Machine Learning for Path Deformation
and
Bayesian Data Analysis
in
Selected Lattice Field Theories

Dissertation
zur
Erlangung des Doktorgrades (Dr. rer. nat.)
der
Mathematisch-Naturwissenschaftlichen Fakultät
der
Rheinischen Friedrich-Wilhelms-Universität Bonn

vorgelegt von

Marcel Rodekamp

aus
Bielefeld

Bonn, 2024

Abstract

Our collective understanding of the laws of nature has a long history of an intricate interplay between *theoretical* considerations and *experimental* falsification. As computational power increases, *simulations*, at the interface between theory and experiment, have taken an increasing role in scientific discovery. In particular, first-principles calculations are indispensable for systems with non-perturbative behavior, requiring simulations to test models against experiment. One widely accepted and deployed method involves formulating the theory on a finite lattice and then applying a Monte Carlo simulation. However, with increasing interest in such simulations practical and fundamental challenges arise, such as computational demand and the numerical sign problem.

In the following, I discuss selected aspects for simulations of strongly correlated systems, namely the Hubbard model and lattice quantum chromodynamics. This encompasses methods to mitigate the numerical sign problem and (Bayesian) analysis of simulation results, in particular fitting methods and the treatment of excited state contamination.

The Hubbard model describes systems of strongly correlated electrons and is used often in studying chemical compounds. First principle studies of this model are almost exclusively done using Monte Carlo techniques, with the exception being very small systems where direct diagonalization methods are feasible. However, away from half filling, Monte Carlo methods struggle because of the numerical sign problem. While the sign problem is unlikely to be completely solved, methods that reduce its impact are very valuable in expanding the computable parameter space. Leveraging theoretical developments on path deformations, I demonstrate that machine learning techniques can be used to mitigate the sign problem. In particular, I train complex-valued neural networks to serve as a parameterization of a sign-optimized manifold related to Lefschetz thimbles. These methods were developed and tested on doped graphene sheets, modelled by a small number of ions with periodic boundary conditions, at fixed temporal discretization and temperature.

Renewable energy is a critical aspect of modern research to reduce effects of climate change. Despite the enormous energy cost of producing solar panels, they are a valuable element in the electricity production. Organic solar cells show great promise in reducing costs and allowing for flexibility. Unfortunately, to date their efficiency falls behind their silicon-based competitors. By studying the electronic structure of certain chemical compounds that are usable for organic solar cells, further development in this area can be fostered. This motivates my work in the molecule $C_{20}H_{12}$ perylene, which can be used as an acceptor material in organic solar cells. This molecule is typically not at half filling, so any simulation requires methods to mitigate the sign problem. The study of perylene shown here requires the analysis of a large data set, of the order of $O(2000)$ correlators, which is only feasible with an automated analysis procedure. In this thesis I present such an automatic routine based on Bayesian analysis using the Akaike information criterion.

Finally, I shift the focus to particle physics to calculate aspects of the internal structure of hadrons. Hadrons are primarily governed by the strong interaction, i.e. described by quantum chromodynamics (QCD). In this theory, the internal structure is modelled by the correlation between spatial and momentum distributions of all constituents. Many details of these distributions remain to be calculated. In this thesis, I use lattice QCD to calculate the 2nd moment of parton distribution functions (PDFs) for the nucleon. These are the average momentum fractions carried by the considered parton of the nucleon. I analyze two ensembles at the physical pion mass to obtain the moments of unpolarized, polarized, and transversity PDF for the nucleon.

Acknowledgments

Contents

List of Contributions

Presented Contributions

Marcel Rodekamp, Evan Berkowitz, Christoph Gäntgen, Stefan Krieg, Thomas Luu, and Johann Ostmeyer. "Mitigating the Hubbard sign problem with complex-valued neural networks." In: *Phys. Rev. B* 106 (12 Sept. 2022), p. 125139. DOI: 10.1103/PhysRevB.106.125139. arXiv: 2203.00390

→ Presented in chapter 2.

Marcel Rodekamp, Evan Berkowitz, Maria Dincă, Christoph Gäntgen, Stefan Krieg, and Thomas Luu. "From Theory to Practice: Applying Neural Networks to Simulate Real Systems with Sign Problems." In: vol. LATTICE2023. 2023. DOI: 10.22323/1.453.0031. arXiv: 2311.18312

→ Presented in chapter 3.

Marcel Rodekamp, Evan Berkowitz, Christoph Gäntgen, Stefan Krieg, Thomas Luu, Johann Ostmeyer, and Giovanni Pederiva. *Single Particle Spectrum of Doped* $C_{20}H_{12}$-*Perylene*. June 2024. arXiv: 2406.06711

→ Presented in chapter 4.

Marcel Rodekamp, Michael Engelhardt, Jeremy R. Green, Stefan Krieg, Simonetta Liuti, Stefan Meinel, John W. Negele, Andrew Pochinsky, and Sergey Syritsyn. "Moments of nucleon unpolarized, polarized, and transversity parton distribution functions from lattice QCD at the physical point." In: *Physical Review D* 109.7 (Apr. 2024), p. 074508. ISSN: 2470-0010, 2470-0029. DOI: 10.1103/PhysRevD.109.074508. arXiv: 2306.14431v2

→ Presented in chapter 5.

Additional Contributions and Co-Authorships

Marcel Rodekamp and Christoph Gäntgen. "Mitigating the Hubbard Sign Problem. A Novel Application of Machine Learning." In: vol. LATTICE2022. 2023. DOI: 10.22323/1.430.0032. arXiv: 2211.09584

→ Presented in chapter A.

Marcel Rodekamp, Michael Engelhardt, Jeremy R. Green, Stefan Krieg, Stefan Meinel, John W. Negele, Andrew Pochinsky, and Sergey" Syritsyn. "Moments of Parton Distributions Functions from Lattice QCD at the Physical Point." In: *30th International Workshop on Deep-Inelastic Scattering and Related Subjects*. June 2023. arXiv: 2306.14431

→ Presented in chapter B.

Christoph Gäntgen, Evan Berkowitz, Thomas Luu, Johann Ostmeyer, and Marcel Rodekamp. "Fermionic sign problem minimization by constant path integral contour shifts." In: *Physical Review B* 109.19 (May 2024), p. 195158. ISSN: 2469-9950, 2469-9969. DOI: 10.1103/PhysRevB.109.195158. arXiv: 2307.06785v1

Christoph Gäntgen, Evan Berkowitz, Thomas Luu, Johann Ostmeyer, and Marcel Rodekamp. "Reducing the Sign Problem with simple Contour Deformation." In: 2024. URL: https://pos.sissa.it/453/007/

Michael Engelhardt, Nesreen Hasan, Simonetta Liuti, Stefan Meinel, John W. Negele, Andrew Pochinsky, Marcel Rodekamp, and Sergey Syritsyn. "Quark orbital angular momentum in the proton from a twist-3 generalized parton distribution." In: vol. LATTICE2023. 2023. DOI: https://doi.org/10.22323/1.453.0316. arXiv: 2406.00845

Marcel Rodekamp, Evan Berkowitz, Christoph Gäntgen, Stefan Krieg, Thomas Luu, Giovanni Pederiva, Petar Sinilkov, Marc Túnica, and Keshvi Tuteja. *Nanosystem Simulation Library (NSL)*. URL: https://github.com/Marcel-Rodekamp/NSL

→ Used in chapter 3 and 4.

Marcel Rodekamp and Giovanni Pederiva. *Correlator Analyser*. URL: https://github.com/Marcel-Rodekamp/CorrelatorAnalyser

→ Used in chapter 4.

Chapter 1

Introduction

The study of physical systems necessitates the formulation of an underlying model or theory. In the contemporary era, these are frequently expressed as (Quantum) Field Theories (QFTs), wherein observables are defined by an infinite-dimensional integral, referred to as the path integral

$$\langle O \rangle = \frac{1}{\mathcal{Z}} \int \mathcal{D}\left[\Phi\right] e^{-\mathrm{S}\left[\Phi\right]} O\left[\Phi\right], \tag{1.1}$$

with partition function $\mathcal{Z} = \int \mathcal{D}\left[\Phi\right] e^{-\mathrm{S}\left[\Phi\right]}$. The information about the system is provided by the action $\mathrm{S}\left[\Phi\right]$ and the physical properties are extracted by choosing appropriate observables O. In general, integrations of this form are challenging and analytic solutions are not known for most interesting systems. However, using numerical integration is a well-established method in such situations. To this end, we first introduce a regulator by discretizing space and time onto a lattice Λ, which translates into a path integral measure of the form

$$\mathcal{D}\left[\Phi\right] = \lim_{|\Lambda| \to \infty} C(\Lambda) \prod_{n=0}^{|\Lambda|-1} \mathrm{d}\Phi(\mathbf{x}_n, t_n). \tag{1.2}$$

It is important to exercise caution when taking the limit, which can be done by integrating at finite but increasing volume $|\Lambda|$ and extrapolation.[1] Once translated onto a finite grid of points $(\mathbf{x}_n, t_n)$, one usually refers to a **lattice QFT**. The integral, at finite $|\Lambda|$, is most efficiently calculated using Monte Carlo techniques. The central focus here is to generate random (field-)variables Φ that are distributed according to the Boltzmann distribution

$$\mathrm{p_S}\left[\Phi\right] = \frac{1}{\mathcal{Z}} e^{-\mathrm{S}\left[\Phi\right]}. \tag{1.3}$$

Once an ensemble of configurations $\{\Phi_n \sim \mathrm{p_S}\left[\Phi_n\right]\}_{n=0}^{\mathrm{N_{cfg}}-1}$ is obtained, an observable can be estimated via importance sampling

$$\langle O \rangle = \lim_{\mathrm{N_{cfg}} \to \infty} \sum_{n=0}^{\mathrm{N_{cfg}}-1} O\left[\Phi_n\right]. \tag{1.4}$$

In practice, only a finite number of configurations ($\mathrm{N_{cfg}} < \infty$) can be calculated at a fixed computational cost, which introduces a statistical uncertainty,

$$\sigma_O \propto {}^1\!\big/\!\sqrt{\mathrm{N_{cfg}}}. \tag{1.5}$$

---

[1] Many details are omitted in this expression, for example having multiple fields contributing, the directions can have independent infinite volume limits, the physical volume should be treated properly, and so on. Thus this expression has to be taken with care and has to be explicitly defined for a theory of interest, see subsequent sections.

1

A meticulous data analysis via bootstrap or jackknife is essential to estimate σ_O. It is crucial to recognize, that systematic sources of uncertainty may arise and must be estimated on a case by case basis. The majority of computer time is typically allocated to the generation of ensembles, where numerous algorithms have been developed for diverse use cases. Each algorithm obeys inherent benefits and drawbacks, encompassing aspects such as applicability, scalability, overall cost, and more. Throughout this thesis, different variants of the Hamiltonian Monte Carlo (HMC) are employed due to their potential scalability to large-scale systems. The topic of lattice QFT is exceedingly complex, encompassing a multitude of intricate details Consequently, a comprehensive introduction in this thesis is not feasible. A plethora of textbooks have delved into this process and known pitfalls in more depth, for example [12, 13, 14, 15].

To this point, the discussion has been kept quite general in order to facilitate the two problem sets presented in this thesis. In the following, chapters 2 to 4, aspects of Monte Carlo simulations of the Hubbard model, which describes carbon nanostructures such as graphene and the molecule $C_{20}H_{12}$ perylene, are discussed. In chapter 5, the focus shifts towards lattice Quantum Chromodynamics (QCD) and the data analysis to obtain moments of polarized, unpolarized, and transversity parton distribution functions of nucleons.

1.1 The Numerical Sign Problem

One detail, however, that comes up during chapters 2 to 4, needs to be discussed in more depth, namely the numerical sign problem [16]. Numerical integration is generally unfeasible to apply when the integrand of interest is strongly oscillatory. In such cases, delicate cancellation forces any integration method to evaluate a potentially gigantic number of supporting points, or configurations for Monte Carlo integrations, to achieve a desired precision. Especially, in systems of (strongly) interacting fermions, where the problem is cast into a lattice QFT, the sign problem can become insurmountable with finite computational resources.

To understand the presence of the oscillating integrands one has to consider that fermions obey the Pauli exclusion principle. This means, that the fermionic wave functions change sign when two fermions are interchanged. A changing sign, therefore, means an oscillation which only increases with the number of fermions considered, as the interchange happens more frequently. Now notice, that typical path integrals do not directly involve fermions, but are rather cast into a form that deals only with bosonic fields. This is typically achieved by integrating out the fermionic degrees of freedom using the Matthews–Salam formula [17, 18, 14]

$$\det\{M\} = \int \mathcal{D}\left[\psi, \overline{\psi}\right] e^{\overline{\psi} \cdot M \cdot \psi}. \tag{1.6}$$

Here, $\psi = (\psi_1, \ldots, \psi_{|\Lambda|})$, and $\overline{\psi} = (\overline{\psi}_1, \ldots, \overline{\psi}_{|\Lambda|})$ denote a set of generators for the $2|\Lambda|$-Grassman algebra, representing the fermionic fields. Additionally, the $|\Lambda| \times |\Lambda|$ matrix $M \in \text{Mat}_{|\Lambda|}(\mathbb{C})$ is called the fermion matrix. The latter depends on the action of interest and in particular on the discretization of the continuum theory [19]. Furthermore, depending on the form of the fermion matrix the sign problem may arise as an oscillating complex phase problem due to a complex valued determinants entering the action. [2] This is typically also called the sign problem. Many (quantum) field theories with fermions are plagued by some form of the sign problem. However, in certain situations one can choose or design fermion determinants that are manifestly real and positive. For example, one might argue for using $\det\{MM^\dagger\} = |\det\{M\}|^2$. In particular, in lattice QCD, discussed in chapter 5, this is done by treating the light (up- and down-) quarks with equal mass, namely (2+X) flavors, and exploiting the γ_5-hermiticity of the Wilson fermion matrix [14].

---

[2]It is also possible to find a determinant that is sign indefinite, opposed to complex valued. Also for an indefinite sign delicate cancellations are present requiring a very precise integration. This forms the original meaning of "sign problem". For the sake of the following chapters, we will discuss the complex action case only.

Coming back to the Monte Carlo integration, with a sign problem the Boltzmann distribution is not a proper probability distribution anymore as

$$S[\Phi] \in \mathbb{C} \Rightarrow p_B[\Phi] \sim e^{-\mathrm{Re}\{S[\Phi]\}} e^{-i\,\mathrm{Im}\{S[\Phi]\}} \in \mathbb{C} \not\subseteq [0,1] \tag{1.7}$$

This formal issue, however, can be resolved by implementing the additional phase $e^{-i\,\mathrm{Im}\{S[\Phi]\}}$ into the observable and generating the Markov chain using the real part of the action only. Resulting, the phase reweighted observables

$$\begin{aligned}
\langle O \rangle \equiv \langle O \rangle_S &= \frac{\int \mathcal{D}[\Phi]\, e^{-\mathrm{Re}\{S[\Phi]\}-i\,\mathrm{Im}\{S[\Phi]\}} O[\Phi]}{\int \mathcal{D}[\Phi]\, e^{-\mathrm{Re}\{S[\Phi]\}-i\,\mathrm{Im}\{S[\Phi]\}}} \\[2mm]
&= \frac{1/Z_{\mathrm{Re}\{S\}} \int \mathcal{D}[\Phi]\, e^{-\mathrm{Re}\{S[\Phi]\}-i\,\mathrm{Im}\{S[\Phi]\}} O[\Phi]}{1/Z_{\mathrm{Re}\{S\}} \int \mathcal{D}[\Phi]\, e^{-\mathrm{Re}\{S[\Phi]\}-i\,\mathrm{Im}\{S[\Phi]\}}} \\[2mm]
&= \frac{\left\langle O e^{-i\,\mathrm{Im}\{S\}} \right\rangle_{\mathrm{Re}\{S\}}}{\left\langle e^{-i\,\mathrm{Im}\{S\}} \right\rangle_{\mathrm{Re}\{S\}}},
\end{aligned} \tag{1.8}$$

where the subscript action encodes how the Markov chain is generated. This poses no solution to the sign problem, since the integrand is still oscillating strongly but gives room for improvement. In fact, it has been shown [20] that the statistical uncertainty, compare equation 1.5, is related to the magnitude of the phase as

$$N_{\mathrm{cfg}} \rightarrow N_{\mathrm{cfg}}^{\mathrm{eff}} = \left| \left\langle e^{-i\,\mathrm{Im}\{S\}} \right\rangle \right|^2 N_{\mathrm{cfg}}. \tag{1.9}$$

Noticing that the average phase decays with the extensive phase-quenched potential [20] of the theory, implies an exponential decay in space time volume $\beta|\Lambda|$

$$\left| \left\langle e^{-i\,\mathrm{Im}\{S\}} \right\rangle \right| \sim e^{-\beta|\Lambda|} \tag{1.10}$$

Many techniques have been studied to improve upon the sign problem, including deformations to Lefschetz thimbles [21], complex Langevin [22], and many more. In this thesis, a particular focus lays on contour deformation towards sign-optimized manifolds inspired by Lefschetz thimbles [23]. However, it is expected that the sign problem is NP-complete [16], making it very unlikely that a general solution can be found. Thus, all methods discussed here thus must be understood as partial improvements to equation (1.10) to extend the efficiently-computable parameter space of a given theory.

1.2 The Hubbard Model

The Hubbard model is a quantum model that describes system's of strongly correlated electrons on a fixed spatial lattice. Its Hamiltonian reads

$$\mathcal{H}[\kappa, U, \mu] = -\kappa \sum_{\langle x,y \rangle \in X} \left(p_x^\dagger p_y - h_x^\dagger h_y \right) + \frac{U}{2} \sum_{x \in X} q_x^2 - \mu \sum_{x \in X} q_x. \tag{1.11}$$

Notice, this form makes a particular choice of basis, namely the particle-hole basis. In terms of electron creation and annihilation operators, $e_\sigma^\dagger$ and e_σ, respectively, where σ denotes the spin, $\sigma \in \{\uparrow, \downarrow\}$, the particle and hole operators are

$$p_x^\dagger = e_{\uparrow,x}^\dagger \qquad\qquad p_x = e_{\uparrow,x} \tag{1.12}$$

$$h_x^\dagger = e_{\downarrow,x} \qquad\qquad h_x = e_{\downarrow,x}^\dagger. \tag{1.13}$$

3

We choose this basis as, for $\mu = 0$, the systems ground state is half filled and thus electrically neutral. In particular, carbon-based molecules that are sp^2 hybridized are excellent targets for this model. Here the σ-electrons form the bonds of the ions that give the spatial lattice a fixed form. One should note that the Hubbard model is only an effective approximation, as for example vibrational modes are not included. This approximation holds especially in cold temperatures. The lattice of ions is denoted by X, this can include a lattice with boundary conditions mimicking an infinite system but also a finite set of ions with open boundary conditions. For later reference, the number of ions is typically written as $N_x = |X|$. We refer to ions, as only the σ-electrons are assumed to be fixed leaving possibly empty orbitals. In addition, the π-electrons are free to interact and move on the lattice. Sometimes, within this thesis, a lattice is understood as a graph giving the ions the notion of sites in the lattice.

The Hubbard model incorporates the hopping of electrons between neighboring sites along the bonds with the first term in equation 1.11. Whereby, neighbouring sites are denoted by $\langle x, y \rangle \in X$. Typically, it is called tight-binding, Hückel, or non-interacting term as it corresponds to the non-interacting theory, i.e. at vanishing interaction $U = 0$. The hopping strength, in general depending on the bond length, is denoted by κ. For simulations it is convenient to normalize the hopping strength, $\kappa = 1$, and provide everything in relative units, i.e. U/κ, μ/κ, and so on; this will be suppressed unless otherwise stated. The non-interacting theory can be solved analytically by considering that the hopping terms can be written in matrix form

$$\kappa \sum_{\langle x, y \rangle \in X} \left(p_x^\dagger p_y - h_x^\dagger h_y \right) \rightarrow \sum_{x, y \in X} \left(p_x^\dagger K^{xy} p_y - h_x^\dagger K^{xy} h_y \right) \tag{1.14}$$

with the hopping matrix $K = \kappa \delta_{\langle x, y \rangle}$ formed with the hopping strength and adjacency matrix

$$\delta_{\langle x, y \rangle} = \begin{cases} 1 & x \text{ is nearest neighbour of } y \\ 0 & \text{else.} \end{cases} \tag{1.15}$$

This explicitly assumes uniform hopping strength but can simply be extended to a more general setup by relating $\delta_{\langle x, y \rangle} \sim \kappa_{x,y}$. A few constraints need to be fulfilled such as $K = K^\dagger$. Exemplary hopping matrices, used in the following chapters, are plotted alongside their graph representation in figures 1.1 and 1.2. The former represents a graphene sheet built out of $N_x = 18$ ions with periodic boundary conditions drawn as dotted lines [24, 25]. Additionally, the latter represents the molecule $C_{20}H_{12}$ perylene built upon $N_x = 20$ ions, no periodic boundaries are applied as this is the entire molecule [26, 3]. The blue encoded bonds correspond to the squares on the right side. The position of a colored square indicates a non-zero entry in the adjacency matrix.

Now, finding a unitary transformation that diagonalizes the hopping matrix

$$\mathfrak{U} K \mathfrak{U}^\dagger = \text{diag}(\epsilon_k) \tag{1.16}$$

solves the non-interacting theory. It can be shown that each row of $\mathfrak{U}$ can be associated with an operator $p_k^\dagger = \mathfrak{U}_k^x p_x^\dagger$ which commutes with the non-interacting Hamiltonian $H_0 = -\kappa \sum_{\langle x, y \rangle \in X} \left(p_x^\dagger p_y - h_x^\dagger h_y \right)$

$$\left[H_0, p_k^\dagger \right] = \epsilon_k p_k^\dagger \tag{1.17}$$

For translationally invariant (infinite) systems with periodic boundary conditions, for example the 18-site problem, the unitary transformation can be found by a lattice Fourier transformation defining a lattice momentum k. For systems with open boundary, for example perylene, the unitary transformation can be found by constructing eigenvectors that transform irreducibly under the respective symmetry group, here D_{2h}. This defines indices $k = \Lambda_i$, which label the corresponding irreducible representation Λ and potentially an index i of mixing states.
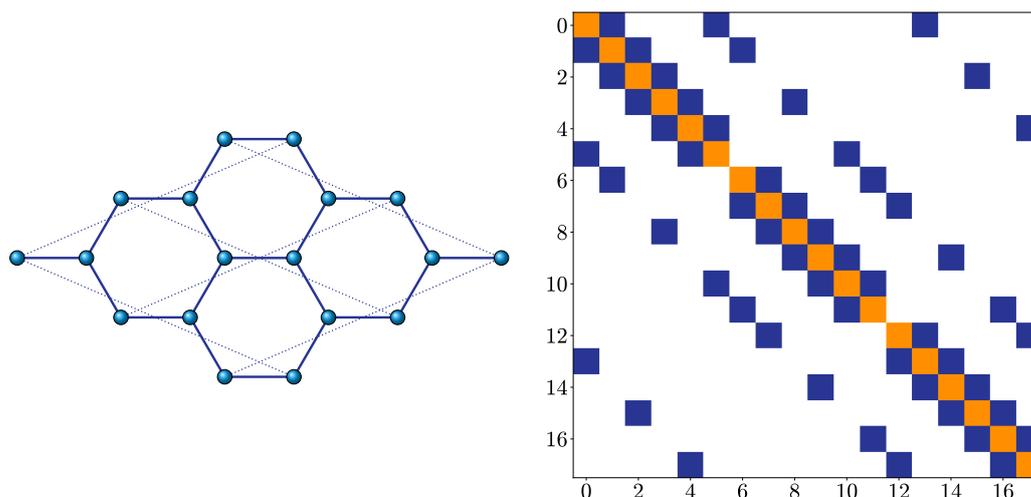
Figure 1.1: Left, lattice of ions for the 18-site problem representing a very small piece of graphene, used in chapter 2. Ions are shown as vertices, light blue balls, and allowed hopping is indicated through edges, blue lines. The dotted lines explicitly show the periodic boundary conditions. Right, the adjacency matrix as heat map, the exact form depends on the numbering of vertices. Blue squares represent non-zero entries in the hopping matrix, while the orange squares indicate those entries affected by chemical potential.
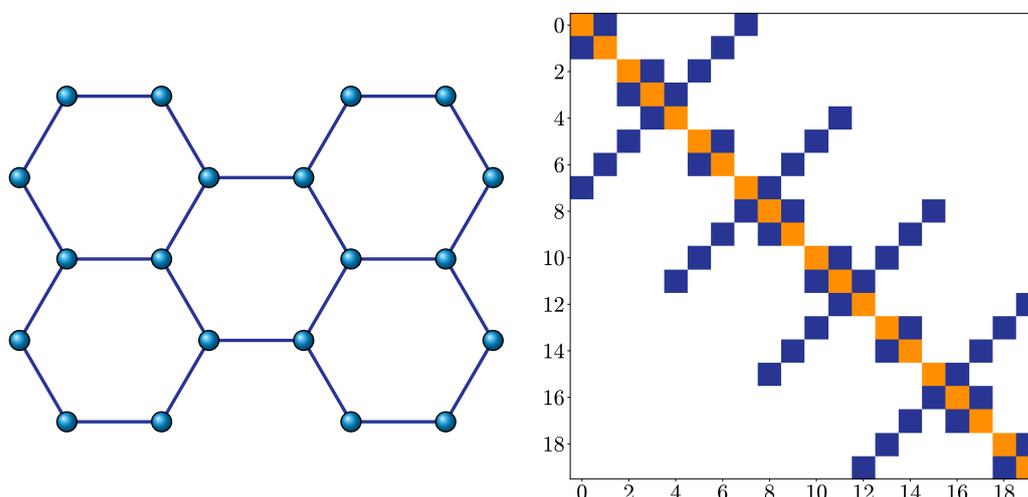
Figure 1.2: Left, lattice of ions for the molecule $C_{20}H_{12}$ perylene, used in chapters 3 and 4. Ions are, again, shown as vertices, light blue balls, and allowed hopping is indicated through edges, blue lines. This molecule, has only a finite amount of ions, hence the lattice is set up with open boundaries. Right, the adjacency matrix as heat map, the exact form depends on the numbering of vertices. Blue squares, again, represent non-zero entries in the hopping matrix, while the orange squares indicate those entries affected by chemical potential.

Defects or doping are modelled by adding a chemical potential, μ, coupling to the charge expressed as the difference between number of holes and particles $q_x = h_x^\dagger h_x - p_x^\dagger p_x$[3]. The third term of the Hamiltonian in equation (1.11) represents this. This does not add any additional complication to the analytical solution as it can be understood as diagonal terms on the hopping matrix, consequently, being drawn in figures 1.1 and 1.2 as orange squares on the diagonal. For simplicity, the chemical potential models an average doping over the entire lattice. However, it is simple to extend this by introducing a site dependent chemical potential at the cost of breaking certain symmetries of the hopping term. This simplification requires to calculate the total system charge

$$\langle Q \rangle = \sum_{x \in X} \langle q_x \rangle, \tag{1.18}$$

to quantify the effect of a certain chemical potential value.

In addition to the tight-binding dynamics, electrons can interact amongst themselves via the Coulomb interaction. This interaction is screened due to the σ electrons and the core ions, and can be modelled as an on-site effective interaction, U, that couples to the square of the local charge, q_x^2. This represents one of the simplest approximations to electron correlations in such low-dimensional systems. Extending the interaction beyond on-site then leads to so called extended Hubbard models, for example the Pariser–Parr–Pople model [27, 28].

Once interactions are introduced, an analytic solution is out of reach for larger systems. This is where Monte Carlo methods may be used to get further understanding of the physics. As a quantum model, this does not provide a path integral out of the box. However, a connection can be made to the thermal trace [29, 30, 30, 31, 24]

$$\langle O \rangle = \frac{\text{tr}\{e^{-\beta \mathcal{H}} O\}}{\text{tr}\{e^{-\beta \mathcal{H}}\}}, \tag{1.19}$$

defining observables. Concentrating on the partition function, the first step is to split the exponential into N_t 'time' slices

$$\mathcal{Z} = \text{tr}\{e^{-\beta \mathcal{H}}\} = \text{tr}\Big\{\underbrace{e^{\delta \mathcal{H}} \cdot e^{\delta \mathcal{H}} \cdots e^{\delta \mathcal{H}}}_{N_t \text{ factors}}\Big\}, \tag{1.20}$$

where the temporal lattice spacing $\delta = \beta/N_t$ is introduced. The thermal trace makes use of Euclidean time, that does not coincide with real time, and implies (anti-)periodic boundary conditions. In comparison to equation (1.2), this forms the only discretization we have to perform as the spatial components are naturally discrete on the set of ions.

The second step in arriving at a path integral is to insert $2N_t$ Grassmanian resolutions of identity, one N_t-set for the particles (Ψ) and one N_t-set for the holes (η), using coherent states

$$\mathbb{1}_t = \int \mathcal{D}\left[\Psi_t^*, \Psi_t, \eta_t^*, \eta_t\right] e^{-\sum_{x \in X} \Psi_{t,x}^* \Psi_{t,x} + \eta_{t,x}^* \eta_{t,x}} |\Psi_t, \eta_t\rangle\langle\Psi_t, \eta_t| \tag{1.21}$$

between the N_t factors and evaluating the trace in this basis. One arrives at [24, 32]

$$\mathcal{Z} = \int \mathcal{D}\left[\Psi^*, \Psi, \eta^*, \eta\right] \prod_{t=0}^{N_t-1} e^{-\sum_{x \in X} \Psi_{t+1,x}^* \Psi_{t+1,x} + \eta_{t+1,x}^* \eta_{t+1,x}} \left\langle\Psi_{t+1,x}, \eta_{t+1,x}\Big|e^{\delta \mathcal{H}}\Big|\Psi_{t,x}, \eta_{t,x}\right\rangle \tag{1.22}$$

To continue we can focus on the matrix elements $\left\langle\Psi_{t+1,x}, \eta_{t+1,x}\big|e^{\delta \mathcal{H}}\big|\Psi_{t,x}, \eta_{t,x}\right\rangle$ first. Inserting the Hubbard Hamiltonian, equation (1.11), into the exponent takes the form

$$\left\langle\Psi_{t+1,x}, \eta_{t+1,x}\big|e^{\delta \mathcal{H}}\big|\Psi_{t,x}, \eta_{t,x}\right\rangle \tag{1.23}$$
$$= \left\langle\Psi_{t+1,x}, \eta_{t+1,x}\big|e^{-\delta \kappa \sum_{\langle x,y \rangle \in X}\left(p_x^\dagger p_y - h_x^\dagger h_y\right) + \frac{\delta U}{2}\sum_{x \in X} q_x^2 - \delta \mu \sum_{x \in X} q_x}\big|\Psi_{t,x}, \eta_{t,x}\right\rangle$$

---

[3]It should be noted, that in the chapters 2 and 3, a different convention is employed, where μ and q_x are expressed with a negative sign. The convention presented here aligns with that of chapter 4.

We can split the exponential using the Baker-Campbell-Dynkin-Hausdorff formula in a specific form. Let $\mathcal{K}$ be the hopping contribution including the chemical potential and $\mathcal{V}$ be the interation part, a symmetric trotterization

$$e^{\delta\mathcal{K}+\delta\mathcal{V}} = e^{\frac{\delta}{2}\mathcal{K}} \cdot e^{\delta\mathcal{V}} \cdot e^{\frac{\delta}{2}\mathcal{K}} + O\left(\delta^3\right) \tag{1.24}$$

leads to order $O\left(\delta^3\right)$ errors. Using this result we can express the exponential of equation (1.23) by

$$e^{-\frac{\delta}{2}\kappa\sum_{\langle x,y\rangle\in X}\left(p_x^\dagger p_y - h_x^\dagger h_y\right) - \frac{\delta}{2}\mu\sum_{x\in X}q_x} \cdot e^{\frac{\delta U}{2}\sum_{x\in X}q_x^2} \cdot e^{-\frac{\delta}{2}\kappa\sum_{\langle x,y\rangle\in X}\left(p_x^\dagger p_y - h_x^\dagger h_y\right) - \frac{\delta}{2}\mu\sum_{x\in X}q_x}, \tag{1.25}$$

where we suppress the error $O\left(\delta^3\right)$. At this point it is important to emphasize that this exponent contains quartic terms of fermionic ladder operators, as $q_x^2 \sim p_x^\dagger p_x h_x^\dagger h_x$. These will later be troublesome and have to be taken care of. Thus a linearization of the q_x's is insurmountable; Using a continuous Hubbard-Stratonovich transformation can achieve this at the cost of introducing an auxiliary field $(\Phi_{t,x}) \in \mathbb{R}^{N_t \times N_x}$,

$$e^{\frac{\delta U}{2}\sum_{x\in X}q_x^2} \propto \int \mathcal{D}\left[\Phi_t\right] e^{\frac{1}{2\delta U}\sum_{x\in X}\Phi_{t,x}^2 - i\sum_{x\in X}\Phi_{t,x}q_x}, \tag{1.26}$$

where the overall normalization is irrelevant as it cancels in the observables as usual. Notice, this is not the only choice of the Hubbard-Stratonovich transformation, naturally one can do it in the real channel $i\Phi \to \Phi$ if the sign of the interaction U changes or in the spin formulation of the Hubbard model, going back to the electron ladder operators. Also a discrete Hubbard-Stratonovich transformation could be utilized. All these result in different formulations of the Hubbard model with different advantages and drawbacks. This particular choice is made as a continuous field allows one to use the Hamiltonian Monte Carlo algorithm to generate configurations; the choice of the imaginary channel allows to relate the sign problem to an oscillating complex phase and not to an undetermined sign, as can be seen later. After Hubbard transformation the exponentials $e^{\delta/2\mathcal{K}}$, from equation (1.25), and the remaining operator exponential $e^{-i\Phi_t\cdot q}$, from equation (1.26), can be recombined by introducing another $O\left(\delta^3\right)$ error, resulting in

$$\int \mathcal{D}\left[\Phi_t\right] e^{-\frac{1}{2\delta U}\sum_{x\in X}\Phi_{t,x}^2} \cdot e^{-\delta\kappa\sum_{\langle x,y\rangle\in X}\left(p_x^\dagger p_y - h_x^\dagger h_y\right) - i\sum_{x\in X}\Phi_x q_x - \delta\mu\sum_{x\in X}q_x} + O\left(\delta^3\right) \tag{1.27}$$

From here we can read off a matrix $A^{xy} = \delta\left(K^{xy} + \mu\delta^{xy}\right) + i\Phi_{t,x}$ such that the exponential is written in the form

$$\int \mathcal{D}\left[\Phi_t\right] e^{-\frac{1}{2\delta U}\sum_{x\in X}\Phi_{t,x}^2} \cdot e^{\sum_{x,y\in X} p_x^\dagger A^{xy} p_y - h_x^\dagger A^{xy} h_y} \tag{1.28}$$

In this form the matrix element (1.23) can be evaluated by using the well known relation between ladder operators, $c \in \{p, h\}$, and coherent states, $|\xi\rangle \in \{|\psi\rangle, |\eta\rangle\}$, [33]

$$\langle\xi|e^{\sum_{x\in X} c_x^\dagger A^{xy} c_y}|\xi'\rangle = e^{\sum_{x\in X} \xi_x^* \left[e^A\right]_{xy} \xi_y'} \tag{1.29}$$

resulting in

$$\langle\psi_{t+1,x}, \eta_{t+1,x}|e^{\delta\mathcal{H}}|\psi_{t,x}, \eta_{t,x}\rangle = \tag{1.30}$$
$$\int \mathcal{D}\left[\Phi_t\right] e^{-\frac{1}{2\delta U}\Phi_t^2} \cdot e^{\sum_{x,y\in X}\psi_{t+1,x}^*\left[e^{\delta K + \delta\mu + i\Phi_t}\right]_{xy}\psi_{ty} + \eta_{t+1,x}^*\left[e^{-\delta K - \delta\mu - i\Phi_t}\right]_{xy}\eta_t} + O\left(\delta^3\right),$$

Inserting this result into the partition function sums up N_t terms of the error leaving a total error of $O(\delta^2)$, and we get

$$\mathcal{Z} = \int \mathcal{D}\left[\Phi\right] \mathcal{D}\left[\psi^*, \psi, \eta^*, \eta\right] \exp\left\{-\frac{1}{2\delta U} \sum_{t,x\in\Lambda} \Phi_{tx}^2\right\}$$

$$\times \exp\left\{ \sum_{t,x;t',y\in\Lambda} \psi_{tx}^* \left(\delta_{xy}\delta_{tt'} - \left[e^{\delta K+\delta\mu+\mathrm{i}\,\Phi_t}\right]_{xy}\delta_{t+1,t'}\right)\psi_{t'y} \right.$$

$$\left. + \sum_{t,x;t',y\in\Lambda} \eta_{tx}^* \left(\delta_{xy}\delta_{tt'} - \left[e^{-\delta K-\delta\mu-\mathrm{i}\,\Phi_t}\right]_{xy}\delta_{t+1,t'}\right)\eta_{t'y} \right\}$$

$$+ O\left(\delta^2\right) \tag{1.31}$$

where the diagonal terms $\delta_{tt'}$ come from the exponential factor in (1.21). I use a typical notation [14], that incorporates the anti-periodic boundary condition into the Kronecker-delta such that

$$\delta_{t+1,t'} \equiv \mathcal{B}_{t'}\delta_{(t+1)\%N_t,\ t'\%N_t} \tag{1.32}$$

where the anti-periodic boundary condition, $\mathcal{B}_{t'} = -1$ if $t' = N_t$ and $\mathcal{B}_{t'} = +1$ otherwise as well as $\delta_{N_t,N_t-1} = \delta_{0,N_t-1}$, is implied. From here we can immediately read off the fermion matrix, in exponential discretization,

$$M\left[\Phi\,|\,K,\mu\right]_{t,x;t',y} = \delta_{xy}\delta_{tt'} - \left[e^{\delta K+\delta\mu}\right]_{xy} e^{\mathrm{i}\,\delta\Phi_{tx}}\delta_{t+1,t'}. \tag{1.33}$$

In various publications, for example [32, 25, 19], a variety of other discretizations have been used with advantages and drawbacks detailed in [19]. These can be derived from here by expanding the exponential to first order $N = 1$

$$\left[e^{\pm\delta K\pm\delta\mu}\right]_{xy} = \left[\sum_{n=0}^{N} \frac{1}{n!}(\pm\delta K \pm \delta\mu)^n\right]_{xy} + O\left(\delta^{N+1}\right). \tag{1.34}$$

Computationally interesting is the case $N = 2$ as it does not increase the error found in the partition function after trotterization (1.31). The sparse spatial part of the fermion matrix can reduce computational cost if implemented properly[4]. As the fermion matrix with $+\Phi$ comes from the ψ coherent states associated with the particle degrees of freedom, it is often referred to as the particle fermion matrix; similarly, the fermion matrix with $-\Phi$ comes from the η resulting in the notion of the hole fermion matrix

$$M^p\left[\Phi\right] = M\left[\ \Phi\,|\ K,\ \mu\right] \tag{1.35}$$

$$M^h\left[\Phi\right] = M\left[-\Phi\,|\,-K,-\mu\right]. \tag{1.36}$$

To continue towards the final path integral form, the fermionic degrees need to be integrated out using the Metthews-Salam formula (1.6) resulting in the Hubbard action

$$\mathrm{S}\left[\Phi\right] = \frac{1}{2\delta U}\Phi^2 - \log\det\{M\left[\Phi|K,\mu\right]\} - \log\det\{M\left[-\Phi|-K,-\mu\right]\}. \tag{1.37}$$

And the partition function reduces, up to an overall irrelevant constant, to

$$\mathcal{Z} = \int \mathcal{D}\left[\Phi\right] e^{-\mathrm{S}[\Phi]} + O\left(\delta^2\right). \tag{1.38}$$

---

[4]This has been discussed within the group, in particular with Dr. Evan Berkowitz, we have not found any reference applying this and plan to test this in the future.
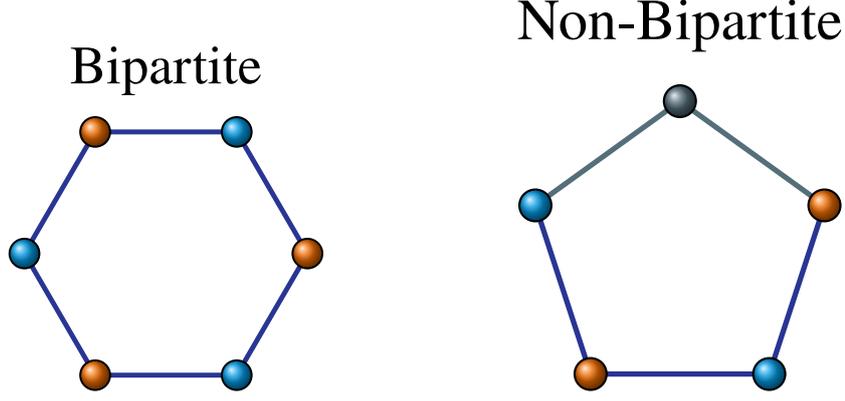
Figure 1.3: Example for bipartite lattice, hexagon on the left, and a non-bipartite lattices, pantagon on the right.

The entire observable is treated in a similar way providing a path integral of the form (1.1).

From the action given in equation (1.37) it is not clear that it is real valued. In fact, it is only in a very specific cases of zero chemical potential on a bipartite lattice. Consider a bipartite lattice, that is the lattice sites can be colored with two colors such that neighbours have different colors, see figure 1.3 for examples. On non-bipartite lattices one typically finds frustrated bonds - colored in gray - which is not the case for bipartite lattices. At zero chemical potential, we can interchange the role of particles and holes leaving the interaction term invariant,

$$(p_x^\dagger p_x - h_x^\dagger h_x)^2 \rightarrow (h_x^\dagger h_x - p_x^\dagger p_x)^2. \tag{1.39}$$

However, the hopping term obtains an additional minus sign,

$$p_x^\dagger K^{xy} p_y - h_x^\dagger K^{xy} h_y \rightarrow -(h_x^\dagger K^{xy} h_y - p_x^\dagger K^{xy} p_y). \tag{1.40}$$

We can further impose $K^{xy} \rightarrow \Sigma^x K^{xy} \Sigma^y$ with a diagonal matrix Σ that has diagonal entries $\Sigma^x = -1$ on one sublattice, defined by the bipartiteness, and $\Sigma^x = +1$ else. Together with the interchange of particles and holes we recover the original Hamiltonian as over the sublattice each term picks up exactly one sign

$$p_x^\dagger K^{xy} p_y - h_x^\dagger K^{xy} h_y \rightarrow p_x^\dagger \Sigma^x K^{xy} \Sigma^y p_y - h_x^\dagger \Sigma^x K^{xy} \Sigma^y h_y = p_x^\dagger K^{xy} p_y - h_x^\dagger K^{xy} h_y \tag{1.41}$$

Thus interchanging particles with holes and transforming $K \rightarrow \Sigma K \Sigma$ leaves the interacting Hamiltonian invariant. Now this implies two important aspects, first, the single particle spectrum at zero chemical potential on a bipartite lattice must be symmetric around zero. Effectively describing half the states with negative energy, this is the Fermi sea. We say the system is half filled because these negative energy states are filled in the ground state. Second, using $\Sigma^2 = \mathbb{1}$ in the fermion matrix determinant allows us to replace $K \rightarrow -K$ in one of the fermion matrices. This together with the fact that $K = K^\dagger$ and complex conjugation sends $\Phi \rightarrow -\Phi$ allows to show that the fermion matrix determinant over both particles and holes is real and positive semi-definite,

$$\det\{M[\Phi|K, \mu = 0] M[-\Phi|-K, \mu = 0] \}$$
$$= \det\{M[\Phi|K, \mu = 0] M[-\Phi| K, \mu = 0] \} \tag{1.42}$$
$$= \det\{M[\Phi|K, \mu = 0] M[\Phi| K, \mu = 0]^\dagger\} \in \mathbb{R}^+.$$

At finite chemical potential or when the lattice is non-bipartite this can not be applied and we encounter a complex valued action.

9

1.2.1 Simple Ways Around the Sign Problem

In the formulation at hand, two simple and semi-analytically determined transformations can mitigate the sign problem, consequently extend the efficiently-computable parameter space. These are used throughout the chapters 2 to 4. The first, and simplest, is what we call tangent plane; the second is the next-to-leading order plane (NLO-plane), being a perturbative correction to the tangent plane. But let us start at the beginning; For a theory with sign problem a method to reduce the sign problem is path deformation onto a manifold $\mathcal{M}$ which has a more optimal phase property

$$\langle O \rangle = \frac{1}{Z} \int_{\mathbb{R}^{|\Lambda|}} \mathcal{D}\left[\Phi\right] e^{-S[\Phi]} O\left[\Phi\right] \tag{1.43}$$

$$= \frac{1}{Z} \int_{\mathcal{M}} \mathcal{D}\left[\varphi\right] e^{-S[\varphi]} O\left[\varphi\right]. \tag{1.44}$$

This works as long as action as well as observable are holomorphic functions in φ. Furthermore, the manifold $\mathcal{M}$ is required to be in the same homology class as $\mathbb{R}^{|\Lambda|}$. These requirements are necessary to apply Cauchys integral theorem [23]. Here the space time lattice is denoted by $\Lambda = [0, N_t - 1] \otimes X$.

One of the most prominent ways of utilizing path deformations are Lefschetz thimbles [21]. Here a complex manifold $\mathcal{M}_\sigma$ is defined as the image under the holomorphic flow equation

$$\frac{\mathrm{d}\Phi_{tx}(\tau_f)}{\mathrm{d}\tau_f} = \pm \left(\frac{\partial S\left[\Phi(\tau_f)\right]}{\partial \Phi_{tx}(\tau_f)} \right)^*, \tag{1.45}$$

where we collect all thimbles that the flow intersects $\mathcal{M} = \cup_\sigma \mathcal{M}_\sigma$. We identify a solution to the holomorphic flow equation by saddle points, that are fixed points, called critical points $\Phi^{\sigma\,\mathrm{cr}}$

$$\left(\frac{\partial S\left[\Phi\right]}{\partial \Phi_{tx}} \right)^* \Bigg|_{\Phi = \Phi^{\sigma\,\mathrm{cr}}} = 0. \tag{1.46}$$

Naively integrating the holomorphic flow equation is analytically infeasible but can be done using for example Runge-Kutta algorithms. This tends to be expensive but for smaller systems greatly reduces the sign problem. There are certain caveats to this, such as choosing the maximal flow time $\tau_f^{\max}$, which will not be discussed here. Instead we focus on a specific case, the classical solution to the flow.

For the Hubbard model described above, constant fields can be seen as classical solutions, in particular the configuration that is zero everywhere, e.g. $\Phi_{tx} = 0 \forall t, x \in \Lambda$ corresponding to the non-interacting, $U = 0$, theory can be used to probe certain properties. Inserting this constrain into the holomorphic flow equation (1.46) yields a purely imaginary solution $\mathrm{i}\,\phi_0$ that can be used to define [7]

$$\Phi_{tx}^{\mathrm{TP}} = \Phi_{tx} + \mathrm{i}\,\phi_0 \tag{1.47}$$

$$\frac{\phi_0}{\delta} = -\frac{U}{N_x} \sum_k \tanh\left\{ \frac{\beta}{2}\left[\epsilon_k + \mu + \frac{\phi_0}{\delta} \right] \right\}. \tag{1.48}$$

This manifold $\mathcal{M}^{\mathrm{TP}}$ is parallel to the real plane and does not impose any computational constraints as it is a simple shift of every configuration into the complex plane. It touches the "main thimble", the thimble contributing most to the path integral, at exactly one point, hence it is called the tangent plane. However, Gängten et.al. went further and included a quantum correction to this saddle point finding an effective action

$$S_{\mathrm{eff}}\left[\Phi^{\mathrm{TP}}\right] = S\left[\Phi^{\mathrm{TP}}\right] - \frac{1}{2}\log\det\left\{ \mathbb{H}_{S\left[\Phi^{\mathrm{TP}}\right]} \right\}, \tag{1.49}$$

with the hessian

$$\mathbb{H}_{S[\Phi^{TP}]} = \frac{\partial}{\partial \Phi_{t'x'}} \frac{\partial}{\partial \Phi_{tx}} S[\Phi]\Big|_{\Phi=\Phi^{TP}}. \tag{1.50}$$

A solution ϕ_1 to the holomorphic flow equation of this effective action, under the constrain of constant fields, can be found numerically in negligible time and serves as another constant shift into the complex plane

$$\Phi^{NLO} = \Phi_{tx} + i\,\phi_1. \tag{1.51}$$

Chapter 4 heavily uses this result.

1.3 Lattice Quantum Chromodynamics

Changing the scale of interest to a single nucleon a vastly different theory is required to be investigated, Quantum Chromodynamics (QCD). The nucleon is governed primarily by the strong interaction, whose degrees of freedom are the fundamental quarks – that are the up (u), down (d), strange (s), charm (c), top (t), and bottom (b) quarks – and the gluons. QCD is described in the continuum by the action

$$S_{QCD}\left[A, \overline{\Psi}, \Psi\right] = \sum_{f=\{u,d,\cdots\}} \int d^4x\, \overline{\Psi}^f(x) \left[\gamma^\mu D_\mu + m^f\right] \Psi^f(x)$$
$$+ \frac{1}{2g^2} \int d^4x\, \mathrm{tr}\left\{F^{\mu\nu}(x) F_{\mu\nu}(x)\right\} \tag{1.52}$$

where the Dirac spinors $\Psi^f(x)^\alpha_a$ come with an index f indicating the quark flavor, a color index a contracted with the gauge field $A \in \mathfrak{su}(3)$, and a Dirac index α contracted with the gamma matrices γ^μ. The latter are solutions of the 4-dimensional Clifford algebra,

$$\{\gamma^\mu, \gamma^\nu\} = 2\eta^{\mu,\nu} \tag{1.53}$$

where we use the Minkowski metric $\eta^{\mu\nu} = \mathrm{diag}(1, -1, -1, -1)$. Furthermore, the fermionic part of the action makes use of the covariant derivative

$$D_\mu = \partial_\mu + i A_\mu(x). \tag{1.54}$$

The field strength tensor used to form the pure gauge part of the action is represented by

$$F_{\mu\nu}(x) = \partial_\mu A_\nu(x) - \partial_\nu A_\mu(x) \tag{1.55}$$

This action can be divided into two parts, the fermion action $S_{fermion}\left[A, \overline{\Psi}, \Psi\right]$, and the gauge action $S_{gauge}[A]$, respectively. These are formulated from the beginning as a quantum field theory, so observables are expressed in terms of the path integral

$$\langle O \rangle = \frac{1}{Z} \int \mathcal{D}\left[A, \overline{\Psi}, \Psi\right] e^{-iS\left[A, \overline{\Psi}, \Psi\right]} O\left[A, \overline{\Psi}, \Psi\right]. \tag{1.56}$$

There is one crucial difference to equation (1.1), that is the Boltzmann weight comes as a phase. This would naturally pose a sign problem. However, for many interesting observables, in particular for hadron spectroscopy and structures, it is sufficient to analytically continue into the Euclidean space effectively mapping the path integral formulation to statistical physics. This analytic continuation is called Wick-rotation and sends

$$t \rightarrow \tau = i x_0, \tag{1.57}$$

mapping the line element in Minkowski space to one in a Euclidean space

$$ds^2 = dx_0^2 - dx_1^2 - dx_2^2 - dx_3^2 \longrightarrow (i\, dx_0)^2 - dx_1^2 - dx_2^2 - dx_3^2 \tag{1.58}$$

where the metric tensor changes from $\eta^{\mu\nu} \rightarrow \delta^{\mu\nu}$. The additional imaginary unit appears in the integral of the QCD action

$$\int d^4x = \int dt\, dx\, dy\, dz \longrightarrow i \int d\tau\, dx\, dy\, dz \equiv i \int d^4x, \tag{1.59}$$

canceling the one in the exponent of the Boltzmann weight.

In many situations lattice QCD has an inherent sign problem, for examples at finite baryon chemical potential or with a Θ-term. However, when these are not considered the sign problem

stemming from fermions can be treated explicitly. To understand this, consider the Wilson fermion matrix, a discretization that will be used in this thesis,

$$\text{M}^f_{\text{Wilson}}[U]_{xy} = \left(m^f + 4/a\right) - \frac{1}{2a}\sum_{\mu=0}^{4}(\mathbb{1} - \gamma^\mu)U_\mu(x)\delta_{x+\hat{\mu},y} + (\mathbb{1} + \gamma^\mu)U_\mu^\dagger(x - \hat{\mu})\delta_{x-\hat{\mu},y}. \quad (1.60)$$

Fermion matrices of this form obey γ^5-hermiticity, $\gamma^5 \text{M} \gamma^5 = \text{M}^\dagger$. Assuming two degenerate flavors, that is two quarks with the same mass, allows to make the fermion determinant positive definite,

$$\det\{\text{M}\} = \det\{\gamma^5 \text{M} \gamma^5\} = \det\{\text{M}^\dagger\}$$
$$\Rightarrow \det\{\text{MM}\} = \det\{\text{MM}^\dagger\} \in \mathbb{R}^+ \quad (1.61)$$

The degenerate flavors can be justified by considering that the light quarks, up and down, have approximately the same mass, $m^u_{\overline{\text{MS}}} = 2.16^{+0.49}_{-0.26}\text{MeV} \approx 4.67^{+0.48}_{-0.17}\text{MeV} = m^d_{\overline{\text{MS}}}$, especially when compared to the strange mass $m^s_{\overline{\text{MS}}} = 93.4^{+8.6}_{-3.4}\text{MeV}$ [34] which modern simulations typically include. For the strange quark, the fermion determinant is usually squared and square-rooted to make it positive definite, yielding a real and positive fermion action [35]. These operations are implemented using an RHMC, see [14] for more details.

To complete the discussion about discretization, the gauge action can be discretized using the Wilson gauge action [14],

$$S_{\text{Wilson-Gauge}}[U] = \frac{2}{g^2}\sum_{x\in\Lambda}\sum_{\substack{\mu,\nu\in[0,3]\\ \mu<\nu}}\text{Re}\{\text{Tr}\{\mathbb{1} - U_{\mu\nu}(x)\}\}, \quad (1.62)$$

where we use the plaquette, $U_{\mu\nu}(x) = U_\mu(x)U_\nu(x+\hat{\mu})U_\mu^\dagger(x+\hat{\nu})U_\mu^\dagger(x)$. A multitude of discretizations for the gauge and fermion action exist, targeting various objectives, including enhanced continuum limits, computational tractability, and more. A comprehensive overview of potential optimisation can be found in chapters 9 and 10 of [14]. Bringing all this together, observables can be expressed on the lattice path integral via

$$\langle O\rangle = \int \mathcal{D}[U]\, e^{-S_{\text{Wilson-Gauge}}[U] - \log\det\{\text{M}^{ud}[U]\text{M}^{ud\dagger}[U]\} - \log\det\{M^s[U]\}} O[U] \quad (1.63)$$

Similar to the Hubbard model, one can apply variants of HMC algorithms to generate a Markov chain of these configurations and measure observables by means of (1.4). It should be noted that the computational demand is considerably higher than that of the Hubbard model! Over the past decades, significant efforts have been made to optimize the algorithms and identify improved discretization. However, only with the advent of supercomputers has it been possible to achieve calculations to great precision.

1.4 Parton Distribution Functions

In 1969, Richard Feynman proposed a model that describes the constituents of hadrons, today known as quarks and gluons. The constituents are typically called partons. This description was formulated as a bag of free quarks that make up the hadron of interest, for example a proton consisting out of three quarks, two up quarks and one down quark. It became clear, that this model was too simple to describe the scattering processes, and later developments then led to QCD. Within QCD, the quantum numbers of hadrons are not only determined by the valence quarks, for the proton the two up quarks and one down quark. Furthermore, these valence quarks are not free but rather bound by the strong interaction carried through gluons. QCD allows for the spontaneous creation of particle anti-particle and successive rapid annihilation at high enough energies, which lead to the

understanding that the hadron consists of the valence quarks and sea quarks, all held together by strong force transmitted via gluons. All these can alter the quantum numbers. In short, it is a very vibrant state with an extremely rich internal structure. Note that today all these constituents, valence quarks, sea quarks as well as the gluons, are still called partons following Feynmans convention. However, the calculation of the inner structure is still an outstanding challenge for modern particle physics, and many theoretical and phenomenological studies as well as experiments are being conducted to gain a glimpse into these systems. All of this work is a great example of how theoretical developments go hand in hand with experimental work, as related measurable quantities can be compared, and thus provide a valuable testing ground for the methods and ideas of particle physics.

The partonic structure of a nucleon is best probed in processes such as Deep Inelastic Scattering (DIS) of leptons – electrons, muons, tauons and their corresponding neutrinos – off hadrons [36, 37]. Here the lepton is assumed to transfer a four-momentum Q to the nucleon in the collision. This generates a resolution of the order of $\hbar/|Q|$ where higher momentum transfers allow for a higher resolution of the inner structure. Of course, all observables related to the parton structure must depend on the momentum transfer, which is typically expressed through the squared momentum transfer

$$Q^2 = (P_f - P_i)^2, \tag{1.64}$$

which makes use of an initial hadron momentum P_i, before the collision, and a final hadron momentum P_f, after the collision, to equivalently express the momentum transfer. However, when studying relativistic quantum systems one must be careful when talking about resolution of the inner structure. Heisenberg's indeterminacy principle prohibits simultaneous knowledge of position and conjugate momentum. A natural resolution barrier is provided by the Compton wavelength; for a hadron H with mass m_H, a system cannot be probed with a higher precision than $O(1/m_H)$ [38]. The fundamental limitation is that for probes with higher energy, particles with mass of that order are potentially produced. In this situation, the resulting measured inner structure is influenced by the interaction with these particles concealing the true inner structure. A way to get an understanding of the dynamics in quantum mechanical system was proposed by Wigner in 1932 [39] through functions today known as Wigner distributions [38]. Investigation on the parton structure of hadrons draw on these ideas in an attempt to find a generalized description.

One particular structure discussed in this thesis, see chapter 5, is the momentum distribution of certain partons of the nucleon – the so-called parton distribution function (PDF). In particular, we focus on the quark content and neglect the gluon PDFs. Considering a fast moving nucleon, where all partons move roughly along the direction of the nucleon. Consequently, the momentum fraction, along the direction of movement, carried by one parton describes a relevant variable in this limit. This is the longitudinal momentum fraction, usually denoted by $x \in [0, 1]$. For a given process with momentum transfer squared Q^2 the PDF is a function $f^q(x, Q^2) \equiv f(x)$, for a given parton q. We differentiate between three PDFs; the unpolarized $q(x)$ PDF, the polarized PDF $\Delta q(x)$, and the transversity PDF $\delta q(x)$. Notice that the dependence on Q^2 is typically suppressed in this notation.

Parton distribution functions can be defined by considering bilocal quark operators [36]

$$O_q^X(x) = \frac{1}{2\pi} \int d\lambda\, e^{i\lambda x} \overline{q}(0) \Gamma^X \mathcal{U} q(\lambda n). \tag{1.65}$$

where $q = u, d, \cdots$ denotes the quark parton of interest, and $n = (n^\mu)$ is a light cone vector, chosen such that

$$n \cdot \frac{P_f + P_i}{2} = 1. \tag{1.66}$$

with a typical choice of $n = (1, 0, 0, -1)$ [38]. The gauge link $\mathcal{U}$ connects to the quark and extends along the light cone to make the operator gauge invariant. From this we can define the PDF by taking the matrix element for a process $H(P_i) \rightarrow H(P_f)$ [38]

$$f^q(x) = \frac{1}{2p^+} \left\langle H(P_i) \middle| O_q^X(x) \middle| H(P_f) \right\rangle. \tag{1.67}$$

In particular we can access the polarized PDF, the unpolarized PDF and the transversity PDF by inserting the vector (V), axial (A), or tensor (T) Dirac structure, respectively,

$$\left(\Gamma^V\right)^\mu = \gamma^\mu \qquad \& \qquad \left(\Gamma^A\right)^\mu = \gamma^\mu \gamma_5 \qquad \& \qquad \left(\Gamma^T\right)^{\mu\nu} = \gamma^{\mu\nu}. \tag{1.68}$$

Unfortunately, lattice QCD is calculated in Euclidean space, which makes the direct access to operators of this form difficult. There have been several developments in recent years to extract parton distribution functions from the lattice. In particular, the quasi-PDF methods [40, 41], allows to calculated the x dependence more directly.

However, in this thesis I discuss the more traditional approach to calculate moments of PDF [42, 36, 37, 43, 44, 36, 45, 46, 47, 48, 49]. Instead of calculating the entire x-dependence only moments of the PDF are calculated. Since these can be related to matrix elements of local leading twist operators, significantly less momentum is required compared to the quasi-PDF methods. Using the operator product expansion it is possible to relate the nth Mellin moment,

$$\langle x_q \rangle^n = \int dx \, x^{n-1} f^q(x) \tag{1.69}$$

to local leading twist operators with n-derivative operators

$$O^X_{\alpha,\mu_1,\cdots,\mu_n} = \overline{q}\Gamma^X_{\{\alpha}\overleftrightarrow{D}_{\mu_1}\cdots\overleftrightarrow{D}_{\mu_n\}}q. \tag{1.70}$$

where we make use of the left and right acting Dirac operator $\overleftrightarrow{D} = 1/2(\overrightarrow{D} - \overleftarrow{D})$, discretized on the lattice. The relation for the nth moment and the matrix element is linear, for the nucleon we have

$$\langle N(p)|O^X_{\{\alpha\mu_1,\cdots,\mu_n\}}|N(p)\rangle = \langle x \rangle \, \overline{q}_{N(p)}\Gamma^X_{\{\alpha} \, i\, p_{\mu_1}\cdots i\, p_{\mu_n\}}q_{N(p)}. \tag{1.71}$$

Notice that the indices $\alpha, \mu_1, \cdots, \mu_n$ are symmetrized and traces are removed, indicated by $\{\cdots\}$. The kinematic factor can be determined analytically. However, it is highly non trivial to calculate the left-hand-side and we use lattice QCD to determine it.

To calculate matrix elements of local leading twist operators (1.70), one first generates an ensemble of gauge configuration with the favourite lattice discretization. Then two-point correlators,

$$C_{2\text{pt}}(\tau) = \int d^3 y \, e^{-i\vec{p}\vec{y}} \, \text{Tr}\left\{\Gamma_{\text{pol}}\left\langle \chi(\vec{y},\tau)\overline{\chi}(\vec{0},0)\right\rangle\right\}, \tag{1.72}$$

and three-point correlators,

$$C_{3\text{pt}}^{O^X}(T,\tau) \quad = \quad \int d^3 y \, d^3 z \left[e^{-i\vec{p}'\vec{y}}e^{i(\vec{p}'-\vec{p})\vec{z}} \, \text{Tr}\left\{\Gamma_{\text{pol}}\left\langle \chi(\vec{y},T) \, O^X(\vec{y},\tau)\overline{\chi}(\vec{0},0)\right\rangle\right\}\right], \tag{1.73}$$

have to be calculated on each gauge configuration. A graphical representation of this can be found in chapter 5 in figure B.1. Notice, these equation use certain interpolating operators χ that need to be constructed for the hadron at hand. Further, definitions of the components used here can be found in chapter 5 section B.2. Now the matrix element can be extracted by taking the ratio of these correlators in the limit of large times and source sink separations

$$\langle H(P')|O|H(P)\rangle = \lim_{T-\tau,T\to\infty} \frac{C_{3\text{pt}}^{O^X}(T,\tau)}{C_{2\text{pt}}(T)}. \tag{1.74}$$

Since infinite lattices, in particular time extents, are practically impossible, a detailed analysis of the excited state contamination must be performed. This procedure has been shown to work well for the first few methods, which is discussed in more detail in chapter 5. Unfortunately, high moments become increasingly difficult to calculate because they become more noisy, requiring more

configurations, and have more excited state contamination, requiring larger source sink separations making the estimates even more noisy. In addition, the renormalization of more derivative operators becomes increasingly difficult as mixing occurs.

All this is done to extract the momentum distribution of the partons inside a hadron. However, spatial distributions, charges and currents, can be obtained in a similar manner. They relate to a 0-derivative matrix elements, see for example [50]. All this, and more, can be summarized in a more extensive picture that, currently, provides the most complete description of the inner structure of Hadrons. The so-called generalized parton distribution functions (GPDs) allow to access the dynamical correlation of spatial and momentum distributions. Consider the generalized parton distribution function, typically denoted by $H, \tilde{H}, E$, and $\tilde{E}$ [51, 38, 36], as well as the transverse GPDs $H_T, \tilde{H}_T$ [52, 53]. These depend, besides the renormalization scale, on the partons longitudinal momentum fraction x, the parton longitudinal momentum transfer ξ, and the momentum transfer to the target t. It should be emphasized that the dependence on x and ξ only encode the longitudinal structure while the dependence on t allows to access information about the transverse momentum. To argue that generalized parton distributions encode information about spatial and momentum distributions, three important relations can be considered [53, 51].

First, integrating out the momentum fraction, results in the form factor limit

$$\int \mathrm{d}x \, H(x, \xi, t) = F(t). \tag{1.75}$$

The Dirac form factor is governed by H, the Pauli form factor by E, the axial vector form factor by $\tilde{H}$, and the pseudoscalar form factor by $\tilde{E}$.

Second taking the forward limit, $\xi, t \rightarrow 0$, result in the parton distributions

$$\lim_{\xi, t \rightarrow 0} H(x, \xi, t) = q(x) \tag{1.76}$$

$$\lim_{\xi, t \rightarrow 0} \tilde{H}(x, \xi, t) = \Delta q(x) \tag{1.77}$$

$$\lim_{\xi, t \rightarrow 0} H_T(x, \xi, t) = \delta q(x) \tag{1.78}$$

which are detailed above.

Last, there is Ji's sum rule making a connection to the total angular momentum J_q,

$$\lim_{t \rightarrow 0} \frac{1}{2} \int \mathrm{d}x \, x (H + E)(x, \xi, t) = J_q. \tag{1.79}$$

The first two rules demonstrate clearly, how the notion of a generalized parton distribution function provides a framework that combines the information about the spatial and momentum distributions similar to the Wigner distributions. In fact, generalizing the Wigner function to a quantum field theory, allows to use it as a generating functional for GPDs [38].

Chapter 2

Mitigating the Hubbard Sign Problem with Complex-Valued Neural Networks

This chapter is based on [1]:

In order for the Boltzmann distribution, compare equation (1.3), to be interpret as a proper probability distribution the action must be real-valued. This is true for some relevant quantum field theories. Consequently, the widely applied Markov chain methods can be used to generate random variables distributed according to the Boltzmann distribution, which, subsequently are used to compute the observables, compare equation (1.1) using the Monte Carlo approach efficiently. Unfortunately, for many interesting physical systems the discretization yields a generally complex-valued action. Examples for such cases are lattice QCD at finite baryon-chemical potential or with a Θ-term. Furthermore, systems of strongly correlated electrons described by a Hubbard model away from half filling typically form a complex action. Additionally, when the Hubbard model is used to describe systems with frustrated bonds a sign problem is found even at zero chemical potential. Finally, any simulation with a real valued action that is not Wick rotated into euclidean time suffers from a sign problem as the Boltzmann weight is $\sim e^{-iS}$. In all these cases, and many more, the Boltzmann distribution can only be understood as a pseudo-distribution and Monte Carlo become inapplicable. This phenomenon is known as the oscillating complex phase or sign problem.

Phase-reweighting allows observables to be re-expressed in a way that enables Monte Carlo methods to be applied without change, compare the result in equation 1.8. The basic idea is to put the complex phase, coming from the imaginary part of the action, into the observable and sampling the Markov chain according to the real part of the action. The equation (1.8) is exact and does not depend on the algorithm that generates the configurations. However, it comes at the cost of a widely oscillating integrand, $e^{-i \text{Im}\{S\}}$ which tends to be infeasible to compute. It is anticipated that the sign problem is NP-complete [20], thus my objective is to ameliorate its effects to extend the computable parameter space. Many formal and practical developments have been investigated for various theories. In particular, path deformations have acquired great interest [21, 54, 55, 56, 57, 58, 59]. The use of neural networks to identify path deformations has been explored for various theories but typically suffers from bad volume scaling [60, 61].

This chapter examines the potential of complex-valued neural networks to mitigate the sign problem of the Hubbard model. The Hubbard Hamiltonian,

$$\mathcal{H}[\kappa, U, \mu] = -\kappa \sum_{\langle x,y \rangle \in X} \left(p_x^\dagger p_y - h_x^\dagger h_y \right) + \frac{U}{2} \sum_{x \in X} q_x^2 + \mu \sum_{x \in X} q_x, \qquad (2.1)$$

is treated with a trotterization to obtain an action on continues degrees of freedom, compare equation 1.37. Notice, that the sign convention compared to equation 1.11 is different assuming that the number difference between particles and holes, $q_x = p_x^\dagger p_x - h_x^\dagger h_x$, corresponds to electrons having a positive charge. This convention is the typically chosen for the Hubbard model at zero chemical potential as the sign does not affect the interaction. The convention can be translated by sending $\mu \to -\mu$. Once the action is set up, the simulation using HMC obeys a sign problem at finite chemical potential μ. Complex-valued neural networks, build upon affine coupling layers that change half the input based on a coupling c that relates the input partitions Φ_A, Φ_B,

$$(\Phi_A, \Phi_B) \mapsto \begin{pmatrix} c(\Phi_A, \Phi_B) \\ \Phi_B \end{pmatrix}, \qquad (2.2)$$

are used to identify the transformation towards Lefschetz thimbles. Affine coupling layers traditionally allow for the computation of the Jacobian determinant in $O(V)$ time, which is a significant improvement over a naive determinant scaling as $O(V^3)$. For complex-valued networks a subtlty appears that is discussed in this paper with further information at the end of this chapter in section 2.6. We demonstrate the efficacy of this method by successfully applying it to systems of different size. Each is intractable with a naive HMC implementation due to its sever sign problem. With simulations on the tangent plane, a classical solution to the holomorphic flow equation that is parallel to the – original – real plane, the smaller systems can be resolved. However, as the sign problem becomes worse with larger number of ions this is not enough anymore. Consequently, the complex valued neural network discussed in this chapter is used to transform to a less simple

manifold on which the sign problem becomes further reduced. The boundary is further pushed to a 8- and 18-site system, actually showing some graphene like system, on which the tangent plane fails to resolve single particle correlators. A study of physical properties is beyond this chapter as only a single ensemble is considered. We discuss more physics in chapter 4 on the molecule $C_{20}H_{12}$ perylene which is build out of 20 ions.

The idea to use affine coupling layers was proposed by Wynen et.al. [61] and then executed by me. Many helpful discussions with the co-authors guiding the process.

2.1 Introduction

The computational sign problem encumbers successful importance sampling from complex-valued distributions with Markov Chain Monte Carlo algorithms such as Hybrid Monte Carlo (HMC). Sampling from the configuration space of a wide variety of interesting physical systems suffers such a difficulty, ranging from lattice QCD at finite baryon chemical potential and doped condensed matter systems in equilibrium to the real-time evolution of quantum systems.

By deforming the real manifold of integration for a path integral of interest into complex variables, one may reduce the sign problem substantially [62, 23, 63, 64]. In the last few years, new formal developments have inspired investigation into leveraging Lefschetz thimbles [21, 54, 55, 56, 57, 58, 59]—high-dimensional analogues of contours of steepest descent which can be located by holomorphic flow. In [54], for example, fluctuations about the saddle point of each thimble were sampled to simulate the 0+1 dimensional Thirring model, something much akin to the method of steepest descent. In practice the determination of the precise location of each thimble's saddle point, or critical point, as well as the relevant sampling 'direction' about these points, is numerically costly and prohibitive. An alternative method is to train neural networks to learn the map from some starting manifold to any beneficial manifold, including one that approximates the thimbles that contribute to the integral [60, 65, 61].

In our previous work [61] we were limited by the computational cost of incorporating the Jacobian determinant of this map into our importance sampling. In this paper we leverage complex-valued neural networks built of affine coupling layers to reduce the scaling of the Jacobian determinant cost. We focus on the Hubbard model on a honeycomb lattice away from half-filling and compare methods by computing single-particle correlation functions.

This paper is organized in the following way. In Section A.2, a brief recap of the Hubbard model and basic notation is given. After that, some prior methods to alleviate or remove the sign problem and usage within HMC are discussed. In Section 2.3, we describe the new neural network architecture. In Section 2.4, we show a numerical test of the network on three systems where we can exactly diagonalize the Hamiltonian, and one larger system beyond our ability to exactly diagonalize.

2.2 Formalism

The Hubbard model [66, 30, 31] describes a fixed spatial lattice X on which particles can move and interact. In the particle-hole basis it is described by Hamiltonian

$$\mathcal{H}[K,V,\mu] = -\sum_{x,y\in X}\left(p_x^\dagger K^{xy}p_y - h_x^\dagger K^{xy}h_y\right) + \frac{1}{2}\sum_{x,y\in X}\rho_x V^{xy}\rho_y + \mu\sum_{x\in X}\rho_x, \quad (2.3)$$

where the amplitudes in K encode the hopping of fermionic particles p and holes h, the potential V describes the interactions between charges

$$\rho_x = p_x^\dagger p_x - h_x^\dagger h_x \quad (2.4)$$

19

(a) 2 Sites   (b) 4 Sites   (c) 8 Sites

(d) 18 Sites (boundary suppressed)
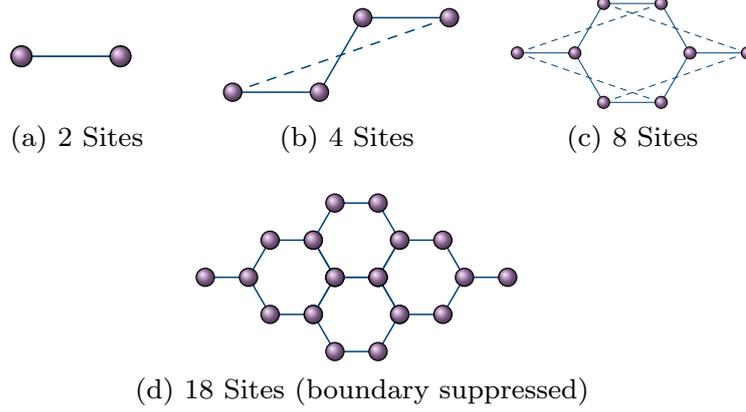
Figure 2.1: Graphical representation of the arrangement of ions considered in the numerical investigation. Each node corresponds to an ion while each edge indicates an allowed particle/hole hopping. The dashed lines represent the periodic boundary.

and the chemical potential μ incentivizes charge. By adjusting K and V this model can describe a wide variety of physical systems. We restrict our attention to the case where K encodes a honeycomb structure with nearest-neighbor hopping and the interaction V is local,

$$K = \kappa\delta_{\langle xy \rangle} \quad \& \quad V = U\delta_{xy}; \tag{2.5}$$

the bipartiteness of the honeycomb permits a signed sublattice transformation that flips the sign of the hopping of holes. As we are focusing on algorithmic issues we focus on only the four systems displayed in Figure 2.1. These—the 2, 4, 8, and 18 site models—are examples of the honeycomb lattice with periodic boundary conditions.

Our aim is to compute observables O according to the thermal trace

$$\langle O \rangle = \frac{1}{\mathcal{Z}} \operatorname{tr}\left\{\left[O e^{-\beta H}\right]\right\}. \tag{2.6}$$

where the partition function $\mathcal{Z}$ is the trace without the observable and β is the inverse temperature, the euclidean time extent. Trotterizing into N_t timeslices, inserting Grassmannian resolutions of the identity, and linearizing the interaction via the Hubbard-Stratonovich transformation [29] leads to the action

$$S\left[\Phi \mid K, V, \mu\right] = \frac{1}{2} \sum_t \sum_{x,y \in X} \Phi_{tx}(\delta V^{-1})^{xy}\Phi_{ty} - \log\det\{M\left[\Phi \mid K, \mu\right] \cdot M\left[-\Phi \mid -K, -\mu\right]\}, \tag{2.7}$$

where $\Phi \in \mathbb{R}^{|\Lambda|}$ is an auxiliary field on the spacetime lattice $\Lambda = [0, N_t - 1] \otimes X$ and $\delta = \beta/N_t$. We use the exponential discretization [19] for the fermion matrices

$$M\left[\Phi \mid K, \mu\right]_{x't';xt} = \delta_{x'x}\delta_{t't} - \left(e^{\delta(K+\mu)}\right)_{x'x} e^{+i\Phi_{xt}} \mathcal{B}_{t'}\delta_{t'(t+1)} \tag{2.8}$$

where $\mathcal{B}$ encodes the antiperiodic boundary conditions in time. On a bipartite lattice we may replace the $-K$ in the holes' fermion matrix with $+K$; then when $\mu = 0$ the determinant may be made manifestly positive-semidefinite. When μ is finite S is complex; a great deal of recent effort has been made in the computational physics community to understand this case [67, 68, 69, 70].

The transformation of the thermal average (2.6) leads to the path integral

$$\langle O \rangle = \frac{1}{\mathcal{Z}} \int \mathcal{D}\left[\Phi\right] e^{-\beta S[\Phi]} O\left[\Phi\right] \equiv \int \mathcal{D}\left[\Phi\right] p_S\left[\Phi\right] O\left[\Phi\right] \tag{2.9}$$

20

where the partition function $\mathcal{Z}$ is the integral without the observable O. When the action is real importance-sampling methods draw N_{cfg} configurations according to the Boltzmann distribution

$$p_S[\Phi] = \frac{1}{\mathcal{Z}} e^{-S[\Phi]} \tag{2.10}$$

and estimate observables (A.5) by an unweighted average. Any practical calculation samples only finitely many configurations N_{cfg} and the resulting statistical uncertainties scale like $N_{cfg}^{-1/2}$ as long as the configurations are independent.

At finite μ a complex-valued action yields an oscillating integrand and p_S (2.10) can no longer be interpreted as a standard probability density, rendering a straightforward application of importance sampling impossible.

To recover an importance-sampling algorithm we can separate the real and imaginary parts of the action $S = \text{Re}\{S\} + i\,\text{Im}\{S\}$ and rewrite the partition function

$$\mathcal{Z} = \int \mathcal{D}[\Phi]\, e^{-S} = \int \mathcal{D}[\Phi]\, e^{-\text{Re}\{S\}} e^{-i\,\text{Im}\{S\}} \propto \left\langle e^{-i\,\text{Im}\{S\}} \right\rangle_{\text{Re}\{S\}} \equiv \Sigma \tag{2.11}$$

where the expectation value is with respect to the real part of the action and we call Σ the statistical power. So, by sampling according to $p_{\text{Re}\{S\}}$ we can estimate

$$\langle O \rangle = \frac{\left\langle e^{-i\,\text{Im}\{S\}} O \right\rangle_{\text{Re}\{S\}}}{\left\langle e^{-i\,\text{Im}\{S\}} \right\rangle_{\text{Re}\{S\}}} = \frac{1}{\Sigma} \left\langle e^{-i\,\text{Im}\{S\}} O \right\rangle_{\text{Re}\{S\}}. \tag{2.12}$$

When the statistical power Σ (3.5) cannot be reliably distinguished from zero the sign problem is too strong and the whole procedure fails [20, 61, 54, 71]. [20] showed that the effective number of configurations

$$N_{cfg}^{eff} = |\Sigma|^2 \cdot N_{cfg} \tag{2.13}$$

controls the scaling of statistical errors $\sim \left(N_{cfg}^{eff}\right)^{-1/2}$.

It is widely expected that the statistical power shrinks exponentially with spacetime volume $\beta|X|$. Because the power is the ratio of the full and phase-quenched partition functions it should be exponential in a difference of free energies, which is extensive in the spacetime volume [72]. For small nonbipartite examples we have previously confirmed the exponential dependence on β [61].

A promising alternative to simple reweighting is to complexify the domain of integration and transform $\phi \in \mathcal{M}_{\mathbb{R}} = \mathbb{R}^{|\Lambda|}$ to a manifold $\Phi \in \mathcal{M} \subset \mathbb{C}^{|\Lambda|}$. As long as $\mathcal{M}$ is in the same homology class, the analogue of the Cauchy integral theorem ensures that the partition function is unchanged [23],

$$\mathcal{Z} = \int_{\mathcal{M}} \mathcal{D}[\Phi]\, e^{-S[\Phi]}. \tag{2.14}$$

Parametrizing the manifold $\mathcal{M}$ by the real fields induces a Jacobian determinant, yielding [54]

$$\mathcal{Z} = \int_{\mathcal{M}_{\mathbb{R}}} \mathcal{D}[\phi]\, e^{-S[\Phi(\phi)] + \log\det\{J[\Phi(\phi)]\}} \tag{2.15}$$

and observables are computed on the manifold $O[\Phi(\phi)]$.

A judicious choice of the manifold $\mathcal{M}$ can diminish or completely remove the sign problem [73, 23]. Even when sampling according to $p_{\text{Re}\,S^{eff}}$ with an imperfect manifold with a complex effective action

$$S^{eff}[\phi] = S[\Phi(\phi)] - \log\det J[\Phi(\phi)] \qquad J_{ij} = \frac{\partial \Phi_i}{\partial \phi_j} \tag{2.16}$$

21

if the statistical power Σ (3.5) is sufficiently improved we can reweight (A.6) with the imaginary part Im S^{eff}.

There are many strategies for picking target manifolds [74]. One choice is to try to approximate the Lefschetz thimbles – high-dimensional manifolds analogous to contours of steepest descent, which have constant imaginary action and therefore have a much-reduced sign problem [73]. Each thimble contains a critical point Φ_{crit} that satisfies

$$\left. \frac{\partial S\,[\Phi]}{\partial \Phi} \right|_{\Phi=\Phi_{\text{crit}}} = 0 \tag{2.17}$$

and is therefore a fixed point of the holomorphic flow

$$\frac{d\Phi(\tau)}{d\tau} = \left(\frac{\partial S\,[\Phi(\tau)]}{\partial \Phi(\tau)} \right)^* \tag{2.18}$$

as a function of the fictitious flow time τ and initial condition $\Phi(0) = \phi$. We can trace trajectories under the flow using the integrator

$$\mathfrak{I}_\tau^\pm\,[\phi] \equiv \int\limits_0^{\pm\tau} \left(\frac{\partial S\,[\Phi(\tau_f)]}{\partial \Phi(\tau_f)} \right)^* \mathrm{d}\tau_f \,. \tag{2.19}$$

A thimble is the set of complexified configurations that flow to a critical point under downward flow $\mathfrak{I}_\infty^-$.

There may be many thimbles in $\mathbb{C}^{|\Lambda|}$ and only some might contribute. The upward flow $\mathfrak{I}_\infty^+$ discovers these thimbles automatically. After enough flow time τ the integrator $\mathfrak{I}_\tau^+$ drives any $\Phi(0)$ to either a place on a thimble or to neverland – any place where thimbles of different imaginary action meet and therefore must have zero weight. When Φ starts on a valid integration manifold its image under $\mathfrak{I}_\infty^+$ is on a thimble that contributes to the integral or is in neverland. For an approachable discussion and proof, see the recent review [23].

Therefore, we can try to evaluate the path integral (2.15) on the manifold given by $\Phi\,(\phi) = \mathfrak{I}_\infty^+\,[\phi]$ for each ϕ on any valid starting manifold $\mathcal{M}_0$, such as $\mathcal{M}_\mathbb{R}$. Though this seems to make sign problem free simulations possible, two issues remain. While integrating the flow (2.19) is cheap, performing molecular dynamics integration on the thimbles at first glance involves the costly computation of the Hessian $\partial_{\Phi_i} \partial_{\Phi_j} S\,[\Phi]$ due to the appearance of the Jacobian determinant of the flow in the effective action (A.8), though some ideas for quickly estimating the Jacobian have been proposed [75] and recent work [76] shows how to accelerate this for sparse, local (bosonic) actions. The Jacobian determinant has to be evaluated at any accept-reject step with computational cost scaling like $|\Lambda|^3$.

Second, because thimbles only touch at places of zero weight, algorithms like HMC [77] which use a smooth update of the fields Φ would be encumbered by an ergodicity problem. The severity of this issue is ameliorated in two ways. As any practical integrator $\mathfrak{I}_\tau^+$ necessarily approximates the flow, the resulting integration manifold is only approximately the union of contributing thimbles. Additionally, we do not need to flow for very much time. Both of these mean that the important configurations are smoothly connected, though the imaginary part of the action is not perfectly piecewise constant. In practice, picking a τ is a tradeoff between reducing the computational cost of the flow and an improvement of the statistical power.

The cost of the flow and the associated Jacobian determinant is such that it is beneficial to train a neural network to learn the map $\mathfrak{I}_\tau^+ : \mathcal{M}_0 \to \tilde{\mathcal{M}}$. In the next section we explain our network's architecture.

Of course, understanding neural networks as general function approximators yields an interpretation of any (numerical) integrator as a network, though it is parameter-free and needs no training—its layers, given by some discretization of the flow equations (A.9), are exactly known. Just as we can produce training configurations closer to the thimbles with a more precise integrator, by adding additional layers we may train the network to reproduce the integrated flow more accurately. So,

one expects a trade-off between the nearness to the thimbles (thinking of the number of layers as a proxy) and the effort required to train. The algorithm we describe is exact, even in the case where the network does not offer an acceleration, since the network produces a manifold with the correct homology class regardless of its fidelity to the thimbles.

Because we can integrate on any manifold in the same homology class as $\mathbb{R}^{|\Lambda|}$, it may be beneficial to find simple manifolds that can improve the statistical power without the computational cost of flowing [78, 79]. One such manifold is the tangent (hyper-)plane $\Phi \in \mathcal{M}_T$ [54, 61, 23], a hyperplane parallel to the real manifold offset by a constant imaginary piece so that it intersects the critical-point image of the zero configuration $i\Phi^0_{\text{crit}} = \mathfrak{I}^+_\infty(0)$

$$\Phi(\phi) = \phi + i\Phi^0_{\text{crit}} \tag{2.20}$$

for all $\phi \in \mathcal{M}_\mathbb{R}$. For many smaller systems this transformation already reduces the sign problem enough that reweighting can be applied. However, in our larger examples the tangent plane gives no appreciable statistical power. Nevertheless, we can reduce the cost and potentially increase the potency of flowing if we start from the tangent plane [61].

One obvious approach to constructing an HMC-like algorithm is to attempt molecular dynamics on the target manifold $\tilde{\mathcal{M}}$ given by $\tilde{\Phi}$; in our case, an approximation of the thimbles. However, remaining on the manifold is not so simple [80, 81, 82, 76].

In contrast, performing HMC on the tangent plane *is* simple – when integrating molecular dynamics trajectories simply neglect the imaginary part of the force. Because the real plane suffers from a severe sign problem in the examples we study, we use this tanget-plane HMC as a benchmark. In the remainder of this paper we refer to it simply as "HMC" unless clarification is needed.

For further improvement we do molecular dynamics on the tangent plane $\mathcal{M}_T$ and perform the Metropolis-Hastings accept/reject step on the target manifold $\tilde{\mathcal{M}}$ according to the effective action (A.8). We track the configuration on both the integration manifold $\mathcal{M}_0$ and its image on the target manifold $\tilde{\mathcal{M}}$ to avoid paying the computational cost of applying or inverting the transformation $\tilde{\Phi}$ more than needed. Assuming the numerical implementation of the map $\tilde{\Phi}$ is invertible, proof that this algorithm has detailed balance is provided in [61]. One can use a reversible integrator or an invertible neural network to satisfy this requirement.

2.3 Machine-Learning Method

To accelerate the transformation to the target manifold $\tilde{\mathcal{M}}$, reducing computational complexity, it is possible to define a neural network trained to approximate the integrator (2.19) $\mathcal{NN} \approx \mathfrak{I}^+_\tau$.

One approach is to learn the imaginary part of any configuration on the target manifold $\tilde{\mathcal{M}}$ given its real part[60, 61]

$$\text{SHIFT} : \mathcal{M}_0 \to \tilde{\mathcal{M}}, \ \Phi \mapsto \Phi + i\mathcal{NN}(\text{Re}\{\Phi\}). \tag{2.21}$$

This ansatz has two advantages. First, the ergodicity issue, induced by potential trapping on individual thimbles, is removed [60]. Second, the network can use the well-established methods of real-valued neural networks. Computational costs due to flowing are reduced as the application of the neural network is much cheaper then any numerical integration. However, a major disadvantage is the computational effort and severe volume scaling of the Jacobian determinant [61].

In this work we use complex-valued neural networks – networks with complex parameters – to instead learn the map from the integration manifold $\mathcal{M}_0$ to the target manifold $\tilde{M}$,

$$\tilde{\Phi} = \mathcal{NN}(\phi) \approx \mathfrak{I}^+_\tau(\phi). \tag{2.22}$$

This approach enjoys a significant advantage over the SHIFT network (A.10): given the right network architecture the Jacobian may be evaluated very quickly. Below we will explain our use of

affine coupling layers to reduce the scaling of the Jacobian determinant from a general cubic scaling down to a linear scaling in the volume $|\Lambda|$.

For a recent overview of complex-valued networks see [83]. Typical automatic differentiation algorithms can be applied to complex-valued neural networks in a similar manner as to real-valued ones [84, 85, 86] by switching the differentiation rule to Wirtinger derivatives [84]

$$
\begin{aligned}
\frac{\partial f(z)}{\partial z} &= \frac{1}{2}\left(\frac{\partial f(z)}{\partial \operatorname{Re} z} - i\frac{\partial f(z)}{\partial \operatorname{Im} z}\right) \\
\frac{\partial f(z)}{\partial z^*} &= \frac{1}{2}\left(\frac{\partial f(z)}{\partial \operatorname{Re} z} + i\frac{\partial f(z)}{\partial \operatorname{Im} z}\right).
\end{aligned}
\tag{2.23}
$$

The Wirtinger derivatives have the advantage that they coincide with complex derivatives for holomorphic functions while also extending to non-holomorphic ones. This generalization is required for two reasons. First, loss functions typically are not holomorphic and are not differentiable in the complex sense. Second, Liouville's theorem, stating that bounded entire functions are constant, reduces the usability of any complex-valued neural network if only holomorphic components can be used. As automatic differentiation is possible through backpropagation using Wirtinger derivatives, these restrictions can be overcome and a neural network $\mathcal{NN}: \mathbb{C}^m \to \mathbb{C}^n$ with complex-valued weights can be defined [83]. We want to emphasize that a non holomorphic network can approximate the thimbles even though their definition is manifestly holomorphic. This can be understood by utilizing the universal approximation theorem [87], and realizing that the change of variable requires an embedding which is at least twice-differentiable in the Wirtinger sense. It is expected that such networks have an improved expressivity compared to real valued networks of twice the size – mimicking the real and imaginary parts – as complex networks do not have to learn complex arithmetic [83].

Special care has to be taken when evaluating the Jacobian induced by the parametrization of $\tilde{M}$.[1] The Jacobian in the effective action (A.8) is defined by the derivative of the transformation according to its real parameters – a derivative in the real sense. When applying a non-holomorphic neural network to parametrize the manifold, the Wirtinger derivatives force us to reexpress the derivative in the real sense by combining the two equations of (2.23) and the transformation on the tangent plane (2.20)

$$
J_{ij} \equiv \frac{\partial \mathcal{NN}(\phi + i\Phi_c^0)_i}{\partial \phi_j} \qquad\qquad = \frac{\partial \mathcal{NN}(\Phi)_i}{\partial \Phi_j} + \frac{\partial \mathcal{NN}(\Phi)_i}{\partial \Phi_j^*}.
\tag{2.24}
$$

To identify an architecture with an efficiently-computable Jacobian determinant, split the network into L constituent layers:

$$
\begin{aligned}
\Phi_0(\phi) &= \phi + i\Phi_c^0 \\
\Phi_{\ell>0}(\phi) &= \mathcal{NN}_\ell(\Phi_{\ell-1}(\phi)) \\
&= \left(\mathcal{NN}_\ell \circ \mathcal{NN}_{\ell-1} \circ \cdots \circ \mathcal{NN}_1\right)(\phi) \\
\tilde{\Phi}(\phi) &= \Phi_L(\phi) = \mathcal{NN}_L(\phi) \equiv \mathcal{NN}(\phi).
\end{aligned}
\tag{2.25}
$$

The Jacobian determinant of the neural network[2] is then given as the product of the Jacobian determinants of each layer

$$
\det\{J\} = \prod_{\ell=1}^{L} \det\{J_{\mathcal{NN}_\ell}\}.
\tag{2.26}
$$

Consequently, we focus on layers with computationally simple Jacobian determinants. Coupling layers

$$
\mathcal{NN}_\ell(\Phi) = \begin{cases} c_\ell\,[\Phi_A,\,\Phi_B] & A_\ell \text{ components} \\ \Phi_B & B_\ell \text{ components} \end{cases}
\tag{2.27}
$$

---

[1] For a more thorough discussion on this result, please refer to 2.6

[2] Note that this requires the input and output dimension of each layer to be equal.
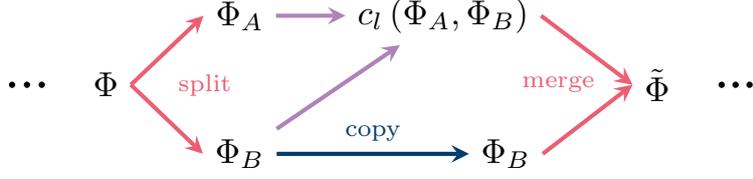
Figure 2.2: Pictorial representation of one coupling layer (A.12). First the input configuration Φ is split into two partitions Φ_A and Φ_B. The corresponding A components are then changed accoring to the prescribed coupling $c_l (\Phi_A, \Phi_B)$ while the B components are untouched. We utilize an affine transformation (A.13) for the coupling c_l. The resulting output vector $\tilde{\Phi}$ is then constructed from the transformed A components and unchanged B components.

fulfill this requirement [88, 89]. Here A and B are layer-specific partitions of the input vector Φ of equal cardinality $\frac{1}{2}|\Lambda|$, and $\Phi_{A,B}$ are the components of the input belonging to the indicated partition. If the coupling layer $c_\ell [\Phi_A, \Phi_B]$ acts elementwise and is holomorphic in the components Φ_A

$$\frac{\partial c_\ell [\Phi_A, \Phi_B]}{\partial \Phi_A^*} = 0 \,, \tag{2.28}$$

the Jacobian determinant of each layer is given by

$$\det\{J_{\mathcal{NN}_\ell}(\Phi)\} = \prod_{i=0}^{|A|-1} \frac{\partial c_\ell [\Phi_A, \Phi_B]}{\partial (\Phi_A)_i} \,. \tag{2.29}$$

Furthermore, using an affine coupling [88]

$$c_\ell [\Phi_A, \Phi_B] = e^{m_\ell(\Phi_B)} \odot \Phi_A + a_\ell (\Phi_B) \tag{2.30}$$

with arbitrary differentiable functions $m_\ell, a_\ell : \mathbb{C}^{|\Lambda|/2} \to \mathbb{C}^{|\Lambda|/2}$ acting on the B indices of the input configuration Φ, yields a computationally cheap (log) Jacobian determinant

$$\log \det\{J_{\mathcal{NN}}(\phi)\} = \sum_{\ell=1}^{L} \sum_{i=0}^{|A|-1} m_\ell (\Phi_{\ell-1}(\phi)_B)_i \,. \tag{2.31}$$

The expressivity of the neural network is controlled by the trainable parameters in the coupling functions m_ℓ, a_ℓ. If f denotes an affine transformation

$$f(\Phi) = \omega \cdot \Phi + b \tag{2.32}$$

and g the nonlinear "softsign" function

$$g(z) = \frac{z}{1 + |z|} \tag{2.33}$$

we take the coupling functions to be

$$a_\ell, m_\ell = g \circ f \circ g \circ f \tag{2.34}$$

with independent complex weight matrices ω and bias vectors b. The softsign function is non-holomorphic, requiring us to consider the Jacobian in the Wirtinger sense (2.24). Due to the structure of the Jacobian matrix, the nonzero non-holomorphic components $\partial_{\Phi_B^*} c_\ell$ do not contribute to the determinant (A.14). A graphical representation of this architecture is displayed in Figure 2.2.

We add layers in pairs so that L is even. Each pair shares their partitioning. In each pair the first layer modifies the A partition (A.12) and the next modifies the B partition using the same ansatz with independent weights and biases. Notice, the Jacobian determinant can be implemented so it is evaluated during the forward pass [88] which reduces the required additional cost to only the sums of equation (A.14). Consequently, the Jacobian determinant in the effective action (A.8) only adds a computational complexity linear in the volume $|\Lambda|$.

The training setup was kept simple, allowing for further improvements in the future. A standard L_1 loss function and the ADAM algorithm implemented in PyTorch [90] was used to train the network. We kept the ADAM specific hyper parameters – running average coefficients $\beta_i = (0.9, 0.999)$, denominator shift $\varepsilon = 1 \times 10^{-8}$ as well as weight decay $w = 0$ – at the standard values. The training data comprised 10 000 (16 000 for the 18 Sites) configurations drawn from normal distributions $\phi \sim \mathcal{N}_{0,\sigma}$, with σ uniformly sampled between $\sqrt{U/(1+16/N_t)}$ and $\sqrt{U}$ [61], as input. The "labels" consist of the corresponding flowed configurations $\mathfrak{I}_\tau^+(\phi)$, where the integration is performed using an adaptive Runge Kutta method of 4th order. A similar setup is used for the validation and testing data but only for 2000 configurations each. To avoid learning features of the thimbles irrelevant to the integral [61, 58, 54, 71], only configurations that did not flow to neverland are included in the training.

The network $\mathcal{NN}$ with 2 pairs of coupling layers was initialized to the identity so that before training it reproduced the tangent plane configurations which were fed into it. We experimented with learning $\mathfrak{I}_\tau^+$ different flow times $\tau \in \{1 \times 10^{-6}, 1 \times 10^{-5}, 1 \times 10^{-3}, 1 \times 10^{-2}, 1 \times 10^{-1}\}$. We computed both the statistical power and measured correlators, as in Figure 2.4. If we flow too much most configurations flow to neverland and training becomes expensive; if we flow too little the statistical power hardly improves. The results shown in the next section have a flow-time $\tau = 1 \times 10^{-1}$.

Unfortunately picking a fixed flow time of this size was not feasable for the 18 sites problem. Instead, we defined a window of flow times $\tau \in [0, 0.1]$ on which the flow is performed, as was originally done in [61]. In this manner, fixed flow-time configurations which would have flowed to neverland and thus have been rejected could still be used if their configurations remained valid within the window of flow times. It was found in [61] that this method greatly decreased the cost generating training data. In future work will continue to investigate more efficient ways of generating training data, and the training process itself, including by sampling one training point from the steps along a holomorphic flow to $\tau = \infty$ according to the real part of the step's action.

2.4 Results

We simulate the Hubbard model on the honeycomb lattices of 2, 4, 8 and 18 sites shown in 2.1, using configurations obtained on the tangent plane and via our neural network $\mathcal{NN}$, at inverse temperature $\beta = 4$, $N_t = 32$ timeslices, on-site coupling $U = 4$, and chemical potential $\mu = 3$. To compare the machine learning enhanced HMC to other implementations such as the real-plane (standard) HMC with molecular dynamics on $\mathcal{M}_\mathbb{R}$ and the tangent plane HMC on $\mathcal{M}_\mathbb{T}$ we consider the statistical power Σ. A suitable algorithm will have $|\Sigma|$ close to 1, whereas low values indicate a less suitable algorithm, since considerably more statistics would be required (A.7). Figure 2.3 shows estimates of $|\Sigma|$ with different numbers of configurations for the three mentioned HMC variants. The ML HMC is shown in blue, the tangent plane HMC in orange and the real plane HMC in red. The ML HMC outperforms the two other algorithms in every case. Moreover, in the case of 8 sites enormous statistics are required to even get a reasonable estimate of the statistical power for the real- and tangent-plane HMCs while the power of the ML HMC stabilizes with far fewer samples. For 18 sites it was not feasible to simulate with the real plane HMC thus it is not shown here. We can see that the tangent plane HMC does not get any reliable value for the statistical power while the MLHMC converges relatively fast.

We show the efficacy of our method by computing euclidean-time correlators for a single

Figure 2.3: The statistical power $|\Sigma| = \left|\left\langle e^{-i\,\mathrm{Im}\{S\}}\right\rangle\right|$ is plotted against the number of configurations. Three different algorithms are compared, the real plane (standard) HMC (orange), the tangent plane HMC (red) and the ML HMC (blue). For 18 sites the real plane was totally noisy, and it is left out here. It can be seen that the ML HMC outperforms both real and tangent plane HMCs.

particle or single hole created at time 0 and site y and destroyed at time t and site x.

$$C_{xy}^{p}(t) = \left\langle p_x^{\dagger}(t) p_y(0) \right\rangle = \left\langle M[+\Phi| + \mu]_{xt;y0}^{-1} \right\rangle \tag{2.35}$$

$$C_{xy}^{h}(t) = \left\langle h_x^{\dagger}(t) h_y(0) \right\rangle = \left\langle M[-\Phi| - \mu]_{xt;y0}^{-1} \right\rangle \tag{2.36}$$

To improve our signal we average on time slices in $t \in [\delta, \beta - \delta]$,

$$C_{xy}(t) = \frac{1}{2} \left(C_{xy}^{p}(t) + C_{xy}^{h*}(\beta - t) \right) ; \tag{2.37}$$

addends equal by symmetry even when $\mu \neq 0$. We then project both spatial indices to the same momentum k to construct $C_k(t)$ for each momentum allowed by the lattice, and average correlators whose momenta are equal by rotational symmetry.

The match of our correlators in 2.4 with the exact results demonstrate that our algorithm is sampling the correct distribution. Each row of the figure corresponds to one of the exactly-diagonalizable system sizes and each column restricts the number of configurations N_{cfg} used to estimate the correlators. The red correlators are determined using a tangent plane HMC, the blue ones using ML HMC. Finally, the black dashed lines correspond to the correlators obtained by an exact diagonalization procedure. For the smaller examples the statistical errors of ML-HMC are much smaller, especially with fewer samples, as is expected from their respective statistical powers shown in 2.3. The worst sign problem can be found in the 8 sites case. Here the tangent plane HMC fails even for $N_{conf} = 100\,000$ and the statistical uncertainty in the correlators is essentially 100%. ML HMC obtains a weak signal at $N_{conf} = 50\,000$ configurations and improves with greater statistics.

Finally, we compute correlators for a system with 18 sites and the same parameters but with $U = 3$ which is not tractable by exact diagonalization. As shown in the statistical power plot 2.3 this model has a severe sign problem which could not be previously overcome. Again comparing tangent plane and ML HMC in 2.5 it can be seen that the ML HMC outperforms the tangent plane HMC and with the $100\,000$ measurements quite a good signal is obtained.

In all cases we measured on every 10th configuration such that no appreciable autocorrelation is found. All these simulations indicate that the neural network improves the statistical power and uncertainty in observables quite drastically even when using a simple architecture. We anticipate further improvements of our network by incorporating additional layers or incorporating knowledge of the problem's symmetries using equivariant layers [92, 93, 94].

The main advantage of our new complex architecture lies in the efficiency of the Jacobian determinant (A.14) calculation. The form of the determinant (A.14) shows that it can be computed during the forward pass, reusing intermediate results from the application of the network, and is linear in the volume $|\Lambda|^{\alpha_{NN}}$

$$\alpha_{NN} = 1 . \tag{2.38}$$

The calculation of the determinant using a SHIFT layer [60, 61] with the implementation of PyTorch [90] (through LU-decomposition) scales with the third power of volume, $|\Lambda|^{\alpha_{\text{SHIFT}}}$ i.e. $\alpha_{\text{SHIFT}} = 3$. Measurements of the execution times of the determinant for the two neural network architectures are compared in Figure 2.6. The left panel shows the execution time per layer of $\log \det\{J\}$ for different artificial system volumes. These volumes define the size of the configuration Φ which is randomly sampled and then passed to the networks. On the log-log plot the linear behavior in the region $|\Lambda| > 1 \times 10^7$ – for α_{NN} – and $|\Lambda| > 7 \times 10^2$ – for α_{SHIFT} – determines the algorithms' scaling. A simple least square fit provides the scaling exponents

$$\begin{aligned} \alpha_{\text{SHIFT}} &= 2.955(1) \\ \alpha_{NN} &= 1.008(1), \end{aligned} \tag{2.39}$$

confirming our expected scaling behavior. We then calculate the speedup achieved with the complex over the SHIFT network architecture in the right panel of figure 2.6. The expected quadratic speedup

is confirmed by the benchmark result of

$$\alpha_{\text{SHIFT}} - \alpha_{\mathcal{NN}} = 1.947(2).$$

(2.40)
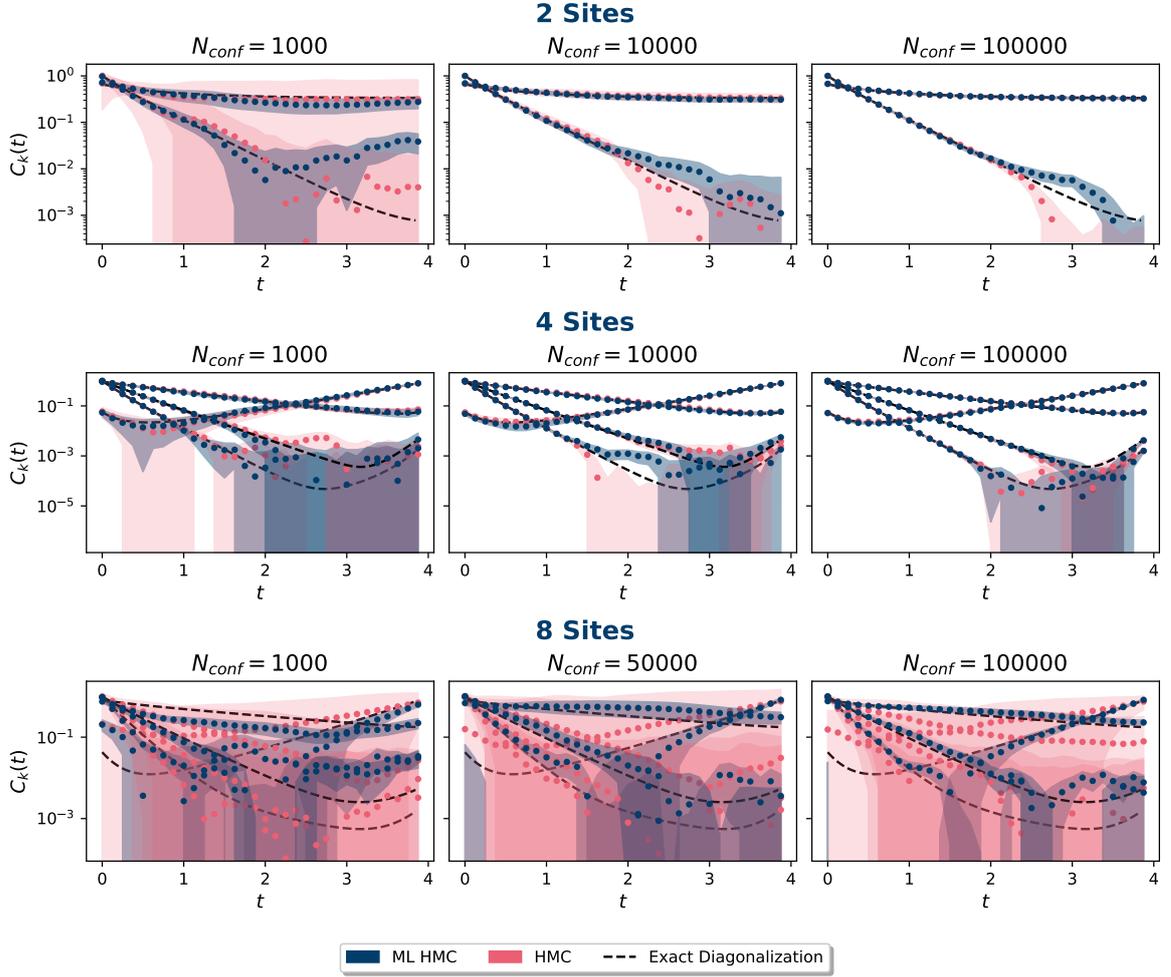
Figure 2.4: Momentum-projected correlation functions, measured using tangent plane HMC and ML HMC are shown in red and blue, respectively. These correlators were calculated with an inverse temperature $\beta = 4$, $N_t = 32$ time slices, on-site coupling $U = 4$, and chemical potential $\mu = 3$. The dashed black lines were determined by exact diagonalization. Each row corresponds to different number of ions (as in Fig. 2.1) increasing from top to bottom. Each column uses a different number of configurations N_{cfg} to estimate the correlators, increasing from left to right. Comparing the statistical power per N_{cfg} from figure 2.3 suggests to use $N_{cfg} = 1000, 10\,000, 100\,000$ for 2 and 4 sites as the uncertainty strongly differs. However, for 8 sites there is not much difference in the uncertainty of $|\Sigma|$ between $N_{cfg} = 1000$ and $10\,000$; we show $N_{cfg} = 1000, 50\,000, 100\,000$ instead. ML HMC's improved statistical power shown in Fig. 2.3 is reflected in the accuracy and uncertainty of these correlators. The sign problem of the 2 site and 4 site models is mild enough such that the tangent plane gives fairly good results, but the ML HMC gives more precise results with fewer configurations. For 8 sites the tangent plane HMC completely fails even at $N_{cfg} = 100\,000$ while ML HMC succeeds.
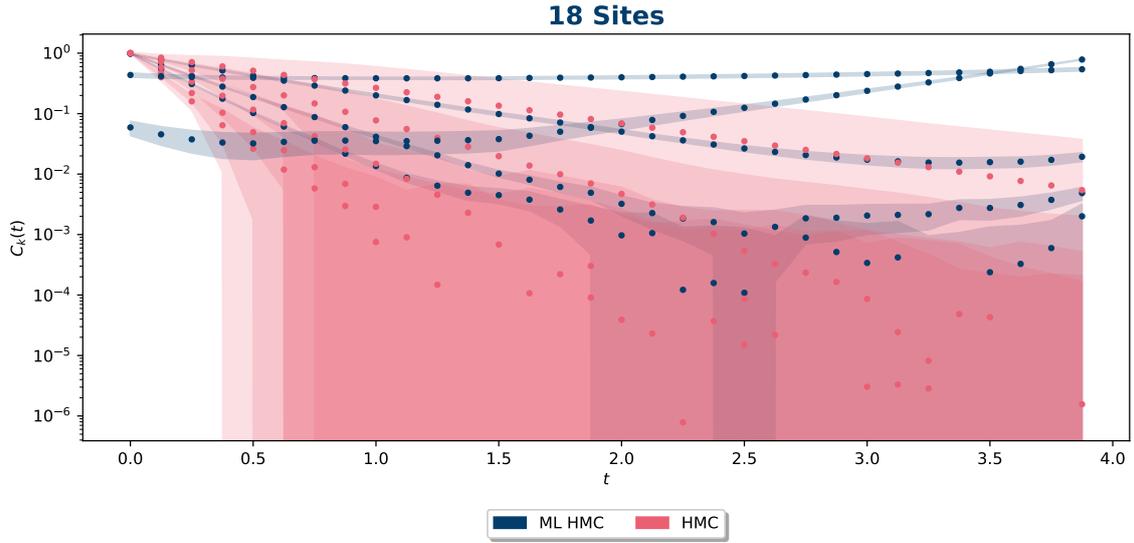
18 Sites

Figure 2.5: The single-particle correlators are displayed for the 18-site model at an inverse temperature $\beta = 4$, with $N_t = 32$ number of time slices, with $U = 3$ and $\mu = 3$. The correlators have been measured with $N_{cfg} = 100\,000$ configurations. Again, the tangent plane HMC (red) does not provide any insight while the ML HMC resolves the correlators well. Assuming similarity to the smaller $U = 4$ examples, ML HMC clearly determines the low-energy correlator while tangent plance HMC fails to find it at all.
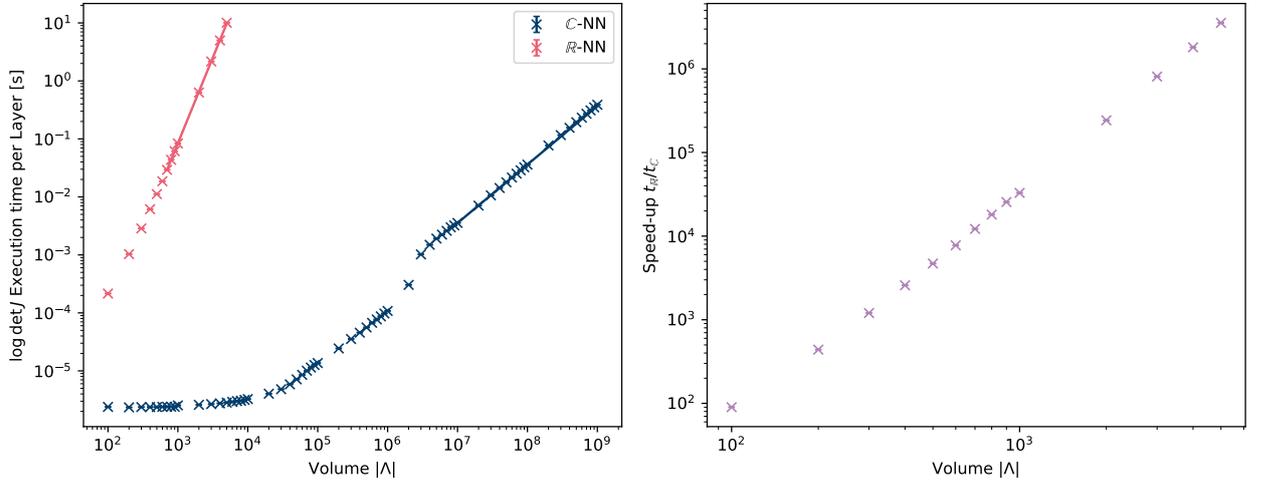
Figure 2.6: The left panel shows the scaling behavior per layer of $\log\det\{J\}$ for the previously used SHIFT neural network, $\tilde{\Phi} = \Phi + iNN(\Phi)$ (red), and for the complex-valued paired affine coupling neural network, $\tilde{\Phi} = \mathcal{NN}(\Phi)$ (blue). On the right panel, we show the speedup for the different system volumes. Theoretically the Jacobian determinant scaling of the SHIFT network is expected to be cubic in the system volume while the $\mathcal{NN}$ is expected to scale linearly, resulting in a quadratic speed-up. The solid lines, on the left panel, represent log-log fits whose slopes determine the measured scaling orders. We find for the slopes of the SHIFT layer (red) a value of $2.955(1)$ and for $\mathcal{NN}$ (blue) a value of $1.008(1)$, resulting in a scaling improvement of power $1.947(2)$. The timing measurements were performed on JURECA [91] using one AMD EPYC CPU.

31

2.5 Conclusions

Mitigating the sign problem induced by a complex action is a major target of algorithmic development for simulating quantum-mechanical systems. The application of neural networks approximating Lefschetz thimbles have shown great promise in the past. We show that the supervised training of a simple complex-valued neural network architecture – paired affine coupling layers with complex weights and biases – allows for the successful simulation of systems with increasingly severe sign problems. Our ML HMC approach reduces the sign problem sufficiently and enjoys a statistical power much greater than vanilla real-plane or tangent-plane HMC, as shown in Figure 2.3, improving the reliability of the correlator estimators in Figure 2.4. We demonstrated the fidelity and correctness of our method by simulating 2, 4 and 8 site models and comparing our results to that obtained from direct diagonalization, obtaining excellent agreement. We then applied our method to the 18 sites problem where direct diagonalization is not realizable. Our results here thus represent predictions for this system in a regime where standard Monte Carlo methods are not possible due to the severity of the sign problem.

Our results were made possible due to the favorable volume scaling of our new method. Compared to previous methods we drastically reduced the computational cost of the Jacobian determinant from a general cubic scaling down to linear in the volume. This has been numerically tested and demonstrated in Figure 2.6. Our computational complexity is therefore dominated by the application of the neural network itself, and can be further improved by using sparse methods, convolutional layers, or other layer architectures. We are actively investigating such possibilities.

Acknowledgments

2.6 Closing Remarks

The definition of the Jacobian for a $\mathbb{R} \to \mathbb{C}$ function is not straightforward in the case of a non-holomorphic function. This section presents additional calculations that did not appear in the published paper [1].

Consider a cycle $\mathcal{M}$, which is parametrized by a function $f : \mathbb{R} \to \mathcal{M}$. In equation 2.24 we explicitly recreate the derivative with respect to the real variable, inspired by combining the two derivatives (2.23). This argument can also made by considering the complexification of $\phi \to \Phi = \phi + i\,\Phi_c^0$.

$$\frac{\partial f(\Phi)}{\partial \phi} = \frac{\partial f(\Phi)}{\partial \Phi} \underbrace{\frac{\partial \Phi}{\partial \phi}}_{=\mathbb{1}} + \frac{\partial f(\Phi)}{\partial \Phi^*} \underbrace{\frac{\partial \Phi^*}{\partial \phi}}_{=\mathbb{1}} \tag{2.41}$$

$$= \frac{\partial f(\Phi)}{\partial \Phi} + \frac{\partial f(\Phi)}{\partial \Phi^*} \tag{2.42}$$

However, the Wirtinger derivative is formally defined in a $\mathbb{C} \to \mathbb{C}$ context. so we have to analytically continue first, $\phi \in \mathbb{R} \to \Phi \in \mathbb{C}$, and parameterize $\mathcal{M}$ by $g : \mathbb{C} \to \mathbb{C}$. Then a natural extension of

the induced Jacobian is simply

$$\frac{\partial g(\Phi + \mathrm{i}\,\Phi_c^0)}{\partial \Phi} = \frac{\partial g(\varphi)}{\partial \varphi}\underbrace{\frac{\partial \varphi}{\partial \Phi}}_{=\mathbb{1}}\bigg|_{\varphi = \Phi + \mathrm{i}\,\Phi_c^0} + \frac{\partial g(\varphi)}{\partial \varphi^*}\underbrace{\frac{\partial \varphi^*}{\partial \Phi}}_{=0}\bigg|_{\varphi = \Phi + \mathrm{i}\,\Phi_c^0} \tag{2.43}$$

When we treat the complexification $\phi \to \Phi(\phi) = \phi + \mathrm{i}\,\Phi_c^0$ of (2.41) as an analytic function, the second term in (2.41) vanishes and we recover the result (2.43). The argument extends to $\mathbb{C}^V$.

The next consideration is the Jacobian of the network, after (2.43) suppressing the evaluation of the linear shift. We start with a single layer

$$\frac{\partial \mathcal{NN}_\ell(\Phi)}{\partial \Phi} = \begin{pmatrix} \frac{\partial c_\ell(\Phi_A,\Phi_B)}{\partial \Phi_A} & 0 \\ \frac{\partial c_\ell(\Phi_A,\Phi_B)}{\partial \Phi_B} & \mathbb{1} \end{pmatrix} \tag{2.44}$$

Now its determinant is not specific to the use of Wirtinger derivatives as the only contributing parts are defined to be analytic, condition (2.28). Thus we recover (A.14) for a single layer.

This alone is not very useful, as only half the configuration is transformed. Thus we consider a pair of two layers with alternating use of the partition of A & B, suppressing the layer index ℓ, we have

$$\mathfrak{P}\mathcal{NN}(\Phi) = \mathcal{NN}_{BA}(\mathcal{NN}_{AB}(\Phi)) = \begin{pmatrix} c_1(\Phi_A, \Phi_B) \\ c_2(\Phi_B, c_1(\Phi_A, \Phi_B)) \end{pmatrix}. \tag{2.45}$$

Using this, we can find the jacobian

$$J_{\mathfrak{P}\mathcal{NN}} \equiv \begin{pmatrix} J_{11} & J_{12} \\ J_{21} & J_{22} \end{pmatrix} = \begin{pmatrix} \frac{\partial c_1(\Phi_A,\Phi_B)}{\partial \Phi_A} & \frac{\partial c_1(\Phi_A,\Phi_B)}{\partial \Phi_B} \\ \frac{\partial c_2(\Phi_B,c_1(\Phi_A,\Phi_B))}{\partial \Phi_A} & \frac{\partial c_2(\Phi_B,c_1(\Phi_A,\Phi_B))}{\partial \Phi_B} \end{pmatrix} \tag{2.46}$$

$$= \begin{pmatrix} \mathbb{1} & 0 \\ J_{21}J_{11}^{-1} & \mathbb{1} \end{pmatrix} \cdot \begin{pmatrix} J_{11} & 0 \\ 0 & J_{22} - J_{21}J_{11}^{-1}J_{12} \end{pmatrix} \cdot \begin{pmatrix} \mathbb{1} & J_{11}^{-1}J_{12} \\ 0 & \mathbb{1} \end{pmatrix}, \tag{2.47}$$

where the second line is a (block) LU decomposition that simplifies the calculation of the determinant as the L,U have determinant 1. This is only applicable if $J_{11} = \frac{\partial c_1(\Phi_A,\Phi_B)}{\partial \Phi_A}$ is invertible. Indeed, the structure of $c_\ell(A,B)$, compare equation (A.13), implies $J_{\ell\ell} = e^{m_\ell(\Phi_B)}$, which is invertible. Taking the determinant of (2.47) factorizes to

$$\det\{J_{\mathfrak{P}\mathcal{NN}}\} = \det\{J_{11}\}\det\{J_{22}\}\det\{\mathbb{1} - J_{22}^{-1}J_{21}J_{11}^{-1}J_{12}\} \tag{2.48}$$

$$= \det\left\{\frac{\partial c_1(\Phi_A,\Phi_B)}{\partial \Phi_A}\right\}\det\left\{\frac{\partial c_2(\Phi_B,c_1)}{\partial \Phi_B}\right\} \tag{2.49}$$

$$\times \det\left\{\mathbb{1} - \left[\frac{\partial c_2(\Phi_B,c_1)}{\partial \Phi_B}\right]^{-1}\frac{\partial c_2(\Phi_B,c_1)}{\partial c_1^*}\frac{\partial c_1^*(\Phi_A,\Phi_B)}{\partial \Phi_B}\right\}. \tag{2.50}$$

Here the first two factors, (2.49), give rise to the Jacobian factorization as desired in (A.14). In the case of holomorphic couplings, the third factor, (2.50), reduces to one since $\frac{\partial c_\ell^*(A,B)}{\partial B} = 0$. Unfortunately, this is not true for non-holomorphic functions couplings.

We can take advantage of this result by defining an effective Jacobian containing only the first two factors and reweight according to the non-holomorphic contribution.

$$\langle O \rangle_{\mathrm{S}} = \frac{\langle O \cdot \det\{\mathbb{1} - J_{22}^{-1}J_{21}J_{11}^{-1}J_{12}\}\rangle_{\mathrm{S_{eff}}}}{\langle \det\{\mathbb{1} - J_{22}^{-1}J_{21}J_{11}^{-1}J_{12}\}\rangle_{\mathrm{S_{eff}}}} \tag{2.51}$$

This involves sampling according to the effective action (A.8) using only the efficient to compute jacobians (2.49). The advantage of this approach is that the entire Jacobian, which is expensive to

compute, does not need to be evaluated during trajectory generation, but only for the configurations that are used for measuring observables.

In figure 2.7, the difference between the log det of the effective Jacobian and the full Jacobian J' is plotted per configuration. Additionally, figure 2.8 shows the evolution of the action, log dets of the effective Jacobian and the full Jacobian. These demonstrate, that the contribution of the non-holomorphic part, which is of order $O(10^{-3})$, is negligible compared to the effective Jacobian. Furthermore, the calculation of this non-holomorphic contribution can only be avoided with holomorphic couplings. An example of this is discussed in the next chapter on the example of perylene described by the Hubbard model.
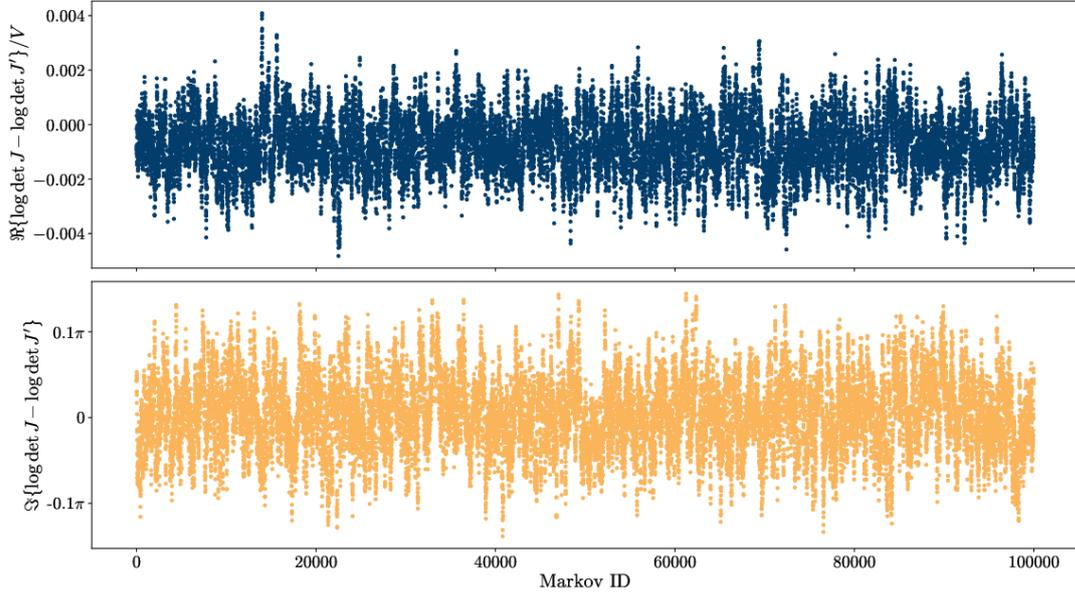
Figure 2.7: Difference between the log-determinant of the effective Jacobian 2.49 and the full Jacobian. This equals the logarithm of the non-holomorphic contribution, equation 2.50. It appears negligible compared to the overall magnitude of the Jacobian shown in figure 2.8. These are evaluated on the 4-site model described in this chapter.
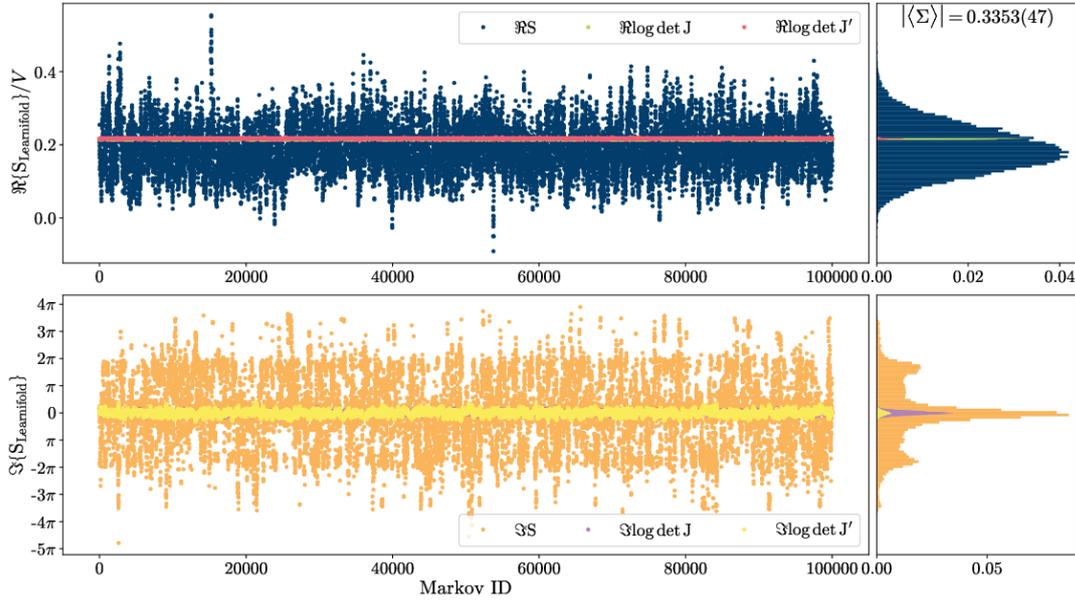
Figure 2.8: Evolution of real and imaginary part of the action S, compare A.3, blue and orange respectively, effective Jacobian, green and purple respectively, and full Jacobian, red and yellow respectively. These are evaluated on the 4-site model described in this chapter.

Chapter 3

Applying Neural Networks to Simulate Real Systems with Sign Problems

This chapter is based on [2]:

This chapter continues the work on complex valued neural networks to reduce the sign problem of the Hubbard model away from half filling. To avoid calculating the overall Jacobian on the measured configurations, a holomorphic architecture is tested. Holomorphic architectures do not lead to general function approximators [87]. However, the rationale behind this approach is that an exact approximation of the flow is not necessary; rather, a transformation that reduces the sign problem is desired.

Non-holomorphic activation functions were used to escape Liouville's theorem, that is all entire bounded complex functions are constant. For a holomorphic neural network, it is consequently inevitable to have it unbounded, which necessitates the implementation of measures ensuring the preservation of homology. It is clear, that a projection to the real or tangent plane for diverging configurations is enough to ensure homology [23]. These projections, with the use of a tunable parameter, can be pushed to large enough values such that they have practically no influence on any finite simulation.

Furthermore, different methods for generating training data are explored. The algorithm presented here employs a Runge-Kutta to flow exponentially-sampled configurations. This idea is adapted from the algorithm in the previous chapter. Under holomorphic flow, the real part of the action monotonically increases, and almost all configuration flow towards neverland – a region of infinite action where different thimbles meet [23]. To reduce the number of configuration close to neverland, the real part of the action is tracked during evolution. Once the change starts to diverge, as the derivative becomes larger and larger, the evolution is stopped. This procedure ensures that the training data contains more relevant configurations in relation to the true Boltzmann distribution than it would be the case if the data were just evolved. This represents an advancement over the process used in the previous chapter.

Discussed architecture is tested in a molecule with $N_x = 20$ ions, namely the molecule $C_{20}H_{12}$ Perylene. Perylene has a bipartite lattice [95, 96, 97], a sign problem thus appears only away from half filling. As perylene naturally appears in derivatives with a perylene core, it naturally is not found at half filling [98]. Contrary to graphene, discussed in the previous chapter, a spatial infinite volume limit is not required for this molecule making it a perfect target to study the efficiency of the sign problem algorithms with direct practical results.

Perylene has several applications in industry and other fields of physics. For example it can be used as a dopant in organic light-emitting diodes (OLEDs) [99], as organic semiconductors [26], or as an acceptor material in organic solar cells [100]. It is further interesting for astronomy as it is found in interstellar gases and nebulae [101, 102]. Thus, understanding the electronic properties of this molecule is of great value.

In this chapter, a chemical potential scan, i.e. a scan over doping, is performed at fixed discretization, $N_t = 32$, and inverse temperature $\beta = 4$. The simulations are done on the tangent plane, compare equation 1.47, where most chemical potentials exhibit a mild sign problem at this temperature. However, to improve the resolution of the total system charge, a single point, specifically at $\mu = 1.8$, was treated with the afore mentioned path deformation, learned by the holomorphic neural network. Significant reduction of the statistical uncertainty is found using this neural network. A variational basis of single particle operators is computed to provide insight into the single particle spectrum. Effective masses have been computed for a selected set of chemical potentials to provide a first estimate of energies.

A more thorough analysis of the molecule including a continuum extrapolation and careful analysis of the energy levels can be found in the next chapter 4.

Ideas, development, and writing is primarily done by myself, based on the results of the previous chapter. Co-authors were involved through discussions and advice. Maria Dincă, as DAAD Rise student, helped testing and debugging components of the architecture.
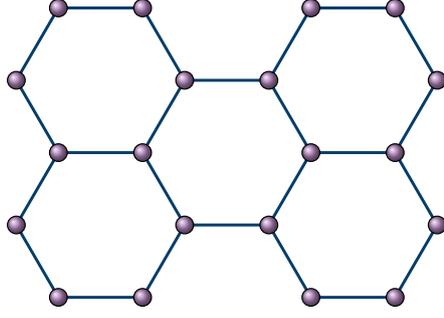
Figure 3.1: Graphical representation of the perylene molecule. The sites represent carbon-ions, while links indicate allowed hopping. External hydrogen atoms are not drawn.

3.1 Introduction

The challenge of the computational sign problem poses obstacles for effective importance sampling when dealing with complex-valued distributions in Markov-Chain Monte-Carlo algorithms, like Hybrid Monte Carlo (HMC). This challenge is pervasive when sampling from the configuration space of many physical systems, including but not limited to lattice QCD at finite baryon chemical potential, doped condensed matter systems in equilibrium, and the real-time evolution of quantum systems. A method to substantially alleviate the sign problem involves deforming the original (real) manifold of integration into a sign improved (complex) one [62, 23, 63, 64]. Formal developments have spurred exploration into leveraging Lefschetz thimbles [21, 54, 55, 56, 57, 58, 59] —these are high-dimensional counterparts to contours of steepest descent and can be located through holomorphic flow. However, pinpointing the location of each relevant thimble's saddle point, or critical point, and determining the relevant sampling 'direction' around these points is prohibitive. An alternative approach is to employ neural networks to learn the mapping from an initial manifold to a beneficial one, including one that approximates the contributing thimbles to the integral [60, 65, 61, 1].

In this work, we explore the use of this method for perylene, $C_{20}H_{12}$ [97]. The molecule is (almost) planar and sp^2-hybridized making it a perfect target for the Hubbard model [95]. We show some preliminary results for a chemical potential (doping) scan. Practical applications range from organic semiconductors [26], over organic light emitting diodes (OLEDs) [99], to organic solar cells [100]. Perylene also falls under the class of so-called polycyclic aromatic hydrocarbon (PAH) molecules thought to be ubiquitous in interstellar gases and nebulae [101]. A detailed accounting of electronic properties of PAHs, like perylene, could help constrain interstellar gas models [102].

3.2 Formalism

Given a fixed spatial arrangement of ions X, the Hubbard model in particle-hole basis reads [66, 30, 31]

$$\mathcal{H}[K, V, \mu] = - \sum_{x,y \in X} \left(p_x^\dagger K^{xy} p_y - h_x^\dagger K^{xy} h_y \right) + \frac{1}{2} \sum_{x,y \in X} \rho_x V^{xy} \rho_y + \mu \sum_{x \in X} \rho_x, \quad (3.1)$$

where the amplitudes in K encode the hopping of fermionic particles p and holes h, interactions between these are modelled by the potential V, $\rho_x = p_x^\dagger p_x - h_x^\dagger h_x$ is the net charge per site, and the chemical potential μ incentivizes charge. We restrict our attention to the case where $K = \kappa \delta_{\langle xy \rangle}$ encodes the structure of perylene, compare figure 3.1, with nearest-neighbor hopping, and an on-site interaction $V = U \delta_{xy}$. In this case, when $\mu = 0$, the Hamiltonian in (A.1) corresponds to the neutral, 'half-filled' system.

Trotterizing the thermal trace into N_t timeslices, inserting Grassmannian resolutions of the identity, and linearizing the interaction via the Hubbard-Stratonovich transformation [29] exposes the Hubbard action

$$S[\Phi \,|\, K, V, \mu] = \frac{1}{2} \sum_{\substack{x,y \in X \\ t \in [0, N_t - 1]}} \Phi_{tx} (\delta V^{-1})^{xy} \Phi_{ty} - \log \det\{M[\Phi \,|\, K, \mu] \, M[-\Phi \,|\, -K, -\mu]\},$$

(3.2)

where $\Phi \in \mathbb{R}^{|\Lambda|}$ is the (auxiliary) hubbard field on the spacetime lattice $\Lambda = [0, N_t - 1] \otimes X$ and $\delta = \beta/N_t$. We use the exponential discretization [19] for the fermion matrices

$$M[\Phi \,|\, K, \mu]_{x't';xt} = \delta_{x'x}\delta_{t't} - \left(e^{\delta(K+\mu)}\right)_{x'x} e^{+i\Phi_{xt}}\delta_{t'(t+1)}$$

(3.3)

with antiperiodic boundary conditions in time. With this we can express the thermal trace as path integral

$$\langle O \rangle = \frac{1}{\mathcal{Z}} \text{Tr}\{O e^{-\beta \mathcal{H}}\} = \frac{1}{\mathcal{Z}} \int \mathcal{D}[\Phi] \, e^{-S[\Phi]} O[\Phi] .$$

(3.4)

The partition function $\mathcal{Z}$ is the trace/integral without the observable. On a bipartite lattice we may replace the $-K$ in the holes' fermion matrix with $+K$ leading to a real and positive determinant at vanishing chemical potential. When μ is finite the determinant is complex valued and results in a sign problem.

3.3 Method

As in our previous work [61, 1, 7], for a Monte-Carlo algorithm with complex action we can separate the real and imaginary parts, $S = \text{Re}\{S\} + i\,\text{Im}\{S\}$, and rewrite the path integral as

$$\mathcal{Z} = \int \mathcal{D}[\Phi] \, e^{-S} = \int \mathcal{D}[\Phi] \, e^{-\text{Re}\{S\}} e^{-i\,\text{Im}\{S\}} \propto \left\langle e^{-i\,\text{Im}\{S\}} \right\rangle_{\text{Re}\{S\}} \equiv \Sigma$$

(3.5)

where $|\Sigma| \in [0, 1]$ is the statistical power. Sampling according to the real part of the action and then applying reweighting,

$$\langle O \rangle = \frac{\left\langle e^{-i\,\text{Im}\{S\}} O \right\rangle_{\text{Re}\{S\}}}{\left\langle e^{-i\,\text{Im}\{S\}} \right\rangle_{\text{Re}\{S\}}} = \frac{1}{\Sigma} \left\langle e^{-i\,\text{Im}\{S\}} O \right\rangle_{\text{Re}\{S\}} ,$$

(3.6)

allows the computation of observables. This procedure fails if the statistical power $|\Sigma|$ cannot be reliably distinguished from zero [20, 61, 54, 71].

Combining reweighting with path deformation techniques offers a promising approach to mitigating the sign problem. The primary objective is to expand the accessible parameter space for computation, ultimately allowing extrapolations to the continuum limit, $\delta \to 0$, and/or the zero-temperature limit, $\beta \to \infty$. The core idea involves manipulating the integration contour in a way that enhances the statistical power. One noteworthy method in this context is the use of Lefschetz thimbles, which eliminate fluctuations in the imaginary part of the action. However, it is essential to acknowledge that these thimbles are notoriously challenging to compute in practical simulations. As a result, the focus shifts towards harnessing a more versatile transformation to a sign-optimized manifold, denoted as $\tilde{\mathcal{M}}$. Crucially, Cauchy's theorem assures us that expectation values of holomorphic observables remains unchanged under this transformation[1]:

$$\begin{aligned} \langle O \rangle &= \frac{1}{\mathcal{Z}} \int_{\tilde{\mathcal{M}}} \mathcal{D}[\tilde{\Phi}] \, e^{-S[\tilde{\Phi}]} O[\tilde{\Phi}] \\ &= \frac{1}{\mathcal{Z}} \int \mathcal{D}[\Phi] \, e^{-S[\tilde{\Phi}(\Phi)] + \log\det\{J_{\tilde{\Phi}}[\Phi]\}} O[\tilde{\Phi}(\Phi)] \end{aligned}$$

(3.7)

---

[1] Strictly speaking, this holds when no singularities are crossed when performing the contour deformation, which is the case in our problem.

A simple transformation is the mapping to the tangent plane $\tilde{\Phi}[\Phi] = \Phi + i\,\phi_0$. As outlined in [7] a transcendental equation

$$\phi_0/\delta = -\frac{U}{N_x} \sum_k \tanh\left(\frac{\beta}{2}\left[\epsilon_k + \mu + \phi_0/\delta\right]\right) \tag{3.8}$$

can be used to identify this plane. The sum is over the eigenvalues ϵ_k of the hopping matrix K. This adds no additional computational cost to the simulation and thus forms our starting point for further transformations.

Based on our earlier machine learning approach [1], we can use a neural network to map to a sign optimized manifold $\tilde{\mathcal{M}}$, i.e. $\tilde{\Phi}[\Phi] = \mathcal{NN}\Phi + i\phi_0$. We continue to use coupling networks to enable a tractable Jacobian determinant. However, we changed the activation function to be holomorphic compared to our earlier approach which simplified the training procedure significantly. A single layer thus reads

$$\mathrm{NN}_l\,[\Psi]_{t,x} = \begin{cases} e^{s_l(\Psi_B)} \odot \Psi_A + t_l\,(\Psi_B) & (t,x) \in A \\ \Psi_B & (t,x) \in B \end{cases} \tag{3.9}$$

where $s, t : \mathbb{C}^{|\Lambda|/2} \to \mathbb{C}^{|\Lambda|/2}$, $\Psi_B \mapsto w' \cdot P\,[w \cdot \Psi_B + b] + b'$ are linear networks; the activation $P\,[\cdot]$ is a polynomial of degree 3. We always pair two of these layers to form a single transformation that allows us to change every value of the input configuration. Notice, this network is unbounded at $\Psi \to \pm\infty$ and thus escapes Liouville's theorem [103, 104]. Furthermore, the homology class is completely determined by the asymptotic behavior of the network. This means that any parametrization of the contour $\tilde{\mathcal{M}}$ preserves the homology class if it approaches a constant asymptotically. To achieve this, we implement a projection layer

$$\mathcal{P}_\sigma\,[\Psi]_i = \mathrm{Re}\{\Psi\}_i + i\,\mathrm{Im}\{\Psi\}_i e^{-\frac{|\Psi_i|^2}{\sigma^2}} \tag{3.10}$$

which pushes the imaginary part down towards the real manifold. The parameter σ can be chosen arbitrarily. Choosing $\sigma \gg 1$ but finite, the projection virtually becomes the identity up to $O(\sigma^{-2})$ and the Jacobian remains unchanged $J_{\tilde{\Phi}}[\Phi] = \det\{\mathcal{NN}\Phi\} + O(\sigma^{-2})$. Care has to be taken if the configuration value $\mathcal{NN}\Phi$ becomes comparable to σ. Naturally, such configuration values have small weights; thus, they practically never appear.

We train the parameters – $w, w' \in \mathrm{Mat}_{|\Lambda|/2}\,[\mathbb{C}]$, $b, b' \in \mathbb{C}^{|\Lambda|/2}$ per layer – of the network using training data sets of flowed configurations [61]. The training data is generated by integrating the holomorphic flow equation, starting from configurations sampled with a normal-distribution $\mathcal{N}(\mu = 0, \sigma = \delta U)$ and mapped to the tangent plane, with a Runge Kutta and accepting only configurations that preserve the imaginary part of the action, are flowed enough, and are not in neverland [2]. A graphical representation of this can be found in figure 3.2. Once the data is gathered we initialize the network with random parameters $\sim U(0.01, 0.01)$ [105], and train it using the Adam optimizer [106] with a learning rate of 10^{-3}. For the first iterations we use a plain L_2-loss to measure the distance between the training data and the network output. After that a couple of iterations are trained by comparing the real part of the action plus the preservation of the imaginary part, $L\left(\tilde{\Phi}, \mathcal{NN}\Phi + i\phi_0\right) = |\Delta\,\mathrm{Re}\{S\}| + \left|1 - e^{i\,\Delta\,\mathrm{Im}\{S\}}\right|$. Once a desired precision ($< 1e - 4$) is reached, we perform a short HMC simulation, $N_{\mathrm{cfg}} = 1000$, and measure the statistical power $|\Sigma|$. If the statistical power is increased, we accept the new network parameters and continue the HMC simulation resulting in the observables presented in the next section.

---

[2] By neverland we denote regions of infinite $\mathrm{Re}\{S\}$. Points where Lefshetz thimbles meet have diverging $\mathrm{Re}\{S\}$ and are not relevant for the training.
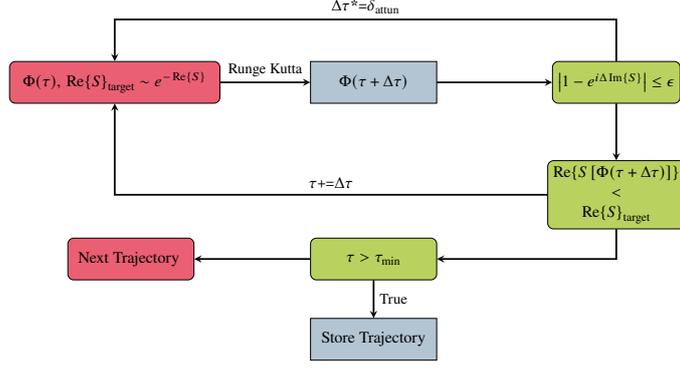
Figure 3.2: Algorithm to generate training data. Each step integrates the flow equation for a time $\Delta\tau$ and checks that the imaginary part of the action is preserved for the step size. The step size becomes updated accordingly. Further, only data that stems to be relevant for training, i.e. is not flowed to little and is not in neverland, is accepted.

3.4 Results

3.4.1 Correlation Functions

The basic building block of Monte-Carlo simulations for physical systems is the correlation function

$$C^{x,y}(\tau) = \left\langle p_x(\tau)p_y^\dagger(0)\right\rangle = \left\langle M^{-1}_{(\tau,x);(0,y)}\left[\Phi\,|\,K,\mu\right]\right\rangle, \qquad (3.11)$$

as it encodes the physics of a single excitation-annihilation process. In the following we block diagonalize the correlation matrix using the symmetry eigenspaces of the ion lattice X

$$C^k(\tau) = \delta^{k,k'}\overline{\mathfrak{u}}_{k,x}C^{x,y}(\tau)\mathfrak{u}_{y,k'}, \qquad (3.12)$$

where the unitary matrix $\mathfrak{u}$ is determined by the irreducible representations of the point group D_{2h} of X. In principle, a further diagonalization within the irreducible blocks is required, however, we found that the diagonal terms are dominant in the cases we have studied thus far and so only take the diagonal correlators for study in these proceedings. To classify the molecule and understand the entirety of its (single particle) spectrum a full parameter scan in the interaction strength and its doping needs to be performed as well as the continuum and zero temperature limit ($\beta, N_t \rightarrow \infty$). In this preliminary results we will focus on the application of the neural network to the sign problem. Hence, we fix $N_t = 32$ and $\beta = 4$, while the interaction and chemical potential strength are indicated at the specific results. In Figures 3.3 we compare correlator measurements from HMC on the tangent plane, on the left, and from machine learning enhanced HMC (MLHMC), on the right. These are computed at fixed $U = 2$ and $\mu = 1.7$. As expected the MLHMC resolves the correlator much better than the HMC on the tangent plane. This chemical potential is the first where the computational extra cost of the neural network can be justified.

3.4.2 Charge

The chemical potential is an effective description of the doping of the system. We translate this by measuring the total system charge

$$Q = \sum_{x=0}^{N_x}\langle\rho_x\rangle = \sum_{x=0}^{N_x}C^{x,x}(\tau=0). \qquad (3.13)$$

In Figure 3.4 we present the total charge as a function of the chemical potential at fixed $U = 2$. The interaction strength is the same as for the correlators in Figure 3.3. We can identify, that
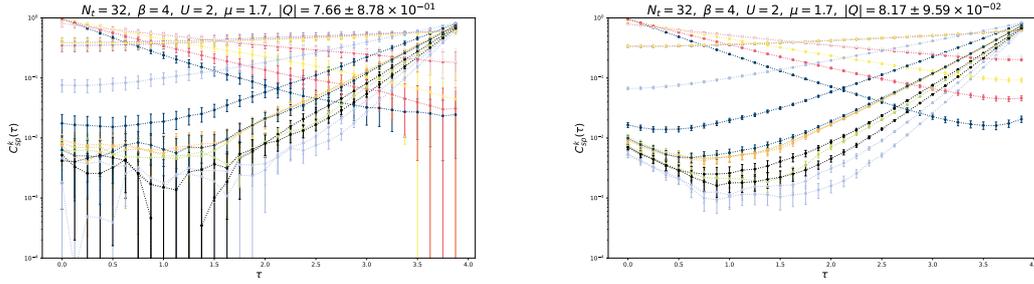
Figure 3.3: Perylene correlator for $U = 2$, $\mu = 1.7$ simulated using HMC on the tangent plane (top) and machine learning guided HMC (bottom)
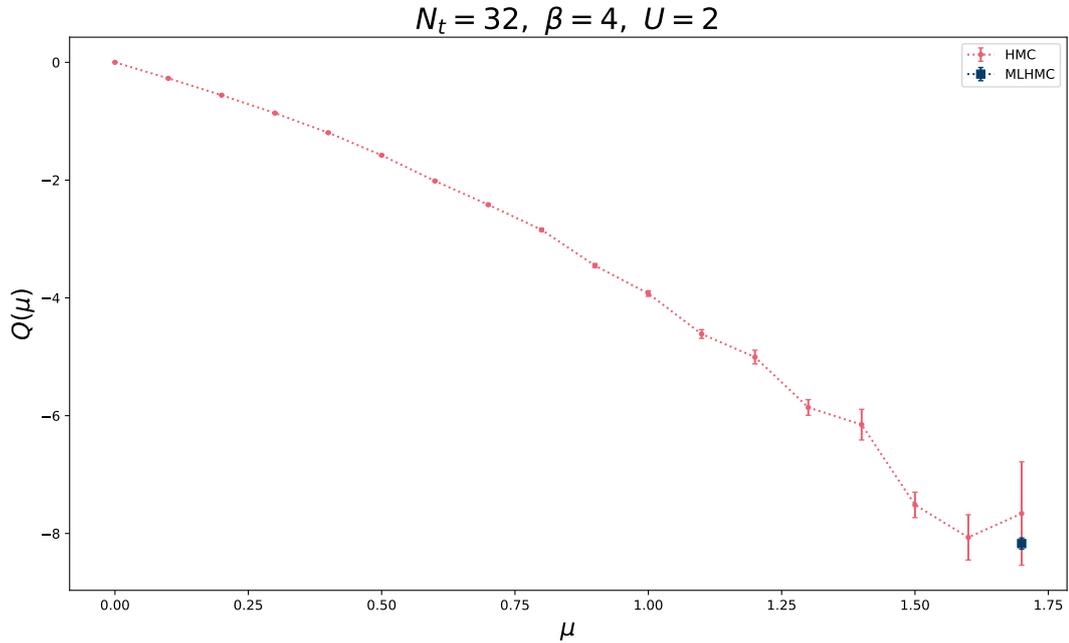
Figure 3.4: Perylene total charge μ-scan at $U = 2$

the statistical error at $\mu = 1.7$ is much smaller for the MLHMC. At cold temperatures, $\beta \to \infty$, we expect a stepwise behaviour for integer charge changes at certain chemical potentials. This is washed out here due to the small value of β. To fully classify the doping of the molecule a zero-tempreature limit must be taken which will increase the use of the neural network as the sign problem exponentially becomes worse with increasing β.

3.4.3 Effective Masses

To study the molecule for various applications like energy production in organic solar cells it is required to get an understanding of the low lying energy spectrum of a particle-hole excitation. At non-interacting systems $U = 0$, this can be done analytically as the single-particle energy is given by the eigenvalues of the hopping matrix $E_{sp}^k = \epsilon^k$. Doping simply shifts these $E_{sp/sh}^k = \epsilon^k \pm \mu$. Further, if particles and holes do not interact the exciton energy is simply the sum $E^k = E_{sp}^k + E_{sh}^k$, which, noticeably, is independent of doping. As soon as interactions are turned on this is not true anymore and the exciton energy needs to be measured explicitly involving computations of 4 point functions. For this proceedings we take a first step into that direction and measure the single particle

(a) Perylene effective mass $U = 2$, $\mu = 0$ (HMC) (b) Perylene effective mass $U = 4$, $\mu = 0$ (HMC)

(c) Perylene effective mass $U = 2$, $\mu = 0.4$ (HMC) (d) Perylene effective mass $U = 2$, $\mu = 1.7$ (MLHMC)
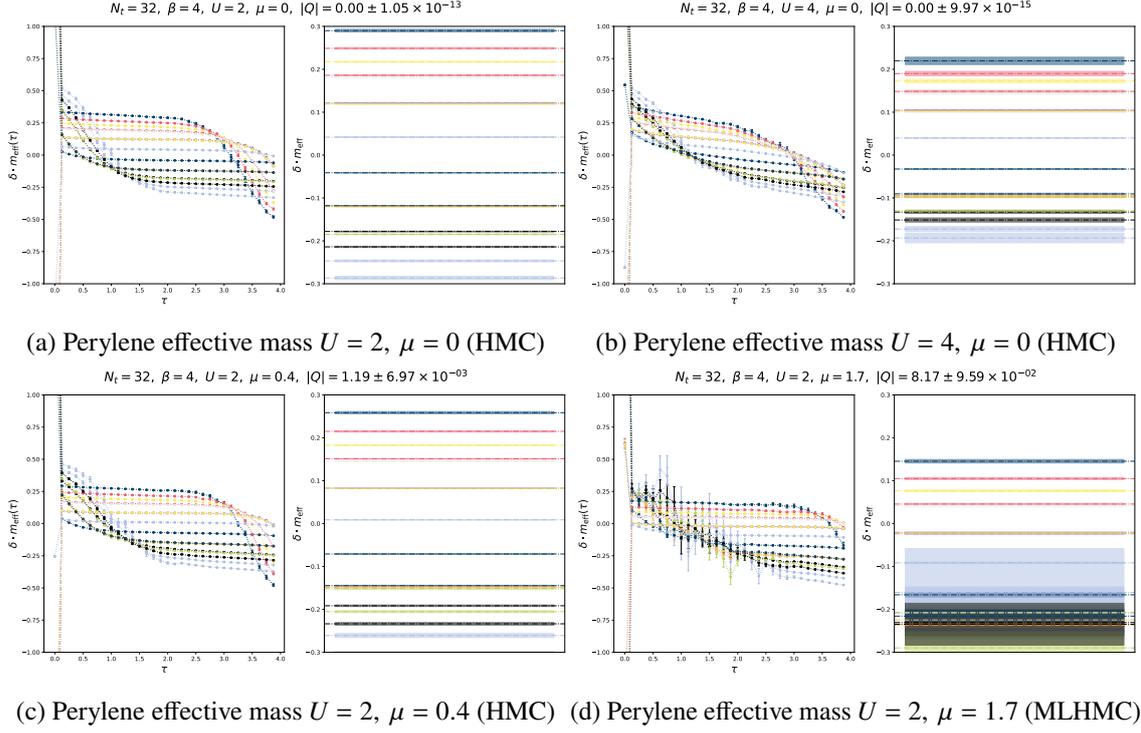
energy at various interaction strengths and chemical potentials. In the low-temperature spectral decomposition $C^k(\tau) = A_0 e^{-E_0^k \tau} + O\left(e^{-E_1^k \tau}\right)$ one expects ground state contributions in the large τ limit, defining the effective mass

$$m_{\text{eff}}^k\big|_\tau = \frac{\log C^k(\tau + \delta) - \log C^k(\tau)}{\delta}. \tag{3.14}$$

The effective mass serves as a simple estimator for the energy $E_{sp}^k \approx m_{\text{eff}}^k\big|_{\tau=\beta/2}$ that neglects excited states. In figures 3.5a to 3.5d we present a selection of effective mass plots for two interaction strengths $U = 2, 4$ (3.5a,3.5b) at zero chemical potential and two chemical potentials $\mu = 0.4, 1.7$ at fixed $U = 2$ (3.5c,3.5d). These plots suggest that the interaction strength has a great effect on the excited states, as the plateau becomes less pronounced, but only little on the ground state energy, i.e. the smallest positive energy value. Further, we can identify that the gap between states opens as the interaction strength is increased. Additionally, the change of chemical potential clearly shifts the ground state gap, which, together indicate a rich energy spectrum of the molecule.

3.5 Summary

The numerical sign problem affects Monte-Carlo simulations of various systems. In this contribution we modelled the electronic structure of the sp^2-hybridized molecule $C_{20}H_{12}$ (perylene) with a contact-interaction Hubbard model which is then simulated using the Hybird/Hamilton Monte-Carlo algorithm. At finite chemical potential, a parameter to control doping, a sign problem appears which we tackle using reweighting and contour deformations. Starting with simple deformations, the tangent plane, and continuing with neural networks at larger μ we are able to resolve the total system charge and single particle energy spectrum of perylene. To limit thermal contamination of the energy states, we require that $\beta E_0 \gg 1$, where E_0 is the lowest single particle excitation energy in the system. In our case we have $\beta E_0 \sim 1.5$, which should be increased in future calculations. We plan to extend the parameter space and perform a more detailed analysis of the system, including the study of particle-hole excitations which are relevant for practical application of the molecule.

Acknowledgments

Chapter 4

Single Particle Spectrum of Doped $C_{20}H_{12}$-Perylene

This chapter is based on [3]:

In this chapter, a Hamiltonian Monte Carlo study of doped $C_{20}H_{12}$ perylene is presented. It tight-binding the work published in [2], specifically chapter 3, to encompass lower temperatures, a continuum extrapolation, as well as a thorough analysis of the single particle spectrum.

Following the previous chapter 3, the molecule can be modelled using the Hubbard Hamiltonian, due to its sp^2 hybridized bonds [95]. As mentioned, the electronic properties of this molecule are of great interest to industry to improve and study organic semiconductors [26], organic solar cells [100], or organic light emitting diodes [99], but also to astronomy research as the polycyclic aromatic hydrocarbon (PAH) molecules, that perylene forms, are found in interstellar gases and nebulae [101, 102]

A particular focus of this study is the systematic scan over chemical potential, which corresponds to a scan over doping. This is particularly relevant as perylene naturally comes with groups attached – as perylene derivatives. These groups alter the amount of charge carriers in the system which we model effectively with the average chemical potential μ. Consequently, one primary objective of this study is to quantify the amount of doping by examining the total system charge $\langle Q \rangle$ as a function of chemical potential. Understanding this relationship allows to classify many electronic properties through the single particle spectrum. To this end, the spectrum is calculated as a function of chemical potential by calculating correlators of operators from a variational basis.

In order to facilitate this the analysis of said correlators an automatic fitting procedure is developed. This procedure requires only a set of fit models and some prior information on the system which we obtain through the analytically calculable non-interacting Hubbard model. With this procedure the single particle energy spectrum, relative to the ground state, is evaluated. Given the number of correlators to be analysed, an automatic procedure is necessary. After all, this study includes 108 ensembles divided into twelve different chemical potentials, $\mu = 0, 0.1, \ldots, 1.1$, across three inverse temperatures $\beta = 4, 6, 8$, and with three discretization, $N_t = 32, 64, 96$, to control continuum extrapolations. The on-site interaction is kept at $U = 2$. Notice, all model parameters are provided in lattice units, $\kappa = 1$. On each ensemble a total of 20 correlators are analysed to obtain values for the single particle spectrum. In total this means to analyse 2160 correlators. The results of this study demonstrate significant deviation from the non-interacting tight-binding model especially at large chemical potentials, which highlights the importance of encompassing interactions to accurately describe electronic properties of perylene.

As previously discussed, Monte Carlo simulations of the Hubbard model at finite chemical potential simulations suffer from the notorious sign problem. Unfortunately, the vast parameter space explored in this study did not permit the training of neural networks to mitigate this issue. In the previous section it was discussed that the sign problem is mitigated sufficiently on the tangent plane for a large range of chemical potentials. This conclusion assumed inverse temperature of $\beta = 4$ but does not extend to colder temperatures discussed here. Therefore, the next simple step is to simulate on the next-to-leading order plane discussed around equation (1.51), for more details see reference [7]. This ansatz allows to calculate a simple manifold by a numerical minimization of the effective action (1.49). It is simple enough to be applied in such a parameter scan. Further, within the ensemble of simulations, we have identified a robust signal using the next-to-leading order plane.

Also here, the work presented is primarily done by me. The used simulation code, the Nanosystem simulation library (NSL) [10], is developed by me with additional help by the contributors seen on the github page. Further, the analysis package [11] is primarily developed by me, with extensive help, in debugging and testing, from Dr. Giovanni Pederiva. The next-to-leading order planes where calculated with a python code developed in part by Cristoph Gäntgen and adapted by me. The other co-authors contributed through discussions and in running simulations.
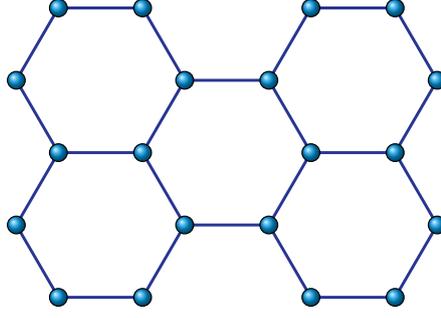
Figure 4.1: Graphical representation of the perylene molecule. The sites represent carbon-ions, while links indicate allowed hopping. External hydrogen atoms are not drawn.

4.1 Introduction

The perylene molecule $C_{20}H_{12}$, pictured in Fig. 4.1, has attracted great interest in various technological applications, ranging from organic semiconductors [26, 107], organic light emitting diodes (OLEDs) [99], to organic solar cells [100, 108, 109]. As it is a polycyclic aromatic hydrocarbon, it is also of great interest to astronomy; perylene and its derivatives have been found in interstellar gases and nebulae [110, 101, 102].

The ionization energy and electron affinity of perylene is well studied experimentally [111, 112]. Kinetic Monte Carlo simulations have also been conducted involving ensembles of perylene molecules, see e.g. [113, 114]. Theoretical studies of the electronic structure of perylene have been performed using various methods, for example density functional theory (DFT) [110, 115] and DMRG [96].

In derivatives of perylene the π orbitals of the sp^2-hybridized valence orbitals will not be half-filled; additional bonded groups may supply or draw away electrons. However, to our knowledge, little is theoretically known about the electronic structure of a single doped perylene molecule. We therefore model perylene's π electrons using the Hubbard model and perform ab-intio grand-canonical Monte Carlo simulations to map the single-electron spectrum as a function of the electron chemical potential μ. We describe this model in Sec. 4.2.1.

We describe our computational approach in Sec. 4.2.2. In particular, at non-zero μ our system is not half-filled and our simulations are afflicted by a numerical sign problem. We briefly describe the issue and how we leverage recent developments to nevertheless get reliable statistical estimates [67, 59, 116, 69].

We measure the global charge and single-particle (and single-hole) euclidean-time correlation functions from which we extract energy spectra. In section 4.3 we explain how this analysis is performed but relegate many details to Appendix 4.A and further results to Appendix 4.B. Finally, we summarize our findings in section 4.4.

4.2 Formalism

4.2.1 Modelling Perylene

Perylene consists of sp^2-hybridized carbon atoms arranged in five hexagons [95, 97], giving N_x=20 ions as shown in figure 4.1, and twelve hydrogen atoms bonded to the carbons on the boundary (which are not shown in fig. 4.1). The hybridized nature of the carbon bonds allows the valence π electrons to hop along the bonds. We model the kinematics and interactions of these π-electrons with the Hubbard model

$$\mathcal{H}\left[\kappa, U, \mu\right] = -\kappa \sum_{\langle x,y \rangle \in X} \left(p_x^\dagger p_y - h_x^\dagger h_y \right) + \frac{U}{2} \sum_{x \in X} q_x^2 - \mu \sum_{x \in X} q_x. \tag{4.1}$$

The hopping strength κ (which we take to be bond-independent) is the amplitude for a free electron to traverse the bond between nearest neighbors $\langle x, y \rangle$. We work in the particle/hole basis for computational reasons [19]; the p_x (h_x) represents a particle (hole) annihilation operator. We denote the collection of ions by X. The strength of interaction depends on the charge per site $q_x = h_x^\dagger h_x - p_x^\dagger p_x$ (so that particles represent electrons with negative electric charge), and is controlled by the onsite term U; a more realistic two-body interaction $\sum_{xy} q_x V_{xy} q_y$ can be easily incorporated into our simulations.

Typical applications of perylene involve attaching additional chemical structures to a perylene core [117, 118]. To model the electrons in these chemical derivatives in our simulations, we apply a homogeneous effective chemical potential μ coupling to the total system charge. For simplicity, we will provide all physical quantities in units of the hopping strength, i.e. U/κ, μ/κ, E/κ, etc. and in what follows, we will express these quantities already rescaled by κ. Following [96], we can reintroduce physical units setting $\kappa = 2.4\,\text{eV}$.

The point symmetry group of perylene is typically identified as D_{2h}. Our Hamiltonian (A.1), however, treats the ions as a fixed graph with no knowledge of its three-dimensional embedding, and we can split the symmetry into the dihedral group D_2 and a $\mathbb{Z}_2$ whose only action is to flip spin components (which amounts to an exchange of particles and holes). Hamiltonian eigenstates will have definite spin and will transform in the A, B_1, B_2, and B_3 representations of D_2, which are all one-dimensional.

We can perform a basis transformation of the 20 single-particle position-space operators. The vector space defined on the 20 sites can be decomposed into invariant subspaces on which the action of the D_2 symmetries act irreducibly as A, B_1, B_2, and B_3; in a slight but common abuse of language we identify these invariant subspaces as the irreps themselves. The irreps have multiplicity 6, 4, 6, and 4, respectively.

We can arrange for this transformation to diagonalize the hopping matrix $K = \kappa \delta_{\langle x, y \rangle}$ These operators are shown in detail in Appendix 4.A.4; each operator has definite irrep and tight-binding energy ϵ. In the non-interacting $U = 0$ case these irreducible operators carry definite energy and satisfy $[H, p_{\Lambda_i}^\dagger] = \epsilon_{\Lambda_i} p_{\Lambda_i}^\dagger$ where the state is labelled by irrep Λ and an index i. The same transformation can be made to the holes; the only difference arises from the sign of the hopping term for the holes in the Hamiltonian (A.1). Some operators have positive tight-binding energy and others have negative tight-binding energy; in the non-interacting case the global ground state consists of every negative-energy operator applied to the Fock vacuum.

4.2.2 Simulation Methods

We compute observables O expressed through the thermal trace over all Fock space states,

$$\langle O \rangle = \frac{1}{\mathcal{Z}} \text{Tr}\left\{ O e^{-\beta \mathcal{H}} \right\}. \tag{4.2}$$

Here the partition function $\mathcal{Z} = \text{Tr}\left\{ e^{-\beta \mathcal{H}} \right\}$ and $\beta = 1/T$ is the inverse temperature in natural units, $c = k_B = \hbar = 1$. We Trotterize β into N_t timeslices each separated by the temporal lattice spacing $\delta = \beta/N_t$. We introduce a continuous auxiliary field Φ on every site of the spacetime lattice via a Hubbard-Stratonovich transformation [24, 119, 120, 32] $\Phi = (\Phi_{tx}) \in \mathbb{R}^{|\Lambda|}$, with indices on the spacetime lattice $\Lambda = [0, N_t - 1] \otimes X$. Exactly integrating out the fermions transforms our problem from a discrete sum over Fock states into a path integral [29, 66, 30, 31, 25, 61],

$$\langle O \rangle = \frac{1}{\mathcal{Z}} \int \mathcal{D}[\Phi] \, e^{-S[\Phi]} O[\Phi] \tag{4.3}$$

where the action S is

$$S[\Phi \,|\, \kappa, U, \mu] = \frac{\Phi^2}{2\delta U} - \log \det\{M[\Phi \,|\, \kappa, \mu]\} - \log \det\{M[-\Phi \,|\, -\kappa, -\mu]\}, \tag{4.4}$$

and the Gaussian piece can be replaced by $1/2\, \Phi(\delta V)^{-1}\Phi$ for a more generic interaction, as long as the interaction matrix V_{xy} is positive definite. The fermion matrices are in the exponential discretization [19]

$$M\left[\Phi\mid K,\mu\right]_{x't';xt} = \delta_{x'x}\delta_{t't} - \left(e^{\delta(K-\mu)}\right)_{x'x} e^{+i\Phi_{xt}}\mathcal{B}_{t'}\delta_{t'(t+1)} \tag{4.5}$$

where $\mathcal{B}$ encodes the anti-periodic boundary conditions in time. We perform the path integral stochastically using the Hybrid/Hamilton Monte Carlo (HMC) algorithm [77].

At finite chemical potential the fermionic part of the action S can become complex, and removes any ergodicity problem [19]. However, it also introduces the so-called 'sign problem' since e^{-S} can oscillate. A severe sign problem ultimately results in unreliable statistical estimates of observables with finite statistics.

Complex actions and integrand oscillations can arise across a wide set of computational models and approaches, ranging across ϕ^4 theory [56, 80], topological (Chern-Simons) models [58], molecular systems [121] and lattice QCD [72, 73], for example. In recent years there has been a great push to leverage contour deformation to mitigate the sign problem in all these theories. In addition to trying to deform the contour integration onto Lefschetz thimbles [55, 75, 54, 68, 82], machine learning methods [54, 65, 122, 123, 23] can often but not always [124] locate integration contours with much more modest problems. Related deformations to complex Langevin methods [20, 76] are also undergoing rapid development. Moreover, the signal-to-noise problem present for many observables in Markov Chain Monte Carlo simulations can be improved with a similar approach [63, 64].

Leveraging experience gained while developing these methods for the Hubbard model [19, 61, 1, 5, 7, 8, 2], we perform a simple and cost-efficient transformation by incorporating a spacetime constant imaginary shift ϕ_c

$$\Psi(\Phi) = \Phi + i\,\phi_c. \tag{4.6}$$

Such a shift represents an integration manifold in the complex plane that is parallel to the real plane. For this investigation we utilize the next-to-leading order (NLO) plane [7], whereby ϕ_c is determined by including quantum (thermal) corrections to the saddle-point approximation of S. We briefly motivate this method in Appendix 4.C. Even with this shift in the integration contour the action remains complex and we perform HMC changing the real part of Φ according to the real part of the HMC force, accepting proposed changes according to the real part of the action, and reweighting with the imaginary part of the action as described in Appendix 4.A.1.

4.3   Analysis

The goal of this investigation is to assess the single particle spectrum in relation to the system's total charge, as a measure of doping. These two quantities can be obtained by calculating the euclidean time single particle (p) and hole (h) correlators

$$
\begin{aligned}
C^{\mathrm{sp}}_{x,y}(\tau) &= \left\langle p_x(\tau)p_y^\dagger(0)\right\rangle = \left\langle M^{-1}_{x,\tau;y,0}\left[\,\Phi\mid \kappa,\,\mu\right]\right\rangle, \\
C^{\mathrm{sh}}_{x,y}(\tau) &= \left\langle h_x(\tau)h_y^\dagger(0)\right\rangle = \left\langle M^{-1}_{x,\tau;y,0}\left[\text{-}\Phi\mid\text{-}\kappa,\text{-}\mu\right]\right\rangle,
\end{aligned}
\tag{4.7}
$$

which we can analyze using the standard spectral decomposition (Appendix 4.A.3).

After averaging particles and time-reversed holes we have a 20×20 matrix of correlators for each ensemble. The irreducible representation is a good quantum number, allowing us to block-diagonalize to four small correlators, one for each A (6×6), B_1 (4×4), B_2 (6×6), and B_3 (4×4) using the irreducible single-particle operators. Interactions can mix the operators within an irrep and we variationally extract the six or four interacting energy levels closest to the fully interacting ground state as explained in Appendix 4.A.4.

The chemical potential μ controls the total charge of the system. To quantify its effect, we compute the total system charge by

$$
\begin{aligned}
\langle Q \rangle = \sum_{x \in X} \langle q_x \rangle &= \sum_x \left(\langle h_x^\dagger h_x - \langle p_x^\dagger p_x \rangle \rangle \right) \\
&= \sum_x \left(\langle p_x p_x^\dagger \rangle - \langle h_x h_x^\dagger \rangle \right) \\
&= \sum_x \left(C_{x,x}^{\mathrm{sp}}(0) - C_{x,x}^{\mathrm{sh}}(0) \right) ,
\end{aligned}
\tag{4.8}
$$

as a function of μ.

In the non-interacting case we can compute the total charge

$$
\langle Q \rangle \big|_{U=0} = 2 \sum_{\Lambda_i} \frac{1}{e^{-\beta(\epsilon_{\Lambda_i} + \mu)} + 1} - N_x,
\tag{4.9}
$$

$$
\lim_{\beta \to \infty} \langle Q \rangle \big|_{U=0} = 2 \sum_{\Lambda_i} \Theta \left(\epsilon_{\Lambda_i} + \mu \right) - N_x.
\tag{4.10}
$$

The factor of two comes from the spin degeneracy and the subtraction by N_x ensures that $Q = 0$ when $\mu = 0$.

At non-zero interaction, $U \neq 0$, observables are computed using the NLO-plane HMC algorithm as discussed in the previous section. This alleviates the sign problem sufficiently to allow us to extract statistically meaningful quantities. Further details on the analysis steps can be found in appendix 4.A.

We perform our studies using an on-site interaction of $U = 2$. This provides us with an initial qualitative behavior of perylene's charge Q as a function of μ. In the future we aim to tune this on-site coupling to a more realistic value or use a more realistic two-body interaction.

To access different total charges, we scan over the chemical potential $\mu = 0, 0.1, \ldots, 1.1$. This choice is inspired by the non-interacting charges discussed in section 4.3.2. We control the temporal continuum limit using three time discretizations $N_t = 32, 64, 96$ and study the temperature dependence with $\beta = 4, 6, 8$. For each parameter combination we measure a total of $N_{\mathrm{cfg}} = 10\,000$ configurations.

4.3.1 Statistical Power

Before discussing the analysis of the physical observables, i.e. (4.7) and (4.8), it is important to map out the severity of the sign problem. A typical measure is the absolute average phase, called the statistical power,

$$
|\langle \Sigma \rangle| = \left| \langle e^{-i \, \mathrm{Im}\{S[\Phi]\}} \rangle \right|.
\tag{4.11}
$$

A value of 1 for the statistical power implies no sign problem, whereas a value of 0 represents the most severe sign problem. One can further relate the statistical power to an effective number of configurations $N_{\mathrm{cfg}}^{\mathrm{eff}} \propto |\langle \Sigma \rangle|^2 N_{\mathrm{cfg}}$ [20]; when the statistical power is small each configuration is worth less. The average phase appears in the denominator when reweighting (Appendix 4.A.1) and, therefore, for small, hard-to-estimate statistical powers, stochastic estimates of observables become unreliable.

In figure 4.2 we show the statistical power as a function of μ plotted for the various β and N_t. With $N_{\mathrm{cfg}} = 10\,000$ configurations, simulations with $|\langle \Sigma \rangle| \lesssim 0.1$ become unreliable. We emphasize that without the contour deformation (4.6) the statistical power is indistinguishable from 0 for almost all of the $\mu \neq 0$ ensembles shown.

We observe that the total system charge (Sec. 4.3.2) is less susceptible to statistical noise which allows us to access it over all considered chemical potentials. In contrast the single particle energy spectrum (Sec. 4.3.5) is more susceptible to the noise resulting in significant uncertainty

at $\beta = 8$ with $\mu = 0.9, 1$. At $\mu = 1.1$ more data is required to reliably estimate the larger energies. Consequently, we remove this point from the analysis.
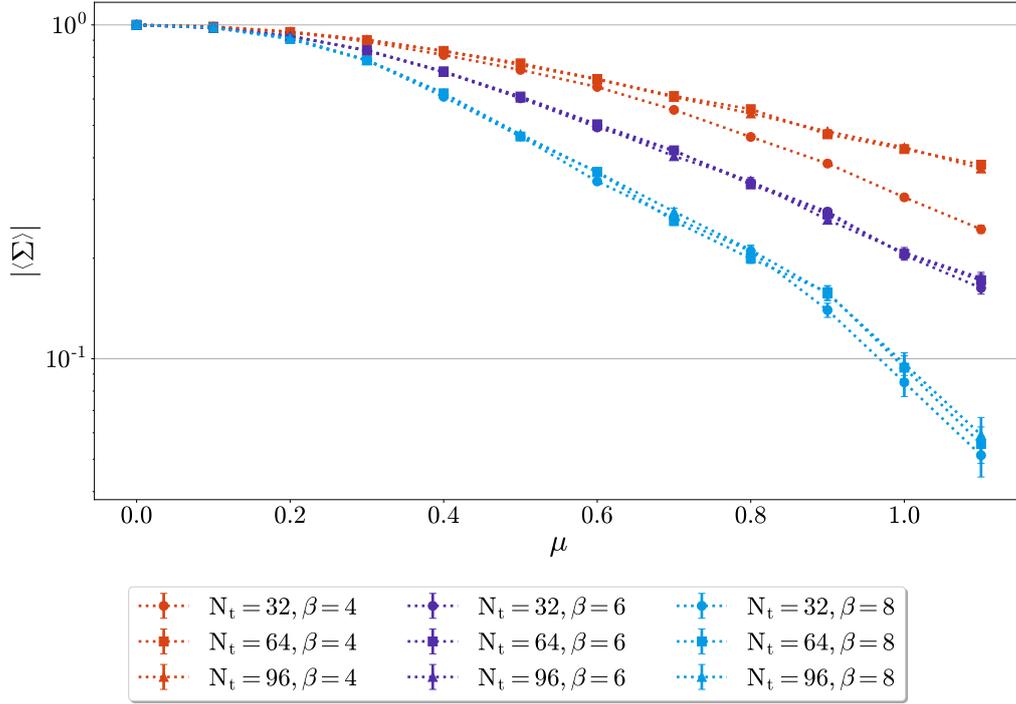
Figure 4.2: Statistical power $|\langle\Sigma\rangle|$ as a function of the chemical potential μ. With the given amount of configurations, beyond $|\langle\Sigma\rangle| \lessapprox 0.1$ simulations are unreliable.

4.3.2 Total System Charge
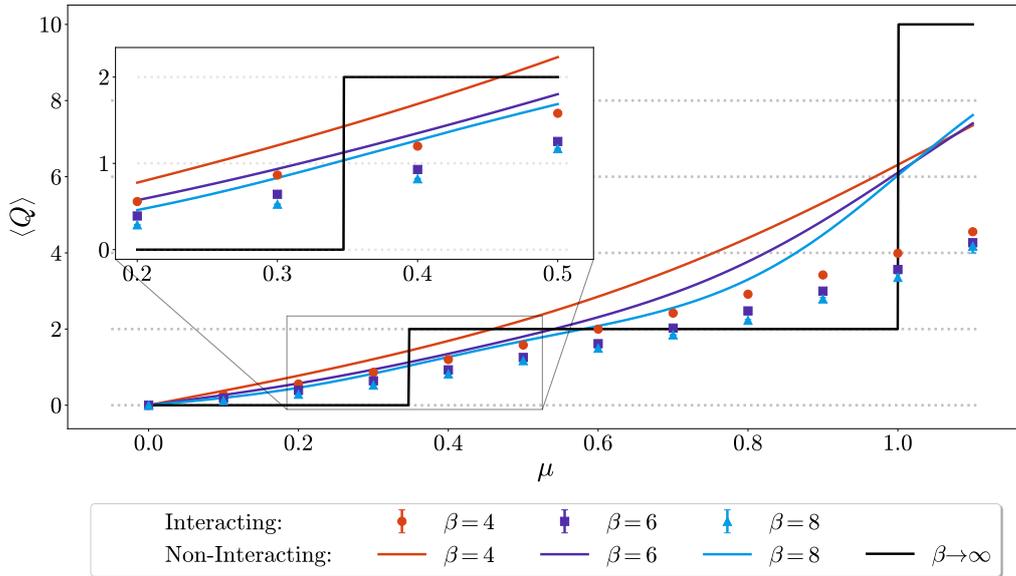
Figure 4.3: Plot of the total system charge as a function of the chemical potential μ.

In figure 4.3 we show the charge as a function of μ. Solid lines are exact non-interacting $U = 0$ results, while the continuum-extrapolated ($\delta \to 0$) total system charge measured with $U = 2$ is shown as points with uncertainties. As shown in appendix 4.A.7 our data is close to the continuum

limit and we simply fit a constant to the charge at all three N_t values.

Focusing on the non-interacting result, colored lines are at the simulated β while the black line represents the zero-temperature limit. For the latter we observe a sudden jump of $\Delta Q = 2$ at $\mu \approx 0.347$ which corresponds to the smallest single particle energy $\Delta E_{U=0}^{B_3^3}(\mu = 0)$; the jump corresponds to two electrons moving out of the Fermi sea. The difference in charge must be a multiple of two due to the spin-degeneracy preserved in the Hubbard model. In the non-interacting case a further jump of $\Delta Q = 8$ appears at $\mu = 1$ corresponding to the next (accidentally-quadruply-degenerate) single particle energy. Corresponding single particle energies can be found in the first panel of figure 4.8.

Finite temperature washes out the step function (4.9) and we draw colored solid lines for each temperature we simulated. They cross the $\Delta Q = 2$ threshold necessarily at higher chemical potentials due to finite temperature effects. Furthermore, at finite temperature $Q \neq 0$ states are partly populated and we can cross $Q = 1$ below the free zero-temperature single-particle threshold.

The circles, squares and triangles in figure 4.3 display the continuum limit of the charge at finite temperature, $\beta = 4, 6, 8$ respectively. The temperatures are too hot to identify a clear charge jump, however, the $\beta = 6, 8$ data go through $Q = 1$ between $\mu = 0.4$ and 0.5, later than the free system.

Comparing the finite temperature interacting and non-interacting results shows a growing deviation as we increase μ. Already, for the first charge jump a significant change is deduced suggesting a noticeable influence from the interactions. Furthermore, as we will see in section 4.3.5, the 4-fold degeneracy around $E \sim 1$ splits, and we expect the jump of $\Delta Q = 8$ to break into jumps of size $\Delta Q = 2$. A final assessment on the importance of the interaction in this molecule, however, cannot be made, as only one non-physical, interaction value is considered.

Furthermore, since a typical level of doping is expected to be only a few elemental charges [98, 125], we argue that the NLO-plane HMC provides an acceptable signal at values of μ in the relevant range for perylene.

4.3.3 Extracting Energies

Each ensemble, fixed by a choice for N_t, β, and μ, results in 20 correlators; a total of 1980 correlators need to be analyzed. Using the fitting routine described in Appendix 4.A, we perform about 30 to 100 fits (depending on the fit intervals and the minimum of the correlator) with either two or three exponential terms in the model for the central value and for each of the N_{bst} bootstrap samples. With $N_{bst} = 500$ this results in $O(10^8)$ fits. This sheer number emphasizes that an automatic fitting procedure with well formulated criteria is needed. In this section we discuss a selection of correlators and how their corresponding energies are extracted. We focus in particular on the finest lattice spacing ($N_t = 96$) and the lowest temperature ($\beta = 8$).

As discussed in Appendix 4.A.3 the single particle spectrum contains positive and negative energies and the spectral decomposition can be split into increasing and decreasing exponentials. This motivates the fit model

$$C_{\Lambda_i}(\tau) = z_0^L e^{-E_0^L \tau} + z_0^R e^{E_0^R(\tau-\beta)} + \sum_{n>0}^{N_L} z_n^L e^{-\left(\Delta E_n^L + E_{n-1}^L\right)\tau} + \sum_{n>0}^{N_R} z_n^R e^{\left(\Delta E_n^R + E_{n-1}^R\right)(\tau-\beta)} \quad (4.12)$$

where the L and R labels indicate whether the contribution is large at small or large τ and we have dropped the state label on the fit parameters. Notice that the parameters $E_n^{L/R}$ and the respective splittings $\Delta E_n^{L/R} = E_n^{L/R} - E_{n-1}^{L/R}$ are positive. Thus accessing the desired energy requires us to identify the dominant contribution and assign

$$E_{U=2}^{\Lambda_i} = E_0^L \text{ or } - E_0^R . \quad (4.13)$$

For more details please refer to appendix 4.A.5.
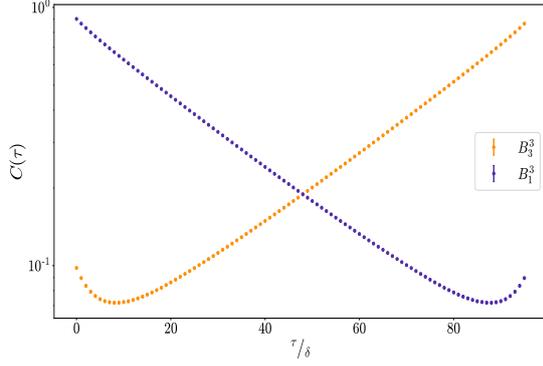
Figure 4.4a: Correlators corresponding to states B_3^3 & B_1^3, orange and blue respectively. These are estimated at $N_t = 96$, $\beta = 8$, $U = 2$, $\mu = 0$ and correspond to the smallest, in magnitude, energy of the system. Uncertainties are less then 1% making them hard to spot.
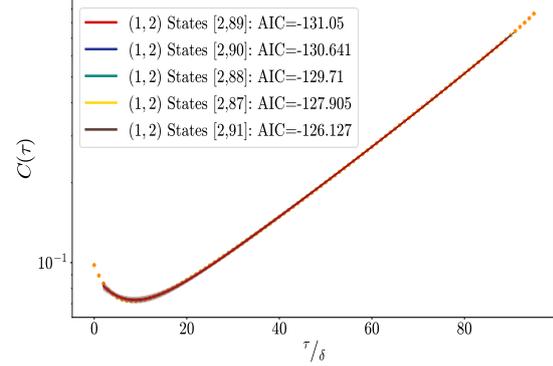
Figure 4.4b: Best B_3^3 fits. The data points are the same as in figure 4.4a. For each fit, two confidence bands are plotted corresponding to one and two σ. These best fits are performed according to the model (4.12) with ($N_L = 1$, $N_R = 2$). No $N_R = 1$ fit is in the best five. The fit range is indicated in the square brackets expressing values of τ/δ comprising almost the entire correlator.

$\mu = 0$

The smallest energy, in magnitude, is most interesting as it moves across zero for finite chemical potential first, indicating the previously discussed charge jump. These energies come from the state B_3^3 (negative energy) and B_1^3 (positive energy).

The B_3^3 and B_1^3 correlators at $\mu = 0$ are displayed in figure 4.4a. The uncertainties at each time point are less then 1% which results from relatively high statistics $\sigma \sim O(1/\sqrt{10000})$ and the fact that its decay is relatively mild. Especially for larger energy correlators we find a signal-to-noise problem around the minimal point. As the energies of the B_3^3 and B_1^3 correlators differ only in sign, we find them equal up to time reversal. Furthermore, on a log scale they appear extremely straight for a large range of euclidean time τ indicating little excited state contamination.

In figure 4.4b the 5 best fits are plotted on top of the correlator. The data points represent the correlator, the solid lines are the fits colour-coded as indicated in the legend, and the bands indicate the one- and two-σ confidence interval on the fit. All these fits have two exponentials on the right while the left side has one exponential. We also performed fits with only one exponential on the right but none are among the 5 best fits shown here. Visually all these fits are extremely close to the data points; quantitatively the $\chi^2/\text{dof} \sim 1$ as desired for good fits. Appendix 4.A.6 explains how we model average fits. Furthermore, the best fit resulting over a fitting range of $\tau/\delta \in [2, 89]$ with $\chi^2/\text{dof} = 0.53$. Its result is displayed in table 4.1. From here we see that excited states are clearly distinguished providing additional evidence for a reliable estimate.

E_0^L	$=$	$1.749(63)$	z_0^L	$=$	$0.04277(84)$
E_0^R	$=$	$0.3228(93)$	z_0^R	$=$	$0.590(56)$
ΔE_1^R	$=$	$0.270(29)$	z_1^R	$=$	$0.290(55)$.

Table 4.1: Best fit results of the B_3^3 correlator displayed in figure 4.4b. Uncertainties are determined through bootstrap while central values come from a fit to the central values of the data.

We can further assess the stability of the fitting procedure by considering the overview plots in figure 4.5. The main body for each figure shows the value of the fit parameter as a function of the model probability given the data [126, 127, 128]

$$p(m|D) \sim e^{-\frac{1}{2}\text{AIC}}, \tag{4.14}$$

55

where AIC is the Akaike information criterion, as explained in Appendix 4.A.6. A model m is defined by the number of exponentials (N_L, N_R) in the fit function (4.12) and the range of euclidean time it is evaluated on. In figure 4.5 the $(1, 1)$ and $(1, 2)$ state fits are plotted as circles and pluses, respectively. These points represent the central value fit, uncertainties are not drawn. We find the correlator to be predominantly increasing, resulting in the choice of varying N_R and identifying the lowest energy to be negative. We highlight the $p(m|D)$-weighted (model) average, (4.32), of each parameter with a solid line and the uncertainty as a band. This uncertainty is obtained by the standard deviation of the model average over all bootstrap samples. For the E_0^R the absolute value of the non-interacting energy is added as a grey dashed line to provide a reference. Attached to the ordinate and abscissa are the counts of the parameters and model weights (histograms). They visualize the distribution of the fit results. The total number of fits done is indicated in the lower right corner. This number is naturally smaller for the parameters only appearing in the two state fits.

Overall, we find great stability in these fits, as evidenced by the string of points converging towards larger weights. The two bands in all figures originate from the two allowed fit interval starting points at $\tau/\delta = 1, 2$ for the fits. As the AIC penalizes additional parameters, we find significant support for the $(1, 2)$ state fits; their respective mode is strongly correlated with the mode of the weights. The fact that the best fits almost span over the entire abscissa strengthens this even further.

Finally, this fitting procedure results in the model averaged energy

$$E_{U=2}^{B_3^3} (N_t = 96, \beta = 8 \,|\, \mu = 0) = -0.3230(64).$$ (4.15)

Transition of the Smallest Energy at $\mu \neq 0$

The interacting energies change with μ, and sometimes a state's energy changes sign. This happens, for instance, between $\mu = 0.4$ and 0.5, where the B_3^3 state's energy crosses 0. To illustrate the effect of the chemical potential on the correlator, figure 4.6a again shows the two states B_3^3 and B_1^3; between $\mu = 0.4$ (the upper panel) and $\mu = 0.5$ (lower panel) the B_3^3 correlator goes from predominantly decreasing to predominantly increasing, indicating an energy crossing 0. Even at these chemical potentials we find a great resolution accounting for a statistical power of $|\langle \Sigma \rangle| = 0.6228(46), 0.4707(59)$ at $\mu = 0.4, 0.5$ respectively. We emphasize that without alleviating the sign problem with a contour deformation these correlators are overwhelmed by noise and no results can be extracted.

In figure 4.6b, similar plots for the 5 best fits to the B_3^3 correlator are shown. As in figure 4.6a, the upper panel shows $\mu = 0.4$, while the lower panel shows $\mu = 0.5$. The best of these fits have $\chi^2/\text{dof} = 0.11, 0.075$. The overview plots exhibit the same nice features as in the $\mu = 0$ case and we omit them for concision.

For these two fits we find model average energies,

$$E_{U=2}^{B_3^3} (N_t = 96, \beta = 8 \,|\, \mu = 0.4) = -0.0480(37),$$ (4.16)

$$E_{U=2}^{B_3^3} (N_t = 96, \beta = 8 \,|\, \mu = 0.5) = +0.0427(73).$$ (4.17)

4.3.4 Continuum Limit

To remove the systematic errors introduced by discretizing the thermal trace we must perform a continuum limit $\delta \to 0$. Given our data, at each β we can fit a constant, as shown for the B_3^3 state in figure 4.7. The inverse temperature β increases across the columns and the chemical potential increases down the rows. In each row the ordinate maintains the same scale to provide a rough idea of the β dependence. A triangle at $\delta = 0$ indicates the continuum value; a corresponding solid line is put to guide the eye toward larger δ.

The legend gives the χ^2/dof for the constant fit. All states give values between $\chi^2/\text{dof} = 3.5 \times 10^{-4}$ and 0.8. Overall, the residuals are significantly smaller than one would expect for an ideal fit
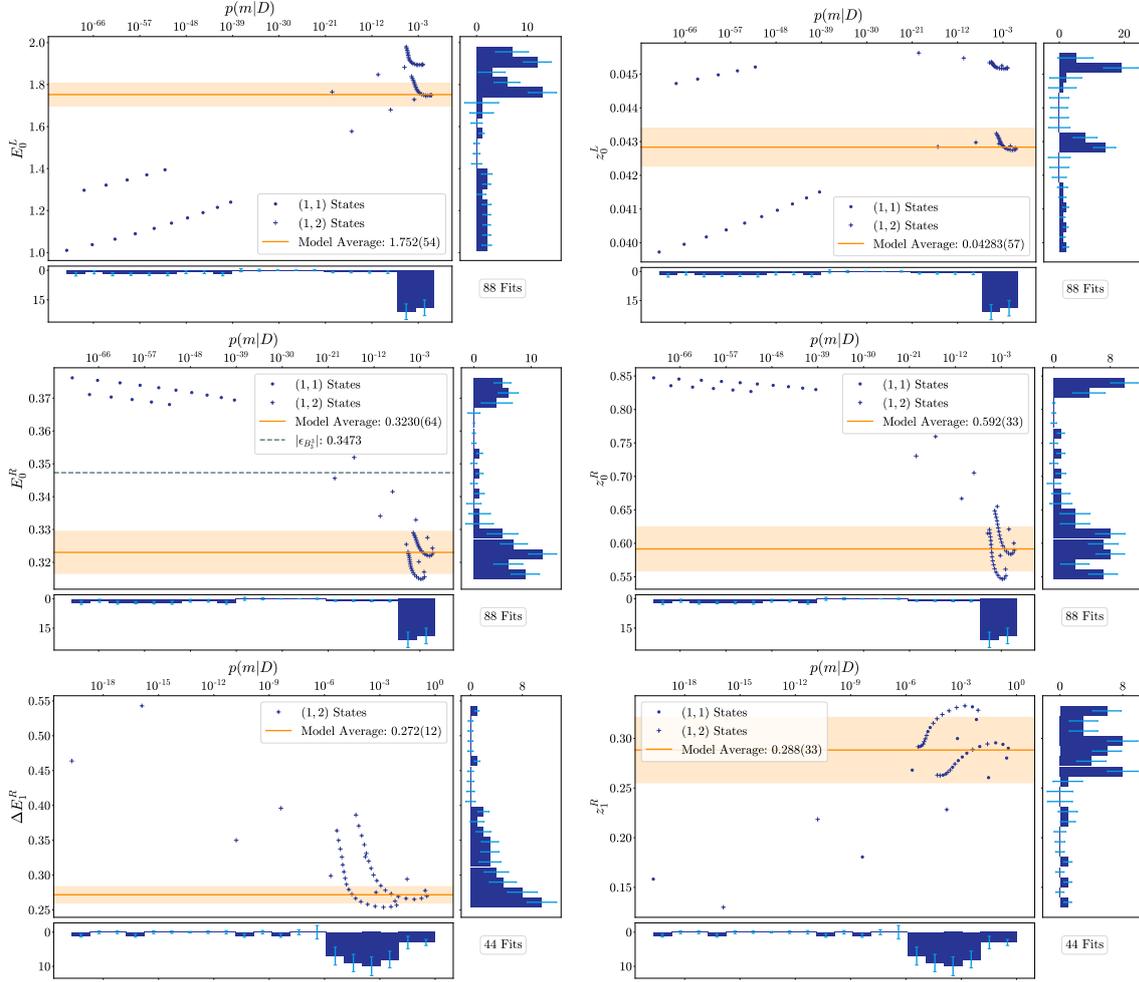
Figure 4.5: Fit overview plots for a fit to the B_3^3 correlator at $N_t = 96$, $\beta = 8$, $U = 2$, $\mu = 0$. A subfigure is dedicated for each parameter in the fit model (4.12) as a function of the model weight $p(m|D) \sim \exp(-1/2\text{AIC})$. One-state fits, $(N_L = 1, N_R = 1)$, are indicated with circles, while two-state fits, $(N_L = 1, N_R = 2)$, are plotted with pluses. The model average is indicated through a solid line with adjacent uncertainty determined by the standard deviation of the model average on each bootstrap sample. Attached to the axes are counts of the fit results (unweighted) and the model weights. The correlation of the mode of the fit results with the mode of the weights indicate the support for the two-state fits. Uncertainties are only displayed on the counts, computed by bootstrapping the heights on fixed bin widths. Fits with AIC > 200 are not shown. Finally the total number of fits is shown in the lower right corner, with less fits for parameters only available in the two state fits.
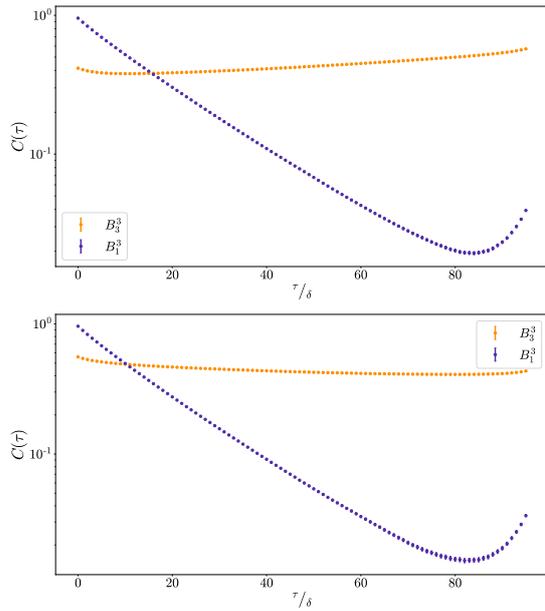
Figure 4.6a: Correlators corresponding to states B_3^3 (orange) and B_1^3 (blue) estimated with $N_t = 96$, $\beta = 8$, $U = 2$. While figure 4.4a shows $\mu = 0$, the upper panel shows $\mu = 0.4$ and the lower panel $\mu = 0.5$. The transition of the smallest negative energy to a positive energy happens in between these two values. The trend of both correlators is toward a more negative slope with increasing chemical potential, as expected given our sign convention for μ.
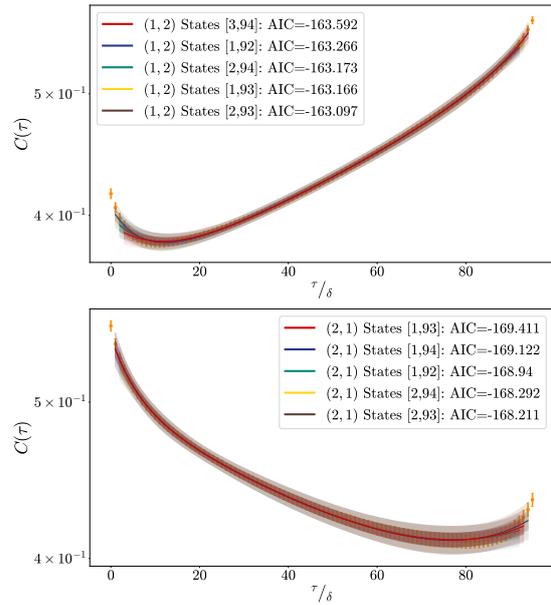
Figure 4.6b: Best B_3^3 Fits. The data points are the same as in figure 4.6a, at again $\mu = 0.4$ (upper) and $\mu = 0.5$ (lower). Similarly, each fit is plotted as a solid line with two confidence bands corresponding to one and two σ. Further, the best fits are again with two states ($N_L = 1$, $N_R = 2$) and ($N_L = 2$, $N_R = 1$) respectively. The fit range is indicated in the square brackets expressing values of τ/δ comprising almost the entire correlator. The transition from negative to positive energy can be seen better on this scale.

$\chi^2/\text{dof} \approx 1$. In particular the very small χ^2-values are governed by the increased uncertainties at larger chemical potentials. At this point we want to emphasize that the purely statistical uncertainties on the best fits are significantly smaller. However, due to the bootstrap over model averages we include systematics, from the choice of fit model, in the uncertainties. This conservative error estimation allows us to be very confident about the correctness of our results within the provided uncertainty range.

We find that all ensemble's extrapolations are extremely flat, showing little dependence on the lattice spacing at the chosen parameters. Linear contributions are not well-supported by the data, see Appendix 4.A.7 for a thorough discussion. We discard the spectrum at $\mu = 1.1$ as the noise is too large to extract the higher energies for $\beta = 8$ reliably without additional samples.

4.3.5 Spectrum

We can now collect all continuum energies and plot them as a function of chemical potential. We present this result in two ways, first with the barcode plot in figure 4.8 that provides an overview on how the spectrum behaves as function of μ. Each panel in this figure details 20 single-particle states in $\beta = 8$ spectrum at fixed chemical potential. The first two panels offer a comparison between non-interacting and interacting spectra at $\mu = 0$; each shows the expected symmetric spectrum, providing a check on the analysis. While the small energies are very close and the ends of the spectra differ more meaningfully, we can see that the interactions split the accidental quadruplets of states at $E = \pm 1$. Lower panels have increasing chemical potential and the energies grow with μ as expected. In particular, the least negative state B_3^3 moves closer and closer to zero, changing sign after $\mu = 0.4$ as expected from the $\langle Q \rangle = 1$ crossing in figure 4.3. Up to $\mu = 0.8$ the signal is good to resolve all energies with great precision. Starting at $\mu = 0.9$ the sign problem becomes prevalent, providing statistical powers smaller than $|\langle \Sigma \rangle| \leq 0.1547(75)$ resulting in significantly larger uncertainties. To map out the second big transition, expected after $\mu = 1$ from figure 4.3, more statistics are required. Appendix 4.B details the same plots for $\beta = 4, 6$ obeying a similar behaviour.

Second, figure 4.9 details the μ dependence for each state's energy, which makes it easier to compare to the non-interacting finite-μ result. In each panel the solid black line represents the non-interacting result, while the data points display the interacting result. For most states a significant divergence from the non-interacting result can be seen. As the chemical potential increases the behaviour of a given state is expected to change as the ground state changes. Indeed we observe slightly different slopes for all states after $\mu = 0.4$. This is more pronounced at larger β pointing towards a non-trivial zero temperature limit. Finally, the energy levels at $\beta = 8$ are detailed in table 4.2.
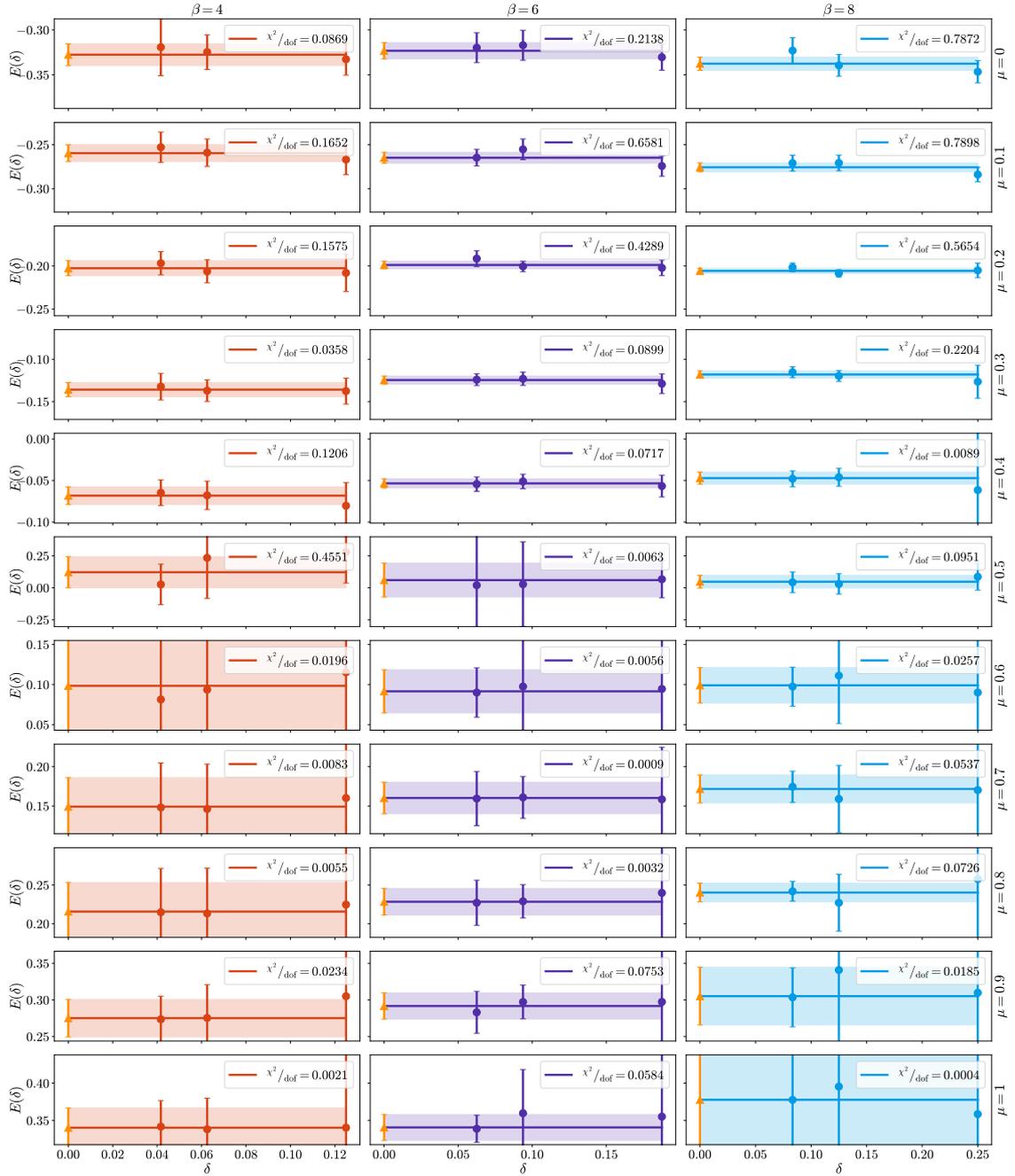
Figure 4.7: Continuum limit of the B_3^3 as a function of chemical potential (top to bottom) for the three available temperatures left to right. The scale of the ordinate is kept to provide a feeling on the β dependence. The title of each subplot shows the validity via a χ^2/dof. As these are just fits to a constant, smaller values stem from uncertainties thus all fits are very good.
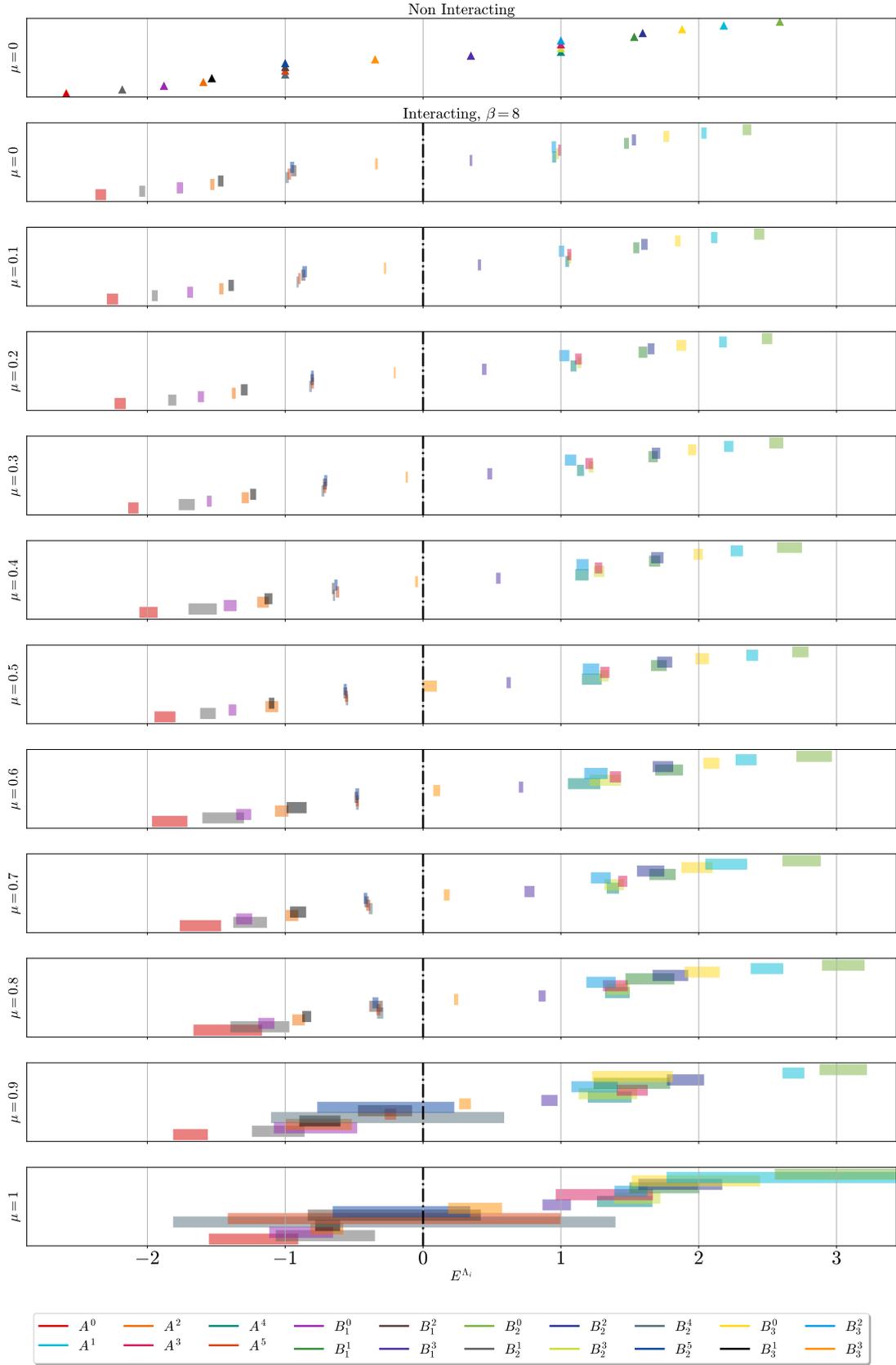
Figure 4.8: Spectrum plot at $\beta = 8$ shows the single particle spectrum as a function of chemical potential. The first panel indicates the non-interacting result at $\mu = 0$, all subsequent panels are at finite interaction, $U = 2$. Notably, the accidental degeneracy at $E_{U=0} = \pm 1$ is split into eight energy levels. Following the chemical potential, B_3^3 can be observed moving towards zero energy and transitioning out of the fermi sea past $\mu = 0.4$ as expected from the total system charge discussed in figure 4.3. A second transition is not resolved at available statistics.
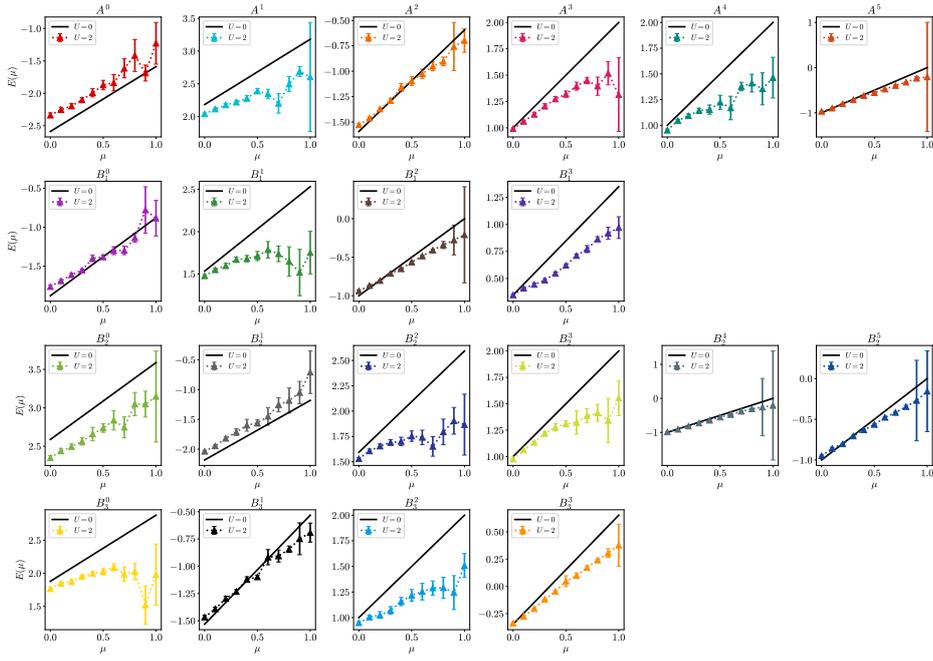
Figure 4.9: Single particle energy overview as a function of chemical potential at $\beta = 8$. Each panel shows the energy of a particular state. The non-interacting energy is provided as a solid black line while the data points are at $U = 2$. Difference to the non-interacting energy is more pronounced towards larger chemical potentials.

μ	A^0	A^1	A^2	A^3	A^4	A^5	B_1^0	B_1^1	B_1^2	B_1^3
0	-2.337(36)	2.038(16)	-1.529(12)	0.9917(85)	0.951(12)	-0.9696(92)	-1.763(20)	1.476(14)	-0.938(16)	0.3477(68)
0.1	-2.253(37)	2.112(19)	-1.463(13)	1.061(12)	1.046(11)	-0.8969(50)	-1.689(18)	1.547(19)	-0.869(12)	0.4090(86)
0.2	-2.197(38)	2.176(25)	-1.373(11)	1.127(21)	1.092(19)	-0.8010(69)	-1.611(20)	1.596(29)	-0.8041(92)	0.445(13)
0.3	-2.102(35)	2.218(31)	-1.289(23)	1.205(25)	1.143(23)	-0.7123(75)	-1.551(14)	1.669(31)	-0.712(11)	0.484(15)
0.4	-1.992(63)	2.276(41)	-1.160(39)	1.273(24)	1.152(45)	-0.6201(94)	-1.400(43)	1.680(38)	-0.6515(59)	0.546(13)
0.5	-1.872(73)	2.388(40)	-1.097(43)	1.320(30)	1.224(69)	-0.5555(82)	-1.383(24)	1.711(54)	-0.5635(99)	0.621(13)
0.6	-1.84(13)	2.344(73)	-1.025(46)	1.396(38)	1.17(11)	-0.4762(78)	-1.301(51)	1.785(98)	-0.4848(96)	0.711(13)
0.7	-1.61(15)	2.20(15)	-0.951(44)	1.449(30)	1.377(41)	-0.398(12)	-1.298(55)	1.737(93)	-0.409(13)	0.772(33)
0.8	-1.42(25)	2.50(12)	-0.903(42)	1.395(87)	1.410(87)	-0.323(12)	-1.137(56)	1.65(17)	-0.341(46)	0.864(22)
0.9	-1.69(12)	2.687(76)	-0.76(24)	1.52(11)	1.35(16)	-0.236(39)	-0.78(30)	1.52(27)	-0.28(19)	0.917(56)
1	-1.23(32)	2.60(83)	-0.70(12)	1.32(35)	1.46(20)	-0.2(1.2)	-0.88(23)	1.75(25)	-0.21(62)	0.97(10)
	B_2^0	B_2^1	B_2^2	B_2^3	B_2^4	B_2^5	B_3^0	B_3^1	B_3^2	B_3^3
0	2.350(28)	-2.037(18)	1.530(12)	0.9758(77)	-0.9870(96)	-0.949(12)	1.766(18)	-1.467(17)	0.949(13)	-0.3376(74)
0.1	2.438(34)	-1.946(18)	1.606(20)	1.0622(78)	-0.9099(44)	-0.859(14)	1.848(17)	-1.392(16)	1.005(17)	-0.2757(50)
0.2	2.496(35)	-1.819(26)	1.655(22)	1.134(12)	-0.8163(65)	-0.8027(72)	1.873(32)	-1.297(21)	1.025(33)	-0.2057(30)
0.3	2.563(47)	-1.714(55)	1.691(28)	1.219(15)	-0.7259(76)	-0.7061(85)	1.953(26)	-1.233(18)	1.070(37)	-0.1180(44)
0.4	2.658(87)	-1.599(99)	1.700(42)	1.277(35)	-0.6455(53)	-0.6311(86)	1.997(30)	-1.121(25)	1.157(41)	-0.0472(72)
0.5	2.737(56)	-1.561(53)	1.753(51)	1.311(30)	-0.5514(59)	-0.5643(83)	2.025(46)	-1.099(17)	1.218(57)	0.048(50)
0.6	2.84(13)	-1.45(15)	1.740(71)	1.32(11)	-0.4763(68)	-0.477(11)	2.092(56)	-0.917(70)	1.254(81)	0.099(22)
0.7	2.75(14)	-1.25(12)	1.651(96)	1.386(67)	-0.380(10)	-0.417(10)	1.99(11)	-0.906(55)	1.289(69)	0.172(18)
0.8	3.05(15)	-1.18(21)	1.79(13)	1.415(81)	-0.311(19)	-0.346(18)	2.03(12)	-0.844(29)	1.29(10)	0.240(12)
0.9	3.05(17)	-1.05(19)	1.90(13)	1.34(21)	-0.26(84)	-0.27(49)	1.52(29)	-0.75(15)	1.24(17)	0.305(39)
1	3.15(59)	-0.71(36)	1.87(30)	1.55(17)	-0.2(1.6)	-0.16(50)	1.98(46)	-0.692(88)	1.51(12)	0.38(19)

Table 4.2: Values of the energy levels at $\beta = 8$. These numbers correspond to the squares or points displayed in 4.8 and 4.9 respectively.

4.4 Conclusions

In this work we have performed an initial Monte Carlo study of the electronic structure of a single doped perylene $C_{20}H_{12}$ molecule described with the Hubbard model. We treated discretization errors by simulating at three discretizations and performing a continuum limit extrapolation. The effect of temperature is studied qualitatively at three values. Central to this study is the scan over chemical potential starting at half filling ($\mu = 0$), including the first doping transition ($0.4 < \mu < 0.5$), and stretching further out to $\mu = 1.1$. We quantify the doping by calculating the total system charge, providing evidence for the position of the transition. We map out the low single particle energy spectrum at each chemical potential, backing the transition with a negative energy state moving out of the Fermi sea. Throughout all results, we find significant divergence from the non-interacting model. In particular, the point of transition moves to larger chemical potentials and an additional splitting of accidentally degenerate energy states emerges. For technological applications to perylene-derived molecules we can easily leverage a more accurate interaction. We also plan to compute charge-neutral excitations, responses to external electromagnetic sources, and to carefully study the cold regime.

Acknowledgments

4.A Analysis Details

In this appendix we describe in detail each step of the analysis.

4.A.1 Reweighting

When dealing with systems obeying a complex valued action a way to utilize Monte Carlo integration is reweighting. For this, the Markov Chain is generated by sampling according to the Boltzmann distribution originating from the real part of the action effectively treating the complex phase $e^{-i\,\text{Im}\{S\}}$ as part of the observable. In order to generate the intended observables the relation

$$\langle O \rangle = \frac{\langle O e^{-i\,\text{Im}\{S\}} \rangle_{\text{Re}\{S\}}}{\langle e^{-i\,\text{Im}\{S\}} \rangle_{\text{Re}\{S\}}}. \tag{4.18}$$

has to be evaluated. Under a bootstrap analysis each resample is evaluated in this way maintaining the correlations and fluctuations of the observables with the phase.

4.A.2 Autocorrelation

When estimating statistical uncertainty of observables, especially with bootstrap based analysis, the observables need to be statistically independent between configurations. This naively is not the case for Markov Chain algorithms. Yet, we can ensure statistical independence by various means for example by striding – only measuring on every n$^{\text{th}}$ trajectory with n big enough. A post-processing option is to evaluate the autocorrelation function

$$\Gamma_O(\nu) \propto \sum_{n=0}^{N_{\text{cfg}}-\nu} \left(O\left[\Phi_{n+\nu}\right] - \langle O \rangle \right) \left(O\left[\Phi_n\right] - \langle O \rangle \right)^* \tag{4.19}$$

normalized by $\Gamma_O(0)$, and estimating the integrated autocorrelation time [129, 130],

$$\tau_{\text{int}}^O = \frac{1}{2} + \sum_{\nu=1}^{M} \Gamma_O(\nu). \tag{4.20}$$

One can find the cut-off $M \ll N_{\text{cfg}}$ by searching for the smallest number such that $M \leq 10 \cdot \tau_{\text{int}}^O$ [129]. For the analysis discussed here we measure on every 10$^{\text{th}}$ trajectory and subsequently identify the largest autocorrelation time over all our considered observables (the set of correlators $C_{x,y}^{\text{sp}}(\tau)$). To ensure no observable is autocorrelated, we use this largest integrated autocorrelation as a stride between measurements resulting in $N_{\text{cfg}}^{\text{indep}} = N_{\text{cfg}}/2\max_O\{\tau_{\text{int}}^O\}$ independent samples[1]. We find that most of the time $\tau_{\text{int}}^O \approx 0.5$. In exceptional cases, we find $\tau_{\text{int}}^O \approx 1$.

4.A.3 Spectral Decomposition

By inserting complete sets of Hamiltonian eigenstates into the thermal trace defining the single-particle (and -hole) correlators (4.7) we find the spectral decomposition

$$C_{xy}^{\text{sp}}(\tau) = \frac{1}{Z} \sum_{\alpha n} z_{\alpha x n} z_{\alpha y n}^* e^{-E_n \tau} e^{-E_\alpha(\beta-\tau)} , \tag{4.21}$$

$$Z = \sum_n e^{-E_n \beta} \tag{4.22}$$

---

[1] For convenience, we denote the number of independent samples simply by N_{cfg} from here on.
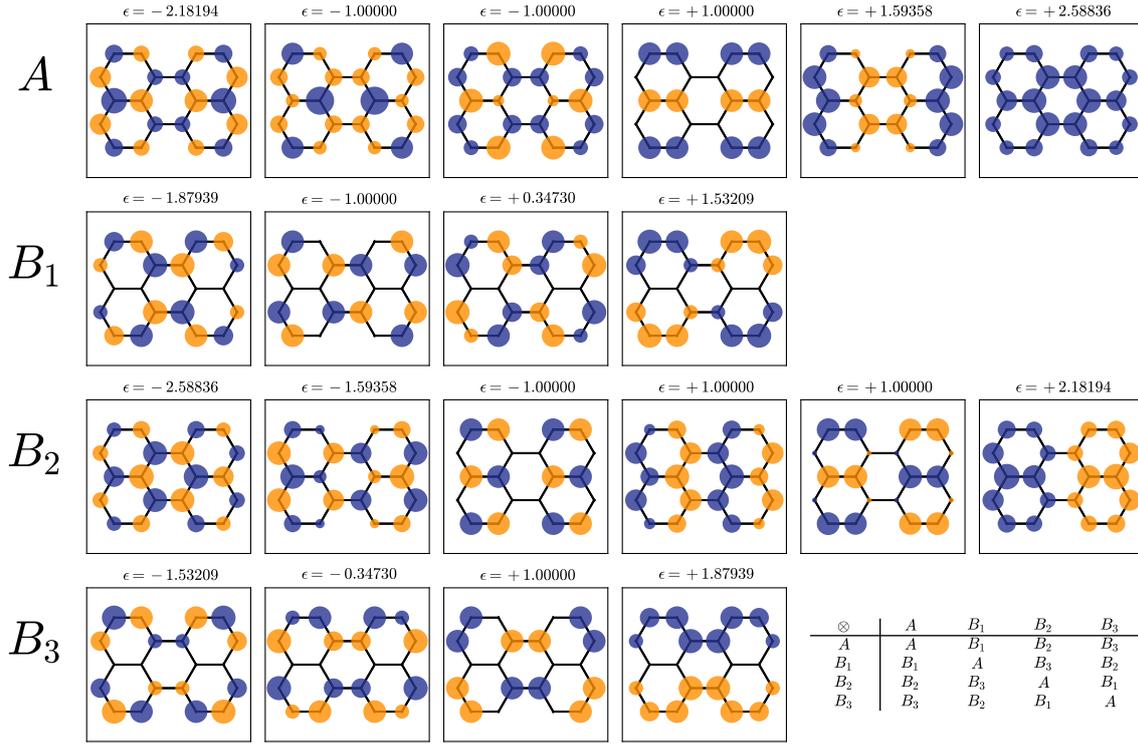
Figure 4.10: Single-particle eigenoperators $\sum_x c_x^* p_x^\dagger$ of the perylene tight-binding Hamiltonian. The circle on each site is scaled according to $|c|$, the absolute value of operator's amplitude there, and colored according to its sign (dark blue is positive, light red is negative). Each row is an irreducible representation, each irrep is sorted by the non-interacting tight-binding energy eigenvalue ϵ labelling the operators. In the lower-right we show the D_2 product table.

where we define the overlap factors

$$z_{\alpha x n} = \langle \alpha | p_x | n \rangle \tag{4.23}$$

and α and n label many-body energy eigenstates that differ by the quantum numbers of a single particle.

In the large-β limit the spectral decomposition simplifies to

$$C_{xy}^{\text{sp}}(\tau) = \sum_n z_{\Omega x n} z_{\Omega y n}^* e^{-(E_n - E_\Omega)\tau} \tag{4.24}$$

with $|\Omega\rangle$ the many-body ground state (if multiple states are degenerate, the decomposition is the obvious sum). By analyzing the spectral decomposition we can find energy differences from the ground state; at finite chemical potential $\mu \neq 0$ the eigenvalues E are of $H - \mu Q$.

4.A.4 Diagonalizing Correlators

An analogue of CPT symmetry allows us to average the single-particle and the time-reversed single-hole correlators; this helps us increase statistics and reduce the amount of required analysis, and we henceforth drop the single-particle superscript on C.

The point symmetry group of perylene is typically identified as D_{2h}. Our Hamiltonian, however, treats the ions as fixed, and we can split the symmetry into the dihedral group D_2 and a $\mathbb{Z}_2$ whose only action is to flip spin components; we already average over particles and holes leveraging the equivalent of CPT, so this $\mathbb{Z}_2$ is accounted for. The D_2 symmetry can be understood as a combination of reflections across the two principle axes; the A irrep is even under both reflections,

the B_1 irrep is odd under top-to-bottom reflections and odd under left-to-right reflections, B_2 is even/odd, and B_3 is odd/even.

We can perform a basis transformation on the correlation functions (4.7) to compute correlators of $\sum_x c_x^* p_x^\dagger$ with the amplitudes c defined on every site. The vector space defined on the 20 sites can be decomposed into invariant subspaces on which the action of the D_2 symmetries act irreducibly as A, B_1, B_2, and B_3; in a slight but common abuse of language we identify these invariant subspaces as the irreps themselves. The irreps are all one-dimensional and have multiplicity 6, 4, 6, and 4, respectively. Fig. 4.10 shows an orthonormal basis of operators for each irrep, chosen to diagonalize the tight-binding ($U = 0$) problem.

We can divide the ions of the lattice in A and B sublattices such that neighbours are always in the different sublattice. If we multiply all the fermion operators on a single sublattice by -1, the tight-binding Hamiltonian flips, because every possible hopping picks up exactly one sign, and we see that the tight-binding spectrum is symmetric around zero.

However, this sublattice symmetry does not commute with the D_2 point group, so the operators with opposite tight-binding energies (related by staggering the amplitudes' signs on one sublattice) appear in different irreps; this is particularly clear in Fig. 4.10 for the B_1 and B_3 irreps which have no accidental degeneracies. Another good example is the highest-energy A operator (with uniformly-signed amplitudes) and the lowest-energy B_2 operator (with corresponding staggered amplitudes). The ion-independent Hubbard interaction does not break the D_2 symmetry.

We can use the amplitudes c to construct a unitary matrix that block-diagonalizes the correlator C,

$$C_{\Lambda'_i, \Lambda_j}(\tau) = \sum_{xy} U_{\Lambda'_i, x} C_{xy}(\tau) (U^\dagger)_{y, \Lambda_j} = C_{ij}^\Lambda(\tau) \delta_{\Lambda'\Lambda} \tag{4.25}$$

where Λ and Λ' label the D_2 irreps and i and j operators of the respective irrep. Because our Hamiltonian has D_2 symmetry the irrep is conserved and the transformed correlator is block diagonal, as shown in the second equality (4.25). Each block $C^\Lambda(\tau)$ has a spectral decomposition (4.24) which sums over only states n that differ from the ground state by irrep Λ; put another way in the full spectral decomposition (4.21) the D_2 Wigner-Eckhart theorem states that $\alpha = \Lambda \otimes n$ using the D_2 product table in Fig. 4.10 where α and n are the irreps of their respective states.

When the interaction is weak, the single-particle correlation function transformed into this basis is nearly diagonal because the basis diagonalizes the tight-binding problem; when the interaction is strong, it remains block diagonal in irrep but within an irrep the operators can mix. Because every off-diagonal entry has differing contributions from excited states, no single unitary transformation diagonalizes an irreducible block for every time τ. We can nevertheless diagonalize each time slice independently.

Many diagonalization routines sort eigenvalues, which can lead to misidentifying the time dependence when correlators cross and cause trouble under a bootstrap analysis. A variety of sorting methods that can help to avoid this misidentification are discussed in Ref. [131]. To maintain the ordering of states and avoid said ambiguity, we diagonalize using a Jacobi method based on Givens rotation: the largest off-diagonal elements are iteratively rotated into the diagonal. By tracking these rotations we can also find the linear combination of operators that yield a diagonalized time slice.

However, this tracking procedures fail when correlators within an irrep cross; if we diagonalize timeslice-by-timeslice the crossings have level repulsion and introduce unphysical discontinuities in the resulting correlators. These crossings frequently appear, rendering a perfect timeslice-by-timeslice diagonalization inaccessible. This numerical problem stems from using only the 20 single particle operators, which do not constitute a complete basis of the spin-half $Q = 1$ sector. For example, we do not include in our calculation operators which have the same quantum numbers as our single-particle interpolators, like $p^\dagger h^\dagger h$. Interacting eigenstates mix $p^\dagger$ with all such operators, but our irreducible blocks are truncated to only the single-particle interpolators. If we would measure a much bigger correlator built from a complete basis of the single-particle sector the timeslice-by-timeslice diagonalization would produce perfect correlators with no repulsion.

Rather than grapple with these discontinuities, we instead adopt a variational approach. Given N_t unitaries U_t, one for each timeslice, we select the one that best diagonalizes all other time slices,

$$\tau = \min_{t \neq t'} \left\| U_t^\dagger \cdot U_{t'} - \mathbb{1} \right\|, \tag{4.26}$$

and use it to approximately diagonalize the blocks. This unitary can be thought of as variationally selecting a linear combination of the tight-binding eigenoperators shown in figure 4.10. From these mostly-diagonalized blocks we simply take the diagonal elements, resulting in a set of 20 correlators $C_{\Lambda_i}(\tau)$ where Λ labels an irrep and i is just an index. From these variationally-diagonalized correlators we are ultimately interested in the lowest energy—or more precisely, the energy closest to zero—in the spectral decomposition (4.24).

4.A.5 Fitting Energies

In order to systematically reduce the effect of excited states, we can fit correlators to a truncated spectral decomposition. The fit program proceeds with three steps; First, decide on a fit model, including number of states – terms in (4.24) – and fit range as well as identify prior-knowledge; second perform a Bayesian fit; and last measure how well the fit did.

As mentioned before the spectrum contains positive and negative energies. Therefore, the spectral decomposition can be split into two contributions, decaying (z_n^L, E_n^L) and increasing exponentials (z_n^R, E_n^R), suppressing the state label Λ_i for clarity. To further stabilize the fit and ensuring that $E_0^{L/R}$ is the smallest energy, the model is recast with relative energy differences $E_n^{R/L} \to \Delta E_n^{R/L}$ such that $\Delta E_n^{L/R} = E_n^{L/R} - E_{n-1}^{L/R} > 0$ resulting in the fit model (4.12). With this fit model, and the variational basis constructed in the previous section, we can identify the energy gap and overlap by

$$E^{\Lambda_i} = E_0^L \text{ or } -E_0^R, \tag{4.27}$$

$$\left| z_{\Omega \Lambda_i} \right|^2 = z_0^L \text{ or } z_0^R. \tag{4.28}$$

If the correlator $C_{\Lambda_i}(\tau)$ is primarily decaying take z_0^L, E_0^L otherwise z_0^R, E_0^R. This choice is made based on the fact that the slowest decay/increase of the correlator comes from the lowest energy, consequently we treat the other as excited state contamination. We truncate the spectral decomposition (4.12) after $N_{\text{states}} = 1, 2$ on the longer part of the correlator and keep $N_{\text{states}} = 1$ on the shorter end.

The contribution from excited states is different from time slice to time slice. Thus, it is advisable to include different fit intervals $\tau \in \delta[\tau_{\text{start}}, \tau_{\text{end}}]$. These are chosen by identifying the minimal point of the correlator, $\tau_{\text{min}} = \min_\tau(|C_{\Lambda_i}(\tau)|)$ and taking all possible combinations of $\tau_{\text{start}} < \tau_{\text{min}} < \tau_{\text{end}}$. For many correlators, the center part is relatively flat due to overlaps of exponentials causing artificially small energies $E_0^{L/R}$. To prevent this behaviour, the space of fit intervals is truncated to always take at least 75% of the subintervals to the left and right, i.e. $\tau_{\text{start}} < 0.75 \cdot (\tau_{\text{min}} - 1)$, $0.75 \cdot (N_t - 1 - \tau_{\text{min}}) < \tau_{\text{end}}$.

The last ingredients are the priors to the fit. As discussed previously, the non-interacting energy spectrum can be accessed analytically through ϵ_{Λ_i}. Though we expect divergence from this, it at least provides a good order of magnitude of the energies of the interacting simulations. Therefore, we use this information in combination with a log-normal prior for the 1-state fits,

$$E_0^{L/R} \sim \log \mathcal{N} \left(\left| E_{U=0}^{\Lambda_i}(\mu) \right|, \left| E_{U=0}^{\Lambda_i}(\mu) \right| \right) \tag{4.29}$$

In case a zero crossing is expected $(E_{U=0}^{\Lambda_i}(\mu) = 0)$, we simply use a gaussian prior with mean 0 and standard deviation 10. Considering the form of the correlator, especially its magnitudes at the end, we expect that the overlaps are $\mathcal{O}(1)$. This is encoded with a gaussian-prior with mean and standard

deviation equal 1. The variationally-diagonalized correlators are positive-definite, so too large a standard deviation would allow unphysical results.

For two-state fits the priors are partially determined by the one-state fit results we have already obtained. We utilize the model average, discussed in the next section 4.A.6. The central value serves as a mean to the (log-)gaussian prior while the standard deviation is determined by the maximum of 5σ and 10% of the central value, giving the fitter enough freedom to adjust the fit result. For the first two-state fit two additional parameter $z_1^{L/R}, \Delta E_1^{L/R}$, that can not be obtained from the one state fit results, are using a flat prior.

This fitting procedure is done on the central values of the correlator to provide central values for the energies. Furthermore, it is performed on each bootstrap sample to provide uncertainties on the energies. The fits are done in an uncorrelated manner, as the correlation is being tracked through the bootstraps.

4.A.6 Model Averaging

This procedure results in a high number of fits obtained using `lsqfit` [132]. For each, we compute the Akaike information criterion [126, 127, 128]

$$\text{AIC} = \chi^2 + 2\text{N}_{\text{params}} - 2\left|\tau_e - \tau_s\right|, \tag{4.30}$$

This measure penalizes the number of parameters and smaller fit range which is exactly what we are varying. A thorough discussion on this criterion in comparison to others can be found in Ref. [127]. With that we weight each fit result by the associated probability

$$P\left(\text{model}|\text{data}\right) \propto e^{-\frac{1}{2}\text{AIC}}. \tag{4.31}$$

to obtain the final parameter value $\langle p_n \rangle$, $p_n^i \in \{z_n^{L/R}, E_n^{L/R}\}_{n=0}^{\text{N}_{\text{states}}-1}$ where i labels the different results,

$$\langle p_n \rangle = \frac{\sum_i e^{-\frac{1}{2}\text{AIC}_i} p_n^i}{\sum_i e^{-\frac{1}{2}\text{AIC}_i}}. \tag{4.32}$$

4.A.7 Continuum Limit

Once the charges and model averaged energies for a given set of parameters $(\text{N}_{\text{t}}, \beta, \mu)$ are obtained a continuum limit has to be performed, $\delta = \beta/\text{N}_{\text{t}} \to 0$. The temperatures considered are too high for a reliable zero-temperature limit. We follow a similar approach as outlined in [25]. Expanding the correlator (4.7) in a geometric sum and expanding in small δ suggests a polynomial in δ. This results in a expansion for the total charge, estimated from the 0th time slice,

$$\langle Q(\delta, \beta) \rangle = Q_0(\beta) + \sum_{d=1}^{D} \delta^d Q_d(\beta) + O\left(\delta^D\right) \tag{4.33}$$

Usually, a control point is beneficial as otherwise priors can strongly bias fits of this form leaving us with $D = 1$ (2 parameters). Following the string of chemical potentials, the slopes Q_1 are distributed without a clear trend suggesting that discretization effects can be neglected – we are deep into the scaling regime. Consequently, we perform the continuum limit with only the constant piece, $D = 0$. figure 4.11 provides an overview of the continuum limits for the total system charges discussed in section 4.3.2. We find good fits across all systems, with some divergence on the coarsest lattices ($\text{N}_{\text{t}} = 32$).

Developing this expansion into the spectral expansion of the correlator maintains this relation.

$$E_0^{\Lambda_i}(\delta, \beta) = E_0^{\Lambda_i}(\beta) + \sum_{d=1}^{D} \delta^d E_d^{\Lambda_i}(\beta) + O\left(\delta^D\right) \tag{4.34}$$
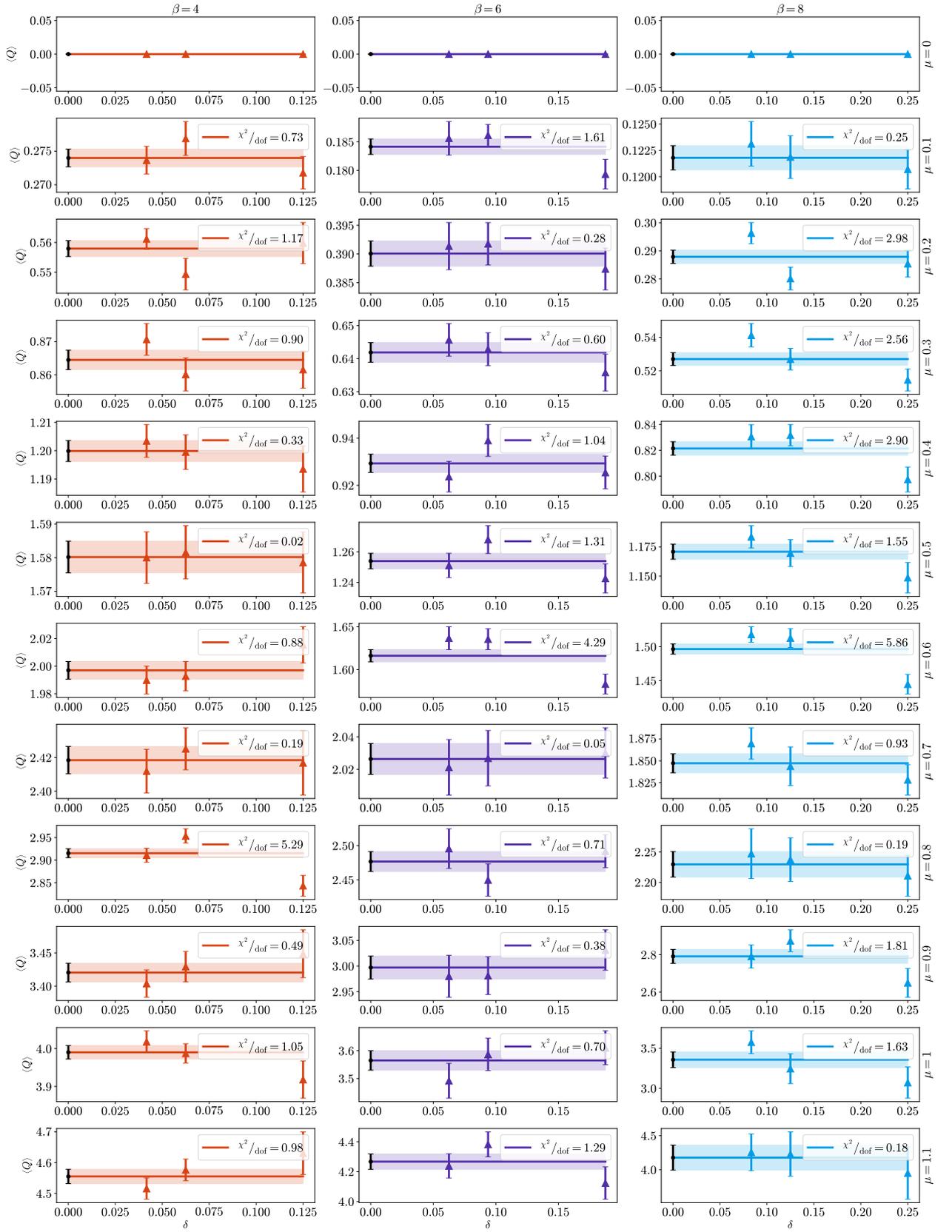
Figure 4.11: Continuum limit for the total system charge Q.

μ	A^0	A^1	A^2	A^3	A^4	A^5	B_1^0	B_1^1	B_1^2	B_1^3
0	-2.202(59)	1.861(49)	-1.420(21)	0.945(10)	0.911(15)	-0.910(15)	-1.627(37)	1.348(28)	-0.902(15)	0.338(16)
0.1	-2.182(46)	1.933(44)	-1.334(21)	1.004(12)	0.970(16)	-0.841(12)	-1.537(31)	1.373(35)	-0.822(14)	0.386(14)
0.2	-2.161(37)	2.025(44)	-1.288(19)	1.068(14)	1.019(18)	-0.775(11)	-1.535(24)	1.471(31)	-0.772(13)	0.444(16)
0.3	-2.025(33)	1.968(49)	-1.224(16)	1.126(14)	1.064(21)	-0.7187(61)	-1.479(19)	1.486(34)	-0.7135(89)	0.493(17)
0.4	-1.916(50)	1.973(76)	-1.155(15)	1.163(19)	1.115(23)	-0.6385(77)	-1.378(23)	1.450(49)	-0.6473(79)	0.556(16)
0.5	-1.847(42)	2.106(48)	-1.083(14)	1.243(18)	1.168(25)	-0.5663(66)	-1.313(22)	1.489(51)	-0.5750(78)	0.608(16)
0.6	-1.814(39)	2.197(58)	-1.007(15)	1.277(22)	1.218(30)	-0.4947(74)	-1.233(20)	1.559(46)	-0.5001(91)	0.681(16)
0.7	-1.694(46)	2.170(59)	-0.944(14)	1.339(24)	1.274(30)	-0.420(10)	-1.175(20)	1.557(60)	-0.4380(91)	0.742(16)
0.8	-1.667(46)	2.237(71)	-0.868(18)	1.386(26)	1.327(31)	-0.351(13)	-1.120(20)	1.695(45)	-0.368(11)	0.803(18)
0.9	-1.572(48)	2.154(85)	-0.813(18)	1.429(28)	1.385(31)	-0.286(16)	-1.055(23)	1.681(53)	-0.303(16)	0.873(16)
1	-1.546(41)	2.19(11)	-0.742(18)	1.491(31)	1.464(31)	-0.218(23)	-0.975(25)	1.691(64)	-0.223(24)	0.929(17)

μ	B_2^0	B_2^1	B_2^2	B_2^3	B_2^4	B_2^5	B_3^0	B_3^1	B_3^2	B_3^3
0	2.208(62)	-1.861(43)	1.435(21)	0.927(13)	-0.932(12)	-0.896(15)	1.647(36)	-1.305(30)	0.917(14)	-0.328(12)
0.1	2.306(51)	-1.802(42)	1.470(25)	0.980(15)	-0.8605(94)	-0.833(13)	1.663(39)	-1.237(29)	0.952(19)	-0.2596(96)
0.2	2.402(62)	-1.796(35)	1.530(29)	1.041(17)	-0.7890(90)	-0.766(12)	1.797(31)	-1.233(21)	1.022(19)	-0.2026(87)
0.3	2.276(63)	-1.715(23)	1.552(31)	1.107(16)	-0.7275(52)	-0.7031(93)	1.762(44)	-1.198(13)	1.052(22)	-0.1359(84)
0.4	2.27(11)	-1.582(39)	1.583(38)	1.142(21)	-0.6488(54)	-0.6356(87)	1.739(66)	-1.099(16)	1.102(25)	-0.068(11)
0.5	2.385(64)	-1.533(29)	1.607(40)	1.224(20)	-0.5674(62)	-0.5630(88)	1.813(52)	-1.040(14)	1.151(25)	0.12(12)
0.6	2.481(91)	-1.480(28)	1.650(43)	1.251(26)	-0.4904(73)	-0.4955(96)	1.794(73)	-0.962(16)	1.236(25)	0.098(71)
0.7	2.449(95)	-1.415(24)	1.732(41)	1.304(26)	-0.4208(92)	-0.430(11)	1.892(66)	-0.908(16)	1.278(29)	0.149(37)
0.8	2.608(78)	-1.350(25)	1.761(45)	1.355(29)	-0.346(13)	-0.357(16)	2.027(50)	-0.828(19)	1.345(27)	0.216(37)
0.9	2.651(95)	-1.288(24)	1.756(54)	1.402(29)	-0.275(17)	-0.297(17)	1.955(68)	-0.760(24)	1.406(26)	0.275(26)
1	2.35(18)	-1.216(31)	1.899(47)	1.468(32)	-0.211(30)	-0.229(21)	2.004(73)	-0.698(24)	1.443(33)	0.340(26)

Table 4.3: Values of the energy levels at $\beta = 4$. These numbers correspond to the squares or points displayed in 4.12 and 4.13 respectively.

Where this sum is truncated to some power D. We truncate at $D = 0$ similar to the total charge.

In figure 4.7 the continuum limit for B_3^3 is shown. This flat extrapolation is typical extrapolation for all states and we do not show them here. The results are further summarized in table 4.3 for $\beta = 4$ and 4.4 for $\beta = 6$.

4.B   More Spectrum

We provide the $\beta = 4, 6$ spectra in figures 4.12 and 4.14 and summarize the values in the tables 4.3 and 4.4 respectively.

4.C   Complex Contour

Here we provide a short explanation for our choice of imaginary offset.

In lattice field theory it has been known for a while that a contour deformation to the tangent plane of the main saddle point of the action, i.e. the one with the greatest statistical weight, reduces the sign problem. This point in $\mathbb{C}^{|\Lambda|}$ fulfils $\left(\partial_{x,t} S[\phi]\right)\big|_{\phi=\phi_c} = 0$. For the Hubbard model this tangent plane turns out to be parallel to the real axis due to symmetry, hence we are talking about an imaginary shift. Intuitively this improvement makes sense, because the integration manifold would be closer to the Lefschetz Thimbles. The novelty of our recently developed contour deformation is the expansion of the action around said saddle point making it an effective action. This follows a standard practice in QFT and is equivalent to taking into account one-particle irreducible diagrams. We call this the next to leading order approximation (NLO). Because the linear term vanishes we expand until second order and get

$$S_{\text{eff}}[\phi_c] = S[\phi_c] + \frac{1}{2} \log \det \mathbb{H}_{S[\phi_c]} . \tag{4.35}$$

as the new function to be minimized, where $\mathbb{H}$ is the hessian. This can be done numerically along the imaginary axis, i.e. $\phi_c = i\phi_1$. By including the expansion we take into account the curvature of a saddle point, which shifts the classical (tangent) offset towards the optimal sign minimizing plane.
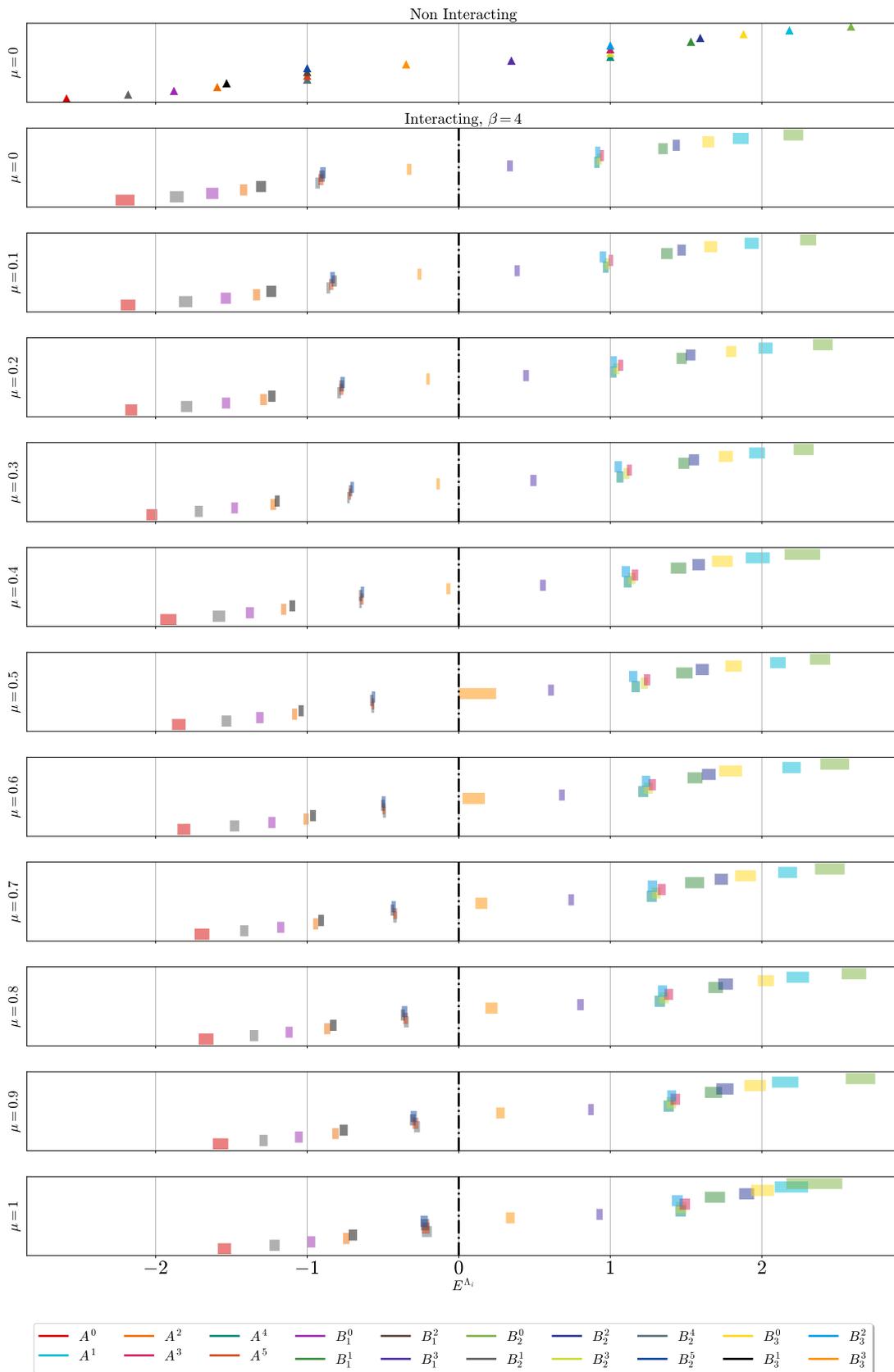
Figure 4.12: Similar spectrum as in figure 4.8 with $\beta = 4$. The sign problem is significantly less sever than at $\beta = 8$ consequently giving better estimates past $\mu = 0.8$.
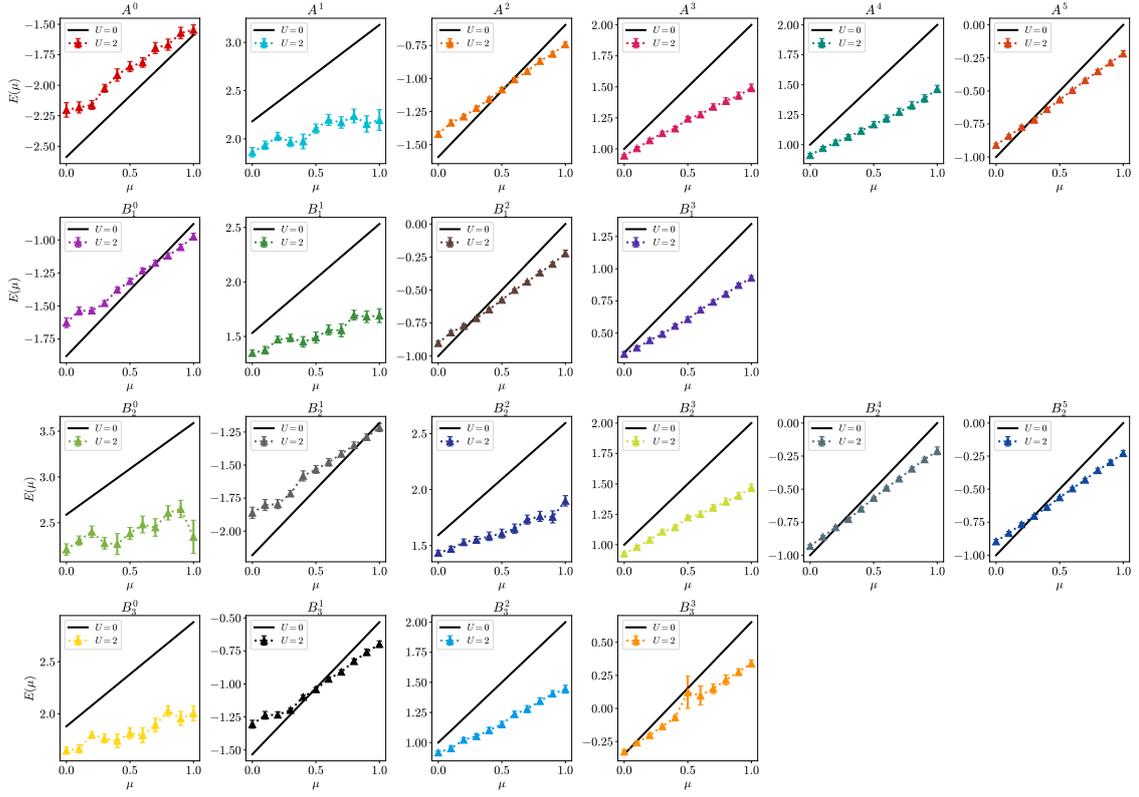
Figure 4.13: Similar plots as in figure 4.9 with different $\beta = 4$.

μ	A^0	A^1	A^2	A^3	A^4	A^5	B_1^0	B_1^1	B_1^2	B_1^3
0	-2.324(44)	1.991(34)	-1.486(17)	0.9809(67)	0.9507(93)	-0.9555(98)	-1.735(26)	1.414(31)	-0.904(15)	0.3368(88)
0.1	-2.264(29)	2.083(27)	-1.407(20)	1.036(11)	1.003(13)	-0.8759(77)	-1.663(22)	1.486(23)	-0.855(11)	0.393(10)
0.2	-2.142(41)	2.106(30)	-1.339(16)	1.099(13)	1.064(14)	-0.7941(69)	-1.569(25)	1.531(26)	-0.8071(64)	0.443(12)
0.3	-2.047(42)	2.163(45)	-1.2889(87)	1.182(13)	1.089(20)	-0.7137(70)	-1.497(26)	1.523(46)	-0.7194(76)	0.493(14)
0.4	-2.005(39)	2.181(54)	-1.167(17)	1.232(17)	1.157(25)	-0.6405(44)	-1.437(19)	1.619(39)	-0.6403(77)	0.547(14)
0.5	-1.886(55)	2.217(54)	-1.101(15)	1.290(20)	1.194(26)	-0.5551(62)	-1.344(23)	1.633(46)	-0.5662(73)	0.618(13)
0.6	-1.905(38)	2.306(63)	-1.021(17)	1.330(43)	1.266(33)	-0.4720(78)	-1.277(28)	1.696(53)	-0.4917(82)	0.683(13)
0.7	-1.724(85)	2.309(68)	-0.934(24)	1.367(47)	1.323(36)	-0.4106(67)	-1.217(27)	1.684(72)	-0.4274(58)	0.746(13)
0.8	-1.59(17)	2.29(15)	-0.850(38)	1.483(72)	1.358(82)	-0.330(12)	-1.111(50)	1.620(92)	-0.352(15)	0.829(14)
0.9	-1.63(12)	2.10(20)	-0.797(29)	1.509(59)	1.42(10)	-0.268(15)	-1.038(62)	1.72(13)	-0.284(27)	0.885(17)
1	-1.675(67)	2.20(20)	-0.787(22)	1.531(97)	1.43(12)	-0.197(39)	-1.021(58)	1.64(18)	-0.23(12)	0.965(24)

μ	B_2^0	B_2^1	B_2^2	B_2^3	B_2^4	B_2^5	B_3^0	B_3^1	B_3^2	B_3^3
0	2.327(41)	-1.976(32)	1.502(19)	0.9671(82)	-0.9688(89)	-0.925(12)	1.749(29)	-1.425(21)	0.931(13)	-0.3233(91)
0.1	2.399(38)	-1.938(24)	1.526(27)	1.027(11)	-0.8959(57)	-0.857(10)	1.804(28)	-1.369(15)	0.989(15)	-0.2649(62)
0.2	2.414(44)	-1.869(17)	1.560(36)	1.100(12)	-0.8120(54)	-0.7867(81)	1.841(37)	-1.296(13)	1.042(17)	-0.1988(44)
0.3	2.496(49)	-1.745(24)	1.595(43)	1.137(17)	-0.7390(30)	-0.7130(70)	1.889(45)	-1.194(18)	1.097(22)	-0.1245(48)
0.4	2.524(58)	-1.684(37)	1.635(45)	1.200(20)	-0.6436(36)	-0.6326(70)	1.952(40)	-1.132(15)	1.131(26)	-0.0535(56)
0.5	2.610(58)	-1.546(33)	1.620(62)	1.253(25)	-0.5659(38)	-0.5485(85)	1.962(57)	-1.063(13)	1.216(21)	0.06(13)
0.6	2.54(10)	-1.523(34)	1.663(83)	1.327(28)	-0.4799(47)	-0.4776(88)	2.01(11)	-0.954(22)	1.236(32)	0.092(27)
0.7	2.685(95)	-1.406(59)	1.786(78)	1.362(40)	-0.3944(81)	-0.4224(85)	2.09(10)	-0.930(13)	1.285(55)	0.160(20)
0.8	2.66(12)	-1.31(12)	1.846(71)	1.430(62)	-0.326(12)	-0.352(13)	1.97(14)	-0.826(27)	1.391(52)	0.228(17)
0.9	2.31(23)	-1.20(13)	1.74(18)	1.463(55)	-0.255(22)	-0.284(14)	1.99(13)	-0.808(42)	1.427(69)	0.292(18)
1	3.106(35)	-1.288(71)	1.85(13)	1.462(87)	-0.180(41)	-0.202(53)	1.65(24)	-0.720(40)	1.43(13)	0.341(17)

Table 4.4: Values of the energy levels at $\beta = 6$. These numbers correspond to the squares or points displayed in 4.14 and 4.15 respectively.
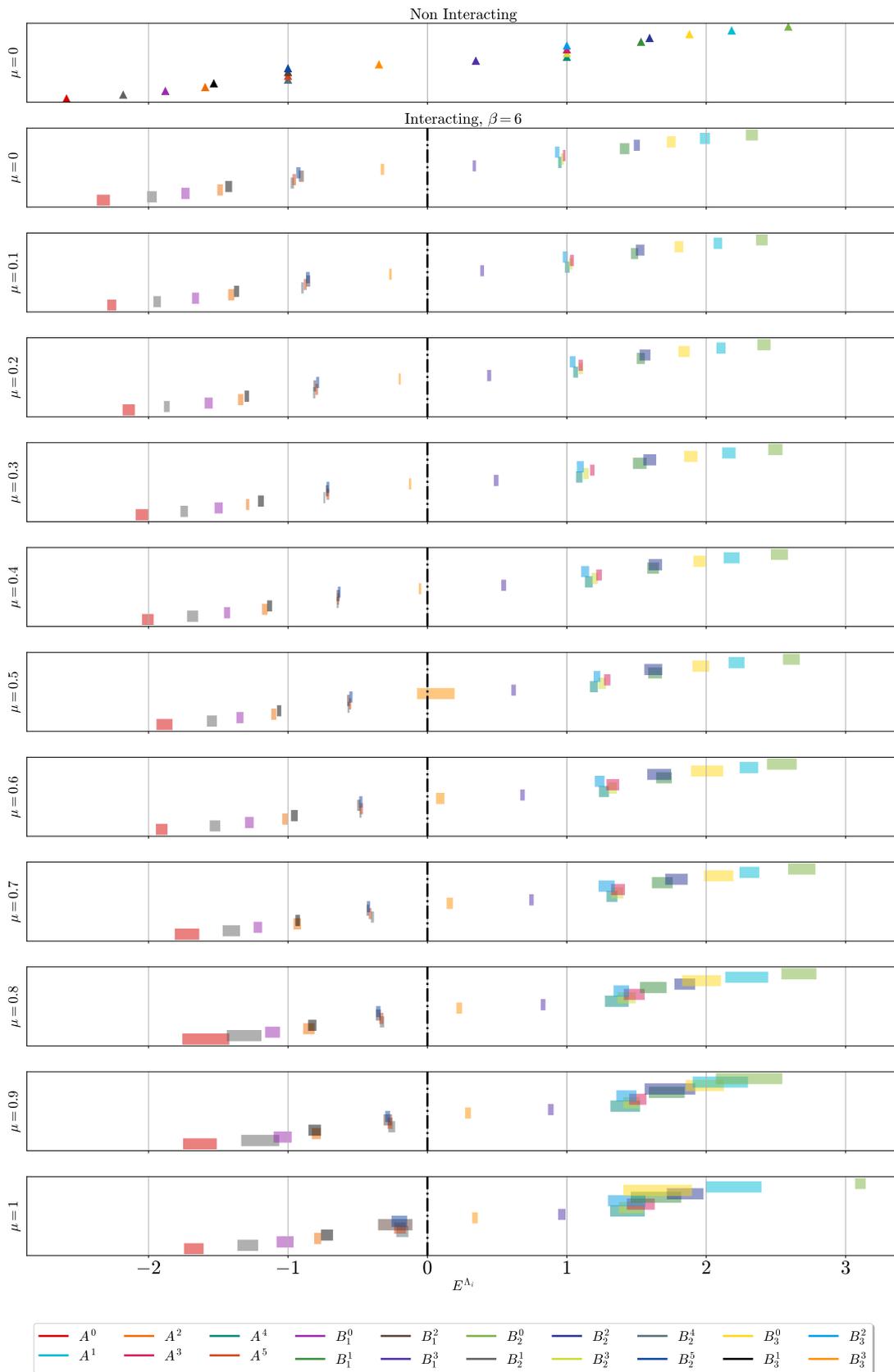
Figure 4.14: Similar spectrum as in figure 4.8 with different $\beta = 6$. The sign problem is less severe than at $\beta = 8$ consequently giving better estimates past $\mu = 0.8$.

73
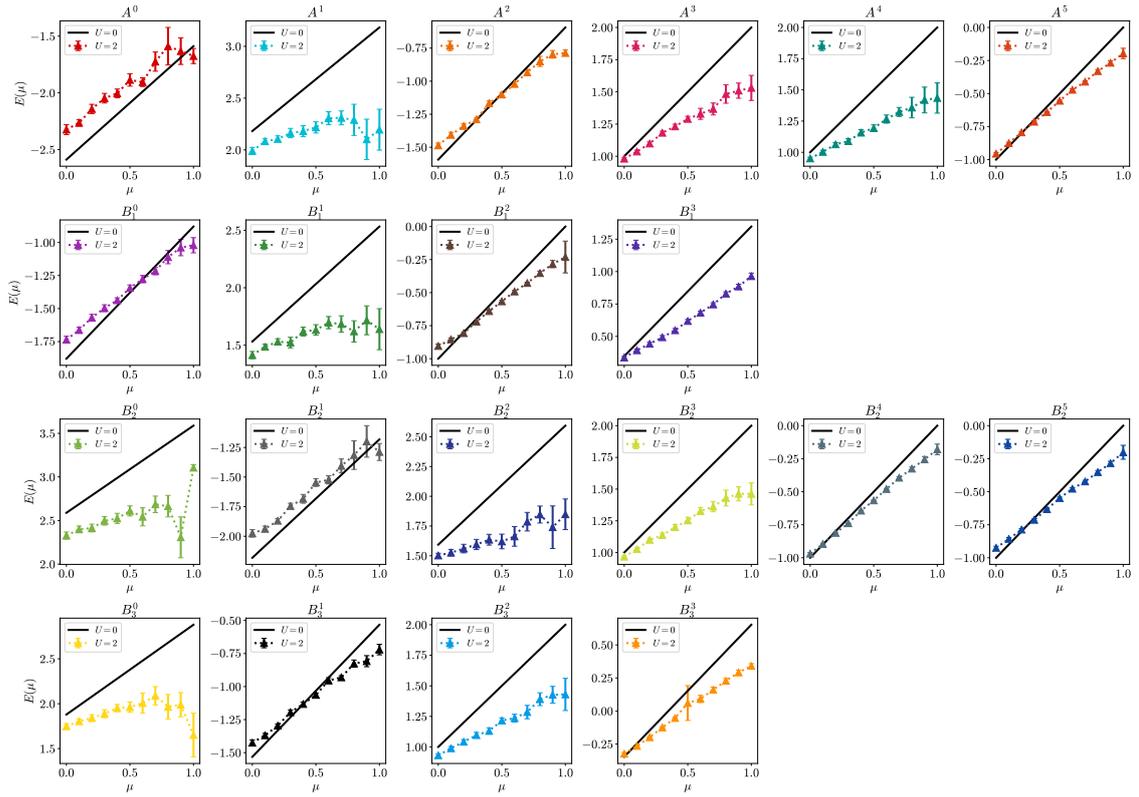
Figure 4.15: Similar plots as in figures 4.9, 4.12 with different $\beta = 6$.

We observe only small ranges of μ where it performs worse due to over-correction of steep regions in the action landscape.

Further details on the derivation and other optimizations can be found in [7, 2, 8].

Chapter 5

Moments of Nucleon Unpolarized, Polarized, and Transversity Parton Distribution Functions from Lattice QCD at the Physical Point

This chapter is based on [4]:

Marcel Rodekamp, Michael Engelhardt, Jeremy R. Green, Stefan Krieg, Simonetta Liuti, Stefan Meinel, John W. Negele, Andrew Pochinsky, and Sergey Syritsyn. "Moments of nucleon unpolarized, polarized, and transversity parton distribution functions from lattice QCD at the physical point." In: *Physical Review D* 109.7 (Apr. 2024), p. 074508. ISSN: 2470-0010, 2470-0029. DOI: 10.1103/PhysRevD.109.074508. arXiv: 2306.14431v2

In this chapter, the focus shifts away from the Hubbard model and sign problem towards the study of the internal structure of nucleons.

Nucleons are composite particles build from quarks and gluons. The underlying theory of quarks and gluons is Quantum Chromodynamics (QCD). Consequently, first principle studies of nuclear matter require calculation of observables within QCD. Due to the non-pertubative nature of QCD, an analytic treatment appears beyond reach for regimes of interest. Lattice QCD, which discretizes (Euclidean) space-time, formulates a systematic approach to calculate many relevant observables, inter alia related to nucleons. Once discretized, and put within a finite volume, Monte Carlo techniques can be employed to compute observables. Versions of the Hamiltonian Monte Carlo (HMC) are widely adapted and developed for the specific needs of the theory. A comprehensive introduction to the techniques of Lattice QCD is beyond the scope of this chapter; however, it can be found in several well-established textbooks, including [12, 13, 14, 15]. A brief overview is provided in the introductory section 1.3. As discussed there, a sign problem does not appear as the fermion determinant is set up to be real. Even without a sign problem calculations with lattice QCD are computationally challenging due to the vast amount of degrees of freedom.

The internal structure of Hadrons is a fascinating open question of particle physics from theoretical, phenomenological, as well as experimental point of view. Naturally, a great deal of work from the lattice QCD community goes into the evaluation of adjacent observables. In particular, the description of the nucleon structure through generalized parton distribution functions provides consistent mapping of the correlation between spatial and momentum information of the constituents, called the partons, of a fast moving nucleon. In the forward limit, these reduce to the ordinary parton distribution function encoding the momentum distribution. Traditionally, one can extract so-called Mellin moments of the parton distribution functions [42, 36, 37, 43, 44, 36, 45, 46, 47, 48, 49] by relating them to matrix element of local leading twist operators. this procedure works considerably well for the first few moments, however, becomes increasingly cumbersome for higher moments due to statistical noise and mixing with lower dimensional operators. Other methods have been proposed, based on the quasi-distribution, which require to calculate correlators at significantly larger momenta, effectively limited by statistical noise [40, 41].

In the following chapter a calculation of the second Mellin moments $\langle x \rangle$ of the nucleon's unpolarized, polarized, and transversity parton distribution functions is presented. This calculation is based on two lattice QCD ensembles at the physical pion mass, which were previously generated [133, 134]. Typical improved actions are employed to reduce discretization effect, in particular, the tree-level Symanzik-improved gauge action and for the 2+1 flavor use tree-level improved Wilson Clover fermions coupling via 2-level HEX-smearing. The moments are extracted from forward matrix elements of a set of local leading twist operators. Furthermore, the renomalization factors are determined in RI-(S)MOM and matched to $\overline{\text{MS}}$ at a scale 2 GeV.

The analysis is done by calculating ratios of two and three point functions effectively extracting the matrix elements in the large source-sink, previously denoted by T, and large separation, previously denoted by τ, limit. For this multiple methods are employed, first using the summed ratios and second through a direct fit of a two-state expansion of the ratio. Usually, summed ratios have reduced excited state contamination compared to the usual ratios, however, the extraction of matrix elements only takes the first state into account. Consequently, an analysis involving two-states is performed. The results are averaged and a systematic error coming from the truncation to one and two states is estimated.

The presented results indicate that operators exhibiting vanishing kinematics at zero momentum can have significantly reduced excited-state contamination. The resulting polarized moment is used to quantify the longitudinal contribution to the quark spin-orbit correlation. All results are in agreement within two sigma with previously determined lattice results.

The analysis was developed and executed primarily by myself, including the investigation of different operators. The required building blocks were calculated by Dr. Nesreen Hasan [50], yet they had not been analysed previously. The calculation of the renormalization factors was done by

Dr. Jeremy Green. All co-authors contributed with valuable discussions.

5.1 Introduction

The distribution of the momentum and spin within a hadron is encoded by parton distribution functions (PDFs). Determining the PDFs is thus an indispensable ingredient to our understanding of the structure of hadrons [135, 136, 40]. There have been various efforts of extracting the PDFs from global fits, for a recent summary see [137]. The Lattice QCD community has also achieved remarkable strides in the computation of PDFs over the recent years [37, 41, 138].

In this study[1], our focus centers on the evaluation of the second Mellin moment, denoted as $\langle x \rangle$ [42, 36, 37, 43], of unpolarized, polarized, and transversity PDFs. We achieve this through the examination of matrix elements of local twist-two operators [44, 36, 45, 46, 47, 48, 49]. This method does not require high momenta to suppress higher-twist contributions, as is needed in calculations that use non-local operators, for example the widely used quasi-PDF method [40, 41]. One of our objectives is to gain insights into the contamination stemming from excited states for different matrix elements and constraining the resulting uncertainty. To attain this objective, a comprehensive investigation of matrix elements at finite but modest momenta becomes imperative, as certain operators have nonvanishing matrix elements exclusively at nonzero momentum. Although the exploration of forward matrix elements of local operators at non-zero momentum is somewhat unconventional, references [139, 140, 141] have previously ventured into this territory.

This paper is organized as follows. In section B.2 we explain our analysis chain and discuss in detail which operators are considered. The different steps of the analysis to extract the matrix elements are shown in section 5.3. We continue the computation of moments in section 5.4 where they become renormalized and averaged over the different results. Further, our findings are put in relation to other Lattice QCD results and global fits. In section 5.5 we utilize the moment of the polarized PDF to compute the quark spin-orbit correlation. Last, in section B.4 we summarize our findings. There are three appendices: Appendix 5.A shows the extraction of the matrix element for each operator these results get summarized in plots shown in appendix 5.B and finally the calculation of renormalization factors is discussed in appendix 5.C.

5.2 Method

Moments of PDFs can be obtained by calculating forward matrix elements of local leading-twist operators [142, 143, 42, 45]

$$O^X \equiv O^X_{\{\alpha,\mu\}} = \overline{q}\Gamma^X_{\{\alpha} \overleftrightarrow{D}_{\mu\}} q. \tag{5.1}$$

Here, the symbol X denotes either V, A, or T, corresponding to the vector, axial, or tensor channels, respectively, and in the tensor case $\Gamma_\alpha = \sigma_{\beta\gamma}$ so that α is a compound index. These channels are associated with unpolarized, polarized, or transversity PDFs. Symmetrizing the indices and subtracting traces is indicated by braces, $\{\alpha, \mu\}$. We specifically focus on the isovector channel, which involves the difference between O^X for up and down quarks, $O^X(q = u) - O^X(q = d)$, to avoid calculating disconnected diagrams. The left-right acting covariant derivative $\overleftrightarrow{D} = \frac{1}{2}(\overrightarrow{D} - \overleftarrow{D})$ is constructed on the Euclidean lattice using central finite differences between neighboring lattice points, connected by appropriate gauge links.

It is well understood that these forward matrix elements are proportional to the desired moment $\langle x \rangle$ [44, 36, 45]. The matrix element is given by

$$\langle N(p)|O^X_{\{\alpha,\mu\}}|N(p)\rangle = \langle x \rangle \overline{u}_{N(p)}\Gamma^X_{\{\alpha} i p_{\mu\}} u_{N(p)}. \tag{5.2}$$

In this equation, p represents the 4-momentum of the nucleon.

---

[1] Preliminary results were reported in [6].

In the continuum, the operators described in Equation (B.1) form irreducible representations of the Lorentz group. However, in the context of Euclidean space, the Lorentz group is replaced by the orthogonal group [42]. When we transition to the lattice, the orthogonal group further reduces to the hypercubic group $H(4)$. This reduced symmetry can lead to certain operators mixing with lower-dimensional ones. Fortunately, for the specific one-derivative operator studied, such mixing does not occur [144].

Nevertheless, it is important to note that the Euclidean irreducible representations to which our operators belong are divided into multiple hypercubic irreducible representations. In our work, we adopt the common notation, where $\tau_a^{(b)}$ represents the ath irreducible representation of dimension b. Each of these hypercubic irreducible representations necessitates a distinct renormalization factor. To keep renormalization diagonal, we construct operators with well-defined hypercubic irreducible representations, as suggested by Göckeler et al. [42].

In practical terms, this implies that for each $\tau_a^{(b)}$, we must compute the corresponding renormalization factor $Z_{\tau_a^{(b)}}$. This factor is subsequently applied to the matrix elements of an operator that transforms irreducibly under the given representation. As a result, we denote the renormalization factor for the operator O^X as Z_{O^X}, equivalent to $Z_{\tau_a^{(b)}}$.

The matrix element described in Equation (B.2) can be determined on the lattice by considering the ratios of three-point and two-point correlation functions, as previously discussed in the literature, e.g. [44, 45]. The two-point correlation function, denoted as

$$C_{2\text{pt}}(\tau) = \int d^3 y \, e^{-i\vec{p}\vec{y}} \, \text{Tr}\left\{ \Gamma_{\text{pol}} \left\langle \chi(\vec{y}, \tau) \overline{\chi}(\vec{0}, 0) \right\rangle \right\}, \tag{5.3}$$

quantifies the correlation between a nucleon source and a nucleon sink separated by a time interval τ. Here we use[2] $\Gamma_{\text{pol}} = P_+ [1 - i\gamma_1\gamma_2]$ with $P_+ = {}^{(1+\gamma_4)}/2$ and a nucleon interpolating operator of the form

$$\chi_\alpha = \epsilon_{abc} \left(\tilde{u}_a^T C \gamma_5 P_+ \tilde{d}_b \right) \tilde{u}_{c,\alpha} \tag{5.4}$$

with smeared quark fields $\tilde{q}$.

The three-point correlation function, denoted as

$$C_{3\text{pt}}^{O^X}(T, \tau) = \int d^3 y \, d^3 z \left[e^{-i\vec{p}'\vec{y}} e^{i(\vec{p}'-\vec{p})\vec{z}} \, \text{Tr}\left\{ \Gamma_{\text{pol}} \left\langle \chi(\vec{y}, T) \, O^X(\vec{y}, \tau) \, \overline{\chi}(\vec{0}, 0) \right\rangle \right\} \right], \tag{5.5}$$

separates the source and sink nucleons by a time interval T while incorporating the operator of interest, O^X, at time τ. From here on we let $\vec{p}' = \vec{p}$ as indicated in (B.2). A visual representation is given by Figure B.1. The matrix element is extracted in the limit where

$$\mathcal{M} \equiv \lim_{T-\tau, \tau \to \infty} R(T, \tau) \equiv \lim_{T-\tau, \tau \to \infty} \frac{C_{3\text{pt}}^{O^X}(T, \tau)}{C_{2\text{pt}}(T)}. \tag{5.6}$$

Once the matrix element is obtained, we can compute the moment by simply dividing the kinematic factor, $\langle x \rangle K = \mathcal{M}$, with

$$K = \frac{1}{2E_N(p)} \frac{\text{Tr}\left\{ \Gamma_{\text{pol}} \left(-i\gamma_\mu p^\mu + m_N \right) \left[a^{\alpha,\mu} \Gamma_\alpha^X p_\mu \right] \left(-i\gamma_\mu p^\mu + m_N \right) \right\}}{\text{Tr}\left\{ \Gamma_{\text{pol}} \left(-i\gamma_\mu p^\mu + m_N \right) \right\}}, \tag{5.7}$$

where the $a^{\alpha,\mu} \in \mathbb{R}$ are appropriate factors to express the symmetrization and removal of traces discussed above;

This analysis involves a spectral decomposition of the ratio, which allows us to isolate the matrix element of the ground state:

$$R(T, \tau) = \mathcal{M} + \text{excited states}. \tag{5.8}$$

---

[2]The same results can also be obtained using P_+ by itself as the spin projector in $C_{2\text{pt}}$.
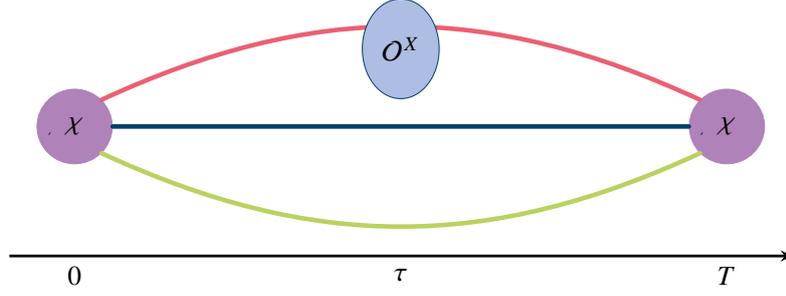
Figure 5.1: Graphical representation of $C_{3pt}^{O^X}(T, \tau)$: a source nucleon inserted at time $t = 0$ and a sink nucleon removed at time $t = T$. A local leading twist operator (B.1) is inserted on a given time slice τ. The nucleons on the lattice are represented by interpolating operators χ (5.4)

To account for the influence of the first excited state, we expand the expression, obtaining the leading contribution from excited states

$$\mathcal{M} \frac{1 + R_1 e^{-\frac{T}{2}\Delta E} \cosh\left[(T/2 - \tau)\Delta E\right] + R_2 e^{-T\Delta E}}{1 + R_3 e^{-T\Delta E}}, \tag{5.9}$$

where ΔE represents the energy difference between the first excited state and the ground state ($\Delta E = E_1 - E_0$). In principle, one would aim to consider large values of T and τ to approach the limit defined in Equation (B.3). However, it is important to note that as T increases, so does the statistical noise.

The constants in the numerator, R_1, R_2, are dependent on the specific operator O^X, and their values influence the extent of excited-state contamination in the matrix element. Smaller values of these constants or the presence of certain symmetries can lead to reduced excited-state contamination in the final result.

In the sum of ratios

$$S(T, \tau_{\text{skip}}) = a \sum_{\tau=\tau_{\text{skip}}}^{T-\tau_{\text{skip}}} R(T, \tau) = \mathcal{M}\left(T - \tau_{\text{skip}}\right) + \text{excited states}, \tag{5.10}$$

excited-state contamination is exponentially suppressed with T compared to $T/2$ for the ratios themselves [145, 146]. Increasing τ_{skip} reduces excited-state contamination. Following the proportionality relation of the ratios and desired matrix element (B.2), we can extract the latter by use of a finite difference. Neglecting excited states, one finds

$$\mathcal{M} = \frac{S(T + \delta, \tau_{\text{skip}}) - S(T, \tau_{\text{skip}})}{\delta}. \tag{5.11}$$

Due to the available data we use a combination of $\delta/a \in \{1, 2, 3\}$ depending on whether a neighbour $T + \delta$ is available.

The analysis is outlined as follows:

Estimation of Ratios: First, we calculate the ratios $R(T, \tau)$ and ratio sums $S(T, \tau_{\text{skip}})$ for each operator. In the unpolarized (V) case we use

1. $\tau_1^{(3)}$ $\frac{1}{2}\left[\frac{O_{11}^V + O_{22}^V + O_{33}^V}{3} - O_{44}^V\right]$,

2. $\tau_1^{(3)}$ $\frac{1}{\sqrt{2}}\left[O_{33}^V - O_{44}^V\right]$,

3. $\tau_3^{(6)}$ $\frac{1}{\sqrt{2}}\left[O_{14}^V + O_{41}^V\right]$,

further, in the polarized (A) case we use

$$
\begin{array}{lll}
1. & \tau_4^{(6)} & \frac{1}{\sqrt{2}}\left[O_{13}^A + O_{31}^A\right], \\[4pt]
2. & \tau_4^{(6)} & \frac{1}{\sqrt{2}}\left[O_{34}^A + O_{43}^A\right],
\end{array}
$$

and finally for the transversity (T) case we use

$$
\begin{array}{lll}
1. & \tau_1^{(8)} & O_{211}^T - O_{244}^T, \\[4pt]
2. & \tau_1^{(8)} & O_{233}^T - O_{244}^T, \\[4pt]
3. & \tau_2^{(8)} & O_{124}^T - O_{241}^T, \\[4pt]
4. & \tau_2^{(8)} & O_{142}^T + O_{421}^T - 2O_{214}^T.
\end{array}
$$

These have been carefully chosen to have nonzero kinematic factors, compare Equation (B.2), and to be linearly independent.

Matrix Element Extraction: In the next step we extract matrix elements $\mathcal{M}$ using two different methods. *Method 1*: We extract the slope via finite differences at a specific source-sink separation $T = T'$, compare (B.7). *Method 2*: We obtain the matrix element from a simultaneous (over all source-sink separations) and fully correlated fit to the 2-state form, equation (B.5). A matrix element obtained through either method is denoted as $\mathcal{M}|_{T',\mathfrak{m}}$, where $\mathfrak{m}$ represents the extraction method. For the second method the T' index can be ignored.

From fitting the $C_{2\text{pt}}(\tau)$ we can obtain the ground-state energy E_0 which is used to calculate the kinematic factor. After this, we calculate the unrenormalized moment as

$$
\mathfrak{X}_{O^X,p,\mathfrak{m}}(T') = \frac{\mathcal{M}|_{T',\mathfrak{m}}}{\bar{u}_{N(p)}\Gamma_{\{\alpha}^X \, i\, p_{\mu\}}u_{N(p)}}. \tag{5.12}
$$

To simplify the following equations, we define a compound index $j = (O^X, p, \mathfrak{m})$ that runs over all operators and momenta with nonzero kinematic factors as well as the different methods to obtain the matrix element.

Renormalization Factors: We determine the renormalization factors in RI-(S)MOM and match them to $\overline{\text{MS}}(2\,\text{GeV})$; for details see appendix 5.C. This allows us to express the renormalized moment as $\mathfrak{X}_j^{\text{ren}}(T') = Z_{O^X} \cdot \mathfrak{X}_j(T')$.

Moment of PDF: To obtain the second moments of PDFs, we define the central value as the weighted average of the different results:

$$
\langle x \rangle^{\text{ren}} = \sum_{j,T' \geq T_{\text{plat}}^j} \mathfrak{W}_j(T')\mathfrak{X}_j^{\text{ren}}(T'). \tag{5.13}
$$

Here T_{plat}^j denotes the smallest source-sink separation such that $\mathfrak{X}_j(T')$ agree for all $T' \geq T_{\text{plat}}^j$. Naturally, the sum over T' does not apply for the second method, where we fit the 2-state function, as there is no T' to consider. The weights $\mathfrak{W}_j(T') \propto 1/\sigma_j^2(T')$ are normalised in such a way that weights associated to sum ratios sum to $\frac{1}{2}$ as do the weights for the 2-state fit. The used variances are estimated via bootstrap over $\mathfrak{X}_j(T')$ and the errors of the renormalization constants are propagated.

Systematic Error Estimation: Finally, we estimate a systematic error, constraining the uncertainty from excited state contamination, by calculating the weighted standard deviation over the different results:

$$
\sigma_{\text{syst}}^2 = \sum_{j,T' \geq T_{\text{plat}}^j} \mathfrak{W}_j(T')\left[\mathfrak{X}_j^{\text{ren}}(T') - \langle x \rangle^{\text{ren}}\right]^2. \tag{5.14}
$$

Again the sum over T' is not applied for the 2-state fit.

Relation to Quark Spin-Orbit Correlations: The longitudinal quark spin-orbit correlation $L_\ell^q S_\ell^q$ in the proton (where the subscript ℓ denotes alignment with the direction of motion of the

Table 5.1: Details of the used ensembles. The ensembles are at the physical pion mass, $m_\pi \approx m_\pi^{phys}$. A larger and a smaller lattice spacing, labelled as "Coarse" and "Fine" respectively, are available. The ensembles were generated with a tree-level Symanzik-improved gauge action with 2+1 flavour tree-level improved Wilson Clover fermions coupled via 2-level HEX-smearing [133, 134, 50]. Furthermore, the available source-sink separations (T) and momenta (p_x) which are used in the calculation of the ratios, Equation (B.3), are displayed.

Ensemble	Size	β	a[fm]	m_π[MeV]	$m_\pi L$	T/a	$p_x[2\pi/L]$	N_{cfg}
Coarse	48^4	3.31	0.1163(4)	136(2)	3.9	3, 4, 5, 6, 7, 8, 10, 12	0, −2	212
Fine	64^4	3.5	0.0926(6)	133(1)	4.0	10, 13, 16	0, −1	427

proton) is related to the generalized transverse momentum-dependent parton distribution (GTMD) G_{11}^q [147] as in Equation (5.15), which in turn can be related to the generalized parton distributions (GPDs) $\widetilde{H}^q$, H^q, E_T^q and $\widetilde{H}_T^q$ [148, 149] as in Equation (5.16),

$$2 L_\ell^q S_\ell^q = \int_{-1}^{1} dx \int d^2 k_T \frac{k_T^2}{m_N^2} G_{11}^q \tag{5.15}$$

$$= \frac{1}{2} \int_{-1}^{1} dx\, x\widetilde{H}^q - \frac{1}{2} \int_{-1}^{1} dx\, H^q$$

$$+ \frac{m_q}{2m_N} \int_{-1}^{1} dx \left(E_T^q + 2\widetilde{H}_T^q \right) , \tag{5.16}$$

where all distribution functions are quoted according to the nomenclature of [150] and are taken in the forward limit; k_T denotes the quark transverse momentum. $\widetilde{H}^q$ is the standard chiral-even helicity GPD and H^q is the standard chiral-even unpolarized GPD; E_T^q and $\widetilde{H}_T^q$ are chiral-odd GPDs. The longitudinal quark spin-orbit correlation has been evaluated according to Equation (5.15) in Ref. [151]; on the other hand, the results of the present work can be used complementarily to access the correlation via Equation (5.16), which can be viewed as the axial analogue of Ji's sum rule for orbital angular momentum: At the physical pion mass, the term proportional to m_q/m_N is negligible. In the forward limit, $\int dx\, H^q$ corresponds to the number of valence quarks, i.e., unity in the isovector, $u - d$ quark case considered here. Therefore, to an excellent approximation, one has

$$2 L_\ell^q S_\ell^q = \frac{1}{2} \left(\langle x \rangle_A^{ren} - 1 \right) , \tag{5.17}$$

where $\int dx\, x\widetilde{H}^q = \langle x \rangle_A$ in the forward limit has been identified. The results obtained in the following section will be used to quantify the longitudinal quark spin-orbit correlation and will also be confronted with the results of Ref. [151].

Simulation Parameters: We use a tree-level Symanzik-improved gauge action with 2+1 flavour tree-level improved Wilson Clover fermions coupling via 2-level HEX-smearing. Detailed information about the simulation setup can be found in references [133, 134, 50]. Key simulation parameters are summarized in Table B.1. Two ensembles, coarse and fine, at the physical pion mass are used. These ensembles correspond to lattice spacings of 0.1163(4) fm and 0.0926(6) fm, respectively. As described in [50], the smearing is done using Wuppertal smearing [152] – $\tilde{q} \propto (1 + \alpha H)^N q$ with H being the nearest-neighbor gauge-covariant hopping matrix – at $\alpha = 3$ and $N = 60, 100$ for the coarse and fine ensemble, respectively. For each ensemble, two-point and three-point correlation functions are calculated. These calculations involve source-sink separations ranging from approximately 0.3 fm to 1.4 fm for the coarse ensemble and approximately 0.9 fm to 1.5 fm for the fine ensemble. Furthermore, we consider two different momenta: $\vec{p} = (p_x, 0, 0)$ with $p_x = 0, -2[2\pi/L]$ for the coarse ensemble, and with $p_x = 0, -1[2\pi/L]$ for the fine ensemble.
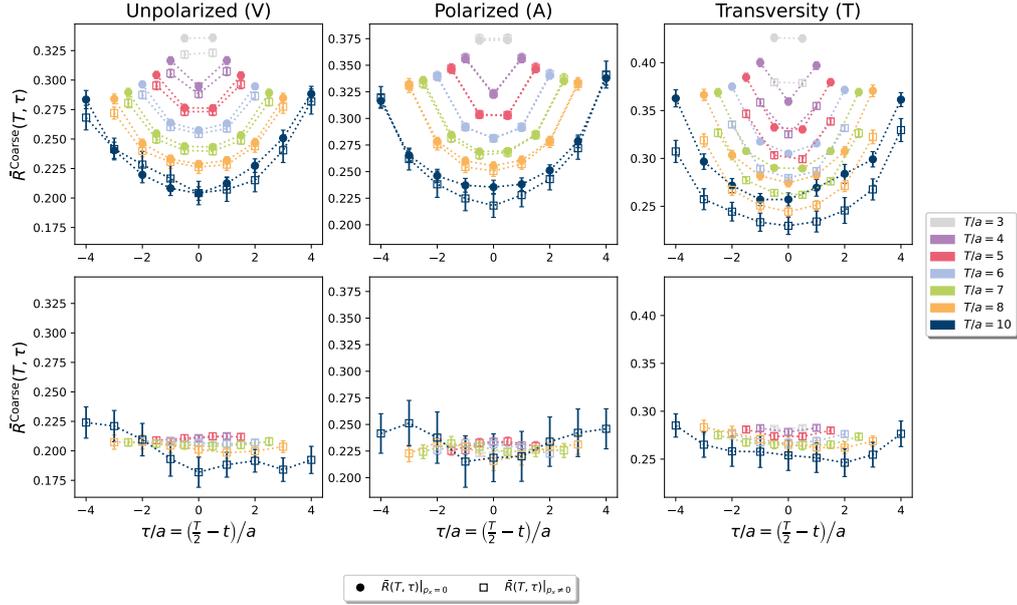
Figure 5.2: Ratios, cf. Equation (B.3), for the coarse ensemble. Various source-sink separations T are represented by different colors, while the two momenta are distinguished using filled circles and unfilled squares. Each subplot corresponds to a different operator from the different channels organized by column. For the Unpolarized (V) case we display operators 2. and 3.; for the polarized (A) case we display operators 1. and 2.; and for the transversity (T) case we display operators 3. and 1. for the upper and lower panel, respectively.

5.3 Estimation of Matrix Elements

In Figure B.2, we present results obtained from the coarse ensemble, using two different operators O^X per channel, as shown in the upper and lower rows. Each column is dedicated to a particular channel: From top to bottom we display the operators 2. and 3. (unpolarized), 1. and 2. (polarized), and 2. and 3. (transversity). Different source-sink separations are represented by various colours, while momenta are distinguished using filled circles for zero momentum and unfilled squares for finite momentum; this is kept consistent throughout all figures. A plateau in these plots corresponds to the matrix element of the shown operator. To simplify comparison we directly translate this to the bare moment, by multiplying with the kinematic factor $\overline{R}(T, \tau) = 1/\kappa \cdot R(T, \tau)$. It is worth noting that we exclude the largest source-sink separation from these plots due to its substantial statistical uncertainty.

These operator choices are intentionally selected to illustrate the extreme variability of the excited-state contamination. While the upper row has a clearly visible cosh behavior – as expected from the 2-state function (B.5) – the lower row remains perfectly flat within statistics. Moreover, we observe that the convergence in source-sink separation is much faster for the lower row. For instance, in the lower row the plateau already converges after $T/a = 3$ while the upper row requires $T/a \geq 8$ in these particular examples. This rapid convergence in the lower row is noteworthy, but it also comes with a drawback, which we observe in general, compare analysis summary plots in appendix 5.A for the other operators: Operators that exhibit such flat behavior at small source-sink separations have a vanishing kinematic factor at zero momentum, making them computationally more challenging to handle.

In Figure B.3, we present sum ratios, using the same operators as in Figure B.2 but put into one subplot. The upper and lower row represent the coarse and fine ensemble, respectively. The value of $\tau_{\text{skip}} = 1$ is fixed as the slope of the summed ratios did not change for larger values. The presence of excited-state contamination is subtly hinted at by the slight curvature observed in the
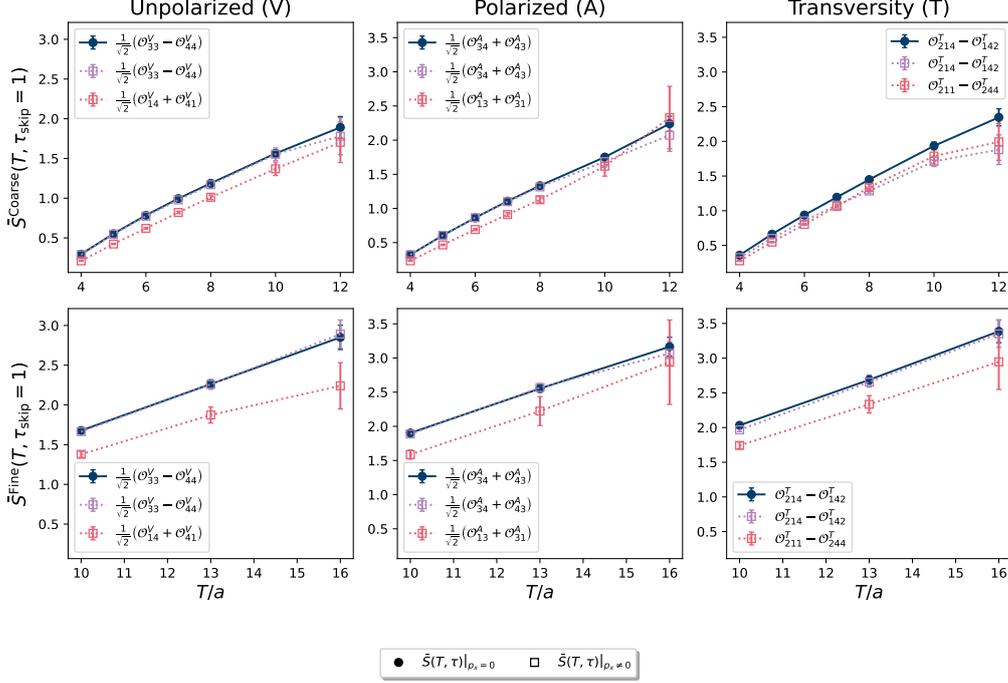
Figure 5.3: Ratio sums $\overline{S}(T, \tau_{\text{skip}})$ on the coarse and fine lattice, employing the same operators as in B.2. Each $\overline{S}(T, \tau_{\text{skip}})$ is plotted at fixed $\tau_{\text{skip}} = 1$. As in Fig. B.2, different momenta are displayed with hollow and filled markers.

data, although it is considerably less pronounced compared to the ratios.

In Figure 5.4, we present the result of the matrix element extraction for the same operators as displayed in Figure B.2. Similar plots for all used operators can be found in appendix 5.A. We plot horizontal lines to represent the average (over $T' \geq T_{\text{plat}}^j$) slope of the summed ratios, divided by the kinematic factor. These slopes are extracted with the finite difference approach (B.7). As the matrix element is given by a plateau of the ratios, the expectation is that the plotted slope agrees at least with the central points $\tau \sim 0$ of large source-sink separations, which can be verified for all operators within uncertainty. Again, those operators which are already flat match this expectation for more points and at smaller source sink separation.

Following the axolotl-like shape of the ratios, the solid, i.e. zero momentum, and dashed, i.e. finite momentum, lines indicate the central value 2-state fit result, using the form (B.5). The area around these indicate statistical uncertainty obtained via fitting on each bootstrap sample. We use all data points that are covered by the best fit plot in a (T, τ)-simultaneous fit. This presents a fit interval in τ/a which has been chosen by minimizing a χ^2/dof. The smaller source-sink separations for the coarse ensemble are excluded by this condition, as no points were left in the fit interval.

Considering all fits, values of χ^2/dof range from 0.4 to 2.7. Correlations which go into these were estimated over the bootstrap samples of the included points and then kept fixed for the central value fit as well as the fits per bootstrap sample.

Notably, the values of the matrix element obtained from summed ratios and 2-state fits always agree within statistics. The latter has reduced statistical uncertainty.

5.4 Moments of PDFs

In Figure B.4, we illustrate the results for the renormalized moments, which are extracted from the summed ratios (shown in grey, defined in Equation (B.6)) and the 2-state fits (displayed in red, as
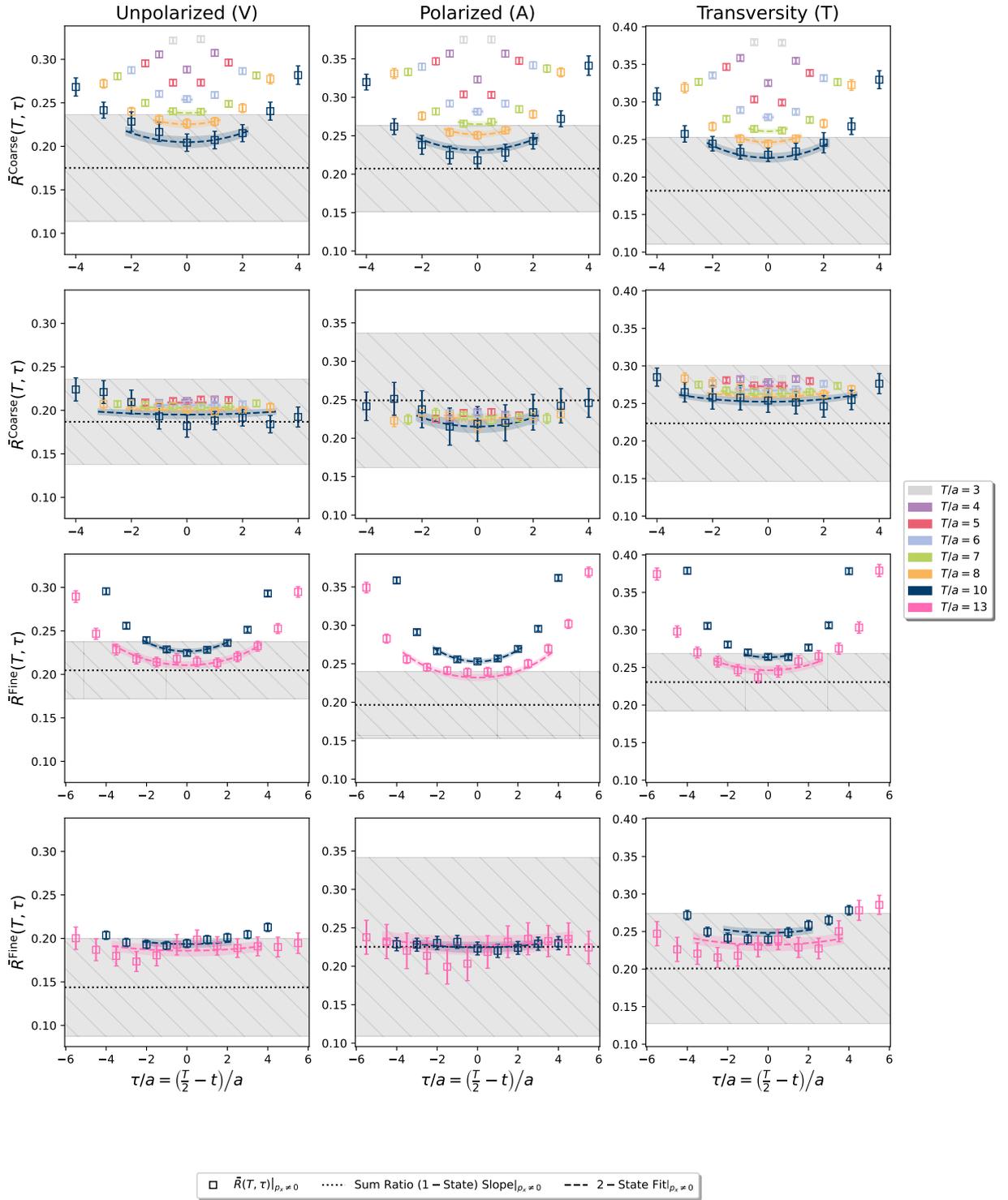
Figure 5.4: Extraction result for the matrix elements, cf. Equation (B.8), plotted on top of the original ratio data. The same operators and layout as in Figure B.2 are used. Coarse and fine ensemble results are displayed in the first two and second two rows, respectively. Dot-dashed and dotted horizontal lines represent the average slope of the summed ratios divided by the kinematic factor. Solid and dashed lines represent the simultaneous and fully correlated central value fit to the ratios using the 2-state form (B.5). Surrounding colored areas represent bootstrap uncertainties.
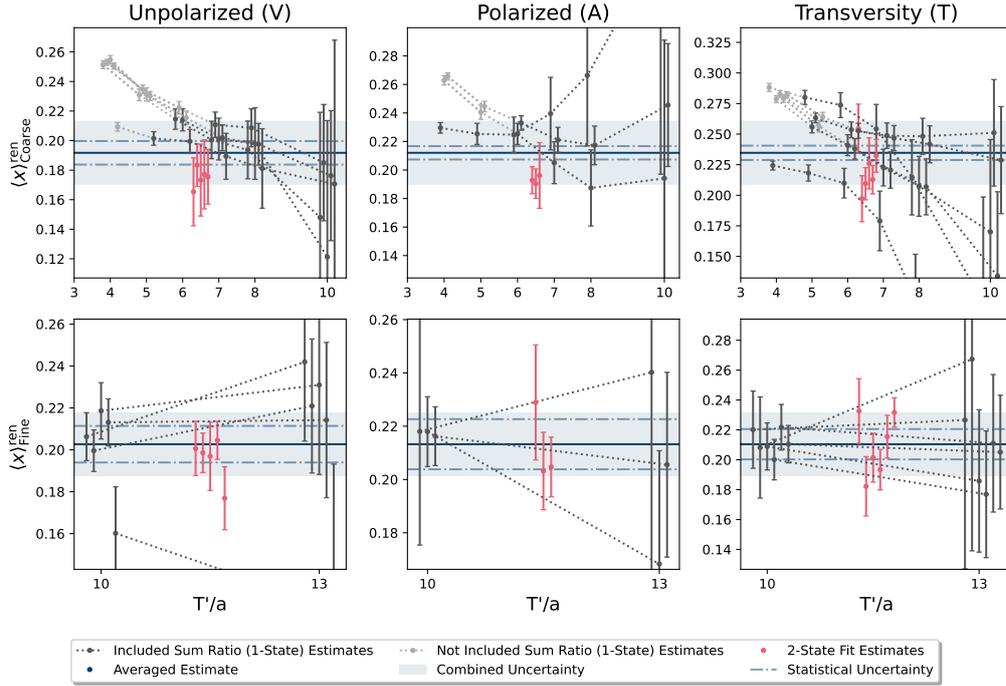
Figure 5.5: Renormalized moments calculated from the summed ratio (1-state) (grey) (B.6), and 2-state fit (red) (B.5). The final average is displayed as a blue solid line while its statistical uncertainty is indicated via the blue dot-dashed line. The blue band represents the statistical and systematic uncertainty, cf. Equation (B.10), added in quadrature. The light gray points are not included in the average as per the T_{plat}^{j} constraint, compare (B.9), to reduce excited-state effects. The ordinate limit is cut at 4σ around the final average to increase resolution of the relevant points.

defined in Equation (B.5))[3]. The final average is denoted by the blue solid line, while its statistical uncertainty is indicated by the blue dot-dashed lines. The blue band represents the combined statistical and systematic uncertainty, as outlined in Equation (B.10), added in quadrature. The light gray points are not included in the average, in accordance with the T_{plat}^{j} constraint. To enhance the resolution of the relevant data points, the ordinate limit is truncated at 4σ and centered around the final average. The numerical values of the final averages can be found in Table 5.2.

Comparing the two ensembles we find agreement within statistics indicating a flat continuum extrapolation. With only two points a reliable extrapolation is not possible. The best we can do is to interpret the data points as Gaussian distributions, with mean equaling the central value and width given by the uncertainties added in quadrature, and perform a Bayesian fit. The relevant scale of discretization effects [144, 153] is $a\Lambda_{\text{QCD}}$ resulting in a term proportional to $\alpha_s a\Lambda_{\text{QCD}}$. The operators themselves have tree-level quadratic discretization effects, resulting in the extrapolation

$$\langle x \rangle^{\text{ren}}(a) = \langle x \rangle^{\text{ren}}_{\text{cont}} \left(1 + m_1 \alpha_s a\Lambda_{\text{QCD}} + m_2 \left(a\Lambda_{\text{QCD}} \right)^2 \right) \tag{5.18}$$

We use Gaussian priors for the coefficients, $p_{m_i} = \mathcal{N}(0, 2)$ and no prior on the continuum value $\langle x \rangle^{\text{ren}}_{\text{cont}}$. We approximate $\alpha_s \approx 0.3$ which is sufficient due the fact that the coefficients m_i are mainly constrained by the prior. The continuum-extrapolated results are likewise given in Table 5.2.

Our results are in good agreement, at the level of one to two standard deviations, with moments previously computed by other Lattice QCD collaborations [154, 46, 47, 48, 155, 49]. Moreover, confronting our moments with phenomenological extractions, the comparison is quite favorable in the case of the axial moment, with Ref. [156] giving $\langle x \rangle_A = 0.190 \pm 0.008$. On the other hand, in

---

[3]Summary plots, showing these results separated and labeled by their corresponding operators, momenta, methods and source-sink separations can be found in appendix 5.B.

	Ensemble	$\langle x \rangle^{ren}$
Unpolarized (V)	Coarse	0.192(08)(20)
	Fine	0.203(09)(12)
	Continuum	0.200(17)
Polarized (A)	Coarse	0.212(05)(21)
	Fine	0.213(09)(07)
	Continuum	0.213(16)
Transversity (T)	Coarse	0.235(06)(25)
	Fine	0.210(10)(18)
	Continuum	0.219(21)

Table 5.2: Final averages for the second moments of PDFs in the unpolarized, polarized and transversity channels, compare Figure B.4. For the coarse and fine ensemble results, the central value is obtain as a weighted average over the different operators, momenta, and extraction methods, cf. Equation (B.9). Further, the statistical uncertainty (first uncertainty) comes from a bootstrap over the original ensemble, while the systematic uncertainty (second uncertainty) is computed using the weighted standard deviation over the same set of results, cf. Equation (B.10). We extrapolated the two points to the continuum limit using a Bayesian fit approach assuming them to be independent and Gaussian distributed with mean equaling the central value and standard deviation coming from the combined statistical and systematical uncertainty, compare Figure 5.6.

the unpolarized case, a certain tension between lattice and phenomenological results remains, with the recent determination in Ref. [137], $\langle x \rangle_V = 0.143(5)$, differing from our result by about three standard deviations.

5.5 Quark spin-orbit correlation

With the results from Table 5.2 we can calculate the longitudinal quark spin-orbit correlation in the proton according to Equation (5.17). The obtained values can be found in Table 5.3, along with the result obtained using the GTMD approach, Equation (5.15), in Ref. [151]. The results are in good agreement. As discussed in more detail in Ref. [151], the magnitude of this direct correlation between the spin and the orbital angular momentum of a quark significantly exceeds the correlation induced by the quark being embedded in a polarized proton environment. There is, therefore, a strong direct dynamical coupling between quark orbital angular momentum and spin, reminiscent of the jj coupling scheme in atomic physics, rather than the Russell-Saunders coupling scheme.

Table 5.3: Deduced isovector longitudinal quark spin-orbit correlation, estimated using the results for the polarized (A) moment shown in Table 5.2 and relation (5.17).

Ensemble	$2L_\ell^q S_\ell^q$ (σ_{stat}) (σ_{syst})
Coarse	−0.394(02)(10)
Fine	−0.393(05)(0)
Continuum	−0.393(08)
GTMD$\vert_{a=0.114\text{fm}}$ [151]	−0.40(2)
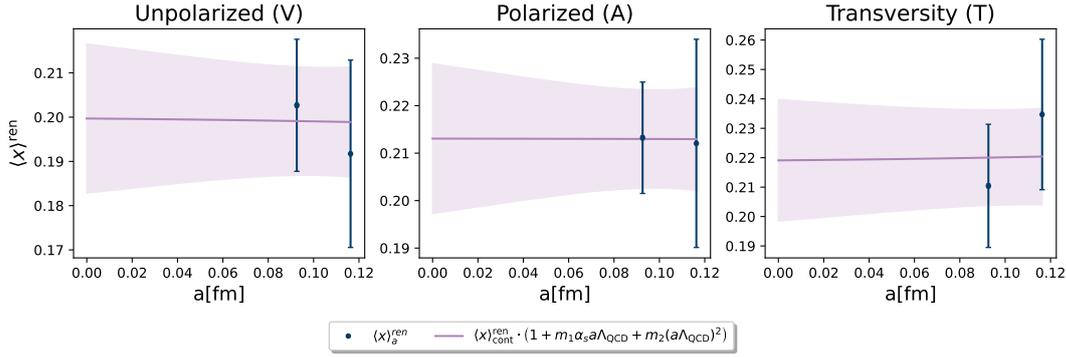
Figure 5.6: Continuum extrapolation using a Bayesian fit with the model described in Equation (5.18). The limited amount of data makes this extrapolation strongly dependent on the chosen priors for the coefficients m_i. Coming from a power counting these are expected to be of $O(1)$, reflected in Gaussian priors of mean zero and variance 2. Resulting estimates are listed in Table 5.2.

5.6 Summary

In this study, we compute the second Mellin moment $\langle x \rangle$ of the unpolarized, polarized, and transversity parton distribution functions using two lattice QCD ensembles at the physical pion mass. Our approach involves extracting forward nucleon matrix elements at both zero and finite momentum, boosted in the x-direction. Through the finite momentum data, we identify operators that exhibit remarkably small excited-state contamination. Given the two ensembles a reliable continuum extrapolation is not accessible. Regardless, we apply a Bayesian fit, accepting a strong dependence on the choice of priors, to provide a continuum estimate. The resulting values are in agreement with both individual ensembles: $\langle x \rangle_{u^+-d^+} = 0.200(17)$, $\langle x \rangle_{\Delta u^- - \Delta d^-} = 0.213(16)$, and $\langle x \rangle_{\delta u^+ - \delta d^+} = 0.219(21)$. Furthermore, we extract the isovector longitudinal quark spin-orbit correlation in the proton using the moment of the polarized PDF, $2L_\ell^q S_\ell^q = -0.393(08)$. We find good agreement with earlier calculations based on GTMDs [151].

Acknowledgments

5.A   Results Per Operator

In this appendix we show the analysis summary resolved per operator. As before, coarse and fine ensemble results are displayed in the first and second row, respectively. Different colors represent different source sink separations and the horizontal dash-dotted, i.e. zero momentum, and dotted, i.e. finite momentum, lines represent the average (over $T' \geq T_{\text{plat}}^{j}$) slope of the summed ratios, divided by the kinematic factor. These slopes are extracted with the finite difference approach (B.7). The solid and dashed curves are the best-fit result of the 2-state fit to (B.5), the surrounding band corresponds to the bootstrap uncertainty of the fit. Figure 5.7 displays the analysis of the operators corresponding to the unpolarized (vector) PDFs. Figure 5.8 displays the analysis of the operators corresponding to the polarized (axial) PDFs. Figures 5.9 and 5.10 display the analysis of the operators corresponding to the transversity (tensor) PDFs. As mentioned in Section 5.3, agreement of the slope of summed ratios with the plateau region expected around $\tau = 0$ is given for all operators within one sigma. Corresponding best 2-state fit lines are in perfect agreement with the data points.

5.B   Summary Plots

We present summary plots of the moments for the coarse 5.11 and fine 5.12 ensemble. The three channels, unpolarized (V), polarized (A), and transversity (T) are shown in the columns. Each result, i.e. the different operators and momenta, is displayed in the panels separated by the dotted and dashed lines. The solid black line separates the sum-ratio method, points in purple, and the 2-state fit method, points in red. For the sum-ratio method the different T' are spread across the abscissa. As a point of reference, the average over the points, as described in equation (B.9), is shown by the horizontal blue line, with the statistical uncertainty shown by the dotted dashed line and the combined uncertainty by the blue band. This corresponds to the blue line in figure B.4. A (strong) dependence on the source-sink separation can be seen in the sum-ratio related points.
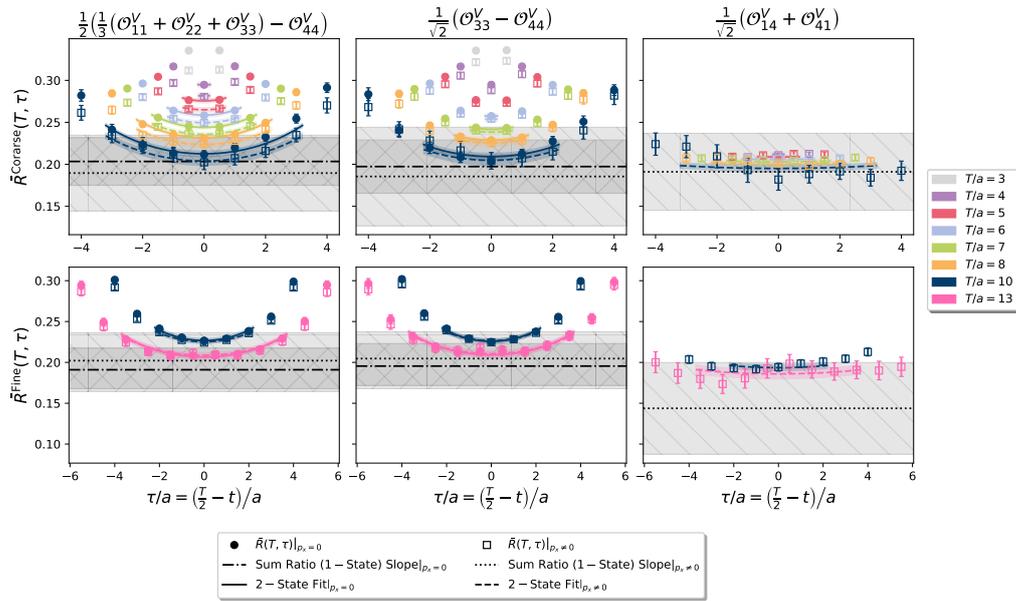
Figure 5.7: Analysis results of the ratios (points), slope of summed ratios (horizontal lines) and 2-state fit results (curves) for the operators 1., 2. and 3. corresponding to the unpolarized (vector) PDF.
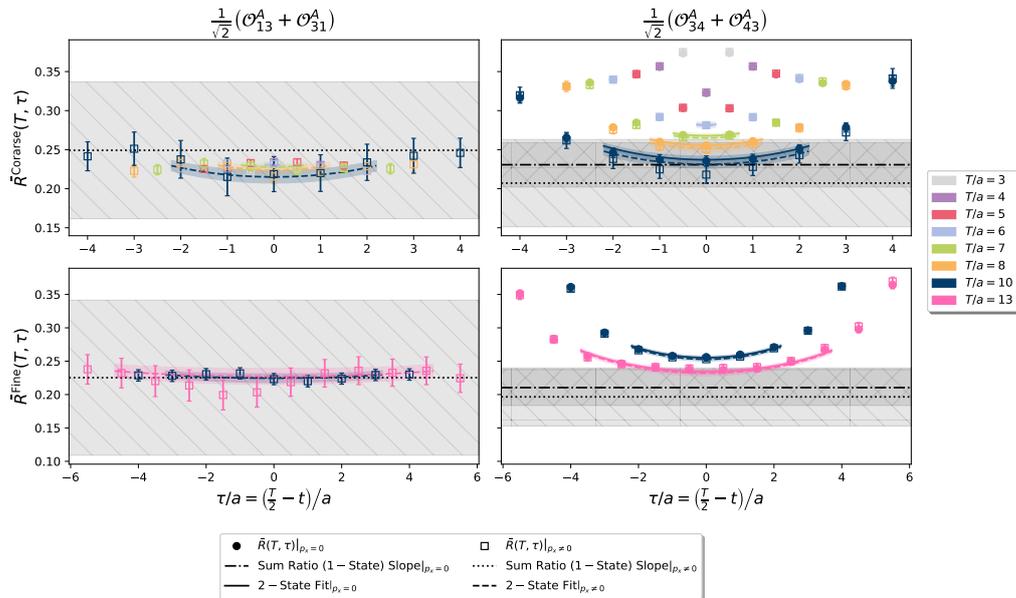
Figure 5.8: Analysis results of the ratios (points), slope of summed ratios (horizontal lines) and 2-state fit results (curves) for the operators 1. and 2. corresponding to the polarized (axial) PDF.
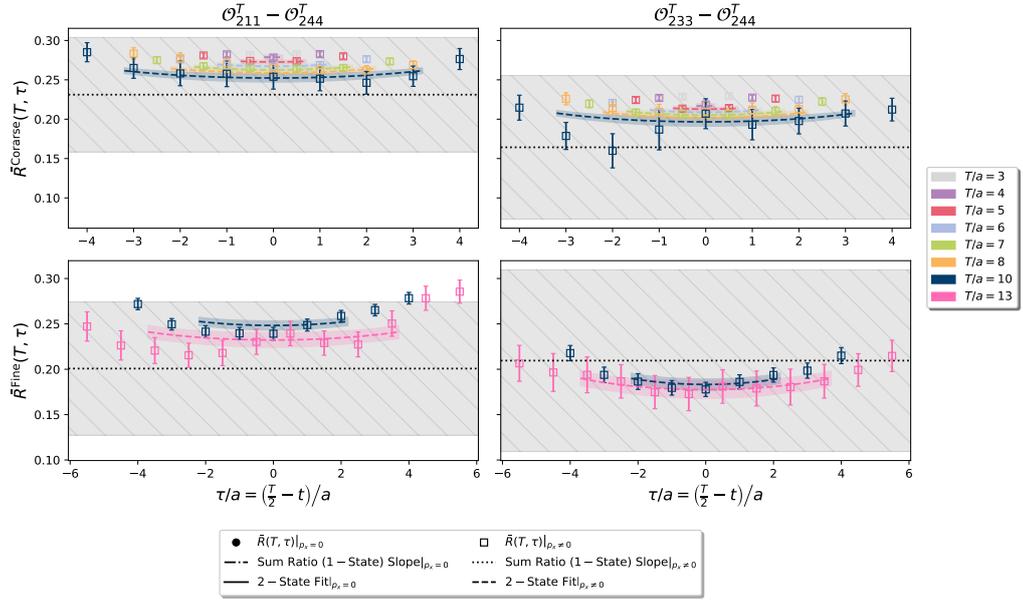
Figure 5.9: Analysis results of the ratios (points), slope of summed ratios (horizontal lines) and 2-state fit results (curves) for the operators 1. and 2. corresponding to the transversity (tensor) PDF.
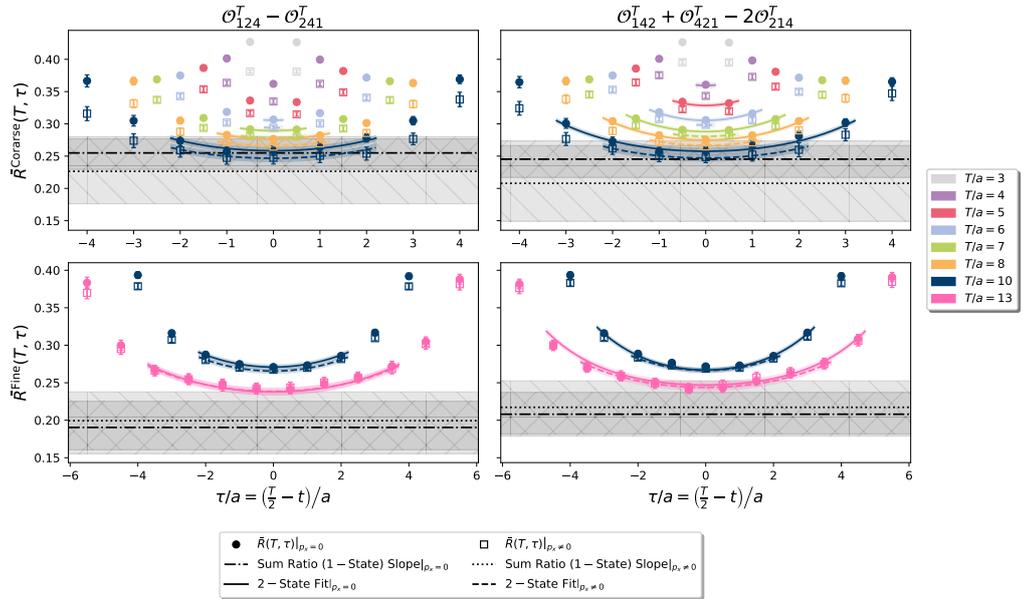
Figure 5.10: Analysis results of the ratios (points), slope of summed ratios (horizontal lines) and 2-state fit results (curves) for the operators 3. and 4. corresponding to the transversity (tensor) PDF.
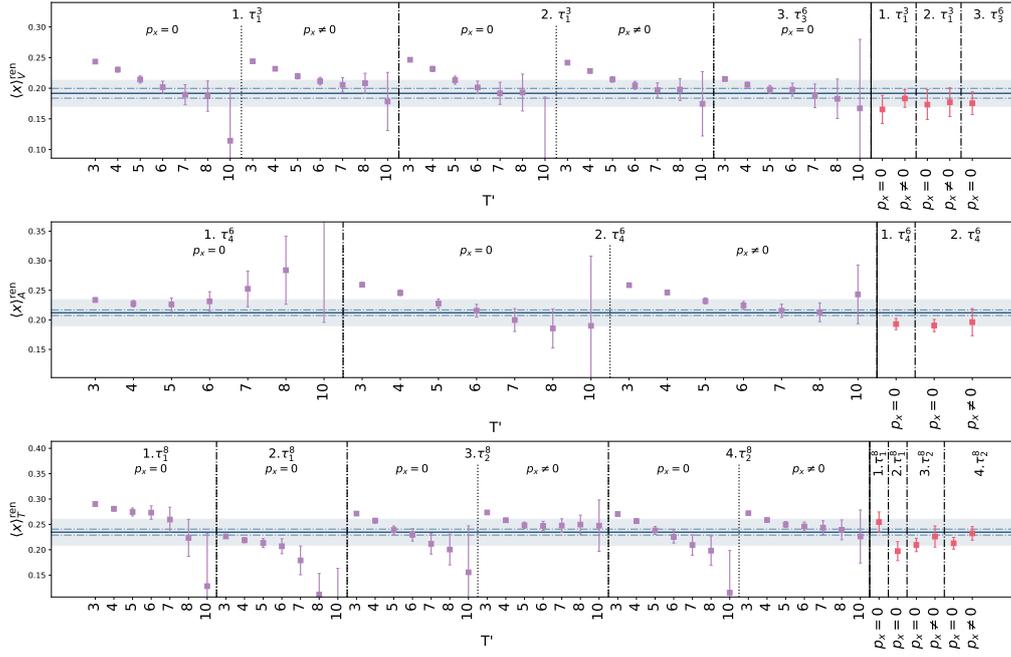
Figure 5.11: Summary plot of the renormalized moment of PDF of the coarse ensemble resolved for each operator O^X, represented by the corresponding ID and irrep, and momentum p_x. The purple points correspond to results obtained by the sum-ratio method evaluated at a source-sink separation T'. The red points are obtained using the two-state fit. The blue solid line corresponds to the overall average as described by equation (B.9) with the dashed line indicating statistical uncertainty and the blue area indicating statistical and systematic uncertainty added in quadrature.
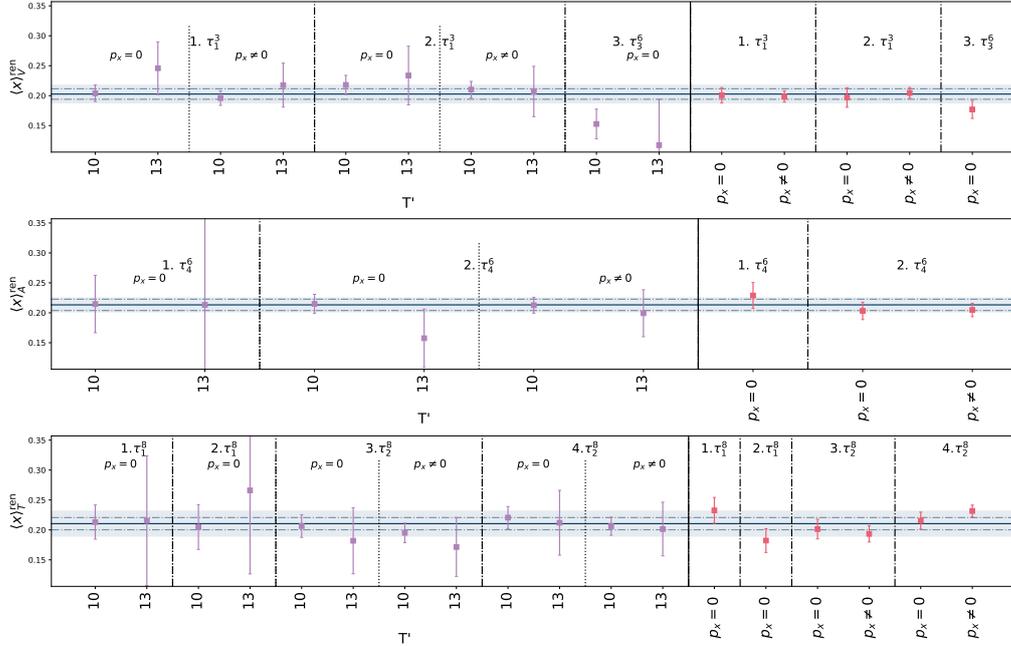
Figure 5.12: Summary plot of the renormalized moment of PDF of the fine ensemble, similar to figure 5.11

5.C   Nonperturbative Renormalization

We determine renormalization factors for isovector vector, axial, and tensor one-derivative twist-two operators using the nonperturbative Rome-Southampton approach [162], in both RI′-MOM [162, 163] and RI-SMOM schemes [164], and convert and evolve to the $\overline{\text{MS}}$ scheme at scale 2 GeV using perturbation theory. We label these renormalization factors Z_{DV}^{ρ}, Z_{DA}^{ρ}, and Z_{DT}^{ρ} for the one-derivative vector, axial, and tensor operators, respectively, with ρ denoting the irreducible representation of the hypercubic group that takes on two possible values in each case.

We largely follow our earlier work that used operators with no derivatives [50]. Our primary data are the Landau-gauge quark propagator,

$$S(p) = \int d^4x \, e^{-ip \cdot x} \langle u(x) \overline{u}(0) \rangle, \tag{5.19}$$

and the Landau-gauge Green's functions for operator O,

$$G_O(p', p) = \int d^4x' \, d^4x \, e^{-ip' \cdot x'} e^{ip \cdot x} \langle u(x') O(0) \overline{u}(x) \rangle. \tag{5.20}$$

Here O is an isovector quark bilinear with one derivative, yielding one Wick contraction: a connected diagram. We evaluate these objects using four-dimensional volume plane wave sources, yielding an effectively large sample size from the volume average. From these, we construct our main objects, the amputated Green's functions,

$$\Lambda_O(p', p) = S^{-1}(p') G_O(p', p) S^{-1}(p). \tag{5.21}$$

Provided that O belongs to a definite irreducible representation of the hypercubic group, these renormalize diagonally: $\Lambda_O^R = (Z_O/Z_\psi) \Lambda_O$. To avoid determining Z_ψ directly, we will form ratios to determine Z_O/Z_V and take Z_V computed from pion three-point functions in Ref. [50].

5.C.1   Conditions and matching

The RI′-MOM scheme uses kinematics $p' = p$, whereas RI-SMOM uses $p^2 = (p')^2 = q^2$ with $q = p' - p$. In both cases, the scale is defined as $\mu^2 = p^2$. For the vector current, we impose the conditions listed in Ref. [50] on $\Lambda_{V_\mu}^R$ to determine Z_V/Z_ψ.

For the one-derivative operators, we start with the continuum decomposition of the amputated Green's function $\Lambda_{O_{\mu\nu\dots}}(p', p)$ into a sum of products of $O(4)$-invariant functions $\Sigma_O^{(i)}(p^2)$ and simple kinematic tensors $\Lambda_{O_{\mu\nu\dots}}^{(i)}(p', p)$. We then decompose the operator and its kinematic tensors into irreducible representations ρ, replacing $\mu\nu\dots$ with ρn, where n ranges from 1 to the dimension of ρ. Tracing the amputated Green's function with each of the tensors within each irrep, we get

$$\sum_n \text{Tr} \left[\Lambda_{O,\rho n}^{(i)}(p', p) \Lambda_{O,\rho n}(p', p) \right] = M_\rho^{ij}(p', p) \Sigma_{O,\rho}^{(j)}(p^2), \tag{5.22}$$

where

$$M_\rho^{ij}(p', p) = \sum_n \text{Tr} \left[\Lambda_{O,\rho n}^{(i)}(p', p) \Lambda_{O,\rho n}^{(j)}(p', p) \right] \tag{5.23}$$

is a known kinematic matrix. Inverting M, we obtain the $O(4)$-invariant functions computed within each irrep ρ, $\Sigma_{O,\rho}^{(i)}(p^2)$. Our choice of decomposition, given below, is such that at tree level, $\Sigma_O^{(i)}(p^2) = \delta^{i1}$, and our renormalization conditions will all be of the form $\Sigma_{O^R,\rho}^{(1)}(\mu^2) = 1$. Basing this condition on a $O(4)$-invariant function computed within each irrep ensures that the ratio of renormalization factors for two different lattice irreps of the same continuum operator is scale and scheme-invariant.

The one-derivative vector operator is

$$O_{\mu\nu}^V = \mathcal{S}\overline{\psi}\tau_3\gamma_\mu\overleftrightarrow{D}_\nu\psi, \tag{5.24}$$

where $\mathcal{S}$ takes the symmetric traceless part of the tensor:

$$\mathcal{S}T_{\mu\nu} = \frac{1}{2}\left(T_{\mu\nu} + T_{\nu\mu}\right) - \frac{1}{4}\delta_{\mu\nu}T_{\alpha\alpha}. \tag{5.25}$$

Our decomposition for the RI′-MOM scheme is a scaled version of the one used by Gracey [165]:

$$\Lambda_{O_{\mu\nu}^V}^{(1)}(p,p) = \mathcal{S}\gamma_\mu p_\nu, \tag{5.26}$$

$$\Lambda_{O_{\mu\nu}^V}^{(2)}(p,p) = \mathcal{S}\frac{p_\mu p_\nu}{p^2}\slashed{p}, \tag{5.27}$$

where here and below we neglect tensors of opposite chirality. For RI-SMOM, the derivative in the operators basis used by Gracey [166] did not yield a definite C-symmetry, unlike our operator containing $\overleftrightarrow{D}$. This allows us to use half as many tensors as Gracey; see also [167, 168]. Defining $\overline{p} = (p' + p)/2$, our tensors and their relation to Gracey's tensors $P_{(i)}^{W_2}$ are the following:

$$\Lambda_{O_{\mu\nu}^V}^{S(1)}(p',p) = \mathcal{S}\overline{p}_\mu\gamma_\nu = \frac{1}{4}\left(P_{(2)}^{W_2} - P_{(1)}^{W_2}\right), \tag{5.28}$$

$$\Lambda_{O_{\mu\nu}^V}^{S(2)}(p',p) = \mathcal{S}\frac{\overline{p}_\mu\overline{p}_\nu}{\overline{p}^2}\slashed{\overline{p}} = -\frac{1}{6}\sum_{i=3}^{8}(-1)^i P_{(i)}^{W_2}, \tag{5.29}$$

$$\Lambda_{O_{\mu\nu}^V}^{S(3)}(p',p) = \mathcal{S}\frac{\overline{p}_\mu q_\nu}{q^2}\slashed{q} = \frac{1}{2}\left(P_{(3)}^{W_2} - P_{(5)}^{W_2} + P_{(6)}^{W_2} - P_{(8)}^{W_2}\right), \tag{5.30}$$

$$\Lambda_{O_{\mu\nu}^V}^{S(4)}(p',p) = \mathcal{S}\frac{\overline{q}_\mu q_\nu}{q^2}\slashed{\overline{p}} = \frac{1}{2}\left(\sum_{i=3}^{5}P_{(i)}^{W_2} - \sum_{i=6}^{8}P_{(i)}^{W_2}\right), \tag{5.31}$$

$$\Lambda_{O_{\mu\nu}^V}^{S(5)}(p',p) = \mathcal{S}\frac{\overline{p}_\alpha q_\beta}{q^2}\gamma_{[\mu}\gamma_\alpha\gamma_{\beta]}\overline{p}_\nu = \frac{1}{2}\left(P_{(10)}^{W_2} - P_{(9)}^{W_2}\right), \tag{5.32}$$

where $\overline{p}^2 = \frac{3}{4}\mu^2$ and the square brackets denote antisymmetrization. The one-derivative axial operator,

$$O_{\mu\nu}^A = \mathcal{S}\overline{\psi}\tau_3\gamma_\mu\gamma_5\overleftrightarrow{D}_\nu\psi, \tag{5.33}$$

is related to the vector operator by chiral symmetry and its tensor structures correspond to those of the vector operator, multiplied by γ_5. We use the four-loop anomalous dimension [169, 170, 171][4] and three-loop matching to $\overline{\text{MS}}$ [165, 168].

The one-derivative tensor operator is

$$O_{\mu\nu\rho}^T = \mathcal{S}\overline{\psi}\tau_3\sigma_{\mu\nu}\overleftrightarrow{D}_\rho\psi, \tag{5.34}$$

where the symmetrization and trace subtraction has the form [165]

$$\mathcal{S}T_{\mu\nu\rho} = \frac{1}{2}\left(\tilde{T}_{\mu\nu\rho} + \tilde{T}_{\mu\rho\nu}\right) - \frac{1}{3}\delta_{\nu\rho}\tilde{T}_{\mu\alpha\alpha} + \frac{1}{6}\left(\delta_{\mu\nu}\tilde{T}_{\rho\alpha\alpha} + \delta_{\mu\rho}\tilde{T}_{\nu\alpha\alpha}\right) \tag{5.35}$$

with $\tilde{T}_{\mu\nu\rho} = \frac{1}{2}(T_{\mu\nu\rho} - T_{\nu\mu\rho})$. Choosing to start by antisymmetrizing $\mu\nu$ leaves us with only two tensor structures in the RI′-MOM scheme, compared with Gracey's three:

$$\Lambda_{O_{\mu\nu\rho}^T}^{(1)}(p,p) = \mathcal{S}\sigma_{\mu\nu}p_\rho, \tag{5.36}$$

$$\Lambda_{O_{\mu\nu\rho}^T}^{(2)}(p,p) = \mathcal{S}\frac{1}{p^2}\sigma_{\mu\alpha}p_\alpha p_\nu p_\rho. \tag{5.37}$$

---

[4]Ref. [172] reports that Ref. [170] contains a misprint.

For RI-SMOM, the supplementary data of Ref. [173] lists 30 structures. First antisymmetrizing $\mu\nu$ reduces this to 16 and charge conjugation further reduces the number to 8. We write the first six as

$$\Lambda^{S(1)}_{O^T_{\mu\nu\rho}}(p',p) = \mathcal{S}\sigma_{\mu\nu}\overline{p}_\rho, \tag{5.38}$$

$$\Lambda^{S(2)}_{O^T_{\mu\nu\rho}}(p',p) = \mathcal{S}\frac{1}{\overline{p}^2}\sigma_{\mu\alpha}\overline{p}_\alpha\overline{p}_\nu\overline{p}_\rho, \tag{5.39}$$

$$\Lambda^{S(3)}_{O^T_{\mu\nu\rho}}(p',p) = \mathcal{S}\frac{1}{q^2}\sigma_{\mu\alpha}q_\alpha q_\nu\overline{p}_\rho, \tag{5.40}$$

$$\Lambda^{S(4)}_{O^T_{\mu\nu\rho}}(p',p) = \mathcal{S}\frac{1}{q^2}\sigma_{\mu\alpha}q_\alpha\overline{p}_\nu q_\rho, \tag{5.41}$$

$$\Lambda^{S(5)}_{O^T_{\mu\nu\rho}}(p',p) = \mathcal{S}\frac{1}{q^2}\sigma_{\mu\alpha}\overline{p}_\alpha q_\nu q_\rho, \tag{5.42}$$

$$\Lambda^{S(6)}_{O^T_{\mu\nu\rho}}(p',p) = \mathcal{S}\frac{1}{q^2\overline{p}^2}\sigma_{\alpha\beta}\overline{p}_\alpha q_\beta\overline{p}_\mu q_\nu\overline{p}_\rho. \tag{5.43}$$

The last two tensors involve γ_5 or the identity, and they have vanishing trace with each of the first six, so they can be neglected. We use the three-loop anomalous dimension [165], the three-loop matching from RI′-MOM [165], and the two-loop matching from RI-SMOM [173].

5.C.2 Calculation

Our numerical setup follows Ref. [50], extended to include both sets of momenta on both ensembles. We use partially twisted boundary conditions, namely periodic in time for the valence quarks. The plane wave sources are given momenta either along the four-dimensional diagonal $p^{(')} = \frac{2\pi}{L}(k,k,k,\pm k)$ or along the two-dimensional diagonal $p,p' \in \{\frac{2\pi}{L}(k,k,0,0), \frac{2\pi}{L}(k,0,k,0)\}$, with $k = 2,3,\ldots,\frac{L}{4a}$. By contracting them in different combinations, we get data for both RI′-MOM kinematics, $p'-p=0$, and RI-SMOM kinematics, $p'-p = \frac{2\pi}{L}(0,0,0,\pm 2k)$ or $\pm\frac{2\pi}{L}(0,k,-k,0)$. We used 54 gauge configurations from each ensemble; however, on one configuration of the coarse ensemble the valence twisted boundary condition yielded a near-singular Dirac operator and the multigrid solver was unable to converge. We omitted this configuration, using only 53 on the coarse ensemble.

Perturbatively matching from RI′-MOM or RI-SMOM to the $\overline{\text{MS}}$ scheme and evolving to scale 2 GeV does not eliminate dependence on the initial scale μ: there are remaining effects from lattice artifacts, truncation of the perturbative series, and nonperturbative contributions. To control these artifacts, we perform fits including terms polynomial in μ^2 and also a pole term, following Ref. [174]. These fits have the form $A + B\mu^2 + C\mu^4 + D/\mu^2$, where the constant term A serves as our estimate of the renormalization factor ratio Z_O/Z_V. Results for this ratio, choosing one irrep for each of the three operator types, are shown in Fig. 5.13. For each operator and scheme, we perform three fits: using the 4D data with two different fit ranges $\mu^2 \in [4,20]$ and $[10,30]$ GeV^2 and using the 2D data with $\mu^2 \in [4,15]$ GeV^2. In some cases (particularly using the very precise RI-SMOM data), the fit quality is very poor and thus we scale the statistical uncertainty by $\sqrt{\chi^2/\text{dof}}$ whenever this is greater than one. Following the same prescription as in Ref. [50], we combine the results first within each scheme and then for our final result using both schemes, estimating the systematic uncertainty (which is dominant) from the scatter of results and conservatively taking the maximum statistical uncertainty. In all cases, there is good agreement between the two schemes.

Renomalization-group-invariant ratios of renormalization factors in different irreps $Z^{\rho'}_O/Z^{\rho}_O$ are shown in Fig. 5.14. Note that it is not possible to isolate the chosen $O(4)$-invariant function for the tensor operator in irrep $\tau_2^{(8)}$ in the RI-SMOM scheme using the 4D kinematics. Because in the continuum and infinite volume these ratios are independent of μ^2, we can fit using much lower momenta and only avoid the region $\mu^2 < 1$ GeV^2 due to finite-volume effects. We also omit the pole term, i.e. set $D = 0$. In all cases, we obtain results within a few percent of unity.

Our final values for the ratios of renormalization factors are given in Table 5.4.

94
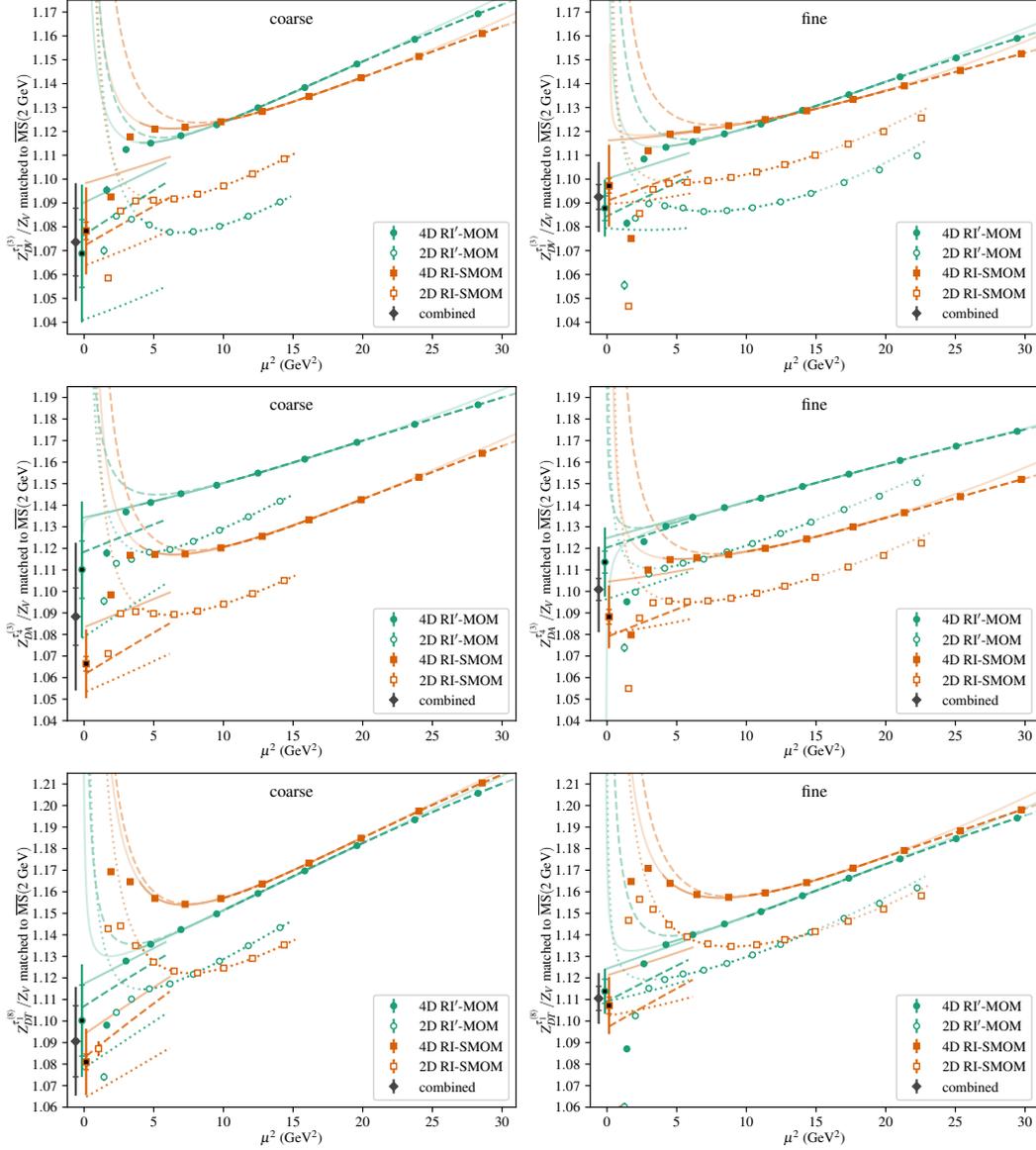
Figure 5.13: Ratios of renormalization factors $Z_{DV}^{\tau_1^{(3)}}/Z_V$, $Z_{DA}^{\tau_4^{(3)}}/Z_V$, $Z_{DT}^{\tau_1^{(8)}}/Z_V$ on the coarse (left) and fine (right) ensembles, determined using the RI′-MOM (green circles) and RI-SMOM (orange squares) intermediate schemes together with momenta along the four-dimensional diagonal (filled symbols) and two-dimensional diagonal (open symbols) and then matched to $\overline{\text{MS}}$ at scale 2 GeV. For most points, the statistical uncertainty is smaller than the plotted symbol. The solid curves are fits to the 4D data in the μ^2 range from 4 to 20 GeV^2, the dashed curves in the range from 10 to 30 GeV^2, and the dotted curves are fits to the 2D data in the range from 4 to 15 GeV^2. The fit curves without the pole term are also plotted in the range $0 < \mu^2 < 6$ GeV^2. To reduce clutter, uncertainties on the fits are not shown. The symbols filled with black near $\mu^2 = 0$ provide the final estimat for each intermediate scheme; their outer (without end cap) and inner (with end cap) error bars show the total and statistical uncertainties. The filled dark gray diamonds are the final estimates that combine both schemes.

Figure 5.14: Scale and scheme-invariant ratios of renormalization factors $Z_{DV}^{\tau_3^{(6)}}/Z_{DV}^{\tau_1^{(3)}}$, $Z_{DA}^{\tau_4^{(6)}}/Z_{DA}^{\tau_4^{(3)}}$, and $Z_{DT}^{\tau_2^{(8)}}/Z_{DT}^{\tau_1^{(8)}}$, determined using the RI'-MOM (green circles) and RI-SMOM (orange squares) intermediate schemes together with momenta along the four-dimensional diagonal (filled symbols) and two-dimensional diagonal (open symbols) and then matched to $\overline{\mathrm{MS}}$ at scale 2 GeV. For most points, the statistical uncertainty is smaller than the plotted symbol. The solid curves are fits to the 4D data in the μ^2 range from 1 to 8 GeV^2 and the dotted curves are fits to the 2D data in the range from 1 to 5 GeV^2. To reduce clutter, uncertainties on the fits are not shown. The symbols filled with black near $\mu^2 = 0$ provide the final estimate for each intermediate scheme; their outer (without end cap) and inner (with end cap) error bars show the total and statistical uncertainties. The filled dark gray diamonds are the final estimates that combine both schemes.

	coarse	fine
Z_V	0.9094(36)	0.9438(1)
$Z_{DV}^{\tau_1^{(3)}}/Z_V{}^1$	1.0736(142)(202)	1.0925(52)(137)
$Z_{DV}^{\tau_3^{(6)}}/Z_{DV}^{\tau_1^{(3)}}$	1.0232(36)(63)	1.0167(29)(27)
$Z_{DA}^{\tau_4^{(3)}}/Z_V{}^1$	1.0883(113)(316)	1.1009(51)(192)
$Z_{DA}^{\tau_4^{(6)}}/Z_{DA}^{\tau_4^{(3)}}$	1.0058(28)(50)	1.0074(40)(16)
$Z_{DT}^{\tau_1^{(8)}}/Z_V{}^1$	1.0906(165)(191)	1.1105(56)(104)
$Z_{DT}^{\tau_2^{(8)}}/Z_{DT}^{\tau_1^{(8)}}$	1.0034(35)(38)	1.0016(134)(19)

Table 5.4: Vector current renormalization factor from Ref. [50] and ratios of renormalization factors computed in this work.

Chapter 6

Summary

This thesis included work on a set of topics. First, in chapter 2 and 3 with an extension in the appendix A, aspects of the sign problem within the Hubbard model were discussed to extend the computable parameter space of a Monte Carlo simulation. Second, in chapter 3, these results where used to get a first understanding of the total system charge of the molecule $C_{20}H_{12}$ perylene as a function of chemical potential at a hot temperature. During this work, it became clear that a non automatic procedure dealing with the sign problem can be used for single ensembles but needs further improvements to be used in a full parameter scan where $O(100)$ ensembles are required. Consequently, in chapter 4, the sign problem reduction is done only by a simple shift, that has been discussed in an adjacent work [7]. In this chapter, a greater focus was placed on the analysis of physical observables for the molecule $C_{20}H_{12}$ perylene extending the total charge calculation of chapter 3 and additionally extracting the single particle spectrum. Finally, in chapter 5, the focused shifted towards a Monte Carlo study of aspects on hadron structure, namely moments of parton distribution functions.

6.1   Software

In addition to the content and results of this thesis, software had to be implemented to perform the required calculations and analysis. I want to point out two major packages, that were mainly developed by me.

First, the software package "Nanosystem Simulation Library", or short NSL [10] provides a full GPU enabled package to perform Monte Carlo simulations of low dimensional systems. It uses an abstraction layer above the C++ implementation of PyTorch [90], called libtorch, to handle memory management as well as linear algebra operations. This package implements various discretizations for the Hubbard model, in particular the exponential discretization of the fermion matrix, but is not limited to this model. An entire infrastructure is created to facility all kinds of action with a particular focus on usage-simplicity. This infrastructure defines possible actions, a handling for runtime parameters and file-IO as well as implementations of HMC with more Monte Carlo algorithms in development phase. It is intended as a replacement for isle [175] that was used in chapter 2. One current draw back is in the realm of small systems, where additional overhead created by the interface to libtorch, slows down computation. However, especially the possibility to go onto GPU allows for a improved volume scaling compared to isle [175].

Second, I developed a correlator analysis program [11] based on lsqfit [132] and ideas of model averaging through the Baysian Akkaike information criterion, see appendix 4.A.6 for details. This work was done together with Dr. Giovanni Pederiva, who helped debugging and testing the code. We found, that a model averaging approach on well defined criterions is beneficial compared to a discussion on best fits, as it automatically includes systematic uncertainties from the model choice.

6.2 Aspects of the Sign Problem in the Hubbard Model

As mentioned, the first chapters of the thesis discuss the sign problem of the Hubbard model. At non-zero chemical potential, the Hubbard action, as derived in (1.37), is generally complex valued causing the Boltzmann weight to oscillate strongly. Using reweighting, it is still possible to estimate observables, however, at the cost of reduced statistics, compare equation (1.9). Mitigating the sign problem by means of increasing the statistical power is still a major target of algorithmic development. In particular, in the first two chapters I tested and discussed a method proposed by Wynen et.al. [61]. A neuronal network, that intrinsically complexifies configurations, is defined to learn a path deformation towards Lefschetz thimbles, compare equations (1.44) and (1.45). The proposed architecture are paired affine coupling layers with complex valued trainable parameters, i.e. the weights and biases. Chapter 2 shows that the supervised training of this architecture allows for the successful simulation of systems with increasingly severe sign problems. The discussed machine learning enhanced HMC approach reduces the sign problem sufficiently and enjoys a statistical power much greater than original real-plane or even tangent-plane HMC. This is explicitly shown in figure 2.3. Further the effect is shown on correlators, where the statistical noise is drastically reduced, compare figure 2.4. We demonstrated the fidelity and correctness of our method by simulating 2, 4 and 8 site models and comparing our results to that obtained from direct diagonalization, obtaining excellent agreement. A note about reweighting is added in this thesis arguing that the non-holomorphic contribution to the Jacobian is negligible to obtained precision. The tested method is then applied to the 18 sites problem, a small piece of a graphene sheet, where direct diagonalization is not possible. The correlators, show significantly reduced statistical noise compared to the tangent plane, arguing for possible extended parameter space. Adapting the reweighting of section 2.6, shows favorable volume scaling for the generation of configurations.

To avoid naively calculating Jacobian determinants on measurements, a further test is executed using holomophic architectures in chapter 3. Additionally, further improvements on the generation of data are implemented that track the explosion of the real part of the action, compare algorithm 3.2. This allows to maintain a more relevant training data set compared to chapter 2 as the configurations with smaller action have more weight in the path integral. Unfortunately, all these training procedure require a considerable amount of testing making this approach hard to use for a large number of ensembles. In the future, I aim to tackle this problem tying to identify a more automatic procedure of training. The tested system, $C_{20}H_{12}$ perylene, is slightly larger than the 18 site problem from chapter 2, but does not require a infinite volume, in space, limit. This makes it a great target to test the developed methods for obtaining physically relevant observables that classify the molecule. Therefore, for a set of chemical potentials at fixed temperature β and distretization N_t, are simulated and correlators are obtained. Overall, the reduction of statistical noise is self-evident as seen in figures 3.3 and 3.4. Finally, effective masses are extracted, from the correlators, as a function of chemical potential μ, compare figures 3.5a to 3.5d. This provides a first estimate of the single particle spectrum. In conclusion, the combination of measures against the sign problem, namely the tangent plane and the path deformations through neural networks, can be effective to perform a charge scan of this molecule. However, for a full continuum study constant, semi-analytic, shifts are beneficial due to the large number of ensembles.

6.3 The Single Particle Spectrum of Perylene

All this work converges, as presented in chapter 4, in an ab initio study of the single particle spectrum of perylene described by the Hubbard model. As mentioned, the sign problem is reduced by simulating on the next-to-leading order plane, compare equation (1.51), which allows to compute the molecule at chemical potentials $\mu = 0, 0.1, \cdots, 1.1$. The statistical power, shown in figure 4.2, quantifies the severity of the sign problem as a function of the chemical potential at different values for discretizations, N_t, and inverse temperatures, β, on this simple manifold.

The focus of this study lays on the continuum analysis of physical quantities. To control the continuum extraploation, $\delta \to 0$, three discretizations, $N_t = 32, 64, 96$, are chosen. All these are calculated at three different inverse temperatures, $\beta = 4, 6, 8$, to provide a rough idea on the temperature dependence. All continuum extrapolations, exemplarily shown on the single particle spectrum in figure 4.7 and on the charge in figure 4.11, show little dependence on the discretization suggesting that the chosen values for N_t are already far in the scaling regime for the given precision.

All observables are computed non-interactingly, i.e. $U = 0$, and compared with the interacting, $U = 2$, result. The charge is explicitly quantified as a function of the chemical potential, see figure 4.3. Considering the non-interacting result for the charge, one finds an exact jump at $\mu = 0.34730$. For the interacting theory, the charge jump is suggested between $\mu = 0.4$ and 0.5. Furthermore, comparing interacting results and non-interacting ones shows stronger discrepancy with increasing chemical potential. After quantifying the meaning of chemical potential with the total system charge, Then a variational basis of single particle operators is computed from which the single particle spectrum is estimated. This requires a fitting procedure, that works automatically, as the number of correlators to fit is large. A Baysian fitting procedure with model averaging based on the Akkaike information criterion is employed. The continuum single particle energies are mapped as a function of chemical potential, compare figures 4.8 and 4.9. It can be observed that a negative energy crosses zero between chemical potentials of 0.4 and 0.5 as suggested by the total charge. Notably, the energies show significant divergence from the non-interacting result growing towards larger chemical potential.

6.4   Moments of Parton Distribution Functions

Parallel, to the work in the Hubbard model, interesting questions arise about the structure of a nucleon. In the last chapter, 5, an analysis for the second Mellin moment, $\langle x \rangle$, of the unpolarized, polarized, and transversity parton distribution functions using two lattice QCD ensembles at the physical pion mass. These ensembles were generated using a tree-level Symanzik-improved gauge action and 2+1 flavor tree-level improved Wilson Clover fermions coupling via 2-level HEX-smearing. Forward matrix elements of a set of local leading twist operators are used to determine the moments and renomalization factors in are determined in RI-(S)MOM and match to $\overline{\text{MS}}$ at scale 2GeV.

Certain operators have only a contribution at non vanishing momentum due to the kinematic factors relating moments and matrix elements. I identified operators that, at finite momentum, have remarkably small excited-state contamination, compare figure B.2. Excitingly, these have vanishing kinematics at zero momentum.

A continuum extrapolation can not reliably be performed, given that only two ensembles are available. Regardless, within a Bayesian approach a continuum estimate can be obtained, by treating the results of the individual ensemble as Gaussian distributions. Naturally, this estimate is strongly biased by the chosen priors. After, a relatively conservative choice of priors, a continuum value is extracted that agrees with the two individual ensembles points, resulting in $\langle x \rangle_{u^+ - d^+} = 0.200(17)$, $\langle x \rangle_{\Delta u^- - \Delta d^-} = 0.213(16)$, and $\langle x \rangle_{\delta u^+ - \delta d^+} = 0.219(21)$. Furthermore, the isovector longitudinal quark spin-orbit correlation in the proton is obtained by using the moment of the polarized PDF, $2L_\ell^q S_\ell^q = -0.393(08)$. This is in good agreement with previous results based on GTMDs [151].

Appendix A

Mitigating the Hubbard sign Problem. A Novel Application of Machine Learning

This appendix chapter is based on [5]:

Marcel Rodekamp and Christoph Gäntgen. "Mitigating the Hubbard Sign Problem. A Novel Application of Machine Learning." In: vol. LATTICE2022. 2023. DOI: 10.22323/1.430.0032. arXiv: 2211.09584

In this appendix chapter the results of chapter 2 are extended. Several systems of strongly correlated electrons described by the Hubbard model are discussed. Due to the sign problem of these systems, the use of machine learning enhanced HMC (ML HMC) is demonstrated. The difference between the original formulation of [61] and the network discussed in chapter 2 is discussed. To complete the story, some results from chapter 2 are reused as indicated. Additional ensembles, calculated for the master thesis [176], is reanalyzed, extending the 18-site model analysis over a wide range of chemical potentials. Additional examples, the C_{20} and C_{60} buckminsterfullerene molecules, are added. These have a sign problem even at zero chemical potential due to their non-bipartite geometry. The investigation on these molecules was later continued in [7] using the NLO plane and optimized shifts.

A.1   Introduction

Markov Chain Monte Carlo algorithms (MCMC) enjoy great success in simulating many theories from the Ising model up to Lattice QCD. Albeit the potential, MCMC has a hard time whenever the action becomes complex-valued due to the associated Boltzmann weight loosing its interpretability as probability distribution.

Using MCMC, we focus on the Hubbard model capturing electronic properties of systems with strongly interacting electrons propagating on a fixed spatial lattice of ions. Examples for such systems are carbon nano structures like Graphene and Fullerene C_n. In the Hubbard model the sign problem is observed at finite chemical potential as well as on non-bipartite lattices[1]. Reweighting can treat the complex-valued Boltzmann weight though, at the same time, introducing an exponentially hard to estimate normalization.

Deforming the region of integration onto Manifolds with an almost constant imaginary action showed great promise in reducing the sign problem substantially [62, 23, 63, 64]. Practically, this deformation requires numerical integration of differential equations which becomes infeasible for larger systems. We aim to identify efficient Neural Network architectures to learn such beneficial deformations. This removes the cost of numerically integrating configurations and enables simulations of large systems with a sign problem beyond the standard reweighting approach.

In this proceedings, we collect material from our earlier publications [61, 1] and a master thesis [176]. The manuscript is organized in the following way. In section A.2 a brief introduction to the Hubbard model and the tested system is presented along a short discussion of the sign problem. This discussion is then followed by the definition of the neural network architectures as published in [61, 1]. Further, in section A.3 correlator results are presented and we discuss the obtained charge density of one of the larger systems.

A.2   Formalism

The Hubbard model [66, 30, 31] describes a fixed spatial lattice X of ions on which electrons can move and interact. Its Hamiltonian, in particle-hole basis, is

$$\mathcal{H}[K, V, \mu] = - \sum_{x,y \in X} \left(p_x^\dagger K^{xy} p_y - h_x^\dagger K^{xy} h_y \right) + \frac{1}{2} \sum_{x,y \in X} \rho_x V^{xy} \rho_y + \mu \sum_{x \in X} \rho_x,$$

where the amplitudes in K encode the hopping of fermionic particles p and holes h, the potential V describes the interactions between charges

$$\rho_x = p_x^\dagger p_x - h_x^\dagger h_x \tag{A.1}$$

and the chemical potential μ incentivizes charge.

---

[1]Bipartite describes lattice geometries at which the sites can be two coloured such that no neighbouring sites have the same colour. For example, the square is bipartite while the triangle is non-bipartite.

(a) 2 Sites   (b) 8 Sites   (c) 18 Sites (boundary suppressed)

(d) Fullerene C_{20} (e) Fullerene C_{60}
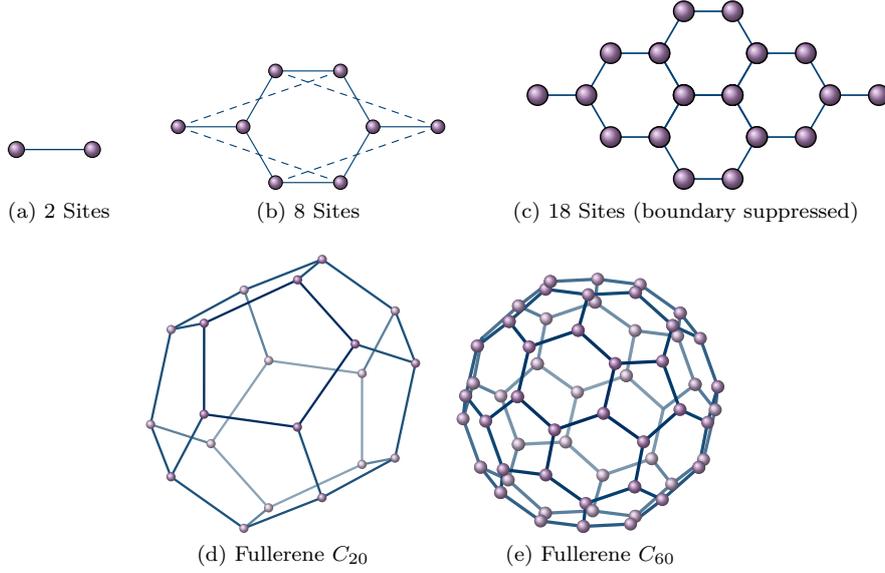
Figure A.1: Showing the different geometries considered in this proceedings. Each vertex represents an ion and each (dashed) line depicts the nearest-neighbor hopping that is allowed by the Hubbard model. Dashed lines indicate periodic boundary condition where possible.

By adjusting the hopping and lattice symmetry K as well as the interaction V this model can describe a wide variety of physical systems. In the following investigation, five systems are considered as displayed in figure A.1. The 2 site system describes one unit cell of the honeycomb structure used for Graphene type models which we successively built up with the 8 and 18 site ones. Secondly, we present preliminary results for fullerenes C_{20} and C_{60} at zero chemical potential. In all cases K encodes nearest-neighbor hopping and we assume an on-site interaction,

$$K = \kappa \delta_{\langle xy \rangle} \quad \& \quad V = U \delta_{xy}. \tag{A.2}$$

In figure A.1 the sites, i.e. ions and their nearest neighbor connections are depicted. Lines stretching out display periodic boundary of the spatial lattice (suppress in the 18 site case).

A.2.1 Simulation Setup

Calculating observables follows the standard procedure [24] of evaluating the thermal trace. After trotterizing into N_t time slices, inserting Grassmannian resolutions of the identity and linearizing the interaction via the Hubbard-Stratonovich transformation [29] the Hamiltonian is transformed into the action

$$S[\Phi \,|\, K, V, \mu] = -\log\det\{M[\Phi \,|\, K, \mu] \cdot M[-\Phi \,|\, -K, -\mu]\} + \frac{1}{2}\sum_{t=0}^{N_t-1}\sum_{x,y \in X} \Phi_{tx} \delta V^{-1\,xy} \Phi_{ty}, \tag{A.3}$$

where $\Phi \in \mathbb{R}^{|\Lambda|}$ is an auxiliary field on the spacetime lattice $\Lambda = [0, N_t - 1] \otimes X$ and $\delta = \beta/N_t$ is the (temporal) lattice spacing controlled by the inverse temperature β. The fermion matrix is not uniquely defined on the lattice, we choose the exponential discretization [19]

$$M[\Phi \,|\, K, \mu]_{x't';xt} = \delta_{x'x}\delta_{t't} - \left(e^{\delta(K+\mu)}\right)_{x'x} e^{+i\Phi_{xt}} \mathcal{B}_{t'} \delta_{t'(t+1)} \tag{A.4}$$

where $\mathcal{B}$ encodes the anti-periodic boundary conditions in time. For bipartite systems we may replace $-K$ with $+K$ in the holes' fermion matrix [24]. The thermal trace for this is expressed as path integral

$$\langle O \rangle = \frac{1}{\mathcal{Z}} \int \mathcal{D}[\Phi]\, e^{-S[\Phi]} O[\Phi] = \int \mathcal{D}[\Phi]\, p_S[\Phi]\, O[\Phi] \tag{A.5}$$

For cases of real action we can apply MCMC to generate N_{cfg} configurations according to the Boltzmann distribution $p_S[\Phi]$ and estimate observables (A.5) by an unweighted expectation value. If the action is complex valued we apply reweighting

$$\langle O \rangle = \frac{\left\langle e^{-i\,\text{Im}\{S\}} O \right\rangle_{\text{Re}\{S\}}}{\left\langle e^{-i\,\text{Im}\{S\}} \right\rangle_{\text{Re}\{S\}}} = \frac{1}{\Sigma} \left\langle e^{-i\,\text{Im}\{S\}} O \right\rangle_{\text{Re}\{S\}}. \tag{A.6}$$

and sample configurations according to the Boltzmann distribution under the real part of the action. It has been shown [20] that an effective number of configurations

$$N_{\text{cfg}}^{\text{eff}} = |\Sigma|^2 \cdot N_{\text{cfg}} \tag{A.7}$$

controls the scaling of statistical errors $\sim \left(N_{\text{cfg}}^{\text{eff}} \right)^{-1/2}$. This translates the sign problem to the ability of calculating the denominator Σ, i.e. the statistical power, reliably [20, 19, 54, 71].

A promising approach to mitigate, or even remove, the sign problem is to deform the region of integration $\Phi \in \mathcal{M}_{\mathbb{R}} = \mathbb{R}^{|\Lambda|}$ onto a manifold on which the imaginary part of the action is (nearly[2]) constant [73, 23], $\mathcal{M} = \{\Phi \in \mathbb{C}^{|\Lambda|} \mid \text{Im}\{S[\Phi]\} = const.\}$. If $\mathcal{M}$ is in the same homology class as $\mathcal{M}_{\mathbb{R}}$ an analogue of the Cauchy integral theorem ensures that the observables are unchanged. Parametrizing fields $\Phi(\phi) \in \mathcal{M}$ then adds a Jacobian defining the used effective action

$$S[eff]\,\phi = S[\Phi(\phi)] - \log \det\{J[\phi]\}, \quad J_{ij} = \frac{\partial \Phi(\phi)_i}{\partial \phi_j} \tag{A.8}$$

There are many strategies for picking target manifolds $\mathcal{M}$ [74]. One choice is to try to approximate the Lefschetz thimbles [73]. Each thimble contains a critical point Φ_{cr} that is a fixed point of the holomorphic flow

$$\frac{d\Phi(\tau)}{d\tau} = \pm \left(\frac{\partial S[\Phi(\tau)]}{\partial \Phi(\tau)} \right)^*, \quad \Phi(0) = \phi \tag{A.9}$$

introducing the fictitious flow time τ. A thimble is the set of complexified configurations that flow to a critical point under downward flow, i.e. $-$ in (A.9). An integrator for (A.9) will always be computationally expensive [23, 1, 61]. However, the non-interacting solution $\phi = 0$ for (A.9) assuming a constant field $\Phi_{t,x} = \Phi_{t',x'}$ defines a (hyper-) plane parallel to the real plane $\mathcal{M}_{\mathbb{R}}$ that goes through the main critical point $i\Phi_c^0$. This so called tangent plane $\mathcal{M}_T = \{\phi + i\Phi_c^0 \mid \forall \phi \in \mathcal{M}_{\mathbb{R}}\}$ showed promise in sufficiently mitigating the sign problem, at least for smaller systems [54, 78, 79, 61, 1].

Neural Network Architectures

To improve beyond the tangent plane it seems plausible to identify a transformation that transforms a given configuration $\phi \in \mathcal{M}_{\mathbb{R}}$ to a target manifold $\tilde{\mathcal{M}}$ that may be closer to $\mathcal{M}$ than the tangent plane. Such a transformation may be parametrized by a neural network

$$\text{SHIFT} : \mathcal{M}_{\mathbb{R}} \to \tilde{\mathcal{M}}, \ \phi \mapsto \phi + i\left(\Phi_c^0 + \mathcal{NN}(\phi) \right). \tag{A.10}$$

---

[2]If $\text{Im}\{S\} \approx const$, the statistical power $|\Sigma| = \left| \left\langle e^{i\,\text{Im}\{S[eff]\}} \right\rangle \right| \approx \left| \left\langle e^{i \cdot const} \right\rangle \right| \approx 1$ yielding nearly no reduction in effective number of configurations $N_{\text{cfg}}^{\text{eff}} \approx N_{\text{cfg}}$.

For the neural network part $\mathcal{NN}$ we pick two linear layers $\omega\phi + b$ with real trainable weights ω and biases b which are separated by a leacky-ReLU. As the effective action (A.8) suggests the defining transformations Jacobian needs to be computed

$$\log\det\{J_{\mathrm{SHIFT}}[\phi]\} = \log\det\left\{\mathbb{1} + i\frac{\partial\mathcal{NN}(\phi)}{\partial\phi}\right\} \tag{A.11}$$

which requires an $O(V^3)$ algorithm for the determinant calculation. This scaling is not feasible for large scale systems but still much cheaper than applying a Runge-Kutta — or similar algorithms — to integrate the holomorphic flow equations.

To improve the scaling, we identify a neural network that has a cheaper determinant. One such architectures is given by Affine Coupling Layers [88, 89] that approximate the integrator $\Phi(\phi) \approx \mathcal{NN}(\phi)$

$$\mathcal{NN}_\ell(\Phi) = \begin{cases} c_\ell\,[\Phi_A,\,\Phi_B] & A_\ell \text{ components} \\ \Phi_B & B_\ell \text{ components} \end{cases} \tag{A.12}$$

Here A and B are layer-specific partitions of the input vector Φ of equal cardinality $1/2|\Lambda|$. $\Phi_{A,B}$ are the components of the input belonging to the indicated partition. We apply the affine coupling [88]

$$c_\ell\,[\Phi_A, \Phi_B] = e^{m_\ell(\Phi_B)} \odot \Phi_A + a_\ell\,(\Phi_B)\,. \tag{A.13}$$

The coupling functions $m_\ell, a_\ell : \mathbb{C}^{|\Lambda|/2} \to \mathbb{C}^{|\Lambda|/2}$ are again two linear layers separated by the non-linear Softsign function. To ensure that the neural network produces a complex configuration as is required by the holomorphic flow, the weights and biases need to be complex valued which is discussed in more detail in [1]. A single layer just transforms half of the configuration, we thus pair it up with a second layer that is set up in the same way but with the roles of A and B interchanged. This setup allows to express the Jacobian, with $L/2$ of these pairs, in the form

$$\log\det\{J_{\mathcal{NN}}(\phi)\} = \sum_{\ell=1}^{L}\sum_{i=0}^{|A|-1} m_\ell\,(\Phi_{\ell-1}(\phi)_B)_i\,. \tag{A.14}$$

which adds only an $O(V)$ cost to the application of the transformation.

For any of these architectures we perform Molecular Dynamics on $\mathcal{M}_\mathbb{R}$ using a standard leapfrog algorithm and then apply the network to move onto $\tilde{\mathcal{M}}$. Accept/Reject is then performed according to the effective action (A.8) using the transformed configuration from the Network and the Jacobian defined by the network. This machine learning enhanced Hybrid Monte Carlo is referred to as ML HMC.

A.2.2 Observables

We are interested in the electronic properties of a given system. Euclidean correlation functions of a single particle or a single hole

$$C_{xy}^{\overset{p}{h}}(t) = \left\langle p_x^\dagger(t)p_y(0)\right\rangle = \left\langle M[\pm\Phi|\pm K, \pm\mu]_{xt;y0}^{-1}\right\rangle, \tag{A.15}$$

momentum projected and averaged we obtain $C_{sp}(t)$ [1]. In the future we aim to match the parameters $U/\kappa, \mu$ to real-world systems and extract the band-gap $C_{xy}(t) \sim e^{-t\cdot\Delta E}$. Further, the charge density is defined by

$$\rho(\mu) = \frac{1}{|X|}\sum_{x\in X}\langle\rho_x\rangle = \frac{1}{|X|}\sum_{x\in X} C_{x,x}^p(0) - C_{x,x}^h(0). \tag{A.16}$$

It is point symmetric around the electric neutral half-filling point, $\mu = 0$, due to particle-hole symmetry. For very large $\mu \to \pm\infty$ the charges (+) or holes (-) are favoured yielding a charge density of ± 1. Qualitatively, it is expected that the system's charge equals integer multiples of the electric charge $n \cdot e^-$ with $n \in [-N_x, N_x]$, i.e. $\rho(\mu) = n/N_x$.
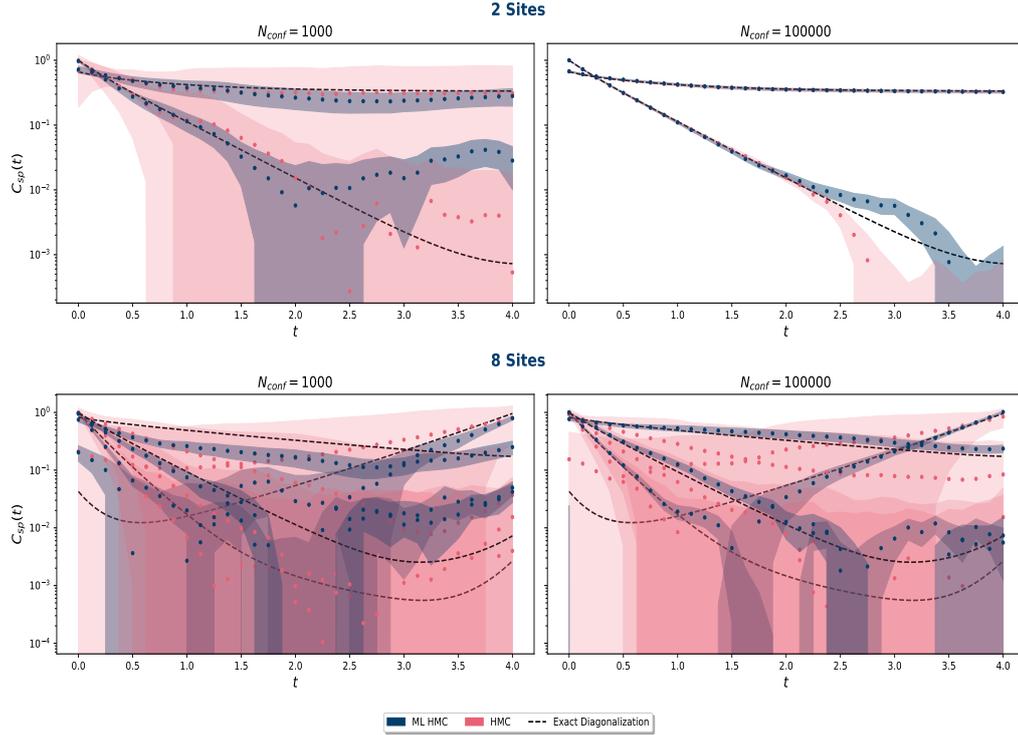
Figure A.2: Each row in this figure shows the correlators of a system with 2 (upper row) and 8 (lower row) sites [1]. The different columns correspond to a number of configurations, $N_{cfg} \in \{1000, 100\,000\}$, used to estimate the correlators. Three methods — ML HMC with coupling layer $\mathcal{NN}$ (blue), HMC on the Tangent Plane (red) and exact diagonalization (black) — are compared to show the effectiveness and correctness of the introduced machine learning enhanced method. The sign problem manifests as a loss of signal, i.e. small number of effective configurations N_{cfg}^{eff} (A.7), and greatly increases as the number of sites expands. It can be seen that the ML HMC has a substantially reduced sign problem.

A.3 Numerical Results

We experimented with different training setups. Foremost, supervised training using ADAM to minimize the $L1-$Loss. The training data consists out of $\mathcal{O}(10\,000)$ pairs $(\phi, \Phi(\phi))$ obtained by a Runge-Kutta of 4th order. For further details consider [61, 1].

In figure A.2 correlators for systems with 2 and 8 sites are compared using the different algorithms HMC (red) — on the tangent plane — ML HMC (blue), applying the coupling network $\mathcal{NN}$, and exact diagonalization (black) [1]. We use $N_{cfg} \in \{1000, 100\,000\}$ to portray the effect of the statistical power on the effective number of configurations. Corresponding statistical powers $|\Sigma|$ can be found in [1]. The system parameters — $N_t = 32$, $\beta = 4$, $U = 4$ and $\mu = 3$ — are kept fix. The ML HMC outperforms the HMC in terms of signal. The 8 site system has a stronger sign problem to an extend that HMC retrieves no signal. If a signal is obtained, both algorithms agree with the exact diagonalization. In figure A.3 the correlators for the system with 18 sites are displayed [1]. The system is computed at $N_t = 32$, $\beta = 4$, $U = 3$ and $\mu = 3$. The sign problem is much stronger than in the previous cases due to the larger volume. Nevertheless the ML HMC extracts a good signal for the correlators. Similar to the 8 site case HMC can't keep up. Due to the number of sites exact diagonalization is not feasible.

Continuing the 18 site model — with $U = 2$, $\beta = 5$, $N_t = 32$ — we want to study the charge density (A.16) subjected to the chemical potential. This can be seen in figure A.4 [176]. We compare tangent plane HMC (red) and ML HMC (blue) using the SHIFT network. Varying the chemical potential has shown that for small and large values the sign problem is mild (on the tangent plane).

18 Sites
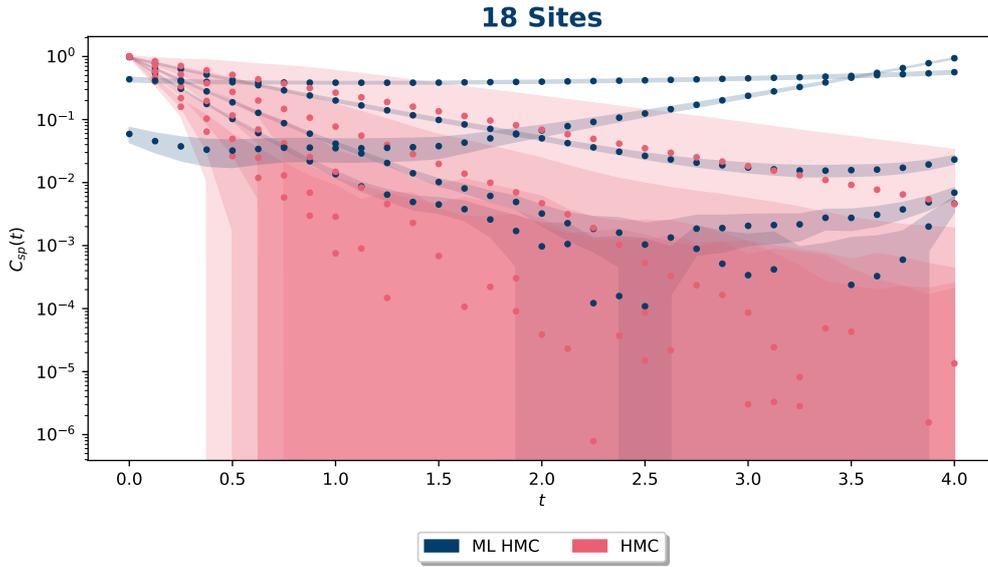
Figure A.3: The correlators of Graphene with 18 ions are shown [1]. 100 000 measurements are taken. With this larger lattice direct diagonalization as in figure A.2 is not tractable any more hence only the two statistical methods ML HMC using the coupling network $\mathcal{NN}$ (blue) and HMC on the tangent plane (red) are compared. As expected the ML HMC improves the signal drastically.
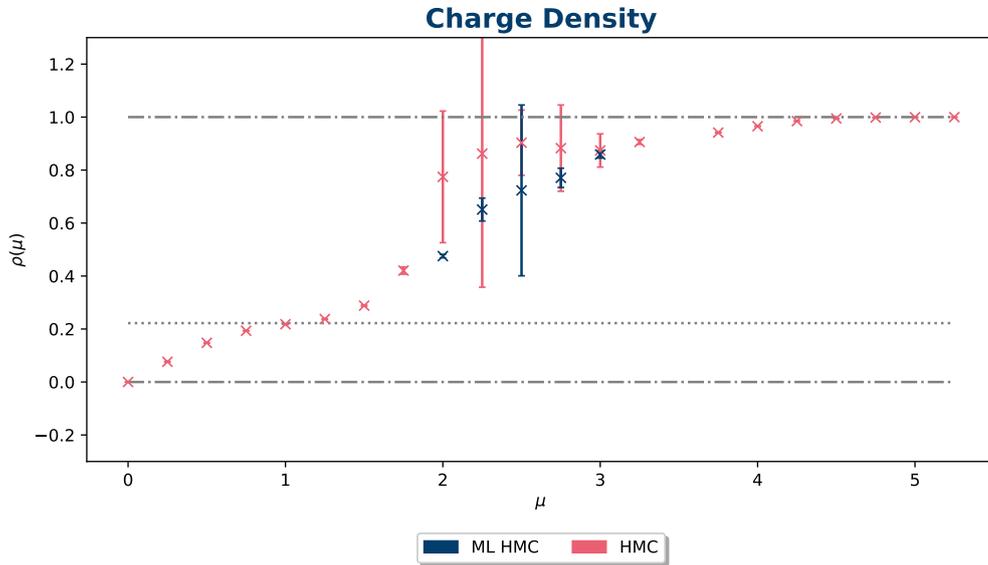
Charge Density

Figure A.4: We computed the charge density for several values of the chemical potential $\mu \in [0, 5.2]$ for an 18 site Graphene sheet [176]. For most smaller and larger values of μ the sign problem is small enough that estimation with HMC on the tangent plane (red) is sufficient. However, in the region $\mu \in [2, 3]$ an ML HMC (blue) is used to narrow particular values for which the sign problem becomes untraceable. The features at $\mu = 0$ and $\mu \to \infty$ are captured as expected. Finding the charge plateaus at $\rho(\mu) \sim n$, however, is yet unavailable due to the small β. The dashed line at $\rho(\mu) = {}^4/_{18}$ indicate an expected plateau which may be surmised around $\mu \approx 1$.
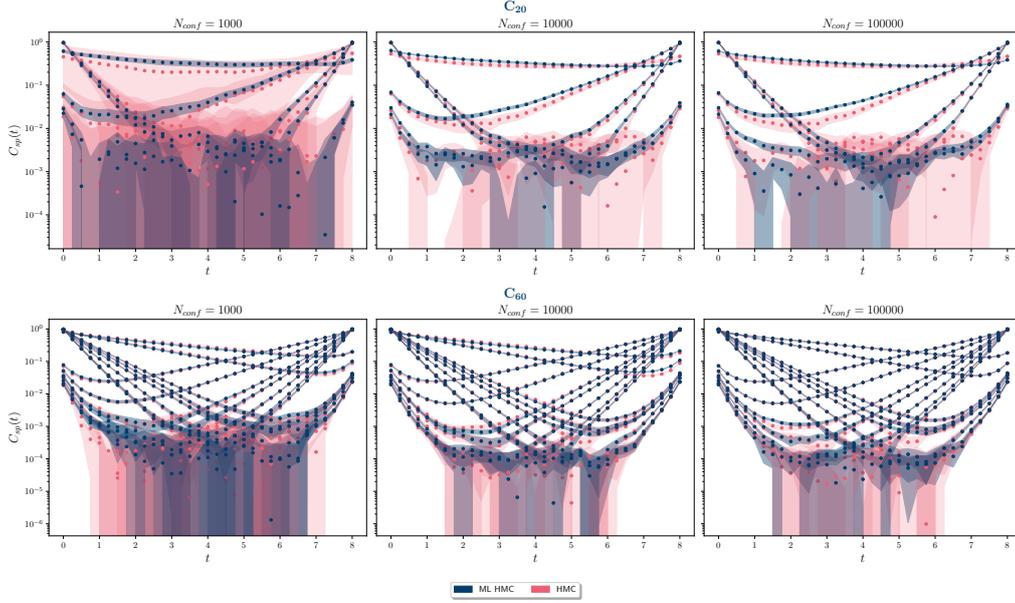
Figure A.5: The correlators of Fullerene C_{20} (upper row) and C_{60} (lower row) are shown. A real plane (red) and a tangent plane (blue) HMC at inverse temperature $\beta = 8$ and $N_t = 32$ time slices are compared. We consider an on-site interaction $U = 3$ and zero chemical potential. The, here negligible, sign problem solely stems from the non-bipartiteness of the system due to the pentagonal rings. Already, at small number of configurations $N_{cfg} \leq 10\,000$ the signals are very good.

However, in the intermediate range of $\mu \in [2, 3]$ the tangent plane is not sufficient for a reasonable estimate, where we apply ML HMC with the SHIFT network. The point at $\mu = 2.5$ requires more attention and we plan to address it with the coupling network in the future expecting much better results. The expected behaviour of the charge density, $\rho(\mu = 0) = 0$ and $\rho(\mu \to \infty) \to 1$, is found numerically. The dashed line exemplarily indicates an expected plateau at $\langle \rho(\mu) \rangle = 4/18$. As it can be seen, this plateau is not fully deducible but may be surmised around $\mu \approx 1$. Studies of smaller systems, see [176], indicate increasing β makes these plateaus more pronounced.

To probe our method in physically more relevant systems than the 18 Site Graphene sheet, we started an investigation of Fullerene C_{20} and C_{60}. The correlators at $N_t = 32$, $\beta = 8$, $U = 3$ and zero chemical potential are displayed in figure A.5. The mild sign problem is solely due to the non-bipartiteness of the lattice structure. We compare standard HMC (red) with tangent plane HMC (blue) to show that the tangent plane obtains a good signal already at small number of configurations $N_{cfg} = 1000$ in both systems. For C_{60} the sign problem is negligible and the real plane HMC gives a good signal too. For finite chemical potential the sign problem aggravates as it imposes a second source. We are currently working on this particular lattice geometry.

A.4   Conclusions

Simulating systems with strongly correlated electrons is a rather challenging task due to the innate sign problem for doped systems. Current methods, like deformation onto Lefschetz thimbles, suffer from a very difficult scaling in computational cost. We overcome this issue by identifying efficient Neural Network architectures and incorporating them in a HMC algorithm. We present first studies of doped Graphene sheets using this enhanced HMC and demonstrate a substantial improvement of the signal, effectively mitigating the sign problem. Considering systems with increasing volume illustrates some stability of this method for larger volumes. Further, preliminary simulation of Fullerene C_{20} and C_{60} at vanishing chemical potential are shown. In the near future we will apply the neural network enhanced HMC also at finite chemical potential.

Acknowledgements

Appendix B

Moments of Parton Distributions Functions from Lattice QCD at the Physical Point

This appendix chapter is based on (not peer-reviewed) [6]:

Marcel Rodekamp, Michael Engelhardt, Jeremy R. Green, Stefan Krieg, Stefan Meinel, John W. Negele, Andrew Pochinsky, and Sergey" Syritsyn. "Moments of Parton Distributions Functions from Lattice QCD at the Physical Point." In: *30th International Workshop on Deep-Inelastic Scattering and Related Subjects*. June 2023. arXiv: 2306.14431

These proceedings are not peer-reviewed, but show preliminary results of the publication discussed in chapter 5. A Lattice QCD calculation of the second Mellin moments of the nucleon unpolarized, polarized, and transversity parton distribution functions. As before, the calculation is performed at the physical pion mass with two different lattice spacings, and includes both zero and non-zero nucleon momenta. A stronger emphasis, compared to chapter 5, lies in the identification of operators that greatly reduce excited state contamination. Therefore, the set of operators does not completely match the set of independent operators shown in chapter 5.

B.1 Introduction

Parton distribution functions (PDFs) have proved to be a valuable tool in describing the structure of hadrons and making predictions for high-energy processes at hadron colliders. First-principles calculations of PDFs are very difficult due to their non-perturbative nature. Lattice QCD provides a way of calculating (non-perturbative) observables by introducing a four-dimensional Euclidean hypercubic lattice to discretise the space-time, serving as a regulator. The path integral is then calculated with a Monte Carlo algorithm.

In the past years the Lattice QCD community has made tremendous progress in calculating PDFs by directly assessing their Bjorken-x dependence from the leading-twist contribution to bilocal matrix elements at high momentum. In this work, we concentrate on the second Mellin moment $\langle x \rangle$ [36, 37, 43] via matrix elements of local twist-two operators, which does not require large momenta to suppress higher-twist contributions and thus simplifies the numerical estimation. We aim to understand the excited-state contamination and identify a set of matrix elements that have particularly low contributions from exited states. This requires the study of matrix elements at finite but small momenta as some have contributions only at non-zero momentum. The study of forward matrix elements of local operators at non-zero momentum is uncommon but has been done in references [139, 140, 141].

This contribution is organized as follows. In section B.2 we explain our analysis chain and discuss in detail which operators are considered. In section B.3 we show our preliminary results of the different steps of the analysis and discuss their significance in terms of excited-state contamination. Last, in B.4 we summarize our findings.

B.2 Method

Moments of PDFs can be obtained by calculating forward matrix elements of local leading twist operators [142, 143, 42, 45]

$$O^X \equiv O^X_{\{\alpha,\mu\}} = \overline{q} \Gamma^X_{\{\alpha} \overleftrightarrow{D}_{\mu\}} q, \tag{B.1}$$

where $X = V, A, T$ indicates the vector, axial or tensor channel leading to unpolarized, polarized or transversity PDFs respectively. We symmetrize the indices and take the traceless part, denoted by $\{\alpha, \mu\}$, and restrict ourselves to the isovector channel, $O^X(q = u) - O^X(q = d)$, to avoid calculating disconnected diagrams. The left-right acting covariant derivative $\overleftrightarrow{D}$ is constructed on the Euclidean lattice by finite differences of neighbouring points connected by an appropriate gauge link $U_\mu(\mathbf{x})$. One can show that the forward matrix element is proportional to the desired moment $\langle x \rangle$ [44, 36, 45]

$$\mathcal{M} \equiv \langle N(p)|O^X_{\{\alpha,\mu\}}|N(p)\rangle = \langle x \rangle \overline{u}_{N(p)} \Gamma^X_{\{\alpha} \mathrm{i} \, p_{\mu\}} u_{N(p)}, \tag{B.2}$$

where p is the nucleon's 4-momentum.

In the continuum, the operators (B.1) are classified according to irreducible representations of the Lorentz group, which in Euclidean space is replaced by the orthogonal group [42]. On the lattice, the orthogonal group is further reduced to the hypercubic group $H(4)$. This explicit breaking causes some operators to mix with lower-dimensional ones; however, for a one-derivative
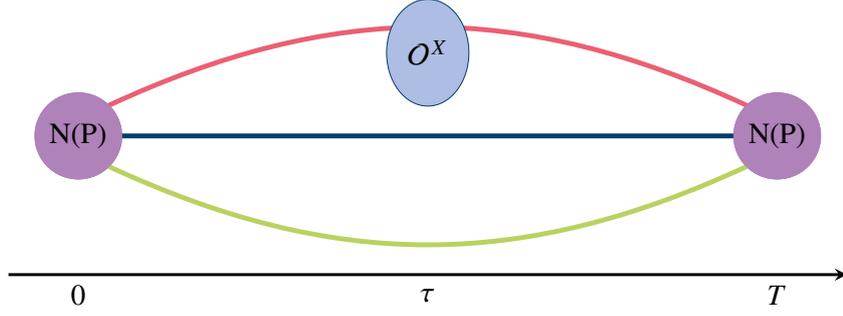
Figure B.1: Graphical representation of $C_{3pt}^{O^X}(T, \tau)$, a source nucleon inserted at time $t = 0$ and a sink nucleon removed at time $t = T$. A local leading twist operator (B.1) is inserted on a given time slice τ. The nucleons on the lattice are represented by interpolating operators while O^X is determined by finite differences connected with gauge links.

operator as used here this does not happen. Still, the Euclidean irreducible representations to which our operators belong split into multiple hypercubic irreps; we use the typical notation where $\tau_a^{(b)}$ denotes the ath b-dimensional irrep. Each of the latter has a different renormalization factor, so we construct operators with definite hypercubic irreducible representation to keep the renormalization diagonal [42]. In practice this means for each $\tau_a^{(b)}$ we have to calculate the renomalization factor $Z_{\tau_a^{(b)}}$ to multiply matrix elements of an operator that transforms irreducibly under it, consequently we denote $Z_{O^X} \equiv Z_{\tau_a^{(b)}}$.

The matrix element of (B.2) can be obtained from the lattice by considering ratios of three-point and two-point correlation functions [44, 45]. The two-point correlation function $C_{2pt}(\tau)$ measures the correlation of a nucleon source and a nucleon sink separated by a time τ, while the three-point correlation function $C_{3pt}^{O^X}(T, \tau)$ separates the source and sink nucleons by a time T and inserts an operator of interest, here O^X, at time τ. For a graphical representation consider figure B.1. The matrix element is then obtained in the limit

$$\mathcal{M} = \lim_{T-\tau, \tau \to \infty} R(T, \tau) \equiv \lim_{T-\tau, \tau \to \infty} \frac{C_{3pt}^{O^X}(T, \tau)}{C_{2pt}(T)}. \tag{B.3}$$

Doing a spectral analysis of this ratio reveals the matrix element of the ground state

$$R(T, \tau) = \mathcal{M} + \text{Excited States}. \tag{B.4}$$

Expanding further, including the first excited state, shows the dominant excited-state contamination

$$\mathcal{M} \frac{1 + c_1 e^{-\frac{T}{2}\Delta E} \cosh\left[(T/2 - \tau)\Delta E\right] + c_2 e^{-T\Delta E}}{1 + c_3 e^{-T\Delta E}}, \tag{B.5}$$

where we use $\Delta E = E_1 - E_0$. Naturally, one would consider large T, τ approaching the limit of (B.3). The statistical noise increases with T, implying increased numerical costs for this approach. The constants c_i depend on the operator O^X; thus, if they appear to be small or obey some symmetry, the excited-state contamination of the matrix element is further reduced.

Considering the sum of ratios

$$S(T, \tau_{\text{skip}}) = a \sum_{\tau=\tau_{\text{skip}}}^{T-\tau_{\text{skip}}} R(T, \tau) = \mathcal{M}(T - \tau_{\text{skip}}) + c + \text{Excited States}, \tag{B.6}$$

excited-state contamination is exponentially suppressed with T compared to $T/2$ for the ratios themselves [145, 146]. Increasing τ_{skip} reduces excited-state contamination, here typically a value

of $\tau_{\text{skip}}/a = 1, 2, 3$ is enough. In order to extract the matrix element from the ratio sums, up to excited states, we can either fit the slope according to (B.6) or use finite differences

$$\mathcal{M} = \frac{S(T + \delta, \tau_{\text{skip}}) - S(T, \tau_{\text{skip}})}{\delta}. \tag{B.7}$$

Due to the available data we use a combination of $\delta/a \in \{1, 2, 3\}$ depending on whether a neighbour $T + \delta$ is available.

Having the basic quantities of interest, we can summarize the analysis as follows. First estimate the ratios $R(T, \tau)$ and ratio sums $S(T, \tau_{\text{skip}})$. Currently, we extract matrix elements $\mathcal{M}$ in two ways. First, fitting the slope of $S(T, \tau_{\text{skip}})$ at fixed τ_{skip} limiting $T \geq T'$ for various minimal source-sink separations T'. Second, extracting the slope via finite differences at a source-sink separation $T = T'$. A matrix element extracted with either those is denoted by $\mathcal{M}|_{T',\mathfrak{m}}$ where $\mathfrak{m}$ denotes one of the two above methods. For both methods, as we increase T' excited states are expected to decay. Dividing the kinematic factor results in a T'-dependent moment for a given operator O^X and momentum p using the matrix element extraction method $\mathfrak{m}$

$$\mathfrak{X}_{O^X, p, \mathfrak{m}}(T') = \frac{\mathcal{M}|_{T',\mathfrak{m}}}{\overline{u}_{N(p)} \Gamma^X_{\{\alpha} \, \mathrm{i} \, p_{\mu\}} u_{N(p)}}. \tag{B.8}$$

To simplify the following equations, we define a compound index $j = (O^X, p, \mathfrak{m})$ that runs over all operators and momenta with nonzero kinematic factors as well as the different methods to obtain the matrix element. Determining the renormalization factors in RI-(S)MOM and matching it to $\overline{\text{MS}}$(2 GeV) allows us to express the renormalized moment $\mathfrak{X}^{\text{ren}}_j(T') = Z_{O^X} \cdot \mathfrak{X}_j(T')$. With these we define the central value as weighted average of the different results

$$\langle x \rangle^{\text{ren}} = \sum_{j, T' \geq T^j_{\text{plat}}} \mathfrak{W}_j(T') \mathfrak{X}^{\text{ren}}_j(T'). \tag{B.9}$$

Here T^j_{plat} denotes the smallest source-sink separation such that $\mathfrak{X}_j(T')$ agree for all $T' \geq T^j_{\text{plat}}$. The weights $\mathfrak{W}_j(T') \propto 1/\sigma^2_j(T')$ are normalised such that they sum to 1. The used variances are estimated via bootstrap over $\mathfrak{X}_j(T')$ and the errors of the renormalization constants are propagated. Last we estimate a systematic error by taking the weighted standard deviation over the different results

$$\sigma^2_{syst} = \sum_{j, T' \geq T^j_{\text{plat}}} \mathfrak{W}_j(T') \left[\mathfrak{X}^{\text{ren}}_j(T') - \langle x \rangle^{\text{ren}} \right]^2. \tag{B.10}$$

B.3 Results

We use a tree-level Symanzik-improved gauge action with 2+1 flavour tree-level improved Wilson Clover fermions coupling via 2-level HEX-smearing. The details can be found in [133, 134, 50] and relevant parameters are summarized in table B.1. Two ensembles, coarse and fine, have been generated at the physical pion mass corresponding to lattice spacings of 0.1163(4) fm and 0.0926(6) fm respectively. On each ensemble two-point and three-point correlation functions are calculated with source-sink separations ranging from ≈ 0.3 fm to 1.4 fm and ≈ 0.9 fm to ≈ 1.5 fm. For each ensemble two momenta are chosen; $\vec{p} = (p_x, 0, 0)$ with $p_x = 0, -2[2\pi/L]$ and $p_x = 0, -1[2\pi/L]$ respectively.

Figures B.2 to B.4 show the different steps of the analysis. For a given channel X, the figures B.2a and B.2b show results using one possible operator $O^X_{\{\alpha,\mu\}}$. Here we multiply with the kinematic factor $\overline{R}(T, \tau) = 1/\overline{u}_{N(p)} \Gamma^X_{\{\alpha} \, \mathrm{i} \, p_{\mu\}} u_{N(p)} \cdot R(T, \tau)$ such that a plateau corresponds to the bare moment. These plots omit the largest source sink separation due to its enormous statistical uncertainty. Two different operators $O^X_{\{\alpha,\mu\}}$ are used for each channel X going from

Ensemble	Size	β	$a\,[\text{fm}]$	$m_\pi\,[\text{MeV}]$	$m_\pi L$	T/a	$p_x\,[2\pi/L]$	N_{cfg}
Coarse	48^4	3.31	0.1163(4)	136(2)	3.9	3, 4, 5, 6, 7, 8, 10, 12	0, −2	212
Fine	64^4	3.5	0.0926(6)	133(1)	4.0	10, 13, 16	0, −1	427

Table B.1: Details of the used ensembles. The ensembles are at the physical pion mass, $m_\pi \approx m_\pi^{phys}$. A larger and a smaller lattice spacing, labelled as "Coarse" and "Fine" respectively, are available. The ensembles are generated with a tree-level Symanzik-improved gauge action with 2+1 flavour tree-level improved Wilson Clover fermions coupled via 2-level HEX-smearing [133, 134, 50]. Furthermore, the available source-sink separations (T) and momenta (p_x) which are used in the calculation of the ratios, equation (B.3), are displayed.

figure B.2a to B.2b. For the axial channel ($X = A$) both operators transform under $\tau_4^{(6)}$. The lower excited-state contamination for some operators can be deduced directly from these figures as B.2a obey the cosh behaviour of (B.5) while B.2b are perfectly flat within uncertainty. Notably, those operators have a contribution only at finite momentum ($p_x \neq 0$ here) which increases the statistical noise.

A similar rescaling has been done for the ratio sums $\overline{S}(T, \tau_{\text{skip}}) = 1/\overline{u}_{N(p)} \Gamma_{\{\alpha}^X i\,p_{\mu\}} u_{N(p)} \cdot S(T, \tau_{\text{skip}})$ in figure B.3. The exited state contamination is indicated by the slight curvature though much more obscured compared to the ratios. The slopes of these lines are used in the current analysis shown in figure B.4. In future work we want to include a 2-state analysis as in (B.5), improving on the central value (B.9) as well as the systematic error estimation (B.10).

In Figure B.4 the gray points correspond to the different renormalized moments $\mathfrak{X}_j^{\text{ren}}(T')$ from finite differences plotted against T' but slightly displaced to increase readability. The blue points represent the preliminary result, computed using (B.9). The inner errorbars represent the statistical – bootstrap – uncertainty while the outer ones add the estimate of systematic errors, $\sqrt{\sigma_{stat}^2 + \sigma_{syst}^2}$. The upper and lower row collect results from the coarse and fine ensemble respectively. Encouragingly, the central values agree within the uncertainties.

B.4   Summary

We calculate the second Mellin moment $\langle x \rangle$ of axial, vector and tensor PDFs from lattice QCD with two lattice spacings at the physical pion mass. The study includes nucleon matrix elements at zero and finite momentum, boosted in the x-direction. We identified a set operators that contribute only at finite momentum and have particularly low excited-state contamination. For the future, we are working on a direct 2-state analysis of the ratios to improve the quantitative analysis of the excited-state contamination.

Acknowledgments

(a)

(b)

Figure B.2: Ratios, cf. eq. (B.3), for the coarse ensemble. Different source-sink separations T are shown in different colours and the two momenta, in B.2a, are distinguished by hollow and filled markers. Different sets of operators were chosen for the two subfigures.
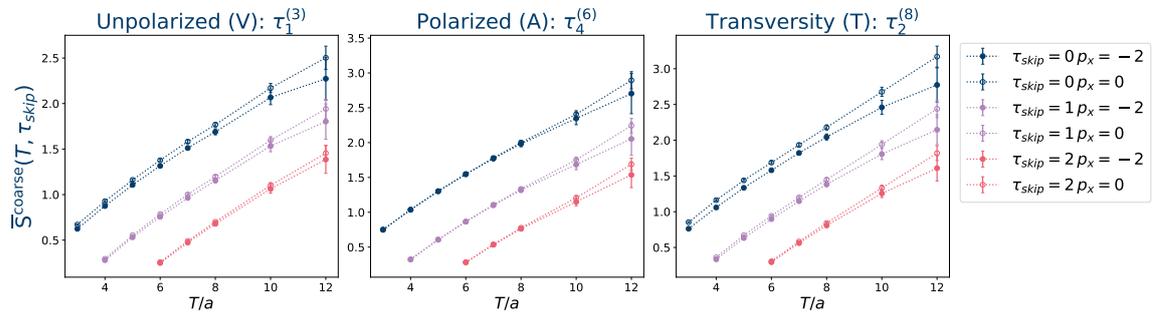
Figure B.3: ratio sums $\overline{S}(T, \tau_{\mathrm{skip}})$ on the coarse lattice. The chosen operators O^X are the same as in B.2a. Each $\overline{S}(T, \tau_{\mathrm{skip}})$ is plotted at fixed τ_{skip}, indicated by colour, over different source-sink separations. As in B.2a different momenta are displayed with hollow and filled markers.
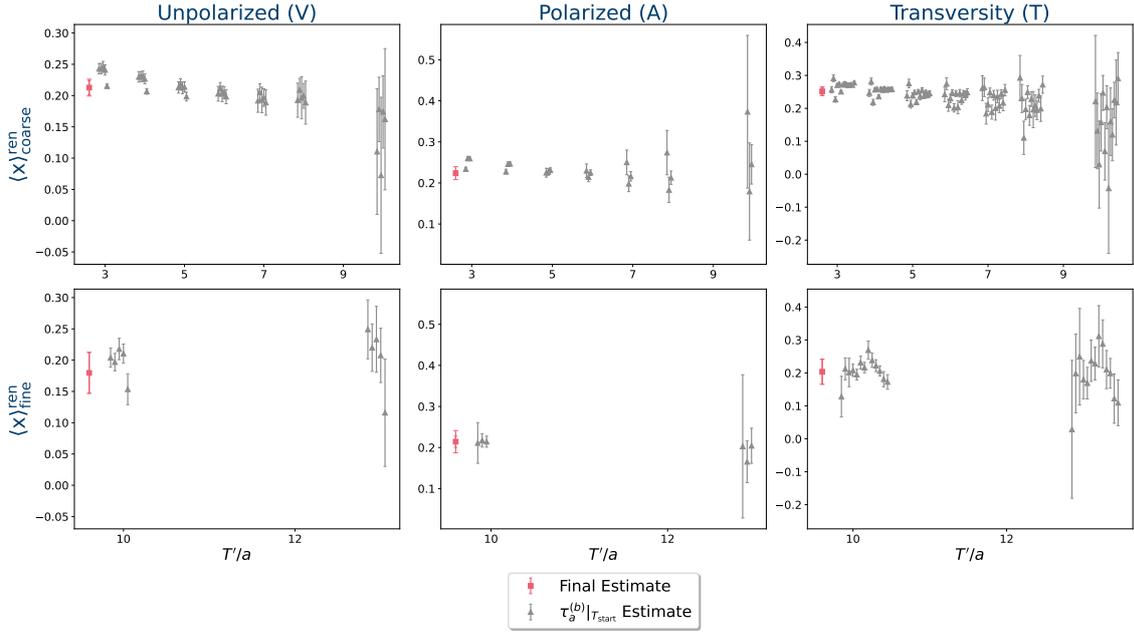
Figure B.4: Results for the renormalized moments computed from the ratio sums (B.6). Moments are computed from finite differences at fixed T' and $\tau_{\text{skip}} = 1$. The grey points represent $\mathfrak{X}_j^{\text{ren}}(T')$. The moments are plotted against T' and slightly displaced for clarity. The red points represent the preliminary result obtained using (B.9). The inner and outer errorbars indicate statistical and total uncertainty respectively.

Bibliography

[1] Marcel Rodekamp et al. "Mitigating the Hubbard sign problem with complex-valued neural networks." In: *Phys. Rev. B* 106 (12 Sept. 2022), p. 125139. DOI: 10.1103/PhysRevB.106.125139. arXiv: 2203.00390.

[2] Marcel Rodekamp et al. "From Theory to Practice: Applying Neural Networks to Simulate Real Systems with Sign Problems." In: vol. LATTICE2023. 2023. DOI: 10.22323/1.453.0031. arXiv: 2311.18312.

[3] Marcel Rodekamp et al. *Single Particle Spectrum of Doped* $C_{20}H_{12}$-*Perylene*. June 2024. arXiv: 2406.06711.

[4] Marcel Rodekamp et al. "Moments of nucleon unpolarized, polarized, and transversity parton distribution functions from lattice QCD at the physical point." In: *Physical Review D* 109.7 (Apr. 2024), p. 074508. ISSN: 2470-0010, 2470-0029. DOI: 10.1103/PhysRevD.109.074508. arXiv: 2306.14431v2.

[5] Marcel Rodekamp and Christoph Gäntgen. "Mitigating the Hubbard Sign Problem. A Novel Application of Machine Learning." In: vol. LATTICE2022. 2023. DOI: 10.22323/1.430.0032. arXiv: 2211.09584.

[6] Marcel Rodekamp et al. "Moments of Parton Distributions Functions from Lattice QCD at the Physical Point." In: *30th International Workshop on Deep-Inelastic Scattering and Related Subjects*. June 2023. arXiv: 2306.14431.

[7] Christoph Gäntgen et al. "Fermionic sign problem minimization by constant path integral contour shifts." In: *Physical Review B* 109.19 (May 2024), p. 195158. ISSN: 2469-9950, 2469-9969. DOI: 10.1103/PhysRevB.109.195158. arXiv: 2307.06785v1.

[8] Christoph Gäntgen et al. "Reducing the Sign Problem with simple Contour Deformation." In: 2024. URL: https://pos.sissa.it/453/007/.

[9] Michael Engelhardt et al. "Quark orbital angular momentum in the proton from a twist-3 generalized parton distribution." In: vol. LATTICE2023. 2023. DOI: https://doi.org/10.22323/1.453.0316. arXiv: 2406.00845.

[10] Marcel Rodekamp et al. *Nanosystem Simulation Library (NSL)*. URL: https://github.com/Marcel-Rodekamp/NSL.

[11] Marcel Rodekamp and Giovanni Pederiva. *Correlator Analyser*. URL: https://github.com/Marcel-Rodekamp/CorrelatorAnalyser.

[12] Istvan Montvay and Gernot Münster. *Quantum Fields on a Lattice*. 1st ed. Cambridge University Press, Mar. 1994. ISBN: 978-0-521-40432-7. DOI: 10.1017/CBO9780511470783.

[13] Thomas DeGrand and Carleton DeTar. *Lattice Methods for Quantum Chromodynamics*. WORLD SCIENTIFIC, Sept. 2006. ISBN: 978-981-256-727-7. DOI: 10.1142/6065.

[14] Christof Gattringer and Christian B. Lang. *Quantum Chromodynamics on the Lattice: An Introductory Presentation*. Vol. 788. Lecture Notes in Physics. Berlin, Heidelberg: Springer Berlin Heidelberg, 2010. ISBN: 978-3-642-01849-7. DOI: 10.1007/978-3-642-01850-3.

[15] Francesco Knechtli, Michael Günther, and Michael Peardon. *Lattice Quantum Chromodynamics*. SpringerBriefs in Physics. Dordrecht: Springer Netherlands, 2017. ISBN: 978-94-024-0997-0. DOI: 10.1007/978-94-024-0999-4.

[16] Matthias Troyer and Uwe-Jens Wiese. "Computational complexity and fundamental limitations to fermionic quantum Monte Carlo simulations." In: *Physical Review Letters* 94.17 (May 2005), p. 170201. ISSN: 0031-9007, 1079-7114. DOI: 10.1103/PhysRevLett.94.170201. arXiv: cond-mat/0408370.

[17] P. T. Matthews and A. Salam. "The Green's functions of quantised fields." en. In: *Il Nuovo Cimento* 12.4 (Oct. 1954), pp. 563–565. ISSN: 0029-6341, 1827-6121. DOI: 10.1007/BF02781302. (Visited on 06/15/2024).

[18] P. T. Matthews and A. Salam. "Propagators of Quantized Field." In: *Il Nuovo Cimento* 2.1 (July 1955), pp. 120–134. ISSN: 0029-6341, 1827-6121. DOI: 10.1007/BF02856011.

[19] Jan-Lukas Wynen et al. "Avoiding Ergodicity Problems in Lattice Discretizations of the Hubbard Model." In: *Physical Review B* 100.7 (Aug. 2019), p. 075141. ISSN: 2469-9950, 2469-9969. DOI: 10.1103/PhysRevB.100.075141. arXiv: 1812.09268.

[20] Casey E. Berger et al. "Complex Langevin and Other Approaches to the Sign Problem in Quantum Many-Body Physics." In: *Physics Reports* 892 (Jan. 2021), pp. 1–54. ISSN: 03701573. DOI: 10.1016/j.physrep.2020.09.002. arXiv: 1907.10183.

[21] Solomon Lefschetz. "On certain numerical invariants of algebraic varieties with application to abelian varieties." In: *Transactions of the American Mathematical Society* 22.3 (1921), pp. 327–406. ISSN: 0002-9947, 1088-6850. DOI: 10.1090/S0002-9947-1921-1501178-3.

[22] Gert Aarts. "Can Stochastic Quantization Evade the Sign Problem? – The Relativistic Bose Gas at Finite Chemical Potential." In: *Physical Review Letters* 102.13 (Apr. 2009), p. 131601. ISSN: 0031-9007, 1079-7114. DOI: 10.1103/PhysRevLett.102.131601. arXiv: 0810.2089.

[23] Andrei Alexandru et al. "Complex paths around the sign problem." In: *Reviews of Modern Physics* 94.1 (Mar. 2022), p. 015006. ISSN: 0034-6861, 1539-0756. DOI: 10.1103/RevModPhys.94.015006. arXiv: 2007.05436v1.

[24] Richard C. Brower, David Schaich, and Claudio Rebbi. "Hybrid Monte Carlo simulation on the graphene hexagonal lattice." In: *Proceedings of XXIX International Symposium on Lattice Field Theory — PoS(Lattice 2011)*. Squaw Valley, Lake Tahoe, California: Sissa Medialab, July 2012, p. 056. DOI: 10.22323/1.139.0056. arXiv: 1204.5424.

[25] Johann Ostmeyer et al. "Semimetal–Mott insulator quantum phase transition of the Hubbard model on the honeycomb lattice." In: *Physical Review B* 102.24 (Dec. 2020), p. 245105. ISSN: 2469-9950, 2469-9969. DOI: 10.1103/PhysRevB.102.245105. arXiv: 2005.11112v3.

[26] Ananth Dodabalapur, Howard E. Katz, and Luisa Torsi. "Molecular orbital energy level engineering in organic transistors." In: *Advanced Materials* 8.10 (Oct. 1996), pp. 853–855. ISSN: 0935-9648, 1521-4095. DOI: 10.1002/adma.19960081021.

[27] Rudolph Pariser and Robert G. Parr. "A Semi-Empirical Theory of the Electronic Spectra and Electronic Structure of Complex Unsaturated Molecules. I." In: *The Journal of Chemical Physics* 21.3 (Mar. 1953), pp. 466–471. ISSN: 0021-9606, 1089-7690. DOI: 10.1063/1.1698929.

[28] Rudolph Pariser and Robert G. Parr. "A Semi-Empirical Theory of the Electronic Spectra and Electronic Structure of Complex Unsaturated Molecules. II." In: *The Journal of Chemical Physics* 21.5 (Mar. 1953), pp. 767–776. ISSN: 0021-9606, 1089-7690. DOI: 10.1063/1.1699030.

[29] John Hubbard. "Calculation of Partition Functions." In: *Physical Review Letters* 3.2 (July 1959), pp. 77–78. ISSN: 0031-9007. DOI: 10.1103/PhysRevLett.3.77.

[30] John Hubbard. "Electron correlations in narrow energy bands. II. The degenerate band case." In: *Proceedings of the Royal Society of London. Series A. Mathematical and Physical Sciences* 277.1369 (Jan. 1964), pp. 237–259. ISSN: 0080-4630, 2053-9169. DOI: 10.1098/rspa.1964.0019.

[31] John Hubbard. "Electron correlations in narrow energy bands III. An improved solution." In: *Proceedings of the Royal Society of London. Series A. Mathematical and Physical Sciences* 281.1386 (Sept. 1964), pp. 401–419. ISSN: 0080-4630, 2053-9169. DOI: 10.1098/rspa.1964.0190.

[32] Thomas Luu and Timo A. Lähde. "Quantum Monte Carlo Calculations for Carbon Nanotubes." In: *Physical Review B* 93.15 (Apr. 2016), p. 155106. ISSN: 2469-9950, 2469-9969. DOI: 10.1103/PhysRevB.93.155106. arXiv: 1511.04918.

[33] M. Lüscher. "Construction of a Selfadjoint, Strictly Positive Transfer Matrix for Euclidean Lattice Gauge Theories." In: *Communications in Mathematical Physics* 54.3 (Oct. 1977), pp. 283–292. ISSN: 0010-3616, 1432-0916. DOI: 10.1007/BF01614090.

[34] R. L. Workman et al. "Review of Particle Physics." In: *PTEP* 2022 (2022), p. 083C01. DOI: 10.1093/ptep/ptac097.

[35] Constantia Alexandrou et al. "The Deconfinement Phase Transition in One-Flavour QCD." In: *Physical Review D* 60.3 (June 1999), p. 034504. ISSN: 0556-2821, 1089-4918. DOI: 10.1103/PhysRevD.60.034504. arXiv: hep-lat/9811028.

[36] Philipp Hägler. "Hadron structure from lattice quantum chromodynamics." In: *Physics Reports* 490.3-5 (May 2010), pp. 49–175. ISSN: 03701573. DOI: 10.1016/j.physrep.2009.12.008. arXiv: 0912.5483v1.

[37] Huey-Wen Lin et al. "Parton distributions and lattice QCD calculations: A community white paper." In: *Progress in Particle and Nuclear Physics* 100 (May 2018), pp. 107–160. ISSN: 01466410. DOI: 10.1016/j.ppnp.2018.01.007. arXiv: 1711.07916v3.

[38] Xiangdong Ji. "GENERALIZED PARTON DISTRIBUTIONS." In: *Annual Review of Nuclear and Particle Science* 54.1 (Dec. 2004), pp. 413–450. ISSN: 0163-8998, 1545-4134. DOI: 10.1146/annurev.nucl.54.070103.181302.

[39] E. Wigner. "On the Quantum Correction For Thermodynamic Equilibrium." In: *Physical Review* 40.5 (June 1932), pp. 749–759. ISSN: 0031-899X. DOI: 10.1103/PhysRev.40.749.

[40] Xiangdong Ji. "Parton Physics on a Euclidean Lattice." In: *Physical Review Letters* 110.26 (June 2013), p. 262002. ISSN: 0031-9007, 1079-7114. DOI: 10.1103/PhysRevLett.110.262002. arXiv: 1305.1539v1.

[41] Krzysztof Cichy and Martha Constantinou. "A Guide to Light-Cone PDFs from Lattice QCD: An Overview of Approaches, Techniques, and Results." In: *Advances in High Energy Physics* 2019 (June 2019), pp. 1–68. ISSN: 1687-7357, 1687-7365. DOI: 10.1155/2019/3036904. arXiv: 1811.07248v2.

[42] Meinulf Göckeler et al. "Lattice operators for moments of the structure functions and their transformation under the hypercubic group." In: *Physical Review D* 54.9 (Nov. 1996), pp. 5705–5714. ISSN: 0556-2821, 1089-4918. DOI: 10.1103/PhysRevD.54.5705. arXiv: hep-lat/9602029v1.

[43] Martha Constantinou. "The x-dependence of hadronic parton distributions: A review on the progress of lattice QCD." In: *The European Physical Journal A* 57.2 (Feb. 2021), p. 77. ISSN: 1434-6001, 1434-601X. DOI: 10.1140/epja/s10050-021-00353-7. arXiv: 2010.02445.

[44] Meinulf Göckeler et al. "Quark helicity flip generalized parton distributions from two-flavor lattice QCD." In: *Physics Letters B* 627.1-4 (Oct. 2005), pp. 113–123. ISSN: 03702693. DOI: 10.1016/j.physletb.2005.09.002. arXiv: hep-lat/0507001.

[45] Constantia Alexandrou. "Parton Distribution Functions from Lattice QCD." In: *Few-Body Systems* 57.8 (Aug. 2016), pp. 621–626. ISSN: 0177-7963, 1432-5411. DOI: 10.1007/s00601-016-1073-5. arXiv: 1602.08726v1.

[46] Tim Harris et al. "Nucleon isovector charges and twist-2 matrix elements with N f = 2 + 1 dynamical Wilson quarks." In: *Physical Review D* 100.3 (Aug. 2019), p. 034513. ISSN: 2470-0010, 2470-0029. DOI: 10.1103/PhysRevD.100.034513. arXiv: 1905.01291v2.

[47] Gunnar Bali et al. "Nucleon generalized form factors from two-flavor lattice QCD." In: *Physical Review D* 100.1 (July 2019), p. 014507. ISSN: 2470-0010, 2470-0029. DOI: 10.1103/PhysRevD.100.014507. arXiv: 1812.08256v2.

[48] Constantia Alexandrou et al. "Moments of nucleon generalized parton distributions from lattice QCD simulations at physical pion mass." In: *Physical Review D* 101.3 (Feb. 2020), p. 034519. ISSN: 2470-0010, 2470-0029. DOI: 10.1103/PhysRevD.101.034519. arXiv: 1908.10706.

[49] Santanu Mondal et al. "Moments of nucleon isovector structure functions in 2 + 1 + 1 -flavor QCD." In: *Physical Review D* 102.5 (Sept. 2020), p. 054512. ISSN: 2470-0010, 2470-0029. DOI: 10.1103/PhysRevD.102.054512. arXiv: 2005.13779v2.

[50] Nesreen Hasan et al. "Nucleon axial, scalar, and tensor charges using lattice QCD at the physical pion mass." In: *Physical Review D* 99.11 (June 2019), p. 114505. ISSN: 2470-0010, 2470-0029. DOI: 10.1103/PhysRevD.99.114505. arXiv: 1903.06487v3.

[51] Michel Garcon. "An Introduction to the Generalized Parton Distributions." In: *The European Physical Journal A* 18.2-3 (Nov. 2003), pp. 389–394. ISSN: 1434-6001, 1434-601X. DOI: 10.1140/epja/i2002-10242-2. eprint: hep-ph/0210068.

[52] Robert L. Jaffe. *Spin, Twist and Hadron Structure in Deep Inelastic Processes*. Feb. 1996. arXiv: hep-ph/9602236.

[53] M. Diehl. "Generalized Parton Distributions." In: *Physics Reports* 388.2-4 (Dec. 2003), pp. 41–277. ISSN: 03701573. DOI: 10.1016/j.physrep.2003.08.002. arXiv: hep-ph/0307382.

[54] Andrei Alexandru, Gokce Basar, and Paulo Bedaque. "A Monte Carlo algorithm for simulating fermions on Lefschetz thimbles." In: *Physical Review D* 93.1 (Jan. 2016), p. 014504. ISSN: 2470-0010, 2470-0029. DOI: 10.1103/PhysRevD.93.014504. arXiv: 1510.03258.

[55] Marco Cristoforetti et al. "An efficient method to compute the residual phase on a Lefschetz thimble." In: *Physical Review D* 89.11 (June 2014), p. 114505. ISSN: 1550-7998, 1550-2368. DOI: 10.1103/PhysRevD.89.114505. arXiv: 1403.5637.

[56] Marco Cristoforetti et al. "Monte Carlo simulations on the Lefschetz thimble: taming the sign problem." In: *Physical Review D* 88.5 (Sept. 2013), p. 051501. ISSN: 1550-7998, 1550-2368. DOI: 10.1103/PhysRevD.88.051501. arXiv: 1303.7204.

[57] Abhishek Mukherjee, Marco Cristoforetti, and Luigi Scorzato. "Metropolis Monte Carlo on the Lefschetz thimble: application to a one-plaquette model." In: *Physical Review D* 88.5 (Sept. 2013), p. 051502. ISSN: 1550-7998, 1550-2368. DOI: 10.1103/PhysRevD.88.051502. arXiv: 1308.0233.

[58] Takuya Kanazawa and Yuya Tanizaki. "Structure of Lefschetz thimbles in simple fermionic systems." In: *Journal of High Energy Physics* 2015.3 (Mar. 2015), p. 44. ISSN: 1029-8479. DOI: 10.1007/JHEP03(2015)044. arXiv: 1412.2802.

[59] Yuya Tanizaki, Yoshimasa Hidaka, and Tomoya Hayata. "Lefschetz-thimble analysis of the sign problem in one-site fermion model." In: *New Journal of Physics* 18.3 (Mar. 2016), p. 033002. ISSN: 1367-2630. DOI: 10.1088/1367-2630/18/3/033002. arXiv: 1509.07146.

[60] Andrei Alexandru et al. "Deep learning beyond Lefschetz thimbles." In: *Physical Review D* 96.9 (Nov. 2017), p. 094505. ISSN: 2470-0010, 2470-0029. DOI: 10.1103/PhysRevD.96.094505. arXiv: 1709.01971v1.

[61] Jan-Lukas Wynen et al. "Leveraging Machine Learning to Alleviate Hubbard Model Sign Problems." In: *Physical Review B* 103.12 (Mar. 2021), p. 125153. ISSN: 2469-9950, 2469-9969. DOI: 10.1103/PhysRevB.103.125153. arXiv: 2006.11221.

[62] Kouji Kashiwa, Yuto Mori, and Akira Ohnishi. "Control the model sign problem via path optimization method: Monte-Carlo approach to QCD effective model with Polyakov loop." In: *Physical Review D* 99.1 (Jan. 2019), p. 014033. ISSN: 2470-0010, 2470-0029. DOI: 10.1103/PhysRevD.99.014033. arXiv: 1805.08940.

[63] William Detmold et al. "Path integral contour deformations for noisy observables." In: *Physical Review D* 102.1 (July 2020), p. 014514. ISSN: 2470-0010, 2470-0029. DOI: 10.1103/PhysRevD.102.014514. arXiv: 2003.05914.

[64] William Detmold et al. "Path integral contour deformations for observables in $SU(N)$ gauge theory." In: *Physical Review D* 103.9 (May 2021), p. 094517. ISSN: 2470-0010, 2470-0029. DOI: 10.1103/PhysRevD.103.094517. arXiv: 2101.12668.

[65] Yuto Mori, Kouji Kashiwa, and Akira Ohnishi. "Application of neural network to sign problem via path optimization method." In: *Progress of Theoretical and Experimental Physics* 2018.2 (Feb. 2018). ISSN: 2050-3911. DOI: 10.1093/ptep/ptx191. arXiv: 1709.03208.

[66] John Hubbard. "Electron correlations in narrow energy bands." In: *Proceedings of the Royal Society of London. Series A. Mathematical and Physical Sciences* 276.1365 (Nov. 1963), pp. 238–257. ISSN: 0080-4630, 2053-9169. DOI: 10.1098/rspa.1963.0204.

[67] Abhishek Mukherjee and Marco Cristoforetti. "Lefschetz thimble Monte Carlo for many-body theories: A Hubbard model study." In: *Physical Review B* 90.3 (July 2014), p. 035134. ISSN: 1098-0121, 1550-235X. DOI: 10.1103/PhysRevB.90.035134. arXiv: 1403.5680v1.

[68] Masafumi Fukuma, Nobuyuki Matsumoto, and Naoya Umeda. "Applying the tempered Lefschetz thimble method to the Hubbard model away from half-filling." In: *Physical Review D* 100.11 (Dec. 2019), p. 114510. ISSN: 2470-0010, 2470-0029. DOI: 10.1103/PhysRevD.100.114510. arXiv: 1906.04243,2001.01665v1.

[69] Maksim Ulybyshev, Christopher Winterowd, and Savvas Zafeiropoulos. "Lefschetz thimbles decomposition for the Hubbard model on the hexagonal lattice." In: *Physical Review D* 101.1 (Jan. 2020), p. 014508. ISSN: 2470-0010, 2470-0029. DOI: 10.1103/PhysRevD.101.014508. arXiv: 1906.07678.

[70] Manuel Schneider et al. "Simulating both parity sectors of the Hubbard Model with Tensor Networks." In: *Physical Review B* 104.15 (Oct. 2021), p. 155118. ISSN: 2469-9950, 2469-9969. DOI: 10.1103/PhysRevB.104.155118. arXiv: 2106.13583.

[71] Yuto Mori, Kouji Kashiwa, and Akira Ohnishi. "Lefschetz thimbles in fermionic effective models with repulsive vector-field." In: *Physics Letters B* 781 (June 2018), pp. 688–693. ISSN: 03702693. DOI: 10.1016/j.physletb.2018.04.018. arXiv: 1705.03646.

[72] K. Splittorff and J. J. M. Verbaarschot. "Phase of the Fermion Determinant at Nonzero Chemical Potential." In: *Physical Review Letters* 98.3 (Jan. 2007), p. 031601. ISSN: 0031-9007, 1079-7114. DOI: 10.1103/PhysRevLett.98.031601. arXiv: hep-lat/0609076.

[73] AuroraScience Collaboration et al. "New approach to the sign problem in quantum field theories: High density QCD on a Lefschetz thimble." In: *Physical Review D* 86.7 (Oct. 2012), p. 074506. ISSN: 1550-7998, 1550-2368. DOI: 10.1103/PhysRevD.86.074506. arXiv: 1205.3996.

[74] Yuya Tanizaki, Hiromichi Nishimura, and Jacobus J. M. Verbaarschot. "Gradient flows without blow-up for Lefschetz thimbles." In: *Journal of High Energy Physics* 2017.10 (Oct. 2017), p. 100. ISSN: 1029-8479. DOI: 10.1007/JHEP10(2017)100. arXiv: 1706.03822.

[75] Andrei Alexandru et al. "Fast Estimator of jacobians in Monte Carlo Integration on Lefschetz Thimbles." In: *Physical Review D* 93.9 (May 2016), p. 094514. ISSN: 2470-0010, 2470-0029. DOI: 10.1103/PhysRevD.93.094514. arXiv: 1604.00956.

[76] Genki Fujisawa et al. "Backpropagating Hybrid Monte Carlo algorithm for fast Lefschetz thimble calculations." In: *Journal of High Energy Physics* 2022.4 (Apr. 2022), p. 179. ISSN: 1029-8479. DOI: 10.1007/JHEP04(2022)179. arXiv: 2112.10519.

[77] Simon Duane et al. "Hybrid Monte Carlo." In: *Physics Letters B* 195.2 (Sept. 1987), pp. 216–222. ISSN: 03702693. DOI: 10.1016/0370-2693(87)91197-X.

[78] Andrei Alexandru et al. "Fermions at Finite Density in (2+1)d with Sign-Optimized Manifolds." In: *Physical Review Letters* 121.19 (Nov. 2018), p. 191602. ISSN: 0031-9007, 1079-7114. DOI: 10.1103/PhysRevLett.121.191602. arXiv: 1808.09799.

[79] Neill C Warrington. "Taming the Sign Problem in Lattice Field Theory With Deformed Path Integral Contours." PhD thesis. Maryland: Maryland, 2019. DOI: 10.13016/F2WT-JEN0.

[80] H. Fujii et al. "Hybrid Monte Carlo on Lefschetz Thimbles – A study of the residual sign problem." In: *Journal of High Energy Physics* 2013.10 (Oct. 2013), p. 147. ISSN: 1029-8479. DOI: 10.1007/JHEP10(2013)147. arXiv: 1309.4371.

[81] H. Fujii, S. Kamata, and Y. Kikukawa. "Lefschetz thimble structure in one-dimensional lattice Thirring model at finite density." In: *Journal of High Energy Physics* 2015.11 (Nov. 2015), p. 78. ISSN: 1029-8479. DOI: 10.1007/JHEP11(2015)078. arXiv: 1509.08176.

[82] Masafumi Fukuma, Nobuyuki Matsumoto, and Naoya Umeda. *Implementation of the HMC algorithm on the tempered Lefschetz thimble method*. Feb. 2020. arXiv: 1912.13303.

[83] Joshua Bassey, Lijun Qian, and Xianfang Li. *A Survey of Complex-Valued Neural Networks*. Jan. 2021. arXiv: 2101.12249.

[84] P. Bouboulis. *Wirtinger's Calculus in general Hilbert Spaces*. May 2010. arXiv: 1005.5170.

[85] D.H. Brandwood. "A complex gradient operator and its application in adaptive array theory." In: *IEE Proceedings H Microwaves, Optics and Antennas* 130.1 (1983), p. 11. ISSN: 01437097. DOI: 10.1049/ip-h-1.1983.0004.

[86] Ken Kreutz-Delgado. *The Complex Gradient Operator and the CR-Calculus*. June 2009. arXiv: 0906.4835.

[87] Felix Voigtlaender. "The universal approximation theorem for complex-valued neural networks." In: *Applied and Computational Harmonic Analysis* 64 (May 2023), pp. 33–61. ISSN: 10635203. DOI: 10.1016/j.acha.2022.12.002. arXiv: 2012.03351v2.

[88] Michael S. Albergo et al. *Introduction to Normalizing Flows for Lattice Field Theory*. Aug. 2021. arXiv: 2101.08176.

[89] Sam Foreman et al. "HMC with Normalizing Flows." In: *Proceedings of The 38th International Symposium on Lattice Field Theory — PoS(LATTICE2021)*. Online@MIT: Sissa Medialab, May 2022, p. 073. DOI: 10.22323/1.396.0073. arXiv: 2112.01586v1.

[90] Adam Paszke et al. "PyTorch: An Imperative Style, High-Performance Deep Learning Library." In: *Advances in Neural Information Processing Systems 32*. Ed. by H. Wallach et al. Curran Associates, Inc., 2019, pp. 8024–8035. URL: http://papers.neurips.cc/paper/9015-pytorch-an-imperative-style-high-performance-deep-learning-library.pdf.

[91] Jülich Supercomputing Centre. "JURECA: Modular supercomputer at Jülich Supercomputing Centre." In: *Journal of large-scale research facilities* 4.A132 (2018). DOI: 10.17815/jlsrf-4-121-1.

[92] Matteo Favoni et al. "Lattice gauge equivariant convolutional neural networks." In: *Physical Review Letters* 128.3 (Jan. 2022), p. 032003. ISSN: 0031-9007, 1079-7114. DOI: 10.1103/PhysRevLett.128.032003. arXiv: 2012.12901.

[93] Di Luo et al. "Gauge equivariant neural networks for quantum lattice gauge theories." In: *Physical Review Letters* 127.27 (Dec. 2021), p. 276402. ISSN: 0031-9007, 1079-7114. DOI: 10.1103/PhysRevLett.127.276402. arXiv: 2012.05232.

[94] Gurtej Kanwar et al. "Equivariant flow-based sampling for lattice gauge theory." In: *Physical Review Letters* 125.12 (Sept. 2020), p. 121601. ISSN: 0031-9007, 1079-7114. DOI: 10.1103/PhysRevLett.125.121601. arXiv: 2003.06413.

[95] D. M. Donaldson, John Monteath Robertson, and John White. "The crystal and molecular structure of perylene." In: *Proceedings of the Royal Society of London. Series A. Mathematical and Physical Sciences* 220.1142 (Dec. 1953), pp. 311–321. ISSN: 0080-4630, 2053-9169. DOI: 10.1098/rspa.1953.0189.

[96] Geetanjali Giri et al. "A Model Exact Study of the Properties of Low-Lying Electronic States of Perylene and Substituted Perylenes." In: *The Journal of Physical Chemistry A* 122.43 (Nov. 2018), pp. 8650–8658. ISSN: 1089-5639, 1520-5215. DOI: 10.1021/acs.jpca.8b08656. arXiv: 1809.05909v1.

[97] Mark Botoshansky, Frank H. Herbstein, and Moshe Kapon. "Towards a Complete Description of a Polymorphic Crystal: The Example of Perylene: Redetermination of the Structures of the (Z =2 and 4) Polymorphs." In: *Helvetica Chimica Acta* 86.4 (Apr. 2003), pp. 1113–1128. ISSN: 0018-019X, 1522-2675. DOI: 10.1002/hlca.200390097.

[98] Brian A. Gregg and Russell A. Cormier. "Doping Molecular Semiconductors: n-Type Doping of a Liquid Crystal Perylene Diimide." In: *Journal of the American Chemical Society* 123.32 (Aug. 2001), pp. 7959–7960. ISSN: 0002-7863, 1520-5126. DOI: 10.1021/ja016410k.

[99] Y. Sato, S. Ichinosawa, and H. Kanai. "Operation characteristics and degradation of organic electroluminescent devices." In: *IEEE Journal of Selected Topics in Quantum Electronics* 4.1 (Feb. 1998), pp. 40–48. ISSN: 1077260X. DOI: 10.1109/2944.669464.

[100] C. W. Tang. "Two-layer organic photovoltaic cell." In: *Applied Physics Letters* 48.2 (Jan. 1986), pp. 183–185. ISSN: 0003-6951, 1077-3118. DOI: 10.1063/1.96937.

[101] Farid Salama. "PAHs in Astronomy - A Review." In: *Proceedings of the International Astronomical Union* 4.S251 (Feb. 2008), pp. 357–366. ISSN: 1743-9213, 1743-9221. DOI: 10.1017/S1743921308021960.

[102] Aigen Li. "Spitzer's perspective of polycyclic aromatic hydrocarbons in galaxies." In: *Nature Astronomy* 4.4 (Mar. 2020), pp. 339–351. ISSN: 2397-3366. DOI: 10.1038/s41550-020-1051-1. arXiv: 2003.10489v1.

[103] Akira Hirose, ed. *Complex-Valued Neural Networks: Advances and Applications*. 1st ed. Wiley, Apr. 2013. ISBN: 978-1-118-34460-6. DOI: 10.1002/9781118590072.

[104] ChiYan Lee, Hideyuki Hasegawa, and Shangce Gao. "Complex-Valued Neural Networks: A Comprehensive Survey." In: *IEEE/CAA Journal of Automatica Sinica* 9.8 (Aug. 2022), pp. 1406–1426. ISSN: 2329-9266, 2329-9274. DOI: 10.1109/JAS.2022.105743.

[105]   Xavier Glorot and Yoshua Bengio. "Understanding the difficulty of training deep feedforward neural networks." In: *Proceedings of the Thirteenth International Conference on Artificial Intelligence and Statistics*. Vol. 9. Proceedings of Machine Learning Research. Chia Laguna Resort, Sardinia, Italy: PMLR, May 2010, pp. 249–256. URL: https://proceedings.mlr.press/v9/glorot10a.html.

[106]   Diederik P. Kingma and Jimmy Ba. *Adam: A Method for Stochastic Optimization*. Jan. 2017. arXiv: 1412.6980.

[107]   Qiang Guo et al. "Effect of Energy Alignment, Electron Mobility, and Film Morphology of Perylene Diimide Based Polymers as Electron Transport Layer on the Performance of Perovskite Solar Cells." In: *ACS Applied Materials & Interfaces* 9.12 (Mar. 2017), pp. 10983–10991. ISSN: 1944-8244, 1944-8252. DOI: 10.1021/acsami.7b00902.

[108]   Wenjun Ni, Licheng Sun, and Gagik G. Gurzadyan. "Ultrafast spectroscopy reveals singlet fission, ionization and excimer formation in perylene film." In: *Scientific Reports* 11.1 (Mar. 2021), p. 5220. ISSN: 2045-2322. DOI: 10.1038/s41598-021-83791-z.

[109]   Jin Cao and Shaopeng Yang. "Progress in perylene diimides for organic solar cell applications." In: *RSC Advances* 12.12 (2022), pp. 6966–6973. ISSN: 2046-2069. DOI: 10.1039/D1RA08484D.

[110]   Thomas M. Halasinski et al. "Electronic Absorption Spectra of Neutral Perylene ($C_{20} H_{12}$), Terrylene ($C_{30} H_{16}$), and Quaterrylene ($C_{40} H_{20}$) and Their Positive and Negative Ions: Ne Matrix-Isolation Spectroscopy and Time-Dependent Density Functional Theory Calculations." In: *The Journal of Physical Chemistry A* 107.19 (May 2003), pp. 3660–3669. ISSN: 1089-5639, 1520-5215. DOI: 10.1021/jp027394w.

[111]   Maria I. Shchuka, Andrea L. Motyka, and Michael R. Topp. "Two-photon threshold ionization spectroscopy of perylene and van der waals complexes." In: *Chemical Physics Letters* 164.1 (Dec. 1989), pp. 87–95. ISSN: 00092614. DOI: 10.1016/0009-2614(89)85208-X.

[112]   J. Schiedt and R. Weinkauf. "Photodetachment photoelectron spectroscopy of perylene and CS2: two extreme cases." In: *Chemical Physics Letters* 274.1-3 (Aug. 1997), pp. 18–22. ISSN: 00092614. DOI: 10.1016/S0009-2614(97)00648-9.

[113]   A. Manian et al. "Singlet Exciton Dynamics of Perylene Diimide- and Tetracene-Based Hetero/Homogeneous Substrates via an *Ab Initio* Kinetic Monte Carlo Model." In: *The Journal of Physical Chemistry C* 125.43 (Nov. 2021), pp. 23646–23656. ISSN: 1932-7447, 1932-7455. DOI: 10.1021/acs.jpcc.1c07929.

[114]   Gabriele D'Avino et al. "Electron–Hole Separation in Perylene Diimide Based Self-Assembled Nanostructures: Microelectrostatics Analysis and Kinetic Monte Carlo Simulations." In: *The Journal of Physical Chemistry C* 126.23 (June 2022), pp. 9762–9776. ISSN: 1932-7447, 1932-7455. DOI: 10.1021/acs.jpcc.2c00527.

[115]   Aurora E. Clark, Changyong Qin, and Alexander D. Q. Li. "Beyond Exciton Theory: A Time-Dependent DFT and Franck-Condon Study of Perylene Diimide and Its Chromophoric Dimer." In: *Journal of the American Chemical Society* 129.24 (June 2007), pp. 7586–7595. ISSN: 0002-7863, 1520-5126. DOI: 10.1021/ja0687724.

[116]   Maksim Ulybyshev, Christopher Winterowd, and Savvas Zafeiropoulos. *Taming the sign problem of the finite density Hubbard model via Lefschetz thimbles*. June 2019. arXiv: 1906.02726.

[117]   Yang Zhao et al. "Theoretical study of one-photon and two-photon absorption properties of perylene tetracarboxylic derivatives." In: *The Journal of Chemical Physics* 129.1 (July 2008), p. 014301. ISSN: 0021-9606, 1089-7690. DOI: 10.1063/1.2938374.

[118] Yiding Lai et al. "Synthesis of Ultrathin Nanosheets of Perylene." In: *Crystal Growth & Design* 15.3 (Mar. 2015), pp. 1011–1016. ISSN: 1528-7483, 1528-7505. DOI: 10.1021/cg5015016.

[119] Dominik Smith and Lorenz Smekal. "Monte-Carlo simulation of the tight-binding model of graphene with partially screened Coulomb interactions." In: *Physical Review B* 89.19 (May 2014), p. 195429. ISSN: 1098-0121. DOI: 10.1103/PhysRevB.89.195429. arXiv: 1403.3620.

[120] Maksim Ulybyshev et al. "Monte-Carlo study of the semimetal-insulator phase transition in monolayer graphene with realistic inter-electron interaction potential." In: *Physical Review Letters* 111.5 (July 2013), p. 056801. ISSN: 0031-9007. DOI: 10.1103/PhysRevLett.111.056801. arXiv: 1304.3660.

[121] Mario Motta and Shiwei Zhang. "Ab initio computations of molecular systems by the auxiliary-field quantum Monte Carlo method." In: *WIREs Computational Molecular Science* 8.5 (2018), e1364. DOI: https://doi.org/10.1002/wcms.1364. arXiv: 1711.02242. URL: https://wires.onlinelibrary.wiley.com/doi/abs/10.1002/wcms.1364.

[122] Kouji Kashiwa, Yuto Mori, and Akira Ohnishi. "Application of the path optimization method to the sign problem in an effective model of QCD with a repulsive vector-type interaction." In: *Physical Review D* 99.11 (June 2019), p. 114005. ISSN: 2470-0010, 2470-0029. DOI: 10.1103/PhysRevD.99.114005. arXiv: 1903.03679.

[123] Scott Lawrence. "Beyond Thimbles: Sign-Optimized Manifolds for Finite Density." In: *Proceedings of The 36th Annual International Symposium on Lattice Field Theory — PoS(LATTICE2018)*. Michigan State University, East Lansing, Michigan, USA.: Sissa Medialab, May 2019, p. 149. DOI: 10.22323/1.334.0149. arXiv: 1810.06529v2.

[124] Scott Lawrence and Yukari Yamauchi. *Convex optimization of contour deformations*. Nov. 2023. arXiv: 2311.13002.

[125] Ian E. Jacobs and Adam J. Moulé. "Controlling Molecular Doping in Organic Semiconductors." In: *Advanced Materials* 29.42 (Nov. 2017), p. 1703063. ISSN: 0935-9648, 1521-4095. DOI: 10.1002/adma.201703063.

[126] William I. Jay and Ethan T. Neil. "Bayesian model averaging for analysis of lattice field theory results." In: *Physical Review D* 103.11 (June 2021), p. 114502. ISSN: 2470-0010, 2470-0029. DOI: 10.1103/PhysRevD.103.114502. arXiv: 2008.01069. (Visited on 06/02/2024).

[127] Ethan T. Neil and Jacob W. Sitison. "Improved information criteria for Bayesian model averaging in lattice field theory." In: *Physical Review D* 109.1 (Jan. 2024), p. 014510. ISSN: 2470-0010, 2470-0029. DOI: 10.1103/PhysRevD.109.014510. arXiv: 2208.14983.

[128] Ethan T. Neil and Jacob W. Sitison. "Model averaging approaches to data subset selection." In: *Physical Review E* 108.4 (Oct. 2023), p. 045308. ISSN: 2470-0045, 2470-0053. DOI: 10.1103/PhysRevE.108.045308. arXiv: 2305.19417.

[129] A. Sokal. "Monte Carlo Methods in Statistical Mechanics: Foundations and New Algorithms." In: *Functional Integration*. Ed. by Cecile DeWitt-Morette, Pierre Cartier, and Antoine Folacci. Vol. 361. Boston, MA: Springer US, 1997, pp. 131–192. ISBN: 978-1-4899-0321-1. DOI: 10.1007/978-1-4899-0319-8_6.

[130] Ulli Wolff. "Monte Carlo errors with less errors." In: *Computer Physics Communications* 176.5 (Mar. 2007), p. 383. ISSN: 00104655. DOI: 10.1016/j.cpc.2006.12.001. arXiv: hep-lat/0306017.

[131] M. Fischer et al. "On the generalised eigenvalue method and its relation to Prony and generalised pencil of function methods." In: *The European Physical Journal A* 56.8 (Aug. 2020), p. 206. ISSN: 1434-6001, 1434-601X. DOI: 10.1140/epja/s10050-020-00205-w. arXiv: 2004.10472.

[132] Peter Lepage and Christoph Gohlke. *gplepage/lsqfit: lsqfit version 13.0.1*. 2023. DOI: 10.5281/ZENODO.7931361. URL: https://zenodo.org/record/7931361.

[133] S. Durr et al. "Lattice QCD at the physical point: Simulation and analysis details." In: *Journal of High Energy Physics* 2011.8 (Aug. 2011), p. 148. ISSN: 1029-8479. DOI: 10.1007/JHEP08(2011)148. arXiv: 1011.2711.

[134] S. Durr et al. "Lattice QCD at the physical point: light quark masses." In: *Physics Letters B* 701.2 (July 2011), pp. 265–268. ISSN: 03702693. DOI: 10.1016/j.physletb.2011.05.053. arXiv: 1011.2403.

[135] George Sterman et al. "Handbook of perturbative QCD." In: *Reviews of Modern Physics* 67.1 (Jan. 1995), pp. 157–248. ISSN: 0034-6861, 1539-0756. DOI: 10.1103/RevModPhys.67.157.

[136] Hung-Liang Lai et al. "New parton distributions for collider physics." In: *Physical Review D* 82.7 (Oct. 2010), p. 074024. ISSN: 1550-7998, 1550-2368. DOI: 10.1103/PhysRevD.82.074024. arXiv: 1007.2241v3.

[137] Shujie Li et al. *Extraction of the neutron F2 structure function from inclusive proton and deuteron deep-inelastic scattering data*. Oct. 2023. arXiv: 2309.16851.

[138] Martha Constantinou et al. *Lattice QCD Calculations of Parton Physics*. Feb. 2022. arXiv: 2202.07193.

[139] Jonathan D. Bratt et al. "Nucleon structure from mixed action calculations using 2+1 flavors of asqtad sea and domain wall valence fermions." In: *Physical Review D* 82.9 (Nov. 2010), p. 094502. ISSN: 1550-7998, 1550-2368. DOI: 10.1103/PhysRevD.82.094502. arXiv: 1001.3620.

[140] Gunnar S. Bali et al. "Nucleon axial structure from lattice QCD." In: *Journal of High Energy Physics* 2020.5 (May 2020), p. 126. ISSN: 1029-8479. DOI: 10.1007/JHEP05(2020)126. arXiv: 1911.13150.

[141] Lorenzo Barca, Gunnar Bali, and Sara Collins. "Toward N to $N\pi$ matrix elements from lattice QCD." In: *Physical Review D* 107.5 (Mar. 2023), p. L051505. ISSN: 2470-0010, 2470-0029. DOI: 10.1103/PhysRevD.107.L051505. arXiv: 2211.12278.

[142] Guido Martinelli and Christopher T. Sachrajda. "A lattice calculation of the pion's form factor and structure function." In: *Nuclear Physics B* 306.4 (Sept. 1988), pp. 865–889. ISSN: 05503213. DOI: 10.1016/0550-3213(88)90445-2.

[143] Guido Martinelli and Christopher T. Sachrajda. "A lattice study of nucleon structure." In: *Nuclear Physics B* 316.2 (Apr. 1989), pp. 355–372. ISSN: 05503213. DOI: 10.1016/0550-3213(89)90035-7.

[144] Meinulf Göckeler et al. "A lattice determination of moments of unpolarized nucleon structure functions using improved Wilson fermions." In: *Physical Review D* 71.11 (June 2005), p. 114511. ISSN: 1550-7998, 1550-2368. DOI: 10.1103/PhysRevD.71.114511. arXiv: hep-ph/0410187v1.

[145] Stefano Capitani et al. "Systematic errors in extracting nucleon properties from lattice QCD." In: *Proceedings of The XXVIII International Symposium on Lattice Field Theory — PoS(Lattice 2010)*. Villasimius, Sardinia Italy: Sissa Medialab, June 2011, p. 147. DOI: 10.22323/1.105.0147. arXiv: 1011.1358v1.

[146] John Bulava, Michael Donnellan, and Rainer Sommer. "The B*Bpi coupling." In: *Proceedings of The XXVIII International Symposium on Lattice Field Theory — PoS(Lattice 2010)*. Villasimius, Sardinia Italy: Sissa Medialab, June 2011, p. 303. DOI: 10.22323/1.105.0303. arXiv: 1011.4393v1.

[147] Cédric. Lorcé and B. Pasquini. "Quark Wigner distributions and orbital angular momentum." In: *Physical Review D* 84.1 (July 2011), p. 014015. ISSN: 1550-7998, 1550-2368. DOI: 10.1103/PhysRevD.84.014015. arXiv: 1106.0139v1.

[148] Cédric Lorcé. "Spin–orbit correlations in the nucleon." In: *Physics Letters B* 735 (July 2014), pp. 344–348. ISSN: 03702693. DOI: 10.1016/j.physletb.2014.06.068. arXiv: 1401.7784v2.

[149] Abha Rajan, Michael Engelhardt, and Simonetta Liuti. "Lorentz Invariance and QCD Equation of Motion Relations for Generalized Parton Distributions and the Dynamical Origin of Proton Orbital Angular Momentum." In: *Physical Review D* 98.7 (Oct. 2018), p. 074022. ISSN: 2470-0010, 2470-0029. DOI: 10.1103/PhysRevD.98.074022. arXiv: 1709.05770.

[150] Stephan Meißner, Andreas Metz, and Marc Schlegel. "Generalized parton correlation functions for a spin-1/2 hadron." In: *Journal of High Energy Physics* 2009.08 (Aug. 2009), pp. 056–056. ISSN: 1029-8479. DOI: 10.1088/1126-6708/2009/08/056. arXiv: 0906.5323v2.

[151] Michael Engelhardt et al. "Quark spin-orbit correlations in the proton." In: *Proceedings of The 38th International Symposium on Lattice Field Theory — PoS(LATTICE2021)*. Zoom/Gather@MIT: Sissa Medialab, May 2022, p. 413. DOI: 10.22323/1.396.0413. arXiv: 2112.13464v1.

[152] Stephan Güsken. "A study of smearing techniques for hadron correlation functions." In: *Nuclear Physics B - Proceedings Supplements* 17 (Sept. 1990), pp. 361–364. ISSN: 09205632. DOI: 10.1016/0920-5632(90)90273-W.

[153] Stefano Capitani et al. "Renormalisation and off-shell improvement in lattice perturbation theory." In: *Nuclear Physics B* 593.1-2 (Jan. 2001), pp. 183–228. ISSN: 05503213. DOI: 10.1016/S0550-3213(00)00590-3. arXiv: hep-lat/0007004v1.

[154] Yi-Bo Yang et al. "Proton Mass Decomposition from the QCD Energy Momentum Tensor." In: *Physical Review Letters* 121.21 (Nov. 2018), p. 212001. ISSN: 0031-9007, 1079-7114. DOI: 10.1103/PhysRevLett.121.212001. arXiv: 1808.08677v2.

[155] C. Alexandrou et al. "Complete flavor decomposition of the spin and momentum fraction of the proton using lattice QCD simulations at physical pion mass." In: *Physical Review D* 101.9 (May 2020), p. 094513. ISSN: 2470-0010, 2470-0029. DOI: 10.1103/PhysRevD.101.094513. arXiv: 2003.08486v1.

[156] Johannes Blümlein and Helmut Böttcher. "QCD analysis of polarized deep inelastic scattering data." In: *Nuclear Physics B* 841.1-2 (Dec. 2010), pp. 205–230. ISSN: 05503213. DOI: 10.1016/j.nuclphysb.2010.08.005. arXiv: 1005.3113v2.

[157] Andrew Pochinsky. *Qlua*. URL: https://usqcd.lns.mit.edu/qlua.

[158] Ronald Babich et al. "Adaptive Multigrid Algorithm for the Lattice Wilson-Dirac Operator." In: *Physical Review Letters* 105.20 (Nov. 2010), p. 201602. ISSN: 0031-9007, 1079-7114. DOI: 10.1103/PhysRevLett.105.201602. arXiv: 1005.3043v2.

[159] James Osborn et al. *QOPQDP*. URL: https://usqcd-software.github.io/qopqdp/.

[160] Jülich Supercomputing Centre. "JUQUEEN: IBM Blue Gene/Q Supercomputer System at the Jülich Supercomputing Centre." In: *Journal of large-scale research facilities* 1.A1 (2015). DOI: 10.17815/jlsrf-1-18.

[161] Jülich Supercomputing Centre. "JUWELS: Modular Tier-0/1 Supercomputer at the Jülich Supercomputing Centre." In: *Journal of large-scale research facilities* 5.A135 (2019). DOI: 10.17815/jlsrf-5-171.

[162] G. Martinelli et al. "A general method for non-perturbative renormalization of lattice operators." In: *Nuclear Physics B* 445.1 (July 1995), pp. 81–105. ISSN: 05503213. DOI: 10.1016/0550-3213(95)00126-D. arXiv: hep-lat/9411010v1.

[163] Meinulf Göckeler et al. "Nonperturbative Renormalisation of Composite Operators in Lattice QCD." In: *Nuclear Physics B* 544.3 (Apr. 1999), pp. 699–733. ISSN: 05503213. DOI: 10.1016/S0550-3213(99)00036-X. arXiv: hep-lat/9807044.

[164] Christian Sturm et al. "Renormalization of quark bilinear operators in a momentum-subtraction scheme with a nonexceptional subtraction point." In: *Physical Review D* 80.1 (July 2009), p. 014501. ISSN: 1550-7998, 1550-2368. DOI: 10.1103/PhysRevD.80.014501. arXiv: 0901.2599.

[165] John A. Gracey. "Three loop anomalous dimension of the second moment of the transversity operator in the and RI' schemes." In: *Nuclear Physics B* 667.1-2 (Sept. 2003), pp. 242–260. ISSN: 05503213. DOI: 10.1016/S0550-3213(03)00543-1. arXiv: hep-ph/0306163v2.

[166] John A. Gracey. "Two loop renormalization of the n = 2 Wilson operator in the RI'/SMOM scheme." In: *Journal of High Energy Physics* 2011.3 (Mar. 2011), p. 109. ISSN: 1029-8479. DOI: 10.1007/JHEP03(2011)109. arXiv: 1103.2055v1.

[167] Jonathan Flynn et al. "Nonperturbative renormalisation for low moments of light-meson distribution amplitudes." In: *Proceedings of The 32nd International Symposium on Lattice Field Theory — PoS(LATTICE2014)*. Columbia University New York, NY: Sissa Medialab, May 2015, p. 168. DOI: 10.22323/1.214.0168.

[168] Bernd A. Kniehl and Oleg L. Veretin. "Moments n = 2 and n = 3 of the Wilson twist-two operators at three loops in the RI′/SMOM scheme." In: *Nuclear Physics B* 961 (Dec. 2020), p. 115229. ISSN: 05503213. DOI: 10.1016/j.nuclphysb.2020.115229. arXiv: 2009.11325v2.

[169] Pavel A. Baikov and Konstantin G. Chetyrkin. "New Four Loop Results in QCD." In: *Nuclear Physics B - Proceedings Supplements* 160 (Oct. 2006), pp. 76–79. ISSN: 09205632. DOI: 10.1016/j.nuclphysbps.2006.09.031.

[170] Vitaly N. Velizhanin. "Four loop anomalous dimension of the second moment of the non-singlet twist-2 operator in QCD." In: *Nuclear Physics B* 860.2 (July 2012), pp. 288–294. ISSN: 05503213. DOI: 10.1016/j.nuclphysb.2012.03.006. arXiv: 1112.3954v2.

[171] P.A. Baikov, K.G. Chetyrkin, and J.H. Kühn. "Massless Propagators, R(s) and Multiloop QCD." In: *Nuclear and Particle Physics Proceedings* 261–262 (Apr. 2015), pp. 3–18. ISSN: 24056014. DOI: 10.1016/j.nuclphysbps.2015.03.002. arXiv: 1501.06739v1.

[172] Vitaly N. Velizhanin. "Four-loop anomalous dimension of the third and fourth moments of the nonsinglet twist-2 operator in QCD." In: *International Journal of Modern Physics A* 35.32 (Nov. 2020), p. 2050199. ISSN: 0217-751X, 1793-656X. DOI: 10.1142/S0217751X20501997. arXiv: 1411.1331v2.

[173] Vladimir M. Braun et al. "The ρ-meson light-cone distribution amplitudes from lattice QCD." In: *Journal of High Energy Physics* 2017.4 (Apr. 2017), p. 82. ISSN: 1029-8479. DOI: 10.1007/JHEP04(2017)082. arXiv: 1612.02955v2.

[174] Ph. Boucaud et al. "Artifacts and $<A^2>$ power corrections: Reexamining $Z_\psi(p_2)$ and Z_V in the momentum-subtraction scheme." In: *Physical Review D* 74.3 (Aug. 2006), p. 034505. ISSN: 1550-7998, 1550-2368. DOI: 10.1103/PhysRevD.74.034505. arXiv: hep-lat/0504017v1.

[175] Jan-Lukas Wynen and Evan Berkowitz. *jl-wynen/isle: Release 0.3*. May 2020. DOI: `10.5281/zenodo.3834729`.

[176] Christoph Gäntgen. "Investigating Doped Systems with Holomorphic Flow and Machine Learning." Access upon request (`c.gaentgen@fz-juelich.de`). Rheinischen Friedrich-Wilhelms-Universität, 2021.

[177] Shailesh Chandrasekharan and Uwe-Jens Wiese. "Meron-Cluster Solution of Fermion Sign Problems." In: *Physical Review Letters* 83.16 (Oct. 1999), pp. 3116–3119. ISSN: 0031-9007, 1079-7114. DOI: `10.1103/PhysRevLett.83.3116`. arXiv: `cond-mat/9902128`.

[178] F.-M. Dittes et al. "The Altarelli-Parisi Kernel as Asymptotic Limit of an Extended Brodsky-Lepage Kernel." In: *Physics Letters B* 209.2-3 (Aug. 1988), pp. 325–329. ISSN: 03702693. DOI: `10.1016/0370-2693(88)90955-0`.

[179] Xiangdong Ji. "Gauge-Invariant Decomposition of Nucleon Spin and Its Spin-Off." In: *Physical Review Letters* 78.4 (Jan. 1997), pp. 610–613. ISSN: 0031-9007, 1079-7114. DOI: `10.1103/PhysRevLett.78.610`. arXiv: `hep-ph/9603249`.

[180] John C. Collins and Davison E. Soper. "Parton Distribution and Decay Functions." In: *Nuclear Physics B* 194.3 (Jan. 1982), pp. 445–492. ISSN: 05503213. DOI: `10.1016/0550-3213(82)90021-9`.

[181] K. K. L. B. Adikaram et al. "Data Transformation Technique to Improve the Outlier Detection Power of Grubbs' Test for Data Expected to Follow Linear Relation." In: *Journal of Applied Mathematics* 2015 (2015), pp. 1–9. ISSN: 1110-757X, 1687-0042. DOI: `10.1155/2015/708948`.

[182] Constantia Alexandrou et al. "Moments of nucleon generalized parton distributions from lattice QCD." In: *Physical Review D* 83.11 (June 2011), p. 114513. ISSN: 1550-7998, 1550-2368. DOI: `10.1103/PhysRevD.83.114513`. arXiv: `1104.1600`.

[183] Constantia Alexandrou et al. "Light-Cone Parton Distribution Functions from Lattice QCD." In: *Physical Review Letters* 121.11 (Sept. 2018). Preprint Title: *Reconstruction of light-cone parton distribution functions from lattice QCD simulations at the physical point*, p. 112001. ISSN: 0031-9007, 1079-7114. DOI: `10.1103/PhysRevLett.121.112001`. arXiv: `1803.02685`.

[184] Constantia Alexandrou et al. "Parton distribution functions from lattice QCD using Bayes-Gauss-Fourier transforms." In: *Physical Review D* 102.9 (Nov. 2020), p. 094508. ISSN: 2470-0010, 2470-0029. DOI: `10.1103/PhysRevD.102.094508`. arXiv: `2007.13800`.

[185] Andrei Alexandru et al. "Structure Factors of Neutron Matter at Finite Temperature." In: *Physical Review Letters* 126.13 (Apr. 2021), p. 132701. ISSN: 0031-9007, 1079-7114. DOI: `10.1103/PhysRevLett.126.132701`. arXiv: `2008.02824`.

[186] C. R. Allton et al. "The QCD thermal phase transition in the presence of a small chemical potential." In: *Physical Review D* 66.7 (Oct. 2002), p. 074507. ISSN: 0556-2821, 1089-4918. DOI: `10.1103/PhysRevD.66.074507`. arXiv: `hep-lat/0204010`.

[187] Manjunath Bhat et al. "Continuum limit of parton distribution functions from the pseudo-distribution approach on the lattice." In: *Physical Review D* 106.5 (Sept. 2022), p. 054504. ISSN: 2470-0010, 2470-0029. DOI: `10.1103/PhysRevD.106.054504`. arXiv: `2205.07585`.

[188] Shohini Bhattacharya et al. "Generalized Parton Distributions from Lattice QCD with Asymmetric Momentum Transfer: Unpolarized Quarks." In: *Physical Review D* 106.11 (Dec. 2022), p. 114512. ISSN: 2470-0010, 2470-0029. DOI: `10.1103/PhysRevD.106.114512`. arXiv: `2209.05373`.

[189] Christopher M. Bouchard et al. "B s → K ℓ ν form factors from lattice QCD." In: *Physical Review D* 90.5 (Sept. 2014), p. 054506. ISSN: 1550-7998, 1550-2368. DOI: 10.1103/PhysRevD.90.054506. arXiv: 1406.2279v2.

[190] Philippe de Forcrand and Owe Philipsen. "The QCD phase diagram for small densities from imaginary chemical potential." In: *Nuclear Physics B* 642.1-2 (Oct. 2002), pp. 290–306. ISSN: 05503213. DOI: 10.1016/S0550-3213(02)00626-0. arXiv: hep-lat/0205016.

[191] S. Ejiri et al. "The QCD phase transition at high temperature and low density." In: *Nuclear Physics B - Proceedings Supplements* 119 (May 2003), pp. 538–540. ISSN: 09205632. DOI: 10.1016/S0920-5632(03)01613-X. arXiv: hep-lat/0209012.

[192] Xiang Gao et al. "Lattice QCD Determination of the Bjorken-x Dependence of Parton Distribution Functions at Next-to-next-to-leading Order." In: *Physical Review Letters* 128.14 (Apr. 2022), p. 142003. ISSN: 0031-9007, 1079-7114. DOI: 10.1103/PhysRevLett.128.142003. arXiv: 2112.02208.

[193] Xiang Gao et al. "Unpolarized proton PDF at NNLO from lattice QCD with physical quark masses." In: *Physical Review D* 107.7 (Apr. 2023), p. 074509. ISSN: 2470-0010, 2470-0029. DOI: 10.1103/PhysRevD.107.074509. arXiv: 2212.12569.

[194] Frank E. Grubbs. "Sample Criteria for Testing Outlying Observations." In: *The Annals of Mathematical Statistics* 21.1 (Mar. 1950), pp. 27–58. ISSN: 0003-4851. DOI: 10.1214/aoms/1177729885.

[195] Rajan Gupta. *Introduction to Lattice QCD*. Lecture Notes. July 1998. arXiv: hep-lat/9807028.

[196] Joseph Karpie et al. "The Continuum and Leading Twist Limits of Parton Distribution Functions in Lattice QCD." In: *Journal of High Energy Physics* 2021.11 (Nov. 2021), p. 24. ISSN: 1029-8479. DOI: 10.1007/JHEP11(2021)024. arXiv: 2105.13313.

[197] Huey-Wen Lin. "Nucleon Helicity Generalized Parton Distribution at Physical Pion Mass from Lattice QCD." In: *Physics Letters B* 824 (Jan. 2022), p. 136821. ISSN: 03702693. DOI: 10.1016/j.physletb.2021.136821. arXiv: 2112.07519.

[198] Marius Löffler. "Moments of parton distribution functions for the pion and rho meson from Nf = 2+1 lattice QCD." PhD thesis. Regensburg: Regensburg, 2022. DOI: 10.5283/epub.52287.

[199] E. Y. Loh et al. "Sign problem in the numerical simulation of many-electron systems." In: *Physical Review B* 41.13 (May 1990), pp. 9301–9307. ISSN: 0163-1829, 1095-3795. DOI: 10.1103/PhysRevB.41.9301.

[200] Santanu Mondal et al. "Nucleon Momentum Fraction, Helicity and Transversity from 2+1-flavor Lattice QCD." In: *Journal of High Energy Physics* 2021.4 (Apr. 2021), p. 44. ISSN: 1029-8479. DOI: 10.1007/JHEP04(2021)044. arXiv: 2011.12787.

[201] Aurora Scapellato et al. "Proton generalized parton distributions from lattice QCD." In: *Suplemento de la Revista Mexicana de Física* 3.3 (June 2022). ISSN: 2683-2585. DOI: 10.31349/SuplRevMexFis.3.0308104. arXiv: 2201.06519v1.

[202] Andrea Shindler, Thomas Luu, and Jordy de Vries. "The nucleon electric dipole moment with the gradient flow: the θ-term contribution." In: *Physical Review D* 92.9 (Nov. 2015), p. 094518. ISSN: 1550-7998, 1550-2368. DOI: 10.1103/PhysRevD.92.094518. arXiv: 1507.02343.

[203] Z. G. Soos and S. Ramasesha. "Valence bond approach to exact nonlinear optical properties of conjugated systems." In: *The Journal of Chemical Physics* 90.2 (Jan. 1989), pp. 1067–1076. ISSN: 0021-9606, 1089-7690. DOI: 10.1063/1.456160. Corrected in "Erratum: Valence bond approach to exact nonlinear optical properties of conjugated systems [J. Chem. Phys. **9 0** , 1067 (1989)]." In: *The Journal of Chemical Physics* 92.8 (Apr. 1990), pp. 5166–5167. ISSN: 0021-9606, 1089-7690. DOI: 10.1063/1.458623.

[204] Z. G. Soos and S. Ramasesha. "Erratum: Valence bond approach to exact nonlinear optical properties of conjugated systems [J. Chem. Phys. **9 0** , 1067 (1989)]." In: *The Journal of Chemical Physics* 92.8 (Apr. 1990), pp. 5166–5167. ISSN: 0021-9606, 1089-7690. DOI: 10.1063/1.458623.

[205] Jian-Hui Zhang et al. "First direct lattice-QCD calculation of the x-dependence of the pion parton distribution function." In: *Physical Review D* 100.3 (Aug. 2019), p. 034505. ISSN: 2470-0010, 2470-0029. DOI: 10.1103/PhysRevD.100.034505. arXiv: 1804.01483.

List of Figures

List of Tables