Solving SU(3) Yang-Mills theory on the lattice: a calculation of selected gauge observables with gradient flow

by

Hans Mathias Mamen Vege

Thesis

for the degree of

Master of Science

Faculty of Mathematics and Natural Sciences
University of Oslo

April 2019

Abstract

In this thesis we develop GLAC, a C++ code capable of generating pure SU(3) Yang-Mills gauge configurations and applying gradient flow. Using GLAC we generate five main ensembles with $\beta = 6.0$, 6.1, 6.2, and $\beta = 6.45$. Four of the ensembles, A, B, C, and D_2, are taken to be of approximately equal volumes, with an additional ensemble D_1 for $\beta = 6.45$ with a smaller volume. Using the energy $t_f^2 \langle E \rangle$ we perform a continuum extrapolation yielding a reference scale $t_{0,\text{cont}}/r_0^2 = 0.11087(50)$ and $w_{0,\text{cont}} = 0.1695(5)$ fm. When investigating the topological susceptibility we are able to perform a similar continuum extrapolation to retrieve $\chi_{t_f}^{1/4} = 0.179(10)$ GeV for the ensembles A, B, C, and D_2, and $\chi_{t_f}^{1/4} = 0.186(6)$ GeV for the ensembles A, B, C, and D_1. Using the Witten-Veneziano formula we extract respectively $N_f = 3.75(29)$ and $N_f = 3.21(25)$, close to the expected $N_f = 3$ number of flavors for the η'-meson. We also look into the fourth cumulant of the topological charge, as well as the glueball mass for the pseudoscalar state 0^{-+}.

To grandpa and both my grandmas,

who saw me begin this journey, but not end it.

Acknowledgements

I want to thank my supervisor, Andrea Shindler, for aiding and guiding me through my thesis. His ability to answer any question I would have, and to deepen my knowledge of any given topic has been greatly appreciated.

A huge thanks also to my co-supervisor, Morten Hjorth-Jensen, without whom I would not be doing lattice QCD nor be studying Computational Physics. Thank you for believing in me and other students, thus pushing us to do our best.

Further, I greatly appreciate both the opportunity and time I spent at Michigan State University and the Lattice group I got to know there. In particular, I want to thank Jack Dragos for sharing his knowledge and providing us with data to verify our codes.

A thank you goes out to the students at Lillefy as well, as I had a great time studying there during my B.Sc. degree.

I want to thank the Computational Physics master group for all the discussions and meetings. I could not have had better co-students to work with.

I also want to thank Reinert for helpful discussions on topology, as I found them highly enlightening.

A thank you goes to Eirik, Magnus, and Simen for the countless hours we worked together and learned from each other on our B.Sc. degree. I had the best of times working with you.

I am grateful towards Giovanni Pederiva for our time working together and exploring the field of lattice QCD. If you had not decided to work on lattice QCD, I would most likely not have known about it. I greatly appreciate all of our discussions and I had a great time working with you.

Thank you to my parents and sister for being there and always believing in me, without you I would not have made it.

Finally, I want to give a special thanks to Mari for being there and listening to me after both success and defeat. You cheered me on to do my best, and knowing you always had my back gave me more support than you can imagine. Thank you from the bottom of my heart.

The lattice QCD simulations was performed on the Abel Cluster, owned by the University of Oslo and Uninett/Sigma2, and operated by the Department for Research Computing at USIT, the University of Oslo IT-department. `http://www.hpc.uio.no/`.

This work was also supported in part by Michigan State University through computational resources provided by the Institute for Cyber-Enabled Research.

Supporting runs were made at the Smaug computing cluster at the Computational Physics group at the University of Oslo in Norway.

In the end, I want to thank the Norwegian state and every taxpayer for the support. I couldn't have done it without you.

Contents

Chapter 1

Introduction

Our understanding of particle physics as it stands today is largely governed by the Standard Model of physics. The Standard Model(hereby abbreviated by *SM*), saw it's final form in the 1970s, once the electroweak forces were brought together with the strong nuclear force. Although the SM is not unifying the electroweak and strong nuclear force under a common symmetry in a similar manner that was possible to do for the weak nuclear force and electromagnetic nuclear force, the SM describes interactions between them, and has given rise to some of the best predictions physics has to offer such as the W and Z bosons, gluons and quarks.

The standard model is not without its faults, as it fails to explain phenomena such as gravity, the matter-antimatter asymmetry and many more. When the goal is to understand nature, we must ask ourselves if the SM is the correct theory. It is in this endeavor the study of the strong nuclear force, popularly called Quantum Chromo Dynamics or just *QCD*, comes into play. QCD is the theory of the strong force of quarks and gluons, which is the fundamental building blocks of hadrons i.e. protons and neutrons. As it turns out, QCD is highly non-trivial in the low-energy limit due to phenomena such as asymptotic freedom. Due to asymptotic freedom, we have that the theory is perturbative only in the high-energy regime, excluding the possibility for analytical results in the low-energy regime. This is problematic as much of the relevant physics can be found in the non-perturbative low energy regime.

Case and point being the binding energy for a nucleus. For the one down quark and two up quarks in a proton we have

$$m_p \neq m_u + m_u + m_d,$$
$$936 \text{ MeV} \neq 3 \text{ MeV} + 3 \text{ MeV} + 6 \text{ MeV}.$$

We are immediately made aware of the embarrassingly large amount of over 900 MeV that is missing. From this, we infer that a large amount of energy must be stowed away in the binding energy between the quarks through the gluons.

QCD also contains other phenomena such as color confinement in which we see no free quarks. Further, QCD is a highly nonlinear theory with three- and four-gluon vertices

which complicates the analytical investigation. A possible way out appeared almost immediately after QCD saw the light of day by Kenneth G. Wilson [76]. In what allegedly was just a pastime project during the summer, Wilson invented the field of physics now known as *lattice QCD*. Lattice QCD can be in its most basic form be described as an attempt to check if QCD is the correct theory of the strong force. One of the hopes is that lattice QCD will help bridge the gap between nuclear physics, effective field theories, and QCD.

The initial celebration of lattice QCD was quickly dampened once one realized the intense computational resources needed in order to obtain good results. Wilson himself said in 1989 that we would need a 10^8 increase in computing power as well as algorithmical advantages [77]. It is only during the last couple decades the computational power required for lattice QCD has expanded to point of true viability.

Wilson's idea of lattice QCD(abbreviated as LQCD) is to discretize Euclidean spacetime such that the fermion fields are defined at the lattice points, and the gluon fields are defined as *links* between the lattice points. The lattice will then be of the shape of a Euclidean hypercube(a cube in four dimensions) of a finite lattice spacing. From this, one can build correlators and observables to be studied. In this thesis, a pure SU(3) *Yang-Mills theory* will be studied. All fermion contributions will be left out, such that we are only studying the gluon dynamics. Even though we are not including the effects of the fermions, we can still study important phenomena such as the running coupling and topological effects such as instantons.

The primary goal of this thesis has been to develop a pure Yang-Mills lattice code from the bottom up and retrieve usable results in lattice gauge theory. We have created a code capable of generating gauge field configurations and applying gradient flow on them, while also being able to be scalable enough for running on high-performance computing clusters. Particular to this thesis, we chose to focus on topological quantities. In this thesis we will present results on the topological charge and related quantities, such as topological susceptibility, the fourth cumulant, as well as the glueball 0^{-+} state. Comparisons between other papers are made, and we hope to shed some light on a few peculiarities of these quantities.

The next chapter, chapter 2, will deal with Quantum Chromodynamics. The chapter will serve as an introduction to the standard model, then move on to QCD where we will cover some of the experimental proofs of QCD, symmetries in QCD, the topology of a SU(3) gauge theory, and other topics relevant to this thesis.

Chapter 3 will deal with lattice QCD and attempt to provide an intuitive understanding of LQCD, starting with the path integral in quantum mechanics and working our way up to topics such as the gradient flow and basic lattice observables.

After that, chapter 4 will cover numerical details related to LQCD such as the Metropolis algorithm and Runge Kutta 3 applied as a Lie group integrator.

In chapter 5 we will present the code developed for this thesis, as well as a few benchmarks for the code.

Chapter 6 will begin with presenting the scale setting t_0 and w_0, before moving on to topological quantities such as the charge, susceptibility and fourth cumulant. We will also take a look at the topological charge correlator and pseudoscalar glueball state, before

finally ending with chapter 7 - the conclusion and future outlooks.

At the very end, an extensive appendix is included, covering all from notational conventions to statistics to Lie groups to some of the more cumbersome calculations related to this thesis.

As is customary, we begin by going through the theory and required background for looking at and understanding the results obtained - QCD.

Chapter 2

Quantum Chromodynamics

In order to understand lattice QCD, we need to get a grip on the context in which it appears. It follows that in order to start this journey, we are required to look at QCD. We will begin by delving into the Standard Model and from that QCD. This will lead us to look into the interactions and many of the idiosyncrasies of QCD. We will go through asymptotic freedom, quark confinement, and experimental proof, before specifying the QCD path integral and the symmetries which appear in relation to QCD. We will make a distinction between QCD and pure SU(3) Yang-Mills theory, for then to cover the topic of topology in relation to a Yang-Mills SU(3) theory.

2.1 The Standard Model

As previously mentioned, the SM is a quantum field theory combining the weak nuclear force, the strong nuclear force, and the electromagnetic force under a single equation, and is responsible for most of the modern day understanding of particle physics. The SM is a local gauge symmetry of

$$\mathrm{SU}(3)_C \times \mathrm{SU}(2)_L \times \mathrm{U}(1)_Y, \tag{2.1}$$

where $\mathrm{SU}(3)_C$ is the color symmetry of the strong force, $\mathrm{SU}(2)_L$ is the weak isospin gauge symmetry, and $\mathrm{U}(1)_Y$ is the hypercharge gauge symmetry. From these symmetries, we have the SM Lagrangian,

$$\mathcal{L}_{\mathrm{SM}} = -\frac{1}{4}F_{\mu\nu}F^{\mu\nu} + \bar{\psi}_i i \not{D} \psi_i + h.c. - \psi_i y_{ij} \psi_j \phi + h.c. + |D_\mu \phi|^2 - V(\phi), \tag{2.2}$$

where the first term contains the gauge fields, the next two terms gives rise to the fermion propagators as well as their interaction with the gauge fields. The next two terms after that are the Yukawa couplings, which is responsible for fermions interacting with the Higgs field and subsequently resulting in their masses. The next term contains the Higgs propagator and the interaction between the Higgs field and gauge fields. The final term is the Higgs potential which is responsible for the spontaneous symmetry breaking.

Now, the SM Lagrangian in eq. (2.2) on the preceding page is a slight misdirection, as each and every term is far more complex than it looks. Expanding the field tensor $\frac{1}{4}F_{\mu\nu}F^{\mu\nu}$ gives us

$$\mathcal{L}_{\text{SM}} \subset -\frac{1}{4}B_{\mu\nu}B^{\mu\nu} - \frac{1}{2}\text{tr}W_{\mu\nu}W^{\mu\nu} - \frac{1}{2}\text{tr}G_{\mu\nu}G^{\mu\nu}. \tag{2.3}$$

The first two components of this equation go under the electroweak sector, while the latter goes under the QCD sector. The trace indicates a summation over the group generators. For instance, the gluon field strength tensor becomes $-\frac{1}{2}\text{tr}G_{\mu\nu}G^{\mu\nu} = -\frac{1}{4}G^a_{\mu\nu}G^{a\mu\nu}$ where the eight gluon fields enumerated by a is traced.

The Lagrangian can be expanded further, and if we break the electroweak symmetries $\text{SU}(2)_L \times \text{U}(1)_Y$ we end up with the fundamental particles as seen in fig. 2.1. As we can see, there are three generations of leptons and quarks, the gauge bosons of $W^{\pm}$, the Z, the photon γ and eight gluons g.

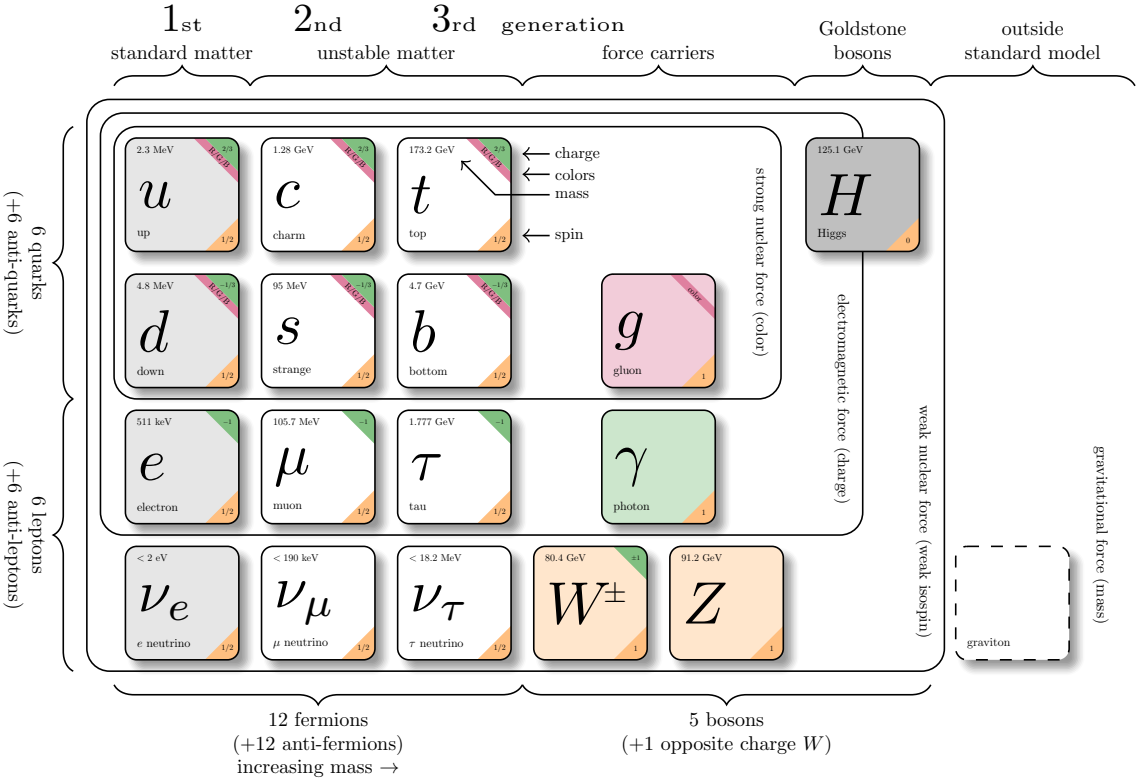


Figure 2.1: The standard model, as given under the gauge symmetries of eq. (2.1) on the previous page. Figure created by Carsten Burgard and posted here with his permission. Retrieved from `http://www.texample.net`.

Since we will hone in on the QCD sector, it is worth noting a few features of the quarks. The first being their masses and how the up u and down d quarks are of roughly the same,

small mass. We then have the quark masses in increasing order strange s, charm c, bottom b and top t. Their electric charges are fractional, with u, c, t having $+2/3$ and d, s, b having $-1/3$. We will refer to the number of quark flavors as N_f.

After expanding the Lagrangian in eq. (2.2) on page 5 and invoking the Higgs mechanism, it will be possible to extract the QCD sector[1], which is what we are interested in.

2.2 The QCD Lagrangian

As already discussed, QCD is the theory governing the strong force of quarks and gluons, and exhibits a $\mathrm{SU}(3)_C$ gauge symmetry. The Lagrangian in Minkowski space is given as

$$\mathcal{L}_{\mathrm{QCD}} = \sum_{f=1}^{N_f} \bar{\psi}^{(f)} \left(i\not{D} - m^{(f)} \right) \psi^{(f)} - \frac{1}{4} G^a_{\mu\nu} G^{a\mu\nu}, \tag{2.4}$$

where the first term is the kinetic term running over all of the N_f quark flavors: up, down, charm, strange, bottom, top. We have so far omitted labeling the quark spinors $\psi^{(f)}$ with spinor indices α and color indices c. Including these gives, $\psi^{(f)}_{\alpha c}$. The bar in ψ is a shorthand for $\bar{\psi} \equiv \psi^\dagger \gamma_0$. $\not{D}$ is the covariant derivative and is given as

$$\not{D} = D_\mu \gamma^\mu = \left(\partial_\mu - i g_s t^a A^a_\mu \right) \gamma^\mu, \tag{2.5}$$

where t^a are the generators of the Lie algebra of $\mathrm{SU}(3)_C$ known as the Gell-Mann matrices, and g_s is the strong coupling. See appendix B on page 139 for more details on the SU(3) Lie group. The spinor fields ψ transform under a local $\mathrm{SU}(3)_C$ gauge transformation $V(x) = \mathrm{e}^{i\alpha^a(x)t^a}$ as

$$\psi(x) \to \psi'(x) = V(x)\psi(x). \tag{2.6}$$

Since the covariant derivative D_μ is required to be invariant under a gauge transformation, the gauge field A^a_μ must have a finite transformation,

$$A^a_\mu(x) t^a \to A'^a_\mu(x) t^a = V(x) \left(A^a_\mu(x) t^a + \frac{i}{g_s} \partial_\mu \right) V^\dagger(x), \tag{2.7}$$

The last part of QCD Lagrangian in eq. (2.4) contains the kinetic term of the gluons, with the field strength tensor(or curvature) defined as

$$G_{\mu\nu} = \partial_\mu t^a A^a_\nu - \partial_\nu t^a A^a_\mu + i g_s t^a f^{abc} A^b_\mu A^c_\nu, \tag{2.8}$$

and follows from the commutator

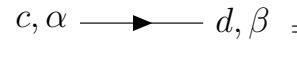
$$[D_\mu, D_\nu] = -i g_s G_{\mu\nu}. \tag{2.9}$$

For more details on the Lie group generators see appendix B on page 139.

[1] Any introductory textbook in quantum field theory will go through such mechanics, see Peskin and Schroeder [51], Willenbrock [75] or Weinberg [74] for some hands-on examples.

2.3 Interactions of QCD

From the QCD Lagrangian in eq. (2.4) on the preceding page, one can extract the Feynman rules for strong interactions. We start by writing the quark propagator,



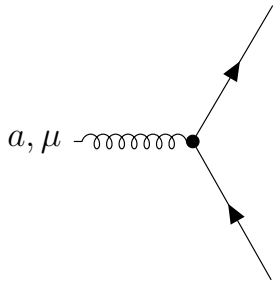
$$= \frac{i\delta^{ab}\delta^{\alpha\beta}}{\not{p} - m^{(f)}}, \tag{2.10}$$

where α, β is the Dirac indices, $f = 1, 2, \ldots, 6$ is the fermion flavor index and $a, b = 1, 2, 3$ is the color indices. The $\not{p} = \gamma^\mu p_\mu$ is the momentum of the fermion. The gluon propagator is given as

$$= \frac{-ig_{\mu\nu}}{k^2}\delta^{ab}, \tag{2.11}$$

where $g_{\mu\nu}$ is the metric tensor.

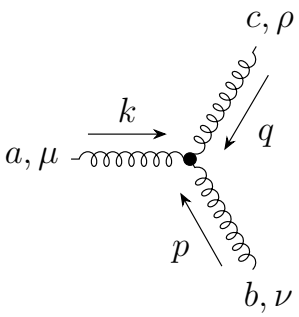
The interaction between gluons and fermions is given by



$$= ig_s\gamma^\mu t^a, \tag{2.12}$$

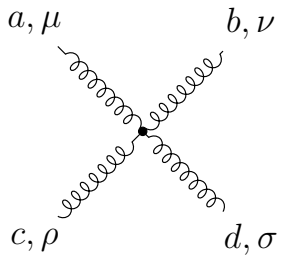
where g_s is the strong coupling.

Some of perhaps the most interesting features of a Yang-Mills theory - which QCD builds upon, is the self-interactions of the gauge fields. The first being the three gluon vertex,



$$= g_s f^{abc}[g^{\mu\nu}(k-p)^\rho + g^{\nu\rho}(p-q)^\mu + g^{\rho\mu}(q-k)^\nu], \tag{2.13}$$

and second the four gluon vertex,



$$\begin{aligned} & -ig_s^2[f^{abe}f^{cde}(g^{\mu\rho}g^{\nu\sigma} - g^{\mu\sigma}g^{\nu\rho}) \\ = \quad & + f^{ace}f^{bde}(g^{\mu\nu}g^{\rho\sigma} - g^{\mu\sigma}g^{\nu\rho}) \, . \\ & + f^{ade}f^{bce}(g^{\mu\nu}g^{\rho\sigma} - g^{\mu\rho}g^{\nu\sigma})] \end{aligned} \tag{2.14}$$

These two self-interaction terms of the gluon complicate matters, in the sense that the entire theory becomes highly non-linear.

Since we are dealing with a Yang-Mills theory, this theory also contains Faddev-Poppov ghosts which is needed in order to get rid of unphysical degrees of freedom [51, see ch. 16.2]. These will be skipped as they are not relevant to this thesis, and is usually ignored in lattice QCD calculations in general.

Having now covered the Feynman rules, we now need to look into two defining features of QCD. Namely, asymptotic freedom and quark confinement.

2.4 Asymptotic freedom

The asymptotic freedom of QCD is the phenomena in which the running coupling decreases in strength as the energy increases. This runs contrary to what is known for the electromagnetic and weak nuclear force, and as a consequence perturbation theory is only allowed in the high energy limit [51, chapter 17.1]. For instance, in QED the coupling will decrease at shorter distances, while in QCD it is reversed(see fig. 2.2 on the next page).

An expression for the running coupling $\alpha_S(Q^2)$ is realized by renormalization in the high-energy limit of perturbation theory. The one-loop solution for the running coupling for some energy scale Q^2 in the UV can be expressed as [16]

$$\alpha_S(Q^2) = \frac{\pi^2}{\beta_0 \ln(Q^2/\Lambda^2)}, \tag{2.15}$$

where β_0 is obtained from a perturbative series and Λ is the scale parameter. Both the β_0 and Λ changes depending on the theory we are in and the number of flavors, and we usually denote Λ as Λ_{QCD} or Λ_{YM}, depending on if we are in QCD or pure gauge Yang-Mills theory.

2.5 Quark confinement

One fundamental trait of QCD is that all observable hadrons are colorless. There are as of today no theoretical proof of color confinement, but all experimental data points to this

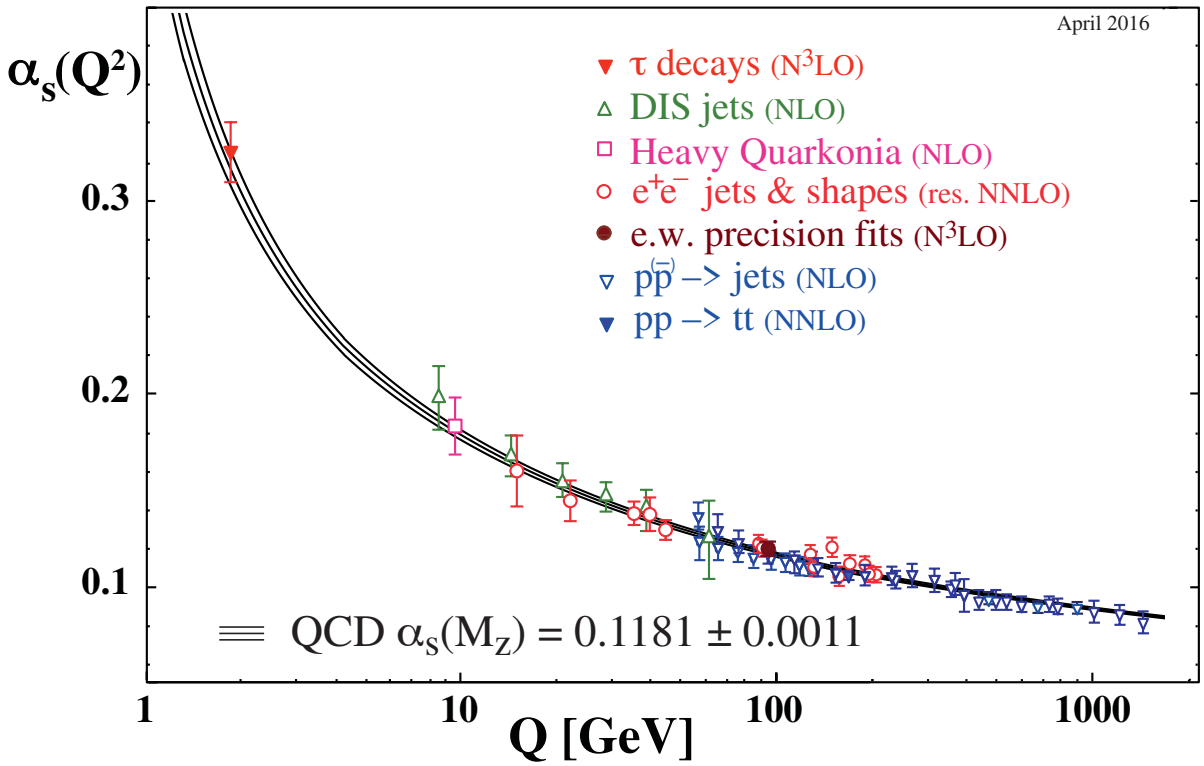


Figure 2.2: The asymptotic freedom of the strong coupling $\alpha_s(Q^2)$ with Q^2 being momentum. As discussed in section 2.4 on the preceding page, we can observe that contrary to QED, which has a coupling that increases in strength, we now have that for high energies, the QCD coupling becomes weak. The implication is profound, suggests that our method of probing low-energy QCD cannot be done analytically through perturbation theory but through numerical studies such as lattice QCD or effective field theories(e.g. chiral perturbation theory). Figure retrieved from Tanabashi et al. [68, chapter 9].

being the case [68]. There are eight gluons in the $SU(3)_C$ symmetry forming a color octet in the adjoint representation(there is no singlet, since one can only write out eight matrices that fulfills this requirement), and none of these are colorless.

A common example of how confinement works, is to visualize a meson(a quark and antiquark in a bound state) which we try to pull apart. As we pull them apart the binding energy increases and a tube of gauge fields forms between them. Once they break free, there will have been generated so much potential energy that a quark-antiquark pair is

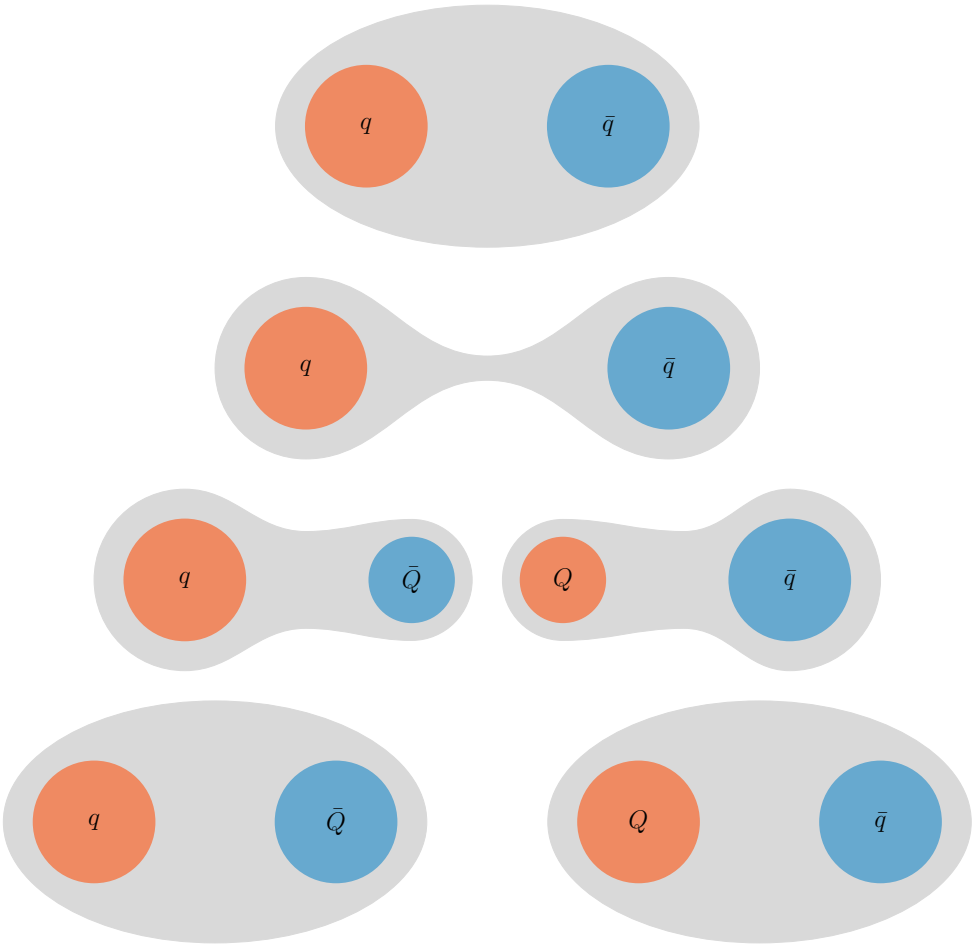


Figure 2.3: The confinement of quarks, here illustrated by a meson. As we try to pull the quarks $q\bar{q}$ apart, the potential energy between them increases until we spontaneously create a new pair of quarks $\bar{Q}Q$, that results in two new mesons, $q\bar{Q}$ and $Q\bar{q}$.

created and the two new quarks now form two new mesons with the two quarks we just pulled apart. This can be seen illustrated in fig. 2.3. There exist no theoretical proof for this as of today, but it is an observed fact one can never observe free quarks. Understanding this phenomenon is one of the avenues of investigation in lattice QCD.

2.6 Experimental proof of QCD

An outsider might worry that since it appears to be impossible to observe free quarks and gluons, QCD is a theory that *spontaneously* leaped into existence to generate funds and please the minds of physicists. Luckily - this is not the case. The asymptotic freedom which we have already seen in fig. 2.2 on the preceding page, is one of the experimental pieces of evidence for QCD. There are also several other lines of evidence for QCD.

In particle accelerators there are phenomenas called jets, which are cones of hadrons and other particles originating from the creation of quarks and gluons. Due to confinement, we know that quarks and gluons cannot exist separately, such that they will spontaneously pair up with quarks created by the vacuum in order to remain colorless. The jets that is produced by these events can be differentiated depending on the origin of production, leading to an argument for the existence of quarks and gluons.

From deep inelastic scattering experiments, the form factors reveal that the proton has a inner structure [53]. This leads to the argument that the fundamental building blocks of nature is not protons and neutrons, but rather quarks.

The cross section for the decay $\pi_0 \to \gamma\gamma$ is shown to quadratically depend on the number of colors N_c, leading to strong experimental confirmation of there existing colors in nature[2].

Another piece of strong evidence comes from looking at electrons decaying into hadrons and muons. The reaction R is defined as the cross section fraction of electron-positron annihilating into hadrons, and electron-positron annihilating into muons,

$$R \equiv \frac{\sigma(e^+e^- \to \text{hadrons})}{\sigma(e^+e^- \to \mu^+\mu^-)}. \tag{2.16}$$

We have that the first reaction is given by

$$\sigma(e^+e^- \to \text{hadrons}) = \sigma_0 \cdot 3 \cdot \sum_f Q_f^2, \tag{2.17}$$

with σ_0 being the QED cross section for $e^+e^- \to \mu^+\mu^-$. The factor 3 in front comes from the fact that there are three colors for each flavor. Factoring out $\sigma_0 = \sigma(e^+e^- \to \mu^+\mu^-)$ from the ratio, we get the fraction

$$R = 3\sum_f Q_f^2, \tag{2.18}$$

Since it is kinematically forbidden to create quark pairs $q\bar{q}$ with energies(and thus masses) larger than E(energy we put into the system), we will have that certain reactions become available to us for increasing energies. The first reaction will only include muons and the up and down quarks. Later, we will have reactions including the charm and tau, then bottom and finally the top quark. The ratio in eq. (2.16) should thus increase as more quarks come into play as the energy increases. For example, for u, d and s the ratio R in eq. (2.18) becomes

$$R = 3\left[Q_d^2 + Q_u^2 + Q_s^2\right] = 3\left[\left(\frac{2}{3}\right)^2 + \left(-\frac{1}{3}\right)^2 + \left(-\frac{1}{3}\right)^2\right] = 2, \tag{2.19}$$

where Q_d, Q_u, and Q_s are the fractional charges of the quarks.

[2]See Donoghue et al. [17, chapter VI-5] for a calculation of the cross section.

If we crank up the energy to include the c-quark we get $R = 3.33$, and cranking up further yields the b-quark with $R = 3.67$. If we were to include the top quark t as well, we would get $R = 5$. However, since the top quark is too heavy to form bound states, it is absent from data(and mostly irrelevant in lattice QCD as a whole) and we can safely ignore it.

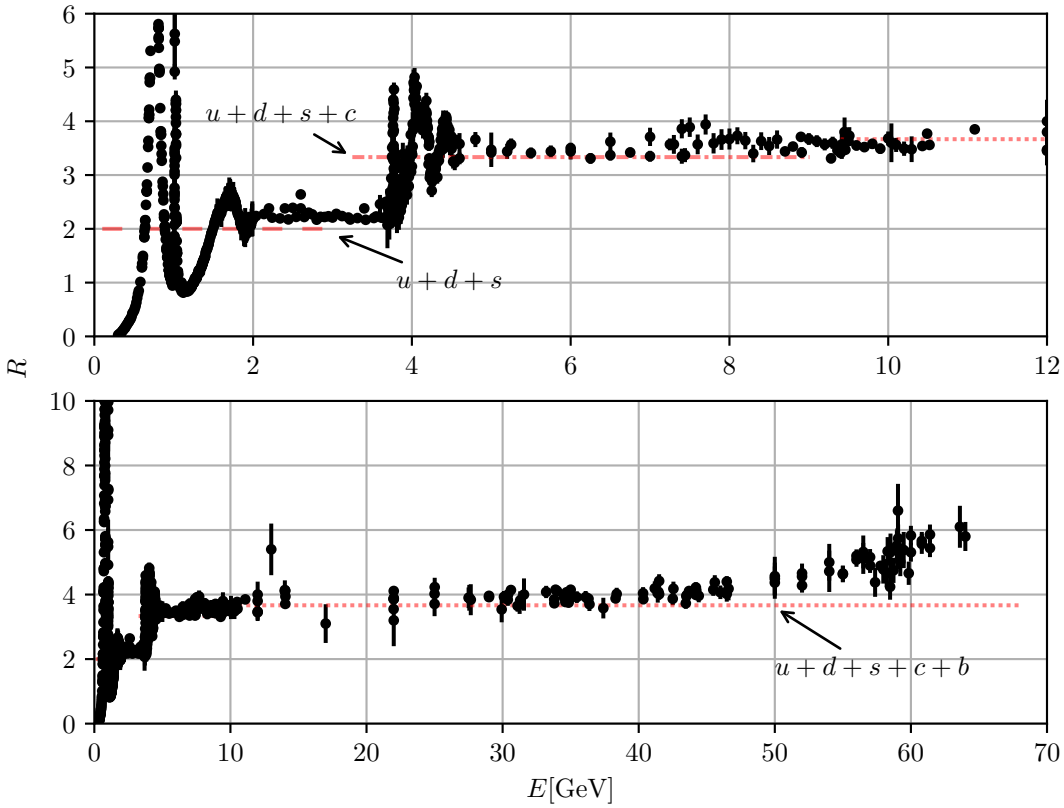


Figure 2.4: The R cross section as given from eq. (2.16) on the preceding page, with an plotted overlay of corresponding R-values for different number of quarks. Data retrieved from Tanabashi et al. [68, chapter 9] and has been compiled into a figure.

This behavior can be seen in fig. 2.4. As is clearly visible, the line is not perfect, which has to do with resonances at certain peaks, and that the quarks are virtual, such that they are not in their final state. One important takeaway from fig. 2.4, is that the factor 3 in the reaction ratio in eq. (2.18) on the facing page gives meaning to the result. As in that in without it, the prediction would be wildly off. This indicates that color is a fundamental symmetry, even though it is not directly observable.

2.7 The QCD Path Integral

Having motivated the existence of QCD, it is time to introduce the main tool and method of exploring quantum field theories and eventually lattice QCD - the path integral. In the next chapter, chapter 3, we will cover how the path integral can be discretized, as well as give a slightly more hands-on introduction to path integrals in general. As is customary and necessary when working lattice QCD and many of the topics are about to cover, we will perform a *Wick rotation*, in which we let time become imaginary such that we get what is called *Euclidean space*(see appendix A.2 on page 135). This is essential to how we are able to calculate anything in lattice QCD. Since we now will be working Euclidean space unless stated otherwise, we will be using lowered tensors as contravariant and covariant tensors are equal in Euclidean space. With this in mind, the path integral for any expectation value O in Minkowski space is

$$\langle O \rangle = \frac{1}{Z} \int \mathcal{D}A\mathcal{D}\bar{\psi}\mathcal{D}\psi O[\psi, \bar{\psi}, A] \mathrm{e}^{iS_G^M[A] + iS_F^M[\psi, \bar{\psi}, A]}, \tag{2.20}$$

where we labeled the action in Minkowski space as S^M. Applying the Wick transformation seen in eq. (A.4) on page 135 in appendix A.2 such that $A_0 \to iA_0$, $\mathrm{d}t \to -i\mathrm{d}t$, yields

$$\langle O \rangle = \frac{1}{Z} \int \mathcal{D}A\mathcal{D}\bar{\psi}\mathcal{D}\psi O[\psi, \bar{\psi}, A] \mathrm{e}^{-S_G[A] - S_F[\psi, \bar{\psi}, A]}. \tag{2.21}$$

Notice we have dropped any notational indication on the action that we are in Euclidean space, as that will be the default from now on.

We see that integration measures in the path integral from eq. (2.21) is given as

$$\mathcal{D}A = \prod_{\mu,a,x} \mathrm{d}A_\mu^a(x), \tag{2.22}$$

$$\mathcal{D}\psi = \prod_{\beta,c,x,f} \mathrm{d}\psi_{\beta c}^{(f)}(x), \tag{2.23}$$

$$\mathcal{D}\bar{\psi} = \prod_{\alpha,c,x,f} \mathrm{d}\bar{\psi}_{\alpha c}^{(f)}(x), \tag{2.24}$$

where $\alpha = 1, 2, 3, 4$ is the Dirac indices for the fermion spinors, $a, c = 1, 2, 3$ is the color indices, $\mu = 1, 2, 3, 4$ is the Lorentz indices for Euclidean space(hence no $\mu = 0$) and f is the fermion flavor.

We have also split the QCD action into its fermion- and gluon-sector, $S_{\mathrm{QCD}} = S_F + S_G$. The partition function Z in eq. (2.21) is given as

$$Z = \int \mathcal{D}A\mathcal{D}\bar{\psi}\mathcal{D}\psi e^{-S_G[A] - S_F[\psi, \bar{\psi}, A]}. \tag{2.25}$$

The action of a Lagrangian is given as $S = \int \mathrm{d}^4\mathcal{L}$, which for the gluonic action(or gauge action) becomes

$$S_G[A] = \frac{1}{2} \int \mathrm{d}^4x \mathrm{tr}\left[G_{\mu\nu}G_{\mu\nu}\right], \tag{2.26}$$

and for the fermionic action

$$\begin{aligned} S_F[\psi,\bar{\psi},A] &= \sum_{f=1}^{N_f}\int \mathrm{d}^4x\bar{\psi}^{(f)}(x)\left(\gamma_\mu(\partial_\mu+iA_\mu(x))+m^{(f)}\right)\psi^{(f)}(x) \\ &= \sum_{f=1}^{N_f}\int \mathrm{d}^4x\bar{\psi}^{(f)}_{\alpha c}(x)\left((\gamma_\mu)_{\alpha\beta}(\delta_{cd}\partial_\mu+iA_\mu(x)_{cd})+m^{(f)}\delta_{\alpha\beta}\delta_{cd}\right)\psi^{(f)}_{\beta d}(x). \end{aligned} \tag{2.27}$$

The flavor-, spinor- and color-indices will be assumed implicitly, with latter given as matrix/vector notation in future discussions.

Worth noting is that the fermion fields ψ is Grassmann valued fields. This means we can integrate out their contributions in the shape of determinants[3], such that the partition function becomes

$$Z = \int \mathcal{D}Ae^{-S_G[A]}\prod_f \det\left[D^{(f)}\right], \tag{2.28}$$

where $\det\left[D^{(f)}\right]$ is the fermion determinant for a given flavor f.

Before moving on, we can introduce the two-point correlator,

$$\langle O_2(t)O_1(0)\rangle = \frac{1}{Z}\int \mathcal{D}A\mathcal{D}\bar{\psi}\mathcal{D}\psi O_2[\psi,\bar{\psi},A]O_1[\psi,\bar{\psi},A]\mathrm{e}^{-S_G[A]-S_F[\psi,\bar{\psi},A]}. \tag{2.29}$$

The O_1 and O_2 are operators taken at some specific time, in which the operator O_2 usually is at some Euclidean time t and is called the *sink* and O_1 at Euclidean time $t=0$ called the *source*. How the path integral and the two-point correlator is applied to lattice QCD will be explained in chapter 3.

2.8 Additional symmetries of QCD

As discussed earlier, the rules of the game in QCD is governed by the local gauge symmetry $\mathrm{SU}(3)_C$. In addition to this, follow a series of "accidental" symmetries, which has several important consequences. These symmetries are *global*, contrary to the local $\mathrm{SU}(3)_C$ color symmetry. The most important and relevant for this thesis is the global chiral symmetries for N_f flavors for a massless fermion Lagrangian.

2.8.1 Chiral symmetries

The axial symmetry of QCD is an approximate symmetry. If we assume that we are working in the *chiral limit* - that is letting the quark masses $m_f \to 0$. Ignoring the gauge

[3]See [24, 51] or any standard quantum field theory book for details on how this procedure is done.

fields for now gives us the Lagrangian

$$\mathcal{L}(\psi, \bar{\psi}, A) = \sum_{f=1}^{N_f} \bar{\psi}^{(f)} \not{D} \psi^{(f)}, \tag{2.30}$$

with $\not{D}$ defined as in eq. (2.5) on page 7, and f being quark flavor.

In order to illustrate chiral symmetry breaking we consider a *single* flavor in the fermionic section of the QCD Lagrangian in eq. (2.4) on page 7. We check if the QCD Lagrangian is invariant under a chiral rotation,

$$\psi \to \psi' = \mathrm{e}^{i\alpha\gamma_5}\psi, \quad \bar{\psi} \to \bar{\psi}' = \bar{\psi}\mathrm{e}^{i\alpha\gamma_5}. \tag{2.31}$$

By investigation of this property(see appendix B.5 on page 144 for an explicit calculation of eq. (2.32) and eq. (2.33)) it turns out that we can decouple the fermionic section into right- and left-handed components,

$$\mathcal{L}(\psi, \bar{\psi}, A) = \bar{\psi}_R \not{D} \psi_R + \bar{\psi}_L \not{D} \psi_L, \tag{2.32}$$

which obeys the chiral symmetry in eq. (2.31). That is, as long as we have no mass term. Including a mass term gives us

$$m\bar{\psi}\psi = m\left(\bar{\psi}_L \psi_R + \bar{\psi}_R \psi_L\right), \tag{2.33}$$

which breaks chiral symmetry. It is important to stress that this example is only for one flavor, as we in full QCD have that γ_5 is anomalous.

When including N_f flavors, the Lagrangian is invariant under vector transformations. If we add up all of the symmetries for N_f flavors, we get a $\mathrm{U}(N_f)_L \times \mathrm{U}(N_f)_R$ global chiral symmetry. Since $\mathrm{U}(N) = \mathrm{SU}(N) \times \mathrm{U}(1)$, this can be composed into

$$\mathrm{U}(N_f)_L \times \mathrm{U}(N_f)_R \simeq \mathrm{SU}(N_f)_L \times \mathrm{SU}(N_f)_R \times \mathrm{U}(1)_V \times \mathrm{U}(1)_A, \tag{2.34}$$

where the symmetry group $\mathrm{SU}(N_f)_L \times \mathrm{SU}(N_f)_R = \mathrm{SU}(N_f)_V \times \mathrm{SU}(N_f)_A$ spontaneously breaks to $\mathrm{SU}(N_f)_V$. Assuming $N_f = 3$(the three lightest quark masses), we have that the breaking of the $\mathrm{SU}(3)_A$ leads to the eight pseudoscalar mesons(approximate Goldstone bosons). The $\mathrm{SU}(3)_V$ is a vector symmetry[4] that is softly broken, leading to approximate symmetries of isospin and strangeness. The $\mathrm{U}(1)_V$ symmetry holds, and corresponds the conservation of baryon numbers. One would expect that the remaining axial $\mathrm{U}(1)_A$ symmetry to be broken spontaneously in the chiral limit. Estimations by Weinberg [73] indicated that the mass of this isosinglet pseudoscalar Goldstone boson should have a mass less that $\sqrt{3}m_\pi$, with m_π being the mass of a pion. The two candidates available was $\eta(549)$ and $\eta'(985)$, where the former is already explained for by the octet, but the latter has a mass too high for the bound. This discrepancy is known as the $U(1)_A$ problem, and was eventually solved as it turned out to be violated by the axial anomaly. The axial anomaly leads to a change in the integration measure due to effects of the QCD vacuum [66, 67].

[4]The Noether current of the vector symmetry is a vector current, and it does not distinguish between handedness.

2.9 From QCD to pure Yang-Mills SU(3) theory

Since this thesis will be focusing on a pure Yang-Mills SU(3) lattice simulation, it is appropriate to make it formal what this entails. We will effectively neglect any quark contributions in the Lagrangian, such that the theory we are left with then is a fermion free theory describing gluon dynamics. The action and Lagrangian describing pure Yang-Mills gauge theory becomes

$$S_{\mathrm{YM}}[A] = \int \mathrm{d}^4x \mathcal{L}_{\mathrm{YM}} = \int \mathrm{d}^4x \frac{1}{4} G^a_{\mu\nu} G^{a\mu\nu}, \tag{2.35}$$

where a is a trace of the color indices. The Feynman rules for pure gluon interactions and dynamics remain the same. There is still a plethora of theory relevant to QCD that can be explored in this approximation, e.g. the topology of the gauge fields and the coupling $\alpha_{\mathrm{YM}}(Q^2)$.

2.10 The topology of a SU(3) gauge theory

Investigating the topology of gauge theories can provide a deep and profound insight into the theory we are investigating. In the context of QCD, we wish to investigate how topology is related to the QCD vacuum. It turns out that configurations[5] of the gauge field can be classified based on their topological properties. These properties, such as their *winding number*, provides a classification of the gauge fields. Configurations of the gauge field have local minimums of the action S, which in Euclidean spacetime are known as *instantons*. Further, these instantons leads to the introduction of a possible CP-violating θ term, and lifts the vacua to new states that are a linear combination of old states shifted by a phase θ.

In order to better appreciate the results presented in this thesis and properly understand instantons and the θ-vacuum, we need to build an adequate understanding of the basics of topology.

2.10.1 A brief introduction to topology

Topology is the study of general properties of mathematical objects or spaces. Since the theory we are exploring is a global SU(3) gauge symmetry, we are dealing with a *manifold*. A manifold $\mathscr{M}$ is a topological space that locally looks like $\mathbb{R}^n$ Euclidean space [49]. What we are interested in is how we can classify different topological spaces. A topological space is a set of points that together with a set of neighborhoods for each point, satisfies a set of axioms relating points and neighborhoods.

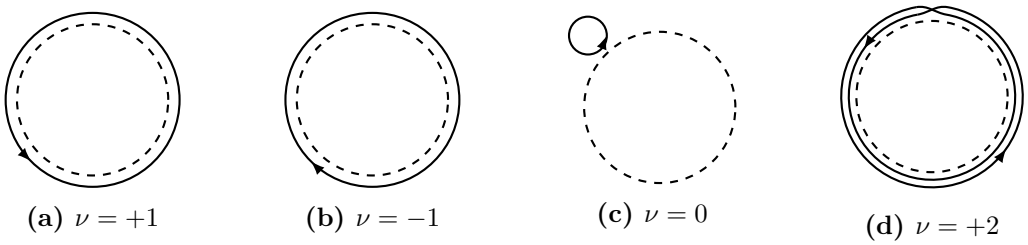
Two topological spaces is equal, if they can be continuously deformed into each other leaving their properties invariant under such a transformation. We then call the objects

[5] A configuration is simply a given alignment or state(not in the quantum mechanical sense) of a field. In LQCD we generate configurations, or simply a 4D hypercube cube of oriented SU(3) matrices.

homeomorphic. The visually pleasing example of this is imagining a mug or a cup, and a donut(torus). We can transform the mug into a torus and correspondingly the other way around, without losing the hole in neither, making them homeomorphic to each other.

We can now further classify objects of into *equivalence classes* of the same homotopy group $\pi_k(\mathscr{M})$, with k being the dimension. Such a mapping can, for instance, be of a simple circle or 1-sphere S^1, leading to a non-trivial homotopy $\pi_1(S^1)$. Given a function f which parametrizes $f : S^1 \to S^1$, we have that there exist a denumerable(one-to-one) number of classes which can be labeled by a natural number, an integer $\nu \in \mathbb{Z}$, which counts how many times the mapping f wrap around the circle. We call this number ν the *winding number*. In fig. 2.5 we see an example of such a mapping of four circles with winding numbers $\nu = -1, 0, +1, +2$.

Two mappings, f and g, are homotopic to each other if they can be continuously deformed into each other without leaving their equivalence class. The set of the classes in such a manifold $\mathscr{M}$ forms a homotopy group $\pi_k(\mathscr{M})$.



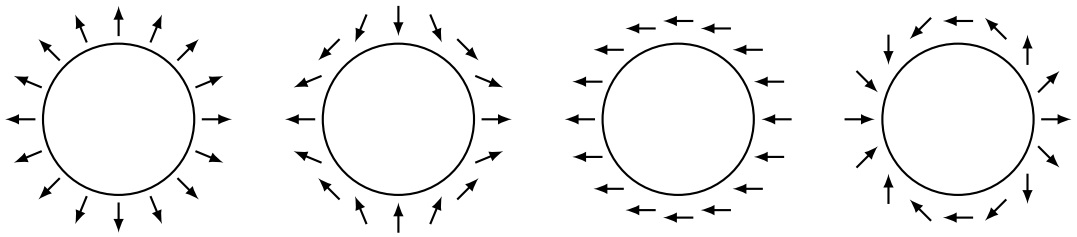
(a) $\nu = +1$ **(b)** $\nu = -1$ **(c)** $\nu = 0$ **(d)** $\nu = +2$

Figure 2.5: An illustration of how one can view the winding number given a function f that parametrizes a path around a circle S^1. Given that it starts and ends at the same point, we have that the number of times it wraps around the circle gives us the winding number. The figure is taken from [21, p. 32].

A similar example can be given by viewing a vector field and imagining taking a path around such a vector field. From fig. 2.6 on the facing page, we see that the winding number tells us the number of times a vector rotates on the boundary rotates. From Stokes theorem, we can think of the winding number as something that counts the number of field vortices inside the boundary [27].

In general, for a gauge group G, the set of classes of a topologically distinct mapping $S^{d-1} \mapsto G$ where S^{d-1} is the sphere of dimension $d-1$, is known as $\pi_{d-1}(G)$. For all simple Lie groups G[6], $\pi_3(G) = \mathbb{Z}$, this includes SU(3) as seen in [74]. The π_3 indicates the group we are mapping from - in this case S^3. The winding number ν goes under many different names, topological charge, Pontryagin number, winding number, and instanton number. For *how* the winding number can provide us insight into the structure of the gauge fields, we will have to introduce instantons and the concept of topological charge, Q. We have that two field configurations belong to the same homotopy class if they can be continuously deformed into each other, leaving their properties invariant under this transformation(i.e. not passing through a forbidden region where the action S is infinite).

[6]By simple Lie group, we mean a compact, non-Abelian, connected Lie Group G, e.g. SU(3) [49].



(a) $\nu = +1$ **(b)** $\nu = -1$ **(c)** $\nu = 0$ **(d)** $\nu = +2$

Figure 2.6: The winding number illustrated by the number of rotations a vector field makes on some boundary The initial orientation of the vectors is arbitrary, as only the rotation is counted. The figure is taken from [27, 36].

2.10.2 Instantons

Instantons can be viewed as local minimums of the action in Euclidean space, as they by definition are localized finite action solutions to the classical Euclidean equations of motion [21, 36, 49, 54, 58, 69, 70],

$$D_\mu G_{\mu\nu} = 0. \tag{2.36}$$

Further, they can be viewed as pseudoparticles in Euclidean time [2], as they induce excitations to the vacuum.

To understand how instantons arise, let us recall the Euclidean Yang-Mills path integral in eq. (2.21) on page 14 and write out the Euclidean action $S_G[A]$,

$$S_G[A] = \frac{1}{2}\int \mathrm{d}^4x \mathrm{tr}[G_{\mu\nu}(x)G_{\mu\nu}(x)] = \frac{1}{4}\int \mathrm{d}^4x G^a_{\mu\nu}(x)G^a_{\mu\nu}(x). \tag{2.37}$$

In order for $S_G[A]$ to be finite the gauge fields must approach pure gauge at $|x| \to \infty$,

$$it_a A_{a\mu} \overset{|x|\to\infty}{=} V(\hat{x})^\dagger \partial_\mu V(\hat{x}), \tag{2.38}$$

with $V(\hat{x}) \in \mathrm{SU}(3)$ as a group element of SU(3) and $\hat{x}$ as a directional dependent element of gauge group SU(3).

If we now rewrite the action in eq. (2.37) using the identity as listed in appendix B.6, we get

$$S = \frac{1}{4}\int \mathrm{d}^4x \left[\pm G^a_{\mu\nu}\tilde{G}^a_{\mu\nu} + \frac{1}{2}(G^a_{\mu\nu} \mp \tilde{G}^a_{\mu\nu})^2\right]. \tag{2.39}$$

From this we can get a lower bound on the field strength tensor [6], by requiring it to be self dual or anti-self dual(see eq. (B.33) on page 146),

$$S \geq \pm\frac{1}{4}\int \mathrm{d}^4x G^a_{\mu\nu}\tilde{G}^a_{\mu\nu}. \tag{2.40}$$

One can show that this equals the *topological charge* [74],

$$S[A] = \frac{8\pi^2}{g_S^2}|Q|. \tag{2.41}$$

The $1/g_s^2$ factor comes from the convention of letting A_μ and thus $F_{\mu\nu}$ carry and extra $1/g_s$ factor. We will come back to this in chapter 3.

When inserting the self dual tensor from eq. (B.33) on page 146 in the action in eq. (2.40) on the previous page and using that the right-hand side equals eq. (2.41), the topological charge is given as

$$\frac{8\pi^2}{g_S^2}Q = \frac{1}{8}\epsilon_{\mu\nu\sigma\rho}\int \mathrm{d}^4x G^a_{\mu\nu}G^a_{\sigma\rho}, \tag{2.42}$$

which becomes

$$Q = \frac{g_S^2}{64\pi^2}\int \mathrm{d}^4x \epsilon_{\mu\nu\sigma\rho}G^a_{\mu\nu}G^a_{\sigma\rho}. \tag{2.43}$$

We can rewrite this equation in terms of the topological charge density $q(x)$,

$$Q = \int \mathrm{d}^4x q(x), \tag{2.44}$$

where

$$q(x) = \frac{g_S^2}{32\pi^2}\epsilon_{\mu\nu\sigma\rho}\mathrm{tr}\left[G_{\mu\nu}(x)G_{\rho\sigma}(x)\right]. \tag{2.45}$$

Every instanton contributes to one unit of topological charge, while anti-instantons contribute to one unit of negative topological charge. This can also be seen by virtue of the index theorem [21],

$$Q = n_- - n_+, \tag{2.46}$$

where n_- and n_+ corresponds to left- and right-handed zero modes of the Dirac operator $\not{D}$.

The introduction of instantons provides a solution to the $\mathrm{U}(1)_A$ problem, as they change the integration measure for quark fields when undergoing a global, chiral U(1) transformation. This invariance is referred to as the *chiral anomaly*, as it is the anomaly that breaks the $\mathrm{U}_A(1)$ symmetry.

2.10.3 θ-vacuum

Because of tunneling between topological sectors, we have that the vacuum becomes a linear combination of all vacua [58],

$$|\mathrm{vac}\rangle \equiv |\theta\rangle = \sum_n \mathrm{e}^{in\theta}|n\rangle. \tag{2.47}$$

The θ-term can be derived from either Gauss' law or from cluster decomposition [21, 74]. This ground state redefinition will lead to a modification of the integration measure. To illustrate this we can imagine us to take the expectation value of some observable O. In order to account for instantons and tunneling between topological sectors, we add some weighting factor $w(Q)$ to the expectation value, summing up all contributions from the topological charge in the Euclidean spacetime volume Ω. This gives us

$$\langle O\rangle_\Omega = \langle 0|O|0\rangle_\Omega = \frac{\sum_{Q=-\infty}^{\infty} w(Q)\int \mathcal{D}\psi\mathcal{D}\bar{\psi}\mathcal{D}A_Q O[\psi,\bar{\psi},A]\mathrm{e}^{-S[\psi,\bar{\psi},A]_\Omega}}{\sum_{Q=-\infty}^{\infty} w(Q)\int \mathcal{D}\psi\mathcal{D}\bar{\psi}\mathcal{D}A_Q \mathrm{e}^{-S[\psi,\bar{\psi},A]_\Omega}}, \tag{2.48}$$

where $S[\psi,\bar{\psi},A]_\Omega$ is taken over the entire the Euclidean spacetime volume Ω.

We can now split the Euclidean spacetime volume into two sectors, $\Omega = \Omega_1 + \Omega_2$, with Ω_1 containing the observable O. This splits the contribution of the topological charge from the gauge field into $Q = Q_1 + Q_2$ as well, giving us two quantities that not necessarily are integer valued[7]. The action will also split into $S[\psi,\bar{\psi},A]_\Omega = S[\psi,\bar{\psi},A]_{\Omega_1} + S[\psi,\bar{\psi},A]_{\Omega_2}$. If we now take a look at the integral measure, we can try to decompose it into one section for Ω_1 and one for Ω_2.

Consider a path integral measure where we use that $\Omega = \Omega_1 + \Omega_2$,

$$\begin{aligned}\mathcal{D}\phi^{(\Omega)} &= \prod_{x\in\Omega}\int \mathrm{d}\phi(x) = \prod_{x_1\in\Omega_1}\int \mathrm{d}\phi(x_1)\prod_{x_2\in\Omega_2}\int \mathrm{d}\phi(x_2)\\ &= \int \mathcal{D}\phi^{(\Omega_1)}\int \mathcal{D}\phi^{(\Omega_2)}.\end{aligned} \tag{2.49}$$

Following closely what is done in [21], we have that since only the gluon gauge fields affect the integral measure, we get

$$\begin{aligned}\sum_{Q=-\infty}^{\infty} w(Q)\int \mathcal{D}A_Q^{(\Omega)} &= \sum_{Q=-\infty}^{\infty} w(Q)\sum_{Q_1=-\infty}^{\infty}\int \mathcal{D}A_{Q_1}^{(\Omega_1)}\sum_{Q_2=-\infty}^{\infty}\int \mathcal{D}A_{Q_2}^{(\Omega_2)}\delta_{Q,Q_1+Q_2}\\ &= \sum_{Q=-\infty}^{\infty}\left(\sum_{Q_1,Q_2=-\infty}^{\infty} w(Q)\int \mathcal{D}A_{Q_1}^{(\Omega_1)}\int \mathcal{D}A_{Q_2}^{(\Omega_2)}\delta_{Q,Q_1+Q_2}\right),\end{aligned} \tag{2.50}$$

where we use that the path integral measure is equal to the decomposition into *two* path integral measures, as long as we only count their overall contribution to the first sum once using the Kronecker delta. Applying the Kronecker delta on the first sum, we end up with

$$\sum_{Q_1,Q_2=-\infty}^{\infty} w(Q_1+Q_2)\int \mathcal{D}A_{Q_1}^{(\Omega_1)}\int \mathcal{D}A_{Q_2}^{(\Omega_2)}. \tag{2.51}$$

[7]Since Q is defined at the boundary of Ω, this is the case. When considering the entire spacetime volume, Q should be integer valued [54].

Cluster decomposition[8] now enters the picture, and allows us to make the following assumption given that the two events are spatially separated in Ω_1 and Ω_2 from initial to a final state at all times,

$$w(Q_1 + Q_2) = w(Q_1)w(Q_2). \tag{2.52}$$

By demanding this, we get that the only choice for $w(Q)$ is to be an exponential,

$$w(Q) = \mathrm{e}^{i\theta Q}. \tag{2.53}$$

Inserting this into eq. (2.48) on the previous page leads to a change in the integration measure,

$$\mathcal{D}\psi\mathcal{D}\bar{\psi}\mathcal{D}A \to \exp\left(-\frac{i\theta g_S^2}{64\pi^2}\int \mathrm{d}x^4 \mathrm{tr} G_{\mu\nu}\tilde{G}_{\mu\nu}\right)\mathcal{D}\psi\mathcal{D}\bar{\psi}\mathcal{D}A. \tag{2.54}$$

Changing the integration measure gives rise to the axial anomaly current,

$$\partial_\mu j_5^\mu(x) = i2N_f q(x). \tag{2.55}$$

The inclusion of the weighting factor in eq. (2.53) in the integration measure is equivalent with adding a θ-term to the QCD Lagrangian in eq. (2.4) on page 7,

$$\mathcal{L}_\theta = -\frac{i\theta g_S^2}{64\pi^2}\mathrm{tr} G_{\mu\nu}\tilde{G}_{\mu\nu} = -i\theta Q. \tag{2.56}$$

As it turns out, this term is CP-violating. If it exists, it should be observable through the EDM(electric dipole moment) of the neutron,

$$d_n \approx \bar{\theta} e \frac{m_\pi^2}{m_N^3} \approx \bar{\theta} \times 10^{-16} e\,\mathrm{cm}. \tag{2.57}$$

The current upper limit on experiments is $d_n \leq 3 \times 10^{-26} e\,\mathrm{cm}(90\ \%\ \mathrm{C.L.})$. This puts the θ-value at [59],

$$\bar{\theta} \lesssim 10^{-10}, \tag{2.58}$$

where the bar indicates that it contains a CP-violating phase from weak interaction. Similar results from lattice QCD have recently been presented by Dragos et al. [18] that puts the neutron EDM at $|\bar{\theta}| < 1.61(51) \times 10^{-10}$.

In principle, the θ term could have any value between 0 and 2π. The mechanism behind why the θ term is this small and close to zero remains unknown, and is referred to as the *strong CP problem*. One of the proposed solutions to this problem suggest that we promote θ to a dynamic variable, thus becoming a field Peccei and Quinn [50] and leading to the introduction of particles called *axions*. If the strong CP problem is solved, it could provide a clue to why there is so much matter and so little anti-matter in the universe.

[8] Cluster decomposition can be summed up as following: given that at initial time two sufficiently spaced subsystems can be factorized and they remain spatially separated through at all time, then the final state can also be factorized [63]. For a more accessible approach to this topic, please see [1].

2.10.4 Witten-Veneziano formula

Witten and Veneziano resolved the issue of the large η' mass [71, 78], by showing that in the limit of a large number of colors $N_c \to \infty$ and to leading order of $\mathcal{O}(1/N_c)$, the chiral anomaly vanishes, so that the symmetry $U_A(1)$ is restored. This gave rise to the formula connecting the η' mass and the topological susceptibility [15, 25, 71, 72, 78],

$$m_{\eta'}^2 = \frac{2N_f}{f_\pi^2}\chi_{\text{top}}, \tag{2.59}$$

with N_f being the number of flavors, f_π is the pion decay constant given at $f_\pi \approx 92$ MeV and χ_{top} is the topological susceptibility. Equation (2.59) is accurate up to $\mathcal{O}\left(1/N_c^2\right)$. Recalling the topological charge in eq. (2.43) on page 20 and the topological charge density from eq. (2.45) on page 20, the topological susceptibility takes the shape as

$$\chi_{\text{top}} = \int \mathrm{d}^4x \langle q(x)q(0)\rangle, \tag{2.60}$$

where $q(x)$ is the topological charge density in the two point correlator $\langle q(x)q(0)\rangle$.

One interesting aspect to note is that the r.h.s and l.h.s. of the WV-formula in eq. (2.59) represent separate theories. That is, χ_{top} is taken in pure gauge Yang-Mills theory, while the right-hand side is taken in full QCD.

2.10.5 The glueball and topological charge correlators

Having briefly touched upon the two-point correlator for the topological charge in the previous section, it is worth introducing the glueball. Glueballs are gluons in bound, massive, colorless states. In pure Yang-Mills gauge theory, we can have stable glueball states, while in full QCD theory we will have mixing with quark states, making it a difficult phenomenon to explore. The gluonic currents with quantum numbers J^{PC} for the lowest glueball states is given as [58, p. 88]

$$S = 0^{++}: \quad j_S = (G^a_{\mu\nu})^2, \tag{2.61}$$

$$P = 0^{-+}: \quad j_P = \frac{1}{2}\epsilon_{\mu\nu\rho\sigma}G^a_{\mu\nu}G^a_{\rho\sigma}, \tag{2.62}$$

$$T = 2^{++}: \quad j_T = \frac{1}{4}(G^a_{\mu\nu})^2 - G^a_{0\alpha}G^a_{0\alpha}. \tag{2.63}$$

The state of interest to use will be the $P = 0^{-+}$ state, which can be investigated using the correlation function

$$C(r) = \langle q(x)q(y)\rangle, \quad \text{with} \quad r = |x - y|, \tag{2.64}$$

where the $q(x)$ is the topological charge.

Chapter 3

Lattice QCD

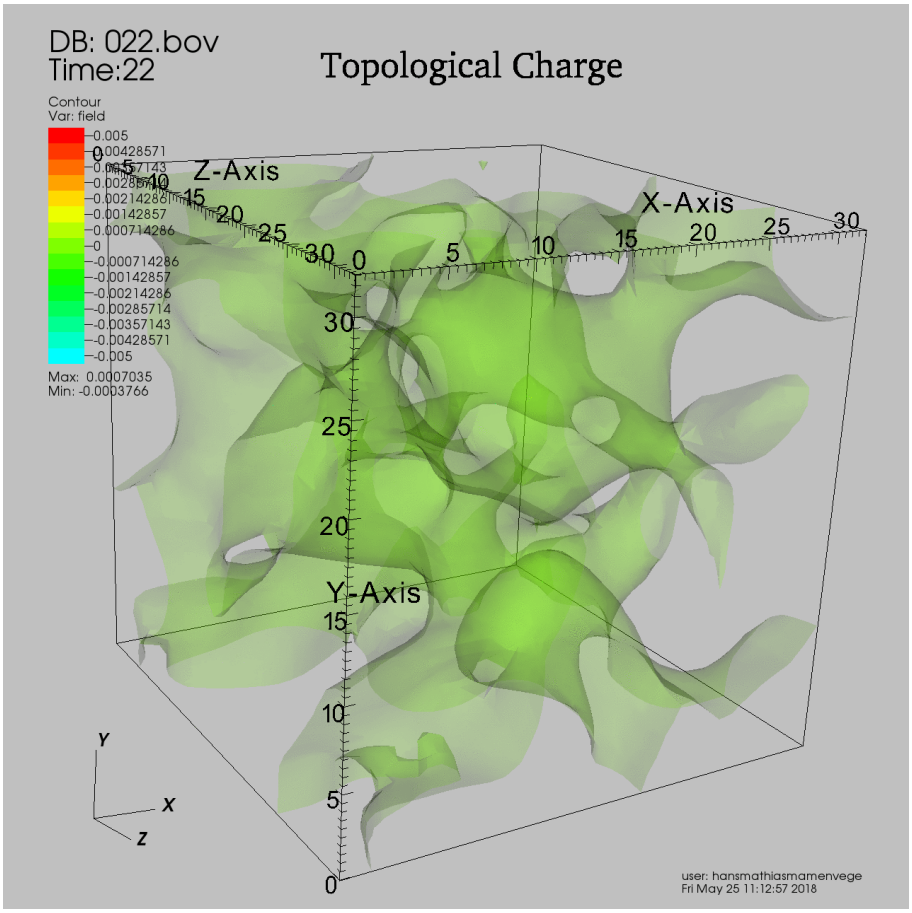


Figure 3.1: A visualization of the topological charge density in eq. (3.90) on page 42 for a $32^3 \times 64$ lattice with $\beta = 6.2$ and flow time $\sqrt{8t_f} = 0.54$ fm. The image was created by LatViz, a program created in a joint collaboration between Giovanni Pederiva and the author, Mathias M. Vege. The code for LatViz is available at GitHub: `https://github.com/hmvege/LatViz`.

Lattice QCD saw its light of day in 1974, in the seminal paper by Kenneth G. Wilson [76]. In his paper, Wilson showed how one could discretize a Euclidean spacetime on a lattice and maintain gauge invariance, albeit giving up Lorentz- and rotational invariance. By Wilson's method of discretizing spacetime we are introduced to lattice QCD and allowed to explore non-perturbative aspects of QCD. The introduction of a lattice also serves as UV cutoff proportional to the inverse lattice spacing. As a motivating example, let us begin by introducing the path integral in the context of quantum mechanics.

3.1 A motivational example in quantum mechanics

Having already introduced the path integral in eq. (2.21) on page 14 in section 2.7, it is time to show how it can be applied to lattice QCD. We begin by introducing it in a context of quantum mechanics. To goal here is not to derive the path integral, but to give an intuitive idea for how one can discretize a path integral into something which can be calculated numerically[1].

In the following discussion, we are working in Euclidean time, as a Wick rotation has been performed. For details on Wick rotation, see appendix A.2 on page 135 for details.

3.1.1 The Euclidean quantum mechanical correlator

We start by considering a general Hamiltonian in quantum mechanics,

$$\hat{H} = \frac{\hat{p}^2}{2m} + \hat{V}(x), \tag{3.1}$$

where $\hat{p}$ is the momentum operator, m is some mass and $\hat{V}(x)$ is a potential. We then apply the Hamiltonian on a ket in some state $|n\rangle$,

$$\hat{H}\,|n\rangle = E_n\,|n\rangle\,. \tag{3.2}$$

Since E_n are eigenvalues of the Hamiltonian, ordered, and real numbers we have

$$E_0 \leq E_1 \leq E_2 \ldots. \tag{3.3}$$

Given that the basis $|n\rangle$ of $\hat{H}$ is an orthonormal one, we get

$$\begin{aligned}\sum_n \langle n|\, \mathrm{e}^{-T\hat{H}}\, |n\rangle &= \sum_n \mathrm{e}^{-TE_n} \langle n\,|n\rangle = \sum_n \mathrm{e}^{-TE_n} \times 1 \\ &= \mathrm{e}^{-TE_0} + \mathrm{e}^{-TE_1} + \mathrm{e}^{-TE_2} + \ldots,\end{aligned} \tag{3.4}$$

which is a sum over the energy spectrum.

[1] Most textbooks on quantum field theory(and usually books on advanced quantum mechanics) contain some introduction on path integrals, and thus there are many good to chose from [24, 34, 37, 51, 60].

The next step is to evaluate two observables at time $t_i = 0$ and $t_f = t$ such that $T = t_f - t_i$,

$$\left\langle \hat{O}_2(t)\hat{O}_1(0) \right\rangle_T = \frac{1}{Z_0} \sum_m \langle m| \, \mathrm{e}^{-(T-t)\hat{H}} \hat{O}_2 \mathrm{e}^{-t\hat{H}} \hat{O}_1 \, |m\rangle \tag{3.5}$$

$$= \frac{1}{Z_0} \sum_m \langle m| \, \mathrm{e}^{-(T-t)\hat{H}} \hat{O}_2 \sum_n |n\rangle \langle n| \, \mathrm{e}^{-t\hat{H}} \hat{O}_1 \, |m\rangle \tag{3.6}$$

$$= \frac{1}{Z_0} \sum_{m,n} \mathrm{e}^{-(T-t)E_m} \langle m| \, \hat{O}_2 \, |n\rangle \, \mathrm{e}^{-tE_n} \langle n| \, \hat{O}_1 \, |m\rangle \tag{3.7}$$

$$= \frac{\sum_{m,n} \langle m| \, \hat{O}_2 \, |n\rangle \langle n| \, \hat{O}_1 \, |m\rangle \, \mathrm{e}^{-tE_n} \mathrm{e}^{-(T-t)E_m}}{\mathrm{e}^{-TE_0} + \mathrm{e}^{-TE_1} + \mathrm{e}^{-TE_2} + \dots}, \tag{3.8}$$

where we inserted a complete set of states in eq. (3.6). We can now divide by the ground state e^{-TE_0} such that $\Delta E_n = E_n - E_0$,

$$\left\langle \hat{O}_2(t)\hat{O}_1(0) \right\rangle_T = \frac{\sum_{m,n} \langle m| \, \hat{O}_2 \, |n\rangle \langle n| \, \hat{O}_1 \, |m\rangle \, \mathrm{e}^{-t\Delta E_n} \mathrm{e}^{-(T-t)\Delta E_m}}{1 + \mathrm{e}^{-T\Delta E_1} + \mathrm{e}^{-T\Delta E_2} + \dots}. \tag{3.9}$$

We can now let $T \to \infty$ and relabel by convention $\Delta E_n \to E_n$,

$$G(t) \equiv \lim_{T\to\infty} \left\langle \hat{O}_2(t)\hat{O}_1(0) \right\rangle_T = \sum_n \langle 0| \, \hat{O}_2 \, |n\rangle \langle n| \, \hat{O}_1 \, |0\rangle \, \mathrm{e}^{-tE_n}, \tag{3.10}$$

where the exponential $\mathrm{e}^{-(T-t)\Delta E_m}$ became $\mathrm{e}^{-(T-t)\times 0}$, leaving us with m as the only state which survives due to being the ground state $m = 0$. We now have an expression for the correlator, $G(t)$. We can write out the sum as

$$G(t) = A_0 \mathrm{e}^{-tE_0} \left(1 + \mathcal{O}(\mathrm{e}^{-t\Delta E_1})\right), \tag{3.11}$$

where A_0 is a constant. There now is a real danger of confusion as we have re-introduced ΔE_1, only this time as the difference between E_0 and E_1 as the first excited state. One quick observation to take home, is that we can only investigate energy differences.

Picking up from eq. (3.10) and we assume that $E_1 \gg E_2$ such that if we take the correlator at two different points t and $t + a$ where a is simply one step length forward,

$$\log\left(\frac{G(t)}{G(t+a)}\right) = \log\left(\frac{A_0 \mathrm{e}^{-t\Delta E_1}}{A_0 \mathrm{e}^{-(t+a)\Delta E_1}}\right) = a\Delta E_1, \tag{3.12}$$

we have an expression for extracting the first energy level. While this technique is crude and limited in scope, it serves as a simple gateway to how one can extract masses and energy levels from correlators.

3.1.2 The path integral in quantum mechanics

The next step now is to connect the quantum mechanical propagator in eq. (3.10) on the previous page with the path integral. In order to build an intuitive understanding of the path integral, we imagine that we are drawing all possible paths a particle can take from an initial position (x_i, t_i) to final position (x_f, t_f). We then find the action $S[x(t)]$ for each of these paths and use that as a weight. Summing all of the contributions gives us the propagator $U(x_f, t_f; x_i, t_i)$ that describes the most likely path a particle will take,

$$U(x_f, t_f; x_i, t_i) = \sum_{\text{all paths}} \mathrm{e}^{-(\text{phase})}. \tag{3.13}$$

Instead of writing this as a sum over all paths, we can rewrite this in terms of an integral,

$$U(x_f, t_f; x_i, t_i) = \int \mathcal{D}x(t) \mathrm{e}^{-S[x(t)]}. \tag{3.14}$$

The $x(t)$ describes the path for $t = t_i \to t_f$, and $x(t_i) = x_i$ and $x(t_f) = x_f$ with $T = t_f - t_i$. The action in eq. (3.14) is given as

$$S[x] = \int_0^T \mathrm{d}t \left[\frac{m}{2} \dot{x}(t)^2 + V(x(t)) \right]. \tag{3.15}$$

As it turns out eq. (3.13) is equal to the quantum mechanical propagator in position space[51, 60],

$$\langle x_f | \, \mathrm{e}^{-T\hat{H}} \, | x_i \rangle = \int \mathcal{D}x(t) \mathrm{e}^{-S[x]}. \tag{3.16}$$

Given that we start and end at equal times, i.e. $x_i = x_f$, we have that eq. (3.16) is equal to the partition function Z. Using this, the quantum mechanical propagator can be written as

$$\langle x(t_2) x(t_1) \rangle = \frac{1}{Z} \int \mathcal{D}x(t) \mathrm{e}^{-S[x]} x(t_2) x(t_1). \tag{3.17}$$

In order to extract the energies from this expression as we did in the preceding section with eq. (3.11) on the previous page, we set $x_i = x_f$ and start discretizing the action.

$$\begin{aligned} S[x] &= \int \mathrm{d}t L[x(t), \dot{x}(t)] \\ &= \epsilon \sum_{j=0}^{N_{T-1}} \left(\frac{m}{2} \left(\frac{x_{j+1} - x_j}{\epsilon} \right)^2 + V(x_j) \right), \end{aligned} \tag{3.18}$$

where $\epsilon = T/N_T$ and is equal eq. (3.15) up to $\mathcal{O}(\epsilon)$. We used a simple *forward Euler* approximation to the time derivative, $\dot{x}$. To reduce the discretization error, we can for

instance instead use a symmetric derivative such as $\dot{x} = (x_{j+1} - x_{j-1})/2\epsilon$ such that the discretization error becomes $\mathcal{O}(\epsilon^2)$.

Discretizing the path from $t_i = t_0$ to $t_f = t_{N_T}$ changes the path integral measure to

$$\int \mathcal{D}x(t) \to \prod_{j=0}^{N_T-1} \int \mathrm{d}x_j, \tag{3.19}$$

such that the final path integral propagator to solve becomes,

$$\langle x| \, \mathrm{e}^{-T\hat{H}} \, |x\rangle = \frac{1}{Z} \prod_{j=0}^{N_{T-1}} \int \mathrm{d}x_j \exp\left(\epsilon \sum_{j=0}^{N_{T-1}} \left[\frac{m}{2} \left(\frac{x_{j+1} - x_j}{\epsilon} \right)^2 + V(x_j) \right] \right). \tag{3.20}$$

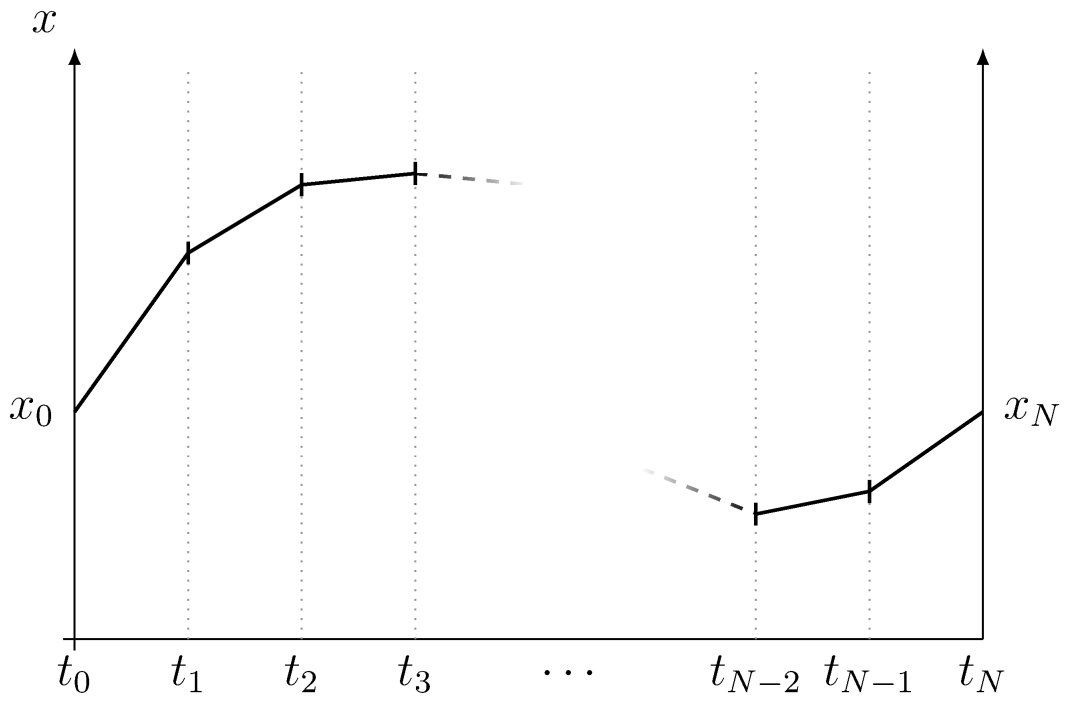


Figure 3.2: An example of the discretized path integral, going from time t_0 to t_{N_T}, where the end points is taken to be equal, $x_0 = x_{N_T}$. We integrate over all of space at each time t_i finding the most likely position at a given time.

A pictorial view of the path integral in eq. (3.20) can be seen in fig. 3.2. In this perspective, we only need to sample the path some N times at every time step, which can be done numerically[2].

[2]A quick implementation where we solve the quantum mechanical harmonic oscillator using the Metropolis algorithm can be found at the authors GitHub page `https://github.com/hmvege/QMPathIntegralSolver`.

To extract energies we use that the right hand side of eq. (3.17) on page 28 equals

$$\begin{aligned}\langle x(t)x(0)\rangle &= \frac{1}{Z}\int \mathrm{d}x\, \langle x|\, \mathrm{e}^{-(T-t)\hat{H}}\hat{x}\mathrm{e}^{-t\hat{H}}\hat{x}\,|x\rangle \\ &= \frac{\sum_n \mathrm{e}^{-TE_n}\,\langle E_n|\,\hat{x}\mathrm{e}^{-(\hat{H}-E_n)t}\hat{x}\,|E_n\rangle}{\sum_n \mathrm{e}^{-TE_n}}.\end{aligned} \tag{3.21}$$

Through an equivalent procedure as in the preceding section, we can extract the energy ΔE_1. We assume that the ground state will dominate, and by looking at different positions t in $x(t)$, we get

$$G(t) = \langle x(t)x(0)\rangle \overset{E_1 \gg E_2}{\to} a\Delta E_1 = \log\left(\frac{G(t)}{G(t+a)}\right). \tag{3.22}$$

Having reduced the path integral to a study of numerical integration, we are ready to move on the lattice QCD.

3.2 From QCD to LQCD: building a lattice field theory

Having shown how one can discretize a path integral in quantum mechanics, we are ready to scale up to the QCD path integral. While the discretization itself is more or less analogous to that of quantum mechanics, our point of attack will be in maintaining local gauge invariance as most of the prominent features of QCD are all bound up in the gauge invariance.

We begin by defining our spacetime to be a four-dimensional lattice, Λ,

$$\Lambda = \{n = (n_1, n_2, n_3, n_4) | n_1, n_2, n_3 = 0, \ldots, N-1\ ;\ n_4 = 0, \ldots, N_T - 1\}\,, \tag{3.23}$$

where $n = (n_1, n_2, n_3, n_4)$ defines a point on the lattice, N is the number of spatial points and N_T is the number of temporal points. The size of the lattice in terms of lattice sites is then given by

$$|\Lambda| = N^3 \times N_T. \tag{3.24}$$

Recalling the fermion fields we introduced in the previous chapter, we define fermions to live on the lattice points,

$$\psi(n)\ ,\ \bar{\psi}(n)\ ,\ n \in \Lambda, \tag{3.25}$$

where the physical position x can be recovered by multiplying n with the *lattice spacing* a, $x = an$. The determination of the lattice spacing will be made explicit in section 3.6.

Before we begin in earnest, we will also state that we are performing a field redefinition of absorbing the g_S into the gauge fields A_μ,

$$g_S A_\mu \to A_\mu. \tag{3.26}$$

We shall see that part of this motivation is that it makes g_S explicit in the lattice gauge action, which is useful when performing lattice calculations. It is also worth noting that g_S is the bare coupling, i.e. it has not been renormalized.

3.2.1 Deriving lattice action: First attempt

We begin by recalling the expression of the fermionic action in eq. (2.27) on page 15,

$$S_F[\psi, \bar{\psi}, A] = \int \mathrm{d}^4x \bar{\psi}(\gamma_\mu \partial_\mu + iA_\mu + m)\psi, \tag{3.27}$$

where we have omitted the flavor-, spinor- and color-indices f, α and c. We then split the fermion action into an interacting part and a non-interacting part,

$$\begin{aligned} S_F[\psi, \bar{\psi}, A] &= \int \mathrm{d}^4x \bar{\psi}(\gamma_\mu \partial_\mu + m)\psi + \int \mathrm{d}^4x \bar{\psi}(iA_\mu)\psi \\ &= S_F^0 + S_F^I. \end{aligned} \tag{3.28}$$

From this we discretise the fermion derivative with the symmetric derivative,

$$\partial_\mu \psi(x) \to \frac{\psi(n+\hat{\mu}) - \psi(n-\hat{\mu})}{2a}, \tag{3.29}$$

where a is the lattice spacing, and $\hat{\mu}$ is a directional index[3]. The non-interacting fermion action in eq. (3.28) becomes

$$\begin{aligned} S_F^0 &= a \sum_{n_1=0}^{N} a \sum_{n_2=0}^{N} a \sum_{n_3=0}^{N} a \sum_{n_4=0}^{N_T} \bar{\psi}(n) \left(\sum_{\mu=1}^{4} \gamma_\mu \frac{\psi(n+\hat{\mu}) - \psi(n-\hat{\mu})}{2a} + m\psi(n) \right) \\ &= a^4 \sum_{n \in \Lambda} \bar{\psi}(n) \left(\sum_{\mu=1}^{4} \gamma_\mu \frac{\psi(n+\hat{\mu}) - \psi(n-\hat{\mu})}{2a} + m\psi(n) \right). \end{aligned} \tag{3.30}$$

We are now ready to try to enforce gauge invariance. We can now attempt to apply the same gauge transformation as we performed in eq. (2.6) on page 7 to see if eq. (3.30) preserves gauge invariance. Let us define $\Omega(n) \in \mathrm{SU}(3)$, as opposite to $V(x) \in \mathrm{SU}(3)$ which we used for a non-discretized gauge transformation. The fields thus transform as

$$\psi(n) \to \psi'(n) = \Omega(n)\psi(n), \quad \bar{\psi}(n) \to \bar{\psi}'(n) = (\Omega(n)\psi(n))^\dagger \gamma_0 = \bar{\psi}(n)\Omega^\dagger(n). \tag{3.31}$$

If we now look back at the fermion derivative in eq. (3.29) we see that it is defined at neighboring points to n. In order for the non-interacting fermion action in eq. (3.30) to be gauge invariant, all the terms in the lattice action must be invariant. While the mass term is trivially invariant,

$$m\bar{\psi}(n)\psi(n) \to m\bar{\psi}'(n)\psi'(n) = m\bar{\psi}(n)\Omega^\dagger(n)\Omega(n)\psi(n) = m\bar{\psi}(n)\psi(n), \tag{3.32}$$

[3]One could think of $\hat{\mu}$ as a unit lattice vector of length a.

this is not the case for the discretized derivative,

$$\begin{aligned}\bar{\psi}(n)\psi(n+\hat{\mu}) &\to \bar{\psi}'(n)\psi'(n+\hat{\mu}) = \bar{\psi}(n)\Omega^\dagger(n)\Omega(n+\hat{\mu})\psi(n+\hat{\mu}),\\ \bar{\psi}(n)\psi(n-\hat{\mu}) &\to \bar{\psi}'(n)\psi'(n-\hat{\mu}) = \bar{\psi}(n)\Omega^\dagger(n)\Omega(n-\hat{\mu})\psi(n-\hat{\mu}),\end{aligned} \tag{3.33}$$

is we clearly do not retain gauge invariance.

Introducing the link variable U_μ

A remedy for maintaining gauge invariance was suggested in Wilson's famous paper [76, section III.A], which involves the introduction of *link variables*, $U_\mu(n)$. We define a link $U_\mu(n)$ as an object connecting two sites on the lattice, see fig. 3.3. The link gauge transform as

$$U_\mu(n) \to U'_\mu(n) = \Omega(n)U_\mu(n)\Omega(n+\hat{\mu})^\dagger, \tag{3.34}$$

connecting two points on the lattice.

(a) **(b)**

Figure 3.3: A link in the positive $\hat{\mu}$ direction fig. 3.3a, and a link in the negative $\hat{\mu}$ direction fig. 3.3b.

For a link going from n to $n-\hat{\mu}$, we define

$$U_{-\mu}(n) = U_\mu(n-\hat{\mu})^\dagger, \tag{3.35}$$

which transforms as

$$U_{-\mu}(n) \to U'_{-\mu}(n) = [\Omega(n-\hat{\mu})U_\mu(n-\hat{\mu})\Omega(n)^\dagger]^\dagger = \Omega(n)U_\mu(n-\hat{\mu})^\dagger\Omega(n-\hat{\mu})^\dagger. \tag{3.36}$$

If we now go back to our definition of the fermion derivative in eq. (3.29) on the previous page and require the discretized derivative to be a product of the link variable $U_\mu(n)$,

$$\partial_\mu\psi(x) \to \frac{U_\mu(n)\psi(n+\hat{\mu}) - U_{-\mu}(n)\psi(n-\hat{\mu})}{2a}, \tag{3.37}$$

we can try to apply a gauge transformation on the now modified expression and its constituent parts. For the forward derivative we get

$$\begin{aligned}\bar{\psi}(n)U_\mu(n)\psi(n+\hat{\mu}) &\to \bar{\psi}'(n)U'_\mu(n)\psi'(n+\hat{\mu})\\ &= \bar{\psi}(n)\Omega^\dagger(n)\Omega(n)U_\mu(n)\Omega(n+\hat{\mu})^\dagger\Omega(n+\hat{\mu})\psi(n+\hat{\mu})\\ &= \bar{\psi}(n)U_\mu(n)\psi(n+\hat{\mu}),\end{aligned} \tag{3.38}$$

and correspondingly for the backwards derivative,

$$\begin{aligned}\bar{\psi}(n)U_{-\mu}(n)\psi(n-\hat{\mu}) &\to \bar{\psi}'(n)U'_{-\mu}(n)\psi'(n-\hat{\mu}) \\ &= \bar{\psi}(n)\Omega^{\dagger}(n)\Omega(n)U_{\mu}(n-\hat{\mu})^{\dagger}\Omega(n-\hat{\mu})^{\dagger}\Omega(n-\hat{\mu})\psi(n-\hat{\mu}) \\ &= \bar{\psi}(n)U_{-\mu}(n)\psi(n-\hat{\mu}). \end{aligned} \tag{3.39}$$

With this redefinition of the derivative in eq. (3.37) on the facing page, our hope is that we can recover the full fermion action in the continuum $S_F = S_F^0 + S_F^I$. In general, this is naturally an expectation for every expression on the lattice. In order to investigate this, we need to properly define what the link $U_\mu(n)$ is.

3.2.2 The Gauge Link

The link variable $U_\mu(n)$ stems from the *gauge transporter*(or Wilson line) [51, ch. 15][24],

$$G(x,y) = \mathcal{P}\exp\left[i\int_P \mathrm{d}x^\mu A_\mu(x)\right], \tag{3.40}$$

where x and y are to points connected by some path P, while $\mathcal{P}$ path orders A_μ along the integration path. We have that this object transforms as

$$G(x,y) \to G'(x,y) = V(x)G(x,y)V(x)^{\dagger}. \tag{3.41}$$

From this follows that all objects consisting of closed paths P - called *Wilson loops*, are gauge invariant. To retrieve an expression for the link variable, we can start by taking the gauge transporter from x to $x+a$ and approximate the integral using the trapezoidal rule,

$$\begin{aligned} G(x,x+a) &= \exp\left[i\int_x^{x+a}\mathrm{d}^4x A_\mu(x)\right] \\ &= \exp\left[i\left((x_\mu+a)-x_\mu\right)\frac{1}{2}\left(A_\mu(x+\mu)+A_\mu(x)\right)\right] \\ &= \exp\left[\frac{ia}{2}\left(2A_\mu(x)+a\partial_\nu A_\mu(x)+A_\mu+\mathcal{O}(a^2)\right)\right] \\ &= \exp\left[iaA_\mu(x)+\mathcal{O}(a^2)\right], \end{aligned} \tag{3.42}$$

where we used that Taylor expanding around x in the gauge field for some a in a direction $\hat{\nu}$ is given as

$$A_\mu(x+a\hat{\nu}) = A_\mu(x) + a\partial_\nu A_\mu(x) + \mathcal{O}(a^2), \tag{3.43}$$

but to leading order $\mathcal{O}(a^2)$.

Moving on to the lattice, we can take the path from point $x=n$ to point $n+\hat{\mu}$,

$$G(n,n+\hat{\mu}) = \exp\left[iaA_\mu(n)+\mathcal{O}(a)\right], \tag{3.44}$$

from which we have the link given as

$$U_\mu(n) = \exp\left[iaA_\mu(n)\right]. \tag{3.45}$$

Performing an expansion of the exponential in $U_{-\mu}(n)$ and $U_\mu(n)$ in terms of a gives us

$$\begin{aligned} U_\mu(n) &= \exp\left[iaA_\mu(n)\right] = 1 + iaA_\mu(n) + \mathcal{O}(a^2) \\ U_{-\mu}(n) &= U_\mu(n-\hat{\mu})^\dagger = \exp\left[-iaA_\mu^\dagger(n-\hat{\mu})\right] = 1 - iaA_\mu(n-\hat{\mu}) + \mathcal{O}(a^2), \end{aligned} \tag{3.46}$$

where we used that any element of the Lie algebra SU(N) has the property of being Hermitian, $A_\mu^\dagger(x) = A_\mu(x)$. The expansion of the exponential in eq. (3.46) and the link definition in eq. (3.45) will be our primary tools for making sure there is a correspondence between the lattice expressions and the continuum ones.

Gauge invariant objects on the lattice

We have that fully connected paths $\mathcal{P}$ of links

$$P[U] = U_{\mu_0}(n_0)U_{\mu_1}(n_0 \pm \hat{\mu}_0)\cdots U_{\mu_{k-1}}(n_1 \pm \hat{\mu}_{k-1}) \equiv \prod_{n,\mu\in\mathcal{P}} U_\mu(n), \tag{3.47}$$

transform according to

$$P[U] \to P[U'] = \Omega(n_0)P[U]\Omega(n_1)^\dagger. \tag{3.48}$$

This is easy to see, since every link transform according to eq. (3.34) on page 32. Due to how we defined the links, if this object connects two fermions on the lattice $\bar{\psi}(n_0)$ and $\psi(n_1)$ it will be gauge invariant,

$$\bar{\psi}'(n_0)P[U']\psi'(n_1) = \bar{\psi}(n_0)P[U]\psi(n_1). \tag{3.49}$$

Another way of constructing gauge invariant objects os to build *closed* loops of links. If we define any closed loop $\mathcal{L}$ as

$$L[U] = \mathrm{tr}\left[\prod_{(n,\mu\in\mathcal{L})} U_\mu(n)\right], \tag{3.50}$$

we have that this is gauge invariant since any section will have

$$U_{\mu_k}(n_k \pm \mu_k)U_{\mu_{k+1}}(n_{k+1} \pm \hat{\mu}_{k+1}) \tag{3.51}$$

$$\to \Omega(n_k)U_{\mu_k}(n_k \pm \mu_k)\Omega(n_k \pm \hat{\mu}_k)^\dagger\Omega(n_k \pm \hat{\mu}_k)U_{\mu_{k+1}}(n_{k+1} \pm \hat{\mu}_{k+1})\Omega(n_{k+1} \pm \hat{\mu}_{k+1})^\dagger \tag{3.52}$$

$$= \Omega(n_k)U_{\mu_k}(n_k \pm \hat{\mu}_k)U_{\mu_{k+1}}(n_{k+1} \pm \hat{\mu}_{k+1})\Omega(n_{k+1} \pm \hat{\mu}_{k+1})^\dagger. \tag{3.53}$$

Since we are tracing, the cyclic property of matrix under trace will ensure that the end points will cancel as well, such that $L[U'] = L[U]$.

3.2.3 The naive fermion action

Let us begin by performing the much-awaited substitution of the fermion derivative in eq. (3.37) on page 32. This gives us what is called the *naive fermion action*,

$$S_F[\psi,\bar{\psi},U] = a^4 \sum_{n\in\Lambda} \bar{\psi}\left(\sum_{\mu=1}^{4}\gamma_\mu \frac{U_\mu(n)\psi(n+\hat{\mu}) - U_{-\mu}(n)\psi(n-\hat{\mu})}{2a} + m\psi(n)\right). \tag{3.54}$$

While we have shown that the constituent parts of this equation are gauge invariant in eq. (3.38) on page 32, eq. (3.39) on page 33 and in eq. (3.32) on page 31, we now need to show that we indeed recover the continuum expression once $a \to 0$. Having just introduced the link variable properly, this is an easy task to perform. If we insert the Taylor expanded link variables from eq. (3.46) on the facing page into the naive fermion action in eq. (3.54),

$$\begin{aligned}
&\frac{U_\mu(n)\psi(n+\hat{\mu}) - U_{-\mu}(n)\psi(n-\hat{\mu})}{2a} \\
&\quad= \frac{(1+iaA_\mu(n))\psi(n+\hat{\mu}) - (1-iaA_\mu(n-\hat{\mu}))\psi(n-\hat{\mu})}{2a} \\
&\quad= \frac{\psi(n+\hat{\mu}) - \psi(n-\hat{\mu})}{2a} + i\frac{A_\mu(n)\psi(n+\hat{\mu}) + A_\mu(n-\hat{\mu})\psi(n-\hat{\mu})}{2} \\
&\quad= \frac{\psi(n+\hat{\mu}) - \psi(n-\hat{\mu})}{2a} + i\frac{A_\mu(n)\psi(n) + A_\mu(n)\psi(n)}{2} + \mathcal{O}(a) \\
&\quad= \frac{\psi(n+\hat{\mu}) - \psi(n-\hat{\mu})}{2a} + iA_\mu(n)\psi(n) + \mathcal{O}(a).
\end{aligned} \tag{3.55}$$

In the second to last line we used that one can Taylor expand around n, such that $\psi(n\pm\hat{\mu}) = \psi(n) + \mathcal{O}(a)$ and $A(n\pm\hat{\mu}) = A(n) + \mathcal{O}(a)$.

If we now insert eq. (3.55) into the naive fermion action in eq. (3.54) and then send $a \to 0$, we get

$$S_F[\psi,\bar{\psi},U] = a^4 \sum_{n\in\Lambda} \bar{\psi}\left(\sum_{\mu=1}^{4}\gamma_\mu \frac{\psi(n+\hat{\mu}) - \psi(n-\hat{\mu})}{2a} + iA_\mu(n)\psi(n) + m\psi(n)\right) + \mathcal{O}(a) \tag{3.56}$$

$$\overset{a\to 0}{=} \int \mathrm{d}^4x \bar{\psi}(\not{D} + m)\psi, \tag{3.57}$$

which recovers the continuum action in eq. (2.27) on page 15.

Before we move on to see if we can perform a similar link expansion of the gauge action S_G, it is worth mentioning that the name of the naive fermion action stems the fact that it gives rise to *doublers*. These can be seen if one Fourier transforms the naive fermion action in eq. (3.54) to momentum space and we get 15 additional mass terms [24].

3.2.4 The Wilson gauge action

We begin by recalling the pure gauge action of QCD in eq. (2.26) on page 14 with the field strength tensor explicitly written out as

$$S_G[A] = \frac{1}{2g_S^2}\int \mathrm{d}^4x \mathrm{tr} F_{\mu\nu}^2 = \frac{1}{2g_S^2}\int \mathrm{d}^4x \mathrm{tr}\left(\partial_\mu A_\nu - \partial_\nu A_\mu + i[A_\mu, A_\nu]\right), \tag{3.58}$$

where we have a factor $1/g_S^2$ in front of the integral due to the field redefinition $g_S A_\mu \to A_\mu$. Since we now are moving onto the lattice we will denote the lattice field strength tensor as $F_{\mu\nu}(n)$, where n is the lattice site.

To build a lattice gauge action corresponding to eq. (3.58), we start with the simplest possible gauge invariant object we can imagine - a square.

$$\begin{aligned} P_{\mu\nu}(n) &= U_\mu(n)U_\nu(n+\hat{\mu})U_{-\mu}(n+\hat{\mu}+\hat{\nu})U_{-\nu}(n+\hat{\nu}) \\ &= U_\mu(n)U_\nu(n+\hat{\mu})U_\mu(n+\hat{\nu})^\dagger U_\nu(n)^\dagger, \end{aligned} \tag{3.59}$$

and is called the *plaquette*. A pictorial view of how it is built can be seen in fig. 3.4.

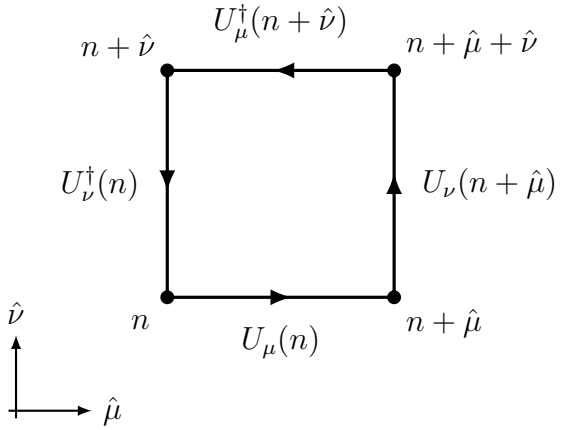


Figure 3.4: The plaquette as given by eq. (3.59).

Our hope is that using the link definition in eq. (3.45) on page 34 we can recover the continuum gauge action. We begin by substituting the terms in the plaquette from eq. (3.59) by the link definition in eq. (3.45) on page 34,

$$\begin{aligned} P_{\mu\nu}(n) &= U_\mu(n)U_\nu(n+\hat{\mu})U_\mu(n+\hat{\nu})^\dagger U_\nu(n)^\dagger \\ &= \exp\left(iaA_\mu(n)\right)\exp\left(iaA_\nu(n+\hat{\mu})\right)\exp\left(-iaA_\mu(n+\hat{\nu})\right)\exp\left(-iaA_\nu(n)\right). \end{aligned} \tag{3.60}$$

The next trick is to use the Baker-Campbell-Hausdorff formula for two matrices A and B,

$$\exp A \exp B = \exp\left(A + B + \frac{1}{2}[A,B] + \ldots\right), \tag{3.61}$$

in order to join the exponentials of eq. (3.60) on the facing page. We get

$$(3.60) = \exp\left(iaA_\mu(n) + iaA_\nu(n+\hat{\mu}) + \frac{1}{2}\left[iaA_\mu(n), iaA_\nu(n+\hat{\mu})\right] + \cdots\right) \times \exp\left(-iaA_\mu(n+\hat{\nu}) - iaA_\nu(n) + \frac{1}{2}\left[-iaA_\mu(n+\hat{\nu}), -iaA_\nu(n)\right] + \cdots\right). \tag{3.62}$$

If we now use that we are only keeping terms less than $\mathcal{O}(a^3)$ and that the $A_\mu(n+\hat{\nu})$ can be Taylor expanded as

$$A_\mu(n+\hat{\nu}) = A_\mu(n) + a\partial_\nu A_\mu(n) + \mathcal{O}(a^2), \tag{3.63}$$

which similar to the continuum expression in eq. (3.43) on page 33, we can join the exponentials in eq. (3.62) to

$$(3.62) = \exp\Bigg(iaA_\mu(n) + iaA_\nu(n) + ia^2\partial_\nu A_\mu(n) - \frac{a^2}{2}\left[A_\mu(n), A_\nu(n)\right] - iaA_\mu(n) - ia^2\partial_\nu A_\mu(n) - iaA_\nu(n) - \frac{a^2}{2}\left[A_\mu(n), A_\nu(n)\right] + \frac{a^2}{2}\left[A_\mu(n), A_\mu(n)\right] + \frac{a^2}{2}\left[A_\mu(n), A_\nu(n)\right] + \frac{a^2}{2}\left[A_\nu(n), A_\mu(n)\right] + \frac{a^2}{2}\left[A_\nu(n), A_\nu(n)\right] + \mathcal{O}(a^3)\Bigg), \tag{3.64}$$

which after canceling terms and using $[A,B] = -[B,A]$ reduces to

$$(3.64) = \exp\left[ia^2\left(\partial_\nu A_\mu(n) - \partial_\nu A_\mu(n) + i\left[A_\mu(n), A_\nu(n)\right]\right)\right] \tag{3.65}$$

$$= \exp\left[ia^2 F_{\mu\nu}(n) + \mathcal{O}(a^3)\right]. \tag{3.66}$$

Taylor expanding this expression gives us

$$P_{\mu\nu} = 1 + ia^2 F_{\mu\nu}(n) - \frac{a^4}{2}F_{\mu\nu}^2(n) + \mathcal{O}(a^6). \tag{3.67}$$

From this we can can observe that taking the real and imaginary parts yields

$$\mathrm{Re}(P_{\mu\nu}) = 1 - \frac{a^4}{2}F_{\mu\nu}^2(n) + \mathcal{O}(a^6), \tag{3.68}$$

$$\mathrm{Im}(P_{\mu\nu}) = a^2 F_{\mu\nu}(n) + \mathcal{O}(a^6). \tag{3.69}$$

Since we are only interested in the field strength tensor squared, and we know that the trace of the field strength tensor yields the gauge action, we can take unity subtracted by the plaquette as the lattice gauge action,

$$S_G[U] = \frac{2}{g_S^2}\sum_{n\in\Lambda}\sum_{\mu<\nu}\mathrm{Re}\,\mathrm{tr}\left[1 - P_{\mu\nu}(n)\right] \tag{3.70}$$

$$= \frac{a^4}{2g_S^2}\sum_{n\in\Lambda}\sum_{\mu,\nu}\left(\mathrm{tr}\left[F_{\mu\nu}^2(n)\right] + \mathcal{O}(a^2)\right), \tag{3.71}$$

in which the extra factor 2 in eq. (3.70) on the previous page comes from using that

$$\mathrm{Re}\,\mathrm{tr}(P_{\mu\nu}(n)) = \mathrm{Re}\,\mathrm{tr}(P_{\nu\mu}(n)), \tag{3.72}$$

and thus reducing the plaquettes needed to six. Inserting the plaquette and taking the real part as in eq. (3.68) on the preceding page gives us eq. (3.71) on the previous page where we can see that $a \to 0$ recovers the continuum limit of the gauge action.

Usually, we denote the coupling constant as $\beta = 2N_c/g_S^2$, which gives us a final expression for the gauge action for $N_c = 3$,

$$S_G[U] = \frac{\beta}{3}\sum_{n\in\Lambda}\sum_{\mu<\nu}\mathrm{Re}\,\mathrm{tr}\left[1 - P_{\mu\nu}(n)\right], \tag{3.73}$$

This action is also sometimes called the *Wilson gauge action* [76], and is the simplest way of expressing the gauge action.

Improvements to the action

As a final note, it is worth mentioning that the gauge action in eq. (3.73) can be improved in several manners, most notable is through the Symanzik improvement program [24, ch. 9]. More details on this will come towards the end of this chapter in section 3.7 on page 54.

3.3 The LQCD partition function

Having introduced the lattice action $S[\bar{\psi}, \psi, U]$ for both fermions $S_F[\bar{\psi}, \psi, U]$ and gluons $S_G[U]$, it is time to make the path integral on the lattice explicit,

$$Z = \int \mathcal{D}\bar{\psi}\mathcal{D}\psi\mathcal{D}U\mathrm{e}^{-S_F[\bar{\psi},\psi,U]-S_G[U]}, \tag{3.74}$$

where the Haar measure ensure that $\mathcal{D}(U)$ is gauge invariant [24, ch. 3.1].

As we already have discussed in chapter 2 on page 5, the fermions are Grassmann valued numbers and we can integrate them out of the partition function and expectancy-value as determinants, with one determinant for each fermion flavor. The partition function becomes

$$Z = \int \mathcal{D}U\mathrm{e}^{-S_G[U]}\prod_f \det\left[D^{(f)}\right], \tag{3.75}$$

where $D^{(f)}$ is the Dirac operator containing the fermion mass.

The expectation value when including fermion fields is given by

$$\langle O\rangle = \frac{1}{Z}\int \mathcal{D}U\mathrm{e}^{-S_G[U]}\mathcal{D}\psi\mathcal{D}\bar{\psi}\mathrm{e}^{-S_F[\psi,\bar{\psi},U]}O[\psi,\bar{\psi},U]. \tag{3.76}$$

Integrating out the Grassmann-valued fields ψ and $\bar{\psi}$ gives us

$$\begin{aligned} \langle O \rangle &= \frac{1}{Z} \int \mathcal{D}U \mathrm{e}^{-S_G[U]} O[U] \det D[U], \\ Z &= \int \mathcal{D}U \mathrm{e}^{-S_G[U]} \det D[U], \end{aligned} \tag{3.77}$$

where $D[U]$ is a matrix of $12|\Lambda|$ rows and columns, dependent on the quark flavor f. If we wish to compute correlators, we will have to take the inverse of the Dirac operator $D[U]$ as our propagator.

Since we are only interested in pure gauge theory the fermion determinants are set to one, such that we are essentially overlooking any fermionic loops. This is called the *quenched approximation*. While one can still calculate correlators that include fermions, the effects of sea quarks or dynamical fermions are neglected(ignoring vacuum fluctuations of quarks). The quarks used as propagators is called valence quarks. For the two lightest quarks, u and d, we have that their masses are taken to be approximately equal, thus reducing the number of determinants needed to compute to one, as they can be taken as equal[4].

The integral we are left with after setting the determinants to one in eq. (3.77), is an integral over the gauge fields. For this thesis, all quark effects will be ignored, such that we essentially have a purely gluonic theory, or pure gauge or Yang-Mills theory. Our goal will be to generate gauge configurations such that we can use those when sampling expectation values $O[U]$.

3.4 Correlators and observables

Having briefly introduced fermions on the lattice just to wave them away, we can begin to look at some of the observables and correlators that will be investigated in this thesis. When ignoring quarks, the expectation value for O as we saw in eq. (3.77), becomes

$$\begin{aligned} \langle O \rangle &= \frac{1}{Z} \int \mathcal{D}U \mathrm{e}^{-S_G[U]} O[U], \\ Z &= \int \mathcal{D}U \mathrm{e}^{-S_G[U]}, \end{aligned} \tag{3.78}$$

where U indicates all of the links on the lattice and $S_G[U]$ is given by eq. (3.73) on the preceding page.

[4]The number of determinants/quarks in the simulation is denoted by $N_f = n_i + n_j + \cdots$. Each standalone number indicates a unique determinant, while the number itself indicates the number of times we are to include the determinant. E.g. $N = 2 + 1 + 1$ means we are to take the u/d determinants to be equal(only need to calculate one determinant and square it), while s and c are unique and each needs to be calculated explicitly.

A basic example of an observable is the plaquette observable. Having introduced the plaquette in eq. (3.59) on page 36, the plaquette on the lattice is given by

$$P = \frac{1}{16|\Lambda|} \sum_{n\in\Lambda} \sum_{\mu<\nu} \mathrm{Re}\,\mathrm{tr} P_{\mu\nu}, \tag{3.79}$$

where $|\Lambda|$ is the total number of lattice points. The final result will be an average of all of the gauge configurations.

As discussed in chapter 2 on page 5, understanding topology in QCD can give profound insights into its nature. In this section, we will reintroduce quantities such as topological charge and susceptibility, only this time as they are defined on the lattice. For this task, we will rely on an improved lattice field strength tensor, *the clover*.

3.4.1 The clover field strength tensor

The clover consists of four plaquettes organized in a square, resembling a clover(see fig. 3.5 on the facing page). It has discretization errors of $\mathcal{O}(a^2)$, and is given as following [3, 61]

$$\begin{aligned} C_{\mu\nu} \propto \Big(& U_\mu(n) U_\nu(n+\hat{\mu}) U_\mu^\dagger(n+\hat{\nu}) U_\nu^\dagger(n) \\ & - U_\mu(n) U_\nu^\dagger(n+\hat{\mu}-\hat{\nu}) U_\mu^\dagger(n-\hat{\nu}) U_\nu(n-\hat{\nu}) \\ & - U_\mu^\dagger(n-\hat{\mu}) U_\nu^\dagger(n-\hat{\mu}-\hat{\nu}) U_\mu(n-\hat{\mu}-\hat{\nu}) U_\nu(n-\hat{\nu}) \\ & + U_\mu^\dagger(n-\hat{\mu}) U_\nu(n-\hat{\mu}) U_\mu(n-\hat{\mu}+\hat{\nu}) U_\nu^\dagger(n) \Big), \end{aligned} \tag{3.80}$$

from which we construct an anti-Hermitian clover term,

$$\tilde{C}_{\mu\nu} \propto -i \left[\left(C_{\mu\nu}(n) - C_{\mu\nu}^\dagger(n) \right) - \frac{1}{3} \mathrm{tr} \left(C_{\mu\nu}(n) - C_{\mu\nu}^\dagger(n) \right) \right]. \tag{3.81}$$

The proportionality sign in eq. (3.80) indicates we need to divide by 4 since we have four plaquettes, while in eq. (3.81) we pick up a factor $\frac{1}{2}$ since we are making it anti-Hermitian. Bringing these factors into a single term, the clover field strength tensor becomes

$$F_{\mu\nu}^{\mathrm{clov}}(n) = \frac{1}{8a^2} \tilde{C}_{\mu\nu}, \tag{3.82}$$

with $F_{\mu\nu}^{\mathrm{clov}} = -F_{\nu\mu}^{\mathrm{clov}}$.

With the clover in place, we can move on to define all of the relevant gauge observables.

3.4.2 Energy density

While we won't go into detail on the energy density(also known as the action density), we will use it for studying scale setting on the lattice. We will return to the scale setting in section 3.6.3 on page 53.

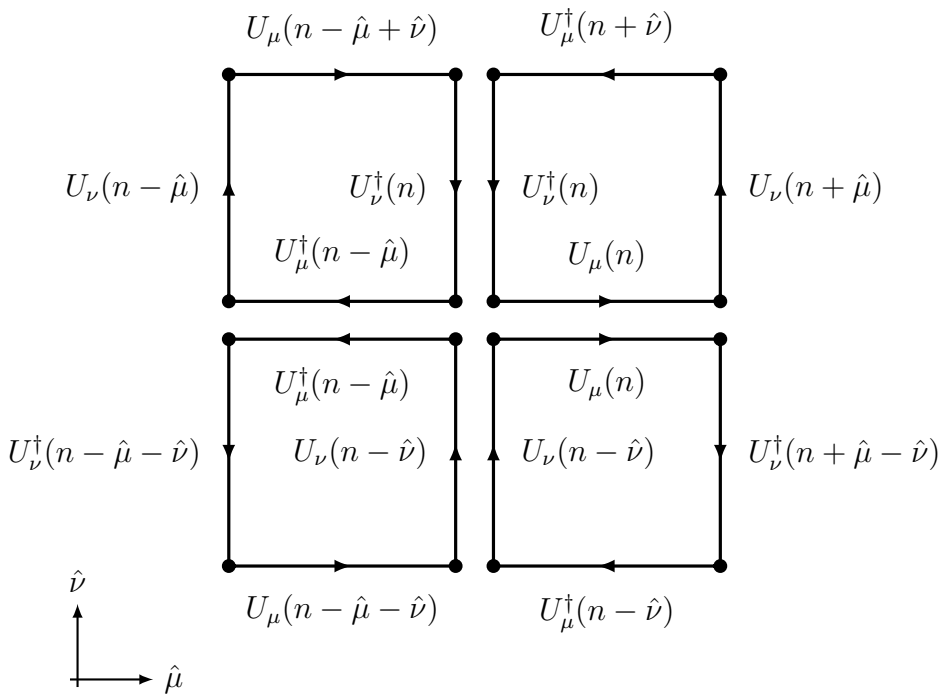


Figure 3.5: The maximal symmetric clover expression for the field strength tensor from eq. (3.82) on the facing page.

The energy density is given as

$$\begin{aligned} E &= \frac{a^4}{2|\Lambda|}\sum_{n\in\Lambda}\sum_{\mu,\nu}\left(F_{\mu\nu}^{\text{clov}}(n)\right)^2 \\ &= \frac{a^4}{|\Lambda|}\sum_{n\in\Lambda}\sum_{\mu<\nu}\left(\frac{1}{64a^4}\tilde{C}_{\mu\nu}\tilde{C}^{\mu\nu}\right)^2 \\ &= -\frac{1}{64|\Lambda|}\sum_{n\in\Lambda}\sum_{\mu<\nu}\left(\tilde{C}_{\mu\nu}(n)\right)^2, \end{aligned} \tag{3.83}$$

where we in the second line inserted the clover field strength terms from eq. (3.82) on the facing page. We get an extra factor 2 from only needing to sum over each clover term once. The definition we use for the energy density on the lattice is given by eq. (3.83).

3.4.3 Topological charge

The topological charge in eq. (2.44) on page 20 on the lattice is given as

$$Q = a^4\sum_{n\in\Lambda} q(n), \tag{3.84}$$

with the charge density given by

$$q(n) = \frac{1}{32\pi^2}\epsilon_{\mu\nu\rho\sigma}\text{tr}\left[F_{\mu\nu}(n)F_{\rho\sigma}(n)\right]. \tag{3.85}$$

The g_S in the continuum expression in eq. (2.45) on page 20 has been absorbed by the field redefinition from eq. (3.26) on page 30. Due to the fact that

$$\text{tr}[F_{\mu\nu}F_{\rho\sigma}] = \text{tr}[F_{\rho\sigma}F_{\mu\nu}] = \text{tr}[F_{\sigma\rho}F_{\nu\mu}] \tag{3.86}$$
$$\text{tr}[F_{\mu\nu}F_{\sigma\rho}] = \text{tr}[F_{\sigma\rho}F_{\mu\nu}] = \text{tr}[F_{\rho\sigma}F_{\nu\mu}], \tag{3.87}$$

the number of tensor index combinations we need to calculate goes from 24 to 3, and $q(n)$ picks up a factor 8. Writing out all of the Levi-Civita combinations and inserting the clover field strength in eq. (3.80) on page 40, gives us

$$q(n) = \frac{1}{4\pi^2}\text{tr}\left[F_{01}^{\text{clov}}(n)F_{23}^{\text{clov}}(n) + F_{02}^{\text{clov}}(n)F_{13}^{\text{clov}}(n) + F_{03}^{\text{clov}}(n)F_{12}^{\text{clov}}(n)\right]. \tag{3.88}$$

The topological charge density is then

$$q(n) = \frac{1}{4\pi^2}\text{tr}\left[\frac{1}{(8a^2)^2}\left(\tilde{C}_{01}(n)\tilde{C}_{23}(n) + \tilde{C}_{02}(n)\tilde{C}_{13}(n) + \tilde{C}_{03}(n)\tilde{C}_{12}(n)\right)\right] \tag{3.89}$$
$$q(n) = -\frac{1}{a^4}\frac{1}{256\pi^2}\text{tr}\left(\tilde{C}_{01}(n)\tilde{C}_{23}(n) + \tilde{C}_{02}(n)\tilde{C}_{13}(n) + \tilde{C}_{03}(n)\tilde{C}_{12}(n)\right). \tag{3.90}$$

Summing over the lattice gives us the expression for the total topological charge,

$$\begin{aligned} Q &= a^4\sum_{n\in\Lambda} q(n) \\ &= -\frac{1}{256\pi^2}\sum_{n\in\Lambda}\text{tr}\left(\tilde{C}_{01}(n)\tilde{C}_{23}(n) + \tilde{C}_{02}(n)\tilde{C}_{13}(n) + \tilde{C}_{03}(n)\tilde{C}_{12}(n)\right). \end{aligned} \tag{3.91}$$

Momenta of topological charge

The topological charge posses several interesting properties as discussed in section 2.10, and we can through investigations of the topological charge cumulants achieve some key insights. For an explicit derivation the cumulants, see appendix E on page 159. The first cumulant - the average, is given as

$$\langle Q\rangle = \frac{1}{N_{\text{MC}}}\sum_{i=0}^{N_{\text{MC}}} Q_i. \tag{3.92}$$

It follows from Q being parity odd, that the expectancy value becomes zero. If the average topological charge is not zero, it may be an indication that we are not sampling the configuration space correctly.

Topological susceptibility and the Witten-Veneziano formula

The second cumulant correspond to topological susceptibility which we introduced in section 2.10.2 on page 19. Using the normalized second cumulant[5], we have [24]

$$\chi_{\text{top}} = \frac{1}{V}\left\langle Q^2 \right\rangle = \frac{a^4}{|\Lambda|}\sum_{m,n}\langle q(m)q(n)\rangle = a^4\sum_{n}\langle q(0)q(n)\rangle\,, \tag{3.93}$$

where the $V = a^4|\Lambda|$ is the physical volume of the lattice[6]. Taking the quartic root of the susceptibility gives us

$$\chi_{\text{top}}^{1/4} = \frac{1}{V^{1/4}}\left\langle Q^2 \right\rangle^{1/4}. \tag{3.94}$$

Applying this to the Witten-Veneziano formula from eq. (2.59) on page 23 is straight forward, and can provide insight into the η' mass or the number of flavors N_f involved.

The Fourth Cumulant

While the third cumulant is zero since $\langle Q\rangle = 0$, and all the terms that are given in eq. (E.12) on page 160 contains the first moment, the *fourth cumulant* in eq. (E.16) on page 161 is not zero. The fourth cumulant has been investigated in several papers [13, 19, 26, 72] and is given as

$$\left\langle Q^4 \right\rangle_c = \frac{1}{V^2}\left(\left\langle Q^4\right\rangle - 3\left\langle Q^2\right\rangle^2\right). \tag{3.95}$$

From this, we can also measure the ratio R,

$$R = \frac{\left\langle Q^4\right\rangle_c}{\frac{1}{V}\left\langle Q^2\right\rangle} = \frac{1}{V}\frac{\left\langle Q^4\right\rangle - 3\left\langle Q^2\right\rangle^2}{\left\langle Q^2\right\rangle}, \tag{3.96}$$

which can provide insight into the distribution of Q and whether or not it is Gaussian. If this ratio is far away from 1, one can rule out the θ-behavior predicted by the dilute instanton gas model. It implies that the fluctuations of the topological charge are quantum non-perturbative and compatible with being suppressed by $1/N_c^2$ in the large N_c limit [13], i.e. QCD can not be described by the dilute instanton gas model.

3.4.4 Correlators

Correlators in lattice QCD provides an entry point to investigating the masses and excited states of objects on the lattice. We can take the expression for the quantum mechanical correlator and instead of time t look at the temporal points on the lattice $t = an_t$,

$$C(n_t) = \left\langle \hat{O}_2(\mathbf{0}, n_t)\hat{O}_1(\mathbf{0}, 0)\right\rangle = \sum_k \langle 0|\,\hat{O}_2\,|k\rangle\,\langle k|\,\hat{O}_1\,|0\rangle\,\mathrm{e}^{-n_t E_k}, \tag{3.97}$$

[5]The topological charge is an extensive quantity, which is why we divide by the lattice volume.

[6]If we wish to convert to physical units, we can use the conversion unit given in eq. (A.3) on page 135.

in which we can extract the ground states similar to eq. (3.12) on page 27,

$$C(n_t) = A_0 \mathrm{e}^{-n_t E_0} + A_1 \mathrm{e}^{-n_t E_1} + \dots . \tag{3.98}$$

There exist a small collection of more advanced methods for extracting energy levels. An example is the variational method, which converts to extraction issue into an eigenvalue problem [44]. We refer the reader to books such as Gattringer and Lang [24, ch. 6] for an overview on the different methods available.

The topological charge density correlator

If we take the glueball correlation function in eq. (2.64) from section 2.10.5, we can extract an effective mass for the state $P = 0^{-+}$,

$$am_{\text{eff}} = \log\left(\frac{C(n_t)}{C(n_t+1)}\right), \tag{3.99}$$

where $C(n_t)$ is taken to be

$$C(n_t) = \langle q(n_t) q(0) \rangle \,, \tag{3.100}$$

with $q(n_t)$ being the topological charge density from eq. (3.85) on page 42 in Euclidean time. Note that we have summed $q(n_t)$ in x, y and z, leaving Euclidean time t_e available to us. Another way of representing this correlator is given as

$$C(t_e) = \langle q(t_e) q(t_{e,0}) \rangle \,, \tag{3.101}$$

where we use $C(t_e)$ as a function of Euclidean time t_e, and the source $q(t_{e,0})$ is frozen at some euclidean time $t_{e,0}$. The value t_e is $t_e = an_t$.

3.5 Gradient flow

A problem with studying gauge configurations is that they contain noise and that several observables exhibit UV divergencies. An example of this would be the topological susceptibility, which at zero flow time contains a UV divergence [55]. A solution to this problem is to apply *gradient flow*. The idea behind gradient flow is to introduce a fifth dimension or scale, called *flow time* t_f which has dimension two, and let the fields evolve according to what in the crudest manner can be seen as a "diffusion" equation. This makes makes the gauge field into a smooth, renormalized field at $t_f > 0$ [40]. Applying gradient flow separates the gauge field into topologically distinct sectors, and allows us to sample quantities of topological charge without having to apply a renormalization procedure.

The flow of the SU(3) gauge fields are denoted by $B_\mu(x, t_f)$ which are Lie algebra valued gauge fields,

$$\frac{\mathrm{d}}{\mathrm{d}t_f} B_\mu(x, t_f) = D_\nu G_{\nu\,\mu}(x, t_f), \tag{3.102}$$

$$D_\mu = \partial_\mu + [B_\mu(x, t_f), \cdot] \,, \tag{3.103}$$

$$G_{\mu\nu}(x, t_f) = \partial_\mu B_\nu(x, t_f) - \partial_\nu B_\mu(x, t_f) - i[B_\mu(x, t_f), B_\nu(x, t_f)], \tag{3.104}$$

with the initial condition of eq. (3.102) on the facing page being the fundamental gauge field,

$$B_\mu(x,t_f)|_{t_f=0} = A_\mu(x). \tag{3.105}$$

The $\cdot$ of the covariant derivative refers to the derivative with respect to gauge fields.

We have that eq. (3.102) on the preceding page is proportional to the gradient of the action [42],

$$\frac{\mathrm{d}}{\mathrm{d}t_f}B_\mu(x,t_f) \sim -\frac{\partial S_G(B)}{\partial B_\mu(x,t_f)}, \tag{3.106}$$

such that the system will be driven towards the stationary points of the Yang-Mills gauge action, making the gauge field smoother. This allows us to investigate topological properties such as instanton effects as they by definition are local minimums of the gauge field. From eq. (3.102) on the facing page we see that flow time t_f is required to have dimensionality of -2 since D_ν and $G_{\nu,\mu}(x,t_f)$ have respectively dimension $+1$ and $+2$, and $B_\mu(x,t_f)$ has dimension $+1$.

We will later solve the flow differential equation numerically using Runge-Kutta 3 in section 4.2 on page 64, but before that it is worth taking a look at another possibility. We can attempt to solve the flow differential equation by invoking perturbation theory and reintroducing the coupling g_S into the gauge fields. Then, by expanding B_μ in an asymptotic series $B_\mu = \sum_{k=1}^{\infty} g_S^k B_{\mu,k}$ where $B_{\mu,k}|_{t=0} = \delta_{k1}A_\mu$, the leading order coefficient $B_{\mu,1}$ is equivalent with a heat equation with the solution [40]

$$B_{\mu,1}(x,t_f) = \int \mathrm{d}^4 y K_{t_f}(x-y)A_\mu(y), \tag{3.107}$$

with $K_{t_f}(x-y)$ being the heat kernel,

$$K_{t_f}(z) = \int \frac{\mathrm{d}p^4}{(2\pi)^4}\mathrm{e}^{ipz}\mathrm{e}^{-t_f p^2}. \tag{3.108}$$

Completing the square of the exponentials,

$$ipz - t_f p^2 = ipz - t_f p^2 + \frac{z^2}{4t_f} - \frac{z^2}{4t_f} \tag{3.109}$$

$$= \left(\frac{z}{2\sqrt{t_f}} + i\sqrt{t_f}p\right)^2 - \frac{z^2}{4t_f}, \tag{3.110}$$

allows us to rewrite eq. (3.108) as

$$K_{t_f}(z) = \mathrm{e}^{-\frac{z^2}{4t_f}} \int \frac{\mathrm{d}p^4}{(2\pi)^4} \exp\left(\frac{z}{2\sqrt{t_f}} + i\sqrt{t_f}p\right)^2. \tag{3.111}$$

If we now use the Gaussian integral,

$$\int_{-\infty}^{\infty} \mathrm{d}x \mathrm{e}^{-a(x+b)^2} = \sqrt{\frac{\pi}{a}}, \tag{3.112}$$

we get that eq. (3.111) on the previous page becomes

$$\begin{aligned} K_{t_f}(z) &= \frac{\mathrm{e}^{-z^2/4t_f}}{(2\pi)^4} \frac{\pi^2}{t_f^2} \\ &= \frac{\mathrm{e}^{-z^2/4t_f}}{(4\pi t_f)^2}. \end{aligned} \tag{3.113}$$

We recognize eq. (3.113) to be a Gaussian distribution, such that $\sigma^2 = 2t_f$ is the variance in four dimensions. The smearing radius is given by the root mean square radius of the heat kernel [40, p. 4],

$$\sqrt{\langle z^2 \rangle} = \left(\int \mathrm{d}^4 z \frac{z^2 \mathrm{e}^{-z^2/4t_f}}{(4\pi t_f)^2} \right)^{1/2} \tag{3.114}$$

$$= \left(\frac{1}{(4\pi t_f)^2} \int \mathrm{d}^4 z z^2 \mathrm{e}^{-z^2/4t_f} \right)^{1/2}, \tag{3.115}$$

where $z^2 = z_1^2 + z_2^2 + z_3^2 + z_4^2$. Using the Gaussian integral listed earlier in eq. (3.112) with $b = 0$ and

$$\int_{-\infty}^{\infty} \mathrm{d}x x^2 \mathrm{e}^{-ax^2} = \frac{1}{2}\sqrt{\frac{\pi}{a^3}}, \tag{3.116}$$

we get when we set $a = 1/4t_f$,

$$\sqrt{\langle z^2 \rangle} = \left[\frac{1}{(\pi/a)^2} \int \mathrm{d}z_1 \mathrm{d}z_2 \mathrm{d}z_3 \mathrm{d}z_4 \left(z_1^2 + z_2^2 + z_3^2 + z_4^2 \right) \mathrm{e}^{-a\left(z_1^2+z_2^2+z_3^2+z_4^2\right)} \right]^{1/2} \tag{3.117}$$

$$= \left[\left(\frac{a}{\pi} \right)^2 4 \left(\frac{1}{2} \sqrt{\frac{\pi}{a^3}} \left(\sqrt{\frac{\pi}{a}} \right)^3 \right) \right]^{1/2} \tag{3.118}$$

$$= \left[2 \left(\frac{a}{\pi} \right)^2 \frac{\pi^2}{a^3} \right]^{1/2} \tag{3.119}$$

$$= \sqrt{8t_f}. \tag{3.120}$$

This is our flow smearing radius $\sqrt{8t_f}$, which we will use to measure the how far we have let the gradient flow run. If we let the flow run long enough, the smearing $\sqrt{8t_f}$ will at some point be larger than the lattice itself. This will result in unphysical results, and is something we will avoid.

As it turns out, gauge correlation functions at $t_f > 0$ are finite at all Euclidean distances, such that no renormalization except for the gauge coupling(and quark masses) is required [42].

On the lattice the flow equation from eq. (3.102) on page 44 becomes

$$\dot{V}_{t_f}(x,\mu) = -g_S^2 \left\{\partial_{x,\mu} S_G[V_{t_f}]\right\} V_{t_f}(x,\mu), \quad V_{t_f}(x,\mu)|_{t_f=0} = U(x,\mu), \tag{3.121}$$

where $V_{t_f}(x,\mu)$ is the flow field associated with the fundamental gauge field $U_\mu(x)$. $\partial_{x,\mu}$ is the $\mathfrak{su}(3)$ differential operator, acting on the action. Our goal will be to integrate this function. To do so, several ingredients are needed. We need to set up a numerical integration scheme of choice - in our case, this will be the *Runge-Kutta 3* method which will be discussed in the next chapter. Another important component is the derivative of the gauge action.

It should be mentioned that it is also possible to flow quarks on the lattice, but since we are dealing with a pure gauge theory we will simply refer the reader to [41] for a treatment of this topic.

3.5.1 The action derivative

In eq. (3.121) we are to take derivative of the gauge action using the $\mathfrak{su}(3)$ valued differential operator. We start with defining the differential operator acting on some function $f(U)$ of the gauge field U,

$$\partial_{x,\nu} f(U) = t^a \partial^a_{x,\mu} f(U), \tag{3.122}$$

with the algebra-valued differential operator being

$$\partial^a_{x,\mu} f(U) = \frac{\mathrm{d}}{\mathrm{d}s} f\left(\mathrm{e}^{sX} U\right)\big|_{s=0}, \quad \text{with} \quad X(y,\nu) = \begin{cases} t^a & \text{if} \quad (y,\nu) = (x,\mu), \\ 0 & \text{else.} \end{cases} \tag{3.123}$$

For notation we will use x, y instead of n, m for lattice locations.

Taking the derivative of $S_G[U]$ yields

$$\partial_{x,\mu} S_G[U] = t^a \partial^a_{x,\mu} S_G[U] = \frac{\mathrm{d}}{\mathrm{d}s} S_G\left[\mathrm{e}^{sX} U\right]\big|_{s=0} \tag{3.124}$$

$$= t^a \frac{\mathrm{d}}{\mathrm{d}s} \left\{ \frac{\beta}{3} \sum_y \sum_{\alpha<\beta} \mathrm{Re}\,\mathrm{tr}\left(1 - \mathrm{e}^{sX} U_\alpha(y) U_\beta(y+\hat{\alpha}) U^\dagger_\alpha(y+\hat{\beta}) U^\dagger_\beta(y)\right)\right\}\Bigg|_{s=0} \tag{3.125}$$

$$= t^a \frac{\beta}{3} \sum_y \sum_{\alpha<\beta} \mathrm{Re}\,\mathrm{tr}\left(-X \mathrm{e}^{sX} U_\alpha(y) U_\beta(y+\hat{\alpha}) U^\dagger_\alpha(y+\hat{\beta}) U^\dagger_\beta(y)\right)\Bigg|_{s=0} \tag{3.126}$$

$$= t^a \frac{\beta}{3} \sum_y \sum_{\alpha<\beta} \mathrm{Re}\,\mathrm{tr}\left(-X U_\alpha(y) U_\beta(y+\hat{\alpha}) U^\dagger_\alpha(y+\hat{\beta}) U^\dagger_\beta(y)\right). \tag{3.127}$$

Since X leaves all terms different than (x,μ) as zero, we are only left with the two links at y.

$$\begin{aligned}(3.127) = -t^a\frac{\beta}{3}\sum_{\nu\neq\mu}\mathrm{Re\,tr}\Big(&t^aU_\mu(x)U_\nu(x+\hat{\mu})U_\mu^\dagger(x+\hat{\nu})U_\nu^\dagger(x)\\ &+U_\mu(x-\hat{\nu})U_\nu(x-\hat{\nu}+\hat{\mu})\left(t^aU_\mu(x)\right)^\dagger U_\nu^\dagger(x-\hat{\nu})\Big).\end{aligned}\tag{3.128}$$

We can now use that $\mathrm{Re\,tr}(P_{\mu\nu}) = \mathrm{Re\,tr}(P_{\mu\nu}^\dagger)$, $\mathrm{tr}(A+B) = \mathrm{tr}A + \mathrm{tr}B$ and the cyclic property of the trace $\mathrm{tr}(ABCD) = \mathrm{tr}(DABC) = \cdots = \mathrm{tr}(BCDA)$ on the last part of the sum in eq. (3.128),

$$\begin{aligned}&\mathrm{Re\,tr}\big(U_\mu(x-\hat{\nu})U_\nu(x-\hat{\nu}+\hat{\mu})\left(t^aU_\mu(x)\right)^\dagger U_\nu^\dagger(x-\hat{\nu})\big)\\ &\quad=\mathrm{Re\,tr}\left(\Big[U_\mu(x-\hat{\nu})U_\nu(x-\hat{\nu}+\hat{\mu})\left(t^aU_\mu(x)\right)^\dagger U_\nu^\dagger(x-\hat{\nu})\Big]^\dagger\right)\\ &\quad=\mathrm{Re\,tr}\left(U_\nu^\dagger(x-\hat{\nu})t^aU_\mu(x)U_\nu(x-\hat{\nu}+\hat{\mu})^\dagger U_\mu(x-\hat{\nu})^\dagger\right)\\ &\quad=\mathrm{Re\,tr}\left(t^aU_\mu(x)U_\nu(x-\hat{\nu}+\hat{\mu})^\dagger U_\mu(x-\hat{\nu})^\dagger U_\nu^\dagger(x-\hat{\nu})\right).\end{aligned}\tag{3.129}$$

By inserting eq. (3.129) into eq. (3.128), we can pull out $t^aU_\mu(x)$,

$$\begin{aligned}\partial_{x,\mu}S_G[U] = -t^a\frac{\beta}{3}\sum_{\nu\neq\mu}\mathrm{Re\,tr}\big[&t^aU_\mu(x)\big(U_\nu(x+\hat{\mu})U_\mu^\dagger(x+\hat{\nu})U_\nu^\dagger(x)\\ &+U_\nu(x-\hat{\nu}+\hat{\mu})^\dagger U_\mu(x-\hat{\nu})^\dagger U_\nu^\dagger(x-\hat{\nu})\big)\big].\end{aligned}\tag{3.130}$$

To further clean this up, we can define the two plaquette-segments as a *staple*, such that the combined parts become a rectangle which we label as $X_{\mu\nu}(x)$,

$$\partial_{x,\mu}S_G[U] = -t^a\frac{\beta}{3}\sum_{\nu\neq\mu}\mathrm{Re\,tr}\left[t^aU_\mu(x)X_{\mu\nu}(x)\right],\tag{3.131}$$

$$X_{\mu\nu}(x) = U_\nu(x+\hat{\mu})U_\mu^\dagger(x+\hat{\nu})U_\nu^\dagger(x) + U_\nu(x-\hat{\nu}+\hat{\mu})^\dagger U_\mu(x-\hat{\nu})^\dagger U_\nu^\dagger(x-\hat{\nu}).\tag{3.132}$$

An schematic view of the staple can be seen in fig. 3.6 on the next page.

The next step is to get rid of the generators t^a in eq. (3.131). We can now either make an explicit calculation using the generators t^a with symbolic manipulation software[7], or we can try to rewrite the expression even further.

We can start by making the matrix $Y_\mu(x,t_f) = U_\mu(x)X_\mu(x)$ real,

$$-g_S^2\left\{\partial_{x,\mu}S_G[V_{t_f}]\right\} = 2\sum_a t^a\,\mathrm{Re\,tr}\left\{t^aY_\mu(x,t_f)\right\}\tag{3.133}$$

$$= 2\sum_a t^a\mathrm{tr}\left\{t^a\frac{-i}{2}\left(Y_\mu(x,t_f) - Y_\mu^\dagger(x,t_f)\right)\right\},\tag{3.134}$$

[7]The class `WilsonExplicitDer` does this explicitly, and has shown to be about 5% faster than the ordinary method in eq. (3.138) on the next page.

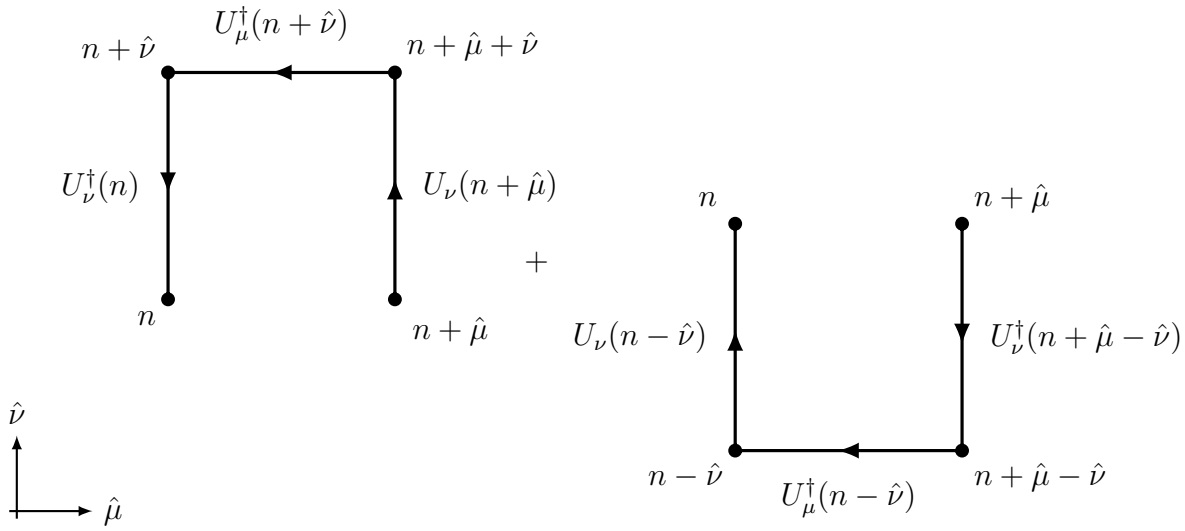


Figure 3.6: The two staples in $X_{\mu\nu}(n)$ from eq. (3.132) on the facing page used in the action derivative in eq. (3.131) on the preceding page.

where we used that writing a complex matrix A as $A = -\frac{i}{2}(B - B^\dagger)$ ensures that we are only selecting the real components. We have that a 3×3 matrix A can be written as [4]

$$A = a_0 I + a^i t^i, \quad \text{with} \quad a_0 = \frac{1}{3}\text{tr}A, \quad a^i = 2\text{tr}(t^i A), \tag{3.135}$$

which can be rearranged to

$$\text{tr}(t^i A)t^i = \frac{1}{2}A - \frac{1}{6}\text{tr}A. \tag{3.136}$$

Using eq. (3.136) for the trace in eq. (3.134) on the preceding page gives us

$$-g_S^2 \left\{\partial_{x,\mu} S_G[V_{t_f}]\right\} = -i \sum_a t^a \text{tr}\left[t^a \left(Y_\mu(x,t_f) - Y_\mu^\dagger(x,t_f)\right)\right] \tag{3.137}$$

$$= -i\left(\frac{1}{2}\left(Y_\mu(x) - Y_\mu^\dagger(x)\right) - \frac{1}{6}\text{tr}\left(Y_\mu(x) - Y_\mu^\dagger(x)\right)\right), \tag{3.138}$$

which is our final expression for the pure gauge action derivative.

3.5.2 Gradient flow and topological charge

A useful aspect of the topological charge is that it is independent of flow time [13]. To see this, we begin by restating the topological charge density in terms of $G_{\mu\nu}$,

$$q = \frac{1}{32\pi^2}\epsilon_{\mu\nu\rho\sigma}\text{tr}\left[G_{\mu\nu}G_{\rho\sigma}\right], \tag{3.139}$$

for which we perform a small variation on B_μ. We get

$$\delta q = \frac{1}{32\pi^2}\epsilon_{\mu\nu\rho\sigma}\left\{\mathrm{tr}\left[\delta G_{\mu\nu}G_{\rho\sigma}\right] + \mathrm{tr}\left[G_{\mu\nu}\delta G_{\rho\sigma}\right]\right\}. \tag{3.140}$$

Moving the indices of the Levi-Civita, we have $\epsilon_{\mu\nu\rho\sigma} = \epsilon_{\rho\sigma\mu\nu}$ which remains positive due to an even number of permutations. Also interchanging $\mu \leftrightarrow \rho$ and $\nu \leftrightarrow \sigma$ and using the cyclicity of the trace,

$$\mathrm{tr}[ABC] = \mathrm{tr}[CAB] = \mathrm{tr}[BCA], \tag{3.141}$$

we get

$$\delta q = \frac{1}{16\pi^2}\epsilon_{\mu\nu\rho\sigma}\mathrm{tr}[G_{\mu\nu}\delta G_{\rho\sigma}]. \tag{3.142}$$

If we now look at the variation of the field strength tensor,

$$\begin{aligned}\delta G_{\mu\nu} &= \delta\left(\partial_\mu B_\nu - \partial_\nu B_\mu - i[B_\mu, B_\nu]\right)\\ &= \partial_\mu \delta B_\nu - \partial_\nu \delta B_\mu - i\left([\delta B_\mu, B_\nu] + [B_\mu, \delta B_\nu]\right)\\ &= \partial_\mu \delta B_\nu - \partial_\nu \delta B_\mu - i\left([B_\mu, \delta B_\nu] - [B_\nu, \delta B_\mu]\right),\end{aligned}$$

We can insert this into eq. (3.142),

$$\delta q = \frac{1}{16\pi^2}\big\{\epsilon_{\mu\nu\rho\sigma}\mathrm{tr}(G_{\mu\nu}\partial_\rho\delta B_\sigma - i[B_\rho, \delta B_\sigma]) \tag{3.143}$$

$$+ \epsilon_{\mu\nu\rho\sigma}\mathrm{tr}(G_{\mu\nu}\partial_\sigma\delta B_\rho - i[B_\sigma, \delta B_\rho])\big\}, \tag{3.144}$$

and then use $\epsilon_{\mu\nu\rho\sigma} = -\epsilon_{\mu\nu\sigma\rho}$ and $\rho \leftrightarrow \sigma$,

$$\delta q = \frac{1}{8\pi^2}\epsilon_{\mu\nu\rho\sigma}\big\{\mathrm{tr}\left(G_{\mu\nu}\partial_\rho\delta B_\sigma\right) - i\mathrm{tr}\left(G_{\mu\nu}[B_\rho, \delta B_\sigma]\right)\big\}. \tag{3.145}$$

We can now use the Jacobi identity(or Bianchi identity),

$$\epsilon_{\mu\nu\rho\sigma}D_\rho G_{\mu\nu} = 0, \tag{3.146}$$

which is valid for all ρ. Inserting the definition of the covariant derivative from eq. (3.103) on page 44,

$$\epsilon_{\mu\nu\rho\sigma}\left(\partial G_{\mu\nu} - i[B_\rho, G_{\mu\nu}]\right) = 0, \tag{3.147}$$

we get

$$\epsilon_{\mu\nu\rho\sigma}\partial G_{\mu\nu} = i\epsilon_{\mu\nu\rho\sigma}[B_\rho, G_{\mu\nu}]. \tag{3.148}$$

Before using eq. (3.148) on the facing page, we must reorganize the last term in eq. (3.145) on the preceding page. We use that a matrix commutator can be written as

$$[A, BC] = [A, B]C + B[A, C]. \tag{3.149}$$

Tracing both sides and using that the trace of a commutator is zero, we are left with

$$[A, B]C = -B[A, C], \tag{3.150}$$

which we can insert into the last term of eq. (3.145) on the facing page,

$$\delta q = \frac{1}{8\pi^2}\epsilon_{\mu\nu\rho\sigma}\left\{\mathrm{tr}\left(G_{\mu\nu}\partial_\rho\delta B_\sigma\right) + i\mathrm{tr}\left([B_\rho, G_{\mu\nu}]\delta B_\sigma\right)\right\}. \tag{3.151}$$

Applying eq. (3.148) on the preceding page on eq. (3.151) yields

$$\begin{aligned}\delta q &= \frac{1}{8\pi^2}\epsilon_{\mu\nu\rho\sigma}\left\{\mathrm{tr}\left[G_{\mu\nu}\partial_\rho\delta B_\sigma\right] + \epsilon_{\mu\nu\rho\sigma}\mathrm{tr}\left[\partial_\rho G_{\mu\nu}\delta B_\sigma\right]\right\} \\ &= \partial_\rho\left(\frac{1}{8\pi^2}\epsilon_{\mu\nu\rho\sigma}\mathrm{tr}\left[G_{\mu\nu}\delta B_\sigma\right]\right) \\ &= \partial_\rho\tilde{w}_\rho.\end{aligned} \tag{3.152}$$

If we now consider the variation from the gradient flow on the gauge field and on the charge density,

$$\partial B_\sigma = \partial_t B_\sigma\delta t, \quad \delta q = \partial_t q\delta t, \tag{3.153}$$

the change in $\partial_t q$ is equivalent with what we have found,

$$\begin{aligned}\partial_t q\delta t &= \frac{1}{8\pi^2}\epsilon_{\mu\nu\rho\sigma}\mathrm{tr}\left[G_{\mu\nu}\delta B_\sigma\delta t\right] \\ &= \frac{1}{8\pi^2}\epsilon_{\mu\nu\rho\sigma}\mathrm{tr}\left[G_{\mu\nu}D_\lambda G_{\lambda\sigma}\right]\delta t \\ &= w_\rho\partial t,\end{aligned} \tag{3.154}$$

where we in the second to the last term used the flow equation from eq. (4.27) on page 65. This gives us $\partial_t q = \partial_\rho w_\rho$. From this follows that the topological charge

$$Q = \int \mathrm{d}^4x q, \tag{3.155}$$

can now be written as a gradient of the fields,

$$\partial_t Q = \int \mathrm{d}^4\partial_t q = \int \mathrm{d}^4x\partial_\rho w_\rho. \tag{3.156}$$

Given that the Yang-Mills action has to vanish at the border when tending to infinity, the change in topological charge in flow time must be zero,

$$\partial_t Q = 0. \tag{3.157}$$

This fact allows us to explore the gauge fields certain that they will retain their topological properties, such as their winding numbers.

3.6 Scale setting: connecting with the real world

Since all observables are given in units of $\hbar c$ and the action is dimensionless, we need to relate them to physical quantities in order for them to make any sense. Take for instance the dimensionless quantity am_{eff}, where a is the lattice spacing and m_{eff} is some effective mass. In order to connect the mass to actual physics, we must first connect the lattice spacing a to a physical unit. In non-quenched calculations this is usually done through physical pion masses. Since we are working in pure gauge, we will instead look at some quenched approaches and how one can relate a to the gauge coupling β.

3.6.1 The Sommer parameter

One way of connecting the lattice spacing a to the gauge coupling is through a physical distance r_0, the *Sommer parameter* [28, 62]. The Sommer parameter is taken to be approximately $r_0 \simeq 0.5$ fm, which is what we will take it to be for this thesis. In order to find the Sommer parameter exactly, one needs to calculate the *static quark potential* through the briefly mentioned Wilson loops. Given the vacuum the expectation value for the Wilson loop

$$\langle W_{\mathcal{C}} \rangle = \frac{1}{Z} \int \mathcal{D}U \mathrm{e}^{-S_G} \mathrm{tr} \left[\prod_{l \in \mathcal{C}} U_l \right], \tag{3.158}$$

we take a contour $\mathcal{C}$ of size $n_r \times n_t$, forming a rectangle. This potential corresponds to the force between two static quarks [24, ch. 3.4] and can be expanded and rewritten as

$$\langle W_{\mathcal{C}} \rangle = C \exp(-tV(r)), \tag{3.159}$$

with $r = na$ and $t = n_t a$. $V(r)$ is what is called the static quark potential, and C is some proportionality constant. Extracting $V(r)$, we will parametrize the potential as

$$V(r) = A + \frac{B}{r} + \sigma r. \tag{3.160}$$

We can connect $V(r)$ with experimental data of the static force at $F(r = r_0)$,

$$F(r_0) r_0^2 = 1.65, \tag{3.161}$$

which corresponds to $r_0 \simeq 0.5$ fm. Connecting the static force with the static potential is done by simply taking the derivative of $V(r)$,

$$F(r) = \frac{\mathrm{d}}{\mathrm{d}r} V(r) = -\frac{B}{r^2} + \sigma. \tag{3.162}$$

Since the lattice spacing a is unknown, we extract $aV(an)$ and perform a fit on

$$aV(an) = Aa + \frac{B}{n} + \sigma a^2 n. \tag{3.163}$$

When combined with experimental results for the static force in eq. (3.161) on the facing page we have that

$$F(r_0)r_0^2 = -B + \sigma r_0^2, \tag{3.164}$$

which suggests that

$$r_0 = \sqrt{(1.65 + B)\sigma}. \tag{3.165}$$

In lattice units this is

$$\frac{r_0}{a} = \sqrt{\frac{1.65 + B}{\sigma a^2}}. \tag{3.166}$$

In order to determine r_0, we first have to perform a fit for the static quark potential. We perform several Wilson loops at varying length $r = an$, and then extract $aV(an)$ based on eq. (3.159) on the preceding page.

If we wish to determine a from the static quark potential, we first determine B and σa^2. Then we calculate $X = r_0/a$ from eq. (3.166). The lattice spacing is then simply $a = 0.5/X$ fm.

3.6.2 Calculating the lattice spacing for pure gauge theories

When the lattice spacing a becomes small, the coupling β increases. For the gauge action in eq. (3.73) on page 38, a scale has been determined [28]. Given that β is between $5.7 \leq \beta \leq 6.92$, the β-dependency on a was found to be

$$a = r_0 \exp\left(-1.6804 - 1.7331(\beta - 6) + 0.7849(\beta - 6)^2 - 0.4428(\beta - 6)^2\right). \tag{3.167}$$

The paper by Guagnelli et al. [28] they determine the uncertainty of this lattice spacing as 0.3% for $\beta_{\text{low}} = 5.7$, growing linearly to 0.6% for $\beta_{\text{high}} = 6.57$. Thus, an estimation of the error in a can be parametrized as

$$\sigma_a(\beta) = a(\alpha_1\beta + \alpha_0), \quad \alpha_1 = \frac{(0.6 - 0.3)}{100(\beta_{\text{high}} - \beta_{\text{low}})}, \quad \alpha_0 = \frac{0.3}{100} - \alpha_1\beta_{\text{low}}. \tag{3.168}$$

3.6.3 Setting scale with gradient flow

Using gradient flow, new options for setting the scale is possible [40]. Since $t_f^2 \langle E(t) \rangle$ is a dimensionless quantity, proportional to the gauge coupling, we can define a reference scale t_0 as

$$\left\{t_f^2 \langle E(t) \rangle\right\}_{t_f = t_0} = 0.3. \tag{3.169}$$

The energy is given from eq. (3.83) on page 41. After finding t_0 for different lattice spacing a, one can perform a simple extrapolation for $\sqrt{8t_0}/r_0$ versus t_0/a^2 to the continuum, $a = 0$. The benefit of using the t_0 reference scale versus the r_0 reference scale is that we require far fewer configurations for good statistics and it involves no complicated fitting procedure. Further, we get $\langle E(t) \rangle$ more of less for free if we calculate the topological charge since we have already calculated the leafs of the clover.

The w_0 reference scale

There is also the w_0 scale first described in [5]. The w_0 scale is defined through the derivative of $t_f^2 \langle E(t) \rangle$,

$$W(t)|_{t=w_0^2} = 0.3, \tag{3.170}$$

where

$$W(t) \equiv t_f \frac{\mathrm{d}}{\mathrm{d}t_f} \left\{ t_f^2 \langle E \rangle \right\}. \tag{3.171}$$

The benefit of the w_0 scale is that $W(t)$ incorporates information of smaller scales around $\mathcal{O}(1/\sqrt{t_f})$, while $t_f^2 \langle E(t) \rangle$ captures information for scales larger than $\mathcal{O}(1/\sqrt{t_f})$. This can be highly useful since at small flow times $t \sim a^2$ we are subject to discretization effects of the lattice [5].

There is also the advantage for both t_0 and w_0 that they appear to not require any renormalization [40].

3.7 Recovering the continuum limit

The results obtained are non-perturbative and exact for a given lattice spacing a in the limit of $N_{\mathrm{MC}} \to \infty$ down to an error $\mathcal{O}(a^2)$ for the Wilson gauge action and observable. When extrapolating to the continuum, we will have no error in the lattice spacing a, but the errors of each of the values we use to extrapolate will have $\mathcal{O}(a^n)$ errors. These errors depend on the action and the observable and will give contributions to the overall uncertainty in the extrapolation. To determine the leading contributions to the error we need to perform a Symanzik analysis, which allows us to identify different contributions and add counter terms to promptly remove them [64, 65]. Performing a Symanzik analysis will help us determine if our leading order stems from the classical Taylor expansion and discretization or if they are quantum contributions. Typically, we have $\mathcal{O}(a)$ discretization effects for fermions and $\mathcal{O}(a^2)$ for gauge fields [24, ch. 9]. We know that from the Taylor expansion, we have a $\mathcal{O}(a^2)$ error contribution. Removing this error contribution could improve the overall error as the Taylor expansion may contribute more than other sources of error, even if they too are of order $\mathcal{O}(a^2)$. When introducing corrections one will need to take into account both quantum corrections as well as classical corrections. An example of an improved gauge action is the Lüscher-Weisz action [38] that removes an $\mathcal{O}(a^2)$ effect and adds a first-order loop correction.

As mentioned previously, when we are calculating expectation values on the lattice one is only calculating that value as it occurs for a given lattice spacing a. In order to properly study the real world value, we need to extrapolate to the continuum $a \to 0$. This process usually involves making a linear fit in a^2 and setting $a = 0$. The details surrounding a linear fit is covered in appendix D.5 on page 156. In order to perform the continuum extrapolation, we will need results for several different lattice spacings a. However, the

finer spacing a means a smaller overall volume. To remedy this we increase the number of lattice points, which drives the computational cost up.

A final consideration to make when extrapolation to the continuum, is at which flow time we chose to extrapolate from. The rule of thumb is that we wish to extrapolate at a point where we have no discretization effects, i.e. the values have plateaued. We also seek to extrapolate from a point where we have a smearing radius $\sqrt{8t_f}$ of hadronic scales. This is typically around $0.4 - 0.5$ and onwards. One can also extrapolate from the reference scale t_0 or w_0.

Chapter 4

Numerical and algorithmical details

So far we have covered QCD and how it translates to the lattice, without specifying how one actually generates gauge configurations or flow the configurations. This is a strenuous task with several numerical and programming challenges. The main challenge is generating independent gauge configurations utilizing a Markov Chain Monte Carlo method(MCMC), such that the configurations remain uncorrelated with each other throughout the Monte Carlo history. For this we can choose between several algorithms, the most famous(and easiest to implement) is perhaps the Metropolis-Hastings algorithm. Other noteworthy choices are the much used hybrid Monte Carlo(HMC) method, heat bath, overrelaxation and microcanonical methods [24].

Statistical details on data analysis can be seen in appendix D on page 149.

4.1 The Metropolis-Hastings algorithm

Our MCMC method of choice for creating configurations of the gauge field is to use the Metropolis-Hastings algorithm [29, 45]. Given a probability distribution function(PDF) $P_i^{(n)}$ for state i at time step n, we seek to end up at a final distribution p_i as $n \to \infty$ from which we can sample gauge configurations. To achieve this goal we wish to apply the Metropolis-Hastings algorithm which relies on two fundamental premises,

1. **Existence of a stationary distribution.** There must exist a stationary distribution p_i. This is often summarized by the *detailed balance condition* which requires a process to be reversible.

2. **Uniqueness of stationary position.** The stationary distribution p_i must be unique. This is true as long as the Markov chain is ergodic, that is that in the space of every possible state, we are always able to visit or revisit a state in a finite number of steps.

In order to define the probability of going from a time step $n-1$ to n, we have to introduce the transitioning probability $T_{i\to j}$ and the probability of accepting a proposed move $A_{i\to j}$.

Our goal is to find p_i as $P_i^{n\to\infty} \to p_i$. The probability $P_i^{(n)}$ for a given state i at time n is

$$P_i^{(n)} = \sum_j \left[P_j^{(n-1)} T_{j\to i} A_{j\to i} + P_i^{(n-1)} T_{i\to j}(1 - A_{i\to j}) \right], \tag{4.1}$$

where the first part is the probability of going from a state j to i, and the latter part is from starting at state i and not transitioning to a state j. We can rewrite this as

$$\begin{aligned} P_i^{(n)} &= \sum_j \left[P_j^{(n-1)} T_{j\to i} A_{j\to i} - P_i^{(n-1)} T_{i\to j} A_{i\to j} + P_i^{(n-1)} T_{i\to j} \right] \\ &= \sum_j \left[P_j^{(n-1)} T_{j\to i} A_{j\to i} - P_i^{(n-1)} T_{i\to j} A_{i\to j} \right] + \sum_j P_i^{(n-1)} T_{i\to j} \\ &= \sum_j \left[P_j^{(n-1)} T_{j\to i} A_{j\to i} - P_i^{(n-1)} T_{i\to j} A_{i\to j} \right] + P_i^{(n-1)}, \end{aligned} \tag{4.2}$$

where we in the last line used that the total transitioning probability from $i \to j$ is $\sum_j T_{i\to j} = 1$. If we now let $P_i^{n\to\infty} \to p_i$, we get

$$\begin{aligned} p_i &= \sum_j \left[p_j T_{j\to i} A_{j\to i} - p_i T_{i\to j} A_{i\to j} \right] + p_i \\ 0 &= \sum_j \left[p_j T_{j\to i} A_{j\to i} - p_i T_{i\to j} A_{i\to j} \right], \end{aligned}$$

which applies for the system as a whole when $n \to \infty$. Since we want to find the probability for a single state we look at

$$\begin{aligned} & p_j T_{j\to i} A_{j\to i} - p_i T_{i\to j} A_{i\to j} = 0 \\ & \to p_j T_{j\to i} A_{j\to i} = p_i T_{i\to j} A_{i\to j}. \end{aligned} \tag{4.3}$$

Rearranging the terms in eq. (4.3) and we get

$$\frac{A_{i\to j}}{A_{j\to i}} = \frac{p_j T_{j\to i}}{p_i T_{i\to j}}, \tag{4.4}$$

which is called *detailed balance*. Detailed balance tells us that the probability of accepting any new move is proportional to the probability of generating a state i or j and transitioning from a state j or i.

Since we now want to maximize the chance of accepting a new configuration, we write the probability $A_{i\to j}$ as a choice between

$$A_{i\to j} = \min\left(1, \frac{p_j T_{j\to i}}{p_i T_{i\to j}}\right). \tag{4.5}$$

Since a probability can never be larger than 1, we choose 1 if $p_i T_{i\to j} < p_j T_{j\to i}$ else we choose $\frac{p_j T_{j\to i}}{p_i T_{i\to j}}$. Equation (4.5) is called the *Metropolis choice*. The advantage now is that we only need to find an expression for $\frac{p_j T_{j\to i}}{p_i T_{i\to j}}$.

The Metropolis-Hastings algorithm so far is summed up in algorithm 1 on the next page.

Algorithm 1 The basic Metropolis algorithm, as applied to a general system generating N_{MC} Monte Carlo samples.

repeat
  Randomly generate a candidate state j with probability $T_{i\to i}$.
  Calculate $A_{i\to j}$ given by eq. (4.5) on the facing page.
  Generate random number r from a uniform distribution, $u \in [0,1]$.
  if $u \leq A_{i\to j}$ **then**
    Accept new state j.
  else if $u > A_{i\to j}$ **then**
    Reject new state j and retain the old state i.
  end if
until N_{MC} samples are generated.

4.1.1 Applying the Metropolis-Hastings algorithm to LQCD

The next step now is to figure out p_i and $T_{j\to i}$ as applied to lattice QCD. Recalling the QCD expectation value from eq. (3.78) on page 39 we can define a PDF as

$$\mathrm{d}P(U) = \frac{\mathrm{e}^{-S[U]}\mathcal{D}U}{\int \mathcal{D}U \mathrm{e}^{-S[U]}}. \tag{4.6}$$

Discretized, we have that the partition function becomes an average over N_{MC} Monte Carlo samples,

$$\langle O \rangle = \lim_{N_{\mathrm{MC}}\to\infty} \frac{1}{N_{\mathrm{MC}}} \sum_{i}^{N_{\mathrm{MC}}} O[U_i], \tag{4.7}$$

where U_i are random configurations of the lattice. Discretizing the PDF to a single sample,

$$P[U_i] = \frac{\mathrm{e}^{-S[U_i]}}{\int \mathcal{D}U \mathrm{e}^{-S[U]}} = \frac{1}{Z}\mathrm{e}^{-S[U_i]} \equiv p_i, \tag{4.8}$$

allows us to insert p_i into the Metropolis choice in eq. (4.5) on the facing page,

$$A_{i\to j} = \min\left(1, \frac{\mathrm{e}^{-S[U_j]}T_{j\to i}}{\mathrm{e}^{-S[U_i]}T_{i\to j}}\right). \tag{4.9}$$

Assuming that the transitioning probability is fully symmetric, so that our algorithm becomes the standard Metropolis algorithm, $T_{j\to i} = T_{i\to j}$, the acceptance probability becomes

$$A_{i\to j} = \min\left(1, \mathrm{e}^{-\Delta S}\right), \tag{4.10}$$

where $\Delta S = S[U_j] - S[U_i]$. Since we now have an expression for the acceptance probability $A_{i\to j}$, we only need to find what ΔS is.

From now on, we denote a candidate link as U' instead of U_j.

4.1.2 Finding the change in action ΔS

Recalling the lattice gauge action in eq. (3.73) on page 38, we can write out every plaquette associated with a link $U_\mu(n)$. A plaquette can be oriented in six directions around a link variable $U_\mu(n)$. Without summing over μ and ν, the two plaquettes in the plane becomes

$$\begin{aligned} P_{\mu\nu}(n) + P_{\mu,-\nu}(n) &= U_\mu(n)U_\nu(n+\hat{\mu})U_\mu^\dagger(n+\hat{\nu})U_\nu^\dagger(n) \\ &\quad + U_\mu(n)U_\nu^\dagger(n+\hat{\mu}-\hat{\nu})U_\mu^\dagger(n-\hat{\nu})U_\nu(n-\hat{\nu}) \end{aligned} \tag{4.11}$$

$$\begin{aligned} &= U_\mu(n)\big[U_\nu(n+\hat{\mu})U_\mu^\dagger(n+\hat{\nu})U_\nu^\dagger(n) \\ &\qquad + U_\nu^\dagger(n+\hat{\mu}-\hat{\nu})U_\mu^\dagger(n-\hat{\nu})U_\nu(n-\hat{\nu})\big] \end{aligned} \tag{4.12}$$

$$= U_\mu(n)X_{\mu\nu}(n), \tag{4.13}$$

which allows us to rewrite the action to

$$S_G[U] = \frac{\beta}{3}\sum_{n\in\Lambda}\sum_{\mu\neq\nu}\mathrm{Re}\,\mathrm{tr}\left[1 - U_\mu(n)X_{\mu\nu}(n)\right]. \tag{4.14}$$

If we use the staple definition(see fig. 3.6 on page 49) of the gauge action when taking the difference in action between $U'_\mu(n)$ and $U_\mu(n)$, all terms except those involved directly in the candidate link dies,

$$\Delta S = -\frac{\beta}{3}\,\mathrm{Re}\,\mathrm{tr}\left[\big(U'_\mu(n) - U_\mu(n)\big)\,X_\mu(n)\right], \tag{4.15}$$

with $X_\mu(n)$ being the sum of the 6 staples oriented around $U_\mu(n)$,

$$X_\mu(n) = \sum_{\mu\neq\nu} X_{\mu\nu}(n). \tag{4.16}$$

4.1.3 Generating a candidate link

In the Metropolis algorithm 1 on the previous page an essential step is generating a candidate configuration. We will generate a candidate link $U'_\mu(n)$ to use when comparing the change in action. The way a candidate is suggested is to suggest one similar to the old one, which is done by rotating the existing link by a random SU(3) matrix X,

$$U'_\mu(n) = XU_\mu(n). \tag{4.17}$$

Since we want to suggest an update $U'_\mu(n)$ that is not too dissimilar to the old one $U_\mu(n)$, we start by presenting a method of generating random SU(3) matrices close to unity. A method that generates general SU(3) matrices is presented in appendix C on page 147.

Generating a random matrix near unity, *RST*

We start by generating three SU(2) matrices populated with four random numbers each, in an interval $r_i \in (-\frac{1}{2}, \frac{1}{2})$, for embedding in three SU(3) matrices [8]. Then, the SU(2) matrix is given by

$$\mathbf{x} = \varepsilon_{\text{rnd}} \frac{\mathbf{r}}{|\mathbf{r}|}, \quad x_0 = \text{sign}(r_0)\sqrt{1 - \varepsilon_{\text{rnd}}}, \tag{4.18}$$

where ε_{rnd} is a parameter that controls the spread of the update matrix X around 1. For populating the SU(2) matrices r, s, t we get

$$r, s, t = \begin{pmatrix} x_0 & x_1 \\ x_2 & x_3 \end{pmatrix}. \tag{4.19}$$

The SU(2) matrices r, s, t are then embedded in the 3×3 matrices,

$$R = \begin{pmatrix} r_{11} & r_{12} & 0 \\ r_{21} & r_{22} & 0 \\ 0 & 0 & 1 \end{pmatrix}, \; S = \begin{pmatrix} s_{11} & 0 & s_{12} \\ 0 & 1 & 0 \\ s_{21} & 0 & s_{22} \end{pmatrix}, \; T = \begin{pmatrix} 1 & 0 & 0 \\ 0 & t_{11} & t_{12} \\ 0 & t_{21} & t_{22} \end{pmatrix}, \tag{4.20}$$

which yields the final update matrix

$$X = RST. \tag{4.21}$$

The matrix X is now close to unity. In order to ensure that we have an equal probability of choosing the inverse, we test with a random number if we have to invert the matrix X.

A summary of the matrix generation algorithm can be seen in algorithm 2 on the next page.

4.1.4 Initial conditions

When applying the Metropolis algorithm on some system, one usually have two options for the initial configuration - *hot* where all the elements are randomly oriented, and *cold* where all elements are set to unity. We will differentiate between to options for hot initial configuration, one where we use the *RST* method for generating matrices close to unity(see section 4.1.3), and one where we use a random initial configuration as described in appendix C on page 147. In summary, we have three possible choices for the initial conditions,

- **Cold start**. Start with all links set to unity.
- **Hot start**. Start with matrices set to fully random SU(3) matrices, as given in appendix C on page 147.
- **Hot *RST* start**. Start with links as random SU(3) matrices close to unity, as given in section 4.1.3.

Algorithm 2 Algorithm for generating a SU(3) matrix close to unity. Takes only ϵ_{rnd} as an input that controls the spread of the matrix.

procedure GENERATERSTMATRIX(ϵ_{rnd})
    for SU(2) matrix r, s, t **do**
        Generate 4 random numbers $\mathbf{r}$, r_0 from a uniform $(-1/2, 1/2)$ distribution.
        Find $\mathbf{x}$, x_0 according to eq. (4.18) on the preceding page.
        Populate the SU(2) matrix according to eq. (4.19) on the previous page.
        Embed the SU(2) matrix in their corresponding SU(3) matrix from eq. (4.20) on the preceding page.
    end for
    Generate a random number u from a uniform distribution $(-1/2, 1/2)$.
    if $u \leq 0.5$ **then**
        return $X = RST$.
    else
        return $X = (RST)^{-1}$.
    end if
end procedure

4.1.5 Algorithm for generating and flowing gauge configurations

The full algorithm for generating gauge configurations is summarized in algorithm 3 on the next page. We start by initializing the lattice at either unity or with random valued SU(3) matrices. Then, we need to *thermalize* the lattice. That is, we skip N_{therm} steps, since at the beginning we have not yet reached a stationary point from which we can start sampling.

In order to minimize autocorrelation[1] between the different configurations, we will let the system evolve a number of steps in between every sampling in order to allow it to sufficiently change a previous configuration. We denote the number of so-called *correlation updates* as N_{corr}. I.e. we perform N_{corr} sweeps on the lattice before we sample it.

As a final trick in order to reduce the autocorrelation, we can utilize the fact we have an expression for the staple in eq. (4.16) on page 60, that only needs to be calculated once at a given link. With the staple, we can run several *local updates* on that exact link. We denote the number of updates we perform on a single link as N_{up}.

The total number of Monte Carlo samples we generate is denoted as N_{MC}. The total number of sweeps on the lattice is $N_{\text{MC}} \times N_{\text{corr}} + N_{\text{therm}}$. We define one *sweep* as having iterated through all the links on the lattice performing N_{up} updates on each link, $U_\mu(n)$.

If one wishes to flow gauge configurations as well, one does this N_{flow} times.

[1] See appendix D.3 on page 150 for the details on autocorrelation.

Algorithm 3 The full algorithm for generating (and if needed, flowing) gauge configurations.

Initialize random number generators with ϵ_{rnd}.
Initialize lattice size, N, N_T.
Initialize other parameters, N_{MC}, N_{therm}, N_{corr}, N_{up}.
Initialize N_{flow} and ϵ_f if desired.
Give the lattice in initial condition of either cold, random hot, or RST-random hot.
for i_{therm} in N_{therm} thermalization steps **do**
    Perform a full sweep of the lattice, and update each link N_{up} times.
end for
for i_{MC} in N_{MC} **do**
    for i_{corr} in N_{corr} **do**
        Perform a full sweep of the lattice, and update each link N_{up} times.
    end for
    Sample observables on the lattice, and store them in memory.
    Write field configuration to file.
    Perform N_{flow} updates on the lattice if that is desired. Observables from the flow will be written to file after the flowing is complete.
end for
Write all observable data to file.

4.1.6 Boundary conditions

An important, but so far overlooked topic is that of boundary conditions. Say we are at an edge of the lattice, and the action $S[U]$ or observable $O[U]$ is requiring the $U_\mu(n+\hat{\mu})$ link. The standard operating procedure is then to implement periodic boundary conditions. If we write the full position of the link variable as $n = (n_1, n_2, n_3, n_4)$ the periodic boundary conditions become

$$\begin{aligned} U_\mu(N, n_2, n_3, n_4) &= U_\mu(0, n_2, n_3, n_4), \\ U_\mu(n_1, N, n_3, n_4) &= U_\mu(n_1, 0, n_3, n_4), \\ U_\mu(n_1, n_2, N, n_4) &= U_\mu(n_1, n_2, 0, n_4), \\ U_\mu(n_1, n_2, n_3, N_T) &= U_\mu(n_1, n_2, n_3, 0). \end{aligned} \tag{4.22}$$

4.1.7 Critical slowdown

Once we start generating configurations, we will discover that as the lattice spacing decreases, the autocorrelation increases. This is a well-known phenomena called called *critical slowdown* [14, 57]. As mentioned in the QCD chapter 2 on page 5, two configurations are equal if they can be continuously transformed into one another, which is possible given they have the same topological charge. If the charges are different this is not possible, as we would require an infinite amount of energy in the continuum to go from one instanton sector to another. Thus, as we shrink the lattice spacing a, the energy required to overcome

the potential barrier increases as well, making it difficult for the algorithm to change from one configuration to a new one that is independent from the previous one.

Once we start presenting our results in chapter 6 on page 89 we will see that the ensembles for $\beta = 6.45$, i.e. the smallest lattice spacing, is far less random and appear to change more slowly in Monte Carlo time than those with larger lattice spacing. Due to critical slowdown, we will in the next chapter 5 on page 67 look into methods for decreasing autocorrelation.

4.2 Integrating gradient flow with Runge-Kutta of third order

Runge-Kutta methods provide us with a framework of introducing high-order methods by allowing us to evaluate several derivatives at a given step. The method is considered first introduced by Runge [56] and later refined by Kutta [32] into its modern shape.

Given a differential equation on the form

$$\dot{y}(t_n) = f(t_n, y_n), \quad \text{with} \quad y(t_0) = y_0, \tag{4.23}$$

where f is some function of y and t, we can write the general numerical solver for Runge-Kutta as

$$y(t_{n+1}) = y(t_n) + \epsilon \sum_{i=1}^{s} b_i k_i, \tag{4.24}$$

with

$$\begin{aligned} k_1 &= f(t_n, y(t_n)), \\ k_2 &= f(t_n + \epsilon c_2, y(t_n) + \epsilon a_{21} k_1), \\ k_3 &= f(t_n + \epsilon c_3, y(t_n) + \epsilon a_{31} k_1 + \epsilon a_{32} k_2), \\ &\vdots \\ k_s &= f\left(t_n + \epsilon c_s, y(t_n) + \epsilon \sum_{j=1}^{s-1} a_{sj} k_j\right). \end{aligned} \tag{4.25}$$

Where s is the number of *stages*, which dictates the number of intermediate steps k_i we include and define the number of coefficients a_{ij}, c_i, and b_j that we will have to choose. These can also be represented in *Butcher tableaux'* [7], whose values can be found by Taylor expanding around t_0 and the k_i values by using $y(t_{n+1}) = y(t_0 + \epsilon)$. If we rewrite

$$\dot{y} = \frac{\mathrm{d}y(t)}{\mathrm{d}t} = f\left(y(t), t\right) \equiv f, \tag{4.26}$$

we can perform second order derivatives $y'' = f_t + f_y f$. This process can be continued for higher order derivatives. All that remains to do is to compare the Taylor terms of the Runge-Kutta solution from eq. (4.24) with the terms from the Taylor expansion of $y(t_0+\epsilon)$, and we have the coefficients for the RK-method we are going to use.

4.2.1 Integrating the flow

We want to apply the Runge-Kutta 3 method in order to solve the flow equation in eq. (3.121) on page 47,

$$\dot{V}_{t_f} = Z(V_{t_f})V_{t_f}, \tag{4.27}$$

where $V_{t_f} \in \mathrm{SU}(3)$ and $Z(V_{t_f}) \in \mathfrak{su}(3)$.

Crouch and Grossman [12] and Munthe-Kaas [48] have looked into structure preserving Runge-Kutta methods for evolving differential equations on Lie groups. Based on their work, Celledoni et al. [9] found a method that has the advantage that requiring the computation of far fewer exponentials and commutators. Their general solution can be written as shown in algorithm 4. The coefficients associated with the Runge-Kutta method for Lie

Algorithm 4 Runge-Kutta for Lie groups $\mathrm{SU}(N)$ as given by Celledoni et al. [9]. p is here the initial configuration of the lattice, while J counts the number of exponentials of each stage and is typically $r-1$. $E_i = e_i(y)\cdot y$ where e_i is the basis for the Lie algebra $\mathfrak{su}(N)$.

1: **for** r = 1:s **do**
2: $\quad Y_r = \exp\left(\sum_k \alpha^k_{rJ} F_k\right) \cdots \exp\left(\sum_k \alpha^k_{r1} F_k\right) p.$
3: $\quad F_r = \epsilon F_{Y_r} = h \sum_i f_i(Y_r) E_i.$
4: **end for**
5: $y_1 = \exp\left(\sum_k \beta^k_J F_k\right) \cdots \exp\left(\sum_J^k F_k\right) \cdots \left(\sum_k \beta^k_1 F_k\right) p.$

groups in algorithm 4 can be determined in a similar fashion to the regular Runge-Kutta seen at the beginning of section 4.2 on the preceding page, but is slightly trickier.

We will utilize coefficients presented by Lüscher [40] for Runge-Kutta of third order. The integration step size is denoted by ϵ_f, such that the flow time will be given by $t_f = i_f \epsilon_f$ with $i_f = 1, 2, 3, \ldots$, and we use that $Y_r = W_r$ and $F_r = Z_r$. For a single time step ϵ_f, the integration is then given by

$$\begin{aligned}
W_0 &= V_{t_f}, \\
W_1 &= \exp\left[\frac{1}{4} Z_0\right] W_0, \\
W_2 &= \exp\left[\frac{8}{9} Z_1 - \frac{17}{36} Z_0\right] W_1, \\
V_{t_f+\epsilon_f} &= \exp\left[\frac{3}{4} Z_2 - \frac{8}{9} Z_1 + \frac{17}{36} Z_0\right] W_2,
\end{aligned} \tag{4.28}$$

with

$$Z_i = \epsilon_f Z(W_i), \quad i = 0, 1, 2. \tag{4.29}$$

The global error for Runge-Kutta 3 goes as $\mathcal{O}(\epsilon^4)$.

A method for exponentiating Hermitian matrices can be found in appendix B.4 on page 141, together with an analysis of its error and performance.

4.2.2 Dealing with systematic uncertainties

We have a few potential sources of systematic uncertainties. An obvious contribution to a systematic error is from the Runge-Kutta 3 integration scheme, which we will investigate in the next chapter. As we will see, this error is of order $\mathcal{O}(\epsilon_f^4)$ is small enough to be considered negligible.

Another source for systematic uncertainties can come from how we extract masses and perform certain linear fits. If we are fitting in a given region and with a variable fit range, we can estimate the systematic error from taking the width of the distribution. The distribution will be generated from bootstrapped samples.

Chapter 5

A new code for generating $SU(3)$ Yang-Mills gauge fields

One of the challenges with developing a program for generating pure gauge configurations is the sheer size of a configuration. A lattice Λ consists of $N^3 \times N_T$ lattice points, with each point having four SU(3) matrices associated to them. Since a single SU(3) matrix is a 3×3 matrix consisting of nine complex numbers or 18 real numbers. The total number of real numbers needed for a single configuration thus becomes

$$\underbrace{N^3}_{\text{Spatial}} \times \underbrace{N_T}_{\text{Temporal}} \times \underbrace{4}_{\text{Links}} \times \underbrace{9}_{\text{SU(3) matrix}} \times \underbrace{2}_{\mathbb{C}\text{-numbers}} = 72N^3N_T, \tag{5.1}$$

which translates to $8 \times 72N^3N_T$ bytes. This requires us to have an efficient method of writing gauge configurations to file, as well as having enough storage space if we wish to store the configuration more permanently.

The size of the lattice is also informing on the computational complexity. With this many points, efficiency is an absolute requirement and something which must be paid close attention to in the development of any lattice QCD code.

5.1 Developing a program for generating Yang-Mills gauge fields

A major point of this thesis has been to create a code capable of generating and flowing gauge configuration from scratch. Another important requirement for this code was that it had to be able to run on a High-Performance Computing cluster, i.e. be able to run in parallel in an efficient manner. The code we developed ended up being on around 15000 lines, excluding any whitespace or comments. We chose the arbitrary name of *GLAC*, short for Gluon Action[1].

[1]The full program can be seen on GitHub `https://github.com/hmvege/GLAC`, together with guides for setup and running the code.

5.1.1 A guide to GLAC

The program was structured such that it takes in a `.json` parameter file specifying all relevant settings, then it sets up the lattice and the correct geometry, and initiates the gauge configuration generation and/or gradient flow. The program allows one to generate configurations and then flow them, or load a folder of configurations and then flow them, or continue generating configurations (and/or flow) from a specified configuration. The latter to avoid re-thermalizing and breaking the ergodicity of the Markov chain.

Another goal of GLAC was to have a code that was easily extendable in order to add new actions or observables, requiring the user to simply inherit a base action or base observable class, and only needing to define a minimal amount of class methods.

During development, we chose to make it so that the main user interaction went through a Python program[2], which generated a configuration file based on a simpler configuration file. In retrospect, this may seem like a bad idea and a convoluted way of doing things. The reasoning behind this choice was to have a user-friendly interface, so that elements such as the lattice parallelization geometry setup where not required to be specified in detail and default parameters would be automatically passed. While also generating a `.json` configuration to be passed on to GLAC, `createJobs.py` also submits jobs to either *Slurm* or *Torque* in the appropriate format. Another useful aspect of running everything through a Python script is that we could implement a command line argument interface, helping the setup process significantly.

In fig. 5.1 on the next page we see an overview of the program structure of GLAC. The top row of red boxes consists of external libraries, which in this case is a `.json` reader[3] and MPI [22]. The second row consists of two orange boxes each containing classes and methods of similar functionality. The box on the left contains methods that do not depend on external libraries such as a SU(3) matrix class for performing matrix operations, while the box to the right contains methods such as communication between processors, parameters, input, and storage, that do rely on the external libraries. The third row of yellow boxes consists of methods that rely on all of the previous modules and libraries. I.e. observables, gauge actions, flow, input/output and a class for `Lattice` objects. The next row of green boxes contains most notably the `System` class, which controls the program flow, and implements all of the preceding methods. It also contains modules for unit testing and performance testing, as those are standalone methods utilizing previous parts and are independent of the `System` method. Finally, GLAC is the program which combines the previous rows into one package of accessibility.

5.1.2 Parallelization

One major challenge that has already been alluded to, is parallelization. Due to the sheer size of the lattices, parallelization of the lattice calculations was a non-negotiable requirement, as there simply is not enough memory when running on a single processor

[2]See the `createJobs.py` Python 3 program.

[3]See `https://github.com/nlohmann/json`.

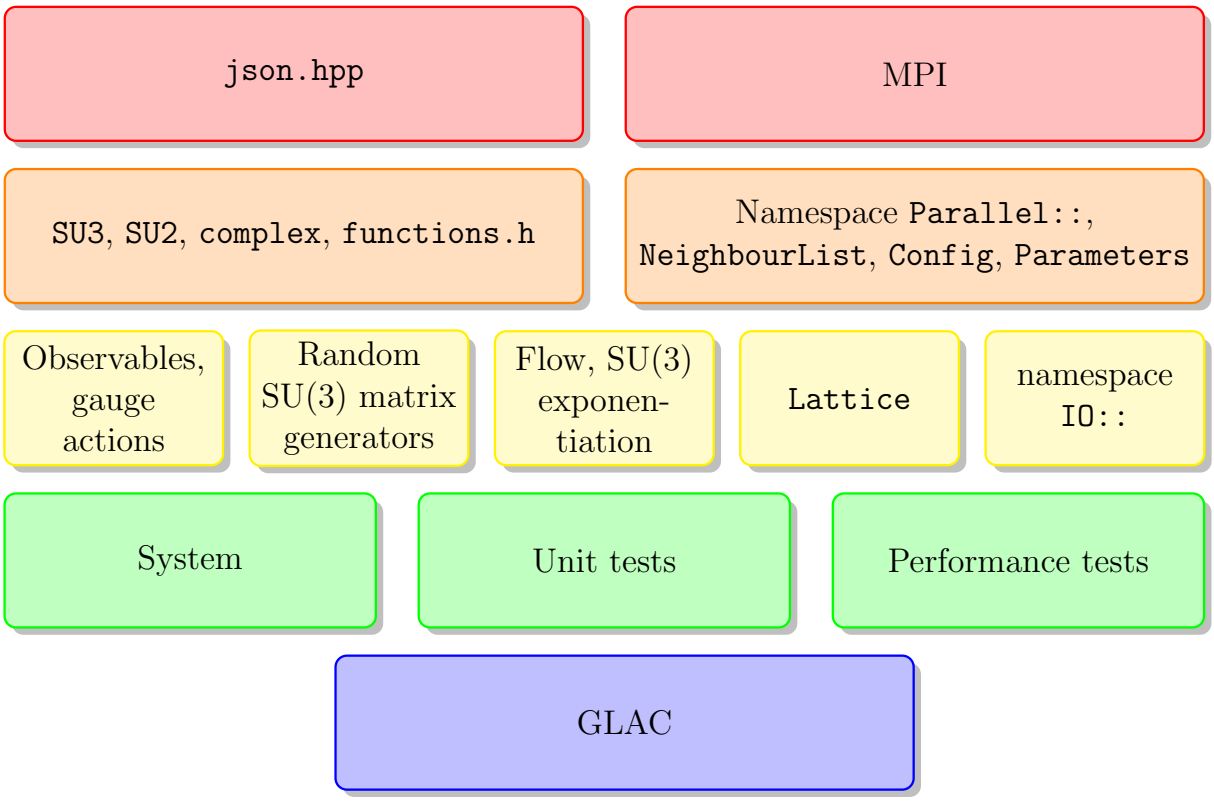


Figure 5.1: The program structure of GLAC. The top row of red boxes consists of external libraries, the second orange row contains methods that are either standalone or depend on external libraries. The next row of yellow boxes contains a method that relies on the previous method, forming more complex methods. The green boxes in the next row contain methods that compile the previous methods into a more overarching framework, while the last one is GLAC itself.

for the largest lattices. The solution is to divide the lattice into *sub-lattices*. This imposes a 4D geometry on the cores which have to be implemented in order for us to share links with the correct processor.

A goal of *any* attempt in parallelization is to have as minimal communication as possible. Further, when forced to communicate we wish to share as much as possible. Since we were working in C++ and we are required to communicate between nodes(one node consists of many threads with shared memory), we landed on using OpenMPI, short for Message Passing Interface [22, 23] for parallelization.

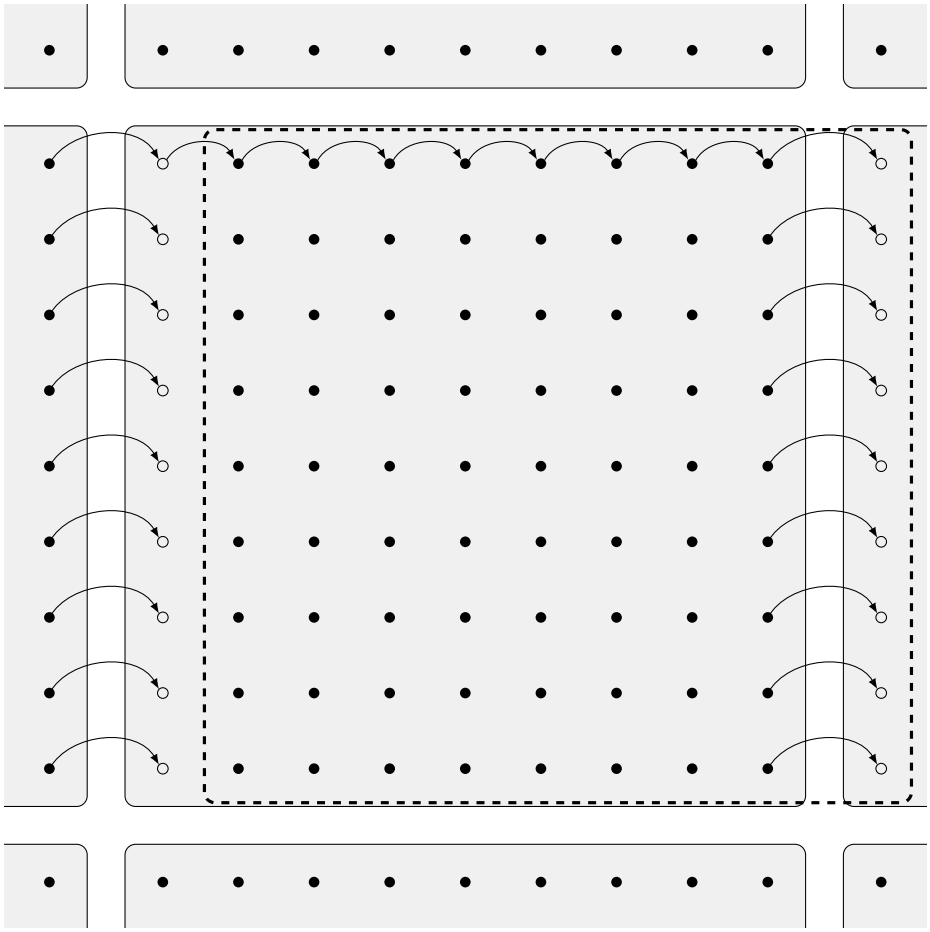


Figure 5.2: An illustration of the lattice shift. The links U_ν of the lattice are copied over to a temporary lattice shifted in direction $\hat{\mu}$. The face that is shifted over to an adjacent sub-lattice is shared through a non-blocking MPI call, while we copy the links to the temporary lattice.

The guiding principles of parallelization led us down several avenues, with the first and simplest approach being the *single link sharing*. That is, we performed a MPI call sharing the single SU(3) matrix between the threads. This approach was used when generating gauge configurations since we were updating every single link many times over and only

required to communicate every N_{up} updates.

As the name implies, single link sharing is rather inefficient since we only share a single link and not for instance an entire face of the hypercube. One attempt to remedy some of this inefficiency was to include buffer zones - or *halos*. That is, the lattice was padded with an extra layer of points at each face of the hypercube. The hope of this approach was that we could perform one large and efficient MPI call before each lattice sweep, but we quickly realized that this was not feasible as when calculating a staple $X_\mu(n)$ from eq. (4.16) on page 60 we would update the corners and edges of the lattice. Since the corners and edges change during a lattice sweep, the halo would have to be shared after each update, making it too often to offer any improvement in parallelization.

With the halos approach laid to waste, we decided on implementing *lattice shifts*. A lattice shift works by copying the entire lattice but shifted in one direction $\hat{\mu}$. Since we are shifting the lattice in one direction $\hat{\mu}$, we have that a face of the lattice will move into the domain of another processor. We then exchange the face of the lattice that has been shifted. A pictorial view of the shift can be seen in fig. 5.2 on the preceding page.

For example, take the plaquette observable in eq. (3.59) on page 36 where we can have two links which are shifted, $U_\nu(n+\hat{\mu})$ and $U_\mu^\dagger(n+\hat{\nu})$. For the first link, $U_\nu(n+\hat{\mu})$ we will share the face of $n = N + \hat{\mu}$ with adjacent processor in direction $\hat{\mu}$, and retrieve corresponding $n = N + \hat{\mu}$ for $n = 0$. The advantage of this method is two-fold: we share a lot of links at once, and it allows for simplified syntax when sharing. As an example, the plaquette code is written as,

```
void Plaquette::calculate(Lattice<SU3> *lattice,
   unsigned int iObs) {
     Lattice<SU3> m_temp;
     m_temp.allocate(lattice.m_dim);
     double tempObservable = 0;
     for (int mu = 0; mu < 4; mu++) {
          for (int nu = mu+1; nu < 4; nu++) {
               m_temp = lattice[mu];
               m_temp *= shift(lattice[nu],FORWARDS,mu);
               m_temp *=
                  inv(shift(lattice[mu],FORWARDS,nu));
               m_temp *= inv(lattice[nu]);
               tempObservable += sumRealTrace(m_temp);
          }
     }
     tempObservable *= 1/(16*m_latticeSize);
     (*m_observable)[iObs] = tempObservable;
}
```

This code snippet is quite close to the actual equation for the plaquette in eq. (3.79) on page 40. We pass in a pointer, `lattice` consisting of 4 `Lattice` objects, one for each

dimension, to the plaquette function. This expression is almost written exactly as it is stated in eq. (3.59) on page 36, and we see that the `shift`-operator takes a lattice, shift direction(`FORWARDS` or `BACKWARDS`), and the direction index `nu`. The function `inv` simply refers to us taking the inverse of the lattice. Having `Lattice` objects allows us to overload operators and abstract the code. Since loops over the dimensions can be hidden by overloading, we are interfacing much closer with the actual equations we wish to work with.

Lattice shifts do not make sense when generating gauge configurations, as we are updating every single link and we share a lot of unused links with lattice shifts. However, when calculating observables or flowing, this is the method of choice.

5.1.3 Scaling

An import issue when discussing any parallel program is scaling. The primary concern for a parallel program is the communication overhead between nodes. If there is negligible communication between nodes, we will have perfect scaling. In the case of there being communication between nodes and the communication increases for the system size, we will at some point see a plateau for the optimal number of cores we can use. There exist several different approaches to measure the parallel scaling of a program[4], the most obvious one is the speedup, efficiency, strong scaling, and weak scaling. Since we are dealing with high-performance computing, we will look specifically at strong scaling, weak scaling, and speedup.

We expect the speedup and scaling of a program to plateau given an increase in the number cores, as the increased communication will outweigh any performance gains. The scaling tests performed in this section where all performed on the Abel cluster.

Strong scaling

Strong scaling is defined as having a *fixed problem* and a *variable N_p cores*. The strong scaling of GLAC can be seen in fig. 5.3 on the next page. We performed the strong scaling with a fixed total lattice size of $N = 16$ and $N_T = 32$, since these is pure multiples of two which allows for ease in parallelization.

Weak scaling

Weak scaling is defined as having a *fixed problem per processor* and a *variable N_p cores*. The weak scaling of GLAC can be seen in fig. 5.4 on page 74. We performed the weak scaling test with a fixed sub-lattice size of 4^4.

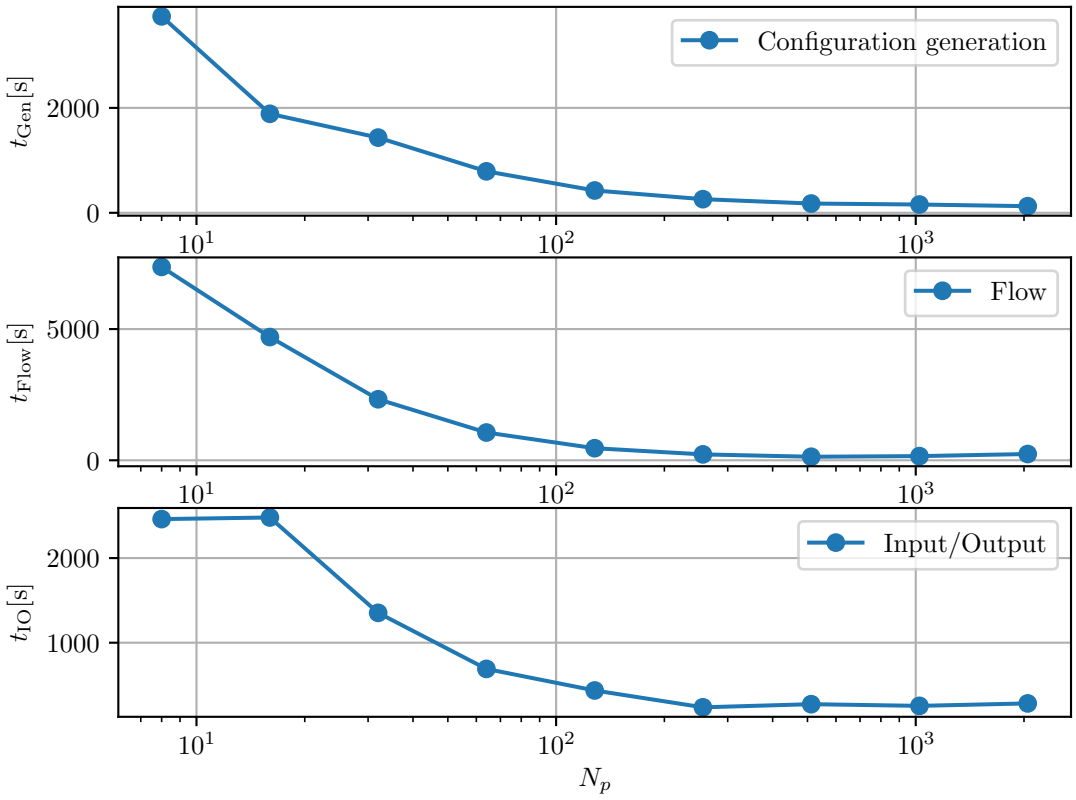


Figure 5.3: Strong scaling of the configuration generation, flowing procedure, and input/output operations. Initialization time is included. The number of cores goes from $N_p = 8$ to $N_p = 2048$. The lattice size was $N^3 \times N_T = 16^3 \times 32$. The flow scaling loaded and flowed a single configuration with $N_{\mathrm{flow}} = 1000$ flow steps, the IO scaling wrote ten configurations to file and the configuration generation scaling generated a single configuration with $N_{\mathrm{corr}} = 600$ and $N_{\mathrm{up}} = 30$.

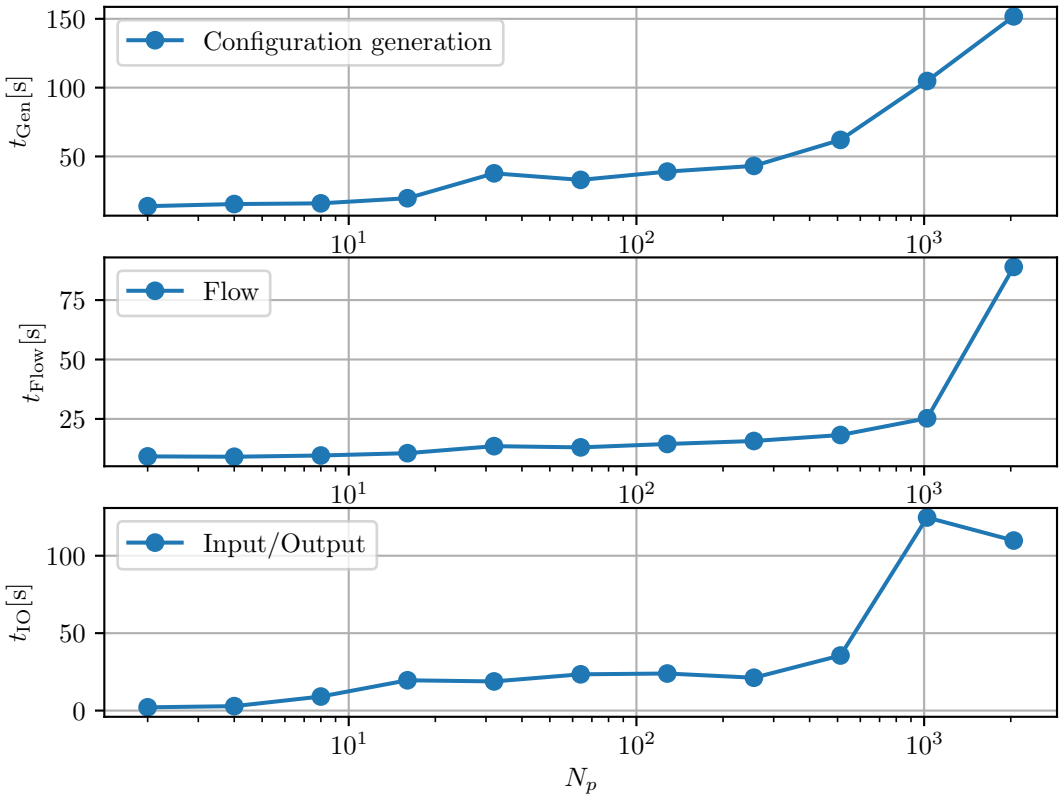


Figure 5.4: Weak scaling of the configuration generation, flowing procedure, and input/output operations. Initialization time is included. The number of cores goes from $N_p = 2$ to $N_p = 2048$. The lattice size was $N^3 \times N_T = 16^3 \times 32$. The flow scaling loaded and flowed a single configuration with $N_{\mathrm{flow}} = 1000$ flow steps, the IO scaling wrote ten configurations to file and the configuration generation scaling generated a single configuration with $N_{\mathrm{corr}} = 600$ and $N_{\mathrm{up}} = 30$.

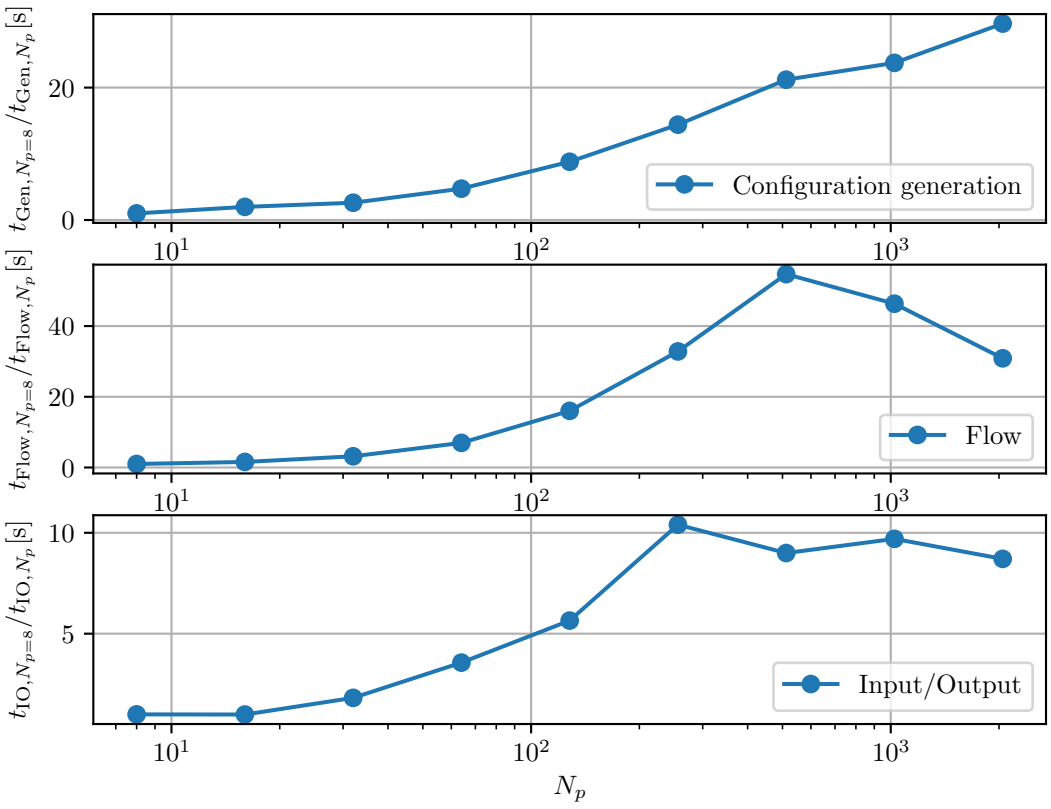


Figure 5.5: The speedup of the configuration generation, flowing, and IO. The speedup is calculated by dividing the run time of each N_p run, with the run time of the run with the least number of processors, $N_p = 8$. The parameters of the speedup is identical to those in fig. 5.3 on page 73 for the strong scaling.

Speedup

The speedup of the program is defined as

$$S(p) = \frac{t_{N_{p,0}}}{t_{N_p}}, \tag{5.2}$$

where $t_{N_{p,0}}$ is the timed run with the least threads, while t_{N_p} is all of the other times. Setting $N_{p_0} = 8$, we get what that is seen in fig. 5.5 on the previous page, where it appears we are plateauing around $N_p = 256$ cores for IO[5], $N_p = 512$ for flowing, while for configuration generation we appear to still be gaining improvements in speed by adding more cores.

5.2 Thermalization

Since we have three options of lattice initialization, cold start, random hot start and random hot start with matrices close to unity (see section 4.1.4 on page 61), runs was made into investigating effects of these initializations. The setup for these runs was $\beta = 6.0$ and a lattice size of $N = 24$ and $N_T = 48$

In fig. 5.6 on the next page we see how topological charge evolves for the first 10^4 Monte Carlo updates. What we see is that there is really not much of an evolution in terms of the value of Q. This makes somewhat sense considering the expectation value for $\langle Q \rangle$ is zero. However, for the energy, we see in fig. 5.7 on page 78 that the cold start actually approaches the stabilized value somewhat quicker. This is confirmed when plotting the relative difference $|E - \langle E \rangle| / \langle E \rangle$ in fig. 5.8 on page 79.

Our goal is not to optimize the thermalization time, but to assure ourselves that we spend enough time thermalizing the system so that it is safe to begin sampling. We use a minimum of $N_{\text{therm}} = 2 \times 10^4$ thermalization updates. Although we only tested this for $\beta = 6.0$, we assume that 2×10^4 updates are enough Monte Carlo updates for thermalization, even though larger lattices and smaller lattice spacings may need longer time thermalizing given the increase in autocorrelations caused by critical slowdown.

5.3 Random matrix generation step size

Different ϵ_{rnd} was tested in order to judge the effects of different step sizes when generating random matrices close to unity. The different ϵ_{rnd} we tested for can be seen in table 5.1 on page 81.

In fig. 5.9 on page 80 we see how the τ_{int} for topological charge behave for different ϵ_{rnd} at flow time $t_f/a^2 = 4.0$. What we see is that for $\epsilon_{\text{rnd}} = 0.24$ the autocorrelation appears to be at an absolute minimum. An exploration into several ϵ_{rnd} between 0.2-0.3

[4]See Xavier and Iyengar [80, chap. 1.6]

[5]After the scaling tests had been performed, an improvement were found for the IO, which lead to a factor ten in performance. The scaling results shown here should thus be taken merely as suggestive for the IO performence.

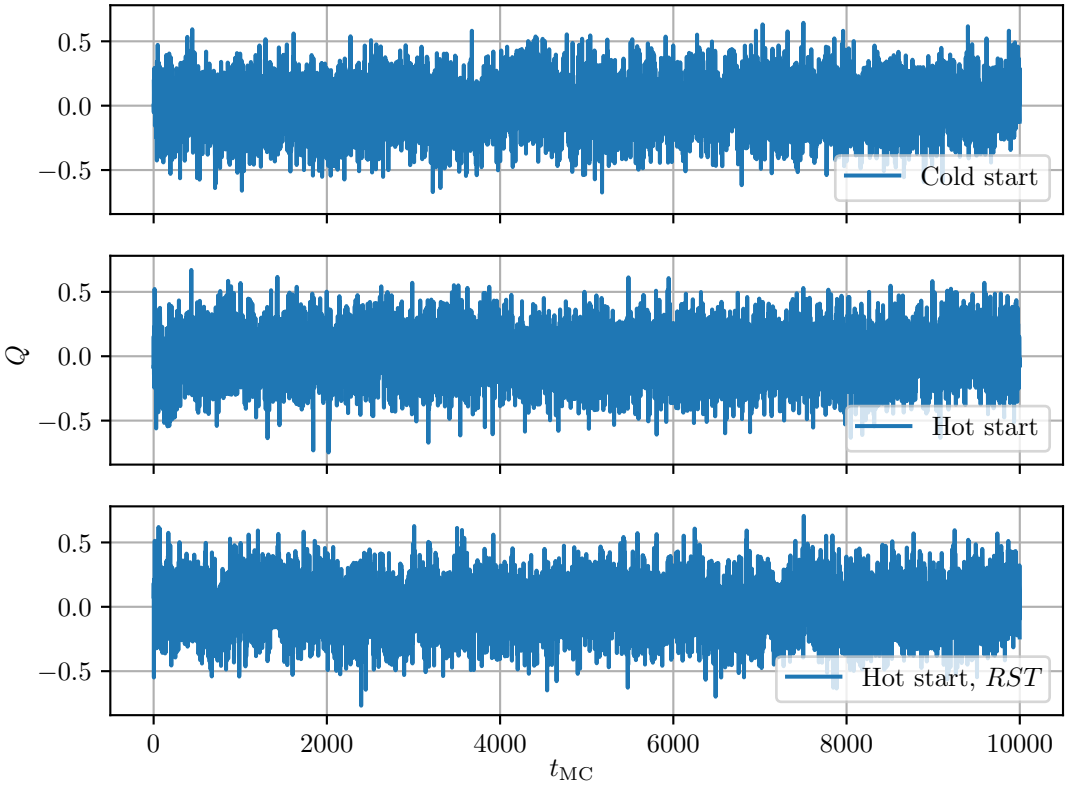


Figure 5.6: Thermalization of topological charge Q from eq. (3.91) on page 42 versus Monte Carlo time for cold start, hot random start and hot RST start.

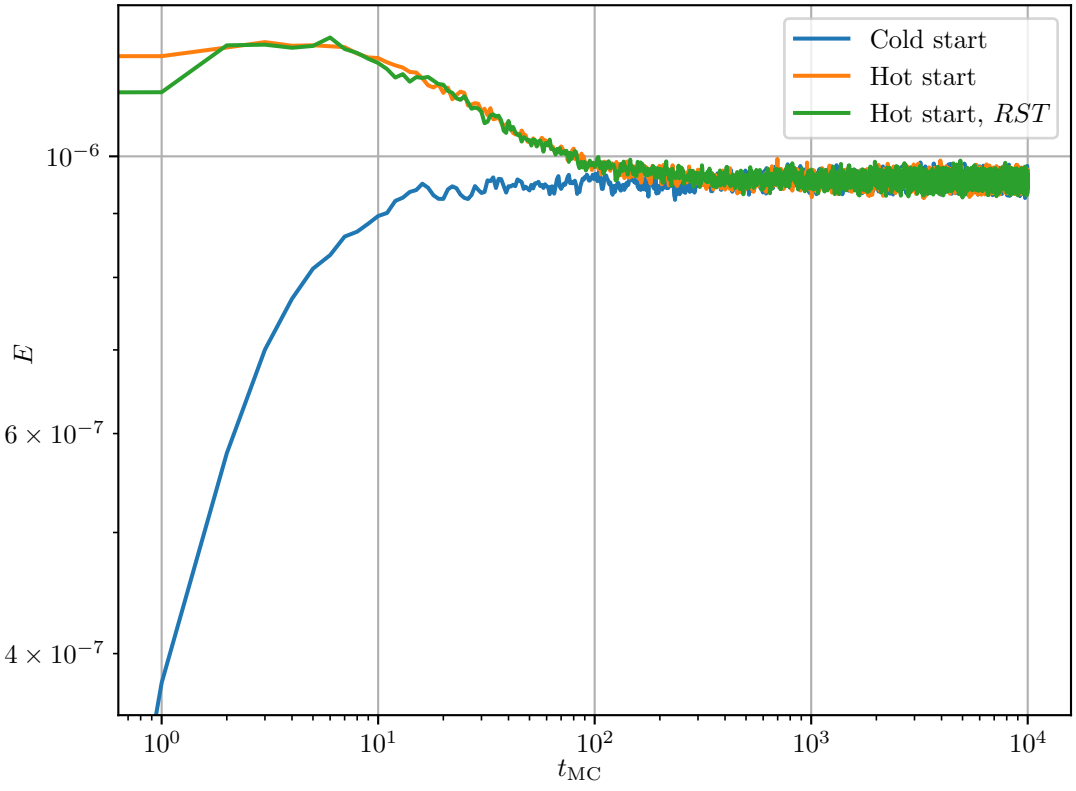


Figure 5.7: Thermalization of the energy E from eq. (3.83) on page 41 versus Monte Carlo time for cold start, hot random start and hot RST start.

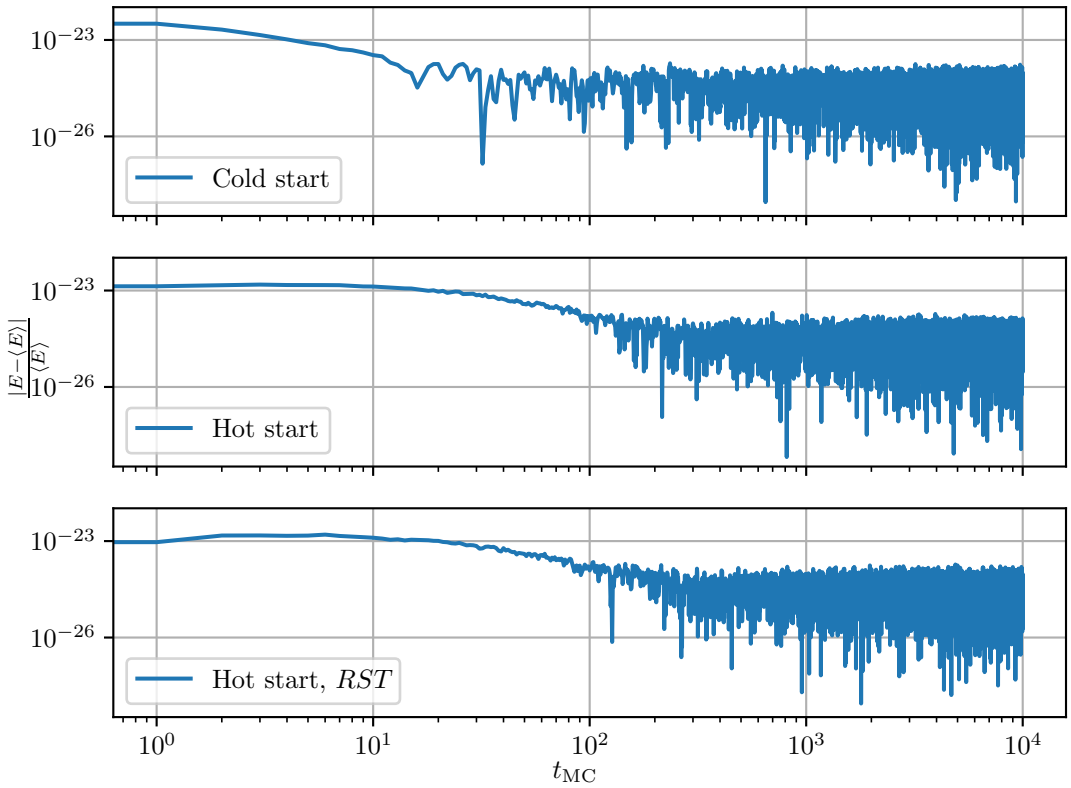


Figure 5.8: Thermalization of the relative difference between E and $\langle E\rangle$ as defined by eq. (3.83) on page 41 versus Monte Carlo time for cold start, hot random start and hot RST start.

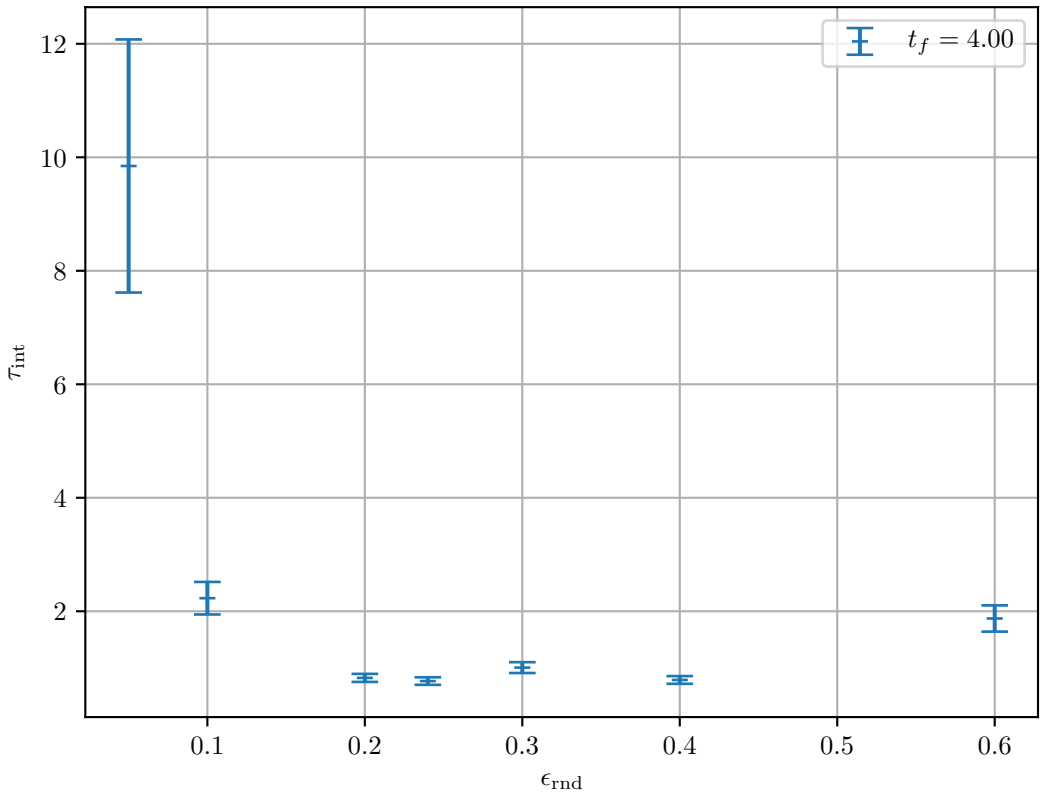


Figure 5.9: The integrated autocorrelation time τ_{int} for $\beta = 6.0$, $N = 6$ and $N_T = 12$ for the topological charge at flow time $t_f/a^2 = 4.0$ with different ϵ_{rnd} flow steps.

Table 5.1: Different ϵ_{rnd} SU(3) matrix generation update lengths.

ϵ_{rnd}	0.05	0.10	0.20	0.24	0.30	0.40	0.60

could be done in order to determine exactly what ϵ_{rnd} that minimize the autocorrelation, but for this thesis we will settle on using $\epsilon_{\text{rnd}} = 0.24$ since that appears to minimize the autocorrelation from the tested values.

5.4 Flow step size

An investigation into different flow integration step sizes was performed in order to verify the choice of $\epsilon_f = 0.01$ and $\epsilon_f = 0.02$. The different ϵ_f investigated can be seen in table 5.2. The tests for different ϵ_f-sizes were all run for a single lattice configuration of size $N = 24$ and $N_T = 48$, and $\beta = 6.0$.

Table 5.2: Different flow integration steps ϵ_f tested for the numerical Runge-Kutta 3 integrator.

ϵ_f	0.001	0.005	0.007	0.009	0.01	0.02	0.03	0.05	0.1	0.5

In fig. 5.10 on the next page we see how the flow behaves for different flow steps. It is worth noticing that $\epsilon_f = 0.5$ completely misses the integration path, which tells us we are outside of Runge-Kutta 3's stability region.

The reason for the sparsity of points seen in fig. 5.11 on page 83 has to do with the constraint of only selecting points that are close to each other in order to compare relatively equal values of ϵ_f. This could have been partly remedied if we let the flow run significantly longer for the smallest ϵ_f.

For this thesis, we will be using $\epsilon_f = 0.01$ and $\epsilon_f = 0.02$. Inspecting fig. 5.11 on page 83 and the absolute difference between the smallest flow time step tested $\epsilon_f = 0.001$ and the rest listed in table 5.2, we see that there is at most one order of magnitude in difference between $\epsilon_f = 0.01$ and $\epsilon_f = 0.02$. Since the difference between the smallest step size $\epsilon_f = 0.001$ and $\epsilon_f = 0.01$ and $\epsilon_f = 0.02$ is already small, we will consider it negligible for the rest of the thesis.

5.5 Minimizing the autocorrelation

An overarching goal when generating configurations is always to minimize the autocorrelation. As we will see in the next chapter on results, some quantities such as topological charge tend to be far more autocorrelated than other observables, i.e. the energy or the plaquette. The tools at our disposal for reducing autocorrelations is currently that we can

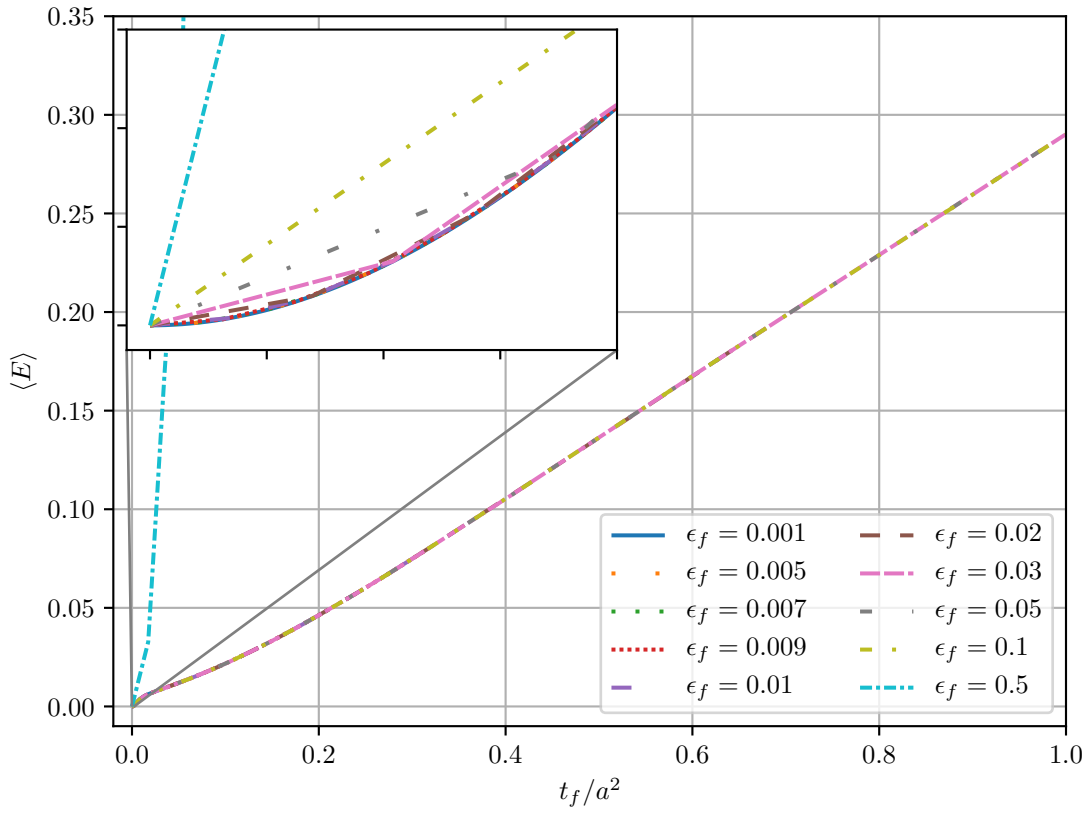


Figure 5.10: The energy for different flow steps as listed in table 5.2 on the previous page on a lattice of size $N^3 \times N_T = 24^3 \times 48$ with $\beta = 6.0$.

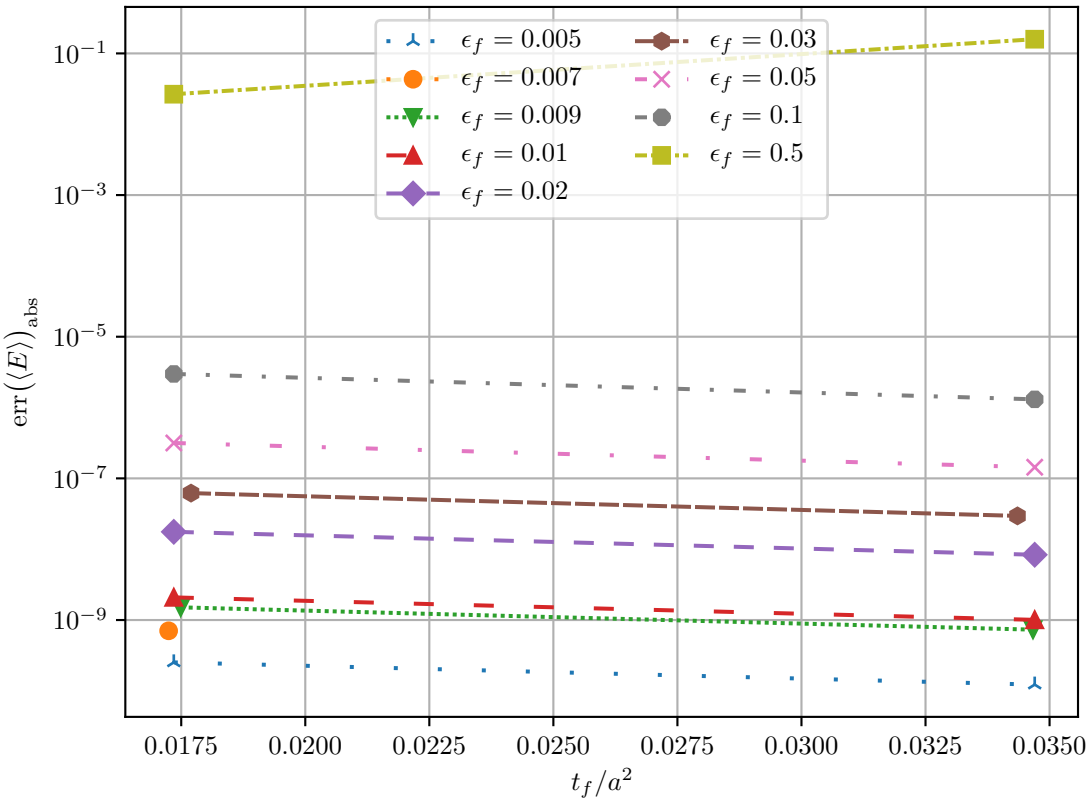


Figure 5.11: The absolute difference between the smallest flow time $\epsilon_f = 0.001$ and those listed in table 5.2 on page 81. The lattice was of size $N^3 \times N_T = 24^3 \times 48$ with $\beta = 6.0$. The reason for only having two points is due to the fact that we are only comparing points that are close to each other in flow time. If we were to have more points, we would have to double the number of flow time steps for the smallest lattices.

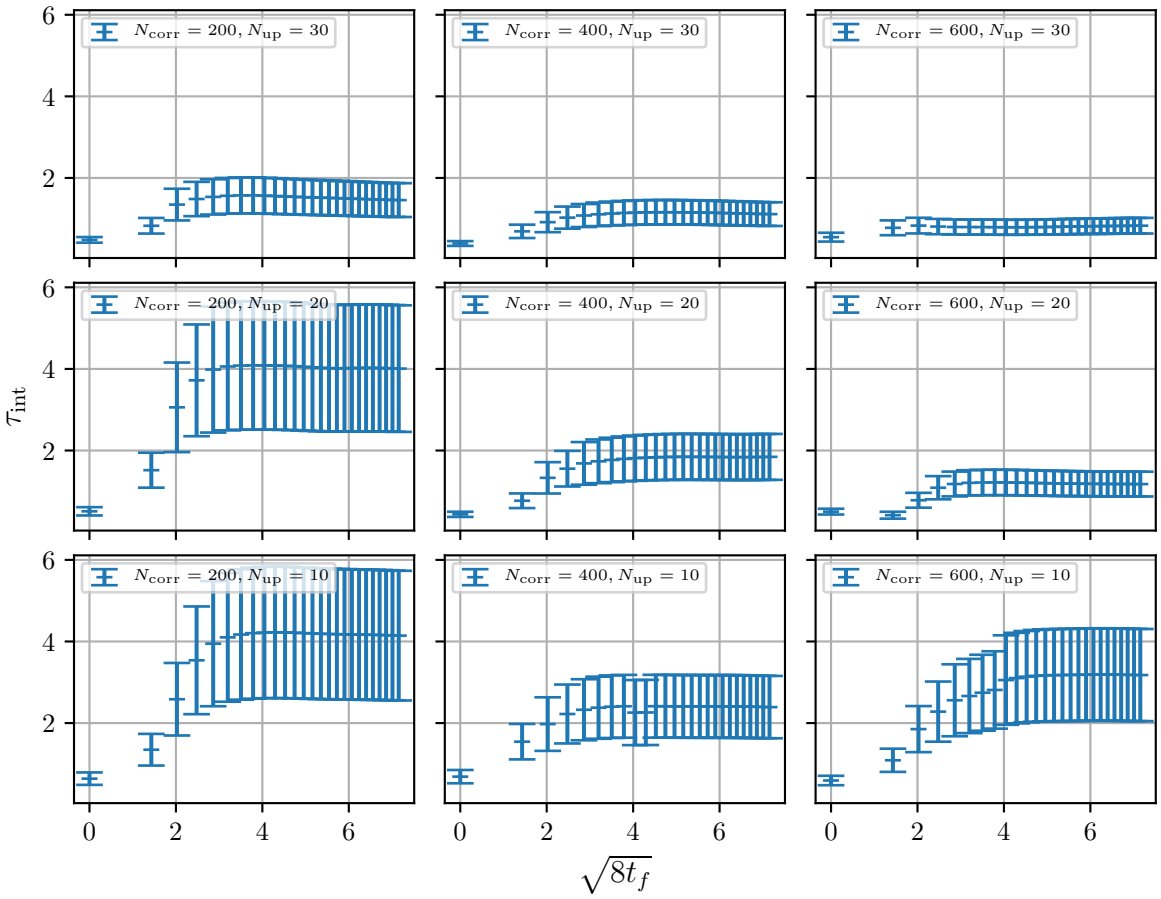


Figure 5.12: The integrated autocorrelation time for topological charge $\langle Q \rangle$ for a lattice of size $N = 16$ and $N_T = 32$ with $\beta = 6.0$ for combinations of $N_{\text{corr}} \in [200, 400, 600]$ and $N_{\text{up}} \in [10, 20, 30]$, plotted against flow time $\sqrt{8t_f}$.

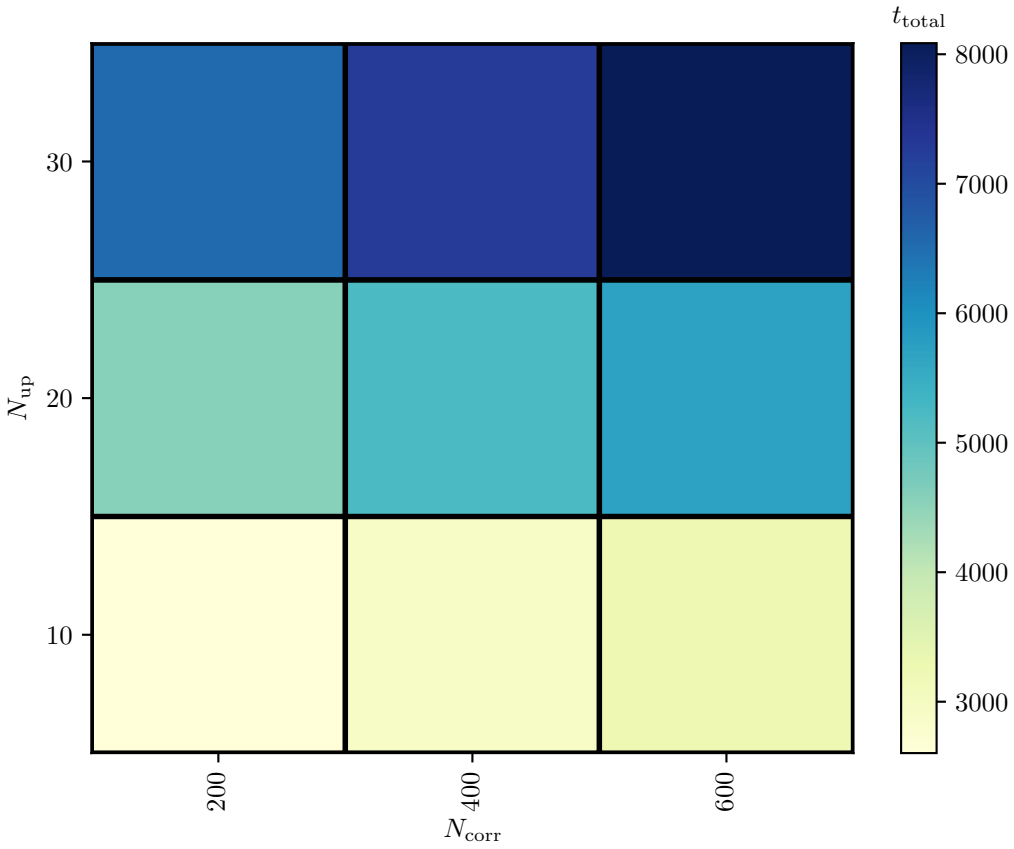


Figure 5.13: The time taking to generate 200 configurations and flowing them $N_{\text{flow}} = 250$ flow steps for a lattice of size $N = 16$ and $N_T = 32$, with $\beta = 6.0$ for combinations of $N_{\text{corr}} \in [200, 400, 600]$ and $N_{\text{up}} \in [10, 20, 30]$.

increase the number of correlation updates, N_{corr} and the number of single link updates N_{up}. The cost, however, is computational time.

An investigation has been made into combining different values of $N_{\text{corr}} \in [200, 400, 600]$ and $N_{\text{up}} \in [10, 20, 30]$ on a lattice of size $N = 16$ and $N_T = 32$ with $\beta = 6.0$. For each of the nine possible combinations, we generated 200 configurations and flowed them $N_{\text{flow}} = 250$ steps with $\epsilon_f = 0.01$.

In fig. 5.12 on page 84 we see how the integrated autocorrelation time behaves in flow time for the topological charge. Inferring from this plot, we see that the autocorrelation seems to decrease the strongest when we increase the number of single link updates and be at an absolute minimum when the number of correlation steps is increased to $N_{\text{corr}} = 600$ together with $N_{\text{up}} = 30$.

From fig. 5.13 on the preceding page we see the behavior for the same set of N_{corr} and N_{up} as we saw in fig. 5.12 on page 84, but instead looking at the total computational time for generating 200 gauge configurations and flowing them $N_{\text{flow}} = 250$ steps. What is immediately clear is that N_{up} costs less computational time than full lattice sweeps such as N_{corr}. For this thesis we chose the standard of $N_{\text{corr}} = 600$ and $N_{\text{up}} = 30$ when generating gauge configurations(except for the smallest lattice spacings). Were we to recommend ways of reducing autocorrelation, increasing the number of single link lattice updates would be it, due to its relatively small impact on CPU time.

5.6 Verifications of the code

As with all code going into production, three types of tests should be implemented,

- **Unit testing**. Testing single functions, such as SU(3) matrix multiplication, inversion and so on, in order to verify that it returns the correct results. In GLAC all of the operations of the SU(2)- and SU(3)-objects are unit tested.

- **Integration testing**. Testing combinations and groups of functions. Ensuring that they work together as intended. In GLAC we test composite elements such as the Lattice objects and the lattice shifts.

- **Validation testing**. Testing that the software returns the desired results based on its specifications. This would be to have actual results we can compare with, such as the output of a similar program or checking that the lattice is gauge invariant. The latter can also be seen as an integration testing, since it including SU(3) matrix objects as well as Lattice objects.

An important part of the validation testing was comparing the output of the topological charge observables with that of *FlowOps*, which is another lattice QCD code based on *Chroma* [20], when flowing the same gauge configuration. We reproduced the output of FlowOps down to machine precision, which tellls us that our implementation is correct[6].

[6]The exact results can be found on the GitHub address of GLAC, `https://github.com/hmvege/GLAC`.

As of today[7] and at the time of generating and flowing the gauge configurations, all of the unit tests, integration tests, and validation tests were passed.

5.7 Future developments and improvements

The code developed for this thesis still has a lot of potential for continued development. Several paths and branches are available for us to expand in. For instance would I have liked to implement many more types of field strength tensors, such as $\mathcal{O}(a^4)$ and $\mathcal{O}(a^6)$ improved clovers as well as improved gauge action. Further, adding fermions seems like a natural extension once this is done. Once fermions are added, usage of GPUs to find the fermion determinants and matrix inverses could also be implemented.

Another addition could be to make the code general to SU(N). This would require changing the methods directly involving SU(N) matrices such as the random matrix generation, but should otherwise not be too difficult to generalize.

5.7.1 Improvements to code

As mentioned, interfacing with GLAC is done through a Python script[8]. This is perhaps something which could be streamlined, such that we for instance would not require multiple `.json` files(one for `createJobs.py` and one for GLAC), but a single one that is updated by `createJobs.py` instead.

Another improvement would be to try and avoid allocating temporary faces when using the shift method as seen in fig. 5.2 on page 70. This could be done by defining the faces that is shared as static objects, only requiring them to be initialized at program start. This could increase the performance of lattice shifts, as we would not have to resize the vector containers for the buffers for every shift call.

When we developed this code using fairly common MPI functionalities. An improvement would be to utilize the capabilities of MPI 3.0, such as local node memory. If we then optimized the lattice geometry such that the threads on each node where neighboring in the hypercube, we could save some time on communication. While on the subject of MPI, we could also switch to using the MPI functionality for setting up a processor grid, thus taking care of the communication structure and removing a possible source of future errors.

Finally, better optimization for SIMD(Single Instruction, Multiple Data) and the cache could be implemented, since we currently did not write the program with this in mind. It should be noted that a few improvements were added for this during the later stages of development, and there might not be that many optimizations available to us.

[7] April 12, 2019

[8] See `https://github.com/hmvege/GLAC/createJobs.py` on GitHub

Chapter 6

Results and discussion

While the focal point of this thesis has been on developing a code capable of generating and flowing gauge configurations, the emphasis when presenting the results will be put on analyzing the topological charge and its properties. We start by going through the ensembles which have been used in the analysis, before moving on to setting a reference scale t_0 and w_0 from the energy. After that, we will go through topological charge, topological susceptibility, and the fourth cumulant. Finally, we will present results for the topological charge correlator and we will attempt at extracting the glueball mass for the state 0^{-+}.

6.1 Production runs

For the main part of the results, five ensembles with different volumes were investigated. Subsequent supporting ensembles were also created in order to verify code or enlighten other aspects of the physics.

Table 6.1: The main ensembles made for this thesis. Every configuration was flown with $N_{\text{flow}} = 1000$ flow steps.

Ensemble	β	N	N_T	N_{cfg}	N_{corr}	N_{up}	ϵ_{flow}	Config. size[GB]
A	6.0	24	48	1000	600	30	0.01	0.356
B	6.1	28	56	1000	600	30	0.01	0.659
C	6.2	32	64	2000	600	30	0.01	1.125
D_1	6.45	32	32	1000	1600	30	0.02	0.563
D_2	6.45	48	96	250	1600	30	0.02	5.695

For all of the main ensembles, we used $N_{\text{therm}} = 20000$ sweeps to thermalize the lattice in order to make sure we had no contamination from the initial state. Each lattice was initialized with a hot start consisting of RST random matrices(see algorithm 2 on page 62), with $\epsilon_{\text{rnd}} = 0.24$, consistent with the findings of the ϵ_{rnd} analysis seen in fig. 5.9 on page 80.

Each of the configurations generated was flown $N_{\text{flow}} = 1000$ steps. Further details on each of the ensembles can be seen in table 6.1 on the previous page.

Most of the ensembles were generated on the Abel supercomputer at the University of Oslo. The D_2 ensemble was generated on the Laconia supercomputer at Michigan State University.

As we will see when investigating quantities such as the topological susceptibility, the D_2 ensemble will somewhat diverge compared to what we would expect from ensemble A, B, and C. Due to this observation, the D_1 ensemble with $\beta = 6.45$ and the lattice size 32^4 was generated in order to illuminate some of the D_2 idiosyncrasies and see if what we were experiencing was just a sampling and low statistics issue. This ensemble was generated on Abel.

Table 6.2: Supporting ensembles made on Smaug. All ensembles were flown $N_{\text{flow}} = 1000$ steps with $\epsilon_{\text{flow}} = 0.01$.

Ensemble	N	N_T	N_{cfg}	N_{corr}	N_{up}	a [fm]	L [fm]
E	8	16	8135	600	30	0.0931(4)	0.745(3)
F	12	24	1341	200	20	0.0931(4)	1.118(5)
G	16	32	2000	400	20	0.0790(3)	1.265(6)

In addition to the ensembles listed in table 6.1 on the preceding page, several supporting ensembles were created on the Smaug cluster, available to M.Sc. students at Computational Physics at the University of Oslo. These can be seen in table 6.2. For the smallest lattices, we are flowing beyond their lattice volumes which we have to keep in mind when analyzing their data. The smallest lattice $8^3 \times 16$ has a side length of $L = 0.75$ fm, and since we are flowing $N_{\text{flow}} = 1000$ steps using a step length $\epsilon_{\text{flow}} = 0.01$, we will smear with a maximal radius of $\sqrt{8t_{f,\max}} = 0.83$ fm which is well beyond $L = 0.75$ fm. For all of the other ensembles in this thesis, the lattice size is larger than the flow radius.

When sampling the ensembles, we exclusively used the clover field strength tensor from eq. (3.80) on page 40, in both the energy in eq. (3.83) on page 41 and the topological charge definition in eq. (3.85) on page 42.

6.2 Setting a reference scale

As mentioned in chapter 3 on page 25, we need to introduce a reference scale in order to compare our results with the physical world. We will start by looking at the t_0 flow reference scale, which is retrieved from the energy at some flow time t_f(see eq. (3.169) on page 53).

After finding t_0, we will look at the flow time derivative of the energy $W(t_f)$ from eq. (3.171) on page 54 and the associated reference scale w_0(see eq. (3.170) on page 54).

6.2.1 The energy

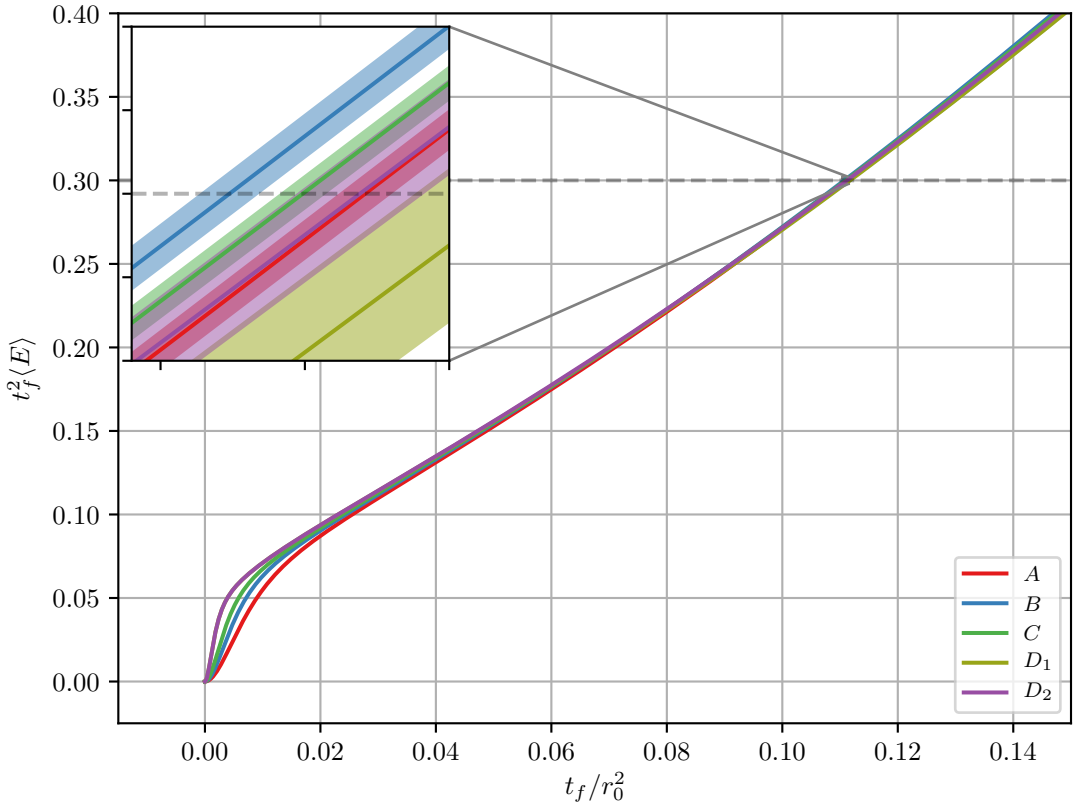


Figure 6.1: The energy results as given by eq. (3.83) on page 41 for table 6.1 on page 89.

As a warm-up, let us inspect fig. 6.1 for the energy from eq. (3.83) on page 41 and how it evolves in flow time t_f. A line has been drawn at $t_f^2 \{\langle E\rangle\}|_{t_f=t_0} = 0.3$, and we have zoomed in to see the behavior around the intersection between $\langle E\rangle$ and this line.

When the flow smearing exceeds the lattice spacing, such that $\sqrt{8t_f}/a \gg 1$, the discretization effects will be less visible. This can be seen if we compare ensemble A with ensemble D_2 in fig. 6.1, where we see that it takes longer for A to be smeared so that we no longer have any discretization effects compared to D_2.

In order to exactly select the flow time t_0 of $t_f^2 \langle E\rangle$ that intersects 0.3, we performed a linear fit using the procedure seen in appendix D.5 on page 156 for $N_{\text{bs}} = 500$ bootstrap samples. For each of the fitted bootstrap samples we invert and select the exact value that intersects 0.3, and finally take the mean to get an estimate for t_0. From the estimated t_0-values seen in table 6.3 on the next page. Using only ensembles with matching volumes, i.e. A, B, C and D_2, we perform a continuum extrapolation which can be seen in fig. 6.2 on

Table 6.3: Extrapolation results for t_0, where we retrieved the exact point of intersection between $t_f^2 \langle E \rangle$ and 0.3 using $N_{\text{bs}} = 500$ bootstrap fits. Extrapolating to the continuum gives us $t_{0,\text{cont}}/r_0^2 = 0.11087(50)$.

Ensemble	t_0[fm^2]	t_0/a^2	t_0/r_0^2	L/a	L [fm]	a [fm]
A	0.02780(2)	3.20(3)	0.11121(9)	24	2.235(9)	0.0931(4)
B	0.02769(2)	4.43(4)	0.11075(10)	28	2.214(10)	0.0791(3)
C	0.02775(2)	6.01(6)	0.11099(8)	32	2.17(1)	0.0679(3)
D_1	0.02779(5)	12.2(1)	0.1112(2)	32	1.530(9)	0.0478(3)
D_2	0.02794(9)	12.2(1)	0.1117(3)	48	2.29(1)	0.0478(3)

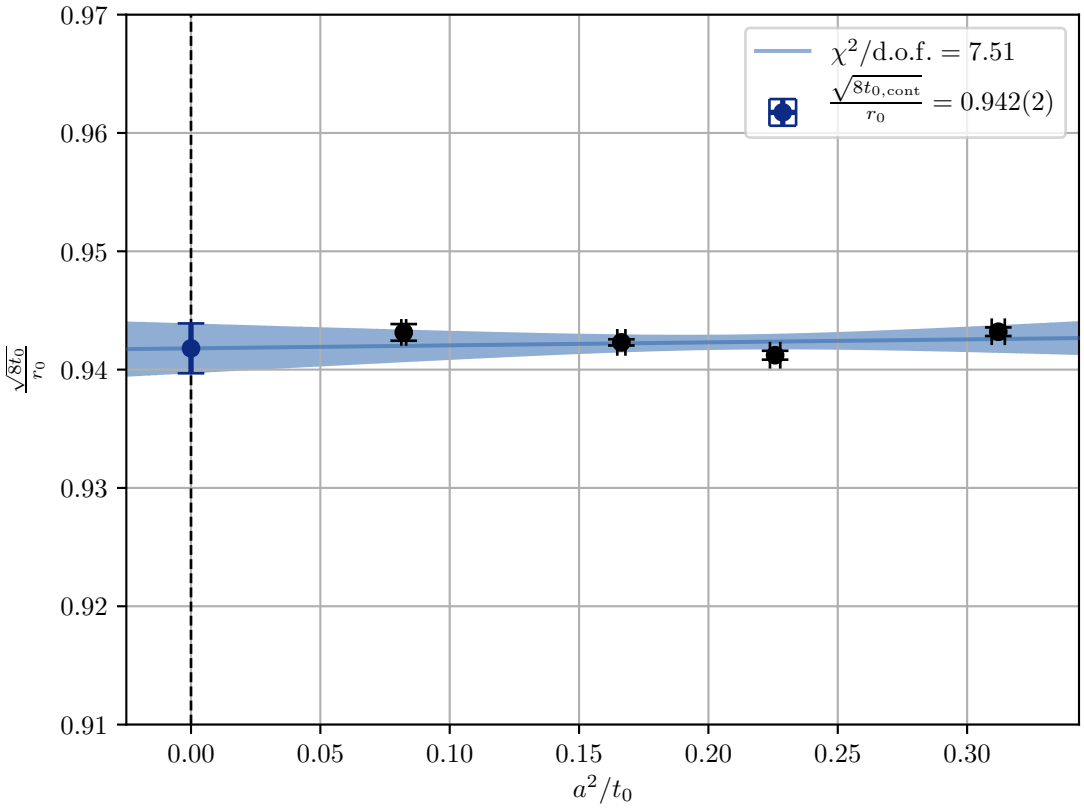


Figure 6.2: The continuum extrapolation $a \to 0$ for t_0 of the four ensembles A, B, C, and D_2 seen in table 6.1 on page 89.

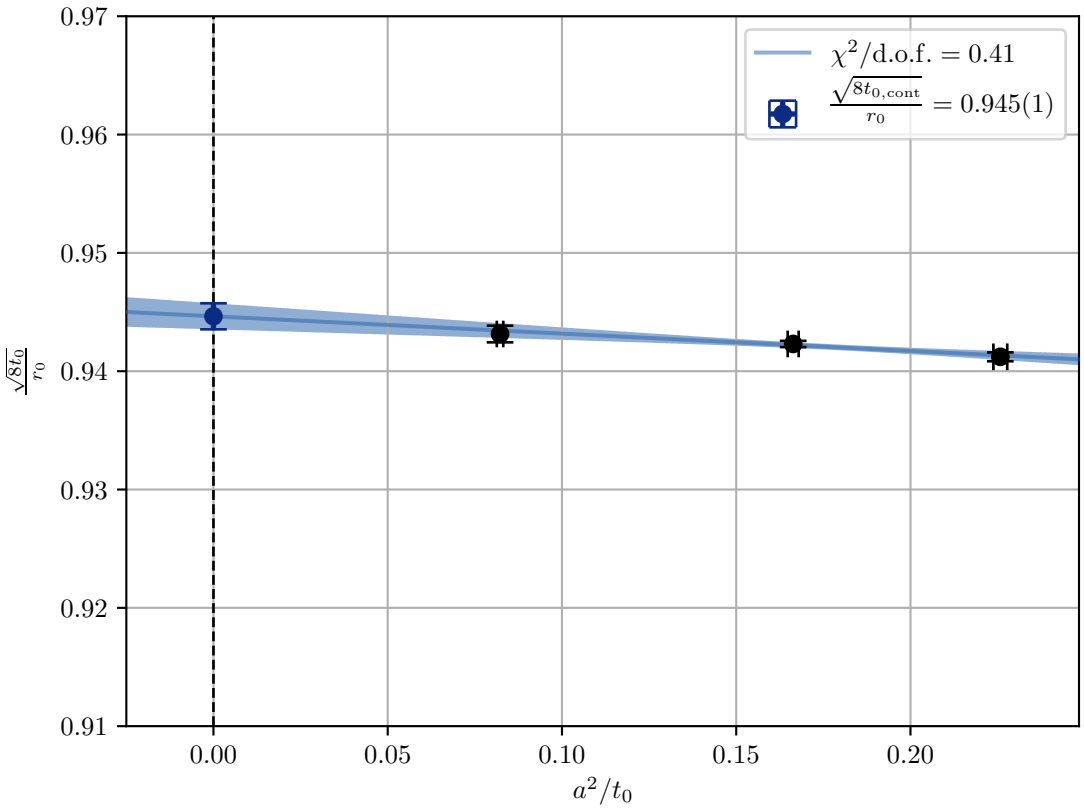


Figure 6.3: An extrapolation of t_0 to the continuum $a \to 0$ using the ensembles B, C, and D_2.

page 92. Again, we use the method for line fitting presented in appendix D.5 on page 156. The resulting t_0 is then

$$\frac{t_{0,\text{cont}}}{r_0^2} = 0.11087(50). \tag{6.1}$$

Comparing our continuum extrapolation seen in fig. 6.2 on page 92, we see we have a good match with that presented by Lüscher [40].

Removing ensemble A from the continuum extrapolation in fig. 6.2 on page 92 gave us

$$\frac{t_{0,\text{cont}}}{r_0^2} = 0.1115(3), \tag{6.2}$$

and a far better continuum fit with $\chi^2/\text{d.o.f.} = 0.41$. The extrapolation can be seen in fig. 6.3 on the preceding page.

Performing the same analysis, only with D_1 instead of D_2 gives us

$$\frac{t_{0,\text{cont}}}{r_0^2} = 0.1119(6), \tag{6.3}$$

with a $\chi^2/\text{d.o.f.} = 0.88$.

Measuring autocorrelation in the energy

The integrated autocorrelation τ_{int} for the main ensembles can be seen in fig. 6.4 on the next page. The sudden shifts in the τ_{int} is due to the automatic windowing procedure cutoff criteria for changing. See algorithm 8 on page 155 for details on how this procedure is done.

When later comparing with the autocorrelation for the topological charge and susceptibility we will see that τ_{int} in fig. 6.4 on the next page comparatively low to that of other quantities.

6.2.2 The energy derivative

From the energy flow time derivative $W(t_f)$ in eq. (3.171) on page 54, we can define a scale w_0 similar to that of t_0. The derivative $W(t_f)$ can be seen in fig. 6.5 on page 96, where we have visibly more pronounced discretization effects for small flow times. When implementing the numerical derivative for $E(t_f)$, we used a symmetric definition for the derivative discretization.

We extract the w_0-values in the same manner as we did for the t_0-values. The different extracted w_0 for each of the ensembles can be seen in table 6.4 on the facing page, and a continuum extrapolation of these(only including the ensembles of similar volume, thus excluding D_1) can be seen in fig. 6.6 on page 97. The extrapolation is performed using the fitting procedure given in appendix D.5 on page 156. The resulting w_0 was found to be

$$w_{0,\text{cont}} = 0.1695(5) \text{ fm}. \tag{6.4}$$

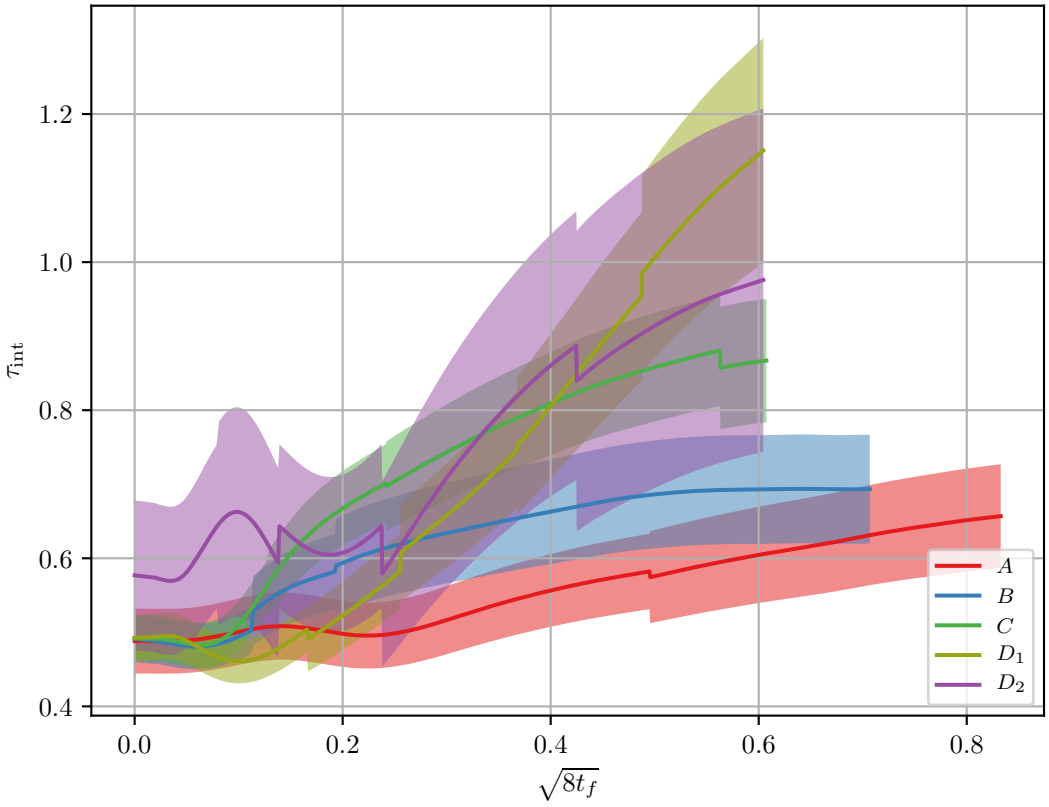


Figure 6.4: The autocorrelation of the energy, calculated using the method presented in appendix D.3.3 on page 153. A value of $\tau_{\text{int}} = 0.5$ indicates that we have zero autocorrelation.

Table 6.4: Extrapolation results for w_0, where we retrieved the exact point of intersection between $t_f W(t_f)$ and 0.3 using $N_{\text{bs}} = 500$ bootstrap fits. Extrapolating to the continuum for the ensembles A, B, C and D_2, gives us $w_{0,\text{cont}} = 0.1695(5)$ fm.

Ensemble	w_0[fm]	w_0/a	L/a	L [fm]	a [fm]
A	0.16787(9)	1.802(7)	24	2.235(9)	0.0931(4)
B	0.16796(9)	2.124(9)	28	2.214(10)	0.0791(3)
C	0.16860(6)	2.48(1)	32	2.17(1)	0.0679(3)
D_1	0.1700(3)	3.55(2)	32	1.530(9)	0.0478(3)
D_2	0.1693(2)	3.54(2)	48	2.29(1)	0.0478(3)

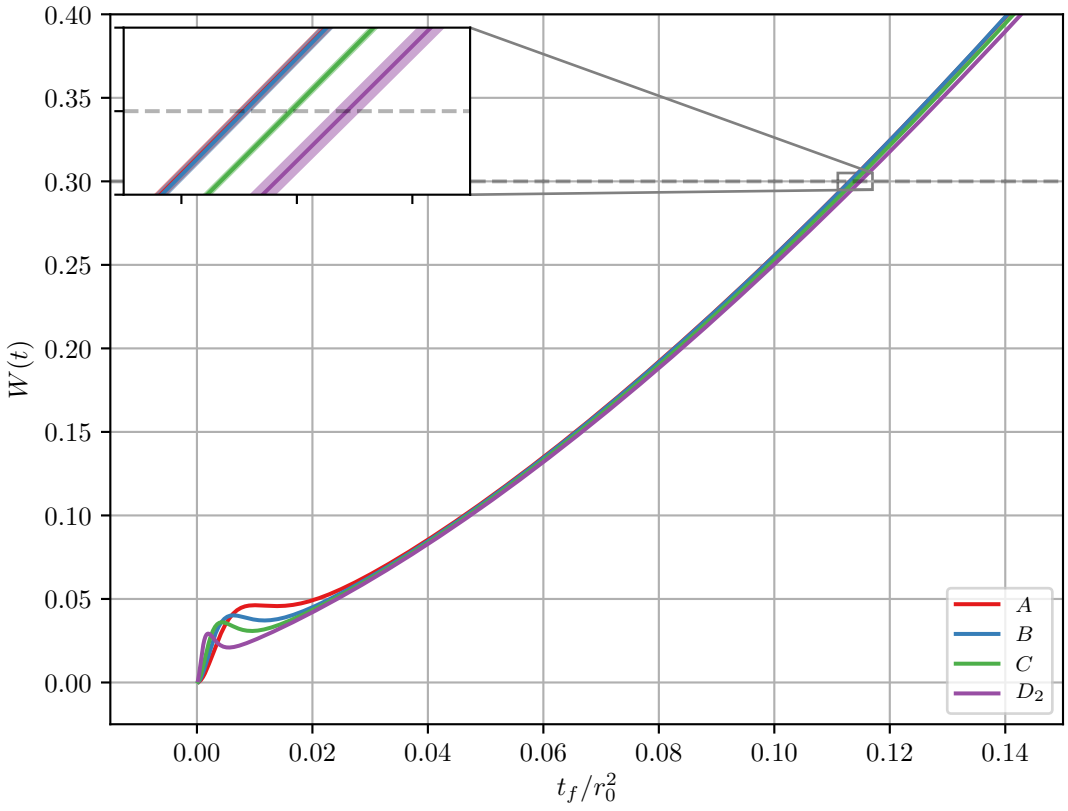


Figure 6.5: The energy derivative $W(t_f)$ as given by eq. (3.171) on page 54.

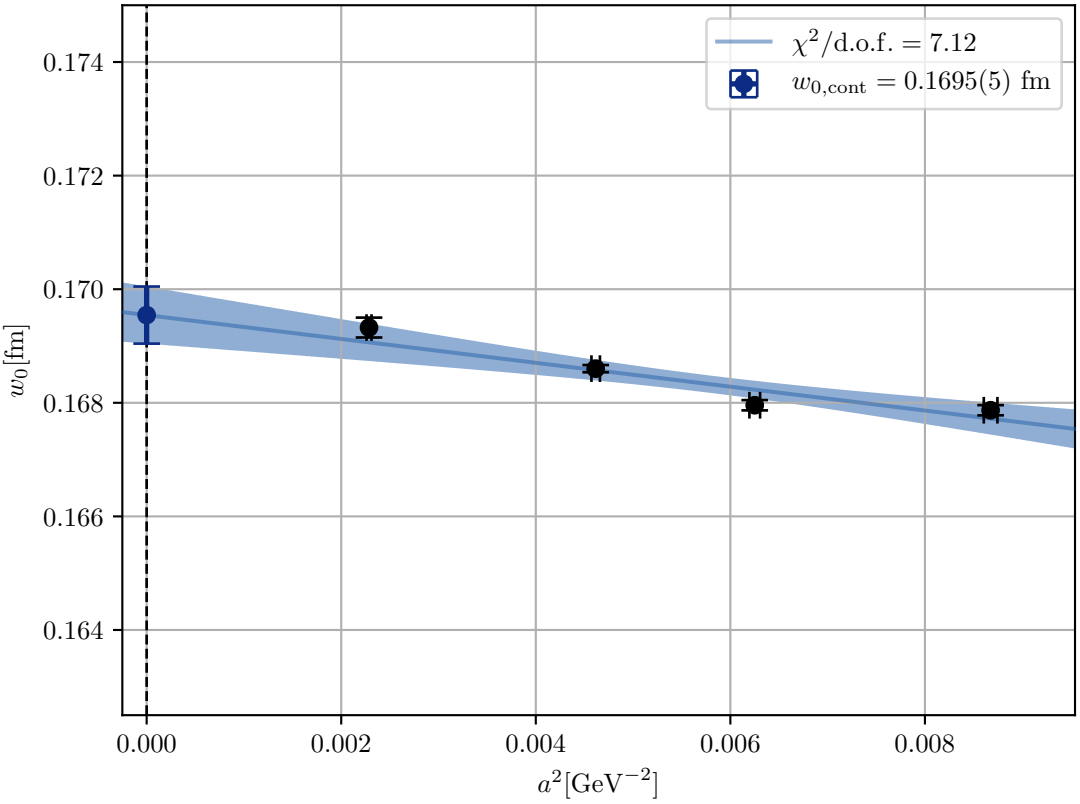


Figure 6.6: An extrapolation of w_0 to the continuum using ensemble A, B, C, and D_2.

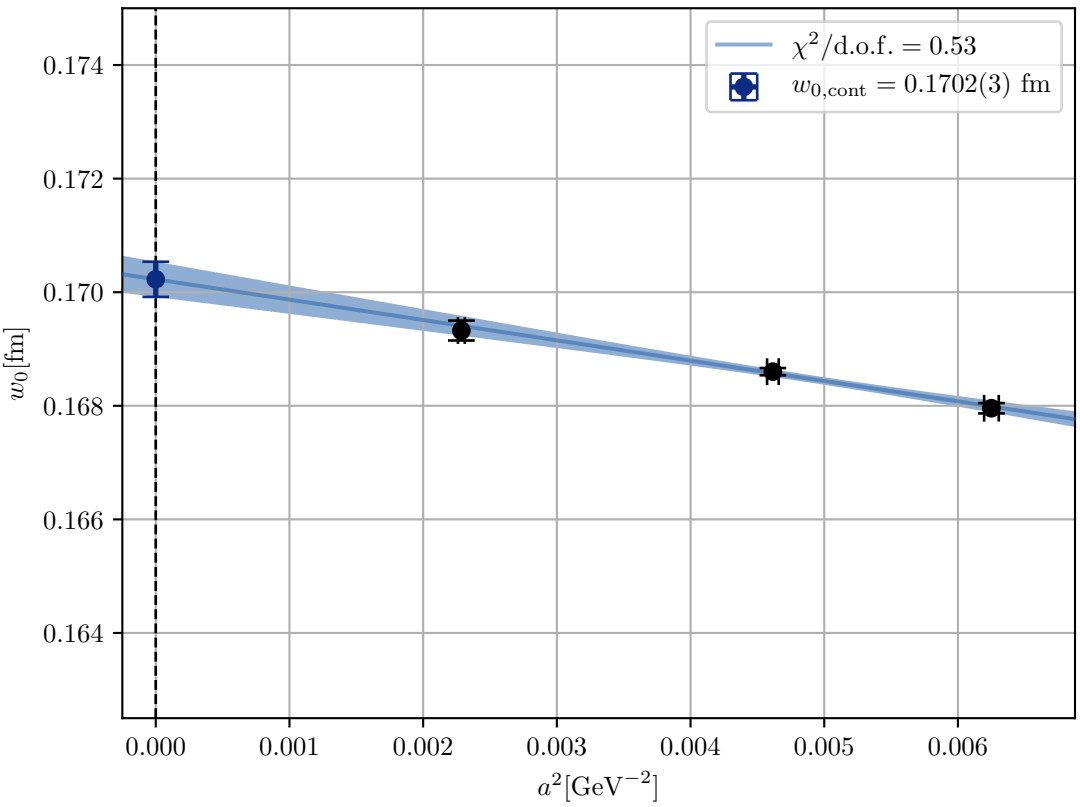


Figure 6.7: An extrapolation of w_0 to the continuum using ensemble B, C, and D_2.

As in a similar fashion for t_0, we remove the A ensemble from the continuum extrapolation in fig. 6.6 on page 97. This gives us

$$w_{0,\text{cont}} = 0.1702(3) \text{ fm}, \tag{6.5}$$

and $\chi^2/\text{d.o.f.} = 0.53$. This extrapolation can be seen in fig. 6.7 on the facing page.

As before, we perform the same analysis only with D_1 instead of D_2. This gives us

$$w_{0,\text{cont}} = 0.1706(6) \text{ fm}, \tag{6.6}$$

and $\chi^2/\text{d.o.f.} = 0.86$.

The extrapolated value for w_0 seem to have a weak dependence on fermions, as our results is comparable to that presented by Borsanyi et al. [5] which included dynamical fermions, with $w_{0,\text{cont}} = 0.1755(18)(04)$ fm.

6.3 Topological charge

In the continuum we expect the topological charge to take on discrete numbers. On the lattice, this will not be the case, as many of the topological properties which we seek to explore is suppressed at $t_f/a^2 = 0$. The topological charge will contain cut-off effects and for the topological susceptibility we even have a UV divergence at zero flow time [55]. Further, these effects will be more pronounced for the lattices with smaller lattice spacings as these are closer to the continuum. In order to remedy these issues, we seek to apply gradient flow.

6.3.1 Smearing effects on small lattices

The number of zero-modes available to the topological charge is highly dependent on the size of the lattice. With a set of supporting runs listed in table 6.2 on page 90, this effect was investigated in fig. 6.8 on the following page. For the smallest ensemble E, with a lattice of size $8^3 \times 16$, we only have $Q \in [-1, 0, 1]$. For the larger lattices F and G, the number of Q available rapidly increases.

Focusing on ensemble G with the $16^3 \times 32$ lattice, we can in fig. 6.9 on page 101 inspect how the flow smears the topological charge such that topological features become more prominent for increasing flow time. The reason for the topological charge not taking on exact integers as one would expect from theory is due to the fact that we have a finite lattice spacing and a chirality breaking action.

6.3.2 Topological charge evolved in flow time

In fig. 6.10 on page 102 we see $\langle Q \rangle$ evolved in flow time t_f for the main ensembles seen in table 6.1 on page 89. Assuringly, we see that for ensemble A, B, C and D_2 the topological charge is compatible with zero within one standard deviation. These results have

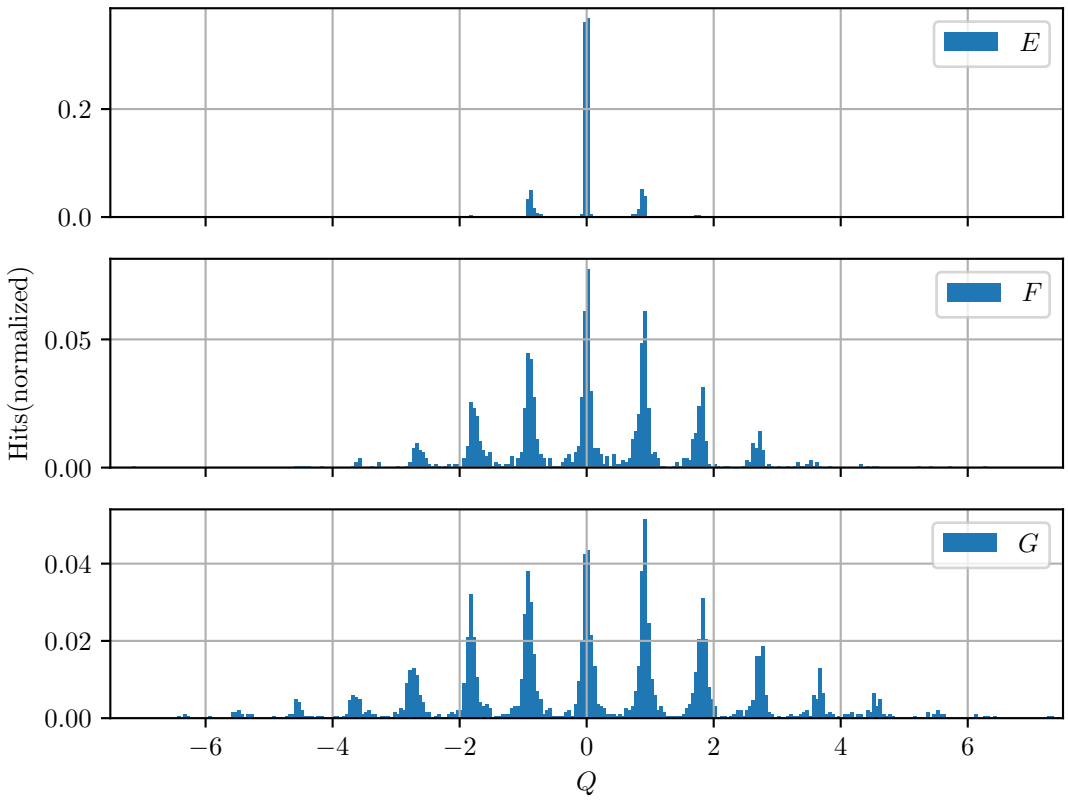


Figure 6.8: Histograms for the topological charge for the ensembles seen in table 6.2 on page 90 at $t_f/a^2 = 0.25$ fm.

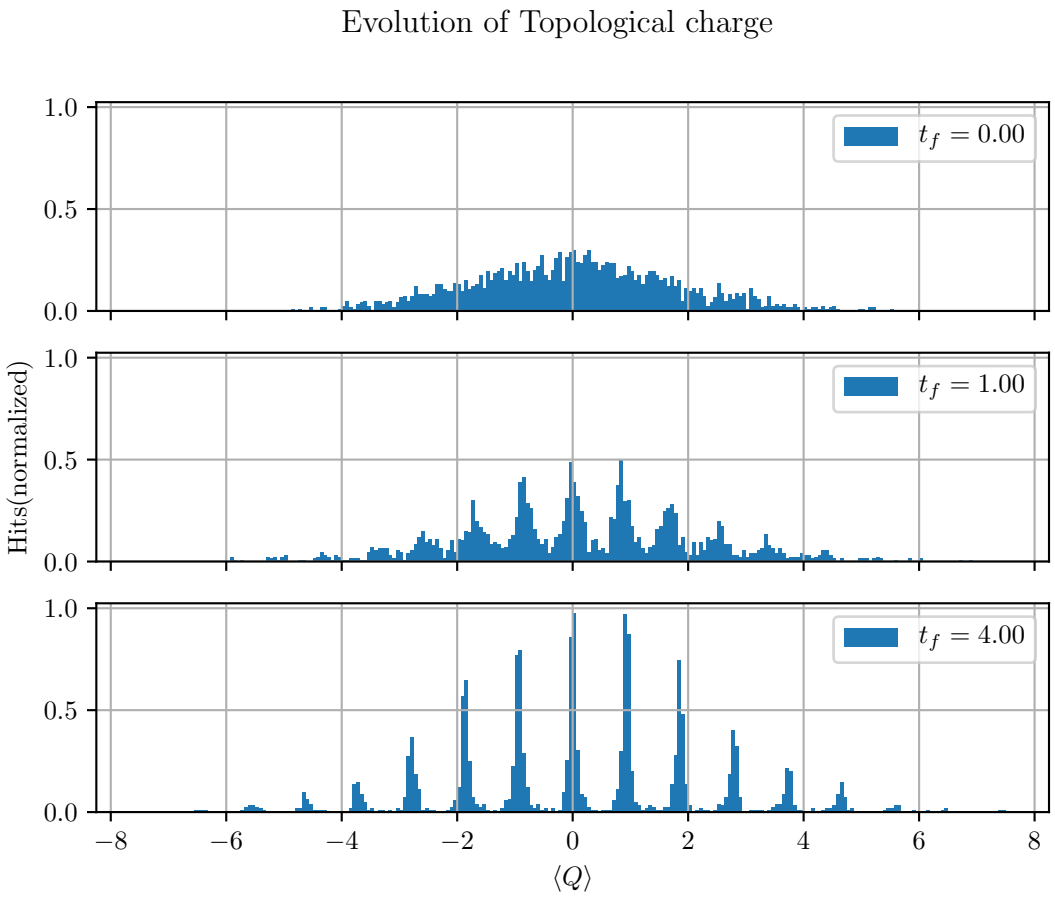


Figure 6.9: Histograms for the topological charge for ensemble G with a lattice of size $N^3 \times N_T = 16^3 \times 32$ with $\beta = 6.1$, taken at different flow times $t_f/a^2 = 0.0$, 1.0, 4.0 fm.

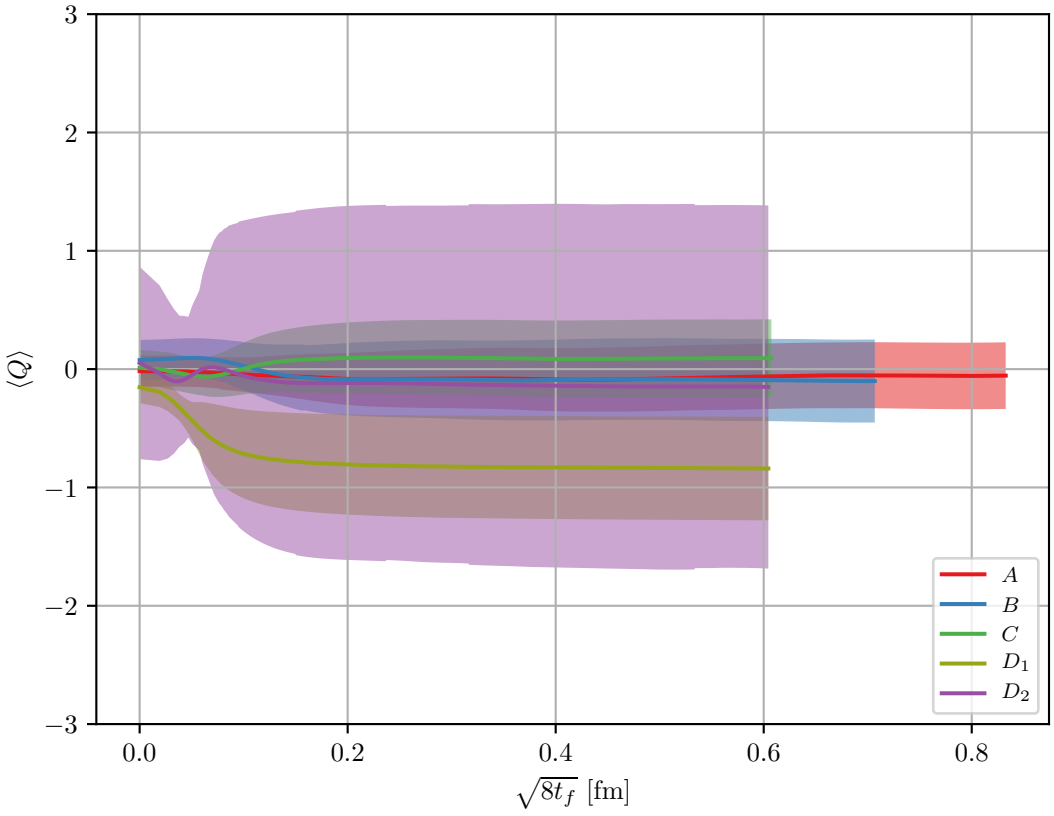


Figure 6.10: Topological charge Q as evolved in flow time for the five ensembles seen in table 6.1 on page 89. $N_{\text{bs}} = 500$ bootstraps have been performed on each ensemble, corrected for autocorrelations with $\sigma = \sqrt{2\tau_{\text{int}}}\sigma_0$.

been corrected for by the integrated autocorrelation time τ_{int} as in equation eq. (D.36) on page 155 and bootstrapped 500 times.

Inspecting fig. 6.10 on the preceding page, we notice that the ensemble D_1 is not centered at zero, but compatible within two standard deviations. While this is not necessarily a tremendous problem, it indicates the need for large statistics and good methods of estimating the error. As we shall see, this ensemble is also highly autocorrelated.

Using the ability to store the charge of the lattice at every point in the hypercube in GLAC, we have flown a configuration with $\beta = 6.2$ and $32^3 \times 64$ ensemble and generated a visualization using LatViz seen in fig. 6.11 on the following page. The frames of the configuration are taken at the Euclidean time $n_T/a^2 = 21$ and at flow times $t_f/a^2 = 0.0, 2.0, 4.0, 6.0$. The visualization allows for an intuitive understanding of gradient flow, as we see that the smearing leaves only the most prominent features of a configuration intact.

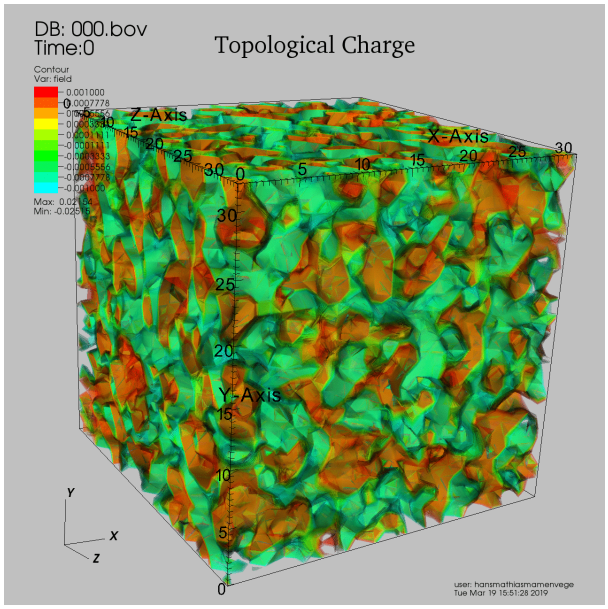
6.3.3 Autocorrelation for topological charge

The integrated autocorrelation τ_{int} can be seen in fig. 6.12 on page 105. We immediately see that the topological charge is quantity far more prone to exhibiting autocorrelations than the energy if we compare with fig. 6.4 on page 95. Considering we will mainly be investigating the topological charge and aspects thereof, correct error estimations in the face of highly correlated data becomes paramount.

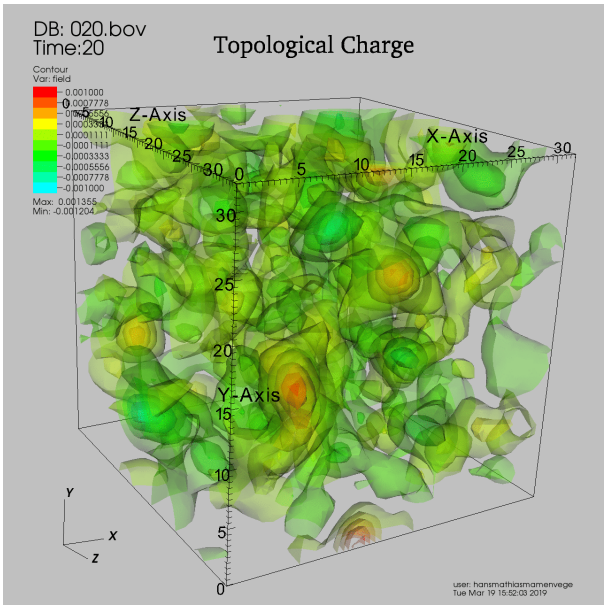
6.3.4 Monte Carlo history for topological charge

With the results for the autocorrelation for the topological charge in mind, we can take a look at the Monte Carlo history at two flow times. In fig. 6.13 on page 106 and in fig. 6.14 on page 107 we see the Monte Carlo history for flow time $\sqrt{8t_f} = 0.0$ fm and $\sqrt{8t_f} = 0.6$ fm. When comparing fig. 6.14 on page 107 with fig. 6.13 on page 106 we see a sharp increase in correlations, as the Monte Carlo history of Q show a clear trend in the data. This is particularly visible for the ensembles D_1 and D_2, reinforcing what we already known from previous sections.

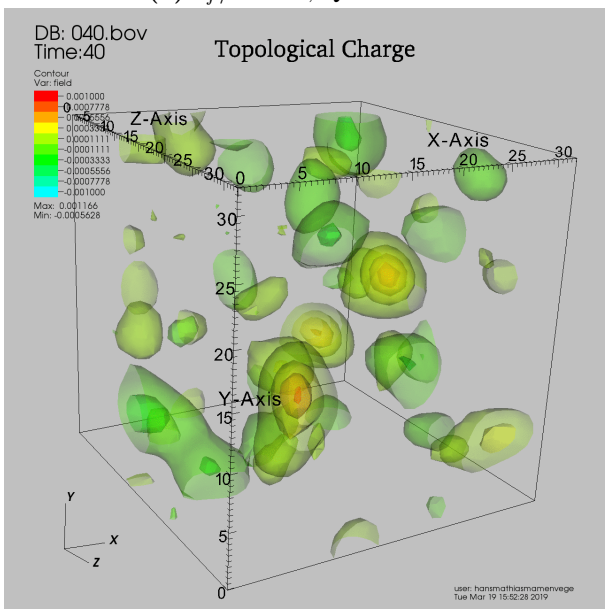
Before investigating the topological susceptibility, it is worth inspecting the autocorrelation in fig. 6.12 on page 105 and in particular for D_1 and D_2. We see that for D_1 the autocorrelation is significantly higher than that of D_2, something which contrasts the errors on the charge seen in fig. 6.10 on the facing page. The error bands of the charge in fig. 6.10 on the preceding page have been corrected by the autocorrelation, so in order to better understand why the errors on D_2 is far higher than that of D_1, we need to take a look at table 6.5 on page 105. Here we have the value of the charge $\langle Q \rangle$ without any bootstrap sampling and autocorrelation correction. We see that for D_2 the error is almost five times larger than that of D_1 at $\sqrt{8t_f} = 0.6$ fm. As we see in the column for the autocorrelation correction factor $\sqrt{2\tau_{\text{int}}}$ in table 6.5 on page 105, the D_1 and D_2 at $\sqrt{8t_f} = 0.6$ fm have respectively $\sqrt{2\tau_{\text{int}}} = 6.49(1.19)$ and $\sqrt{2\tau_{\text{int}}} = 4.72(1.05)$, something which is not sufficient to offset the inherent error of the D_2 set.



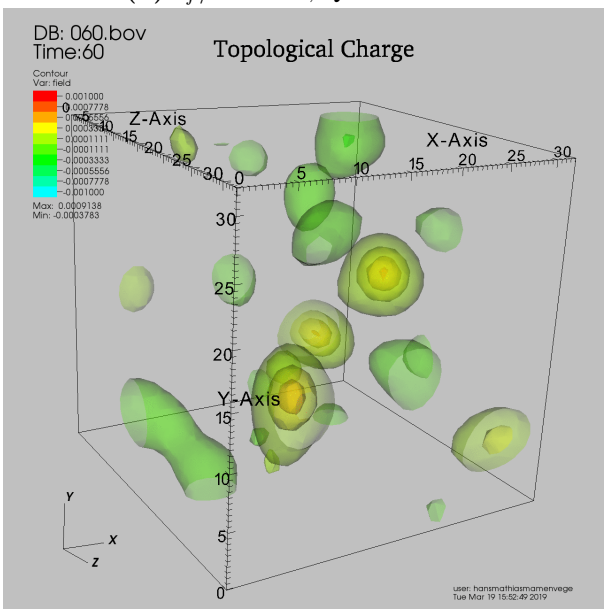
(a) $t_f/a^2 = 0$, $Q = 4.8201$.



(b) $t_f/a^2 = 2.0$, $Q = -8.9930$.



(c) $t_f/a^2 = 4.0$, $Q = -8.9991$.



(d) $t_f/a^2 = 6.0$, $Q = -9.0002$.

Figure 6.11: Stills from a LatViz animation, presenting the smearing of a configuration for a $\beta = 6.2$ with size $32^3 \times 64$ ensemble. The configuration was flowed from $t_f/a^2 = 0.0$ to $t_f/a^2 = 6.0$. The gauge configuration used for this visualization was provided Giovanni Pederiva.

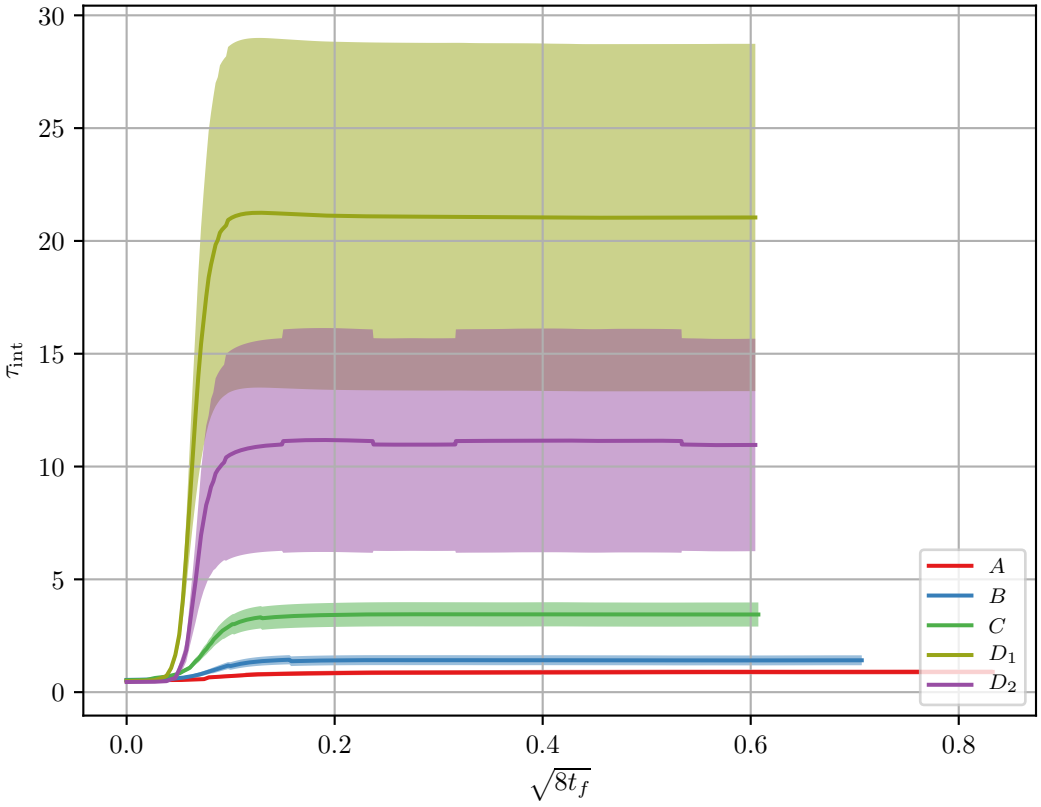


Figure 6.12: The integrated autocorrelation τ_{int} for topological charge for the five ensembles seen in table 6.1 on page 89.

Table 6.5: The topological charge for the ensembles seen in table 6.1 on page 89 at flow time $\sqrt{8t_f} = 0.6$ fm. The third column has had its error corrected by $\sqrt{\tau_{\text{int}}}$. Extra precision is left in some values for making the comparison between non-corrected and corrected values. None of the results have been bootstrapped. A value of $\sqrt{\tau_{\text{int}}} = 1$ would indicate that the ensemble contain no autocorrelation.

Ensemble	$\langle Q \rangle$	$\langle Q \rangle$, corrected	$\sqrt{\tau_{\text{int}}}$
A	$-0.06(20)$	$-0.06(27)$	1.335(80)
B	$-0.08(20)$	$-0.08(33)$	1.68(13)
C	0.09(13)	0.09(35)	2.62(20)
D_1	$-0.84(7)$	$-0.84(45)$	6.49(1.19)
D_2	$-0.14(33)$	$-0.14(1.57)$	4.72(1.05)

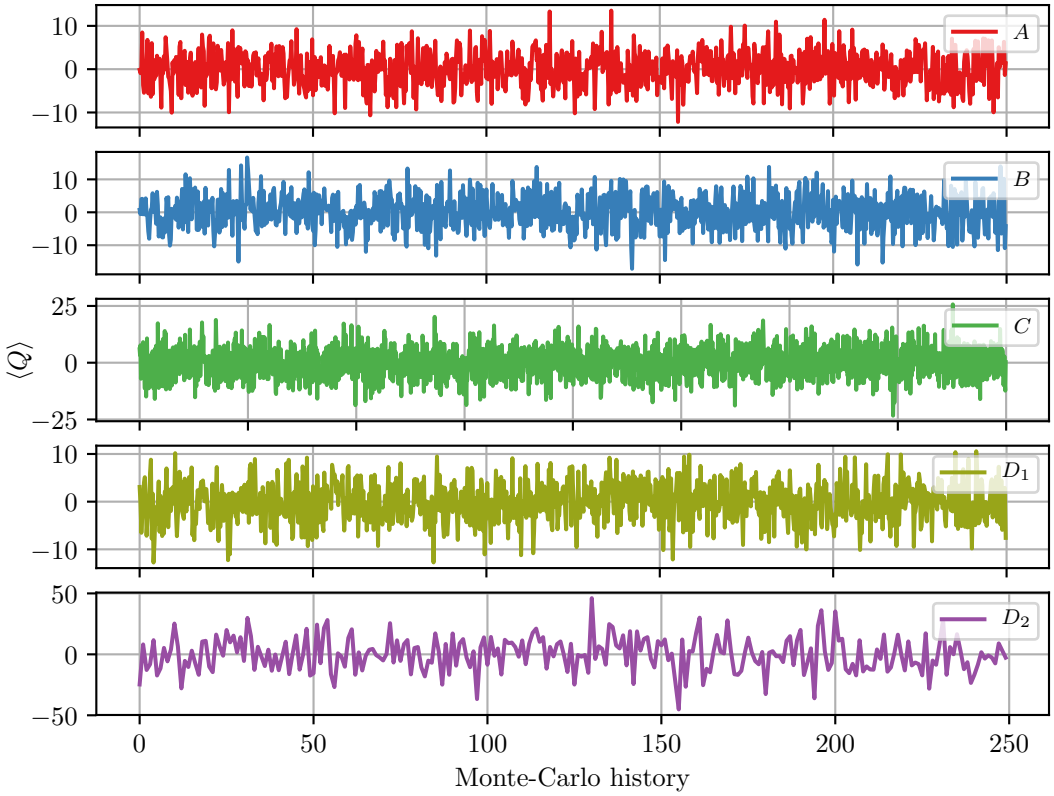


Figure 6.13: The Monte Carlo history for the five ensembles seen in table 6.1 on page 89 at $\sqrt{8t_f} = 0.0$ fm.

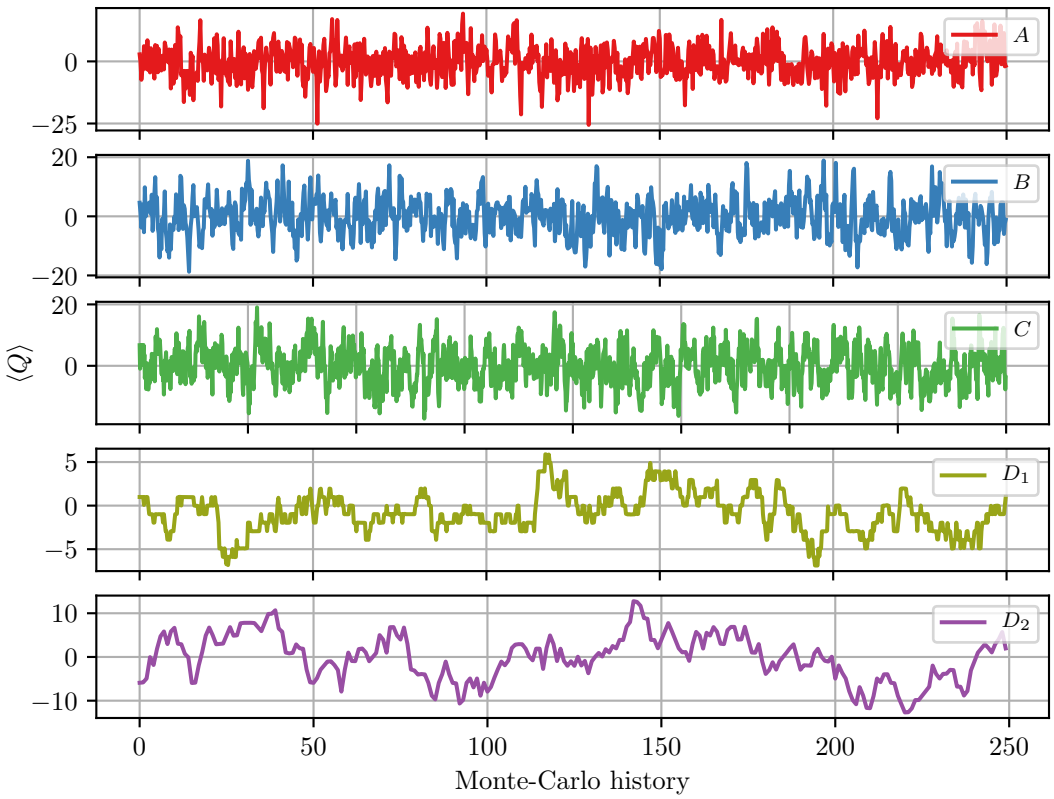


Figure 6.14: The Monte Carlo history for the five ensembles seen in table 6.1 on page 89 at $\sqrt{8t_f} = 0.6$ fm.

It is also likely that the autocorrelation for D_2 should be higher, due to the D_1 dataset being of the same β-value and a longer Markov-chain. A long Markov chain would increase the autocorrelation and thus give a better estimate of the error, with the increased statistics leading to more precise results in general.

Topological charge for different intervals in Monte Carlo time.

Table 6.6: The intervals the topological charge is split into in fig. 6.15 on the facing page for all of our main ensembles seen in table 6.1 on page 89.

Ensemble	MC Intervals
A	$[000-250], [250-500], [500-750], [750-1000]$
B	$[000-250], [250-500], [500-750], [750-1000]$
C	$[000-500], [500-1000], [1000-1500], [1500-2000]$
D_1	$[000-250], [250-500], [500-750], [750-1000]$
D_2	$[000-062], [062-124], [124-186], [186-248]$

To highlight some of the troubles when working with small ensembles, we can take a look at the topological charge Q at four different intervals in Monte Carlo time, as seen in table 6.6. The resulting charge seen in fig. 6.15 on the next page shows us that had we only had, say the 500 first configurations of the C ensemble, we would have gotten an average charge of around $\langle Q \rangle \sim 1.5$. This leads us to believe that with better statistics we would have a centering of the charge for the D_1 ensemble.

6.4 Topological susceptibility

The topological susceptibility for the ensembles listed in table 6.1 on page 89 can be seen in fig. 6.16 on page 110. The integrated autocorrelation can be seen in fig. 6.17 on page 111.

As discussed in [13], the topological susceptibility appears to only depend on the volume up to $L \sim 1.2$ fm to $L \sim 1.4$ fm, and after that, the results are size independent. Thus, as long as we have a lattice of volume larger than $L \approx 1.4$ fm we should get results which are accurate.

Later on in fig. 6.16 on page 110 we see that the D_2 ensemble is quite dissimilar from the other ensembles at flow times larger than $\sqrt{8t_f} > 0.15$. Due to this difference, we generated the D_1 ensemble with the same $\beta = 6.45$ as D_2, but with a different volume of $N^3 \times N_T = 32^4$. Since the side of ensemble D_1 is $L = 1.530(9)$ fm, it should be unproblematic to use.

In fig. 6.16 on page 110 we see that the topological susceptibility is increasingly divergent for smaller lattice spacings at flow time $\sqrt{8t_f} = 0$. This effect is particularly apparent for the D_1 and D_2 ensembles. As mentioned earlier, this is due to the topological susceptibility having a UV divergence at $t_f = 0$ fm [55], which becomes more pronounced for smaller

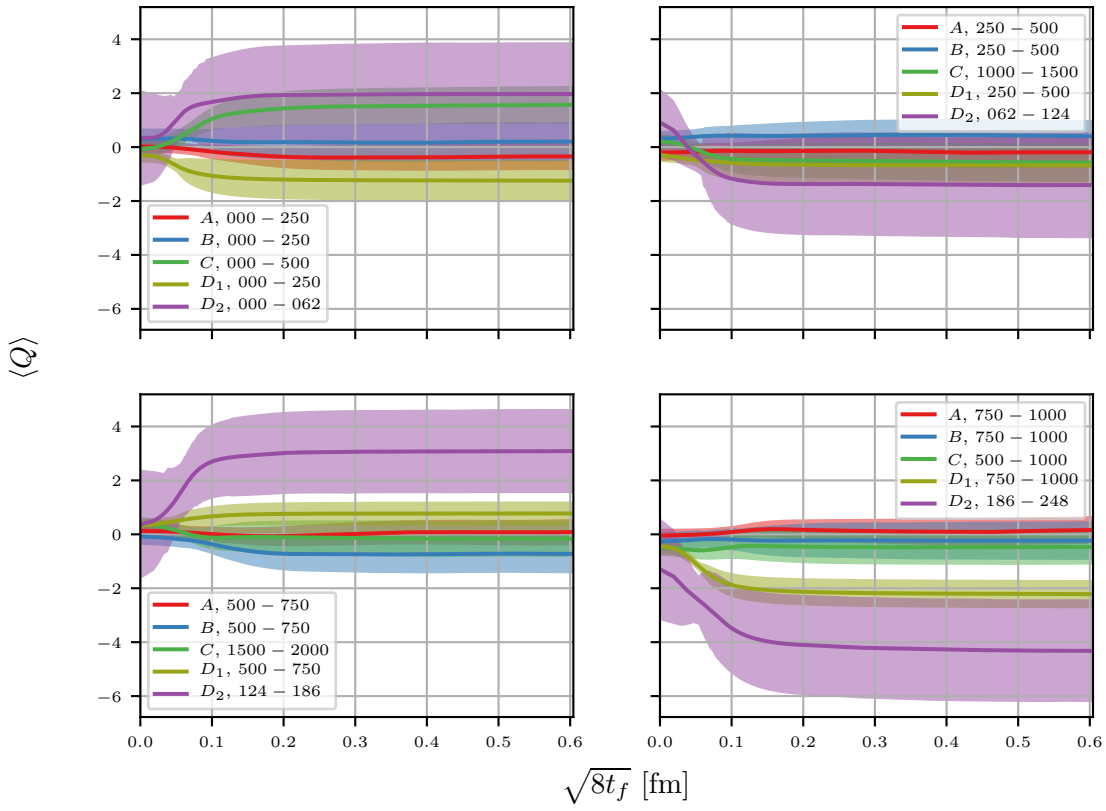


Figure 6.15: The topological charge for each ensemble in table 6.1 on page 89 for four different intervals in Monte Carlo time. The ensembles were analyzed using 500 bootstrap samples and corrected for by the integrated autocorrelation time, $\sqrt{2\tau_{\text{int}}}$.

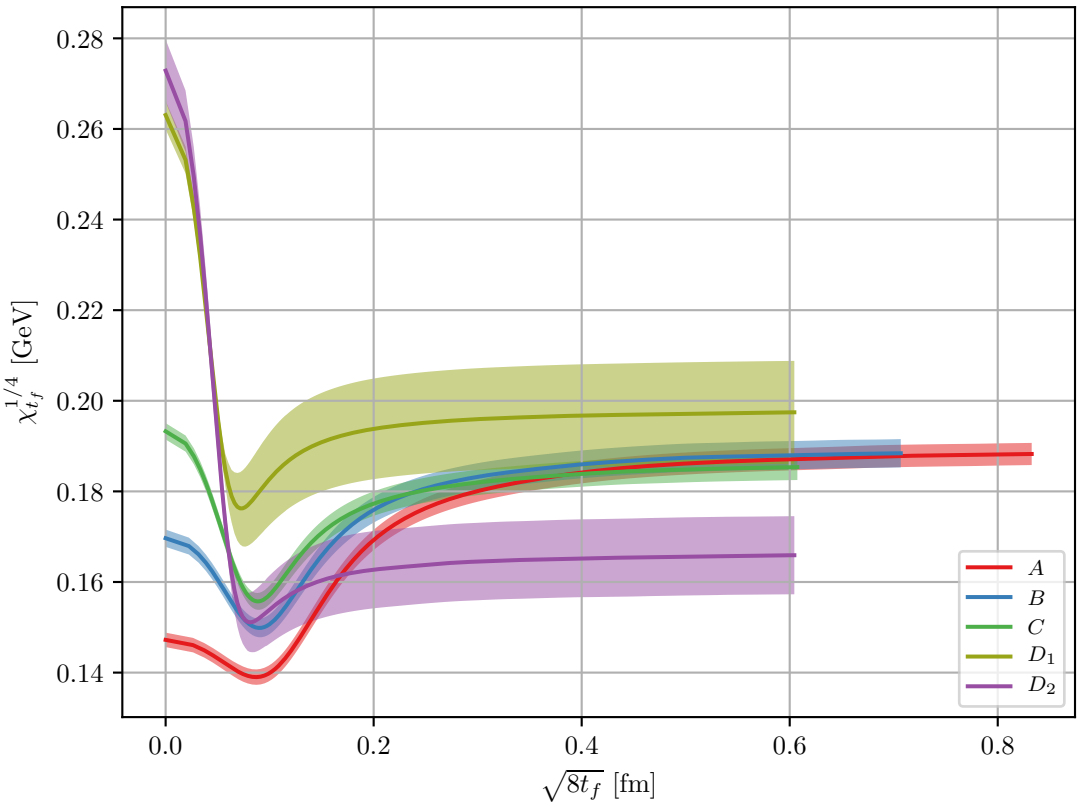


Figure 6.16: The topological susceptibility $\chi_{t_f}^{1/4}$ from eq. (3.94) on page 43. Bootstrapped $N_{\text{bs}} = 500$ times, and corrected for autocorrelations with $\sigma = \sqrt{2\tau_{\text{int}}}\sigma_0$.

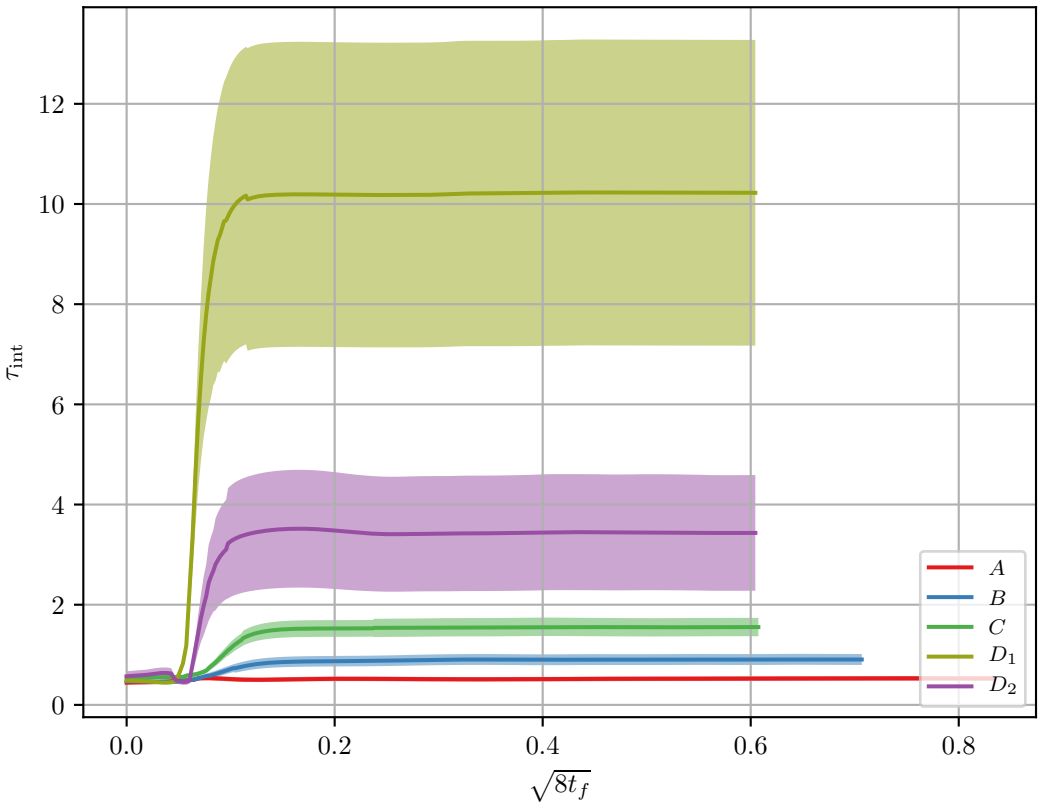


Figure 6.17: τ_{int} for the topological susceptibility $\chi_{t_f}^{1/4}$ seen in fig. 6.16 on the preceding page.

lattice spacings a as they approach the continuum. While topological charge itself do not contain any divergences(but do need renormalization at zeroth flow time), it will become divergence once we have two charges at the same point.

Table 6.7: The topological susceptibility for the ensembles seen in table 6.1 on page 89 together with the correction factor from the integrated autocorrelation time seen in eq. (D.36) on page 155. The second column have not had its results corrected by $\sqrt{2\tau_{\text{int}}}$. None of the results have been analyzed with bootstrapping.

Ensemble	$\chi_{t_f}^{1/4}$ [GeV]	$\chi_{t_f}^{1/4}$ [GeV], corrected	$\sqrt{2\tau_{\text{int}}}$
A	0.1877(23)	0.1877(24)	1.028(46)
B	0.1880(21)	0.1880(29)	1.346(81)
C	0.1853(14)	0.1853(24)	1.762(104)
D_1	0.1971(22)	0.1971(101)	4.523(675)
D_2	0.1656(33)	0.1656(86)	2.624(441)

From the error estimation in the topological charge in fig. 6.10 on page 102, our intuition would tell us that the D_2 ensemble would contain far larger error bars than that of D_1. However, as seen in fig. 6.16 on page 110 this is clearly not the case for the topological susceptibility. In fact the D_1 ensemble contains larger errors than the D_2 ensemble. If we take a look at table 6.7, we see that the errors are roughly on the same order, but that the integrated autocorrelation is far higher for the D_1 ensemble. This is similar to the autocorrelation we observed for the topological charge in fig. 6.12 on page 105. Another contribution to the large error in D_1 is its small lattice size. Since we are dividing by the volume V in the definition of the topological susceptibility in eq. (3.94) on page 43, a smaller volume V will lead to a greater error.

6.4.1 Continuum extrapolation

When extrapolating $\chi_{t_f}^{1/4}$ to the continuum, we use the same procedure as discussed for extracting the reference scales t_0 and w_0, except we do not need to find the inverse point. We perform linear fits on the bootstrap samples around the extraction point and extract the exact $\chi_{t_f}^{1/4}$ at $t_{f,\text{extrap}}$ for each of the fits. We then use the mean and standard deviation from the bootstrap samples in the continuum extrapolation. Performing a continuum extrapolation for the ensembles seen in table 6.1 on page 89 while excluding D_1, gives us

$$\chi_{t_f}^{1/4}\left(\langle Q^2\rangle\right) = 0.179(10)\text{ GeV}. \tag{6.7}$$

This continuum extrapolation can be seen in fig. 6.18 on the next page. If we instead perform a continuum extrapolation excluding the D_2 ensemble, we get

$$\chi_{t_f}^{1/4}\left(\langle Q^2\rangle\right) = 0.186(6)\text{ GeV}, \tag{6.8}$$

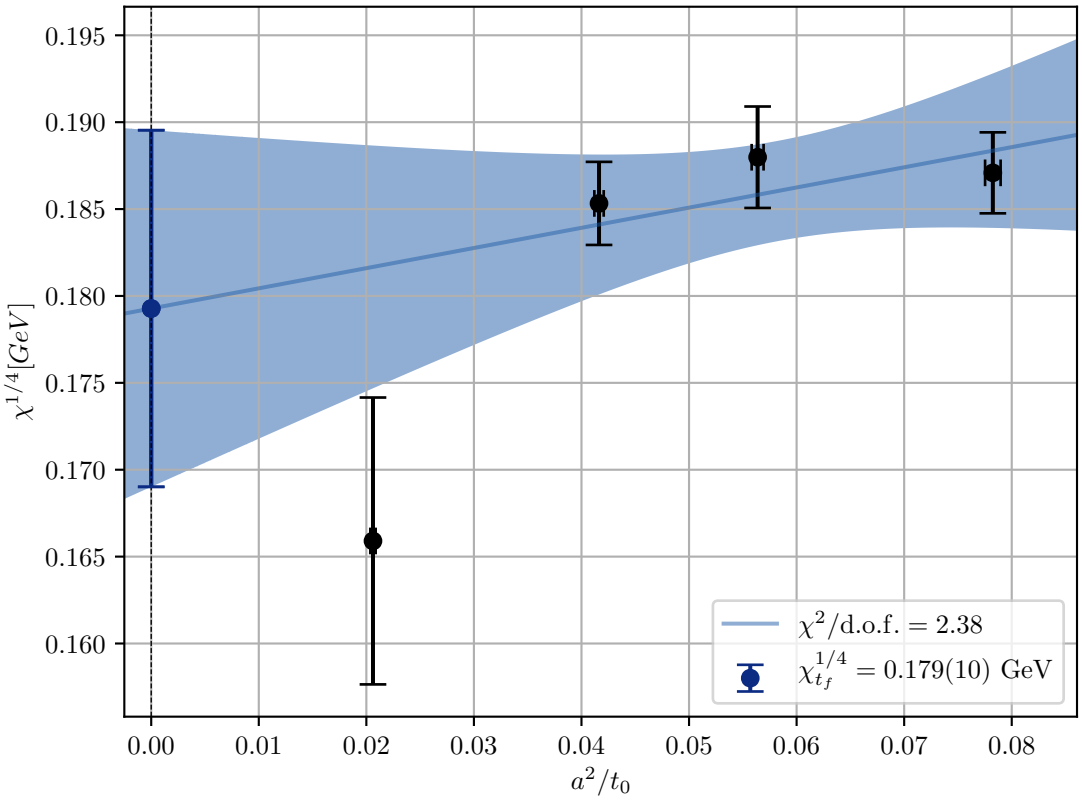


Figure 6.18: A continuum extrapolation of the topological susceptibility $\chi_{t_f}^{1/4}$ for the ensembles in table 6.1 on page 89 excluding the D_1 ensemble. The points for $\chi_{t_f}^{1/4}$ is taken at $\sqrt{8t_{f,0}} = 0.6$ fm. The exact value of the continuum extrapolation is seen in eq. (6.7) on the preceding page.

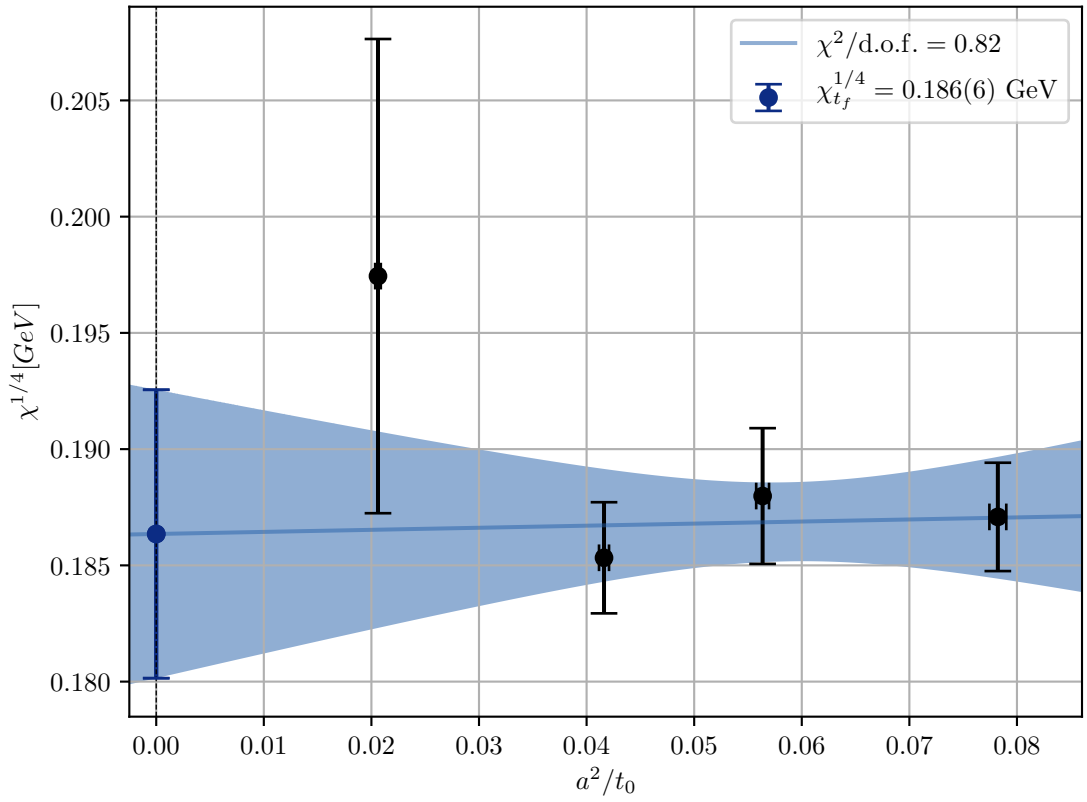


Figure 6.19: A continuum extrapolation of the topological susceptibility $\chi_{t_f}^{1/4}$ for the ensembles in table 6.1 on page 89 excluding the D_2 ensemble. The points for $\chi_{t_f}^{1/4}$ is taken at $\sqrt{8t_{f,0}} = 0.6$ fm. The exact value of the continuum extrapolation is seen in eq. (6.8) on page 112.

where the extrapolation can be seen in fig. 6.19 on the facing page. The point for extrapolation were chosen such that we have few discretization effects and the topological susceptibility have plateaued. Incidentally, the extrapolation point chosen is at a hadronic scale of flow time $\sqrt{8t_f} = 0.6$ fm.

Another approach is to extrapolate from the reference time scale, t_0 or w_0. This can either be the continuum scale of t_0 and w_0, or the individual reference scales. Results using different extrapolation points can with the ensembles A, B, C and D_1 can be seen in table 6.8 on page 125, and in table 6.9 on page 126, we see the continuum extrapolations with for the ensembles A, B, C and D_2.

Using a time series bootstrap as described in appendix D.4 on page 156, we achieved results similar to those of regular bootstrap that was corrected by $\sqrt{2\tau_{\text{int}}}$. These results can be seen in table 6.10 on page 127, table 6.11 on page 128.

The reason for investigating an alternate ensemble D_1 of $\beta = 6.45$ is now clear. If we inspect the D_2 ensemble in fig. 6.16 on page 110, it is easy to suspect there to be something incorrect in the calculation of the D_2 ensemble. However, considering the D_1 ensemble in fig. 6.16 on page 110, we see that the reason for this off-shoot by the D_2 ensemble is most likely just due to low statistics, as the D_1 ensemble carries none of the same idiosyncrasies.

An added bonus of using the A, B, C, and D_1 ensembles for the continuum extrapolation is that the fit is better with $\chi^2/\text{d.o.f.} = 0.82$ versus $\chi^2/\text{d.o.f.} = 2.38$ for A, B, C, and D_2.

6.4.2 Topological susceptibility with one charge fixed

To investigate the independence of the susceptibility in regards to the charge and flow time different from zero, we take a look at the topological susceptibility with one charge frozen at a selected flow time. We write this as

$$\chi_{t_f}^{1/4} \left(\left\langle Q_{t_f} Q_{t_{f,0}} \right\rangle \right), \tag{6.9}$$

where $Q_{t_{f,0}}$ indicates that the charge is set to flow time $t_{f,0}$. The result of freezing a charge at $t_{f,0}/a^2 \in [0.0, 0.2, 0.4, 0.6]$ can be seen in fig. 6.20 on the following page. We see that with the exception of zero flow time, we have that the topological susceptibility appears to be independent of the flow time. As we mentioned earlier, the topological charge becomes divergent once two charges is located at the same Euclidean time at zero flow time $t_f = 0$. The differing topological susceptibility with for $Q_{t_{f,0}/a^2=0.0}$ can thus be understood due to the fact that one charge is taken at $t_f = 0$, giving us a UV divergence. In order for the plot in the upper left corner of fig. 6.20 on the next page to not be divergent at $\sqrt{8t_f} = 0$, we would have to renormalize the charge.

6.4.3 Extracting the number of flavors

Using the Witten-Veneziano relation from eq. (2.59) on page 23, we can try to extract the number of flavors N_f by using the experimental values for the mass of η' and the

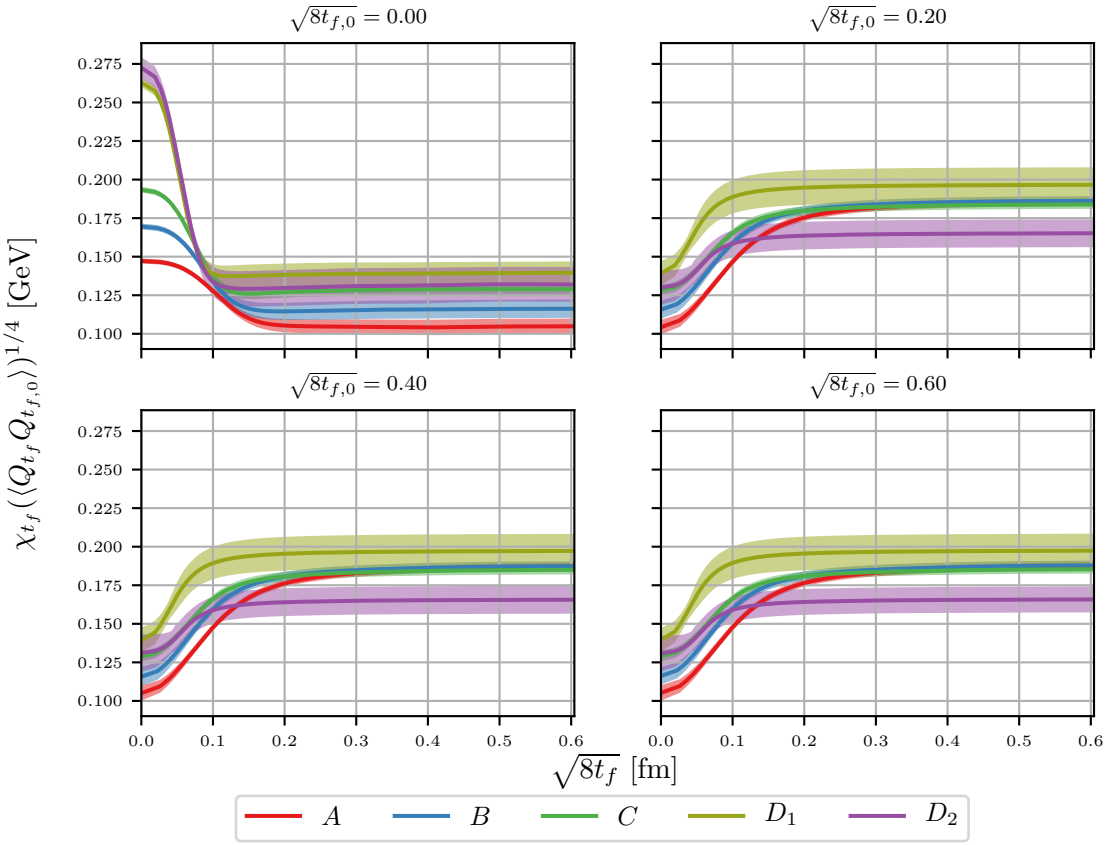


Figure 6.20: The topological susceptibility as given by eq. (6.9) on the previous page with one charge set at a flow time $\sqrt{8t_{f,0}} \in [0.0, 0.2, 0.4, 0.6]$. The figure in the upper left corner with one charge set at $\sqrt{8t_f} = 0.0$ will need to have its charge renormalized in order give the same result as the others.

π^0 decay constant. For this analysis, we use the values of $F_\pi = 0.130(5)/\sqrt{2}$ GeV and $m_{\eta'} = 0.95778(6)$ GeV [68].

From the continuum extrapolation seen in fig. 6.18 on page 113 using the ensembles A, B, C and D_2 of the topological susceptibility seen in eq. (6.7) on page 112, we obtain that the number of flavors to be

$$N_f = 3.75(29). \tag{6.10}$$

Equivalently, for the continuum extrapolation seen in fig. 6.19 on page 114 for the ensembles A, B, C and D_1 we get

$$N_f = 3.21(25). \tag{6.11}$$

which is within one standard deviation of the expected value of $N_F = 3$. Considering the Witten-Veneziano relation bridges two different theories, this is quite remarkable.

6.5 The fourth cumulant

For investigating the fourth cumulant, we will compare with the results presented in Cè et al. [13]. The ratio R as given by eq. (3.96) on page 43 is plotted for all our ensembles in table 6.1 on page 89 in fig. 6.21 on the following page. As mentioned in chapter 3 on page 25, an $R \not\approx 1$ implies that QCD is more complicated than what is predicted by the dilute instanton gas model. It is, however, challenging to discern anything of meaning, as the R-values for each ensemble is highly dissimilar. Investigating table 6.12 on page 129 in conjecture with fig. 6.21 on the next page this concern is amplified. If anything, we seem to have a dependence on the number of lattice points. As with the results for the topological charge and susceptibility, this might just be an artifact of low statistics.

6.5.1 A comparison of the fourth cumulant

In order to investigate our results further, we will compare our results with those presented by Cè et al. [13]. Their ensembles can be seen in table 6.13 on page 129. We had to perform a normalization of their results as they used a different normalization than what we used, which can be seen in table 6.14 on page 129.

For comparing our results with those by Cè et al. [13], we look at the ratio between our results presented in table 6.12 on page 129 divided by those with similar volume seen in table 6.14 on page 129. The resulting ratios can be seen in table 6.15 on page 130. For the $\langle Q^2 \rangle$ and $\langle Q^4 \rangle$ we have a quite good match with our similar-sized results, but this is not the case for $\langle Q^4 \rangle_C$ and R. With the notable exception of B and $\tilde{D}_2$, there is quite a disconnect between our results for $\langle Q^4 \rangle_C$ and R and those presented table 6.14 on page 129. The good agreement between $\langle Q^2 \rangle$ and $\langle Q^4 \rangle$ indicates that this disconnect appears be due to the particular sensitivity of $\langle Q^4 \rangle_C$ and R. It is likely that higher statistics would have remedied this issue.

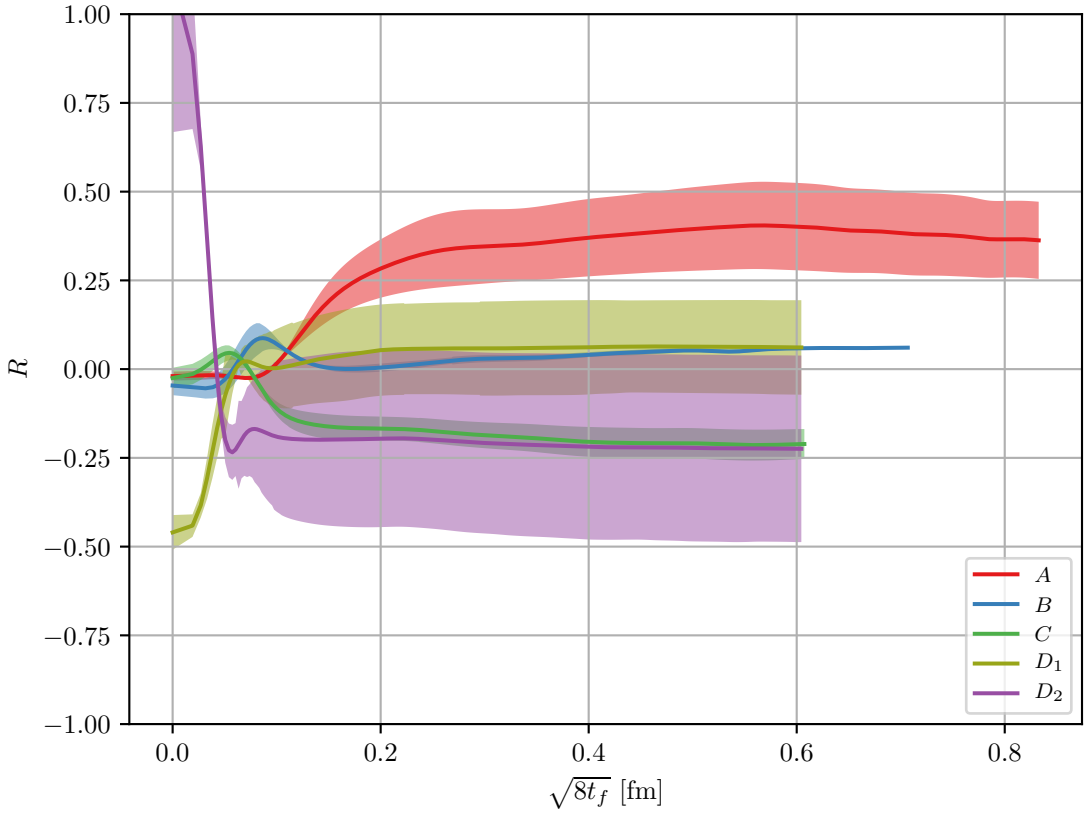


Figure 6.21: The fourth cumulant ratio $R = \langle Q^4 \rangle_C / \langle Q^2 \rangle$ for the ensembles seen in table 6.1 on page 89. The results was analyzed using $N_{\text{bs}} = 500$ bootstrap samples, with the error corrected for by $\sqrt{2\tau_{\text{int}}}$.

6.6 The topological charge correlator

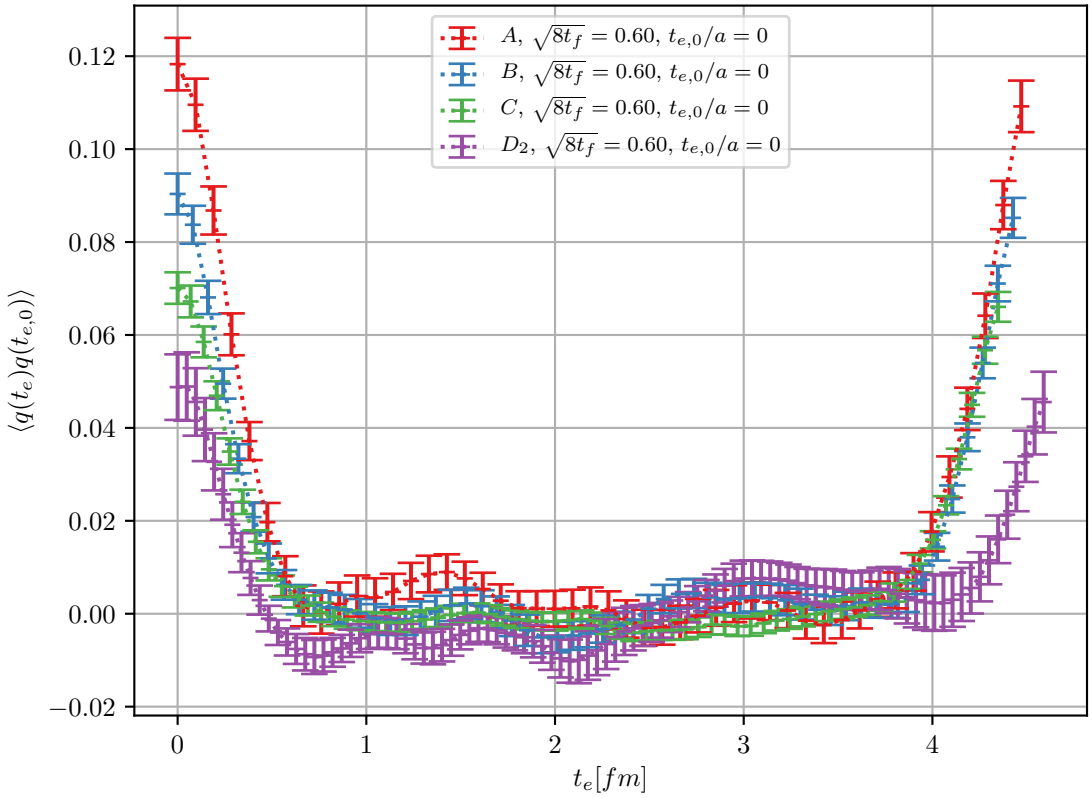


Figure 6.22: The topological charge correlator for all of the ensembles except D_1 seen in table 6.1 on page 89. The x-axis contains the sink-source separation, as the source $q(0)$ is placed at $t_e = 0$ fm, and the sink $q(t_e)$ is taken at t_e.

In fig. 6.22 we see the topological charge correlator from eq. (3.100) on page 44 with one charge at $t_e/a = 0$ for the ensembles A, B, C and D_2. In fig. 6.23 on the following page we see the topological charge correlator for the D_1 ensemble, taken in a separate plot as the Euclidean length is far shorter than the ensembles. In both cases, the correlator have been aligned around $t_e/a = 0$.

From eq. (3.98) on page 44 we expect there to be an exponential dampening in the signal, something which is hard to discern among the noise in both fig. 6.22 and in fig. 6.23 on the next page. An signal would likely be found in the range between 0.6 fm and 1 fm, given that we had higher statistics.

In fig. 6.24 on page 121 we the smearing effect of gradient flow. The source in $t_{e,0}/a$ in $C(n_t) = \langle q(t_e/a)q(t_{e,0}/a)\rangle$ is taken to be at the midpoint in Euclidean time for all of the ensembles, and we can observe how the gradient flow smears out much of the fluctuations

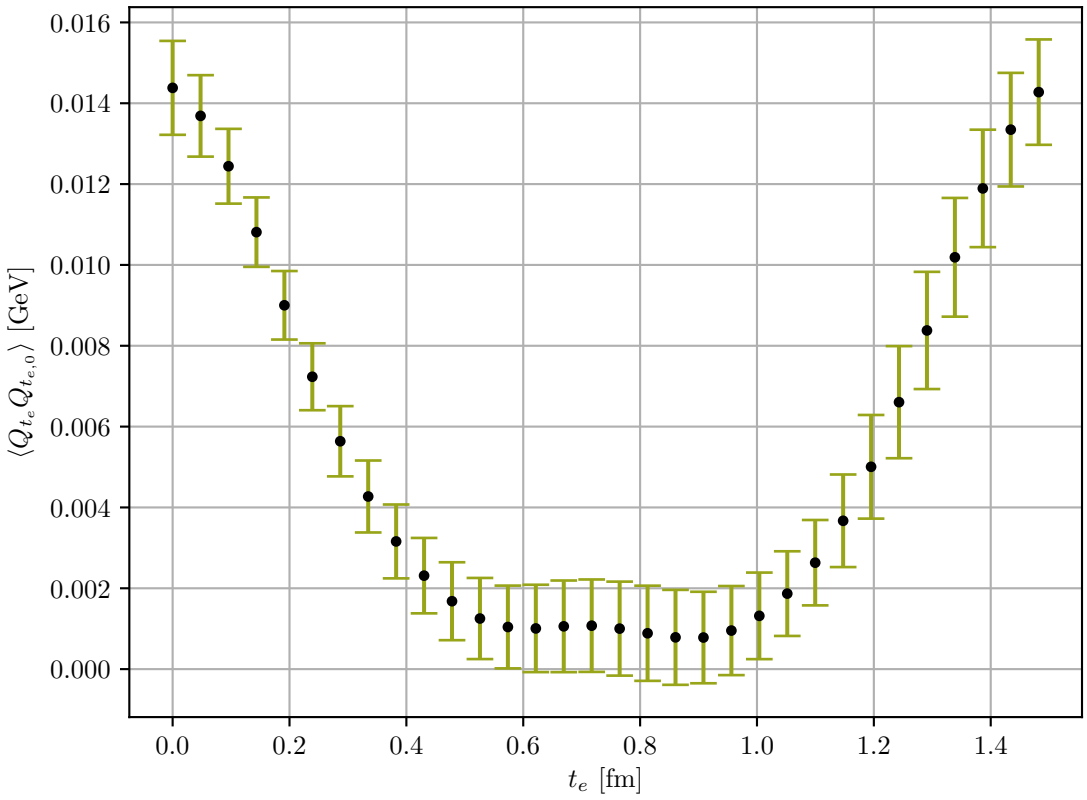


Figure 6.23: The topological charge correlator for the D_1 ensembles seen in table 6.1 on page 89. The source $q(0)$ is placed at $t_e = 0$ fm and the sink $q(t_e)$ is taken at t_e.

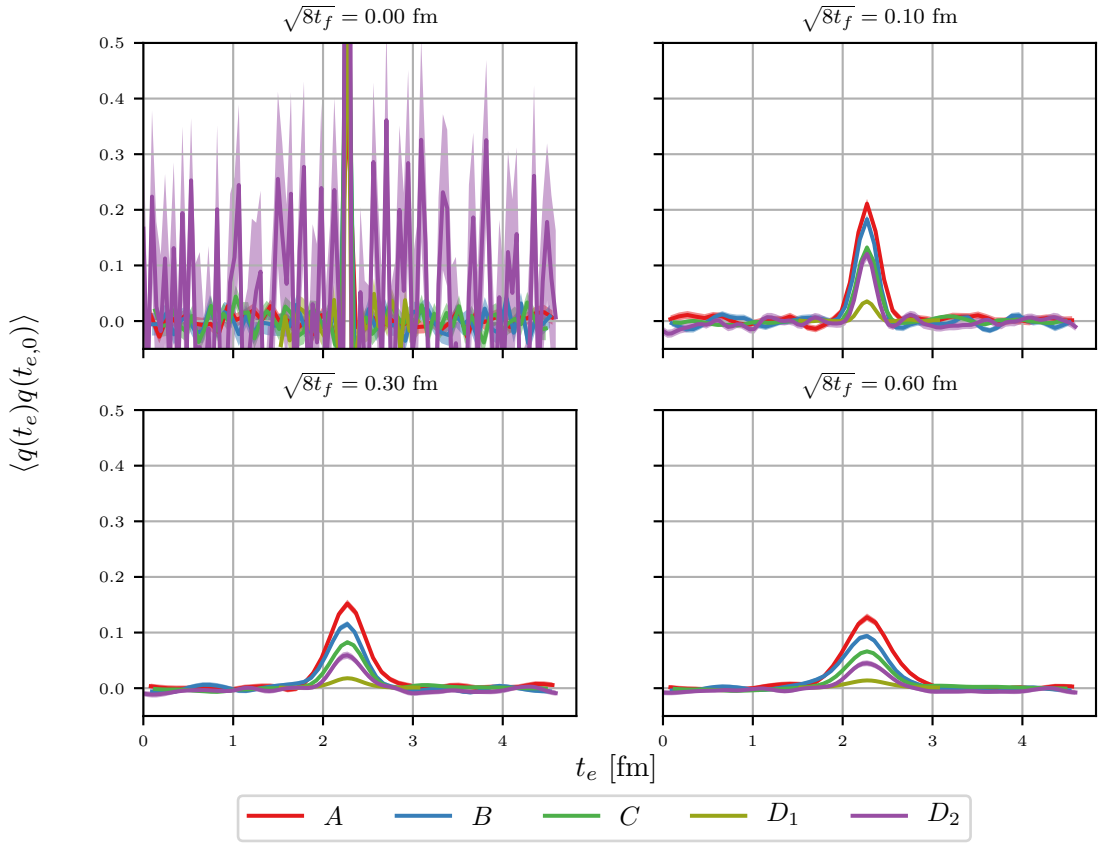


Figure 6.24: The topological charge correlator taken the center-most Euclidean time t_e for all of the ensembles seen in table 6.1 on page 89. The error on the correlator is shown using a band instead of bars in order to avoid cluttering. A danger with limit statistics is that we might lose any signal of an exponentially dampened signal which can make any extraction of states problematic.

present in the correlator. Another issue with the smearing which we can see in fig. 6.24 is that we might lose signal of an exponential dampening for any exited states from eq. (6.13).

6.6.1 The effective mass of the glueball

The next step after investigating the topological charge correlator is to try to extract excited states. In our case, this would be to try to extract the O^{-+} state(see section 2.10.5 on page 23). We start by folding the correlator,

$$\begin{aligned}\big\{C(0), C(1), &C(2), \ldots, C(N_T-1), C(N_T)\big\} \\ &\rightarrow \quad \Big\{C(0), \frac{1}{2}\left(C(1)+C(N_T-1)\right), \frac{1}{2}\left(C(2)+C(N_T-2)\right), \ldots, \\ &\qquad \frac{1}{2}\left(C(N_T/2-1)+C(N_T/2)\right)\Big\},\end{aligned} \tag{6.12}$$

such that we get twice the statistics. We now compute the effective mass as given by eq. (3.99) on page 44, which can be seen in fig. 6.25 on the facing page. We see that for all of the available ensembles, no plateau is reached.

A comparison can still be made by between our results and those seen in table 6.16 on page 130. In fig. 6.25 on the facing page of the glueball mass we made an overlay of the mass values from the literature listed in table 6.16 on page 130.

Even though we did not reach and plateau, we can still attempt to extract a mass in the region $t_e \in [0.3, 0.6]$. Extrapolating to the continuum, the mass is found to be

$$m_{\text{eff}} = 1.284(046)(215)\ \text{GeV}, \tag{6.13}$$

where the systematic uncertainty in the last parenthesis was determined by performing several plateau fits in the region $t_e \in [0.1, 0.8]$ for different interval sizes, and then taking the width of that distribution to be our systematic error.

The effective mass from eq. (6.13) is expected to be grossly incorrect, as no plateau is reached. This is conformed from comparing with the values by other articles seen in table 6.16 on page 130. The region we would like to explore from 0.6 to 1.0 fm never experiences any plateau as most of the correlators disappear after around 0.4 fm due to negative values in the logarithm in eq. (6.13). At short Euclidean times we also appear to have contamination from other states which further complicates the extraction process.

Due to gradient flow, we might smear our lattice to such a degree that the signal is lost. To investigate this, we have looked at the correlator for four different flow times which can be seen in fig. 6.26 on page 124. From the effective mass plots in the four figures, we see that none of them exhibits any proper plateauing and we still lack the statistics for a glueball mass signal.

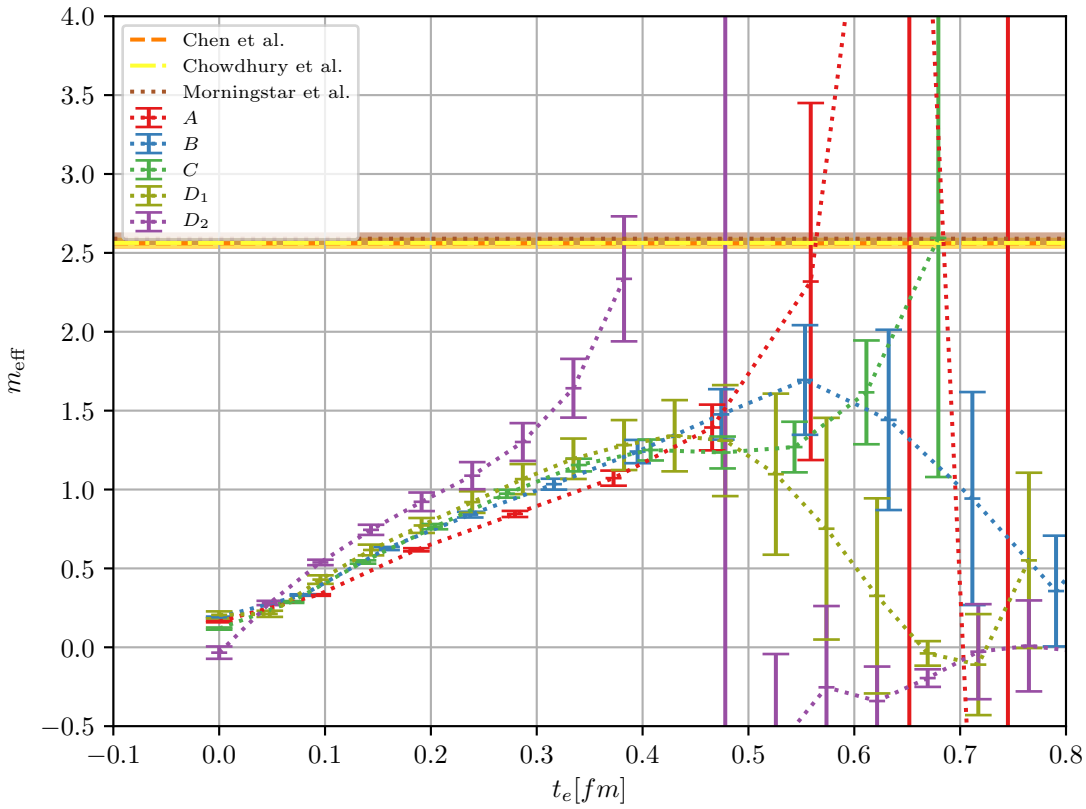


Figure 6.25: The effective mass of the glueball, as extracted from the topological charge correlator in Euclidean time.

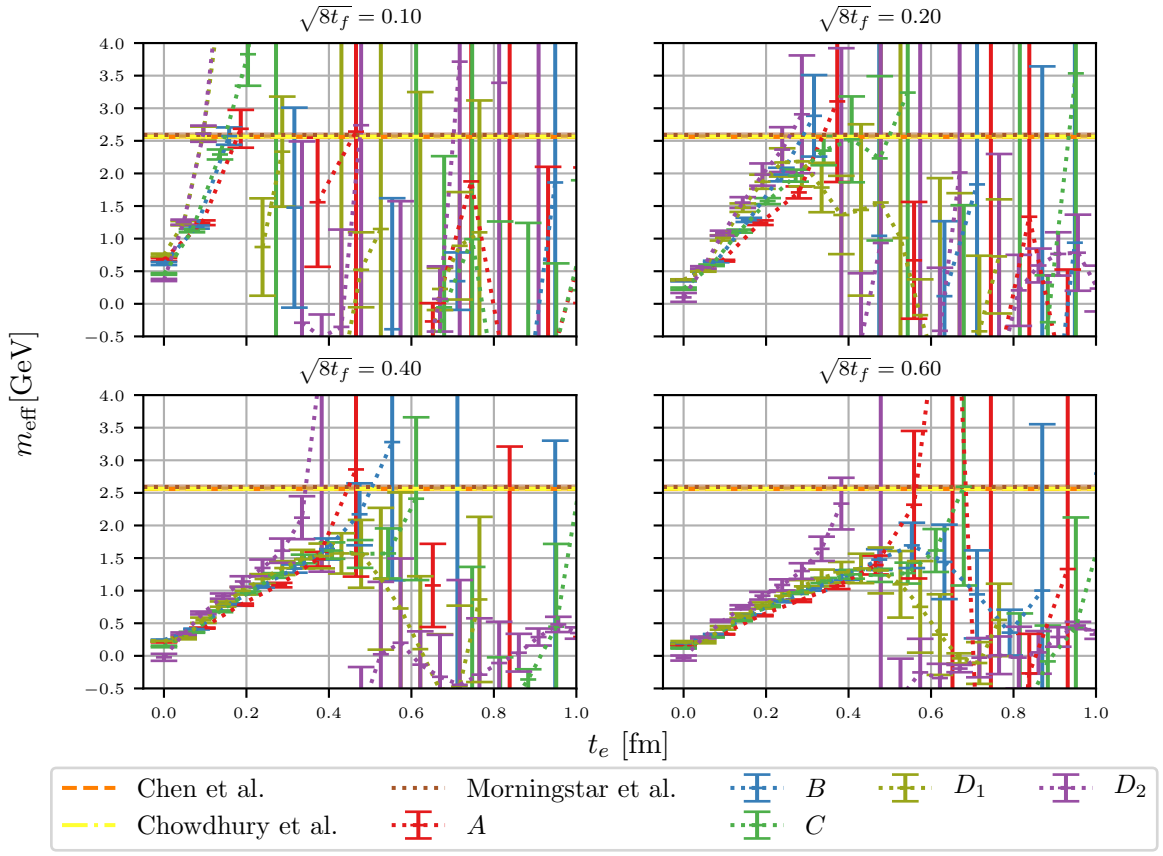


Figure 6.26: The effective mass of the glueball seen at four different flow times, $\sqrt{8t_{f,0}} \in [0.1, 0.2, 0.3, 0.4, 0.6]$.

Table 6.8: The topological susceptibility χ_{t_f} as extrapolated from different $t_{f,\text{extrap.}}$, together with the corresponding N_f. The results are extrapolated using the ensembles seen in the first column. The second column specifies what definition of the topological susceptibility we used, the third column specifies the extrapolation flow time and the fourth column specifies the fixed flow time as seen in section 6.4.2 on page 115. The fifth, sixth and seventh columns contains goodness of the fit, the continuum extrapolation of the topological susceptibility and lastly the number of flavors. The data were analyzed with bootstrapping that was corrected by the integrated autocorrelation, $\sqrt{2\tau_{\text{int}}}$.

Ensemble	$\chi_{t_f}^{1/4}$-variant	$\sqrt{8t_{f,0,\text{extrap}}}$ [fm]	Fixed flow time $t_{f,0}$ [fm]	χ^2/d.o.f.	$\chi_{t_f}^{\frac{1}{4}}$ [GeV]	N_F
A, B, C, D_1	$\chi_{t_f}^{1/4}(\langle Q^2\rangle)$	$\sqrt{8t_f} = 0.50$		0.83	0.186(6)	3.21(25)
	$\chi_{t_f}^{1/4}\langle Q_{t_f}Q_{t_{f,0}}\rangle$		0.10	0.80	0.185(6)	3.29(25)
			0.20	0.83	0.185(6)	3.28(25)
			0.30	0.89	0.185(6)	3.28(25)
			0.40	0.82	0.186(6)	3.22(25)
			0.50	0.89	0.186(6)	3.24(25)
			0.60	0.84	0.186(6)	3.22(25)
	$\chi_{t_f}^{1/4}(\langle Q^2\rangle)$	$\sqrt{8t_f} = 0.60$		0.82	0.186(6)	3.21(25)
	$\chi_{t_f}^{1/4}\langle Q_{t_f}Q_{t_{f,0}}\rangle$		0.10	0.79	0.185(6)	3.28(25)
			0.20	0.82	0.186(6)	3.27(25)
			0.30	0.89	0.185(6)	3.27(25)
			0.40	0.82	0.186(6)	3.23(25)
			0.50	0.89	0.186(6)	3.25(25)
			0.60	0.84	0.186(6)	3.23(25)
B, C, D_1	$\chi_{t_f}^{1/4}(\langle Q^2\rangle)$	$\sqrt{8t_f} = 0.50$		1.66	0.186(24)	3.21(25)
		$\sqrt{8t_f} = 0.60$		1.63	0.187(24)	3.18(24)
A, B, C	$\chi_{t_f}^{1/4}(\langle Q^2\rangle)$	$\sqrt{8t_f} = 0.50$		0.38	0.184(6)	3.37(26)
		$\sqrt{8t_f} = 0.60$		0.33	0.184(6)	3.37(26)

Table 6.9: The topological susceptibility χ_{t_f} as extrapolated from different extrapolation flow times $t_{f,\text{extrap.}}$, together with the corresponding N_f. The results are extrapolated using the ensembles seen in the first column. The description of each of the columns is given in table 6.8 on the preceding page. The data were analyzed with bootstrapping that was corrected by the integrated autocorrelation, $\sqrt{2\tau_{\text{int}}}$. Note that this tables show a different set of ensembles than those listed in table 6.8 on the previous page.

Ensemble	$\chi_{t_f}^{1/4}$-variant	$\sqrt{8t_{f,0,\text{extrap}}}$ [fm]	Fixed flow time $t_{f,0}$ [fm]	χ^2/d.o.f.	$\chi_{t_f}^{\frac{1}{4}}$ [GeV]	N_F
A, B, C, D_2	$\chi_{t_f}^{1/4}(\langle Q^2\rangle)$	$\sqrt{8t_f} = 0.60$		2.38	0.179(10)	3.75(29)
	$\chi_{t_f}^{1/4}\langle Q_{t_f}Q_{t_{f,0}}\rangle$		0.10	2.15	0.179(10)	3.78(29)
			0.20	2.05	0.179(10)	3.79(29)
			0.30	2.29	0.179(10)	3.79(29)
			0.40	2.04	0.179(10)	3.79(29)
			0.50	2.12	0.179(10)	3.77(29)
			0.60	2.39	0.179(10)	3.79(29)
	$\chi_{t_f}^{1/4}(\langle Q^2\rangle)$	$\sqrt{8t_{0,\text{cont}}} = 0.4709$		2.48	0.181(11)	3.64(28)
	$\chi_{t_f}^{1/4}\langle Q_{t_f}Q_{t_{f,0}}\rangle$		0.10	2.27	0.180(10)	3.72(29)
			0.20	2.19	0.179(10)	3.74(29)
			0.30	2.43	0.179(10)	3.74(29)
			0.40	2.15	0.180(10)	3.73(29)
			0.50	2.22	0.180(10)	3.72(29)
			0.60	2.47	0.179(10)	3.74(29)
	$\chi_{t_f}^{1/4}(\langle Q^2\rangle)$	$\sqrt{8w_{0,\text{cont}}^2} = 0.4795$		2.42	0.179(10)	3.74(29)
	$\chi_{t_f}^{1/4}\langle Q_{t_f}Q_{t_{f,0}}\rangle$		0.10	2.15	0.179(10)	3.79(29)
			0.20	2.06	0.179(10)	3.80(29)
			0.30	2.31	0.179(10)	3.79(29)
			0.40	2.08	0.179(10)	3.77(29)
			0.50	2.15	0.179(10)	3.75(29)
			0.60	2.42	0.179(10)	3.78(29)
B, C, D_2	$\chi_{t_f}^{1/4}(\langle Q^2\rangle)$	$\sqrt{8t_f} = 0.50$		2.02	0.166(24)	5.08(39)
		$\sqrt{8t_f} = 0.60$		2.05	0.166(24)	5.06(39)

Table 6.10: The topological susceptibility χ_{t_f} as extrapolated from different $t_{f,\text{extrap.}}$, together with the corresponding N_f. The results are extrapolated using the ensembles seen in the first column. The description of each of the columns is given in table 6.8 on page 125. The data were analyzed with a time series bootstrap, thus handling the autocorrelation automatically as opposed to the results in and table 6.8 on page 125 and table 6.9 on the preceding page

Ensemble	$\chi_{t_f}^{1/4}$-variant	$\sqrt{8t_{f,0,\text{extrap}}}$ [fm]	Fixed flow time $t_{f,0}$ [fm]	χ^2/d.o.f.	$\chi_{t_f}^{\frac{1}{4}}$ [GeV]	N_F
A, B, C, D_1	$\chi_{t_f}^{1/4}(\langle Q^2\rangle)$	$\sqrt{8t_f} = 0.50$		0.17	0.185(4)	3.29(25)
	$\chi_{t_f}^{1/4}\langle Q_{t_f}Q_{t_{f,0}}\rangle$		0.10	0.16	0.185(4)	3.33(26)
			0.20	0.15	0.185(4)	3.30(25)
			0.30	0.15	0.185(4)	3.32(26)
			0.40	0.15	0.185(4)	3.30(25)
			0.50	0.17	0.185(4)	3.28(25)
			0.60	0.17	0.185(4)	3.32(26)
	$\chi_{t_f}^{1/4}(\langle Q^2\rangle)$	$\sqrt{8t_f} = 0.60$		0.15	0.185(4)	3.30(25)
	$\chi_{t_f}^{1/4}\langle Q_{t_f}Q_{t_{f,0}}\rangle$		0.10	0.15	0.185(4)	3.33(26)
			0.20	0.15	0.185(4)	3.30(25)
			0.30	0.14	0.185(4)	3.32(26)
			0.40	0.14	0.185(4)	3.31(25)
			0.50	0.17	0.185(4)	3.30(25)
			0.60	0.17	0.185(4)	3.34(26)
B, C, D_1	$\chi_{t_f}^{1/4}(\langle Q^2\rangle)$	$\sqrt{8t_f} = 0.50$		0.18	0.179(13)	3.75(29)
		$\sqrt{8t_f} = 0.60$		0.18	0.180(13)	3.72(29)
A, B, C	$\chi_{t_f}^{1/4}(\langle Q^2\rangle)$	$\sqrt{8t_f} = 0.50$		0.22	0.185(7)	3.32(26)
		$\sqrt{8t_f} = 0.60$		0.20	0.185(6)	3.33(26)

Table 6.11: The topological susceptibility χ_{t_f} as extrapolated from different extrapolation flow times $t_{f,\text{extrap.}}$, together with the corresponding N_f. The ensembles used in the extrapolation is seen in the first column. The description of the rest of the columns is seen in table 6.8 on page 125. The data were analyzed with a time series bootstrap.

Ensemble	$\chi_{t_f}^{1/4}$-variant	$\sqrt{8t_{f,0,\text{extrap}}}$ [fm]	Fixed flow time $t_{f,0}$ [fm]	χ^2/d.o.f.	$\chi_{t_f}^{\frac{1}{4}}$ [GeV]	N_F
A, B, C, D_2	$\chi_{t_f}^{1/4}(\langle Q^2\rangle)$	$\sqrt{8t_f} = 0.60$		0.48	0.183(7)	3.46(27)
	$\chi_{t_f}^{1/4}\langle Q_{t_f} Q_{t_{f,0}}\rangle$		0.10	0.44	0.183(7)	3.49(27)
			0.20	0.45	0.183(7)	3.48(27)
			0.30	0.47	0.182(7)	3.50(27)
			0.40	0.44	0.183(7)	3.48(27)
			0.50	0.45	0.183(7)	3.47(27)
			0.60	0.44	0.183(7)	3.49(27)
	$\chi_{t_f}^{1/4}(\langle Q^2\rangle)$	$\sqrt{8t_{0,\text{cont}}} = 0.4708$		0.53	0.184(7)	3.36(26)
	$\chi_{t_f}^{1/4}\langle Q_{t_f} Q_{t_{f,0}}\rangle$		0.10	0.48	0.183(7)	3.42(26)
			0.20	0.49	0.184(7)	3.42(26)
			0.30	0.50	0.183(7)	3.44(26)
			0.40	0.47	0.183(7)	3.42(26)
			0.50	0.47	0.184(7)	3.41(26)
			0.60	0.46	0.183(7)	3.44(26)
	$\chi_{t_f}^{1/4}(\langle Q^2\rangle)$	$\sqrt{8w_{0,\text{cont}}^2} = 0.4794$		0.49	0.183(7)	3.44(26)
	$\chi_{t_f}^{1/4}\langle Q_{t_f} Q_{t_{f,0}}\rangle$		0.10	0.44	0.182(7)	3.49(27)
			0.20	0.46	0.183(7)	3.48(27)
			0.30	0.48	0.183(7)	3.49(27)
			0.40	0.45	0.183(7)	3.47(27)
			0.50	0.46	0.183(7)	3.45(27)
			0.60	0.45	0.183(7)	3.47(27)
B, C, D_2	$\chi_{t_f}^{1/4}(\langle Q^2\rangle)$	$\sqrt{8t_f} = 0.50$		0.38	0.172(17)	4.39(34)
		$\sqrt{8t_f} = 0.60$		0.39	0.173(18)	4.37(34)

Table 6.12: The fourth cumulant is taken at their individual reference scales seen in the third column. $\langle Q^4 \rangle_C$ is given by eq. (3.95) on page 43, and R is given by eq. (3.96). The data were analyzed with using a bootstrap analysis of $N_{bs} = 500$ samples, with error corrected by the integrated autocorrelation, $\sqrt{2\tau_{\text{int}}}$.

Ensemble	L/a	t_0/a^2	$\langle Q^2 \rangle$	$\langle Q^4 \rangle$	$\langle Q^4 \rangle_C$	R
A	2.24	3.20(3)	0.78(4)	2.13(27)	0.282(67)	0.359(65)
B	2.21	4.43(4)	0.81(5)	1.98(23)	0.036(11)	0.044(11)
C	2.17	6.01(6)	0.77(4)	1.6(2)	−0.174(40)	−0.226(64)
D_1	1.53	12.2(1)	1.00(20)	3.01(1.07)	0.03(12)	0.03(12)
D_2	2.29	12.2(1)	0.497(100)	0.64(20)	−0.103(95)	−0.21(23)

Table 6.13: Parameters of the ensembles presented by Cè et al. [13]. The first column is the ensemble name from the article. The letter indicates the volume, while the subindex indicates the β value. We see that the ensembles of similar letters keep approximately the same length L.

Ensemble	β	L/a	L [fm]	a [fm]	t_0/a^2	t_0/r_0^2	N_{cfg}
F_1	5.96	16	1.632	0.102	2.7887(2)	0.1113(9)	1 440 000
B_2	6.05	14	1.218	0.087	3.7960(12)	0.1114(9)	144 000
$\tilde{D}_2$		17	1.479		3.7825(8)	0.1110(9)	
B_3	6.13	16	1.232	0.077	4.8855(15)	0.1113(10)	144 000
$\tilde{D}_3$		19	1.463		4.8722(11)	0.1110(10)	
B_4	6.21	18	1.224	0.068	6.2191(20)	0.1115(11)	144 000
$\tilde{D}_4$		21	1.428		6.1957(14)	0.1111(11)	

Table 6.14: Results as presented by Cè et al. [13], normalized by the lattice volume.

Ensemble	$\langle Q^2 \rangle_{\text{normed}}$	$\langle Q^4 \rangle_{\text{normed}}$	$\langle Q^4 \rangle_{C,\text{normed}}$	R_{normed}
F_1	0.728(1)	1.608(4)	0.016(1)	0.022(1)
B_2	0.772(3)	1.873(19)	0.085(4)	0.110(5)
$\tilde{D}_2$	0.770(3)	1.817(17)	0.037(4)	0.048(5)
B_3	0.760(3)	1.805(17)	0.074(3)	0.097(4)
$\tilde{D}_3$	0.769(3)	1.801(14)	0.027(1)	0.035(1)
B_4	0.776(3)	1.874(18)	0.069(3)	0.089(4)
$\tilde{D}_4$	0.785(3)	1.891(17)	0.040(4)	0.052(5)

Table 6.15: A comparison between the results obtained in this thesis on the fourth cumulant, and by those similar in volume form Cè et al. [13] seen in table 6.14 on the preceding page. *Ratio* indicates that we are dividing our results by the ones of in table 6.14 on the previous page.

Article	Thesis	Ratio($\langle Q^2 \rangle$)	Ratio($\langle Q^4 \rangle$)	Ratio($\langle Q^4 \rangle_C$)	Ratio(R)
F_1	A	1.08(6)	1.34(18)	19.03(5.81)	17.64(4.48)
B_2	A	1.02(5)	1.15(15)	3.60(1.09)	3.54(90)
	B	1.04(6)	1.06(11)	0.480(74)	0.46(4)
$\tilde{D}_2$	A	1.02(5)	1.19(15)	8.31(1.99)	8.15(1.56)
	B	1.05(6)	1.10(12)	1.1(1)	1.06(3)
B_3	B	1.06(6)	1.10(12)	0.550(86)	0.52(5)
$\tilde{D}_3$	B	1.05(6)	1.11(12)	1.51(23)	1.4(1)
B_4	C	0.99(5)	0.86(8)	−2.32(46)	−2.35(59)
$\tilde{D}_4$	C	0.98(5)	0.85(8)	−3.95(96)	−4.05(1.19)

Table 6.16: The glueball mass from the three listed papers, which will be used when comparing our results to that of existing literature.

Article	m_{eff} [GeV]
Chen et al. [10]	2.560(35)(120)
Chowdhury et al. [11]	2.563(34)
Morningstar and Peardon [47]	2.590(40)(130)

Chapter 7

Conclusion and final remarks

The main focus of this thesis has been on developing a code capable of generating and flowing gauge configurations. Several programming challenges have had to be overcome, particularly those related to ensuring that the code remains scalable for large systems.

In the scalability analysis of the weak and strong scaling seen in section 5.1.3 on page 72, we found that for flow and input/output[1] we seemed to reach a plateau in performance around $N_p = 512$ processors. This was however not the case for the configuration generation, as that did not appear to plateau in performance. An increased number of processors when generating should thus be feasible.

Subsequent analysis related to the configuration generation and minimization of autocorrelation were performed, where different values for single link updates N_{up} and correlation lengths N_{corr} were investigated. We found that for the integrated autocorrelation in fig. 5.12 on page 84 and for the computational time in fig. 5.13 on page 85 we still have a lot of room for $N_{\text{up}} > 30$ single link updates, as that has a far lesser computational cost and a larger impact on the autocorrelation compared to increased correlation lengths N_{corr}.

GLAC has been extensively unit tested and integration tested as well as validation tested with results from Chroma [20], indicating we have a code which produces the expected output.

While the goal of this thesis has been to create a code for generating and flowing gauge configurations, we utilized the code to investigate a selection of gauge observables. Among the gauge observables we investigated, we focused mostly on the topological charge and related quantities.

We began by extracting two reference scales. The first being the t_0-scale which we extracted from the energy in fig. 6.1 on page 91 and extrapolated to the continuum in fig. 6.2 on page 92. Using ensembles A, B, C and D_2 we got the continuum value $t_{0,\text{cont}}/r_0^2 = 0.11087(50)$, which agrees well with that found in [40, fig. 3, p. 12].

The same goes for w_0, where we got $w_{0,\text{cont}} = 0.1695(5)$ fm which is comparable to $w_{0,\text{cont}} = 0.1755(18)(04)$ fm by Borsanyi et al. [5] when considering they included dynamical fermions in their configurations.

[1] We remind the reader that improvements to the IO were found late in the process of writing the thesis which, lead to a factor ten in performance.

When investigating the topological charge seen in fig. 6.10 on page 102 we saw that for some of the ensembles, the averaged topological charge were not centered at zero, indicating an uneven sampling of the configuration space. The issue of low statistics is something which we ran into again for both the topological susceptibility in fig. 6.16 on page 110 and the fourth cumulant ratio R in fig. 6.21 on page 118. This was particularly apparent for the D_2 ensemble. Using the ensembles A, B, C, and D_2 we retrieved a continuum value of $\chi_{t_f}^{1/4} = 0.179(10)$ GeV, which through the Witten-Veneziano formula in eq. (2.59) on page 23 corresponds to $N_f = 3.75(29)$.

Investigating the integrated autocorrelation of the topological charge in fig. 6.12 on page 105, we see that the topological charge for the D_2 ensemble has a $\tau_{\text{int}} = 11.0(4.7)$ for the final flow time, telling us that we only have around 22 independent gauge configurations of the 250 generated. In order to illuminate the details surrounding our sampling, we generated the ensemble D_1. While the D_1 ensemble is more autocorrelated with $\tau_{\text{int}} = 21.0(7.7)$ thus only giving us around 48 independent gauge configurations of 1000 generated, it helped us to determine that our results are most likely due to a short Markov chain.

Relying on the results seen in [13], which shows that the topological susceptibility becomes independent of volume in lattices larger than $\sim 1.2-1.4$ fermi, we found that extrapolating to the continuum with the ensembles A, B, C and D_1 yielded $\chi_{t_f}^{1/4} = 0.186(6)$ GeV and a corresponding $N_f = 3.21(25)$. This corresponds very well with the value of $N_f = 3$ that we would expect from the η'-meson.

An issue remains regarding the autocorrelation of D_1 and D_2. Since we used the same N_{up} and N_{corr} parameters and β-values, we would expect the autocorrelation to be similar. The reason for them not being similar may be due to having a limited number of gauge configurations. If this is the case, we might be underestimating the error and may indeed only have around ten actually independent gauge configurations in the D_2 ensemble, assuming a similar integrated autocorrelation τ_{int} as that to D_1. Accounting for this we might expect the estimated $\chi_{t_f}^{1/4}$ for the ensembles A, B, C and D_2 to be much closer to that of ensembles A, B, C and D_1.

Regarding the fourth cumulant, we were able to achieve correspondence between our results in table 6.12 on page 129 and those presented by Cè et al. [13] in table 6.14 on page 129, but only for $\langle Q^2 \rangle$ and $\langle Q^4 \rangle$. Under closer inspection of $\langle Q^4 \rangle_C$ and R it appears that both are highly susceptible to low statistics and small variations. Thus, the same level of agreement were not achieved.

Finally, we looked at the correlator $C(n_t) = \langle q(n_t)q(0) \rangle$, and attempted to extract the glueball mass for the 0^{-+} state. Extracted at flow time $\sqrt{8t_f} = 0.6$ fm, the resulting mass of $m_{\text{eff}} = 1.284(046)(215)$ GeV differs significantly from what other papers report in table 6.16 on page 130. There are several possible explanations available as to why this is the case. The most obvious one is the available statistics. We see that at large euclidean times in fig. 6.25 on page 123 the signal is dominated by noise, and we never experience any true plateau. Take for instance the D_2 ensemble, where after 0.4 fm, we no longer have any signal(i.e. we have negative values in the log). A similar problem exist at short euclidean times, where the signal is contaminated by exited states. It is likely, that higher statistics for all of the ensembles would have helped separate the different exited states

and made the plateauing more prominent.

Another possibility which works in conjunction with the available statistics, is that $\sqrt{8t_f} = 0.6$ fm may be extracting the effective mass at a too large flow time. To investigate this we retrieved the effective mass at four different flow times $\sqrt{8t_{f,0}} \in [0.1, 0.2, 0.3, 0.4, 0.6]$ fm in fig. 6.26 on page 124. In particular, we see that for $\sqrt{8t_f} = 0.2$ fm we have something which resembles a plateau and that is in the vicinity of the reported effective masses in the literature table 6.16. With greater statistics, this is something that would have been interesting to investigate further.

In the ensembles for the smallest lattice spacing a, our results were affected by critical slowdown as they exhibited strong autocorrelations(see fig. 6.12 on page 105). This can be seen directly in the Monte Carlo history in fig. 6.14 on page 107, where we for increasing β(or decreasing lattice spacing a) see that the ensemble becomes far less random and appear to change more slowly.

The issue of critical slow down is something which is inherent to all lattice calculation. The simplest fix is to increase the statistics or try to change the number of N_{corr} of N_{up} updates. One could also test other types of algorithms and try to improve the results in that way, although critical slowdown is a problem in all algorithms related to configuration generation. The Metropolis algorithm that we used here, while simple in implementation, offers us a profound insight into the mechanics of a lattice QCD calculation and serves as solid stepping stone to more advanced methods. This is particularly true for implementing fermions, as many of the tools needed have been developed in this thesis.

Implementing the much-celebrated gradient flow has been a significant component of this thesis and deserves to be mentioned. One of the conclusions we perhaps can draw is that applying gradient flow on an ensemble of configurations can provide insight into whether or not one has sampled configuration space properly. Thus, the analysis performed in this thesis can be applied to other sets of configurations and reveal the *goodness* of an ensemble.

7.1 Future prospects and improvements

Having already covered some of the possible code improvements in section 5.7 on page 87, we will only reiterate some of the general findings in this section. If we were to quickly improve our results, the brute force way would be to increase the statistics by generating more configurations, particularly for ensemble D_1 and D_2 for reasons already discussed in the previous. From inspecting the Monte-Carlo evolution in fig. 6.14 on page 107 we see how for different blocks in Monte-Carlo time the charge is centered around wildly different values. Thus, increasing the statistics would help with the aforementioned critical slowdown.

Looking back over the presented results and in particular comparisons between the set of ensembles A, B, C, D_1 and A, B, C, D_2, it is likely that we, for the most part, are dealing with low statistics. This is reinforced by having applied gradient flow on a configuration generated by Chroma [20] and reproducing Chroma's output down to machine precision.

A better utilization of the ensembles E, F, and G should also be performed.

In generating new ensembles, we would recommend generating new configurations with more N_{up} link updates in between each sampling to achieve better statistics. As we saw in the fig. 5.12 on page 84, what had the largest impact on the autocorrelation appeared to be an increased number of single link updates. Coincidentally, this has (perhaps unsurprisingly) a smaller computational cost compared to increasing the number of correlation updates N_{corr}, as we see in fig. 5.13 on page 85.

Another option for reducing the error would be to implement higher order actions and improved field strength tensors for our gauge observables. Two candidates would be the Lüscher-Weisz action [38] and the $\mathcal{O}(a^4)$-improved field strength tensor [3]. Since all of the framework required for retrieving shifted fields and links are implemented, this would be one of the most obvious and quickest ways of improving the results, next to increasing the statistics and the number of single link updates N_{up}.

The most tantalizing path to take is perhaps to implement fermions. This would require the implementation of a Hybrid Monte Carlo(HMC) type of algorithm, as our current type of Metropolis algorithm is not suitable for fermions. Fortunately, most of the tools can be reused, but implementations of linear algebra methods and γ-matrix arithmetic are needed in order to be able to calculate fermion determinants and matrices. While it certainly would take time to fully integrate fermions in the code, the development of this code, GLAC, serves as a first step towards developing a code capable of solving full QCD numerically.

Appendix A

Conventions

A.1 Notational conventions

In this thesis, repeated indices indicate summation, i.e. Einstein's summation convention. Further, natural units are assumed,

$$\hbar = c = 1. \tag{A.1}$$

This allows us to express every unit in terms of a single unit,

$$[\text{length}] = [\text{time}] = [\text{energy}]^{-1} = [\text{mass}]^{-1}, \tag{A.2}$$

such that when we say something has unit -1 or $+2$ we mean in units of energy.

With this, we get that the rest energy of a particle mc^2 is equal to just m. A useful unit of conversion is given by

$$(1\text{ GeV})^{-1}(\hbar c) = 0.1973 \times 10^{-15}\text{ m} = 0.1973\text{ fm}, \tag{A.3}$$

which yields the useful conversion unit $1 = 197.3$ MeVfm. 1 fm is taken to mean one unit of Fermi, which is 10^{-15} meters, while 1 eV is taken to mean one electron volt, which is the energy acquired by an electron when exposed to 1 Volt.

A.2 Practicals on Wick rotation

A Wick rotation is when we are going from real time in Minkowski space, to imaginary time in Euclidean space. With τ being real time,

$$\tau \to -it, \tag{A.4}$$

we get Euclidean space where t is now imaginary time. The choice of $-i$ instead of just i ensures that the propagator e^{-tH} is bounded.

Since we $x^2 = x_0^2 - |\mathbf{x}|^2$, we get the following quantities

$$\begin{aligned} p_0^M &= ip_4, \quad (p^M)^2 = -(p_4^2 + \mathbf{p}^2) = -p^2, \quad \mathrm{d}^4 p^M = i\mathrm{d}p_4 \mathrm{d}^3 p = i\mathrm{d}^4 p, \\ x_0^M &= -ix_4, \quad (x^M)^2 = -(t^2 + \mathbf{x}^2) = -x^2, \quad \mathrm{d}^4 x^M = -i\mathrm{d}x_4 \mathrm{d}^3 x = -i\mathrm{d}^4 x, \\ \partial_0^M &= i\partial_4, \quad (\partial^M)^2 = -(\partial_t^2 + \boldsymbol{\nabla}) = -\partial^2, \\ A_0^M &= iA_4. \end{aligned} \tag{A.5}$$

A.3 Gamma matrices

The Minkowski gamma matrices γ_μ^M with $\mu = 0, 1, 2, 3$ is given as

$$\{\gamma_\mu^M, \gamma_\nu^M\} = 2g_{\mu\nu} \times 1_4. \tag{A.6}$$

1_4 simply refers to the identity matrix of size 4×4. The subscript 4 assumed and thus be dropped in further discussions. Note we label the gamma matrices by M, and that this is not the case in the presentation of QCD chapter 3 on page 25. The metric is given by $g_{\mu\nu} = g^{\mu\nu} = \mathrm{diag}(1, -1, -1, -1)$. The corresponding Euclidean matrices are constructed from multiplying the spatial gamma matrices by $-i$,

$$\gamma_1 = -i\gamma_1^M, \ \gamma_1 = -i\gamma_2^M, \ \gamma_3 = -i\gamma_3^M, \ \gamma_4 = \gamma_0^M. \tag{A.7}$$

The Euclidean commutator becomes

$$\{\gamma_\mu, \gamma_\nu\} = 2\delta_{\mu\nu} \times 1. \tag{A.8}$$

Just as $\gamma_5^M = \prod_{i=0}^3 \gamma_i^M$, the Euclidean chirality operator is given by $\gamma_5 = \prod_{i=1}^4 \gamma_i$.

There are several representations available to us, i.e. the Dirac-, chiral- or Majorana-basis. For lattice QCD, one usually uses the chiral representation in which γ_5 is diagonal [24, 31]. The γ-matrices are then given as

$$\gamma_i = \begin{bmatrix} 0 & -i\sigma_i \\ i\sigma_i & 0 \end{bmatrix}, \ \gamma_4 = \begin{bmatrix} 0 & 1_2 \\ 1_2 & 0 \end{bmatrix}, \ \gamma_5 = \begin{bmatrix} 1_2 & 0 \\ 0 & -1_2 \end{bmatrix}, \tag{A.9}$$

which when explicitly written out becomes

$$\begin{gathered} \gamma_1 = \begin{bmatrix} 0 & 0 & 0 & -i \\ 0 & 0 & -i & 0 \\ 0 & i & 0 & 0 \\ i & 0 & 0 & 0 \end{bmatrix}, \quad \gamma_2 = \begin{bmatrix} 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ -1 & 0 & 0 & 0 \end{bmatrix}, \quad \gamma_3 = \begin{bmatrix} 0 & 0 & -i & 0 \\ 0 & 0 & 0 & i \\ i & 0 & 0 & 0 \\ 0 & -i & 0 & 0 \end{bmatrix}, \\ \gamma_4 = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}, \quad \gamma_5 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix}. \end{gathered} \tag{A.10}$$

From these matrices follows several relations. First off we can note that $\gamma_5^2 = 1$ and $(\gamma_5)^\dagger = \gamma_5$. Further, it follows that γ_μ is its own inverse for $\mu = 1, 2, 3, 4, 5$,

$$\gamma_\mu = \gamma_\mu^\dagger = \gamma_\mu^{-1}. \tag{A.11}$$

We also have that a γ_μ and γ_5 anti-commute,

$$\{\gamma_\mu, \gamma_5\} = 0. \tag{A.12}$$

There are several other matrix relations and identities, but these are mostly related to taking traces and calculating cross sections, thus not as relevant to us.

Appendix B

Lie groups

A brief overview of the basic definitions and properties of the Lie groups of SU(N), the special unitary groups, is given here. We will be using the fundamental representation of the SU(N) groups.

B.1 Representation

The representation of SU(N) is given by complex $N \times N$ unitary matrices,

$$\Omega^\dagger \Omega = \Omega \Omega^\dagger = 1, \tag{B.1}$$

with the Hermitian conjugate given as the inverse $\Omega^\dagger = (\Omega^*)^T = \Omega^{-1}$, and whose determinant is one, $\det \Omega = 1$. Worth noting, is that two elements of the group do not commute, making it non-Abelian.

Given two elements Ω_1, Ω_2 obeying eq. (B.1) and their determinants is 1, we have their multiplication $\Omega_1 \cdot \Omega_2$ should also belong to SU(N). This is the case if their product satisfies unitarity,

$$(\Omega_1 \Omega_2)^\dagger \Omega_1 \Omega_2 = \Omega_2^\dagger \underbrace{\Omega_1^\dagger \Omega_1}_{1} \Omega_2 = \Omega_2^\dagger \Omega_2 = 1. \tag{B.2}$$

Further, we have that the associative and the identity element is trivial, together with their inverse element. Thus, since the elements are closed under multiplication, and the unit matrix is contained in SU(N), SU(N) forms a group.

B.1.1 The Lie algebra and the fundamental representation

The Lie algebra $\mathfrak{su}(N)$ of SU(N) consist of t^a generators for $a = 1, 2, \ldots N^2 - 1$. This allows us to represent represent $\Omega \in \mathrm{SU}(N)$ by an exponential sum in the t^a generators of SU(N). The generators t^a are traceless, complex and hermitian $N \times N$ matrices with the normalization condition of

$$\mathrm{tr}\left[t^a t^b\right] = \frac{1}{2}\delta^{ab}. \tag{B.3}$$

The structure constants f^{abc} are defined through

$$[t^a, t^b] = i f^{abc} t^c, \tag{B.4}$$

and are anti-symmetric.

An element Ω of the Lie algebra can be written as

$$\Omega = \exp\left(i \sum_{j=1}^{N^2-1} \omega^{(j)} T_j \right). \tag{B.5}$$

The ω^j are real numbers that parametrize Ω.

B.2 Generators of SU(2)

The generators of SU(2) are given by the 2×2 matrices,

$$t_a = \frac{1}{2}\sigma_a, \tag{B.6}$$

with σ_a being the Pauli matrices,

$$\sigma_1 = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \quad \sigma_2 = \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix}, \quad \sigma_3 = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}. \tag{B.7}$$

The structure constants are given by the anti-symmetric Levi-Civita tensor,

$$[\sigma_a, \sigma_b] = f_{abc} t_c = 2i\epsilon_{abc} t_c. \tag{B.8}$$

B.3 Generators of SU(3)

The generators of the SU(3) group are the 3×3 matrices,

$$t_a = \frac{1}{2}\lambda_a, \tag{B.9}$$

with λ_a being the Gell-Mann matrices,

$$\begin{gathered} \lambda_1 = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \quad \lambda_2 = \begin{bmatrix} 0 & -i & 0 \\ i & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \quad \lambda_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \\ \lambda_4 = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix}, \quad \lambda_5 = \begin{bmatrix} 0 & 0 & -i \\ 0 & 0 & 0 \\ i & 0 & 0 \end{bmatrix}, \quad \lambda_6 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}, \\ \lambda_7 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & -i \\ 0 & i & 0 \end{bmatrix}, \quad \lambda_8 = \frac{1}{\sqrt{3}} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -2 \end{bmatrix}. \end{gathered} \tag{B.10}$$

B.4 Exponentiation of a $\mathfrak{su}(3)$ matrix

In order to be able to calculate anything using gradient flow and the Runge-Kutta method of third order(see section 4.2 on page 64), we need to figure out how one exponentiates the matrices Z_i. The simplest approach would be to make a Taylor expansion and simply add enough terms, but luckily a better way of SU(3) matrix exponentiation have been found by Morningstar and Peardon [46]. Given a traceless Hermitian 3×3 matrix Q,

$$c_0 = \det Q = \frac{1}{3}\mathrm{tr}(Q^3), \tag{B.11}$$

$$c_1 = \frac{1}{2}\mathrm{tr}(Q^2), \tag{B.12}$$

with the max of c_0 given by[1]

$$c_0^{\max} = 2\left(\frac{c_1}{3}\right)^{3/2}, \tag{B.13}$$

the exponentiation is given by

$$\mathrm{e}^{iQ} = f_0 I + f_1 Q + f_2 Q^2, \tag{B.14}$$

such that we will have to find f_0, f_1 and f_2. We start by defining

$$u = \sqrt{\frac{1}{3}c_1}\cos\left(\frac{1}{3}\theta\right), \tag{B.15}$$

$$w = \sqrt{c_1}\sin\left(\frac{1}{3}\theta\right), \tag{B.16}$$

$$\theta = \arccos\left(\frac{c_0}{c_0^{\max}}\right). \tag{B.17}$$

From this, we have that f_j with $j = 0, 1, 2$ is defined as

$$f_j = \frac{h_j}{9u^2 - w^2}, \tag{B.18}$$

with

$$h_0 = (u^2 - w^2)\mathrm{e}^{2iu} + \mathrm{e}^{-iu}\left[8u^2\cos(w) + 2iu(3u^2 + w^2)\xi_0(w)\right], \tag{B.19}$$

$$h_1 = 2u\mathrm{e}^{2iu} - \mathrm{e}^{-iu}\left[2u\cos(w) - i(3u^2 - w^2)\xi_0(w)\right], \tag{B.20}$$

$$h_2 = \mathrm{e}^{2iu} - \mathrm{e}^{-iu}\left[\cos(w) + 3iu\xi_0(w)\right]. \tag{B.21}$$

The function $\xi_0(w)$ is given by

$$\xi_0(w) = \begin{cases} \frac{\sin(w)}{w} & \text{if} \quad |w| > 0.05 \\ 1 - \frac{1}{6}w^2(1 - \frac{1}{20}w^2(1 - \frac{w^2}{42})) & \text{else} \end{cases} \tag{B.22}$$

[1] A max of c_1 is also presented in [46], but it is not needed in order to exponentiate a SU(3) matrix.

where we in the case of $w \to 0$ and $c_0 \to c_0^{\max}$ uses a numerically safe approximation to $\sin(w)/w$.

In the limit of $w \to 3u \to \sqrt{3}/2$ as $c_0 \to -c_0^{\max}$(i.e. having $c_0 < 0$), we use the symmetry of $c_0 \to -c_0$ and take the complex conjugate of f_j,

$$f_j(-c_0, c_1) = (-1)^j f_j(c_0, c_1). \tag{B.23}$$

All that remains is thus to insert f_j into eq. (B.14) on the preceding page, and we will have exponentiated the matrix Q.

B.4.1 Analyzing the matrix exponentiation

An analysis was performed in order to determine the error of the matrix exponentiation and the computational cost. We will use the method described by Morningstar and Peardon [46] and by Lüscher [39, appendix A] compared against regular Taylor expansions.

Timing analysis

We performed a timing analysis where we ran 10^6 exponentiations for each method. The results for the Morningstar method, the Lüscher method and a few select Taylor expansions(denoted by Taylor(N) with N indicating degree), can be seen in table B.1.

Table B.1: The time of performing 10^6 exponentiations with a select method.

Method	Timing[seconds]
Morningstar	0.38
Lüscher	1.82
Taylor(4)	0.40
Taylor(8)	0.79
Taylor(16)	1.55

The results how much time the Taylor exponentiation takes for different polynomial degrees and how it compares against the Morningstar method, can both be seen in fig. B.1 on the facing page. At $y = 1$ a line is drawn, which indicates where N Taylor expansions become computationally more expensive than the method presented by Morningstar and Peardon [46].

If the method by Morningstar and Peardon [46] turns out to be equal or more accurate than Taylor expanding of order $N = 4$, it will be the obvious method of choice for exponentiating SU(3) matrices.

Accuracy analysis

Similar to the previous section, an analysis of the accuracy has been made investigating the differences between the Morningstar method and the Taylor expansions at different

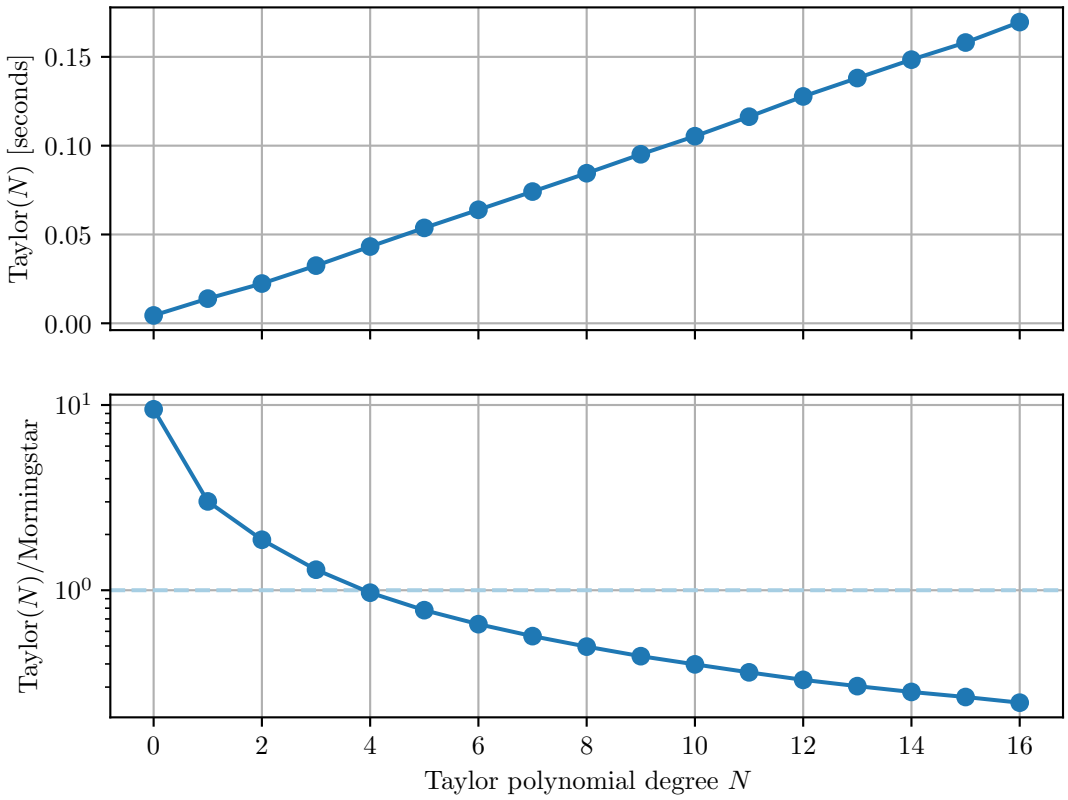


Figure B.1: In the upper figure, we see how the time evolves for different expansion degrees N. In the bottom figure we see how the Morningstar method measures up against the time it takes to calculate the Taylor expansion of degree N. A line at $y = 1$ marks the point where Taylor expansions become computationally more expensive than the method presented in [46].

Taylor expansion degrees N. The absolute difference between a Taylor expansion of degree N and the method by Morningstar and Peardon [46] and a Taylor expansion of degree 16 can be seen in fig. B.1.

In fig. B.2 on the following page we found that after a Taylor expansion of degree $N = 10$, the improvement in accuracy appeared to flatten. This indicates that the numerical precision of the Taylor expansion is too crude to tell us anything of significance, other than that we can expect the method by Morningstar to be accurate up to order 10. In order to implement a stronger analysis, steps to avoid numerical imprecision would have to be implemented in the Taylor-expansions.

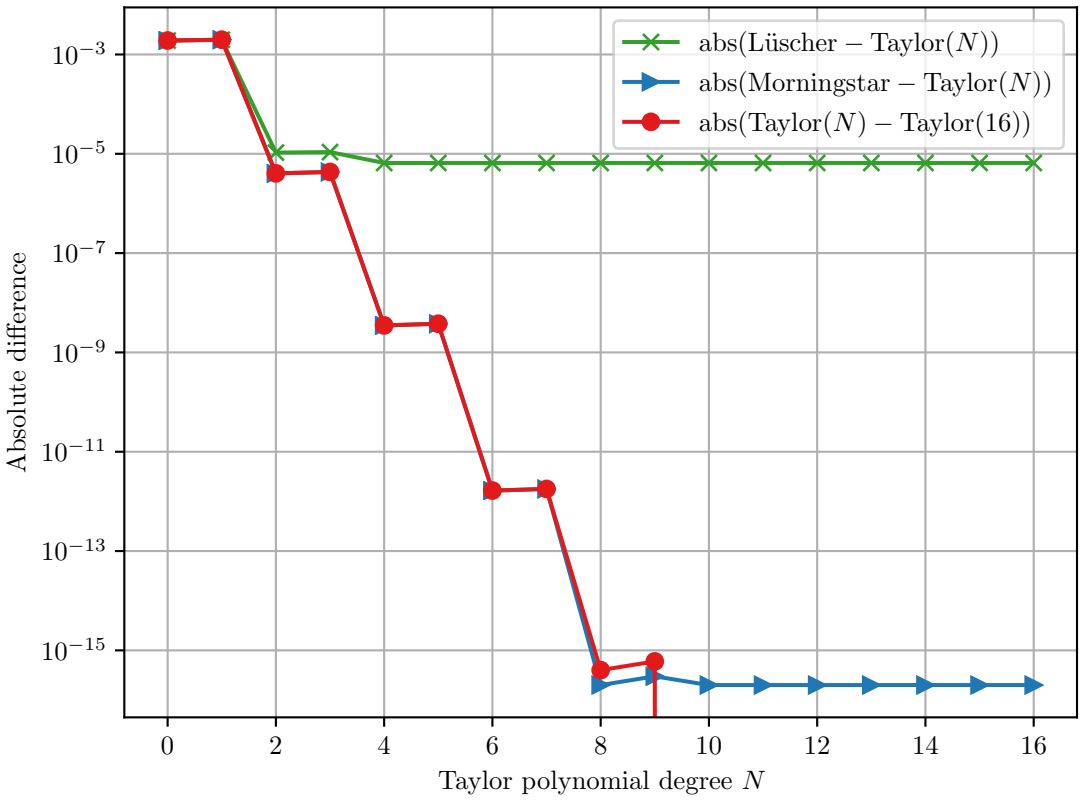


Figure B.2: The absolute difference between the norms of the first element in SU(3) matrix expansion methods by Morningstar and Peardon [46], Lüscher [39, appendix A] and Taylor expansions of degree 16 for different Taylor polynomials of degree N.

B.5 Notes on chirality projectors

Recalling the QCD Lagrangian for a single flavor in eq. (2.30) on page 16 $N_f = 1$ in Euclidean space,

$$\mathcal{L}(\psi, \bar{\psi}, A) = \bar{\psi}\gamma_\mu(\partial_\mu + iA_\mu)\psi, \tag{B.24}$$

we can attempt to split it into left- and right-handed components such that these are invariant under a chiral symmetry,

$$\psi \to \psi' = \mathrm{e}^{i\alpha\gamma_5}\psi, \quad \bar{\psi} \to \bar{\psi}' = \bar{\psi}\mathrm{e}^{i\alpha\gamma_5}. \tag{B.25}$$

We write the chiral projectors as

$$P_R = \frac{1+\gamma_5}{2}, \quad P_L = \frac{1-\gamma_5}{2}. \tag{B.26}$$

We can now define left- and right-handed fields by applying the chiral projectors in eq. (B.26) on the facing page on the fermion fields in eq. (B.24) on the preceding page,

$$\psi_R = P_R\psi, \quad \psi_L = P_L\psi. \tag{B.27}$$

Their Hermitian conjugates are given as

$$\begin{aligned}\bar{\psi}_R = \overline{P_R\psi} = (P_R\psi)^\dagger\gamma_0 = \psi^\dagger P_R^\dagger\gamma_0 = \psi^\dagger\gamma_0 P_L = \bar{\psi}P_L,\\ \bar{\psi}_L = \bar{\psi}P_R,\end{aligned} \tag{B.28}$$

where we used that $P_R^\dagger = P_R$, $P_L^\dagger = P_L$ and the anti-commutation of γ_5.

Before we move on, we note that $P_RP_L = P_LP_R = 0$ commutes, $\gamma_\mu P_{R/L} = P_{L/R}\gamma_\mu$, and that $1 = P_R + P_L$. We start with inserting unity in the Lagrangian seen in eq. (B.24) on the facing page,

$$\begin{aligned}\bar{\psi}1\gamma_\mu(\partial_\mu + iA_\mu)1\psi &= \bar{\psi}(P_R + P_L)\gamma_\mu(\partial_\mu + iA_\mu)(P_R + P_L)\psi\\ &= \bar{\psi}P_R\gamma_\mu(\partial_\mu + iA_\mu)P_R\psi + \bar{\psi}P_R\gamma_\mu(\partial_\mu + iA_\mu)P_L\psi\\ &+ \bar{\psi}P_L\gamma_\mu(\partial_\mu + iA_\mu)P_R\psi + \bar{\psi}P_L\gamma_\mu(\partial_\mu + iA_\mu)P_L\psi\\ &= \bar{\psi}\gamma_\mu(\partial_\mu + iA_\mu)P_LP_R\psi + \bar{\psi}_L\gamma_\mu(\partial_\mu + iA_\mu)\psi_L\\ &+ \bar{\psi}_R\gamma_\mu(\partial_\mu + iA_\mu)\psi_R + \bar{\psi}\gamma_\mu(\partial_\mu + iA_\mu)P_RP_L\psi\\ &= \bar{\psi}_L\gamma_\mu(\partial_\mu + iA_\mu)\psi_L + \bar{\psi}_R\gamma_\mu(\partial_\mu + iA_\mu)\psi_R.\end{aligned} \tag{B.29}$$

Using the aforementioned identities, we have thus split the massless section of the Lagrangian in eq. (B.24) on the preceding page into left- and right-handed projections. If we now go to the massive section,

$$\begin{aligned}m\bar{\psi}\psi &= m\bar{\psi}(P_R + P_L)\psi = m\psi^\dagger 1\gamma_0 P_R\psi + m\psi^\dagger 1\gamma_0 P_L\psi\\ &= m\psi^\dagger(P_R + P_L)\gamma_0 P_R\psi + m\psi^\dagger(P_R + P_L)\gamma_0 P_L\psi\\ &= m\psi^\dagger\gamma_0 P_LP_R\psi + m\psi^\dagger\gamma_0 P_RP_R\psi + m\psi^\dagger\gamma_0 P_LP_L\psi + m\psi^\dagger\gamma_0 P_RP_L\psi\\ &= 0 + m\bar{\psi}P_RP_R\psi + m\bar{\psi}P_LP_L\psi + 0\\ &= m(\bar{\psi}_L\psi_R + \bar{\psi}_R\psi_L),\end{aligned} \tag{B.30}$$

where we see we get a mixing of left- and right-handed fermion fields. Thus we have that when we are *not* in the limit of $m_f \to 0$, we lose the chiral symmetry.

This can also be seen when trying to transform the fields according to eq. (B.25) on the facing page,

$$m\bar{\psi}'\psi' = m\bar{\psi}\mathrm{e}^{i2\alpha\gamma_5}\psi, \tag{B.31}$$

where the mass term clearly don't transform under a chiral rotation.

B.6 Proof of field strength tensor identity

We wish to show that,

$$F^a_{\mu\nu}F^a_{\mu\nu} = \pm F^a_{\mu\nu}\tilde{F}^a_{\mu\nu} + \frac{1}{2}(F^a_{\mu\nu} \mp \tilde{F}^a_{\mu\nu})^2. \tag{B.32}$$

Using that the (anti) dual field strength tensor is defined as

$$\tilde{F}_{\mu\nu} = \pm\frac{1}{2}\epsilon_{\mu\nu\sigma\rho}F_{\sigma\rho}, \tag{B.33}$$

we can start with the right hand side of eq. (B.32),

$$\pm F^a_{\mu\nu}\tilde{F}^a_{\mu\nu} + \frac{1}{2}(F^a_{\mu\nu} \mp \tilde{F}^a_{\mu\nu})^2 = \pm F^a_{\mu\nu}\tilde{F}^a_{\mu\nu} + \frac{1}{2}F^a_{\mu\nu}F^a_{\mu\nu} \mp \frac{1}{2}\left(2F^a_{\mu\nu}\tilde{F}^a_{\mu\nu}\right) + \frac{1}{2}\tilde{F}^a_{\mu\nu}\tilde{F}^a_{\mu\nu}. \tag{B.34}$$

If we now use that the Levi-Civita tensor squares gives $\epsilon^{\mu\nu\sigma\rho}\epsilon_{\mu\nu\sigma\rho} = 4$, we get that the last part of eq. (B.34) can be rewritten,

$$\text{(B.34)} = \frac{1}{2}F^a_{\mu\nu}F^a_{\mu\nu} + \frac{1}{2}F^a_{\mu\nu}F^a_{\mu\nu} = F^a_{\mu\nu}F^a_{\mu\nu}, \tag{B.35}$$

and we have proved the identity eq. (B.32).

Appendix C

Generating random SU(3) matrices

An alternative method to the *RST*-method of generating SU(3) matrices in section 4.1.3 on page 61 is given here. While this method does not provide us with matrices close to unity, it still provides us with an alternate method of generating random SU(3) matrices. Given three rows in a matrix, $\mathbf{u}$, $\mathbf{v}$ and $\mathbf{w}$, we start by populating the first two rows with random numbers in an interval of $r \in (-\epsilon_{\text{rnd}}, \epsilon_{\text{rnd}})$ where ϵ_{rnd} determines the spread of the random numbers. The second row $\mathbf{v}$ is then orthogonalized using Gram-Schmitt,

$$\begin{aligned} \mathbf{u}_{\text{new}} &= \frac{\mathbf{u}}{|\mathbf{u}|}, \\ \mathbf{v}_{\text{new}} &= \frac{\mathbf{v}'}{|\mathbf{v}'|} \quad \text{where} \quad \mathbf{v}' = \mathbf{v} - \mathbf{u}_{\text{new}}(\mathbf{v} \cdot \mathbf{u}_{\text{new}}), \end{aligned} \tag{C.1}$$

with the third row $\mathbf{w}$ simply being the cross product of the first two rows. This procedure is summed up in algorithm 5.

Algorithm 5 Algorithm for generating a random SU(3) matrix. Takes only ϵ_{rnd} as an input that controls the spread of the matrix.

1: **procedure** GenerateRandomMatrix(ϵ_{rnd})
2: Generate six random complex numbers(in total 12 random numbers) from the uniform distribution $(-\epsilon_{\text{rnd}}, \epsilon_{\text{rnd}})$, and populate $\mathbf{u}$ and $\mathbf{v}$.
3: Normalize $\mathbf{u}_{\text{new}} = \frac{\mathbf{u}}{|\mathbf{u}|}$.
4: Make $\mathbf{v}$ orthogonal to $\mathbf{u}$ using eq. (C.1).
5: Take $\mathbf{w} = \mathbf{u}_{\text{new}} \times \mathbf{v}_{\text{new}}$.
6: **return** matrix $X = [\mathbf{u}_{\text{new}}, \mathbf{v}_{\text{new}}, \mathbf{w}]$
7: **end procedure**

Appendix D

Statistical analysis

In this section, we will present the most pertinent sections of statistics relevant to this thesis. One of the issues of lattice QCD is the tremendous computational cost and the consequentally induced autocorrelation between the gauge configurations for small lattice spacing a. For the first part, we can partly remedy the issue of having a small dataset by performing bootstrapping and jackknifing. For the latter, we can find the autocorrelation to get a better estimate for the error. An automated blocking method such as the one described in [30] were considered, but due to our small ensembles this path was discarded.

D.1 Bootstrapping

Bootstrapping is one of the first tools we turn to when tackling either a low-statistics dataset or when attempting to estimate errors of convoluted observables[1].

Given a dataset $\{x_n\}$ of N samples, we will create N_{bs} bootstrap samples $\{\tilde{x}_i\}$, where each $\tilde{x}_i$ is built from randomly drawing with replacement N samples from $\{x_n\}$ and taking the mean of these randomly drawn samples. Thus, we are forming a new data set $\{\tilde{x}_i\}$. From this, we get a final estimator

$$\langle \tilde{x} \rangle = \frac{1}{N_{\text{bs}}} \sum_{i=1}^{N_{\text{bs}}} \tilde{x}_i, \tag{D.1}$$

with the standard deviation

$$\sigma_{\tilde{x}} = \frac{N_{\text{bs}} - 1}{N_{\text{bs}}} \sum_{i=1}^{N_{\text{bs}}} \big(\tilde{x}_i - \langle \tilde{x} \rangle \big)^2. \tag{D.2}$$

The $\langle \ldots \rangle$ indicates an average, as per convention.

A quick sketch of bootstrapping can be seen in algorithm 6 on the next page.

[1]The latter, while not relevant for us, could be estimating errors on linear regression coefficients.

Algorithm 6 An algorithmic description of bootstrapping a dataset $\{x_n\}$.

1: **for** $i = 0 : N_{\text{bs}}$ **do**
2: With replacement, randomly draw N_{obs} elements, forming temporary dataset $\{\tilde{\tilde{x}}_n\}$.
3: Take the mean of the new data set, $\tilde{x}_i = \frac{1}{N_{\text{obs}}} \sum_{n=1}^{N_{\text{obs}}} \tilde{\tilde{x}}_n$.
4: Add $\tilde{x}_i$ to a new dataset of bootstrapped values.
5: **end for**
6: Compute statistics on the bootstrapped dataset $\{\tilde{x}_i\}$.

D.2 Jackknifing

An alternative to bootstrapping is *jackknifing*. The name stems from the fact that we are removing every jth element when creating a "new" dataset $\{\tilde{x}_n\}$ of size N. Each element in the dataset $\tilde{x}_n$ is a mean of the $N-1$ elements of the original dataset, excluding the nth element. The final variance of jackknifing is given as

$$\sigma_{\hat{x}} = \frac{N-1}{N} \left(\tilde{x}_i - \langle \hat{x} \rangle\right)^2 . \tag{D.3}$$

The final unbiased estimator is reported as $\langle \hat{x} \rangle = \langle x \rangle - (N-1)(\langle \tilde{x} \rangle - \langle x \rangle)$, with $\langle \tilde{x} \rangle$ given by

$$\langle \tilde{x} \rangle = \frac{1}{N} \sum_{n=1}^{N} \tilde{x}_n . \tag{D.4}$$

An algorithmic implementation of jackknifing can be seen in algorithm 7.

Algorithm 7 An algorithmic description of jackknifing a dataset $\{x_n\}$.

1: **for** $j = 0 : N_{\text{obs}}$ **do**
2: Take the mean $\tilde{x}_n$ of $\{x_n\}$, excluding the jth element.
3: Add the mean $\tilde{x}_n$ to a new dataset, $\{\tilde{x}_n\}$.
4: **end for**
5: Take the mean in eq. (D.4) and the variance in eq. (D.3) of the new dataset $\{\tilde{x}_n\}$.
6: Take the mean of the original dataset, $\langle x \rangle$.
7: Remove the bias by $\langle \hat{x} \rangle = \langle x \rangle - (N-1)(\langle \tilde{x} \rangle - \langle x \rangle)$, to get the final mean.

D.3 Autocorrelation

Autocorrelation is a measure of correlations within a time series signal. Since we are dealing with a Monte Carlo time series, we wish to measure if what we are sampling is periodically self-similar in Monte Carlo time. If this is the case, our statistics may be far worse than what we are lead to believe due to the periodically repeating signal.

The method we will utilize is the one presented by Wolff [79], as it is applicable to a wide range of problems and applicable to propagated signals. We first present an alternative method by Lüscher [39] that is easier to implement for finding the autocorrelation, which serves as a gateway to the more general method by Wolff [79].

D.3.1 Notational conventions for autocorrelation

Notational conventions for this section will be close to that presented by Wolff [79],

- **Observables** are denoted by a Greek index, such as $\alpha, \beta, ...$
- **Measurements** are denoted by Latin letters, $i, j, ...$ with $i = 1, 2, \ldots, N_r$, where N_r denotes the number of measurements for a replicum.
- **Replicums** are denoted by $r = 1, 2, \ldots, R$ and indicate the number of statistically independent replica, i.e. sets of measurements. R is the total number of sets of measurements.

A single measurement of an observable α is denoted by $a_\alpha^{i,r}$, and have a per-replicum average of

$$\bar{a}_\alpha^r = \frac{1}{N_r} \sum_{i=1}^{N_r} a_\alpha^{i,r}, \tag{D.5}$$

with the total average of all replicums given as

$$\bar{\bar{a}}_\alpha = \frac{1}{N} \sum_{r=1}^{R} N_r \bar{a}_\alpha^r. \tag{D.6}$$

The total number of measurements N is given as

$$N = \sum_{r=1}^{R} N_r. \tag{D.7}$$

The exact statistical mean of α is given as A_α.

The autocorrelation function between two observables α and β(these may be equal in order to investigate self-periodicity) is given as

$$\Gamma_{\alpha\beta}(t) = \left\langle (a_\alpha^i - A_\alpha)(a_\beta^{i+t} - A_\beta) \right\rangle = \Gamma_{\beta\alpha}(-t), \tag{D.8}$$

where t is the *lag*, which is the number of measurements we are measuring correlation between.

D.3.2 Single-variable autocorrelation

A method by Lüscher [39] of measuring the autocorrelation for a single variable and single replicum is presented. Given a single observable $\alpha = 1$ and a single replicum $r = 1$, we have that the exact autocorrelation function is given by

$$\Gamma(t) = \Gamma(-t) = \langle (a_i - a)(a_{i+t} - a) \rangle \,, \tag{D.9}$$

where $a = \langle \bar{a} \rangle$ is the true mean since $\bar{a}$ is a stochastic variable. $\bar{a}$ is given as

$$\bar{a} = \frac{1}{N} \sum_{i=1}^{N} a_i, \tag{D.10}$$

and its variance is given by

$$\sigma^2 = \left\langle (a_i - \bar{a})^2 \right\rangle = \frac{1}{N^2} \sum_{i,j=1}^{N} \Gamma(i-j). \tag{D.11}$$

This can be rewritten to

$$\sigma^2 = 2\tau_{\text{int}} \sigma_0^2, \tag{D.12}$$

where $\sigma_0^2 = \Gamma_{\alpha\beta}(0)/N$, and is how we will correct for autocorrelation in results. τ_{int} is the *integrated autocorrelation time* and is given by

$$\tau_{\text{int}} = \frac{1}{2} + \sum_{t=1}^{\infty} \frac{\Gamma(t)}{\Gamma(0)}. \tag{D.13}$$

The autocorrelation from eq. (D.8) on the previous page can be approximated by

$$\bar{\Gamma}(t) = \frac{1}{N-t} \sum_{i=1}^{N-t} (a_i - \bar{a}) \, (a_{i+t} - \bar{a}) \,. \tag{D.14}$$

Substituting $\bar{\rho}(t) = \bar{\Gamma}(t)/\bar{\Gamma}(0)$ one has that the variance becomes

$$\left\langle \delta\rho(t)^2 \right\rangle \simeq \frac{1}{N} \sum_{k=1}^{t+\Lambda} \left\{ \bar{\rho}(k+t) + \rho(k^{-} - t) - 2\bar{\rho}(k)\bar{\rho}(t) \right\}^2 , \tag{D.15}$$

where Λ is a cutoff, and is usually larger than 100. The reason for this is that after some lag t the autocorrelation will be dominated by fluctuations, and hence become unreliable. This expression can be used to determine the integrated autocorrelation time,

$$\tau_{\text{int}} = \frac{1}{2} + \sum_{t=1}^{W} \bar{\rho}(t), \tag{D.16}$$

where W is chosen to be the first lag t where

$$\bar{\rho}(t) \leq \left\langle \delta\rho(t)^2 \right\rangle^{1/2}. \tag{D.17}$$

The uncertainty of the integrated autocorrelation is given by the Madras-Sokal formula [43],

$$\left\langle \delta\tau_{\text{int}}^2 \right\rangle \approx \frac{2(W+1)}{N}\tau_{\text{int}}^2. \tag{D.18}$$

D.3.3 Propagated autocorrelation

In this section we present a general method by Wolff [79] for calculating the autocorrelation when it is propagated through a function. Given some function f which propagates $\bar{\bar{a}}_\alpha$, where the double bar is given by eq. (D.6) on page 151,

$$\bar{\bar{F}} = f(\bar{\bar{a}}_\alpha), \tag{D.19}$$

such that

$$\bar{F} = \frac{1}{N}\sum_{r=1}^{R} N_r f(\bar{a}_\alpha^r). \tag{D.20}$$

We define derivatives of f as

$$f_\alpha = \frac{\partial f}{\partial A_\alpha}. \tag{D.21}$$

If no analytical evaluation is available, one may use a symmetric numerical approximation as of $\mathcal{O}(h_\alpha^2)$,

$$\bar{\bar{f}}_\alpha \approx \frac{1}{2h_\alpha}\left[f(\bar{\bar{a}}_1, \bar{\bar{a}}_2, \dots, \bar{\bar{a}}_\alpha + h_\alpha, \dots) - f(\bar{\bar{a}}_1, \bar{\bar{a}}_2, \dots, \bar{\bar{a}}_\alpha - h_\alpha, \dots)\right] \tag{D.22}$$

with the step size given as

$$h_\alpha = \sqrt{\frac{\bar{\bar{\Gamma}}_{\alpha\alpha}(0)}{N}}. \tag{D.23}$$

Given some measured data $a_\alpha^{i,r}$, we define the autocorrelation function $\bar{\bar{\Gamma}}_{\alpha\beta}$ as

$$\bar{\bar{\Gamma}}_{\alpha\beta}(t) = \frac{1}{N - Rt}\sum_{r=1}^{R}\sum_{i=1}^{N_r - t}\left(a_\alpha^{i,r} - \bar{\bar{a}}_\alpha\right)\left(a_\beta^{i+t,r} - \bar{\bar{a}}_\beta\right). \tag{D.24}$$

Given a derived quantity F, the full autocorrelation becomes

$$\bar{\bar{\Gamma}}_F(t) = \sum_{\alpha\beta}\bar{\bar{f}}_\alpha\bar{\bar{f}}_\beta\bar{\bar{\Gamma}}_{\alpha\beta}(t), \tag{D.25}$$

where $\bar{\bar{f}}_\alpha$ is given by eq. (D.21) on the previous page and evaluated at every α, $\bar{\bar{a}}_1, \bar{\bar{a}}_2, \ldots$.

As a practical side note, in the case we only need to calculate the error propagation of a single observable, A_α we can simply project the error directly on the measurements,

$$a_\alpha^{i,r} \to a_f^{i,r} = \sum_a \bar{\bar{f}}_\alpha a_\alpha^{i,r}, \tag{D.26}$$

and we will not have to calculate the full autocorrelation matrix $\bar{\bar{\Gamma}}_{\alpha\beta}(t)$.

We have that τ_{int} may be estimated as

$$\bar{\bar{\tau}}_{\text{int},F}(W) = \frac{\bar{\bar{C}}_F(W)}{2\bar{\bar{v}}_F}, \tag{D.27}$$

where $\bar{\bar{C}}_F$ is given by

$$\bar{\bar{C}}_F(W) = \left[\bar{\bar{\Gamma}}_F(0) + 2\sum_{t=1}^{W} \bar{\bar{\Gamma}}_F(t)\right], \tag{D.28}$$

and $\bar{\bar{v}}_F$ is given as

$$\bar{\bar{v}}_F = \bar{\bar{\Gamma}}_F(0). \tag{D.29}$$

The error of $\bar{\bar{\tau}}_{\text{int},F}(W)$ is given as

$$\left\langle \delta\bar{\bar{\tau}}^2_{\text{int},F} \right\rangle = \left\langle (\bar{\bar{\tau}}_{\text{int},F}(W) - \tau_{\text{int},F})^2 \right\rangle \approx \frac{4}{N}\left(W + 1/2 - \tau_{\text{int},F}\right)\tau^2_{\text{int},F}. \tag{D.30}$$

The final error estimate now becomes

$$\bar{\bar{\sigma}}^2_F = \frac{\bar{\bar{C}}(W)}{N}, \tag{D.31}$$

where we have removed the bias in $\bar{\bar{C}}(W)$ by a rescaling of

$$\bar{\bar{C}}(W) \to \bar{\bar{C}}(W)\left(1 + \frac{2W+1}{N}\right). \tag{D.32}$$

The next challenge is to find W. W is defined as the value that minimizes the sum of absolute errors when looking at $\bar{\bar{C}}_F(W)$. For finding W we will use an automatic windowing procedure, in which Wolff [79] proposed to start from the hypothesis that $\tau \sim S\tau_{\text{int},F}$ where S is a tunable factor. This is stated as a problem where we wish to solve

$$2\bar{\bar{\tau}}_{\text{int},F}(W) = \sum_{t=-\infty}^{\infty} \exp\left(-S|t|/\bar{\bar{\tau}}(W)\right), \tag{D.33}$$

for $\bar{\bar{\tau}}$. If we assume small $\bar{\bar{\tau}}$, we can make an approximation,

$$\frac{S}{\bar{\bar{\tau}}(W)} = \ln\left(\frac{2\bar{\bar{\tau}}_{\text{int},F}(W)+1}{\bar{\bar{\tau}}_{\text{int},F}(W)-1}\right). \tag{D.34}$$

Solving for $\bar{\bar{\tau}}$, we will set it to a small positive value if $\bar{\bar{\tau}}_{\text{int},F} \leq 1/2$. Then, for $W = 1, 2, \ldots$ we calculate

$$g(W) = \exp\left[-\frac{W}{\bar{\bar{\tau}}(W)} - \frac{\bar{\bar{\tau}}(W)}{\sqrt{WN}}\right], \tag{D.35}$$

until $g(W)$ is negative. The corresponding W for this is taken to be our integration window. The value of $S \in [1, 2]$ is tuned such that $\bar{\bar{\tau}}_{\text{int},F}(W)$ exhibits a plateau. That is, we integrate the autocorrelation time until we begin exhibiting large statistical fluctuations. The optimal W will then give us the true integrated autocorrelation time $\bar{\bar{\tau}}_{\text{int},F}$.

The autocorrelation corrected variance is given by

$$\sigma^2 = 2\tau_{\text{int}}\sigma_0^2. \tag{D.36}$$

An algorithm for the automatic windowing procedure can be seen in algorithm 8.

Algorithm 8 An algorithmic description of the automatic windowing procedure for retrieving the optimal cut-off W for $\tau_{\text{int},F}(W)$, using the method presented in [79]. Note that $\tau_{\text{int},F}(W)$ is a vector given by eq. (D.27) on the facing page.

procedure AUTOMATICWINDOWING($\tau_{\text{int},F}(W), S, \epsilon = 10^{-7}$)
  Set up an empty vector τ to populate.
  for $\tau^i_{\text{int},F}(W)$ in $\tau_{\text{int},F}(W)$ **do**
    if $\tau^i_{\text{int},F}(W) \leq 0.5$ **then**
      $\tau^i = \epsilon$.
    else
      $\tau^i = S\left(\log\left[\frac{2\tau^i_{\text{int},F}(W)+1}{2\tau^i_{\text{int},F}(W)-1}\right]\right)^{-1}$.
    end if
  end for
  for $i = 1 : \text{length}(\tau)$ **do**
    if $\exp\left[-\frac{i}{\tau^i} - \frac{\tau^i}{\sqrt{iN}}\right] < 0$ **then**
      $W = i$.
      Break.
    end if
  end for
  Return W.
end procedure

D.4 Bootstrapping a time series

Since we have a time series that exhibits autocorrelations, we need to account for these in the bootstrap in order to properly account for autocorrelation. This can be done by what is called *bootstrap blocking* or simply time series bootstrap [33, 52]. The idea behind it is intuitive. Given a dataset $\{x_n\}$ of N samples, we will as in basic bootstrapping create N_{bs} bootstrap samples. However, when generating a bootstrap sample we will populate it with "chunks" of size b as this will preserve autocorrelations in the sampling. The total number of these "chunks" will be $k = N/b$, and they can be overlapping. The size of the chunk b is determined from where we get a maximal standard deviation or have that the autocorrelation $\bar{\bar{\Gamma}}(t) \to 0$. An example of this can be seen in fig. D.1 on the next page, where a dashed, gray line has been drawn at where the block size b was found to become negative. As we can see, the choice of the block size seems to maximize the error to a reasonable degree.

A quick sketch of bootstrap blocking can be seen in algorithm 9.

Algorithm 9 An algorithmic description of bootstrap blocking a dataset $\{x_n\}$.

1: Find the optimal block size b from the autocorrelation, where b is the index for when $\bar{\bar{\Gamma}}(t)$ first becomes negative(see appendix D.3.2 on page 152).
2: **for** $i = 0 : N_{\text{bs}}$ **do**
3: Prepare an empty array $\tilde{\tilde{x}}$ of size N.
4: With replacement, draw $k \in [0, N - b)$ random numbers $\{j\}$.
5: **for** each j in $\{j\}$ **do**
6: Populate $\{\tilde{\tilde{x}}\}$ with chunks of size $\tilde{\tilde{x}}_j = x_j, \ldots, x_{j+b}$.
7: **end for**
8: Take the mean(or chosen bootstrap statistic) of $\{\tilde{\tilde{x}}\}$, and use that to set all $\tilde{x}_i$. If $\tilde{\tilde{x}}$ exceed N, the excess elements will be removed.
9: Add $\tilde{x}_i$ to a new dataset of bootstrapped values $\{\tilde{x}_i\}$.
10: **end for**
11: Compute statistics on the bootstrapped dataset $\{\tilde{x}_i\}$.

D.5 Line fitting

In order to extrapolate our results to the continuum, a proper method for linear extrapolation is needed. We present the line fitting procedure given by Lavagnini and Magno [35]. In the paper, two methods are presented - one unweighted, and one weighted. We will settle on presenting the weighted method(although both has been implemented), and follow the notation used in the paper.

We will try to fit a line to

$$\hat{y}_{jw} = b_{0w} + b_{1w} x_j, \tag{D.37}$$

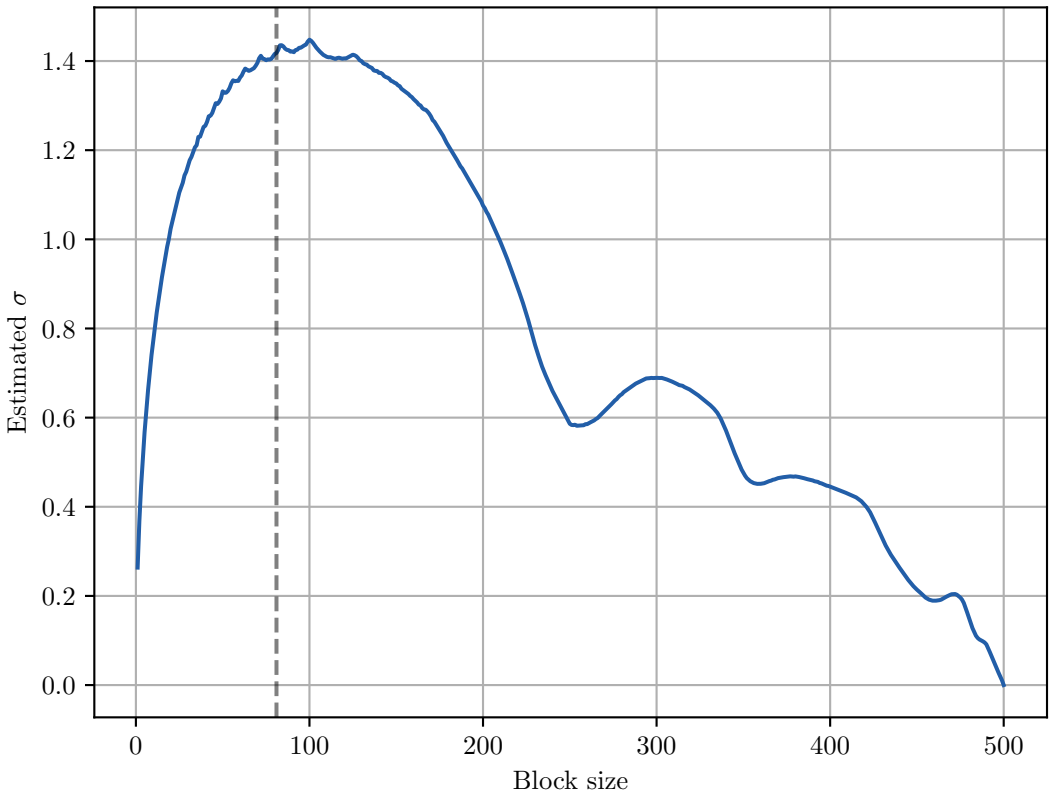


Figure D.1: Standard deviation versus the block size in the time series bootstrapping. The gray line indicates where the autocorrelation first turns negative and is where we will choose to set our block size b. In this case, that was at $b = 81$. The data presented here consists 500 topological charge samples at $t_f/a^2 = 10.0$ at $\beta = 6.2$, generated such that we get strong correlations in Monte Carlo time. $N_{\text{bs}} = 10^6$ bootstrap samples were used in order to get a smooth curve, although far fewer samples should be required.

where $\hat{y}_{jw}$ is the line we will predict given input points x_j. Assuming that we are going to fit to $\bar{y}_w$ and $\bar{x}_w$, we solve for b_{0w},

$$b_{0w} = \bar{y}_w - b_{1w}\bar{x}_w, \tag{D.38}$$

where the subscript w symbolizes that we are performing a weighted fit and $\bar{y}_w$ is given by

$$\bar{y}_w = \frac{\sum_{i=1}^{n} w_i y_i}{\sum_{i=1}^{n} w_i}, \tag{D.39}$$

and $\bar{x}_w$ is given by

$$\bar{x}_w = \frac{\sum_{i=1}^{n} w_i x_i}{\sum_{i=1}^{n} w_i}. \tag{D.40}$$

b_{1w} is given by

$$b_{1w} = \frac{\sum_{i=1}^{n} w_i \left(x_i - \bar{x}_w\right) y_i}{\sum_{i=1}^{n} w_i (x_i - \bar{x}_w)^2}, \tag{D.41}$$

and the weights w_i is given by

$$w_i = \frac{1}{s_i^2}, \tag{D.42}$$

where $s_i = \sigma_i$. We denote the error in a term by s.

The next step is now to retrieve the error in b_{0w}, b_{1w}, and the slope, in order to retrieve the error bands of y_{jw}, where j indicates a set of we will predict a new line for. The error in b_{0w} is given by

$$s_{b_{0w}}^2 = \left(\frac{1}{\sum_{i=1}^{n} w_i} + \frac{\bar{x}_w^2}{\sum_{i=1}^{n} w_i (x_i - \bar{x}_w)^2} \right) \left(s_{y/x}\right)_w^2. \tag{D.43}$$

Next, the error in b_{1w} is given by

$$s_{b_{1w}}^2 = \frac{1}{\sum_{i=1}^{n} w_i (x_i - \bar{x}_w)^2} \left(s_{y/x}\right)_w^2, \tag{D.44}$$

where the $(s_{y/x})_w^2$ is the error in the slope,

$$(s_{y/x})_w^2 = \frac{\sum_{i=1}^{n} w_i \left(y_i - \hat{y}_{iw}\right)^2}{n-2}. \tag{D.45}$$

From this we can calculate the upper and lower limits of $\bar{y}_{jw}$,

$$\bar{y}_{jw}^{\pm} = b_{0w} + b_{1w} x_j \pm t(1 - \alpha/2, n-2) \tag{D.46}$$

$$\times \left(s_{y/x}\right)_w \left(\frac{1}{\sum_{i=1}^{n} w_i} + \frac{\left(x_j - \bar{x}_w\right)^2}{\sum_{i=1}^{n} w_i \left(x_i - \bar{x}_w\right)^2} \right), \tag{D.47}$$

where $t(1 - \alpha/2, n-2)$ is the constant of $(1-\alpha/2)100\%$ percentage point of Student's t-distribution in a one-sided interval. We will set $\alpha = 2 \times 0.32$, such that we pick up one standard deviation of the distribution.

Appendix E

Cumulants

A short introduction to cumulants is presented in this section, as it is used in the investigation of the topological charge. Given a probability distribution $P(x)$, we wish to describe how some stochastic variable X takes on x. That is, how x behaves in the $P(x)$ distribution. For this purpose, let us introduce a characteristic function $G(k)$,

$$G(k) = \left\langle \mathrm{e}^{ikx} \right\rangle = \int \mathrm{d}x \mathrm{e}^{ikx} P(x), \tag{E.1}$$

with of course $\int \mathrm{d}x P(x) = 1$. Note that $G(0) = 1$. We then expand the exponential,

$$G(k) = \int \mathrm{d}x \sum_{n=0}^{\infty} \frac{(ik)^n}{n!} x^n P(x) = \sum_{n=0}^{\infty} \frac{(ik)^n}{n!} \left\langle x^n \right\rangle, \tag{E.2}$$

such that in order to generate the *moments* of the distribution, we take the nth derivative,

$$\left. \frac{\mathrm{d}^n G(k)}{\mathrm{d}k^n} \right|_{k=0} = i^n \left\langle x^n \right\rangle. \tag{E.3}$$

To generate the cumulants, we take the log of this expression(note that the first term of the sum which is 1, dies) and get the cumulants C_n,

$$\ln G(k) = \sum_{n=1}^{\infty} \frac{(ik)^n}{n!} C_n. \tag{E.4}$$

Isolating the cumulants can be done if we take the nth derivative,

$$C_n = (-i)^n \left. \frac{\mathrm{d}^n \ln G(k)}{\mathrm{d}k^n} \right|_{k=0}. \tag{E.5}$$

With this definition in place, let us go ahead and make the first four calculations explicitly.

E.1 First cumulant

The first cumulant $n = 1$ is given by

$$C_1 = (-i)^1 \frac{1}{G(k)} \frac{\mathrm{d}^1 G(k)}{\mathrm{d}k^1}\bigg|_{k=0} = (-i)i \langle x \rangle = \langle x \rangle \,, \tag{E.6}$$

where after taking the derivative of the log, one can simply use eq. (E.3) on the preceding page for the moment. We end up with the first cumulant being the mean of the distribution.

E.2 Second cumulant

The second cumulant $n = 1$ is given as

$$C_2 = (-i)^2 \frac{\mathrm{d}}{\mathrm{d}k} \left[\frac{\mathrm{d}}{\mathrm{d}k} \ln G(k) \right] \bigg|_{k=0} = -\frac{\mathrm{d}}{\mathrm{d}k} \left[\frac{1}{G(k)} \ln G(k) \right] \bigg|_{k=0} \tag{E.7}$$

$$= \left[\frac{1}{G(k)^2} \left[\frac{\mathrm{d}}{\mathrm{d}k} G(k) \right]^2 - \frac{1}{G(k)} \frac{\mathrm{d}^2}{\mathrm{d}k^2} G(k) \right] \bigg|_{k=0} \tag{E.8}$$

$$= \langle x \rangle^2 - \left\langle x^2 \right\rangle . \tag{E.9}$$

E.3 Third cumulant

Taking the derivative of eq. (E.8) gives us the third cumulant,

$$C_3 = (-i)^3 \frac{\mathrm{d}}{\mathrm{d}k} \left[\frac{1}{G(k)^2} \left[\frac{\mathrm{d}}{\mathrm{d}k} G(k) \right]^2 - \frac{1}{G(k)} \frac{\mathrm{d}^2}{\mathrm{d}k^2} G(k) \right] \bigg|_{k=0} \tag{E.10}$$

$$= i \left[\frac{2}{G(k)^3} \left[\frac{\mathrm{d}}{\mathrm{d}k} G(k) \right]^3 - \frac{2}{G(k)^2} \left[\frac{\mathrm{d}}{\mathrm{d}k} G(k) \right] \left[\frac{\mathrm{d}^2}{\mathrm{d}^2 k} G(k) \right] + \frac{1}{G(k)} \frac{\mathrm{d}}{\mathrm{d}k^3} G(k) \right] \bigg|_{k=0} \tag{E.11}$$

$$= 2 \langle x \rangle^3 - 3 \langle x \rangle \left\langle x^2 \right\rangle + \left\langle x^3 \right\rangle . \tag{E.12}$$

E.4 Fourth cumulant

The fourth cumulant is generated from taking the next kth derivative of eq. (E.11) on the facing page.

$$C_4 = (-i)^4 \frac{\mathrm{d}^4}{\mathrm{d}k^4} \left[\ln G(k)\right]\big|_{k=0} \tag{E.13}$$

$$= \frac{\mathrm{d}}{\mathrm{d}k}\left[\frac{2}{G(k)^3}\left[\frac{\mathrm{d}}{\mathrm{d}k}G(k)\right]^3 - \frac{2}{G(k)^2}\left[\frac{\mathrm{d}}{\mathrm{d}k}G(k)\right]\left[\frac{\mathrm{d}^2}{\mathrm{d}^2 k}G(k)\right] + \frac{1}{G(k)}\frac{\mathrm{d}}{\mathrm{d}k^3}G(k)\right]\Bigg|_{k=0} \tag{E.14}$$

$$= \left[\frac{-6}{G(k)^4}\left[\frac{\mathrm{d}}{\mathrm{d}kG(k)}\right]^4 + \frac{12}{G(k)^3}\left[\frac{\mathrm{d}}{\mathrm{d}k^2}G(k)\right]\left[\frac{\mathrm{d}}{\mathrm{d}k}G(k)\right]^2 - \frac{3}{G(k)^2}\left[\frac{\mathrm{d}^2}{\mathrm{d}k^2}G(k)\right]^2 - \frac{4}{G(k)^2}\left[\frac{\mathrm{d}}{\mathrm{d}}G(k)\right]\left[\frac{\mathrm{d}^3}{\mathrm{d}k^3}G(k)\right] + \frac{1}{G(k)}\left[\frac{\mathrm{d}^4}{\mathrm{d}k^4}G(k)\right]\right]\Bigg|_{k=0} \tag{E.15}$$

$$= \left\langle x^4 \right\rangle - 4\left\langle x^3 \right\rangle \left\langle x \right\rangle - 3\left\langle x^2 \right\rangle^2 + 12\left\langle x^2 \right\rangle \left\langle x \right\rangle^2 - 6\left\langle x \right\rangle^4 . \tag{E.16}$$

Bibliography

[1] Alvise Bastianello and Spyros Sotiriadis. Cluster expansion for ground states of local Hamiltonians. *Nuclear Physics B*, 909:1020–1078, August 2016. ISSN 05503213. doi: 10.1016/j.nuclphysb.2016.06.021. URL `http://arxiv.org/abs/1601.02532`. arXiv: 1601.02532.

[2] A.A. Belavin, A.M. Polyakov, A.S. Schwartz, and Yu.S. Tyupkin. Pseudoparticle solutions of the Yang-Mills equations. *Physics Letters B*, 59(1):85–87, October 1975. ISSN 03702693. doi: 10.1016/0370-2693(75)90163-X. URL `http://linkinghub.elsevier.com/retrieve/pii/037026937590163X`.

[3] S. O. Bilson-Thompson, D. B. Leinweber, and A. G. Williams. Highly-improved lattice field-strength tensor. *Annals of Physics*, 304(1):1–21, March 2003. ISSN 00034916. doi: 10.1016/S0003-4916(03)00009-5. URL `http://arxiv.org/abs/hep-lat/0203008`. arXiv: hep-lat/0203008.

[4] V. I. Borodulin, R. N. Rogalyov, and S. R. Slabospitsky. CORE 2.1 (COmpendium of RElations, Version 2.1). *arXiv:hep-ph/9507456*, July 1995. URL `http://arxiv.org/abs/hep-ph/9507456`. arXiv: hep-ph/9507456.

[5] S. Borsanyi, S. Durr, Z. Fodor, C. Hoelbling, S. D. Katz, S. Krieg, T. Kurth, L. Lellouch, T. Lippert, C. McNeile, and K. K. Szabo. High-precision scale setting in lattice QCD. *Journal of High Energy Physics*, 2012(9), September 2012. ISSN 1029-8479. doi: 10.1007/JHEP09(2012)010. URL `http://arxiv.org/abs/1203.4469`. arXiv: 1203.4469.

[6] Falk Bruckmann. Topological objects in QCD. *The European Physical Journal Special Topics*, 152(1):61–88, December 2007. ISSN 1951-6355, 1951-6401. doi: 10.1140/epjst/e2007-00377-2. URL `http://arxiv.org/abs/0706.2269`. arXiv: 0706.2269.

[7] J. C. Butcher. *Numerical Methods for Ordinary Differential Equations*. John Wiley & Sons, Ltd, Chichester, UK, March 2008. ISBN 978-0-470-75376-7 978-0-470-72335-7. doi: 10.1002/9780470753767. URL `http://doi.wiley.com/10.1002/9780470753767`.

[8] Nicola Cabibbo and Enzo Marinari. A new method for updating SU(N) matrices in computer simulations of gauge theories. *Physics Letters B*, 119(4):387–390, De-

cember 1982. ISSN 0370-2693. doi: 10.1016/0370-2693(82)90696-7. URL http://www.sciencedirect.com/science/article/pii/0370269382906967.

[9] Elena Celledoni, Arne Marthinsen, and Brynjulf Owren. Commutator-free Lie group methods. *Future Generation Computer Systems*, 19(3):341–352, April 2003. ISSN 0167-739X. doi: 10.1016/S0167-739X(02)00161-9. URL http://www.sciencedirect.com/science/article/pii/S0167739X02001619.

[10] Y. Chen, A. Alexandru, S. J. Dong, T. Draper, I. Horvath, F. X. Lee, K. F. Liu, N. Mathur, C. Morningstar, M. Peardon, S. Tamhankar, B. L. Young, and J. B. Zhang. Glueball Spectrum and Matrix Elements on Anisotropic Lattices. *Physical Review D*, 73(1), January 2006. ISSN 1550-7998, 1550-2368. doi: 10.1103/PhysRevD.73.014516. URL http://arxiv.org/abs/hep-lat/0510074. arXiv: hep-lat/0510074.

[11] Abhishek Chowdhury, A. Harindranath, and Jyotirmoy Maiti. Correlation and localization properties of topological charge density and the pseudoscalar glueball mass in SU(3) lattice Yang-Mills theory. *Physical Review D*, 91(7), April 2015. ISSN 1550-7998, 1550-2368. doi: 10.1103/PhysRevD.91.074507. URL http://arxiv.org/abs/1409.6459. arXiv: 1409.6459.

[12] P. E. Crouch and R. Grossman. Numerical integration of ordinary differential equations on manifolds. *Journal of Nonlinear Science*, 3(1):1–33, December 1993. ISSN 0938-8974, 1432-1467. doi: 10.1007/BF02429858. URL http://link.springer.com/10.1007/BF02429858.

[13] Marco Cè, Cristian Consonni, Georg P. Engel, and Leonardo Giusti. Non-Gaussianities in the topological charge distribution of the SU(3) Yang–Mills theory. *Physical Review D*, 92(7), October 2015. ISSN 1550-7998, 1550-2368. doi: 10.1103/PhysRevD.92.074502. URL http://arxiv.org/abs/1506.06052. arXiv: 1506.06052.

[14] Luigi Del Debbio, Gian Mario Manca, and Ettore Vicari. Critical slowing down of topological modes. *Physics Letters B*, 594(3-4):315–323, August 2004. ISSN 03702693. doi: 10.1016/j.physletb.2004.05.038. URL http://arxiv.org/abs/hep-lat/0403001. arXiv: hep-lat/0403001.

[15] Luigi Del Debbio, Leonardo Giusti, and Claudio Pica. Topological Susceptibility in SU(3) Gauge Theory. *Physical Review Letters*, 94(3), January 2005. ISSN 0031-9007, 1079-7114. doi: 10.1103/PhysRevLett.94.032003. URL https://link.aps.org/doi/10.1103/PhysRevLett.94.032003.

[16] A. Deur, S. J. Brodsky, and G. F. de Teramond. The QCD Running Coupling. *Progress in Particle and Nuclear Physics*, 90:1–74, September 2016. ISSN 01466410. doi: 10.1016/j.ppnp.2016.04.003. URL http://arxiv.org/abs/1604.08082. arXiv: 1604.08082.

[17] John F. Donoghue, Eugene Golowich, and Barry R. Holstein. *Dynamics of the Standard Model*. Cambridge University Press, Cambridge, 1992. ISBN 978-0-511-52437-0. doi: 10.1017/CBO9780511524370. URL http://ebooks.cambridge.org/ref/id/CBO9780511524370.

[18] Jack Dragos, Thomas Luu, Andrea Shindler, Jordy de Vries, and Ahmed Yousif. Confirming the Existence of the strong CP Problem in Lattice QCD with the Gradient Flow. *arXiv:1902.03254 [hep-lat, physics:hep-ph]*, February 2019. URL http://arxiv.org/abs/1902.03254. arXiv: 1902.03254.

[19] Stephan Durr, Zoltan Fodor, Christian Hoelbling, and Thorsten Kurth. Precision study of the SU(3) topological susceptibility in the continuum. *Journal of High Energy Physics*, 2007(04):055–055, April 2007. ISSN 1029-8479. doi: 10.1088/1126-6708/2007/04/055. URL http://arxiv.org/abs/hep-lat/0612021. arXiv: hep-lat/0612021.

[20] Robert G. Edwards and Balint Joo. The Chroma Software System for Lattice QCD. *Nuclear Physics B - Proceedings Supplements*, 140:832–834, March 2005. ISSN 09205632. doi: 10.1016/j.nuclphysbps.2004.11.254. URL http://arxiv.org/abs/hep-lat/0409003. arXiv: hep-lat/0409003.

[21] Hilmar Forkel. A Primer on Instantons in QCD. *arXiv:hep-ph/0009136*, September 2000. URL http://arxiv.org/abs/hep-ph/0009136. arXiv: hep-ph/0009136.

[22] Message P Forum. MPI: A Message-Passing Interface Standard. Technical report, University of Tennessee, Knoxville, TN, USA, 1994.

[23] Edgar Gabriel, Graham E. Fagg, George Bosilca, Thara Angskun, Jack J. Dongarra, Jeffrey M. Squyres, Vishal Sahay, Prabhanjan Kambadur, Brian Barrett, Andrew Lumsdaine, Ralph H. Castain, David J. Daniel, Richard L. Graham, and Timothy S. Woodall. Open MPI: Goals, Concept, and Design of a Next Generation MPI Implementation. In *Proceedings, 11th European PVM/MPI Users' Group Meeting*, pages 97–104, Budapest, Hungary, September 2004.

[24] Christof Gattringer and Christian B. Lang. *Quantum Chromodynamics on the Lattice*, volume 788 of *Lecture Notes in Physics*. Springer Berlin Heidelberg, Berlin, Heidelberg, 2010. ISBN 978-3-642-01849-7 978-3-642-01850-3. doi: 10.1007/978-3-642-01850-3. URL http://link.springer.com/10.1007/978-3-642-01850-3.

[25] L. Giusti, G. C. Rossi, M. Testa, and G. Veneziano. The U_a(1) Problem on the Lattice with Ginsparg-Wilson Fermions. *Nuclear Physics B*, 628(1-2):234–252, April 2002. ISSN 05503213. doi: 10.1016/S0550-3213(02)00093-7. URL http://arxiv.org/abs/hep-lat/0108009. arXiv: hep-lat/0108009.

[26] Leonardo Giusti, Silvano Petrarca, and Bruno Taglienti. Theta dependence of the vacuum energy in the SU(3) gauge theory from the lattice. *Physical Review D*, 76 (9), November 2007. ISSN 1550-7998, 1550-2368. doi: 10.1103/PhysRevD.76.094510. URL http://arxiv.org/abs/0705.2352. arXiv: 0705.2352.

[27] Florian Gruber. *Topology in dynamical Lattice QCD simulations*. Universität Regensburg, 2013. URL https://epub.uni-regensburg.de/27631/.

[28] Marco Guagnelli, Rainer Sommer, and Hartmut Wittig. Precision computation of a low-energy reference scale in quenched lattice QCD. *Nuclear Physics B*, 535(1-2): 389–402, December 1998. ISSN 05503213. doi: 10.1016/S0550-3213(98)00599-9. URL http://linkinghub.elsevier.com/retrieve/pii/S0550321398005999.

[29] W. K. Hastings. Monte Carlo sampling methods using Markov chains and their applications. *Biometrika*, 57(1):97–109, April 1970. ISSN 1464-3510, 0006-3444. doi: 10.1093/biomet/57.1.97. URL https://academic.oup.com/biomet/article/57/1/97/284580.

[30] Marius Jonsson. Standard error estimation by an automated blocking method. *Physical Review E*, 98(4):043304, October 2018. doi: 10.1103/PhysRevE.98.043304. URL https://link.aps.org/doi/10.1103/PhysRevE.98.043304.

[31] Francesco Knechtli, Michael Günther, and Michael Peardon. *Lattice Quantum Chromodynamics*. SpringerBriefs in Physics. Springer Netherlands, Dordrecht, 2017. ISBN 978-94-024-0997-0 978-94-024-0999-4. doi: 10.1007/978-94-024-0999-4. URL http://link.springer.com/10.1007/978-94-024-0999-4.

[32] Wilhelm Kutta. Beitrag zur näherungweisen Integration totaler Differentialgleichungen. *Z. Math. Phys.*, 46:435–453, 1901.

[33] S. N. Lahiri. *Resampling Methods for Dependent Data*. Springer Series in Statistics. Springer New York, New York, NY, 2003. ISBN 978-1-4419-1848-2 978-1-4757-3803-2. doi: 10.1007/978-1-4757-3803-2. URL http://link.springer.com/10.1007/978-1-4757-3803-2.

[34] Tom Lancaster and Stephen Blundell. *Quantum field theory for the gifted amateur*. Oxford University Press, Oxford, first edition edition, 2014. ISBN 978-0-19-969933-9 978-0-19-969932-2. OCLC: ocn859651399.

[35] Irma Lavagnini and Franco Magno. A statistical overview on univariate calibration, inverse regression, and detection limits: Application to gas chromatography/mass spectrometry technique. *Mass Spectrometry Reviews*, 26(1):1–18, January 2007. ISSN 02777037, 10982787. doi: 10.1002/mas.20100. URL http://doi.wiley.com/10.1002/mas.20100.

[36] Frieder Lenz. Topological Concepts in Gauge Theories. *arXiv:hep-th/0403286*, March 2004. doi: 10.1007/978-3-540-31532-2_2. URL http://arxiv.org/abs/hep-th/0403286. arXiv: hep-th/0403286.

[37] G. Peter Lepage. Lattice QCD for Novices. *arXiv:hep-lat/0506036*, June 2005. URL http://arxiv.org/abs/hep-lat/0506036. arXiv: hep-lat/0506036.

[38] M. Lüscher and P. Weisz. On-shell improved lattice gauge theories. *Communications in Mathematical Physics*, 97(1):59–77, March 1985. ISSN 1432-0916. doi: 10.1007/BF01206178. URL https://doi.org/10.1007/BF01206178.

[39] Martin Lüscher. Schwarz-preconditioned HMC algorithm for two-flavour lattice QCD. *Computer Physics Communications*, 165(3):199–220, February 2005. ISSN 00104655. doi: 10.1016/j.cpc.2004.10.004. URL http://arxiv.org/abs/hep-lat/0409106. arXiv: hep-lat/0409106.

[40] Martin Lüscher. Properties and uses of the Wilson flow in lattice QCD. *Journal of High Energy Physics*, 2010(8), August 2010. ISSN 1029-8479. doi: 10.1007/JHEP08(2010)071. URL http://arxiv.org/abs/1006.4518. arXiv: 1006.4518.

[41] Martin Lüscher. Chiral symmetry and the Yang–Mills gradient flow. *Journal of High Energy Physics*, 2013(4), April 2013. ISSN 1029-8479. doi: 10.1007/JHEP04(2013)123. URL http://arxiv.org/abs/1302.5246. arXiv: 1302.5246.

[42] Martin Lüscher and Peter Weisz. Perturbative analysis of the gradient flow in non-abelian gauge theories. *Journal of High Energy Physics*, 2011(2), February 2011. ISSN 1029-8479. doi: 10.1007/JHEP02(2011)051. URL http://link.springer.com/10.1007/JHEP02(2011)051.

[43] Neal Madras and Alan D. Sokal. The pivot algorithm: A highly efficient Monte Carlo method for the self-avoiding walk. *Journal of Statistical Physics*, 50(1-2):109–186, January 1988. ISSN 0022-4715, 1572-9613. doi: 10.1007/BF01022990. URL http://link.springer.com/10.1007/BF01022990.

[44] Benjamin J. Menadue, Waseem Kamleh, Derek B. Leinweber, and M. S. Mahbub. Extracting Low-Lying Lambda Resonances Using Correlation Matrix Techniques. *arXiv:1102.3492 [hep-lat]*, pages 213–215, 2011. doi: 10.1063/1.3587609. URL http://arxiv.org/abs/1102.3492. arXiv: 1102.3492.

[45] Nicholas Metropolis, Arianna W. Rosenbluth, Marshall N. Rosenbluth, Augusta H. Teller, and Edward Teller. Equation of State Calculations by Fast Computing Machines. *The Journal of Chemical Physics*, 21(6):1087–1092, June 1953. ISSN 0021-9606, 1089-7690. doi: 10.1063/1.1699114. URL http://aip.scitation.org/doi/10.1063/1.1699114.

[46] Colin Morningstar and Mike Peardon. Analytic smearing of SU (3) link variables in lattice QCD. *Physical Review D*, 69(5), March 2004. ISSN 1550-7998, 1550-2368. doi: 10.1103/PhysRevD.69.054501. URL https://link.aps.org/doi/10.1103/PhysRevD.69.054501.

[47] Colin J. Morningstar and Mike Peardon. The glueball spectrum from an anisotropic lattice study. *Physical Review D*, 60(3), July 1999. ISSN 0556-2821, 1089-4918. doi: 10.1103/PhysRevD.60.034509. URL http://arxiv.org/abs/hep-lat/9901004. arXiv: hep-lat/9901004.

[48] Hans Munthe-Kaas. Runge-Kutta methods on Lie groups. *BIT Numerical Mathematics*, 38(1):92–111, March 1998. ISSN 0006-3835, 1572-9125. doi: 10.1007/BF02510919. URL http://link.springer.com/10.1007/BF02510919.

[49] Mikio Nakahara. *Geometry, topology, and physics*. Graduate student series in physics. Institute of Physics Publishing, Bristol ; Philadelphia, 2nd ed edition, 2003. ISBN 978-0-7503-0606-5. OCLC: ocm52622950.

[50] R. D. Peccei and Helen R. Quinn. CP Conservation in the Presence of Pseudoparticles. *Physical Review Letters*, 38(25):1440–1443, June 1977. ISSN 0031-9007. doi: 10.1103/PhysRevLett.38.1440. URL https://link.aps.org/doi/10.1103/PhysRevLett.38.1440.

[51] Michael E. Peskin and Daniel V. Schroeder. *An Introduction to quantum field theory*. Addison-Wesley, Reading, USA, 1995. ISBN 978-0-201-50397-5 0-201-50397-2. URL http://www.slac.stanford.edu/mpeskin/QFT.html.

[52] Dimitris N. Politis. The Impact of Bootstrap Methods on Time Series Analysis. *Statistical Science*, 18(2):219–230, May 2003. ISSN 0883-4237. doi: 10.1214/ss/1063994977. URL http://projecteuclid.org/euclid.ss/1063994977.

[53] V. Punjabi, C. F. Perdrisat, M. K. Jones, E. J. Brash, and C. E. Carlson. The Structure of the Nucleon: Elastic Electromagnetic Form Factors. *arXiv:1503.01452 [nucl-ex, physics:nucl-th]*, March 2015. URL http://arxiv.org/abs/1503.01452. arXiv: 1503.01452.

[54] R. Rajaraman. *Solitons and instantons: an introduction to solitons and instantons in quantum field theory*. North-Holland, Amsterdam, 4. impr; digital printing 2005 edition, 2005. ISBN 978-0-444-87047-6. OCLC: 255035301.

[55] M. Reuter. Renormalization of the Topological Charge in Yang-Mills Theory. *Modern Physics Letters A*, 12(36):2777–2802, November 1997. ISSN 0217-7323, 1793-6632. doi: 10.1142/S0217732397002922. URL http://arxiv.org/abs/hep-th/9604124. arXiv: hep-th/9604124.

[56] Carl Runge. Über die numerische Auflösung von Differentialgleichungen. *Mathematische Annalen*, 46(2):167–178, 1895.

[57] Stefan Schaefer, Rainer Sommer, and Francesco Virotta. Critical slowing down and error analysis in lattice QCD simulations. *Nuclear Physics B*, 845(1):93–119, April 2011. ISSN 05503213. doi: 10.1016/j.nuclphysb.2010.11.020. URL `http://arxiv.org/abs/1009.5228`. arXiv: 1009.5228.

[58] T. Schaefer and E. Shuryak. Instantons in QCD. *Reviews of Modern Physics*, 70(2): 323–425, April 1998. ISSN 0034-6861, 1539-0756. doi: 10.1103/RevModPhys.70.323. URL `http://arxiv.org/abs/hep-ph/9610451`. arXiv: hep-ph/9610451.

[59] P. Schmidt-Wellenburg. The quest to find an electric dipole moment of the neutron. *arXiv:1607.06609 [hep-ex, physics:nucl-ex, physics:physics]*, July 2016. URL `http://arxiv.org/abs/1607.06609`. arXiv: 1607.06609.

[60] R. Shankar. *Principles of Quantum Mechanics*. Springer US, Boston, MA, 1994. ISBN 978-1-4757-0578-2 978-1-4757-0576-8. doi: 10.1007/978-1-4757-0576-8. URL `http://link.springer.com/10.1007/978-1-4757-0576-8`.

[61] B. Sheikholeslami and R. Wohlert. Improved continuum limit lattice action for QCD with wilson fermions. *Nuclear Physics B*, 259(4):572–596, September 1985. ISSN 0550-3213. doi: 10.1016/0550-3213(85)90002-1. URL `http://www.sciencedirect.com/science/article/pii/0550321385900021`.

[62] R. Sommer. A New Way to Set the Energy Scale in Lattice Gauge Theories and its Application to the Static Force and α_s in SU(2) Yang–Mills Theory. *Nuclear Physics B*, 411(2-3):839–854, January 1994. ISSN 05503213. doi: 10.1016/0550-3213(94)90473-1. URL `http://arxiv.org/abs/hep-lat/9310022`. arXiv: hep-lat/9310022.

[63] F. Strocchi. Local and covariant gauge quantum field theories. Cluster property, superselection rules, and the infrared problem. *Physical Review D*, 17(8):2010–2021, April 1978. ISSN 0556-2821. doi: 10.1103/PhysRevD.17.2010. URL `https://link.aps.org/doi/10.1103/PhysRevD.17.2010`.

[64] K. Symanzik. Continuum limit and improved action in lattice theories: (I). Principles and phi4 theory. *Nuclear Physics B*, 226(1):187–204, September 1983. ISSN 0550-3213. doi: 10.1016/0550-3213(83)90468-6. URL `http://www.sciencedirect.com/science/article/pii/0550321383904686`.

[65] K. Symanzik. Continuum limit and improved action in lattice theories: (II). O(N) non-linear sigma model in perturbation theory. *Nuclear Physics B*, 226(1):205–227, September 1983. ISSN 0550-3213. doi: 10.1016/0550-3213(83)90469-8. URL `http://www.sciencedirect.com/science/article/pii/0550321383904698`.

[66] G. 't Hooft. Computation of the quantum effects due to a four-dimensional pseudoparticle. *Phys. Rev. D*, 14(12):3432–3450, December 1976. doi: 10.1103/PhysRevD.14.3432. URL `https://link.aps.org/doi/10.1103/PhysRevD.14.3432`.

[67] Gerard 't Hooft. Symmetry Breaking Through Bell-Jackiw Anomalies. *Phys. Rev. Lett.*, 37:8–11, 1976. doi: 10.1103/PhysRevLett.37.8.

[68] M. Tanabashi, K. Hagiwara, K. Hikasa, K. Nakamura, Y. Sumino, F. Takahashi, J. Tanaka, K. Agashe, G. Aielli, C. Amsler, M. Antonelli, D. M. Asner, H. Baer, Sw. Banerjee, R. M. Barnett, T. Basaglia, C. W. Bauer, J. J. Beatty, V. I. Belousov, J. Beringer, S. Bethke, A. Bettini, H. Bichsel, O. Biebel, K. M. Black, E. Blucher, O. Buchmuller, V. Burkert, M. A. Bychkov, R. N. Cahn, M. Carena, A. Ceccucci, A. Cerri, D. Chakraborty, M.-C. Chen, R. S. Chivukula, G. Cowan, O. Dahl, G. D'Ambrosio, T. Damour, D. de Florian, A. de Gouvêa, T. DeGrand, P. de Jong, G. Dissertori, B. A. Dobrescu, M. D'Onofrio, M. Doser, M. Drees, H. K. Dreiner, D. A. Dwyer, P. Eerola, S. Eidelman, J. Ellis, J. Erler, V. V. Ezhela, W. Fetscher, B. D. Fields, R. Firestone, B. Foster, A. Freitas, H. Gallagher, L. Garren, H.-J. Gerber, G. Gerbier, T. Gershon, Y. Gershtein, T. Gherghetta, A. A. Godizov, M. Goodman, C. Grab, A. V. Gritsan, C. Grojean, D. E. Groom, M. Grünewald, A. Gurtu, T. Gutsche, H. E. Haber, C. Hanhart, S. Hashimoto, Y. Hayato, K. G. Hayes, A. Hebecker, S. Heinemeyer, B. Heltsley, J. J. Hernández-Rey, J. Hisano, A. Höcker, J. Holder, A. Holtkamp, T. Hyodo, K. D. Irwin, K. F. Johnson, M. Kado, M. Karliner, U. F. Katz, S. R. Klein, E. Klempt, R. V. Kowalewski, F. Krauss, M. Kreps, B. Krusche, Yu. V. Kuyanov, Y. Kwon, O. Lahav, J. Laiho, J. Lesgourgues, A. Liddle, Z. Ligeti, C.-J. Lin, C. Lippmann, T. M. Liss, L. Littenberg, K. S. Lugovsky, S. B. Lugovsky, A. Lusiani, Y. Makida, F. Maltoni, T. Mannel, A. V. Manohar, W. J. Marciano, A. D. Martin, A. Masoni, J. Matthews, U.-G. Meißner, D. Milstead, R. E. Mitchell, K. Mönig, P. Molaro, F. Moortgat, M. Moskovic, H. Murayama, M. Narain, P. Nason, S. Navas, M. Neubert, P. Nevski, Y. Nir, K. A. Olive, S. Pagan Griso, J. Parsons, C. Patrignani, J. A. Peacock, M. Pennington, S. T. Petcov, V. A. Petrov, E. Pianori, A. Piepke, A. Pomarol, A. Quadt, J. Rademacker, G. Raffelt, B. N. Ratcliff, P. Richardson, A. Ringwald, S. Roesler, S. Rolli, A. Romaniouk, L. J. Rosenberg, J. L. Rosner, G. Rybka, R. A. Ryutin, C. T. Sachrajda, Y. Sakai, G. P. Salam, S. Sarkar, F. Sauli, O. Schneider, K. Scholberg, A. J. Schwartz, D. Scott, V. Sharma, S. R. Sharpe, T. Shutt, M. Silari, T. Sjöstrand, P. Skands, T. Skwarnicki, J. G. Smith, G. F. Smoot, S. Spanier, H. Spieler, C. Spiering, A. Stahl, S. L. Stone, T. Sumiyoshi, M. J. Syphers, K. Terashi, J. Terning, U. Thoma, R. S. Thorne, L. Tiator, M. Titov, N. P. Tkachenko, N. A. Törnqvist, D. R. Tovey, G. Valencia, R. Van de Water, N. Varelas, G. Venanzoni, L. Verde, M. G. Vincter, P. Vogel, A. Vogt, S. P. Wakely, W. Walkowiak, C. W. Walter, D. Wands, D. R. Ward, M. O. Wascko, G. Weiglein, D. H. Weinberg, E. J. Weinberg, M. White, L. R. Wiencke, S. Willocq, C. G. Wohl, J. Womersley, C. L. Woody, R. L. Workman, W.-M. Yao, G. P. Zeller, O. V. Zenin, R.-Y. Zhu, S.-L. Zhu, F. Zimmermann, P. A. Zyla, J. An-

derson, L. Fuller, V. S. Lugovsky, and P. Schaffner. Review of Particle Physics. *Phys. Rev. D*, 98(3):030001, August 2018. doi: 10.1103/PhysRevD.98.030001. URL https://link.aps.org/doi/10.1103/PhysRevD.98.030001.

[69] Stefan Vandoren and Peter van Nieuwenhuizen. Lectures on instantons. *arXiv:0802.1862 [hep-th]*, February 2008. URL http://arxiv.org/abs/0802.1862. arXiv: 0802.1862.

[70] A. I. Vaĭnshteĭn, Valentin I. Zakharov, Viktor A. Novikov, and Mikhail A. Shifman. ABC of instantons. *Soviet Physics Uspekhi*, 25(4):195, 1982. ISSN 0038-5670. doi: 10.1070/PU1982v025n04ABEH004533. URL http://iopscience.iop.org/article/10.1070/PU1982v025n04ABEH004533/meta.

[71] G. Veneziano. Goldstone mechanism from gluon dynamics. *Physics Letters B*, 95(1): 90–92, September 1980. ISSN 0370-2693. doi: 10.1016/0370-2693(80)90406-2. URL http://www.sciencedirect.com/science/article/pii/0370269380904062.

[72] Ettore Vicari and Haralambos Panagopoulos. Theta dependence of SU(N) gauge theories in the presence of a topological term. *Physics Reports*, 470(3-4):93–150, January 2009. ISSN 03701573. doi: 10.1016/j.physrep.2008.10.001. URL http://arxiv.org/abs/0803.1593. arXiv: 0803.1593.

[73] S. Weinberg. The U(1) problem. *Physical Review D*, 11(12):3583–3593, June 1975. ISSN 0556-2821. doi: 10.1103/PhysRevD.11.3583. URL https://link.aps.org/doi/10.1103/PhysRevD.11.3583.

[74] S. Weinberg. *The Quantum Theory of Fields*, volume 2 of *The Quantum Theory of Fields 3 Volume Hardback Set*. Cambridge University Press, 1995. ISBN 978-0-521-55002-4. URL https://books.google.no/books?id=sn9QvU5dmBQC.

[75] Scott Willenbrock. Symmetries of the Standard Model. *arXiv:hep-ph/0410370*, October 2004. URL http://arxiv.org/abs/hep-ph/0410370. arXiv: hep-ph/0410370.

[76] Kenneth G. Wilson. Confinement of quarks. *Physical Review D*, 10(8):2445–2459, October 1974. ISSN 0556-2821. doi: 10.1103/PhysRevD.10.2445. URL https://link.aps.org/doi/10.1103/PhysRevD.10.2445.

[77] Kenneth G. Wilson. Ab initio quantum chemistry: A source of ideas for lattice gauge theorists. *Nuclear Physics B - Proceedings Supplements*, 17:82–92, September 1990. ISSN 09205632. doi: 10.1016/0920-5632(90)90223-H. URL http://linkinghub.elsevier.com/retrieve/pii/092056329090223H.

[78] E. Witten. Current algebra theorems for the U(1) "Goldstone boson". *Nuclear Physics B*, 156(2):269–283, September 1979. ISSN 05503213. doi: 10.1016/0550-3213(79)90031-2. URL http://linkinghub.elsevier.com/retrieve/pii/0550321379900312.

[79] Ulli Wolff. Monte Carlo errors with less errors. *Computer Physics Communications*, 176(5):383, March 2007. ISSN 00104655. doi: 10.1016/j.cpc.2006.12.001. URL `http://arxiv.org/abs/hep-lat/0306017`. arXiv: hep-lat/0306017.

[80] C. Xavier and S.S. Iyengar. *Introduction to Parallel Algorithms*. A Wiley interscience publication. Wiley, 1998. ISBN 978-0-471-25182-8. URL `https://books.google.no/books?id=W3Ld65MnwgkC`.