

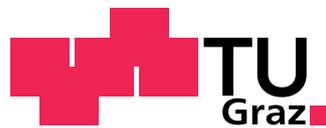
Werner Dobrautz, BSc

Application of the Full Configuration Interaction Quantum Monte Carlo algorithm to the two-dimensional fermionic Hubbard model

Master Thesis

For obtaining the academic degree
Diplom-Ingenieur

Master Programme of
Technical Physics



Graz University of Technology

Supervisor:

Univ.-Prof. Dipl-Phys. Dr.rer.nat. Wolfgang von der Linden

Institut für Theoretische Physik - Computational Physics

Graz, March 2014

Statutory Declaration

I declare that I have authored this thesis independently, that I have not used other than the declared sources/resources, and that I have explicitly marked all material which has been quoted either literally or by content from the used sources.

Graz, _____

Date

Signature

Eidesstattliche Erklärung¹

Ich erkläre an Eides statt, dass ich die vorliegende Arbeit selbstständig verfasst, andere als die angegebenen Quellen/Hilfsmittel nicht benutzt, und die den benutzten Quellen wörtlich und inhaltlich entnommenen Stellen als solche kenntlich gemacht habe.

Graz, am _____

Datum

Unterschrift

¹ Beschluss der Curricula-Kommission für Bachelor-, Master- und Diplomstudien vom 10.11.2008; Genehmigung des Senates am 1.12.2008

Abstract

The Hubbard model, one of the simplest models in many-body physics, describes highly correlated electrons on a lattice. It is believed to be able to give a description of the copper oxide planes responsible for the high- T_C superconductivity of the cuprates. But despite its seeming simplicity the exponential growth of the Hilbert space restricts standard numerical methods to very small system sizes.

The Full Configuration Interaction Quantum Monte Carlo (FCIQMC) algorithm, a recently developed adaptation of the Diffusion Monte Carlo method by the group of Prof. Ali Alavi allows an exact calculation of the ground state of a quantum mechanical system. It avoids the necessity of a prior knowledge of the nodal structure of the fermionic ground state by formulation in the anti-symmetric Hilbert space of Slater determinants. Record breaking results, regarding system sizes and accuracy, in quantum chemistry calculations were achieved by the use of this method.

In this thesis the application of the FCIQMC method to the two-dimensional fermionic Hubbard model is investigated. The differences of formulation in restricted and unrestricted Hartree-Fock single particle basis and the influences of different boundary conditions (periodic and twisted) are analyzed. To compare results, small systems were solved with exact diagonalization with extensive use of the symmetries of the model.

Kurzfassung

Das Hubbard Modell, eines der simpelsten Modelle in der Viel-Teilchen Physik, beschreibt stark korrelierte Elektronen auf einem Gitter. Man glaubt, dass es in der Lage ist eine Beschreibung der Kupfer-Oxid Ebenen, die für die Hochtemperatursupraleitung der Kuprate verantwortlich sind, zu liefern. Aber trotz der anscheinenden Einfachheit schränkt der exponentielle Wachstum des Hilbert Raums übliche numerische Methoden auf sehr kleine Systemgrößen ein.

Der 'Full Configuration Interaction Quantum Monte Carlo' (FCIQMC) Algorithmus, eine kürzlich von der Gruppe von Prof. Ali Alavi entwickelte Weiterentwicklung der 'Diffusion Monte Carlo' Methode, erlaubt eine exakte Berechnung des Grundzustandes eines quantenmechanischen Systems. Er umgeht die Notwendigkeit einer vorherigen Kenntnis der Vorzeichenstruktur des fermionischen Grundzustandes durch eine Formulierung im anti-symmetrischen Hilbert Raum von Slater Determinanten. Rekordbrechende Resultate, Systemgrößen und Genauigkeit betreffend, in quantenchemischen Berechnung wurde mit dieser Methode erzielt.

In dieser Arbeit wird die Anwendung der FCIQMC Methode auf das zweidimensionale fermionische Hubbard Modell untersucht. Die Unterschiede der Formulierung in der 'restricted' und 'unrestricted' Hartree-Fock Einteilchen Basis und die Einflüsse von verschiedenen Randbedingungen (periodisch and verdreht) wird analysiert. Für den Vergleich von Ergebnissen werden kleine System mit exakter Diagonalisierung, mit umfangreicher Benutzung der Symmetrien des Modells, gelöst.

Acknowledgments

The work on this thesis was supported by Prof. Ali Alavi and members of his group at the Department of Chemistry of the University of Cambridge, a cooperation initiated by Prof. Hans Gerd Evertz.

The FCIQMC simulations in this work were performed with the stand-alone FCIQMC code NECI [1] available at:

https://github.com/ghb24/NECI_STABLE

FCIQMC code developed by George Booth and Ali Alavi, 2013

Copyright (c) 2013, Ali Alavi.

The code was provided, explained and modified for this work by Prof. Ali Alavi and Dr. George H. Booth.

Acknowledgments

First and foremost I would like to thank my supervisor Prof. Wolfgang von der Linden for giving me the possibility to work on this topic and his assistance and guidance during my work on it. And I am especially grateful that he made it possible for me to get in contact with Prof. Ali Alavi and allowing me to visit his group in the beautiful city of Cambridge. I am very thankful for the warm welcome and the hospitality of Prof. Alavi and his group at the department of chemistry during my two visits there. The expertise and assistance of Prof. Alavi, George Booth and Michael Möckel, which went beyond my stay at their group, were of great help for the advancement of my work, which would not have been possible without it.

Additionally I would like to thank my family for their unconditional support and all the great people I met and enjoyed spending time with during my studies in Graz. Especially I want to thank Veronika Obersteiner for sharing our journey during our final phase of university education, giving me motivation when needed, understanding the struggles related to the work on a master thesis and being a great friend in general.

Contents

Abstract	iv
1 Introduction	1
1.1 The Hubbard model	1
1.2 High- T_c cuprate superconductors	1
1.3 Transition to the theoretical model	3
1.4 Full configuration interaction Quantum Monte Carlo	7
1.5 Outline	8
2 Exact Diagonalization	9
2.1 Analytic solution of limiting cases	9
2.2 Basis construction	12
2.3 Calculation of the Hamilton matrix	16
2.3.1 No translational and point-group symmetry	16
2.3.2 Fermionic phase factors	17
2.3.3 Calculation of off-diagonal matrix elements	18
2.4 Application of translational and point-group symmetries	22
2.4.1 Hamilton matrix	28
2.5 Results	31
3 Mean-field treatment	36
3.1 Hartree-Fock approximation	36
3.1.1 Restricted Hartree-Fock	37
3.1.2 Unrestricted Hartree-Fock	38
3.1.3 Configuration interaction	39
3.2 Mean-field approximation to the Hubbard model	41
3.2.1 Restricted Hartree-Fock	42
3.2.2 Self-consistent unrestricted Hartree-Fock	43
3.2.3 Analytic unrestricted Hartree-Fock	47
3.2.4 Tilted square lattice	51

3.3	Basis transformation	53
3.3.1	Transforming the Hamiltonian	53
3.3.2	Transformation of the ground state	55
3.4	Comparison of different bases	57
4	Quantum Monte Carlo	65
4.1	Variational QMC	67
4.2	Diffusion QMC	67
4.2.1	The fixed-node approximation in DMC	70
4.3	The sign problem	70
5	Full configuration interaction QMC	73
5.1	The algorithm	75
5.2	Population dynamics	81
5.3	The initiator approximation	82
5.4	Energy estimators	83
5.5	Error analysis	84
5.6	FCIQMC run properties for the 4×4 -square and 18-site tilted lattice	86
5.7	Outlook on larger lattices	105
6	Outlook: Drude and superfluid weight	108
	Bibliography	113

1 Introduction

1.1 The Hubbard model

Introduced in 1963 as a model for the description of the strong electronic correlation effects in the narrow d - and f -bands of the transition metals [2], the Hubbard model until today is in the focus of attention of modern theoretical research.

With the discovery high- T_c cuprate superconductors, [3], this interest has been revitalized, as it is believed that the Hubbard model is the simplest model able to describe the physics of the two-dimensional copper-oxide planes [4, 5, 6], responsible for the unconventional superconductivity of the cuprates.

1.2 High- T_c cuprate superconductors

The first discovered high- T_c superconductor was the ceramic barium-doped lanthanum-copper oxide $\text{La}_{2-x}\text{Ba}_x\text{CuO}_4$ in 1986 [3], with the record-breaking critical temperature of 30 K. Shortly afterwards the first superconductor with a critical temperature above the technical important boiling point of liquid nitrogen of 77 K, was found in yttrium-barium-copper oxide $\text{YBa}_2\text{Cu}_3\text{O}_7$ (YBCO) [7].

As an example for most cuprates the crystal structure of YBCO, a layered perovskite-like structure, is shown in figure (1.1a) [8].

The CuO_2 planes, responsible for the superconducting properties of YBCO, are separated by insulating layers of yttrium or barium. These layers, besides conventional doping and oxygen defects, determine the hole-doping of the CuO_2 layers, which is crucial for the electronic properties of the materials.

Band-structure calculations for $\text{La}_{2-x}\text{Ba}_x\text{CuO}_4$, as discussed below, show that the $3d$ -copper- and $2p$ -oxygen-orbitals in the two-dimensional CuO_2 planes are the dominant factors for its electronic properties [10], [11].

In fig.(1.2) a typical phase diagram for superconductors of cuprate type is displayed, taken from [9]. For low doping values around half-filling the low temperature phase has antiferromagnetic order, due to Mott-localization of the $3d$ copper electrons. By adding holes to the system the antiferromagnetic long-range order of the localized spins is destroyed.

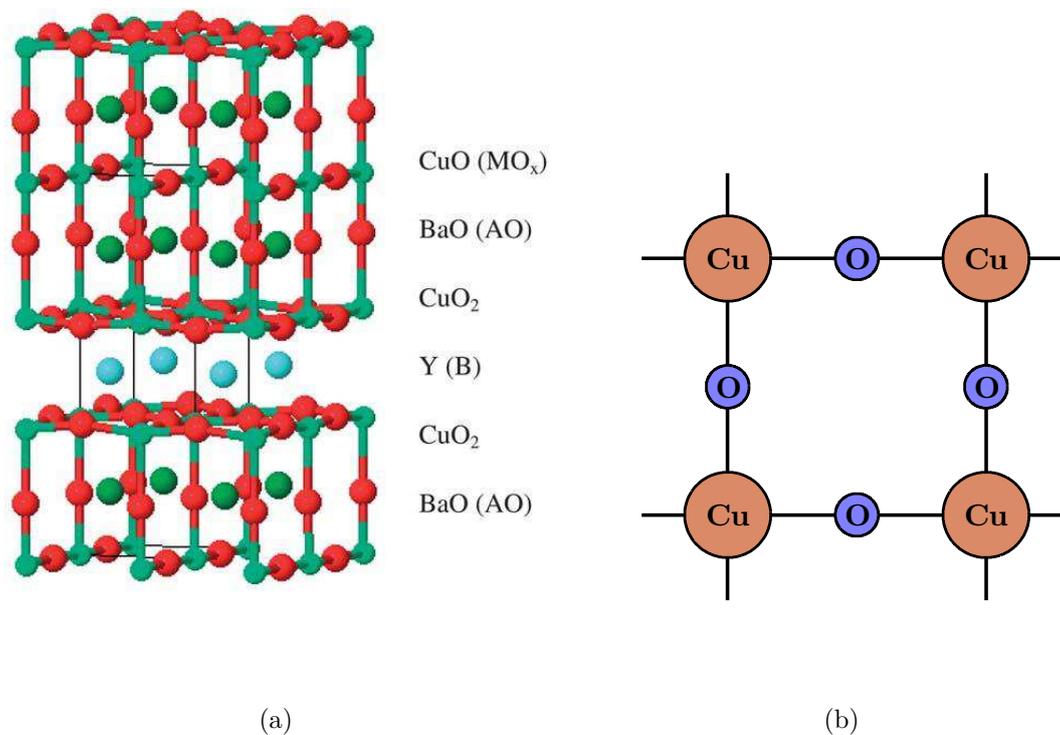


Figure 1.1: General chemical structure of high- T_c cuprates using the example of YBCO. (a) the crystal structure of $\text{YBa}_2\text{Cu}_3\text{O}_7$ taken from [9]. (b) A schematic two-dimensional CuO_2 plane most important for the electronic properties of the cuprates.

With increasing hole doping the system changes to a pseudogap state below the critical superconducting temperature. This pseudogap state is characterized by the existence of very few states near the Fermi level [12], [13]. Upon a specific value of hole-doping the superconducting state emerges, although the possible reasons for it are still subject of debate. Experimental evidence strongly suggests that, also for the cuprates, the formation of Cooper-pairs in the copper-oxide layers is the reason for superconductivity. But the pairing mechanism is not as in BCS-theory, the process of saving ionic kinetic energy through formation of Cooper-pairs, but most likely a consequence of strong electron-electron interaction [14].

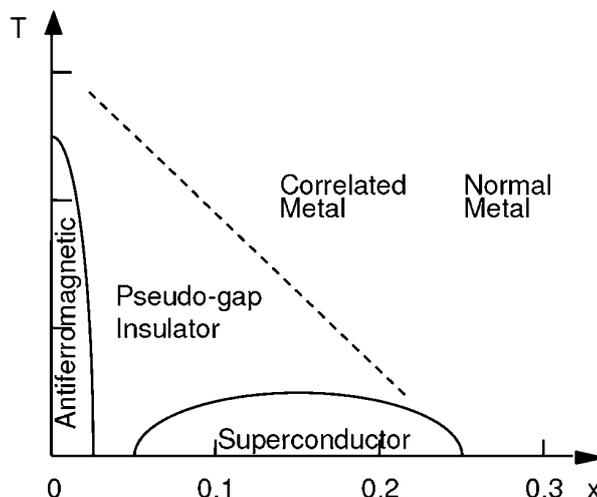


Figure 1.2: General phase diagram example of cuprate-superconductors with T being the temperature and x the hole doping in the CuO_2 plane. Taken from [9].

1.3 Transition to the theoretical model

Taking barium-doped lanthanum-copper oxide $\text{La}_{2-x}\text{Ba}_x\text{CuO}_4$ as an example it can be concluded from its stoichiometry that the copper ion is in the valency Cu^{2+} , implying a one hole- $3d^9$ electronic configuration. While the oxygen has a filled $2p$ -orbital configuration with a valency O^{2-} . As seen in the LAPW band-structure of $\text{La}_{2-x}\text{Ba}_x\text{CuO}_4$ in fig. (1.3) [10], there are only two of the $\text{Cu}(3d)$ - $\text{O}(2p)$ bands of importance. These are the filled bonding(B) and half-filled anti-bonding(A) $\text{Cu}(3d_{x^2-y^2})$ - $\text{O}(2p_{x/y})$ σ -bonds in the xy -plane, [11]. The dominant electronic behavior of the CuO_2 planes can be seen as there is almost no dispersion for any orbital along the Λ -path between Γ and Z in fig. (1.3), which is out of the CuO_2 in the z -direction of the Brillouin zone.

A schematic representation of the bonding and anti-bonding configurations of the $x^2 - y^2$ -symmetric $3d$ -copper orbitals and the corresponding non-orthogonal p_x - or p_y -oxygen orbitals are displayed in fig. (1.4)

Although there are these three orbitals ($\text{Cu}(3d_{x^2-y^2})$, $\text{O}(2p_{x/y})$) in the CuO_2 unit-cell, essentially requiring a three-band model for description, angle resolved photoemission spectroscopy data suggest that only one band is present at the Fermi surface of superconducting cuprates, [15]. This means the electronic properties of these CuO_2 planes of the cuprates can be described by an effective one band model on a two-dimensional lattice.

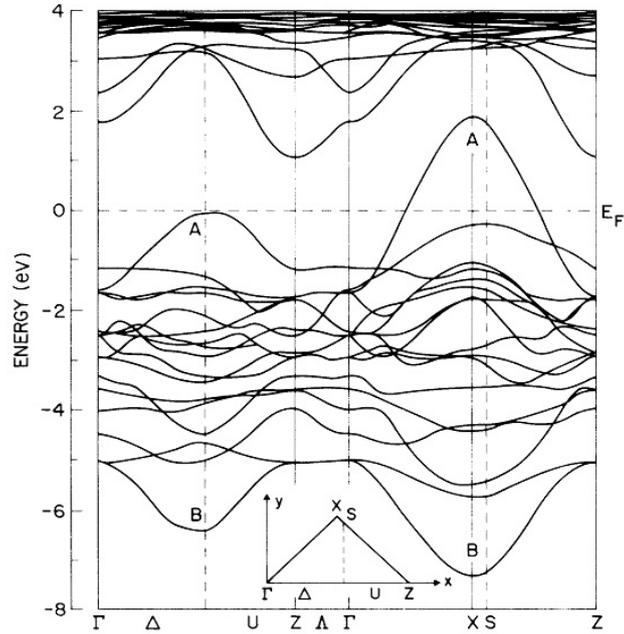


Figure 1.3: LAPW band structure of $\text{La}_{2-x}\text{Ba}_x\text{CuO}_4$. Except the filled bonding(B) and half-filled anti-bonding(A) $\text{Cu}(3d_{x^2-y^2})\text{-O}(2p_{x/y})$ in-plane σ -bonds of the CuO_2 plane, no other orbitals display a significant dispersion. This figure is taken from [10].

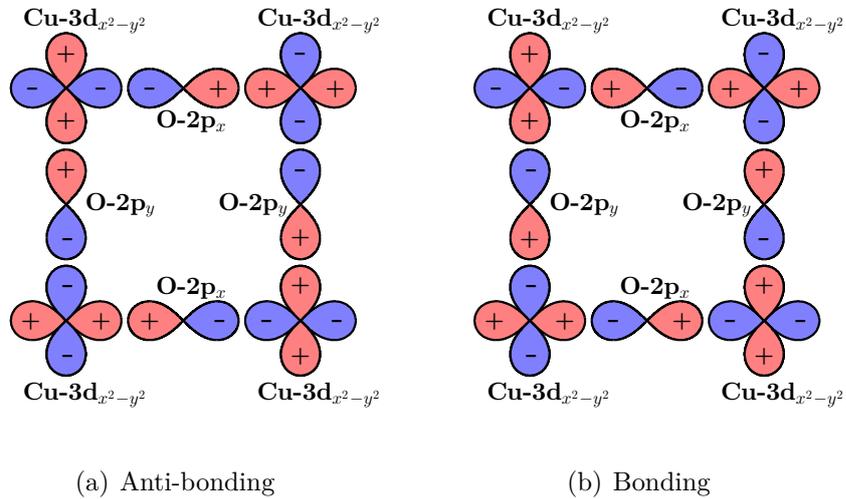


Figure 1.4: Arrangement of the electronically active orbitals in the CuO_2 plane. (a) anti-bonding configuration. (b) bonding configuration

The model

Since the Hubbard model is the simplest electronic lattice model which may possess a superconducting phase in two dimensions and is able to describe the one- and two-particle dynamics of electrons in the orbitals of the CuO_2 -planes it is in the spotlight as possible theoretical description of the superconductivity of the cuprates.

In real space and second quantization notation it is defined as

$$H = -t \sum_{\langle i,j \rangle, \sigma} c_{i\sigma}^\dagger c_{j\sigma} + U \sum_i n_{i,\uparrow} n_{i,\downarrow}. \quad (1.1)$$

$c_{i,\sigma}^\dagger$ ($c_{i,\sigma}$) are the fermionic creation (annihilation) operators, which create (destroy) a fermion with spin σ on site i and n_i^\uparrow (n_i^\downarrow) are the corresponding number-operators. Due to the anti-symmetric properties of fermions, the fermionic creation and annihilation operators $c_{i,\sigma}^\dagger$ and $c_{i,\sigma}$ obey the canonical anticommutation relations (CAR):

$$\{c_{i,\sigma}^\dagger, c_{j,\sigma'}^\dagger\} = [c_{i,\sigma}^\dagger, c_{j,\sigma'}^\dagger]_+ = [c_{i,\sigma}, c_{j,\sigma'}]_+ = 0 = \delta_{i,j} \delta_{\sigma,\sigma'} \quad (1.2)$$

and

$$\{c_{i,\sigma}^\dagger, c_{j,\sigma'}\} = [c_{i,\sigma}^\dagger, c_{j,\sigma'}]_+ = \delta_{i,j} \delta_{\sigma,\sigma'}, \quad (1.3)$$

where $\{A,B\} = [A,B]_+ = AB + BA$. In this work the summation in the first term of (1.1) is restricted to nearest-neighbors sites only and corresponds to the one-particle kinetic energy ('hopping') part of the model, identical to the tight-binding Hamiltonian. It describes the movement of electron between neighboring lattice sites. The second part in (1.3) accounts for the very short-ranged on-site two-particle interaction. In the repulsive $U > 0$ fermionic Hubbard model this corresponds to the Coulomb repulsion of two electrons with opposite spin residing on the same lattice site i . In fig. (1.5) the two possible effects occurring within the Hubbard model are displayed. A hopping of an electron with spin- σ to an nearest-neighbor site, previously unoccupied by an spin- σ electron, with kinetic energy $-t$. And the repulsive on-site repulsion of two electrons with opposite spin on the same site in units of U/t .

Since the model (1.3) possesses only two parameters, the hopping strength t and on-site repulsion U , it is possible to fix one value and measure every quantity of interest in units of it. In this work the hopping parameter was set to $t = 1$ and only parameter U was allowed to change.

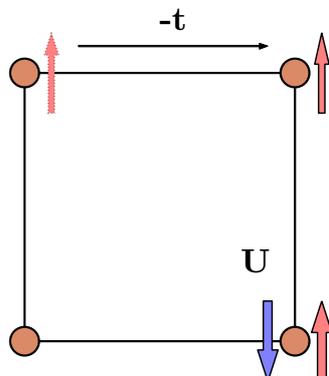


Figure 1.5: Example of the possible processes of the effective two-dimensional one-band Hubbard model on a 2×2 -lattice: the possible on-site repulsion of strength U and nearest-neighbor hopping with magnitude $-t$.

Optical lattices

A further spark of interest in lattice models, and especially the Hubbard model, was caused by the advent of precisely controllable ultra-cold atom gases on optical lattices used as quantum simulators [16, 17, 18, 19, 20, 21]. Where the ultra-cold atoms are trapped on certain lattice sites and the strength of the corresponding hopping and repulsion parameters are experimentally tunable, and thus, in principle, allowing the physical simulation of the theoretical model.

The Hubbard model was exactly solved in one-dimension through the use of the Bethe-Ansatz by Lieb and Wu [22]. And in two dimension the two limiting cases $t = 0$, which is already diagonal in the real-space representation, and $U = 0$, which is identical to the tight-binding model and diagonalized through a momentum-space transformation (2.1), are also easily solved. And yet, due to the dynamic interplay of the kinetic hopping term and the Coulomb-like on-site interaction part, the Hubbard model, nearly 50 years after its introduction, is still not solved and not even thoroughly understood for a wide range of the parameter U in dimensions greater than one. And due to the exponentially growing Hilbert-space with lattice sites and number of electrons an exact solution in the classical sense is out of reach for even intermediate system sizes.

The Hubbard model is not only of interest due to its physical applications in the above mentioned cases, but also from a pure theoretical standpoint, since it is the simplest possible model for strongly interacting fermions, only taking into account the motion of electrons through lattice sites and short range interaction of opposite

spin fermions. Nevertheless it exhibits many interesting complex physical phenomena, like metal-insulator transitions, antiferromagnetism and superconductivity, naming only a few examples. Also other prominent theoretical models like the $t - J$ and its special case at half filling, the Heisenberg model, can be derived in the strong coupling $U \gg t$ limit of the Hubbard model [23]. Any advancement in the solutions for bigger lattice sizes, in dimensions greater than one, would be a great starting point in the understanding of many interesting physical phenomena.

1.4 Full configuration interaction Quantum Monte Carlo

In this work the recently developed full configuration interaction quantum Monte Carlo algorithm (FCIQMC) by G.H. Booth, A. Thom and A. Alavi [24] and its shortly afterwards adapted initiator-approximation version [25] were applied to the two dimensional Hubbard model. The FCIQMC algorithm set new records, concerning accuracy and system size, in calculation of chemical systems [26, 27], in the field of quantum chemistry and was recently also applied to the homogeneous electron gas [28] and real solids [29]. The in chapter 5 more thoroughly described FCIQMC method is an adaptation of the diffusion quantum Monte Carlo method which projects an arbitrary starting state $|\Psi\rangle$ onto the ground state of a system by imaginary-time evolution

$$|\Psi_0\rangle = \lim_{\tau \rightarrow \infty} e^{-\tau(H-E_0)} |\Psi\rangle. \quad (1.4)$$

In the FCIQMC method the application of the projector $\exp(-\tau(H - E_0))$ is simulated by population dynamic rules for a set of, so called, walkers living in the Hilbert-space in which the wave function is expanded in

$$|\Psi\rangle = \sum_i c_i |\Phi_i\rangle. \quad (1.5)$$

The sign problem of the DMC algorithm, due to the unknown nodal structure of the true ground state wave function, is circumvented in the FCIQMC algorithm by expanding the wave function (1.5) in Slater determinants $|\Phi_i\rangle$. The weight c_i of each basis state is given by the number of signed walkers inhabiting it and which evolve according to rules set by the underlying Hamiltonian of the system. Annihilation processes of walkers with different sign inhabiting the same Slater determinant $|\Phi_i\rangle$ allow the true ground state wave function to emerge during simulation.

The choice of the single orbital basis in which the wave function is represented is of crucial importance to the performance of the FCIQMC algorithm. In the above

mentioned successful applications to chemical system the basis states in which the wave function were expanded were highly sophisticated, but very specific to the investigated system, basis sets prevalent in the quantum chemistry community. For system defined on periodic lattice systems, as the Hubbard model in this work, the choice of basis is more generic. In an optimal basis as few basis states $|\Phi_i\rangle$ as possible would contain most of the important information of the system.

In this work the performance of the FCIQMC method applied to the Hubbard Hamiltonian formulated in the **momentum-space** and **unrestricted Hartree-Fock** basis was investigated.

For FCIQMC calculations in this work the openly available stand-alone FCIQMC code NECI [1] was utilized, which was provided and explained to me by developers Prof. Ali Alavi and George H. Booth during two visits at their workgroup and in ongoing correspondence.

1.5 Outline

The remaining structure of this thesis is as follows: In chapter 2 the exact diagonalization of the Hubbard model for small systems is described. The exact solution is on the one hand used to compare results obtained from FCIQMC simulations and on the other to assert the validity and investigate properties of the momentum-space and Hartree-Fock basis transformations, which are described in the following chapter 3. Also in chapter 3 the underlying method of full configuration interaction and basic terminology used in quantum chemistry, and thus in the FCIQMC formulation, is described.

Chapter 4 gives a short review to the general quantum Monte Carlo technique and serves as a transition to the introduction of the main topic of this thesis, the full configuration quantum Monte Carlo method, described in chapter 5. Chapter 5 also contains results on which factors the performance of the FCIQMC method depends when applied to the Hubbard model on a periodic lattice.

The concluding chapter 6 contains a discussion and outlook on how to conclude superconducting properties from the Drude and superfluid weight of a system, obtained from ground state energies as a function of an applied external field.

2 Exact Diagonalization

2.1 Analytic solution of limiting cases

For $t = 0$ the Hubbard Hamiltonian (1.1) is already diagonal in the real-space basis. In the non-interacting $U = 0$ limit the Hubbard Hamiltonian is identical to the tight-binding Hamiltonian and can be diagonalized by a momentum space transformation

$$c_{\mathbf{k},\sigma}^\dagger = \frac{1}{\sqrt{N}} \sum_{\mathbf{r}} e^{i\mathbf{k}\mathbf{r}} c_{\mathbf{r},\sigma}^\dagger, \quad c_{\mathbf{k},\sigma} = \frac{1}{\sqrt{N}} \sum_{\mathbf{r}} e^{-i\mathbf{k}\mathbf{r}} c_{\mathbf{r},\sigma}, \quad (2.1)$$

which is also one of the bases in which the FCIQMC algorithm is applied to the Hubbard model. Inserting the inverse transformation of (2.1) into the real-space Hubbard Hamiltonian (1.1) yields

$$H = \frac{-t}{N} \sum_{\mathbf{k},\mathbf{k}',\sigma} \sum_{\langle \mathbf{r},\mathbf{r}' \rangle} e^{-i(\mathbf{k}\mathbf{r}-\mathbf{k}'\mathbf{r}')} c_{\mathbf{k},\sigma}^\dagger c_{\mathbf{k}',\sigma} \\ + \frac{U}{N^2} \sum_{\mathbf{k}\mathbf{l}\mathbf{m}\mathbf{n},\mathbf{r}} e^{-i\mathbf{r}(\mathbf{k}-\mathbf{l}+\mathbf{m}-\mathbf{n})} c_{\mathbf{k},\uparrow}^\dagger c_{\mathbf{l},\uparrow} c_{\mathbf{m},\downarrow}^\dagger c_{\mathbf{n},\downarrow}. \quad (2.2)$$

With the nearest neighbor conditions

$$r_x = r'_x + 1 \vee r'_x - 1 \quad \text{for } r_y = r'_y \quad (2.3)$$

$$r_y = r'_y + 1 \vee r'_y - 1 \quad \text{for } r_x = r'_x \quad (2.4)$$

and

$$\frac{1}{N} \sum_{\mathbf{r}} e^{i\mathbf{r}(\mathbf{k}-\mathbf{k}')} = \delta_{\mathbf{k},\mathbf{k}'} \quad (2.5)$$

equation (2.2) becomes

$$\begin{aligned}
H &= -t \sum_{\mathbf{k}, \mathbf{k}', \sigma} \left(e^{ik_x} + e^{-ik_x} + e^{ik_y} + e^{-ik_y} \right) c_{\mathbf{k}, \sigma}^\dagger c_{\mathbf{k}', \sigma} \delta_{\mathbf{k}, \mathbf{k}'} \\
&+ \frac{U}{N} \sum_{\mathbf{k}, \mathbf{l}, \mathbf{m}, \mathbf{n}} c_{\mathbf{k}, \uparrow}^\dagger c_{\mathbf{l}, \downarrow} c_{\mathbf{m}, \uparrow}^\dagger c_{\mathbf{n}, \downarrow} \delta_{\mathbf{k}, \mathbf{l} + \mathbf{n} - \mathbf{m}} \\
&= \sum_{\mathbf{k}, \sigma} \varepsilon(\mathbf{k}) n_{\mathbf{k}, \sigma} + \frac{U}{N} \sum_{\mathbf{l}, \mathbf{m}, \mathbf{n}} c_{\mathbf{l} + \mathbf{n} - \mathbf{m}, \uparrow}^\dagger c_{\mathbf{l}, \uparrow} c_{\mathbf{n}, \downarrow}^\dagger c_{\mathbf{m}, \downarrow},
\end{aligned} \tag{2.6}$$

which is the Hubbard model formulated in momentum space representation. For $U = 0$ it is diagonal with the tight binding single particle energy $\varepsilon(\mathbf{k}) = -2t(\cos(k_x) + \cos(k_y))$ in two dimensions. But compared to the real-space representation (1.1) the former diagonal two-particle interaction part now takes on a more complicated form.

Exact diagonalization

For every other values of U and t , due to the interplay of the one particle hopping part and the two-particle interaction term, no such trivial solution is possible. To obtain an exact solution of a quantum-mechanical problem, like the Hubbard model, the associated Hamiltonian, formulated in a suitable basis, has to be diagonalized. This exact diagonalization (ED) then yields the complete spectrum of eigenvalues and eigenvectors and hence complete knowledge of the investigated system. But the complete diagonalization of a matrix is a computationally very expensive procedure, generally scaling like $\mathcal{O}(N^3)$, where N is the exponentially growing size of the Hilbert space of the problem. If one is only interested in the low-energy, zero-temperature physics of a model, where only the ground-state and a few excited-states are of major importance to the properties of the system, less costly methods like the Lanczos-algorithm [30] can be used.

For exact diagonalization one has to construct a numerical representation of the basis states of the Hilbert space of the system. In the fermionic Hubbard model (1.1), due to the Pauli exclusion principle, there are four states per lattice site possible. With arbitrary choice of the z -axis as spin-quantization axis the possible states of a site can be seen in fig. (2.1).

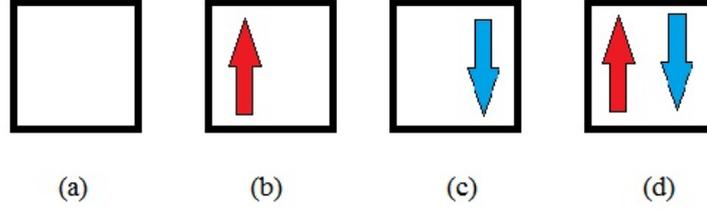


Figure 2.1: Possible states on a site i :

- (a) $|0\rangle \dots$ empty site i ,
- (b) $c_{i,\uparrow}^\dagger|0\rangle \dots$ one spin-up electron occupying site i ,
- (c) $c_{i,\downarrow}^\dagger|0\rangle \dots$ one spin-down electron occupying site i ,
- (d) $c_{i,\uparrow}^\dagger c_{i,\downarrow}^\dagger \dots$ two electrons with opposite spin on site i .

This leads to a 4^L -exponential growth of the Hilbert space with the number of lattice sites L . Without use of symmetries this limits the exact diagonalization application to $L \approx 16$, even on modern supercomputers [31]. Since the Hamiltonian will not mix eigenstates belonging to different symmetry subspaces, represented by operators A commuting with H , a separate solution of the eigenvalue problem within each symmetry subspace is possible. The most important symmetries for basis state construction of the Hubbard Hamiltonian (1.1) are the following:

- **Particle-number/charge conservation:**

The Hubbard Hamiltonian does not change the total number of particles of a basis state

$$N = \sum_{i,\sigma} n_{i,\sigma}. \quad (2.7)$$

- **Magnetization conservation:**

The total magnetization projected on the chosen quantization axis is conserved.

$$S_{tot}^z = \frac{1}{2} \sum_i (n_{i,\uparrow} - n_{i,\downarrow}) \quad (2.8)$$

- **Translational invariance:**

For periodic boundary conditions the Hubbard Hamiltonian commutes with

the translation operator

$$Tc_{i,\sigma}^{(\dagger)} = c_{i+1,\sigma}^{(\dagger)} \quad (2.9)$$

- **Point-group symmetry:**

Depending on the momentum of a state, certain underlying symmetries of the lattice point-group are conserved.

- **Full $SU(2)$ spin symmetry:**

In the Hubbard model all components of the total spin

$$S^\alpha = \frac{1}{2} \sum_{i,\mu,\nu} c_{i,\mu}^\dagger \sigma_{\mu,\nu}^\alpha c_{i,\nu} \quad (2.10)$$

with σ^α being the Pauli-matrices, are conserved. As opposed to the aforementioned symmetries the $SU(2)$ is computationally more complicated to implement and not used in practical computations most of the times.

In this work the use of the first four symmetries was implemented, allowing for an exact diagonalization of the Hubbard Hamiltonian up to the 16-site 4×4 -square lattice at half filling. Besides the dramatic reduction Hilbert space size, the use of symmetries in exact diagonalization calculations also allows an association of the solution with the conserved quantum numbers of the model.

2.2 Basis construction

The real-space basis is the most obvious basis to represent the Hubbard Hamiltonian in, and given by all possible combinations of distributing a given number of \uparrow - and \downarrow -spin -electrons on a lattice with N sites. Each basis state is uniquely identified by the positions of \uparrow - and \downarrow -spin electrons

$$|\Psi\rangle_i = \prod_{j \in \{I_\uparrow\}} c_{j,\uparrow}^\dagger \prod_{j' \in \{I_\downarrow\}} c_{j',\downarrow}^\dagger |0\rangle, \quad (2.11)$$

where $I_{\uparrow/\downarrow}$ are index lists of the corresponding occupied sites. Since we are dealing with fermions an important detail is that one has to define the normal order of the creation operators $c_{i,\sigma}^\dagger$ for the uniqueness of basis states. In this work the operator were first sorted by spin (the spin-down operators were applied first) and then by decreasing linear lattice index. Sorting first by spin has the advantage in the Hubbard model that the nearest-neighbor hopping processes, since they only involve

electron of same spin, do not cause inter-spin sign factors. Since application of the down-spin hopping part to an arbitrary state (2.11)

$$c_{i,\downarrow}^\dagger c_{j,\downarrow} \prod_{l \in \{I_\uparrow\}} c_{l,\uparrow}^\dagger \prod_{l' \in \{I_\downarrow\}} c_{l',\downarrow}^\dagger |0\rangle, \quad (2.12)$$

yields the same amount of phase factors for the creation and annihilation operator through the fermionic anti-commutation relations

$$(-1)^{2n_l} \prod_{l \in \{I_\uparrow\}} c_{l,\uparrow}^\dagger c_{i,\downarrow}^\dagger c_{j,\downarrow} \prod_{l' \in \{I_\downarrow\}} c_{l',\downarrow}^\dagger |0\rangle. \quad (2.13)$$

The use of the particle number and total magnetization conservation of the Hubbard Hamiltonian is readily implemented by only creating the basis states for certain fixed number of spin-up N_\uparrow and spin-down N_\downarrow electrons on a lattice with N sites. The size of these symmetry subspaces is determined by the number of possibilities to distribute N_\uparrow spin-up electrons and N_\downarrow spin-down electrons on N sites and thus

$$\binom{N}{N_\uparrow} \binom{N}{N_\downarrow}. \quad (2.14)$$

The biggest subspace with fixed particle number and total spin magnetization is the half-filled sector with $N_\uparrow = N_\downarrow = \frac{N}{2}$. Which is approximately a factor $\frac{\pi L}{2}$ smaller than the un-symmetrized case of 4^L . An efficient numerical implementation is to represent the spin-up and spin-down of an basis state (2.11)

$$|\Psi\rangle_i = |\Psi_\uparrow\rangle_i \otimes |\Psi_\downarrow\rangle_i \quad (2.15)$$

by a binary representation with two bit strings, where a 1 indicates an occupied site by either a \uparrow - or \downarrow -spin and a 0 an unoccupied one. For example a linear 4-site lattice with two spin- \uparrow electrons on site 1 and 3 and two spin- \downarrow electrons on site 3 and 4:

$$\begin{aligned} |\Psi\rangle &= c_{1,\uparrow}^\dagger c_{3,\uparrow}^\dagger c_{3,\downarrow}^\dagger c_{4,\downarrow}^\dagger |0\rangle \hat{=} |\uparrow, 0, \uparrow\downarrow, \downarrow\rangle \\ &\hat{=} |\uparrow, 0, \uparrow, 0\rangle_\uparrow \otimes |0, 0, \downarrow, \downarrow\rangle_\downarrow \hat{=} |1010\rangle_\uparrow \otimes |0011\rangle_\downarrow. \end{aligned} \quad (2.16)$$

This representation is efficiently storable as the corresponding integer value, where the spin- \uparrow part is treated as the 'most significant bit' part. For example (2.16) this would yield

$$\text{Int}(|\Psi\rangle) = 2^4 \times (2^3 + 1^1) + 1^1 + 1^0 = 163 \quad (2.17)$$

Another advantage of this bit-string integer representation is a fast calculation of the action of the Hamiltonian on a basis state through bit-operations, further explained in section 2.3. To get the full Hubbard basis all possible states $|\Psi_{\uparrow}\rangle_i$ with N_{\uparrow} \uparrow -spin electrons on N sites and $|\Psi_{\downarrow}\rangle_j$ with N_{\downarrow} \downarrow -spin electron on N sites have to be created and all combinations of $|\Psi_{\uparrow}\rangle_i \otimes |\Psi_{\downarrow}\rangle_j$ have to be calculated. For efficient searching within the basis set, required in the Hubbard Hamiltonian calculation, it is important to create the basis in an ordered fashion. An efficient way to create an ordered list of integers with a fixed number of bits set to 1 is as follows:

1. Start with lowest possible binary integer representation for N_{σ} electrons on N sites. E.g.:

$$00111 \hat{=} 7$$

2. To get the next state find the least significant 0 with at least one 1 right of it and set it to 1. Set all bits to the right of this bit to 0 except the $(n - 1)$ lowest bits to 1, where n is the number of 1s right of the initial 0. Leave the rest of the bitstring unchanged.

$$00 \underbrace{111}_n \rightarrow 00 \mathbf{10} \underbrace{11}_{n-1} \hat{=} 11$$

3. continue until there is no 0 with 1s right of it.

$$\begin{aligned} 00111 &\hat{=} 7 \\ 01011 &\hat{=} 11 \\ 01101 &\hat{=} 13 \\ 01110 &\hat{=} 14 \\ 10011 &\hat{=} 19 \\ &\dots \\ 11100 &\hat{=} 28 \end{aligned}$$

It is only necessary to store the \uparrow - and \downarrow -spin basis states separately and implicitly remember that all combinations are needed for the full Hubbard basis and treat one spin channel (in this work the \uparrow -spin) as the most significant bit part. As an illustrative example the stored basis of the $N_{\uparrow} = 3$, $N_{\downarrow} = 2$, $N = 4$ site lattice system is given in table (2.1).

Table 2.1: \uparrow - and \downarrow -spin Hubbard basis parts for $N_\uparrow = 3, N_\downarrow = 2$ on a $L = 4$ site linear lattice

$ \uparrow\rangle$	$2^L \times \text{Int}(\uparrow\rangle)$	\otimes	$ \downarrow\rangle$	$\text{Int}(\downarrow\rangle)$
$ 0111\rangle$	$2^4 \times 7 = 112$		$ 0011\rangle$	3
$ 1011\rangle$	$2^4 \times 11 = 176$		$ 0101\rangle$	5
$ 1101\rangle$	$2^4 \times 13 = 208$		$ 0110\rangle$	6
$ 1110\rangle$	$2^4 \times 14 = 224$		$ 1001\rangle$	9
			$ 1010\rangle$	10
			$ 1100\rangle$	12

For the full basis each pair of basis states has to be combined, and since every possible \downarrow -spin integer is smaller than the smallest possible \uparrow -spin integer times 2^L , the basis stays in ordered form if the full basis is created by first combining the first \uparrow -spin state with each \downarrow -spin state, then the second \uparrow -spin state and so on as seen in table (2.2).

Table 2.2: Combined \uparrow - and \downarrow -spin Hubbard basis of table (2.1)

$ \uparrow\rangle$	$ \downarrow\rangle$	$\text{Int}(\uparrow\rangle \downarrow\rangle) = 2^L \text{Int}(\uparrow\rangle) + \text{Int}(\downarrow\rangle)$	Index in combined basis
$ 0111\rangle$	$ 0011\rangle$	$2^4 \times 7 + 3 = 115$	$0 \times 6 + 1 = 1$
$ 0111\rangle$	$ 0101\rangle$	$2^4 \times 7 + 5 = 117$	$0 \times 6 + 2 = 2$
$ 0111\rangle$	$ 0110\rangle$	$2^4 \times 7 + 6 = 118$	$0 \times 6 + 3 = 3$
$ 0111\rangle$	$ 1001\rangle$	$2^4 \times 7 + 9 = 121$	$0 \times 6 + 4 = 4$
$ 0111\rangle$	$ 1010\rangle$	$2^4 \times 7 + 10 = 122$	$0 \times 6 + 5 = 5$
$ 0111\rangle$	$ 1100\rangle$	$2^4 \times 7 + 12 = 124$	$0 \times 6 + 6 = 6$
$ 1011\rangle$	$ 0011\rangle$	$2^4 \times 11 + 3 = 179$	$1 \times 6 + 1 = 7$
$ 1011\rangle$	$ 0101\rangle$	$2^4 \times 11 + 5 = 181$	$1 \times 6 + 2 = 8$
$ 1011\rangle$	$ 0110\rangle$	$2^4 \times 11 + 6 = 182$	$1 \times 6 + 3 = 9$
...
$ 1110\rangle$	$ 1010\rangle$	$2^4 \times 14 + 10 = 234$	$3 \times 6 + 1 = 23$
$ 1110\rangle$	$ 1100\rangle$	$2^4 \times 14 + 12 = 236$	$3 \times 6 + 6 = 24$

2.3 Calculation of the Hamilton matrix

2.3.1 No translational and point-group symmetry

Without further use of symmetries, except the already implemented particle number and magnetization conservation, an efficient way of calculating the matrix representation of the Hubbard Hamiltonian (1.1) is to apply it on all basis states and efficiently relate the created states to the ordered list of basis states. The action of the Hamilton operator (1.1) can be split up in a diagonal and an off-diagonal part

$$H = \underbrace{-t \sum_{\langle i,j \rangle, \sigma} c_{i,\sigma}^\dagger c_{j,\sigma}}_{\text{off-diagonal}} + U \underbrace{\sum_i n_{i,\uparrow} n_{i,\downarrow}}_{\text{diagonal}}, \quad (2.18)$$

since the on site interaction term does not change basis states in real-space representations. It only counts the number of doubly occupied sites for each basis state. To implement this with the two-integer-table basis representation is to loop over both spin bases and sum the amount of 1s resulting from the bitwise AND-operation of \uparrow and \downarrow -spin bitstring basis multiplied by the on-site interaction parameter U . E.g.:

$$\begin{array}{rcl} |\uparrow\rangle: & 2^5 \times 7 & \hat{=} \quad 00111 \\ |\downarrow\rangle: & 11 & \hat{=} \quad 01011 \\ & \text{AND} & 00011 \end{array}$$

yields two doubly occupied sites.

The off-diagonal hopping part of eq. (2.18) can be further split up into an \uparrow - and \downarrow -spin part, since the hopping is independent of the opposed spin channel.

$$H_{hop} = H_{\uparrow}^h + H_{\downarrow}^h = -t \sum_{\langle i,j \rangle} \left(\underbrace{c_{i,\uparrow}^\dagger c_{j,\uparrow}}_{H_{\uparrow}} + \underbrace{c_{i,\downarrow}^\dagger c_{j,\downarrow}}_{H_{\downarrow}} \right). \quad (2.19)$$

Because as mentioned above, by using spin-first normal ordering of basis states, there are no inter-spin phase factors in the hopping processes. An electron on site i is in two-dimensions able to hop to its four neighboring sites in the one-band Hubbard model on a square lattice. The information of possible hopping destinations can be saved in a list with all linear indices of nearest-neighbor sites of each lattice site. This is essentially the only lattice information needed if there is even lattice spacing between each site, which is assumed throughout this work. To avoid double counting of hopping processes only two of the four possible hoppings per site must be counted.

The labeling conventions of the two-dimensional lattice throughout this work are of column-major order and the single site index I is defined as

$$I = I_x + I_y(L_x - 1), \quad (2.20)$$

where $I_x(I_y)$ are the $x(y)$ indices and L_x the linear size of the lattice in x -index direction, as seen at the example of a 4×4 -lattice in figure (2.2).

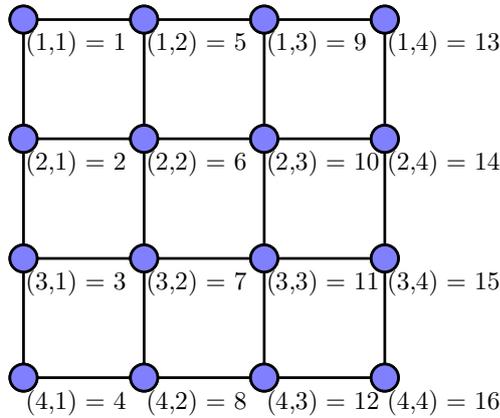


Figure 2.2: Lattice indication convention using the example of a 4×4 -lattice.

2.3.2 Fermionic phase factors

For a system with periodic (PBC) or twisted boundary conditions (TBC) fermionic phase factors have to be considered, within each spin channel, if a hopping over a boundary occurs, since the creation operators have to be brought into normal order. For a one-dimensional system these are easy to calculate, since fermionic phase factors only occur if an electron jumps over the boundary and the arrangement of operators have to be brought into normal order. But for a two-dimensional system, since the two lattice indices have to be mapped on a single index representation for a bit-string integer representation like (2.20), there are not only fermionic phases factors when there is hopping over a boundary, but also within the lattice if a hop corresponds to jump 'over' another electron in the linear lattice index. So if for example we have a state with \uparrow -spin electrons on lattice site 2 and 5, a hop from

the electron from site 5 to 1

$$c_{1,\uparrow}^\dagger \underbrace{c_{5,\uparrow} c_{2,\uparrow} c_{5,\uparrow}^\dagger}_{-c_{2,\uparrow}^\dagger c_{5,\uparrow}} |0\rangle = -c_{1,\uparrow}^\dagger c_{2,\uparrow}^\dagger \underbrace{c_{5,\uparrow} c_{5,\uparrow}^\dagger}_{|0\rangle} |0\rangle$$

yields a negative sign. The reversed process, a jump from an electron from site 1 to site 5 while an electron is at site 2 yields the same negative sign, due to normal ordering of the operators. In general the phase p_h of a hopping process is determined by the number of creation operators between the two sites ($i < j$) involved in the linear lattice index

$$p_h = (-1)^{\sum_{i < k < j} n_k}. \quad (2.21)$$

Since each \uparrow -spin basis state has the same \downarrow -spin basis part in the combined full Hubbard basis and \uparrow - and \downarrow -spin hopping does not interfere with each other, the application of each spin-hopping part of the Hamiltonian only has to be applied once to the respective spin-channel basis states to get the matrix representation of H_\uparrow and H_\downarrow from (2.19). The resulting indices have to be replicated accordingly to the whole basis.

Another major advantage of the bit-wise integer representation of the basis states is the very efficient calculation of the hopping process. An application of a bit-wise exclusive-OR (XOR) operation on a basis state and a bit-mask, with 1s on sites where hopping occurs, gives the desired hopped state.

001011	state
001100	bitmask
000111	XOR

The bit mask for site i is obtained through the previously mentioned nearest-neighbor index list kept for each site i .

2.3.3 Calculation of off-diagonal matrix elements

In this work following method was used to calculate the off-diagonal hopping matrices H_σ^h :

1. For a given ordered single-spin basis $|\Psi_\sigma\rangle$ each possible state $|\Psi'_\sigma\rangle = H_\sigma |\Psi_\sigma\rangle$ is calculated. This is achieved by first calculating, for each basis state $|\Psi_\sigma\rangle_i$, the position and number of possible hoppings by a XOR operation of state $|\Psi_\sigma\rangle_i$

and the nearest-neighbor state. E.g. in one dimension for the nearest-neighbor to the right of each site with periodic boundary conditions:

$$|\Psi_i\rangle_\sigma = |001011\rangle \quad (2.22)$$

$$\text{nearest-neighbors} : |010110\rangle \quad (2.23)$$

$$\text{XOR} : |011101\rangle \quad (2.24)$$

$$\text{bitmasks} : |011000\rangle, |001100\rangle, |000110\rangle, |100001\rangle \quad (2.25)$$

In two dimensions a loop over the dimensions is performed. Each 1 after the XOR operation in (2.25) and the site to the right of it indicate a possible hopping process.

2. The states $H_\sigma|\Psi_\sigma\rangle$ are calculated by a XOR operation with $|\Psi_\sigma\rangle$ and each of the bitmasks from (2.25). The resulting integers are stored in a list.
3. A ordered list I_x , the x-index of the H_σ , indicating the position of a state $|\Psi_\sigma\rangle_i$ in the basis, is also kept. Indices of states with multiply hopping possibilities get replicated accordingly.
4. The phase factor of each hopping process, calculated according to (2.21), is also stored.
5. The integer list of resulting states $I_y = \text{Int}(|\Psi'_\sigma\rangle)$ is appended by integer values of states not reached by a hopping process.
6. The resulting integer list and corresponding phases are sorted. The original positions of the now sorted integers in I_y are the y-indices of the H_σ matrix.

The H_σ matrices are now represented in the single spin-channel basis states. This is also the reason why they are easily storable, since the single spin basis size is tremendously smaller than the whole Hubbard basis. To get the indices of the full hopping-matrix, the \downarrow -spin indices have to be replicated and adjusted to the whole combined basis. Since the \uparrow -spin basis is the MSB part of the bit representation and each \uparrow -spin basis state has the same \downarrow -spin part, to the matrix indices of H_\downarrow a factor $n \cdot G_\downarrow$, where $n = 0, 1, \dots, G_\uparrow - 1$, and G_σ the number of basis states in the respective spin-channel, is added to obtain the indices of the whole Hubbard Hamilton matrix.

The matrix indices of H_\uparrow are the MSB-indices. To get the corresponding indices of the full basis the replicated factor $n = 1, 2, \dots, G_\downarrow$ has to be added to the index corresponding to $(H_\uparrow - 1) \cdot G_\downarrow$.

For 'very large' systems, $L > 16$, the storage of the Hamilton matrix is not possible anymore because of memory constraints due to the enormous size of the Hilbert space. In this case only the action of the matrix on a vector $H|\Psi\rangle = |\Psi'\rangle$ has to be programmed as actually only this is needed in the Lanczos algorithm to calculate the ground state of the system.

But in this work the Hamilton matrix was always stored in its sparse form (only storing the non-zero elements and corresponding indices), which is possible for lattices up to $L = 16$ sites with use of symmetries, due to its sparseness. There are very little non-zero matrix elements because of the few hopping possibilities, which are at most $N_\sigma \cdot N_c$ for each spin channel, where N_c is the coordination number of the lattice. This leads to at most $N_\uparrow \cdot N_\downarrow \cdot N_c^2$ non-zeros elements per state, but most of the time even less due to prohibited hoppings due to the Pauli exclusion principle. Figure (2.3) shows matrix representation of the non-zero elements of the $N_\uparrow = N_\downarrow = 5$ 3×3 -square lattice Hubbard Hamiltonian. Although the Hilbert space is of size $\binom{9}{5}^2 = 15876$ there are only 333396 of possible $15876^2 = 0.25 \cdot 10^9$ matrix elements non-zero, which corresponds to only 0.13%.

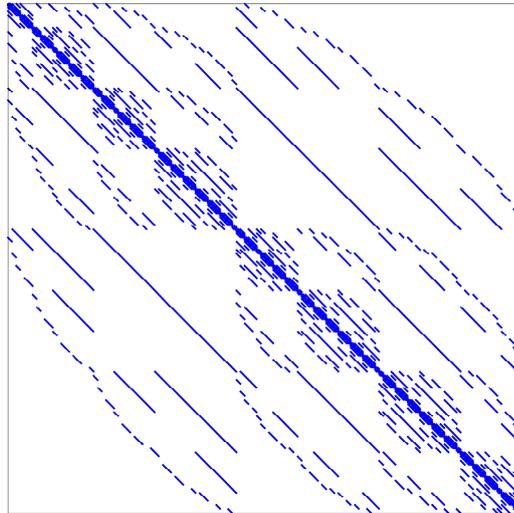


Figure 2.3: Matrix representation of the Hubbard Hamiltonian on a 3×3 -square lattice with $N_\uparrow = N_\downarrow = 5$ \uparrow - and \downarrow -spin electrons and periodic boundary conditions. Only $0.33 \cdot 10^6$ of possible $0.25 \cdot 10^9$ matrix elements, 0.13%, are non-zero.

By constructing the Hamilton matrix and storing it for further computation, memory becomes the restricting factor in the ED implementation. The biggest system calculated without the use of additional symmetries was the 4×4 -square lattice with $N_\uparrow = 5, N_\downarrow = 7$ electrons with a Hilbert space dimension of $N = 49.97 \cdot 10^6$, with the demanding memory cost to store the Hamilton matrix of 52.12 Gb.

But even in the matrix-free implementation of the Lanczos algorithm, where only three vectors with length of the Hilbert space size need to be stored in memory, due to the exponential growth memory rapidly becomes the restricting factor. In table (2.3) the required memory to store three vectors of *double* data type of the length of the Hilbert space size and the Hamilton matrix in sparse form with two integer index lists of *single* data type and a list of *double* type for the matrix elements. For an estimate of the length of the sparse-matrix lists the worst-case half-filled system where every electron in each state can hop to all the nearest neighbor sites is assumed. This causes $4 \dim \mathcal{H} N_\uparrow N_\downarrow$ off-diagonal matrix elements in the two-dimensional case, where $\dim \mathcal{H}$ is the size of the Hilbert space.

Table 2.3: Size of the Hilbert space and memory requirements of saving three vectors $|\Psi\rangle$ and the Hamilton matrix in sparse form. For system sizes $L_x \cdot L_y > 18$ without use of symmetry and $L_x \cdot L_y > 20$ with symmetries memory restrictions even cause the matrix-free Lanczos implementation to cease to work. The maximum system size, without use of symmetries, which could be handled in this work, where the Hamilton matrix also got stored, was the quarter filled 4×4 -lattice. With use of symmetries the half-filled 4×4 -lattice was manageable.

L_x	L_y	N_\uparrow	N_\downarrow	$\dim \mathcal{H}$	memory $3 \times \Psi\rangle$	memory H
2	2	2	2	36	865 b	9.00 kb
3	3	5	5	15876	372.09 kb	24.22 Mb
4	4	4	4	$3.31 \cdot 10^6$	75.81 Mb	3.16 Gb
4	4	4	6	$14.57 \cdot 10^6$	222.39 Mb	10.42 Gb
4	4	5	7	$49.97 \cdot 10^6$	762.48 Mb	52.12 Gb
4	4	6	6	$64.13 \cdot 10^6$	1.47 Gb	137.60 Gb
4	4	8	8	$165.64 \cdot 10^6$	3.71 Gb	631.85 Gb
18	1	9	9	$2.36 \cdot 10^9$	52.83 Gb	5.57 Tb
19	1	10	10	$8.53 \cdot 10^9$	190.74 Gb	24.84 Tb
5	4	10	10	$34.13 \cdot 10^9$	762.97 Gb	198.69 Tb
21	1	11	11	$124.41 \cdot 10^9$	2.72 Tb	438.11 Tb
5	5	13	13	$27.04 \cdot 10^{12}$	590.29 Tb	259.79 Pb

2.4 Application of translational and point-group symmetries

For periodic boundary conditions the most obvious symmetry to further reduce the size of the Hilbert space is translational invariance. Due to the fact that the translation operator T commutes with H , for periodic boundary conditions, the Hamilton matrix acquires block structure in a eigenbasis of T . In one dimension the application of the translation operator on a general state is the displacement of each electron one lattice site in the positive x -direction

$$T c_{i,\sigma}^\dagger c_{i'<i,\sigma}^\dagger \dots c_{j,\bar{\sigma}}^\dagger c_{j'<j,\bar{\sigma}}^\dagger \dots |0\rangle = c_{i+1,\sigma}^\dagger c_{i'+1,\sigma}^\dagger \dots c_{j+1,\bar{\sigma}}^\dagger c_{j'+1,\bar{\sigma}}^\dagger \dots |0\rangle \quad (2.26)$$

with possible phases factor if a hopping over the boundary $i > L$ occurs. For eigenstates $|\Psi\rangle$ of T

$$T|\Psi\rangle = e^{ik} |\Psi\rangle \quad (2.27)$$

the discrete lattice momenta or wave vectors k are determined by the fact

$$T^L = \mathbb{1} \quad \rightarrow \quad k = \frac{2\pi n}{L}, \quad n = 0, 1, \dots, L-1. \quad (2.28)$$

In on dimension the eigenbasis of T can be constructed by the application of the projector [32]

$$P_k = \frac{1}{L} \sum_{j=0}^{L-1} e^{\frac{2i\pi jk}{L}} T^j, \quad (2.29)$$

which commutes with the Hamiltonian

$$[P_k, H] = 0, \quad (2.30)$$

since T commutes with H . And the projector property $P_k^2 = P_k$ can also easily be verified

$$P_k^2 = \frac{1}{L^2} \sum_{j,l}^{L-1} e^{\frac{2i\pi(j+l)k}{L}} T^{j+l} \underbrace{=}_{l \rightarrow -l} \frac{1}{L} \sum_j e^{\frac{2i\pi jk}{L}} T^j = P_k, \quad (2.31)$$

The projector applied to an general basis state $P_k|\Psi\rangle$ is an eigenstate of T with eigenvalue $e^{-\frac{2\pi ik}{L}}$ since

$$TP_k|\Psi\rangle = \frac{1}{L} \sum_{j=0}^{L-1} e^{\frac{2\pi ijk}{L}} T^{j+1}|\Psi\rangle = e^{-2\pi ik} P_k|\Psi\rangle. \quad (2.32)$$

The generalization to two dimension is straightforward given by

$$[T_x, H] = [T_y, H] = [T_x, T_y] = 0, \quad (2.33)$$

$$P_{\mathbf{k}} = \frac{1}{\sqrt{L_x L_y}} \sum_{r_x=0}^{L_x-1} \sum_{r_y=0}^{L_y-1} e^{-i(k_x r_x + k_y r_y)} T_y^{r_y} T_x^{r_x}, \quad (2.34)$$

$$k_{x/y} = \frac{2\pi n_{x/y}}{L_{x/y}}, \quad n_{x/y} = 0, 1, \dots, L_{x/y} - 1, \quad (2.35)$$

with $L_{x/y}$ being the linear lattice sizes in x/y -direction. The details of the procedure of constructing these eigenstates is described below in combination with point-group symmetries. Represented in this basis the original Hamilton matrix decomposes into $L = L_x \cdot L_y$ sub-matrices, associated with an discrete lattice momentum \mathbf{k} , which are generally a factor $1/L$ smaller than the original matrix. The $1/L$ is not strictly the Hilbert space reduction since there are possible states which are incompatible with certain lattice momenta, as discussed below. In figure (2.4) the Hamilton matrix of the above mentioned (2.3) system (3×3 -square lattice, $N_\uparrow = N_\downarrow = 5$) with (b) and without (a) use of translation symmetry is displayed.

To implement both translational and lattice point group symmetries, eigenstates of the combined symmetry group have to be constructed. The combined spatial symmetry group G of a system is the direct product of the set of translations TG and point group elements PG [33]

$$G = TG \otimes PG. \quad (2.36)$$

Since the translational group is abelian, it has a one dimensional irreducible representation [34], which can be associated to a wave vector \mathbf{k} . The character of this representation is

$$\chi_{\mathbf{k}}(t) = e^{i\mathbf{k} \cdot \mathbf{r}(t)}, \quad \forall t \in TG, \quad (2.37)$$

with $\mathbf{r}(t)$ being a translation operation of the translation group. For the combination of translational and point-group symmetries only those elements of PG , are allowed

which leave a given reciprocal lattice vector \mathbf{k} invariant

$$PG_{\mathbf{k}} = \{p \in PG | p(\mathbf{k} = \vec{k})\}. \quad (2.38)$$

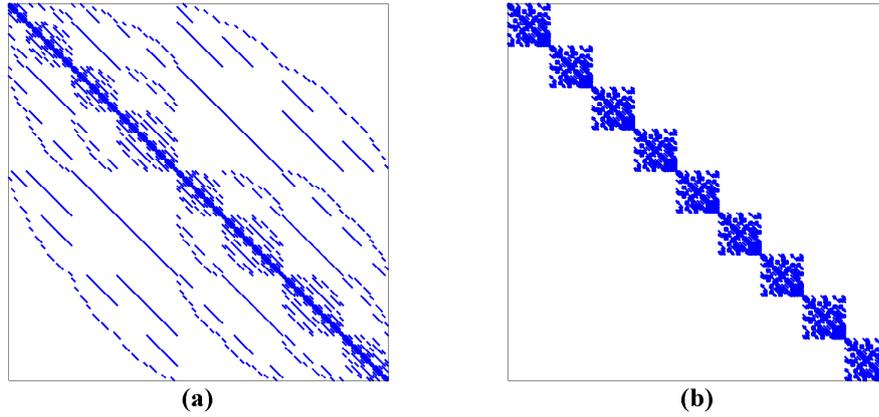


Figure 2.4: Matrix representation of the Hubbard Hamiltonian on a 3×3 -square lattice with $N_{\uparrow} = N_{\downarrow} = 5$ \uparrow - and \downarrow -spin electrons and periodic boundary conditions. (a) Without use of translation symmetry the full $N \times N$ matrix with $N = 15876$ would have to be used in the Lanczos diagonalization. (b) In the eigenbasis of the translation operator T the original matrix decomposes into blocks of size, in this case exactly, $N/L \times N/L$ with $N/L = 15876/9 = 1764$.

In this work only the two-dimensional square lattice was considered, which technically possesses the C_{4v} , in Schönflies, or $4mm$ in Hermann-Mauguin notation, point group. This point group consists of the identity operation E , the clock and counter-clock-wise rotation of 90° C_4^\pm , the 180° rotation C_2 and the two vertical reflections with respect to the x - and y -axis and the two dihedral reflections along the diagonal ($x = y$)- and the orthogonal ($x = -y$)-axis. But as mentioned above for a given wave vector \mathbf{k} only point group symmetries conserving \mathbf{k} are applicable. The Brillouin zone of a two-dimensional square lattice with lattice spacing $a = 1$ is again a square lattice and is displayed in fig. (2.5). The compatible point group symmetries for a given wave vector are listed in table (2.4).

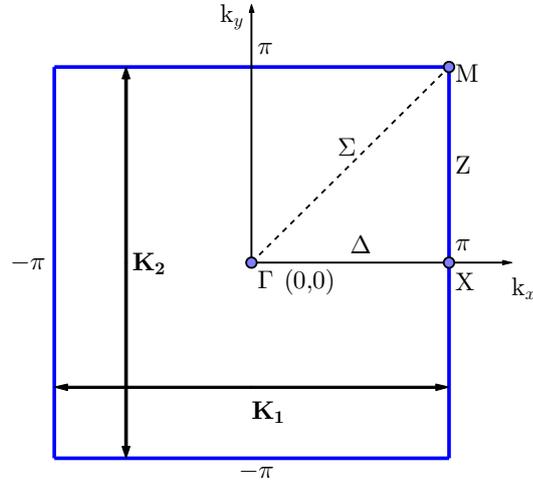


Figure 2.5: Brillouin zone of the two-dimensional square lattice with lattice spacing $a = 1$. Points of high symmetry are displayed with the usual notation. Since every \mathbf{k} vector is equivalent to \mathbf{k}' vectors differing by a reciprocal lattice vector \mathbf{K} e.g. points along the line of symmetry Z between the points $M = (\pi, \pi)$ and $X = (\pi, 0)$ are equivalent to $\mathbf{k} - \mathbf{K}_1$ -vectors, and hence also possess σ_x symmetry.

Table 2.4: Compatible point-group symmetries for a given \mathbf{k} -vector in the two-dimensional square lattice.

\mathbf{k}	applicable symmetry operations	corresponding point group
$\begin{pmatrix} k_x \\ k_y \end{pmatrix}$	E	C_1
$\begin{pmatrix} k_x \\ k_x \end{pmatrix}$	E, σ_d	C_s
$\begin{pmatrix} k_x \\ -k_x \end{pmatrix}$	E, σ_e	C_s
$\begin{pmatrix} 0 \\ k_y \end{pmatrix}, \begin{pmatrix} \pi \\ k_y \end{pmatrix}$	E, σ_x	C_s
$\begin{pmatrix} k_x \\ 0 \end{pmatrix}, \begin{pmatrix} k_x \\ \pi \end{pmatrix}$	E, σ_y	C_s
$\begin{pmatrix} \pi \\ 0 \end{pmatrix}, \begin{pmatrix} \pi \\ \pi \end{pmatrix}$	$E, \sigma_x, \sigma_y, C_2$	C_{2v}
$\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \pi \\ \pi \end{pmatrix}$	$E, \sigma_x, \sigma_y, C_2, C_4^\pm, \sigma_d, \sigma_e$	C_{4v}

Since a wave vector \mathbf{k} is only determined up to addition of a primitive reciprocal lattice vector the wave vectors $\begin{pmatrix} \pi \\ k_y \end{pmatrix}$ and $\begin{pmatrix} k_x \\ \pi \end{pmatrix}$ are also invariant to a σ_x , respectively σ_y , reflection. To construct the corresponding symmetry-adapted basis states for a

given \mathbf{k} vector the irreducible representations (irreps) of the compatible symmetry group have to be known. Those are tabulated and easily available for every important symmetry group, mostly in form of character tables. For a given irrep, denoted by \mathbf{k} and a second quantum number $\mathbf{p}_\mathbf{k}$ for distinction within the compatible point group, a symmetry-adapted basis state for a single spin-channel is obtained, similar to the translation-only case, by applying a projector onto a basis state $|\Psi_\sigma\rangle_i$

$$|\Phi_\sigma\rangle_i = \frac{1}{N_{\Psi_{i\sigma}, \mathbf{k}, \mathbf{p}_\mathbf{k}}} \sum_{T \in TG, P \in PG_\mathbf{k}} \chi_\mathbf{k}(T) \chi_{\mathbf{p}_\mathbf{k}}(P) (T \circ P) |\Psi_\sigma\rangle_i. \quad (2.39)$$

$\chi_{\mathbf{p}_\mathbf{k}}(P)$ is the tabulated character of the irrep $\mathbf{p}_\mathbf{k}$ within the applicable point-group and $N_{\Psi_{i\sigma}, \mathbf{k}, \mathbf{p}_\mathbf{k}}$ the, later discussed, normalization factor. Such states are invariant under the corresponding symmetry operation. The character tables of the compatible point groups in the two-dimensional square lattice (2.4) are given in table (2.5)

Table 2.5: Character tables of the point groups C_s , C_{2v} and C_{4v} . The header of each table denotes the symmetry classes of the group. Integer values in the header of character tables indicate the number of elements of this symmetry class, so e.g. the C_{4v} symmetry group consist of two symmetry operations of class C_4 being the counter- and clock-wise 90° rotations C_4^+ and C_4^- . The row entries denote the irreducible representations within each symmetry group. Roman letters A and B denote one-dimensional irreps, while E specifies a two-dimensional irrep in C_{4v} .

C_s			C_{2v}					C_{4v}				
E	C_h		E	C_2	σ_x	σ_y	E	$2C_4$	C_2	$2\sigma_v$	$2\sigma_d$	
A'	1	1	A_1	1	1	1	1	1	1	1	1	
A''	1	-1	A_2	1	1	-1	-1	A_2	1	1	-1	
			B_1	1	-1	1	-1	B_1	1	-1	1	
			B_2	1	-1	-1	1	B_2	1	-1	-1	
								E	2	0	-2	
										0	0	

The two-dimensional irrep E , not to be confused with the identity operation, can be reduced to the two one-dimensional irreps B_1 and B_2 of the subgroup C_{2v} by the subduction chain $C_{4v} \supset C_{2v}$ because the characters of E of C_{4v} can be obtained by an direct sum of $B_1 \oplus B_2$ of C_{2v} [32]:

		E	C_2	σ_x	σ_y		
C_{2v} :	B_1	1	-1	1	-1		
	B_2	1	-1	-1	1		
C_{4v} :	E	2	-2	0	0	$B_1 \oplus B_2$	

For a given \mathbf{k} vector and an irrep of the compatible symmetry group the application of the projector (2.39) on an original basis state, to form a symmetry-adapted basis, yields a sum of symmetry connected states. It is sufficient to only store one element of (2.39) as the representative of this so called 'cycle', since all other elements can be constructed from it. The calculation of the normalization is more involved, since it is not just the number of elements of a cycle. For certain electron configurations a given symmetry operation only acts like the identity operation, so only the number of unique elements of a cycle count towards the normalization. Additionally there is also the possibility of cancellation of states in subsets $G' \subset G$ when the characters χ cancel each other

$$\sum_{T \in TG, P \in PG_{\mathbf{k}}} \chi_{\mathbf{k}}(T) \chi_{\mathbf{p}_{\mathbf{k}}}(P) = 0, \quad \text{for } (T \circ P)|\Psi\rangle = |\Psi\rangle, \forall T, P \in G'. \quad (2.40)$$

Such states are incompatible to the given symmetry operation and must not be used in the basis. This means for a given basis set only the representative states which are not related by the given symmetry operation and the normalization factor are enough to represent the new symmetrized basis set. It is essentially arbitrary which element to store as representative, but again for reason of faster searching it is optimal to store the state with the lowest corresponding integer representation. For the Hubbard Hamiltonian with the two distinct basis state sets $|\Psi_{\uparrow}\rangle$ and $|\Psi_{\downarrow}\rangle$ which form the full Hubbard basis $|\Psi\rangle = |\Psi_{\uparrow}\rangle \otimes |\Psi_{\downarrow}\rangle$ a possible calculation of the whole basis set is as follows [35]:

Creating the symmetry-adapted basis

1. For a given symmetry sector $(\mathbf{k}, \mathbf{p}_{\mathbf{k}})$ loop over the \uparrow -spin basis state set $|\Psi_{\uparrow}\rangle_i$ and calculate the symmetry-adapted basis according to (2.39). Only store the unique lowest integer basis states $|R_{\uparrow}\rangle$ as representatives of each symmetrized cycle and keep track of the subset of symmetry operations leaving the representatives invariant

$$E_{\mathbf{k}} \subset G_{\mathbf{k}} = \{(T \circ P) \in G_{\mathbf{k}} | (T \circ P)|R_{\uparrow}\rangle = |R_{\uparrow}\rangle\}. \quad (2.41)$$

For faster calculation of the Hamilton matrix it is advantageous to save a list I_\uparrow relating the elements of the \uparrow -spin basis states to their representatives containing the position of the representative, the fermionic phase factor, acquired due the symmetry operation, and the characters $\chi_{\mathbf{k}}(T)\chi_{\mathbf{p}_{\mathbf{k}}}(P)$ of the symmetry operation relating them. The calculation of the normalization factor and the elimination of incompatible states are not yet done at this stage.

2. The whole Hubbard basis is created by all possible combination of the \uparrow -spin representatives $|R_\uparrow\rangle$ and the \downarrow -spin basis states $|\Psi\rangle = |R_\uparrow\rangle \otimes |\Psi_\downarrow\rangle$. The symmetry operations $E_{\mathbf{k}}$ which leave the \uparrow -spin representatives invariant are then applied to the \downarrow -spin basis states and only lowest integer representatives of these cycles are stored. Since for the full Hubbard basis all \uparrow -spin representatives have to be combined with each possible \downarrow -spin state the combination of $|R_\uparrow\rangle_i \otimes |\Psi_\downarrow\rangle_j$ is automatically a representative of a cycle of the corresponding symmetry operation, and due to the ordering of both sets even the lowest integer one, except for symmetry operations which leave $|R_\uparrow\rangle_i$ invariant. The normalization $N_{\Psi_{\sigma_i, \mathbf{k}, \mathbf{p}_{\mathbf{k}}}}$ of the combined basis states $|\Psi\rangle = |R_\uparrow\rangle \otimes |\Psi_\downarrow\rangle$ is calculated and incompatible states (2.40) are eliminated at this step. Again for faster calculation of the Hamilton matrix a list of indices now relating the \downarrow -spin states to their representative is also stored.

2.4.1 Hamilton matrix

For the calculation of the Hamilton matrix the Hamilton operator can again be split in diagonal and spin-dependent off-diagonal parts as in (2.18). Since the projection operator P_k commutes with H (2.30) and fulfills $P_k^2 = P_k$ (2.31) the diagonal elements calculation reduces to the counting of doubly occupied sites of the representatives of a cycle

$$\begin{aligned} \langle \Phi | H_d | \Phi \rangle &= \frac{\langle \Psi | P_k^\dagger H_d P_k | \Psi \rangle}{\langle \Psi | P_k^\dagger P_k | \Psi \rangle} = \frac{\langle \Psi | P_k^2 H_d | \Psi \rangle}{\langle \Psi | P_k^2 | \Psi \rangle} \\ &= h_d(|\Psi\rangle) \frac{\langle \Psi | P_k | \Psi \rangle}{\langle \Psi | P_k | \Psi \rangle} = h_d(|\Psi\rangle). \end{aligned} \quad (2.42)$$

With $h_d(|\Psi\rangle)$ denoting the diagonal matrix element of representative $|\Psi\rangle$ and is calculated as in (2.3.1).

The off-diagonal matrix calculation is more complicated when translational and point-group symmetries are used. The Hamiltonian is again split up into \uparrow -spin and \downarrow -spin part (2.19). The full Hubbard basis now consists of representatives of

\uparrow -spin states $|\uparrow_R\rangle$ and if there are no symmetry operations $E_{\mathbf{k}}$ leaving them invariant (2.41) the whole down spin basis set $|\uparrow_R\rangle|\downarrow\rangle$ and other wise representatives of the corresponding cycle $|\downarrow_R\rangle = E_{\mathbf{k}}|\downarrow\rangle$. The off-diagonal \downarrow -spin hopping matrix elements for the case that an \uparrow -spin representative has no symmetry operation leaving it invariant is given by

$$\begin{aligned}\langle\Phi'|H_{\downarrow}|\Phi\rangle &= N_{\Phi'}N_{\Phi}\langle\uparrow'_R|\langle\downarrow'|P_kH_{\downarrow}P_k|\uparrow_R\rangle|\downarrow\rangle \\ &= N_{\Phi'}N_{\Phi}\langle\uparrow'_R|\langle\downarrow'|P_kH_{\downarrow}|\uparrow_R\rangle|\downarrow\rangle \\ &= h_{\downarrow}(\Phi'')N_{\Phi'}N_{\Phi}\langle\uparrow'_R|\langle\downarrow'|P_k|\uparrow_R\rangle|\downarrow''\rangle.\end{aligned}\quad (2.43)$$

And since in the case $E_{\mathbf{k}} = \mathbb{1}$ each resulting combination $|\uparrow_R\rangle|\downarrow''\rangle$ is again a representative of a cycle

$$\langle\uparrow_R|\langle\downarrow|P_k|\uparrow'_R\rangle|\downarrow''\rangle = N_{\Phi}\delta_{\uparrow'_R,\uparrow_R}\delta_{\downarrow,\downarrow''}.\quad (2.44)$$

So in this case the off-diagonal Hamiltonian calculation in principle reduces to the same calculation as without the use of spatial symmetries, except the \uparrow -spin basis part only consists of the representatives.

If there are symmetry operations leaving the \uparrow -spin representatives invariant, implying there can be \downarrow -spin states being non-representatives, the application $H_{\downarrow}|\downarrow_R\rangle$ can lead to states being non-representatives of its corresponding cycle. These states have to be linked to their representatives, including phase factors, through the corresponding list stored in the basis creation process mentioned above.

The \uparrow -spin hopping matrix elements are given by:

$$\begin{aligned}\langle\Phi'|H_{\uparrow}|\Phi\rangle &= N_{\Phi'}N_{\Phi}\langle\uparrow'_R|\langle\downarrow'_R|P_kH_{\uparrow}|\uparrow_R\rangle|\downarrow_R\rangle \\ &= h_{\uparrow}(\Phi'')N_{\Phi'}N_{\Phi}\langle\uparrow'_R|\langle\downarrow'_R|P_k|\uparrow''\rangle|\downarrow_R\rangle,\end{aligned}\quad (2.45)$$

where it is now possible that $|\uparrow''\rangle$ is not a representative of a \uparrow -spin cycle anymore and has to be linked to it through the list stored in the basis creation process. The symmetry operations relating $|\uparrow''\rangle$ and its representative $|\uparrow''_R\rangle$

$$|\uparrow_R\rangle = (T_i \circ P_j)|\uparrow\rangle,\quad (2.46)$$

also stored in a list during the basis creation, and possible symmetry operations $E_{\mathbf{k}}$ leaving $|\uparrow_R\rangle$ invariant have to be applied to the \downarrow -spin basis part $|\downarrow_R\rangle$ in (2.45) to

get the correct representative.

$$(T_i \circ P_j) |\uparrow''\rangle |\downarrow_R\rangle = p(|\uparrow''\rangle) |\uparrow''_R\rangle (E_{\mathbf{k}} \circ T_i \circ P_j) |\downarrow_R\rangle = \underbrace{p(|\uparrow''\rangle) p(|\downarrow_R\rangle)}_{p(\Phi'')} |\uparrow''_R\rangle |\downarrow''_R\rangle, \quad (2.47)$$

with $p(\Phi'')$ containing all possible phase factors from the application of the symmetry operations. This leads for eq.(2.45) to

$$\langle \Phi' | H_{\uparrow} | \Phi \rangle = h_{\uparrow}(\Phi'') N_{\Phi'} N_{\Phi} p(\Phi'') \underbrace{\langle \uparrow'_R | \langle \downarrow' | P_k | \uparrow''_R \rangle | \downarrow''_R \rangle}_{N_{\Phi''} \delta_{\Phi', \Phi''}} \quad (2.48)$$

In figure (2.6) the influence of applied point-group symmetries on the block-structure of the previously compared (2.4) 3×3 -lattice with $N_{\uparrow} = N_{\downarrow} = 5$ electron is displayed. The smallest block with use of point-group symmetries is 220×220 which is about 8 times smaller than the translation symmetry block size (1764×1764) and about 72 times smaller than the original 15876×15876 Hamilton matrix without use of any symmetries.

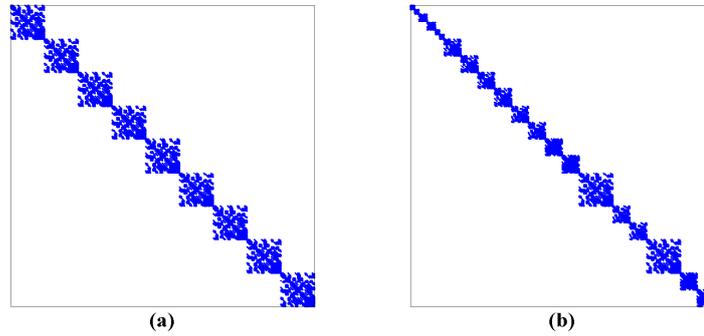


Figure 2.6: Hamilton matrix of the 3×3 -lattice with $N_{\uparrow} = N_{\downarrow} = 5$ electrons. (b) Shows the further reduction of block sizes through the use of lattice point-group symmetries. (a) Is the same as in fig. (2.4b) in the eigenbasis of the translation operator alone. As can be seen in (b) there are some lattice momenta \mathbf{k} which are incompatible with every point-group symmetry and thus no further block size reduction occurs. The smallest block in (b) is 220×220 compared to the 1764×1764 block sizes in (a).

Depending on the compatible point group symmetries $PG_{\mathbf{k}}$ the sizes of the Hamilton matrix in the direct product $G = TG \otimes PG_{\mathbf{k}}$ symmetry-adapted basis is about a factor $N_{\mathbf{k}}N_{\nu}$, where N_{ν} is the number of irreps of the compatible point group and $N_{\mathbf{k}}$ the number of irreps of the translation group, hence the number of lattice sites, smaller as the original matrix. Thus the half-filled 4×4 -lattice is manageable with storage of the Hamilton matrix. The reduced blocks are not exactly a factor $N_{\mathbf{k}}N_{\nu}$ smaller since there still can be incompatible states for a given irrep of the symmetry group.

2.5 Results

In table (2.6) the exact ground state results for various system sizes and parameter $U = 4$ are displayed. Additionally to the ground state energy, with reference results where available, for each system, the total calculation time from the start of the basis calculation until the finish of the Lanczos diagonalization t_{tot} , the block size of the ground state symmetry subspace N_0 , the adjusted calculation time $t_{adj} = t_{tot}N_0/\dim \mathcal{H}$ and the memory needed for the storage of the Hamilton matrix in sparse form for the use of no symmetry, translation symmetry only and combined translation and compatible point group symmetry are listed. Due to the increased complexity in the creation of basis states and calculation of Hamilton matrix elements the total calculation time t_{tot} increases with more extensive symmetry use. Whereas the needed memory to store the Hamilton matrix, which is in the end the limiting factor of this implementation, declines with additional symmetry use.

In most cases a very symmetric \mathbf{k} -vector is associated to the ground state and thus allowing a efficient further reduction of the ground state subspace size through use point group symmetries. But there are cases, e.g. the $L = 16, N_{\uparrow} = 4, N_{\downarrow} = 6$ system, where the \mathbf{k} -point of the ground state has no compatible point group symmetries and thus allows no further reduction of the translational invariant subspace. The maximal manageable memory requirement were the 52.12 Gb needed to store the Hamilton matrix of the $L = 16, N_{\uparrow} = 5, N_{\downarrow} = 7$ without use of symmetry. While the application of symmetries allowed the calculation of the half filled $L = 16$ system.

The adjusted calculation time t_{adj} has to be taken with a grain of salt, since it is only naively assumed as the fraction of the total calculation time corresponding to the ground state subspace size. First of all does the calculation time not linearly scale with subspace size, although this probably favors the adjusted time, since in most cases the ground state symmetry belongs to a highly symmetric and thus highly block diagonalized subspace. But the symmetry of the ground state is generally not known prior to the calculation and one has to diagonalize all subsystems first to

find the subspace of the groundstate.

Another advantage, as mentioned above, is the association of calculated energies to specific irreducible representations of the underlying symmetry group, which consecutively can be associated to good quantum number of the Hamiltonian in question and thus allowing for a more sophisticated physical analysis of system properties. And additionally by calculating the lowest energy states of symmetry subspaces not containing the actual ground state, it is possible to easily calculate low-lying excited states system otherwise not so easily implemented with the Lanczos algorithm.

In figure (2.7) the total needed computational time of the exact diagonalization of the Hubbard model for various system sizes and fillings for $U = 4$ is displayed in a logarithmic plot. The first entries of the x -axis labels denotes the system size $L = L_x \cdot L_y$ and the second the number of \uparrow - and \downarrow -spin electrons $N_\uparrow + N_\downarrow$. The reason for the non strictly monotonic behavior is that the calculations were performed on different powerful computers. But the general trend of exponential scaling with system size is clearly observable. Also the impact of the previously discussed increased implementation complexity of symmetry-adapted bases is identifiable, although the difference in computation decreases with system size. Unfortunately no bigger systems than $L = 16, N_\uparrow + N_\downarrow = 12$ were computable without the use of symmetry to further investigate this behavior.

In figure (2.8) the corresponding memory costs to store Hamilton matrix of the symmetry subspace containing the ground state for the same system sizes and symmetry used as in fig. (2.7) are displayed. The non monotonic behavior for the 'no-symmetry' and translation symmetry case occurs since the system sizes are first ordered by lattice size and then by the filling factor. And although the Hilbert space of the $L = 9, N_\uparrow + N_\downarrow = 2$ system is bigger than the corresponding $L = 4, N_\uparrow + N_\downarrow = 2$ Hilbert space there are less hopping possibilities and thus less off-diagonal matrix elements, which cause the memory requirements to be lower. The smaller memory requirements for the $L = 16, N_\uparrow + N_\downarrow = 5$ compared to the $L = 9, N_\uparrow + N_\downarrow = 8$ system for used translation symmetry also stem from the fact that there are more \mathbf{k} vectors in the bigger $L = 16$ system and thus the Hamilton matrix is more blocked. As mentioned above the varying memory requirements for the use of compatible point group symmetry depend on the \mathbf{k} -point symmetry associated to the ground state of the system. The red line in fig. (2.8) denotes the maximum available RAM memory for computations, and as can be seen the $L = 16, N_\uparrow = 6, N_\downarrow = 6$ is the first non-computable system without use of symmetries, whereas both translation- and translation combined with compatible point group symmetry allow an ED of the half-filled 4×4 lattice.

Table 2.6: Exact diagonalization results of the Hubbard model on the square lattice for various system sizes ($L = L_x \cdot L_y$) and number of electron (N_\uparrow, N_\downarrow) for $U = 4$. $\dim \mathcal{H}$ denotes the size of the Hilbert space. G is the above described application of the direct product combination of the translation group and compatible point groups $G = TG \otimes PG_{\mathbf{k}}$. t_{tot} denotes the total time needed from basis creation until the finish of the Lanczos algorithm to solve for the lowest energy eigenvalue. N_0 is the dimension of the symmetry subspace containing the ground state $|\Psi_0\rangle$. $\frac{t_{tot}N_0}{\dim \mathcal{H}}$ is the theoretical fraction of time needed to diagonalize the ground state symmetry subspace and the last column is the needed memory to store the Hamilton matrix in sparse form. Reference energies from the literature, where available, are documented below the calculated energies.

L	N_\uparrow	N_\downarrow	$\dim \mathcal{H}$	E_0/N	Symmetry	t_{tot} [s]	N_0	$\frac{t_{tot}N_0}{\dim \mathcal{H}}$ [s]	Memory
4	2	2	36	-1.4142	-	0.06	36	0.06	4.5 kb
					TG	1.17	12	0.39	1.4 kb
					G	1.80	6	0.30	768 b
9	3	3	7056	-1.1417	-	0.30	7056	0.30	1.94 Mb
					TG	3.83	792	0.43	222.75 kb
					G	9.96	101	0.15	28.41 kb
9	5	5	15876	-0.699	-	0.47	15876	0.47	12.11 Mb
					TG	3.85	1764	0.43	1.35 Mb
					G	8.65	232	0.14	181.25 kb
16	2	2	14400	-0.7206	-	0.61	14400	0.61	1.76 Mb
			Ref. [36]:	-0.72064	TG	10.21	912	0.65	114.0 kb
					G	33.06	144	0.33	18.0 kb
16	3	3	$0.31 \cdot 10^6$	-0.946	-	6.8	$0.31 \cdot 10^6$	6.8	85.14 Mb
			Ref. [36]:	-0.946	TG	45.4	$19.6 \cdot 10^3$	2.8	5.38 Mb
					G	101.9	$2.47 \cdot 10^3$	0.8	694 kb
16	4	4	$3.31 \cdot 10^6$	-1.0959	-	71.5	$3.31 \cdot 10^6$	71.5	1.58 Gb
			Ref. [36]:	-1.0959	TG	317.7	$0.21 \cdot 10^6$	19.9	102.54 Mb
					G	532.5	$26.07 \cdot 10^3$	4.2	12.73 Mb
16	4	6	$14.58 \cdot 10^6$	-1.1359	-	415	$14.58 \cdot 10^6$	415	10.43 Gb
					TG	1313	$0.91 \cdot 10^6$	82	666.5 Mb
					G	2210	$0.91 \cdot 10^6$	138	666.5 Mb
16	5	7	$49.97 \cdot 10^6$	-1.1053	-	2378	$49.97 \cdot 10^6$	2378	52.12 Gb
					TG	4458	$3.12 \cdot 10^6$	279	3.25 Gb
					G	5849	$1.56 \cdot 10^6$	183	1.63 Gb
16	6	6	$64.13 \cdot 10^6$	-1.1081	-	N/A	$64.13 \cdot 10^6$	N/A	68.8 Gb
			Ref.[36]:	-1.1081	TG	8985	$4.01 \cdot 10^6$	562	4.3 Gb
					G	11163	$0.50 \cdot 10^6$	87	549.32 Mb
16	8	8	$0.17 \cdot 10^9$	-0.85137	-	N/A	$0.17 \cdot 10^9$	N/A	324.25 Gb
			Ref. [36],[37]:	-0.8514	TG	12916	$10.35 \cdot 10^6$	807	19.74 Gb
					G	16896	$1.30 \cdot 10^6$	132	2.48 Gb

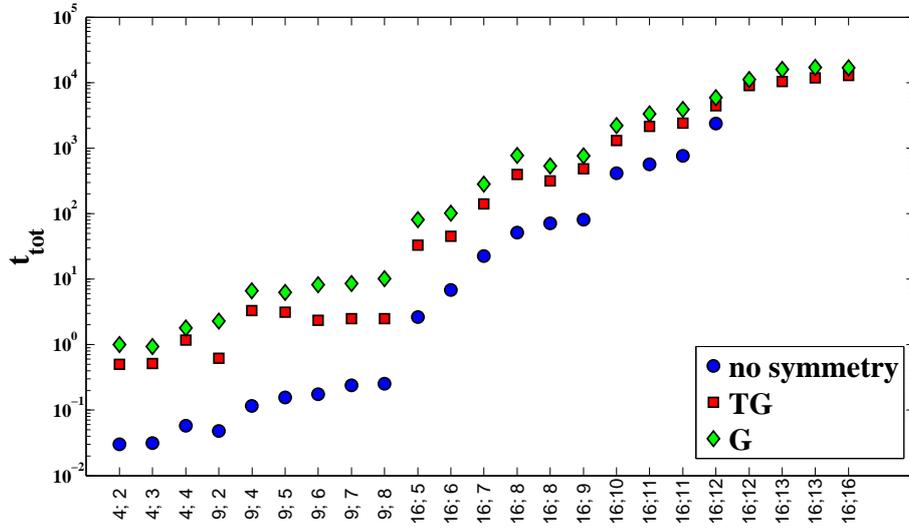


Figure 2.7: Total needed computational time of the exact diagonalization of various systems with and without the use of symmetry-adapted bases.

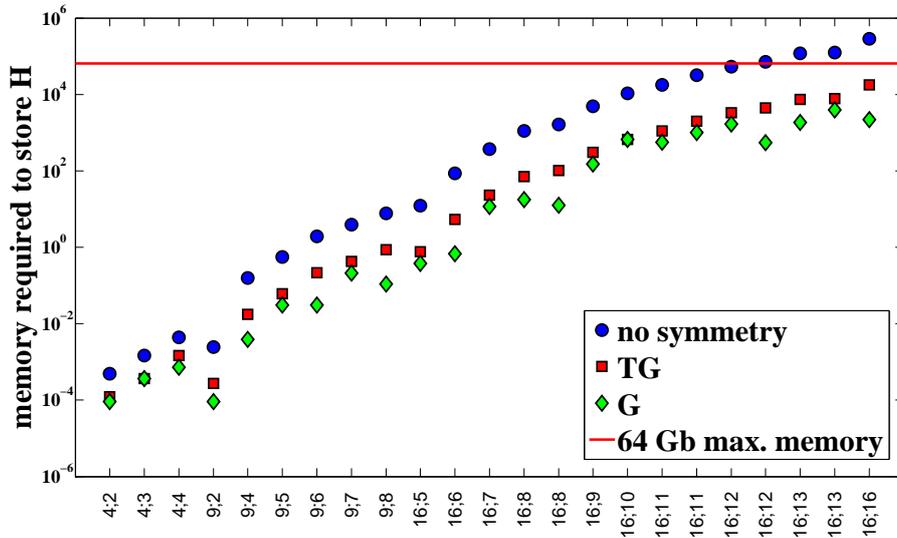


Figure 2.8: Memory requirements to store the sparse-Hamilton matrix in the ground state symmetry subspace. *TG* denotes the use of translation symmetry only and *G* marks the results of combined translation and point-group symmetry.

In fig. (2.9) the memory reduction due to translation- and combined translation-

point-group symmetry compared to the full Hilbert space size are displayed. The translation symmetry implementation generally reduces the size of the subspace by a factor of L , but the efficiency of point-group symmetry usage depends on the symmetry of the ground-state wave vector \mathbf{k} as mentioned above.

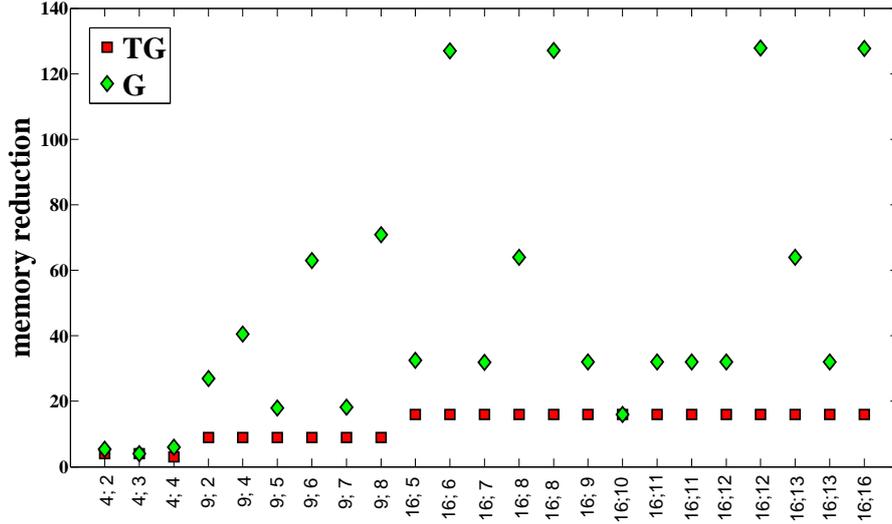


Figure 2.9: Reduction of memory requirements to store the Hamilton matrix of the ground state symmetry subspace with use of symmetries compared to the full Hilbert space. TG denotes the use of translation symmetry only and G marks the results of combined translation and point-group symmetry.

3 Mean-field treatment

With an obtained exact solution for small model systems (2.5) the results of the used momentum-space transformation, introduced in the previous chapter 2.1 and the Hartree-Fock transformation can be analyzed.

3.1 Hartree-Fock approximation

The Hartree-Fock approximation is one of the most prominent approximation techniques and starting point for many more evolved ones. In first quantization formulation the Hartree-Fock (HF) ground state of a Hamiltonian is given by the single wave function $|\Psi_0\rangle$ which variationally minimizes the energy expectation value

$$E_0^{HF} = \langle \Psi_0 | H | \Psi_0 \rangle. \quad (3.1)$$

The wave function $|\Psi_0\rangle = |\chi_1 \chi_2 \dots \chi_N\rangle$ is the total anti-symmetrized N -electron Slater determinant

$$|\Psi(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N)\rangle = \frac{1}{\sqrt{N!}} \begin{vmatrix} \chi_1(\mathbf{x}_1) & \chi_1(\mathbf{x}_2) & \cdots & \chi_1(\mathbf{x}_N) \\ \chi_2(\mathbf{x}_1) & \chi_2(\mathbf{x}_2) & \cdots & \chi_2(\mathbf{x}_N) \\ \vdots & \vdots & \ddots & \vdots \\ \chi_N(\mathbf{x}_1) & \chi_N(\mathbf{x}_2) & \cdots & \chi_N(\mathbf{x}_N) \end{vmatrix}, \quad (3.2)$$

with $\chi_i(\mathbf{x}_j) = \phi_i^\sigma(\mathbf{x}_j)$ being spin orbital states describing an spatial electron orbital ϕ and spin state $\sigma = \{\uparrow, \downarrow\}$. The spatial orbitals are optimized to find the best possible approximation to (3.1). The formulation as a Slater determinant ensures the correct fermionic anti-symmetry properties with regard to particle exchange due to the Pauli exclusion principle.

The general non relativistic N -electron many-body Hamiltonian in the Born-Oppenheimer approximation in Hartree atomic units ($e = m_e = \hbar = 4\pi\epsilon_0 = 1$) reads as

$$H = - \sum_{i=1}^N \frac{1}{2} \nabla_i^2 - \sum_{i=1}^N \sum_{A=1}^M \frac{Z_A}{r_{i,A}} + \sum_{i=1, j>i}^N \frac{1}{r_{i,j}}, \quad (3.3)$$

where the first term describes the kinetic motion of the electrons, the second the attractive interaction between the electron with the M fixed nuclei with charge Z_A and the third the electron-electron repulsion. The minimization of the expectation value of (3.3) by variation of the spin-orbital states $\chi_i(\mathbf{x})$ of the Slater determinant (3.2) leads to the self-consistent Hartree-Fock eigenvalue problem [38]

$$f(i)\chi(\mathbf{x}_i) = \varepsilon_i\chi(\mathbf{x}_i). \quad (3.4)$$

With $f(i)$ being the single-electron Fock operator for electron i

$$f(i) = -\frac{1}{2}\nabla_i^2 - \sum_{A=1}^M \frac{Z_A}{r_{i,A}} + v_i^{HF}, \quad (3.5)$$

where two-particle electron-electron term of (3.3), responsible for the exponential complexity, is approximated by an averaged potential v_i^{HF} , caused by all other electrons, acting on electron i . As a matter of fact the mean field potential v_i^{HF} depends on the spin orbitals of all other $N - 1$ electrons, which means (3.4) has to be solved iteratively until self-consistency. The form of v_i^{HF} depends on the basic symmetry assumptions to the spin-orbitals $\chi_i(\mathbf{x})$ in the Slater determinant (3.2) and the finite basis-set in which they are expressed.

3.1.1 Restricted Hartree-Fock

If the spatial part of the spin-orbital basis functions χ_i are assumed to be the same for the \uparrow - and \downarrow -spin part $\chi_i(\mathbf{x}_j) = \phi_i(\mathbf{x}_j) \cdot \sigma$, the solution to (3.4) is called the restricted Hartree-Fock (RHF) approximation. In this case the orthogonality of two spin orbital χ_i, χ_j with the same spatial orbital ϕ is only ensured by the spin part σ

$$\int \chi_i^* \chi_j d\Omega = \int \phi^* \phi \sigma_i \sigma_j d\Omega = \delta_{\sigma_i, \sigma_j} \int |\phi|^2 d\mathbf{x} = \delta_{\sigma_i, \sigma_j} \quad (3.6)$$

In this case the averaged potential v_i^{HF} of the Fock operator $f(i)$ in (3.5) takes the form [39]

$$v_i^{HF} = \sum_j \int d\mathbf{x}' \frac{|\phi_j(\mathbf{x}')|^2}{|\mathbf{x} - \mathbf{x}'|} - \sum_j \delta_{\sigma_i, \sigma_j} \int d\mathbf{x}' \frac{\phi_j^*(\mathbf{x}') \phi_i(\mathbf{x}')}{|\mathbf{x} - \mathbf{x}'|}. \quad (3.7)$$

Since the Fock operator (3.5) in (3.4) does not explicitly depend on the spin part of χ_i it has to be only solved once for both \uparrow - and \downarrow -spin part. For a given basis set expansion with K spatial function ϕ_i there are then $2K$ corresponding spin

orbitals. From the solution to (3.4) the energetically N -lowest spin orbitals χ_i , the so called occupied orbitals, determine the single Slater determinant ground state, the Hartree-Fock (HF) determinant $|\Psi_0\rangle$. The corresponding $2K - N$ energetically higher orbitals are called the unoccupied orbitals. This is schematically displayed in fig. (3.1a).

3.1.2 Unrestricted Hartree-Fock

A more general approach is to place no restrictions on the spatial parts $\phi_i(\mathbf{x})$ of the spin orbitals χ_i and allow them to differ for opposite spin: $\chi_i(\mathbf{x}_j) = \phi_i^\sigma(\mathbf{x}_j)$. This unrestricted Hartree-Fock (UHF) approach leads to two coupled eigenvalue equations from (3.4) with a Fock operator [40]

$$v_{i,\sigma}^{UHF} = \sum_{\tau=\uparrow,\downarrow} \sum_j \int d\mathbf{x}' \frac{|\phi_j^\tau(\mathbf{x}')|^2}{|\mathbf{x} - \mathbf{x}'|} - \sum_{\tau=\uparrow,\downarrow} \sum_j \phi_j^\sigma(\mathbf{x}) \int \frac{(\phi_j^\tau(\mathbf{x}))^* \phi_i^\tau(\mathbf{x})}{|\mathbf{x} - \mathbf{x}'|} \quad (3.8)$$

for each spin channel. The next iterative solution for spin-part $\sigma(t+1)$ depends on the current solution of the opposing spin part $\sigma'(t)$ and vice versa. As in the RHF case the N energetically lowest orbitals constitute the HF determinant, which now can differ for the two spin channels, schematically displayed in fig. (3.1b) Since there is more freedom in the UHF basis functions opposed to the RHF functions, the variational principle leads to strictly lower energy solutions, except when the two solutions are identical. But a major drawback in the UHF approach is the fact that it breaks the spin-symmetry of the Hubbard Hamiltonian, allowing a spatial overlap between wave functions with different spin

$$\int \phi_i^\sigma \phi_j^{\sigma'} d\Omega \neq \delta_{i,j} \delta_{\sigma,\sigma'}, \quad (3.9)$$

called the spin contamination. And additionally the translational and point group symmetry of the system is also lost in the UHF approach. Although in the half-filled case it is possible to regain partial translation symmetry for the UHF transformation as described below in (3.2.3).

The quality of both the RHF and UHF and approximation depends on the size of the spatial basis set. In the limit of complete basis sets in which the spin orbitals are expanded the HF procedure converges to the so called Hartree-Fock limit energy E_0^{HF} and not the exact ground state of the system.

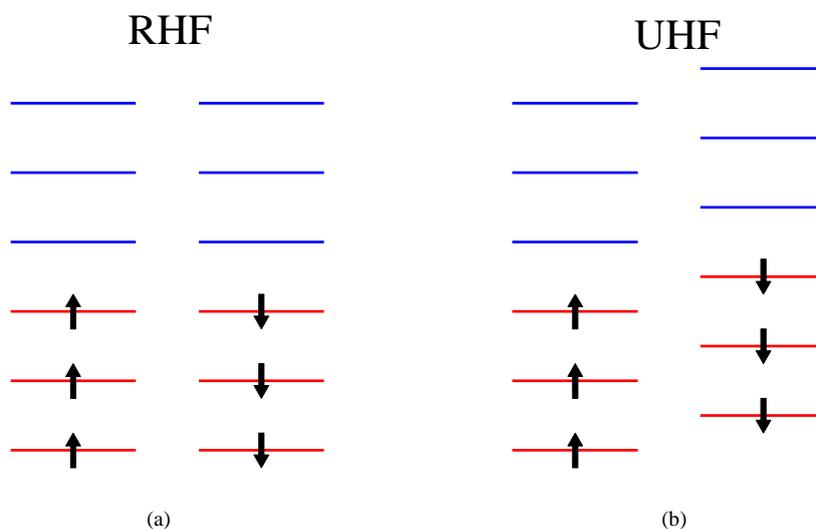


Figure 3.1: Schematic representation of the occupied spin-orbitals of the Hartree-Fock determinant marked red in the RHF (a) and UHF (b) approximation. The energies of the different spin channels are allowed to vary in the UHF approximation leading to a variationally lower energy compared to the RHF approximation.

3.1.3 Configuration interaction

Both in the RHF and UHF approximation there are $2K$ spin orbitals in the basis, from which the N energetically lowest are occupied in the HF determinant. There are however all possible excitations from this HF determinant possible. Including the HF ground state there are totally $\binom{2K}{N}$ possible configurations of occupied and unoccupied orbitals possible. By taking the HF determinant $|\Psi_0\rangle$ as reference determinant an excited determinant is denoted by

$$|\Psi_{a,b,\dots}^{r,s,\dots}\rangle = |\chi_1\chi_2\cdots\chi_r\chi_s\cdots\chi_N\rangle, \quad (3.10)$$

where spin orbitals (r,s,\dots) are occupied as opposed to (a,b,\dots) in the HF determinant. An example is displayed in fig. (3.2). For a complete basis set of spin orbitals $\{\chi_i\}$ all possible Slater determinants (3.10) also form a complete basis set, in which any wave function can be expanded

$$|\Phi\rangle = c_0|\Psi_0\rangle + \sum_{ra} c_a^r |\Psi_a^r\rangle + \sum_{a<b,r<s} c_{ab}^{rs} |\Psi_{ab}^{rs}\rangle + \dots \quad (3.11)$$

The method of expressing and diagonalizing the matrix representation of a Hamiltonian (3.3) in the Hilbert space of all possible Slater determinants (3.10) is called **full configuration interaction (FCI)**. This procedure is basically identical to exact diagonalization as described in chapter 2. Since the FCI procedure is exact the difference of the exact lowest energy eigenvalue of the FCI approach E_0^{CI} and the HF-limit energy E_0^{HF} , obtained by neglecting electron correlation,

$$E_{corr} = E_0^{CI} - E_0^{HF} \quad (3.12)$$

is called the correlation energy. But like ED the FCI approach suffers from exponential scaling with system size, hence it is not applicable in practice, except for very small systems. In chapter 5 the full configuration interaction quantum Monte Carlo (FCIQMC) approach is introduced, which tries to handle this exponential scaling of the FCI procedure by applying quantum Monte Carlo techniques within the subspace of Slater determinants of a system.

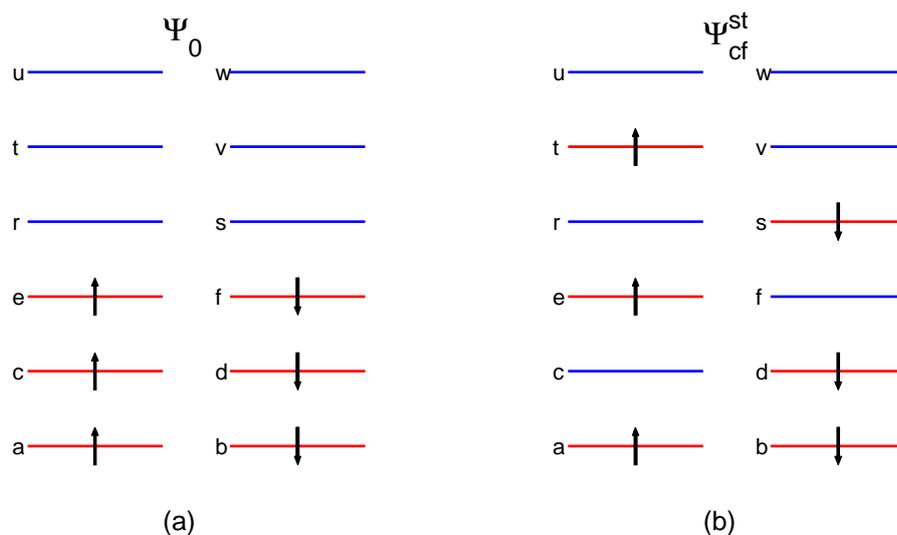


Figure 3.2: An example of an excitation of the HF determinant $|\Psi_0\rangle$ (a), where the previously unoccupied orbitals s, t are now occupied as opposed to orbitals c, f .

3.2 Mean-field approximation to the Hubbard model

The Hartree-Fock approximation is applied to the Hubbard Hamiltonian (3.1.2), in second quantization formulation,

$$H = -t \sum_{\langle i,j \rangle, \sigma} c_{i\sigma}^\dagger c_{j,\sigma} + U \sum_i n_{i,\uparrow} n_{i,\downarrow}. \quad (3.13)$$

by approximating the electron number operator $n_{i,\sigma}$ by its expectation value $\langle n_{i,\sigma} \rangle$ and derivations $\delta n_{i,\sigma}$ thereof

$$n_{i,\sigma} \approx \langle n_{i,\sigma} \rangle + \delta n_{i,\sigma}. \quad (3.14)$$

The real-space basis states (2.11) used in the exact diagonalization section 2 to calculate the matrix representation of the Hubbard Hamiltonian (1.1) can be interpreted as Slater determinants. The fermionic anti-symmetry is automatically included in the second quantization formulation through the anti-commutation relations of the fermionic creation and annihilation operators (1.3). For finite lattice systems the basis states (2.11) are also a complete basis set and thus the obtained energy will be the Hartree-Fock limit energy.

By inserting the mean-field approximation (3.14) into the Hamiltonian (3.13)

$$H^{HF} = -t \underbrace{\sum_{\langle i,j \rangle, \sigma} c_{i\sigma}^\dagger c_{j,\sigma}}_T + U \sum_i (\langle n_{i,\uparrow} \rangle + \delta n_{i,\uparrow}) (\langle n_{i,\downarrow} \rangle + \delta n_{i,\downarrow}) \quad (3.15)$$

and neglecting the $\delta n_{i,\uparrow} \delta n_{i,\downarrow}$ -term yields

$$\begin{aligned} H^{HF} &= T + U \sum_i (\langle n_{i,\uparrow} \rangle \langle n_{i,\downarrow} \rangle + \langle n_{i,\uparrow} \rangle \delta n_{i,\downarrow} + \langle n_{i,\downarrow} \rangle \delta n_{i,\uparrow}) \\ &\stackrel{\delta n_{i,\sigma} = n_{i,\sigma} - \langle n_{i,\sigma} \rangle}{=} T + U \sum_{i,\sigma} (\langle n_{i,\sigma} \rangle n_{i,\bar{\sigma}} - \langle n_{i,\uparrow} \rangle \langle n_{i,\downarrow} \rangle). \end{aligned} \quad (3.16)$$

Through the HF approximation the former two particle electron-electron Coulomb interaction term in (3.13) is replaced by an mean-field electron term $\langle n_{i,\sigma} \rangle$ which can be interpreted as the electron density of spin- σ on site i . The Hamiltonian is now a sum of one-particle operators for each spin σ

$$H^{HF} = H_\uparrow^{HF} + H_\downarrow^{HF} = \sum_{\langle i,j \rangle, \sigma} (t_{ij} + U \langle n_{i,\bar{\sigma}} \rangle \delta_{i,j}) c_{i,\sigma}^\dagger c_{j,\sigma} - U \sum_i \langle n_{i,\uparrow} \rangle \langle n_{i,\downarrow} \rangle, \quad (3.17)$$

where $t_{ij} = -t$ for nearest neighbors (i,j) and zero else and the last term is just a constant factor. As in (3.4) the eigenvalue equation corresponding to (3.17)

$$H_{\sigma}^{HF} |\Psi\rangle = E_{\sigma}^k |\Psi\rangle, \quad (3.18)$$

depends on the solution of the opposite spin solution through the electron density term $\langle n_{\bar{\sigma}} \rangle$ and the coupling constant $\langle n_{i,\uparrow} \rangle \langle n_{i,\downarrow} \rangle$ and has to be solved iteratively until self-consistency.

3.2.1 Restricted Hartree-Fock

For the restricted Hartree-Fock solution of (3.17) the electron density solution is assumed to be independent of lattice site i , and for $N_{\uparrow} = N_{\downarrow}$ identical for both spin channels, thus $\langle n_{i,\uparrow} \rangle = \langle n_{i,\downarrow} \rangle = \langle n \rangle = N_{\sigma}/L$. This yields a Hartree-Fock Hamiltonian independent of spin

$$H_{\sigma}^{HF} = \sum_i (t_{ij} + U \langle n \rangle \delta_{i,j}) c_{i,\sigma}^{\dagger} c_{i,\sigma} - U \sum_i \frac{\langle n \rangle^2}{2}. \quad (3.19)$$

and similar to the tight-binding Hamiltonian t_{ij} only with an additional diagonal on-site terms $U \langle n \rangle$. Thus it can be analytically diagonalized by a momentum space transformation. For the one-dimensional case and omitting spin labels for simplicity, with the transformation

$$c_j^{\dagger} = \frac{1}{\sqrt{L}} \sum_k e^{ikj} b_k^{\dagger} \quad \text{and} \quad c_l = \frac{1}{\sqrt{L}} \sum_{k'} e^{-ik'l} b_{k'}, \quad (3.20)$$

the single spin parts of (3.19) become

$$H^{RHF} = \sum_{k,k'} \frac{1}{L} \sum_{j,l} (t_{jl} + U \langle n \rangle \delta_{j,l}) e^{i(jk-lk')} b_k^{\dagger} b_{k'} - U \sum_i \frac{\langle n \rangle^2}{2}. \quad (3.21)$$

For the generality of the derivation phase factors $e^{\pm iA_x}$ due to later introduced twisted boundary conditions (TBC) are included in the nearest neighbor hopping

terms in one dimension $t_{jl} = -t(\delta_{j,j+1} e^{iA_x} + \delta_{j,j-1} e^{-iA_x})$. This leads to

$$\begin{aligned}
H_a^{RHF} &= \frac{1}{L} \sum_{k,k'} \sum_{j,l} \left[-t (\delta_{j,j+1} e^{iA_x} + \delta_{j,j-1} e^{-iA_x}) + U \langle n \rangle \right] e^{i(jk-lk')} b_k^\dagger b_{k'} \\
&\quad - U \sum_i \frac{\langle n \rangle^2}{2} \\
&= \sum_{kk'} \left[-t (e^{-i(k'-A_x)} + e^{i(k'-A_x)}) + U \langle n \rangle \right] b_k^\dagger b_{k'} \underbrace{\frac{1}{L} \sum_j e^{ij(k-k')}}_{\delta_{k,k'}} \\
&\quad - U \sum_i \frac{\langle n \rangle^2}{2} \\
&= \sum_k \underbrace{(-2t \cos(k - A_x) + U \langle n \rangle)}_{\varepsilon(k)} \hat{n}_k - U \sum_i \frac{\langle n \rangle^2}{2}, \tag{3.22}
\end{aligned}$$

which is similar to the tight binding solution, except an electron density dependent energy offset $U \langle n \rangle$ and a shift of the k values due to the twisted boundary conditions. The total RHF energy is, due to identical spin solutions, given by the N_σ lowest single orbital energies minus the energy offset

$$E_{tot}^{RHF} = 2 \sum_k^{N_\sigma} \varepsilon(k) - U \sum_i^L \langle n \rangle^2. \tag{3.23}$$

3.2.2 Self-consistent unrestricted Hartree-Fock

For the UHF solution there is no constraint on the electron densities $\langle n_{i,\sigma} \rangle$, leading to the coupled eigenvalue problems

$$\left[\sum_{i,j} (t_{ij} + U \langle n_{i,\bar{\sigma}} \rangle \delta_{i,j}) c_{i,\sigma}^\dagger c_{j,\sigma} - \sum_i \frac{\langle n_{i,\uparrow} \rangle \langle n_{i,\downarrow} \rangle}{2} \right] |\Psi_\sigma\rangle = E_\sigma^k |\Psi_\sigma\rangle, \tag{3.24}$$

which have to be solved simultaneously in iterative fashion. For a starting density $\langle n_{i,\uparrow/\downarrow}^0 \rangle$ the matrix (3.24) is diagonalized for the \uparrow - and \downarrow -spin parts and from the N_σ energetically lowest spin orbital solutions Ψ_σ^k , with eigenvalues E_σ^k , the densities

for the next step are constructed by

$$\langle n_{i,\sigma}^{t+1} \rangle = \sum_k^{N_\sigma} (\Psi_{i,\sigma}^k)^* \cdot \Psi_{i,\sigma}^k, \quad (3.25)$$

This process is repeated until the total UHF energy

$$E_{tot}^{UHF} = \sum_k^{N_\uparrow} E_\uparrow^k + \sum_k^{N_\downarrow} E_\downarrow^k \quad (3.26)$$

and the electron densities $\langle n_{i,\sigma} \rangle$ are converged.

The problem of calculating self consistent solutions iteratively is the risk of getting stuck in local minima during the calculations. To counteract to this behavior, inspired by [41], a simulated annealing procedure was coupled to the iterative calculation, by randomly, but dependent on the simulated temperature, modify already converged density solutions.

In table (3.1) various UHF ground state results for half-filled systems are compared to references from literature [42], [43] and as can be seen there is excellent agreement.

In the half-filled case for $U = 4$ the UHF solution adopts an anti-ferromagnetic shape as seen in fig. (3.3) for the 4×4 -lattice and in fig. (3.4) for the 16×16 -lattice, where in the left panel the local difference in \uparrow - and \downarrow -spin densities, the local spin-density $m_i = \langle n_{i,\uparrow} \rangle - \langle n_{i,\downarrow} \rangle$, and in the right panel the local charge density $\rho_i = \langle n_{i,\uparrow} \rangle + \langle n_{i,\downarrow} \rangle$ is displayed. For the half filled case the charge density on each site is exactly equal to one in both examples. Whereas the absolute value of the local spin density $|m_i|$ for both cases is $|m_i^{4 \times 4}| = 0.70449$ and $|m_i^{16 \times 16}| = 0.69066$ uniformly for each lattice site but with alternating signs between nearest-neighbor sites.

Table 3.1: UHF results for half filling and various system sizes with reference results from [42],[43].

L_x	L_y	U	E_{ref}^{UHF}	E_{calc}^{UHF}
4	4	2	-17.55618	-17.556175
4	4	4	-12.56655	-12.566554
4	4	6	-9.37989	-9.379897
4	4	8	-7.38963	-7.389626
4	4	10	-6.06641	-6.066413
8	8	2	-72.39542	-72.395414
8	8	4	-50.99788	-50.997884
8	8	6	-37.93309	-37.933094
8	8	8	-29.81631	-29.816313
8	8	10	-24.42973	-24.429730
10	10	2	-113.54082	-113.540828
10	10	4	-79.70094	-79.700944
10	10	6	-59.26953	-59.269530
10	10	8	-46.58779	-46.587789
10	10	10	-38.17143	-38.171434
12	12	2	-163.75967	-163.759676
12	12	4	-114.77146	-114.771461
12	12	6	-85.34817	-85.3481676
12	12	8	-67.08643	-67.086430
12	12	10	-54.96687	-54.966867
16	16	2	-291.43655	-291.436547
16	16	4	-204.03940	-204.039401
16	16	6	-151.73009	-151.730089
16	16	8	-119.26476	-119.264764
16	16	10	-97.71887	-97.718875

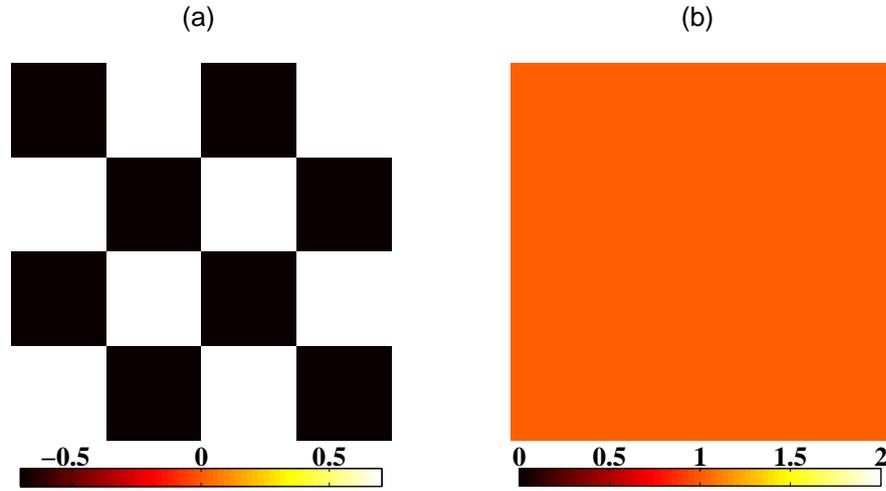


Figure 3.3: (a) Local spin density $m_i = \langle n_{i,\uparrow} \rangle - \langle n_{i,\downarrow} \rangle$ and (b) local charge density $\rho_i = \langle n_{i,\uparrow} \rangle + \langle n_{i,\downarrow} \rangle$ for the 4×4 -lattice at $U = 4$. The charge density ρ_i assumes uniformly the value 1 on each lattice site, whereas the charge density $m_i = \pm 0.70449$ on every site with alternating signs between nearest-neighbor sites.

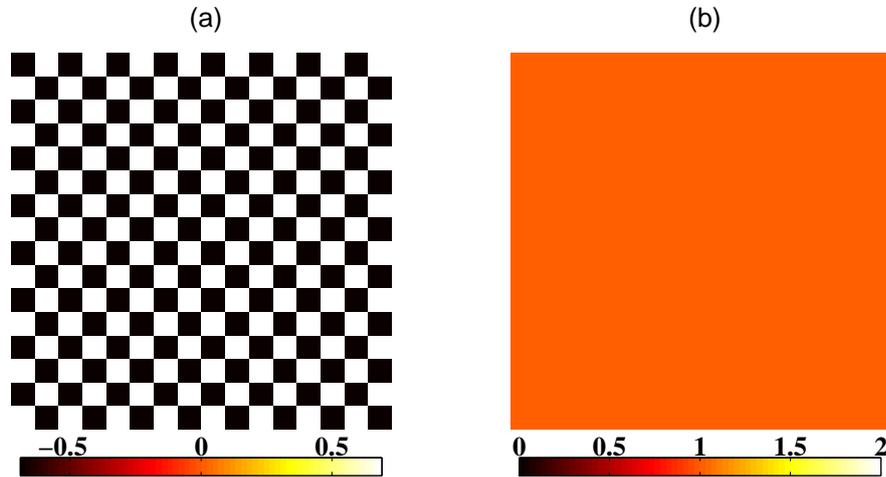
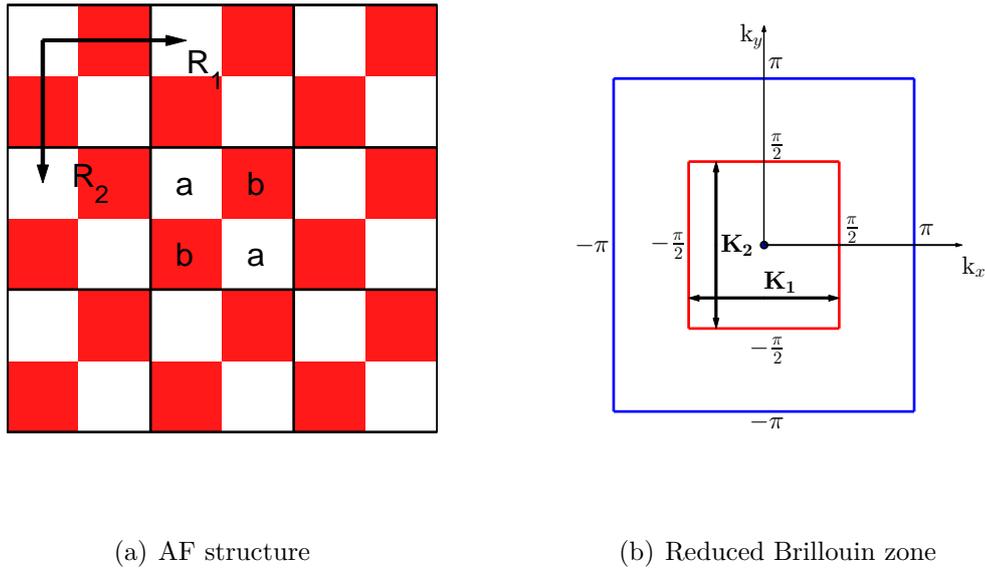


Figure 3.4: (a) Local spin density $m_i = \langle n_{i,\uparrow} \rangle - \langle n_{i,\downarrow} \rangle$ and (b) local charge density $\rho_i = \langle n_{i,\uparrow} \rangle + \langle n_{i,\downarrow} \rangle$ for the 16×16 -lattice at $U = 4$. The charge density ρ_i assumes uniformly the value 1 on each lattice site, whereas the charge density $m_i = \pm 0.69066$ on every site with alternating signs between nearest-neighbor sites.

3.2.3 Analytic unrestricted Hartree-Fock

Due to the anti-ferromagnetic (AF) order of the UHF solution of the Hubbard model in the half-filled case there is the possibility of an analytic formulation of it. Since the solution assumes a periodic form with a repeating 2×2 sub-structure on the square lattice, as seen in fig. (3.3) and (3.4), the solution of the whole $L_x \times L_y$ -lattice is determinable by the solution of the 2×2 structure [44]. The solution of the 2×2 system can be obtained by a Fourier transformation with a unit cell of size 2×2 and twice as large primitive lattice vectors \mathbf{T}'_i as depicted in fig. (3.5a). Due to the AF order in the half-filled case the electronic densities on the sublattice are given by $a_\sigma = \bar{n}_\sigma + \Delta n_\sigma$ and $b_\sigma = \bar{n}_\sigma - \Delta n_\sigma$ with $\bar{n}_\sigma = \frac{N_\sigma}{N}$ being the mean spin- σ electron density per site. For generality of the derivation the hopping matrix t_{ij} is again modified with phase factors $e^{iA_{x/y}}$ due to twisted boundary condition explained below in chapter 6.



(a) AF structure

(b) Reduced Brillouin zone

Figure 3.5: Sublattice structure due to the AF character of the UHF solution on the square lattice in the half filled case. In (a) the color red(white) and letter a(b) indicate the spin density on the different lattice sites. The new primitive lattice vectors $\mathbf{R}_1, \mathbf{R}_2$ are twice as large as in the original lattice. This causes the reduced Brillouin zone (indicated red in (b)) to be quarter as large as the original BZ (blue in (b))

The UHF Hamiltonian (3.13) is mapped on the 2×2 system by associating each

creation (annihilation) operator with a index I denoting the unit cell and second index α for the position within sublattice. Omitting spin labels this transforms

$$c_i^{(\dagger)} \rightarrow c_{I,\alpha}^{(\dagger)} \quad \text{and} \quad t_{ij} \rightarrow t_{\alpha,\alpha'}^{I,J}, \quad (3.27)$$

with same column-first indexing for the sublattice index α as used in the exact diagonalization chapter 2. Within a sublattice $\Delta I = I - J = 0 \rightarrow \mathbf{R} = 0$ the nearest neighbor hopping terms are, as previously, given by $-t e^{\pm i A_{x/y}}$, with a phase factor due to TBC depending on the direction of the hopping process. For inter-sublattice hoppings the unit cells have to be nearest-neighbors $|\Delta I| = 1$ with corresponding primitive lattice vector $\mathbf{R} \in \left\{ \begin{pmatrix} 0 \\ 2 \end{pmatrix}, \begin{pmatrix} -2 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ -2 \end{pmatrix}, \begin{pmatrix} 2 \\ 0 \end{pmatrix} \right\}$. The Fourier transformation of the operators in the UHF Hamiltonian, where the constant energy shift $-U \sum_i \langle n_{i,\uparrow} \rangle \langle n_{i,\downarrow} \rangle$ is pulled into H for now, in terms of the sublattice operators

$$H_\sigma^{UHF} = \sum_{I,J} \sum_{\alpha,\alpha'} \left(t_{\alpha,\alpha'}^{I,J} + U \langle n_{I,\alpha,\bar{\sigma}} \rangle \delta_{I,J} \delta_{\alpha,\alpha'} \right) c_{I,\alpha,\sigma}^\dagger c_{J,\alpha',\sigma} \quad (3.28)$$

is then given by

$$b_{K,\alpha,\sigma} = \frac{1}{\sqrt{N/4}} \sum_I^{N/4} e^{i\mathbf{K}\cdot\mathbf{R}} c_{I,\alpha,\sigma}, \quad (3.29)$$

where \mathbf{R} is the primitive lattice vector associated to the sublattice index I . The wave vectors \mathbf{K} now belong to the reduced Brillouin as depicted in red in fig.(3.5b) which is only quarter as large as the original BZ due to the 2×2 unit cell. But due to the association of the solutions to specific wave vectors, translational symmetry can at least be partially retained [45]. The transformed sublattice Hamiltonian for a given wave vector \mathbf{K} for one spin channel is then given by

$$H_{\alpha,\alpha'}^\sigma(\mathbf{K}) = \begin{pmatrix} U(\bar{n} + \Delta n) & -t e^{-i A_y} (1 + e^{2i K_y}) & -t e^{i A_x} (1 + e^{-2i K_x}) & 0 \\ -t e^{i A_y} (1 + e^{-2i K_y}) & U(\bar{n} - \Delta n) & 0 & -t e^{i A_x} (1 + e^{-2i K_x}) \\ -t e^{-i A_x} (1 + e^{2i K_x}) & 0 & U(\bar{n} - \Delta n) & -t e^{-i A_y} (1 + e^{2i K_y}) \\ 0 & -t e^{-i A_x} (1 + e^{2i K_x}) & -t e^{i A_y} (1 + e^{-2i K_y}) & U(\bar{n} + \Delta n) \end{pmatrix} \quad (3.30)$$

The corresponding matrix for the opposite spin channel $\bar{\sigma}$ is given by exchanging $\Delta n \leftrightarrow -\Delta n$. Independent of spin the analytic eigenvalue solutions in both cases are

given by

$$\lambda_{\sigma}^{1,2,3,4}(\mathbf{k}', \Delta n) = \frac{U}{2} \pm \sqrt{(\Delta n U)^2 + 4t(\cos(k'_x) \pm \cos(k'_y))^2}, \quad (3.31)$$

with the shifted wave vectors $\mathbf{k}' = \mathbf{K} + \mathbf{A}$ caused by the twisted boundary conditions. The total UHF energy of (3.17) for half filling is given by filling up the $N_{\sigma} = N/2$ lowest energy orbitals of (3.31) for both \uparrow - and \downarrow -spin electrons. In the half filled case this are exactly all the energy eigenvalues in the two energetically lower branches of (3.31)

$$E_{tot}^{UHF} = 2 \sum_{\mathbf{k}'}^{N_{\sigma}} \left[U\bar{n} - \sqrt{(\Delta n U)^2 + 4t(\cos(k'_x) \pm \cos(k'_y))^2} \right] - U \sum_i^N \langle n_{i,\uparrow} \rangle \langle n_{i,\downarrow} \rangle. \quad (3.32)$$

The factor 2 in front of the sum stems from the fact that the corresponding sum for the opposite spin electrons is identical, since all eigenvalues are exactly the same as seen in (3.31). In the analytic UHF solution the constant factor at the end of (3.32) can be evaluated. Due to the antiferromagnetic order the electron density on each site i is given by

$$\langle n_{i,\sigma} \rangle = \bar{n} \pm \Delta n \quad \text{and} \quad \langle n_{i,\bar{\sigma}} \rangle = \bar{n} \mp \Delta n. \quad (3.33)$$

By pulling out the constant $U\bar{n}$ factor in each addend in (3.32) and inserting (3.33) into it the total UHF energy as function of electron density difference Δn is given by

$$E_{tot}^{UHF}(\Delta n) = UN(\bar{n} - \bar{n}^2 + \Delta n^2) - 2 \sum_{\mathbf{k}'}^{N_{\sigma}} \sqrt{(\Delta n U)^2 + 4t(\cos(k'_x) \pm \cos(k'_y))^2}. \quad (3.34)$$

The analytic UHF solution is given by the electron density difference Δn minimizing $E_{tot}^{UHF}(\Delta n)$.

In figure (3.6a) the analytic UHF energy as function of spin density difference for the 4×4 -square lattice is plotted. It exactly coincides with the half of the local spin density $|m_i^{4 \times 4}|/2 = 0.35224$ from the self consistent solution from above (3.3). For zero spin density difference the UHF solution would be equivalent to the RHF solution of the system. In figure (3.6b-c) the energy minimizing density difference Δn_{UHF} as function of parameter U for the 4×4 - (b) and 16×16 - (c) lattice is displayed. As can be seen for this system the UHF solution is energetically favorable for infinitesimally $U > 0$ but the magnitude of the jump at $U = 0$ decreases with increasing system size. In (3.6d) the thermodynamic limit $\lim_{L \rightarrow \infty}$ behavior of the density jump at zero U is displayed, and from the perfect quadratic fit to the data

points it can be concluded that the discontinuity of Δn is only a finite size effect.

A sublattice decomposition with a 1×2 unit cell and primitive lattice vectors $\mathbf{R}_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$, $\mathbf{R}_2 = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$ on the square lattice would also be possible. This would cause the reduced Brillouin zone to only be half as small as the original one but at the same time would cause ambiguities which \mathbf{K} vectors belong to the BZ. So for sake of simplicity the 2×2 -decomposition was chosen.

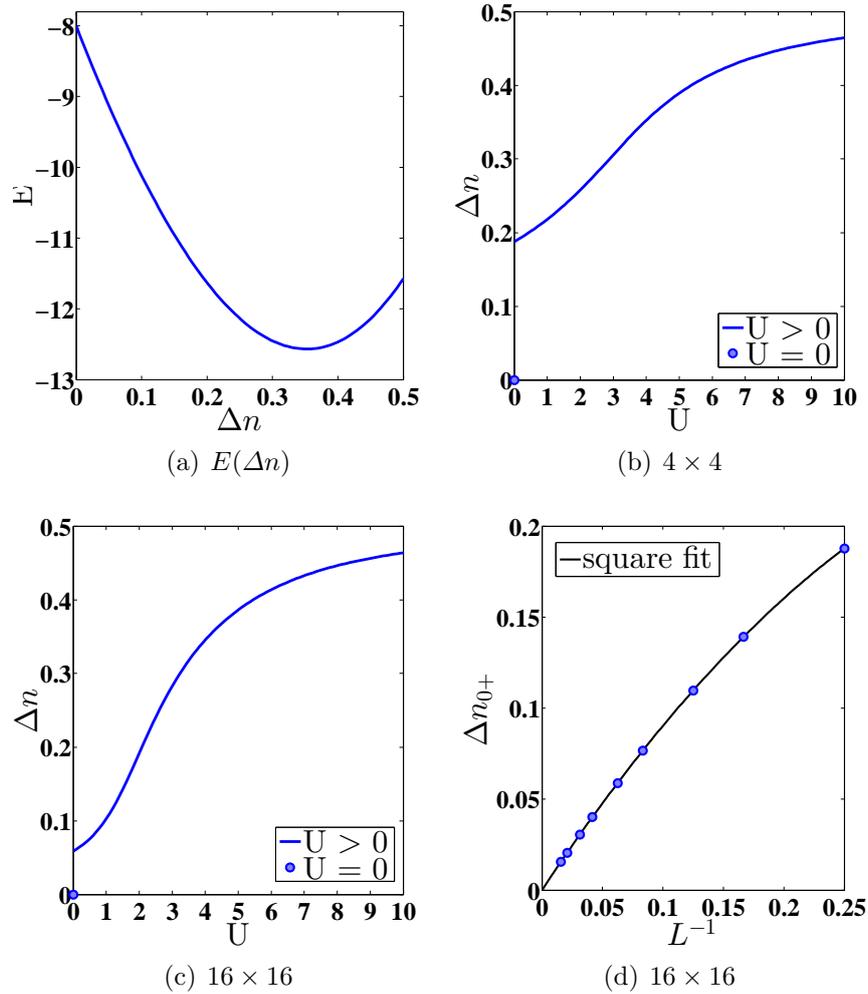


Figure 3.6: (a) Total UHF energy as function of electron density difference for $U = 4$ in the 4×4 -square lattice. Energy minimizing electron density as function of U for the 4×4 - (b) and 16×16 -square lattice (c). (d) Thermodynamic limit behavior of the discontinuity of the energy minimizing electron density difference at $U = 0^+$.

3.2.4 Tilted square lattice

The second lattice used in this work was the tilted square lattice. It is basically a regular square lattice with the same nearest-neighbor hopping possibilities but with different periodic boundary conditions. In fig. (3.7) the 3×3 tilted square lattice is displayed with its primitive lattice vectors $\mathbf{T}_1 = \begin{pmatrix} L_1 \\ L_2 \end{pmatrix} = \begin{pmatrix} 3 \\ 3 \end{pmatrix}$, $\mathbf{T}_2 = \begin{pmatrix} 3 \\ -3 \end{pmatrix}$. As seen in fig. (3.7) the $L_1 \times L_2$ tilted square lattice has $N = \sqrt{2}L_1\sqrt{2}L_2 = 2L_1L_2$ lattice sites.

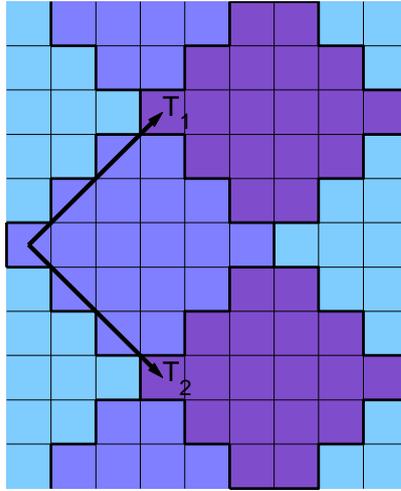


Figure 3.7: Periodicity of the 3×3 -tilted square lattice with 18 sites. The only difference to the regular square lattice are different boundary conditions.

As opposed to the regular square lattice the sublattice transformation of the UHF Hamiltonian (3.24) and subsequent Fourier transformation is unambiguously possible with a 1×2 -unit cell for the tilted square lattice

Due to the 1×2 unit cell there are $N/2$ \mathbf{K} -points in the reduced BZ as seen in (3.8). In the transformed sublattice Hamiltonian there are now one intra-cell in positive or negative x -direction $-te^{\pm iA_x}$ and three inter-cell hoppings possible with associated lattice vectors $\mathbf{R} \in \left\{ \begin{pmatrix} -1 \\ 1 \end{pmatrix}, \begin{pmatrix} -1 \\ -1 \end{pmatrix}, \begin{pmatrix} -2 \\ 0 \end{pmatrix} \right\}$. The Fourier transformed sublattice UHF Hamiltonian, again including phase factors from additional twisted boundary conditions, is given by

$$H_{\alpha,\alpha'}^{\sigma}(\mathbf{K}) = \begin{pmatrix} U(\bar{n} + \Delta n) & T \\ T^* & U(\bar{n} + \Delta n) \end{pmatrix}, \quad (3.35)$$

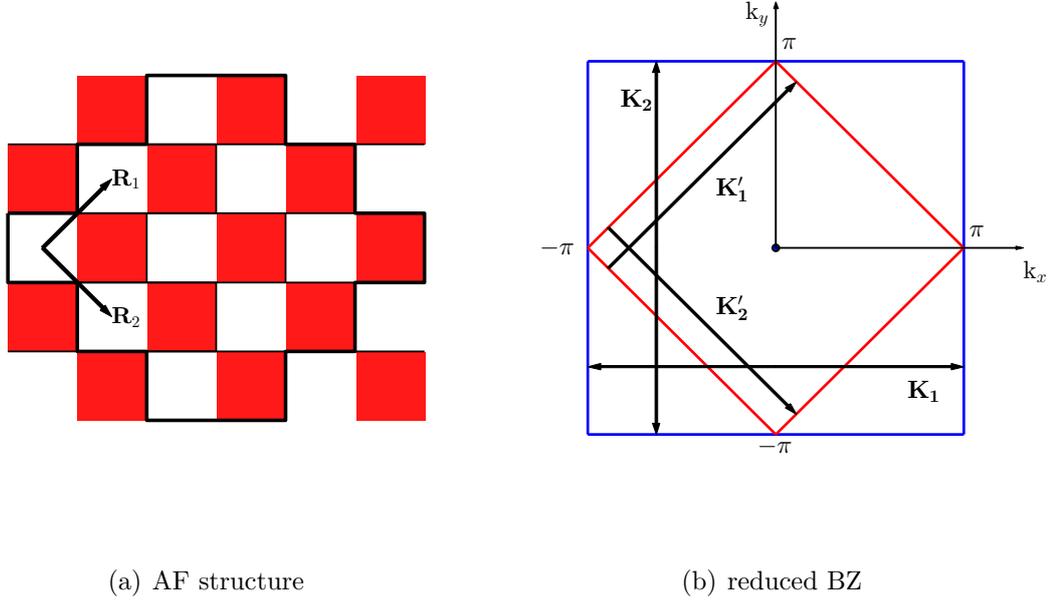


Figure 3.8: (a) 1×2 sublattice structure due to AF order in the tilted square lattice. (b) Original BZ (blue) and half as large reduced BZ (red) of the tilted square lattice due to the 1×2 unit cell.

where the off-diagonal hopping entries are given by

$$T = -t((e^{iA_x}) + e^{iA_y} e^{i(-K_x+K_y)} + e^{-iA_y} e^{i(-K_x-K_y)} + e^{-iA_x} e^{-2iK_x}). \quad (3.36)$$

As above in the regular square lattice the eigenvalues of both spin matrices are identical and given by

$$\lambda_\sigma^{1,2}(\mathbf{k}', \Delta n) = U\bar{n} \pm \sqrt{(\Delta n U)^2 + 4t(\cos(k'_x) + \cos(k'_y))^2}, \quad (3.37)$$

where again the effect of the twisted boundary conditions is a shift of the \mathbf{K} values $\mathbf{k}' = \mathbf{K} + \mathbf{A}$. And for half filling since there are exactly $N/2 = N_\sigma$ \mathbf{k} -points in the reduced BZ this leads to a total UHF energy of

$$E_{tot}^{UHF}(\Delta n) = UN(\bar{n} - \bar{n}^2 + \Delta n^2) - 2 \sum_{\mathbf{k}'}^{N_\sigma} \sqrt{(\Delta n U)^2 + 4t(\cos(k'_x) + \cos(k'_y))^2}. \quad (3.38)$$

The minimization of (3.38) with regard to the density difference Δn again yields the anti-ferromagnetic UHF solution.

The eigenstates, called orbitals from now on, for the UHF basis transformation are obtained by diagonalizing the $2 \times 2/4 \times 4$, depending on the unit cell size, for each \mathbf{K} -point in the reduced BZ after the optimal UHF electron density difference Δn_{UHF} is obtained. The $2 \times 2/4 \times 4$ solutions are replicated accordingly to the whole lattice with their corresponding Bloch phase factor $e^{i\mathbf{K}\mathbf{R}}$. In the 4×4 matrix corresponding to the 2×2 -unit cell for the regular square lattice there are degeneracies in the eigenvalue solutions possible leading to unfavorable linear combinations in degenerate subspace eigenstate solutions.

3.3 Basis transformation

The matrix of eigenvectors U_σ obtained by the solution of the Hartree-Fock eigenvalue problem

$$H_\sigma^{HF} |\Psi\rangle = [t_{ij} + U \langle n_{i,\bar{\sigma}} \rangle \delta_{ij}] |\Psi\rangle = E_\sigma^k |\Psi\rangle \quad (3.39)$$

can be viewed as an unitary basis transformation between the real-space and restricted and unrestricted Hartree-Fock-space. As mentioned the RHF transformation corresponds to the momentum space transformation. These are the bases in which the FCIQMC algorithm was formulated to be applied to the Hubbard model.

3.3.1 Transforming the Hamiltonian

To compare ground state properties in the real-space and Hartree-Fock-space basis representation a transformation and exact diagonalization of the Hubbard Hamiltonian in these bases can be performed. With the unitary transformation of the real-space creation and annihilation operators

$$\vec{b}_\sigma^\dagger = U_\sigma^\dagger \vec{a}_\sigma^\dagger, \quad \text{and} \quad \vec{b}_\sigma = \vec{a}_\sigma U_\sigma \quad (3.40)$$

creation and annihilation operators of Hartree-Fock single orbital states are defined. With the inverse transformation of (3.40)

$$a_{i,\sigma}^\dagger = \sum_l U_{il,\sigma} b_{l,\sigma}^\dagger, \quad \text{and} \quad a_{i,\sigma} = \sum_l (U_\sigma^\dagger)_{li} b_{l,\sigma} \quad (3.41)$$

the real-space Hubbard Hamiltonian can be transformed into Hartree-Fock representation

$$\begin{aligned}
H &= \sum_{ij,\sigma} t_{ij} a_{i,\sigma}^\dagger a_{i,\sigma} + U \sum_i n_{i,\sigma} n_{i,\bar{\sigma}} \\
&= \sum_{ij,\sigma} t_{ij} \sum_{l'} U_{il,\sigma} U_{jl',\sigma}^* b_{l,\sigma}^\dagger b_{l',\sigma} \\
&\quad + U \sum_{i,jklm} U_{ij,\sigma} U_{ik,\sigma}^* U_{il,\bar{\sigma}} U_{im,\bar{\sigma}}^* b_{j,\sigma}^\dagger b_{k,\sigma} b_{l,\bar{\sigma}}^\dagger b_{m,\bar{\sigma}}.
\end{aligned} \tag{3.42}$$

The exact diagonalization of the Hubbard Hamiltonian in HF-space (3.42) is more involved than in its real-space counterpart due to the more general one- and two-particle terms. The one-particle term of (3.42) can be split up into an diagonal $l = l'$ term

$$H_1^{diag} = \sum_{l,\sigma} \left[U_{il,\sigma} t_{ij} U_{jl,\sigma}^* \right] n_{l,\sigma}, \tag{3.43}$$

where the prefactors can be computed as the diagonal elements of the matrix multiplication $(U_\sigma^\dagger t U_\sigma)_{ll}$, and a hopping term

$$H_1^{hop} = \sum_{l',\sigma} \left[U_\sigma^\dagger t U_\sigma \right]_{ll'} b_{l,\sigma}^\dagger b_{l',\sigma}. \tag{3.44}$$

The quantity in the square brackets of eq. (3.44) is again meant to be understood as a matrix multiplication and this hopping term is now not constraint to nearest neighbors, as in the real-space Hubbard Hamilton representation, but a hopping from each HF-state l to every other l' is possible, with a weight factor determined by the matrix element $\left[U_\sigma^\dagger t U_\sigma \right]_{ll'}$. This is still computable with the method described in chapter 2.3, albeit the computational cost is much higher due to the greatly increased hopping possibilities. The increase in computational complexity is even more severe for the in real-space representation originally diagonal on-site interaction term. For optimized implementation the possible terms can be broke up in several terms. For $j = k, l = m$ the 2-particle H_2 term of (3.42) becomes

$$H_2^{diag} = U \sum_{jl} \left[\sum_i U_{ij,\sigma} U_{ij,\sigma}^* U_{il,\bar{\sigma}} U_{il,\bar{\sigma}}^* \right] n_{j,\sigma} n_{l,\bar{\sigma}}, \tag{3.45}$$

which is fully diagonal for both spin channels and is given by the sum of doubly occupied sites times the weight factor $\sum_i U_{ij,\sigma} U_{ij,\sigma}^* U_{il,\bar{\sigma}} U_{il,\bar{\sigma}}^*$.

For $j = k, l \neq m$ a term coined semi-diagonal \uparrow -spin part

$$H_{2,\uparrow}^{semi-diag} = U \sum_{jlm} \left[\sum_i U_{ij,\sigma} U_{ij,\sigma}^* U_{il,\bar{\sigma}} U_{im,\bar{\sigma}}^* \right] n_{j,\sigma} b_{l,\bar{\sigma}}^\dagger b_{m,\bar{\sigma}} \quad (3.46)$$

and the corresponding semi-diagonal \downarrow -spin part for $j \neq k, l = m$

$$H_{2,\downarrow}^{semi-diag} = U \sum_{jkl} \left[\sum_i U_{ij,\sigma} U_{ik,\sigma}^* U_{il,\bar{\sigma}} U_{il,\bar{\sigma}}^* \right] b_{j,\sigma}^\dagger b_{k,\sigma} n_{l,\bar{\sigma}}. \quad (3.47)$$

are obtained. Those terms are diagonal in one spin channel, thus named semi-diagonal, and describe a hopping of the opposite spin electrons. These can also be calculated by the hopping algorithm described in section 2.3 with additionally taking into account the prefactor due to the unitary transformation matrices U_{ij} and the number of electrons with opposite spin due to the $n_{i,\sigma/\bar{\sigma}}$ term.

Until now, expect for the increased density of the Hamilton matrix due to more allowed hoppings no general change of the exact diagonalization procedure, described in chapter 2, occurred. But the computational most costly new term is the fully off-diagonal $j \neq k, l \neq m$ part

$$H_2^{off-diag} = U \sum_{jklm} \left[\sum_i U_{ij,\sigma} U_{ik,\sigma}^* U_{il,\bar{\sigma}} U_{im,\bar{\sigma}}^* \right] b_{j,\sigma}^\dagger b_{k,\sigma} b_{l,\bar{\sigma}}^\dagger b_{m,\bar{\sigma}}, \quad (3.48)$$

which takes into account all combinations of all possible \uparrow - and \downarrow -spin hopping possibilities. The meliorating fact is that if one keeps track over the possible hoppings in the calculation of (3.46) and (3.47) the calculation of (3.48) reduces to the calculation of all combinations of these hoppings with the corresponding weight prefactors.

Due to the increase in computational cost and more severe memory demands due to a higher density of the Hamilton matrix and the loss of applicable translational and point-group symmetries the possible system size in which the Hubbard Hamiltonian can be exactly diagonalized is greatly reduced. For our implementation the ED of a 4×4 system was only possible up to $N_\uparrow = N_\downarrow = 3$ electrons.

3.3.2 Transformation of the ground state

Since we are only interested in the differences of coefficient magnitudes and other properties of the ground state of the Hubbard model in various bases, it is also possible to not transform the whole Hamiltonian, but only the ground state alone.

The ground state of the original Hamiltonian is given by a linear combination of real space basis states

$$|\Psi_0\rangle = \sum_{\Gamma_\uparrow\Gamma_\downarrow} c_{\Gamma_\uparrow\Gamma_\downarrow} |\Gamma_\uparrow\rangle |\Gamma_\downarrow\rangle, \quad (3.49)$$

where the coefficients $c_{\Gamma_\uparrow\Gamma_\downarrow}$ are obtained from the exact diagonalization. The σ -spin part of the real-space Hubbard basis $|I_\sigma^i\rangle$ is given by

$$|I_\sigma^i\rangle = \prod_{\nu}^{N_\sigma} a_{\Gamma_\nu^i, \sigma}^\dagger |0\rangle. \quad (3.50)$$

Γ_ν^i, σ denotes a index list of sites occupied by σ -spin electrons. Transforming the creation operators in (3.50) with the unitary basis transformation obtained from the solution of the Hartree-Fock eigenvalue problem (3.39)

$$a_{i, \sigma}^\dagger = \sum_l U_{il, \sigma} b_{l, \sigma}^\dagger \quad (3.51)$$

yields [44]

$$|I_\sigma^i\rangle = \prod_{\nu}^{N_\sigma} \sum_l U_{\Gamma_\nu^i l, \sigma} b_{l, \sigma}^\dagger |0\rangle = \sum_{l_1 l_2 \dots l_N} \left(\prod_{\nu} U_{\Gamma_\nu^i l_\nu, \sigma} \right) \prod_{\nu} b_{l_\nu, \sigma}^\dagger |0\rangle. \quad (3.52)$$

The index combination (Γ_ν^i, l_ν) denotes occupied states in the original and transformed basis, since the last term in (3.52)

$$|L_\sigma^{l_\nu}\rangle := \prod_{\nu} b_{l_\nu, \sigma}^\dagger |0\rangle \quad (3.53)$$

is a specific basis state in the new HF basis. Eq. (3.52) can be rewritten as a summation over all possible configurations of occupied states and their permutations, and according signs, in the transformed basis as [44]

$$|I_\sigma^i\rangle = \sum_L \underbrace{\sum_{\mathbb{P}} \text{sign}(\mathbb{P}) \prod_{\nu} U_{\Gamma_\nu^i l_{\nu, \mathbb{P}}},}_A |L_\sigma\rangle, \quad (3.54)$$

where the term A is equivalent to the Leibniz determinant formula $\det U_{\Gamma^i L^i, \sigma}$, of the minor of the transformation matrix U_{ij} containing the rows and columns (Γ^i, L^i) determined by the occupied states in the original basis state $|I_\sigma^i\rangle$ and the transformed

state $|L_\sigma^i\rangle$.

Inserting (3.54) into (3.49) yields the representation of the exact ground state in the new basis

$$|\Psi'_0\rangle = \sum_{\Gamma_\uparrow\Gamma_\downarrow} c_{\Gamma_\uparrow\Gamma_\downarrow} \sum_{L_\uparrow L_\downarrow} \det U_{\Gamma_\uparrow L_\uparrow}^\sigma \det U_{\Gamma_\downarrow L_\downarrow}^{\bar{\sigma}} |L_\uparrow\rangle |L_\downarrow\rangle = \sum_{L_\uparrow L_\downarrow} d_{L_\uparrow L_\downarrow} |L_\uparrow\rangle |L_\downarrow\rangle. \quad (3.55)$$

Unfortunately the calculation of the new ground state coefficients $d_{L_\uparrow L_\downarrow}$ through this approach is still of order $\mathcal{O}(N^2)$ with N the Hilbert space of the problem.

3.4 Comparison of different bases

Inspired by [46] following quantities of the exact ground-state wave function of the Hubbard model in real-space, momentum-space and unrestricted Hartree-Fock-space representation were compared to determine the optimal basis:

- **Coefficient magnitudes** $|c_i|$

The simplest quantity to compare are the absolute values of the ground state wave function weights $c_{\Gamma_\uparrow\Gamma_\downarrow}$ and $d_{L_\uparrow L_\downarrow}$ from (3.49) and (3.55).

- **Energy Contribution** ΔE_i

A more elaborate quantity is the energy contribution of basis states to the exact ground state energy defined by [47]

$$\Delta E_i = \frac{(E_{ex} - H_{ii})|c_i|^2}{1 - |c_i|^2}, \quad (3.56)$$

where E_{ex} is the exact ground state energy and H_{ii} the diagonal Hamilton matrix element and c_i the ground state coefficient of the i th basis state. Although more elaborate than bare comparison of ground state coefficients, the diagonal matrix elements H_{ii} in each basis are additionally needed to compute (3.56). This energy contribution ignores cumulative effects when discarding states in a wavefunction $|\Psi'\rangle$ with an energy contribution below a certain threshold for the energy calculation $E' = \langle\Psi'|H|\Psi'\rangle$.

- **Truncated energy ratio** R

To account for the cumulative effects the Hubbard Hamiltonian was diagonalized for different truncated basis sets. States with an energy contribution ΔE

below certain threshold K were discarded and the ratio

$$R = \frac{\langle \Psi_t^K | H^K | \Psi_t^K \rangle}{\langle \Psi_t^K | \Psi_t^K \rangle} / E_{ex}, \quad (3.57)$$

where $|\Psi_t^K\rangle$ denotes the threshold dependent truncated basis, was calculated for different threshold values K . Unfortunately this method requires the full knowledge of the Hamiltonian in each basis representation, which restricts this comparison, as mentioned above, to at most $N_\uparrow = N_\downarrow = 3$ electrons for the 4×4 -lattice due to increased memory requirements to store the Hamiltonian in the UHF basis.

The bases in which the ground state properties were compared, were the real-space \mathcal{R} , in which the Hubbard Hamiltonian was diagonalized, the momentum-space (RHF) \mathcal{K} and UHF basis \mathcal{U} in which the ground state was transformed through the use of eq. (3.55). Since systems off half filling were investigated the self-consistent UHF solution (3.2.2) was used to transform the real-space basis into the UHF basis.

The sorted absolute values of the ground state coefficients of the 4×4 , $N_\uparrow = N_\downarrow = 2$ system, with a Hilbert space size of $\mathcal{H} = 14400$, for the three compared bases for various values of the on-site repulsion U are displayed in fig. (3.9). As can be seen both the UHF, denoted with \mathcal{U} , and the momentum space, or RHF, (\mathcal{K}) ground states have far less coefficients with a significant weight than in the real-space representation \mathcal{R} . For the momentum- and UHF basis only a few of the 14400 states have a coefficient greater than 10^{-2} . For the UHF case more than a third, and in the momentum space the half, of all states have a coefficient magnitude smaller than 10^{-5} . This is more clearly displayed in fig. (3.10), where for $U = 2$ and $U = 8$ the number of coefficients $N(C)$ with a greater weight than C is plotted versus C in a double logarithmic scale. In both UHF and RHF representation there are less than 100 states with a coefficient magnitude greater than 10^{-2} whereas there are over 1000 of those in the real-space representation, where almost every state has a magnitude of at least 10^{-4} . The number of states on the left end of both figures (3.9) are different because the x-axis is cut off at 10^{-10} and the remaining UHF- and RHF-states have an even lower weight. In this analysis of the ground state properties there is little difference between results for different values of parameter U .

In fig. (3.11) the number of states $N(\Delta E)$ with an higher energy contribution than ΔE (3.56) is plotted against ΔE and basically shows the same behavior as the analysis in fig. (3.10).

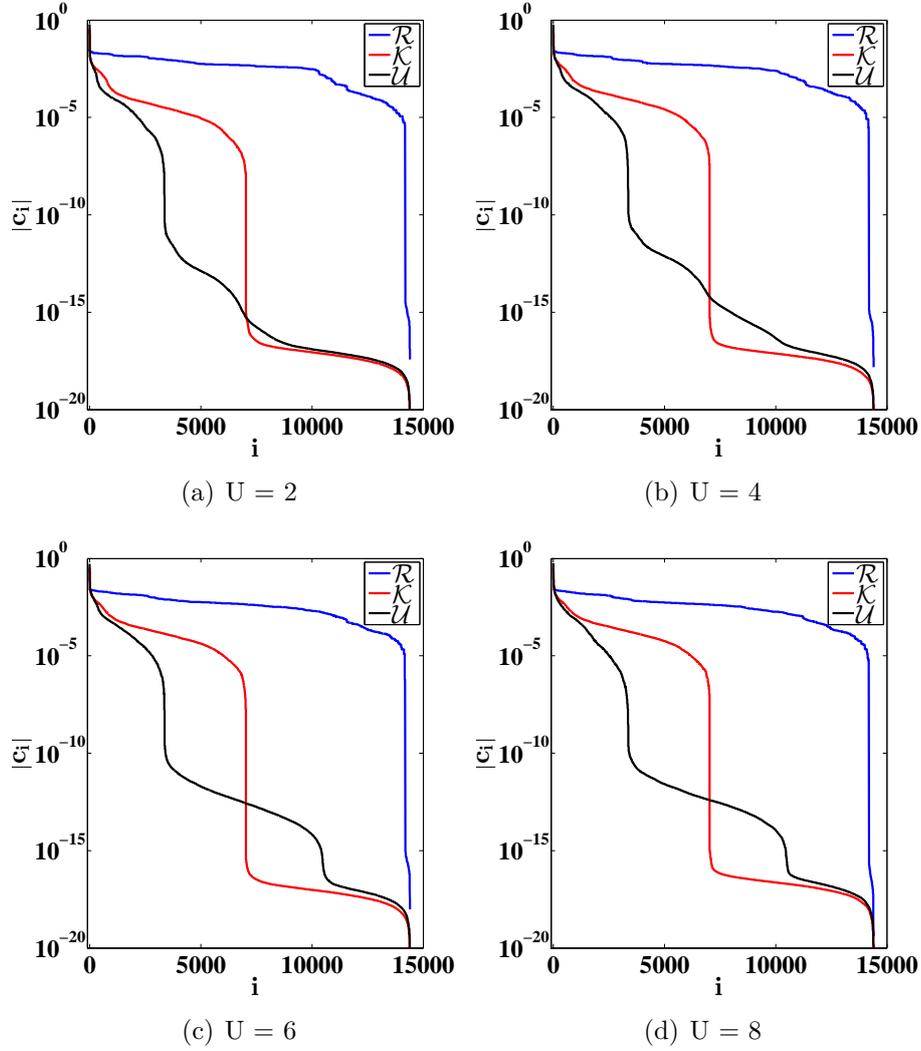


Figure 3.9: Sorted absolute values of the ground state coefficients $|c_i|$ of the 4×4 , $N_\uparrow = N_\downarrow = 2$ system in real- (\mathcal{R}), RHF- (\mathcal{K}) and UHF-space (\mathcal{U}) for on-site repulsion $U = \{2, 4, 6, 8\}$. Both the momentum- and even more the UHF-space ground states have far less states with significant weight compared to the real-space representation. The differences with increasing U are marginal.

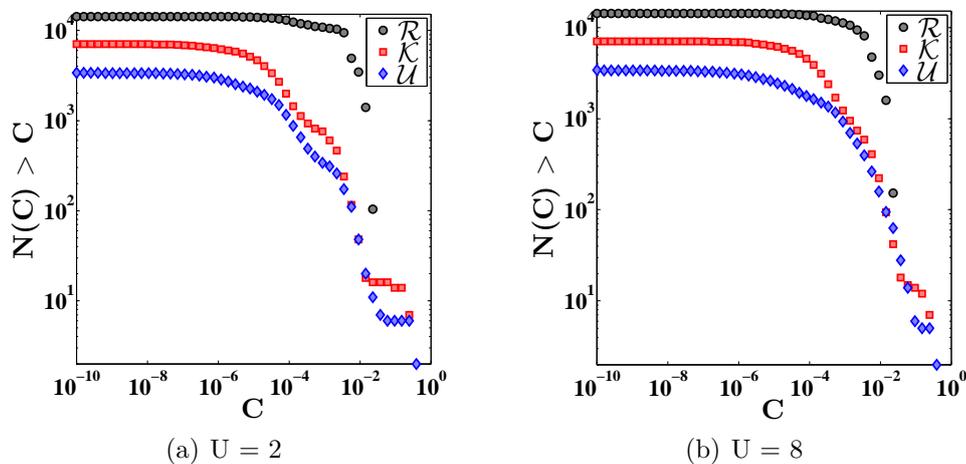


Figure 3.10: The number ground state coefficients $N(C)$ with a greater weight than C versus C . There are under 100 states in the RHF- and UHF-space with a weight greater than 10^{-2} whereas there are over 1000 of those in the real-space representation. The differences between $U = 2$ and $U = 4$ are very little.

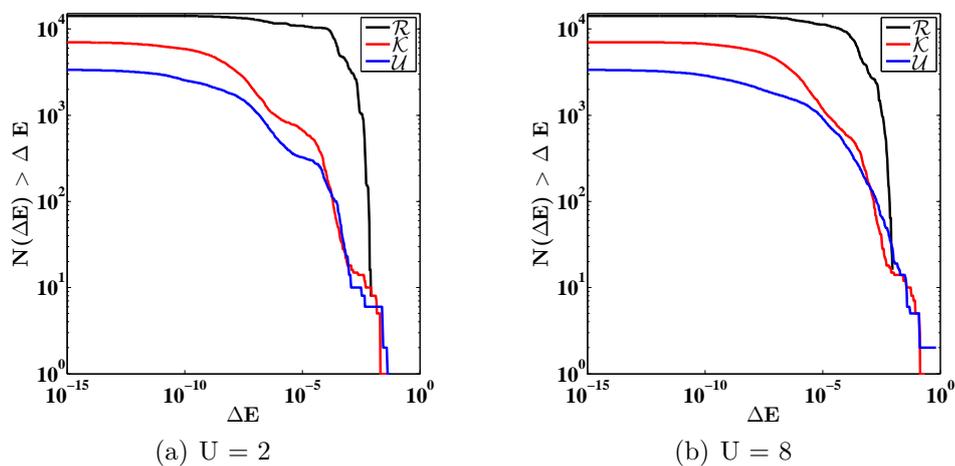


Figure 3.11: Number of ground states $N(\Delta E)$ with an higher energy contribution than ΔE versus ΔE .

In fig. (3.12a) the energy ratio (3.57), which takes into account cumulative effects of discarding basis states as described above, versus the number of regarded states with an energy contribution higher than a certain threshold for the $U = 2, 4 \times 4, N_{\uparrow} = N_{\downarrow} = 2$ system is plotted. Compared to the real-space basis in both the UHF- and RHF-space representation very little states are required to regain almost completely the exact ground state energy. The zoomed in version in the right panel of (3.12) shows that for $U = 2$ in the UHF- and RHF-basis less than 10 states contain over 99% of the exact ground state energy information, whereas in the real-space representation almost all basis states are needed to achieve this accuracy. Additionally opposed to the previous ground state comparisons a stronger influence of the parameter U is noticeable, as seen in fig. (3.13). In tile (a)-(c) again the energy ratio R as function of regarded number of states for values of $U = \{4, 6, 8\}$ is displayed and in tile (d) the number of required states to achieve an energy ratio $R > 0.99$ $N_{\mathcal{R}(R>0.99)}$ as function of U for the UHF- and RHF-basis is displayed. The number of important states definitely increases with U for both the UHF- and RHF-basis. In the $4 \times 4, N_{\uparrow} = N_{\downarrow} = 2$ system the number of needed states is higher in the RHF- than in the UHF-basis, but this is no general conclusion as can be seen compared to the $4 \times 4, N_{\uparrow} = N_{\downarrow} = 3$ system seen below in fig. (3.14). The non-monotonic growth with U is probably due to numerical issues during calculation, and a general monotonic behavior can be assumed. In comparison the number of needed states in real-space representation was at a constant value of about $N_{\mathcal{R}(R>0.99)} = 9700$ for all values of U .

In fig. (3.14) the sorted absolute values of the ground state coefficients $|c_i|$, the truncated energy ratio R and the required number of states to achieve a ratio $R > 0.99$ for the $N_{\uparrow} = N_{\downarrow} = 3, 4 \times 4$ -lattice system for various values of the parameter U and all compared bases are displayed. The Hilbert space size of this system is $\dim \mathcal{H} = 313600$. In the UHF- and RHF-basis for low values of $U = 1 - 4$ below 1% of the Hilbert space size is needed to achieve an accuracy in energy of above 99%. As mentioned, opposed to the $N_{\uparrow} = N_{\downarrow} = 2$ system the UHF basis generally needs more states than the momentum space basis. The non-monotonic behavior again is due to numerical issues.

For certain values of U and filling factors the UHF- and momentum-space transformation provide a drastically improved basis, usable in FCIQMC algorithm, compared the real-space representation. For low values of U the momentum space transformation provides a good basis for the Hubbard model since the kinetic hopping term in the Hamiltonian (3.13) dominates the on-site repulsion. This is also the fact for low fillings ($\hat{=}$ high hole-doping) due to decreased electron-electron interactions, which is also the reason why the momentum-space ground state properties analyzed above remain so beneficial even for high U values in this case. For low U values the

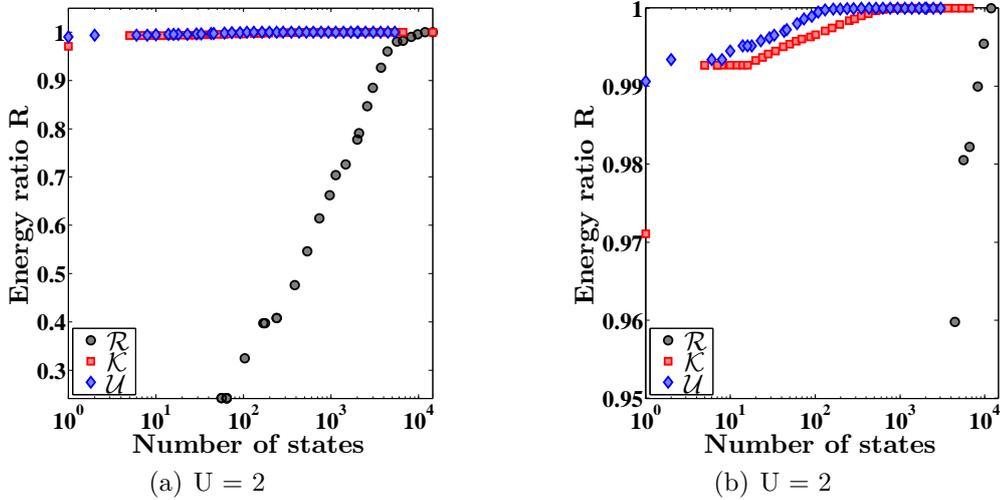


Figure 3.12: Truncated energy ratio R for the $U = 2, N_{\uparrow} = N_{\downarrow} = 2, 4 \times 4$ -lattice system. The right panel (b) is an enhancement of the 0.95 – 1 y-axis region of the left panel (a). As seen in (b) in the UHF and momentum space basis less than 10 states of the 14400-state big Hilbert space are needed to retrieve over 99% of the exact ground state energy.

UHF transformation works so well, since it greatly resembles a momentum space transformation due to the dominating kinetic hopping term in the approximation (3.17). Since the UHF approximation favors anti-ferromagnetic order, which is naturally established by more separated electrons due to high on-site repulsion, it maintains its favorable ground state properties for increasing U . But since the approximation the UHF approach makes, by linearizing the two-particle interaction operator in (3.13), is directly proportional to U it gets inherently worse with an increasing value of it.

Since the majority of the ground state energy information is compressed in only a tiny fraction of the total number of states compared to real-space formulation, both the momentum-space and the unrestricted Hartree-Fock basis seem to be prime candidates to express the FCIQMC algorithm in.

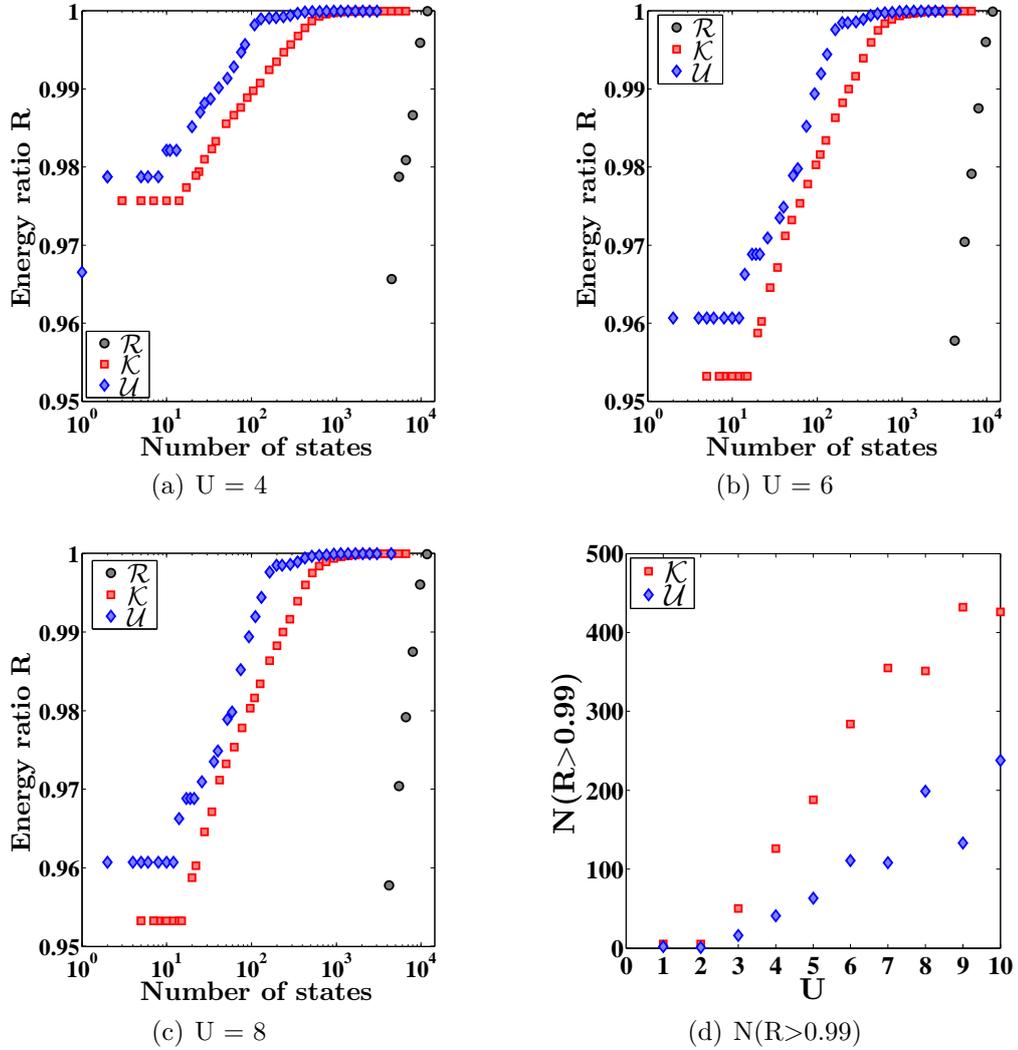


Figure 3.13: Truncated energy ratios R for $U = \{4, 6, 8\}$ of the $N_{\uparrow} = N_{\downarrow} = 2, 4 \times 4$ -lattice system (a-c) and in (d) the number of states needed to consider to achieve a ratio of $R > 0.99$ versus parameter U is plotted. As can be seen, in general for increasing U the number of states needed increases and at a faster pace for the momentum space compared to the UHF basis in the $N_{\uparrow} = N_{\downarrow} = 2, 4 \times 4$ system.

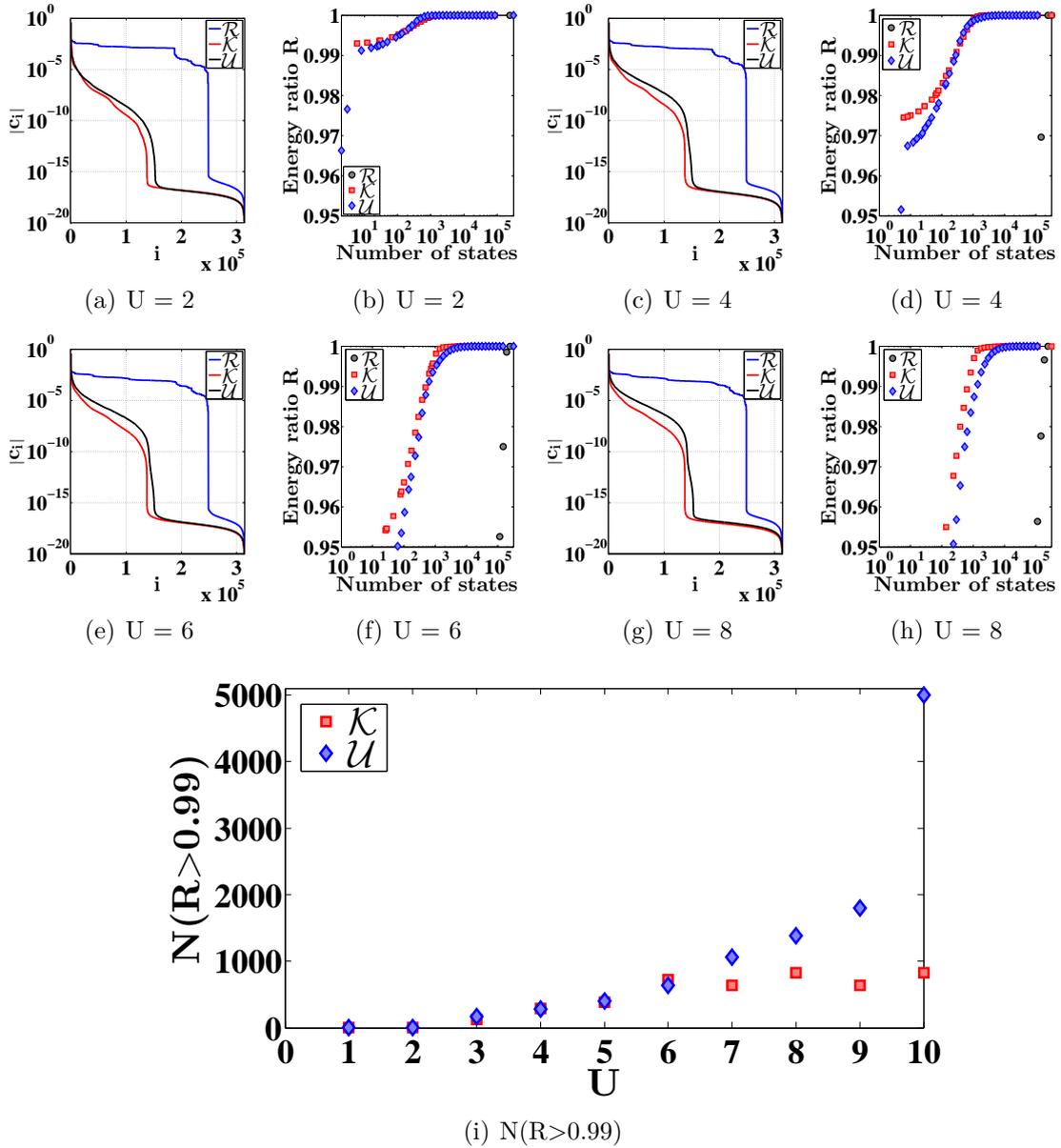


Figure 3.14: Sorted absolute values of ground state coefficients $|c_i|$ and truncated energy ratios R of the 4×4 -lattice with $N_{\uparrow} = N_{\downarrow} = 3$ electrons for parameter $U = \{2, 4, 6, 8\}$ (a-h) and required number of states to achieve a ratio $R > 0.99$ (i). As opposed to the $N_{\uparrow} = N_{\downarrow} = 2$ system, see figures (3.12, 3.13) there are now generally fewer states needed in the RHF basis.

4 Quantum Monte Carlo

Before introducing the FCIQMC algorithm a short review of the basics of Monte Carlo integration techniques, Markov chains and the FCIQMC predecessor diffusion quantum Monte Carlo is given.

The Monte Carlo method allows the evaluation of multidimensional integrals by stochastically sampling the integrand and averaging over the results with only polynomially scaling with integral dimensions. Following [39], consider a three dimensional system with N particles, and denote a possible configuration of the particles as $\mathbf{R} = (\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_N)$, with the probability of this configuration

$$\mathcal{P}(\mathbf{R}) \geq 0 \quad \text{and} \quad \int d\mathbf{R} \mathcal{P}(\mathbf{R}) = 1. \quad (4.1)$$

As described below (4.3), due to the anti-symmetry of fermionic wave-functions, the quantity $\mathcal{P}(\mathbf{R})$ can assume negative values for electrons and thus not be interpreted as probability density. This is a manifestation of the renowned sign problem for fermionic system in QMC.

Nevertheless for bosonic systems the mean value of the function $f(\mathbf{R}_M)$ of M uncorrelated, according to distribution $\mathcal{P}(\mathbf{R})$, randomly distributed configurations \mathbf{R}_i can be interpreted as a random variable

$$Z_f = \frac{1}{M} \sum_{i=1}^M f(\mathbf{R}_i). \quad (4.2)$$

Due to the central limit theorem, Z_f is normally distributed with mean

$$\mu_f = \int d\mathbf{R} f(\mathbf{R}) \mathcal{P}(\mathbf{R}) \quad (4.3)$$

and standard deviation σ_f/\sqrt{M}

$$\sigma_f^2 = \int d\mathbf{R} (f(\mathbf{R}) - \mu_f)^2 \mathcal{P}(\mathbf{R}) \quad (4.4)$$

of the function $f(\mathbf{R})$ for large values of M [48]. As discussed below this estimate of

the statistical error $\Delta\mu_f = \sqrt{\sigma_f^2/M}$ is only valid for truly, practically never achieved, uncorrelated configurations.

Nevertheless this means that for a large enough number of measurements of configurations M , the mean value of these measurements will be a good estimator for the mean of this function, with the standard deviation decreasing as $M^{-1/2}$. This allows an evaluation of integrals

$$I = \int d\mathbf{R}g(\mathbf{R}) \quad (4.5)$$

by rewriting the integrand with $f(\mathbf{R}) = g(\mathbf{R})/\mathcal{P}(\mathbf{R})$ to

$$I = \int d\mathbf{R}f(\mathbf{R})\mathcal{P}(\mathbf{R}). \quad (4.6)$$

With $\mathcal{P}(\mathbf{R})$ obeying the positive-definiteness condition (4.1) the integral can be evaluated approximately by sampling the mean value of a large number of measurements of configurations \mathbf{R} distributed according to $\mathcal{P}(\mathbf{R})$

$$I \approx \frac{1}{M} \sum_i f(\mathbf{R}_i). \quad (4.7)$$

Allowing a calculation of the integral in the exponentially growing phase space with particle number N with a only polynomially scaling algorithm with N .

A efficient method of generating random samples of configurations according to $\mathcal{P}(\mathbf{R})$ is by creating Markov chains with the Metropolis-Hastings algorithm [49]. A Markov chain is a series of random states, where the next state only depends on the current state and not on the history of the series. For an arbitrary starting state \mathbf{R}_i a random trial state \mathbf{R}_t according to a symmetric, ergodic and non periodic probability distribution $\mathcal{Q}(\mathbf{R}_t|\mathbf{R}_i)$ is proposed as next state. In the Metropolis implementation the trial state \mathbf{R}_t is accepted as the next state of the Markov chain with a probability

$$\mathcal{N}(\mathbf{R}_i \rightarrow \mathbf{R}_t) = \min \left(1, \frac{\mathcal{P}(\mathbf{R}_t) \mathcal{Q}(\mathbf{R}_i|\mathbf{R}_t)}{\mathcal{P}(\mathbf{R}_i) \mathcal{Q}(\mathbf{R}_t|\mathbf{R}_i)} \right), \quad (4.8)$$

elsewise \mathbf{R}_i is the next state. Since only the ratio of the probabilities $\mathcal{P}(\mathbf{R}_t)/\mathcal{P}(\mathbf{R}_i)$ is needed the generally hard to calculate normalization of the probability density drops out of the calculation. This process creates a Markov chain with states distributed according to $\mathcal{P}(\mathbf{R})$. But since the next Markov chain element depends on the current one the created configurations are correlated. The influences of these

autocorrelations within the Markov chain manifest themselves in an additional factor in the estimated standard deviation (4.4)

$$\Delta I = \frac{\sqrt{\sigma_I^2(2\tau_A + 1)}}{M}, \quad (4.9)$$

where τ_A is the integrated autocorrelation time [50]. In this work the calculation of the adjusted standard deviation, due to correlations, was done by a blocking analysis [51] discussed in section 5.5.

4.1 Variational QMC

A quantum Monte Carlo algorithm closely related to the above described procedure is variational quantum Monte Carlo (VMC). The starting point of VMC is a given trial wave function $|\Psi(\mathbf{R})_t\rangle$ which should already be a rather good approximation to the true ground state wave function of the system. The energy expectation value

$$E_t = \frac{\langle \Psi_t | H | \Psi_t \rangle}{\langle \Psi_t | \Psi_t \rangle} = \frac{\int d\mathbf{R} \Psi_t^*(\mathbf{R}) H \Psi_t(\mathbf{R})}{\int d\mathbf{R} \Psi_t^*(\mathbf{R}) \Psi_t(\mathbf{R})} \quad (4.10)$$

is a variational upper bound to the true ground state energy E_0 . To evaluate the integral in (4.10) stochastically it is rewritten to [39]

$$E_t = \underbrace{\frac{1}{\int |\Psi_t(\mathbf{R})|^2 d\mathbf{R}}}_{\mathcal{P}(\mathbf{R})} \int |\Psi_t(\mathbf{R})|^2 \overbrace{(\Psi_t^{-1}(\mathbf{R}) H \Psi_t(\mathbf{R}))}^{E_l(\mathbf{R})} d\mathbf{R}, \quad (4.11)$$

where the first term can be interpreted as probability density and $E_l(\mathbf{R})$ is the local energy, which mean sample value is an estimate of the trial state energy E_t

$$E_t \approx \frac{1}{M} \sum_i^M E_l(\mathbf{R}_i). \quad (4.12)$$

4.2 Diffusion QMC

The Diffusion quantum Monte Carlo algorithm (DMC) is a projector method related to the Power method, which determines the true ground state of a system by repeated

application of an operator related to the time-dependent Schrödinger equation

$$i\hbar \frac{\partial |\Psi\rangle}{\partial t} = H|\Psi\rangle \quad (4.13)$$

to an arbitrary starting state. It is closely related to the Green's function Monte Carlo technique (GFMC)[52, 52], as it also uses a functional of the Hamiltonian, repeatedly applied to an arbitrary state, with a non zero overlap with the true ground state, to project out the ground state. In the GFMC technique the projector G applied is the first order difference approximation of the propagator $(1 + \tau(H - \omega))^{-1}$ [53]

$$|\Psi^{(n+1)}\rangle = \underbrace{[1 - \tau(H - \omega)]}_G |\Psi^{(n)}\rangle, \quad (4.14)$$

hence the name Green's function Monte Carlo. As seen below in chapter 5 this choice of approximation is also used in the FCIQMC algorithm. In order for the projection operator G to filter out only the ground state component of $|\Psi\rangle$ the magnitude of the eigenvalue of G corresponding to the ground state $|\Phi_0\rangle$ must be equal to one and all eigenvalues of G associated to excited states must have an absolute value less than one, causing τ to be bound by $\tau \leq 2/(E_{max} - E_0)$, where E_{max} is the largest eigenvalue of H [54]. In this case there is no time-step error if τ is small enough and the spectrum of H is bound from above and below [55], which is asserted for a discrete lattice Hamiltonian. In the FCIQMC algorithm E_{max} can be estimated by the excited determinant containing the energetically highest single particle orbitals [24].

The projector of the DMC method is obtained from the imaginary time Schrödinger equation by performing a so called Wick rotation of time $\tau = it$ in (4.13) leading to

$$\frac{\partial |\Psi\rangle}{\partial \tau} = -H|\Psi\rangle \quad (4.15)$$

in Hartree atomic units. The name of this method stems from the fact that for a general non-relativistic Hamiltonian H eq. (4.15)

$$-\frac{\partial |\Psi(\mathbf{R}, \tau)\rangle}{\partial \tau} = -\frac{1}{2} \nabla^2 |\Psi(\mathbf{R}, \tau)\rangle + V(\mathbf{R}) |\Psi(\mathbf{R}, \tau)\rangle \quad (4.16)$$

can be viewed as a diffusion equation, when the potential term is neglected and $|\Psi\rangle$ is interpreted as diffusive particles, or walkers as they are called in the FCIQMC formalism. By expanding the arbitrary starting wave function $|\Psi\rangle$ in eigenstates of

the Hamiltonian

$$|\Psi\rangle = \sum_i c_i |\Phi_i\rangle, \quad \text{with} \quad H|\Phi_i\rangle = E_i|\Phi_i\rangle, \quad (4.17)$$

a formal solution to (4.15) is given by

$$|\Psi(\tau + \delta\tau)\rangle = e^{-H\delta\tau} |\Psi(\tau)\rangle = \sum_i c_i e^{-E_i\delta\tau} |\Phi_i\rangle. \quad (4.18)$$

For a discrete spectrum with a non-degenerate ground state energy $E_0 < E_1 \leq E_2 \dots$ this time evolution will converge to the ground state in the long time limit

$$\lim_{\tau \rightarrow \infty} |\Psi(\tau)\rangle = c_0 e^{-E_0\tau} |\Phi_0\rangle, \quad (4.19)$$

with all other excited states decaying exponentially. Although the decay of the ground state component $|\Phi_0\rangle$ could be avoided by constant renormalization it is favorable to shift the energy $E_i \rightarrow E_i - E_S$ in the exponential for the long time limit of $|\Psi\rangle$ to remain finite. This causes equation (4.16) to become

$$-\frac{\partial |\Psi(\mathbf{R}, \tau)\rangle}{\partial \tau} = -\frac{1}{2} \nabla^2 |\Psi(\mathbf{R}, \tau)\rangle + (V(\mathbf{R}) - E_S) |\Psi(\mathbf{R}, \tau)\rangle. \quad (4.20)$$

Since (4.19) will only remain finite for $E_S = E_0$ the energy shift E_S has to be varied during the calculation.

As mentioned above, when the potential term in (4.20) is neglected it reduces to a diffusion equation. And if otherwise the kinetic term is neglected it is identical to a rate equation. Hence it is possible to simulate (4.20) by population dynamics of walkers on configurations $|\Psi(\mathbf{R})\rangle$. Simulating the diffusive process by particle movement and the rate term by death and birth of walkers.

When neither potential nor kinetic term are neglected an approximation to the Green's function corresponding to the Hamiltonian (4.20) can be obtained by a Trotter-Suzuki decomposition of $e^{-\tau(T+V)}$, where T denotes the kinetic and V the potential part of (4.20). For small τ an approximate Green's function is given by [56]

$$G(\mathbf{R}, \mathbf{R}', \tau) \approx \underbrace{\frac{1}{(2\pi\tau)^{3N/2}} \exp\left(\frac{(\mathbf{R} - \mathbf{R}')^2}{2\tau}\right)}_A \underbrace{\exp\left(-\tau \frac{V(\mathbf{R}) + V(\mathbf{R}') - 2E_S}{2}\right)}_B. \quad (4.21)$$

Where term A can be interpreted as a diffusion governing part and term B is a

time-dependent reweighting of part A [39], interpretable again as the birth/death of walkers. So even for both kinetic and potential term considered concurrently it is still possible to simulate eq. (4.20) exactly by population dynamics of a large ensemble of walkers, with a small enough timestep τ . As the energy shift E_S appears in the birth/death term B of eq. (4.21) it is possible to control the total population by adjusting it during the calculation.

4.2.1 The fixed-node approximation in DMC

According to [39] the above described implementation of the DMC algorithm assumes a strictly positive wavefunction $|\Psi(\mathbf{R})\rangle$ in the entire phase space, hence effectively making it only applicable to bosonic systems. Since due to the anti-symmetry property of every fermionic wave function it has to have regions in phase space with positive and negative sign. As described in section 4.3 this causes the Monte Carlo scheme to break down for fermionic systems. An approach to this problem within DMC is the fixed-node approximation. In this approximation the nodal surface, this is the multidimensional surface on which the wave function changes sign, of a trial wave function $|\Psi_t\rangle$ is used as a reference for the lowest energy wave function obtained by DMC. In the long-time limit the obtained wave function is then the best possible solution for the given nodal surface. The DMC algorithm effectively calculates the exact solution within each area enclosed by the nodal surface, subject to the boundary condition $|\Psi(\mathbf{R})\rangle = 0$. It is variational in the sense that if $|\Psi_t\rangle$ possesses exactly the true ground state nodal surface, DMC will produce the true ground state.

4.3 The sign problem

The sign problem of a fermionic system, stemming from the anti-symmetry property of the fermionic wave-function, manifests itself in the calculation of any expectation value through (4.6)

$$\langle \mathcal{O} \rangle = \frac{\int d\mathbf{R} \mathcal{O}(\mathbf{R}) \mathcal{P}(\mathbf{R})}{\int d\mathbf{R} \mathcal{P}(\mathbf{R})}. \quad (4.22)$$

For a bosonic system, with symmetric wave functions regarding particle exchange, the quantity $\mathcal{P}(\mathbf{R})$ can always be chosen positive-definite and thus be regarded as probability density, as assumed in the start of chapter 4. Allowing for a polynomially scaling evaluation technique of exponential scaling integrals. But for a fermionic system, due to the negative sign obtained through particle exchange, the quantity

$\mathcal{P}(\mathbf{R})$ can be negative valued and therefor not be regarded as probability density anymore. Thus the general method outlined in the beginning of chapter 4 is not straightforward applicable to fermionic systems.

A way to circumvent negative $\mathcal{P}(\mathbf{R})$ is to sample from the modified 'bosonic' absolute values of the probability density [57]

$$\mathcal{P}'(\mathbf{R}) = \frac{|\mathcal{P}(\mathbf{R})|}{\int |\mathcal{P}(\mathbf{R})|} \quad (4.23)$$

and regard the sign s of $\mathcal{P}(\mathbf{R})$ as part of the measured quantity

$$\langle \mathcal{O} \rangle = \frac{\int \mathcal{O}(\mathbf{R}) \text{sign}[\mathcal{P}(\mathbf{R})] \mathcal{P}'(\mathbf{R})}{\int \text{sign}[\mathcal{P}(\mathbf{R})] \mathcal{P}'(\mathbf{R})} = \frac{\langle \mathcal{O}_s \rangle}{\langle s \rangle}. \quad (4.24)$$

While in principle allowing the Monte Carlo technique to be applied, this approach suffers from the fact that for a set of M measurements the estimate for the average sign (4.3) and standard deviation (4.4) is

$$\langle \bar{s} \rangle = \langle s \rangle \pm \frac{\sigma_s}{\sqrt{M}}. \quad (4.25)$$

If the average of the sign is very small $\langle s \rangle = \varepsilon \ll 1$ its variance σ_s^2 has to be approximately 1, causing the relative error of (4.25) to be

$$\frac{\Delta \bar{s}}{\langle \bar{s} \rangle} \approx \frac{1}{\sqrt{M} \varepsilon}. \quad (4.26)$$

According to [50] for a general quantum mechanical system the mean values of the sign $\langle s \rangle$ is given by the ratio of the fermionic Z and bosonic Z' partition functions

$$\langle s \rangle = \frac{Z}{Z'} = \frac{\int \mathcal{P}(\mathbf{R})}{\int |\mathcal{P}(\mathbf{R})|}. \quad (4.27)$$

Since the partition functions are the exponentials of the corresponding free energies f the average sign is the exponential of the difference Δf ,

$$\langle s \rangle = e^{-\beta N \Delta f}. \quad (4.28)$$

This causes the number of measurements needed to achieve a certain accuracy (4.26) to increase exponentially with system size N and inverse temperature β . This means for systems with a sign problem the favorable polynomial scaling of the Monte Carlo

approach is lost. There is no general solution to the sign problem, and since it has even been shown by Troyer and Wiese to be a NP -hard problem [50], a solution in polynomial computational effort would imply the answer to one of the millenium problems: $NP = P$?

5 Full configuration interaction QMC

The full configuration quantum Monte Carlo (FCIQMC) algorithm, developed recently by G.H. Booth, A. Thom and A. Alavi [24], is an adaptation of the diffusion quantum Monte Carlo (DMC) approach discussed in section 4.2, which avoids the necessity of knowledge of the nodal structure of the fermionic ground state wave function to converge to it. This is achieved by sampling in the complete anti-symmetric Hilbert space of the full configuration interaction Slater determinant expansion (3.1.3) of the system. The starting point is, as in the previously discussed projector-like GFMC and DMC methods, the imaginary-time Schrödinger equation (4.15)

$$\frac{\partial|\Psi\rangle}{\partial\tau} = -(H - E_S)|\Psi\rangle, \quad (5.1)$$

where an energy shift E_S is introduced to ensure convergence to a finite long-time limit value. As in the GFMC method the projector corresponding to (5.1) which filters out the ground state component of any starting state $|\Phi\rangle$ in the long-time limit

$$|\Psi_0\rangle = \lim_{\tau \rightarrow \infty} e^{-\tau(H-E_S)} |\Phi\rangle \quad (5.2)$$

can be approximated in first order finite difference by

$$e^{-\tau(H-E_S)} \approx 1 - \tau(H - E_S). \quad (5.3)$$

The long-time limit application of this projector (5.3) converges to the ground state if $E_S = E_0$, but since E_0 is not known until convergence the shift parameter E_S is adjusted simulation. Additionally, as described later, it is also used for walker population control. As mentioned in section 4.2, this first-order finite difference approximation restricts the time step τ to

$$\tau \leq \frac{2}{E_{max} - E_0} \quad (5.4)$$

to avoid a time-step error.

As opposed to DMC and GFMC the Hilbert space in which FCIQMC is expressed in is the fully anti-symmetric space of Slater determinants formed from a finite single-particle spin orbital basis set, as discussed in chapter 3. A single Slater determinant in second quantization is given by

$$|\Psi_i\rangle = a_{\phi_1}^\dagger a_{\phi_2}^\dagger \dots a_{\phi_N}^\dagger |0\rangle \quad (5.5)$$

where $a_{\phi_i}^\dagger$ is the fermionic creation operator of an electron in orbital ϕ_i , which is part of a $2M$ spin orbital single particle basis set. As mentioned in chapter 3 the basis sets used in this work were the momentum-space and unrestricted Hartree-Fock basis. As discussed in section 3.1.3, for a complete basis set $\{\phi_i\}$ every wave function can be expressed as linear combination

$$|\Phi\rangle = \sum_i c_i |\Psi_i\rangle \quad (5.6)$$

and the lowest energy solution to the eigenvalue problem obtained when the Hamiltonian H is expressed in this basis (5.5)

$$\sum_j c_j \langle \Psi_i | H | \Psi_j \rangle = E_0 c_i \quad (5.7)$$

is the exact ground state solution of the system. The idea of the stochastic FCIQMC algorithm is now to simulate the imaginary-time Schrödinger equation (5.1) in its first order finite-difference approximation (5.3) through the population dynamics of an ensemble of particles, called walkers, inhabiting the Slater determinant space (5.5). By inserting the general state expressed in Slater determinants (5.6) in the imaginary-time Schrödinger equation (5.1) with the first order finite-difference approximation (5.3) one obtains

$$c_i^{(n+1)} = [1 - \tau (H_{ii} - E_S)] c_i^{(n)} - \tau \sum_{j \neq i} H_{ij} c_j^{(n)}. \quad (5.8)$$

This equation governs the dynamics of the walker evolution with time and if iterated until convergence the Monte Carlo averages of the number of walkers inhabiting determinant $|\Psi_i\rangle$ represent the coefficients c_i of the FCI ground state wave function

$$c_i = \frac{\langle N_i \rangle}{\langle N_{tot} \rangle} \quad (5.9)$$

where N_i is the number of walkers on determinant $|\Psi_i\rangle$ and N_{tot} the total number of

walkers. Same as in the DMC method (4.2) a diffusional process, implemented as **spawning step** of new walkers on different determinants, and a **death/cloning step** of walkers mimic the action of (5.8), more thoroughly described below. Additionally, as was shown by Booth et al. [24] and analyzed by Spencer et al. [55], an essential ingredient to enable the wave function to converge to the fermionic ground state is to assign each walker an additional sign property and allow for annihilation of walkers with different sign occupying the same determinant $|\Psi_i\rangle$. The number of signed walkers on each component c_i in each time step gives an 'instantaneous' representation of the wave function $|\Phi\rangle$.

5.1 The algorithm

For FCIQMC calculations in this work the openly available stand-alone FCIQMC code NECI [1] was used, which was provided by developers Prof. Ali Alavi and George Booth. The following description of the specific algorithmic implementation is from [58] and personal communication with the code creators George H. Booth and Ali Alavi.

The main input to the NECI program is a file containing the one- and two- particle overlaps of the single particle orbitals of the basis the Hamiltonian is expressed in. For the Hubbard Hamiltonian in terms of the single particle basis functions ϕ_{ij}^σ they are given by

$$\langle k|t|l\rangle = \sum_{ij} t_{ij} (\phi_{ik}^\sigma)^* \phi_{jl}^{\sigma'} \delta_{\sigma,\sigma'} \quad (5.10)$$

$$\langle jk|lm\rangle = U \sum_i (\phi_{ij}^{\sigma_j})^* (\phi_{ik}^{\sigma_k})^* \phi_{il}^{\sigma_l} \phi_{im}^{\sigma_m} \delta_{\sigma_j,\sigma_l} \delta_{\sigma_k,\sigma_m}, \quad (5.11)$$

with the hopping matrix $t_{ij} = -t \exp(\pm i A_{x/y})$ for nearest-neighbor indices i,j and zero else.

Additionally, if available, also symmetry information of each single particle orbital like associated \mathbf{k} -points used to allow the application of translational symmetry in the FCIQMC calculation are given as input.

A NECI calculation is usually started with a single walker with positive sign residing on the Hartree-Fock (HF) determinant and a high enough value of the energy shift E_S to ensure an exponential walker growth.

Within NECI a main list of occupied determinants specified by the N of $2M$ occupied spin orbitals, where N is the number of electrons and M is the number of single particle spatial orbitals, is stored in bit string integer representation. Similar to the basis state storage described in the exact diagonalization chapter 2. A second

integer list stores the walker number, sign and additional information like symmetry values, and when in use the 'initiator' flag, described below. And for decreasing computational cost the diagonal Hamilton matrix elements H_{ii} of each occupied determinant are also stored. As only information of occupied determinants is stored, at no point of the computation vectors of the size of the Hilbert-space need to be handled.

To simulate eq. (5.8) following set of rules are applied to the set of N_w walkers each time step [58]:

The spawning step

To simulate the diffusional off-diagonal part of (5.8)

$$c_{i,off}^{(n+1)} = -\tau \sum_{j \neq i} H_{ij} c_j^{(n)}, \quad (5.12)$$

for each walker α on every determinant $|\Psi_i\rangle$ a possible connected determinant $|\Psi_j\rangle$ with probability $p_{gen}(j|i_\alpha)$ is randomly chosen. In general a possible connected determinant of the source determinant $|\Psi_i\rangle$, for a Hamiltonian containing maximal a two particle term, as the Hubbard Hamiltonian, can differ at most by two occupied and unoccupied orbitals. For computational reasons the process of finding a possible connected determinant $|\Psi_j\rangle$ and the calculation of the Hamilton matrix element $\langle \Psi_j | H | \Psi_i \rangle$, which can still be zero, is disjoint in the NECI implementation.

Actually a walker would have to try to spawn progeny on all connected determinants with probability $\tau |H_{ij}|$, but since this is computationally unfavorable to implement, each walker only tries to spawn on one connected determinant, modified by the probability of choosing this specific determinant $\tau |H_{ij}| / p_{gen}(j|i)$. The process of choosing a connected determinant is termed '*random excitation generation*' and implemented as follows [58]:

In NECI each determinant is represented bitwise by its occupied and unoccupied single particle orbitals. For the selected source determinant $|\Psi_s\rangle$ two lists containing the numbers of occupied and unoccupied orbitals, with corresponding symmetry information, is created. The symmetry information in this case is associated to the one-particle symmetries of the specific orbitals and can be of spin, point-group and/or translational type. This information is the same for each walker occupying $|\Psi_i\rangle$.

The first step of selecting a specific connected determinant is to decide whether a single p_{single} or a double p_{double} excitation of the occupied single particle orbitals is considered. The values of p_{single}/p_{double} are determined by ratio of possible single and doubly excitations. For ergodicity it is not necessary to exactly account for the

ratio of possible single and double excitation, although being as close as possible is favorable, only the sum of both processes has to be normalized $p_{single} + p_{double} = 1$. Hence specific values for $p_{single/double}$, from the Hartree-Fock determinant excitation possibilities, are set for the whole run.

Single excitation

In the case of a single excitation a specific occupied orbital i from the previously created list is randomly chosen to excite from with a probability $p(i) = 1/N$. For the selected orbital i the number of unoccupied symmetry-compatible orbitals M_{allow} to excite to is calculated. For an excitation to orbital a to be symmetry-compatible

$$\Gamma_i \otimes \Gamma_a \ni A_1, \quad (5.13)$$

where A_1 is the totally symmetric representation, has to be fulfilled [58].

In the case where no excitation is possible, the next walker is regarded. Otherwise a compatible orbital a is chosen and the corresponding bit representation is just given by an exchange of the i and a bit in $|\Psi_s\rangle$ and corresponding fermionic phases due to the $i \leftrightarrow a$ exchange have also to be considered. The overall probability of a specific single excitation is given by

$$p_{gen}(a|i) = p_{single} \times p(i) \times p(a|i) = \frac{p_{single}}{NM_{allow}}. \quad (5.14)$$

Double Excitation

With probability p_{double} two distinct occupied orbitals (i,j) are chosen randomly with probability $p(i,j) = \frac{2}{N(N-1)}$ to excite from. Due to spin conservation of the Hamiltonian the first picked unoccupied orbital a depends on the spins of the chosen occupied orbitals (i,j) . If both are of the same spin, orbital a also has to have this spin, otherwise there is no restriction. For an alike spin σ pair (i,j) there are $M - N_\sigma$ possibilities and for unlike spins $2M - N$. For more effective sampling the number δ_a of unoccupied orbitals a which would not allow any second orbital b to be picked due to spin and symmetry restrictions is calculated. This allows a normalized re-picking of a orbitals, which increases the efficiency of the spawning step. according to [58] the possible symmetry-compatible b orbitals are uniquely defined through the triplet $(i,j;a)$ through the condition

$$\Gamma_a \otimes \Gamma_b = \Gamma_i \otimes \Gamma_j. \quad (5.15)$$

From the allowed number of b orbitals $M_{allow}(b|a)$ one is selected at random and the corresponding probability if the pair (a,b) had been chosen in reversed order is calculated, since generally $M_{allow}(a|b) \neq M_{allow}(b|a)$. The total probability of the

specific double excitation is given by

$$p_{gen}(a,b|i,j) = p_{double} \cdot p(i,j) \cdot (p(a|i,j)p(b|a,i,j) + p(b|i,j)p(a|b,i,j)) \cdot \frac{M_{allow}(a)}{M_{allow}(a) - \delta_d}, \quad (5.16)$$

where the last factor is due to renormalization if multiple a orbitals are drawn until there are compatible b orbitals. (5.16) simplifies to

$$p_{gen}(a,b|i,j) = \frac{2p_{double}}{N(N-1)(M_{allow}(a) - \delta_d)} \left(\frac{1}{M_{allow}(a|b)} + \frac{1}{M_{allow}(b|a)} \right). \quad (5.17)$$

According to eq. (5.8) the probability to spawn a new walker on the chosen connected determinant $|\Psi_j\rangle$ is then given by

$$p_{spawn} = \frac{\tau|H_{ij}|}{p_{gen}(j|i_\alpha)} \quad (5.18)$$

and the sign of the new walker is given by

$$\text{sign}(|\Psi_j\rangle) = -\text{sign}(|\Psi_i\rangle)\text{sign}(H_{ij}). \quad (5.19)$$

If the probability (5.18) exceeds 1, the amount of walkers corresponding to the integer part $\lfloor p_{spawn} \rfloor$ are spawned definitely and an additional walker with probability $p_{spawn} - \lfloor p_{spawn} \rfloor$.

Matrix elements

Within NECI the calculation of the matrix elements of an Hamilton operator with a one- and two-electron part H_{ij} happen on the single particle orbital level through the use of the Condon-Slater rules [38]

$$\langle \Psi | H | \Psi \rangle = \frac{1}{2} \sum_{i < j}^N (\langle ij | \hat{g} | ij \rangle - \langle ij | \hat{g} | ji \rangle) + \sum_i^N \langle i | \hat{h} | i \rangle \quad (5.20)$$

$$\langle \Psi | H | \Psi_a^r \rangle = \sum_i^N (\langle ai | \hat{g} | ri \rangle - \langle ai | \hat{g} | ir \rangle) + \langle a | \hat{h} | r \rangle \quad (5.21)$$

$$\langle \Psi | H | \Psi_{ab}^{rs} \rangle = \langle ab | \hat{g} | rs \rangle - \langle ab | \hat{g} | sr \rangle. \quad (5.22)$$

\hat{g} denotes the one- and \hat{h} the two-electron part of a general non-relativistic Hamiltonian and $|\Psi_{a(b)}^{r(s)}\rangle$, as described in section 3.1.3, is the single(double) excited Slater

determinant compared to $|\Psi\rangle$. As mentioned above these overlap integrals of the single particle orbitals are given as input to the NECI program.

If a spawning attempt was successful in a list, separate from the main walker list, the newly inhabited determinants with the corresponding number of walkers, sign and diagonal Hamilton matrix elements are stored.

The death/cloning step

After every walker on a determinant has tried to spawn progeny on a connected determinant, a second population dynamic step corresponding to the diagonal part of (5.8)

$$c_{i,diag}^{(n+1)} = [1 - \tau (H_{ii} - E_S)] c_i^{(n)} \quad (5.23)$$

is carried out. Each walker on a determinant 'tries to die' with probability

$$p_{death}(i) = \tau (H_{ii} - E_S). \quad (5.24)$$

Walker with a successful death event are immediately removed from the simulation and do not participate in the later described third 'annihilation' step of the simulation. If $p_{death} < 0$ a walker is cloned instead with probability $|p_{death}|$. Cloning events only occur for $E_S > H_{ii}$ so, as later described in section 5.2, the shift parameter E_S can be used to control the total population of walkers. The sign of cloned walkers remain the same.

A major advantage of the up until now described algorithm is that no communication between different walkers need to occur, allowing for easy and effective parallelization of the code.

The annihilation step

But unfortunately this implementation alone would not yet converge to the true FCI ground state. As shown by Booth et al. [24] a third step, termed 'annihilation' step, in the algorithm is needed to correctly simulate the master equation (5.8) and allow it to converge to the true fermionic ground state. As shown by Spencer et al. [55] without annihilation the evolution of walkers with positive c_i^+ and negative c_i^-

sign occurs according to the coupled differential equations

$$\frac{\partial c_i^+}{\partial \tau} = \sum_j (T_{ij}^+ c_j^+ + T_{ij}^- c_j^-) \quad (5.25)$$

$$\frac{\partial c_i^-}{\partial \tau} = \sum_j (T_{ij}^+ c_j^- + T_{ij}^- c_j^+). \quad (5.26)$$

$T = T^+ + T^-$ is the transition matrix from (5.8)

$$T = -(H - E_S), \quad (5.27)$$

where $T^{+/-}$ now only contains the positive/negative off-diagonal elements. This is because the sign of a spawned walker depends on the off-diagonal matrix elements (5.19) and without annihilation adding and subtracting c_i^- from (5.26) the coupled equations can be decoupled to

$$\frac{\partial c_i^+ + c_i^-}{\partial \tau} = \sum_j (T_{ij}^+ + T_{ij}^-) (c_j^+ + c_j^-) \quad (5.28)$$

$$\frac{\partial c_i^+ - c_i^-}{\partial \tau} = \sum_j (T_{ij}^+ - T_{ij}^-) (c_j^+ - c_j^-). \quad (5.29)$$

And since the largest eigenvalue of $T^+ + T^-$ is always greater than the largest eigenvalue of $T^+ - T^-$ the simulation converges to the $c^+ + c^-$ solution in the long-time limit in absence of annihilation [55]. This solution corresponds to the unphysical additive combination of walker of opposite signs inhabiting the same determinant. Although it is not a 'bosonic' solution as in DMC since the anti-symmetric behavior of the solution is ensured by the formulation in the Slater determinant space.

The annihilation step consists of removing walkers of opposite sign residing on the same Slater determinant. For the annihilation step communication between the different walkers is needed, as the list of newly spawned walkers and the original walker list have to be merged and opposite signed walkers on same determinants have to be removed.

There is a system specific number of walkers and time needed to ensure enough occurrences of annihilation processes to allow the true fermionic ground state to emerge during a simulation. This dependence on annihilation processes to converge to the true fermionic ground state is the manifestation of the sign problem in the FCIQMC algorithm.

5.2 Population dynamics

The preliminary described implementation of the FCIQMC algorithm poses typical behavior throughout the simulation [24]. Typically a simulation is started with a single walker with positive sign on the Hartree-Fock determinant. The shift values E_S is kept at a constant value above H_{ii} to ensure exponential walker growth. As mentioned above, if a large enough system dependent walker population is reached, the constructive off-diagonal spawning and cloning events and the destructive death and annihilation events cancel each other. This phase of almost constant walker number is called the annihilation plateau, during which the correct sign structure of the ground state wave function emerges through annihilation processes. The number of walkers in this simulation phase compared to the total Hilbert space size of the system is an indication for the difficulty of the system to be simulated through the FCIQMC method [55]. There have to be enough walkers in a well connected subspace of determinants which contribute significantly to the ground state wave function. This implies the difficulty of a system is basis dependent. The formulation of FCIQMC in the discrete basis of Slater determinants increases the efficiency of the annihilation step significantly compared to first quantized [59] or even earlier attempts in continuous basis sets for DMC and GFMC [60], [61], [62].

After a system dependent duration the simulation exits this annihilation phase, when the sign structure of ground state wave function has been agreed on, and a second exponential growth phase occurs. If the, until now constant, energy shift E_S is not adjusted this second growth is unbound. A constant walker population can be achieved by varying the shift according to [24]

$$E_S(n) = E_S(n - A) - \frac{\zeta}{A\tau} \ln \left(\frac{N_w(n)}{N_w(n - A)} \right), \quad (5.30)$$

where ζ is a damping parameter to ensure more smooth behavior and A is a predefined parameter how often E_S is updated. This dynamic updating causes the shift value E_S to fluctuate around the true ground state energy E_0 causing its mean value to be an estimation for E_0 .

Results and analysis for the Hubbard model with this FCIQMC implementation for different parameters U and both momentum-space and UHF-space single particle bases are found in section 5.6.

5.3 The initiator approximation

Shortly after the publication of the FCIQMC algorithm the 'initiator' approximation i -FCIQMC was proposed [25]. In this variant an additional restriction is imposed on the spawning step of the original implementation. The spawning of walkers onto previously unoccupied determinants is only allowed by a specific set of determinants, called initiators. Determinants become initiators if the number of walkers occupying it exceeds a predefined threshold n_i and also loose this property if the occupation drops below this value. The occupied determinants when the simulation is started, usually only the HF determinant, are set as initiators. Algorithmically this is implemented in the merging of the new and main walker list in the annihilation step, by checking a flag for initiator property, if the determinant of the new walker is not in the main list[58].

The idea of the initiator approach is to avoid the system dependent annihilation plateau phase in which the correct sign structure of the ground state wave function emerges. This is achieved by only allowing spawning events of already sufficiently high occupied determinants, implying that the sign of this determinant already should be correct. Hence the progeny of walkers from these determinants also should possess the correct sign for the final ground state wave function. Additionally an incorrect occupation, of eventually unimportant determinants in the converged ground state, during the simulation is also suppressed with the initiator approximation.

There is in fact no plateau phase in the i -FCIQMC implementation but a slower 'sign coherent' initial exponential walker growth up to a fixed maximum walker number N_{max} . The maximum walker number needed in i -FCIQMC calculations to obtain comparable ground state energies is in most cases substantially smaller as the height of the annihilation plateau in standard FCIQMC calculations [25], [63].

But at the same time the disregard of specific spawning events between connected determinants $|\Psi_i\rangle \rightarrow |\Psi_j\rangle$ is equivalent of applying a truncated Hamiltonian where the overlap matrix element is set to zero $H_{ij} = 0$. Hence it restricts the reachable Hilbert space for low walker number, essentially violating the ergodicity of the algorithm and leads to a non variational 'initiator error'. In the large walker limit the error due to the initiator approximation vanishes as eventually all determinants will become initiators. The only possibility to check if an applied initiator approximation influences the calculated energy is by checking its behavior with increasing maximum walker number. Optimally the ground state energy converges with a much smaller walker number to its correct FCIQMC value.

Results and analysis of the influence of the initiator approximation of the FCIQMC algorithm applied to the Hubbard model are found in section 5.6.

5.4 Energy estimators

As mentioned above a possible estimator for the ground state energy is the long-time limit average value of the applied energy shift $\langle E_s \rangle$. When the simulation is in constant walker mode the shift E_s oscillates according (5.30) around the ground state energy E_0 if the correct ground state wave function emerged during the simulation.

The second method of energy calculation is the projected energy of the instantaneous representation of the evolved wave function $|\Psi(\tau)\rangle$. For the exact converged ground state wave function $|\Phi_0\rangle$ the ground state energy would be given by

$$E_0 = \frac{\langle \Psi_0 | \overbrace{H}^{E|\Phi_0\rangle} | \Phi_0 \rangle}{\langle \Psi_0 | \Phi_0 \rangle} = E_0 \frac{\langle \Psi_0 | \Phi_0 \rangle}{\langle \Psi_0 | \Phi_0 \rangle}, \quad (5.31)$$

where $|\Psi_0\rangle$ is a single determinant component of $|\Phi_0\rangle$ called reference determinant. During the simulation an estimator for the ground state energy is then given by

$$\begin{aligned} \langle E_{proj} \rangle(\tau) &= \frac{\langle \Psi_0(\tau) | H | \Psi(\tau) \rangle}{\langle \Psi_0(\tau) | \Psi(\tau) \rangle} \\ &= \sum_i \frac{c_i(\tau)}{c_0(\tau)} \langle \Psi_0(\tau) | H | \Psi_i(\tau) \rangle \\ &= \sum_i \frac{\langle N_w^i(\tau) \rangle}{\langle N_w^0(\tau) \rangle} \langle \Psi_0(\tau) | H | \Psi_i(\tau) \rangle, \end{aligned} \quad (5.32)$$

where the instantaneous wave function is expanded in its occupied Slater determinants $|\Psi_i\rangle$ and $|\Psi_0\rangle$ is a specific highly occupied reference determinant, usually the HF determinant, but adjustable on the fly during the calculation if another determinant becomes significantly higher occupied. A high walker number is desirable since the ratio of coefficients c_i/c_0 is calculated by the number of walkers N_w^i/N_w^0 residing on the corresponding determinants. And thus the higher the number of walkers on the reference determinant, the better the statistics of the measured quantity. To avoid statistical cancellation the ensemble average of the numerator and denominator of (5.32) are calculated separately. The matrix elements $\langle \Psi_0 | H | \Psi_i \rangle$ are calculated according to the Condon-Slater rules (5.22) so only singly and doubly excited determinants $|\Psi_i\rangle$ from the reference determinant have to be considered. Since any operator which commutes with H possesses the same eigenvector spectrum, expectation values of these operator are also calculable with equation (5.32).

Whereas ground state expectation values of other operators

$$\langle O \rangle = \frac{\langle \Phi_0 | O | \Phi_0 \rangle}{\langle \Phi_0 | \Phi_0 \rangle} = \frac{\sum_{ij} \langle \Psi_i | O | \Psi_j \rangle c_i c_j}{\sum_i c_i^2} \quad (5.33)$$

not commuting with H are not possible to calculate within this framework, since the off diagonal elements O_{ij} corresponding to all determinants ever occupied would be needed, and not only the current wavefunction representation.

Due to the fact that the energy from the shift $\langle E_s \rangle$ depends on the whole walker population and the projected energy estimator $\langle E_{proj} \rangle$ only on the number of walker of the reference determinant and single and double excitations thereof, these two values provide two uncorrelated energy estimates [24].

5.5 Error analysis

As in every Monte Carlo method a careful analysis of the statistical error of the correlated data set is mandatory. The used NECI program [1] automatically applies a **blocking error analysis** of the energy shift $\langle E_s \rangle$ and projected energy $\langle E_{proj} \rangle$ estimators [51].

Opposed to the description in chapter 4 the subsequent configurations during a Monte Carlo simulation are never completely uncorrelated and this has an effect on the statistical properties of the obtained estimators. Following [51] the estimator of the variance $\sigma^2(m)$ of the average value of n measurements x_i is given by

$$\sigma^2(m) = \langle m^2 \rangle - \langle m \rangle^2, \quad m = \frac{1}{n} \sum_{i=1}^n x_i. \quad (5.34)$$

Inserting for m yields

$$\sigma^2(m) = \frac{1}{n^2} \sum_{i,j} \gamma_{i,j}, \quad (5.35)$$

with

$$\gamma_{i,j} = \langle x_i x_j \rangle - \langle x_i \rangle \langle x_j \rangle \quad (5.36)$$

being the correlation function of the data set $\{x_i\}$. The correlation function is invariant under time translation and hence only dependent on the distance of the

two data points $\gamma_{i,j} = \gamma_{|i-j|}$. This causes eq. (5.35) with $t = |i - j|$ to become

$$\sigma^2(m) = \frac{1}{n}\gamma_0 + \frac{2}{n^2} \sum_{i>j}^n \gamma_t. \quad (5.37)$$

And since there are $(n - 1)$ terms with $t = 1$, $(n - 2)$ terms with $t = 2$ and so on, the variance estimator is given by

$$\sigma^2(m) = \frac{1}{n^2} \sum_{i,j=1}^n \gamma_{|i-j|} = \frac{1}{n} \left[\gamma_0 + 2 \sum_{t=1}^{n-1} \left(1 - \frac{t}{n}\right) \gamma_t \right], \quad (5.38)$$

The direct evaluation of this quantity is computationally too costly and has to be calculated approximately. The blocking method for the calculation of the variance estimator works as follows [51]:

If the data set $\{x_i\}$ for which σ^2 has to be calculated is transformed to the half as big data set $\{x'_i\}$ by averaging over two subsequent values

$$x'_i = \frac{x_{2i-1} + x_{2i}}{2}, \quad n' = \frac{n}{2} \quad (5.39)$$

the average value $m' = m$ of the new data set remains unchanged. Also, since $\langle x + x' \rangle = \langle x \rangle + \langle x' \rangle$, the correlation function (5.36) of the transformed data set (5.39) is given by

$$\gamma'_t = \begin{cases} \frac{\gamma_0 + \gamma_1}{2}, & \text{for } t = 0. \\ \frac{\gamma_{2t-1} + 2\gamma_{2t} + \gamma_{2t+1}}{4}, & \text{for } t > 0, \end{cases} \quad (5.40)$$

and hence the variance of the new data set

$$\begin{aligned} \sigma^2(m') &= \frac{1}{n'^2} \sum_{i,j}^{n'} \gamma'_{i,j} \\ &= \frac{2}{n} \left[\frac{\gamma_0 + \gamma_1}{2} + \frac{2}{4} \sum_{t=1}^{n/2-1} \left(1 - \frac{2t}{n}\right) (\gamma_{2t-1} + 2\gamma_{2t} + \gamma_{2t+1}) \right] \\ (\text{with } t' = 2t) &= \sigma^2(m) \end{aligned} \quad (5.41)$$

remains unchanged. Since $\gamma_t \geq 0$ equation (5.38) indicates

$$\sigma^2(m) \geq \gamma_0/n \quad (5.42)$$

and unless $\gamma_1 = 0$ γ_0/n increases everytime the data set is blocked due to eq. (5.40) and $n' = n/2$. In the limit of repeated blocking transformations the variance estimate is exactly given by γ_0/n [51]. To evaluate γ_0 at every blocking step it is approximated by the naive variance

$$c_t = \frac{1}{n-t} \sum_{k=1}^{n-t} (x_k - \bar{x})(x_{k+1} - \bar{x}), \quad (5.43)$$

which is a biased estimator for the correlation function γ_t . Biased because it has an expectation value of [51]

$$\langle c_t \rangle = \gamma_t - \sigma^2(m) + \Delta_t, \quad \text{with } \Delta_0 = 0, \quad (5.44)$$

what can be used to eliminate γ_0 in (5.42) to obtain

$$\sigma^2(m) \geq \frac{\langle c_0 \rangle}{n-1}. \quad (5.45)$$

To compute $\sigma^2(m)$ for each blocked data set $\{x_i\}$ the naive variance c_0 with (5.43) is computed and used as an estimate for $\langle c_0 \rangle / (n-1)$ and this is repeated until the number of elements of the final available data set $n' = 2$ is reached. The values of $c_0/(n-1)$ converge to its fixed point value with increased blocking of the data set, which is the estimate for the variance of the data set $\sigma^2(m)$.

5.6 FCIQMC run properties for the 4×4 -square and 18-site tilted lattice

Standard-FCIQMC

This section deals with the results of the standard FCIQMC implementation *without the initiator approximation* applied to the 4×4 square- and 18-site tilted square-lattice at half-filling for different values of U with the UHF and momentum-space basis, also called RHF basis throughout the rest of this chapter. The time-step of the calculation was set to $\tau = 0.0001$ for all calculations throughout this chapter. The maximum allowed number of walkers was adjusted individually for each system to be larger than the system specific annihilation plateau height. The Hilbert-space size of the 4×4 system at half-filling is $\dim \mathcal{H} = 165.6 \cdot 10^6$.

Figure (5.1) shows the walker population dynamics (a-b) and total energy obtained with the projected energy estimator (c-d) for the RHF (left) and UHF (right) basis for $U = 2$. As described in section 5.2 the walker population without initiator approx-

imation shows an initial fast exponential growth until a certain walker threshold, the so called initiator plateau, is reached. During this phase constructive creation and destructive annihilation and death process balance each other and the ground state wave function emerges. After this phase a second exponential growth occurs which can be regulated by adjusting the global energy shift parameter during simulation to keep the walker number at a given maximum value. The intermediate walker growth at the annihilation plateau for the UHF basis in fig. (5.1b) was caused by a stop and restart of the FCIQMC simulation and has no special meaning.

The walker population and duration in the annihilation plateau indicate the difficulty to simulate a given system with the FCIQMC algorithm and are given in table (5.1) for FCIQMC runs in this section.

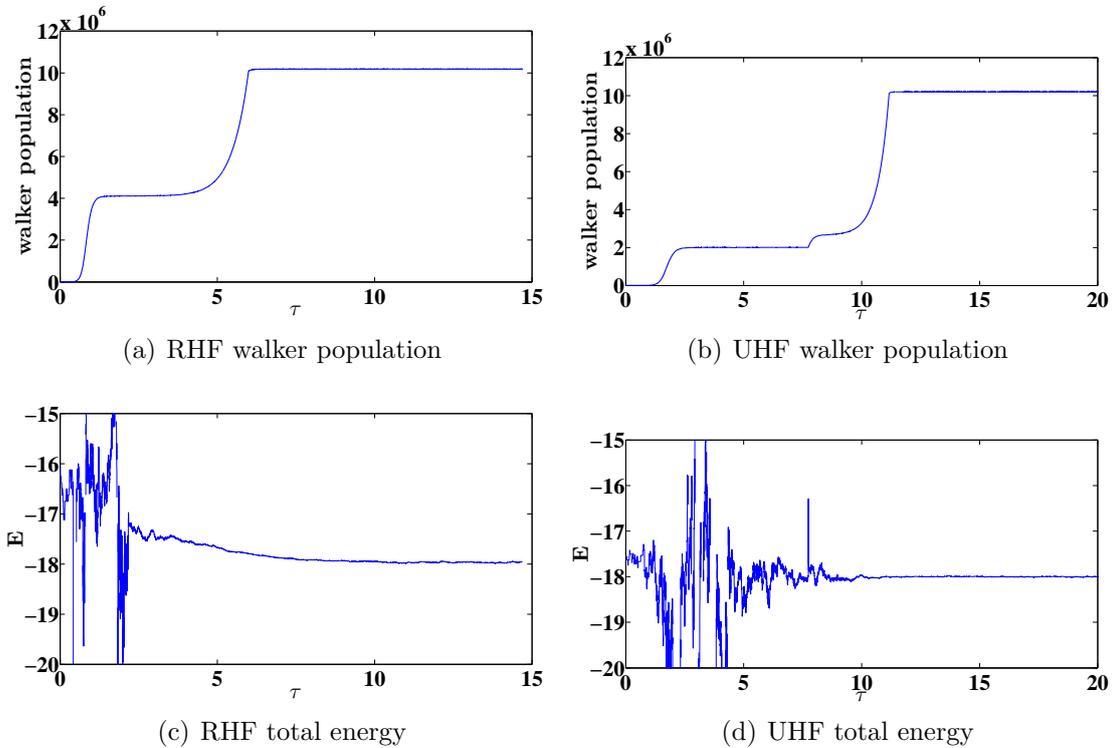


Figure 5.1: Walker population dynamics (a-b) and total projected energy estimator (c-d) for the $U = 2,4 \times 4$ square lattice for the RHF (left) and UHF (right) basis. The weird looking growth phase at the annihilation plateau for the UHF basis (b) happened due to a stop and restart of the FCIQMC simulation.

The standard-FCIQMC results for $U = 4$ are displayed in fig. (5.3), where additionally the weight of the HF determinant is shown in the last row (e-f).

The jump in the total energy estimator for the RHF basis obtained through the projection to a reference determinant can be explained by the fact that the reference determinant can be dynamically adjusted during the simulation to be the most populated one. This happened for the RHF run as can be seen in tile (e) of fig. (5.3). The corresponding reference energy is the triple excited determinant with occupied orbitals: (1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 15, 16, 17, 18, 22). The detailed results can again be found in table (5.1). Walker population and energy results for the $U = 6$ case are displayed in figure (5.2) and can be found in table (5.1).

As seen in figure (5.4) for the investigated range of parameter U there seems to be a linear increase of the annihilation plateau with U in the RHF basis. In the UHF basis there seems to be no linear behavior, but to conclude a specific functional behavior more data points would be needed. In the 4×4 -square lattice significantly less walkers are needed in the UHF basis as in the RHF basis to reach the annihilation phase.

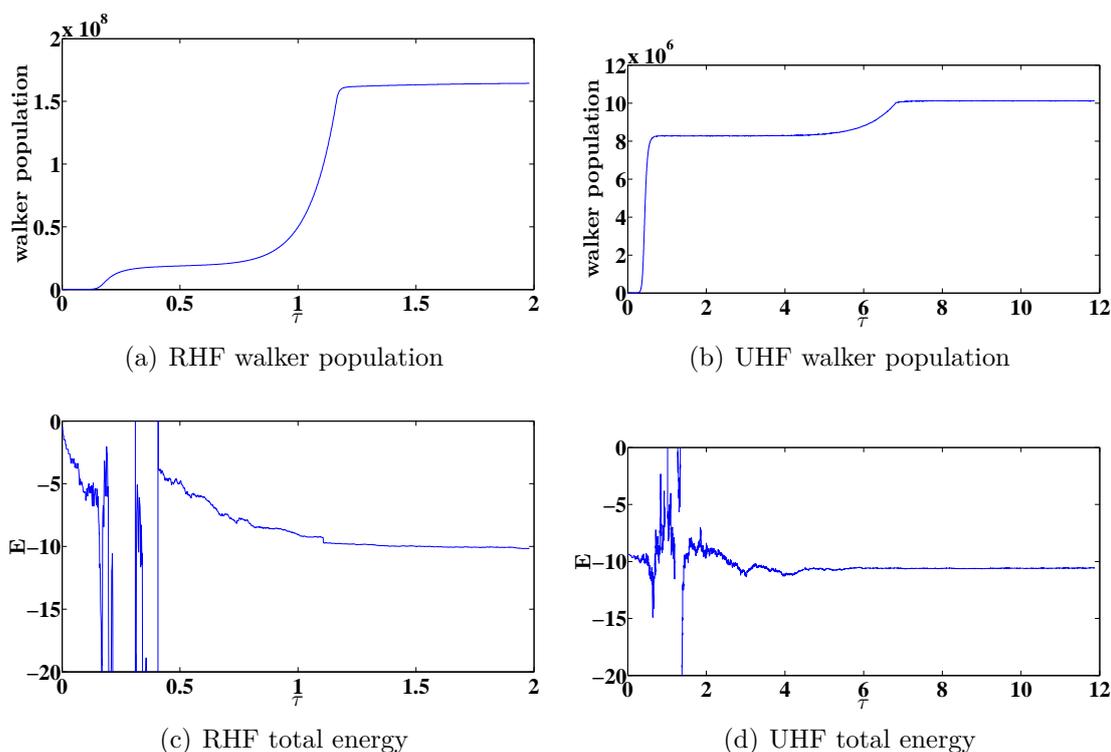


Figure 5.2: Walker population dynamics (a-b) and total projected energy estimator (c-d) for the $U = 6$, 4×4 square lattice for the RHF (left) and UHF (right) basis.

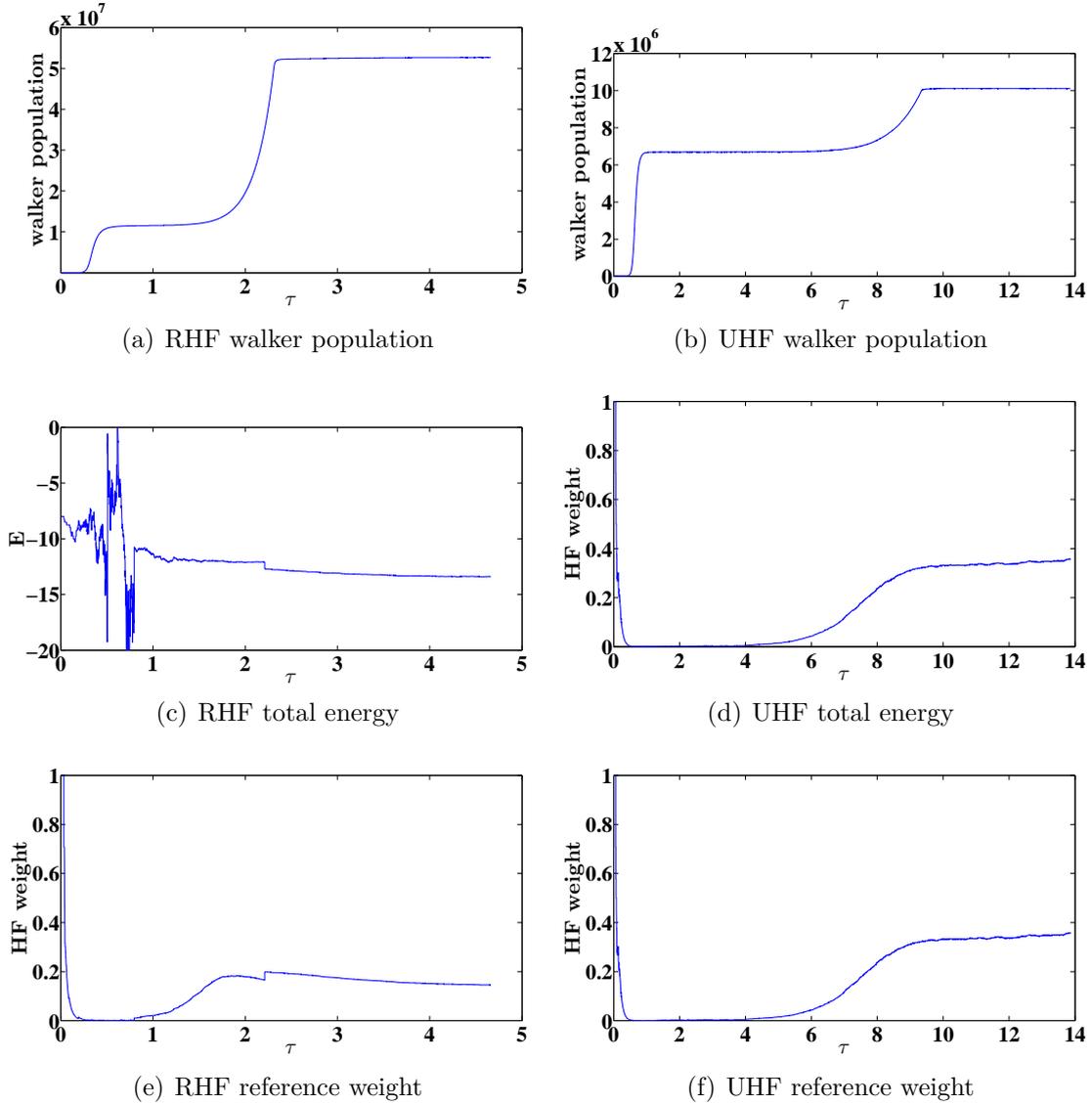


Figure 5.3: Walker population dynamics (a-b), total projected energy estimator (c-d) and weight of the highest occupied determinant (e-f) for the $U = 4$, 4×4 square lattice for the RHF (left) and UHF (right) basis.

The second investigated lattice was the previously described 18-site tilted square lattice with an Hilbert space size of $2.4 \cdot 10^9$ at half filling. For $U = 4$ the height of annihilation plateau in the RHF basis is $129.2 \cdot 10^6$ walkers. In the UHF basis even a maximum number of 500 million walkers was not enough to reach the annihilation plateau and since this already accounts for 21.2% of the Hilbert space size no further

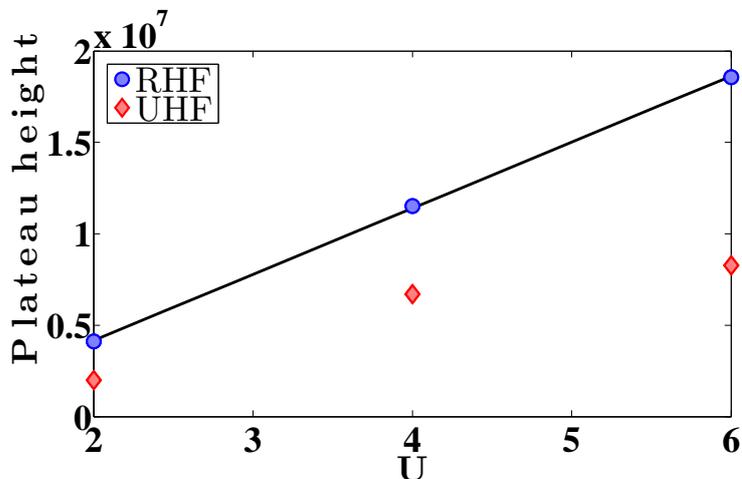


Figure 5.4: Number of needed walkers to reach the annihilation phase in the standard FCIQMC implementation for the 4×4 -square lattice as function of U for the RHF and UHF basis. In the investigated parameter U range the walker population in the RHF basis shows a linear behavior as function of U .

run attempt with a higher walker population was done.

Spencer et al. [55] also investigated the application of the FCIQMC algorithm to the Hubbard model on the 18-site tilted square lattice in momentum-space representation. Similar to our results for the 4×4 -square lattice they observed a linear increase of the walker population at the annihilation plateau with parameter U . And our plateau height of $129.2 \cdot 10^6$ walkers for the $U = 4$ case is in good agreement with their results.

The results of the standard-FCIQMC runs for the square- and tilted square lattice for the RHF- and UHF-basis for different values of U are summed up in table (5.1). The simulation on the tilted lattice in the RHF basis was still in the annihilation phase when this report was written. But as mentioned, an extensive investigation of this system was done by Spencer et al. in [55].

For the 4×4 -lattice the height of the annihilation plateau, and hence the difficulty to simulate the system with the FCIQMC algorithm, rises with increasing U values. As seen in fig. (5.4) linearly in the RHF basis. Additionally the number of walkers needed to reach the annihilation phase is higher and the convergence of the energy is slower when using the RHF basis compared to the UHF basis.

Table 5.1: Standard-FCIQMC run results for the 4×4 -square lattice for $U = \{2,4,6\}$ and the 18-site tilted square-lattice for $U = 4$ for the RHF and UHF basis. The reference exact energy results for the 4×4 -lattice were obtained through exact diagonalization.

Lattice	U	basis	walker pop.	fraction dim \mathcal{H}	E_{proj}	ΔE	weight
4×4	2	RHF	$4.11 \cdot 10^6$	2.42%	-17.96	0.10	0.17
	2	UHF	$2.00 \cdot 10^6$	1.21%	-18.00	0.02	0.50
	4	RHF	$11.51 \cdot 10^6$	7.95%	-13.42	0.20	0.15
	4	UHF	$6.70 \cdot 10^6$	4.05%	-13.61	0.01	0.35
	6	RHF	$18.57 \cdot 10^6$	11.21%	-10.15	0.40	0.12
	6	UHF	$8.27 \cdot 10^6$	5.99%	-10.60	0.05	0.27
18-sites	4	RHF	$129.2 \cdot 10^6$	5.57%	N/A		
	4	UHF	$> 500 \cdot 10^6$	$> 21\%$	N/A		

Initiator-FCIQMC

The effects of the initiator approximation for the square- and tilted square-lattice for different values of U are studied in this section. The threshold number of walkers on a determinant to become an initiator determinant was set to 2. As in the previous section the simulation time-step was $\tau = 0.0001$. The influence of the initiator approximation was studied as function of maximum allowed number of walkers, which was increased in steps: $100 \cdot 10^3 \rightarrow 500 \cdot 10^3 \rightarrow 1 \cdot 10^6 \rightarrow 2 \cdot 10^6 \rightarrow 4 \cdot 10^6 \rightarrow 8 \cdot 10^6$.

4×4 , $U = 2$: In fig. (5.5) the controlled walker population (a-b) and averaged energy shift during the different phases (c-d) for the RHF- (left) and UHF-basis (right) for the 4×4 -square lattice and $U = 2$ is displayed. As mentioned in section 5.3 the population dynamics with applied initiator approximation do not possess a two-stage exponential growth but only one sign coherent, albeit slower, exponential growth. The spikes in the energy shift in (c-d) are due to the increase of the maximum walker number during simulation. For the RHF-basis the energy shift for $U = 2$ already converges to its final value at a walker population of $500 \cdot 10^3$. The results for the same settings for the UHF basis are displayed on the right of fig. (5.5). Also for the UHF basis the averaged energy shift does not change anymore above $500 \cdot 10^3$ walkers.

The total energy of the i -FCIQMC simulation for the RHF basis E_i^{RHF} has the same slow convergence to the exact energy $E_{exact} = -18.02$ as in the standard FCIQMC run above. It is also given by $E_i^{RHF} = -17.96$ for this duration of simulation. The total energy in the UHF basis for this i -FCIQMC simulation is given by $E_i^{UHF} = -18.02$ which is in good agreement with the exact energy.

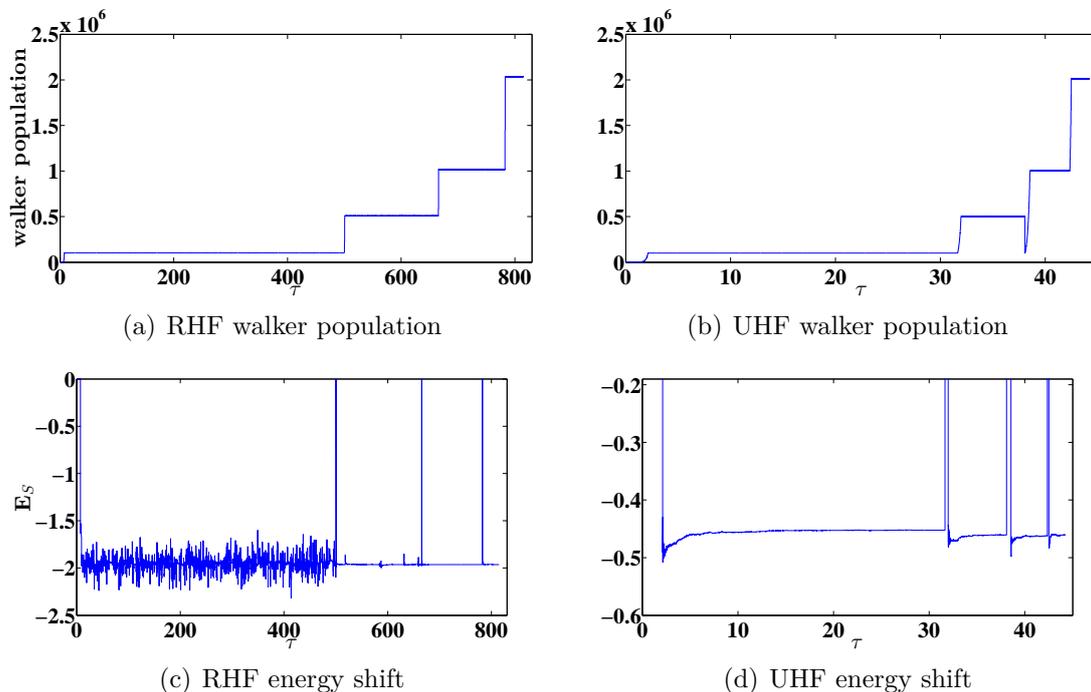


Figure 5.5: Controlled maximum number of walkers (a-b) and energy shift (c-d) for the $U = 2$, 4×4 -square lattice i -FCIQMC calculation for the RHF- (left) and UHF-basis (right). The energy shift for the RHF and UHF basis converges to its final value already at $500 \cdot 10^3$ walkers

4×4 , $U = 4$: The results of the i -FCIQMC simulation for $U = 4$ are displayed in fig. (5.6). The needed walker population for a converged energy shift in the RHF basis is $2 \cdot 10^6$ and for the UHF basis $4 \cdot 10^6$. As in the $U = 2$ case, (5.5), the convergence of the total energy in the RHF basis, like in the standard FCIQMC implementation, is very slow. It is given by $E_i^{RHF} = -13.46$ against the exact energy of $E_{exact} = -13.62$. And as in the $U = 2$ case and in the standard FCIQMC calculation the total energy with UHF basis converges much faster to the exact ground state with a value of $E_i^{UHF} = -13.61$ versus the exact value of $E_{exact} = -13.62$.

4×4 , $U = 6$: For $U = 6$ the RHF energy shift does not change anymore above 4 million walkers, whereas the UHF shift is still not converged at 8 million walkers. Nevertheless the total energy of the UHF run is, with a value of $E_i^{UHF} = -10.54$ compared to $E_i^{RHF} = -10.31$, closer to the exact ground state energy of $E_{exact} = -10.55$.

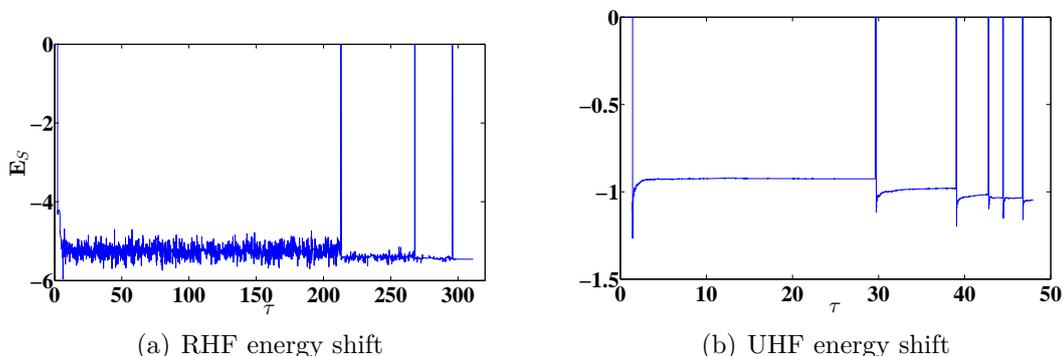


Figure 5.6: Energy shift for the $U = 4$, 4×4 -square lattice i -FCIQMC calculations for the RHF- (left) and UHF-basis (right). The maximum walker number increases same as in the $U = 2$ case in fig. (5.5) from $100 \cdot 10^3 \rightarrow 500 \cdot 10^3 \rightarrow 1 \cdot 10^6 \rightarrow 2 \cdot 10^6$.

18-site tilted: As seen in figure (5.7a) already 2 million walkers are sufficient for the $U = 4$, 18-site tilted square lattice in the RHF basis for the average shift value to not change anymore. This accounts for only 0.08% of the total Hilbert space size of $2.36 \cdot 10^9$ and only 1.6% of the annihilation plateau height of $129.2 \cdot 10^6$ walkers of the system. For the UHF basis 8 million walkers were not enough for convergence of the average shift value. But as 500 million walkers were not enough to reach the annihilation plateau phase with the UHF basis in the standard FCIQMC implementation in this system, a high needed maximum walker number in the initiator implementation is to be expected.

Different to the 4×4 -square lattice system the convergence properties of the total energy in the RHF basis are more favorable in the tilted lattice. For this comparably short run the total energy for the RHF basis is already given by $E_i^{RHF} = -17.23$ compared to the exact value of $E_{exact} = -17.25$ [45]. Whereas the total energy estimated from the energy shift for the UHF basis is given by $E_i^{UHF} = -17.07$.

In table (5.2) the results for the initiator error study for the square- and tilted square-lattice are gathered. It can be seen that, although the number of needed states to overcome the initiator error in the 4×4 -square lattice is lower for the RHF basis, the convergence of the total energy is better in the UHF basis. And as seen for the $U = 6$ UHF case even if the initiator error is still there its magnitude can still be very low.

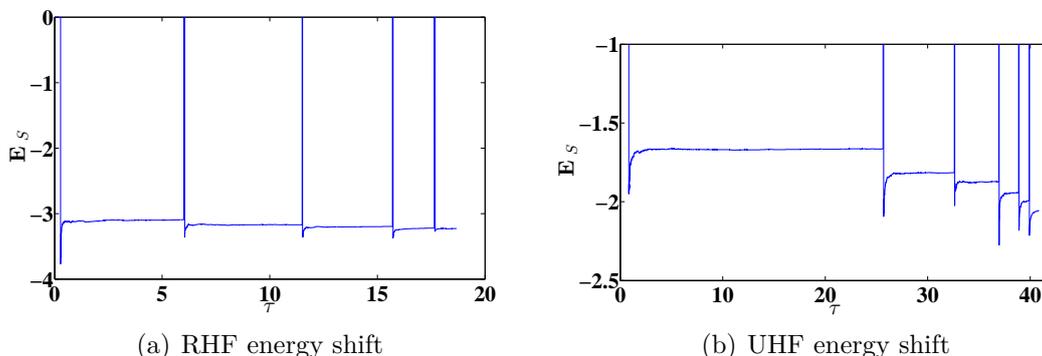


Figure 5.7: Energy shift for the $U = 4$, 18-site tilted lattice i -FCIQMC calculations for the RHF- (left) and UHF-basis (right). The maximum number of walkers increases now from $100 \cdot 10^3 \rightarrow 500 \cdot 10^3 \rightarrow 1 \cdot 10^6 \rightarrow 2 \cdot 10^6$ up to $4 \cdot 10^6$ in the RHF case (a) and up to $8 \cdot 10^6$ in the UHF case (b).

Table 5.2: Fraction of Hilbert space needed to overcome the error due to the initiator approximation. Additionally the fraction of the needed walkers in i -FCIQMC and the annihilation plateau height in the standard FCIQMC implementation and the energy difference to the correct ground state energy due to not yet converged simulations are given.

lattice	U	basis	fraction dim \mathcal{H}	fraction annihilation plateau	ΔE
4×4 square	2	RHF	0.30%	12.2%	0.06
		UHF	0.30%	25.0%	0.00
	4	RHF	1.21%	17.4%	0.20
		UHF	2.42%	59.7%	0.01
	6	RHF	2.42%	21.5%	0.24
		UHF	N/A	N/A	0.01
18-site tilted	4	RHF	0.08%	1.6%	0.02
		UHF	N/A		

Single orbital properties

This section is devoted to draw conclusions between the differences of the above FCIQMC simulations and the corresponding single orbital properties of the basis used. There are fundamental differences between the RHF and UHF basis on the single orbital level in the square- and tilted square-lattice. The single orbital energies of the RHF and UHF basis in both lattice types are displayed in fig. (5.8). As can be seen the RHF basis in the 4×4 -square lattice has an open-shell configuration, no

energy gap between the highest occupied and lowest unoccupied orbital, in the half-filled case and this degeneracy causes an ambiguity of the Hartree-Fock determinant. All other cases possess a closed-shell configuration in the half-filled case. As can be seen in all FCIQMC calculations with an open-shell HF configuration the run qualities are inherently poorer. There are many more equally important excited determinants, since already the HF determinant is degenerate. So in systems where the RHF basis possesses an open-shell configuration the UHF basis shows better simulation properties.

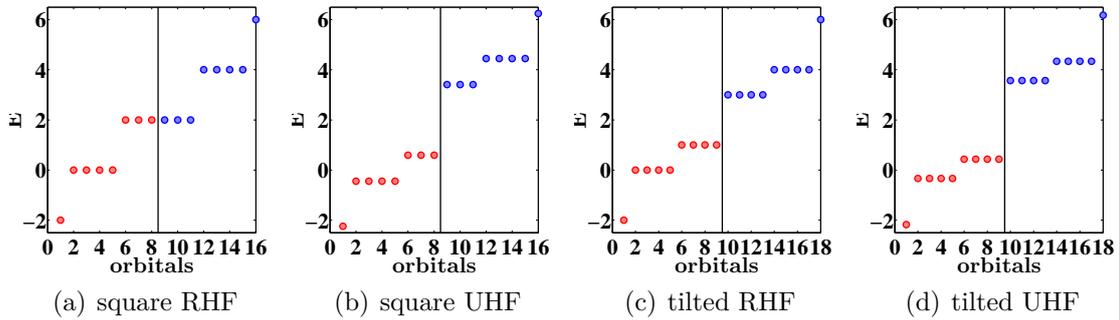


Figure 5.8: Single orbital energies of the RHF and UHF basis in the square- and tilted square-lattice.

In fig. (5.9a) the energy gap between the highest occupied and lowest unoccupied orbital E_G of the UHF basis in the half filled case divided by the bandwidth E_W of the whole single orbital spectrum as function of U is displayed. For low values of U the energy gap tends to zero whereas in the high U limit the ratio converges to 1.

In fig. (5.9b) the mean value of the absolute values of overlap of opposite-spin orbitals of the UHF basis, the so called spin-contamination as discussed in section 3.1.2, is displayed. It is an indication of the additional symmetry breaking properties of the UHF- compared to the RHF-basis. For U exactly zero the UHF solution is equivalent to the RHF solution and hence the opposite-spin wavefunctions are identical. Which for increasing values of U become more distinct, breaking the inherent spin-symmetry of the Hubbard Hamiltonian, which explains that the FCIQMC run properties in the UHF basis deteriorate with increasing U .

As mentioned above for the tilted square lattice there exists a finite value U_C below which the RHF solution is the energetically more favorable one. And hence the energy gap/bandwidth ratio and spin-contamination possess a constant value even for the UHF basis below U_C , as seen in fig. (5.10c-d).

So although the existence of an energy gap between the highest occupied and lowest unoccupied orbital seems to be favorable for FCIQMC calculations, increasing

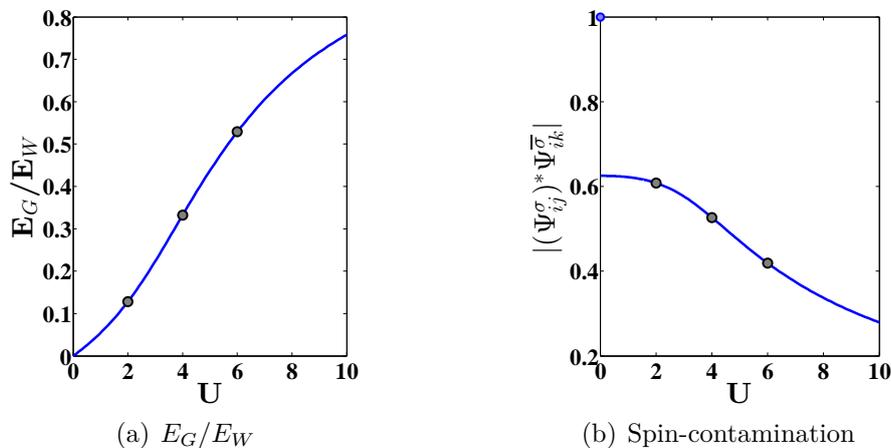


Figure 5.9: (a) Energy gap between highest occupied and lowest unoccupied orbital divided by the bandwidth of the energy spectrum as function of U for the UHF basis in the 4×4 -square lattice. (b) Mean value of the absolute values of the orbital overlap, the so called spin-contamination, of the UHF basis states.

values with U thereof do not automatically imply better simulation qualities. As at the same time other orbital properties, as the spin-contamination, deteriorate.

Another defining difference between the two bases is the quality of the approximations and parameter U range where they coincide. As seen in section 3.2.3 in the 4×4 -lattice the UHF solution is energetically more favorable for all infinitesimal $U > 0$, whereas in the 18-site tilted lattice, below a critical value $U_C = 2.87$, the RHF solution is more favorable, as can be seen in fig. (5.10a). Tile (b) of fig. (5.10) shows that in the thermodynamic limit $\lim_{L \rightarrow \infty} U_C$ possesses a finite value. In table (5.3) the energy differences for the RHF and UHF approximation for different U for each lattice are found. As mentioned in the tilted case there is no difference for $U = 2$ since it is below U_C . Even compared to UHF solution the RHF approximation gets very inaccurate already at intermediate values of U , implying worse FCIQMC simulation properties.

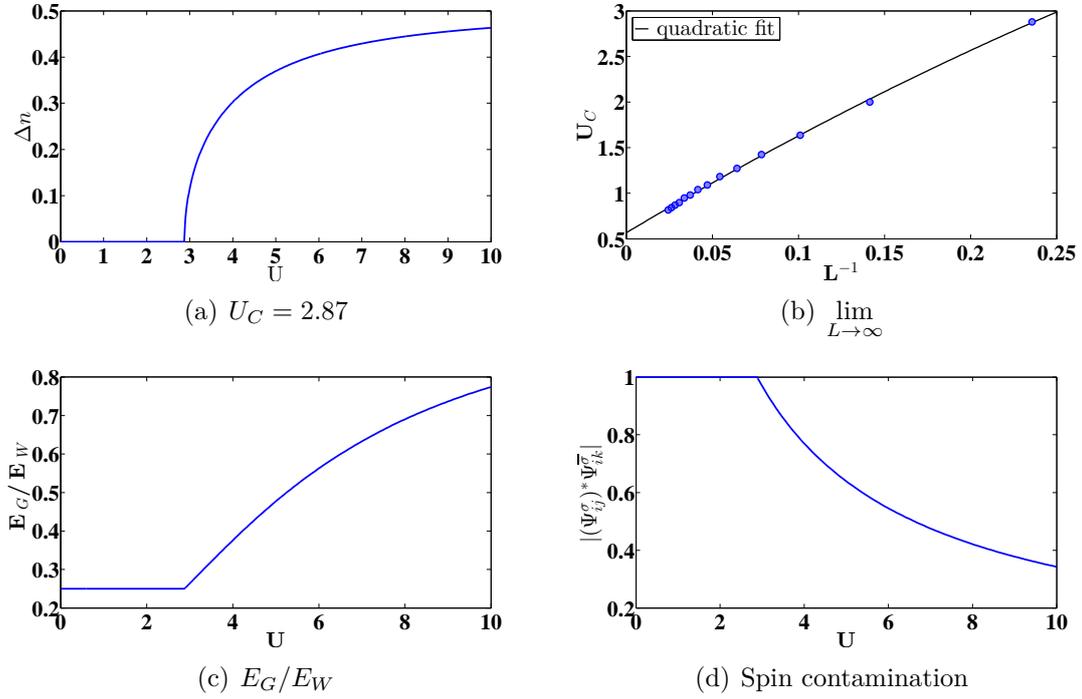


Figure 5.10: Single orbital properties of the UHF basis for the 18-site tilted square lattice. (a) Energy minimizing spin density difference as function of U . For $U < U_C = 2.87$ the RHF solution is energetically more favorable. (b) Thermodynamic limit behavior of the critical U_C with an residual value $U_C^\infty = 0.56651$. The non-perfect fit behavior is due to numerical inaccuracies. (c) Ratio of energy gap and bandwidth as function of U . (d) Mean spin contamination value as function of U .

Table 5.3: Energies of the RHF and UHF approximation for different U on the square and tilted square lattice.

lattice	U	E_{UHF}	E_{RHF}	ΔE_{HF}	E_{exact}
4×4 -square	2	-17.556	-16.0	1.5562	-18.02
	4	-12.567	-8.0	4.5666	-13.46
	6	-9.3799	0.0	9.3799	-10.55
18-site tilted	4	-15.0279	-14	1.0279	-17.25
	6	-10.8453	-5	5.8453	N/A

Non-optimal UHF-basis

In this section the question is discussed whether there exists a basis with better FCIQMC run properties between the RHF and UHF solution. The non-optimal UHF basis is defined by, not finding the energy minimizing spin-density difference Δn_{UHF} , but by enforcing a certain difference between \uparrow - and \downarrow -spin densities Δn for each lattice site.

In fig. (5.11) and (5.12a-d) the *i*-FCIQMC results for the 18-site tilted square lattice with $U = 4$ and a maximum number of one million walkers, corresponding to only 0.04% of the Hilbert space size, for different Δn values are displayed. On the left side the total energy calculated with the projected energy estimator $\langle E \rangle_{proj}$ and on the right side the proportion of total walkers on the current reference determinant to calculate the projected energy is shown. When not mentioned explicitly the reference determinant is always the Hartree-Fock determinant, hence the axis caption 'HF weight'. A high occupation of the reference determinant ensures better statistical properties of the projected energy estimator since it reduces noise of the measurement. And if the highest occupied determinant happens to be the HF determinant the weight of it is also an indication of how good the basic HF approximation already describes the interacting system.

In tile (e) of fig. (5.12) the converged energies of the different Δn runs are displayed and it can be seen, as in the initiator error section above, that the limited number of walkers is too little for the UHF-basis, and even more the $\Delta n = 0.5$ variant, to yield good results compared to the exact energy of $E_{ex} = -17.25$ [45]. But for bases with $\Delta n \leq 0.2$ and especially the RHF basis already a walker population corresponding to only 0.04% of the Hilbert space size is enough yield satisfying energy values in this case.

For the 18-site tilted lattice, tile (g-h) of fig. (5.12) show the energy gap/bandwidth ratio and spin-contamination of the UHF-like basis as function of spin-density difference.

The corresponding energy and HF weight results, only for the RHF- and UHF-basis, for different values of U are displayed in fig. (5.13a-b). In all cases either the RHF- or UHF-basis yielded the best results and no non-optimal UHF-like basis in between. Although the energy results for high U values, e.g. $U = 10$, are similar in both bases used, as seen in tile (a), the FCIQMC simulation properties are significantly worse with RHF-basis as seen in tiles (c-d).

As the initiator error study above showed that for the UHF basis a rather high number of walkers is needed to overcome the initiator error, it is not surprising that the converged total energy of the $U = 10$ UHF run with only 1 million maximum walkers $E_{UHF} = -7.87$ is not yet in good agreement with the exact ground state energy of $E_{exact} = -8.07$ [45].

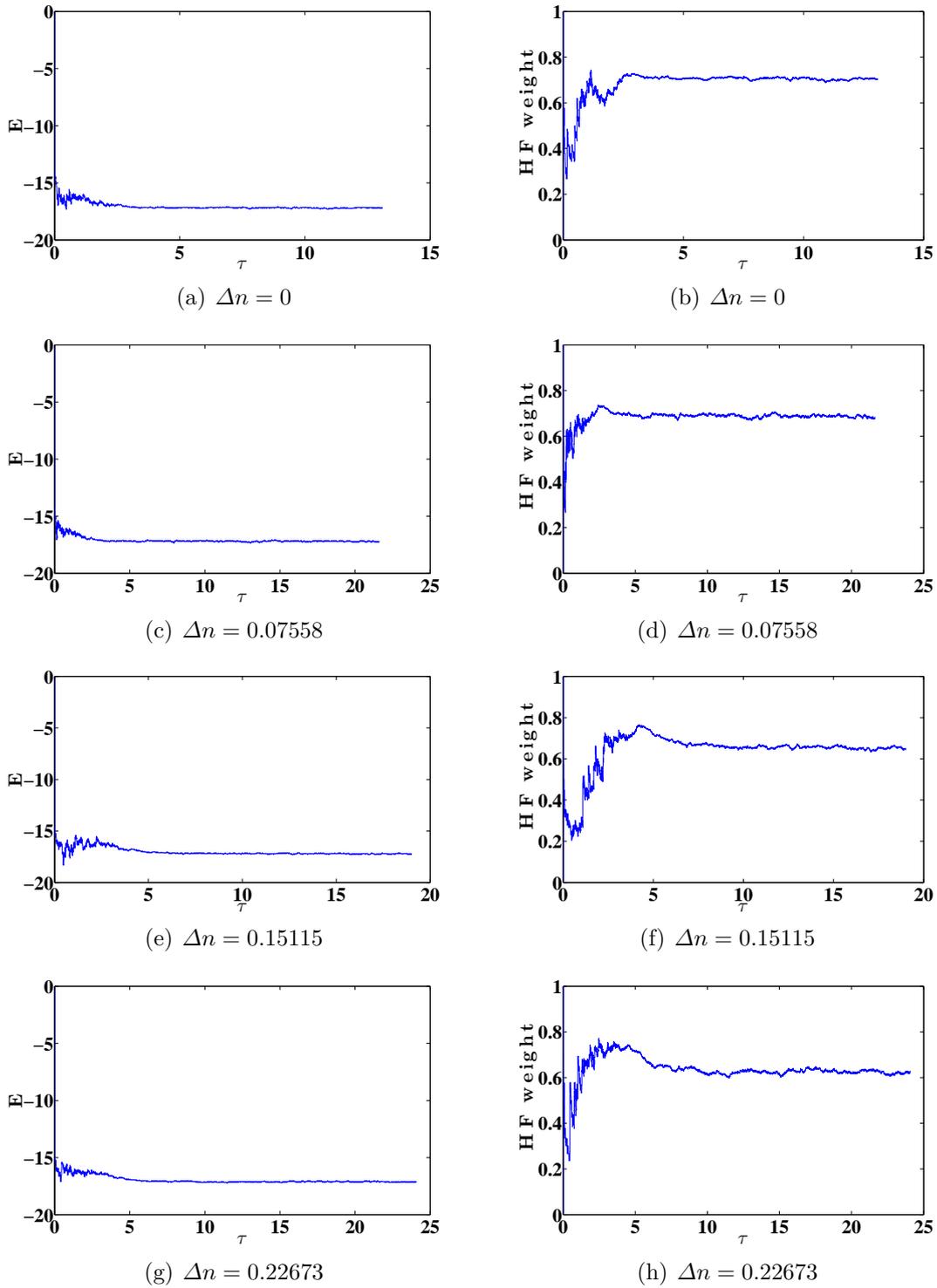


Figure 5.11: $U = 4$, 18-site tilted square lattice i -FCIQMC total energy and reference determinant weight (called HF weight generically) results for the RHF- and non-optimal UHF-basis.

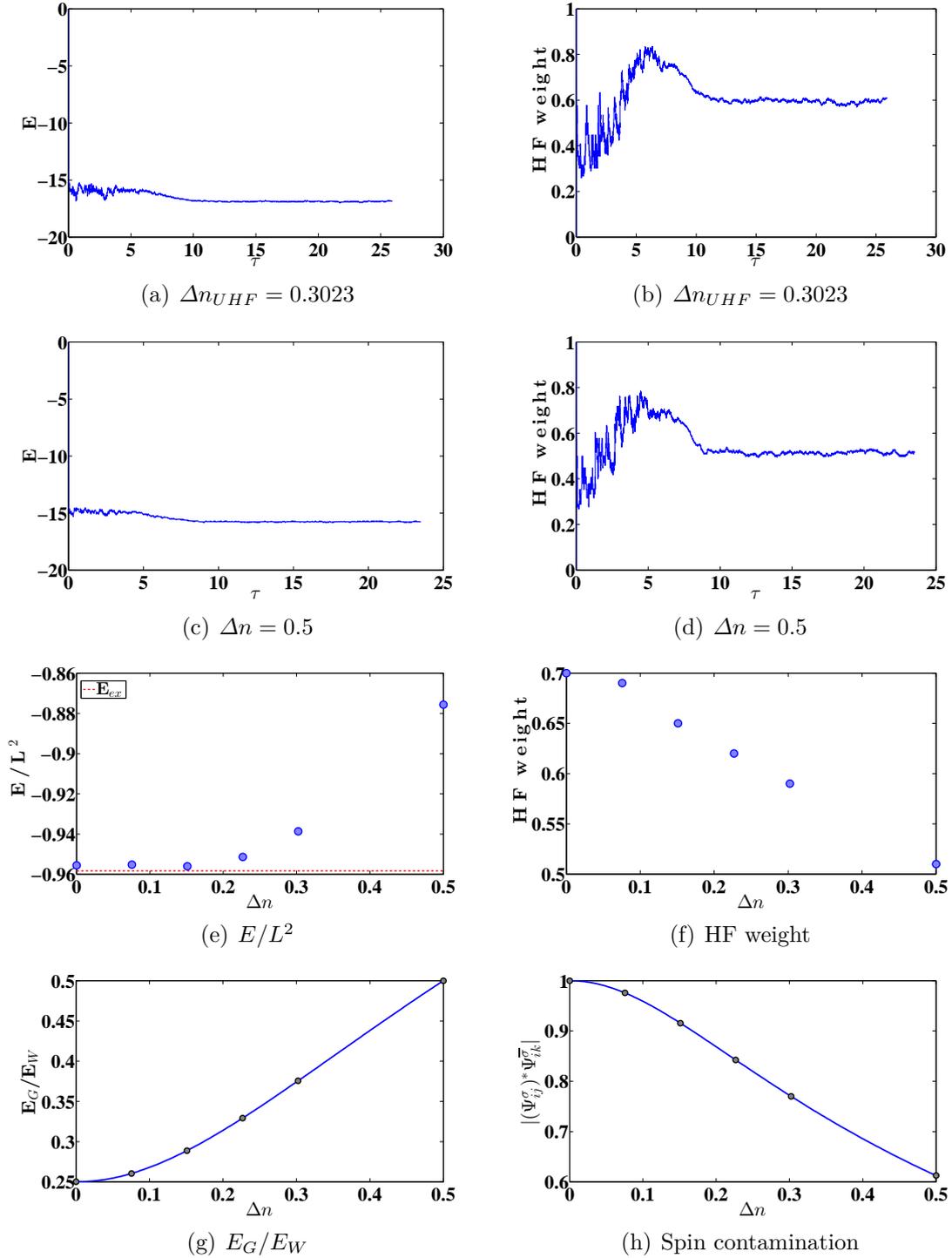


Figure 5.12: $U = 4$, 18-site tilted square lattice i -FCIQMC total energy and reference determinant weight results for the UHF-basis (a-b) and non optimal $\Delta n = 0.5$ basis (c-d). In tile (e) the energy per site of the FCIQMC runs is plotted as a function of the density difference. The red line corresponds to the exact energy solution of $E_{ex} = -17.25$ [45]. (f) shows the weight of the reference determinant and (g) the energy gap and (h) the spin contamination of the UHF-like basis as function of density difference Δn .

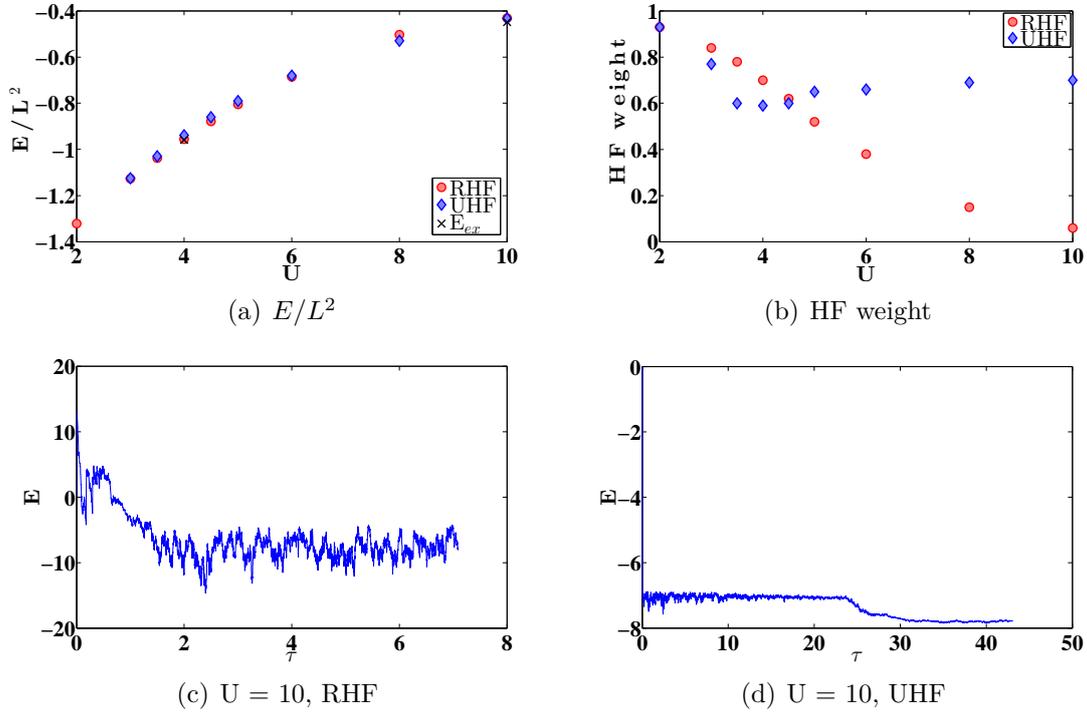


Figure 5.13: (a) Energy per site for the UHF and RHF basis on the 18-site tilted square lattice as function of U . Exact energy result are given for $U = 4$ and $U = 10$ [45]. For values of $U < 8$ the RHF-basis yields better results but as expected for larger values of U this behavior switches. The inspected non-optimal UHF bases always showed worse behavior than the corresponding optimal basis. (b) shows the weight of the reference determinant for RHF and UHF as function of U . The UHF weight starts off similar to the RHF basis, but increases again after a slight decrease for intermediate U values. In (c-d) the vastly different simulation properties of the total energy for $U = 10$ in the RHF and UHF basis are shown. Compared to the very noisy RHF run the UHF run, although the simulation time is very long, shows good convergence even for the small used maximum walker number of 10^6 .

Influence of twisted boundary conditions

Although the reasoning behind introducing twisted boundary conditions (TBC) is only described in the next chapter 6 the effects of them on FCIQMC runs are analyzed in this section.

The most obvious effect of twisted boundary conditions is the lifting of degeneracies within the single orbital energies due to shift of the associated \mathbf{k} -vectors. In fig. (5.14) the single orbital energies for both the square and tilted square lattice and RHF and UHF bases for an arbitrary twist value are displayed. With TBC also the

RHF basis in the square lattice for half-filling possesses a closed-shell structure, as every other system.

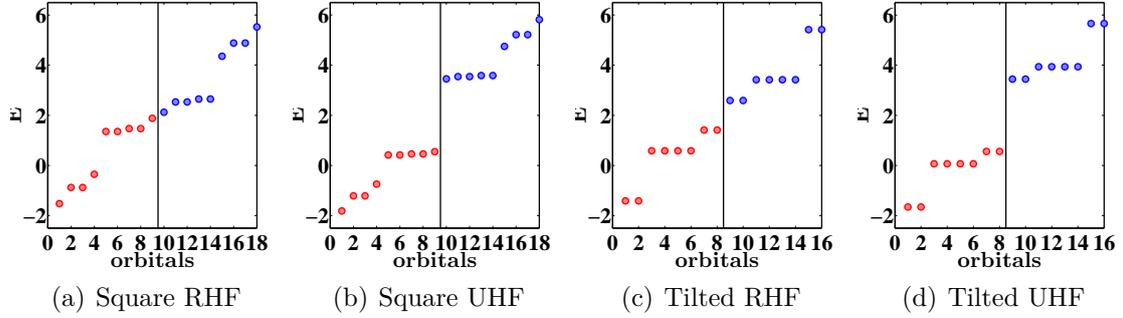


Figure 5.14: Lifted degeneracies of single orbital energies due to applied twisted boundary conditions.

In figure (5.15) the energy and reference determinant weight results of *i*-FCIQMC runs with a maximum walker number of 10^6 for different values of U and twist values along the x-axis for the RHF and UHF basis in the 18-site tilted square lattice are shown. As can be seen the RHF results get worse with increasing twist values Φ_x compared to the UHF basis.

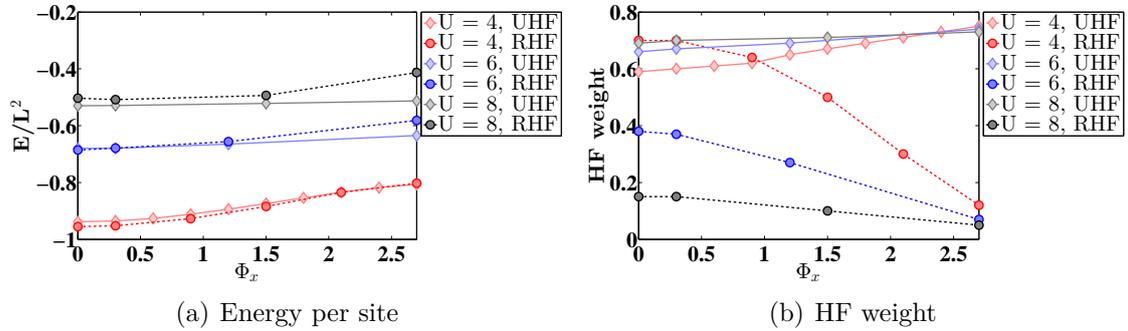


Figure 5.15: (a) Energy and (b) reference determinant weight as function of twist Φ_x for different values of U for the RHF and UHF basis in the 18-site tilted square lattice.

This behavior can be explained by the fact that the energy gap in the half-filled case of the tilted square lattice for the RHF basis decreases with applied twist Φ_x as seen in fig. (5.16a). The energy minimizing spin-density difference as function of U and twist Φ_x is displayed in fig. (5.16b). It can be seen that for a certain twist value of $\Phi_x = \pi$, corresponding to anti-periodic boundary conditions, the UHF solution is

the energetically lower one for every $U > 0$, as in the square lattice with no twist applied. The energy gap/bandwidth ratio for the UHF basis is displayed as function of U for different specific twist values in fig. (5.16b). For the ratio and for the spin contamination of the UHF orbitals, fig. (5.16d), it can be seen that the transition between UHF and RHF properties as function of U changes from discontinuous to continuous behavior for a certain value of twist Φ_x .

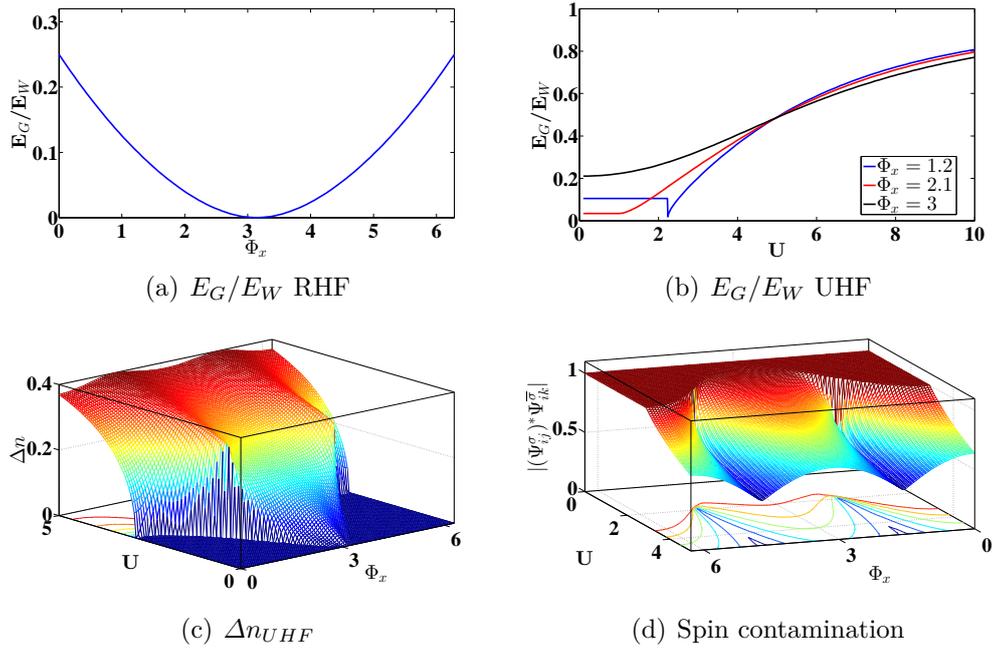


Figure 5.16: Single orbital properties of the RHF and UHF basis in the 18-site tilted square lattice. (a) Energy gap between highest occupied and lowest unoccupied orbital divided by the width of the energy spectrum as function of applied twist Φ_x for the RHF basis. The ratio for RHF case is independent of U , whereas the ratio for the UHF basis (b) is displayed as function of U for specific twist values. (c) The energy minimizing density as function of applied twist and parameter U and the corresponding spin contamination of the UHF basis (d). As can be seen in (c) the maximal value of U_C for the tilted lattice is at zero applied twist and actually goes to zero for a twist $\Phi_x = \pi$ corresponding to anti-periodic boundary conditions.

Fig. (5.17) shows the energy and reference determinant weight of different non-optimal UHF-bases for no applied twist (blue) and a $\Phi_x = 3$ twist value at $U = 4$ in the 18-site tilted square lattice. It can be concluded that also for twisted boundary conditions either the RHF- or the UHF- basis is the optimal one and no UHF-like

basis in between.

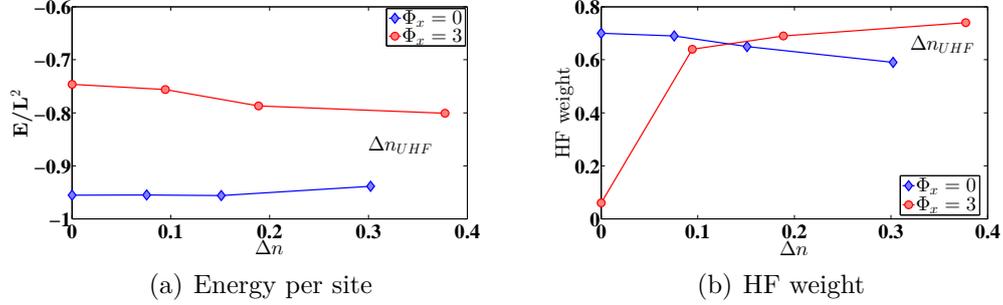


Figure 5.17: (a) Energy per site and (b) reference determinant weight for different twist values as function of spin-density difference for $U = 4$ on the 18-site tilted square lattice.

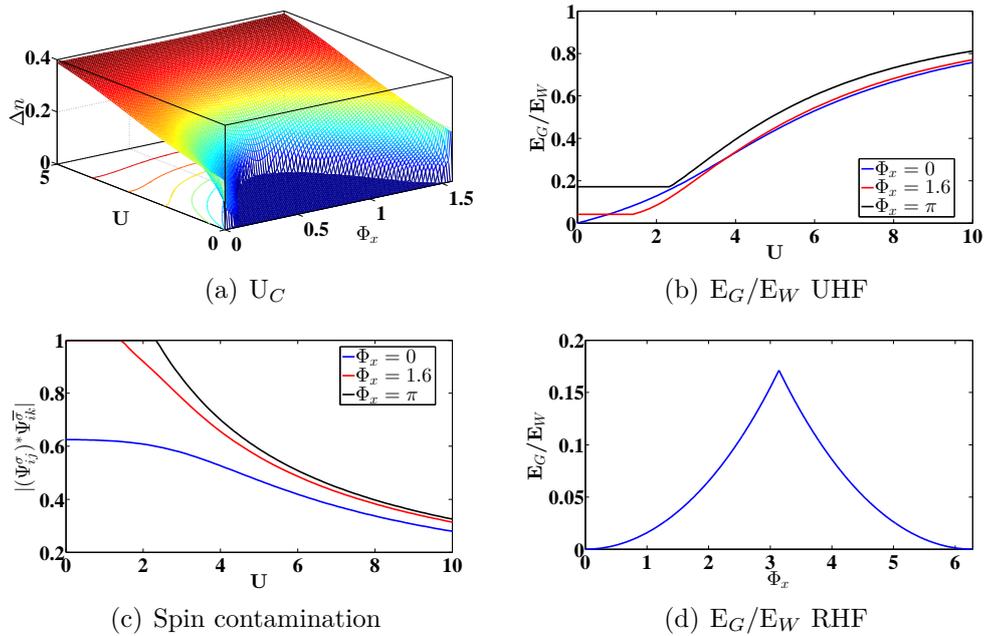


Figure 5.18: Single orbital properties for as function of applied twist for the 4×4 -square lattice. (a) shows the energy minimizing spin-density difference as function of twist and U . In (b) the ratio energy gap/spectrum width as function of U for different twists for the UHF basis is shown and in (c) the corresponding spin contamination. The U -independent E_G/E_W for the RHF basis as function of twist Φ_x is displayed in (d).

In fig. (5.18) the single orbital properties of the 4×4 -square lattice for TBC are displayed. Opposed to the tilted lattice case, the minimum critical $U_C = 0$ is at zero applied twist for the square lattice. And the maximum of the energy gap/bandwidth ratio for the RHF basis is at the anti-periodic $\Phi_x = \pi$ boundary condition.

Conclusions:

It can be concluded that for the RHF and UHF basis a distinct energy gap between the highest occupied and lowest unoccupied orbital greatly improves the FCIQMC simulation properties. And due to the shift of \mathbf{k} -vectors for systems with twisted boundary conditions there are more filling-factors with closed-shell configurations than in the non-twisted case.

Also in each investigated case either the RHF or the UHF basis possessed better simulation properties compared to a basis with a non-optimal electron density difference Δn .

For closed-shell systems, like the half-filled 18-site tilted lattice, and low values of U the RHF basis shows better energy convergence and annihilation plateau height properties compared to the UHF basis. The additional breaking of the spin symmetry of the Hubbard model in the UHF basis seems to deteriorate FCIQMC simulation properties in these systems. Simulations in the RHF basis in systems with open-shell configuration and with increasing values of U get more difficult, meaning a higher number of needed walkers, worse statistical properties and sometimes even no convergence of the simulation at all. But in these system where the RHF basis is in an open-shell configuration, like the half filled 4×4 -square lattice, the UHF basis provides an improvement concerning simulation properties. But the use of the UHF basis may become troublesome in systems off half-filling, due to the fact that the analytic UHF solution is not valid anymore in these systems. But although the iterative self-consistent calculation becomes more problematic for bigger lattices, there have been recent advancements in this topic by Xu et al. [64]. Unfortunately there is not even a partially conserved translation symmetry anymore for the self-consistent UHF solution off half filling.

These problems are topic of ongoing investigations.

5.7 Outlook on larger lattices

The 6×6 -square lattice system with $N_\uparrow = N_\downarrow = 9$ electrons, corresponding to a closed-shell configuration, has an Hilbert space size of $8.9 \cdot 10^{15}$. But only a maximum walker number of one million walkers, corresponding to only $1.1 \cdot 10^{-8}\%$ of the Hilbert space size, were enough to simulate the system up to $U = 4$ without a apparent

initiator error. And for $U = 5$ and $U = 6$ from only 5 million walkers upwards no initiator error was detectable. The total energy and reference determinant occupation for the $U = 1$ and $U = 4$ case are displayed in fig. (5.19). Although the total energy in the $U = 4$ case gets noticeable noisier the value it fluctuates around does not change for a higher number of maximum walkers. The astonishing high reference determinant weight, which is the HF determinant in this case, of over 98% in the $U = 1$ case shows that this system is already good described with the RHF approximation.

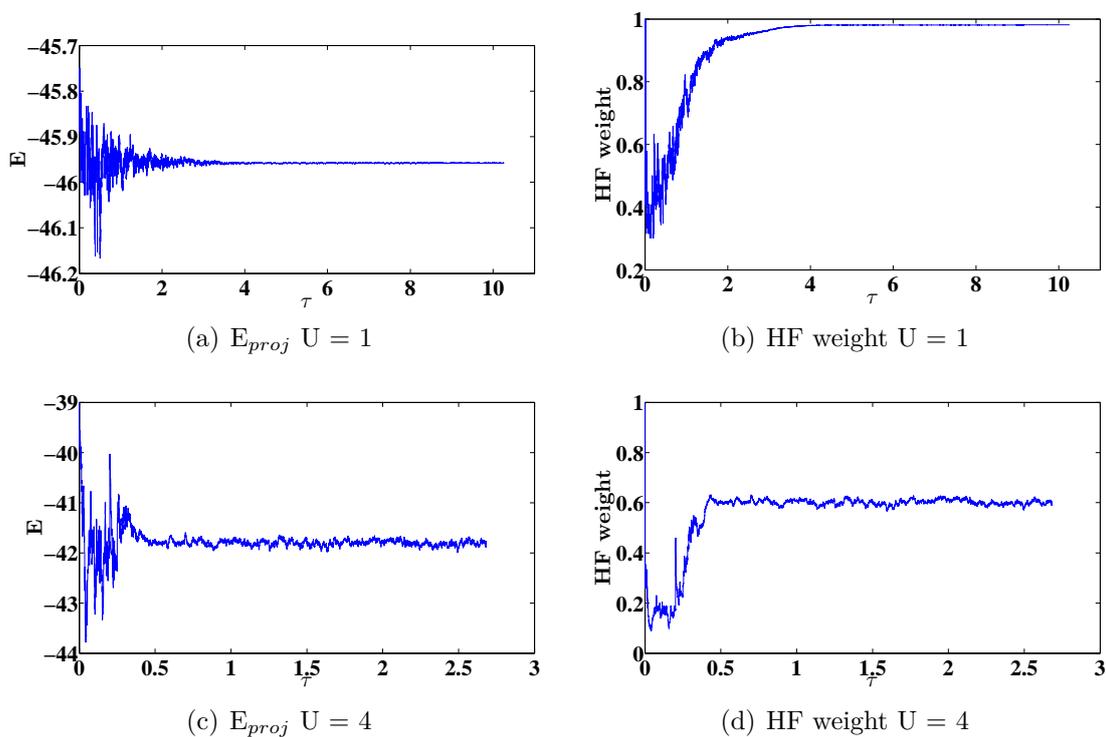


Figure 5.19: Energy and reference determinant weight of the 6×6 -square lattice with $N_{\uparrow} = N_{\downarrow} = 9$ for $U = 1$ (top) and $U = 4$ (bottom) with 1 million maximum walkers.

For $N_{\uparrow} = N_{\downarrow} = 13$ in the 6×6 -square lattice up to $U = 4$ is easily simulated with a fairly low maximum walker number of 5 million. But for $U > 4$ a smaller timestep of $\tau = 0.00001$ and a higher maximum walker number of 10 million upwards was needed to simulate these systems.

The biggest system simulated with only 1 million maximum walkers was the 8×8 -square lattice with $N_{\uparrow} = N_{\downarrow} = 13$ at $U = 1$. The Hilbert space size of this

system is $1.7 \cdot 10^{26}$ meaning only a fraction of $5.8 \cdot 10^{-19}\%$ walkers are sufficient to simulate this system. The weight of the Hartree-Fock determinant in this system is also at an extremely high value of 0.95.

The problem is that only systems which are already well described by the Hartree-Fock approximation, indicated by an high reference determinant weight, are so easily simulated with the *i*-FCIQMC method in the RHF basis. For higher values of U or unfavorable filling factors the simulation time, the needed time step and the maximum number of walkers increase significantly.

And, as mentioned, although the UHF basis provides an improvement over the RHF basis in the half-filled 4×4 -square lattice case, the analytic calculation, based on the anti-ferromagnetic structure of the solution, is not valid off half-filling. And the self-consistent solutions also has convergence problems for larger lattice sizes and additionally does not provide the useful translational invariance.

6 Outlook: Drude and superfluid weight

From the response of a system to an applied phase twisting field Φ , manifesting itself in the ground states energy $E_0(\Phi)$, conclusions can be drawn if a system is in a metallic, insulating or superconducting phase [65]. The phase twist can be interpreted as twisted boundary conditions of the system, meaning a translation of an electron over the boundaries cause a phase factor in the wavefunction:

$$T_x^{L_x} T_y^{L_y} |\Psi\rangle = e^{i\Phi_x} e^{i\Phi_y} |\Psi\rangle, \quad (6.1)$$

where $T_x^{L_x}$ and $T_y^{L_y}$ is the translation in x/y-direction for the linear number of lattice sites of a $L_x \times L_y$ -lattice and $\Phi_{x/y}$ is the value of the phase twist.

Disregarding the phase twist in y-axis direction for now and omitting the x-axis subscript, the twisted boundary conditions can be mapped onto the hopping-matrix of the Hubbard Hamiltonian by a pseudo-gauge transformation,[66]. Where each hopping amplitude is modified by a Peierls phase $e^{\pm iA}$, with the vector potential $A = \frac{\Phi}{L}$ being the fraction of the phase twist and the linear lattice size L . With alternate signs depending on the direction of the electron movement t is transformed to

$$-t \rightarrow -te^{\pm i\frac{\Phi}{L}} = -te^{\pm iA}. \quad (6.2)$$

Applying this transformation to the hopping parts in x-direction of the Hubbard Hamiltonian 1.1

$$\begin{aligned} H(A) = -t \sum_{x,y,\sigma} & \left(e^{iA} c_{(x+1,y),\sigma}^\dagger c_{(x,y),\sigma} + e^{-iA} c_{(x,y),\sigma}^\dagger c_{(x+1,y),\sigma} \right. \\ & \left. + c_{(x,y+1),\sigma}^\dagger c_{(x,y),\sigma} + c_{(x,y),\sigma}^\dagger c_{(x,y+1),\sigma} \right) + U \sum_{i,\sigma} n_{i,\sigma} n_{i,\bar{\sigma}}, \end{aligned} \quad (6.3)$$

where (x,y) in the hopping part are now the two-dimensional indices of the lattice. As seen in section 3.2.3 the effect of this phase factor in the hopping matrix t leads only to an additive shift of the reciprocal lattice vectors in k_x direction.

The Drude weight, or charge stiffness, of a system can be obtained from the second derivative of the ground state energy with respect to the applied phase twist in the

thermodynamic limit [67],[65]

$$\rho_d = \lim_{L \rightarrow \infty} \frac{\pi}{L^2} \frac{\partial^2 E^L}{\partial A^2}. \quad (6.4)$$

It is the ratio of the mobile charge carrier density to their mass

$$\rho_d = \pi e^2 \frac{n}{m} \quad (6.5)$$

and the δ -function coefficient of the long wavelength ($\mathbf{q} \rightarrow 0$) static ($\omega \rightarrow 0$) limit of the real part of the optical conductivity [67]. Hence a finite value in the thermodynamic limit ($L \rightarrow \infty$) would characterize a zero resistance state, while for an insulator it would vanish.

It is important to note that the Drude weight has to be obtained from the curvature of the non-degenerate ground state at zero applied field. But with increasing phase twist level crossings of the many body eigenstates occur. So one has to adiabatically follow the ground state energy to the desired phase twist. Adiabatic following means a slow enough change of outside conditions, so that the system is able to adapt to them. In this case the system remains in an initial eigenstate if there is a gap between the corresponding eigenvalue and the rest of the system [68].

If the twist is applied non-adiabatically, since level crossing of the single many body states occurs, it is not ensured that the system is in the original ground state at a specific Φ . A possible estimation for the critical phase twist Φ_c at which the first level crossing occurs is as follows [65]:

From the second order perturbed Hubbard Hamiltonian (6.3) by expansion of the Peierls phase factor $e^{\pm i \frac{\Phi}{L}}$, an estimate for the change of the ground state energy due to the phase twist is

$$E_0(\Phi) - E_0(0) \approx \Phi^2 + \mathcal{O}(\Phi^4). \quad (6.6)$$

While the typical level spacing δE for many-body eigenstates scales as L^{-d} [65], where d is the dimensionality of the system, in this case $d = 2$. This leads to an estimate for the critical phase

$$\Phi_c \approx L^{-1}. \quad (6.7)$$

This means that Φ_c vanishes in the thermodynamic limit $L \rightarrow \infty$, which causes the order of calculating the curvature $\frac{\partial^2 E^L}{\partial \Phi^2}$ for a finite system and taking the limit $L \rightarrow \infty$ to matter in the calculation of the Drude weight (6.4).

A possibility to obtain the Drude weight is to calculate the curvature of the ground state energy for phase twists below the critical $\Phi_c \approx L^{-1}$ for finite systems first,

$$\rho_d^L = \frac{\pi}{L^2} \frac{\partial^2 E^L}{\partial A^2}, \quad A < \frac{\Phi_x}{L}. \quad (6.8)$$

and afterwards take the thermodynamic limit

$$\rho_d = \lim_{L \rightarrow \infty} \rho_d^L. \quad (6.9)$$

If the weight is calculated the other way around, by first taking the thermodynamic limit of the ground state energies at a finite phase twist

$$E^\infty(A) = \lim_{L \rightarrow \infty} \frac{\pi}{L^2} E^L(A) \quad (6.10)$$

and afterwards calculating the curvature of the infinite-size limit energies

$$\rho_s = \frac{\partial^2 E^\infty(A)}{\partial A^2} \quad (6.11)$$

the resulting quantity can be associated with the superfluid weight of the system [65]. In this case the contributions from the different lowest energy levels, after level crossing, are being averaged and the curvature of the envelope of them is obtained.

The superfluid weight (6.11) is the ratio of the superfluid density n_s and the electron mass m of a system [65]

$$\rho_s = \pi e^2 \frac{n_s}{m}. \quad (6.12)$$

And the superfluid density on the other hand can be related to the penetration depth λ of the Meissner effect [69]

$$\frac{1}{\lambda^2} = \frac{2\pi n_s e^2}{mc^2}. \quad (6.13)$$

The behavior of both the Drude and superfluid weight in the thermodynamic bulk limit $L \rightarrow \infty$ gives insight on whether the ground state of the system is insulating, metallic or superconducting [65]. For an insulator both ρ_d and ρ_s are expected to go to 0. Indicating a resistive behavior with infinite magnetic penetration depth. While a finite value of ρ_d and $\rho_s = 0$ characterizes a metal, since no expulsion of a magnetic field occurs. And finally finite values for both ρ_d and ρ_s imply a superconducting ground state with zero-resistance and an expulsion of a magnetic field up to the

penetration depth λ .

Additionally it has been shown [70] that the Drude weight, opposed to other quantities used to draw conclusions on the electronic behavior of a system, e.g. the compressibility, is less sensitive to finite-size effects. These are the effects from artificial gaps in the energy spectrum of a system due to its finiteness. Since for certain system sizes at filling factors corresponding to a closed-shell configuration, these finite-size gaps can become of the same order of magnitude as gaps of the thermodynamic limit system responsible for certain electronic behaviors. This property of the Drude weight is advantageous to our calculations, as it has been shown in the last chapter 5 that the FCIQMC simulations behave better in systems with a closed-shell configuration.

The problem of the scaling of the critical phase twist where level crossings occur with lattice size (6.7) is that the differences in the ground state energies at different phase twists get very small with increasing lattice sizes. Hence it may become impossible to reliably calculate the curvature of the energy for larger systems due to the fact that the energy differences become of the same order of magnitude as the statistical errors from the Monte Carlo runs. The level crossing between ground state belonging to different \mathbf{k} -sectors can be avoided by use of translational symmetry [71], possible with both RHF and partially with analytically calculated UHF basis states. But this does not prevent level crossing of states within specific \mathbf{k} -sectors.

A workaround of this problem may be the recent advances of Hetényi [72, 73] to associate the Drude and superfluid weight, or transport susceptibilities of a system in general, to eigenstates of the reduced density matrix of a specific order. In [72] it is shown that the superfluid weight can be associated to eigenstates of the one-particle reduced density matrix (1-RDM) of the system, due to off-diagonal long range order (ODLRO) in it. In this case the superfluid weight is given by

$$\rho_s = N \left(1 - \sum_j R_j \right), \quad (6.14)$$

where N is the number of particles of the system and R_j the weight of the eigenstates of the one-particle density matrix, the natural orbitals, which are not simultaneously eigenstates of the one-particle momentum operator to the eigenvalue zero [72]. Further in this paper the two-particle Cooper pairing mechanism, associated with the Meissner weight, is similarly given by eigenstates of the two-particle reduced density matrix, due to ODLRO, not simultaneously being eigenstates of the two-particle momentum operator to the eigenvalue zero.

Hetényi argues that, as it has been proven [74] that if ODLRO exists for a given

order of the reduced p -particle density matrix it exists for all s -RDM $s > p$, it is sufficient to calculate the first finite weight associated with p -RDM. And if p is of order of the total particle number of the system it can be associated with the Drude weight and hence regular conductance of the system. Whereas if all weights associated with reduced density matrices of order p are zero the system can be classified as insulator [72].

As natural orbital basis states, and hence reduced density matrix eigenstates, are currently investigated by the group of Alavi as possible basis for FCIQMC calculations in the Hubbard model [45] the possibility of simultaneous calculation of quantities for description of electronic transport properties is a promising outlook for further studies.

Bibliography

- [1] G. H. Booth and A. Alavi, “Codebase for NECI. FCIQMC code developed by George Booth and Ali Alavi, 2013 Copyright (c) 2013, Ali Alavi.”
URL:https://github.com/ghb24/NECI_STABLE.git.
- [2] J. Hubbard, “Electron correlations in narrow energy bands,” *Proceedings of the Royal Society of London. Series A. Mathematical and Physical Sciences*, vol. 276, pp. 238–257, Nov. 1963.
- [3] J. G. Bednorz and K. A. Müller, “Possible high T_C superconductivity in the Ba-La-Cu-O system,” *Zeitschrift für Physik B Condensed Matter*, vol. 64, pp. 189–193, June 1986.
- [4] F. C. Zhang and T. M. Rice, “Effective Hamiltonian for the superconducting Cu oxides,” *Physical Review B*, vol. 37, pp. 3759–3761, Mar. 1988.
- [5] P. W. Anderson, “The Resonating Valence Bond State in La_2CuO_4 and Superconductivity,” *Science*, vol. 235, pp. 1196–1198, Mar. 1987.
- [6] V. J. Emery, “Theory of high- t_c superconductivity in oxidesw,” *Physical Review Letters*, vol. 58, pp. 2794–2797, June 1987.
- [7] M. K. Wu, J. R. Ashburn, C. J. Torng, P. H. Hor, R. L. Meng, L. Gao, Z. J. Huang, Y. Q. Wang, and C. W. Chu, “Superconductivity at 93 K in a new mixed-phase Y-Ba-Cu-O compound system at ambient pressure,” *Physical Review Letters*, vol. 58, pp. 908–910, Mar. 1987.
- [8] A. Simon, “C. P. Poole jr., T. Datta, H. A. Farach, M. M. Rigney, C. R. Sanders: Copper-oxide superconductors, John Wiley & Sons, New York, Chichester, Brisbane, Toronto, Singapore 1988. 289 Seiten, Preis: £ 22.40,” *Berichte der Bunsengesellschaft für physikalische Chemie*, vol. 93, no. 5, p. 644–645, 1989.
- [9] J. P. Attfield, “Chemistry and high temperature superconductivity,” *Journal of Materials Chemistry*, vol. 21, pp. 4756–4764, Mar. 2011.
- [10] L. F. Mattheiss, “Electronic band properties and superconductivity in $\text{La}_{2-y}\text{X}_y\text{CuO}_4$,” *Physical Review Letters*, vol. 58, pp. 1028–1030, Mar. 1987.

- [11] J. Yu, A. J. Freeman, and J. H. Xu, “Electronically driven instabilities and superconductivity in the layered $\text{La}_{2-x}\text{Ba}_x\text{CuO}_4$ perovskites,” *Physical Review Letters*, vol. 58, pp. 1035–1037, Mar. 1987.
- [12] T. Timusk and B. Statt, “The pseudogap in high-temperature superconductors: an experimental survey,” *Reports on Progress in Physics*, vol. 62, p. 61, Jan. 1999.
- [13] J. W. Loram, K. A. Mirza, J. R. Cooper, and W. Y. Liang, “Electronic specific heat of $\text{YBa}_2\text{Cu}_3\text{O}_{6+x}$ from 1.8 to 300 K,” *Physical Review Letters*, vol. 71, pp. 1740–1743, Sept. 1993.
- [14] A. J. Leggett, “What do we know about high T_c ?,” *Nature Physics*, vol. 2, no. 3, p. 134–136, 2006.
- [15] D. L. Feng, A. Damascelli, K. M. Shen, N. Motoyama, D. H. Lu, H. Eisaki, K. Shimizu, J.-i. Shimoyama, K. Kishio, N. Kaneko, M. Greven, G. D. Gu, X. J. Zhou, C. Kim, F. Ronning, N. P. Armitage, and Z.-X. Shen, “Electronic Structure of the Trilayer Cuprate Superconductor $\text{Bi}_2\text{Sr}_2\text{Ca}_2\text{Cu}_3\text{O}_{10+\delta}$,” *Physical Review Letters*, vol. 88, p. 107001, Feb. 2002.
- [16] R. Jördens, N. Strohmaier, K. Günter, H. Moritz, and T. Esslinger, “A Mott insulator of fermionic atoms in an optical lattice,” *Nature*, vol. 455, pp. 204–207, Sept. 2008.
- [17] D. Jaksch, C. Bruder, J. I. Cirac, C. W. Gardiner, and P. Zoller, “Cold Bosonic Atoms in Optical Lattices,” *Physical Review Letters*, vol. 81, pp. 3108–3111, Oct. 1998.
- [18] J. K. Chin, D. E. Miller, Y. Liu, C. Stan, W. Setiawan, C. Sanner, K. Xu, and W. Ketterle, “Evidence for superfluidity of ultracold fermions in an optical lattice,” *Nature*, vol. 443, pp. 961–964, Oct. 2006.
- [19] W. Hofstetter, J. I. Cirac, P. Zoller, E. Demler, and M. D. Lukin, “High-Temperature Superfluidity of Fermionic Atoms in Optical Lattices,” *Physical Review Letters*, vol. 89, p. 220407, Nov. 2002.
- [20] I. Bloch, J. Dalibard, and W. Zwerger, “Many-body physics with ultracold gases,” *Reviews of Modern Physics*, vol. 80, pp. 885–964, July 2008.
- [21] U. Schneider, L. Hackermüller, S. Will, T. Best, I. Bloch, T. A. Costi, R. W. Helmes, D. Rasch, and A. Rosch, “Metallic and Insulating Phases of Repulsively Interacting Fermions in a 3D Optical Lattice,” *Science*, vol. 322, pp. 1520–1525, May 2008.

- [22] E. H. Lieb and F. Y. Wu, “Absence of Mott Transition in an Exact Solution of the Short-Range, One-Band Model in One Dimension,” *Physical Review Letters*, vol. 20, pp. 1445–1448, June 1968.
- [23] J. Spalek, “t-J model then and now: A personal perspective from the pioneering times,” *arXiv:0706.4236 [cond-mat]*, June 2007.
- [24] 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,” *The Journal of Chemical Physics*, vol. 131, p. 054106, Aug. 2009.
- [25] D. Cleland, G. H. Booth, and A. Alavi, “Communications: Survival of the fittest: Accelerating convergence in full configuration-interaction quantum Monte Carlo,” *The Journal of Chemical Physics*, vol. 132, no. 4, p. 041103, 2010.
- [26] D. Cleland, G. H. Booth, C. Overy, and A. Alavi, “Taming the First-Row Diatomics: A Full Configuration Interaction Quantum Monte Carlo Study,” *Journal of Chemical Theory and Computation*, vol. 8, pp. 4138–4152, Nov. 2012.
- [27] G. H. Booth, D. Cleland, A. J. W. Thom, and A. Alavi, “Breaking the carbon dimer: The challenges of multiple bond dissociation with full configuration interaction quantum Monte Carlo methods,” *The Journal of Chemical Physics*, vol. 135, p. 084104, Aug. 2011.
- [28] J. J. Shepherd, G. Booth, A. Grüneis, and A. Alavi, “Full configuration interaction perspective on the homogeneous electron gas,” *Physical Review B*, vol. 85, Feb. 2012.
- [29] G. H. Booth, A. Grüneis, G. Kresse, and A. Alavi, “Towards an exact description of electronic wavefunctions in real solids,” *Nature*, vol. 493, pp. 365–370, Dec. 2012.
- [30] C. Lanczos, *An iteration method for the solution of the eigenvalue problem of linear differential and integral operators*. National Bureau of Standards, 1950.
- [31] A. Weiße and H. Fehske, “Exact Diagonalization Techniques,” in *Computational Many-Particle Physics* (H. Fehske, R. Schneider, and A. Weiße, eds.), no. 739 in Lecture Notes in Physics, pp. 529–544, Springer Berlin Heidelberg, Jan. 2008.
- [32] A. W. Sandvik, “Computational studies of quantum spin systems,” *arXiv preprint arXiv:1101.3281*, 2011.

- [33] A. M. Läuchli, “Numerical Simulations of Frustrated Systems,” in *Introduction to Frustrated Magnetism*, p. 481–511, Springer, 2011.
- [34] P. W. Atkins and R. Friedman, *Molecular quantum mechanics*. Oxford; New York: Oxford University Press, 2011.
- [35] N. Laflorencie and D. Poilblanc, “Simulations of pure and doped low-dimensional spin-1/2 gapped systems,” *arXiv:cond-mat/0408363*, vol. 645, pp. 227–252.
- [36] E. Dagotto, A. Moreo, F. Ortolani, D. Poilblanc, and J. Riera, “Static and dynamical properties of doped Hubbard clusters,” *Physical Review B*, vol. 45, pp. 10741–10760, May 1992.
- [37] H. Lin, “Ground-state properties of the two-dimensional Hubbard model,” *Physical Review B*, vol. 44, pp. 7151–7154, Oct. 1991.
- [38] A. Szabo and N. S. Ostlund, *Modern quantum chemistry: introduction to advanced electronic structure theory*. Dover, 1996.
- [39] W. M. C. Foulkes, L. Mitas, R. J. Needs, and G. Rajagopal, “Quantum Monte Carlo simulations of solids,” *Reviews of Modern Physics*, vol. 73, pp. 33–83, Jan. 2001.
- [40] M. L. Volker Bach, Elliott H. Lieb and J. P. Solovej, “There are no unfilled shells in unrestricted Hartree-Fock theory,” *Physical Review Letters*, vol. 72, no. 19, pp. 2981–2983, 1994.
- [41] K. H. Lang, “Diffusion Monte Carlo applied to the two-dimensional fermion Hubbard model,” Master’s thesis, Graz University of Technology, 2012.
- [42] E. Louis, F. Guinea, M. L. Sancho, and J. A. Vergés, “Configuration-interaction approach to hole pairing in the two-dimensional Hubbard model,” *Physical Review B*, vol. 59, no. 21, p. 14005, 1999.
- [43] E. Louis, G. Chiappe, F. Guinea, J. A. Vergés, and E. V. Anda, “Multipole-polaron description of the wave function of a single hole in Hubbard clusters of the square lattice,” *Physical Review B*, vol. 48, no. 13, p. 9581, 1993.
- [44] “Private Communication Wolfgang von der Linden.”
- [45] “Private communications Ali Alavi.”
- [46] I. Shavitt, B. J. Rosenberg, and S. Palalikit, “Comparison of configuration interaction expansions based on different orbital transformations,” *International Journal of Quantum Chemistry*, vol. 10, no. S10, p. 33–46, 1976.

- [47] H. F. Schaefer, *Methods of electronic structure theory*. New York: Plenum Press, 1977.
- [48] W. Feller, *An introduction to probability theory and its applications. Volume I*. New York; London; Sydney: J. Wiley & sons, 1968.
- [49] N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, and E. Teller, "Equation of State Calculations by Fast Computing Machines," *Journal of Chemical Physics*, vol. 21, pp. 1087–1092, June 1953.
- [50] M. Troyer and U.-J. Wiese, "Computational Complexity and Fundamental Limitations to Fermionic Quantum Monte Carlo Simulations," *Physical Review Letters*, vol. 94, p. 170201, May 2005.
- [51] H. Flyvbjerg and H. G. Petersen, "Error estimates on averages of correlated data," *The Journal of Chemical Physics*, vol. 91, pp. 461–466, July 1989.
- [52] M. H. Kalos, "Monte Carlo Calculations of the Ground State of Three- and Four-Body Nuclei," *Physical Review*, vol. 128, pp. 1791–1795, Nov. 1962.
- [53] W. von der Linden, "A quantum Monte Carlo approach to many-body physics," *Physics Reports*, vol. 220, pp. 53–162, Nov. 1992.
- [54] N. Trivedi and D. M. Ceperley, "Ground-state correlations of quantum anti-ferromagnets: A Green-function Monte Carlo study," *Physical Review B*, vol. 41, pp. 4552–4569, Mar. 1990.
- [55] J. S. Spencer, N. S. Blunt, and W. M. Foulkes, "The sign problem and population dynamics in the full configuration interaction quantum Monte Carlo method," *The Journal of Chemical Physics*, vol. 136, no. 5, p. 054110, 2012.
- [56] P. J. Reynolds, D. M. Ceperley, B. J. Alder, and W. A. L. Jr, "Fixed-node quantum Monte Carlo for molecules," *The Journal of Chemical Physics*, vol. 77, pp. 5593–5603, Dec. 1982.
- [57] H. De Raedt and A. Lagendijk, "Monte Carlo Calculation of the Thermodynamic Properties of a Quantum Model: A One-Dimensional Fermion Lattice Model," *Physical Review Letters*, vol. 46, pp. 77–80, Jan. 1981.
- [58] G. H. Booth, S. D. Smart, and A. Alavi, "Linear-scaling and parallelizable algorithms for stochastic quantum chemistry," *arXiv:1305.6981 [cond-mat, physics:physics]*, May 2013.
- [59] M. H. Kolodrubetz, J. S. Spencer, B. K. Clark, and W. M. Foulkes, "The effect of quantization on the full configuration interaction quantum Monte Carlo sign problem," *The Journal of Chemical Physics*, vol. 138, no. 2, p. 024110, 2013.

- [60] M. H. Kalos, “Correlated pairs in model fermion Monte Carlo calculations,” *Physical Review E*, vol. 53, pp. 5420–5427, May 1996.
- [61] S. Zhang and M. H. Kalos, “Exact Monte Carlo calculation for few-electron systems,” *Physical Review Letters*, vol. 67, pp. 3074–3077, Nov. 1991.
- [62] J. Carlson and M. H. Kalos, “Mirror potentials and the fermion problem,” *Physical Review C*, vol. 32, pp. 1735–1741, Nov. 1985.
- [63] J. J. Shepherd, L. R. Schwarz, R. E. Thomas, G. Booth, D. Frenkel, and A. Alavi, “Emergence of Critical Phenomena in Full Configuration Interaction Quantum Monte Carlo,” *arXiv preprint arXiv:1209.4023*, 2012.
- [64] J. Xu, C.-c. Chang, E. J. Walter, and S. Zhang, “Spin- and charge-density waves in the hartree-fock ground state of the two-dimensional hubbard model,” vol. 23, p. 1–13.
- [65] D. J. Scalapino, S. R. White, and S. Zhang, “Insulator, metal, or superconductor: The criteria,” *Physical Review B*, vol. 47, no. 13, p. 7995, 1993.
- [66] N. Byers and C. N. Yang, “Theoretical Considerations Concerning Quantized Magnetic Flux in Superconducting Cylinders,” *Physical Review Letters*, vol. 7, pp. 46–49, July 1961.
- [67] W. Kohn, “Theory of the Insulating State,” *Physical Review*, vol. 133, pp. A171–A181, Jan. 1964.
- [68] M. Born and V. Fock, “Beweis des Adiabatenatzes,” *Zeitschrift für Physik*, vol. 51, pp. 165–180, Mar. 1928.
- [69] J. R. Schrieffer, *Theory of Superconductivity*.
Advanced Book Program, Perseus Books, 1983.
- [70] R. Mondaini, K. Bouadim, T. Paiva, and R. R. d. Santos, “Finite-size effects in transport data from Quantum Monte Carlo simulations,” *Physical Review B*, vol. 85, Mar. 2012.
- [71] D. Poilblanc, “Twisted boundary conditions in cluster calculations of the optical conductivity in two-dimensional lattice models,” *Physical Review B*, vol. 44, pp. 9562–9581, Nov. 1991.
- [72] B. Hetényi, “Drude weight, Meissner weight, rotational inertia of bosonic superfluids: how are they distinguished?,” *Journal of the Physical Society of Japan*, vol. 83, p. 034711, Oct. 2014.
- [73] “Private communications Balázs Hetényi.”

-
- [74] C. N. Yang, “Concept of Off-Diagonal Long-Range Order and the Quantum Phases of Liquid He and of Superconductors,” *Reviews of Modern Physics*, vol. 34, pp. 694–704, Oct. 1962.