i]. \quad (5.28)$$

If the new entry is rejected, make another entry for ϕ . In the limit $\delta\tau \rightarrow 0$, also $\Delta\mathcal{H} \rightarrow 0$ and hence $P_{\text{acc}} \rightarrow 1$. However, even at finite $\Delta\mathcal{H}$ it can be proven that $\langle e^{-\Delta\mathcal{H}} \rangle = 1$ holds for a Metropolis-like algorithm. This method generates a transition probability that satisfies *detailed balance*.

For a badly chosen step-size $\delta\tau$ it can happen that large energy violations occur. In dynamical fermion simulations this is often due to large fermionic forces when solving the equations of motion, thus rendering the integrator unstable at this specific choice of $\delta\tau$. When setting up a Monte-Carlo simulation it is essential to tune the simulation parameters in order to obtain a stable run. Furthermore the system has to be evolved for a sufficiently large time such that *thermal equilibrium* is established. This process is called *thermalisation* and necessary to obtain the right probability distribution. As should be clear after this discussion, the step-size is directly related to the computer time one has to spent to measure an observable within certain bounds of precision. Because computer time is limited one needs to speed up such computations. There exist several ways to achieve that, starting with the implementation of optimization techniques in the most expensive part of an existing algorithm. This is directly related to the used hardware/machine on which the simulation is performed. Also developing new algorithms is a fast growing field since years. It was also realized that badly chosen simulation parameters can lengthen the run time of a simulation. Therefore different pre-conditioning techniques are usually used to speed up

the convergence process. Without going into any details we now want to describe the particular algorithm that is used to perform our dynamical fermion simulations.

Our specific algorithmic choice

We follow [106, 107, 158] where the pseudo-fermion part in our algorithm was already discussed to some extent. The fermion matrix is defined by

$$a(D + \delta D + m_0) = (2\kappa)^{-1}M \quad (5.29)$$

D is the Wilson–Dirac operator and δD the $O(a)$ correction containing the clover term (3.37) and the boundary term (3.38). First, a Schur complement decomposition of the hermitean Dirac operator $\gamma_5 M$ is performed, giving

$$\gamma_5 M = \gamma_5 \begin{pmatrix} M_{ee} & M_{eo} \\ M_{oe} & M_{oo} \end{pmatrix} = \begin{pmatrix} \gamma_5 M_{ee} & 0 \\ \gamma_5 M_{oe} & 1 \end{pmatrix} \begin{pmatrix} 1 & M_{ee}^{-1} M_{eo} \\ 0 & \gamma_5 \{M_{oo} - M_{oe} M_{ee}^{-1} M_{eo}\} \end{pmatrix}. \quad (5.30)$$

It naturally splits the fermion matrix into an asymmetric and symmetric part, \widehat{Q}_A and \widehat{Q} respectively. This is called *even-odd preconditioning* [108] because it decomposes the fermion matrix in pieces according to their interaction with neighbouring sites. The parity of a lattice site is even or odd with respect to the sum of its coordinates $x_0 + x_1 + x_2 + x_3$. For us the most important quantities are:

$$\widehat{Q}_A = \widehat{c} \gamma_5 (M_{oo} - M_{oe} M_{ee}^{-1} M_{eo}), \quad \widehat{Q} = M_{oo}^{-1} \widehat{Q}_A, \quad \widehat{c} = (1 + 64\kappa^2)^{-1}. \quad (5.31)$$

While \widehat{Q}_A is hermitean, \widehat{Q} is not. However, the product $\widehat{Q}^\dagger \widehat{Q}$ still gives a positive-definite matrix. The pseudo-fermion and determinant part of the effective action now become

$$\mathcal{S}_{\text{PF}} = \phi_0^\dagger \left[\widehat{Q} \widehat{Q}^\dagger + \rho^2 M_{oo}^{-2} \right]^{-1} \phi_0 + \phi_1^\dagger \left[\rho^{-2} + \widehat{Q}_A^{-2} \right] \phi_1, \quad (5.32)$$

$$\mathcal{S}_{\text{det}} = (-2) \ln \det M_{ee} + (-2) \ln \det M_{oo}. \quad (5.33)$$

The determinants appearing here are computed exactly. Also *mass preconditioning* as introduced by Hasenbusch [109, 110] is applied, yielding the new parameter ρ . This additional parameter can be further tuned to reduce the condition number and hence the computing time. If we denote the minimal and maximal eigenvalues by $\lambda_{\min} \equiv \lambda_{\min}(\widehat{Q})$ and $\lambda_{\max} \equiv \lambda_{\max}(\widehat{Q})$, respectively, then an optimal choice for the virtual mass parameter ρ was found to be [106]

$$\rho_{\text{opt}} \equiv \{\lambda_{\max} \lambda_{\min}\}^{1/4}. \quad (5.34)$$

The proportionality is valid for the Schrödinger functional at fixed renormalized coupling due to (3.34). For the computation of the lowest eigenvalue the Kaltreuter-Simma-algorithm [111] is implemented.

The molecular dynamics is characterized by the integrator which can be leap-frog (LF) or Sexton-

L/a	$[\tau/\delta\tau_1; \delta\tau_1/\delta\tau_0]$	ρ	$\langle e^{-\Delta\mathcal{H}} \rangle$	$\langle U_p \rangle$	$\langle N_{CG}^{(0)} \rangle$	$\langle N_{CG}^{(1)} \rangle$	P_{acc}
20	14 ; 14	0.1	0.9932(79)	0.662308(11)	37	286	89%
16	16 ; 16	0.1	0.9977(45)	0.654403(16)	37	244	94%
12	14 ; 14	0.1	1.0022(20)	0.641353(18)	37	189	96%
10	12 ; 12	0.1	0.9988(8)	0.632308(13)	34	163	96%

Table 7: Algorithmic parameters for computations in volume L_0 with $T = 3L_0/2$. Here the Omelyan integrator was used with a trajectory length $\tau = 1$ in MD units. For the gauge force the SW integrator with $\delta\tau_0/\delta\tau_g = 4$ was used. (β, κ) are given in table 4.

L/a	Int.	$[\tau/\delta\tau_1; \delta\tau_1/\delta\tau_0]$	ρ	$\langle e^{-\Delta\mathcal{H}} \rangle$	$\langle U_p \rangle$	$\langle N_{CG}^{(0)} \rangle$	$\langle N_{CG}^{(1)} \rangle$	P_{acc}
40	SW	64 ; 5	0.0450	1.0032(38)	0.662409(2)	81	533	92%
32	SW	56 ; 4	0.0651	1.0009(48)	0.651034(3)	63	442	88%
24	SW	50 ; 5	0.0755	1.0006(45)	0.636756(8)	56	346	90%
20	SW	40 ; 4	0.0828	0.9896(84)	0.630066(13)	51	292	87%

Table 8: Algorithmic parameters for computations in volume L_1 with $T = L_1$ and $\theta = 0.5$. In all cases the trajectory length was set to $\tau = 2$ in MD units and for the gauge force, the SW integrator with $\delta\tau_0/\delta\tau_g = 4$ was used. (β, κ) are those of table 5.

Weingarten (SW)²⁰ and by the set of parameters $[\tau; \tau/\delta\tau_1; \delta\tau_1/\delta\tau_0; \delta\tau_0/\delta\tau_g]$. Here τ is the trajectory length, $\delta\tau_1$ and $\delta\tau_0$ the step-sizes of the second and first pseudo-fermion respectively and $\delta\tau_g$ is the step-size for the gauge part. The update of the SU(3) gauge fields is done by using SU(2) subgroups as proposed by Cabibbo–Marinero [112]. In table 7 and 8 we summarise some important algorithmic parameters at which the simulations specified in table 4 and 5 are performed, respectively. Most of the simulations reported here were carried out on the apeNEXT supercomputers systems provided by DESY/NIC²¹ in Zeuthen, Germany. See also appendix C for more details concerning the system architecture. As suggested in [107], all simulations are performed using trajectory length $\tau = 2$. It was found that this rather long trajectory length reduces the autocorrelation of successive measurements of physical observables in $N_f = 0, 2$ lattice QCD substantially.

For the simulations in the matching volume L_1 the parameter ρ was chosen close to its theoretical optimum, except for $L_1/a = 40$ where a more conservative value was taken for stability reasons. The acceptance rate P_{acc} for those runs is about 90% and the expectation value $\langle e^{-\Delta\mathcal{H}} \rangle$ for the three largest resolutions is identical to one as it should be. On the smallest lattice $\langle e^{-\Delta\mathcal{H}} \rangle$ is a bit more than one sigma away from one which is not that alarming. After a fixed number of trajectories we also performed reversibility test which have not shown any unexpected deviations. Furthermore, we list the plaquette average $\langle U_p \rangle$ and the averaged number of iterations of the solver for pseudo-fermion one $\langle N_{CG}^{(0)} \rangle$ and two $\langle N_{CG}^{(1)} \rangle$.

²⁰ also referred to as Multiple Time Scale Integration (MTSI)

²¹ DESY: Deutsches Elektron-Synchrotron, Hamburg (Germany)
NIC: John von Neumann-Institute for Computing

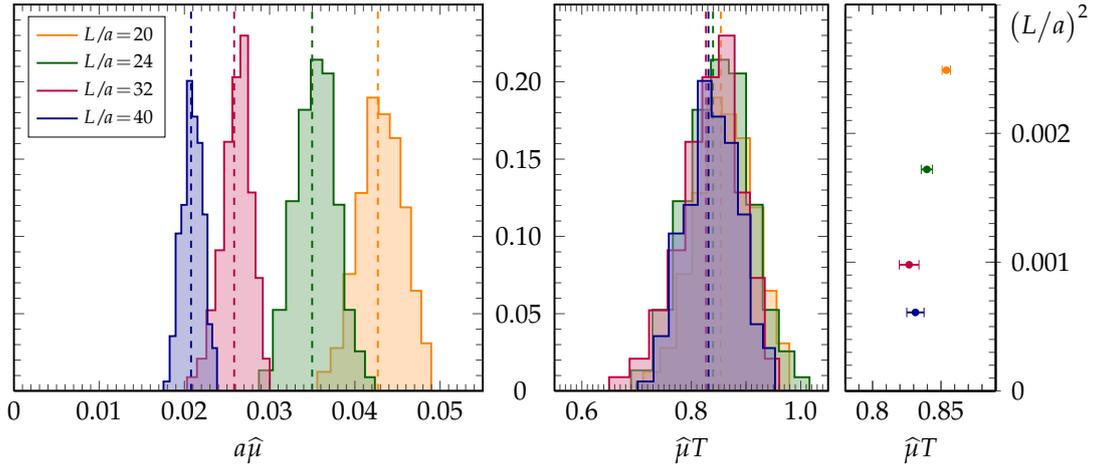


Figure 7: Normalised lowest-eigenvalue distribution of simulations at $L = T = L_1$ (left) and after a rescaling with T (center). The right plot shows the mean value of $\hat{\mu}T$ with the corresponding error as obtained from the MC history using the Γ -method. Each histogram consists of ten blocks.

Measurements

Due to high cost of dynamical fermion simulations the extraction of physical observables is done in two steps. First, one performs the dynamical fermion simulation as previously described. While the HMC algorithm evolves the system towards thermal equilibrium, one monitors the molecular dynamics history of certain observables. By a rule of thumb one usually continues the MD trajectory to a point where the systematics observed at the beginning accounts for approximately one third of the total Monte-Carlo history. This usually allows to estimate the corresponding autocorrelations near thermal equilibrium by analysing the observables starting at half of the total Monte-Carlo history. The autocorrelation times obtained give a hint which gauge configurations in molecular dynamic units can be expected to be statistically independent. Then a cost-benefit analysis gives an optimal value for the *save frequency* in MD units after which subsequent gauge configurations are stored to disk. In the best case they are really statistical independent. But this strongly depends on the algorithm, the point in parameter space where the simulation takes place and naturally on the observable in question. Among other things, autocorrelations usually scale to some power of the volume. However, the second step is to read the gauge field configurations into a measurement program that computes all desired correlation functions. We report about this in more detail in appendix B.5 and C.

In [113] the spectral gap of the hermitean Wilson–Dirac operator with periodic boundaries was used to diagnose the stability of the HMC algorithm and in [158] we reported about scaling tests performed with the same algorithm as discussed here but at non-vanishing quark mass in a large (fixed) volume of $2.4 \text{ fm} \times (1.8 \text{ fm})^3$. Here we can complement this picture by looking at the spectral gap of the Schrödinger functional in a physically small volume at vanishing quark mass.

The spectral gap revisited

We focus on the largest lattices simulated for this work, the matching volume L_1 with $L_1/a \in \{20, 24, 32, 40\}$ and run parameters as listed in table 5 and 8. For the choice of the Wilson–Dirac

(a) Simulations in volume L_1 :				(b) Simulations in volume L_2 :			
L/a	β	κ_c	am	L/a	β	κ_c	am
06	5.2638	0.135985	-0.00585(18)	12	5.2638	0.135985	+0.00798(14)
08	5.4689	0.136700	-0.00339(13)	16	5.4689	0.136700	+0.00044(8)
10	5.6190	0.136785	-0.00260(9)	20	5.6190	0.136785	-0.00103(20)
12	5.7580	0.136623	+0.00040(6)	24	5.7580	0.136623	+0.00136(6)

Table 9: HQET simulations in physical volumes L_1 and L_2 , see figure 5. Kinematic parameters are set to $T = L$, $\theta = 0.5$ and the bare couplings are such that $\bar{g}^2(L_1/2) = 2.989$. am refers to the PCAC mass corresponding to the hopping parameters given.

operator in the Schrödinger functional as recently discussed, the *spectral gap* is given by

$$a\hat{\mu} = \frac{1 + 64\kappa^2}{4\kappa} \sqrt{\lambda_{\min}(\hat{Q}\hat{Q}^+)}. \quad (5.35)$$

The leading normalisation appears to correctly match the quark mass to that of the free theory with periodic boundary conditions. Since the renormalized coupling and light quark mass are kept fixed, we expect the lowest eigenvalue to obey the scaling law $\lambda_{\min}(\hat{Q}\hat{Q}^+) \propto 1/T^2$ which will differ from the naive expectation (3.34) due to fluctuating gauge fields and the non-vanishing value $\theta = 0.5$. The result is presented on the left of figure 7 as normalized distribution of the spectral gap $a\hat{\mu}$. Obviously the gap decreases with increasing lattice resolution and also the width of the histogram gets smaller. A multiplication with T/a to cancel the expected scaling behaviour yields the plot in the middle. The agreement of the low-lying eigenvalue distributions is remarkable and shows the stability of the algorithm and very well confirms the classical expectations [71, 22] in a non-perturbative fashion. The rightmost plot finally shows the median values as obtained with the Γ -method. As estimate for the spectral gap I quote the mean value

$$\hat{\mu} T = 0.838(19), \quad (5.36)$$

with an uncertainty such that the outermost error bars in the data are enclosed.

Comments

Even if it is beyond the scope of this work, let me mention that the tuning of the run parameters for simulating the HQET side (c.f. fig. 5) was finished and most of the production runs already reached a target statistic of 8000 gauge field configurations in total for each lattice [114]. Only the largest lattices are still pending. The corresponding parameters are given in table 9. Note that to apply a step-scaling procedure the bare couplings and the critical hopping parameters have to be held fixed from $L_1/a \rightarrow 2L_1/a$. To follow the second strategy proposed in [12] the same lattices but with $T = L/2$ have been simulated. We already measure HQET correlation functions to subleading order on those lattices and a corresponding first trial analysis is started.

6 Computation of improvement and renormalization factors

The goal of this section is to non-perturbatively compute the improvement factors $b_m(g_0)$ and $[b_A - b_P](g_0)$ as well as the renormalization constant $Z(g_0)$. For this purpose we closely follow the strategy proposed for the quenched case in [115, 97], extended to the parameter region typically used in $N_f = 2$ finite volume simulations. In order to study the heavy mass dependence of some QCD observables towards their HQET limit, we will only need b_m and Z in subsequent sections.

To not overload the following equations and as the renormalization scale dependence is unimportant for the purpose of this section, we write m_R as shorthand for $\bar{m}(\mu)$.

6.1 The PCAC relation

Mass-degenerate quarks

Using renormalized currents in a lattice regularised theory, the PCAC relation reads

$$\langle \tilde{\partial}_\mu (A_R)_\mu^a(x) \mathcal{O} \rangle = 2m_R \langle (P_R)^a(x) \mathcal{O} \rangle + \mathcal{O}(a), \quad (6.1)$$

as long as x is not in the support of \mathcal{O} . Otherwise contact terms would appear. Provided that the action and currents are properly $\mathcal{O}(a)$ -improved, the associated renormalized current quark mass is

$$m_R = \frac{1}{2} \frac{\langle \tilde{\partial}_\mu (A_R)_\mu^a(x) \mathcal{O} \rangle}{\langle (P_R)^a(x) \mathcal{O} \rangle} + \mathcal{O}(a^2). \quad (6.2)$$

As \mathcal{O} appears here in the nominator and denominator, and so the corresponding renormalization and improvement factors do, m_R is independent from the choice of a composite field operator up to terms of order a^2 which then mix with that from the currents itself.²² With the renormalization of the improved currents as given in eq. (2.44), the renormalization of the current quark mass is

$$m_R = \frac{Z_A(1 + b_A am_q)}{Z_P(1 + b_P am_q)} m, \quad m = \frac{1}{2} \frac{\langle \tilde{\partial}_\mu (A_I)_\mu^a(x) \mathcal{O} \rangle}{\langle (P_I)^a(x) \mathcal{O} \rangle} + \mathcal{O}(a^2), \quad (6.3)$$

where m is a bare current quark mass defined through some matrix element of the improved²³ PCAC relation. With the mass parameter renormalization of eq. (2.43) at hand, we can equate both renormalized masses and obtain the bare current quark mass in terms of the subtracted bare

²²Also the concrete choice of improvement conditions imposed to determine c_A introduce higher order lattice artefacts.

²³Please note that we are a bit sloppy here. While corrections in eq. (6.2) are truly of order a^2 , the bare current quark mass m of eq. (6.3) is not. It misses the factors containing the improvement coefficients b_χ which contribute as $\mathcal{O}(am_q)$ terms. Nevertheless, we always call m an $\mathcal{O}(a)$ -improved mass which we obtain from improved correlation functions.

quark mass:

$$m = \frac{Z_m Z_P}{Z_A} \times \frac{(1 + b_m a m_q)(1 + b_P a m_q)}{(1 + b_A a m_q)} m_q + \mathcal{O}(a^2) \quad (6.4)$$

$$= \frac{Z_m Z_P}{Z_A} \left[1 + (b_m + b_P - b_A) a m_q + \mathcal{O}([a m_q]^2) \right] m_q + \mathcal{O}(a^2). \quad (6.5)$$

The renormalization factor defined by

$$Z(\tilde{g}_0^2) = \frac{Z_m(\tilde{g}_0^2, a\mu) Z_P(\tilde{g}_0^2, a\mu)}{Z_A(\tilde{g}_0^2)}, \quad (6.6)$$

is finite as the divergence of Z_m is canceled by that of Z_P . Therefore Z is a function of \tilde{g}_0^2 only, but the second factor in (6.4) is not. The perturbative estimate of b_g in table 2 shows a weak dependence on g_0^2 and we could expand $Z(\tilde{g}_0^2)$ in terms of g_0^2 for small $g_0^2 b_g a m_q$,

$$Z(\tilde{g}_0^2) = Z(g_0^2) + \frac{\partial Z}{\partial g_0^2} [g_0^2 b_g a m_q] + \mathcal{O}([g_0^2 b_g a m_q]^2), \quad (6.7)$$

which would cause another contribution to (6.5),

$$m = Z \left[1 + \left(b_m + b_P - b_A + g_0^2 \frac{\partial \ln Z}{\partial g_0^2} b_g \right) a m_q + \mathcal{O}([a m_q]^2) \right] m_q + \mathcal{O}(a^2), \quad (6.8)$$

where Z is a function of g_0^2 now. In quenched lattice QCD one does not have to deal with \tilde{g}_0^2 because $b_g(g_0)$ vanishes in that case. Within our special case of $N_f = 2$ dynamical fermion simulations we still rely on $\tilde{g}_0^2 = g_0^2$ since we perform our simulations at vanishing sea quark mass while the valence quarks are treated as quenched. Without these simplification one would be forced to introduce further improvement terms as proposed in [116] for a more general treatment of this problem. In our case we do not need to care about b_g . To disentangle b_m from $b_P - b_A$ which appear at the same order, one needs to consider non-degenerated current quark masses. The key observation is the following: each quark flavour i gets separately improved but shares a common renormalization factor as in the degenerated case,

$$\tilde{m}_{q,i} = m_{q,i} (1 + b_m a m_{q,i}), \quad m_{R,i} = Z_m \tilde{m}_{q,i}. \quad (6.9)$$

Then a third renormalized mass defined as the mean of two given ones,

$$\frac{1}{2}(m_{R,1} + m_{R,2}) = Z_m \left\{ \frac{1}{2}(m_{q,1} + m_{q,2}) + b_m a \frac{1}{2}(m_{q,1}^2 + m_{q,2}^2) \right\}, \quad (6.10)$$

differs from that obtained by the renormalization of a mass whose mean, $m_{0,3} = \frac{1}{2}(m_{0,1} + m_{0,2})$, was taken at the level of bare quark masses,

$$m_{R,3} \equiv Z_m m_{q,3} (1 + b_m a m_{q,3}) = Z_m \left\{ \frac{1}{2}(m_{q,1} + m_{q,2}) + b_m a \frac{1}{4}(m_{q,1} + m_{q,2})^2 \right\}. \quad (6.11)$$

The difference of the quadratic term in eq. (6.10) and (6.11) can be used to estimate b_m and shows that building appropriate differences with non-degenerate quarks, allows to construct estimators for the improvement coefficient b_m and therefore $b_A - b_P$ separately.

Off-diagonal bilinear fields for non-degenerate quarks

The presence of non-degenerate masses breaks isospin symmetry. In this case one considers the *off-diagonal bilinear fields* we introduced in section 3.6. The PCAC relation now reads

$$\langle \tilde{\partial}_\mu (A_R)_\mu^\pm(x) \mathcal{O} \rangle = (m_{R,i} + m_{R,j}) \langle (P_R)^\pm(x) \mathcal{O} \rangle, \quad (6.12)$$

up to terms of $\mathcal{O}(a^2)$ with

$$(A_R)^\pm(x) = Z_A(g_0^2) \left[1 + b_A(g_0) \frac{1}{2} (am_{q,i} + am_{q,j}) \right] (A_I)_\mu^\pm(x), \quad (6.13)$$

$$(P_R)^\pm(x) = Z_P(g_0^2, a\mu) \left[1 + b_P(g_0) \frac{1}{2} (am_{q,i} + am_{q,j}) \right] (P_I)_\mu^\pm(x). \quad (6.14)$$

The involved renormalization constants Z_A and Z_P are the same as those in the theory with two mass-degenerate quarks. Using the corresponding $\mathcal{O}(a)$ improved versions of the SF correlation functions (3.55) the bare current quark masses are given by

$$m_{ij}(x_0) = \frac{\tilde{\partial}_0 f_A^{ij}(x_0) + ac_A \partial_0^* \partial_0 f_P^{ij}(x_0)}{2 f_P^{ij}(x_0)}, \quad (6.15)$$

and depend on the kinematic parameters $(L/a, T/L, \theta)$. Again the standard lattice derivatives are those of appendix B.3.1 but following [117, 115, 97], we will also use *derivative improved current quark masses*, defined by the replacements

$$\tilde{\partial}_0 \rightarrow \tilde{\partial}_0 \left(1 - \frac{1}{6} a^2 \partial_0^* \partial_0 \right), \quad \partial_0^* \partial_0 \rightarrow \partial_0^* \partial_0 \left(1 - \frac{1}{12} a^2 \partial_0^* \partial_0 \right), \quad (6.16)$$

to estimate b_m , $b_A - b_P$ and Z . When acting on smooth functions, these derivatives have an error of $\mathcal{O}(g_0^2 a^2, a^4)$. I provide further details in appendix B.3.1.

6.2 Strategy to compute $b_A - b_P$, b_m and Z

In case of non-degenerate quarks the bare current quark mass of eq. (6.4) now reads

$$m_{ij} = Z \frac{\left(1 + \frac{1}{2} b_P a (m_{q,i} + m_{q,j}) \right) \cdot \left(\frac{1}{2} (m_{q,i} + m_{q,j}) + \frac{1}{2} b_m a (m_{q,i}^2 + m_{q,j}^2) \right)}{1 + \frac{1}{2} b_A a (m_{q,i} + m_{q,j})} + \mathcal{O}(a^2) \quad (6.17)$$

$$\begin{aligned} &= Z \left[\frac{1}{2} (m_{q,i} + m_{q,j}) + \frac{1}{2} b_m (am_{q,i}^2 + am_{q,j}^2) \right. \\ &\quad \left. - \frac{1}{4} (b_A - b_P) a (m_{q,i} + m_{q,j})^2 \right] + \mathcal{O}(a^2), \quad (6.18) \end{aligned}$$

and reduces to the current quark mass of a single quark flavour, i.e. at $i = j$, where the associated hopping parameters κ_i and κ_j become equal. Now, forming ratios of suitable combinations of

degenerate and non-degenerate current quark masses in this representation, enables to derive direct estimators for $b_A - b_P$, b_m and Z :

An estimator for $b_A - b_P$

To isolate the coefficient $b_A - b_P$ in (6.18) one considers the combination,

$$2am_{12} - (am_{11} + am_{22}) = f(am_{q,1}, am_{q,2}) \quad (6.19)$$

which is an analytic function of the subtracted bare quark masses involved. Due to the obvious symmetry properties of (6.19), namely $f(x, y) = f(y, x)$ and $f(x, x) = 0$, which follows that of eq. (6.17), its series expansion can be cast in the form

$$f(x, y) = (x - y)^2 \sum_{n,k=0}^{\infty} c_{nk} (x - y)^{2n} (x + y)^k, \quad (6.20)$$

for real coefficients c_{nk} . In fact all c_{nk} include the factor c_{00} . An explicit calculation shows

$$f(am_{q,1}, am_{q,2}) = c_{00} (am_{q,1} - am_{q,2})^2 \left[1 - \left(\frac{3}{2}b_A - b_m\right) (am_{q,1} + am_{q,2}) + O([am_q]^2) \right], \quad (6.21)$$

$$c_{00} = Z \frac{1}{2} (b_A - b_P). \quad (6.22)$$

Here only the quark mass dependent cutoff effects are written explicitly. In practice, all coefficients c_{nk} – if one would be able to work in the desired precision at all – still suffer from $O(a)$ terms which do not depend on the quark masses. One might think that first one needs to know the value of Z to get $b_A - b_P$ from c_{00} , but as it gets multiplied by $a(m_{q,1} - m_{q,2})$, one can cancel it by²⁴

$$am_{11} - am_{22} = Z (am_{q,1} - am_{q,2}) \left[1 - (b_A - b_P - b_m) (am_{q,1} + am_{q,2}) + O([am_q]^2) \right]. \quad (6.23)$$

Therefore, an estimator for $b_A - b_P$ is given by

$$R_{AP} \equiv \frac{2(2m_{12} - m_{11} - m_{22})}{(m_{11} - m_{22})(am_{q,1} - am_{q,2})}, \quad (6.24)$$

the expansion of which reads

$$R_{AP} = (b_A - b_P) \cdot \left[1 - \left(\frac{1}{2}b_A + b_P\right) (am_{q,1} + am_{q,2}) + O([am_q]^2) \right], \quad (6.25)$$

again up to mass independent $O(a)$ corrections. Accidentally, the $O(am_q)$ term does not depend on b_m but higher orders do.

²⁴ In analogy to $f(x, y)$ a second function $g(am_{q,1}, am_{q,2}) = am_{11} - am_{22}$ was introduced in the literature, c.f. [115]. It has the very same expansion but the leading power on $(x - y)$ is one instead of two. In principle, one could define such expansion function due to the symmetry properties for various combinations of masses and ratios thereof. But as we are only interested in the leading approximations this is without further insights.

An estimator for b_m

As stated before one needs to introduce a third mass in accordance with

$$m_{0,3} = \frac{1}{2}(m_{0,1} + m_{0,2}) \quad \Leftrightarrow \quad m_{q,3} = \frac{1}{2}(m_{q,1} + m_{q,2}) , \quad (6.26)$$

to disentangle the coefficient b_m . It suggests itself to look at

$$am_{12} - am_{33} = \frac{1}{4}b_m Z (am_{q,1} - am_{q,2})^2 \times \frac{1 + \frac{1}{2}b_P(am_{q,1} + am_{q,2})}{1 + \frac{1}{2}b_A(am_{q,1} + am_{q,2})} \quad (6.27)$$

$$= \frac{1}{4}b_m Z (am_{q,1} - am_{q,2})^2 \left[1 - \frac{1}{2}(b_A - b_P)(am_{q,1} + am_{q,2}) + \mathcal{O}([am_q]^2) \right] , \quad (6.28)$$

and shows that b_m can be estimated through

$$R_m \equiv \frac{4(m_{12} - m_{33})}{(m_{11} - m_{22})(am_{q,1} - am_{q,2})} . \quad (6.29)$$

Thus, its expansion neglecting quark mass independent lattice artefacts of $\mathcal{O}(a)$ becomes

$$R_m = b_m \left[1 - (b_m - \frac{1}{2}(b_A - b_P))(am_{q,1} + am_{q,2}) + \mathcal{O}([am_q]^2) \right] . \quad (6.30)$$

An estimator for Z

It is worth noting that an estimator for the combination $b_A - b_P - b_m$ directly follows from the definitions of R_{AP} and R_m , namely

$$R_{AP} - R_m \equiv \frac{2(2m_{33} - m_{11} - m_{22})}{(m_{11} - m_{22})(am_{q,1} - am_{q,2})} \quad (6.31)$$

$$= (b_A - b_P - b_m) + \mathcal{O}(am_{q,1} + am_{q,2}) , \quad (6.32)$$

and involves only correlation functions of mass degenerate quarks. The quark mass dependent cut-off effect in $R_{AP} - R_m$ follows immediately from the difference (6.25)–(6.30). With eq. (6.23) in mind, one already has an estimator for Z , namely $(m_{11} - m_{22})/(m_{q,1} - m_{q,2})$. But the knowledge of $b_A - b_P - b_m$ allows to cancel the leading mass-dependent cutoff effect by introducing the estimator

$$R_Z \equiv \frac{m_{11} - m_{22}}{m_{q,1} - m_{q,2}} + (b_A - b_P - b_m)(am_{11} + am_{22}) = Z + \mathcal{O}(a^2) . \quad (6.33)$$

This holds up to $\mathcal{O}(a^2)$ corrections if the correct value for $b_A - b_P - b_m = R_{AP} - R_m$ is inserted.

Concluding remarks on R_X

Since the bare current quark masses m_{ij} are functions of the time coordinate x_0 where the operator insertion of the currents A_0 and P takes place, so do the estimators R_X , $X = AP, m, Z$ depend on x_0 . Therefore, one has to specify which time slices are exactly used to compute the estimators R_X .

This choice must be considered as part of the *improvement conditions* which uniquely define the ratios R_X as smooth functions of g_0 .

Within the Schrödinger functional formalism usually a good plateau shows up for the bare PCAC masses starting at the middle of the time extension, $T/2$. This must not necessarily be the case for observables derived from these masses, especially when investigating improvement coefficients that naturally 'live' at order a . At least higher fluctuations are expected to show up, which depend on the quantity and the region in bare parameter space where the simulation takes place. However, as in the quenched case we choose our estimators to be averaged over time slices of the middle third of the time extension,

$$R_X \equiv \frac{a}{T - L/2 + a} \sum_{x_0=L/2a}^{(T-L/2)/a} R_X(x_0), \quad X \in \{AP, m, Z\}. \quad (6.34)$$

Note that the main reason for this is rather to be less dependent on a special time slice – which means to get a smoother and more reliable signal – than just decreasing the size of the error. Furthermore, since the number of time slices used for the plateau average scales with the lattice size, the physical size of this plateau is kept constant and hence respects our requirement of a constant physics condition to be used below.

It is a priori not clear how large intrinsic $O(a)$ ambiguities really are for the improvement coefficients, because any other estimate \tilde{R}_X , $X \neq Z$, would lead to another dependence on g_0 and therefore a/L . However, the difference $R_X - \tilde{R}_X$ must vanish in the continuum limit with a rate proportional to a/L for the improvement coefficients or even $(a/L)^2$ for the renormalization constant. These intrinsic ambiguities imply that the essential information lies in the correct g_0^2 -dependence resulting from working at fixed physics while varying β .

6.2.1 Setup of Improvement conditions at constant physics

We have two applications in mind. First, we want to non-perturbatively improve observables relevant in finite volume QCD simulations in the bare parameter space of the matching volume L_1 . Secondly, we want to provide the same estimators in a parameter region typical for large volume applications. This is where charm physics computations usually take place and where the impact of R_X is more pronounced than in our small volume simulations. We already know that the definition of the PCAC mass depends on the kinematic parameters of the Schrödinger functional. It was found to be advantageous to choose different values for each of the above mentioned parameter regions. Hence they are fixed to²⁵

$$C_k = C'_k = 0, \quad T/L = 3/2, \quad \theta = \begin{cases} 0.5 & \text{at } L = L_0 \text{ for B-physics applications} \\ 0.0 & \text{at } L = L_2 \text{ for charm physics appl.} \end{cases}. \quad (6.35)$$

²⁵In principle one could choose other values of θ for the valence quarks. But this would be some sort of mixed action and could introduce further ambiguities. Therefore we always use the sea-quark value of θ for the valence quarks here.

Table 10: Results after tuning the hopping parameters in our small volume computations in L_0 to different values of the dimensionless averaged PCAC masses ($L_0 m_{11}, L_0 m_{22}$) as specified in eq. (6.36) and (6.37). Set 1 and set 2 share the same sea quark content, given by $\kappa_l = \kappa_1 \approx \kappa_c$. The hopping parameters marked by a hat belong to set 2. In both sets the third hopping parameter is given exactly by eq. (6.39) and therefore available up to the intrinsic machine precision.

L/a	T/a	β	κ_i		Lm_{ii}	
					impr.lat.der.	std.lat.der.
10	15	6.1906	κ_1	0.1360160	-0.0053(7)	- 0.0054 (7)
			κ_2	0.1343182	+0.4929(6)	+ 0.5004 (6)
			κ_3	0.1351618	+0.2477(6)	+0.2505(6)
			$\hat{\kappa}_2$	0.1276218	+2.2909(5)	+ 2.5004 (7)
			$\hat{\kappa}_3$	0.1316853	+1.2313(6)	+1.2749(6)
12	18	6.3158	κ_1	0.1357930	-0.0062(10)	- 0.0062 (10)
			κ_2	0.1343782	+0.4952(9)	+ 0.5008 (9)
			κ_3	0.1350819	+0.2478(10)	+0.2499(10)
			$\hat{\kappa}_2$	0.1287549	+2.3475(7)	+ 2.5007 (10)
			$\hat{\kappa}_3$	0.1321803	+1.2453(8)	+1.2770(9)
16	24	6.5113	κ_1	0.1354410	-0.0101(13)	- 0.0101 (13)
			κ_2	0.1343872	+0.4922(13)	+ 0.4949 (13)
			κ_3	0.1349120	+0.2434(13)	+0.2444(13)
			$\hat{\kappa}_2$	0.1301456	+2.4068(11)	+ 2.4955 (13)
			$\hat{\kappa}_3$	0.1327405	+1.2552(12)	+1.2725(12)
20	30	6.6380	κ_1	0.1351630	+0.0177(10)	+ 0.0176 (9)
			κ_2	0.1343562	+0.5005(9)	+ 0.5024 (9)
			κ_3	0.1347584	+0.2608(9)	+0.2616(9)
			$\hat{\kappa}_2$	0.1309650	+2.4433(8)	+ 2.5039 (10)
			$\hat{\kappa}_3$	0.1330310	+1.2757(8)	+1.2878(9)

The second choice with $\theta = 0$ can be used in simulations like that in [118, 119] and we can take advantage of some of the gauge configurations already used for the computation of Z_A [78].

Even if it is clear from the context, I refer to the different choices on $\theta \in \{0.5, 0\}$ as condition 1 and condition 2 in the following. To complete the definition of the line of constant physics, values for the bare PCAC masses of the valence quarks must be selected. In order to get an impression of the quark mass dependence in physical observables, we consider in analogy to [97] two sets,

$$\text{set 1 :} \quad Lm_{11} \approx 0, \quad Lm_{22} \approx 0.5, \quad (6.36)$$

$$\text{set 2 :} \quad Lm_{11} \approx 0, \quad Lm_{22} \approx 2.5, \quad (6.37)$$

where throughout this section these bare current quark masses refer itself to the average

$$m_{ii} \equiv \frac{a}{T - L/2 + a} \sum_{x_0=L/2a}^{(T-L/2)/a} m_{ii}(x_0), \quad (6.38)$$

over the local $m_{ii}(x_0)$ estimates obtained by using the standard lattice derivatives in the PCAC relation instead of the improved ones. This has the advantage that they remain directly comparable

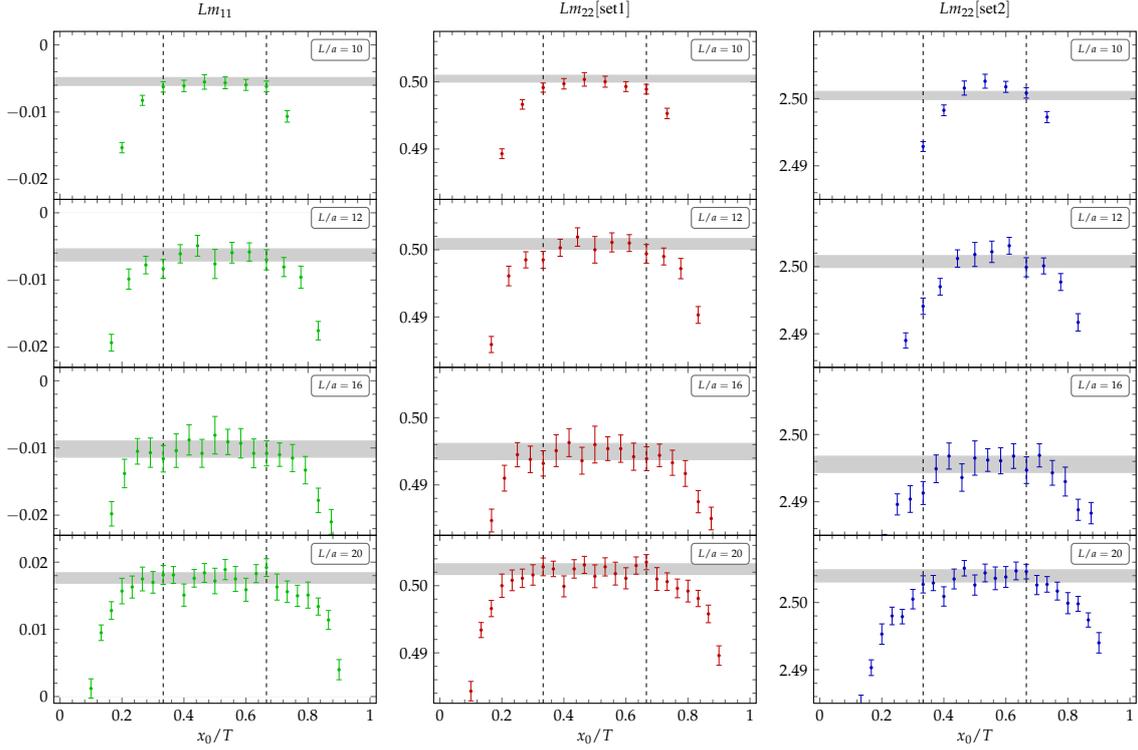


Figure 8: Time dependence of dimensionless PCAC masses obtained in small volume QCD simulation after tuning the hopping parameters of the valence quarks as given in table 10. The dashed lines define the plateau region used for averaging. Corresponding averaged masses are shown as error bands in each subplot. $L_0 m_{11} \approx 0$, $L_0 m_{22} \approx 0.5$ for set 1 and $L_0 m_{22} \approx 2.5$ for set 2 are achieved within the desired accuracy. For a direct comparison the scales are made equal.

to the (standard) PCAC masses one usually employs in other calculations.

The choice on Lm_{22} in set 1 is motivated by a quenched investigation [115]. It was argued based on experiences from PT to be advantageous with respect to the size of $O(a)$ ambiguities encountered. In contrast, the choice of $Lm_{22} \approx 2.5$ in set 2 is closer to the typical b-quark region itself.

After the gauge field configurations with $N_f = 2$ massless dynamical quarks were produced at values of $(L_0/a, \beta)$ as given in table 4, one needs one or two steps of tuning to properly adjust the hopping parameters κ_2 satisfying the conditions imposed on $L_0 m_{22}$.²⁶ κ_2 is responsible for the mass value of the heavy valence quark flavour. Furthermore, the hopping parameter κ_3 that controls the third mass that one needs to extract b_m , is defined by the requirement

$$m_{0,3} = \frac{1}{2}(m_{0,1} + m_{0,2}) \quad \Leftrightarrow \quad \kappa_3 = \frac{2\kappa_1\kappa_2}{\kappa_1 + \kappa_2} \quad (6.39)$$

and was directly implemented in the measurement code to avoid rounding errors. The final parameter values for calculating the ratios R_X in $L = L_0$ are given in table 10. Furthermore, we

²⁶ Similar to the situation in [97, 115], this is to sufficient precision equivalent to keeping the corresponding renormalized masses $L_0 Z_A m / Z_P$ fixed, as for the considered couplings the entering overall renormalization constant barely varies.

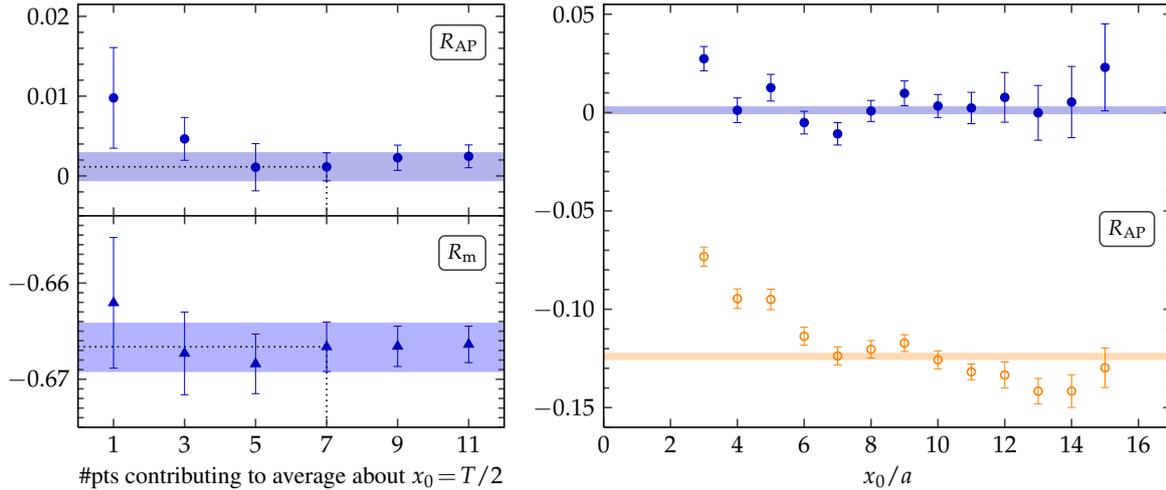


Figure 9: A typical data set, showing R_{AP} and R_m data at $\beta = 6.3158$ for set 1. *Left:* Dependence on the plateau size with preferred choice marked. *Right:* Time-slice dependence of $R_X(x_0)$ if plateau average in masses is omitted. Open symbols denote results using standard derivatives while filled symbols are obtained with improved lattice derivatives, eq. (6.16). Also the (symmetric) error bands are shown which correspond to results obtained with plateau averaged masses.

use the additional time-reflection symmetry as mentioned in section 3.6 and sum over forward and backward correlation functions without further notice.

6.3 Results for estimators R_X

Our method of choice for estimating observables in this section is the Γ -method as described in section D.2. As underlying statistic we produced 200 – 300 gauge field configurations at the parameters of table 5 but with $T = 3L/2$ and a save frequency of 5 or 10 in MD trajectory units. A larger time extent is chosen such that the plateau average, eq. (6.34), is less affected by finite-size effects coming from the boundaries. The estimators R_X , $X \in \{AP, m, Z\}$ are defined in terms of PCAC masses which itself are secondary observables build from SF correlation functions by means of eq. (6.15). The results for estimator with improvement condition 1 for set 1 and set 2 are listed in table 11.

The two plots on the right of figure 9 show the time slice dependence $R_{AP}(x_0)$ and $R_m(x_0)$ for set 1 at $\beta = 6.3158$. The averaged estimators which are our main results are shown as error bands. The blue points in those plots refer to the improved time derivative according to the re-

set	L/a	β	$b_A - b_P$	b_m	Z	$b_A - b_P - b_m$
1	10	6.1906	+0.0027(11)	-0.6576(15)	+1.10418(19)	+0.6637(8)
	12	6.3158	+0.0011(37)	-0.6666(26)	+1.10522(29)	+0.6653(13)
	16	6.5113	-0.006(11)	-0.6637(36)	+1.10395(29)	+0.6614(22)
	20	6.6380	-0.0050(29)	-0.6636(48)	+1.10333(23)	+0.6648(21)
2	10	6.1906	+0.07261(41)	-0.56417(38)	+1.09522(11)	+0.63820(37)
	12	6.3158	+0.05177(61)	-0.57800(66)	+1.09747(13)	+0.62971(60)
	16	6.5113	+0.02950(60)	-0.5955(10)	+1.09945(13)	+0.62617(69)
	20	6.6380	+0.02101(68)	-0.6077(11)	+1.09997(11)	+0.62914(71)

Table 11: Summary of improvement coefficients in volume $L = L_0$ for both sets of the heavy quark mass using plateau averaged masses and improved derivatives.

placement (6.16) while the orange data points are obtained using the standard derivatives. Their difference can become quite large especially for R_{AP} as can be inferred from the plot. This is expected to be a pure cutoff effect and underlines the necessity to use improved derivatives. A plateau in $R_{AP}(x_0)$ is visible in the data for both, standard and improved derivatives but it is more pronounced for the latter one.

For a careful error analysis we also looked at the dependence of the mean value and uncertainty on the number of points entering the plateau average about $x_0 = T/2$. This is shown in the left plot of figure 9. There the leftmost point is the one at $x_0 = T/2$ in the plots on the right. In the examples given here the plateau average is fast approached. This also nicely shows that a plateau average is less affected by local fluctuations which can distort the mean value by more than the associated uncertainty. If a plateau is visible, averaging gives a more reliable estimate for the mean value and the corresponding error. Furthermore, our results have been checked with a Jackknife error analysis.

The dependence of our main estimates R_X is visualized in the left plot of figure 10. The non-perturbative data points show a smooth dependence on g_0^2 as expected. Red data points correspond to our preferred choice, set 1. The blue points are those with improvement conditions of set 2. The qualitative dependence on g_0^2 is comparable to the quenched case [97]. However, note that agreement at the quantitative level is not expected since the simulations there are performed in a physical volume of approximately half of the present extent. Furthermore, effects of dynamical quarks which are taken into account here, always shift the values of the bare coupling by an approximate constant amount, thus making a direct comparison even less useful.

In our fixed volume L_0 the lattice resolution and thus the lattice size gets smaller with decreasing g_0^2 . Also the difference between both data sets decreases in this direction as it should, because their difference is dominated by mass-dependent cutoff effects. Except for R_{AP} the data points of set 1 show a significant deviation from the 1-loop perturbative estimate which is shown as dashed line. For R_m it is of $O(10\%)$ and for R_Z about 1%. Furthermore, we can provide a reliable error estimate that can not be given by perturbation theory. Computed on the same gauge field configurations the results of set 2 have a smaller error throughout. This is due to the much larger

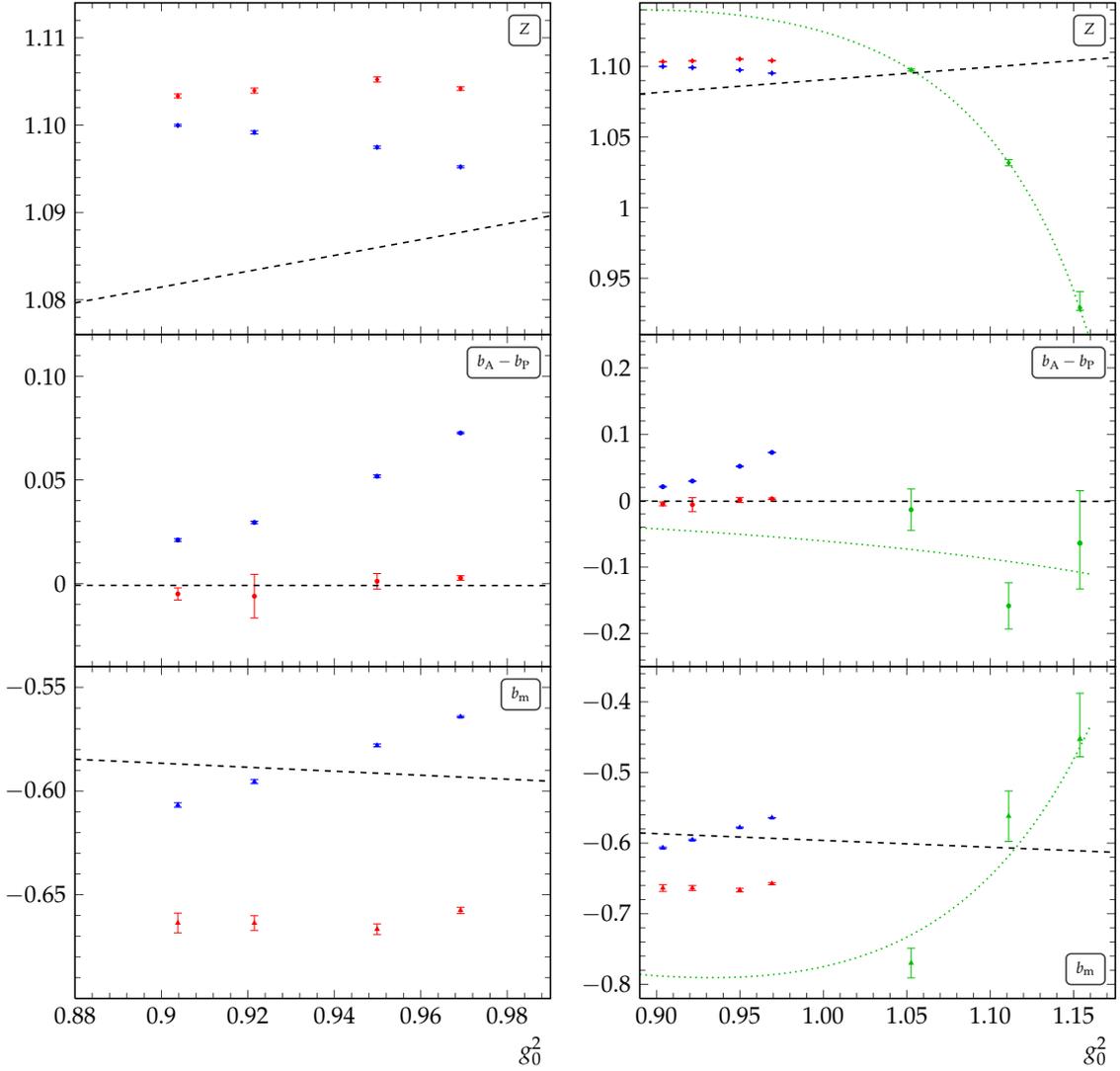
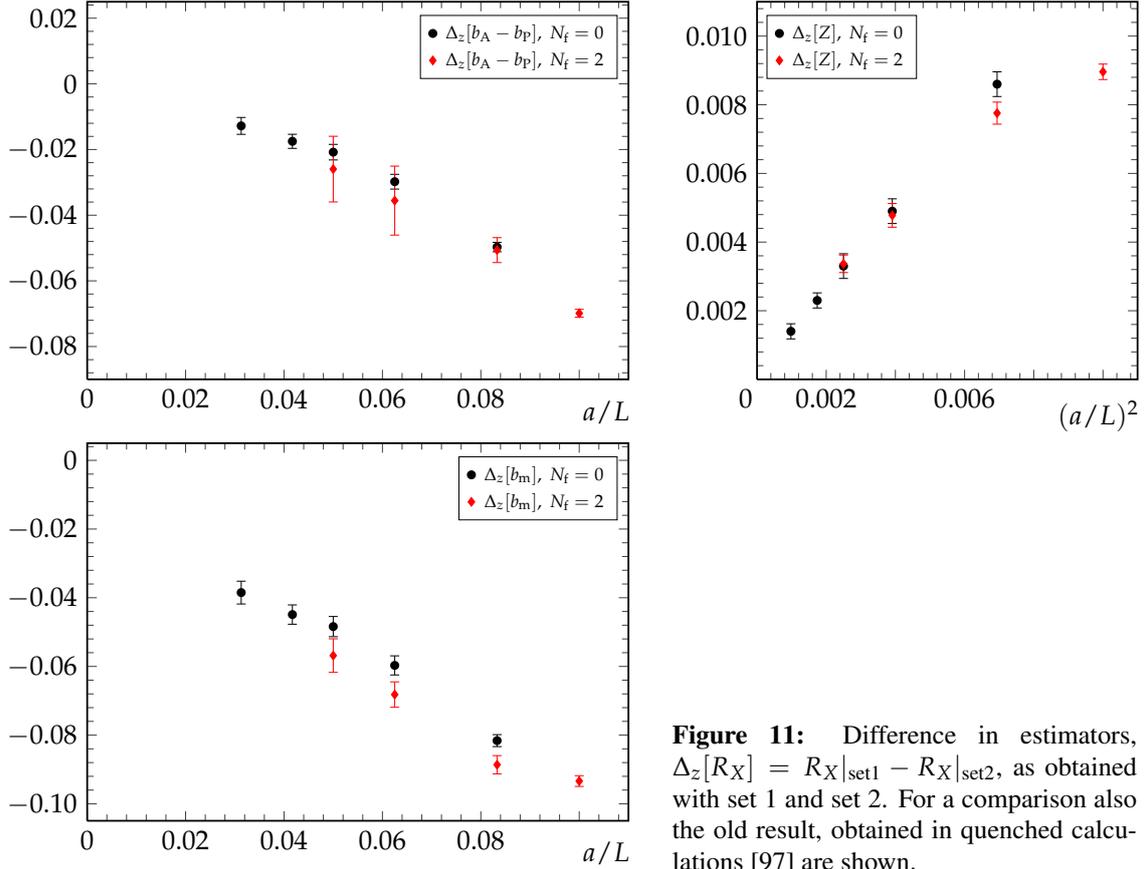


Figure 10: From top to bottom results for Z , $b_A - b_P$ and b_m are shown. Note that in each column the scale of the plots for the improvement coefficients was set to be equal. The dashed lines show the corresponding 1-loop perturbative estimation as given in table 2. *Left:* Results for simulations in small volume QCD at condition 1 for set 1 (red points) and set 2 (blue points). *Right:* Results of the plot on the left together with additional results obtained by simulations in the region of larger coupling, i.e. set 1 at condition 2 (green points).

mass, $L_0 m_{22}|_{\text{set2}} \approx 5 L_0 m_{22}|_{\text{set1}}$, where the dynamics is governed by the heavy quark flavour and fluctuations due to the dynamical massless quarks in the gauge background are not dominating anymore. Also the slope or curvature in the range of bare couplings considered here gets larger for data set 2.

Cutoff dependence

We mentioned that cutoff effects are already visible between the use of standard and improved time derivatives which are used to build PCAC masses. The reader may have wondered about the large differences between these two choices of a lattice derivative in figure 9. But one should keep



in mind that the results shown there are obtained at an intermediate lattice spacing with $L/a = 12$. We also performed the full analysis for each data set with the standard derivatives. Their difference was found to vanish in the limit $a \rightarrow 0$ as expected.

Furthermore, we had to make a choice for the non-vanishing quark mass in order to extract the estimators R_X . The reasons to compute all those estimators at two different sets of improvement conditions was also to obtain an impression of the intrinsic cutoff ambiguities [115]. To this end we define

$$\Delta_z[R_X(g_0)] \equiv R_X(g_0)|_{\text{set1}} - R_X(g_0)|_{\text{set2}}. \quad (6.40)$$

From eq. (6.25) and (6.30) one expects a linear a -dependence as mixtures of $\mathcal{O}(a/L)$ and $\mathcal{O}(am_q)$ terms in R_{AP} and R_m . According to eq. (6.33) the cutoff-dependence of R_Z should be $\mathcal{O}(a^2)$. For the values listed in table 11 we show the results for $\Delta_z[R_X]$ as red points (diamonds) in figure 11. The additionally shown black points (circles) are values obtained in the aforementioned quenched analysis [97]. Surprisingly, their absolute values are of the same magnitude. However, for the quantities under consideration the light quark content is the same for both sets and the differences (6.40) are taken at the fixed g_0 . Thus and in view of the expansions for distinct R_X made previously, we expect the contributions of the light quarks to largely cancel.

The scaling behaviour observed in the various differences of estimators (6.40) hence confirms the theoretical behaviour and shows the reliability of the applied non-perturbative method also in the

Table 12: Simulation parameters to compute improvement coefficients and renormalization constants in the charm region. This time θ has been chosen to be zero.

L/a	T/a	β	κ	#config \times #rep	total
12	18	5.20	0.135800	250×16	4000
			0.135700	250×16	4000
			0.135500	250×16	4000
16	24	5.40	0.136645	750×1	750
24	36	5.70	0.136704	328×2	656

	$T/3$ avr'd	$x_0 = T/2$	$T/3$ avr'd	$x_0 = T/2$	$T/3$ avr'd	$x_0 = T/2$
κ_1		0.135800		0.135700		0.135500
Lm_{11}	0.1051(30)	0.1072(47)	0.1584(32)	0.1668(47)	0.2495(30)	0.2547(43)
Lm_{22}	0.5015(23)	0.5015(37)	0.4956(28)	0.5016(39)	0.5006(28)	0.5041(37)
$b_A - b_P$	-0.098(50)	-0.078(95)	-0.040(42)	-0.048(74)	-0.086(25)	-0.108(45)
b_m	-0.453(28)	-0.413(43)	-0.513(25)	-0.511(43)	-0.485(20)	-0.424(36)
Z	0.9262(23)	0.9194(38)	0.9305(21)	0.9225(35)	0.9191(16)	0.9058(31)

Table 13: Results at $\beta = 5.2$ before extrapolating to vanishing light PCAC mass, $Lm_{11} \approx 0$.

case of $N_f = 2$ dynamical quarks. It also checks the universality of the continuum limit.

Determination for charm physics applications

Beside the estimators discussed in the previous sections which are of immediate use to us we also computed them in a parameter range which is relevant for charm physics computations in a large physical volume. These large volume simulations [118] are performed at slightly different couplings due to the rescaling in fixed volume L_2 as described in section 4.4 and shown in figure 5. Thus we will provide functions in g_0^2 for our estimators R_X that allow to interpolate to β -values relevant for large volume computations. In table 12 we summarise the simulation parameters relevant to us. At the largest bare coupling, i.e. at $\beta = 5.2$ and $L/a = 12$, three simulations only differing by the hopping parameter of the dynamical sea quark doublet are still available through the work reported in [78]. This parameter set $(\beta, L/a)$ lies in a region where it is impossible to simulate the Schrödinger functional at vanishing quark mass. Thus, one needs three different light quark masses to extrapolate the results to the chiral limit ($am_q \rightarrow 0$). To extract our improvement estimators R_X we need the second PCAC quark mass Lm_{22} to be tuned again to approximately 0.5. This could be achieved quite well as can be read off from table 13. The now non-vanishing values of Lm_{11} are also listed there. For this lattice we do not observe a pronounced plateau. Thus we list in table 13 results coming from our standard definition with an average taken over time-slices of the middle third together with the corresponding results obtained at $x_0 = T/2$ only. We investigated fits of R_X to a constant as well as linear and quadratic fits using two or all three points. Almost all quadratic fits have shown a strong curvature resulting in large uncertainty in the chiral limit with a mean value that differs up to a factor two from the linear and constant estimates for R_m and R_{AP} . The most stable estimates are obtained by a weighted average of the data using two

L/a	β	Lm_{22}	$b_A - b_P$	b_m	Z
12	5.20	0.4986(36)	-0.064^{+79}_{-69}	-0.486^{+65}_{-25}	0.9285^{+117}_{-16}
16	5.40	0.4996(20)	$-0.159(37)$	$-0.562(36)$	$+1.0319(22)$
24	5.70	0.5082(9)	$-0.014(21)$	$-0.770(21)$	$+1.0976(8)$

Table 14: Results of $b_A - b_P$, b_m and Z in the charm region. Asymmetric errors are our (conservative) estimates after extrapolation to $Lm_{11} = 0$.

or three points. Thus we choose as our favourite values at $Lm_{11} \approx 0$ the ones which comes from a weighted average of the two lightest points of the x_0 -averaged quantities. This is in accordance with the choice made in [78] for the computation of Z_A . The error is estimated such that it includes the range of the 3pt-linear extrapolations. Results for all values of β studied here are summarised in table 14. The other two simulations at $\beta = 5.4$ and 5.7 have a much smaller statistic compared to the previous ones. An analysis has shown that we cannot decrease the uncertainty in R_X much further within a reasonable time scale. Thus we decided to stop them at that point. As a result the estimates obtained here have an uncertainty which in general is one order of magnitude larger than in our small volume simulations. Even if we cannot compare both results obtained through different improvement conditions directly (due to different θ), we show both together on the right in figure 10. The trend in $R_Z(g_0^2)$ qualitatively follows that observed in small volume, but a fit to the data constraint with the one-loop PT value is away from those points. This is due to the fact that the three data points do not constrain the maximum of $Z(g_0^2)$ which is expected in view of the data and the perturbative estimate which is only approached at small g_0^2 . Note that we are only interested in a smooth description of the data with the correct $g_0^2 \rightarrow 0$ limit rather than seeking an analytic formula. The preferred choice to describe the data is by means of a Padé fit as done in the quenched case [115]. We do not obtain a meaningful fit to our data if the same function is used. Thus we tried other variations which do not affect the leading asymptotics predicted by PT, see also table 2. The fit

$$Z(g_0^2) = (1 + 0.090514g_0^2) \times \frac{1 - 0.447749g_0^4 - 0.181007g_0^6}{1 - 0.638143g_0^4} \quad (6.41)$$

describes the data in the corresponding range of g_0^2 very well. Note that for smaller couplings the fit result can vary quite drastically when changing only the highest power in the nominator. The results of the corresponding fits in the data range $5.7 \geq \beta \geq 5.2$ are remarkably stable. However, these variations are meaningless without further data points and the stable behaviour is due to the use of a Padé fit and the fact that no degrees of freedom are left. Due to the scatter in our data for R_{AP} our best fit estimate is

$$[b_A - b_P](g_0^2) = -0.00093g_0^2 \times \frac{1 + 23.5939g_0^2}{1 - 0.6235g_0^2}, \quad (6.42)$$

and represents the data within two sigma of its related uncertainty. Note that accidentally the parameter in the nominator is close to the quenched value, 23.306. Furthermore, note that even an unconstrained linear fit to the data points could represent the data within these error bounds. Also

set	L/a	β	$b_A - b_P$	b_m	Z	Z_P
1	10	6.1569	-0.0000(12)	-0.6633(12)	1.10443(17)	0.6065(9)
	12	6.2483	-0.0016(8)	-0.6661(9)	1.10475(12)	0.5995(8)
	16	6.4574	-0.0050(17)	-0.6674(23)	1.10455(17)	0.5941(10)
	20	6.6380	-0.0045(28)	-0.6692(27)	1.10379(17)	0.5949(12)
2	10	6.1569	+0.07852(53)	-0.56196(38)	1.09488(13)	
	12	6.2483	+0.06284(33)	-0.57145(29)	1.09632(8)	
	16	6.4574	+0.03567(51)	-0.59147(63)	1.09888(10)	
	20	6.6380	+0.02150(58)	-0.60763(78)	1.10021(10)	

Table 15: Summary of interpolated improvement coefficients and renormalization constants in volume $L = L_0$. The lattice results at $L/a = 20$ do not need to be extrapolated. Z_P was determined independently following the NP renormalization procedure of [29]. Both are just listed for completeness.

the interpolating fit

$$b_m(g_0^2) = (-0.5 - 0.09623g_0^2) \times \frac{1 - 0.5778g_0^6}{1 - 0.64216g_0^4} \quad (6.43)$$

describes the data within a confidence level of 95%. The high power in the nominator is needed to closer represent the charm data, especially at $\beta = 5.7$. Actually, there is no big difference to an interpolation by a straight line. But in this case no connection to PT at low g_0^2 can be established. After this side step to the charm region we now come back to the results obtained in L_0 and immediately apply them to be prepared for non-perturbative computations in small volume QCD, which can be used for the NP determination of HQET parameters as well as for the tests of HQET reported in section 7.

6.4 Hopping parameters at fixed RGI heavy quark mass $z = LM$

After setting up our line of constant physics by fixing $\bar{g}^2(L_0) = 2.989$ and $m_1 = 0$ in a finite volume of extend $L_0 = L_1/2 \approx 0.25$ fm, we produced appropriate gauge field configurations and measured the improvement coefficient $b_m(g_0)$ and the renormalization factor $Z(g_0)$. Because fixing the line of constant physics in L_0 is crucial for the next steps, a refined analysis has shown that the renormalized coupling was not matched within the desired accuracy. Hence, before starting the productions of gauge field configurations in volume L_1 we had to estimate the β -values for the three smallest resolutions $L_1/a \in \{20, 24, 32\}$ more carefully. This is reflected in the mismatch for $\beta = g_0^2/6$ between table 4 and 5. This forced us to inter-/extrapolate the non-perturbative estimates²⁷ in table 11 to the new values for β . For each set we have chosen 2-point linear and quadratic fits in g_0^2 as well as in β . Furthermore a global quadratic fit was applied. All results are in good agreement as can be imagined from the smooth g_0^2 -dependence of the data in figure 10.

²⁷ Actually, these estimates are our final values while those used in the fit were only computed on a subset. However, it has been checked that the variations observed between both results are negligible compared to the error that is associated to the values of z which is of approximately one percent.

	z	L/a			
		20	24	32	40
κ_s	0	0.1360536	0.1359104	0.1355210	0.135192285
κ_1	4	0.1327278	0.1332121	0.1335643	0.133651098
κ_2	6	0.1309498	0.1317899	0.1325495	0.132858327
κ_3	7	0.1300226	0.1310561	0.1320315	0.132455684
κ_4	9	0.1280709	0.1295337	0.1309715	0.131636658
κ_5	11	0.1259456	0.1279214	0.1298749	0.130797412
κ_6	13	0.1235550	0.1261898	0.1287348	0.129935253
κ_7	15	0.1206872	0.1242898	0.1275422	0.129046886
κ_8	18	—	0.1208919	0.1256259	0.127655927
κ_9	21	—	0.1151926	0.1234913	0.126177452

Table 16: Results for the hopping parameters corresponding to fixed dimensionless RGI heavy quark mass $z = LM$ in physical volume $L = L_1$ for different lattice resolutions. The hopping parameter of the sea quark κ_s are those in table 5.

We only list the results of the 2-pt linear fits in table 15 for further use. With those at hand we are now able to finish our discussion started in section 5.1 to fix the dimensionless RGI heavy quark mass $z = LM_h$ to some selected values for different lattice resolutions in physical volume $L_1 \approx 0.5$ fm. The important equations are

$$z = L_1 M_h = L_1 Z_M (1 + b_m am_{q,h}) am_{q,h}, \quad Z_M = h(L_0) \frac{ZZ_A}{Z_P(L_0)}, \quad (6.44)$$

with the universal coefficient $h(L_0) = 1.521(14)$. The corresponding hopping parameter reads

$$\kappa_h(z, L_1) = \left[\frac{1}{\kappa_c} - \frac{1}{b_m} \left(1 - \sqrt{1 + z \cdot \frac{4b_m}{[L_1/a]Z_M}} \right) \right]^{-1}. \quad (6.45)$$

The dependence of Z_A on g_0 was given in eq. (3.65) and we also computed Z_P non-perturbatively in course of our simulations in $L = L_0$ by the method described in section 3.7. Its (β -interpolated) values are listed in table 15, together with those for b_m and Z . Upper bounds on z restrict the largest values for $L_1 M_h$ that are possible on each lattice separately. They are given by $z_{\max} \approx 17.6, 21.3, 28.8, 36.3$ for $L_1/a = 20, 24, 32, 40$, respectively. Thus we choose as fixed values of the RGI heavy quark mass

$$z \in \{4, 6, 7, 9, 11, 13, 15, 18, 21\}, \quad (6.46)$$

where obviously $z = 18, 21$ cannot be simulated at $L_1/a = 20$. The corresponding hopping parameters on each lattice resolution are summarised in table 16.

Error of $z(g_0, L_1)$

The estimation of the dimensionless RGI heavy quark mass z depends on many quantities and their error, $z \equiv z[h, Z, Z_A, Z_P, b_m]$, which themselves depend on the parameters g_0 and maybe L_1 . Thus we expect by standard Gaussian error propagation the uncertainty in our estimation of z

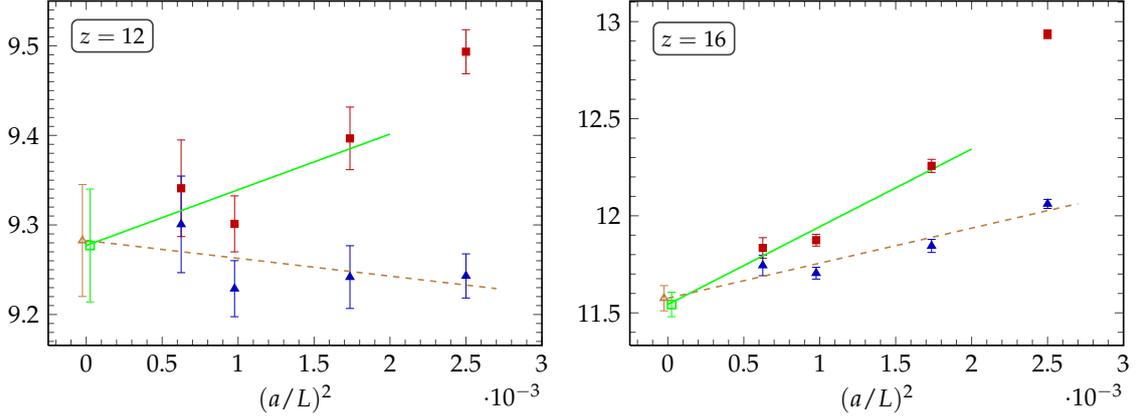


Figure 12: Unconstrained continuum extrapolations of $L_1\Gamma_{\text{PS}}$ computed on a subset of the available gauge configurations with different b_m and Z . The data obtained with set 1 (red squares) and data with set 2 (blue triangles) is fitted using three points (solid line) and four points (dashed line), respectively. Note the different scales in the plots.

to be

$$(\Delta z)^2 = \sum_i \left(\frac{\partial z}{\partial x_i} \right)^2 (\Delta x_i)^2 + \left(\frac{\partial z}{\partial h} \right)^2 (\Delta h)^2, \quad x_i \in \{b_m, Z, Z_A, Z_P\}. \quad (6.47)$$

The relative error is straightforward to compute:

$$\left(\frac{\Delta z}{z} \right)^2 = \left(\frac{b_m^2 am_{q,h}}{1 + b_m am_{q,h}} \right)^2 \left(\frac{\Delta b_m}{b_m} \right)^2 + \left(\frac{\Delta Z}{Z} \right)^2 + \left(\frac{\Delta Z_A}{Z_A} \right)^2 + \left(\frac{\Delta Z_P}{Z_P} \right)^2 + \left(\frac{\Delta h}{h} \right)^2. \quad (6.48)$$

With the estimate $\Delta Z_A/Z_A \approx 0.36\%$ as given in [78] and $\Delta Z_P/Z_P$, $\Delta Z/Z$ as well as $\Delta b_m/b_m$ taken from table 15, we obtain an accumulated relative error in the range $0.38\% \leq (\Delta z/z) \leq 0.41\%$. This holds for all values of z and lattice resolutions L_1/a without the contribution of $h(L_0)$. The latter contributes with $\Delta h/h = 0.92\%$ and hence is the main source of uncertainty here. This can in principle be reduced by increasing the precision of the step scaling functions involved in its definition [120]. As total uncertainty in the definition of fixed z we quote $\Delta z/z = 1.33\%$. The error of z has to be taken into account for any quantity regarded as function of z in such a way that the uncertainty on $h(L_0)$ is added in quadrature after a continuum limit of the corresponding quantity was taken.

O(a) ambiguities revisited

As a last more sophisticated check for the vanishing of $O(a)$ ambiguities through different improvement conditions and the universality of the continuum limit, we here consider the effective energy associated to the time component of the axial vector current in the matching volume, $L_1\Gamma_{\text{PS}}(z, g_0^2)$ to be defined in equation (7.1). We confront its lattice spacing dependence using b_m and Z of set 1 with that using b_m and Z of set 2 to fix the dimensionless RGI heavy quark mass to $z = 12, 16$.

Both data sets and its unconstrained continuum extrapolations linear in $(a/L)^2$ are displayed in

figure 12 and extrapolate to the same continuum limit. The assumed quadratic scaling behaviour is clearly violated for the coarsest lattice at $z = 16$ in set 1. For comparison we thus discard that point in our fits of data set 1. The fact that cutoff effects in $L_1\Gamma_{\text{PS}}$ for both cases are larger for set 1 is not surprising, since the improvement condition of set 2 with $L_0m_{22} \approx 2.5$ and thus $L_1m_{22} \approx 5$ is much closer to the line in parameter space with $z = 12, 16$ along which $L_1\Gamma_{\text{PS}}$ is computed. This general behaviour present in our new data was also observed in the quenched case [97]. Different scales used in both plots blur the fact that the slope of the green solid line at $z = 16$ is approximately 6.5 times larger than the corresponding one at $z = 12$. However, the excellent agreement of the continuum limits in both cases infer once more that our results correctly model the g_0^2 -dependence. No matter what set of improvement conditions is used, both entail convergence to the continuum limit with leading corrections of $\mathcal{O}(a^2)$. In consequence of this behaviour we conclude that the ambiguity introduced by choosing either set 1 or set 2 vanishes also for other values of z and other observables when the continuum limit is taken.

7 Non-perturbative tests of HQET

With the correct setup of an on-shell $O(a)$ improved lattice theory as discussed in the foregoing sections we are now able to build appropriate QCD observables to test the predictions made by HQET in the large quark mass limit.

7.1 Observables for non-perturbative tests of HQET

To study the heavy quark mass dependence of certain observables in ground state meson systems, we implement the representatives of eq. (4.5) and (4.6). According to section 3.8 the *effective pseudo-scalar mass* ($\sim m_B$) and *effective vector mass* ($\sim m_{B^*}$) read

$$\Gamma_{\text{PS}}(x_0) = -\tilde{\partial}_0 \ln [f_A^I(x_0)] , \quad \Gamma_V(x_0) = -\tilde{\partial}_0 \ln [k_V^I(x_0)] , \quad (7.1)$$

in terms of the improved SF correlation functions (3.61) and (3.62). The corresponding improvement coefficient c_A is known non-perturbatively (2.40), while for c_V we have to rely on its one-loop estimate (2.39a). Together, Γ_{PS} and Γ_V constitute the *spin-averaged effective mass*

$$\Gamma_{\text{av}}(L, M, \theta) \equiv \frac{1}{4} [\Gamma_{\text{PS}}(x_0) + 3\Gamma_V(x_0)] \Big|_{x_0=T/2} . \quad (7.2)$$

Note that $L_1 \Gamma_{\text{av}} \equiv \Phi_3^{\text{QCD}}(L_1, M)$ is the matching observable in (4.26,4.31). These observables depend on the θ -parameter as usual. Furthermore, we define the renormalized (finite) ratios

$$Y_{\text{PS}}(L, M, \theta) \equiv + \left[\frac{f_A^I(T/2, \theta)}{\sqrt{f_1(\theta)}} \right]_{\text{R}} , \quad Y_V(L, M, \theta) \equiv - \left[\frac{k_V^I(T/2, \theta)}{\sqrt{k_1(\theta)}} \right]_{\text{R}} , \quad (7.3a)$$

$$R_{\text{PS/P}}(L, M, \theta) \equiv - \left[\frac{f_A^I(T/2, \theta)}{f_P^I(T/2, \theta)} \right]_{\text{R}} , \quad R_{\text{PS/V}}(L, M, \theta) \equiv - \left[\frac{f_A^I(T/2, \theta)}{k_V^I(T/2, \theta)} \right]_{\text{R}} , \quad (7.3b)$$

$$R_{\text{spin}}(L, M, \theta) \equiv \frac{1}{4} \ln \left[\frac{f_1(\theta)}{k_1(\theta)} \right]_{\text{R}} , \quad R'_{\text{spin}}(L, M, \theta) \equiv \frac{1}{4} \left[\frac{f_1(\theta)}{k_1(\theta)} - 1 \right]_{\text{R}} . \quad (7.3c)$$

By using the notation $[\bullet]_{\text{R}}$ we do not list the renormalization and improvement constants that appear here explicitly. All renormalization constants are non-perturbatively known: $Z_A(g_0)$ and $Z_V(g_0)$ are taken from equation (3.65) and (3.66), respectively; $Z_P(\mu, g_0)$ was computed as explained in section 3.7 with the results listed in table 17. While for the improvement coefficient $[b_A - b_P](g_0)$ appearing in $R_{\text{PS/P}}$, we can use the non-perturbative estimate obtained in the last section, table 15, for $b_A(g_0)$ and $b_V(g_0)$ we use the one-loop estimates in table 2.

These test observables are supposed to be computed in finite volume QCD with $T = L = L_1$ for values of the dimensionless RGI heavy quark mass $z = L_1 M$ as given in eq. (6.46) at kinematic parameters $\theta \in \{0, 0.5, 1\}$. The time dependent correlation functions are evaluated at $x_0 = T/2$. In the *large volume limit* $L \rightarrow \infty$ the ratios Y_{PS} and Y_V become proportional to the *pseudo-scalar* and *vector heavy-light decay* constant, respectively. R_{spin} and R'_{spin} are two different estimates proportional to the *spin splitting* term. As discussed in section 4.3, these observables are well

L/a	Z_P	$c(L/a)$
40	0.5167(44)	0.999910000
32	0.5161(16)	0.999633931
24	0.5182(17)	0.999349405
20	0.5310(22)	0.999063427

Table 17: List of renormalization factors Z_P obtained from the relativistic runs at $L = L_1$ together with the applied tree-level normalization factor $c(L/a)$. c at $L/a = 40$ was not available and is an extrapolation. Its impact is beyond the accuracy that was reached in the MC simulation.

defined and approximated by their counterparts in the effective theory when approaching the static limit $1/M \rightarrow 0$. In fact, after inserting (4.21) into (7.3) and expanding to subleading order by taking account of (4.18), one gets

$$Y_{\text{PS}} \simeq X_{\text{R}} \left\{ 1 + \frac{Z_{\text{A}}^{(1)}}{Z_{\text{A}}^{\text{stat}}} + c_{\text{A}}^{\text{HQET}} \frac{f_{\delta\text{A}}^{\text{stat}}}{f_{\text{A}}^{\text{stat}}} + \omega_{\text{kin}} \left[\frac{f_{\text{A}}^{\text{kin}}}{f_{\text{A}}^{\text{stat}}} - \frac{f_1^{\text{kin}}}{2f_1^{\text{stat}}} \right] + \omega_{\text{spin}} \left[\frac{f_{\text{A}}^{\text{spin}}}{f_{\text{A}}^{\text{stat}}} - \frac{f_1^{\text{spin}}}{2f_1^{\text{stat}}} \right] \right\}, \quad (7.4a)$$

$$Y_{\text{V}} \simeq X_{\text{R}} \left\{ 1 + \frac{Z_{\text{V}}^{(1)}}{Z_{\text{A}}^{\text{stat}}} + c_{\text{V}}^{\text{HQET}} \frac{f_{\delta\text{A}}^{\text{stat}}}{f_{\text{A}}^{\text{stat}}} + \omega_{\text{kin}} \left[\frac{f_{\text{A}}^{\text{kin}}}{f_{\text{A}}^{\text{stat}}} - \frac{f_1^{\text{kin}}}{2f_1^{\text{stat}}} \right] - \frac{1}{3} \omega_{\text{spin}} \left[\frac{f_{\text{A}}^{\text{spin}}}{f_{\text{A}}^{\text{stat}}} - \frac{f_1^{\text{spin}}}{2f_1^{\text{stat}}} \right] \right\}, \quad (7.4b)$$

$$R_{\text{PS/P}} \simeq 1 + \frac{Z_{\text{A}}^{(1)} - Z_{\text{P}}^{(1)}}{Z_{\text{A}}^{\text{stat}}} + (c_{\text{A}}^{\text{HQET}} - c_{\text{P}}^{\text{HQET}}) f_{\delta\text{A}}^{\text{stat}} / f_{\text{A}}^{\text{stat}}, \quad (7.4c)$$

$$R_{\text{PS/V}} \simeq 1 + \frac{Z_{\text{A}}^{(1)} - Z_{\text{V}}^{(1)}}{Z_{\text{A}}^{\text{stat}}} + (c_{\text{A}}^{\text{HQET}} - c_{\text{V}}^{\text{HQET}}) f_{\delta\text{A}}^{\text{stat}} / f_{\text{A}}^{\text{stat}} + \frac{4}{3} \omega_{\text{spin}} f_{\text{A}}^{\text{spin}} / f_{\text{A}}^{\text{stat}}, \quad (7.4d)$$

$$R_{\text{spin}} \simeq \frac{1}{3} \omega_{\text{spin}} \frac{f_1^{\text{spin}}}{f_1^{\text{stat}}} \left\{ 1 - \omega_{\text{kin}} \frac{f_1^{\text{kin}}}{f_1^{\text{stat}}} - \frac{1}{3} \omega_{\text{spin}} \frac{f_1^{\text{spin}}}{f_1^{\text{stat}}} \right\}, \quad (7.4e)$$

$$R'_{\text{spin}} \simeq \frac{1}{3} \omega_{\text{spin}} \frac{f_1^{\text{spin}}}{f_1^{\text{stat}}} \left\{ 1 - \omega_{\text{kin}} \frac{f_1^{\text{kin}}}{f_1^{\text{stat}}} + \frac{1}{3} \omega_{\text{spin}} \frac{f_1^{\text{spin}}}{f_1^{\text{stat}}} \right\}, \quad (7.4f)$$

with abbreviation for the *static light decay constant* $X_{\text{R}} = Z_{\text{A}}^{\text{stat}} f_{\text{A}}^{\text{stat}} / \sqrt{f_1^{\text{stat}}}$. Yet we only considered the correspondence of the effective theory and QCD at the *classical level* and verified the static limits

$$\lim_{M \rightarrow \infty} Y_{\text{PS}}(L, M) \equiv X_{\text{R}}(L) \equiv \lim_{M \rightarrow \infty} Y_{\text{V}}(L, M), \quad (7.5)$$

$$\lim_{M \rightarrow \infty} R_{\text{PS/P}}(L, M) \equiv 1 \equiv \lim_{M \rightarrow \infty} R_{\text{PS/V}}(L, M), \quad (7.6)$$

$$\lim_{M \rightarrow \infty} R_{\text{spin}}(L, M) \equiv 0 \equiv \lim_{M \rightarrow \infty} R'_{\text{spin}}(L, M), \quad (7.7)$$

because $\omega_{\text{spin}} \propto 1/M \rightarrow 0$. As expected Y_{PS} and Y_{V} approach the *static-light decay constant* in finite volume and the spin splitting disappears.

In *quantum theory*, the scale dependent renormalization of the effective theory introduces logarithmic modifications to such relations. An example is the axial current in the effective theory [26], where the renormalized ratio X_{R} depends logarithmically on the chosen renormalization scale μ

$$X_{\text{R}}(L, \mu) = Z_{\text{A}}^{\text{stat}}(\mu) X_{\text{bare}}(L), \quad X_{\text{bare}}(L) = f_{\text{A}}^{\text{stat}} / \sqrt{f_1^{\text{stat}}}, \quad (7.8a)$$

$$X_{\text{RGI}}(L) = \lim_{\mu \rightarrow \infty} \left\{ [2b_0 \bar{g}^2(\mu)]^{-\gamma_0/(2b_0)} X_{\text{R}}(L, \mu) \right\} \quad \gamma_0 = -1/(4\pi^2), \quad (7.8b)$$

$$\equiv Z_{\text{RGI}}^{\text{stat}} X_{\text{bare}}(L). \quad (7.8c)$$

The scale dependence is removed explicitly by passing to the RGI matrix element X_{RGI} . The renormalization factor $Z_{\text{RGI}}^{\text{stat}}$ was non-perturbatively computed for $N_f = 2$ in [155]. Also the spin operator gets renormalized in the effective theory and by observation, eq. (7.4) and $\omega_{\text{spin}} = 1/(2M)$, we could define it as

$$Z_{\text{R}}^{\text{spin}}(L, \mu) = Z_{\text{spin}}(\mu) X_{\text{bare}}^{\text{spin}}(L), \quad X_{\text{bare}}^{\text{spin}}(L) = f_1^{\text{stat}}/(6f_1^{\text{stat}}), \quad (7.9a)$$

$$X_{\text{RGI}}^{\text{spin}}(L) = \lim_{\mu \rightarrow \infty} \{ [2b_0 \bar{g}^2(\mu)]^{-\gamma_0^{\text{spin}}/(2b_0)} Z_{\text{R}}^{\text{spin}}(L, \mu) \} \quad (7.9b)$$

$$\equiv Z_{\text{RGI}}^{\text{spin}} X_{\text{bare}}^{\text{spin}}(L), \quad (7.9c)$$

with some other qualified constant γ_0^{spin} . After expressing our QCD test observables through the corresponding RGIs and choosing the dimensionless RGI heavy quark mass to be $z = L_1 M$, their large mass behaviour is driven by the RGIs of the effective theory together with so-called *conversion functions*, C . In section 4.6 we already discussed how to obtain these functions in our matching scheme. They encode the full leading logarithmic mass dependence of the associated operators in the effective theory. As arguments of the latter we choose the ratio of RGIs M/Λ , since it can be fixed on the lattice without perturbative uncertainties [29]. The $1/z$ -expansions of our test observables now read as follows:

$$Y_X(L, M) \stackrel{M \rightarrow \infty}{\sim} C_X(M/\Lambda_{\overline{\text{MS}}}) \left[X_{\text{RGI}}(L) \right] \left(1 + \mathcal{O}(1/z) \right), \text{ for } X = \text{PS}, \text{V}, \quad (7.10a)$$

$$R_X(L, M) \stackrel{M \rightarrow \infty}{\sim} C_X(M/\Lambda_{\overline{\text{MS}}}) \left[1 \right] \left(1 + \mathcal{O}(1/z) \right), \text{ for } X = \text{PS}/\text{V}, \text{PS}/\text{P}, \quad (7.10b)$$

$$R_{\text{spin}}(L, M) \stackrel{M \rightarrow \infty}{\sim} C_{\text{spin}}(M/\Lambda_{\overline{\text{MS}}}) \left[X_{\text{RGI}}^{\text{spin}}(L)/z \right] \left(1 + \mathcal{O}(1/z) \right), \quad (7.10c)$$

$$L\Gamma_{\text{av}}(L, M) \stackrel{M \rightarrow \infty}{\sim} C_{\text{mass}}(M/\Lambda_{\overline{\text{MS}}}) \left[z \right] \left(1 + \mathcal{O}(1/z) \right). \quad (7.10d)$$

Note that the function C_{mass} was already introduced in eq. (4.41). The other functions are defined by eq. (4.37). Beside the overall logarithmic mass dependence governed by the conversion functions the leading matrix elements in the effective theory are enclosed in square brackets. The former follow from renormalization group considerations (c.f. [127]²⁸) and read in case of Y_{PS} for instance,

$$C_{\text{PS}}(M/\Lambda_{\overline{\text{MS}}}) \stackrel{M \rightarrow \infty}{\sim} \left(\ln \frac{M}{\Lambda_{\overline{\text{MS}}}} \right)^{-\gamma_0^{\text{PS}}/2b_0} \left\{ 1 + \mathcal{O}\left(\frac{\ln[\ln(M/\Lambda_{\overline{\text{MS}}})]}{\ln(M/\Lambda_{\overline{\text{MS}}})} \right) \right\}. \quad (7.11)$$

Due to the symmetries of the static effective theory, the conversion function $C_{\text{V}}(M/\Lambda_{\overline{\text{MS}}})$ shares the asymptotic behaviour to this order. To finally extract $(1/M)$ -corrections in HQET from our test observables as predicted by the asymptotics in eq. (7.10) the following steps are involved:

1. compute QCD test observables defined in eq. (7.2) and (7.3) non-perturbatively in L_1 at fixed z
2. compute perturbative conversion functions $C(M/\Lambda_{\overline{\text{MS}}})$, $M/\Lambda_{\overline{\text{MS}}} \propto z$, to high accuracy

²⁸or combine eqs. (1.21) and (7.8c)

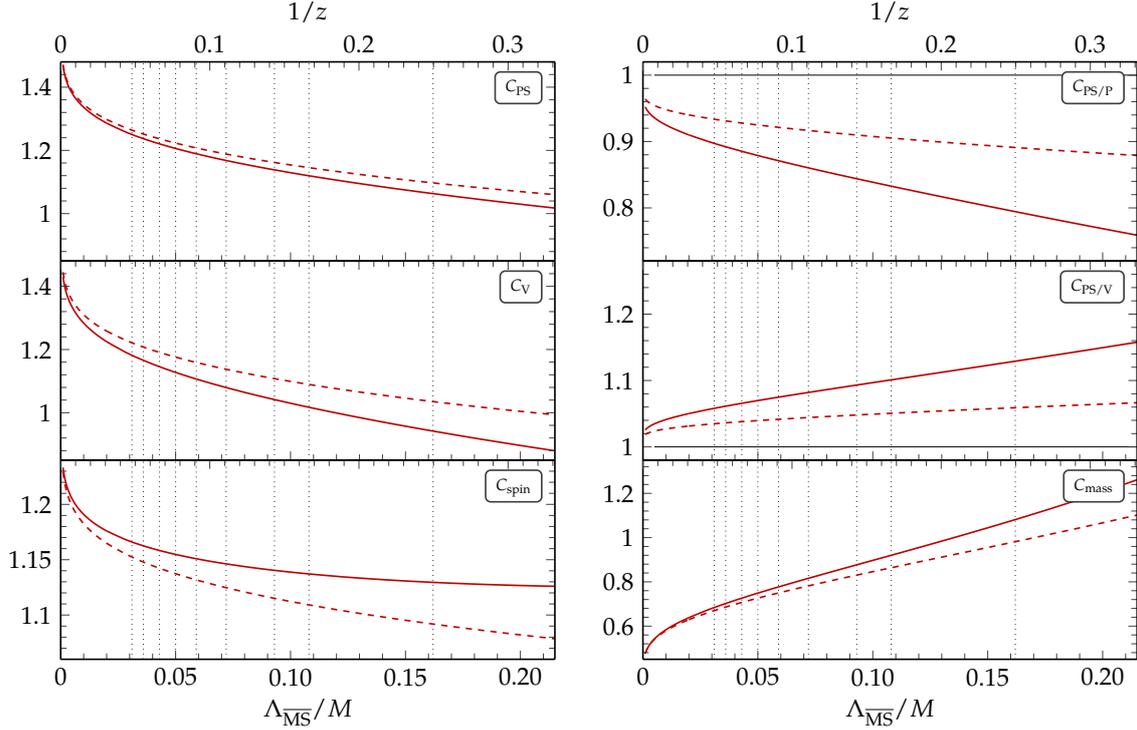


Figure 13: Quark mass dependence of various conversion factors C_X according to eq. (7.13)–(7.18). The solid and dashed lines corresponds to the 3-loop and 2-loop order anomalous dimensions in the matching scheme. The scale on top of each subplot is $1/z$ and $\Lambda_{\overline{\text{MS}}}/M$ at the bottom. Furthermore the investigated z values are plotted as vertical lines starting from $z = 4$ at the rightmost. Note that for comparison the plot range was chosen to be equal for C_{PS} and C_V as well as $C_{\text{PS}/P}$ and $C_{\text{PS}/V}$.

3. divide the test observables by the corresponding conversion function to get the leading matrix elements in the effective theory; only (7.10c) and (7.10d) need a further correction by the appropriate leading power of z
4. repeat this for all fixed values of the heavy quark mass z and extract the $1/z$ -dependence in the effective theory.

The perturbative conversion functions $C_X(M/\Lambda_{\overline{\text{MS}}})$, $X \in \{\text{PS}, \text{V}, \text{PS}/P, \text{PS}/V, \text{spin}, \text{mass}\}$ will be numerically computed along the lines of ref. [14] by solving the corresponding perturbative renormalization group equations.

7.2 Conversion functions between QCD and HQET

To compute the conversion functions appearing in (7.10) in our matching scheme, we need to evaluate (4.38). It is crucial to know the conversion functions to high accuracy in perturbation theory. Fortunately, through computations in recent years, all anomalous dimensions are now known up to three-loop. The only but not very appealing way to get an impression of the error introduced at some order of perturbation theory is to compare the two highest loop orders that are known. Thus we also compute the conversion factors using the two-loop estimates of the anomalous dimensions

(AD). In each case we use the four-loop β -function for the coupling. The part about heavy-light currents as summarized in [154] is of importance here and can be used as a guide to the original literature. All perturbative coefficients as they are used in our double precision calculations are listed in appendix E: the matching coefficients $\{c_i^{\mathcal{O}}\}$ of eq. (4.34) and (4.38) for heavy-light currents are listed in table 30. The coefficients $\{k_i\}$ appearing through the mass reparametrization (4.39) are summarized in table 29. The leading, universal anomalous dimensions $\{\gamma_0^{\mathcal{O}}\}$ are listed in table 27, while the two- and three-loop AD in the $\overline{\text{MS}}$ scheme can be found in table 28. As can be inferred from (4.38c) and (4.38d), we note again that an n -loop anomalous dimension only presupposes an $(n - 1)$ -loop matching coefficient. The parametrization of C_{PS} for the $N_f = 2$ theory was already given in [155]. It is convenient to represent the conversion coefficients C_X as smooth functions in terms of the variable

$$x \equiv 1 / \ln \left[M / \Lambda_{\overline{\text{MS}}} \right], \quad \frac{M}{\Lambda_{\overline{\text{MS}}}} = \frac{z}{L^* \Lambda_{\overline{\text{MS}}} \cdot r} = \frac{z}{0.64881}, \quad (7.12)$$

where $L^* \Lambda_{\overline{\text{MS}}}$ and r are known from (5.4) and (5.5), respectively. The functions decompose into a pre-factor which encodes the leading asymptotics as $x \rightarrow 0$, multiplied by a polynomial in x :

$$C_{\text{PS}}(x) = \begin{cases} x^{\gamma_0^{\text{PS}}/(2b_0)} \{1 - 0.107x + 0.093x^2\} & : \text{2-loop } \gamma^{\text{PS}} \\ x^{\gamma_0^{\text{PS}}/(2b_0)} \{1 - 0.118x - 0.010x^2 + 0.043x^3\} & : \text{3-loop } \gamma^{\text{PS}} \end{cases}, \quad (7.13)$$

$$C_V(x) = \begin{cases} x^{\gamma_0^{\text{V}}/(2b_0)} \{1 - 0.239x + 0.153x^2\} & : \text{2-loop } \gamma^{\text{V}} \\ x^{\gamma_0^{\text{V}}/(2b_0)} \{1 - 0.266x - 0.178x^2 + 0.193x^3\} & : \text{3-loop } \gamma^{\text{V}} \end{cases}, \quad (7.14)$$

$$C_{\text{spin}}(x) = \begin{cases} x^{\gamma_0^{\text{spin}}/(2b_0)} \{1 + 0.043x + 0.09x^2\} & : \text{2-loop } \gamma^{\text{spin}} \\ x^{\gamma_0^{\text{spin}}/(2b_0)} \{1 + 0.044x + 0.179x^2 - 0.099x^3\} & : \text{3-loop } \gamma^{\text{spin}} \end{cases}, \quad (7.15)$$

$$C_{\text{PS/P}}(x) = \begin{cases} 1 - 0.266x + 0.123x^2 & : \text{2-loop } \gamma^{\text{PS,P}} \\ 1 - 0.293x - 0.304x^2 + 0.284x^3 & : \text{3-loop } \gamma^{\text{PS,P}} \end{cases}, \quad (7.16)$$

$$C_{\text{PS/V}}(x) = \begin{cases} 1 + 0.136x - 0.052x^2 & : \text{2-loop } \gamma^{\text{PS,V}} \\ 1 + 0.142x + 0.250x^2 - 0.148x^3 & : \text{3-loop } \gamma^{\text{PS,V}} \end{cases}, \quad (7.17)$$

$$C_{\text{mass}}(x) = \begin{cases} x^{d_0/(2b_0)} \{1 + 0.373x + 0.176x^2\} & : \text{2-loop } \tau \\ x^{d_0/(2b_0)} \{1 + 0.287x + 0.752x^2 + 0.011x^3\} & : \text{3-loop } \tau \end{cases}. \quad (7.18)$$

The functional dependence is shown in figure 13. Solid curves represent the matching functions to highest available accuracy, i.e. three-loop anomalous dimensions of the involved operators and the four-loop β -function. The dashed curves are those obtained with two-loop anomalous dimensions. Furthermore, as vertical dotted lines we plot the values of z which have been fixed in order to non-perturbatively compute our test observables. For the wide range of masses covered in this work, the differences between the results for C_X obtained using two-loop or three-loop AD

vary between $\mathcal{O}(1\%)$ and $\mathcal{O}(10\%)$. Up to C_{PS} and C_{spin} , all conversion functions crack the 10% barrier at the smallest mass considered, $z = 4$. This is not surprising for $C_{\text{PS}/P}$ and $C_{\text{PS}/V}$ since their anomalous dimensions in the $\overline{\text{MS}}$ scheme of HQET are equal and only the corresponding perturbative matching coefficients c_i define the AD in our matching scheme. However, in the b-quark region itself, the difference between two- and three-loop AD is about 1% and contributions from higher order anomalous dimensions are expected to be even smaller. We apply the conversion factor after the continuum limit of each test observable was taken.

From our discussion at the end of section 6.4 where we exemplarily studied continuum extrapolations of $L\Gamma_{\text{PS}}$ at fixed $z = 12, 16$, we expect quite large cutoff effects especially at high values of z . Thus it is worth to look at another technique not mentioned so far which may improve the scaling behaviour to the continuum limit of our test observables.

7.3 Tree-level improvement of test observables

Symanzik's investigations of the cutoff dependence in field theories [121] can be generalized to the present case. In any lattice scheme a renormalizable observable $\Omega(\bar{g}^2, a/L)$ approaches its continuum counterpart $\Omega(\bar{g}^2)$ by means of

$$\Omega(\bar{g}^2, a/L) = \Omega(\bar{g}^2) \left\{ 1 + \delta_0(a/L) + \delta_1(a/L)\bar{g}^2 + \mathcal{O}(\bar{g}^4) \right\}, \quad (7.19)$$

$$\lim_{a/L \rightarrow 0} \delta_\ell(a/L) \equiv 0. \quad (7.20)$$

For the purpose of the present discussion we assume that the renormalized coupling in the continuum \bar{g}^2 has been fixed for all values of a/L . The lattice spacing effects are encoded in the functions $\delta_\ell(a/L)$ whose asymptotic expansion is given by

$$\delta_\ell(a/L) \stackrel{a/L \rightarrow 0}{\sim} \sum_{n=1}^{\infty} \left(\frac{a}{L}\right)^n \sum_{k=0}^{\ell} c_{n,k} \ln^k(a/L). \quad (7.21)$$

The coefficients $c_{n,k}$ which are different for each observable further depend on the gauge and fermion action used.²⁹ The convergence of $\Omega(\bar{g}^2, a/L)$ to the continuum limit can be perturbatively improved by canceling the cutoff effects $\delta_\ell(a/L)$ successively to each order of \bar{g}^2 . Obviously the ratio

$$\frac{\Omega(\bar{g}^2, a/L) - \Omega(\bar{g}^2)}{\Omega(\bar{g}^2)} = \delta_0(a/L) + \mathcal{O}(\bar{g}^2) \quad (7.22)$$

yields the tree-level corrections $\delta_0(a/L)$ in the limit $\bar{g}^2 \rightarrow 0$. The new observable, perturbatively improved to loop order l is then given by

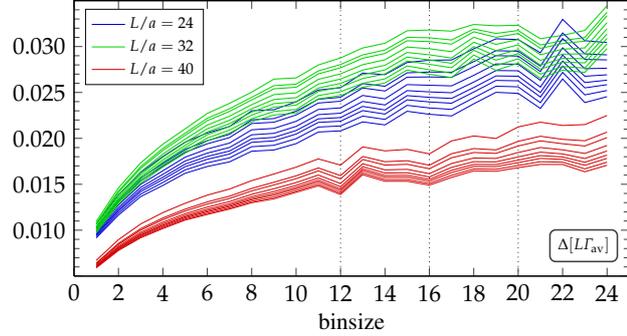
$$\Omega^{(l)}(\bar{g}^2, a/L) = \frac{\Omega(\bar{g}^2, a/L)}{1 + \sum_{\ell=0}^l \delta_\ell(a/L)\bar{g}^{2\ell}}. \quad (7.23)$$

²⁹In the quenched approximation, this dependence reduces to the gauge action.

Table 18: Statistic used to compute test observables in QCD. After pre-binning, 50 Jackknife-samples were used in the final analysis.

L_1/a	20	24	32	40
total meas.	400	632	800	1044
skipped	0	32	0	44
bin-size	8	12	16	20

Figure 14: Dependence of the Jackknife error $\Delta[LG_{\text{av}}]$ on the bin-size. Shown are all values of z for lattices $L_1/a \in \{24, 32, 40\}$. As the error should reach a plateau, one can infer an error of the error below 0.005.



One should notice the following point here: perturbative improvement only works *up to a fixed order in \bar{g}^2* with the advantage that *all* of the cutoff dependence gets removed to this order. By contrast, according to the discussion in section 2.2.1 non-perturbative $\mathcal{O}(a)$ -improvement removes the linear cutoff effects completely, i.e. including the perturbative contributions. Because we are working in an on-shell $\mathcal{O}(a)$ -improved lattice theory, the knowledge of the observables in perturbation theory at finite lattice spacing has to take care of the various improvement coefficients that appear in its definition. This is crucial to not reintroduce terms which are already canceled in the $\mathcal{O}(a)$ improved theory. As we expect from forgoing discussions large cutoff effects especially in observables at high values of the heavy quark mass, we also apply *tree-level improvement* [114] to our test observables, given by

$$\Omega^{(0)}(\bar{g}^2, a/L) = \frac{\Omega(\bar{g}^2, a/L)}{1 + \delta_0(a/L)}. \quad (7.24)$$

The correction factors $\delta_0(a/L)$ also depend on the kinematic parameters of the SF, i.e. basically on $\theta \in \{0, 0.5, 1\}$ here. Note that according to (7.21) no logarithmic contributions appear at tree-level.

7.4 Results

The relativistic heavy-light test observables are measured in the matching volume L_1 , setting the hopping parameter of the light quark to that of the sea quarks: $\kappa_1 \equiv \kappa_c$ as given in table 5. The heavy valence quark mass in our heavy-light currents is defined by setting the corresponding hopping parameter κ_h to the values in table 16, thus fixing $z = L_1 M$ for each lattice resolution $L_1/a \in \{20, 24, 32, 40\}$. The underlying statistic is given in table 18. To perform a subsequent analysis of the $1/z$ -asymptotics in our observables, we use the Jackknife method in this section throughout. To this end and to reduce autocorrelations, the available data was pre-binned for each L_1/a to get an overall statistic of 50 Jackknife samples. We take care of increased autocorrelations for increasing L_1/a by adjusting the block size in the binning procedure.

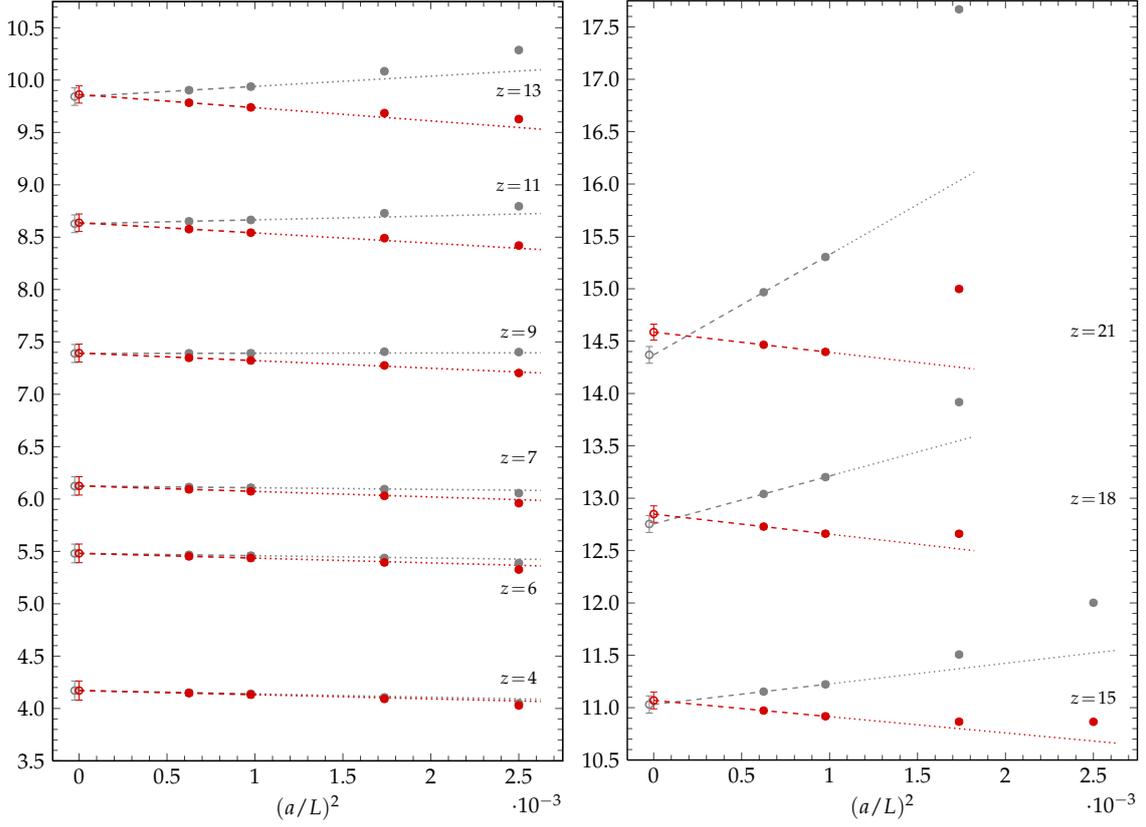


Figure 15: Continuum extrapolations linear in $(a/L)^2$ of the spin-averaged mass $L\Gamma_{\text{av}}(z, \theta = 0.5)$. The raw lattice data points are the gray ones while the tree-level improved results are plotted in red. The error of the lattice data is too small to be reasonably resolved at this scale, for details see table 32. Dotted lines extend the fit to those points which were not used for fitting. The continuum extrapolated points of raw data are slightly displaced for better visibility.

As an example for the observables under study, the continuum extrapolations linear in $(a/L)^2$ of $L\Gamma_{\text{av}}$ for $\theta = 0.5$ and all values of fixed z are presented in figure 15. The gray points refer to the (raw) observable $\Omega_{\text{av}} = L\Gamma_{\text{av}}(z, a/L)$, while the red points show its tree-level improved version $\Omega_{\text{av}}^{(0)}$ in accordance to eq. (7.24). The open symbols in the respective colour represents the continuum limit at fixed z . For consistency and comparability the continuum extrapolations presented only use the two largest lattice resolutions $L_1/a = 40, 32$. Thus they are shown as dashed lines while to guide the eye, the dotted lines just extend it to the data points excluded. The raw data points show the expected strong dependence on the heavy quark mass z , reflected by very large cutoff effects. At finite a/L the impact of tree-level improvement is already visible at moderate values of z . As is well-known, linear fits with only two data points do not have any degrees of freedom left. In those cases the continuum extrapolations are quite sensitive to cutoff effects. But as the data shows a remarkable consistency of the continuum limits between both data sets up to our second largest mass, $z = 18$, we can infer that the continuum limit is trustworthy and well-controlled. The behaviour of $\Omega_{\text{av}}(z, a/L)$ for $\theta = 0, 1$ is fully comparable. The continuum limits of the raw data are summarized in table 19 and the data points obtained at finite a/L_1 are

θ	z	$LI_{\text{av}}(L, M)$	$R_{\text{spin}}(L, M)$	$Y_{\text{PS}}(L, M)$	$Y_{\text{V}}(L, M)$	$R_{\text{PS/V}}(L, M)$	$R_{\text{PS/P}}(L, M)$
0	4	3.94(11)	0.02115(91)	-0.5175(37)	-0.5829(20)	0.9262(38)	0.851(16)
	6	5.25(10)	0.01667(72)	-0.5418(32)	-0.5898(19)	0.9499(31)	0.908(14)
	7	5.90(10)	0.01506(65)	-0.5515(30)	-0.5929(19)	0.9588(28)	0.931(13)
	9	7.17(10)	0.01261(54)	-0.5674(27)	-0.5985(19)	0.9725(24)	0.971(12)
	11	8.41(10)	0.01084(46)	-0.5798(26)	-0.6034(19)	0.9824(21)	1.003(10)
	13	9.63(10)	0.00951(40)	-0.5898(25)	-0.6075(20)	0.9898(18)	1.031(9)
	15	10.82(10)	0.00849(36)	-0.5977(24)	-0.6109(20)	0.9954(16)	1.055(9)
	18	12.54(9)	0.00683(76)	-0.6133(10)	-0.6224(9)	0.9985(6)	1.083(8)
	21	14.16(9)	0.00600(65)	-0.6248(10)	-0.6305(10)	1.0020(5)	1.107(7)
0.5	4	4.17(9)	0.01855(81)	-0.4805(33)	-0.5567(17)	0.8969(42)	0.807(13)
	6	5.48(9)	0.01472(64)	-0.5084(28)	-0.5646(16)	0.9281(34)	0.873(12)
	7	6.12(9)	0.01332(58)	-0.5192(26)	-0.5680(16)	0.9396(30)	0.900(11)
	9	7.39(9)	0.01118(49)	-0.5367(24)	-0.5737(16)	0.9569(25)	0.944(10)
	11	8.63(8)	0.00963(42)	-0.5501(22)	-0.5786(17)	0.9692(22)	0.980(9)
	13	9.84(8)	0.00847(37)	-0.5606(21)	-0.5828(17)	0.9783(19)	1.010(8)
	15	11.03(8)	0.00756(33)	-0.5690(21)	-0.5862(18)	0.9852(17)	1.035(7)
	18	12.75(8)	0.00591(63)	-0.5850(9)	-0.5974(9)	0.9903(5)	1.066(6)
	21	14.37(8)	0.00521(54)	-0.5966(9)	-0.6052(9)	0.9951(4)	1.091(6)
1	4	4.97(12)	0.01358(61)	-0.3965(26)	-0.5060(16)	0.8058(57)	0.677(16)
	6	6.24(11)	0.01095(50)	-0.4343(20)	-0.5163(15)	0.8611(43)	0.768(13)
	7	6.87(11)	0.00996(46)	-0.4483(19)	-0.5202(15)	0.8805(38)	0.804(12)
	9	8.11(10)	0.00843(39)	-0.4702(17)	-0.5267(16)	0.9093(31)	0.862(10)
	11	9.34(10)	0.00730(34)	-0.4865(16)	-0.5320(16)	0.9294(26)	0.907(9)
	13	10.54(9)	0.00644(30)	-0.4992(15)	-0.5363(17)	0.9440(22)	0.945(8)
	15	11.72(9)	0.00577(27)	-0.5091(16)	-0.5398(17)	0.9551(19)	0.976(8)
	18	13.43(9)	0.00431(49)	-0.5269(8)	-0.5505(9)	0.9665(7)	1.014(7)
	21	15.04(9)	0.00383(42)	-0.5395(8)	-0.5581(9)	0.9750(6)	1.044(6)

Table 19: Continuum extrapolations of QCD test observables (raw data)

listed in table 31 to 33 for $\theta = 0, 0.5, 1$, respectively. Furthermore, we can say that the perturbative tree-level improvement on top of our non-perturbatively $\mathcal{O}(a)$ improved observables does a quite good job and is obviously necessary above $z = 18$. From the distribution of the data points we can also reason that the improved data set may allow to incorporate the points at $L_1/a = 24$ to the continuum extrapolation up to $z = 18$. For the raw data this seems to be possible only to about $z = 12$. At $z = 21$ the effect of the additional improvement becomes most impressive. There, the relative deviation of the $L_1/a = 24$ raw data point from the corresponding continuum limit is about 20.6% and reduces to about 3.5% for the improved one.

In general the behaviour is as follows. Our volume of extent $L_1 \approx 0.5$ fm admits to reach heavy quark masses up to $M \approx 1.3M_b$. The $a \rightarrow 0$ extrapolations appear to be well controllable, provided that one accounts for the growing (heavy) quark mass in lattice units at given a/L as z is increased. This is similar to the quenched work [14] and realized by imposing a cut on aM ($aM \lesssim 0.7$) that translates, for given z , into the coarsest resolutions which may still be included in the continuum extrapolations. For all observables considered, the deviation of the extrapolating fit function from the result at the respective coarsest available lattice resolution grows with increasing z . This suggests that between $z = 18$ and $z = 21$ the $\mathcal{O}(a)$ -improvement and thus the a -expansion

has broken down for our lattices. But we should have in mind that the coarsest lattice at $z = 21$ had the very close theoretical limit $z_{\max} \approx 21.3$. Hence this behaviour is not really surprising as it just tells us quantitatively that $a \ll 1/M$ is not well-fulfilled any longer. This can be easily seen by noting $aM = z(a/L_1) = 21/24 = 0.875$ which also clearly violates our cut $aM \leq 0.7$. However, the continuum limits obtained are under reasonable control and hence we can use all points to extract the large $1/z$ -asymptotics. The results of the continuum limit extrapolations for the tree-level improved data set is now used and divided by the conversion functions as well as by the leading non-vanishing power of z to obtain non-trivial results. This is only necessary for $L\Gamma_{\text{av}}$ and R_{spin} as can be seen in eq. (7.10). The appropriate observables become

$$\Omega(z, \theta) \in \left\{ \left[L\Gamma_{\text{av}}(z)/(zC_{\text{mass}}(z)) \right], \left[(zR_{\text{spin}}(z))/C_{\text{spin}}(z) \right], \left[Y_{\text{PS}}(z)/C_{\text{PS}}(z) \right], \right. \\ \left. \left[Y_{\text{V}}(z)/C_{\text{V}}(z) \right], \left[R_{\text{PS/P}}(z)/C_{\text{PS/P}}(z) \right], \left[R_{\text{PS/V}}(z)/C_{\text{PS/V}}(z) \right] \right\}. \quad (7.25)$$

We aim for a description of our data in the full mass range considered, but also on a description which only incorporates the leading $(1/z)$ -asymptotics. Our polynomial fit ansatz reads

$$\Omega(z, \theta) = a_0(\theta) + a_1(\theta) \cdot x + a_2(\theta) \cdot x^2, \quad x = 1/z. \quad (7.26)$$

The fit constants $\{a_i\}$ implicitly depend on phase angle θ and it may be possible to perform a global fit. But this is not our concern here because computations of HQET observables to sub-leading order are also restricted to the fixed values of θ investigated here. Thus we focus on fits with θ kept fixed. All fits are performed within the 50 Jackknife samples for each observable in (7.25) separately and the error of the continuum results reported earlier are taken into account by appropriate weights. Depending on the finite-volume continuum observable $\Omega(z, \theta)$, the leading fit constant a_0 may be constrained to its value in the effective theory, i.e. $a_0 = 1$ for $\Omega \in \{L\Gamma_{\text{av}}/(zC_{\text{mass}}), R_{\text{PS/P}}/C_{\text{PS/P}}, R_{\text{PS/V}}/C_{\text{PS/V}}\}$. The results are listed in table 20 and shown in the plots of figure 16 and 17. Vertical shaded areas signal the region where the RGI bottom quark mass $z_{\text{b}} = L_1 M_{\text{b}}$ or the RGI charm quark mass $z_{\text{c}} = L_1 M_{\text{c}}$ are expected to lie. To this end I use the quenched results $M_{\text{b}} = 6.758(68)$ GeV and $M_{\text{c}} = 1.60(2)$ GeV as obtained in [12, 119], together with the approximately known scale $L^* \approx 0.63$ fm and eq. (5.5). The outcome is $z_{\text{b}} \approx 17.5$ and $z_{\text{c}} \approx 4.13$.

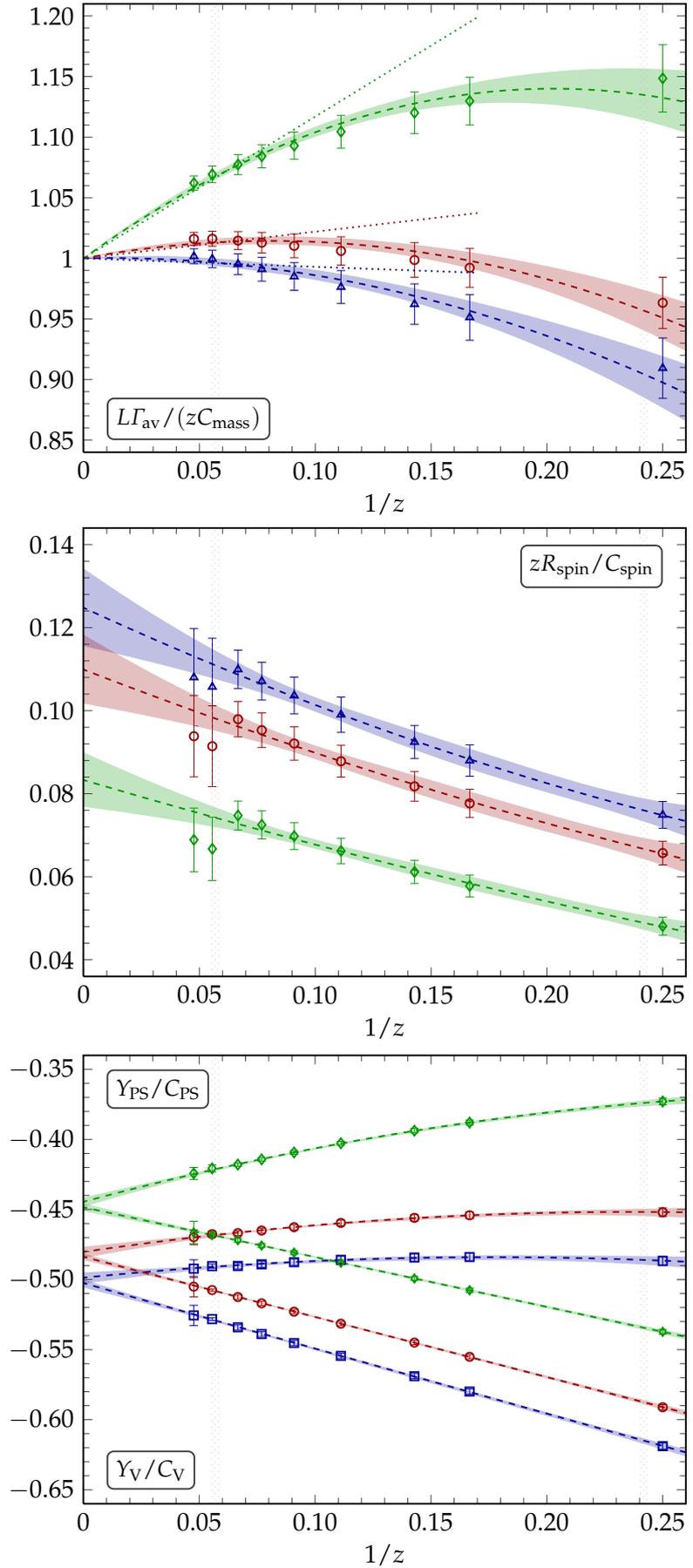
All continuum data points obtained show a very smooth dependence on $1/z$. The $(1/z)$ -asymptotic in the averaged mass for instance is very well consistent with the expected behaviour, shown on top of figure 16. A global constrained quadratic fit plotted as dashed line with corresponding error band describes all data points within their accuracy. We only show the tree-level improved data points and not the raw ones, but as can be inferred from the foregoing discussion, they are consistent to each other. Obviously, the slope in our data points increases with increasing values for θ , starting from 0.04 at $\theta = 0$ (□) and reaching 1.38 at $\theta = 1$ (◇). However, the error associated to our data points is nearly identical for each value of θ at fixed z . This is also reflected in a comparable error estimate across the fit parameters obtained. While the leading order asymptotics has a clear trend, the heavy quark dependence is smallest over the whole range of masses, $z = 4 - 21$, for $\theta = 0.5$ (○).

Figure 16:

$1/z$ -asymptotics of our relativistic heavy-light test observables in finite volume QCD. The leading logarithmic running as predicted by HQET is canceled by the corresponding conversion function C_χ . To obtain a finite result in the limit $1/z \rightarrow 0$ also leading mass dependence is canceled by the appropriate power of z . The dashed lines with corresponding error bands show a weighted quadratic fit to all data points at fixed θ , with $(\square, \circ, \diamond)$ corresponding to $\theta = (0, 0.5, 1)$.

Dotted lines show our estimates for the $1/z$ corrections as given in table 20. From top to bottom we have:

- (1) Spin-averaged, dimensionless heavy quark mass $L\Gamma_{\text{av}}/(zC_{\text{mass}})$.
- (2) Spin-splitting term.
- (3) The effective heavy-light meson decay constant in the pseudo-scalar and vector channel.



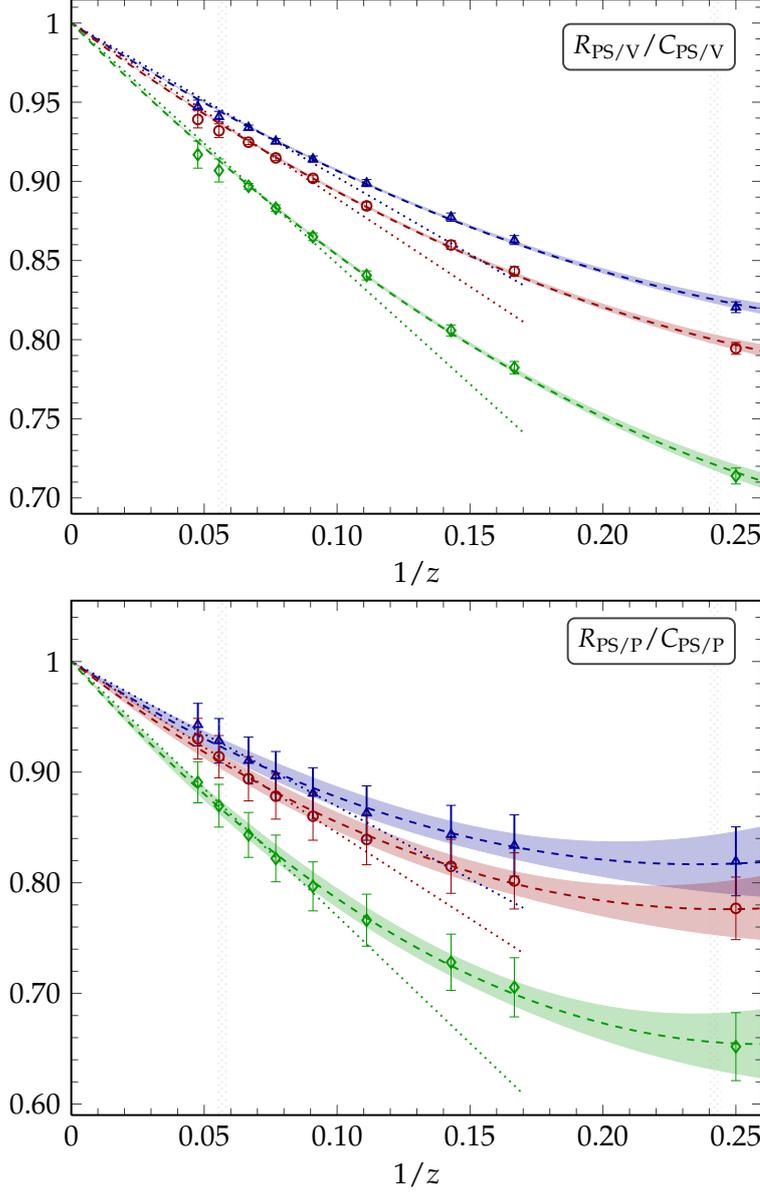


Figure 17: $1/z$ -asymptotics of ratio $R_{\text{PS}/\text{V}}/C_{\text{PS}/\text{V}}$ and $R_{\text{PS}/\text{P}}/C_{\text{PS}/\text{P}}$. The assignments are equal to those described in figure 16.

The case of the spin splitting terms is more intricate. First, the tree-level improvement term vanishes in that case so that no comparisons can be made in this direction. Secondly, we do not have a prediction from HQET because $X_{\text{RGI}}^{\text{spin}}$, eq. (7.9), is not known. Furthermore, an explicit value for this matrix element would also depend on θ . Up to the two largest masses all continuum extrapolations here are done with the three largest lattice resolutions and only two for $z = 18, 21$. This is consistent with the aforementioned cut $aM \approx 0.7$. The corresponding systematic is clearly visible in the plot and taken into account by increasing the error of the continuum extrapolation at $z = 18, 21$ by the difference of its mean to the data point discarded. By this method the free quadratic fit is less affected by these two points which are still consistent with the fit. Nonetheless, all other points are very well described by the quadratic fit and also consistent with a linear fit. Actually the curvature in the quadratic fit is almost vanishing within the error for all θ .

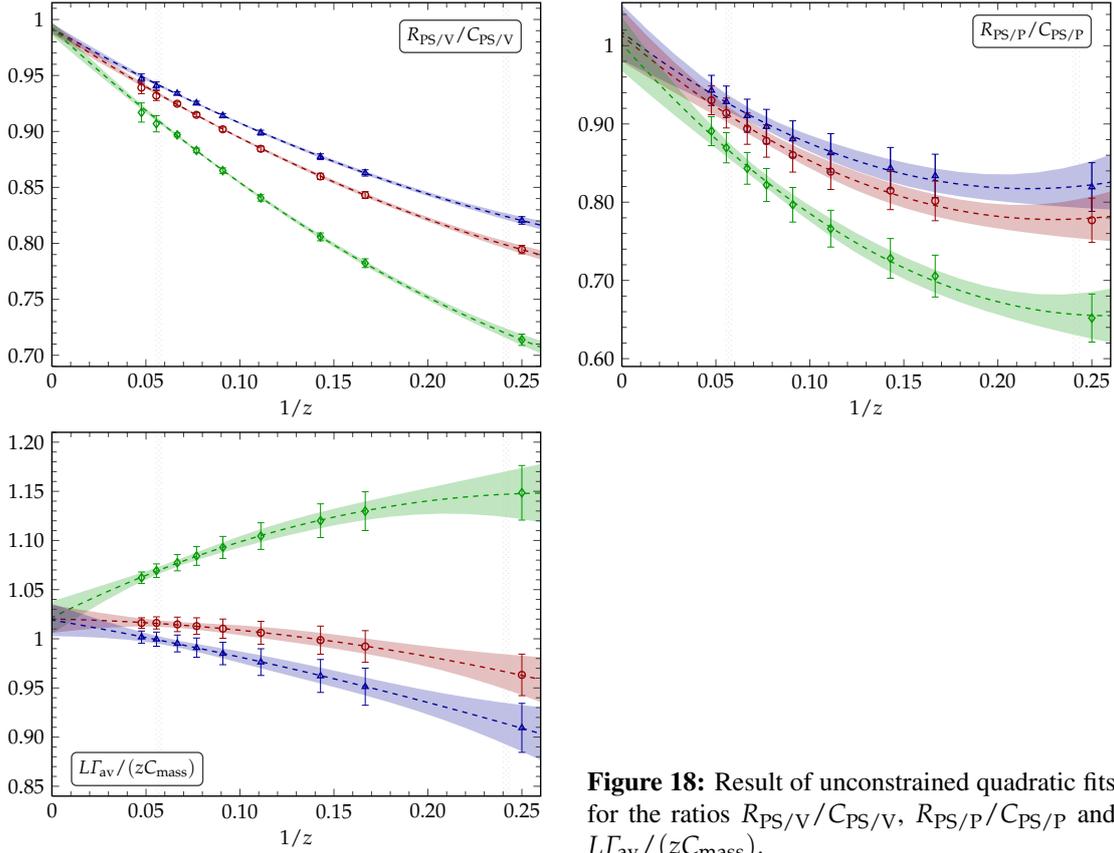


Figure 18: Result of unconstrained quadratic fits for the ratios $R_{PS/V}/C_{PS/V}$, $R_{PS/P}/C_{PS/P}$ and $L\Gamma_{av}/(zC_{mass})$.

In the bottom plot of figure 16 we show the small volume representative of the heavy-light decay constant for both, the pseudo-scalar and the vector channel. The results in the pseudo-scalar channel approach their static limit from above with a slight curvature. Also the unconstrained fits to the vector channel data points approach the same static limits, which as expected are different for each θ . Furthermore, the curvature vanishes for five of six fits as can be inferred from table 20. The consistency here is remarkable and shows that we understand the asymptotics which also legitimates the application of the effective theory approach. In principle, their common limit $1/z \rightarrow 0$ for fixed θ can be constrained by the static theory as well, because the renormalization factor of the associated static axial current matrix element is already known non-perturbatively [155]. The computation of the bare matrix element to obtain the static result for X_{RGI} , eq. (7.8), is in progress. For the ratios $R_{PS/V}/C_{PS/V}$ and $R_{PS/P}/C_{PS/P}$ as shown in figure 17, the constrained quadratic fits do interpolate our data very well. Especially, the agreement for the quantity in the pseudo-scalar channel with its quite large error is remarkable. We should bear in mind that only the matching coefficients c_1 and c_2 have an influence here. This suggests that the remaining perturbative uncertainty of the conversion functions originating from unknown higher orders is much smaller than the precision of our lattice data. Thus a study of the $1/z^n$ -corrections becomes feasible.

The additional dotted lines in the plots represent the constrained linear fits where only a_1 was estimated. For comparison the data points of the five largest masses are incorporated here. In general we can conclude that corrections to the observed linear ($1/z$)-asymptotics become significant

about $1/z \approx 0.1$, i.e. $z \approx 10$, or $M \approx 3.87 \text{ GeV}$ if we reverse the discussion about $M_{b,c}$ from above. An exception is given by the small volume vector decay constant where a deviation from the linear $1/z$ behaviour is not visible at all.

The goodness of the non-perturbative data obtained becomes even more impressive, if we do not constrain the fits to their particular static limit. For the free fit case the results are shown in figure 18. In the case of $R_{PS/V}/C_{PS/V}$ (top left plot), the static limit as predicted by HQET is slightly below 1% for all values of θ . In fact, all three absolute estimates only differ below one per mil to each other. Without the small systematics which may be visible at $z = 18, 21$, one would expect that these values are even closer to one in the static limit. Somewhat different is the behaviour of $R_{PS/V}/C_{PS/V}$. While for $\theta = 1$ the static limit is practically one, the other values are one to two per cent away. However, within the given error they agree with the static limit. The asymptotics of the averaged mass (bottom left plot) is quite similar to that. While all three estimates of the static limit agree by one per mil, they miss the predicted static limit by only 2%. In view of figure (15) one might think to choose another continuum extrapolation at $z = 15, 18, 21$. The corresponding trend introduced to the $(1/z)$ -asymptotics would result in a static limit that is even closer to one. However, we should also not forget that z is afflicted with a relative error of about 1.33% as computed in section 6.4. All unconstrained quadratic fits represent our non-perturbative data very accurately. Actually only the averaged mass is much better represented by the unconstrained fit.

Thus we can conclude that the heavy quark mass dependence in our observables towards the static limit as predicted by HQET is well understood and a controlled error estimate can be given.

7.5 An outlook

At the end of section 5.2, we already reported about ongoing computations on the HQET side. Here we want to mention some preliminary results. Some of the observables used in the non-perturbative determination of the bottom quark mass to subleading order [12] are available now. Even at low statistic they are comparable to the old values which means that the effect of dynamical fermions in some of these observables is maybe not visible. Furthermore some of the step scaling functions involved in this strategy like that of the kinetic term in HQET also show dependence on the lattice spacing which is also comparable to the quenched case. We plan to improve such observables by means of the perturbative improvement discussed in section 7.3 Analogously to our case of QCD observables, the improvement obtained is consistent to the original continuum limit but shows a much better convergence rate.

observable	θ	a_0	a_1	a_2	#points
$L\Gamma_{\text{av}}/(z C_{\text{mass}})$	0	1	-0.11(14)	-0.99(20)	raw/all
		1	+0.04(14)	-1.80(18)	tree/all
		+1.0191(8)	-0.34(15)	-0.42(19)	tree/all
		1	-0.07(13)	—	tree/5
	0.5	1	+0.21(12)	-1.43(18)	raw/all
		1	+0.36(12)	-2.21(17)	tree/all
		+1.0201(8)	-0.037(13)	-0.77(17)	tree/all
		1	+0.22(11)	—	tree/5
	1	1	+1.23(13)	-2.57(17)	raw/all
		1	+1.38(13)	-3.40(16)	tree/all
		+1.0217(9)	+0.95(14)	-1.77(16)	tree/all
		1	+1.17(12)	—	tree/5
$zR_{\text{spin}}/C_{\text{spin}}$	0	+0.1248(78)	-0.257(70)	+0.23(19)	all
	0.5	+0.1099(70)	-0.216(64)	+0.15(17)	all
	1	+0.0833(56)	-0.165(52)	+0.09(14)	all
$Y_{\text{PS}}/C_{\text{PS}}$	0	-0.5010(17)	+0.199(23)	-0.572(39)	raw/all
		-0.4987(15)	+0.166(43)	-0.47(11)	tree/all
	0.5	-0.4816(15)	+0.263(20)	-0.580(33)	raw/all
		-0.4803(11)	+0.242(29)	-0.512(63)	tree/all
	1	-0.4445(16)	+0.449(17)	-0.652(27)	raw/all
		-0.4445(12)	+0.446(18)	-0.636(30)	tree/all
Y_V/C_V	0	-0.5039(18)	-0.450(08)	-0.043(13)	raw/all
		-0.5025(10)	-0.469(28)	+0.017(73)	tree/all
	0.5	-0.4841(17)	-0.424(08)	-0.017(15)	raw/all
		-0.4837(8)	-0.430(23)	+0.00(6)	tree/all
	1	-0.4478(17)	-0.361(10)	+0.015(20)	raw/all
		-0.4484(8)	-0.356(16)	-0.00(4)	tree/all
$R_{\text{PS}/V}/C_{\text{PS}/V}$	0	1	-1.802(24)	+1.487(49)	all
		+0.9922(7)	-0.955(33)	+1.069(74)	all
		1	-0.974(21)	—	tree/5
		0.5	1	-1.240(24)	+1.717(45)
	0.5	1	-1.232(26)	+1.676(51)	tree/all
		+0.9917(6)	-1.098(34)	+1.23(7)	tree/all
		1	-1.110(22)	—	tree/5
	1	1	-1.720(21)	+2.381(44)	raw/all
		1	-1.686(30)	+2.203(54)	tree/all
		+0.9920(7)	-1.554(38)	+1.77(7)	tree/all
		1	-1.521(26)	—	tree/5
	$R_{\text{PS}/P}/C_{A/P}$	0	1	-1.56(12)	+3.30(26)
+1.0172(17)			-1.85(9)	+4.28(17)	all
1			-1.31(11)	—	tree/5
0.5		1	-1.83(10)	+3.73(23)	raw/all
		1	-1.82(10)	+3.71(23)	tree/all
		+1.0131(16)	-2.05(8)	+4.45(14)	tree/all
		1	-1.55(9)	—	tree/5
1		1	-2.66(11)	+5.13(22)	raw/all
		1	-2.65(11)	+5.08(22)	tree/all
		+1.0021(16)	-2.68(9)	+5.19(14)	tree/all
		1	-2.30(10)	—	tree/5

Table 20: Fit results for the $(1/z)$ -asymptotics in our QCD test observables. Fit parameter a_i without error estimate have been fixed. Results are given for fits to the raw data and to tree-level improved observables, if available. While the error bands shown in the plots are derived from the covariance matrix of the corresponding fit, the table shows only the errors associated to the diagonal elements.

8 Summary

The subject of the first part of this work was to set up a non-perturbatively on-shell $\mathcal{O}(a)$ improved lattice theory. This was achieved within the Schrödinger functional renormalization scheme by computing the relevant improvement coefficients and renormalization constants which became important in the matching volume L_1 . There we prepared lattice simulations which can be used along the lines of a general *non-perturbative matching of HQET to QCD in a finite volume* [11], and which have been used to perform non-perturbative tests of Heavy Quark Effective Theory using relativistic heavy-light observables. To this end the dimensionless renormalization group invariant quark mass $z = L_1 M$ has been fixed to various values in the continuum. Imposing this condition in the continuum translates to different hopping parameters for the fixed mass of the heavy quark flavour at different lattice resolutions. They could be estimated with the previously determined improvement coefficients and renormalization constant. We have chosen nine different continuum values for the heavy quark mass z to enclose the charm and bottom quark mass region. The improvement coefficients and renormalization constant have shown a smooth dependence on the bare coupling g_0^2 after a *constant physics condition* was imposed. Furthermore, a universality test confirmed that the ambiguity introduced by fixing a specific improvement condition vanishes in the continuum limit. We found the theoretically expected cutoff dependence in these quantities. Thus the lattice spacing effects observed are under good control. For further purposes we also provided a smooth interpolating formula for b_m , $b_A - b_P$ and Z which can be used in charm physics application by means of lattice computations.

The second part was devoted to the computation of relativistic heavy-light observables in the continuum of the matching volume to non-perturbatively test predictions made by HQET. These observables were built from heavy-light currents of the pseudo-scalar and vector channel. The light quark mass was set to the sea quark mass and the heavy quark mass to the previously fixed values of z . To relate the QCD observables at finite mass z (and thus finite M) to the renormalization group invariants of the effective theory, we had to suitably parametrize and evaluate the conversion functions C_χ perturbatively. It was done using known results for the appropriate anomalous dimensions. This could be achieved to high accuracy, i.e. using the three-loop anomalous dimension of the corresponding HQET current and the associated Wilson coefficient to two-loop order. The fit results applied to our data are remarkable compatible with the expected $1/z^n$ power corrections, showing us that they are dominating over the perturbative ones. Thus the effective theory especially at the scale of the b-quark is very useful. Actually, the data is very consistent over the whole range of masses even if the difference in the conversion functions with two-loop and three-loop as highest order gets larger. The heavy quark mass dependence in our observables for all three values of the Schrödinger functional parameter θ is well-behaved and a smooth interpolations seems possible if needed. In general we observed larger cutoff effects than in the quenched case which is not even comparable due to different physical parameters. For this reason we also applied tree-level improvement to all quantities under consideration to remove all cutoff-effects appearing at tree-level on top of our non-perturbatively improved data set. Only for the spin splitting term there is no contribution at tree-level. The outcome confirmed in almost all cases the continuum limit obtained earlier which is also under good control up to $z \approx 20$. The

leading $(1/z)$ -corrections encountered are at the expected order for each quantity and also reliable estimates for $(1/z^2)$ -corrections could be given. We can conclude that the heavy quark mass dependence is well well compatible with the asymptotic behaviour predicted by HQET. In all parts of this work we have verified the expected theoretical behaviour and we have a really good control of systematic errors. Thus this goal has been accomplished. What remains is to combine predictions from HQET in the static limit with our data for $X_{\text{RGI}}^{\text{spin}}(L)$ and X_{RGI} for all values of θ .

Beyond the general interest of providing non-perturbative test of HQET, our results are important in view of the non-perturbative matching programme [11]. We outlined this strategy as it is currently applied to the case of two dynamical flavour of quarks. Especially, the observable $\Phi_3^{\text{QCD}}(L_1, M) = L_1 \Gamma_{\text{av}}(L_1, M)$ plays an important role in an accurate estimation of the bottom quark mass to subleading order in HQET [12]. This work is part of a much larger collaborative effort which is technically advanced and computer-intensive. Also simulations in physically large volume are involved [142] and under progress.

A The Renormalization group

A bare vertex function $\Gamma_0^{(n)}$ depends on momenta p_i and the bare parameters of the theory. In this section I skip the momenta and just use one coupling and mass parameter. A generalisation to more than those bare parameters is easy by introducing appropriate indices. Then the connection between the bare and the renormalized vertex function is given by

$$\Gamma_0^{(n)}(g_0, m_0) = Z^{-n/2} \Gamma^{(n)}(\bar{g}, \bar{m}, \mu), \quad (\text{A.1})$$

with a field renormalization constant Z and renormalized parameters \bar{g}, \bar{m} which depend on the subtraction point μ , introduced by the renormalization prescription. By definition bare quantities are independent of μ what means that a change of the subtraction point in eq. (A.1) which changes the value of $\Gamma^{(n)}$ has to be compensated by a change in Z . This statement can be explored by applying the dimensionless total derivative operator $\mu d/d\mu$,

$$\begin{aligned} 0 &\equiv \mu \frac{d}{d\mu} \left(Z^{-n/2} \Gamma^{(n)}(\bar{g}, \bar{m}, \mu) \right) \\ &= \frac{\mu}{Z^{-n/2}} \frac{d}{d\mu} \left(Z^{-n/2} \right) \Gamma^{(n)}(\bar{g}, \bar{m}, \mu) + \mu \frac{d}{d\mu} \left(\Gamma^{(n)}(\bar{g}, \bar{m}, \mu) \right) \\ &= \left[\frac{\mu}{Z^{-n/2}} \frac{d}{d\mu} \left(Z^{-n/2} \right) + \mu \frac{d}{d\mu} \right] \Gamma^{(n)}(\bar{g}, \bar{m}, \mu) \\ &= \left[-\frac{n}{2} \frac{d \ln Z}{d \ln \mu} + \mu \left(\frac{\partial}{\partial \mu} + \frac{\partial \bar{g}}{\partial \mu} \frac{\partial}{\partial \bar{g}} + \frac{\partial \bar{m}}{\partial \mu} \frac{\partial}{\partial \bar{m}} \right) \right] \Gamma^{(n)}(\bar{g}, \bar{m}, \mu). \end{aligned} \quad (\text{A.2})$$

In the last line the operator is just expanded due to the chain rule with respect to the independent variables μ, \bar{g} and \bar{m} . This statement can obviously be made for an arbitrary physical quantity Q that depends on the renormalized parameters and the subtraction point of the theory,

$$0 \equiv \mu \frac{d}{d\mu} Q = \left[\mu \frac{\partial}{\partial \mu} + \beta(\bar{g}) \frac{\partial}{\partial \bar{g}} + \tau(\bar{g}) \bar{m} \frac{\partial}{\partial \bar{m}} + \gamma_Q(\bar{g}) \right] Q \quad (\text{A.3})$$

where the following definitions apply,

$$\beta(\bar{g}) = \frac{\partial \bar{g}}{\partial \ln \mu}, \quad \tau(\bar{g}) = \frac{\partial \ln \bar{m}}{\partial \ln \mu}, \quad \gamma_Q(\bar{g}) = -\frac{n}{2} \frac{\partial \ln Z}{\partial \ln \mu}. \quad (\text{A.4})$$

Note, in order to determine the running of Q , eq. (A.3), the first task is to estimate the running of the parameters of the theory, namely the couplings and masses, encoded by the (Gell-Mann–Low) β -function and the mass anomalous dimension τ .

Renormalization Group Invariants

When solving the (first order) RGE beyond the formal level, one has to introduce one constant of integration. The RG is a non-perturbative concept and for our lack of (non-perturbative) analytical tools we cannot entirely solve it. At least, we can apply perturbation theory near a Gaussian fixed point – zero(s) of the β -function – and hope that all (at some scale relevant) physical effects are

taken into account by just increasing the order of the perturbative expansion. For asymptotically free theories such a fixed point exists at zero length scale, or energy scale $\mu_0 \rightarrow \infty$. In this limit, where the asymptotic of the renormalized coupling $\bar{g}(\mu_0) \rightarrow 0$ is known, all physical quantities become (trivially) scale independent. Hence, it is natural to associate the integration constant to this fixed point. In turn, after fixing the integration constant which we then call Renormalization Group Invariant (RGI), we would exactly know the scaling of that observable if we possess their non-perturbative anomalous dimension.

Let us go through these steps in case of the QCD coupling. With respect to some reference scale μ_0 , the general solution of the Callan-Symanzik equation for the coupling (A.4) reads

$$\ln \frac{\mu_0}{\mu} = \int_{\bar{g}(\mu)}^{\bar{g}(\mu_0)} \frac{d\bar{g}}{\beta(\bar{g})}, \quad \text{or} \quad \frac{\mu_0}{\mu} = \exp \left\{ - \int_{\bar{g}(\mu_0)}^{\bar{g}(\mu)} \frac{d\bar{g}}{\beta(\bar{g})} \right\}, \quad (\text{A.5})$$

with a general scale $\mu < \mu_0$. Now the properties of the underlying theory come into play by

$$\beta(\bar{g}) \stackrel{\bar{g} \rightarrow 0}{\sim} -\bar{g}^3 (b_0 + b_1 \bar{g}^2 + \dots). \quad (\text{A.6})$$

The universal expansion coefficients b_0 , b_1 and higher order ones are computed by evaluating appropriate Feynman diagrams. The next step would be to connect the general solution (A.5) to the Gaussian fixed point, but it diverges in the limit $\mu_0 \rightarrow \infty$ due to (A.6). This divergence just emerges as a consequence how we have written down the equation as already the LHS trivially diverges. To solve this problem one formally integrates out the asymptotically divergent part at finite μ_0 . The leading order (LO) divergence is

$$- \int \frac{1}{\beta(\bar{g})} \Big|_{\text{LO}} d\bar{g} = \frac{1}{b_0} \int \frac{d\bar{g}}{\bar{g}^3} = -\frac{1}{2b_0 \bar{g}^2}. \quad (\text{A.7})$$

After subtracting this LO in the integral kernel from the next-to-leading order (NLO) insertion of the perturbative β -function, a logarithmic divergent part for $\bar{g} \rightarrow 0$ remains,

$$- \int \left[\frac{1}{\beta(\bar{g})} \Big|_{\text{NLO}} + \frac{1}{b_0 \bar{g}^3} \right] d\bar{g} = -\frac{b_1}{2b_0^2} \left[\ln \bar{g}^2 - \ln b_0 - \ln \left(1 + \frac{b_1}{b_0} \bar{g}^2 \right) \right]. \quad (\text{A.8})$$

Note that the constant term proportional to $\ln b_0$ on the RHS, cancels after making the integral definite. There are no further divergent parts coming in higher order PT. Separating only the encountered asymptotically divergent parts from the integral kernel in eq. (A.5) gives the identity

$$\frac{\mu_0}{\mu} = \exp \left\{ - \int_{\bar{g}(\mu_0)}^{\bar{g}(\mu)} \left[\frac{1}{\beta(\bar{g})} + \frac{1}{b_0 \bar{g}^3} - \frac{b_1}{b_0^2 \bar{g}} \right] d\bar{g} - \frac{1}{2b_0 \bar{g}^2} \Big|_{\bar{g}(\mu_0)}^{\bar{g}(\mu)} + \frac{b_1}{2b_0^2} \ln \bar{g}^2 \Big|_{\bar{g}(\mu_0)}^{\bar{g}(\mu)} \right\} \quad (\text{A.9})$$

$$= \left(\frac{k \bar{g}^2(\mu)}{k \bar{g}^2(\mu_0)} \right)^{-\frac{b_1}{2b_0^2}} \times \exp \left\{ \frac{1}{2b_0} \left[\frac{1}{\bar{g}^2(\mu_0)} - \frac{1}{\bar{g}^2(\mu)} \right] \right\} \\ \times \exp \left\{ - \int_{\bar{g}(\mu_0)}^{\bar{g}(\mu)} \left[\frac{1}{\beta(\bar{g})} + \frac{1}{b_0 \bar{g}^3} - \frac{b_1}{b_0^2 \bar{g}} \right] d\bar{g} \right\}. \quad (\text{A.10})$$

Note that the constant factor $k > 0$ introduced in the first term of eq. (A.10) has no effect here. Rearranging such that the divergent parts in μ_0 appear on the LHS and all μ -dependent quantities on the RHS gives

$$\begin{aligned} \mu_0 [k \bar{g}^2(\mu_0)]^{-\frac{b_1}{2b_0^2}} e^{-\frac{1}{2b_0 \bar{g}^2(\mu_0)}} & \quad (A.11) \\ = \mu [k \bar{g}^2(\mu)]^{-\frac{b_1}{2b_0^2}} e^{-\frac{1}{2b_0 \bar{g}^2(\mu)}} \exp \left\{ - \int_{\bar{g}(\mu_0)}^{\bar{g}(\mu)} \left[\frac{1}{\beta(\bar{g})} + \frac{1}{b_0 \bar{g}^3} - \frac{b_1}{b_0^2 \bar{g}} \right] d\bar{g} \right\}. \end{aligned}$$

Taking the limit $\mu_0 \rightarrow \infty$ with $\bar{g} \rightarrow 0$, the RHS converges and so does the LHS. By the convenient choice $k = b_0$ the constant contribution in equation (A.8), which is proportional to $\ln b_0$, is canceled. With this we choice we obtain the RGI of the QCD coupling, the Λ -parameter of QCD:

$$\Lambda_{\text{QCD}} \equiv \lim_{\mu_0 \rightarrow \infty} \mu_0 [b_0 \bar{g}^2(\mu_0)]^{-\frac{b_1}{2b_0^2}} e^{-\frac{1}{2b_0 \bar{g}^2(\mu_0)}}, \quad (A.12a)$$

$$\Lambda_{\text{QCD}} = \mu [b_0 \bar{g}^2(\mu)]^{-\frac{b_1}{2b_0^2}} e^{-\frac{1}{2b_0 \bar{g}^2(\mu)}} \exp \left\{ - \int_0^{\bar{g}(\mu)} \left[\frac{1}{\beta(\bar{g})} + \frac{1}{b_0 \bar{g}^3} - \frac{b_1}{b_0^2 \bar{g}} \right] d\bar{g} \right\}. \quad (A.12b)$$

Other renormalization group invariants like that of the mass M or some local multiplicative composite operator Φ_{RGI} are obtained in the same way. They read

$$M \equiv \lim_{\mu_0 \rightarrow \infty} \bar{m}(\mu_0) [2b_0 \bar{g}^2(\mu_0)]^{-\frac{d_0}{2b_0}}, \quad (A.13a)$$

$$M = \bar{m}(\mu) [2b_0 \bar{g}^2(\mu)]^{-\frac{d_0}{2b_0}} \exp \left\{ - \int_0^{\bar{g}(\mu)} d\bar{g} \left[\frac{\tau(\bar{g})}{\beta(\bar{g})} - \frac{d_0}{b_0 \bar{g}} \right] \right\}, \quad (A.13b)$$

and

$$\Phi_{\text{RGI}} \equiv \lim_{\mu_0 \rightarrow \infty} \Phi_{\text{RGI}}(\mu_0) [2b_0 \bar{g}^2(\mu_0)]^{-\frac{d_0}{2b_0}}, \quad (A.14a)$$

$$\Phi_{\text{RGI}} = \Phi_{\text{RGI}}(\mu) [2b_0 \bar{g}^2(\mu)]^{-\frac{\gamma_0}{2b_0}} \exp \left\{ - \int_0^{\bar{g}(\mu)} d\bar{g} \left[\frac{\gamma(\bar{g})}{\beta(\bar{g})} - \frac{\gamma_0}{b_0 \bar{g}} \right] \right\}. \quad (A.14b)$$

Remarks

Choosing a scheme means to declare a definition for the coupling $\bar{g}(\mu)$ accompanied by a renormalization condition for the coupling. Thus it should be clear that Λ_{QCD} is inherently scheme dependent. Suppose we have two different schemes, characterized by their Λ -parameters Λ_a and Λ_b . With the well-defined expressions (A.12a) for each of them, we can write at arbitrary, finite $\mu_0^a = \mu_0^b = \mu_0$,³⁰

$$2b_0 \ln \left[\frac{\Lambda_b}{\Lambda_a} \right] = -\frac{b_1}{b_0} \ln \left[\frac{\bar{g}_b^2}{\bar{g}_a^2} \right] + \left(\frac{1}{\bar{g}_a^2} - \frac{1}{\bar{g}_b^2} \right). \quad (A.15)$$

³⁰ This discussion remains true in case of $\mu_0^a \neq \mu_0^b$. Then logarithms $\ln(\mu_0^a/\mu_0^b)$ appear and the limit $\mu_0 \rightarrow \infty$ has to be taken more carefully.

At any fixed scale μ_0 the couplings can be related by $\bar{g}_b^2 = \bar{g}_a^2 (1 + c_1^{(ba)} \bar{g}_a^2 + \mathcal{O}(\bar{g}_a^4))$ which gives

$$2b_0 \ln \left[\frac{\Lambda_b}{\Lambda_a} \right] = -\frac{b_1}{b_0} \ln \left[1 + c_1^{(ba)} \bar{g}_a^2 + \mathcal{O}(\bar{g}_a^4) \right] + c_1^{(ba)} \left(\frac{1 + \mathcal{O}(\bar{g}_a^2)}{1 + c_1^{(ba)} \bar{g}_a^2 + \mathcal{O}(\bar{g}_a^4)} \right). \quad (\text{A.16})$$

Obviously, the only term surviving the limit $\mu_0 \rightarrow \infty$ is $c_1^{(ba)}$, emerging from the leading order asymptotics in the definition of Λ . Therefore, the same result is obtained by using the LO definition of Λ in eq. (A.12a). Also even higher order terms will have no effect on the result,

$$\Lambda_b / \Lambda_a = \exp \left\{ c_1^{(ba)} / 2b_0 \right\}, \quad (\text{A.17})$$

which *exactly* relates two different schemes. It reflects the fact that all possible (massless) renormalization schemes form a one parameter family. On the other hand, ratios like M_b/M_a or $\Phi_{\text{RGI}}^b/\Phi_{\text{RGI}}^a$ between two different schemes are exactly one and hence *scheme independent*!

B Definitions and conventions

Following the usual conventions I use $\mu, \nu, \dots \in \{0, \dots, 3\}$ and $i, j, \dots \in \{1, 2, 3\}$ to denote space-time and space indices respectively. The same symbol with a hat on it stands for the unit vector in the corresponding direction within that scope. For the internal symmetry group I take the first lowercase letters from the Latin alphabet, a, b, \dots , independent of the group it belongs to. Upper case letters like A, B, \dots are reserved for the components of Dirac matrices. The symbol $\mathbb{1}$ stands for the unit element or unit matrix in the corresponding context and labeling the dimension can be omitted. For repeated indices the summation convention apply. Any deviations from this conventions will be stated explicitly. Furthermore a serves as the lattice spacing.

B.1 The group $\text{SU}(N)$

The *special unitary group* $\text{SU}(N)$ is a non-Abelian Lie-group of $N^2 - 1$ real dimensions. It is *compact* and *simply connected* such that a group volume can be defined. $\text{SU}(N)$ has a natural realisation as $N \times N$ matrix group. Thus, an element U in the *fundamental representation* reads

$$U = \exp(i\alpha^a T_a), \quad \alpha^a \in \mathbb{R}, a = 1, \dots, N^2 - 1, \quad (\text{B.1})$$

with traceless *generators* $\{T_a\}$. Here, they are Hermitean $T^\dagger = T$, but can also be chosen to be anti-Hermitean, $T \mapsto -iT$ in eq. (B.1). Thus all elements of $\text{SU}(N)$ are parametrized in terms of the generators of the corresponding Lie algebra, $\mathfrak{su}(N)$, defined by

$$[T_a, T_b] = if_{abc} T_c, \quad (\text{B.2})$$

$$[T_a, [T_b, T_c]] + [T_b, [T_c, T_a]] + [T_c, [T_a, T_b]] = 0. \quad (\text{B.3})$$

The (real) *structure constants* f_{abc} define the group multiplication table as the commutator serves as multiplication symbol. The *Jacobi identity*, equation (B.3) holds for every Lie-group. The

generators can always be chosen such that f_{abc} is total antisymmetric under permutation of the indices and the following normalization holds,

$$\{T_a, T_b\} = \frac{1}{N}\delta_{ab}\mathbb{1} + d_{abc}T_c \quad \Leftrightarrow \quad \text{Tr}(T_a T_b) = \frac{1}{2}\delta_{ab}, \quad (\text{B.4})$$

$$f_{abc} = -2i \text{Tr}([T_a, T_b] T_c), \quad d_{abc} = 2 \text{Tr}(\{T_a, T_b\} T_c). \quad (\text{B.5})$$

with total symmetric constants d_{abc} . In the *adjoint representation* the generators $(T_a)_{bc} = -if_{abc}$ are $(N^2 - 1) \times (N^2 - 1)$ matrices defined by the structure constants.

SU(2)

The three generators of $SU(2)$ are given in terms of the Pauli matrices,

$$\tau^1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \tau^2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \tau^3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \quad (\text{B.6})$$

originally introduced by W. Pauli to describe the spin of a particle, [134]. With the notation of the preceding section one has

$$T_a = \tau^a / 2, \quad f_{abc} = \varepsilon_{abc}, \quad d_{abc} = 0, \quad a, b, c \in \{1, 2, 3\}, \quad (\text{B.7})$$

where the total antisymmetric symbol in three dimensions with $\varepsilon_{123} = 1$ appears as structure constant and the symmetric constants vanishes. One furthermore defines

$$\tau^\pm \equiv \tau^1 \pm i\tau^2 \quad \Leftrightarrow \quad \tau^+ = \begin{pmatrix} 0 & 2 \\ 0 & 0 \end{pmatrix}, \quad \tau^- = \begin{pmatrix} 0 & 0 \\ 2 & 0 \end{pmatrix}, \quad (\text{B.8})$$

which are also known as ladder operators from the non-relativistic description of the spin.

SU(3)

A conventional set of eight generators $\{T^a = \lambda^a / 2 \mid a = 1, \dots, 8\}$ for the *colour gauge group* $SU(3)$ is given by the so-called *Gell-Mann matrices* [133],³¹

$$\begin{aligned} \lambda_b &= \begin{pmatrix} \tau^b & 0 \\ 0 & 0 \end{pmatrix}, & \lambda_6 &= \begin{pmatrix} 0 & 0 \\ 0 & \tau^1 \end{pmatrix}, & \lambda_7 &= \begin{pmatrix} 0 & 0 \\ 0 & -\tau^2 \end{pmatrix}, \\ \lambda_4 &= \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}, & \lambda_5 &= \begin{pmatrix} 0 & 0 & -i \\ 0 & 0 & 0 \\ i & 0 & 0 \end{pmatrix}, & \lambda_8 &= \frac{1}{\sqrt{3}} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -2 \end{pmatrix}, \end{aligned} \quad (\text{B.9})$$

³¹Originally introduced as $SU(3)$ -isospin symmetry group that enlarges the formerly known $SU(2)$ -isospin group by a new additive quantum number, called strangeness. Because of large mass differences ($m_K > 3m_\pi$) within the same multiplet, this symmetry is poor realised but leads to the development of the quark picture with exact colour symmetry. Today, one rather groups the three lightest (elementary) quarks u , d and s into a $SU(3)$ -flavour symmetry.

with $b = 1, 2, 3$. The τ 's are those of eq. (B.6). The only non-vanishing *structure constants* are³²

$$\begin{aligned} f_{123} &= 1, & f_{458} &= f_{678} = \sqrt{3}/2, \\ f_{147} &= -f_{156} = f_{246} = f_{345} = f_{257} = -f_{367} = 1/2. \end{aligned} \quad (\text{B.10})$$

and the symmetric constants follow to

$$\begin{aligned} -d_{118} &= -d_{228} = -d_{338} = d_{888} = 2d_{448} = 2d_{558} = 2d_{668} = 2d_{778} = -1/\sqrt{3}, \\ d_{146} &= d_{157} = -d_{247} = d_{256} = d_{344} = d_{355} = -d_{366} = -d_{377} = 1/2. \end{aligned} \quad (\text{B.11})$$

B.2 The Dirac algebra

For a Euclidean description of a four-dimensional relativistic QFT with fermions one also has to change from a Clifford algebra with respect to the metric $(g_{\mu\nu}) = \text{diag}(1, -1, -1, -1)$ to the Euclidean one, $(\delta_{\mu\nu})$. This is achieved by the substitution $(\gamma_0, \gamma)^M \rightarrow (\gamma_0, \gamma) = (\gamma_0, -i\gamma)^M$ where M is short for Minkowski space. In terms of the Euclidean metric the Dirac matrices are now defined by

$$\{\gamma_\mu, \gamma_\nu\} = 2\delta_{\mu\nu}, \quad \gamma_5 = \gamma_0\gamma_1\gamma_2\gamma_3, \quad (\text{B.12})$$

and obey the following relations

$$\gamma_\mu^\dagger = \gamma_0\gamma_\mu\gamma_0, \quad (\gamma_\mu)^2 = \mathbb{1}, \quad (\text{B.13})$$

$$\gamma_5^\dagger = \gamma_5, \quad (\gamma_5)^2 = \mathbb{1}, \quad \{\gamma_5, \gamma_\nu\} = 0. \quad (\text{B.14})$$

To simplify matters one usually prefers to work in a *hermitean representation* with $\gamma_\mu^\dagger \equiv \gamma_\mu$. This is the case in the present work by using the *chiral representation*, given by

$$\gamma_\mu = \begin{pmatrix} 0 & e_\mu \\ e_\mu^\dagger & 0 \end{pmatrix}, \quad (e_\mu) = (e_0, \mathbf{e}) = (-\mathbb{1}, -i\boldsymbol{\tau}), \quad \gamma_5 = \begin{pmatrix} \mathbb{1} & 0 \\ 0 & -\mathbb{1} \end{pmatrix}, \quad (\text{B.15})$$

with a tuple $\boldsymbol{\tau} = (\tau^1, \tau^2, \tau^3)$ composed of the standard Pauli matrices (B.6). Furthermore one needs the total antisymmetric tensor,

$$\sigma_{\mu\nu} = \frac{i}{2} [\gamma_\mu, \gamma_\nu], \quad (\text{B.16})$$

which in this representation is explicitly given by

$$\sigma_{0k} = \begin{pmatrix} \tau^k & 0 \\ 0 & -\tau^k \end{pmatrix}, \quad \sigma_{ij} = -\epsilon_{ijk} \begin{pmatrix} \tau^k & 0 \\ 0 & \tau^k \end{pmatrix}, \quad (\text{B.17})$$

³²Except those obtained by applying the symmetry properties of the indices.

where ϵ_{ijk} is the totally antisymmetric symbol with $\epsilon_{123} = 1$. The hermiticity property,

$$\sigma_{\mu\nu}^\dagger = \gamma_0 \sigma_{\mu\nu} \gamma_0 = \sigma_{\mu\nu} , \quad (\text{B.18})$$

follows directly from that of the γ_μ . The (linearly independent) ordered set of matrices $\{\mathbb{1}, \gamma_\mu, \sigma_{\mu\nu}, \gamma_\mu \gamma_5, \gamma_5\}$ with $16 = (1 + 4 + 6 + 4 + 1)$ independent components can serve as a basis $\{\Gamma\}$ of the $2^{(d=4)}$ dimensional (closed) Dirac-Clifford algebra. It is an orthonormal basis with respect to the scalar product

$$(\Gamma_i, \Gamma_j) \equiv \frac{1}{4} \text{Tr} \{ \Gamma_i^\dagger \Gamma_j \} = \delta_{ij} , \quad i, j = 1, \dots, 16 , \quad (\text{B.19})$$

where all matrices except the unit element are traceless, $\text{Tr} \mathbb{1} = 4$, $\text{Tr} \Gamma_i = 0$. As a consequence of the closeness of that algebra an arbitrary matrix can be decomposed by

$$M = \sum_{i=1}^{16} C_i^M \Gamma_i , \quad C_i^M = \frac{1}{4} \text{Tr}(\Gamma_i^\dagger M) , \quad (\text{B.20})$$

and any product of such matrices can again be expressed by the basis matrices.

B.3 Definitions on the lattice

B.3.1 Lattice derivatives

In this section we work at fixed μ , i.e. no summation over repeated indices is applied. The standard lattice *forward* and *backward derivatives*, ∂_μ and ∂_μ^* , act as finite-difference operators on functions $\phi(x) = \phi_x$ as

$$\partial_\mu \phi(x) = \frac{1}{a} [\phi(x + a\hat{\mu}) - \phi(x)] , \quad (\text{B.21a})$$

$$\partial_\mu^* \phi(x) = \frac{1}{a} [\phi(x) - \phi(x - a\hat{\mu})] , \quad (\text{B.21b})$$

which is equal to

$$\partial_\mu = \frac{1}{a} [\delta_{x+a\hat{\mu},x} - \delta_{x,x}] , \quad \partial_\mu^* = \frac{1}{a} [\delta_{x,x} - \delta_{x-a\hat{\mu},x}] . \quad (\text{B.22})$$

The *gauge covariant lattice derivative* operators acting on fermion and anti-fermion fields are

$$\nabla_\mu \psi(x) = \frac{1}{a} [\lambda_\mu U(x, \mu) \psi(x + a\hat{\mu}) - \psi(x)] , \quad (\text{B.23a})$$

$$\nabla_\mu^* \psi(x) = \frac{1}{a} [\psi(x) - \lambda_\mu^{-1} U(x - a\hat{\mu}, \mu)^{-1} \psi(x - a\hat{\mu})] , \quad (\text{B.23b})$$

$$\overleftarrow{\nabla}_\mu \bar{\psi}(x) = \frac{1}{a} [\bar{\psi}(x + a\hat{\mu}) U(x, \mu)^{-1} \lambda_\mu^{-1} - \bar{\psi}(x)] , \quad (\text{B.23c})$$

$$\overleftarrow{\nabla}_\mu^* \bar{\psi}(x) = \frac{1}{a} [\bar{\psi}(x) - \bar{\psi}(x - a\hat{\mu}) U(x - a\hat{\mu}, \mu) \lambda_\mu] , \quad (\text{B.23d})$$

or

$$\nabla_\mu = \frac{1}{a} [\lambda_\mu U_{x,\mu} \delta_{x+a\hat{\mu},x} - \delta_{x,x}] , \quad \nabla_\mu^* = \frac{1}{a} [\delta_{x,x} - \lambda_\mu^{-1} U_{x-a\hat{\mu},x}^{-1} \delta_{x,x-a\hat{\mu}}] , \quad (\text{B.24})$$

$$\overleftarrow{\nabla}_\mu = \frac{1}{a} [\delta_{x,x+a\hat{\mu}} U_{x,\mu}^{-1} \lambda_\mu^{-1} - \delta_{x,x}] , \quad \overleftarrow{\nabla}_\mu^* = \frac{1}{a} [\delta_{x,x} - \delta_{x,x-a\hat{\mu}} U_{x-a\hat{\mu},\mu} \lambda_\mu] , \quad (\text{B.25})$$

respectively. As already explained in section 3.3 the periodicity angles or phase factors

$$\lambda_\mu \equiv e^{ia\theta_\mu/L}, \quad \theta_0 = 0, \quad -\pi < \theta_k \leq \pi, \quad (\text{B.26})$$

appearing in the generalized fermion field boundary conditions of eq. (3.33), are conventionally included in the gauge covariant derivatives. Their standard form with $\lambda_\mu \equiv \mathbb{1}$ is recovered in the infinite volume limit, or for $\theta_k \equiv 0$ of course.³³ *Symmetrized lattice derivatives* are given by their mean according to

$$\begin{aligned} \tilde{\partial}_\mu &= \frac{1}{2}(\partial_\mu + \partial_\mu^*) & \tilde{\nabla}_\mu &= \frac{1}{2}(\nabla_\mu + \nabla_\mu^*) \\ &= \frac{1}{2a}[\delta_{x,x+a\hat{\mu}} - \delta_{x,x-a\hat{\mu}}] & &= \frac{1}{2a}[U_{x\mu}\delta_{x,x+a\hat{\mu}} - U_{x-a\hat{\mu},x}^{-1}\delta_{x,x-a\hat{\mu}}]. \end{aligned} \quad (\text{B.27})$$

Furthermore one needs an approximation for the second derivative, or Laplacian if more than one dimension is involved. According to the first derivatives introduced above, various combinations are possible again. The common choice is to use

$$\partial_\mu^* \partial_\mu \phi(x) \equiv \frac{1}{a^2}[\phi(x+a\hat{\mu}) + \phi(x-a\hat{\mu}) - 2\phi(x)]. \quad (\text{B.28})$$

This follows directly by applying the original definitions to the test function and it can be easily proven that $\partial_\mu^* \partial_\mu = \partial_\mu \partial_\mu^*$ holds. The only other symmetric choice for a second lattice derivative is $\tilde{\partial}_\mu \tilde{\partial}_\mu$. This would introduce next to next-neighbour contributions and hence increase the contaminations by higher excited states what is what one usually wants to avoid.

Improved derivatives

Any finite-difference operator Δ_x applied to a function ϕ at lattice site x_i , approximates the unique continuum derivative to some order in a ,

$$\frac{\partial\phi(x_i)}{\partial x} \approx \Delta_x \phi(x_i). \quad (\text{B.29})$$

If we choose Δ_x to be the forward derivative ∂_μ for instance, we get from their definition (B.21a) and a Taylor expansion in a ,

$$a\Delta_x \phi(x) \equiv \phi(x+a) - \phi(x) = (e^{a\partial_x} - 1)\phi(x) \quad (\text{B.30a})$$

$$= \left(1 + a\partial_x + \frac{1}{2}a^2\partial_x^2 + \mathcal{O}(a^3) - 1\right)\phi(x) \quad (\text{B.30b})$$

$$= a\partial_x \left(1 + \frac{1}{2}a\partial_x + \frac{1}{6}a^2\partial_x^2\right)\phi(x) + \mathcal{O}(a^4), \quad (\text{B.30c})$$

³³ Even if the phase angles θ_i were used since a long time within the Schrödinger functional picture, they received a much broader attention through the lattice community only in recent years, see for example [136] and references therein. They go under the name *twisted boundary conditions*.

which shows that the forward derivative approximates the continuum derivative up to corrections of order a . Needless to say, this is true for the backward derivative $\Delta_x \equiv \partial_\mu^*$ too:

$$a\Delta_x\phi(x) \equiv \phi(x) - \phi(x-a) = \left(1 - e^{-a\partial_x}\right)\phi(x) \quad (\text{B.31a})$$

$$= \left(1 - 1 + a\partial_x - \frac{1}{2}a^2\partial_x^2 + \mathcal{O}(a^3)\right)\phi(x) \quad (\text{B.31b})$$

$$= a\partial_x\left(1 - \frac{1}{2}a\partial_x + \frac{1}{6}a^2\partial_x^2\right)\phi(x) + \mathcal{O}(a^4). \quad (\text{B.31c})$$

It makes clear that the symmetrized derivative $\Delta_x \equiv \tilde{\partial}_\mu$ is the best approximation to the continuum one that includes only nearest neighbour lattice sites,

$$2a\Delta_x\phi(x) \equiv \phi(x+a) - \phi(x-a) = (e^{a\partial_x} - e^{-a\partial_x})\phi(x) \quad (\text{B.32a})$$

$$= a\partial_x\left(1 + \frac{1}{6}a^2\partial_x^2\right)\phi(x) + \mathcal{O}(a^5), \quad (\text{B.32b})$$

and has errors of $\mathcal{O}(a^2)$. Clearly no odd powers of a appear as corrections to the leading approximation. In order to achieve an improvement of the symmetrized lattice derivative one has to add a counterterm to their definition in eq. (B.27) that cancels the second term in (B.32b). This leads to the following representation of a *symmetric improved first order lattice derivative*,

$$\tilde{\partial}_\mu\left(1 - \frac{1}{6}a^2\partial_\mu^*\partial_\mu\right)\phi(x) = \frac{\phi(x-2a\hat{\mu}) - 8\phi(x-a\hat{\mu}) + 8\phi(x+a\hat{\mu}) - \phi(x+2a\hat{\mu})}{12a}, \quad (\text{B.33})$$

and brings us to the expansion of the representation $\Delta_x^{(2)} \equiv \partial_\mu\partial_\mu^*$ of a second order derivative $\Delta_x^{(2)}$, that approximates the continuum one $\partial_x^2\phi(x)$ to some order in a ,

$$a^2\Delta_x^{(2)}\phi(x) \equiv \phi(x+a) + \phi(x-a) - 2\phi(x) = (e^{a\partial_x} + e^{-a\partial_x} - 2)\phi(x) \quad (\text{B.34a})$$

$$= \left(a^2\partial_x^2 + \frac{1}{12}a^4\partial_x^4 + \mathcal{O}(a^6)\right)\phi(x) \quad (\text{B.34b})$$

$$= a^2\partial_x^2\left(1 + \frac{1}{12}a^2\partial_x^2\right)\phi(x) + \mathcal{O}(a^6). \quad (\text{B.34c})$$

On the one hand this shows that $\partial_\mu^*\partial_\mu$ is a good candidate for the second derivative in (B.32b) as the higher order contribution only affects the $\mathcal{O}(a^5)$ term there. On the other hand it leads us again by introducing the appropriate counterterm to a representation of a *symmetric improved second order lattice derivative*,

$$\begin{aligned} & \partial_\mu^*\partial_\mu\left(1 - \frac{1}{12}a^2\partial_\mu^*\partial_\mu\right)\phi(x) \\ &= -\frac{\phi(x+2a\hat{\mu}) - 16\phi(x+a\hat{\mu}) + 30\phi(x) - 16\phi(x-a\hat{\mu}) + \phi(x-2a\hat{\mu})}{12a^2}. \end{aligned} \quad (\text{B.35})$$

If we apply the same expansions to the improved derivatives as we did it for the standard ones we get,

$$\tilde{\partial}_\mu \left(1 - \frac{1}{6}a^2 \partial_\mu^* \partial_\mu\right) \phi(x) = \partial_x \left(1 - \frac{1}{30}a^4 \partial_x^4 + \mathcal{O}(a^6)\right) \phi(x) = \partial_x \phi(x) + \mathcal{O}(a^4) , \quad (\text{B.36})$$

$$\partial_\mu^* \partial_\mu \left(1 - \frac{1}{12}a^2 \partial_\mu^* \partial_\mu\right) \phi(x) = \partial_x^2 \left(1 - \frac{1}{90}a^4 \partial_x^4 + \mathcal{O}(a^6)\right) \phi(x) = \partial_x^2 \phi(x) + \mathcal{O}(a^4) . \quad (\text{B.37})$$

As one expects, the improved lattice derivatives approximate their continuum counterpart up to corrections of $\mathcal{O}(a^4)$, appropriate smoothness of the functions ϕ assumed.

B.4 Finite volume continuum gauge fields

In a (continuum) $SU(N)$ gauge field theory with finite volume, the gauge connection or vector field $A_\mu(x)$ has values in the corresponding Lie algebra $\mathfrak{su}(N)$ with generators $\{T^a\}$ and may thus be written as

$$A_\mu(x) = A_\mu^a(x) T^a , \quad (\text{B.38})$$

for real components $A_\mu^a(x)$. In this case the (anti-hermitean) fundamental representation of the generators is defined by

$$(T^a)^\dagger = -T^a , \quad \text{Tr } T^a T^b = -\frac{1}{2} \delta^{ab} . \quad (\text{B.39})$$

The field strength tensor is

$$F_{\mu\nu}(x) = \partial_\mu A_\nu(x) - \partial_\nu A_\mu(x) + [A_\mu(x), A_\nu(x)] , \quad (\text{B.40})$$

and the right/left acting form of the covariant derivative on fermion/anti-fermion fields is defined by

$$D_\mu \psi(x) = (\partial_\mu + A_\mu + i\theta_\mu/L) \psi(x) , \quad (\text{B.41})$$

$$\bar{\psi}(x) \overleftarrow{D}_\mu = \bar{\psi}(x) (\overleftarrow{\partial}_\mu - A_\mu - i\theta_\mu/L) . \quad (\text{B.42})$$

The phase factors λ_μ in the lattice theory appear here in the continuum of volume L^4 as Abelian gauge field $i\theta_\mu/L$.

B.5 Schrödinger functional correlation functions

A complete list of Wick contractions among fermion fields in the Schrödinger functional as defined in section 3.5, is given by [74]

$$[\psi(x) \bar{\psi}(y)]_{\text{F}} = S(x, y), \quad (\text{B.43a})$$

$$[\psi(x) \bar{\zeta}(y)]_{\text{F}} = \tilde{c}_t S(x, y) U(y - a\hat{0}, 0)^{-1} P_+ \Big|_{y_0=a}, \quad (\text{B.43b})$$

$$[\psi(x) \bar{\zeta}'(y)]_{\text{F}} = \tilde{c}_t S(x, y) U(y, 0) P_- \Big|_{y_0=T-a}, \quad (\text{B.43c})$$

$$[\zeta(x) \bar{\psi}(y)]_{\text{F}} = \tilde{c}_t P_- U(x - a\hat{0}, 0) S(x, y) \Big|_{x_0=a}, \quad (\text{B.43d})$$

$$[\zeta'(x) \bar{\psi}(y)]_{\text{F}} = \tilde{c}_t P_+ U(x, 0)^{-1} S(x, y) \Big|_{x_0=T-a}, \quad (\text{B.43e})$$

$$\begin{aligned} [\zeta(x) \bar{\zeta}(y)]_{\text{F}} &= \tilde{c}_t^2 P_- U(x - a\hat{0}, 0) S(x, y) U(y - a\hat{0}, 0)^{-1} P_+ \Big|_{y_0=a}^{x_0=a} \\ &\quad - \frac{1}{2} \tilde{c}_s P_- \gamma_k (\nabla_k^* + \nabla_k) a^{-2} \delta_{xy}, \end{aligned} \quad (\text{B.43f})$$

$$[\zeta(x) \bar{\zeta}'(y)]_{\text{F}} = \tilde{c}_t^2 P_- U(x - a\hat{0}, 0) S(x, y) U(y, 0) P_- \Big|_{y_0=T-a}^{x_0=a}, \quad (\text{B.43g})$$

$$[\zeta'(x) \bar{\zeta}(y)]_{\text{F}} = \tilde{c}_t^2 P_+ U(x, 0)^{-1} S(x, y) U(y - a\hat{0}, 0)^{-1} P_+ \Big|_{y_0=a}^{x_0=T-a}, \quad (\text{B.43h})$$

$$\begin{aligned} [\zeta'(x) \bar{\zeta}'(y)]_{\text{F}} &= \tilde{c}_t^2 P_+ U(x, 0)^{-1} S(x, y) U(y, 0) P_- \Big|_{y_0=T-a}^{x_0=T-a} \\ &\quad - \frac{1}{2} \tilde{c}_s P_+ \gamma_k (\nabla_k^* + \nabla_k) a^{-2} \delta_{xy}. \end{aligned} \quad (\text{B.43i})$$

The γ_5 -hermiticity of the propagator $S(x, y)$, $0 < x_0, y_0 < T$ together with $\gamma_5 P_{\pm} = P_{\mp} \gamma_5$ and $(P_{\pm})^{\dagger} = P_{\pm}$ implies $\gamma_5 [\zeta(\mathbf{z}) \bar{\psi}(x)]_{\text{F}} \gamma_5 = \{[\psi(x) \bar{\zeta}(\mathbf{z})]_{\text{F}}\}^{\dagger}$. As usual we have $U^{-1} = U^{\dagger}$.

The standard case

With the definitions (3.55) the correlation functions $C \in \{f_A, f_P, k_V, k_T\}$ and $C_1 \in \{f_1, k_1\}$ are generally given by

$$C^{ij}(x_0) = K_C a^9 \sum_{\mathbf{xyz}} \langle \bar{\psi}_i(x) \Gamma_C \psi_j(x) \bar{\zeta}_j(\mathbf{y}) \Gamma \zeta_i(\mathbf{z}) \rangle, \quad (\text{B.44a})$$

$$C_1^{ij} = K_{C_1} a^{12} \sum_{\mathbf{y}'\mathbf{z}'\mathbf{yz}} \langle \bar{\zeta}_i(\mathbf{y}') \Gamma \zeta_j(\mathbf{z}') \bar{\zeta}_j(\mathbf{y}) \Gamma \zeta_i(\mathbf{z}) \rangle, \quad (\text{B.44b})$$

with appropriate normalization constants K_C , K_{C_1} and Dirac-structures $\Gamma_C \in \{\gamma_5 \gamma_0, \gamma_5, \gamma_k, i\sigma_{0k}\}$, $\Gamma \in \{\gamma_5, \gamma_k\}$, respectively. In case of the vector insertions a summation over the spatial components k is implied. By construction it is easy to apply Wick's theorem. The result for $i \neq j$ is

$$C^{ij}(x_0) = -K_C a^9 \sum_{\mathbf{xyz}} \langle \text{Tr} \left\{ [\zeta_i(\mathbf{z}) \bar{\psi}_i(x)]_{\text{F}} \Gamma_C [\psi_j(x) \bar{\zeta}_j(\mathbf{y})]_{\text{F}} \Gamma \right\} \rangle, \quad (\text{B.45a})$$

$$C_1^{ij} = -K_{C_1} a^{12} \sum_{\mathbf{y}'\mathbf{z}'\mathbf{yz}} \langle \text{Tr} \left\{ [\zeta_i(\mathbf{z}) \bar{\zeta}_i(\mathbf{y}')]_{\text{F}} \Gamma [\zeta_j(\mathbf{z}') \bar{\zeta}_j(\mathbf{y})]_{\text{F}} \Gamma \right\} \rangle. \quad (\text{B.45b})$$

Now one inserts the Wick contractions (B.43) for each flavour and gets

$$C^{ij}(x_0) = -K_C a^3 \sum_{\mathbf{x}} \left\langle \text{Tr} \left\{ \tilde{c}_t a^3 \sum_{\mathbf{z}} P_- U(z - a\hat{0}, 0) S_i(z, \mathbf{x}) \Big|_{z_0=a} \Gamma_C \right. \right. \\ \left. \left. \tilde{c}_t a^3 \sum_{\mathbf{y}} S_j(\mathbf{x}, \mathbf{y}) U(\mathbf{y} - a\hat{0}, 0)^{-1} P_+ \Big|_{y_0=a} \Gamma \right\} \right\rangle, \quad (\text{B.46a})$$

$$C_1^{ij} = -K_{C_1} \left\langle \text{Tr} \left\{ \tilde{c}_t^2 a^6 \sum_{\mathbf{y}'\mathbf{z}} P_- U(z - a\hat{0}, 0) S_i(z, \mathbf{y}') U(\mathbf{y}', 0) P_- \Big|_{y'_0=T-a}^{z_0=a} \Gamma \right. \right. \\ \left. \left. \tilde{c}_t^2 a^6 \sum_{\mathbf{z}'\mathbf{y}} P_+ U(\mathbf{z}', 0)^{-1} S_j(\mathbf{z}', \mathbf{y}) U(\mathbf{y} - a\hat{0}, 0)^{-1} P_+ \Big|_{y_0=a}^{z'_0=T-a} \Gamma \right\} \right\rangle. \quad (\text{B.46b})$$

The summation over the boundary fields is trivial and suggest to define for arbitrary flavour i ,

$$\bar{S}_i(x) = \tilde{c}_t a^3 \sum_{\mathbf{y}} S_i(\mathbf{x}, \mathbf{y}) U(\mathbf{y} - a\hat{0}, 0)^{-1} P_+ \Big|_{y_0=a}, \quad (\text{B.47a})$$

$$\bar{S}_{T,i} = \tilde{c}_t a^3 \sum_{\mathbf{z}'} P_+ U(\mathbf{z}', 0)^{-1} \bar{S}_i(\mathbf{z}') \Big|_{z'_0=T-a}, \quad (\text{B.47b})$$

the *boundary-to-bulk* and *boundary-to-boundary propagator*, \bar{S} and \bar{S}_T respectively. These are the propagators corresponding to zero-momentum sources at the boundary $x_0 = 0$. Using the γ_5 -hermiticity property of the propagators, the correlation functions become

$$C^{ij}(x_0) = -K_C \left\langle \text{Tr} \left\{ \bar{S}_i^\dagger(x) \gamma_5 \Gamma_C \bar{S}_j(x) \Gamma \gamma_5 \right\} \right\rangle, \quad (\text{B.48a})$$

$$C_1^{ij} = -K_{C_1} \left\langle \text{Tr} \left\{ \bar{S}_{T,i}^\dagger \gamma_5 \Gamma \bar{S}_{T,j} \Gamma \gamma_5 \right\} \right\rangle. \quad (\text{B.48b})$$

Note that the propagators \bar{S} and \bar{S}_T are obtained by the same inversion because the latter is just the continuation of \bar{S} to the boundary at $x_0 = T$. \bar{S} on the other hand is the propagator in the bulk, defined by the usual lattice Dirac equation

$$(2\kappa)^{-1} M(x, y)_{AB}^{ab} \equiv [aD + a\delta D + am_0](x, y)_{AB}^{ab}, \quad (\text{B.49a})$$

$$a^4 \sum_y M(x, y)_{AB}^{ab} S(y, z)_{BC}^{bc} = 2\kappa \delta_{x,z} \delta_{A,C} \delta_{a,c}, \quad (\text{B.49b})$$

attached to the boundary at $x_0 = 0$. To get the source vector for the propagator $\bar{S}(x)$ one just applies M to equation (B.47a),

$$a^4 \sum_y M(x, y)_{AB}^{ab} \bar{S}(y)_{BC}^{bc} = \tilde{c}_t a^7 \sum_{\mathbf{z}, \mathbf{y}} M(x, y)_{AB}^{ab} [S(y, z) U(z - a\hat{0}, 0)^{-1} P_+]_{BC}^{bc} \Big|_{z_0=a} \\ = \tilde{c}_t a^7 \sum_{\mathbf{z}, \mathbf{y}} M(x, y)_{AB}^{ab} S(y, z)_{BD}^{bd} [U(z - a\hat{0}, 0)^{-1} P_+]_{DC}^{dc} \Big|_{z_0=a} \\ = \tilde{c}_t a^3 \sum_{\mathbf{z}} 2\kappa \delta_{x,z} \delta_{A,D} \delta_{a,d} [U(z - a\hat{0}, 0)^{-1} P_+]_{DC}^{dc} \Big|_{z_0=a} \\ = 2\kappa \tilde{c}_t \delta_{x_0, a} [U^\dagger((0, \mathbf{x}), 0) P_+]_{AC}^{ac}. \quad (\text{B.50})$$

We did not mentioned the backward correlation functions yet, but it should be clear from the discussion above that by applying the very same steps one gets the corresponding backward prop-

agator \bar{R} fulfilling

$$a^4 \sum_y M(x, y)_{AB}^{ab} \bar{R}(y)_{BC}^{bc} = 2\kappa \tilde{c}_t \delta_{x_0, T-a} [U^\dagger((T-a, \mathbf{x}), 0) P_-]_{AC}^{ac}, \quad (\text{B.51a})$$

$$\bar{R}_i(x) = \tilde{c}_t a^3 \sum_y S_i(x, y) U(y, 0) P_- \Big|_{y_0=T-a}, \quad (\text{B.51b})$$

$$\bar{R}_{T,i} = \tilde{c}_t a^3 \sum_{z'} P_- U(z' - a\hat{0}, 0)^{-1} \bar{R}_i(z') \Big|_{z'_0=a}. \quad (\text{B.51c})$$

For completeness we also list the backward boundary-to-boundary propagator \bar{R}_T . It already appears in eq. (B.46b) which hence could be written as $C_1^{ij} = -K_{C_1} \langle \text{Tr} \{ \bar{R}_{T,i} \Gamma \bar{S}_{T,j} \Gamma \} \rangle$, showing that $\bar{R}_T = \gamma_5 \bar{S}_T^\dagger \gamma_5$ holds. This explicitly verifies our expectation that the boundary-to-boundary correlation functions are not special to any of the boundaries. However, to compute backward correlation functions we still have to solve eq. (B.51a). They are obtained by substituting the forward propagator in favour of the backward ones in eq. (B.48b). For vanishing boundary fields the forward and backward correlation functions obey time reflection symmetry at $x_0 = T/2$ and in some cases we will average them accordingly to reduce the statistical noise in the Monte-Carlo simulation further. Actually, such a symmetry that holds at the classical level is only expected to be recovered in the statistical mean, i.e. after gauge-averaging. An important observation is that

$$\bar{S}(x) P_- = 0 = \bar{R}(x) P_+ \quad (\text{B.52})$$

holds. Thus, the number of independent Dirac-components of the propagators is reduced by a factor of two and only half of the usual inversions has to be done explicitly.

The extended case

In [77, 155] it was advantageous to use a wave function smearing at the boundary sources. Thus, we also extended our code in this regard for future purposes. To this end we introduce real (scalar) wave functions ω at the boundary sources by writing

$$\mathcal{O}_{ji}(\omega) \equiv a^6 \sum_{\mathbf{y}\mathbf{z}} \bar{\zeta}_j(\mathbf{y}) \omega(\mathbf{y} - \mathbf{z}) \Gamma \zeta_i(\mathbf{z}), \quad (\text{B.53a})$$

$$\mathcal{O}'_{ji}(\omega) \equiv a^6 \sum_{\mathbf{y}'\mathbf{z}'} \bar{\zeta}'_j(\mathbf{y}') \omega(\mathbf{y}' - \mathbf{z}') \Gamma \zeta'_i(\mathbf{z}'). \quad (\text{B.53b})$$

But this interrelates the previously independent contractions in (B.44a) and especially (B.44b). Since global sums are expensive it is desirable to disentangle those additional convolutions. This is partially achieved by introducing a noisy estimator or *stochastic source* σ for the spatial components. It is implicitly defined by

$$\delta_{\mathbf{y},\mathbf{z}} = \langle \sigma(\mathbf{y}) \sigma(\mathbf{z}) \rangle, \quad a^3 \sum_{\mathbf{z}} \delta_{\mathbf{y},\mathbf{z}} = 1, \quad (\text{B.54})$$

where $\langle \bullet \rangle$ means the average with respect to the field configurations σ . To use this at the level of correlation functions we (formally) have to introduce one additional sum. We only discuss this for

the boundary-to-bulk correlation function, which now reads

$$C^{ij}(x_0, \omega) = -K_C a^9 \sum_{\mathbf{r}\mathbf{y}\mathbf{z}} \left\langle \text{Tr} \left\{ [\zeta_i(\mathbf{r}) \bar{\psi}_i(x)]_{\text{F}} \Gamma_C [\psi_j(x) \bar{\zeta}_j(\mathbf{y})]_{\text{F}} \Gamma \right\} \delta_{\mathbf{r},\mathbf{z}} \omega(\mathbf{y} - \mathbf{z}) \right\rangle, \quad (\text{B.55})$$

$$= -K_C a^9 \sum_{\mathbf{r}\mathbf{y}\mathbf{z}} \left\langle \text{Tr} \left\{ [\zeta_i(\mathbf{r}) \bar{\psi}_i(x)]_{\text{F}} \sigma(\mathbf{r}) \Gamma_C [\psi_j(x) \bar{\zeta}_j(\mathbf{y})]_{\text{F}} \omega(\mathbf{y} - \mathbf{z}) \sigma(\mathbf{z}) \Gamma \right\} \right\rangle. \quad (\text{B.56})$$

Obviously, the sum between both Wick contractions is not correlated anymore. Again, the next step is to insert (B.43) and identify a new form for the propagators as done in (B.47). But finally this is equivalent to choosing another RHS for the modified Dirac equation, which now read

$$a^4 \sum_y M(x, y)_{AB}^{ab} S(y, \sigma)_{BC}^{bc} = 2\kappa \tilde{c}_t \delta_{x_0, a} \sigma(\mathbf{x}) [U^\dagger((0, \mathbf{x}), 0) P_+]_{AC}^{ac}, \quad (\text{B.57})$$

$$a^4 \sum_y M(x, y)_{AB}^{ab} S(y, \omega_\sigma)_{BC}^{bc} = 2\kappa \tilde{c}_t \delta_{x_0, a} \omega_\sigma(\mathbf{x}) [U^\dagger((0, \mathbf{x}), 0) P_+]_{AC}^{ac}, \quad (\text{B.58})$$

$$\omega_\sigma(\mathbf{y}) = a^3 \sum_{\mathbf{z}} \omega(\mathbf{y} - \mathbf{z}) \sigma(\mathbf{z}). \quad (\text{B.59})$$

Here ω_σ is the convolution of the wave function ω and the stochastic source σ . The convolution ω_σ is clearly independent of the flavour and can be computed in advance. With these propagators the correlation functions become

$$C^{ij}(x_0) = -K_C \left\langle \text{Tr} \left\{ S_i^\dagger(x, \sigma) \gamma_5 \Gamma_C S_j(x, \omega_\sigma) \Gamma \gamma_5 \right\} \right\rangle, \quad (\text{B.60a})$$

$$C_1^{ij} = -K_{C_1} a^6 \sum_{\mathbf{y}'\mathbf{z}'} \left\langle \text{Tr} \left\{ S_{T,i}^\dagger(\mathbf{y}', \sigma) \omega(\mathbf{y}' - \mathbf{z}') \gamma_5 \Gamma S_{T,j}(\mathbf{z}', \omega_\sigma) \Gamma \gamma_5 \right\} \right\rangle, \quad (\text{B.60b})$$

with

$$S_T(\mathbf{z}', \star) = \tilde{c}_t P_+ U(\mathbf{z}', 0)^{-1} S(\mathbf{z}', \star) \Big|_{z'_0 = T-a}, \quad \star = \sigma, \omega_\sigma. \quad (\text{B.61})$$

C Optimization techniques

For this work all simulations and measurements are done on two generations of APE (Array Processor Experiment) massively parallel computers [137], which are custom designed and optimized for lattice gauge theory simulations. All runs for the production of gauge field configurations are done on the latest generation, the apeNEXT computers while only a minor part of the measurement was performed on the foregoing generation, APEmille. An important difference between these two generations is that the processors on APEmille with SIMD architecture are synchronized while those on apeNEXT with a MIMD architecture are running asynchronously.³⁴ Both are using the VLIW³⁵ control style [141] to execute sequences of compiler prepared microcode words. Using VLIW allows to reduce the hardware's complexity (and hence costs) by putting most of the complexity of instruction scheduling onto the compiler. Furthermore, *pipelining* becomes

³⁴SIMD: Single Instruction, Multiple Data; MIMD: Multiple Instruction, Multiple Data

³⁵Very Long Instruction Word

Table 21: A typical run-time profile of the ALPHA HMC code on apeNEXT ($24^3 \times 32$ lattice, 1 traj.)

routine	calls	time	$\langle \text{CMA} \rangle$	Σ lines
Dirac operator (3 variants)	80844	58 %	$O(350)$	$O(2000)$
Linear algebra (3 routines)	60736	26 %	$O(100)$	$O(1000)$
Gauge forces + update	320	8 %	$O(2000)$	$O(2000)$
Global sum (crate, 128 bit)	83554	0.4 %	20	$O(200)$
Others (≈ 70 routines)		7 %		$O(15000)$

possible. That is a simultaneous execution of sub-steps of sequential instructions, resulting in a performance gain. The special design of APE-computers came along with a proprietary high-level programming language called TAO. Furthermore, the Zz dynamic parser [138] that works on top of a given programming language allows for a dynamic extension of the TAO syntax and grammar at compile time. Thus an object-oriented programming style becomes possible and enables among other things operator overloading which is of crucial importance for writing readable code and dealing with special structures that appear in QCD. For apeNEXT also a C-language compiler exists because the MIMD architecture enables a MPI-based parallelism. First, our focus was on the optimization of the ALPHA-collaboration TAO program code for the production of $N_f = 2$ dynamical gauge configurations. In particular the optimization of linear algebra routines is crucial as they will be called several times by the matrix inversion routine like the *Conjugate Gradient* (CG) or *stabilized Conjugate Gradient* (BiCGStab) *algorithm*, cf. table 21. APE-machines are traditionally partitioned in a powerful 3-dimensional communication network and the smallest hardware building block is a board which consists of $2 \times 2 \times 2 = 8$ or $4 \times 2 \times 2 = 16$ nodes for APEmille or apeNEXT respectively. Four boards make one unit and four units one crate. Depending on the lattice size and the hardware topology that is used, one usually gets *replica* of a simulation. That are copies of the same program that run simultaneously on non-interfering sub-partitions of the machine. Some remarks about the node functionality of apeNEXT [140] are in order:

- Each processing node, fig. 19, is a fully independent processor with a flow-control and number-crunching unit. The latter consists of a floating point, integer and logical unit the FPU, INT and LU, respectively. Furthermore, a special function unit (LUT) is implemented. All operands for the arithmetic unit (P0, P1, P2) are provided by the register file and results (P3) are written back there.
- The node memory controller (MC) supports a (DDR-SDRAM) memory bank of size 256-1024 MB where data *and* program instructions are stored. This is opposed to APEmille where data and program memory was separated. Nodes (asynchronously) execute their own copy of the program and get implicitly synchronized only when performing a data-exchange operation. A feature to boost (sustained) performance is the address-generation unit (AGU) which concurrently to the arithmetic box computes addresses for memory access.
- The intrinsic data format is 64-bit (double) precision with all floating-point (FP) data represented in the IEEE standard. A *normal operation* $\pm a \times b \pm c$ for complex numbers a , b and c (CNORM) can be started at each clock cycle with bus frequency of 200 MHz. It provides

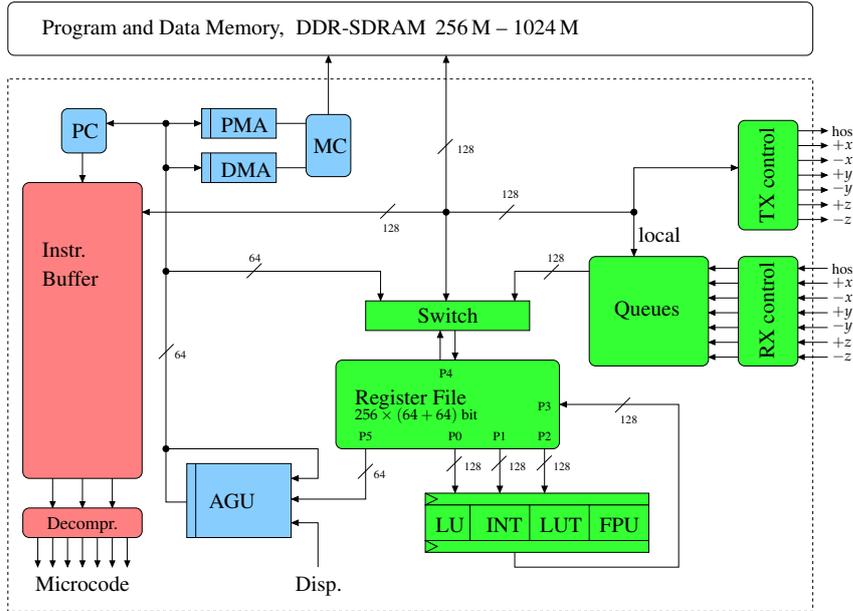


Figure 19: apeNEXT processor design

a maximal throughput of eight FP operation per clock cycle, resulting in a peak performance of 1.6 GFlops³⁶ per node. There is a large number of 256 paired 64-bit registers (complex or integer) and a program cache, the *instruction buffer*, is under full software (SW) control. For the purpose of optimization the hardware (HW) also supports compression of microcode, i.e. the de-compression of instructions is performed on-the-fly by dedicated hardware.

- Network communications and processor operations are working concurrently and independent after a communication request is queued. A bi-directional connection between nodes is established by two separated network controllers, the transmitter (TX) and receiver (RX). Both are equipped with seven links, six for the 3-dimensional node-to-node communication between neighbours and one for communicating with the host subsystem (I/O operations, etc.). A *queue* on each node allows to load *local* (1024 complex) and *remote* (6×128 bit) data. The queue is one important architectural enhancement since it is possible to route all (local or remote) read memory accesses through it. Later, the data in the queue can be accessed by the processor with zero latency and hence exploits the concurrency between memory access, arithmetic operations and network transfers.

Performing data pre-fetch is most important in critical kernel loops, e.g. especially in QCD simulations when a multiplication with the hopping term appears because it involves remote communications. In table 21 we list a decomposition of our HMC code in different parts showing that most of the time is spent in routines concerning the Dirac operator and linear algebra. Thus the performance gain is largest if we focus on optimizations in those parts of the code. In the following I will describe some standard optimization techniques as well as apeNEXT specific optimizations which were investigated to increase the performance of our code.

³⁶Flops: Floating-point operations per second

The main tool for optimization was a *static performance analysis* using the programs `nperf` and `dispmunit` provided by the APE-team. `nperf` gives the number of instructions used in different blocks of the program and furthermore a rough estimate how well the arithmetic unit was utilised. Suppose we want to optimize a linear algebra routine which adds a rescaled spinor to another one and stores the result in a third spinor:

```

1 subroutine linalg_lc2( integer i1, integer i2, integer k, complex cst)
  register su3_spinor p1, p2, p3

  do t=0, TIME_SIZE-1
    do x=0, VOL3-1
6      p1 = phi[i1,t,x]      !! load operand p1
      p2 = phi[i2,t,x]      !! load operand p2
      p3 = p1 + p2*cst      !! compute
      phi[k,t,x] = p3      !! store result p3
    enddo
11  enddo
end

```

Here the global array `phi` contains all spinors. Each appearance of `[i1,t,x]` causes a memory address computation for either a subsequent read or store operation. Furthermore, the integers used in this address computation are located in the main memory and thus have to be loaded for each iteration again. But explicit memory access is usually a slow operation and always causes a new address computation. This bottleneck can be easily avoided by storing the integers `i1`, `i2` and `k` in a physical register before entering the loop. Inside the loop one uses the corresponding `physreg` variable which is stored in the cache and hence can be accessed very fast. Beside other optimizations this was done in the following optimized routine. I directly show its typical output as produced by `nperf`:

```

1008b99 | subroutine linalg_c2_impr( complex z2, integer i1, integer i2, integer k )
  --> GL_0x1019 (L)
3 |     physreg integer pix, pk, pi1, pi2, pitt
  |     physreg complex rz2
  |     physreg su3_spinor_block psi1, psi2, chi
  |
  |     rz2 = z2
  |     pk = k
8 |     pi1 = i1
  |     pi2 = i2
272 | 1 % C: 0 F: 0 M: 0 X: 0 L: 0 I: 2 IQO: 7/7 1/1 0/0

1008c21 | cache do pitt=1, TIME_SIZE-1 --> GL_0x108b (L)
13 | 28 7 % C: 0 F: 0 M: 0 X: 0 L: 0 I: 2 IQO: 0/0 0/0 2/2

1008c3d | do pix=0, VOL3-1, LINALG_SPACE_BLOCK --> GL_0x108d (L)
  |
  |     extract psi1 from phi[ pi1, pitt, pix ]
18 |     extract psi2 from phi[ pi2, pitt, pix ]
  |     /for b=0 to LINALG_SPACE_BLOCK-1 {
  |         chi.[b] = psi1.[b] + rz2*psi2.[b]
  |     }
  |     replace chi into phi[ pk, pitt, pix ]
23 |
140 | 18 % C: 24 F: 0 M: 0 X: 0 L: 0 I: 1 IQO: 0/0 48/2 24/1

```

abbrev.	operation	abbrev.	operation
C	CNORM		
F	FNORM	I	integer
M	LUTMOVE		MTR-input/load bursts
X	LUTCROSS	IQQ	QTR input/MTQ load bursts
L	LUT		RTM output/ store

Table 22: Abbreviations appearing in the static performance analysis. CNORM: *complex normal operation*; FNORM: *FP vector normal operation*; LUT*: *operations on data in registers*

operation	abbreviation	description
M2R	Memory-To-Register	generic data transfer from memory to register file
M2Q	Memory-To-Queue	memory to pre-fetch queue data transfer
R2M	Register-To-Memory	RegisterFile to memory data transfer
R2Q	Register-To-Queue	RegisterFile to pre-fetch queue data transfer
Q2R	Queue-To-Register	fetch data from queue and write it to RegisterFile
A2RF	AGU-To-RegisterFile	performs AGU-to-RegisterFile data transfer

Table 23: apeNEXT address generation and memory access instructions

```

1008cc9 |      enddo                                --> GL_0x108e (L)
      36  0 % C: 0 F: 0 M: 0 X: 0 L: 0 I: 0 IQO: 2/2  0/0  0/0
28
1008ced |      enddo                                --> GL_0x108c (L)
      46  0 % C: 0 F: 0 M: 0 X: 0 L: 0 I: 0 IQO: 0/0  0/0  0/0
1008d1b |      end                                  --> GL_0x108f (L)

```

Here line 24 is the most important one because it contains informations about the performance of the innermost (volume) loop. First, the number of microcode instructions are listed, followed by the theoretical performance of the corresponding code segment. The abbreviations appearing in the following are summarized in table 22. Explicitly, this line tells us that 24 complex normal operation are performed, because each spinor has 12 components and the compile time variable `LINALG_SPACE_BLOCK` was set to 2. Setting this variable to 4 doubles the number of normal operations as well as the IO operations:

```

212 23 % C: 48 F: 0 M: 0 X: 0 L: 0 I: 1 IQO: 0/0 96/2 48/1

```

In line 15 of the static performance analysis output, the inner loop has the label `GL_0x108d` attached. With `dispmnit` the exact execution order of microcode instructions can be looked up to get further hints for optimization. A part of it is shown in table 24. The subroutine under consideration uses burst memory access. First the routine gets initialized and performs some memory access computations. After the `M2Q` instruction was given, 15 cycles later it starts at address `01008C69` with reading a bunch of 48 data words as expressed by `Q2R`. 26 cycles later the first complex normal operations `CN00` is performed, starting at address `01008C83`. This nicely shows pipelining. Only two cycles after the last `Q2R` access, all 24 complex normal operations are performed.

Table 24: Microcode example on apeNEXT. The lines are executed at successive cycles. For abbreviations see main text, table 23 and figure 19.

ADDR	DISP	C5	MCC	STKC	FLW	IOC	AGU	ASEL	BS5	P5	BS4	C4	P4	MPC	BS3	C3	P3	P2	P1	P0	ZADDR
<i>STARTPMA: 0x00FFBCDC ENDFMA: 0x00FFBD16 LEN: 140</i>																					
4	01008C3D: 0d08cc9	0	-	-	-	-	-	-	0	00	0	0	00	0	0	00	00	00	11		0x00FFBCDC
	01008C3E: 00000000	0	-	-	-	-	-	LFPC	0	00	0	0	00	-	0	00	00	00	00		0x00FFBCDD
.....																					
	01008C59: 00000000	0	-	-	-	-	-	-	0	00	0	0	00	-	0	00	00	00	00	2	0x00FFBCE6
	01008C5A: 00000000	0	M2Q	-	-	-	RA	-	0	00	0	0	00	-	0	00	00	00	00		0x00FFBCE7
9	01008C5B: 00000000	0	-	-	-	-	-	-	0	00	3	0	1d	-	0	00	00	00	00	1	0x00FFBCE7
	01008C5C: 00000000	0	-	-	-	-	-	-	0	1d	0	0	00	-	0	00	00	00	00	2	0x00FFBCE7
	01008C5D: 00000000	0	-	-	-	-	-	-	0	00	0	0	00	-	0	00	00	00	00	3	0x00FFBCE7
	01008C5E: 000000d8	0	-	-	-	-	RXE	-	0	18	0	0	00	-	0	00	00	00	00		0x00FFBCE8
	01008C5F: 00000000	0	-	-	-	-	-	-	0	00	0	0	00	-	0	00	00	00	00	1	0x00FFBCE8
14	01008C60: 00000000	0	-	-	-	-	A2RF	RA	-	0	00	0	00	-	0	00	00	00	00	2	0x00FFBCE8
	01008C61: 00000000	0	-	-	-	-	-	-	0	00	3	0	1d	-	0	00	00	00	00	3	0x00FFBCE8
.....																					
	01008C67: 00000019	0	-	-	-	-	RXE	LAL	0	1c	0	0	00	-	0	00	00	00	00	1	0x00FFBCEA
	01008C68: 00000000	0	-	-	-	-	-	-	0	00	0	0	00	-	0	00	00	00	00	2	0x00FFBCEA
19	01008C69: 00000000	0	-	-	-	-	Q2R	RA	-	0	00	0	00	-	0	00	00	00	00		0x00FFBCEB
	01008C6A: 00000000	0	-	-	-	-	Q2R	-	0	00	3	0	45	-	0	00	00	00	00	1	0x00FFBCEB
	01008C6B: 00000000	0	-	-	-	-	Q2R	-	0	00	3	0	46	-	0	00	00	00	00	2	0x00FFBCEB
	01008C6C: 00000000	0	-	-	-	-	Q2R	-	0	00	3	0	47	-	0	00	00	00	00		0x00FFBCEC
	01008C6D: 00000000	0	-	-	-	-	Q2R	-	0	00	3	0	48	-	0	00	00	00	00	1	0x00FFBCEC
24	01008C6E: 00000000	0	-	-	-	-	Q2R	-	0	00	3	0	49	-	0	00	00	00	00	2	0x00FFBCEC
	01008C6F: 00000000	0	-	-	-	-	Q2R	-	0	00	3	0	4a	-	0	00	00	00	00		0x00FFBCEC
	01008C70: 00000000	0	-	-	-	-	Q2R	-	0	00	3	0	4b	-	0	00	00	00	00	1	0x00FFBCEC
	01008C71: 00000000	0	-	-	-	-	Q2R	-	0	00	3	0	4c	-	0	00	00	00	00	2	0x00FFBCEC
	01008C72: 00000000	0	M2Q	-	-	-	Q2R	-	0	00	3	0	4d	-	0	00	00	00	00		0x00FFBCEE
29	01008C73: 00000000	0	-	-	-	-	Q2R	-	0	00	3	0	4e	-	0	00	00	00	00	1	0x00FFBCEE
.....																					
	01008C82: 00000000	0	-	-	-	-	Q2R	-	0	00	3	0	5d	-	0	00	00	00	00	1	0x00FFBCF3
	01008C83: 00000000	0	-	-	-	-	Q2R	-	0	00	3	0	5e	CN00	0	00	00	44	5d		0x00FFBCF4
	01008C84: 00000000	0	-	-	-	-	Q2R	-	0	00	3	0	5f	CN00	0	00	00	44	5e	1	0x00FFBCF4
34	01008C85: 00000000	0	-	-	-	-	Q2R	-	0	00	3	0	60	CN00	0	00	45	44	5f		0x00FFBCF5
	01008C86: 00000000	0	-	-	-	-	Q2R	-	0	00	3	0	61	CN00	0	00	46	44	60	1	0x00FFBCF5
	01008C87: 00000000	0	-	-	-	-	Q2R	-	0	00	3	0	62	CN00	0	00	47	44	61		0x00FFBCF6
.....																					
	01008C98: 00000000	0	-	-	-	-	Q2R	-	0	00	3	0	73	CN00	3	80	58	44	72		0x00FFBD04
39	01008C99: 00000000	0	-	-	-	-	A2RF	-	0	00	3	0	74	CN00	3	81	59	44	73		0x00FFBD05
	01008C9A: 00000000	0	-	-	-	-	-	-	0	00	3	0	1d	CN00	3	82	5a	44	74		0x00FFBD06
	01008C9B: 00000000	0	-	-	-	-	-	-	0	1d	0	0	00	-	3	83	5b	00	00	1	0x00FFBD06
	01008C9C: 00000000	0	-	-	-	-	-	-	0	00	0	0	00	-	3	84	5c	00	00		0x00FFBD07
	01008C9D: 000000d8	0	-	-	-	-	RXE	-	0	18	0	0	00	-	3	85	00	00	00	1	0x00FFBD07

Some optimization techniques which were tested and partially mixed are in order:

- *instruction caching*: While loading instructions into the cache, no other operations can be performed. Since instruction loading competes with data loading (both access the same memory banks) it is advantageous to hold the loop body in the cache. This means that the processor has to load instructions only at the beginning of the loop instead of every time the loop body is executed. Above this is implicit in the statement `cache do . . . enddo`.
- *loop-unrolling*: If the loop body itself is very small, the overhead which the loop produces at the beginning and end can become significant. In this case it is helpful to duplicate the whole body several (`LINALG_SPACE_BLOCK`) times. This is hidden in the `Zz`-statement `/for . . . { . . . }` above.
- *burst memory access*: For each memory access instruction, the processor has to load the address into a register before initiating the memory access itself. This takes at least four clock cycles. It is possible to instruct the processor to load more data into the register with only one instruction. This is done with the `extract` statement by making use of a new data type which concatenates a basic data type into consecutive blocks of size `LINALG_SPACE_BLOCK`. Thus it is reasonable to use loop-unrolling and burst memory access together.
- *partial sum*: While the `apeNEXT` processor is able to compute one `CNORM` operation per clock cycle, the result is not immediately available. Thus a small overhead is introduced. One can get rid of it by removing some dependencies in the calculation. An example to do that is to use several register variables which can independently be summed within the loop body. The remaining summations among the register variables can then be done outside of the loop body.
- *prefetching*: As already mentioned, the waiting time for local or remote memory accesses can almost be hidden by making use of the pre-fetch queue. To this end one prepares a memory access operation at least one loop iteration before the data is really needed. The queue works as FIFO (First In, First Out) and is thus easy to control. In TAO one just uses the statements `queue=. . .` and `. . .=queue` to submit data to or receive data from the queue, respectively.

Because it is cumbersome and time consuming to run tests using the full production code for optimization purposes, we decided to write a small wrapper program. It provided a minimal environment to test the linear algebra routines under real conditions and allowed a fast compilation and executing on the available `apeNEXT` test-board. As standard candles some routines already used on `APEmille` have been used. In table 25 and 26 we list run-time measurements for several Dirac and linear algebra routines, respectively.

For the measurement program our starting point was the one used to compute HQET correlation functions [12] on `APEmille`. Thus we had to implement the relativistic forward/backward SF correlation functions. Another major component was to adopt the code such that it compiles

Table 25: Run-time (in ms) of some selected Dirac routines for different APE generations on a 8^4 lattice with $2 \times 2 \times 2$ topology. If the MLAT flag is set to two, our code doubles the number of total replica by introducing a second replica on the software side. APEmille64 is just APEmille but with 64-bit software precision.

routine	apeNEXT		APEmille		APEmille64	
	1	2	1	2	1	2
MLAT:						
qphi8	1.37	1.78	2.98	3.61	8.32	9.94
qphi10	0.43	0.56	1.00	1.30	2.09	2.58
qphi15	1.27	1.65	2.86	3.46	7.20	8.69
qphi16	1.43	1.85	3.24	3.87	9.2	11.0
qphi18	0.43	0.52	0.99	1.19	2.11	2.54
lc2o+qphi10	0.65	0.81	1.44	1.77	2.63	3.23
qphi19	1.75	2.24	4.06	4.90	–	–
sp2o+qphi16	2.14	2.72	4.19	5.02	10.3	12.4

Table 26: Run-time (in ms) of some selected linear algebra routines for different APE generations on a $8^3 \times 32$ lattice with $2 \times 2 \times 2$ topology. [lc: linear combination, sp: scalar product, o: odd sites only]

routine	apeNEXT		APEmille		APEmille64	
	1	4	1	4	1	4
block_size:						
lc1o	0.70	0.35	1.38	0.66	1.79	1.09
lc2o	0.92	0.48	1.80	0.91	2.39	1.50
lc2o*	0.83	0.46	1.51	0.84	2.12	1.45
lc2do	3.72	2.76	4.85	3.97	5.23	4.18
2*lc2o+sp2o	4.88	3.35	7.84	5.47	9.78	6.38
sp1o	5.40	4.73	7.94	7.42	8.34	6.07
sp2o	3.03	2.38	4.24	3.64	5.00	3.38
sp1	5.52		4.41		8.43	

and runs also on apeNEXT. The solver in our measurement program is that of [139] with SSOR preconditioning. We introduced the standard and extended correlation functions of appendix B.5 and also implemented the multiple sources technique used in [122].

D Error estimation techniques

It exists several methods to estimate the uncertainty of observables obtained by Monte-Carlo methods.³⁷ They are all based on an re-sampling of the original data set. We use the Jackknife and the Γ -method [132] as described in the next two subsections. Which method we exactly apply for the different observables is explained in the corresponding sections and in some cases we apply both to check that one method alone does not over- or underestimate the statistical error.

D.1 The Jackknife method

The Jackknife method only deals with one ensemble, i.e. if we get R replica from our Monte-Carlo simulation, we concatenate them into one long history of length $N = \sum_{r=1}^R N_r$. We usually deal with replica of the same length, i.e. $N_1 = N_2 = \dots = N_R$. If we are in thermal equilibrium and N_r is large enough, this naturally separates any remnant correlation among different replica which were generated at the same molecular dynamics time. For more general considerations we divide N into N_B blocks of consecutive measurements of length B . By reducing our data set in each replica by a small amount we can always make B a divisor of N . We label the Monte-Carlo estimates for a full set of N_α *primary observables* (correlation functions) by $\{a_\alpha^i | i = 1, \dots, N; \alpha = 1, \dots, N_\alpha\}$. Their *mean value* is denoted as

$$\bar{a}_\alpha = \frac{1}{N} \sum_{i=1}^N a_\alpha^i . \quad (\text{D.1})$$

Each *secondary observable* is derived of those by a function f . Its best estimate is given by

$$\bar{f} = f(\bar{a}_\alpha) , \quad (\text{D.2})$$

and becomes exact for independent measurements in the limit $N \rightarrow \infty$. With the *block average* of the primary observables,

$$b_\alpha^k = \frac{1}{B} \sum_{j=1}^B a_\alpha^{j+(k-1)B} , \quad k = 1, \dots, N_B , \quad (\text{D.3})$$

the *Jackknife bins* or Jackknife samples are defined by

$$c_\alpha^k \equiv \frac{1}{N-B} \left[\sum_{i=1}^N a_\alpha^i - \sum_{j=1}^B a_\alpha^{j+(k-1)B} \right] = \frac{1}{N-B} \left[N\bar{a} - Bb_\alpha^k \right] . \quad (\text{D.4})$$

Thus each Jackknife bin c_α^k is an average of the block-subtracted original data set. The blocking procedure reduces the correlation among the c_α^k -samples compared to the a_α^i -samples. This gives us a tool to study the correlation by looking at the uncertainty of an observable while increas-

³⁷The term *Monte Carlo method* was coined by N. Metropolis and S. Ulam [105] who were working amongst other theoreticians on problems about nuclear fission at the *Los Alamos National Laboratory* during and after World War II. For a historical retrospection see for instance *Los Alamos Science*, No. 15, 1987, *Special Issue*, S. Ulam.

ing the block size B . A reliable error estimate is given if the uncertainty is stable with varying block size which also depends on the real amount of the available statistic. $B = 1$ marks the standard Jackknife procedure which is in one-to-one correspondence with the original one. If the autocorrelation time among the primary observables is not too large, blocking allows to work with completely independent measurements subsequently. For two secondary observables f and g that were computed on the same gauge configuration, their correlation is given by the Jackknife covariance,

$$\text{Cov}_J(f, g) = \frac{N_B - 1}{N_B} \sum_{k=1}^{N_B} \left(f(c_\alpha^k) - \bar{f} \right) \left(g(c_\alpha^k) - \bar{g} \right), \quad (\text{D.5})$$

which for $g = f$ reduces as usual to the Jackknife variance of one observable,

$$\bar{\sigma}_J^2(f) = \frac{N_B - 1}{N_B} \sum_{k=1}^{N_B} \left(f(c_\alpha^k) - \bar{f} \right)^2. \quad (\text{D.6})$$

D.2 The Gamma method

The Γ -method explicitly takes correlations between different replica into account. Hence we label the primary observables $a_\alpha^{i,r}$ using an additional replica index r . The number of measurements per replica do not have to be equal. Here the key quantity is the *autocorrelation function* $\Gamma_{\alpha\beta}$ due to Sokal [130], defined by

$$\left\langle \left(a_\alpha^{i,r} - A_\alpha \right) \left(a_\beta^{j,s} - A_\beta \right) \right\rangle = \delta_{r,s} \Gamma_{\alpha\beta}(j - i), \quad \Gamma_{\alpha\beta}(n) = \Gamma_{\beta\alpha}(-n). \quad (\text{D.7})$$

Obviously, $n = j - i$ is proportional to the molecular dynamics time and replica $r, s = 1, \dots, R$ are naturally treated as statistical numerical experiments of the same system with independent random numbers and initial states. The A_α are understood as exact statistical mean values of the primary observables whose natural estimators \bar{a}_α are defined by the *replica-weighted mean*

$$\bar{a}_\alpha \equiv \frac{1}{N} \sum_{r=1}^R N_r \bar{a}_\alpha^{i,r}, \quad \bar{a}_\alpha^r = \frac{1}{N_r} \sum_{i=1}^{N_r} a_\alpha^{i,r}, \quad N = \sum_{r=1}^R N_r. \quad (\text{D.8})$$

The covariance matrix associated to these estimators is given by

$$\text{Cov}(\bar{a}_\alpha, \bar{a}_\beta) = \frac{1}{N} C_{\alpha\beta} \times \{1 + \mathcal{O}(\tau R/N)\}, \quad C_{\alpha\beta} = \sum_{t=-\infty}^{\infty} \Gamma_{\alpha\beta}(t), \quad (\text{D.9})$$

i.e. dominated by the asymptotic part of the autocorrelation function, $C_{\alpha\beta}$. The autocorrelation time scale τ is characteristic to the measured observable because it gives its exponential decay rate through $\Gamma_{\alpha\beta}(t) \stackrel{t \rightarrow \infty}{\sim} e^{-|t|/\tau}$. Our best estimator for the exact mean values A_α is clearly

$$\bar{f} = f(\bar{a}_\alpha). \quad (\text{D.10})$$

This can be used to get the corresponding estimator for the autocorrelation function,

$$\bar{\Gamma}_{\alpha\beta}(t) = \frac{1}{N - Rt} \sum_{r=1}^R \left[\sum_{i=1}^{N_r-t} (a_{\alpha}^{i,r} - \bar{a}_{\alpha}) (a_{\beta}^{i+t,r} - \bar{a}_{\beta}) \right] \quad (\text{D.11})$$

and to define a *projected autocorrelation function* by

$$\bar{\Gamma}_f(t) = \sum_{\alpha\beta} \bar{f}_{\alpha} \bar{\Gamma}_{\alpha\beta}(t) \bar{f}_{\beta}, \quad \bar{f}_{\alpha} = \frac{\partial \bar{f}}{\partial \bar{a}_{\alpha}}. \quad (\text{D.12})$$

The derivative $\partial_{\alpha} \bar{f}$ is meant to be taken at \bar{a}_{α} . Note that $\bar{\Gamma}_f(0)$ is the variance of f disregarding autocorrelations. Now the variance $\bar{\sigma}_f^2$ of any observable defined by the mapping f of primary observables is given in terms of the *integrated autocorrelation time* τ_{int} ,

$$\bar{\sigma}_f^2 = \frac{2\tau_{\text{int},f}}{N} \bar{\Gamma}_f(0), \quad (\text{D.13})$$

$$\tau_{\text{int},f} = \frac{1}{2\bar{\Gamma}_f(0)} \sum_{t=-\infty}^{\infty} \sum_{\alpha\beta} \bar{f}_{\alpha} \bar{\Gamma}_{\alpha\beta}(t) \bar{f}_{\beta} = \frac{1}{2\bar{\Gamma}_f(0)} \left[\bar{\Gamma}_f(0) + 2 \sum_{t=1}^{\infty} \bar{\Gamma}_f(t) \right]. \quad (\text{D.14})$$

We need to introduce a cutoff on the time summation because our statistic is finite and more importantly the ratio $\bar{\Gamma}_f(t)/\bar{\Gamma}_f(0)$ is noise dominated for $|t| \gg \tau$. On the other hand, a cutoff that is too small $\sim \tau$ would lead to a large systematic error. Thus a summation window W is introduced that has to be balanced between these two extremes. This is done by an automatic windowing procedure. A software package for this kind of error analysis is provided by the authors of [132] where further details can be found. We finally quote the optimal *variance estimate*,

$$\bar{\sigma}_f^2 = \frac{1}{N} \left[\bar{\Gamma}_f(0) + 2 \sum_{t=1}^W \bar{\Gamma}_f(t) \right]. \quad (\text{D.15})$$

E Tables for matching QCD and HQET

In this section I list various perturbative coefficients that are used in computations described in the main text. Due to dimensional regularisation the Riemann zeta function $\zeta(s)$ appears in higher loop computations in all $\overline{\text{MS}}$ -renormalization schemes. The following abbreviations are used:

$$\zeta_2 \equiv \zeta(2) = \pi^2/6, \quad I = \pi^2 \ln(2) - 3\zeta_3/2, \quad (\text{E.1})$$

$$\zeta_3 \equiv \zeta(3) \approx 1.2020569032. \quad (\text{E.2})$$

Beta function

There is no unique convention for coefficients of perturbative expansions as already the defining equations for anomalous dimensions can differ. In [20] for instance the β -function is defined by

$$\frac{\partial a_s}{\partial \ln \mu^2} \equiv \beta(a_s), \quad \beta(a_s) = -a_s^2(\beta_0 + \beta_1 a_s + \beta_2 a_s^2 + \beta_3 a_s^3) + \mathcal{O}(a_s^6), \quad (\text{E.3})$$

in which $a_s = \alpha_s/4\pi = \bar{g}^2/(4\pi)^2$ and $\bar{g} = \bar{g}(\mu^2)$ is the renormalized strong coupling constant of the QCD Lagrangian at renormalization point μ in the $\overline{\text{MS}}$ -scheme. From this definition, easy algebra for arbitrary order n ,

$$\begin{aligned} \frac{\partial a_s}{\partial \ln \mu^2} \equiv \beta(a_s) &\Leftrightarrow \frac{1}{(4\pi)^2} \frac{\partial \bar{g}^2(\mu^2)}{\partial \ln \mu^2} = -\frac{\bar{g}^4}{(4\pi)^4} \cdot \sum_{i=0}^n \frac{\beta_i}{(4\pi)^{2i}} \bar{g}^{2i} \\ &\Downarrow \\ \mu \frac{\partial \bar{g}(\mu)}{\partial \mu} \equiv \beta(\bar{g}) &\Leftrightarrow \mu \frac{\partial \bar{g}(\mu)}{\partial \mu} = -\bar{g}^3 \cdot \sum_{i=0}^n \frac{\beta_i}{(4\pi)^{2(i+1)}} \bar{g}^{2i} \end{aligned} \quad (\text{E.4})$$

shows the relation to the coefficients b_i which are used in this work:

$$b_i = \beta_i/(4\pi)^{2(i+1)}, \quad i \in \{1, \dots, n\}. \quad (\text{E.5})$$

See for instance section 1 or appendix A. The perturbative coefficients of the $\overline{\text{MS}}$ β -function are known up to four loops [20]. For generic group $\text{SU}(N)$ they read

$$\beta_0 = \frac{11}{3}C_A - \frac{4}{3}T_F N_f \quad (\text{E.6a})$$

$$\beta_1 = \frac{34}{3}C_A^2 - \left(4C_F + \frac{20}{3}C_A\right)T_F N_f \quad (\text{E.6b})$$

$$\beta_2 = \frac{2857}{54}C_A^3 + \left(2C_F^2 - \frac{205}{9}C_F C_A - \frac{1415}{27}C_A^2\right)T_F N_f + \left(\frac{44}{9}C_F + \frac{158}{27}C_A\right)T_F^2 N_f^2 \quad (\text{E.6c})$$

N_f dependence	
b_0	$(11 - 2N_f/3)/(4\pi)^2$
b_1	$(102 - 38N_f/3)/(4\pi)^4$
d_0	$8/(4\pi)^2$
$\gamma_0^{\text{PS}} = \gamma_0^{\text{V}}$	$-1/(4\pi)^2$
γ_0^{spin}	$-2/(4\pi)^2$

Table 27: Summary of scheme independent coefficients of perturbation theory, where the b_i , d_i , γ_i belong to the expansion of the coupling, mass and axial current respectively.

$\overline{\text{MS}}$	N_f dependence
b_2	$(22.3203 - 4.36892N_f + 0.0940394N_f^2)/(4\pi)^3$
b_3	$(114.23 - 27.1339N_f + 1.58238N_f^2 + 0.0058567N_f^3)/(4\pi)^3$
d_1	$2(4.20833 - 0.138889N_f)/(4\pi^2)$
d_2	$2(19.5156 - 2.28412N_f - 0.0270062N_f^2)/(4\pi^2)^2$
d_3	$2(98.9434 - 19.1075N_f + 0.276163N_f^2 + 0.00579322N_f^3)/(4\pi^2)^3$
γ_1^{PS}	$([-254/9 - 56\pi^2/27] + (20/9)N_f)/(4\pi)^4$
γ_2^{PS}	$64(-12.941 + 1.55406N_f + 0.0270062N_f^2)/(4\pi)^6$

Table 28: Some scheme dependent perturbative coefficients in the $\overline{\text{MS}}$ scheme after conversion with respect to our conventions. The $\{b_i\}$ are taken from [20] and $\{d_i\}$ from [131]. To each loop order in HQET one has $\gamma_n^{\text{PS}} = \gamma_n^{\text{V}}$ in the $\overline{\text{MS}}$ scheme. Values are taken from [126].

i	$k_i \cdot (4\pi^2)^i$
1	$4/3$
2	$-1.0414N_f + 13.4434$
3	$0.6527N_f^2 - 26.655N_f + 190.595$

Table 29: Perturbative factors k_i which originates from a reparametrisation of $m_Q \rightarrow \bar{m}_*$, see eq. (4.39). Values are taken from [128]. Again, N_f is the number of light quark flavours and set to two.

$$\begin{aligned}
 \beta_3 = & \left(\frac{150653}{486} - \frac{44}{9}\zeta_3 \right) C_A^4 + \left[\left(-\frac{39143}{81} + \frac{136}{3}\zeta_3 \right) C_A^3 + \left(\frac{7073}{243} - \frac{656}{9}\zeta_3 \right) C_A^2 C_F \right. \\
 & + \left. \left(-\frac{4204}{27} + \frac{352}{9}\zeta_3 \right) C_A C_F^2 + 46C_F^3 \right] T_F N_f + \left[\left(\frac{7930}{81} + \frac{224}{9}\zeta_3 \right) C_A^2 \right. \\
 & + \left. \left(\frac{17152}{243} + \frac{448}{9}\zeta_3 \right) C_A C_F + \left(\frac{1352}{27} - \frac{704}{9}\zeta_3 \right) C_F^2 \right] T_F^2 N_f^2 + \left[\frac{424}{243} C_A \right. \\
 & + \left. \frac{1232}{243} C_F \right] T_F^3 N_f^3 + \frac{N^2(N^2+36)}{24} \left(-\frac{80}{9} + \frac{704}{3}\zeta_3 \right) + \frac{N(N^2+6)}{48} \left(\frac{512}{9} - \frac{1664}{3}\zeta_3 \right) N_f \\
 & + \frac{N^4-6N^2+18}{96N^2} \left(-\frac{704}{9} + \frac{512}{3}\zeta_3 \right) N_f^2. \tag{E.6d}
 \end{aligned}$$

For colour gauge group SU(3) with $C_F = 4/3$, $C_A = 3$ and $T_F = 1/2$ the values of the non-universal coefficients are summarized in table 28.

Also the mass anomalous dimension was perturbatively computed in the $\overline{\text{MS}}$ scheme [131] using a slightly different convention for the renormalization group equation. Thus the extra factors in front of the brackets in table 28.

Table 30: Coefficients c_i relevant for the matching of the effective theory operators renormalized in the $\overline{\text{MS}}$ scheme to the physical ones in QCD, from [124].

X	Γ	n	η	$c_1^X \cdot (4\pi^2)$	$c_2^X \cdot (4\pi^2)^2$
PS		1	+1	-2/3	-4.198369830 + 0.4373501223 N_f
V		1	-1	-4/3	-11.50298789 + 0.7892019741 N_f
P		0	-1	+2/3	+10.91923705 - 0.6040167889 N_f
mag		2	± 1	-4/3	-16.19329152 + 1.1301514380 N_f

Heavy-Light currents

The matching of heavy-light currents between QCD and HQET was described in [123, 124, 125]. A general expression for the two-loop matching coefficients of heavy-light currents is given by

$$C_\Gamma(m, m) = 1 + C_F \left(\frac{\overline{g}(m)}{4\pi} \right)^2 [3(n-2)^2 + (2-\eta)(n-2) - 4] + C_F \left(\frac{\overline{g}(m)}{4\pi} \right)^4 \left[C_F a_F + C_A a_A + T_F a_h + T_F N_f a_1 + \sum_{i=1}^{N_f} \Delta(m_i/m) \right]. \quad (\text{E.7})$$

The two-loop contributions are

$$a_F = \frac{1}{24} (317 - 80 \zeta_2) (n-2)^4 + 11(n-2)^3 - \frac{11}{2} \eta (n-2)^3 + \frac{1}{6} (-253 + 288 \zeta_2 - 32 I) (n-2)^2 - 2\eta (n-2)^2 - 20(n-2) + \frac{1}{3} (32 - 64 \zeta_2 + 8) \eta (n-2) + \frac{689}{16} - 81 \zeta_2 - 8 \zeta_3 + 12 I \quad (\text{E.8})$$

$$a_A = \frac{1}{12} (-43 + 16 \zeta_2) (n-2)^4 - 2(n-2)^3 + \eta (n-2)^3 + \frac{1}{216} (9491 - 3744 \zeta_2 + 576 I) (n-2)^2 + \frac{143}{18} (n-2) + \frac{1}{18} (-281 + 144 \zeta_2 - 24 I) \eta (n-2) - \frac{29017}{432} + 29 \zeta_2 + 2 \zeta_3 - 6 I \quad (\text{E.9})$$

$$a_1 = \frac{1}{54} (-445 - 144 \zeta_2) (n-2)^2 - \frac{2}{9} (n-2) + \frac{38}{9} \eta (n-2) + \frac{1745}{108} + \frac{20}{3} \zeta_2 \quad (\text{E.10})$$

$$a_h = \frac{59}{54} (n-2)^2 - \frac{2}{9} (n-2) + \frac{1}{9} (-82 + 48 \zeta_2) \eta (n-2) + \frac{809}{27} - \frac{56}{3} \zeta_2 \quad (\text{E.11})$$

in a notation taken from [154]. The Γ structure of the current is represented by the integer parameters n and η and the ones used to get the matching functions for the currents under consideration are listed together with the resulting one- and two-loop coefficients in table 30.

Equation (E.7) is separated in parts of quenched - a_F, a_A - and unquenched, light-quark a_1 and heavy-quark a_h , contributions. Because the heavy quark is treated as a quenched quark in the simulations one has to neglect the term proportional to a_h . Furthermore, the finite mass correction terms, $\Delta(m_i/m)$, do not contribute as the mass of the two light quark flavours is set to zero and $\Delta(0) = 0$ holds.

F Further figures & tables

L/a	z	$L\Gamma_{\text{av}}$	R_{spin}	Y_{PS}	Y_{V}	$R_{\text{PS/V}}$	$R_{\text{PS/P}}$
20	4	+3.866(25)	+0.01700(54)	-0.5228(20)	-0.5813(11)	+0.9305(19)	+0.8694(35)
	6	+5.204(24)	+0.01323(41)	-0.5447(18)	-0.5875(11)	+0.9520(16)	+0.9248(30)
	7	+5.872(24)	+0.01186(37)	-0.5537(17)	-0.5907(11)	+0.9599(15)	+0.9481(28)
	9	+7.221(23)	+0.00975(30)	-0.5691(15)	-0.5974(11)	+0.9715(13)	+0.9886(25)
	11	+8.614(22)	+0.00816(25)	-0.5827(14)	-0.6047(11)	+0.9795(11)	+1.0237(22)
	13	+10.106(22)	+0.00688(21)	-0.5956(14)	-0.6130(11)	+0.9852(9)	+1.0557(19)
	15	+11.822(22)	+0.00576(18)	-0.6095(13)	-0.6233(12)	+0.9892(8)	+1.0871(17)
24	4	+3.904(25)	+0.01877(52)	-0.5201(20)	-0.5811(12)	+0.9291(16)	+0.8668(34)
	6	+5.235(24)	+0.01472(40)	-0.5431(17)	-0.5878(11)	+0.9516(13)	+0.9239(30)
	7	+5.894(24)	+0.01326(36)	-0.5524(16)	-0.5910(10)	+0.9598(12)	+0.9478(28)
	9	+7.208(23)	+0.01101(30)	-0.5681(15)	-0.5975(10)	+0.9721(11)	+0.9891(25)
	11	+8.532(23)	+0.00935(25)	-0.5813(14)	-0.6040(10)	+0.9806(9)	+1.0241(22)
	13	+9.889(22)	+0.00804(21)	-0.5930(13)	-0.6107(10)	+0.9867(8)	+1.0548(20)
	15	+11.313(22)	+0.00697(18)	-0.6040(12)	-0.6180(10)	+0.9911(7)	+1.0829(18)
	18	+13.725(22)	+0.00560(15)	-0.6209(12)	-0.6307(11)	+0.9955(6)	+1.1231(16)
	21	+17.478(21)	+0.00416(11)	-0.6461(12)	-0.6527(12)	+0.9981(5)	+1.1729(13)
32	4	+3.953(33)	+0.01989(69)	-0.5188(24)	-0.5821(12)	+0.9274(24)	+0.8599(49)
	6	+5.272(32)	+0.01571(55)	-0.5425(20)	-0.5891(11)	+0.9503(19)	+0.9164(43)
	7	+5.921(31)	+0.01419(50)	-0.5520(18)	-0.5923(11)	+0.9588(18)	+0.9400(40)
	9	+7.206(30)	+0.01187(42)	-0.5679(16)	-0.5984(11)	+0.9717(15)	+0.9805(35)
	11	+8.480(30)	+0.01016(36)	-0.5808(14)	-0.6042(11)	+0.9809(13)	+1.0143(31)
	13	+9.754(29)	+0.00885(31)	-0.5917(13)	-0.6098(12)	+0.9875(11)	+1.0434(28)
	15	+11.039(29)	+0.00779(28)	-0.6013(13)	-0.6154(12)	+0.9925(10)	+1.0691(25)
	18	+13.020(28)	+0.00653(23)	-0.6142(13)	-0.6237(13)	+0.9977(8)	+1.1031(22)
	21	+15.123(27)	+0.00553(20)	-0.6262(13)	-0.6325(13)	+1.0011(7)	+1.1336(19)
40	4	+3.948(28)	+0.02025(67)	-0.5185(25)	-0.5820(15)	+0.9277(29)	+0.8567(41)
	6	+5.265(27)	+0.01590(53)	-0.5423(22)	-0.5887(15)	+0.9510(23)	+0.9133(35)
	7	+5.914(26)	+0.01434(47)	-0.5519(21)	-0.5918(15)	+0.9596(21)	+0.9368(33)
	9	+7.193(25)	+0.01196(39)	-0.5676(19)	-0.5976(15)	+0.9728(18)	+0.9770(29)
	11	+8.455(25)	+0.01022(34)	-0.5803(19)	-0.6030(16)	+0.9822(15)	+1.0104(26)
	13	+9.709(25)	+0.00890(29)	-0.5908(18)	-0.6081(16)	+0.9891(14)	+1.0388(23)
	15	+10.960(24)	+0.00786(26)	-0.5999(18)	-0.6129(17)	+0.9942(12)	+1.0636(21)
	18	+12.848(24)	+0.00664(22)	-0.6115(18)	-0.6199(18)	+0.9998(10)	+1.0959(19)
	21	+14.776(23)	+0.00570(19)	-0.6218(19)	-0.6267(18)	+1.0035(9)	+1.1239(16)

Table 31: Non-perturbative results for QCD test observables in volume L_1 at $\theta = 0$.

θ	z	$L\Gamma_{av}$	R_{spin}	Y_{PS}	Y_V	$R_{PS/V}$	$R_{PS/P}$
20	4	+4.049(20)	+0.01490(50)	-0.4852(14)	-0.5541(9)	+0.9021(17)	+0.8322(27)
	6	+5.388(19)	+0.01169(39)	-0.5107(12)	-0.5614(8)	+0.9311(14)	+0.8952(24)
	7	+6.056(19)	+0.01051(35)	-0.5208(11)	-0.5649(8)	+0.9414(12)	+0.9212(22)
	9	+7.403(19)	+0.00868(28)	-0.5377(10)	-0.5719(9)	+0.9568(10)	+0.9658(20)
	11	+8.795(18)	+0.00730(24)	-0.5523(10)	-0.5792(9)	+0.9675(9)	+1.0039(17)
	13	+10.287(18)	+0.00618(20)	-0.5659(10)	-0.5875(9)	+0.9751(8)	+1.0383(16)
	15	+12.002(18)	+0.00519(17)	-0.5801(10)	-0.5977(10)	+0.9808(7)	+1.0719(14)
24	4	+4.106(26)	+0.01649(48)	-0.4820(18)	-0.5544(10)	+0.8985(19)	+0.8263(35)
	6	+5.436(24)	+0.01305(38)	-0.5088(15)	-0.5622(10)	+0.9290(15)	+0.8916(30)
	7	+6.094(24)	+0.01179(34)	-0.5193(14)	-0.5657(10)	+0.9399(14)	+0.9184(28)
	9	+7.407(23)	+0.00983(28)	-0.5367(13)	-0.5724(10)	+0.9562(11)	+0.9641(25)
	11	+8.729(23)	+0.00837(24)	-0.5509(12)	-0.5790(10)	+0.9675(10)	+1.0023(22)
	13	+10.084(22)	+0.00722(20)	-0.5633(11)	-0.5858(10)	+0.9756(9)	+1.0356(20)
	15	+11.508(22)	+0.00626(18)	-0.5748(11)	-0.5929(10)	+0.9816(8)	+1.0656(18)
	18	+13.917(21)	+0.00505(14)	-0.5922(11)	-0.6055(10)	+0.9879(6)	+1.1084(15)
	21	+17.668(21)	+0.00376(11)	-0.6175(12)	-0.6269(11)	+0.9924(5)	+1.1610(12)
32	4	+4.140(28)	+0.01772(69)	-0.4826(16)	-0.5565(10)	+0.8985(18)	+0.8219(37)
	6	+5.459(27)	+0.01412(56)	-0.5098(14)	-0.5645(10)	+0.9290(15)	+0.8861(33)
	7	+6.108(27)	+0.01280(51)	-0.5205(13)	-0.5680(10)	+0.9401(14)	+0.9124(31)
	9	+7.392(26)	+0.01075(43)	-0.5378(11)	-0.5744(11)	+0.9567(12)	+0.9570(28)
	11	+8.665(25)	+0.00922(37)	-0.5517(11)	-0.5804(11)	+0.9683(10)	+0.9937(25)
	13	+9.938(25)	+0.00804(32)	-0.5633(11)	-0.5860(12)	+0.9763(9)	+1.0251(23)
	15	+11.222(25)	+0.00709(28)	-0.5734(11)	-0.5915(13)	+0.9832(8)	+1.0525(21)
	18	+13.201(24)	+0.00595(24)	-0.5867(12)	-0.5998(14)	+0.9900(7)	+1.0886(18)
	21	+15.303(24)	+0.00504(20)	-0.5991(12)	-0.6085(14)	+0.9946(6)	+1.1209(16)
40	4	+4.151(27)	+0.01771(52)	-0.4801(20)	-0.5550(13)	+0.8963(30)	+0.8165(40)
	6	+5.467(26)	+0.01401(42)	-0.5077(17)	-0.5627(13)	+0.9278(24)	+0.8814(34)
	7	+6.114(25)	+0.01267(38)	-0.5183(16)	-0.5660(13)	+0.9392(22)	+0.9078(31)
	9	+7.391(24)	+0.01060(32)	-0.5357(15)	-0.5721(13)	+0.9564(18)	+0.9523(27)
	11	+8.652(24)	+0.00909(27)	-0.5493(15)	-0.5776(14)	+0.9685(16)	+0.9887(24)
	13	+9.903(23)	+0.00793(24)	-0.5605(15)	-0.5827(14)	+0.9773(14)	+1.0195(22)
	15	+11.153(23)	+0.00701(22)	-0.5700(15)	-0.5875(15)	+0.9840(12)	+1.0461(20)
	18	+13.040(23)	+0.00593(19)	-0.5822(15)	-0.5944(15)	+0.9912(10)	+1.0805(17)
	21	+14.967(22)	+0.00510(16)	-0.5927(16)	-0.6011(16)	+0.9962(9)	+1.1102(15)

Table 32: Non-perturbative results for QCD test observables in volume L_1 at $\theta = 0.5$.

θ	z	$L\Gamma_{av}$	R_{spin}	Y_{PS}	Y_V	$R_{PS/V}$	$R_{PS/P}$
20	4	+4.720(23)	+0.01128(37)	-0.3992(10)	-0.5020(10)	+0.8134(28)	+0.7176(34)
	6	+6.037(21)	+0.00900(30)	-0.4350(8)	-0.5117(9)	+0.8655(20)	+0.8024(27)
	7	+6.696(20)	+0.00815(27)	-0.4485(8)	-0.5158(9)	+0.8837(18)	+0.8365(25)
	9	+8.029(19)	+0.00679(22)	-0.4703(7)	-0.5235(9)	+0.9107(14)	+0.8936(21)
	11	+9.410(18)	+0.00575(19)	-0.4881(7)	-0.5312(9)	+0.9296(11)	+0.9411(18)
	13	+10.893(18)	+0.00490(16)	-0.5041(7)	-0.5395(9)	+0.9435(9)	+0.9830(16)
	15	+12.600(17)	+0.00414(14)	-0.5201(8)	-0.5494(9)	+0.9545(8)	+1.0234(14)
24	4	+4.783(32)	+0.01245(37)	-0.3971(15)	-0.5032(10)	+0.8092(32)	+0.7112(42)
	6	+6.092(29)	+0.01001(30)	-0.4342(12)	-0.5133(9)	+0.8629(24)	+0.7982(34)
	7	+6.741(28)	+0.00909(27)	-0.4480(11)	-0.5175(9)	+0.8817(21)	+0.8329(31)
	9	+8.040(27)	+0.00764(23)	-0.4702(10)	-0.5249(8)	+0.9095(16)	+0.8908(27)
	11	+9.351(25)	+0.00654(19)	-0.4876(9)	-0.5319(9)	+0.9288(14)	+0.9380(24)
	13	+10.698(25)	+0.00566(17)	-0.5022(9)	-0.5387(9)	+0.9429(11)	+0.9783(21)
	15	+12.114(24)	+0.00493(15)	-0.5154(9)	-0.5458(9)	+0.9536(10)	+1.0143(19)
	18	+14.515(23)	+0.00399(12)	-0.5346(10)	-0.5581(10)	+0.9656(8)	+1.0646(16)
	21	+18.257(22)	+0.00298(9)	-0.5612(10)	-0.5786(11)	+0.9757(6)	+1.1255(13)
32	4	+4.781(36)	+0.01352(67)	-0.3997(15)	-0.5049(11)	+0.8134(40)	+0.7125(47)
	6	+6.083(32)	+0.01096(54)	-0.4366(12)	-0.5155(11)	+0.8658(29)	+0.7970(38)
	7	+6.725(31)	+0.00998(49)	-0.4504(11)	-0.5197(11)	+0.8842(25)	+0.8308(34)
	9	+7.998(29)	+0.00844(41)	-0.4723(11)	-0.5270(11)	+0.9116(20)	+0.8867(30)
	11	+9.262(28)	+0.00728(35)	-0.4892(11)	-0.5333(12)	+0.9308(16)	+0.9318(26)
	13	+10.528(27)	+0.00637(31)	-0.5030(11)	-0.5392(13)	+0.9447(14)	+0.9695(23)
	15	+11.806(26)	+0.00563(27)	-0.5146(12)	-0.5448(14)	+0.9553(12)	+1.0022(21)
	18	+13.779(26)	+0.00474(23)	-0.5297(13)	-0.5530(15)	+0.9669(10)	+1.0445(18)
	21	+15.875(25)	+0.00402(19)	-0.5433(14)	-0.5616(16)	+0.9752(8)	+1.0819(16)
40	4	+4.850(37)	+0.01308(37)	-0.3951(16)	-0.5045(14)	+0.8039(40)	+0.6997(49)
	6	+6.140(34)	+0.01052(31)	-0.4329(13)	-0.5144(14)	+0.8594(31)	+0.7867(41)
	7	+6.777(33)	+0.00956(28)	-0.4469(12)	-0.5183(14)	+0.8789(28)	+0.8211(38)
	9	+8.039(31)	+0.00806(24)	-0.4689(11)	-0.5250(14)	+0.9077(23)	+0.8777(32)
	11	+9.289(30)	+0.00695(21)	-0.4857(11)	-0.5308(14)	+0.9279(19)	+0.9230(29)
	13	+10.532(29)	+0.00609(19)	-0.4991(12)	-0.5360(15)	+0.9426(16)	+0.9605(26)
	15	+11.774(28)	+0.00540(17)	-0.5102(12)	-0.5408(15)	+0.9537(14)	+0.9926(23)
	18	+13.654(27)	+0.00458(15)	-0.5242(13)	-0.5477(16)	+0.9659(12)	+1.0334(20)
	21	+15.574(27)	+0.00395(13)	-0.5359(14)	-0.5542(16)	+0.9747(10)	+1.0682(18)

Table 33: Non-perturbative results for QCD test observables in volume L_1 at $\theta = 1$.

References

The features of electronic publishing via pdf files makes it possible to integrate hyperlinks to the world wide web. The following list of references provides such hyperlinks to one of the mirrors of the e-print archive, namely <http://lanl.arxiv.org>. Beside various other so-called *open access* archives that are free available to everyone, the arXiv started in 1991 by Paul Ginsparg, was the very first of them.

But not all papers are available through arXiv. In those cases (and if possible) I provide a link by use of the Digital Object Identifier (DOI) system which assigns a time independent number to the original paper wherever it may be stored in the web. Unfortunately the access to some of those papers is restricted by the publisher.

- [1] **Particle Data Group** Collaboration, C. Amsler *et. al.*, *Review of particle physics*, Phys. Lett. **B667** (2008) 1.
- [2] K. G. Wilson, *Confinement of quarks*, Phys. Rev. **D10** (1974) 2445–2459.
- [3] J. H. Christenson, J. W. Cronin, V. L. Fitch and R. Turlay, *Evidence for the 2π Decay of the K_2^0 Meson*, Phys. Rev. Lett. **13** (1964) 138–140.
- [4] M. Kobayashi and T. Maskawa, *CP Violation in the Renormalizable Theory of Weak Interaction*, Prog. Theor. Phys. **49** (1973) 652–657.
- [5] N. Cabibbo, *Unitary Symmetry and Leptonic Decays*, Phys. Rev. Lett. **10** (1963) 531–533.
- [6] **SuperKEKB Physics Working Group** Collaboration, A. G. Akeroyd *et. al.*, *Physics at super B factory*, hep-ex/0406071.
- [7] E. Eichten, *Heavy quarks on the lattice*, Nucl. Phys. Proc. Suppl. **4** (1988) 170. [FERMILAB-PUB-87-158-T].
- [8] E. Eichten and B. Hill, *An effective field theory for the calculation of matrix elements involving heavy quarks*, Phys. Lett. **B234** (1990) 511. [FERMILAB-PUB-89-184-T].
- [9] E. Eichten and B. Hill, *Static effective field theory: $1/m$ corrections*, Phys. Lett. **B243** (1990) 427. [FERMILAB-PUB-90-054-T].
- [10] L. Maiani, G. Martinelli and C. T. Sachrajda, *Nonperturbative subtractions in the heavy quark effective field theory*, Nucl. Phys. **B368** (1992) 281–292.
- [11] **ALPHA** Collaboration, J. Heitger and R. Sommer, *Non-perturbative heavy quark effective theory*, J. High Energy Phys. **02** (2004) 022 [hep-lat/0310035].
- [12] **ALPHA** Collaboration, M. Della Morte, N. Garron, M. Papinutto and R. Sommer, *Heavy quark effective theory computation of the mass of the bottom quark*, J. High Energy Phys. **01** (2007) 007 [hep-ph/0609294].
- [13] **ALPHA** Collaboration, B. Blossier, M. Della Morte, N. Garron and R. Sommer, *Heavy-light decay constant at the $1/m$ order of HQET*, PoS **LAT2007** (2007) 245 [arXiv:0710.1553].
- [14] **ALPHA** Collaboration, J. Heitger, A. Jüttner, R. Sommer and J. Wennekers, *Non-perturbative tests of heavy quark effective theory*, J. High Energy Phys. **11** (2004) 048 [hep-ph/0407227].
- [15] G. 't Hooft, *Dimensional regularization and the renormalization group*, Nucl. Phys. **B61** (1973) 455–468.

- [16] D. J. Gross and F. Wilczek, *Ultraviolet Behavior of Non-Abelian Gauge Theories*, Phys. Rev. Lett. **30** (1973) 1343–1346.
- [17] H. D. Politzer, *Reliable Perturbative Results for Strong Interactions?*, Phys. Rev. Lett. **30** (1973) 1346–1349.
- [18] W. E. Caswell, *Asymptotic Behavior of Nonabelian Gauge Theories to Two Loop Order*, Phys. Rev. Lett. **33** (1974) 244.
- [19] O. V. Tarasov, A. A. Vladimirov and A. Y. Zharkov, *The Gell-Mann-Low Function of QCD in the Three Loop Approximation*, Phys. Lett. **B93** (1980) 429–432.
- [20] T. van Ritbergen, J. A. M. Vermaseren and S. A. Larin, *The four-loop β -function in Quantum Chromodynamics*, Phys. Lett. **B400** (1997) 379 [hep-ph/9701390].
- [21] ALPHA Collaboration, S. Sint and P. Weisz, *The running quark mass in the SF scheme and its two-loop anomalous dimension*, Nucl. Phys **B545** (1999) 529 [hep-lat/9808013].
- [22] S. Sint and R. Sommer, *The running coupling from the QCD Schrödinger functional: A one-loop analysis*, Nucl. Phys. **B465** (1996) 71 [hep-lat/9508012].
- [23] J. Gasser and H. Leutwyler, *Quark Masses*, Phys. Rept. **87** (1982) 77.
- [24] J. Gasser and H. Leutwyler, *Chiral perturbation theory to one loop*, Ann. Phys. **158** (1984) 142.
- [25] J. Gasser and H. Leutwyler, *Chiral perturbation theory: Expansions in the mass of the strange quark*, Nucl. Phys. **B250** (1985) 465.
- [26] ALPHA Collaboration, M. Kurth and R. Sommer, *Renormalization and $O(a)$ improvement of the static axial current*, Nucl. Phys. **B597** (2001) 488 [hep-lat/0007002].
- [27] G. M. Prosperi, M. Raciti and C. Simolo, *On the running coupling constant in QCD*, Prog. Part. Nucl. Phys. **58** (2007) 387–438 [hep-ph/0607209].
- [28] ALPHA Collaboration, M. Della Morte, R. Frezzotti, J. Heitger, J. Rolf, R. Sommer and U. Wolff, *Computation of the strong coupling in QCD with two dynamical flavours*, Nucl. Phys. **B713** (2005) 378 [hep-lat/0411025].
- [29] ALPHA Collaboration, M. Della Morte, R. Hoffmann, F. Knechtli, J. Rolf, R. Sommer, I. Wetzorke and U. Wolff, *Non-perturbative quark mass renormalization in two-flavor QCD*, Nucl. Phys. **B729** (2005) 117 [hep-lat/0507035].
- [30] A. S. Wightman, *Quantum Field Theory in Terms of Vacuum Expectation Values*, Phys. Rev. **101** (1956) 860–866.
- [31] K. Osterwalder and R. Schrader, *Axioms for Euclidean Green's Functions*, Commun. Math. Phys. **31** (1973) 83–112. [euclid:1103858969].
- [32] K. Osterwalder and R. Schrader, *Axioms for Euclidean Green's Functions. 2*, Commun. Math. Phys. **42** (1975) 281. [euclid:1103899050].
- [33] M. Kac, *On Distributions of Certain Wiener Functionals*, Transactions of the American Mathematical Society **65** (1949), no. 1 1–13.
- [34] K. G. Wilson, *Quarks and Strings on a Lattice*, . New Phenomena In Subnuclear Physics. Part A. Proceedings of the First Half of the 1975 International School of Subnuclear Physics, Erice, Sicily, July 11 - August 1, 1975, ed. A. Zichichi, Plenum Press, New York, 1977, p. 69, CLNS-321.

- [35] M. Lüscher, *Construction of a Selfadjoint, Strictly Positive Transfer Matrix for Euclidean Lattice Gauge Theories*, Commun. Math. Phys. **54** (1977) 283. [euclid:1103900872].
- [36] T. Reisz, *Renormalization of Feynman integrals on the lattice*, Commun. Math. Phys. **117** (1988) 79. [euclid:1104161595].
- [37] T. Reisz, *Renormalization of lattice Feynman integrals with massless propagators*, Commun. Math. Phys. **117** (1988) 639. [euclid:1104161822].
- [38] J. E. Mandula, G. Zweig and J. Govaerts, *Representations of the rotation reflection symmetry group of the four-dimensional cubic lattice*, Nucl. Phys. **B228** (1983) 91.
- [39] M. Lüscher, *Volume dependence of the energy spectrum in massive quantum field theories. I: Stable particle states.*, Commun. Math. Phys. **104** (1986) 177–206. [euclid:1104114998].
- [40] M. Lüscher, *Volume dependence of the energy spectrum in massive quantum field theories. II. Scattering states*, Comm. Math. Phys. **105** (1986), no. 2 153–188. [euclid:1104115329].
- [41] L. H. Karsten and J. Smith, *Lattice fermions: Species doubling, chiral invariance and the triangle anomaly*, Nucl. Phys. **B183** (1981) 103–140.
- [42] H. B. Nielsen and M. Ninomiya, *A no-go theorem for regularizing chiral fermions*, Phys. Lett. **B105** (1981), no. 2-3 219–223.
- [43] H. B. Nielsen and M. Ninomiya, *Absence of neutrinos on a lattice : (I). Proof by homotopy theory*, Nucl. Phys. **B185** (1981) 20–40. [Erratum *ibid*: B195 (1982) 541].
- [44] H. B. Nielsen and M. Ninomiya, *Absence of neutrinos on a lattice : (II). Intuitive topological proof*, Nucl.Phys. **B193** (1981) 173–194.
- [45] M. Bochicchio and L. Maiani and G. Martinelli and G. C. Rossi and M. Testa, *Chiral symmetry on the lattice with Wilson fermions*, Nucl. Phys. **B262** (1985) 331.
- [46] K. Symanzik, *Continuum limit and improved action in lattice theories: 1. Principles and ϕ^4 theory*, Nucl. Phys. **B226** (1983) 187.
- [47] K. Symanzik, *Continuum limit and improved action in lattice theories: 2. $O(N)$ non-linear sigma model in perturbation theory*, Nucl. Phys. **B226** (1983) 205.
- [48] M. Lüscher, S. Sint, R. Sommer and P. Weisz, *Chiral symmetry and $O(a)$ improvement in lattice QCD*, Nucl. Phys. **B478** (1996) 365 [hep-lat/9605038].
- [49] B. Sheikholeslami and R. Wohlert, *Improved continuum limit lattice action for QCD with Wilson fermions*, Nucl. Phys. **B259** (1985) 572.
- [50] N. Cundy *et. al.*, *Non-perturbative improvement of stout-smearred three flavour clover fermions*, 0901.3302.
- [51] ALPHA Collaboration, M. Lüscher, S. Sint, R. Sommer, P. Weisz and U. Wolff, *Non-perturbative $O(a)$ improvement of lattice QCD*, Nucl. Phys. **B491** (1997) 323 [hep-lat/9609035].
- [52] ALPHA Collaboration, K. Jansen and R. Sommer, *$O(a)$ improvement of lattice QCD with two flavors of Wilson quarks*, Nucl. Phys. **B530** (1998) 185 [hep-lat/9803017].
- [53] ALPHA Collaboration, F. Tekin and R. Sommer. work in progress.

- [54] **ALPHA** Collaboration, S. Sint and P. Weisz, *Further results on $O(a)$ improved lattice QCD to one-loop order of perturbation theory*, Nucl. Phys. **B502** (1997) 251 [hep-lat/9704001].
- [55] S. Sint and P. Weisz, *Further one-loop results in $O(a)$ improved lattice QCD*, Nucl. Phys. Proc. Suppl. **63** (1998) 856 [hep-lat/9709096].
- [56] Y. Taniguchi and A. Ukawa, *Perturbative calculation of improvement coefficients to $O(g^2 a)$ for bilinear quark operators in lattice QCD*, Phys. Rev. **D58** (1998) 114503 [hep-lat/9806015].
- [57] Y. Taniguchi, S. Aoki, K.-i. Nagai and A. Ukawa, *One-loop renormalization factors and mixing coefficients of bilinear quark operators for improved gluon and quark actions*, Nucl. Phys. Proc. Suppl. **73** (1999) 912–914 [hep-lat/9809070].
- [58] **ALPHA** Collaboration, M. Della Morte, R. Hoffmann and R. Sommer, *Non-perturbative improvement of the axial current for dynamical Wilson fermions*, J. High Energy Phys. **03** (2005) 029 [hep-lat/0503003].
- [59] **ALPHA** Collaboration, M. Guagnelli and R. Sommer, *Non-perturbative $O(a)$ improvement of the vector current*, Nucl. Phys. Proc. Suppl. **63** (1998) 886 [hep-lat/9709088].
- [60] K. Symanzik, *Schrödinger representation and Casimir effect in renormalizable quantum field theory*, Nucl. Phys. **B190** (1981) 1.
- [61] U. Wolff, *Anomalous scale transformations and functional integrals*, Nucl. Phys. **B265** (1986) 506.
- [62] U. Wolff, *Monte Carlo estimate of the background field coupling renormalization in the Heisenberg model*, Nucl. Phys. **B265** (1986) 537.
- [63] M. Lüscher, P. Weisz and U. Wolff, *A numerical method to compute the running coupling in asymptotically free theories*, Nucl. Phys. **B359** (1991) 221.
- [64] M. Lüscher, R. Narayanan, P. Weisz and U. Wolff, *The Schrödinger functional: A renormalizable probe for non-abelian gauge theories*, Nucl. Phys. **B384** (1992) 168 [hep-lat/9207009].
- [65] R. Sommer, *Non-perturbative QCD: renormalization, $O(a)$ -improvement and matching to Heavy Quark Effective Theory*, Lectures given at *ILFTN Workshop on “Perspectives in Lattice QCD”*, Nara, Japan, 31 October – 11 November 2005 [hep-lat/0611020].
- [66] A. A. Belavin, A. M. Polyakov, A. S. Shvarts and Y. S. Tyupkin, *Pseudoparticle solutions of the Yang-Mills equations*, Phys. Lett. **B59** (1975) 85–87.
- [67] B. S. DeWitt, *Quantum theory of gravity. II. The manifestly covariant theory*, Phys. Rev. **162** (1967) 1195–1239.
- [68] M. Lüscher, R. Sommer, P. Weisz and U. Wolff, *A precise determination of the running coupling in the $SU(3)$ Yang-Mills theory*, Nucl. Phys. **B413** (1994) 481 [hep-lat/9309005].
- [69] E. Bilgici *et al.*, *A new scheme for the running coupling constant in gauge theories using Wilson loops*, 0902.3768.
- [70] **ALPHA** Collaboration, A. Bode, P. Weisz and U. Wolff, *Two loop computation of the Schrödinger functional in lattice QCD*, Nucl. Phys. **B576** (2000) 517 [hep-lat/9911018]. Errata: *ibid.* **B600** (2001) 453, *ibid.* **B608** (2001) 481.

- [71] S. Sint, *On the Schrödinger functional in QCD*, Nucl. Phys. **B421** (1994) 135 [hep-lat/9312079].
- [72] S. Sint, *One-loop renormalization of the QCD Schrödinger functional*, Nucl. Phys. **B451** (1995) 416 [hep-lat/9504005].
- [73] G. M. de Divitiis, R. Petronzio and N. Tantalo, *On the discretization of physical momenta in lattice QCD*, Phys. Lett. **B595** (2004) 408–413 [hep-lat/0405002].
- [74] M. Lüscher and P. Weisz, *$O(a)$ improvement of the axial current in lattice QCD to one-loop order of perturbation theory*, Nucl. Phys. **B479** (1996) 429 [hep-lat/9606016].
- [75] ALPHA Collaboration, M. Guagnelli, J. Heitger, R. Sommer and H. Wittig, *Hadron masses and matrix elements from the QCD Schrödinger functional*, Nucl. Phys. **B560** (1999) 465 [hep-lat/9903040].
- [76] 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** (2009) 094 [0902.1265].
- [77] ALPHA Collaboration, M. Della Morte, R. Hoffmann, F. Knechtli, R. Sommer and U. Wolff, *Non-perturbative renormalization of the axial current with dynamical Wilson fermions*, J. High Energy Phys. **07** (2005) 007 [hep-lat/0505026].
- [78] M. Della Morte, R. Sommer and S. Takeda, *On cutoff effects in lattice QCD from short to long distances*, Phys. Lett. **B672** (2009) 407–412 [0807.1120].
- [79] ALPHA Collaboration, S. Capitani, M. Lüscher, R. Sommer and H. Wittig, *Non-perturbative quark mass renormalization in quenched lattice QCD*, Nucl. Phys. **B544** (1999) 669 [hep-lat/9810063].
- [80] B. Grinstein, *The static quark effective theory*, Nucl. Phys. **B339** (1990) 253.
- [81] J. G. Korner and G. Thompson, *The Heavy mass limit in field theory and the heavy quark effective theory*, Phys. Lett. **B264** (1991) 185–192.
- [82] M. Neubert, *Heavy quark symmetry*, Phys. Rept. **245** (1994) 259 [hep-ph/9306320].
- [83] M. E. Luke and A. V. Manohar, *Reparameterization invariance constraints on heavy particle effective field theories*, Phys. Lett. **B286** (1992) 348 [hep-ph/9205228].
- [84] Y.-Q. Chen, *On the reparameterization invariance in heavy quark effective theory*, Phys. Lett. **B317** (1993) 421 – 427.
- [85] R. Sundrum, *Reparameterization invariance to all orders in heavy quark effective theory*, Phys. Rev. **D57** (1998) 331 [hep-ph/9704256].
- [86] A. F. Falk and M. Neubert, *Second-order power corrections in the heavy quark effective theory. 1. Formalism and meson form-factors*, Phys. Rev. **D47** (1993) 2965 [hep-ph/9209268].
- [87] ALPHA Collaboration, M. Della Morte, A. Shindler and R. Sommer, *On lattice actions for static quarks*, J. High Energy Phys. **08** (2005) 051 [hep-lat/0506008].
- [88] K. G. Wilson, *Non-Lagrangian Models of Current Algebra*, Phys. Rev. **179** (1969) 1499–1512.
- [89] K. G. Wilson, *Operator-Product Expansions and Anomalous Dimensions in the Thirring Model*, Phys. Rev. **D2** (1970) 1473–1477.

- [90] M. Neubert, *Short distance expansion of heavy quark currents*, Phys. Rev. **D46** (1992) 2212–2227.
- [91] M. Neubert, *Short distance expansion of heavy-light currents at order $1/m_Q$* , Phys. Rev. **D49** (1994) 1542 [hep-ph/9308369].
- [92] M. Beneke and V. M. Braun, *Heavy quark effective theory beyond perturbation theory: Renormalons, the pole mass and the residual mass term*, Nucl. Phys. **B426** (1994) 301–343 [hep-ph/9402364].
- [93] I. I. Y. Bigi, M. A. Shifman, N. G. Uraltsev and A. I. Vainshtein, *The Pole mass of the heavy quark. Perturbation theory and beyond*, Phys. Rev. **D50** (1994) 2234–2246 [hep-ph/9402360].
- [94] K. G. Chetyrkin, *Quark mass anomalous dimension to $O(\alpha_s^4)$* , Phys. Lett. **B404** (1997) 161 [hep-ph/9703278].
- [95] N. Gray, D. J. Broadhurst, W. Grafe and K. Schilcher, *Three-loop relation of quark (modified) \overline{MS} and pole masses*, Z. Phys. **C48** (1990) 673.
- [96] L. Del Debbio, L. Giusti, M. Luscher, R. Petronzio and N. Tantalo, *QCD with light Wilson quarks on fine lattices (I): first experiences and physics results*, JHEP **02** (2007) 056 [hep-lat/0610059].
- [97] ALPHA Collaboration, J. Heitger and J. Wennekers, *Effective heavy-light meson energies in small-volume quenched QCD*, J. High Energy Phys. **02** (2004) 064 [hep-lat/0312016].
- [98] S. Duane, A. D. Kennedy, B. J. Pendleton and D. Roweth, *Hybrid Monte Carlo*, Phys. Lett. **B195** (1987) 216–222.
- [99] D. J. E. Callaway and A. Rahman, *Microcanonical Ensemble Formulation of Lattice Gauge Theory*, Phys. Rev. Lett. **49** (1982) 613.
- [100] D. J. E. Callaway and A. Rahman, *Lattice Gauge Theory in Microcanonical Ensemble*, Phys. Rev. **D28** (1983) 1506.
- [101] R. Frezzotti and K. Jansen, *The phmc algorithm for simulations of dynamical fermions. i: Description and properties*, Nucl. Phys. **B555** (1999) 395–431 [hep-lat/9808011].
- [102] A. D. Kennedy, *Algorithms for dynamical fermions*, hep-lat/0607038.
- [103] I. P. Omelyan, I. M. Mryglod and R. Folk, *Construction of high-order force-gradient algorithms for integration of motion in classical and quantum systems*, Phys. Rev. **E66** (2002) 026701.
- [104] J. C. Sexton and D. H. Weingarten, *Hamiltonian evolution for the hybrid Monte Carlo algorithm*, Nucl. Phys. **B380** (1992) 665–678.
- [105] N. Metropolis and S. Ulam, *The Monte Carlo Method*, Journal of the American Statistical Association **44** (1949) 335–341.
- [106] ALPHA Collaboration, M. Della Morte, F. Knechtli, J. Rolf, R. Sommer, I. Wetzorke and U. Wolff, *Simulating the Schrödinger functional with two pseudofermions*, Comput. Phys. Commun. **156** (2003) 62 [hep-lat/0307008].
- [107] ALPHA Collaboration, H. B. Meyer, H. Simma, R. Sommer, M. Della Morte, O. Witzel and U. Wolff, *Exploring the HMC trajectory-length dependence of autocorrelation times in lattice QCD*, Comput. Phys. Commun. **176** (2007) 91 [hep-lat/0606004].

- [108] K. Jansen and C. Liu, *Implementation of Symanzik's improvement program for simulations of dynamical Wilson fermions in lattice QCD*, Comput. Phys. Commun. **99** (1997) 221 [hep-lat/9603008].
- [109] M. Hasenbusch, *Speeding up the Hybrid-Monte-Carlo algorithm for dynamical fermions*, Phys. Lett. **B519** (2001) 177 [hep-lat/0107019].
- [110] M. Hasenbusch and K. Jansen, *Speeding up lattice QCD simulations with clover-improved Wilson fermions*, Nucl. Phys. **B659** (2003) 299 [hep-lat/0211042].
- [111] T. Kalkreuter and H. Simma, *An Accelerated conjugate gradient algorithm to compute low lying eigenvalues: A Study for the Dirac operator in SU(2) lattice QCD*, Comput. Phys. Commun. **93** (1996) 33–47 [hep-lat/9507023].
- [112] N. Cabibbo and E. Marinari, *A New Method for Updating SU(N) Matrices in Computer Simulations of Gauge Theories*, Phys. Lett. **B119** (1982) 387–390.
- [113] L. Del Debbio, L. Giusti, M. Lüscher, R. Petronzio and N. Tantalo, *Stability of lattice QCD simulations and the thermodynamic limit*, JHEP **02** (2006) 011 [hep-lat/0512021].
- [114] ALPHA Collaboration, M. Della Morte *et al.* work in progress.
- [115] ALPHA Collaboration, M. Guagnelli, R. Petronzio, J. Rolf, S. Sint, R. Sommer and U. Wolff, *Non-perturbative results for the coefficients b_m and $b_A - b_P$ in $O(a)$ improved lattice QCD*, Nucl. Phys. **B595** (2001) 44 [hep-lat/0009021].
- [116] T. Bhattacharya, R. Gupta, W. Lee, S. R. Sharpe and J. M. S. Wu, *Improved bilinears in lattice QCD with non-degenerate quarks*, Phys. Rev. **D73** (2006) 034504 [hep-lat/0511014].
- [117] G. M. de Divitiis and R. Petronzio, *Non-perturbative renormalization constants on the lattice from flavor non-singlet ward identities*, Phys. Lett. **B419** (1998) 311 [hep-lat/9710071].
- [118] G. von Hippel, R. Sommer, J. Heitger, S. Schaefer and N. Tantalo, *D_s physics from fine lattices*, PoS LATTICE2008 (2008) 227 [0810.0214].
- [119] J. Heitger and A. Jüttner, *Lattice cutoff effects for F_{D_s} with improved Wilson fermions - a final lesson from the quenched case*, JHEP **05** (2009) 101 [0812.2200].
- [120] ALPHA Collaboration, J. Heitger, D. Hesse and R. Sommer. work in progress.
- [121] K. Symanzik, *Some Topics in Quantum Field Theory*, Lecture Notes in physics **153** (1982) 47. Presented at 6th Int. Conf. on Mathematical Physics, Berlin, West Germany, Aug 11-21, 1981.
- [122] A. Billoire, E. Marinari and G. Parisi, *Computing the hadronic mass spectrum. Eight is better than one*, Phys. Lett. **B162** (1985) 160.
- [123] Broadhurst, David J. and Gray, N. and Schilcher, K., *Gauge invariant on-shell Z(2) in QED, QCD and the effective field theory of a static quark*, Z. Phys. **C52** (1991) 111–122.
- [124] D. J. Broadhurst and A. G. Grozin, *Matching QCD and HQET heavy-light currents at two loops and beyond*, Phys. Rev. **D52** (1995) 4082 [hep-ph/9410240].
- [125] A. G. Grozin, *Decoupling of heavy quark loops in light-light and heavy-light quark currents*, Phys. Lett. **B445** (1998) 165 [hep-ph/9810358].

- [126] K. G. Chetyrkin and A. G. Grozin, *Three-loop anomalous dimension of the heavy-light quark current in HQET*, Nucl. Phys. **B666** (2003) 289 [hep-ph/0303113].
- [127] ALPHA Collaboration, J. Heitger, M. Kurth and R. Sommer, *Non-perturbative renormalization of the static axial current in quenched QCD*, Nucl. Phys. **B669** (2003) 173 [hep-lat/0302019].
- [128] K. Melnikov and T. v. Ritbergen, *The three-loop relation between the \overline{MS} and the pole quark masses*, Phys. Lett. **B482** (2000) 99–108 [hep-ph/9912391].
- [129] Ji, X.-D., *Exact matching condition for matrix elements in lattice and \overline{MS} schemes*, MIT CTP-preprint no. 2447 (1995) [hep-lat/9506034].
- [130] A. D. Sokal Monte Carlo methods in Statistical Mechanics: Foundations and new algorithms, in *Cours de Troisième Cycle de la Physique en Suisse Romande.
- [131] J. A. M. Vermaseren, S. A. Larin and T. van Ritbergen, *The four-loop quark mass anomalous dimension and the invariant quark mass*, Phys. Lett. **B405** (1997) 327 [hep-ph/9703284].
- [132] ALPHA Collaboration, U. Wolff, *Monte Carlo errors with less errors*, Comput. Phys. Commun. **156** (2004) 143 [hep-lat/0306017].
- [133] M. Gell-Mann, *Symmetries of Baryons and Mesons*, Phys. Rev. **125** (1962) 1067.
- [134] W. Pauli, *Zur Quantenmechanik des magnetischen Elektrons*, Z. Phys. **43** (1927) 601–623.
- [135] M. Testa, *Some observations on broken symmetries*, JHEP **04** (1998) 002 [hep-th/9803147].
- [136] C. T. Sachrajda and G. Villadoro, *Twisted boundary conditions in lattice simulations*, Phys. Lett. **B609** (2005) 73–85 [hep-lat/0411033].
- [137] R. Tripiccion, *LGT simulations on APE machines*, Comp. Phys. Commun. **139** (2001) 55.
- [138] S. Cabasino, P. S. Paolucci and G. M. Todesco, *Dynamic parsers and evolving grammars*, SIGPLAN Not. **27** (1992), no. 11 39–48.
- [139] ALPHA Collaboration, M. Guagnelli and J. Heitger, *SSOR preconditioning in simulations of the QCD Schroedinger functional*, Comput. Phys. Commun. **130** (2000) 12–21 [hep-lat/9910024].
- [140] F. Bodin *et. al.*, *The apeNEXT project*, hep-lat/0306018.
- [141] J. A. Fisher, *Very Long Instruction Word architectures and the ELI-512*, ISCA '83: Proceedings of the 10th annual international symposium on Computer architecture (1983) 140–150.
- [142] CLS Collaboration. Coordinated Lattice Simulations;
<https://twiki.cern.ch/twiki/bin/view/CLS/WebHome>.

Textbooks

- [143] M. E. Peskin and D. V. Schroeder, *An Introduction to Quantum Field Theory*, Addison-Wesley (1995) 842pp.
- [144] S. Weinberg, *The Quantum Theory of Fields 3 volume set*, Cambridge University Press (2005) 1600pp.

- [145] T. P. Cheng and L. F. Li, *Gauge Theory of Elementary Particle Physics*, Oxford University Press (1985) 536pp.
- [146] E. Zeidler, *Basics in Mathematics and Physics*, Springer, Quantum Field Theory: Basics in Mathematics and Physics **1** (2009) 1020pp.
- [147] E. Zeidler, *Quantum Electrodynamics*, Springer, Quantum Field Theory: Basics in Mathematics and Physics **2** (2009) 1102pp.
- [148] R. A. Bertlmann, *Anomalies in Quantum Field Theory*, International Series of Monographs on Physics **91** (2000) 584pp.
- [149] J. Zinn-Justin, *Quantum Field Theory and Critical Phenomena*, International Series of Monographs on Physics **no.113** (2002) 1074pp.
- [150] I. Montvay and G. Münster, *Quantum Fields on a Lattice*, Cambridge Monographs on Mathematical Physics (1994) 508pp.
- [151] H. J. Rothe, *Lattice Gauge Theories – An Introduction*, World Scientific Lecture Notes in Physics, Vol. 74 (2005) 608pp.
- [152] T. DeGrand and C. DeTar, *Lattice methods for Quantum Chromodynamics*, World Scientific Publishing Company (2006) 364pp.
- [153] A. V. Manohar and M. B. Wise, *Heavy quark physics*, Camb. Monogr. Part. Phys. Nucl. Phys. Cosmol. **10** (2000) 1–191.
- [154] A. G. Grozin, *Heavy Quark Effective Theory*, Springer Tracts in Mod. Phys. **201** (2004) 213pp.

Publications

- [155] **ALPHA** Collaboration, M. Della Morte, P. Fritzsche and J. Heitger, *Non-perturbative renormalization of the static axial current in two-flavour QCD*, J. High Energy Phys. **02** (2007) 079 [hep-lat/0611036].
- [156] **ALPHA** Collaboration, M. Della Morte, P. Fritzsche, J. Heitger, H. B. Meyer, H. Simma and R. Sommer, *Towards a non-perturbative matching of HQET and QCD with dynamical light quarks*, PoS **LAT2007** (2007) 246 [0710.1188].
- [157] **ALPHA** Collaboration, M. Della Morte, P. Fritzsche, B. Leder, H. B. Meyer, H. Simma, R. Sommer, S. Takeda, O. Witzel and U. Wolff, *Preparing for $N_f = 2$ simulations at small lattice spacings*, PoS **LAT2007** (2007) 255 [0710.1263].
- [158] **ALPHA** Collaboration, M. Della Morte, P. Fritzsche, H. Meyer, H. Simma, R. Sommer, S. Takeda, O. Witzel and U. Wolff, *Scaling test of two-flavor $O(a)$ -improved lattice QCD*, JHEP **07** (2008) 037 [0804.3383].
- [159] **ALPHA** Collaboration, M. Della Morte, P. Fritzsche, J. Heitger and R. Sommer, *Non-perturbative quark mass dependence in the heavy-light sector of two-flavour QCD*, PoS **LAT2008** (2008) 226 [0810.3166].

Acknowledgments

First of all I want to express most of my hearty gratitude to my family, especially to my parents Gunter and Steffi Fritzsche who always encouraged me to do my job and life in balanced modality. I think that would not be the case if they were not there. The same holds for my brother Christopher and my grandparents. All of you made my life eventful. But there is more to say:

- This work would not be in existence without the help of Priv.-Doz. Dr. Jochen Heitger. He brought this interesting subject to me and thus allowed me to get a deeper understanding of the physics behind it and the tools we use than it was possible during my graduate studies. Jochen, you were always there if I had a physical question or other problems to solve. Thank you very much.
- Furthermore, I thank the Deutsche Forschungsgemeinschaft (German Research Foundation) for two years of financial support under grant HE 4517/2-1 as well as the European Community through EU Contract No. MRTN-CT-2006-035482, "FLAVIANet". Without their support it had not been possible for me to participate in a lot of conferences during the last years and finish this thesis abroad in Southampton, UK.
- Thank you Thomas and Michael for the enjoyable time during and off the work. After those many years it is time to sail away.
- Furthermore I want to thank M. Della Morte, R. Sommer and other members of the ALPHA collaboration for valuable support over the years and discussion about the physics we really do. I am glad that I was part of this journey.
- Special thanks go to Hubert Simma who told me everything I know about the apeNEXT computer and its optimization.
- I thank N. Tantalo for collaboration and the production of some gauge field configurations on apeNEXT machines in Rome.
- I am deeply grateful to D. Brömmel, J. Flynn and C.T. Sachrajda for their valuable support and patience in Southampton.
- Last but not least I appreciate DESY/NIC for providing computing time and resources and especially the IT team in Zeuthen for very good maintenance of the systems and their fast help when issues appeared.

Selbstständigkeitserklärung

Hiermit erkläre ich, die vorliegende Arbeit selbstständig und ohne fremde Hilfe verfasst und nur die angegebene Literatur und Hilfsmittel verwendet zu haben.

PATRICK FRITZSCH
MÜNSTER, 2. JULI 2009