TECHNISCHE
UNIVERSITÄT
DARMSTADT

From Spectral Reconstruction to Chaos in Quantum Chromodynamics

---

vom Fachbereich Physik
der Technischen Universität Darmstadt

zur Erlangung des akademischen Grades
Doctor rerum naturalium
(Dr. rer. nat.)

genehmigte Dissertation von
Jonas Turnwald

Referent: Prof. Ph.D. Guy D. Moore
Korreferent: Prof. Dr. Jens Braun

Darmstadt 2025

From Spectral Reconstruction to Chaos in Quantum Chromodynamics
Von Spektralrekonstruktion zu Chaos in der Quantenchromodynamik

Dissertation von Jonas Turnwald

Tag der mündlichen Prüfung: 14.07.2025
Jahr der Veröffentlichung auf TUprints: 2025

Darmstadt, Technische Universität Darmstadt

Bitte zitieren Sie dieses Dokument als:
URN: urn:nbn:de:tuda-tuprints-310211
URI: `https://tuprints.ulb.tu-darmstadt.de/id/eprint/310211`

ABSTRACT

Understanding the dynamic properties of quantum chromodynamics (QCD) and other strongly coupled field theories has remained a challenging task. Yet, accurate knowledge of their real-time behavior is crucial for describing the evolution of heavy-ion collisions, particularly the strongly coupled quark-gluon plasma phase. To this end, we develop a method based on Gaussian processes in order to reconstruct spectral functions from Euclidean correlation functions. This approach enables the extraction of real-time information from non-perturbative calculations, either on the lattice or within functional approaches. We introduce several extensions of this framework that enable the incorporation of additional physical input into the reconstruction of the spectral function, such as normalization constraints, asymptotic behavior, and higher-dimensional data. These extensions lead to a versatile method that can be applied to a wide range of different problems, and we have therefore implemented it in a Python package called *fredipy*.

We present multiple applications of this method to QCD problems. First, we reconstruct the spectral functions of four-point gluon correlation functions in pure Yang-Mills theory, allowing us to extract the glueball masses of the scalar and pseudoscalar glueball states, which are in very good agreement with results from lattice and other functional methods. In a second application of the method, we reconstruct the strong coupling in QCD and compare it to a direct calculation using previously reconstructed ghost and gluon propagators. This is useful for calculating S-matrix elements of non-perturbative QCD processes, and can be used in further calculations of spectral functions or bound states. Finally, as a third application, we reconstruct the thermal photon rate in QCD from lattice data. We compare different reconstruction methods and find excellent agreement between them. Together, these applications demonstrate that the Gaussian process reconstruction method effectively captures the key features of spectral functions while offering significant flexibility to incorporate a wide range of physical prior information.

Additionally, we investigate out-of-time ordered correlators in $\mathrm{O}(N)$ models. These correlators quantify the scrambling of quantum information, and hence we expect that thermalization time scales are related to their exponential growth. As a first exploration, we study their behavior using classical statistical lattice simulations around the phase transition and diagrammatic weak coupling methods, with a special focus on the spatial spreading of information. We can relate the resulting Lyaponov exponents to more familiar quantities in thermal field theories.

ZUSAMMENFASSUNG

Die genaue Bestimmung dynamischer Eigenschaften der Quantenchromodynamik und anderer stark gekoppelter Feldtheorien ist nach wie vor eine Herausforderung für theoretische Methoden. Die genaue Kenntnis ihres dynamischen Verhaltens ist jedoch wichtig für die Beschreibung der zeitlichen Entwicklung von Schwerionenkollisionen, insbesondere während der stark gekoppelten Quark-Gluon-Plasmaphase. In dieser Arbeit entwickeln wir eine auf Gaußschen Prozessen basierende Methode zur Rekonstruktion von Spektralfunktionen aus euklidischen Korrelationsfunktionen. Dieser Ansatz ermöglicht die Berechnung dynamischer Observablen aus euklidischen Daten von nicht-perturbativen Gittersimulationen oder funktionaler Methoden. Wir stellen verschiedene Erweiterungen dieser Methode vor, die es ermöglichen, physikalische Eigenschaften direkt in die Rekonstruktion einzubinden, wie z. B. Normalisierung, asymptotisches Verhalten und Daten in mehreren Dimensionen. Diese Erweiterungen führen zu einer vielseitigen Methode, die auf ein breites Spektrum physikalischer Problemstellungen angewendet werden kann. Daher haben wir sie in einem Python-Paket namens *fredipy* implementiert.

Wir präsentieren mehrere Anwendungen dieser Methode auf QCD-Probleme. Als erste Anwendung rekonstruieren wir die Spektralfunktionen von Vierpunkt-Gluon-Korrelationsfunktionen in reiner Yang-Mills-Theorie. Dadurch können wir die Glueballmassen der skalaren und pseudoskalaren Glueballzustände extrahieren, die sehr gut mit Ergebnissen aus Gitter- und anderen funktionalen Methoden übereinstimmen. In einer zweiten Anwendung der Methode rekonstruieren wir die starke Kopplung der QCD und vergleichen sie mit einer direkten Berechnung unter Verwendung zuvor rekonstruierter Ghost- und Gluonpropagatoren. Dies ist nützlich für die Berechnung von S-Matrix-Elementen nicht-perturbativer QCD-Prozesse oder kann für weitere Berechnungen von Spektralfunktionen oder gebundenen Zuständen verwendet werden. Schließlich rekonstruieren wir in einer dritten Anwendung die thermische Photonenrate in der QCD aus Gitterdaten. Wir vergleichen verschiedene Rekonstruktionsmethoden und stellen eine hervorragende Übereinstimmung fest. Zusammen zeigen diese Anwendungen, dass die Gaußschen Prozess Rekonstruktionsmethode die wichtigsten Merkmale von Spektralfunktionen effektiv erfasst und gleichzeitig eine hohe Flexibilität bietet, um ein breites Spektrum physikalischer Vorinformationen zu berücksichtigen.

Zusätzlich untersuchen wir zeitlich nicht geordnete Korrelatoren in O(N)-Modellen. Diese Korrelatoren quantifizieren die Ausbreitung von Quanteninformation. Daher erwarten wir, dass die Thermalisierungszeitskalen mit ihrem exponentiellen Wachstum zusammenhängen. Zunächst untersuchen wir ihr Verhalten mithilfe klassischer statistischer Gittersimulationen im Umfeld des Phasenübergangs sowie diagrammatischer Methoden bei schwacher Kopplung. Die daraus resultierenden Lyapunov-Exponenten lassen sich mit bekannten Größen aus der thermischen Feldtheorie in Beziehung setzen.

CONTENTS

1. INTRODUCTION

Relativistic heavy-ion collisions allow the study of matter under extreme conditions in the laboratory, most importantly at high temperature and density. Such conditions have existed only in the early universe, a few microseconds after the Big Bang [4]. During this period, quarks and gluons constituted the fundamental degrees of freedom, forming a *quark-gluon plasma* (QGP) [5]. After cooling down, the QGP undergoes a phase change at around 155 MeV [6–8] to a hadronic phase, where quarks and gluons are confined into hadrons, such as protons and neutrons, which are the building blocks of atomic nuclei.

The existence of the QGP was predicted in the mid-1970s [9–11] and was first experimentally confirmed by four collaborations at the Relativistic Heavy Ion Collider (RHIC) [12–15]. The phase diagram of the underlying theory, quantum chromodynamics (QCD), in the plane of temperature and baryon chemical potential, has been studied extensively in recent years, both theoretically and experimentally. In this phase diagram of QCD, multiple interesting features are predicted to exist, such as the critical end point, where the crossover ends in a first-order phase transition line [16]. This critical point is expected to be located at a high baryon chemical potential of $600-700$ MeV and a temperature of around 110 MeV [17–20]. We show a schematic overview of the QCD phase diagram in Figure 1.1. Other proposed phenomena are a color superconducting phase [21–23] with crystalline [24] or other inhomogeneous features, and a moat regime [25] at high chemical potentials, where possible spatial modulations can be explained via the presence of a non-trivial minimum in dispersion relations. The QCD phase diagram, characterized by this wealth of nontrivial phenomena, constitutes an area of active and ongoing research, with numerous fundamental questions yet to be resolved through theoretical and experimental approaches [26].

Experimentally, the phase diagram of QCD is studied at RHIC in Brookhaven National Laboratory (BNL) and the Large Hadron Collider (LHC) at CERN in different heavy-ion collision (HIC) systems, such as gold-gold (Au-Au) and lead-lead (Pb-Pb) collisions. Currently, the FAIR experiment at GSI Darmstadt is under construction, which will further explore the

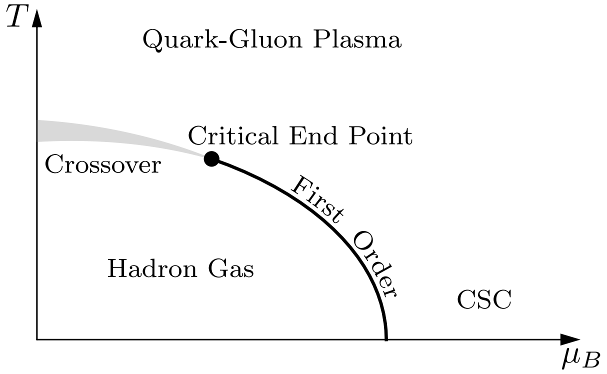


Figure 1.1: Projected phase diagram of QCD in the plane of temperature T and baryon chemical potential μ_{B}. At small baryon chemical potential, there is a crossover transition from the QGP to the hadronic phase. At higher chemical potentials, a first-order phase transition is expected to occur, ending in a critical point. For very high chemical potentials, a color superconducting (CSC) phase is expected to exist.

QCD phase diagram, especially at high baryon chemical potential in order to probe the region of the critical end point [27]. Such collisions produce extremely hot and dense matter, that quickly expands and cools down. Therefore, during the evolution after the collision, the resulting matter traverses the phase diagram of QCD and carries information about the different phases and their transitions. Understanding the exact evolution of the produced matter is crucial for understanding the QCD phase diagram.

These collisions of heavy ions are complex processes, consisting of several different stages, illustrated in Figure 1.2. For a review on HICs, see [28]. Directly after the collision, the system is far from equilibrium. Since the initial system has a high energy density and temperature, weak coupling descriptions have been successfully applied to early stages [29]. These have been expanded significantly since then to include more sophisticated methods and give more quantitative predictions [30, 31]. After this short period of pre-equilibrium dynamics, that lasts around 1 fm/c, the system approaches a prethermal or hydrodynamic attractor [32, 33]. From this point on, the system is dominated by hydrodynamic modes, which describe the collective evolution of the system of the initial hot QGP. During the hydrodynamic evolution, the system cools down and expands. It therefore traverses the crossover phase transition line and hadronizes. Consequently, the fluid is converted into hadron distribution functions, described by a hadron gas [34, 35]. Then it continues to cool down and reaches the *thermal freeze out*, where the produced hadrons then freely stream to the detectors. Since the early 2000s, predictions based on hydrodynamic evolution have been remarkably successful in describing final state observables of heavy-ion collisions, such as particle spectra and elliptic flow [36–38]. However, describing the full evolution of HICs is still an active area of research and several frameworks are being developed to model the different stages of the evolution, see

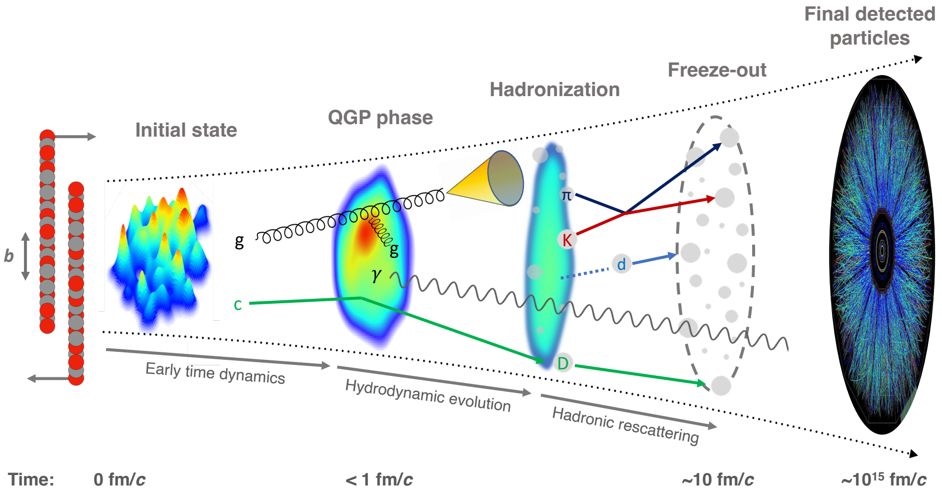


Figure 1.2: Schematic overview of the evolution of a heavy-ion collision and the different stages that occur during the evolution. The figure is taken from [26].

e.g. [39–41].

For such frameworks, a wide range of theoretical input is required to describe the dynamics of such a complex system. On the one hand, it is essential to understand the spectrum of the involved microscopic degrees of freedom, such as hadrons or also other more exotic states of matter. On the other hand, collective degrees of freedom emerge during the hydrodynamic evolution, which need to be described; different orders of the hydrodynamic expansion are necessary to describe the system and their respective transport coefficients have to be computed from the underlying microscopic theories. These theoretical inputs need to be calculated from first-principles methods and serve as an important theory input for these simulations.

Since QCD is a strongly coupled theory, first-principle calculations of such quantities are not possible using perturbation theory. Only at very large scales, e.g. a large temperature or chemical potential, perturbation theory can be applied, since QCD becomes weakly coupled. For other regions of the phase diagrams, the regions of interest for HICs, non-perturbative methods are needed. Very successful first-principle methods used to study QCD are lattice QCD and non-perturbative functional methods, such as the functional renormalization group (fRG), Dyson-Schwinger equations (DSE), or nPI methods. In lattice QCD, the QCD Lagrangian is discretized on a four-dimensional Euclidean lattice with a finite volume and Monte Carlo methods are used to sample field configurations in order to compute the path integral. For reviews on different aspects of lattice QCD see [42–45].

These non-perturbative methods are excellent for computing static properties of QCD such as the position of the crossover phase transition at zero chemical potential. Such quantities can be computed in Euclidean space-time, or equivalently imaginary time, where these methods are much more efficient. However, transport properties of QCD or the early-time behavior in HICs are inherently dynamical objects and need to be understood in a real-time framework, while these non-perturbative methods are formulated in Euclidean space-time. When going to real time in lattice QCD, the sign problem is encountered: the path integral becomes complex and the Monte Carlo sampling of field configurations becomes infeasible, as sampling oscillatory contributions becomes numerically NP-hard [46]. There are several approaches to improving the lattice algorithm for studying QCD in real time, such as complex Langevin [47–49] or utilizing Lefschetz thimbles to perform the complex integration [50, 51]. However, both methods have severe challenges, such as not converging to physical solutions or being numerically expensive. Functional methods, such as the fRG do not suffer from such conceptional problems as the sign problem in lattice methods and real-time computations are generally feasible. Although the formulation of functional equations in Minkowski space-time is significantly more involved, a lot of progress has been made in recent years in this area. Some advancements include the calculation of spectral functions from spectral functional approaches [52–57], analytically continued fRG calculations [58–60] or formulations of fRG directly in real time on the Keldysh-Schwinger closed time-path [61–73]. Although these methods yield impressive results for critical dynamics and real-time correlation functions, fully dynamical QCD calculations remain highly challenging to this day.

In this work, we will focus on a method that does not attempt to solve the theory directly in Minkowski space-time. Instead, we will utilize spectral reconstruction in order to extract real-time correlation functions from non-perturbative Euclidean correlation functions. However, this is still a hard problem, since reconstructing the spectral function from Euclidean correlation functions is – technically speaking – an ill-conditioned problem. By performing the reconstruction of the real-time correlator from data at imaginary time, we can capitalize on the breadth of available non-perturbative data from lattice QCD and functional methods.

From QCD spectral functions, transport coefficients are directly accessible, such as dilepton or photon rates, diffusion coefficients, or shear viscosity [74]. Since around the phase transition, QCD is strongly coupled, these serve as essential inputs for the hydrodynamic evolution of HICs, with the goal of understanding the phase structure of QCD. High precision across a broad range of temperatures and chemical potentials is crucial for accurately predicting final-state observables, which in turn enables a meaningful comparison between theoretical predictions and experimental data. It is therefore essential to employ spectral reconstruction methods that are able to consistently produce reliable results, by incorporating different data sets and the full extent of available theoretical input. In this work, we develop a method based on Gaussian process models to reconstruct spectral functions, and extend this framework to include multiple different constraints and data sets. Utilizing this method, we will reconstruct different spectral functions important for QCD applications in order to compute observables like the glueball masses or the thermal photon rate.

Another question that we will investigate in this thesis, is the process of thermalization in quantum systems, such as in HICs. In order to understand the process of thermalization, we will utilize a well-understood concept from classical mechanics, namely chaos. In general, classical many-body systems that thermalize, usually do so via chaotic dynamics. We want to apply this concept to quantum systems by utilizing the proposed out-of-time ordered correlators (OTOC) [75] and the associated Lyapunov exponent in order to measure and quantify the chaotic behavior of quantum field theories. The underlying idea is that the Lyapunov exponent can be related to the fast relaxing modes of the system, describing the timescales of the initial relaxation towards a prethermal state or hydrodynamic attractor. We would expect that the Lyapunov exponent leads to time scales that are comparable to the hydrodynamization timescale [40]. As a first step towards this goal, we will study the OTOC in the O(4) model. While this does not provide us with direct quantitative results relevant for HICs, it serves as a first step towards understanding OTOCs and their behavior in a system that has the same critical dynamics as chirally symmetric QCD.

This thesis is organized as follows: In Chapter 2, we introduce the necessary fundamental concepts and establish a consistent notation that will be used throughout the work. Chapter 3 presents the method for reconstructing spectral functions from Euclidean correlation functions using Gaussian processes. We detail the underlying mathematical framework, the implementation of the method, and discuss several extensions that allow for the incorporation of additional physical properties. Furthermore, we introduce the Python package *fredipy*, which provides an accessible implementation of the Gaussian process reconstruction method. In Chapter 4, we apply this method to compute glueball masses in pure Yang-Mills theory using four-point gluon correlation functions. Chapter 5 focuses on reconstructing the strong coupling in QCD from ghost and gluon propagators, comparing two distinct reconstruction approaches. In Chapter 6, we analyze the thermal photon rate in QCD using lattice data and assess the consistency of various reconstruction methods. Chapter 7 explores out-of-time ordered correlators in the O(N) model, investigating their behavior and their relation to quantum chaos in field theories. Finally, Chapter 8 summarizes the main findings of this thesis and outlines potential directions for future research.

List of Publications

The results presented in this thesis have been obtained in collaboration with different co-authors. The publications in the context of my thesis are listed below.

[1] Title: "Yang-Mills glueball masses from spectral reconstruction"
Authors: Jan M. Pawlowski, Coralie S. Schneider, Jonas Turnwald, Julian M. Urban and Nicolas Wink
Published in: Phys. Rev. D 108, 076018 (2023)
e-Print: arXiv:2212.01113 [hep-ph]

[2] Title: "Nonperturbative strong coupling at timelike momenta"
Authors: Jan Horak, Jan M. Pawlowski, Jonas Turnwald, Julian M. Urban, Nicolas Wink and Savvas Zafeiropoulos
Published in: Phys. Rev. D 107, 076019 (2023)
e-Print: arXiv:2301.07785 [hep-ph]

[3] Title: "Lattice QCD estimates of thermal photon production from the QGP"
Authors: Ali Sajid, Dibyendu Bala, Anthony Francis, Greg Jackson, Olaf Kaczmarek, Jonas Turnwald, Tristan Ueding and Nicolas Wink
Published in: Phys. Rev. D 110, 054518 (2024)
e-Print: arXiv:2403.11647 [hep-lat]

Additionally, the following software package was developed in the context of this thesis and is presented in Section 3.4

– Title: "*fredipy*: Fredholm Inversion with Python"
Authors: Jonas Turnwald, Julian M. Urban and Nicolas Wink
available at: github.com/JonasTurnwald/fredipy
and: pypi.org/project/fredipy

In [1], the author of this thesis has performed all calculations. These results were already partially presented in the Master's thesis of the author; however, large parts of the error analysis, the exploration of different optimization methods, and the writing of the manuscript were done during the doctoral project. In [2], the author has performed all the numerical GP reconstructions and contributed significantly to the manuscript, while Jan Horak performed the calculation of the strong coupling spectral function directly from the ghost and gluon propagators. In [3], the author has performed the GP reconstructions, contributed significantly to the analysis of the results, and to the relevant parts of the manuscript. The other reconstructions were performed by the co-authors, with leading contributions from Dibyendu Bala, and the author contributed to these mainly through discussions, contributing ideas, and taking part in writing these parts of the manuscript. Hence, these results are only briefly discussed in this thesis, while the focus is on the GP reconstruction method. The perturbative analysis was performed by Greg Jackson. This is not part of the author's doctoral project, and hence is not discussed in this thesis, apart from comparing the results in the final figures. For the software package *fredipy*, the author has taken a leading role, written most of the code, and has written the documentation.

In Chapter 7, we present two unpublished projects on OTOCs. The project on OTOCs with classical statistical simulations was initiated in collaboration with Nicolas Wink. All calculations and the manuscript were completed by the author. The project on OTOCs at weak coupling was initiated in collaboration with Derek Teaney. The author has performed all calculations and written to the manuscript.

2. FUNDAMENTALS

Quantum field theory (QFT) is a versatile framework, that can describe many different physical systems, ranging from solid-state systems to high-energy particle physics. In this chapter, we will introduce some fundamental concepts of QFT dynamics and establish the notation to be used throughout this thesis. We will also give a brief overview of quantum chromodynamics, the theory of the strong interaction, which is the main focus of this thesis.

2.1 DYNAMICS OF QUANTUM FIELDS

In the path integral formalism, the central object of interest is the generating functional. For a general quantum field theory, it is given by the path integral over different field configurations in the presence of a source

$$Z[J] = \int \mathcal{D}\phi \, e^{iS[\phi]+i\int d^4x J(x)\phi(x)} \,, \tag{2.1}$$

where $S[\phi]$ is the action of the field theory, ϕ is the quantum field[1], and J is a source term. The path integral is a functional integral over all possible field configurations ϕ, and by taking the functional derivative of the generating functional w.r.t. the source term, all n-point correlation functions can be obtained. In particular, the two-point function is given by

$$\langle T\phi(x)\phi(y)\rangle = \frac{1}{Z[0]} \frac{\delta^2 Z[J]}{\delta J(x)\delta J(y)}\bigg|_{J=0} \,, \tag{2.2}$$

where T is the time ordering operator. In order to facilitate real-time computations, we will use the Keldysh-Schwinger formalism [77, 78], which is a general approach to computing correlation functions, even in non-equilibrium situations. For introductions to the formalism, that are the basis for the following overview, see [76, 79–82]. Here, we will focus on sketching

---

[1] In this section, we will derive the dynamics for bosonic fields only. When discussing fermionic fields, we obtain very similar equations, often with different signs or Fermi-Dirac statistics instead of Bose-Einstein statistics, due to the anti-commutation relations of fermions. These relations can be found in, e.g. [76].

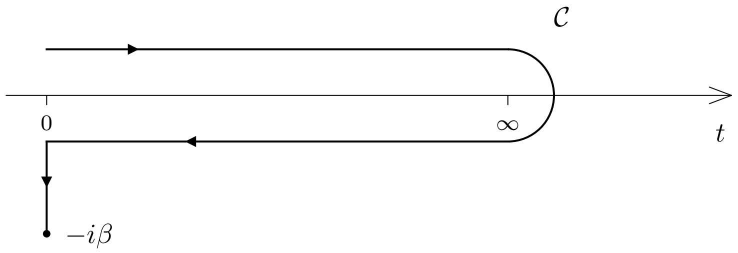


Figure 2.1: Keldysh-Schwinger contour $\mathcal{C}$ in the complex time plane.

the main ideas and introduce the notation. When studying the dynamics of QFT, we want to compute the time evolution of operators, denoted by $\hat{O}$. Their time-dependent expectation value is given in the Heisenberg representation by

$$\langle \hat{O}(t) \rangle \;=\; \mathrm{tr}\left(\rho_0\, \hat{O}_H(t)\right) \;=\; \mathrm{tr}\left(\rho_0\, \hat{U}(0,t)\, \hat{O}\, \hat{U}(t,0)\right) \,, \tag{2.3}$$

where ρ_0 is the initial density matrix at an initial time $t_0 = 0$, and we define $\hat{U}(t_0, t) = T \exp(-i \int_{t_0}^{t} dt'\, \hat{H}(t'))$ as the time evolution operator with the (possibly time-dependent) Hamiltonian operator $\hat{H}$ of the system. If we assume that the system is initially in a thermal state, the initial density matrix is given by $\rho_0 = \exp(-\beta \hat{H})/Z$, where $\beta = 1/T$ is the inverse temperature, and $Z = Z[0]$. For a constant Hamiltonian, the density matrix can be written in terms of the time evolution operator as $\rho_0 = \hat{U}(-i\beta, 0)/Z$ and the expectation value can be expressed as

$$\langle \hat{O}(t) \rangle \;=\; \frac{1}{Z} \mathrm{tr}\left(\hat{U}(-i\beta, 0)\, \hat{U}(0,t)\, \hat{O}\, \hat{U}(t,0)\right) \,. \tag{2.4}$$

From this, we can realize that the time evolution, from right to left in (2.4), can be written as a closed time contour $\mathcal{C}$ in the complex time plane. It goes from $t = 0$ to some final time t and back to $t = 0$, and then from 0 to $-i\beta$ and since we can insert any identity operator in terms of time evolution operators, the time contour can be extended to ∞ and consequently closed. We show such a Schwinger-Keldysh contour in Figure 2.1. Such a contour is often referred to as *close time path* contour when closed at infinity or *Kadanoff-Baym* contour when a thermal branch is added along the imaginary axis. The expectation value of the operator is then given on such a contour as

$$\langle \hat{O}(t) \rangle \;=\; \frac{1}{Z} \mathrm{tr}\left(T_{\mathcal{C}}\, e^{-i \int_{\mathcal{C}} \hat{H}}\, \hat{O}(t)\right) \,, \tag{2.5}$$

where $T_{\mathcal{C}}$ is the time ordering operator on the closed contour and $\int_{\mathcal{C}}$ denotes the integration over the closed contour. Here, the operator $\hat{O}(t)$ is not the time-dependent operator in the Heisenberg representation, but rather the operator inserted at the time t on the closed contour.

We can therefore compute expectation values by inserting operators on the closed contour. For example, when inserting an operator on the complex, temperature part of the contour, we recover the equilibrium thermal expectation value of the operator.

The main advantage of this formalism is that it enables us to compute correlation functions away from thermal equilibrium. In usual path integral computations, the initial and final eigenstates are only allowed to differ by a phase factor. This holds in equilibrium but breaks down for non-equilibrium initial conditions. The Keldysh-Schwinger formalism allows for the computation of correlation functions in such cases, by only ever making a reference to the initial state.

We can then define the Green's functions or two-point correlators, the central objects for solving QFTs, on the Keldysh-Schwinger contour. In order to distinguish between the different branches of the contour, we will denote the upper branch of the contour as $\mathcal{C}_1$ and the lower branch as $\mathcal{C}_2$. The Green's function then has four different choices of operator insertions on either of the segments of the contour, which we denote as G_{ij}, where $i, j = 1, 2$ are the indices of the contour segments. The four different Green's functions are then given by

$$G_{\mathcal{C}} = \begin{pmatrix} G_{11} & G_{12} \\ G_{21} & G_{22} \end{pmatrix} = \begin{pmatrix} G_F & G^{<} \\ G^{>} & G_{\bar{F}} \end{pmatrix} , \tag{2.6}$$

where $G_{F/\bar{F}}$ are the Feynman and anti-Feynman propagators, respectively, and $G^{</>}$ are the lesser/greater propagators that capture cross-contour correlations. They are explicitly given as

$$\begin{aligned} iG_F(t,t') &= \langle T(\phi(t)\phi(t'))\rangle , &(2.7)\\ iG_{\bar{F}}(t,t') &= \langle \bar{T}(\phi(t)\phi(t'))\rangle , &(2.8)\\ iG^{<}(t,t') &= \langle \phi(t')\phi(t)\rangle , &(2.9)\\ iG^{>}(t,t') &= \langle \phi(t)\phi(t')\rangle , &(2.10)\end{aligned}$$

where $\bar{T}$ denotes the anti-time ordering operator, and we have suppressed the spatial arguments for clarity. From these correlators, the well-known retarded and advanced Green's functions and the Keldysh function can be defined as

$$\begin{aligned} G_R(t,t') &= -i\theta(t-t')\langle[\phi(t),\phi(t')]\rangle = G_F - G^{<} , &(2.11)\\ G_A(t,t') &= i\theta(t'-t)\langle[\phi(t),\phi(t')]\rangle = G_F - G^{>} , &(2.12)\\ G_K(t,t') &= -i\langle\{\phi(t),\phi(t')\}\rangle = G^{<} + G^{>} = G_F + G_{\bar{F}} , &(2.13)\end{aligned}$$

where $[\cdot,\cdot]$ and $\{\cdot,\cdot\}$ are the commutator and anti-commutator.

2.1.1 The Spectral Function

With these definitions, we can define the central quantity of interest for this thesis, the spectral function, which is defined as

$$\rho(t,t') \,=\, i(G^{<}(t,t') - G^{>}(t,t')) \,=\, i(G_A(t,t') - G_R(t,t')) \,=\, \langle[\phi(t),\phi(t')]\rangle\,, \tag{2.14}$$

where we again suppress the spatial arguments.

In order to gain a better physical intuition of the spectral function, we will first focus on the spectral function in vacuum, i.e. in the absence of any external sources or heat baths. Hence, we take a complete set of states, which are time-independent, and are denoted by $|n\rangle$ with energy eigenvalues E_n. We insert this into the expectation value of the commutator in (2.14), and take the expectation value w.r.t. the vacuum state $|0\rangle$. This leads to

$$\rho(t,t') \,=\, \sum_n \left(e^{-iE_n(t-t')} - e^{iE_n(t-t')}\right) |\langle 0|\phi(0)|n\rangle|^2 \,. \tag{2.15}$$

The Fourier transform of the spectral function w.r.t. the time difference is then given by

$$\rho(\omega) \,=\, 2\pi\sum_n \left(\delta(\omega - E_n) - \delta(\omega + E_n)\right) |\langle 0|\phi(0)|n\rangle|^2 \,, \tag{2.16}$$

where we immediately see, that the spectral function is an odd function of ω. Writing the spectral function in terms of eigenstates gives a clear interpretation of the spectral function. The spectral function describes the density of states of the system. It gives the probability of exciting a state with energy ω from the vacuum state. This is illustrated in Figure 2.2, where we show an example spectral function for a theory with a physical mass peak – a single-particle state – and a scattering continuum. In vacuum, isolated peaks in the spectral function correspond to states with a physical mass given by the position of the peak. When the energy is sufficient to allow for higher-order processes, the spectral function develops a continuum, since the final state spatial momenta can take any value. In a scalar ϕ^4 theory without symmetry breaking this threshold is given by 3 times the physical mass in order to have a $1 \to 3$ scattering process, see e.g. [56]. At finite temperature, energy can be taken or deposed from the heat bath, leading to a broadening of the delta peaks in the spectral function.

The complex argument Green's function is defined in terms of the spectral function as

$$G(\omega) \,=\, \int_{\mathbb{R}} \frac{dz}{2\pi}\,\frac{1}{z-\omega}\rho(z)\,. \tag{2.17}$$

This is called the *spectral representation* of the Green's function. It is defined for all $\omega \in \mathbb{C}$, and is analytic everywhere except on the real axis, where it can have poles and cuts. All the other Green's functions can be obtained by evaluating this Green's function in this complex plane. The retarded propagator is obtained, by evaluating the complex Green's function on the upper half plane, i.e. $G_R(\omega) = G(\omega + i\epsilon)$, and the advanced propagator is obtained by

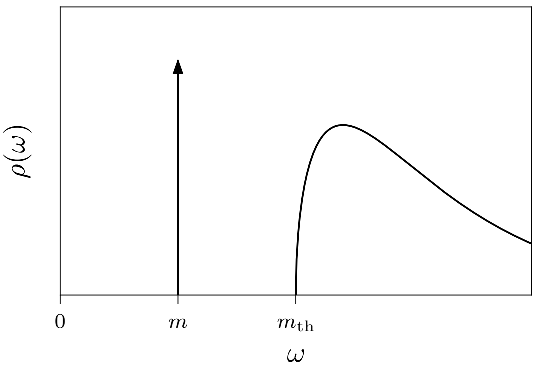


Figure 2.2: Illustration of some spectral function $\rho(\omega)$ in vaccum. At the physical mass m, we have a delta peak, which corresponds to the single-particle state. We have then plotted a continuum of states starting at some mass threshold m_{th}, which corresponds to the lowest energy that is necessary to create multiple particles. Since these particles can carry space-like momentum, the spectrum becomes continuous. For finite temperature, all features of the spectral function are broadened due to the thermal fluctuations, and delta peaks are smeared into Breit-Wigner distributions.

evaluating the complex Green's function on the lower half plane, i.e. $G_A(\omega) = G(\omega - i\epsilon)$, while the Feynman propagator comes from both these contributions $G_F(\omega) = G(\omega + i\epsilon\,\text{sign}(\omega))$. These prescriptions then lead to the right causality properties of the respective propagators, as defined in (2.7), (2.11) and (2.12). From the spectral representation (2.17), we immediately obtain the spectral function as[2]

$$\rho(\omega) \;=\; 2\,\text{Im}(G(\omega + i\epsilon))\,. \tag{2.18}$$

This relation between the imaginary part of the complex argument Green's function and the spectral function is known as the Kramers-Kronig relation, which is a general statement about complex functions with analyticity requirements, cf. [76].

In non-perturbative simulations, the Euclidean propagator is typically computed, as calculating correlators in real time is extremely challenging or even infeasible. This Euclidean version arises from a Wick rotation, replacing real time t with imaginary time τ via $t \to -i\tau$. In vacuum, the Euclidean propagator in momentum space is given by evaluating the complex Green's function on the imaginary axis, i.e. $G_E(\omega) \equiv G(i\omega)$. The spectral function can be connected to the Euclidean propagator via the *Källén-Lehmann spectral representation* [83, 84], that follows from the spectral representation (2.17) by utilizing the antisymmetry of the spectral function. It follows

$$G_E(\omega) \;=\; \int_0^\infty \frac{d\omega'}{\pi}\, \frac{\omega'}{\omega'^2 + \omega^2}\, \rho(\omega')\,. \tag{2.19}$$

---

[2] using the identity $\frac{1}{x \pm i\epsilon} = \mp i\pi\delta(x) + \text{P.V.}\frac{1}{x}$

Note, that in vacuum, we have Lorentz symmetry. Therefore, the spectral function and the propagators only depend on sign(ω), and the four-momentum squared. We can therefore write the spectral representation in vacuum (2.19), also in terms of the – Euclidean or Minkowskian[3] – four-momenta.

Going to finite temperature $T = 1/\beta$, the Euclidean propagator is defined on a finite interval $\tau \in [0, \beta]$, and is periodic in imaginary time. Therefore, when going to frequency space, the Euclidean propagator is evaluated on discrete Matsubara frequencies, which are given by $\omega_n = \frac{2\pi n}{\beta}$ for bosons and $\omega_n = \frac{(2n+1)\pi}{\beta}$ for fermions. The Källén-Lehmann representation (2.19) is still valid by simply replacing the continuous ω with the discrete Matsubara frequencies ω_n, and we obtain

$$G_E(\omega_n) = \int_0^\infty \frac{d\omega'}{\pi} \frac{\omega'}{\omega'^2 + \omega_n^2} \rho(\omega') \,. \tag{2.20}$$

The relation between the Euclidean propagator at imaginary time and the spectral function requires some care. A detailed derivation is provided in Appendix A, from which we obtain the spectral representation

$$G_E(\tau) = \int_0^\infty \frac{d\omega}{2\pi} \frac{\cosh\left(\omega(\tau - \beta/2)\right)}{\sinh\left(\omega\beta/2\right)} \rho(\omega) \quad \text{for } 0 \leq \tau \leq \beta \,. \tag{2.21}$$

The factor of $1/(2\pi)$ is a normalization factor, that can differ depending on the conventions. Note that in Chapter 6, we will use a different normalization, which is given by $1/\pi$, since this is consistent with other works on thermal photon rates.

Calculating these spectral functions from Euclidean propagators is one of the central objectives of this thesis. We will discuss the difficulties of this task in Section 3.1 and present a method to reconstruct the spectral function from Euclidean correlators in Section 3.3.

2.1.2 KMS Relations

In thermal equilibrium, the different Green's functions are related to each other via the Kubo-Martin-Schwinger (KMS) relations. These provide a powerful tool to relate the different real-time Green's functions to each other and to the spectral function. Using the cyclicity of the trace and $\rho_0^{-1}\phi(t)\rho_0 = \phi(t - i\beta)$, we arrive at the KMS relation for the lesser and greater Green's functions

$$\begin{aligned} G^<(t,t') &= \operatorname{tr}\left(\rho_0 \phi(t')\phi(t)\right) \\ &= \operatorname{tr}\left(\rho_0 \rho_0^{-1}\phi(t)\rho_0\phi(t')\right) = G^>(t + i\beta, t') \,. \end{aligned} \tag{2.22}$$

The KMS relation states that the equilibrium Green's functions are periodic in imaginary time and consequently lead to discrete Matsubara frequencies in momentum space. When

---

[3] We use the "mostly minus" convention for the metric: $g_{\mu\nu} = \text{diag}(1, -1, -1, -1)$

taking the Fourier transform w.r.t. the time difference, we can write the KMS relation as

$$G^{<}(\omega) \,=\, e^{-\beta\omega} G^{>}(\omega)\,. \tag{2.23}$$

This KMS relation can be used to relate Green's functions and provides an abstract definition of equilibrium states, i.e. states that fulfill the KMS relations are in equilibrium. Additionally, the fluctuation-dissipation theorem can be immediately derived from this relation as

$$G_K(\omega) \,=\, \left(G_R(\omega) - G_A(\omega)\right) \coth\left(\frac{\beta\omega}{2}\right)\,. \tag{2.24}$$

This can then be connected to the spectral function, by noting that $G_R(\omega)^* = G_A(\omega)$. We can then write the spectral function in frequency space as

$$\rho(\omega) \,=\, i G_K(\omega) \tanh\left(\frac{\beta\omega}{2}\right) \,=\, i G^{<}(\omega)\left(e^{\beta\omega} - 1\right)\,. \tag{2.25}$$

Consequently, in equilibrium, all the propagators and the spectral function are related to each other via the KMS relations. The spectral function – or, by that logic, any other of the Green's functions – contains all the information about the dynamics.

2.2 Fundamentals of Quantum Chromodynamics

Quantum chromodynamics describes the strong interaction, one of the four fundamental forces in the standard model, which acts between quarks and gluons. QCD is a non-abelian gauge theory based on the gauge group SU(3). The non-abelian nature of QCD allows the force carriers, the gluons, to have self-interactions, leading to rich and complex dynamics. Already, pure SU(3) Yang-Mills (YM) theory, i.e. QCD without fermions and only considering gluons, is a non-trivial theory, which is strongly coupled at low energies and asymptotically free at high energies. This leads to the phenomenon of color confinement: the asymptotic states of QCD are color singlets. Quarks and gluons, which are charged under SU(3), are never observed in the physical spectrum. Confinement can be associated with the conjectured mass gap of YM theory, although the fundamental degrees of freedom are massless on the level of the classical field equations. A rigorous proof of the existence of YM theory and the mass gap is still an open problem in mathematical physics and is one of the millennium prize problems, see [85] and references therein. However, the emergence of a dynamical mass gap in the gauge sector of QCD, which physically manifest as the lightest glueballs, has been observed in multiple lattice simulations, see e.g. [86–91], and in continuum methods, see e.g. [92–103]. In Chapter 4, we will present a study of the glueball spectrum of YM theory using spectral reconstruction techniques and functional renormalization group methods.

We obtain the full QCD Lagrangian similarly as in other gauge theories, such as quantum electrodynamics (QED), by requiring local gauge invariance of the action. In the case of QED, the gauge group is U(1), here we consider the gauge group SU(3). When we have the Dirac action for the free quarks of the form

$$S_{\text{Dirac}} = \int d^4x\, \bar{\psi}(i\partial\!\!\!/ - m)\psi \,, \tag{2.26}$$

where $\partial\!\!\!/ = \gamma^\mu \partial_\mu$ and γ^μ are the Dirac matrices, the action is not invariant under local SU(3) transformations of the form $\psi \to \mathcal{U}\psi$, where $\mathcal{U} \in \mathrm{SU}(3)$, since the partial derivative acts on the local transformation. To make the action gauge invariant, we replace the partial derivative with the gauge-covariant derivative given by

$$D_\mu = \partial_\mu - igA_\mu \,, \tag{2.27}$$

where $A_\mu \in \mathfrak{su}(3)$ is the gauge field, and g is the gauge coupling. The gauge fields A_μ are in the adjoint representation of the gauge group and elements of the Lie algebra $\mathfrak{su}(3)$ with generators t^a, such that

$$A_\mu = A_\mu^a t^a, \quad \text{with} \quad a = 1, \ldots, N_c^2 - 1 = 8 \,. \tag{2.28}$$

For SU(3), the generators are the Gell-Mann matrices, and in general, they satisfy the commutation relations

$$[t^a, t^b] = if^{abc}t^c \,, \tag{2.29}$$

where f^{abc} are the structure constants of the Lie algebra, and they are normalized such that $\mathrm{tr}(t^a t^b) = \frac{1}{2}\delta^{ab}$, where the trace is in the fundamental representation. The structure constants fully determine the Lie algebra, and consequently, the group structure. As the name suggests, the covariant derivative (2.27) transforms covariantly under gauge transformations $\mathcal{U} \in \mathrm{SU}(3)$, from which, we can derive the transformations of the gauge field

$$D_\mu \to \mathcal{U} D_\mu \mathcal{U}^\dagger \quad \text{and} \quad A_\mu \to \mathcal{U} A_\mu \mathcal{U}^\dagger + \frac{i}{g}\mathcal{U}\partial_\mu \mathcal{U}^\dagger \,. \tag{2.30}$$

We introduce the field strength tensor or curvature tensor $F_{\mu\nu}$, which is defined as the commutator of the covariant derivative

$$F_{\mu\nu} = \frac{i}{g}[D_\mu, D_\nu] = \partial_\mu A_\nu - \partial_\nu A_\mu - ig[A_\mu, A_\nu] \,. \tag{2.31}$$

The field strength tensor also transforms covariantly under gauge transformations. The full action of QCD is then given by

$$S_{\text{QCD}} = \int d^4x \left[\bar{\psi}(iD\!\!\!\!/ - m)\psi - \frac{1}{4}\mathrm{tr} F_{\mu\nu}F^{\mu\nu} \right] , \tag{2.32}$$

where we have a term that describes the dynamics of the gauge fields in terms of the field strength tensor, and we take the trace w.r.t. the color indices. Due to the cyclicity of the trace, the action is invariant under gauge transformations. When considering abelian gauge theories, the commutator of the gauge fields vanishes in (2.31) and the product of the field strength tensors does not produce any self-interactions of the gauge fields, i.e. there are no photon self-interactions on the level of the classical field equations in QED. However, in non-abelian gauge theories, the YM term leads to self-interactions of the gauge fields – on the level of the classical action we have three and four-point interactions – which are responsible for the rich dynamics of QCD.

Since the YM action (and also the full QCD action) is gauge invariant, all physical observables are gauge invariant. However, when computing the generating functional of YM theory, an integration over all possible gauge field configurations is performed. These configurations include infinitely many physically equivalent configurations, that are related by gauge transformations. This problem can be immediately seen in the breakdown of the classical propagator of the gauge fields; due to its transversality, the two-point function has a vanishing eigenvalue and can therefore not be inverted, in order to obtain the gluon propagator. In methods that rely on the gauge-dependent fields as the fundamental degrees of freedom, such as functional methods, it is therefore necessary to fix a gauge. There are different gauge fixing conditions, such as covariant gauges, axial gauges, and Coulomb gauges, which are used in different contexts. In this work, we will use the Landau gauge, which is a covariant gauge, and is defined by the condition

$$\partial_\mu A^\mu = 0\,. \tag{2.33}$$

Such a gauge fixing condition corresponds to choosing a representative from each gauge orbit, up to possible Gribov copies [104]. In general, gauge fixing conditions split the general path integral measure dA into an integration over physically inequivalent, gauge fixed contributions A_{gf} and the corresponding gauge transformation $\mathcal{U}$, such that

$$dA = J\, dA_{\text{gf}}\, d\mathcal{U}\,, \tag{2.34}$$

where J is the Jacobian of the integral transformation and $d\mathcal{U}$ is the Haar measure of the gauge group. The computation of such a coordinate transformation, including the Jacobian is done via Faddeev-Popov quantization, introduced in [105]. This leads to the introduction of ghost fields c, Grassmann valued fields that are unphysical degrees of freedom emerging from the gauge fixing procedure. The gauge fixed YM action then reads

$$S_{\text{YM}} = \int d^4x \left[-\frac{1}{4} F^a_{\mu\nu} F^{a,\mu\nu} + \frac{1}{2\xi} (\partial_\mu A^{a,\mu})^2 + \bar{c}^a \partial_\mu D^{ab,\mu} c^b \right] , \tag{2.35}$$

where ξ is the gauge fixing parameter. Choosing the Landau gauge corresponds to setting $\xi = 0$. This can be understood intuitively: In the path integral, all gauge field configurations with $\partial_\mu A^\mu \neq 0$ are suppressed by a weight $\lim_{\xi\to 0} \exp(-\frac{1}{2\xi}(\partial_\mu A^\mu)^2)$ and consequently only gauge configurations with $\partial_\mu A^\mu = 0$ contribute. Other gauges, such as Feynman gauge ($\xi = 1$) do not have such simple interpretations, but they can be useful in different contexts. For example, Feynman gauge simplifies perturbative calculations. By construction, the gauge fixing procedure, and specifically the gauge fixing term, explicitly breaks gauge invariance. However, this information about the gauge symmetry of YM is not lost, it is encoded in the BRST (Becchi-Rouet-Stora-Tyutin) symmetry. This symmetry can be understood heuristically as a symmetry that leads to a compensation of the gauge variance of the gauge fixing term by the ghost fields.

As mentioned before, QCD is strongly coupled at low energies and asymptotically free at high energies. This is a consequence of the running of the strong coupling g^2 given by the β function at one-loop order as

$$\beta(g) = -\frac{g^3}{(4\pi)^2}\left(\frac{11}{3}N_c - \frac{2}{3}N_f\right) , \tag{2.36}$$

where N_c is the number of colors and N_f is the number of flavors. For sufficiently small N_f, the β function is negative, which means that the coupling decreases with increasing energy scale. This phenomenon of asymptotic freedom has been first described by Gross, Wilczek, and Politzer in 1973 [106, 107] and has been awarded the Nobel prize in physics in 2004. Since then, the perturbative running of this coupling has been confirmed by multiple collider experiments [108]. However, at small energies the coupling grows, perturbation theory encounters a Landau pole and perturbation theory breaks down. The theory becomes strongly coupled, and non-perturbative methods are necessary. Many advancements have been made in the last decades by performing studies of QCD on the lattice and with functional methods such as Dyson-Schwinger equations and the functional renormalization group. For the current status of QCD calculations with fRG methods, see [109], and references therein, and for reviews on lattice QCD, see [42–45].

At small energies, where the coupling becomes large, the gluons and quarks are confined to bound states, which are color singlets. The physical spectrum of QCD below the phase transition is therefore composed of hadrons. We can distinguish between two types of hadrons: One type is baryons, which consist of an odd number of quarks, which include protons and neutrons, and therefore make up most of the observed mass in the universe, apart from dark matter. The second type is mesons, which are composite states made of an even number of quarks, usually quark-antiquark pairs. They are usually short-lived, and in contrast to baryons, they do not obey fermionic statistics, but have an integer valued spin and are therefore bosons. As bosons, they can be viewed as force carriers of a residual strong or electroweak force between baryons, which is mediated by the exchange of mesons. The lightest mesons are pions. They appear as the Goldstone modes of the spontaneously broken chiral symmetry. Since this symmetry is also explicitly broken due to finite quark masses, the pions are not massless but have a small mass of about 135 MeV for the π^0. Still, they are significantly lighter than the other hadrons. Hence, at low energies the pions are the most relevant degree of freedom, dominating the low energy dynamics of QCD. Knowing the spectral function of the lightest mesons is therefore crucial for understanding the low energy dynamics of QCD [110, 111].

When knowing the spectral function of the important degrees of freedom – below the phase transition these are the light mesons and above the phase transition the conserved currents – we can use these to extract bound state masses, decay widths, and other important dynamical properties of the system. Notably, using Kubo relations [112, 113], we can relate real-time correlation functions – and therefore spectral functions – to transport coefficients using the fluctuation-dissipation theorem (2.25). We conclude, that the spectral function is a central object in QCD, and knowing it enables us to extract important physical properties of the system. Therefore, this will be a large focus in the following chapters.

3. SPECTRAL RECONSTRUCTION

As discussed in the previous chapter, the spectral function is a crucial quantity in many fields of physics. When knowing the spectral functions, the whole imaginary plane of the Green's function is available, and the spectral function can be used to calculate many important real-time observables. In non-perturbative calculations, the spectral function must be reconstructed from the Euclidean correlator, which is connected to the spectral function via the inverse of the Källén-Lehmann spectral representation (2.19).

In this chapter, we will discuss the difficulties of spectral reconstruction and give an overview of different approaches developed in the context of QCD in Section 3.1. We will focus on a method that is supposed to systematically alleviate some of the problems that arise in spectral reconstruction, namely Gaussian processes (GPs). To do so, we will introduce the basic concepts of GPs in Section 3.2, and discuss how they can be used for simple regression problems. The explicit application of GPs to inverse problems, such as spectral reconstruction, is then discussed in Section 3.3. We will demonstrate how GPs can be used to reconstruct the spectral function from the Euclidean correlator and discuss the advantages of this method and the difficulties associated with spectral reconstruction. In Section 3.4, we will introduce the Python package *fredipy*, which is a tool for solving linear inverse problems, especially spectral reconstruction, using GPs. We will discuss specific spectral reconstruction examples and provide a brief overview of the package. In the following Chapters 4 – 6, we will then apply this method to different problems in QCD.

3.1 SPECTRAL RECONSTRUCTION: AN ILL-CONDITIONED PROBLEM

The problem of spectral reconstruction is an inverse problem, where the objective is to invert the Källén-Lehmann spectral representation, cf. (2.19) and (2.21), to extract the spectral function from the Euclidean correlator. Interestingly, the problem of spectral reconstruction in the context of QFTs is a special problem in the more general class of problems with the purpose of inverting Fredholm integral equations of the first kind, commonly defined as

$$\int_a^b dy\, K(x,y) f(y) \;=\; g(x)\,, \tag{3.1}$$

where $K(x,y)$ is some known continuous kernel function, $f(y)$ is the unknown function we want to reconstruct, and $g(x)$ provides the data we want to fit. This is a ubiquitous problem in many fields of physics and engineering, such as optics, geophysics, and medical imaging [114]. In general, many of such problems are ill-posed, meaning that the solution is not unique [115]. Due to the inherent difficulty of such problems, a variety of methods have been developed to address them. For some general overviews, see [114–117].

As already established earlier, in the context of QFTs, the spectral function is a crucial quantity when studying dynamical observables, such as transport coefficients, or when calculating the bound states of the theory. Although it can be shown, that the spectral function can be uniquely determined from the Euclidean propagator in the limit of infinite data points with infinite precision [118], this limit is usually not feasible. Consequently, the inversion of the discretized version of the Källén-Lehmann spectral representation is an ill-conditioned problem. Due to the presence of close-to-zero modes in the operator that is inverted, small changes in the data can lead to significant changes in the prediction. In practice, however, the Euclidean data comes from lattice or functional QFT simulations. Such simulations inherently involve statistical and numerical noise, and in the case of lattice simulations, the number of available data points is strictly constrained by the finite lattice size. When having finite, noisy data, the inversion of the Källén-Lehmann spectral representation produces an amplification of the noise up to several orders of magnitude. Consequently, the spectral representation can not be naively inverted, and we need to perform some regularization procedure to stabilize the inversion.

One of the most straightforward approaches to regularize the inversion is to use a Tikhonov regularization, which adds a penalty term to the inversion problem. The most naive realization of Tikhonov regularization in the context of discrete inversion, i.e. matrix inversion, is to add a scalar multiple of the identity matrix before the inversion, such that

$$A\mathbf{x} \;=\; \mathbf{b} \quad \Rightarrow \quad \mathbf{x} \sim (A + \lambda \mathbb{1})^{-1}\mathbf{b}\,, \tag{3.2}$$

where A is some near-singular, and in this example, square matrix. This improves the stability of the inversion, but it can be hard to quantify the effect of the regularization on the reconstructed function. One of the goals of this thesis is to introduce a more systematic approach to regularizing this problem.

Since the problem of spectral reconstruction is present in many important areas of QCD, a wide range of methods have been developed and utilized to tackle this problem. The aforementioned Tikhonov regularization method is one of the most widely used methods [119–121]. However, in the context of spectral reconstruction, in some cases, a significant amount of prior information can be inferred about the spectral function, which strictly limits the possible solutions, and the functional form of the spectral function can be assumed. Therefore, explicit fits to some functional form [122–125] or expansions in some basis functions [126, 127] are often applied. Another method for regularizing the inversion is the Backus-Gilbert method [128–132], where the problem is regularized by reconstructing a smoothed version of the spectral function. In recent years, statistical or Bayesian methods have been gaining popularity, such as the maximum entropy method [133–136], Bayesian inference techniques [137–139] or kernel ridge regression related methods [140, 141]. Recently, neural network approaches have become a popular tool for spectral reconstruction [142–152], as the forward problem is efficient and stable and therefore allows the construction of large datasets, that can be used to train the neural networks. Some spectral reconstruction approaches additionally make use of spatial correlator data [153, 154], which can effectively constrain the properties of the spectral function.

Instead of using the Källén-Lehmann spectral representation, one can also use the Padé approximation [155, 156] to perform the analytic continuation (2.18) directly. Other methods make use of Nevanlinna theory [157–161], constructing a continuous fraction expansion for the analytic continuation. The methods mentioned above are only a small selection of the many approaches that have been developed to tackle the problem of spectral reconstruction.

One important aspect, that most of the methods mentioned above have in common, is that it can be difficult to quantify the systematic uncertainty introduced by the regularization procedure, either through some smoothing or the assumption of a fixed functional basis. This generally introduces uncontrollable systematic uncertainties. Additionally, depending on the method it can be difficult, or even impossible, to include additional prior information about the spectral function. For example, it is currently not possible to include information about the normalization of the spectral function in Backus-Gilbert methods. In order to alleviate some of these problems, we will introduce a new method for spectral reconstruction based on Gaussian processes in the following sections.

3.2 GAUSSIAN PROCESSES

In this section, we will introduce the basic concepts of Gaussian processes (GPs) on a technical level, and understand how they can be used for simple regression problems. Since spectral reconstruction is, in essence, a regression problem with no direct information about the spectral function, but only about some integrated correlator, we will first focus on direct regression. With these examples, we will be able to understand the basic concepts of Gaussian processes, and only later apply them to inverse problems.

Originally, Gaussian processes were developed in the field of Geophysics, in order to interpolate irregularly distributed and noisy data, that is abundant in experimental measurements over different spatial and/or temporal scales. Today, Gaussian processes are used in many fields, such as machine learning, statistics, and physics. Especially in the context of machine learning Gaussian processes have been experiencing a surge in interest in the last two decades. One of the connections between GPs and machine learning is the fact, that in the limit of infinite width, a single-layer fully connected neural network with an i.i.d. Gaussian prior on the weights is equivalent to a Gaussian process [162]. Furthermore, GPs can be used as supervised non-parametric models for regression and classification in different machine learning tasks and deep Gaussian process models, cf. [163–166] for some notable examples in the extensive literature.

In the following, we will give a brief introduction to Gaussian processes. We mainly follow [163], [167] and [168], but a wide range of high-quality literature exists, that have excellent introductions to the broad field of Gaussian processes such as [169–172].

Gaussian processes are a type of stochastic process commonly used to model and predict an observable variable y based on input variables x. As in general Bayesian models, the conditional distribution $p(y\,|\,x)$ can then describe this connection, and if $y \in \mathbb{R}$, such a model is called a regression model. In this work, we will only discuss regression models, but GPs can be used in classification as well, see e.g. [163]. Instead of doing a direct regression, we assume that we can separate the model into a systematic and a random part. The systematic part is given by a function $f(x)$ and the sampling distribution or likelihood is then

$$p(y\,|\,f(x))\,. \tag{3.3}$$

We therefore assume that y and x are connected by some latent function f and the probabilistic part captures the uncertainty in the model. Note, that the conditional distribution does not explicitly depend on the input x, only indirectly through the function $f(x)$. Additional parameters $\boldsymbol{\sigma}$ can be added to the regression model. One common parameter is the noise level since in most regression problems we have some experimental or numerical data that does not exactly reproduce the function $f(x)$. For example, when performing regression, where we have some normal noise with variance σ_n^2 on the data, the likelihood is extended to

$$p(y\,|\,f(x),\boldsymbol{\sigma}) \;=\; \mathcal{N}(y\,|\,f(x),\sigma_n^2)\,, \tag{3.4}$$

where we denote a normal distribution with mean μ and covariance Σ as $\mathcal{N}(\mu,\Sigma)$.

The aim of regression, i.e. the goal of Gaussian process regression, is then to find the form of the function f based on input data and prior assumptions. There are (infinitely) many approaches to any regression problem, but we will focus on two ways of thinking about this problem. One very straightforward way is to restrict ourselves to a finite set of basis functions, such as a finite number of polynomials, e.g. $\boldsymbol{\phi}(x) = (1,\, x,\, x^2,\, x^3,\, \dots)^\top$. Finding each coefficient w_i of the basis functions then gives an estimate of the desired function as $f(x) = \sum_i w_i \phi_i(x)$. This is in essence, what is done in Bayesian linear regression: we map the input to a set of basis functions and then perform regression on the level of the parameters of the basis functions, such as the prefactors of a polynomial model. However, such a parametric regression model has the disadvantage that the accuracy at which we can predict the underlying function f is bounded by the function having exactly that structure. So when we do not know prior information that highly restricts the number of basis functions, we can artificially restrict ourselves to potentially wrong models.

Another approach is to perform regression not on the level of parameters, but instead on the level of the function directly. Gaussian process models utilize Gaussian processes as a prior on f, giving a prior distribution over all possible functions. Such a Gaussian process is non-parametric in the sense that we formulate a prior on the function values directly instead of assuming some finite set of parameters that are essentially fitted to the data. This is done via the *kernel trick*, where we define a covariance function that describes the similarity between points in the input space, which allows for models that do not explicitly formulate an underlying basis of functions. This is discussed in detail in Section 3.2.3. In essence, the Gaussian process is a distribution over functions, and the regression problem is then to find the posterior distribution over functions that are consistent with the data. Gaussian processes are the generalization of Gaussian probability *distributions* to infinite dimensions. Instead of describing the probability distribution over a set of points $\mathbf{x}$, the evaluation of a stochastic process returns a function value $f(x)$ for every $x \in \mathcal{X} \subseteq \mathbb{R}$.

A Gaussian process is with a mean $\mu(x)$ and a covariance function, or *kernel*, $k(x, x')$ is then denoted as

$$f(x) \sim \mathcal{GP}(\mu(x),\, k(x, x'))\,. \tag{3.5}$$

Such a Gaussian process can be evaluated at every continuous point $x \in \mathcal{X}$, the evaluation at a finite set of points $\tilde{\mathbf{x}} = (x_1, ..., x_n)$ is given by a multivariate normal distribution, i.e.

$$p(\mathbf{f} \,|\, \tilde{\mathbf{x}}) = \mathcal{N}(\boldsymbol{\mu},\, \mathbf{K})\,, \tag{3.6}$$

where we denote $\mathbf{f} = (f(x_1), ..., f(x_n))^\top$ with the mean $\boldsymbol{\mu} = (\mu(x_1), ..., \mu(x_n))^\top$, and the covariance $\mathbf{K} = (k(x_i, x_j))_{i,j=1,...,n}$. From this, we can straightforwardly see that a Gaussian process is an infinite-dimensional generalization of a multivariate normal distribution: Instead of a finite number of points, a GP can be evaluated at every point in the input space. The covariance function or kernel of the Gaussian process usually introduces additional parameters, that we denote by $\boldsymbol{\theta}$. We will discuss in Section 3.2.1, how the choice of kernel and its parameters impacts the regression model and in Section 3.2.3 specify how the choice of the kernel function can restrict or broaden the set of possible functions are described by the Gaussian process. Optimizing the kernel choice and the associated parameters is a pivotal step in the regression process. However, we will defer this discussion to later sections and focus on the basic concepts of Gaussian processes in this section. For now, we will just extend the distributions by a parameter $\boldsymbol{\theta}$, and assume a fixed model that is specified by this parameter(s).

We can use Gaussian processes to construct a probabilistic model and condition such a model on some data set $\mathcal{D} = \{(x_i, y_i)\,|\, i = 1, ..., n'\}$, where n' is the number of samples. With data at hand, we can then calculate the conditional posterior distribution using Bayes' rule

$$p(\mathbf{f}\,|\,\mathbf{y}, \boldsymbol{\theta}) \;=\; \frac{p(\mathbf{y}\,|\,\mathbf{f})\,p(\mathbf{f}\,|\,\boldsymbol{\theta})}{p(\mathbf{y}\,|\,\boldsymbol{\theta})}\,, \tag{3.7}$$

where $\mathbf{f} = (f(x_1), ..., f(x_{n'}))^\top$ is now evaluated at the sample points. In Bayesian statistics, $p(\mathbf{f}\,|\,\boldsymbol{\theta})$ is the *prior* distribution, which quantifies the probability of the model producing a certain draw. The term $p(\mathbf{y}\,|\,\mathbf{f})$ is called the *likelihood*, it quantifies how likely some random draw from the prior reproduces the data. And the denominator $p(\mathbf{y}\,|\,\boldsymbol{\theta})$ is the *marginal likelihood*, which is the probability of the model reproducing the data. It is called marginal likelihood, since it is the likelihood of the data marginalized over the latent function values, i.e. the numerator integrated over all possible latent function values $\mathbf{f}$. Since the posterior of the process is a pivotal ingredient of the regression, we will outline the calculation of it in the following. Note, that the posterior predictive distribution $p(\mathbf{f}\,|\,\mathbf{y}, \boldsymbol{\theta})$ is also often written as $p(\mathbf{f}\,|\,\mathcal{D}, \boldsymbol{\theta})$, where, in a sense, $\mathbf{x}$ appears on both sides and is therefore redundant. However, the meaning is the same, and we will use both notations interchangeably.

The simplest case for deriving the posterior distribution is to assume that the data is noise-free and reproduces the function exactly $y = f(x)$. In this case the posterior on $\mathbf{f}$ is $p(\mathbf{f}\,|\,\mathbf{y}) = \delta(\mathbf{f} - \mathbf{y})$. However, in most practical cases, we expect some noise on the data. In order for the model to remain analytically traceable, we have to assume that the noise is Gaussian distributed. The observed data often given by $y = f(x) + \epsilon$, where the noise $\epsilon \sim \mathcal{N}(0,\, \sigma_n^2 \mathbb{1})$ is i.i.d. Gaussian distributed. In general, we can also assume correlated Gaussian noise on the data set and have a full covariance matrix $\mathbf{C}_n$, possibly with non-zero off-diagonal entries. We therefore reintroduce the additional parameter $\mathbf{C}_n$ to the likelihood, see (3.4). The posterior on $\mathbf{f}$ is then given by Bayes' rule (3.7) as

$$p(\mathbf{f}\,|\,\mathbf{y},\, \boldsymbol{\theta},\, \mathbf{C}_n) \;\sim\; p(\mathbf{y}\,|\,\mathbf{f},\, \mathbf{C}_n)\, p(\mathbf{f}\,|\,\mathbf{x},\, \boldsymbol{\theta})\,. \tag{3.8}$$

For a Gaussian process, the likelihood of f with some noise, as discussed above, is given by

$$p(\mathbf{y}\,|\,\mathbf{f},\, \mathbf{C}_n) \;=\; \mathcal{N}\left(\mathbf{y}\,|\,\mathbf{f},\, \mathbf{C}_n\right)\,. \tag{3.9}$$

The prior distribution in (3.8) has already been discussed above, e.g. in (3.6). In order to simplify the derivation we will set the prior mean to zero, i.e. $\mu(x) = 0$. Choosing any different, non-zero mean function merely introduces a shift in the results. One can equivalently subtract the mean function from the data, perform the regression assuming a zero mean, and then add the mean back to the predictions, yielding the same outcome as incorporating the mean function directly into the regression. With a zero mean the prior distribution of $\mathbf{f}$ is then given by

$$p(\mathbf{f}\,|\,\mathbf{x},\, \boldsymbol{\theta}) \;=\; \mathcal{N}(\mathbf{0},\, \mathbf{K})\,. \tag{3.10}$$

With these two distributions, we can calculate the posterior distribution

$$\begin{aligned} p(\mathbf{f}\,|\,\mathbf{y},\,\boldsymbol{\theta}) &\sim \exp\left(-\frac{1}{2}(\mathbf{y}-\mathbf{f})^\top\mathbf{C}_n^{-1}(\mathbf{y}-\mathbf{f})-\frac{1}{2}\mathbf{f}^\top\mathbf{K}^{-1}\mathbf{f}\right) \\ &\sim \exp\left(-\frac{1}{2}\mathbf{f}^\top\left(\mathbf{C}_n^{-1}+\mathbf{K}^{-1}\right)\mathbf{f}+\mathbf{f}^\top\mathbf{C}_n^{-1}\mathbf{y}\right) \\ &\sim \exp\left(-\frac{1}{2}(\mathbf{f}-\mathbf{u})^\top\boldsymbol{\Sigma}^{-1}(\mathbf{f}-\mathbf{u})\right)\,, \end{aligned} \tag{3.11}$$

where the covariance $\boldsymbol{\Sigma}$ is

$$\begin{aligned} \boldsymbol{\Sigma} &= (\mathbf{C}_n^{-1}+\mathbf{K}^{-1})^{-1} \\ &= \mathbf{C}_n\mathbf{K}(\mathbf{C}_n+\mathbf{K})^{-1}\,, \end{aligned} \tag{3.12}$$

and the mean $\mathbf{u}$ is

$$\mathbf{u} = \mathbf{C}_n^{-1}\boldsymbol{\Sigma}\mathbf{y} = \mathbf{K}(\mathbf{C}_n+\mathbf{K})^{-1}\mathbf{y}\,. \tag{3.13}$$

Since the product of two multivariate normal distributions is again a multivariate normal distribution, the right normalization immediately follows and the posterior distribution of the Gaussian process is

$$p(\mathbf{f}\,|\,\mathbf{y},\,\boldsymbol{\theta}) = \mathcal{N}(\mathbf{u},\,\boldsymbol{\Sigma}) = \mathcal{N}\left(\mathbf{K}(\mathbf{K}+\mathbf{C}_n)^{-1}\mathbf{y},\,\sigma^2\mathbf{K}(\mathbf{K}+\mathbf{C}_n)^{-1}\right)\,. \tag{3.14}$$

From this posterior distribution, the posterior predictive distribution can be calculated in order to make predictions for the function f at new points $\mathbf{x}_*$. The vector of latent function values is denoted by $\mathbf{f}_*$, and its predictive distribution is obtained by marginalizing over the posterior uncertainty in the latent function values

$$p(\mathbf{f}_*\,|\,\mathcal{D},\,\mathbf{x}_*,\,\boldsymbol{\theta}) = \int d\mathbf{f}\;p(\mathbf{f}_*\,|\,\mathbf{f},\,\mathbf{x}_*,\,\boldsymbol{\theta})\,p(\mathbf{f}\,|\,\mathbf{y},\,\boldsymbol{\theta})\,, \tag{3.15}$$

where on the left-hand side, we restored the dependence on the data $\mathcal{D}$, and not only $\mathbf{y}$, since the data locations are not already specified in the latent function values. Next, we calculate the first term in the integral, which describes the dependency of $\mathbf{f}_*$ on $\mathbf{f}$. This captures the dependency of the regression on the GP prior. We calculate this conditional distribution from the joint prior distribution of $\mathbf{f}$ and $\mathbf{f}_*$. This is again a multivariate normal distribution and given by

$$p(\mathbf{f},\,\mathbf{f}_*\,|\,\mathbf{x},\,\mathbf{x}_*,\,\boldsymbol{\theta}) = \mathcal{N}\left(\begin{pmatrix}\mathbf{f}\\ \mathbf{f}_*\end{pmatrix}\middle|\begin{pmatrix}\boldsymbol{\mu}\\ \boldsymbol{\mu}_*\end{pmatrix},\begin{pmatrix}\mathbf{K} & \mathbf{K}_*\\ \mathbf{K}_*^\top & \mathbf{K}_{**}\end{pmatrix}\right)\,, \tag{3.16}$$

where $\mathbf{K}$ and $\mathbf{K}_{**}$ are the prior covariance matrices of $\mathbf{f}$ and $\mathbf{f}_*$, respectively, and $\mathbf{K}_*$ is the covariance matrix between $\mathbf{f}$ and $\mathbf{f}_*$. Obtaining the conditional distribution $\mathbf{f}_*\,|\,\mathbf{f}$ from the joint

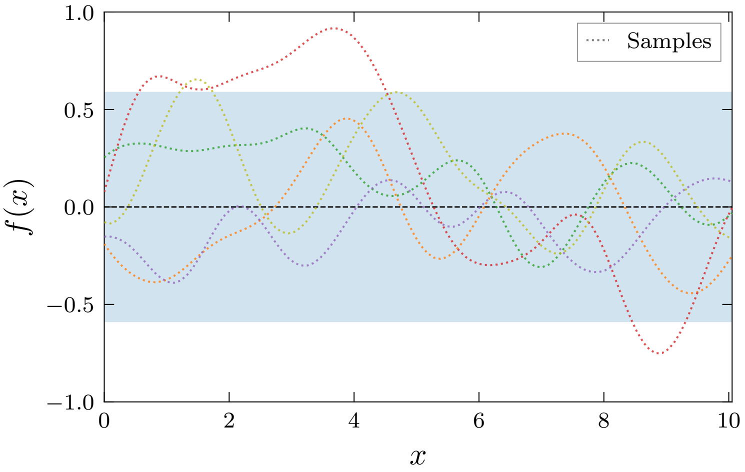


Figure 3.1: Plot of a Gaussian process prior with mean zero, $\mu(x) = 0$ and a squared exponential kernel $k(x, x') = \sigma^2 \exp\left(-\frac{1}{2\ell^2}(x - x')^2\right)$, with $\sigma = 0.3$ and $\ell = 0.8$. The blue shaded area indicates the 95% confidence interval. We have drawn some samples from the Gaussian process, in order to illustrate the fact that Gaussian processes are distributions over functions.

prior distribution is a standard result in the context of multivariate normal distributions, see e.g. [170], and given by

$$p(\mathbf{f}_* \,|\, \mathbf{f},\, \mathbf{x}_*,\, \boldsymbol{\theta}) \;=\; \mathcal{N}(\mathbf{f}_* \,|\, \boldsymbol{\mu}_* + \mathbf{K}_*^\top \mathbf{K}^{-1}(\mathbf{f} - \boldsymbol{\mu}),\, \mathbf{K}_{**} - \mathbf{K}_*^\top \mathbf{K}^{-1} \mathbf{K}_*)\,. \tag{3.17}$$

The second part under the integral of the posterior predictive distribution in (3.15), i.e. the posterior distribution of the process depends on the given data $\mathcal{D}$. Finally, collecting all the equations and inserting (3.17) and (3.14) into (3.15), a calculation presented in Appendix B.1 leads to the posterior predictive distribution of the Gaussian process

$$p(\mathbf{f}_* \,|\, \mathcal{D},\, \mathbf{x}_*,\, \boldsymbol{\theta},\, \boldsymbol{\sigma}) \;=\; \mathcal{N}\left(\mathbf{f}_* \,|\, \mathbf{K}_*^\top (\mathbf{K} + \mathbf{C}_n)^{-1} \mathbf{y},\, \mathbf{K}_{**} - \mathbf{K}_*^\top (\mathbf{K} + \mathbf{C}_n)^{-1} \mathbf{K}_*\right)\,. \tag{3.18}$$

For regression problems, we use this posterior predictive distribution to make predictions at new points $\mathbf{x}_*$, and the covariance matrix quantifies the uncertainty of the prediction. The impressive feature of this result is that the predictive distribution is analytically traceable. We can calculate the mean and variance via simple matrix operations instead of numerically optimizing parameters.

The above argument directly generalizes to arbitrary inputs $\mathbf{x}_*$ and therefore, we can write the predictive distribution as a Gaussian process, that can be evaluated at arbitrary $x \in \mathcal{X}$

$$f(x) \,|\, \mathcal{D} \;\sim\; \mathcal{GP}(\mu_*(x),\, k_*(x, x'))\,. \tag{3.19}$$

The mean and covariance directly generalize from (3.18) to

$$\mu_*(x) \;=\; \mathbf{k}(x)^\top (\mathbf{K} + \mathbf{C}_n)^{-1} \mathbf{y} \tag{3.20}$$

$$k_*(x, x') \;=\; k(x, x') - \mathbf{k}(x)^\top (\mathbf{K} + \mathbf{C}_n)^{-1} \mathbf{k}(x')\,, \tag{3.21}$$

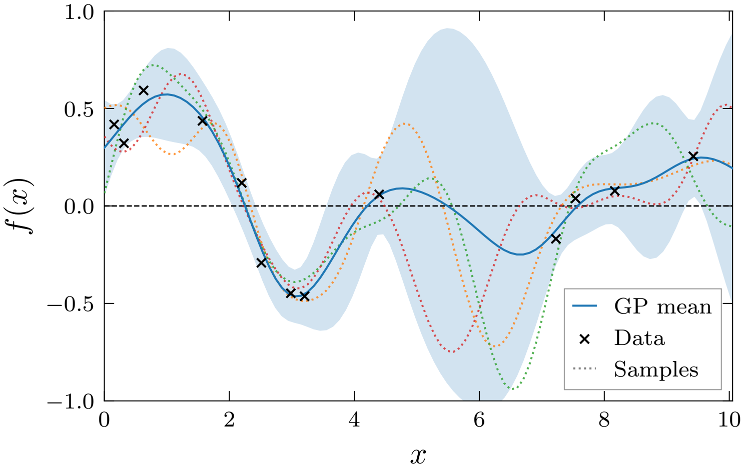


Figure 3.2: Plot of the Gaussian process posterior conditioned on some noisy data, denoted by black crosses. We perform this regression with the GP prior presented in Figure 3.1,i.e. with mean zero, $\mu(x) = 0$ and a squared exponential kernel $k(x, x') = \sigma^2 \exp\left(-\frac{1}{2\ell^2}(x - x')^2\right)$, with $\sigma = 0.3$ and $\ell = 0.8$. The blue solid line is the posterior mean of the GP and the blue shaded area indicates the 95% confidence interval. We have drawn some samples from the Gaussian process, in order to illustrate the fact that Gaussian processes are distributions over functions.

where $\mathbf{k}(x) = (k(x_1, x), ..., k(x_n, x))^\top$ is the vector of covariances between the data and the test point(s). Therefore, we can make predictions at any point x utilizing the posterior of the Gaussian process and immediately propagate the error on the data through the process. Note, that in practice we do not compute the explicit inverse of the covariance matrix in (3.20) and (3.21), but instead use the Cholesky decomposition to solve the linear system, for details see Appendix B.2. This provides a numerically more stable and efficient algorithm since we invert a positive definite and symmetric matrix [163].

In Figure 3.1 we show an example of a Gaussian process prior (3.10) with a zero mean and some kernel function $k(x, x')$. To gain some intuition, we can think about the Gaussian process as a normal distribution at each point in the input space, so at each x–slice we have a normal distribution, with the variance given by the diagonal entries of the kernel function. If they were uncorrelated Gaussian distributions at each point, we would have uncorrelated random draws at each x. However, in contrast to normal distributions, draws at different values of x are correlated, since the off-diagonal entries of the kernel function are usually non-zero. The length and structure of such correlations is precisely quantified by the kernel function. In this example, we have chosen a squared exponential kernel, meaning that points that are close in the input space are in return close in the output space and vice versa. As a result, we obtain smooth functions, when we take draws from the Gaussian process. Another point illustrated by these samples is that a Gaussian process represents a distribution over functions. However, in subsequent plots, we will omit the samples and focus solely on the mean and variance.

We then use this GP prior and condition it on some noisy data. The resulting posterior predictive distribution is shown in Figure 3.2, where we can see how the mean interpolates the data while the diagonal of the covariance matrix gives the uncertainty of the prediction. Again, we can draw samples from the GP posterior. These samples are functions that, with a high probability, are consistent with the data. This visualizes that the GP posterior is a distribution over the functions in the GP prior that are consistent with the data.

3.2.1 COVARIANCE FUNCTIONS

In the previous section, we have seen that the choice of the covariance function, the kernel, is the only prior information that we have to specify in order to perform regression with Gaussian processes. Therefore, it is important to carefully consider the influence of the kernel and understand the potential biases it may introduce into the model. In the following, we will introduce the general properties of kernel function and discuss some common kernel choices. Afterwards, we will discuss, how kernels can be extended and how kernel parameters are optimized.

The kernel captures the dependencies between function values, i.e. the covariance between $f(x)$ and $f(x')$

$$k(x,x') = \operatorname{cov}(f(x), f(x'))\,. \tag{3.22}$$

As mentioned beforehand, the kernel quantifies the strength and structure of the correlation between the function values.

For k to be a valid kernel, it needs to map the pair of inputs (x, x') into $\mathbb{R}$, and it has to be symmetric and positive semi-definite

$$k(x,x') = k(x',x) \quad \text{and} \quad \int dx dx'\, k(x,x') f(x) f(x') \geq 0\,, \tag{3.23}$$

for all $f \in L^2$. If a kernel function fulfills these properties, a valid corresponding covariance matrix $(\mathbf{K})_{ij} = k(x_i, x_j)$ can be constructed.

A kernel is called *stationary* if it only depends on the difference of the inputs, i.e. $k(x,x') = k(x-x')$. More specifically, if the kernel only depends on the (Euclidean) distance between the inputs $|x-x'|$, it is called *isotropic*. Since the kernel only depends on the distance, it is invariant under translations and rotations and often called *radial basis function* (RBF). The squared exponential kernel, introduced in the examples in Figures 3.1 and 3.2, is a widely used RBF (and therefore isotropic and stationary) kernel.

An instructive example of a GP kernel function is the linear kernel $k(x,x') = \sigma_1 + \sigma_2(x - c)(x' - c)$. In this case, the Gaussian process coincides with a Bayesian linear regression model. Calculating the posterior predictive distribution, we find that the mean is a linear fit of the data. Of course, using GPs for linear regression is not the most computationally efficient way, but such kernels can be used to construct more complex kernels. Additionally, we can see that other, more specialized, regression models can be constructed as special cases of Gaussian processes, by choosing the appropriate kernel function.

In general, it can be difficult to show that a given kernel is positive semi-definite. There are some general rules, however, that allow for new kernels k to be constructed from existing, positive semi-definite kernels k_1 and k_2, such as

$$\begin{aligned} k(x,x') &= c_1 k_1(x,x') + c_2 k_2(x,x') \quad \text{for } c_1, c_2 \geq 0 \\ k(x,x') &= k_1(x,x') k_2(x,x') \,. \end{aligned} \tag{3.24}$$

Another useful transformation is to apply a function to the inputs, i.e. $k(\phi(x), \phi(x'))$, where the only requirement is that k is positive semi-definite on the target space of ϕ. Using this, we can perform prior data analysis, and for example, transform the input space into a space where the data is more easily separable.

In the following, we will discuss some common kernel functions that are used in practice and discuss their properties and use cases. For an interactive introduction to some of these kernel classes, an excellent is provided exists in [172].

The aforementioned squared exponential kernel is given by

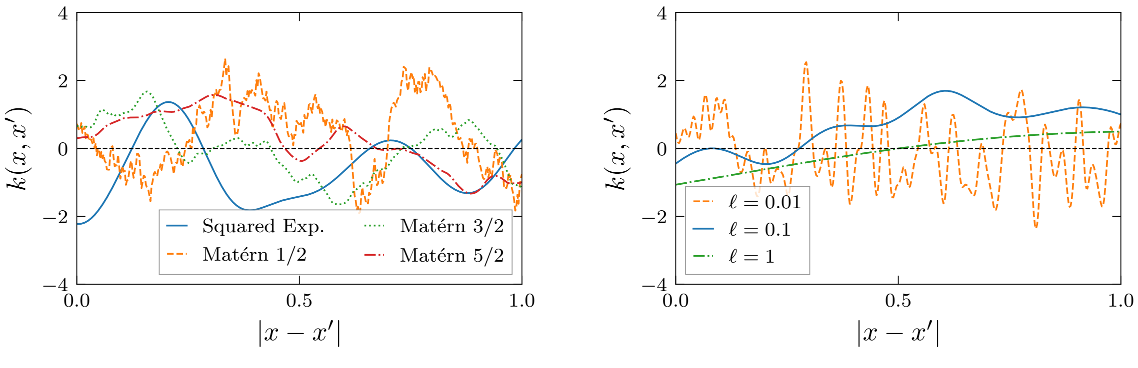
$$k(x,x') = \sigma^2 \exp\left(-\frac{1}{2\ell^2}(x-x')^2\right) , \tag{3.25}$$

where σ controls the variance and ℓ is some length scale. Especially the length scale ℓ has an intuitive interpretation in the context of regression. It controls the smoothness of the function, smaller values of ℓ lead to more oscillations in the interpolation or regression, while larger values of ℓ lead to smoother functions. We compare function draws from a GP with different length scales in Figure 3.3b. When doing regression, this kernel is often used as a default choice. It correlates points that are close in the input space, while points that are far away in x-direction are uncorrelated, which is usually a good assumption for many problems. This leads to interpolations that are smooth functions and otherwise do not assume other prior information. Note that the squared exponential kernel is infinitely differentiable, which leads to draws from the GP being in C^∞, see the discussion in Section 3.2.3.

Since such a strong smoothness assumption can be too restrictive for some problems in regression, another widely used kernel is the Matérn kernel class. It is given by

$$k(x,x') = \sigma^2 \frac{2^{1-\nu}}{\Gamma(\nu)} \left(\frac{\sqrt{2\nu}}{\ell}|x-x'|\right)^\nu K_\nu\left(\frac{\sqrt{2\nu}}{\ell}|x-x'|\right) , \tag{3.26}$$

where K_ν is a modified Bessel function of the second kind and Γ is the gamma function. The variance σ, the length scale ℓ, and the smoothness parameter ν are positive parameters. While σ and ℓ have the same interpretation as in the squared exponential kernel, the parameter ν



(a) (b)

Figure 3.3: Plots of different functions drawn from Gaussian processes with differing kernels and zero mean. In (a), we compare sample paths from the squared exponential kernel and the Matérn kernel with $\nu = 1/2$, $3/2$ and $5/2$. In (b), we show the behavior of the squared exponential kernel for different length scales. The variance is set constant for all draws, $\sigma = 1$.

controls the differentiability of the function. A GP with a Matérn kernel in $\nu - 1$ times differentiable, meaning that small values of ν lead to rougher functions. For $\nu \to \infty$, the Matérn kernel converges to the squared exponential kernel. The Matérn kernel is mostly used at small half-integer values of ν, such as $\nu \in \{1/2, 3/2, 5/2\}$, since these values are computationally more efficient to calculate and are distinct from the squared exponential kernel. In these cases the kernel is

$$k_{\nu=1/2}(x, x') = \sigma^2 \exp\left(-|x - x'|/\ell\right), \tag{3.27}$$

$$k_{\nu=3/2}(x, x') = \sigma^2 \left(1 + \sqrt{3}|x - x'|/\ell\right) \exp\left(-\sqrt{3}|x - x'|/\ell\right), \tag{3.28}$$

$$k_{\nu=5/2}(x, x') = \sigma^2 \left(1 + \sqrt{5}|x - x'|/\ell + 5|x - x'|^2/(3\ell^2)\right) \exp\left(-\sqrt{5}|x - x'|/\ell\right). \tag{3.29}$$

In Figure 3.3a, we show the different behavior of the Matérn kernel for different values of ν.

3.2.2 Model Selection and Parameter Optimization

In the previous section, we have discussed the GP kernel and some examples of common kernels. We have seen that different kernels or different kernel parameters, i.e. different regression models, lead to widely different results, see Figure 3.3. It is therefore crucial to develop consistent methods for model selection and find objects that quantify the quality of the regression model. In general, model selection incorporates the choice of the kernel function and the kernel parameters. In this work, however, we will focus on the kernel parameters, since the choice of the kernel class in the interpolation problems treated here, has turned out to be less crucial, or we have explicit prior information about the functional form and want to explicitly keep the kernel class fixed, see Section 3.3.3. Generally useful kernels, like the squared exponential kernel, have been proven powerful in many problems, and changing this kernel usually does not lead to significant improvements in the regression quality, apart from specialized problems, e.g. periodic functions. Still, different methods have been developed for automatic kernel selection, see e.g. [173, 174] and might be a future extension to this work. The parameter optimization of the GP kernel is often called *training*, especially in the context of machine learning.

Since we only treat the kernel parameters as unknowns, we distinguish different models by the kernel parameters denoted by $\boldsymbol{\theta}$. We use the *marginal likelihood* $p(\mathcal{D}\,|\,\boldsymbol{\theta})$, i.e. the denominator in Bayes' rule (3.7), to compare different models. The marginal likelihood quantifies the probability of generating the observed data $\mathcal{D}$ given the model parameters $\boldsymbol{\theta}$ and given by marginalizing over the likelihood times the prior

$$p(\mathcal{D}\,|\,\boldsymbol{\theta}) \;=\; \int d\mathbf{f}\; p(\mathcal{D}\,|\,\mathbf{f})\, p(\mathbf{f}\,|\,\boldsymbol{\theta})\,. \tag{3.30}$$

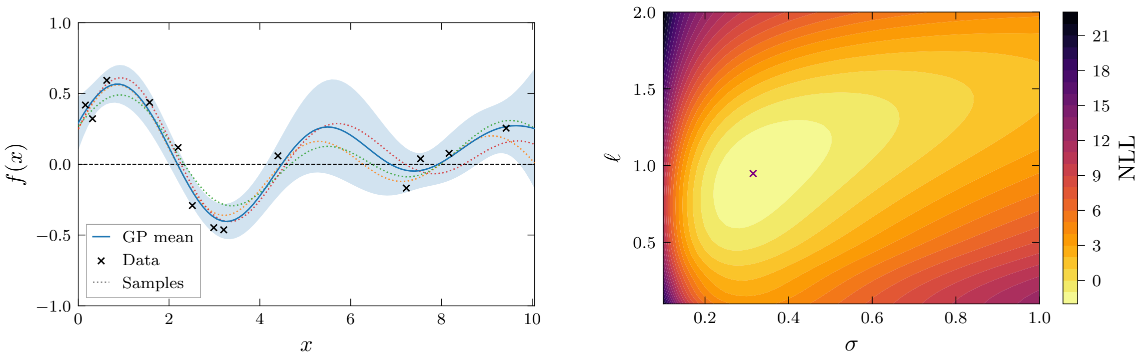
As discussed previously the prior of the GP is given by $\mathbf{f}\,|\,\boldsymbol{\theta} \sim \mathcal{N}(\mathbf{0},\, \mathbf{K})$ and the likelihood is given by $\mathcal{D}\,|\,\mathbf{f} \sim \mathcal{N}(\mathbf{f},\, \mathbf{C}_n)$. The log marginal likelihood is then given by

$$\begin{aligned}
\log p(\mathcal{D}\,|\,\boldsymbol{\theta}) \;&=\; \log \int d\mathbf{f}\; \mathcal{N}(\mathbf{f}\,|\,\mathbf{0},\, \mathbf{K})\, \mathcal{N}(\mathcal{D}\,|\,\mathbf{f},\, \mathbf{C}_n) \\
&=\; \log \mathcal{N}\left(\mathcal{D}\,|\,\mathbf{0},\, \mathbf{K} + \mathbf{C}_n\right) \\
&=\; -\frac{1}{2}\mathbf{y}^\top (\mathbf{K} + \mathbf{C}_n)^{-1}\mathbf{y} - \frac{1}{2}\log|\mathbf{K} + \mathbf{C}_n| - \frac{n}{2}\log 2\pi\,.
\end{aligned} \tag{3.31}$$

These three terms have intuitive interpretations in the context of optimization. The first term is the only one that explicitly depends on the observed targets $\mathbf{y}$ and quantifies the quality of the data fit. The second term acts as a complexity penalty and only depends on the kernel function. It penalizes models that are too complex, e.g. very small length scales in the squared exponential kernel leading to almost diagonal $\mathbf{K}$, and prevents overfitting. The last term is a normalization term only depending on the number of data points n, and is therefore not relevant to parameter optimization.

Note that the optimized kernel parameters can include the noise covariance matrix $\mathbf{C}_n$, or in simpler cases the signal variance σ_n with $\mathbf{C}_n = \sigma_n^2\mathbb{1}$. However, it has been shown, that optimization on the noise parameter σ_n and the variance of the kernel $\sigma^2 k(x, x')$ simultaneously, can lead to numerical instabilities and especially overfitting [175]. Therefore, we will keep the noise fixed and only optimize the kernel parameters in the context of this work.

In Figure 3.4, we show an example of hyperparameter optimization. In the simple example of two kernel parameters, we can use common minimizers to find the minimum of the negative log-marginal likelihood. Or for very simple problems, we can perform scans of the two-dimensional parameter space, as shown in the figure. For more advanced models, the number of parameters can grow rapidly, leading to a high-dimensional optimization problem.



(a) (b)

Figure 3.4: GP regression with optimized hyperparameters with the same data as in Figure 3.2. In (a), we show the GP regression with the likelihood-optimized hyperparameters $\ell = 0.32$ and $\sigma = 0.95$ for the squared exponential kernel. The blue solid line is the posterior mean of the GP and the blue shaded area indicates the 95% confidence interval. In (b), we show a parameter scan of the corresponding negative log-marginal likelihood (NLL). The red cross indicates the optimized hyperparameters.

Additionally, apart from simple examples, the optimization problem is non-convex, and the optimization algorithm can get stuck in local minima. This makes finding the right hyperparameters typically the most involved part of the regression process. In order to calculate the log-marginal likelihood (3.31), we need to invert the covariance matrix in each optimization step. The computational complexity of matrix inversion grows as $\mathcal{O}(n^3)$, where n is the number of data points. This is the main drawback of Gaussian processes: the computational cost grows rapidly with the number of data points, while especially in machine learning applications, the number of data points is generally very large[1]. However, the optimization can be performed more efficiently with gradient-based optimization methods. The gradient of the marginal likelihood w.r.t. the hyperparameters is given by

$$\frac{\partial}{\partial\theta_i}\log p(\mathcal{D}|\boldsymbol{\theta}) \;=\; \frac{1}{2}\mathbf{y}^\top(\mathbf{K}+\mathbf{C}_n)^{-1}\frac{\partial\mathbf{K}}{\partial\theta_i}(\mathbf{K}+\mathbf{C}_n)^{-1}\mathbf{y} - \frac{1}{2}\mathrm{tr}\left((\mathbf{K}+\mathbf{C}_n)^{-1}\frac{\partial\mathbf{K}}{\partial\theta_i}\right)\,, \tag{3.32}$$

with an additional computational cost of calculating the gradient of the order $\mathcal{O}(n^2)$. Thus, choosing gradient-based optimization methods is generally advantageous.

3.2.3 KERNEL EIGENFUNCTIONS

In this section, we will have a close look at the mathematical structure underlying Gaussian processes. This will help to understand the impact of the kernel function on the regression and lead to modifications that incorporate prior information about the function. As another

---

[1] A major part of research on GPs is devoted to finding ways around this problem by using methods like sparse Gaussian processes, for a review on different methods, see [176].

perspective on Gaussian processes, we will introduce *reproducing kernel Hilbert spaces* (RKHS) and lay out how these concepts are related. For a thorough introduction to the general theory of RKHS and kernel eigenfunctions, we refer to [177], while [169] gives an introduction with a stronger connection to GPs. For a more detailed and mathematical introduction [178] is a great resource, focussing on more general kernel methods.

We use the definition of a kernel function as a symmetric, positive semi-definite function that maps the input space into $\mathbb{R}$, i.e. $k : \mathcal{X} \times \mathcal{X} \to \mathbb{R}$. We can then define the RKHS as
Definition 3.2.1 (Reproducing Kernel Hilbert Space)**.** *A Hilbert space $\mathcal{H}_k$ of functions on $\mathcal{X}$ equipped with an inner product $\langle \cdot, \cdot \rangle_{\mathcal{H}_k}$ is called a reproducing kernel Hilbert space with reproducing kernel k, if for all $x \in \mathcal{X}$, we have $k(x, \cdot) \in \mathcal{H}_k$ and for all $f \in \mathcal{H}_k$ and $x \in \mathcal{X}$, we have $f(x) = \langle f, k(x, \cdot) \rangle_{\mathcal{H}_k}$.*

The RKHS can be constructed from some kernel k as

$$\mathcal{H}_k = \left\{ f(x) = \sum_{i=1}^{\infty} \alpha_i k(x, x_i) \,|\, \alpha_i \in \mathbb{R}, x_i \in \mathcal{X}, \text{ such that } \langle f, f \rangle < \infty \right\}, \tag{3.33}$$

where the inner product is given by

$$\langle f, g \rangle_{\mathcal{H}_k} = \sum_{i=1}^{\infty} \sum_{j=1}^{\infty} \alpha_i \beta_j k(x_i, x_j) \,. \tag{3.34}$$

From this construction, we can see that the bases of the RKHS are the kernel functions $k(x, x_i)$ and the functions in the RKHS inherit the properties from the kernel, such as smoothness or differentiability. Additionally, an RKHS $\mathcal{H}_k$ is unique for every positive semi-definite kernel function k by the Moore-Aronszajn theorem [179].

It follows immediately that a kernel function defining the covariance of a Gaussian process has an associated RKHS. When examining the posterior predictive mean of a Gaussian process (3.20), we can see that the mean can be written as a linear combination of the kernel functions

$$\mu_*(x) = \sum_{i=1}^{n} \alpha_i \, k(x_i, x) \,, \tag{3.35}$$

where the coefficients $\boldsymbol{\alpha} = (\mathbf{K} + \mathbf{C}_n)^{-1} \mathbf{y}$ are constants w.r.t. the test point x and only depend on the training data $\mathbf{y}$ and the kernel function at the training points. The posterior mean of the Gaussian process is therefore in the RKHS of the kernel function. So it is pivotal to understand the properties of the RKHS in order to understand how regression with Gaussian processes behaves[2].

This can be further understood by looking at Mercer's theorem [181]:

---

[2] Note, that while the posterior mean lies in the RKHS of the GP kernel, sample paths of the GP prior do not necessarily – and for many kernels almost surely not – lie within the RKHS. For a GP, a sample path does not belong to the corresponding RKHS with probability 1, if $\mathcal{H}_k$ is infinitely dimensional [169]. Sample paths are generally rougher or less regular than RKHS functions. However, they are still connected, as GP sample spaces can be constructed as powers of RKHS [180].

Theorem 3.2.1 (Mercer's Theorem)**.** *Suppose that a continuous, symmetric, positive semi-definite kernel function* $k : \mathcal{X} \times \mathcal{X} \to \mathbb{R}$ *is bounded, i.e.* $|k(x, x')| < \infty$ *for all* $x, x' \in \mathcal{X}$*. We then define the integral operator* T_k *as*

$$T_k f(x) \,=\, \int_{\mathcal{X}} k(x, x') f(x') dx' \,. \tag{3.36}$$

Such operators are called Hilbert-Schmidt operators. Since T_k *is compact, positive, and self-adjoint, the spectral theorem (see e.g. [178], Theorem A.5.13) guarantees that an eigendecomposition* (ϕ_i, λ_i)*, where* ϕ_i *are the eigenfunctions and* λ_i *the eigenvalues, exists*

$$T_k \phi_i \,=\, \lambda_i \phi_i \,. \tag{3.37}$$

The number of eigenvalues is equal to the dimension of the associated RKHS, and they are positive. The kernel function k *can then be expressed in terms of the eigenvalues and eigenfunctions as*

$$k(x, x') \,=\, \sum_{i=1} \lambda_i \phi_i(x) \phi_i(x') \,. \tag{3.38}$$

The proof of this theorem can be found in [182].

The eigenfunctions are therefore an orthogonal basis of the RKHS. Since the posterior mean is in the RKHS, we can understand the interpolation that is performed by the Gaussian process as a projection of the data onto the eigenfunctions of the kernel. As a result, the behavior of the interpolation can be understood by analyzing the set of eigenfunctions of the kernel.

Before we continue the discussion of the eigenfunctions of specific kernels, we will discuss the *kernel trick*, which is a powerful tool in machine learning and statistics. It is, in a sense, the reverse of the above constructions, where we construct a RKHS from a kernel function. We define a *feature map* $\varphi : \mathcal{X} \to \mathcal{H}$, where $\mathcal{H}$ is a Hilbert space called *feature space.* A kernel is then defined as the inner product in feature space

$$k(x, x') \,=\, \langle \varphi(x), \varphi(x') \rangle_{\mathcal{H}} \,, \tag{3.39}$$

where k is symmetric and positive semi-definite by the definition of an inner product. The natural example that arises from Mercer's theorem is $\varphi_i(x) = \sqrt{\lambda_i}\phi_i(x)$, with an inner product

$$\langle \boldsymbol{\varphi}(x), \boldsymbol{\varphi}(x') \rangle_{\mathcal{H}} = \sum_i \lambda_i \phi_i(x) \phi_i(x') = k(x, x') \,. \tag{3.40}$$

However, in general, we construct kernels without explicitly knowing the eigenfunctions of the feature map. This has a major computational advantage: we do not have to calculate high orders in the feature space, e.g. $\varphi_i(x) = c_i x^i, \;\, i \in \{0, \dots, N\}$ for polynomial kernels up to order N, but can implicitly formulate our problem in terms of kernels. This is widely used in machine learning, where the feature space can be of very high dimension or even infinite, and the kernel trick allows for efficient calculations. In the context of interpolation, if we know the basis of the function that we interpolate, we can construct a kernel that is the inner product of the basis functions, and we can perform the interpolation in the RKHS of the kernel. We discussed this in Section 3.2.1, where GP regression reproduces linear regression for a linear kernel, and we will revisit this when constructing special kernels that incorporate asymptotics of the function in Section 3.3.3.

We have seen, that RKHS can have very different sizes; small RKHS can be spanned by a few eigenfunctions, in order to prevent overfitting when we have clear prior information about the underlying data. Another powerful class of kernels referred to as *universal kernels* are defined as:

Definition 3.2.2 (Universal Kernels)**.** *A kernel k on $\mathcal{X}$ is called universal if the associated RKHS $\mathcal{H}_k$ is dense in $C(\mathcal{X})$, the space of continuous functions on $\mathcal{X}$, i.e. for all $f \in C(\mathcal{X})$ and $\epsilon > 0$, there exists a function $g \in \mathcal{H}_k$ such that*

$$\|f - g\|_\infty < \epsilon\,. \tag{3.41}$$

Such kernels allow for the approximation of any continuous function with arbitrary precision. They are extremely flexible and can be used to continuously interpolate data without any prior assumptions.

It can be shown (from [183], Theorem 9), that kernels of the Taylor type are universal.

Theorem 3.2.2 (Universal Taylor Kernels)**.** *Let $f : \mathbb{R} \to \mathbb{R}$ be a C^∞ function that can be Taylor expanded at 0 as*

$$f(x) = \sum_{i=0}^{\infty} a_i x^i\,. \tag{3.42}$$

If we have $a_i > 0$ for all i, and some scalar product $\langle \cdot, \cdot \rangle : \mathcal{X} \times \mathcal{X} \to \mathbb{R}^+$, then the kernel k given by

$$k(x, x') = f(\langle x, x' \rangle)\,, \tag{3.43}$$

is a universal kernel.

This means that the exponential kernel $k(x, x') = \exp(-\|x - x'\|/\ell)$ and the squared exponential kernel $k(x, x') = \exp(-\|x - x'\|^2/(2\ell^2))$, presented in Section 3.2.1, are universal kernels. This is one of the reasons, such kernels are so widely used: They can approximate any continuous function, and we do not make prior assumptions.

Note that, although the squared exponential kernel is a universal kernel, its associated RKHS does not have to contain *all* continuous functions. For example, the constant, non-zero function, and any polynomial function are not contained in the RKHS [184, 185]. However, for practical purposes, this usually does not pose a problem since the RKHS is dense in the space of continuous functions. Hence, we can get arbitrarily close to these functions. This can become problematic in the limit of an infinite number of data points and the error approaching zero. In such cases, the convergence towards the true function can be slow.

3.3 Spectral Reconstruction with Gaussian Process Regression

In the previous section, we introduced Gaussian processes as a tool for regression and discussed the mathematical properties of GPs and their kernels. Since the spectral reconstruction problem is, at heart, a regression problem, it is natural to consider GPs for spectral reconstruction. In this section, we will introduce the last prerequisites for tackling spectral reconstruction problems with GPs and derive the predictive posterior distribution for spectral reconstruction. This will allow the systematic incorporation of prior information about the spectral function, such as asymptotic behavior or normalization information, into the inversion.

The inversion of Fredholm integral equations of the first kind with Gaussian processes has been introduced in [168, 186] and independently developed in [187] and [139, 188, 189]. Apart from the applications presented in this thesis, this method has been applied to several inversion problems in the context of QCD in recent years [190–194].

3.3.1 Gaussian Processes under Linear Transformations

As in the section beforehand, we start by assuming a function $f(x)$, that we can model with a Gaussian process as

$$f(x) \sim \mathcal{GP}(\mu(x), k(x,x'))\,. \tag{3.44}$$

We then define a linear operator $\hat{L}$ that maps f to a new function $g(x)$, i.e. $g(x) = \hat{L}[f](x)$. For example, such linear operators can be derivatives or integrals of the function. Since Gaussian distributions are closed under linear transformations, meaning that Gaussian statistics are preserved under linear transformation of the random variables, the new function $g(x)$ can also be described by a Gaussian process

$$g(x) \sim \mathcal{GP}(\tilde{\mu}(x), \tilde{k}(x,x'))\,, \tag{3.45}$$

with a new mean and a new kernel, that are given by

$$\begin{aligned} \tilde{\mu}(x) &= \hat{L}[\mu](x) \\ \tilde{k}(x,x') &= \hat{L}^2[k](x,x')\,. \end{aligned} \tag{3.46}$$

Where the notation $\hat{L}^2$ for the transformed covariance indicates, that one operator acts on the first argument and the other on the second argument.

A widely used example is the derivative of a Gaussian process [163]. We consider the derivative operator $\hat{L} = \partial_x$ and the above GP of $f(x)$. Then the GP of the derivative of f is given by

$$\partial_x f(x) \sim \mathcal{GP}(\partial_x \mu(x), \partial_x \partial_{x'} k(x, x')) \,. \tag{3.47}$$

With this, we can then condition the posterior on some function value data, but also on the derivative of the function at some points. Additionally, we can directly extract the derivative interpolation of the function from the GP, even when only having information about the function and not its derivatives, including error estimates at every point.

Similarly, the Källén-Lehmann spectral representation (2.19) is a linear transformation of the spectral function[3]

$$G(p_0) = \int_0^\infty \frac{d\omega}{\pi} \frac{\omega}{\omega^2 + p_0^2} \, \rho(\omega) \equiv \int_0^\infty d\omega \; \mathcal{K}(\omega, p_0) \, \rho(\omega) \,, \tag{3.48}$$

where we will call $\mathcal{K}$ the *kernel* of the integral transformation or, more specifically, the *kernel* of the spectral representation. In contrast to the previous section, we will now call the dataset $\mathcal{D} = \{(p_i, G_i) | i = 1, \dots, n\}$, where $G_i = G(p_i)$, since in practice, we have data on the Euclidean correlator at different momenta. As we have done previously, we now assume that the spectral function can be written as a GP, i.e.

$$\rho(\omega) \sim \mathcal{GP}(\mu(\omega), k(\omega, \omega')) \,. \tag{3.49}$$

As already noted in (3.45), the prior for the Euclidean propagator can then be written as a GP with the correct integral insertions in terms of the spectral function GP as

$$G_i \sim \mathcal{N}\left(\int_0^\infty d\omega \, \mathcal{K}(\omega, p_i) \mu(\omega), \int_0^\infty \int_0^\infty d\omega \, d\omega' \, \mathcal{K}(\omega, p_i) k(\omega, \omega') \mathcal{K}(\omega', p_j) \right) \,. \tag{3.50}$$

Then, we can write down the joint prior distribution of the data and the spectral function, while accounting for some Gaussian noise $\mathbf{C}_n$ on the data and a prior mean $\mu(\omega)$, as

$$\begin{pmatrix} \mathbf{G} \\ \rho(\omega) \end{pmatrix} \sim \mathcal{N}\left(\begin{pmatrix} \boldsymbol{\kappa} \\ \mu(\omega) \end{pmatrix}, \begin{pmatrix} \mathbf{K}_{**} + \mathbf{C}_n & \mathbf{K}_*(\omega') \\ \mathbf{K}_*^\top(\omega) & k(\omega, \omega') \end{pmatrix} \right) \,, \tag{3.51}$$

where we have defined

$$\begin{aligned} (\boldsymbol{\kappa})_i &= \int_0^\infty d\omega' \, \mathcal{K}(\omega', p_i) \mu(\omega') \,, \\ (\mathbf{K}_*(\omega))_i &= \int_0^\infty d\omega' \, \mathcal{K}(\omega', p_i) k(\omega', \omega) \,, \\ (\mathbf{K}_{**})_{ij} &= \int_0^\infty \int_0^\infty d\omega' \, d\omega'' \, \mathcal{K}(\omega', p_i) \mathcal{K}(\omega'', p_j) k(\omega', \omega'') \,. \end{aligned} \tag{3.52}$$

---

[3] Here, we present the reconstruction for the spectral function at zero temperature. The exact same is true for the finite temperature spectral representation for the Euclidean propagator in imaginary time (2.21), e.g. in Chapter 6, since this is also a linear transformation of the spectral function.

As already discussed in Section 3.2, we can now directly calculate the posterior predictive distribution for the spectral function $\rho(\omega)$ from the joint prior distribution and arrive at [168, 187]

$$\begin{aligned}\rho(\omega)|\mathbf{G} \sim \mathcal{GP}\big(\mu(\omega) + \mathbf{K}_*^\top(\omega)(\mathbf{K}_{**} + \mathbf{C}_n)^{-1}(\mathbf{G} - \boldsymbol{\kappa}),\\ k(\omega, \omega') - \mathbf{K}_*^\top(\omega)(\mathbf{K}_{**} + \mathbf{C}_n)^{-1}\mathbf{K}_*(\omega')\big)\,. \end{aligned} \tag{3.53}$$

This is the main result of this section. We can encode our knowledge about the underlying spectral function $\rho(\omega)$ in terms of correlator data $\mathbf{G}$. This poster distribution is the distribution over all possible spectral functions that are consistent with the correlator data. The mean then gives the most probable spectral function and the covariance provides the uncertainty of the prediction. When we do not have a clear constraint on the spectral function from the correlator data, we recover the prior distribution of the GP spectral function. In essence, we perform all the different possible interpolations of the spectral function at the same time and obtain the distribution over these fits. As discussed in Section 3.2.3, GPs with *universal* kernels can approximate any continuous function. Consequently, using GPs for spectral reconstruction makes minimal prior assumptions about the spectral function.

Doing spectral reconstruction with GPs shifts the objective from finding the best fit parameters and having a distribution over these fit parameters to having a distribution over the spectral functions themselves. In a sense, this embraces the ill-conditioning of the inversion problem, discussed in Section 3.1, as the posterior distribution contains all possible functions that the model can fit to the given correlator data. However, in practice, the problem of ill-conditioning is not solved by this. We usually obtain wide ranges of possible spectral functions, that are consistent with the data, but do not provide any meaningful predictions, since the deviation from the prior is often minimal. GPs only provide a consistent tool that does not arbitrarily regularize our problem. The remaining challenge is to equip the model with enough prior information to make the prediction meaningful by constraining the possible spectral functions to a physically motivated subset.

Note, that here we take a different approach for deriving the predictive posterior distribution compared to Section 3.2, where we derived the posterior from Bayesian statistics. Here, we have used the well-known results for marginalizing over Gaussian joint distributions [170]. Although this is less intuitive, it allows for easier generalization and the inclusion of additional information in the model. For example, we can assume that we have some direct data on the spectral function $\mathcal{D}_{\text{direct}} = \{(\tilde{\omega}_i, \tilde{\rho}_i)|i = 1, \ldots, m\}$ with some Gaussian noise with covariance $\mathbf{C}_{\text{direct}}$. Additionally, we can also add first derivative data of the spectral function $\mathcal{D}_{\text{deriv}} = \{(\bar{\omega}_i, \bar{\gamma}_i)|i = 1, \ldots, k\}$ with Gaussian noise covariance $\mathbf{C}_{\text{deriv}}$ at some points to complement the data on the correlator. For simplicity, we set the prior mean to zero in the following. We can then include this additional data in the joint prior distribution as

$$\begin{pmatrix} \mathbf{G} \\ \bar{\gamma} \\ \tilde{\boldsymbol{\rho}} \\ \rho(\omega) \end{pmatrix} \sim \mathcal{N}\left(\mathbf{0}\,, \begin{pmatrix} \mathbf{K}_{**} + \mathbf{C}_n & \mathbf{K}_{*\dagger} & \mathbf{K}_{*}(\tilde{\boldsymbol{\omega}}^\top) & \mathbf{K}_{*}(\omega') \\ \mathbf{K}_{*\dagger}^\top & \mathbf{K}_{\dagger\dagger} + \mathbf{C}_{\text{deriv}} & \mathbf{K}_{\dagger}(\tilde{\boldsymbol{\omega}}^\top) & \mathbf{K}_{\dagger}(\omega') \\ \mathbf{K}_{*}^\top(\tilde{\boldsymbol{\omega}}) & \mathbf{K}_{\dagger}^\top(\tilde{\boldsymbol{\omega}}) & k(\tilde{\boldsymbol{\omega}}, \tilde{\boldsymbol{\omega}}^\top) + \mathbf{C}_{\text{direct}} & k(\tilde{\boldsymbol{\omega}}, \omega') \\ \mathbf{K}_{*}^\top(\omega) & \mathbf{K}_{\dagger}^\top(\omega) & k(\omega, \tilde{\boldsymbol{\omega}}^\top) & k(\omega, \omega) \end{pmatrix}\right), \tag{3.54}$$

where a separate (possibly full) covariance matrix is included for all the different data sets. The matrices are defined analogously to the previous case, with the additional data points included and defined as

$$\begin{aligned} (\mathbf{K}_{*\dagger})_{ij} &= \int_0^\infty d\omega'\, \mathcal{K}(\omega', p_i)\, \partial_2\, k(\omega', \bar{\omega}_j)\,, \\ (\mathbf{K}_{\dagger\dagger})_{ij} &= \partial_1 \partial_2\, k(\bar{\omega}_i, \bar{\omega}_j)\,, \\ (\mathbf{K}_{\dagger})_i(\omega) &= \partial_1\, k(\bar{\omega}_i, \omega)\,, \end{aligned} \tag{3.55}$$

where ∂_1 and ∂_2 denote the partial derivatives w.r.t. the first and second argument of the kernel respectively, and when the kernel is evaluated at vectors of data, the result is a vector or a matrix, e.g. $(k(\tilde{\boldsymbol{\omega}}, \tilde{\boldsymbol{\omega}}^\top))_{ij} = k(\tilde{\omega}_i, \tilde{\omega}_j)$. We can now collect the different matrices and vectors of the data covariances and means into a single matrix and vector each and use the same formula as in (3.53) to calculate the posterior distribution of the spectral function. Any linearly connected data can be included in this way and the posterior distribution then gives a consistent estimate of the spectral function that is consistent with all the different constraints. Therefore, it is straightforward to include additional data about the spectral function, such as normalization or the behavior at $\omega = 0$, prior information that is often known from the definition of the spectral function.

3.3.2 Spectral Reconstruction in Higher Dimensions

In general, the spectral reconstruction problem can be a multidimensional problem, since the spectral function can change with temperature T, chemical potential μ_B, or spatial momentum $\mathbf{p}$. The spectral function is a function of multiple variables: $\rho(\omega, T, \mu_B, \mathbf{p})$. For most practical cases, we expect the spectral function to change continuously with these additional parameters. We can use this to our advantage and include additional data points in the Gaussian process regression while demanding that the spectral function is continuous in the additional parameters. This can additionally regularize the problem and therefore improve the reconstruction.

In order to be general, we will introduce the additional dimensions as a vector $\mathbf{X}$, such that $\rho(\omega, \mathbf{X})$. In that case, we have correlator data $(\mathbf{G})_i = G(p_i, \mathbf{X}_i)$. The kernel function now quantifies correlations in a higher dimensional space, i.e. $k(\omega, \mathbf{X};\, \omega', \mathbf{X}')$. Generally, we want to modify the kernel, to respect that the spectral function can have different parameters in different directions. For example, the length scale can differ, or the data can be more sparse in certain directions. Here, we will exemplify this on the squared exponential kernel, where we introduce a length scale for each dimension

$$k(\omega, \mathbf{X}; \omega', \mathbf{X}') = \sigma^2 \exp\left(-\frac{(\omega - \omega')^2}{2\ell_\omega^2} - \sum_{k=1}^{\dim(\mathbf{X})} \frac{(X_k - X_k')^2}{2\ell_{X_k}^2} \right) . \tag{3.56}$$

We can now insert the data $\mathbf{G}$ and the new kernel into (3.51), and calculate the posterior distribution of the spectral function. In order to obtain the different covariances in the joint prior, we only integrate over the first dimension, i.e. ω, while keeping the other dimensions fixed. Note, that we can also include additional data points in the higher dimensional space, e.g. direct spectral function information $\mathcal{D}_{\text{direct}} = \{(\tilde{\omega}_i, \mathbf{X}_i; \tilde{\rho}_i)\}$, or derivatives, to further constrain the spectral function.

3.3.3 Asymptotic Behavior of the Spectral Function

In many cases of spectral reconstruction, the asymptotic behavior in the ultraviolet (UV) or infrared (IR) regime is known. Either perturbation theory or operator product expansions can often provide information about the behavior of the spectral function in the UV. We aim to incorporate this information into the Gaussian process model to improve the reconstruction.

In Section 3.2.1, we have seen that the kernel function of a GP can be seen as the inner product of the basis functions of the RKHS. Since the posterior mean of the GP is a function in the RKHS, we can understand the interpolation as a projection of the data onto the basis functions. In the case of universal kernels, the RKHS is dense in the space of continuous functions; such kernels allow for very flexible interpolation. On the contrary, in certain regions, we want to restrict the possible spectral functions to a subset that is physically motivated.

We therefore introduce separate kernels for different regions of the spectral function. For example, when we have information about the spectral function asymptotics at high or small energies, as $\rho_{\text{IR/UV}}(\omega)$, we can construct the kernel as

$$k_{\text{IR/UV}}(\omega, \omega') \sim \rho_{\text{IR/UV}}(\omega)\rho_{\text{IR/UV}}(\omega') . \tag{3.57}$$

There is usually an additional intermediate region, where we do not have any prior information, and we can use some universal kernel $k_{\text{universal}}$ in this region. Since these kernels dominate different regimes, we can construct the full kernel as a sum of the individual contributions, while defining some function that interpolates between the different regions, e.g. a step function or a sigmoid function. In order to avoid any non-differentiable points in the spectral function, in this work we take the smooth step function as a transition function with the form

$$\theta^{\pm}(\omega; \mu, \ell) = \frac{1}{1 + \exp(\pm 2(\omega - \mu)/\ell)} . \tag{3.58}$$

This function has a midpoint at μ and the steepness of the transition is controlled by ℓ. The full kernel can then be written simply as a sum of the individual contributions

$$k(\omega,\omega') \;=\; k_{\text{intermediate}}(\omega,\omega') + k_{\text{IR}}(\omega,\omega') + k_{\text{UV}}(\omega,\omega')\,, \tag{3.59}$$

where

$$\begin{aligned}
&k_{\text{intermediate}}(\omega,\omega';\sigma,\mu_{\text{IR}},\ell_{\text{IR}},\mu_{\text{UV}},\ell_{\text{UV}})\\
&\quad=\;\theta^{+}_{\text{IR}}(\omega;\mu_{\text{IR}},\ell_{\text{IR}})\,\theta^{+}_{\text{IR}}(\omega';\mu_{\text{IR}},\ell_{\text{IR}})\,\theta^{-}_{\text{UV}}(\omega;\mu_{\text{UV}},\ell_{\text{UV}})\,\theta^{-}_{\text{UV}}(\omega';\mu_{\text{UV}},\ell_{\text{UV}})\,k_{\text{universal}}(\omega,\omega';\sigma)\\
&k_{\text{IR}}(\omega,\omega';\mu_{\text{IR}},\ell_{\text{IR}})\;=\;\theta^{-}_{\text{IR}}(\omega;\mu_{\text{IR}},\ell_{\text{IR}})\,\theta^{-}_{\text{IR}}(\omega';\mu_{\text{IR}},\ell_{\text{IR}})\,\rho_{\text{IR}}(\omega)\,\rho_{\text{IR}}(\omega')\\
&k_{\text{UV}}(\omega,\omega';\mu_{\text{UV}},\ell_{\text{UV}})\;=\;\theta^{+}_{\text{UV}}(\omega;\mu_{\text{UV}},\ell_{\text{UV}})\,\theta^{+}_{\text{UV}}(\omega';\mu_{\text{UV}},\ell_{\text{UV}})\,\rho_{\text{UV}}(\omega)\,\rho_{\text{UV}}(\omega')\,.
\end{aligned}$$

The midpoints of the transition functions $\theta_{\text{IR/UV}}$ are specified by $\mu_{\text{IR/UV}}$ and their steepness is controlled by $\ell_{\text{IR/UV}}$. The set of universal kernel parameters is denoted by σ.

By having a smooth transition function between the different regions, we can ensure that the interpolation is differentiable. Moreover, in practice, the asymptotics do not abruptly start at some point but do gradually dominate the functional form of the spectral function. A smooth spectral function incorporates that and also allows for some deviation in the asymptotic regimes due to the remnant of the universal kernel. Such a construction is able to capture the asymptotic behavior while allowing for some flexibility in order to limit the amount of prior bias introduced by the modifications of the kernel.

Since each of the smooth step functions introduced here has two free parameters, μ and ℓ, we can include these in the hyperparameters of the GP and optimize them along with the other hyperparameters. This allows for the data to determine the exact form of the transition between the different regions. In particular, optimizing the onset of the asymptotic behavior can have interesting physical interpretations and can be determined by comparing the fit quality by minimizing the negative log-likelihood or the χ^2 error of the fit. By finding the distribution of transition parameters, we can also quantify the systematic error that is introduced to the GP by the restriction of the interpolation basis to the asymptotic behavior. This discussion has centered on asymptotic behavior, but in a similar manner, prior knowledge can be incorporated into the Gaussian process through combinations of various kernels and appropriate transition functions. Also, different transition functions can be used for different regions or different directions, as long as the full kernel remains symmetric. Consequently, there is a large flexibility, that can be adjusted to the specific problem at hand, where, for example, the tails of the universal kernel behavior can have algebraic instead of exponential falloff into the tails of the reconstructed spectral function.

3.4 NUMERICAL IMPLEMENTATION: AN INTRODUCTION TO *fredipy*

Motivated by the wide range of applications, not only in QCD but also across other fields of science, where solutions to linear inverse problems are needed, we have developed a Python package based on Gaussian process regression, called *Fredholm inversion in Python: fredipy*. The goal is to invert Fredholm integral equations of the first kind, i.e. equations of the form

$$g(x) = \int dx' \, K(x, x') f(x') \,, \tag{3.60}$$

where K is some known function, and the objective is to find the function $f(x)$ using some data on g. As discussed in Section 3.1, this is a common problem in many fields of science and engineering. The Python package *fredipy* is designed to make the method of Gaussian process regression for solving such problems accessible to a wider audience, while still providing flexible tools in order to address different inverse problems. In the following, we provide an overview of the package, discuss relevant numerical subtleties, and present simple examples demonstrating the application of Gaussian process regression to spectral reconstruction.

Since Gaussian processes are a widely applied method in machine learning and probabilistic modeling, many packages provide tools for Gaussian process regression. For *python*, such packages include *GPy* [195], a native Python implementation, or *GPflow* [196], that is build on the utilities of *TensorFlow* [197], or *GPyTorch* [198], that is build on *PyTorch* [199], or *GPJax* [200], that utilizes *JAX* [201]. Other packages that are not specialized toward Gaussian processes, but general machine learning or Bayesian inference applications often provide some basic GP functionality, such as *scikit-learn* [202], *pystan* [203] or *pymc* [204].

Implementing methods for solving Fredholm integral equations of the first kind with Gaussian processes as outlined in Section 3.3 in these frameworks is not straightforward, since integration operators are not a standard utility and the APIs generally do not provide enough flexibility to include these constraints manually. The package *fredipy* is designed to provide a simple interface to solve this and allow for the inclusion of different constraints. However, in contrast to many of the aforementioned packages, *fredipy* is not designed to be a general-purpose machine learning package but is specialized in (inverse) regression problems. To install *fredipy*, one can use the Python package manager *pip* by running

```
pip install fredipy
```

or directly clone the repository from *GitHub*[4] and install it manually. The only dependencies are the standard Python packages *numpy* [205] and *scipy* [206] in order to perform efficient matrix operations and random sampling.

To introduce the package, we will present some examples of reconstructing Breit-Wigner peaks, which provide a good model for resonances in relativistic high-energy physics. We will invert the zero temperature Källén-Lehmann spectral representation

$$G_{\mathrm{BW}}(p) = \int_0^\infty \frac{d\omega}{\pi} \frac{\omega}{\omega^2 + p^2} \rho_{\mathrm{BW}}(\omega) \,, \tag{3.61}$$

---

[4] `https://github.com/JonasTurnwald/fredipy`

where the Breit-Wigner propagator G_{BW} is given by

$$G_{\mathrm{BW}}(p) \;=\; \frac{A}{m^2 + (p + \Gamma)^2}\,, \tag{3.62}$$

and the associated spectral function ρ_{BW} is analytically known as

$$\rho_{\mathrm{BW}}(\omega) \;=\; \frac{4A\Gamma\omega}{4\Gamma^2\omega^2 + (\Gamma^2 + m^2 - \omega^2)^2}\,. \tag{3.63}$$

We can proceed to generate mock data and use the package to reconstruct the spectral function. The following code snippet shows how to use the package to perform the reconstruction. First, we define our data as a dictionary where we have some noisy data on the correlator (3.62) denoted by `G_err` at momenta `p` and an error estimate `err`, which, in this example, is taken constant for all data.

```python
data = {
    'x': p,
    'y': G_err,
    'cov_y': err**2}
```

This data array always has the same structure and has to be defined for every type of constraint such as derivative or direct constraints separately. It always needs to contain the data points `'x'`, the data `'y'`. The data can also be given as a `list` or `numpy.ndarray` of these data arrays. If no explicit error is given, the error is taken to be zero. The error can also be given as a full covariance matrix if the errors are correlated. In the next steps, we will perform the reconstructing by importing the *fredipy* package, defining the kernel, the type of constraint, and then performing the GP regression with the given data:

```python
import fredipy as fp

# define the kernel
kernel = fp.kernels.RadialBasisFunction(variance=0.3, lengthscale=0.4)

# define the integrator method to use, with upper and lower bounds and number of points
integrator = fp.integrators.Riemann_1D(omega_min=0, omega_max=100, int_n=1000)

# define the operator, the integration with the källen lehmann kernel
integral_op = fp.operators.Integral(kl_kernel, integrator)

# define the full constraint using the data on the propagator
constraints = [fp.constraints.LinearEquality(integral_op, data)]

# now we can define the model using the constraints and the GP kernel
model = fp.models.GaussianProcess(kernel, constraints)

# ... and do a prediction on the points w_pred
_rho, _rho_err = model.predict(w_pred)
```

The `fp.kernels` class provides different kernels that can be used for regression. Here we use the squared exponential or radial basis function kernel (3.25) with some fixed variance and length scale. The Matérn kernels introduced in (3.27) – (3.29) are also predefined in the package. Additional kernels can be manually defined by inheriting from the `fp.kernels.Kernel` class and implementing the `make` method that returns the matrix of the kernel function between two input arrays.

In line 11, we define the integration method to use. The package provides two different integrators, the Riemann sum and the Simpson rule, both in general dimensions as `Riemann` and `Simpson`. In one dimension, we can make significant speed improvements by using the `Riemann_1D` and `Simpson_1D` integrators, which are optimized for this case. We do not use general-purpose integrators like `scipy.integrate.quad` for two reasons. First, we can significantly improve the speed by using a small number of matrix operations, instead of doing a large number of similar integrations over the kernel function. In this way, the overhead of general-purpose integrators can be avoided. Second, but most importantly, using integrators that do not have a fixed set of grid points that are used for all integrations, can lead to inconsistencies in the GP regression. Since the inversion of the kernel matrix can have a large condition number, the resulting difference due to integration errors does not cancel exactly and the GP regression can become unstable. This can result in isolated spikes in the regression; however, such artifacts can be entirely avoided by using the exact same integration routine consistently across all steps of the reconstruction. Note, that we perform the numerical integration up to an upper limit of 100 for the examples shown in this section. This is a reasonable choice, since the integrand decreases with $1/\omega^4$ for large ω, and the integral converges quickly. For other applications, it might be useful to transform the integration variables, e.g. $\tilde{\omega} = \frac{\omega}{1-\omega}$, in order to have a finite domain of integration. The usefulness of such transformations heavily depends on the specific problem and must be implemented by the user.

Next, we define the operator that connects the data to the interpolated function in line 14. In this case, the spectral function is connected to the correlator data via integration over the Källén-Lehmann kernel (3.61), here denoted by `kl_kernel`. This kernel is defined by the user and must be a function that takes the integration variable as the last argument and the data point as the first argument. In our case, the kernel is defined as

```python
def kl_kernel(p, omega):
    return omega / (omega**2 + p**2) / np.pi
```

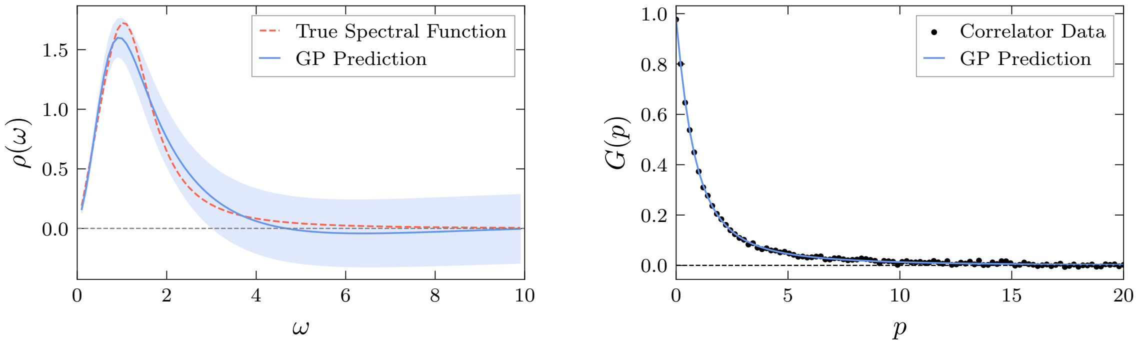
The kernel has to take a `numpy.ndarray` as input and return the same type. We then define the operator by passing the kernel and the integrator to the `fp.operators.Integral` class. There are different operators that are currently implemented, such as the derivative operator, the integral operator, or the direct operator, that can be used to include different types of constraints on the regression. As a first introduction, we only use integral constraints, but we will extend this in the course of the section.

We can now define an array of different constraints. Here we only have one constraint, the integral constraint, but we can extend the list to incorporate different data that might be available. Note that we additionally wrap the operator in a `fp.constraints.LinearEquality` class. This means, that we have data that is linearly connected to the spectral function and that comes in the form of equality constraints. In the published version of this package, no other type constraints are available, but as described in Section 3.5, inequality constraints are an important extension. They are currently not available in the official branch of the package, since it remains a challenge to implement them in a way that consistently provides meaningful results.

The final GP model is then built by passing the kernel and the constraints to the predefined `fp.models.GaussianProcess` class. The kernel of the Gaussian process must be in the form of a `fp.kernels.Kernel` object, while the constraints must be a list or an array of `fp.constraints.Constraint` objects. We can make predictions on the spectral function at some points `w_pred` by calling the `predict` method of the model. The result is the mean and the standard deviation of the spectral function at the given points. Again, the explicit computation is performed by Cholesky decomposition of the kernel matrix, as described in Appendix B.2. With the keyword argument `full_cov=True`, the full covariance matrix of the prediction can be obtained. We show the result of such a reconstruction in Figure 3.5a with some mock data shown in Figure 3.5b. The reconstruction is done with the squared exponential kernel and the integral operator, as described above. We see that the reconstruction of such a simple spectral function is rather accurate, however, the propagator data does not constrain the large ω regime of the spectral function. In this region, we obtain the GP prior and consequently do not learn much about the spectral function. This is a common problem in spectral reconstruction, where the data is insufficient to constrain the spectral function in all regions. Including additional constraints can help to improve the reconstruction, as we will see in the following examples. In order to make sure that the reconstruction actually reproduces the data, we can also use *fredipy* to predict the data points that we have used for the reconstruction. This is done by calling the `predict_data` method of the model

```
_G, _G_err = model.predict_data()
```

where we can again use the keyword argument `full_cov=True` to obtain the full covariance matrix of the prediction. The resulting prediction is shown in Figure 3.5b, and the data points are well reproduced by the GP model, confirming that the reconstruction is consistent with the data. In the following, we will subsequently improve the reconstruction by adding additional constraints to the model and show the effects of these constraints on the reconstruction.



(a) Reconstruction of a Breit-Wigner peak.

(b) Interpolation of the correlator data.

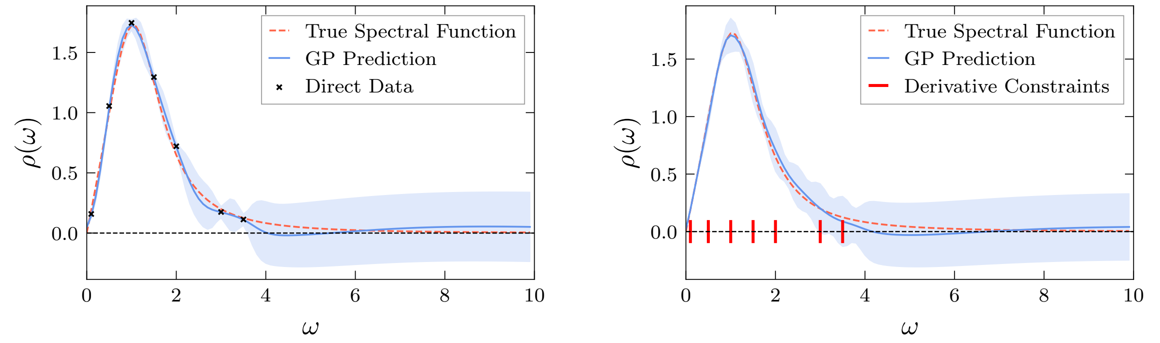
Figure 3.5: Reconstruction of a Breit-Wigner peak using *fredipy* with parameters $A = 1.6,\ m = 1$ and $\Gamma = 0.8$. In (a), we show the reconstruction of the spectral function as a solid line and the variance of the prediction as the shaded area. The reconstruction compares well with the true spectral function, denoted by the dashed line. In (b), we show the associated interpolation of the correlator data. The variance of the correlator reconstruction is too small to be visible.

3.4.1 Additional Constraints in *fredipy*

As mentioned before, we can include a wide variety of additional constraints that are linearly connected to the underlying spectral function in the reconstruction. There are currently three different types of linear equality constraints that can be implemented in *fredipy*: direct constraints, derivative constraints, and integral constraints, which were already introduced. Direct constraints are constraints on the spectral function itself, i.e. data on the spectral function, while derivative constraints are constraints on the derivative of the spectral function. For both types of constraints, we have to define a data dictionary, as we have done for the correlator data in line 1. Here, we construct them as `data_direct` and `data_deriv` respectively. We can then include these constraints in the list of previous constraints that we pass to the model as

```
24  identity_op = fp.operators.Identity()
25  derivative_op = fp.operators.Derivative()
26
27  constraints = [fp.constraints.LinearEquality(integral_op, data),
28                 fp.constraints.LinearEquality(identity_op, data_direct),
29                 fp.constraints.LinearEquality(derivative_op, data_deriv)]
```

Here we define the identity operator `identity_op` that is used for direct constraints since they are connected to the spectral function via the identity, and the derivative operator `derivative_op` that is used for derivative constraints. Both of these operators are predefined in the package and can be used directly. The derivative operator calls the derivative of the GP kernel function as introduced in (3.55). Therefore, the derivative of the kernel function must



(a) Reconstruction with additional direct constraints.

(b) Reconstruction with additional derivative constraints.

Figure 3.6: Reconstruction of a Breit-Wigner peak using *fredipy* with the same correlator data used in the reconstruction presented in Figure 3.5. In (a), we have added direct data on the spectral function with some Gaussian noise, denoted by the black crosses. In (b), we have added derivative data of the spectral function with some Gaussian noise. The position in ω of these data points is denoted by the red lines on the ω-axis. For both examples, we can observe that the additional data points significantly improve the reconstruction in the regions where they are added while having an insignificant impact on the rest of the spectral function.

be implemented in the `fp.kernels.Kernel` class. At the moment, this is only implemented for the squared exponential kernel but can be easily extended to other kernels, by defining the `dK_dx`, `dK_dy` and `dK_dxdy` methods in the kernel class. The full joint prior covariance matrix is then built as described in (3.54) and the posterior distribution is calculated as before. The `constraints` list can be extended as needed, and the constraints can be combined in any way, as long as the type of constraint is implemented. One can also add multiple constraints of the same type, e.g. when having multiple separate data sets of the same type. We continue with the Breit-Wigner example from the section before and show the effects of the reconstruction with additional direct and derivative constraints in Figure 3.6.

3.4.2 ASYMPTOTIC BEHAVIOR OF THE SPECTRAL FUNCTION IN *fredipy*

Another important extension of the GP model is the inclusion of asymptotics of the spectral function in the model. This can be done by modifying the kernel function as described in Section 3.3.3. This is implemented in *fredipy* by defining a kernel that includes the asymptotic behavior of the spectral function and then performing the reconstruction.

```python
# define the functional form of the asymptotics
def uv_asymptotics(w):
    return 1 / w**3

# define the kernel function of the non-asymptotic regimes
```

```
6  kernel = fp.kernels.RadialBasisFunction(variance=0.3, lengthscale=0.3)
7
8  # define the asymptotic kernel
9  asymptotic_kernel = fp.kernels.AsymptoticKernel(kernel)
10
11 # add the known asymptotic form to the kernel
12 asymptotic_kernel.add_asymptotics(region='UV', asymptotics=uv_asymptotics)
13 asymptotic_kernel.set_params([1.0, 3.0])
14
15 # now we can define the model using the constraints from before and additionally
16 # the asymptotic GP kernel
17 model = fp.models.GaussianProcess(asymptotic_kernel, constraints)
```

In this example, the asymptotic behavior of the Breit-Wigner spectral function is $1/\omega^3$ in the UV. We define a function that returns the asymptotic behavior for a given ω in line 2. The kernel is then defined by passing the non-asymptotic kernel (in this example, the squared exponential kernel) to the `fp.kernels.AsymptoticKernel` class. We can then add the asymptotic behavior to the kernel by calling the `add_asymptotics` method. The keyword argument `region` specifies the region of the asymptotics, which can be either `'UV'` or `'IR'` and the keyword argument `asymptotics` specifies the functional form. The implementation of the asymptotics is done via smooth step functions and follows the description in Section 3.3.3 exactly. In line 13, we set the parameters of the step functions, i.e. $\mu = 3.0$ and $\ell = 1.0$ in (3.58). We can then perform the reconstruction as before by using the previously defined `constraints` and the `asymptotic_kernel` in order to define the GP model and perform the prediction at some test points with the `predict` method. In Figure 3.7, we show the result of the reconstruction with the asymptotic kernel. As expected, the inclusion of the asymptotic behavior significantly improves the reconstruction in the asymptotic regime. However, one has to be careful with interpreting the error band in this case. Whereas the GP model with a universal kernel was extremely flexible and introduced a minimal prior bias, the asymptotic kernel introduces a strong bias in the regimes where it applies. Therefore, the statistical error that is shown here, can be misleading, as its model dependence is far greater and a significant systematic error can be introduced. To achieve a more reliable error estimate, one would have to perform a more detailed analysis by varying the parameters of the asymptotic kernel and, for example, average over the resulting spectral functions. We will discuss the optimization of the asymptotic parameters in more detail when we discuss applications in Chapters 5 and 6.

3.4.3 Higher Dimensional Reconstruction in *fredipy*

As discussed in Section 3.3.2, GPs are easily extended to interpolation in more than one dimension. In *fredipy* we have implemented such an extension by allowing for additional dimensions in the data and the kernel function. As a practical example, we can simply modify the Breit-Wigner example from before to include an additional dimension k in the

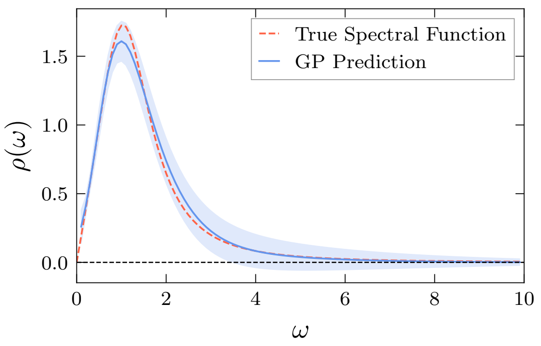


Figure 3.7: Reconstruction of a Breit-Wigner peak using *fredipy* with the same correlator data used in the reconstruction presented in Figure 3.5. We have added the asymptotic behavior of the spectral function $\sim 1/\omega^3$ in the UV regime to the kernel. The reconstruction is highly improved in the UV regime, where otherwise the correlator data does not provide much information.

spectral function and the correlator as

$$G_{\text{mod}}(p,k) = (1+k^2)G_{\text{BW}}(p), \quad \rho_{\text{mod}}(\omega,k) = (1+k^2)\rho_{\text{BW}}(\omega)\,. \tag{3.64}$$

The code for constructing the appropriate arrays and performing the reconstruction is presented in Appendix B.3. In essence, the only modification that has to be made is to define the data and test arrays in the appropriate dimensions. Additionally, the kernel function can now have multiple length scales, one for each dimension. The reconstruction of the modified spectral function ρ_{mod} is then performed as before, and the result is shown in Figure 3.8. In general, the inclusion of additional dimensions improves the posterior error of the reconstruction, as expected, since we have more data that constrains the spectral function. We emphasize, that although we only show the reconstruction of the spectral function in two dimensions, the package is designed to be able to handle any number of additional dimensions.

Currently, the only implementation is for additional dimensions like the modification in (3.64), i.e. additional dimensions in the argument of the correlator and the spectral functions. Other types of additional dimensions, such as modifications to the kernel function, are not implemented in the package at the moment. This can become important when considering the reconstruction of spectral functions with an additional temperature direction, since both the Källén-Lehmann kernel and the spectral function become dependent on the temperature, see (2.21). Such a modification provides no conceptual challenge, and we will implement this when the need arises.

3.4.4 Hyperparameter Optimization in *fredipy*

As discussed in Section 3.2.2, the optimization of the hyperparameters of the GP model is an important step in the reconstruction process. However, optimizing the kernel hyperparame-

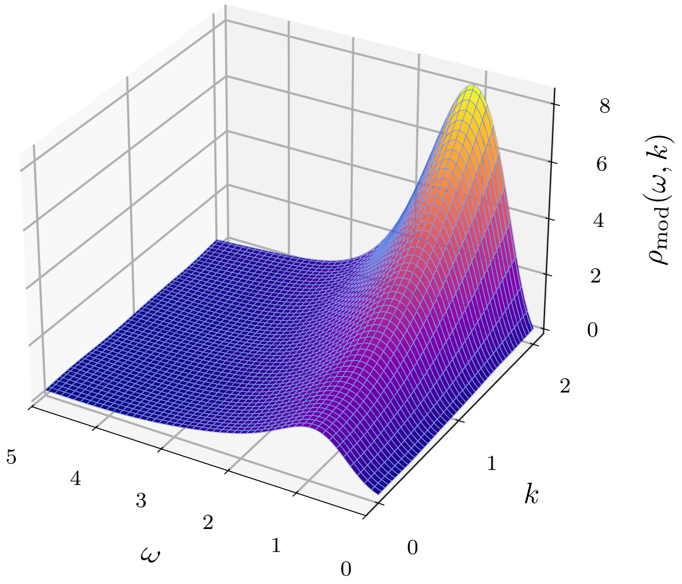


Figure 3.8: Reconstruction of the modified Breit-Wigner spectral function (3.64) using *fredipy* with the same parameters as the examples above in Figures 3.5−3.7. The reconstruction is performed in two dimensions, and we recover the examples above at $k = 0$. For visual clarity, we do not show the reconstruction error here, but it is of the order of the other examples, albeit smaller than the reconstruction error of the one-dimensional example without additional derivative or direct constraints.

ters is a highly problem-specific task and different methods can lead to significantly different results. In *fredipy*, we therefore do not provide a general optimization routine, since the results can be highly misleading when not carefully interpreted. Instead, we provide users with the tools to perform the optimization in a manner tailored to their specific problem. The `fp.models.GaussianProcess` class provides the `log_likelihood` method that returns the logarithm of the marginal likelihood of the data given the model, given in (3.31). We can maximize this likelihood (or minimize the negative log-likelihood) in order to find the optimal hyperparameters of the model. For a small number of hyperparameters, e.g. two parameters in case of the standard squared exponential kernel, we can use simple grid scans to find the optimal hyperparameters, see Figure 3.4b. The complexity grows quickly when introducing additional dimensions with separate parameters, as in Sections 3.3.2 and 3.4.3 or when including asymptotic behavior, as in Sections 3.3.3 and 3.4.2. In these cases, the optimization routines have to find the minimum in high-dimensional parameter space and it becomes more likely for local minima in the negative log-likelihood to develop. In Appendix B.4, we have provided a simple example of how to perform the optimization of the hyperparameters of the kernel function in *fredipy* using the `scipy.optimize.minimize` routine.

Since we update the parameters of the model after each optimization step using the `set_params` method, we need to compute the inverse of the kernel matrix after each step. As discussed in Section 3.2.2 this can be computationally expensive, especially for large data sets, since the inversion of the kernel matrix scales as $\mathcal{O}(N^3)$, where N is the number of data points. One method to speed up the optimization is to use the gradient of the log-likelihood, defined in (3.32), which is given by the `log_likelihood_grad` method of the model. Computing this gradient makes use of the precomputed inverse of the kernel matrix and only scales as $\mathcal{O}(N^2)$ per hyperparameter. It is therefore highly recommended to use gradient-based optimization routines.

In the chapters, where Gaussian processes are applied to spectral reconstruction problems, cf. Chapters 4 – 6, we will apply and discuss different methods of hyperparameter optimization. However, this is not conclusive, as these optimization routines can still be improved and are in general highly problem-specific. There exist many other, very sophisticated and useful methods for optimizing the hyperparameters of a GP model. A comprehensive overview of these methods, along with their respective advantages and disadvantages, is beyond the scope of this work.

3.5 INEQUALITY CONSTRAINTS

One important extension of the GP model is the inclusion of inequality constraints, meaning that the reconstructed spectral function should be constrained by inequalities of the form

$$a(\omega) \leq \rho(\omega) \leq b(\omega)\,, \tag{3.65}$$

where a and b are some known functions. In many applications, especially in the context of spectral reconstruction in high-energy physics, many of the reconstructed spectral functions are positive definite, i.e. $\rho(\omega) \geq 0$. This is the case for all spectral functions associated with a physical degree of freedom, such as the spectral function of the pion or the vector current correlator in QCD. However, demanding such constraints on a GP model that are not equality constraints does not come naturally. Gaussian processes excel when providing linearly connected data points, but inequality information is a non-linear constraint leading to many challenges when implementing this into the current framework.

We have implemented two different methods that attempt to include inequality constraints in the GP model. The first method is to introduce *virtual observations*. It was first introduced in [207] in the context of standard GP regression with monotonicity constraints. Another method is to utilize non-Gaussian processes and reconstruct the logarithm of the spectral function, or the square of the spectral function, and then transform the result back after the reconstruction. This results in a non-linear reconstruction problem. We will discuss both methods in the following and show the results of the reconstruction of a simple spectral function with inequality constraints.

3.5.1 Linear Inequality Constraints with Virtual Observations

The method of virtual observations provides a very flexible framework to include inequality constraints in the GP model. The idea provided in [207] is as follows: When having inequality constraints of the form (3.65), we assume that these constraints do not necessarily have to hold for all ω, but only for a finite set of points ω_i^v for $i = 1, \dots, N$, called *virtual observation locations*. For every virtual observation location we introduce $c(\omega_i^v) = \rho(\omega_i^v) + \varepsilon_i^v$, a finite width on the spectral function at these locations, with $\varepsilon_i^v \sim \mathcal{N}\left(\mathbf{0},\, \sigma_v^2 \mathbb{1}\right)$. Then, we define the event, that the constraint $a(\omega) \leq c(\omega) \leq b(\omega)$ is satisfied at all virtual observation locations $\boldsymbol{\omega^v} = (\omega_1^v, \dots, \omega_N^v)$, denoted as

$$C(\boldsymbol{\omega}^v) \,=\, \cap_{i=1}^N \left\{a(\omega_i^v) \leq c(\omega_i^v) \leq b(\omega_i^v)\right\} \,. \tag{3.66}$$

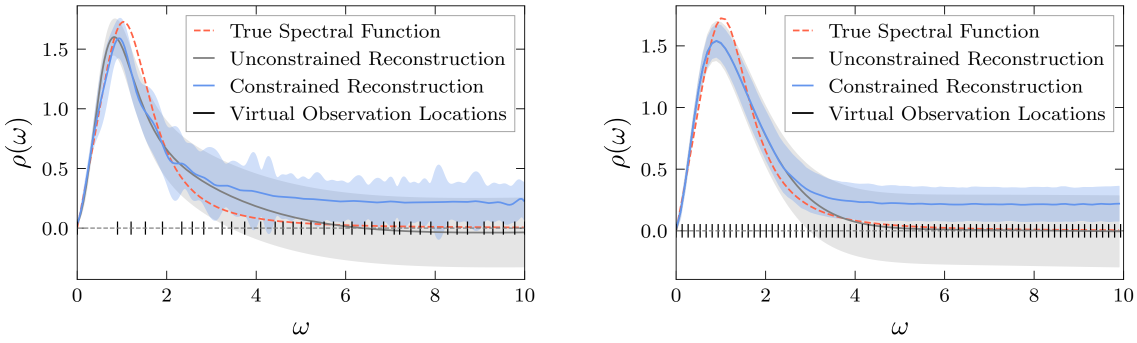
We therefore have to find the posterior distribution of the GP model conditioned not only on the propagator data $\mathbf{G}$, but also conditioned on the event $C(\boldsymbol{\omega}^v)$ as

$$p(\rho(\omega) \,|\, \mathbf{G},\, C(\boldsymbol{\omega}^v)) \,=\, \mathcal{N}\left(\mu(\tilde{C}),\, \Sigma\right) \,. \tag{3.67}$$

Now, the posterior mean depends on the additional constraint, that the inequality constraints are satisfied at the virtual observation locations, i.e. $\tilde{C} = c \,|\, \mathbf{G},\, C(\boldsymbol{\omega}^v)$. The distribution $c \,|\, \mathbf{G}$ is still a Gaussian distribution since the GP model is linear in the data points. The nonlinearity is then introduced by the inequality constraints: $c \,|\, \mathbf{G}$ is conditioned to exclusively lie on the hyperrectangle defined by the inequality constraints at the virtual observation locations. Therefore, the result is a truncated Gaussian

$$p(c \,|\, \mathbf{G},\, C(\boldsymbol{\omega}^v)) \,=\, \mathcal{TN}\left(\mu^*(\omega^v),\, \Sigma^*(\omega^v, \omega^{'v}), a(\omega^v), b(\omega^v)\right) \,, \tag{3.68}$$

where μ^* and Σ^* are standard posterior mean and covariance, e.g. as described in (3.20), (3.21) and (3.53), and $\mathcal{TN}(\cdot, \cdot, a, b)$ denotes the truncated Gaussian distribution to the interval $[a, b]$. The posterior predictive distribution (3.67) of such a constrained GP model is not a purely Gaussian distribution anymore, but a *compound* or *mixture* Gaussian distribution; its mean depends on a truncated Gaussian. The explicit derivation of the expressions for the mean and covariance of the posterior distribution is, in principle, very similar to the derivation of the marginal likelihood with additional constraints from a joint prior distribution described in Section 3.3.1. However, since the resulting distribution is not strictly Gaussian anymore, the evaluation of the posterior can not be done analytically anymore. We therefore sample the posterior of the constraint distribution (3.68) and then sample the posterior (3.67) using these samples in the mean of the GP model. The full expressions for the posterior distribution can be found in [207], Lemma 1 & 2.



(a) Optimized VO

(b) Non-optimized VO

Figure 3.9: Reconstruction of the Breit-Wigner spectral function with the same data as in previous examples, see Figure 3.5b. In both reconstructions, virtual observation locations are added to the model in order to achieve local positivity, where the location of these virtual observations is indicated by vertical dashes on the ω-axes. The results are compared to both, the true spectral function and the unconstrained reconstruction. In (a), the virtual observation locations are optimized to a target probability of satisfying the constraint with $p_{\text{target}} = 0.8$. With these optimized virtual observation locations, a positive spectral function is achieved, while adding significant numerical and systematic noise to the reconstruction. For a larger number of virtual observation locations, in (b), a smoother reconstruction is achieved, while adding a significantly higher numerical cost to the computation while avoiding the optimization procedure. The propagator data compared to the integral on the constrained spectral functions shows significant deviation, as discussed in the main text.

Up to now, we have assumed some fixed virtual observation locations $\boldsymbol{\omega}^v$. However, the choice of these locations is crucial for the regression or reconstruction problem. It is not possible to have infinitely dense virtual observation locations, since the computational cost of the GP model scales as $\mathcal{O}(N^3)$, and the virtual observations add to these costs. Following [207], we implement an optimization procedure to find the optimal virtual observation locations: We compute the probability of the spectral function to satisfy the inequality constraints as

$$p_c(\omega) \;=\; p\left(a(\omega) \leq (\rho(\omega)\,|\,\mathbf{G},\, C(\boldsymbol{\omega}^v)) \leq b(\omega)\right)\,. \tag{3.69}$$

We then find the minimum of the probability $p_c(\omega)$ for the given inequality constraints. When the minimum is below a certain target probability p_{target}, we add a new virtual observation at this location. This procedure is repeated until the minimum of the probability p_c is above the target probability.

In Figure 3.9, we show the result of the BW reconstruction example from the previous section, with the inclusion of the introduced virtual observations. We can see, that the reconstructed spectral function is now positive. The optimization procedure introduces a significant numerical cost to the reconstruction since the probability of the spectral function to satisfy the inequality constraints is computed at every step of the optimization, which has to be performed at every addition of a virtual observation. The noise on the reconstruction with the optimized virtual observation locations in Figure 3.9a is due to the irregularity in the positions of the virtual observation locations. This can be avoided by using a fixed set of virtual observation locations, as shown in Figure 3.9b, which also circumvents the optimization procedure. However, there are two major drawbacks to this method: We must select a finite interval for the virtual observation locations; in the example above, this interval was chosen as $[0, 10]$. However, the positivity constraint in this example and in many practical spectral reconstruction applications is a global constraint. Especially when reconstructing a spectral function with an asymptotic approach to zero, as in the example, a global optimization of virtual observations will not converge. Otherwise, adding a large number of equally spaced virtual observation locations adds significant numerical cost to the reconstruction, and only works on finite intervals. A second, and even more drastic, drawback of this method is the observation, that the reconstructed spectral function is not guaranteed to satisfy the inequality constraints at all points, i.e. the data is not as well reproduced by the integrated mean of the posterior spectral functions in Figure 3.9, compared to the unconstrained reconstruction. This is due to the fact, that this method is essentially a resampling of the GP model under the constraint of positivity, or in practice under the constraint of positivity at the virtual observation locations. Naively, we could take the unconstrained GP model, take samples, and only accept those samples that satisfy the inequality constraints. This, however, is highly inefficient when acceptance rates become small, and the method of virtual observations provides a computationally efficient way to sample such a distribution. In the above example, we mostly sample the edges of the original distribution, since positive spectral functions are not very common in the unconstrained GP model. Much more common are spectral functions that oscillate around zero for large ω. Therefore, the resulting mean of the constrained posterior distribution lies further away from the input data, compared to the unconstrained GP model. One reason the constrained model's posterior mean fails to reproduce the data accurately is the assumption that the spectral function follows a Gaussian model. For a model of a spectral function to approach zero while being positive, intuitively, we would expect other distributions to describe the spectral function better, that are able to have zero, or close to zero mean values, while random samples only lie above the positive axis. Such distributions then have to be non-Gaussian, as will be discussed next.

3.5.2 Reconstruction with Non-Gaussian Processes

An alternative method to include inequality constraints in GP models is the use of non-Gaussian processes. Here, we will restrict ourselves to the case of positivity constraints, i.e. $\rho(\omega) \geq 0$, but this can be generalized to any constant inequality constraint. This method can therefore not cover all cases that can be considered from (3.65) and the method utilizing virtual observations. However, constant inequality constraints, and especially positivity constraints, are the most important constraints in the practical application of spectral reconstruction in high-energy physics-related problems.

In order to achieve a non-negative spectral function, we can reconstruct a transformed spectral function $\varrho(\omega)$, which is not constrained to be positive. When inverting the transformation, we can then obtain the positive spectral function $\rho(\omega)$, i.e.

$$\rho(\omega) = f(\varrho(\omega)) \quad \text{with} \quad f : \mathbb{R} \to \mathbb{R}^+ \,. \tag{3.70}$$

The transformation f is chosen such that the positivity constraint on ρ is satisfied, while the function ϱ does not have to satisfy any additional constraints. Possible choices for the transformation are e.g.

$$f(\varrho) = \exp(\varrho) \quad \text{or} \quad f(\varrho) = \varrho^2 \,. \tag{3.71}$$

Such transformations have already been used for standard GP regression, in order to produce non-negative predictions, log-Gaussian processes e.g. in [208] and squared GPs in [209]. Recently, the special case of the squared spectral function has also been discussed in [194] in the context of the reconstruction of PDFs. This transformation, however, introduces new difficulties: The reconstruction problem becomes non-linear, since such transformations are, as expected for inequality constraints, non-linear. In the case of the exponential transformation, the posterior distribution of the GP model is not Gaussian anymore, but a log-Gaussian distribution. However, the logarithm maps spectral function values approaching zero to negative infinity. This can lead to numerical instabilities in the sampling of the GP model. Since many physical spectral functions have either explicit zeros, e.g. $\rho(0) = 0$, or asymptotically approach zero for large ω, the logarithm transformation is generally not suitable for spectral reconstruction.

Reconstructing the square root of the spectral function, i.e. $\varrho = \sqrt{\rho}$, however, is a more suitable transformation for the problem at hand. We define the transformed spectral function to be described by a Gaussian processes

$$\varrho(\omega) \sim \mathcal{GP}\left(\mu(\omega),\, k(\omega, \omega')\right) \,. \tag{3.72}$$

The objective is then to obtain the posterior distribution of the transformed spectral function, conditioned on some propagator data $\mathbf{G}_{\text{data}}$ and some noise covariance matrix $\mathbf{C}_d$, as described in Sections 3.2 and 3.3. From Bayes' theorem (3.7), the posterior distribution of the transformed spectral function is given by

$$p(\boldsymbol{\varrho} \,|\, \mathbf{G}) \sim p(\mathbf{G} \,|\, \boldsymbol{\varrho})\, p(\boldsymbol{\varrho}) \,. \tag{3.73}$$

The prior $p(\varrho)$ is still a Gaussian distribution, by definition of the GP model (3.72), and only depends on the model hyperparameters. However, the likelihood $p(\mathbf{G} \,|\, \boldsymbol{\varrho})$ is now a non-Gaussian distribution of the transformed spectral function. Since the propagator data is assumed to have Gaussian noise the integral of the spectral function has to obey this distribution and is therefore given by a Gaussian distribution in terms of the propagator data as

$$p(\mathbf{G}\,|\,\boldsymbol{\varrho}) \sim \exp\left(-\frac{1}{2}(\mathbf{G}_* - \mathbf{G}_{\text{data}})^\top \mathbf{C}_d^{-1}(\mathbf{G}_* - \mathbf{G}_{\text{data}})\right), \tag{3.74}$$

where we have defined

$$\begin{aligned}(\boldsymbol{\varrho})_i &= \varrho(\omega_i)\\ (\mathbf{G}_*)_i &= \int d\omega\, \mathcal{K}(\omega, p_i)\, f(\varrho(\omega))\,.\end{aligned} \tag{3.75}$$

The likelihood has therefore a non-Gaussian contribution $\sim \exp(-f(\boldsymbol{\varrho})\mathbf{C}_d^{-1} f(\boldsymbol{\varrho}))$. The full posterior distribution of the transformed spectral function is then given, similar to the derivation for the usual GP regression (3.11), by

$$p(\boldsymbol{\varrho}\,|\,\mathbf{G}) \sim \exp\left(-\frac{1}{2}\boldsymbol{\varrho}^\top \mathbf{K}^{-1}\boldsymbol{\varrho} - \frac{1}{2}(\mathbf{G}_* - \mathbf{G}_{\text{data}})^\top \mathbf{C}_d^{-1}(\mathbf{G}_* - \mathbf{G}_{\text{data}})\right), \tag{3.76}$$

where $(\mathbf{K})_{ij} = k(\omega_i, \omega_j)$ is the kernel matrix of the transformed spectral function. Since the resulting posterior distribution is non-Gaussian, it is not possible to evaluate the posterior mean and covariance analytically. As a result, we must resort to numerical methods to sample the posterior distribution. The computation of the posterior mean becomes a $\dim(\boldsymbol{\varrho})$-dimensional optimization problem, i.e. the number of points in the evaluated spectral function. This becomes computationally expensive, especially since the likelihood depends on an integral over the function $f(\varrho)$. Therefore, we need to evaluate the function $f(\varrho)$ on enough points ω to obtain a good approximation of the integral, and $\dim(\boldsymbol{\varrho})$ is typically of the order of 10^2 to 10^3.

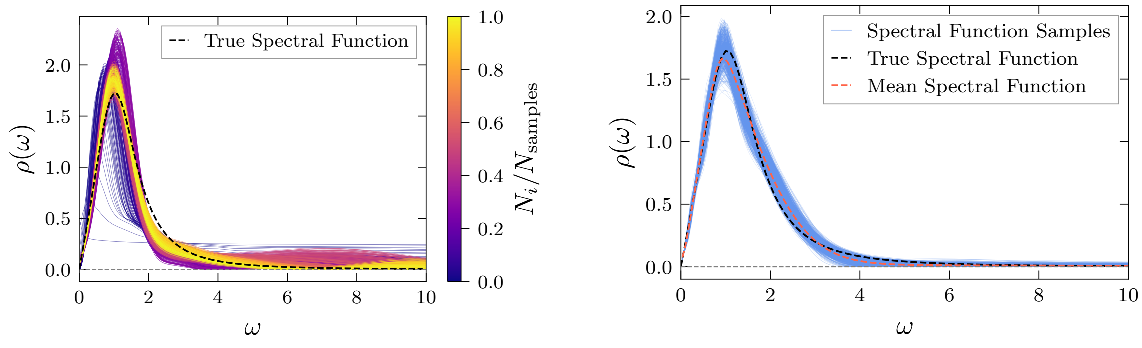
We sample from this distribution using a custom Hamiltonian Monte Carlo (HMC) implementation. We perform the optimization on the level of the transformed spectral function, and then transform the result back to the spectral function using the inverse of the transformation f^{-1}. The action of the HMC procedure is given by the negative log-posterior distribution (3.76), so we have $p(\boldsymbol{\varrho}\,|\,\mathbf{G}) \sim \exp(-S(\boldsymbol{\varrho}))$. Since HMC is a gradient-based method, we have to compute the gradient of the action, which is given by

$$\nabla_{\boldsymbol{\varrho}} S(\boldsymbol{\varrho}) = \mathbf{K}^{-1}\,\boldsymbol{\varrho} + 2\,\tilde{\mathbf{G}}_*\,\mathbf{C}_d^{-1}(\mathbf{G}_* - \mathbf{G}_{\text{data}})\,, \tag{3.77}$$

where

$$(\tilde{\mathbf{G}}_*)_{ij} = \mathcal{K}(\omega_i, p_j)\varrho(\omega_i)\,. \tag{3.78}$$

In other common HMC implementations, the initial momentum is drawn from a Gaussian distribution at each step of the HMC sampling procedure. However, the Gaussian process prior (3.72) demands that the resulting transformed spectral function is smooth, and therefore



(a) HMC burn-in

(b) HMC samples

Figure 3.10: Reconstruction of the Breit-Wigner spectral function from previous examples with the non-Gaussian process method. The spectral function is reconstructed by sampling the posterior distribution (3.76) with HMC as described in the text. In (a), we show the burn-in phase of the HMC sampling, where the samples are not yet converged to the posterior distribution. The burn-in phase is shown for the first $N_{\text{samples}} = 10000$ samples and the samples are color-coded by their iteration number. We initialize the spectral function relatively far away from its true value with $\varrho_0(\omega) = 0.5$ and still achieve convergence to the true spectral function. In (b), we show $N_{\text{samples}} = 10000$ samples after the burn-in phase. Again, the samples agree with the true spectral function. Particularly in the large ω region, where the error of the reconstruction was usually large, we achieve excellent agreement with the true spectral function.

the momentum has to be drawn from a distribution that respects this smoothness. Otherwise, neighboring points in the spectral function become very different, e.g. when the initial momentum has different signs for neighboring points ω_i and ω_{i+1}, and the acceptance rate of such an HMC sampling is usually close to zero, as the GP prior excludes such solutions. We therefore draw the initial momentum from a multivariate Gaussian distribution with a covariance matrix given by the squared exponential kernel.

In Figure 3.10, we show the result of the reconstruction of the BW spectral function using the non-Gaussian process method. We can see that the samples of the posterior distribution agree very well with the true spectral function, and the resulting spectral function has the correct, positive behavior. The agreement of the correlator data with the integral of the spectral function is also within errors, in contrast to the reconstruction with the virtual observation method. Even when starting away from the true spectral function, the HMC sampling can converge to the true spectral function, as shown in Figure 3.10a. It is important to note that this is not always the case, as the posterior distribution often contains local minima where the HMC sampling can become trapped and fail to converge to the true spectral function. For illustrative purposes, we have selected an example where the method works well and the optimization remains relatively stable, though this stability cannot be guaranteed in general. HMC parameters, such as the step size and the number of leapfrog steps, must be carefully tuned for each problem, and some initial guesses for the spectral function may prevent convergence altogether. This challenge is common to many optimization problems but can be mitigated by employing more robust optimization techniques. Additionally, a significant issue arises due to the symmetry introduced when taking the square root of the spectral function, complicating the sampling process further. We therefore have, by construction, two degenerate minima at $\varrho = \pm\sqrt{\rho}$. This produces local minima, where the transformed spectral function ϱ has positive and negative regions. The resulting zero-crossing introduces a kink in the spectral function, which is not physical. Depending on the sampling parameters, this can lead to a large number of such non-physical samples and distort the final result. Again, in Figure 3.10, we have chosen an example where the sampling has worked out well, but this is not guaranteed in general.

Due to the aforementioned subtleties, we have not yet published the associated code in the context of the *fredipy* package, however, this approach is very promising, and we will continue to work on this method in the future. The primary goal is to improve the optimization algorithm and prevent the doubling problem in the transformed spectral function.

3.6 CONCLUSION

In this chapter, we have discussed the spectral reconstruction problem and presented a method that offers a consistent approach to this problem. Since many methods have significant problems of either providing realistic systematic error estimates or are not able to incorporate the full range of available prior information, the need for further improvements in the spectral reconstruction problem is evident. To address some of these issues, we have presented a method based on Gaussian processes that consistently incorporates the available, and physically motivated prior information and enables more reliable quantification of systematic errors. We have reviewed the central properties of Gaussian processes and have significantly expanded on the method of using Gaussian processes for spectral reconstruction.

We have presented the implementation of the Gaussian process method in the *fredipy* package, which is a Python package for spectral reconstruction and analysis of spectral functions. This package provides utilities in order to account for the different prior information, that is usually available in the spectral reconstruction problem in QCD, namely data on the spectral function, normalization, asymptotic information, and higher-dimensional reconstruction.

One important aspect of spectral functions, that is not yet implemented in the package, is the conditioning on inequality constraints. We have discussed two different methods to include such constraints in the GP model, one based on virtual observations and one based on non-Gaussian processes. These methods both provide promising results and a path forward to include such constraints. However, they both come with their specific challenges and are therefore not yet implemented in the *fredipy* package. Especially the non-Gaussian process method is able to provide promising results and by improving the optimization algorithm, we expect to be able to include this method in the package in the future. With these improvements, we expect to be able to provide a consistent and flexible framework for spectral reconstruction in QCD and other applications, that can incorporate all available prior information and provide realistic systematic error estimates.

In the following chapters, we apply the Gaussian process method to various spectral reconstruction problems in QCD to demonstrate its flexibility and reliability. We present different parameter optimization methods, each becoming more sophisticated with subsequent applications. Generally, identifying the appropriate optimization routine for the hyperparameters is the most challenging aspect of the reconstruction process and offers the greatest potential for improvement.

4. Yang-Mills Glueball Masses

This chapter is based on the publication [1].

In this chapter, we present the first application of the Gaussian process framework for spectral reconstruction, which was established in the previous chapter. We utilize this framework for the determination of glueball masses in Yang-Mills theory. The hadronic spectrum of Yang-Mills theory and QCD includes purely gluonic bound state contributions, the glueballs. The experimental verification of their existence is an important test of QCD; however, it is not yet conclusive [210–213] as these states are difficult to access due to their large overlap with other hadronic resonances. Different possible experimental candidates have been proposed, including the various f_0 states, some of which are expected to appear in decay channels of J/ψ [214, 215]. The overlap with other states also complicates their theoretical determination when considering QCD; for corresponding lattice calculations see [216–218]. In Yang-Mills theory, the situation is much simpler and the first few lightest states are well known; for lattice results see e.g. [86–91]. For computations with functional approaches – in particular with a combination of Dyson-Schwinger equations (DSE) and Bethe-Salpeter equations (BSE) – see e.g. [96, 98–103].

Here, we put forward a self-consistent functional ansatz for computing masses of bound states by exploiting their overlap with resonant interaction channels of gauge-fixed correlation functions. The approach is then used to determine the masses of the scalar ($J^{PC} = 0^{++}$) and pseudo-scalar ($J^{PC} = 0^{-+}$) Yang-Mills glueballs, utilizing the fact that these states have overlap with channels of the four-gluon vertex that carry the respective symmetries, where they appear as peaks of the corresponding spectral functions. We use the Gaussian process framework presented in the previous Chapter 3 to compute these spectral functions by reconstructing Euclidean correlators obtained within the functional renormalization group framework in [219].

This chapter is organized as follows. In Section 4.1, we discuss the spectral representation of Euclidean dressing functions in Yang-Mills theory. In Section 4.2, we present the four-gluon

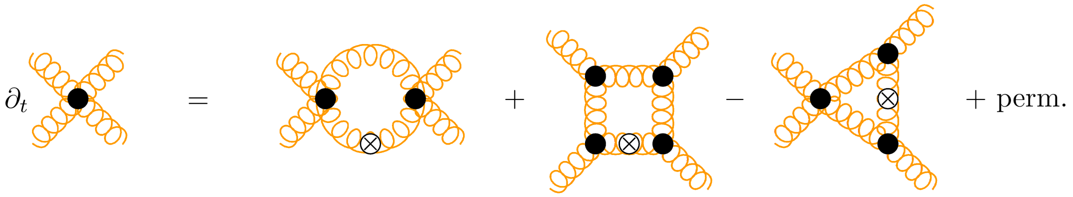


Figure 4.1: fRG equation for the s-channel four-gluon vertex dressing. Wiggly orange lines correspond to fully dressed gluon propagators; black dots indicate fully dressed vertices. Permutations include the various possible configurations of external legs as well as permutations of the regulator insertion (indicated by a crossed circle). Due to projections onto specific tensor structures, the ghost loop analytically drops out.

correlation function, its spectral representation, and the tensor structures that are relevant for the glueball. An overview of the fRG setup used to compute the four-gluon vertex dressing functions is given in Section 4.3. In Section 4.4, the resulting spectral functions are presented, the GP parameter optimization is discussed, and we report the masses of the scalar and pseudo-scalar glueballs. We conclude in Section 4.5.

4.1 SPECTRAL REPRESENTATIONS OF YANG-MILLS CORRELATION FUNCTIONS

As previously discussed in Section 2.1, correlation functions of physical states, and in particular the two-point functions, admit a spectral representation. For the propagator – the inverse 1PI two-point function – this is the Källén-Lehmann (KL) representation (2.19). The spectral functions of asymptotic states are positive semi-definite and admit the interpretation of a probability density. In gauge theories, however, the situation becomes more complicated. To begin with, even the existence of a KL representation is not settled for ghost and gluon propagators, and may feature additional structures in the complex momentum plane; for a detailed discussion see [219, 220]. Moreover, in the Landau gauge, the gluon and ghost spectral functions exhibit negative infrared and ultraviolet tails; see [127]. These properties can be inferred from the respective IR and UV asymptotic behavior of the Euclidean correlation functions [127, 221, 53, 219], and these relations also hold true for the present analysis involving four-gluon vertices. Note also that while gauge-fixed correlation functions may not permit a KL representation, the scattering matrix elements are directly constructed in terms of these correlators and obey the KL representation. Hence, features that are in direct correspondence to observables – such as bound states – can still be extracted from such gauge-fixed correlation functions.

4.2 Four-Gluon Correlation Function

Here, we consider single interaction channels that have overlap with the bound states of interest. In these channels, the respective bound states appear as resonances in the spectral functions. The spectral representations of these channels follow directly from the structure of the full, analytically continued correlation functions; see e.g. [222]. Up to minor modifications, they are given by the Källén-Lehman spectral representation (2.19): for the relevant scalar dressings of the four-gluon vertex, we use [52]

$$\lambda_{A^4}(p^2) \,=\, \lambda_{A^4,0} + \int_0^\infty \frac{\mathrm{d}\omega}{\pi} \frac{\omega}{p^2+\omega^2}\, \rho_{A^4}(\omega)\,. \tag{4.1}$$

The constant part $\lambda_{A^4,0}$ accounts for the classical contribution.

In order to access the masses of the scalar $J^{PC} = 0^{++}$ and pseudo-scalar $J^{PC} = 0^{-+}$ glueball, we have to determine tensor structures and momentum channels that overlap with these states. In general, it is desirable that the chosen channels have overlap only with the states of interest, as any reconstruction method faces increasing problems with multi-peak structures due to the exponential suppression of heavier states in the Euclidean data. Accordingly, their resolution requires an exponentially increasing accuracy, contributing to the ill-conditioned nature of the reconstruction problem.

For the scalar glueball, the above requirement is particularly simple to satisfy, since it is the lightest excitation and the classical tensor structure suffices, i.e.

$$\begin{aligned} \tau^{abcd}_{\mathrm{cl},\mu\nu\rho\sigma} \,=&\, f^{abe} f^{cde} (\delta_{\mu\rho}\delta_{\nu\sigma} - \delta_{\mu\sigma}\delta_{\nu\rho}) \\ &+ f^{ace} f^{bde} (\delta_{\mu\nu}\delta_{\rho\sigma} - \delta_{\mu\sigma}\delta_{\nu\rho}) \\ &+ f^{ade} f^{bce} (\delta_{\mu\nu}\delta_{\rho\sigma} - \delta_{\mu\rho}\delta_{\nu\sigma})\,. \end{aligned} \tag{4.2}$$

In Landau gauge, the longitudinal part of the correlation functions is related to the redundant, gauge fixing degrees of freedom, while the transversal part carries the physical information [223, 219]. Additionally, the transversal part of the flow equation of any n-point correlation function is closed, i.e. it only depends on other, fully transversal, correlation functions. We therefore project onto the transversal part of the 4-gluon vertex, with the transverse projection operator given by

$$\Pi^\perp_{\mu\nu}(p) \,=\, \delta_{\mu\nu} - \frac{p_\mu p_\nu}{p^2}\,. \tag{4.3}$$

The full projection operator for obtaining the scalar glueball mass is simply a contraction with the transverse part of the classical tensor structure, given by

$$\mathcal{P}^{abcd}_{s,\mu\nu\rho\sigma}(p,p,-p,-p) \,=\, \frac{\left[\Pi^\perp(p)\Pi^\perp(p)\Pi^\perp(-p)\Pi^\perp(-p)\tau_{A^4,cl}(p,p,-p)\right]^{abcd}_{\mu\nu\rho\sigma}}{\Pi^\perp(p)\Pi^\perp(p)\Pi^\perp(-p)\Pi^\perp(-p)\tau_{A^4,cl}(p,p,-p)\tau_{A^4,cl}(p,p,-p)}\,. \tag{4.4}$$

Indices are suppressed for simplicity and the external momenta are already matched to the momentum parametrization of the four-gluon vertex,

$$p_1 = p_2 = -p_3 = -p_4 \equiv p\,. \tag{4.5}$$

We have restricted ourselves to a single exchange momentum and the external (incoming and outgoing) momenta are chosen to have the same magnitude, and for the scalar glueball, they are chosen to be parallel. For the calculation of the mass of the pseudoscalar glueball, we use

$$\tau^{abcd}_{ps,\mu\nu\rho\sigma}(p_1,p_2) \;=\; \frac{\epsilon_{\mu\nu\alpha\beta}p_1^\alpha p_2^\beta}{\sqrt{p_1^2p_2^2}}\frac{\epsilon_{\rho\sigma\gamma\delta}p_1^\gamma p_2^\delta}{\sqrt{p_1^2p_2^2}}(\delta^{ab}\delta^{cd}+\delta^{ac}\delta^{bd}+\delta^{ad}\delta^{bc}) \tag{4.6}$$

as a projection operator, which has the correct transformation properties (see e.g. [96, 101]), and does not overlap with the scalar glueball. In (4.6), $\varepsilon_{\mu\nu\rho\sigma}$ denotes the fully antisymmetric tensor and the momenta are chosen to be orthogonal, $p_1 \cdot p_2 = 0$. It therefore projects onto the part of the four-gluon vertex that carries the quantum numbers of the pseudoscalar glueball. The pseudo-scalar projection operator is defined by the tensor structure (4.6) and given by

$$\mathcal{P}^{abcd}_{ps,\mu\nu\rho\sigma}(p_1,p_2,-p_1,-p_2) \;=\; \frac{\left[\Pi^\perp(p_1)\Pi^\perp(p_2)\Pi^\perp(-p_1)\Pi^\perp(-p_2)\tau_{A^4,ps}(p_1,p_2)\right]^{abcd}_{\mu\nu\rho\sigma}}{\Pi^\perp(p_1)\Pi^\perp(p_2)\Pi^\perp(-p_1)\Pi^\perp(-p_2)\tau_{A^4,ps}(p_1,p_2)\tau_{A^4,ps}(p_1,p_2)}\,, \tag{4.7}$$

where the in- and outgoing external momenta are chosen to be orthogonal,

$$p_1 = -p_3, \quad p_2 = -p_4, \tag{4.8}$$

$$p_1 \cdot p_2 = 0, \quad p_1^2 = p_2^2 = p^2\,. \tag{4.9}$$

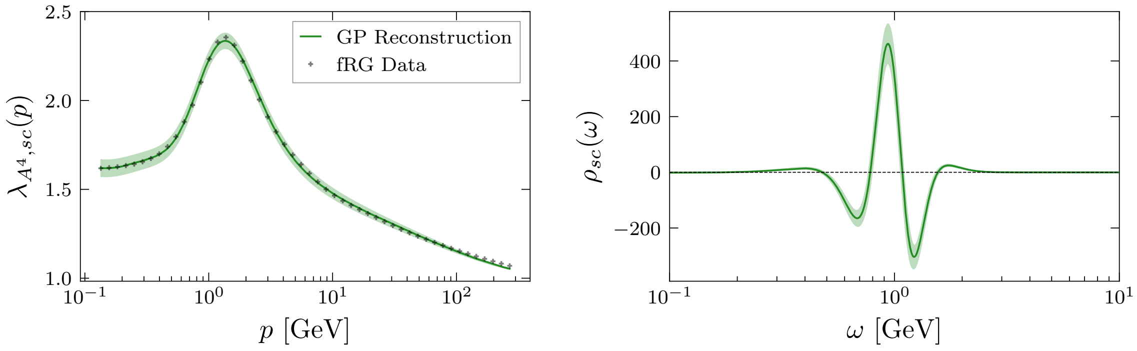
We note that the projection onto the ghost loop part of the flow analytically vanishes for both projections on the momentum configurations under consideration. This was observed for a similar momentum configuration in [224].

4.3 fRG Setup

The master equation of the fRG is the flow equation of the scale-dependent 1PI effective action. It is obtained by introducing an IR cutoff with a cutoff scale k via a momentum-dependent mass function $R_k(p^2)$ that is added to the inverse propagator. The respective flow equation is derived by taking a derivative of the generating functions w.r.t. the cutoff scale k,

$$\partial_t\Gamma_k[\Phi] \;=\; \frac{1}{2}\mathrm{Tr}\frac{1}{\Gamma_k^{(2)}[\Phi]+R_k}\partial_t R_k, \qquad t = \log\left(\frac{k}{\Lambda}\right)\,, \tag{4.10}$$

where t is the RG-time, and the trace in (4.10) sums over species of fields, space-time (momentum), Lorentz indices, and group indices. The regulator functions carry the classical dispersion



(a) (b)

Figure 4.2: In (a), we show the Euclidean dressing of the four-gluon vertex $\lambda_{A^4,sc}$ with the projection to obtain the scalar glueball mass, see Section 4.2, from the fRG (black crosses). This is compared to the reconstruction from the GP (green line). The corresponding spectral function $\rho_{A^4,sc}$ over frequency ω obtained with GPR is shown in (b). The light green band represents the 1σ region.

of the ghost and gluon fields as well as a dimensionless shape function. The present results are computed with the usual exponential shape function,

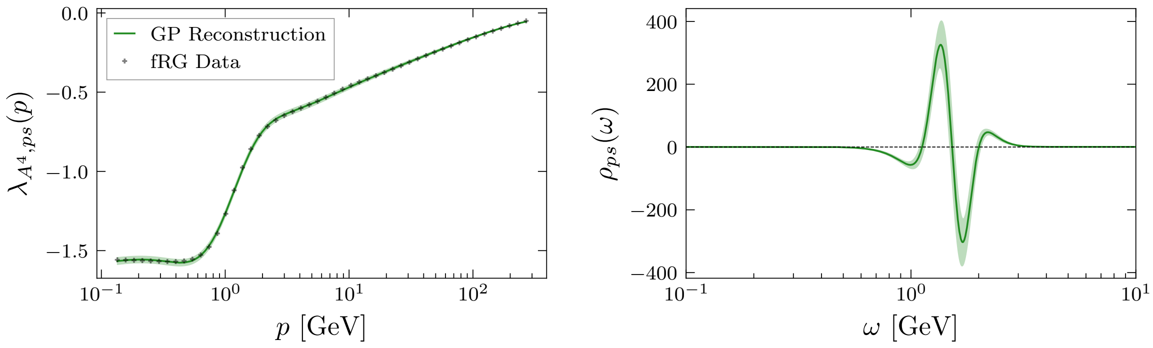
$$r(p^2/k^2) = \frac{\mathrm{e}^{-p^2/k^2}}{1-\mathrm{e}^{-p^2/k^2}}\,, \tag{4.11}$$

and an additional wave function renormalization $Z_{A,k}$ or $Z_{c,k}$; for more details see [219]. For a recent review of the fRG see e.g. [225] and references therein.

Our general setup in Landau gauge Yang-Mills theory follows [226, 219]. The flow of the four-point vertex is obtained by taking the fourth derivative of (4.10) w.r.t. the gluon field. In this work, we are only interested in certain channels of the four-gluon vertex. Hence, we do not solve the full system self-consistently, but take the flow of all other correlation functions such as the gluon propagators and the gluon vertex dressings from [219] as input.

The fRG equation for the four-gluon vertex solved in the present work is depicted in Figure 4.1. This flow is integrated on the solution of the correlation functions obtained in [219]. There, different IR closures of correlation functions in the Landau gauge have been computed, and the present work utilizes the *scaling* solution. The independence of this choice has recently been shown in [101], where both solutions – decoupling and scaling – were considered in the context of glueballs. The approximation used in [219] only includes the primitively divergent (classical) tensor structures, which leads to semi-quantitative results. Further details can be found in [219].

We use the k-dependent dressing functions from [219] as input. Their parametrizations are



(a)

(b)

Figure 4.3: In (a), we show the Euclidean dressing of the four-gluon vertex $\lambda_{A^4,ps}$ with the projection to obtain the pseudo-scalar glueball mass, see Section 4.2, from the fRG (black crosses). This is compared to the reconstruction from the GP (green line). The corresponding spectral function $\rho_{A^4,ps}$ over frequency ω obtained with GPR is shown in (b). The light green band represents the 1σ region.

given by

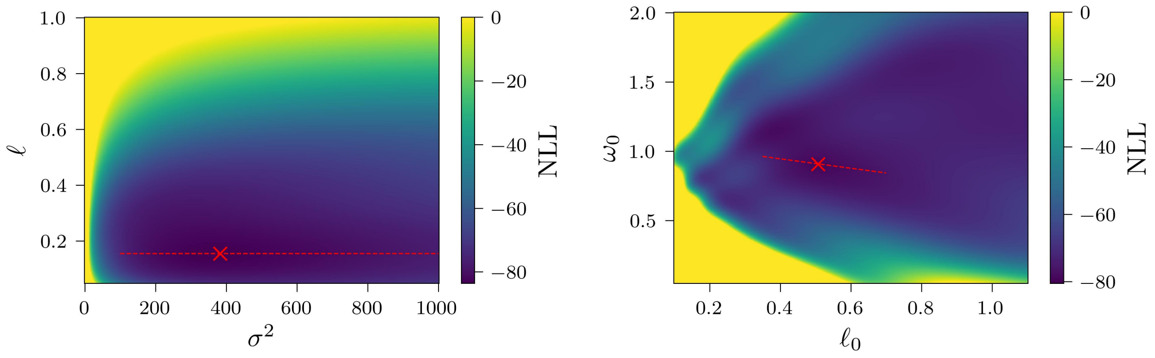
$$\begin{aligned} \Gamma^{(2),ab}_{AA,\mu\nu}(p) &= \delta^{ab}\Pi^{\perp}_{\mu\nu}(p)Z_A(p)(p^2+m_T^2), \\ \Gamma^{(3),abc}_{A^3,\mu\nu\rho}(p_1,p_2) &= if^{abc}\lambda_{A^3}(\bar{p})\Big[(p_1-p_2)_\rho\delta_{\mu\nu}+\text{perm.}\Big], \\ \Gamma^{(4),abcd}_{A^4,\mu\nu\rho\sigma}(p_1,p_2,p_3) &= \lambda_{A^4}(\bar{p})\left[f^{abn}f^{cdn}\delta_{\mu\rho}\delta_{\nu\sigma}+\text{perm.}\right], \end{aligned} \tag{4.12}$$

where we approximate the full momentum dependence of the vertices with the symmetric point configuration $\bar{p}$, see e.g. [226], defined by

$$\bar{p}^2 = \frac{1}{n}\sum_{i=1}^{n} p_i^2\,, \tag{4.13}$$

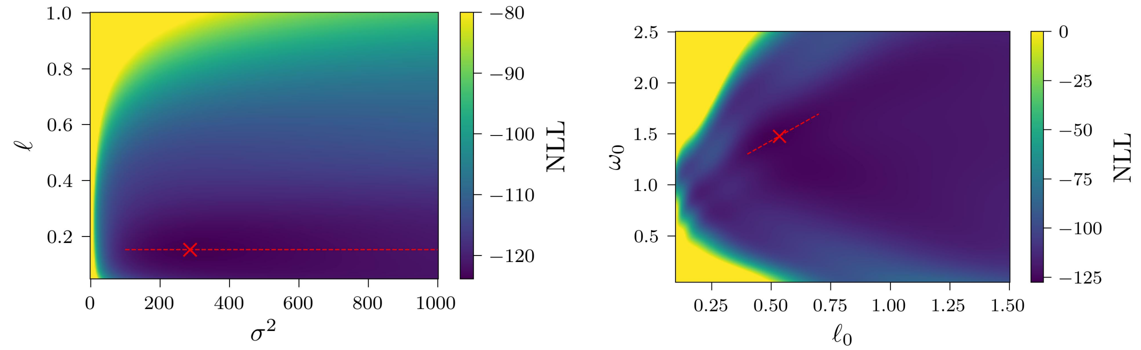
with $n = 3, 4$ is the number of external momenta.

We remark in this context that correlation functions in Landau gauge Yang-Mills theory computed within sophisticated truncations to the fRG pass all available lattice benchmark tests; see [226, 219]. This concerns in particular the ghost and gluon propagators, whereas lattice results for vertices still exhibit large uncertainties. Nevertheless, since state-of-the-art functional results for correlation functions fully agree with lattice calculations within statistical errors, any reconstruction based on the former approaches is consistent with the latter.



(a) Kernel parameters, scalar channel.

(b) Bias parameters, scalar channel.



(c) Kernel parameters, pseudo-scalar channel.

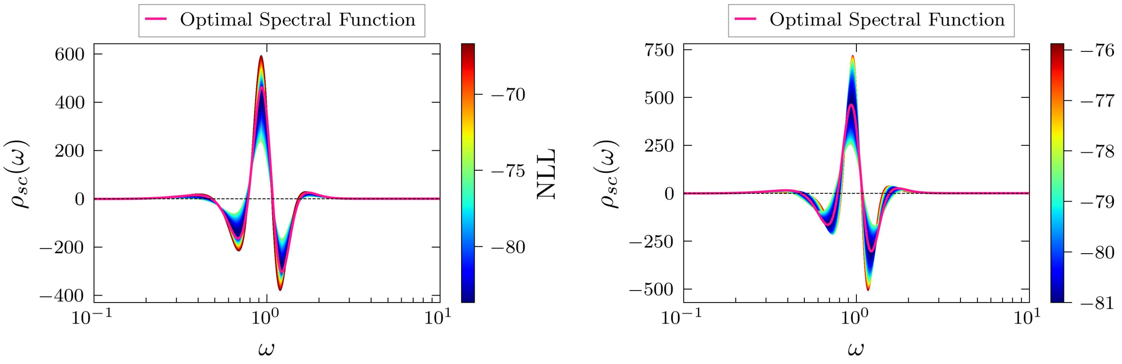
(d) Bias parameters, pseudo-scalar channel.

Figure 4.4: Grid scans of the NLL (3.31) of the reconstructions for both channels. Note that the optimizations are performed subsequently, starting with the squared exponential kernel and followed by the bias parameters. The red lines indicate the trajectories in parameter space used for comparing the variance in the spectral functions; see Figures 4.5 and 4.6. The red cross indicates the NLL optimized parameters; see Table 4.1.

4.4 RESULTS

We calculate the vertex dressings with the fRG as was just outlined in Section 4.3. The resulting Euclidean dressing functions for the scalar and pseudo-scalar projections are shown in Figures 4.2a and 4.3a, respectively.

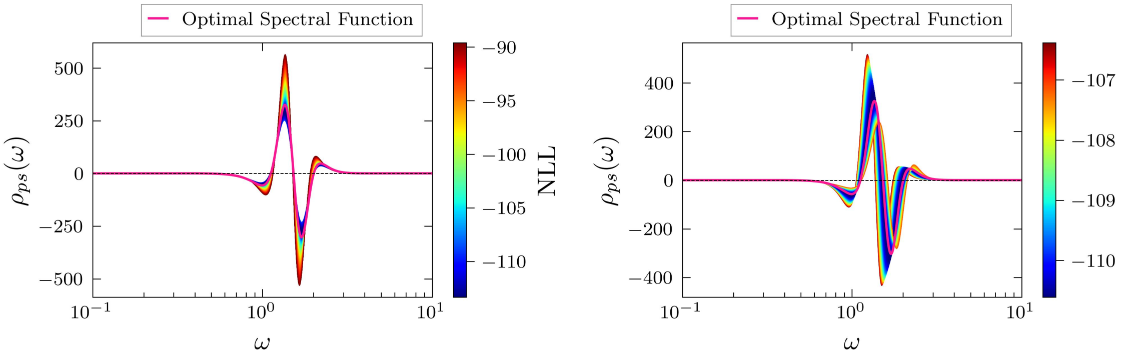
As mentioned before, in the channels considered here, the ghost loops drop out due to the combination of projection and momentum configuration; similarly to [224]. Hence, these channels are free of the IR divergences that are in general present in the four-gluon vertex and we can utilize the constraint $\rho_{A^4}(0) = 0$ in the GP reconstruction. We use the standard squared exponential kernel (3.25) for the GP prior. Furthermore, an additional bias is introduced in order to suppress unphysical oscillations at the tails of the spectral function. Similar to the



(a) Scan of the flat direction in the kernel parameters.

(b) Scan around optimal bias parameters.

Figure 4.5: Spectral function of the scalar channel. The bands show the variance with respect to the flat directions of the parameter space; see Figures 4.4a and 4.4b. The peak position is observed to be robust, even under large parameter changes. The overall magnitude on the other hand shows considerable variation for both sets of parameters.



(a) Scan of the flat direction in the kernel parameters.

(b) Scan around optimal bias parameters.

Figure 4.6: Spectral function of the pseudo-scalar channel. The bands show the variance with respect to the flat directions of the parameter space; see Figures 4.4c and 4.4d. The peak position is observed to be robust under variations of the standard squared exponential kernel parameters but shows significantly stronger deviations compared to the scalar channel for different values of the bias parameters.

procedure applied in [187], this is achieved by rescaling the frequency with a soft step function

$$\omega \to \tilde{\omega} = \frac{1}{\exp(-2(\omega-\omega_0)/\ell_0)+1}\,, \tag{4.14}$$

where the parameter ℓ_0 controls the steepness and ω_0 the position of the midpoint. This rescaling can be understood as the introduction of a frequency-dependent length scale in the standard squared exponential kernel, with smaller values around ω_0 and larger values at the tails of the spectral function. We note that the resonances of interest are already observed at the same position without introducing this additional bias. However, the peaks are enhanced by this procedure while the reconstruction of the correlator remains in good agreement with the input data. We have extensively tested the impact of this rescaling and found that changes to the functional form of the kernel do not significantly alter the peak position of the spectral function, as long as the qualitative features remain similar. While this parametrization suppresses additional structures (such as excited glueball states at higher energies; see e.g. [101]), even without the rescaling (4.14) no additional features beyond the dominant peak corresponding to the bound state are observed, apart from the usual oscillatory behavior at the tail of the spectral function, which can not be distinguished from reconstruction artifacts. This implies that higher excited states exhibit at most subleading contributions to these vertex projections. Resolving these structures therefore requires either more sophisticated projections of the tensor structures or a significantly higher precision in the calculation of the vertex itself.

The parameters of the standard squared exponential kernel and frequency rescaling are optimized by minimizing an objective function, conventionally taken to be the negative log-likelihood (NLL) (3.31). The optimal parameters are found by performing grid scans in the parameter space of the kernel and bias parameters. When optimizing the parameters of the squared exponential kernel (3.25) and the frequency rescaling bias (4.14) simultaneously, the parameters favor values that tend to nullify the bias, such as ℓ_0 becoming large. Hence, the parameters are first optimized only considering the bare squared exponential kernel in order to obtain baseline values. Subsequently, the bias is introduced, and its parameters are optimized given the squared exponential kernel calculated beforehand. This way, the position and size of the dynamical region of the spectral function are also subject to optimization. Unsurprisingly, the NLL shows a flat direction where some parameters are unconstrained; see Figure 4.4. This can be interpreted as a manifestation of the ill-conditioned nature of the inverse problem, and may be treated by imposing a hyperprior; a similar procedure has been performed in [187]. We observe that changing the parameters in this direction has negligible impact on the resonant peak position; see Figures 4.5 and 4.6. Hence, the seemingly heuristic use of a generic hyperprior is well justified in this context as it does not introduce a bias for the quantity of interest. However, the overall magnitude of the spectral function is sensitive to these parameters and consequently, this quantity is not well constrained by the model. Predictions of other features such as the peak position, width, and overall shape are robust as the NLL diverges quickly when considering non-optimal parameters. Scanning the spectral functions in the plane of the bias parameters on the other hand reveals a more drastic change in the peak position, as observed in Figures 4.5 and 4.6. However, these parameters are restricted to a much smaller region by the likelihood and the stability of the peak position is retained. The optimal hyperparameters are given in Table 4.1 below.

J^{PC}	σ	ℓ	ω_0	ℓ_0
0^{++}	19.56	0.155	0.908	0.508
0^{-+}	16.93	0.152	1.476	0.534

Table 4.1: Optimized GP hyperparameters for the kernel and rescaling functions, (3.25) and (4.14).

The intrinsic error estimate of the GP posterior is fixed to $\sigma_n = 10^{-2}$, corresponding to an upper bound on the uncertainty of the fRG calculation. σ_n is not optimized as this diminishes the significance of the likelihood for the other parameters [175], as already discussed in Section 3.2.2. We have extensively tested different parameter optimization strategies, that attempt to optimize the parameters of the squared exponential kernel and the bias simultaneously, or in a different order. However, we find that the above procedure yields the most stable results, in particular for the peak position of the spectral function. Methods, that optimize the squared exponential kernel and the bias simultaneously intuitively have a similar problem as the optimization of variance and error estimate. These parameters have strong interdependencies, and optimization will never converge, due to cancellations in the NLL and resulting flat direction. We can also observe, that the NLL has multiple local minima in the two-dimensional scan, especially for the bias parameters, cf. Figures 4.4b and 4.4d. As we go to higher dimensional parameter spaces, such local minima pose a more significant problem, as they become harder to identify.

The reconstructed dressings are compared to the fRG input data in Figures 4.2a and 4.3a, with the associated spectral functions shown in Figures 4.2b and 4.3b, respectively. From a Hubbard–Stratonovich transformation it can be inferred that the vertex dressing corresponds to the negative dressing function of the bound state under consideration. Hence, the spectral functions $\rho_{sc/ps}$ are computed from the negative vertex dressing. Consequently, the positive peak indicates an asymptotic state that is interpreted as the respective glueball resonance. We also observe negative structures in the spectral function since the four-gluon vertex itself is not a gauge-invariant object. The reconstruction of the vertex dressing largely reproduces the fRG data within errors. For high momenta, the result deviates more strongly, in particular for the scalar glueball. This is due to the additional bias introduced to the kernel that specifically suppresses any dynamics in the UV regime.

The glueball masses are extracted from the dominant peak positions of the spectral functions. We obtain $\hat{\omega}_{sc} = 0.93\,\text{GeV}$ for the scalar and $\hat{\omega}_{ps} = 1.35\,\text{GeV}$ for the pseudo-scalar channel. Since we work within the s-channel approximation and have two incoming momenta each with the magnitude p, the peak position corresponds to the half of the glueball mass, i.e. $m_{sc/ps} = 2\,\hat{\omega}_{sc/ps}$. Hence, we obtain the masses $m_{sc} = 1870(75)\,\text{MeV}$ for the scalar and $m_{ps} = 2700(120)\,\text{MeV}$ for the pseudo-scalar glueball. The reported errors are a combination

J^{PC}	lattice		DSE-BSE		this work
0^{++}	1760(70)	[88]	1850(130)	[101]	1870(75)
	1740(70)	[89]	1640	[98]	
	1651(23)	[90]			
	1618(26)(25)	[91]			
0^{-+}	2650(60)	[88]	2580(180)	[101]	2700(120)
	2610(70)	[89]	4530	[98]	
	2600(40)	[90]			
	2483(61)(55)	[91]			

Table 4.2: Comparison of scalar ($J^{PC} = 0^{++}$) and pseudo-scalar ($J^{PC} = 0^{-+}$) glueball masses from different methods. The results of [88, 89] are rescaled to match [90, 91] with $r_0 = 1/418(5)\,\mathrm{MeV}$. The errors of [88, 89] are a combination of statistical as well as systematic uncertainties stemming from the lattice anisotropy and the scale r_0. The errors for [90] are statistical only. For [91], the quoted values are the statistical as well as systematic uncertainties for the continuum extrapolation, respectively. For [101], the error comes from the extrapolation method.

of the standard deviations computed from the GP posterior and an additional 3% error from the scale setting procedure of the input data. We compare our results with masses obtained from independent lattice and DSE/BSE studies of the glueball spectrum in Table 4.2 and find them to be in reasonable agreement, in particular for the pseudo-scalar channel where they match well within the provided uncertainties.

4.5 Conclusion

We put forward a self-consistent approach for the extraction of bound state information from gauge-fixed correlation functions. Key to this framework is the spectral reconstruction of interaction channels in Euclidean space-time that have overlap with the corresponding gauge-invariant bound state. The method is applied to low-lying glueball states in Yang-Mills theory, extracted from the dressing functions of the Euclidean four-gluon vertex. With appropriate projection operators of the four-gluon vertex, we obtain access to the masses of the scalar and pseudo-scalar glueballs.

The Euclidean dressings are obtained with the functional renormalization group, utilizing earlier results for correlation functions from [219]. The respective spectral functions are then computed via Gaussian process regression and their resonance peaks are identified with the glueball masses. For the scalar and pseudo-scalar glueballs, we arrive at 1870(75) MeV and 2700(120) MeV, respectively. The results agree well with independent studies of the glueball spectrum, lending further credibility to our proposed method of computing bound state properties from vertex dressing functions via spectral reconstruction. The present approach can also be directly applied to higher glueball states in Yang-Mills theory, as well as glueball and other hadronic states in QCD.

5. STRONG COUPLING AT TIMELIKE MOMENTA

This chapter is based on the publication [2].

As laid out before, knowing the spectral function of a correlator, i.e. knowing the correlator in the full complex plane, allows the computation of other important quantities using Bethe-Salpeter equations in order to determine bound states, or serves as an input for spectral functional approaches that enable the computation of new spectral functions of different degrees of freedom of the theory. Additionally, in the calculation of physical scattering amplitudes, the strong coupling in Minkowski space can be a necessary ingredient.

Since we have established, the spectral function of the ghost and the gluon coupling in an earlier work, cf. [187], we can use this knowledge to compute the spectral function of other QCD observables. Specifically, we establish a spectral representation for the strong coupling constant and compute its spectral function. In this chapter, we will compare results from the spectral reconstruction of the strong coupling constant with results from a direct calculation of the spectral function, utilizing the ghost and gluon spectral functions as input. This non-trivial benchmark of our reconstruction method yields remarkable agreement between the direct and indirect results, thereby making a strong case for the reliability of spectral reconstruction via probabilistic inversion with GP models. For the first time, we will apply the method of including asymptotics in the Gaussian process, cf. Section 3.3.3, to the spectral reconstruction.

This chapter is organized as follows. In Section 5.1, we give a motivation for the calculation of the strong coupling at timelike momenta, by calculating a standard S-matrix element and derive the spectral representation for the strong coupling constant. The reconstruction procedure, the specifics of the inclusion of the asymptotic information, and the resulting spectral functions are presented in Section 5.2. We conclude in Section 5.3.

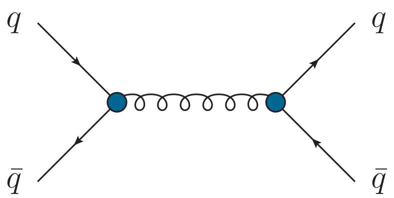


Figure 5.1: $q\bar{q}$-scattering process with a one-gluon exchange. At sufficiently large timelike exchange momenta, this process plays an important role in its respective S-matrix elements. Consequently, all internal quantities are dressed. Blue blobs represent full vertices, and the wiggly internal line is a full gluon propagator.

5.1 Scattering Processes & the Timelike QCD Coupling

Scattering processes and decays in QCD are described in terms of S-matrix elements. At low energies, the operators of the physical in and out states are complicated objects in terms of the fundamental QCD degrees of freedom. For instance, a description of the Compton scattering of protons requires the definition of the proton or, more generally, the nucleon operator in terms of its partonic constituents. Since, on the fundamental level, the partons are related to quarks and gluons, the building blocks of the respective S-matrix elements are quark-gluon and quark-photon scattering processes.

In most partonic models the fundamental scattering processes are approximated by effective models for the exchange process, such as one-gluon exchange potentials that carry the qualitative property of the gluon mass gap in QCD in terms of an effective mass. Ideally, however, they should be constructed from tree-level processes in QCD with full propagators and vertices, both of which carry on-shell, timelike, and spacelike momenta. The final S-matrix is gauge-invariant, while the tree-level components making up the individual S-matrix element contributions are not. Moreover, the S-matrix admits a spectral representation, which is not necessarily present for the gauge-fixed correlation functions.

5.1.1 Cross-Section of Quark–Anti-Quark Scattering Events and the S-Matrix Element

In the present work, we undertake a first step towards such a determination of non-perturbative S-matrix building blocks in QCD. To that end, we compute the timelike strong coupling in 2+1 flavor QCD that governs the quark–anti-quark scattering process depicted in Figure 5.1. This diagram is at the core of many of the scattering processes used to determine the strong running coupling, defined as

$$\alpha_s(p) = \frac{g_s^2(p)}{4\pi} \,. \tag{5.1}$$

It is also one of the fundamental building blocks of scattering processes in the Pomeron model [227–230] – such as the aforementioned Compton scattering of the proton – where it is typically estimated by one-gluon exchange models. For a review, see [231]; for a recent application related to the present work, see [232].

Assuming the incoming and outgoing quarks $q(p)$ and anti-quarks $\bar{q}(\bar{p})$ to be on-shell, $q\bar{q}$-scattering is similar to e^+e^- scattering. We expect this analogy to hold for sufficiently large timelike exchange momenta $p^2 \gtrsim 1\,\mathrm{GeV}^2$, whereas for $p^2 \lesssim 1\,\mathrm{GeV}^2$ we enter the hadronic, strongly correlated regime. There, the non-trivial embedding of the scattering quarks and anti-quarks in hadrons becomes increasingly relevant, and quark–anti-quark scattering should be also considered off-shell alongside further, more complicated processes; for a formulation in functional approaches, see [229].

Here, we concentrate on the one-gluon exchange diagram as one of the building blocks of the full S-matrix element. The associated tree-level process shown in Figure 5.1 consists of two full quark-gluon vertices, $\Gamma^{(3)}_{q\bar{q}A}(p_1, p_2, p)$ with the on-shell momenta p_1, p_2 for the incoming as well as $\Gamma^{(3)}_{q\bar{q}A}(p_3, p_4, -p)$ with on-shell $-p_3, -p_4$ for the outgoing quark and anti-quark, respectively. The relative minus sign is due to the notational convention in functional computations treating all momenta as incoming. The momentum p is that of the exchange gluon with the full gluon propagator $G_A(p)$. In combination, this process can be expressed as

$$\begin{aligned}\langle q(p_3)\bar{q}(p_4)|\, S\, |q(p_1)\bar{q}(p_2)\rangle \simeq \prod_{i=1}^{4} Z_q^{-1/2}(p_i) \times \Bigg\{ \Big[\bar{u}_q(p_3)\Gamma^{(3)}_{q\bar{q}A}(p_3, p_4, p) v_q(p_4)\Big]^a_\mu \\ G_A(p)\,\delta^{ab}\left(g^{\mu\nu} - \frac{p^\mu p^\nu}{p^2}\right)\Big[\bar{v}_q(p_2)\Gamma^{(3)}_{q\bar{q}A}(p_1, p_2, -p) u_q(p_1)\Big]^b_\nu \Bigg\}\,, \end{aligned} \tag{5.2}$$

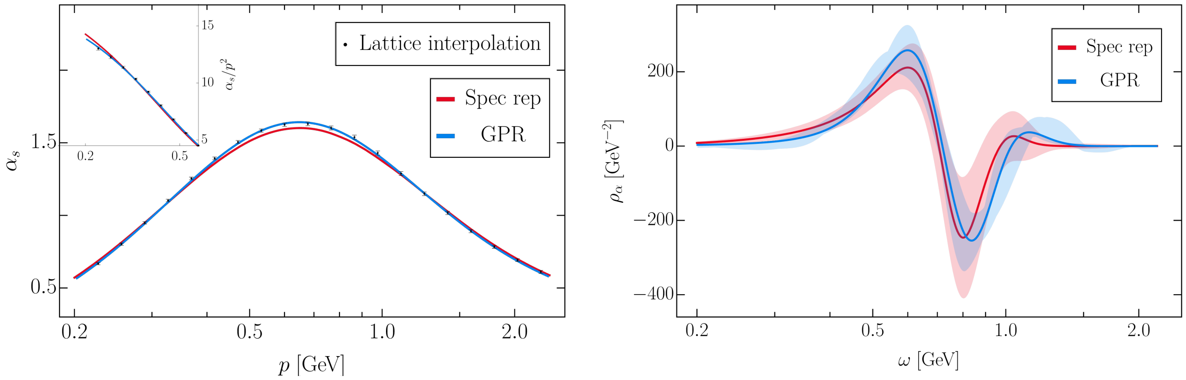
where the (on-shell) quark wave functions Z_q originate in the LSZ reduction formula. Note that the quark and gluon wave functions are defined such that the quark and gluon propagators $G_q(p), G_A(p)$ are proportional to $1/Z_q(p), 1/Z_A(p)$, respectively. The scalar parts of the Euclidean propagators read

$$G_A(p) = \frac{1}{Z_A(p)}\frac{1}{p^2}\,, \qquad G_q(p) = \frac{1}{Z_q(p)}\frac{1}{p^2 + M_q^2(p)}\,, \tag{5.3}$$

where the full propagators are proportional to the identity in color space in the adjoint (gluon) and fundamental (quark) representations. The gluon propagator in the Landau gauge also carries the projection operator on the transverse subspace (see (5.2)), and the quark propagator is multiplied by $\mathrm{i}\not{p} + M_q(p)$. With (5.3), the standard LSZ factors carrying the pole residues are simply $Z_q^{-1/2}$, as already used in (5.2).

The S-matrix element (5.2) is renormalization group (RG) invariant, as required. To see this explicitly, we reparametrize the vertices in terms of wave functions of the legs and an RG-invariant core,

$$\Gamma^{(3)}_{q\bar{q}A}(p_i, p_j, p) = Z_q^{\frac{1}{2}}(p_i) Z_q^{\frac{1}{2}}(p_j) Z_A^{\frac{1}{2}}(p)\, \bar{\Gamma}^{(3)}_{q\bar{q}A}(p_i, p_j, p)\,, \tag{5.4}$$



(a) Spacelike Taylor coupling $\alpha_s(p)$.

(b) Spectral function $\rho_\alpha(\omega)$ of $\alpha_s(p)$.

Figure 5.2: Spacelike Taylor coupling α_s in QCD (a) and its spectral function $\rho_\alpha(\omega)$ (b). We compare the spectral function computed directly via (5.14) (red) to that obtained via reconstruction with GPs (blue). The direct calculation uses the reconstruction results for gluon and ghost spectral functions from [187]. For the reconstruction, we use the gluon and ghost propagator data in 2+1 flavor lattice QCD from [233, 234]. Both the input spectral functions and the corresponding lattice data are displayed in Figure 5.3. The coupling spectral functions obtained via these two complementary approaches share all qualitative features, such as peak positions and heights as well as asymptotic behavior. The peak structure can be connected to the respective peak structure of the gluon spectral function; see Figure 5.3b. The error band of the reconstruction result accounts for the change in the spectral function when varying the GP kernel parameters, whereas that of the direct calculation originates from propagating the uncertainty of the input. The Euclidean lattice data for the Taylor coupling α_s are displayed as gray squares in Figure 5.2a. We compare it to the data from its spectral representation (5.12) (red) as well as the reconstruction result (blue), showing that the representation holds and that the reconstruction accurately describes the lattice data.

where $\bar{\Gamma}^{(3)}_{q\bar{q}A}$ has the transformation properties of a running coupling and naturally occurs in the S-matrix element. Inserting (5.4) into the S-matrix element (5.2) leads us to

$$\begin{aligned}&\langle q(p_3)\bar{q}(p_4)|\, S\,|q(p_1)\bar{q}(p_2)\rangle \simeq \\ &\quad\left[\bar{u}_q(p_3)\bar{\Gamma}^{(3)}_{q\bar{q}A}(p_3,p_4,p)v_q(p_4)\right]^a_\mu \frac{1}{p^2}\,\delta^{ab}\left(g^{\mu\nu}-\frac{p^\mu p^\nu}{p^2}\right)\left[\bar{v}_q(p_2)\bar{\Gamma}^{(3)}_{q\bar{q}A}(p_1,p_2,-p)u_q(p_1)\right]^b_\nu .\end{aligned} \tag{5.5}$$

We restrict ourselves to the limit of large transfer momentum $p^2 \equiv s$ of the scattering event with $p_1p_3 = p_2p_4 = s(1-\cos\theta)/4$ and scattering angle $\cos\theta = \boldsymbol{p}_1\boldsymbol{p}_3/(|\boldsymbol{p}_1||\boldsymbol{p}_3|)$. For small s, we approach the reliability limit of our approximations. We return to the respective discussion after deriving our results.

Additionally, in a last approximation step, we concentrate on the classical tensor structure $\gamma_\mu T^a$ in the full quark-gluon vertex,

$$\left[\bar{\Gamma}^{(3)}_{q\bar{q}A}(p_i,p_j,p)\right]^a_\mu \approx \mathrm{i}\,\gamma_\mu T^a\sqrt{4\pi\alpha_s(s)}\,. \tag{5.6}$$

Here, T^a is the $SU(3)$ generator in the fundamental representation, and $\alpha_s(s)$, defined in (5.1), is the strong coupling of the quark-gluon scattering process in the s-channel. On the equation of motion, the $\not{p}_i$ terms vanish, and we obtain

$$\left|\bar{u}_q(p_3)\gamma_\mu T^a v_q(p_4)\,\bar{v}_q(p_2)\gamma_\mu T^a u_q(p_1)\right|^2 \to \frac{s^2}{9}\left(1+\cos^2\theta\right)\,, \tag{5.7}$$

in the high energy limit. In (5.7), we have performed an average/sum over spins and color in the initial/final state. With (5.5) and (5.7), we arrive at

$$\left|\langle q(p_3)\bar{q}(p_4)|\,S\,|q(p_1)\bar{q}(p_2)\rangle\right|^2 \to \frac{1}{9}\left[4\pi\alpha_s(s)\right]^2\left(1+\cos^2\theta\right)\,. \tag{5.8}$$

with $\alpha_s(p)$ defined in (5.1). Equation (5.8) highlights the importance of the strong coupling constant $\alpha_s(s)$ for physical scattering processes. For the remainder of this work, we adopt the linear momentum argument $p=\sqrt{s}$ for the coupling.

In the present work, we shall compute the strong coupling $\alpha_s(p)$ and, hence, the above S-matrix element from its spectral representation for general complex frequencies, including the timelike momenta relevant for (5.8). We utilize that the strong coupling can be computed from the quark-gluon vertex, the three- and four-gluon vertices, as well as the ghost-gluon vertex. The computation involves the wave functions $Z_q(p), Z_A(p)$ of quarks and gluons as defined in (5.3) and the ghost wave function $Z_c(p)$ from

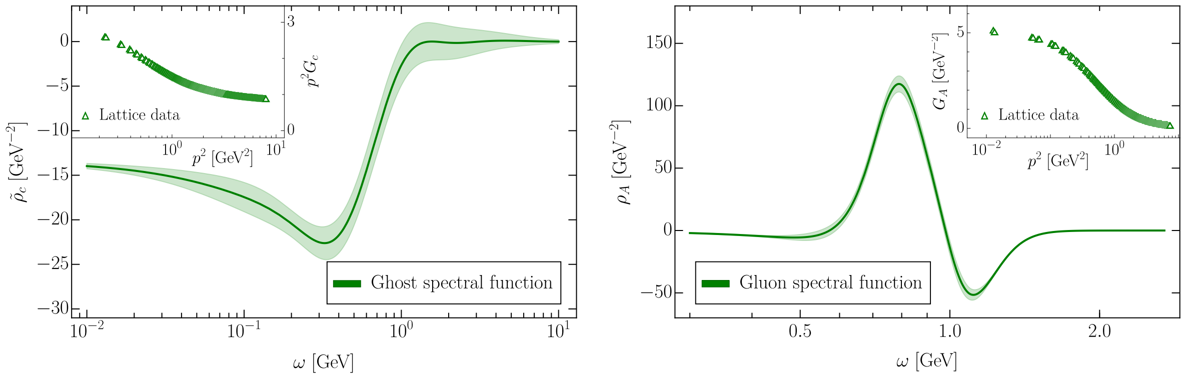
$$G_c(p) = \frac{1}{Z_c(p)}\frac{1}{p^2}\,. \tag{5.9}$$

The avatars of the strong couplings are then defined as the (symmetric point) dressings of the classical tensor structures, see (5.4) and (5.6).

A final definition of the strong coupling in the Landau gauge is given by the propagator or Taylor coupling, which utilizes Taylor's non-renormalization theorem for the ghost-gluon vertex. This leads to the Taylor coupling, solely defined by the ghost and gluon dressing functions,

$$\alpha_s(p) = \frac{g_s^2}{4\pi}\frac{1}{Z_A(p)Z_c^2(p)}\,. \tag{5.10}$$

All strong coupling avatars have the same universal two-loop running but differ for infrared momenta; see [235]. For an evaluation of the infrared differences between the Taylor coupling and the quark-gluon coupling, see [236]. The latter regime is not accessible within the present approximation. Hence, we use the Taylor coupling (5.10) for the evaluation of (5.8). Its corresponding spectral function ρ_α is depicted in Figure 5.2. It allows us to compute the coupling $\alpha_s(p)$ for complex frequencies including timelike momenta; see Figure 5.6. Timelike results for the strong coupling in the perturbative domain can be found, e.g., in [237, 238].



(a) Ghost spectral function.

(b) Gluon spectral function.

Figure 5.3: The continuous parts of the ghost (a) and gluon (b) spectral functions obtained in [53] (see also Section 5.1.3), used here as input for the calculation of the coupling spectral function shown in Figure 5.2a via its spectral representation (5.12). Shaded areas represent 1σ-bands of the statistical error of the mean prediction based on the available observations and precision. Note that for the calculation of the gluon spectral function, the UV and IR asymptotic regimes are assumed to be maximally large. This leads to a small reconstruction error without accounting for systematics; see Section 5.2 for a detailed discussion.

5.1.2 Spectral Representation

For the computation of (5.10), and hence of (5.8), we require the ghost and gluon propagators for timelike momenta, the Källén-Lehmann spectral representation (2.19). Note, that for propagators of 'unphysical' fields, such as gauge fields, positive semi-definiteness is no longer required and the spectral representation reduces to a statement about the analytic structure of the corresponding correlation function; see, e.g., [239–241, 127, 221, 53, 220]. The ghost propagator is known to exhibit a massless particle pole in the origin, entailing a delta pole at vanishing frequency in its spectral function ρ_c [53]. The gluon spectral function ρ_A is continuous along the whole real frequency axis and is not expected to show distributional contributions. Taking into account the explicit forms of the spectral functions, the ghost and gluon dressing functions can be expressed as

$$\begin{aligned}\frac{1}{Z_A(p)} &= p^2 \int_0^\infty \frac{\mathrm{d}\omega}{\pi} \frac{\omega\, \rho_A(\omega)}{\omega^2 + p^2}\,, \\ \frac{1}{Z_c(p)} &= \frac{1}{Z_c^0} + p^2 \int_0^\infty \frac{\mathrm{d}\omega}{\pi} \frac{\omega\, \tilde{\rho}_c(\omega)}{\omega^2 + p^2}\,,\end{aligned} \tag{5.11}$$

where $1/Z_c^0$ is the residue of the massless delta pole of ρ_c, and $\tilde{\rho}_c$ denotes the continuous part of the spectral function.

Given the existence of a spectral representation – for a deeper discussion on this, we refer to [220] – the associated correlation function must obey certain symmetries and fulfill requirements about its infrared and ultraviolet asymptotics. It can be shown that the existence of spectral representations for the ghost and gluon propagators implies the existence of such a representation also for the Taylor coupling as defined in (5.10). Specifically, it is given by

$$\alpha_s(p) \,=\, p^2 \int_0^\infty \mathrm{d}\omega \, \frac{\omega}{\omega^2 + p^2} \, \rho_\alpha(\omega) \,. \tag{5.12}$$

With (5.12), the strong coupling spectral function is obtained from its retarded correlator via

$$\rho_\alpha(\omega) \,=\, -\frac{2}{\omega^2} \mathrm{Im} \, \alpha_s\big(- \mathrm{i}(\omega + \mathrm{i}0^+)\big) \,. \tag{5.13}$$

Now we use the definition of the Taylor coupling (5.10) and insert the spectral representations of ghost and gluon dressing functions (5.11). Then, the spectral function (5.13) of the coupling can be written as

$$\begin{aligned} \rho_\alpha(\omega) \,=\, &- 2 \, \mathrm{Im} \left[\left(\int_0^\infty \frac{\mathrm{d}\lambda}{\pi} \frac{\lambda \, \rho_A(\lambda)}{\lambda^2 - \omega^2 + \mathrm{i}0^+} \right) \right. \\ &\left. \times \left(\frac{1}{Z_c^0} - \omega^2 \int_0^\infty \frac{\mathrm{d}\lambda}{\pi} \frac{\lambda \, \tilde{\rho}_c(\lambda)}{\lambda^2 - \omega^2 + \mathrm{i}0^+} \right)^2 \right] . \end{aligned} \tag{5.14}$$

Since the Taylor coupling decays logarithmically in the UV, its spectral function obeys a superconvergence condition [53, 221], given by

$$\int_0^\infty \mathrm{d}\omega \, \omega \rho_\alpha(\omega) \,=\, 0 \,. \tag{5.15}$$

In the case of the gluon propagator, this is the well-known Oehme-Zimmermann condition [242, 243].

A treatment of the analytic low-frequency behavior of continuous parts of the spectral functions has been initiated in [127]. In particular, it was shown that for correlation functions obeying a KL representation, a simple relation between the IR asymptotics of the correlator and its spectral function can be derived by differentiating with respect to the frequency. For the Taylor coupling, we explicitly find

$$\lim_{\omega \to 0^+} \partial_\omega \rho_\alpha(\omega) \,=\, -2 \lim_{p \to 0^+} \partial_p \frac{\alpha_s(p)}{p^2} \,. \tag{5.16}$$

Hence, if the coupling approaches zero in the origin faster than p^2, we expect the spectral function to approach zero from below, and vice versa.

For the gluon and the ghost propagators, the leading order IR and UV asymptotics are known analytically; see [127] and references therein. In the infrared, the decoupling solution of the ghost is characterized by a constant propagator dressing $Z_c \equiv Z_c(p=0)$. On the other hand, the gluon propagator is dominated by the ghost loop polarization diagram in the IR, since the gluon propagator itself has a mass gap and decouples in the infrared. This results in a $p^2 \log p^2$ contribution in the IR regime; for a detailed discussion thereof, see [127].

Using the definition of the strong coupling (5.10), we see that it has the same, but negative, IR behavior as the inverse gluon dressing, up to a constant contribution from the ghost dressing. From (5.16), we can then infer the asymptotic behavior of the spectral function as

$$\rho_{\alpha,\mathrm{IR}}(\omega) \sim \omega^2 \,, \tag{5.17}$$

analogously to [127]. The UV asymptotic behavior of the strong coupling is well-known from perturbative calculations and reads

$$\alpha_{s,\mathrm{UV}}(p) \sim \frac{1}{\log(p^2)} \,. \tag{5.18}$$

The asymptotic behavior of the spectral function follows directly from (5.13), and we obtain

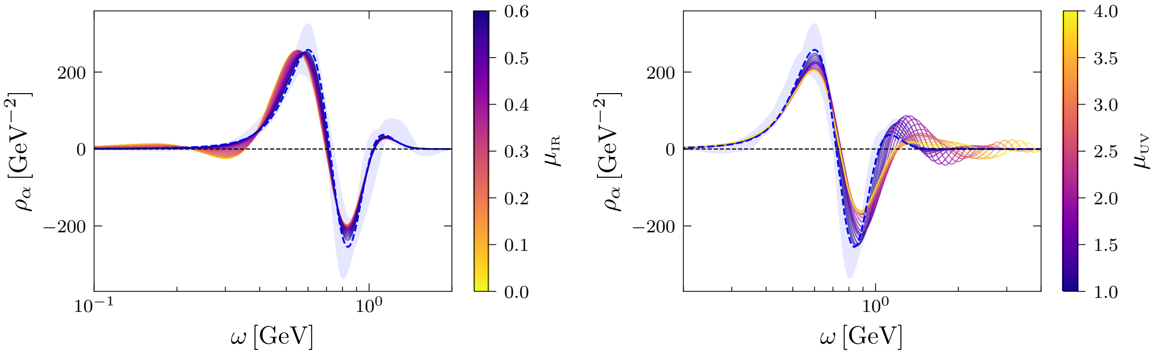
$$\rho_{\alpha,\mathrm{UV}}(\omega) \sim -\frac{1}{\omega^2(\pi^2 + \log(\omega^2)^2)} \,. \tag{5.19}$$

5.1.3 Lattice Data

During the past two decades, lattice QCD results for Landau gauge two-point functions have advanced to an impressive quantitative level of precision; see, e.g. [244–259]. A recent review of lattice and functional results can be found in [260]. The lattice ghost dressing function and gluon propagator data used in this work have been obtained from recent calculations with 2+1 dynamical fermion flavors at the physical point [233, 234]. In particular, the ensembles of gauge configurations were generated by the RBC/UKQCD collaboration in [261–265], leveraging the Iwasaki gauge action [266] and the domain wall fermion action [267, 268] with a pion mass of 139 MeV.

The continuum limit of the lattice data is only obtained with proper treatment of discretization effects. For the Landau gauge propagators, this is achieved by an analysis of the physical scaling violation as described in [269], leading to continuum extrapolated propagators with the correct momentum running. The resulting gluon propagator and ghost dressing data are displayed in the insets of Figure 5.3. These data combined with results from functional Yang-Mills theory and QCD [226, 235, 17, 236] have also been reconstructed in [187].

Since the lattice data for the propagators are available only on different momentum grids, the coupling as defined in (5.10) is computed from interpolations of the respective dressings. These are obtained by direct GP reconstruction and therefore assume no general features of the underlying correlators apart from continuity. We compute the coupling including errors for 600 logarithmically spaced points between 0.23 GeV and 2.69 GeV. For technical convenience, the coupling is extended perturbatively at large momenta in order to control the amplitude of the UV asymptotics; see Section 5.1.2. A subset of these data is shown in Figure 5.2a. Here, we replace the error with the difference between the values computed as described above, and the coupling obtained from the product of the ghost and gluon spectral functions, described around (5.14) and in Section 5.2.



(a) Scan of ρ_α along the flat μ_{IR} direction

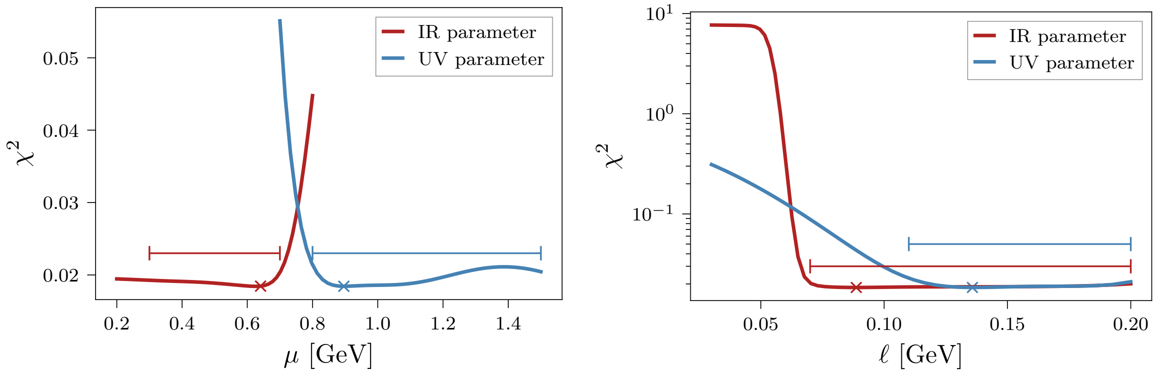
(b) Scan of ρ_α along the flat μ_{UV} direction

Figure 5.4: Behavior of the spectral function when varying the midpoints $\mu_{\mathrm{IR/UV}}$ of the transition kernels to the asymptotic IR (a) and UV regimes (b). The respective values of the parameters are color-coded. The resulting scan of the spectral functions is compared to the final result with maximally enhanced asymptotics, displayed with a dashed blue line. The error band obtained by varying the parameters of the asymptotics – as indicated in Figure 5.5 – is given by the shaded blue area.

5.2 RESULTS

Our main result, the spectral function of the Taylor coupling (5.13) in QCD, is displayed in Figure 5.2b. It shows two variants: $\rho_\alpha^{\mathrm{GP}}$ from the reconstruction of the lattice QCD data via GP reconstruction, and $\rho_\alpha^{\mathrm{spec}}$ from the direct calculation based on the spectral representations of ghost and gluon propagators (5.14). The associated input spectral functions are shown in Figure 5.3. In this context, we have improved the reconstruction of the gluon propagator reported in [187] by explicitly incorporating the known IR and UV asymptotics with the method described in Section 3.3.3. The error band of $\rho_\alpha^{\mathrm{spec}}$ is obtained by propagating the errors of these input data. Importantly, the coupling spectral functions obtained via these two different approaches agree well within errors and share all qualitative features, such as peaks and asymptotic behavior. In both results, we can identify two prominent peaks of similar size in positive and negative direction at roughly $\sim 0.6\,\mathrm{GeV}$ and $\sim 0.8\,\mathrm{GeV}$, along with a smaller positive peak at $\sim 1.1\,\mathrm{GeV}$. The spectral function $\rho_\alpha^{\mathrm{spec}}$ (5.14) allows for a direct interpretation of this behavior: it is connected to the peak structure of the gluon spectral function, which carries information about the gluon mass gap; see Figure 5.3b. This information is extracted reliably from the lattice data with the GP reconstruction.

In the reconstruction of the coupling, the correct asymptotic behavior is enforced by smooth step functions at transition points μ_{IR} and μ_{UV}, while fully retaining the flexibility in the enclosed region where the GP kernel remains unrestricted and universal, see Section 3.3.3 for details. This procedure has also been applied to the reanalysis of the gluon spectral functions used here. It significantly enhances the stability and reliability of the prediction by connect-

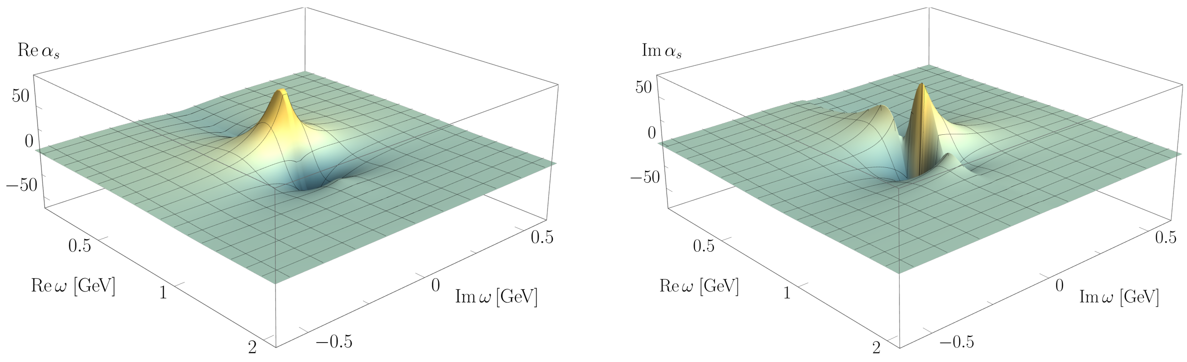


(a) χ^2 of the reconstruction under the variation of $\mu_{\mathrm{IR/UV}}$.

(b) χ^2 of the reconstruction under the variation of $\ell_{\mathrm{IR/UV}}$.

Figure 5.5: Scans of the bias parameters defined in (3.59). We compare the quality of the dressing reconstruction – quantified by χ^2 – when varying the midpoint positions of the bias transition $\mu_{\mathrm{IR/UV}}$ (a) as well as its steepness $\ell_{\mathrm{IR/UV}}$ (b). The values of the bias parameters chosen for the reconstruction are marked by crosses. This choice maximizes the size of the regions dominated by the coupling infrared and ultraviolet asymptotics while producing small χ^2 reconstructions of the data. Additionally, the parameters are then scanned in the flat directions, indicated by the horizontal bars, in order to obtain the error estimation for the reconstruction results shown in Figure 5.2.

ing it to analytic results at low and high frequencies, ensuring agreement with functional and perturbative results in the relevant limits without reducing the expressivity of the GP model in the domain of interest. While the prediction shows some variation with the choice of the transition midpoints, the peak positions and heights remain remarkably stable; see Figure 5.4. The kernel hyperparameters are optimized as follows: First, the parameters of the squared exponential are optimized by minimizing the NLL, cf. (3.31), of the GP without any asymptotics. Then, the asymptotics, specified in Section 5.1.2, are added to the kernel. The parameters of the asymptotic kernel are then optimized by minimizing the χ^2 of the reconstruction while keeping the parameters of the squared exponential fixed. We observe, that the χ^2 of the reconstruction remains relatively flat for a large range of parameters and has a shallow minimum, before rapidly increasing, see Figure 5.5. We therefore choose the optimal parameters to be in the minimum of the χ^2 scan, indicated by the crosses in Figures 5.5a and 5.5b respectively. As an error estimate, we vary the asymptotic parameters over the region, where the χ^2 remains small, indicated by the horizontal bars in Figures 5.5a and 5.5b. Changing the parameters controlling the transition to the asymptotic behavior accounts for the majority of the variation in the spectral function while changing the parameters of the RBF kernel produces errors at least one order of magnitude smaller. We therefore do not include the latter error in the final error band. The numerical values of all kernel hyperparameters are listed in Table 5.1. We can observe, that the parameters tend to



(a) Real part of the Taylor coupling $\alpha_s(\omega)$.

(b) Imaginary part of the Taylor coupling $\alpha_s(\omega)$.

Figure 5.6: Taylor coupling $\alpha_s(\omega)$ of 2+1 flavour QCD defined in (5.10) in the complex frequency right half plane (positive real frequencies), real (a) and imaginary part (b). The imaginary part explicitly shows the branch cut along the real frequency axis. The spectral function corresponds to the imaginary part of α_s at the upper half plane boundary of the branch cut, divided by ω^2. Both, the real and imaginary parts, exhibit distinctive peaks that can be connected to the peak structure of the gluon spectral function; see Figure 5.3b. The coupling decays logarithmically for increasing $|\omega|$.

Parameter	σ	l	μ_{UV}	ℓ_{UV}	μ_{IR}	ℓ_{IR}
Value	67.399	0.074	0.890	0.137	0.637	0.090

Table 5.1: Hyperparameters for the combined RBF and fixed-asymptotics kernel, as defined in (3.25) and (3.59).

favor values maximize that the region of asymptotic behavior. Accordingly, the size of the dynamical region carrying information about the QCD mass gap is minimized, supporting the gluonic quasi-particle picture employed in various applications such as bound state studies and transport computations. Specifically, this suggests dismissing smaller negative peaks close to the dominant quasi-particle peak – they merely reflect the asymptotic behavior and the superconvergence condition (5.15). As such, they are sensitive to changes in the gauge fixing parameter and infrared closure. This suggests that they carry physically relevant information only on a subleading level.

In Figure 5.2a, we compare the reconstructed Euclidean Taylor couplings to the result computed from the lattice data for the ghost and gluon propagators, as described in Section 5.1.3. Using the dressing function data obtained in this way, the resulting coupling is shown to decay towards small and large momenta. In correspondence to the scale of the peaks of the spectral function – reflecting the mass gap of the theory – also the peak of the coupling itself appears at $\sim 0.6\,\mathrm{GeV}$.

The blue curve in Figure 5.2a represents the GP reconstruction of the Taylor coupling lattice data, corresponding to $\rho_\alpha^{\mathrm{GP}}$. The red curve represents the coupling obtained via its spectral representation (5.12) using the directly computed spectral function $\rho_\alpha^{\mathrm{spec}}$. The calculation involves finite precision, both in the input data and in the integration. Hence, we expect a small, but not negligible, relative error. The decent agreement between this result and the lattice/GP reconstruction result provides a highly non-trivial benchmark check. The error is well within our expectations since the result obtained from the directly computed spectral function depends on the reconstructions of the gluon and ghost propagators. If the ghost and gluon spectral functions were describing their respective propagator data to infinite precision, we would also expect perfect agreement from analytic considerations. Hence, the small difference can be attributed to systematic uncertainties present in the calculation. Please note that they do not contribute to the error bands, corresponding to the purely statistical error, shown in Figure 5.2.

In the inset of Figure 5.2a, we also show the Taylor coupling divided by p^2 for small Euclidean momenta p. The derivative of this quantity is connected to the asymptotic behavior of the spectral function in the IR by (5.16). We observe that in the region where lattice data are available, the slope of α_s/p^2 is negative. In accordance with the analytic requirement (5.16), the slope of the spectral function is observed to be positive in this regime.

Finally, in Figure 5.6 we display the real and imaginary parts of the coupling in the full complex momentum plane. The data are obtained by evaluating the coupling spectral representation (5.12) with the directly calculated spectral function $\rho_\alpha^{\mathrm{spec}}$ in the complex plane. The branch cut in the imaginary part, responsible for the spectral representation, is clearly visible. As expected, no further non-analyticities in the complex plane are encountered and the coupling shows the expected decay behavior towards large frequencies.

5.3 Conclusion

In this chapter, we have presented results for the spectral function of the strong coupling constant in QCD obtained through a direct calculation as well as a reconstruction via Gaussian processes. Assuming spectral representations for the ghost and gluon, we have derived the spectral representation of the Taylor coupling, which is fully determined by the ghost and gluon dressing functions. With this relation, we have calculated the associated spectral function as well as the coupling itself in the full complex plane; see Figures 5.2 and 5.6. The required ghost and gluon spectral functions have been obtained using the same reconstruction method, explicitly taking into account the known asymptotic IR and UV behavior of the gluon spectral function; see Figure 5.3. This modification substantially improves the reliability of the approach by properly encoding the analytically tractable regimes into the prediction while preserving the expressivity and universality of the GP model in the region of interest.

A comparison of the results from the direct calculation and GP reconstruction shows excellent agreement between both approaches; see Figure 5.2b. This independent verification provides strong support for the accuracy of the computed spectral function and also underlines the power of probabilistic inversion with GP models as a spectral reconstruction approach. In particular, the findings demonstrate that uncertainty estimates obtained within this framework are reasonable, allowing to reliably quantify the expected errors in potential downstream applications based on reconstruction results. Possible future extension can entail to also analyze the quark-gluon vertex coupling directly based on available lattice data; see e.g. [270].

Our results find direct application in the calculation of non-perturbative, physical scattering processes, where the strong coupling constant needs to be known at timelike momenta. While neglecting angular dependencies, the Taylor coupling considered here carries the correct RG running and hence scale-dependence of the strong coupling constant. Furthermore, it encodes genuine non-perturbative information through the input ghost and gluon dressing functions obtained from 2+1 flavor lattice QCD. This work hence paves the way for incorporating non-perturbative information from lattice field theory to functional methods in the calculation of timelike scattering processes.

6. Thermal Photon Rate

This chapter is based on parts of the publication [3].

In this chapter, we will present results from reconstructions of the thermal photon rate from the quark-gluon plasma (QGP). As has been discussed in Chapter 1, above the pseudo-critical temperature of $T_{\mathrm{pc}} = 156.5 \pm 1.5\,\mathrm{MeV}$ [6–8], QCD matter can be described by a quark-gluon plasma (QGP), where quarks and gluons are the relevant degrees of freedom. Since the photons and dileptons that are produced in heavy-ion collisions travel from the production point to the detector with relatively little interaction with the medium, they can provide valuable information about the properties of the QGP [271]. Importantly, they carry local information about temperature and density during the QGP phase. However, photons and dileptons are produced not only at high energies but also at lower energies, where the formation of a QGP is not expected These electromagnetic probes are produced in collisions of charged particles or from hadronic decays such as π^0, $\eta \to \gamma\,\gamma$ and contribute to a large portion of photons measured in experiment [272].

Experimental setups for heavy-ion collisions, namely Au-Au collisions at PHENIX [273] and Pb-Pb collisions at ALICE [274] have revealed large yields of this hadronic background. These *direct* photons have to be separated from the *thermal* photons that are produced in the QGP in order to extract the properties of interest [275]. Especially at high transverse momentum p_T, these direct photons dominate the spectrum. However, at low p_T an excess of photons is observed, which is attributed to in-medium effects from the QGP. Measurements of direct photons and dileptons also reveal a large azimuthal anisotropy, quantified by v_2 and often referred to as elliptical flow [276–279]. Connecting theoretical calculations to these dynamic observables has been a challenge for many years [280]. Simulations have revealed that the 'slopes' of the p_T spectrum (for photons) and the invariant mass spectrum (for dileptons) can be systematically related to an average temperature of the fireball [281–283]. The anisotropic flow of direct photons and dileptons has been shown to be sensitive to the initial conditions and shear viscosity [284–286]. Such calculations require the differential photon or dilepton production rate at every stage of the evolution as an input to hydrodynamic simulations. Since hydrodynamic expansion assumes local thermal equilibrium, such hydrodynamic quantities can be extracted from equilibrium field theory calculations at the appropriate temperatures.

Numerous studies have investigated the thermal photon rate, both perturbatively and non-perturbatively, using lattice techniques. Perturbative results were established in [287, 288], and later extended to establish the full $\mathcal{O}(\alpha_{\text{em}}\alpha_s)$ result [289, 290]. More recently, next-to-leading-order results have been obtained in [291].

For lattice QCD calculations, the challenge is similar to other hydrodynamic observables: they require real-time information about QCD correlators and are therefore not directly calculable on the lattice. As in the sections above, we therefore apply spectral reconstruction techniques to obtain the thermal photon rate. In a previous study [123], the photon rate was estimated using the vector channel correlator, wherein the UV part of the spectral function was calculated perturbatively, while the IR part was fixed using lattice data. Since the UV tail diverges, the information about the relevant region of the spectral function is extremely hard to quantify. Therefore, a new correlator was proposed in [292], based on the difference between the transverse and longitudinal channels. Such a construction leads to a correlator that is free of the UV divergence, and the spectral function can be calculated in a more controlled way. A similar method was employed in another study [293], where the photon rate was estimated from the transverse channel correlator. Recently, a novel idea has been proposed for calculating the photon rate. In this method, the photon production rate is studied using the imaginary momentum correlator, which does not require spectral reconstruction [294, 295]. With this method, the first two moments of the spectral function are calculated at lightlike kinematics in Euclidean spacetime, which is realized by employing an imaginary spatial momentum, and the photon rate is estimated from these moments. These can provide further constraints on the photon rate in the future.

Here, following [292], we estimate the thermal photon rate from the QGP using the difference between the transverse and longitudinal correlators. For the spectral reconstruction, we present various techniques, ranging from physics-informed model fits and the Backus-Gilbert method to Gaussian processes. Here, we will focus mainly on the reconstructions using Gaussian processes and compare them with other methods. With these methods, we calculate the thermal photon rate from the lattice correlator for light quarks in two different scenarios: Quenched QCD at a temperature of $470\,\text{MeV}$ $(1.5\,T_c)$, and (2+1)-flavor QCD at a temperature of $220\,\text{MeV}$ $(1.22\,T_{pc})$. For the quenched QCD scenario, there is continuum extrapolated correlator data, which means that our prediction for the photon production rate is also at the continuum. For the full QCD scenario, our results are at a finite lattice spacing. We investigate the dependence of the photon rate on the lattice spacing for quenched QCD and from this infer its dependence on the lattice spacing for full QCD.

This chapter is organized as follows. After introducing the thermal photon rate and its connection to the spectral function in Section 6.1, we will discuss the lattice details in Section 6.2. In Section 6.3, we will present the spectral reconstruction methods that are used to reconstruct the thermal photon rate from this data and compare the results in Section 6.4.

6.1 THERMAL PHOTON RATE FROM QCD SPECTRAL FUNCTIONS

The thermal photon rate is defined as the number of photons radiated from the plasma per unit time and per unit volume at a fixed temperature T. In the thermalized plasma, the photon production is due to the $N_f = 3$ degenerated quark flavors that couple to the photons via the electromagnetic current

$$J^{\mu} = e \sum_f Q_f \bar{\psi}_f \gamma^{\mu} \psi_f \,, \tag{6.1}$$

where Q_f is the quark charge fraction and ψ_f is the quark field of flavor f. The square of the S-matrix element of a transition of an initial hadronic state to a final hadronic state with the emission of a photon is then related to the correlator of such a current [296], i.e.$\langle J^{\mu}(x) J^{\nu}(0) \rangle_{\beta}$.

If we consider the three flavor case, $N_f = 3$, where $f = u, d, s$, then $\sum_f Q_f = 0$ and $\sum_f Q_f^2 = \frac{2}{3}$. Under these conditions, to leading order in the QED coupling α_{em}, this rate can be calculated as [297]

$$\frac{\mathrm{d}\Gamma_{\gamma}}{\mathrm{d}^3\mathbf{k}} = -\frac{\alpha_{\text{em}} n_b(k)}{2\pi^2 k} \Big\{ \textstyle\sum_{i=1}^{N_f} Q_i^2 \Big\} g^{\mu\nu} \rho_{\mu\nu}\big(\, \omega = |\mathbf{k}|,\, \mathbf{k}\,\big)\,, \tag{6.2}$$

where $n_b(\omega) = 1/(e^{\omega/T} - 1)$ is the Bose distribution function. In this expression, $\rho_{\mu\nu}$ denotes the connected part of the vector current spectral function for a single flavor. Since we are interested in the thermal photon rate, and photons are massless, the spectral function is evaluated on the light cone, i.e. $\omega = |\mathbf{k}|$. Note that other interesting quantities such as the dilepton rate are frequency dependent and require the spectral function at momenta inside the light cone. Although this expression of the thermal photon rate is only valid up to leading order in the electromagnetic coupling it captures all orders in the strong coupling.

As introduced in Section 2.1.1, the spectral function $\rho_{\mu\nu}(\omega, \mathbf{k})$ can be calculated through analytic continuation, see (2.18), from the Matsubara frequency modes ($\omega_n = 2\pi T n \,,\, n \in \mathbb{Z}$), i.e. $\rho_{\mu\nu}(\omega, \mathbf{k}) = \mathrm{Im}[G^E_{\mu\nu}(\omega_n \to -i(\omega + i0^+), \mathbf{k})]$, where the Euclidean current correlator is given by

$$G^E_{\mu\nu}(\omega_n, \mathbf{k}) = \int \mathrm{d}^3\mathbf{x} \int_0^{\beta} \mathrm{d}\tau \, e^{i(\omega_n \tau - \mathbf{k}\cdot\mathbf{x})} \langle J_{\mu}(\tau, \mathbf{x}) J_{\nu}(0, \mathbf{0}) \rangle \,, \tag{6.3}$$

and $\beta = 1/T$ is the temporal extent.

On the lattice, real-time calculations are not feasible due to the sign problem, so we calculate the Euclidean correlation function $G^E_{\mu\nu}(\tau, \mathbf{k}) = T \sum_n \exp(i\omega_n \tau) G^E_{\mu\nu}(\omega_n, \mathbf{k})$. As introduced before, cf. (2.21), the correlator can be written in terms of its associated finite temperature spectral function as

$$G^E_{\mu\nu}(\tau,\mathbf{k}) \;=\; \int_0^\infty \frac{\mathrm{d}\omega}{\pi}\frac{\cosh\left(\omega(\tau-\beta/2)\right)}{\sinh\left(\omega\beta/2\right)}\,\rho_{\mu\nu}(\omega,\mathbf{k})\,. \tag{6.4}$$

One major challenge is therefore, similar to the sections before, the extraction of the spectral function from the Euclidean correlator. In the case of the extraction of the contracted (or "vector channel") spectral function $g^{\mu\nu}\rho_{\mu\nu}$, this inversion does not only suffer from the ill-conditioned nature of the problem but also from the fact that the spectral function is proportional to ω^2 in the ultraviolet domain. As a result, the lattice correlator, determined by the integration in (6.4), receives a significant contribution from the UV part of the spectral function. This presents an additional challenge for reconstructing the spectral function, as the thermal photon rate is primarily governed by its infrared component. Therefore, the interesting information about the IR part is buried under the UV tail that dominates the correlator data. When reconstructing the thermal photon rate directly from such a correlator, small deviations in the UV tail would lead to significant deviations in the interesting IR regime. This consequently produces large errors, often spanning large parts of the interesting regime, see e.g. [123–125, 298] for reconstructions encountering this obstacle.

Recently, an approach has been proposed to circumvent this problem by constructing a correlator that is free of the UV divergence [292]. The idea is to consider the difference between the transverse and longitudinal part of the spectral function, which leads to a spectral function that is heavily suppressed in the UV. This special combination, referred to as the *T-L spectral function*, has a corresponding Euclidean correlator that is governed by the IR part of the spectral function, and the actual information we want to access has a significantly larger contribution to the correlator.

In order to obtain the T-L part, the spectral function $\rho_{\mu\nu}$ is decomposed in terms of ρ_T (the transverse component) and ρ_L (the longitudinal component), leading to

$$\rho_{\mu\nu}(\omega,\mathbf{k}) = P^T_{\mu\nu}\rho_T(\omega,\mathbf{k}) + P^L_{\mu\nu}\rho_L(\omega,\mathbf{k})\,. \tag{6.5}$$

Here, $P^T_{\mu\nu}$ and $P^L_{\mu\nu}$ are the transversal and longitudinal projection operators, given explicitly by

$$\begin{aligned}
P^T_{ij}(\omega,\mathbf{k}) \;&=\; -g_{ij} - \frac{k_i k_j}{\mathbf{k}^2}\,,\\
P^T_{0i} \;=\; P^T_{i0} \;&=\; P^T_{00} \;=\; 0\,,\\
P^L_{\mu\nu}(\omega,\mathbf{k}) \;&=\; -g_{\mu\nu} + \frac{K_\mu K_\nu}{K^2} - P^T_{\mu\nu}(\omega,\mathbf{k})\,,
\end{aligned} \tag{6.6}$$

where $K = (\omega,\mathbf{k})$ is the associated four-momentum. The projection operators satisfy the following relations: $P_T^2 = P_T$, $P_L^2 = P_L$, and $P_T P_L = P_L P_T = 0$. In terms of these components, the vector channel spectral function is given by

$$g^{\mu\nu}\rho_{\mu\nu} \;=\; \rho^{\mu}_{\mu} \;=\; -2\rho_T - \rho_L\,. \tag{6.7}$$

On the light cone, i.e. $\mathbf{k}^2 = \omega^2$, we can see, that $g^{\mu\nu}P^L_{\mu\nu}(\omega = |\mathbf{k}|, \mathbf{k}) = 0$. This means, that the longitudinal part of the vector channel spectral function vanishes on the light cone, which is a manifestation of current conservation [292]. The photon production rate therefore only depends on the transverse part of the current spectral function $d\Gamma_\gamma \sim \rho_T(\omega = |\mathbf{k}|, \mathbf{k})$. This allows us to calculate the photon production using any combination of the transversal and longitudinal parts of the spectral function. A practical choice is the T-L spectral function, given by

$$\rho_H(\omega, \mathbf{k}) \;=\; 2\,[\,\rho_T(\omega, \mathbf{k}) - \rho_L(\omega, \mathbf{k})\,]\;. \tag{6.8}$$

In the zero temperature limit, which is the same as the infinite ω limit, we recover Lorentz invariance. Together with current conservation, this implies that [292]

$$\omega^2\rho^{00}(\omega, \mathbf{k}) - k^i k^j \rho^{ij}(\omega, \mathbf{k}) = 0\,. \tag{6.9}$$

Therefore, the T-L spectral function vanishes for all ω and $\mathbf{k}$ at zero temperature and displays purely thermal effects at finite temperature. At finite temperature, no new divergences are introduced and the T-L spectral function stays finite. Therefore, we can do an Operator Product Expansion (OPE): by power counting, it can be shown that in the domain where $\omega \gg k, \pi T$, this spectral function behaves asymptotically like [299, 300]

$$\rho_H(\omega, \mathbf{k}) \;\sim\; \frac{\mathbf{k}^2\langle\mathcal{O}_4\rangle}{\omega^4}\,, \quad \omega \gg k, \pi T\,, \tag{6.10}$$

where $\mathcal{O}_4$ is a dimension four operator $\mathcal{O}_4$. Another consequence of this is that this spectral function satisfies a sum rule

$$\int_0^\infty \mathrm{d}\omega\,\omega\,\rho_H(\omega) \;=\; 0\,. \tag{6.11}$$

We should note that the photon production rate (6.2) is often (equivalently) expressed as

$$\frac{\mathrm{d}\Gamma_\gamma}{\mathrm{d}^3\mathbf{k}} \;=\; \frac{\alpha_{\mathrm{em}} n_b(k)\chi_q}{\pi^2}\left\{\textstyle\sum_{i=1}^{N_f} Q_i^2\right\} D_{\mathrm{eff}}(k)\,, \tag{6.12}$$

where the effective diffusion coefficient is defined by

$$D_{\mathrm{eff}}(k) \;\equiv\; \frac{\rho_H\big(|\mathbf{k}|, k\big)}{2\chi_q|\mathbf{k}|}\,. \tag{6.13}$$

Here χ_q is the quark number susceptibility. In the hydrodynamic regime of large wavelengths, $D_{\mathrm{eff}}(k)$ approaches the well-known diffusion coefficient D, i.e. $\lim_{k\to 0} D_{\mathrm{eff}}(k) = D$.

6.2 Lattice Details

Here, we briefly describe the lattice details for the calculation of the T-L propagator. For a more detailed discussion, we refer to the accompanying publication [3].

In lattice calculations, space-time is discretized and has a finite spatial extent (L). Therefore, the (spatial) momenta are also discrete and given by $\mathbf{k} = \left(\frac{2\pi n_x}{L}, \frac{2\pi n_y}{L}, \frac{2\pi n_z}{L}\right)$, where $(n_x, n_y, n_z) \in \mathbb{Z}$. On the lattice, we then calculate the correlator in different directions, i.e. G_{xx}, G_{yy}, G_{zz}, and $G_{\tau\tau}$. If the momentum is in the x direction i.e. $\mathbf{k} = (k, 0, 0)$, then the correlator for ρ_H is given by [300]

$$\begin{aligned} G_H(\tau, k_x = k) =& G_{yy}(\tau, k_x = k) + G_{zz}(\tau, k_x = k) \\ &- 2\Big(G_{xx}(\tau, k_x = k) - G_{\tau\tau}(\tau, k_x = k)\Big) . \end{aligned} \tag{6.14}$$

This correlator is symmetric under $\tau \to \beta - \tau$ and invariant under rotations, i.e. $G_H(\tau, k_x = k) = G_H(\tau, k_y = k) = G_H(\tau, k_z = k)$. Therefore, we can compute the T-L correlator as the average of all these different combinations as

$$G_H(\tau, k) = \frac{1}{6}\Big(G_H(\tau, k_x = k) + G_H(\tau, k_y = k) + G_H(\tau, k_z = k) + (\tau \to \beta - \tau)\Big) . \tag{6.15}$$

We calculate this correlator for two different scenarios: quenched QCD at a temperature of 470 MeV (1.5 T_c), and (2+1)-flavor QCD at a temperature of 220 MeV (1.22 T_{pc}). An overview of the lattice parameters can be found in Table 6.1.

The quenched configurations are generated using a standard Wilson gauge action with heat-bath and overrelaxation updates. The correlator is then calculated using Clover-improved Wilson fermions and the scale has been set by $r_0 = 0.47$ fm [301] and the T_c is obtained by $T_c = 0.7457/r_0$ [302]. The pure gauge configurations are generated for three different lattice spacings, which allows us to perform a continuum extrapolation of the correlators.

In the case of full QCD, we use gauge field configurations generated by the HotQCD collaboration using Highly Improved Staggered Quarks (HISQ) [303] and a tree-level improved Lüscher-Weisz gauge action [304, 305]. The hopping parameter value κ is tuned to match the pion mass, which in this study is unphysical and equal to 320 MeV with $m_l = m_s/5$. The scale is determined by $r_1 = 0.3106$ fm and $T_{pc} = 180$ MeV is obtained from disconnected chiral susceptibility, cf. [306], as in [298].

6.3 Spectral Reconstruction

In this section, we present the spectral reconstruction techniques used to obtain the T-L spectral function from the lattice correlator. Since the reconstruction problem is ill-conditioned, there are generally large uncertainties in the reconstructed spectral function. As discussed earlier, this issue can be addressed using inversion methods that regularize the problem, such

N_f	T	β_0	κ	$N_\sigma^3 \times N_\tau$	confs	a^{-1} [GeV]
0	$1.5\,T_c$	7.394	0.13407	$120^3 \times 30$	1950	14.1
		7.192	0.13440	$96^3 \times 24$	2000	11.3
		7.035	0.13467	$80^3 \times 20$	1824	9.39
$2+1$	$1.2\,T_{pc}$	8.249	0.13515	$96^3 \times 32$	1750	7.04

Table 6.1: Details of the lattices studied, where β_0 is the coefficient of the Wilson plaquette and κ is the hopping parameter. The critical temperature $T_c = 313\,\text{MeV}$ for the quenched, $N_f = 0$, case and the pseudo critical temperature is $T_{pc} = 180\,\text{MeV}$ for the QCD, $N_f = 2+1$, case.

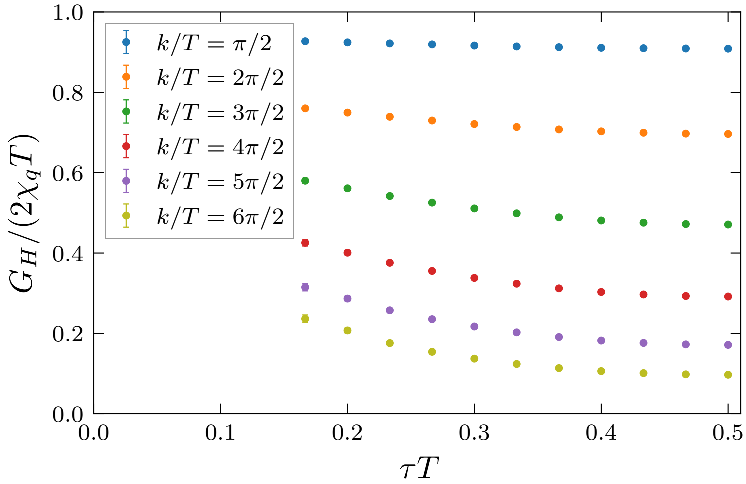


Figure 6.1: Continuum extrapolated lattice data for the Euclidean T-L correlator G_H for the quenched, $N_f = 0$, case at $1.5\,T_c$ for different spatial momenta k/T.

as Tikhonov regularization, or by fitting a functional form to the data. The drawback of such regularization techniques is that they introduce bias, which can result in significant systematic uncertainties that are often difficult to quantify. For instance, when fitting a functional model to the data, obtaining a realistic estimate of systematic uncertainty would require considering a wide range of models or using a model with many parameters. However, these approaches are typically impractical, as they are either susceptible to overfitting or computationally demanding.

Instead of doing an error analysis of the systematic uncertainties, we compare different reconstruction methods and use the spread of the results as an estimate of the systematic uncertainties. In this work, we compare three different methods: two different physics-informed model fits, the Backus-Gilbert method, and the Gaussian process regression. Note that while the fits and the Backus-Gilbert method do not provide systematic uncertainties, the recon-

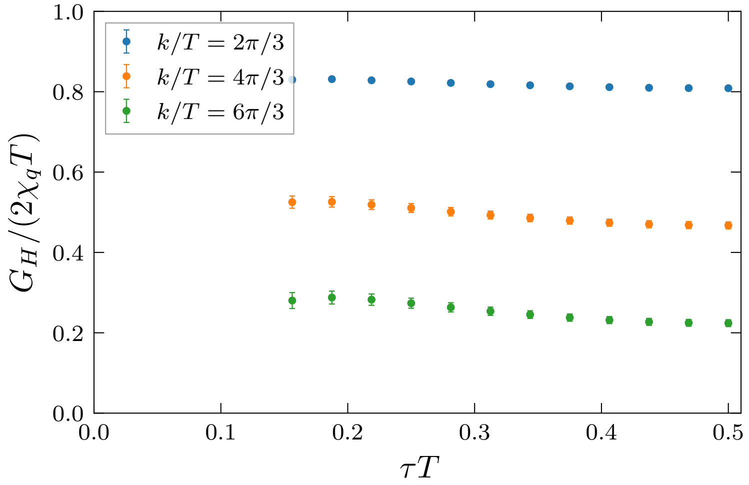


Figure 6.2: Lattice data for the Euclidean T-L correlator G_H for the full QCD, $N_f = 2+1$, case at $1.22\,T_{pc}$ at finite lattice spacing for different spatial momenta k/T.

struction with GPs covers the full uncertainty of the reconstruction. This will be discussed in greater detail in Section 6.3.3. Before that, we will briefly introduce the model fits and the Backus-Gilbert method.

6.3.1 PHYSICS-INFORMED MODEL FITS

We consider two different model fits to the lattice data. Firstly, a polynomial fit, that separates the spectral function into two parts separated at some fixed ω_0. The ultraviolet region $\omega \geq \omega_0$ is modeled by $\rho^{\text{poly}}_{\text{UV}}$, which is constructed from inverse even powers of ω (starting with $1/\omega^4$) in order to satisfy the large ω OPE result (6.10). The infrared region $\omega \leq \omega_0$ is modeled by $\rho^{\text{poly}}_{\text{IR}}$, where we adopt the same polynomial put forward in [123]. The two regions are matched continuously

$$\rho^{\text{poly}}_H(\omega) \;=\; \rho^{\text{poly}}_{\text{IR}}(\omega)\,\Theta(\omega_0-\omega) + \rho^{\text{poly}}_{\text{UV}}(\omega)\,\Theta(\omega-\omega_0)\,, \tag{6.16}$$

where Θ is the Heaviside step function. The two parts of the spectral function are given by

$$\rho^{\text{poly}}_{\text{IR}}(\omega) = \frac{\beta\omega^3}{2\omega_0^3}\left(5-3\frac{\omega^2}{\omega_0^2}\right) - \frac{\gamma\omega^3}{2\omega_0^2}\left(1-\frac{\omega^2}{\omega_0^2}\right) + \frac{\delta\,\omega}{\omega_0}\left(1-\frac{\omega^2}{\omega_0^2}\right)^2\,, \tag{6.17}$$

$$\rho^{\text{poly}}_{\text{UV}}(\omega) = -\frac{\beta\,\omega_0^4}{7\,\omega^4}\left(54\frac{\omega_0^4}{\omega^4} - 94\frac{\omega_0^2}{\omega^2} + 33\right) + \frac{\gamma\,\omega_0^5}{140\,\omega^4}\left(-81\frac{\omega_0^4}{\omega^4} + 92\frac{\omega_0^2}{\omega^2} - 11\right) \tag{6.18}$$

$$-\frac{16\,\delta\,\omega_0^4}{35\,\omega^4}\left(1-\frac{\omega_0^2}{\omega^2}\right)^2\,, \tag{6.19}$$

where β, γ, and δ are free parameters. Since the spectral function is expected to be smooth and differentiable across the lightcone [307], the two parts of the spectral function are chosen to match at ω_0 and its first derivative is continuous. The parameter β is the value of the spectral function at ω_0, while γ is the value of the first derivative at ω_0. The slope at $\omega = 0$ is controlled by δ, and therfore $\delta > 0$. Finally, the error of the reconstruction is obtained by varying $\omega_0 = \sqrt{k^2 + \nu(\pi T)^2}$, between different values of $\nu = \{1, 2, 5\}$.

The second model of the T-L spectral function is a Padé-like ansatz that has already been applied to the reconstruction problem in [292] and therefore allows for a direct comparison between the different lattice computations. This spectral function model is given by

$$\rho_H^{\text{Padé}}(\omega) \;=\; \frac{A\tanh(\omega\beta/2)\big(1 + B\omega^2\big)}{(\omega^2 + a^2)\big[(\omega + \omega_0)^2 + b^2\big]\big[(\omega - \omega_0)^2 + b^2\big]}\,. \tag{6.20}$$

In the infrared regime this model reproduces the hydrodynamic expectation $\rho_H(\omega) \sim \frac{A\omega}{\omega^2 + a^2}$ [74]. The remaining part of the spectral function is inspired by qualities from AdS/CFT (see e.g. [308]) and is consistent with the OPE asymptotics (6.10) at large ω. In order to satisfy the sum rule, B becomes a function of ω_0, a and b (here the parameter ω_0 is not the same as the one from (6.16)). The fit then is performed with respect to A, a, b, ω_0 by minimizing the uncorrelated χ^2 in each of the bootstrap samples. The error of the reconstruction is obtained by the largest and smallest values of D_{eff} obtained from the different bootstrap samples.

For more details on the model fits, we refer to [3], where different reconstructed spectral functions are shown, and compared to the perturbative predictions.

6.3.2 BACKUS-GILBERT METHOD

As a second approach to the reconstruction, we use the Backus-Gilbert (BG) method. This method has been, similarly to Gaussian processes, developed in the context of geophysics [128]. The BG method is an inversion technique that allows for the reconstruction of a function from noisy data, and has been commonly applied to QCD problems in the past, see the discussion in Section 3.1. In order to regularize the ill-conditioned problem, the idea is to compute a smeared estimator for the spectral function $\rho^{\text{BG}}(\omega, \mathbf{k})$, by introducing a set of weights $\mathbf{q}(\omega, \mathbf{k})$ such that

$$\sum_i q_i(\omega, \mathbf{k}) G_H(\tau_i, \mathbf{k}) \;=\; \int_0^\infty d\omega'\, \rho_H(\omega', \mathbf{k}) \sum_i q_i(\omega, \mathbf{k}) K(\omega', \tau_i) \;\equiv\; \rho^{\text{BG}}(\omega, \mathbf{k})\,, \tag{6.21}$$

where

$$\delta(\omega, \omega') \;=\; \sum_i q_i(\omega, \mathbf{k}) K(\omega', \tau_i)\,, \tag{6.22}$$

is called the resolution function. This means that the smeared version ρ^{BG} is obtained by convoluting the true spectral function with the resolution function. By tuning the weights $\mathbf{q}$, the implied averaging kernel can be varied, and the resolution function can be made more or less narrow. Usually, the optimization objective for finding these weights is to bring the smeared estimator as close as possible to the true spectral function. This is achieved by

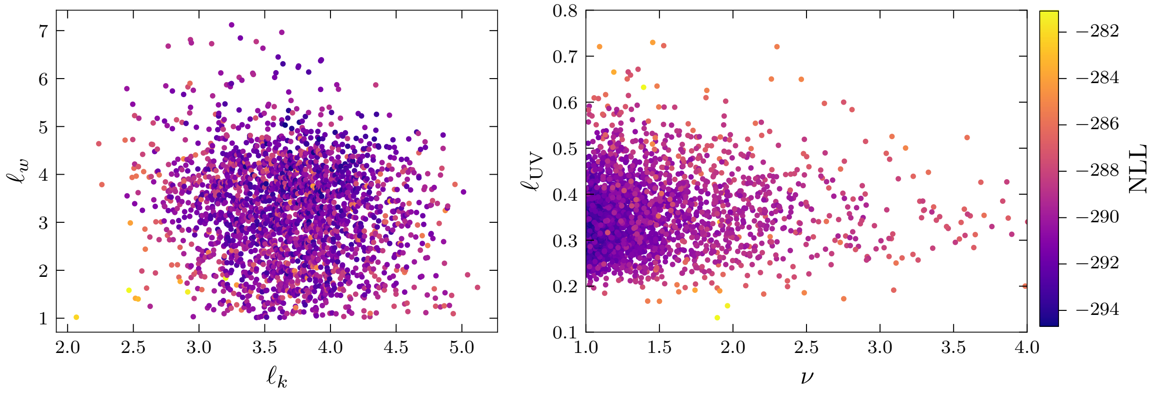


Figure 6.3: Projection of the Metropolis-Hastings scan of the negative log-likelihood for the GP reconstruction of the T-L spectral function. The scan is performed in the five-dimensional parameter space of the GP-kernel. Here, we only show the projections onto the ℓ_ω-ℓ_k plane on the left and onto the ℓ_ω-σ plane on the right. The color coding indicates the value of the NLL. We can see a clear minimum in the length scale parameters of the squared exponential kernel, while the onset of the UV kernel, parametrized by ν, has an open direction towards 0 and an explicit cutoff has to be introduced.

bringing the resolution function as close as possible to the delta function. For this, we define some measure of the distance between the resolution function and the delta function and minimize this distance. For the explicit details, we refer to [3]. We note here, that the Backus-Gilbert method and Gaussian Process regression can be, under specific choices of the respective kernels and distance metric, equivalent [168, 189].

In order to include information about the asymptotics of the spectral function, we rescale the Källén-Lehmann kernel $K(\omega,\tau)$ by a factor $f(\omega,\mathbf{k})$ such that

$$G_H(\tau,\mathbf{k}) \;=\; \int_0^\infty \mathrm{d}\omega\, \frac{\rho(\omega,\mathbf{k})}{f(\omega,\mathbf{k})}\, \tilde{K}(\omega,\tau)\,, \tag{6.23}$$

where

$$\tilde{K}(\omega,\tau) \;\equiv\; \frac{f(\omega,\mathbf{k})\cosh\left(\omega(\beta/2-\tau)\right)}{\pi\sinh(\omega\beta/2)}\,. \tag{6.24}$$

The function f incorporates prior information about the asymptotic behavior of the spectral function and is therefore chosen as

$$f(\omega,\mathbf{k}) \;=\; \left(\frac{\omega_0}{\omega}\right)^4 \tanh\left(\frac{\omega}{\omega_0}\right)^5\,, \tag{6.25}$$

where we take $\omega_0 = \sqrt{k^2+\nu(\pi T)^2}$ as for the model fits and we vary $\nu = \{0,\,0.5,\,1,\,2,\,5,\,8\}$. Again, the explicit spectral functions can be found in [3].

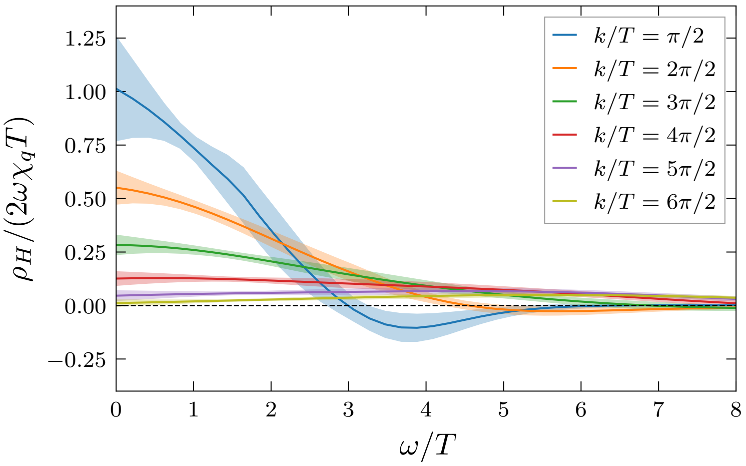


Figure 6.4: Reconstructed spectral functions for the quenched, $N_f = 0$ case at $1.5\,T_c$ using Gaussian processes. The associated lattice data is shown in Figure 6.1. The shaded area represents the error of the reconstruction, which is a combination of the propagated error on the lattice data and the systematic error from the GP reconstruction.

6.3.3 RECONSTRUCTION WITH GAUSSIAN PROCESSES

Since we have data for several spatial momenta, the reconstruction of the spectral function is not only performed in the ω, but also simultaneously in the k-direction. This ensures continuity in both directions and consequently increases the stability of the reconstruction. However, apart from continuity, we do not assume any additional structure of the spectral function in the momentum direction. For the reconstruction, we include the lattice data for all the available spatial momenta as well as the sum rule (6.11).

Similarly to the previous reconstructions, the squared exponential kernel (3.25) is used. In order to include the known UV asymptotics, $\rho_H^{\mathrm{UV}}(\omega) \sim 1/\omega^4$, the GP-kernel is modified in the UV regime to restrict the functional basis to the asymptotic behavior resulting in a UV part of the spectral function given by

$$k_{\mathrm{UV}}(\omega,\omega') \;=\; (\omega\,\omega')^{-4}\,. \tag{6.26}$$

The transition from the universal squared exponential kernel to the UV kernel is controlled by smooth step functions, as described in Section 3.3.3. Combining the squared exponential kernel in two dimensions with the asymptotic kernel results in a total number of five GP-kernel parameters: ℓ_ω, ℓ_k, σ, ℓ_{UV}, ν. The lenghtscale in k and ω direction is controlled by ℓ_k and ℓ_ω respectively, while σ gives a prior estimate on the variance of the GP. The position of the transition from the squared exponential kernel towards the UV kernel is controlled by $\omega_0 = \sqrt{k^2 + \nu(\pi T)^2}$, where ν is varied, in analogy to the polynomial fit ansatz, and ℓ_{UV} is the smoothness of the transition between the two kernels.

Again, these parameters are optimized by minimizing the associated negative log-likelihood

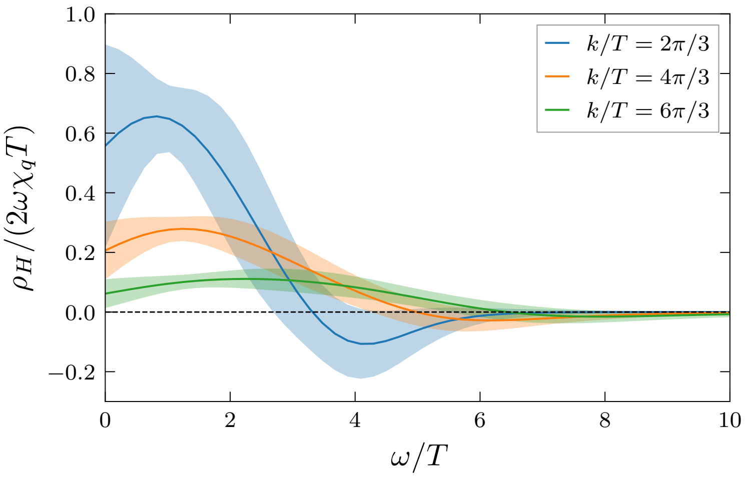


Figure 6.5: Reconstructed spectral functions for the full QCD, $N_f = 2+1$ case at $1.22\,T_{pc}$ using Gaussian processes. The associated lattice data is shown in Figure 6.2. The shaded area represents the error of the reconstruction, which is a combination of the propagated error on the lattice data and the systematic error from the GP reconstruction.

as introduced in Section 3.2.2. When optimizing the parameters for the one-dimensional GP, i.e. for one fixed spatial momentum, we notice that the overall magnitude σ of the kernel decreases with $1/k^2$, while the other parameters fluctuate, but do not have a qualitatively different behavior. We therefore rescale the magnitude parameter for the two-dimensional reconstruction as $\sigma \to \sigma/k^2$. Since this optimization is performed in a five-dimensional parameter space and the parameters are not fully independent, this optimization generally does not converge consistently, and we find local minima in the NLL. We perform a Metropolis-Hastings sampling of the NLL in the five-dimensional parameter space, see Figure 6.3 for a projection of the parameter scan on two different planes. The parameter space is limited to values of $\nu \geq 1$, since the perturbative behavior is not expected to reach this far into the IR. However, these solutions are still represented by the universal squared exponential kernel; we simply aim to avoid imposing overly restrictive constraints on the kernel in this region of the spectral function. Additionally, the length scale of the transition is also restricted from below, $\ell_{\mathrm{UV}} \geq 0.1$, in order to avoid sharp edges in the reconstructed spectral function at the transition point. These open directions in the parameter space can be attributed to the inversion problem being ill-conditioned. Additionally, other choices of the transition function might be more suitable for the transition, especially functions that become exactly zero faster than the smooth step function (3.58). We find a mild dependence of the spectral function on the length scales and the magnitude of the squared exponential kernel, e.g. ℓ_ω, ℓ_k, σ, while the dependence on the asymptotic parameters, as expected, introduces the majority of the systematic error. While the error on the lattice data is propagated through the GP and is given by the covariance of the posterior predictive distribution of the GP, the systematic error is estimated by the uncertainty in the parameters while minimizing the associated NLL. The

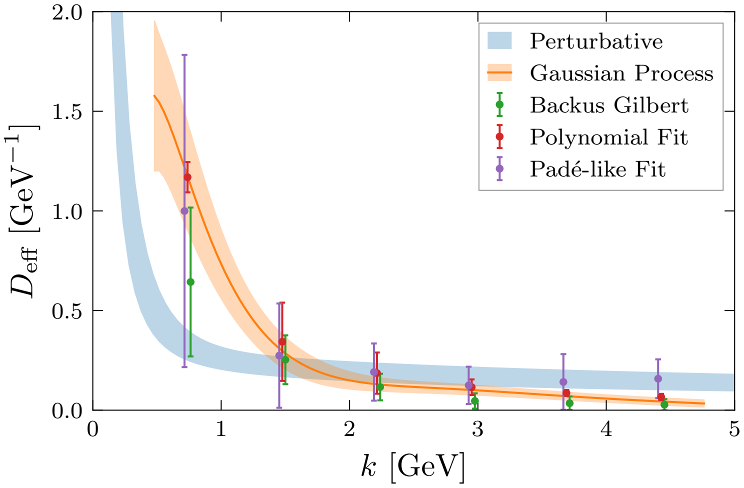


Figure 6.6: Effective diffusion coefficient D_{eff} for the quenched, $N_f = 0$, case at $T = 470$ MeV $= 1.5\,T_c$. We compare the perturbative result at NLO + LPM$^{\text{LO}}$ from [309], indicated by the blue shaded area with the spectral reconstructions from non-perturbative lattice calculations. The orange line represents the result from the GP reconstruction, which provides a continuous estimate of D_{eff}, since the reconstruction was performed in ω and k-direction simultaneously. The other reconstruction methods are shown as individual data points, where for better visibility, they are slightly shifted in k-direction.

systematic error, captured by varying these hyperparameters has similar magnitudes, as the error propagation on the error of the underlying lattice data, i.e. the statistical error. The final error estimate on the spectral function from the GP reconstruction is given by a combination of these two contributions.

The resulting spectral functions corresponding to the quenched lattice correlator data are presented in Figure 6.4 and the full QCD results can be found in Figure 6.5. In order to compare with the reconstructions from the fits and the BG method, we have evaluated the spectral functions at the spatial momenta k, where the lattice data is available. In general, however, the GP provides a statistical estimate for all momenta in between the available data. This provides a continuous estimate of the thermal photon rate not only at the given spatial momentum values but also in between. The validity of the interpolation between data points has been additionally confirmed by reconstructing the effective diffusion coefficient while disregarding lattice data for a single spatial momentum. This results, as expected, in very similar values for the effective diffusion coefficient with an increased error around this momentum value. Although the interpolation in momentum direction gives reliable estimates, the extrapolation towards smaller or higher momenta comes with large uncertainties. Ultimately, when extrapolating towards $k = 0$, to extract the diffusion coefficient, we recover the GP prior as the current assumptions about the spectral function do not allow for a systematic extrapolation.

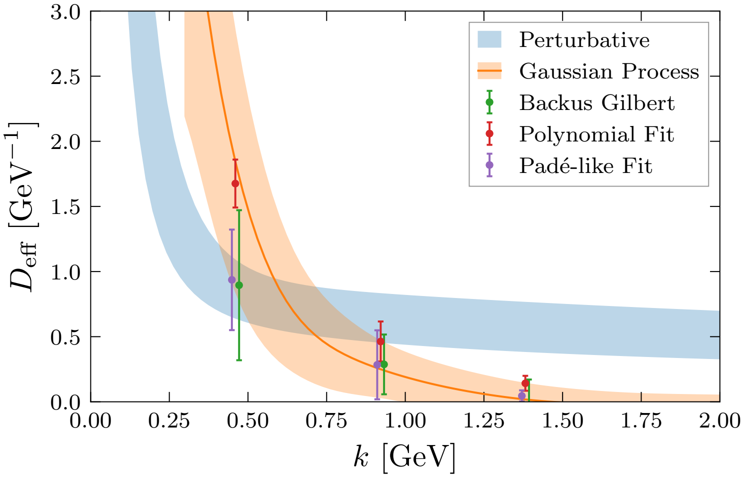


Figure 6.7: Effective diffusion coefficient $D_{\text{eff}}T$ for the quenched, $N_f = 0$, case at $1.5\,T_c$. We compare the perturbative result at NLO+LPM$^{\text{LO}}$ from [309], indicated by the blue shaded area with the spectral reconstructions from non-perturbative lattice calculations. The orange line represents the result from the GP reconstruction, which provides a continuous estimate of D_{eff}, since the reconstruction was performed in ω and k-direction simultaneously. The other reconstruction methods are shown as individual data points, where for better visibility, they are slightly shifted in k-direction.

6.4 RESULTS

With these methods, we reconstructed the T-L spectral function for the quenched and full QCD cases. As discussed in Section 6.1, the effective Diffusion coefficients $D_{\text{eff}}(k)$ can be extracted from the spectral function on light cone, i.e. $\rho_H(\omega = |\mathbf{k}|, \mathbf{k})$. We compare the extracted values of D_{eff} from the different methods in Figure 6.6 for the quenched case and in Figure 6.7 for the full QCD case. In both cases, we also compare the reconstructed effective diffusion coefficient with a perturbative computation. This perturbative calculation of the spectral functions is at next-to-leading order (NLO) in the coupling constant away from the light cone, while leading order Landau-Pomeranchuk-Migdal resummation (LPM$^{\text{LO}}$), as proposed in [289, 310], was used near the light cone. The full procedure is described in [309] and more details can be found in the accompanying publication [3].

In Figure 6.6 and Figure 6.7, we present the results of the effective diffusion coefficient D_{eff} for the quenched and full QCD cases, respectively. For both lattice data sets, the reconstruction becomes more difficult with smaller momenta. This is expected, as for smaller momentum the T-L correlator becomes completely flat, and the spectral function ρ_H approaches a delta function at zero frequency. Since the slope of the spectral function grows, the error on the reconstruction increases. This also makes an extrapolation of the diffusion coefficient towards $k = 0$ not feasible.

In the case of quenched QCD, see Figure 6.6, we find that the GP reconstruction and the polynomial fit ansatz yield consistent results. We find an excellent agreement in both the high and low momentum regimes, while the Backus-Gilbert method agrees with both of these results, except for the lowest momentum. However, all these methods agree within error bars. As expected the errors shrink drastically with increasing momentum. However, while the Padé-like ansatz agrees with the other methods within error bars, it gives a greater error estimate. This is due to the observation that the distribution of the spectral functions over the bootstrap samples is broad, indicating a near-degenerate minimum in χ^2. A similar observation has been made in [300], so instead of taking the standard mean and variance as the prediction and error, we take the maximum and minimum of the resulting D_{eff} as the error estimate. Consequently, this results in larger errors for the Padé-like ansatz. When comparing the results to the perturbative estimate, we find a qualitative agreement; the perturbative estimate has a similar shape and partly agrees with the non-perturbative results within error bars. However, the non-perturbative results still significantly deviate from the perturbative estimate, especially at small momenta. Since, on the light cone, the four-momentum squared is zero, the perturbative scale is set only by the temperature, the perturbative estimate and the lattice reconstruction are therefore not expected to coincide when going to higher spatial momenta.

For the full QCD case, see Figure 6.7, we do not have continuum extrapolated lattice data, as we have for the quenched case. Currently, we only have data for a single lattice spacing, in comparison to the quenched case, where we have data for three different lattice spacings, see Table 6.1. In order to estimate the error coming from the finite lattice spacing, we have performed the reconstruction using the GP reconstruction method and the polynomial fits for these different lattice spacing for the quenched data. The specific results are presented in Appendix C. There, we can see, that for small momenta, the finite lattice reconstructions agree extremely well and the systematic error due to finite lattice spacing is small. For larger momenta, the deviation at finite lattice spacing increases. This is expected as with increasing momentum, cutoff effects become more pronounced, see the discussion in [3]. Therefore, we have only considered relatively small spatial momenta for the full QCD reconstruction in this study. We find a good agreement between the different methods, most notably the Padé-like ansatz now has a similar uncertainty as the other methods. We find that the GP result has a bigger relative error compared to the quenched case, due to fewer data being available in the k-direction. This is mostly due to the systematic error; we find a larger parameter space, and more local minima that are able to correctly model the lattice data. When comparing with the perturbative prediction, we find that, similarly to the quenched case, a qualitative agreement, but the effective diffusion coefficient decreases significantly faster with increasing spatial momentum. This aligns with the fact that, in units of the respective transition temperature, quenched QCD is at approximately 25% higher temperature than full QCD. As a result, one would expect that the quenched QCD result will be closer to

k/T	Polynomial	Padé	Backus-Gilbert	Gaussian process
1.57	0.549 ± 0.035	0.469 ± 0.368	0.302 ± 0.175	0.552 ± 0.129
3.14	0.161 ± 0.092	0.129 ± 0.123	0.119 ± 0.057	0.140 ± 0.039
4.71	0.087 ± 0.048	0.089 ± 0.067	0.054 ± 0.031	0.061 ± 0.016
6.28	0.054 ± 0.018	0.058 ± 0.044	0.022 ± 0.018	0.048 ± 0.012
7.85	0.041 ± 0.007	0.066 ± 0.066	0.016 ± 0.016	0.033 ± 0.010
9.42	0.031 ± 0.007	0.074 ± 0.045	0.013 ± 0.013	0.020 ± 0.009

Table 6.2: Extracted values of $D_{\text{eff}}T$ from different methods for $N_f = 0$.

k/T	Polynomial	Padé	Backus-Gilbert	Gaussian process
2.09	0.369 ± 0.040	0.206 ± 0.085	0.197 ± 0.127	0.404 ± 0.219
4.19	0.102 ± 0.034	0.062 ± 0.058	0.063 ± 0.047	0.055 ± 0.048
6.28	0.031 ± 0.013	0.010 ± 0.010	-0.011 ± 0.047	0.004 ± 0.027

Table 6.3: Extracted values of $D_{\text{eff}}T$ from different methods for $N_f = 2 + 1$.

the perturbative result than the full QCD case.

Qualitatively, we can compare our dynamical results with other lattice studies of D_{eff}. Our results are in qualitative agreement with the study in [292] at a temperature of 254 MeV (1.2 T_{pc}). At the lowest momentum ($k = 0.4\,\text{GeV}$), D_{eff} ranges from $0.5\,\text{GeV}^{-1}$ to $2.7\,\text{GeV}^{-1}$, which agrees with the results obtained in this work. At the highest available momentum, the results are consistent with zero within the error bars. However, the error on D_{eff} presented in [292] is significantly larger. Although one of the reconstruction methods used in this work, the Padé-like ansatz, is very similar to the one used in [292], we find these deviations. This could be due to the different systematics involved in the spectral reconstruction methods, as well as the fact that the results in [292] are from continuum-extrapolated 2-flavor QCD correlators, whereas the dynamical results in this paper are from (2+1)-flavor QCD at a fixed lattice spacing. In [293], the photon production rate is estimated from the transverse channel of the correlator, i.e. only the ρ_T part in (6.8), with the same lattice configurations as in [292]. Similarly, we find a good agreement at the lowest momentum, while the results at the larger momenta are larger than the results obtained here.

6.5 CONCLUSION

In this chapter, we have presented the reconstruction of the thermal photon rate from Euclidean lattice correlators for quenched QCD ($T = 470\,\mathrm{MeV}$) and full QCD ($T = 220\,\mathrm{MeV}$) with an unphysical pion mass of $320\,\mathrm{MeV}$. As discussed in Section 3.1, different methods exist for reconstruction, but many of them introduce an uncontrollable bias that can lead to large systematic uncertainties. To mitigate such biases, we employ different methods and from the spread of the results, we can make an estimate of the possible systematic uncertainties. We compare physics-informed model fits, the Backus-Gilbert method, and spectral reconstruction with Gaussian processes. As discussed in the previous chapters, GP reconstruction stands out as a method that is able to simultaneously include all the different available data points and constraints on the spectral function, such as the sum rule, asymptotics, and the lattice data in two dimensions. This already gives a reliable result for the thermal photon rate, formulated in terms of the effective diffusion coefficient, D_{eff}, at the available spatial momenta and in between. The comparison shows that all methods yield consistent results, especially at higher momenta. Polynomial fits in particular reproduce the spectral function well, whereas the Padé-like ansatz exhibits larger uncertainties. For a more in-depth discussion of the quality of the polynomial fits, especially when compared to perturbative computations, we refer to the accompanying publication [3]. The quenched lattice data used here are continuum extrapolated, while the full QCD data are currently available only at a single lattice spacing. Nevertheless, the small-momentum regime shows no significant discretization effects, allowing for a reliable reconstruction of the spectral function even in the full QCD case at finite lattice spacing. We have presented the resulting effective diffusion coefficient for the quenched and full QCD cases and compared them to perturbative estimates in Figures 6.6 and 6.7, while we give the extracted values in Tables 6.2 and 6.3 respectively. The results show a qualitative agreement with the perturbative estimates, but the effective diffusion coefficient decreases significantly faster with increasing spatial momentum. The disagreement with the perturbative estimate is expected, since this estimate becomes more reliable only with increasing temperature, and the temperatures considered here, $1.5\,T_c$ and $1.2\,T_{pc}$, are still comparatively small. The results are in qualitative agreement with other lattice studies of D_{eff}, but the error on the results is smaller than in other studies. This work has shown the potential of Gaussian processes to capture the full error of the reconstruction, statistical and systematic, and therefore provide reliable estimates for QCD observables from lattice data.

Future directions are straightforward. Continuum extrapolated results for full QCD with a physical pion mass are the first steps that can be made. An important first step is obtaining continuum-extrapolated results for full QCD with a physical pion mass. In addition, accessing dilepton rates, which are also important for describing the evolution of the quark-gluon plasma, requires the reconstruction of the original current correlator. This process involves addressing the divergent tail of the correlator, which remains a significant challenge.

7. OUT-OF-TIME ORDERED CORRELATORS

Since the first description of chaos in the field of meteorology, more specifically, the chaotic behavior of the Lorenz equations, a system of coupled differential equations, chaos has been a topic of interest in many different fields. In the context of classical mechanics, the concept of chaos builds a bridge between classical and statistical mechanics. The chaotic evolution of a classical system leads to *ergodicity* of the system, meaning that the system explores its entire phase space, and therefore, the time averages of observables are equal to the ensemble averages. Ergodicity or, equivalently, the description of a system in terms of statistical ensembles, is the definition of thermal equilibrium. Consequently, chaos is an important indicator of thermalization in classical systems. Chaos also implies a certain timescale, at which the system forgets about its initial conditions. This is a central concept for kinetic Boltzmann equations, that describe the time evolutions of thermodynamic systems. With the assumption of *molecular chaos*, the collision kernel of the Boltzmann equation only depends on two-body collisions.

Since classical chaos plays such a crucial role in the thermalization of classical systems, it is natural to ask if there is a similar concept in quantum systems. A precise understanding of these microscopic mechanisms responsible for the local equilibration of energy and momentum in heavy-ion collisions is crucial for describing how and when these systems thermalize. First explorations of chaos in Yang-Mills theories have been performed in [311–314]. However, in recent years, with rapid developments in many-body theory, new concepts, such as eigenstate thermalization and out-of-time ordered correlators (OTOCs), have been introduced to the field of quantum chaos. In this chapter, we will introduce the concept of quantum chaos and demonstrate how out-of-time ordered correlators can be used to quantify the exponential spread of information in quantum systems. The goal is to compute Lyapunov exponents in order to estimate transport and thermalization timescales in quantum field theories.

In this work, we will focus on the O(4) model in 3+1 space-time dimensions, which can exhibit a second-order phase transition at finite temperature. This model is relevant in the context of

QCD, since in the limit of two massless flavors, the chiral phase transition is expected to be in the O(4) universality class [315]. Additionally, lattice QCD simulations have suggested that the scaling of the chiral susceptibilities is consistent with O(4) scaling behavior [316]. Note that recent functional QCD studies have suggested that this scaling regime may in fact be significantly smaller than previously thought [317]. The O(N) model is not only relevant in the context of QCD but also serves as a description of many other physical systems, such as general spin systems like the Ising or Heisenberg model, or as a model for the Higgs sector of the Standard model. We will therefore, as a first exploration, compute the Lyapunov exponent in the O(4) model at finite temperature with two different methods and try to connect the results to the thermalization of the system.

We will first introduce the concept of quantum chaos and OTOCs in Section 7.1. Then, we will discuss the classical statistical simulations of the O(4) model in Section 7.2, where we compute the OTOCs and the corresponding Lyapunov exponents. These results will be compared to diagrammatic, weak coupling expansions in Section 7.3, where we will compute the Lyapunov exponent in the O(4) model at finite spatial momentum. Finally, we will conclude with a summary and an outlook in Section 7.4.

7.1 QUANTUM CHAOS AND OTOCS

The work of Wigner [318] on the connection between the eigenstate distribution of the Hamiltonian of a quantum system and the eigenvalues of random matrices played a pioneering role in the study of quantum chaos. The idea originated in the context of trying to understand the spectra of heavy nuclei. Since it is difficult to calculate the full spectrum of such complex quantum systems analytically, the idea was to describe them statistically. When looking at a small energy window, the density of states is approximately constant. Such a spectrum is then described by the eigenspectrum of a random matrix [319]. Therefore, studying the rather generic properties of random matrices, that still adhere to the fundamental symmetries of the Hamiltonian, can give us insight into the properties of complex quantum systems. Explicitly, the level spacing distribution of a chaotic system is expected to behave according to the Wigner-Dyson distribution. This idea led to the description of quantum systems in terms of random matrix theory (RMT) and was experimentally confirmed for spectra of heavy nuclei, see e.g. [320]. Intuitively, this can be understood as follows: In a chaotic system, it is not possible to find a set of quantum numbers such that the associated energy levels are independent. This leads to off-diagonal matrix elements and consequently, to an interaction between the levels. A consequence of this interaction is that the levels have a minimum spacing, depending on the strength of the interaction – the levels repel each other. The level statistics of non-integrable many-body quantum systems can then be described by the eigenvalue distribution of matrices with random entries, i.e. RMT, which is characterized by the Wigner-Dyson distribution. This realization was generalized to any complex system, via the Bohigas-Giannoni-Schmit (BGS) conjecture [321], which states that the eigenvalue distribution of a quantum system whose classical counterpart is chaotic will be described by random matrix theory.

Complementary to this, the level spacing of integrable quantum systems is expected to behave according to Poisson statistics, as stated in the Berry-Tabor conjecture [322]. The levels of integrable systems are expected to behave like a sequence of independent random variables since they do not interact with each other, and therefore the level spacing distribution is expected to be Poissonian.

Strongly connected to RMT and the level-spacing statistics is the *eigenstate thermalization hypothesis* (ETH) [323–327]. It relates the matrix elements of the observables in the energy eigenstates of the system with the microcanonical ensemble. It conjectures that the diagonal entries of the matrix elements of observables are equal to their thermal averages and vary little between different energy eigenstates. Time averages are therefore described by a thermal ensemble, and we recover the definition of thermal equilibrium. Additionally, off-diagonal matrix elements are exponentially suppressed. The off-diagonal matrix elements are not necessary for obtaining absolute statements about thermalization, but describe the approach to equilibrium [327]. In terms of matrix elements of observables $O_{mn} = \langle E_m|\hat{O}|E_n\rangle$, the ETH is given by

$$O_{mn} \; = \; \langle \hat{O} \rangle_{\mathrm{micro},E_n} \delta_{mn} + e^{-S((E_n+E_m)/2)} f(E_n, E_m) R_{nm} \,, \tag{7.1}$$

where $\langle \cdot \rangle_{\mathrm{micro},E_n}$ denotes the microcanonical ensemble average at energy E_n, S is the entropy, f is a smooth function of the energy that goes towards zero for large energy differences, and R_{nm} is a random variable with zero mean and unit variance. The ETH is expected to hold for all physical observables in generic non-integrable quantum systems [328] and has been tested numerically in many-body quantum systems [325, 329]. Recently, progress has been made towards understanding the ETH in non-abelian gauge theories, cf. [330].

While all of these classifications of thermalization in quantum systems are very useful and powerful, eigenstate spectra are not an intuitive concept in the context of relativistic quantum field theories. Although level-spacing statistics and ETH are used as indicators for so-called quantum chaos, they do not share many conceptual similarities with classical chaos. In quantum systems, localized trajectories in phase space are not possible due to Heisenberg's uncertainty principle. Therefore, naively generalizing the classical concept of chaos, namely the divergence of trajectories with infinitesimally small differences in initial conditions, to quantum systems is generally not possible. However, in recent years, a new concept of quantum chaos has emerged, which is based on out-of-time-order correlators [75]. OTOCs regained interest in the context of black hole information scrambling [331–334], and subsequently became a popular measure of chaos in many-body quantum systems and QFTs.

Out-of-time-order correlators are defined – as the name suggests – as correlators of operators that are not time-ordered, compared to the typical set of time-ordered correlators, which are used to solve quantum field theories and compute all fundamental correlation functions. In order to diagnose the effect of perturbations of some operator $\hat{V}$ on later measurements of some other operator $\hat{W}$, we define the OTOC in terms of the commutator. Additionally, since we generally take thermal averages, we take the absolute value, or the squared value of the commutator in order to avoid cancellations. This leads to the definition of the OTOC as [334]

$$C(t) = -\langle[\hat{W}(t), \hat{V}(0)]^2\rangle \,. \tag{7.2}$$

In the late-time limit, the exponential growth of the OTOC is connected to quantum chaos and the exponentially fast scrambling of information, with a characteristic Lyapunov exponent λ_L.

In order to gain intuition about such an – admittedly quite obscure – quantity, we can look at the classical limit of the OTOC, cf. [335]. Classical chaos is defined as the sensitivity of the system to perturbations in the initial conditions. This can be expressed as a derivative of the trajectory $x(t)$ with respect to the initial conditions x_0

$$\left|\frac{\partial x(t)}{\partial x_0}\right| \sim e^{\lambda_L t} \,, \tag{7.3}$$

where $\lambda_L > 0$ is the characteristic Lyapunov exponent. When expressing the derivative as a classical Poisson bracket, this becomes

$$\frac{\partial x(t)}{\partial x_0} = \{x(t), p_0\}_{\text{poisson}} \,, \tag{7.4}$$

and canonically quantizing the Poisson bracket leads to the commutator and the momentum and position are promoted to operators

$$[\hat{x}(t), \hat{p}_0] \,. \tag{7.5}$$

In order to avoid cancellations, we take the square of the commutator, which corresponds to the absolute value in (7.3). When generalizing this to any arbitrary operators, instead of the position and momentum operators, we then obtain the definition of the OTOC (7.2). Consequently, in the classical limit, the OTOC is a measure of chaos. In quantum systems, the OTOC is a measure of the spread of information in a quantum system, which is often referred to as information scrambling [336, 332, 337]. This can be intuitively understood by looking at a system with one spatial dimension, cf. [338, 339]. When the two operators $\hat{W}$ and $\hat{V}$ are separated in space, their initial commutator will be zero

$$[\hat{W}(t=0,x), \hat{V}(0,y)] = 0 \quad \text{for } |x-y| > 0 \,. \tag{7.6}$$

The time evolution of $\hat{W}$ in the Heisenberg picture can be expanded in terms of nested commutators as [338]

$$\hat{W}(t) = e^{i\hat{H}t}\hat{W}(0)e^{-i\hat{H}t} = \sum_{n=0}^{\infty} \frac{(it)^n}{n!}[\hat{H}, \ldots [\hat{H}, \hat{W}] \ldots] \,. \tag{7.7}$$

In a system with, for example, nearest-neighbor interactions, each order in this expansion adds more interacting sites. With growing time, higher orders of this expansion become relevant, and more sites will become affected by the initial operator $\hat{W}$. When the OTOC of two spatially separated operators becomes non-zero after some time, this indicates that the information about the initial state has spread to the other operator. Complementary to the Lyapunov exponent, the speed of this information spreading is then defined as the butterfly velocity v_B. This butterfly velocity, the speed limit for information spreading, is of course bound by the speed of light in relativistic systems or by the Lieb-Robinson bound in non-relativistic spin systems [340, 341]. The expected behavior of an OTOC, a function of position and time, in a chaotic system at late or intermediate times is therefore proposed to be given by [342]

$$C(t,x) \sim e^{\lambda_L (t - x^2/(v_B^2 t))} \,. \tag{7.8}$$

After Fourier transforming the spatial part of the OTOC, this becomes

$$C(t,k) \sim e^{\lambda_L(k)t} \quad \text{with} \quad \lambda_L(k) = \lambda_L(k=0) - D_L k^2 \,, \tag{7.9}$$

where we have defined a prefactor D_L with the dimension of a diffusion coefficient, given by

$$D_L = \frac{v_B^2}{4\lambda_L(k=0)} \,. \tag{7.10}$$

Note, that depending on the convention, the Lyapunov exponent can differ by a factor of two, since the OTOC is defined as the square of the commutator. Additionally, the spatial dependence of the OTOC, proposed in (7.8), is not a priori given and is merely an observation. We will therefore numerically confirm this behavior in the context of the O(4) model. In holographic models, it has been shown that the butterfly velocity controls diffusive transport [343–345], and the thermal diffusivity is given by $D_E \sim v_B^2/(2\pi T)$. The diffusion coefficient is an important quantity for understanding the thermalization of systems, and it is a crucial input for the hydrodynamic description of heavy-ion collisions [346, 130]. See also the discussion in Chapter 6.

In [334], a universal bound on the Lyapunov exponent has been conjectured for general quantum systems, that only depends on the temperature of the system

$$\lambda_L \leq \frac{2\pi}{\beta} \,. \tag{7.11}$$

The bound is saturated by different systems, most prominently holographic models with gravity, describing a black hole in Einstein gravity [331, 332, 347] and the SYK model [348–350]. The SYK model is a $0+1$ dimensional model of $N \gg 1$ fermions with random quartic interactions between all pairs of fermions. The realization, that the bound is saturated by both such theories, which are very different in nature, suggested for many, that large N systems saturating this bound have an Einstein gravity dual, at least in some regions [334]. Up to

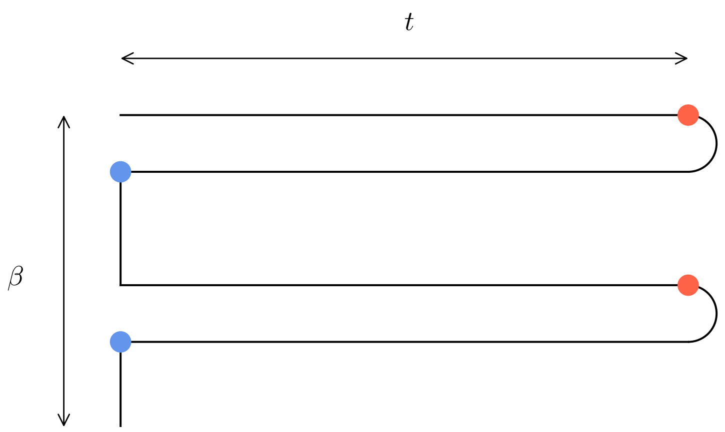


Figure 7.1: Doubled Keldysh-Schwinger contour in order to accommodate for the out-of-time ordering of the operators. We denote the two operators at different times by the two different colors, i.e. $W(t)$ is denoted by the red dot, while the operator $V(0)$ is denoted by the blue dot. There are 3 additional permutations of these operators resulting from the square of the commutator in (7.2), with different signs, that are not shown here. Following standard conventions in the OTOC literature (cf. [334, 351, 352]), we separate the two timefolds by half the thermal circle, leading to $C(t) = -\mathrm{tr}\left(\sqrt{\rho}[\hat{W}(t), \hat{V}(0)]\sqrt{\rho}[\hat{W}(t), \hat{V}(0)]\right)$.

now, this duality has not been proven, and it remains an open question if the saturation of this bound is a sufficient condition for such a duality to exist.

Since there has been an ongoing debate about the validity of the OTOC as a measure of quantum chaos, we will attempt to clarify some of the ambiguities in the following. As noted, in the classical limit, the OTOC grows exponentially, with a Lyapunov exponent. However, an absolute upper bound on (7.2) can be asserted by using the triangular identity ($\|ab\| \leq \|a\|\,\|b\|$), namely

$$C(t) \leq 4\|V\|\,\|W\|\,. \tag{7.12}$$

This means, that for bounded operators, the OTOC reaches an upper bound, after some *scrambling time* t^* [334]. It is therefore, in contrast to the classical case, where the phase-space trajectories diverge exponentially with no late-time saturation, only a transient effect. When investigating local observables in finite fermionic systems, operators are often strictly bounded and the exponential growth of the OTOC can be hard to observe [353]. However, for systems with a well-defined classical limit or in the thermodynamic limit, the OTOC can be used as a measure of quantum chaos to quantify the scrambling of information and the time scales at which a system thermalizes. It can also be, in some cases, directly connected to the ETH, a well-understood indicator of quantum chaos [354, 355].

The introduction of out-of-time ordered operators leads to significant complications, as the Keldysh-Schwinger formalism, introduced in Section 2.1, can only account for time-ordered operators. This anomalous time-ordering needs to be accounted for, by augmenting the Keldysh-

Schwinger contour with an additional timefold, see [352]. The square of the commutator in (7.2) results in a total of 4 terms, where we show one exemplary combination of operators on the appropriate contour in Figure 7.1. In the case of the squared commutator (7.20), we need one additional branch to account for the time-ordering of the four operators. Any arbitrary time-ordering of four operators needs up to one additional timefold, six operators need up to two additional timefolds, and so on. This can be generalized to any number of out-of-time ordered operators, cf. [356, 357] for systematic studies.

In the context of relativistic quantum field theories, OTOCs remain relatively unexplored. However, many concepts and even results in the context of many-body quantum systems can be transferred to field theories of interest in the context of QCD. Early works consist of weak coupling $\mathrm{O}(N)$ symmetric matrix models [351, 358, 359], that utilize diagrammatic resummation techniques, leading to a Boltzmann equation, similar to earlier works on the calculation of transport coefficients, cf. [360]. The Lyapunov exponents can then be extracted from the eigenvalue spectrum of this equation. This was extended to a strongly coupled $\mathrm{O}(N)$ model in [361], where the equations for a Lyapunov exponent at finite spatial momentum were introduced in two dimensions. In this work, we will continue this line of research and compute the Lyapunov exponent in the $\mathrm{O}(N)$ model in $d = 3$ spatial dimensions, and away from the critical point in the symmetric phase. We will present a first comprehensive study of the momentum dependence of the Lyapunov exponent in a relativistic quantum field theory.

Additionally, in [362], OTOCs were computed in a $d = 2$ dimensional $\mathbb{Z}_2$ scalar model at finite temperature, using classical statistical simulations. We have followed this approach and computed the OTOCs in the $\mathrm{O}(N)$ model in $d = 3$ spatial dimensions, with the goal of obtaining a clearer picture on the behavior of OTOCs around the phase transition. During the writing of this thesis, a similar approach was recently published in [363], for a $d = 3$ dimensional $\mathbb{Z}_2$ scalar model around the phase transition.

The last and overarching goal of this chapter is to gain a more intuitive understanding of how the OTOCs relate to known concepts in thermal field theories. In the current literature, OTOCs remain a rather abstract quantity that is not well-connected to intuitive physical concepts. As this work could not be completed in its entirety, we present preliminary results here, highlighting some key findings and outlining directions for future work.

7.2 Classical Statistical Simulations

As a first step, we will perform classical statistical simulations of the $\mathrm{O}(N)$ model in 3 spatial dimensions. Its Lagrangian is given by

$$\mathcal{L} = \frac{1}{2}\partial_\mu\varphi^a\partial^\mu\varphi^a - \frac{1}{2}m^2\varphi^a\varphi^a - \frac{\lambda}{4!N}(\varphi^a\varphi^a)^2\,, \tag{7.13}$$

where φ^a is a N-component scalar field and λ is the coupling constant. We discretize this Lagrangian on a 3-dimensional lattice with N_L^3 sites, with a finite lattice spacing a_s. Since we perform a classical simulation, we introduce a momentum field π^a conjugate to the field φ^a. The Hamiltonian of the system is then given by

$$\begin{aligned} H \;=\; \sum_x a_s^3 \Bigg[& \frac{1}{2}\pi_a(x)\pi_a(x) - \frac{1}{2a_s^2}\varphi_a(x)\sum_{e_i}(\varphi_a(x+e_i) - 2\varphi_a(x) + \varphi_a(x-e_i)) \\ & + \frac{m^2}{2}\varphi_a(x)\varphi_a(x) + \frac{\lambda}{4!N}(\varphi_a(x)\varphi_a(x))^2 \Bigg] , \end{aligned} \tag{7.14}$$

where the sum runs over all lattice sites x and e_i are the unit vectors in the spatial directions, in order to capture the nearest neighbor interactions. The time evolution of the system is then given by the Hamilton equations of motion, which are given by the functional derivatives w.r.t. the fields

$$\begin{aligned} \dot{\pi}_a(x) \;=& - \frac{\delta H}{\delta \varphi_a(x)} \\ =& \, \Delta\varphi(x) - m^2\varphi_a(x) - \frac{\lambda}{3!N}(\varphi_b(x)\varphi_b(x))\varphi_a(x) \, , \\ =& \, \frac{1}{a_s^2}\sum_{e_i}(\varphi_a(x+e_i) - 2\varphi_a(x) + \varphi_a(x-e_i)) \\ & - m^2\varphi_a(x) - \frac{\lambda}{3!N}(\varphi_b(x)\varphi_b(x))\varphi_a(x) \, , \\ \dot{\varphi}_a(x) \;=& \, \frac{\delta H}{\delta \pi_a(x)} \;=\; \pi_a(x) \, . \end{aligned} \tag{7.15}$$

These equations govern the classical time evolution of the system and can be solved numerically on the lattice. Such simulations of the classical equations of motion of the O(N) model have been performed in order to study the thermalization of the system [364–367] and the behavior of the spectral functions [368–373]. In classical statistical simulations, commutator relations for propagators or the spectral functions or the propagators, cf. Section 2.1, become Poisson brackets, as already introduced in Section 7.1. Consequently, the spectral functions are given by, cf. [362]

$$\rho_{ab}(t,x) \;=\; -\langle \{\varphi_a(t,x)\varphi_b(0,0)\}_{\text{Poisson}}\rangle_{\text{cl}} \;=\; \left\langle \frac{\delta\varphi_a(t,x)}{\delta\pi_b(0,0)} \right\rangle_{\text{cl}} , \tag{7.16}$$

where $\langle\cdots\rangle_{\text{cl}}$ denotes the average over the ensembles in the classical statistical simulation[1]. Similarly, the OTOC is then given in terms of Poisson brackets as

$$C_{ab}(t,x) \;=\; -\langle \{\varphi_a(t,x),\varphi_b(0,0)\}^2_{\text{Poisson}}\rangle_{\text{cl}} \;=\; \left\langle \left(\frac{\delta\varphi_a(t,x)}{\delta\pi_b(0,0)}\right)^2 \right\rangle_{\text{cl}} . \tag{7.18}$$

---

[1] It is usually more elegant to compute the spectral functions using the classical fluctuation-dissipation relation [368, 370]

$$\rho_{ab}(t,x) \;=\; -\frac{1}{T}\partial_t\langle\varphi_a(t,x)\varphi_b(0,0)\rangle_{\text{cl}} \, . \tag{7.17}$$

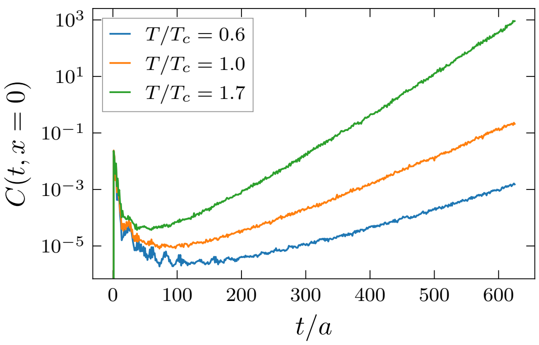


Figure 7.2: Time evolution of OTOCs for the O(4) model at different temperatures around the phase transition. After relaxation of the initial conditions, the OTOCs show a clear exponential growth at late times, which is consistent with the expected behavior of a chaotic system.

The only difference between the spectral function and the OTOCs is therefore, the square of the Poisson bracket, which is taken before taking the ensemble average. We can therefore, in the classical case, interpret the OTOC as a measure of the statistical fluctuations of the ensembles, quantifying the variance of the spectral function in the late-time limit, as the spectral function becomes zero. In the following, we will only compute the OTOCs for the diagonal components, i.e. $a = b$.

Following [362], we compute the OTOC (7.18) by evolving the linearized equations of motion for some small perturbation $\delta\varphi_a$ in parallel

$$\partial_t^2 \delta\varphi_a(x) \;=\; \Delta\delta\varphi_a(x) - m^2\delta\varphi_a(x) + \frac{\lambda}{3!}(\varphi_b(x)\varphi_b(x))\delta\varphi_a(x) + \frac{\lambda}{3}(\varphi_b(x)\delta\varphi_b(x))\varphi_a(x)\,. \quad (7.19)$$

With a perturbation in the initial momenta $\delta\pi_a(x)$, we can then compute the OTOC (7.18) and average over initial conditions. While the FDR does only hold in equilibrium – as already mentioned in Section 2.1 – computing the spectral function or the OTOC via the Poisson bracket is defined irrespective of the equilibrium state of the system.

7.2.1 Numerical Implementation

We perform the classical statistical simulations of the O(4) model in 3+1 dimensions on a lattice with N_L^3 sites and periodic boundary conditions. The simulations are performed using the programming language `Julia` [374] and the simulations are performed on GPUs on the Lichtenberg cluster at the TU Darmstadt. The lattice is distributed over the GPUs using the Julia packages `ParallelStencil.jl` [375], a package that parallelizes the lattice over the GPU – or also CPUs – optimizing for the memory layout. Additionally, we use `ImplicitGlobalGrid.jl` [376], which allows for parallelization of the lattice over multiple GPUs, with efficient MPI communication, where an outer shell of the stencil is computed first, and the MPI communication happens, while the inner part is computed.

We calculate the configurations on a square lattice with periodic boundary conditions. First, we compute the equilibrium configurations of the system using a Hybrid Monte Carlo (HMC) algorithm. We follow the standard procedure of drawing some initial momentum configurations $\pi_a(x)$ from a Gaussian distribution, and then performing a time evolution via the classical equations of motion, cf. (7.15), with a Leapfrog time-stepping algorithm. With an accept/reject step, we then make sure to sample the right ensemble given by the action and consequently arrive at the equilibrium configurations of the full quantum system.

After obtaining the equilibrium configurations, we compute the classical evolution of the system using the classical equations of motion (7.15). Parallel to the time evolution of the fields $\varphi_a(x)$, we also compute the time evolution of the perturbations $\delta\varphi_a(x)$ and $\delta\pi_a(x)$ using (7.19). We initialize the perturbations with a delta function in the momentum perturbation, i.e. $\delta\pi_a(x, t = 0) = c\delta_{x0}$, where c is a random number. In order to improve statistics, we apply the same perturbations to all field directions and later average over the different field components. We then compute the observables, most importantly the OTOC, by averaging over different time evolutions with independent initial equilibrium configurations and with random initial perturbations.

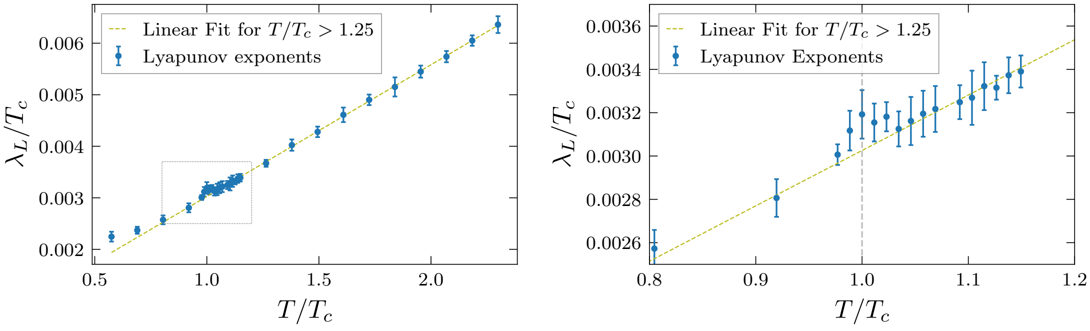
In order to estimate the Lyapunov exponent, we fit the late-time behavior of the OTOC to an exponential function. We perform this for several independent batches of the resulting OTOC samples and then take the variance and mean of the resulting Lyapunov exponents as the final result.

7.2.2 RESULTS

In Figure 7.2, we show the time evolution of the OTOC for different temperatures around the phase transition. The OTOCs show a clear exponential growth at late times, and the Lyapunov exponent can be extracted with an exponential fit to the late-time behavior of the OTOC. These samples were computed on a 126^3 lattice[2], a lattice spacing of $a = 0.4$ and a time step of $a_t/a = 0.005$. The coupling was chosen as $\lambda = 1$ and the dimensionless mass was set to $m^2a^2 = -1$. With these parameters, the critical temperature was confirmed to be $aT_c \approx 17.4$ [370]. In Figure 7.3, we show the Lyapunov exponents for different temperatures around the phase transition. We note a very clear linear growth of the Lyaponov exponents with temperature above the critical temperature, i.e. in the symmetric phase. The linear growth of the Lyapunov exponent is small, and with a value of $\lambda_L \sim 0.00256\,T$, significantly smaller than the conjectured upper bound (7.11). Around the phase transition, the Lyapunov exponent shows a local maximum. A similar behavior was independently found in [363] around the phase transition of a $\mathbb{Z}_2$ model in three spatial dimensions, but more pronounced than the results for the O(4) model.

---

[2] We reserve the boundaries of the lattice for the effective communication of the boundary conditions. Therefore, all lattice sizes in this work have the unusual size of $2^n - 2$ with $n \in \mathbb{N}$.



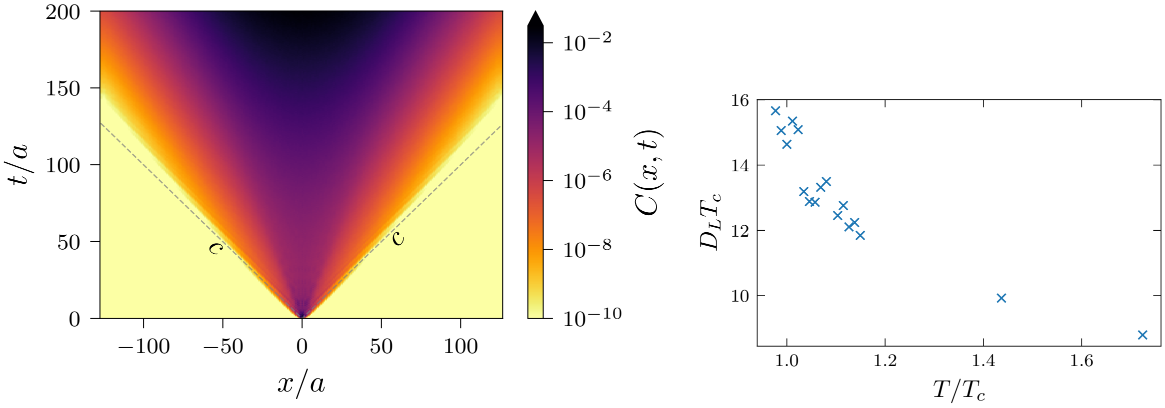
(a)

(b)

Figure 7.3: Lyapunov Exponents for the O(4) model at finite temperature, computed using classical statistical simulations at $m^2a^2 = -1$ and $\lambda = 1$. (a) shows the Lyapunov exponents for different temperatures around the phase transition. Lyapunov exponents, obtained by fitting the lattice simulation results are shown as blue dots with 2σ error bars. We can immediately see a linear dependence on the temperature in the symmetric phase, where a linear fit of these data points is shown as a dashed line. Around the critical temperature, some small deviations from this linear behavior appear. The box around the region in (a) is shown in (b), where we can see that the Lyapunov exponent has a local maximum at the critical temperature.

When measuring the Lyapunov exponent, we take relatively small lattices and still compute the OTOC up to late times $t > L$. Due to the periodic boundary conditions, at such late times, we measure higher n-point functions. However, we have checked for larger lattices, and the exponential growth and the value of the Lyapunov exponent do not change.

When computing the butterfly velocity, i.e. when investigating the spatial dependence of the OTOC, we need larger lattices, since can only use data on the OTOC up to times $t < L/2$. Otherwise, the spatially spreading OTOC overlaps with itself, due to the periodicity of the boundary conditions. We do this for the same parameters as above, i.e. $m^2a^2 = -1$, $\lambda = 1$ and $a_s = 0.4$, however, we take a larger lattice with $N_L = 254$. Note, that the statistics for these runs are worse, and for these runs, we currently do not have the full error analysis available, hence we do not show error bars here. In Figure 7.4, we show the ballistic spread of the OTOC at $T/T_c = 1.15$. We can clearly see the exponential growth towards late times and the onset of chaotic exponential growth at approximately the light cone. For every time-slice in Figure 7.4a, we then fit different functions to the spatial direction and find that the OTOC is described by the proposed form (7.8). We then extract the variance σ from the Gaussian fit and compute the diffusion coefficient D_L from the time-dependent variance $\sigma(t) = 1/(4D_Lt)$. For the specific range of parameters used, we find that the diffusion coefficient decays with temperature as $D_L \sim 1/T$. Together with the Lyapunov exponent, we can use (7.10) to compute the butterfly velocity. Since the Lyapunov exponent grows approximately linearly with temperature in this parameter regime, we find that the butterfly velocity is approximately



(a) (b)

Figure 7.4: In (a), we show the OTOC at $T/T_c = 1.15$ as a function of time and distance from the initial perturbation. The OTOC shows a clear exponential growth at late times, with a ballistic spread of the OTOC in space. We show the light cone as a dashed line and find that the butterfly velocity approximately coincides with the speed of light c. In (b), we show the diffusion coefficient D_L as a function of temperature. We find an approximate inverse dependence of the diffusion coefficient on the temperature, $D_L \sim 1/T$.

constant, with a value of $v_B \approx c$. Due to statistical noise and potential errors introduced by the fit procedures, we find a spread of the butterfly velocity of approximately $0.96c - 1.06c$. Since the butterfly velocity should be bounded by the speed of light, the error on this result should therefore be larger than 6%. The results do not indicate any temperature dependence of the butterfly velocity, however, this needs to be confirmed with a more thorough error analysis and better statistics.

Since the OTOC essentially measures, how perturbations on the initial conditions evolve, they can also measure lattice artifacts. We therefore find a dependence on the time step, even for small time steps, and hence take the very small step size of $a_t/a = 0.005$. We have checked that at this point the results do not change significantly when compared to the statistical error. Alternatively, as an improvement, one could use more sophisticated higher-order symplectic time-stepping algorithms, to reduce the time step dependence of the results. Additionally, considering next-to-nearest neighbor interactions would be an important improvement, especially around the critical point, where the correlation length diverges.

7.3 OTOCs at Weak Coupling

In the previous section, we have computed the OTOCs in the O(N) model at finite temperature using classical statistical simulations. This has produced Lyapunov exponents and OTOCs that are consistent with the expected behavior of a chaotic system. However, the classical statistical simulations do not give a good intuition about the general behavior of

OTOCs. In order to gain a better intuition about OTOCs and their behavior, we will now compute the OTOCs in the O(N) model, given in (7.13), at weak coupling using diagrammatic resummation techniques.

We will follow the approach of [351, 345, 361, 342, 359], where the OTOCs are computed using a Boltzmann equation. Here, we will briefly recap the results of [351] in the context of the O(N) model. The OTOC, we are interested in is defined as

$$C(t,\mathbf{x}) \;=\; -\frac{1}{N^2}\sum_{ab}\mathrm{tr}\left(\sqrt{\rho}\,[\varphi_a(t,\mathbf{x}),\,\varphi_b(0,\mathbf{0})]\,\sqrt{\rho}\,[\varphi_a(t,\mathbf{x}),\,\varphi_b(0,\mathbf{0})]\right)\,, \tag{7.20}$$

where ρ is the density matrix. Splitting the density matrix into two factors and putting the two commutators on opposite sides of the thermal circle simplifies the calculation, while not affecting the Lyapunov exponent [334, 351]. Otherwise, short-distance divergences can appear in the OTOC, and this prescription regulates them.

In order to calculate the late-time behavior of (7.20), we compute the OTOC at different orders of the coupling λ. As an essential ingredient, we need the retarded and Wightman propagators, G_R and G_W respectively, which are given in momentum space as

$$G_R(k) \;=\; \int\frac{d\omega}{2\pi}\,\frac{\rho(\omega,|\mathbf{k}|)}{k^0-\omega+i\epsilon}\,, \tag{7.21}$$

$$G_W(k) \;=\; \frac{\rho(k)}{2\sinh(k^0\beta/2)}\,. \tag{7.22}$$

Here, we use the (2.11) and (2.17) for the definition of the retarded propagator. The Wightman propagator is defined similar to the greater propagator (2.10) but with the density matrix split into two factors, as in the definition of the OTOC (7.20). In momentum space, this leads to the relation $G_W(k) = e^{\beta k^0/2}G^>(k)$, and using KMS relations (2.25) to the form of the Wightman propagator given above.

In the free theory limit, the spectral function is given by

$$\rho(\omega,\mathbf{k}) \;=\; \frac{\pi}{E_{\mathbf{k}}}\left(\delta(\omega-E_{\mathbf{k}})-\delta(\omega+E_{\mathbf{k}})\right)\,, \tag{7.23}$$

where $E_{\mathbf{k}} = \sqrt{\mathbf{k}^2+m^2}$ is the dispersion relation of the free theory. In this case, we simply sum over the four contributions, one of these shown in Figure 7.1, and the OTOC is then given by

$$C(t,\mathbf{x}) \;=\; -\frac{1}{N}G_R(t,\mathbf{x})^2\,. \tag{7.24}$$

At order λ we simply get one-loop corrections to the retarded propagators, resulting in a renormalization of the mass. The only consequence of this is that the mass in the final expressions has to be understood as the renormalized, and not the bare mass.

At order λ^2, we get different contributions. First, when the two vertices are on the same branch of the contour, we get an additional, albeit small contribution to the mass from the real part.

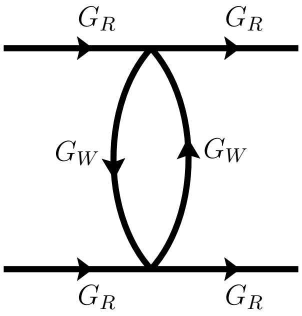


Figure 7.5: One rung contribution to the OTOC, see (7.26).

From the imaginary part, we get an exponential decay of the correlation function due to single particle decay, quantified by the decay width Γ. This leads to a retarded propagator of the form

$$G_R(k) \approx \frac{1}{2E_{\mathbf{k}}}\left(\frac{1}{k^0 - E_{\mathbf{k}} + i\Gamma_{\mathbf{k}}} - \frac{1}{k^0 + E_{\mathbf{k}} + i\Gamma_{\mathbf{k}}}\right). \tag{7.25}$$

The first interesting term arises when the two vertices are on different branches of the contour. It connects the two branches of the contour via a *rung*, as shown in Figure 7.5. This diagram is given by

$$C(k) = \int \frac{d^4\ell}{(2\pi)^4}\frac{d^4\ell'}{(2\pi)^4}\, G_R(k-\ell)G_R(\ell)R(\ell-\ell')G_R(k-\ell')G_R(\ell')\,, \tag{7.26}$$

where the rung function R is the internal loop of the diagram, and the two ends of the diagram in Figure 7.5 are closed. The rung function, as indicated in Figure 7.5, only has Wightman propagators as internal lines and is given by

$$R(\ell) = \frac{\lambda^2}{N}\int \frac{d^4p}{(2\pi)^4}\, G_W(\ell/2-p)G_W(\ell/2+p)\,. \tag{7.27}$$

Note, that depending on the theory, and the number of fields – and, of course, the convention of the coupling normalization – the prefactor of the rung function can differ, cf. [351, 359]. At higher orders in λ, we get additional contributions, that can be computed by considering dressed ladders. These are diagrams, similar to the one shown in Figure 7.5, but with additional, identical rungs. These then give corrections at least to order λ^{2n} with n being the number of rungs. There are diagrams that are not accounted for, e.g. interactions between the rungs, but we drop these contributions here, for a discussion, see [351]. In order to perform this ladder summation, we define the function f as

$$C(k) = \frac{1}{N}\int \frac{d^4p}{(2\pi)^4}\, f(k,p)\,. \tag{7.28}$$

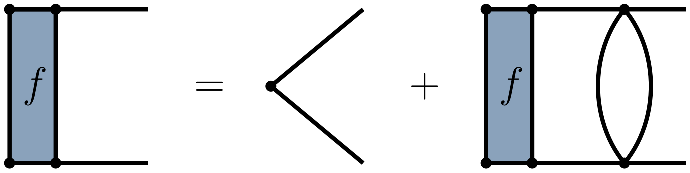


Figure 7.6: Implicit equation for the OTOC, (7.29), resulting from summing over the dressed ladder diagrams. The inhomogeneous term is neglected at late times, and the equation then states that the late-time behavior is unchanged if we add one extra rung.

The resulting sum of ladders is shown in Figure 7.6. This Bethe-Salpeter equation captures the chaotic behavior that is seeded by the interaction via the rungs. The ladder summation leads to

$$f(k,p) \;=\; G_R(p)G_R(k-p)\left[1+\int\frac{d^4\ell}{(2\pi)^4}R(\ell-p)f(k,p)\right]. \tag{7.29}$$

At large t, since we expect exponential growth from the rungs, we can neglect the inhomogeneous, zero–rung, term. The equation then states that the late-time behavior is unchanged if we add one extra rung. We will now first establish the equations for computing the Lyapunov exponents at zero momentum, which were already derived in [351], and afterward extend this to finite momentum.

This structure of the equation, consisting of retarded and advanced correlators that interact via the rungs, already gives some intuition about the OTOC. In essence, the OTOC measures how the system behaves when we perturb it after some time evolution, and then evolve the system back in time. This leads to experimental proposals for measuring OTOCs such as [377, 378], by observing the change of some observable after such a forward and then a backward time evolution.

7.3.1 Zero Spatial Momentum

For calculating the Lyapunov exponent at zero momentum, i.e. $k=(k^0,0,0,0)$, we first need to understand the dominant contributions at late times. When looking at the first rung (7.26), we can see that the late-time behavior is dominated by poles in the propagators. Explicitly, performing the p^0 integration over the product of the two propagators in (7.26) is given by the sum of the residues of the poles, and leads to terms proportional to $(k^0)^{-1}$, $(k^0-2E_{\mathbf{p}})^{-1}$ and $(k^0+2E_{\mathbf{p}})^{-1}$. The late-time behavior is dominated by the $(k^0)^{-2}$ contribution from the product of the two pairs of retarded propagators. It leads to a late-time growth proportional to t, while the other, non-singular terms do not lead to late-time growth. We, therefore, make the replacement

$$G_R(p^0,\mathbf{p})G_R(k^0-p^0,\mathbf{p}) \;\to\; \frac{i\pi}{2E_{\mathbf{p}}^2}\left(\frac{\delta(p^0-E_{\mathbf{p}})+\delta(p^0+E_{\mathbf{p}})}{k^0-i\epsilon}\right)\,, \tag{7.30}$$

where we refer to [351] for more details. In order to account for the self-energy we again make the replacement of the $i\epsilon$ with the decay width $i2\Gamma_{\mathbf{p}}$.

The integral equation (7.29) can then be written for zero spatial momentum as

$$(-ik^0+2\Gamma_{\mathbf{p}})f(k^0,p) \;\approx\; \frac{\pi}{E_{\mathbf{p}}}\,\delta((p^0)^2-E_{\mathbf{p}}^2)\int\frac{d^4\ell}{(2\pi)^4}\,R(\ell-p)f(k^0,p)\,. \tag{7.31}$$

Because of the delta functions coming from the pairs of propagators, $f(k^0,p)$ will entirely be supported on-shell, and we replace

$$f(k^0,p) \;=\; f(k^0,\mathbf{p})\,\delta\left((p^0)^2-E_{\mathbf{p}}^2\right)\,. \tag{7.32}$$

Going to real time by replacing $-ik^0\to\partial_t$ and evaluating the delta functions, leads to the final Boltzmann equation at zero spatial momentum as

$$\begin{aligned}\partial_t f(t,\mathbf{p}) \;=&\; -\,2\Gamma_{\mathbf{p}}f(t,\mathbf{p})\\ &+\int\frac{d^3\ell}{(2\pi)^3}\frac{1}{4E_{\mathbf{p}}E_{\boldsymbol{\ell}}}\,\left(R(E_{\mathbf{p}}+E_{\boldsymbol{\ell}},\mathbf{p}-\boldsymbol{\ell})+R(E_{\mathbf{p}}-E_{\boldsymbol{\ell}},\mathbf{p}-\boldsymbol{\ell})\right)\,f(t,\boldsymbol{\ell})\,.\end{aligned} \tag{7.33}$$

The real parts of the eigenvalues of this equation then dominate the late-time behavior of this equation, and if there are positive real parts, these are identified as Lyapunov exponents. Note, that the rung function is isotropic, i.e. $R(\ell)=R(|\ell^0|,|\boldsymbol{\ell}|)$, and $f(t,|\boldsymbol{p}|)$ as well. We can therefore perform the angular integral numerically and the remaining one-dimensional integral over the magnitude of the momentum, when discretized, leads to an eigenvalue problem, $\partial_t f = Mf$, where M is the matrix of the discretized integral. The Lyapunov exponent is then given by the largest eigenvalue of this matrix. For additional details on the angular integration, we refer to [351], Appendix C.

The Boltzmann equation (7.33) has an interpretation of information spreading by the kinetic collision of particles, which is similar to epidemic spreading [379, 351, 358]. When we add some perturbation to the system and follow the information spreading in the system, we can also think of it as adding an infected particle to an ensemble and following the spread of the infection, via kinetic collisions. We can interpret $f(t,\mathbf{p})$ as the probability of finding an infected particle with momentum $\mathbf{p}$ at time t. The Lyapunov exponent λ_L then describes the rate at which the infected particles spread through this ensemble. The finite decay width $\Gamma_{\mathbf{p}}$ can suppress this growth, by removing infected particles. The sum of rungs then describes the creation of infected particles/holes via scattering with particles in the ensemble. It can therefore be interpreted as the gross exchange of particles in a thermal ensemble, that grows exponentially in time, up to the point, where the whole system is infected. For further discussion of the epidemic analogy, we refer to [351, 358].

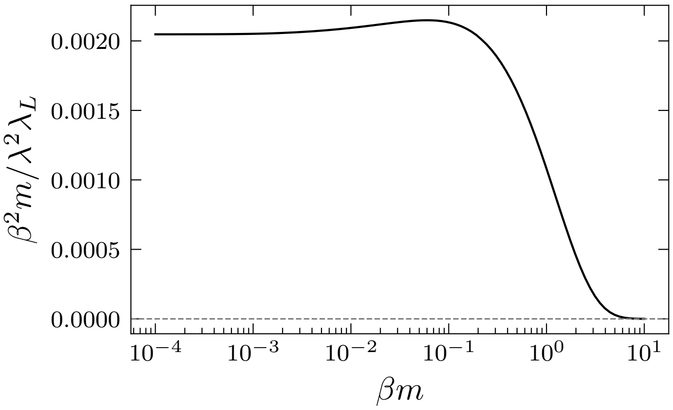


Figure 7.7: Lyapunov exponent λ_L at zero spatial momentum for the O(4) model at weak coupling. The results are consistent with the results of [351] for a matrix model. At small masses or large temperatures, $\beta m \ll 1$, the functional form of the Lyapunov exponent is given by (7.34). For masses larger than the temperature, the Lyapunov exponent is exponentially suppressed.

We have calculated the Lyapunov exponent at zero spatial momentum for different values of the coupling λ and the mass m. The results are shown in Figure 7.7 and reproduce the results of [351] for a matrix model. We can see that for a small mass compared to the temperature, $\beta m \ll 1$, we find

$$\lambda_L \approx 0.008 \frac{\lambda^2}{N\beta^2 m} . \tag{7.34}$$

Since in this approximation, the only difference between the Boltzmann equations of the two models is the number of diagrams contributing to the rungs, the only difference is the prefactor in (7.27). When comparing to the existing literature, we find that the functional form of the Lyapunov exponent at large temperatures coincides with the plasmon damping rate [380–382] which is exactly $\gamma \sim \lambda^2/(\beta^2 m)$ for a classical $\mathbb{Z}_2$ scalar theory. This gives a direct interpretation of the results: The Lyapunov at zero momentum describes the fastest possible decay time of the system. We sum over the whole space and therefore consider all possible modes. We therefore find the decay time of the fastest decaying mode of the classical system, the collective mode of the plasmon.

For small masses, we find that the Lyapunov exponent would diverge. However, for exactly these values, where the classical mass becomes small compared to the temperature, $\beta m \ll 1$, the temperature contribution to the mass has to be taken into account [351]. By including the one-loop thermal mass $m_{\text{th}}^2 = \lambda/24\beta^2$ [113], we find the Lyapunov exponent in the case of (almost) massless fields as

$$\lambda_L \approx 0.039 \frac{\lambda^{3/2}}{N\beta} . \tag{7.35}$$

This result is proportional to the thermal width of particles in hot scalar theories [360, 383] for soft on-shell momenta and to the one-loop improved plasmon damping rate at high temperatures [380, 382]. This correspondence is already suggested by the interpretation of this equation as the spreading of an infection in a thermal ensemble, via the kinetic collision of particles. The timescale between collisions is given by the inverse thermal width, and therefore the Lyapunov exponent is given by the inverse of this timescale.

7.3.2 FINITE MOMENTUM

A similar study has been performed in [361] for a strongly coupled O(N) model at finite momentum in two spatial dimensions. Here, we build on this and extend these results to three spatial dimensions and weak coupling.

One of the changes when going to finite spatial momenta, is in the product of the propagators. Similar to (7.30), we replace the product of the two propagators with

$$\begin{aligned} &G_R(p^0,\mathbf{p})G_R(k^0-p^0,\mathbf{k}-\mathbf{p}) \\ &\qquad \to \frac{i\pi}{2E_\mathbf{p}E_{\mathbf{k}-\mathbf{p}}}\left(\frac{\delta(p^0-E_\mathbf{p})}{k^0-(E_\mathbf{p}-E_{\mathbf{k}-\mathbf{p}})+i\epsilon}+\frac{\delta(p^0+E_\mathbf{p})}{k^0+(E_\mathbf{p}-E_{\mathbf{k}-\mathbf{p}})+i\epsilon}\right). \end{aligned} \tag{7.36}$$

Note, that there is a larger number of terms, but we only consider the most divergent here, in order to capture the late-time behavior. As before, we can replace the $i\epsilon$ with the decay width $2\Gamma_{\mathbf{p}-\mathbf{k}}$, in order to account for the finite decay width. From the two delta functions, f is only supported on-shell, and we can write, where we now have to separate the two delta functions as

$$f(k^0,\mathbf{k},p) = \frac{f_+(k^0,\mathbf{k},\mathbf{p})}{2E_\mathbf{p}}\delta(p^0-E_\mathbf{p})+\frac{f_-(k^0,\mathbf{k},\mathbf{p})}{2E_\mathbf{p}}\delta(p^0+E_\mathbf{p})\,. \tag{7.37}$$

We insert the propagators and the on-shell ansatz into the Boltzmann equation (7.29), and obtain

$$\begin{aligned} &\frac{f_+(k^0,\mathbf{k},\mathbf{p})}{2E_\mathbf{p}}\delta(p^0-E_\mathbf{p})+\frac{f_-(k^0,\mathbf{k},\mathbf{p})}{2E_\mathbf{p}}\delta(p^0+E_\mathbf{p}) = \\ &\frac{\pi i}{2E_\mathbf{p}E_{\mathbf{k}-\mathbf{p}}}\left[\frac{\delta(p^0-E_\mathbf{p})}{k^0-(E_\mathbf{p}-E_{\mathbf{k}-\mathbf{p}})+2i\Gamma_{\mathbf{p}-\mathbf{k}}}+\frac{\delta(p^0+E_\mathbf{p})}{k^0+(E_\mathbf{p}-E_{\mathbf{k}-\mathbf{p}})+2i\Gamma_{\mathbf{p}-\mathbf{k}}}\right] \\ &\int\frac{d^4\ell}{(2\pi)^4}R(\ell-p)\left[\frac{f_+(k^0,\mathbf{k},\boldsymbol{\ell})}{2E_{\boldsymbol{\ell}}}\delta(\ell^0-E_{\boldsymbol{\ell}})+\frac{f_-(k^0,\mathbf{k},\boldsymbol{\ell})}{2E_{\boldsymbol{\ell}}}\delta(\ell^0+E_{\boldsymbol{\ell}})\right]. \end{aligned} \tag{7.38}$$

From the delta functions, we then get two separate equations for the two contributions, one for f_+

$$\begin{aligned} &(-ik^0+i(E_\mathbf{p}-E_{\mathbf{k}-\mathbf{p}})+2\Gamma_{\mathbf{p}-\mathbf{k}})f_+(k^0,\mathbf{k},\mathbf{p}) = \\ &\frac{1}{2E_{\mathbf{k}-\mathbf{p}}}\int\frac{d^3\ell}{(2\pi)^3}\left[\frac{f_+(k^0,\mathbf{k},\boldsymbol{\ell})}{2E_{\boldsymbol{\ell}}}R(E_{\boldsymbol{\ell}}-E_\mathbf{p},\boldsymbol{\ell}-\mathbf{p})+\frac{f_-(k^0,\mathbf{k},\boldsymbol{\ell})}{2E_{\boldsymbol{\ell}}}R(-E_{\boldsymbol{\ell}}-E_\mathbf{p},\boldsymbol{\ell}-\mathbf{p})\right], \end{aligned} \tag{7.39}$$

and one for f_-

$$
\begin{aligned}
&(-ik^0 - i(E_{\mathbf{p}} - E_{\mathbf{k-p}}) + 2\Gamma_{\mathbf{p-k}})f_-(k^0, \mathbf{k}, \mathbf{p}) \;=\\
&\frac{1}{2E_{\mathbf{k-p}}}\int \frac{d^3\ell}{(2\pi)^3}\left[\frac{f_+(k^0,\mathbf{k},\boldsymbol{\ell})}{2E_{\boldsymbol{\ell}}}R(E_{\boldsymbol{\ell}}+E_{\mathbf{p}},\boldsymbol{\ell}-\mathbf{p}) + \frac{f_-(k^0,\mathbf{k},\boldsymbol{\ell})}{2E_{\boldsymbol{\ell}}}R(-E_{\boldsymbol{\ell}}+E_{\mathbf{p}},\boldsymbol{\ell}-\mathbf{p})\right] \,. \qquad (7.40)
\end{aligned}
$$

These equations can be decoupled by interpreting the f_+ and f_- as an analogue of the particle and hole densities respectively. Then, we can make the ansatz, that the density of holes and particles at opposite momenta are equal, i.e. $f_-(k^0, \mathbf{k}, \mathbf{p}) = f_+(k^0, \mathbf{k}, -\mathbf{p})$. This leads to the following equation for f_+

$$
\begin{aligned}
&(-ik^0 + i(E_{\mathbf{p}} - E_{\mathbf{k-p}}) + 2\Gamma_{\mathbf{p-k}})f_+(k^0, \mathbf{k}, \mathbf{p}) \;=\\
&\int \frac{d^3\ell}{(2\pi)^3}\frac{1}{4E_{\boldsymbol{\ell}}E_{\mathbf{k-p}}}\left[R(E_{\boldsymbol{\ell}}-E_{\mathbf{p}},\boldsymbol{\ell}-\mathbf{p}) + R(E_{\boldsymbol{\ell}}+E_{\mathbf{p}},\boldsymbol{\ell}+\mathbf{p})\right] f_+(k^0,\mathbf{k},\boldsymbol{\ell})\,, \qquad (7.41)
\end{aligned}
$$

and a similar equation for f_-. In order to numerically solve this equation, we need to write out the full angle dependence of (7.41). The spatial momenta are parametrized by the standard spherical coordinates, and we note that the rung functions only depend on the absolute value of the spatial momentum. We fix $\mathbf{k} = (0, 0, k)^\top$ and $\varphi_p = 0$, leading to $\mathbf{p} = p\,(\sin\theta_p, 0, \cos\theta_p)^\top$. From this, we obtain the following combination of momenta

$$
\begin{aligned}
|\boldsymbol{\ell} \pm \mathbf{p}| \;&=\; \left(\ell^2 + p^2 \pm 2p\ell\cos\theta_\ell\cos\theta_p \pm 2p\ell\cos\varphi_\ell\sin\theta_\ell\sin\theta_p\right)^{1/2}\\
|\mathbf{k} - \mathbf{p}| \;&=\; \left(k^2 + p^2 - 2kp\cos\theta_p\right)^{1/2} \,. \qquad (7.42)
\end{aligned}
$$

Leading to the following Boltzmann equation for f_+

$$
\begin{aligned}
(-ik^0+i(E_{\mathbf{p}} - E_{\mathbf{k-p}}) + 2\Gamma_{\mathbf{p-k}})f_+(k^0, k, p, \theta_p) \;&=\; \int_0^\infty d\ell\,\ell^2\int_0^{2\pi} d\varphi_\ell\int_0^\pi d\theta_\ell\,\sin\theta_\ell\\
&\frac{1}{(2\pi)^3}\frac{1}{4E_{\boldsymbol{\ell}}E_{\mathbf{k-p}}}\left[R(E_{\boldsymbol{\ell}}-E_{\mathbf{p}},|\boldsymbol{\ell}-\mathbf{p}|) + R(E_{\boldsymbol{\ell}}+E_{\mathbf{p}},|\boldsymbol{\ell}+\mathbf{p}|)\right] f_+(k^0,k,\ell,\theta_l)\,. \qquad (7.43)
\end{aligned}
$$

As before, we can replace $-ik^0$ with ∂_t, and compute the eigenvalues of this equation. Note, that we have explicitly checked, that the equation for f_- leads to the same Lyapunov exponent, as the equation for f_+.

As opposed to the case of zero spatial momenta, f now depends on the angle θ_p. We perform the angular integral over φ_ℓ numerically and are left with a two-dimensional integral over the momentum magnitude and the angle. This is again written as a matrix equation, and we compute the eigenvalues of this matrix, in order to find the Lyapunov exponent. The integral is discretized over $N_p = 1000$ points in the momentum space and $N_\theta = 20$ points in the angular direction. Note, that since we now have two integrals, the eigenvalue problem becomes significantly larger: we have to compute the eigenvalue of an $N \times N$ matrix with $N = N_P \cdot N_\theta$ and the computational complexity is $O(N^3)$.

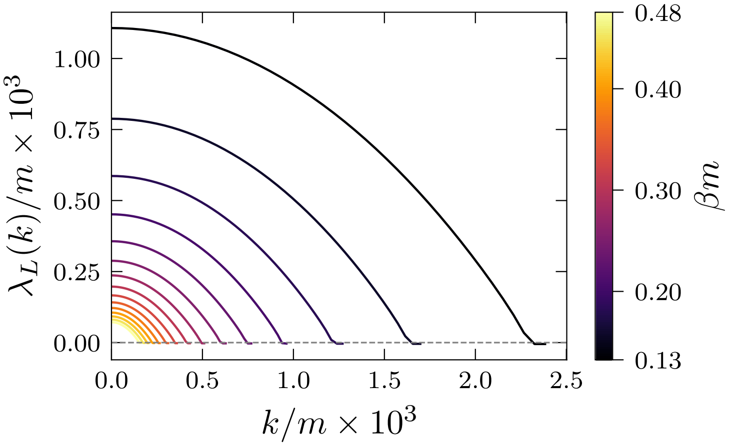


Figure 7.8: Lyapunov exponent λ_L at finite spatial momentum for the O(4) model at $\lambda = 0.1$ for different temperatures larger than the mass, i.e. $\beta m < 1$. The Lyapunov exponent shows the same functional form for all temperatures, i.e. $\lambda_L(k) = \lambda_L(k=0) - D_L k^2$. The resulting diffusion coefficient D_L is shown in Figure 7.9.

Another complication, that arises, is the observation, that the coupling λ is not only a prefactor, as in the case of zero spatial momentum, where we can factor out the λ^2 from the eigenvalue problem. We therefore have to calculate the Lyapunov exponent for fixed values of λ and explicitly investigate the exact dependencies numerically.

7.3.3 RESULTS

In Figure 7.8, we show the Lyapunov exponent at finite spatial momentum for different values of the spatial momentum k at fixed coupling $\lambda = 0.1$. We find, that the momentum dependent Lyapunov exponent has the same form we already confirmed with classical statistical simulations, see (7.9) and Section 7.2.2, i.e.

$$\lambda_L(k) \;=\; \lambda_L(k=0) - D_L k^2 \,. \tag{7.44}$$

Note, that we only compute the largest eigenvalue of the matrix, which becomes negative for large enough momenta. We have confirmed, that it remains negative for larger momenta, so we interrupt the calculation as soon as the largest eigenvalue becomes negative. From this data, we can extract the prefactor D_L, by fitting the data in Figure 7.8 to a quadratic function. For a fixed $\lambda = 0.1$, we find an approximate behavior of the Lyapunov exponent

$$D_L \;\approx\; 1.5 \cdot 10^4 \cdot \beta^2 m \,. \tag{7.45}$$

Additionally, we compute the Lyapunov exponent for different values of the coupling λ at a fixed $\beta m = 0.1$. These results are shown in Figure 7.10. We find again, that for the parameters investigated the Lyapunov exponent has the same functional form as in Figure 7.8 and decreases quadratically with the momentum. The resulting diffusion coefficients D_L are

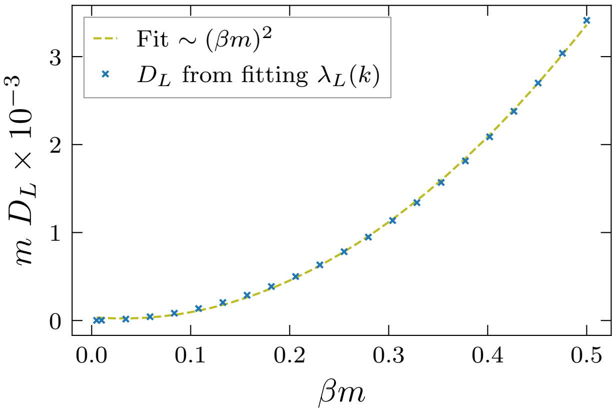


Figure 7.9: Diffusion coefficient D_L, obtained from fitting the momentum dependent Lyapunov exponent $\lambda_L(k)$ to a quadratic function, as shown in Figure 7.8, for different values of the dimensionless combination βm. We find a clear quadratic dependence of $m\ D_L$ on βm, and show the fit as a dashed line.

shown in Figure 7.11 and show an inverse quadratic dependence on the coupling λ. Note, that at such large values of the coupling, higher order-corrections to the correlators can become important, and the current results are not valid anymore. Nevertheless, the parameter dependence of the equations remains a valid observation. We therefore find – for small couplings and the investigated range of βm – the approximate behavior of the diffusion coefficient

$$D_L \approx 29.3 \cdot \frac{N\beta^2 m}{\lambda^2}\,. \tag{7.46}$$

For all momentum dependent Lyapunov exponents, we find that the butterfly velocity, given by $v_B = \sqrt{4D_L\lambda_L}$, cf. (7.10), is close to the speed of light, $v_B \approx 0.95 - 0.97\ c$, for $\beta m < 0.5$ and the couplings λ investigated here. As a result, the diffusion coefficient is approximately inversely proportional to the Lyapunov exponent, $D_L \approx 1/4\lambda_L$. The small changes in the butterfly velocity do not significantly affect the diffusion coefficient in this regime, and this approximation holds.

Since we expect there to be a qualitative difference between the high and low temperature regime, i.e. $\beta m \ll 1$ and $\beta m \gg 1$, we also compute the Lyapunov exponent for larger values of βm. We present these results in Figure 7.12. At these larger values of βm, or equivalently, smaller temperatures, we find that the butterfly velocity is no longer constant, but decreases with increasing mass of decreasing temperature. With the currently available data, we cannot determine the exact functional form of this dependence, however, it is clear that the butterfly velocity decreases with increasing mass and is compatible with $1/\sqrt{\beta m}$. The Lyapunov exponent at zero momentum is exponentially suppressed for larger values of βm, as shown in Figure 7.7. This makes it numerically challenging to compute the Lyapunov exponent at larger values of βm since the numerical precision is bound by the finite size of the matrix.

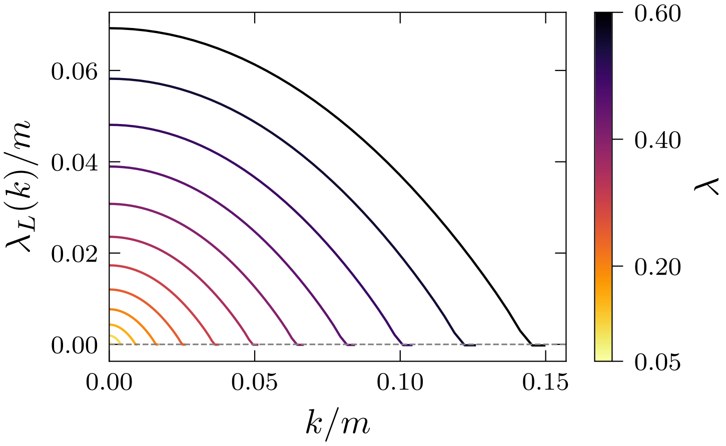


Figure 7.10: Lyapunov exponent λ_L at finite spatial momentum for the O(4) model at $\beta m = 0.1$ for different values of the coupling λ. The Lyapunov exponent shows the functional form $\lambda_L(k) = \lambda_L(k = 0) - D_L k^2$ for couplings approximately smaller than 0.6. The resulting diffusion coefficients D_L are shown in Figure 7.11.

7.4 Conclusion and Outlook

In this chapter, we have explored the behavior of OTOCs in a relativistic framework. In the context of relativistic systems, OTOCs remain relatively unexplored, and we have therefore focused on gaining first insights by understanding the general behavior of OTOCs and how they relate to physical quantities important for the evolution of quantum systems. We have investigated the exponential growth of OTOCs in the O(4) model using classical statistical lattice simulations and diagrammatic weak coupling methods, and find that the O(4) model exhibits exponentially fast information spreading and classical chaos, as diagnosed by the OTOC.

In the classical statistical simulations, we observe that for temperatures above the phase transition, the Lyapunov exponent increases approximately linearly with temperature, see Figure 7.3. However, near the phase transition, we find that the Lyapunov exponent reaches a local maximum. This non-trivial behavior suggests interesting dynamics at the critical point and qualitatively agrees with the $\mathbb{Z}_2$ results, that were independently obtained in [363]. However, we note that, in order to make conclusive statements about the Lyapunov exponent at the critical point, one would need to extend the lattice simulation analysis. This entails verifying time-step independence and ensuring that the results are not affected by finite-size effects, especially in the vicinity of the phase transition. At present, a full analysis of these dependencies is still lacking; they are assessed only qualitatively and exclusively in the symmetric phase.

Next, we compute the Lyapunov exponent at zero spatial momentum using the diagrammatic weak coupling approach, similar to [351]. The behavior of the Lyapunov exponent for weak

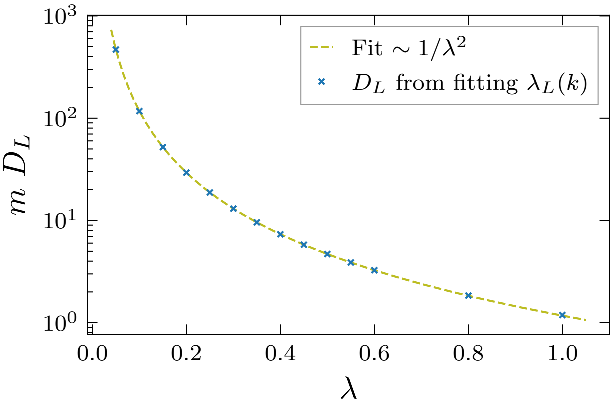


Figure 7.11: Diffusion coefficient D_L, obtained from fitting the momentum dependent Lyapunov exponent $\lambda_L(k)$ to a quadratic function, as shown in Figure 7.10, for different values of the coupling λ and $\beta m = 0.1$. We find an inverse quadratic dependence of D_L on λ, and show the fit as a dashed line.

coupling has two qualitatively different regimes, the high temperature and the low temperature regime, i.e. $\beta m \ll 1$ and $\beta m \gg 1$. In the high temperature regime, the one-loop thermal mass dominates, and the zero momentum Lyapunov exponent grows linearly with the temperature, $\lambda_L \sim T$, cf. (7.35). This dependence agrees with the results of the lattice simulations. In the low temperature regime, this Lyapunov exponent is exponentially suppressed, as shown in Figure 7.7. Note that the results obtained for the Lyapunov exponents at zero spatial momentum have the same parameter dependence, but are around 40 times larger than the plasmon damping rate for a classical $\mathbb{Z}_2$ scalar theory [381, 382].

When comparing the two methods, it is important to note that the Lyapunov exponents are computed in different parameter regimes when comparing the two methods. The weak coupling results apply in the regime of small coupling, and either at large mass or at temperatures where the one-loop thermal mass dominates, see (7.35). These are typically the conditions under which the Lyapunov exponent is small. Measuring such small exponents via classical statistical simulations is challenging, as the signal is small and oscillations emerging from the slow relaxation of initial conditions dominate.

At higher temperatures, Lyapunov exponents increase and become easier to measure in classical statistical simulations. Additionally, thermal effects dominate, and we would expect the classical simulation to approach the full quantum description. While we compute the Lyapunov exponent for different small couplings in this regime, the results do not allow for a quantitative comparison with the weak coupling results, as the Lyapunov exponents rapidly deviate from each other for larger couplings. We find similar results for the Lyapunov exponents at very small couplings, $\lambda < 0.1$, but direct quantitative comparison between the two approaches is not possible at this stage. Nevertheless, finding common ground between these two approaches remains an important direction for future research.

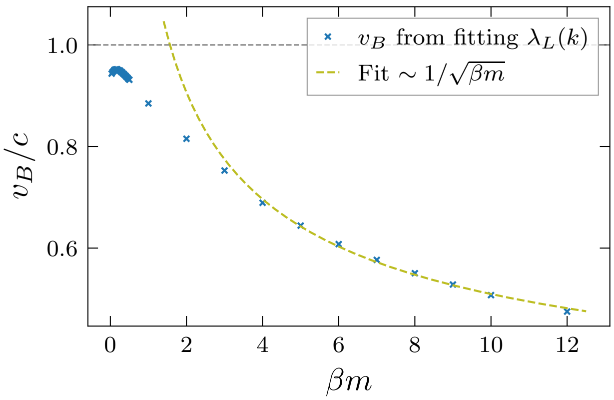


Figure 7.12: Butterfly velocity $v_B = \sqrt{4D_L\lambda_L(k=0)}$ for the O(4) model at different values of βm for a fixed $\lambda = 0.1$. We find a qualitative difference between the high temperature regime, $\beta m \ll 1$, and the low temperature regime, $\beta m \gg 1$. In the high temperature regime, the butterfly velocity is close to the speed of light, while in the low temperature regime, the butterfly velocity decreases with increasing mass or decreasing temperature. This decrease aligns with the intuitive expectation that the butterfly velocity should scale with the average velocity of the particles in a gas, $\bar{v} \sim 1/\sqrt{\beta m}$. We show a fit of the data for $\beta m \geq 4$ to a function proportional to $1/\sqrt{\beta m}$ as a dashed line. The data allows for different negative powers of βm, so we can not make a conclusive statement about the exact functional form, but we can exclude a linear dependence.

We compute the spatial spread of the OTOC using both classical statistical simulations and the diagrammatic method, and find that in the parameter regimes investigated, the momentum dependent Lyapunov exponent has the form

$$\lambda_L(k) \;=\; \lambda_L(k=0) - D_L k^2\,. \tag{7.47}$$

From this, we can extract the butterfly velocity, defined as $v_B = \sqrt{4D_L\lambda_L(k=0)}$. At high temperatures, with both approaches, we find that the butterfly velocity is approximately constant and close to the speed of light, $v_B \approx c$, as shown in Figure 7.12. This can be explained by the fact that at these temperatures, the mass of the particles is small when compared to the temperature, and they can propagate at almost the speed of light. Additionally, for small couplings, the mean free time of these particles is large, and they can propagate over large distances before scattering. We expect, that the Lyapunov exponent is inversely proportional to the mean free time, and therefore would expect a modification to the butterfly velocity for faster scrambling systems.

At the phase transition and in the symmetry broken regime, using classical statistical simulations, we also obtain a butterfly velocity that is close to the speed of light. In these regimes, we have massless Goldstone modes, and at the critical point, the correlation length diverges. This means that information is able to spread over the whole system at the speed of light.

Only in the low temperature regime, i.e. $\beta m \gg 1$, using the diagrammatic weak coupling method, we find that the butterfly velocity is no longer at almost the speed of light, but decreases with increasing mass, as shown in Figure 7.12. A naive expectation would be that the butterfly velocity is proportional to the average velocity of the particles in a gas, $\bar{v} \sim 1/\sqrt{\beta m}$. Such a functional form agrees with the results shown in Figure 7.12, however, the data allows for different negative powers of βm, so we can not make a conclusive statement here.

The diffusion coefficient D_L quantifies the diffusion of information over the system. In strongly coupled holographic models, it was found that this diffusion coefficient is exactly the energy or charge diffusion coefficient, or at least proportional to it [343, 384, 345]. It would be interesting to investigate whether this is also the case in the O(4) model.

At the critical point, the theory becomes strongly coupled, and the diagrammatic weak coupling results are not valid. However, the diagrammatic strong coupling analysis in [361] suggests that the Lyapunov exponent is significantly larger than the weak coupling results, although these results are obtained in 2+1 dimensions and therefore not directly quantitatively comparable to the results presented in this chapter. Additionally, it would be interesting to extend the diagrammatic analysis to the symmetry-broken regime. Since the weak coupling results for the Lyapunov exponent are inversely proportional to the mass, and the Goldstone modes do not acquire a thermal mass, some different scaling must emerge. The first steps have been taken towards calculating the Lyapunov exponent at small coupling with both massive and massless fields, but final results could not be obtained before the writing of this thesis.

There are several highly interesting extensions to this work. The diagrammatic approach can be extended to QCD at weak coupling, and since QCD has massless modes, the behavior of the Lyapunov exponent is particularly interesting. However, an OTOC of fundamental degrees of freedom in QCD, such as the quark or gluon fields, is not a gauge invariant quantity. It therefore needs to be carefully evaluated, which operators to use, in order to obtain physically meaningful quantities. Especially then, it would be possible to compare the Lyapunov exponents that have been obtained for classical dynamics of Yang-Mills theory by earlier studies [311–314] and more recently [363]. However, in the case of the gauge theory, the exponential growth of a distance measured between two gauge configurations was used as a measure of chaos, and it is not a priori clear how and if the respective Lyapunov exponent is related to the growth of OTOCs. Additionally, the computation of OTOCs can be extended to pion effective theories [385, 386], low energy effective theories of QCD, in order to study the Lyapunov exponent in this physically relevant phase.

The Bethe-Salpeter equation for the OTOC, i.e. Figure 7.6 and (7.29), can be extended to not only include the two-loop propagators, but also higher order corrections. Utilizing the spectral function of non-perturbative propagators and vertices, it is possible to calculate the OTOC away from the weak coupling limit. For example, since we have already calculated the spectral function of ghost and gluon propagators in [187], and the spectral function of the strong coupling in Chapter 5, we could use these results to compute the OTOC in the strong coupling regime of YM theory. Of course, this would again be a gauge-dependent quantity, and its physical interpretation would have to be carefully evaluated. Still, this can give insight into the scrambling behavior of non-perturbative QCD and help us understand the thermalization behavior of the theory.

Another interesting research direction would be to investigate prethermalization [32, 387, 388], and how OTOCs behave, when the system approaches such an attractor solution away from equilibrium. This can provide insight into which modes are equilibrating first, and which modes survive until the system is described by hydrodynamics. First studies have been performed in this context, e.g. [389], and it would be interesting to investigate this from a relativistic QFT perspective.

Beyond OTOCs, there exist many different diagnostics for quantum chaos, as already mentioned in Section 7.1. One of these, which should be suitable in the context of relativistic QFTs, is the spectral form factor [390, 319]. The spectral form factor essentially captures the fluctuation of the partition function, and a linear time dependence growth at intermediate times is a signature of *spectral rigidity*. This essentially means that long range fluctuations in the eigenvalue spectrum are suppressed, and the system is chaotic. Understanding the concepts developed in the context of quantum many-body systems, such as ETH, the spectral form factor, Krylov complexity, or fidelity susceptibility can be interesting to understand the thermalization of relativistic quantum systems.

8. Summary and Outlook

In this thesis, we have explored various approaches to understanding the dynamics of quantum field theories. While a particular emphasis was placed on QCD in the context of heavy-ion collisions, the methods developed here are broadly applicable to a wide range of physical systems. To investigate the dynamical properties of QCD, access to real-time information, specifically the evolution of two-point correlation functions, is essential. In the first part of the thesis, we introduced a method for reconstructing spectral functions from Euclidean correlation functions. This approach enables the extraction of real-time dynamics from non-perturbative calculations performed in Euclidean spacetime, such as those used in lattice and functional methods, by solving an inverse problem.

In Chapter 3, we have introduced a method based on Gaussian processes in order to approach this ill-conditioned inverse problem. This method allows to utilize a wide range of prior information about the Euclidean correlation functions, while simultaneously providing a probabilistic framework to quantify the uncertainty on the reconstructed spectral function. We extended this method to include higher-dimensional data, normalization conditions, and the known asymptotic behavior of the spectral function. This leads to a versatile method, that does not make any strong a priori assumptions about the spectral function but only relies on physical information in order to constrain the solution space of possible spectral functions.

To make this method available to a wider audience, we have implemented this method in a Python package called *fredipy* in Section 3.4. This package is available on GitHub and PyPI, and can be used to reconstruct spectral functions from Euclidean correlation functions in a user-friendly way, and implements all of the methods presented beforehand.

While Gaussian processes can be conditioned on many different prior information, inequality constraints are not straightforward to implement. These non-linear constraints, however, are essential to describe some of the physically relevant spectral functions. In Section 3.5, we have tested two different methods to implement such constraints. Especially the second method, using non-Gaussian processes in order to capture the non-linear constraints, has shown to be a promising approach. Due to the non-Gaussian nature of the posterior, it can not be described analytically, but the distribution over spectral function predictions has to be sampled numerically. This increases the computational cost of the method but leads to positive definite spectral functions, which is essential for the physical interpretation of the results. Although further optimizations are still necessary, and its success still depends on careful tuning of prior parameters for the sampling, it has already produced promising results in the examples discussed in this section. This method is not yet included in the current public release of *fredipy*, but the package is under active development and may be extended to support this approach in the future.

In the subsequent three chapters, we apply the GP reconstruction method to different problems in QCD. In Chapter 4, we have applied the GP reconstruction method to the calculation of glueball masses in pure Yang-Mills theory. For this purpose, we have calculated the four-point gluon correlation function using the fRG for special projections onto tensor structures that have the relevant quantum numbers of the glueball states. Since we expect the resonant glueball states to appear as peaks in the spectral functions of the four-point functions, we reconstruct the spectral functions using the GP method. We observe that these correlators show clearly dominant peaks, from which we can extract the glueball masses, for the two lowest lying spin zero glueball states, the scalar and the pseudoscalar glueball. These masses are in very good agreement with results from lattice and other functional methods.

Secondly, we have applied the reconstruction method to the calculation of the real-time behavior of the strong coupling in QCD in Chapter 5. This can be useful, when calculating S-matrix elements of non-perturbative QCD processes, or can be used in further calculation of spectral function or bound states, for example using Bethe-Salpeter equations. We pursue two approaches: in the first, we compute the strong coupling using the spectral functions of the gluon and ghost propagators, reconstructed from lattice data via GPs; in the second, we directly reconstruct the strong coupling from the Taylor coupling, which is derived from the same propagators. Directly reconstructing this data leads to very similar results, and shows that the GP reconstruction method is able to give consistent results.

As a third application, in Chapter 6, we have applied the GP reconstruction method to the calculation of the thermal photon rate in QCD. This is a relevant quantity in the context of heavy-ion collisions, as it carries important information about the quark-gluon plasma. For this, we use the T-L current correlator computed for quenched and full QCD on the lattice. This particular correlator has the advantage, that it does not have the usual divergent UV tail, which significantly simplifies the reconstruction of the spectral function, while still yielding the right thermal photon rate when evaluated on the lightcone. We apply multiple reconstruction techniques, namely functional fits, Backus-Gilbert inversion, and the GP method. For both data sets, we find excellent agreement between the different methods, showing that the data quality is very high, and the T-L correlator significantly improves the reconstruction quality. This is an important cross-check for the GP method and the functional fits, as it shows that these two methods, which are very different in their approach, lead to consistent results. GPs only make very general assumptions about the spectral function, while functional fits rely on a specific functional form, and therefore highly restrict the range of possible spectral functions. This shows that the GP method is able to capture the relevant features of the spectral function, even when no specific functional form is assumed. Another important physical observable for describing the evolution of the QGP is the thermal dilepton rate. However, this rate is not accessible from this T-L correlator, and the full current spectral function has to be reconstructed. Therefore, exactly and consistently implementing the divergent UV behavior of the spectral function is an important step for future work.

In Chapter 7, we focused on information scrambling in relativistic quantum field theories by studying out-of-time ordered correlators as diagnostics of quantum chaos. While OTOCs remain relatively unexplored in QCD and relativistic systems more broadly, we investigated their properties in the simpler scalar $\mathrm{O}(N)$ model as a prototypical example, with particular attention to their spatial dependence.

To this end, we employed two methods. First, classical-statistical lattice simulations were used to compute OTOCs near the $\mathrm{O}(4)$ critical point. These simulations revealed a local maximum of the Lyapunov exponent near the phase transition, while more generally, the Lyapunov exponent grows approximately linearly with temperature above the critical point. This suggests interesting dynamics near criticality that deserve further numerical investigation.

Second, using diagrammatic weak-coupling methods, we computed Lyapunov exponents, and performed a first comprehensive study of their momentum dependence. We found that the Lyapunov exponent scales proportionally with temperature, in regimes where the thermal mass dominates, consistent with the classical-statistical results in the symmetric phase. Crucially, we connected the zero-momentum Lyapunov exponent to physical quantities such as the plasmon damping rate and the thermal width of excitations in the $\mathrm{O}(N)$ model. This connection anchors the OTOCs to established thermalization and transport scales, providing a pathway to a more intuitive picture of scrambling in these systems.

Beyond temporal behavior, we studied the spatial spread of OTOCs by analyzing the momentum dependence of the Lyapunov exponent. The data fits well to a negative quadratic dependence on the momentum, indicating diffusive spreading of information characterized by a diffusion coefficient. From this, we extracted the butterfly velocity, which approaches the speed of light at high temperatures and decreases as the mass surpasses the thermal scale. This suppression aligns with the intuitive picture of finite particle velocities in a gas, scaling roughly as the inverse square root of the mass.

Overall, our results indicate a strong interplay between thermal properties and quantum chaos: information scrambling as measured by OTOCs is closely related to physical thermal scales and transport properties of the system.

Looking ahead, many open questions remain. Extending these analyses to strongly coupled systems such as QCD is a promising but challenging direction. Further development of diagrammatic approaches incorporating non-perturbative spectral functions – potentially reconstructed with Gaussian process techniques – may enable the study of Lyapunov exponents beyond weak coupling.

In summary, this thesis has presented a versatile method to reconstruct spectral functions from Euclidean correlation functions using Gaussian processes, has applied this method to various problems in QCD, and has explored the properties of out-of-time ordered correlators as a measure of information scrambling.

A. Spectral Representation for Propagators in Euclidean Time

Here, we provide some details on the derivation of the spectral representation of the Euclidean correlator in imaginary time, and at finite temperature, cf. (2.21).

The Euclidean correlator in imaginary time is given by

$$G_E(\tau) \,=\, \frac{1}{\beta}\sum_{n=-\infty}^{\infty} \exp(-i\omega_n\tau) G_E(\omega_n)\,, \tag{A.1}$$

where ω_n are the Matsubara frequencies. We can then express the Euclidean correlator in terms of spectral representation (2.17) using the spectral function $\rho(\omega)$ as

$$G_E(\tau) \,=\, \frac{1}{\beta}\sum_{n=-\infty}^{\infty} \exp(-i\omega_n\tau) \int_{\mathbb{R}} \frac{d\omega}{2\pi}\, \frac{1}{\omega - i\omega_n}\rho(\omega)\,. \tag{A.2}$$

Hence, we want to calculate

$$K_\beta(\tau,\omega) \,=\, \frac{1}{\beta}\sum_{n=-\infty}^{\infty} \exp(-i\omega_n\tau)\, \frac{1}{\omega - i\omega_n} \tag{A.3}$$

We use the well-known trick of replacing the infinite sum by a complex integral with poles at the Matsubara frequencies, i.e.

$$K_\beta(\tau,\omega) \,=\, \oint_C \frac{dz}{2\pi i}\, \frac{\exp(z\tau)}{\omega - z} n_B(z)\,, \tag{A.4}$$

where C is a contour in the complex plane that encloses the poles at $z = i\omega_n$. This is possible, since the Bose distribution function $n_B(z) = 1/(\exp(\beta z) - 1)$ has simple poles at $z = i\omega_n = i2\pi n/\beta$ – the bosonic Matsubara frequncies – with residues equal to $\frac{1}{\beta}$, and is analytic everywhere else in the complex z-plane. Now we can deform the contour: instead of closing the contour around the poles on the imaginary axis in z, we can close it around the poles on the real axis in ω. We close the contour around the positive, real half-plane and use the residue theorem to evaluate the integral at the pole $z = \omega$. This leads to

$$K_\beta(\tau,\omega) \,=\, \exp(\omega\tau) n_B(\omega)\,, \tag{A.5}$$

Plugging this back into the spectral representation of the Euclidean correlator leads to

$$\begin{aligned} G_E(\tau) &= \int_{\mathbb{R}} \frac{d\omega}{2\pi} \frac{\exp(\omega\tau)}{\exp(\omega\beta)-1}\,\rho(\omega) \\ &= \int_0^\infty \frac{d\omega}{2\pi} \frac{\exp(\omega\tau)+\exp(-\omega\tau+\omega\beta)}{\exp(\omega\beta)-1}\,\rho(\omega) \\ &= \int_0^\infty \frac{d\omega}{2\pi} \frac{\cosh(\omega(\tau-\beta/2))}{\sinh(\beta\omega/2)}\,\rho(\omega)\,. \end{aligned} \tag{A.6}$$

In the second line, we have used the antisymmetry of the spectral function $\rho(\omega) = -\rho(-\omega)$, and in the last line, we just use the definition of the hyperbolic cosine and sine functions. This is the spectral representation of the Euclidean correlator in imaginary time, which is used in the reconstruction methods discussed in this thesis. This derivation can also, and more quickly, be performed, by using the KMS relation (2.25), doing a Fourier transform, and performing a Wick rotation to imaginary time.

B. GAUSSIAN PROCESS DETAILS

B.1 GAUSSIAN PROCESS POSTERIOR

Here we sketch the derivation of the posterior predictive distribution of a Gaussian process. To simplify notation, we denote

$$a = \mathbf{K}^{-1}\mathbf{K}_* \,, \tag{B.1}$$

$$b = \mathbf{K}_{**} - \mathbf{K}_*^\top \mathbf{K}^{-1}\mathbf{K}_* \,, \tag{B.2}$$

$$c = \mathbf{K}(\mathbf{K} + \mathbf{C}_n)^{-1}\mathbf{y} \,, \tag{B.3}$$

$$d = \sigma^2 \mathbf{K}(\mathbf{K} + \mathbf{C}_n)^{-1} \,. \tag{B.4}$$

Inserting (3.17) and (3.14) into (3.15) leads to

$$\begin{aligned}
p(\mathbf{f}_* \,|\, \mathcal{D},\, \mathbf{x}_*,\, \boldsymbol{\theta},\, \boldsymbol{\sigma}) &\sim \int d\mathbf{f} \exp\left(-\frac{(\mathbf{f} - a^\top \mathbf{f})^2}{2b} - \frac{(\mathbf{f} - c)^\top d^{-1} (\mathbf{f} - c)}{2}\right) && \text{(B.5)} \\
&\sim \int d\mathbf{f} \exp\left(-\frac{(\mathbf{f}_*)^2}{2b} - \frac{1}{2}\left(\mathbf{f}^\top \left(d^{-1} + \frac{a a^\top}{b}\right) - 2\left(d^{-1} c + \mathbf{f}_* \frac{a}{b}\right) \mathbf{f}\right)\right) && \text{(B.6)} \\
&\sim \exp\left(-\frac{1}{2}\left(\frac{1 - a^\top (b d^{-1})^{-1} a}{b} (\mathbf{f}_*)^2 - 2 a^\top (b d^{-1} + a a^\top)^{-1} d^{-1} c \mathbf{f}_*\right)\right) \,, && \text{(B.7)}
\end{aligned}$$

where we have calculated a rather tedious Gaussian integral. We can see that the above is a Gaussian distribution with mean and covariance given by

$$\mu_* = \frac{b a^\top (b d^{-1} a a^\top)^{-1} d^{-1} c \mathbf{f}_*}{1 - a^\top (b d^{-1} a a^\top)^{-1} a} \,, \tag{B.8}$$

$$\sigma_* = \frac{1}{1 - a^\top (b d^{-1} a a^\top)^{-1} a} \,. \tag{B.9}$$

We use the *Sherman-Morrison-Woodbury* formula given by

$$(A + U V^\top)^{-1} = A^{-1} - A^{-1} U (1 + V^\top A^{-1} U)^{-1} V^\top A^{-1} \,, \tag{B.10}$$

to rewrite

$$(bd^{-1} + aa^\top)^{-1} = \frac{1}{b}\left(d - \frac{daa^\top d}{b + a^\top da}\right). \tag{B.11}$$

Inserting this into (B.8) and doing some algebra leads to

$$\mu_* = a^\top c\mathbf{f}_* = \mathbf{K}_*^\top (\mathbf{K} + \mathbf{C}_n)^{-1}\,\mathbf{y}\,. \tag{B.12}$$

And the covariance (B.9) can be rewritten as

$$\begin{aligned} \sigma_* &= b + a^\top da \qquad (B.13)\\ &= \mathbf{K}_{**} - \mathbf{K}_*^\top (\mathbf{K} + \mathbf{C}_n)^{-1}\,\mathbf{K}_*\,, \qquad (B.14) \end{aligned}$$

where in between we had to use the Sherman-Morrison-Woodbury formula (B.10) again. The expressions (B.12) and (B.13) are the posterior predictive mean and covariance of the Gaussian process.

B.2 CHOLESKY DECOMPOSITION

Instead of directly inverting the matrix in the posterior of the Gaussian process, we use Cholesky decomposition to solve the linear system of equations. We precompute the quantities

$$\begin{aligned} L &= \text{cholesky}(\mathbf{K} + \mathbf{C}_n)\,, \qquad (B.15)\\ \boldsymbol{\alpha} &= L^\top \backslash (L \backslash \mathbf{y})\,, \qquad (B.16)\\ \mathbf{v}(x) &= L \backslash \mathbf{k}(x)\,, \qquad (B.17) \end{aligned}$$

where L is the lower triangular matrix from the Cholesky decomposition, $\mathbf{C}_n$ is the noise covariance matrix, $\mathbf{y}$ is the data vector, and $\mathbf{k}(x)$ is the kernel vector for a new point x. The notation $x = A \backslash b$ denotes the solution of the linear system $Ax = b$. The Cholesky decomposition is a numerical method to decompose a positive definite matrix $\mathbf{K} + \mathbf{C}_n$ into the product of a lower triangular matrix L and its transpose, such that $\mathbf{K} + \mathbf{C}_n = LL^\top$. This decomposition is numerically stable and efficient, especially for large matrices. This allows us to compute the posterior mean and covariance in (3.20) and (3.21) as

$$\begin{aligned} \mu_*(x) &= \mathbf{k}(x)^\top \boldsymbol{\alpha}\,, \qquad (B.18)\\ k_*(x, x') &= k(x, x') - \mathbf{v}(x)^\top \mathbf{v}(x')\,. \qquad (B.19) \end{aligned}$$

B.3 HIGHER DIMENSIONAL RECONSTRUCTION CODE

Here we show an example code for the higher dimensional reconstruction of spectral functions using the *fredipy* package. Note, that the momentum array needs to be carefully constructed.

```python
import numpy as np
import fredipy as fp

# prepare the arrays for the 2D case

# ... in k direction
k = np.linspace(0, 2, 20)
# ... and in w direction
w = np.linspace(0.1, 10, 100)

# mesh them together
ww, wk = np.meshgrid(w, k)
w_pred = np.c_[ww.ravel(), wk.ravel()]

# ... and in p direction
p = np.linspace(0.1, 10, 30)
pp, kk = np.meshgrid(p, k)
p_data = np.c_[pp.ravel(), kk.ravel()]

# example data
a, m, g = 1.6, 1, 0.8
rho = get_rho2d(w_pred[:,0], w_pred[:,1], a, m, g)
G = get_BW2d(p_data[:,0],p_data[:,1], a, m, g)

err = 1e-4

data2d = {
    'x': p_data,
    'y': G + err * np.random.randn(len(G)),
    'dy': err}

# we do the same as before
kernel = fp.kernels.RadialBasisFunction(0.3, [0.3, 0.3])
integrator = fp.integrators.Riemann(0.01, 20, 1000)
integral_op = fp.operators.Integral(kl_kernel, integrator)
constraints = [fp.constraints.LinearEquality(integral_op, data2d)]
model = fp.models.GaussianProcess(kernel, constraints)

# ... and do a prediction
_rho, _rho_err = model.predict(w_pred)
```

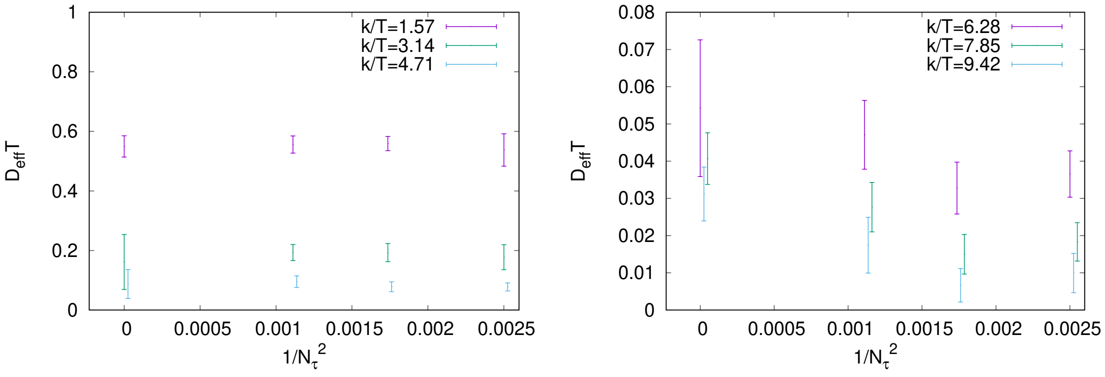
B.4 OPTIMIZATION OF KERNEL PARAMETERS CODE

Here we present a minimal example for optimizing the kernel parameters of a Gaussian process model by minimizing the NLL (3.31) using the *fredipy* package. We simply use the *scipy* package to perform the optimization using standard minimization routines. In general, however, more advanced optimization methods can be used, see the discussions in the respective chapters, where reconstructions are explicitly performed, e.g. Chapter 6.

```python
from scipy import optimize

def optimize_kernel_parameters(
        gp: fp.models.GaussianProcess, # a Gaussian Process Model
        guess: list[float],  # some initial guess on the parameters
        bound_min: float = 1e-4 # minimal value for the parameters (we want to avoid 0)
        ):
    # wrapper function
    def optimized_function(params):
        gp.set_kernel_params(params)
        return - gp.log_likelihood()

    num_params = len(guess)
    bounds = [(bound_min, None)]*num_params
    res = optimize.minimize(optimized_function, guess, bounds = bounds, method = 'L-BFGS-B')

    if res.success:
        print(' Optimized parameters are: ', res.x)
        gp.set_kernel_params(res.x)
    else:
        print('No convergence in optimization!')

optimize_kernel_parameters(model, [0.2, 0.1])
```

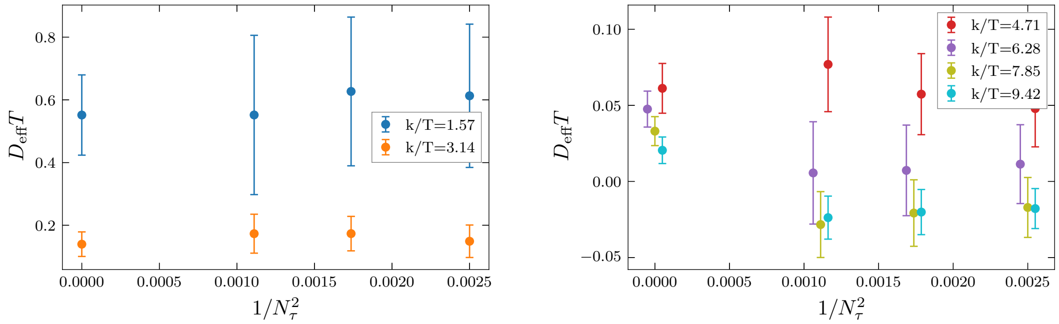
C. Thermal Photon Rate at Finite Lattice Spacing

In this appendix, we will compare the estimate of the effective diffusion coefficient from different lattice spacings in order to estimate the cutoff effects on the reconstruction. This will be illustrated for the polynomial ansatz of the spectral function and the GPR reconstruction. A comparison of the different D_{eff} values from the polynomial ansatz can be found in Figure C.1, while the same comparison for the GPR reconstruction can be found in Figure C.2. We observe that for both methods the results for small spatial momenta agree remarkably well, as different values remain consistent in the margin of error. For the polynomial ansatz and spatial momenta $k/T > 4.71$, we observe mild cutoff effects on D_{eff}, while in the GP reconstruction, the error generally grows for higher momenta. This outcome aligns with the expectation, considering that cutoff effects are already minimal in the correlator.



(a) (b)

Figure C.1: Effective diffusion coefficient, D_{eff} obtained from the polynomial estimate for different lattice spacing along with continuum extrapolated value. Here $N_f = 0$.



(a) (b)

Figure C.2: Effective diffusion coefficient, D_{eff} obtained from the GP regression model for different lattice spacing along with continuum extrapolated value. Here $N_f = 0$.

BIBLIOGRAPHY

[1] J. M. Pawlowski, C. S. Schneider, J. Turnwald, J. M. Urban, and N. Wink, *Yang-Mills glueball masses from spectral reconstruction*, Phys. Rev. D **108**, 076018 (2023), arXiv:2212.01113 [hep-ph].

[2] J. Horak, J. M. Pawlowski, J. Turnwald, J. M. Urban, N. Wink, and S. Zafeiropoulos, *Nonperturbative strong coupling at timelike momenta*, Phys. Rev. D **107**, 076019 (2023), arXiv:2301.07785 [hep-ph].

[3] S. Ali, D. Bala, A. Francis, G. Jackson, O. Kaczmarek, J. Turnwald, T. Ueding, and N. Wink (HotQCD), *Lattice QCD estimates of thermal photon production from the QGP*, Phys. Rev. D **110**, 054518 (2024), arXiv:2403.11647 [hep-lat].

[4] D. Boyanovsky, H. J. de Vega, and D. J. Schwarz, *Phase transitions in the early and the present universe*, Ann. Rev. Nucl. Part. Sci. **56**, 441–500 (2006), arXiv:hep-ph/0602002.

[5] J. I. Kapusta, *Quantum Chromodynamics at High Temperature*, Nucl. Phys. B **148**, 461–498 (1979).

[6] A. Bazavov *et al.* (HotQCD), *Chiral crossover in QCD at zero and non-zero chemical potentials*, Phys. Lett. B **795**, 15–21 (2019), arXiv:1812.08235 [hep-lat].

[7] A. Bazavov *et al.* (HotQCD), *Equation of state in (2+1)-flavor QCD*, Phys. Rev. D **90**, 094503 (2014), arXiv:1407.6387 [hep-lat].

[8] S. Borsanyi, Z. Fodor, C. Hoelbling, S. D. Katz, S. Krieg, and K. K. Szabo, *Full result for the QCD equation of state with 2+1 flavors*, Phys. Lett. B **730**, 99–104 (2014), arXiv:1309.5258 [hep-lat].

[9] J. C. Collins and M. J. Perry, *Superdense Matter: Neutrons Or Asymptotically Free Quarks?*, Phys. Rev. Lett. **34**, 1353 (1975).

[10] N. Cabibbo and G. Parisi, *Exponential Hadronic Spectrum and Quark Liberation*, Phys. Lett. B **59**, 67–69 (1975).

[11] E. V. Shuryak, *Quark-Gluon Plasma and Hadronic Production of Leptons, Photons and Psions*, Phys. Lett. B **78**, 150 (1978).

[12] I. Arsene *et al.* (BRAHMS), *Quark gluon plasma and color glass condensate at RHIC? The Perspective from the BRAHMS experiment*, Nucl. Phys. A **757**, 1–27 (2005), arXiv:nucl-ex/0410020.

[13] B. B. Back *et al.* (PHOBOS), *The PHOBOS perspective on discoveries at RHIC*, Nucl. Phys. A **757**, 28–101 (2005), arXiv:nucl-ex/0410022.

[14] J. Adams *et al.* (STAR), *Experimental and theoretical challenges in the search for the quark gluon plasma: The STAR Collaboration's critical assessment of the evidence from RHIC collisions*, Nucl. Phys. A **757**, 102–183 (2005), arXiv:nucl-ex/0501009.

[15] K. Adcox *et al.* (PHENIX), *Formation of dense partonic matter in relativistic nucleus-nucleus collisions at RHIC: Experimental evaluation by the PHENIX collaboration*, Nucl. Phys. A **757**, 184–283 (2005), arXiv:nucl-ex/0410003.

[16] M. A. Stephanov, K. Rajagopal, and E. V. Shuryak, *Signatures of the tricritical point in QCD*, Phys. Rev. Lett. **81**, 4816–4819 (1998), arXiv:hep-ph/9806219.

[17] W.-j. Fu, J. M. Pawlowski, and F. Rennecke, *QCD phase structure at finite temperature and density*, Phys. Rev. D **101**, 054032 (2020), arXiv:1909.02991 [hep-ph].

[18] F. Gao and J. M. Pawlowski, *QCD phase structure from functional methods*, Phys. Rev. D **102**, 034027 (2020), arXiv:2002.07500 [hep-ph].

[19] F. Gao and J. M. Pawlowski, *Chiral phase structure and critical end point in QCD*, Phys. Lett. B **820**, 136584 (2021), arXiv:2010.13705 [hep-ph].

[20] P. J. Gunkel and C. S. Fischer, *Locating the critical endpoint of QCD: Mesonic back-coupling effects*, Phys. Rev. D **104**, 054022 (2021), arXiv:2106.08356 [hep-ph].

[21] M. G. Alford, K. Rajagopal, and F. Wilczek, *QCD at finite baryon density: Nucleon droplets and color superconductivity*, Phys. Lett. B **422**, 247–256 (1998), arXiv:hep-ph/9711395.

[22] M. G. Alford, K. Rajagopal, and F. Wilczek, *Color flavor locking and chiral symmetry breaking in high density QCD*, Nucl. Phys. B **537**, 443–458 (1999), arXiv:hep-ph/9804403.

[23] M. G. Alford, A. Schmitt, K. Rajagopal, and T. Schäfer, *Color superconductivity in dense quark matter*, Rev. Mod. Phys. **80**, 1455–1515 (2008), arXiv:0709.4635 [hep-ph].

[24] M. G. Alford, J. A. Bowers, and K. Rajagopal, *Crystalline color superconductivity*, Phys. Rev. D **63**, 074016 (2001), arXiv:hep-ph/0008208.

[25] R. D. Pisarski and F. Rennecke, *Signatures of Moat Regimes in Heavy-Ion Collisions*, Phys. Rev. Lett. **127**, 152302 (2021), arXiv:2103.06890 [hep-ph].

[26] M. Arslandok *et al.*, *Hot QCD White Paper*, (2023), arXiv:2303.17254 [nucl-ex].

[27] T. Ablyazimov *et al.* (CBM), *Challenges in QCD matter physics –The scientific programme of the Compressed Baryonic Matter experiment at FAIR*, Eur. Phys. J. A **53**, 60 (2017), arXiv:1607.01487 [nucl-ex].

[28] W. Busza, K. Rajagopal, and W. van der Schee, *Heavy Ion Collisions: The Big Picture, and the Big Questions*, Ann. Rev. Nucl. Part. Sci. **68**, 339–376 (2018), arXiv:1802.04801 [hep-ph].

[29] R. Baier, A. H. Mueller, D. Schiff, and D. T. Son, *'Bottom up' thermalization in heavy ion collisions*, Phys. Lett. B **502**, 51–58 (2001), arXiv:hep-ph/0009237.

[30] S. Schlichting and D. Teaney, *The First fm/c of Heavy-Ion Collisions*, Ann. Rev. Nucl. Part. Sci. **69**, 447–476 (2019), arXiv:1908.02113 [nucl-th].

[31] J. Berges, M. P. Heller, A. Mazeliauskas, and R. Venugopalan, *QCD thermalization: Ab initio approaches and interdisciplinary connections*, Rev. Mod. Phys. **93**, 035003 (2021), arXiv:2005.12299 [hep-th].

[32] J. Berges, S. Borsanyi, and C. Wetterich, *Prethermalization*, Phys. Rev. Lett. **93**, 142002 (2004), arXiv:hep-ph/0403234.

[33] M. P. Heller and M. Spalinski, *Hydrodynamics Beyond the Gradient Expansion: Resurgence and Resummation*, Phys. Rev. Lett. **115**, 072501 (2015), arXiv:1503.07514 [hep-th].

[34] R. Hagedorn, *Statistical thermodynamics of strong interactions at high-energies*, Nuovo Cim. Suppl. **3**, 147–186 (1965).

[35] F. Cooper and G. Frye, *Single-particle distribution in the hydrodynamic and statistical thermodynamic models of multiparticle production*, Phys. Rev. D **10**, 186–189 (1974).

[36] D. Teaney, J. Lauret, and E. V. Shuryak, *Flow at the SPS and RHIC as a quark gluon plasma signature*, Phys. Rev. Lett. **86**, 4783–4786 (2001), arXiv:nucl-th/0011058.

[37] H. Song, S. A. Bass, and U. Heinz, *Viscous QCD matter in a hybrid hydrodynamic+Boltzmann approach*, Phys. Rev. C **83**, 024912 (2011), arXiv:1012.0555 [nucl-th].

[38] U. Heinz and R. Snellings, *Collective flow and viscosity in relativistic heavy-ion collisions*, Ann. Rev. Nucl. Part. Sci. **63**, 123–151 (2013), arXiv:1301.2826 [nucl-th].

[39] J. Weil *et al.* (SMASH), *Particle production and equilibrium properties within a new hadron transport approach for heavy-ion collisions*, Phys. Rev. C **94**, 054905 (2016), arXiv:1606.06642 [nucl-th].

[40] A. Kurkela, A. Mazeliauskas, J.-F. Paquet, S. Schlichting, and D. Teaney, *Matching the Nonequilibrium Initial Stage of Heavy Ion Collisions to Hydrodynamics with QCD Kinetic Theory*, Phys. Rev. Lett. **122**, 122302 (2019), arXiv:1805.01604 [hep-ph].

[41] A. Kurkela, A. Mazeliauskas, J.-F. Paquet, S. Schlichting, and D. Teaney, *Effective kinetic description of event-by-event pre-equilibrium dynamics in high-energy heavy-ion collisions*, Phys. Rev. C **99**, 034910 (2019), arXiv:1805.00961 [hep-ph].

[42] A. S. Kronfeld, *Twenty-first Century Lattice Gauge Theory: Results from the QCD Lagrangian*, Ann. Rev. Nucl. Part. Sci. **62**, 265–284 (2012), arXiv:1203.1204 [hep-lat].

[43] Z. Fodor and C. Hoelbling, *Light Hadron Masses from Lattice QCD*, Rev. Mod. Phys. **84**, 449 (2012), arXiv:1203.4789 [hep-lat].

[44] H.-T. Ding, F. Karsch, and S. Mukherjee, *Thermodynamics of strong-interaction matter from Lattice QCD*, Int. J. Mod. Phys. E **24**, 1530007 (2015), arXiv:1504.05274 [hep-lat].

[45] C. Ratti, *Lattice QCD and heavy ion collisions: a review of recent progress*, Rept. Prog. Phys. **81**, 084301 (2018), arXiv:1804.07810 [hep-lat].

[46] M. Troyer and U.-J. Wiese, *Computational complexity and fundamental limitations to fermionic quantum Monte Carlo simulations*, Phys. Rev. Lett. **94**, 170201 (2005), arXiv:cond-mat/0408370.

[47] G. Parisi, *On complex probabilities*, Phys. Lett. B **131**, 393–395 (1983).

[48] G. Aarts, E. Seiler, and I.-O. Stamatescu, *The Complex Langevin method: When can it be trusted?*, Phys. Rev. D **81**, 054508 (2010), arXiv:0912.3360 [hep-lat].

[49] E. Seiler, D. Sexty, and I.-O. Stamatescu, *Gauge cooling in complex Langevin for QCD with heavy quarks*, Phys. Lett. B **723**, 213–216 (2013), arXiv:1211.3709 [hep-lat].

[50] E. Witten, *Analytic Continuation Of Chern-Simons Theory*, AMS/IP Stud. Adv. Math. **50**, 347–446 (2011), arXiv:1001.2933 [hep-th].

[51] A. Alexandru, G. Basar, P. F. Bedaque, and N. C. Warrington, *Complex paths around the sign problem*, Rev. Mod. Phys. **94**, 015006 (2022), arXiv:2007.05436 [hep-lat].

[52] J. Horak, J. M. Pawlowski, and N. Wink, *Spectral functions in the ϕ^4-theory from the spectral DSE*, Phys. Rev. D **102**, 125016 (2020), arXiv:2006.09778 [hep-th].

[53] J. Horak, J. Papavassiliou, J. M. Pawlowski, and N. Wink, *Ghost spectral function from the spectral Dyson-Schwinger equation*, Phys. Rev. D **104**, 074017 (2021), arXiv:2103.16175 [hep-th].

[54] J. Braun *et al.*, *Renormalised spectral flows*, SciPost Phys. Core **6**, 061 (2023), arXiv:2206.10232 [hep-th].

[55] J. Horak, J. M. Pawlowski, and N. Wink, *On the quark spectral function in QCD*, SciPost Phys. **15**, 149 (2023), arXiv:2210.07597 [hep-ph].

[56] J. Horak, F. Ihssen, J. M. Pawlowski, J. Wessely, and N. Wink, *Scalar spectral functions from the spectral functional renormalization group*, Phys. Rev. D **110**, 056009 (2024), arXiv:2303.16719 [hep-th].

[57] J. M. Pawlowski and J. Wessely, *The causal structure of the quark propagator*, (2024), arXiv:2412.12033 [hep-ph].

[58] S. Floerchinger, *Analytic Continuation of Functional Renormalization Group Equations*, JHEP **05**, 021, arXiv:1112.4374 [hep-th].

[59] K. Kamikado, N. Strodthoff, L. von Smekal, and J. Wambach, *Real-time correlation functions in the $O(N)$ model from the functional renormalization group*, Eur. Phys. J. C **74**, 2806 (2014), arXiv:1302.6199 [hep-ph].

[60] R.-A. Tripolt, D. H. Rischke, L. von Smekal, and J. Wambach, *Fermionic excitations at finite temperature and density*, Phys. Rev. D **101**, 094010 (2020), arXiv:2003.11871 [hep-ph].

[61] T. Gasenzer and J. M. Pawlowski, *Towards far-from-equilibrium quantum field dynamics: A functional renormalisation-group approach*, Phys. Lett. B **670**, 135–140 (2008), arXiv:0710.4627 [cond-mat.other].

[62] J. M. Pawlowski and N. Strodthoff, *Real time correlation functions and the functional renormalization group*, Phys. Rev. D **92**, 094009 (2015), arXiv:1508.01160 [hep-ph].

[63] D. Mesterházy, J. H. Stockemer, L. F. Palhares, and J. Berges, *Dynamic universality class of Model C from the functional renormalization group*, Phys. Rev. B **88**, 174301 (2013), arXiv:1307.1700 [cond-mat.stat-mech].

[64] C. Duclut and B. Delamotte, *Frequency regulators for the nonperturbative renormalization group: A general study and the model A as a benchmark*, Phys. Rev. E **95**, 012107 (2017), arXiv:1611.07301 [cond-mat.stat-mech].

[65] J. M. Pawlowski, N. Strodthoff, and N. Wink, *Finite temperature spectral functions in the O(N)-model*, Phys. Rev. D **98**, 074008 (2018), arXiv:1711.07444 [hep-th].

[66] S. Huelsmann, S. Schlichting, and P. Scior, *Spectral functions from the real-time functional renormalization group*, Phys. Rev. D **102**, 096004 (2020), arXiv:2009.04194 [hep-ph].

[67] J. V. Roth, D. Schweitzer, L. J. Sieke, and L. von Smekal, *Real-time methods for spectral functions*, Phys. Rev. D **105**, 116017 (2022), arXiv:2112.12568 [hep-ph].

[68] J. V. Roth and L. von Smekal, *Critical dynamics in a real-time formulation of the functional renormalization group*, JHEP **10**, 065, arXiv:2303.11817 [hep-ph].

[69] Y.-r. Chen, Y.-y. Tan, and W.-j. Fu, *Critical dynamics within the real-time FRG approach*, Phys. Rev. D **109**, 094044 (2024), arXiv:2312.05870 [hep-ph].

[70] Y.-r. Chen, Y.-y. Tan, and W.-j. Fu, *Critical dynamics of model H within the real-time FRG approach*, Phys. Rev. D **111**, 094025 (2025), arXiv:2406.00679 [hep-ph].

[71] J. V. Roth, Y. Ye, S. Schlichting, and L. von Smekal, *Dynamic critical behavior of the chiral phase transition from the real-time functional renormalization group*, JHEP **01**, 118, arXiv:2403.04573 [hep-ph].

[72] J. V. Roth, Y. Ye, S. Schlichting, and L. von Smekal, *Universal critical dynamics near the chiral phase transition and the QCD critical point*, Phys. Rev. D **111**, L111901 (2025), arXiv:2409.14470 [hep-ph].

[73] J. V. Roth, Y. Ye, S. Schlichting, and L. von Smekal, *Effects of dissipation on phase diagram and bosonic excitations in the quark-meson model*, (2025), arXiv:2503.21746 [hep-ph].

[74] H. B. Meyer, *Transport Properties of the Quark-Gluon Plasma: A Lattice QCD Perspective*, Eur. Phys. J. A **47**, 86 (2011), arXiv:1104.3708 [hep-lat].

[75] A. I. Larkin and Y. N. Ovchinnikov, *Quasiclassical method in the theory of superconductivity*, Sov Phys JETP **28**, 1200–1205 (1969).

[76] J. Rammer, *Quantum Field Theory of Non-equilibrium States* (Cambridge University Press, 2007).

[77] J. S. Schwinger, *Brownian motion of a quantum oscillator*, J. Math. Phys. **2**, 407–432 (1961).

[78] L. V. Keldysh, *Diagram technique for nonequilibrium processes*, Zh. Eksp. Teor. Fiz. **47**, 1515–1527 (1964).

[79] J. Berges, *Introduction to nonequilibrium quantum field theory*, AIP Conf. Proc. **739**, 3–62 (2004), arXiv:hep-ph/0409233.

[80] R. van Leeuwen, N. Dahlen, G. Stefanucci, C.-O. Almbladh, and U. von Barth, Introduction to the keldysh formalism, in *Time-Dependent Density Functional Theory*, edited by M. A. Marques, C. A. Ullrich, F. Nogueira, A. Rubio, K. Burke, and E. K. U. Gross (Springer Berlin Heidelberg, Berlin, Heidelberg, 2006) pp. 33–59.

[81] J. Maciejko, *An introduction to nonequilibrium many-body theory*, Lecture Notes, Springer **9**, 104 (2007).

[82] L. M. Sieberer, M. Buchhold, and S. Diehl, *Keldysh Field Theory for Driven Open Quantum Systems*, Rept. Prog. Phys. **79**, 096001 (2016), arXiv:1512.00637 [cond-mat.quant-gas].

[83] G. Kallen, *On the definition of the Renormalization Constants in Quantum Electrodynamics*, Helv. Phys. Acta **25**, 417 (1952).

[84] H. Lehmann, *Über eigenschaften von ausbreitungsfunktionen und renormierungskonstanten quantisierter felder*, Il Nuovo Cimento **11**, 342–357 (1954).

[85] A. M. Jaffe and E. Witten, *Quantum Yang-Mills theory*, (2000).

[86] C. Michael and M. Teper, *The Glueball Spectrum in SU(3)*, Nucl. Phys. B **314**, 347–362 (1989).

[87] G. S. Bali, K. Schilling, A. Hulsebos, A. C. Irving, C. Michael, and P. W. Stephenson (UKQCD), *A Comprehensive lattice study of SU(3) glueballs*, Phys. Lett. B **309**, 378–384 (1993), arXiv:hep-lat/9304012.

[88] C. J. Morningstar and M. J. Peardon, *The Glueball spectrum from an anisotropic lattice study*, Phys. Rev. D **60**, 034509 (1999), arXiv:hep-lat/9901004.

[89] Y. Chen *et al.*, *Glueball spectrum and matrix elements on anisotropic lattices*, Phys. Rev. D **73**, 014516 (2006), arXiv:hep-lat/0510074.

[90] A. Athenodorou and M. Teper, *The glueball spectrum of SU(3) gauge theory in 3 + 1 dimensions*, JHEP **11**, 172, arXiv:2007.06422 [hep-lat].

[91] K. Sakai and S. Sasaki, *Glueball spectroscopy in lattice QCD using gradient flow*, Phys. Rev. D **107**, 034510 (2023), arXiv:2211.15176 [hep-lat].

[92] A. Szczepaniak, E. S. Swanson, C.-R. Ji, and S. R. Cotanch, *Glueball spectroscopy in a relativistic many body approach to hadron structure*, Phys. Rev. Lett. **76**, 2011–2014 (1996), arXiv:hep-ph/9511422.

[93] A. P. Szczepaniak and E. S. Swanson, *The Low lying glueball spectrum*, Phys. Lett. B **577**, 61–66 (2003), arXiv:hep-ph/0308268.

[94] V. Mathieu, N. Kochelev, and V. Vento, *The Physics of Glueballs*, Int. J. Mod. Phys. E **18**, 1–49 (2009), arXiv:0810.4453 [hep-ph].

[95] D. Dudal, M. S. Guimaraes, and S. P. Sorella, *Glueball masses from an infrared moment problem and nonperturbative Landau gauge*, Phys. Rev. Lett. **106**, 062003 (2011), arXiv:1010.3638 [hep-th].

[96] J. Meyers and E. S. Swanson, *Spin Zero Glueballs in the Bethe-Salpeter Formalism*, Phys. Rev. D **87**, 036009 (2013), arXiv:1211.4648 [hep-ph].

[97] I. C. Cloet and C. D. Roberts, *Explanation and Prediction of Observables using Continuum Strong QCD*, Prog. Part. Nucl. Phys. **77**, 1–69 (2014), arXiv:1310.2651 [nucl-th].

[98] H. Sanchis-Alepuz, C. S. Fischer, C. Kellermann, and L. von Smekal, *Glueballs from the Bethe-Salpeter equation*, Phys. Rev. D **92**, 034001 (2015), arXiv:1503.06051 [hep-ph].

[99] E. V. Souza, M. Narciso Ferreira, A. C. Aguilar, J. Papavassiliou, C. D. Roberts, and S.-S. Xu, *Pseudoscalar glueball mass: a window on three-gluon interactions*, Eur. Phys. J. A **56**, 25 (2020), arXiv:1909.05875 [nucl-th].

[100] L. P. Kaptari and B. Kämpfer, *Mass Spectrum of Pseudo-Scalar Glueballs from a Bethe–Salpeter Approach with the Rainbow–Ladder Truncation*, Few Body Syst. **61**, 28 (2020), arXiv:2004.06523 [hep-ph].

[101] M. Q. Huber, C. S. Fischer, and H. Sanchis-Alepuz, *Spectrum of scalar and pseudoscalar glueballs from functional methods*, Eur. Phys. J. C **80**, 1077 (2020), arXiv:2004.00415 [hep-ph].

[102] M. Q. Huber, C. S. Fischer, and H. Sanchis-Alepuz, *Higher spin glueballs from functional methods*, Eur. Phys. J. C **81**, 1083 (2021), arXiv:2110.09180 [hep-ph].

[103] M. Q. Huber, C. S. Fischer, and H. Sanchis-Alepuz, *Apparent convergence in functional glueball calculations*, (2025), arXiv:2503.03821 [hep-ph].

[104] V. N. Gribov, *Quantization of Nonabelian Gauge Theories*, Nucl. Phys. B **139**, 1 (1978).

[105] L. D. Faddeev and V. N. Popov, *Feynman Diagrams for the Yang-Mills Field*, Phys. Lett. B **25**, 29–30 (1967).

[106] D. J. Gross and F. Wilczek, *Ultraviolet Behavior of Nonabelian Gauge Theories*, Phys. Rev. Lett. **30**, 1343–1346 (1973).

[107] H. D. Politzer, *Reliable Perturbative Results for Strong Interactions?*, Phys. Rev. Lett. **30**, 1346–1349 (1973).

[108] S. Navas *et al.* (Particle Data Group), *Review of particle physics*, Phys. Rev. D **110**, 030001 (2024).

[109] F. Ihssen, J. M. Pawlowski, F. R. Sattler, and N. Wink, *Towards quantitative precision in functional QCD I*, (2024), arXiv:2408.08413 [hep-ph].

[110] R.-A. Tripolt, N. Strodthoff, L. von Smekal, and J. Wambach, *Spectral Functions for the Quark-Meson Model Phase Diagram from the Functional Renormalization Group*, Phys. Rev. D **89**, 034010 (2014), arXiv:1311.0630 [hep-ph].

[111] C. Jung, F. Rennecke, R.-A. Tripolt, L. von Smekal, and J. Wambach, *In-Medium Spectral Functions of Vector- and Axial-Vector Mesons from the Functional Renormalization Group*, Phys. Rev. D **95**, 036020 (2017), arXiv:1610.08754 [hep-ph].

[112] R. Kubo, *Statistical mechanical theory of irreversible processes. 1. General theory and simple applications in magnetic and conduction problems*, J. Phys. Soc. Jap. **12**, 570–586 (1957).

[113] M. L. Bellac, *Thermal Field Theory*, Cambridge Monographs on Mathematical Physics (Cambridge University Press, 2011).

[114] J. Mroczka and D. Szczuczyński, *Inverse problems formulated in terms of first-kind fredholm integral equations in indirect measurements*, Metrology and Measurement Systems **16**, 333–357 (2009).

[115] S. I. Kabanikhin, *Definitions and examples of inverse and ill-posed problems*, Journal of Inverse and Ill-posed Problems **16**, 317–357 (2008).

[116] J. G. McWhirter and E. R. Pike, *On the numerical inversion of the laplace transform and similar fredholm integral equations of the first kind*, Journal of Physics A: Mathematical and General **11**, 1729 (1978).

[117] A. Neumaier, *Solving ill-conditioned and singular linear systems: A tutorial on regularization*, SIAM Review **40**, 636–666 (1998).

[118] G. Cuniberti, E. De Micheli, and G. A. Viano, *Reconstructing the thermal Green functions at real times from those at imaginary times*, Commun. Math. Phys. **216**, 59–83 (2001), arXiv:cond-mat/0109175.

[119] M. Ulybyshev, C. Winterowd, and S. Zafeiropoulos, *Collective charge excitations and the metal-insulator transition in the square lattice Hubbard-Coulomb model*, Phys. Rev. B **96**, 205115 (2017), arXiv:1707.04212 [cond-mat.str-el].

[120] D. Dudal, O. Oliveira, M. Roelfs, and P. Silva, *Spectral representation of lattice gluon and ghost propagators at zero temperature*, Nucl. Phys. B **952**, 114912 (2020), arXiv:1901.05348 [hep-lat].

[121] D. Dudal, O. Oliveira, and M. Roelfs, *Källén–Lehmann spectral representation of the scalar SU(2) glueball*, Eur. Phys. J. C **82**, 251 (2022), arXiv:2103.11846 [hep-lat].

[122] H.-T. Ding, O. Kaczmarek, and F. Meyer, *Thermal dilepton rates and electrical conductivity of the QGP from the lattice*, Phys. Rev. D **94**, 034504 (2016), arXiv:1604.06712 [hep-lat].

[123] J. Ghiglieri, O. Kaczmarek, M. Laine, and F. Meyer, *Lattice constraints on the thermal photon rate*, Phys. Rev. D **94**, 016005 (2016), arXiv:1604.07544 [hep-lat].

[124] L. Altenkort, A. M. Eller, O. Kaczmarek, L. Mazur, G. D. Moore, and H.-T. Shu, *Sphaleron rate from Euclidean lattice correlators: An exploration*, Phys. Rev. D **103**, 114513 (2021), arXiv:2012.08279 [hep-lat].

[125] L. Altenkort, O. Kaczmarek, R. Larsen, S. Mukherjee, P. Petreczky, H.-T. Shu, and S. Stendebach (HotQCD), *Heavy Quark Diffusion from 2+1 Flavor Lattice QCD with 320 MeV Pion Mass*, Phys. Rev. Lett. **130**, 231902 (2023), arXiv:2302.08501 [hep-lat].

[126] Y. Burnier, M. Laine, and L. Mether, *A Test on analytic continuation of thermal imaginary-time data*, Eur. Phys. J. C **71**, 1619 (2011), arXiv:1101.5534 [hep-lat].

[127] A. K. Cyrol, J. M. Pawlowski, A. Rothkopf, and N. Wink, *Reconstructing the gluon*, SciPost Phys. **5**, 065 (2018), arXiv:1804.00945 [hep-ph].

[128] G. Backus and F. Gilbert, *The Resolving Power of Gross Earth Data*, Geophys. J. Int. **16**, 169–205 (1968).

[129] B. B. Brandt, A. Francis, H. B. Meyer, and D. Robaina, *Pion quasiparticle in the low-temperature phase of QCD*, Phys. Rev. D **92**, 094510 (2015), arXiv:1506.05732 [hep-lat].

[130] B. B. Brandt, A. Francis, B. Jäger, and H. B. Meyer, *Charge transport and vector meson dissociation across the thermal phase transition in lattice QCD with two light quark flavors*, Phys. Rev. D **93**, 054510 (2016), arXiv:1512.07249 [hep-lat].

[131] M. Hansen, A. Lupo, and N. Tantalo, *Extraction of spectral densities from lattice correlators*, Phys. Rev. D **99**, 094508 (2019), arXiv:1903.06476 [hep-lat].

[132] C. Alexandrou *et al.* (Extended Twisted Mass Collaboration (ETMC)), *Probing the Energy-Smeared R Ratio Using Lattice QCD*, Phys. Rev. Lett. **130**, 241901 (2023), arXiv:2212.08467 [hep-lat].

[133] M. Jarrell and J. E. Gubernatis, *Bayesian inference and the analytic continuation of imaginary-time quantum Monte Carlo data*, Phys. Rept. **269**, 133–195 (1996).

[134] M. Asakawa, T. Hatsuda, and Y. Nakahara, *Maximum entropy analysis of the spectral functions in lattice QCD*, Prog. Part. Nucl. Phys. **46**, 459–508 (2001), arXiv:hep-lat/0011040.

[135] M. Haas, L. Fister, and J. M. Pawlowski, *Gluon spectral functions and transport coefficients in Yang–Mills theory*, Phys. Rev. D **90**, 091501 (2014), arXiv:1308.4960 [hep-ph].

[136] A. Rothkopf, *Bayesian inference of real-time dynamics from lattice QCD*, Front. Phys. **10**, 1028995 (2022), arXiv:2208.13590 [hep-lat].

[137] Y. Burnier and A. Rothkopf, *Bayesian Approach to Spectral Function Reconstruction for Euclidean Quantum Field Theories*, Phys. Rev. Lett. **111**, 182003 (2013), arXiv:1307.6106 [hep-lat].

[138] A. Rothkopf, *Bayesian inference of nonpositive spectral functions in quantum field theory*, Phys. Rev. D **95**, 056016 (2017), arXiv:1611.00482 [hep-ph].

[139] L. Del Debbio, T. Giani, and M. Wilson, *Bayesian approach to inverse problems: an application to NNPDF closure testing*, Eur. Phys. J. C **82**, 330 (2022), arXiv:2111.05787 [hep-ph].

[140] S. Offler, G. Aarts, C. Allton, J. Glesaaen, B. Jäger, S. Kim, M. P. Lombardo, S. M. Ryan, and J.-I. Skullerud, *News from bottomonium spectral functions in thermal QCD*, PoS **LATTICE2019**, 076 (2019), arXiv:1912.12900 [hep-lat].

[141] S. Offler, G. Aarts, C. Allton, B. Jäger, S. Kim, M.-P. Lombardo, B. Page, S. M. Ryan, J.-I. Skullerud, and T. Spriggs, *Reconstruction of bottomonium spectral functions in thermal QCD using Kernel Ridge Regression*, PoS **LATTICE2021**, 509 (2022), arXiv:2112.02116 [hep-lat].

[142] L.-F. Arsenault, R. Neuberg, L. A. Hannah, and A. J. Millis, *Projected regression method for solving fredholm integral equations arising in the analytic continuation problem of quantum physics*, Inverse Problems **33**, 115007 (2017), arXiv:1612.04895 [cond-mat.str-el].

[143] L. Kades, J. M. Pawlowski, A. Rothkopf, M. Scherzer, J. M. Urban, S. J. Wetzel, N. Wink, and F. P. G. Ziegler, *Spectral Reconstruction with Deep Neural Networks*, Phys. Rev. D **102**, 096001 (2020), arXiv:1905.04305 [physics.comp-ph].

[144] J. Karpie, K. Orginos, A. Rothkopf, and S. Zafeiropoulos, *Reconstructing parton distribution functions from Ioffe time data: from Bayesian methods to Neural Networks*, JHEP **04**, 057, arXiv:1901.05408 [hep-lat].

[145] R. Fournier, L. Wang, O. V. Yazyev, and Q. S. Wu, *Artificial neural network approach to the analytic continuation problem*, Phys. Rev. Lett. **124**, 056401 (2020), arXiv:1810.00913 [physics.comp-ph].

[146] H. Yoon, J.-H. Sim, and M. J. Han, *Analytic continuation via domain knowledge free machine learning*, Phys. Rev, B **98**, 245101 (2018), arXiv:1806.03841 [cond-mat.str-el].

[147] M. Zhou, F. Gao, J. Chao, Y.-X. Liu, and H. Song, *Application of radial basis functions neutral networks in spectral functions*, Phys. Rev. D **104**, 076011 (2021), arXiv:2106.08168 [hep-ph].

[148] L. Wang, S. Shi, and K. Zhou, *Reconstructing spectral functions via automatic differentiation*, Phys. Rev. D **106**, L051502 (2022), arXiv:2111.14760 [hep-ph].

[149] L. Del Debbio, T. Giani, J. Karpie, K. Orginos, A. Radyushkin, and S. Zafeiropoulos, *Neural-network analysis of Parton Distribution Functions from Ioffe-time pseudodistributions*, JHEP **02**, 138, arXiv:2010.03996 [hep-ph].

[150] T. Lechien and D. Dudal, *Neural network approach to reconstructing spectral functions*

and complex poles of confined particles, SciPost Phys. **13**, 097 (2022), arXiv:2203.03293 [hep-lat].

[151] M. Buzzicotti, A. De Santis, and N. Tantalo, *Teaching to extract spectral densities from lattice correlators to a broad audience of learning-machines*, Eur. Phys. J. C **84**, 32 (2024), arXiv:2307.00808 [hep-lat].

[152] A. Tewari, T. Yin, G. Cazenavette, S. Rezchikov, J. Tenenbaum, F. Durand, B. Freeman, and V. Sitzmann, *Diffusion with forward models: Solving stochastic inverse problems without direct supervision*, Advances in Neural Information Processing Systems **36**, 12349–12362 (2023), arXiv:2306.11719 [cs.CV].

[153] P. Lowdon and O. Philipsen, *Pion spectral properties above the chiral crossover of QCD*, JHEP **10**, 161, arXiv:2207.14718 [hep-lat].

[154] D. Bala, O. Kaczmarek, P. Lowdon, O. Philipsen, and T. Ueding, *Pseudo-scalar meson spectral properties in the chiral crossover region of QCD*, JHEP **05**, 332, arXiv:2310.13476 [hep-lat].

[155] R.-A. Tripolt, P. Gubler, M. Ulybyshev, and L. Von Smekal, *Numerical analytic continuation of Euclidean data*, Comput. Phys. Commun. **237**, 129–142 (2019), arXiv:1801.10348 [hep-ph].

[156] D. Bala, O. Kaczmarek, R. Larsen, S. Mukherjee, G. Parkar, P. Petreczky, A. Rothkopf, and J. H. Weber (HotQCD), *Static quark-antiquark interactions at nonzero temperature from lattice QCD*, Phys. Rev. D **105**, 054513 (2022), arXiv:2110.11659 [hep-lat].

[157] J. Fei, C.-N. Yeh, and E. Gull, *Nevanlinna analytical continuation*, Phys. Rev. Lett. **126**, 056402 (2021), arXiv:2010.04572 [cond-mat.str-el].

[158] J. Fei, C.-N. Yeh, D. Zgid, and E. Gull, *Analytical continuation of matrix-valued functions: Carathéodory formalism*, Phys. Rev. B **104**, 165111 (2021).

[159] K. Nogaki, J. Fei, E. Gull, and H. Shinaoka, *Nevanlinna.jl: A julia implementation of nevanlinna analytic continuation*, SciPost Physics Codebases 10.21468/scipost-physcodeb.19 (2023).

[160] K. Nogaki and H. Shinaoka, *Bosonic nevanlinna analytic continuation*, Journal of the Physical Society of Japan **92**, 035001 (2023), https://doi.org/10.7566/JPSJ.92.035001.

[161] T. Bergamaschi, W. I. Jay, and P. R. Oare, *Hadronic structure, conformal maps, and analytic continuation*, Phys. Rev. D **108**, 074516 (2023), arXiv:2305.16190 [hep-lat].

[162] R. M. Neal, *Priors for infinite networks*, Bayesian learning for neural networks , 29–53 (1996).

[163] C. E. Rasmussen and C. K. I. Williams, *Gaussian Processes for Machine Learning* (MIT Press, 2006).

[164] Y. Engel, S. Mannor, and R. Meir, *Reinforcement learning with gaussian processes*, Proceedings of the 22nd International Conference on Machine Learning ICML '05, 201–208 (2005).

[165] A. Damianou and N. D. Lawrence, *Deep Gaussian processes*, *Proceedings of the Sixteenth International Conference on Artificial Intelligence and Statistics*, Proceedings of Machine Learning Research, **31**, 207–215 (2013), arXiv:1211.0358 [stat.ML].

[166] M. Raissi, P. Perdikaris, and G. E. Karniadakis, *Machine learning of linear differential equations using gaussian processes*, Journal of Computational Physics **348**, 683–693 (2017).

[167] M. Kuss, *Gaussian Process Models for Robust Regression, Classification, and Reinforcement Learning*, Ph.D. thesis, Technische Universität Darmstadt (2006).

[168] A. P. Valentine and M. Sambridge, *Gaussian process models—i. a framework for probabilistic continuous inverse theory*, Geophysical Journal International **220**, 1632–1647 (2020).

[169] M. Kanagawa, P. Hennig, D. Sejdinovic, and B. K. Sriperumbudur, *Gaussian processes and kernel methods: A review on connections and equivalences*, (2018), arXiv:1807.02582 [stat.ML].

[170] K. P. Murphy, *Machine learning: a probabilistic perspective* (MIT press, 2012).

[171] C. M. Bishop, *Pattern Recognition and Machine Learning (Information Science and Statistics)* (Springer-Verlag, Berlin, Heidelberg, 2006).

[172] J. Görtler, R. Kehlbeck, and O. Deussen, *A visual exploration of gaussian processes*, Distill 10.23915/distill.00017 (2019).

[173] D. Duvenaud, J. Lloyd, R. Grosse, J. Tenenbaum, and G. Zoubin, *Structure discovery in nonparametric regression through compositional kernel search*, *Proceedings of the 30th International Conference on Machine Learning*, Proceedings of Machine Learning Research, **28**, 1166–1174 (2013).

[174] F. Bach, *Exploring large feature spaces with hierarchical multiple kernel learning*, *Proceedings of the 22nd International Conference on Neural Information Processing Systems*, NIPS'08, 105–112 (2008).

[175] S. W. Ober, C. E. Rasmussen, and M. van der Wilk, *The promises and pitfalls of deep kernel learning*, *Proceedings of the Thirty-Seventh Conference on Uncertainty in Artificial Intelligence (UAI)*, Proceedings of Machine Learning Research, **161**, 1206–1216 (2021).

[176] H. Liu, Y.-S. Ong, X. Shen, and J. Cai, *When gaussian process meets big data: A review of scalable gps*, IEEE transactions on neural networks and learning systems **31**, 4405–4423 (2019), arXiv:1807.01065 [stat.ML].

[177] A. Berlinet and C. Thomas-Agnan, *Reproducing kernel Hilbert spaces in probability and statistics* (Springer Science & Business Media, 2011).

[178] I. Steinwart and A. Christmann, *Support vector machines* (Springer Science & Business Media, 2008).

[179] N. Aronszajn, *Theory of reproducing kernels*, Transactions of the American mathematical society **68**, 337–404 (1950).

[180] I. Steinwart, *Convergence types and rates in generic karhunen-loeve expansions with applications to sample path properties*, Potential Analysis **51**, 361–395 (2019), arXiv:1403.1040 [math.PR].

[181] J. Mercer, *Xvi. functions of positive and negative type, and their connection the theory of integral equations*, Philosophical transactions of the royal society of London. Series A, containing papers of a mathematical or physical character **209**, 415–446 (1909).

[182] F. Riesz and B. Sz.-Nagy, *Functional Analysis* (Dover Books on Mathematics, Dover, 1990).

[183] I. Steinwart, *On the influence of the kernel on the consistency of support vector machines*, J. Mach. Learn. Res. **2**, 67–93 (2002).

[184] I. Steinwart, D. Hush, and C. Scovel, *An explicit description of the reproducing kernel hilbert spaces of gaussian rbf kernels*, IEEE Transactions on Information Theory **52**, 4635–4643 (2006).

[185] H. Q. Minh, *Some properties of gaussian reproducing kernel hilbert spaces and their implications for function approximation and learning theory*, Constructive Approximation **32**, 307–338 (2010).

[186] A. P. Valentine and M. Sambridge, *Gaussian process models—ii. lessons for discrete inversion*, Geophysical Journal International **220**, 1648–1656 (2020).

[187] J. Horak, J. M. Pawlowski, J. Rodríguez-Quintero, J. Turnwald, J. M. Urban, N. Wink, and S. Zafeiropoulos, *Reconstructing QCD spectral functions with Gaussian processes*, Phys. Rev. D **105**, 036014 (2022), arXiv:2107.13464 [hep-ph].

[188] A. Candido, L. Del Debbio, T. Giani, and G. Petrillo, *Inverse Problems in PDF determinations*, PoS **LATTICE2022**, 098 (2023), arXiv:2302.14731 [hep-lat].

[189] A. Lupo, L. Del Debbio, M. Panero, and N. Tantalo, *Bayesian interpretation of Backus-Gilbert methods*, PoS **LATTICE2023**, 004 (2024), arXiv:2311.18125 [hep-lat].

[190] C. Huang, Y.-y. Tan, R. Wen, S. Yin, and W.-j. Fu, *Reconstruction of baryon number distributions**, Chin. Phys. C **47**, 104106 (2023), arXiv:2303.10869 [hep-ph].

[191] H. Dutrieux, J. Karpie, K. Orginos, and S. Zafeiropoulos, *Simple nonparametric reconstruction of parton distributions from limited Fourier information*, Phys. Rev. D **111**, 034515 (2025), arXiv:2412.05227 [hep-lat].

[192] L. Del Debbio, M. Naviglio, and F. Tarantelli, *Neural Networks Asymptotic Behaviours for the Resolution of Inverse Problems*, (2024), arXiv:2402.09338 [physics.comp-ph].

[193] L. Del Debbio, A. Lupo, M. Panero, and N. Tantalo, *Bayesian solution to the inverse problem and its relation to Backus–Gilbert methods*, Eur. Phys. J. C **85**, 185 (2025), arXiv:2409.04413 [hep-lat].

[194] A. Candido, L. Del Debbio, T. Giani, and G. Petrillo, *Bayesian inference with Gaussian processes for the determination of parton distribution functions*, Eur. Phys. J. C **84**, 716 (2024), arXiv:2404.07573 [hep-ph].

[195] GPy: A gaussian process framework in python, `http://github.com/SheffieldML/GPy` (2012).

[196] A. G. d. G. Matthews, M. van der Wilk, T. Nickson, K. Fujii, A. Boukouvalas, P. León-Villagrá, Z. Ghahramani, and J. Hensman, *GPflow: A Gaussian process library using TensorFlow*, Journal of Machine Learning Research **18**, 1–6 (2017).

[197] M. Abadi, A. Agarwal, P. Barham, E. Brevdo, Z. Chen, C. Citro, G. S. Corrado, A. Davis, J. Dean, M. Devin, S. Ghemawat, I. Goodfellow, A. Harp, G. Irving, M. Isard, Y. Jia, R. Jozefowicz, L. Kaiser, M. Kudlur, J. Levenberg, D. Mané, R. Monga, S. Moore, D. Murray, C. Olah, M. Schuster, J. Shlens, B. Steiner, I. Sutskever, K. Talwar, P. Tucker, V. Vanhoucke, V. Vasudevan, F. Viégas, O. Vinyals, P. Warden, M. Wattenberg, M. Wicke, Y. Yu, and X. Zheng, TensorFlow: Large-scale machine learning on heterogeneous systems (2015).

[198] J. R. Gardner, G. Pleiss, D. Bindel, K. Q. Weinberger, and A. G. Wilson, *Gpytorch: Blackbox matrix-matrix gaussian process inference with gpu acceleration* Advances in Neural Information Processing Systems 31, 10.5555/3327757.3327857 (2018), arXiv:1809.11165 [cs.LG].

[199] A. Paszke, S. Gross, F. Massa, A. Lerer, J. Bradbury, G. Chanan, T. Killeen, Z. Lin, N. Gimelshein, L. Antiga, A. Desmaison, A. Kopf, E. Yang, Z. DeVito, M. Raison, A. Tejani, S. Chilamkurthy, B. Steiner, L. Fang, J. Bai, and S. Chintala, *Pytorch: An imperative style, high-performance deep learning library*, Advances in Neural Information Processing Systems 32, 8024–8035 (2019), arXiv:1912.01703 [cs.LG].

[200] T. Pinder and D. Dodd, *Gpjax: A gaussian process framework in jax*, Journal of Open Source Software **7**, 4455 (2022).

[201] J. Bradbury, R. Frostig, P. Hawkins, M. J. Johnson, C. Leary, D. Maclaurin, G. Necula, A. Paszke, J. VanderPlas, S. Wanderman-Milne, and Q. Zhang, JAX: composable transformations of Python+NumPy programs (2018).

[202] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and E. Duchesnay, *Scikit-learn: Machine learning in Python*, Journal of Machine Learning Research **12**, 2825–2830 (2011).

[203] A. Riddell, A. Hartikainen, and M. Carter, pystan (3.0.0), PyPI (2021).

[204] O. Abril-Pla, V. Andreani, C. Carroll, L. Dong, C. J. Fonnesbeck, M. Kochurov, R. Kumar, J. Lao, C. C. Luhmann, O. A. Martin, *et al.*, *Pymc: a modern, and comprehensive probabilistic programming framework in python*, PeerJ Computer Science **9**, e1516 (2023).

[205] C. R. Harris, K. J. Millman, S. J. van der Walt, R. Gommers, P. Virtanen, D. Cournapeau, E. Wieser, J. Taylor, S. Berg, N. J. Smith, R. Kern, M. Picus, S. Hoyer, M. H. van Kerkwijk, M. Brett, A. Haldane, J. F. del Río, M. Wiebe, P. Peterson, P. Gérard-Marchant, K. Sheppard, T. Reddy, W. Weckesser, H. Abbasi, C. Gohlke, and T. E. Oliphant, *Array programming with NumPy*, Nature **585**, 357–362 (2020).

[206] P. Virtanen, R. Gommers, T. E. Oliphant, M. Haberland, T. Reddy, D. Cournapeau, E. Burovski, P. Peterson, W. Weckesser, J. Bright, S. J. van der Walt, M. Brett, J. Wilson, K. J. Millman, N. Mayorov, A. R. J. Nelson, E. Jones, R. Kern, E. Larson, C. J. Carey, İ. Polat, Y. Feng, E. W. Moore, J. VanderPlas, D. Laxalde, J. Perktold, R. Cimrman, I. Henriksen, E. A. Quintero, C. R. Harris, A. M. Archibald, A. H. Ribeiro, F. Pedregosa, P. van Mulbregt, and SciPy 1.0 Contributors, *SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python*, Nature Methods **17**, 261–272 (2020).

[207] C. Agrell, *Gaussian processes with linear operator inequality constraints*, Journal of Machine Learning Research **20**, 1–36 (2019).

[208] M. Osborne, R. Garnett, Z. Ghahramani, D. K. Duvenaud, S. J. Roberts, and C. Rasmussen, *Active learning of model evidence using bayesian quadrature*, Advances in Neural Information Processing Systems **25** (2012).

[209] T. Gunter, M. A. Osborne, R. Garnett, P. Hennig, and S. J. Roberts, *Sampling for inference in probabilistic models with fast bayesian quadrature*, Advances in neural information processing systems **27** (2014), arXiv:1411.0439 [stat.ML].

[210] E. Klempt and A. Zaitsev, *Glueballs, Hybrids, Multiquarks. Experimental facts versus QCD inspired concepts*, Phys. Rept. **454**, 1–202 (2007), arXiv:0708.4016 [hep-ph].

[211] V. Crede and C. A. Meyer, *The Experimental Status of Glueballs*, Prog. Part. Nucl. Phys. **63**, 74–116 (2009), arXiv:0812.0600 [hep-ex].

[212] W. Ochs, *The Status of Glueballs*, J. Phys. G **40**, 043001 (2013), arXiv:1301.5183 [hep-ph].

[213] E. Klempt, *Glueballs, a fulfilled promise of QCD?*, (2022), arXiv:2211.12901 [hep-ph].

[214] A. V. Sarantsev, I. Denisenko, U. Thoma, and E. Klempt, *Scalar isoscalar mesons and the scalar glueball from radiative J/ψ decays*, Phys. Lett. B **816**, 136227 (2021), arXiv:2103.09680 [hep-ph].

[215] E. Klempt and A. V. Sarantsev, *Singlet-octet-glueball mixing of scalar mesons*, Phys. Lett. B **826**, 136906 (2022), arXiv:2112.04348 [hep-ph].

[216] E. Gregory, A. Irving, B. Lucini, C. McNeile, A. Rago, C. Richards, and E. Rinaldi, *Towards the glueball spectrum from unquenched lattice QCD*, JHEP **10**, 170, arXiv:1208.1858 [hep-lat].

[217] F. Chen, X. Jiang, Y. Chen, K.-F. Liu, W. Sun, and Y.-B. Yang, *Glueballs at physical pion mass**, Chin. Phys. C **47**, 063108 (2023), arXiv:2111.11929 [hep-lat].

[218] R. Brett, J. Bulava, D. Darvish, J. Fallica, A. Hanlon, B. Hörz, and C. Morningstar, *Spectroscopy From The Lattice: The Scalar Glueball*, AIP Conf. Proc. **2249**, 030032 (2020), arXiv:1909.07306 [hep-lat].

[219] J. M. Pawlowski, C. S. Schneider, and N. Wink, *On Gauge Consistency In Gauge-Fixed Yang-Mills Theory*, (2022), arXiv:2202.11123 [hep-th].

[220] J. Horak, J. M. Pawlowski, and N. Wink, *On the complex structure of Yang-Mills theory*, (2022), arXiv:2202.09333 [hep-th].

[221] A. Bonanno, T. Denz, J. M. Pawlowski, and M. Reichert, *Reconstructing the graviton*, SciPost Phys. **12**, 001 (2022), arXiv:2102.02217 [hep-th].

[222] T. S. Evans, *N point finite temperature expectation values at real times*, Nucl. Phys. B **374**, 340–370 (1992).

[223] C. S. Fischer, A. Maas, and J. M. Pawlowski, *On the infrared behavior of Landau gauge Yang-Mills theory*, Annals Phys. **324**, 2408–2437 (2009), arXiv:0810.1987 [hep-ph].

[224] A. K. Cyrol, M. Q. Huber, and L. von Smekal, *A Dyson–Schwinger study of the four-gluon vertex*, Eur. Phys. J. C **75**, 102 (2015), arXiv:1408.5409 [hep-ph].

[225] N. Dupuis, L. Canet, A. Eichhorn, W. Metzner, J. M. Pawlowski, M. Tissier, and N. Wschebor, *The nonperturbative functional renormalization group and its applications*, Phys. Rept. **910**, 1–114 (2021), arXiv:2006.04853 [cond-mat.stat-mech].

[226] A. K. Cyrol, L. Fister, M. Mitter, J. M. Pawlowski, and N. Strodthoff, *Landau gauge Yang-Mills correlation functions*, Phys. Rev. D **94**, 054005 (2016), arXiv:1605.01856 [hep-ph].

[227] P. V. Landshoff and O. Nachtmann, *Vacuum Structure and Diffraction Scattering*, Z. Phys. C **35**, 405 (1987).

[228] A. Donnachie and P. V. Landshoff, *Gluon Condensate and Pomeron Structure*, Nucl. Phys. B **311**, 509–518 (1989).

[229] C. Ewerz and O. Nachtmann, *Towards a nonperturbative foundation of the dipole picture. I. Functional methods*, Annals Phys. **322**, 1635–1669 (2007), arXiv:hep-ph/0404254.

[230] C. Ewerz and O. Nachtmann, *Towards a nonperturbative foundation of the dipole picture. II. High energy limit*, Annals Phys. **322**, 1670–1726 (2007), arXiv:hep-ph/0604087.

[231] S. Donnachie, H. G. Dosch, O. Nachtmann, and P. Landshoff, *Pomeron physics and QCD*, Vol. 19 (Cambridge University Press, 2004).

[232] G. B. Bopsin, E. G. S. Luna, A. A. Natale, and M. Peláez, *Nonperturbative gluon exchange in pp elastic scattering at TeV energies*, Phys. Rev. D **107**, 114011 (2023), arXiv:2212.04007 [hep-ph].

[233] S. Zafeiropoulos, P. Boucaud, F. De Soto, J. Rodríguez-Quintero, and J. Segovia, *Strong Running Coupling from the Gauge Sector of Domain Wall Lattice QCD with Physical Quark Masses*, Phys. Rev. Lett. **122**, 162002 (2019), arXiv:1902.08148 [hep-ph].

[234] Z.-F. Cui, J.-L. Zhang, D. Binosi, F. de Soto, C. Mezrag, J. Papavassiliou, C. D. Roberts, J. Rodríguez-Quintero, J. Segovia, and S. Zafeiropoulos, *Effective charge from lattice QCD*, Chin. Phys. C **44**, 083102 (2020), arXiv:1912.08232 [hep-ph].

[235] A. K. Cyrol, M. Mitter, J. M. Pawlowski, and N. Strodthoff, *Nonperturbative quark, gluon, and meson correlators of unquenched QCD*, Phys. Rev. D **97**, 054006 (2018), arXiv:1706.06326 [hep-ph].

[236] F. Gao, J. Papavassiliou, and J. M. Pawlowski, *Fully coupled functional equations for the quark sector of QCD*, Phys. Rev. D **103**, 094013 (2021), arXiv:2102.13053 [hep-ph].

[237] K. A. Milton and O. P. Solovtsova, *Analytic perturbation theory: A New approach to the analytic continuation of the strong coupling constant alpha-s into the timelike region*, Phys. Rev. D **57**, 5402–5409 (1998), arXiv:hep-ph/9710316.

[238] A. I. Alekseev, *Strong coupling constant to four loops in the analytic approach to QCD*, Few Body Syst. **32**, 193–217 (2003), arXiv:hep-ph/0211339.

[239] A. Cucchieri, T. Mendes, and A. R. Taurines, *Positivity violation for the lattice Landau gluon propagator*, Phys. Rev. D **71**, 051902 (2005), arXiv:hep-lat/0406020.

[240] P. Lowdon, *Non-perturbative constraints on the quark and ghost propagators*, Nucl. Phys. B **935**, 242–255 (2018), arXiv:1711.07569 [hep-th].

[241] P. Lowdon, *On the analytic structure of QCD propagators*, PoS **Confinement2018**, 050 (2018), arXiv:1811.03037 [hep-th].

[242] R. Oehme and W. Zimmermann, *Gauge Field Propagator and the Number of Fermion Fields*, Phys. Rev. **D21**, 1661 (1980).

[243] R. Oehme, *On superconvergence relations in quantum chromodynamics*, Phys. Lett. **B252**, 641–646 (1990).

[244] F. D. R. Bonnet, P. O. Bowman, D. B. Leinweber, and A. G. Williams, *Infrared behavior of the gluon propagator on a large volume lattice*, Phys. Rev. D **62**, 051501(R) (2000), arXiv:hep-lat/0002020.

[245] A. Sternbeck, E. M. Ilgenfritz, M. Muller-Preussker, and A. Schiller, *Towards the infrared limit in SU(3) Landau gauge lattice gluodynamics*, Phys. Rev. D **72**, 014507 (2005), arXiv:hep-lat/0506007.

[246] P. Boucaud, J. P. Leroy, A. Le Yaouanc, A. Y. Lokhov, J. Micheli, O. Pene, J. Rodriguez-Quintero, and C. Roiesnel, *Asymptotic behavior of the ghost propagator in SU3 lattice gauge theory*, Phys. Rev. D **72**, 114503 (2005), arXiv:hep-lat/0506031.

[247] P. J. Silva and O. Oliveira, *Infrared Gluon Propagator from lattice QCD: Results from large asymmetric lattices*, Phys. Rev. D **74**, 034513 (2006), arXiv:hep-lat/0511043.

[248] A. Cucchieri, A. Maas, and T. Mendes, *Exploratory study of three-point Green's functions in Landau-gauge Yang-Mills theory*, Phys. Rev. D **74**, 014503 (2006), arXiv:hep-lat/0605011.

[249] A. Cucchieri, A. Maas, and T. Mendes, *Three-point vertices in Landau-gauge Yang-Mills theory*, Phys. Rev. D **77**, 094510 (2008), arXiv:0803.1798 [hep-lat].

[250] O. Oliveira and P. J. Silva, *Does The Lattice Zero Momentum Gluon Propagator for Pure Gauge SU(3) Yang-Mills Theory Vanish in the Infinite Volume Limit?*, Phys. Rev. D **79**, 031501(R) (2009), arXiv:0809.0258 [hep-lat].

[251] I. L. Bogolubsky, E. M. Ilgenfritz, M. Muller-Preussker, and A. Sternbeck, *Lattice gluodynamics computation of Landau gauge Green's functions in the deep infrared*, Phys. Lett. B **676**, 69–73 (2009), arXiv:0901.0736 [hep-lat].

[252] T. Iritani, H. Suganuma, and H. Iida, *Gluon-propagator functional form in the Landau gauge in SU(3) lattice QCD: Yukawa-type gluon propagator and anomalous gluon spectral function*, Phys. Rev. D **80**, 114505 (2009), arXiv:0908.1311 [hep-lat].

[253] A. Ayala, A. Bashir, D. Binosi, M. Cristoforetti, and J. Rodriguez-Quintero, *Quark flavour effects on gluon and ghost propagators*, Phys. Rev. D **86**, 074512 (2012), arXiv:1208.0795 [hep-ph].

[254] A. Athenodorou, D. Binosi, P. Boucaud, F. De Soto, J. Papavassiliou, J. Rodriguez-Quintero, and S. Zafeiropoulos, *On the zero crossing of the three-gluon vertex*, Phys. Lett. B **761**, 444–449 (2016), arXiv:1607.01278 [hep-ph].

[255] P. Boucaud, F. De Soto, J. Rodríguez-Quintero, and S. Zafeiropoulos, *Refining the detection of the zero crossing for the three-gluon vertex in symmetric and asymmetric momentum subtraction schemes*, Phys. Rev. D **95**, 114503 (2017), arXiv:1701.07390 [hep-lat].

[256] A. G. Duarte, O. Oliveira, and P. J. Silva, *Further Evidence For Zero Crossing On The Three Gluon Vertex*, Phys. Rev. D **94**, 074502 (2016), arXiv:1607.03831 [hep-lat].

[257] A. C. Aguilar, F. De Soto, M. N. Ferreira, J. Papavassiliou, J. Rodríguez-Quintero, and S. Zafeiropoulos, *Gluon propagator and three-gluon vertex with dynamical quarks*, Eur. Phys. J. C **80**, 154 (2020), arXiv:1912.12086 [hep-ph].

[258] A. C. Aguilar, F. De Soto, M. N. Ferreira, J. Papavassiliou, and J. Rodríguez-Quintero, *Infrared facets of the three-gluon vertex*, Phys. Lett. B **818**, 136352 (2021), arXiv:2102.04959 [hep-ph].

[259] A. C. Aguilar, C. O. Ambrósio, F. De Soto, M. N. Ferreira, B. M. Oliveira, J. Papavassiliou, and J. Rodríguez-Quintero, *Ghost dynamics in the soft gluon limit*, Phys. Rev. D **104**, 054028 (2021), arXiv:2107.00768 [hep-ph].

[260] M. N. Ferreira and J. Papavassiliou, *Gauge Sector Dynamics in QCD*, Particles **6**, 312–363 (2023), arXiv:2301.02314 [hep-ph].

[261] C. Allton *et al.* (RBC, UKQCD), *2+1 flavor domain wall QCD on a (2 fm)*83 lattice: Light meson spectroscopy with L(s) = 16*, Phys. Rev. **D76**, 014504 (2007), arXiv:hep-lat/0701013 [hep-lat].

[262] C. Allton *et al.* (RBC-UKQCD), *Physical Results from 2+1 Flavor Domain Wall QCD and SU(2) Chiral Perturbation Theory*, Phys. Rev. **D78**, 114509 (2008), arXiv:0804.0473 [hep-lat].

[263] R. Arthur *et al.* (RBC, UKQCD), *Domain Wall QCD with Near-Physical Pions*, Phys. Rev. **D87**, 094514 (2013), arXiv:1208.4412 [hep-lat].

[264] T. Blum *et al.* (RBC, UKQCD), *Domain wall QCD with physical quark masses*, Phys. Rev. **D93**, 074505 (2016), arXiv:1411.7017 [hep-lat].

[265] P. A. Boyle, L. Del Debbio, A. Jüttner, A. Khamseh, F. Sanfilippo, and J. T. Tsang, *The decay constants $\mathbf{f_D}$ and $\mathbf{f_{D_s}}$ in the continuum limit of $\mathbf{N_f = 2+1}$ domain wall lattice QCD*, JHEP **12**, 008, arXiv:1701.02644 [hep-lat].

[266] Y. Iwasaki, *Renormalization Group Analysis of Lattice Theories and Improved Lattice*

Action: Two-Dimensional Nonlinear O(N) Sigma Model, Nucl. Phys. **B258**, 141–156 (1985).

[267] D. B. Kaplan, *A Method for simulating chiral fermions on the lattice*, Phys. Lett. **B288**, 342–347 (1992), arXiv:hep-lat/9206013 [hep-lat].

[268] Y. Shamir, *Chiral fermions from lattice boundaries*, Nucl. Phys. **B406**, 90–106 (1993), arXiv:hep-lat/9303005 [hep-lat].

[269] P. Boucaud, F. De Soto, K. Raya, J. Rodríguez-Quintero, and S. Zafeiropoulos, *Discretization effects on renormalized gauge-field Green's functions, scale setting, and the gluon mass*, Phys. Rev. D **98**, 114515 (2018), arXiv:1809.05776 [hep-ph].

[270] A. Kızılersü, O. Oliveira, P. J. Silva, J.-I. Skullerud, and A. Sternbeck, *Quark-gluon vertex from Nf=2 lattice QCD*, Phys. Rev. D **103**, 114515 (2021), arXiv:2103.02945 [hep-lat].

[271] C. Gale and J. I. Kapusta, *Dilepton radiation from high temperature nuclear matter*, Phys. Rev. C **35**, 2107–2116 (1987).

[272] S. Turbide, R. Rapp, and C. Gale, *Hadronic production of thermal photons*, Phys. Rev. C **69**, 014903 (2004), arXiv:hep-ph/0308085.

[273] A. Adare *et al.* (PHENIX), *Enhanced production of direct photons in Au+Au collisions at $\sqrt{s_{NN}} = 200$ GeV and implications for the initial temperature*, Phys. Rev. Lett. **104**, 132301 (2010), arXiv:0804.4168 [nucl-ex].

[274] M. Wilde (ALICE), *Measurement of Direct Photons in pp and Pb-Pb Collisions with ALICE*, Nucl. Phys. A **904-905**, 573c–576c (2013), arXiv:1210.5958 [hep-ex].

[275] A. Adare *et al.* (PHENIX), *Centrality dependence of low-momentum direct-photon production in Au+Au collisions at $\sqrt{s_{NN}} = 200$ GeV*, Phys. Rev. C **91**, 064904 (2015), arXiv:1405.3940 [nucl-ex].

[276] A. Adare *et al.* (PHENIX), *Azimuthally anisotropic emission of low-momentum direct photons in Au+Au collisions at $\sqrt{s_{NN}} = 200$ GeV*, Phys. Rev. C **94**, 064901 (2016), arXiv:1509.07758 [nucl-ex].

[277] A. Adare *et al.* (PHENIX), *Observation of direct-photon collective flow in $\sqrt{s_{NN}} = 200$ GeV Au+Au collisions*, Phys. Rev. Lett. **109**, 122302 (2012), arXiv:1105.4126 [nucl-ex].

[278] D. Lohner and the ALICE Collaboration, *Measurement of direct-photon elliptic flow in pb-pb collisions at = 2.76 tev*, Journal of Physics: Conference Series **446**, 012028 (2013).

[279] R. Chatterjee, E. S. Frodermann, U. W. Heinz, and D. K. Srivastava, *Elliptic flow of thermal photons in relativistic nuclear collisions*, Phys. Rev. Lett. **96**, 202302 (2006), arXiv:nucl-th/0511079.

[280] H. van Hees, C. Gale, and R. Rapp, *Thermal Photons and Collective Flow at the Relativistic Heavy-Ion Collider*, Phys. Rev. C **84**, 054906 (2011), arXiv:1108.2131 [hep-ph].

[281] C. Shen, U. W. Heinz, J.-F. Paquet, and C. Gale, *Thermal photons as a quark-gluon plasma thermometer reexamined*, Phys. Rev. C **89**, 044910 (2014), arXiv:1308.2440 [nucl-th].

[282] J. Churchill, L. Du, C. Gale, G. Jackson, and S. Jeon, *Dilepton production at next-to-leading order and intermediate invariant-mass observables*, Phys. Rev. C **109**, 044915 (2024), arXiv:2311.06675 [nucl-th].

[283] J. Churchill, L. Du, C. Gale, G. Jackson, and S. Jeon, *Virtual Photons Shed Light on the Early Temperature of Dense QCD Matter*, Phys. Rev. Lett. **132**, 172301 (2024), arXiv:2311.06951 [nucl-th].

[284] A. K. Chaudhuri and B. Sinha, *Direct photon production from viscous QGP*, Phys. Rev. C **83**, 034905 (2011), arXiv:1101.3823 [nucl-th].

[285] J.-F. m. c. Paquet, C. Shen, G. S. Denicol, M. Luzum, B. Schenke, S. Jeon, and C. Gale, *Production of photons in relativistic heavy-ion collisions*, Phys. Rev. C **93**, 044906 (2016).

[286] G. Vujanovic, J.-F. Paquet, G. S. Denicol, M. Luzum, S. Jeon, and C. Gale, *Electromagnetic radiation as a probe of the initial state and of viscous dynamics in relativistic nuclear collisions*, Phys. Rev. C **94**, 014904 (2016), arXiv:1602.01455 [nucl-th].

[287] R. Baier, H. Nakkagawa, A. Niegawa, and K. Redlich, *Production rate of hard thermal photons and screening of quark mass singularity*, Z. Phys. C **53**, 433–438 (1992).

[288] J. I. Kapusta, P. Lichard, and D. Seibert, *High-energy photons from quark - gluon plasma versus hot hadronic gas*, Phys. Rev. D **44**, 2774–2788 (1991), [Erratum: Phys.Rev.D 47, 4171 (1993)].

[289] P. B. Arnold, G. D. Moore, and L. G. Yaffe, *Photon emission from ultrarelativistic plasmas*, JHEP **11**, 057, arXiv:hep-ph/0109064.

[290] P. B. Arnold, G. D. Moore, and L. G. Yaffe, *Photon emission from quark gluon plasma: Complete leading order results*, JHEP **12**, 009, arXiv:hep-ph/0111107.

[291] J. Ghiglieri, J. Hong, A. Kurkela, E. Lu, G. D. Moore, and D. Teaney, *Next-to-leading order thermal photon production in a weakly coupled quark-gluon plasma*, JHEP **05**, 010, arXiv:1302.5970 [hep-ph].

[292] M. Cè, T. Harris, H. B. Meyer, A. Steinberg, and A. Toniato, *Rate of photon production in the quark-gluon plasma from lattice QCD*, Phys. Rev. D **102**, 091501 (2020), arXiv:2001.03368 [hep-lat].

[293] M. Cè, T. Harris, A. Krasniqi, H. B. Meyer, and C. Török, *Photon emissivity of the quark-gluon plasma: A lattice QCD analysis of the transverse channel*, Phys. Rev. D **106**, 054501 (2022), arXiv:2205.02821 [hep-lat].

[294] H. B. Meyer, *Euclidean correlators at imaginary spatial momentum and their relation to the thermal photon emission rate*, Eur. Phys. J. A **54**, 192 (2018), arXiv:1807.00781 [hep-lat].

[295] M. Cè, T. Harris, A. Krasniqi, H. B. Meyer, and C. Török, *Probing the photon emissivity of the quark-gluon plasma without an inverse problem in lattice QCD*, Phys. Rev. D **109**, 014507 (2024), arXiv:2309.09884 [hep-lat].

[296] C. Gale and J. I. Kapusta, *Vector dominance model at finite temperature*, Nucl. Phys. B **357**, 65–89 (1991).

[297] L. D. McLerran and T. Toimela, *Photon and Dilepton Emission from the Quark - Gluon Plasma: Some General Considerations*, Phys. Rev. D **31**, 545 (1985).

[298] L. Altenkort, D. de la Cruz, O. Kaczmarek, R. Larsen, G. D. Moore, S. Mukherjee, P. Petreczky, H.-T. Shu, and S. Stendebach (HotQCD), *Quark Mass Dependence of Heavy Quark Diffusion Coefficient from Lattice QCD*, Phys. Rev. Lett. **132**, 051902 (2024), arXiv:2311.01525 [hep-lat].

[299] S. Caron-Huot, *Asymptotics of thermal spectral functions*, Phys. Rev. D **79**, 125009 (2009), arXiv:0903.3958 [hep-ph].

[300] B. B. Brandt, A. Francis, T. Harris, H. B. Meyer, and A. Steinberg, *An estimate for the thermal photon rate from lattice QCD*, EPJ Web Conf. **175**, 07044 (2018), arXiv:1710.07050 [hep-lat].

[301] R. Sommer, *Scale setting in lattice QCD*, PoS **LATTICE2013**, 015 (2014), arXiv:1401.3270 [hep-lat].

[302] A. Francis, O. Kaczmarek, M. Laine, T. Neuhaus, and H. Ohno, *Critical point and scale setting in SU(3) plasma: An update*, Phys. Rev. D **91**, 096002 (2015), arXiv:1503.05652 [hep-lat].

[303] E. Follana, Q. Mason, C. Davies, K. Hornbostel, G. P. Lepage, J. Shigemitsu, H. Trottier, and K. Wong (HPQCD, UKQCD), *Highly improved staggered quarks on the lattice, with applications to charm physics*, Phys. Rev. D **75**, 054502 (2007), arXiv:hep-lat/0610092.

[304] M. Luscher and P. Weisz, *On-Shell Improved Lattice Gauge Theories*, Commun. Math. Phys. **97**, 59 (1985), [Erratum: Commun.Math.Phys. 98, 433 (1985)].

[305] M. Lüscher and P. Weisz, *Computation of the action for on-shell improved lattice gauge theories at weak coupling*, Physics Letters B **158**, 250–254 (1985).

[306] A. Bazavov *et al.* (MILC), *Results for light pseudoscalar mesons*, PoS **LATTICE2010**, 074 (2010), arXiv:1012.0868 [hep-lat].

[307] S. Caron-Huot, *O(g) plasma effects in jet quenching*, Phys. Rev. D **79**, 065039 (2009), arXiv:0811.1603 [hep-ph].

[308] P. K. Kovtun and A. O. Starinets, *Quasinormal modes and holography*, Phys. Rev. D **72**, 086009 (2005), arXiv:hep-th/0506184.

[309] G. Jackson and M. Laine, *Testing thermal photon and dilepton rates*, JHEP **11**, 144, arXiv:1910.09567 [hep-ph].

[310] P. Aurenche, F. Gelis, G. D. Moore, and H. Zaraket, *Landau-Pomeranchuk-Migdal resummation for dilepton production*, JHEP **12**, 006, arXiv:hep-ph/0211036.

[311] B. Muller and A. Trayanov, *Deterministic chaos in nonAbelian lattice gauge theory*, Phys. Rev. Lett. **68**, 3387–3390 (1992).

[312] C.-q. Gong, *Lyapunov exponent of classical SU(3) gauge theory*, Phys. Lett. B **298**, 257–262 (1993), arXiv:hep-lat/9209018.

[313] T. S. Biro, C. Gong, and B. Muller, *Lyapunov exponent and plasmon damping rate in nonabelian gauge theories*, Phys. Rev. D **52**, 1260–1266 (1995), arXiv:hep-ph/9409392.

[314] T. Kunihiro, B. Muller, A. Ohnishi, A. Schafer, T. T. Takahashi, and A. Yamamoto, *Chaotic behavior in classical Yang-Mills dynamics*, Phys. Rev. D **82**, 114015 (2010), arXiv:1008.1156 [hep-ph].

[315] R. D. Pisarski and F. Wilczek, *Remarks on the Chiral Phase Transition in Chromodynamics*, Phys. Rev. D **29**, 338–341 (1984).

[316] H. T. Ding *et al.* (HotQCD), *Chiral Phase Transition Temperature in (2+1)-Flavor QCD*, Phys. Rev. Lett. **123**, 062002 (2019), arXiv:1903.04801 [hep-lat].

[317] J. Braun *et al.*, *Soft modes in hot QCD matter*, Phys. Rev. D **111**, 094010 (2025), arXiv:2310.19853 [hep-ph].

[318] E. P. Wigner, *Characteristic vectors of bordered matrices with infinite dimensions*, Annals of Mathematics **62**, 548–564 (1955).

[319] F. Haake, *Quantum Signatures of Chaos*, Springer Series in Synergetics (Springer, Berlin, 2010).

[320] R. U. Haq, A. Pandey, and O. Bohigas, *Fluctuation properties of nuclear energy levels: Do theory and experiment agree?*, Phys. Rev. Lett. **48**, 1086–1089 (1982).

[321] O. Bohigas, M. J. Giannoni, and C. Schmit, *Characterization of chaotic quantum spectra and universality of level fluctuation laws*, Phys. Rev. Lett. **52**, 1–4 (1984).

[322] M. V. Berry and M. Tabor, *Level clustering in the regular spectrum*, Proceedings of the Royal Society of London. A. Mathematical and Physical Sciences **356**, 375–394 (1977).

[323] J. M. Deutsch, *Quantum statistical mechanics in a closed system*, Phys. Rev. A **43**, 2046–2049 (1991).

[324] M. Srednicki, *Chaos and quantum thermalization*, Phys. Rev. E **50**, 888–901 (1994).

[325] M. Rigol, V. Dunjko, and M. Olshanii, *Thermalization and its mechanism for generic isolated quantum systems*, Nature **452**, 854–858 (2008), arXiv:0708.1324 [cond-mat.stat-mech].

[326] M. Srednicki, *The approach to thermal equilibrium in quantized chaotic systems*, Journal of Physics A: Mathematical and General **32**, 1163 (1999), arXiv:cond-mat/9809360 [cond-mat.stat-mech].

[327] J. M. Deutsch, *Eigenstate thermalization hypothesis*, Reports on Progress in Physics **81**, 082001 (2018), arXiv:1805.01616 [cond-mat.stat-mech].

[328] L. D'Alessio, Y. Kafri, A. Polkovnikov, and M. Rigol, *From quantum chaos and eigenstate thermalization to statistical mechanics and thermodynamics*, Advances in Physics **65**, 239–362 (2016), arXiv:1509.06411 [cond-mat.stat-mech].

[329] M. Rigol, *Breakdown of thermalization in finite one-dimensional systems*, Phys. Rev. Lett. **103**, 100403 (2009).

[330] L. Ebner, B. Müller, A. Schäfer, C. Seidl, and X. Yao, *Eigenstate thermalization in (2+1)-dimensional SU(2) lattice gauge theory*, Phys. Rev. D **109**, 014504 (2024), arXiv:2308.16202 [hep-lat].

[331] S. H. Shenker and D. Stanford, *Black holes and the butterfly effect*, JHEP **03**, 067, arXiv:1306.0622 [hep-th].

[332] S. H. Shenker and D. Stanford, *Stringy effects in scrambling*, JHEP **05**, 132, arXiv:1412.6087 [hep-th].

[333] A. Kitaev, *Hidden correlations in the hawking radiation and thermal noise*, (2014), talk given at Fundamental Physics Prize Symposium.

[334] J. Maldacena, S. H. Shenker, and D. Stanford, *A bound on chaos*, JHEP **08**, 106, arXiv:1503.01409 [hep-th].

[335] J. S. Cotler, D. Ding, and G. R. Penington, *Out-of-time-order Operators and the Butterfly Effect*, Annals Phys. **396**, 318–333 (2018), arXiv:1704.02979 [quant-ph].

[336] Y. Sekino and L. Susskind, *Fast Scramblers*, JHEP **10**, 065, arXiv:0808.2096 [hep-th].

[337] P. Hosur, X.-L. Qi, D. A. Roberts, and B. Yoshida, *Chaos in quantum channels*, JHEP **02**, 004, arXiv:1511.04021 [hep-th].

[338] B. Swingle, *Unscrambling the physics of out-of-time-order correlators*, Nature Phys. **14**, 988–990 (2018).

[339] I. García-Mata, R. A. Jalabert, and D. A. Wisniacki, *Out-of-time-order correlators and quantum chaos*, Scholarpedia **18**, 55237 (2023), arXiv:2209.07965 [quant-ph].

[340] E. H. Lieb and D. W. Robinson, *The finite group velocity of quantum spin systems*, Commun. Math. Phys. **28**, 251–257 (1972).

[341] D. A. Roberts and B. Swingle, *Lieb-Robinson Bound and the Butterfly Effect in Quantum Field Theories*, Phys. Rev. Lett. **117**, 091602 (2016), arXiv:1603.09298 [hep-th].

[342] A. A. Patel, D. Chowdhury, S. Sachdev, and B. Swingle, *Quantum butterfly effect in weakly interacting diffusive metals*, Phys. Rev. X **7**, 031047 (2017), arXiv:1703.07353 [cond-mat.str-el].

[343] M. Blake, *Universal Charge Diffusion and the Butterfly Effect in Holographic Theories*, Phys. Rev. Lett. **117**, 091601 (2016), arXiv:1603.08510 [hep-th].

[344] S. A. Hartnoll, *Theory of universal incoherent metallic transport*, Nature Phys. **11**, 54 (2015), arXiv:1405.3651 [cond-mat.str-el].

[345] A. A. Patel and S. Sachdev, *Quantum chaos on a critical Fermi surface*, Proc. Nat. Acad. Sci. **114**, 1844–1849 (2017), arXiv:1611.00003 [cond-mat.str-el].

[346] G. Aarts, C. Allton, A. Amato, P. Giudice, S. Hands, and J.-I. Skullerud, *Electrical conductivity and charge diffusion in thermal QCD from the lattice*, JHEP **02**, 186, arXiv:1412.6411 [hep-lat].

[347] D. A. Roberts, D. Stanford, and L. Susskind, *Localized shocks*, JHEP **03**, 051, arXiv:1409.8180 [hep-th].

[348] A. Kitaev, *Simple model of quantum holography*, (2015), talk given at KITP Program: Entanglement in Strongly-Correlated Quantum Matter.

[349] S. Sachdev and J. Ye, *Gapless spin fluid ground state in a random, quantum Heisenberg magnet*, Phys. Rev. Lett. **70**, 3339 (1993), arXiv:cond-mat/9212030.

[350] J. Maldacena and D. Stanford, *Remarks on the Sachdev-Ye-Kitaev model*, Phys. Rev. D **94**, 106002 (2016), arXiv:1604.07818 [hep-th].

[351] D. Stanford, *Many-body chaos at weak coupling*, JHEP **10**, 009, arXiv:1512.07687 [hep-th].

[352] I. L. Aleiner, L. Faoro, and L. B. Ioffe, *Microscopic model of quantum butterfly effect: out-of-time-order correlators and traveling combustion waves*, Annals Phys. **375**, 378–406 (2016), arXiv:1609.01251 [cond-mat.stat-mech].

[353] I. Kukuljan, S. Grozdanov, and T. Prosen, *Weak Quantum Chaos*, Phys. Rev. B **96**, 060301 (2017), arXiv:1701.09147 [cond-mat.stat-mech].

[354] E. B. Rozenbaum, S. Ganeshan, and V. Galitski, *Universal level statistics of the out-of-time-ordered operator*, Phys. Rev. B **100**, 035112 (2019), arXiv:1801.10591 [cond-mat.dis-nn].

[355] M. Brenes, S. Pappalardi, M. T. Mitchison, J. Goold, and A. Silva, *Out-of-time-order correlations and the fine structure of eigenstate thermalization*, Phys. Rev. E **104**, 034120 (2021), arXiv:2103.01161 [cond-mat.stat-mech].

[356] F. M. Haehl, R. Loganayagam, and M. Rangamani, *Schwinger-Keldysh formalism. Part I: BRST symmetries and superspace*, JHEP **06**, 069, arXiv:1610.01940 [hep-th].

[357] F. M. Haehl, R. Loganayagam, P. Narayan, and M. Rangamani, *Classification of out-of-time-order correlators*, SciPost Phys. **6**, 001 (2019), arXiv:1701.02820 [hep-th].

[358] S. Grozdanov, K. Schalm, and V. Scopelliti, *Kinetic theory for classical and quantum many-body chaos*, Phys. Rev. E **99**, 012206 (2019), arXiv:1804.09182 [hep-th].

[359] W. Fischler, T. Guglielmo, and P. Nguyen, *Quantum chaos in a weakly-coupled field theory with nonlocality*, JHEP **09**, 097, arXiv:2111.10895 [hep-th].

[360] S. Jeon, *Hydrodynamic transport coefficients in relativistic scalar field theory*, Phys. Rev. D **52**, 3591–3642 (1995), arXiv:hep-ph/9409250.

[361] D. Chowdhury and B. Swingle, *Onset of many-body chaos in the $O(N)$ model*, Phys. Rev. D **96**, 065005 (2017), arXiv:1703.02545 [cond-mat.str-el].

[362] A. Schuckert and M. Knap, *Many-body chaos near a thermal phase transition*, SciPost Phys. **7**, 022 (2019), arXiv:1905.00904 [cond-mat.stat-mech].

[363] S. Guin, H. Pandey, and S. Sharma, *Understanding thermalization in a non-Abelian gauge theory in terms of its soft modes*, Phys. Lett. B **866**, 139490 (2025), arXiv:2501.04397 [hep-lat].

[364] J. Berges, S. Schlichting, and D. Sexty, *Dynamic critical phenomena from spectral functions on the lattice*, Nucl. Phys. B **832**, 228–240 (2010), arXiv:0912.3135 [hep-lat].

[365] T. Gasenzer, B. Nowak, and D. Sexty, *Charge separation in Reheating after Cosmological Inflation*, Phys. Lett. B **710**, 500–503 (2012), arXiv:1108.0541 [hep-ph].

[366] A. Piñeiro Orioli, K. Boguslavski, and J. Berges, *Universal self-similar dynamics of relativistic and nonrelativistic field theories near nonthermal fixed points*, Phys. Rev. D **92**, 025041 (2015), arXiv:1503.02498 [hep-ph].

[367] G. D. Moore, *Condensates in Relativistic Scalar Theories*, Phys. Rev. D **93**, 065043 (2016), arXiv:1511.00697 [hep-ph].

[368] G. Aarts, *Spectral function at high temperature in the classical approximation*, Phys. Lett. B **518**, 315–322 (2001), arXiv:hep-ph/0108125.

[369] G. Aarts and J. Berges, *Classical aspects of quantum fields far from equilibrium*, Phys. Rev. Lett. **88**, 041603 (2002), arXiv:hep-ph/0107129.

[370] S. Schlichting, D. Smith, and L. von Smekal, *Spectral functions and critical dynamics of the O(4) model from classical-statistical lattice simulations*, Nucl. Phys. B **950**, 114868 (2020), arXiv:1908.00912 [hep-lat].

[371] D. Schweitzer, S. Schlichting, and L. von Smekal, *Spectral functions and dynamic critical behavior of relativistic* Z_2 *theories*, Nucl. Phys. B **960**, 115165 (2020), arXiv:2007.03374 [hep-lat].

[372] A. Florio, E. Grossi, A. Soloviev, and D. Teaney, *Dynamics of the* $O(4)$ *critical point in QCD*, Phys. Rev. D **105**, 054512 (2022), arXiv:2111.03640 [hep-lat].

[373] A. Florio, E. Grossi, and D. Teaney, *Dynamics of the O(4) critical point in QCD: Critical pions and diffusion in model G*, Phys. Rev. D **109**, 054037 (2024), arXiv:2306.06887 [hep-lat].

[374] J. Bezanson, A. Edelman, S. Karpinski, and V. B. Shah, *Julia: A fresh approach to numerical computing*, SIAM Review **59**, 65–98 (2017).

[375] S. Omlin and L. Räss, High-performance xpu stencil computations in julia (2022), arXiv:2211.15634 [cs.DC].

[376] S. Omlin, L. Räss, and I. Utkin, Distributed parallelization of xpu stencil computations in julia (2022), arXiv:2211.15716 [cs.DC].

[377] M. Campisi and J. Goold, *Thermodynamics of quantum information scrambling*, Phys. Rev. E **95**, 062127 (2017), arXiv:1609.05848 [quant-ph].

[378] B. Swingle, G. Bentsen, M. Schleier-Smith, and P. Hayden, *Measuring the scrambling of quantum information*, Phys. Rev. A **94**, 040302 (2016), arXiv:1602.06271 [quant-ph].

[379] H. Kim and D. A. Huse, *Ballistic Spreading of Entanglement in a Diffusive Nonintegrable System*, Phys. Rev. Lett. **111**, 127205 (2013).

[380] R. R. Parwani, *Resummation in a hot scalar field theory*, Phys. Rev. D **45**, 4695 (1992), [Erratum: Phys.Rev.D 48, 5965 (1993)], arXiv:hep-ph/9204216.

[381] E.-k. Wang and U. W. Heinz, *The plasmon in hot phi**4 theory*, Phys. Rev. D **53**, 899–910 (1996), arXiv:hep-ph/9509333.

[382] G. Aarts and J. Smit, *Finiteness of hot classical scalar field theory and the plasmon damping rate*, Phys. Lett. B **393**, 395–402 (1997), arXiv:hep-ph/9610415.

[383] S. Jeon and L. G. Yaffe, *From quantum field theory to hydrodynamics: Transport coefficients and effective kinetic theory*, Phys. Rev. D **53**, 5799–5809 (1996), arXiv:hep-ph/9512263.

[384] Y. Gu, X.-L. Qi, and D. Stanford, *Local criticality, diffusion and chaos in generalized Sachdev-Ye-Kitaev models*, JHEP **05**, 125, arXiv:1609.07832 [hep-th].

[385] D. T. Son and M. A. Stephanov, *Real time pion propagation in finite temperature QCD*, Phys. Rev. D **66**, 076011 (2002), arXiv:hep-ph/0204226.

[386] E. Grossi, A. Soloviev, D. Teaney, and F. Yan, *Soft pions and transport near the chiral critical point*, Phys. Rev. D **104**, 034025 (2021), arXiv:2101.10847 [nucl-th].

[387] T. Mori, T. N. Ikeda, E. Kaminishi, and M. Ueda, *Thermalization and prethermalization in isolated quantum systems: a theoretical overview*, J. Phys. B **51**, 112001 (2018), arXiv:1712.08790 [cond-mat.stat-mech].

[388] K. Mallayya, M. Rigol, and W. D. Roeck, *Prethermalization and Thermalization in Isolated Quantum Systems*, Phys. Rev. X **9**, 021027 (2019).

[389] K. X. Wei, P. Peng, O. Shtanko, I. Marvian, S. Lloyd, C. Ramanathan, and P. Cappellaro, *Emergent prethermalization signatures in out-of-time ordered correlations*, Phys. Rev. Lett. **123**, 090605 (2019), arXiv:1812.04776 [quant-ph].

[390] J. S. Cotler, G. Gur-Ari, M. Hanada, J. Polchinski, P. Saad, S. H. Shenker, D. Stanford, A. Streicher, and M. Tezuka, *Black Holes and Random Matrices*, JHEP **05**, 118, [Erratum: JHEP 09, 002 (2018)], arXiv:1611.04650 [hep-th].

Acknowledgements

First of all, I would like to thank Nicolas Wink for being a great supervisor. I learned a great deal from him and greatly enjoyed our discussions. Thank you for your patience and open-mindedness, which allowed me to develop my own ideas and methods.

I am very grateful to Guy Moore, who was my de jure supervisor and provided invaluable support, especially during the final stages of my PhD. He helped in many ways, from insightful discussions to enabling a research visit and supporting me through the final steps of my thesis.

I would also like to thank Jens Braun for agreeing to be my second referee.

Many thanks to all those I had the opportunity to work with on multiple projects throughout my PhD. In particular, I would like to thank Julian Urban and Dibyendu Bala for being great collaborators, and for the many fruitful discussions and learning experiences.

I would also like to thank all of my colleagues at the Institute for Nuclear Physics, who made my time there very enjoyable. Special thanks go to Andreas Geißel, Hosein Gholami, Simon Stendebach, Jonas Winter, Niklas Zorbach and all the other students who shared this time in Darmstadt with me. Thanks for all the discussions, some about physics and some not so much.

A heartfelt thank you goes to Derek Teaney, who made it possible for me to visit Stony Brook University for a research stay and generously shared his time for discussions. I also want to thank the Nuclear Theory group at Stony Brook for warmly welcoming me and making my stay an enjoyable and enriching experience.

I also thank Andreas, Astrid, Guy, Julian, Nicolas and Simon for providing valuable feedback on my thesis and proofreading.

I am incredibly grateful to my family, who supported me during my PhD and all the years before. Finally, and most importantly, I want to thank Sherin for supporting me in every way possible.

Erklärung laut Promotionsordnung

§ 8 Abs. 1 lit. c PromO

Ich versichere hiermit, dass die elektronische Version meiner Dissertation mit der schriftlichen Version übereinstimmt.

§ 8 Abs. 1 lit. d PromO

Ich versichere hiermit, dass zu einem vorherigen Zeitpunkt noch keine Promotion versucht wurde. In diesem Fall sind nähere Angaben über Zeitpunkt, Hochschule, Dissertationsthema und Ergebnis dieses Versuchs mitzuteilen.

§ 9 Abs. 1 PromO

Ich versichere hiermit, dass die vorliegende Dissertation selbstständig und nur unter Verwendung der angegebenen Quellen verfasst wurde.

§ 9 Abs. 2 PromO

Die Arbeit hat bisher noch nicht zu Prüfungszwecken gedient.

Darmstadt, den

Jonas Turnwald