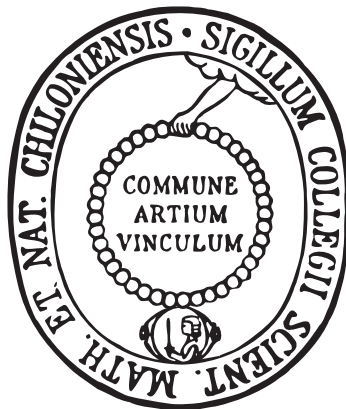Christian-Albrechts-Universität zu Kiel
Institut für Theoretische Physik und Astrophysik

Configuration Path Integral Monte Carlo: Ab initio simulations of fermions in the warm dense matter regime

Dissertation

zur Erlangung des Doktorgrades
der Mathematisch-Naturwissenschaftlichen Fakultät
der Christian-Albrechts-Universität zu Kiel

vorgelegt von

Tim Schoof

Kiel, Dezember 2016

Publications

Journal publications

[1] T. Schoof, M. Bonitz, A. Filinov, D. Hochstuhl, and J. Dufty, "Configuration Path Integral Monte Carlo", Contrib. Plasma Phys. **51**, 687 (2011).

[2] T. Schoof, S. Groth, and M. Bonitz, "Towards ab initio thermodynamics of the electron gas at strong degeneracy", Contrib. Plasma Phys. **55**, 136 (2015).

[3] T. Schoof, S. Groth, J. Vorberger, and M. Bonitz, "*Ab initio* thermodynamic results for the degenerate electron gas at finite temperature", Phys. Rev. Lett. **115**, 130402 (2015).

[4] T. Dornheim, T. Schoof, S. Groth, A. Filinov, and M. Bonitz, "Permutation blocking path integral Monte Carlo approach to the uniform electron gas at finite temperature", J. Chem. Phys. **143**, 204101 (2015).

[5] S. Groth, T. Schoof, T. Dornheim, and M. Bonitz, "*Ab initio* quantum Monte Carlo simulations of the uniform electron gas without fixed nodes", Phys. Rev. B **93**, 085102 (2016).

[6] T. Dornheim, S. Groth, T. Schoof, C. Hann, and M. Bonitz, "*Ab initio* quantum Monte Carlo simulations of the uniform electron gas without fixed nodes: The unpolarized case", Phys. Rev. B **93**, 205134 (2016).

[7] M. Bonitz, E. Pehlke, and T. Schoof, "Attractive forces between ions in quantum plasmas: Failure of linearized quantum hydrodynamics", Phys. Rev. E **87**, 033105 (2013).

[8] M. Bonitz, E. Pehlke, and T. Schoof, "Reply to "Comment on 'Attractive forces between ions in quantum plasmas: Failure of linearized quantum hydrodynamics'"", Phys. Rev. E **87**, 037102 (2013).

[9] M. Bonitz, E. Pehlke, and T. Schoof, "Comment on 'Discussion on novel attractive force between ions in quantum plasmas—failure of simulations based on a density functional approach'", Phys. Scr. **88**, 057001 (2013).

[10]   Z. Moldabekov, T. Schoof, P. Ludwig, M. Bonitz, and T. Ramazanov, "Statically screened ion potential and Bohm potential in a quantum plasma", Physics of Plasmas **22**, 102104 (2015).

Book chapters

[1]   T. Schoof, S. Groth, and M. Bonitz, "Introduction to Configuration Path Integral Monte Carlo", in *Complex plasmas: scientific challenges and technological opportunities*, edited by M. Bonitz, J. Lopez, K. Becker, and H. Thomsen, (Springer, New York, 2014), pp. 153–194.

Selected contributions to workshops and conferences

[1]   T. Schoof, "Ab initio thermodynamic results for the degenerate electron gas at finite temperatures ", contributed talk at the APS March Meeting, Baltimore (2016).

[2]   T. Schoof, "Thermodynamic properties of the homogeneous electron gas - a Configuration Path Integral Monte Carlo approach", contributed talk at the conference SCCS, Santa Fe (2014).

[3]   T. Schoof, "Thermodynamic properties of the homogeneous electron gas - a Configuration Path Integral Monte Carlo approach", poster prize at the CECAM workshop "Stochastic Wavefunction Methods in Quantum Chemistry, Electronic Structure Theory and Condensed Matter Physics", Lausanne (2015).

Abstract

The reliable quantum mechanical description of thermodynamic properties of fermionic many-body systems at high densities and strong degeneracy is of increasing interest due to recent experimental progress in generating systems that exhibit a non-trivial interplay of quantum, temperature, and coupling effects. While quantum Monte Carlo methods are among the most accurate approaches for the description of the ground state, finite-temperature path integral Monte Carlo (PIMC) simulations cannot correctly describe weakly to moderately coupled and strongly degenerate Fermi systems due to the so-called fermion sign problem. By switching from the coordinate representation to a basis of anti-symmetric Slater-determinants, the Configuration Path Integral Monte Carlo (CPIMC) approach greatly reduces the sign problem and allows for the exact computation of thermodynamic properties in this regime.

During this work, the CPIMC algorithm was greatly improved in terms of efficiency and accessible observables. The first successful implementation of the diagrammatic worm algorithm for a general Hamiltonian in Fock space with arbitrary pair interactions gives direct access to the Matsubara Green function. This allows for the reconstruction of dynamic properties from simulations in thermodynamic equilibrium and significantly reduces the statistical variance of derived estimators, such as the one-particle density. The strongly improved MC sampling, the much more efficient calculation of update probabilities, and the successful parallelization to thousands of CPU cores, which have been achieved as part of the new implementation, are essential for the subsequent application of the method to much larger systems than in previous works.

This thesis demonstrates the capabilities of the CPIMC approach for a model system of Coulomb interacting fermions in a two-dimensional harmonic trap. The correctness of the CPIMC implementation is verified by rigorous comparisons with an exact diagonalization method. Benchmark results are presented which reveal large errors of the Hartree-Fock approximation in open shell configurations even for weak coupling strengths and a significant deviation of multi-level blocking PIMC data in the complete basis set limit.

The application of the CPIMC method to the warm dense homogeneous electron

gas (HEG) quantifies the accuracy of recently published restricted PIMC (RPIMC) results which have been the basis for the construction of exchange-correlation free energy functionals to be used in finite-temperature density functional theory calculations of warm dense matter. It is shown that the errors of the RPIMC data exceed 10 % at intermediate densities. Additionally, highly accurate data for the exchange-correlation energy at high densities, which are inaccessible by the RPIMC method, are provided in this work. These results are useful to significantly increase the quality of future exchange-correlation functionals for finite-temperature applications.

The benchmark results of this work have already been used in the development of other methods such as density-matrix QMC. In particular, thorough comparisons have been crucial for the successful verification of the accuracy of the permutation blocking PIMC method, which extends PIMC calculations to much higher densities and degeneracies. The combination of the complementary CPIMC and PBPIMC methods allows for an almost complete description of the HEG at finite temperatures in the thermodynamic limit.

Zusammenfassung

Die zuverlässige quantenmechanische Beschreibung thermodynamischer Eigenschaften fermionischer Vielteilchensysteme bei großen Dichten und starker Entartung ist auf Grund des jüngsten experimentellen Fortschritts bei der Erzeugung von Materiezuständen, die sich durch ein nichttriviales Zusammenspiel von Quanten-, Temperatur- und Kopplungseffekten auszeichnen, von wachsender Bedeutung. Während Quanten-Monte-Carlo-Verfahren zu den genauesten und erfolgreichsten Ansätzen zur Berechnung des Grundzustands gehören, können Pfadintegral-Monte-Carlo-Simulationen (PIMC) bei endlichen Temperaturen schwach bis moderat gekoppelte und stark entartete Fermi-Systeme wegen des sogenannten fermionischen Vorzeichenproblems nicht korrekt beschreiben. Durch den Wechsel von der Ortsdarstellung zu einer Basis aus vollständig antisymmetrisierten Slater-Determinanten kann der Konfigurations-Pfadintegral-Monte-Carlo-Ansatz (CPIMC) das Vorzeichenproblem in diesem Parameterbereich erheblich reduzieren und ermöglicht somit die exakte Berechnung thermodynamischer Erwartungswerte.

Im Rahmen dieser Arbeit wurde die Effizienz des CPIMC-Algorithmus enorm verbessert und weitere Observablen zugänglich gemacht. Die erste erfolgreiche Implementation des diagrammatischen Wurm-Algorithmus für allgemeine Hamilton-Operatoren im Fock-Raum mit beliebiger Zweiteilchenwechselwirkung öffnet den direkten Zugang zu der Matsubara-Green-Funktion, welche die statistische Varianz daraus abgeleiteter Schätzer für Erwartungswerte von Observablen wie die Einteilchendichte signifikant reduziert und die Rekonstruktion dynamischer Eigenschaften aus den Ergebnissen von Simulationen im thermodynamischen Gleichgewicht erlaubt. Die durch die neue Implementation erreichte immense Verbesserung des Monte-Carlo-Algorithmus, einschließlich einer viel effizienteren Berechnung der Übergangswahrscheinlichkeiten, sowie die erfolgreiche Parallelisierung auf tausende CPU-Kerne sind essentiell für die anschließende Anwendung der Methode auf im Vergleich zu vorigen Arbeiten wesentlich größere Systeme.

Die vorliegende Arbeit demonstriert die Fähigkeiten des CPIMC-Ansatzes für ein Modellsystems aus Coulomb-wechselwirkenden Fermionen in einer zweidi-

mensionalen harmonischen Falle. Die Korrektheit der CPIMC-Implementierung wird durch ausgiebige Vergleiche mit einer exakten Diagonalisierungsmethode verifiziert. Zusätzlich werden Benchmark-Ergebnisse präsentiert, welche für Systeme mit offenen Schalen große Fehler der Hartree-Fock-Näherung schon bei geringen Wechselwirkungsstärken belegen und eine signifikante Abweichung der Multi-Level-Blocking-PIMC-Methode im Limes einer vollständigen Basis aufzeigen.

Die Anwendung der CPIMC-Methode auf das warme, dichte homogene Elektronengas (HEG) erbringt den Nachweis von systematischen Ungenauigkeiten von vor Kurzem veröffentlichen Restricted-PIMC (RPIMC) Ergebnissen. Diese wurden bereits zur Parametrisierung der freien Austauschkorrelationsenergie, wie sie als Funktional für Berechnungen im Rahmen der Dichtefunktionaltheorie bei endlichen Temperatur benötigt wird, genutzt. Jedoch konnte gezeigt werden, dass die relativen Abweichung der RPIMC-Daten zu den exakten CPIMC-Resultaten bei mittleren Dichten 10 % übersteigen. Zusammen mit den äußerst genauen Ergebnissen für die Austauschkorrelationsenergie bei hohen Dichten, welche unzugänglich für die RPIMC-Methode sind, können die in dieser Arbeit erzeugten Daten helfen, die Genauigkeit von zukünftigen Austauschkorrelationsfunktionalen bei endlichen Temperaturen signifikant zu erhöhen.

Die Benchmark-Ergebnisse dieser Arbeit wurden bereits für die Entwicklung anderer Zugänge, wie z.B. der Dichtematrix-QMC-Methode, genutzt. Insbesondere waren ausführliche Vergleiche entscheidend für die erfolgreiche Verifikation der Permutation-Blocking-PIMC-Methode (PBPIMC), welche den Anwendungsbereich von PIMC-Berechnungen von Fermionen zu deutlich höheren Dichten und Entartungen hin erweitert. Die Kombination der komplementären CPIMC- und PBPIMC-Methoden erlaubt eine fast vollständige Beschreibung des HEGs bei endlichen Temperaturen im thermodynamischen Limes.

Contents

Acronyms

HF Hartree-Fock

DFT Density Functional Theory

FTDFT Finite-Temperature Density Functional Theory

TDDFT Time-Dependent Density Functional Theory

CI Configuration Interaction

MC Monte Carlo

QMC Quantum Monte Carlo

PIMC Path Integral Monte Carlo

RPIMC Restricted Path Integral Monte Carlo

MLB Multi-Level Blocking

PBPIMC Permutation Blocking Path Integral Monte Carlo

CPIMC Configuration Path Integral Monte Carlo

DiagMC Diagrammatic Monte Carlo

DMC Diffusion Monte Carlo

FCIQMC Full Configuration Interaction Quantum Monte Carlo

DMQMC Density Matrix Quantum Monte Carlo

ONV Occupation Number Vector

CONS Complete Orthonormal System

MGF Matsubara Green Function

HEG Homogeneous Electron Gas

NIF National Ignition Facility

ICF Inertial Confinement Fusion

1 | Introduction

The fundamental laws of physics that govern the behavior of ordinary matter at microscopic scales and describe, in principle, nearly all phenomena that are encountered in everyday life are already known for almost a century [1]. The majority of problems in condensed matter physics, plasma physics, and chemistry can be formulated within the framework of non-relativistic quantum mechanics, which is sufficiently accurate for the predominant energy scales found on earth. However, exactly solving the resulting equations is extremely difficult for all but the most trivial cases. In general, exact analytical and numerical approaches are infeasible for more than two or three particles. Therefore, a large part of theoretical physics has been concerned with the development and improvement of analytical and numerical approximations.

Especially important is the quantum mechanical description of electrons. Quantum effects of electrons are dominantly responsible for the properties of matter, while the much heavier nuclei can often be treated classically. Electrons, among protons and many atoms, belong to a class of particles that are called *fermions* and have half-integer spins as opposed to bosons which possess integer spins. Despite all the efforts of several decades of extensive research, a general and accurate method for the simulation of fermions is still far from being available in the near future. The present work is another step in a long series of achievements towards the goal of reliable, large scale *computer experiments* of fermionic quantum particles.

Before the advent of computational physics, only analytical approximations had been available. Among the earliest developments is the Hartree-Fock (HF) method [2, 3], which is still widely used today. It can be considered as the baseline in a hierarchy of more and more accurate, but also more and more demanding approaches. Often, correlations between particles are defined as all effects not included in the HF picture. With this definition, the HF result is considered to be uncorrelated while the exact solution includes the complete correlation. In general, however, even solving the HF equations is computationally too de-

manding to be carried out without the help of computers and even more simple approximations had to be used. Going beyond the level of HF and including correlations almost always requires numerical computations. Therefore, the rapid development of faster computers was and still is of key relevance for the continuous progress in quantum mechanical calculations. Advanced methods for calculations in the ground-state or in thermodynamic equilibrium include perturbation expansions of various orders, e.g., the Møller-Plesset perturbation theory [4], the random phase approximation (RPA) [5], the Montroll-Ward (MW) and e^4 approximation [6, 7], and the Singwi-Tosi-Land-Sjölander (STLS) approximation [8, 9] to mention a few. Dynamic properties can be obtained by the quantum kinetic theory and non-equilibrium Green function approaches [10, 11].

One of the most popular methods for ground-state calculations of electronic structure is the Density Functional Theory (DFT) [12–14]. It reduces the many-body Schrödinger equations to an equation of functionals of only the one-particle density, which in principle is an exact mapping. However, the so-called exchange-correlation functional is generally unknown and needs to be approximated. Accurate approximations of the exchange-correlation functional require reliable data from other methods [15]. The local density approximation (LDA), which is widely used in solid state physics, relies on exact results for the exchange-correlation energy of the Homogeneous Electron Gas (HEG) [16]. The need for independent input data applies to generalizations like Finite-Temperature Density Functional Theory (FTDFT) [3] and time-dependent DFT [17] as well.

Most of the above approximations have in common that they introduce unknown errors with no *systematic* way for improvement. While their results have been shown to be sufficiently accurate for many applications by comparisons with experiments, their predictive power cannot be established *a priori*, i.e., any of the above methods can yield qualitatively wrong results in some situations. An especially successful class of algorithms that can overcome this problem belong to the so called Monte Carlo (MC) methods. These rely on random numbers to yield results that are only correct on average, with statistical deviations that typically decrease proportional to the square root of the CPU time. Avoiding the calculation of exact quantities, which is unnecessary because in all realistic applications one requires only a finite relative accuracy, greatly increases the efficiency. As the deviations can in principle be made arbitrary small by extended calculations, the results can still be considered to be *quasi-exact*. Well-known examples are the Diffusion Monte Carlo (DMC) method for the ground-state [18] and the Path Integral Monte Carlo (PIMC) method for finite-temperatures [19].

While Quantum Monte Carlo (QMC) methods are very efficients for bosons and can be applied to systems consisting of several thousands of particles [20],

fermionic QMC simulations are hampered by the so-call fermion sign problem, which leads to an exponential increase of the statistical error in dependence of the particle number [21]. The sign-problem is a consequence of the required *anti-symmetry* of the wave function, a fundamental statistical property of fermions which also manifests in the Pauli principle, i.e., the observation that no more than one fermion can occupy the same quantum mechanical state at any instance of time. Unfortunately, a general solution of the fermion sign problem seems to be very unlikely [22]. The usual approach is the introduction of approximations like the *fixed-node* approximation [23] which completely avoids the sign problem. However, such approaches reintroduce unknown, systematic errors. Nevertheless, for the ground-state, fixed-node DMC calculations belong to the most accurate methods available [24], although only biased results can be obtained for observables that do not commute with the Hamiltonian, e.g., the kinetic and interaction energies [25]. At finite-temperatures, the Restricted Path Integral Monte Carlo (RPIMC) method based on the same idea is widely used [26–28], but its reliability is still an open question. Other approaches to reduce the sign problem are the Multi-Level Blocking (MLB) MC [29] and the recent very promising Permutation Blocking Path Integral Monte Carlo (PBPIMC) methods [30] employing the idea of so-called sign blocking.

In the diploma thesis preceding this work, the new Configuration Path Integral Monte Carlo (CPIMC) method for the simulation of highly degenerate fermions at finite-temperatures was developed [31]. Its basic idea is to evaluate the Feynman path-integral in a basis of already properly anti-symmetrized Slater-determinants instead of the usual coordinate presentation, thereby avoiding the additional anti-symmetrization of the wave functions, which is the source of the sign problem in the PIMC method [32]. Complementary to the latter, the CPIMC approach has no sign problem at all for ideal fermions at arbitrary degeneracy but an increasing sign problem for large coupling strengths [33]. For small particle numbers, the complementary behavior allows for exact calculations at any coupling strength using always the most efficient method. For larger particle numbers, it constitutes a valuable benchmark for approximative methods that try to close the gap between the applicable range of exact CPIMC and PIMC methods. While being based on the so-called continuous time and diagrammatic MC (DiagMC) methods [34, 35], which are successfully used for simulations of bosons and fermions with simplified interactions [36, 37], the CPIMC algorithm is the first implementation of that approach for Hamiltonians in continuous space with general pair-interactions including the important case of the full, long range Coulomb potential. A similar idea of performing DMC calculations without fixed-nodes in a basis of Slater-determinants is employed by recent Full Configuration

Interaction Quantum Monte Carlo (FCIQMC) simulations of the ground state [38]. Very recently, the approach was generalized to finite temperatures by the Density Matrix Quantum Monte Carlo (DMQMC) method [39].

Among several drawbacks of the original implementation are the low acceptance ratios of the MC updates and the high variance of the estimator for the one-particle density [31]. The goal of this work is to further improve the CPIMC method, to enable the calculation of new quantities, and to apply the method to current physical problems of high interest. The main effort is a new implementation of a CPIMC algorithm based of the so-called *worm algorithm* which has been proven to be very successful in increasing the efficiency of MC calculations of lattice models and in the coordinate representation [20, 40]. The worm algorithm also enables the direct sampling of grand-canonical expectation values and the Matsubara Green function. The latter is a generalization of the one-particle density matrix and even allows for the reconstruction of dynamical properties from equilibrium simulations [41].

To test the new implementation the method will be applied to fermions in a two-dimensional harmonic potential. This model system is fairly simple, yet its Hamiltonian exhibits many properties of a general many-body system of Coulomb interacting particles in an external potential. Beside being a suitable toy problem for the development of new quantum mechanical methods, the model is also used to describe few-electron quantum dots in nano-structures, which exhibit many interesting phenomena like Wigner crystallization, collective modes, and superfluidity [42–46].

An important system of increasing interest is the HEG at finite temperatures. Recent experimental progress in the investigation of plasmas and laser-excited solids in the so-called warm dense matter regime and open question about the structure of astrophysical objects like gas giants and white dwarfs increase the need for improved finite-temperature simulations of large, complex systems [47–53]. However, the physical properties of warm dense matter are determined by a non-trivial interplay of coupling, temperature, and quantum effects, all of which must be accurately included in its reliable theoretical description. A candidate for such simulations is the FTDFT, which depends on exact data for the exchange-correlation free energy of the HEG at finite temperatures [54]. While accurate QMC results for the ground state have been seminal for the success of the DFT, only recently RPIMC results at finite temperatures have been published [55] and used to construct suitable exchange-correlation functionals [56–58]. Nevertheless, no reliable data is available for strongly degenerate electrons at high densities that occur in the warm dense matter regime and the quality of the available RPIMC data is unknown. Both problems will be addressed in this work.

This work consists of three main chapters. After a short general introduction of the quantum statistics of many-body systems and the fermion sign problem of the PIMC method in coordinate representation in chapter 2, chapter 3 will give a complete and detailed description of the CPIMC method and its implementation. The derivation of the formalism and notes on the development process are kept very concise because both are carefully covered by the master thesis of Simon Groth [59] which was co-supervised during this work. Besides contributing to the design and implementation of several MC updates and the application to the harmonic oscillator, the main result of his master thesis are the development of a greatly improved estimator for the Matsubara Green function and an algorithm to construct an optimized HF basis which reduces the sign problem of the CPIMC method. For completeness, a summary of both these results is included in this thesis as well. Despite being rather technical, the MC updates are introduced in great detail as the algorithm is a central achievement of this work and essential for the results in the following chapters. The corresponding sections 3.4 and 3.9 serve as a documentation of the CPIMC program code, which until now is the only implementation of the worm algorithm for a fully general Hamiltonian in Fock space with arbitrary pair-interactions, but can easily be skipped by the reader. Further, the chapter covers general optimizations and the parallelization of the method.

Chapter 4 is dedicated to the application of the CPIMC method to Coulomb-interacting fermions in a harmonic trap. The chapter demonstrates the correctness of the formalism and the implementation by comparing the results for the expectation values of various observables with Configuration Interaction (CI) calculations, which is an exact diagonalization method and can be considered as an exact reference [60]. The potential of CPIMC simulations for benchmarking purposes is demonstrated by comparisons with the HF and MLB calculations.

The results of the CPIMC method for the HEG at finite-temperatures are presented in chapter 5. The chapter also includes a description of the necessary improvements of the CPIMC method that allow its application to hundreds of particles and the simulation of the HEG in the warm dense matter regime. Again, the correctness of the changes to the implementation are thoroughly verified by comparisons with exact CI and PIMC results. The sign problem is investigated in detail. The main result of this chapter is the exchange-correlation energy of 33 polarized electrons at high densities and low but finite temperatures and the direct comparison with available RPIMC data. Existing finite-size corrections are evaluated and an attempt is made to yield reliable results for the exchange-correlation energy of the HEG in the thermodynamic limit. Finally, recent developments in this direction are summarized.

2 Statistical Physics of Quantum Many-Body Systems

This chapter will provide a concise overview of the basic concepts of statistical quantum mechanics in thermodynamic equilibrium and the prevalent approaches to the numerical calculation of the associated physical properties. Beside introducing the PIMC method and the problems that arise by its application to fermionic systems, which inspired the development of the CPIMC method, this chapter will also establish the notation and terminology used throughout this work.

2.1 Quantum Statistical Mechanics

The central quantity for the description of quantum-mechanical systems in thermodynamic equilibrium is the density operator

$$\hat{\rho} = \sum_i p_i \, |\psi_i\rangle \, \langle\psi_i| \, ,$$

which is defined by the probability p_i to find the system in the many-body state $|\psi_i\rangle$ for all members of an ensemble of possible states $\{|\psi_i\rangle\}$. Being proper probabilities, the numbers p_i satisfy $p_i \geq 0$ and $\sum_i p_i = 1$. Thus, the density operator is an hermitian, $\hat{\rho}^\dagger = \hat{\rho}$, and positive semidefinite, $\langle\phi|\hat{\rho}|\phi\rangle \geq 0$ for all states ϕ, operator with trace 1, i.e., $\mathrm{Tr}\,\hat{\rho} = \sum_i \langle\phi_i|\hat{\rho}|\phi_i\rangle = 1$ for an arbitrary basis $\{|\phi_i\rangle\}$ of the system's Hilbert space. The density operator determines all physical properties of the system, in particular the expectation value of an arbitrary observable $\hat{O}$ can be calculated according to

$$\langle\hat{O}\rangle = \mathrm{Tr}\,\hat{O}\hat{\rho}. \tag{2.1}$$

If the complete information about the system is known, the system is described by the *pure* state $|\Psi\rangle$ and the density operator is reduced to the projection operator $\hat{\rho} = |\Psi\rangle\langle\Psi|$ with $\hat{\rho}^2 = \hat{\rho}$. In that case above equation is equivalent to the well known definition $\langle\hat{O}\rangle_\Psi = \langle\Psi|\hat{O}|\Psi\rangle$ and the relation $\mathrm{Tr}\,\hat{\rho}^2 = 1$ holds. If

Tr $\hat{\rho}^2 < 1$, some information is missing and the micro-state of the system is not completely determined. The system is said to be in a *mixed* state and is described statistically by an ensemble. A thermodynamic ensemble is determined by a set of macroscopic observables. The *canonical* ensemble applies to systems that are weakly coupled to and in thermal equilibrium with a macroscopic reservoir. The thermodynamic variables associated with such a system are the total number of particles N, the absolute Temperature T, and the volume V. Except for a weak energy exchange, the system is isolated from the reservoir. The density operator in the canonical ensemble reads

$$\hat{\rho}(N, \beta, V) = \frac{1}{Z(N, \beta, V)} e^{-\beta \hat{H}},$$

with the partition function Z, the inverse Temperature[1] $\beta = 1/T$, and the Hamiltonian of the system $\hat{H}$. The partition function

$$Z = \text{Tr}\, e^{-\beta \hat{H}}$$

is related to the free energy via $F(N, \beta, V) = -T \ln Z(N, \beta, V)$. All thermodynamic properties of the system are determined by the thermodynamic potential F and can therefore be calculated knowing only the partition function, e.g.,

$$U = \langle \hat{H} \rangle = -T^2 \frac{\partial}{\partial T} \frac{F}{T} \bigg|_{N,V} = -\frac{\partial}{\partial \beta} \ln Z \bigg|_{N,V}$$

is the internal energy. Alternatively, the same relation can be obtained from Eq. (2.1) directly.

If the system is also allowed to exchange particles with the reservoir in addition to exchanging energy, it is described by the *grand canonical* ensemble. The total particle number N is now a fluctuating quantity and is replaced by the chemical potential μ as the thermodynamic variable. The density operator becomes

$$\hat{\rho}(\mu, \beta, V) = \frac{1}{Z(\mu, \beta, V)} e^{-\beta(\hat{H} - \mu \hat{N})},$$

where $\hat{N}$ is the total particle number operator, see 3.1, and the grand canonical partition function is given by

$$Z(\mu, \beta, V) = \text{Tr}\, e^{-\beta(\hat{H} - \mu \hat{N})}. \tag{2.2}$$

Here, the trace is over an arbitrary basis of the Fock space, i.e., over states with varying particle number N. The associated thermodynamic potential is the grand potential $\Omega(\mu, \beta, V) = -T \ln Z(\mu, \beta, V)$. All other thermodynamic properties can be obtained from Ω using the corresponding thermodynamic relations.

---

[1]Throughout this work, natural units with $\hbar = k_B = 1$ are used.

2.2 Metropolis Monte Carlo method

The defining characteristic of MC methods is the use of random samples to obtain approximate numerical results. Due to the central limiting theorem, the uncertainty tends to zero in the limit of large numbers and the exact result is recovered. The idea was first applied to neutron scattering problems by Ulam and von Neumann at the Los Alamos National Laboratory [61]. Monte Carlo methods are especially suited for high-dimensional integrals as they appear in the calculation of expectation values in statistical physics and often are the only method available for solving such integrals.

The expectation value $\langle \hat{O} \rangle$ of an observable $\hat{O}$ in thermodynamic equilibrium, see Eq. (2.1), can be written as a high-dimensional integral or sum over the multi variable x describing the micro state of the system,

$$\langle \hat{O} \rangle = \int \mathrm{d}x \, p(x) O(x).$$

If $p(x)$ is a probability distribution, i.e., if it is positive, $p(x) \geq 0$, and normalized, $\int \mathrm{d}x \, p(x) = 1$, the integral can be approximated by the arithmetic mean

$$\langle \hat{O} \rangle \approx \frac{1}{N_{\mathrm{MC}}} \sum_{i=1}^{N_{\mathrm{MC}}} O(x_i),$$

where the N_{MC} micro states x_i are random samples from the distribution $p(x)$. The statistical error deceases as $O(1/\sqrt{N_{\mathrm{MC}}})$ and the approximation becomes exact for $N_{\mathrm{MC}} \to \infty$. Choosing the samples x_i according to the probability density $p(x)$ is called importance sampling opposed to sampling x_i uniformly in the integration region. The advantage of this method is that its efficiency depends only weakly on the number of dimensions.

There exist several methods for drawing random samples from a given distribution $p(x)$, e.g., the inverse sampling method used in Sec. 3.5 or rejection sampling methods. However, for problems of statistical physics the normalization factor of $p(x)$, i.e., the partition function Z, is usually unknown and most of the methods cannot be applied. A solution was found by Metropolis *et. al.* in 1953 for the simulation of a two dimensional gas of rigid spheres [62]. The Metropolis algorithm, which was later generalized by Hastings [63, 64], constructs a Markov chain by generating a random sample x_{i+1} from a previous sample x_i starting from an arbitrary initial configuration x_0. The procedure of choosing x_{i+1} from a given x_i is also called a MC step. The distribution of the configurations of this Markov process approaches the desired distribution $p(x)$ with increasing

number of samples if two conditions are met. First, the transition probabilities $v(x_i \rightarrow x_{i+1})$ from configuration x_i to configuration x_{i+1} must satisfy the detailed balance equation:

$$p(x_i)v(x_i \rightarrow x_{i+1}) = p(x_{i+1})v(x_{i+1} \rightarrow x_i). \tag{2.3}$$

And second, the transition probabilities must allow to reach any configuration x' starting from any other configuration x within a finite number of MC steps. This condition is called ergodicity and is sometimes difficult to satisfy and often impossible to prove formally. Even if the transition probabilities are all finite, transitions between important regions of the configuration space may happen so infrequently that the process is effectively non-ergodic. In the Metropolis-Hastings algorithm, the transition probability is composed of a proposal density $Q(x_i \rightarrow x'_i)$ for proposing a candidate configuration x'_i given the current configuration x_i and the acceptance probability $A(x_i \rightarrow x'_i)$ for accepting the proposed candidate as the new configuration $x_{i+1} = x'_i$. With probability $1 - A$, the proposed candidate is rejected and the current configuration is chosen instead, i.e., $x_{i+1} = x_i$. Obviously, the resulting transition probability $v(x_i \rightarrow x_{i+1}) = Q(x_i \rightarrow x_{i+1})A(x_i \rightarrow x_{i+1})$ satisfies the detailed balance condition when

$$A(x_i \rightarrow x'_i) = \min\left[1, \frac{Q(x'_i \rightarrow x_i)p(x'_i)}{Q(x_i \rightarrow x'_i)p(x_i)}\right]. \tag{2.4}$$

The advantage of the Metropolis algorithm is that only the ratio of the weights $w(x)$ has to be calculated and the the normalization constant Z of $p(x) = w(x)/Z$ is not needed. Only the proposal densities Q must be normalized as one has to be able to draw random samples from that distribution. The proposal density can be chosen freely as long as the resulting MC steps are ergodic. As a rule of thumb, the algorithm is most efficient if the choice of Q leads to an average acceptance ratio of about 50 %.

The proposal and acceptance probabilities and thus the transition probabilities depend only on the current configuration and not on the history of previous samples, i.e., the Metropolis method has the Markov property and the sampling process is indeed a Markov chain. Due to the dependence on the current configuration, the sampled configurations are autocorrelated, which has to be taken into account when estimating the statistical errors of the results. Usually, the autocorrelation is first reduced by keeping only every n-th configuration and ignoring the rest in order to avoid unnecessary time-consuming calculations of $O(x_i)$. The number n is called the cycle. The optimal cycle depends on the autocorrelation time and the relative cost of computing $O(x_i)$ compared to the cost of a MC step.

Another effect of the autocorrelation is the need for an equilibration time or burn-in period after starting from an initial configuration x_0. Because x_0 can be arbitrary and very different from a typical configuration, it can take several ten to a hundred autocorrelation times until the samples x_i are actually distributed according to $p(x)$. It is therefore necessary to discard an appropriate number of configurations before collecting samples for the calculation of averages.

Note that the detailed balance equation (2.3) is only a sufficient and not a necessary condition and often chosen because of its simplicity. In some cases, using other conditions [65], like the maximum global balance condition, can significantly increase the efficiency.

2.3 Path Integral Monte Carlo

In order to apply the Metropolis Monte Carlo method to the simulation of quantum mechanical systems, the thermodynamic expectation values have to be expressed in terms of weights and estimators that can be efficiently calculated. This is not the case for the trace in Eq. (2.1) as the matrix elements of the density operator are generally unknown. The common approach to obtaining a suitable form is the PIMC method based on the Feynman path integral formulation of quantum mechanics in coordinate representation [66]. Because of the product property of the exponential function, the canonical density operator at inverse temperature β can be written as a product of exponentials at an M-times higher temperature:

$$\hat{\rho} = \frac{1}{Z}e^{-\beta\hat{H}} = \frac{1}{Z}\left[e^{-\frac{\beta}{M}\hat{H}}\right]^M.$$

Expressing the density matrix $\rho(R, R') = \langle R|\hat{\rho}|R'\rangle$ in a basis of N-particle product states of eigenfunctions of the coordinate operator, $|R\rangle = |r_1\rangle \cdots |r_N\rangle$, and inserting $M - 1$ identities of the form $\hat{1} = \int dR\, |R\rangle\langle R|$ yields

$$\rho(R, R', \beta) = \frac{1}{Z(\beta)} \int dR_1 \cdots \int dR_{M-1}\, \langle R|e^{-\frac{\beta}{M}\hat{H}}|R_1\rangle \cdots \langle R_{M-1}|e^{-\frac{\beta}{M}\hat{H}}|R'\rangle.$$

The integrals are over all trajectories or *paths* $\{R, R_1, \ldots, R'\}$ between the fixed starting point R and end point R'. In the limit $M \rightarrow \infty$ the path becomes continuous, $R(\tau)$, and the resulting expression is connected to the Feynman path integral by a Wick rotation, giving rise to the notion of *imaginary time* $\tau = it$. The $M + 1$ hyperplanes containing the points R_i are called *time slices* with an imaginary time step of $\epsilon = \beta/M$.
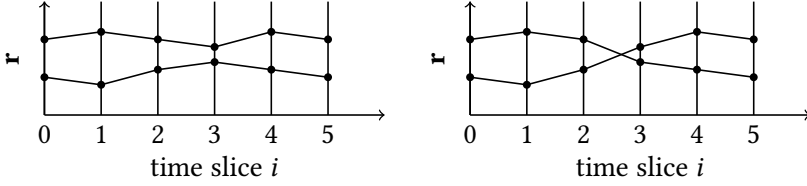
Figure 2.1: Paths in coordinate representation. Shown are two possible paths of two particles with five time slices in one dimension. A path is determined by a coordinate r for each particle and time slice. In the left example, the trajectories of the particles form two separate β-periodic loops, while in the right picture, the path consists of a single loop of length 2β involving both particles due to a pair exchange. For fermions, particle exchange induces a sign change for each transposition.

For indistinguishable particles, the density operator has to be projected on the subspace of states with the correct symmetry, see Sec. 3.1. Bosonic states are completely symmetric under arbitrary particle permutations while fermionic states are completely antisymmetric. In both cases, the projection operator $\hat{P}_\pm = (\mathcal{N}_\pm/N!)\hat{S}_\pm$ satisfies $\hat{P}_\pm^2 = \hat{P}_\pm$ and commutes with the Hamiltonian, $[\hat{P}_\pm, \hat{H}] = 0$. Therefore, only the last state has to be projected:

$$\hat{\rho}_\pm = \hat{P}_\pm^\dagger \hat{\rho} \hat{P}_\pm = \hat{\rho} \hat{P}_\pm.$$

In the following, only the partition function Z is considered as all thermodynamic properties can be calculated from it. With the definition of the (anti-)symmetrization operator, Eq. (3.2), the partition function becomes

$$Z = \frac{1}{N!} \sum_{\sigma \in S_n} (\pm 1)^s \int \mathrm{d}R \int \mathrm{d}R_1 \cdots \int \mathrm{d}R_{M-1} \, \langle R|e^{-\epsilon \hat{H}}|R_1\rangle \cdots \langle R_{M-1}|e^{-\epsilon \hat{H}}|P_\sigma R\rangle,$$

where the sum is over all $N!$ permutations of N particles and the sign of the permutation has to be included in the case of fermions. The contributions to the partition function can be interpreted as collections of closed loops or polymers which represent the trajectories of the particles in imaginary time [67]. Due to the permutation P_σ of the last coordinate, n particles can form a single closed loop of length $n\beta$. An example for two particles and five time slices is shown in Fig. 2.1.

The above expressions are exact for any $M \geq 1$ but cannot be used directly for a MC integration as the matrix elements $\langle R|e^{-\epsilon \hat{H}}|R'\rangle$ are unknown. If M is sufficiently large, the matrix elements can be replaced by a high temperature approximation. The simplest approximation is based on the Trotter decomposition [68],

$$e^{-\beta(\hat{A}+\hat{B})} = \lim_{M \to \infty} \left[e^{-\frac{\beta}{M}\hat{A}} e^{-\frac{\beta}{M}\hat{B}} \right]^M,$$

for any two Hermitian operators $\hat{A}$ and $\hat{B}$. Splitting the Hamiltonian $\hat{H} = \hat{T} + \hat{V}$ into the kinetic and potential energy contributions, which in the coordinate representation are off-diagonal and diagonal, respectively, one obtains

$$\langle R_1|e^{-\epsilon \hat{H}}|R_2\rangle \approx \int dR' \, \langle R_1|e^{-\epsilon \hat{T}}|R'\rangle \langle R'|e^{-\epsilon \hat{V}}|R_2\rangle,$$

with

$$\langle R|e^{-\epsilon \hat{T}}|R'\rangle \approx (4\pi\lambda\epsilon)^{-3N/2} e^{-\frac{(R-R')^2}{4\lambda\epsilon}},$$

$$\langle R|e^{-\epsilon \hat{V}}|R'\rangle = e^{-\epsilon V(R)}\delta(R - R').$$

The approximation for the kinetic matrix elements is only valid if the thermal wavelength $\lambda\epsilon = \epsilon/2m$ is much smaller than the system volume V. Otherwise, a Jacobi theta function has to be used [66]. The resulting expression for the partition function,

$$Z = \frac{1}{N!} \frac{1}{(4\pi\lambda\epsilon)^{3NM/2}} \sum_{\sigma \in S_n} (\pm 1)^s \int dR \int dR_1 \cdots \int dR_{M-1} \, e^{-\sum_{k=1}^{M} \frac{(R_{k-1}-R_k)^2}{4\lambda\epsilon} - \epsilon V(R_k)},$$

converges to the exact value with $O(\beta^2/M)$. Some other approximations of higher order are described in [66]. For example, the PBPIMC method relies on a fourth order approximation developed by Chin *et. al.* [69, 70].

Having expressed the partition function in terms of weights that can be easily calculated for a given path, i.e., a set of coordinates for each particle on each time slice, one can find MC steps that satisfy the detailed balance and ergodicity conditions to efficiently sample the configuration space. In the simplest case, this is a random displacement of a single randomly chosen particle on a random time slice. More advanced methods displace several particles at once and can be found in the literature, e.g, in [66]. The most efficient method, the worm algorithm [20], constructs non-local updates of the configuration space of Z from local MC moves in an enlarged configuration space, thereby avoiding the critical

slowdown in the presence of metastable states. In the case of the PIMC method, it helps to overcome energy barriers between different permutations. An adaption of the worm algorithm to the CPIMC method will be presented in Sec. 3.7.

The PIMC method, which was only briefly introduced here, is the most successful approach for the simulation of bosonic quantum systems at finite temperatures that are subject to non-trivial effects such as superfluidity and Bose-Einstein condensation [66, 71–75]. With advanced sampling techniques it is possible to simulate several thousand particles on modern computers. For fermions, the necessary anti-symmetrization of the density operator introduces an alternating sign. Therefore, the contributions to the partition function cannot be directly used as weights in the Metropolis algorithm. How the method can be adapted to integrals with negative weights will be explained in the next section.

2.4 Fermion Sign Problem

The Metropolis-Hastings algorithm is the best sampling method for the MC integration of high dimensional integrals or sums of the form $\langle \hat{O} \rangle = \int \mathrm{d}x\, O(x)w(x)/Z$ when the normalization Z of the probability $p(x) = w(x)/Z$ is unknown. However, for some systems the weights $w(x)$ are not always positive and $p(x)$ can thus not be interpreted as a probability. This problem arises, e.g., for quantum spin systems with frustrated interaction or, as shown above, for path integral simulations of fermions [22]. In the latter case, the sign of the weights changes due to particle permutations that are necessary for the anti-symmetrization of the product states.

The common solution is to incorporate the sign of the weight $s(x) = \mathrm{sign}\big(w(x)\big)$ into the estimator for the observable and to use the modulus of the weights in the acceptance probabilities of the Metropolis algorithm. The expectation value of the observable then becomes

$$\langle \hat{O} \rangle = \frac{\int \mathrm{d}x\, O(x)s(x)p'(x)}{\int \mathrm{d}x\, s(x)p'(x)} = \frac{\langle \hat{O}s \rangle'}{\langle s \rangle'},$$

where $\langle \cdot \rangle'$ denotes the expectation value in the system where the configurations x are distributed according to the probability $p'(x) = |p(x)|/Z'$ with $Z' = \int \mathrm{d}x\, |p(x)|$. In the case of the PIMC method for fermions, this system is equal to the corresponding bosonic system. The average sign $\langle s \rangle'$ is defined by

$$\langle s \rangle' = \frac{\int \mathrm{d}x\, s(x)|p(x)|}{\int \mathrm{d}x\, |p(x)|} = \frac{Z}{Z'} = e^{-\beta N(f - f')} \tag{2.5}$$

and decreases exponentially with the particle number N, the inverse temperature β, and the difference between the free energy per particle of the two systems, $f - f'$. As a consequence for MC simulations, the estimate of the relative statistical error of the mean increases for small average signs $\langle s \rangle' \ll 1$ like

$$\frac{\Delta O}{\bar{O}} \sim \frac{\Delta s}{\bar{s}} \approx \frac{1}{\sqrt{N_{\mathrm{MC}}}} e^{\beta N \Delta f},$$

which leads to an exponential increase in the number of MC samples M that are necessary for a desired accuracy. This so-called fermion sign problem applies not only to PIMC simulations but to fermionic QMC in general, restricting *ab initio* calculations to small system sizes and/or high temperatures.

The common approach to avoid the sign problem is the RPIMC method or *fixed-node* approximation [66]. The approximation limits the integration to a region where the density matrix is strictly positive. In other words, the particle trajectories are not allowed to cross the nodal surface $\rho(\boldsymbol{R}, \boldsymbol{R}') = 0$. As the exact density matrix is unknown, a trial density matrix is used to determine the nodes, i.e., the points where the density matrix changes its sign. The use of a trial density matrix introduces an uncontrollable systematic error. Additionally, even when the exact nodal surface is known, e.g., for an ideal system, the method does not reproduce the correct fermionic density matrix [76].

The PBPIMC method takes a different approach and analytically sums contributions of opposite sign by writing the sum over all particle permutations as a determinant, which can be calculated efficiently using standard software libraries [30, 77]. Sampling meta-configurations that analytically combine configurations of opposite sign is called *sign blocking*. To increase the effect of blocking, a high order approximation for the density matrix is used, which significantly reduces the necessary number of time slices. Furthermore, the sampling efficiency is increased by a variation of the worm algorithm. The combination of these ideas extends the range of applicability of the PIMC method significantly [78]. A stochastic blocking mechanism is employed by the MLB method [79], but is not widely used due to its complexity.

Since the fermion sign problem is nondeterministic polynomial (NP) hard, finding a general exact solution is very unlikely [22]. Nevertheless, the average sign is not a physical observable and depends on the representation. It can therefore be increased by finding a suitable equivalent reformulation of the expressions for the quantities of interest. The basic idea of the CPIMC method is to switch from continuous product states to a discrete basis of already properly anti-symmetrized Slater-determinants, thereby avoiding the subsequent projection

on the fermionic Hilbert space, which is the origin of the sign problem in the path integral formulation in coordinate representation.

A similar approach is taken by the FCIQMC method [38] and its finite-temperature generalization DMQMC [39, 80]. The FCIQMC method implements the zero-temperature DMC method in a basis of Slater-determinants instead of coordinate product states in a similar way as the CPIMC method applies a basis change to the path integral formulation. The discrete basis allows the introduction of an annihilation step that removes walkers of opposite sign on the same Slater-determinant, thereby drastically reducing the statistical variance associated with the sign problem. By replacing the sampling of the wave function in the Hilbert space of Slater-determinants $|\{n\}\rangle$ with a sampling of the density operator in the operator space of tensor products of the form $|\{n\}\rangle\langle\{n\}'|$, the DMQMC method is able to obtain expectation values at finite temperatures instead of ground state properties. Having a similar range of applicability, DMQMC simulations that make use of the initiator approximation can reach stronger interactions than the CPIMC method, while the latter is more efficient at larger basis sizes necessary for high temperatures and large particle numbers. However, due to the initiator approximation, the DMQMC methods has difficulties to access observables that do not commute with the Hamiltonian.

In some special cases, the sign problem can even be solved completely by choosing an appropriate representation, e.g., for one dimensional systems [81, 82]. Furthermore, if the Hamiltonian can be diagonalized, the trace in its eigenbasis is just a sum over positive exponential functions of the eigenvalues. In general, the cost for diagonalizing the Hamiltonian grows exponentially with system size and can only be applied to very small particle numbers. Exceptions include non-interacting fermions in the canonical ensemble, where the eigenvalues are known but expectation values can still not be calculated analytically for large finite systems.

Finally, it should be mentioned that for the simulation of the real-time evolution of quantum systems a related dynamical sign problem occurs, restricting simulation to short time dynamics [83]. A recent paper [84] claims to overcome the dynamical sign problem and demonstrates a successful application of the so called *inchworm algorithm* to the Anderson impurity model.

3 | Configuration Path Integral Monte Carlo

The main effort of this work was the improvement and extension of the CPIMC method, which I developed in my diploma thesis [31, 33], towards much larger and more realistic systems such as the electron component in current warm dense matter applications. New physical quantities, in particular the Matsubara Green function, became accessible due to the implementation of the worm algorithm. A detailed description of the development process can be found in the master thesis of Simon Groth [59] that I co-supervised during this doctoral research. Since the completion of that thesis, the Monte Carlo steps for general Hamiltonians with arbitrary pair-interactions have undergone only minor changes. More recent advances include efficiency improvements like the heat bath method and massive parallelization.

In this chapter, a detailed and updated description of the current state of the general CPIMC method will be given. For the derivations of the formulas and any reasoning about development choices, reference is made to the previous works. The adaption of the algorithm to the important special case of the HEG is part of chapter 5.

3.1 Second Quantization

The basic idea of the CPIMC approach is to express all quantities, wave functions and observables alike, in terms of operators. The commutation relations of these operators completely define the quantum statistical properties of the particles, i.e., of fermions and bosons. Thus, the necessary antisymmetrization of the density operator in coordinate space, which is the source of the sign problem, can be omitted. This formulation is called second quantization and is fundamental to the understanding of the remaining chapter. Therefore, a brief introduction will be given in this section, which will also serve as a reference to the notation and definitions used throughout this work. Details can be found in [31, 59] and any textbook on the topic, e.g., [60, 85, 86].

3.1.1 Indistinguishable Particles

A solution of the stationary many-body Schrödinger equation of N particles

$$\hat{H} |\Psi\rangle = E |\Psi\rangle ,$$

with the Hamilton operator $\hat{H}$, is an element of the N-particle Hilbert space $\mathcal{H}^N = \bigotimes_{\alpha=1}^{N} \mathcal{H}_\alpha$, which is the tensor product of single-particle Hilbert spaces $\mathcal{H}_\alpha = \mathcal{H}_{\text{space},\alpha} \otimes \mathcal{H}_{\text{spin},\alpha}$. The Hilbert space of each particle α is itself the tensor product of the Hilbert spaces describing spatial and spin degrees of freedom, respectively. A basis of $\mathcal{H}^N$ can be constructed from product states

$$|i_1 i_2 \ldots i_N\rangle = |i_1\rangle_1 |i_2\rangle_2 \cdots |i_N\rangle_N . \tag{3.1}$$

Here, the single-particle states $|i\rangle_\alpha \in \mathcal{H}_\alpha$ are spin orbitals, i.e., the wave function $\langle r\sigma|i\rangle = (\langle r|\langle\sigma|)|i\rangle = \phi_i(r,\sigma)$ depends on the spatial coordinate r and spin coordinate σ. In the following, both coordinates will be combined in the notation $x = \{r, \sigma\}$.

As quantum particles of the same species are indistinguishable, all observables must remain invariant under exchange of two particles. Thus, their state vector can only change up to a phase factor. The spin-statistic theorem connects the spin of the particles with the value of this phase factor. Particles with integer spin are symmetric under particle exchange and are called bosons. Fermions have a half-integer spin and an anti-symmetric state vector.

States with the proper symmetry can be constructed from product states, Eq. (3.1), by applying the (anti-)symmetrization operator $\hat{S}_\pm$

$$
\begin{aligned}
|i_1 i_2 \ldots i_N\rangle_\pm &= \hat{S}_\pm |i_1 i_2 \ldots i_N\rangle \\
&= \frac{1}{\mathcal{N}_\pm} \sum_{\sigma \in S_N} (\pm 1)^s |i_{\sigma(1)}\rangle_1 |i_{\sigma(2)}\rangle_2 \cdots |i_{\sigma(N)}\rangle_N .
\end{aligned} \tag{3.2}
$$

The sum is over all $N!$ elements of the symmetric group S_N and the number s counts the number of transpositions, i.e., the number of exchanges of two particles, that create the permutation σ. For fermions, the sign of each term is determined by the parity of the permutation with a minus sign for odd permutations. The normalization factor is given by

$$\mathcal{N}_\pm = \begin{cases} \sqrt{N! \prod_{i=0}^{\infty} n_i!}, & \text{for bosons} \\ \sqrt{N!}, & \text{for fermions} \end{cases} ,$$

where n_i is the number of times the single-particle state $|i\rangle$ appears in the product state. For Bosons, the sum is often taken only over all different states. In that case, the normalization factor becomes $\mathcal{N}_+ = \sqrt{N!/\prod_{i=0}^{\infty} n_i!}$.

The states (3.2) span the two orthogonal subspaces of the symmetric and anti-symmetric states $\mathcal{H}_\pm^N \subset \mathcal{H}^N$. The projections onto these subspaces are given by the operators $\hat{P}_\pm = (\mathcal{N}_\pm/N!)\hat{S}_\pm$. It is $\hat{P}_\pm^2 = \hat{P}_\pm$ and $\hat{P}_+\hat{P}_- = 0$. The anti-symmetrized states $|\Psi\rangle \in \mathcal{H}_-^N$ can be written as so-called Slater determinants

$$|i_1 i_2 \ldots i_N\rangle_- = \frac{1}{\sqrt{N!}} \begin{vmatrix} |i_1\rangle_1 & |i_2\rangle_1 & \cdots & |i_N\rangle_1 \\ |i_1\rangle_2 & |i_2\rangle_2 & \cdots & |i_N\rangle_2 \\ \vdots & \vdots & & \vdots \\ |i_1\rangle_N & |i_2\rangle_N & \cdots & |i_N\rangle_N \end{vmatrix}.$$

In coordinate representation, this translates to

$$\begin{aligned} \Psi(x_1, x_2, \ldots, x_N)_- &= \langle x_1 x_2 \ldots x_N | i_1 i_2 \ldots i_N \rangle_- \\ &= \frac{1}{\sqrt{N!}} \begin{vmatrix} \phi_{i_1}(x_1) & \phi_{i_2}(x_1) & \cdots & \phi_{i_N}(x_1) \\ \phi_{i_1}(x_2) & \phi_{i_2}(x_2) & \cdots & \phi_{i_N}(x_2) \\ \vdots & \vdots & & \vdots \\ \phi_{i_1}(x_N) & \phi_{i_2}(x_N) & \cdots & \phi_{i_N}(x_N) \end{vmatrix}. \end{aligned}$$

Obviously, there can be no states with two particles occupying the same single-particle state. Therefore, two indistinguishable fermions must differ in at least one quantum number, which is known as the Pauli exclusion principle. For bosons, there is no such principle. Their symmetric states can be represented by a permanent.

3.1.2 Occupation Number Representation

By restricting the (anti-)symmetric states $|i_1 i_2 \ldots i_N\rangle_\pm$ to those following an arbitrary but fixed ordering of the single-particle states $|i\rangle$, e.g., $i_k < i_{k+1} \forall k$, one obtains a basis of the corresponding subspaces $\mathcal{H}_\pm^N$. Because the (anti-)symmetrization procedure (3.2) wipes out any details about which particle occupies which single-particle state, the whole information about the state is given by specifying the occupied states and how many particles they contain. Therefore, a basis state of the (anti-)symmetric N-particle Hilbert space $\mathcal{H}_\pm^N$ can be written as an Occupation Number Vector (ONV)

$$|n_0 n_1 n_2 \ldots\rangle =: |\{n\}\rangle, \tag{3.3}$$

where the occupation number

$$n_i \in \begin{cases} \mathbb{N}_0, & \text{for bosons} \\ \{0, 1\}, & \text{for fermions} \end{cases}$$

specifies the number of particles occupying the single-particle state $|i\rangle$. Note that the particle number $\sum_{i=0}^{\infty} n_i = N$ is fixed. For fermions, the restriction on $n_i = 0$ and $n_i = 1$ automatically satisfies the Pauli exclusion principle and each ONV can be identified with a Slater determinant.

If the states $|i\rangle$ form a Complete Orthonormal System (CONS) of the single-particle Hilbert space $\mathcal{H}$, the states $|\{n\}\rangle^N$ form a CONS of the (anti-)symmetric N-particle Hilbert space $\mathcal{H}_{\pm}^N$ with the orthogonality relations

$$\langle \{n\}|\{\bar{n}\}\rangle = \prod_{i=0}^{\infty} \delta_{n_i, \bar{n}_i} =: \delta_{\{n\}, \{\bar{n}\}}, \tag{3.4}$$

and the completeness relation

$$\sum_{\{n\}} |\{n\}\rangle \langle \{n\}| \, \delta_{\sum_i n_i, N} = \hat{1}_N. \tag{3.5}$$

Here, the generalized Kronecker delta $\delta_{\{n\}, \{\bar{n}\}}$ has been defined and the notation

$$\sum_{\{n\}} := \begin{cases} \sum_{n_0=0}^{\infty} \sum_{n_1=0}^{\infty} \cdots, & \text{for bosons} \\ \sum_{n_0=0}^{1} \sum_{n_1=0}^{1} \cdots, & \text{for fermions} \end{cases}$$

has been introduced. In the following, the Kronecker delta $\delta_{\sum_i n_i, N}$ will often be omitted, when the restriction on a fixed particle number is clear from the context.

If the restriction to a fixed particle number N is dropped, the ONVs $|\{n\}\rangle$ form a basis of the so-called Fock space $\mathcal{F}_{\pm} = \mathcal{H}^0 \oplus \mathcal{H}^1 \oplus \mathcal{H}_{\pm}^2 \cdots$. According to (3.4), the vacuum state $|00\ldots\rangle = |0\rangle$ without any particles is also normalized to unity $\langle 0|0\rangle = 1$. In general, states $|\Psi\rangle \in \mathcal{F}_{\pm}$ do not have a defined particle number.

3.1.3 Creation and Annihilation Operators

The definition of the Fock space allows one to introduce operators that increase or decrease the number of particles of a state. A special case are the pairwise adjoint, so-called creation and annihilation operators. The fermionic creation operator $\hat{a}_i^{\dagger}$ creates a particle in the single-particle state $|i\rangle$. Its action on a ONV is given by

$$\hat{a}_i^{\dagger} |\{n\}\rangle = (1 - n_i)(-1)^{\alpha_{\{n\}, i}} |\ldots, n_i + 1, \ldots\rangle,$$

where the sign is determined by

$$\alpha_{\{n\},i} = \sum_{l=0}^{i-1} n_l,$$

i.e, the number of occupied states before the state $|i\rangle$ in the chosen order of single-particle states. Applying the annihilation operator yields

$$\hat{a}_i \, |\{n\}\rangle = n_i (-1)^{\alpha_{\{n\},i}} \, |\ldots, n_i - 1, \ldots\rangle.$$

Obviously, it is not possible to create a particle in an already occupied single-particle state. Likewise, the annihilator $\hat{a}_i$ vanishes in case of an empty state $|i\rangle$.

The creation and annihilation operators can also be defined by their commutation relations

$$\{\hat{a}_i^\dagger, \hat{a}_j^\dagger\} = \{\hat{a}_i, \hat{a}_j\} = 0,$$
$$\{\hat{a}_i, \hat{a}_j^\dagger\} = \delta_{i,j},$$

where the braces denote the anti-commutator $\{\hat{A}, \hat{B}\} = \hat{A}\hat{B} + \hat{B}\hat{A}$. The quantum statistical properties of fermions are completely determined by these relations, e.g., the Pauli exclusion principle follows directly from the first equation, i.e., $(\hat{a}_i^\dagger)^2 = 0$. The ONVs (3.3) can be constructed by repeatedly applying the creation operators on the vacuum state:

$$|\{n\}\rangle = \left(\prod_{i=0}^{\infty} (\hat{a}_i^\dagger)^{n_i}\right) |\{0\}\rangle.$$

The order of the operators is given by the chosen order of the single-particle state $|i\rangle$.

As will be shown in the next section, not only the states of the Fock space can be expressed in terms of the creation and annihilation operators but also the operators. A special operator is the occupation number operator

$$\hat{n}_i = \hat{a}_i^\dagger \hat{a}_i,$$

which acts on a ONV as follows:

$$\hat{n}_i \, |\{n\}\rangle = n_i \, |\{n\}\rangle.$$

A related operator is the particle number operator

$$\hat{N} = \sum_{i=0}^{\infty} \hat{n}_i. \tag{3.6}$$

Its action is given by

$$\hat{N} \ket{\{n\}} = N \ket{\{n\}}.$$

Thus, the ONVs are common eigenvectors of these operators, and their eigenvalues are the number of particles occupying a single-particle state and the total number of particles, respectively. If the underlying single-particle basis is the eigenbasis of the single-particle Hamiltonian, $\hat{h}\ket{i} = e_i \ket{i}$, then the ONVs are also eigenvectors of the interaction-free many-body Hamiltonian $\hat{H} = \sum_{\alpha=1}^{N} \hat{h}_\alpha$. A transformation from the single-particle basis $\ket{i}$ to a single-particle basis $\ket{v}$ corresponds to the following transformation of the creation and annihilation operators:

$$\hat{a}_v^\dagger = \sum_{i=0}^{\infty} \braket{i|v} \hat{a}_i^\dagger,$$

$$\hat{a}_v = \sum_{i=0}^{\infty} \braket{v|i} \hat{a}_i.$$

This is also valid in case v is a continuous variable, e.g., for the so-called field operators $\hat{\Psi}^\dagger(x)$ and $\hat{\Psi}(x)$.

Replacing the anti-commutator with the commutator $[\hat{A}, \hat{B}] = \hat{A}\hat{B} - \hat{B}\hat{A}$, one obtains the commutation relations of the bosonic creation and annihilation operators

$$[\hat{a}_i^\dagger, \hat{a}_j^\dagger] = [\hat{a}_i, \hat{a}_j] = 0,$$

$$[\hat{a}_i, \hat{a}_j^\dagger] = \delta_{i,j}.$$

Their action on a ONV is given by

$$\hat{a}_i^\dagger \ket{n_0 n_1 \ldots n_i \ldots} = \sqrt{n_i + 1} \ket{n_0 n_1 \ldots n_i + 1 \ldots},$$

$$\hat{a}_i \ket{n_0 n_1 \ldots n_i \ldots} = \sqrt{n_i} \ket{n_0 n_1 \ldots n_i - 1 \ldots}.$$

When constructing the ONVs, one has to account for the multiple occupation of orbitals,

$$\ket{\{n\}} = \frac{1}{\sqrt{\prod_i n_i!}} \left(\prod_{i=0}^{\infty} (\hat{a}_i^\dagger)^{n_i} \right) \ket{\{0\}}.$$

In case of single occupancy, this reduces to the same formula as for fermions. All other expressions of this section remain unchanged.

3.1.4 Operators in Second Quantization

In order to calculate their matrix elements, operators in second quantization are expressed in terms of creation and annihilation operators. For the important cases of one- and two-particle operators, using the fermionic anti-commutation relations, one obtains the Slater-Condon rules [60], which will be given below.

A one-particle operator in first quantization is of the form $\hat{B}_1 = \sum_\alpha \hat{b}_\alpha$, where $\hat{b}_\alpha$ is an operator acting on the one-particle Hilbert space $\mathcal{H}_\alpha$. In second quantization, $\hat{B}_1$ can be expressed as

$$\hat{B}_1 = \sum_{i,j=0}^{\infty} b_{ij} \hat{a}_i^\dagger \hat{a}_j,$$

with the one-particle integrals

$$b_{ij} = \langle i|\hat{b}|j\rangle = \int \mathrm{d}x \, \phi_i^*(x) b(x) \phi_j(x).$$

Note that the integration over $x = \{r, \sigma\}$ combines an integration over the spatial coordinates r and a summation over the spin coordinates σ. The matrix elements of $\hat{B}_1$ can now be calculated from the matrix elements of the creation and annihilation operators. Let $|\{n\}_q^p\rangle$ be the ONV one obtains from the state $|\{n\}\rangle$ by removing a particle from the orbital $|q\rangle$ and adding a particle to the orbital $|p\rangle$, e.g.,

$$|\{n\}_q^p\rangle = |\ldots, n_q - 1, \ldots, n_p + 1, \ldots\rangle.$$

The Slater-Condon rules for one-particle operators are then given by:

$$\langle\{n\}|\hat{B}_1|\{\bar{n}\}\rangle = \begin{cases} \sum_{k=0}^{\infty} b_{kk} n_k, & \{n\} = \{\bar{n}\} \\ b_{pq}(-1)^{\alpha_{\{n\},p,q}}, & \{n\} = \{\bar{n}\}_q^p \\ 0, & \text{otherwise} \end{cases}.$$

The matrix elements are non-zero only if the bra and ket vectors differ by exactly zero or two occupation numbers. For the latter case, the phase factor is determined by the number of occupied orbitals between the orbitals $|p\rangle$ and $|q\rangle$,

$$\alpha_{\{n\},p,q} = \sum_{l=\min(p,q)+1}^{\max(p,q)-1} n_l.$$

A two-particle operator can be written as $\hat{B}_2 = \frac{1}{2} \sum_{\alpha \neq \beta=1}^{N} \hat{b}_{\alpha,\beta}$ in first quantization, where $\hat{b}_{\alpha,\beta}$ acts on the two-particle Hilbert space $\mathcal{H}_\alpha \otimes \mathcal{H}_\beta$. Its representation in second quantization is given by:

$$\hat{B}_2 = \frac{1}{2} \sum_{i,j,k,l=0}^{\infty} b_{ijkl} \hat{a}_i^\dagger \hat{a}_j^\dagger \hat{a}_l \hat{a}_k.$$

Due to the fermionic anti-commutation relations, the order of the indices is important. The two-particle integrals are defined as:

$$b_{ijkl} = \langle ij|\hat{b}|kl \rangle = \int dx \int dy \, \phi_i^*(x)\phi_j^*(y)b(x,y)\phi_k(x)\phi_l(y).$$

For pair-interactions $\hat{W}$, the operator $b(x,y) = w(x,y)$ is symmetric and real, i.e., $w(x,y) = w(y,x)$ and $w^*(x,y) = w(x,y)$. Using these symmetries and defining $w_{ijkl}^- = w_{ijkl} - w_{ijlk}$, the pair-interaction operator can be rewritten:

$$\hat{W} = \sum_{i=0}^{\infty} \sum_{j=i+1}^{\infty} \sum_{k=0}^{\infty} \sum_{l=k+1}^{\infty} w_{ijkl}^- \hat{a}_i^\dagger \hat{a}_j^\dagger \hat{a}_l \hat{a}_k.$$

Calculating the matrix elements of the creation and annihilation operators, one obtains the Slater-Condon rules for the matrix elements of the pair-interaction operator:

$$\langle \{n\}|\hat{W}|\{\bar{n}\}\rangle = \begin{cases} \sum_{i=0}^{\infty} \sum_{j=i+1}^{\infty} w_{ijij}^- n_i n_j, & \{n\} = \{\bar{n}\} \\ \sum_{\substack{i=0 \\ i \neq p,q}}^{\infty} w_{ipiq}^- n_i (-1)^{\alpha_{\{n\},p,q}}, & \{n\} = \{\bar{n}\}_q^p \\ w_{pqrs}^-(-1)^{\alpha_{\{n\},p,q}+\alpha_{\{\bar{n}\},r,s}}, & \{n\} = \{\bar{n}\}_{r<s}^{p<q} \\ 0, & \text{otherwise} \end{cases},$$

where in the state $|\{n\}_{r<s}^{p<q}\rangle$ two particles are excited from the orbitals $|r\rangle$ and $|s\rangle$ with $r < s$ to the orbitals $|p\rangle$ and $|q\rangle$ with $p < q$. There are only three distinct cases where the matrix elements are non-zero: The diagonal elements, i.e., both states are the same, and if they differ in exactly two or four orbitals. In the case of two differing orbitals, the phase factor is the same as for the one-particle operator. In the latter case, the phase factor is given by the sum of occupied orbitals between p and q in the left state $\langle\{n\}|$ and between r and s in the right state $|\{\bar{n}\}\rangle$.

3.2 Derivation of the Partition Function

In [31], the CPIMC representation of the canonical partition function was originally derived analogously to the well-known PIMC method in coordinate representation using the Trotter formula, see Sec. 2.3. It was subsequently found, that equivalent expressions were already obtained before in the interaction picture [31, 59, 87]. This more elegant derivation will be presented here.

In a given basis $|\{n\}\rangle$, any time-independent Hamiltonian $\hat{H}$ can be split into a diagonal and off-diagonal part,

$$\hat{H} = \hat{D} + \hat{Y},$$

with

$$\langle\{n\}|\hat{H}|\{\bar{n}\}\rangle = \begin{cases} \langle\{n\}|\hat{D}|\{\bar{n}\}\rangle, & \{n\} = \{\bar{n}\} \\ \langle\{n\}|\hat{Y}|\{\bar{n}\}\rangle, & \{n\} \neq \{\bar{n}\} \end{cases}.$$

When switching to the interaction picture, the off-diagonal part becomes time-dependent,

$$\hat{H}(t) = \hat{D} + \hat{Y}(t),$$

with the time-dependence being governed by the diagonal part,

$$\hat{Y}(t) = e^{it\hat{D}}\hat{Y}e^{-it\hat{D}}.$$

For the time evolution operator of the Schrödinger picture, the following identity holds:

$$\hat{U}(t, t_0) = e^{-i\hat{D}(t-t_0)}\hat{T}e^{-i\int_{t_0}^{t} dt'\hat{Y}(t')},$$

where $\hat{T}$ denotes the time ordering operator. Its action on the exponential function can be explicitly written as a Dyson series:

$$\hat{T}e^{-i\int_{t_0}^{t} dt'\hat{Y}(t')} = 1 - i\int_{t_0}^{t} dt_1 \hat{Y}(t_1) + (-i)^2 \int_{t_0}^{t} dt_1 \int_{t_0}^{t_1} dt_2 \hat{Y}(t_1)\hat{Y}(t_2)$$

$$+ \cdots + (-i)^k \int_{t_0}^{t} dt_1 \cdots \int_{t_0}^{t_{k-1}} dt_k \hat{Y}(t_1) \cdots \hat{Y}(t_k) + \cdots$$

$$= \sum_{K=0}^{\infty} (-i)^K \int_{t_0}^{t} dt_1 \int_{t_1}^{t} dt_2 \cdots \int_{t_{K-1}}^{t} dt_K \hat{Y}(t_K) \cdots \hat{Y}(t_2)\hat{Y}(t_1).$$

The density operator of the canonical ensemble is connected by a Wick rotation to the time evolution operator in imaginary time $\tau = it$:

$$\hat{\rho}(\beta) = \frac{1}{Z}\hat{U}(-i\beta, 0).$$

Inserting $K - 1$ identity operators of the N-particle anti-symmetric Hilbert space (3.5) in every term $K \geq 2$, using that $\hat{D}|\{n\}\rangle = \hat{D}_{\{n\}}|\{n\}\rangle$ is diagonal, and combining all exponential functions, one obtains for the canonical partition function

$$Z = \text{Tr}\,\hat{U}(-i\beta, 0)$$

$$= \sum_{K=0}^{\infty} \sum_{\{n\}} \sum_{\{n^{(1)}\}} \cdots \sum_{\{n^{(K-1)}\}} \int_0^\beta d\tau_1 \int_{\tau_1}^\beta d\tau_2 \cdots \int_{\tau_{K-1}}^\beta d\tau_K \qquad (3.7)$$

$$\cdot (-1)^K e^{-\sum_{i=0}^K D_{\{n^{(i)}\}}(\tau_{i+1}-\tau_i)} \prod_{i=1}^K Y_{\{n^{(i)}\},\{n^{(i-1)}\}},$$

with $\tau_0 = 0$, $\tau_{K+1} = \beta$, and $\{n^{(0)}\} = \{n^{(K)}\} = \{n\}$. The sums are over ONVs with fixed particle number N. The $K = 1$ term vanishes because $\hat{Y}$ is off-diagonal, i.e., $Y_{\{n\},\{\bar{n}\}} = \langle\{n\}|\hat{Y}|\{\bar{n}\}\rangle = 0$ for $\{n\} = \{\bar{n}\}$. In the following, a compact notation for the multiple integrations over the imaginary time will be used:

$$\int' d^K\tau := \int_0^\beta d\tau_1 \int_{\tau_1}^\beta d\tau_2 \cdots \int_{\tau_{K-1}}^\beta d\tau_K.$$

The sum (3.7) can be interpreted as a sum over all possible β-periodic imaginary-time paths of ONVs $|\{n\}\rangle(\tau)$ in Fock-space. An example path is visualized in Fig. 3.1. Compared to the analogous path integral picture in coordinate space, there are instantaneous changes or *kinks* of the state $|\{n\}\rangle(\tau)$ at distinct positions in imaginary-time τ_i while between two kinks, the state is constant. A path $C = (K, \{n\}, \{n^{(1)}\}, \ldots, \{n^{(K-1)}\}, \tau_1, \ldots, \tau_K)$ is uniquely determined by the number of kinks K, their imaginary times τ_i and the occupation numbers between the kinks $\{n^{(i)}\}$. The weight of a path is given by

$$W(C) = (-1)^K e^{-\sum_{i=0}^K D_{\{n^{(i)}\}}(\tau_{i+1}-\tau_i)} \prod_{i=1}^K Y_{\{n^{(i)}\},\{n^{(i-1)}\}}.$$

The diagonal part of the Hamiltonian $D_{\{n\}}$ determines the weight of the constant parts of a path while each kink corresponds to an off-diagonal matrix element $Y_{\{n\},\{\bar{n}\}}$.
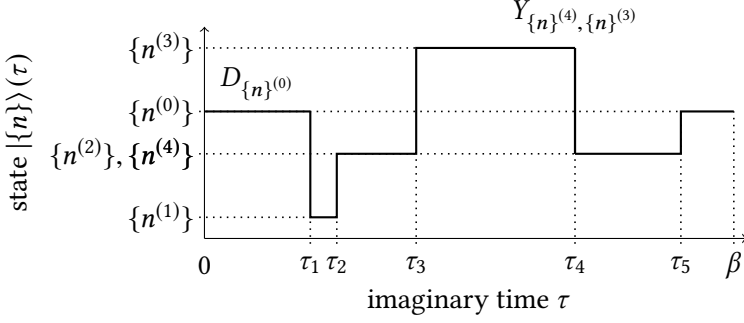
Figure 3.1: Visualization of a path. An N-particle configuration is given by an "imaginary time path" of ONV states $|\{n\}\rangle$ in the interval $[0, \beta]$. The y-axis represents an arbitrary ordering of states. Occupation changes occur at imaginary times $\tau_1, \ldots, \tau_5$, i.e., there are $K = 5$ kinks. From [32].

From the partition function many thermodynamic properties of interest like the total energy can easily be derived. The resulting expressions are suited for the integration by the Metropolis Monte Carlo method and are the starting point for the development of a working implementation in [31]. In the following, a reformulation that allows for more efficient Monte Carlo updates and is better suited for the adaption of the worm algorithm will be presented.

For a Hamiltonian in second quantization

$$\hat{H} = \sum_{i,j} h_{ij} \hat{a}_i^\dagger \hat{a}_j + \sum_{i<j, k<l} w_{ijkl}^- \hat{a}_i^\dagger \hat{a}_j^\dagger \hat{a}_l \hat{a}_k,$$

where $\hat{h}$ is the one-particle part and $\hat{w}$ the pair-interaction, the off-diagonal part follows directly from the Slater-Condon rules:

$$Y_{\{n\}, \{\bar{n}\}} = \sum_{i,j} q_{\{n\}, \{\bar{n}\}}(i, j) + \sum_{i<j, k<l} q_{\{n\}, \{\bar{n}\}}(i, j, k, l),$$

with the definitions

$$q_{\{n\}, \{\bar{n}\}}(p, q) = \left(h_{pq} + \sum_{\substack{i=0 \\ i \neq p, q}} w_{ipiq}^- n_i \right) (-1)^{\alpha_{\{n\}, p, q}} \delta_{\{n\}, \{\bar{n}\}_q^p},$$

$$q_{\{n\}, \{\bar{n}\}}(p, q, r, s) = w_{pqrs}^- (-1)^{\alpha_{\{n\}, p, q} + \alpha_{\{\bar{n}\}, r, s}} \delta_{\{n\}, \{\bar{n}\}_{r<s}^{p<q}}. \tag{3.8}$$

The Kronecker-deltas are meant to be zero if the indices p, q, r, s are not pairwise different and if $p > q$ or $r > s$. The two sums can be combined into one by

introducing the multi-variable s:

$$Y_{\{n\},\{\bar{n}\}} = \sum_{s \in \mathbb{N}^2 \cup \mathbb{N}^4} q_{\{n\},\{\bar{n}\}}(s).$$

Hereby, the matrix elements of the so-called kink operators $\hat{q}(s)$ have been defined. Inserting this into the partition sum (3.7), one obtains

$$Z = \sum_{K=0}^{\infty} \sum_{\{n\}} \sum_{\{n^{(1)}\}} \cdots \sum_{\{n^{(K-1)}\}} \sum_{s_1} \sum_{s_2} \cdots \sum_{s_K} \int' \mathrm{d}^K \tau$$

$$(-1)^K e^{-\sum_{i=0}^{K} D_{\{n^{(i)}\}}(\tau_{i+1}-\tau_i)} \prod_{i=1}^{K} q_{\{n^{(i)}\}\{n^{(i-1)}\}}(s_i).$$

Obviously, the matrix elements of the kink operators $q_{\{n\},\{\bar{n}\}}(s)$ are already completely determined by s and $\{\bar{n}\}$, because $\{n\}$ follows from $\{\bar{n}\}$ by swapping the occupation numbers of the orbitals specified by s (if the orbitals are distinct, otherwise the matrix element is zero). Therefore, all but one of the summations over the occupation numbers $\{n^{(i)}\}$ can be dropped. Additionally, the requirement for β-periodicity fixes the last determinant to be equal to the first, i.e., $\{n^{(K)}\} = \{n\}$, so the summation over s_K is also not needed. This yields the final expression for the canonical partition function in the CPIMC formulation:

$$Z = \sum_{K=0}^{\infty} \sum_{\{n\}} \sum_{s_1} \sum_{s_2} \cdots \sum_{s_{K-1}} \int' \mathrm{d}^K \tau$$

$$(-1)^K e^{-\sum_{i=0}^{K} D_{\{n^{(i)}\}}(\tau_{i+1}-\tau_i)} \prod_{i=1}^{K} q_{\{n^{(i)}\}\{n^{(i-1)}\}}(s_i). \tag{3.9}$$

As before, it is $\tau_0 = 0$, $\tau_{K+1} = \beta$, and $\{n^{(0)}\} = \{n^{(K)}\} = \{n\}$. The last multi-index s_K is uniquely determined by $\{n\}$ and all other indices s_i with $i \neq K$.

In this formulation, a path $C = (K, \{n\}, s_1, \ldots, s_K, \tau_1, \ldots, \tau_K)$ is defined by the number of kinks K, the occupation numbers $\{n\}$ at imaginary time $\tau = 0$, the indices of the kinks s_i and their positions in imaginary time τ_i. Between two kinks at τ_i and τ_{i+1}, a determinant $|\{n^{(i)}\}\rangle$ is determined by $\{n\}$ and all preceding indices s_j with $j \leq i$. A kink with the indices $s_i = (p_i, q_i)$ corresponds to a one-particle excitation while $s_i = (p_i, q_i, r_i, s_i)$ describes a two-particle excitation. Occasionally, a multi-index s itself will be called a kink. The weight of a path is given by

$$W(C) = (-1)^K e^{-\sum_{i=0}^{K} D_{\{n^{(i)}\}}(\tau_{i+1}-\tau_i)} \prod_{i=1}^{K} q_{\{n^{(i)}\}\{n^{(i-1)}\}}(s_i). \tag{3.10}$$
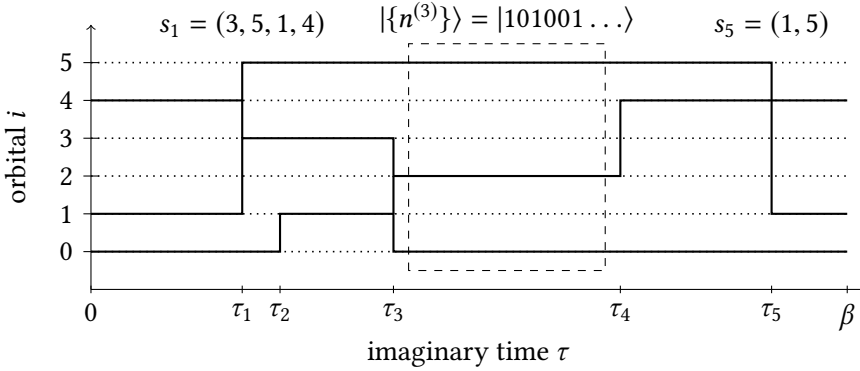
Figure 3.2: Visualization of a path in the kink picture. A configuration is given by the "imaginary time paths" of N-particles in the interval $[0, \beta]$. The y-axis represents the orbitals of the underlying single-particle basis. Thick lines show occupied orbitals while dotted lines indicate empty orbitals. Occupation changes occur at imaginary times $\tau_1, \ldots, \tau_5$, i.e., there are $K = 5$ kinks. Each kink corresponds to a one- or two-particle excitation. Adapted from [119].

Each path can be identified with a path in the previous formulation (3.7) and its weight has the same value. A typical path is sketched in Fig. 3.2. It is instructive to compare this illustration with the interpretation of the path integral in coordinate representation, see Fig. 2.1. Here, the discrete time slices are replaced by an continuous imaginary time while the continuous coordinate $\boldsymbol{r}$ is replaced by the discrete orbital number i.

The partition sum (3.9) will be the basis of the CPIMC method developed in this work. Its form will guide the choice of Monte Carlo steps which will be explained in detail in section 3.4. Before that, the most important estimators that can be derived from the partition function will be summarized in the next section. The introduction of the kink operator q simplifies the development of the worm-algorithm, which is discussed in section 3.7.

As the efficiency of any QMC method depends strongly on the sign problem, it is important to note that there are three different sources of sign changes in the partition function. First, and most obviously, the weight is multiplied by a negative sign for uneven numbers of kinks K. Second, the phase factors present in the Slater-Condon rules may change the sign depending on the number of occupied orbitals between the orbitals affected by the kink. And finally, the matrix elements themselves can have different signs. The impact of these signs on the

sign problem depends on the system and will be investigated in the corresponding chapters.

For a Monte Carlo simulation, the basis size has to be restricted to a finite number of orbitals or basis functions N_B. The single-particle orbitals $|i\rangle$ are ordered according to their orbital energy, i.e., $h_{ii} \leq h_{i+1,i+1}$, and all but the lowest N_B orbitals discarded. This is the only approximation of the CPIMC method and it is easy to control as the convergence to the complete basis set limit $N_B \to \infty$ is well-behaved in general. In a given basis $|i\rangle$, the matrix elements h_{ij} and w_{ijkl}^- can be calculated in advance with arbitrary accuracy. In some cases, like in the HEG, the calculation is so fast that one can recalculate the matrix elements each time they are used, which strongly decreases the memory requirements.

3.3 Expectation Values

Expectation values of thermodynamic observables in equilibrium can be calculated from the canonical partition function. The partition function (3.9) can be written as

$$Z = \sum_C \!\!\!\!\!\!\oint W(C).$$

If the expectation value of an observable $\hat{O}$ can be expressed as

$$\langle \hat{O} \rangle = \frac{1}{Z} \sum_C \!\!\!\!\!\!\oint O(C)W(C),$$

it can be estimated by a Monte Carlo simulation by sampling the value $O(C)$ with the weight $W(C)$. In the context of the Monte Carlo method, the term $O(C)$ is called an estimator. The estimators of some important quantities will be summarized in this section.

A central quantity is the total internal Energy of a system. A straightforward way to obtain an estimator for the energy is to use the well known thermodynamic relationship

$$\langle \hat{H} \rangle = -\frac{\partial}{\partial \beta} \ln Z$$

$$= \frac{1}{Z} \sum_C \!\!\!\!\!\!\oint \left(\sum_{i=0}^{K} D_{\{n^{(i)}\}} \frac{(\tau_{i+1} - \tau_i)}{\beta} - \frac{K}{\beta} \right) W(C).$$

Thus, the contribution of the diagonal part of the Hamiltonian to the estimator of the total energy $E(C)$ is given by the matrix elements $D_{\{n^{(i)}\}}$ of the determinants $|\{n^{(i)}\}\rangle$ that constitute the path C, each weighted by the ratio of the length $(\tau_{i+1} - \tau_i)$ of the corresponding fraction of the path to the total length β of the path. The off-diagonal part of the Hamiltonian contributes to the total energy only indirectly through the number of kinks K. This estimator is unchanged if the partition function in the form (3.7) is used instead.

A related thermodynamic quantity is the heat capacity at constant volume C_V. Its expectation value is given by

$$
\begin{aligned}
C_V &= \frac{\partial}{\partial T} \langle \hat{H} \rangle \Big|_V \\
&= \frac{1}{T^2} \frac{1}{Z} \oint_C \left(\left(E(C) - \langle \hat{H} \rangle \right)^2 - \frac{K}{\beta^2} \right) W(C).
\end{aligned}
$$

This differs from the quadratic fluctuation of the total energy $\langle (\hat{H} - \langle \hat{H} \rangle)^2 \rangle$ by the term $-K/\beta^2$.

The matrix elements of the reduced one-particle density matrix $d_{pq} = \langle \hat{a}_p^\dagger \hat{a}_q \rangle$ can be calculated by a partial differentiation of the partition function with respect to the one-particle matrix element h_{pq}:

$$
d_{pq} = -\frac{1}{\beta} \frac{1}{Z} \frac{\partial Z}{\partial h_{pq}}.
$$

Its diagonal elements d_{pp} are the average occupation numbers

$$
\langle \hat{n}_p \rangle = \frac{1}{Z} \oint_C \left(\sum_{i=0}^{K} n_p^{(i)} \frac{(\tau_{i+1} - \tau_i)}{\beta} \right) W(C),
$$

where, for a path $|\{n\}\rangle(\tau)$, the occupation number $n_p(\tau)$ is simply averaged over β. The off-diagonal case is more complicated. If the weight $W(C)$ does not vanish, i.e., if the matrix elements $q_{\{n^{(i)}\}\{n^{(i-1)}\}}(s_i)$ are nonzero for all i, then the estimator is given by

$$
d_{pq}(C) = -\frac{1}{\beta} \sum_{i=1}^{K} \frac{(-1)^{\alpha_{\{n^{(i)}\},p,q}}}{q_{\{n^{(i)}\}\{n^{(i-1)}\}}(s_i)} \delta_{s_i,(p,q)}. \tag{3.11}
$$

The importance of the reduced one-particle density matrix lies in the fact that the expectation value of any one-body operator $\hat{B}_1$ can be calculated from it:

$$
\langle \hat{B}_1 \rangle = \sum_{i,j=0}^{\infty} b_{ij} d_{ij}.
$$

An example is the one-particle density

$$n(\boldsymbol{r}) = \sum_{\sigma} \langle \hat{\Psi}^\dagger(x)\hat{\Psi}(x) \rangle = \sum_{\sigma} \sum_{i,j=0}^{\infty} \phi_i^*(x)\phi_j(x)d_{ij}. \tag{3.12}$$

Unfortunately, the variance of the estimator (3.11) is usually very high. The reason is that paths which contain kinks with a small weight $q_{\{n\},\{\bar{n}\}}(s)$ will rarely be sampled in a Monte Carlo simulation, but the value of the estimator is large for these paths. This problem is reduced by the introduction of an improved estimator for the worm algorithm, see section 3.7.

Analogously, the expectation value of any two-body operator $\hat{B}_2$ is given by the reduced two-particle density matrix $d_{pqrs} = \langle \hat{a}_p^\dagger \hat{a}_q^\dagger \hat{a}_s \hat{a}_r \rangle$:

$$\langle \hat{B}_2 \rangle = \frac{1}{2} \sum_{i,j,k,l=0}^{\infty} b_{ijkl}d_{ijkl}.$$

In a similar manner to the matrix elements of the one-particle density matrix d_{pq}, the matrix elements d_{pqrs} can be calculated by a differentiation of the partition sum. However, attention must be paid to the order of the indices in the differentiation with respect to w_{pqrs}^-. If $p = r$ and $q = s$ but $q \neq p$, the matrix elements are given by

$$d_{pqpq} = \frac{1}{Z} \oint_C \left(\sum_{i=0}^{K} n_p^{(i)} n_q^{(i)} \frac{(\tau_{i+1} - \tau_i)}{\beta} \right) W(C).$$

In the case of $p = r$ but with all other indices being different, the matrix elements are given by

$$d_{pqps} = -\frac{1}{\beta}\frac{1}{Z} \oint_C \left(\sum_{i=1}^{K} \frac{(-1)^{\alpha_{\{n^{(i)}\},q,s}}}{q_{\{n^{(i)}\}\{n^{(i-1)}\}}(s_i)} n_p^{(i)} \delta_{s_i,(q,s)} \right) W(C). \tag{3.13}$$

This expression is the same as for the off-diagonal element of the one-particle density matrix up to the additional factor $n_p^{(i)}$. If all indices are pairwise different as well as $p < q$ and $r < s$, one obtains

$$d_{pqrs} = -\frac{1}{\beta}\frac{1}{Z} \oint_C \left(\sum_{i=1}^{K} \frac{(-1)^{\alpha_{\{n^{(i)}\},p,q}+\alpha_{\{n^{(i-1)}\},r,s}}}{q_{\{n^{(i)}\}\{n^{(i-1)}\}}(s_i)} \delta_{s_i,(p,q,r,s)} \right) W(C). \tag{3.14}$$

All remaining matrix elements are given by the following symmetry relations:

$$d_{pqrs} = d_{qpsr} = -d_{pqsr} = -d_{qprs}.$$

Obviously, if $p = q$ or $r = s$, the matrix elements vanish, e.g.,

$$d_{pprs} = 0.$$

Note that the expressions for the off-diagonal elements of the d are only valid if the sum does not include terms with $q_{\{n^{(i)}\}\{n^{(i-1)}\}}(s_i) = 0$. This is satisfied in the context of a MC simulation, where such terms do not occur because of their vanishing weight. As this estimator suffers from the same variance problem as the estimator for the one-particle density matrix, it has not been implemented yet.

The two-particle density or pair-distribution function is given by

$$g(\mathbf{r}_1, \mathbf{r}_2) = \sum_{\sigma_1, \sigma_2} \langle \hat{\Psi}^\dagger(x_1)\hat{\Psi}^\dagger(x_2)\hat{\Psi}(x_2)\hat{\Psi}(x_1) \rangle$$

$$= \sum_{\sigma_1, \sigma_2} \sum_{i,j,k,l=0}^{\infty} \phi_i^*(x_1)\phi_j^*(x_2)\phi_k(x_1)\phi_l(x_2)d_{ijkl}. \tag{3.15}$$

It is normalized to $N(N-1)$ and proportional to the probability of finding a particle at coordinate $\mathbf{r}$ and another particle at coordinate $\mathbf{r}'$. In homogeneous systems, this quantity is a function of the relative coordinate $\mathbf{r} = \mathbf{r}_1 - \mathbf{r}_2$ only. In inhomogeneous systems, it is common to integrate over the center of mass coordinate $\mathbf{R} = (\mathbf{r}_1 + \mathbf{r}_2)/2$:

$$g(\mathbf{r}) = \int d\mathbf{R}\, g\left(\mathbf{R} + \frac{\mathbf{r}}{2}, \mathbf{R} - \frac{\mathbf{r}}{2}\right)$$

$$= \iint d\mathbf{r}_1 d\mathbf{r}_2\, \delta(\mathbf{r} - (\mathbf{r}_1 - \mathbf{r}_2))g(\mathbf{r}_1, \mathbf{r}_2). \tag{3.16}$$
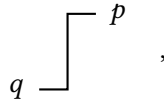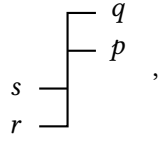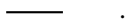
If the system is isotropic, $g(\mathbf{r})$ depends only on the distance r. For the two-dimensional harmonic trap, the integration is carried out in Appendix B.

3.4 Monte Carlo Steps

The high dimensional integrals and sums of the previous sections can be solved using Monte Carlo techniques. The method of choice is the Metropolis algorithm, because the normalization of the weights $W(C)$, i.e., the partition function, is

unknown, cf. section 2.2. The challenge is to find a set of Monte Carlo updates that fulfill the detailed balance equation (2.3) and that are both ergodic and efficient. If the first two conditions are met, the Metropolis algorithm will create a Markov chain where the paths C are distributed according to their probability $p(C) = W(C)/Z$ and the expectation values can be estimated by a simple arithmetic mean over the sampled configurations.

The updates are ergodic if starting from any path with non-vanishing weight C_0 every other path C with initial occupation numbers $\{n\}$, K kinks with excitations s_i and imaginary times τ_i, and non-vanishing weight $W(C)$ can be generated within a finite number of Monte Carlo steps. In the diagrammatic representation this is identical to the ability to construct any β-periodic path consisting of

- an arbitrary number of type 2 kinks $s = (p, q)$:

- an arbitrary number of type 4 kinks $s = (p, q, r, s)$:

- and horizontal lines either connecting these kinks or going from 0 to β undisturbed:

Note, that for an N particle configuration, there are N horizontal lines at each imaginary time τ between two consecutive kinks.

In general, the Hamiltonian $\hat{H}$ is a sparse matrix. Therefore, the matrix element $q_{\{n\},\{\bar{n}\}}(s)$ of most kinks and accordingly the weight of most paths vanishes. This makes it difficult to ensure ergodicity, because the forbidden paths can separate the configuration space into islands that are not connected by a given set of Monte Carlo moves. Thus, the correctness of the current set of updates has to be checked for every Hamiltonian with a new structure by comparing with an exact diagonalization method like CI. For larger systems where a comparison with CI is not possible, artificial weights can be assigned to all kinks with small matrix elements $|q_{\{n\},\{\bar{n}\}}(s)| < q_{\min}$. The artificial weights and the cutoff $q_{\min}$ can be chosen arbitrarily but have to be fixed during a simulation. These so-called virtual kinks can serve as bridges between separated islands in the configuration space and hence restore ergodicity. Only paths with no virtual kinks are valid samples for the calculation of expectation values and all paths containing virtual kinks have to be discarded. The disadvantage of this method is the need to fine-tune the virtual weights for every set of system parameters. If the weight is too large,

the simulation will spend most of its time in virtual configuration and only few samples can be obtained resulting in large statistical errors. On the other hand, if the weight is too low, switches between disconnected regions of the configuration space happen too rarely and the algorithm is effectively non-ergodic. Therefore, a large amount of time was invested in finding a set of Monte Carlo updates that are ergodic for the Hamiltonians considered in this work. The Monte Carlo moves presented here are the result of a lengthy trial-and-error development process, which is described in [59].

While ergodicity is a necessary condition for correctness, efficiency is usually a trade-off between the CPU time of a single MC step, the autocorrelation time of the resulting Markov chain, and the complexity of the implementation. The efficiency of the Monte Carlo moves depends on the structure of the Hamiltonian and can be optimized with detailed knowledge about the system under consideration. The following Monte Carlo updates are designed to work with a general Hamiltonian consisting of arbitrary one-body operators and a two-body operator describing the pair potential. A set of moves that are optimized by taking the special structure of the Hamiltonian into account is presented in chapter 5 for the HEG.

The following rather technical description of the MC moves is included here because the development and implementation of an ergodic and efficient algorithm constitutes a central achievement of this work. The level of detail ensures the reproducibility of the results in the following chapters. Nevertheless, the rest of this section can be easily skip on a first read.

Apart from the Monte Carlo moves of the worm algorithm that are not included here and will be considered in section 3.7, there are six different MC moves implemented in the current version of CPIMC. For each of the first five moves, there are two branches exciting either one or two particles in a specific β-periodic interval. The sixth MC move excites a particle at all times instead. A β-periodic interval is denoted by $[\tau_a, \tau_b]_\beta$. It is

$$\tau \in [\tau_a, \tau_b]_\beta := \begin{cases} \tau \in [\tau_a, \tau_b], & \text{if } \tau_a \leq \tau_b \\ \tau \in [0, \tau_b] \cup [\tau_a, \beta], & \text{if } \tau_b < \tau_a \end{cases}.$$

The length of such an interval is given by

$$L_\beta(\tau_a, \tau_b) := \begin{cases} \tau_b - \tau_a, & \text{if } \tau_a \leq \tau_b \\ \beta - (\tau_a - \tau_b), & \text{if } \tau_b < \tau_a \end{cases}.$$

In the following, $\tau_{\min}$ is defined by the time of the first kink $s_{\min}$ that is on the left of some time τ and that affects some orbital p. Here, the first kink on the

Figure 3.3: Illustration of the MC steps for adding and removing a pair of kinks. The sketches depict only a relevant subset of a path. The orbitals shown here do not necessarily need to be consecutive or in the same order in the full path. Due to the β-periodicity, the time interval that is shown potentially includes the point $\tau = 0$ or $\tau = \beta$. Left: one-particle excitation **1.**A) $\leftrightarrow$ **2.**A). Right: two-particle excitation **1.**B) $\leftrightarrow$ **2.**B).

left of a time τ is the kink which has the smallest interval length $L_\beta(\tau_{\min}, \tau)$ of all kinks affecting the orbital p, i.e., $L_\beta(\tau_{\min}, \tau) < L_\beta(\tau_i, \tau)$ for all other kinks s_i that affect the orbital p at the times τ_i. An orbital p is affected by a kink s if $p \in s$. Analogously, the first kink that is on the right of the time τ and that is affecting the orbital p is denoted by $s_{\max}$ with the time $\tau_{\max}$. It has the smallest interval length $L_\beta(\tau, \tau_{\max})$ of all kinks affecting that orbital.

As the order of the orbitals is only important for the calculation of acceptance ratios and expectation values, the conditions $p < q$ and $r < s$ for kinks of type 4 are ignored in the Monte Carlo algorithm. The Monte Carlo moves are given in detail by:

1. Add two kinks

    A) Excite one particle

        i. Choose a time $\tau_a \in [0, \beta)$ at random.

        ii. At τ_a, randomly choose an occupied orbital $p \in [0, N_B - 1]$, i.e, an orbital with $n_p(\tau_a) = 1$.

        iii. At τ_a, randomly choose an empty orbital $q \in [0, N_B - 1]$, i.e., an orbital with $n_q(\tau_a) = 0$.

        iv. If no kink affects the orbitals p and q:

            a) Choose a time $\tau_b \in [0, \beta)$ at random. The new kinks are given by $s_a = (q, p)$ and $s_b = (p, q)$ at the times τ_a and τ_b, respectively.

        Otherwise choose randomly between b) left or c) right:

            b) Randomly choose $\tau_b \in [\tau_{\min}, \tau_a)_\beta$, where $\tau_{\min}$ is the time of the first kink left of τ_a that affects at least one of the orbitals p and q. The new kinks are given by $s_a = (p, q)$ and $s_b = (q, p)$.

c) Randomly choose $\tau_b \in [\tau_a, \tau_{\max})_\beta$, where $\tau_{\max}$ is the time of the first kink right of τ_a that affects at least one of the orbitals p and q. The new kinks are given by $s_a = (q, p)$ and $s_b = (p, q)$.

B) Excite two particles

i. Choose a time $\tau_a \in [0, \beta)$ at random.

ii. At τ_a, choose two different random occupied orbitals p and q, i.e., $p, q \in [0, N_B - 1]$, $q \neq p$, and $n_p(\tau_a) = n_q(\tau_a) = 1$.

iii. At τ_a, choose two different random empty orbitals r and s.

iv. If no kink affects the orbitals p, q, r and s:

a) Choose a time $\tau_b \in [0, \beta)$ at random. The new kinks are given by $s_a = (r, s, p, q)$ and $s_b = (p, q, r, s)$.

Otherwise choose randomly between b) left or c) right:

b) Randomly choose $\tau_b \in [\tau_{\min}, \tau_a)_\beta$, where $\tau_{\min}$ is the time of the first kink left of τ_a that affects at least one of the orbitals p, q, r and s. The new kinks are given by $s_a = (p, q, r, s)$ and $s_b = (r, s, p, q)$.

c) Randomly choose $\tau_b \in [\tau_a, \tau_{\max})_\beta$, where $\tau_{\max}$ is the time of the first kink right of τ_a that affects at least one of the orbitals p, q, r and s. The new kinks are given by $s_a = (r, s, p, q)$ and $s_b = (p, q, r, s)$.

2. Remove two kinks

Choose a kink s_a at random. There are two possibilities:

A) The kink is given by $s_a = (p, q) \Rightarrow$ one-particle excitation

i. Choose a random direction: a) left or b) right:

a) Find the first kink on the left of s_a affecting either p or q. If it is $s_b = (q, p)$, then the two kinks s_a and s_b form a kink pair and can be removed.

b) Find the first kink on the right of s_a affecting either p or q. Remove both kinks if $s_b = (q, p)$.

B) The kink is given by $s_a = (p, q, r, s) \Rightarrow$ two-particle excitation

i. Choose a random direction: a) left or b) right:

a) Find the first kink on the left of s_a affecting either one of the orbitals p, q, s or r. If this kink, s_b, and s_a form a kink pair, i.e.
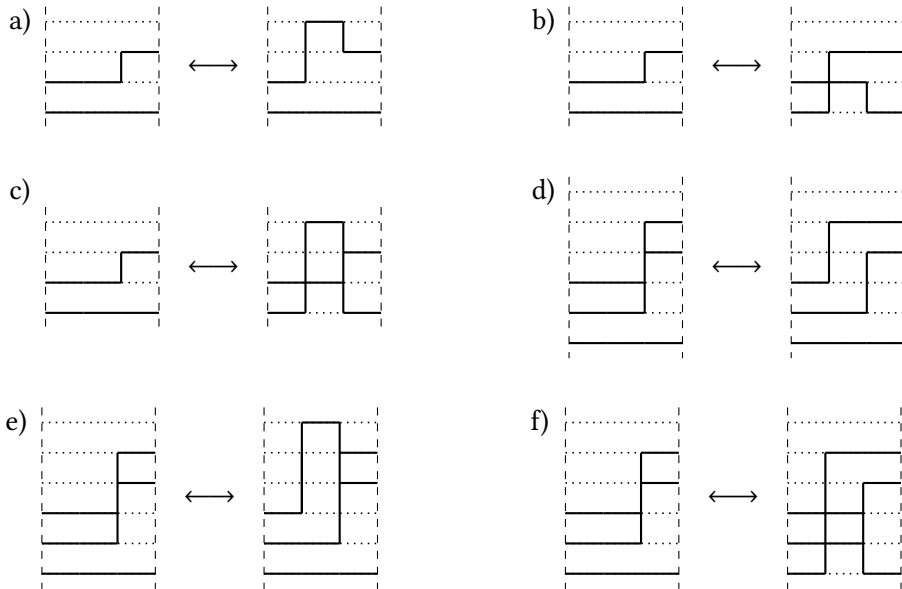
Figure 3.4: Add or remove a single kink, case A). The labels a) to f) correspond to the cases of point v. of step **3.**A) for adding a kink and of point v. of step **3.**B) for removing a kink. See also Fig. 3.3.

if s_b is one of (r, s, p, q), (s, r, p, q), (r, s, q, p) or (s, r, q, p), then remove both kinks.

b) Find s_b, the first kink on the right of s_a affecting either p, q, s or r. If $s_b \in \{(r, s, p, q), (s, q, p, q), (r, s, q, p), (s, r, q, p)\}$, remove both kinks.

3. Add one kink

Choose a kink s_a and a direction at random.

A) Excite one particle left of s_a

i. Choose a random orbital p from the occupied orbitals left of s_a

ii. Choose an empty orbital q. Distinguish the following cases

a) If the kink s_a is of type 4 and $p \notin s_a$, i.e $s_a = (i, j, k, l)$ with $p \neq k$ and $p \neq l$, choose either i or j as the target orbital q

b) In all other cases choose q randomly from all empty orbitals left of s_a. Reject the move if $s_a = (q, p)$.
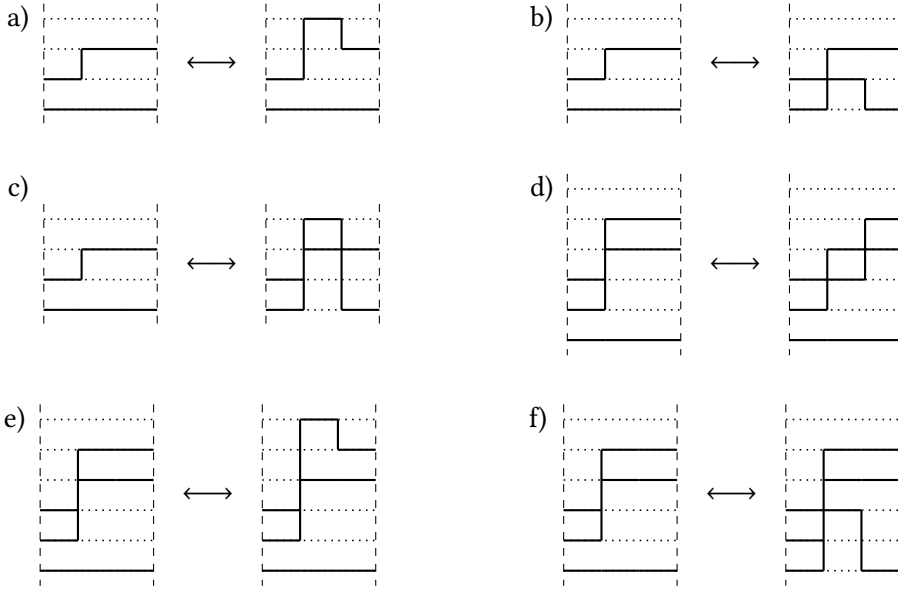
Figure 3.5: Add or remove a single kink, case B). The labels a) to f) correspond to the cases of point v. of step **3.**B) for adding a kink and of point v. of step **3.**A) for removing a kink. See also Fig. 3.3.

    iii. Find the time $\tau_{\min}$ of the first kink left of τ_a affecting either p or q

    iv. Choose a time $\tau_b \in [\tau_{\min}, \tau_a)_\beta$ at random. The new kink is given by
$$s_b = (q, p)$$

    v. The kink s_a is changed to s_a' according to the following cases:

        a) $s_a = (i, p) \;\Rightarrow\; s_a' = (i, q)$

        b) $s_a = (q, i) \;\Rightarrow\; s_a' = (p, i)$

        c) $s_a = (i, j) \;\Rightarrow\; s_a' = (i, p, j, q)$

        d) $s_a \in \{(q, i, p, j), (i, q, p, j), (q, i, j, p), (i, q, j, p)\} \Rightarrow s_a' = (i, j)$

        e) $s_a = (i, j, p, k) \;\Rightarrow\; s_a' = (i, j, q, k)$ or
           $s_a = (i, j, k, p) \;\Rightarrow\; s_a' = (i, j, k, q)$

        f) $s_a = (q, j, k, l) \;\Rightarrow\; s_a' = (p, j, k, l)$ or
           $s_a = (i, q, k, l) \;\Rightarrow\; s_a' = (i, p, k, l)$

B) Excite one particle right of s_a

    i. Choose a random orbital p from the occupied orbitals right of s_a.

    ii. Choose an empty orbital q. Distinguish the following cases:

        a) If s_a is of type 4 and $p \notin s_a$, i.e., $s_a = (i, j, k, l)$ with $p \neq i$ and $p \neq j$, choose either k or l as the empty orbital q.

        b) In all other cases choose q randomly from all empty orbitals right of s_a. Reject the move if $s_a = (p, q)$.

    iii. Find the time $\tau_{\max}$ of the first kink right of τ_a affecting either p or q.

    iv. Randomly choose the time $\tau_b \in [\tau_a, \tau_{\max})_\beta$. The new kink is given by $s_b = (p, q)$.

    v. The kink s_a changes in one of the following ways:

        a) $s_a = (p, i) \implies s_a' = (q, i)$

        b) $s_a = (i, q) \implies s_a' = (i, p)$

        c) $s_a = (i, j) \implies s_a' = (i, q, j, p)$

        d) $s_a \in \{(p, i, q, j), (i, p, q, j), (p, i, j, q), (i, p, j, q)\} \implies s_a' = (i, j)$

        e) $s_a = (p, i, j, k) \implies s_a' = (q, i, j, k)$ or
           $s_a = (i, p, j, k) \implies s_a' = (i, q, j, k)$

        f) $s_a = (i, j, q, l) \implies s_a' = (i, j, p, l)$ or
           $s_a = (i, j, k, q) \implies s_a' = (i, j, k, p)$

    These correspond to the possibilities of point v. of step **3**.A) with the indices of the creation and annihilation operators exchanged.

C) Excite two particles left of s_a

    i. Choose a random orbital p from the occupied orbitals left of s_a.

    ii. Choose a second random orbital $q \neq p$ from the occupied orbitals left of s_a.

    iii. For the choice of an empty orbital r left of s_a consider the following cases:

        a) If $s_a = (i, j)$ with $p \neq j$ and $q \neq j$, it is $r = i$.

        b) If $s_a = (i, j, k, l)$ with either $p \in [k, l]$ or $q \in [k, l]$ but not both, choose either $r = i$ or $r = j$.

        c) If $s_a = (i, j, k, l)$ and both $p \notin [k, l]$ and $q \notin [k, l]$, it is $r = i$.

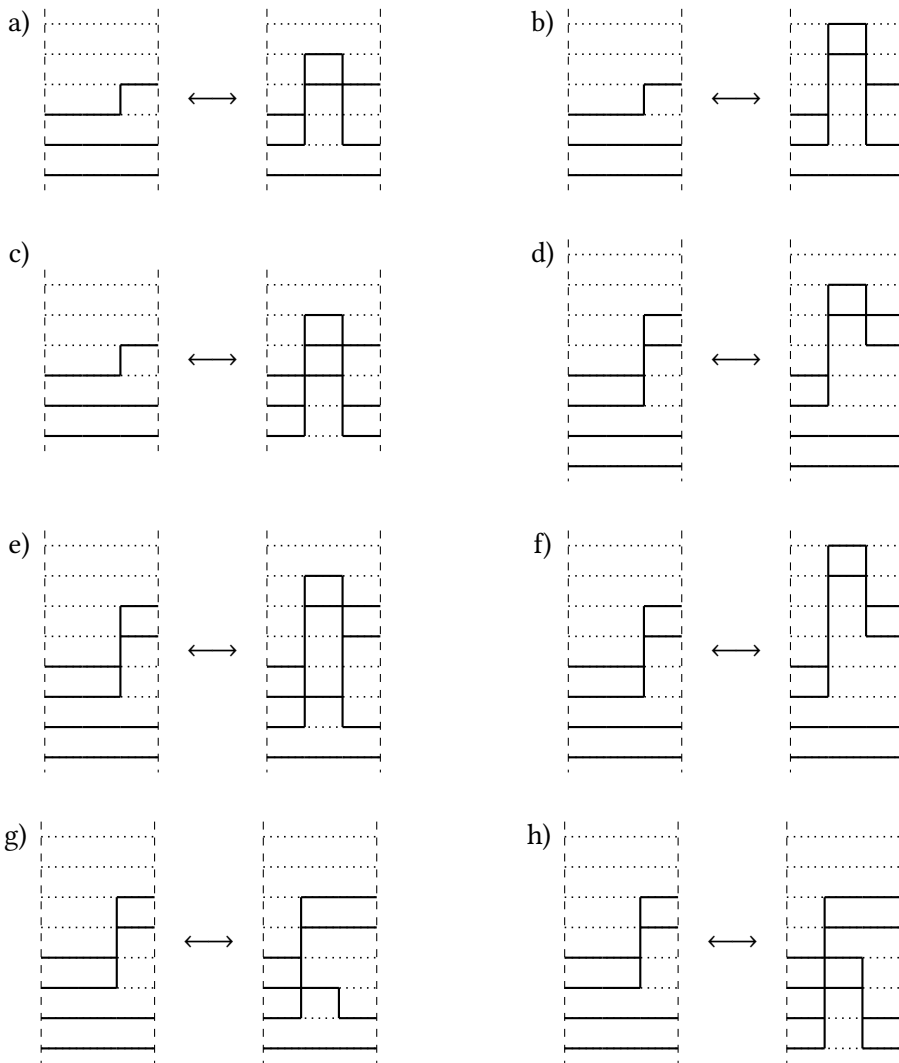   d) In all other cases, choose an empty orbital r left of s_a at random.

Figure 3.6: Add or remove a single kink, case C). The labels a) to h) correspond to the cases of point vii. of step **3.**C) for adding a kink and of point vii. of step **3.**D) for removing a kink. See also Fig. 3.3.

iv. For the choice of an orbital s which is empty left of s_a, the following cases have to be distinguished:

    a) If $s_a = (i, j, k, l)$ and both $p \notin [k, l]$ and $q \notin [k, l]$, it is $s = j$.

    b) In all other cases, choose a random empty orbital left of s_a with $r \neq s$. If $s_a \in [(r, s, p, q), (s, r, p, q), (r, s, q, p), (s, r, q, p)]$, reject the move.

v. From all kinks left of s_a that affect the orbitals p, q, r, or s, find the closest one with time $\tau_{\min}$.

vi. Choose a time $\tau_b \in [\tau_{\min}, \tau_a)_\beta$ at random. The new kink is given by $s_b = (r, s, p, q)$.

vii. The kink s_a changes in one of the following ways:

    a) $s_a = (r, p) \implies s_a' = (q, s)$ or
      $s_a = (r, q) \implies s_a' = (p, s)$ or
      $s_a = (s, p) \implies s_a' = (q, r)$ or
      $s_a = (s, q) \implies s_a' = (p, r)$

    b) $s_a = (i, p) \implies s_a' = (i, q, r, s)$ or
      $s_a = (i, q) \implies s_a' = (i, p, r, s)$

    c) $s_a = (r, j) \implies s_a' = (p, q, j, s)$

    d) $s_a \in [(r, i, p, q), (i, r, p, q), (r, i, q, p), (i, r, q, p)] \implies s_a' = (i, s)$
      or
      $s_a \in [(s, i, p, q), (i, s, p, q), (s, i, q, p), (i, s, q, p)] \implies s_a' = (i, r)$

    e) $s_a = (r, j, p, l) \implies s_a' = (q, j, s, l)$ or
      $s_a = (r, j, k, p) \implies s_a' = (q, j, k, s)$ or
      $s_a = (r, j, q, l) \implies s_a' = (p, j, s, l)$ or
      $s_a = (r, j, k, q) \implies s_a' = (p, j, k, s)$ or
      $s_a = (i, r, p, l) \implies s_a' = (i, q, s, l)$ or
      $s_a = (i, r, k, p) \implies s_a' = (i, q, k, s)$ or
      $s_a = (i, r, q, l) \implies s_a' = (i, p, s, l)$ or
      $s_a = (i, r, k, q) \implies s_a' = (i, p, k, s)$

    f) $s_a \in [(i, j, p, q), (i, j, q, p)] \implies s_a' = (i, j, r, s)$

    g) $s_a \in [(r, s, p, l), (s, r, p, l)] \implies s_a' = (q, l)$ or
      $s_a \in [(r, s, q, l), (s, r, q, l)] \implies s_a' = (p, l)$ or
      $s_a \in [(r, s, k, p), (s, r, k, p)] \implies s_a' = (q, k)$ or
      $s_a \in [(r, s, k, q), (s, r, k, q)] \implies s_a' = (p, k)$

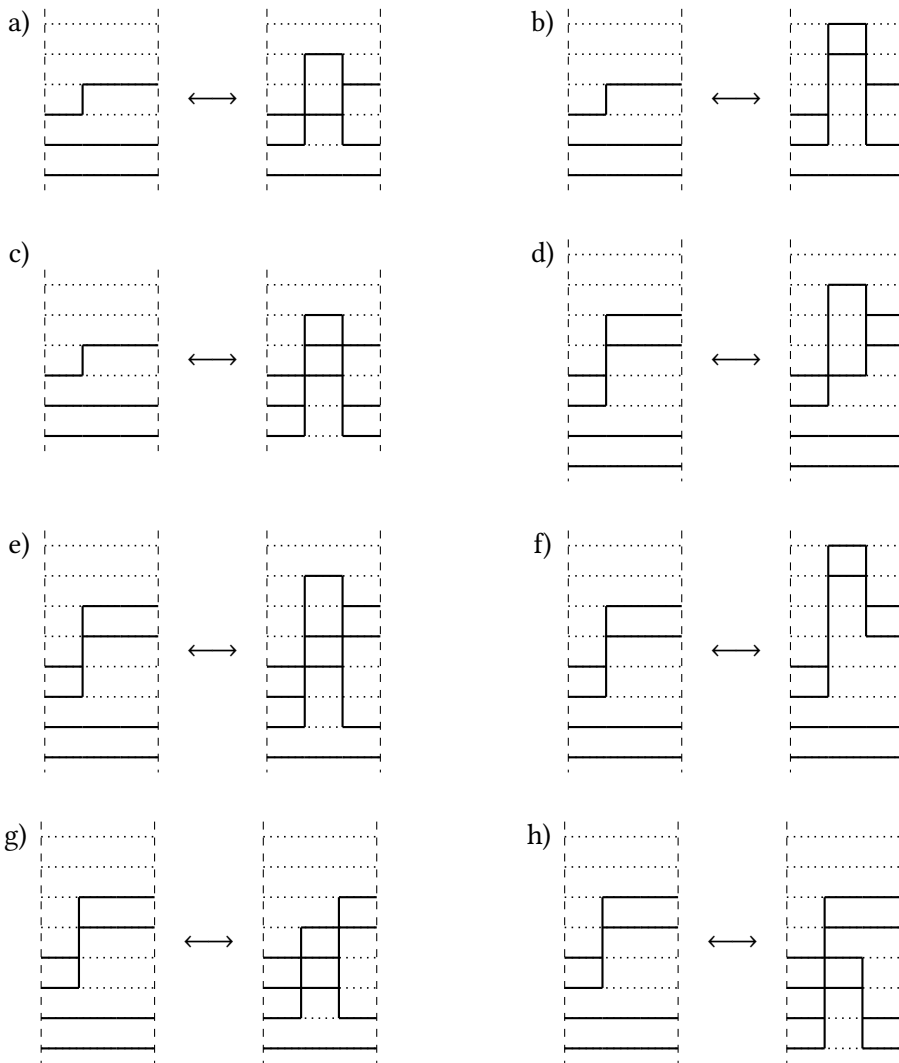   h) $s_a = (r, s, k, l) \implies s_a' = (p, q, k, l)$

Figure 3.7: Add or remove a single kink, case D). The labels a) to h) correspond to the cases of point vii. of step **3.**D) for adding a kink and of point vii. of step **3.**C) for removing a kink. See also Fig. 3.3.

D) Excite two particles right of s_a

   i. Choose a random orbital p from the occupied orbitals right of s_a.

   ii. Choose a second random orbital $q \neq p$ from the occupied orbitals right of s_a.

   iii. For the choice of an empty orbital r right of s_a, consider the following cases:

      a) If $s_a = (i, j)$ with $p \neq i$ and $q \neq i$, it is $r = k$.

      b) If $s_a = (i, j, k, l)$ with either $p \in [i, j]$ or $q \in [i, j]$ but not both, choose either $r = k$ or $r = l$.

      c) If $s_a = (i, j, k, l)$ and both $p \notin [i, j]$ and $q \notin [i, j]$, it is $r = k$.

      d) In all other cases, choose an empty orbital r right of s_a at random.

   iv. For the choice of a second empty orbital s right of s_a, the following cases have to be distinguished:

      a) If $s_a = (i, j, k, l)$ and both $p \notin [i, j]$ and $q \notin [i, j]$, it is $s = l$.

      b) In all other cases, choose a random empty orbital right of s_a with $r \neq s$. If $s_a \in [(p, q, r, s), (p, q, s, r), (q, p, r, s), (q, p, s, r)]$, reject the move.

   v. From all kinks right of s_a that affect the orbitals p, q, r, or s, find the closest one. It defines the time τ_{max}.

   vi. Choose a time $\tau_b \in [\tau_a, \tau_{max})_\beta$ at random. The new kink is given by $s_b = (p, q, r, s)$.

   vii. The kink s_a changes in one of the following ways:

      a) $s_a = (p, r) \implies s_a' = (s, q)$ or
         $s_a = (q, r) \implies s_a' = (s, p)$ or
         $s_a = (p, s) \implies s_a' = (r, q)$ or
         $s_a = (q, s) \implies s_a' = (r, p)$

      b) $s_a = (p, j) \implies s_a' = (r, s, j, q)$ or
         $s_a = (q, j) \implies s_a' = (r, s, j, p)$

      c) $s_a = (i, r) \implies s_a' = (i, s, p, q)$

      d) $s_a \in [(p, q, r, i), (p, q, i, r), (q, p, r, i), (q, p, i, r)] \implies s_a' = (s, i)$
         or
         $s_a \in [(p, q, s, i), (p, q, i, s), (q, p, s, i), (q, p, i, s)] \implies s_a' = (r, i)$

e) $s_a = (p, j, r, l) \implies s'_a = (s, j, q, l)$ or
   $s_a = (p, j, k, r) \implies s'_a = (s, j, k, q)$ or
   $s_a = (q, j, r, l) \implies s'_a = (s, j, p, l)$ or
   $s_a = (q, j, k, r) \implies s'_a = (s, j, k, p)$ or
   $s_a = (i, p, r, l) \implies s'_a = (i, s, q, l)$ or
   $s_a = (i, p, k, r) \implies s'_a = (i, s, k, q)$ or
   $s_a = (i, q, r, l) \implies s'_a = (i, s, p, l)$ or
   $s_a = (i, q, k, r) \implies s'_a = (i, s, k, p)$

f) $s_a \in [(p, q, i, j), (q, p, i, j)] \implies s'_a = (r, s, i, j)$

g) $s_a \in [(p, j, r, s), (p, j, s, r)] \implies s'_a = (j, q)$ or
   $s_a \in [(q, j, r, s), (q, j, s, r)] \implies s'_a = (j, p)$ or
   $s_a \in [(i, p, r, s), (i, p, s, r)] \implies s'_a = (i, q)$ or
   $s_a \in [(i, q, r, s), (i, q, s, r)] \implies s'_a = (i, p)$

h) $s_a = (i, j, r, s) \implies s'_a = (i, j, p, q)$

These correspond to the possibilities of point vii. of step **3.**C) with the indices of the creation and annihilation operators exchanged.

4. Remove one kink

Choose a kink s_a and a direction at random. If the kink is of type 2, its removal will result in a one-particle excitation. If it is of type 4, excite two particles instead. There are four possibilities:

A) Excite one particle right of $s_a = (p, q)$

   i. Find the time $\tau_{\max}$ of the nearest kink that affects one of the orbitals p or q right of s_a.

   ii. Choose a kink s_b at random in the interval $[\tau_a, \tau_{\max}]_\beta$ including only kinks succeeding s_a in the total ordering of kinks.

   iii. Reject the step if $s_b = (q, p)$, otherwise remove s_a.

   iv. The kink s_b changes according to the cases in point v. of step **3.**A). Reject if none of these conditions applies.

B) Excite one particle left of $s_a = (q, p)$

   i. Find the closest kink left of s_a that affects one of the orbitals p or q. It defines the time $\tau_{\min}$.

   ii. Choose a random kink s_b from all kinks preceding s_a in the interval $[\tau_{\min}, \tau_a]_\beta$.

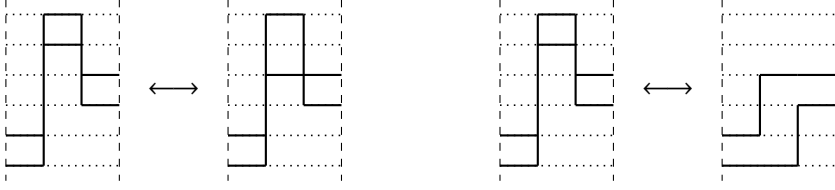   iii. Reject the step if $s_b = (p, q)$. Otherwise, remove s_a.

Figure 3.8: Change two kinks. Due to the large number of different cases, only one example is shown for a one-particle excitation (left) and a two-particle excitation (right), respectively. See also Fig. 3.3.

    iv. The kink s_b changes according to point v. of step **3.**B). Reject if none of these conditions applies.

C) Excite two particles right of $s_a = (p, q, r, s)$

    i. Find the closest kink on the right of s_a that affects one of the orbitals p, q, r, or s. It defines the time τ_{max}.

    ii. Randomly choose a kink in the interval $[\tau_a, \tau_{max}]_\beta$ excluding s_a and kinks preceding it.

    iii. If $s_b \in [(r, s, p, q), (r, s, q, p), (s, r, p, q), (s, r, q, p)]$, reject the step. Remove s_a otherwise.

    iv. The kink s_b changes according to point vii. of step **3.**C). Reject if none of these conditions applies.

D) Excite two particles left of $s_a = (r, s, p, q)$

    i. Find the closest kink left of s_a that affects one of the orbitals p, q, r, or s. It defines the time τ_{min}.

    ii. Choose a random kink s_b from all kinks left of s_a in the interval $[\tau_{min}, \tau_a]_\beta$.

    iii. Reject the step if $s_b \in [(p, q, r, s), (p, q, s, r), (q, p, r, s), (q, p, s, r)]$. Otherwise, remove s_a.

    iv. The kink s_b changes according to the cases in point vii. of step **3.**D). Reject if none of these conditions applies.

5. Change two kinks

A) One-particle excitation

    i. Choose a kink s_a at random.

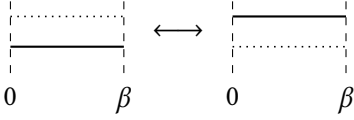   ii. Choose a random orbital p from the occupied orbitals right of s_a.

Figure 3.9: Excite whole orbital. Both orbitals must not contain any kinks in the whole interval $[0, \beta]$.

iii. An empty orbital q is chosen analogously to point ii. of step **3.**B).

iv. Find the time $\tau_{\max}$ of the closest kink at the right of s_a that affects either orbital p or q.

   a) If no such kink exists, randomly choose a kink $s_b \neq s_a$ from all kinks in the current path.

   b) Otherwise, choose a random kink s_b in the range $[\tau_a, \tau_{\max}]_\beta$ excluding kink s_a and all kinks preceding s_a in the total ordering of kinks.

v. The kinks s_a and s_b change according to point v. of step **3.**B) and point v. of step **3.**A), respectively. Reject if none of these conditions applies.

B) Two-particle excitation

   i. Choose a kink s_a at random.

   ii. Choose a random orbital p from the occupied orbitals right of s_a.

   iii. Choose a second random orbital $q \neq p$ from the occupied orbitals right of s_a.

   iv. Choose an empty orbital at the right of s_a as in point iii. of step **3.**D).

   v. Choose a second empty orbital right of s_a as in point iv. of step **3.**D).

   vi. Find the time $\tau_{\max}$ of the closest kink at the right of s_a that affects one of the orbitals p, q, r, or s.

   vii. Choose a kink s_b at random in the range $[\tau_a, \tau_{\max}]_\beta$ other than s_a and its predecessors.

   viii. The kinks s_a and s_b change according to point vii. of step **3.**D) and point vii. of step **3.**C), respectively. Reject if none of these conditions applies.

6. Excite whole orbital

   i. Randomly choose an occupied orbital p that is not affected by any kink.

        ii. Randomly choose an empty orbital q that does not contain any kinks.

        iii. Excite the particle from p to q.

The steps **1.** to **5.** are chosen with equal probability. For steps **1.**, **3.**, and **5.**, the branch is chosen with equal probability as well, while for steps **2.** and **4.** the number of excitations is determined by the chosen kink. With probability $1/K$, step **6.** is chosen instead of step **5.**A). If one of the steps is not possible, e.g., because there are not enough occupied or unoccupied orbitals at the chosen times, the whole update is simply rejected.

    The steps above are able to satisfy the detailed balance condition because for each MC move there exists another move that transforms the path back to the previous configuration when executed consecutively. For all such pairs of MC moves, the acceptance probabilities are determined by the detailed balance equation (2.3) as given below.

    Despite being chosen uniformly in the current implementation as described above, the probability to choose a certain move is denoted by p_{ap} and p_{rp} for adding and removing a pair of kinks, p_{ak} and p_{rk} for adding and removing a single kink, p_{ck} for changing two kinks and p_{eo} for exciting a whole orbital, respectively. This will allow to change the probabilities in future simulations, which may increase their efficiency. In the current program code, it is also possible to adapt the probabilities for choosing between a one and a two-particle excitation, but here this is neglected for simplicity. A configuration or path is denoted by $C = (K, \{n\}, s_1, \ldots, s_K, \tau_1, \ldots, \tau_K)$ and its weight by (cf. Eq. (3.10))

$$W(C) = (-1)^K e^{-\sum_{i=0}^{K} D_{\{n(i)\}}(\tau_{i+1}-\tau_i)} \prod_{i=1}^{K} q_{\{n^{(i)}\}\{n^{(i-1)}\}}(s_i).$$

The number of choices for the orbital i are denoted by N_i while the number of occupied (filled) and empty orbitals are denoted by N_f and N_e, respectively. The orbitals are chosen with an equal probability of $1/N_i$. The total probability to choose all orbitals is denoted by $1/N_o$. A new imaginary time τ can as well be chosen uniformly in the interval $[\tau_{\min}, \tau_{\max})_\beta$, in which case the distribution is given by $p(\tau) = 1/L_\beta(\tau_{\min}, \tau_{\max})$. In addition, it is possible to improve the efficiency by finding a new time τ according to the heat bath idea as explained in the next section. If a kink has to be chosen from the kinks in a certain range, the number of kinks in that range will be denoted with N_K. How that range is found is described in the corresponding step above. In the following, the factors in the detailed balance will be listed in the order of random choices used in the above description of the MC moves.

1.↔2. The detailed balance for the transition between a configuration C with K kinks and a configuration C' with $K + 2$ kinks by adding and removing a pair of kinks as shown in Fig. 3.3 is given by:

$$p_{\mathrm{ap}} \frac{1}{2} \frac{1}{\beta} \frac{1}{N_o} p(\tau_b) |W(C)| \nu(C \to C')$$
$$= p_{\mathrm{rp}} \frac{1}{K+2} |W(C')| \nu(C' \to C),$$

where $p(\tau_b) = 1/\beta$ if no kink affects the orbitals p and q, see case iv.a), and $p(\tau_b) = 1/L_\beta(\tau_{\min}, \tau_a)$ or $p(\tau_b) = 1/L_\beta(\tau_a, \tau_{\max})$ in case iv.b) or iv.c), respectively. The probability to choose a direction does not appear in the equation, because in each case it is canceled by a factor of 2 due to the two different ways to create the same configuration. It is $N_o = N_p N_q = N_f N_e$ for a one-particle excitation, case A), and $N_o = N_p N_q N_r N_s = N_f (N_f - 1) N_e (N_e - 1)$ for a two-particle excitation, case B).

3.↔4. This pair of steps changes the configuration C with K kinks to a configuration C' with $K + 1$ kinks and vice versa via a one-particle excitation. All qualitatively different possibilities for the cases A) to D) are depicted in figures 3.4 to 3.7. The detailed balance equation is given by:

$$p_{\mathrm{ak}} \frac{1}{2} \frac{1}{K} \frac{1}{2} \frac{1}{N_o} p(\tau_b) |W(C)| \nu(C \to C') \tag{3.17}$$
$$= p_{\mathrm{rk}} \frac{1}{K+1} \frac{1}{2} \frac{1}{N_K} |W(C')| \nu(C' \to C).$$

The probability for choosing the time τ_b of the new kink is either $p(\tau_b) = 1/L_\beta(\tau_{\min}, \tau_a)$ in the cases A) and C) or $p(\tau_b) = 1/L_\beta(\tau_a, \tau_{\max})$ in the cases B) and D), while the probability for choosing the orbitals, $1/N_o$, depends on both the randomly chosen and the added kink. For a one-particle excitation, it is $1/N_o = 1/N_p N_q$ with $N_p = N_f$. In addition, it is $N_q = 2$ in case v.f) of A) or B), and $N_q = N_e$ otherwise.

The two-particle excitation in cases C) and D) is chosen with probability $1/N_o = 1/N_p N_q N_r N_s$, where $N_p N_q = N_f (N_f - 1)/2$ always applies. The additional factor of $1/2$ corresponds to the two different ways the same two occupied orbitals can be chosen. The cases a) to h) in point vii. can be

assigned to five groups that determine the remaining two factors:

$$N_r N_s = \begin{cases} N_e(N_e - 1)/2, & \text{cases a), b), d) and f),} \\ 1(N_e - 1), & \text{case c),} \\ 2(N_e - 1), & \text{case e),} \\ 2(N_e - 1)/2, & \text{case g),} \\ 1, & \text{case h).} \end{cases}$$

Again, factors of 1/2 occur when two different choices can lead to the same configuration.

5.↔5. Changing two kinks via a one- or two-particle excitation switches between two configurations, C and C', with the same number of kinks K. The detailed balance can be written as:

$$p_{ck} \frac{1}{2} \frac{1}{K} \frac{1}{N_o} \frac{1}{N_K} |W(C)| v(C \to C')$$
$$= p_{ck} \frac{1}{2} \frac{1}{K} \frac{1}{N_o} \frac{1}{N'_K} |W(C')| v(C' \to C),$$

where in case B) $N_K = N'_K$, because the second kink s_b always affects one of the orbitals p, q, r, and s. The choices of the orbitals are the same as in **3.↔4.** on both sides of the equation and thus cancel each other in the acceptance probability. Two examples for this step are shown in Fig. 3.8.

6.↔6. The excitation of a whole orbital, as illustrated in Fig. 3.9, satisfies the simple detailed balance equation

$$p_{eo} \frac{1}{\tilde{N}_p} \frac{1}{\tilde{N}_q} |W(C)| v(C \to C')$$
$$= p_{eo} \frac{1}{\tilde{N}_p} \frac{1}{\tilde{N}_q} |W(C')| v(C' \to C),$$

where $\tilde{N}_p$ and $\tilde{N}_q$ are the number of occupied and empty orbitals that are not affected by any kink in the current path.

The MC steps presented above show increased acceptance probabilities compared to the steps in [31] because they take into account some properties of the general structure of the Slater-Condon rules. Since all different cases for adding kinks are known, it can be avoided to propose particle excitations that would

result in a difference other than in two or four orbitals between two determinants. Yet, there are still many possible improvements for these steps. For example, in the MC step for changing two kinks, only the first kink is considered when choosing the excitation and the proposal is simply rejected if the second kink cannot be changed accordingly. For removing kinks, an advanced option would be to maintain a list of all removable kinks and kink pairs and randomly choose only from this list instead from all kinks in a configuration. Nevertheless, because of the exponential growth of the sign problem, the slightly larger parameter range for which simulations are feasible may not be worth the increased complexity of the algorithm. Despite being relatively easy to apply graphically, the correct implementation of only the current set of MC moves is already quite involved.

3.5 Heat Bath Method

The heat bath sampling method, also known as Gibbs sampling [64], is a special case of the Metropolis algorithm. Under certain circumstances, the heat bath method can be used to increase the efficiency of the MC algorithm by increasing the acceptance probability. This is achieved by choosing the proposal density of a single variable equal to its conditional probability. Of course, this is only possible, if that conditional probability is known and samples can be drawn from it efficiently.

The detailed balance equation (2.3) determines the acceptance probability for a MC update of a single component x_i of the configuration vector $\boldsymbol{x}$ when all remaining components $\boldsymbol{x}_i^- = (x_1, ..., x_{i-1}, x_{i+1}, ...)$ are fixed:

$$A(x_i \to x_i'|\boldsymbol{x}_i^-) = \min\left[1, \frac{Q(x_i' \to x_i|\boldsymbol{x}_i^-)p(x_i'|\boldsymbol{x}_i^-)}{Q(x_i \to x_i'|\boldsymbol{x}_i^-)p(x_i|\boldsymbol{x}_i^-)}\right].$$

As in Sec. 2.2, Q and p are the proposal and equilibrium probability density, respectively. Note, that this equation is just a special case of Eq. (2.4) for the update of a single component x_i. From the definition of the conditional probability, $p(x_i|\boldsymbol{x}_i^-) = p(\boldsymbol{x})/p(\boldsymbol{x}_i^-)$ and $p(x_i'|\boldsymbol{x}_i^-) = p(\boldsymbol{x}')/p(\boldsymbol{x}_i^-)$, it follows that the common denominator $p(\boldsymbol{x}_i^-)$ cancels in the ratio above.

If the conditional probability $p(x_i|\boldsymbol{x}_i^-)$ is known, one can choose the proposal density $Q(x_i \to x_i'|\boldsymbol{x}_i^-) = p(x_i'|\boldsymbol{x}_i^-)$ and the acceptance probability becomes unity, i.e. every MC update of the component x_i is always accepted. Usually, this means that the normalization factor $\int p(x_i, \boldsymbol{x}_i^-)\mathrm{d}x_i$ has to be calculated in every MC step and a random variable x_i' has to be drawn from the resulting distribution. This is efficient if $p(x_i|\boldsymbol{x}_i^-)$ is a simple analytic expression or if x_i is a discrete random

variable with only a small number of different values so that it can be summed efficiently.

In the case of the CPIMC algorithm, the time τ_l of a kink in the configuration $C = (K, \{n\}, s_1, \ldots, s_K, \tau_1, \ldots, \tau_l, \ldots, \tau_K)$ is distributed in the interval $[\tau_{l-1}, \tau_{l+1})$ according to,

$$p(\tau_l | C_l^-) = \frac{1}{Z_l} |W(C_l^-)| e^{-(D_{\{n^{(l-1)}\}} - D_{\{n^{(l)}\}})(\tau_l - \tau_{l-1})},$$

with the weights

$$W(C_l^-) = (-1)^K e^{-\sum_{\substack{i=0 \\ i \neq l-1, l}}^{K} D_{\{n^{(i)}\}}(\tau_{i+1} - \tau_i) - D_{\{n^{(l)}\}}(\tau_{l+1} - \tau_{l-1})} \prod_{i=1}^{K} q_{\{n^{(i)}\}\{n^{(i-1)}\}}(s_i),$$

and the normalization factor

$$Z_l = |W(C_l^-)| \int_{\tau_{l-1}}^{\tau_{l+1}} e^{-(D_{\{n^{(l-1)}\}} - D_{\{n^{(l)}\}})(\tau_l - \tau_{l-1})} d\tau_l$$

$$= |W(C_l^-)| \frac{1 - e^{-(D_{\{n^{(l-1)}\}} - D_{\{n^{(l)}\}})(\tau_{l+1} - \tau_{l-1})}}{D_{\{n^{(l-1)}\}} - D_{\{n^{(l)}\}}}.$$

As in Eq. (3.10), it is $\tau_0 = 0$ and $\tau_{K+1} = \beta$.

The inverse transform sampling method, see, e.g., [88], can be used to draw a time τ_l from this distribution. The cumulative distribution function of $p(\tau_l | C_l^-)$ is given by

$$F_l(x) = \int_{\tau_{l-1}}^{x} p(\tau | C_l^-) d\tau$$

$$= \frac{1}{Z_l} |W(C_l^-)| \frac{1 - e^{-(D_{\{n^{(l-1)}\}} - D_{\{n^{(l)}\}})(\tau - \tau_{l-1})}}{D_{\{n^{(l-1)}\}} - D_{\{n^{(l)}\}}}$$

for $x \in [\tau_{l-1}, \tau_{l+1}]$. Furthermore, it is $F_l(x) = 0$ for $x \leq \tau_{l-1}$ and $F_l(x) = 1$ for $x \geq \tau_{l+1}$ by definition. This function can easily be inverted yielding

$$F_l^{-1}(y) = \frac{-\ln\big(1 - y Z_l (D_{\{n^{(l-1)}\}} - D_{\{n^{(l)}\}}) / |W(C_l^-)|\big)}{(D_{\{n^{(l-1)}\}} - D_{\{n^{(l)}\}})} + \tau_{l-1}.$$

Given a uniformly distributed random number $y \in [0, 1)$, $F_l^{-1}(y)$ is a random number distributed according to $p(\tau_l | C_l^-)$ in the interval $[\tau_{l-1}, \tau_{l+1})$. Therefore, a MC move that updates $\tau_l \in [\tau_{l-1}, \tau_{l+1})$ can be implemented efficiently by using the heat bath idea resulting in an acceptance probability of $A(\tau_l \rightarrow \tau_l') = 1$.

Compared to the simple example above, the use of the heat bath sampling method for the choice of a new imaginary time τ in the MC steps described in

Sec. 3.4 is more complex for two reasons. First, a new kink is not only added between two adjacent kinks at times τ_i and τ_{i+1} but in a range $[\tau_{\min}, \tau_{\max})_\beta$ spanning multiple kinks. And second, besides τ, the MC moves update multiple variables of the configuration C. For those variables, implementing the heat bath method might be much more difficult or not feasible at all. The latter problem can easily be solved by realizing that the heat bath method is just a special case of the Metropolis algorithm. Thus, the uniform distribution $p(\tau) = 1/L_\beta(\tau_{\min}, \tau_{\max})$ in the detailed balance equations above is just replaced by the conditional probability $p(\tau|C)$ and the rest of the steps remains unchanged. The solution to the former problem will be the focus of the remainder of this section. Due to the difficulties involved in writing down the general case in a concise way, only the special case of adding a kink to the right of an existing kink will be considered in detail. All other cases can be derived analogously and only the end result will be presented.

The detailed balance equation for adding a single kink s_b right of an existing kink s_a, Eq. (3.17), is satisfied by the acceptance probability

$$A(C \to C') = \min\left[1, \frac{p_{\text{rk}}}{p_{\text{ak}}} \frac{2K}{K+1} \frac{N_o}{N_K} \frac{1}{p(\tau_b)} \left|\frac{W(C')}{W(C)}\right|\right].$$

Like in the description of the corresponding step **3.**, τ_a is the time of the chosen kink s_a and s_b is the new kink to be inserted in the range $[\tau_a, \tau_{\max})_\beta$. Because of the asymmetry between the proposal densities for adding and removing a kink, the proposal density for choosing the new imaginary time τ_b will be constructed to be proportional to the ratios of the configuration weights, i.e., $p(\tau_b) \propto |W(C')/W(C)|$.

In the configuration C, the range $[\tau_a, \tau_{\max})_\beta$ is composed of m adjacent intervals $[\tau_a, \tau_{a+1})_\beta, [\tau_{a+1}, \tau_{a+2})_\beta, \ldots, [\tau_{a+m-1}, \tau_{\max})_\beta$ determined by the the times of the kinks $s_{a+1}, \ldots, s_{a+m-1}$ between τ_a and $\tau_{\max}$. If the imaginary time of the new kink s_b is restricted to a single interval $[\tau_i, \tau_{i+1})_\beta$ with $i = a, \ldots, a + m - 1$, the probability density to choose the time $\tau = \tau_b$ in that interval is given by

$$p_i(\tau|C, s_b) = \frac{1}{Z_i}\left|\frac{W(C')}{W(C)}\right| = \frac{1}{Z_i}\gamma_i e^{-\eta_i(\tau-\tau_i)},$$

where η_i denotes the change of the diagonal energy due to the excitation induced by s_b, e.g., in the case $s_b = (p, q)$ one gets

$$\eta_i = h_{qq} - h_{pp} + \sum_{\substack{j=0 \\ j \neq p, q}}^{N_B} (w^-_{jqjq} - w^-_{jpjp})n_j.$$

Furthermore, the definition $\gamma_i = |\pi_i|e^{-\kappa_i}$ is used, with $\kappa_i = \sum_{j=a+1}^{i} \eta_{j-1}(\tau_j - \tau_{j-1})$ summing the corresponding diagonal energy changes in all previous intervals

and with π_i reflecting the change in the weights of the kinks. For example, if $s_b = (p, q)$, the ratio of the weights is given by

$$
\pi_i = \pi_a \frac{\prod_{\substack{j=a+1 \\ s_j=(k_j, l_j)}}^{i} \left(q_{\{n^{(j)}\}, \{n^{(j-1)}\}}(k_j, l_j) + w^-_{qk_jql_j} - w^-_{pk_jpl_j} \right)}{\prod_{\substack{j=a+1 \\ s_j=(k_j, l_j)}}^{i} q_{\{n^{(j)}\}, \{n^{(j-1)}\}}(k_j, l_j)}
$$

with

$$
\pi_a = \frac{q_{\{n^{(a)'}\}, \{n^{(a-1)}\}}(s_a')q_{\{n^{(i)}\}, \{n^{(i)'}\}}(s_b)}{q_{\{n^{(a)}\}, \{n^{(a-1)}\}}(s_a)}.
$$

Note that the products are over all kinks of type 2 that are between s_a and s_b and (k_j, l_j) denotes the one-particle excitation associated with kink s_j. The normalization constant can then be calculated as above:

$$
Z_i = \gamma_i \frac{1 - e^{-\eta_i(\tau_{i+1} - \tau_i)}}{\eta_i}.
$$

Likewise, the cumulative density function and its inverse can be found:

$$
F_i(x) = \frac{1}{Z_i} \gamma_i \frac{1 - e^{-\eta_i(x - \tau_i)}}{\eta_i}, \quad x \in [\tau_i, \tau_{i+1})_\beta,
$$

$$
F_i^{-1}(y) = \frac{-\ln(1 - yZ_i\eta_i/\gamma_i)}{\eta_i} + \tau_i, \quad y \in [0, 1).
$$

Note, that in the limit $\eta_i \to 0$, the inverse density function becomes $F^{-1}(y) \to y(\tau_{i+1} - \tau_i) + \tau_i$, i.e., the uniform distribution $p_i(\tau) = 1/(\tau_{i+1} - \tau_i)$ is recovered. As p_i is the probability to add s_b in the interval $[\tau_i, \tau_{i+1})_\beta$, it vanishes outside of that interval, i.e., $p_i(\tau) = 0$ for $\tau \notin [\tau_i, \tau_{i+1})_\beta$.

The proposal density for adding the kink s_b in the whole range $[\tau_a, \tau_{\max})_\beta$ is the piecewise continuous function

$$
p(\tau|C, s_b) = \frac{1}{Z_{\text{tot}}} \sum_{i=a}^{a+m-1} Z_i p_i(\tau|C, s_b),
$$

with the normalization constant $Z_{\text{tot}} = \sum_{i=a}^{a+m-1} Z_i$. The corresponding cumulative density function is continuous and for $x \in [\tau_a, \tau_{\max})_\beta$ given by

$$
F(x) = \int_{\tau_a}^{x} p(\tau|C, s_b)d\tau
$$

$$
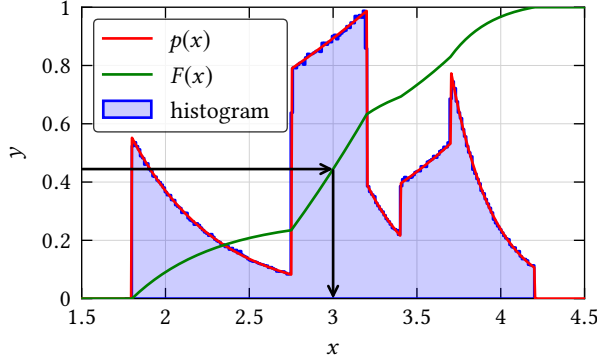= \frac{1}{Z_{\text{tot}}} \left(\sum_{j=a}^{\tau_{i_x}-1} Z_j + Z_{i_x} F_{i_x} \right),
$$

Figure 3.10: Inverse sampling method for a piecewise exponential distribution. The figure shows the probability density $p(x)$ (red) which is proportional to an exponential function in each interval and its cumulative distribution function $F(x)$ (green). The black arrows indicate how a random variable $y \in [0, 1)$ is mapped to a variable x that is distributed according to $p(x)$. The blue filled curve represents a histogram of 1×10^6 random samples.

where i_x is determined by the condition that $x \in [\tau_{i_x}, \tau_{i_x+1})_\beta$. A possible probability density $p(\tau)$ and its cumulative density function $F(x)$ are shown in Fig. 3.10. The inverse of the latter is given by

$$F^{-1}(y) = \frac{-\ln\left(1 - (yZ_{\text{tot}} - \sum_{j=a}^{i_y-1} Z_j)\, \eta_{i_y}/\gamma_{i_x}\right)}{\eta_{i_y}} + \tau_{i_y}, \qquad (3.18)$$

where i_y is determined by $\sum_{j=a}^{i_y-1} Z_j \leq yZ_{\text{tot}} < \sum_{j=a}^{i_y} Z_j$. Fig. 3.10 includes a histogram of one million random samples to show that $F^{-1}(y)$ is indeed distributed according to $p(\tau|C, s_b)$ for uniformly distributed $y \in [0, 1)$.

As γ_i has to be calculated for each of the m intervals in every MC step, it is important to implement this as efficient as possible. Naively using the equations above would require a lot of redundant computations. A better approach is to use the following recursion formulas for $i \in [a + 1, a + m - 1]$ and $m > 1$, which avoid the redundancy:

$$\pi_i = p_{i-1} \begin{cases} \frac{q_{\{n(i)\},\{n(i-1)\}}(k_i, l_i) + w^-_{qk_i ql_i} - w^-_{pk_i pl_i}}{q_{\{n(i)\},\{n(i-1)\}}(k_i, l_i)}, & s_i = (k_i, l_i) \\ 1, & \text{otherwise} \end{cases},$$

$$\kappa_i = \kappa_{i-1} + \eta_{i-1}(\tau_i - \tau_{i-1}),$$

where π_a given above and $\kappa_a = 0$. The full sum in the change of the pair interaction η_i has to be calculated only in the first interval, i.e., for η_a. In each of the following intervals, the immediately preceding kink s_i determines the change $\eta_i - \eta_{i-1}$, thus further reducing the computational time from $O(mN)$ to $O(m+N)$.

A similar recursion relation can be found for adding a kink s_b left of s_a:

$$\pi_i = p_{i+1} \begin{cases} \dfrac{q_{\{n^{(i+1)}\},\{n^{(i)}\}}(k_{i+1}, l_{i+1}) + w^-_{qk_{i+1}ql_{i+1}} - w^-_{pk_{i+1}pl_{i+1}}}{q_{\{n^{(i+1)}\},\{n^{(i)}\}}(k_{i+1}, l_{i+1})}, & s_{i+1} = (k_{i+1}, l_{i+1}) \\ 1, & \text{otherwise} \end{cases},$$

$$\kappa_i = \kappa_{i+1} + \eta_{i+1}(\tau_{i+2} - \tau_{i+1}),$$

for $i \in [a-m, a-2]$ with $\pi_{a-1} = \pi_a$ and $\kappa_{a-1} = \kappa_a$ Here, the algorithm steps through the intervals from the rightmost to the leftmost interval instead of the opposite direction as in the previous case. The inverse of the cumulative distribution function is given by:

$$F^{-1}(y) = \frac{\ln\left((yZ_{\text{tot}} - \sum_{j=i_y+1}^{a-1} Z_j)\eta_{i_y}/\gamma_{i_y} + e^{-\eta_{i_y}(\tau_{i_y+1} - \tau_{i_y})}\right)}{\eta_{i_y}} + \tau_{i_y+1}, \quad (3.19)$$

where the definitions of η_i, γ_i, and Z_i are the same as before[1], $Z_{\text{tot}} = \sum_{j=a-m}^{a-1} Z_j$, and i_y is determined by $\sum_{j=i_y+1}^{a-1} Z_j \le yZ_{\text{tot}} < \sum_{j=i_y}^{a-1} Z_j$. Because of $y = 1-(1-y)$, the expression above can be rewritten to:

$$F^{-1}(y) = \frac{\ln\left(1 - ((1-y)Z_{\text{tot}} - \sum_{j=i_y+1}^{a-1} Z_j)\,\eta_{i_y}/\gamma_{i_x}\right)}{\eta_{i_y}} + \tau_{i_y+1},$$

with $\sum_{j=i_y+1}^{a-1} Z_j < (1-y)Z_{\text{tot}} \le \sum_{j=i_y}^{a-1} Z_j$. This form more closely resembles the formula for adding the new kink right of s_a and avoids the calculation of an additional exponential function.

Moving an existing kink τ_a with imaginary time τ_a, i.e., choosing a new time $\tau'_a \in [\tau_{\min}, \tau_{\max})_\beta$, where $[\tau_{\min}, \tau_{\max})_\beta$ consists of m_l intervals left of τ_a and m_r intervals right of τ_a, can be accomplished by the following procedure: First one calculates $Z_l = \sum_{j=a-m_l}^{a-1} Z_j$ and $Z_r = \sum_{j=a}^{a+m_r-1} Z_j$ using the formulas above for the left and right case, respectively, with an appropriate definition of π_{a-1} and π_a. Then, one uses Eq. (3.19), if $yZ_{\text{tot}} < Z_l$, and Eq. (3.18), otherwise.

The heat bath method for the CPIMC algorithm results in a significant speedup if $\eta L_\beta(\tau_{\min}, \tau_{\max})$ is large. If it is small, this method can even be slower than a simple uniform sampling of the kink times, but only by a small factor. The

---

[1]In the actual implementation, η_i is replaced by $-\eta_i$ in these formulas for historical reasons.

reason is that $W(\tau')/W(\tau)$ has to be calculated for the maximum possible range $[\tau_{\min}, \tau_{\max})_\beta$, instead of only the range that actually changes, i.e., for the range $[\tau, \tau']_\beta$. Nevertheless, the heat bath idea is essential for greatly increasing the sampling efficiency of the Matsubara Green function, see Sec. 3.7.

In a recent paper [89], Holmes *et. al.* propose a heat bath method for choosing the excitation s with a probability proportional to the matrix element $q_{\{n\}, \{n'\}}(s)$. They report large efficiency gains for FCIQMC simulations of different molecular test systems. Compared to choosing uniformly between all possible one- and two-particle excitation as presented in the previous section, this idea could result in similar gains for CPIMC simulations. A potential drawback is that usually more than one matrix element determines the acceptance probability in the CPIMC method. Therefore, the usefulness of the idea for CPIMC simulations remains unclear.

3.6 Parallelization

The high computational demands of the numerical treatment of state-of-the-art quantum mechanical problems cannot be satisfied by serial calculations on a single processor core. Thus, the computational work has to be split among a large amount of processors. On modern high-performance-computing clusters, several thousand processors are available. While Moore's law, which states that the CPU transistor counts double approximately every two years, is still valid, the growth of single-thread performance slowed down significantly in recent years and programs are forced to use the increasing number of cores per processor unit even on desktop machines to benefit from current technological advances. These developments are especially beneficial for MC simulations as they belong to a class of so-called "embarrassingly parallel" computations.

For a problem of fixed size, the effective speedup S of a program that only partly benefits from additional resources can be described by Amdahl's law [90]

$$S = \frac{1}{(1-p) + p/P},\tag{3.20}$$

where P denotes the ideal speedup due the additional resources and p denotes the fraction of the program to which that speedup applies to. In the case of parallelization, P is the number of cores and p is the fraction of the work that can be done in parallel while $1 - p$ is the serial part of the program. Fig. 3.11 shows the speedup of a CPIMC simulation compared to the calculation of an observable with fixed accuracy on a single core. A fit of Amdahl's law to the data reveals
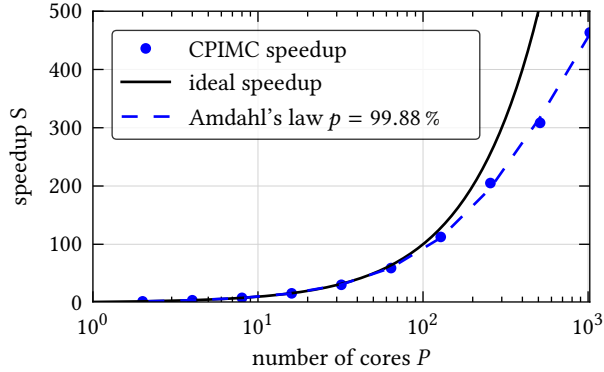
Figure 3.11: Speedup of CPIMC due to parallelization. The test system consists of $N = 33$ particles in the HEG at temperature $\theta = 4$ and $r_s = 2$. Blue dots show the measured speedup with increasing number of cores P for the fixed size problem of calculating the total energy with an relative statistical error of 6×10^{-5}. A fit of Amdahl's law (3.20) to the data is given by the blue dashed line. The ideal speedup is shown in black for comparison.

a parallel fraction of $p = 99.88\,\%$ for this test system of $N = 33$ particles in the HEG at temperature $\theta = 4$ and $r_s = 2$ with a relative accuracy of the total energy of 6×10^{-5}. In this particular test case, the speedup is approximately optimal for up to 100 cores. This nearly ideal behavior with respect to parallelization is inherent to MC methods. Ignoring possible autocorrelation, which is only important for the error estimation, the result of a simulation is just the average over *independent* samples. Thus, producing N_{MC} samples can be split among P processor cores each producing N_{MC}/P samples. The very small serial part consists mainly of the time needed for equilibration and the time to collect and average the results of the independent simulations.

It is important to ensure that the simulations are indeed independent. Because in the current implementation of the CPIMC method all simulations start from the same initial configuration, the equilibration time of each simulation should amount to the order of hundred autocorrelation times. This is the largest contribution to the serial part of the CPIMC method. Furthermore, the random number generator needs to provide independent streams of random numbers that do not overlap. The current implementation relies on the very long period of the Mersenne Twister [91]. The random number generator of each simulation is initialized with a random seed. Typically far less than 1×10^6 streams each

consisting of less than 1×10^{12} random numbers are used. Because of the period of $2^{19937} - 1 \approx 4.3 \times 10^{6001}$ it is very unlikely that an overlap occurs.

As the results of each simulation can be considered being independent, the calculation of statistical errors is even easier than in a serial MC simulation where the autocorrelation time needs to be accurately estimated[2]. However, a potential pitfall arises when averaging the results. Due to the sign problem, all results for physical observables are ratio estimators of the form $r = a/b$. The naive summation, $\bar{r}_1 = \frac{1}{M} \sum_{i=1}^{M} r_i$, will lead to a bias of the average $\bar{r}_1$ that only depends on the relative uncertainty of the denominator b_i, i.e., the average sign of a MC simulation, and not on the number of simulations M. Since the statistical error of $\bar{r}_1$ decreases with $1/\sqrt{M}$, the bias will become significant with increasing number of simulations. A better way to estimate $\bar{r}$ is to average nominator and denominator separately before calculating the ratio:

$$\bar{r}_2 = \frac{\bar{a}}{\bar{b}} = \frac{\frac{1}{M} \sum_{i=1}^{M} r_i b_i}{\frac{1}{M} \sum_{i=1}^{M} b_i}.$$

The bias of the ratio estimator $\bar{r}_2$ will decrease with $1/M$ and is thus asymptotically unbiased. For the estimation of the statistical error of $\bar{r}_2$, one has to take the cross-correlation between a and b into account:

$$\sigma_{\bar{r}_2}^2 = \bar{r}_2^2 \left[\left(\frac{\sigma_{\bar{a}}}{\bar{a}} \right)^2 + \left(\frac{\sigma_{\bar{b}}}{\bar{b}} \right)^2 - 2 \frac{\sigma_{\bar{a}\bar{b}}}{\bar{a}\bar{b}} \right], \tag{3.21}$$

with the sample standard deviation of the mean

$$\sigma_{\bar{x}}^2 = \frac{\sigma_x^2}{M} = \frac{1}{M(M-1)} \sum_{i=1}^{M} (x_i - \bar{x})^2 = \frac{\overline{x^2} - \bar{x}^2}{M-1},$$

and the sample covariance of the means

$$\sigma_{\bar{a}\bar{b}} = \frac{\sigma_{ab}}{M} = \frac{1}{M(M-1)} \sum_{i=1}^{M} (a_i - \bar{a})(b_i - \bar{b}) = \frac{\overline{ab} - \bar{a}\bar{b}}{M-1}.$$

For most applications, the bias of $\bar{r}_2$ can be neglected because the statistical uncertainty is much larger for any reasonable M. However, if several similar values $\bar{r}_2(x)$ for different parameters x are combined, e.g., when integrating a smooth function $r(x)$ using the points $\bar{r}_2(x)$, the bias of the aggregated value may

---

[2]Calculating the autocorrelation time may still be useful to estimate the necessary equilibration time and the most efficient cycle.
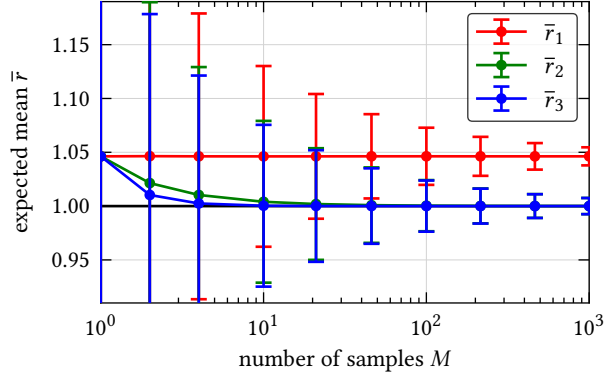
Figure 3.12: Comparison of ratio estimators. The points denote the expected
mean of averaging over M independent samples $r = a/b$ according
to three different estimators $\bar{r}_1$, $\bar{r}_2$, and $\bar{r}_3$. The error bars show the
corresponding expected standard deviation of the mean $\sigma_{\bar{r}}$. The true
mean is $\bar{r} = \bar{a}/\bar{b} = 1$, shown by the black line. The relative error
of a and b of a single sample is 20 %. The naive estimator $\bar{r}_1$ has a
constant bias independent of the number of samples, while the bias
of the standard and Beale's estimator decrease like $1/M$ and $1/M^2$,
respectively. Note that the standard deviation is not a good estimate
for the true error in such an extreme case because the distribution of
r is heavily skewed.

still become significant compared to its statistical error. In that case, one can
either increase the number of samples of the serial simulations to reduce the
relative error of the values b_i at the cost of a longer runtime or use a higher order
estimator, e.g., Beale's estimator [92–94]:

$$\bar{r}_3 = \bar{r}_2 \frac{1 + \dfrac{1}{M} \dfrac{\sigma_{ab}}{\overline{ab}}}{1 + \dfrac{1}{M} \dfrac{\sigma_b^2}{\bar{b}^2}}.$$

The bias of this second order estimator decreases with $1/M^2$. Up to the order of
$1/M$, the statistical error of $\bar{r}_3$ equals that of $\bar{r}_2$, i.e., $\sigma_{\bar{r}_2} \simeq \sigma_{\bar{r}_3}$. As the covariance
has to be already calculated for a reliable error estimation, Beale's estimator
reduces the bias with little overhead. A comparison of the three ratio estimators
$\bar{r}_1$, $\bar{r}_2$, and $\bar{r}_3$ is shown in Fig. 3.12.

3.7 The Worm Algorithm

Proposed for lattice models by Prokof'ev *et. al.* [40] and adapted to continuous-space PIMC simulation by Boninsegni *et. al.* [20], the worm algorithm is one of the most successful MC methods for the simulation of quantum many body problems at finite temperatures. It is widely used for classical and quantum lattice models [95–97], the Anderson impurity model [36, 98, 99], the PIMC method in coordinate representation [74], and Diagrammatic Monte Carlo (DiagMC) simulations of polaron models [35]. Its basic idea is to extend the configuration space of the partition function consisting of closed world lines by introducing two discontinuities called the head and the tail of the worm. The advantage of the worm algorithm is twofold. First, it enables the direct sampling of the Matsubara Green Function (MGF), which is generally difficult to obtain from expansions of the partition function but allows for the calculation of dynamic quantities from equilibrium simulations by a reconstruction of the one-particle spectral function [41]. And second, it greatly increases the sampling efficiency in cases of critical slow down in the presence of metastable states, like in the vicinity of phase transitions. Since MC updates do not have to resemble physical dynamics, *local* updates are performed in the extended configuration space of the MGF resulting in *non-local* updates in the configuration space of the partition function. Several variations that, e.g., restrict the movement of the worm in imaginary time or introduce multiple worms, have been developed [30, 100, 101].

Based on the observed enhancements of other methods by the introduction of the worm algorithm, the first goal of this work was its adaption to the CPIMC method to overcome low acceptance probabilities preventing the simulation of larger systems [31]. Chronologically, the steps in the next section have been implemented before the development of the closed-path updates presented in Sec. 3.4. However, even after the implementation of a second worm, it was not possible to find MC steps that are ergodic due to the sparse structure of the Hamiltonian. Thus, the local updates of Sec. 3.4 are necessary in addition to the worm moves to ensure ergodicity of the algorithm. Additionally, the strong dependence of the number of kinks on the particle number causes a runaway effect for calculations with varying particle numbers, cf. Sec. 5.4. These problems could not be overcome during this work, although ideas exist to partly mitigate them. Therefore, the usefulness of the worm algorithm for the CPIMC method is reduced to the sampling of the MGF and related quantities and it is not used if only canonical expectation values are of interest.

In the grand-canonical ensemble, it is convenient to switch to a modified interaction picture with respect to $\hat{D}' = \hat{D} - \mu\hat{N}$, where $\hat{D}$ is the diagonal part of

the Hamiltonian $\hat{H} = \hat{D} + \hat{Y}$ in the chosen basis as usual, $\hat{N}$ is the total particle number operator, Eq. (3.6), and μ the chemical potential. As it is always clear from the context which ensemble is used, the prime will be dropped in the following. In this modified interaction picture, the MGF can be written as [87]

$$\mathcal{G}_{ij}(\tau, \tau') = \frac{1}{Z} \operatorname{Tr}\left\{ e^{-\beta\hat{D}} \hat{T}\left[e^{-\int_0^\beta \hat{Y}(\bar{\tau})\mathrm{d}\bar{\tau}} \hat{a}_i(\tau)\hat{a}_j^\dagger(\tau') \right] \right\},$$

with the grand-canonical partition function, cf. Eq.(2.2),

$$Z(\mu, \beta, V) = \operatorname{Tr}\left\{ e^{-\beta\hat{D}} \hat{T} e^{-\int_0^\beta \hat{Y}(\bar{\tau})\mathrm{d}\bar{\tau}} \right\}.$$

Note that the trace has to be evaluated in the Fock space. Following the derivation of the CPIMC expression for the canonical partition sum, Eq. (3.9), a similar formula can be found for the so-called Green function sum

$$
\begin{aligned}
G &:= \sum_{i,j=1}^\infty \int_0^\beta \mathrm{d}\tau_h \int_0^\beta \mathrm{d}\tau_t\, Z\mathcal{G}_{ij}(\tau_h, \tau_t) \\
&= \sum_{K=0}^\infty \sum_{\{n\}} \sum_{i,j} \sum_{s_1} \sum_{s_2} \cdots \sum_{s_K} \int_0^\beta \mathrm{d}\tau_h \int_0^\beta \mathrm{d}\tau_t \int' \mathrm{d}^K\tau \\
&\quad (-1)^K \langle\{n\}| e^{-\beta\hat{D}} \hat{T}\left[\hat{q}(s_K, \tau_K) \cdots \hat{q}(s_1, \tau_1)\hat{a}_i(\tau_h)\hat{a}_j^\dagger(\tau_t) \right]|\{n\}\rangle,
\end{aligned}
$$

where $\hat{q}(s, \tau) = e^{\tau\hat{D}} \hat{q}(s) e^{-\tau\hat{D}}$ are the kink operators defined by Eq. (3.8) and the sum over $\{n\}$ includes Slater-determinants with varying particle number. After inserting $K + 1$ identity operators and explicitly applying the time ordering operator, one arrives at

$$
G = \sum_{K=0}^\infty \sum_{\{n\}} \sum_{i,j} \sum_{s_1} \sum_{s_2} \cdots \sum_{s_{K-1}} \int_0^\beta \mathrm{d}\tau_h \int_0^\beta \mathrm{d}\tau_t \int' \mathrm{d}^K\tau \sum_{\sigma \in S_{K+2}} \bar{\theta}(\tau_{\sigma(1)}, \ldots, \tau_{\sigma(K+2)})
$$

$$
(-1)^{K+\theta(\tau_t - \tau_h)} e^{-\sum_{k=0}^{K+2} D_{\{n(\sigma(k))\}}(\tau_{\sigma(k+1)} - \tau_{\sigma(k)})} \prod_{k=1}^{K+2} \bar{q}_{\{n(\sigma(k))\},\{n(\sigma(k-1))\}}(s_{\sigma(k)}),
$$

where $\sigma \in S_{K+2}$ is a permutation of $K + 2$ indices, $\sigma(0) = 0$, $\sigma(K + 3) = K + 3$, $\tau_0 = 0$, $\tau_{K+1} = \tau_h$, $\tau_{K+2} = \tau_t$, $\tau_{K+3} = \beta$, and the sum over s_K was dropped because one kink is uniquely determined by all other kinks under the condition $\{n^{(0)}\} = \{n^{(\sigma(K+2))}\} = \{n\}$, i.e., β-periodicity. Because of the definitions $s_{K+1} = i$ and $s_{K+2} = j$, the variable s_k denotes a multi index $s \in \mathbb{N}^2 \cup \mathbb{N}^4$ only in the case
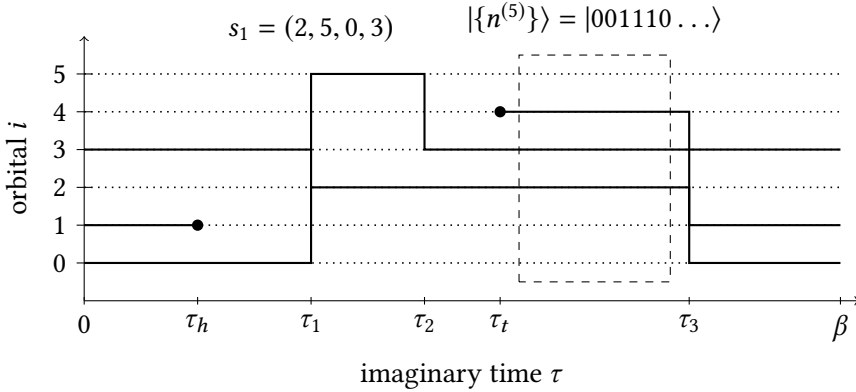
Figure 3.13: Open path in the worm algorithm with $K = 3$ kinks of type 2 and type 4 at τ_1, τ_2, and τ_3. In addition to a closed path as show in Fig. 3.2 an open path contains a single pair of worm operators at τ_h and τ_t, respectively. The particle number now depends on the imaginary time and varies between $N = 2$ and $N = 3$. Note, that due to the time ordering permutation of the kinks it is $\{n^{(\sigma(4))}\} = \{n^{(5)}\}$.

$k \in \{1, \ldots, K\}$. To ensure the correct time ordering the following notations have been introduced:

$$\hat{\bar{q}}(s_{\sigma(k)}, \tau_{\sigma(k)}) := \begin{cases} \hat{a}_i(\tau_h), & \sigma(k) = K + 1 \\ \hat{a}_j^\dagger(\tau_t), & \sigma(k) = K + 2 \\ \hat{q}(s_{\sigma(k)}, \tau_{\sigma(k)}), & \sigma(k) \in \{1, \ldots, K\} \end{cases},$$

which defines the generalized kink operator $\hat{\bar{q}}(s)$, and

$$\bar{\theta}(\tau_{\sigma(1)}, \ldots, \tau_{\sigma(K+1)}) := \theta(\tau_{\sigma(2)} - \tau_{\sigma(1)}) \cdots \theta(\tau_{\sigma(K+2)} - \tau_{\sigma(K+1)}),$$

which is non-zero only if the permutation σ corresponds to the correct time ordering of the integration variables $\tau_1, \ldots, \tau_K, \tau_h, \tau, t$, i.e., for any given set of imaginary times $\bar{\theta}$ picks exactly one permutation of the kink indices.

Due to the time ordering, the Green function sum looks more complicated than the partition sum, Eq. (3.9), but it can be interpreted in the same way with the addition of a single pair of the two new kink operators $\hat{a}_i(\tau_h)$ and $\hat{a}_j^\dagger(\tau_t)$, called the head and tail of the worm, respectively. A visualization of a

typical contribution to G is given in Fig. 3.13. In contrast to the partition sum, where all trajectories are closed loops, the contributions to the Green function sum are β-periodic *open paths*. Such a configuration is uniquely defined by $C_G = (K, \{n\}, i, j, s_1, \ldots, s_K, \tau_h, \tau_t, \tau_1, \tau_K)$ and its weight is given by

$$W_G(C_G) = \sum_{\sigma \in S_{K+2}} \bar{\theta}(\tau_{\sigma(1)}, \ldots, \tau_{\sigma(K+2)})$$

$$(-1)^{K+\theta(\tau_t - \tau_h)} e^{-\sum_{i=0}^{K+2} D_{\{n(\sigma(i))\}}(\tau_{\sigma(i+1)} - \tau_{\sigma(i)})} \prod_{i=1}^{K+2} \bar{q}_{\{n(\sigma(i))\}, \{n(\sigma(i-1))\}}(s_{\sigma(i)})$$

where just one term of the sum is non-zero. The worm's contribution to the weight is only a sign change. The sign depends on the order of the head and the tail as well as the total number of particles in orbitals below the orbitals i and j at the imaginary time τ_h and τ_t, respectively.

3.8 Green Function Estimator

Finding an estimator for the MGF that has a low variance and is free of discretization errors is a difficult task. The trivial estimator that just counts the number of times that the head and the tail of the worm are in the bin associated with the point $\mathcal{G}_{ij}(\tau_h, \tau_t)$ fails at both demands. Increasing the bin size can improve the statistical error but comes at the cost of a large discretization error. On the other hand, a small bin size results in only a few counts per bin and therefore has a large variance. An estimator without discretization error proposed by Boninsegni *et. al.* [20] works well for the PIMC method in coordinate representation but exhibits large statistical fluctuations in the case of CPIMC, rendering it unusable for interesting system sizes. A significantly better estimator, developed together with Simon Groth, will be presented in the following. A detailed comparison of different estimators for the MGF can be found in his master thesis [59].

Comparing the definition of the Green function sum, Eq. 3.7, and its final expression, Eq. 3.7, one finds

$$\mathcal{G}_{ij}(\tau_h, \tau_t) = \frac{1}{Z} \oint_C W_G(C, i, j, \tau_h, \tau_t),$$

with the previously used abbreviation for the summation and integration over all closed paths $C = (K, \{n\}, s_1, \ldots, s_K, \tau_1, \ldots, \tau_K)$. Because the MGF is homoge-

neous in imaginary time, $\mathcal{G}_{ij}(\tau, \tau') = \mathcal{G}_{ij}(\tau - \tau', 0)$, one can write

$$\mathcal{G}_{pq}(\tau, 0) = \mathcal{G}_{pq}(\tau + \tau_t, \tau_t) = \frac{1}{\beta} \sum_{i,j} \int_0^\beta d\tau_t \delta_{i,p} \delta_{j,q} \mathcal{G}_{ij}(\tau + \tau_t, \tau_t)$$

$$= \frac{1}{Z} \frac{1}{\beta} \oint_C \sum_{i,j} \int_0^\beta d\tau_t \delta_{i,p} \delta_{j,q} W_G(C, i, j, \tau + \tau_t, \tau_t).$$

Dropping all arguments of W_G apart from the time of the head τ_h for clarity, the following identities hold:

$$W_G(\tau + \tau_t) = \frac{\int_0^{\prime\beta} d\tau_h |W_G(\tau_h)| W_G(\tau_h) / W_G(\tau_h)}{\int_0^{\prime\beta} d\tau' |W_G(\tau')|} W_G(\tau + \tau_t)$$

$$= \int_0^{\prime\beta} d\tau_h \frac{W_G(\tau + \tau_t) / W_G(\tau_h)}{\int_0^{\prime\beta} d\tau' |W_G(\tau') / W_G(\tau_h)|} W_G(\tau_h).$$

The prime on the integral denotes a restriction of the integration to the region where the weight of the current path is non-zero, $W_G(\tau_h) \neq 0$. Defining the ratio of the weights

$$R(\tau_h \to \tau + \tau_t) = \frac{W_G(C, i, j, \tau + \tau_t, \tau_t)}{W_G(C, i, j, \tau_h, \tau_t)}$$

and multiplying by G/G, one arrives at

$$\mathcal{G}_{pq}(\tau, 0) = \frac{G}{Z} \frac{1}{G} \frac{1}{\beta} \oint_{C_G}' \left(\delta_{i,p} \delta_{j,q} \frac{R(\tau_h \to \tau + \tau_t)}{\int_0^{\prime\beta} d\tau' |R(\tau_h \to \tau')|} \right) W_G(C_G). \qquad (3.22)$$

As the MC algorithm will only sample configurations with non-vanishing weight, the restriction of the integration to $W_G \neq 0$ will be taken into account implicitly and the prime of the outer integral can be dropped. In a given path, the integral in the denominator is at least restricted to the interval $[\tau_l, \tau_r]_\beta$ defined by the times of the nearest kinks left and right of τ_h that affect the orbital i. Depending on $W_G(\tau)$, the integration region may be even smaller. As will be shown in the next section, the ratio R also appears in the acceptance probability for moving the head of the worm in imaginary time. Therefore, the estimator for $\mathcal{G}_{pq}(\tau, 0)$ is related to the normalized probability for moving the head from its position τ_h in the configuration C_G to the new position $\tau + \tau_t$ while all other components of C_G are fixed. Therefore, one can reuse the implementation of the heat bath sampling for the calculation of estimator, see Seq. 3.5.

The evaluation of Eq. (3.22) requires the calculation of the ratio G/Z. While, in general, the Metropolis algorithm cannot calculate the partition functions Z or G separately, their ratio has a rather simple estimator if one samples both the configurations C_G, which contribute to the MGF, and the configurations C, which contribute to thermodynamic expectation values, in a single simulation by constructing a Markov chain in the combined configuration space of the quantity

$$Z_{\text{tot}} = Z + \omega G = \sum_C W(C) + \sum_{C_G} \omega W_G(C_G) =: \sum_{C_{\text{tot}}} W_{\text{tot}}(C_{\text{tot}}).$$

Here, the sum over C_{tot} combines the sums over C and C_G and $W_{\text{tot}}(C_{\text{tot}})$ is either $W(C)$ or $\omega W_G(C_G)$ depending on C_{tot}. The parameter ω can be freely chosen to balance the relative weight between Z and G and thus the time spent by the algorithm in the respective configuration spaces. Using Eq. (2.5) and the relations

$$Z = \sum_{C_{\text{tot}}} \delta_Z(C_{\text{tot}}) W_{\text{tot}}(C_{\text{tot}}),$$

$$G = \sum_{C_{\text{tot}}} \frac{1}{\omega} \delta_G(C_{\text{tot}}) W_{\text{tot}}(C_{\text{tot}}),$$

where $\delta_{Z/G}(C_{\text{tot}}) = 1$ if C_{tot} is a closed (open) path and $\delta_{Z/G}(C_{\text{tot}}) = 0$ otherwise, one gets

$$\frac{G}{Z} = \frac{G'}{Z'} \frac{\langle s \rangle'_G}{\langle s \rangle'_Z} = \frac{1}{\omega} \frac{\langle \delta_G \rangle'_{\text{tot}}}{\langle \delta_Z \rangle'_{\text{tot}}} \frac{\langle s \rangle'_G}{\langle s \rangle'_Z}.$$

In other words, one just has to count the numbers of samples in the respective configuration spaces and measure the corresponding average signs to get an accurate estimate for the desired ratio. The MC steps that sample configurations in G and switch between the configuration space of Z and G are presented in the next section.

The estimator for the MGF still has a relatively large variance because the denominator $\int_0^{'\beta} d\tau' |R(\tau_h \to \tau')|$ can be small if the interval between the kinks adjacent to the head on orbital i is small. This leads to occasionally very large values and the distribution of the samples is heavily skewed. Nevertheless, as shown in [59], it performs significantly better than other estimators and can yield accurate results if the number of kinks is not too large. In particular, Eq. (3.22) is an improvement over the closed path estimator for the off-diagonal part of

the reduced one-particle density matrix, Eq. (3.11). The MGF is connected to the one-particle density matrix by

$$d_{pq} = \lim_{\tau \to 0^-} \mathscr{G}_{qp}(\tau, 0). \tag{3.23}$$

The limit has a defined particle number and thus allows for the sampling of the canonical density matrix in addition to the grand-canonical expectation value.

3.9 Monte Carlo Steps for the Worm Algorithm

For the evaluation of the expectation value of the MGF, the MC algorithm has to sample configurations in the combined configuration space of Z and G. In addition to all β-periodic closed paths constructed from the elements depicted in Sec. 3.4, the algorithm must be able to create all possible open paths containing a so-called worm, i.e., a single pair of a head, —•, and a tail, •—. In contrast to the worm algorithm in lattice systems or in the coordinate representation, no ergodic set of updates involving only the worm could be found for the systems governed by the full Coulomb pair interaction studied in this work. Therefore, the steps presented in the following have to be used in addition to the steps from Sec. 3.4, mitigating the advantage of effective non-local sampling observed in simpler systems. There are four different MC steps involving the head of the worm, i.e, being applicable to open paths only. These steps move the head in imaginary time or between orbitals, thereby creating new kinks or changing existing ones. Due to the symmetry of the MGF, one can restrict these additional MC updates to the head only and does not need to move the tail. Two MC steps add or remove the worm by adding and removing a line segment, i.e., they switch between the configuration spaces of open and closed paths. The steps from Sec. 3.4 are applied to closed and open paths as well. In the following, a detailed description of the six additional steps is presented:

1. Add the worm
Choose a random orbital p and a time τ_t uniformly distributed in $(0, \beta)$. There are two possibilities:

A) At time τ_t the orbital p is occupied

i. Find the time $\tau_{\min}$ of the first kink on the left of τ_t affecting the orbital p if there exists such a kink.

ii. If $\tau_{\min}$ exists, choose a time τ_h in the interval $(\tau_{\min}, \tau_t)_\beta$ at random. Otherwise, choose $\tau_h \in (0, \beta)$. Head and tail of the worm are given

Figure 3.14: Illustration of the MC steps for adding and removing a worm, i.e., a pair of a head operator $\hat{a}_i(\tau_h)$ and a tail operator $\hat{a}_j^\dagger(\tau_t)$. Left: a worm is added into an empty orbital or removed according to case b) of both branches of step **2.**, respectively. Right: a worm is inserted into an occupied orbital or removed according to case a) of both branches of step **2.**, respectively. The sketches depict only a relevant subset of the path. For details see also Fig. 3.3.

by $\hat{a}_p(\tau_h)$ and $\hat{a}_p^\dagger(\tau_t)$, respectively. This corresponds to removing a particle in the interval $[\tau_h, \tau_t]_\beta$.

B) At time τ_t the orbital p is empty

 i. If there are kinks affecting orbital p, find the time $\tau_{\max}$ of first one on the right of τ_t.

 ii. Randomly choose a time τ_h in the interval $(\tau_t, \tau_{\max})_\beta$, or in $(0, \beta)$ if there are no kinks affecting the orbital p. Head and tail of the worm are given by $\hat{a}_p(\tau_h)$ and $\hat{a}_q^\dagger(\tau_t)$, respectively. This corresponds to adding a particle in the interval $[\tau_t, \tau_h]_\beta$.

2. Remove the worm

Head and tail of the worm are given by $\hat{a}_p(\tau_h)$ and $\hat{a}_q^\dagger(\tau_t)$, respectively. Reject the step if $p \neq q$. Otherwise, there are two possibilities:

A) There are no other kinks affecting the orbital p

 i. Choose one of the following two cases at random:

  a) Remove the worm while occupying the interval $[\tau_h, \tau_t]_\beta$.

  b) Deleting the worm removes a particle in the interval $[\tau_t, \tau_h]_\beta$.

B) There is at least one additional kink affecting the orbital p

 i. There are two possibilities:

  a) The head of the worm is the first kink on the left of the tail on the orbital p. Removing both adds a particle in the interval $[\tau_h, \tau_t]_\beta$.

  b) The head is the first kink affecting the orbital p right of the tail of the worm. Delete the worm while removing a particle in the interval $[\tau_t, \tau_h]_\beta$.
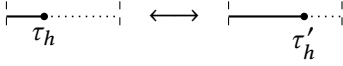
Figure 3.15: Move the head of the worm in imaginary time. This simple MC step is essential for the sampling of the MGF.

3. Move the head in time

The head is given by $\hat{a}_p(\tau_h)$.

    i. On the orbital p, find the time of the first kink on the left of τ_h, $\tau_{\min}$, and the time of the first kink on the right, $\tau_{\max}$. If there is only one other kink affecting orbital p, it is $\tau_{\min} = \tau_{\max}$.

    ii. Randomly find a new time for the head, $\tau'_h \neq \tau_h$, in the interval $(\tau_{\min}, \tau_{\max})_\beta$, or in $(0, \beta)$ if $\tau_{\min} = \tau_{\max}$.

    iii. Move the head of the worm from τ_h to τ'_h. Distinguish the following two cases:

        a) If τ'_h is left of τ_h in the β-periodic time ordering, moving the head corresponds to removing a particle in the interval $[\tau'_h, \tau_h]_\beta$.

        b) If τ'_h is on the right of τ_h, a particle is added in $[\tau_h, \tau'_h]_\beta$ while moving the head.

4. Add a kink

The head is given by $\hat{a}_p(\tau_h)$. Choose an orbital $q \neq p$ at random. There are two possibilities:

  A) The orbital q is occupied at time τ_h

    i. Find the time $\tau_{\max}$ of the first kink right of τ_a affecting either orbital p or q.

    ii. Choose a random time τ_a in the interval $[\tau_h, \tau_{\max})_\beta$. The new kink is given by $s_a = (q, p)$ at τ_a.

    iii. Move the head from orbital p to orbital q, thereby exciting a particle from q to p in the interval $[\tau_h, \tau_{\max}]_\beta$.

  B) At time τ_h, there is no particle on orbital q

    i. From all kinks left of τ_h that affect the orbitals p or q, find the closest one with time $\tau_{\min}$.

    ii. Randomly choose a time τ_a in the interval $(\tau_{\min}, \tau_h]_\beta$. The new kink is given by $s_a = (p, q)$.

    iii. Move the head from orbital p to orbital q. In the interval $[\tau_{\min}, \tau_h]_\beta$, a particle is excited from p to q.

Figure 3.16: Add and remove a kink by moving the head of the worm between orbitals. Left: the head is moved to an empty orbital and thus a kink is added or removed to the left of τ_h. Right: moving the head to an occupied orbital adds or removes a kink to the right of τ_h. In both cases, only type 2 kinks can be added or removed.

5. Remove a kink

The head is given by $\hat{a}_p(\tau_h)$. Choose one of the following two cases at random:

A) Remove a kink left of τ_h

    i. Find the kink s_a that is closest to the left of τ_h of all kinks affecting orbital p.

    ii. Reject the step if $s_a \neq (q, p)$.

    iii. Remove the kink s_a and change the orbital of the head from p to q, thereby exciting a particle from p to q in the interval $[\tau_a, \tau_h]_\beta$.

B) Remove a kink right of τ_h

    i. Find the nearest kink right of τ_h that affects orbital p.

    ii. If the kink is not of type 2, i.e., not given by $s_a = (p, q)$, reject the move.

    iii. Delete the kink s_a and move the head from orbital p to orbital q. This corresponds to a one-particle excitation from q to p in the interval $[\tau_h, \tau_a]_\beta$.

6. Change a kink

The head is given by $\hat{a}_p(\tau_h)$. Choose one of the following two cases at random:

A) Change a kink on the left of τ_h

    i. In the time ordering of all kinks, let s_a be the kink immediately preceding the head of the worm.

    ii. For the choice of a random orbital q, consider the following cases:

        a) If $s_a = (i, j, k, l)$ and $p \neq [i, j]$, choose $q \in [k, l]$.

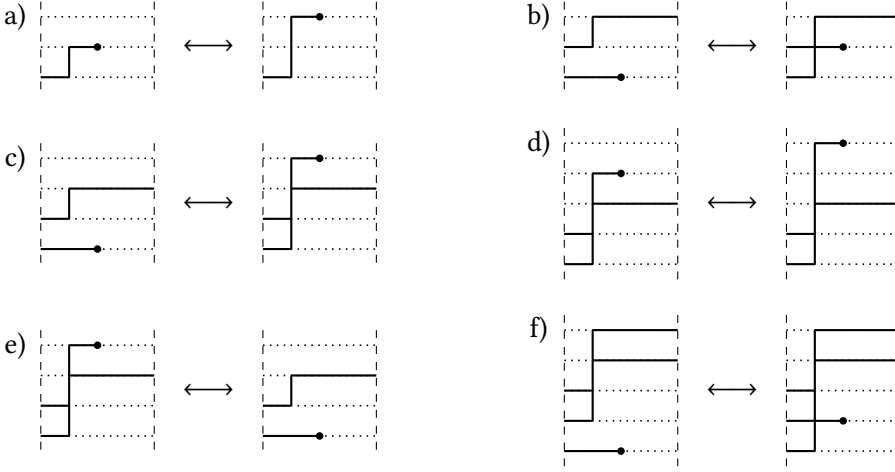   b) In all other cases, choose $q \neq p$ at random from the occupied orbitals at τ_h.

Figure 3.17: Change a kink by moving the head of the worm between empty orbitals, case A). The labels a) to f) correspond to the cases of point iv. Cases c) and e) allow for the insertion of a type 4 kink into the path, but rely on the existence of a matching type 2 kink.

    iii. If $s_a = (p, q)$, reject the step. The orbital of the head changes from p to q otherwise.

    iv. The kink s_a changes in one of the following ways:

        a) $s_a = (p, j) \implies s'_a = (q, j)$

        b) $s_a = (i, q) \implies s'_a = (i, p)$

        c) $s_a = (i, j) \implies s'_a = (i, q, j, p)$

        d) $s_a = (p, j, k, l) \implies s'_a = (q, j, k, l)$ or
           $s_a = (i, p, k, l) \implies s'_a = (i, q, k, l)$

        e) $s_a \in [(p, j, q, l), (j, p, q, l), (p, j, l, q), (j, p, q, l)] \implies s'_a = (j, l)$

        f) $s_a = (i, j, q, l) \implies s'_a = (i, j, p, l)$ or
           $s_a = (i, j, k, q) \implies s'_a = (i, j, k, p)$

B) Change a kink on the right of τ_h

    i. The kink immediately following τ_h in the β-periodic time ordering of all kinks is s_a.

    ii. Choose an orbital q at random depending on the following cases:

        a) If $s_a = (i, j, k, l)$ and $p \neq [i, j]$, choose $q \in [k, l]$.
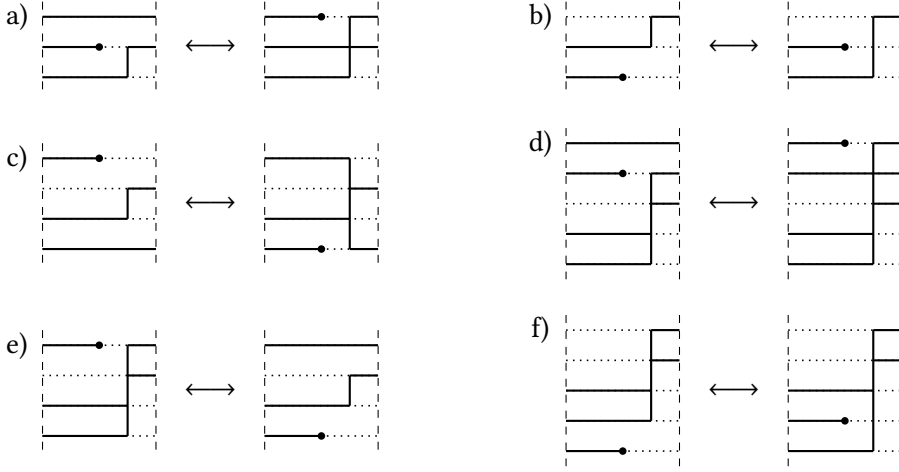
Figure 3.18: Change a kink by moving the head of the worm between occupied orbitals, case B). The labels a) to f) correspond to the cases of point iv. of branch A). Cases c) and e) allow for the insertion of a type 4 kink into the path, but rely on the existence of a matching type 2 kink.

    b) Otherwise, choose $q \neq p$ randomly from all orbitals that are occupied at τ_h.

  iii. Reject the step if $s_a = (p, q)$. Otherwise, move the head from orbital p to orbital q.

  iv. The kink s_a changes according to the rules in point iv. of case A).

The steps **1.** to **6.** are chosen with the probabilities p_{aw} and p_{rw} for adding and removing the worm, p_{mht} for moving the head in time, and p_{hak}, p_{hrk}, p_{hck} for adding, removing, and changing a kink by moving the head in orbital space. In a closed-path configuration, these probabilities are set to zero, except for p_{aw}, which is set to zero in open-path configurations instead. Usually, the non-zero probabilities are set equally to the probabilities of the steps that do not involve the worm. The free parameter ω, which balances the relative weight of open and closed paths, is determined during the equilibration time so that the acceptance ratios for adding and removing the worm are about the same. This choice was found to be sufficiently efficient for all parameters, while strongly biased ratios can lead to wrong results. An open-path configuration is denoted by C_G and its weights by $W_G(C_G)$, while closed-path configuration are described by C and

$W(C)$ as in Sec. 3.4. All other notations used in the following detailed balance equations are defined there as well.

1.↔2. The detailed balance for switching between the configuration spaces of closed and open paths, i.e., between a configuration C without a worm and a configuration $C_G = (C, p, q, \tau_h, \tau_t)$ containing a worm, as shown in Fig. 3.14, is given by:

$$p_{\text{aw}} \frac{1}{\beta} \frac{1}{N_B} p(\tau_h) |W(C)| v(C \rightarrow C_G)$$
$$= p_{\text{rw}} \frac{1}{d} \omega |W_G(C_G)| v(C_G \rightarrow C),$$

where $d = 2$ if the orbital p is not affected by any kink apart from the worm. Otherwise, it is $d = 1$.

3.↔3. This step moves the head of the worm in imaginary time as shown in Fig. 3.15. Moving the head between τ_h in the configuration C_G and τ'_h in C'_G fulfills the detailed balance equation

$$p_{\text{mht}} p(\tau'_h) W_G(C_G) v(C_G \rightarrow C'_G)$$
$$= p_{\text{mht}} p(\tau_h) W_G(C'_G) v(C'_G \rightarrow C_G)$$

Using the heat bath idea and thus setting the distribution of the imaginary time τ'_h to $p(\tau'_h) = |R(\tau_h \rightarrow \tau'_h)| / \int |R(\tau_h \rightarrow \tau'_h)| d\tau'_h$ with the ratio $R(\tau_h \rightarrow \tau'_h) = W_G(C, i, j, \tau'_h, \tau_t) / W_G(C, i, j, \tau_h, \tau_t)$, results in the acceptance probability $A(C_G \rightarrow C'_G) = 1$, i.e., the step is never rejected.

4.↔5. These steps, which are illustrated in Fig. 3.16, move the head of the worm from one orbital to another by changing a configuration C_G with $K + 2$ kinks to a configuration C'_G with $K + 3$ kinks and vice versa. The acceptance probabilities are determined by:

$$p_{\text{hak}} \frac{1}{N_B - 1} p(\tau_a) |W_G|(C_G) v(C_G \rightarrow C'_G)$$
$$= p_{\text{hrk}} \frac{1}{2} |W_G(C'_G)| v(C'_G \rightarrow C_G).$$

6.↔6. Moving the head of the worm in orbital space while changing an existing kink switches between the configurations C_G and C'_G as illustrated in

Figs. 3.17 and 3.18. This leads to the following detailed balance equation:

$$p_{\text{hck}} \frac{1}{2} \frac{1}{N_q} W_G(C_G) v(C_G \to C_G')$$
$$= p_{\text{hck}} \frac{1}{2} \frac{1}{N_q} W_G(C_G') v(C_G' \to C_G),$$

where $N_q = N_e$ in case A)ii.b) and $N_q = N_f$ in case B)ii.b). In case ii.a), it is $N_q = 2$ for both directions.

If not stated otherwise, for the random choice of a new imaginary time, the heat bath method is used, see Sec. 3.5.

Together, the steps presented here and in Sec. 3.4 are ergodic and fulfill the detailed balance, thus creating a Markov chain in the extended configuration space of open and closed paths. The combined sampling is needed to determine the ratio G/Z occurring in the estimator for the MGF. The MGF itself has to be sampled in the subspace of open paths containing any number of particles, while the calculation of thermodynamic estimators in the canonical ensemble, like the ones presented in Sec. 3.3, has to be restricted to closed paths with a given particle number N. As the weight $W_{\text{tot}}(C_N)$ of a configuration with N particles in the total configuration space contributing to Z_{tot} differs from its weight $W_N(C_N)$ in the canonical configuration space contributing to $Z(N)$ only by a normalization factor that is constant for a given N during the simulation, the subset of closed path samples $C_{N,i}$ with particle number N is distributed according to $W_N(C_N)/Z(N)$. Therefore, an estimate for an expectation value $\langle O \rangle_N$ in the canonical ensemble is simply given by the arithmetic mean over $O(C_{N,i})$. Likewise, estimates in the grand-canonical ensemble and of the MGF can be calculated by only considering the subset of samples belonging to the corresponding configuration spaces and ignoring the rest. As each subspace is visited with a different frequency, the autocorrelation time between samples of one subset differs from the autocorrelation time between samples belonging to another one. It is therefore advantageous to use an optimized cycle, cf. Sec. 2.2, for each subset, e.g., a large cycle for samples belonging to the grand-canonical ensemble and a small cycle for samples contributing to an expectation value for a certain particle number N. This non-trivial sampling procedure is demonstrated in Fig. 3.19. Note, that conditional sampling, like taking always the first sample after entering an associated subspace, leads to a wrong result.

As mentioned above, the six MC steps presented in this section are not ergodic on their own. An obvious limitation is that the creation of a type 4 kink with a non-zero weight requires the existence of a matching type 2 kink in the current

	C_1	C_2	C_3	C_4	C_5	C_6	C_7	C_8	C_9	C_{10}	C_{11}	C_{12}	C_{13}	C_{14}	
All samples:	C_1	C_2	C_3	C_4	C_5	C_6	C_7	C_8	C_9	C_{10}	C_{11}	C_{12}	C_{13}	C_{14}	$\cdots$
Contr. to $G(\mu, \beta, V)$:		C_2	C_3	C_4			C_7		C_9	C_{10}				C_{14}	$\cdots$
Contr. to $Z(\mu, \beta, V)$:	C_1				C_5	C_6		C_8			C_{11}	C_{12}	C_{13}		$\cdots$
Contr. to $Z(N_1, \beta, V)$:	C_1							C_8							$\cdots$
Contr. to $Z(N_2, \beta, V)$:					C_5	C_6					C_{11}	C_{12}	C_{13}		$\cdots$
$\vdots$															$\cdots$

Figure 3.19: Example for the sampling procedure in the total configuration space. Shown are the first 14 configurations after the end of the equilibration time. While all configurations contribute to the ratio G/Z, only a certain subset of configurations contributes to other quantities, e.g., only $C_{N_1,1} = C_1$ and $C_{N_1,2} = C_8$ contribute to expectation values in the canonical ensemble with N_1 particles. To reduce the autocorrelation time, only every second sample $C_{N_1,i}$ is taken into account for calculating averages over $O(C_{N_1,i})$ as indicated by the blue background. Because of a larger autocorrelation time, a cycle of $n = 3$ is chosen for the samples contributing to the MGF and the grand-canonical partition function, respectively, and $n = 4$ for G/Z. In real simulations, the cycles are much larger and can span several thousand MC steps.

configuration. For example, in the HEG, because of momentum conservation, type 2 kinks do not exist at all, i.e., all kinks inducing a one-particle excitation have a vanishing weight. For similar reasons, some type 4 kinks are incorrectly missing from all configurations in simulations for other Hamiltonians as well. In an attempt to solve the violation of the ergodicity condition, a second worm was introduced and a step that moves both heads at the same time was implemented. Although this step can add type 4 kinks directly without relying on the existence of a matching type 2 kink in the path, it is not sufficient for the algorithm to be ergodic, as revealed by comparison with simulations involving virtual kinks. The remaining missing configurations include combinations of at least three different type 4 kinks and no variation of the worm updates could be found to construct them, see. [59]. Therefore, the closed path updates of Sec. 3.4 had to be implemented in addition. However, as two worms allow for the direct sampling of the two-particle Green function $G_{ijkl}(\tau_1, \tau_2, \tau_1', \tau_2') = \langle \hat{T} \hat{a}_i(\tau_1) \hat{a}_j(\tau_2) \hat{a}_k^\dagger(\tau_2') \hat{a}_l^\dagger(\tau_1') \rangle$, the step might be useful in the future and will therefore be briefly presented here in detail.
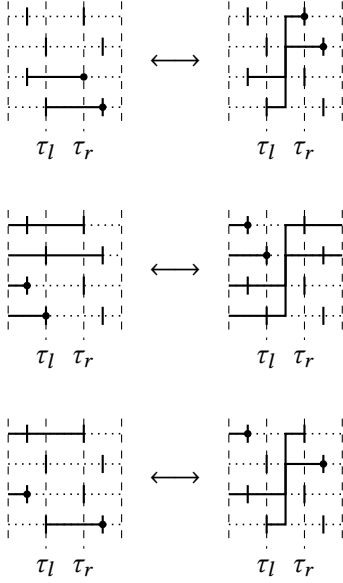
Figure 3.20: Moving the heads of two worms at once, thereby adding or removing a type 4 kink directly. The short horizontal lines mark the times of additional kinks in the path that determine the limits $\tau_{\min}$, $\tau'_{\min}$, $\tau_{\max}$, or $\tau'_{\max}$. The limits τ_l and τ_r of the interval I that are indicated by the horizontal dashed lines determine the possible range for the time τ_a of the kink to be added. Top: The kink s_a is added or removed left of both heads corresponding to the cases **7.**A) and **8.**iv.a), respectively. Middle: The kink s_a is located on the right of both heads corresponding to the cases **7.**B) and **8.**iv.b), respectively. Bottom: The cases **7.**C), **7.**D), **8.**iv.c), and **8.**iv.d) refer to adding or removing the kink s_a between both heads of the worms.

7. Add kink by moving two heads

The first head is given by $\hat{a}_p(\tau_h)$ and the second by $\hat{a}_q(\tau'_h)$. Choose an orbital $r \notin [p,q]$ and an orbital $s \notin [p,q,r]$ among all N_B orbitals. There are three possibilities:

A) Both orbitals r and s are unoccupied at the times τ_h and τ'_h, respectively

    i. Find the time $\tau_{\min}$ of the first kink on the left of τ_h affecting either orbital p or orbital r.

    ii. Find the time $\tau'_{\min}$ of the first kink on the left of τ'_h affecting either orbital p or orbital r.

    iii. Determine the interval $I = (\tau_{\min}, \tau_h)_\beta \cap (\tau'_{\min}, \tau'_h)_\beta$ taking the β-periodicity into account.

B) Orbital r is occupied at τ_h and orbital s is occupied at τ'_h

    i. Find the time $\tau_{\max}$ of the first kink on the right of τ_h affecting either orbital p or orbital r.

    ii. Find the time $\tau'_{\max}$ of the first kink on the right of τ'_h affecting either one of the orbitals q and s.

    iii. Determine the interval $I = (\tau_h, \tau_{\max})_\beta \cap (\tau'_h, \tau'_{\max})_\beta$ taking the β-periodicity into account.

C) Orbital r is empty at τ_h and orbital s is occupied at τ_h'

    i. Find the first kink on the left of τ_h that affects either orbital p or r. Its time defines $\tau_{\min}$.

    ii. Find the time $\tau_{\max}'$ of the first kink on the right of τ_h' affecting either one of the orbitals q and s.

    iii. Determine the interval $I = (\tau_{\min}, \tau_h)_\beta \cap (\tau_h', \tau_{\max}')_\beta$ taking the β-periodicity into account.

D) Orbital r is occupied at τ_h and orbital s is empty at τ_h'

    i. Find the time $\tau_{\max}$ of the first kink on the right of τ_h affecting either one of the orbitals p and r.

    ii. Find the first kink on the left of τ_h' that affects either one of the orbitals q and s. Its time defines $\tau_{\min'}$.

    iii. Determine the interval $I = (\tau_h, \tau_{\max})_\beta \cap (\tau_{\min}', \tau_h')_\beta$ taking the β-periodicity into account.

In all cases, reject the step if I is empty. Choose a time $\tau_a \in I$ at random. The new kink is given by $s_a = (r, s, p, q)$.

8. **Remove kink by moving two heads**
   The first head is given by $\hat{a}_p(\tau_h)$ and the second by $\hat{a}_q(\tau_h')$.

    i. Choose a kink s_a randomly among

        a) The kink immediately preceding τ_h.

        b) The first kink after τ_h.

        c) The latest kink before τ_h'.

        d) The kink immediately succeeding τ_h'.

        In all cases, the position refers to the β-periodic time ordering of all kinks.

    ii. If the kink is not given by $s_a = (p, q, i, j)$ with arbitrary i and j, reject the step.

    iii. Choose either a) or b) at random:

        a) $r = i$ and $s = j$.

        b) $r = j$ and $s = i$.

iv. There are four possibilities:

a) The orbitals r and s are each empty at the times τ_h and τ_h', respectively. If s_a is not the first kink on the left of τ_h that affects the orbitals p or r or not the first kink on the left of τ_h' affecting the orbitals q or s, reject the step.

b) Orbital r is occupied at τ_h and orbital s is occupied at τ_h'. The kink s_a must be the first kink on the right of τ_h affecting the orbitals p or r and the first kink on the right of τ_h' affecting the orbitals q or s. Otherwise, reject the step.

c) Orbital r is empty at τ_h and orbital s is occupied at τ_h'. The kink s_a must be the first kink on the left of τ_h that affects the orbitals p or r and the first kink on the right of τ_h' that affects the orbitals q or s. Reject the step otherwise.

d) Orbital r is occupied while orbital s is empty at their respective times. Reject the step, if the kink s_a is not the first kink on the right of τ_h that affects the orbitals p or r or not the first kink on the left of τ_h' affecting either one of the orbitals q and s.

v. Move the head at τ_h from orbital p to orbital r and the head at τ_h' from q to s, thereby deleting the kink s_a.

For adding the second worm, the step **1.** can be used without changes. For the removal of the second worm, one of the heads and one of the tails is chosen randomly. Afterwards, the step **2.** is executed as before. For all other steps that involve only one worm, one of the two heads is chosen at random before proceeding as described above. With these changes, the MC steps construct a Markov chain in the combined configuration space determined by the quantity $Z_{\text{tot},2} = Z + \omega G + \omega_2 G_2$, where G_2 is the two-particle Green function sum obtained by summing over all indices and integrating over all time arguments. The parameter ω_2 can be chosen freely to balance the relative weight among the three subspaces of paths containing none, one, or two worms.

4 | **The Harmonic Oscillator**

For the development of the CPIMC method, in [31, 33] the one-dimensional harmonic oscillator was chosen as a test system because its Hamiltonian exhibits many basic properties of general Coulomb interacting particles in continuous space while being relatively simple and easy to implement. As a model, the system is investigated in the context of ultracold atoms and electrons in quantum wires [102]. In this work, the application is extended to the two dimensional case, which is of broader interest for the understanding of the fundamental physics of few-body interacting quantum systems found in semiconductor quantum dots, often referred to as artificial atoms [42–46]. The MC steps of the CPIMC method are not tailored to the particular structural properties of the this Hamiltonian apart from the Slater-Condon rules. Therefore, the observations of this chapter should be transferable to systems with more complex Hamiltonians, which can easily be simulated by providing the corresponding one- and two-particle matrix elements as input data to the CPIMC method. In the following, only the fully polarized system will be considered for simplicity.

4.1 System Parameters

A two dimensional system of N Coulomb interacting particles trapped in a harmonic potential with rotational symmetry is described by the Hamiltonian

$$\hat{H} = \frac{1}{2} \sum_{\alpha=1}^{N} (\hat{\boldsymbol{p}}_\alpha^2 + \hat{\boldsymbol{r}}_\alpha^2) + \sum_{1 \leq \alpha < \beta} \frac{\lambda}{|\hat{\boldsymbol{r}}_\alpha - \hat{\boldsymbol{r}}_\beta|} \,,$$

in standard dimensionless oscillator units with length scale $r_0 = \sqrt{1/m\omega}$ and energy scale $E_0 = \omega$. The coupling parameter $\lambda = E_C/E_0$ is the ratio of the characteristic Coulomb energy $E_C = e^2/r_0$ and the characteristic energy of the confinement.
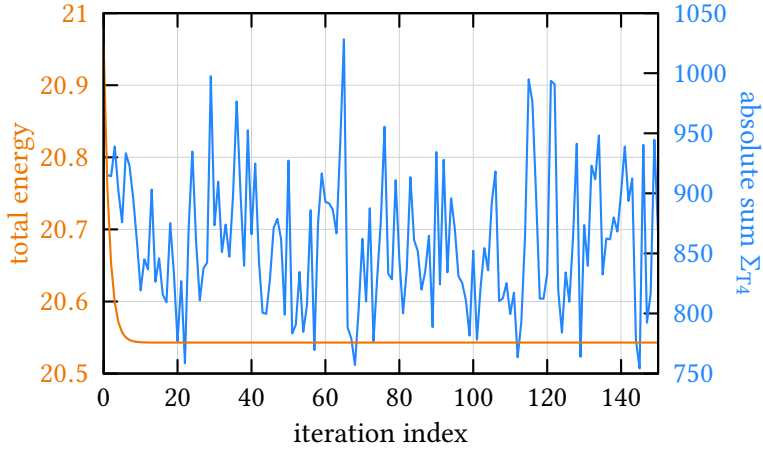
Figure 4.1: Total energy (orange line) and sum over the absolute values of all anti-symmetrized matrix elements with pairwise different indices (blue line) for each of the 150 iterations of a single ground-state HF calculation for $N = 6$ particles with a coupling parameter $\lambda = 0.8$ and $N_B = 45$ basis function. While the energy converges quickly, the large fluctuations of Σ_{T4} indicate that the basis continues to rotate. Data from [59].

The Hamiltonian of the ideal system of N non-interacting particles, corresponding to $\lambda = 0$, can be diagonalized analytically. In polar coordinates, $\boldsymbol{r} = (r, \varphi)$, the eigenstates of the one-particle problem are the well-known Fock-Darwin orbitals $\phi_{n_r m_\varphi}(r, \varphi)$ with radial quantum number $n_r \geq 0$ and angular quantum number $m_\varphi \in \mathbb{Z}$, see, e.g., [103]. The energy eigenvalues are given by $E_{n_r m_\varphi} = 2n_r + |m_\varphi| + 1$. Thus, the eigenvalues are degenerate for all but the lowest energy, giving rise to a shell structure where the i-th shell consists of i eigenstates of equal energy. In this basis, the two-particle interaction matrix elements

$$w_{ijkl} = \langle n_{r,i} m_{\varphi,i} n_{r,j} m_{\varphi,j} | \hat{w} | n_{r,k} m_{\varphi,k} n_{r,l} m_{\varphi,l} \rangle \, \delta_{\sigma_i, \sigma_k} \delta_{\sigma_j, \sigma_l}$$

are calculated in advance of a simulation using a program implemented by David Hochstuhl. The program employs the Talmi-Brody-Moshinsky transformation [104, 105] to reduce the original four dimensional integral to a finite sum of one dimensional integrals which can be numerically solved by half-open Gauss-Hermite integration [106], see App. A.

92
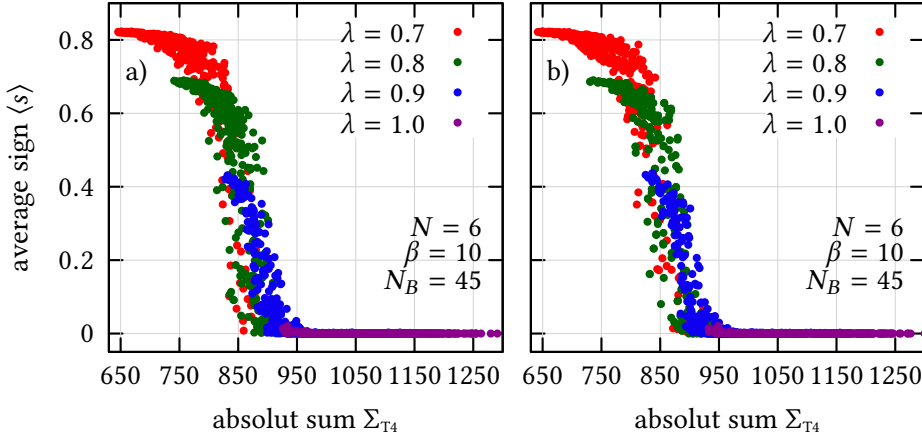
Figure 4.2: Dependence of the average sign on the value of Σ_{T4} for $N = 6$ particles, inverse temperature $\beta = 10$, $N_B = 45$ basis functions, and different coupling parameters λ in the canonical ensemble. Each point corresponds to a single CPIMC simulation using a basis set from a different iteration of a) a ground-state or b) a grand-canonical HF calculation. For each coupling strength, the basis with the smallest value of Σ_{T4} leads to an approximately maximum sign, while the difference between a) and b) is neglectable. Figure taken from [59].

Already in the preceding diploma thesis [31], it was discovered that the use of the HF basis instead of the eigenstates of the ideal Hamiltonian greatly increases the average sign of CPIMC simulations. The HF basis sets are obtained via a self-consistent, iterative diagonalization of the HF operator. However, in subsequent investigation by Simon Groth [59], it was shown that even after the energy and the reduced one-particle density are converged, the approximate eigenstates of the HF operator continue to rotate with every iteration which leads to large fluctuations of the average sign. The fluctuations occur independent of the ensemble at finite temperature but only for closed shell systems at $T = 0$. He found that the sum of the absolute values of all antisymmetrized matrix elements with four pairwise different indices,

$$\Sigma_{T4} := \sum_{\substack{i<j,k<l \\ i \neq k, i \neq l, j \neq k, j \neq l}} |w^-_{ijkl}|,$$

is strongly correlated with the average sign of a CPIMC simulation. The parameter Σ_{T4} can be calculated in every HF iteration and an example for its continued

fluctuation after the convergence of the total energy is depicted in Fig. 4.1. Fig. 4.2 shows that an approximately maximum sign corresponds to a minimum value of Σ_{T4}. This allows to find an approximately optimal set of basis functions by continuing the HF iteration after convergence is reached and picking the basis with the lowest value of the sum Σ_{T4}. Such a basis set can increase the average sign of a CPIMC calculation by up to one order of magnitude, resulting in a reduction of the CPU time by two orders of magnitude for a given accuracy. Simon Groth also showed that the average sign depends only weakly on the temperature of the HF calculation. Therefore, the basis functions from ground state HF calculations can be used for all CPIMC simulations, avoiding the computation of a unique basis set for each temperature.

4.2 Results

In this section, CPIMC results for several Coulomb interacting fermions in a two-dimensional harmonic trap will be presented to verify the correctness of the implementation of the algorithm and to demonstrate the capabilities of the method for simulating a general inhomogeneous many body system where coupling, temperature, and quantum degeneracy effects play an important role simultaneously. More results from CPIMC calculations for this system can be found in [59]. In all cases a HF basis was used as the underlying one-particle basis set. Where necessary, an optimized basis was chosen from additional HF iterations as described in the previous section. CPIMC results are subject to basis incompleteness and statistical errors only. All reported statistical uncertainties correspond to a carefully estimated one σ standard deviation, see Eq. (3.21) and the appendix in [31]. The basis size incompleteness error arises from the use of a finite number of basis functions, N_B, and is the same for otherwise exact CI and approximate HF calculations. The convergence with basis size is generally well-behaved and in most cases the remaining error is smaller than 0.1 %. For the purpose of comparing with CI and HF calculations, convergence towards the complete basis set limit is not always possible due to the restriction of these methods to small basis sizes.

The total energy per particle for two closed shell systems with $N = 3$ and $N = 6$ particles and various system parameters in the canonical ensemble is shown in Fig. 4.3 in dependence on the basis size. CPIMC and exact CI results are given for intermediate interaction strengths between $\lambda = 0.5$ and $\lambda = 1.5$ and two temperatures of $\beta = 1$ and $\beta = 10$, the latter being close to the ground-state. For a single data point, in total 1.2×10^8 samples with a cycle of 200 have been
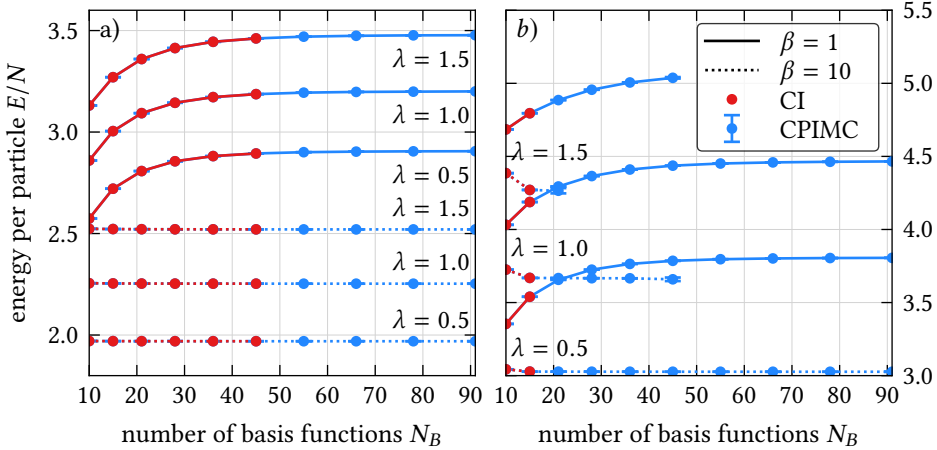
Figure 4.3: Comparison of energies per particle with CI calculations for a) $N = 3$ and b) $N = 6$ particles in the canonical ensemble for different number of basis functions, temperatures and coupling strengths. Red points denote exact CI results and blue points results from CPIMC simulations. Solid lines are guide to the eyes for points corresponding to $\beta = 1$ calculations while results for $\beta = 10$ are connected by dotted lines. Where CI results are available, the points lie exactly on top of each other.

accumulated during a parallel CPIMC run on 12 CPU cores, corresponding to 2.4×10^{10} MC steps and a total CPU time between 6 h and 24 h. The availability of CI calculations depends only on the size of the Hilbert space $N_{\text{det}} = \binom{N_b}{N}$. With the finite temperature CI implementation by David Hochstuhl that is used in this work, system sizes of up to $N_{\text{det}} = 25\,000$ are feasible, restricting calculations to $N_B = 45$ for $N = 3$ particles and $N_B = 15$ for $N = 6$ particles. On the other hand, CPIMC calculations are hampered by the fermion sign problem, which depends on temperature and coupling strength as well and will be investigated below. Where both methods are applicable, the energies are in perfect agreement with relative statistical errors smaller than 5×10^{-5}, rigorously verifying the correctness of the method and the implementation. For $N = 6$, $N_B = 10$, $\beta = 10$, and $\lambda = 0.5$ the relative statistical accuracy is even as high[1] as 2.8×10^{-7}. Compared with CI, the CPIMC method allows for much larger system sizes. For $\lambda = 0.5$ and in the case of $\beta = 1$ also for $\lambda = 1$, Hilbert space sizes up to $N_{\text{det}} = 6.7 \times 10^8$

---

[1] The highest accuracy is found for the more difficult $N = 6$ particle system because the HF basis has not been optimized for $N = 3$ particles.
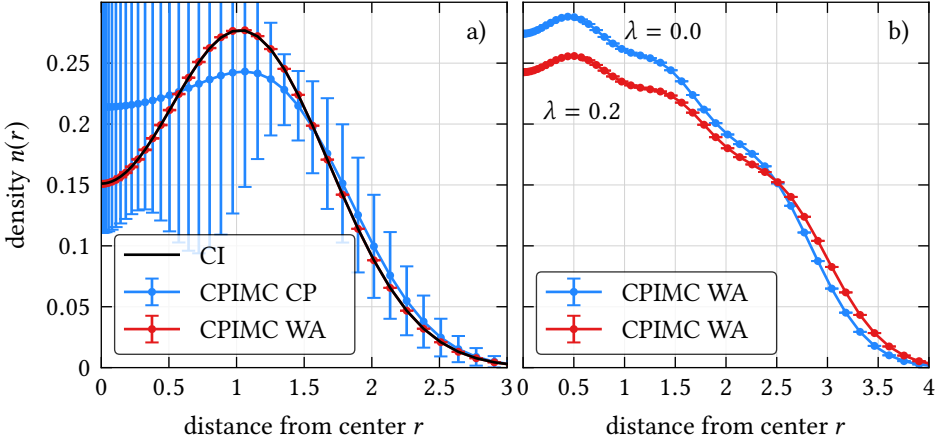
Figure 4.4: One-particle density $n(r)$ in the canonical ensemble. a) Comparison of results from CPIMC simulations using two different estimators with CI calculations (black line) for $N = 3$, $N_B = 28$, $\beta = 10$, and $\lambda = 2$. b) Effect of Coulomb interaction for $N = 20$, $N_B = 91$ and $\beta = 5$. The blue and red curves represent the CPIMC results for the non-interacting and weakly interacting system with $\lambda = 0$ and $\lambda = 0.2$, respectively. In b) the density was rescaled by $1/3.5$.

do not pose a problem. Therefore, for all parameters apart from $N = 6$ and $\lambda = 1.5$, convergence towards the complete basis size limit better than $0.1\,\%$ can be achieved, which is not possible with CI calculations except for small particle numbers and very low temperatures.

Another basic quantity is the reduced one-particle density (3.12), which is shown in Fig. 4.4 for $N = 3$ and $N = 20$ particles in the canonical ensemble. Due to the rotational symmetry of the trap, the density depends only on the radial distance from the center and statistical errors can be reduced by averaging over all angles. In all cases, one observes a rich structure, which is present even in the non-interacting case for $\lambda = 0$ because the Pauli exclusion principle forces the occupation of excited orbitals. In contrast to bosons, which form a simple Gaussian shaped density in the weakly interacting case [30], for fermions, the maximum density is located in a ring around the center, i.e., outside the region with the lowest potential energy. For larger particle numbers, further rings begin to emerge already at a rather weak interaction strength of $\lambda = 0.2$. Fig. 4.4 a) compares the densities obtained from the closed path (CP) estimator, Eq. (3.11), and from the worm algorithm (WA) estimator, Eqs. (3.22) and (3.23), with the exact

CI density for $N = 3$ particles in $N_B = 28$ basis functions, an inverse temperature $\beta = 10$, and a coupling parameter $\lambda = 2$. Both CPIMC results are from the same simulation using 30 CPU hours on a single core. While the CP estimator is useless in this case due to the large variance, the WA estimator agrees nicely with the result from the CI calculation. The relative accuracy is about 0.3 % in the center and smaller than 0.05 % at the first peak at $r_{\text{peak}} \approx 1$. The uncertainty increases rapidly for $r \gtrsim 3$ where the density drops below 0.005 but is always smaller than the accuracy of the CP estimator. This result demonstrates the huge improvement that was achieved by the development of the worm algorithm in this work. The new estimator allows for the calculation of the density for much larger systems as shown in Fig. 4.4 b). For $N = 20$ particles in $N_B = 91$ basis functions, $\beta = 5$, and $\lambda = 0.2$, the variance of the CP estimator is too large to even be estimated reliably so that it is not shown. However, the WA estimator results in a relative accuracy of 0.01 % in the center region, which is higher compared to the $N = 3$ case due to the reduced coupling strength. Despite the weak interaction, the differences to the ideal case with $\lambda = 0$ are obvious as the particles are pushed away from the center of the trap. There, the density is about 13 % lower. Additionally, density fluctuations are more pronounced.

Due to the large system size, CI and canonical HF calculations are infeasible even for $N = 20$ ideal particles. Therefore, in Fig. 4.5, HF and CPIMC densities are compared for smaller system sizes where HF calculations are possible for the grand-canonical as well as the canonical ensemble. For intermediate coupling strengths of $\lambda = 1.5$ and $\lambda = 1$, respectively, and low temperatures close to the ground-state, the HF density is in reasonable agreement with the CPIMC result only for closed shell configurations like in Fig. 4.5 a), where the blue lines show the density of a grand-canonical ensemble with an average particle number of $\langle N \rangle = 3$, while the deviation in the center amounts to 25 % for the open shell configuration in Fig. 4.5 b) with $N = 4$ particles. In both ensembles, HF calculations fail to resolve temperature effects at intermediate temperatures around $\beta = 3$ with deviations of about 10 % in the trap center. For higher temperatures, the system becomes more classical and more ideal and at $\beta = 0.5$, any structure in the density is lost. In that case, the differences between both methods are small. Note, that for such a high temperature a much larger basis would be necessary for a converged result. The comparison shows that the quality of HF results varies strongly with temperature and particle number and an exact method like CPIMC is necessary to identify parameters where HF calculations are problematic.

In addition to an increased accuracy of the one-particle density, the implementation of the worm algorithm for the CPIMC method also enables the calculation
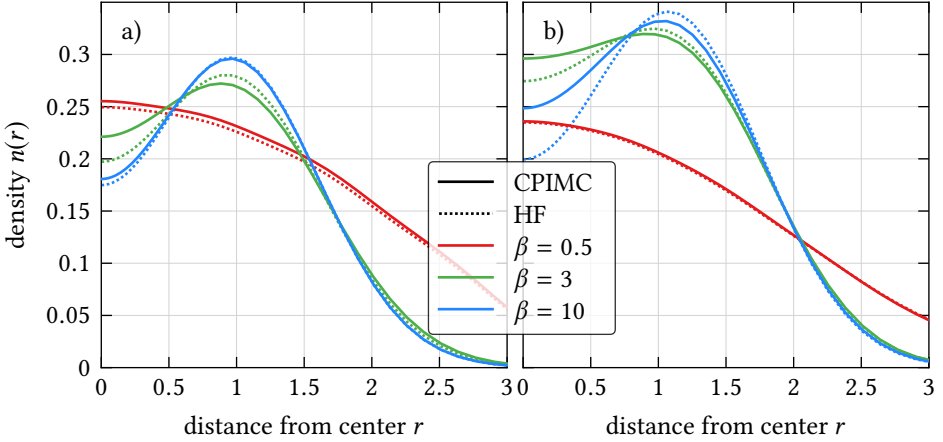
Figure 4.5: Temperature dependence of the one-particle density from CPIMC (solid lines) and HF (dotted lines) calculations. a) Grand-canonical ensemble with $\lambda = 1.5$, $N_B = 28$, and $\mu = 4.4$, corresponding to $\langle \hat{N} \rangle \approx 4.7$ at $\beta = 0.5$ (red lines) and $\langle \hat{N} \rangle \approx 3$ at $\beta = 3$ (green line) and $\beta = 10$ (blue lines). b) Canonical ensemble with $N = 4$, $N_B = 45$, and $\lambda = 1$. The statistical error of CPIMC calculations is smaller than the line width.

of the MGF. Given accurate data for the MGF, different methods allow for the reconstruction of the spectral function $A(q, \omega)$ and thus of dynamical properties from simulations in the thermodynamic equilibrium [41, 107, 108]. As the reconstruction of $A(q, \omega)$ is an ill-posed problem, it is essential for reliable results, that the systematic and statistical errors of the MGF data are below 0.1 %. Fig. 4.6 compares the diagonal elements of the MGF from HF and CPIMC calculations for a system consisting of $N = 4$ particles, a basis size of $N_B = 36$, moderate inverse temperature $\beta = 2$, and coupling strength $\lambda = 0.9$. Except for the nearly unoccupied orbital $i = 10$, for which the agreement between both methods looks reasonable, the deviations are obvious. For each orbital, the HF errors exceed 5 % at some point in imaginary time. This confirms the observation, that HF calculations are not reliable for open shell systems. The statistical errors of the otherwise exact CPIMC results are larger than 0.1 % only for the smallest values, e.g., at $\tau = 2$ for $i = 10$, but never exceed 0.5 %. Of course, the CPU time can always be extended beyond the 48 h used for this plot to further decrease the error. This shows that the CPIMC calculation can yield data with the necessary accuracy for the reliable reconstruction of the single-particle spectral function.
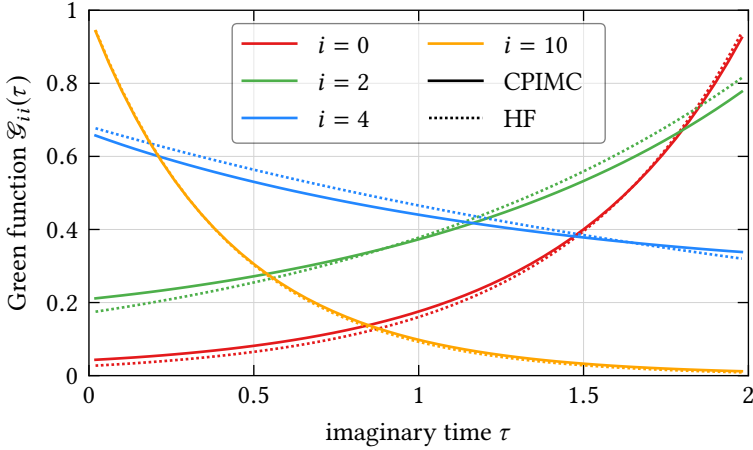
Figure 4.6: Diagonal elements of the Matsubara Green function from CPIMC (solid lines) and HF (dotted lines) calculations over imaginary time for $N = 4$, $N_B = 36$, $\beta = 2$, and $\lambda = 0.9$. The color indicates the orbital number. The statistical error of the CPIMC results is smaller than the line thickness.

Even if the differences between HF and CPIMC results for the energy and single-particle density are small for closed shell systems at low temperature for coupling strengths accessible to CPIMC simulations, the availability of an exact method is still valuable. On the one hand, it can provide benchmark results that allow for the quantification of the errors of approximative methods and help with the development of other new methods [30, 80]. On the other hand, other quantities are much more sensitive to approximations. An example is the single-particle entanglement entropy

$$S_1 := - \sum_i \lambda_i \log \lambda_i,$$

where the λ_i denote the natural orbital occupation numbers, i.e., the eigenvalues of the reduced single-particle density matrix. At $T = 0$, S_1 is a measure for the entanglement of a single particle with all others and hence, the single-particle entanglement entropy of a non-interacting system is zero. Fig. 4.7 shows the dependence of the entanglement entropy on the coupling parameter for a closed shell system in the canonical ensemble with $N = 3$ particles at low, $\beta = 10$, and moderate, $\beta = 3$, temperatures. Starting at nearly zero for $\lambda = 0$ and $\beta = 10$, which shows how close this temperature is to the ground-state, the entanglement
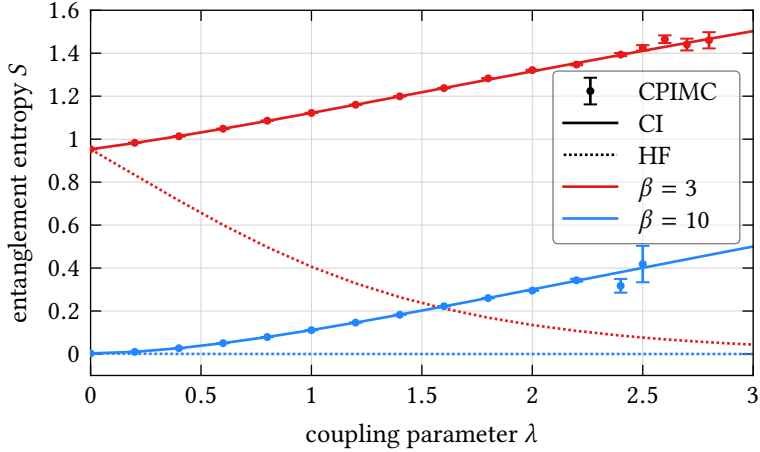
Figure 4.7: Dependence of the entanglement entropy on the coupling strength. Shown are results from exact CI (solid lines), CPIMC (symbols with errorbars) and HF (dotted lines) calculations for $N = 3$ and $N_B = 28$ in the canonical ensemble. Red and blue colors correspond to $\beta = 3$ and $\beta = 10$, respectively.

entropy increases with increasing interaction strength and temperature. Within statistical errors, the CPIMC results are in perfect agreement with the results from CI calculations. However, while for the corresponding system in the grand-canonical ensemble the HF density is in reasonable agreement with the exact CPIMC data, cf. Fig. 4.5 a), HF calculations yield the opposite trend of a decreasing entropy with increasing coupling strength. Because for $\beta = 10$, both ensembles nearly coincide, this behavior cannot be explained with the difference between the canonical and grand-canonical ensembles, but must be attributed to the neglect of correlations in the HF approximation.

Another example demonstrating the potential of the CPIMC method to quantify the accuracy of other methods is depicted in Fig. 4.8. For a system of $N = 3$ particles at a temperature of $\beta = 10$ and a coupling strength of $\lambda = 2$ in the canonical ensemble, the total energy obtained from calculations using the MLB algorithm [29] is compared with the corresponding CPIMC result. As the MLB approach is based on the PIMC method in coordinate representation, it is free of any basis incompleteness error that is inherent for wave-function based methods like CPIMC. Therefore, the CPIMC results have to be extrapolated to the complete basis size limit. Although the extrapolation introduces a small systematic uncertainty because the functional dependency on the number of
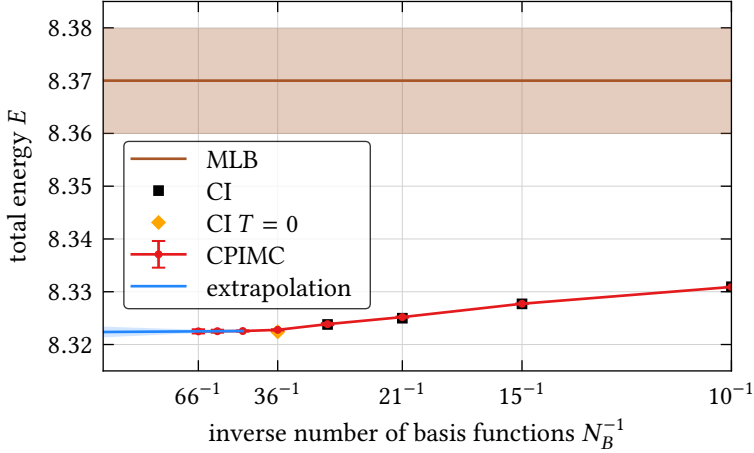
Figure 4.8: Total energy for $N = 3$, $\beta = 10$, and $\lambda = 2$ from CPIMC (red circles) and MLB [29] (brown line) calculations. For CPIMC, the convergence with basis size is shown, while the MLB result is basis size independent. Where available, CI results are given by black squares. At $N_B = 36$, the CI energy for the ground state is denoted by an orange diamond. A linear extrapolation of the CPIMC results to the complete basis size limit is shown by the blue line. Shaded areas indicate statistical uncertainties.

basis functions is unknown, the deviation between the CPIMC and the MLB energy is significant. To further rule out any error in the basis functions that are used in both the CI and CPIMC calculations, the energy from independent ground-state CI calculations [103] is shown for $N_B = 36$. The observed small deviation is expected for a very low but finite temperature of $\beta = 10$, supporting the reliability of the CPIMC results. The most likely explanation for the deviation of the MLB energy is an insufficient convergence with the number discretization factors as only up to $M = 32$ time slices have been used for all results reported in [29]. Alternatively, the number of $K = 600$ samples per time slice, which constitutes an additional approximation in the MLB algorithm, introduces a larger bias than expected for certain system parameters. This example shows how important CPIMC calculations are for providing exact results at system sizes, temperatures, and coupling strengths that cannot be accessed by other methods.

The range of applicability of QMC methods is predominantly determined by the fermion sign problem because the number of samples and thereby the runtime needed for a desired accuracy scales quadratically with the inverse of
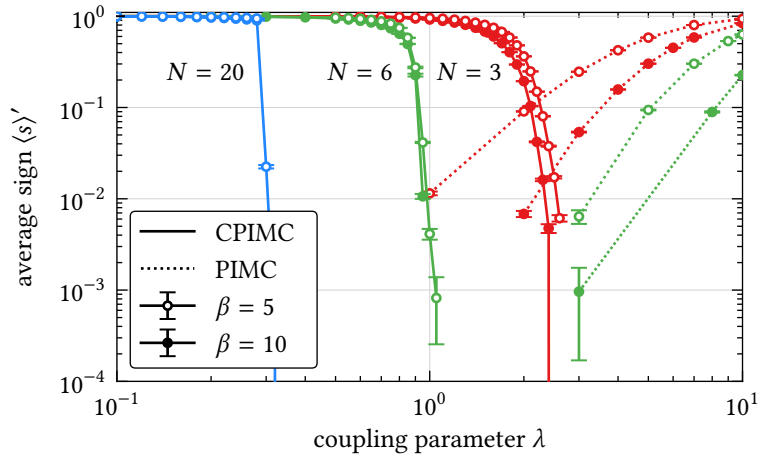
Figure 4.9: Sign problem for CPIMC and PIMC simulations over coupling strength for different particle numbers and temperatures in the canonical ensemble. Solid (dotted) lines represent CPIMC (PIMC) calculations at $\beta = 10$ (filled symbols) and $\beta = 5$ (open symbols). Red, green, and blue colors denote $N = 3$, $N = 6$, and $N = 20$ particles in $N_B = 36$, $N_B = 55$, and $N_B = 91$ basis functions, respectively.
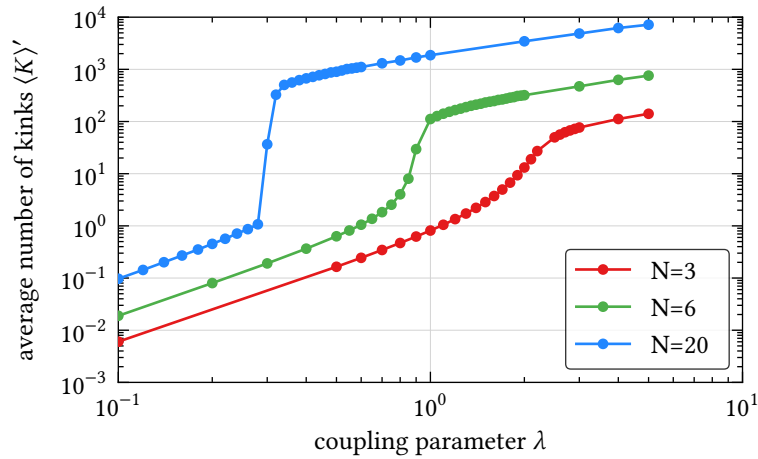

Figure 4.10: Dependence of the average number of kinks on the coupling strength for $\beta = 10$ and the same number of particles and basis functions as in Fig. 4.9.

average sign. When the average sign $\langle s \rangle'$ is lower than 0.01, calculations become very costly, and they become generally infeasible for an average sign below 0.001. Fig. 4.9 shows the average sign of the standard PIMC and the CPIMC method in dependence of the coupling parameter for different particle numbers and temperatures. For strong repulsion, $\lambda \gg 1$, the inter-particle distances are large compared to the spatial extent of the particles and thus the overlap of the wave functions is small and particle exchange, which is the source of the sign problem in PIMC calculations, plays only a minor role. For such systems of nearly classical, point-like particles, the average sign of the PIMC method is close to 1 and simulations are highly efficient. When the coupling strength is reduced, i.e., when the confinement is strengthened compared to the Coulomb interaction, the quantum nature of the particles becomes more important and the average sign decreases exponentially. This decrease of the average sign is more pronounced for larger particle numbers and lower temperatures. Therefore, weakly interacting systems are inaccessible for PIMC calculations.

The CPIMC method, on the other hand, has no sign problem at all in the non-interacting case, but becomes more challenging for larger coupling parameters. At certain critical value of λ, the average sign starts to decrease rapidly, eventually resembling a step function for large N. The critical coupling strength λ_{crit} at which this drop is observed depends mainly on the particle number and to a lesser degree on the temperature. It occurs at weaker interaction for larger particle numbers and lower temperatures. Therefore, the CPIMC method is most efficient for weak to intermediate interaction strengths, making it completely complementary to PIMC methods in the coordinate representation.

The sign problem of the CPIMC method vanishes for the ideal system and becomes more severe towards stronger interactions because the expression (3.9) is a series expansion of the partition function in terms of the coupling parameter λ around $\lambda = 0$. The larger the parameter λ, the more terms in the series are necessary for convergence, resulting in an increasing number of kinks in the simulation, as shown in Fig. 4.10. As each kink potentially changes the total sign of a configuration, the average sign decreases with an increasing number of kinks.

One observes three distinct regions with different functional dependence of the average number of kinks on the coupling parameter. Starting with zero kinks in the non-interacting limit, the number of kinks grows quadratically with λ for weak interactions. The coefficient of this quadratic growth increases with particle number and inverse temperature (not shown here). When the average number of kinks is of the order of one, it starts to increase rapidly until the functional dependence becomes approximately linear for stronger interactions. For large

particle numbers, this transition happens almost instantaneously, increasing the average number of kinks by several orders of magnitude within a rather small region around the critical coupling parameter λ_{crit}. In the case of the HEG, see Sec. 5.4, where much larger basis sizes are available, it was checked carefully that the approximately linear growth for $\lambda \gg 1$ is not an effect of the finite basis size or of the finite simulation time. The observed behavior is similar to a phase transition in a finite-size system where certain parameters like the energy exhibit a discontinuity for $N \rightarrow \infty$. However, since the average number of kinks is not a physical observable, there is no connection to a physical phase transition of the system at λ_{crit}. For a way to obtain accurate results for interaction strengths beyond λ_{crit}, see Sec. 5.5.

5 | The Homogeneous Electron Gas

Despite its apparent simplicity, the accurate investigation of the HEG is of tremendous importance for the understanding of the physics of complex matter. Accurate data for the HEG are an essential ingredient for DFT, one of the most widely used methods in solid-state physics and chemistry. In their seminal paper [109], Ceperley and Alder calculated the exchange-correlation energies for the complete density range of the ground-state HEG using a fixed-nodes QMC method, thus allowing for the construction of the first accurate exchange-correlation functionals in the local density approximation [16] and thereby laying the basis for the large success of DFT at zero temperature.

In recent years, experimental progress, e.g., new inertial confinement fusion (ICF) devices at the National Ignition Facility (NIF) [47, 110, 111], Rochester [48], and Sandia [49, 112], increased the demand for an accurate theoretical understanding of warm dense matter. At such conditions, going beyond the ground state approximation is crucial. DFT for finite temperatures FTDFT is in principle long known [3] but requires data for the HEG over the entire density-temperature plane. However, at finite temperatures, fixed-node QMC methods cannot be applied to the high densities in the warm dense matter regime, cf. available data in Fig. 5.1, and are additionally afflicted with significant uncontrollable systematic errors [55]. Therefore, improving the exchange-correlation functionals which are based on existing data [58] is of key relevance for the development of FTDFT. Furthermore, the HEG in the warm dense matter regime is a fundamental model system to study the non-trivial interplay between quantum degeneracy, correlation, and thermal excitation effects as well as finite-size corrections in periodic boundary conditions [113–115].

5.1 System Parameters

The HEG, also called the uniform electron gas (UEG) or jellium, is a model system of Coulomb-interacting electrons embedded in a homogeneous, positively
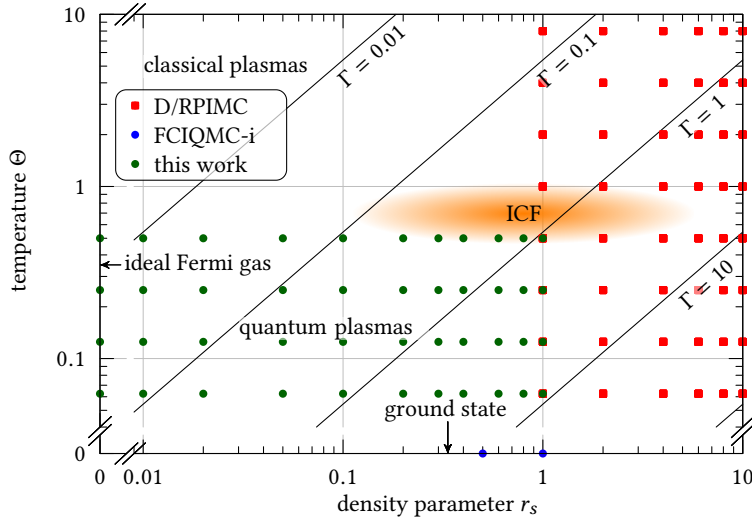
Figure 5.1: Overview over the density-temperature plane of the HEG and the warm dense matter regime. The degeneracy parameter $\Theta = T/E_F$ measures the importance of a quantum mechanical description. For $\Theta < 1$ quantum effects dominate. $\Gamma = e^2/\bar{(r)}T$ is the classical coupling parameter. In the quantum regime, r_s is the relevant measure for the interaction strength. Available RPIMC results for finite-temperatures [55] and FCIQMC results for the ground state [116] are denoted by red squares and blue dots. The green dots show the main results of this work. Typical parameters for ICF experiments [111] are indicated by the orange shaded area. From [117].

charged background ensuring charge neutrality. It is the quantum mechanical analog to the one-component plasma of classical physics [118]. In the macroscopic limit, the HEG is described by only three parameters: the Brückner parameter $r_s = (3/4\pi n)^{1/3}$ which depends on the density $n = N/V$ and corresponds to the mean inter-particle distance, the degeneracy parameter $\Theta = T/T_F$ which is given by the ratio of the temperature T to the Fermi temperature T_F, and the spin polarization $\xi = (n_\uparrow - n_\downarrow)/n$ where $n_\uparrow$ and $n_\downarrow$ denote the separate densities of the electrons with spin up and down, respectively. In the following, unless explicitly mentioned, only the fully spin polarized case, i.e. $\xi = 1$, is studied.

As the CPIMC method cannot simulate the thermodynamic limit directly, a system of N electrons in a cubic box of side length L and volume $V = L^3$ with periodic boundary conditions is considered. Following Fraser *et. al.* [114], the

Hamiltonian in coordinate representation reads[1]:

$$\hat{H} = -\sum_{i=1}^{N} \nabla_i^2 + \sum_{i=1}^{N} \sum_{\substack{j=1 \\ j\neq i}}^{N} \psi(\boldsymbol{r}_i, \boldsymbol{r}_j) + E_M,$$

where the Ewald pair potential $\psi(\boldsymbol{r}_i, \boldsymbol{r}_j)$ accounts for the long-range nature of the Coulomb interaction in periodic boundary conditions and E_M is the Madelung energy due to the self-interaction of the electrons with all their periodic images. The Ewald potential is derived by splitting the sum of all Coulomb interactions in the periodic array of images of the central simulation cell into a short-range and a long-range contribution. Evaluating the short-range part in real space and the long-range part in Fourier space one arrives at

$$\psi(\boldsymbol{r}_i, \boldsymbol{r}_j) = \frac{1}{V} \sum_{\boldsymbol{G}\neq 0} \frac{e^{-\pi^2 G^2/\kappa^2} e^{2\pi i \boldsymbol{G}(\boldsymbol{r}_i - \boldsymbol{r}_j)}}{\pi G^2} - \frac{\pi}{\kappa^2 V} + \sum_{\boldsymbol{R}} \frac{\text{erfc}(\kappa|\boldsymbol{r}_i - \boldsymbol{r}_j + \boldsymbol{R}|)}{|\boldsymbol{r}_i - \boldsymbol{r}_j + \boldsymbol{R}|},$$

where $\boldsymbol{R}$ and $\boldsymbol{G}$ are the real and reciprocal lattice vectors of the periodically repeated simulation cell, satisfying $\boldsymbol{G} \cdot \boldsymbol{R} \in \mathbb{Z}$. The positive constant κ can be chosen arbitrarily to control the speed of convergence. For sensible choices of κ, both sums converge much faster than the direct Coulomb sum[2]. The Madelung energy per particle is given by

$$E_M/N = \frac{1}{V} \sum_{\boldsymbol{G}\neq 0} \frac{e^{-\pi^2 G^2/\kappa^2}}{\pi G^2} - \frac{\pi}{\kappa^2 V} + \sum_{\boldsymbol{R}\neq 0} \frac{\text{erfc}(\kappa|\boldsymbol{R}|)}{|\boldsymbol{R}|} - \frac{2\kappa}{\sqrt{(\pi)}}$$

and can be numerically calculated to $E_M \approx -2.837297 \cdot (3/4\pi)^{1/3} N^{2/3} r_s^{-1}$.

Using a single-particle basis of plane waves, which is the eigenbasis of the ideal, non-interacting system and in coordinate representation given by

$$\phi_i(\boldsymbol{r}, \sigma) = \frac{1}{\sqrt{V}} e^{i\boldsymbol{k}_i \cdot \boldsymbol{r}} \delta_{\sigma_i, \sigma},$$

with wave vector $\boldsymbol{k}_i = \frac{2\pi}{L} \boldsymbol{m}_i$ and $\boldsymbol{m}_i \in \mathbb{Z}^3$, the Hamiltonian can be written in second quantization as

$$\hat{H} = \sum_{ij} h_{ij} \hat{a}_i^\dagger \hat{a}_i + 2 \sum_{\substack{i<j, k<l \\ i\neq k, j\neq l}} w_{ijkl}^- \hat{a}_i^\dagger \hat{a}_j^\dagger \hat{a}_l \hat{a}_k,$$

---

[1] Throughout this chapter Rydberg units are used.

[2] Note that the Coulomb sum of an infinite system of oppositely charged particles is only conditionally convergent, i.e., the result depends on the order of the summation. On the other hand, the limit of taking finite, neutral clusters of larger and larger size is well defined and it is assumed that the Ewald summation is a good approximation of this limit.

where $h_{ij} = k_i^2 \delta_{i,j}$ is the ideal kinetic energy contribution and $w_{ijkl}^- = w_{ijkl} - w_{ijlk}$ with

$$w_{ijkl} = \frac{4\pi}{V(k_i - k_k)^2} \delta_{k_i + k_j, k_k + k_l} \delta_{\sigma_i, \sigma_k} \delta_{\sigma_j, \sigma_l}$$

are the anti-symmetrized two-particle integrals. Here, the total momentum and spin are conserved due to the Kronecker-deltas. Note that the matrix elements of the Ewald interaction are equal to the Fourier components of the bare Coulomb potential. The divergent contributions for $k_i = k_k$ and $k_j = k_l$ are canceled by the interaction with the positive background and are therefore excluded from the sum.

In the plane wave basis, the CPIMC method is also naturally capable of simulating the ideal, non-interacting system, which is recovered for $r_s \to 0$. As the above expression in Rydberg units diverges in that limit, the following expressions are internally used for all numerical simulations instead:

$$\tilde{h}_{ij} = k_i^2 \delta_{i,j},$$

$$\tilde{w}_{ijkl} = \frac{\lambda}{2} \frac{1}{(k_i - k_k)^2} \delta_{k_i + k_j, k_k + k_l} \delta_{\sigma_i, \sigma_k} \delta_{\sigma_j, \sigma_l}, \tag{5.1}$$

with $k_i \in \mathbb{Z}^3$. The internal units implied by these relations are connected to Rydberg units by

$$\lambda = \frac{4}{(2\pi)^3} L = \frac{4}{(2\pi)^3} \left(\frac{4\pi}{3}\right)^{\frac{1}{3}} r_s N^{\frac{1}{3}},$$

$$\tilde{E} = \frac{(2\pi)^4}{16} \lambda^2 E,$$

$$\tilde{T} = \frac{(6\pi^2 N)^{\frac{2}{3}}}{(2\pi)^2} \Theta.$$

In these units it is obvious that the limit $\lambda = r_s = 0$ corresponds to the ideal Fermi gas[3].

5.2 CPIMC for the HEG

The algorithm outlined in Sec. 3.4 is sufficiently general to yield exact results given precomputed tables of the kinetic and interaction matrix elements h_{ij} and w_{ijkl}.

---

[3]It should be noted that the description of the HEG given in this section is exact only at the Schrödinger level. For densities higher than $r_s = 0.1$ the kinetic energy of the electrons corresponds to about 10 % of the speed of light and a relativistic treatment would be more appropriate.
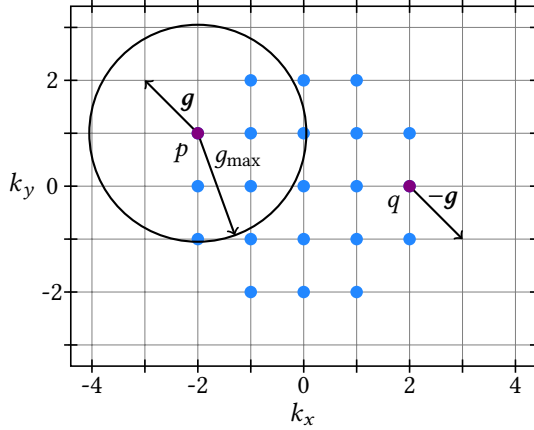
Figure 5.2: Choosing two-particle excitations. The grid shows available wave vectors in the 2D HEG. Blue points indicate occupied orbitals. After two random occupied orbitals p and q have been chosen, an excitation vector $\boldsymbol{g}$ with $g < g_{\max}$ is chosen at random. The new orbitals are then given by $\boldsymbol{k}_r = \boldsymbol{k}_p + \boldsymbol{g}$ and $\boldsymbol{k}_s = \boldsymbol{k}_q - \boldsymbol{g}$.

However, a very large number of basis functions is necessary for convergence to the complete basis set limit. Fortunately, the matrix elements can easily be computed on demand, thereby dramatically reducing the memory requirements of a simulation. Additionally, the simple form of the matrix elements allows to construct MC steps explicitly tailored to the HEG for a greatly enhanced efficiency of the Metropolis algorithm. In the following, the adapted MC steps, which have been published in [119], will be described.

The first property of the HEG Hamiltonian to note is that

$$\langle \{n\}|\hat{H}|\{\bar{n}\}\rangle = 0 \text{ for } \{n\} = \{\bar{n}\}_q^p.$$

In the picture of the CPIMC method, this means that there exist no kinks of type 2, i.e., $q_{\{n\},\{\bar{n}\}}(p,q) = 0$. Therefore, any MC step involving type 2 kinks can be omitted which greatly reduces the number of required updates. The only remaining step involving a one particle excitation is the excitation of a fully occupied orbital, i.e., an orbital that is not affected by any kinks. This step is still necessary because the trace over the antisymmetric subspace of the Hilbert space involves a summation over different total momentums.

Further, the new orbitals in a two-particle excitation were both chosen randomly in correspondence to the general structure of the Slater-Condon rules.

However, given two particles in the occupied orbitals p and q, the momentum conservation in the HEG requires that the chosen empty orbitals r and s satisfy $\mathbf{k}_s = \mathbf{k}_p + \mathbf{g}$ and $\mathbf{k}_r = \mathbf{k}_q - \mathbf{g}$ with $\mathbf{g} \in Z^3$ and $\mathbf{g} \neq 0$. The resulting matrix element is then proportional to $1/\mathbf{g}^2 - 1/(\mathbf{k}_p - \mathbf{k}_q + \mathbf{g})^2$. This allows for two changes of the algorithm that are depicted in Fig. 5.2. First, instead of choosing the orbitals s and r randomly, an excitation vector $\mathbf{g}$ is chosen at random instead. It is then proposed to excite the two particles from the orbitals p and q to the orbitals given by $\mathbf{k}_r = \mathbf{k}_p + \mathbf{g}$ and $\mathbf{k}_s = \mathbf{k}_q - \mathbf{g}$. The step is rejected, if either r or s is affected by a kink or fully occupied. Second, as the matrix elements are non-zero for any vector $\mathbf{g}$, an arbitrary excitation with $|\mathbf{g}| > g_{\max}$ can be achieved by a finite number of excitations where the vectors $\mathbf{g}_i$ of the individual excitations are chosen from only the vectors with $|\mathbf{g}_i| < g_{\max}$. The number $g_{\max}$ can be chosen to maximize the efficiency of the simulations. In the following, $g_{\max} = 2$, as shown in Fig. 5.2, is used for all calculations.

Finally, due to the large number of single-particle orbitals necessary for the convergence to the complete basis size limit, the excitation of a single fully occupied orbital p was improved. All orbitals are grouped into shells $i = 0, 1, \ldots$ with shell i containing orbitals p with energy $e_p \in [i\Delta e, (i+1)\Delta e)$, where the shell width Δe depends on temperature. Instead of choosing a target orbital q among all empty orbitals that are not affected by kinks, only orbitals in the shells $i - 1, i$, and $i + 1$ are considered. Still, arbitrary orbitals can be reached by several consecutive excitations, ensuring ergodicity. This improvement is independent of the HEG and will be useful for any system where a large number of basis functions is needed.

Apart from these changes, it was necessary to improve many more details of the implementation to allow for efficient simulations with basis sizes of $N_B > 10\,000$. Some of these improvements are already included in Sec. 3.5 and Sec. 3.6.

5.3 Comparison with CI

As a consequence of the complexity of the Metropolis algorithm in the case of the CPIMC method, the correctness of the results has to be verified for each different Hamiltonian because the ergodicity of the MC steps cannot be guaranteed. Fortunately, such a test can be performed without difficulties due to the equivalence between CPIMC and CI calculations.

A highly accurate comparison between CPIMC and CI results for a small system of $N = 4$ particles in $N_B = 19$ basis functions is shown in Fig. 5.3. $N = 4$ is the smallest particle number allowing for all MC steps to occur, i.e., comparisons for
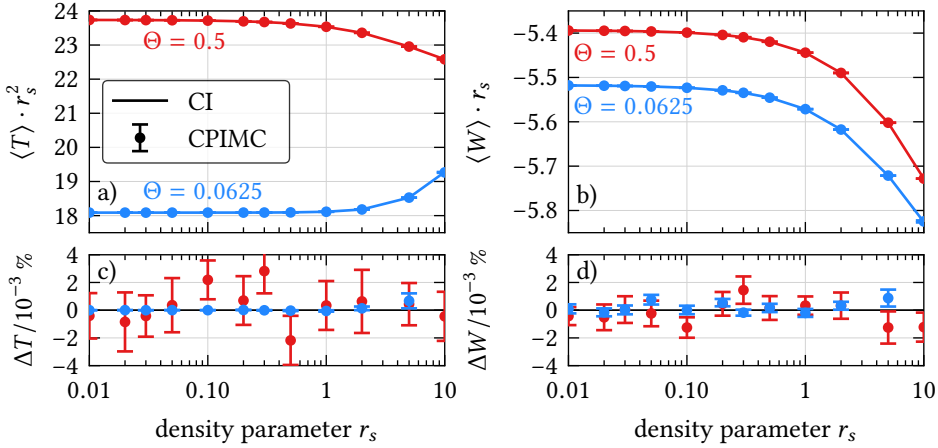
Figure 5.3: Energies for the 4 particle HEG. Shown are a) the total kinetic energy, b) the total interaction energy, and, in c) and d), the corresponding relative deviations from exact CI calculations for $N = 4$ particles in $N_B = 19$ basis functions for various densities. Colors indicate temperatures of $\Theta = 0.0625$ (blue) and $\Theta = 0.5$ (red). Dots with errorbars represent CPIMC results while solid lines connect CI results.

$N = 3$ or less particles cannot detect problems in all of the MC steps. Earlier, slightly less accurate results for the total energy of the $N = 4$ particle system have been published in [119]. Here, for different densities between $r_s = 0.01$ and $r_s = 10$ and temperatures of $\Theta = 0.0625$ and $\Theta = 0.5$, the kinetic energy times r_s^2 and the interaction energy times r_s are compared separately. This is more sensitive to systematic errors because a cancellation of errors may occur for the total energy. For each data point, 1.6×10^8 samples with a cycle $n = 200$ have been used, resulting in a total runtime of about 6 CPUh. As already mentioned in Sec. 4.2, the runtime of the CI calculations depends only on the system size but scales exponentially with the number of basis functions and number of particles. On the other hand, the CPIMC algorithm is applicable to much larger systems, but is afflicted with the sign problem which grows additionally with the interaction strength, i.e., with the density parameter r_s in the case of the HEG. Although the sign problem is less severe for higher temperatures, the variance of the estimators grows with temperature. For high densities, i.e., small r_s, the latter effect can dominate the statistical error, as can be seen in the figure. The sign problem will be investigated in more detail in the next section. The comparison with the CI results reveals a relative accuracy of the CPIMC calculation up to 1×10^{-6} for
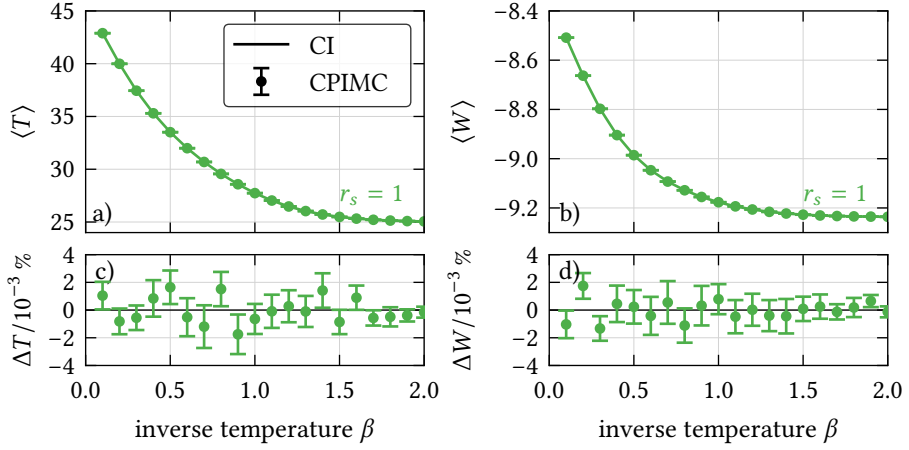
Figure 5.4: Energies for the 7 particle HEG. Shown are the temperature dependence of a) the total kinetic energy, b) the total interaction energy, and, in c) and d), the corresponding relative deviations from exact CI calculations for $N = 7$ particles in $N_B = 19$ basis functions for $r_s = 1$. Dots with errorbars represent CPIMC results while solid lines connect CI results. CI data courtesy of Fionn Malone.

$\Theta = 0.5$ and even 3×10^{-8} for $\Theta = 0.0625$. The relative accuracy of the kinetic energy is comparable to that of the interaction energy for the high temperature but higher by an order of magnitude for the low temperature. For all data points, the agreement is excellent, which confirms the correctness of the algorithm and its implementation.

Another comparison for a slightly larger number of particles is shown in Fig. 5.4. For $N = 7$ particles in $N_B = 19$ basis functions and a density of $r_s = 1$, the kinetic and interaction energy from CI and CPIMC calculation is compared for various temperatures between $\Theta = 0.086$ and $\Theta = 1.71$. Again, for all data points an excellent agreement is achieved with relative deviations smaller than 1.8×10^{-5}. As the CI data points are obtained from independent calculations by Fionn Malone, this comparison additionally verifies the correctness of the Hamiltonian matrix elements, which in the previous figure were computed by a program code that is partly shared between the CI and CPIMC implementation used in this work.
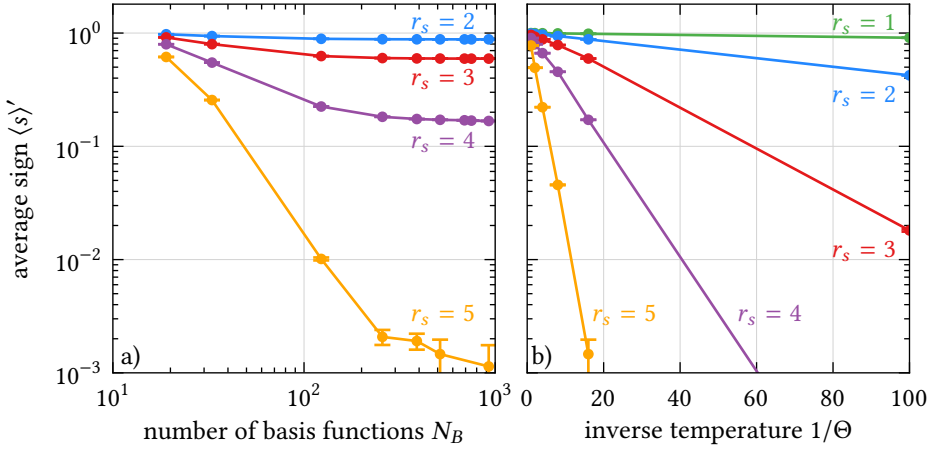
Figure 5.5: Average sign versus a) basis size and b) temperature for $N = 4$ particles and densities given by $r_s \in [1, 2, 3, 4, 5]$, indicated by the colors green, blue, red, violet, and orange, respectively. In a), the temperature is $\Theta = 0.0625$ and in b), $N_B = 515$ basis functions are used. Solid lines are guides to the eye. Note the different scales of the abscissa. Adapted from [119].

5.4 Sign Problem

The main advantage of the CPIMC approach compared to other PIMC methods has already been shown to be the complementary behavior with respect to the sign problem in the case of the harmonic oscillator, cf. Sec. 4.2 and also earlier works [31, 59]. In this section it will be investigated if this advantage applies to the HEG in the same way and how the system parameters, i.e., particle number, density, and temperature, determine the average sign, which is the most important measure for the applicability of any method afflicted with the sign problem.

As will been shown later, for the HEG a large number of basis functions are necessary to achieve an accurate convergence to the complete basis size limit. Therefore, a weak scaling of the average sign with the number of basis functions is important for efficient simulations. As shown in Fig. 5.5 for a system of $N = 4$ particles at a temperature of $\Theta = 0.0625$, the average sign scales favorably and converges towards a minimum value which depends on the density parameter r_s. The larger r_s, the smaller is the infinite basis size limit of the average sign. If the minimum value is sufficiently large for a given density, e.g., higher than 1×10^{-2}, simulations with an almost arbitrary large number of basis functions are feasible,

because the runtime of the current implementation scales only with $O(N_B)$ for a given number of MC steps.

Fig. 5.5 b) shows the dependence of the average sign on the dimensionless inverse temperature $1/\Theta = T_F\beta$ for the same $N = 4$ particle system as in panel a) but with a fixed basis size of $N_B = 515$. For all densities, one observes approximately straight lines. Due to the logarithmic presentation this corresponds to an exponential scaling of the average sign with the inverse temperature in accordance to the general relation, Eq. (2.5). However, the coefficient $f - f'$ in the exponential, i.e., the difference of the free energy per particle in the original and the actually simulated system, depends strongly on the density. Thus, for moderately high densities, even temperatures close to the ground state can be reached with acceptable computational costs.

As mentioned in the introduction to this chapter and shown in Fig. 5.1, prior to this work, QMC data only cover densities with $r_s \geq 1$. For higher densities no reliable results existed at finite temperatures. Therefore, it is of particular interest how the average sign depends on the density. The upper part of Fig. 5.6 shows the dependence on r_s for various particle numbers in $N_B = 2108$ plane waves at $\Theta = 0.125$. Since the density parameter r_s is proportional to the coupling strength λ in the internally used units, see Eq. (5.1), it is not surprising that the curves exhibit the same trend as in the case of the harmonic trap, cf. Fig. 4.9. Again, the CPIMC method has no sign problem at all for the ideal system in the limit of $r_s \to 0$ and its efficiency depends only weakly on r_s for high densities. Then, at a density $r_{s,\text{crit}}$ which is determined by the number of particles, the average sign starts dropping rapidly for lower densities. The critical density is higher for larger numbers of particles and the slope is also steeper. Therefore, the CPIMC method is applicable to high densities and allows to investigate the properties of the HEG at densities complementary to those achievable by PIMC calculations in coordinate representation. However, at low temperatures and strong degeneracy, a comparison with available RPIMC results for $r_s = 1$ and $N = 33$ is impossible without further improvements, because for this particle number and $\Theta \leq 0.5$, the average sign vanishes almost instantaneously at $r_s \approx 0.5$, restricting direct CPIMC calculations to densities higher than $r_s \leq 0.4$.

The explanation for the behavior of the average sign given in Sec. 4.2 applies to the HEG as well. Fig. 5.6 b) shows how the average number of kinks grows with the density parameter r_s for the same system parameters as above. As in the case of the harmonic oscillator, the rapid decrease of the average sign is accompanied by a similarly rapid increase of the average number of kinks. In the CPIMC simulation, a configuration with a large number of kinks represents a high order term in the series expansion of the partition sum with respect to the coupling
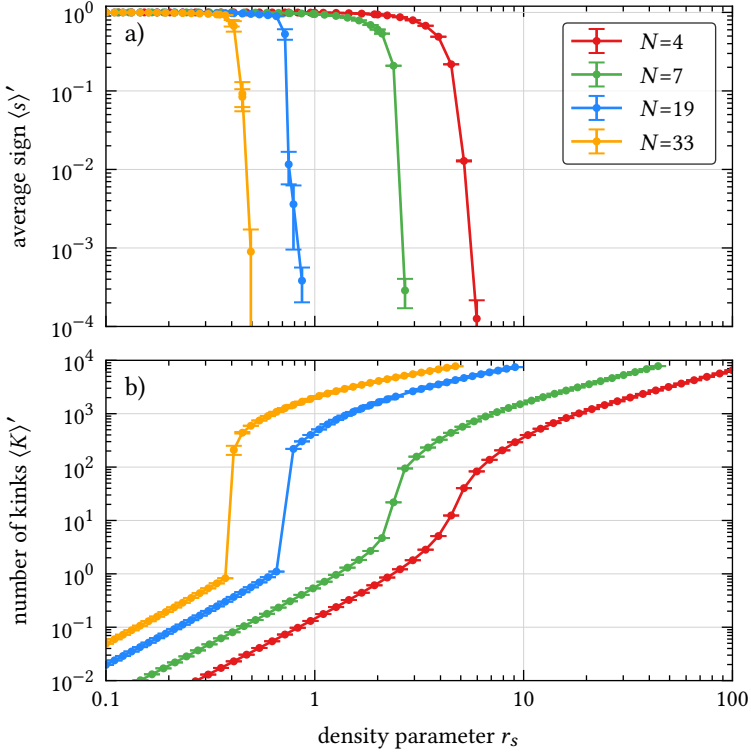
Figure 5.6: Density and particle dependence of a) the average sign and b) the average number of kinks for $\Theta = 0.125$ and $N_B = 2109$. Particle number are indicated by colors from $N = 4$ (red, rightmost curve) to $N = 33$ (orange, leftmost curve). Solid lines are guides to the eye. Note the logarithmic scale. From [120].

parameter λ, cf. Eq. (3.9). The larger λ, the larger the magnitude of high order terms which corresponds to a larger weight of configurations with many kinks. As each kink is a possible source of a sign change, this causes the dependence of the average sign on the coupling strength. However, it is an open question why the average number of kinks grows by several orders of magnitudes around the critical density $r_{s,\mathrm{crit}}$ but then continues to grow approximately linearly with r_s for lower densities. The different slope for $N = 33$ particles in Fig. 5.6 b) is likely an effect of the finite basis or runtime.

Note that for the application of the Metropolis algorithm it is necessary to take the absolute values of the expansion terms in the Dyson series into account. If the
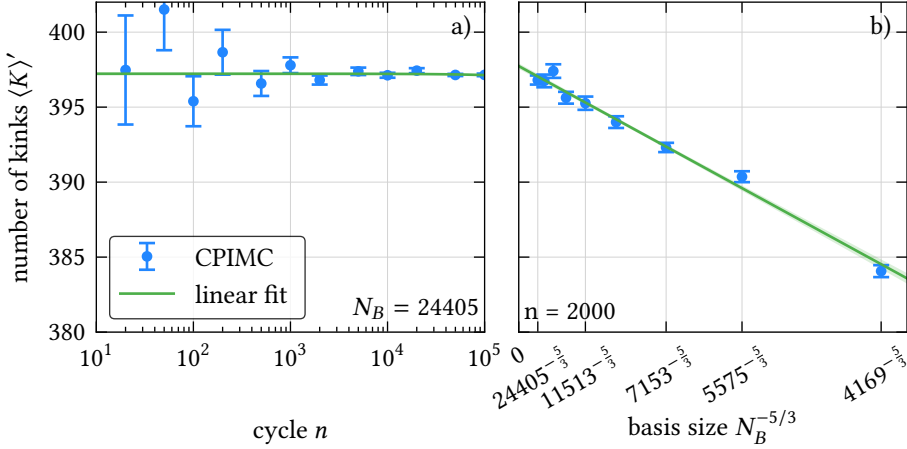
Figure 5.7: Convergence test of the average number of kinks with a) runtime and b) basis size for $N = 4$ particles at $r_s = 10$ and $\Theta = 0.125$. The results are averages of 1×10^6 measurements. The cycle n determines the number of skipped MC steps between two measurements and thereby the runtime. In each simulation, the first 1×10^5 measurements are thrown away, i.e., the equilibration time is always about 10 % of the total runtime. In a), the basis size is fixed at $N_B = 24405$, while in b), a cycle of $n = 2000$ is used. The solid lines are linear fits to all data points.

partition function was only conditionally convergent, the sum of absolute values might not even converge at all. To exclude the possibility that the observed finite number of kinks is only an artifact of the finite basis size or the finite simulation time, the convergence of the average number of kinks with both variables is depicted in Fig. 5.7 for $N = 4$ particles at $r_s = 10$ and $\Theta = 0.125$ which is beyond the the rapid increase and at the beginning of the approximately linear regime. Each data point corresponds to a simulation with 1×10^6 measurements and an equilibration time of 1×10^5 full cycles. Thus, the total runtime is implicitly determined by the cycle n, which is the number of skipped MC steps between two measurements. For $N_B = 24405$ basis functions, the left panel a) demonstrates that even for a small cycle of $n = 100$ the equilibration time is sufficiently long to yield a statistically correct result, although the autocorrelation time is much larger as can be seen by the decreasing error bars for longer cycles. An estimate of the integrated autocorrelation time is $\tau_{\text{int}} \approx 6.5 \times 10^6$. Using a cycle of $n = 2000$, which yields reasonable small error bars, the convergence

with basis size is shown in panel b) of the figure. It is already known [80] that the total energy converges with $N_B^{-5/3}$, which will also be demonstrated later. Here, the convergence behavior of the average number of kinks is in very good agreement with the same power law. Therefore, it can be concluded that neither the finite basis size nor the finite simulation time is the reason for the observed finite average number of kinks in the CPIMC simulations, and thus, the CPIMC formulation of the partition sum is likely absolutely convergent. However, a deeper understanding of the reason for the observed behavior is still needed as further insight might reveal strategies to reduce the sign problem and allow simulation of larger particle numbers and lower densities.

In the case of the harmonic oscillator, switching to the HF basis significantly improved the average sign of the simulations, see Sec. 4.2. However, for the HEG, the plane wave basis, i.e., the eigenbasis of the ideal system, coincides with the HF basis and also with the natural orbitals, i.e., the reduced single-particle density matrix is diagonal in this basis. Thus, there is no obvious choice of another single-particle basis set that could enable CPIMC simulations at lower densities. An explicitly correlated plane wave basis is proposed in [121], but has not been implemented yet for the CPIMC method due to the increased complexity. Instead, a different approach is explored in the next section.

5.5 Kink Potential

In order to limit the number of kinks and to increase the average sign, an additional factor $V_K(\kappa, \delta) = 1/(e^{-\delta(\kappa - K + 0.5)} + 1)$ is introduced in the partition sum:

$$Z(\kappa, \delta) = \sum_{K=0}^{\infty} \sum_{\{n\}} \sum_{s_1} \sum_{s_2} \cdots \sum_{s_{K-1}} \int' \mathrm{d}^K \tau$$
$$V_K(\kappa, \delta) W(K, \{n\}, s_1, \ldots, s_K, \tau_1, \ldots, \tau_K).$$

The factor $V_K(\kappa, \delta)$ was chosen to strongly reduce the weight of configurations with a lot of kinks but to converge to the exact solution in the limit $\kappa \to \infty$. The parameter δ controls the smoothness of the potential. In the limit $\delta \to \infty$, the potential $V_K(\kappa, \delta)$ reduces to a step function that excludes all configurations with $K > \kappa$. The additional constant 0.5 was added for symmetry. Because the exponential decrease of V_K with K is not strong enough to prevent the jump in the average number of kinks for all densities, the potential is set to zero if $V_K < V_c$ with an arbitrarily chosen constant cutoff V_c. A small value for this constant, e.g, $V_c = 1e - 9$, is usually sufficient. Performing simulations with different values of

κ allows to extrapolate the results to the exact limit with $V_K = 1$. This so-called kink potential was developed and compared to similar summation techniques for divergent series like Lindelöf and Cesàro [35, 122] by Simon Groth during his doctoral research. His investigation surprisingly revealed that the ad-hoc choice of a Fermi function generally results in a more favorable convergence behavior allowing a more reliable extrapolation compared to the other methods.

Plotting the total energy for a fixed set of simulation parameters for different values of $1/\kappa$ and $\delta = 1$, one observes a smooth S-shaped curve. For small values of κ, the curve is concave and shows an accelerated decrease of the energy with increasing κ. At some point, the decrease of the energy begins to decelerate until it converges to the exact value for $\kappa \to \infty$, resulting in a convex tail of the curve for large κ. For a reliable extrapolation, the error bars have to be sufficiently small for the onset of the convergence after the inflection point to be clearly visible. Fig. 5.8 shows the extrapolation of the kink potential parameter κ to the limit $1/\kappa \to 0$ for three especially difficult cases. For all three curves, the condition of a distinguishable convex tail is fulfilled. The top panel a) of Fig. 5.8 demonstrates the convergence of the total energy to the correct limit for a small system of $N = 4$ particles in $N_B = 19$ basis functions where an exact result from a CI calculation is available. At a temperature of $\Theta = 0.5$, the density parameter $r_s = 40$ is at the edge of the range of applicability of direct CPIMC calculations with an average sign of $\langle s \rangle' = 0.00417(8)$. Nevertheless, the extrapolated black point includes the exact CI result shown by an orange point within the errorbars given by the extrapolation procedure explained below. A second demonstration of the reliability of this method is shown in panel b) for a large system of $N = 33$ particles with a temperature of $\Theta = 8$ and a density of $r_s = 10$. The large basis size of $N_B = 44473$ ensures that the difference to the complete basis set limit is negligible relative to the statistical error. For these parameters, the system is almost classical so that direct PIMC simulations are available [55]. The systematic error of a finite number of time slices can be easily controlled and the results can be considered exact. Again, the extrapolation of the CPIMC results for the different values of κ shown by the black point with error bars agrees nicely with orange PIMC data point. The relative errors between the extrapolated estimates and the exact reference results amount to 0.15 % and 0.06 % for the comparisons with CI and PIMC, respectively. As a final example, the most challenging extrapolation for a system without an exact reference energy is shown in the bottom panel of Fig. 5.8. For $N = 33$ particles in $N_B = 2109$ basis functions with $r_s = 1$ and $\Theta = 0.0625$ only RPIMC results with an uncontrollable systematic error are available, which will be shown in the next section. The shape of the curve for these parameters closely resembles the one shown in panel
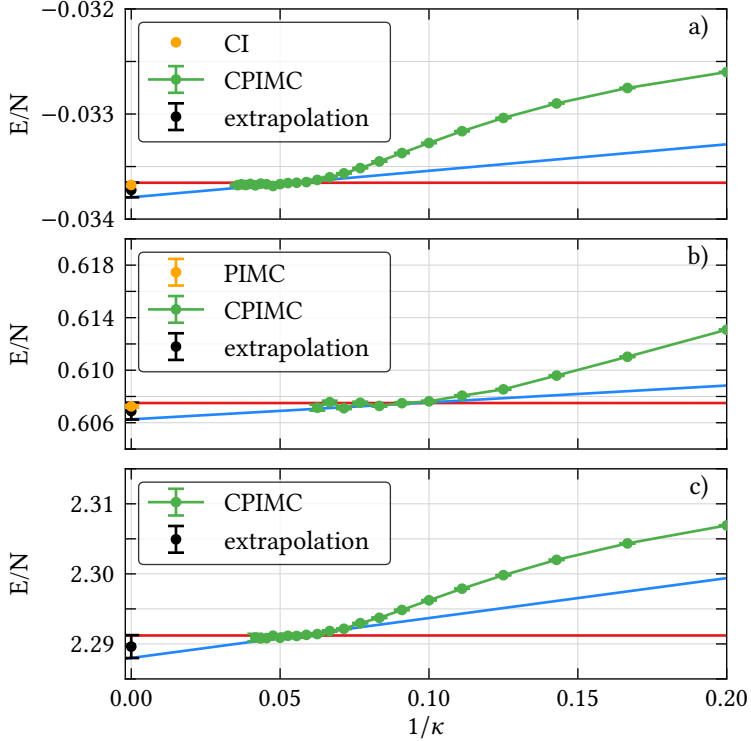
Figure 5.8: Extrapolation with respect to the kink potential parameter κ. Green dots with error bars are CPIMC results. The red lines and blue lines are horizontal and linear worst case fits (see text). The black dot is the mean value of the extreme cases, its error bar is given by their difference. a) For $N = 4$ particles in $N_B = 19$ basis functions the extrapolated value can be compared to exact CI results (orange dot). Shown is an example for $\Theta = 0.5$ and $r_s = 40$. Taken from [120]. b) The same as in a), but for $N = 33$, $N_B = 44473$, $\Theta = 8$, and $r_s = 10$, where an exact PIMC data point is available [55]. In both cases, the exact results lie between the error bars of the extrapolated value. c) Extrapolation from [117] for $N = 33$, $N_B = 2109$, $\Theta = 0.0625$, and $r_s = 1$. At the time of that writing, no exact value for comparison existed.

a). Therefore, one can safely assume that the extrapolated result with a relative error of 0.07 % is a reliable estimate for the correct energy. In this case, the total runtime for all points amounts to 20 000 CPUh. For all other densities and temperatures in this work the sign problem is less severe and the extrapolation is even more accurate.

The extrapolation to the exact limit $1/\kappa \to 0$ is performed as follows. After discarding all points with an statistical error exceeding 1 %, the remaining points are shifted upwards by $1\,\sigma$ and a horizontal line is fitted to the 5 points with the largest value of κ. To find a worst case estimate for the upper bound, further consecutive points with smaller κ are included in the horizontal fit one by one. The first time a newly added point deviates more than $4\,\sigma$ from the resulting fit the previous value of the horizontal fit is taken as the upper bound $E_{\max}$. The fit is considered unreliable if less than 6 points are included in the horizontal fit. The lower bound is found by shifting the point with an error smaller than 1 % downwards by $1\,\sigma$ and performing a linear fit to the consecutive points with κ values included in the horizontal fit. Additional linear fits are performed with further points added one by one. The smallest intersect of all linear fits with fewer than 4 points deviating more than $2\,\sigma$ and fewer than 2 points deviating more than $3\,\sigma$ from the fit is taken as the lower bound $E_{\min}$. The mean of these two extreme values is taken as the estimate with half of the difference as the error bar. For some quantities, e.g., for the kinetic energy, the results for different κ converge to the exact limit from below. In such cases, the above procedure is inverted. By the above approach, conservative estimates of the worst case errors of the extrapolation are obtained, always enclosing the correct value for $V_K = 1$ within the resulting error bars. The uncertainty and the small bias of the estimated confidence region can be further reduced by relaxing some of the above conditions, but that requires a more elaborate detection of the onset of the convergence to exhibit a comparable robustness.

With the introduction of the kink potential $V_K(\kappa, \delta)$, the range of applicability of the CPIMC method could be extended from $r_s = 0.4$ to $r_s = 1$ for $N = 33$ and $\Theta = 0.0625$ and up to $r_s = 10$ for $\Theta = 8$. This result is even more remarkable if one considers that in the last case, a nearly converged result for the total energy can be obtained with $\kappa = 10$, corresponding to an average number of kinks of $\langle K \rangle' = 8.4169(25)$ and an average sign of $\langle s \rangle' = 0.076918(43)$, whereas the direct simulation of the same system involves on average configurations with more than 1000 kinks leading to an average sign that is indistinguishable from zero within statistical error bars. This discrepancy in the average number of kinks of more than 2 orders of magnitude can be explained by a nearly complete cancellation of high order contributions with opposite signs to the series expansion of the

partition sum. It is in the special case of the CPIMC simulations, where only the absolute weights of the configurations are take into account and consequently determine the distribution of the number of kinks, that such high order terms are important for the convergence of the series. By reducing the weight of high order contributions, only the most important configuration of a given order occur during a simulation while higher orders are even completely removed by a non-vanishing cut-off V_c. This corresponds to the assumption, that all neglected terms approximately sum to zero when their sign is taken into account. The error of this assumption can be reliably estimated by the extrapolation procedure described above.

5.6 *Ab initio* results for 33 polarized electrons

With the introduction of the kink potential extrapolation technique it is now possible to close the gap between direct CPIMC calculations and available RPIMC results for $N = 33$ particles for a large range of finite temperatures. The results shown in the following are restricted to the fully polarized system, i.e., $\xi = 1$, at low temperatures with $\Theta \leq 0.5$. For particles with only a single spin state, $N = 33$ corresponds to a closed shell with respect to the degenerate energy states of the ideal system. It is expected that this choice minimizes finite basis set errors, which are investigated in the next section. Results for higher temperatures and the unpolarized system will be outlined in the last section of this chapter.

Simulations in the coordinate representation, such as used by the PIMC and RPIMC methods, are automatically performed in the complete basis set limit. For a meaningful comparison, the error that is introduced by using only a finite basis set in wave-function-based methods like CPIMC has to be accurately quantified. It is already known that for a large number of basis functions, the convergence of the total energy follows a power law of $N_B^{-5/3}$ in the fully polarized system [80]. Because it constitutes a valuable test for the CPIMC method at large basis sizes, the convergence behavior is thoroughly investigated here as well. Fig. 5.9 shows the incomplete basis set error of the total energy over $N_B^{-5/3}$ for a system of $N = 4$ particles close to the ground state at $\Theta = 0.0625$ and two densities of $r_s = 0.5$ and $r_s = 2$ where CPIMC calculations with the necessary accuracy are possible. While for a small basis set with $N_B < 257$ the error deviates from a straight line and thus from the expected power law, the inset clearly demonstrates that for large basis sizes the CPIMC results are in good agreement with a $N_B^{-5/3}$ convergence as shown by the linear fit depicted by the dotted line. It can also be seen that the absolute finite basis set error is larger for higher densities. The reason is
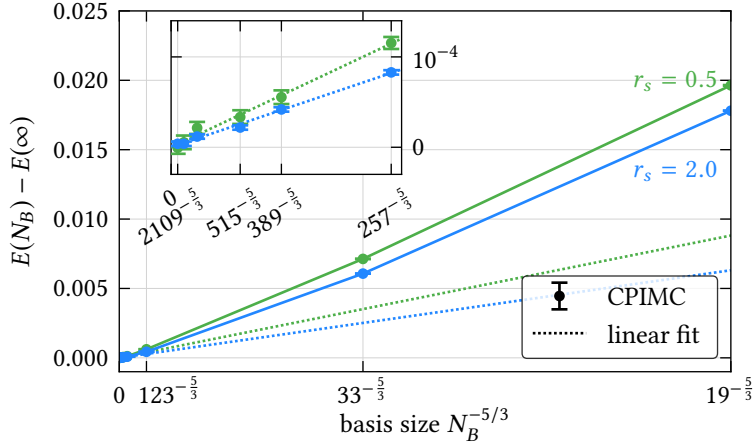
Figure 5.9: Basis size incompleteness error of the total energy for $N = 4$ particles, $\Theta = 0.0625$, and $r_s = 0.5$ (green, upper curve) and $r_s = 2$ (blue, lower curve). respectively. The dashed lines are linear fits to data points with $N_B \geq 257$, see inset. The scaling with $N_B^{-5/3}$ [80] is accurately reproduced for all but the smallest basis sizes. The solid lines are guides to the eye.

the much larger total energy at higher densities in Rydberg units. Conversely, the relative finite basis size error is larger for lower densities. The relatively slow convergence to the complete basis set limit in the plane wave basis is an effect of the slow convergence of the Fourier transform of the Coulomb potential. Fortunately, due to the similar convergence of the average sign, cf. Fig. 5.5, and the ability to calculate the matrix elements during the calculation, the efficiency of the CPIMC algorithm depends only weakly on the number of basis function in the case of the HEG and therefore allows for basis sizes orders of magnitude larger than what is possible for other exact wave-function-based methods.

The influence of finite temperatures on the convergence behavior is illustrated in Fig. 5.10. In addition to the curve for $r_s = 2$ and $\Theta = 0.0625$ from Fig. 5.9, the finite basis set error at $\Theta = 0.5$ is shown for the same system. For large basis sizes, the results are in agreement with a power law of $N_B^{-5/3}$. However, due to the increased variance of the estimator at high temperatures the error bars are too large to determine a specific functional dependence. At small basis sizes, the convergence behavior is fundamentally different. While the total energy is slightly overestimated at low temperatures, it is far too low at high temperatures. The reason is that high temperatures lead to the occurrence of highly excited
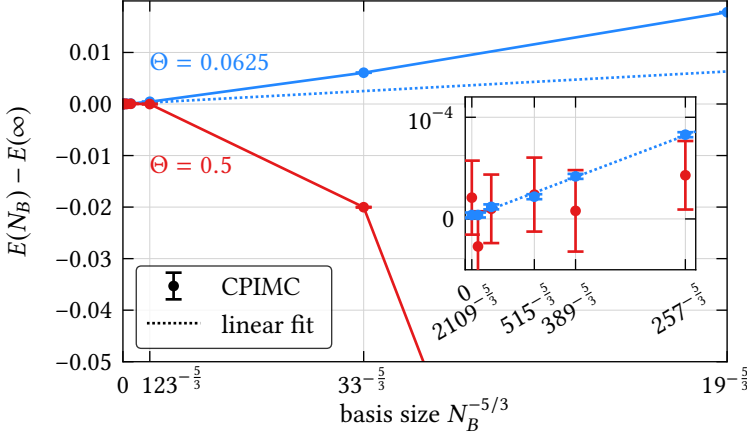
Figure 5.10: Influence of temperature on the basis size incompleteness error for $N = 4$ particles and $r_s = 2$. The blue curve and the dashed line for $\Theta = 0.0625$ are the same as in Fig. 5.9. For $\Theta = 0.5$ (red curve), the basis size incompleteness error scales exponentially for small numbers of basis functions, but converges within errorbars towards the $N_B^{-5/3}$ scaling of the ground state, see inset.

states which cannot be represented by a small number of basis function. In the ideal Fermi gas in the grand-canonical ensemble, the occupation numbers are given by the Fermi distribution, corresponding to an exponentially decreasing finite basis set error. In the canonical ensemble and for finite coupling strengths, the Fermi distribution is still applicable as a qualitative approximation. The total convergence can thus qualitatively be described by a superposition of the power law observed at finite r_s close to the ground state and the exponential convergence caused by the Fermi distribution.

The insights gained in the above considerations can be used to estimate the finite basis set errors arising in CPIMC simulations of the polarized HEG with $N = 33$ electrons. Fig. 5.11 shows the extrapolation to the complete basis set limit according to $N_B^{-5/3}$ for $r_s = 0.4$ and different temperatures. Due to the large error bars, the slope of the linear fits are not reliable and all curves are consistent with a horizontal line for basis sizes larger than $N_B = 925$. For basis sizes with $N_B \geq 2109$, the finite basis set error is negligible compared to the statistical error. Although the relative finite basis set error increases for larger rs, it remains small relative to the uncertainties of the CPIMC results which grow much faster due to the necessary use of the kink potential $V_K(\kappa, \delta)$ and the extrapolation to the limit
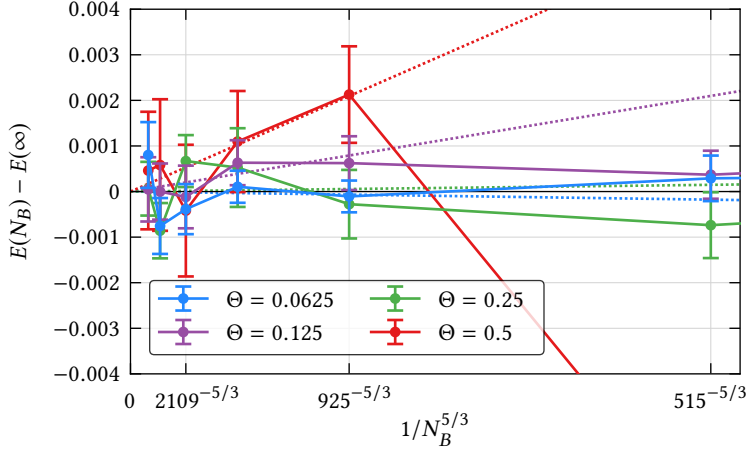
Figure 5.11: Accuracy of incomplete basis calculations for $N = 33$ particles and $r_s = 0.4$. Colors denote different temperatures. The figures demonstrates that the statistical error dominates over the basis size incompleteness error for large basis sizes $N_B \geq 2109$. In all cases, the slope is consistent with a horizontal within error bars. Taken from [117].

$\kappa \to \infty$. Therefore, one can avoid a costly extrapolation towards the complete basis set limit for all combinations of temperature and density and simply use $N_B = 2109$ basis functions for all simulation with $r_s \leq 1$ and $\Theta \leq 0.5$.

Having thoroughly verified the accuracy the of the CPIMC method for the HEG at finite temperatures, it can now be applied to calculate the exchange-correlation energy $E_{XC} = E_{tot} - E^{id}$ for $N = 33$ electrons over the whole density range between the known HF limit for $r_s \to 0$ and the available RPIMC results for $r_s \geq 1$ at finite temperatures with $\Theta \leq 0.5$. The results for the exchange-correlation energy per particle times r_s, which are a starting point for the construction of exchange-correlation free energy functionals used in FTDFT calculations, are shown in Fig. 5.12. The plot reveals a significant mismatch between the two methods at $r_s = 1$ for all depicted temperatures. While the CPIMC results reproduce the constant HF limit for $r_s < 0.1$ and can be smoothly connected to the RPIMC data for $r_s \gtrsim 6$ as indicated by the monotonically decreasing blue dotted line, the RPIMC curves exhibit a positive slope between $r_s = 1$ and $r_s = 2$ that would indicate a non-monotonic convergence to the HF limit. As the uncertainties of the CPIMC method are well controlled, the discrepancy must be attributed to the RPIMC results [55] which are afflicted with an unknown systematic error
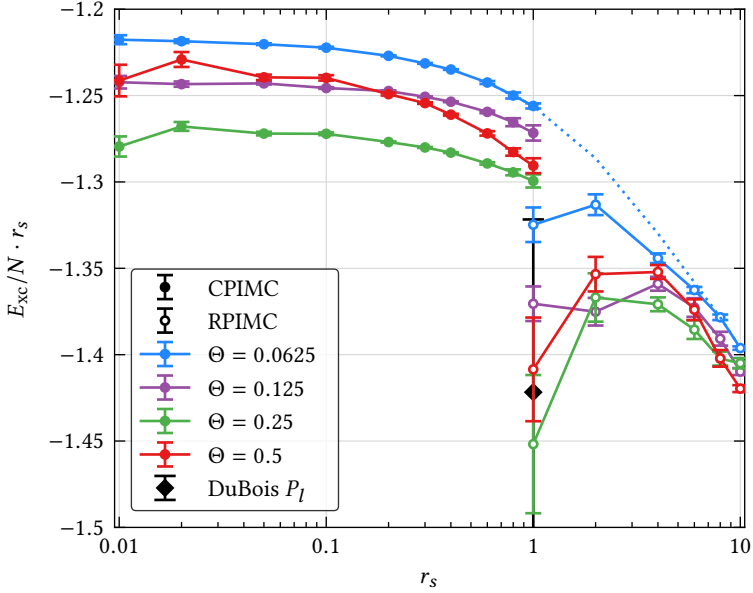
Figure 5.12: Exchange-correlation energy times r_s for 33 particles in $N_B = 2109$ basis functions for various densities and temperatures. CPIMC results denoted by filled symbols are compared to RPIMC results [55] represented by open symbols. Colors indicate temperatures from $\Theta = 0.0625$ (blue) to $\Theta = 0.5$ (red). The dotted line is a fourth order spline interpolation for $\Theta = 0.0625$ between the CPIMC results and RPIMC data with $r_s \geq 8$ (RPIMC data for $r_s > 10$ not shown here). The black diamond shows a data point taken from [123]. Figure taken from [117].

caused by the fixed-node approximation[4]. Another data point for the exchange-correlation energy by DuBios *et. al.* [123] for $r_s = 1$ and $\Theta = 0.125$ has a too large statistical error to be of any significance for this comparison. Interestingly, both methods agree in the non-monotonic behavior of the exchange-correlation energy with respect to the temperature which indicates the presence of non-trivial physics in this warm dense matter regime. The increasingly large error bars of the CPIMC results for $r_s < 1$ stem from the fact that the exchange-correlation energy is only a tiny fraction of the total energy for such nearly ideal systems,

---

[4]Recently the accuracy of the CPIMC results has been confirmed by independent DMQMC calculations [124].

especially at higher temperatures. The relative statistical uncertainty of the total energy at $r_s = 0.01$ is not larger than 0.0015 %.

5.7 Finite-Size Corrections

The CPIMC results for the exchange-correlation energy E_{XC} for $N = 33$ polarized electrons establish a valuable benchmark for other methods and could already quantify the problems of the fixed-node approximation that biases the available RPIMC results for the HEG at finite temperatures, but this is only an intermediate step for providing input data for FTDFT calculations that rely on an accurate exchange-correlation functional for the HEG in the thermodynamic limit, i.e., for $N \to \infty$ and $V \to \infty$ with $N/V = $ const. Using a finite number of particles in periodic boundary conditions inevitably introduces systematic errors that have to be corrected. For example, the energy in the thermodynamic limit is given by $E_\infty = E(N) + \Delta E$ where $E(N)$ is the energy of the periodic system and ΔE is the so-called finite-size error. Finding accurate corrections for these errors is a general problem for any method that approximates a macroscopic system by a small simulation box in periodic boundary conditions. However, existing finite-size corrections have been developed only for the ground state or at low densities [55, 113–115, 125] and their validity for finite temperatures and densities larger than r_s has to be tested.

A finite-size correction for the ground state was proposed by Chiesa *et. al.* [125] and is valid at low densities. Drummond *et. al.* [115] added a $1/r_s^2$ dependent extension for intermediate densities. The resulting corrections are given by

$$\Delta T = \frac{\omega_p}{4} + \frac{C_{3D}}{\pi r_s^2 (2N)^{1/3}} \frac{1}{(1 + \xi)^{2/3} + (1 - \xi)^{2/3}},$$
$$\Delta V = \frac{\omega_p}{4} \tag{5.2}$$

where ΔT and ΔV are the corrections to be applied to the kinetic and potential energy, respectively, ξ is the spin polarization, and $\omega_p = 2\sqrt{3/r_s^3}$ is the plasma frequency. For a simple cubic simulation cell, they find $C_{3D} = 5.264$.

For the finite-size correction of their RPIMC results at finite temperatures, Brown *et. al.* [55] proposed to multiply these expressions by a temperature dependent factor derived in the random phase approximation (RPA). For the kinetic energy the additional factor is given by $\tanh(\beta\omega_p/2)$ and for the potential energy by $\coth(\beta\omega_p/2)$. The finite-size correction will be applied to CPIMC results below. As will be shown, both are incorrect at high densities. Due to the

lack of a better alternative, a direct extrapolation to the macroscopic limit will be attempted in the following.

An advantage of the CPIMC method is, that it can be directly applied to the ideal system where the thermodynamic limit is known analytically. Note that for the ideal, finite-size HEG in the canonical ensemble, there exists no exact analytic formula that is applicable to large systems with more than $N = 30$ particles. Therefore, numerical MC simulations like CPIMC are the only available tool that can provide unbiased results for this system. The limit is identical to the ideal Fermi gas with the total energy per particle given by

$$E_\infty^{\text{id}} = (2 - \xi)\frac{r_s^3}{3\pi\beta^{5/2}}I_{3/2}(\alpha_{\text{id}})$$

with the Fermi integral I_ν defined as

$$I_\nu(\alpha) = \int_0^\infty \mathrm{d}x\,\frac{x^\nu}{e^{-\alpha+x} + 1}.$$

The ideal chemical potential $\alpha_{\text{id}} = \beta\mu_{\text{id}}$ is determined by $I_{1/2}(\alpha_{\text{id}}) = \frac{2}{3}\Theta^{-3/2}$.

For high densities, the volume $V = L^3$ of the simulation cell is very small for the particle numbers in typical QMC calculations. In fact, in the ideal limit $r_s \to 0$ the cell is even infinitesimal small, but it can be rescaled to a finite system with non-interacting electrons. Therefore, the main source of the finite-size error in the nearly ideal HEG is the discretization of the plane wave basis giving rise to large shell effects for small system sizes. This discretization error mainly effects the kinetic energy and can be greatly reduced by so-called *twist averaging* [113]. Recall that a wave function in periodic boundary conditions is only determined up to a phase factor

$$\Psi(\boldsymbol{r}_1 + L\boldsymbol{m}, \boldsymbol{r}_2, \ldots, \boldsymbol{r}_N) = e^{2\pi i\boldsymbol{\theta} L\boldsymbol{m}}\Psi(\boldsymbol{r}_1, \boldsymbol{r}_2, \ldots, \boldsymbol{r}_N)$$

with $\boldsymbol{m} \in \mathbb{Z}^3$ and $\boldsymbol{\theta} \in \mathbb{R}^3$. The so called twist angle $\boldsymbol{\theta}$ can be freely chosen as only physical observables like the density have to be periodic with L. For plane waves, above condition is satisfied by a discrete grid of wave vectors

$$\boldsymbol{k}_i = \frac{2\pi}{L}\boldsymbol{m}_i + \frac{2\pi}{L}\boldsymbol{\theta},$$

where the usual definition is obtained for $\boldsymbol{\theta} = 0$, cf. Sec. 5.1. Independent of $\boldsymbol{\theta}$, the grid is very coarse if the length of the cell L is small, e.g., for small particle number N or high densities. In the gound state, this results in a bad approximation of the
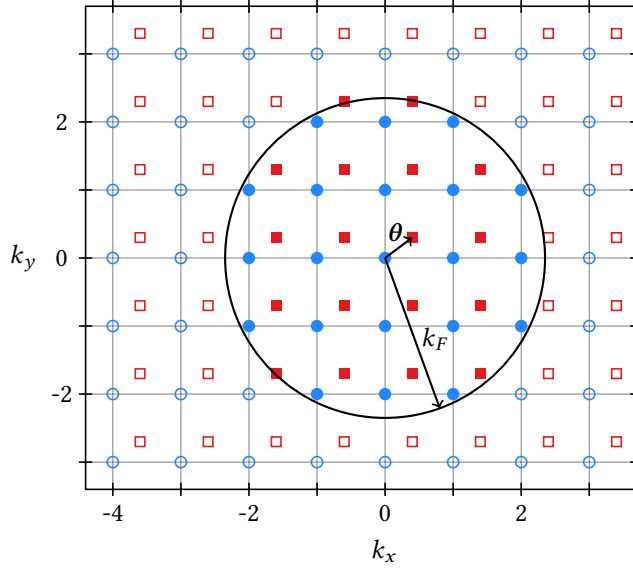
Figure 5.13: Finite twist angle. Shown is the k_x-k_y-plane of the 3D HEG in internal units (5.1). Blue circles indicate allowed wave vectors when the twist angle is zero, while the red squares correspond to a finite twist angle $\theta = (0.4, 0.3, 0)$. Filled symbols represent occupied states in the ground state of the ideal Fermi gas with $N = 57$ particles. The Fermi surface of the macroscopic system is also shown.

Fermi spheres, which is given by the spherical volume of occupied wave vectors in the macroscopic system. Fig. 5.13 shows the grid of possible wave vectors and the occupied states in the k_x-k_y plane for the ground state of the ideal HEG with $N = 57$ particles for two different twist angles and the Fermi sphere of the corresponding macroscopic system. Such a coarse approximation will lead to aforementioned shell effects, i.e., large fluctuations of expectation values with particle number and twist angle θ. At finite but low temperatures, the transition from occupied to unoccupied state is still sharp and shell effects will occur in this case, too. Because for finite system sizes the expectation value of an observable $\hat{O}$ depends on θ, an improved approximation with reduced fluctuations can be obtained by averaging over all distinct twist angles

$$\langle \hat{O} \rangle \approx \int_{-0.5}^{0.5} \mathrm{d}^3\theta \, \langle \hat{O} \rangle (\boldsymbol{\theta}). \tag{5.3}$$

Note that each twist angle can be mapped to an equivalent twist angle inside the
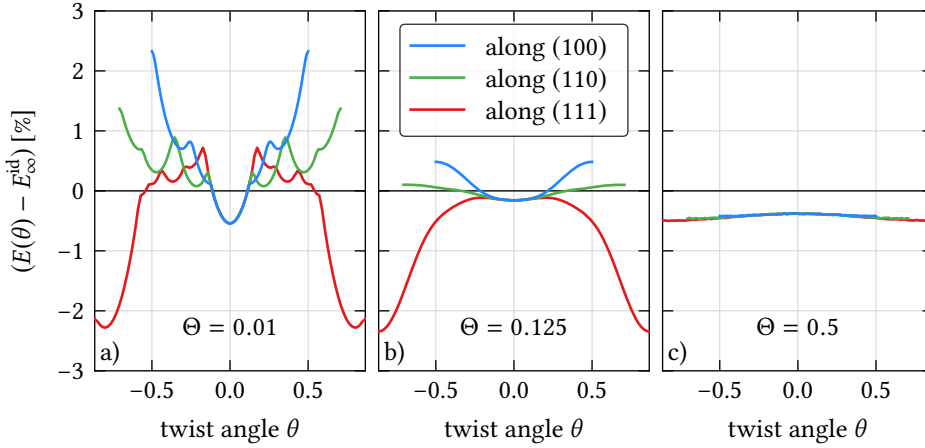
Figure 5.14: Twist angle dependence of the total energy of the ideal Fermi gas with $N = 33$ particles relative to the macroscopic limit E_∞^{id}. The energy deviation is shown along different directions and for temperatures of a) $\Theta = 0.01$, b) $\Theta = 0.125$, and c) $\Theta = 0.5$. The twist angle is given by $\boldsymbol{\theta} = \theta\boldsymbol{d}$ for $\boldsymbol{d} = (1, 0, 0)$ (blue), $\boldsymbol{d} = (1, 1, 0)$ (green), and $\boldsymbol{d} = (1, 1, 1)$ (red).

cube defined by $-0.5 \leq \theta_i < 0.5$ for $i = x, y, z$. It can be shown that this integral is exact for ideal particles in the grand-canonical ensemble [113].

The dependence of the total energy on the twist angle $\boldsymbol{\theta}$ relative to the energy of the macroscopic system is shown in Fig. 5.14 for $N = 33$ ideal particles. Each curve represents the relative energy deviation for values of $\boldsymbol{\theta} = \theta\boldsymbol{d}$ along a different direction $\boldsymbol{d}$. At low and intermediate temperatures, panel a) and b), the energy for different twist angles can deviate by several percent. The strong fluctuations at $\Theta = 0.01$ are the result of allowed wave vectors crossing the sharp Fermi surface. At $\Theta = 0.5$, panel c), the transition between occupied and unoccupied states is so smooth, that the twist angle has only a small influence on the total energy.

At low temperatures, integrating a complicated function as in Fig. 5.14 a) needs many twist angles to ensure a high degree of accuracy and hence, greatly increases the computational costs of deterministic methods like HF and CI. Fortunately, the calculation of (5.3) can be combined with the CPIMC integration of the expectation values with relatively little overhead. Similar to the ways parallel computations are performed, cf. 3.6, accumulating results that include statistical errors will further reduce the total error even if the results are obtained for dif-
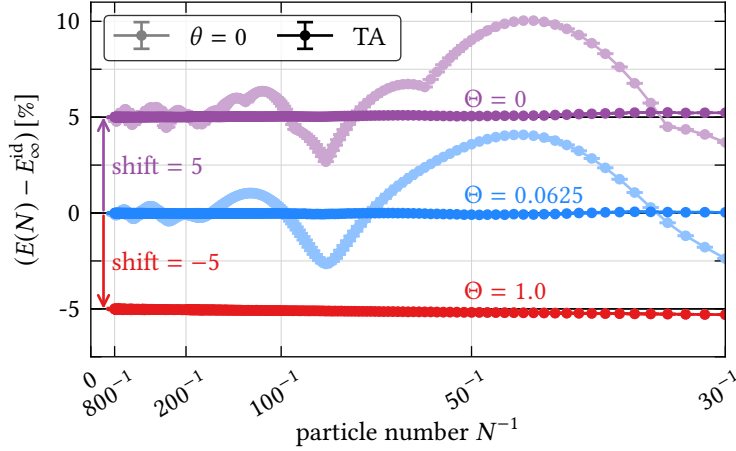
Figure 5.15: Effect of twist averaging on the total energy for the ideal Fermi gas at different temperatures. Shown is the convergence with particle number to the macroscopic limit E_∞^{id} for $\Theta = 0$ (violet), $\Theta = 0.0625$ (blue), and $\Theta = 1.0$ (red). Light colors denote results without twist averaging (hidden by the twist averaged results in case of $\Theta = 1.0$). Results denoted by normal colors have been averaged over 1×10^4 twist angles. Data for $\Theta = 0$ ($\Theta = 1.0$) are shifted upwards (downwards) by 5 percentage points.

ferent twist angles. In this work, a simple MC integration is used. This is most efficient if the error of the individual MC calculations is larger than the statistical error of the integration over $\boldsymbol{\theta}$. Thus, each individual MC simulation should be made as short as possible without wasting a significant amount of time with the equilibration steps that are necessary for each new twist angle. However, special care must be taken when averaging results from methods afflicted with a sign problem. For small average signs, the results are biased. As explained in Sec. 3.6, the bias can only be neglected for individual calculations, because it decreases linearly with the number of sample N_{MC} while the statistical error is reduced with $\sqrt{N_{\mathrm{MC}}}$, but the bias becomes significant when the many independent results are averaged. Therefore, one either has to increase the number of samples per twist angle, which reduces the number of different twist angles that can be calculated within a given amount of time, or use a high order ratio estimator, e.g., Beale's estimator.

The effect of twist averaging on the finite-size error is demonstrated in Fig. 5.15 which shows the relative deviation of the total energy in dependence of the
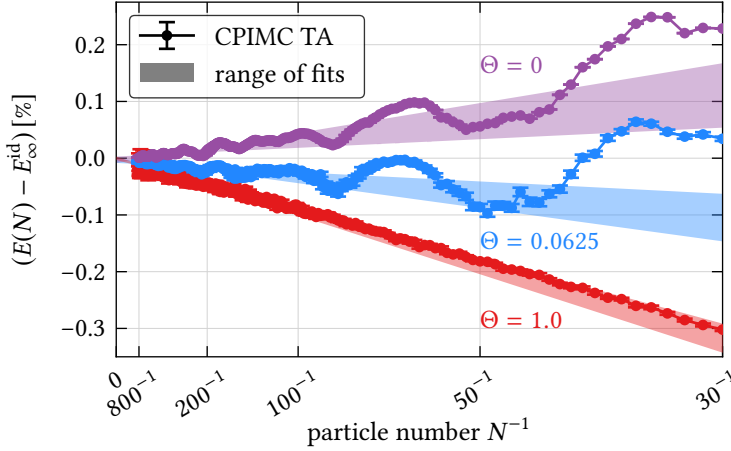
Figure 5.16: Linear extrapolation of the total energy to the thermodynamic limit for the ideal Fermi gas. Shown are the twist averaged results from Fig. 5.15 (without shift). The shaded areas mark the range of possible linear fits with every second number between 100 and 200 as starting point and every fifth number between 500 and 800 as end point.

particle number for different temperatures with and without twist averaging. In the latter case, the energy can change by several percent for different particle numbers N when the temperature is low, similar to the observed fluctuations with θ. Even for $N = 800$ particles, the fluctuations are too large to allow for a reliable extrapolation to the thermodynamic limit given by $N \to \infty$ for fixed r_s. With twist averaging, the curves are much smoother and are almost indistinguishable from a horizontal line with the correct limit at this scale. On the other hand, twist averaging has no visible effect for high temperature as shown by the red curve for $\Theta = 1$.

Zooming in reveals that even after averaging over 1×10^4 twist angles significant fluctuations of the total energy remain at low temperatures, see Fig. 5.16. As indicated by the small errorbars, averaging over more twist angles would have a negligible effect. At a relatively high temperature of $\Theta = 1$, the results form an approximately monotonic curve but have relatively large error bars for large particle numbers. In all cases, assuming a convergence with $1/N$ seems justified, but due to the fluctuations, the extrapolation to the macroscopic limit depends on the starting- and endpoint of the linear fit. Therefore, a separate fit is performed for every even particle number between $N = 100$ and $N = 200$ as the starting point and every fifth particle number between $N = 500$ and
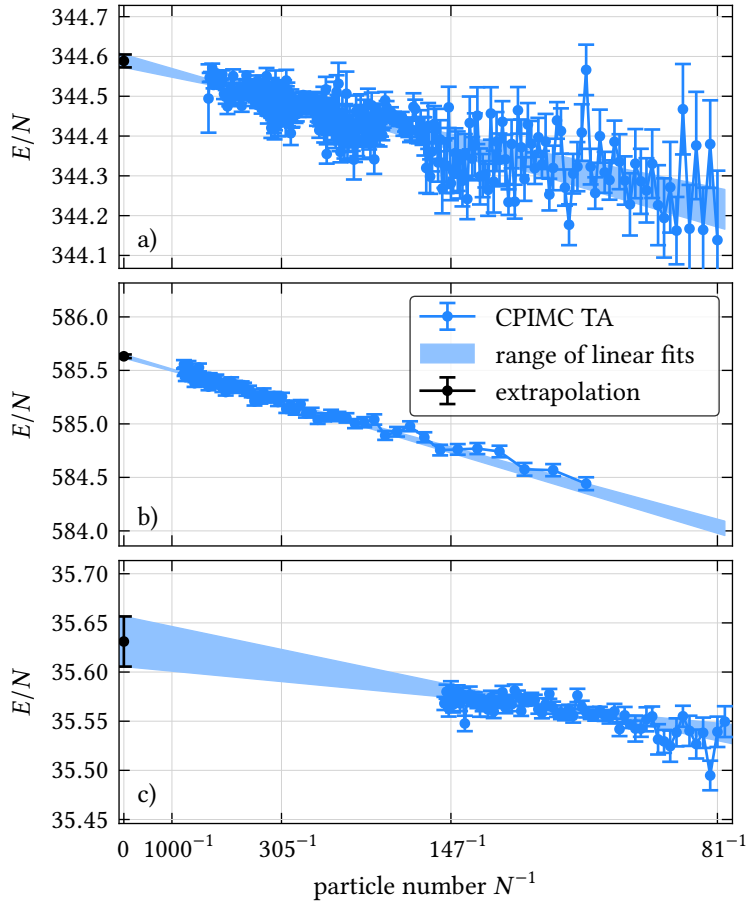
Figure 5.17: Linear extrapolation of the total energy to the thermodynamic limit. Blue points with errorbars denote CPIMC data and the shaded area marks the range of possible fits with starting points N_s and end points N_e. The black point with errorbars is an estimate from the extrapolation. a) $r_s = 0.1$, $\Theta = 0.0625$, $N_s \in [81, 200]$, and $N_e \in [400, 568]$. b) $r_s = 0.1$, $\Theta = 0.5$, $N_s \in [100, 200]$, and $N_e \in [600, 800]$. c) $r_s = 0.3$, $\Theta = 0.0625$, $N_s \in [80, 100]$, and $N_e \in [120, 150]$. In the last case, an approximation using only pairs of kinks has been used, cf. main text. The calculations have been performed in a variable number of basis function up to $N_B = 24405$. Adapted from [126].

$N = 800$ as the endpoint. The two fits with the minimum and maximum intercept are then used to draw the shaded region shown in the figure. For all three temperatures, the energy E_∞^{id} of the thermodynamic limit is enclosed between the two intercepts with relative deviations smaller than 0.01 %, confirming the validity of the asymptotic convergence with $1/N$ for the non-interacting system. However, for particle numbers $N < 100$ there are significant deviations from the linear fits. Therefore, extrapolations relying on a $1/N$ convergence for small particle numbers have only a limited reliability.

The linear convergence with $1/N$ for large particle numbers is also found for low densities as can be seen in the leading terms in Eq. (5.2). This functional dependence has been shown to be inaccurate for the kinetic energy at intermediate densities around $r_s = 1$ and thus a higher order term proportional to $1/N^{1/3}$ has been added [115]. Nevertheless, a linear extrapolation over $1/N$ is attempted for twist averaged results at high densities as shown in Fig. 5.17. For $r_s = 0.1$ and $\Theta = 0.0625$, panel a), the sign problem restricts the CPIMC calculation to less than $N = 600$ particles. Despite the large errorbars and using a wide range of starting- and endpoints, the difference between the intercepts of the fits is relative small with 0.01 %. The extrapolation for the same density but $\Theta = 0.5$ in b) closely resembles the red curve for $\Theta = 1$ in the ideal system depicted in Fig. 5.16. Here, particle numbers between $N = 100$ and $N = 200$ and $N = 600$ and $N = 800$ have been used as the starting- and endpoint of the fit, respectively. The uncertainty from the fit procedure is 0.006 %.

At lower densities, the sign problem of the CPIMC method prevents a direct extrapolation. Instead, Fig. 5.17 c) shows approximate results obtained by only allowing to add and remove pairs of kinks. As such pairs always have a sign of 1, this approximation does not suffer from the sign problem. Investigations by Simon Groth have shown, that the systematic error of this approach is small for high densities and low temperatures and should be negligible compared to the uncertainty of the finite-size extrapolation. Including up to $N = 150$ particles into the fit results in an expectedly large uncertainty of 0.15 %.

Another approach for a finite-size correction that was investigated during this work is based on the idea of using the difference of HF at finite sizes and the well known analytic HF solution for the macroscopic limit as an approximation for the finite-size effects of exact CPIMC simulations. However, there apparently exist no HF calculations for finite temperatures in the canonical but only in the grand canonical ensemble. For the particle number accessible by the CPIMC methods, the difference between the canonical and grand-canonical ensemble at finite temperatures introduces another unknown systematic error in addition to the finite-size effects that affect canonical CPIMC calculations. Thus, no finite-size
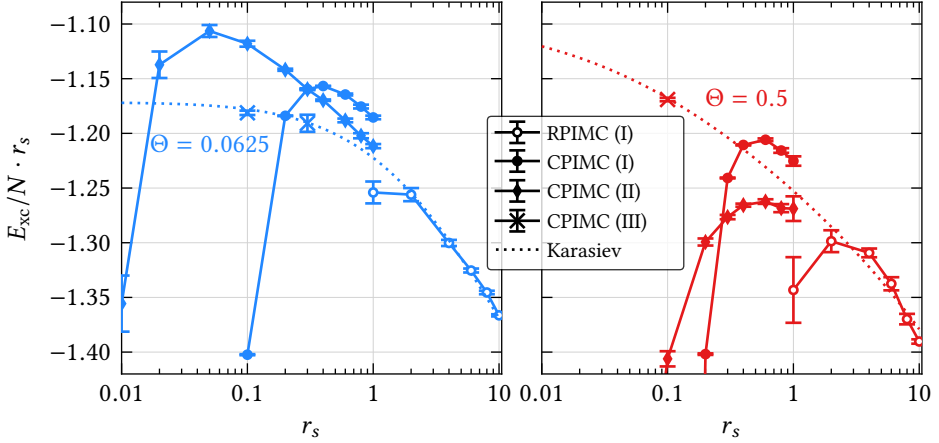
Figure 5.18: Comparison of finite-size corrections of the exchange-correlation energy for $\Theta = 0.0625$ (left) and $\Theta = 0.5$ (right). Filled symbols denote CPIMC results and open symbols correspond to RPIMC datas. (Open) circles use the finite-size correction[5] of Brown *et. al.* [55], diamonds use the ground-state finite-size correction of Drummond *et. al.* [115], and crosses show the extrapolated values from Fig.5.17. The dotted line is a fit to the RPIMC data by Karasiev *et. al.* [58]. Adapted from [117].

correction with the necessary accuracy could be obtained. It is an open question, if this approach will be useful for grand canonical CPIMC simulations using the worm algorithm.

In Fig. 5.18, the three different finite-size corrections are compared to a fit to the available RPIMC data [58] for two temperatures of $\Theta = 0.0625$ and $\Theta = 0.5$. CPIMC (I), shown by closed circles, denotes the finite temperature extension by Brown *et. al.* [55] applied to the CPIMC results from the previous section[5]. The same correction is included in the RPIMC results (open circles) that were used as a basis for the fit shown by the dotted line. The fit incorporates the correct HF limit for $r_s \to 0$. For both temperatures, the correction seems to overestimate the finite-size error at $r_s = 1$, thereby canceling the systematic error of the RPIMC calculations to some degree. For higher densities, the trend is completely unreasonable. The data points denoted CPIMC (II) have been obtained by applying

---

[5]The data in the supplementary material of [55] deviates from the formula given in the main paper. The data for the finite-size corrections at high densities have been obtained by private communication with Ethan Brown.

the ground state corrections of Drummond *et. al.* , Eq. (5.2), to properly twist averaged CPIMC results. For $r_s = 1$ at $\Theta = 0.0625$ and, interestingly, also for $\Theta = 0.5$, the results differ only slightly from the fit, albeit in different directions. However, also this finite-size correction is not applicable at high densities. The failure of existing finite-size corrections at high densities led to the attempt of directly extrapolating the CPIMC result for different particle numbers to the macroscopic limit as described above. The resulting 3 data points from Fig. 5.17 are denoted CPIMC (III) and shown by crosses. Except for an insignificant deviation at $r_s = 0.1$ and $\Theta = 0.0625$, the agreement with the fit is surprisingly good, considering that the original RPIMC data is systematically too low at intermediate densities. Still, the error bars are too large for a definitive statement about the quality of the fit and the derived exchange-correlation free energy functional.

Although the reliability of its error estimation is not entirely certain, the direct extrapolation scheme appears to yield reasonable results at high densities. However, extending this approach to lower densities and increasing its accuracy will be challenging because of the sign problem and the unknown convergence behavior in the warm dense matter regime around $r_s = 1$. Therefore, better methods for the correction of finite-size errors are urgently needed.

5.8 Recent developments

Recently, as part of this thesis, several contributions have been made to the works of Simon Groth and Tobias Dornheim, who extended the investigation of the HEG towards higher temperature and arbitrary polarizations [78, 120]. The combination of the CPIMC and PBPIMC methods yields thermodynamic results for a broad range of system parameters with an unprecedented accuracy. For the unpolarized case, all MC steps had to be adapted to avoid proposals of updates that would change the value of the total spin. For completeness, the results of these works are presented here, although the main effort has to be ascribed to my colleagues mentioned above.

Due to the sign problem, no single method is able to yield results with the necessary accuracy over the whole range of densities and temperatures in the warm dense matter regime. While the CPIMC approach is applicable to high densities, direct PIMC calculations are restricted to low densities. With the development of the PBPIMC algorithm, which extends the PIMC methods towards lower densities with a much higher accuracy than the RPIMC approximation, it is now possible to close the gap between the different approaches and provide
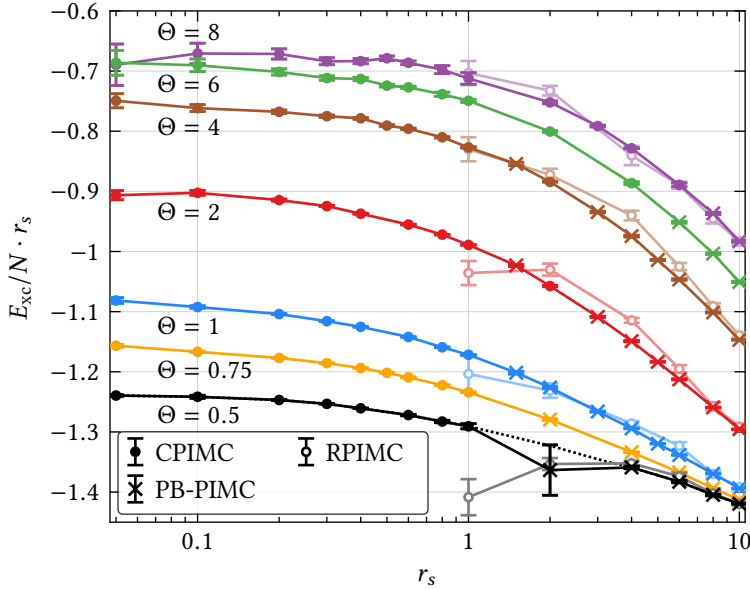
Figure 5.19: Exchange-correlation energies times r_s versus density for temperatures of $\Theta \leq 0.5$ and $N = 33$ particles. Filled circles with error bars denote CPIMC data and crosses with error bars mark results from PBPIMC calculations. Colors are associated with temperature. At each Θ and r_s value, the most accurate result of these two methods have been chosen. Additionally, open circles with error bars (light colors) show RPIMC data for comparison. For $\Theta = 0.5$, the dotted line is an interpolation between the CPIMC and PBPIMC results. Solid lines are guides to the eye. Adapted from [120].

highly accurate results for the exchange correlation energy of the finite size HEG with $N = 33$ polarized electrons for all densities and temperatures with $\Theta \geq 0.5$. Each normal colored data point in Fig. 5.19 shows the result with the smaller statistical error of either the CPIMC or the PBPIMC method. It was carefully checked that both methods are in perfect agreement with each other where there ranges of applicability overlap, cf. [77]. Note that for $\Theta = 8$, the CPIMC method is most accurate up to $r_s = 6$. For comparison, RPIMC results are shown in lighter colors, revealing significant deviations even for intermediate temperatures and densities. At $\Theta = 0.5$, the small remaining gap could be closed by a fourth order spline interpolation.

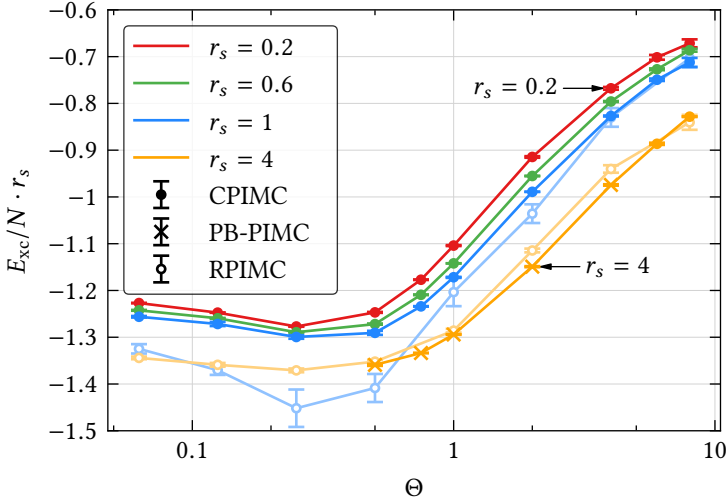In the same way, the dependence of the exchange correlation energy per

Figure 5.20: Exchange-correlation energies times r_s versus temperature for various densities distinguished by color and $N = 33$ particles. Shown are results from CPIMC (filled circles with error bars) and PBPIMC (crosses with error bars) simulations For $r_s = 1$ and $r_s = 4$, these results are compared to RPIMC data (open circles with error bars, light colors). Solid lines are guides to the eye. From [120].

particle on the degeneracy parameter Θ is shown in Fig. 5.20 for various densities. For $r_s \leq 1$, CPIMC calculations are feasible even close to the ground at $\Theta = 0.0625$. For lower densities, no accurate results are available as the PBPIMC method is restricted to temperatures with $\Theta \gtrsim 0.5$. In accordance with Figs. 5.19 and 5.12, significant, systematic deviations of the RPIMC to the CPIMC and PBPIMC results can be observed for almost all temperatures, albeit the sign of the deviations is inconsistent for different densities. The deviations are most pronounced for $r_s = 1$ and $\Theta = 0.25$. This is in similar to the non-monotonic behavior of the exchange correlation energy reaching its maximum magnitude also around intermediate temperatures between $\Theta = 0.25$ and $\Theta = 0.5$.

So far, only results for fully polarized systems have been shown. This is not a principle limitation of the CPIMC methods but was chosen only for simplicity. In fact, the formulation in Sec. 3.2 is fully general with respect to the single-particle basis of spin orbitals $|i\rangle$. However, there are two complications that need to be considered: first, not using the spin symmetry of the Hamiltonians studied in this work would result in eight times larger memory requirement
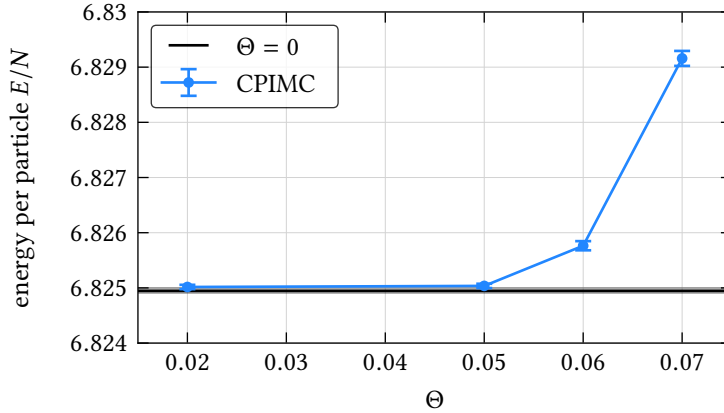
Figure 5.21: Comparison of *unpolarized* CPIMC results for $N = 14$ particles and $r_s = 0.5$ with FCIQMC data for the ground-state energy [127]. Blue circles with error bars, connected by the blue solid line, denote results from finite-temperature CPIMC calculations. The black solid line shows the FCIQMC ground-state energy per particle. The gray shaded area is the corresponding statistical uncertainty. From [126].

because the interaction matrix w_{ijkl} is a 4-tensor. Second, the trace over the full antisymmetric subspace of the Hilbert space in Eq. (2.1) automatically includes a summation over all possible values of the total spin, but typically simulations are performed with a fixed number of spin-up and spin-down particles. Therefore, several improvements of the CPIMC implementation have been made to address these issues, as mentioned above. A first test of the correctness of these changes is shown in Fig. 5.21 for the unpolarized HEG with $N = 14$ particles at $r_s = 0.5$. The convergence of the total energy per particle with temperature to the ground state energy is compared with a result from a highly accurate FCIQMC calculation [127]. As can be seen, the CPIMC energies converge nicely to the correct limit. At $\Theta = 0.02$, both methods agree perfectly within error bars of similar size. Thus, for not too large particle numbers and coupling strength, the CPIMC method is capable of providing results for zero temperature with comparable accuracy to explicit ground state methods. Of course, the correctness of the method for the unpolarized system is also confirmed by further comparisons with CI results for smaller particle numbers not shown here.

Similar to the case of the fully polarized HEG the complementarity of the CPIMC and the PBPIMC approaches allows for the calculation of highly accurate
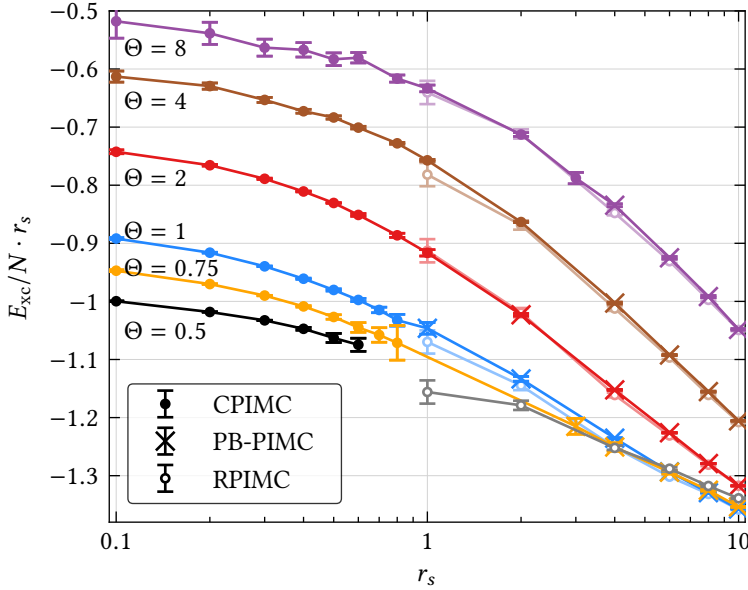
Figure 5.22: Exchange-correlation energy times r_s versus density parameter for the *unpolarized* HEG with $N = 66$ particles. Temperatures are distinguished by color. Filled circles and crosses with error bars mark CPIMC and PBPIMC results, respectively. For comparison, RPIMC results are shown by open circles with error bars and light colors. Solid lines are guides to the eye. From [78].

exchange-correlation energies of $N = 66$ unpolarized particles in the entire density range for $\Theta \geq 1$. Because of the larger particle number and a more severe sign problem of the CPIMC method in case of unpolarized particles, the gap between the approaches cannot be closed for lower temperatures. At $\Theta = 0.75$, CPIMC results with reasonable errorbars and computation times are available for $r_s \lesssim 0.7$ and PBPIMC results for $r_s \gtrsim 3$. At $\Theta = 0.5$, PBPIMC simulations are infeasible for the whole depicted density regime. The comparison with RPIMC results shows that the fixed-node error is not significant for the available temperatures, albeit at lower temperatures there are signs for a systematic drop of the exchange-correlation energy as has been found in the polarized system.

To conclude, in this chapter it has been demonstrated that the CPIMC method is a valuable tool for the investigation of the HEG at finite temperatures. The complementary behavior of the sign problem compared to PIMC methods in coordinate representation enables highly accurate, unbiased simulations of finite

polarized and unpolarized systems at high densities in the warm dense matter regime. With the help of the kink potential extrapolation technique, its range of applicability can be extended by a factor of two to three with respect to r_s while still providing small and reliable errorbars. Depending on the particle number, the CPIMC method can cover a wide range of parameters, from highly degenerate systems indistinguishable from the ground state at $\Theta = 0.01$ and $r_s = 0.5$ to almost classical systems at $\Theta = 8$ and $r_s = 10$. Because of the weak scaling of the runtime with the number of basis functions, large basis sizes can be used, yielding an incomplete basis size error that is negligible compared to the statistical error. For $N = 33$ polarized and $N = 66$ unpolarized particles, the first quasi-exact exchange-correlation energies for densities larger than $r_s = 1$ at finite temperatures could be obtained, revealing significant deviations of the existing RPIMC results. By combining the best results from CPIMC and PBPIMC calculations, the complete density range is accessible for temperatures with $\Theta \gtrsim 1$. The CPIMC method can also help to study finite-size effects in periodic boundary conditions at high densities as simulations with hundreds of particles in thousands of basis functions are feasible, corresponding to Hilbert spaces of up to $N_{\text{det}} \approx 10^{1527}$ slater determinants. If existing finite-size corrections, which are inapplicable in the high density regime, can be improved, a complete and accurate exchange-correlation free energy of the HEG at finite temperatures usable by FTDFT calculations will be within reach in the near future. For temperatures of $\Theta \gtrsim 0.5$ this was recently achieved by Simon Groth and Tobias Dornheim using a combination of the CPIMC and the PBPIMC methods with the analytic finite-temperature Singwi-Tosi-Land-Sjölander (STLS) approximation [128].

6 | Conclusions

In this work, the worm algorithm was successfully implemented for the CPIMC method[1]. In conjunction with the improvements of massively parallel computation, as well as higher acceptance ratios of the MC steps and their much more efficient calculation, this refined CPIMC method is several orders of magnitude faster for large systems than the original implementation. While its previously observed advantages for the efficiency of the MC sampling cannot be utilized to the full extent as local updates in the configuration space of the partition function are still necessary to ensure ergodicity for general Hamiltonians with arbitrary pair interactions, the worm algorithm enables grand canonical simulations with varying particle number and the direct sampling of the Matsubara Green function, which gives access to new observables and leads to a greatly improved estimator for the one-particle density. Additionally, the accuracy of the results for the Matsubara Green function will allow for the reconstruction of dynamic properties like the spectral functions from equilibrium calculations. Furthermore, variants of the MC steps that are specifically tailored to the properties of the HEG Hamiltonian resulted in an additional significant increase of the efficiency of at least the order of the basis size in the case of the HEG. The introduction of the kink potential extrapolation technique by Simon Groth further extended the range of applicability to larger coupling strengths. These enormous advances have been essential for the simulations of systems at high temperatures and the investigation of finite size errors presented in this work. The results of these applications feature an unprecedented accuracy unattainable with conventional finite-temperature methods.

The correctness of the new algorithm has been demonstrated by rigorous comparisons with CI calculations for Coulomb interacting fermions in a two-dimensional harmonic trap as well as the HEG at finite temperatures. For all investigated system parameters, the accuracy of the algorithm is verified by the

---

[1]With additional contributions from Simon Groth during his master thesis which was co-supervised as part of this research.

perfect agreement within relative statistical errors as small as 10^{-8}. In the case of the two-dimensional harmonic trap, the improved benchmarking capabilities of the CPIMC method revealed large deviations of the HF approximations for open shell configurations even at small coupling strengths and a significant deviation of the MLB method in the complete basis set limit.

As a major achievement of this work, the first quasi-exact results of the HEG at finite temperatures in the warm dense matter regime could be obtained, thereby providing the first accurate assessment of the fixed-node error of the RPIMC method. It was discovered that the systematic biases of the available RPIMC data are significant even at intermediate densities of $r_s \approx 5$ and exceed 10 % at $r_s \approx 1$. An investigation of the finite size effects utilizing extensive twist-averaging revealed a rather low quality of the available formulas for the correction of finite size errors at high densities. An attempt of a direct extrapolation to the thermodynamic limit with system sizes more than 10^{1527} determinants offered accurate estimates at $r_s = 0.1$ as a stepping stone for the necessary development of improved finite size corrections and the subsequent construction of accurate free energy exchange-correlation functionals that are essential for the successful application of FTDFT. In the light of the recent developments, which are, in parts, based on the CPIMC method and include contributions from this work, this can be considered a significant advancement for simulations of fermions in the warm dense matter regime.

Being a versatile method for gauging the accuracy of other approaches, CPIMC calculations support the search for quantum statistical simulation techniques, in particular, efficient approximations applicable to real physical systems. The exact benchmark results obtained during this work already facilitated the development of two innovative approaches for accurate simulations of fermionic quantum systems at warm dense matter conditions. Without rigorous comparisons with CPIMC data for a wide range of densities and temperatures, the quality of PBPIMC calculations using just two time slices could not have been established [30]. Likewise, large scale comparisons stimulated improvements of the DMQMC method while perfectly verifying the correctness of the CPIMC method [124]. Nevertheless, due to the initiator approximation which is necessary for DMQMC calculations of large systems, the CPIMC method remains the most accurate approach with acceptable computational requirements for a wide range of system parameters. Finally, the errors of the Montroll-Ward, the e^4, the random phase, and the Singwi-Tosi-Land-Sjölander approximation for the HEG were reliably quantified for the first time [117, 128].

Future work will include a generalization of the CPIMC implementation to several particle species and arbitrary spin polarizations, the development and

testing of approximations within the CPIMC formulation to allow for larger particle numbers and coupling strengths while still providing a high degree of accuracy, the calculation of important observables such as the momentum and pair distribution function as well as reconstructed dynamical properties using the Matsubara Green function, and, last but not least, the investigation of more complex Hamiltonians that contain, e.g., magnetic or relativistic effects.

While this thesis was focused on the application of the CPIMC method to fully spin-polarized systems, the formalism is fully general with regard to the simulation of arbitrarily fixed and also alternating spin configurations. The recent applications to the unpolarized HEG have been very successful but revealed an increased fermion sign problem compared to the polarized case. Since kinks that differ only with respect to the spin of the involved particles (but are otherwise identical) have opposite signs, the so-called blocking approach could be used to reduce the sign problem in this case.

Even if blocking for arbitrary spin polarizations is successfully employed, exact CPIMC simulations will still be restricted to intermediate densities, $r_s \lesssim 1$, for low temperatures and particle numbers required for reliable calculations of observables in the thermodynamic limit. More than doubling the range of applicability appears only to be possible with additional approximations. One approach is to propose only certain classes of diagrams in the MC updates. Very promising is the restriction to add and remove only pairs of kinks, which completely avoids the sign problem and is very accurate at low temperatures. The weights of the remaining diagrams can be either set to zero or strongly decreased. The latter technique would allow for a systematic improvement of the errors by reducing the suppression of the weights at the cost of a smaller average sign. Subsequently, one could attempt an extrapolation to the exact limit without any constraints, similar to the kink potential extrapolation.

An interesting first step towards the investigation of real two-component plasmas would be the simulation of a single atom at finite temperatures, possibly embedded in a HEG, described by an external Coulomb potential. Screening effects, the dependence of the ionization on temperature and density, and the exact potential of ions in the HEG are still open questions, which recently stimulated controversial discussions in the literature [129–137]. Exact CPIMC calculations could verify the approximate DFT results [132, 134] and yield *ab initio* data of the decreased ionization energies of atoms in warm dense matter. A challenge is the accurate description of the continuum and bound states with a single basis set. Using a mixed basis as in Refs. [138, 139] could help to alleviate this problem.

The direct simulation of the two-component plasma without approximations is extremely challenging because of the large mass difference of electrons and ions.

While in the warm dense matter regime the electrons are highly degenerate and weakly coupled and therefore well suited for the CPIMC method, the ions are still nearly classical and strongly coupled. Pure CPIMC calculations would only be feasible at much higher densities, which can be found in the crust of neutron stars or in situations where the electrons form a nearly ideal, relativistic Fermi gas. Alternatively, the CPIMC method can be applied to a dense plasma of electrons and positrons [140], which have equal mass, coupling strengths, and degeneracy. In any case, the attractive Coulomb potential does not need a special treatment due to the non-local basis function used in the CPIMC method. Relativistic effects can be added approximately, e.g., by simply using the relativistic energy-momentum dispersion relation, which can be directly employed in the CPIMC algorithm.

A | Matrix Elements of the 2D Harmonic Oscillator

In Cartesian coordinates, the eigenstates of the Hamiltonian of a single particle in a two dimensional harmonic confinement are simple product states $\phi_{n_x n_y}(x, y) = \phi_{n_x}(x)\phi_{n_y}(y)$ of the well-known solution for the one dimensional problem with the Hamiltonian $\hat{h} = \frac{1}{2}(\hat{p}^2 + \hat{x}^2)$:

$$\phi_n(x) = \sqrt{\frac{1}{n!2^n\sqrt{\pi}}}\mathcal{H}_n(x)e^{-x^2/2},$$

with the quantum number $n \geq = 0$ and the Hermite polynomials

$$\mathcal{H}_n(x) = (-1)^n e^{x^2} \frac{d^n}{dx^n} e^{-x^2},$$

and the energy eigenvalues $E = n + \frac{1}{2}$. Calculating the two-particle integrals of the Coulomb interaction in this basis is numerically challenging and it is advantageous to switch to polar coordinates with $\boldsymbol{r} = (r, \varphi)$. One then obtains the eigenstates of the two-dimensional harmonic oscillator which are known as the Fock-Darwin orbitals, see e.g., [103], and can be written as

$$\phi_{n_r m_\varphi}(r, \varphi) = R_{n_r m_\varphi}(r)\frac{e^{im_\varphi \varphi}}{\sqrt{2\pi}}, \tag{A.1}$$

$$R_{n_r m_\varphi}(r) = (-1)^{n_r}\sqrt{2}\sqrt{\frac{n_r!}{(n_r + |m_\varphi|)!}} r^{|m_\varphi|}\mathcal{L}_{n_r}^{|m_\varphi|}(r^2)e^{-r^2/2},$$

with the quantum numbers $n_r \geq 0$, $m_\varphi \in \mathbb{Z}$, the generalized Laguerre polynomials $\mathcal{L}$, and the energy $E_{n_r m_\varphi} = 2n_r + |m_\varphi| + 1$. For the sign of $R_{n_r n_\varphi}$ and all following expressions, the convention given in [105] is used. Switching to relative and center-of-mass coordinates, $\boldsymbol{R} = (\boldsymbol{r}_1 + \boldsymbol{r}_2)/\sqrt{2}$ and $\boldsymbol{r} = (\boldsymbol{r}_1 - \boldsymbol{r}_2)/\sqrt{2}$, and dropping the indices of the quantum numbers, $n_r = n$ and $m_\varphi = m$, the spacial part of

two-particle integrals can be simplified:

$$
w_{n_1 m_1 n_2 m_2 n_3 m_3 n_4 m_4}
$$

$$
= \iint \phi^*_{n_1 m_1}(\boldsymbol{r}_1)\phi^*_{n_2 m_2}(\boldsymbol{r}_2)\frac{1}{|\boldsymbol{r}_1 - \boldsymbol{r}_2|}\phi_{n_3 m_3}(\boldsymbol{r}_1)\phi_{n_4 m_4}(\boldsymbol{r}_2)\mathrm{d}\boldsymbol{r}_1\mathrm{d}\boldsymbol{r}_2
$$

$$
= \sum_{NMnm}\sum_{N'M'n'm'} M^{*NMnm}_{n_1 m_1 n_2 m_2}M^{N'M'n'm'}_{n_3 m_3 n_4 m_4}
$$

$$
\int \phi^*_{NM}(\boldsymbol{R})\phi_{N'M'}(\boldsymbol{R})\mathrm{d}\boldsymbol{R} \int \phi^*_{nm}(\boldsymbol{r})\frac{1}{\sqrt{2}|\boldsymbol{r}|}\phi_{n'm'}(\boldsymbol{r})\mathrm{d}\boldsymbol{r}
$$

$$
= \sum_{NMnmn'm'} M^{*NMnm}_{n_1 m_1 n_2 m_2}M^{NMn'm'}_{n_3 m_3 n_4 m_4}v_{nmn'm'}, \tag{A.2}
$$

where the Talmi-Brody-Moshinsky transformation coefficients are defined by, see, e.g., [104],

$$
\phi_{n_1 m_1}(\boldsymbol{r}_1)\phi_{n_2 m_2}(\boldsymbol{r}_2) = \sum_{NMnm} M^{NMnm}_{n_1 m_1 n_2 m_2}\phi_{NM}(\boldsymbol{R})\phi_{nm}(\boldsymbol{r}).
$$

The functions on the left- and right-hand side of that equation are both of the form (A.1) and only a finite number of coefficients are non-zero. Because the eigenfunctions form a basis, the integral over the center-of-mass coordinate is simply given $\delta_{N,N'}\delta_{M,M'}$. Further, the integral over the relative coordinate can be reduced to

$$
v_{nmn'm'} = \int_0^{2\pi} \frac{e^{i(m-m')\varphi}}{2\pi}\mathrm{d}\varphi \frac{1}{\sqrt{2}}\int_{-\infty}^{\infty} R_{nm}(r)R_{n'm'}(r)\frac{1}{r}r\mathrm{d}r
$$

$$
= \delta_{mm'}\frac{1}{\sqrt{2}}\int_{-\infty}^{\infty} R_{nm}(r)R_{n'm'}(r)\mathrm{d}r.
$$

The remaining integral can be numerically solved by half-open Gauss-Hermite integration [106]. An explicit expression for the Talmi-Brody-Moshinsky coefficients is given by [105]:

$$
M^{NMnm}_{n_1 m_1 n_2 m_2} = M^{N_+ n_+}_{n_{+,1} n_{+,2}}M^{N_- n_-}_{n_{-,1} n_{-,2}} \tag{A.3}
$$

with an alternative definition of the polar quantum numbers

$$
n_+ = (2n + |m| + m)/2,
$$

$$
n_- = (2n + |m| - m)/2,
$$

and the coefficients

$$M_{n_1 n_2}^{Nn} = \delta_{N+n,\, n_1+n_2} \sqrt{\frac{n_1! n_2!}{N! n!}} \left(\frac{1}{\sqrt{2}}\right)^{N+n} (-1)^{n_2} \sum_{k=\max(0,\, n_2-n)}^{\min(N,n)} (-1)^k \binom{N}{k}\binom{n}{n_2-k}.$$

The Kronecker-deltas in the coefficients reflect the conservation of energy, $E_{n_1, m_1} + E_{n_2, m_2} = E_{N, M} + E_{n, m}$, and conservation of angular momentum, $m_1 + m_2 = M + m$. Together with $\delta_{m, m'}$, the latter also implies $m_1 + m_2 = m_3 + m_4$. Using this procedure, the calculation of the two-particle matrix elements is several orders of magnitude faster than the direct integration in Cartesian coordinates.

B | Pair-distribution of the 2D Harmonic Oscillator

The Talmi-Brody-Moshinsky transformation that is described in the previous section can also be used to calculate the integral over the center-of-mass coordinate $\boldsymbol{R}$ of the pair-distribution function, Eq. (3.16), in a two-dimensional harmonic trap. Expressing the pair-distribution function $g(\boldsymbol{r}_1, \boldsymbol{r}_2)$ in terms of the reduced two-particle density matrix d_{ijkl}, Eq. (3.15), in the basis of Fock-Darwin orbitals $\phi_{n_i m_i}(\boldsymbol{r})\delta_{\sigma,\sigma_i}$, one gets

$$
g(\boldsymbol{r}) = \iint \mathrm{d}\boldsymbol{r}_1 \mathrm{d}\boldsymbol{r}_2 \, \delta\big(\boldsymbol{r} - (\boldsymbol{r}_1 - \boldsymbol{r}_2)\big) g(\boldsymbol{r}_1, \boldsymbol{r}_2)
$$

$$
= \sum_{n_1 m_1, n_2 m_2, n_3 m_3, n_4 m_4, \sigma_1, \sigma_2} d_{n_1 m_1 \sigma_1, n_2 m_2 \sigma_2, n_3 m_3 \sigma_1, n_4 m_4 \sigma_2} I_{n_1 m_1, n_2 m_2, n_3 m_3, n_4 m_4}(\boldsymbol{r}),
$$

where the integral is given by

$$
I_{n_1 m_1, n_2 m_2, n_3 m_3, n_4 m_4}(\boldsymbol{r})
$$

$$
= \iint \phi^*_{n_1 m_1}(\boldsymbol{r}_1)\phi^*_{n_2 m_2}(\boldsymbol{r}_2)\delta\big(\boldsymbol{r} - (\boldsymbol{r}_1 - \boldsymbol{r}_2)\big)\phi_{n_3 m_3}(\boldsymbol{r}_1)\phi_{n_4 m_4}(\boldsymbol{r}_2)\mathrm{d}\boldsymbol{r}_1\mathrm{d}\boldsymbol{r}_2
$$

$$
= \sum_{NMnmn'm'} M^{*NMnm}_{n_1 m_1 n_2 m_2} M^{NMn'm'}_{n_3 m_3 n_4 m_4} v_{nmn'm'}(\boldsymbol{r}).
$$

The last step is an analog switch to center-of-mass and relative coordinates as in Eq. (A.2) with the Talmi-Brody-Moshinsky coefficients given by Eq. (A.3). While the center-of-mass integration is again trivial, the remaining integral over the relative coordinate reads

$$
v_{nmn'm'}(\boldsymbol{r}) = \int \phi^*_{nm}(\boldsymbol{r}')\delta(\boldsymbol{r} - \sqrt{2}\boldsymbol{r}')\phi_{n'm'}(\boldsymbol{r}')\mathrm{d}\boldsymbol{r}'
$$

$$
= \frac{1}{2}\frac{e^{i(m'-m)\varphi}}{2\pi}R_{nm}\left(\frac{r}{\sqrt{2}}\right)R_{n'm'}\left(\frac{r}{\sqrt{2}}\right).
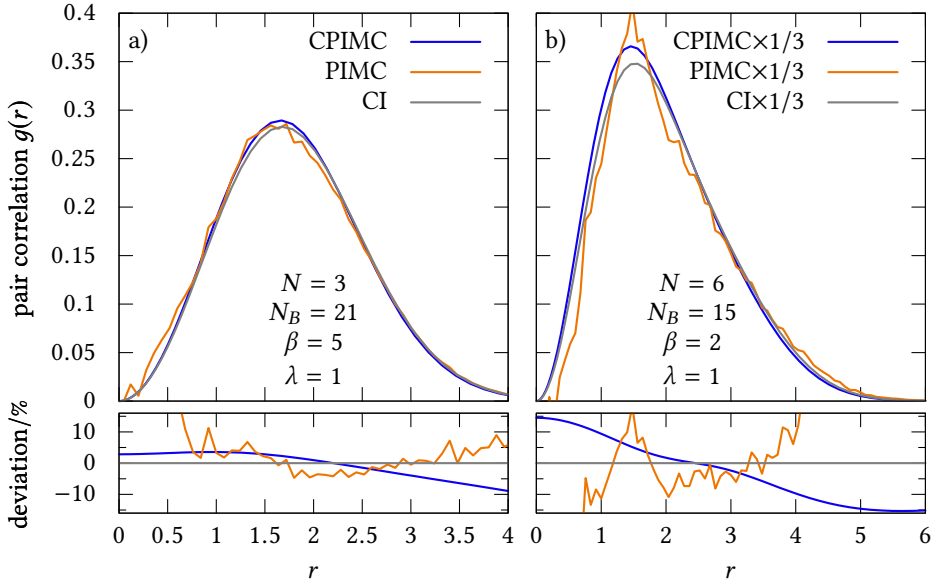$$

Figure B.1: Pair-distribution of a) $N = 3$ and b) $N = 6$ fermions in a 2D harmonic
trap with coupling parameter $\lambda = 1$ and inverse temperature $\beta = 5$
and $\beta 2$, respectively. Shown are results from CPIMC (blue line), CI
(gray line), and PIMC (orange line) calculations, the former two using
$N_B = 21$ and $N_B = 15$ basis functions, respectively. The bottom panels
show the respective deviation from the CI data. The CPIMC results
have been obtained with the implementation from [31] and the PIMC
data are from Alexey Filinov.

As the reduced two-particle density matrix has the same symmetry as the Hamil-
tonian, it is subject to the same constraints. In particular, for all terms that lead
to $m \neq m'$, it is $d = 0$. Therefore, one can replace $v_{nmn'm'}(r)$ with

$$v_{nmn'm'}(r) = \delta_{m,m'} \frac{1}{4\pi} R_{nm}\left(\frac{r}{\sqrt{2}}\right) R_{n'm'}\left(\frac{r}{\sqrt{2}}\right)$$

in above expression and the relative pair-distribution becomes a function of the
particle distance only, $g(r) = g(r)$.

A comparison of the pair-distribution function from CPIMC, CI, and PIMC
calculations for $N = 3$ and $N = 6$ particles is shown in Fig. B.1. The curves
are in reasonable agreement with each other and clearly show the correlation
hole at $r = 0$ due to the Pauli blocking. In contrast to a homogeneous system,

the pair-distribution approaches zero for large distances in the finite trap. The statistical errors of the CPIMC and PIMC data are of the same order of magnitude. For $N = 6$, a part of the deviation of the PIMC data can be explained by the small basis size of only $N_B = 15$, which is a restriction of the CI method. The reason for the relatively large deviation of the CPIMC results are the large fluctuations of the estimator for the reduced two-particle density, Eqs. (3.13) and (3.14). It should be possible to derive a much better estimator from the two-particle Green function, which can be sampled using two worms as explained in Sec. 3.9.

Bibliography

[1]   P. A. M. Dirac, "Quantum Mechanics of Many-Electron Systems", Proceedings of the Royal Society of London. Series A, Containing Papers of a Mathematical and Physical Character **123**, 714 (1929).

[2]   V. Fock, "Näherungsmethode zur Lösung des quantenmechanischen Mehrkörperproblems", Z. Physik **61**, 126 (1930).

[3]   N. D. Mermin, "Thermal Properties of the Inhomogeneous Electron Gas", Phys. Rev. **137**, A1441 (1965).

[4]   C. Møller, and M. S. Plesset, "Note on an Approximation Treatment for Many-Electron Systems", Phys. Rev. **46**, 618 (1934).

[5]   D. Bohm, and D. Pines, "A Collective Description of Electron Interactions: III. Coulomb Interactions in a Degenerate Electron Gas", Phys. Rev. **92**, 609 (1953).

[6]   E. W. Montroll, and J. C. Ward, "Quantum Statistics of Interacting Particles; General Theory and Some Remarks on Properties of an Electron Gas", Physics of Fluids (1958-1988) **1**, 55 (1958).

[7]   J. Vorberger, M. Schlanges, and W. D. Kraeft, "Equation of state for weakly coupled quantum plasmas", Phys. Rev. E **69**, 046407 (2004).

[8]   K. S. Singwi, M. P. Tosi, R. H. Land, and A. Sjölander, "Electron Correlations at Metallic Densities", Phys. Rev. **176**, 589 (1968).

[9]   S. Tanaka, and S. Ichimaru, "Thermodynamics and Correlational Properties of Finite-Temperature Electron Liquids in the Singwi-Tosi-Land-Sjölander Approximation", J. Phys. Soc. Jpn. **55**, 2278 (1986).

[10] M. Bonitz, *Quantum kinetic theory*, 2nd ed. (Springer, 2016), 406 pp.

[11] K. Balzer, and M. Bonitz, *Nonequilibrium green's functions approach to inhomogeneous systems* (Springer, Dec. 14, 2012).

[12] P. Hohenberg, and W. Kohn, "Inhomogeneous Electron Gas", Phys. Rev. **136**, B864 (1964).

[13] W. Kohn, and L. J. Sham, "Self-Consistent Equations Including Exchange and Correlation Effects", Phys. Rev. **140**, A1133 (1965).

[14] W. Kohn, "Nobel Lecture: Electronic structure of matter—wave functions and density functionals", Rev. Mod. Phys. **71**, 1253 (1999).

[15] A. J. Cohen, P. Mori-Sánchez, and W. Yang, "Challenges for Density Functional Theory", Chem. Rev. **112**, 289 (2012).

[16] J. P. Perdew, and A. Zunger, "Self-interaction correction to density-functional approximations for many-electron systems", Phys. Rev. B **23**, 5048 (1981).

[17] E. Runge, and E. K. U. Gross, "Density-Functional Theory for Time-Dependent Systems", Phys. Rev. Lett. **52**, 997 (1984).

[18] D. Ceperley, G. Chester, and M. Kalos, "Monte Carlo study of the ground state of bosons interacting with Yukawa potentials", Physical Review B **17**, 1070 (1978).

[19] E. Pollock, and D. Ceperley, "Simulation of quantum many-body systems by path-integral methods", Physical Review B **30**, 2555 (1984).

[20] M. Boninsegni, N. Prokof'ev, and B. Svistunov, "Worm Algorithm for Continuous-Space Path Integral Monte-Carlo Simulations", Phys. Rev. Lett. **96**, 070601 (2006).

[21] E. Loh, J. Gubernatis, R. Scalettar, S. White, D. Scalapino, and R. Sugar, "Sign problem in the numerical simulation of many-electron systems", Physical Review B **41**, 9301 (1990).

[22] M. Troyer, and U.-J. Wiese, "Computational Complexity and Fundamental Limitations to Fermionic Quantum Monte Carlo Simulations", Phys. Rev. Lett. **94**, 170201 (2005).

[23] D. M. Ceperley, "Fermion nodes", J. Stat. Phys. **63**, 1237 (1991).

[24] R. J. Needs, M. D. Towler, N. D. Drummond, and P. L. Ríos, "Continuum variational and diffusion quantum Monte Carlo calculations", J. Phys.: Condens. Matter **22**, 023201 (2010).

[25] I. G. Gurtubay, R. Gaudoin, and J. M. Pitarke, "Benchmark quantum Monte Carlo calculations of the ground-state kinetic, interaction and total energy of the three-dimensional electron gas", J. Phys.: Condens. Matter **22**, 065501 (2010).

[26] D. M. Ceperley, "Path-integral calculations of normal liquid ^{3}He", Phys. Rev. Lett. **69**, 331 (1992).

[27] B. Militzer, and D. M. Ceperley, "Path Integral Monte Carlo Calculation of the Deuterium Hugoniot", Phys. Rev. Lett. **85**, 1890 (2000).

[28] K. P. Driver, and B. Militzer, "All-Electron Path Integral Monte Carlo Simulations of Warm Dense Matter: Application to Water and Carbon Plasmas", Phys. Rev. Lett. **108**, 115502 (2012).

[29] R. Egger, W. Häusler, C. H. Mak, and H. Grabert, "Crossover from Fermi Liquid to Wigner Molecule Behavior in Quantum Dots", Phys. Rev. Lett. **82**, 3320 (1999).

[30] T. Dornheim, S. Groth, A. Filinov, and M. Bonitz, "Permutation blocking path integral Monte Carlo: a highly efficient approach to the simulation of strongly degenerate non-ideal fermions", New J. Phys. **17**, 073017 (2015).

[31] T. Schoof, *Thermodynamische Eigenschaften entarteter, korrelierter Fermionen*, Diplomarbeit (Kiel, 2011).

[32] T. Schoof, S. Groth, and M. Bonitz, "Introduction to Configuration Path Integral Monte Carlo", in *Complex plasmas: scientific challenges and technological opportunities*, edited by M. Bonitz, J. Lopez, K. Becker, and H. Thomsen, (Springer, New York, 2014), pp. 153–194.

[33] T. Schoof, M. Bonitz, A. Filinov, D. Hochstuhl, and J. Dufty, "Configuration Path Integral Monte Carlo", Contrib. Plasma Phys. **51**, 687 (2011).

[34] N. V. Prokof'ev, B. V. Svistunov, and I. S. Tupitsyn, "Exact quantum Monte Carlo process for the statistics of discrete systems", Jetp Lett. **64**, 911 (1996).

[35] N. V. Prokof'ev, and B. V. Svistunov, "Bold diagrammatic Monte Carlo: A generic sign-problem tolerant technique for polaron models and possibly interacting many-body problems", Phys. Rev. B **77**, 125101 (2008).

[36] E. Gull, A. J. Millis, A. I. Lichtenstein, A. N. Rubtsov, M. Troyer, and P. Werner, "Continuous-time Monte Carlo methods for quantum impurity models", Rev. Mod. Phys. **83**, 349 (2011).

[37] K. Van Houcke, F. Werner, E. Kozik, N. Prokof'ev, B. Svistunov, M. J. H. Ku, A. T. Sommer, L. W. Cheuk, A. Schirotzek, and M. W. Zwierlein, "Feynman diagrams versus Fermi-gas Feynman emulator", Nature Physics **8**, 366 (2012).

[38] G. H. Booth, A. J. W. Thom, and A. Alavi, "Fermion Monte Carlo without fixed nodes: A game of life, death, and annihilation in Slater determinant space", J. Chem. Phys. **131**, 054106 (2009).

[39] N. S. Blunt, T. W. Rogers, J. S. Spencer, and W. M. C. Foulkes, "Density-matrix quantum Monte Carlo method", Phys. Rev. B **89**, 245124 (2014).

[40] N. V. Prokof'ev, B. V. Svistunov, and I. S. Tupitsyn, ""Worm" algorithm in quantum Monte Carlo simulations", Phys. Lett. A **238**, 253 (1998).

[41] A. Filinov, and M. Bonitz, "Collective and single-particle excitations in two-dimensional dipolar Bose gases", Phys. Rev. A **86**, 043628 (2012).

[42] A. V. Filinov, M. Bonitz, and Y. E. Lozovik, "Wigner Crystallization in Mesoscopic 2D Electron Systems", Phys. Rev. Lett. **86**, 3851 (2001).

[43] S. M. Reimann, and M. Manninen, "Electronic structure of quantum dots", Rev. Mod. Phys. **74**, 1283 (2002).

[44] A. Ghosal, A. D. Güçlü, C. J. Umrigar, D. Ullmo, and H. U. Baranger, "Correlation-induced inhomogeneity in circular quantum dots", Nat Phys **2**, 336 (2006).

[45] Y. Yan, and D. Blume, "Abnormal Superfluid Fraction of Harmonically Trapped Few-Fermion Systems", Physical Review Letters **112**, 235301 (2014).

[46] J. W. Abraham, M. Bonitz, C. McDonald, G. Orlando, and T. Brabec, "Quantum breathing mode of trapped systems in one and two dimensions", New J. Phys. **16**, 013001 (2014).

[47] O. A. Hurricane, et al., "Fuel gain exceeding unity in an inertially confined fusion implosion", Nature **506**, 343 (2014).

[48] R. Nora, et al., "Gigabar Spherical Shock Generation on the OMEGA Laser", Phys. Rev. Lett. **114**, 045001 (2015).

[49] P. F. Schmit, et al., "Understanding Fuel Magnetization and Mix Using Secondary Nuclear Reactions in Magneto-Inertial Fusion", Phys. Rev. Lett. **113**, 155004 (2014).

[50] R. Ernstorfer, M. Harb, C. T. Hebeisen, G. Sciaini, T. Dartigalongue, and R. J. D. Miller, "The Formation of Warm Dense Matter: Experimental Evidence for Electronic Bond Hardening in Gold", Science **323**, 1033 (2009).

[51] M. D. Knudson, M. P. Desjarlais, R. W. Lemke, T. R. Mattsson, M. French, N. Nettelmann, and R. Redmer, "Probing the Interiors of the Ice Giants: Shock Compression of Water to 700 GPa and 3.8 g/cm^3", Phys. Rev. Lett. **108**, 091102 (2012).

[52] N. Nettelmann, A. Becker, B. Holst, and R. Redmer, "Jupiter Models with Improved Ab Initio Hydrogen Equation of State (H-REOS.2)", ApJ **750**, 52 (2012).

[53] P. F. Bedaque, E. Berkowitz, and S. Sen, "Thermodynamics of nuclear condensates and phase transitions in white dwarfs", Phys. Rev. D **89**, 045010 (2014).

[54] V. V. Karasiev, L. Calderín, and S. B. Trickey, "Importance of finite-temperature exchange correlation for warm dense matter calculations", Phys. Rev. E **93**, 063207 (2016).

[55] E. W. Brown, B. K. Clark, J. L. DuBois, and D. M. Ceperley, "Path-Integral Monte-Carlo Simulation of the Warm Dense Homogeneous Electron Gas", Phys. Rev. Lett. **110**, 146405 (2013).

[56] E. W. Brown, J. L. DuBois, M. Holzmann, and D. M. Ceperley, "Exchange-correlation energy for the three-dimensional homogeneous electron gas at arbitrary temperature", Phys. Rev. B **88**, 081102 (2013).

[57] T. Sjostrom, and J. Dufty, "Uniform electron gas at finite temperatures", Phys. Rev. B **88**, 115123 (2013).

[58] V. V. Karasiev, T. Sjostrom, J. Dufty, and S. B. Trickey, "Accurate Homogeneous Electron Gas Exchange-Correlation Free Energy for Local Spin-Density Calculations", Phys. Rev. Lett. **112**, 076403 (2014).

[59] S. Groth, *Strongly Degenerate Nonideal Fermi Sytems: Configuration Path Integral Monte Carlo Simulation*, Master Thesis (Kiel, 2014).

[60] T. Helgaker, P. Jorgensen, and J. Olsen, *Molecular Electronic-Structure theory*, 1st ed. (Wiley, 2000).

[61] C. Andrieu, N. de Freitas, A. Doucet, and M. I. Jordan, "An Introduction to MCMC for Machine Learning", Mach. Learn. **50**, 5 (2003).

[62] N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, and E. Teller, "Equation of State Calculations by Fast Computing Machines", J. Chem. Phys. **21**, 1087 (1953).

[63] W. K. Hastings, "Monte Carlo sampling methods using Markov chains and their applications", Biometrika **57**, 97 (1970).

[64] W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, *Numerical recipes 3rd edition: the art of scientific computing*, 3 edition (Cambridge University Press, Cambridge, UK ; New York, Sept. 10, 2007), 1256 pp.

[65] M. Michel, S. C. Kapfer, and W. Krauth, "Generalized event-chain Monte Carlo: Constructing rejection-free global-balance algorithms from infinitesimal steps", J. Chem. Phys. **140**, 054116 (2014).

[66] D. M. Ceperley, "Path integrals in the theory of condensed helium", Rev. Mod. Phys. **67**, 279 (1995).

[67] D. Chandler, and P. G. Wolynes, "Exploiting the isomorphism between quantum theory and classical statistical mechanics of polyatomic fluids", J. Chem. Phys. **74**, 4078 (1981).

[68] H. F. Trotter, "On the Product of Semi-Groups of Operators", Proc. Amer. Math. Soc. **10**, 545 (1959).

[69] S. A. Chin, and C. R. Chen, "Gradient symplectic algorithms for solving the Schrödinger equation with time-dependent potentials", J. Chem. Phys. **117**, 1409 (2002).

[70] K. Sakkos, J. Casulleras, and J. Boronat, "High order Chin actions in path integral Monte Carlo", J. Chem. Phys. **130**, 204109 (2009).

[71] W. Krauth, "Quantum Monte Carlo Calculations for a Large Number of Bosons in a Harmonic Trap", Phys. Rev. Lett. **77**, 3695 (1996).

[72] F. Mezzacapo, and M. Boninsegni, "Structure, superfluidity, and quantum melting of hydrogen clusters", Phys. Rev. A **75**, 033201 (2007).

[73] H. Li, N. Blinov, P.-N. Roy, and R. J. L. Roy, "Path-integral Monte Carlo simulation of $v3$ vibrational shifts for $CO2$ in $(He)n$ clusters critically tests the He–$CO2$ potential energy surface", J. Chem. Phys. **130**, 144305 (2009).

[74] A. Filinov, N. V. Prokof'ev, and M. Bonitz, "Berezinskii-Kosterlitz-Thouless Transition in Two-Dimensional Dipole Systems", Phys. Rev. Lett. **105**, 070401 (2010).

[75] T. Dornheim, A. Filinov, and M. Bonitz, "Superfluidity of strongly correlated bosons in two- and three-dimensional traps", Phys. Rev. B **91**, 054503 (2015).

[76] V. S. Filinov, "Analytical contradictions of the fixed-node density matrix", High Temp **52**, 615 (2014).

[77]   T. Dornheim, T. Schoof, S. Groth, A. Filinov, and M. Bonitz, "Permutation blocking path integral Monte Carlo approach to the uniform electron gas at finite temperature", J. Chem. Phys. **143**, 204101 (2015).

[78]   T. Dornheim, S. Groth, T. Schoof, C. Hann, and M. Bonitz, "*Ab initio* quantum Monte Carlo simulations of the uniform electron gas without fixed nodes: The unpolarized case", Phys. Rev. B **93**, 205134 (2016).

[79]   C. H. Mak, R. Egger, and H. Weber-Gottschick, "Multilevel Blocking Approach to the Fermion Sign Problem in Path-Integral Monte Carlo Simulations", Phys. Rev. Lett. **81**, 4533 (1998).

[80]   F. D. Malone, N. S. Blunt, J. J. Shepherd, D. K. K. Lee, J. S. Spencer, and W. M. C. Foulkes, "Interaction picture density matrix quantum Monte Carlo", J. Chem. Phys. **143**, 044116 (2015).

[81]   T. Nakamura, "Vanishing of the negative-sign problem of quantum Monte Carlo simulations in one-dimensional frustrated spin systems", Phys. Rev. B **57**, R3197 (1998).

[82]   A. P. Lyubartsev, "Simulation of excited states and the sign problem in the path integral Monte Carlo method", J. Phys. A: Math. Gen. **38**, 6659 (2005).

[83]   T. L. Schmidt, P. Werner, L. Mühlbacher, and A. Komnik, "Transient dynamics of the Anderson impurity model out of equilibrium", Phys. Rev. B **78**, 235110 (2008).

[84]   G. Cohen, E. Gull, D. R. Reichman, and A. J. Millis, "Taming the Dynamical Sign Problem in Real-Time Evolution of Quantum Many-Body Problems", Phys. Rev. Lett. **115**, 266802 (2015).

[85]   F. Schwabl, *Quantenmechanik (QM I): Eine Einführung*, 7. Aufl. (Springer, Berlin, 2007).

[86]   F. Schwabl, *Quantenmechanik für Fortgeschrittene (QM II)*, 5., erw. u. aktual. Aufl. (Springer, Berlin, 2008).

[87]   N. V. Prokof'ev, B. V. Svistunov, and I. S. Tupitsyn, "Exact, complete, and universal continuous-time worldline Monte Carlo approach to the statistics of discrete quantum systems", J. Exp. Theor. Phys. **87**, 310 (1998).

[88]   M. Bonitz, and D. Semkat, eds., *Introduction to computational methods in many body physics* (Rinton Press Inc, 2006).

[89]   A. A. Holmes, H. J. Changlani, and C. J. Umrigar, "Efficient Heat-Bath Sampling in Fock Space", J. Chem. Theory Comput. **12**, 1561 (2016).

[90] D. A. Patterson, G. Gibson, and R. H. Katz, "A Case for Redundant Arrays of Inexpensive Disks (RAID)", SIGMOD Rec. **17**, 109 (1988).

[91] M. Matsumoto, and T. Nishimura, "Mersenne Twister: A 623-dimensionally Equidistributed Uniform Pseudo-random Number Generator", ACM Trans. Model. Comput. Simul. **8**, 3 (1998).

[92] E. M. L. Beale, "Some uses of computers in operational research", Industrielle Organisation **31**, 27 (1962).

[93] M. Tin, "Comparison of Some Ratio Estimators", J. Am. Stat. Assoc. **60**, 294 (1965).

[94] R. C. Ogliore, G. R. Huss, and K. Nagashima, "Ratio estimation in SIMS analysis", Nucl. Instr. Meth. Phys. Res. B **269**, 1910 (2011).

[95] N. Prokof'ev, and B. Svistunov, "Worm Algorithms for Classical Statistical Models", Phys. Rev. Lett. **87**, 160601 (2001).

[96] S. Trotzky, L. Pollet, F. Gerbier, U. Schnorrberger, I. Bloch, N. V. Prokof'ev, B. Svistunov, and M. Troyer, "Suppression of the critical temperature for superfluidity near the Mott transition", Nat Phys **6**, 998 (2010).

[97] L. Pollet, "Recent developments in quantum Monte Carlo simulations with applications for cold gases", Rep. Prog. Phys. **75**, 094501 (2012).

[98] E. Gull, D. R. Reichman, and A. J. Millis, "Bold-line diagrammatic Monte Carlo method: General formulation and application to expansion around the noncrossing approximation", Phys. Rev. B **82**, 075109 (2010).

[99] P. Gunacker, M. Wallerberger, E. Gull, A. Hausoel, G. Sangiovanni, and K. Held, "Continuous-time quantum Monte Carlo using worm sampling", Phys. Rev. B **92**, 155102 (2015).

[100] K. Van Houcke, S. M. A. Rombouts, and L. Pollet, "Quantum Monte Carlo simulation in the canonical ensemble at finite temperature", Phys. Rev. E **73**, 056703 (2006).

[101] T. Ohgoe, and N. Kawashima, "Quantum Monte Carlo method for pairing phenomena: Supercounterfluid of two-species Bose gases in optical lattices", Phys. Rev. A **83**, 023622 (2011).

[102] K. Balzer, M. Bonitz, R. van Leeuwen, A. Stan, and N. E. Dahlen, "Nonequilibrium Green's function approach to strongly correlated few-electron quantum dots", Phys. Rev. B **79**, 245306 (2009).

[103] M. Rontani, C. Cavazzoni, D. Bellucci, and G. Goldoni, "Full configuration interaction approach to the few-electron problem in artificial atoms", J. Chem. Phys. **124**, 124102 (2006).

[104] L. M. Robledo, "Separable approximation to two-body matrix elements", Phys. Rev. C **81**, 044312 (2010).

[105] J. C. Cremon, *Quantum few-body physics with the configuration interaction approach: method development and application to physical systems*, OCLC: 838063222, PhD thesis (Div. of Mathematical Physics, LTH, Lund University, Lund, 2010).

[106] J. Ball, "Half-Range Generalized Hermite Polynomials and the Related Gaussian Quadratures", SIAM J. Numer. Anal. **40**, 2311 (2002).

[107] A. S. Mishchenko, N. V. Prokof'ev, A. Sakamoto, and B. V. Svistunov, "Diagrammatic quantum Monte Carlo study of the Fröhlich polaron", Phys. Rev. B **62**, 6317 (2000).

[108] E. Vitali, M. Rossi, L. Reatto, and D. E. Galli, "Ab initio low-energy dynamics of superfluid and solid ^{4}He", Phys. Rev. B **82**, 174510 (2010).

[109] D. M. Ceperley, and B. J. Alder, "Ground State of the Electron Gas by a Stochastic Method", Phys. Rev. Lett. **45**, 566 (1980).

[110] J. D. Lindl, P. Amendt, R. L. Berger, S. G. Glendinning, S. H. Glenzer, S. W. Haan, R. L. Kauffman, O. L. Landen, and L. J. Suter, "The physics basis for ignition using indirect-drive targets on the National Ignition Facility", Phys. Plasmas **11**, 339 (2004).

[111] S. X. Hu, B. Militzer, V. N. Goncharov, and S. Skupsky, "First-principles equation-of-state table of deuterium for inertial confinement fusion applications", Phys. Rev. B **84**, 224109 (2011).

[112] M. R. Gomez, et al., "Experimental Demonstration of Fusion-Relevant Conditions in Magnetized Liner Inertial Fusion", Phys. Rev. Lett. **113**, 155003 (2014).

[113] C. Lin, F. H. Zong, and D. M. Ceperley, "Twist-averaged boundary conditions in continuum quantum Monte Carlo algorithms", Phys. Rev. E **64**, 016702 (2001).

[114] L. M. Fraser, W. M. C. Foulkes, G. Rajagopal, R. J. Needs, S. D. Kenny, and A. J. Williamson, "Finite-size effects and Coulomb interactions in quantum Monte Carlo calculations for homogeneous systems with periodic boundary conditions", Phys. Rev. B **53**, 1814 (1996).

[115]   N. D. Drummond, R. J. Needs, A. Sorouri, and W. M. C. Foulkes, "Finite-size errors in continuum quantum Monte Carlo calculations", Phys. Rev. B **78**, 125106 (2008).

[116]   J. J. Shepherd, G. Booth, A. Grüneis, and A. Alavi, "Full configuration interaction perspective on the homogeneous electron gas", Phys. Rev. B **85**, 081103 (2012).

[117]   T. Schoof, S. Groth, J. Vorberger, and M. Bonitz, "*Ab initio* thermodynamic results for the degenerate electron gas at finite temperature", Phys. Rev. Lett. **115**, 130402 (2015).

[118]   S. Ichimaru, H. Iyetomi, and S. Tanaka, "Statistical physics of dense plasmas: Thermodynamics, transport coefficients and dynamic correlations", Phys. Rep. **149**, 91 (1987).

[119]   T. Schoof, S. Groth, and M. Bonitz, "Towards ab initio thermodynamics of the electron gas at strong degeneracy", Contrib. Plasma Phys. **55**, 136 (2015).

[120]   S. Groth, T. Schoof, T. Dornheim, and M. Bonitz, "*Ab initio* quantum Monte Carlo simulations of the uniform electron gas without fixed nodes", Phys. Rev. B **93**, 085102 (2016).

[121]   A. Grüneis, J. J. Shepherd, A. Alavi, D. P. Tew, and G. H. Booth, "Explicitly correlated plane waves: Accelerating convergence in periodic wavefunction expansions", J. Chem. Phys. **139**, 084112 (2013).

[122]   N. Prokof'ev, and B. Svistunov, "Fermi-polaron problem: Diagrammatic Monte Carlo method for divergent sign-alternating series", Phys. Rev. B **77**, 020408 (2008).

[123]   J. L. DuBois, E. W. Brown, and B. J. Alder, "Overcoming the fermion sign problem in homogeneous systems", arXiv:1409.3262 [cond-mat] (2014).

[124]   F. D. Malone, N. S. Blunt, E. W. Brown, D. K. K. Lee, J. S. Spencer, W. M. C. Foulkes, and J. J. Shepherd, "Accurate Exchange-Correlation Energies for the Warm Dense Electron Gas", Phys. Rev. Lett. **117**, 115701 (2016).

[125]   S. Chiesa, D. M. Ceperley, R. M. Martin, and M. Holzmann, "Finite-Size Error in Many-Body Simulations with Long-Range Interactions", Phys. Rev. Lett. **97**, 076404 (2006).

[126]   Supplementary material of Ref. [117].

[127] J. J. Shepherd, G. H. Booth, and A. Alavi, "Investigation of the full configuration interaction quantum Monte Carlo method using homogeneous electron gas models", J. Chem. Phys. **136**, 244101 (2012).

[128] T. Dornheim, S. Groth, T. Sjostrom, F. D. Malone, W. M. C. Foulkes, and M. Bonitz, "*Ab initio* quantum Monte Carlo simulation of the warm dense electron gas in the thermodynamic limit", Phys. Rev. Lett. **117**, 156403 (2016).

[129] P. K. Shukla, and B. Eliasson, "Novel Attractive Force between Ions in Quantum Plasmas", Phys. Rev. Lett. **108**, 165007 (2012), errata: **108**, 219902(E) (2012); **109**, 019901(E) (2012).

[130] M. Bonitz, E. Pehlke, and T. Schoof, "Reply to "Comment on 'Attractive forces between ions in quantum plasmas: Failure of linearized quantum hydrodynamics'"", Phys. Rev. E **87**, 037102 (2013).

[131] P. K. Shukla, B. Eliasson, and M. Akbari-Moghanjoughi, "Comment on "Attractive forces between ions in quantum plasmas: Failure of linearized quantum hydrodynamics"", Phys. Rev. E **87**, 037101 (2013).

[132] M. Bonitz, E. Pehlke, and T. Schoof, "Attractive forces between ions in quantum plasmas: Failure of linearized quantum hydrodynamics", Phys. Rev. E **87**, 033105 (2013).

[133] P. K. Shukla, B. Eliasson, and M. Akbari-Moghanjoughi, "Discussion on 'Novel attractive force between ions in quantum plasmas—failure of simulations based on a density functional approach'", Phys. Scr. **87**, 018202 (2013).

[134] M. Bonitz, E. Pehlke, and T. Schoof, "Comment on 'Discussion on novel attractive force between ions in quantum plasmas—failure of simulations based on a density functional approach'", Phys. Scr. **88**, 057001 (2013).

[135] M. Akbari-Moghanjoughi, "Hydrodynamic limit of Wigner-Poisson kinetic theory: Revisited", Physics of Plasmas (1994-present) **22**, 022103 (2015), erratum: **22**, 039904 (2015).

[136] D. Michta, F. Graziani, and M. Bonitz, "Quantum Hydrodynamics for Plasmas – a Thomas-Fermi Theory Perspective", Contrib. Plasma Phys. **55**, 437 (2015).

[137] Z. Moldabekov, T. Schoof, P. Ludwig, M. Bonitz, and T. Ramazanov, "Statically screened ion potential and Bohm potential in a quantum plasma", Physics of Plasmas **22**, 102104 (2015).

[138] D. Hochstuhl, and M. Bonitz, "Time-dependent restricted-active-space configuration-interaction method for the photoionization of many-electron atoms", Phys. Rev. A **86**, 053424 (2012).

[139] D. Hochstuhl, C. M. Hinz, and M. Bonitz, "Time-dependent multiconfiguration methods for the numerical simulation of photoionization processes of many-electron atoms", Eur. Phys. J. Spec. Top. **223**, 177 (2014).

[140] N. D. Drummond, P. López Ríos, R. J. Needs, and C. J. Pickard, "Quantum Monte Carlo Study of a Positron in an Electron Gas", Phys. Rev. Lett. **107**, 207402 (2011).

Danksagung

Selbständigkeitserklärung

Ich erkläre, dass die vorliegende Dissertation nach Inhalt und Form meine eigene Arbeit darstellt und unter Einhaltung der Regeln guter wissenschaftlicher Praxis der Deutschen Forschungsgemeinschaft entstanden ist. Ausgewählte Ergebnisse wurden in den auf Seite 3 angegebenen Fachartikeln publiziert, was an entsprechender Stelle in dieser Arbeit gekennzeichnet wurde. Teile der von mir im Rahmen dieser Arbeit erlangten Ergebnisse wurden bereits vorab in der Masterarbeit "Strongly Degenerate Nonideal Fermi Systems: Configuration Path Integral Monte Carlo Simulations" (Simon Groth, Universität Kiel, 2014) dargestellt und dort entsprechend gekennzeichnet. Weiterhin bestätige ich, dass diese Arbeit weder ganz noch in Teilen in einem anderen Prüfungsverfahren zur Begutachtung vorgelegen hat.

Kiel, den 22. Dezember 2016,

_____

Tim Schoof