BIELEFELD UNIVERSITY

DISSERTATION

Topological Aspects in Lattice QCD

Author:
Lukas MAZUR

Theoretical High Energy Physics Group
Faculty of Physics

April 22, 2021

Supervisor and 1st Referee:	Dr. Olaf KACZMAREK
2nd Referee:	Prof. Dr. Frithjof KARSCH

List of Publications

Part of the results of this thesis are included in the following papers, which are at the time of writing under peer review in the journal *Physical Review D*:

- Olaf Kaczmarek, Lukas Mazur, and Sayantan Sharma. "Eigenvalue spectra of QCD and the fate of $U_A(1)$ breaking towards the chiral limit". Feb. 2021. arXiv: 2102.06136 [hep-lat]

- Luis Altenkort, Alexander M. Eller, Olaf Kaczmarek, Lukas Mazur, Guy D. Moore, and Hai-Tao Shu. "The sphaleron rate from Euclidean lattice correlators: an exploration". Dec. 2020. arXiv: 2012.08279 [hep-lat]

Preliminary results have been presented at the following conferences and proceedings:

- Lukas Mazur, Luis Altenkort, Olaf Kaczmarek, and Hai-Tao Shu. "Euclidean correlation functions of the topological charge density". In: *PoS* LATTICE2019 (2020), 219. DOI: 10.22323/1.363.0219

- Lukas Mazur, Olaf Kaczmarek, Edwin Laermann, and Sayantan Sharma. "The fate of axial U(1) in 2+1 flavor QCD towards the chiral limit". In: *PoS* LATTICE2018 (2019), 153. DOI: 10.22323/1.334.0153

In addition the author of this thesis contributed to the following publications, which are not part of this thesis:

- Luis Altenkort, Alexander M. Eller, Olaf Kaczmarek, Lukas Mazur, Guy D. Moore, and Hai-Tao Shu. "Heavy quark momentum diffusion from the lattice using gradient flow". In: *Phys. Rev. D* 103.1 (2021), 014511. DOI: 10.1103/PhysRevD.103.014511

- Adrien Florio, Olaf Kaczmarek, and Lukas Mazur. "Open-Boundary Conditions in the Deconfined Phase". In: *Eur. Phys. J. C* 79.12 (2019), 1039. DOI: 10.1140/epjc/s10052-019-7564-z

- Luis Altenkort, Olaf Kaczmarek, Lukas Mazur, and Hai-Tao Shu. "Color-electric correlation functions under gradient flow". In: *PoS* LATTICE2019 (2019), 204. DOI: 10.22323/1.363.0204

- Yannis Burnier, Adrien Florio, Olaf Kaczmarek, and Lukas Mazur. "Thermal Simulations, Open Boundary Conditions and Switches". In: *EPJ Web Conf.* 175 (2018). Ed. by M. Della Morte, P. Fritzsch, E. Gámiz Sánchez, and C. Pena Ruano, 07004. DOI: 10.1051/epjconf/201817507004

Contents

List of Publications **3**

1 Introduction **7**

2 QCD and its Symmetries **11**
2.1 The QCD Lagrangian and the path integral formulation 11
2.2 Gauge symmetry 12
2.3 Chiral symmetry 13
2.4 The axial anomaly 14
2.5 QCD phase diagram in the chiral limit 15

3 Topology **19**
3.1 Compactifying spacetime 19
3.2 Classification of topological sectors 21
3.3 Instantons 23
3.4 Properties of the topological charge 25
3.5 The sphaleron rate 26

4 Lattice QCD **29**
4.1 Discretizing QCD 29
4.2 Fermion doublers and the Wilson discretization 33
4.3 The staggered discretization 34
4.4 The Ginsparg-Wilson relation and the overlap Dirac operator 36
4.5 The discretized field strength tensor 39
4.6 Topological charge 40
4.7 The gradient flow 42

5 The ParallelGPUCode **47**
5.1 Memory management 48
5.2 Indexing 49
5.3 Parallelizing using functor syntax 53
5.4 Communication techniques 54
5.5 SiteComm and derived classes 55
5.6 Lattice container 56
5.7 Coalesced memory access 57
5.8 Getting started 58
5.9 Status summary and benchmarks 59

6 Topological Susceptibility at Varying Light Quark Masses **63**
6.1 Lattice setup 63
6.2 Extracting the topological susceptibility 66
6.3 Volume dependence 66
6.4 Mass dependence 68

6.5 Temperature dependence 69
6.6 Discretization effects 70
6.7 Conclusion 70

7 The Overlap Dirac Eigenvalue Spectrum 73
7.1 Numerical setup 74
7.1.1 Accelerating the overlap Dirac eigenvalue measurements 75
7.2 Analyzing the overlap eigenvalue spectrum 76
7.3 Quantifying $\mathrm{U}_A(1)$ breaking in the chiral limit 79
7.4 Conclusion 83

8 The qq Correlation Function and the Sphaleron Rate 85
8.1 Numerical setup 86
8.2 Determining the topological charge density 87
8.3 Continuum and flow-time extrapolation 89
8.3.1 Interpolation 90
8.3.2 Continuum extrapolation 91
8.3.3 Flow-time extrapolation 92
8.4 The sphaleron rate 94
8.5 Conclusion 98

9 Conclusion 101

A Appendix 103
A.1 Gamma matrices 103
A.2 Pure gauge 104

Bibliography 111

Acknowledgements 125

Chapter 1

Introduction

Physics is a natural science that studies fundamental phenomena of nature and tries to explain their properties and behavior on the basis of quantitative models and laws. One of the most successful and most verified theories in physics is the Standard Model of particle physics. It incorporates three of the four fundamental interactions of nature through the framework of quantum field theory (QFT): The Electromagnetic, strong, and weak interactions. The strong interaction, which we deal with in this thesis, is described by quantum chromodynamics (QCD). It covers the dynamics of quarks and gluons, which are the constituents of hadrons such as the proton, neutron, or pion.

Quarks are fermions which carry spin and color charge. Today six different quark species are known which are referred to as *flavors*. Gluons on the other hand are gauge bosons which mediate strong interactions of QCD and carry themselves superpositions of color states. The gauge group of QCD is the non-abelian group $\mathrm{SU}(3)$, which implies that gauge fields do not compute. Hence, these fields interact with each other, leading to strong interactions between quarks. In fact, the interaction strength even increases with increasing distances. Consequently, if two quarks have been separated enough, it becomes energetically favored to create a quark-antiquark pair. This explains why under normal conditions no free quarks have been observed. They are always bound into color-neutral hadrons such as baryons or mesons. This phenomenon of QCD is called *confinement*. On the other hand, at high energies, quarks and gluons interact very weakly and are no longer bound inside hadrons. Hence, they behave like free particles leading to a quark gluon plasma (QGP). This phenomenon is called *asymptotic freedom*.

In this work we will investigate different topological aspects of QCD. Topology is a field of mathematics that has many important implications in modern physics. It deals with properties of mathematical structures which are preserved under continuous invertible deformations. Such continuous mappings are called *homeomorphisms*. A very popular example of a homeomorphism is the mapping between a mug and a donut (torus). The mug can be continuously reshaped to the form of a donut and vice versa, without losing the hole in the handle. Such structures carry a specific integer number called the *winding number*, which characterizes their topological properties.

In QCD, field configurations of different winding numbers exist which are separated by topological barriers. Between these configurations, tunneling effects occur. These effects are called *instantons* and carry a topological charge of $Q = 1$. An observable which is sensitive to the fluctuations of instantons is the topological susceptibility. At high temperatures, the topological susceptibility is expected to be proportional to $\chi_t = m_u m_d m_s / T^3$ [1], while at low temperatures below the phase transition, the topological susceptibility is expected to be proportional to $\chi_t \propto (1/m_u + 1/m_d + 1/m_s)^{-1}$ [2, 3]. Hence the expectation would be that in both regimes the susceptibility should vanish as the quark masses are driven to zero.

Instantons are more and more suppressed with increasing temperature. A dilute instanton gas approximation (DIGA) has shown that at high temperatures the susceptibility follows a power law dependence on temperature [1, 4]. However, at high temperatures topological transitions occur more often by jumping over the topological barrier due to thermal fluctuations. These transitions are called sphalerons and induce chiral imbalances in the quark gluon plasma. This for example may influence the chiral magnetic effect [5, 6]. Non-perturbative studies at finite temperatures on the $\mathrm{SU}(3)$ sphaleron rate are rather limited. A recent study can be found in [7].

Topology is related to the restoration and breaking of the chiral and axial symmetry. The behavior of topological observables is dictated by chiral symmetry breaking at low temperatures and dilute-instanton models at very high temperatures [8–10]. In general, the QCD Lagrangian has a $\mathrm{U}_L(2) \times \mathrm{U}_R(2)$ chiral symmetry for two massless flavors of quarks. However, the $\mathrm{SU}_V(2) \times \mathrm{SU}_A(2) \times \mathrm{U}_V(1)$ subgroup is spontaneously broken to $\mathrm{SU}_V(2) \times \mathrm{U}_V(1)$ in the hadronic phase, which leads to pions much lighter than nucleons. The subgroup $\mathrm{U}_A(1)$ is somewhat special. Classically it is a symmetry of the Lagrangian, but in a quantized theory it is broken by the measure of the path integral. This so-called *axial anomaly* [11–13] is believed to affect the nature of the chiral phase transition in QCD. The divergence of the Noether current of the axial symmetry in full quantum theory turns out to be equal to the topological charge density. It follows that any instanton field with topological charge Q must be accompanied by a corresponding change in axial charge as well as fermion zero modes in the background of these fields. Hence, there is a non-trivial connection between the axial anomaly and topology. Whether or not the anomalous $U_A(1)$ is effectively restored in the chiral limit is subject of current research [14–28].

At very high energies QCD can be treated perturbatively in an expansion in terms of the gauge coupling constant. At low energies, however, the coupling is strong and perturbation theory is not applicable. Instead, non-perturbative methods are required. A very successful non-perturbative approach which is used extensively in this work is lattice QCD. In this approach spacetime is discretized on a Euclidean hypercubic lattice with a finite lattice spacing. The quark fields are located on the sites of the lattice while the gauge fields are defined on the links connecting neighboring sites. Expectation values are then calculated using observables which are formulated in terms of lattice dimensions. Ultimately, a continuum extrapolation can be performed, which extrapolates the results to zero lattice spacing. Such a discretized theory can be computed numerically on a computer.

This is however a highly non-trivial task, since a typical lattice QCD calculation requires a lot of computational power. To get the best possible performance, these calculations must be implemented in such a way that they utilize all available capacities of the underlying hardware. Usually, lattice QCD calculations are carried out on large supercomputers. These machines are optimized for parallel computations on CPUs, GPUs or even both. One of the world's most powerful GPU based supercomputer at the time of writing this thesis is Summit at Oak Ridge National Laboratory, USA (see TOP500, November 2020 [29]). In this work, most of the calculations were performed on the Bielefeld GPU cluster [30]. Lattice QCD is very well suited for parallel computations, since the lattice can be domain decomposed and distributed across different processors. The computations on each sub-lattice are then performed on each processor in parallel. However, implementing parallel algorithms is quite challenging. Apart from the usual challenges which appear in sequential programming, one additionally has to deal with concepts such as inter-process communication, asynchronous computation or coalesced memory access. Moreover, CPUs and GPUs are being controlled through different programming languages. Also, processors from

different vendors use different interfaces to communicate with each other. Hence, if algorithms share similar patterns (as e.g. in lattice QCD) it is highly recommended to implement algorithms in a clean and structured fashion, such that basic functionalities can be used across different algorithms.

The outline of this thesis is as follows: In chapter 2 we will recap some basics of QCD and give an introduction to some important symmetries in QCD. Chapter 3 will introduce the concept of topology and draw the connection to QCD and the axial anomaly. An overview of the general formalism of lattice QCD is given in chapter 4. There we will also introduce some important lattice QCD discretizations such as the Zeuthen flow, the topological charge and the overlap Dirac operator, which will be needed in later chapters.

Chapter 5 is dedicated to the technical implementation of the ParallelGPUCode. It is a highly parallelized multi-GPU framework for lattice QCD calculations, which has mainly been developed within this work. We will discuss the general structure and design of this framework and give an introduction on how to implement custom routines. Additionally, we show results of some benchmarks which have been performed on the Bielefeld GPU cluster and at the GPU supercomputer Summit at Oak Ridge National Laboratory.

In chapter 6 we will measure the topological susceptibility on $2+1$ highly improved staggered quark (HISQ) [31, 32] ensembles around the pseudo-critical temperature T_c, which have a fixed physical strange quark mass but equal or lower-than-physical light quark masses. The gradient flow [33–36] will be used to smooth out UV fluctuations before measuring the topological charge. We will investigate the mass dependence as well as the volume and cut-off effects of the topological susceptibility on these configurations. Additionally, we will check the temperature dependence of the susceptibility in that temperature range.

In chapter 7 we will study the overlap Dirac eigenvalue spectrum [37, 38] on a subset of the same HISQ ensemble as in the previous chapter. The eigenvalue measurements in this chapter have been accelerated by using gradient flow as a noise reduction method. The results of the topological charge of the previous chapter have been used to predict the chiralities of the zero modes in the measurements. We will study the light quark mass dependence of the eigenvalue spectrum near T_c as it is reduced towards the chiral limit and show results for the renormalized observable $m_l^2(\chi_\pi - \chi_\delta)/T^4$, which is sensitive to $\mathrm{U}_A(1)$. Additionally, we perform a chiral extrapolation of this observable to see whether it survives in the chiral limit, which would indicate whether or not the $\mathrm{U}_A(1)$ is effectively restored at the chiral phase transition.

Chapter 8 will be dedicated to the topological charge density correlation function and the sphaleron rate [7, 39–41]. We measure the topological charge density correlation function on large, fine, pure gauge configurations at $1.5\,T_c$. We use the gradient flow as a noise reduction method to improve the signal of the correlation function. The sphaleron rate is then determined by analytical continuation of the Euclidean topological charge density two-point function. We perform a continuum extrapolation at fixed physical flow-time and then extrapolate the continuum results to zero flow-time. We use the extrapolated correlators to study the sphaleron rate by spectral reconstruction based on perturbatively motivated models.

Finally, in chapter 9 we conclude this thesis by summarizing all results that have been obtained throughout this work.

Chapter 2

QCD and its Symmetries

The Standard Model of particle physics describes particle physics as it stands today. It summarizes all known elementary particles and the important interactions between them through the framework of quantum field theory (QFT). The part of the model this thesis deals with is quantum chromodynamics (QCD). It is a non-abelian $\mathrm{SU}(3)$ gauge theory which describes the strong interactions between quarks and gluons, which are the fundamental building blocks of atomic nuclei. An important concept not only in QCD but in modern physics are symmetries. By Noether's Theorem, each continuous symmetry of the Lagrangian corresponds to a conserved current; by Goldstone's Theorem, each spontaneously broken generator of a continuous symmetry corresponds to a Goldstone boson. From this property it is possible to extract physical laws. In this chapter, we therefore introduce a few important symmetries as well as their special implications for QCD. Most of the topics in this chapter can also be found in textbooks [42–47].

2.1 The QCD Lagrangian and the path integral formulation

In lattice QCD we use importance sampling to select the gauge configurations in the Monte Carlo simulation (more on that in section 4.1). This requires the Boltzmann weight of the path integral to be purely real, which is only possible in Euclidean spacetime. Hence, throughout this thesis we are dealing with the Euclidean metric, if not stated otherwise, i.e. an analytic continuation to imaginary (Euclidean) time is performed via Wick rotation $t \to -i\tau$.

The Euclidean Lagrange density of QCD is given by [45]

$$\mathcal{L}_{\mathrm{QCD}} = \sum_f \bar{\psi}_f \left(\gamma_\mu D_\mu + m_f\right) \psi_f - \frac{1}{4g^2} F_{\mu\nu} F_{\mu\nu}, \tag{2.1}$$

where ψ_f are the fermionic fields, m_f the corresponding masses and $F_{\mu\nu}$ the field strength tensor. In general, the sum over f goes over all six presently known flavors u, d, s, c, b, t. Additionally, the quark fields hold a color- and spinor-index which are suppressed for simplicity. The covariant derivative D_μ is given by

$$D_\mu = \partial_\mu - iA_\mu, \tag{2.2}$$

with gluon fields A_μ which are 3×3 matrices living in the adjoint representation of the $SU(3)$ gauge group. In this work we follow the notation that the coupling constant g is absorbed by the gauge fields, i.e. $gA_\mu \to A_\mu$. The field strength tensor $F_{\mu\nu}$ can be expressed as

$$F_{\mu\nu}(x) = \partial_\mu A_\nu - \partial_\nu A_\mu - i[A_\mu A_\nu] \equiv F^a_{\mu\nu} t^a, \qquad t^a = \frac{\lambda^a}{2}. \tag{2.3}$$

Now, as the QCD Lagrangian is set, it is natural to search for a solution of its equations of motion. Unfortunately, an analytical solution has not been discovered yet. Therefore, other methods are required to gain more information from the theory. The path integral formalism turns out to be a great tool for estimating expectation values. The Euclidean path integral for evaluating expectation values is defined as

$$\langle O \rangle = \frac{1}{Z} \int \mathcal{D}\psi \mathcal{D}\bar{\psi} \mathcal{D}A \, \mathrm{e}^{-S_{\mathrm{QCD}}\left[\psi,\bar{\psi},A\right]} O\left[\psi,\bar{\psi},A_\mu\right], \tag{2.4}$$

where Z is the partition function

$$Z = \int \mathcal{D}\psi \mathcal{D}\bar{\psi} \mathcal{D}A \, \mathrm{e}^{-S_{\mathrm{QCD}}\left[\psi,\bar{\psi},A\right]}. \tag{2.5}$$

A Euclidean correlator is constructed in a similar manner,

$$\langle O_1(t) O_2(0) \rangle = \frac{1}{Z} \int \mathcal{D}\psi \mathcal{D}\bar{\psi} \mathcal{D}A \, \mathrm{e}^{-S_{\mathrm{QCD}}\left[\psi,\bar{\psi},A\right]} O_1\left[\psi,\bar{\psi},A_\mu\right] O_2\left[\psi,\bar{\psi},A_\mu\right]. \tag{2.6}$$

The action is obtained by

$$S_{\mathrm{QCD}}\left[\psi,\bar{\psi},A_\mu\right] = \int \mathrm{d}^4x \; \mathcal{L}_{\mathrm{QCD}}\left(\psi,\bar{\psi},A_\mu\right). \tag{2.7}$$

It can be shown, using the partition function from statistical mechanics [45], that the above expression corresponds to a system with zero temperature. Non-zero temperature is achieved by setting the integration interval of the temporal direction to be finite,

$$S_{\mathrm{QCD}}\left[\psi,\bar{\psi},A_\mu\right] = \int_0^{1/T} \mathrm{d}t \int \mathrm{d}^3x \; \mathcal{L}_{\mathrm{QCD}}\left(\psi,\bar{\psi},A_\mu\right), \tag{2.8}$$

where the inverse temporal extent equals the temperature.

2.2 Gauge symmetry

The most important symmetry in QCD is the local gauge symmetry. The fermion fields ψ and $\bar{\psi}$ transform under gauge transformation as

$$\psi'(x) = U(x)\psi(x), \qquad \bar{\psi}'(x) = \bar{\psi}(x)U^\dagger(x), \tag{2.9}$$

where $U(x)$ is a 3×3 special unitary matrix which mixes the different color components. The gauge fields A_μ, on the other hand, transform as

$$A'_\mu(x) = U(x)A_\mu(x)U^\dagger(x) + i\left(\partial_\mu U(x)\right)U^\dagger(x). \tag{2.10}$$

From this it follows that the transformation of the covariant derivative is given by

$$D'_\mu(x) = U(x)D_\mu(x)U^\dagger(x). \tag{2.11}$$

Making use of (2.10) and (2.11), we see that the fermionic term in the Lagrangian (2.1) is invariant under gauge transformation. The transformation of the field strength tensor in the second term of (2.1) can easily be derived by using (2.11). Similar to

the transformations above it changes as

$$F'_{\mu\nu}(x) = U(x)F_{\mu\nu}(x)U^{\dagger}(x). \tag{2.12}$$

It follows that also the gluonic part of the Lagrangian does not change under gauge transformations. Any physical result obtained from QCD should not depend on the gauge, in which it was calculated. Therefore, calculations can make use of the freedom of choosing a gauge to simplify the algebraic manipulations.

2.3 Chiral symmetry

The chiral symmetry is an approximate symmetry of the QCD Lagrangian. It is approximate because it is a symmetry only at zero quark masses, i.e. in the chiral limit. We know, however, that the masses of the u-quark, $m_u = 2.3(7)\mathrm{MeV}$, and d-quark, $m_d = 4.8(5)\mathrm{MeV}$ (in $\overline{\mathrm{MS}}$-scheme, $\mu = 2\mathrm{GeV}$) are much smaller than the typical scale of QCD $\sim 1\mathrm{GeV}$. Thus, they can be treated as small perturbations. In order to discuss this symmetry, we decompose the Dirac fermions into left-handed and right-handed parts,

$$\psi_L = P_L\psi, \qquad \psi_R = P_R\psi, \tag{2.13}$$

with the projection operators,

$$P_L = \frac{1}{2}\left(1 - \gamma_5\right), \qquad P_R = \frac{1}{2}\left(1 + \gamma_5\right), \tag{2.14}$$

which satisfy the necessary conditions $P_{L/R}^2 = P_{L/R}$, $P_L P_R = 0$ and $P_L + P_R = 1$. The Lagrange density can therefore be written in terms of these fields,

$$\mathcal{L}_{\mathrm{QCD}} = \bar{\psi}_L\gamma_\mu D_\mu\psi_L + \bar{\psi}_R\gamma_\mu D_\mu\psi_R + \bar{\psi}_R m\psi_L + \bar{\psi}_L m\psi_R - \frac{1}{4g^2}F_{\mu\nu}F_{\mu\nu}, \tag{2.15}$$

where the sum over the flavors is suppressed for simplicity. In the chiral limit it now follows that the left- and right-handed quarks completely decouple and behave as independent degrees of freedom,

$$\mathcal{L}_{\mathrm{QCD}} = \bar{\psi}_L\gamma_\mu D_\mu\psi_L + \bar{\psi}_R\gamma_\mu D_\mu\psi_R - \frac{1}{4g^2}F_{\mu\nu}F_{\mu\nu}. \tag{2.16}$$

Moreover, it is observed that the Lagrangian (2.16) has a global symmetry,

$$\begin{aligned} \psi'_L &= U_L\psi_L, \quad U_L \in \mathrm{SU}\left(N_f\right)_L, \\ \psi'_R &= U_R\psi_R, \quad U_R \in \mathrm{SU}\left(N_f\right)_R, \end{aligned} \tag{2.17}$$

where $U_{L/R}$ are special unitary $N_f \times N_f$ matrices which act in flavor space. Furthermore, the symmetries (2.17) contain two special symmetries: The isospin symmetry $\mathrm{U}(1)_V$ and the axial vector symmetry $\mathrm{U}(1)_A$. $\mathrm{U}(1)_V$ enters, when U_L and U_R are diagonal and equal. This symmetry even holds for degenerate masses and represents the quark number conservation in strong interactions. For the $\mathrm{U}(1)_A$ symmetry U_L and U_R are diagonal as well but represent rotations in the opposite directions. Altogether, the massless Lagrangian (2.16) is invariant under

$$\mathrm{SU}(N_f)_L \times \mathrm{SU}(N_f)_R \times \mathrm{U}(1)_V \times \mathrm{U}(1)_A. \tag{2.18}$$

This is equal to the more compact form $\mathrm{U}_L(N_f) \times \mathrm{U}_R(N_f)$, which is also used in this thesis. However, it turns out, if a full quantized theory is considered, the $\mathrm{U}(1)_A$ symmetry breaks in the measure of the path integral. This is the so called *axial anomaly*, which we discuss in the next section. Therefore, the true symmetry in the chiral limit is given by

$$\mathrm{SU}(N_f)_L \times \mathrm{SU}(N_f)_R \times \mathrm{U}(1)_V. \tag{2.19}$$

This group is sometimes written in the basis of the vector and axial flavor sub-group $\mathrm{SU}(N_f)_V \times \mathrm{SU}(N_f)_A \times \mathrm{U}(1)_V$.

2.4 The axial anomaly

As described in section 2.3 the QCD Lagrange density obeys a property called chiral symmetry when including massless fermions. It collectively describes several symmetries among rotations of multiple quark flavor. One such symmetry is the axial symmetry $U(1)_A$ which not only mixes the flavors but also the Dirac components. Its transformation can be represented as

$$\psi' = \mathrm{e}^{i\alpha\gamma_5}\psi, \qquad \bar{\psi}' = \bar{\psi}\mathrm{e}^{i\gamma_5\alpha} \quad \text{with } \alpha \in \mathbb{R}. \tag{2.20}$$

However, as already stated previously in section 2.3, this symmetry is broken in the quantized theory, because as one computes expectation values the measure of the path integral changes upon axial rotations. In particular under infinitesimal axial transformation, the measure of the path integral reads

$$\mathcal{D}[\psi, \bar{\psi}] = \mathcal{D}\left[\psi', \bar{\psi}'\right]\left(1 - 2\mathrm{i}\varepsilon N_f Q + \mathcal{O}\left(\varepsilon^2\right)\right), \tag{2.21}$$

where Q denotes the topological charge which we will discuss in more detail in the next chapter. Hence, expectation values are not invariant under this symmetry. Another way to view this is by computing the axial current. The axial current of the fermionic part in the Lagrangian reads [4, 48–50]

$$J^5_\mu = \bar{\psi}\gamma_\mu\gamma_5\psi, \tag{2.22}$$

which is classically conserved $\partial_\mu J^5 = 0$. Due to Noether's theorem it implies that the classical theory is invariant under $\mathrm{U}(1)_A$ transformations. However, at the quantum level, the four-divergence of the current eq. (2.22) becomes non-zero [4, 51],

$$\partial_\mu J^5_\mu = \frac{N_F}{16\pi^2} F^a_{\mu\nu}\tilde{F}^a_{\mu\nu}. \tag{2.23}$$

We will see in the next chapter that the right-hand side of this equation is the topological charge density.

To summarize, QCD is not invariant under axial rotation, even though the Lagrangian for massless quarks obeys the $\mathrm{U}_A(1)$ symmetry. The symmetry breaking comes from the measure of the path integral and is connected to the topology of the gauge field configurations. Physically, the axial anomaly implies that the chiral flavor singlet symmetry cannot be broken spontaneously. That explains why the η' meson cannot be considered a Goldstone boson and has almost twice the mass of the η meson.

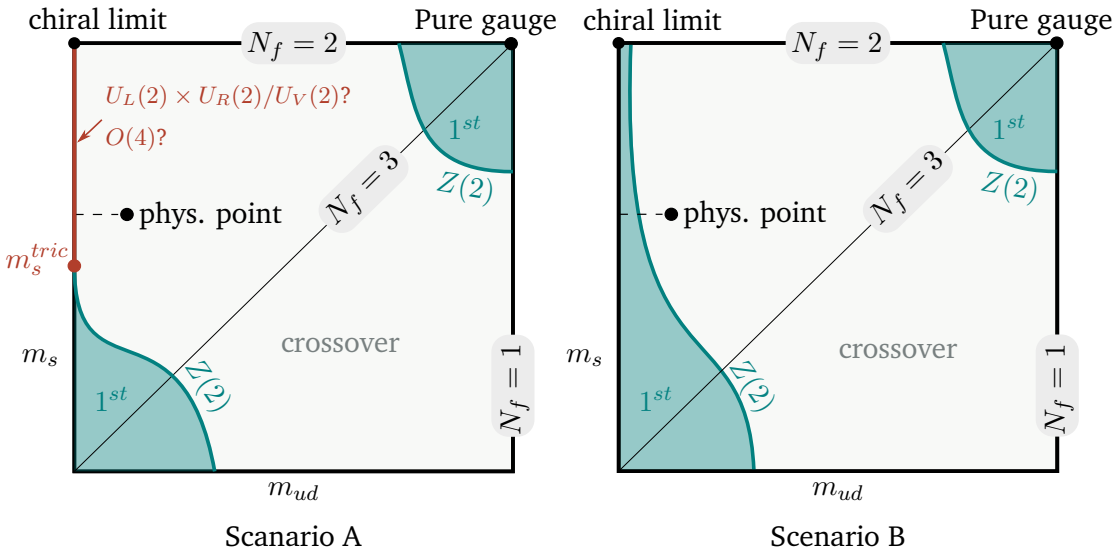


FIGURE 2.1: Two possible scenarios of the QCD phase diagram in the plane of the light quark masses m_{ud} and the strange quark mass m_s. *Left:* Phase diagram where the universality class of the phase transition in the chiral limit depends on the strange quark mass. For strange quark mass below the tricritical point the transition is of first order. Above the tricritical point it is of second order belonging either to the $\mathrm{O}(4)$ or the $\mathrm{U}_L(2) \times \mathrm{U}_R(2)/\mathrm{U}_V(2)$ universality class. *Right:* Phase diagram where the transition in the chiral limit is of first order. The first order region and the crossover region in both scenarios is separated by a line of second order transitions belonging to the $\mathrm{Z}(2)$ universality class.

2.5 QCD phase diagram in the chiral limit

In a chirally symmetric world the proton mass would be degenerate. However, the chiral symmetry is spontaneously broken at low temperatures, which leads to non-degenerate masses between a nucleon (≈ 940 MeV) and its negative parity partner, N^* (≈ 1535 MeV). Spontaneously broken symmetries are related to phase transitions. The chiral symmetry is of main interest when studying the phase diagram as a function of chemical potential and temperature. The order parameter of the chiral symmetry breaking is the *chiral condensate*, defined as

$$\langle \bar{\psi}_l \psi_l \rangle = \frac{T}{V} \frac{\partial \ln Z}{\partial m_l}, \tag{2.24}$$

where m_l is the mass of the light quarks. This term is not invariant under chiral rotations and transforms like a mass term. Hence, chiral symmetry is broken when the chiral condensate is non-zero.

Of particular interest is the order of the transition in the chiral limit, i.e. at vanishing light quark masses. In case it is of second order, observables close to the transition temperature can be described using *universal scaling* [52]. Universal scaling groups phase transitions into universality classes. The phase transitions of superficially different physical phenomena can fall into the same universality class, which means that certain divergent quantities will also diverge with the same critical exponents as one approaches the transition point. For example if the chiral transition in the continuum limit belongs to the $\mathrm{O}(4)$ universality class, that would be the same universality class as that of an $\mathrm{O}(4)$ spin system. Currently, the universality class of

the chiral phase transition is however unclear, which prevents us in doing any further scaling analysis. Figure 2.1 shows two possible scenarios on the plane of the strange quark mass and the light quark mass. In both scenarios the phase transition at infinite mass limit as well as the 3-flavor chiral limit is of first order [53, 54]. For the two-flavor chiral limit with an additional strange quark the situation is unclear. In scenario A, the region of first order transition at low strange quark masses end in a tricritical point. In scenario B it continues to two-flavor QCD. In both cases the physical point is located within the crossover region. The region above the tricritical point in scenario A belongs to the $\mathrm{O}(4)$ or $\mathrm{U}_L(2) \times \mathrm{U}_R(2)/\mathrm{U}_V(2)$ universality class which leads to a second order chiral phase transition. The position of the tricritical point in scenario A determines which universality class is more relevant for the physical point. In case it is lying below the physical point, the relevant universality class is either $\mathrm{O}(4)$ or $\mathrm{U}_L(2) \times \mathrm{U}_R(2)/\mathrm{U}_V(2)$. If it is lying above the physical point, remnants of the $\mathrm{Z}(2)$ symmetry become more important.

Due to a finite light quark mass m_l, the singular part of the free energy which carries information about the universal critical behavior gets overwhelmed by the regular part which is an analytic function in m_l^2 [55]. Thus there is no phase transition characterizing chiral symmetry restoration in QCD with physical light quark masses but a smooth crossover [56–62] which should go over to an exact phase transition only in the chiral limit. However, recent lattice studies have revealed signatures of $\mathrm{O}(4)$ scaling in chiral observables as one lowers the light quark masses towards the chiral limit along the line of constant physical strange quark mass [63]. Within the current precision, it is possible to rule out $Z(2)$ scaling [63] (although to establish these results one ultimately needs to perform a continuum extrapolation).

In order to figure out which scenario is taking place, the restoration of the anomalous $\mathrm{U}_A(1)$ symmetry provides a useful tool. The $\mathrm{U}_A(1)$, though not an exact symmetry, is believed to affect the nature of the phase transition of QCD with two degenerate light quark favours. From the renormalization group studies of model quantum field theories with same symmetries as QCD with two degenerate massless quark flavors, it is known that the existence of a critical point at vanishingly small baryon density depends crucially on the magnitude of the $\mathrm{U}_A(1)$ anomaly breaking near the chiral symmetry restoration temperature [54]. If the magnitude of the $\mathrm{U}_A(1)$ breaking term is comparable to its zero temperature value even at T_c, then the phase transition is of second order with $\mathrm{O}(4)$ critical exponents [54, 64–66]. If on the other hand the $\mathrm{U}_A(1)$ is approximately restored as the chiral symmetry restoration occurs, then the phase transition from the hadron phase to the quark-gluon plasma phase is expected to be either of first order [54, 64] or second order of $\mathrm{U}_L(2) \times \mathrm{U}_R(2)/\mathrm{U}_V(2)$ universality class [65, 67].

In such model quantum field theories however, the coefficient of the $\mathrm{U}_A(1)$ breaking term is a parameter whose magnitude can only be estimated from nonperturbative studies of QCD. Currently, lattice regularization is the most practical method which can provide a reliable answer to such a question.

In the limit of two massless flavors, it is not possible to use a local expectation value such as the chiral condensate to investigate the anomalous $\mathrm{U}_A(1)$ symmetry. Therefore, whether or not the anomalous $\mathrm{U}_A(1)$ is broken or not has to be estimated from other quantities. It turns out that the difference of two meson states indicates whether a certain symmetry is broken or not. Some examples of scalar and pseudo-scalar mesons which are related by symmetry transformations are illustrated in figure 2.2. Thus, to study the restoration of the $\mathrm{U}_A(1)$ symmetry one may look at the susceptibilities of the pion and delta meson. These mesons are built from up and down quarks. They

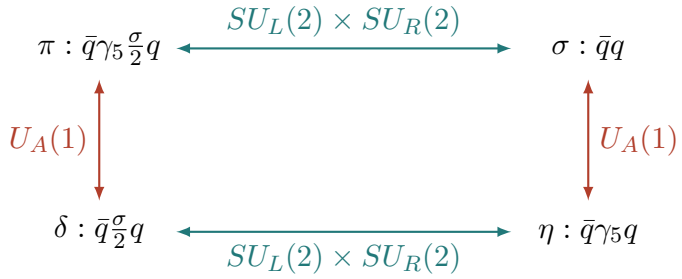


FIGURE 2.2: Illustration of scalar and pseudo-scalar mesons which are related by symmetry transformations. If a symmetry is restored then the corresponding states connected with arrows will become degenerate.

are described by the (local) operators [45, 68]

$$\delta^i(x) = \frac{1}{\sqrt{2}}\bar{\psi}(x)\sigma^i\psi(x), \tag{2.25}$$

$$\pi^i(x) = \frac{i}{\sqrt{2}}\bar{\psi}(x)\sigma^i\gamma_5\psi(x), \tag{2.26}$$

where σ^i is a Pauli matrix (see eq. (A.5)). It acts on the up and down flavor components of ψ, for example, $\delta^3(x) = (\bar{u}(x)u(x) - \bar{d}(x)d(x))/\sqrt{2}$. Using a $\mathrm{U}_A(1)$ rotation with $\alpha = \pi/4$ the pion can be transformed into a delta meson [68]

$$\begin{aligned}\delta^i(x) = \frac{1}{\sqrt{2}}\bar{\psi}(x)\sigma^i\psi(x) &\to \frac{1}{\sqrt{2}}\bar{\psi}(x)\mathrm{e}^{i\alpha\gamma_5}\sigma^i\mathrm{e}^{i\alpha\gamma_5}\psi(x)\\ &= \frac{1}{2\sqrt{2}}\bar{\psi}(x)\left(1+i\gamma_5\right)\sigma^i\left(1+i\gamma_5\right)\psi(x)\\ &= \frac{i}{\sqrt{2}}\bar{\psi}(x)\sigma^i\gamma_5\psi(x) = \pi^i(x).\end{aligned} \tag{2.27}$$

Hence, if the $\mathrm{U}_A(1)$ symmetry is restored the corresponding correlators should be degenerate. By integrating the correlation function of these operators we obtain the susceptibilities. The difference of the isospin-triplet pion ($\bar{\psi}\sigma_2\gamma_5\psi$) and delta ($\bar{\psi}\sigma_2\psi$) meson susceptibilities is then given by [69]

$$\chi_\pi - \chi_\delta = \int \mathrm{d}^4x \left[\langle i\pi^+(x)i\pi^-(0)\rangle - \langle\delta^+(x)\delta^-(0)\rangle\right]. \tag{2.28}$$

This can be written in terms of the density $\rho(\lambda, m_f)$ of the Dirac operator [70]

$$\chi_\pi - \chi_\delta \overset{V\to\infty}{\longrightarrow} \int_0^\infty \mathrm{d}\lambda \frac{4m_f^2\rho(\lambda, m_f)}{(\lambda^2+m_f^2)^2}. \tag{2.29}$$

In this work we will compute the Dirac density $\rho(\lambda, m_f)$ using lattice QCD methods.

From eq. (2.29) it is clear that whether or not the anomalous $\mathrm{U}_A(1)$ is effectively restored or broken depends also on Dirac eigenvalue density. There are many possible scenarios how $\rho(\lambda, m_f)$ affects the correlation functions. For instance if the eigenvalue spectrum approaches the trivial limit $\lim_{m_f\to 0}\rho(0, m_f) \to 0$, the $\mathrm{U}_A(1)$ would be effectively restored. From studies of up to six-point correlation functions in the scalar and pseudo-scalar channels [71], another sufficient condition for the effective restoration of $\mathrm{U}_A(1)$ in $N_f = 2$ QCD is that the eigenvalue spectrum behaves as

$\lim_{m_f \to 0} \rho(\lambda, m_f) \sim \lambda^3$. However, it has been shown in studies of up to 2-point correlation functions [15, 16] that when chiral symmetry is restored, the $\mathrm{U}_A(1)$ can still be broken if the infrared part of eigenvalue density goes as $\lim_{\lambda \to 0} \rho(\lambda, m_f) = \delta(\lambda) m_f^\alpha$ with $1 < \alpha < 2$

In previous studies the Dirac density has been investigated at almost physical quark masses [68, 72]. The infrared part of $\rho(\lambda, m_f)$ had both a non-analytic contribution which forms a near-zero peak and an analytic contributions in λ which forms the bulk eigenvalue spectrum. Also, it was shown that $\rho(\lambda, m_f) \sim \lambda^2$ at $T = 1.2 T_c$. Following the above statements this lead to $\mathrm{U}_A(1)$ still being broken above T_c. In this work, we want to see if this analytic dependence survives in the chiral limit. Also we want to check if the non-analytic part survives in the chiral limit.

Chapter 3

Topology

Topology studies the general properties of mathematical objects or spaces, which are invariant under continuous transformations. We call two topological objects equivalent if they are homeomorphic, i.e. there exists a continuous, invertible mapping between them. In this work we are working on Euclidean spaces. In general, any metric space such as the Euclidean space is a topological space which allows defining continuous deformation of subspaces. Hence, in what follows we introduce the concept of topology in the context of QCD in addition to some useful observables which are connected to topology. Most of the topics in this chapter can also be found for example in the literature [4, 51, 73–77].

3.1 Compactifying spacetime

In order to understand the topological properties of the QCD vacuum, we need to find those classical fields which minimize the Euclidean Yang-Mills action. In the literature such fields are often referred to as *classical vacua*. The Euclidean Yang-Mills action reads

$$S = \frac{1}{4g^2} \int \mathrm{d}^4x \, F^a_{\mu\nu} F^a_{\mu\nu} \tag{3.1}$$

where the gluon field strength tensor $F^a_{\mu\nu}$ is defined in eq. (2.3) and g is the coupling constant. Contrary to the Minkowskian Yang-Mills action the Euclidean action eq. (3.1) is non-negative. Therefore, to obtain its absolute minima we need to find the gluon fields which yield zero action. These can be obtained by requiring the gluon field strength tensor to be

$$F_{\mu\nu} = 0. \tag{3.2}$$

One way to achieve that is by setting the gauge potential configuration

$$A_\mu = 0. \tag{3.3}$$

However, this does not capture all possibilities which lead to zero action. The transformation law for $\mathrm{SU}(3)$ gauge theories is given by:

$$A_\mu \to A'_\mu = U A_\mu U^\dagger - iU\partial_\mu U^\dagger. \tag{3.4}$$

with $U \in \mathrm{SU}(3)$. Setting $A_\mu = 0$ in (3.4) leads to:

$$A^{pg}_\mu = -iU\partial_\mu U^\dagger. \tag{3.5}$$

These fields are the *pure gauges*. It can be shown explicitly that using that definition of the gauge potential (3.5) the field strength tensor vanishes (see appendix A.2). At first glance, it seems that these fields A^{pg}_μ form a unique QCD ground state. However,

by taking a closer look, we find that these fields consist of an enumerable infinity of topological equivalence classes. This requires a careful investigation of the gauge transformations.

In order to see this, let us start by choosing the temporal gauge

$$A_0 = 0. \tag{3.6}$$

The gauge transformation should conserve this condition, so that they are also time-independent (such transformations are also called residual gauge transformations). Moreover, for the purpose of the following topological (homotopy) classification it is sufficient to only consider those gauge transformations that become a constant (chosen to be unity) at spatial infinity [4, 51, 74], i.e.

$$U(\vec{x}) \to 1 \quad \text{for} \quad |\vec{x}| \to \infty. \tag{3.7}$$

This restriction guarantees that the gauge fields satisfy definite boundary conditions at the surface of a large box (which should not affect the local physics inside).

This condition effectively means that the range where $\vec{x}$ is defined is now $\mathbb{R}^3 \cup \{\infty\}$. With the inclusion of the point at infinity we can treat that range as an element of S^3. The reason for this is that $U(\vec{x})$ evaluated at boundary points $\vec{x}$ with $|\vec{x}| \to \infty$ can not be distinguished, and thus can be identified. This can be easier understood in one dimension [78]. The x-axis defines $\mathbb{R}$ and as $x \to \pm\infty$, a map $f(x)$ becomes constant. Therefore, these points x which give the same $f(x)$ in that boundary region can be treated as one point at infinity. The x-axis can thus be deformed into a circle S^1 which includes a point at infinity. Figure 3.1 illustrates that concept.

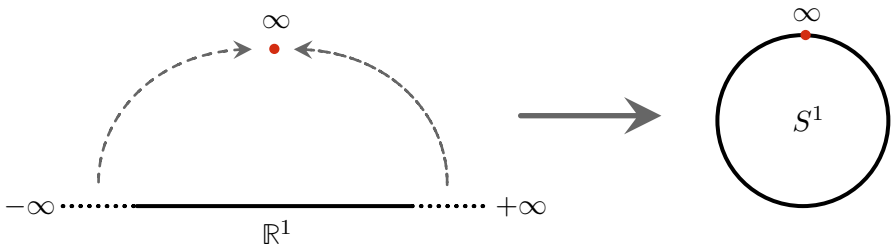


FIGURE 3.1: Illustration of compactifying $\mathbb{R}^1 \cup \{\infty\}$ on S^1

Consequently, the residual gauge transformations $U(\vec{x})$ define a map

$$S^3 \to \mathrm{SU}(3). \tag{3.8}$$

According to a theorem by Raoul Bott [79] the *topologically active* part of a Lie group which contains $\mathrm{SU}(2)$ as a subgroup is only that of the $\mathrm{SU}(2)$ subgroup. Thus, it is sufficient to only consider the transformations

$$S^3 \to \mathrm{SU}(2). \tag{3.9}$$

In $\mathrm{SU}(2)$ any element can be represented by the exponential

$$U = \exp(i\alpha\hat{r}\vec{\sigma}), \tag{3.10}$$

with an angle $\alpha \in \mathbb{R}$, the Pauli matrices $\vec{\sigma} = (\sigma_1, \sigma_2, \sigma_3)$ (see equation (A.5) for an explicit representation), and a unit vector $\hat{r}$. Rewriting that in terms of $\sin$ and $\cos$

yields

$$U = \cos(\alpha)\mathbb{1} + i\hat{r}\vec{\sigma}\sin(\alpha). \tag{3.11}$$

From this equation we can read off the quaternion representation of any $U \in \mathrm{SU}(2)$ by defining $u_0 \equiv \cos(\alpha)$ and $\vec{u} \equiv \sin(\alpha)\hat{r}$:

$$U = u_0\mathbb{1} + i\vec{u}\vec{\sigma} \qquad \text{with} \quad u_0^2 + \vec{u}^2 = 1. \tag{3.12}$$

This shows that also the group $\mathrm{SU}(2)$ can be mapped onto the 3-sphere S^3. Therefore, in the sense of homotopy theory, the map (3.7) is topologically equivalent to

$$S^3_{spatial} \to S^3_{group}. \tag{3.13}$$

i.e. the residual gauge transformations can be viewed as maps from a spatial 3-sphere into a group 3-sphere.

3.2 Classification of topological sectors

A crucial point now is that in general two such transformations $U \in \mathrm{SU}(2)$ can not be continuously deformed into each other. This can be seen by making the angle in equation (3.11) x-dependent:

$$U^{(n)}(x) = \cos(\alpha_n(x))\mathbb{1} + i\hat{r}\vec{\sigma}\sin(\alpha_n(x)). \tag{3.14}$$

where n is an integer. The condition that $U(x) \to 1$ at infinity (see eq. (3.7)) requires $\alpha_n(x)$ to reach a value which is periodic in n (since $\sin(2\pi n) = 0$ and $\cos(2\pi n) = 1$). Therefore, at $|x| \to \infty$, the angle has to become $\alpha_n(x) \to 2\pi n$. A smooth deformation between two transformations $U^{(n)}$ and $U^{(m)}$, i.e. with different parameter $n \neq m$, is only possible if their angles $\alpha_n(x)$ and $\alpha_m(x)$ deform smoothly at all x. However, that can not be fulfilled at $|x| \to \infty$, since there the restriction $\alpha_n(x) \to 2\pi n$ allows only discrete deformations. This is illustrated in figure 3.2. Hence, elements of $\mathrm{SU}(2)$

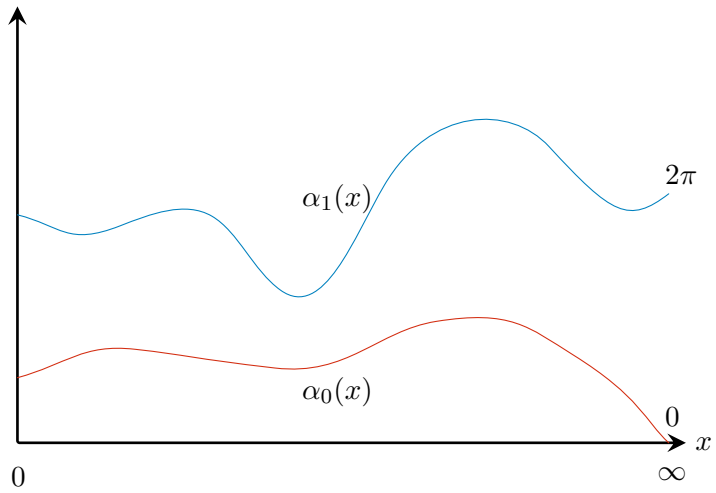


FIGURE 3.2: Illustration of two transformations. The endpoints have to fulfill $\lim_{x\to\infty}\alpha_n(x) = 2\pi n$. Therefore no smooth deformation between $\alpha_0(x)$ and $\alpha_1(x)$ possible.

with the restriction of equation (3.7) are topologically distinguishable and can be classified by the integer number n. This number also states how many elements from

the domain space $x \in \mathrm{SU}(2)$ are mapped to the same element in the target space $U(x) \in \mathrm{SU}(2)$. In order to get a visual understanding of that [80], let us simplify the problem by just considering the mapping

$$S^1 \to S^1, \qquad \mathrm{e}^{i\alpha} \to \mathrm{e}^{i\phi(\alpha)} \tag{3.15}$$

i.e. maps between two circles with the angle $\alpha \in [0, 2\pi]$. Note, that these phase fields can as well be divided into various topological sectors labeled by an integer n. Thus, that mapping can be represented by

$$U^{(n)}(\alpha) = \mathrm{e}^{in\alpha} \in \mathrm{U}(1). \tag{3.16}$$

Now consider the topological sector $n = 2$:

$$U^{(2)}(\alpha) = \mathrm{e}^{i2\alpha}. \tag{3.17}$$

This yields for two different angles of the domain space the same element in the target space:

$$U^{(2)}(\pi/2) = \mathrm{e}^{i\pi} = U^{(2)}(3\pi/2) = \mathrm{e}^{i2\pi}\mathrm{e}^{i1\pi} = -1 \tag{3.18}$$

Similar results can be observed for any integer n. In that sense, if we walk around the domain circle, the integer n counts how often we spot an angle that leads to the same target element $U \in \mathrm{U}(1)$. This integer is called the *winding number*. An illustration of that is shown in figure 3.3.

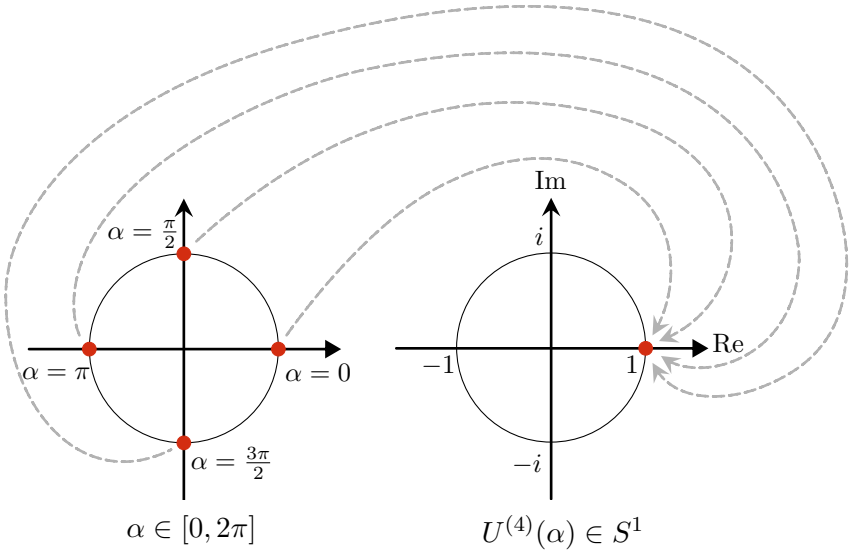


FIGURE 3.3: Illustration of a one dimensional mapping with winding number $n = 4$.

We have made it clear that maps of that kind (3.15) can be characterized by their topological properties. To be more specific, the maps (3.15) fall into an enumerable infinity of disjoint classes classified by an integer winding number $n \in \mathbb{Z}$. In that context it is useful to consider the notion of *homotopy*. Homotopy classifies maps between topological spaces. It builds up an equivalence class, while different equivalence classes form the so-called *homotopy group* $\pi_d(S^d) = \mathbb{Z}$. d denotes the dimension of the domain sphere, while the argument indicates the target space. Thus, the mapping (3.15) is in the homotopy group

$$\pi_1(S^1) = \mathbb{Z} \tag{3.19}$$

Since elements of $\mathrm{SU}(2)$ can be as well compactified into maps between spheres (3.13), we conclude that these elements are in the homotopy group

$$\pi_3(S^3) = \mathbb{Z}. \tag{3.20}$$

For a gauge theory in 4 dimensions with $d = 3$ such as $\mathrm{SU}(2)$, we can also define a winding number. It is the so-called *Cartan-Maurer integral invariant* [81, 82]

$$n_w \equiv \frac{1}{24\pi^2} \int \mathrm{d}^3x\, \epsilon_{ijk}\, \mathrm{tr}\left[(U\,\partial_i U^\dagger)(U\,\partial_j U^\dagger)(U\,\partial_k U^\dagger)\right]. \tag{3.21}$$

Rewriting that equation in terms of the gauge field we obtain the *3-d Chern-Simons number*

$$n_{CS} = \int \mathrm{d}^3x\, J_0^{CS} = \frac{1}{16\pi^2} \int \mathrm{d}^3x\, \epsilon_{ijk} \left(A_i^a \partial_j A_k^a + \frac{1}{3} f^{abc} A_i^a A_j^b A_k^c \right) \tag{3.22}$$

where J_μ^{CS} is the Chern-Simons current

$$J_\mu^{CS} = \frac{1}{16\pi^2} \epsilon_{\mu\alpha\beta\gamma} \left(A_\alpha^a \partial_\beta A_\gamma^a + \frac{1}{3} \epsilon^{abc} A_\alpha^a A_\beta^b A_\gamma^c \right). \tag{3.23}$$

Let us briefly summarize what we have shown so far: Pure gauge fields (3.5) form the QCD vacuum. They are constructed from gauge transformations $U(x)$ which satisfy the boundary conditions (3.7). The transformations fall into disjoint homotopy classes characterized by an integer winding number n, and thus the pure gauge field configurations inherit that property.

3.3 Instantons

So far we have shown that it is not possible to perform a continuous deformation between two transformations $U(x) \in \mathrm{SU}(2)$ of different winding number. However, if we consider the gauge fields $A_\mu(x)$ that is not the case. It is actually possible to transform a pure-gauge field with winding number n by continuous deformation into one of winding number $m \neq n$. However, it is only possible if the shape of the fields change as well. In other words, a continuous deformation between the fields $A^{(n)}$ and $A^{(m)}$ can only be realized by leaving pure gauge, i.e. the fields can not stay always in the form of eq. (3.5). Consequently, we encounter fields with non-zero action while transforming (recall that only pure gauge field configurations yield zero action).

To make that more clear, consider a pure-gauge classical vacuum in the $n = 1$ topological sector, i.e. $A_i^{(1)}$. As in the previous section, let us work in the gauge $A_0 = 0$. A continuous deformation of that vacuum into one of winding number $n = 0$ can be described by [75]

$$A_i^{(\beta)}(x) = \beta A_i^{(1)}(x) \tag{3.24}$$

where β is a real parameter $0 \leq \beta \leq 1$. For $\beta = 0$ we get the pure gauge field $A_i = 0$ with winding number $n = 0$, while for $\beta = 1$, we get the pure gauge field $A_i = A_i^{(1)}$ with winding number $n = 1$. In both cases, the classical energy vanishes, since the field strength tensor of these fields is zero. However, in the transition region, i.e. for $0 < \beta < 1$, the fields $A_i^{(\beta)}(x)$ are not in pure gauge. The electric part of the field strength tensor (i.e. F^{0i}) is still zero, because $\dot{A}_0^{(\beta)} = 0$ and $A_i^{(\beta)}$ is time-independent. The magnetic part (i.e. $B_i = \frac{1}{2}\varepsilon_{ijk}F_{jk}$) on the other hand does not vanish. This can be

seen by calculating the field strength tensor from eq. (3.24):

$$F_{jk}^{(\beta)} = \beta(\partial_j A_k^{(1)} - \partial_k A_j^{(1)}) - i\beta^2[A_j^{(1)}, A_k^{(1)}] \tag{3.25}$$

$$= i(\beta - \beta^2)[A_j^{(1)}, A_k^{(1)}] \tag{3.26}$$

$$\neq 0, \qquad \text{for } 0 < \beta < 1, \tag{3.27}$$

where in the first line we make use of the fact that $F_{\mu\nu}^{(1)} = 0$. This shows, that the energy which is proportional to $\int \mathrm{tr}(F_{jk}F_{jk})d^3x$ is non-zero. Moreover, it is not only non-zero but also finite, since $A_i^{(\beta)}$ vanishes sufficiently fast at the boundaries. The fields $A_i^{(\beta)}$ can now be viewed as a curve in field space as β varies from 0 to 1. This curve connects two topological sectors which correspond to absolute minima of the action. In between these minima the curve exbibits non-zero but finite action barriers: the so-called *sphaleron barrier*. This example can trivially be extended to any two topological sectors.

Since we are dealing with quantum theory, we might expect that quantum tunneling occurs through that barrier. A tunneling path in gauge theory which connects different classical vacua can be found by looking at classical solutions of the Euclidean equation of motion. The best path is the solution with minimal Euclidean action connecting vacua with different Chern-Simons Number. Such a solution can be found by writing the action in terms of the dual field strength tensor $\tilde{F}_{\mu\nu} = 1/2\epsilon_{\rho\sigma\mu\nu}F_{\rho\sigma}$. Note that $F_{\mu\nu}F_{\mu\nu} = \tilde{F}_{\mu\nu}\tilde{F}_{\mu\nu}$. Thus, we get the identity [4]

$$\frac{1}{2}\,\mathrm{tr}\left(\tilde{F}_{\mu\nu} \mp F_{\mu\nu}\right)^2 = \mathrm{tr}\, F_{\mu\nu}F_{\mu\nu} \pm \mathrm{tr}\,\tilde{F}_{\mu\nu}F_{\mu\nu}. \tag{3.28}$$

Since the left hand side is always non-negative it follows that

$$\int \mathrm{d}^4x\, \mathrm{tr}\, F_{\mu\nu}F_{\mu\nu} \geq \left|\int \mathrm{d}^4x\, \mathrm{tr}\,\tilde{F}_{\mu\nu}F_{\mu\nu}\right|. \tag{3.29}$$

Making use of this relation in the action (3.1) we then obtain the so called *Bogomolny bound*

$$S = \frac{1}{4g^2}\int \mathrm{d}^4x\, F_{\mu\nu}^a F_{\mu\nu}^a \geq \frac{1}{4g^2}\left|\int \mathrm{d}^4x\, \mathrm{tr}\,\tilde{F}_{\mu\nu}F_{\mu\nu}\right| = \frac{8\pi^2|Q|}{g^2}, \tag{3.30}$$

where the *topological charge* Q is defined as the integral over the *topological charge density* $q(x)$

$$Q = \int \mathrm{d}^4x\, q(x), \qquad q(x) = \frac{1}{16\pi^2}\,\mathrm{tr}\,\tilde{F}_{\mu\nu}F_{\mu\nu}. \tag{3.31}$$

It is now clear that the action is minimal if the field strength tensor is (anti-) self-dual

$$F_{\mu\nu}^a = \pm\tilde{F}_{\mu\nu}^a. \tag{3.32}$$

Note that this self-duality condition implies the Euclidean equation of motion (the reverse is not true)

$$\frac{\delta S}{\delta A_\nu} = D_\mu F_{\mu\nu} = 0. \tag{3.33}$$

However, using the self-duality equation (3.32) is more convenient, since in contrast to the equation of motion it is a first order differential equation.

By solving the self-duality equation (3.32) or the Yang-Mills equation (3.33) we can find the QCD *instantons*. An example for a solution is the BPST instanton [83]

$$A^a_\mu(x) = \frac{2\eta^a_{\mu\nu}(x-z)_\nu}{(x-z)^2+\rho^2} \tag{3.34}$$

with the 't Hooft symbol $\eta^a_{\mu\nu}$ [84, 85]. This is a solution where an instanton is localized at the center z_μ with size ρ.

3.4 Properties of the topological charge

From equation (3.30) we know that the action of a self-dual field (i.e. the instanton solution) is determined by the topological charge Q, (see eq. (3.31)). A crucial point now is that the topological charge is an integer and classifies the topology of gauge fields. To see that, we first need to specify the boundary of the Euclidean spacetime. For finite action configurations, the field strength tensor $F_{\mu\nu}$ needs to be square integrable ($\int F^2_{\mu\nu} < \infty$), i.e it must vanish towards the boundary of Euclidean spacetime. We have seen previously that the action has its absolute minimum when the gauge potentials are pure gauge. Thus, at infinity we require [51]

$$\lim_{x^2\to\infty} A_\mu = -iU^{(n)}(x)\partial_\mu U^{(n)\dagger}(x), \tag{3.35}$$

where the gauge group element $U^{(n)} \in \mathcal{G}$ has the winding number n. Equation (3.35) is showing us that all finite-action gauge fields can be classified topologically. This classification is similar to the topology discussion we did in the last section. However, the physical setting is different. Previously we were looking at a winding number n which classifies maps $U^{(n)} \in \mathrm{SU}(2)$ from a compactified space into the gauge group. Now, equation (3.35) defines a map from the boundary of S^3 of Euclidean spacetime into the gauge group. The salient point now is that such maps fall into disjoint homotopy classes according to the third homotopy group $\pi_3(\mathcal{G})$. For $\mathcal{G} = \mathrm{SU}(N)$ with $N = 3$, we have $\pi_3(\mathrm{SU}(N)) = \mathbb{Z}$ as in the previous section. By rewriting eq. (3.31) as a total derivative we get

$$Q = \frac{1}{16\pi^2}\int \mathrm{d}^4x\, \mathrm{tr}\, \tilde{F}_{\mu\nu}F_{\mu\nu} = \int \mathrm{d}^4x\, \partial_\mu J^{CS}_\mu = \int_{S^3} \mathrm{d}\sigma_\mu\, J^{CS}_\mu \tag{3.36}$$

where J_μ is the topological current defined in eq. (3.23). If we insert eq. (3.35) in eq. (3.36) we find that $Q = n$, and thus the topological charge is an integer. Consequently, the winding number which correspond to finite-action gauge fields (see eq. (3.35)) is the topological charge Q.

Since the last term in equation (3.36) is a surface integral over the 3-sphere, it follows that the topological charge is completely determined by the asymptotic behavior of the gauge fields. Additionally, the topological charge turns out to be invariant under small variations δA_μ of the gauge fields [73], since

$$\delta Q = \frac{1}{32\pi^2}\int_{S^3} \mathrm{d}\sigma_\mu \delta J_\mu = \frac{1}{8\pi^2}\int_{S^3} \mathrm{d}\sigma_\mu \tilde{F}^a_{\mu\nu}\delta A^a_v = 0, \tag{3.37}$$

where we made use of the fact that $\tilde{F}^a_{\mu\nu}$ vanishes at spatial infinity. Hence, it does not depend on the local behavior of the gauge fields. Different topological sectors are thus truly disparate.

Another property can be observed by writing the topological charge in terms of the Chern-Simons number [4, 51, 74]

$$Q = \int \mathrm{d}^4x\, \partial_\mu J^{CS}_\mu = \int \mathrm{d}^4x\, (\partial_0 J^{CS}_0 + \partial_i J^{CS}_i) = \int \mathrm{d}^3x\, \mathrm{d}\tau\, \partial_0 J^{CS}_0$$
$$= \int \mathrm{d}^3x\, J^{cs}_0 \Big|^{\tau=+\infty}_{\tau=-\infty} = n_{CS}(\tau=\infty) - n_{CS}(\tau=-\infty), \tag{3.38}$$

where in the first line we make use of the fact that the integral over the spatial part of the Chern-Simons current vanishes in the temporal gauge $A_0 = 0$. Equation (3.38) shows that field configurations with $Q \neq 0$ connect different topological vacua. Thus, a gauge field configuration of that kind starts out at $\tau \to -\infty$ as pure gauge field with a winding number n, and ends up at $\tau \to \infty$ as a pure gauge field with a winding number of m. The crucial point now is that these configurations start in a vacuum state (i.e. zero field energy) and still manage to pass through a barrier to reach the other vacuum state. Classically, that would require energy which is not available in the vacuum. However, since we are dealing with quantum theory, such processes are allowed. Hence, instantons are tunneling phenomena which mediate through the sphaleron barrier.

As one probably already noticed, the axial anomaly in eq. (2.23) is proportional to the topological charge density in eq. (3.31). Hence, computing the change in axial charge ΔQ_5, i.e. integrating this anomaly over spacetime leads to the topological charge Q multiplied by $2N_F$,

$$\Delta Q_5 = \int \mathrm{d}^4x \partial_\mu j^5_\mu = \int \mathrm{d}^4x \frac{N_F}{16\pi^2} F^a_{\mu\nu} \tilde{F}^a_{\mu\nu} = 2N_F Q. \tag{3.39}$$

Furthermore, the left-hand side in (2.23) can be integrated in terms of (2.22). Making use of (3.39), it can be shown that [4, 48–50]

$$Q = n_L - n_R, \tag{3.40}$$

where $n_{L/R}$ is the number of left/right-handed zero modes of localized fermion states. This relates the topological charge Q of the gauge field to the number of unpaired quark zero modes of the Dirac operator in its background. It is also referred to a special case of the Atiyah-Singer index theorem [86] for the Dirac operator. Hence, any instanton field with topological charge Q must come along with a corresponding change in axial charge $\Delta Q_5 = 2N_F Q$ which involves the existence of fermion zero modes in the background of these fields.

3.5 The sphaleron rate

In the previous section we have introduced instantons which turned out to be tunneling processes which mediate through the sphaleron barrier. The height of that barrier is of order the QCD scale Λ_{QCD} over the strong coupling constant α_s [87]. Thus, since that potential barrier is relatively high, these transitions are highly suppressed at low temperatures [84, 88]. However, at high temperatures it is possible to jump over that barrier, which is not suppressed anymore. These transitions are called *sphalerons* [40, 41] and the quark gluon plasma is the best place to find direct experimental evidence of such configurations with non-zero winding number. Mathematically speaking, these objects are static, unstable, finite-energy solutions of the classical field equations in

Minkowski time (In contrast, instantons are localized, finite-action solutions of the classical field equations in Euclidean time).

The sphaleron rate is defined in Minkowski time by the correlation function

$$\begin{aligned}\Gamma_{\text{sphal}} &= \lim_{t\to\infty} \frac{(n_{CS}(t) - n_{CS}(0))^2}{Vt} \\ &= \int \mathrm{d}^4x \left\langle \frac{1}{32\pi^2} F^a_{\mu\nu}\tilde{F}^{\mu\nu}_b(x) \frac{1}{32\pi^2} F^a_{\alpha\beta}\tilde{F}^{\alpha\beta}_b(0) \right\rangle . \end{aligned} \tag{3.41}$$

However, in lattice QCD we can only compute Euclidean correlators. Fortunately, the Minkowski correlator can be related to the Euclidean correlator. The zero-momentum Euclidean correlation function that can be calculated on the lattice is defined by

$$G(\tau) = \int \mathrm{d}^3x \left\langle q(0,0)\, q(\vec{x},\tau) \right\rangle , \tag{3.42}$$

where τ is the temporal distance between the two charge-density operators. The spectral density ρ_q of this integral can be extracted using

$$G(\tau) = -\int_0^\infty \frac{\mathrm{d}\omega}{\pi} \rho_q(\omega) \frac{\cosh\left[\omega\left(\frac{1}{2T} - \tau\right)\right]}{\sinh\frac{\omega}{2T}}, \tag{3.43}$$

where T is the temperature and the unusual minus sign arises because q is a time reversal odd operator. A so-called *Kubo formula* [89] can be derived, which returns the sphaleron rate in terms of the low-frequency part of this spectral function at zero spatial momentum,

$$\Gamma_{\text{sphal}} = \lim_{\omega\to 0} \frac{2T\rho_q(\omega)}{\omega}. \tag{3.44}$$

Extracting $\rho_q(\omega)$ from the Euclidean correlator is not straightforward. The problem is, that we can only get the correlator G in our lattice simulations, which is defined as an integral over the desired spectral function. Thus, it is an ill-posed inversion problem. The same problem arises as well in determining other transport coefficients of the quark gluon plasma, e.g. the shear viscosity. Hence, there exist methods which offer possibilities to extract this quantity [90–92].

The sphaleron rate was first studied in the context of electroweak baryogenesis [39, 93, 94]. It has been well understood for electroweak interacting matter and determined using Bödeker's effective theory [95–98] in the weak-coupling regime. This progress cleared the path for a non-perturbative semiclassical real-time evaluation on a Minkowski lattice [99, 100].

Due to its success in electroweak baryogenesis, the sphaleron rate also received attention within QCD [93, 101, 102]. More recently it has raised interest because magnetic phenomena in heavy ion collisions might give axial charge density an important role to play in, for example, the chiral magnetic effect [5, 6]. At high temperatures, where the QCD coupling is weak, there are semiclassical approaches to determine the sphaleron rate, but for the physically interesting coupling regime a direct evaluation in Minkowski space is impossible [39]. Also results of non-perturbative lattice studies are rather limited. A recent lattice QCD study on the SU(3) sphaleron rate can be found in [7].

Chapter 4

Lattice QCD

Computing the solution of the equation of motion from the Yang-Mills Lagrangian requires too much computational power. Therefore, computing the time evolution of observables at high temperatures is currently not feasible. However, it is possible to compute expectation values of observables in statistical physics according to the path integral formulation in Euclidean spacetime. The theory behind that is called lattice QCD. In this formulation spacetime is discretized in a finite volume by introducing a finite lattice spacing a such that the path integral has finite dimensions. This allows us to compute observables numerically on computers using statistical Monte-Carlo techniques. Computing observables this way results in some uncertainties. To control these discretization effects, a continuum extrapolation $a \to 0$ is done in a final step. In what follows we only give a brief explanation of the above mentioned steps. More detailed explanations can be found in the literature e.g. [45–47, 76, 77].

4.1 Discretizing QCD

The discrete spacetime is given by an isotropic 4D lattice,

$$\Lambda = \{x = \{x_0, x_1, x_2, x_3\} \mid x_i \in \mathbb{N}; x_0 \leq N_\tau - 1; x_1, x_2, x_3 \leq N_\sigma - 1\}\,, \tag{4.1}$$

where N_τ is the temporal and N_σ the spatial extent of the lattice. All sites are separated by the lattice spacing a. Then, each spacetime point x is connected to a spacetime point in the continuum via $x = an$. All resulting lattice quantities are then expressed in units of a. Spinors are placed only at sites, which means that the fermionic degrees of freedom are represented as

$$\psi(x),\ \bar{\psi}(x),\ x \in \Lambda. \tag{4.2}$$

They carry the same color, Dirac, and flavor indices as in the continuum. Local gauge invariance of the action is achieved by introducing $\mathrm{SU}(3)$-valued matrices, the so-called link variables $U_\mu(x)$. They are gauge transporters and are related to gauge fields by

$$U_\mu(x) = \exp\left(iaA_\mu(x)\right), \tag{4.3}$$

where $A_\mu(x)$ are the Lie algebra-valued gauge fields. On the lattice these quantities (4.3) are links between two adjacent sites x and $x + \hat{\mu}$. Moreover, they own an orientation, i.e. the link $U_{-\mu}(x)$, which points from x to $x - \hat{\mu}$ is related to the positively oriented link variable $U_\mu(x - \hat{\mu})$ via

$$U_{-\mu}(x) = U_\mu^\dagger(x - \hat{\mu}). \tag{4.4}$$

Under local gauge transformations a link transforms as

$$U_\mu(x) \to U_\mu(x)' = \Omega(x) U_\mu(x) \Omega^\dagger(x+\hat{\mu}), \tag{4.5}$$

where $\Omega(x)$ denotes a $\mathrm{SU}(3)$ matrix at each site x. This allows terms of the form $\bar{\psi}(x) U_\mu(x) \psi(x+\hat{\mu})$ to be gauge invariant, since the fermion fields transform as

$$\psi(x) \to \psi(x)' = \Omega \psi(x), \qquad \bar{\psi}(x) \to \bar{\psi}(x)' = \bar{\psi}(x) \Omega^\dagger. \tag{4.6}$$

The lattice version of the gauge action can be constructed in terms of a gauge invariant product of link variables. Gauge invariance is achieved by arranging the link variables into a closed loop. After taking the trace of that structure, gauge transformations vanish under cyclic permutation. The shortest, nontrivial closed loop on the lattice is called the plaquette. It is constructed as

$$P_{\mu\nu}^{(1\times1)}(x) = U_\mu(x) U_\nu(x+\hat{\mu}) U_\mu^\dagger(x+\hat{\nu}) U_\nu^\dagger(x). \tag{4.7}$$

From this definition we create the gauge action by summing up all plaquette variables on the lattice. To be more precise,

$$S_{\mathrm{G}}^{\text{Wilson}}[U] = \frac{2}{g^2} \sum_{x\in\Lambda} \sum_{\mu<\nu} \mathrm{Re}\,\mathrm{tr}\left[\mathbb{1} - P_{\mu\nu}^{(1\times1)}\right], \tag{4.8}$$

whose error is of order $\mathcal{O}\left(a^2\right)$. This is the so-called *Wilson gauge action*. A more sophisticated discretization of the action is the so called Symanzik gauge action [103, 104]

$$S_{\mathrm{G}}^{\text{Symanzik}} = \frac{2}{g^2} \sum_{x\in\Lambda} \sum_{\mu<\nu} \left\{ \frac{5}{3} \mathrm{Re}\,\mathrm{tr}\left(\mathbb{1} - P_{\mu\nu}^{(1\times1)}\right) - \frac{1}{12} \mathrm{Re}\,\mathrm{tr}\left[2\cdot\mathbb{1} - \left(P_{\mu\nu}^{(1\times2)} + P_{\mu\nu}^{(2\times1)}\right)\right] \right\}, \tag{4.9}$$

where we additionally introduced rectangle link structures

$$P_{\mu\nu}^{(1\times2)}(x) = U_\mu(x) U_\nu(x+\hat{\mu}) U_\nu(x+\hat{\mu}+\hat{\nu}) U_\mu^\dagger(x+2\hat{\nu}) U_\nu^\dagger(x+\hat{\nu}) U_\nu^\dagger(x), \tag{4.10}$$

$$P_{\mu\nu}^{(2\times1)}(x) = U_\mu(x) U_\mu(x+\hat{\mu}) U_\nu(x+2\hat{\mu}) U_\mu^\dagger(x+\hat{\mu}+\hat{\nu}) U_\mu^\dagger(x+\hat{\nu}) U_\nu^\dagger(x). \tag{4.11}$$

Contrary to the Wilson gauge action, the Symanzik gauge action eliminates errors of order $\mathcal{O}\left(a^2\right)$.

A simple discretization of the free fermion action is given by

$$S_F^{\text{free}} = a^4 \sum_{n\in\Lambda} \bar{\psi}(an) \left(\sum_{\mu=0}^{3} \gamma_\mu \frac{\psi(an+a\hat{\mu}) - \psi(an-a\hat{\mu})}{2a} + m\psi(an) \right), \tag{4.12}$$

with the unit vector $\hat{\mu}$ pointing in the direction μ. This definition is not gauge invariant. Rescaling the fields $a^{3/2}\psi(an) \to \psi(n)$ and the mass $am \to m$ and using the gauge transformation property (4.5), we get the gauge invariant fermionic action

$$S_F = \sum_{n\in\Lambda} \bar{\psi}(n) \left(\sum_{\mu=0}^{3} \gamma_\mu \frac{U_\mu(n)\psi(n+\hat{\mu}) - U_{-\mu}(n)\psi(n-\hat{\mu})}{2} + m\psi(n) \right), \tag{4.13}$$

where we use the abbreviation $U_{-\mu}(x) = U_\mu^\dagger(x-\hat{\mu})$. Using the Taylor expansion of the gauge transporter, we receive the continuum expression up to order $\mathcal{O}(a)$.

With equation (4.8) and (4.13) we have discretized the QCD action in terms of the link variables U_μ and the *Grassmann*-valued fields $\hat{\psi}$ and ψ. We can now define the path integral measure

$$\int \mathcal{D}\bar{\psi}\mathcal{D}\psi \int \mathcal{D}U \coloneqq \int \prod_{n\in\Lambda} \mathrm{d}\bar{\psi}(n)\mathrm{d}\psi(n) \int \prod_{n\in\Lambda}\prod_{\mu=0}^{3} \mathrm{d}U_\mu(n) \tag{4.14}$$

with the Grassmann-valued integration measure $\mathrm{d}\psi(n)$ and the *Haar* measure $\mathrm{d}U_\mu$ for $\mathrm{SU}(3)$ matrices. Since the Grassmann-valued fermionic part can be integrated out, the partition function can be written entirely in terms of the link variables and the fermion determinant

$$\begin{aligned} Z &= \int \prod_f \mathcal{D}\bar{\psi}_f \mathcal{D}\psi_f \mathcal{D}U \mathrm{e}^{S_F[U,\bar{\psi}_f,\psi_f,m_f]} \mathrm{e}^{-S_G[U]} \\ &= \int \mathcal{D}U \prod_f \det D[U,m_f] \mathrm{e}^{-S_G[U]}, \end{aligned}$$

where $D[U,m_f]$ is the *Dirac matrix*. It is represented on the discrete spacetime as

$$D_{nm}[U] = \sum_{\mu=0}^{3} \gamma_\mu \frac{U_\mu(n)\delta_{n,m-\mu} - U_\mu^\dagger(n-\hat{\mu})\delta_{n,m+\mu}}{2} + m\delta_{n,m}. \tag{4.15}$$

The Dirac operator is a normal operator, which means it is γ_5-Hermitian,

$$\gamma_5 D \gamma_5 = D^\dagger. \tag{4.16}$$

This ensures that eigenvectors of the Dirac operator form an orthogonal basis. Additionally, it also leads to the fermion determinant to be real,

$$\det[D]^* = \det\left[D^\dagger\right] = \det[\gamma_5 D \gamma_5] = \det[D]. \tag{4.17}$$

The partition function is a key ingredient in statistical physics, and thus important in lattice QCD. However, the direct computation of the partition function on current supercomputer is not feasible. A better way is to calculate the partition function in the path integral formalism in terms of finite number of fields on the lattice. This method is called *importance sampling*. Consider the expectation values of observables in the path integral representation according to

$$\langle O \rangle = \frac{1}{Z} \int \prod_f \mathcal{D}\bar{\psi}_f \mathcal{D}\psi_f \mathcal{D}U \mathcal{O}[U,\bar{\psi}_f,\psi_f] \mathrm{e}^{-S_F[U,\bar{\psi}_f,\psi_f,m_f]} \mathrm{e}^{-S_G[U]}. \tag{4.18}$$

Using *Wick's theorem* the Grassmann-valued fermionic integrals can be computed analytically for most observables [45], which results in

$$\langle O \rangle = \frac{1}{Z} \int \mathcal{D}U \mathcal{O}_F[U] \mathrm{e}^{-S_G[U]} \prod_f \det D[U,m_f], \tag{4.19}$$

with $\mathcal{O}_F$ given by

$$\mathcal{O}_F = \langle \mathcal{O} \rangle_F = \frac{1}{Z_F[U]} \int \prod_f \mathcal{D}\bar{\psi}_f \mathcal{D}\psi_f \mathcal{O}[U,\bar{\psi}_f,\psi_f] \mathrm{e}^{-S_F[U,\bar{\psi}_f,\psi_f,m_f]}, \tag{4.20}$$

where Z_F is defined as

$$Z_F[U] = \int \prod_f \mathcal{D}\bar{\psi}_f \mathcal{D}\psi_f \mathrm{e}^{-S_F[U,\bar{\psi}_f,\psi_f,m_f]}. \tag{4.21}$$

The expectation value in equation (4.19) can now be calculated by performing importance sampling. This is achieved by using so-called *gauge configurations* C_i, which are one possible realization of the gauge field in the path integral. By distributing the gauge configurations C_i according to the probability distribution

$$\mathrm{d}P[C] = \frac{\prod_f \det D[C, m_f]}{Z} \mathrm{e}^{-S_G[C]} \mathrm{d}C, \tag{4.22}$$

we may approximate the expectation value of an operator in the continuum as

$$\langle \mathcal{O} \rangle \approx \frac{1}{N_{\text{conf}}} \sum_{i=1}^{N_{\text{conf}}} O[C_i] \tag{4.23}$$

with N_{conf} the total number of gauge configurations sampled on the lattice. In the limit of large N_{conf} it matches the continuum value.

The Dirac matrix has the dimension $|\Lambda| \times |\Lambda|$, which makes the computation of its determinant not feasible. Fortunately, methods have been developed in the past that avoid the calculation of the determinant. The so-called Hybrid Monte Carlo [105] is widely used to generate gauge configurations according to the probability distribution in equation (4.22) for QCD with dynamical fermions. Although computers became a lot more powerful in the past few decades, it is still not feasible to perform lattice QCD computations including fermions with physical masses on lattices of sizes larger than $|\Lambda| \sim 100^4$. Hence, for qualitative understanding of QCD we sometimes compute observables in the so called *quenched approximation* [53]. In that approximation the determinant of the Dirac matrix is set to unity $\det D[U, m_f] = \mathbb{1}$, which also implies an infinite quark mass limit $m_f \to \infty$. In particular, this is equivalent to neglecting all kinds of quark loops. Even in this limit the pure gauge theory has a first order phase transition [45]. Another advantage is that setting the determinant to one accelerates the calculations dramatically. Therefore, the quenched approximation is a good starting point to test and develop new lattice QCD methods. In the quenched approximation, a combination of overrelaxation [106, 107] and heat bath updates [108, 109] turns out to be most efficient to generate gauge configurations according to the probability distribution in equation (4.22).

The way we approximate the path integral to compute observables on the lattice does not come without uncertainties. The approximation in equation (4.23) shows already that the expectation value will necessary come with a statistical error, since we can only use a finite set of configurations. Additionally, observables and correlation functions suffer from systematic errors due to the finite lattice spacing. This error is usually of order $\mathcal{O}(a)$ or $\mathcal{O}(a^2)$ depending on the discretization scheme and may be controlled by performing a continuum extrapolation. Performing the limit $a \to 0$ one needs to keep physical quantities fixed. Therefore, it is important to increase N_σ and N_τ so that $\beta = N_\tau a$ and $L_\sigma = N_\sigma a$ remain constant.

Furthermore, it is necessary to make sure that uncertainties introduced by finite volume effects are controlled when choosing the size of the lattices. We find that for most thermodynamic quantities an aspect ratio of $N_\sigma / N_\tau \geq 4$ is sufficient.

4.2 Fermion doublers and the Wilson discretization

Discretizing the fermion action on the lattice as in the previous section has the problem that we end up having unphysical fermion doublers which change the underlying physics. In order to see this, let us have a look at the Fourier transform of the lattice Dirac operator for free lattice fermions:

$$\begin{aligned}\tilde{D}(p|q) &= \frac{1}{|\Lambda|}\sum_{n,m\in\Lambda} \mathrm{e}^{-ipna} D(n|m)\mathrm{e}^{ipma} \\ &= \delta(p-q)\tilde{D}(p),\end{aligned} \tag{4.24}$$

from which it follows

$$\tilde{D}(p) = m\mathbb{1} + \frac{i}{a}\sum_{\mu=1}^{4}\gamma_\mu \sin(p_\mu a). \tag{4.25}$$

The quark propagator is related to the inverse of eq. (4.25), which reads

$$\tilde{D}(p)^{-1} = \frac{m\mathbb{1} - ia^{-1}\sum_\mu \gamma_\mu \sin(p_\mu a)}{m^2 + a^{-2}\sum_\mu \sin(p_\mu a)^2}. \tag{4.26}$$

In the limit $a \to 0$, this propagator has a pole at $p = 0$. This means in the continuum theory the propagator corresponds to a physical particle, satisfying the Dirac equation. However, for finite a with massless fermions, eq. (4.26) shows multiple poles everywhere where $p_\mu a = \pi$. In particular, 15 so called *fermion doublers* arise that correspond to extra particles. In the free theory they do not matter, since they only increase the effective degrees of freedom, but in the interacting theory there may be gluons of momentum π/a, which will mediate interactions among the doublers. A way to remove these doublers is by introducing an extra term in our discretized theory. At finite lattice spacing it makes the doublers heavy, but in the continuum limit it still leads to the correct physics. Such an extra term was proposed by Wilson [110] and it reads

$$\delta D_{n,m}[U] = -\sum_{\mu}^{3}\frac{U_\mu(n)\delta_{n+\hat{\mu},m} - 2\mathbb{1}\delta_{n,m} + U_{-\mu}(n)\delta_{n-\hat{\mu},m}}{2}. \tag{4.27}$$

This so called *Wilson term* leads to an additional term in the lattice momentum space Dirac operator

$$\tilde{D}(p) = m\mathbb{1} + \frac{i}{a}\sum_\mu \gamma_\mu \sin(p_\mu a) + \frac{1}{a}\sum_\mu \mathbb{1}\big(1-\cos(p_\mu a)\big). \tag{4.28}$$

It can be shown by calculating the corresponding momentum space propagator $\tilde{D}^{-1}$ that the unwanted doublers become infinitely massive, and hence decouple in the continuum limit. The complete Dirac operator using this formulation becomes [45]

$$D(x|y)^f_{\alpha\beta c_1 c_2} = \left(m^f + \frac{4}{a}\right)\delta_{\alpha\beta}\delta_{c_1c_2}\delta_{xy} - \frac{1}{2a}\sum_{\mu=\pm1}^{\pm4}(\mathbb{1}-\gamma_\mu)_{\alpha\beta}U_\mu(x)_{c_1c_2}\delta_{x+a\hat{\mu},y}, \tag{4.29}$$

where

$$\gamma_{-\mu} \equiv -\gamma_\mu. \tag{4.30}$$

This formulation of fermions is called *Wilson fermions*. A drawback of this term is that it breaks the chiral symmetry, which was given in the naive discretization in the

previous section. This can be seen by rewriting the additional term in eq. (4.28),

$$\frac{1}{a}\sum_{\mu}\mathbb{1}\big(1-\cos(p_\mu a)\big) \tag{4.31}$$

in terms of exponentials and doing the inverse Fourier transformation,

$$a\sum_{\mu=1}^{4}\frac{1}{2a^2}\left(2\delta_{x,y}-\delta_{x,y-a\hat{\mu}}-\delta_{x,y+a\hat{\mu}}\right). \tag{4.32}$$

This term does not anti-commute with γ_5 (see appendix A.1), and hence breaks the chiral symmetry.

4.3 The staggered discretization

Instead of eliminating doublers, another approach would be to interpret them as additional flavors. However, such a 16-flavor QCD is not physical, so we need to reduce the number of flavors. The *staggered action* is one possibility to do so. It reduces the number of degrees of freedom to 4 flavors by introducing a lattice structure which mixes spinor and spacetime indices. The remaining 4 flavors are then getting removed by taking the root of the staggered fermion determinant in a final step. The basic idea is to use the following transformation to replace quark fields:

$$\psi(n)=\gamma_1^{n_1}\gamma_2^{n_2}\gamma_3^{n_3}\gamma_4^{n_4}\psi'(n),\quad \bar{\psi}(n)=\bar{\psi}'(n)\gamma_4^{n_4}\gamma_3^{n_3}\gamma_2^{n_2}\gamma_1^{n_1}. \tag{4.33}$$

with n_i being the coordinates of the lattice points. Since these transformed fields come with power of gamma matrices and since the square of gamma matrices yields unity (see section A.1), the naive action will be diagonalized in Dirac space. In particular, when inserting that transformation in the naive discretization of the fermion action eq. (4.13) one eventually interchanges gamma matrices. Each interchange of the gamma matrices results in a phase factor of (-1). Hence, we find relations of that form

$$\bar{\psi}(n)\gamma_\mu\psi(n\pm\hat{\mu})=\eta_\mu\bar{\psi}(n)'\mathbb{1}\psi(n\pm\hat{\mu})', \tag{4.34}$$

where the staggered phases η_μ are products of the (-1) factors

$$\eta_\mu=(-1)^{\sum_{\nu<\mu}n_\nu}. \tag{4.35}$$

In that step the gammas disappear, and thus the action becomes diagonal in Dirac space. The diagonal of the Dirac matrix carries four copies of the same equation. Since these copies carry no additional information, we can drop three of them and keep only one copy $\chi(n)$, and the number of doublers is reduced. The staggered fermionic action is then given by

$$S_F=\sum_{n\in\Lambda}\bar{\chi}(n)\left(\sum_{\mu=0}^{3}\eta_\mu(n)\frac{U_\mu(n)\chi(n+\hat{\mu})-U_{-\mu}(n)\chi(n-\hat{\mu})}{2}+m\chi(n)\right) \tag{4.36}$$

By removing three of the four spinor components one may argue that the information of the original Dirac structure must be hidden in the staggered phases. However, that is not possible, since these phases are just scalar values, and thus can not carry the information of the γ-matrices at each site. This information is actually spread

over the lattice within hypercubes of size 2^4. This can be seen by projecting the one dimensional quark fields back to the spinor entries using

$$\psi^t_\alpha(N) = \frac{1}{8}\sum_\rho \Omega_{\alpha t,\rho}\chi(N+\rho), \tag{4.37}$$

where we label the coordinates of each hypercube by $N = (\vec{N}, N_{\hat{\tau}})$. Ω is defined as

$$\Omega_{\alpha t,\rho} = (\gamma_1^{\rho_1}\gamma_2^{\rho_2}\gamma_3^{\rho_3}\gamma_4^{\rho_4})_{\alpha t}, \tag{4.38}$$

where the sub-vector ρ_i is pointing to one of the corners of the hypercube. The hypercubes have 16 sites, but the spinor structure has only four dimensions. Therefore, we had to introduce that additional index t into the spinor. This additional structure is called *taste* (in analogy to flavor). With that back-transformation, we finally get for the total QCD action [45]

$$S_F = b^4\sum_N \left(\sum_{t=0}^{3}\left(m\bar{\psi}^t(N)\psi^t(N) + \sum_{\mu=0}^{3}\bar{\psi}^t(N)\gamma_\mu\nabla_\mu\psi^t(N)\right)\right. \tag{4.39}$$

$$\left. - \frac{b}{2}\sum_{t,t'=0}^{3}\sum_{\mu=0}^{3}\bar{\psi}^t(N)\gamma_5(\tau_5\tau_\mu)_{t,t'}(\nabla)^2_\mu\psi^{t'}(N)\right), \tag{4.40}$$

where $\tau_\mu = \gamma_\mu^T$,∇_μ is the discretized derivative on the lattice of hypercubes and b is the lattice spacing on the lattice of hypercubes. The third term is called the *taste-breaking* term. Similar to the Wilson term from the previous section, it removes the unwanted fermion doublers and vanishes in the continuum limit. However, it allows for interactions between fermions of different taste, which is unphysical. This *taste mixing* term reduces the symmetry of the kinetic term and is only invariant under a remaining $\mathrm{U}(1) \times \mathrm{U}(1)$ symmetry. In particular this term is invariant under transformations

$$\psi' = \mathrm{e}^{i\omega}\psi, \qquad \bar{\psi}' = \bar{\psi}\mathrm{e}^{-i\omega} \tag{4.41}$$

and

$$\psi' = \mathrm{e}^{i\omega\gamma_5\otimes\tau_5}\psi, \qquad \bar{\psi}' = \bar{\psi}\mathrm{e}^{i\omega\gamma_5\otimes\tau_5}. \tag{4.42}$$

These transformations correspond to a subgroup of the symmetry group $\mathrm{SU}(N_t)$, where N_t is the number of tastes. Another drawback of that taste structure is that the staggered discretization comes with three extra unphysical flavors. To eliminate them we have to take the root of the fermion determinant. For example in $2+1$ flavor QCD the path integral may be defined as

$$\langle O\rangle = \frac{1}{Z}\int \mathcal{D}U\mathcal{O}[U](\det D[U, m_{u/d}])^{1/2}(\det D[U, m_s])^{1/4}\mathrm{e}^{-S_G[U]}. \tag{4.43}$$

This rooting of the fermion determinant, however, is controversial. It is not yet clear whether the staggered action converges to the true continuum limit, even though the additional term in eq. (4.39) vanishes and the tastes decouple in the continuum [111, 112]. So far, however, reasonable results have been obtained with the staggered action and no evidence has been found that the continuum limit is not yielding true physics. A way to reduce the effect of the taste mixing may be realized by introducing extra terms into the staggered action. The *Highly Improved Staggered Quark (HISQ)* action is one way of doing that [31]. Apart from suppressing taste-changing interactions it reduces

also the order of the discretization error to $\mathcal{O}(a^2)$. The details of the implementation of the HISQ action can be found in [32].

4.4 The Ginsparg-Wilson relation and the overlap Dirac operator

Implementing fermions with exact chiral symmetry on the lattice used to be a big problem. Nielsen and Ninomiya formulated a no-go theorem [113] which states that chiral symmetry can not be established on the lattice without violating either of these properties like translational invariance, ultra-locality or the absence of doublers. Only later it has been shown [114], that there exists a corresponding exact symmetry on the lattice if the Dirac operator satisfies the so called *Ginsparg-Wilson (GW) relation* [115], which becomes the standard chiral symmetry in the continuum limit. The GW relation is defined as

$$D\gamma_5 + \gamma_5 D = \frac{1}{r} D\gamma_5 D, \tag{4.44}$$

where D is some Dirac operator and r is a parameter to fix the scale of the Dirac operator. If a Dirac operator which obeys the GW equation is multiplied by a factor α, then that new operator is a solution of the GW equation where r is replaced by αr. The Ginsparg-Wilson relation eq. (4.44) guarantees that any Dirac operator which satisfies that relation is chirally symmetric. A property that the Wilson and staggered fermions share is that they are γ_5-Hermitian ($D^\dagger = \gamma_5 D \gamma_5$). The solution of the GW equation, which we are going to derive later, will share that property too. Thus, let us make use of the γ_5-hermiticity and multiply eq. (4.44) with γ_5 from either side

$$D + D^\dagger = \frac{1}{r} D^\dagger D = \frac{1}{r} D D^\dagger. \tag{4.45}$$

It follows that $[D, D^\dagger] = 0$ if D is γ_5-Hermitian. Operators which are of that kind are called *normal*, which means that they have an orthonormal set of eigenvectors that form a basis for the vector space. D can be diagonalized in such a way that eigenvectors of D with eigenvalue $\lambda = x + iy$ are also eigenvectors of $D^\dagger$ with eigenvalue $\lambda^\dagger = x - iy$ [45]. Multiplying eq. (4.45) with a normalized eigenvector v_λ from the right and with $v_\lambda^\dagger$ from the left, we get

$$\begin{aligned} 2rx &= x^2 + y^2 \\ \Leftrightarrow (x-r)^2 + y^2 &= r^2, \end{aligned} \tag{4.46}$$

which shows that the eigenvalues lie on a circle in the complex plane with radius r and the center $r + i0$. This *Ginsparg-Wilson circle* can be parametrized as

$$\lambda = r\left(1 - \mathrm{e}^{-i\phi}\right) \quad \text{with} \quad \phi \in (-\pi, \pi]. \tag{4.47}$$

The eigenvalues can therefore become purely real at $\phi = 0$ and $\phi = \pi$, i.e $\lambda = 0$ and $\lambda = 2r$ respectively. Another consequence of the γ_5-hermiticity can be seen in the eigensystem. Calculating the characteristic polynomial $P(\lambda)$ of D, we find,

$$P(\lambda) = \det\left[D - \lambda\mathbb{1}\right] = \det\left[\gamma_5^2(D - \lambda\mathbb{1})\right] = \det\left[\gamma_5(D - \lambda\mathbb{1})\gamma_5\right] \tag{4.48}$$

$$= \det\left[D^\dagger - \lambda\mathbb{1}\right] = \det\left[D - \lambda^*\mathbb{1}\right]^* = P(\lambda^*)^*, \tag{4.49}$$

where $\mathbb{1} = \gamma_5^2$ has been used. Since the eigenvalues λ are obtained by solving $P(\lambda) = 0$, eq. (4.48) implies that if λ is a zero, then λ^* is also a zero. Therefore, the eigenvalues for a γ_5-hermitian Dirac operator are either real, or come in complex conjugate pairs. A further interesting consequence of the γ_5-hermiticity can be obtained from the γ_5 matrix element of the eigenvectors. Defining the dot product $u^\dagger v \equiv (u, v)$, we get for the γ_5 matrix element,

$$\lambda(v_\lambda, \gamma_5 v_\lambda) = (v_\lambda, \gamma_5 D v_\lambda) = (v_\lambda, D^\dagger \gamma_5 v_\lambda) = (D v_\lambda, \gamma_5 v_\lambda) = \lambda^*(v_\lambda, \gamma_5 v_\lambda), \tag{4.50}$$

and therefore it follows

$$(v_\lambda, \gamma_5 v_\lambda) = 0, \quad \text{unless} \quad \lambda \in \mathbb{R}. \tag{4.51}$$

An eigenvector is chiral if this matrix element is nonzero, i.e. $(v_\lambda, \gamma_5 v_\lambda) \neq 0$. Hence, from eq. (4.50) it follows that an eigenvector can have non-vanishing chirality only when its eigenvalues are real. We have seen that the only possible real values are $\lambda = 0$ and $\lambda = 2r$. $\lambda = 0$ are the exact zero modes. $\lambda = 2r$ on the other hand are the doubler partner of the zero modes [45].

A general solution of the Ginsparg-Wilson relation, eq. (4.44) can be written as,

$$D = r(1 + V) \quad \text{with} \quad V^\dagger V = \mathbb{1}, \tag{4.52}$$

where the unitary operator V can be written in terms of a sign function of some normal operator W,

$$V = \text{sgn}(W) = \frac{W}{\sqrt{W^\dagger W}} \tag{4.53}$$

This solution was first discussed by Neuberger which he named as the *overlap-Dirac operator* D_{ov} [37, 38]. To fulfill the requirement that the Dirac operator is γ_5-Hermitian, it is sufficient to choose W to be γ_5-Hermitian. Thus, we can define $K = \gamma_5 W$ to be Hermitian and the Dirac operator takes the form,

$$D_{\text{ov}} = r(1 + \gamma_5 \sigma) = r\left(1 + \gamma_5 \frac{K}{\sqrt{K}}\right) \tag{4.54}$$

This also fulfills the Ginsparg-Wilson equation,

$$\{\gamma_5, D_{\text{ov}}\} - \frac{1}{r} D_{\text{ov}} \gamma_5 D_{\text{ov}} = r\left(2\gamma_5 + \sigma + \gamma_5 \sigma \gamma_5 - (1 + \gamma_5 \sigma)\gamma_5 (1 + \gamma_5 \sigma)\right) = 0, \tag{4.55}$$

where $\sigma^2 = \mathbb{1}$ and $\gamma_5^2 = \mathbb{1}$ have been used. Another important property is that the Dirac operator is local in order to respect causality in QFT. This means that the interaction length should be fixed in lattice units and infinitesimal in physical units when the continuum limit is performed, i.e. $|D_{\text{ov}}(n|m)|$ should fall exponentially with the distance $|n - m|$. However, due to the square root in the denominator, D_{ov} will involve more than just a few lattice points in the intermediate neighborhood, i.e. D_{ov} will not be ultra-local. Fortunately, the sign function is defined by its polynomial expansion. Thus, a way to tackle that issue is to demand that the kernel W is ultra-local. The *overlap Dirac operator* is the solution of the GW equation where the Kernel is chosen to be the Wilson Dirac operator with negative mass, i.e. $W = D^W_{-M}$. The Wilson Dirac operator satisfies the locality property, which has been shown in [116].

Furthermore, D_{ov} should have no doubler and reproduce continuum physics in the continuum limit. For the overlap Dirac operator one can show that easily in the non-interacting case. To see that, let us start with the Wilson Dirac operator in

momentum space,

$$W \propto \sum_{\mu} (1 - \cos(ap_\mu) + i\gamma_\mu \sin(ap_\mu)) - aM. \tag{4.56}$$

In the vicinity of a pole of the naive Dirac propagator, the momentum can be expressed as $ap_\mu = \pi n_\mu + \varepsilon_\mu$ where $n_\mu = \{0, 1\}$ and $n = \sum_\mu = n_\mu$ [68]. An expansion in ε then gives

$$\begin{aligned} \cos(ap_\mu) &= 1 - 2\delta_{1n_\mu} + \mathcal{O}(\varepsilon^2), \\ \sin(ap_\mu) &= (1 - 2\delta_{1n_\mu})\varepsilon_\mu + \mathcal{O}(\varepsilon^3). \end{aligned}$$

In order to check whether there are no doublers, only terms independent of ε are important. Using that in the overlap Dirac operator yields

$$D_{\text{ov}} = r\left(1 + \frac{2n - aM}{|2n - aM|}\right) = \begin{cases} 0 & 2n < aM \\ 2r & 2n > aM \end{cases}. \tag{4.57}$$

Therefore, aM has to be within $[0, 2)$ such that the Dirac operator vanishes at the proper pole $n = 0$, but not at doublers $n \geq 1$. Moreover, in order for the doublers to decouple in the continuum limit, r should be of order $1/a$. To compare it with the continuum Dirac operator we keep the linear terms and restrict our calculation to the proper pole. This gives

$$W \propto -M + i\gamma_\mu p_\mu + \mathcal{O}(p^2), \tag{4.58}$$

where we omit the summation sign. Plugging that into the overlap Dirac operator gives

$$\begin{aligned} D_{\text{ov}} &= r\left(1 + \frac{-M + i\gamma_\mu p_\mu}{\sqrt{M^2}}\right) + \mathcal{O}(p^2) \\ &= \frac{r}{M} i\gamma_\mu p_\mu + \mathcal{O}(p^2). \end{aligned} \tag{4.59}$$

Setting $r = M$ will then give us the correct continuum limit $i\gamma_\mu p_\mu$. The best result is obtained by selecting aM close to 2, as the mass of the doublers is given by $2r$. However, the exact border where the doublers become massless again may get excited due to interactions. Hence, aM should not be too close to 2. In non-perturbative simulations values in the range $aM = [1.4, 1.8]$ have proven to give good results [68]. In this work, we have chosen $aM = 1.8$ most of the time. Unless the Dirac operator is non-analytic, a continuous deformation of the gauge field does not change the topological charge. If M is chosen in such a way that K^2 has zero modes, D_{ov} becomes non-analytic. However, in that case the sign function becomes ill-defined. In particular, if M is changing continuously such that the lowest eigenvalue of K^2 first reaches zero and then rises again to some non-zero value, the topological charge will have changed. The numerical approximation of the sign function becomes difficult when an eigenvalue of K^2 is close to zero and eigenmodes of K^2 with low precision result in a badly satisfied Ginsparg-Wilson equation. Therefore, a way to tackle this issue is to compute the lowest eigenvalues of K^2 before computing the actual eigenvalue spectrum of the overlap operator. If the lowest eigenvalue in that measurement was exceptionally small, then the Wilson spectrum should be recomputed with a lower value of aM, e.g. 1.7.

As mentioned before, r and M have to be equal, such that the overlap Dirac operator eq. (4.54) matches the normalization of the continuum Dirac operator.

However, while calculating expectation values of fermion observables, the choice of r is irrelevant. Eigenvectors and eigenvalues computed by different choices of r are trivially related, and therefore we simply used $r = 1$ in the implementation.

The overlap operator eq. (4.54) and its Hermitian conjugate can be written as

$$D_{\text{ov}} = 1 + W I(K^2) \tag{4.60}$$

$$D_{\text{ov}}^\dagger = 1 + I(K^2) W^\dagger, \tag{4.61}$$

where $I(K^2) = 1/\sqrt{K^2}$. The function $I(K^2)$ can be approximated as [68]

$$I(K^2) \approx \sum_{i=1}^{n} \frac{1}{\sqrt{\kappa_i}} k_i k_i^\dagger + Q(K^2) \left(1 - \sum_{i=1}^{n} k_i k_i^\dagger\right) \tag{4.62}$$

$$\text{with} \qquad Q(t) = \sum_{i=1}^{m} \frac{c_i}{d_i + t}, \tag{4.63}$$

where k_i are the eigenvectors and κ_i the eigenvalues of K^2. The first n eigenvectors of the operator $I(K^2)$ are calculated exactly and the remaining higher eigenvectors are approximated through the Zolotarev rational function Q of order $[m-1, m]$ [117, 118]. The coefficients c_i and d_i are functions of $\lambda_{\text{min}}/\lambda_{\text{max}}$, where λ_{max} is obtained by the lowest eigenvalue of $-K^2$ (which is the highest eigenvalue of K^2) and λ_{min} is the smallest eigenvalue of the first n explicitly computed eigenvalues of K^2. Therefore, all eigenvalues of K^2 which are not explicitly treated should be between λ_{min} and λ_{max} [68]. The most computationally expensive part is finding the solution of the set of linear equations

$$(K^2 + d_i) y_i = x_i, \tag{4.64}$$

such that $Q(K^2)x = \sum_{i=1}^{m} c_i y_i$. The operator K^2 is hermitian, and thus a *conjugate gradient* (CG) solver can be used. In fact, a *multi-shift CG* [119] can be used, because the operators differ only by their shifts d_i, and therefore belong to the same Krylov space. This way the CG basically only solves the smallest shift, while the other solutions can be obtained by performing some linear algebra, which reduces the amount of CG iterations dramatically.

The code used in this work to compute the overlap eigenvalues is part of a framework which has been developed by the Bielefeld lattice QCD group, called *Bielefeld ParallelLatticeCode*. The code computes the eigenvalue spectrum using the Kalkreuter-Simma Ritz algorithm [120] and the QUDA library [121] in the back-end to accelerate the computations on the GPU. The numerical details of the implementation can be found in [68].

4.5 The discretized field strength tensor

Later in this work we need to compute the topological charge density on the lattice. Since the charge is constructed by the field strength tensor as in eq. (3.31), we need to use a discretized version of the tensor for our lattice calculations. By expanding the plaquette eq. (4.7), one can derive a simple lattice representation of the field strength tensor

$$F_{\mu\nu}^{\text{plaq}}(x) = \frac{1}{2a^2} \left(P_{\mu\nu}^{(1\times1)}(x) - P_{\mu\nu}^{(1\times1)\dagger}(x) \right)_{\text{AH}} + \mathcal{O}\left(a^2\right), \tag{4.65}$$

where AH is the projection on the traceless antihermitian part. This is required, since the field strength tensor is an element of the Lie Algebra $\mathfrak{su}(3)$. However, $F_{\mu\nu}^{\text{plaq}}(x)$

has the drawback, that it is centered in the middle of the $\mu - \nu$ plaquette. This leads to $\mathcal{O}(a)$ errors in the definition of the topological charge density. In this work we use a Symanzik $\mathcal{O}(a^2)$-improver field strength tensor following ref. [122–124]. This definition is based on a clover leaf shaped arrangement of plaquettes and rectangles around a lattice site x. We define two different clover shaped discretizations, where one is based on plaquettes eq. (4.7) and the other one is based on rectangles (4.10)

$$C_{\mu\nu}^{(1,1)}(x) = \frac{1}{4}\left(P_{\mu\nu}^{(1\times1)}(x) + P_{\nu-\mu}^{(1\times1)}(x) + P_{-\mu-\nu}^{(1\times1)}(x) + P_{-\nu\mu}^{(1\times1)}(x)\right), \tag{4.66}$$

$$\begin{aligned} C_{\mu\nu}^{(1,2)}(x) = \frac{1}{8}\Big(&P_{\mu\nu}^{(2\times1)}(x) + P_{\nu-\mu}^{(2\times1)}(x) + P_{-\mu-\nu}^{(2\times1)}(x) + P_{-\nu\mu}^{(2\times1)}(x) \\ &+P_{\mu\nu}^{(1\times2)}(x) + P_{\nu-\mu}^{(1\times2)}(x) + P_{-\mu-\nu}^{(1\times2)}(x) + P_{-\nu\mu}^{(1\times2)}(x)\Big). \end{aligned} \tag{4.67}$$

Additionally, we define the combinations

$$\Omega_{\mu\nu}^{(1,1)}(x) = \frac{1}{2a^2}\left(C_{\mu\nu}^{(1,1)}(x) - C_{\mu\nu}^{(1,1)\dagger}(x)\right)_{\mathrm{AH}}, \tag{4.68}$$

$$\Omega_{\mu\nu}^{(1,2)}(x) = \frac{1}{2a^2}\left(C_{\mu\nu}^{(1,2)}(x) - C_{\mu\nu}^{(1,2)\dagger}(x)\right)_{\mathrm{AH}}, \tag{4.69}$$

where AH is again the projection on the traceless antihermitian part. Expanding those expressions for small a yields

$$\begin{aligned} \Omega_{\mu\nu}^{(1,1)}(x) =& F_{\mu\nu}(x) + \frac{1}{6}a^2\left(\partial_\mu^2 + \partial_\nu^2\right)F_{\mu\nu}(x) + \frac{1}{120}a^4\left(\partial_\mu^4 + \partial_\nu^4\right)F_{\mu\nu}(x) \\ &+ \frac{1}{36}a^4\left(\partial_\mu^2\partial_\nu^2\right)F_{\mu\nu}(x) + \mathcal{O}\left(a^6\right), \end{aligned} \tag{4.70}$$

$$\begin{aligned} \Omega_{\mu\nu}^{(1,2)}(x) =& 2F_{\mu\nu}(x) + \frac{5}{6}a^2\left(\partial_\mu^2 + \partial_\nu^2\right)F_{\mu\nu}(x) + \frac{17}{120}a^4\left(\partial_\mu^4 + \partial_\nu^4\right)F_{\mu\nu}(x) \\ &+ \frac{2}{9}a^4\left(\partial_\mu^2\partial_\nu^2\right)F_{\mu\nu}(x) + \mathcal{O}\left(a^6\right). \end{aligned} \tag{4.71}$$

The $\mathcal{O}(a^2)$ contributions can now be removed by a proper linear combination of $\Omega_{\mu\nu}^{(1,1)}(x)$ and $\Omega_{\mu\nu}^{(1,2)}(x)$

$$\begin{aligned} F_{\mu\nu}^{\mathrm{Imp}}(x) &= \frac{5}{3}\Omega_{\mu\nu}^{(1,1)}(x) - \frac{1}{3}\Omega_{\mu\nu}^{(1,2)}(x) \\ &= F_{\mu\nu}(x) - \frac{1}{30}a^4\left(\partial_\mu^4 + \partial_\nu^4\right)F_{\mu\nu}(x) - \frac{1}{36}a^4\left(\partial_\mu^2\partial_\nu^2\right)F_{\mu\nu}(x) + \mathcal{O}\left(a^6\right), \end{aligned} \tag{4.72}$$

which is the final $\mathcal{O}(a^2)$-improved field strength tensor.

4.6 Topological charge

The topological charge can be measured on the lattice in various ways. A very precise approach would be to make use of the index theorem and measure the topological charge according to equation (3.40), i.e. via the difference of the number of left- and right-handed zero modes of the Dirac operator. However, this approach is computationally very expensive, since the precise determination of chiral zero modes of the Dirac operator requires a lattice discretization of the Dirac operator which obeys chiral symmetry. One such discretization of the Dirac operator is the overlap Dirac operator which we have discussed in section 4.4.

Another way to measure the topological charge on the lattice is by using the field theoretical definition eq. (3.31). Its lattice version is given by

$$Q = a^4 \sum_x q(x) \tag{4.73}$$

where $q(x)$ is the topological charge density and a the lattice spacing. The density $q(x)$ is defined as in equation (3.31), but we replace the continuum field strength tensor with the discretized version as in eq. (4.72). After applying the rules of the epsilon tensor, the topological charge density reads

$$\begin{aligned} q(x) =& -\frac{1}{32\pi^2}\epsilon_{\mu\nu\rho\sigma}\,\mathrm{tr}\left(\hat{F}_{\mu\nu}(x)\hat{F}_{\rho\sigma}(x)\right) \\ =& \frac{1}{8\pi^2}\left(\hat{F}^a_{01}(x)\hat{F}^a_{23}(x) + \hat{F}^a_{02}(x)\hat{F}^a_{31}(x) + \hat{F}^a_{03}(x)\hat{F}^a_{12}(x)\right). \end{aligned} \tag{4.74}$$

The topological susceptibility is the variance of Q divided by the four-volume V,

$$\chi_{\text{top}} = \frac{\langle Q^2\rangle - \langle Q\rangle^2}{V}. \tag{4.75}$$

This gluonic definition of the topological charge, however, is highly contaminated by ultra-violet fluctuations which can lead to entirely wrong results for the charge. To eliminate these fluctuations one should apply a smoothing method on the gauge fields before measuring $q(x)$. A very robust smoothing method, called the *gradient flow*, will be discussed in section 4.7.

Contrary to the fermionic approach to determine the topological charge, the gluonic definition on the lattice does not give truly integer values, due to discretization effects. In this work, however, we suppress these effects by using an $\mathcal{O}(a^2)$-improved lattice discretization of the field strength tensor (see section 4.5) instead of the standard plaquette discretization in the definition of the topological charge density.

In the continuum configurations with different topological charge are fully disparate and only connected via configurations with infinite action (see eq. (3.37)). Such configurations can not be deformed into each other by continuous transformations. Hence, configurations generated in Monte Carlo simulations should get trapped in a specific topological sector. However, on the lattice these infinite action barriers become just configurations with large values of the action. Therefore, instantons can tunnel through such non-continuum like configurations in Monte Carlo simulations. Unfortunately, with decreasing lattice spacings these finite action barriers get more and more strongly suppressed. As a consequence, also lattice field configurations evolve slowly in Q, as the lattice spacing decreases. This effect is called "topological freezing" and leads to large autocorrelation times which deteriorate the statistical power of the final results. There are some methods to reduce autocorrelations of the topological charge [125]. For example in [126] the authors propose to use open boundary conditions on the lattice instead of periodic boundaries. Open boundaries lift the topological barrier, and hence allow for a better topological sampling. This approach, however, introduces a boundary zone which need to be considered in any measurements. In ref. [127] we have analyzed the influence of the temperature on the boundary effects and observed a noticeable temperature dependence. Another method was proposed in ref. [128, 129], where the authors propose to suppress the topological freezing by applying a reweighting technique in terms of the topological charge.

4.7 The gradient flow

Due to high-frequency fluctuations of the gauge fields on the lattice, the quantities considered in this work are noisy. Therefore, some noise reduction method is required. The gradient flow [33–36] is a method which has a valid definition in terms of the continuum field theory. Additionally, we have a good analytical understanding of how it affects the gauge fields. This makes it more preferable to traditional approaches such as cooling [130, 131], or APE [132], stout [133], or HYP [134] smearing. Since its introduction, the gradient flow has proven to be useful for a variety of issues in lattice QCD [135–139]. For example it is useful for investigating how instantons emerge in the continuum limit of lattice QCD [7, 140–142]. The field smearing nature of the gradient flow enables the creation of smooth gauge configurations on the lattice, from which a well-defined topological charge can be obtained [35, 143]. Additionally, it also accelerates the convergence of the conjugate gradient, which we use for calculating eigenvalues.

By introducing an extra coordinate τ_F, called the flow-time, the flow fields $B_\mu(x,\tau_\mathrm{F})$ are defined by

$$\dot{B}_\mu(x,\tau_\mathrm{F}) = D_\nu G_{\nu\mu}(x,\tau_\mathrm{F}), \qquad B_\mu(x,\tau_\mathrm{F})|_{\tau_\mathrm{F}=0} = A_\mu(x),$$
$$D_\mu = \partial_\mu + [B_\mu(x,\tau_\mathrm{F}),\ \cdot\ ], \tag{4.76}$$

where the dot denotes a derivative as a function of the flow-time τ_F. $G_{\nu\mu}(x,\tau_\mathrm{F})$ is the field strength tensor on the field $B_\mu(x,\tau_\mathrm{F})$,

$$G_{\mu\nu}(x,\tau_\mathrm{F}) = \partial_\mu B_\nu(x,\tau_\mathrm{F}) - \partial_\nu B_\mu(x,\tau_\mathrm{F}) + [B_\mu(x,\tau_\mathrm{F}), B_\nu(x,\tau_\mathrm{F})]\,. \tag{4.77}$$

It has been shown that the transformations are invertible, which means that the flow equation can be integrated backwards from $\tau_\mathrm{F} \to 0$ [35]. Gradient flow can be understood as a modification of the operators used in the measurement, replacing the elementary links with unitarized averages over many paths, an extreme form of the use of "fat links". The smearing radius can be obtained by considering leading order perturbation theory in the bare coupling g. The linear term of the smeared field $B_\mu(x,\tau_\mathrm{F})$ can be expressed in terms of the initial field $B_\mu(x,0) = A_\mu(x)$ by [35]

$$B_{\mu,1}\left(\tau_\mathrm{F},x\right) = g\int d^4x\, K_{\tau_\mathrm{F}}\left(x-y\right) A_\mu\left(y\right), \tag{4.78}$$

where K_{τ_F} is given by the Gaussian distribution,

$$K_{\tau_\mathrm{F}}\left(z\right) = \frac{e^{-z^2/4\tau_\mathrm{F}}}{\left(4\pi\tau_\mathrm{F}\right)^2}. \tag{4.79}$$

From this distribution we can read off the mean-square of the variance $\sigma^2 = 2\tau_\mathrm{F}$ in four dimensions. At lowest perturbative order the radius is therefore equivalent to replacing the gauge fields with their averages over a Gaussian with width $2\sigma = \sqrt{8\tau_\mathrm{F}}$.

Correlation functions constructed from these smoothed fields turn out to have a well-defined continuum limit [35]. Furthermore, the flow suppresses the ultra-violet contributions of correlation functions at distances $\tau \lesssim \sqrt{8\tau_\mathrm{F}}$, e.g. by reducing the spectral weight in the spectral representation of the correlator. On the other hand, the flow leaves its infrared contributions at distances $\tau \gg \sqrt{8\tau_\mathrm{F}}$ almost unchanged [144], which corresponds to an invariant spectral function in that regime. Thus, it also becomes a great tool for investigating the transport properties of the QGP, such as

for example the sphaleron rate.

The lattice formulation of the flow equation (4.76) in terms of the gradient flow link variables $V_\mu(x,\tau_\mathrm{F})$ of the lattice gauge may be defined by the so called *Zeuthen flow* [145]. It is a specific, a^2-improved lattice discretization of the gradient flow and defined by

$$a^2\left(\partial_{\tau_\mathrm{F}} V_\mu(x,\tau_\mathrm{F})\right) V_\mu^\dagger(x,\tau_\mathrm{F}) = -g^2\left(1+\frac{a^2}{12}\nabla_\mu^*\nabla_\mu\right)\partial_{x,\mu} S_\mathrm{G}^{\text{Symanzik}}[V], \tag{4.80}$$

where $S_\mathrm{G}^{\text{Symanzik}}$ is the Symanzik gauge action as in eq. (4.9). $\partial_{x,\mu}$ is a $\mathfrak{su}(3)$-valued differential operator defined as

$$\partial_{x,\mu}^a f\left(U_\nu(y)\right) = \frac{\mathrm{d}}{\mathrm{d}\epsilon} f\left(U_\nu^\epsilon(y)\right)\Big|_{\epsilon=0}, \quad U_\nu^\epsilon(y) = \begin{cases} \mathrm{e}^{\epsilon T^a} U_\mu(x), & (\nu,y)=(\mu,x) \\ U_\nu(y), & \text{otherwise} \end{cases}, \tag{4.81}$$

and ∇_μ, ∇_μ^* are the lattice forward and backward covariant derivatives

$$a\nabla_\mu f(x) = U_\mu(x) f(x+\hat{\mu}) U_\mu^\dagger(x) - f(x), \tag{4.82}$$

$$a\nabla_\mu^* f(x) = f(x) - U_\mu^\dagger(x-\hat{\mu}) f(x-\hat{\mu}) U_\mu(x-\hat{\mu}). \tag{4.83}$$

In order to implement the Zeuthen flow in our lattice framework, we need to compute the explicit form of the differential operator [122, 145]

$$\mathcal{F}_\mu^\mathrm{S}(x) \equiv -g^2 T^a \partial_{x,\mu}^a S_\mathrm{G}^{\text{Symanzik}}[U], \tag{4.84}$$

which is also known as the *gauge force*. The Symanzik action can be divided into two parts

$$\begin{aligned} S_\mathrm{G}^{\text{Symanzik}} &= \frac{2}{g^2}\sum_{x\in\Lambda}\sum_{\mu<\nu}\left\{\frac{5}{3}\mathrm{Re\,tr}\left(\mathbb{1}-P_{\mu\nu}^{(1\times1)}\right) - \frac{1}{12}\mathrm{Re\,tr}\left[2\cdot\mathbb{1} - \left(P_{\mu\nu}^{(1\times2)} + P_{\mu\nu}^{(2\times1)}\right)\right]\right\} \\ &= \frac{5}{3} S_\mathrm{G}^{\text{Wilson}} - \frac{1}{12} S_\text{rect}. \end{aligned} \tag{4.85}$$

We first carry out the force of the Wilson gauge action. The derivative yields

$$\begin{aligned} \partial_{x,\mu}^a S_\mathrm{G}^{\text{Wilson}} =& \frac{\mathrm{d}}{\mathrm{d}\epsilon} S_\mathrm{G}^{\text{Wilson}}\left[U^\epsilon\right] \\ =& -\frac{2}{g^2}\sum_n\sum_{\sigma<\nu}\mathrm{Re\,tr}\left\{\frac{\mathrm{d}}{\mathrm{d}\epsilon}\left[U_\sigma^\epsilon(n) U_\nu^\epsilon(n+\hat{\sigma}) U_\sigma^{\epsilon\dagger}(n+\hat{\nu}) U_\nu^{\epsilon\dagger}(n)\right]\right\}_{\epsilon=0} \\ =& -\frac{2}{g^2}\mathrm{Re\,tr}\left\{T^a\left[\sum_{\mu<\nu}\left(P_{\mu\nu}^{(1\times1)}(x) + P_{\mu-\nu}^{(1\times1)}(x)\right)\right.\right. \\ & \left.\left. + \sum_{\nu<\mu}\left(P_{\mu\nu}^{(1\times1)}(x) + P_{\mu-\nu}^{(1\times1)}(x)\right)\right]\right\} \\ =& -\frac{2}{g^2}\mathrm{Re\,tr}\left\{T^a\sum_{\nu\neq\mu}\left(P_{\mu\nu}^{(1\times1)}(x) + P_{\mu-\nu}^{(1\times1)}(x)\right)\right\}, \end{aligned} \tag{4.86}$$

where we applied the $\mathfrak{su}(3)$-valued differential operator (4.81) and the chain rule. For our convenience we re-express the sum of the two plaquettes by

$$\partial^a_{x,\mu} S_{\mathrm{G}}^{\text{Wilson}} = -\frac{2}{g^2}\mathrm{Re}\,\mathrm{tr}\left\{T^a \Omega_\mu(x)\right\}, \tag{4.87}$$

$$\text{where} \quad \Omega_\mu(x) \equiv \sum_{\nu\neq\mu}\left(P^{(1\times1)}_{\mu\nu}(x) + P^{(1\times1)}_{\mu-\nu}(x)\right). \tag{4.88}$$

Furthermore, the real trace of $\partial^a_{x,\mu} S_{\mathrm{G}}^{\text{Wilson}}$ yields

$$\begin{aligned}\mathrm{Re}\,\mathrm{tr}\left\{T^a\Omega_\mu\right\} &= \frac{1}{2}\left(\mathrm{tr}\left\{T^a\Omega_\mu\right\} + \mathrm{tr}\left\{T^a\Omega_\mu\right\}^\dagger\right)\\ &= \frac{1}{2}\left(\mathrm{tr}\left\{T^a\Omega_\mu\right\} + \mathrm{tr}\left\{\Omega_\mu^\dagger T^{a\dagger}\right\}\right)\\ &= \frac{1}{2}\mathrm{tr}\left\{T^a\left(\Omega_\mu - \Omega_\mu^\dagger\right)\right\}.\end{aligned} \tag{4.89}$$

The force of the Wilson gauge action thus reads

$$\begin{aligned}\mathcal{F}^{\mathrm{W}}_\mu(x) &= -g^2 T^a \partial^a_{x,\mu} S_{\mathrm{G}}^{\text{Wilson}} = T^a \mathrm{tr}\left\{T^a\left(\Omega_\mu(x) - \Omega_\mu^\dagger(x)\right)\right\}\\ &= \frac{1}{2}\left(\Omega_\mu(x) - \Omega_\mu^\dagger(x)\right) - \frac{1}{6}\mathrm{tr}\left(\Omega_\mu(x) - \Omega_\mu^\dagger(x)\right)\mathbb{1}.\end{aligned} \tag{4.90}$$

Carrying out the derivative of the rectangular part of the action is analogue to the Wilson part [122, 145]

$$\begin{aligned}\partial^a_{x,\mu} S_{\text{rect}} = -\frac{2}{g^2}\mathrm{Re}\,\mathrm{tr}\Bigg\{T^a \sum_{\nu\neq\mu}\Big(&P^{(1\times2)}_{\mu\nu}(x) + P^{(2\times1)}_{\mu\nu}(x) + P^{(1\times2)}_{\mu-\nu}(x) + P^{(2\times1)}_{\mu-\nu}(x)\\ &+\tilde{P}^{(1,2)}_{\mu\nu}(x) + \tilde{P}^{(1,2)}_{\mu-\nu}(x)\Big)\Bigg\},\end{aligned} \tag{4.91}$$

where we introduced another type of rectangle

$$\tilde{P}^{(1,2)}_{\mu\nu}(x) = U_\mu(x)U_\nu(x+\hat{\mu})U^\dagger_\mu(x+\hat{\nu})U^\dagger_\mu(x+\hat{\nu}-\hat{\mu})U^\dagger_\nu(x-\hat{\mu})U_\mu(x-\hat{\mu}). \tag{4.92}$$

By re-expressing the sum of rectangles as

$$\Theta_\mu(x) \equiv \sum\left(P^{(1\times2)}_{\mu\nu}(x) + P^{(2\times1)}_{\mu\nu}(x) + P^{(1\times2)}_{\mu-\nu}(x) + P^{(2\times1)}_{\mu-\nu}(x) + \tilde{P}^{(1,2)}_{\mu\nu}(x) + \tilde{P}^{(1,2)}_{\mu-\nu}(x)\right), \tag{4.93}$$

and eliminating the real part in the same way as in the derivative of the Wilson action, the rectangular force reads

$$\mathcal{F}^{\text{rect}}_\mu(x) = \frac{1}{2}\left(\Theta_\mu(x) - \Theta^\dagger_\mu(x)\right) - \frac{1}{6}\mathrm{tr}\left(\Theta_\mu(x) - \Theta^\dagger_\mu(x)\right)\mathbb{1} \tag{4.94}$$

Putting all together, we get for the Symanzik force

$$\begin{aligned}\mathcal{F}^{\mathsf{S}}_\mu(x) =& \frac{1}{2}\left[c_0\left(\Omega_\mu(x) - \Omega^\dagger_\mu(x)\right) + c_1\left(\Theta_\mu(x) - \Theta^\dagger_\mu(x)\right)\right]\\ &- \frac{1}{6}\mathrm{tr}\left[c_0\left(\Omega_\mu(x) - \Omega^\dagger_\mu(x)\right) + c_1\left(\Theta_\mu(x) - \Theta^\dagger_\mu(x)\right)\right]\end{aligned} \tag{4.95}$$

However, we still need to apply the covariant derivatives from the Zeuthen flow equation

$$\tilde{\mathcal{F}}^{\mathrm{S}}_{\mu}(x) = \left(1 + \frac{a^2}{12}\nabla^*_{\mu}\nabla_{\mu}\right)\mathcal{F}^{\mathrm{S}}_{\mu}(x). \tag{4.96}$$

Using the definition of the covariant derivatives (4.82), we obtain

$$\begin{aligned} a\nabla_{\mu}\mathcal{F}^{\mathrm{S}}_{\mu}(x) &= U_{\mu}(x)\mathcal{F}^{\mathrm{S}}_{\mu}(x+\hat{\mu})U^{\dagger}_{\mu}(x) - \mathcal{F}^{\mathrm{S}}_{\mu}(x) \\ a^2\nabla^*_{\mu}\nabla_{\mu}\mathcal{F}^{\mathrm{S}}_{\mu}(x) &= U_{\mu}(x)\mathcal{F}^{\mathrm{S}}_{\mu}(x+\hat{\mu})U^{\dagger}_{\mu}(x) - 2\mathcal{F}^{\mathrm{S}}_{\mu}(x) + U^{\dagger}_{\mu}(x-\hat{\mu})\mathcal{F}^{\mathrm{S}}_{\mu}(x-\hat{\mu})U_{\mu}(x-\hat{\mu}). \end{aligned} \tag{4.97}$$

Finally, the total force of the Zeuthen flow reads

$$\tilde{\mathcal{F}}^{\mathrm{S}}_{\mu}(x) = \frac{5}{6}\mathcal{F}^{\mathrm{S}}_{\mu}(x) + \frac{1}{12}\left[U_{\mu}(x)\mathcal{F}^{\mathrm{S}}_{\mu}(x+\hat{\mu})U^{\dagger}_{\mu}(x) + U^{\dagger}_{\mu}(x-\hat{\mu})\mathcal{F}^{\mathrm{S}}_{\mu}(x-\hat{\mu})U_{\mu}(x-\hat{\mu})\right]. \tag{4.98}$$

In the numerical integration the flow-time τ_{F} can be increased in infinitesimally small steps, which ensures that the length scale over which the gauge fields are smeared can be appropriately adjusted as we vary the lattice spacing and/or temperature. At zero temperature, the flow-time has to be sufficiently large to get rid of ultraviolet noise but typically has to be smaller than $1/\Lambda^2_{QCD}$. For non-zero temperature, additionally, one has to choose the flow-time such that $\sqrt{8\tau_{\mathrm{F}}}$ is smaller than $1/T$ [87].

The Zeuthen flow has been implemented in a multi-GPU framework called *Parallel-GPUCode*. The development of that framework was also part of this work and will be discussed in the next chapter. The integration of the Zeuthen flow equation is done with a 3$^{\mathrm{rd}}$ order Runge-Kutta algorithm for Lie groups [35, 146, 147] and an adaptive step-size method [148, 149].

Chapter 5

The ParallelGPUCode

The main task in lattice QCD is to compute very large dimensional integrals numerically on computers. In order to perform these calculations, code has to be developed which implements the desired lattice QCD equations. Therefore, a substantial amount of work in this field is dedicated to programming. A major task in this work was to develop a new code base for carrying out lattice QCD calculations. It is called the *ParallelGPUCode*.

The name lattice QCD comes from the fact that the space over which these integrals are performed is a n-dimensional lattice. With the increasing demand for better precision, the size of these lattices has become quite large. Currently, we are at a stage where the required sizes are so large that the time needed for a naive sequential calculation on these lattices is not feasible anymore. Thus, methods from *high performance computing* (HPC) have been utilized to increase performance. The computations have been parallelized by domain-decomposing the lattices into sub-lattices equal to the number of available processes. The computations are then performed on each sub-lattice in parallel.

Nowadays, there are many different programming languages, and each language has its advantages and disadvantages. The ParallelGPUCode has been developed in C++. In that programming language, developers have the advantage to write code which is closely tied to hardware level. Direct access to hardware facilities enables many different possibilities to optimize code. Therefore, it is very well suited for HPC, and thus also for lattice QCD applications.

The underlying hardware which is used for these calculations plays an important role. The two important processor types on which these calculations can be made are CPUs and GPUs. Even though CPUs nowadays can have up to 64 cores on which these computations can be performed in parallel, their main strength is that they perform sequential calculations very fast. GPUs on the other hand are compared to CPUs rather slow in sequential calculations. However, they have thousands of processing cores, which allows for a lot more parallelization. In general, CPUs are optimized for latency, while GPUs are optimized for throughput. Code which is written for CPUs is highly different from code which written for GPUs. Moreover, GPUs of different vendors use different programming languages. There are some libraries (e.g. OpenCL) which can be used to produce hardware independent code, but these come along with a loss of performance if one does not implements cumbersome vendor-specific extensions.

The ParallelGPUCode has been chosen to run mainly on multiple NVIDIA GPUs. Therefore, GPU kernel have been written in the CUDA programming language. The code is also able to run parallel calculations on CPUs, but that has not yet been optimized with respect to memory layout and CPU vectorization at the time of this thesis.

During the development of that code, several technical issues had to be tackled: Solutions for different indexing problems had to be found; communication interfaces

which work with different hardware implementations needed to be programmed; data synchronization had to be guaranteed at any time and many more concepts had to be developed to establish a stable multi-GPU interface. Consequently, all these implementations came along with a dramatic increase of source code. A very common problem with large growing source code is that it may become quite unreadable if no clear structure is given from the beginning. Since unreadable code is usually very error prone and makes it hard to extend the existing feature set, a concept had to be found that provides ways to structure code, such that complex algorithms are more readable.

This is another reason why C++ has been chosen to develop this framework. C++ is a *object-oriented programming* (OOP) language, which enables the programmer to write source code in a more human-readable fashion. The idea is to structure the code by encapsulating algorithms into meaningful objects. These objects can carry attributes and methods, which can also be inherited by other objects. The *ParallelGPUCode* relies heavily on these features, which means that developers need to understand C++ at least on an intermediate level.

The goal was to establish a multi-GPU interface, which makes it easy for developers to implement lattice QCD calculations, while still providing the best possible performance.

5.1 Memory management

Data used to perform calculations are usually loaded into computer memory (the *random-access memory*). The CPU as well as the GPU have their own memory. The maximum amount of data which can be stored inside the memory is finite. Therefore, any program which processes large amounts of data needs to keep track of its memory usage in order to not exceed the hardware capabilities. Lattice QCD applications belong to that category. Computations in lattice QCD can rapidly reach these hardware limits, since the memory which is needed basically scales with the size of the lattice. Thus, the goal of this section is to introduce the concept and the advantages of having a centralized memory management unit.

The ParallelGPUCode so far consists of more than 100 different classes. The majority of these classes operate at the back-end, which is not visible in the main program. Many of these classes, including the main, need to allocate and deallocate memory for different purposes. Some classes, like for example the `Gaugefield` class, need to allocate memory for the entire runtime of the program. In these cases the memory just needs to be allocated at initialization of the program and cleaned up when the program finishes. Other routines like the reduction or communication methods, however, need to allocate a lot of memory only temporarily every time when these methods are called.

The `MemoryManagement` class has been developed to manage dynamically allocated memory in the code. If dynamic memory is needed, the way to get it is to ask the `MemoryManagement` class to allocate memory. For a memory request, one needs to pass a string which is used to identify that allocation. Optionally, one can add the prefix `"SHARED_"` to the name, which states whether this memory allocation can be shared among several routines or not. This is especially useful when large amounts of temporary memory are needed very frequently in different objects. Internally the `MemoryManagement` class will count how many objects are still using this shared allocation. After requesting memory, the `MemoryManagement` class will return a copy of a `gMemoryPtr` object. The `gMemoryPtr` is a smart pointer which owns all information

of a memory allocation, i.e. the actual pointer, the size and whether it is located on the CPU or GPU memory. If this smart pointer gets destroyed, for example by leaving a scope or if the sharing counter has reached zero, the allocated memory will automatically be freed. This way, routines which need memory become less error prone, since memory leaks are effectively impossible. Also redundant clean up code does not need to be implemented which otherwise would deteriorate code quality for example in highly nested algorithms. The `MemoryManagement` class has been made static and globally accessible, such that only one instance of it can exist at any time. This has the advantage that this class always knows where and how much memory is currently allocated. To obtain an overview of all allocations, the `MemoryManagement` class offers a method that prints a summary of all names which have been used to allocate memory as well as their sizes.

5.2 Indexing

In lattice QCD, we want to discretize the (gauge / spinor) fields on a Cartesian isometric 4-dimensional lattice. However, since computer memory usually can only be addressed linearly (there is an exception in CUDA, but we will not address it here), we need an indexing method that translates 4-dimensional coordinates into 1-dimensional coordinates. A simple approach would be the lexical index:

$$\text{index}_{lex} = x + y \cdot N_\sigma + z \cdot N_\sigma^2 + t \cdot N_\sigma^3, \tag{5.1}$$

where $0 \leq x, y, z \leq N_\sigma - 1$ and $0 \leq t \leq N_\tau - 1$. In many numerical tasks (like for example for Krylov solvers) it is useful to split the indices such that all even coordinates are located in the first half of the memory and odd indices are located in the second half. This can be achieved by rewriting equation (5.1) using modular arithmetic:

$$\begin{aligned} \text{index}_{e/o} =& (x + y \cdot N_\sigma + z \cdot N_\sigma^2 + t \cdot N_\sigma^3)/2 \\ &+ N_\sigma^3 \cdot N_\tau/2 \cdot (x + y + z + t) \bmod 2. \end{aligned} \tag{5.2}$$

For gauge fields it makes sense to index the links instead of the sites. This can for example be achieved by adding an additional dimension μ:

$$\begin{aligned} \text{index}_{e/o} =& (x + y \cdot N_\sigma + z \cdot N_\sigma^2 + t \cdot N_\sigma^3)/2 \\ &+ N_\sigma^3 \cdot N_\tau/2 \cdot (x + y + z + t) \bmod 2 + \mu \cdot N_\sigma^3 \cdot N_\tau. \end{aligned} \tag{5.3}$$

This structure ensures that all links which point in the same direction are next to each other in memory, which is advantageous for many calculations on the gauge field.

These indexing methods on their own would be enough to address all sites of the lattice that are located in the memory of a single process. However, the ParallelGPUCode should utilize multiple processors by distributing the lattice across them. Hence, when the lattice is distributed across `N` processes, also `N` different arrays located in the memory of each process need to be indexed.

Another issue is that most computations are stencil based operations, which means that a calculation on a certain lattice site needs to have information about neighboring sites. At the boundaries of the sub-lattices this is problematic because some information of the neighboring sides are stored in the memory of other processors. A solution to that problem is to communicate the information between these sub-lattices. One way to do that is to introduce so-called *inner and outer halos* (see figure 5.1).

Inner halos are referred to the sites which are located at the border of a sub-lattice. *Outer halos* on the other hand are artificial extensions of the sub-lattices beyond their borders. They hold copies of the *inner halos* of neighboring sub-lattices, which are communicated before (or while) computations are made. The depth of these halos depends on the "radius" of the stencil computation. In the ParallelGPUCode the *inner and outer halos* have been made part of the sub-lattice. This means that the sites of a sub-lattice are indexed by:

$$\begin{aligned}\text{index}_{e/o} =& (x + y \cdot N_x + z \cdot N_x N_y + t \cdot N_x N_y N_z)/2 \\ &+ N_x N_y N_z N_t/2 \cdot (x + y + z + t) \bmod 2,\end{aligned} \tag{5.4}$$

or for indexing link variables:

$$\begin{aligned}\text{index}_{e/o} =& (x + y \cdot N_x + z \cdot N_x N_y + t \cdot N_x N_y N_z)/2 \\ &+ N_x N_y N_z N_t/2 \cdot (x + y + z + t) \bmod 2 + \mu \cdot N_x N_y N_z N_t,\end{aligned} \tag{5.5}$$

with

$$\begin{aligned}N_i =& N_\sigma + H_i, \qquad i \in x, y, z, \\ N_t =& N_\tau + H_t,\end{aligned}$$

where H_i, H_t are the halo depths in different directions. Note that depending on how the lattice is split, lattice fields do not necessarily have halos in all directions, such that halo depths in some directions may be zero.

By using the index computation (5.4) due to the additional halo offset, the maximum amount of sub-lattice indices change as well. This in turn affects all loops which iterate through the sub-lattices. Every loop will depend on the halo size and splitting of the lattice field, since all physics-related calculations should only iterate over the non-halo part of the lattice (i.e. loops have to skip the additional halo sites). In the ParallelGPUCode these iterations have been automated and hidden behind the functor syntax, which will be discussed in section 5.3.

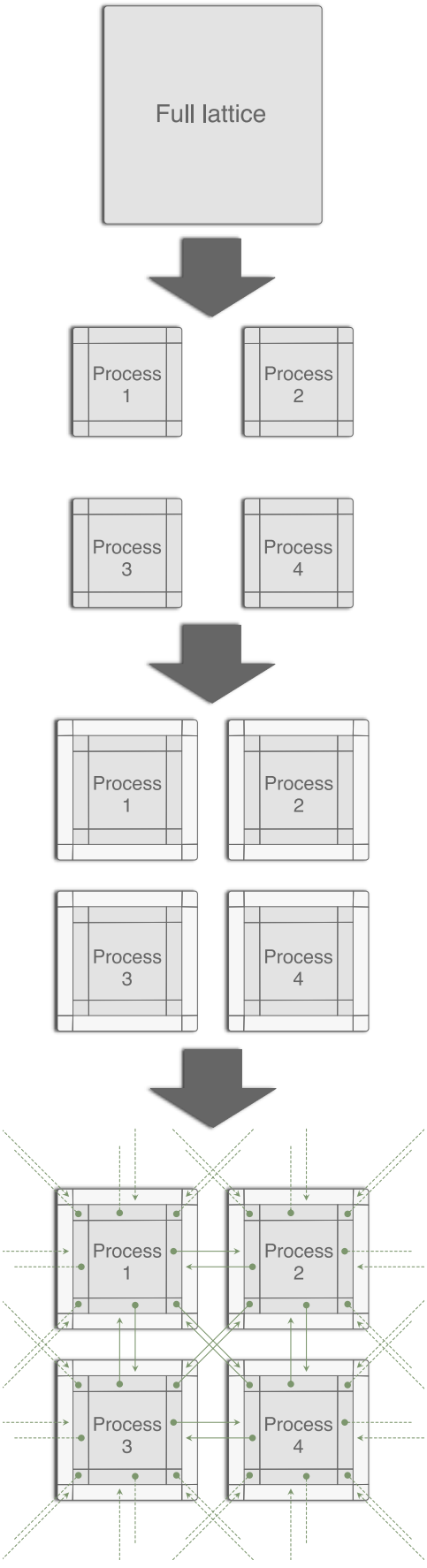
Another problem arises when different lattice fields require different halo sizes. If two fields have different halo sizes, their indices computed by equation (5.4) will not correspond to the same 4-dimensional coordinate. This would be an issue for example in the computation of the plaquette action. The gauge field on which the plaquette is computed needs halos, but the field in which the intermediate results (the summands) are stored does not require halos. Thus, the index of the gauge field can not be reused for the intermediate field, such that additional code is required for translating indices and coordinates of the two fields. This can deteriorate the code quality dramatically. In order to keep the lattice QCD algorithms as readable as possible, this additional code has been automated and encapsulated inside the different `accessor` classes. With these accessors one can set or get elements at the desired sites or links of the various lattice fields (`Gaugefield`, `Spinorfield`, `LatticeContainer`, ...). Every field type has its own custom accessor type. Due to the previously described problems these accessors do not accept an integer index computed by (5.4) as a parameter. Instead they need `gSite` objects. These objects carry more information that is required by the accessors, such as the `bulkindex` (index computed with halo size zero) or the `fullindex` (index computed with non-zero halo size) of the site/link as well as the actual 4-dimensional coordinates. `gSite` objects are provided in each iteration through the lattice field, due to the functor syntax (see section 5.3). Also all site/link arithmetic methods that are provided by the `GIndexer` class return `gSite` objects. In the ParallelGPUCode it is mandatory to work only with `gSite` objects instead of bare integers. Otherwise, the entire multi process functionality could break.

In addition to the `gSite` object, also `gSiteMu` and `gSiteStack` objects exist, which inherit from `gSite`. `gSiteMu` additionally holds the `mu`-direction of a gauge field, while `gSiteStack` holds the *stack number* of a spinorfield.

How data that should be communicated is stored in memory has an effect on performance. The best performance can be achieved when the data is contiguously stored inside a single buffer. If the data is distributed in `N` different locations in memory, also `N` different send/receive calls would be required to perform the communication. Clearly, the performance would suffer if that is the case. Unfortunately, with the indexing structure above, the relevant sites/links inside the halo area which should be communicated are not lying contiguously in memory. Thus, the halos have to be copied into single separate buffers, before communication takes place. Also after the communication, the transferred halos have to be copied back into the field. In a Cartesian grid, a 4-dimensional cube has 80 different halo segments with different sizes which can be sent in different directions. Table 5.1 lists all the different halo shapes that encapsulate a 4-dimensional cube. All these different halo shapes require a custom indexer in order to identify the sites/links in the halo buffers. The `HIndexer` class offers that possibility. It indexes the sites/links according to equation (5.1) using the dimensions given in table 5.1. Additionally, in order to be able to copy the sites/links back and forth between halo buffer and lattice field, routines are required which translate the halo buffer indices into lattice field indices. These routines have been implemented by the methods `ExtractInnerHaloSeg` and `InjectOuterHaloSeg`. `ExtractInnerHaloSeg` copies a single halo segment from the inner halo of the lattice field into the halo buffer, while `InjectOuterHaloSeg` copies a single halo segment from the halo buffer into the outer halo of the lattice field.

Type No.	Name	Direction	Count	Dimension
1	Hyperplane	X	2	$N_y \times N_z \times N_t \times H_x$
2	Hyperplane	Y	2	$N_x \times N_z \times N_t \times H_y$
3	Hyperplane	Z	2	$N_x \times N_y \times N_t \times H_z$
4	Hyperplane	T	2	$N_x \times N_y \times N_z \times H_t$
5	Plane	XY	4	$N_z \times N_t \times H_x \times H_y$
6	Plane	XZ	4	$N_y \times N_t \times H_x \times H_z$
7	Plane	XT	4	$N_y \times N_z \times H_x \times H_t$
8	Plane	YZ	4	$N_x \times N_t \times H_y \times H_z$
9	Plane	YT	4	$N_x \times N_z \times H_y \times H_t$
10	Plane	ZT	4	$N_x \times N_y \times H_z \times H_t$
11	Stripe	XYZ	8	$N_t \times H_x \times H_y \times H_z$
12	Stripe	XYT	8	$N_z \times H_x \times H_y \times H_t$
13	Stripe	XZT	8	$N_y \times H_x \times H_z \times H_t$
14	Stripe	YZT	8	$N_x \times H_y \times H_z \times H_t$
15	Corner	$XYZT$	16	$H_x \times H_y \times H_z \times H_t$

TABLE 5.1: Different types of halo segments on a 4-dimensional cube. N_x, N_y, N_z and N_t are the lattice dimensions. H_x, H_y, H_z and H_t are the halo depths.



Sketch of a two dimensional lattice which should be parallelized on 4 processes.

Splitting up the previous lattice into 4 sub-lattices across 4 processes. The boundary area are called *inner halos* in the ParallelGPUCode. This area is necessary for stencil operations on the lattice.

In order to make these *inner halos* available for the neighboring processes the sub-lattices have to be enlarged. In that enlarged area (the white surfaces) the *inner halos* from the neighboring sub-lattices should be copied. This enlarged area is referred to as the *outer halo* in the ParallelGPUCode.

Communication between all processes has to be established. The green arrows show which *inner halo* elements are copied into the *outer halo* elements of the neighboring processes. The green dotted arrows represent the communication due to periodic boundary conditions.

FIGURE 5.1: Halo pattern of a 2-dimensional lattice field.

5.3 Parallelizing using functor syntax

The way that computations are done on the lattice follows most of the time the same rules:

- Iterate through all lattice sites/links.
- Compute a quantity on each site/link.
- Eventually sum up results of all sites/links.

These steps can be implemented in different ways. The easiest way would be to just iterate sequentially in a loop over all sites and perform the required computation on each site in each iteration. However, on modern computer this way would be quite inefficient. A better way would be to parallelize that iteration. Depending on the processor (CPU or GPU) different options are available. Nowadays, a single CPU can have up to 64 cores. Additionally, HPC centers usually have multiple compute-nodes connected to each other, such that CPUs on different nodes are also able to communicate with each other. In principle each of these cores can be dedicated to iterate over a certain sub-lattice. Depending on the task, that can speed up the computation significantly. If the total amount of cores would be N, then in an ideal situation a speed up of factor N could be realized. In fact, due to hardware limitations and communication overhead, this speed up is never exactly N. The goal is to bring the speedup as close to N as possible. If each core would also have access to a GPU, then the computation could be further parallelized, since GPUs have thousands of cores.

To ensure that a parallelized routine works on any possible combination of lattice splitting, number of processors and whether or not a GPU is used, the implementation can become quite complex. As a consequence the actual lattice calculation can easily get lost inside highly nested parallelization routines. This makes the code barely readable and highly error prone. For that reason the concept of *functors* has been introduced. With that concept the entire iteration and parallelization procedure is hidden inside an abstract base class from which all fields inherit. To perform a lattice calculation which changes the values on each site, the lattice equation just needs to be wrapped inside a `struct` and then be passed to the field. The field then automatically figures out how to parallelize the lattice and performs the computation. Alternatively, standalone iteration functions exist that can be used for field independent tasks. The classes which are passed to the fields are referred to as *functors*. Every field supports a variety of iteration methods. The simplest iterator available iterates over the full sub-lattice, i.e all sites/links including the halo sites/links. This iterator should only be used if no information about neighboring sites/links is needed. An example would be if a field just needs to be set to a particular value. This way, no halo communication is required afterwards, which would otherwise affect the performance. Another available iterator is iterating only over the bulk part of the sub-lattices, i.e all sites/links excluding the halo sites/links. In all stencil-based calculations this is the preferred iterator, since in such calculations the halo region needs to be communicated in the next step anyway, i.e. the halo region will be overwritten. Depending on the underlying field many more iterators exist for different tasks. The implementation of these iterators are comparably easy, such that additional iterators can be implemented with minimal effort.

5.4 Communication techniques

As has been mentioned in the previous sections, some kind of communication is necessary in order to perform a parallelized lattice computation. There are different ways how communication can be achieved. Among them there is:

1) One-sided communication,

2) Two-sided communication.

With one-sided communication the sending process has access to an assigned part of the memory of the receiving process. Communication is then achieved by copying some data of the memory of the sending process into the memory of the receiving process. In that procedure only one of the processes is actively participating (the sending process), such that the communication is performed from *one side* only.

In two-sided communication both processes have to participate. This means that every time before some information should be exchanged, the sending process has to send a communication request while the receiving process has to actively wait for an incoming communication request. As long as this *handshake* procedure has not been finished, no information can be exchanged.

This is also the biggest disadvantage of two-sided communication compared to one-sided communication, since that handshake procedure is effectively slowing down the communication. On the other hand since both processes are participating in two-sided communication, it is easier to synchronize both processes. With the direct access to the memory of the receiving process in case of one-sided communication, the receiving end is not notified when the copy process is finished as long as the notification has not been performed manually.

A library which supports both communication types is for example the *Message Passing Interface (MPI)*. The ParallelGPUCode is only taking advantage of the two-sided communication feature of MPI, since current MPI implementations are not very well optimized for one-sided communication. However, another technique is used for peer-to-peer communication between GPUs: *Cuda GPUDirect P2P*. This is an one-sided communication interface, which allows a GPU to have read and write access to the memory of another GPU.

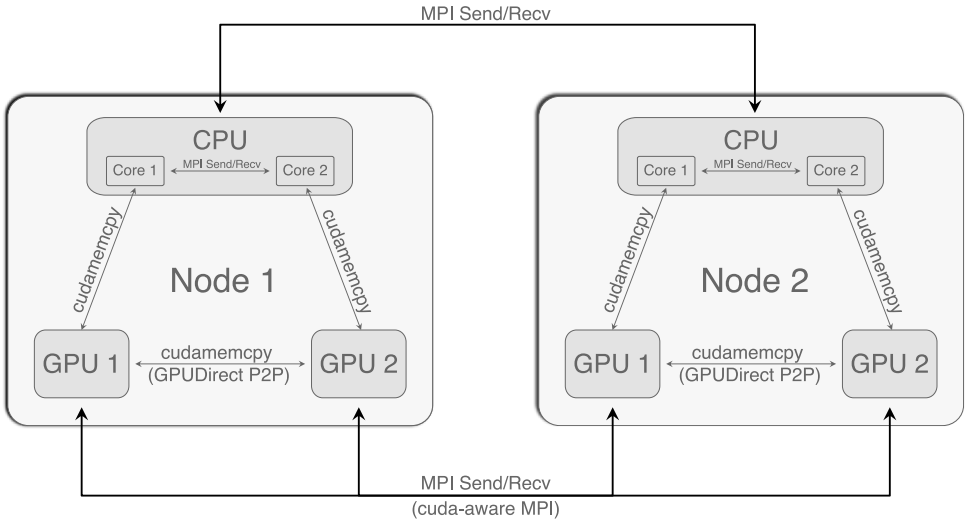


FIGURE 5.2: Illustration showing different communication channel.

The ParallelGPUCode is supposed to run on different systems with different hardware specifications (e.g. number of nodes, GPUs per node, connection between GPUs etc.), such that it adapts its communication system to the hardware capabilities. The logic for that is implemented in the `CommunicationBase` class. Figure 5.2 illustrates the topology of an fictitious system which consists of two nodes with two GPUs per node. The arrows represent all possible communication channels. The `CommunicationBase` always automatically figures out the fastest path between the different processors (as long as this feature is not turned off manually during compilation). If an application is purely running on the GPU, then the fastest path is GPUDirect P2P for intranode communication and cuda-aware MPI for internode communication. If GPUDirect P2P and cuda-aware MPI is not available or turned off, then the entire communication is performed over the CPUs. This means that the data is first copied from the GPU to the CPU, then communicated over MPI to another CPU and then copied from the receiving CPU to the desired GPU. Consequently, this slows down the application, since the communication overhead is much higher.

In case of halo communication, the method used affects the memory usage on the GPU and CPU. If GPUDirect P2P or cuda-aware MPI communication (or both) is turned on, then 2 halo buffers are required on the GPU. One buffer acts as a sending buffer and the other one as a receiving buffer. If the two communication methods are turned off, then only one halo buffer is required on the GPU, since the sending and receiving buffer are outsourced to the CPU memory. Therefore, in case memory on the GPU is not enough memory and performance is not so important, turning off GPUDirect P2P or cuda-aware MPI communication might be a good compromise.

5.5 SiteComm and derived classes

With the memory management class, the indexing classes, the halo injection/extraction methods, the communication classes, the functors and the accessors, we have enough to be able to communicate halo segments of any field which can be represented by a Cartesian grid. In lattice QCD, we work with different kind of fields. While a spinor field can hold multiple 3-vectors on each site (thus it is also called stacked spinor field in the ParallelGPUCode), a gauge field connects each site with neighboring sites by SU(3) matrices.

The construction of the halo communication of these two fields would be basically the same. Only the types and sizes of the field elements change. To simplify that construction, the class `SiteComm` has been introduced. It is an abstract class which basically combines all previously mentioned classes and methods into a single `updateAll` method. This method performs the entire halo communication process. During instantiation the `SiteComm` sets up the halo logic and allocates the required halo buffer. Depending on which kind of connections (MPI, cuda-aware MPI or GPUDirectP2P) to the neighboring processes are available, the `SiteComm` automatically chooses the best between them. Additionally, it also overlaps communication with the extraction/injection procedure when possible to accelerate the halo update even further. Classes like the `Gaugefield` or the `Spinorfield` inherit from `SiteComm`, such that they own the `updateAll` method by construction.

Another optimization can be achieved when calling `updateAll`. By passing a `COMM` parameter to that method, it is possible to overlap halo communication with computation outside of the `siteComm`. The idea is to call `updateAll` twice. The first time just to trigger the communication and the second time to synchronize the

communication. In between these calls some different computation can be performed, which does not alter the halo data that is currently communicated (see listing 5.1).

```
// This call will start communication and immediately return.
// Communication is still ongoing after this call!
gaugeFieldObj.updateAll(COMM_START);

// Perform computation which is unrelated
// to the currently communicated halos.
int result = a * x + 100;

// Wait until communication is done with this call.
gaugeFieldObj.updateAll(COMM_FINISH);
```

LISTING 5.1: Asynchronous halo communication

The `SiteComm` is able to communicate the entire surface of a 4-dimensional cube. The halos on that surface can be split into four major shapes: Hyperplane, plane, stripe and corner (see table 5.1). However, sometimes it is not necessary to communicate everything. For example, if a stencil based lattice operator just needs its nearest neighbors that are just connected via one link, only Hyperplanes would need to be communicated. This would save 72 communication calls and therefore increase performance. The `updateAll` method also offers that possibility. In the above case one just needs to pass the parameter `hyperplane` to the `updateAll` method. A combination with `COMM_START/COMM_FINISH` flags can be achieved as well:

```
gaugeFieldObj.updateAll(COMM_START | Hyperplane);
// ... Perform independent computations ...
gaugeFieldObj.updateAll(COMM_FINISH | Hyperplane);
```

More surfaces (`Plane`, `Stripe`, `Corner`) can be included in the same way with the bitwise `OR` operator `|`. If no surfaces are specified, all surfaces are communicated.

The reason why the `SiteComm` has its name is that it actually communicates only data located at sites. Consequently, the `Gaugefield` has to bundle 4 links per site in order to perform a halo update, since it inherits from `SiteComm`. An illustration of that is shown in figure 5.3. Obviously, this could be further optimized by implementing a `LinkComm` class that would communicate links instead of sites. However, in the current version of the code this class is not available.

5.6 Lattice container

When observables are computed in lattice QCD, usually some intermediate results per lattice site need to be computed as well. Followed by that, these intermediate results are then in some way reduced over the whole lattice (for example by a sum over the whole lattice). On a lattice which is split up on multiple processes, the procedure would be a bit more complicated:

1) Compute intermediate results on all sites of all sub-lattices in parallel.

2) Perform a reduction on each process into a single quantity per process.

3) Perform a final reduction across all processes into a single quantity and distribute that quantity over all processes if required.

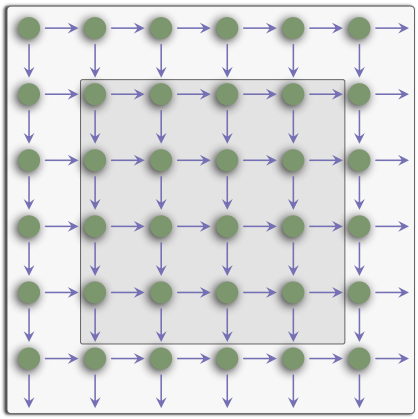


FIGURE 5.3: Illustration showing which data is communicated on a 2D lattice. The outer rectangle represents the inner halo region which is communicated. The green circles represent the sites of the lattice and the arrows represent the links between the sites. A spinor field would just communicate the green circles within the halo region, while a gauge field would communicate the arrows within the halo region.

Since this is a very common task, this procedure has been merged together inside a single reduction call in the `LatticeContainer` class. This class is as the name suggests a container class which can hold any element type and is spread over the processes in a similar way as the `Gaugefield` or `Spinorfield`. This means that the `LatticeContainer` also allocates memory, where intermediate results can be stored. This memory allocation is per default shared with the memory of the halo buffer of the other lattice fields, since in most cases the intermediate results have to be recalculated after a halo update has been performed. Apart from a standard sum over all elements, the `LatticeContainer` class also offers the possibility to reduce only time slices of the lattice, such that the result would be a vector of length N_τ. This is especially useful, if correlation functions in time direction need to be computed. Additionally, there is also the stacked reduction, which can be used to reduce the intermediate results of a stacked spinor. With the current version this class does not support the exchange of halos. However, if it is needed, it could be made available by letting the `LatticeContainer` inherit from `SiteComm`.

5.7 Coalesced memory access

The way the `Spinorfield` or `Gaugefield` class structures the fields in memory has significant performance implications. The `Gaugefield` for example could allocate its array of matrices by using an "array of structures" layout. In this layout each element in the array corresponds to a matrix which is contiguously stored in memory. However, this is inefficient on the GPU because the memory accessed by multiple threads is not contiguous. This results in much more memory access instructions on the GPU. Another way is to allocate an array by using a "structure of arrays" layout. In this layout, the elements of all matrices are stored contiguously in memory, not the matrices themselves. In other words, the first part of the array contains the first elements of all matrices, followed by the next part, which contains the second elements of all matrices, and so on. This layout is much better for coalesced memory access, i.e. simultaneously running threads access memory that is nearby, resulting in much fewer memory access instructions. Therefore, in the ParallelGPUCode the lattice fields

are allocated in the "structure of arrays" layout. The logic behind this is hidden in the `stackedArray` and `accessor` classes.

5.8 Getting started

In order to get started with the ParallelGPUCode, it is very helpful to have a look at the examples in the repository. Also the testing applications can help to understand many concepts of the framework (e.g. the `GeneralFunctorTest`).

Nevertheless, in this section we describe how a very simple application can be programmed. As described in section 5.3, the way how to perform a computation on the lattice is by wrapping the lattice equation inside a `struct` and then passing the `struct` to an iterator of a field. Listing 5.2 shows an example of a `struct` that computes the plaquette action. This kernel should iterate over a `LatticeContainer` object which can be ether located on the CPU or on the GPU memory. The template parameter `onDevice` is a boolean which need to be set accordingly. This means that if the field is on the GPU then `onDevice` should be `true`, otherwise it should be set to `false`.

In order to compute the plaquette, it is necessary to access the link variables of a gauge field. Therefore the constructor need a `Gaugefield` object to initialize the `gaugeAccessor` member of that struct. The lattice equation is implemented in the `operator()` method. The function type qualifiers `__device__` specifies that this method is going to be compiled for the GPU. Similar the function type qualifiers `__host__` specifies that this method will be compiled for the CPU. If both qualifier are given, then the kernel will be compiled for both processors. As runtime arguments the `operator()` provides a `gSite` object which represents the lattice site of the current iteration.

Inside the loop over all directions this site is getting passed to the `getLinkPath` method of the `gaugeAccessor` together with the directions which shape a plaquette. The `getLinkPath` method is accessing the gauge link variables on the given site in the specified directions and multiplying them in the same order as the directions given in the arguments (the amount of direction variables which can be passed to this function is variable). The result is a SU(3) matrix (precisely a `GSU3<floatT>` object) which is then passed to the `tr_d` function. This function computes the real part of the trace of this matrix. Apart from these runtime arguments, it also takes two template parameter. The first template parameter indicates whether we are indexing a field which holds only even, only odd or all sites. The gauge field in this example holds all sites, thus the argument `All` has been passed. The second argument specifies the halo width which is used by the gauge field. This is required because the `getLinkPath` method is calling the indexer internally to access gauge link variables on the gauge field. The reason why the indexer needs to know the halo width has been explained in section 5.2.

After the loop over all directions the result will then just get returned. The iterator of the `LatticeContainer` object is then storing the result automatically inside its array at the same coordinate where the gauge link variables were located in the gauge field.

To perform the iteration from the main, the iterator of the `LatticeContainer` object should be called as follows:

```
gauge.updateAll();
latticeContainerObj.template iterateOverBulk<All, HaloDepth>(
                        CalcPlaq<floatT, true, HaloDepth>(gauge));
```

Since this iteration runs only on the bulk part of the sub-lattices, a halo update should be performed before the iteration starts, in order to make sure that the link variables on the halos are valid. In this example, the `LatticeContainer` object (`latticeContainerObj`) and the `Gaugefield` object (`gauge`) have been created on the GPU. Thus, the template parameter `onDevice` of the kernel have been set to `true` as well.

So far we have only computed the summands of the plaquette action and stored them into the `LatticeContainer`. A final sum can be obtained by calling the `reduce` method of the `latticeContainerObj`, followed by a proper normalization. The full example of that program is available in the example folder of the ParallelGPUCode.

```
template<class floatT, bool onDevice, size_t HaloDepth>
struct CalcPlaq{

  gaugeAccessor<floatT> gaugeAccessor;

  CalcPlaq(Gaugefield<floatT,onDevice,HaloDepth> &gauge)
  : gaugeAccessor(gauge.getAccessor()){
  }

  __device__ __host__ floatT operator()(gSite site) {

    floatT result = 0;
    for (int nu = 1; nu < 4; nu++) {
        for (int mu = 0; mu < nu; mu++) {
            result += tr_d(
              gaugeAccessor.template getLinkPath<All, HaloDepth>(site,
                                                    mu,
                                                    nu,
                                                    Back(mu),
                                                    Back(nu)));
        }
    }

    return result;
  }
};
```

LISTING 5.2: A simple plaquette kernel

5.9 Status summary and benchmarks

At the time of writing this thesis, the ParallelGPUCode is in active use and development by various users in different projects. Therefore this status summary only provides a snapshot of the state of development.

The first fully working and tested program was the gradient flow [35]. It supports the Wilson and Zeuthen flow [145] discretization. The integration is done with a 3rd order Runge-Kutta algorithm for Lie group methods [35, 146, 147] and an adaptive step-size method [149]. Observables or correlation functions can be measured at any flow-time on the fly or later by saving the smoothed configurations. A heat bath [108, 109] and overrelaxation [106, 107] code has been implemented to generate

pure gauge configurations. For configurations with dynamic fermions in the Highly Improved Staggered Quarks discretization (HISQ) [31] a rational hybrid Monte Carlo algorithm (RHMC) [150] has been implemented.

To analyze the scaling behavior of the ParallelGPUCode we need to measure the *speedup*. The speedup of parallel programs is defined by

$$\text{speedup} = T_0/T_N, \tag{5.6}$$

where T_0 is the runtime of the program with the minimal required amount of processors and T_N is the runtime of the program with N processors. In high performance computing there are two common notions of scalability, i.e. strong and weak scaling. With strong scaling, the runtime of a program is measured for a total problem size which is fixed while the number of processors increases. A perfect strong scaling parallel program would for example be twice as fast if the number of used processors is doubled. In weak scaling, on the other hand, the problem size grows with the number of processors, such that the problem size per processor is fixed. The speedup of a perfect weak scaling program would stay constant with increasing number of processors and problem size. In general, due to communication overhead and hardware limitations, perfect scaling can never be reached in strong and weak scaling benchmarks.

We have checked the strong scaling behavior of the gradient flow and the RHMC on the Bielefeld GPU cluster. The previously mentioned problem size in lattice QCD is the size of the lattice. The gradient flow has been applied on a random $64^3 \times 16$ configuration up to a flow-time radius of $\sqrt{8\tau_\mathrm{F}}T = 0.3$ using a fixed step size. At each flow-time step the plaquette and the color electric and magnetic correlation function has been measured. In the RHMC benchmark a single trajectory has been computed on a $64^3 \times 16$ lattice (starting from a random configuration) with physical strange and light quarks masses. In figure 5.4 we show the results of these strong scaling benchmarks. The gradient flow and the RHMC show very good strong scaling behavior with up to 4 GPUs. We observe a speedup of ~ 3.4 with 4 GPUs in both benchmarks. However, going from 4 to 8 GPUs the speedup is rather low. One reason is the total lattice size of $64^3 \times 16$ in both cases. With 8 GPUs, the sub-lattice size is so small that the communication overhead becomes much more significant. Another reason is the NVLink topology of the nodes at the Bielefeld GPU cluster. On each node, groups of four GPUs are fully interconnected with at least one NVLink link. However, these two groups are only connected by four NVLink connections. Hence, with 8 GPUs some pairs of GPUs can only communicate over PCIe express, which is slower.

Similar benchmarks have been performed on Summit in Oak Ridge National Laboratory, USA. There we have checked the multi-node scaling behavior. This time the gradient flow has been applied on a random $96^3 \times 24$ configuration which required at least one full node of six GPUs. In the case of the RHMC, a single trajectory on a 96^4 lattice has been computed. This required in total 36 GPUs which are given by 6 nodes. Since communication between nodes traverses over Infiniband instead of NVLink, we do not expect the same scaling behavior as in the single node benchmarks. In figure 5.5 we show the results of these benchmarks. We see that with 4 nodes the gradient flow still achieves a speedup of ~ 3. With 36 nodes the speedup drops to ~ 15. In case of the RHMC, if we double the nodes from 6 to 12 nodes, we get a speedup of ~ 1.6. If we increase the number of nodes to 36 we just get a speedup of about ~ 2.8. Also here the argument applies that the communication overhead overwhelms the actual lattice computations with too many nodes, due to the small local lattice sizes.

We have also performed a weak scaling test of the gradient flow on the Bielefeld

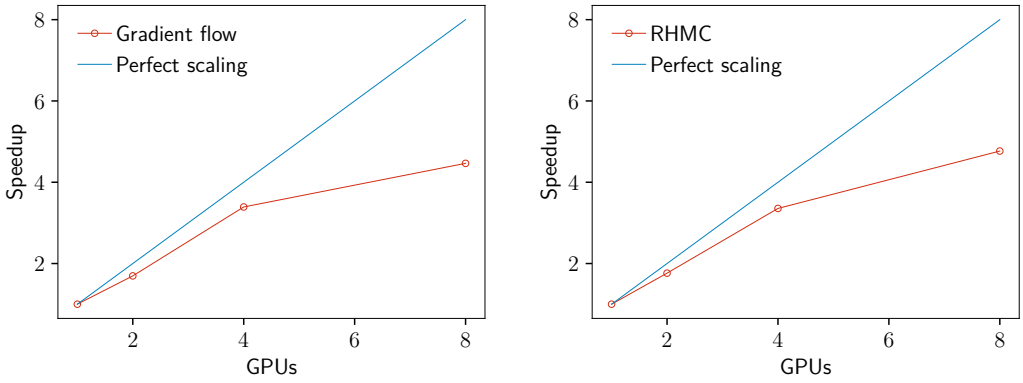


FIGURE 5.4: Single-Node strong scaling plots. These benchmarks have been performed on the Bielefeld GPU cluster. *Left:* Performance of the gradient flow program for a $64^3 \times 16$ lattice. *Right:* Performance of the HISQ gauge field generation program based on a rational hybrid Monte Carlo algorithm (RHMC) for a $64^3 \times 16$ lattice.

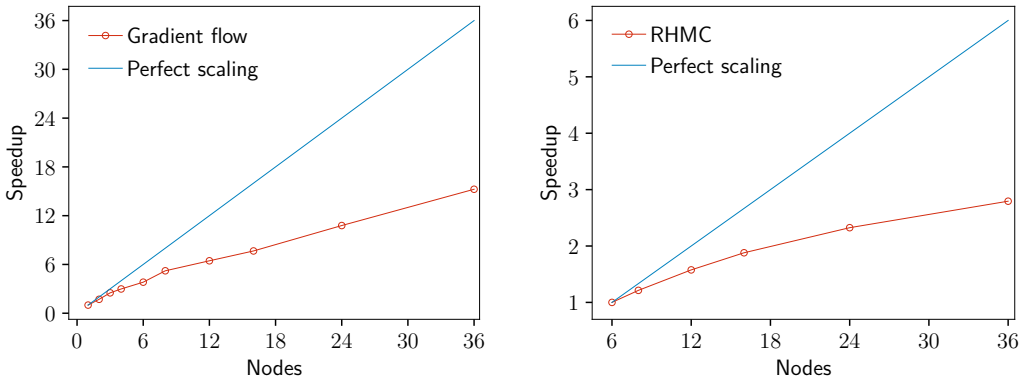


FIGURE 5.5: Multi-Node strong scaling plots. These benchmarks have been performed on Summit in Oak Ridge National Laboratory, USA. *Left:* Performance of the gradient flow program for a $96^3 \times 24$ lattice. *Right:* Performance of the HISQ gauge field generation program based on a rational hybrid Monte Carlo algorithm (RHMC) for a 96^4 lattice.

cluster. The results are shown in the left plot in figure 5.6. The size of the local lattice on a single GPU has been fixed to $64^3 \times 16$. The weak scaling behavior of the gradient flow turns out to be very good. With 2 and 4 GPUs the speedup is at around ~ 0.9. The speedup drops with 8 GPUs to ~ 0.8. This is also due to the already mentioned NVLink topology of the nodes in Bielefeld.

The RHMC is not suitable to perform weak scaling benchmarks, since the number of CG iterations changes with the size of the lattice. Therefore, instead of the full RHMC we have benchmarked 500 iterations of the DSlash, which is the most expensive part of the CG in the RHMC. The runs have been performed also on the Bielefeld GPU cluster. On the right hand side in figure 5.6 we show the results of this benchmark. The speedup of the DSlash is with almost ~ 0.8 for 2,4 and 8 GPUs a bit lower than for the gradient flow, but still in an acceptable region.

Overall, we are convinced that no more significant performance increase can be achieved by further optimizations, so that we are ready to perform multi-GPU production runs with this code.

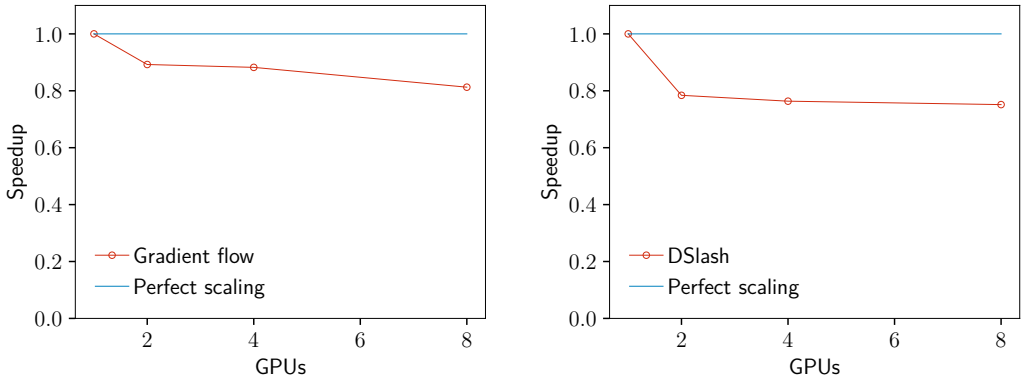


FIGURE 5.6: Single-Node weak scaling plots. These benchmarks have been performed on the Bielefeld GPU cluster. *Left:* Performance of the gradient flow. The local lattice size on a single GPU has been fixed to $64^3 \times 16$. *Right:* Performance of the DSlash. The local lattice size on a single GPU has been fixed to $80^3 \times 20$.

Chapter 6

Topological Susceptibility at Varying Light Quark Masses

As introduced in chapter 3, topology has many interesting implications in QCD. It is closely related to the chiral anomaly and the explicit breaking of $\mathrm{U}_A(1)$. The topological susceptibility measures the fluctuations of the topological charge. In this chapter we take a first look at the topological susceptibility measured on $2+1$ HISQ configurations around the critical temperature T_c.

6.1 Lattice setup

In order to calculate the topological susceptibility, we first need to compute the topological charge. Measuring the topological charge using the gluonic definition, eq. (3.31), is not straightforward. In general, the topological information is hidden behind non-topological noise on the gauge fields. Therefore, a smoothing technique has to be applied to dampen this noise before measuring the topological charge. In this work we choose the gradient flow for this procedure as introduced in section 4.7. It is implemented in the Zeuthen flow discretization using a 3$^{\text{rd}}$ order Runge-Kutta with an adaptive step-size algorithm as an integration method and is part of the ParallelGPUCode. As sea quarks we use $2+1$ HISQ configurations of different masses generated by the HotQCD collaboration [62, 63]. The strange quark mass of these configurations is fixed to its physical value m_s, while the light quark masses are varied such that $m_s/m_l = 27, 40, 80, 160$ corresponding to pion masses of $m_\pi = 135, 110, 80, 55$ MeV respectively. The different configurations as well as the statistics are listed in table 6.1. The lattice implementation of the gluonic definition of the topological charge density $q(x)$ is based on an a^2-improved implementation of the field strength tensor, see eq. (4.74). In contrast to the standard "clover" leaf definition, which involves just a sum of four square plaquettes, we use a combination of square and 1×2 rectangular plaquettes (see section 4.5). By integrating $q(x)$ over the spacetime volume we get the topological charge Q. In eq. (4.75) we then defined the topological susceptibility χ_t as

$$\chi_t = \frac{\langle Q^2\rangle - \langle Q\rangle^2}{V}, \tag{6.1}$$

where V is the four-volume and Q is the topological charge (see eq (4.73)).

The topological charge is known to evolve slowly on finer lattices, i.e. the topological charge freezes in Monte Carlo time. This leads to large autocorrelation times and the requirement for large statistics to sample fluctuations between different topological sectors. To check whether our configurations are affected by this freezing, we plot the trajectory of Q of our finest lattices in figure 6.1. The charges have been extracted

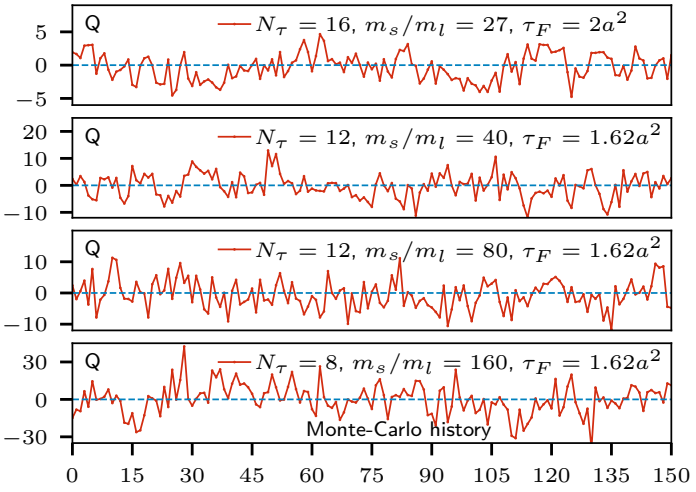


FIGURE 6.1: Trajectory of the topological charge at temperature $T \approx 166$MeV. The x-axis shows the configuration number. Subsequent configurations are separated by 50 hybrid Monte Carlo time steps. Different streams have been concatenated. The topological charge is not always close to integer values.

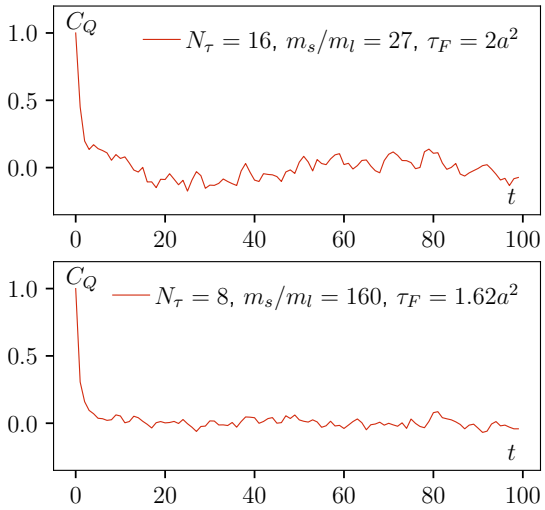


FIGURE 6.2: Autocorrelation times of the topological charge at temperature $T \approx 166$MeV. Subsequent configurations are separated by 50 hybrid Monte Carlo time steps. Different streams have been concatenated.

at a flow-time at which the topological susceptibility settles into a plateau. More details about this plateau criterion will be discussed in the next section. Figure 6.1 shows a large amount of fluctuations, and hence our configurations do not suffer from topological freezing. In figure 6.2 we show the autocorrelation time of the topological charge Q of two lattices. The autocorrelation time of the $N_\tau = 8$ lattice is smaller than 50 Monte Carlo steps. The $N_\tau = 16$ lattice on the other hand shows a very small autocorrelation. However, it should not be a problem, since in our error calculations we choose the block sizes in our Jackknife routine to be at least 40.

In the trajectories in figure 6.2 we noticed that the topological charge is not always close to integer values. This might indicate that the topological charge suffers from strong lattice artifacts.

$m_s/m_l = 27$

N_σ	N_τ	T[MeV]	χ_t/T	#conf
32	8	156	0.89(1)	1196
32	8	166	0.81(1)	1329
32	8	175	0.735(7)	1784
64	16	146	0.69(3)	244
64	16	152	0.624(5)	3928
64	16	158	0.58(2)	207
64	16	163	0.54(2)	597
64	16	167	0.50(2)	831

$m_s/m_l = 80$

N_σ	N_τ	T[MeV]	χ_t/T	#conf
32	8	137	1.05(2)	1200
32	8	140	1.01(2)	1221
32	8	143	0.988(9)	1758
32	8	145	0.97(2)	2341
32	8	147	0.97(2)	2951
32	8	151	0.933(9)	3807
32	8	154	0.89(2)	1200
32	8	162	0.829(9)	1200
32	8	166	0.809(8)	1200
40	8	140	1.03(2)	1200
40	8	143	1.01(2)	1200
40	8	145	1.01(2)	1200
40	8	147	0.94(2)	1200
40	8	151	0.94(2)	1200
40	8	154	0.895(9)	1200
40	8	156	0.88(1)	1200
40	8	162	0.83(2)	1200
40	8	166	0.80(2)	1199
56	8	140	1.05(2)	1199
56	8	143	1.02(2)	1200
56	8	145	1.00(2)	1200
56	8	147	0.97(2)	1199
56	8	151	0.91(2)	1200
56	8	154	0.89(2)	1200
56	8	156	0.86(1)	1200
56	8	162	0.83(1)	1199
56	8	166	0.80(1)	1200
48	12	129	1.00(1)	1200
48	12	135	0.93(2)	1200
48	12	140	0.861(9)	1200
48	12	145	0.813(7)	1200
48	12	148	0.79(1)	1200
48	12	151	0.75(1)	1200
60	12	129	1.03(2)	1200
60	12	135	0.93(2)	1200
60	12	140	0.87(1)	1200
60	12	145	0.82(1)	1200
60	12	151	0.762(8)	1199
60	12	157	0.712(7)	1199
60	12	162	0.647(6)	1199
60	12	166	0.632(9)	1200

$m_s/m_l = 40$

N_σ	N_τ	T[MeV]	χ_t/T	#conf
24	8	137	1.06(2)	2220
24	8	140	1.04(2)	2220
24	8	143	1.01(1)	2219
24	8	145	0.99(1)	2219
24	8	147	0.97(2)	3220
24	8	151	0.932(8)	3219
24	8	152	0.93(1)	3218
24	8	156	0.887(8)	4220
24	8	162	0.832(7)	4220
32	8	137	1.06(2)	3445
32	8	140	1.03(2)	3449
32	8	143	1.00(2)	3458
32	8	145	0.98(2)	3449
32	8	147	0.97(2)	3589
32	8	151	0.94(1)	3593
32	8	152	0.92(1)	3592
32	8	156	0.877(9)	3607
32	8	162	0.841(8)	2825
32	8	166	0.801(7)	2900
40	8	137	1.07(2)	1200
40	8	140	1.04(2)	1200
40	8	143	1.01(2)	1200
40	8	145	0.99(2)	1200
40	8	147	0.97(2)	1200
40	8	151	0.946(9)	1200
40	8	152	0.92(1)	1200
40	8	156	0.88(1)	1200
40	8	162	0.84(1)	1199
40	8	166	0.81(2)	1200
42	12	129	0.99(1)	1199
42	12	135	0.94(2)	1200
42	12	140	0.87(2)	1200
42	12	145	0.82(2)	1200
42	12	151	0.748(9)	1200
42	12	157	0.701(8)	1200
42	12	162	0.664(9)	1200
42	12	166	0.626(8)	1200
60	12	129	1.01(2)	1199
60	12	135	0.95(2)	1200
60	12	140	0.869(8)	1200
60	12	145	0.828(8)	1200
60	12	151	0.77(2)	1200
60	12	157	0.713(7)	1200
60	12	162	0.669(9)	1199
60	12	166	0.625(5)	1200

$m_s/m_l = 160$

N_σ	N_τ	T[MeV]	χ_t/T	#conf
56	8	143	1.00(2)	1200
56	8	145	0.98(2)	1200
56	8	147	0.95(2)	1200
56	8	151	0.92(2)	1200
56	8	154	0.91(2)	1124
56	8	156	0.87(2)	1199
56	8	162	0.85(2)	1200
56	8	166	0.800(9)	1199

TABLE 6.1: Results of the extracted topological susceptibility. The temperatures have been determined using the f_k-scale [59] with data from [61].

6.2 Extracting the topological susceptibility

The gradient flow smoothes the gauge field by integrating the flow equation eq. (4.80) in the direction of the flow-time τ_{F} up to a particular value of τ_{F}. Thus, all observables we measure on these gauge fields will in principle depend on the flow-time τ_{F}. To get physically meaningful quantities, the usual procedure is to first perform a continuum extrapolation of observables at $\tau_{\mathrm{F}} > 0$, and then a flow-time extrapolation $\tau_{\mathrm{F}} \to 0$. The topological susceptibility, however, is somewhat special. In the continuum it has been shown that cumulants of the topological charge, as for example the topological susceptibility, are flow-time independent [140]. On the lattice it has been observed that after some initial flow-time steps, i.e. when the UV-fluctuations have been smoothed out, the topological susceptibility also settles in a plateau [87, 122, 151].

This behavior can also be seen in our measurements. Figure 6.3 shows the topological susceptibility as a function of the flow-time of some selected lattices and each of our mass ratios m_s/m_l. Each plot shows the susceptibility curve of the lowest and highest temperature in this setup. We notice a first peak at $\tau_{\mathrm{F}} = 0$ across all lattices, which is larger on finer lattices than on coarser lattices. This is not surprising since at zero flow-time the gauge fields are affected by high frequency fluctuations and finer lattices should capture more UV-fluctuations than coarser lattices. Beyond this peak all our lattices, except the lattices which have a temporal extent of $N_\tau = 8$ and a low temperature, show a clear plateau which starts around $\tau_{\mathrm{F}} = 1.15a^2$. In what follows we will use this value of the flow-time as the lower limit for the plateau region in the determination of the topological susceptibility.

We choose as an upper limit in flow-time the maximum smoothing radius to be half of the temporal extent of the lattice, i.e. $\sqrt{8\tau_{\mathrm{F}}}T = 0.5$, in order to avoid any potential disturbing effects from the periodic boundaries of the gauge fields. We see a clear plateau region on all lattices except on the lattices with a temporal extent of $N_\tau = 8$ at low temperatures. This indicates that these lattices might be still too coarse to measure the topological susceptibility correctly. The same observation has also been made in [122]. However, the lower flow-time limit $\tau_{\mathrm{F}} = 1.15a^2$, which we have chosen based on the other lattices, is very close to the upper maximum flow-time of the $N_\tau = 8$ lattices, as it can be seen in figure 6.3. Since this interval is so narrow, it might explain why we do not see a clear plateau on these lattices. Thus, although the flow-time interval on the $N_\tau = 8$ lattice is rather tight, the correct susceptibility might still lie within this range. Therefore, we use this interval also on the $N_\tau = 8$ lattices at low temperatures to extract the topological susceptibility.

In most studies the susceptibility is arbitrarily chosen at one flow-time on the plateau. However, we still observe some small visible changes of the susceptibility in this region. In order to account for this, we decided to extract it from the entire plateau. This is achieved by drawing Gaussian noise around each of the data points on the plateau. The resulting distributions are then added to an overall final distribution. The median of this distribution is the final topological susceptibility with the error from a 68% percentile around the median of the distribution. The results are shown as a horizontal green band in each plot in figure 6.3. The extracted values are listed in table 6.1.

6.3 Volume dependence

As we perform computations on finite lattices, we have to make sure that uncertainties introduced by finite volume effects are under control. To suppress finite volume

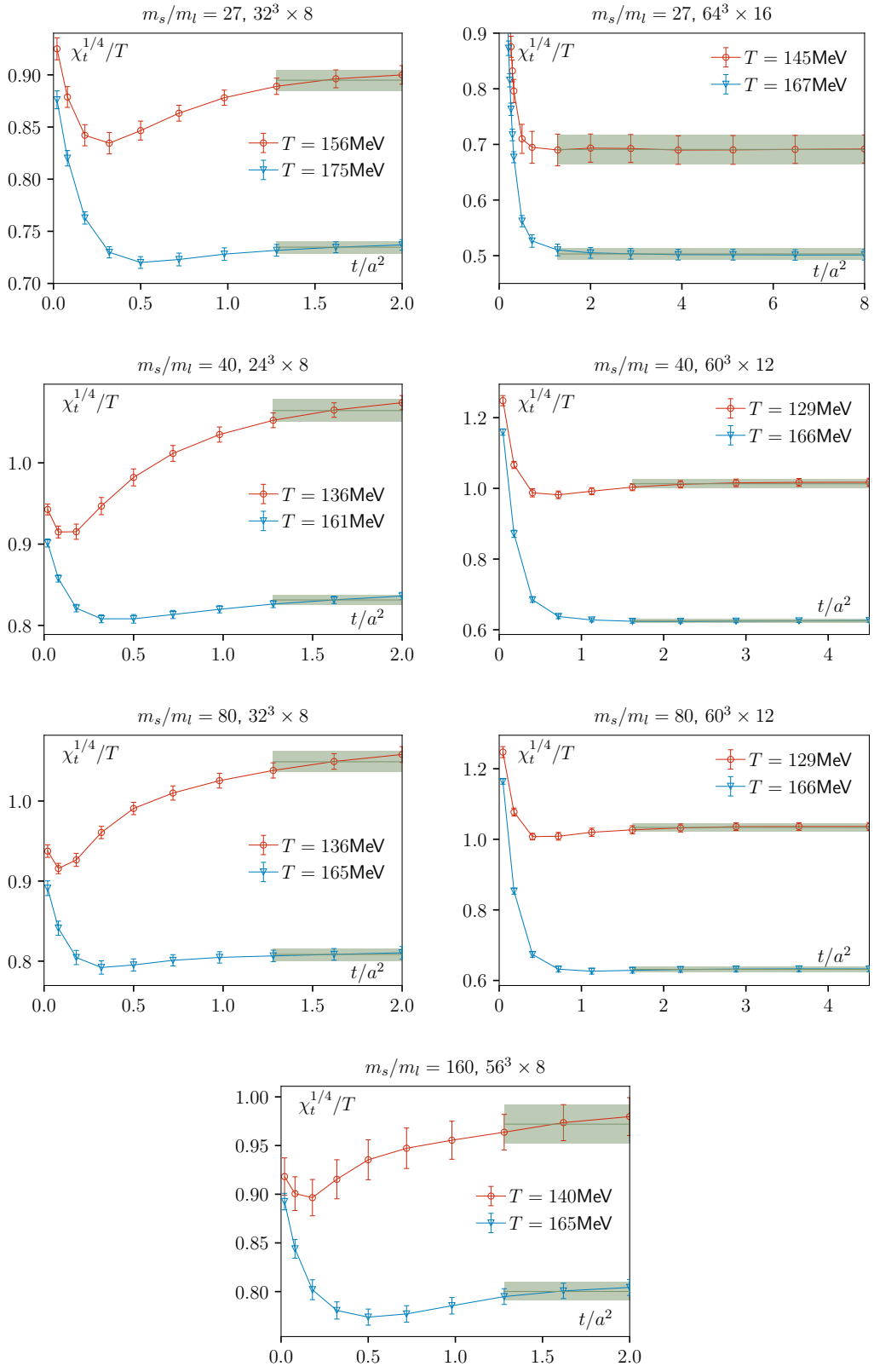


FIGURE 6.3: Topological susceptibility as a function of the flow-time. The green linear error bands show the plateau region which we use to extract the topological susceptibility.

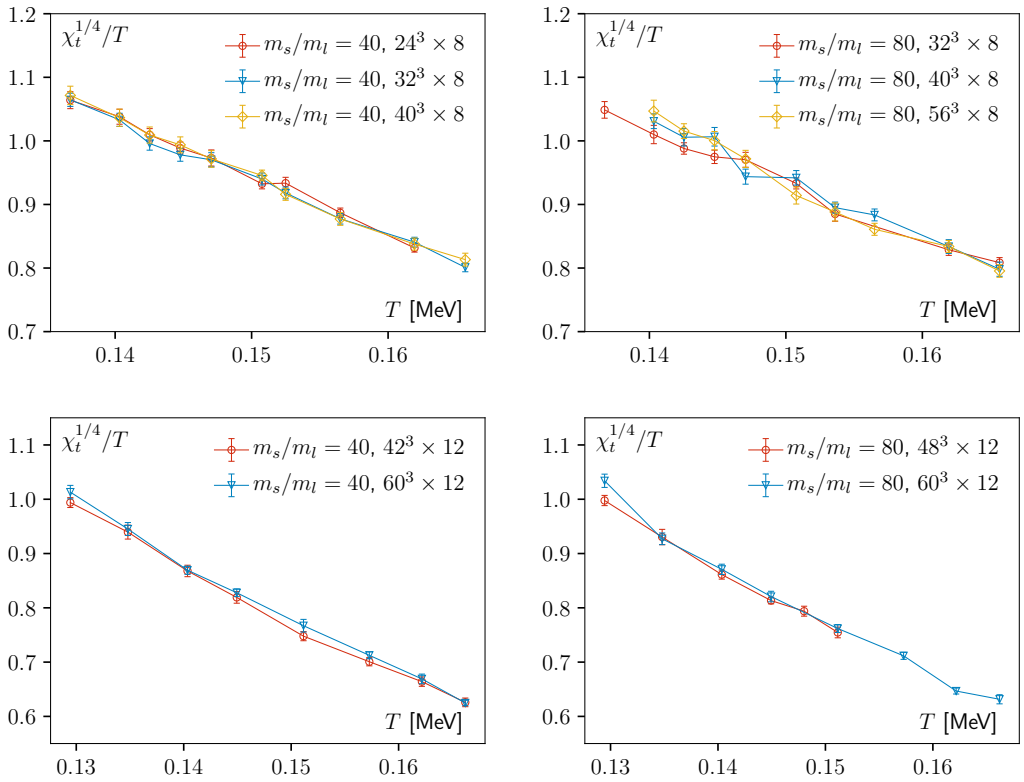


FIGURE 6.4: Topological susceptibilities as a function of the temperature to show whether some volume dependence is visible. Different plots show different temporal extent N_τ and different light quark masses. All curves overlap within two errorbars. Hence, we do not see clear finite volume effects.

effects in our setup the lattice dimensions are chosen such that $m_\pi N_\sigma a \gtrsim 2$. In figure 6.4 we show plots of the topological susceptibility as a function of the temperature for different spatial volumes. It can be seen that no clear volume dependence is discernible within our statistics. Independent of the given masses and temporal extent, the curves agree within two error bars. Hence, the aspect ratios which we are using are acceptable.

6.4 Mass dependence

An interesting question is how the topological susceptibility changes towards the chiral limit. At low temperatures below the phase transition, the topological susceptibility is expected to be proportional to $\chi_t \propto (1/m_u + 1/m_d + 1/m_s)^{-1}$ [2, 3], i.e. it vanishes as the quark masses are taken to zero. At high temperatures, arguments based on a dilute-instanton model show that the topological susceptibility is proportional to $\chi_t = m_u m_d m_s / T^3$ [1], i.e. it should as well vanish in the chiral limit. Figure 6.5 compares the susceptibility of configurations with different light quark masses but fixed strange quark mass. In order to focus only on the mass dependence and not on the a-dependence, the left plot shows the results of the $N_\tau = 8$ lattices while the right plot shows the results of the $N_\tau = 12$ lattices. In both cases the results at different masses overlap within two error bars. Hence, the susceptibility does not show any

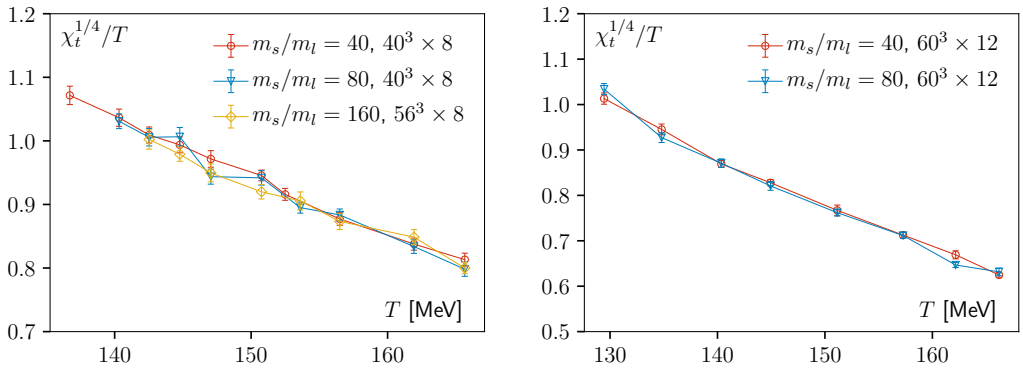


FIGURE 6.5: Comparison of the topological susceptibility as a function of the temperature at different light quark masses but fixed physical strange quark mass. Within two error bars, the curves overlap. Hence, no clear mass dependence visible within our statistics.

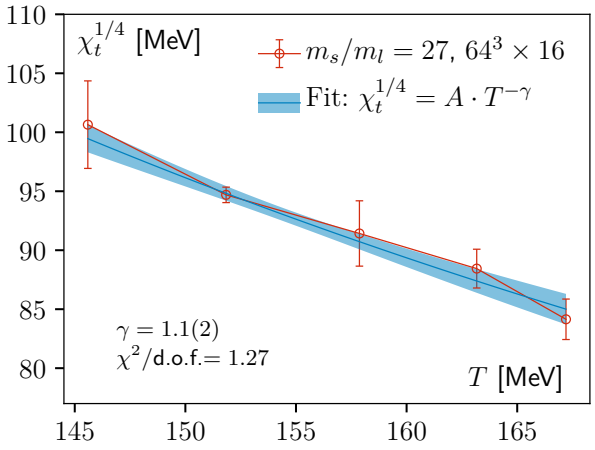


FIGURE 6.6: Topological susceptibility as a function of the temperature of $2+1$ HISQ configurations at physical quark masses. The error band of the fit has been computed using gaussian bootstrap samples.

mass dependence within our statistics, which is in contrast to the previously mentioned expectations.

6.5 Temperature dependence

Chiral perturbation theory predicts that the topological susceptibility should be $\chi_t^{1/4} = 77.8(4)$MeV at zero temperature in QCD with two degenerate flavours [2, 3, 152]. In figure 6.6 we show the topological susceptibility as a function of the temperature of the finest lattices at physical quark masses, i.e. $64^3 \times 16$, $m_s/m_l = 27$. The susceptibility at our lowest temperature is at $\chi_t^{1/4} = 101(4)$MeV, which is quite far away from the chiral perturbation prediction. An explanation for that might be that our values are not continuum extrapolated and the topological susceptibility suffers from large cut-off effects.

Overall the topological susceptibility shows a falling trend in temperature. This behavior has also been observed in previous studies [87, 153]. This is expected, since

with increasing temperature the amount of instantons decreases. We see that behavior not only on configurations with physical masses, i.e. figure 6.6, but also on configurations with non-physical light quark masses as can be seen in various plots within this chapter. At high temperatures, the *dilute instanton gas approximation* (DIGA) [1, 4] suggests that the topological susceptibility might have a power law dependence on temperature that is characterized by an exponent γ, such that $\chi_t^{1/4} \propto T^{-\gamma}$. From DIGA calculations the exponent is expected to be $\gamma = 8$. Similar calculations as in this work with $2+1$ HISQ configurations but in the temperature range $165\text{MeV} < T < 240\text{MeV}$ gave an exponent of $\gamma = 1.96(22)$ [87]. If we apply this fitting ansatz on our $N_\tau = 16$ data (see figure 6.6), we obtain an exponent of $\gamma = 1.1(2)$. However, in that temperature region that fitting ansatz might not be suitable anymore. Also the fact that we have not continuum extrapolated values might be problematic.

6.6 Discretization effects

Physical observables which are computed on the lattice are affected by discretization errors. In other words, these observables come with systematic errors which shift them away from the true physical value. How strongly they are affected by these cut-off effects depends on the underlying discretization scheme, such as the action and the observable itself. In order to check how strong the susceptibility is affected by cut-off effects, we show in figure 6.7 the topological susceptibility as a function of the temperature for different lattice spacings. It can be seen that the curve of the finer lattices $N_\tau = 12, 16$ lie considerably lower than the curve of the coarser lattices $N_\tau = 8$. The differences are about $\Delta\chi_t/T \approx 0.17$. Hence, the topological susceptibility suffers from strong cut-off effects. The same observation has also been made in [87], where $2+1$ HISQ configurations with a mass ratio of $m_s/m_l = 20$ have been used. To account for these systematic errors, one has to perform the thermodynamic and continuum limit. This is achieved by taking the infinite limit of the lattice dimensions, i.e. $N_\sigma \to \infty$ and $N_\tau \to \infty$, while sending the lattice spacing to zero, i.e. $a \to 0$. However, in this study we do not have enough lattices at our disposal to perform this task.

6.7 Conclusion

We have measured the topological susceptibility using the gradient flow on $2+1$ HISQ configurations with varying light quark masses and fixed physical strange quark mass.

Instead of choosing an arbitrary flow-time at which we extract the topological susceptibility, we obtained it by averaging over a flow-time plateau. Unfortunately, the $N_\tau = 8$ lattices at low temperatures did not show a clear plateau. Since all our $N_\tau > 8$ lattices have shown a plateau from $\tau_\text{F} > 1.15a^2$, and since the upper flow-time limit for a $N_\tau = 8$ lattice is $\tau_\text{F} = 2a^2$ (which corresponds to a smoothing range of $\sqrt{8\tau_\text{F}}T = 0.5$), we suspect that these problematic lattices are not large enough to show a straight susceptibility plateau on the flow-time axis. On the other hand, even though no clear plateau is visible, the correct value might still be within that flow-time interval. This motivated us to extract it from that region in the same way as we did for all other lattices. We have checked whether the volumes we are using are large enough to establish topology on these lattices and did not find any finite volume effects within our statistics. Although we were not able to perform a continuum extrapolation, the results provide us some initial indications on how the topological susceptibility behaves towards the chiral limit. Contrary to our intuition and previous

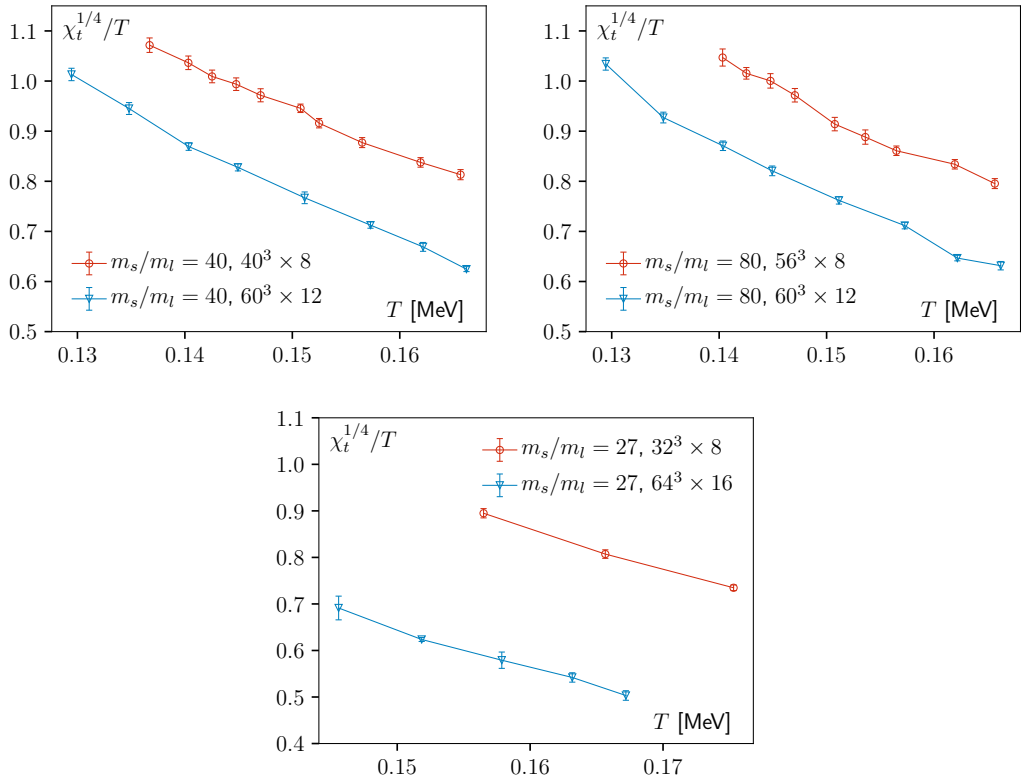


FIGURE 6.7: Comparison of topological susceptibility as a function of the temperature at different lattice spacings. Strong cut-off effects are visible.

studies, we did not see any mass dependence in the susceptibility around the pseudo-critical temperature. However, we have seen that the susceptibility gives very large discretization errors, and therefore a proper continuum extrapolation might reveal a small mass dependence. Note, that the pseudo-critical temperature is different for different light quark masses, which effectively shift the susceptibilities in T/T_{pc}, and hence doing a continuum extrapolation at the right temperature is not trivial. Finally, we found that the overall temperature trend of the susceptibility seems to agree with previous studies and expectations, i.e. it decreases with increasing temperature.

Chapter 7

The Overlap Dirac Eigenvalue Spectrum

In chapter 2 we discussed the concept and importance of symmetries in QCD. For two massless flavors of quarks, the QCD Lagrangian has a $\mathrm{U}_L(2) \times \mathrm{U}_R(2)$ chiral symmetry. The subgroup $\mathrm{SU}_V(2) \times \mathrm{SU}_A(2) \times \mathrm{U}_V(1)$ is spontaneously broken to $\mathrm{SU}_V(2) \times \mathrm{U}_V(1)$ in the hadron phase giving rise to pions which are much lighter than the nucleons. An essentially non-perturbative feature of QCD which arises due to strong color interactions is the axial anomaly $\mathrm{U}_A(1)$. It is not an exact symmetry in QCD but broken due to quantum effects [11–13]. However, this anomaly is of special interest, since it is believed to affect the nature of the chiral phase transition in QCD. In this chapter we are therefore interested in whether or not an effective $\mathrm{U}_A(1)$ restoration occurs at the chiral phase transition. This question can only be answered non-perturbatively. This chapter is based on the paper in ref. [154], which is at the time of writing this thesis under peer review in the journal *Physical Review D*.

Since the up and down quark masses are light compared to the intrinsic scale of QCD, the $\mathrm{U}_L(2) \times \mathrm{U}_R(2)$ symmetry is only mildly broken. Therefore, if we calculate observables that measure the $\mathrm{U}_A(1)$ breaking in $2+1$ flavor QCD with physical u, d quark masses near the chiral crossover transition and reduce $m_{u,d}$, we can smoothly reach the chiral limit. If indeed $\mathrm{U}_A(1)$ is broken, signatures of the $\mathrm{O}(4)$ second order line could be observed by reducing $m_{u,d}$. On the other hand if $\mathrm{U}_A(1)$ is effectively restored we should approach the $\mathrm{Z}(2)$ line or a second order line of the $\mathrm{U}_L(2) \times \mathrm{U}_R(2)/\,\mathrm{U}_V(2)$ universality class [65]. To investigate this we study in detail the eigenvalue spectrum of $2+1$ QCD, where we keep the strange quark mass fixed to its physical value and successively lower the light quark mass to effectively approach the two flavor chiral limit of QCD. We then study its effects on $\mathrm{U}_A(1)$ breaking observables as a function of light quark mass near the chiral crossover transition.

In recent years lattice studies have provided some interesting initial insights about the fate of the anomalous $\mathrm{U}_A(1)$ subgroup of the chiral symmetry in $2+1$ flavor QCD with physical quark masses. A $\mathrm{U}_A(1)$ breaking near and above the chiral crossover region has been reported by studies on the eigenvalue spectra of $2+1$ flavor QCD using domain wall fermions [15, 16, 155], highly improved staggered quark discretization [17, 18] and by using the twisted mass Wilson fermion discretization [21]. On the other hand many recent studies for two flavor QCD, with physical and heavier than physical light quarks and infinitely heavy strange quark, using overlap fermions [22], re-weighted spectra of the domain wall fermions [23–26], improved domain wall fermions [27] as well as from non-perturbatively $\mathcal{O}(a)$ improved Wilson fermions [28] have reported effective restoration of the $\mathrm{U}_A(1)$ near T_c. A detailed understanding of the near-zero mode spectrum of the staggered fermion configurations was achieved using overlap fermions [72] which show similarities to the pure staggered spectrum

T/T_c	m_s/m_l	β	N_σ	N_τ	#conf
0.97	27	6.390	32	8	45
1.05	27	6.445	32	8	108
1.09	27	6.500	32	8	69
0.99	40	6.390	32	8	28
1.03	40	6.423	32	8	52
1.05	40	6.445	32	8	154
1.05	80	6.423	56	8	60

TABLE 7.1: The details of the HISQ configurations analyzed in this work.

on finer lattices, closer to the continuum [156]. The near-zero modes contribute dominantly towards the $\mathrm{U}_A(1)$ breaking [16, 72].

7.1 Numerical setup

The gauge configurations which we use in this chapter are taken from the same $2+1$ flavor HISQ ensemble as in the previous chapter 6. The strange quark mass is set to its physical value and the two light quark flavors are degenerate with their mass varied such that $m_s/m_l = 27, 40, 80$. These choices of the light quark mass correspond to Goldstone pion masses with $\sim$ $135, 110, 80$ MeV, respectively. We focus on a temperature range which is between T_c and $1.1\ T_c$, where T_c is the pseudo-critical temperature (which is a function of the pion mass). The values of T_c are $\sim 158, 157, 154$ MeV in the f_K scale [61] for pion masses $135, 110, 80$ MeV, respectively. We have chosen the lattices with a temporal extent of $N_\tau = 8$. The aspect ratios are chosen such that $N_\sigma/N_\tau = 4$ for the $m_s/m_l = 27, 40$ gauge configurations and $N_\sigma/N_\tau = 7$ for the $m_s/m_l = 80$ gauge configurations. This ensures that the corresponding lattice extent along the spatial directions are large enough $m_\pi L \sim 2.7$-3.5 to minimize the finite volume effects. The details of the configurations used in this chapter are given in Table 7.1. We use the overlap Dirac operator [37, 157] to measure the eigenvalues of the HISQ sea configurations since it has an exact index theorem on the lattice [158], and hence can resolve the small eigenvalues efficiently (see section 4.4 for more details). Resolving the infrared eigenvalue spectrum of the HISQ configurations with a HISQ operator on relatively coarser lattices may be difficult due to the breaking of continuum flavor symmetries [18]. However, on finer lattices which are closer to the continuum, a peak of]near-zero modes is observed. This infrared peak can be very efficiently resolved using the overlap operator on the HISQ sea configurations even on coarser lattices [156]. We perform a proper tuning of the valence overlap quark mass and the sea HISQ quark masses and then measure the appropriately renormalized eigenvalue spectrum and observables sensitive to $\mathrm{U}_A(1)$ breaking to ameliorate effects of the mixed Dirac operator set-up used here.

The overlap operator D_{ov} is defined as in equation (4.54). It was realized by calculating the sign function exactly in terms of the first 50 eigenvectors of the square of the Wilson-Dirac operator $D_W^\dagger D_W$ and then representing the contribution of the higher eigenvalues through a Zolotarev Rational function with 25 terms. The resultant norm of the sign function $D_W/\sqrt{D_W^\dagger D_W}$ deviated from unity on average by about 10^{-9}. The overlap operator satisfied the Ginsparg-Wilson relation up to a numerical precision of 10^{-9} or even lower. The parameter M appearing in the overlap operator

was chosen to be 1.8 since it gave the minimal violation of the Ginsparg-Wilson relation and approximated the sign function to the best numerical precision on the majority of the gauge configurations studied.

We then calculate the lowest eigenvalues of the overlap operator on the HISQ sea-ensembles using the Kalkreuter-Simma Ritz algorithm [120]. For gauge ensembles with pion masses $135, 110$ MeV we have measured the first 100 eigenvalues but increased the number of eigenvalues to 200 for configurations with pion mass of 80 MeV, due to the increasing density of the low-lying eigenvalues. The diagonalization of the Dirac operator becomes numerically quite expensive for the gauge ensembles with lighter sea quark masses. This is due to the fact that the number of zero modes increases, and they need to be calculated with very high precision. We have implemented a novel procedure to circumvent this problem which we describe in the following section.

7.1.1 Accelerating the overlap Dirac eigenvalue measurements

The conjugate gradient algorithm which is used to calculate the eigenvectors only works on Hermitian operators. The overlap Dirac matrix D_{ov}, however, is a normal matrix. To account for that, one usually diagonalizes instead the square of the overlap Dirac matrix $D_{\text{ov}}^{\dagger} D_{\text{ov}}$, which is Hermitian. The eigenvalues of this Hermitian operator come in degenerate pairs with opposite chiralities. The zero modes however are non-degenerate with distinct chirality and their number and the chirality depends on the topological charge of the gauge configurations. A significant amount of time of the diagonalization routine is spent on measuring the zero modes with a reasonable precision. We therefore projected our Dirac operator to measure only those eigenmodes which have a chirality that is of opposite to those of the zero modes. The corresponding eigenspace has no zero modes, and leaving them out accelerates the diagonalization routine significantly. The zero-modes do not contribute to the physical observables in the thermodynamic limit; thus measuring the eigenvalue spectrum without zero-modes allows for a significant speedup of our calculations. This is especially true for the gauge ensembles with sea-quark masses towards the chiral limit where the probability of occurrence of zero modes increases. We have explicitly checked on a few configurations that for the lattice volumes we have considered the contribution from the zero modes to the renormalized observables is negligibly small.

Restricting the measurements to consider eigenvalues of only one chirality is achieved by using appropriate projection operators

$$P_{\pm} = \frac{1}{2} \left(\mathbb{1} \pm \gamma_5 \right), \tag{7.1}$$

which project the overlap Dirac matrix onto a vector space of one chirality, i.e. $P_{\pm} D_{\text{ov}} P_{\pm}$. This projected Dirac operator is Hermitian, since

$$\left(P_{\pm} D_{\text{ov}} P_{\pm} \right)^{\dagger} = P_{\pm} D_{\text{ov}}^{\dagger} P_{\pm} = P_{\pm} \gamma_5 D_{\text{ov}} \gamma_5 P_{\pm} = \left(\pm P_{\pm} \right) D_{\text{ov}} \left(\pm P_{\pm} \right) = P_{\pm} D_{\text{ov}} P_{\pm}. \tag{7.2}$$

This property can further be used to accelerate our algorithm, since we can now use the Hermitian operator $P_{\pm} D_{\text{ov}} P_{\pm}$ instead of $D_{\text{ov}}^{\dagger} D_{\text{ov}}$ in the conjugate gradient. However, in order to project the overlap Dirac matrix onto a vector space which is devoid of any zero modes, we need to know which chirality these modes have. Thanks to the index theorem this prior knowledge can be extracted from the topological charge Q using its field-theoretical definition. The sign of the topological charge corresponds to the chirality of the zero-modes. For that we use the topological charges which we have already extracted in chapter 6 using the Zeuthen flow and the improved

discretization of the topological charge density. The chiralities have been extracted from the topological charge at a flow-time of $\sqrt{8\tau_{\rm F}}T = 0.45$. We have also confirmed already in the previous chapter that these configurations do not suffer from topological freezing which assures that we have ergodically sampled a sufficient number of independent configurations for this study.

In the measurements of the overlap eigenvalues for the $m_s/m_l = 80$ ensemble, we have observed a significant slowing down of the algorithm to converge to the desired precision. The reason for that was that these gauge configurations tend to have significantly more small eigenvalues, some of which are localized on the scale of the lattice spacing. In order to improve the convergence, we systematically removed these ultra-violet defects by smoothing the configurations using the Zeuthen flow up to a flow-time of $\tau_{\rm F} = 0.32a^2$ before measuring the eigenvalue spectrum with the overlap Dirac operator. The smoothing of the gauge fields has been used earlier in the context of measuring the hadron spectrum using valence overlap fermions [159, 160].

7.2 Analyzing the overlap eigenvalue spectrum

In the section 2.5 we have shown some scenarios where one can also deduce from the Dirac eigenvalue density itself whether the anomalous $\mathrm{U}_A(1)$ is restored or broken. Hence, in this section we cover the general features of the eigenvalue spectrum of the QCD Dirac operator. In particular, we study the spectrum of our ensemble near and above the chiral crossover transition. The eigenvalue distributions $\rho(\lambda)$ can be characterized by essentially three distinct features: a non-analytic dependence in λ which forms a zero mode peak and a near-zero mode accumulation as well as a regular analytic dependence in λ which forms the bulk eigenvalue spectrum [72]. Since we do not measure exact zero modes our plots do not show the corresponding peak in $\rho(\lambda)$. The non-analytic near-zero region is denoted as the infrared peak. In figure 7.1 we show the eigenvalue density $\rho(\lambda)$ as a function of the imaginary part of the eigenvalues λ of the overlap Dirac operator for temperatures near T_c and for physical quark masses. We observe that the near-zero modes and the bulk modes overlap considerably. Also the near-zero modes appear to develop a peak towards the infrared region. Similar features have also been observed in ref. [72] for HISQ ensembles with heavier than physical light quark masses $m_s/m_l = 20$. Figure 7.2 shows a plot of the eigenvalue density for lower than physical quark masses $m_s/m_l = 40$, for different temperatures above T_c. Qualitatively, these distributions show similar features of the eigenvalue spectrum as for the physical or heavier than physical quark mass distributions. The presence of the infrared region is easily distinguishable from the bulk modes in the chiral crossover region. Remarkably this near-zero peak becomes more prominent as the temperature is increased gradually from T_c. This is due to less contamination by the bulk modes, whose density shifts further towards the larger eigenvalues.

At this point we want to remind the reader that the bare HISQ eigenvalue spectrum computed without the overlap operator as a valence or probe operator, does not show any peak in the infrared region [17, 18] on coarser $N_\tau = 8$ lattices. However, such a peak appears in bare HISQ eigenvalue spectrum when one goes towards more finer $N_\tau = 16$ lattices [156]. By using the overlap Dirac matrix as the valence operator on the HISQ sea ensembles, this lack of an exact index theorem for the HISQ operator is corrected, such that the infrared peak can be extracted even on the coarser $N_\tau = 8$ lattices. Therefore, this non-analytic peak in the eigenvalue spectrum is not a lattice artifact, which has been discussed earlier in the context of domain wall fermions on comparably small lattice volumes [24, 25]. In fact the presence of such a peak just

above T_c can be explained by an interacting ensemble of topological clusters [161]. In the high temperature phase a dilute gas of instantons [15, 72, 162, 163] may contribute to this peak. It has been recently argued that the existence of such near-zero modes in the chiral symmetry broken phase of QCD with massless quarks, can lead to the disappearance of Goldstone excitations at finite temperature [164]. Hence, we will study the sensitivity of the slope of the bulk region as well as the overlap between the near-zero and the bulk modes to the change in temperature.

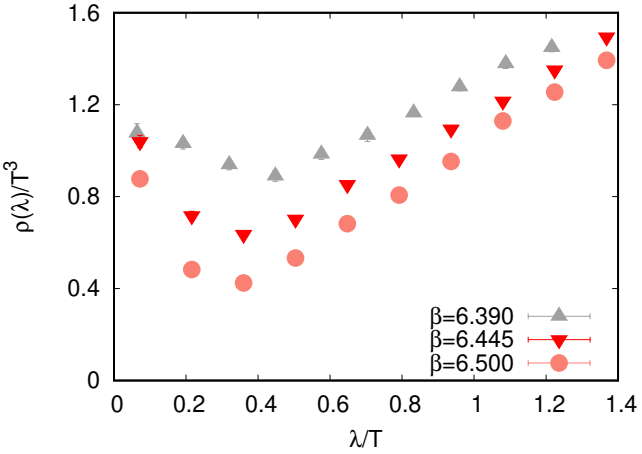


FIGURE 7.1: The eigenvalue density of the massless valence overlap Dirac operator measured on HISQ sea configurations with $m_s/m_l = 27$, as a function of temperature, near and just above T_c.

Calculations based on chiral Ward identities of up to four-point correlation functions in the pseudo-scalar and scalar meson channels have shown that if the eigenvalue distribution for QCD with two light quark flavors is an analytic function in m_l^2, the analytic (bulk) part of the Dirac eigenvalue density in the chirally symmetric phase should be proportional to [71]

$$\lim_{m_l \to 0} \rho(\lambda, m_l) \sim \lambda^3 + \mathcal{O}(\lambda^4). \tag{7.3}$$

With this constraint it was further shown that all up to six-point correlation functions that are related through the $\mathrm{U}_A(1)$ symmetry will be degenerate. Consequently the anomalous breaking of $\mathrm{U}_A(1)$ would be invisible in these correlation functions [71].

We are therefore interested in the bulk region of the eigenvalue spectrum as a function of the sea-quark mass at different lattice spacings to understand the fate of $\mathrm{U}_A(1)$ just above T_c. Motivated by ref. [72], we fit the eigenvalue distribution at different temperatures to the ansatz,

$$\rho(\lambda) = \frac{\rho_0 A}{A^2 + \lambda^2} + c(m_l)\Theta(\lambda - \lambda_0)|\lambda|^{\gamma(m_l)} , \tag{7.4}$$

where $\gamma(m_l)$ characterizes the leading order analytic dependence of the eigenvalue density on λ and can be in general a function of m_l. The exponent γ is extracted by choosing a threshold λ_0 in the eigenvalues that separates the non-analytic peak from the analytic part. In equation (7.4) this is implemented through a Heaviside step function Θ in the second term of the fit ansatz. Since we have measured only $\mathcal{O}(100)$ eigenvalues per configuration, we can study only the leading analytic behavior for the infrared part of the eigenvalue spectrum for $\lambda > \lambda_0$. In table 7.2 we summarize the

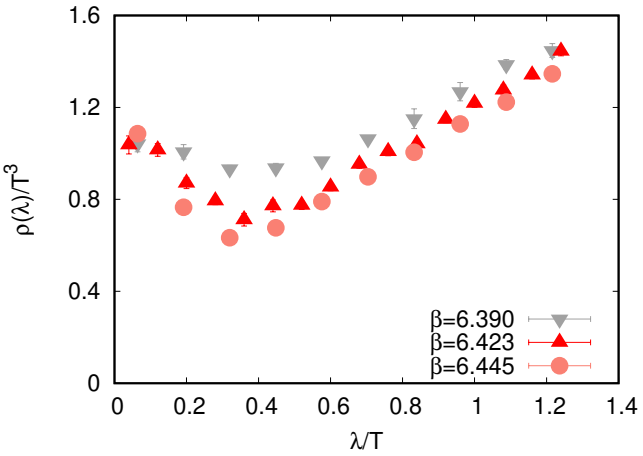


FIGURE 7.2: The eigenvalue density of the valence massless overlap Dirac operator measured on the HISQ sea configurations with $m_s/m_l = 40$.

results of the fit, the exponent γ for different quark masses, including the values of the cut-off λ_0, temperatures and the goodness of the fits.

At $T_c \leq T \leq 1.1\ T_c$ we observe that γ is independent on the sea-quark mass, which is consistent with the results obtained previously on HISQ fermions [72] with heavier than physical quark masses. In fact, the exponent turned out to be $\gamma \sim 1$ for the temperature range we studied so far. This is consistent with predictions from chiral perturbation theory [165, 166]. In ref. [71] it has been argued that the coefficient $c(m_l)$ in equation (7.4) goes as m_l^2 in the chiral limit for two flavor QCD. This in turn implies that the eigenvalue spectrum in the chiral symmetry restored phase would not show any linear dependence on λ. In the context of the Columbia plot in figure 2.1, the two flavor limit is approached along the $m_s = \infty$ line. We check the prediction from ref. [71] instead by approaching the two flavor chiral limit along the line of constant physical value of m_s. This is achieved by studying the dependence of $c(m_l)$ as a function of the light quark mass m_l. Since we are interested in the temperature regime where the chiral symmetry is restored, we neglect the lowest two β values for light sea-quark masses $m_s/27, m_s/40$ respectively. Considering the dimensionless ratio

$$\frac{\rho(\lambda)}{T^3} = \frac{c(m_l)}{T^2} \cdot \frac{\lambda}{T}, \tag{7.5}$$

the idea is to perform a linear fit to $c(m_l)/T^2$ as a function of m_l^2/T^2. If $c(m_l)$ indeed goes as m_l^2 to leading order in the light sea-quark mass m_l, then this fit should have a zero intercept along the vertical axis. Figure 7.3 shows the resulting $c(m_l)/T^2$ extracted from the eigenvalue densities for $T \gtrsim T_c$ and different light quark masses. The fit has a constant intercept with the vertical axis when a chiral extrapolation has been performed which even dominates over the usual $\mathcal{O}(m_l^2/T^2)$ term. As a consequence, the eigenvalue density to leading order should go as $\mathcal{O}(\lambda)$ instead of $\mathcal{O}(\lambda^3)$ just above above T_c even in the chiral limit.

We also observe that no gap opens up in the infrared part of the eigenvalue spectrum. Hence, we can conclude that even when we approach the chiral limit, the $\mathrm{U}_A(1)$ remains broken. This is in addition to the contribution to the $\mathrm{U}_A(1)$ breaking that comes due to the non-analytic peak in the eigenvalue spectrum.

m_s/m_l	β	T/T_c	λ_0/T	γ	χ^2/dof
40	6.390	0.99	0.45	1.09(22)	0.70
40	6.423	1.03	0.5	0.94(23)	0.99
40	6.445	1.05	0.5	1.08(15)	0.66
27	6.390	0.97	0.4	1.03(18)	0.66
27	6.445	1.03	0.5	1.09(11)	0.90
27	6.500	1.09	0.5	1.03(12)	0.94

TABLE 7.2: The temperature (T), the exponent γ characterizing the leading order λ^γ rise of the bulk eigenvalues λ and the goodness of fits performed on eigenvalue densities for different choices of the light sea-quarks and physical value of strange sea-quarks.

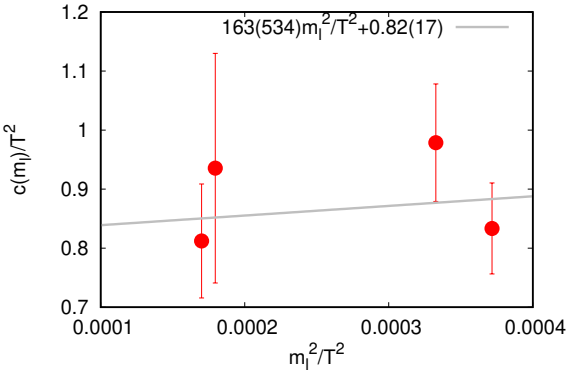


FIGURE 7.3: The coefficient $c(m_l)$ of the λ dependent term of the eigenvalue density shown as a function of light sea quark mass squared m_l^2. Both axes are normalized by appropriate powers of the temperature. The χ^2 per degree of freedom of the fit is 0.95.

7.3 Quantifying $\mathrm{U}_A(1)$ breaking in the chiral limit

The low sensitivity of the exponent γ of the bulk modes to the sea-quark mass indicated already that the $\mathrm{U}_A(1)$ remains broken when we approach the chiral limit. In this section we will provide a more quantitative estimate of the $\mathrm{U}_A(1)$ breaking towards the chiral limit. In particular, we want to compare the eigenvalue spectra at different quark masses and also compare them with the earlier results obtained with the overlap fermions on HISQ configurations [72] for heavier than physical quark masses. It is however important to renormalize the eigenvalue spectra for such a comparison, since the eigenvalue density is not a renormalization group invariant quantity. Here we choose a renormalized eigenvalue density using the valence strange quark mass. The valence overlap quark mass has been tuned to the HISQ sea quark masses by matching the renormalized quantity Δ, which is defined as

$$\Delta = \frac{m_s\langle\bar{\psi}\psi\rangle_l - m_l\langle\bar{\psi}\psi\rangle_s}{T^4} \,, \tag{7.6}$$

where $\langle\bar{\psi}\psi\rangle$ is the chiral condensate (see eq. (2.24)), which can be re-expressed in terms of the overlap Dirac operator $D_m = D_{\text{ov}}(1 - am/2M) + am$, i.e.

$$\langle\bar{\psi}\psi\rangle(m) = \frac{T}{V}\left\langle \text{tr}\left(D_m^{-1}\frac{\partial}{\partial D_m}\right)\right\rangle \tag{7.7}$$

with a (valence) quark mass m. By rewriting eq. (7.7) in terms of overlap eigenvalues we obtain the formula,

$$a^3\langle\bar{\psi}\psi\rangle(m) = \frac{1}{N_\sigma^3 N_\tau}\left[\frac{\langle|Q|\rangle}{am} + \left\langle\sum_{\tilde{\lambda}\neq 0}\frac{2am(4M^2 - |\tilde{\lambda}|^2)}{|\tilde{\lambda}|^2(4M^2 - (am)^2) + 4(am)^2M^2}\right\rangle\right] , \tag{7.8}$$

where Q is the topological charge and $\tilde{\lambda}$ is the eigenvalue of D_{ov}, which is scaled by the defect height parameter M. The term which arises due to the zero modes on the right hand side of this expression has been neglected, since it is a finite volume artefact to the above observable. To make sure that the finite volume effects are under control in our tuning procedure, we have checked explicitly that this neglected term provides negligibly small corrections to Δ. When tuning m_s using the observable Δ we keep the ratio m_l/m_s fixed for the valence and sea quark sectors. With that constraint we calculate Δ in the valence quark sector from the first $\mathcal{O}(100)$ eigenvalues of the overlap Dirac operator measured on the HISQ ensembles. For the sea quark sector we calculate the exact results for Δ obtained by the inversion of the HISQ operator using stochastic sources on the same ensemble. Then we match the values of Δ obtained from the valence quark sector with the values of Δ obtained from the sea quark sector to extract the tuned m_s of the valence quarks that correspond to the HISQ sea quark masses which have been tuned to the physical ones before. With these values we can describe the physics of the underlying sea quarks using the valence overlap fermions. Table 7.3 shows the results for the tuned strange quark masses for different ensembles.

m_s/m_l	β	m^s_{sea}	$m^s_{\text{val}}(\Delta)$
80	6.423	0.0670	0.025
40	6.390	0.0694	0.090
40	6.423	0.0670	0.058
40	6.445	0.0652	0.038
27	6.390	0.0694	0.098
27	6.445	0.0652	0.051
27	6.500	0.0614	0.032

TABLE 7.3: The valence strange quark masses obtained by matching the observable Δ measured using the eigenvalue density of the valence overlap to that measured by inversion of the sea HISQ Dirac operator.

Figure 7.4 shows a comparison of the renormalized eigenvalue density $m_s\rho(\lambda)/T^4$ as a function of λ/m_s. The number of bins of these renormalized densities were kept fixed for different ensembles. The bulk part of the renormalized spectrum shows a linear rise with increasing light quark mass, while the sensitivity of the near-zero peak is very weak to that change in quark mass. The spectrum of the $m_l = m_s/80$ gauge configurations is only shown up to $\lambda/m_s \sim 2.5$ with the first 200 eigenvalues we have measured, since as mentioned earlier, these configurations have an increasing density of these small eigenmodes.

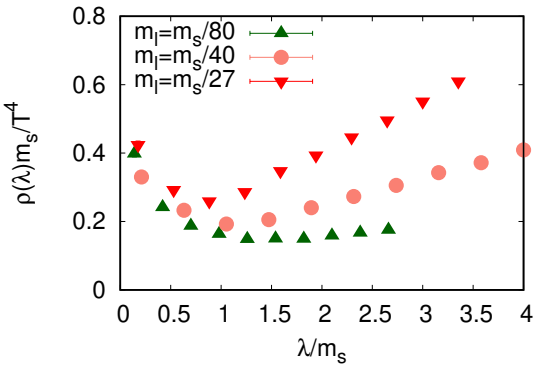


FIGURE 7.4: The renormalized eigenvalue density of the QCD ensembles at $T = 1.05\ T_c$ generated using HISQ discretization and measured using an appropriately mass-tuned valence overlap operator. These are shown for three different choices of the light quarks.

Next we proceed to estimate the sensitivity of $\mathrm{U}_A(1)$ to the light quark mass and whether it is effectively restored above the crossover transition using the tuned m_s. However, $\mathrm{U}_A(1)$ has no unique observable that is sensitive to its restoration, because it is not an exact symmetry. One observable which we introduced in section 2.5 as a measure of $\mathrm{U}_A(1)$ is the difference of the integrated two-point correlators of isospin-triplet pion and delta mesons, $\chi_\pi - \chi_\delta$ (see eq. (2.28)). In fact one needs to further look at the degeneracy between higher point correlation functions for different meson quantum number channels [71, 167, 168]. In this work as a first test, we focus on this specific two-point correlation function. Through a chiral Ward identity $\chi_\pi - \chi_\delta$ can also be obtained by [72]

$$\chi_\pi - \chi_\delta = \frac{\langle\bar{\psi}\psi\rangle}{m} - \chi_{\mathrm{conn}} \tag{7.9}$$

where $\langle\bar{\psi}\psi\rangle$ is the chiral condensate as defined in eq. (7.7). The connected chiral susceptibility χ_{conn} is defined as

$$\chi_{\mathrm{conn}} = \frac{T}{V}\left\langle \partial_m \,\mathrm{tr}\left(D_m^{-1}\partial_m D_m\right)\right\rangle, \tag{7.10}$$

where $D_m = D_{\mathrm{ov}}(1 - am/2M) + am$ is the Dirac operator for overlap quarks with a (valence) quark mass m. Thus, $\omega \equiv \chi_\pi - \chi_\delta$ can be expressed in terms of the eigenvalues of the overlap Dirac operator [72]

$$a^2\omega(m) = \frac{1}{N_\sigma^3 N_\tau}\left[\frac{\langle|Q|\rangle}{(am)^2} + \left\langle \sum_{\tilde{\lambda}\neq 0} \frac{2(am)^2(4M^2 - |\tilde{\lambda}|^2)^2}{\left[|\tilde{\lambda}|^2(4M^2 - (am)^2) + 4(am)^2M^2\right]^2}\right\rangle\right]. \tag{7.11}$$

We used the first $\mathcal{O}(100)$ eigenvalues of the overlap Dirac operator at the values of the tuned valence quark masses to measure the quantity in eq. (7.11). Chiral Ward identities ensure that $\chi_\pi - \chi_\delta = \chi_{\mathrm{disc}}$, where χ_{disc} is the disconnected part of the integrated iso-singlet scalar meson correlator. This relation can be used to verify the quality of our quark mass tuning by checking whether it is satisfied using our data. For χ_{disc} we use the previously measured data for physical quark masses from

ref. [62]. This data has been obtained by the inversion of the HISQ Dirac operator using stochastic sources on $N_\tau = 8,\ 12,\ 16$ lattices. Using that data, we perform a continuum extrapolation of $m_l^2\chi_{\rm disc}/T^4$ and compare it to the observable $m_l^2\omega/T^4$, which we have calculated using the eigenvalues of the valence overlap Dirac operator on the same HISQ ensembles using the tuned valence quark masses. The results are are shown in figure 7.5. We observe a reasonably good agreement of these renormalized quantities, which gave us further confidence on our quark mass tuning procedure.

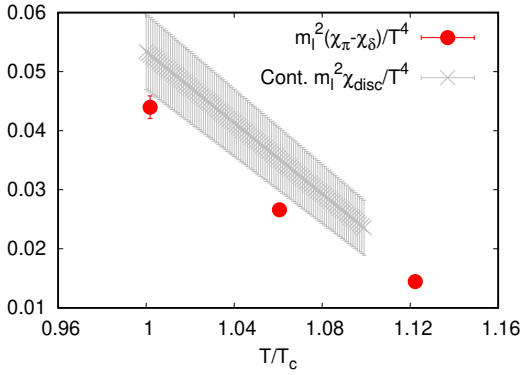


FIGURE 7.5: Comparison of $\chi_\pi - \chi_\delta$ measured using the overlap eigenvalues to the continuum estimates of $\chi_{\rm disc}$ using data from ref. [62], shown for physical quark masses.

With this confidence, we study now the quark mass dependence of the appropriately renormalized $\mathrm{U}_A(1)$ breaking observable $m_l^2\omega/T^4$. The anomalous $\mathrm{U}_A(1)$ is effectively restored in the chiral limit if it has a trivial dependence on the light quark mass (as also expected in the perturbative regime), i.e. if $\omega \sim m_l^2$. On the other hand if the $\mathrm{U}_A(1)$ is broken, then the leading order behavior is $\omega \sim \mathcal{O}(m_l^0)$, which gives us the effective magnitude or the strength of $\mathrm{U}_A(1)$ breaking. We want to remind the reader that we have calculated only the first $\mathcal{O}(100)$ of the $\mathcal{O}(10^6)$ of eigenvalues of the QCD Dirac operator. However, these infrared eigenvalues contribute significantly to the $\mathrm{U}_A(1)$ breaking. When chiral symmetry is effectively restored, our data can be fit to the ansätze,

$$\frac{m_l^2(\chi_\pi - \chi_\delta)}{T^4} = \frac{m_l^2\omega}{T^4} = a_1\frac{m_l^2}{T^2} + a_2\frac{m_l^4}{T^4}, \tag{7.12}$$

$$= b_1\frac{m_l^4}{T^4} + b_2\frac{m_l^6}{T^6}. \tag{7.13}$$

The first ansatz denotes $\mathrm{U}_A(1)$ breaking whereas the second ansatz is valid on its effective restoration. In figure 7.6 we show our calculated $m_l^2\omega/T^4$ at $1.05\ T_c$ for three different tuned light valence quark masses. The red band in this figure corresponds to a fit using equation (7.12). The data fit quite well to that ansatz. The largest contribution to the uncertainty comes from the value corresponding to the lowest quark mass. On the other hand, our data disfavors the second fit ansatz in equation (7.12), since the quality of the fit was $\chi^2/\text{d.o.f.} \approx 5$, which is almost a factor 2.5 larger than that corresponding to the first ansatz.

The extracted magnitude of $(\chi_\pi - \chi_\delta)/T^2$ in the chiral limit is $a_1 = 156 \pm 13$. This value is clearly finite and non-zero within the current uncertainties, and therefore we conclude that $\mathrm{U}_A(1)$ is broken above the chiral crossover temperature for the $N_\tau = 8$

lattices we have studied so far. This is even the case when we approach the chiral limit along the line of constant physical value of m_s.

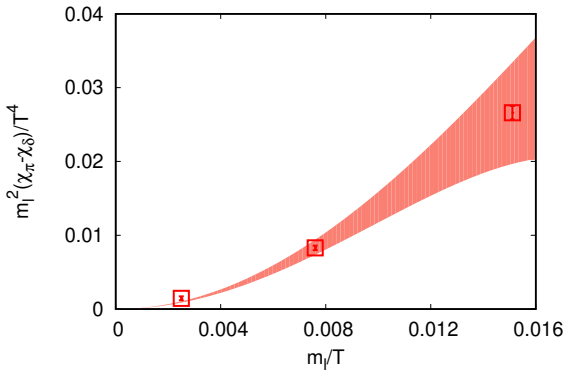


FIGURE 7.6: The renormalized $\mathrm{U}_A(1)$ breaking parameter shown as a function of light quark mass at $1.05\ T_c$.

7.4 Conclusion

In this chapter we have calculated the eigenvalue spectrum in the chiral symmetry restored phase of QCD and estimated the fate of anomalous $\mathrm{U}_A(1)$ symmetry as we approach the chiral limit along the line of constant physical strange quark mass. The calculations have been carried out on selected $N_\tau = 8$ lattices from the HISQ ensemble, which have also been used in the previous chapter. Since the HISQ operator does not realize all the continuum flavor and anomalous symmetries on a finite lattice, we use the overlap Dirac matrix as the valence or probe operator to correctly measure the number and density of the near-zero eigenmodes of these ensembles. In order to accelerate the measurements, we project the overlap Dirac operator to a vector space which is devoid of its zero modes using the chiralities extracted from the topological charge. The valence quark masses have been tuned to the sea quark mass to obtain physical and renormalized results even with different valence and sea quark actions. From the appropriately renormalized eigenvalue density of the QCD Dirac operator we have observed that the eigenvalue density can be represented as $\rho(\lambda) \sim \lambda$ in the chiral limit at a temperature of $1.05\ T_c$. This is in contrast to the expectations of ref. [71], where the leading order behavior of the eigenvalue spectrum of QCD was derived to be $\rho(\lambda) \sim \lambda^3$ for $N_f = 2$ QCD in the chiral limit. We finally observe that the renormalized observable $m_l^2(\chi_\pi - \chi_\delta)/T^4$ results in a non-zero value. Therefore, we conclude that $\mathrm{U}_A(1)$ is broken when one approaches the chiral limit along the line of constant physical m_s.

So far we only studied a single lattice spacing, i.e. $N_\tau = 8$. Although computationally demanding, analyzing remaining cut-off effects due to the HISQ sea quark sector or even continuum extrapolations remain for future studies.

Chapter 8

The qq Correlation Function and the Sphaleron Rate

With increasing temperature, topological tunneling through the sphaleron barrier, i.e. instantons, become more and more suppressed. On the other hand, the temperature acts as a catalyst which allows transitions between two vacua by jumping over that potential barrier instead of tunneling through it. Hence, these jumping phenomena, which are called sphalerons become more present (see section 3.5). This induces chiral imbalances in the quark gluon plasma through the axial anomaly which may affect for example the chiral magnetic effect [5, 6]. There are semiclassical approaches to determine the sphaleron rate at high temperatures [39], but they are not reliable in the relevant coupling region. Since non-perturbative studies on the $\mathrm{SU}(3)$ sphaleron rate at finite temperature are rather limited, it is therefore well-motivated to investigate the sphaleron rate using lattice QCD methods. A recent study can be found in [7].

The sphaleron rate is defined as the mean-squared change of the Chern-Simons number over the 4-volume in Minkowski time (see eq. (3.41)). For our purposes, however, it is more useful to write the sphaleron rate in terms of the low-frequency part of the spectral function of the two-point topological charge density correlation function at zero spatial momentum

$$\Gamma_{\text{sphal}} = \lim_{\omega \to 0} \frac{2T\rho_q(\omega)}{\omega}, \tag{8.1}$$

since it is related to the Euclidean topological charge density correlation function through an integral relation (see eq. (3.42) and (3.43)). This allows us to compute the sphaleron rate non-perturbatively via the Euclidean topological charge density correlation function.

At first sight the above mentioned relations sound straightforward to implement. However, a couple of obstacles need to be addressed. First of all the topological charge density as well as its correlation function are highly affected by high frequency fluctuations on the lattice. Therefore, a noise reduction technique, such as the gradient flow, needs to be applied before measuring the correlation function. Next, a proper continuum and flow-time extrapolations need to be performed to account for lattice artifacts and smoothing effects. Finally, the extraction of the spectral function is a challenging task, since it is hidden inside the integral which we measure.

In this chapter we address all these issues and develop a method to measure the sphaleron rate non-perturbatively on the lattice based on continuum and flow-time extrapolated correlation functions. A similar study has been done previously in the context of color-electric two-point correlation function [169]. This chapter is based on the paper in ref. [170], which is at the time of writing this thesis under peer review in the journal *Physical Review D*.

8.1 Numerical setup

The lattices which we use within this chapter are generated in the quenched approximation using the standard Wilson gauge action. Initially, we have performed 5000 heat bath sweeps to make sure that the configurations are sampled from thermal equilibrium. The thermalized configurations are then used as initial configurations to generate more configurations using a combined sweep of one heat bath and four overrelaxation steps. We stored the configurations of only every 500 combined sweeps to account for possible autocorrelation effects between configurations. In this way we eliminated the autocorrelations of all our measured quantities except the topological charge, which we will discuss later.

In order to perform a proper continuum extrapolation, we have chosen five different isotropic lattices with the temperature chosen to be $T = 1.5T_c$, where the critical temperature T_c is defined via the Sommer parameter $r_0 T_c = 0.7457(45)$ [171]. We used the Sommer parameter r_0 [172] as the scale reference to determine the temperatures with parameterization from [171] and updated coefficients from [173]. The generated sample which we use in this chapter is the same as in our recent paper [169]. A summary of the configurations is listed in table 8.1.

a (fm)	a^{-1} (GeV)	N_σ	N_τ	β	T/T_c	#conf.
0.0262	7.534	64	16	6.8736	1.51	10000
0.0215	9.187	80	20	7.0350	1.47	10000
0.0178	11.11	96	24	7.1920	1.48	10000
0.0140	14.14	120	30	7.3940	1.51	10000
0.0117	16.88	144	36	7.5440	1.50	10000

TABLE 8.1: Lattice spacings, lattice extents, β values and statistics of configurations in this work. The lattice spacing a is determined by the Sommer scale (see [172]).

For the topological charge density $q(x)$ we use the discretization in eq. (4.74) and a a^2-improved implementation of the field strength tensor using a mixture of square and 1×2 rectangular plaquettes [124] (see section 4.5).

Just measuring the topological charge density $q(x)$ as in eq. (4.74) on our generated configurations will however result in values which do not only capture topological information. The topological charge density is in general always contaminated by non-topological high-dimension operators, which are suppressed by powers of the lattice spacing a. If fluctuations with wave number $k \sim 1/a$ are present, these operators contaminate our measurement with non-topological effects. Only smooth gauge fields allow for a well-defined definition of $q(x)$ that accurately captures topological information. Therefore, a smoothing technique needs to be applied before measuring $q(x)$ to dampen the highest-frequency fluctuations.

In this work we use the gradient flow as a noise reduction method. As described in section 4.7 a Symanzik improved discretization of the gradient flow is used. It is called the Zeuthen flow [145] and introduces no new $\mathcal{O}(a^2)$ discretization errors. Integrating the discretized flow equation is achieved by a 3$^{\text{rd}}$ order Runge-Kutta with an adaptive step-size algorithm. The same numerical implementation of the gradient flow is also used in [169]. The observables are measured on flow-times $\sqrt{8\tau_F}T \in \{0, 0.001, \ldots, 0.2, 0.205, \ldots, 0.3\}$.

The gradient flow improves our operators in two ways. First, with increasing flow-time UV fluctuations get eliminated and correlation functions become less noisy.

Second, the gradient flow improves the topological behavior of $q(x)$ by eliminating such short-distance fluctuations. In the next sections we will see that there is a lattice spacing dependent flow-time above which $q(x)$ correctly captures topological information. Values which are smaller than that can not be trusted, and hence must be avoided. The correlation function, however, is τ_{F}-dependent and only the $\tau_{\mathrm{F}} \to 0$ limit (after performing the continuum limit) gives the desired correlation function. Therefore, a flow-time extrapolation has to be performed in order to obtain this limit. However, with increasing flow-time, the correlation functions eventually differ so much that the finite τ_{F} correlator is not useful anymore for the $\tau_{\mathrm{F}} \to 0$ extrapolation, since very coarse lattices lie outside the scaling region.

The gradient flow corrections of a correlation function measured over a separation τ are expected to be dependent on τ_{F}/τ^2. Only sufficiently small τ_{F}/τ^2 values will be in a useful scaling window where they help to determine the $\tau_{\mathrm{F}} \to 0$ extrapolation. For larger values physical information is lost. It has been shown that above a certain ratio of τ_{F}/τ^2, the correlation function even has the wrong sign [174]. Hence there is also a τ-dependent upper limit on available τ_{F} values. In [174] the authors conservatively advise $\sqrt{8\tau_{\mathrm{F}}} < 0.33\tau$ based on a leading order perturbative calculation.

8.2 Determining the topological charge density

To compute the sphaleron rate we need to compute the topological charge density $q(x)$, eq. (3.31). Therefore, it is necessary to verify whether $q(x)$ is actually measuring the topological charge density correctly.

Since the gluonic definition of the topological charge density returns only useful results on smooth configurations, we need to apply a smoothing procedure on the gauge field before measuring $q(x)$. The gradient flow, which we have chosen for this task, smoothes the field with increasing flow-time τ_{F}. Therefore, the question arises: What is the minimum flow-time we need to obtain reliable values of $q(x)$ that capture topological information of the gauge field? To tackle this question, we compute the topological susceptibility χ_t

$$\chi_t = \frac{\langle Q^2 \rangle - \langle Q \rangle^2}{V}, \tag{8.2}$$

where V is the four-volume and Q is the topological charge (see eq (4.73)). We compute this quantity on each lattice as a function of the flow-time. The result can be seen in figure 8.1. We find that the topological susceptibility stabilizes on all lattices after the flow-time $\tau_{\mathrm{F}} = 0.5a^2$. The susceptibility at smaller flow-times is still contaminated by non-topological artifacts. Hence, we use only χ_t at flow-times $\tau_{\mathrm{F}} \geq 0.5a^2$ to be sure that $q(x)$ represents the topological charge density.

Another question is related to a general problem of topology on the lattice, i.e. how does topological freezing affects our measurements? It is well known that the topological charge Q evolves slowly on finer lattices, which results in large autocorrelation times and poor statistics. Since the topological charge Q is obtained by integrating the density $q(x)$ over the spacetime volume, it makes sense to ask the question whether the correlation function of $q(x)$ might suffer on topological freezing as well. Figure 8.1 already shows that the susceptibility is much larger for the finest lattices which is an indication of topological freezing. Indeed, as can be seen from the trajectory in figure 8.2, the topological charge Q of our configurations is affected by strong topological freezing as we go to finer lattices. However, it is not obvious whether the correlation function of $q(x)$ should be sensitive to the Euclidean topological sector, especially since the sphaleron rate is not related to the topological susceptibility at all. Also at

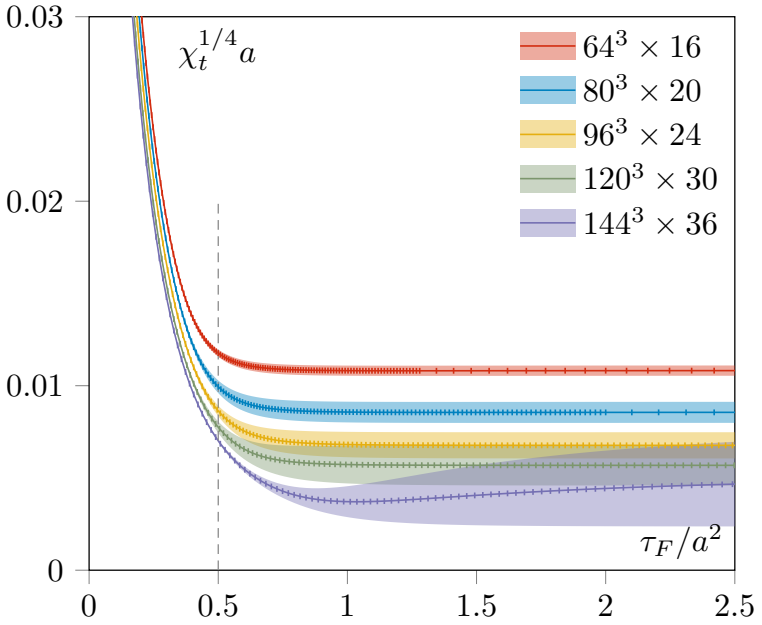


FIGURE 8.1: Topological susceptibility measured at different flow-times. The vertical line is the flow-time which we have chosen as a lower limit to ensure that q properly captures topology.

weak coupling, Euclidean topology and instantons are completely unrelated to the sphaleron rate [175].

In figure 8.2 we show the trajectory of Q of our configurations. We indeed see that the topological charge freezes more and more as we go to finer lattices. Therefore, we need to investigate the extent to which the topological charge density correlators are affected by the topological sectors and, therefore, how much of a problem of topological freezing may be. We do this by measuring the qq correlators separately in different topological sectors, and then analyzing how much a misweighting or exclusion of the $Q \neq 0$ sectors might affect our results. We choose the $N_\tau = 16$ lattice to analyze this, since the topological susceptibility is relatively well determined there as it can be seen in figure 8.1. In order to improve our determination of the properties of the $Q \neq 0$ sector, we perform a Markov chain simulation in which we prevent the $Q = 0$ sector via an extra accept-reject step, to obtain a sample of only $Q \neq 0$ configurations. We then compare the $Q = 1$ correlation function with both the $Q = 0$ sector, and the proper mixture of $Q = 0$ and $Q \neq 0$, i.e.

$$G(\tau) = \frac{Z_{Q=0} G_{Q=0}(\tau) + Z_{Q=1} G_{Q=1}(\tau)}{Z_{Q=0} + Z_{Q=1}}, \tag{8.3}$$

where $Z_{Q=0}$ and $Z_{Q=1}$ are the numbers of configurations of the $Q = 0$ and $Q = 1$ sample, respectively. $G_{Q=0}(\tau)$ is the correlator computed on the $Q = 0$ sample, while $G_{Q=1}(\tau)$ is computed on the $Q = 1$ sample. The results are shown in Figure 8.3. The $Q = 1$ sector disagrees with the $Q = 0$ sector by more than their error bars, but nevertheless the difference is so small that, given the small fraction of $Q = 1$ configurations in the full sample, the difference is smaller than our statistical errors even for the largest flow time where the errors are the smallest. This implies that a misweighting of the $Q = 1$ sector, even by a factor of 2 relative to its correct weighting, will change the correlation functions of interest by less than their statistical errors and can therefore be ignored. Since the finer lattices have a still less topological topological

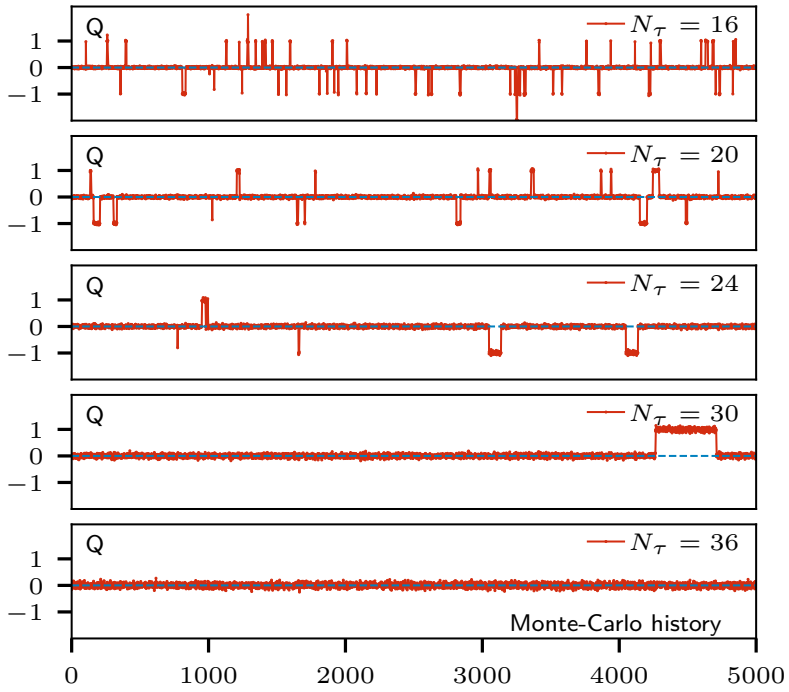


FIGURE 8.2: Trajectory of the topological charge at $\tau_F \approx a^2$. The x-axis shows the configuration number. Subsequent configurations are separated by 500 sweeps. Different streams have been concatenated.

fluctuations, and therefore even less $Q = 1$ configurations, we expect that this result holds for those lattices as well.

Furthermore, we check whether the correlation function is affected by autocorrelations, due to the topological freezing as seen in figure 8.2. We have computed the autocorrelation time of the correlation function at separation $\tau T = 0.5$ of the finest lattice $N_\tau = 36$. The result is seen in figure 8.4. We observe that the correlation function does not suffer from autocorrelations at all.

In summary, from the comparison of the correlation function between different topological sectors on coarser lattices as well as the autocorrelation times of the correlation function on finer lattices we do not expect that topological freezing contaminates our results. Hence, we will proceed without considering topological freezing further.

8.3 Continuum and flow-time extrapolation

When we compute quantities on the lattice we obtain values which are dependent on the lattice spacing. However, physics only resides at zero lattice spacing. Hence, we need to perform a continuum extrapolation of the lattice data. Similarly, also correlation functions measured on gauge fields with non-zero flow-time need to be extrapolated to $\tau_F \to 0$ to make them physically meaningful, especially since the integral eq. (3.43) which we use to extract the spectral function only holds for the correlation function at $\tau_F = 0$. Therefore, we proceed as follows: First, we perform the continuum extrapolation at each flow-time and then use these continuum extrapolated values of $G_{\tau_F}(\tau)$ to extrapolate the flow-time to $\tau_F \to 0$.

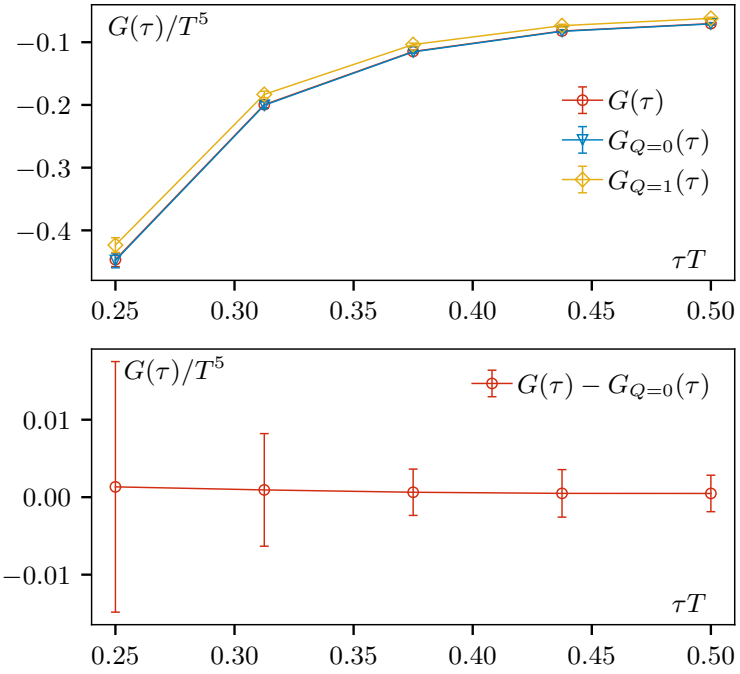


FIGURE 8.3: Topological charge density correlator of the $N_\tau = 16$ lattice at flow times $\sqrt{8\tau_F^{\max}}T = 0.5220\tau T$ (the largest flow-times used in our analysis). Top: Comparison of the full correlator with the correlators of the $Q = 0$ and $Q = 1$ sectors. Bottom: Difference between the full correlator and the correlator computed using only the $Q = 0$ sector, in comparison to the statistical error bars.

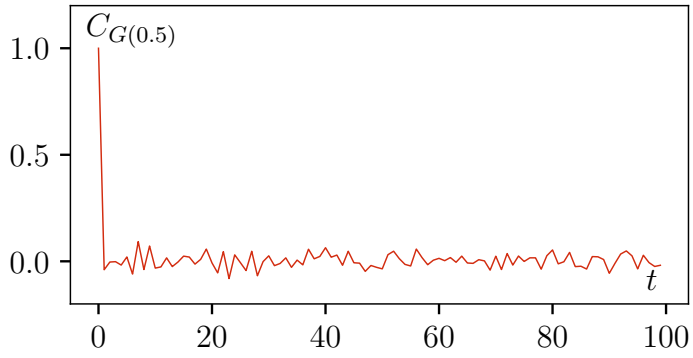


FIGURE 8.4: Autocorrelation time of the topological charge density correlator at separation $\tau T = 0.5$ of the finest lattice $N_\tau = 36$.

8.3.1 Interpolation

Before we can perform a continuum extrapolation, we have to make sure that the correlation functions of all our lattices provide the same separations. This is however not the case, since our lattices have different values of N_τ. Therefore, we first need to interpolate the coarser lattices in order to obtain intermediate values at separations which match with the separations of the finest lattice. The interpolation is performed by using cubic splines, where we constrained the first derivative at $\tau T = 0.5$ to be zero. This constraint should reflect the symmetry of the correlation function on the lattice around the midpoint. The correlation function at separations smaller than $\tau T = 0.166$

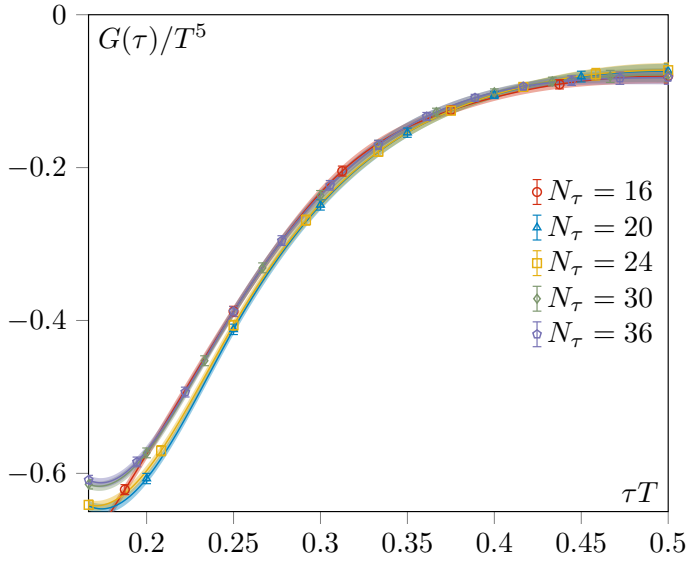


FIGURE 8.5: Topological charge density correlation function at flow-time $\sqrt{8\tau_F}T = 0.15$ for all lattice spacings. Cubic splines have been used to perform these interpolations. The error bands have been calculated during the interpolation procedure.

increases rapidly to values much higher than at larger separations. This makes the cubic spline interpolations unstable at separations larger than $\tau T = 0.166$. However, the shorter distances are not very helpful in extracting the low-frequency spectral function. Also our scaling window of useful τ_F values closes up at small τT, which makes any attempt to study smaller τT values useless. The interpolation range has therefore been chosen to be $\tau T \in [0.166, 0.5]$. Examples of a set of interpolations at one selected flow-time are shown in figure 8.5.

8.3.2 Continuum extrapolation

Using the interpolated values we performed the continuum extrapolation by linear fits in $1/N_\tau^2$ at all available separations of the finest lattice ($N_\tau = 36$) using the ansatz

$$\frac{G_{\tau,\tau_F}(N_\tau)}{T^5} = m \cdot N_\tau^{-2} + b, \tag{8.4}$$

where b is the continuum correlator normalized by T^5.

In the next subsection, a minimum and maximum range for τ_F will be defined that restricts the $\tau_F \to 0$ extrapolation to a correct scaling window. However, that window can interfere in some cases with the condition $\tau_F > 0.5a^2$, which we obtained in the previous section 8.2 in order to get a proper definition of $q(x)$. To account for that, we exclude all lattices where this conflict occurs from the continuum extrapolation. Or in other words, for each τ value, we include only lattices in the continuum extrapolation whose minimum from section 8.2 is smaller than the scaling window minimum which will be introduced in the next subsection.

The statistical error is estimated using bootstrap resampling with 10000 samples, where each sample consists of 10000 configurations. The interpolations and extrapolations are performed within these samples.

In figure 8.5 we see that the relative cutoff effects of the lattices are small at $\sqrt{8\tau_F}T = 0.15$. The same is also visible in figure 8.6. There we show the correlator

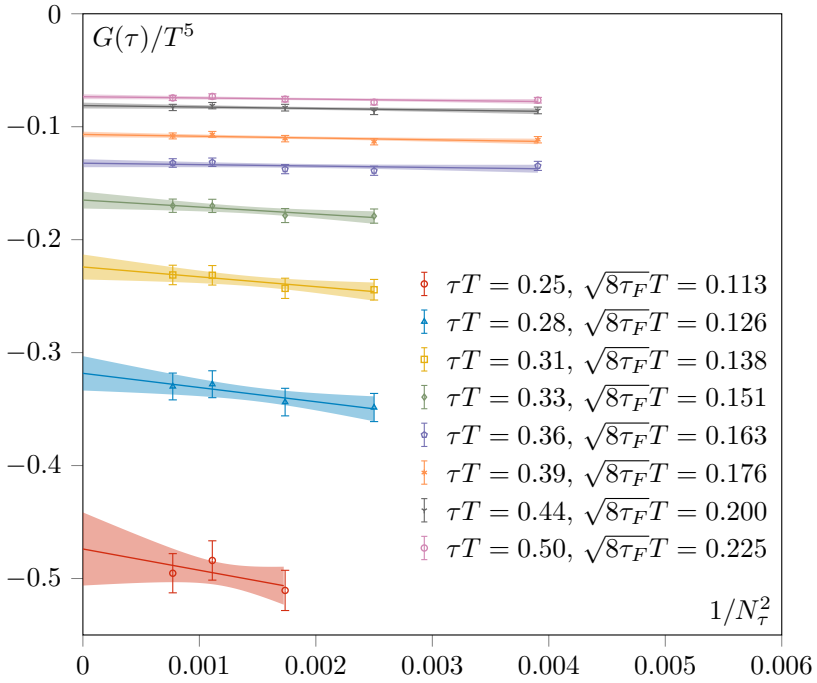


FIGURE 8.6: Continuum extrapolation of some selected separations at flow-times halfway between the beginning and the end of each flow range (8.5).

as a function of $1/N_\tau^2$ for some selected separations at flow times chosen from a flow range which we will define in the next section (see Eq. (8.5)). The two smallest separations exclude the coarsest lattice, which is a result of the minimum-flow-time-condition conflict, which we discussed earlier. In all cases the continuum extrapolated values agree with the lattice data within error bars. Since the gradient flow produces renormalized operators that are insensitive to lattice-scale fluctuations, these findings are not surprising.

8.3.3 Flow-time extrapolation

For the flow-time extrapolation $\tau_\mathrm{F} \to 0$ we need to decide on a useful flow-time range. A lower limit is required, since the correlation function at very low flow-times suffers from UV fluctuations, which would then also deteriorate the flow-time extrapolation. Values extracted at very large flow-times, on the other hand, are also unreliable, since the gradient flow has most likely smoothed out relevant physical information at these flow-times. Therefore, an upper limit for the flow-time extrapolation is also required. According to the small flow-time expansion (SFtX) method [176], the main modifications by gradient flow could be captured by an extrapolation linear in flow-time. A separation-dependent criterion for the upper flow-time limit has been proposed in [174] based on a leading order perturbation calculation. It has been determined by allowing the leading-order term of $\langle F^a_{\mu\nu}\tilde{F}^a_{\mu\nu}(\tau)F^b_{\alpha\beta}\tilde{F}^b_{\alpha\beta}(0)\rangle$ to differ from its non-flowed counterpart by at most 1%.

Unfortunately, this constraint turned out to be too strict for our non-perturbative results. Using this upper limit in flow-time we do not observe that linear-in-flow-time behavior at all. Hence, we decided to loosen that criterion to a 20% deviation. Values below that upper limit in flow-time, namely $\sqrt{8\tau_\mathrm{F}^\mathrm{max}}T = 0.5220\tau T$, show the required linear behavior for our flow-time extrapolation.

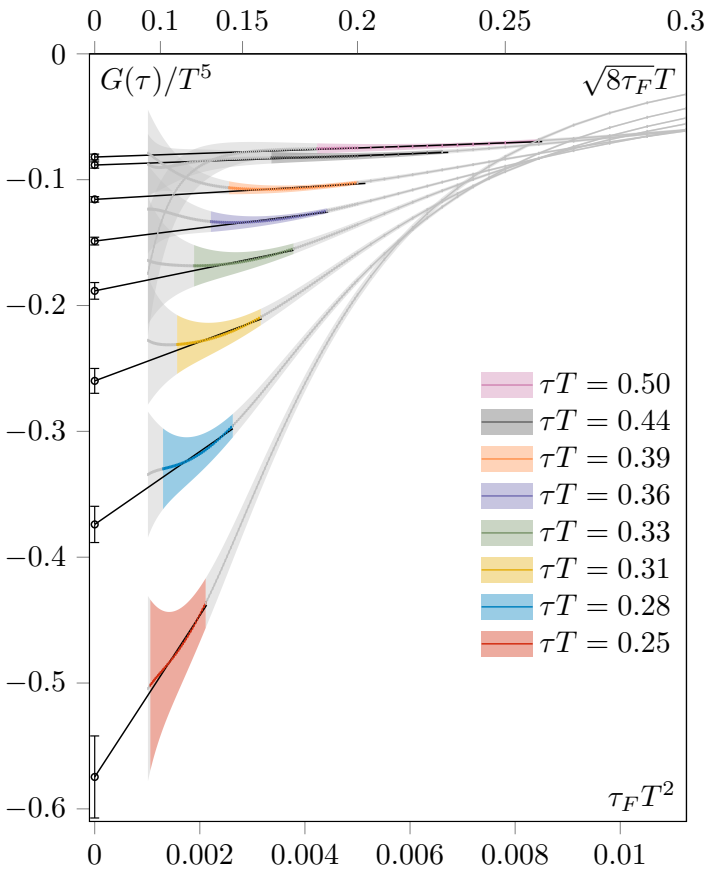


FIGURE 8.7: Continuum-extrapolated topological charge density correlator as a function of flow-time. The colored parts are the regions which are used to perform the flow-time extrapolations, while the black lines on top indicate the linear extrapolations. The gray regions are not considered in the flow-time extrapolation.

For the lower limit in flow-time we chose it to be half of the upper limit. We thus perform a linear $\tau_{\mathrm{F}} \to 0$ extrapolation in the range

$$\tau_{\mathrm{F}} \in [0.5\tau_{\mathrm{F}}^{\mathrm{max}}, \tau_{\mathrm{F}}^{\mathrm{max}}] \tag{8.5}$$

A general analysis of continuum operators on gradient-flowed field configurations shows that, in terms of unflowed operators, they can be expanded as an operator product expansion, and correspond to the desired operator, possibly with a renormalization factor, plus a series of high-dimension operators with compensating positive powers of τ_{F} as determined by operator dimension, see for instance refs [136, 177]. In our case we know that the topological charge does not renormalize (as we easily verify by seeing that it integrates to an integer) and that the high-dimension operators must vanish on space integration, implying that they are of form, e.g., $\tau_{\mathrm{F}} D^2 q$. Such a contaminating high-dimension operator does not affect the determined total topology $\int q d^4x$, but it does affect the correlation functions, leading to corrections of order τ_{F}/τ^2 based on dimensional reasoning. Therefore, an appropriate Ansatz for our correlation function at small τ_{F}, incorporating the lowest-order corrections, is

$$\frac{G_\tau(\tau_{\mathrm{F}})}{T^5} = c \cdot \tau_{\mathrm{F}} T^2 + d, \tag{8.6}$$

where d is the correlator at zero flow-time normalized by T^5.

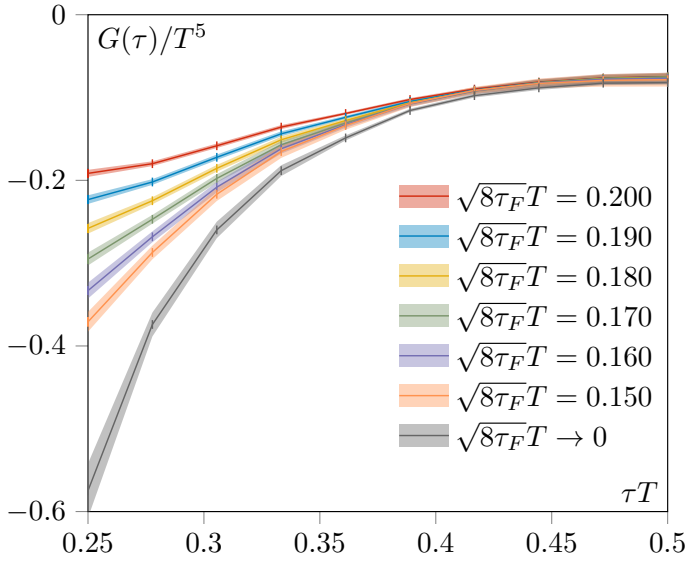


FIGURE 8.8: Continuum extrapolated topological charge density correlator as well as the final flow-time extrapolated correlator.

In figure 8.7 we illustrate the extrapolation procedure. Each curve in this plot represents the continuum-extrapolated $G_{\tau_F}(\tau)$ at a fixed separation τ as a function of τ_F. The colored regions show the ranges which are used for the flow-time-to-zero extrapolation according to eq. (8.5), while the gray regions are not used in the extrapolation. The straight lines in black indicate the linear $\tau_F \to 0$ extrapolation, where the final extrapolated values are the discrete points at $\tau_F = 0$. We see that the extrapolated value and the value at the largest-used τ_F differ by at most 20%, indicating that a small-τ_F expansion should still be valid and that higher-order τ_F^2 terms should only give a few percent corrections and are not (yet) needed.

Our final continuum and flow-time extrapolated correlation function is shown in figure 8.8. We observe that the correlator is negative in the range we are analyzing, that is $\tau T > 0$. Since the topological charge density $q(x)$ itself is odd under time reflections and due to arguments based on reflection positivity [178, 179], the continuum correlation function is expected to be negative for all nonzero separations $\tau T \neq 0$ [178, 180].

8.4 The sphaleron rate

As summarized in section 3.5, the sphaleron rate is determined by the small frequency limit of spectral function, eq. (8.1). However, the high-frequency part of the spectral function is expected to give large contributions to the Euclidean correlation function at small separations. Therefore, in order to extract the sphaleron rate properly, it is important to include as much information as possible about this high-frequency region. To our advantage, perturbation theory can provide us information for the high-frequency region of the spectral function. The spectral function has been computed both at leading order (LO) and next-to-leading order (NLO) in the coupling in

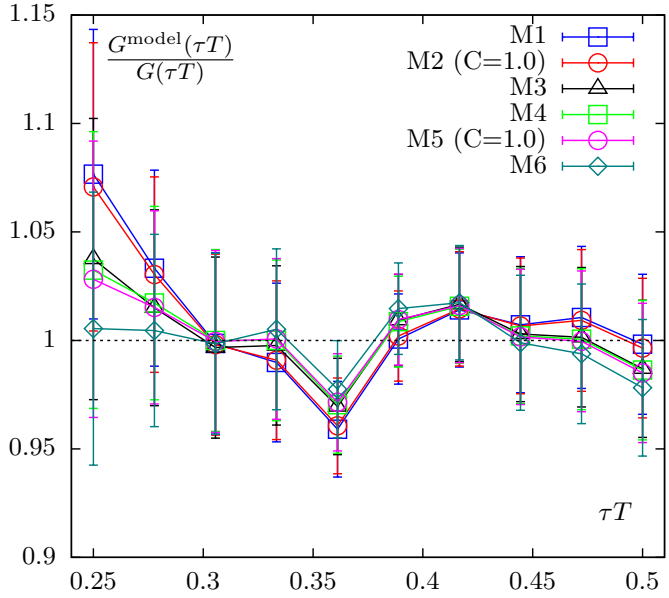


FIGURE 8.9: The ratio of fit correlators to lattice data for different models.

ref. [181]:

$$\rho^{\mathrm{LO}}(\omega) = \frac{d_A \omega^4 g^4}{2^{12}\pi^5} \coth\left(\frac{\omega}{4T}\right), \tag{8.7}$$

$$\rho^{\mathrm{NLO}}(\omega) = \rho^{\mathrm{LO}}(\omega) + \frac{d_A \omega^4}{2^{12}\pi^5} \coth\left(\frac{\omega}{4T}\right) \frac{g^6(\bar{\mu}) N_c}{(4\pi)^2} \times \left[\frac{22}{3} \ln\frac{\bar{\mu}^2}{\omega^2} + \frac{97}{3} + 8\,\phi_T(\omega)\right]. \tag{8.8}$$

Note that our definition of the spectral function differs from that in ref. [181] by a relative minus sign. Here $d_{\mathrm{A}} = N_c^2 - 1 = 8$ is the dimension of the adjoint representation, which counts the number of gluon states. At leading order one does not obtain a prescription on determining the value of the running coupling. We therefore make an educated guess of the renormalization point by choosing the value from the 1-loop order "EQCD" setup yielding (eq.(5.26) in [181])

$$\ln\left(\bar{\mu}^{\mathrm{opt}(T)}\right) \equiv \ln(4\pi T) - \gamma_{\mathrm{E}} - \frac{1}{22}\,. \tag{8.9}$$

Using this relation the coupling is fixed to the value $g^2\left(\bar{\mu}^{\mathrm{opt}(T)}\right) = 2.2346$ at $T = 1.5T_c$, where we use an updated relation $T_c = 1.24\Lambda_{\overline{\mathrm{MS}}}$ [171]. In order to compensate for the possibly bad choice of value of the coupling constant, we introduce an overall scaling coefficient B in the LO models. At NLO, the optimization of the scale $\bar{\mu}$ and the running of the strong coupling constant with ω is possible in the regime $\omega \gg \pi T$, as determined in [181]. In this regime the function $\phi_T(\omega)$, defined in eq.(4.4) of [181], is small which allows to define the optimized renormalization point as a function of ω as (eq.(5.25) in [181])

$$\ln\left(\bar{\mu}^{\mathrm{opt}(\omega)}\right) \equiv \ln(\omega) - \frac{97}{44}\,. \tag{8.10}$$

For values of ω outside this regime one again falls back to the renormalization point given by (8.9). Following the prescription given in [181], one uses the larger value of (8.9) and (8.10) for given ω. The switch between the two formulations happens at $\omega/T = 19.456\pi$.

In order to compensate for higher order corrections to the value of the renormalization point as well as other uncertainties in renormalization, we additionally introduce the scaling parameter B in the NLO models. Since at this temperature the perturbative series is not yet rapidly converging, we fit the lattice data to both the leading and next-to-leading order spectral function and consider the difference as an estimate of the uncertainties which arise due to our incomplete knowledge of the spectral function's high frequency functional form.

We have first performed a fit to the leading-order spectral function. This fit turned out to be rather poor, with a sum of residuals squared $\chi^2/\mathrm{d.o.f} = 68.6$. Slightly better, but still not in agreement with our data, was the fit to the next-to-leading order spectral function. Fitting eq. (8.8) with an additional multiplicative overall rescaling factor B gave $\chi^2/\mathrm{d.o.f.} = 33.2$. Thus, we had to add an additional structure to our fitting models, which represents a low frequency contribution to the spectral function. Note that the perturbative spectral functions are expected to not capture the low frequency part correctly. Since our theoretical understanding of the behavior of this particular spectral function is rather limited, we will consider three possibilities.

The first model is motivated by the appearance of a sharp structure in perturbative calculations [182]. We implement this feature by using a simple δ-peak in ρ/ω where we treat its overall coefficient A/T^4 as another fit parameter. We use that modification in both leading- and next-to-leading order. The combination of that δ-peak with B times eq. (8.7) is denoted as model M1, while the combination with B times eq. (8.8) is called model M4. The later includes B as an overall normalization factor in the high-frequency function.

Another candidate for the transport peak is a Breit-Wigner distribution, e.g.

$$\rho_{\mathrm{peak}}/\omega = (A/T^4)CT^2/(C^2T^2 + \omega^2). \tag{8.11}$$

In contrast to the δ-peak, the Breit-Wigner distribution is much broader. Since the actual coupling is rather large, a broader peak might be more reliable, and hence that distribution is a good alternative. Similar to the δ peak models we combine the Breit-Wigner distribution with eq. (8.7) and call this Model M2. The combination with eq. (8.8) is called M5. In these two models, however, we consider a few distinct C values, varying from a rather narrow to a quite wide structure. Additionally, we treat A/T^4 as a fitting parameter. Model M3 and M6 differ from model M2 and M5 by the large-width limit of the Breit-Wigner distribution, i.e. $\rho_{\mathrm{peak}}/\omega = A/T^4$.

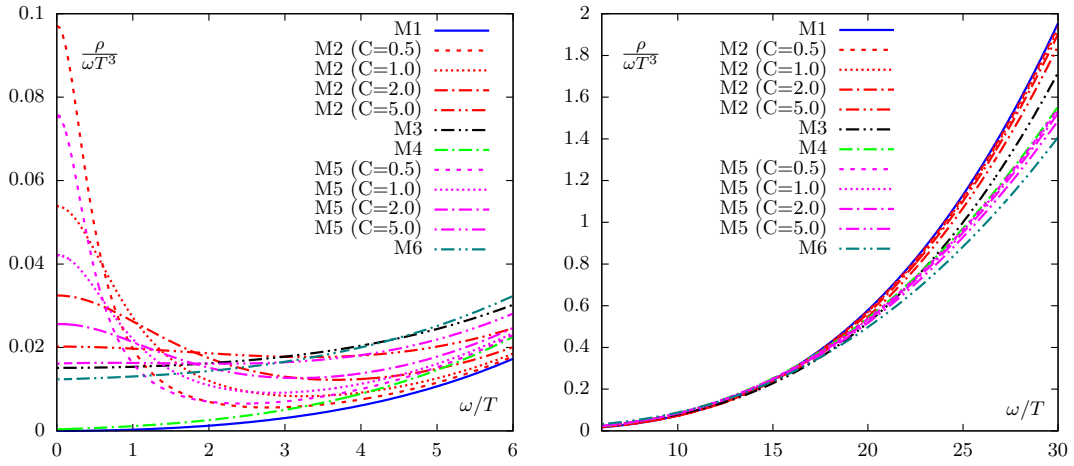


FIGURE 8.10: *Left:* Small frequency parts of the fit spectral functions from different models. For M1 and M4 there is a δ-peak at zero frequency, which is invisible in the plot. *Right:* Same as the left panel but in larger ω/T range.

To summarize, our models are:

$$\begin{aligned}
\text{M1}: \frac{\rho(\omega)}{\omega T^3} &= \frac{A}{T^4}\delta\left(\frac{\omega}{T}\right) + B\frac{\rho^{\text{LO}}(\omega)}{\omega T^3} \\
\text{M2}: \frac{\rho(\omega)}{\omega T^3} &= \frac{A}{T^4}\frac{CT^2}{C^2T^2+\omega^2} + B\frac{\rho^{\text{LO}}(\omega)}{\omega T^3} \\
\text{M3}: \frac{\rho(\omega)}{\omega T^3} &= \frac{A}{T^4} + B\frac{\rho^{\text{LO}}(\omega)}{\omega T^3} \\
\text{M4}: \frac{\rho(\omega)}{\omega T^3} &= \frac{A}{T^4}\delta\left(\frac{\omega}{T}\right) + B\frac{\rho^{\text{NLO}}(\omega)}{\omega T^3} \\
\text{M5}: \frac{\rho(\omega)}{\omega T^3} &= \frac{A}{T^4}\frac{CT^2}{C^2T^2+\omega^2} + B\frac{\rho^{\text{NLO}}(\omega)}{\omega T^3} \\
\text{M6}: \frac{\rho(\omega)}{\omega T^3} &= \frac{A}{T^4} + B\frac{\rho^{\text{NLO}}(\omega)}{\omega T^3}.
\end{aligned} \tag{8.12}$$

The final sphaleron rate is then $2T^4$ times the value of the right-hand side at $\omega = 0$. For M2 and M5 the sphaleron rate can be calculated as $\Gamma_{\text{sphal}}/T^4 = 2A/CT^4$ while for M3 and M6 it is $\Gamma_{\text{sphal}}/T^4 = 2A/T^4$. For M1 and M4, the δ-peak leads to an infinite sphaleron rate.

As a fitting procedure we use the Levenberg-Marquardt algorithm in the χ^2-fitting, where the iterations are stopped as soon as the relative tolerance 10^{-7} is reached for the fit parameters or the $\chi^2/\text{d.o.f.}$ The fits are only using data beyond $\tau T = 0.25$, since the flow-time window over which we can perform the $\tau_{\text{F}} \to 0$ extrapolation closes up as we go to shorter and shorter separations. Also, information about the transport peak is mainly encoded in the correlator at large separations.

We summarize our fit results in table 8.2. It seems that the lattice data tightly constrains the fit parameters in all our models. The quality of the fit ($\chi^2/\text{d.o.f.}$) is also fairly good. The fits based on the next-to-leading order spectral function turned out to be better than the ones with only the leading order spectral function. Looking at the low-frequency behavior, however, no model is clearly preferred. The resulting spectral functions are shown in figure 8.10. For comparison we also show the fitted correlators divided by the lattice data in figure 8.9.

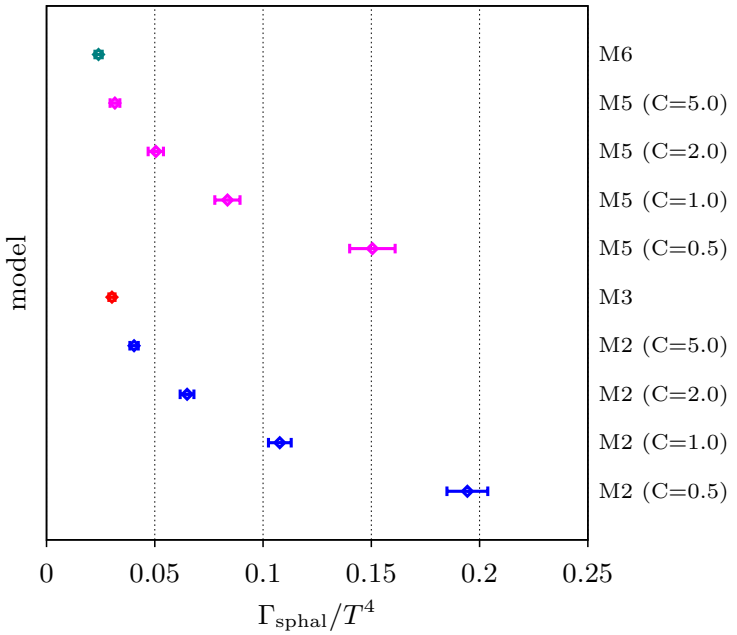


FIGURE 8.11: The sphaleron rate from different models at $1.5\,T_c$.

Finally, we extract the sphaleron rate for the different models according to eq. (8.1). The results are shown in figure 8.11. Our analysis shows, that a linear-in-frequency transport peak as well as a δ-like transport peak are just special cases of a Breit-Wiger transport peak with zero and infinite width. Within the set of fit models we have considered, we find that the sphaleron rate varies within the range at $1.5\,T_c$

$$\begin{aligned} \Gamma_{\rm spha}/T^4 &\geq 0.030(2), \quad \text{based on M1-M3}, \\ \Gamma_{\rm spha}/T^4 &\geq 0.024(2), \quad \text{based on M4-M6}. \end{aligned} \tag{8.13}$$

The higher bound in the first line uses the results of the models based on the LO perturbative series, where the coupling constant is fixed. The results of the models based on the NLO perturbative series are used in the second line. Here, the coupling constant runs with scale.

We see in our results that fitting a spectral function with only a high-frequency part is not sufficient. Hence, there *must* be a low-frequency structure in addition. Two extreme forms have been considered in this section: An infinitely sharp peak (M1/M4) and an infinitely broad peak (M3/M6). Therefore we consider them to span the range of likely functional forms for a "peak", and we propose that the lower value found can be viewed as a lower limit on the sphaleron rate. We are aware that this claim depends somewhat on the choice of fitting functions considered. However, we believe our bound is reasonable, since we have considered many peak structures ranging from sharp to perfectly wide.

8.5 Conclusion

We have computed the topological charge density correlation functions using the gradient flow on 5 different large and fine isotropic lattices in the quenched approximation. The gradient flow eliminates the non-topological contributions in $q(x)$ and reduces short-distance fluctuations in the correlation functions. We developed a methodology

Ansatz	$A/T^4 \times 10$	B	C	χ^2/d.o.f	$\Gamma_{\mathrm{sphal}}/T^4 \times 10$
M1	0.68(4)	2.27(7)		1.86	∞
M2	0.49(3)	2.25(7)	0.5	2.07	1.94(10)
M2	0.54(3)	2.24(7)	1.0	2.02	1.08(6)
M2	0.65(4)	2.21(7)	2.0	1.93	0.65(4)
M2	1.01(5)	2.14(7)	5.0	1.76	0.40(2)
M3	0.15(1)	1.98(8)		1.36	0.30(2)
M4	0.53(4)	1.25(4)		1.35	∞
M5	0.38(3)	1.25(4)	0.5	1.53	1.50(11)
M5	0.42(3)	1.24(4)	1.0	1.51	0.84(6)
M5	0.50(4)	1.23(4)	2.0	1.48	0.50(4)
M5	0.79(6)	1.20(4)	5.0	1.47	0.32(3)
M6	0.12(1)	1.12(5)		1.29	0.24(2)

TABLE 8.2: Fitted parameters and $\chi^2/\mathrm{d.o.f}$ for different ansätze. The numbers in the parentheses are statistical uncertainties from a bootstrap analysis. The sphaleron rate $\Gamma_{\mathrm{sphal}}/T^4$ is calculated from the fitted parameters A/T^4 and C. The meaning of each fit parameter and how the sphaleron rate is determined in each ansatz can be found in (8.4).

to perform reliable continuum- and flow-time extrapolations for the topological charge density correlation function. These double-extrapolated correlation functions are then used in the spectral reconstruction, where we used models inspired by perturbation theory to account for uncertainties from different sources.

Using either the LO or the NLO spectral function *alone* leads to a very poor fit; the data demand the addition of a low-frequency structure. In our spectral analysis we considered a range of spectral functions. The fits with the next-to-leading order spectral function, for the case where the added structure is completely flat, gave us a lower bound for the sphaleron rate of at least $\Gamma_{\mathrm{spha}}/T^4 \geq 0.024$ at $1.5T_c$. Low-frequency structures containing an actual peak give higher values for Γ_{spha}. The leading-order results gave us also a slightly higher value of $\Gamma_{\mathrm{spha}}/T^4 \geq 0.030$.

To correctly capture the topological charge, it is essential to use the gradient flow. However, the $\tau_{\mathrm{F}} \to 0$ extrapolation can only be done on a limited flow-time range to obtain proper extrapolated values. Coarse lattices have the problem that we can not fulfill these constraints at small separations. Therefore, we are driven to volumes with very fine lattice spacings and can only make use of larger separations. Even though the continuum and flow-time extrapolations of the lattice data returned values with only a few percent statistical error bars, the subsequent fitting of the spectral function was lacking clear theoretical guidance, resulting in still rather large errors in the final sphaleron rate. Although we have found a lower bound for the sphaleron rate, we still need an upper bound. Therefore, a better theoretical understanding of the expected form for the topological charge-density spectral function would be very

useful, especially in the low-frequency region. Having such improvements may allow for a reliable determination of the sphaleron rate.

Chapter 9

Conclusion

In this final chapter we summarize and conclude the major results of this work. More detailed conclusions can be found in the previous chapters.

A novel multi-GPU framework has been developed that is capable of performing large scale lattice QCD computations in parallel on multiple GPUs. The framework encapsulates frequently used routines in classes and provides physicists with an interface that allows them to implement their own lattice QCD calculations in a simple and performant way. Base classes such as `Gaugefield`, `Spinorfield`, `CommunicationBase`, `GIndexer`, `LatticeContainer`, `MemoryManagement` and many more have been highly optimized for memory access, cache flow and peer-to-peer communication such that the best performance on current HPC hardware can be achieved. The two biggest modules which have been implemented within this framework are the RHMC and the gradient flow. The later has extensively been used in this work to smooth out fluctuation on already existing configurations, while the former will be used in the near future to generate large HISQ configurations.

We have measured the topological susceptibility on $2+1$ HISQ configurations with varying light quark masses and fixed physical strange quark mass in the temperature range $137-167$ MeV. In order to do this we have smeared the configurations using the gradient flow and then extracted the topological susceptibility from a flow-time plateau. Our results did not show any mass dependence. This is contrary to the results which have been reported in ref. [183], where the topological susceptibility has been measured in 2-flavor QCD using Möbius domain-wall fermions and the gluonic definition of the topological charge via gradient flow as well as the fermionic definition using the overlap Dirac operator in the temperature range $190-220$ MeV. However, since our results have shown large cut-off effects, we would need to compute the topological susceptibility in the future on special choices of lattice spacings and then perform a continuum extrapolation to get a final conclusive evidence on the mass dependence of the susceptibility. Considering the temperature behavior of the topological susceptibility, we have observed that the susceptibility decreases with increasing temperature. This is in agreement with previous studies and expectations.

We have additionally measured the Dirac eigenvalue spectrum on some selected HISQ configurations with mass ratios $m_s/m_l = 27, 40, 80$ of the previously mentioned ensemble. The results have been used to estimate the fate of anomalous $\mathrm{U}_A(1)$ symmetry as we approach the chiral limit along the line of constant physical strange quark mass. Since we were using different valence and sea quark actions, we have tuned the overlap valence quark masses to the HISQ sea quark mass and used these tuned values to appropriately renormalize the eigenvalue density. It turned out that the eigenvalue density can be represented as $\rho(\lambda) \sim \lambda$ in the chiral limit at a temperature of $1.05\ T_c$, which is contrary to the expectations of ref. [71]. There, the leading order behavior of the eigenvalue spectrum was derived to be $\rho(\lambda) \sim \lambda^3$ for two flavor QCD in the chiral limit. Finally, the value of our renormalized observable

$m_l^2(\chi_\pi - \chi_\delta)/T^4$ is non-zero, and hence we conclude that $\mathrm{U}_A(1)$ is broken when we approach the chiral limit along the line of constant physical strange quark mass. For final conclusive evidence it is necessary to perform a continuum extrapolation of our observables. However, so far that was not possible within this work, since the required configurations with smaller lattice spacings were not available. Nevertheless, this is one of the first studies investigating the fate of the $\mathrm{U}_A(1)$ anomalous symmetry, for light quark masses as low as $m_s/m_l = 80$ (i.e. a pion mass of ~ 80 MeV) just above the chiral crossover temperature, which may suggest that we would eventually encounter the $\mathrm{O}(4)$ second order line of the phase transitions.

The final study in this work was an exploratory work on the QCD sphaleron rate. We have measured the topological charge density correlation functions on fine isotropic lattices in the quenched approximation. We used the gradient flow to eliminate non-topological contributions in $q(x)$ as well as to reduce short-distance fluctuations in the correlation functions. On this lattice data we have performed a continuum extrapolation and a flow-time-to-zero extrapolation. After this double extrapolation, we have extracted the spectral function using different models motivated by perturbation theory and a transport peak. With these results we have obtained lower bounds for the sphaleron rate. Using a next-to-leading order ansatz, we obtained $\Gamma_{\mathrm{spha}}/T^4 \geq 0.024$ at $1.5T_c$, while for the leading-order results we obtained $\Gamma_{\mathrm{spha}}/T^4 \geq 0.030$. An upper bound is still missing. For a more robust determination of the sphaleron rate we need a better theoretical understanding of the expected form for the topological charge-density spectral function as well as a clear theoretical guidance on how to fit the spectral function. So far this study was limited to the quenched approximation, but provides a methodology that can be extended to full QCD, e.g. based on gauge field configurations generated with the multi-GPU code developed in this work.

Appendix A

Appendix

This appendix includes some basic notes, which have not been written out earlier in this thesis.

A.1 Gamma matrices

In Euclidean spacetime, the γ_μ matrices with $\mu = 1, 2, 3, 4$ are defined by

$$\{\gamma_\mu, \gamma_v\} = \gamma_\mu\gamma_v + \gamma_\nu\gamma_\mu = 2\delta_{\mu v}\mathbb{1}_{4\times4}. \tag{A.1}$$

In addition to those we define a fifth matrix

$$\gamma_5 = \gamma_1\gamma_2\gamma_3\gamma_4, \tag{A.2}$$

which anti-commutes with all other gamma matrices $\gamma_\mu, \mu = 1, 2, 3, 4$. The Euclidean gamma matrices are traceless and obey

$$\gamma_\mu^2 = \mathbb{1}_{4\times4}, \quad \gamma_5^2 = \mathbb{1}_{4\times4}, \quad \gamma_\mu^\dagger = \gamma_\mu, \quad \gamma_5^\dagger = \gamma_5. \tag{A.3}$$

The chiral representation of the gamma matrices, where γ_5 is diagonal, reads:

$$\gamma_{1,2,3} = \begin{bmatrix} 0 & -\mathrm{i}\sigma_{1,2,3} \\ \mathrm{i}\sigma_{1,2,3} & 0 \end{bmatrix}, \gamma_4 = \begin{bmatrix} 0 & \mathbb{1}_{2\times2} \\ \mathbb{1}_{2\times2} & 0 \end{bmatrix}, \gamma_5 = \begin{bmatrix} \mathbb{1}_{2\times2} & 0 \\ 0 & -\mathbb{1}_{2\times2} \end{bmatrix}. \tag{A.4}$$

where the σ_j are the Pauli matrices

$$\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_2 = \begin{pmatrix} 0 & -\mathrm{i} \\ \mathrm{i} & 0 \end{pmatrix}, \quad \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{A.5}$$

In a more explicit form the Euclidean gamma matrices read

$$\gamma_1 = \begin{bmatrix} 0 & 0 & 0 & -\mathrm{i} \\ 0 & 0 & -\mathrm{i} & 0 \\ 0 & \mathrm{i} & 0 & 0 \\ \mathrm{i} & 0 & 0 & 0 \end{bmatrix}, \gamma_2 = \begin{bmatrix} 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ -1 & 0 & 0 & 0 \end{bmatrix}, \gamma_3 = \begin{bmatrix} 0 & 0 & -\mathrm{i} & 0 \\ 0 & 0 & 0 & \mathrm{i} \\ \mathrm{i} & 0 & 0 & 0 \\ 0 & -\mathrm{i} & 0 & 0 \end{bmatrix},$$
$$\gamma_4 = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}, \gamma_5 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix}. \tag{A.6}$$

A.2 Pure gauge

The gluon field strength tensor reads:

$$G_{\mu\nu}(x) = \partial_\mu A_\nu - \partial_\nu A_\mu + ig[A_\mu, A_\nu] \tag{A.7}$$

Let us insert the pure gauge $\tilde{A}_\mu = -\frac{i}{g} U \partial_\mu U^\dagger$. Note that using the product rule, $\partial_\mu \tilde{A}_\nu$ becomes:

$$\partial_\mu \tilde{A}_\nu = -\frac{i}{g} \partial_\mu U \partial_\nu U^\dagger - \frac{i}{g} U \partial_\mu \partial_\nu U^\dagger. \tag{A.8}$$

Therefore for the first two terms we get:

$$\partial_\mu \tilde{A}_\nu - \partial_\nu \tilde{A}_\mu = \frac{i}{g} \partial_\nu U \partial_\mu U^\dagger - \frac{i}{g} \partial_\mu U \partial_\nu U^\dagger \tag{A.9}$$

The last term of equation (A.7) transforms to:

$$ig[\tilde{A}_\mu, \tilde{A}_\nu] = \frac{i}{g} U(\partial_\nu U^\dagger) U \partial_\mu U^\dagger - \frac{i}{g} U(\partial_\mu U^\dagger) U \partial_\nu U^\dagger \tag{A.10}$$

Note that $U(\partial_\mu U^\dagger)U = -\partial_\mu U$, since

$$\begin{aligned} 0 &= \partial_\mu \mathbb{1} = \partial_\mu U U^\dagger \\ 0 &= (\partial_\mu U) U^\dagger + U(\partial_\mu U^\dagger) \\ 0 &= (\partial_\mu U) + U(\partial_\mu U^\dagger) U \\ &\Rightarrow U(\partial_\mu U^\dagger) U = -\partial_\mu U \end{aligned} \tag{A.11}$$

Inserting (A.11) in (A.10) results in

$$ig[\tilde{A}_\mu, \tilde{A}_\nu] = -\frac{i}{g} \partial_\nu U \partial_\mu U^\dagger + \frac{i}{g} \partial_\mu U \partial_\nu U^\dagger, \tag{A.12}$$

and therefore the full gluon field strength tensor vanishes in pure gauge.

List of Figures

2.1 Two possible scenarios of the QCD phase diagram in the plane of the light quark masses m_{ud} and the strange quark mass m_s. *Left:* Phase diagram where the universality class of the phase transition in the chiral limit depends on the strange quark mass. For strange quark mass below the tricritical point the transition is of first order. Above the tricritical point it is of second order belonging either to the $\mathrm{O}(4)$ or the $\mathrm{U}_L(2) \times \mathrm{U}_R(2)/\,\mathrm{U}_V(2)$ universality class. *Right:* Phase diagram where the transition in the chiral limit is of first order. The first order region and the crossover region in both scenarios is separated by a line of second order transitions belonging to the $\mathrm{Z}(2)$ universality class. . . . 15
2.2 Illustration of scalar and pseudo-scalar mesons which are related by symmetry transformations. If a symmetry is restored then the corresponding states connected with arrows will become degenerate. . . . 17

3.1 Illustration of compactifying $\mathbb{R}^1 \cup \{\infty\}$ on S^1 20
3.2 Illustration of two transformations. The endpoints have to fulfill $\lim_{x\to\infty} \alpha_n(x) = 2\pi n$. Therefore no smooth deformation between $\alpha_0(x)$ and $\alpha_1(x)$ possible. 21
3.3 Illustration of a one dimensional mapping with winding number $n = 4$. 22

5.1 Halo pattern of a 2-dimensional lattice field. 52
5.2 Illustration showing different communication channel. 54
5.3 Illustration showing which data is communicated on a 2D lattice. The outer rectangle represents the inner halo region which is communicated. The green circles represent the sites of the lattice and the arrows represent the links between the sites. A spinor field would just communicate the green circles within the halo region, while a gauge field would communicate the arrows within the halo region. 57
5.4 Single-Node strong scaling plots. These benchmarks have been performed on the Bielefeld GPU cluster. *Left:* Performance of the gradient flow program for a $64^3 \times 16$ lattice. *Right:* Performance of the HISQ gauge field generation program based on a rational hybrid Monte Carlo algorithm (RHMC) for a $64^3 \times 16$ lattice. 61
5.5 Multi-Node strong scaling plots. These benchmarks have been performed on Summit in Oak Ridge National Laboratory, USA. *Left:* Performance of the gradient flow program for a $96^3 \times 24$ lattice. *Right:* Performance of the HISQ gauge field generation program based on a rational hybrid Monte Carlo algorithm (RHMC) for a 96^4 lattice. . . . 61
5.6 Single-Node weak scaling plots. These benchmarks have been performed on the Bielefeld GPU cluster. *Left:* Performance of the gradient flow. The local lattice size on a single GPU has been fixed to $64^3 \times 16$. *Right:* Performance of the DSlash. The local lattice size on a single GPU has been fixed to $80^3 \times 20$. 62

6.1 Trajectory of the topological charge at temperature $T \approx 166$MeV. The x-axis shows the configuration number. Subsequent configurations are separated by 50 hybrid Monte Carlo time steps. Different streams have been concatenated. The topological charge is not always close to integer values. 64
6.2 Autocorrelation times of the topological charge at temperature $T \approx 166$MeV. Subsequent configurations are separated by 50 hybrid Monte Carlo time steps. Different streams have been concatenated. 64
6.3 Topological susceptibility as a function of the flow-time. The green linear error bands show the plateau region which we use to extract the topological susceptibility. 67
6.4 Topological susceptibilities as a function of the temperature to show whether some volume dependence is visible. Different plots show different temporal extent N_τ and different light quark masses. All curves overlap within two errorbars. Hence, we do not see clear finite volume effects. 68
6.5 Comparison of the topological susceptibility as a function of the temperature at different light quark masses but fixed physical strange quark mass. Within two error bars, the curves overlap. Hence, no clear mass dependence visible within our statistics. 69
6.6 Topological susceptibility as a function of the temperature of $2+1$ HISQ configurations at physical quark masses. The error band of the fit has been computed using gaussian bootstrap samples. 69
6.7 Comparison of topological susceptibility as a function of the temperature at different lattice spacings. Strong cut-off effects are visible. 71

7.1 The eigenvalue density of the massless valence overlap Dirac operator measured on HISQ sea configurations with $m_s/m_l = 27$, as a function of temperature, near and just above T_c. 77
7.2 The eigenvalue density of the valence massless overlap Dirac operator measured on the HISQ sea configurations with $m_s/m_l = 40$. 78
7.3 The coefficient $c(m_l)$ of the λ dependent term of the eigenvalue density shown as a function of light sea quark mass squared m_l^2. Both axes are normalized by appropriate powers of the temperature. The χ^2 per degree of freedom of the fit is 0.95. 79
7.4 The renormalized eigenvalue density of the QCD ensembles at $T = 1.05\ T_c$ generated using HISQ discretization and measured using an appropriately mass-tuned valence overlap operator. These are shown for three different choices of the light quarks. 81
7.5 Comparison of $\chi_\pi - \chi_\delta$ measured using the overlap eigenvalues to the continuum estimates of χ_{disc} using data from ref. [62], shown for physical quark masses. 82
7.6 The renormalized $\mathrm{U}_A(1)$ breaking parameter shown as a function of light quark mass at $1.05\ T_c$. 83

8.1 Topological susceptibility measured at different flow-times. The vertical line is the flow-time which we have chosen as a lower limit to ensure that q properly captures topology. 88

8.2 Trajectory of the topological charge at $\tau_{\mathrm{F}} \approx a^2$. The x-axis shows the configuration number. Subsequent configurations are separated by 500 sweeps. Different streams have been concatenated. 89
8.3 Topological charge density correlator of the $N_\tau = 16$ lattice at flow times $\sqrt{8\tau_F^{\max}}T = 0.5220\tau T$ (the largest flow-times used in our analysis). Top: Comparison of the full correlator with the correlators of the $Q = 0$ and $Q = 1$ sectors. Bottom: Difference between the full correlator and the correlator computed using only the $Q = 0$ sector, in comparison to the statistical error bars. 90
8.4 Autocorrelation time of the topological charge density correlator at separation $\tau T = 0.5$ of the finest lattice $N_\tau = 36$. 90
8.5 Topological charge density correlation function at flow-time $\sqrt{8\tau_{\mathrm{F}}}T = 0.15$ for all lattice spacings. Cubic splines have been used to perform these interpolations. The error bands have been calculated during the interpolation procedure. 91
8.6 Continuum extrapolation of some selected separations at flow-times halfway between the beginning and the end of each flow range (8.5). . 92
8.7 Continuum-extrapolated topological charge density correlator as a function of flow-time. The colored parts are the regions which are used to perform the flow-time extrapolations, while the black lines on top indicate the linear extrapolations. The gray regions are not considered in the flow-time extrapolation. 93
8.8 Continuum extrapolated topological charge density correlator as well as the final flow-time extrapolated correlator. 94
8.9 The ratio of fit correlators to lattice data for different models. 95
8.10 *Left:* Small frequency parts of the fit spectral functions from different models. For M1 and M4 there is a δ-peak at zero frequency, which is invisible in the plot. *Right:* Same as the left panel but in larger ω/T range. 97
8.11 The sphaleron rate from different models at $1.5\,T_c$. 98

List of Tables

5.1 Different types of halo segments on a 4-dimensional cube. N_x, N_y, N_z and N_t are the lattice dimensions. H_x, H_y, H_z and H_t are the halo depths. 51

6.1 Results of the extracted topological susceptibility. The temperatures have been determined using the f_k-scale [59] with data from [61]. . . 65

7.1 The details of the HISQ configurations analyzed in this work. 74
7.2 The temperature (T), the exponent γ characterizing the leading order λ^γ rise of the bulk eigenvalues λ and the goodness of fits performed on eigenvalue densities for different choices of the light sea-quarks and physical value of strange sea-quarks. 79
7.3 The valence strange quark masses obtained by matching the observable Δ measured using the eigenvalue density of the valence overlap to that measured by inversion of the sea HISQ Dirac operator. 80

8.1 Lattice spacings, lattice extents, β values and statistics of configurations in this work. The lattice spacing a is determined by the Sommer scale (see [172]). 86
8.2 Fitted parameters and $\chi^2/\mathrm{d.o.f}$ for different ansätze. The numbers in the parentheses are statistical uncertainties from a bootstrap analysis. The sphaleron rate $\Gamma_{\mathrm{sphal}}/T^4$ is calculated from the fitted parameters A/T^4 and C. The meaning of each fit parameter and how the sphaleron rate is determined in each ansatz can be found in (8.4). 99

Bibliography

[1] David J. Gross, Robert D. Pisarski, and Laurence G. Yaffe. “QCD and Instantons at Finite Temperature”. In: *Rev. Mod. Phys.* 53 (1981), 43. DOI: 10.1103/RevModPhys.53.43.

[2] H. Leutwyler and Andrei V. Smilga. “Spectrum of Dirac operator and role of winding number in QCD”. In: *Phys. Rev. D* 46 (1992), 5607–5632. DOI: 10.1103/PhysRevD.46.5607.

[3] P. Di Vecchia and G. Veneziano. “Chiral Dynamics in the Large n Limit”. In: *Nucl. Phys. B* 171 (1980), 253–272. DOI: 10.1016/0550-3213(80)90370-3.

[4] Thomas Schäfer and Edward V. Shuryak. “Instantons in QCD”. In: *Rev. Mod. Phys.* 70 (1998), 323–426. DOI: 10.1103/RevModPhys.70.323.

[5] Dmitri E. Kharzeev, Larry D. McLerran, and Harmen J. Warringa. “The Effects of topological charge change in heavy ion collisions: ’Event by event P and CP violation’”. In: *Nucl. Phys. A* 803 (2008), 227–253. DOI: 10.1016/j.nuclphysa.2008.02.298.

[6] Kenji Fukushima, Dmitri E. Kharzeev, and Harmen J. Warringa. “The Chiral Magnetic Effect”. In: *Phys. Rev. D* 78 (2008), 074033. DOI: 10.1103/PhysRevD.78.074033.

[7] A. Yu. Kotov. “Sphaleron Transition Rate in Lattice Gluodynamics”. In: *JETP Letters* 108.6 (2018), 352–355. DOI: 10.1134/S0021364018180078.

[8] Maria Paola Lombardo and Anton Trunin. “Topology and axions in QCD”. In: *Int. J. Mod. Phys. A* 35.20 (2020), 2030010. DOI: 10.1142/S0217751X20300100.

[9] Giovanni Grilli di Cortona et al. “The QCD axion, precisely”. In: *JHEP* 01 (2016), 034. DOI: 10.1007/JHEP01(2016)034.

[10] Marco Gorghetto and Giovanni Villadoro. “Topological Susceptibility and QCD Axion Mass: QED and NNLO corrections”. In: *JHEP* 03 (2019), 033. DOI: 10.1007/JHEP03(2019)033.

[11] Stephen L. Adler. “Axial vector vertex in spinor electrodynamics”. In: *Phys. Rev.* 177 (1969), 2426–2438. DOI: 10.1103/PhysRev.177.2426.

[12] J.S. Bell and R. Jackiw. “A PCAC puzzle: $\pi^0 \to \gamma\gamma$ in the σ model”. In: *Nuovo Cim. A* 60 (1969), 47–61. DOI: 10.1007/BF02823296.

[13] Kazuo Fujikawa. “Path Integral Measure for Gauge Invariant Fermion Theories”. In: *Phys. Rev. Lett.* 42 (1979), 1195–1198. DOI: 10.1103/PhysRevLett.42.1195.

[14] Shailesh Chandrasekharan et al. “Anomalous chiral symmetry breaking above the QCD phase transition”. In: *Phys. Rev. Lett.* 82 (1999), 2463–2466. DOI: 10.1103/PhysRevLett.82.2463.

[15] A. Bazavov et al. “The chiral transition and $U(1)_A$ symmetry restoration from lattice QCD using Domain Wall Fermions”. In: *Phys. Rev. D* 86 (2012), 094503. DOI: 10.1103/PhysRevD.86.094503.

[16] Michael I. Buchoff et al. “QCD chiral transition, U(1)A symmetry and the dirac spectrum using domain wall fermions”. In: *Phys. Rev. D* 89.5 (2014), 054514. DOI: 10.1103/PhysRevD.89.054514.

[17] H. Ohno et al. “Eigenvalue distribution of the Dirac operator at finite temperature with (2+1)-flavor dynamical quarks using the HISQ action”. In: *PoS* LATTICE2011 (2011). Ed. by Pavlos Vranas, 210. DOI: 10.22323/1.139.0210.

[18] H. Ohno et al. “$U_A(1)$ breaking at finite temperature from the Dirac spectrum with the dynamical HISQ action”. In: *PoS* LATTICE2012 (2012). Ed. by Derek Leinweber et al., 095. DOI: 10.22323/1.164.0095.

[19] Alexei Bazavov et al. “Meson screening masses in (2+1)-flavor QCD”. In: *Phys. Rev. D* 100.9 (2019), 094510. DOI: 10.1103/PhysRevD.100.094510.

[20] Florian Burger et al. “Chiral observables and topology in hot QCD with two families of quarks”. In: *Phys. Rev. D* 98.9 (2018), 094501. DOI: 10.1103/PhysRevD.98.094501.

[21] Lukas Holicki, Ernst-Michael Ilgenfritz, and Lorenz von Smekal. “The Anderson transition in QCD with $N_f = 2 + 1 + 1$ twisted mass quarks: overlap analysis”. In: *PoS* LATTICE2018 (2018), 180. DOI: 10.22323/1.334.0180.

[22] Guido Cossu et al. “Finite temperature study of the axial U(1) symmetry on the lattice with overlap fermion formulation”. In: *Phys. Rev. D* 87.11 (2013). [Erratum: Phys.Rev.D 88, 019901 (2013)], 114514. DOI: 10.1103/PhysRevD.87.114514.

[23] Akio Tomiya et al. “Effects of near-zero Dirac eigenmodes on axial U(1) symmetry at finite temperature”. In: *PoS* LATTICE2014 (2015), 211. DOI: 10.22323/1.214.0211.

[24] A. Tomiya et al. “Evidence of effective axial U(1) symmetry restoration at high temperature QCD”. In: *Phys. Rev. D* 96.3 (2017). [Addendum: Phys.Rev.D 96, 079902 (2017)], 034509. DOI: 10.1103/PhysRevD.96.034509.

[25] Kei Suzuki et al. “Axial U(1) symmetry and Dirac spectra in high-temperature phase of $N_f = 2$ lattice QCD”. In: *PoS* LATTICE2018 (2018), 152. DOI: 10.22323/1.334.0152.

[26] Kei Suzuki et al. “Axial U(1) symmetry and mesonic correlators at high temperature in $N_f = 2$ lattice QCD”. In: *PoS* LATTICE2019 (2020), 178. DOI: 10.22323/1.363.0178.

[27] Ting-Wai Chiu et al. “Chiral symmetry and axial U(1) symmetry in finite temperature QCD with domain-wall fermion”. In: *PoS* LATTICE2013 (2014), 165. DOI: 10.22323/1.187.0165.

[28] Bastian B. Brandt et al. “On the strength of the $U_A(1)$ anomaly at the chiral phase transition in $N_f = 2$ QCD”. In: *JHEP* 12 (2016), 158. DOI: 10.1007/JHEP12(2016)158.

[29] “The TOP500 List and Progress in High-Performance Computing”. URL: https://top500.org/.

[30] “Bielefeld GPU cluster”. URL: https://www2.physik.uni-bielefeld.de/gpu-cluster.html.

[31] Eduardo Follana et al. “Further improvements to staggered quarks”. In: *Nucl. Phys. B Proc. Suppl.* 129 (2004). Ed. by S. Aoki et al., 447–449. DOI: 10.1016/S0920-5632(03)02610-0.

[32] Patrick Steinbrecher. “The QCD crossover up to $O(\mu_B^6)$ from Lattice QCD”. PhD thesis. Bielefeld University, 2018.

[33] Lukas Mazur. “Applications of the Gradient Flow method in Lattice QCD”. Master’s thesis. Bielefeld University, 2017.

[34] Martin Luscher and Peter Weisz. “Perturbative analysis of the gradient flow in non-abelian gauge theories”. In: *JHEP* 02 (2011), 051. DOI: 10.1007/JHEP02(2011)051.

[35] Martin Lüscher. “Properties and uses of the Wilson flow in lattice QCD”. In: *JHEP* 08 (2010). [Erratum: JHEP 03, 092 (2014)], 071. DOI: 10.1007/JHEP08(2010)071.

[36] R. Narayanan and H. Neuberger. “Infinite N phase transitions in continuum Wilson loop operators”. In: *JHEP* 03 (2006), 064. DOI: 10.1088/1126-6708/2006/03/064.

[37] Herbert Neuberger. “Exactly massless quarks on the lattice”. In: *Phys. Lett. B* 417 (1998), 141–144. DOI: 10.1016/S0370-2693(97)01368-3.

[38] Herbert Neuberger. “More about exactly massless quarks on the lattice”. In: *Phys. Lett. B* 427 (1998), 353–355. DOI: 10.1016/S0370-2693(98)00355-4.

[39] Guy D. Moore and Marcus Tassler. “The Sphaleron Rate in SU(N) Gauge Theory”. In: *JHEP* 02 (2011), 105. DOI: 10.1007/JHEP02(2011)105.

[40] N.S. Manton. “Topology in the Weinberg-Salam Theory”. In: *Phys. Rev. D* 28 (1983), 2019. DOI: 10.1103/PhysRevD.28.2019.

[41] Frans R. Klinkhamer and N.S. Manton. “A Saddle Point Solution in the Weinberg-Salam Theory”. In: *Phys. Rev. D* 30 (1984), 2212. DOI: 10.1103/PhysRevD.30.2212.

[42] Steven Weinberg. *The Quantum theory of fields. Vol. 1: Foundations*. Cambridge University Press, 2005.

[43] Michel Le Bellac. *Thermal Field Theory*. Cambridge Monographs on Mathematical Physics. Cambridge University Press, 2011. DOI: 10.1017/CBO9780511721700.

[44] Joseph I. Kapusta. *Finite Temperature Field Theory*. Vol. 360. Cambridge Monographs on Mathematical Physics. Cambridge University Press, 1989.

[45] Christof Gattringer and Christian B. Lang. *Quantum chromodynamics on the lattice*. Vol. 788. Springer, 2010. DOI: 10.1007/978-3-642-01850-3.

[46] I. Montvay and G. Munster. *Quantum fields on a lattice*. Cambridge Monographs on Mathematical Physics. Cambridge University Press, 1997. DOI: 10.1017/CBO9780511470783.

[47] H.J. Rothe. *Lattice Gauge Theories: An Introduction*. World Scientific Lecture Notes In Physics. World Scientific Publishing Company, 1992.

[48] G. ’t Hooft. “Symmetry Breaking through Bell-Jackiw Anomalies”. In: *Phys. Rev. Lett.* 37 (1 1976), 8–11. DOI: 10.1103/PhysRevLett.37.8.

[49] Sidney R. Coleman. “The Uses of Instantons”. In: *Subnucl. Ser.* 15 (1979), 805.

[50] Stefan Vandoren and Peter van Nieuwenhuizen. “Lectures on instantons”. 2008. arXiv: 0802.1862 [hep-th].

[51] H. Forkel. “A Primer on instantons in QCD”. 2000. arXiv: hep-ph/0009136.

[52] J. Binney et al. *The Theory of Critical Phenomena: An Introduction to the Renormalization Group*. Oxford Science Publ. Clarendon Press, 1992.

[53] L.G. Yaffe and B. Svetitsky. “First Order Phase Transition in the SU(3) Gauge Theory at Finite Temperature”. In: *Phys. Rev. D* 26 (1982), 963. DOI: 10.1103/PhysRevD.26.963.

[54] Robert D. Pisarski and Frank Wilczek. “Remarks on the Chiral Phase Transition in Chromodynamics”. In: *Phys. Rev. D* 29 (1984), 338–341. DOI: 10.1103/PhysRevD.29.338.

[55] S. Ejiri et al. “On the magnetic equation of state in (2+1)-flavor QCD”. In: *Phys. Rev. D* 80 (2009), 094505. DOI: 10.1103/PhysRevD.80.094505.

[56] C. Bernard et al. “QCD thermodynamics with three flavors of improved staggered quarks”. In: *Phys. Rev. D* 71 (2005), 034504. DOI: 10.1103/PhysRevD.71.034504.

[57] M. Cheng et al. “The Transition temperature in QCD”. In: *Phys. Rev. D* 74 (2006), 054507. DOI: 10.1103/PhysRevD.74.054507.

[58] Y. Aoki et al. “The Order of the quantum chromodynamics transition predicted by the standard model of particle physics”. In: *Nature* 443 (2006), 675–678. DOI: 10.1038/nature05120.

[59] A. Bazavov et al. “The chiral and deconfinement aspects of the QCD transition”. In: *Phys. Rev. D* 85 (2012), 054503. DOI: 10.1103/PhysRevD.85.054503.

[60] Tanmoy Bhattacharya et al. “QCD Phase Transition with Chiral Quarks and Physical Quark Masses”. In: *Phys. Rev. Lett.* 113.8 (2014), 082001. DOI: 10.1103/PhysRevLett.113.082001.

[61] A. Bazavov et al. “Equation of state in (2+1)-flavor QCD”. In: *Phys. Rev. D* 90 (2014), 094503. DOI: 10.1103/PhysRevD.90.094503.

[62] A. Bazavov et al. “Chiral crossover in QCD at zero and non-zero chemical potentials”. In: *Phys. Lett. B* 795 (2019), 15–21. DOI: 10.1016/j.physletb.2019.05.013.

[63] H.T. Ding et al. “Chiral Phase Transition Temperature in (2+1)-Flavor QCD”. In: *Phys. Rev. Lett.* 123.6 (2019), 062002. DOI: 10.1103/PhysRevLett.123.062002.

[64] Agostino Butti, Andrea Pelissetto, and Ettore Vicari. “On the nature of the finite temperature transition in QCD”. In: *JHEP* 08 (2003), 029. DOI: 10.1088/1126-6708/2003/08/029.

[65] Andrea Pelissetto and Ettore Vicari. “Relevance of the axial anomaly at the finite-temperature chiral transition in QCD”. In: *Phys. Rev. D* 88.10 (2013), 105018. DOI: 10.1103/PhysRevD.88.105018.

[66] Mara Grahl and Dirk H. Rischke. “Functional renormalization group study of the two-flavor linear sigma model in the presence of the axial anomaly”. In: *Phys. Rev. D* 88.5 (2013), 056014. DOI: 10.1103/PhysRevD.88.056014.

[67] Yu Nakayama and Tomoki Ohtsuki. “Bootstrapping phase transitions in QCD and frustrated spin systems”. In: *Phys. Rev. D* 91.2 (2015), 021901. DOI: 10.1103/PhysRevD.91.021901.

[68] Viktor Dick. "The chiral anomaly of Quantum Chromodynamics at high temperatures: lattice investigation of the overlap Dirac spectrum". PhD thesis. Bielefeld University, 2016.

[69] Edward V. Shuryak. "Which chiral symmetry is restored in hot QCD?" In: *Comments Nucl. Part. Phys.* 21.4 (1994), 235–248. arXiv: hep-ph/9310253.

[70] Tetsuo Hatsuda, Su Houng Lee, and Hiroyuki Shiomi. "QCD sum rules, scattering length and the vector mesons in nuclear medium". In: *Phys. Rev.* C52 (1995), 3364–3372. DOI: 10.1103/PhysRevC.52.3364.

[71] Sinya Aoki, Hidenori Fukaya, and Yusuke Taniguchi. "Chiral symmetry restoration, eigenvalue density of Dirac operator and axial U(1) anomaly at finite temperature". In: *Phys. Rev. D* 86 (2012), 114512. DOI: 10.1103/PhysRevD.86.114512.

[72] Viktor Dick et al. "Microscopic origin of $U_A(1)$ symmetry violation in the high temperature phase of QCD". In: *Phys. Rev. D* 91.9 (2015), 094504. DOI: 10.1103/PhysRevD.91.094504.

[73] M. Shifman. *Instantons in Gauge Theories*. World Scientific, 1994.

[74] M. Srednicki. *Quantum field theory*. Cambridge University Press, 2007.

[75] R. Rajaraman. *Solitons and Instantons: An Introduction to Solitons and Instantons in Quantum Field Theory*. North-Holland personal library. North-Holland Publishing Company, 1982.

[76] David Clarke. "Scale Setting and Topological Observables in Pure SU(2) LGT". Other thesis. 2018. arXiv: 1901.03200 [hep-lat].

[77] Peter Thomas Jahn. "The Topological Susceptibility of QCD at High Temperatures". PhD thesis. Darmstadt, Tech. Hochsch., 2019. DOI: 10.25534/tuprints-00009485.

[78] W. Dittrich and M. Reuter. *Selected Topics in Gauge Theories*. Lecture notes in physics. Springer-Verlag, 1986.

[79] R. Bott. "An Application of Morse theory to the topology of Lie groups". In: *Bull. Soc. Math. Fr.* 84 (1956), 251–281.

[80] Luis Alvarez-Gaume and Miguel A. Vazquez-Mozo. *An invitation to quantum field theory*. Vol. 839. Lecture Notes in Physics. Springer Berlin Heidelberg, 2012. DOI: 10.1007/978-3-642-23728-7.

[81] Sidney Coleman. *Aspects of Symmetry: Selected Erice Lectures*. Cambridge University Press, 1985. DOI: 10.1017/CBO9780511565045.

[82] Steven Weinberg. *The quantum theory of fields. Vol. 2: Modern applications*. Cambridge University Press, 2013.

[83] A.A. Belavin et al. "Pseudoparticle Solutions of the Yang-Mills Equations". In: *Phys. Lett. B* 59 (1975). Ed. by J.C. Taylor, 85–87. DOI: 10.1016/0370-2693(75)90163-X.

[84] Gerard 't Hooft. "Symmetry Breaking Through Bell-Jackiw Anomalies". In: *Phys. Rev. Lett.* 37 (1976). Ed. by Mikhail A. Shifman, 8–11. DOI: 10.1103/PhysRevLett.37.8.

[85] Curtis G. Callan Jr., Roger F. Dashen, and David J. Gross. "Toward a Theory of the Strong Interactions". In: *Phys. Rev. D* 17 (1978). Ed. by Mikhail A. Shifman, 2717. DOI: 10.1103/PhysRevD.17.2717.

[86] M.F. Atiyah and I.M. Singer. "The Index of elliptic operators. 1". In: *Annals Math.* 87 (1968), 484–530. DOI: 10.2307/1970715.

[87] Peter Petreczky, Hans-Peter Schadler, and Sayantan Sharma. "The topological susceptibility in finite temperature QCD and axion cosmology". In: *Phys. Lett. B* 762 (2016), 498–505. DOI: 10.1016/j.physletb.2016.09.063.

[88] Gerard 't Hooft. "Computation of the Quantum Effects Due to a Four-Dimensional Pseudoparticle". In: *Phys. Rev. D* 14 (1976). Ed. by Mikhail A. Shifman. [Erratum: Phys.Rev.D 18, 2199 (1978)], 3432–3450. DOI: 10.1103/PhysRevD.14.3432.

[89] Mikko Laine and Aleksi Vuorinen. *Basics of Thermal Field Theory*. Vol. 925. Springer, 2016. DOI: 10.1007/978-3-319-31933-9.

[90] M. Asakawa, T. Hatsuda, and Y. Nakahara. "Maximum entropy analysis of the spectral functions in lattice QCD". In: *Prog. Part. Nucl. Phys.* 46 (2001), 459–508. DOI: 10.1016/S0146-6410(01)00150-8.

[91] Sebastian Fuchs, Thomas Pruschke, and Mark Jarrell. "Analytic continuation of quantum Monte Carlo data by stochastic analytical inference". In: *Physical Review E* 81.5 (2010). DOI: 10.1103/physreve.81.056701.

[92] Hauke Sören Sandmeyer. "Hadronic correlators from heavy to very light quarks: Spectral and transport properties from lattice QCD". PhD thesis. Bielefeld University, 2019. DOI: 10.4119/unibi/2936264.

[93] G.F. Giudice and Mikhail E. Shaposhnikov. "Strong sphalerons and electroweak baryogenesis". In: *Phys. Lett. B* 326 (1994), 118–124. DOI: 10.1016/0370-2693(94)91202-5.

[94] Dmitri Kharzeev, Edward Shuryak, and Ismail Zahed. "Sphalerons, baryogenesis, and helical magnetogenesis in the electroweak transition of the minimal standard model". In: *Phys. Rev. D* 102.7 (2020), 073003. DOI: 10.1103/PhysRevD.102.073003.

[95] Dietrich Bodeker. "On the effective dynamics of soft nonAbelian gauge fields at finite temperature". In: *Phys. Lett. B* 426 (1998), 351–360. DOI: 10.1016/S0370-2693(98)00279-2.

[96] Dietrich Bodeker. "Diagrammatic approach to soft nonAbelian dynamics at high temperature". In: *Nucl. Phys. B* 566 (2000), 402–422. DOI: 10.1016/S0550-3213(99)00582-9.

[97] Dietrich Bodeker. "From hard thermal loops to Langevin dynamics". In: *Nucl. Phys. B* 559 (1999), 502–538. DOI: 10.1016/S0550-3213(99)00435-6.

[98] Peter Brockway Arnold and Laurence G. Yaffe. "High temperature color conductivity at next-to-leading log order". In: *Phys. Rev. D* 62 (2000), 125014. DOI: 10.1103/PhysRevD.62.125014.

[99] Guy D. Moore. "The Sphaleron rate: Bodeker's leading log". In: *Nucl. Phys. B* 568 (2000), 367–404. DOI: 10.1016/S0550-3213(99)00746-4.

[100] D. Bodeker, Guy D. Moore, and K. Rummukainen. "Chern-Simons number diffusion and hard thermal loops on the lattice". In: *Phys. Rev. D* 61 (2000), 056003. DOI: 10.1103/PhysRevD.61.056003.

[101] Larry D. McLerran, Emil Mottola, and Mikhail E. Shaposhnikov. "Sphalerons and Axion Dynamics in High Temperature QCD". In: *Phys. Rev. D* 43 (1991), 2027–2035. DOI: 10.1103/PhysRevD.43.2027.

[102] Sergei A. Voloshin. "Testing the Chiral Magnetic Effect with Central U+U collisions". In: *Phys. Rev. Lett.* 105 (2010), 172301. DOI: 10.1103/PhysRevLett.105.172301.

[103] P. Weisz. "Continuum Limit Improved Lattice Action for Pure Yang-Mills Theory. 1." In: *Nucl. Phys. B* 212 (1983), 1–17. DOI: 10.1016/0550-3213(83)90595-3.

[104] P. Weisz and R. Wohlert. "Continuum Limit Improved Lattice Action for Pure Yang-Mills Theory. 2." In: *Nucl. Phys. B* 236 (1984). [Erratum: Nucl.Phys.B 247, 544 (1984)], 397. DOI: 10.1016/0550-3213(84)90563-7.

[105] S. Duane et al. "Hybrid Monte Carlo". In: *Phys. Lett. B* 195 (1987), 216–222. DOI: 10.1016/0370-2693(87)91197-X.

[106] Michael Creutz. "Overrelaxation and Monte Carlo Simulation". In: *Phys. Rev. D* 36 (1987), 515. DOI: 10.1103/PhysRevD.36.515.

[107] Stephen L. Adler. "An Overrelaxation Method for the Monte Carlo Evaluation of the Partition Function for Multiquadratic Actions". In: *Phys. Rev. D* 23 (1981), 2901. DOI: 10.1103/PhysRevD.23.2901.

[108] N. Cabibbo and E. Marinari. "A New Method for Updating SU(N) Matrices in Computer Simulations of Gauge Theories". In: *Phys. Lett. B* 119 (1982), 387–390. DOI: 10.1016/0370-2693(82)90696-7.

[109] A.D. Kennedy and B.J. Pendleton. "Improved Heat Bath Method for Monte Carlo Calculations in Lattice Gauge Theories". In: *Phys. Lett. B* 156 (1985), 393–399. DOI: 10.1016/0370-2693(85)91632-6.

[110] Kenneth G. Wilson. "Confinement of quarks". In: *Phys. Rev. D* 10 (8 1974), 2445–2459. DOI: 10.1103/PhysRevD.10.2445.

[111] Thomas DeGrand and Carleton E. Detar. *Lattice methods for quantum chromodynamics*. World Scientific, 2006.

[112] Andreas S. Kronfeld. "Lattice Gauge Theory with Staggered Fermions: How, Where, and Why (Not)". In: *PoS* LATTICE2007 (2007). Ed. by Gunnar Bali et al., 016. DOI: 10.22323/1.042.0016.

[113] Holger Bech Nielsen and M. Ninomiya. "No Go Theorem for Regularizing Chiral Fermions". In: *Phys. Lett. B* 105 (1981), 219–223. DOI: 10.1016/0370-2693(81)91026-1.

[114] Martin Luscher. "Exact chiral symmetry on the lattice and the Ginsparg-Wilson relation". In: *Phys. Lett. B* 428 (1998), 342–345. DOI: 10.1016/S0370-2693(98)00423-7.

[115] Paul H. Ginsparg and Kenneth G. Wilson. "A Remnant of Chiral Symmetry on the Lattice". In: *Phys. Rev. D* 25 (1982), 2649. DOI: 10.1103/PhysRevD.25.2649.

[116] Pilar Hernandez, Karl Jansen, and Martin Luscher. "Locality properties of Neuberger's lattice Dirac operator". In: *Nucl. Phys. B* 552 (1999), 363–378. DOI: 10.1016/S0550-3213(99)00213-8.

[117] J. van den Eshof et al. "Numerical methods for the QCD overlap operator. I. Sign function and error bounds". In: *Comput. Phys. Commun.* 146 (2002), 203–224. DOI: 10.1016/S0010-4655(02)00455-1.

[118] P.P. Petrushev and V.A. Popov. *Rational Approximation of Real Functions*. Vol. 28. Encyclopedia of Mathematics and Its Applications. Cambridge University Press, 1987.

[119] Beat Jegerlehner. "Krylov space solvers for shifted linear systems". 1996. arXiv: hep-lat/9612014.

[120] Thomas Kalkreuter and Hubert Simma. "An Accelerated conjugate gradient algorithm to compute low lying eigenvalues: A Study for the Dirac operator in SU(2) lattice QCD". In: *Comput. Phys. Commun.* 93 (1996), 33–47. DOI: 10.1016/0010-4655(95)00126-3.

[121] M.A. Clark et al. "Solving Lattice QCD systems of equations using mixed precision solvers on GPUs". In: *Comput. Phys. Commun.* 181 (2010), 1517–1528. DOI: 10.1016/j.cpc.2010.05.002.

[122] Peter Thomas Jahn, Guy D. Moore, and Daniel Robaina. "Estimating χ_{top} lattice artifacts from flowed SU(2) calorons". In: *Eur. Phys. J. C* 79.6 (2019), 510. DOI: 10.1140/epjc/s10052-019-7008-9.

[123] Guy D. Moore. "Improved Hamiltonian for Minkowski Yang-Mills theory". In: *Nucl. Phys. B* 480 (1996), 689–728. DOI: 10.1016/S0550-3213(96)00497-X.

[124] Sundance O. Bilson-Thompson, Derek B. Leinweber, and Anthony G. Williams. "Highly improved lattice field strength tensor". In: *Annals Phys.* 304 (2003), 1–21. DOI: 10.1016/S0003-4916(03)00009-5.

[125] Adrien Florio. "Aspects of Gauge Theories Vacua". PhD thesis. Ecole Polytechnique, Lausanne, 2020. DOI: 10.5075/epfl-thesis-8139.

[126] Martin Luscher and Stefan Schaefer. "Lattice QCD without topology barriers". In: *JHEP* 07 (2011), 036. DOI: 10.1007/JHEP07(2011)036.

[127] Adrien Florio, Olaf Kaczmarek, and Lukas Mazur. "Open-Boundary Conditions in the Deconfined Phase". In: *Eur. Phys. J. C* 79.12 (2019), 1039. DOI: 10.1140/epjc/s10052-019-7564-z.

[128] Peter Thomas Jahn, Guy D. Moore, and Daniel Robaina. "$\chi_{\text{top}}(T \gg T_{\text{c}})$ in pure-glue QCD through reweighting". In: *Phys. Rev. D* 98.5 (2018), 054512. DOI: 10.1103/PhysRevD.98.054512.

[129] P. Thomas Jahn, Guy D. Moore, and Daniel Robaina. "Improved Reweighting for QCD Topology at High Temperature". 2020. arXiv: 2002.01153 [hep-lat].

[130] B. Berg. "Dislocations and Topological Background in the Lattice O(3) σ Model". In: *Phys. Lett. B* 104 (1981), 475–480. DOI: 10.1016/0370-2693(81)90518-9.

[131] Y. Iwasaki and T. Yoshiè. "Instantons and topological charge in lattice gauge theory". In: *Physics Letters B* 131.1 (1983), 159–164. DOI: https://doi.org/10.1016/0370-2693(83)91111-5.

[132] M. Albanese et al. "Glueball Masses and String Tension in Lattice QCD". In: *Phys. Lett. B* 192 (1987), 163–169. DOI: 10.1016/0370-2693(87)91160-9.

[133] Colin Morningstar and Mike J. Peardon. "Analytic smearing of SU(3) link variables in lattice QCD". In: *Phys. Rev. D* 69 (2004), 054501. DOI: 10.1103/PhysRevD.69.054501.

[134] Anna Hasenfratz and Francesco Knechtli. "Flavor symmetry and the static potential with hypercubic blocking". In: *Phys. Rev. D* 64 (2001), 034504. DOI: 10.1103/PhysRevD.64.034504.

[135] Masakiyo Kitazawa et al. "Equation of State for SU(3) Gauge Theory via the Energy-Momentum Tensor under Gradient Flow". In: *Phys. Rev. D* 94.11 (2016), 114512. DOI: 10.1103/PhysRevD.94.114512.

[136] Masakiyo Kitazawa et al. “Correlations of the energy-momentum tensor via gradient flow in SU(3) Yang-Mills theory at finite temperature”. In: *Phys. Rev. D* 96.11 (2017), 111502. DOI: 10.1103/PhysRevD.96.111502.

[137] Szabolcs Borsanyi et al. “High-precision scale setting in lattice QCD”. In: *JHEP* 09 (2012), 010. DOI: 10.1007/JHEP09(2012)010.

[138] A. Bazavov et al. “Gradient flow and scale setting on MILC HISQ ensembles”. In: *Phys. Rev.* D93.9 (2016), 094510. DOI: 10.1103/PhysRevD.93.094510.

[139] Mattia Dalla Brida and Alberto Ramos. “The gradient flow coupling at high-energy and the scale of SU(3) YangMills theory”. In: *Eur. Phys. J.* C79.8 (2019), 720. DOI: 10.1140/epjc/s10052-019-7228-z.

[140] Marco Cè et al. “Non-Gaussianities in the topological charge distribution of the SU(3) Yang–Mills theory”. In: *Phys. Rev. D* 92.7 (2015), 074502. DOI: 10.1103/PhysRevD.92.074502.

[141] Yusuke Taniguchi et al. “Topological susceptibility in finite temperature (2+1)-flavor QCD using gradient flow”. In: *Phys. Rev. D* 95.5 (2017), 054502. DOI: 10.1103/PhysRevD.95.054502.

[142] Lukas Mazur et al. “Euclidean correlation functions of the topological charge density”. In: *PoS* LATTICE2019 (2020), 219. DOI: 10.22323/1.363.0219.

[143] Martin Luscher. “Topology, the Wilson flow and the HMC algorithm”. In: *PoS* LATTICE2010 (2010). Ed. by Giancarlo Rossi, 015. DOI: 10.22323/1.105.0015.

[144] Simon Wolfgang Mages. “Properties of the Quark Gluon Plasma from Lattice QCD”. PhD thesis. 2015. DOI: 10.5283/epub.31409.

[145] A. Ramos and S. Sint. “Symanzik improvement of the gradient flow in lattice gauge theories”. In: *Eur. Phys. J. C* 76.1 (2016), 15. DOI: 10.1140/epjc/s10052-015-3831-9.

[146] Hans Munthe-Kaas. “Runge-Kutta methods on Lie groups”. In: *BIT Numerical Mathematics* 38.1 (1998), 92–111. DOI: https://doi.org/10.1007/BF02510919.

[147] Elena Celledoni, Arne Marthinsen, and Brynjulf Owren. “Commutator-free Lie group methods”. In: *Future Generation Computer Systems* 19.3 (2003). Special Issue on Geometric Numerical Algorithms, 341–352. DOI: https://doi.org/10.1016/S0167-739X(02)00161-9.

[148] Marcel Rodekamp. “Multi-GPU Implementation of the Wilson Flow in Lattice QCD”. Bachelor’s thesis. Bielefeld University, 2018.

[149] Patrick Fritzsch and Alberto Ramos. “The gradient flow coupling in the Schrödinger Functional”. In: *JHEP* 10 (2013), 008. DOI: 10.1007/JHEP10(2013)008.

[150] M. A. Clark and A. D. Kennedy. “The RHMC algorithm for two flavors of dynamical staggered fermions”. In: *Nucl. Phys. B Proc. Suppl.* 129 (2004). Ed. by S. Aoki et al., 850–852. DOI: 10.1016/S0920-5632(03)02732-4.

[151] Constantia Alexandrou et al. “Comparison of topological charge definitions in Lattice QCD”. In: *Eur. Phys. J. C* 80.5 (2020), 424. DOI: 10.1140/epjc/s10052-020-7984-9.

[152] Claudio Bonati et al. “Recent progress on QCD inputs for axion phenomenology”. In: *EPJ Web Conf.* 137 (2017). Ed. by Y. Foka, N. Brambilla, and V. Kovalenko, 08004. DOI: 10.1051/epjconf/201713708004.

[153] Claudio Bonati et al. “Axion phenomenology and θ-dependence from $N_f = 2+1$ lattice QCD”. In: *JHEP* 03 (2016), 155. DOI: 10.1007/JHEP03(2016)155.

[154] Olaf Kaczmarek, Lukas Mazur, and Sayantan Sharma. “Eigenvalue spectra of QCD and the fate of $U_A(1)$ breaking towards the chiral limit”. 2021. arXiv: 2102.06136 [hep-lat].

[155] Shailesh Chandrasekharan and Norman H. Christ. “Dirac spectrum, axial anomaly and the QCD chiral phase transition”. In: *Nucl. Phys. B Proc. Suppl.* 47 (1996). Ed. by T.D. Kieu, B.H.J. McKellar, and A.J. Guttmann, 527–534. DOI: 10.1016/0920-5632(96)00115-6.

[156] Sayantan Sharma. “The fate of $U_A(1)$ and topological features of QCD at finite temperature”. In: *11th International Workshop on Critical Point and Onset of Deconfinement*. 2018. arXiv: 1801.08500 [hep-lat].

[157] Rajamani Narayanan and Herbert Neuberger. “Chiral fermions on the lattice”. In: *Phys. Rev. Lett.* 71.20 (1993), 3251. DOI: 10.1103/PhysRevLett.71.3251.

[158] Peter Hasenfratz, Victor Laliena, and Ferenc Niedermayer. “The Index theorem in QCD with a finite cutoff”. In: *Phys. Lett. B* 427 (1998), 125–131. DOI: 10.1016/S0370-2693(98)00315-3.

[159] A. Li et al. “Overlap Valence on 2+1 Flavor Domain Wall Fermion Configurations with Deflation and Low-mode Substitution”. In: *Phys. Rev. D* 82 (2010), 114501. DOI: 10.1103/PhysRevD.82.114501.

[160] M. Lujan et al. “The Δ_{mix} parameter in the overlap on domain-wall mixed action”. In: *Phys. Rev. D* 86 (2012), 014501. DOI: 10.1103/PhysRevD.86.014501.

[161] Takuya Kanazawa and Naoki Yamamoto. “Quasi-instantons in QCD with chiral symmetry restoration”. In: *Phys. Rev. D* 91 (2015), 105015. DOI: 10.1103/PhysRevD.91.105015.

[162] Robert G. Edwards et al. “Chiral condensate in the deconfined phase of quenched gauge theories”. In: *Phys. Rev. D* 61 (2000), 074504. DOI: 10.1103/PhysRevD.61.074504.

[163] H. -T. Ding et al. “Correlated Dirac Eigenvalues and Axial Anomaly in Chiral Symmetric QCD”. In: *Phys. Rev. Lett.* 126.8 (2021), 082001. DOI: 10.1103/PhysRevLett.126.082001.

[164] Matteo Giordano. “Localized Dirac eigenmodes, chiral symmetry breaking, and Goldstone’s theorem”. 2020. arXiv: 2009.00486 [hep-th].

[165] Andrei V. Smilga and J. Stern. “On the spectral density of Euclidean Dirac operator in QCD”. In: *Phys. Lett. B* 318 (1993), 531–536. DOI: 10.1016/0370-2693(93)91551-W.

[166] Jacobus J. M. Verbaarschot. “Spectrum of the Dirac operator in a QCD instanton liquid: Two versus three colors”. In: *Nucl. Phys. B* 427 (1994), 534–544. DOI: 10.1016/0550-3213(94)90638-6.

[167] Su H. Lee and T. Hatsuda. “U-a(1) symmetry restoration in QCD with N(f) flavors”. In: *Phys. Rev. D* 54 (1996), R1871–R1873. DOI: 10.1103/PhysRevD.54.R1871.

[168] Michael C. Birse, Thomas D. Cohen, and Judith A. McGovern. “U(1)-A symmetry and correlation functions in the high temperature phase of QCD”. In: *Phys. Lett. B* 388 (1996), 137–140. DOI: 10.1016/0370-2693(96)01151-3.

[169] Luis Altenkort et al. "Heavy quark momentum diffusion from the lattice using gradient flow". In: *Phys. Rev. D* 103.1 (2021), 014511. DOI: 10.1103/PhysRevD.103.014511.

[170] Luis Altenkort et al. "The sphaleron rate from Euclidean lattice correlators: an exploration". 2020. arXiv: 2012.08279 [hep-lat].

[171] A. Francis et al. "Critical point and scale setting in SU(3) plasma: An update". In: *Phys. Rev. D* 91.9 (2015), 096002. DOI: 10.1103/PhysRevD.91.096002.

[172] R. Sommer. "A New way to set the energy scale in lattice gauge theories and its applications to the static force and alpha-s in SU(2) Yang-Mills theory". In: *Nucl. Phys. B* 411 (1994), 839–854. DOI: 10.1016/0550-3213(94)90473-1.

[173] Y. Burnier et al. "Thermal quarkonium physics in the pseudoscalar channel". In: *JHEP* 11 (2017), 206. DOI: 10.1007/JHEP11(2017)206.

[174] Alexander M. Eller and Guy D. Moore. "Gradient-flowed thermal correlators: how much flow is too much?" In: *Phys. Rev.* D97.11 (2018), 114507. DOI: 10.1103/PhysRevD.97.114507.

[175] Peter Brockway Arnold and Larry D. McLerran. "The Sphaleron Strikes Back". In: *Phys. Rev. D* 37 (1988), 1020. DOI: 10.1103/PhysRevD.37.1020.

[176] Hiroshi Suzuki. "Energymomentum tensor from the YangMills gradient flow". In: *PTEP* 2013 (2013). [Erratum: PTEP 2015, 079201 (2015)], 083B03. DOI: 10.1093/ptep/ptt059.

[177] Hiroshi Suzuki and Hiromasa Takaura. "$t \to 0$ extrapolation function in SFtX method for the energy-momentum tensor". 2021. arXiv: 2102.02174 [hep-lat].

[178] Erhard Seiler. "Some more remarks on the Witten-Veneziano formula for the eta-prime mass". In: *Phys. Lett. B* 525 (2002), 355–359. DOI: 10.1016/S0370-2693(01)01469-1.

[179] Ettore Vicari. "The Euclidean two point correlation function of the topological charge density". In: *Nucl. Phys. B* 554 (1999), 301–312. DOI: 10.1016/S0550-3213(99)00297-7.

[180] I. Horvath et al. "The Negativity of the overlap-based topological charge density correlator in pure-glue QCD and the non-integrable nature of its contact part". In: *Phys. Lett. B* 617 (2005), 49–59. DOI: 10.1016/j.physletb.2005.04.076.

[181] M. Laine, A. Vuorinen, and Y. Zhu. "Next-to-leading order thermal spectral functions in the perturbative domain". In: *JHEP* 09 (2011), 084. DOI: 10.1007/JHEP09(2011)084.

[182] Peter Arnold, Dam Son, and Laurence G. Yaffe. "The hot baryon violation rate is $O(\alpha_w^5 T^4)$". In: *Phys. Rev. D* 55 (10 1997), 6264–6273. DOI: 10.1103/PhysRevD.55.6264.

[183] S. Aoki et al. "Study of axial U(1) anomaly at high temperature with lattice chiral fermions". 2020. arXiv: 2011.01499 [hep-lat].

Declaration of Authorship

I hereby declare that the work done in this thesis is that of the author alone with the help of no more than the mentioned literature and auxiliary means.

Bielefeld, April 22, 2021

Lukas Mazur

Acknowledgements

First of all, I would like to thank my supervisor Olaf Kaczmarek who motivated me for high energy physics and high performance computing. He always had time to help me and talk about my projects in physics. He also made it possible for me to travel around the world to attend summer schools, conferences, and workshops. I had the opportunity to see behind the scenes of scientific work, which was a truly life-changing experience, and for that I am very grateful. I would also like to thank Fritjhof Karsch for agreeing to be the second referee for this dissertation and for his support during my stay at Brookhaven National Laboratory.

Thanks to my office colleagues, Hauke Sandmeyer, David Clarke and Anna-Lena Lorenz, for the time we spent together. I really enjoyed the countless inspiring discussions with Hauke about code design and programming in general. Also special thanks to David. Besides the very helpful physics discussions, it is always very entertaining to talk with him about the important things in life, such as questionable German grammar rules or why pineapples do not belong on pizza. I would further like to thank Jishnu Goswami, who was always willing to answer my physics questions. Plus, his amazingly dry sense of humor is truly unique. Another thank you goes to Adrien Florio for the great collaboration and for inviting me to Switzerland. I had a really great time there. I then want to thank Sayantan Sharma. His experience and knowledge was very helpful in our project. I also want to thank the ParallelGPUCode programming crew for their great contributions, i.e. Hauke, David, Dennis Bollweg, Luis Altenkort, Hai-Tao Shu, Phillip Scior, Hai-Tao Shu, Rasmus Larsen and Marcel Rodekamp. A big thank you goes to my colleagues who are (or were) part of the Bielefeld high energy physics group, i.e. Giuseppe Gagliardi, Lorenzo Dini, Guido Nicotra, Anirban Alahiri, Patrick Steinbrecher, Marc Klegrewe, Marc Borell, Marius Neumann, Felix Ziesche, Simon Dentiger.

At this point I also want to mention (again) the people who have helped proofread this dissertation, i.e. Olaf, David, Jishnu, Sayantan and Luis.

Last but not least, I am extremely grateful to my mother Mariola and my siblings Sabine, Andreas and Eva for always believing in me. Without their support I would not have made it this far.