Quantum Monte Carlo Calculation of the Imaginary-Time Green's Function in the Hubbard Model

---

A Thesis
Presented to
The Division of Mathematics and Natural Sciences
Reed College

In Partial Fulfillment
of the Requirements for the Degree
Bachelor of Arts

Huy Nguyen

May 2014

Approved for the Division
(Physics)

_____

Darrell Schroeter

Acknowledgements

First and foremost, I would like to thank my thesis advisor, Prof. Darrell Schroeter, for his wonderful guidance, support, encouragement and wisdom. He helped me through difficult derivations, helped me see the big picture and was very supportive in helping me obtain funding for this thesis. I could not have asked for a better advisor.

I would like to thank my collaborator, Prof. Shiwei Zhang from the College of William and Mary, for mentoring me during two summer REUs, for suggesting this thesis topic to me, and for being an invaluable source of ideas. This thesis would not have been possible without his expertise. I would like to thank Hao Shi who explained many concepts in quantum Monte Carlo to me when I first started.

I also would like to thank professors Lucas Illing, Jerry Shurman and Virginia Hancock for serving on my orals board.

I would like to acknowledge financial support from Reed College through an Undergraduate Research Initiative Grant and from the Reed physics department. I would also like to acknowledge computational support from the SciClone high-performance computing cluster at the College of William and Mary.

Finally, I would like to thank my family for their constant emotional support.

Contents

Introduction **1**

Chapter 1: Non-interacting systems **7**
1.1 Second quantization 7
1.1.1 System of identical particles 7
1.1.2 Creation and annihilation operators 10
1.1.3 Many-body operators 12
1.2 Non-interacting electrons 12
1.2.1 Equal-time Green's function 14
1.2.2 Unequal-time Green's function 16
1.2.3 Non-interacting Hubbard model in an external potential 18

Chapter 2: The Hubbard Model **21**
2.1 The Hubbard Hamiltonian 21
2.2 Some properties of the Hubbard Hamiltonian 22
2.2.1 Conservation of particle numbers 22
2.2.2 Independence of the spin sectors in a Slater determinant 22
2.3 Exact Diagonalization 23

Chapter 3: Constrained Path Monte Carlo **27**
3.1 Slater determinant space 27
3.2 The Hubbard Hamiltonian 30
3.3 Ground-state projection 31
3.4 Suzuki-Trotter decomposition 32
3.5 The Hubbard-Stratonovich transformation 33
3.6 A toy model for illustration 34
3.7 Random walk in Slater determinant space 35
3.8 Importance sampling 36
3.8.1 Importance-sampled random walkers 37
3.8.2 Modified Hirsch spin transformation 37
3.9 The sign problem and the constrained path approximation 39
3.10 Measurement of physical observables 39
3.10.1 The mixed estimator and energy 40
3.10.2 Non-commuting observables and back propagation 41
3.11 Unequal-time Green's function 42
3.11.1 Non-interacting electrons in an external potential 43
3.11.2 Hubbard interaction 50

3.12 Other implementation issues . . . 52
3.12.1 Population control . . . 52
3.12.2 Gram-Schmidt reorthonormalization . . . 52
3.13 Algorithm . . . 53
3.14 Results . . . 54

Chapter 4: Applications . . . **57**
4.1 Application of the imaginary-time Green's function . . . 57
4.2 Neutron scattering by a crystal . . . 57
4.3 Magnetic structure factor . . . 60

Appendix A: Hubbard-Stratonovich transformation . . . **65**

Appendix B: Hirsh transformations . . . **69**

Appendix C: Overlap of two Slater determinants . . . **73**

Appendix D: Thouless theorem . . . **75**

Appendix E: Correctness of the importance-sampled propagator . . . **81**

Appendix F: Equal-time Green's function . . . **85**

Appendix G: Metropolis AFQMC . . . **89**

Appendix H: Unequal-time Green's function . . . **93**

Appendix I: Back-propagation . . . **101**

Appendix J: Stabilization of matrix multiplication . . . **107**

Appendix K: Two-body operators . . . **109**

Abstract

In this thesis, we develop an extension to the Constrained Path Monte Carlo method to calculate the imaginary-time Green's function in the one-band one-dimensional Hubbard model. The results of the Monte Carlo calculations agree very well with exact results (for systems where exact results are available). We then use the calculated Green's functions to probe the magnetic ordering of a large system beyond the reach of exact calculations.

Introduction

The strong correlation problem

The accurate prediction of material properties requires robust and reliable calculations at the microscopic level. At this level, the behavior of particles (specifically electrons in this thesis) is governed by the many-body Schrödinger equation. Given a potential (rules governing how particles interact with each other and their surroundings), this partial differential equation describes how the quantum state of a physical system changes with time. It plays a role similar to Newton's second law in classical mechanics which relates force to mass and acceleration. Despite the Schrödinger equation's simplicity, we only know its analytical solutions in a small number of cases: particles in an infinitely-deep potential well, in a spring-like harmonic potential and so on. Beyond these simple models, we have to resort to numerical methods.

According to Bajdich and Mitas [1], there are four fundamental challenges to the quantum many-body problem. The first challenge is the large number of particles in a typical system of interest: a moderate size solid contains on the order of 10^{23} quantum particles. For systems containing a small enough number of particles, we can use exact numerical methods, such as exact diagonalization [2]. However, these methods are limited by their *exponential* computational complexity: if N is the number of particles in a system, the time taken to compute some property is proportional to e^N. Algorithms with exponential complexity are considered computationally prohibitive except for small systems.

To circumvent this exponential scaling, early quantum mechanical studies of large systems rely on simpler independent-particle methods, known as "mean-field" methods. An example is the Hartree-Fock method [3, pp. 39-230]. This method assumes that electrons in a solid feel an "average" effective potential caused by other surrounding nuclei and electrons, and that the true potential does not deviate too far from this average potential. By replacing the true many-body interactions between electrons with this average potential, the method reduces a many-electron problem to a one-electron problem that can be solved more easily.

The second challenge comes from quantum interactions (or "correlations") between particles. These interactions create novel electronic and magnetic properties in strongly correlated materials i.e. those in which the behavior of electrons cannot be described in terms of non-interacting entities. These materials include [4] high-temperature superconductors (zero resistance at unusually high temperatures), heavy-fermion metals (mobile electrons with effective mass thousands of times that of normal electrons), colossal magnetoresistance (great sensitivity of resistance to small changes in an applied magnetic field) and so on. Strongly-correlated systems is one of the most exciting areas of modern condensed matter physics. Hartree-Fock-like methods do not treat correlation well because all electrons are subject to the same potential. Density functional theory, a widely used and successful class of methods, partly overcomes this problem by estimating the many-body correlation with an exchange correlation functional.

The third challenge is the high accuracy required for predictions that are comparable to experiments. In strongly-correlated materials, even small errors in the electronic correlation can lead to crucial and qualitative differences in physical properties. Despite the success of density functional theory, its treatment of electronic correlation is entirely approximate [5, p. 1]. Thus, alternatives to independent-particle theories are needed, among which quantum Monte Carlo (QMC) is a promising candidate.

Unlike other methods, QMC [5] is stochastic in nature i.e its results have associated statistical uncertainties. However, these uncertainties are inversely proportional to the square root of the number of simulations and, more importantly, independent of the number of particles. By using stochastic sampling instead of explicit integration over phase space (which scales exponentially), QMC has computational costs that scale polynomially (about N^3) with the number of particles N and enables computation of very large systems. The N-dependence, in the case of CPMC, comes from the complexity of regular matrix operations. Due to the nature of Monte Carlo integration, QMC methods have exceptional potential for parallel scaling which enables them to easily take advantage of current advances in high-performance computing like general-purpose computing on GPUs (graphics processing units).

The fourth challenge comes from the required symmetry of the solution as dictated by quantum mechanics, a fundamental departure from classical systems. For example, a solution for electrons has to be antisymmetric i.e. if we switch the labels of two electrons in the system, the solution must negate itself. This antisymmetry requirement causes the "fermion sign problem" [6] that afflicts all fermionic QMC methods. Fortunately, CPMC is able to control this problem with the constrained path approximation, discussed in Section 3.9.

The Hubbard model

Although the ideal goal of condensed matter physics is the treatment of electrons in a solid, the full interaction problem between all electrons and all ions is too difficult to tackle directly. Nevertheless, much insight can be gained from the analysis of simplified models that retain the essential physics, one of which is the Hubbard model, which was first written down by John Hubbard in 1963 [7].

The first simplification in the Hubbard model is the Born-Oppenheimer approximation (also used widely in quantum chemistry). Because the nuclei in a solid are a few thousand times heavier than electrons, we can think of the nuclei as forming a static lattice. One can also assume that the electrons only have short-range interactions. This is often a good approximation in practice because the screening by electrons is enough to render long range interaction negligible.

These two assumptions give rise to the Hubbard model to describe electron interactions on a lattice. There are two contributions to the energy in the model. The kinetic energy term allows electrons to move between localized states on adjacent sites of the lattice. These electrons interact in the form of a Coulomb repulsion introduced by the potential energy term when they meet on the same lattice site (due to the Pauli exclusion principle, there can be at most two electrons on a lattice site). These two terms compete because the kinetic part favors electrons being as mobile as possible while the repulsive potential part encourages electrons to stay apart from each other, i.e. localized on different atomic sites. This competition is at the heart of the electronic many-body problem.

The Hubbard model is the simplest model for many-electron interactions, an important motivation for its studies. By varying the strength of the repulsion, we have a continuous

regime from weak to strong correlation. It has been used in attempts to describe [8, p. 5] the electronic properties of solids with narrow bands; band magnetism in iron, cobalt, nickel; the Mott metal-insulator transition; and electronic properties of high-temperature cuprate superconductors in the normal state.

Despite its apparent simplicity, no consistent treatment of the model exists for lattices in more than one dimension. Fortunately in one-dimensional Hubbard lattices, many properties can be calculated exactly by a number of methods. For example, in 1968, just a few years after Hubbard came up with the model,[1] an analytic formula for the ground state energy was found by Lieb and Wu [9] using the Bethe ansatz.

Quantum Monte Carlo

Quantum Monte Carlo (QMC) is a class of stochastic algorithms that use the Monte Carlo technique to compute properties of quantum systems. The stochasticity means that instead of giving a definite numerical result, QMC calculations give results that have associated statistical uncertainties that can be reduced algebraically with more computer run time. There are many QMC algorithms, which can be classified based on whether they work in ground state (zero-temperature) or finite-temperature; are variational or projective and if projective, the projection operator used ($1 - \tau\widehat{H}$, $\frac{1}{1-\tau(E_T-\widehat{H})}$ or $e^{-\tau(\widehat{H}-E_T)}$); work in continuum or lattice; and whether they take place in first quantization or second quantization. The terms highlighted in red are the characteristics of CPMC.

We will now provide a brief literature review leading to the work in this thesis. The technique of Markov chain Monte Carlo, which is used by CPMC and most QMC methods, can be traced back to the invention of the Metropolis algorithm[2] by Metropolis *et al.* [11] in 1953. This algorithm allows us to sample any probability distribution function without knowing its normalization by generating a Markov chain whose stationary distribution obeys the desired distribution function. This technique gave rise to Variational Monte Carlo (VMC), the first QMC method, by McMillan [12] in 1965. VMC uses the Metropolis algorithm to sample the squared magnitude of a trial wave function (which is the probability of finding an electron at a point in space) and then evaluates the expectation value for some observable with that trial wave function. Because of the large number of particle coordinates, Monte Carlo integration is much more efficient than conventional quadrature e.g. summing the areas of rectangular boxes under a curve.

Nevertheless, the accuracy of VMC depends entirely on the accuracy of the trial wave function. Projective Monte Carlo methods, including CPMC, overcome this limitation by using a projection technique to enhance the ground-state component of a starting trial wave function. A very popular projective Monte Carlo method is the Diffusion Monte Carlo (DMC) method, pioneered by Anderson [13] in 1975. This ground state method works in configuration space and uses an open-ended branching random walk in imaginary time to sample the ground state. The transition probability for this random walk is constructed by invoking the similarity between the Schrödinger equation in imaginary time and the diffusion equation with a source term.

[1] In fact, the title of the seminal paper by Lieb and Wu [9] referred to the model only as "the short-range, one-band model in one dimension" instead of "the Hubbard model," possibly because the name Hubbard has yet to be widely associated with the model at that time.

[2] The Metropolis (also known as M(RT)^2) algorithm was named one of the top 10 algorithms of the 20th century by The Society for Industrial and Applied Mathematics [10] in 2000.

All the aforementioned methods work in first quantization and now we will introduce QMC methods that work in second quantization. The field of auxiliary-field quantum Monte Carlo (AFQMC)—the introduction of auxiliary fields through the Hubbard-Stratonovich transformation to convert an interacting system into many non-interacting systems—was pioneered in 1981 by Blankenbecler *et al.* [14] and Scalapino and Sugar [15] with a finite-temperature algorithm and was followed up in 1986 by Sugiyama and Koonin [16] with a zero-temperature (i.e. ground state) algorithm. Between these two developments, in 1983, Hirsch [17] invented a discrete form of the Hubbard-Stratonovich transformation for the Hubbard Hamiltonian. This transformation became very popular in latter AFQMC calculations of the Hubbard model and Hirsch [18] himself used this transformation to study many physical properties of the Hubbard model at finite temperature. All these methods use the Metropolis formalism.

However, all these early pioneers had to contend with the fermion sign problem. This problem arises because of the antisymmetry of fermionic wave functions and causes large statistical errors at low temperature, large imaginary times or strong correlations, precisely the regimes that yield interesting physics. In a very insightful paper in 1991, Fahy and Hamann [19] showed how the introduction of auxiliary fields through the Hubbard-Stratonovich transformation sets up a diffusion equation (with branching and drift) on the Grassmann manifold of normalized Slater determinants. Based on this observation, they proposed [20] the positive projection approximation to reduce or eliminate the sign problem. However, this approximation is very computationally expensive when integrated into existing AFQMC methods because they all use the Metropolis formalism.

The ground-state CPMC method was pioneered by Zhang *et al.* [6] in 1997. It combines techniques from DMC (open-ended branching random walk in imaginary time, importance sampling) and Metropolis AFQMC (Hubbard-Stratonovich transformation, Slater determinants) to do ground state projection. The major breakthrough here is that the branching random walk allows for a very easy implementation of the aforementioned positive projection approximation by Fahy and Hamann [20], now called the constrained path approximation. In 1999, Zhang [21] extended CPMC to treat systems at finite-temperature. Subsequently in 2003, Zhang and Krakauer [22] extended CPMC again to treat any interactions. This most general formulation of CPMC is called the phaseless AFQMC method.

In 2001, Feldbacher and Assaad [23] invented an elegant and efficient way to calculate the unequal-time Green's function in Metropolis AFQMC. Earlier this year, Motta *et al.* [24] used the phaseless AFQMC and the aforementioned method by Feldbacher and Assaad [23] to calculate the unequal-time Green's function of the jellium model. This last paper comes the closest to the objective of this thesis: **to develop an extension to the constrained path Monte Carlo method by incorporating the method by Feldbacher and Assaad [23] into the back-propagation technique by Zhang *et al.* [6] in order to calculate the unequal-time Green's function[3] in the one-band one-dimensional Hubbard model**. The longer-term goal is to perform these calculations in the two-dimensional Hubbard model where there is no analytic solution. As the starting point for implementation, we will use CPMC-Lab, a MATLAB package developed in Nguyen *et al.* [25] that can perform ground-state CPMC calculations to find the ground state energy of the Hubbard model in up to three dimensions.

[3] In this thesis, we use the term "imaginary-time Green's function" to include both the equal-time and unequal-time Green's function. The equal-time Green's function has already been calculated in CPMC using the back-propagation technique by Zhang *et al.* [6].

Layout of the thesis

This thesis is organized as follows. In chapter 1, we introduce the formalism of second quantization, which is used extensively in many-body quantum theory. We study purely non-interacting electron systems and then non-interacting systems but subject to an external potential. These systems are easy to solve and will serve as a testbed to help us understand the numerical behavior of the extension to the CPMC method that we will develop in chapter 3. They are also used to check the results of the non-interacting case of the algorithm developed in chapter 2.

In chapter 2, we will introduce the Hamiltonian of the one-band one-dimensional Hubbard model, the principal subject of this thesis. We will then discuss a number of useful properties of this Hamiltonian and state the exact diagonalization (ED) algorithm to obtain the unequal-time Green's function of the model exactly.

In chapter 3, we describe the existing CPMC algorithm and the extension we have developed in detail.

Finally, in chapter 4, we discuss the applications of the imaginary-time Green's function.

We want to emphasize that the appendices constitute an important part of this thesis and could have functioned perfectly as parts of the main body. We have elected to put them in appendices in order to not interrupt the flow of ideas in the main body. They mostly contain proofs of theorems or properties that are referred to in the main body. Because many of these theorems and properties are standard results, they are routinely stated without proofs in the literature, which might frustrate a beginner entering the field. This thesis hopes to remedy this situation, to some extent.

Although this thesis is about "quantum Monte Carlo," we will spend most of the time developing the "quantum" part, that is, writing the quantum mechanical quantities we wish to compute in a form that is amenable to a Monte Carlo solution and then showing the Monte Carlo algorithms for such computations. We will not spend a lot time on the "Monte Carlo" part because there are many excellent books that discuss the technique, such as Kalos and Whitlock [26].

The most up-to-date version of this thesis (in PDF format) can always be obtained from the following website:

www.huy.dev/Reed-thesis.

For all inquiries, questions, comments and errors, please contact the author at the following email address:

me@huy.dev.

All journal articles in the bibliography are hyperlinked and can be accessed by clicking on the journal citations.

Chapter 1

Non-interacting systems

In this chapter, we will introduce the formalism of second quantization, which is used extensively in many-body quantum theory. After that, we will derive expressions for the equal- and unequal-time Green's functions for purely non-interacting electron systems and then non-interacting systems but subject to an external potential. These systems are easy to solve and will serve as a testbed to help us understand the numerical behavior of the extension to the CPMC method that we will develop in chapter 3.

1.1 Second quantization

Quantum mechanics of a single particle is usually formulated in terms of the position operator $\widehat{x}$ and momentum operator $\widehat{p}$. All physical observables can be expressed in terms of $\widehat{x}$ and $\widehat{p}$ and a natural representation for quantum mechanics, the coordinate representation, is defined in terms of eigenfunctions of $\widehat{x}$.

Similarly, the quantum mechanics of many identical particles is described by 'second quantization' which is formulated in terms of creation and annihilation operators. All operators of interest can be expressed in terms of these creation and annihilation operators. This section has been adapted from Negele and Orland [27, ch. 1].

1.1.1 System of identical particles

The Hilbert space of states for a system of N identical particles is the space $\mathcal{H}_N$ of complex square integrable functions $\Psi_N(\vec{r}_1, \ldots, \vec{r}_N)$ which represents the probability amplitude of finding particles at one particular configuration of the N positions $\vec{r}_1, \ldots, \vec{r}_N$. $\mathcal{H}_N$ is simply the N-th tensor product of the single particle Hilbert space:

$$\mathcal{H}_N = \mathcal{H} \otimes \mathcal{H} \otimes \cdots \otimes \mathcal{H} \,. \tag{1.1}$$

If $\{|\alpha\rangle\}$ is an orthonormal basis of $\mathcal{H}$, the canonical basis of $\mathcal{H}_N$ is constructed form the tensor products

$$|\alpha_1 \ldots \alpha_N) \equiv |\alpha_1\rangle \otimes |\alpha_2\rangle \otimes \cdots \otimes |\alpha_N\rangle \,. \tag{1.2}$$

Note that these basis states of $\mathcal{H}_N$ utilize parentheses in the ket symbol. The wave functions

associated with these basis states are

$$\psi_{\alpha_1\ldots\alpha_N}(\vec{r}_1,\ldots,\vec{r}_N) = (\vec{r}_1\ldots\vec{r}_N|\alpha_1\ldots\alpha_N) \tag{1.3}$$

$$= \Big(\langle\vec{r}_1|\otimes\cdots\otimes\langle\vec{r}_N|\Big)\Big(|\alpha_1\rangle\otimes\cdots\otimes|\alpha_N\rangle\Big) \tag{1.4}$$

$$= \phi_{\alpha_1}(\vec{r}_1)\phi_{\alpha_2}(\vec{r}_2)\ldots\phi_{\alpha_N}(\vec{r}_N)\,. \tag{1.5}$$

The inner product of two vectors in this basis obeys the usual rules for tensor products:

$$(\alpha_1\ldots\alpha_N|\alpha'_1\ldots\alpha'_N) = \Big(\langle\alpha_1|\otimes\cdots\otimes\langle\alpha_N|\Big)\Big(|\alpha'_1\rangle\otimes\cdots\otimes|\alpha'_N\rangle\Big) \tag{1.6}$$

$$= \langle\alpha_1|\alpha'_1\rangle\ldots\langle\alpha_N|\alpha'_N\rangle \tag{1.7}$$

and the completeness of this basis follows from the completeness of the $\{|\alpha\rangle\}$ basis:

$$\sum_{\alpha_1,\ldots,\alpha_N} |\alpha_1\ldots\alpha_N)\,(\alpha_1\ldots\alpha_N| = 1\,, \tag{1.8}$$

where 1 is the unit operator in $\mathcal{H}_N$. From Eqs. (1.5) and (1.8), we can see that the space $\mathcal{H}_N$ is generated by linear combinations of products of single-particle wave functions.

So far the symmetry property of the wave functions have not been considered in the construction of $\mathcal{H}_N$. However, the spin statistics theorem requires that wave functions for fermions be totally antisymmetric under particle exchange. That is, if P is a permutation of the set $(1,2,\ldots,N)$ then a fermionic wave function satisfies

$$\psi(\vec{r}_{P1},\ldots,\vec{r}_{PN}) = (-1)^P\,\psi(\vec{r}_1,\ldots,\vec{r}_N) \tag{1.9}$$

where $(-1)^P$ denotes the parity of the permutation P. To take symmetry into account, we define the antisymmetrization operator $\mathcal{P}$ that, when acting on a many-particle wave function, creates a linear combination out of all possible permutations of the particle labels and gives each term an appropriate sign so as to make the combination antisymmetric under exchange of particle labels.

$$\mathcal{P}\Psi(\vec{r}_1,\ldots,\vec{r}_N) = \frac{1}{N!}\sum_P (-1)^P\Psi(\vec{r}_{P1},\ldots,\vec{r}_{PN}) \tag{1.10}$$

For example, for two fermions

$$\mathcal{P}\psi(\vec{r}_1,\vec{r}_2) = \frac{1}{2}\left[\psi(\vec{r}_1,\vec{r}_2) - \psi(\vec{r}_2,\vec{r}_1)\right]\,. \tag{1.11}$$

The antisymmetry requirement restricts the Hilbert space of N-particle fermionic wave functions to a smaller space $\mathcal{F}_N$. Since $\mathcal{P}$ is a projector [27, p.5], $\mathcal{F}_N$ is in fact the projection of the full Hilbert space $\mathcal{H}_N$ under $\mathcal{P}$:

$$\mathcal{F}_N = \mathcal{P}\,\mathcal{H}_N\,. \tag{1.12}$$

Using the operator $\mathcal{P}$, a system of fermions with one particle in state α_1, one particle in state α_2 and so on is represented as follows:

$$|\alpha_1\ldots\alpha_N\rangle \equiv \sqrt{N!}\mathcal{P}\,|\alpha_1\ldots\alpha_N) \tag{1.13}$$

$$= \frac{1}{\sqrt{N!}}\sum_P (-1)^P\,|\alpha_{P1}\rangle\otimes\cdots\otimes|\alpha_{PN}\rangle\,. \tag{1.14}$$

As expected, these antisymmetrized states obey the Pauli exclusion principle. Suppose two states are identical $|\alpha_1\rangle = |\alpha_2\rangle = |\alpha\rangle$ then

$$|\alpha_1\alpha_2\alpha_3\ldots\alpha_N\rangle = \sqrt{N!}\mathcal{P}\,|\alpha_1\alpha_2\alpha_3\ldots\alpha_N) \tag{1.15}$$

$$= -\sqrt{N!}\mathcal{P}\,|\alpha_2\alpha_1\alpha_3\ldots\alpha_N) \quad \text{swapping changes the parity cf. Eq. (1.14)} \tag{1.16}$$

$$= 0 \tag{1.17}$$

and no acceptable many-fermion state exists in this case.

From Eq. (1.12), it follows that if $|\alpha_1\ldots\alpha_N)$ is a basis of the Hilbert space $\mathcal{H}_N$ then $\mathcal{P}\,|\alpha_1\ldots\alpha_N)$ spans $\mathcal{F}_N$. Furthermore, if the basis $|\alpha\rangle$ is orthogonal in $\mathcal{H}$ then the basis $|\alpha_1\ldots\alpha_N)$ is orthogonal in $\mathcal{H}_N$ and the basis $|\alpha_1\ldots\alpha_N\rangle$ is orthogonal in $\mathcal{F}_N$. To see why, we evaluate the inner product of two such antisymmetric states.

$$\langle\alpha'_1\ldots\alpha'_N|\alpha_1\ldots\alpha_N\rangle = N!\,(\alpha'_1\ldots\alpha'_N|\mathcal{P}^2|\alpha_1\ldots\alpha_N) \tag{1.18}$$

$$= N!\,(\alpha'_1\ldots\alpha'_N|\mathcal{P}|\alpha_1\ldots\alpha_N) \qquad \text{since } \mathcal{P} \text{ is a projector} \tag{1.19}$$

$$= \sum_P (-1)^P \langle\alpha'_1|\alpha_{P1}\rangle\ldots\langle\alpha'_N|\alpha_{PN}\rangle \tag{1.20}$$

Because of the orthogonality of the basis $|\alpha\rangle$, the states $|\alpha'_i\rangle$ must be a permutation of the states $|\alpha_i\rangle$. Furthermore, there is at most one particle per state $|\alpha_i\rangle$, no two identical states can be present in the set $\{\alpha_1,\ldots,\alpha_N\}$. Thus there is at most one permutation P that transforms $\alpha_1,\ldots,\alpha_N$ into $\alpha'_1,\ldots,\alpha'_N$ and the sum in Eq. (1.20) reduces to a single term. Assuming the states $|\alpha_i\rangle$ are normalized, we obtain:

$$\langle\alpha'_1\ldots\alpha'_N|\alpha_1\ldots\alpha_N\rangle = (-1)^P\,, \tag{1.21}$$

which immediately gives us the norm of $|\alpha_1\ldots\alpha_N\rangle$:

$$\langle\alpha_1\ldots\alpha_N|\alpha_1\ldots\alpha_N\rangle = 1\,. \tag{1.22}$$

So these antisymmetric states are already normalized! Thus the completeness relation in $\mathcal{F}_N$ is

$$\frac{1}{N!}\sum_{\alpha_1,\ldots,\alpha_N} |\alpha_1\ldots\alpha_N\rangle\langle\alpha_1\ldots\alpha_N| = 1\,. \tag{1.23}$$

The overlap between a simple tensor product $|\beta_1\ldots\beta_N)$ and the antisymmetrized state $|\alpha_1\ldots\alpha_N\rangle$ is

$$(\beta_1\ldots\beta_N|\alpha_1\ldots\alpha_N\rangle = \frac{1}{\sqrt{N!}}\sum_P (-1)^P \langle\beta_1|\alpha_{P1}\rangle\langle\beta_2|\alpha_{P2}\rangle\ldots\langle\beta_N|\alpha_{PN}\rangle \tag{1.24}$$

$$= \frac{1}{\sqrt{N!}}\,\mathrm{Det}\,(\langle\beta_i|\alpha_j\rangle) \tag{1.25}$$

where the notation $\mathrm{Det}\,(\langle\beta_i|\alpha_j\rangle)$ denotes the determinant of a matrix whose ij-th entries are $\langle\beta_i|\alpha_j\rangle$. If we let $(\beta_1\ldots\beta_N| = (\vec{r}_1\ldots\vec{r}_N|$ in Eq. (1.24), we obtain the relationship between the normalized antisymmetric states $|\alpha_1\ldots\alpha_N\rangle$ and the single-particle wave functions $\phi_{\alpha_i}(\vec{x}_j)$:

$$\Psi_{\beta_1,\ldots,\beta_N}(\vec{r}_1,\ldots,\vec{r}_N) = (\vec{r}_1\ldots\vec{r}_N\,|\alpha_1\ldots\alpha_N\rangle \tag{1.26}$$

$$\Psi_{\beta_1,\ldots,\beta_N}(\vec{r}_1,\ldots,\vec{r}_N) = \frac{1}{\sqrt{N!}}\,\mathrm{Det}\,(\phi_{\alpha_i}(\vec{r}_j)) \tag{1.27}$$

where we call $\mathrm{Det}\left(\phi_{\alpha_i}(\vec{r}_j)\right)$ a *Slater determinant*. These Slater determinants form a basis for $\mathcal{F}_N$ in coordinate representation.

The overlap of two antisymmetric fermion states is

$$\langle\beta_1 \dots \beta_N|\alpha_1 \dots \alpha_N\rangle = \mathrm{Det}(\langle\beta_i|\alpha_j\rangle)\,. \tag{1.28}$$

1.1.2 Creation and annihilation operators

Now we introduce creation and annihilation operators, which provide a convenient representation of the many-body states in Section 1.1.1 and many-body operators in Section 1.1.3. For each single-particle state $|\lambda\rangle$ of the single-particle space $\mathcal{H}$, we define a creation operator $a^\dagger_\lambda$ by its action on any antisymmetrized state $|\lambda_1 \dots \lambda_N\rangle$ of $\mathcal{F}_N$ where $\{|\lambda_i\rangle\}$ is an orthonormal basis as follows:

$$a^\dagger_\lambda\,|\lambda_1 \dots \lambda_N\rangle = \begin{cases} |\lambda\lambda_1 \dots \lambda_N\rangle & \text{if the state } |\lambda\rangle \text{ is not present in } |\lambda_1 \dots \lambda_N\rangle \\ 0 & \text{if the state } |\lambda\rangle \text{ is present in } |\lambda_1 \dots \lambda_N\rangle \end{cases} \tag{1.29}$$

As the name suggests, the operator $a^\dagger_\lambda$ adds a particle in state $|\lambda\rangle$ to the state on which it operates and antisymmetrizes the new state. In addition to the many-particle states we have used so far, it is useful to define a state with zero particles or *vacuum state*, denoted $|0\rangle$. Note that the vacuum state $|0\rangle$, a state with no particles, is different from the zero of the Hilbert space. A natural extension of Eq. (1.29) is the definition that $a^\dagger_\lambda$ creates a particle in the state $|\lambda\rangle$ when acting on the vacuum $|0\rangle$:

$$a^\dagger_\lambda\,|0\rangle = |\lambda\rangle\ . \tag{1.30}$$

The creation operators $a^\dagger_\lambda$ do not operate within one space $\mathcal{F}_N$ but rather operate from a N-particle space $\mathcal{F}_N$ to a $(N+1)$-particle space $\mathcal{F}_{N+1}$. The Fock space is defined as the direct sum of the Fermion spaces

$$\mathcal{F} = \mathcal{F}_0 \oplus \mathcal{F}_1 \oplus \mathcal{F}_2 \oplus \cdots = \oplus_{n=0}^{\infty}\mathcal{F}_n \tag{1.31}$$

where by definition

$$\mathcal{F}_0 = |0\rangle \tag{1.32}$$

and

$$\mathcal{F}_1 = \mathcal{H}\,. \tag{1.33}$$

Any basis vector $|\lambda_1 \dots \lambda_N\rangle$ may be generated by repeated action of the creation operator on the vacuum $|0\rangle$.

$$|\lambda_1 \dots \lambda_N\rangle = a^\dagger_{\lambda_1} a^\dagger_{\lambda_2} \dots a^\dagger_{\lambda_N}\,|0\rangle \tag{1.34}$$

A general state $|\phi\rangle$ of the Fock space is a linear combination of states with any number of particles. For example, the state

$$|\phi\rangle = \frac{1}{2}\,|0\rangle + \frac{1}{\sqrt{2}}\,|\lambda\rangle + \frac{1}{2}\,|\lambda\mu\rangle \tag{1.35}$$

represents a system of particles in which there is probability 1/4 for having no particles, probability 1/2 for having one particle in state $|\lambda\rangle$ and a probability 1/4 for having two particles in state $|\lambda\mu\rangle$. The completeness relation in Fock space is

$$1 = |0\rangle\langle 0| + \sum_{N=1}^{\infty}\frac{1}{N!}\sum_{\lambda_1 \dots \lambda_N}|\lambda_1 \dots \lambda_N\rangle\langle\lambda_1 \dots \lambda_N|\ . \tag{1.36}$$

The antisymmetry of the many-fermion states impose anticommutation relations between the creation operators. For any state $|\lambda_1 \dots \lambda_N\rangle$ and any single particle states $|\lambda\rangle$ and $|\mu\rangle$, we obtain

$$a_\lambda^\dagger a_\mu^\dagger |\lambda_1 \dots \lambda_N\rangle = |\lambda\mu\lambda_1 \dots \lambda_N\rangle \tag{1.37}$$

$$= -|\mu\lambda\lambda_1 \dots \lambda_N\rangle \quad \text{antisymmetry of wave function} \tag{1.38}$$

$$= -a_\mu^\dagger a_\lambda^\dagger |\lambda_1 \dots \lambda_N\rangle \,. \tag{1.39}$$

Since the above equation holds for any state, the creation operators satisfy the following equation:

$$a_\lambda^\dagger a_\mu^\dagger + a_\mu^\dagger a_\lambda^\dagger = 0\,. \tag{1.40}$$

Defining the *anticommutator* of $a_\lambda^\dagger$ and $a_\mu^\dagger$ as

$$\{a_\lambda^\dagger, a_\mu^\dagger\} = a_\lambda^\dagger a_\mu^\dagger + a_\mu^\dagger a_\lambda^\dagger \tag{1.41}$$

we can rewrite Eq. (1.40) in a more compact form:

$$\{a_\lambda^\dagger, a_\mu^\dagger\} = 0\,. \tag{1.42}$$

Since the operators $a_\lambda^\dagger$ are not self-adjoint, we define the annihilation operators a_λ as the adjoints of the creation operators $a_\lambda^\dagger$. The anticommutation relation of annihilation operators follow from adjoint of Eq. (1.40):

$$a_\lambda a_\mu + a_\mu a_\lambda \equiv \{a_\lambda, a_\mu\} = 0 \tag{1.43}$$

The action of a_λ on a state can be shown [27, p. 13-4] to be

$$a_\lambda |\beta_1 \dots \beta_n\rangle = \begin{cases} (-1)^{i-1} |\beta_1 \dots \widehat{\beta_\lambda} \beta_n\rangle & \text{if } |\beta_\lambda\rangle \text{ is occupied} \\ 0 & \text{if } |\beta_\lambda\rangle \text{ is unoccupied} \end{cases} \tag{1.44}$$

where i is the number of single-particle states before β_λ and the notation $\widehat{\beta_\lambda}$ means β_λ has been removed. As expected, if there are no particles in state $|\lambda\rangle$ to be annihilated, a_λ yields zero.

It is also useful to define the particle number operator

$$\widehat{n}_\alpha = a_\alpha^\dagger a_\alpha\,. \tag{1.45}$$

Using Eqs. (1.29) and (1.44), we observe that the action of $\widehat{n}_\alpha$ on a state $|\phi\rangle$ is to count the number of particles in state $|\alpha\rangle$:

$$\widehat{n}_\alpha |\alpha_1 \dots \alpha_N\rangle = a_\alpha^\dagger a_\alpha |a_1 \dots a_N\rangle \tag{1.46}$$

$$= \sum_{i=1}^{N} (-1)^{i-1} \delta_{\alpha\alpha_i} a_\alpha^\dagger |\alpha_1 \dots \widehat{\alpha_i} \dots \alpha_N\rangle \tag{1.47}$$

$$= \left(\sum_{i=1}^{N} \delta_{\alpha\alpha_i} \right) |a_1 \dots \alpha_i \dots a_N\rangle \,, \tag{1.48}$$

where $\sum_{i=1}^{N} \delta_{\alpha\alpha_i}$ yields the number of particles in state $|\alpha\rangle$ in $|a_1 \dots a_N\rangle$.

1.1.3 Many-body operators

We now consider many-body operators in the canonical basis of $\mathcal{H}_N$ in Eq. (1.2). If $\widehat{O}$ is any operator in $\mathcal{F}_N$, the indistinguishability of fermions implies that $\widehat{O}$ is invariant under any permutation of the particles. Thus, for any state and any permutation P:

$$(\beta_{P1} \dots \beta_{PN} | \widehat{O} | \beta'_{P1} \dots \beta'_{PN}) = (\beta_1 \dots \beta_N | \widehat{O} | \beta'_1 \dots \beta'_N) \; . \tag{1.49}$$

An operator $\widehat{U}$ is a one-body operator if the action of $\widehat{U}$ on a state $|\alpha_1 \dots \alpha_N\rangle$ of N particles is the sum of the action of $\widehat{U}$ on each particle

$$\widehat{U} \, |\alpha_1 \dots \alpha_N\rangle = \sum_{i=1}^{N} \widehat{U}_i \, |\alpha_1 \dots \alpha_N\rangle \; , \tag{1.50}$$

where the operator $\widehat{U}_i$ only acts on the i-th particle. It can be proved [27, p. 9] that the action of a one-body operator is entirely defined by its matrix elements $\langle \alpha | \widehat{U} | \beta \rangle$ in the single-particle Hilbert space $\mathcal{H}$. We can express any one-body operator in terms of creation and annihilation operators:

$$\widehat{U} = \sum_{\alpha,\beta} \langle \alpha | \widehat{U} | \beta \rangle \, a^\dagger_\alpha a_\beta \tag{1.51}$$

Similarly, an operator $\widehat{V}$ is a two-body operator if the action of $\widehat{V}$ on a state $|\alpha_1 \dots \alpha_N\rangle$ of N particles is the sum of the action of $\widehat{V}$ on all distinct pairs of particles.

$$\widehat{V} \, |\alpha_1 \dots \alpha_N\rangle = \frac{1}{2} \sum_{i \neq j} \widehat{V}_{ij} \, |\alpha_1 \dots \alpha_N\rangle \tag{1.52}$$

where $\widehat{V}_{ij}$ only acts on particles i and j. The symmetry requirement in Eq. (1.49) requires that $\widehat{V}_{ij} = \widehat{V}_{ji}$. The action of a two-body operator is entirely defined by its matrix elements $\langle \lambda\mu | \widehat{V} | \nu\rho \rangle$ in the Hilbert space $\widehat{H}_2$ of two-particle systems. Any two-body operator can be expressed in terms of creation and annihilation operators:

$$\widehat{V} = \sum_{\lambda\mu\nu\rho} \langle \lambda\mu | \widehat{V} | \nu\rho \rangle \, a^\dagger_\lambda a^\dagger_\mu a_\rho a_\nu \, . \tag{1.53}$$

Do note the reversed order of ν and ρ in the matrix element compared to that in the annihilation operators.

To recap, in this section, we have given an overview of the notation of second quantization and how that relates to the familiar "first quantization" notation used in introductory quantum mechanics. From this point on, we will use $c^\dagger_i$ and c_i to denote creation and annihilation operators instead of $a^\dagger_i$ and a_i.

1.2 Non-interacting electrons

We will now use the second quantized notation developed in the previous section to study non-interacting electrons. This should be familiar to readers who have taken introductory quantum mechanics and should "ease" them into working with second quantization.

The non-interacting Hamiltonian represents the kinetic energy of electrons as they move around a lattice

$$\widehat{H} = -t \sum_{\sigma, \langle i,j \rangle} c^\dagger_{i\sigma} c_{j\sigma} \, . \tag{1.54}$$

where t is a parameter describing how easy it is for electrons to move between sites. The notation $\langle i, j\rangle$ denotes nearest-neighbor hopping. We will only deal with one-dimensional systems in this thesis.

Because the system is non-interacting, we can use the Fourier transform of the creation and annihilation operators to diagonalize the Hamiltonian. We define the momentum-space creation and annihilation operators to be:

$$c_{k\sigma}^\dagger = \frac{1}{\sqrt{N}} \sum_\ell \mathrm{e}^{\mathrm{i}k\ell} c_{\ell\sigma}^\dagger \tag{1.55}$$

where k are restricted to have discrete values $k_n = \frac{2\pi n}{N}$ because we impose periodic boundary conditions (and tacitly set the lattice spacing to be 1). Thus, inverting Eq. (1.55):

$$c_{\ell\sigma}^\dagger = \frac{1}{\sqrt{N}} \sum_k \mathrm{e}^{-\mathrm{i}k\ell} c_{k\sigma}^\dagger \tag{1.56}$$

and the corresponding annihilation operator is

$$c_{\ell\sigma} = \frac{1}{\sqrt{N}} \sum_k \mathrm{e}^{\mathrm{i}k\ell} c_{k\sigma} \,. \tag{1.57}$$

After the Fourier transform, the non-interacting Hamiltonian in Eq. (1.54) becomes

$$\widehat{H} = -t \sum_{\langle j\ell\rangle\sigma} c_{j\sigma}^\dagger c_{\ell\sigma} \tag{1.58}$$

$$= -t \sum_{\langle j\ell\rangle\sigma} \left[\frac{1}{\sqrt{N}} \sum_k \mathrm{e}^{-\mathrm{i}kj} c_{k\sigma}^\dagger\right] \left[\frac{1}{\sqrt{N}} \sum_{k'} \mathrm{e}^{\mathrm{i}k'\ell} c_{k'\sigma}\right] \quad \text{Eqs. (1.56) and (1.57)} \tag{1.59}$$

$$= -\frac{t}{N} \sum_{\langle j\ell\rangle\sigma} \sum_{kk'} \mathrm{e}^{-\mathrm{i}kj} \mathrm{e}^{\mathrm{i}k'\ell} c_{k\sigma}^\dagger c_{k'\sigma} \tag{1.60}$$

$$= -\frac{t}{N} \sum_{kk'\sigma} c_{k\sigma}^\dagger c_{k'\sigma} \sum_{\langle j\ell\rangle} \left(\mathrm{e}^{-\mathrm{i}kj} \mathrm{e}^{\mathrm{i}k'\ell}\right) \,. \tag{1.61}$$

Because j and l have to be neighboring sites, $j = \ell + 1$ or $j = \ell - 1$. Thus

$$\widehat{H} = -\frac{t}{N} \sum_{kk'\sigma} c_{k\sigma}^\dagger c_{k'\sigma} \sum_\ell \left[\mathrm{e}^{-\mathrm{i}k(\ell+1)} \mathrm{e}^{\mathrm{i}k'\ell} + \mathrm{e}^{-\mathrm{i}k(\ell-1)} \mathrm{e}^{\mathrm{i}k'\ell}\right] \tag{1.62}$$

$$= -\frac{t}{N} \sum_{kk'\sigma} c_{k\sigma}^\dagger c_{k'\sigma} \sum_\ell \left[\mathrm{e}^{\mathrm{i}\ell(k'-k)} \mathrm{e}^{-\mathrm{i}k} + \mathrm{e}^{\mathrm{i}\ell(k'-k)} \mathrm{e}^{\mathrm{i}k}\right] \tag{1.63}$$

$$= -t \sum_{kk'\sigma} c_{k\sigma}^\dagger c_{k'\sigma} \left(\mathrm{e}^{-\mathrm{i}k} + \mathrm{e}^{\mathrm{i}k}\right) \left[\frac{1}{N} \sum_\ell \mathrm{e}^{\mathrm{i}\ell(k'-k)}\right] \tag{1.64}$$

$$= -t \sum_{kk'\sigma} c_{k\sigma}^\dagger c_{k'\sigma} (2\cos k) \delta_{kk'} \quad \text{recognizing the Fourier transform of } \delta \tag{1.65}$$

$$= -2t \sum_{k\sigma} (\cos k) c_{k\sigma}^\dagger c_{k\sigma} \tag{1.66}$$

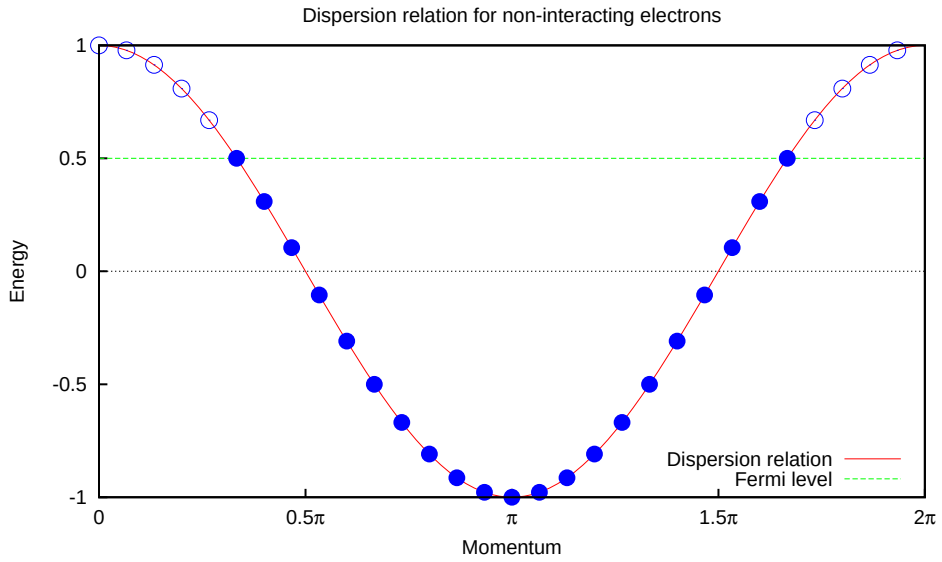


Figure 1.1: The dispersion relation for non-interacting electrons in Eq. (1.68). The solid circles denote filled states (up to the Fermi level) and empty circles denote unfilled states. The Fermi level shown here is arbitrary.

Thus, the non-interacting Hamiltonian in the momentum basis is

$$\widehat{H} \equiv \sum_k \varepsilon_k c_{k\sigma}^\dagger c_{k\sigma} \tag{1.67}$$

where we have defined

$$\varepsilon_k \equiv -2t \cos k \,. \tag{1.68}$$

In this basis, the Hamiltonian matrix is diagonal

$$H = \begin{pmatrix} \varepsilon_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \varepsilon_N \end{pmatrix} . \tag{1.69}$$

We note that Eq. (1.68), as shown in Fig. 1.1, describes the dispersion relation for non-interacting electrons if we require that electrons only move between adjacent nearest-neighboring sites. Other forms of dispersion relations also exist. For example, the dispersion relation for free electrons in continuum with $\hbar = m_\mathrm{e} = 1$ is

$$\varepsilon_k = \frac{k^2}{2} \,. \tag{1.70}$$

1.2.1 Equal-time Green's function

We will now take the first step towards finding the unequal-time Green's function by studying a special case of it: the equal-time Green's function. The operator appearing in the

equal-time Green's function is

$$c_{j\sigma}c^\dagger_{\ell\sigma} = \left(\frac{1}{\sqrt{N}}\sum_k \mathrm{e}^{\mathrm{i}kj}c_{k\sigma}\right)\left(\frac{1}{\sqrt{N}}\sum_{k'} \mathrm{e}^{-\mathrm{i}k'\ell}c^\dagger_{k'\sigma}\right) \tag{1.71}$$

$$= \frac{1}{N}\sum_{k,k'} \mathrm{e}^{\mathrm{i}kj}\mathrm{e}^{-\mathrm{i}k'\ell}c_{k\sigma}c^\dagger_{k'\sigma}\,. \tag{1.72}$$

The expectation for this operator in the ground-state is

$$\langle\Psi_0|c_{j\sigma}c^\dagger_{\ell\sigma}|\Psi_0\rangle = \frac{1}{N}\sum_{k,k'} \mathrm{e}^{\mathrm{i}kj}\mathrm{e}^{-\mathrm{i}k'\ell}\,\langle\Psi_0|c_{k\sigma}c^\dagger_{k'\sigma}|\Psi_0\rangle\,. \tag{1.73}$$

Using the commutation relations of the creation and annihilation operators and the normalization of $|\Psi_0\rangle$, we have

$$\langle\Psi_0|c_{j\sigma}c^\dagger_{\ell\sigma}|\Psi_0\rangle = \frac{1}{N}\sum_{kk'}\left[\mathrm{e}^{\mathrm{i}kj}\mathrm{e}^{-\mathrm{i}k'\ell}\delta_{kk'} - \mathrm{e}^{\mathrm{i}kj}\mathrm{e}^{-\mathrm{i}k'\ell}\,\langle\Psi_0|c^\dagger_{k'\sigma}c_{k\sigma}|\Psi_0\rangle\right]\,. \tag{1.74}$$

Because $|\Psi_0\rangle$ is occupied up to the Fermi level k_F as shown in Fig. 1.1, the annihilation operators $c_{k\sigma}$ with $|k| > k_\mathrm{F}$ acting on $|\Psi_0\rangle$ will yield zero because there are no states with $|k| > k_\mathrm{F}$ in $|\Psi_0\rangle$.

$$\langle\Psi_0|c_{j\sigma}c^\dagger_{\ell\sigma}|\Psi_0\rangle = \frac{1}{N}\left[\sum_k \mathrm{e}^{\mathrm{i}k(j-\ell)} - \sum_{|k|<k_\mathrm{F}} \mathrm{e}^{\mathrm{i}k(j-\ell)}\right] \tag{1.75}$$

$$= \frac{1}{N}\sum_{|k|>k_\mathrm{F}} \mathrm{e}^{\mathrm{i}k(j-\ell)}\,. \tag{1.76}$$

Suppose the lattice is one dimensional with an odd number of sites N and odd number of electrons N_σ with spin σ. Recall that the periodic boundary condition implies that $k_n = \frac{2\pi n}{N}$ where n ranges from $-\frac{N-1}{2}$ to $\frac{N-1}{2}$ i.e. the momentum states are symmetric about $k = 0$. Since the Fermi level is at $n = \frac{N_\sigma-1}{2}$, the k values of unoccupied momentum states satisfy $|n| \geq \frac{N_\sigma+1}{2}$. Thus

$$\langle\Psi_0|c_{j\sigma}c^\dagger_{\ell\sigma}|\Psi_0\rangle = \frac{1}{N}\sum_{n=\frac{N_\sigma+1}{2}}^{\frac{N-1}{2}}\left[\mathrm{e}^{\mathrm{i}\frac{2\pi n}{N}(j-\ell)} + \mathrm{e}^{-\mathrm{i}\frac{2\pi n}{N}(j-\ell)}\right] \tag{1.77}$$

$$= \frac{2}{N}\sum_{n=\frac{N_\sigma+1}{2}}^{\frac{N-1}{2}} \cos\left[\frac{2\pi n}{N}(j-\ell)\right]\,. \tag{1.78}$$

This is the expression for the equal-time Green's function in a non-interacting electron system.

1.2.2 Unequal-time Green's function

The unequal-time Green's function is a generalization of the equal-time Green's function we studied in the previous section. It is formed from the following operator

$$c_\ell(\tau)c_j^\dagger \tag{1.79}$$

where the creation and annihilation operators are in the Heisenberg picture i.e. they evolve in time:

$$c_\ell(t) = \mathrm{e}^{\mathrm{i}Ht}c_\ell(0)\,\mathrm{e}^{-\mathrm{i}Ht} \tag{1.80}$$
$$c_j^\dagger(t) = \mathrm{e}^{\mathrm{i}Ht}c_j^\dagger(0)\,\mathrm{e}^{-\mathrm{i}Ht}\,. \tag{1.81}$$

Note that in real time, we have $[c_\ell(t)]^\dagger = c_\ell^\dagger(t)$.

Now let $\tau = \mathrm{i}t$ be the imaginary time. Then the operator for the unequal-time Green's function becomes $c_\ell(\tau)c_j^\dagger(0)$. In imaginary time, Eqs. (1.80) and (1.81) become

$$c_\ell(\tau) = \mathrm{e}^{\tau\widehat{H}}c_\ell(0)\,\mathrm{e}^{-\tau\widehat{H}} \tag{1.82}$$
$$c_j^\dagger(\tau) = \mathrm{e}^{\tau\widehat{H}}c_j^\dagger(0)\,\mathrm{e}^{-\tau\widehat{H}} \tag{1.83}$$

Note that in imaginary time, $[c_\ell(\tau)]^\dagger \neq c_\ell^\dagger(\tau)$. From now on we will use the convention that if an operator has no time attached to it, it is understood to be at time 0. The ground state expectation value of the operator $c_\ell(\tau)c_j^\dagger$ is

$$\langle c_\ell(\tau)c_j^\dagger\rangle = \langle\Psi_0|\mathrm{e}^{\tau\widehat{H}}c_\ell\,\mathrm{e}^{-\tau\widehat{H}}c_j^\dagger|\Psi_0\rangle\ . \tag{1.84}$$

Since $|\Psi_0\rangle$ is the ground state, acting on the left with $\mathrm{e}^{\tau\widehat{H}}$ gives

$$\langle c_\ell(\tau)c_j^\dagger\rangle = \langle\Psi_0|\mathrm{e}^{\tau E_0}c_\ell\,\mathrm{e}^{-\tau\widehat{H}}c_j^\dagger|\Psi_0\rangle \tag{1.85}$$

and Fourier transforming $c_j^\dagger$ according to Eq. (1.56) gives

$$\langle c_\ell(\tau)c_j^\dagger\rangle = \left\langle\Psi_0\left|\,\mathrm{e}^{\tau E_0}c_\ell\,\mathrm{e}^{-\tau\widehat{H}}\left(\frac{1}{\sqrt{N}}\sum_k \mathrm{e}^{-\mathrm{i}kj}c_{k\sigma}^\dagger\right)\right|\Psi_0\right\rangle\ . \tag{1.86}$$

Because $|\Psi_0\rangle$ is the non-interacting ground state, it is filled up to the Fermi level k_F. Thus all creation operators below the Fermi level yield zero when acting on $|\Psi_0\rangle$:

$$\langle c_\ell(\tau)c_j^\dagger\rangle = \frac{1}{\sqrt{N}}\left\langle\Psi_0\left|\,\mathrm{e}^{\tau E_0}c_\ell\,\mathrm{e}^{-\tau\widehat{H}}\sum_{k>k_\mathrm{F}} \mathrm{e}^{-\mathrm{i}kj}c_{k\sigma}^\dagger\right|\Psi_0\right\rangle\ . \tag{1.87}$$

If we let $c_{k\sigma}$ for $k > k_\mathrm{F}$ act on the ground state $|\Psi_0\rangle$, the resulting state, which (for lack of a better notation) we shall denote $|\Psi_0 + k\rangle$, is the ground state plus a particle in the momentum state k.

$$\langle c_\ell(\tau) c_j^\dagger \rangle = \frac{1}{\sqrt{N}} \left\langle \Psi_0 \left| \mathrm{e}^{\tau E_0} c_\ell \, \mathrm{e}^{-\tau \widehat{H}} \sum_{k>k_\mathrm{F}} \mathrm{e}^{-\mathrm{i}kj} \right| \Psi_0 + k \right\rangle \tag{1.88}$$

$$= \frac{1}{\sqrt{N}} \left\langle \Psi_0 \left| \mathrm{e}^{\tau E_0} c_\ell \sum_{k>k_\mathrm{F}} \mathrm{e}^{-\mathrm{i}kj} \mathrm{e}^{-\tau \widehat{H}} \right| \Psi_0 + k \right\rangle . \qquad \mathrm{e}^{-\mathrm{i}kj} \text{ is a scalar} \tag{1.89}$$

Because the non-interacting Hamiltonian is diagonal in the momentum basis from Eq. (1.67), the energy of $|\Psi_0 + k\rangle$ is that of the ground state plus ε_k, the energy of the momentum state k. Thus acting on $|\Psi_0 + k\rangle$ with $\mathrm{e}^{-\tau \widehat{H}}$ gives

$$\langle c_\ell(\tau) c_j^\dagger \rangle = \frac{1}{\sqrt{N}} \left\langle \Psi_0 \left| \mathrm{e}^{\tau E_0} c_\ell \sum_{k>k_\mathrm{F}} \mathrm{e}^{-\mathrm{i}kj} \mathrm{e}^{-\tau(E_0+\varepsilon_k)} \right| \Psi_0 + k \right\rangle \tag{1.90}$$

$$= \frac{1}{\sqrt{N}} \left\langle \Psi_0 \left| c_\ell \sum_{k>k_\mathrm{F}} \mathrm{e}^{-\mathrm{i}kj} \mathrm{e}^{\tau E_0} \mathrm{e}^{-\tau(E_0+\varepsilon_k)} \right| \Psi_0 + k \right\rangle \tag{1.91}$$

$$= \frac{1}{\sqrt{N}} \left\langle \Psi_0 \left| c_\ell \sum_{k>k_\mathrm{F}} \mathrm{e}^{-\mathrm{i}kj} \mathrm{e}^{-\tau\varepsilon_k} \right| \Psi_0 + k \right\rangle . \tag{1.92}$$

Fourier transforming c_ℓ yields:

$$\langle c_\ell(\tau) c_j^\dagger \rangle = \frac{1}{\sqrt{N}} \left\langle \Psi_0 \left| \left(\frac{1}{\sqrt{N}} \sum_{k'} \mathrm{e}^{\mathrm{i}k'\ell} c_{k'\sigma} \right) \sum_{k>k_\mathrm{F}} \mathrm{e}^{-\mathrm{i}kj} \mathrm{e}^{-\tau\varepsilon_k} \right| \Psi_0 + k \right\rangle \tag{1.93}$$

$$= \frac{1}{N} \left\langle \Psi_0 \left| \sum_{k>k_\mathrm{F}} \sum_{k'} \mathrm{e}^{\mathrm{i}k'\ell} \mathrm{e}^{-\mathrm{i}kj} \mathrm{e}^{-\tau\varepsilon_k} c_{k'\sigma} \right| \Psi_0 + k \right\rangle . \tag{1.94}$$

Once again, because $|\Psi_0 + k\rangle$ is the ground state plus a particle in the momentum state k, $c_{k'\sigma}$ must satisfy either $|k| \leq k_\mathrm{F}$ or $k' = k$. Since k already satisfy $|k| > k_\mathrm{F}$, it must be that $k = k'$. Thus

$$\langle c_\ell(\tau) c_j^\dagger \rangle = \frac{1}{N} \left\langle \Psi_0 \left| \sum_{k>k_\mathrm{F}} \mathrm{e}^{\mathrm{i}k\ell} \mathrm{e}^{-\mathrm{i}kj} \mathrm{e}^{-\tau\varepsilon_k} c_{k\sigma} \right| \Psi_0 + k \right\rangle \tag{1.95}$$

$$= \frac{1}{N} \left\langle \Psi_0 \left| \sum_{k>k_\mathrm{F}} \mathrm{e}^{\mathrm{i}k(\ell-j)} \mathrm{e}^{-\tau\varepsilon_k} \right| \Psi_0 \right\rangle \tag{1.96}$$

$$= \frac{1}{N} \sum_{k>k_\mathrm{F}} \mathrm{e}^{\mathrm{i}k(\ell-j)} \mathrm{e}^{-\tau\varepsilon_k} \langle \Psi_0 | \Psi_0 \rangle \tag{1.97}$$

$$\langle c_\ell(\tau) c_j^\dagger \rangle = \frac{1}{N} \sum_{k>k_\mathrm{F}} \mathrm{e}^{\mathrm{i}k(\ell-j)} \mathrm{e}^{-\tau\varepsilon_k} \tag{1.98}$$

where ε_k is defined as in Eq. (1.68):

$$\varepsilon_k = -2t\cos k\,. \tag{1.99}$$

Assuming periodic boundary condition and an odd number of sites N and spin-σ electrons N_σ, by the same reasoning that leads to Eq. (1.77), Eq. (1.98) becomes

$$\langle c_\ell(\tau)c_j^\dagger\rangle = \frac{2}{N}\sum_{n=\frac{N_\sigma+1}{2}}^{\frac{N-1}{2}} \cos\left[\frac{2\pi n}{N}(j-\ell)\right]\exp\left[2\tau t\cos\left(\frac{2\pi n}{N}\right)\right] \tag{1.100}$$

assuming that the ground state is normalized. Note that this reduces to the equal-time Green's function in Eq. (1.76) when $\tau = 0$ as expected.

1.2.3 Non-interacting Hubbard model in an external potential

We will now increase the complexity of the non-interacting system in the previous sections by including an external potential (Note that the electrons still do not directly interact with each other). This is useful because in chapter 3, we will use the Hubbard-Stratonovich transformation to break up an interacting systems into smaller non-interacting systems living in external potentials, precisely the type of system we study in this section (although the potential is different).

To the non-interacting Hamiltonian in Eq. (1.54) we added an extra term to describe an external potential

$$\widehat{H} = -t\sum_{\sigma,\langle i,j\rangle} c_{i\sigma}^\dagger c_{j\sigma} + \sum_\ell V_\ell c_\ell^\dagger c_\ell\,, \tag{1.101}$$

where V_ℓ describes the potential on site ℓ.

Suppose we have the the ground state $|\Psi_0\rangle$ of Eq. (1.101). We want to find the unequal-time Green's function for this ground state. Let $\tau = it$ be the imaginary time, we have

$$\langle c_{\ell\sigma}(\tau)c_{j\sigma}^\dagger\rangle = \langle\Psi_0|\mathrm{e}^{\tau\widehat{H}}c_{\ell\sigma}\,\mathrm{e}^{-\tau\widehat{H}}c_{j\sigma}^\dagger|\Psi_0\rangle \tag{1.102}$$

$$= \mathrm{e}^{\tau E_0}\,\langle\Psi_0|c_{\ell\sigma}\mathrm{e}^{-\tau\widehat{H}}c_{j\sigma}^\dagger|\Psi_0\rangle\ . \tag{1.103}$$

where E_0 is the ground state energy i.e. the sum of all the filled energy levels. Now let's switch to the eigenbasis of the Hamiltonian operator. Suppose $\widehat{H}$ can be diagonalized as

$$\widehat{H} = \vec{c}^\dagger H\vec{c} \tag{1.104}$$

$$= \underbrace{\vec{c}^\dagger(Q}_{\vec{d}^\dagger}\,D\,\underbrace{Q^\dagger)\vec{c}}_{\vec{d}} \tag{1.105}$$

where D is a diagonal matrix containing the energy eigenvalues E_i (in increasing order i.e. the lowest energy eigenvalue is at the the top left corner of D) and the columns of Q are the eigenvectors of H. This equation also defines new creation and annihilation operators $d^\dagger$ and d in this new basis. These are related to the one old ones by

$$\vec{d}^\dagger = \vec{c}^\dagger Q \implies \vec{c}^\dagger = \vec{d}^\dagger Q^\dagger \tag{1.106}$$

$$\vec{d} = Q^\dagger\vec{c} \implies \vec{c} = Q\vec{d}\,. \tag{1.107}$$

Element-wise, the above two equation give the relationship between individual old and new creation and annihilation operators:

$$c^\dagger_{j\sigma} = \sum_n d^\dagger_{n\sigma}(Q^\dagger)_{nj} = \sum_n Q^*_{jn} d^\dagger_{n\sigma} \tag{1.108}$$

$$c_{\ell\sigma} = \sum_m Q_{\ell m} d_{m\sigma} \,. \tag{1.109}$$

In this new basis, Eq. (1.103) becomes

$$\langle c_{\ell\sigma}(\tau) c^\dagger_{j\sigma}\rangle = \mathrm{e}^{\tau E_0} \sum_{mn} (Q^\dagger)_{nj} Q_{\ell m} \langle \Psi_0 | d_{m\sigma} \mathrm{e}^{-\tau \widehat{H}} d^\dagger_{n\sigma} | \Psi_0\rangle \,. \tag{1.110}$$

Since $d^\dagger_{n\sigma}$ creates a particle in the n-th eigenstate of $\widehat{H}$, $d^\dagger_{n\sigma}$ acting on $|\Psi_0\rangle$ adds a particle in the n-th eigenstate to the ground state $|\Psi_0\rangle$. Let's denote this state by $|\Psi_0 + n\rangle$. We note that because of the Pauli exclusion principle, n must be an unfilled state i.e. $n > N_\sigma$ if we start numbering n from 1. Similarly, $d_{m\sigma}$ acting to the right on $|\Psi_0\rangle$ gives $|\Psi_0 + m\rangle$ for $m > N_\sigma$

$$\langle c_\ell(\tau) c^\dagger_j\rangle = \mathrm{e}^{\tau E_0} \sum_{m,n>N_\sigma}^{N} (Q^\dagger)_{nj} Q_{\ell m} \langle \Psi_0 + m | \mathrm{e}^{-\tau \widehat{H}} | \Psi_0 + n\rangle \tag{1.111}$$

Acting on the left with $\mathrm{e}^{-\tau\widehat{H}}$ and using orthonormality between $|\Psi_0 + m\rangle$ and $|\Psi_0 + n\rangle$ give

$$\langle c_\ell(\tau) c^\dagger_j\rangle = \mathrm{e}^{\tau E_0} \sum_{m,n>N_\sigma} (Q^\dagger)_{nj} Q_{\ell m} \mathrm{e}^{-\tau(E_0+E_n)} \delta_{mn} \tag{1.112}$$

$$= \sum_{n>N_\sigma} Q_{\ell n} \mathrm{e}^{-\tau E_n} (Q^\dagger)_{nj} \tag{1.113}$$

$$= \left[\mathcal{Q}\mathcal{E}\mathcal{Q}^\dagger\right]_{\ell j} \tag{1.114}$$

where

- $\mathcal{Q}^\dagger$ is the bottom $(N - N_\sigma) \times N$ block of $Q^\dagger$.
- $\mathcal{E}$ is the bottom right $(N - N_\sigma) \times (N - N_\sigma)$ block of the diagonal matrix whose diagonal elements are $\mathrm{e}^{-\tau E_i}$ (E_i are the energy eigenvalues).

We will use this result in Section 3.11.1 as a benchmark to understand the numerical behavior of the extension to CPMC that we will develop in this thesis.

Chapter 2

The Hubbard Model

The Hubbard model is named after Hubbard [7] who introduced the Hamiltonian in order to model electronic correlations in narrow energy bands and proposed a number of approximate treatments of the associated many-body problem. In this chapter we will introduce the Hamiltonian of the one-band one-dimensional Hubbard model, the principal subject of this thesis. We will then discuss a number of useful properties of this Hamiltonian and state the exact diagonalization (ED) algorithm to obtain the unequal-time Green's function of the model.

2.1 The Hubbard Hamiltonian

Here we only give a brief introduction to the Hubbard Hamiltonian because the topic has been so extensively covered in the literature. For an excellent and comprehensive derivation, see Essler *et al.* [8, pp. 1-5].

The setting for the one-band one-dimensional Hubbard model is a lattice of M sites. To each lattice site i , we associate a *Wannier wave function*[1] $|\phi_i\rangle = \phi(x - R_i)$ centered on site i that serves as an orbital which can accommodates a maximum of two electrons of opposite spin. Given this setting, the Hubbard Hamiltonian is

$$H = -t \sum_{\langle i,j \rangle} (c^\dagger_{i,a} c_{j,a} + \text{H.c.}) + U \sum_i n_{i\uparrow} n_{i\downarrow} \tag{2.1}$$

where the symbol $\langle i, j \rangle$ denotes summation over ordered pairs of nearest neighbors.

The coefficient $t > 0$ is called the *hopping amplitude* and is defined as

$$t = t_{ij} = \int \mathrm{d}x \, \phi^*(x - R_i) \left(-\frac{\hbar^2 \nabla^2}{2m} \right) \phi(x - R_j) \tag{2.2}$$

for all pairs i and j. The assumption that hopping occurs only between nearest neighbors is justified because the Wannier wave functions are highly localized, making the above integral negligible between sites beyond nearest neighbors.

The coefficient $U > 0$ is called the Coulomb repulsion strength and is defined as

$$U = U_{ii} = \int \mathrm{d}x \, \mathrm{d}x' \; \phi^*_\uparrow(x - R_i) \phi^*_\downarrow(x - R_i) V(|x - x'|) \phi_\downarrow(x - R_i) \phi_\uparrow(x - R_i) \tag{2.3}$$

[1] A Wannier wave function is a Fourier transform of the Bloch wave functions. See Ashcroft and Mermin [28, pp. 187-9].

for any site i where V is the Coulomb potential:

$$V(|x-x'|) = \frac{\mathrm{e}^{-|x-x'|s}}{|x-x'|} \tag{2.4}$$

in which s is called the screening length.

The model only has two parameters: the strength of the interaction U/t and the electron density $(N_\uparrow + N_\downarrow)/M$. In this thesis we will use t as the unit of energy and set $t = 1$.

2.2 Some properties of the Hubbard Hamiltonian

2.2.1 Conservation of particle numbers

One apparent property of the Hubbard Hamiltonian is that it separately conserves the number of particles of spin up and spin down i.e. $\widehat{H}$ commutes with $n_\uparrow = \sum_i n_{i\uparrow}$ and $n_\downarrow = \sum_i n_{i\downarrow}$.

It's easy to see that these number operators commute with the potential energy term because

$$[n_{i\sigma}, n_{j\sigma'}] = 0 \tag{2.5}$$

for all sites i, j and spins σ, σ'. It's also clear that every term in the kinetic energy operator conserves the total number of electrons of a particular spin because $c^\dagger_{i\sigma}c_{j\sigma}$ simply moves an electron of spin-σ from site j to site i without destroying or creating any electrons.

Because the Hubbard Hamiltonian commutes with the total number operator of each spin, we can diagonalize the Hamiltonian for fixed numbers of spin-up and spin-down. In other words, we can choose to work in a sector with a fixed quantum number S_z.[2]

2.2.2 Independence of the spin sectors in a Slater determinant

In this section, we will show that when dealing with Slater determinants in the Hubbard model, we can treat the spin-up and spin-down sectors independently. This will become useful in chapter 3.

Because the Hubbard model does not flip the spins of electrons, the expectation of any one-body operators that contains fermion operators of different spins (e.g. $c^\dagger_{i\uparrow}c_{j\downarrow}$) will be zero.[3] This is quite an important observation whose consequence is that the only operators we have to actually evaluate will have the form $c^\dagger_{i\sigma}c_{j\sigma}$. Consider the effect [2, p. 116] of a term like $c^\dagger_{i\downarrow}c_{j\downarrow}$ on a typical state $|\phi\rangle$ given by

$$|\phi\rangle = [c^\dagger_{j_1\uparrow}\dots c^\dagger_{j_M\uparrow}][c^\dagger_{i_1\downarrow}\dots c^\dagger_{i_N\downarrow}]\,|0\rangle \tag{2.6}$$

The action of the operator gives

$$c^\dagger_{i\downarrow}c_{j\downarrow}\,|\phi\rangle = c^\dagger_{i\downarrow}c_{j\downarrow}[c^\dagger_{j_1\uparrow}\dots c_{j_M\uparrow}][c_{i_1\downarrow}\dots c_{i_N\downarrow}]\,|0\rangle \tag{2.7}$$

$$= [c^\dagger_{j_1\uparrow}\dots c_{j_M\uparrow}](-1)^M(-1)^M c^\dagger_{i\downarrow}c_{j\downarrow}[c_{i_1\downarrow}\dots c_{i_N\downarrow}]\,|0\rangle \tag{2.8}$$

$$= [c^\dagger_{j_1\uparrow}\dots c_{j_M\uparrow}]c^\dagger_{i\downarrow}c_{j\downarrow}[c_{i_1\downarrow}\dots c_{i_N\downarrow}]\,|0\rangle \tag{2.9}$$

[2] The quantum number S_z is one of the eigenvalues of the operator $\widehat{S}_z = \frac{1}{2}\sum_i(n_{i\uparrow} - n_{i\downarrow})$ where we have set $\hbar = 1$ by convention.

[3] We only need to show this is true for any energy eigenstate $|\Psi\rangle$ (because they form a basis for the many-particle Hilbert space) which are also eigenstates of $n_\uparrow$. The state $c^\dagger_{i\uparrow}c_{j\downarrow}\,|\Psi\rangle$ will have one more spin up electron than $|\Psi\rangle$. By basic linear algebra, these two states are necessarily orthogonal i.e. $\langle\Psi|c^\dagger_{i\uparrow}c_{j\downarrow}|\Psi\rangle = 0$ because they are eigenvectors corresponding to different eigenvalues of the operator $n_\uparrow$.

where the two factors of $(-1)^M$ arise when we "move" the $c^\dagger_{i\downarrow}c_{j\downarrow}$ operator through the spin-$\uparrow$ creation operators. Thus the operator $c^\dagger_{i\downarrow}c_{j\downarrow}$ effectively only acts on the spin-down sector. It is easy to see that applying $c^\dagger_{i\uparrow}c_{j\uparrow}$ will only affect the spin-up sector.

2.3 Exact Diagonalization

To benchmark the CPMC extension in chapter 3, we have written a computer program in MATLAB to calculate the imaginary-time Green's function exactly for the one-dimensional Hubbard model with periodic boundary condition. The program takes as arguments the hopping amplitude `t`, the Coulomb repulsion `U`, the number of lattice sites `N_sites`, the number of spin-up and spin-down electrons `N_up` and `N_dn`, and the imaginary time `tau`. The program returns two `N_sites` $\times$ `N_sites` matrices `G_up` and `G_dn` whose (i, j)-th element is the imaginary-time Green's function $\langle c_{i\sigma}(\tau)c^\dagger_{j\sigma}(0)\rangle$. The program uses three "helper" functions:

`generateBasis.m` generates the basis for the Hilbert space.

1. Calculate the total number of basis states in the spin-up and spin-down sector separately. These are given by $\binom{\texttt{N\_sites}}{\texttt{N\_up}}$ and $\binom{\texttt{N\_sites}}{\texttt{N\_dn}}$, respectively.
2. For each spin sector, form all possible basis states of that sector in binary representation where 1 denotes an occupied lattice site and 0 denotes an unoccupied site. For example, 01010 in the spin-up sector corresponds to a 5-site lattice in which sites 2 and 4 are each occupied by a spin-up electron and the other sites are empty and the site numbering increases from left to right.
3. Form the basis for the full Hilbert space by forming all combinations of all spin-up basis states and spin-down basis states. Thus, this full basis consists of $\binom{\texttt{N\_sites}}{\texttt{N\_up}} \times \binom{\texttt{N\_sites}}{\texttt{N\_dn}}$ states.
4. Associate each basis state with a decimal number by converting the binary representation to decimal. This is to facilitate lookup of basis states later.

`hubbardHamiltonian.m` creates the Hubbard Hamiltonian.

1. Create two square matrices, each of size $\binom{\texttt{N\_sites}}{\texttt{N\_up}}\binom{\texttt{N\_sites}}{\texttt{N\_dn}}$ to hold the kinetic and potential Hamiltonians.
2. To form the potential Hamiltonian, loop through each basis state j and use the `bit_and` operation to count the number of sites that are occupied in both the spin-up and spin-down sectors and multiply the total by `U`. This is the value of the (j, j)-th element of the potential Hamiltonian.
3. To form the kinetic Hamiltonian, loop through each basis state m:
   (a) In the spin-up sector, for every occupied site, move the electron on that site to the sites to its left and its right if those are not already occupied. Look up the resulting state in the full basis of the Hilbert space. If the index of this state is i, set the (i, j)-th element of the kinetic Hamiltonian to
   - $+$`t` if the hopping was around the lattice boundary and the total number of spin-up electrons is even.

- $-$`t` otherwise.

(b) Repeat step 3a for the spin down sector.

4. Add the kinetic and potential Hamiltonians to get the total Hamiltonian.

`creationOperator.m` creates the matrix for the creation operator $c^\dagger_{i\sigma}$. We assume $\sigma =\uparrow$ for this description. The procedure for $\sigma =\downarrow$ is very similar.

1. Call `generateBasis.m` twice to generate the full basis for the "source" Hilbert space (for `N_up` spin-up and `N_dn` spin down electrons) and the "target" Hilbert space (for (`N_up` + 1) spin-up and `N_dn` spin-down electrons).
2. Create a non-square matrix of dimension $\binom{\texttt{N\_sites}}{\texttt{N\_up}+1}\binom{\texttt{N\_sites}}{\texttt{N\_dn}}$-by-$\binom{\texttt{N\_sites}}{\texttt{N\_up}}\binom{\texttt{N\_sites}}{\texttt{N\_dn}}$ to hold the resulting matrix of the creation operator.
3. Loop through each basis state m in the "source" Hilbert space:
   (a) If site i is not already occupied by a spin-up electron, create a spin-up electron on that site and look up the resulting state in the basis of the "target" Hilbert space. Let n be the index of this resultant state in the "target" basis.
   (b) Set the (n, m)-th matrix element to 1 if the number of occupied sites from site 1 through site $(i-1)$ is even; otherwise, set the (n, m)-th element to -1. This step is identical for $\sigma =\downarrow$ except that we set the (n, m)-th element of the matrix to 1 if the sum of the number of occupied sites from site 1 through site $(i-1)$ *and* `N_up` is even and set it to -1 otherwise.

With these three functions in place, we are ready to compute the unequal-time Green's function $\langle c_{i\uparrow}(\tau)c^\dagger_{j\uparrow}\rangle$ (the procedure for $\langle c_{i\downarrow}(\tau)c^\dagger_{j\downarrow}\rangle$ is very similar):

1. Create the Hubbard Hamiltonian `firstHamiltonian` for the Hilbert space of `N_up` spin-up and `N_dn` spin-down electrons.
   (a) Diagonalize `firstHamiltonian` to get the lowest energy eigenvector which we call `groundState`.
   (b) Use matrix exponentiation to get the matrix `expmFirstHamiltonian`, defined as
   $$\texttt{expmFirstHamiltonian} = \mathrm{e}^{\texttt{tau}\times I\times\texttt{firstHamiltonian}} \tag{2.10}$$
   where I is the identity matrix.
2. Create the matrix `destructionMatrix` by invoking the function `creationOperator.m` to create the matrix for the operator $c^\dagger_{i\uparrow}$.
3. Create the matrix `creationMatrix` by invoking the function `creationOperator.m` to create the matrix for the operator $c^\dagger_{j\uparrow}$.
4. Create the Hubbard Hamiltonian `secondHamiltonian` for the Hilbert space of (`N_up`+1) spin-up and `N_dn` spin-down electrons.
   (a) Use matrix exponentiation to get the matrix `expmSecondHamiltonian`, defined as
   $$\texttt{expmSecondHamiltonian} = \mathrm{e}^{-\texttt{tau}\times I\times\texttt{secondHamiltonian}} \tag{2.11}$$
   where I is the identity matrix.

5. Evaluate $\langle c_{i\uparrow}(\tau)c^{\dagger}_{j\uparrow}\rangle$ by computing the following matrix multiplication:

$$\begin{aligned}\langle c_{i\uparrow}(\tau)c^{\dagger}_{j\uparrow}\rangle = \texttt{groundState}^{\dagger} \times \texttt{expmFirstHamiltonian}\times \\ \times \texttt{destructionMatrix}^{\dagger} \times \texttt{expmSecondHamiltonian} \\ \times \texttt{creationMatrix} \times \texttt{groundState}\end{aligned} \tag{2.12}$$

where † in the above expression denotes the Hermitian conjugate of a matrix or vector.

We tested to make sure that the results of this program agrees to within numerical precision (10^{-8} to 10^{-9})with Eq. (1.100) when $\texttt{U} = 0$ i.e. for non-interacting electrons. We do note that for large enough values of τ (about $\tau > 4$), this ED algorithm is a little numerically unstable. Although we did not have enough time to investigate the reason, we suspect that it comes from matrix exponentiation. Because we only need $\tau \leq 2$ in this thesis, this numerical issue is not a problem.

We also note that the current ED implementation is not the most optimized because here we have only used the $\widehat{S}_z$ symmetry of the Hubbard Hamiltonian. However, because the Hamiltonian also commutes with the total spin $\widehat{\boldsymbol{S}}^2$, we can reduce the size of the matrix to be diagonalized by a change of basis to make the Hamiltonian block-diagonal where each block corresponds to a fixed eigenvalue of $\widehat{\boldsymbol{S}}^2$. For details, see Jafari [2, p. 118-9].

Chapter 3

Constrained Path Monte Carlo

This chapter is the main chapter of this thesis where we discuss the implementation of the constrained path Monte Carlo (CPMC) method. The ground-state CPMC algorithm has two main components. The *first* component is the formulation of the ground state projection as an open-ended importance-sampled random walk. This random walk takes place in Slater determinant space and uses importance sampling to increase the efficiency. It is exact but suffers from the sign problem. The *second* component is the constraint of the paths of the random walk so that any Slater determinant generated during the random walk maintains an appropriate overlap with a known trial wave function $|\psi_{\mathrm{T}}\rangle$. This constraint eliminates the sign decay, making the CPMC method scale algebraically instead of exponentially, but introduces a systematic error in the algorithm. These two components are independent of each other and can be used separately. We call the combination of these two components the ground-state CPMC algorithm.

This thesis only deals with the application of CPMC to the Hubbard model. This chapter has adapted materials from Zhang *et al.* [6], Nguyen *et al.* [25], and Purwanto and Zhang [29].

3.1 Slater determinant space

Here we define the notations and representations that will be used throughout this chapter. Here we assume a one-particle basis has been chosen.

- M: the number of single-electron basis states i.e. the number of lattice sites in the Hubbard model.
- $|\chi_i\rangle$: the i-th single-particle basis state ($i = 1, 2, \ldots, M$). In this implementation of CPMC, $|\chi_i\rangle$ is the Wannier wave function localized on the i-th lattice site.
- $c_i^\dagger$ and c_i: creation and annihilation operators for an electron in $|\chi_i\rangle$. $n_i \equiv c_i^\dagger c_i$ is the corresponding number operator.
- N: the number of electrons. N_σ is the number of electrons with spin σ ($\sigma =\uparrow$ or $\downarrow$). As expected, $N_\sigma \leq M$.
- φ: a single-particle orbital. The coefficients in the expansion

$$|\varphi\rangle = \sum_i \varphi_i |\chi_i\rangle = \sum_i c_i^\dagger \varphi_i |0\rangle \tag{3.1}$$

in the single particle basis $\{|\chi_i\rangle\}$ can be conveniently expressed as an M-dimensional vector:

$$\begin{pmatrix} \varphi_1 \\ \varphi_2 \\ \vdots \\ \varphi_M \end{pmatrix} . \tag{3.2}$$

- $|\phi\rangle$: a many-body wave function which can be written as a Slater determinant. Given N different single-particle orbitals (for N electrons), we form a many-body wave function from their antisymmetrized product

$$|\phi\rangle \equiv \widehat{\varphi}_1^\dagger \widehat{\varphi}_2^\dagger \cdots \widehat{\varphi}_N^\dagger |0\rangle \tag{3.3}$$

where the operator

$$\widehat{\varphi}_m^\dagger \equiv \sum_i c_i^\dagger \, \varphi_{i,m} \tag{3.4}$$

creates an electron in the m-th single-particle orbital described in Eq. (3.2).

- Φ: an $M \times N$ matrix which represents the Slater determinant $|\phi\rangle$ in Eq. (3.3):

$$\Phi \equiv \begin{pmatrix} \varphi_{1,1} & \varphi_{1,2} & \cdots & \varphi_{1,N} \\ \varphi_{2,1} & \varphi_{2,2} & \cdots & \varphi_{2,N} \\ \vdots & \vdots & & \vdots \\ \varphi_{M,1} & \varphi_{M,2} & \cdots & \varphi_{M,N} \end{pmatrix} . \tag{3.5}$$

Each column of this matrix is an M-dimensional vector and represents a single-particle orbital described by Eq. (3.2). If the real-space coordinates of the electrons are $R = \{\mathbf{r}_1, \mathbf{r}_2, \cdots, \mathbf{r}_N\}$, the configuration space representation of Eq. (3.3) is

$$\langle R|\phi\rangle = \phi(R) = \mathrm{Det} \begin{pmatrix} \varphi_1(\mathbf{r}_1) & \varphi_1(\mathbf{r}_2) & \cdots & \varphi_1(\mathbf{r}_N) \\ \varphi_2(\mathbf{r}_1) & \varphi_1(\mathbf{r}_2) & \cdots & \varphi_2(\mathbf{r}_N) \\ \vdots & \vdots & \ddots & \vdots \\ \varphi_N(\mathbf{r}_1) & \varphi_N(\mathbf{r}_2) & \cdots & \varphi_N(\mathbf{r}_N) \end{pmatrix} \tag{3.6}$$

where

$$\varphi_m(\mathbf{r}) = \sum_i \varphi_{i,m} \chi_i(\mathbf{r}) \, . \tag{3.7}$$

This is essentially Eq. (1.27). This is the more familiar form of a Slater determinant encountered in introductory quantum mechanics.

It is very important to understand the difference between the representation of a Slater determinant in second quantization (Eq. (3.5)) and in first quantization (Eq. (3.6)) even though they represent the same physical object. In first quantization, the Slater determinant is the actual determinant of a square matrix made up of all possible combination of electrons and orbitals. Taking the determinant ensures the anti-symmetry of the wave function, as required by the spin-statistics theorem.

In second quantization, although we refer to Φ as a Slater "determinant," we actually *never* take the determinant of Φ because it is *not* a square matrix. The reason is that

this matrix simply contains the *coefficients* for the creation operators and that the algebra obeyed by these creation operators guarantees the antisymmetry of the wave function.

- $|\Psi\rangle$ (upper case): a many-body wave function which is not necessarily a single Slater determinant, e.g. the many-body ground state $|\Psi_0\rangle$ in Section 3.3.

We list several useful properties of a Slater determinant in the representation of Eq. (3.5). The first property is that for any two non-orthogonal Slater determinants, $|\phi\rangle$ and $|\phi'\rangle$, it can be shown that their overlap integral is a number:

$$\langle\phi|\phi'\rangle = \mathrm{Det}\left(\Phi^\dagger\Phi'\right), \tag{3.8}$$

where $\Phi^\dagger$ is the Hermitian conjugate of the matrix Φ. The proof of Eq. (3.8) is in Appendix C.

The second property is known as *Thouless theorem* because it was first discovered (although not in the present form) by Thouless [30] and restated in the present form by Hamann and Fahy [31]. It says that the operation on any Slater determinant in Eq. (3.3) by the exponential of a one body operator

$$\widehat{B} = \exp\left(\sum_{ij}^{M} c_i^\dagger U_{ij} c_j\right) \tag{3.9}$$

simply leads to another Slater determinant:

$$\widehat{B}\,|\phi\rangle = \widehat{\phi}_1'^\dagger \widehat{\phi}_2'^\dagger \cdots \widehat{\phi}_M'^\dagger\,|0\rangle \equiv |\phi'\rangle \tag{3.10}$$

where the operators

$$\widehat{\phi}_m'^\dagger = \sum_j c_j^\dagger\,\Phi'_{jm} \tag{3.11}$$

create single-particle orbitals just like Eq. (3.4) and the matrix Φ' is related to Φ in a simple way:

$$\Phi' \equiv \mathrm{e}^U \Phi \tag{3.12}$$

where the matrix U is formed from elements U_{ij} in Eq. (3.9).

Since $B \equiv \mathrm{e}^U$ is an $M \times M$ square matrix, the operation of $\widehat{B}$ on $|\phi\rangle$ simply involves multiplying e^U, an $M \times M$ matrix, to Φ, an $M \times N$ matrix. The proof of Eqs. (3.10) to (3.12) is in Appendix D.

As shown in Section 2.2.2, the spin-up and spin-down sectors of a Slater determinant in the Hubbard model are independent. Thus, it is convenient to represent each Slater determinant as two independent spin parts:[1]

$$|\phi\rangle = |\phi^\uparrow\rangle \otimes |\phi^\downarrow\rangle\ . \tag{3.13}$$

The corresponding matrix representation is

$$\Phi = \Phi^\uparrow \otimes \Phi^\downarrow \tag{3.14}$$

[1] Note that the direct product sign $\otimes$ that we use here is somewhat an abuse of notation that is used widely in the literature. Its precise meaning in this context will be explained at the end of the current section.

where $\Phi^{\uparrow}$ and $\Phi^{\downarrow}$ have dimensions $M \times N_{\uparrow}$ and $M \times N_{\downarrow}$, respectively. Because $N_{\sigma} \leq M$, the matrices representing each spin sector are "tall" rather than "wide." The overlap between any two Slater determinants is simply the product of the overlaps of individual spin determinants:

$$\langle \phi | \phi' \rangle = \prod_{\sigma=\uparrow,\downarrow} \langle \phi^{\sigma} | \phi'^{\sigma} \rangle = \mathrm{Det}\left[(\Phi^{\uparrow})^{\dagger} \Phi'^{\uparrow} \right] \cdot \mathrm{Det}\left[(\Phi^{\downarrow})^{\dagger} \Phi'^{\downarrow} \right] . \tag{3.15}$$

Any operator $\widehat{B}$ described by Eq. (3.9) acts independently on the two spin parts:

$$\widehat{B} \, |\phi\rangle = \widehat{B}^{\uparrow} \, |\phi^{\uparrow}\rangle \otimes \widehat{B}^{\downarrow} \, |\phi^{\downarrow}\rangle \, . \tag{3.16}$$

Each of the spin components of $\widehat{B}$ can be represented as an $M \times M$ matrix:

$$B = B^{\uparrow} \otimes B^{\downarrow} \, . \tag{3.17}$$

Applying $\widehat{B}$ to a Slater determinant simply involves matrix multiplications for the $\uparrow$ and $\downarrow$ components separately, leading to another Slater determinant $|\phi'\rangle$ as in Eq. (3.10) i.e. the result is

$$B^{\uparrow} \Phi^{\uparrow} \otimes B^{\downarrow} \Phi^{\downarrow} \, . \tag{3.18}$$

Unless specified, the spin components of $\widehat{B}$ are identical i.e. $B^{\uparrow} = B^{\downarrow}$ (note the absence of a hat on B to denote the matrix of the operator $\widehat{B}$).

Now we will explain why we can separate the spin up and spin down sectors of operators (Eq. (3.17)) and Slater determinants (Eq. (3.14)). Suppose we use a basis consisting of $2M$ states where the first M states correspond to spin-up Wannier orbitals and the last M basis states the spin-down Wannier orbitals. Because the spins are not flipped by the kinetic Hamiltonian, the matrix of $\widehat{K}$ in this basis is block-diagonal

$$K = \left(\begin{array}{c|c} K^{\uparrow} & O \\ \hline O & K^{\downarrow} \end{array} \right) \tag{3.19}$$

where $K^{\uparrow}$ and $K^{\downarrow}$ are identical and O is the zero matrix. Each of these submatrices are square matrices of size $M \times M$. The matrix of a Slater determinant with more spin-up electrons than spin-down electrons has the form

$$\Phi = \left(\begin{array}{c|c} \Phi^{\uparrow} & O \\ \hline O & \Phi^{\downarrow} \end{array} \right) \tag{3.20}$$

where $\Phi^{\uparrow}$ is of size $M \times N_{\uparrow}$ and $\Phi^{\downarrow}$ is of size $M \times N_{\downarrow}$. They are the same matrices that appear in Eq. (3.14). Note that because we have chosen $N_{\uparrow} > N_{\downarrow}$ for illustration purpose, $\Phi^{\uparrow}$ is wider (has more columns) than $\Phi^{\downarrow}$. It can easily be seen that the application of the operator $\widehat{K}$ in Eq. (3.19) to the Slater determinant in Eq. (3.20) retains the block-diagonal structure of Φ. As will be shown later in Eq. (3.40), the matrix of the potential Hamiltonian $\widehat{V}$ also has this block-diagonal structure. Thus we can carry around the two spin parts of Slater determinants and operators separately.

3.2 The Hubbard Hamiltonian

Recall from Section 2.1 that the one-band Hubbard model is a simple paradigm of a system of interacting electrons. Its Hamiltonian is

$$\widehat{H} = \widehat{K} + \widehat{V} = -t \sum_{\langle ij \rangle \sigma} (c^\dagger_{i\sigma} c_{j\sigma} + c^\dagger_{j\sigma} c_{i\sigma}) + U \sum_i n_{i\uparrow} n_{i\downarrow}, \tag{3.21}$$

where t is the hopping matrix element, and $c^\dagger_{i\sigma}$ and $c_{i\sigma}$ are electron creation and destruction operators, respectively, of spin σ on site i. The Hamiltonian is defined on a lattice of dimension d with $M = \prod_d L_d$ (although this thesis only deals with the model in one dimension). The lattice sites serve as the basis functions here i.e. $|\chi_i\rangle = |i\rangle$ denotes a Wannier wave function localized on the site labeled by i. The notation $\langle\ \rangle$ in Eq. (3.21) indicates nearest-neighbors. The on-site Coulomb repulsion is $U > 0$. The model only has two parameters: the strength of the interaction U/t and the electron density $(N_\uparrow + N_\downarrow)/M$. In this paper we will use t as the unit of energy and set $t = 1$.

The difference between the Hubbard Hamiltonian and a general electronic Hamiltonian is in the structure of the matrix elements in $\widehat{K}$ and $\widehat{V}$. In the latter, $\widehat{K}$ is specified by hopping integrals of the form K_{ij}, while $\widehat{V}$ is specified by Coulomb matrix elements of the form $V_{ijk\ell}$, with i, j, k, ℓ in general running from 1 to M. In terms of the CPMC method, the structure of $\widehat{K}$ makes essentially no difference. The structure of $\widehat{V}$, however, dictates the form of the Hubbard-Stratonovich transformation in Section 3.5.

3.3 Ground-state projection

The ground-state wave function $|\Psi_0\rangle$ can be obtained from any trial wave function $|\Psi_\mathrm{T}\rangle$—provided that $|\Psi_\mathrm{T}\rangle$ is not orthogonal to $|\Psi_0\rangle$ and the ground-state is non-degenerate—by repeated applications of the ground-state projection operator

$$\mathcal{P}_\mathrm{gs} = \mathrm{e}^{-\Delta\tau(\widehat{H} - E_\mathrm{T})}, \tag{3.22}$$

where E_T is the best estimate of the ground-state energy. That is, if the wave function at the imaginary time τ is $|\Psi^{(\tau)}\rangle$, the wave function at time $\tau + \Delta\tau$ is given by

$$|\Psi^{(\tau+\Delta\tau)}\rangle = \mathrm{e}^{-\Delta\tau(\widehat{H} - E_\mathrm{T})} |\Psi^{(\tau)}\rangle\ . \tag{3.23}$$

To see why repeated application of $\mathcal{P}_\mathrm{gs}$ in Eq. (3.22) gives the ground-state $|\Psi_0\rangle$, suppose we apply $\mathcal{P}_\mathrm{gs}$ n times to a trial wave function $|\Psi_\mathrm{T}\rangle$ that can be expanded in an eigenbasis $\{|\Psi_j\rangle\}$ of $\widehat{H}$ with coefficients a_j:

$$|\Psi_\mathrm{T}\rangle = \sum_{j=0} a_j |\Psi_j\rangle\ . \tag{3.24}$$

It is easy to see [32] that the application of $\mathrm{e}^{-\tau(\widehat{H} - E_\mathrm{T})}$ where $\tau = n\Delta\tau$ gives

$$\mathrm{e}^{-\tau(\widehat{H} - E_\mathrm{T})} |\Psi_\mathrm{T}\rangle = \sum_{j=0} a_j\, \mathrm{e}^{-\tau(E_j - E_\mathrm{T})} |\Psi_j\rangle \tag{3.25}$$

$$= a_0\, \mathrm{e}^{-\tau(E_0 - E_\mathrm{T})} |\Psi_0\rangle + \sum_{j=1} a_j\, \mathrm{e}^{-\tau(E_j - E_\mathrm{T})} |\Psi_j\rangle\ , \tag{3.26}$$

where we have assumed the non-degeneracy of the lowest energy eigenstate in the last equality. Because E_0 is the smallest of all the E_j, one can infer from Eq. (3.26) the asymptotic behavior of $\mathrm{e}^{-\tau(\widehat{H}-E_\mathrm{T})}\,|\Psi_\mathrm{T}\rangle$ as $\tau \to \infty$ (i.e. $n \to \infty$):

- If $E_\mathrm{T} > E_0$, $\lim_{\tau\to\infty} \mathrm{e}^{-\tau(\widehat{H}-E_\mathrm{T})}\,|\Psi_\mathrm{T}\rangle = \infty$: the wave function diverges exponentially fast.
- If $E_\mathrm{T} < E_0$, $\lim_{\tau\to\infty} \mathrm{e}^{-\tau(\widehat{H}-E_\mathrm{T})}\,|\Psi_\mathrm{T}\rangle = 0$: the wave function vanishes exponentially fast.
- If $E_\mathrm{T} = E_0$, $\lim_{\tau\to\infty} \mathrm{e}^{-\tau(\widehat{H}-E_\mathrm{T})}\,|\Psi_\mathrm{T}\rangle = a_0\,|\Psi_0\rangle$: the wave function converges to the ground state, up to a constant factor $a_0 = \langle\Psi_0|\Psi_\mathrm{T}\rangle$.

This behavior is the basis of the ground-state projection. For $E_\mathrm{T} = E_0$, any trial wave function $|\Psi_\mathrm{T}\rangle$ will converge to the ground state $|\Psi_0\rangle$ as long as there is a numerically significant overlap between them because the excited states are attenuated exponentially. In practice, this overlap is always large enough because $|\Psi_\mathrm{T}\rangle$ is always the ground-state obtained from calculations with non-interacting systems (which is what we use here) or calculations with mean-field theory (in more advanced implementations). These calculations are numerically cheap and simple.

E_T in this projection process essentially plays the role of a normalizing factor. Because $\mathcal{P}_\mathrm{gs}$ is related to the usual quantum mechanical time-evolution operator $\mathrm{e}^{-\mathrm{i}t\widehat{H}}$ by the identification $\tau = \mathrm{i}t$, $\mathcal{P}_\mathrm{gs}$ is also called the imaginary-time propagator. Two essential ingredients are needed in order to evaluate $\mathcal{P}_\mathrm{gs}$ within reasonable computing time: the Suzuki-Trotter decomposition and the Hubbard-Stratonovich transformation.

3.4 Suzuki-Trotter decomposition

For a general Hamiltonian $\widehat{H} = \widehat{K} + \widehat{V}$ where $\widehat{K}$ is a one-body operator and $\widehat{V}$ a two-body operator (specifically, the kinetic and potential operator in the Hubbard model). The Suzuki-Trotter decomposition allows us to treat $\widehat{K}$ and $\widehat{V}$ separately at the cost of an error proportional to the commutator of $\widehat{K}$ and $\widehat{V}$. The first-order non-symmetric form is

$$\mathrm{e}^{-\Delta\tau\widehat{H}} = \mathrm{e}^{-\Delta\tau(\widehat{K}+\widehat{V})} = \mathrm{e}^{-\Delta\tau\widehat{K}}\mathrm{e}^{-\Delta\tau\widehat{V}} + \mathcal{O}(\Delta\tau^2)[\widehat{K},\widehat{V}]\,. \tag{3.27}$$

We can reduce the time step error by using the second-order symmetric version of the decomposition:

$$\mathrm{e}^{-\Delta\tau\widehat{H}} = \mathrm{e}^{-\Delta\tau(\widehat{K}+\widehat{V})} = \mathrm{e}^{-\Delta\tau\widehat{K}/2}\mathrm{e}^{-\Delta\tau\widehat{V}}\mathrm{e}^{-\Delta\tau\widehat{K}/2} + \varepsilon\,. \tag{3.28}$$

The error ε, quoting from Linden [33, p. 68], is

$$\varepsilon \leq (\Delta\tau)^3 \frac{|[\widehat{K},[\widehat{K},\widehat{V}]]| + 2|[\widehat{V},[\widehat{K},\widehat{V}]]|}{24}\,. \tag{3.29}$$

The important thing to note is that the error per time step is proportional to $(\Delta\tau)^3$. Suppose the desired total length of projection is $\tau = n\Delta\tau$ then the error accumulated due to this decomposition along the entire projection interval τ is:

$$\left(\frac{\tau}{n}\right)^3 n = \left(\frac{\tau}{n}\right)^2 \tau = \tau(\Delta\tau)^2 = \mathcal{O}(\Delta\tau^2)\,. \tag{3.30}$$

The residual Trotter error can be removed by, for example, extrapolation with several independent runs of sufficiently small $\Delta\tau$ values. Eq. (3.28) is the version we will use in this implementation.

3.5 The Hubbard-Stratonovich transformation

In Eq. (3.28), the kinetic energy propagator

$$\widehat{B}_{\mathrm{K}/2} \equiv \mathrm{e}^{-\Delta\tau\widehat{K}/2} \tag{3.31}$$

has the same form as Eq. (3.9). However, the potential energy propagator $\mathrm{e}^{-\Delta\tau\widehat{V}}$ does not. In general, a Hubbard-Stratonovich (HS) transformation can be employed to transform $\mathrm{e}^{-\Delta\tau\widehat{V}}$ into the desired form (see Appendix A for more details). In the repulsive Hubbard model, we can use the following special form of the HS transformation, called the *discrete Hirsch spin transformation*, first discovered by Hirsch [18]:

$$e^{-\Delta\tau U n_{i\uparrow} n_{i\downarrow}} = \mathrm{e}^{-\Delta\tau U(n_{i\uparrow}+n_{i\downarrow})/2} \sum_{x_i=\pm 1} p(x_i)\, \mathrm{e}^{\gamma x_i (n_{i\uparrow}-n_{i\downarrow})}\,, \tag{3.32}$$

where γ is given by $\cosh(\gamma) = \exp(\Delta\tau U/2)$. We interpret $p(x_i) = 1/2$ as a discrete probability density function (PDF) with $x_i = \pm 1$. This transformation is proved in Appendix B.

In Eq. (3.32), the exponent on the left, which comes from the interaction term $\widehat{V}$ on the i-th site, is quadratic in n, indicating the interaction of *two* electrons. The exponents on the right, on the other hand, are linear in n, indicating two non-interacting electrons in a common external field characterized by x_i. Thus an interacting system has been converted into a *non-interacting* system living in fluctuating external *auxiliary fields* x_i, and the summation over all such auxiliary-field configurations recovers the many-body interactions. The linearized operator on the right hand side in Eq. (3.32) is the spin $(n_{i\uparrow} - n_{i\downarrow})$ on each site, hence the name of the transformation.

We note in passing that there exist other ways to do the HS transformation (see Appendix A) and that different forms of the HS transformation can have different efficiencies in different situations.

Since we represent a Slater determinant as individual spin determinants in Eq. (3.13), it is convenient to spin-factorize Eq. (3.32) as

$$\mathrm{e}^{-\Delta\tau U n_{i\uparrow} n_{i\downarrow}} = \mathrm{e}^{-\Delta\tau U(n_{i\uparrow}+n_{i\downarrow})/2} \sum_{x_i=\pm 1} p(x_i) \left[\widehat{b}^{\uparrow}_{V}(x_i) \otimes \widehat{b}^{\downarrow}_{V}(x_i)\right] \tag{3.33}$$

where the spin-dependent operator $\widehat{b}^{\sigma}_{V}(x_i)$ on the i-th lattice site is defined as

$$\widehat{b}^{\sigma}_{\mathrm{V}}(x_i) = \mathrm{e}^{s(\sigma)\,\gamma\, x_i c^{\dagger}_{i\sigma} c_{i\sigma}} \tag{3.34}$$

and $s(\uparrow) = +1$ and $s(\downarrow) = -1$. It is easy to see that $\widehat{b}^{\sigma}_{\mathrm{V}}$ is diagonal in the eigenbasis of the operator N_σ. The related operator $\widehat{b}_{\mathrm{V}}(x_i)$ (i.e. without σ) includes both the spin up and spin down parts:

$$\widehat{b}_{\mathrm{V}}(x_i) = \widehat{b}^{\uparrow}_{\mathrm{V}}(x_i) \otimes \widehat{b}^{\downarrow}_{\mathrm{V}}(x_i)\,. \tag{3.35}$$

Below we will use the corresponding symbol without hat (b_{V}) to denote the matrix representation of the operator ($\widehat{b}_{\mathrm{V}}$) associated with that symbol.

The potential energy propagator $\mathrm{e}^{-\Delta\tau\widehat{V}}$ over all sites can easily be seen to be the product of the propagators $\mathrm{e}^{-\Delta\tau U n_{i\uparrow} n_{i\downarrow}}$ over each site:

$$\mathrm{e}^{-\Delta\tau\widehat{V}} = \prod_i \mathrm{e}^{-\Delta\tau U n_{i\uparrow} n_{i\downarrow}} \tag{3.36}$$

$$\mathrm{e}^{-\Delta\tau\widehat{V}} = \mathrm{e}^{-\Delta\tau U(N_\uparrow+N_\downarrow)/2} \sum_{\vec{x}} P(\vec{x}) \prod_{\sigma=\uparrow,\downarrow} \widehat{B}_V^\sigma(\vec{x}) \tag{3.37}$$

where $\vec{x} = \{x_1, \ldots, x_M\}$ is one configuration of auxiliary fields over all M sites and

$$\widehat{B}_V^\sigma(\vec{x}) = \prod_i \widehat{b}_V^\sigma(x_i) \tag{3.38}$$

is the $\vec{x}$-dependent product of the spin-σ propagators over all sites. The overall PDF in Eq. (3.37) is

$$P(\vec{x}) = \prod_i p(x_i) = \left(\frac{1}{2}\right)^M \tag{3.39}$$

to be distinguished from the PDF $p(x_i)$ for one individual auxiliary field x_i in Eq. (3.33).

We now make a quick digression to discuss the structure of $B_V^\sigma(\vec{x})$. In the basis of $2M$ space-spin states (see the discussion leading to Eqs. (3.19) and (3.20)), the matrix for $B_V^\sigma(\vec{x})$ is block-diagonal:

$$B_V^\sigma(\vec{x}) = \left(\begin{array}{c|c} \prod_i \widehat{b}^\uparrow(x_i) & O \\ \hline O & \prod_i \widehat{b}^\downarrow(x_i) \end{array}\right) \tag{3.40}$$

because the operators $\widehat{b}_{\mathrm{V}}^\sigma(x_i)$ as defined in Eq. (3.34) do not flip the spins of electrons.

With Eq. (3.37) in hand, now the projection operator in Eq. (3.22) can be expressed entirely in terms of one-body operators of the form in Eq. (3.9):

$$\mathcal{P}_{\mathrm{gs}} = \mathrm{e}^{\Delta\tau[E_{\mathrm{T}}-U(N_\uparrow+N_\downarrow)/2]} \sum_{\vec{x}} P(\vec{x}) \prod_{\sigma=\uparrow,\downarrow} \widehat{B}_{\mathrm{K}/2}^\sigma \widehat{B}_{\mathrm{V}}^\sigma(\vec{x}) \widehat{B}_{\mathrm{K}/2}^\sigma \,. \tag{3.41}$$

As noted in Eq. (3.16), $B_{\mathrm{K}/2}$ has an $\uparrow$ and a $\downarrow$ component, each of which is an $M \times M$ matrix. Applying each $\widehat{B}_{\mathrm{K}/2}$ to a Slater determinant $|\phi\rangle$ simply involves matrix multiplications with the matrix $B_{\mathrm{K}/2}$ for the $\uparrow$ and $\downarrow$ components of Φ separately, leading to another Slater determinant $|\phi'\rangle$ as in Eq. (3.10).

As it stands now, we can sample configurations $\vec{x}$ by flipping a coin M times and pick the fields accordingly. As discussed later, this simple sampling scheme leads to a very inefficient random walk.

3.6 A toy model for illustration

Let us take, for example, a simple one-dimensional four-site Hubbard model with $N_\uparrow = 2$, $N_\downarrow = 1$ and open boundary condition. The sites are numbered sequentially. First let us examine the trivial case of free electrons i.e. $U = 0$. We can write down the *one-electron*

Hamiltonian matrix, which is of dimension 4×4:

$$H = \begin{pmatrix} 0 & -1 & 0 & 0 \\ -1 & 0 & -1 & 0 \\ 0 & -1 & 0 & -1 \\ 0 & 0 & -1 & 0 \end{pmatrix}. \tag{3.42}$$

Direct diagonalization gives us the eigenstates of H from which we immediately obtain the matrix Φ_0 for the ground-state wave function $|\psi_0\rangle$:

$$\Phi_0 = \begin{pmatrix} 0.3717 & -0.6015 \\ 0.6015 & -0.3717 \\ 0.6015 & 0.3717 \\ 0.3717 & 0.6015 \end{pmatrix} \otimes \begin{pmatrix} 0.3717 \\ 0.6015 \\ 0.6015 \\ 0.3717 \end{pmatrix} \tag{3.43}$$

where the first matrix contains two single-particle orbitals (two columns) for the two $\uparrow$ electrons and the second matrix contains one single-electron orbital for the one $\downarrow$ electron. Each single-electron orbital is an eigenvector of $\widehat{H}$ in Eq. (3.42). The matrix Φ_0 represents $|\phi_0\rangle$ in the same way that Eq. (3.5) represents Eq. (3.3) and is the trial wave function used in our CPMC implementation.

Next we consider the interacting problem, with $U > 0$. Applying the HS transformation of Eq. (3.32) to Eq. (3.28), the projection operator is now

$$\mathcal{P}_{\text{gs}} = \mathrm{e}^{\Delta\tau[E_{\text{T}} - U(N_\uparrow + N_\downarrow)/2]} \sum_{\vec{x}} P(\vec{x}) \left[B_{\text{K}/2} \cdot \begin{pmatrix} \mathrm{e}^{\gamma x_1} & 0 & 0 & 0 \\ 0 & \mathrm{e}^{\gamma x_2} & 0 & 0 \\ 0 & 0 & \mathrm{e}^{\gamma x_3} & 0 \\ 0 & 0 & 0 & \mathrm{e}^{\gamma x_4} \end{pmatrix} \cdot B_{\text{K}/2} \otimes \right.$$
$$\left. \otimes B_{\text{K}/2} \cdot \begin{pmatrix} \mathrm{e}^{-\gamma x_1} & 0 & 0 & 0 \\ 0 & \mathrm{e}^{-\gamma x_2} & 0 & 0 \\ 0 & 0 & \mathrm{e}^{-\gamma x_3} & 0 \\ 0 & 0 & 0 & \mathrm{e}^{-\gamma x_4} \end{pmatrix} \cdot B_{\text{K}/2} \right] \tag{3.44}$$

where $\vec{x} = \{x_1, x_2, x_3, x_4\}$ and $P(\vec{x}) = \left(\frac{1}{2}\right)^4$ and $B_{\text{K}/2}$ is the exponential of the matrix

$$-\frac{\Delta\tau}{2} \begin{pmatrix} 0 & -1 & 0 & 0 \\ -1 & 0 & -1 & 0 \\ 0 & -1 & 0 & -1 \\ 0 & 0 & -1 & 0 \end{pmatrix}. \tag{3.45}$$

Eq. (3.44) is just Eq. (3.41) specialized to a four-site lattice.

3.7 Random walk in Slater determinant space

The first component of the CPMC algorithm is the reformulation of the projection process as branching, open-ended random walks in Slater determinant space. As shown in Eq. (1.27), any antisymmetric wave function can be written as a linear combination of Slater determinants i.e.

$$|\Psi\rangle = \sum_{\phi} \lambda_\Psi(\phi)\, |\phi\rangle \tag{3.46}$$

where the sum is over each member of the Slater determinant basis which is *nonorthogonal* and *overcomplete*.

In CPMC, we represent the wave function at each stage by a finite ensemble of Slater determinants, i.e.

$$|\Psi^{(\tau)}\rangle = \sum_{\mathfrak{j}=1}^{N_{\mathrm{wlkr}}} |\phi_{\mathfrak{j}}^{(\tau)}\rangle \ . \tag{3.47}$$

where the "=" sign denotes equality in the limit of infinite population. Because the random walk is essentially a dynamical process which evolves in imaginary time τ, the parenthesized superscript (τ) differentiates the "generations" of these walkers along the imaginary time axis. The subscript $\mathfrak{j}$ indexes the Slater determinants[2] with the same τ and an overall normalization factor of the wave function has been omitted. These Slater determinants will be referred to as *random walkers* because they evolve according to a random walk.

At any stage of the algorithm, the sum in Eq. (3.47) will be over only part of the basis because, as shown later, the determinants in this sum are statistical samples whose distribution represents the linear coefficients $\lambda_\Psi(\phi)$ in Eq. (3.46). Eq. (3.47) will serve as the definition of the Monte Carlo representation of a wave function in the CPMC method. In this implementation, we will start form an initial ensemble $|\Psi^{(0)}\rangle$ in which all walkers are identical to a single trial Slater determinant $|\phi_{\mathrm{T}}\rangle$ i.e. $|\phi_{\mathfrak{j}}^{(0)}\rangle = |\phi_{\mathrm{T}}\rangle$ for all $\mathfrak{j}$.

Applying the HS-transformed propagator in Eq. (3.41) to one projection step in Eq. (3.23) gives the evolution of a single Slater determinant $|\phi_{\mathfrak{j}}^{(\tau)}\rangle$ in the ensemble in Eq. (3.47):

$$|\phi_{\mathfrak{j}}^{(\tau+\Delta\tau)}\rangle = \mathrm{e}^{\Delta\tau[E_{\mathrm{T}}-U(N_\uparrow+N_\downarrow)/2]} \sum_{\vec{x}} P(\vec{x}) \left[\widehat{B}_{\mathrm{K}/2}\widehat{B}_{\mathrm{V}}(\vec{x})\widehat{B}_{\mathrm{K}/2}\right] |\phi_{\mathfrak{j}}^{(\tau)}\rangle \tag{3.48}$$

where we have defined

$$\widehat{B}_{\mathrm{V}}(\vec{x}) = \prod_{\sigma} \widehat{B}_V^{\sigma}(\vec{x}) \tag{3.49}$$

as we have done for $\widehat{b}_{\mathrm{V}}$ in Eq. (3.35).

According to Kalos and Whitlock [26, pp. 162-3], the iteration of Eq. (3.48) can be achieved stochastically by MC sampling of $\vec{x}$. That is, for each random walker $|\phi_{\mathfrak{j}}^{(\tau)}\rangle$, we choose an auxiliary-field configuration $\vec{x}$ according to the PDF $P(\vec{x})$ and propagate the determinant to a new determinant via

$$|\phi_{\mathfrak{j}}^{(\tau+\Delta\tau)}\rangle = \widehat{B}_{\mathrm{K}/2}\widehat{B}_{\mathrm{V}}(\vec{x})\widehat{B}_{\mathrm{K}/2} |\phi_{\mathfrak{j}}^{(\tau)}\rangle \ . \tag{3.50}$$

We repeat this procedure for *all* walkers in the population. These operations accomplish one step of the random walk. The new population represents $|\Psi^{(\tau+\Delta\tau)}\rangle$ in the sense of Eq. (3.47), i.e.

$$|\Psi^{(\tau+\Delta\tau)}\rangle = \sum_{\mathfrak{j}} |\phi_{\mathfrak{j}}^{(\tau+\Delta\tau)}\rangle \ . \tag{3.51}$$

These steps are iterated indefinitely. After an equilibration phase, all walkers thereon are MC samples of the ground-state wave function $|\Psi_0\rangle$ and ground-state properties can be computed. We will refer to this type of approach as *free projection*.

[2]The indices for Slater determinants in an ensemble are typeset in Euler Fraktur typeface ($\mathfrak{i}$, $\mathfrak{j}$) in order to distinguish them from imaginary unit in roman typeface (i,j) and single-particle basis indices and lattice site indices in default italic typeface (i,j).

Note that the overall prefactor $e^{\Delta\tau[E_T - U(N_\uparrow + N_\downarrow)/2]}$ in Eq. (3.48) is just a scalar and is the same for all walkers. As discussed in Section 3.8, it will be incorporated into the walker's weight.

3.8 Importance sampling

To obtain efficient sampling, the random walkers should give roughly the same contribution i.e. their *weights* should roughly be of the same order. Roughly speaking, the weight is a measure of the "importance" of a walker to the description of the wave function. Those with large weights should be split into several identical walkers with smaller weights. Similarly, those with small weights should be eliminated with a proper probability. This mechanism is called *population control*.

What quantity should define this weight? One choice (though not the best) is the square root of the Slater determinant's magnitude i.e.

$$w_j^{(\tau)} = \sqrt{\langle \phi_j^{(\tau)} | \phi_j^{(\tau)} \rangle}\,. \tag{3.52}$$

With this weight choice, the Monte Carlo representation of a wave function in Eq. (3.47) takes on a different formal appearance:

$$|\Psi^{(\tau)}\rangle = \sum_{j=1}^{N_{\text{wlkr}}} w_j^{(\tau)} \frac{|\phi_j^{(\tau)}\rangle}{\sqrt{\langle \phi_j^{(\tau)} | \phi_j^{(\tau)} \rangle}}\,. \tag{3.53}$$

Note that the weights also carry an imaginary time stamp because they evolve with the walkers and must be carried along with the wave function $\{|\phi_j\rangle\}$ in order to get the correct ground state wave function.

There is one important feature of Eq. (3.53). We can safely scale the walker $|\phi_j^{(\tau)}\rangle$ by any arbitrary constant because this constant will cancel in the ratio $\frac{|\phi_i^{(\tau)}\rangle}{\sqrt{\langle \phi_j^{(\tau)} | \phi_j^{(\tau)} \rangle}}$. The true ground state wave function is fully recoverable as long as the weight factor $w_j^{(\tau)}$ contains the actual value of the normalization constant (i.e. not scaled by any constant).

3.8.1 Importance-sampled random walkers

In practice, the efficiency of the simple random walk described in Section 3.7 is very low because the random walk naively samples the Hilbert space and causes the weights of the walkers to fluctuate greatly, leading to large statistical noise.

We now introduce an importance sampling scheme to increase the walk's efficiency by using the information provided by the trial wave function $|\phi_T\rangle$ to guide the random walk into regions where the expected contribution to the wave function is large. The weight of a walker $|\phi_j^{(\tau)}\rangle$ is defined to be its overlap with the trial wave function $|\phi_T\rangle$ i.e.

$$w_j^{(\tau)} = \langle \phi_T | \phi_j^{(\tau)} \rangle\,. \tag{3.54}$$

The overlap enters to redefine the weight factor such that walkers which have large overlap with $|\phi_T\rangle$ will be considered more "important" and will tend to be sampled more. Such walkers will also have greater contribution in the measured observables.

Like Eq. (3.53), this weight definition leads to a Monte Carlo representation of the wave function of the form

$$|\Psi^{(\tau)}\rangle = \sum_{\mathrm{j}} w_{\mathrm{j}}^{(\tau)} \frac{|\phi_{\mathrm{j}}\rangle^{(\tau)}}{\langle\phi_{\mathrm{T}}|\phi_{\mathrm{j}}^{(\tau)}\rangle} \,. \tag{3.55}$$

Since a walker appears in a ratio $|\phi_{\mathrm{j}}\rangle/\langle\phi_{\mathrm{T}}|\phi_{\mathrm{j}}\rangle$, its normalization is no longer relevant and can be discarded. The only meaningful information in $|\phi_{\mathrm{j}}\rangle$ is its position in the Slater determinant space.

3.8.2 Modified Hirsch spin transformation

With the weight defined in Eq. (3.54), we now iterate a *formally* different but *mathematically* equivalent version of Eq. (3.48), as proved in Appendix E:

$$|\widetilde{\phi}_{\mathrm{j}}^{(\tau+\Delta\tau)}\rangle = \mathrm{e}^{\Delta\tau[E_{\mathrm{T}}-U(N_\uparrow+N_\downarrow)/2]} \sum_{\vec{x}} \widetilde{P}(\vec{x}) \widehat{B}_{\mathrm{K}/2} \widehat{B}_{\mathrm{V}}(\vec{x}) \widehat{B}_{\mathrm{K}/2} \, |\widetilde{\phi}_{\mathrm{j}}^{(\tau)}\rangle \,. \tag{3.56}$$

The importance-sampled walkers are the same as in Eq. (3.55):

$$|\widetilde{\phi}_{\mathrm{j}}^{(\tau)}\rangle = \frac{|\phi_{\mathrm{j}}\rangle^{(\tau)}}{\langle\phi_{\mathrm{T}}|\phi_{\mathrm{j}}^{(\tau)}\rangle} \,. \tag{3.57}$$

The new normalized discrete PDF in Eq. (3.56) is

$$\widetilde{P}(\vec{x}) = \prod_i \left[\frac{\widetilde{p}(x_i)}{\mathfrak{N}_i(\phi_{\mathrm{j}}^{(\tau)})} \right] \tag{3.58}$$

where the probability for sampling the auxiliary field $x_i = \pm 1$ at each lattice is given by

$$\widetilde{p}(x_i) = p(x_i) \frac{\langle\phi_{\mathrm{T}}|\widehat{b}_{\mathrm{V}}(x_i)|\phi_{\mathrm{j},i-1}^{(\tau)}\rangle}{\langle\phi_{\mathrm{T}}|\phi_{\mathrm{j},i-1}^{(\tau)}\rangle} \tag{3.59}$$

and the normalization factor for each site is

$$\mathfrak{N}(\phi_{\mathrm{j},i}^{(\tau)}) = \frac{1}{2} \left[\left\langle \phi_{\mathrm{T}} \left| \widehat{b}_{\mathrm{V}}(x_i = +1) \right| \phi_{\mathrm{j},i-1}^{(\tau)} \right\rangle + \left\langle \phi_{\mathrm{T}} \left| \widehat{b}_{\mathrm{V}}(x_i = -1) \right| \phi_{\mathrm{j},i-1}^{(\tau)} \right\rangle \right] . \tag{3.60}$$

In Eqs. (3.59) and (3.60), the notation

$$|\phi_{\mathrm{j},i-1}^{(\tau)}\rangle = \widehat{b}_V(x_{i-1}) \, \widehat{b}_V(x_{i-2}) \, \cdots \, \widehat{b}_V(x_1) \, |\phi_{\mathrm{j}}^{(\tau)}\rangle \tag{3.61}$$

denotes the state of the j-th walker, $|\phi_{\mathrm{j}}^{(\tau)}\rangle$, just after its first $(i-1)$ fields have been sampled and updated, and

$$|\phi_{\mathrm{j},i}^{(\tau)}\rangle = \widehat{b}_V(x_i) \, |\phi_{\mathrm{j},i-1}^{(\tau)}\rangle \tag{3.62}$$

is the next sub-step after the i-th field is selected and the walker is updated. Thus in this notation:

- $|\phi_{\mathrm{j},0}^{(\tau)}\rangle$ is the walker before any auxiliary field sampling is done at the current time step. Because we apply the propagators in the order specified in Eq. (3.56), applying $\widehat{B}_{\mathrm{K}/2}$ to a walker that has just finished the iteration for $\tau - \Delta\tau$ makes it ready for the sampling of the first auxiliary field in the iteration τ:

$$\widehat{B}_{\mathrm{K}/2} \, |\phi_{\mathrm{j}}^{(\tau-\Delta\tau)}\rangle = |\phi_{\mathrm{j},0}^{(\tau)}\rangle \,. \tag{3.63}$$

- $|\phi_{\mathrm{j},M}^{(\tau)}\rangle$ is the walker after the auxiliary fields on all sites have been sampled. Applying $\widehat{B}_{\mathrm{K}/2}$ to a walker that has just completed the sampling of all auxiliary fields for time τ will complete the iteration i.e. take it to the beginning of the iteration for time $\tau + \Delta\tau$

$$\widehat{B}_{\mathrm{K}/2}\,|\phi_{\mathrm{j},M}^{(\tau)}\rangle = |\phi_{\mathrm{j}}^{(\tau+\Delta\tau)}\rangle\ . \tag{3.64}$$

$\widetilde{P}(\vec{x})$ is a function of both the current and future positions in Slater determinant space and modifies $P(\vec{x})$ such that the probability of sampling $\vec{x}$ is increased when $\vec{x}$ leads to a determinant with larger overlap with $|\phi_{\mathrm{T}}\rangle$ and is decreased otherwise.

In each $\widetilde{p}(x_i)$, x_i can only take the values of $+1$ or -1 and can be sampled by choosing x_i from the PDF $\widetilde{p}(x_i)/\mathfrak{N}(\phi_{\mathrm{j},i}^{(\tau)})$ and multiplying the weight of the walker by $\mathfrak{N}(\phi_{\mathrm{j},i}^{(\tau)})$ (see Appendix E).

We note that the ratio of the overlaps in Eq. (3.59) involves a change of only one row in the matrix representation Φ_{j} of $|\phi_{\mathrm{j}}^{(\tau)}\rangle$ and, in our implementation, is computed quickly using the Sherman-Morrison formula [34]. The inverse of the overlap matrix $\left[(\Phi_{\mathrm{T}})^{\dagger}(\Phi_{\mathrm{j}}^{(\tau)})\right]^{-1}$ is kept and updated after each x_i is selected.[3]

3.9 The sign problem and the constrained path approximation

The sign problem occurs because of the fundamental symmetry between the fermion ground state $|\Psi_0\rangle$ and its negative $-|\Psi_0\rangle$ [35]. This symmetry implies that, for any ensemble of Slater determinants $\{|\phi\rangle\}$ which gives a Monte Carlo representation of the ground-state wave function, there exists another ensemble $\{-|\phi\rangle\}$ which is also a correct representation. In other words, the Slater determinant space can be divided into two degenerate halves ($+$ and $-$) whose bounding surface $\mathcal{N}$ is defined by $\langle\Psi_0|\phi\rangle = 0$. This surface is in general *unknown*. Except for some special cases [18], walkers do cross $\mathcal{N}$ in their propagation by $\mathcal{P}_{\mathrm{gs}}$, causing the sign problem. The constrained-path approximation relies on the observation that at the instant such a walker lands on $\mathcal{N}$, the walker will make no further contribution to the representation of the ground state at any later time because

$$\langle\Psi_0|\phi\rangle = 0 \implies \langle\Psi_0|\mathrm{e}^{-\tau H}|\phi\rangle = 0 \quad \text{for any } \tau\,. \tag{3.65}$$

Paths that result from such a walker have equal probabilities of being in either half of the Slater determinant space [5]. Computed analytically, they would cancel and make no contribution in the ground-state wave function. However, because the random walk has no knowledge of $\mathcal{N}$, these paths continue to be sampled (randomly) in the random walk and become Monte Carlo noise.

To eliminate the decay of the signal-to-noise ratio, we impose the constrained path approximation. It requires that each random walker at each step have a positive overlap with the trial wave function $|\phi_{\mathrm{T}}\rangle$:

$$\langle\phi_{\mathrm{T}}|\phi_{\mathrm{j}}^{(n)}\rangle > 0.\,. \tag{3.66}$$

[3] There are two reasons for keeping the inverse of the overlap matrix $\left[(\Phi_{\mathrm{T}})^{\dagger}(\Phi_{\mathrm{j}}^{(\tau)})\right]^{-1}$ instead of the overlap matrix itself. One is convenience in the Sherman-Morrison update. Another is that the inverse has smaller size ($N_{\uparrow} \times N_{\uparrow}$ or $N_{\downarrow} \times N_{\downarrow}$) than the actual overlap ($M \times M$).

This yields an approximate solution to the ground-state wave function,

$$|\Psi_0^{\mathrm{c}}\rangle = \sum_{\mathrm{j}} |\phi_{\mathrm{j}}\rangle \tag{3.67}$$

in which all Slater determinants $|\phi_{\mathrm{j}}\rangle$ satisfy Eq. (3.66). Note that from Eq. (3.65), the constrained path approximation becomes *exact* for an exact trial wave function $|\psi_{\mathrm{T}}\rangle = |\Psi_0\rangle$. The constrained path approximation is easily implemented by eliminating a walker when its overlap with the trial wave function becomes zero or negative.

3.10 Measurement of physical observables

The ground-state expectation value of an observable $\widehat{A}$ is

$$\langle \widehat{A} \rangle_{\mathrm{gs}} = \frac{\langle \Psi_0 | \widehat{A} | \Psi_0 \rangle}{\langle \Psi_0 | \Psi_0 \rangle} \,. \tag{3.68}$$

In principle, we can use the same Monte Carlo sample as both $|\Psi_0\rangle$ and $\langle\Psi_0|$. A "brute-force" measurement on the population $\{|\phi_{\mathrm{j}}^{(\tau)}\rangle\}$ is given by

$$\langle \widehat{A} \rangle_{\mathrm{bf}}^{(\tau)} = \frac{\sum_{\mathrm{ij}} [w_{\mathrm{j}}^{(\tau)}]^* w_{\mathrm{i}}^{(\tau)} \, \langle \phi_{\mathrm{j}}^{(\tau)} | \widehat{A} | \phi_{\mathrm{i}}^{(\tau)} \rangle}{\sum_{\mathrm{ij}} [w_{\mathrm{j}}^{(\tau)}]^* w_{\mathrm{i}}^{(\tau)} \, \langle \phi_{\mathrm{j}}^{(\tau)} | \phi_{\mathrm{i}}^{(\tau)} \rangle} \tag{3.69}$$

and the estimator $\langle \widehat{A} \rangle_{\mathrm{bf}}$ is the average of such measurements. Because the walkers are nonorthogonal in CPMC, Eq. (3.69) is in principle well-defined. The ground-state energy estimated in this way is variational and converges to the exact value in the limit of large N_{wlkr}, the number of walkers. However, in practice, the use of $\langle \widehat{A} \rangle_{\mathrm{bf}}$ is limited to smaller systems. In general, it will have large variances. Reducing the variance is expensive because the computational cost of $\langle \widehat{A} \rangle_{\mathrm{bf}}$ scales as $\mathcal{O}(N_{\mathrm{wlkr}}^2)$.

3.10.1 The mixed estimator and energy

The simplest way to compute observables is with the so-called *mixed estimator*

$$\langle \widehat{A} \rangle_{\mathrm{mixed}} \equiv \frac{\langle \phi_{\mathrm{T}} | \widehat{A} | \Psi^{(\tau)} \rangle}{\langle \phi_{\mathrm{T}} | \Psi^{(\tau)} \rangle} \,. \tag{3.70}$$

For example, the mixed estimator for the energy is the same as the true estimator and, for an ensemble $\{|\phi\rangle\}$, is given by

$$E_{\mathrm{mixed}} = \frac{\sum_{\mathrm{j}} w_{\mathrm{j}} \, E_{\mathrm{L}} \left[\phi_{\mathrm{T}}, \phi_{\mathrm{j}} \right]}{\sum_{\mathrm{j}} w_{\mathrm{j}}} \,. \tag{3.71}$$

where we have defined the local energy E_{L} as

$$E_{\mathrm{L}} \left[\phi_{\mathrm{T}}, \phi_{\mathrm{j}} \right] = \frac{\langle \phi_{\mathrm{T}} | \widehat{H} | \phi_{\mathrm{j}} \rangle}{\langle \phi_{\mathrm{T}} | \phi_{\mathrm{j}} \rangle} \,. \tag{3.72}$$

The local energy can be easily evaluated for any walker $|\phi\rangle$ as follows. For any pair of Slater determinants $|\phi_{\mathrm{T}}\rangle$ and $|\phi\rangle$, we can calculate the expectation of the operator $c^{\dagger}_{j\sigma}c_{i\sigma}$ as:

$$\langle c^{\dagger}_{j\sigma}c_{i\sigma}\rangle \equiv \frac{\langle\phi_{\mathrm{T}}|c^{\dagger}_{j\sigma}c_{i\sigma}|\phi\rangle}{\langle\phi_{\mathrm{T}}|\phi\rangle} = \left[\Phi^{\sigma}\left[\,(\Phi^{\sigma}_{\mathrm{T}})^{\dagger}\Phi^{\sigma}\,\right]^{-1}(\Phi^{\sigma}_{\mathrm{T}})^{\dagger}\right]_{ij}. \tag{3.73}$$

Eq. (3.73) is proved in Appendix F and immediately enables the computation of the kinetic energy term $\left\langle \phi_{\mathrm{T}} \middle| -t\sum_{\langle ij\rangle\sigma} c^{\dagger}_{i\sigma}c_{j\sigma} \middle| \phi \right\rangle$. The potential energy term $\left\langle \phi_{\mathrm{T}} \middle| U\sum_i c^{\dagger}_{i\uparrow}c_{i\uparrow}c^{\dagger}_{i\downarrow}c_{i\downarrow} \middle| \phi \right\rangle$ does not have the form of Eq. (3.73), but can be reduced to that form by an application of Wick's theorem:[4]

$$\langle c^{\dagger}_{i\uparrow}c_{i\uparrow}c^{\dagger}_{i\downarrow}c_{i\downarrow}\rangle = \langle c^{\dagger}_{i\uparrow}c_{i\uparrow}\rangle\langle c^{\dagger}_{i\downarrow}c_{i\downarrow}\rangle + \langle c^{\dagger}_{i\uparrow}c_{i\downarrow}\rangle\langle c_{i\uparrow}c^{\dagger}_{i\downarrow}\rangle \tag{3.75}$$

$$= \langle c^{\dagger}_{i\uparrow}c_{i\uparrow}\rangle\langle c^{\dagger}_{i\downarrow}c_{i\downarrow}\rangle\ . \tag{3.76}$$

The reduction to the last line occurs because $\langle c^{\dagger}_{i\uparrow}c_{i\downarrow}\rangle$ and $\langle c_{i\uparrow}c^{\dagger}_{i\downarrow}\rangle$ evaluate to zero. The reason is explained in Section 2.2.2.

The mixed estimator for the energy arises naturally from importance sampling, and reduces the statistical variance of the computed result. A drawback of the mixed estimator is that the ground-state energy obtained in AFQMC under the constrained path approximation is not variational [5]. The mixed estimators for observables which do not commute with the Hamiltonian are biased. The back-propagation technique, as discussed in Appendix I, can be used to obtain pure estimates [6, 29].

3.10.2 Non-commuting observables and back propagation

The mixed estimator in Eq. (3.70) is exact only if the operator $\widehat{A}$ commutes with the Hamiltonian. Otherwise, a systematic error arises due to the non-commutativity. One can correct for this systematic error by combining the mixed estimator with the so-called *purely variational estimator*

$$\langle\widehat{A}\rangle_{\mathrm{var}} \equiv \frac{\langle\phi_{\mathrm{T}}|\widehat{A}|\phi_{\mathrm{T}}\rangle}{\langle\phi_{\mathrm{T}}|\phi_{\mathrm{T}}\rangle} \tag{3.77}$$

in the following ways

$$\langle\widehat{A}\rangle_{\mathrm{extrap1}} = 2\,\langle\widehat{A}\rangle_{\mathrm{mixed}} - \langle\widehat{A}\rangle_{\mathrm{var}} \tag{3.78}$$

$$\langle\widehat{A}\rangle_{\mathrm{extrap2}} = \frac{\langle\widehat{A}\rangle^{2}_{\mathrm{mixed}}}{\langle\widehat{A}\rangle_{\mathrm{var}}}\,. \tag{3.79}$$

These corrections are good only if $|\phi_{\mathrm{T}}\rangle$ does not differ significantly from $|\Psi_0\rangle$. In general, we need the back-propagation scheme to obtain the correct ground-state properties where we actually propagate both the wave functions on the right- and left-hand side of the operator $\widehat{A}$ in Eq. (3.68). This gives to the *back-propagated estimator*:

[4] Wick's theorem essentially says that any multi-point Green's function can be decomposed into sums and products of two-point Green's functions. This is a standard result in many-body quantum theory e.g. Negele and Orland [27, pp. 75-8]. The form of the theorem used in this thesis is stated in Santos [36, p. 40]:

$$\langle c^{\dagger}_{x_1}c_{x_2}c^{\dagger}_{x_3}c_{x_4}\rangle_{\vec{X}} = \langle c^{\dagger}_{x_1}c_{x_2}\rangle_{\vec{X}}\langle c^{\dagger}_{x_3}c_{x_4}\rangle_{\vec{X}} + \langle c^{\dagger}_{x_1}c_{x_4}\rangle_{\vec{X}} + \langle c_{x_2}c^{\dagger}_{x_3}\rangle_{\vec{X}} \tag{3.74}$$

where x_1, x_2, x_3 and x_4 are space-spin basis states. We have included the auxiliary-field path $\vec{X}$ to remind the reader that this identity only holds for a *fixed* auxiliary-field path.

$$\langle \widehat{A} \rangle_{\mathrm{bp}} = \frac{\langle \phi_{\mathrm{T}} | \mathrm{e}^{-\tau_{\mathrm{bp}} \hat{H}} \widehat{A} | \Psi_0 \rangle}{\langle \phi_{\mathrm{T}} | \mathrm{e}^{-\tau_{\mathrm{bp}} \hat{H}} | \Psi_0 \rangle} \,. \tag{3.80}$$

The back-propagation technique was first introduced by Zhang *et al.* [6] and discussed in details by Purwanto and Zhang [29]. It reuses the auxiliary field "paths" from different segments of the random walk to obtain $\langle \phi_{\mathrm{T}} | \, \mathrm{e}^{-\tau_{\mathrm{bp}} \widehat{H}}$ while avoiding the $\mathcal{O}(N_{\mathrm{wlkr}}^2)$ scaling of a brute-force evaluation with two separate populations in Eq. (3.69).

The method is derived in detail in Appendix I. The resulting procedure to obtain a back-propagated estimator $\langle \widehat{A} \rangle_{\mathrm{bp}}$ is as follows:

1. At some time τ (which shall be *fixed* in this procedure) of the regular forward walk, the collection of all weights and walkers $\{ |\phi_{\mathrm{j}}^{(\tau)}\rangle \, , w_{\mathrm{j}}^{(\tau)} \}$ are recorded.

2. The forward random walk continues for an imaginary time interval τ_{bp} and the auxiliary-field "path" for each walker is remembered. This path is made of the propagators

   $$\widehat{B}_{\mathrm{j}}^{(\tau)} \equiv \widehat{B}_{\mathrm{K}/2} \widehat{B}_{\mathrm{V}}(\vec{x}_{\mathrm{j}}^{(\tau)}) \widehat{B}_{\mathrm{K}/2} \tag{3.81}$$

   at each time step.

3. Temporarily stop the forward walk when it reaches imaginary time $\tau' = \tau + \tau_{\mathrm{bp}}$. The population of the forward walk at this stage is $\{ |\phi_{\mathrm{j}}^{(\tau')}\rangle \, , w_{\mathrm{j}}^{(\tau')} \}$.

4. Perform an importance-sampled backward random walk

   (a) Initiate the back-propagated population $\{ |\eta_{\mathrm{j}}^{(\theta)}\rangle \, , u_{\mathrm{j}}^{(\theta)} \}$, consisting of the same number of walkers as the population $\{ |\phi_{\mathrm{j}}^{(\tau')}\rangle \}$, to be the same as the trial wave function:

   $$|\eta_{\mathrm{j}}^{(\theta=0)}\rangle = |\phi_{\mathrm{T}}\rangle \tag{3.82}$$

   $$u_{\mathrm{j}}^{(\theta=0)} = \langle \phi_{\mathrm{j}}^{(\tau)} | \phi_{\mathrm{T}} \rangle \tag{3.83}$$

   where θ ranges from 0 to τ_{bp} and is the imaginary time stamp for the backward walk.

   (b) Back-propagate the population $\{ |\eta_{\mathrm{j}}^{(\theta)}\rangle \, , u_{\mathrm{j}}^{(\theta)} \}$ up to time $\theta = \tau_{\mathrm{bp}}$ by applying the propagators $B_{\mathrm{j}}^{(\tau)}$ (as defined in Eq. (3.81)) in *reverse* order in τ. Note that no auxiliary field sampling is necessary here. At every step of the backward walk, update the weight the same way as in a forward walk except that the "guiding" wave function is $|\phi_{\mathrm{j}}^{(\tau)}\rangle$ instead of $|\phi_{\mathrm{T}}\rangle$ where j is the index of the forward walker whose path is being back-traversed by the current backward walker.

5. Calculate the back-propagated estimator as

   $$\langle \widehat{A} \rangle_{\mathrm{bp}} = \frac{\displaystyle\sum_{\mathrm{j}} w_{\mathrm{j}}^{(\tau')} \frac{\langle \eta_{\mathrm{j}}^{(\tau_{\mathrm{bp}})} | \widehat{A} | \phi_{\mathrm{j}}^{(\tau)} \rangle}{\langle \eta_{\mathrm{j}}^{(\tau_{\mathrm{bp}})} | \phi_{\mathrm{j}}^{(\tau)} \rangle}}{\displaystyle\sum_{\mathrm{j}} w_{\mathrm{j}}^{(\tau')}} \,. \tag{3.84}$$

   Note that the walkers in the back-propagated population and the forward-propagated population *must* be matched in a *one-to-one* manner and the weights in Eq. (3.84) are those *at the later time* τ'.

6. Resume the forward random walk that was interrupted in step 3 and repeat this procedure many times to obtain more back-propagated estimators.

The back-propagated estimator in Eq. (3.84) parallels the Metropolis AFQMC estimator for the ground state expectation of $\widehat{A}$ in Eq. (G.30) in Appendix G. The $\langle\eta_{\mathrm{j}}|$'s and $|\phi_{\mathrm{j}}\rangle$'s have similar meanings. The only difference lies in how the paths are generated. In CPMC an open-ended random walk is used to advance an ensemble of paths from τ to τ', which result in fluctuating weights that represent the path distribution. In Metropolis AFQMC, a fixed length path corresponding to $\tau_{\mathrm{bp}}+\tau_{\mathrm{eq}}$ (with τ_{eq} being the minimum time for equilibration) is moved about by the Metropolis algorithm, which estimates branching by the acceptance/rejection step. In other words, the estimators in Eqs. (3.84) and (G.30) are the same except for the weights.

The back-propagated estimator in Eq. (3.84) can be used to calculate the equal-time Green's function $\langle c_i c_j^\dagger\rangle$ because the numerator of Eq. (3.84) contains

$$\frac{\langle\eta_{\mathrm{j}}^{(\tau_{\mathrm{bp}})}|c_{i\sigma}c_{j\sigma}^\dagger|\phi_{\mathrm{j}}^{(\tau)}\rangle}{\langle\eta_{\mathrm{j}}^{(\tau_{\mathrm{bp}})}|\phi_{\mathrm{j}}^{(\tau)}\rangle} = I - \left[\Phi_{\mathrm{j}}^{\sigma,(\tau)}\left[\left(\Omega_{\mathrm{j}}^{\sigma,(\tau_{\mathrm{bp}})}\right)^\dagger\Phi_{\mathrm{j}}^{\sigma,(\tau)}\right]^{-1}\left(\Omega_{\mathrm{j}}^{\sigma,(\tau_{\mathrm{bp}})}\right)^\dagger\right]_{ji} \tag{3.85}$$

where $\Omega_{\mathrm{j}}^{\sigma,(\tau_{\mathrm{bp}})}$ is the matrix of the spin-up sector of $|\eta_{\mathrm{j}}^{(\tau_{\mathrm{bp}})}\rangle$. Eq. (3.85) is proved in Appendix F. The unequal-time Green's function is also a non-commuting observable but because the procedure to calculate it is lengthy, it will be discussed in the next section.

3.11 Unequal-time Green's function

In this section we will develop a procedure to calculate the unequal-time Green's function in CPMC. We will need results proved in Appendix H (a very important appendix). For any two Slater determinants $|\phi_1\rangle$ and $|\phi_2\rangle$ and $\tau > 0$, their unequal-time Green's function

$$G^\sigma(\tau,0)_{\ell j} = \frac{\langle\phi_1|c_{\ell\sigma}(\tau)c_{j\sigma}^\dagger(0)|\phi_2\rangle}{\langle\phi_1|\phi_2\rangle} \tag{3.86}$$

can be expressed as a product of unequal-time Green's functions over L smaller time slices, each of length ε (i.e. $L\varepsilon=\tau$)

$$G(\tau,0) = \prod_{n=0}^{L-1} G\Big((n+1)\varepsilon, n\varepsilon\Big). \tag{3.87}$$

The decomposition into smaller time slices ε is *exact* and helps reduce numerical errors. Each of the unequal-time Green's functions over the ε interval is given by:

$$G\Big((n+1)\varepsilon, n\varepsilon\Big)_{\ell j} = \Big[B\Big((n+1)\varepsilon, n\varepsilon\Big)G(n\varepsilon)\Big]_{\ell j} \tag{3.88}$$

where

- $B\Big((n+1)\varepsilon, n\varepsilon\Big)$ is the product of propagators that evolves a Slater determinant from time $n\varepsilon$ to time $(n+1)\varepsilon$.

- $G(n\varepsilon)$ is the equal-time Green's function for the two Slater determinants $|\phi_1\rangle$ and $|\phi_2\rangle$ at time $n\varepsilon$. Recall that from Appendix F, the equal-time Green's function for any two Slater determinants $|\phi\rangle$ and $|\phi'\rangle$ is

$$G = I - \Phi(\Phi'^{\dagger}\Phi)^{-1}\Phi' \tag{3.89}$$

  where the (i, j)-the element of G is

$$G_{ij} = \frac{\langle\phi'|c_i c_j^{\dagger}|\phi\rangle}{\langle\phi'|\phi\rangle} \tag{3.90}$$

  Also recall from Appendix F that by "equal-time," we mean there are no time evolution operators sandwiched between c_i and $c_j^{\dagger}$.

Because the equal-time Green's function $G(n\varepsilon)$ is well-behaved due do the numerical stabilization scheme outlined in Appendix J, the small time interval ε in Eq. (3.87) is chosen so that the propagator matrix $B\Big((n+1)\varepsilon, n\varepsilon\Big)$ is well-conditioned.

3.11.1 Non-interacting electrons in an external potential

Before implementing Eq. (3.87) to calculate unequal-time Green's function in the Hubbard model (which requires stochastic sampling of auxiliary fields as outlined thus far in this chapter), we want to understand its numerical behavior where exact non-stochastic results are easy to obtain. A non-interacting system subject to an external potential, studied in Section 1.2.3, is an excellent testbed for this purpose because the exact result is easy to obtain and involves no stochasticity.

We will now summarize summarize the results of Section 1.2.3. The Hamiltonian for a non-interacting system in an external potential is

$$\widehat{H} = -t \sum_{\sigma,\langle i,j\rangle} c_{i\sigma}^{\dagger} c_{j\sigma} + \sum_{\ell} V_{\ell} c_{\ell}^{\dagger} c_{\ell}\,. \tag{3.91}$$

Here we assume a one-dimensional lattice of M sites with $N_{\uparrow}$ spin-up electrons, $N_{\downarrow}$ spin-down electrons and periodic boundary conditions. We also assume N_{σ} is odd to avoid energy degeneracy. For this section, we choose the potential to be sinusoidal:[5]

$$V_{\ell} = \cos\left(\frac{2\pi\ell}{M}\right)\,. \tag{3.92}$$

To find the unequal-time Green's function, we first diagonalize the Hamiltonian matrix H as

$$H = QDQ^{\dagger} \tag{3.93}$$

where D is a diagonal matrix containing the energy eigenvalues E_i (in increasing order i.e. the lowest energy eigenvalue is at the the top left corner of D) and the columns of Q are the eigenvectors of H. The unequal-time Green's function is then given by

$$\langle c_{\ell}(\tau) c_j^{\dagger}\rangle = \Big[\mathcal{Q}\mathcal{E}\mathcal{Q}^{\dagger}\Big]_{\ell j} \tag{3.94}$$

where

[5] Furthermore, because potential energy term in Eq. (3.91) does not commute with the kinetic energy term, we will have the the Trotter error in the projection process. This makes the test more "realistic" because we will be forced to choose a small value of $\Delta\tau$ just as we have to in regular projection by Monte Carlo.

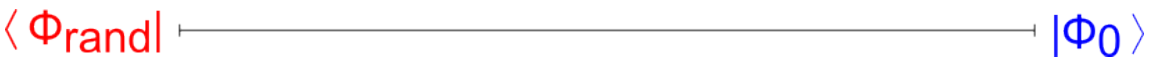


Figure 3.1: Step 1 in the procedure.

- $\mathcal{Q}^\dagger$ is the bottom $(N - N_\sigma) \times N$ block of $Q^\dagger$.
- $\mathcal{E}$ is the bottom right $(N-N_\sigma)\times(N-N_\sigma)$ block of the diagonal matrix whose diagonal elements are $\mathrm{e}^{-\tau E_i}$ (E_i are the aforementioned energy eigenvalues).

Eq. (3.94) provides the benchmark to test the numerical behavior of the extension to CPMC that we will develop in chapter 3. For this test, we will next describe a procedure that looks very contrived to calculate the unequal-time Green's function. Its true utility will become clear in Section 3.11.2.

Procedure

The following temporal quantities are used in this procedure to calculate the unequal-time Green's function $\langle c_{\ell\sigma}(\tau)c_{j\sigma}\rangle$ for a non-interacting system subject to an external sinusoidal potential:

- $\Delta\tau$ is the Trotter time step in the projection operator.
- τ is the imaginary time interval in the unequal-time Green's function $\langle c_{\ell\sigma}(\tau)c^\dagger_{j\sigma}\rangle$.
- ε is the division of τ into L smaller time steps i.e. $L\varepsilon = \tau$ in Eq. (3.87).
- τ_{eq} is the equilibration time for the left-hand walker.

Note that in this procedure, by "propagator" we mean the projection operator in Eq. (3.28) where $\widehat{K}$ and $\widehat{V}$ are the kinetic and potential energy operators, respectively, in the Hamiltonian in Eq. (3.91). $\widehat{K}$ and $\widehat{V}$ are definite (i.e. non-stochastic) in this procedure.

(1) Set the right-hand wave function to the true ground state wave function[6] and set the left-hand wave function to a random matrix of appropriate size.

$$\Phi_{\mathrm{R}} \leftarrow \Phi_0 \tag{3.95}$$
$$\Phi_{\mathrm{L}} \leftarrow \Phi_{\mathrm{rand}} \tag{3.96}$$

See Fig. 3.1.

(2) Because the left-hand wave function is not the ground state, "equilibrate" it to the ground state by performing projection on it for a duration of τ_{eq}.

$$\Phi_{\mathrm{L}} \leftarrow B(\tau_{\mathrm{eq}}, 0)\Phi_{\mathrm{rand}} \tag{3.97}$$

See Fig. 3.2.

(3) Forward-propagate the right-hand wave function for a duration of $(\tau - \varepsilon)$ and forward-propagate the left-hand wave function by ε.

$$\Phi_{\mathrm{R}} \leftarrow B(\tau - \varepsilon, 0)\Phi_0 \tag{3.98}$$
$$\Phi_{\mathrm{L}} \leftarrow B(\tau, \tau - \varepsilon)\Phi_{\mathrm{L}} \tag{3.99}$$

See Fig. 3.3.

[6] Because this is a non-interacting system, the ground state is a single Slater determinant.

Figure 3.2: Step 2 in the procedure.

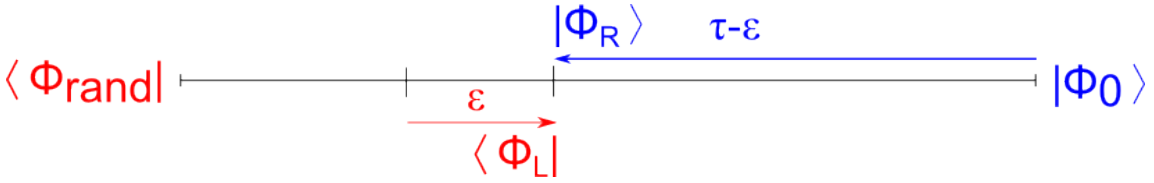


Figure 3.3: Step 3 in the procedure.

(4) Calculate the first small-time unequal-time Green's function as follows (see Fig. 3.4):

- Calculate the equal-time Green's function $G(\tau-\varepsilon)$ for $|\phi_{\rm L}\rangle$ and $|\phi_{\rm R}\rangle$ using Eq. (3.89).
- Calculate the product of propagators $B(\tau,\tau-\varepsilon)$ by multiplying all the propagators between time τ and $\tau-\varepsilon$.
- Calculate the unequal-time Green's function $G(\tau,\tau-\varepsilon)$ by

$$G(\tau,\tau-\varepsilon) = B(\tau,\tau-\varepsilon)G(\tau-\varepsilon)\,. \tag{3.100}$$

(5) Back-propagate $|\phi_{\rm R}\rangle$ by an interval ε and forward-propagate $|\phi_{\rm L}\rangle$ also by an interval ε.

$$\Phi_{\rm R} \leftarrow B^{-1}(\tau-\varepsilon,\tau-2\varepsilon)\Phi_{\rm R} \tag{3.101}$$
$$\Phi_{\rm L} \leftarrow B(\tau-\varepsilon,\tau-2\varepsilon)\Phi_{\rm L} \tag{3.102}$$

See Fig. 3.5.

(6) Calculate the next small-time unequal-time Green's function as follows (see Fig. 3.6):

- Calculate the equal-time Green's function $G(\tau-2\varepsilon)$ for $|\phi_{\rm L}\rangle$ and $|\phi_{\rm R}\rangle$ using Eq. (3.89).
- Calculate the product of propagators $B(\tau,\tau-\varepsilon)$ by multiplying all the propagators between time $\tau-\varepsilon$ and $\tau-2\varepsilon$.

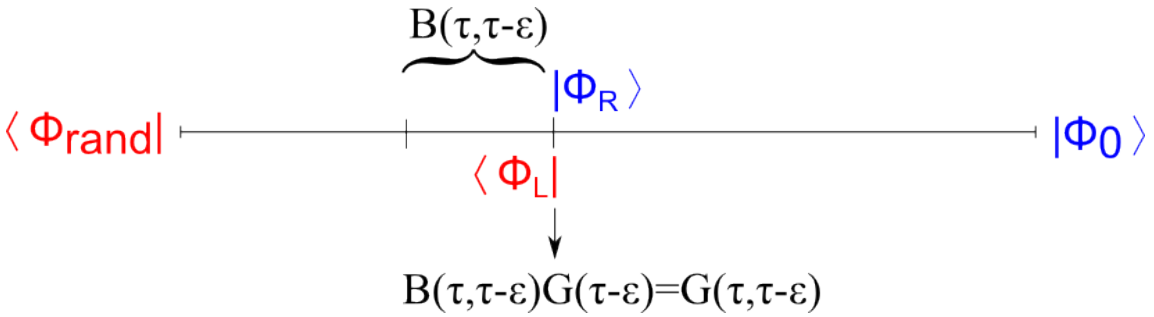


Figure 3.4: Step 4 in the procedure. Note that in these figures, operations on the left-hand wave function is in red and those on the right-hand wave function in blue.

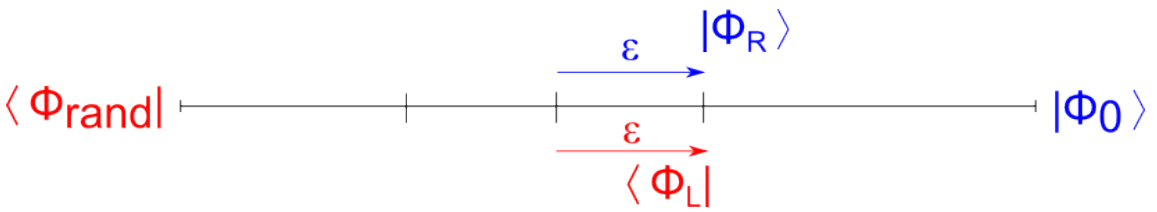


Figure 3.5: Step 5 in the procedure.

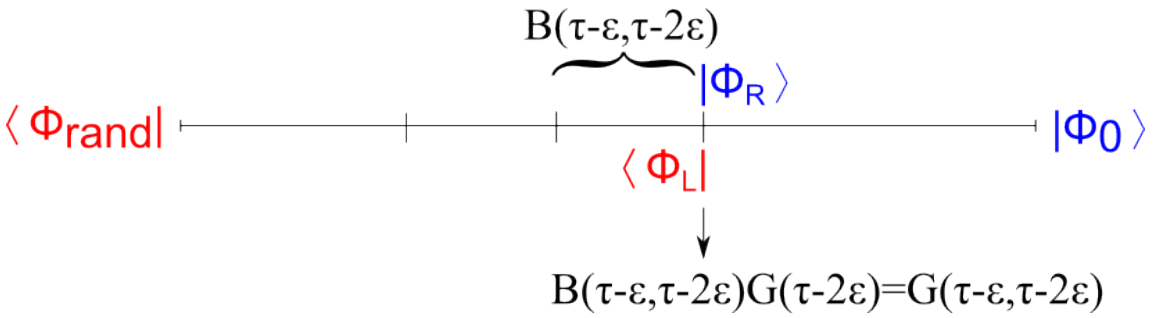


Figure 3.6: Step 6 in the procedure.

- Calculate the unequal-time Green's function $G(\tau-\varepsilon,\tau-2\varepsilon)$ by

$$G(\tau-\varepsilon,\tau-2\varepsilon)=B(\tau-\varepsilon,\tau-2\varepsilon)G(\tau-2\varepsilon)\,. \tag{3.103}$$

(7) Repeat steps 5 to 6 for $(L-2)$ more times.

(8) Calculate the desired unequal-time Green's function as

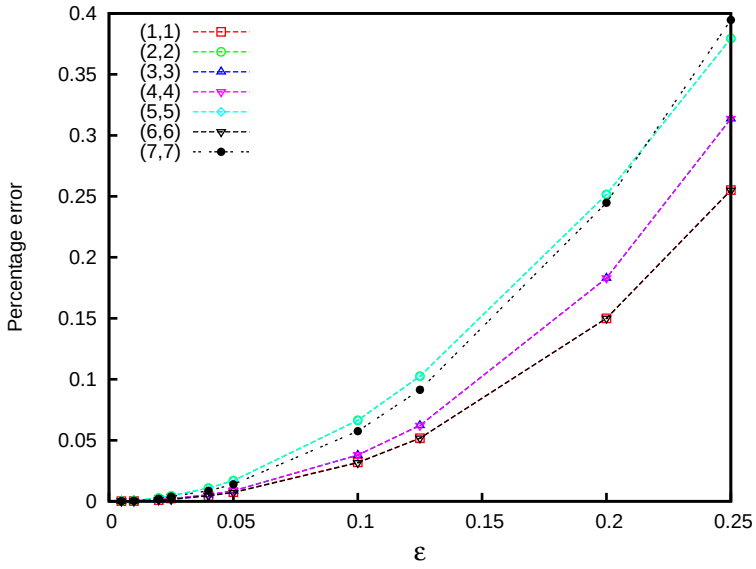
$$G(\tau,0)=G(\tau,\tau-\varepsilon)G(\tau-\varepsilon,\tau-2\varepsilon)\ldots G(\varepsilon,0) \tag{3.104}$$

Note that if we store all the "small-time" unequal-time Green's functions $G\Big((n+1)\varepsilon,n\varepsilon\Big)$ then we can also calculate $G(\tau',0)$ for $0<\tau'<\tau$ and $\tau'=L'\varepsilon$ by the matrix product of L' small-time unequal-time Green's functions.

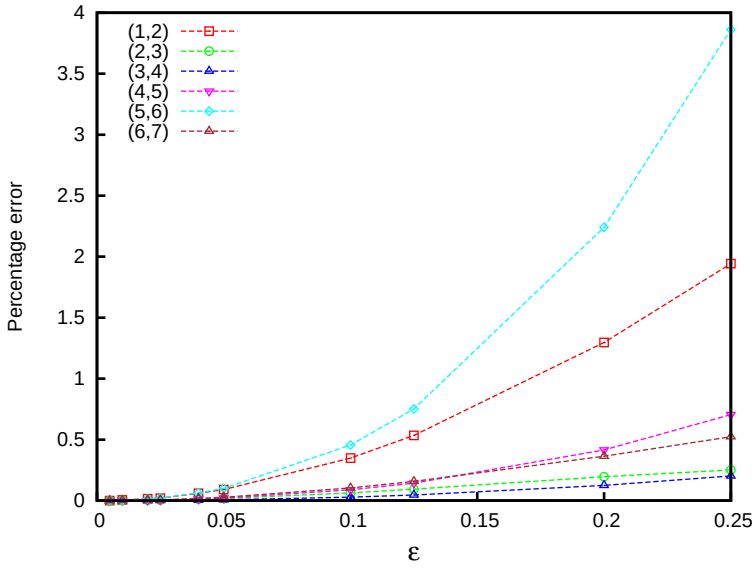
Results

Fig. 3.7 shows the percentage error of the unequal-time Green's function $\langle c^\dagger_{\ell\uparrow}(\tau)c_{j\uparrow}(0)\rangle$ (where $\tau=2$ is constant) as a function of ε for various pairs of (ℓ,j) using the procedure outlined above. The physical system here is a 7-site one-dimensional lattice with 3 spin-$\uparrow$ and 3 spin-$\downarrow$ electrons and $t=1$ and $V=1$. Because the on-diagonal pairs $(\ell=j)$ in general always have smaller errors than the off-diagonal pairs $(\ell\neq j)$, we show them on separate graphs. From Fig. 3.7, we conclude that the procedure above gives very good agreement with exact result if we choose a small enough value of ε e.g. $\varepsilon=0.1$.

On the other hand, Fig. 3.8 plots the imaginary-time Green's function $\langle c^\dagger_{\ell\uparrow}(\tau)c_{j\uparrow}(0)\rangle$ against the imaginary time τ to compare the results obtained with the exact formula in Eq. (1.114) against those obtained by the procedure above. The system is the same as that of Fig. 3.7. Once again we have separate plots for on- and off-diagonal pairs of (ℓ,j). We conclude that the procedure gives very good agreement with exact results.

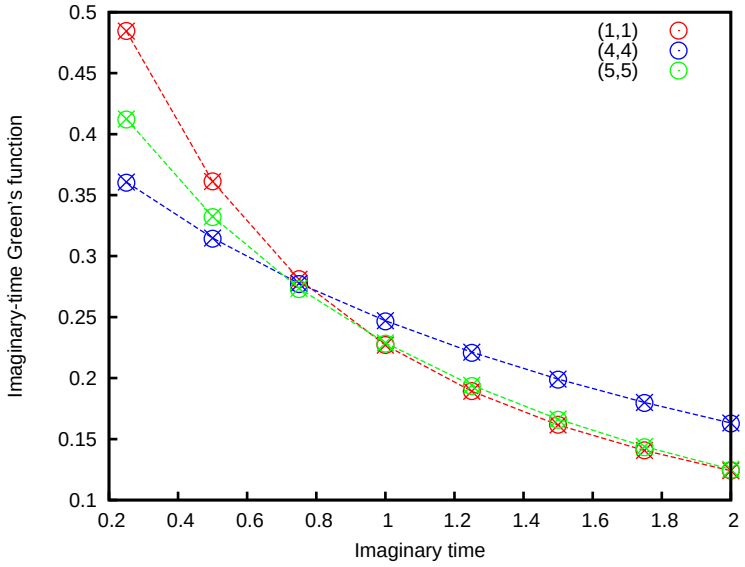


(a) On-diagonal

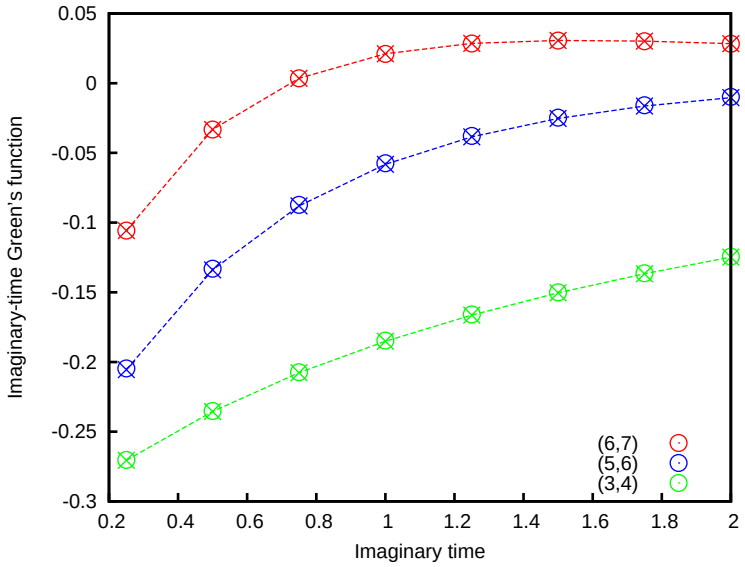


(b) Off-diagonal

Figure 3.7: The percentage error as a function of ε for on- and off-diagonal elements with $\tau = 2$. For example, 0.015 means 1.5% error.



(a) On-diagonal



(b) Off-diagonal

Figure 3.8: Comparison of exact (crosses) and projection (circles) calculations of the unequal-time Green's function $\langle c^\dagger_{\ell\uparrow}(\tau)c_{j\uparrow}(0)\rangle$ against τ for various values of (ℓ, j). The dashed lines connect the points of the exact results and are intended to guide the eye.

Figure 3.9: Step 1 in the procedure.

Figure 3.10: Step 2 in the procedure.

3.11.2 Hubbard interaction

The procedure to obtain a single value of $\dfrac{\langle \eta_j^{(\tau_{\rm bp})} | \widehat{A} | \phi_j^{(\tau)} \rangle}{\langle \eta_j^{(\tau_{\rm bp})} | \phi_j^{(\tau)} \rangle}$ in Eq. (3.84) is almost identical to that in Section 3.11.1. The only difference is that

- $\Phi_{\rm L}$ is initialized to a trial wave function $\Phi_{\rm T}$.
- $\Phi_{\rm R}$ is initialized to be a walker $\Phi_j^{(\tau)}$.
- The (stochastic) propagators used in the procedure are those that make up the remembered path in Eq. (3.81).

Every time back-propagation is carried out, this procedure is performed for every walker in the stored population $\{|\phi_j^{(\tau)}\rangle, w_j^{(\tau)}\}$ in step 1 on page 41.

Before each CPMC run, we have to decide the maximum τ we want in $\langle c_{i\sigma}^\dagger(\tau) c_{j\sigma}(0)\rangle$. This will be called in $\tau_{\rm bp}$ in the following description. We also have to choose an imaginary-time interval $\tau_{\rm eq}$ to equilibrate the trial wave function and the interval ε to stabilize the product. The total "length" of the path (i.e. the number of propagators we have to remember) corresponds to $\tau_{\rm bp} + \tau_{\rm eq}$.

(1) Set the right-hand wave function to the the walker $|\phi_j^{(\tau)}\rangle$ and set the left-hand wave function to the trial wave function $|\phi_{\rm T}\rangle$.

$$\Phi_{\rm R} \leftarrow \Phi_j^{(\tau)} \tag{3.105}$$
$$\Phi_{\rm L} \leftarrow \Phi_{\rm T} \tag{3.106}$$

See Fig. 3.9.

(2) Because the left-hand wave function is not the ground state, "equilibrate" it to the ground state by performing projection on it for a duration of $\tau_{\rm eq}$.

$$\Phi_{\rm L} \leftarrow B(\tau_{\rm eq}, 0)\Phi_{\rm rand} \tag{3.107}$$

See Fig. 3.10.

(3) Forward-propagate the right-hand wave function for a duration of $(\tau_{\rm bp} - \varepsilon)$ and forward-propagate the left-hand wave function by ε.

$$\Phi_{\rm R} \leftarrow B(\tau_{\rm bp} - \varepsilon, 0)\Phi_0 \tag{3.108}$$
$$\Phi_{\rm L} \leftarrow B(\tau_{\rm bp}, \tau_{\rm bp} - \varepsilon)\Phi_{\rm L} \tag{3.109}$$

See Fig. 3.11.

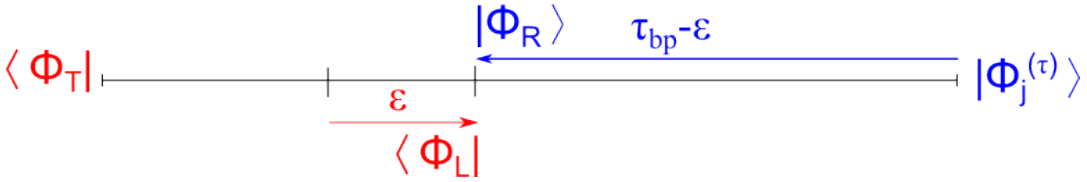


Figure 3.11: Step 3 in the procedure.

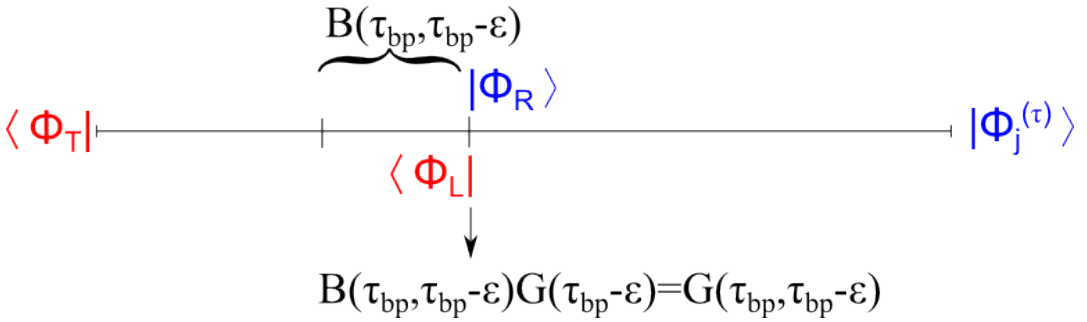


Figure 3.12: Step 4 in the procedure.

(4) Calculate the first small-time unequal-time Green's function as follows (see Fig. 3.12):

- Calculate the equal-time Green's function $G(\tau_{\rm bp}-\varepsilon)$ for $|\phi_{\rm L}\rangle$ and $|\phi_{\rm R}\rangle$ using Eq. (3.89).
- Calculate the product of propagators $B(\tau_{\rm bp}, \tau_{\rm bp}-\varepsilon)$ by multiplying all the propagators between time $\tau_{\rm bp}$ and $\tau_{\rm bp}-\varepsilon$.
- Calculate the unequal-time Green's function $G(\tau_{\rm bp}, \tau_{\rm bp}-\varepsilon)$ by

$$G(\tau_{\rm bp}, \tau_{\rm bp}-\varepsilon) = B(\tau_{\rm bp}, \tau_{\rm bp}-\varepsilon)G(\tau_{\rm bp}-\varepsilon)\,. \tag{3.110}$$

(5) Back-propagate $|\phi_{\rm R}\rangle$ by an interval ε and forward-propagate $|\phi_{\rm L}\rangle$ also by an interval ε.

$$\Phi_{\rm R} \leftarrow B^{-1}(\tau_{\rm bp}-\varepsilon, \tau_{\rm bp}-2\varepsilon)\Phi_{\rm R} \tag{3.111}$$
$$\Phi_{\rm L} \leftarrow B(\tau_{\rm bp}-\varepsilon, \tau_{\rm bp}-2\varepsilon)\Phi_{\rm L} \tag{3.112}$$

See Fig. 3.13.

(6) Calculate the next small-time unequal-time Green's function as follows (see Fig. 3.14):

- Calculate the equal-time Green's function $G(\tau_{\rm bp}-2\varepsilon)$ for $|\phi_{\rm L}\rangle$ and $|\phi_{\rm R}\rangle$ using Eq. (3.89).

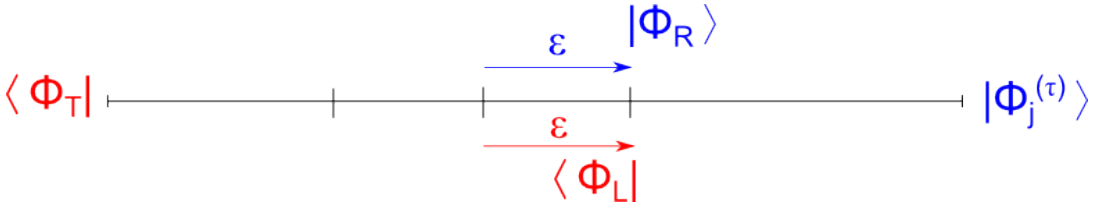


Figure 3.13: Step 5 in the procedure.

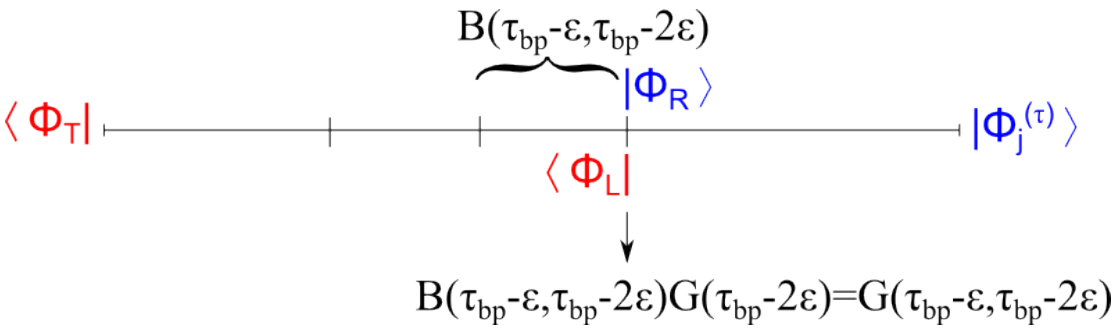


Figure 3.14: Step 6 in the procedure.

- Calculate the product of propagators $B(\tau_{\rm bp}, \tau_{\rm bp} - \varepsilon)$ by multiplying all the propagators between time $\tau_{\rm bp} - \varepsilon$ and $\tau_{\rm bp} - 2\varepsilon$.
- Calculate the unequal-time Green's function $G(\tau_{\rm bp} - \varepsilon, \tau_{\rm bp} - 2\varepsilon)$ by

$$G(\tau_{\rm bp} - \varepsilon, \tau_{\rm bp} - 2\varepsilon) = B(\tau_{\rm bp} - \varepsilon, \tau_{\rm bp} - 2\varepsilon) G(\tau_{\rm bp} - 2\varepsilon)\,. \tag{3.113}$$

(7) Repeat steps 5 to 6 for $(L-2)$ more times.

(8) Calculate the desired unequal-time Green's function as

$$G(\tau_{\rm bp}, 0) = G(\tau_{\rm bp}, \tau_{\rm bp} - \varepsilon) G(\tau_{\rm bp} - \varepsilon, \tau_{\rm bp} - 2\varepsilon) \ldots G(\varepsilon, 0) \tag{3.114}$$

Note that if we store all "small-time" unequal-time Green's functions we can also calculate $G(\tau', 0)$ for $\tau' < \tau_{\rm bp}$ and $\tau' = L'\varepsilon$ by multiplying L' small-time unequal-time Green's functions.

3.12 Other implementation issues

3.12.1 Population control

As the random walk proceeds, some walkers may accumulate very large weights while some will have very small weights. These different weights cause a loss of sampling efficiency because either the algorithm will spend a disproportionate amount of time keeping track of walkers that contribute little to the energy estimate, or the population might perish. Thus a branching scheme is used to "redistribute" the weights without changing the statistical distribution. Walkers with large weights are replicated and walkers with small weights are eliminated with some probability. There are various ways to do branching and the present code uses a simple "combing" method as discussed in Calandra Buonaura and Sorella [37].

3.12.2 Gram-Schmidt reorthonormalization

Repeated multiplications of $B_{\rm K/2}$ and $B_{\rm V}$ to a Slater determinant in Eq. (3.48) lead to numerical instability, such that round-off errors dominate and $|\phi_{\rm j}^{(\tau)}\rangle$ represents an unfaithful propagation of $|\phi_{\rm j}^{(0)}\rangle$. This instability is controlled by periodically applying the modified Gram-Schmidt orthonormalization to each Slater determinant. For each walker $|\phi\rangle$, we factor its corresponding matrix as $\Phi = UDV$ where D is a diagonal matrix, R is an upper

triangular matrix and U is a matrix whose columns are orthonormal vectors representing the re-orthonormalized single-particle orbitals. After this factorization, Φ is replaced by U and the corresponding overlap O_{T} by $O_{\mathrm{T}}/\operatorname{Det}(DV)$. With importance sampling, only the information in U is relevant and D and V can be discarded.

An important point to note here is that this re-orthonormalization does not change the value of any physical observables. See Appendix J for details.

3.13 Algorithm

Finally, we are in a position to state the ground-state CMPC algorithm which contains the back-propagation algorithm in Sections 3.10.2 and 3.11.2. For convenience, in this section, we will use the notation

$$O_{\mathrm{T}} \equiv \langle \phi_{\mathrm{T}} | \phi_{\mathrm{j}}^{(\tau)} \rangle \tag{3.115}$$

to denote the overlap of a walker $|\phi_{\mathrm{j}}^{(\tau)}\rangle$ with the trial wave function $|\phi_{\mathrm{T}}\rangle$.

(1) For each walker, specify its initial state. Here we use the trial wave function Φ_{T} as the initial state and assign the weight w and overlap O_{T} each a value of unity.

(2) If the weight of a walker is nonzero, propagate it via $B_{\mathrm{K}/2}$ as follows:

   (a) Perform the matrix-matrix multiplication

$$\Phi' = B_{\mathrm{K}/2}\Phi \tag{3.116}$$

   (recall the convention that $B_{\mathrm{K}/2}$ denotes the matrix of $\widehat{B}_{\mathrm{K}/2}$) and compute the new overlap as

$$O'_{\mathrm{T}} = O_{\mathrm{T}}(\phi') \,. \tag{3.117}$$

   (b) If $O'_{\mathrm{T}} \neq 0$, update the walker, weight and O_{T} as

$$\Phi \leftarrow \Phi' \,, \quad w \leftarrow w\, O'_{\mathrm{T}}/O_{\mathrm{T}} \,, \quad O_{\mathrm{T}} \leftarrow O'_{\mathrm{T}} \,. \tag{3.118}$$

(3) If the walker's weight is still nonzero, propagate it via $B_{\mathrm{V}}(\vec{x})$ as follows

   (a) Compute the inverse of the overlap matrix

$$O_{\mathrm{inv}} = \left(\Phi_{\mathrm{T}}^{\dagger}\Phi\right)^{-1} \,. \tag{3.119}$$

   (b) For *each* auxiliary field x_i, do the following:

      (i) Compute $\widetilde{p}(x_i)$. See Eq. (3.59).

      (ii) Sample x_i and update the weight as

$$w \leftarrow w\, \mathfrak{N}(\phi_{\mathrm{j},i}^{(\tau)}) \,. \tag{3.120}$$

      See Eq. (3.60).

      (iii) If the weight of the walker is still not zero, propagate the walker by performing the matrix multiplication

$$\Phi' = b_{\mathrm{V}}(x_i)\, \Phi \,. \tag{3.121}$$

      and then update O_{T} and O_{inv}.

(4) Repeat step 2.

(5) Multiply the walker's weight by a normalization factor:

$$w \leftarrow w\, \mathrm{e}^{\Delta\tau E_{\mathrm{T}}} \tag{3.122}$$

where E_{T} is an estimate of the ground state energy E_0.

(6) Repeat steps 2 to 5 for all walkers in the population. This forms one step of the random walk.

(7) If the population of walkers has achieved a steady-state distribution, periodically make measurements of relevant observables using the procedure for back-propagation in Sections 3.10.2 and 3.11.2.

(8) Periodically adjust the population of walkers. (See Section 3.12.1.)

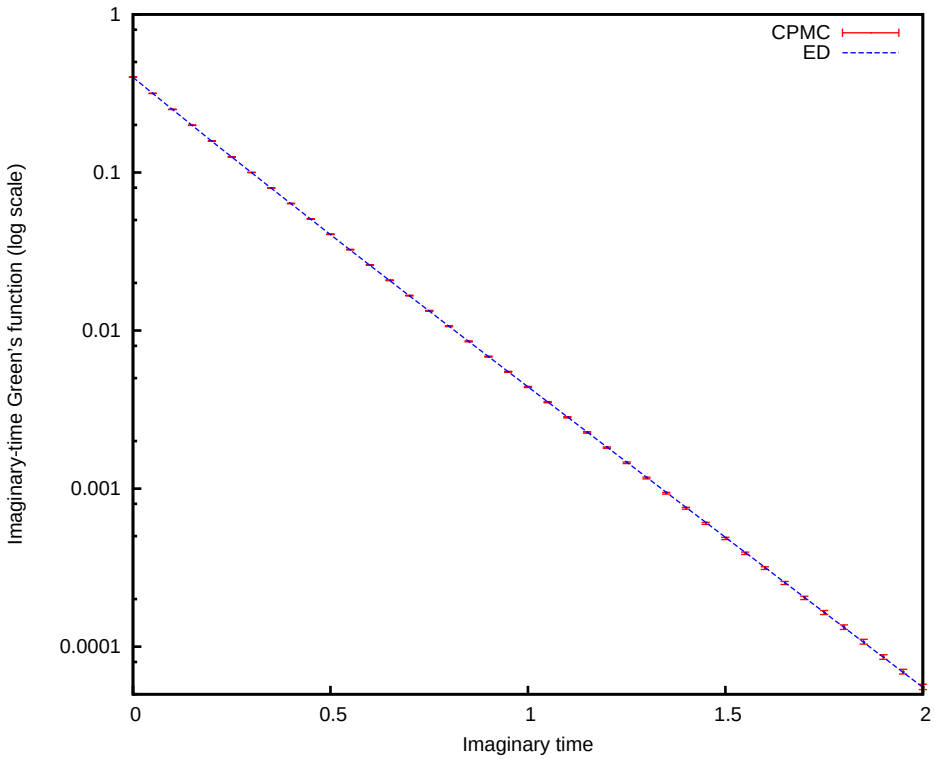
(9) Periodically re-orthonormalize the columns of the matrices Φ representing the walkers. (See Section 3.12.2.)

(10) Repeat this process until an adequate number of measurements have been collected.

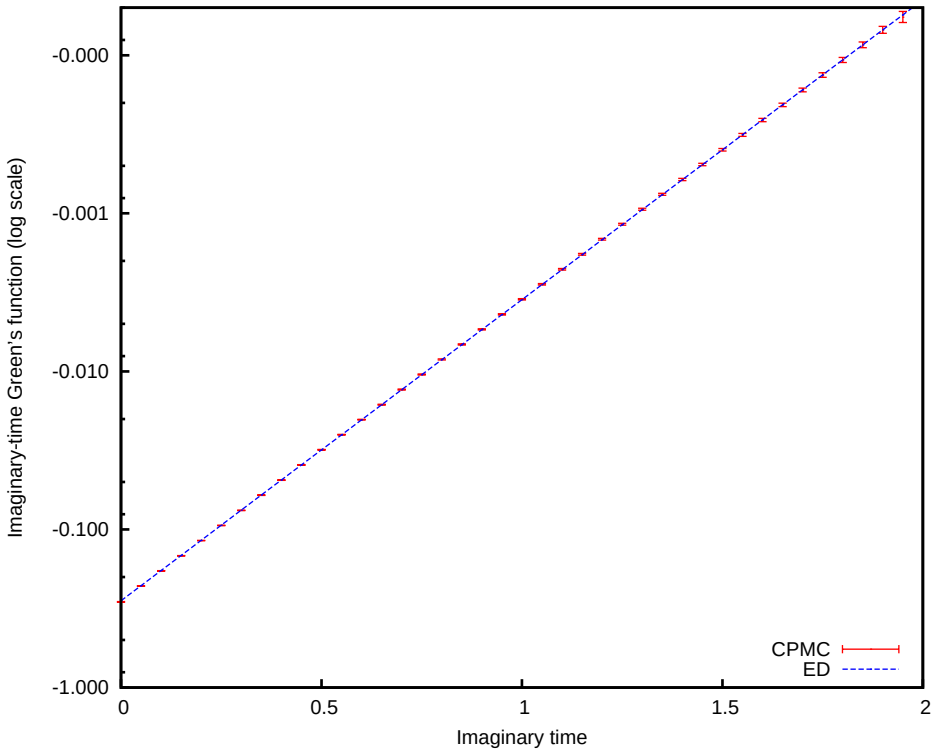
(11) Compute the final average of the measurements of the desired observables and the standard error of this average and then stop.

3.14 Results

Fig. 3.15 compare the CPMC (red with error bars) and ED (blue) calculations of the imaginary-time Green's function $\langle c_{i\sigma}(\tau)c_{j\sigma}(0)\rangle$ against τ from $\tau = 0$ to $\tau = 2$ for two pairs of (i, j). The system is a one-dimensional 5-site lattice with 3 spin-$\uparrow$ and 3 spin-$\downarrow$ electrons with $U = 4$. Although the results are shown for only two pairs of (i, j), the other pairs show the same behavior. The agreement with exact results is excellent.



(a) On-diagonal $\langle c_{i\uparrow}(\tau)c^{\dagger}_{i\uparrow}(0)\rangle$



(b) Off-diagonal $\langle c_{i\uparrow}(\tau)c^{\dagger}_{(i+1)\uparrow}(0)\rangle$

Figure 3.15: Comparison of the ED (blue) and CPMC (red with error bars) calculations of the on- and off-diagonal imaginary-time Green's function vs τ. Note that in the lower figure, the vertical axis indicates $(-1) \times \ln\left[-\langle c_{i\uparrow}(\tau)c^{\dagger}_{(i+1)\uparrow}(0)\rangle\right]$ because the values of $\langle c_{i\uparrow}(\tau)c^{\dagger}_{(i+1)\uparrow}(0)\rangle$ are negative

Chapter 4

Applications

In this chapter, we will show some applications the imaginary-time Green's function.

4.1 Application of the imaginary-time Green's function

In the ground state (i.e. at zero temperature), the dynamical structure factor $S_{\widehat{A},\widehat{B}}(\omega)$ is related to the imaginary-time correlation function by

$$\langle \widehat{A}(\tau)\widehat{B}(0)\rangle = \int_{-\infty}^{\infty} \mathrm{d}\omega\, \mathrm{e}^{-\tau\omega} S_{\widehat{A},\widehat{B}}(\omega) \tag{4.1}$$

where $\langle \widehat{A}(\tau)\widehat{B}(0)\rangle$ in turn can be expressed in terms of unequal-time or equal-time Green's functions using Wick's theorem. An example with the magnetic structure factor is given in Section 4.3. The dynamical structure factor is directly related to quantities that can be measured in experiments. An example with neutron scattering is given in Section 4.2.

Eq. (4.1) is a two-way street. If we have the dynamical structure factor $S(\omega)$ from experiments, we can use Eq. (4.1) to compute the imaginary-time correlation functions and compare to those obtained in theoretical calculations. Conversely, given theoretical calculations of the correlation functions, we can predict the physical properties via the dynamical structure factor. However, the latter procedure requires performing an inverse Laplace transform. There are a number of methods to do this inversion, also known as the *analytic continuation* problem,[1] which are beyond the scope of this thesis. These inversion methods are non-trivial and demand high-quality imaginary-time data which would not be possible without an approach that can control the sign problem, such as the constrained path approximation used in this thesis.

4.2 Neutron scattering by a crystal

In this section, we will show that the neutron scattering cross section is directly proportional to the dynamical structure factor. This section has been adapted from Ashcroft and Mermin [28, pp. 790-2] and Simon [40, pp. 141-4].

[1] There is a sizable literature devoted to this topic. An established and popular method is the Maximum Entropy method as described by Jarrell and Gubernatis [38], a highly recommended reading because of its comprehensive discussion of the relationship between real- and imaginary-time quantities. A more recent inversion method by Vitali *et al.* [39] uses genetic algorithms.

Let a neutron with momentum $\mathbf{p}$ be scattered by a crystal and emerge with momentum $\mathbf{p}'$ and that before the scattering, the ions are in an eigenstate of the crystal Hamiltonian with energy E_i and that after the scattering the ions are in an eigenstate of the crystal Hamiltonian with energy E_f. We describe the initial and final states and energies of the composite neutron-ion system as follows:

Before scattering:

$$\Psi_i = \frac{1}{\sqrt{V}} \mathrm{e}^{\mathrm{i}\mathbf{p}\cdot\mathbf{r}/\hbar} \Phi_i \tag{4.2}$$

$$\varepsilon_i = E_i + \frac{p^2}{2M_n} \,. \tag{4.3}$$

After scattering:

$$\Psi_f = \frac{1}{\sqrt{V}} \mathrm{e}^{\mathrm{i}\mathbf{p}'\cdot\mathbf{r}/\hbar} \Phi_f \tag{4.4}$$

$$\varepsilon_f = E_f + \frac{p'^2}{2M_n} \,, \tag{4.5}$$

where we have assumed the wave functions Ψ_i and Ψ_j are separable as a product of a neutron wave function and ion wave function.

It is convenient to define variables ω and $\mathbf{q}$ in terms of the neutron energy gain and momentum transfer:

$$\hbar\omega = \frac{p'^2}{2M_n} - \frac{p^2}{2M} \tag{4.6}$$

$$\hbar\mathbf{q} = \mathbf{p}' - \mathbf{p} \tag{4.7}$$

It is usually a very good approximation to assume that the scattering potential is the sum over the scattering potentials of the individual atoms in the system so that we can describe the neutron-ion interaction by:

$$U(\mathbf{r}) = \frac{2\pi\hbar^2 a}{M} \sum_j \delta(\mathbf{r} - \mathbf{R}_j) \,. \tag{4.8}$$

The rate at which neutron scatters from $\mathbf{p}$ to $\mathbf{p}'$ by virtue of its interaction with the ions is given by Fermi's "golden rule," a standard result from first-order time-dependent perturbation theory:

$$P = \sum_f \frac{2\pi}{\hbar} \delta(\varepsilon_f - \varepsilon_i) \left| \langle \Psi_i | \widehat{U} | \Psi_f \rangle \right|^2 \,. \tag{4.9}$$

Using Eqs. (4.2), (4.4) and (4.6):

$$P = \sum_f \frac{2\pi}{\hbar} \delta(E_f - E_i + \hbar\omega) \left| \int \mathrm{d}\mathbf{r} \frac{1}{V} \mathrm{e}^{\mathrm{i}(\mathbf{p}'-\mathbf{p})\cdot\mathbf{r}} \, \langle \Phi_i | \widehat{U}(\mathbf{r}) | \Phi_f \rangle \right|^2 \,. \tag{4.10}$$

Using Eqs. (4.7) and (4.8):

$$P = \sum_f \frac{2\pi}{\hbar} \delta(E_f - E_i + \hbar\omega) \left| \int \mathrm{d}\mathbf{r} \frac{1}{V} \mathrm{e}^{\mathrm{i}\mathbf{q}\cdot\mathbf{r}} \left\langle \Phi_i \left| \frac{2\pi\hbar^2 a}{M} \sum_j \delta(\mathbf{r} - \mathbf{R}_j) \right| \Phi_f \right\rangle \right|^2 \,. \tag{4.11}$$

Integrating over $\mathbf{r}$ gives

$$P = \frac{(2\pi\hbar)^3 a^2}{(MV)^2} \sum_f \delta(E_f - E_i + \hbar\omega) \left| \sum_j \left\langle \Phi_i \left| \, \mathrm{e}^{\mathrm{i}\mathbf{q}\cdot\mathbf{R}_j} \, \right| \Phi_f \right\rangle \right|^2 . \tag{4.12}$$

The neutron scattering cross section is proportional to the *dynamical structure factor*, defined as

$$S_i(\mathbf{q},\omega) = \left\{ \frac{1}{N} \sum_f \delta\left(\frac{E_f - E_i}{\hbar} + \omega \right) \left| \sum_j \left\langle \Phi_i \left| \, \mathrm{e}^{\mathrm{i}\mathbf{q}\cdot\mathbf{R}_j} \, \right| \Phi_f \right\rangle \right|^2 \right\} . \tag{4.13}$$

We note that only the index i remains because f has been summed over. To evaluate $S_i(\mathbf{q},\omega)$, we use the following integral representation of the delta function:

$$\delta\left(\frac{E_f - E_i}{\hbar} + \omega \right) = \int_{-\infty}^{\infty} \frac{\mathrm{d}t}{2\pi} \exp\left[\mathrm{i}t \left(\frac{E_f - E_i}{\hbar} + \omega \right) \right] . \tag{4.14}$$

Eq. (4.13) then becomes

$$S_i(\mathbf{q},\omega) = \frac{1}{N} \sum_f \int_{-\infty}^{\infty} \frac{\mathrm{d}t}{2\pi} \mathrm{e}^{\mathrm{i}\omega t} \exp\left[\mathrm{i}t \frac{E_f - E_i}{\hbar} \right] \left| \sum_j \left\langle \Phi_i \left| \, \mathrm{e}^{\mathrm{i}\mathbf{q}\cdot\mathbf{R}_j} \, \right| \Phi_f \right\rangle \right|^2 . \tag{4.15}$$

We can explicitly write out the square:

$$S_i(\mathbf{q},\omega) = \frac{1}{N} \sum_f \int_{-\infty}^{\infty} \frac{\mathrm{d}t}{2\pi} \mathrm{e}^{\mathrm{i}\omega t} \sum_j \left\langle \Phi_i \left| \, \mathrm{e}^{\mathrm{i}\mathbf{q}\cdot\mathbf{R}_j} \, \right| \Phi_f \right\rangle \exp\left[\mathrm{i}t \frac{E_f - E_i}{\hbar} \right] \sum_m \left\langle \Phi_f \left| \, \mathrm{e}^{-\mathrm{i}\mathbf{q}\cdot\mathbf{R}_m} \, \right| \Phi_i \right\rangle . \tag{4.16}$$

We now use the property that

$$\langle \Phi_f | \widehat{A}(t) | \Phi_i \rangle = \exp\left[\mathrm{i}t \frac{(E_f - E_i)}{\hbar} \right] \langle \Phi_f | \widehat{A} | \Phi_i \rangle \tag{4.17}$$

where t is the real time (not the hopping amplitude) and $\widehat{A}(t)$ is the time-dependent operator $\widehat{A}$ in the Heisenberg picture.[2] Plugging this into Eq. (4.16), we obtain

$$S_i(\mathbf{q},\omega) = \frac{1}{N} \sum_f \int_{-\infty}^{\infty} \frac{\mathrm{d}t}{2\pi} \mathrm{e}^{\mathrm{i}\omega t} \sum_j \left\langle \Phi_i \left| \, \mathrm{e}^{\mathrm{i}\mathbf{q}\cdot\mathbf{R}_j} \, \right| \Phi_f \right\rangle \sum_m \left\langle \Phi_f \left| \, \mathrm{e}^{-\mathrm{i}\mathbf{q}\cdot\mathbf{R}_m(t)} \, \right| \Phi_i \right\rangle . \tag{4.18}$$

[2] To see why this is true, use the definition of the time evolution of an operator in the Heisenberg picture:

$$\langle \Phi_f | \widehat{A}(t) | \Phi_i \rangle = \langle \Phi_f | \mathrm{e}^{\mathrm{i}\widehat{H}t/\hbar} \widehat{A} \mathrm{e}^{-\mathrm{i}\widehat{H}t/\hbar} | \Phi_i \rangle$$

Acting with $\mathrm{e}^{-\mathrm{i}\widehat{H}t/\hbar}$ to the right and $\mathrm{e}^{\mathrm{i}\widehat{H}t/\hbar}$ to the left, we have

$$\langle \Phi_f | \widehat{A}(t) | \Phi_i \rangle = \mathrm{e}^{\mathrm{i}E_f t/\hbar} \mathrm{e}^{-\mathrm{i}E_i t/\hbar} \langle \Phi_f | \widehat{A} | \Phi_i \rangle$$

which is the RHS of Eq. (4.17).

Exchange the summations and recognize the completeness relation:

$$S_i(\mathbf{q},\omega) = \frac{1}{N}\int_{-\infty}^{\infty}\frac{\mathrm{d}t}{2\pi}\mathrm{e}^{\mathrm{i}\omega t}\sum_{j,m}\langle\Phi_i|\,\mathrm{e}^{\mathrm{i}\mathbf{q}\cdot\mathbf{R}_j}\left[\sum_f|\Phi_f\rangle\langle\Phi_f|\right]\mathrm{e}^{-\mathrm{i}\mathbf{q}\cdot\mathbf{R}_m(t)}\,|\Phi_i\rangle \tag{4.19}$$

$$= \frac{1}{N}\int_{-\infty}^{\infty}\frac{\mathrm{d}t}{2\pi}\mathrm{e}^{\mathrm{i}\omega t}\sum_{j,m}\langle\Phi_i|\,\mathrm{e}^{\mathrm{i}\mathbf{q}\cdot\mathbf{R}_j}\mathrm{e}^{-\mathrm{i}\mathbf{q}\cdot\mathbf{R}_m(t)}\,|\Phi_i\rangle\ . \tag{4.20}$$

Assume the displacement of the ions about their lattice position is small. Let $\mathcal{R}_i$ be the lattice vector of the i-th ion (so $\mathcal{R}_i$ is constant) and $\mathbf{x}_i$ the ion's displacement so that $\mathbf{R}_i = \mathcal{R}_i + \mathbf{x}_i$. Then

$$S(\mathbf{q},\omega) = \frac{1}{N}\int_{-\infty}^{\infty}\frac{\mathrm{d}t}{2\pi}\mathrm{e}^{\mathrm{i}\omega t}\mathrm{e}^{\mathrm{i}\mathbf{q}\cdot(\mathcal{R}_j-\mathcal{R}_m)}\sum_{j,m}\langle\Phi_0|\,\mathrm{e}^{\mathrm{i}\mathbf{q}\cdot(\mathbf{x}_j-\mathbf{x}_m(t))}\,|\Phi_0\rangle\ , \tag{4.21}$$

where we have replaced the subscript i by 0 to indicate that we are only working in the ground state.

For the sake of completeness, the full relation between the neutron scattering cross section and the dynamical structure factor is

$$\frac{\mathrm{d}\sigma}{\mathrm{d}\Omega\,\mathrm{d}E} = \frac{p'}{p}\frac{Na^2}{\hbar}S(\mathbf{q},\omega) \tag{4.22}$$

where $S(\mathbf{q},\omega)$ is entirely determined by the crystal itself without reference to any properties of the neutrons.

4.3 Magnetic structure factor

In this section, we will show how to obtain an imaginary-time quantity, namely the magnetic structure factor, from the equal-time Green's function. The derivations here have been adapted from Santos [36, p. 43].

First, we give some background discussion. Recall that Wick's theorem allows us to write

$$\langle c^\dagger_{x_1}c_{x_2}c^\dagger_{x_3}c_{x_4}\rangle_{\vec{X}} = \langle c^\dagger_{x_1}c_{x_2}\rangle_{\vec{X}}\langle c^\dagger_{x_3}c_{x_4}\rangle_{\vec{X}} + \langle c^\dagger_{x_1}c_{x_4}\rangle_{\vec{X}}\langle c_{x_2}c^\dagger_{x_3}\rangle_{\vec{X}} \tag{4.23}$$

for a fixed auxiliary-field path $\vec{X}$ where x_1, x_2, x_3 and x_4 are basis space-spin states.

In the notation of second quantization, the operator $\widehat{\mathbf{S}}_i$ that gives the spin in all 3 directions (x, y and z) on site i is

$$\widehat{\mathbf{S}}_i = \frac{1}{2}\sum_{\sigma,\bar{\sigma}=\uparrow,\downarrow}c^\dagger_{i\sigma}\boldsymbol{\sigma}_{\sigma\bar{\sigma}}c_{i\bar{\sigma}} \tag{4.24}$$

where, by convention, we have set $\hbar = 1$ and the components of $\boldsymbol{\sigma}$ are the familiar Pauli matrices:

$$\boldsymbol{\sigma} = (\sigma_x,\sigma_y,\sigma_z)\,. \tag{4.25}$$

For concreteness, we will explicitly write down each component of $\widehat{\mathbf{S}}_i$ in Eq. (4.24):

$$\widehat{S}_i^x = \frac{1}{2}(c_{i\uparrow}^\dagger c_{i\downarrow} + c_{i\downarrow}^\dagger c_{i\uparrow}) \tag{4.26}$$

$$\widehat{S}_i^y = -\frac{\mathrm{i}}{2}(c_{i\uparrow}^\dagger c_{i\downarrow} - c_{i\downarrow}^\dagger c_{i\uparrow}) \tag{4.27}$$

$$\widehat{S}_i^z = \frac{1}{2}(c_{i\uparrow}^\dagger c_{i\uparrow} - c_{i\downarrow}^\dagger c_{i\downarrow})\,. \tag{4.28}$$

Define

$$\widehat{\mathbf{S}}_i = (\widehat{S}_i^x, \widehat{S}_i^y, \widehat{S}_i^z) \tag{4.29}$$

and the dot product in the usual way:

$$\widehat{\mathbf{S}}_i \cdot \widehat{\mathbf{S}}_j = \widehat{S}_i^x \widehat{S}_j^x + \widehat{S}_i^y \widehat{S}_j^y + \widehat{S}_i^z \widehat{S}_j^z\,. \tag{4.30}$$

The static *magnetic structure factor* for a one-dimensional lattice of M sites is defined as

$$S(q) = \frac{1}{M}\sum_{i,j=1}^{M} \mathrm{e}^{\mathrm{i}q(i-j)}\,\langle \widehat{\mathbf{S}}_i \cdot \widehat{\mathbf{S}}_j\rangle \tag{4.31}$$

$$= \frac{1}{M}\sum_{i,j=1}^{M} \mathrm{e}^{\mathrm{i}q(i-j)}\left[\langle \widehat{S}_i^x \widehat{S}_j^x\rangle + \langle \widehat{S}_i^y \widehat{S}_j^y\rangle + \langle \widehat{S}_i^z \widehat{S}_j^z\rangle\right] \tag{4.32}$$

where due to periodic boundary condition, $q = \left(\frac{2\pi}{M}\right) n$ for $n = 0, 1, \ldots, M$. Because $\hbar = 1$, q is both a momentum and a wave vector. We emphasize that the three expectations $\langle \widehat{S}_i^x \widehat{S}_j^x\rangle$, $\langle \widehat{S}_i^y \widehat{S}_j^y\rangle$ and $\langle \widehat{S}_i^z \widehat{S}_j^z\rangle$ must be calculated for each auxiliary field path (i.e. for each walker in each iteration) and averaged over all paths before plugging into Eq. (4.32).

Now we will show how to use Wick's theorem in Eq. (4.23) to write these quantities in terms of Green's function for each auxiliary field path $\overrightarrow{X}$. For example,

$$\langle \widehat{S}_i^x \widehat{S}_j^x\rangle = \frac{1}{4}\left\langle (c_{i\uparrow}^\dagger c_{i\downarrow} + c_{i\downarrow}^\dagger c_{i\uparrow})(c_{j\uparrow}^\dagger c_{j\downarrow} + c_{j\downarrow}^\dagger c_{j\uparrow})\right\rangle \tag{4.33}$$

$$= \frac{1}{4}\left[\langle c_{i\uparrow}^\dagger c_{i\downarrow} c_{j\uparrow}^\dagger c_{j\downarrow}\rangle + \langle c_{i\uparrow}^\dagger c_{i\downarrow} c_{j\downarrow}^\dagger c_{j\uparrow}^\dagger\rangle + \langle c_{i\downarrow}^\dagger c_{i\uparrow} c_{j\uparrow}^\dagger c_{j\downarrow}\rangle + \langle c_{i\downarrow}^\dagger c_{i\uparrow} c_{j\downarrow}^\dagger c_{j\uparrow}\rangle\right] \tag{4.34}$$

$$= \frac{1}{4}\Big\{\Big[\langle c_{i\uparrow}^\dagger c_{i\downarrow}\rangle\,\langle c_{j\uparrow}^\dagger c_{j\downarrow}\rangle + \langle c_{i\uparrow}^\dagger c_{j\downarrow}\rangle\,\langle c_{i\downarrow} c_{j\uparrow}^\dagger\rangle\Big] + \Big[\langle c_{i\uparrow}^\dagger c_{i\downarrow}\rangle\,\langle c_{j\downarrow}^\dagger c_{j\uparrow}\rangle + \langle c_{i\uparrow}^\dagger c_{j\uparrow}\rangle\,\langle c_{i\downarrow} c_{j\downarrow}^\dagger\rangle\Big]$$
$$\Big[\langle c_{i\downarrow}^\dagger c_{i\uparrow}\rangle\,\langle c_{j\uparrow}^\dagger c_{j\downarrow}\rangle + \langle c_{i\downarrow}^\dagger c_{j\downarrow}\rangle\,\langle c_{i\uparrow} c_{j\uparrow}^\dagger\rangle\Big] + \Big[\langle c_{i\downarrow}^\dagger c_{i\uparrow}\rangle\,\langle c_{j\downarrow}^\dagger c_{j\uparrow}\rangle + \langle c_{i\downarrow}^\dagger c_{j\uparrow}\rangle\,\langle c_{i\uparrow} c_{j\downarrow}^\dagger\rangle\Big]\Big\}\,. \tag{4.35}$$

Recall from the discussion in Section 2.2.2 that all correlations involving a spin-up and spin-down operator are zero. That leaves us with only the last term in the second and third square brackets:

$$\langle \widehat{S}_i^x \widehat{S}_j^x\rangle_{\overrightarrow{X}} = \frac{1}{4}\left[\langle c_{i\uparrow}^\dagger c_{j\uparrow}\rangle\,\langle c_{i\downarrow} c_{j\uparrow}^\dagger\rangle + \langle c_{i\downarrow}^\dagger c_{j\downarrow}\rangle\,\langle c_{i\downarrow} c_{j\uparrow}^\dagger\rangle\right] \tag{4.36}$$

$$= \frac{1}{4}\left[\tilde{G}_{ij}^{\uparrow} G_{ij}^{\downarrow} + \tilde{G}_{ij}^{\downarrow} G_{ij}^{\uparrow}\right]_{\overrightarrow{X}} \tag{4.37}$$

where the matrices G and $\tilde{G}$ are related by

$$\tilde{G} = I - G \tag{4.38}$$

and we have inserted a subscript to remind the reader that this equality applies to only a single auxiliary-field path $\vec{X}$. Similarly, we find

$$\langle \widehat{S}_i^y \widehat{S}_j^y \rangle_{\vec{X}} = -\frac{1}{4} \left\langle (c_{i\uparrow}^\dagger c_{i\downarrow} - c_{i\downarrow}^\dagger c_{i\uparrow})(c_{j\uparrow}^\dagger c_{j\downarrow} - c_{j\downarrow}^\dagger c_{j\uparrow}) \right\rangle \tag{4.39}$$

$$= \frac{1}{4} \left[\tilde{G}_{ij}^\uparrow G_{ij}^\downarrow + \tilde{G}_{ij}^\downarrow G_{ij}^\uparrow \right]_{\vec{X}} . \tag{4.40}$$

We make a quick digression to note that here the yy-correlation $\langle \widehat{S}_i^y \widehat{S}_j^y \rangle$ is the same as the xx-correlation. The reason is that correlations between spin-up and spin-down second-quantized operators such as $\langle c_{i\uparrow} c_{j\downarrow}^\dagger \rangle$ are all zero, which in turn happens because the Hubbard Hamiltonian, as it stands, does not flip the spin of electrons (cf. Section 2.2.2). This equality of xx- and yy-correlations will still hold if we add a magnetic field in the z direction because the new Hamiltonian, which has an extra term proportional to $(n_\uparrow - n_\downarrow)$, still does not flip the spins. However, if we add a magnetic field in, say, the x-direction, this equality will not hold anymore[3] because the Hamiltonian will now contain terms like $c_{i\uparrow}^\dagger c_{i\downarrow}$ which flip the spins.

Returning to the main discussion, the last term in Eq. (4.32) can be worked out in a similar way:

$$\langle \widehat{S}_i^z \widehat{S}_j^z \rangle = \frac{1}{4} \left\langle (c_{i\uparrow}^\dagger c_{i\uparrow} - c_{i\downarrow}^\dagger c_{i\downarrow})(c_{j\uparrow}^\dagger c_{j\uparrow} - c_{j\downarrow}^\dagger c_{j\downarrow}) \right\rangle \tag{4.41}$$

$$= \frac{1}{4} \left[\tilde{G}_{ii}^\uparrow \tilde{G}_{jj}^\uparrow + \tilde{G}_{ij}^\uparrow G_{ij}^\uparrow - \tilde{G}_{ii}^\uparrow \tilde{G}_{jj}^\downarrow - \tilde{G}_{ii}^\downarrow \tilde{G}_{jj}^\uparrow + \tilde{G}_{ii}^\downarrow \tilde{G}_{jj}^\downarrow + \tilde{G}_{ij}^\downarrow G_{ij}^\downarrow \right] . \tag{4.42}$$

Shown in Fig. 4.1 is the graph of $S(q)$ vs. q for a 37-site one-dimensional lattice at half-filling (19 spin-$\uparrow$ and 19 spin-$\downarrow$) with $U = 4.0$. The peak at $q = \pi$ means that the dominant wave vector is

$$\frac{2\pi}{\lambda} = \pi \tag{4.43}$$

or

$$\lambda = 2 \tag{4.44}$$

Because we have set the lattice spacing to 1, this means that the spin pattern repeats every other lattice site as shown in Fig. 4.2 and implies that the Hubbard model in one dimension at half filling is an antiferromagnet, consistent with the literature.

[3] A more physical way to see why the equality does not hold is by noting that a magnetic field in the x-direction will break the equivalence between the x- and y-direction in space.

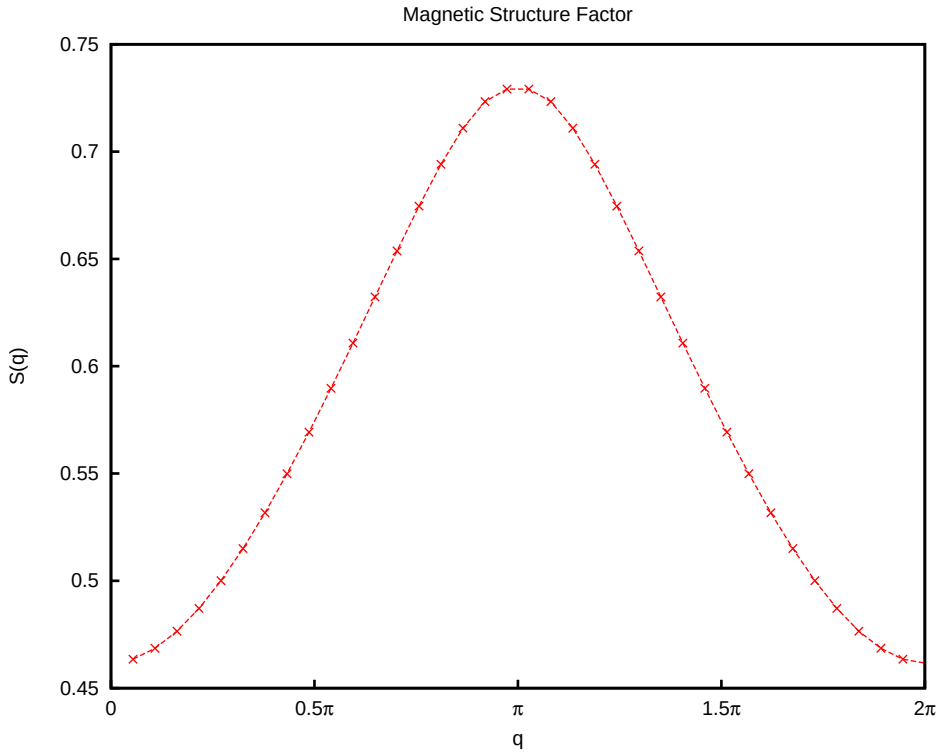


Figure 4.1: The magnetic structure factor for a 37-site one-dimensional lattice at half-filling (19 spin-↑ and 19 spin-↓) with $U = 4.0$.

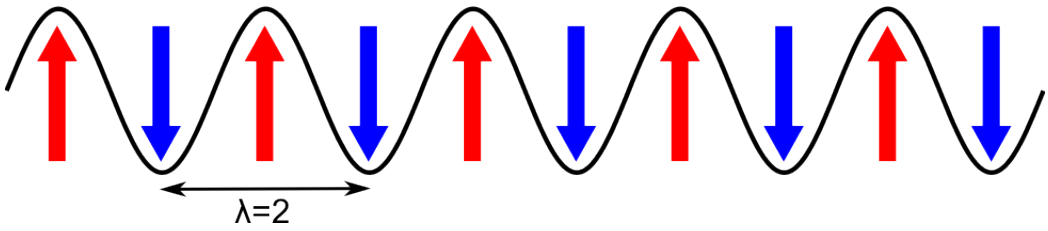


Figure 4.2: The antiferromagnetic ordering of spin-up (red) and spin-down (blue) electrons corresponding to $q = \pi$ on adjacent lattice sites in one dimension.

Appendix A

Hubbard-Stratonovich transformation

This appendix proves and discusses the general Hubbard-Stratonovich transformation and has been adapted from Hirsch [17] and Santos [36]. The transformation changes the exponential of the square of any operator $\widehat{A}$ into an exponential of the operator itself:

$$\mathrm{e}^{\frac{1}{2}\widehat{A}^2} = \int_{-\infty}^{\infty} \mathrm{d}x \left(\frac{\mathrm{e}^{-\frac{1}{2}x^2}}{\sqrt{2\pi}} \right) \mathrm{e}^{-x\widehat{A}} \tag{A.1}$$

at the price of an *auxiliary field* x, which is just a scalar that obeys a Gaussian distribution and couples linearly to $\widehat{A}$. This result can be easily seen by writing the exponential on the RHS as a series:

$$\int_{-\infty}^{\infty} \mathrm{d}x \left(\frac{\mathrm{e}^{-\frac{1}{2}x^2}}{\sqrt{2\pi}} \right) \mathrm{e}^{-x\widehat{A}} = \int_{-\infty}^{\infty} \mathrm{d}x \left(\frac{\mathrm{e}^{-\frac{1}{2}x^2}}{\sqrt{2\pi}} \right) \left[\sum_{n=0}^{\infty} \frac{1}{n!} \left(-x\widehat{A} \right)^n \right] . \tag{A.2}$$

To quote my math professor Jerry Shurman, "if you have a sum and an integral, you *must* exchange them."

$$\int_{-\infty}^{\infty} \mathrm{d}x \left(\frac{\mathrm{e}^{-\frac{1}{2}x^2}}{\sqrt{2\pi}} \right) \mathrm{e}^{-x\widehat{A}} = \sum_{n=0}^{\infty} \frac{1}{n!} \left[\int_{-\infty}^{\infty} \mathrm{d}x \left(\frac{\mathrm{e}^{-\frac{1}{2}x^2}}{\sqrt{2\pi}} \right) \left(-x\widehat{A} \right)^n \right] . \tag{A.3}$$

For n odd, the integrand in the square bracket is an odd function and the symmetric integral evaluates to zero. Thus, only terms with even n remain:

$$\int_{-\infty}^{\infty} \mathrm{d}x \left(\frac{\mathrm{e}^{-\frac{1}{2}x^2}}{\sqrt{2\pi}} \right) \mathrm{e}^{-x\widehat{A}} = \sum_{n=0}^{\infty} \frac{1}{(2n)!} \left[\int_{-\infty}^{\infty} \mathrm{d}x \left(\frac{\mathrm{e}^{-\frac{1}{2}x^2}}{\sqrt{2\pi}} \right) x^{2n} \right] \widehat{A}^{2n} . \tag{A.4}$$

Consulting the formula for even-powered Gaussian integrals in Griffiths [41, backcover], we have

$$\int_{-\infty}^{\infty} \mathrm{d}x \left(\frac{\mathrm{e}^{-\frac{1}{2}x^2}}{\sqrt{2\pi}} \right) \mathrm{e}^{-x\widehat{A}} = \sum_{n=0}^{\infty} \frac{1}{\cancel{(2n)!}} \left[\frac{\cancel{(2n)!}}{n!} \left(\frac{1}{\sqrt{2}} \right)^{2n} \right] \widehat{A}^{2n} \tag{A.5}$$

$$= \sum_{n=0}^{\infty} \frac{1}{n!} \left(\frac{\widehat{A}^2}{2} \right)^n \tag{A.6}$$

$$\int_{-\infty}^{\infty} \mathrm{d}x \left(\frac{\mathrm{e}^{-\frac{1}{2}x^2}}{\sqrt{2\pi}} \right) \mathrm{e}^{-x\widehat{A}} = \mathrm{e}^{\frac{1}{2}\widehat{A}^2} \tag{A.7}$$

as desired.

The HS transformation has been widely used to study models in which the two-body interaction on each site i has the form

$$\widehat{V}_i = U n_{i\uparrow} n_{i\downarrow} \tag{A.8}$$

like the Hubbard model. However, before a transformation of this kind can be applied to $\mathrm{e}^{\widehat{H}_i}$, squares of operators must appear in the argument of the exponential. There are many ways to transform the interaction in Eq. (A.8) to produce the square of an operator.[1] This operator can be either the local magnetic spin $(n_{i\uparrow} - n_{i\downarrow})$ or the local charge $(n_{i\uparrow} + n_{i\downarrow})$ on each site:

$$n_{i\uparrow} n_{i\downarrow} = -\frac{1}{2}(n_{i\uparrow} - n_{i\downarrow})^2 + \frac{1}{2}(n_{i\uparrow} + n_{i\downarrow}) \,, \tag{A.9}$$

$$n_{i\uparrow} n_{i\downarrow} = +\frac{1}{2}(n_{i\uparrow} + n_{i\downarrow})^2 - \frac{1}{2}(n_{i\uparrow} + n_{i\downarrow}) \,, \tag{A.10}$$

$$n_{i\uparrow} n_{i\downarrow} = +\frac{1}{4}(n_{i\uparrow} + n_{i\downarrow})^2 - \frac{1}{4}(n_{i\uparrow} - n_{i\downarrow})^2 \,, \tag{A.11}$$

and can be proved easily by noting that the number operators are idempotent i.e. $n_{i\uparrow}^2 = n_{i\uparrow}$. We can make a few immediate observations about Eqs. (A.9) to (A.11) in relation to the HS transformation:

(i) An auxiliary field will be introduced for each squared operator appearing on the RHS's above. Thus Eqs. (A.9) and (A.10) introduce a single auxiliary field while Eq. (A.11) introduces two auxiliary fields for each lattice site.

(ii) The auxiliary field will couple to the local magnetic spin and local charge when Eq. (A.9) and Eq. (A.10) are used, respectively.

(iii) Because the second terms on the RHS's of Eqs. (A.9) and (A.10) are not squares, they do not produce auxiliary fields and hence require no sampling in QMC and are usually combined into the non-interacting term in the Hamiltonian.

Eq. (A.1) can now be used with Eqs. (A.9) to (A.11) to eliminate the interactions. For example, if Eq. (A.10) is used then

$$\mathrm{e}^{-\Delta\tau V_i} = \mathrm{e}^{U\Delta\tau(n_{i\uparrow}+n_{i\downarrow})/2} \mathrm{e}^{-U\Delta\tau(n_{i\uparrow}+n_{i\downarrow})^2/2} \tag{A.12}$$

$$\mathrm{e}^{-\Delta\tau V_i} = \mathrm{e}^{U\Delta\tau(n_{i\uparrow}+n_{i\downarrow})/2} \int \mathrm{d}x_i \left(\frac{\mathrm{e}^{-\frac{1}{2}{x_i}^2}}{\sqrt{2\pi}} \right) \mathrm{e}^{\sqrt{-U\Delta\tau}\, x_i (n_{i\uparrow}+n_{i\downarrow})} \tag{A.13}$$

[1] Although this appendix only deals with Hubbard-like $\widehat{V}$, the writing of $\widehat{V}$ as a sum of squares of one-body operators can be done for any general two-body interactions. See Appendix K for details.

where we have used Eq. (A.1) by letting

$$\widehat{A} = \sqrt{-\Delta\tau U}(n_{i\uparrow} + n_{i\downarrow})\,. \tag{A.14}$$

Eq. (A.13) is called the continuous Gaussian charge transformation because the auxiliary field x_i obeys a continuous Gaussian distribution and the squared operator is the local charge. Note that from Eq. (A.14), this transformation will produce complex quantities if U is positive.

Since the fermion occupation number can only take the values 0 or 1, Hirsch [17] pointed out that an Ising-like auxiliary field that takes only two values $x_i = \pm 1$ is sufficient to eliminate the fermion interaction. The transformation he proposed is called the discrete Hirsch spin transformation

$$\mathrm{e}^{-\Delta\tau U n_{i\uparrow} n_{i\downarrow}} = \frac{1}{2}\mathrm{e}^{-\Delta\tau U(n_{i\uparrow}+n_{i\downarrow})/2} \sum_{x_i=\pm 1} \mathrm{e}^{x_i\gamma(n_{i\uparrow}-n_{i\downarrow})} \tag{A.15}$$

where $\cosh\gamma = \mathrm{e}^{\Delta\tau U/2}$. The transformation is so named because the auxiliary field x_i on each site is discrete and couples to the magnetic spin $n_{i\uparrow} - n_{i\downarrow}$. For a proof of Eq. (A.15), please refer to Appendix B. Hirsch [17] also invented a slightly different form of the HS transformation, called the discrete Hirsch charge transformation:

$$\mathrm{e}^{-\Delta\tau U n_\uparrow n_\downarrow} = \frac{1}{2}\mathrm{e}^{-\Delta\tau U(n_{i\uparrow}+n_{i\downarrow}-1)/2} \sum_{x_i=\pm 1} \mathrm{e}^{\gamma x_i(n_{i\uparrow}+n_{i\downarrow}-1)} \tag{A.16}$$

where $\cosh\gamma = \mathrm{e}^{-\Delta\tau U/2}$. Because the arccosh function is complex-valued for real inputs with magnitude less than 1, if one wants to work with only real quantities, Eq. (A.15) would be used for the repulsive Hubbard model ($U > 0$) and Eq. (A.16) for the attractive model ($U < 0$).

We want to mention in passing an alternative [42, p. 110] to Eq. (A.1) that could be useful when one wants a HS transformation with discrete auxiliary fields for a general non-Hubbard-like interaction (i.e. not of the form in Eq. (A.8):

$$\mathrm{e}^{\Delta\tau W\widehat{O}^2} = \sum_{\ell=\pm 1,\pm 2} \gamma(\ell)\mathrm{e}^{\sqrt{\Delta\tau W}\eta(\ell)\widehat{O}} + \mathcal{O}(\Delta\tau^4) \tag{A.17}$$

where the coefficients are

$$\gamma(\pm 1) = 1 + \sqrt{6}/3 \qquad \gamma(\pm 2) = 1 - \sqrt{6}/3 \tag{A.18}$$

and

$$\eta(\pm 1) = \pm\sqrt{2(3-\sqrt{6})} \qquad \eta(\pm 2) = \pm\sqrt{2(3+\sqrt{6})}\,. \tag{A.19}$$

Because the per-time-step error is $\mathcal{O}(\Delta\tau^4)$, the error over the whole projection path is $\mathcal{O}(\Delta\tau^3)$ (see Eq. (3.30)) which is not a problem because the Trotter-Suzuki decomposition in Section 3.4 already introduces an error of order $\mathcal{O}(\Delta\tau^2)$.

Appendix B

Hirsh transformations

This section proves the discrete Hirsch spin transformation as stated in Eq. (3.32),

$$e^{-\Delta\tau U n_{i\uparrow} n_{i\downarrow}} = \mathrm{e}^{-\Delta\tau U(n_{i\uparrow}+n_{i\downarrow})/2} \sum_{x_i=\pm 1} p(x_i)\, \mathrm{e}^{\gamma x_i (n_{i\uparrow}-n_{i\downarrow})}\,, \tag{B.1}$$

where γ is given by $\cosh(\gamma) = \exp(\Delta\tau U/2)$ and $p(x_i) = 1/2$ for $x_i = \pm 1$.

First, let $\widehat{H} = (n_\uparrow - n_\downarrow)$. We want an expression for $\mathrm{e}^{\alpha \widehat{H}^2}$, which contains the square of $\widehat{H}$, in terms of $\mathrm{e}^{\beta \widehat{H}}$, which contains $\widehat{H}$ itself.

We make an elementary observation that if an arbitrary quantum operator $\widehat{H}$ has eigenvalues 0, 1 and -1 with corresponding eigenstates $|0_1\rangle$,$|0_2\rangle$ (degenerate), $|1\rangle$ and $|-1\rangle$, its square $\widehat{H}^2$ must have only eigenvalues 0 with degenerate eigenstates $|0_1\rangle$ and $|0_2\rangle$ and 1 with degenerate eigenstates $|1\rangle$ and $|-1\rangle$. Also, if an operator $\widehat{H}$ has eigenvalue a with eigenstate $|a\rangle$ then $\mathrm{e}^{\widehat{H}}$ has eigenvalue e^a because:

$$\mathrm{e}^{\widehat{H}} |a\rangle = \left(1 + \widehat{H} + \frac{\widehat{H}^2}{2} + \dots\right) |a\rangle \tag{B.2}$$

$$= |a\rangle + a\,|a\rangle + \frac{a^2}{2}\,|a\rangle + \dots \tag{B.3}$$

$$= \mathrm{e}^a\,|a\rangle\ . \tag{B.4}$$

Using the completeness relation $H = \sum_i i\,|i\rangle\langle i|$ where i are the eigenvalues, we obtain:

$$\mathrm{e}^{\alpha\widehat{H}^2} = |0_1\rangle\langle 0_1| + |0_2\rangle\langle 0_2| + \mathrm{e}^{\alpha}(\,|1\rangle\langle 1| + |-1\rangle\langle -1|\,) \tag{B.5}$$

$$\mathrm{e}^{\beta\widehat{H}} = |0_1\rangle\langle 0_1| + |0_2\rangle\langle 0_2| + \mathrm{e}^{\beta}\,|1\rangle\langle 1| + \mathrm{e}^{-\beta}\,|-1\rangle\langle -1| \tag{B.6}$$

$$\mathrm{e}^{-\beta\widehat{H}} = |0_1\rangle\langle 0_1| + |0_2\rangle\langle 0_2| + \mathrm{e}^{-\beta}\,|1\rangle\langle 1| + \mathrm{e}^{\beta}\,|-1\rangle\langle -1|\ . \tag{B.7}$$

This is easy to see by noting that $\mathrm{e}^0 = 1$. Add Eqs. (B.6) and (B.7) and multiply by 1/2:

$$\frac{1}{2}\left(\mathrm{e}^{\beta\widehat{H}} + \mathrm{e}^{-\beta\widehat{H}}\right) = |0_1\rangle\langle 0_1| + |0_2\rangle\langle 0_2| + +\frac{1}{2}\left(\mathrm{e}^{\beta} + \mathrm{e}^{-\beta}\right)(|1\rangle\langle 1| + |-1\rangle\langle -1|)\ . \tag{B.8}$$

For Eq. (B.8) to equal Eq. (B.5), we need

$$\mathrm{e}^{\alpha} = \frac{1}{2}\left(\mathrm{e}^{\beta} + \mathrm{e}^{-\beta}\right) = \cosh\beta\,. \tag{B.9}$$

Thus if we require $\cosh\beta = \mathrm{e}^{\alpha}$ then

$$\mathrm{e}^{\alpha\widehat{H}^2} = \frac{1}{2}\left(\mathrm{e}^{\beta\widehat{H}} + \mathrm{e}^{-\beta\widehat{H}}\right) . \tag{B.10}$$

Now let's look at the product of particles on the same lattice site $n_\uparrow n_\downarrow$. Because

$$(n_\uparrow - n_\downarrow)^2 = (n_\uparrow)^2 + (n_\downarrow)^2 - 2n_\uparrow n_\downarrow \quad \text{number operators commute} \tag{B.11}$$

$$= n_\uparrow + n_\downarrow - 2n_\uparrow n_\downarrow , \quad \text{number operators are idempotent} \tag{B.12}$$

we have:

$$n_\uparrow n_\downarrow = \frac{(n_\uparrow + n_\downarrow)}{2} - \frac{(n_\uparrow - n_\downarrow)^2}{2} . \tag{B.13}$$

Thus

$$\mathrm{e}^{-\Delta\tau U n_\uparrow n_\downarrow} = \mathrm{e}^{-\Delta\tau U(n_\uparrow + n_\downarrow)/2}\, \mathrm{e}^{\Delta\tau U\,(n_\uparrow - n_\downarrow)^2/2} . \tag{B.14}$$

Because there are only four possible states on each lattice site, namely $|\cdot\rangle$, $|\uparrow\rangle$, $|\downarrow\rangle$ and $|\uparrow\downarrow\rangle$, the operator $(n_\uparrow - n_\downarrow)$ can only have three eigenvalues 0, 1 and -1. Letting $\widehat{H} = (n_\uparrow - n_\downarrow)$ and $\alpha = \frac{\Delta\tau U}{2}$ in Eq. (B.10), we have[1]

$$\mathrm{e}^{\Delta\tau U\,(n_\uparrow - n_\downarrow)^2/2} = \frac{1}{2}\left(\mathrm{e}^{\gamma(n_\uparrow - n_\downarrow)} + \mathrm{e}^{-\gamma(n_\uparrow - n_\downarrow)}\right) , \tag{B.22}$$

provided that

$$\cosh\gamma = \mathrm{e}^{\Delta\tau U/2} \tag{B.23}$$

[1] As a side note, the form of the Hirsch spin transformation in Eq. (B.1) is not the same as that in the original paper by Hirsch [17]:

$$\mathrm{e}^{-\Delta\tau U n_\uparrow n_\downarrow} = \frac{1}{2}\mathrm{e}^{-\Delta\tau U(n_\uparrow + n_\downarrow)/2} \sum_{x=\pm 1} \mathrm{e}^{2ax(n_\uparrow - n_\downarrow)} \tag{B.15}$$

where the constant a satisfies:

$$\tanh^2 a = \tanh(\Delta\tau U/4) \tag{B.16}$$

However, Eq. (B.15) can be proved to be equivalent to Eq. (A.15) by noting that $\cosh(2a) = \mathrm{e}^{\Delta\tau U/2}$ implies Eq. (B.16) by simple manipulations with hyperbolic functions, and hence $2a$ in Eq. (B.15) is the same as γ in Eq. (A.15). It is easy to show that

$$\cosh^2 a = \frac{1 + \mathrm{e}^{\Delta\tau U/2}}{2} \tag{B.17}$$

$$\sinh^2 a = \frac{\mathrm{e}^{\Delta\tau U/2} - 1}{2} : \tag{B.18}$$

Together, Eqs. (B.17) and (B.18) give

$$\tanh^2 a = \tanh(\Delta\tau U/4) \tag{B.19}$$

For completeness, the original form of the Hirsch charge transformation in Eq. (A.16) is

$$\mathrm{e}^{-\Delta\tau U n_\uparrow n_\downarrow} = \frac{1}{2}\mathrm{e}^{-\Delta\tau U(n_\uparrow + n_\downarrow)/2} \sum_{x=\pm 1} \mathrm{e}^{2bx(n_\uparrow + n_\downarrow - 1)} \tag{B.20}$$

where

$$\tanh^2 b = -\tanh(\Delta\tau U/4) \tag{B.21}$$

Substituting Eq. (B.22) into Eq. (B.14), we obtain the desired expression:

$$\mathrm{e}^{-\Delta\tau U n_\uparrow n_\downarrow} = \frac{1}{2}\mathrm{e}^{-\Delta\tau U(n_\uparrow+n_\downarrow)/2}\left(\mathrm{e}^{\gamma(n_\uparrow-n_\downarrow)} + \mathrm{e}^{-\gamma(n_\uparrow-n_\downarrow)}\right) \tag{B.24}$$

$$= \frac{1}{2}\,\mathrm{e}^{-\Delta\tau U(n_\uparrow+n_\downarrow)/2}\sum_{x=\pm 1}\mathrm{e}^{x\gamma(n_\uparrow-n_\downarrow)}\,, \tag{B.25}$$

which is the expression that we set out to prove.

Appendix C

Overlap of two Slater determinants

In this appendix, we want to prove the formula for the overlap of two Slater determinants as stated in Eq. (3.8):

$$\langle\phi|\phi'\rangle = \mathrm{Det}\left(\Phi^\dagger\Phi'\right) . \tag{C.1}$$

Suppose we have two Slater determinants

$$|\phi'\rangle = \prod_{m=1}^{N}\left(\sum_{\ell=1}^{L} B_{\ell m} c_\ell^\dagger\right)|0\rangle \tag{C.2}$$

$$= \left(\sum_{\ell_1} B_{\ell_1,1} c_{\ell_1}^\dagger\right)\left(\sum_{\ell_2} B_{\ell_2,2} c_{\ell_2}^\dagger\right)\ldots\left(\sum_{\ell_N} B_{\ell_N,N} c_{\ell_N}^\dagger\right)|0\rangle \tag{C.3}$$

$$|\phi\rangle = \prod_{m=1}^{N}\left(\sum_{k=1}^{L} A_{km} c_k^\dagger\right)|0\rangle \tag{C.4}$$

$$= \left(\sum_{k_1} A_{k_1,1} c_{k_1}^\dagger\right)\left(\sum_{k_2} A_{k_2,2} c_{k_2}^\dagger\right)\ldots\left(\sum_{k_N} A_{k_N,N} c_{k_N}^\dagger\right)|0\rangle . \tag{C.5}$$

The adjoint of $|\phi\rangle$ is

$$\langle\phi| = \langle 0|\left(\sum_{k_N} A_{N,k_N}^\dagger c_{k_N}\right)\left(\sum_{k_{N-1}} A_{(N-1),k_{N-1}}^\dagger c_{k_{N-1}}\right)\ldots\left(\sum_{k_1} A_{1,k_1}^\dagger c_{k_1}\right) . \tag{C.6}$$

with the sums in $|\phi\rangle$ applied in reverse order. We want to prove

$$\langle\phi|\phi'\rangle = \mathrm{Det}(A^\dagger B) . \tag{C.7}$$

By taking the inner product of Eqs. (C.3) and (C.6), the LHS of Eq. (C.7) is

$$\langle\phi|\phi'\rangle = \sum_{\{\ell_1,\ldots,\ell_N,k_1,\ldots,k_N\}} A_{1,k_1}^\dagger\ldots A_{N,k_N}^\dagger B_{\ell_1,1}\ldots B_{\ell_N,N}\,\langle 0|\,c_{k_N}\ldots c_{k_1}c_{\ell_1}^\dagger\ldots c_{\ell_N}^\dagger\,|0\rangle \tag{C.8}$$

where the sum is over all the possible combinations of $\{\ell_1, \ldots, \ell_N, k_1, \ldots, k_N\}$. Note that each term in this sum is non-zero only when there is no repetition among the $k_1, \ldots, k_N$ because otherwise, we'll have 2 applications of the same annihilation operator c_j in a row, which certainly gives zero. Furthermore, the $\ell_1, \ldots, \ell_N$ must be some permutation of the $k_1, \ldots, k_N$ because otherwise, some annihilation operator will try to remove an orbital that has not previously been created by a creation operator, which also gives zero. Note that if all the ℓ_i's are some permutation of the k_j's then the inner product $\langle 0| c_{k_N} \ldots c_{k_1} c^\dagger_{\ell_1} \ldots c^\dagger_{\ell_N} |0\rangle$ is ± 1 depending on the permutation P.

$$\langle A|B\rangle = \sum_P (-1)^P \sum_{\text{distinct } k_1,\ldots,k_N} A^\dagger_{1,k_1} \ldots A^\dagger_{N,k_N} B_{P(k_1),1} \ldots B_{P(k_N),N} \tag{C.9}$$

$$= \sum_P (-1)^P \sum_{\text{distinct } k_1,\ldots,k_N}^{L} A^\dagger_{1,k_1} \ldots A^\dagger_{N,k_N} B_{k_1,P(1)} \ldots B_{k_N,P(N)} \tag{C.10}$$

The second equality follows because $\{P(1), \ldots, P(N)\}$ and $\{P(k_1), \ldots, P(k_N)\}$ are both permutations of a set of N elements. The outer sum is over all the permutations P and the inner sum has no repetition among the $k_1, \ldots, k_N$.

On the other hand, the RHS of Eq. (C.7) is

$$\mathrm{Det}(A^\dagger B) = \sum_P (-1)^P (A^\dagger B)_{1,P(1)} \ldots (A^\dagger B)_{N,P(N)} \tag{C.11}$$

$$= \sum_P (-1)^P \left(\sum_{k_1} A^\dagger_{1,k_1} B_{k_1,P(1)} \right) \ldots \left(\sum_{k_N} A^\dagger_{N,k_N} B_{k_N,P(N)} \right) \tag{C.12}$$

$$= \sum_{\substack{k1,\ldots,kN \\ \text{permutations } P}} (-1)^P \left(A^\dagger_{1,k_1} B_{k_1,P(1)} \right) \ldots \left(A^\dagger_{N,k_N} B_{k_N,P(N)} \right) . \tag{C.13}$$

All the indices in Eq. (C.13) must be distinct because if $k_i = k_j = k$ for some term

$$(-1)^P \left(A^\dagger_{i,k} B_{k,P(i)} \right) \ldots \left(A^\dagger_{j,k} B_{k,P(j)} \right) \ldots \tag{C.14}$$

in the summation, then it can easily be seen that there exists another term

$$(-1)^Q \left(A^\dagger_{i,k} B_{k,Q(i)} \right) \ldots \left(A^\dagger_{j,k} B_{k,Q(j)} \right) \ldots \tag{C.15}$$

where the permutation Q satisfies $Q(i) = P(j)$ and $Q(j) = P(i)$. Because the permutations Q and P have opposite signs,[1] the two terms in Eqs. (C.14) and (C.15) cancel. Thus, all the indices in Eq. (C.13) must be distinct.

$$\mathrm{Det}(A^\dagger B) = \sum_P (-1)^P \sum_{\text{distinct } k_1,\ldots,k_N}^{L} \left(A^\dagger_{1,k_1} B_{k_1,P(1)} \right) \ldots \left(A^\dagger_{N,k_N} B_{k_N,P(N)} \right) . \tag{C.16}$$

Because the RHS's of Eqs. (C.10) and (C.16) are identical, Eq. (C.7) has been proved.

[1] Let $Q(i) = P(j) = x$ and $Q(j) = P(i) = y$ Without loss of generality, suppose $x < y$. It takes $(y - x - 1)$ transpositions to bring x to y's position and then it takes $(y - x - 1 - 1)$ transpositions to bring y back to x's original position for a total of $2(y - x - 1) - 1$ transpositions, an odd number.

Appendix D

Thouless theorem

We want to prove Thouless[1] theorem as stated in Eqs. (3.9) to (3.12). Suppose we have a Slater determinant

$$|\phi\rangle \equiv \widehat{\varphi}_1^\dagger \widehat{\varphi}_2^\dagger \cdots \widehat{\varphi}_N^\dagger |0\rangle \tag{D.1}$$

where the operator

$$\widehat{\varphi}_m^\dagger \equiv \sum_i c_i^\dagger \varphi_{i,m} \tag{D.2}$$

creates an electron in the m-th single-particle orbital. Thouless theorem states that the operation of the exponential of a one body operator

$$\widehat{B} = \exp\left(\sum_{ij}^{M} c_i^\dagger T_{ij} c_j\right) \tag{D.3}$$

simply leads to another Slater determinant:

$$\widehat{B}\,|\phi\rangle = \widehat{\phi}_1'^\dagger \widehat{\phi}_2'^\dagger \cdots \widehat{\phi}_M'^\dagger\, |0\rangle \equiv |\phi'\rangle \tag{D.4}$$

where the operators are

$$\widehat{\phi}_m'^\dagger = \sum_j c_j^\dagger \Phi'_{jm} \tag{D.5}$$

like in Eq. (D.2) and the matrix Φ' is related to Φ in a simple way:

$$\Phi' \equiv \mathrm{e}^T \Phi \tag{D.6}$$

where the matrix T is formed from elements T_{ij}.

Let's define a Slater determinant

$$|\Psi\rangle = \prod_{i=1}^{N} \left(\sum_{j=1}^{L} \phi_{ji} c_j^\dagger\right) |0\rangle \tag{D.7}$$

where we have N fermions described by a basis consisting of L states. In the case of the Hubbard model, the basis states are chosen to be Wannier wave functions localized on each lattice site. We must have

$$L \geq N \tag{D.8}$$

[1] Fun fact: my advisor, Prof. Darrell Schroeter, attended Reed with Thouless's daughter.

due to the Pauli exclusion principle. All one-body operators in quantum mechanics have the form

$$\hat{H} = \sum_{i,j=1}^{L} T_{ij} c_i^\dagger c_j \tag{D.9}$$

where L is the number of basis states. This can be viewed as a matrix multiplication

$$\hat{H} = \underbrace{\begin{pmatrix} c_1^\dagger & c_2^\dagger & \dots & c_L^\dagger \end{pmatrix}}_{\vec{c}^\dagger} \boldsymbol{T} \underbrace{\begin{pmatrix} c_1 \\ c_2 \\ \dots \\ c_L \end{pmatrix}}_{\vec{c}} \tag{D.10}$$

where $\boldsymbol{T}$ is a $L \times L$ Hermitian matrix. Since $\widehat{H}$ is Hermitian,[2] it puts a restriction on the elements of T so that we can diagonalize T with a unitary matrix U

$$\widehat{H} = \vec{c}^\dagger \boldsymbol{T} \vec{c} = \underbrace{\vec{c}^\dagger \boldsymbol{U}^\dagger}_{\boldsymbol{\beta}^\dagger} \boldsymbol{D} \underbrace{\boldsymbol{U} \vec{c}}_{\boldsymbol{\beta}} \tag{D.11}$$

where we have defined new vectors:

$$\vec{\beta}^\dagger = \vec{c}^\dagger \boldsymbol{U}^\dagger \implies \beta_\ell^\dagger = \sum_k^L U_{k\ell}^\dagger c_k^\dagger \,. \tag{D.12}$$

Note that the new operators $\beta^\dagger$ and β obey the same anticommutation relations as the original operators c and $c^\dagger$. For example:

$$\{\beta_i^\dagger, \beta_j^\dagger\} = \left\{ \sum_k U_{ki}^\dagger c_k^\dagger, \sum_\ell U_{\ell j}^\dagger c_\ell^\dagger \right\} \tag{D.13}$$

$$= \sum_{k\ell} U_{ki}^\dagger U_{\ell j}^\dagger \{c_k^\dagger, c_\ell^\dagger\} \tag{D.14}$$

$$\{\beta_i^\dagger, \beta_j^\dagger\} = 0 \qquad \text{since } \{c_k^\dagger, c_\ell^\dagger\} = 0 \,. \tag{D.15}$$

On the other hand, $\vec{\beta}^\dagger = \vec{c}^\dagger \boldsymbol{U}^\dagger$ implies

$$\vec{c}^\dagger = \vec{\beta}^\dagger \boldsymbol{U} \implies c_j^\dagger = \sum_k^L U_{kj} \beta_k^\dagger \tag{D.16}$$

Thus, Eq. (D.11) allows us to rewrite Eq. (D.9) as

$$\widehat{H} = \sum_{m=1}^{L} D_m \beta_m^\dagger \beta_m \,. \tag{D.17}$$

[2] If $\widehat{H}$ is skew-Hermitian, this proof is still correct. However, the matrix D will have purely imaginary diagonal elements instead of real diagonal elements.

The action of the exponential of $\hat{H}$ on the Slater determinant $|\Psi\rangle$ is:

$$\mathrm{e}^{\widehat{H}}\,|\Psi\rangle = \exp\left(\sum_{i,j=1}^{L} T_{ij}c_i^\dagger c_j\right)|\Psi\rangle \tag{D.18}$$

$$= \exp\left(\sum_{m=1}^{L} D_m\beta_m^\dagger\beta_m\right)\prod_{i=1}^{N}\left(\sum_{j=1}^{L}\phi_{ji}c_j^\dagger\right)|0\rangle \tag{D.19}$$

$$= \exp\left(\sum_{m=1}^{L} D_m\beta_m^\dagger\beta_m\right)\prod_{i=1}^{N}\left(\sum_{j,k=1}^{L}\phi_{ji}U_{kj}\beta_k^\dagger\right)|0\rangle \quad \text{using Eq. (D.16)}\,. \tag{D.20}$$

Define $Q \equiv U\Phi$ so that $Q_{ki} = (U\Phi)_{ki} = \sum_j U_{kj}\phi_{ji}$. Doing the j-sum in Eq. (D.20) gives

$$\mathrm{e}^{\widehat{H}}\,|\Psi\rangle = \exp\left(\sum_{m=1}^{L} D_m\beta_m^\dagger\beta_m\right)\prod_{i=1}^{N}\left(\sum_{k=1}^{L}Q_{ki}\beta_k^\dagger\right)|0\rangle\,. \tag{D.21}$$

$$\tag{D.22}$$

Just an aside about the notation we will use later. Define:

$$\prod_{i=1}^{N}\underbrace{\left(\sum_{k=1}^{L}Q_{ki}\beta_k^\dagger\right)}_{\widehat{A_i}} = \underbrace{\left(\sum_{k=1}^{L}Q_{k1}\beta_k^\dagger\right)}_{\widehat{A_1}}\underbrace{\left(\sum_{k=1}^{L}Q_{k2}\beta_k^\dagger\right)}_{\widehat{A_2}}\cdots\underbrace{\left(\sum_{k=1}^{L}Q_{kN}\beta_k^\dagger\right)}_{\widehat{A_N}}\,. \tag{D.23}$$

However, the sums should not have the same index of summation. Therefore let the index of the first one be k_1, that of the second one be k_2 and so on up to k_N because there are N terms in the product in Eq. (D.23). Each index k_i runs from 1 to L. Thus

$$\prod_{i=1}^{N}\left(\sum_{k=1}^{L}Q_{ki}\beta_{k_1}^\dagger\right) = \left(\sum_{k_1=1}^{L}Q_{k_11}\beta_{k_2}^\dagger\right)\left(\sum_{k_2=1}^{L}Q_{k_22}\beta_{k_N}^\dagger\right)\cdots\left(\sum_{k_N=1}^{L}Q_{k_NN}\beta_k^\dagger\right) \tag{D.24}$$

$$= \sum_{k_1,k_2,\ldots,k_N=1}^{L} Q_{k_11}Q_{k_22}\ldots Q_{k_NN}\,\beta_{k_1}^\dagger\beta_{k_2}^\dagger\ldots\beta_{k_N}^\dagger \qquad \text{perform the multiplication}\,. \tag{D.25}$$

Put Eq. (D.25) into Eq. (D.21):

$$\mathrm{e}^{\widehat{H}}\,|\Psi\rangle = \exp\left(\sum_{m=1}^{L} D_m \beta_m^\dagger \beta_m\right)\left(\sum_{k_1,k_2,\ldots,k_N=1}^{L} Q_{k_1 1} Q_{k_2 2}\ldots Q_{k_N N}\,\beta_{k_1}^\dagger \beta_{k_2}^\dagger \ldots \beta_{k_N}^\dagger\right)|0\rangle \tag{D.26}$$

$$= \sum_{k_1,k_2,\ldots,k_N=1}^{L}\left[\exp\left(\sum_{m=1}^{L} D_m \beta_m^\dagger \beta_m\right) Q_{k_1 1} Q_{k_2 2}\ldots Q_{k_N N}\,\beta_{k_1}^\dagger \beta_{k_2}^\dagger \ldots \beta_{k_N}^\dagger\right]|0\rangle \tag{D.27}$$

$$= \sum_{k_1,k_2,\ldots,k_N=1}^{L}\Bigg[\underbrace{\exp\left(D_1\beta_1^\dagger\beta_1\right)\exp\left(D_2\beta_2^\dagger\beta_2\right)\ldots\exp\left(D_L\beta_L^\dagger\beta_L\right)}_{\text{there are } L \text{ of these}} Q_{k_1 1} Q_{k_2 2}\ldots Q_{k_N N}\,\underbrace{\beta_{k_1}^\dagger \beta_{k_2}^\dagger \ldots \beta_{k_N}^\dagger}_{\text{there are } N \text{ of these}}\Bigg]|0\rangle\,. \tag{D.28}$$

Note that in the above equation, because of the commutation relations in Eq. (D.15), there can be no duplicates among the $k_1, k_2, \ldots, k_N$ because if so then there will be multiple applications of the creation operator $\beta_{k_i}^\dagger$ to the k_i-th lattice site which gives 0.

Further simplification is possible if we make use of condition in Eq. (D.8). This condition implies that in each term of the above sum there has to be a $\exp\left(D_j\beta_j^\dagger\beta_j\right)$ that does not have a corresponding $\beta_j^\dagger$. Such a $\exp\left(D_j\beta_j^\dagger\beta_j\right)$ simply becomes $e^{D_j(0)}$ which is just the identity. The intuitive reason is that the particle number operator $n_j = \beta_j^\dagger\beta_j$ is acting on a basis state j that has zero particle because there was no creation operator $\beta_j^\dagger$ that acted on $|0\rangle$ to create a particle in the state j. The mathematical way to see this is that we can write:[3]

$$\exp\left(D_i\beta_i^\dagger\beta_i\right) = \exp(D_i)\beta_i^\dagger\beta_i + \beta_i\beta_i^\dagger\,. \tag{D.29}$$

[3] Proof: Let $\widehat{A} = \beta_i^\dagger\beta_i$ and note that $\widehat{A}^2 = \widehat{A}$. Then

$$\begin{aligned}
e^{D_i\widehat{A}} &= 1 + D_i\widehat{A} + \frac{D_i^2\widehat{A}^2}{2} + \frac{D_i^3\widehat{A}^3}{6} + \ldots \\
&= 1 + D_i\widehat{A} + \frac{D_i^2\widehat{A}}{2} + \frac{D_i^3\widehat{A}}{6} + \ldots \\
&= 1 + \widehat{A}\left(D_i + \frac{D_i^2}{2} + \frac{D_i^3}{6} + \ldots\right) \\
&= 1 + \widehat{A}(e^{D_i} - 1) \\
&= e^{D_i}\widehat{A} + (1 - \widehat{A}) \\
&= e^{D_i}\beta_i^\dagger\beta_i + (1 - \beta_i^\dagger\beta_i) \\
&= e^{D_i}\beta_i^\dagger\beta_i + \beta_i\beta_i^\dagger \qquad\qquad \text{since } \{\beta_i, \beta_i^\dagger\} = 1
\end{aligned}$$

Letting this act on $|0\rangle$ gives

$$\exp\left(D_i\beta_i^\dagger\beta_i\right)|0\rangle = \exp(D_i)\beta_i^\dagger\beta_i|0\rangle + \beta_i\beta_i^\dagger|0\rangle \tag{D.30}$$

$$= 0 + 1|0\rangle = 1|0\rangle\ . \tag{D.31}$$

On the other hand, in each term in the sum there are also $\exp\left(D_j\beta_j^\dagger\beta_j\right)$ that has a corresponding $\beta_j^\dagger$.

$$\exp\left(D_i\beta_i^\dagger\beta_i\right)\beta_i^\dagger = \exp(D_i)\beta_i^\dagger\beta_i\beta_i^\dagger + \beta_i\beta_i^\dagger\beta_i^\dagger \tag{D.32}$$

$$= \exp(D_i)\beta_i^\dagger(1) + 0 \tag{D.33}$$

$$= \exp(D_i)\beta_i^\dagger\,. \tag{D.34}$$

The end result is that after applying all the $\exp\left(D_j\beta_j^\dagger\beta_j\right)$ operators to the $\beta_j^\dagger$ operators, there will be N factors of the form $\exp(D_i)\beta_i^\dagger$ and $(L-N)$ factors of 1 which can be ignored

$$\exp\left(D_1\beta_1^\dagger\beta_1\right)\exp\left(D_2\beta_2^\dagger\beta_2\right)\ldots\exp\left(D_L\beta_L^\dagger\beta_L\right)Q_{k_11}Q_{k_22}\ldots Q_{k_NN}\,\beta_{k_1}^\dagger\beta_{k_2}^\dagger\ldots\beta_{k_N}^\dagger \tag{D.35}$$

$$= \underbrace{\exp\left(D_{k_1}\right)\exp\left(D_{k_2}\right)\ldots\exp\left(D_{k_N}\right)}_{N\text{ of these}}\ \underbrace{1(1)\ldots(1)}_{L-N\text{ of these}}\ Q_{k_11}Q_{k_22}\ldots Q_{k_NN}\,\beta_{k_1}^\dagger\beta_{k_2}^\dagger\ldots\beta_{k_N}^\dagger \tag{D.36}$$

$$= \exp\left(D_{k_1}\right)\exp\left(D_{k_2}\right)\ldots\exp\left(D_{k_N}\right)Q_{k_11}Q_{k_22}\ldots Q_{k_NN}\,\beta_{k_1}^\dagger\beta_{k_2}^\dagger\ldots\beta_{k_N}^\dagger\,. \tag{D.37}$$

Note that in obtaining this result we do need the fact that there is no duplication among the k_i. Putting Eq. (D.37) into Eq. (D.28), we have

$$\mathrm{e}^{\widehat{H}}|\Psi\rangle = \sum_{k_1,k_2,\ldots,k_N=1}^{L}\left[\exp\left(D_{k_1}\right)\exp\left(D_{k_2}\right)\ldots\exp\left(D_{k_N}\right)Q_{k_11}Q_{k_22}\ldots Q_{k_NN}\,\beta_{k_1}^\dagger\beta_{k_2}^\dagger\ldots\beta_{k_N}^\dagger\right]|0\rangle\ . \tag{D.38}$$

Letting $k_i = \ell$ allows for a more compact notation:

$$\mathrm{e}^{\widehat{H}} |\Psi\rangle = \sum_{\ell=1}^{N} \left(\prod_{m=1}^{L} \mathrm{e}^{D_\ell} Q_{\ell m} \beta_\ell^\dagger \right) |0\rangle \tag{D.39}$$

$$= \prod_{m=1}^{L} \left(\sum_{\ell=1}^{N} \mathrm{e}^{D_\ell} Q_{\ell m} \beta_\ell^\dagger \right) |0\rangle \qquad \text{switch sum and product} \tag{D.40}$$

$$= \prod_{m=1}^{L} \left(\sum_{k,\ell=1}^{N} \mathrm{e}^{D_\ell} Q_{\ell m} U_{k\ell}^\dagger c_k^\dagger \right) |0\rangle \qquad \text{using Eq. (D.12)} \tag{D.41}$$

$$= \prod_{m=1}^{L} \left[\sum_{k=1}^{L} \left(\sum_{\ell=1}^{L} U_{k\ell}^\dagger \mathrm{e}^{D_\ell} Q_{\ell m} \right) c_k^\dagger \right] |0\rangle \tag{D.42}$$

$$= \prod_{m=1}^{L} \left[\sum_{k=1}^{L} \left(U^\dagger \mathrm{e}^{D} Q \right)_{km} c_k^\dagger \right] |0\rangle \tag{D.43}$$

$$= \prod_{m=1}^{L} \left[\sum_{k=1}^{L} \left(U^\dagger \mathrm{e}^{D} U \Phi \right)_{km} c_k^\dagger \right] |0\rangle \qquad \text{note } Q \equiv U\Phi \tag{D.44}$$

$$= \prod_{m=1}^{L} \left[\sum_{k=1}^{L} \left(\mathrm{e}^{U^\dagger D U} \Phi \right)_{km} c_k^\dagger \right] |0\rangle \tag{D.45}$$

$$= \prod_{m=1}^{L} \left[\sum_{k=1}^{L} (\mathrm{e}^{T} \Phi)_{km} c_k^\dagger \right] |0\rangle \ . \tag{D.46}$$

By comparing Eq. (D.46) to the definition of a Slater determinant in Eq. (D.7), we conclude that applying $\mathrm{e}^{\widehat{H}}$ to a Slater determinant with coefficient matrix Φ gives another Slater determinant with coefficient matrix $\mathrm{e}^{T}\Phi$. Thus, we have proved Thouless theorem.

Interestingly, Rombouts *et al.* [43, p. 272] pointed out that the converse of Thouless theorem is also true: if the effect of applying an operator with non-singular coefficient matrix Q on a Slater determinant is to left-multiply (the coefficient matrix of) that Slater determinant by Q, that operator has to be the exponential of a one-body operator.

Appendix E

Correctness of the importance-sampled propagator

In this appendix, we will show that Eq. (3.56)

$$|\widetilde{\phi}_{\mathrm{j}}^{(\tau+\Delta\tau)}\rangle = \mathrm{e}^{\Delta\tau[E_{\mathrm{T}}-U(N_\uparrow+N_\downarrow)/2]} \sum_{\vec{x}} \widetilde{P}(\vec{x}) \widehat{B}_{\mathrm{K}/2} \widehat{B}_{\mathrm{V}}(\vec{x}) \widehat{B}_{\mathrm{K}/2} \, |\widetilde{\phi}_{\mathrm{j}}^{(\tau)}\rangle \;, \tag{E.1}$$

which describes an importance-sampled random walk, is *mathematically* equivalent to Eq. (3.48)

$$|\phi_{\mathrm{j}}^{(\tau+\Delta\tau)}\rangle = \mathrm{e}^{\Delta\tau[E_{\mathrm{T}}-U(N_\uparrow+N_\downarrow)/2]} \sum_{\vec{x}} P(\vec{x}) \left[\widehat{B}_{\mathrm{K}/2} \widehat{B}_{\mathrm{V}}(\vec{x}) \widehat{B}_{\mathrm{K}/2}\right] |\phi_{\mathrm{j}}^{(\tau)}\rangle \;, \tag{E.2}$$

which describes a regular random walk.

We first motivate and derive the modified propagator $\widetilde{\mathcal{P}}_{\mathrm{gs}}$. It must preserve the representation of $|\Psi_0\rangle$ in Eq. (3.53) which dictates that the walkers propagate in the following manner

$$w_{\mathrm{j}}^{(\tau+\Delta\tau)} \frac{|\phi_{\mathrm{j}}^{(\tau+\Delta\tau)}\rangle}{\langle\phi_{\mathrm{T}}|\phi_{\mathrm{j}}^{(\tau+\Delta\tau)}\rangle} \leftarrow w_{\mathrm{j}}^{(\tau)} \frac{|\phi_{\mathrm{j}}^{(\tau)}\rangle}{\langle\phi_{\mathrm{T}}|\phi_{\mathrm{j}}^{(\tau)}\rangle} \,. \tag{E.3}$$

By bringing the term $\langle\phi_{\mathrm{T}}|\phi_{\mathrm{j}}^{(\tau+\Delta\tau)}\rangle$ in Eq. (E.3) to the RHS, we obtain an overlap ratio

$$\frac{\langle\phi_{\mathrm{T}}|\phi_{\mathrm{j}}^{(\tau+\Delta\tau)}\rangle}{\langle\phi_{\mathrm{T}}|\phi_{\mathrm{j}}^{(\tau)}\rangle} \tag{E.4}$$

that must multiply the weight at every propagation step in order to preserve the representation in Eq. (3.55). As an example of Eq. (E.4), propagation by the kinetic propagator looks like

$$w_{\mathrm{j}}^{(\tau+\Delta\tau)} \, |\phi_{\mathrm{j}}^{(\tau+\Delta\tau)}\rangle = \left[\frac{\langle\phi_{\mathrm{T}}|\widehat{B}_{\mathrm{K}/2}|\phi_{\mathrm{j}}^{(\tau)}\rangle}{\langle\phi_{\mathrm{T}}|\phi_{\mathrm{j}}^{(\tau)}\rangle} w_{\mathrm{j}}^{(\tau)}\right] \widehat{B}_{\mathrm{K}/2} \, |\phi_{\mathrm{j}}^{(\tau)}\rangle \;, \tag{E.5}$$

which immediately gives the evolution of the weight under the kinetic propagator:

$$w_{\mathrm{j}}^{(\tau+\Delta\tau)} \leftarrow \frac{\langle\phi_{\mathrm{T}}|\widehat{B}_{\mathrm{K}/2}|\phi_{\mathrm{j}}^{(\tau)}\rangle}{\langle\phi_{\mathrm{T}}|\phi_{\mathrm{j}}^{(\tau)}\rangle} w_{\mathrm{j}}^{(\tau)} \,. \tag{E.6}$$

We will now focus on the two-body propagator which is evaluated stochastically and is therefore affected in a non-trivial way by importance sampling. The application of the Hirsch transformation in Eq. (3.32) to a walker $|\phi_{\mathrm{j}}^{(\tau)}\rangle$ gives

$$w_{\mathrm{j},i}^{(\tau)}\,|\phi_{\mathrm{j},i}^{(\tau)}\rangle = \mathrm{e}^{-\Delta\tau U n_{i\uparrow}n_{i\downarrow}}\left[w_{\mathrm{j},i-1}^{(\tau)}\,|\phi_{\mathrm{j},i-1}^{(\tau)}\rangle\right]\,. \tag{E.7}$$

The notation $|\phi_{\mathrm{j},i-1}^{(\tau)}\rangle$ is explained in Eqs. (3.61) to (3.64). Because we must multiply by the overlap ratio in Eq. (E.4), in every iteration over the lattice sites i, Eq. (E.7) becomes

$$w_{\mathrm{j},i}^{(\tau)}\,|\phi_{\mathrm{j},i}^{(\tau)}\rangle = \mathrm{e}^{-\Delta\tau U(n_{i\uparrow}+n_{i\downarrow})/2}\sum_{x_i=\pm1}p(x_i)\,\mathrm{e}^{\gamma x_i(n_{i\uparrow}-n_{i\downarrow})}\left[\frac{\langle\phi_{\mathrm{T}}|\phi_{\mathrm{j},i}^{(\tau)}\rangle}{\langle\phi_{\mathrm{T}}|\phi_{\mathrm{j},i-1}^{(\tau)}\rangle}\right]\left[w_{\mathrm{j},i-1}^{(\tau)}\,|\phi_{\mathrm{j},i-1}^{(\tau)}\rangle\right]\,. \tag{E.8}$$

Rearranging gives

$$w_{\mathrm{j},i}^{(\tau)}\,|\phi_{\mathrm{j},i}^{(\tau)}\rangle = \mathrm{e}^{-\Delta\tau U(n_{i\uparrow}+n_{i\downarrow})/2}w_{\mathrm{j},i-1}^{(\tau)}\sum_{x_i=\pm1}\left[p(x_i)\frac{\langle\phi_{\mathrm{T}}|\phi_{\mathrm{j},i}^{(\tau)}\rangle}{\langle\phi_{\mathrm{T}}|\phi_{\mathrm{j},i-1}^{(\tau)}\rangle}\right]\mathrm{e}^{\gamma x_i(n_{i\uparrow}-n_{i\downarrow})}\,|\phi_{\mathrm{j},i-1}^{(\tau)}\rangle\,. \tag{E.9}$$

Because $|\phi_{\mathrm{j},i}^{(\tau)}\rangle$ is the walker $|\phi_{\mathrm{j},i-1}^{(\tau)}\rangle$ after propagation by $\widehat{b}_{\mathrm{V}}(x_i)$, we can rewrite it as

$$w_{\mathrm{j},i}^{(\tau)}\,|\phi_{\mathrm{j},i}^{(\tau)}\rangle = \mathrm{e}^{-\Delta\tau U(n_{i\uparrow}+n_{i\downarrow})/2}w_{\mathrm{j},i-1}^{(\tau)}\sum_{x_i=\pm1}\left[p(x_i)\frac{\langle\phi_{\mathrm{T}}|\widehat{b}_{\mathrm{V}}(x_i)|\phi_{\mathrm{j},i-1}^{(\tau)}\rangle}{\langle\phi_{\mathrm{T}}|\phi_{\mathrm{j},i-1}^{(\tau)}\rangle}\right]\mathrm{e}^{\gamma x_i(n_{i\uparrow}-n_{i\downarrow})}\,|\phi_{\mathrm{j},i-1}^{(\tau)}\rangle\,. \tag{E.10}$$

We can interpret the quantity in the square bracket as a discrete PDF:

$$\widetilde{p}(x_i) = p(x_i)\frac{\langle\phi_{\mathrm{T}}|\widehat{b}_{\mathrm{V}}(x_i)|\phi_{\mathrm{j},i-1}^{(\tau)}\rangle}{\langle\phi_{\mathrm{T}}|\phi_{\mathrm{j},i-1}^{(\tau)}\rangle}\,. \tag{E.11}$$

Note that $\widetilde{p}(x_i)$ modifies the original $p(x_i)$ in such a way that the choice of x_i that leads to a larger overlap with $|\phi_{\mathrm{T}}\rangle$ has higher probability. However, $\widetilde{p}(x_i)$ is not normalized. We can fix this by introducing a normalization factor

$$\mathfrak{N}(\phi_{\mathrm{j},i}^{(\tau)}) = \frac{1}{2}\left[\frac{\left\langle\phi_{\mathrm{T}}\left|\,\widehat{b}_{\mathrm{V}}(x_i=+1)\,\right|\phi_{\mathrm{j},i-1}^{(\tau)}\right\rangle}{\langle\phi_{\mathrm{T}}|\phi_{\mathrm{j},i-1}^{(\tau)}\rangle} + \frac{\left\langle\phi_{\mathrm{T}}\left|\,\widehat{b}_{\mathrm{V}}(x_i=-1)\,\right|\phi_{\mathrm{j},i-1}^{(\tau)}\right\rangle}{\langle\phi_{\mathrm{T}}|\phi_{\mathrm{j},i-1}^{(\tau)}\rangle}\right] \tag{E.12}$$

where $\widehat{b}_{\mathrm{V}}(x_i=+1)$ is the propagator where the auxiliary field has been chosen to be $+1$. Multiplying and dividing by $\mathfrak{N}(\phi_{\mathrm{j},i}^{(\tau)})$ gives:

$$w_{\mathrm{j},i}^{(\tau)}\,|\phi_{\mathrm{j},i}^{(\tau)}\rangle = \mathrm{e}^{-\Delta\tau U(n_{i\uparrow}+n_{i\downarrow})/2}w_{\mathrm{j},i-1}^{(\tau)}\mathfrak{N}(\phi_{\mathrm{j},i}^{(\tau)})\sum_{x_i=\pm1}\left[\frac{\widetilde{p}(x_i)}{\mathfrak{N}(\phi_{\mathrm{j},i}^{(\tau)})}\right]\mathrm{e}^{\gamma x_i(n_{i\uparrow}-n_{i\downarrow})}\,|\phi_{\mathrm{j},i-1}^{(\tau)}\rangle\,. \tag{E.13}$$

It's easy to verify that $\frac{\widetilde{p}(x_i)}{\mathfrak{N}(\phi_{\mathrm{j},i}^{(\tau)})}$ is indeed a normalized (discrete) PDF which we can sample from. We can tidy this up a little:

$$w_{\mathrm{j},i}^{(\tau)} \left|\phi_{\mathrm{j},i}^{(\tau)}\right\rangle = W_{\mathrm{V},\mathrm{j},i} \sum_{x_i=\pm 1} \left[\frac{\widetilde{p}(x_i)}{\mathfrak{N}(\phi_{\mathrm{j},i}^{(\tau)})}\right] \mathrm{e}^{\gamma x_i (n_{i\uparrow}-n_{i\downarrow})} \left|\phi_{\mathrm{j},i-1}^{(\tau)}\right\rangle \tag{E.14}$$

where

$$W_{\mathrm{V},\mathrm{j},i} = \mathrm{e}^{-\Delta\tau U(n_{i\uparrow}+n_{i\downarrow})/2} w_{\mathrm{j},i-1}^{(\tau)} \mathfrak{N}(\phi_{\mathrm{j},i}^{(\tau)}) \tag{E.15}$$

is the aggregate of all the scalar prefactors in Eq. (E.13). Eq. (E.15) very clearly shows the evolution of the weight after propagation by the potential propagator on site i

$$w_{\mathrm{j},i}^{(\tau)} \leftarrow \mathrm{e}^{-\Delta\tau U(n_{i\uparrow}+n_{i\downarrow})/2} w_{\mathrm{j},i-1}^{(\tau)} \mathfrak{N}(\phi_{\mathrm{j},i}^{(\tau)})\,. \tag{E.16}$$

Eq. (E.14) gives the importance-sampled potential propagator over a single lattice site i.

The potential propagator over all sites, as expected, is just the product of the propagators over all M sites:

$$\begin{aligned} w_{\mathrm{j},M}^{(\tau)} \left|\phi_{\mathrm{j},M}^{(\tau)}\right\rangle &= \mathrm{e}^{-\Delta\tau \widehat{V}} w_{\mathrm{j},0}^{(\tau)} \left|\phi_{\mathrm{j},0}^{(\tau)}\right\rangle && \text{(E.17)} \\ &= \prod_i \mathrm{e}^{-\Delta\tau U n_{i\uparrow} n_{i\downarrow}} w_{\mathrm{j},0}^{(\tau)} \left[\left|\phi_{\mathrm{j},0}^{(\tau)}\right\rangle\right] && \text{(E.18)} \\ &= W_{\mathrm{V},\mathrm{j}} \sum_{\vec{x}} \widetilde{P}(\vec{x}) \widehat{B}_{\mathrm{V}}(\vec{x}) \left|\phi_{\mathrm{j},0}^{(\tau)}\right\rangle\,, && \text{(E.19)} \end{aligned}$$

where $\widetilde{P}(\vec{x})$ is the normalized modified PDF over all sites

$$\widetilde{P}(\vec{x}) = \prod_i \left[\frac{\widetilde{p}(x_i)}{\mathfrak{N}_i(\phi_{\mathrm{j}}^{(\tau)})}\right] \tag{E.20}$$

and $\widehat{B}_{\mathrm{V}}(\vec{x})$ was defined in Eq. (3.49). The overall weight $W_{\mathrm{V},\mathrm{j}}$ is simply the product of the weights in Eq. (E.15) over all sites

$$\begin{aligned} W_{\mathrm{V},\mathrm{j}} &= \prod_i W_{\mathrm{V},i,\mathrm{j}} && \text{(E.21)} \\ &= \prod_i \mathrm{e}^{-\Delta\tau U(n_{i\uparrow}+n_{i\downarrow})/2} w_{\mathrm{j},i-1}^{(\tau)} \mathfrak{N}(\phi_{\mathrm{j},i}^{(\tau)}) && \text{(E.22)} \\ &= \mathrm{e}^{-\Delta\tau U(N_\uparrow+N_\downarrow)/2} \prod_i \mathfrak{N}(\phi_{\mathrm{j},i}^{(\tau)}) && \text{(E.23)} \\ &= \mathrm{e}^{-\Delta\tau U(N_\uparrow+N_\downarrow)/2} \mathfrak{N}_{\mathrm{j}}^{(\tau)} && \text{(E.24)} \end{aligned}$$

where, for convenience later, we have defined

$$\mathfrak{N}_{\mathrm{j}}^{(\tau)} = \prod_i \mathfrak{N}(\phi_{\mathrm{j},i}^{(\tau)})\,. \tag{E.25}$$

This gives us the evolution of the weight after propagation by the potential propagator on all M sites:

$$w_{\mathrm{j},M}^{(\tau)} \leftarrow \mathrm{e}^{-\Delta\tau U(N_\uparrow+N_\downarrow)/2} \mathfrak{N}_{\mathrm{j}}^{(\tau)} w_{\mathrm{j},0}^{(\tau)}\,. \tag{E.26}$$

Now we will prove the claim stated at the beginning of this appendix. Referring back to Eq. (E.8), we see that the overlap ratios are implicit in the normalized PDF $\tilde{P}$ and we can see that due to a "telescoping" product over the sites i, $W_{\mathrm{V,j}}$ is:

$$W_{\mathrm{V,j}} = \mathrm{e}^{-\Delta\tau U(N_\uparrow+N_\downarrow)/2}\frac{\langle\phi_{\mathrm{T}}|\phi^{(\tau)}_{\mathrm{j},M}\rangle}{\langle\phi_{\mathrm{T}}|\phi^{(\tau)}_{\mathrm{j},0}\rangle}\,. \tag{E.27}$$

Now we can write the importance-sampled ground-state projection operator:

$$\begin{aligned} w_{\mathrm{j}}^{(\tau+\Delta\tau)}\,|\phi_{\mathrm{j}}^{(\tau+\Delta\tau)}\rangle &= \mathcal{P}_{\mathrm{gs}}\,w_{\mathrm{j}}^{(\tau)}\,|\phi_{\mathrm{j}}^{(\tau)}\rangle && \text{(E.28)}\\ &= \mathrm{e}^{\Delta\tau E_{\mathrm{T}}}\mathrm{e}^{-\Delta\tau\widehat{K}/2}\mathrm{e}^{-\Delta\tau\widehat{V}}\mathrm{e}^{-\Delta\tau\widehat{K}/2} && \text{(E.29)}\\ &= \sum_{\vec{x}}\widetilde{P}(\vec{x})\times\left[\mathrm{e}^{\Delta\tau E_{\mathrm{T}}}w_{\mathrm{j}}^{(\tau)}\frac{\langle\phi_{\mathrm{T}}|\widehat{B}_{\mathrm{K}/2}|\phi^{(\tau)}_{\mathrm{j},M}\rangle}{\langle\phi_{\mathrm{T}}|\phi^{(\tau)}_{\mathrm{j},M}\rangle}W_{\mathrm{V,j}}\frac{\langle\phi_{\mathrm{T}}|\widehat{B}_{\mathrm{K}/2}|\phi^{(\tau)}_{\mathrm{j}}\rangle}{\langle\phi_{\mathrm{T}}|\phi^{(\tau)}_{\mathrm{j}}\rangle}\right]\times\\ &\qquad\times\left[\widehat{B}_{\mathrm{K}/2}\widehat{B}_{\mathrm{V}}(\vec{x})\widehat{B}_{\mathrm{K}/2}\,|\phi_{\mathrm{j}}^{(\tau)}\rangle\right]. && \text{(E.30)}\end{aligned}$$

Using Eqs. (3.63), (3.64) and (E.20) gives

$$\begin{aligned} w_{\mathrm{j}}^{(\tau+\Delta\tau)}\,|\phi_{\mathrm{j}}^{(\tau+\Delta\tau)}\rangle &= \sum_{\vec{x}}\widetilde{P}(\vec{x})\times\left[\mathrm{e}^{\Delta\tau[E_{\mathrm{T}}-U(N_\uparrow+N_\downarrow)/2]}w_{\mathrm{j}}^{(\tau)}\frac{\langle\phi_{\mathrm{T}}|\phi_{\mathrm{j}}^{(\tau+\Delta\tau)}\rangle}{\cancel{\langle\phi_{\mathrm{T}}|\phi^{(\tau)}_{\mathrm{j},M}\rangle}}\frac{\cancel{\langle\phi_{\mathrm{T}}|\phi^{(\tau)}_{\mathrm{j},M}\rangle}}{\cancel{\langle\phi_{\mathrm{T}}|\phi^{(\tau)}_{\mathrm{j},0}\rangle}}\frac{\cancel{\langle\phi_{\mathrm{T}}|\phi^{(\tau)}_{\mathrm{j},0}\rangle}}{\langle\phi_{\mathrm{T}}|\phi^{(\tau)}_{\mathrm{j}}\rangle}\right]\times\\ &\qquad\times\left[\widehat{B}_{\mathrm{K}/2}\widehat{B}_{\mathrm{V}}(\vec{x})\widehat{B}_{\mathrm{K}/2}\,|\phi_{\mathrm{j}}^{(\tau)}\rangle\right] && \text{(E.31)}\\ w_{\mathrm{j}}^{(\tau+\Delta\tau)}\,|\phi_{\mathrm{j}}^{(\tau+\Delta\tau)}\rangle &= \sum_{\vec{x}}\widetilde{P}(\vec{x})\times\left[\mathrm{e}^{\Delta\tau[E_{\mathrm{T}}-U(N_\uparrow+N_\downarrow)/2]}\frac{\langle\phi_{\mathrm{T}}|\phi_{\mathrm{j}}^{(\tau+\Delta\tau)}\rangle}{\langle\phi_{\mathrm{T}}|\phi_{\mathrm{j}}^{(\tau)}\rangle}w_{\mathrm{j}}^{(\tau)}\right]\times\\ &\qquad\times\left[\widehat{B}_{\mathrm{K}/2}\widehat{B}_{\mathrm{V}}(\vec{x})\widehat{B}_{\mathrm{K}/2}\,|\phi_{\mathrm{j}}^{(\tau)}\rangle\right]. && \text{(E.32)}\end{aligned}$$

Moving the term $\langle\phi_{\mathrm{T}}|\phi_{\mathrm{j}}^{(\tau+\Delta\tau)}\rangle$ to the LHS:

$$w_{\mathrm{j}}^{(\tau+\Delta\tau)}\frac{|\phi_{\mathrm{j}}^{(\tau+\Delta\tau)}\rangle}{\langle\phi_{\mathrm{T}}|\phi_{\mathrm{j}}^{(\tau+\Delta\tau)}\rangle}\,.=\mathrm{e}^{\Delta\tau[E_{\mathrm{T}}-U(N_\uparrow+N_\downarrow)/2]}\sum_{\vec{x}}\widetilde{P}(\vec{x})\widehat{B}_{\mathrm{K}/2}\widehat{B}_{\mathrm{V}}(\vec{x})\widehat{B}_{\mathrm{K}/2}\frac{|\phi_{\mathrm{j}}^{(\tau)}\rangle}{\langle\phi_{\mathrm{T}}|\phi_{\mathrm{j}}^{(\tau)}\rangle}\,. \tag{E.33}$$

Thus, the importance-sampled walkers and weights must evolve in the following manner:

$$|\widetilde{\phi}_{\mathrm{j}}^{(\tau+\Delta\tau)}\rangle = \mathrm{e}^{\Delta\tau[E_{\mathrm{T}}-U(N_\uparrow+N_\downarrow)/2]}\sum_{\vec{x}}\widetilde{P}(\vec{x})\widehat{B}_{\mathrm{K}/2}\widehat{B}_{\mathrm{V}}(\vec{x})\widehat{B}_{\mathrm{K}/2}\,|\widetilde{\phi}_{\mathrm{j}}^{(\tau)}\rangle \tag{E.34}$$

$$w_{\mathrm{j}}^{(\tau+\Delta\tau)}\leftarrow\mathrm{e}^{\Delta\tau[E_{\mathrm{T}}-U(N_\uparrow+N_\downarrow)/2]}\frac{\langle\phi_{\mathrm{T}}|\phi_{\mathrm{j}}^{(\tau+\Delta\tau)}\rangle}{\langle\phi_{\mathrm{T}}|\phi_{\mathrm{j}}^{(\tau)}\rangle}w_{\mathrm{j}}^{(\tau)}\,, \tag{E.35}$$

which shows that although the importance-sampled iteration is *formally* different, it is indeed *mathematically* equivalent to the original iteration in Eq. (3.48).

In our CPMC implementation, we simply multiply all the weights by the prefactor $\mathrm{e}^{\Delta\tau[E_{\mathrm{T}}-U(N_\uparrow+N_\downarrow)/2]}$ at the start of every iteration because it is essentially independent of what goes on inside the iteration. The overlap ratio is updated every time a walker is acted on by a propagator.

Appendix F

Equal-time Green's function

In this appendix we will prove a simple formula for calculating the equal-time Green's function for any two Slater determinants

$$\langle c_i c_j^\dagger \rangle = \frac{\langle P | c_i c_j^\dagger | Q \rangle}{\langle P | Q \rangle} \,. \tag{F.1}$$

By "equal-time," we mean that there are no time evolution operators sandwiched between c_i and $c_j^\dagger$. Along the way, we will also prove Eq. (3.73):

$$\langle c_{j\sigma}^\dagger c_{i\sigma} \rangle \equiv \frac{\langle \phi_{\mathrm{T}} | c_{j\sigma}^\dagger c_{i\sigma} | \phi \rangle}{\langle \phi_{\mathrm{T}} | \phi \rangle} = \left[\Phi^\sigma \left[\, (\Phi_{\mathrm{T}}^\sigma)^\dagger \Phi^\sigma \, \right]^{-1} (\Phi_{\mathrm{T}}^\sigma)^\dagger \right]_{ij} \,. \tag{F.2}$$

First we'll prove that

$$\mathrm{Det}(M) = \mathrm{e}^{\mathrm{Tr}(\ln M)} \,. \tag{F.3}$$

Since we only work with matrices, according to Higham [44, p. 17] we can always find the logarithm of M i.e. a matrix A such that $\mathrm{e}^A = M$. With the new matrix A we can easily prove Eq. (F.3):

$$\begin{aligned}
\mathrm{Det}(\mathrm{e}^A) &= \mathrm{Det}[\exp(P^{-1} D P)] && \text{diagonalize } A && \text{(F.4)} \\
&= \mathrm{Det}\left[P^{-1} \mathrm{e}^D P \right] && && \text{(F.5)} \\
&= \mathrm{Det}(\mathrm{e}^D) && \text{cyclic invariance of trace} && \text{(F.6)} \\
&= \mathrm{e}^{D_{11}} \mathrm{e}^{D_{22}} \ldots \mathrm{e}^{D_{nn}} && \text{since } D \text{ is diagonal} && \text{(F.7)} \\
&= \mathrm{e}^{D_{11} + D_{22} + \cdots + D_{nn}} && && \text{(F.8)} \\
&= \mathrm{e}^{\mathrm{Tr}\, D} && && \text{(F.9)} \\
&= \mathrm{e}^{\mathrm{Tr}\, A} && \text{since trace is basis independent} \,. && \text{(F.10)}
\end{aligned}$$

Eq. (F.3) then implies that

$$\ln(\mathrm{Det}\, M) = \mathrm{Tr}(\ln M) \,. \tag{F.11}$$

Now suppose the two aforementioned Slater determinants are

$$|Q\rangle = \prod_{m=1}^{N} \left(\sum_{\ell=1}^{L} Q_{\ell m} c_\ell^\dagger \right) |0\rangle \tag{F.12}$$

$$|P\rangle = \prod_{m=1}^{N} \left(\sum_{k=1}^{L} P_{km} c_k^\dagger \right) |0\rangle \;, \tag{F.13}$$

Our first step is to calculate the following quantity:

$$\tilde{G}_{ij} = \frac{\langle P|c_j^\dagger c_i|Q\rangle}{\langle P|Q\rangle}\,. \tag{F.14}$$

We can replace $c_j^\dagger c_i$ with the second-quantized operator $\widehat{O}$ whose matrix elements are

$$O_{mn} = \delta_{jm}\delta_{ni} \tag{F.15}$$

because

$$c_j^\dagger c_i = \sum_{mn}^{L} \delta_{jm}\delta_{ni} c_m^\dagger c_n \tag{F.16}$$

$$= \sum_{mn} O_{mn} c_m^\dagger c_n \tag{F.17}$$

$$= \widehat{O}\,. \tag{F.18}$$

Next we note that

$$\frac{\partial}{\partial\eta}\ln\langle P|\mathrm{e}^{\eta\widehat{O}}|Q\rangle = \frac{\frac{\partial}{\partial\eta}\langle P|\mathrm{e}^{\eta\widehat{O}}|Q\rangle}{\langle P|\mathrm{e}^{\eta\widehat{O}}|Q\rangle} \quad \text{chain rule} \tag{F.19}$$

$$= \frac{\langle P|\widehat{O}\,\mathrm{e}^{\eta\widehat{O}}|Q\rangle}{\langle P|\mathrm{e}^{\eta\widehat{O}}|Q\rangle} \tag{F.20}$$

and that letting $\eta = 0$ recovers Eq. (F.14)

$$\frac{\partial}{\partial\eta}\ln\langle P|\mathrm{e}^{\eta\widehat{O}}|Q\rangle\Big|_{\eta=0} = \frac{\langle P|\widehat{O}|Q\rangle}{\langle P|Q\rangle} \tag{F.21}$$

because $\mathrm{e}^{\eta\widehat{O}}\big|_{\eta=0}$ is the identity. Thus we can write

$$\tilde{G}_{ij} = \frac{\langle P|\widehat{O}|Q\rangle}{\langle P|Q\rangle} \tag{F.22}$$

$$= \frac{\partial}{\partial\eta}\ln\langle P|\mathrm{e}^{\eta\widehat{O}}|Q\rangle\Big|_{\eta=0} \tag{F.23}$$

$$= \frac{\partial}{\partial\eta}\ln\mathrm{Det}\left[P^\dagger \mathrm{e}^{\eta O}Q\right]\Big|_{\eta=0}\,. \tag{F.24}$$

To go from Eq. (F.23) to Eq. (F.24), we have use the result of Appendix D (Thouless theorem)[1]and then Appendix C (overlap of two Slater determinants). Note the absence of a hat on O in Eq. (F.24) because it is now a matrix with elements given in Eq. (F.15) and not an operator. Now we can continue by using Eq. (F.11)

$$\tilde{G}_{ij} = \frac{\partial}{\partial\eta} \operatorname{Tr} \ln \left[P^\dagger \mathrm{e}^{\eta O} Q \right] \Big|_{\eta=0} \tag{F.25}$$

$$= \operatorname{Tr} \left[\frac{\partial}{\partial\eta} \ln \left(P^\dagger \mathrm{e}^{\eta O} Q \right) \right] \Big|_{\eta=0} \quad \text{trace and derivative commute} \tag{F.26}$$

$$= \operatorname{Tr} \left[\frac{P^\dagger \, O \, \mathrm{e}^{\eta O} Q}{P^\dagger \mathrm{e}^{\eta O} Q} \right] \Big|_{\eta=0} \quad \text{chain rule} \tag{F.27}$$

$$= \operatorname{Tr} \left(\frac{P^\dagger O Q}{P^\dagger Q} \right) \quad \text{letting } \eta = 0 \tag{F.28}$$

$$= \operatorname{Tr} \left[(P^\dagger Q)^{-1} P^\dagger O Q \right] \quad \text{cyclic invariance of trace} \tag{F.29}$$

$$= \operatorname{Tr} \left[Q (P^\dagger Q)^{-1} P^\dagger O \right] \tag{F.30}$$

where the last equality follows from the invariance of the trace under cyclic permutation. Then by letting $N = Q(P^\dagger Q)^{-1} P^\dagger$, we can rewrite Eq. (F.30) as:

$$\tilde{G}_{ij} = \operatorname{Tr}(NO) = \sum_\ell (NO)_{\ell\ell} \tag{F.31}$$

$$= \sum_{k\ell} N_{\ell k} O_{k\ell} \tag{F.32}$$

$$= \sum_{k\ell} N_{\ell k} \delta_{jk} \delta_{\ell i} \quad \text{definition of the matrix } O \tag{F.33}$$

$$= N_{ij} \, . \tag{F.34}$$

In other words:

$$\tilde{G}_{ij} = \frac{\langle P | c_j^\dagger c_i | Q \rangle}{\langle P | Q \rangle} = \left[Q (P^\dagger Q)^{-1} P^\dagger \right]_{ij} \, . \tag{F.35}$$

Do note the different orderings of i and j on the left- and right-hand sides of the last equality.

[1] An attentive reader will notice that the result in Appendix D were proved with the assumption that the operator is Hermitian or skew-Hermitian which $\widehat{O}$, as defined in Eq. (F.15), clearly is not. However, Assaad [42, p. 114] points out that because only terms of order η are relevant in the present proof, we can replace $\mathrm{e}^{\eta\widehat{O}}$ with $\mathrm{e}^{\eta(\widehat{O}+\widehat{O}^\dagger)/2} \mathrm{e}^{\eta(\widehat{O}-\widehat{O}^\dagger)/2}$ which is exact up to order η^2 because

$$\begin{aligned}
\mathrm{e}^{\eta(\widehat{O}+\widehat{O}^\dagger)/2} \mathrm{e}^{\eta(\widehat{O}-\widehat{O}^\dagger)/2} &= \left[1 + \frac{\eta}{2}(\widehat{O}+\widehat{O}^\dagger) + \frac{1}{2}\frac{\eta^2}{4}(\widehat{O}+\widehat{O}^\dagger)^2 + \dots \right] \left[1 + \frac{\eta}{2}(\widehat{O}-\widehat{O}^\dagger) + \frac{1}{2}\frac{\eta^2}{4}(\widehat{O}-\widehat{O}^\dagger)^2 + \dots \right] \\
&= 1 + \eta \left[\frac{\widehat{O}+\widehat{O}^\dagger}{2} + \frac{\widehat{O}-\widehat{O}^\dagger}{2} \right] + \eta^2 \left[\frac{(\widehat{O}-\widehat{O}^\dagger)^2}{8} + \frac{(\widehat{O}+\widehat{O}^\dagger)(\widehat{O}-\widehat{O}^\dagger)}{4} + \frac{(\widehat{O}+\widehat{O}^\dagger)^2}{8} \right] + \mathcal{O}(\eta^3) \\
&= 1 + \eta\widehat{O} + \frac{1}{2}(\eta\widehat{O})^2 + \mathcal{O}(\eta^3) = \mathrm{e}^{\eta\widehat{O}} + \mathcal{O}(\eta^3) \, .
\end{aligned}$$

This replacement allows us to use the result in Appendix D. The main idea here is that any operator can be written as a sum of Hermitian and skew-Hermitian operators.

Finally, the one-particle equal-time Green's function between the two Slater determinants $|P\rangle$ and $|Q\rangle$ is then

$$G_{ij} = \langle c_i c_j^\dagger \rangle = \frac{\langle P|c_i c_j^\dagger|Q\rangle}{\langle P|Q\rangle} \tag{F.36}$$

$$= \frac{\langle P|(\delta_{ij} - c_j^\dagger c_i)|Q\rangle}{\langle P|Q\rangle} \tag{F.37}$$

$$= \delta_{ij} - \tilde{G}_{ij}\,. \tag{F.38}$$

If we think of G_{ij} and $\tilde{G}_{ij}$, respectively, as the (i,j)-th elements of square matrices G and $\tilde{G}$ like we usually do in this thesis, the expression for G follows easily from Eq. (F.38):

$$G = I - \tilde{G} \tag{F.39}$$

where I is the identity matrix. This is a simple formula that allows us to calculate the equal-time Green's function for any two Slater determinants.

Appendix G

Metropolis AFQMC

This appendix gives an overview of the original formulation of the auxiliary-field quantum Monte Carlo (AFQMC) method by Sugiyama and Koonin [16]. This formulation uses the Metropolis algorithm[1] to generate a Metropolis-like random walk, unlike the open-ended branching random walk formulation presented in the rest of the thesis. However, the Metropolis formulation will come in handy when we discuss the computation of the unequal-time Green's function in Appendix H. This appendix has adapted materials from Purwanto and Zhang [29] and Assaad [42].

We are looking to do projection for an imaginary-time interval of $\Lambda = L\Delta\tau$. Let $\ell = 1, \ldots, L$ be the index for the time slices. The total projection over the whole interval τ is

$$\widehat{\mathcal{P}}_\Lambda = \prod_{\ell=L}^{1} \widehat{\mathcal{P}}^{(\ell)} \,. \tag{G.1}$$

where the index ℓ runs backward to indicate that earlier times appear to the right per convention. The ℓ-th step of this projection is

$$\widehat{\mathcal{P}}^{(\ell)} = \mathcal{F} \sum_{\vec{x}^{(\ell)}} \widehat{B}(\vec{x}^{(\ell)}) \,. \tag{G.2}$$

Note that this is exactly the same as Eq. (3.41). The prefactor in Eq. (G.2) is

$$\mathcal{F} = \left(\frac{1}{2}\right)^M , \tag{G.3}$$

and the operator is

$$\widehat{B}(\vec{x}^{(\ell)}) = \prod_{\sigma=\uparrow,\downarrow} \exp\Big(-\frac{\Delta\tau}{2}\sum_{ij} c^\dagger_{i\sigma} K_{ij} c_{j\sigma}\Big) \exp\Big(\sum_i c^\dagger_{i\sigma} V_i^\sigma(\ell) c_{i\sigma}\Big) \exp\Big(-\frac{\Delta\tau}{2}\sum_{ij} c^\dagger_{i\sigma} K_{ij} c_{j\sigma}\Big) \tag{G.4}$$

where $\vec{x}^{(\ell)} = (x_1, x_2, \ldots, x_M)$ is a configuration of auxiliary fields over all M lattice sites, K_{ij} are matrix elements of the Hubbard kinetic operator $\widehat{K}$ and

$$V_i^\sigma(\ell) = -\frac{U\Delta\tau}{2} + s(\sigma)\gamma x_i^{(\ell)} \tag{G.5}$$

[1] Here we use the term Metropolis algorithm to include both the original algorithm by Metropolis *et al.* [11] and similar algorithms that use different acceptance/rejection probabilities, such as the heat-bath algorithm.

where $s(\uparrow) = 1$ and $s(\downarrow) = -1$ and γ is given by $\cosh(\gamma) = \exp(\Delta\tau U/2)$ like in Eq. (3.41). Eq. (G.4) can be very easily derived from Eq. (3.32). Note the time-slice dependence of both $\widehat{\mathcal{P}}$ and $\widehat{B}$ through the auxiliary-field configuration $\vec{x}^{(\ell)}$.

The total projection operator in Eq. (G.1) is

$$\widehat{\mathcal{P}}_\Lambda = \prod_{\ell=L}^{1} \widehat{\mathcal{P}}^{(\ell)} \tag{G.6}$$

$$= \prod_{\ell=L}^{1} \left[\mathcal{F} \sum_{\vec{x}^{(\ell)}} \widehat{B}(\vec{x}^{(\ell)}) \right] \tag{G.7}$$

$$= \mathcal{F}^L \sum_{\vec{X}} \widehat{B}_{\vec{X}}(\Lambda, 0) \tag{G.8}$$

where we have defined the configuration of auxiliary fields over *all lattice sites and time slices* to be

$$\vec{X} = \left\{ \vec{x}^{(L)}, \vec{x}^{(L-1)}, \ldots, \vec{x}^{(1)} \right\} \tag{G.9}$$

and the $\vec{X}$-dependent propagator to be

$$\widehat{B}_{\vec{X}}(\Lambda, 0) = \widehat{B}(\vec{x}^{(L)}) \widehat{B}(\vec{x}^{(L-1)}) \ldots \widehat{B}(\vec{x}^{(1)}) \,. \tag{G.10}$$

Similarly, the projection operator for an interval $\Theta = T\Delta\tau$ from Λ to $\Lambda + \Theta$ is

$$\widehat{\mathcal{P}}_\Theta = \mathcal{F}^T \sum_{\vec{Y}} \widehat{B}_{\vec{Y}}(\Theta + \Lambda, \Lambda) \tag{G.11}$$

where $\vec{Y}$ is the corresponding auxiliary field configuration.

The ground-state expectation of an observable $\widehat{A}$ is

$$\langle \widehat{A} \rangle = \frac{\langle \Psi_0 | \widehat{A} | \Psi_0 \rangle}{\langle \Psi_0 | \Psi_0 \rangle} \tag{G.12}$$

$$= \frac{\langle \phi_{\mathrm{T}} | \widehat{\mathcal{P}}_\Theta \, \widehat{A} \, \widehat{\mathcal{P}}_\Lambda | \phi_{\mathrm{T}} \rangle}{\langle \phi_{\mathrm{T}} | \widehat{\mathcal{P}}_\Theta \, \widehat{\mathcal{P}}_\Lambda | \phi_{\mathrm{T}} \rangle} \,. \tag{G.13}$$

Here Θ and Λ are imaginary-time intervals chosen to be large enough so as to project out the ground state from the left- and right-hand side wave function, respectively. The denominator of Eq. (G.13) is

$$\langle \phi_{\mathrm{T}} | \widehat{\mathcal{P}}_\Theta \, \widehat{\mathcal{P}}_\Lambda | \phi_{\mathrm{T}} \rangle = \langle \phi_{\mathrm{T}} | \left[\mathcal{F}^T \sum_{\vec{Y}} \widehat{B}_{\vec{Y}}(\Lambda + \Theta, \Lambda) \right] \left[\mathcal{F}^L \sum_{\vec{X}} \widehat{B}_{\vec{X}}(\Lambda, 0) \right] | \phi_{\mathrm{T}} \rangle \tag{G.14}$$

$$= \mathcal{F}^{L+T} \sum_{\{\vec{Y}, \vec{X}\}} \left\langle \phi_{\mathrm{T}} \,\middle|\, \widehat{B}_{\vec{Y}}(\Lambda + \Theta, \Lambda) \widehat{B}_{\vec{X}}(\Lambda, 0) \,\middle|\, \phi_{\mathrm{T}} \right\rangle \tag{G.15}$$

$$= \mathcal{F}^{L+T} \sum_{\vec{S}} \left\langle \phi_{\mathrm{T}} \,\middle|\, \widehat{B}_{\vec{S}}(\Lambda + \Theta, \Lambda) \widehat{B}_{\vec{S}}(\Lambda, 0) \,\middle|\, \phi_{\mathrm{T}} \right\rangle \tag{G.16}$$

where we think of the $(L+T)M$-dimensional vector $\vec{S} = (\vec{Y}, \vec{X})$ as $\vec{X}$ concatenated to $\vec{Y}$ and the propagators

$$\widehat{B}_{\vec{S}}(\Lambda + \Theta, \Lambda) = \widehat{B}_{\vec{Y}}(\Lambda + \Theta, \Lambda) \tag{G.17}$$

$$\widehat{B}_{\vec{S}}(\Lambda, 0) = \widehat{B}_{\vec{X}}(\Lambda, 0) \,. \tag{G.18}$$

If we go further and define

$$\widehat{B}_{\vec{S}}(\Lambda+\Theta,0) = \widehat{B}_{\vec{Y}}(\Lambda+\Theta,\Lambda)\widehat{B}_{\vec{X}}(\Lambda,0) \tag{G.19}$$

then

$$\langle\phi_{\mathrm{T}}|\widehat{\mathcal{P}}_{\Theta}\ \widehat{\mathcal{P}}_{\Lambda}|\phi_{\mathrm{T}}\rangle = \mathcal{F}^{L+T}\sum_{\vec{S}}\left\langle\phi_{\mathrm{T}}\,\middle|\,\widehat{B}_{\vec{S}}(\Lambda+\Theta,0)\,\middle|\,\phi_{\mathrm{T}}\right\rangle \tag{G.20}$$

$$= \mathcal{F}^{L+T}\sum_{\vec{S}}\langle\eta_{\vec{S}}\,|\,\phi_{\vec{S}}\rangle \tag{G.21}$$

where we have defined the "left-hand" wave function to be

$$\langle\eta_{\vec{S}}| = \langle\phi_{\mathrm{T}}|\,\widehat{B}_{\vec{S}}(\Lambda+\Theta,\Lambda) \tag{G.22}$$

and the "right-hand" wave function to be

$$|\phi_{\vec{S}}\rangle = \widehat{B}_{\vec{S}}(\Lambda,0)\,|\phi_{\mathrm{T}}\rangle\ . \tag{G.23}$$

The usefulness of this notation will be clear when we evaluate the numerator:

$$\langle\phi_{\mathrm{T}}|\widehat{\mathcal{P}}_{\Theta}\,\widehat{A}\,\widehat{\mathcal{P}}_{\Lambda}|\phi_{\mathrm{T}}\rangle = \mathcal{F}^{L+T}\sum_{\{\vec{Y},\vec{X}\}}\left\langle\phi_{\mathrm{T}}\,\middle|\,\widehat{B}_{\vec{Y}}(\Lambda+\Theta,\Lambda)\widehat{A}\,\widehat{B}_{\vec{X}}(\Lambda,0)\,\middle|\,\phi_{\mathrm{T}}\right\rangle \tag{G.24}$$

$$= \mathcal{F}^{L+T}\sum_{\vec{S}}\left\langle\phi_{\mathrm{T}}\,\middle|\,\widehat{B}_{\vec{S}}(\Lambda+\Theta,\Lambda)\widehat{A}\,\widehat{B}_{\vec{S}}(\Lambda,0)\,\middle|\,\phi_{\mathrm{T}}\right\rangle \tag{G.25}$$

$$= \mathcal{F}^{L+T}\sum_{\vec{S}}\left\langle\eta_{\vec{S}}\,\middle|\,\widehat{A}\,\middle|\,\phi_{\vec{S}}\right\rangle\ . \tag{G.26}$$

Because the prefactor $\mathcal{F}^{L+T}$ in the denominator and numerator cancel, we have

$$\langle\widehat{A}\rangle = \frac{\displaystyle\sum_{\vec{S}}\left\langle\eta_{\vec{S}}\,\middle|\,\widehat{A}\,\middle|\,\phi_{\vec{S}}\right\rangle}{\displaystyle\sum_{\vec{S}}\langle\eta_{\vec{S}}\,|\,\phi_{\vec{S}}\rangle}\ . \tag{G.27}$$

Multiply and divide the numerator by $\langle\phi_{\mathrm{T}}|\widehat{B}_{\vec{S}}(\Lambda+\Theta,0)|\phi_{\mathrm{T}}\rangle$:

$$\langle\widehat{A}\rangle = \frac{\displaystyle\sum_{\vec{S}}\langle\eta_{\vec{S}}|\phi_{\vec{S}}\rangle\left[\frac{\langle\eta_{\vec{S}}|\widehat{A}|\phi_{\vec{S}}\rangle}{\langle\eta_{\vec{S}}|\phi_{\vec{S}}\rangle}\right]}{\displaystyle\sum_{\vec{S}}\langle\eta_{\vec{S}}|\phi_{\vec{S}}\rangle} \tag{G.28}$$

$$= \sum_{\vec{S}} P(\vec{S})\,\langle\widehat{A}\rangle_{\vec{S}} \tag{G.29}$$

where we have defined a PDF for a particular configuration $\vec{S}$ to be

$$P(\vec{S}) = \frac{\langle\eta_{\vec{S}}|\phi_{\vec{S}}\rangle}{\displaystyle\sum_{\vec{S}}\langle\eta_{\vec{S}}|\phi_{\vec{S}}\rangle} \tag{G.30}$$

and

$$\langle \widehat{A} \rangle_{\vec{S}} = \frac{\langle \eta_{\vec{S}} | \widehat{A} | \phi_{\vec{S}} \rangle}{\langle \eta_{\vec{S}} | \phi_{\vec{S}} \rangle} \,. \tag{G.31}$$

Eq. (G.29) can be evaluated by Monte Carlo.[2] We sample configurations $\vec{S}$ from the distribution $P(\vec{S})$ and calculate $\langle \widehat{A} \rangle_{\vec{S}}$ for each of those configurations. The arithmetic mean of the results of those calculations is the Monte Carlo estimator. However, direct sampling of $P(\vec{S})$ is practically impossible because it requires knowing normalization of $P(\vec{S})$ i.e. the denominator of Eq. (G.30). This normalization can only be found by evaluating all the $2^{(L+T)M}$ terms in the summation over $\vec{S}$, clearly impractical in a realistic calculation.

The Metropolis algorithm allows for the sampling of PDF's without knowing their normalizations. An excellent discussion can be found in Kalos and Whitlock [26, pp. 64-74]. Basically, we start with some initial configuration $\vec{S}$, propose a new configuration $\vec{S}'$ and accept or reject $\vec{S}'$ with an appropriate probability that involves the ratio of $P(\vec{S}')$ to $P(\vec{S})$. After obtaining enough samples $\vec{S}$ (with appropriate decorrelation time in between the samples if necessary), we can compute $\langle \widehat{A} \rangle_{\vec{S}}$'s and the Monte Carlo estimator.

What has gone unsaid thus far is that the probability $P(\vec{S})$ is not positive for all $\vec{S}$ because from Appendices C and D, we have

$$\langle \eta_{\vec{S}} | \phi_{\vec{S}} \rangle = \langle \phi_{\mathrm{T}} | \widehat{B}_{\vec{Y}}(\Lambda + \Theta, \Lambda)\, \widehat{B}_{\vec{X}}(\Lambda, 0) | \phi_{\mathrm{T}} \rangle \tag{G.32}$$

$$= \mathrm{Det}\left[(\Phi_{\mathrm{T}})^{\dagger} B_{\vec{Y}}(\Lambda + \Theta, \Lambda)\, B_{\vec{X}}(\Lambda, 0) \Phi_{\mathrm{T}} \right] , \tag{G.33}$$

which is just the determinant of a product of matrices and there's nothing *a priori* to prevent this determinant from being negative.[3]

To recap, in this appendix, we have shown how a different interpretation of the Hirsch transformation leads to an entirely different random walk i.e. that generated by the Metropolis algorithm. The formulas for the Monte Carlo estimator stated in Eqs. (G.28) to (G.31) are the takeaway from this appendix.

[2] In fact, equations of this form are the basis of the Variational Monte Carlo method.

[3] Note that Santos [36] discusses the sign problem in the context of a finite-temperature calculation instead of a ground state calculation.

Appendix H

Unequal-time Green's function

In this appendix we will show how to compute the unequal-time Green's function in a numerically stable way. The expected value of the the unequal-time Green's function for the ground state with a *particular* auxiliary-field path $\vec{S}$ and imaginary times $\tau_1 > \tau_2$ is defined as

$$G(\tau_1, \tau_2)_{\ell j} = \frac{\langle \Psi_0 | c_\ell(\tau_1) c_j^\dagger(\tau_2) | \Psi_0 \rangle}{\langle \Psi_0 | \Psi_0 \rangle} \tag{H.1}$$

$$= \frac{\langle \phi_{\mathrm{T}} | \widehat{\mathcal{P}} \dots \widehat{\mathcal{P}}\, c_\ell\, \widehat{\mathcal{P}} \dots \widehat{\mathcal{P}} c_j^\dagger \widehat{\mathcal{P}} \dots \widehat{\mathcal{P}} | \phi \rangle}{\langle \phi_{\mathrm{T}} | \widehat{\mathcal{P}} \dots \widehat{\mathcal{P}}\, \widehat{\mathcal{P}} \dots \widehat{\mathcal{P}} | \phi \rangle}\,, \tag{H.2}$$

where $\widehat{\mathcal{P}}$ is the ground state projection operator in Eq. (G.2). Then

$$G(\tau_1, \tau_2)_{\ell j} = \frac{\langle \phi_{\mathrm{T}} | \widehat{B}_{\vec{S}}(\Theta, \tau_1)\, c_\ell\, \widehat{B}_{\vec{S}}(\tau_1, \tau_2) c_j^\dagger\, \widehat{B}_{\vec{S}}(\tau_2, 0) | \phi \rangle}{\langle \phi_{\mathrm{T}} | \widehat{B}_{\vec{S}}(\Theta, 0) | \phi \rangle} \tag{H.3}$$

where the simplification to the last line happens in the same manner as the derivation leading to Eq. (G.27). Note that in Eq. (H.3), there is an extra projection interval $\widehat{B}_{\vec{S}}(\Theta, \tau_1)$ on the left hand side of $c_\ell(\tau_2)$ to project out the ground state from $|\phi_{\mathrm{T}}\rangle$. Because in the algorithm used in this thesis, $|\phi\rangle$ will be a walker from the equilibrated ensemble, there is no need for such a projection interval for the right hand side wave function.

Our strategy of computing the unequal-time Green's function is to manipulate Eq. (H.3) so as to put the creation and annihilation operators c_ℓ and $c_j^\dagger$ next to each other, which will allow us to use the formula for the equal-time Green's function in Appendix F. Readers are encouraged to read Appendix G before reading this appendix. This appendix is very important to this thesis and elaborates on the discussion in Assaad [42, pp.123-5].

Let $\widehat{H}$ be the Hamiltonian with coefficient matrix H:

$$\widehat{H} = \sum_{ij} c_i^\dagger H_{ij} c_j \tag{H.4}$$

$$= \vec{c}^{\,\dagger} H \vec{c}\,, \tag{H.5}$$

where in the second equality, $\vec{c}^{\,\dagger}$ should be thought of as a row matrix of creation operators and $\vec{c}$ a column matrix of annihilation operators.

First let's evaluate the following commutator:

$$\left[\vec{c}^{\,\dagger} H \vec{c}, c_y\right] = \sum_{ij} H_{ij} \left[c_i^\dagger c_j, c_y\right] . \tag{H.6}$$

Using the following identity for operators $\widehat{X}$,$\widehat{Y}$ and $\widehat{Z}$:

$$[\widehat{X}\widehat{Y}, \widehat{Z}] = \widehat{X}[\widehat{Y}, \widehat{Z}] + [\widehat{X}, \widehat{Z}]\widehat{Y} , \tag{H.7}$$

we have

$$\left[\vec{c}^{\,\dagger} H \vec{c}, c_y\right] = \sum_{ij} H_{ij} \left(c_i^\dagger [c_j, c_y] + [c_i^\dagger, c_y] c_j\right) . \tag{H.8}$$

Using the anticommutation relations for second-quantized operators, the commutators in Eq. (H.8) are

$$[c_j, c_y] = -2c_y c_j \tag{H.9}$$

$$[c_i^\dagger, c_y] = \delta_{iy} - 2c_y c_i^\dagger . \tag{H.10}$$

Putting Eqs. (H.9) and (H.10) into Eq. (H.8):

$$\left[\vec{c}^{\,\dagger} H \vec{c}, c_y\right] = \sum_{ij} H_{ij} \left(-2c_i^\dagger c_y c_j + \delta_{iy} c_j - 2c_y c_i^\dagger c_j\right) . \tag{H.11}$$

Grouping the first and last term in the sum and recognizing the anticommutator of $c_i^\dagger$ and c_y:

$$\left[\vec{c}^{\,\dagger} H \vec{c}, c_y\right] = \sum_{ij} H_{ij} \left(-2\left\{c_i^\dagger, c_y\right\} c_j + \delta_{iy} c_j\right) \tag{H.12}$$

$$= \sum_{ij} H_{ij} \left(-2\delta_{iy} + \delta_{iy}\right) c_j \tag{H.13}$$

$$\left[\vec{c}^{\,\dagger} H \vec{c}, c_y\right] = -\sum_{j} H_{yj} c_j . \tag{H.14}$$

We now turn our attention to the time evolution of the annihilation operator c_y:

$$c_y(\tau) = \mathrm{e}^{\tau \widehat{H}} c_y \, \mathrm{e}^{-\tau \widehat{H}} = \mathrm{e}^{\tau \vec{c}^{\,\dagger} H \vec{c}} c_y \, \mathrm{e}^{-\tau \vec{c}^{\,\dagger} H \vec{c}} . \tag{H.15}$$

Because $\vec{c}^{\,\dagger} H \vec{c}$ is just a "scalar" (in the sense that it is a 1×1 matrix), we can differentiate Eq. (H.15) with respect to τ:

$$\frac{\partial c_y(\tau)}{\partial \tau} = \frac{\partial [\mathrm{e}^{\tau \widehat{H}} c_y \, \mathrm{e}^{-\tau \widehat{H}}]}{\partial \tau} \tag{H.16}$$

$$= \left(\frac{\partial \mathrm{e}^{\tau \vec{c}^{\,\dagger} H \vec{c}}}{\partial \tau}\right) c_y \, \mathrm{e}^{-\tau \vec{c}^{\,\dagger} H \vec{c}} + \mathrm{e}^{\tau \vec{c}^{\,\dagger} H \vec{c}} c_y \left(\frac{\partial \mathrm{e}^{-\tau \vec{c}^{\,\dagger} H \vec{c}}}{\partial \tau}\right) \tag{H.17}$$

$$= \left(\vec{c}^{\,\dagger} H \vec{c}\right) \mathrm{e}^{\tau \vec{c}^{\,\dagger} H \vec{c}} c_y \, \mathrm{e}^{-\tau \vec{c}^{\,\dagger} H \vec{c}} + \mathrm{e}^{\tau \vec{c}^{\,\dagger} H \vec{c}} c_y \left(-\vec{c}^{\,\dagger} H \vec{c}\right) \mathrm{e}^{-\tau \vec{c}^{\,\dagger} H \vec{c}} . \tag{H.18}$$

Because $\vec{c}^\dagger H \vec{c}$ commutes with itself in the first term:

$$\frac{\partial c_y(\tau)}{\partial \tau} = \mathrm{e}^{\tau \vec{c}^\dagger H \vec{c}} \left(\vec{c}^\dagger H \vec{c} \right) c_y \, \mathrm{e}^{-\tau \vec{c}^\dagger H \vec{c}} - \mathrm{e}^{\tau \vec{c}^\dagger H \vec{c}} c_y \left(\vec{c}^\dagger H \vec{c} \right) \mathrm{e}^{-\tau \vec{c}^\dagger H \vec{c}} \tag{H.19}$$

$$= \mathrm{e}^{\tau \vec{c}^\dagger H \vec{c}} \left[\vec{c}^\dagger H \vec{c}, c_y \right] \mathrm{e}^{-\tau \vec{c}^\dagger H \vec{c}} \, . \tag{H.20}$$

Now we can use Eq. (H.14) to write

$$\frac{\partial c_y(\tau)}{\partial \tau} = \mathrm{e}^{\tau \vec{c}^\dagger H \vec{c}} \left(- \sum_j H_{yj} c_j \right) \mathrm{e}^{-\tau \vec{c}^\dagger H \vec{c}} \tag{H.21}$$

$$= - \sum_j H_{yj} \left(\mathrm{e}^{\tau \vec{c}^\dagger H \vec{c}} c_j \mathrm{e}^{-\tau \vec{c}^\dagger H \vec{c}} \right) \tag{H.22}$$

$$= - \sum_j H_{yj} c_j(\tau) \tag{H.23}$$

$$\frac{\partial c_y(\tau)}{\partial \tau} = - \left[H \vec{c}(\tau) \right]_y \, . \tag{H.24}$$

We can rewrite Eq. (H.24) as a matrix differential equation:

$$\frac{\partial}{\partial \tau} \vec{c}(\tau) = -H \vec{c}(\tau) \, , \tag{H.25}$$

whose solution is

$$\vec{c}(\tau + \Delta\tau) = \mathrm{e}^{-\Delta\tau H} \vec{c}(\tau) \, . \tag{H.26}$$

Element-wise, this means

$$c_y(\tau + \Delta\tau) = \left(\mathrm{e}^{-\Delta\tau H} \vec{c}(\tau) \right)_y \, . \tag{H.27}$$

On the other hand, the time evolution of the operator c_y in the Heisenberg picture is

$$c_y(\tau + \Delta\tau) = \mathrm{e}^{\Delta\tau \widehat{H}} c_y(\tau) \mathrm{e}^{-\Delta\tau \widehat{H}} \, . \tag{H.28}$$

Together, Eqs. (H.27) and (H.28) say

$$\mathrm{e}^{\Delta\tau \widehat{H}} c_y(\tau) \mathrm{e}^{-\Delta\tau \widehat{H}} = \left(\mathrm{e}^{-\Delta\tau H} \vec{c}(\tau) \right)_y \tag{H.29}$$

Creation operators obey a similar differential equation:[1]

$$\frac{\partial}{\partial \tau} \vec{c}^\dagger(\tau) = \vec{c}^\dagger(\tau) H \, , \tag{H.30}$$

whose solution is

$$\mathrm{e}^{-\Delta\tau \widehat{H}} c_y^\dagger(\tau) \mathrm{e}^{\Delta\tau \widehat{H}} = \left(\vec{c}^\dagger(\tau) \, \mathrm{e}^{-\Delta\tau H} \right)_y \, . \tag{H.31}$$

[1] An intermediate step is proving $[\vec{c}^\dagger H \vec{c}, c_y^\dagger] = \sum_i c_i^\dagger H_{iy}$.

Now let the propagator that propagates a Slater determinant from imaginary time τ_2 to time τ_1 for a fixed auxiliary-field path $\vec{s}$ (with $\tau_1 > \tau_2$) be

$$\hat{B}_{\vec{s}}(\tau_1, \tau_2) = \prod_{\ell=L}^{1} \mathrm{e}^{-\Delta\tau \hat{H}(\vec{x}^{(\ell)})}, \tag{H.32}$$

where ℓ indexes the time slices between τ_1 and τ_2 and $\vec{x}^{(\ell)}$ is the collection of all auxiliary fields over all sites at one particular time slice ℓ. Technically, the equality in Eq. (H.32) is not correct because as shown in Eq. (G.8), $\mathrm{e}^{-\Delta\tau \hat{H}(\vec{x}^{(\ell)})}$ also has a scalar prefactor. However, because we always calculate the unequal-time Green's function as a ratio (Eq. (H.3)), that prefactor will cancel out in the end. In Eq. (H.32), ℓ runs backward from L to 1 to denote that earlier times appear to the right per convention. The inverse propagator that takes a determinant from time τ_1 back to time τ_2 is

$$\hat{B}^{-1}_{\vec{s}}(\tau_1, \tau_2) = \prod_{\ell=1}^{L} \mathrm{e}^{\Delta\tau \hat{H}(\vec{x}^{(\ell)})}, \tag{H.33}$$

where the small-time operators are simply applied in reverse order. Note that for Eqs. (H.32) and (H.33) to be inverses of each other, the auxiliary field configurations $\vec{x}^{(\ell)}$ of each time slice ℓ must match. This is automatically satisfied because we are dealing with a *single* auxiliary-field configuration $\vec{s}$ here. In the notation of Eq. (3.48), the matrix version of Eq. (H.32) is

$$B_{\vec{s}}(\tau_1, \tau_2) = \prod_{\ell=L}^{1} B(\vec{x}^{(\ell)}). \tag{H.34}$$

Using the definition of propagators in Eqs. (H.32) and (H.33), we can compute the following quantity

$$\hat{B}^{-1}_{\vec{s}}(\tau_1, \tau_2) c_y \hat{B}_{\vec{s}}(\tau_1, \tau_2) = \underbrace{\left[\mathrm{e}^{\Delta\tau \hat{H}_1} \dots \mathrm{e}^{\Delta\tau \hat{H}_{L-1}} \mathrm{e}^{\Delta\tau \hat{H}_L}\right]}_{L \text{ times}} c_y \underbrace{\left[\mathrm{e}^{-\Delta\tau \hat{H}_L} \mathrm{e}^{-\Delta\tau \hat{H}_{L-1}} \dots \mathrm{e}^{-\Delta\tau \hat{H}_1}\right]}_{L \text{ times}} \tag{H.35}$$

$$= \mathrm{e}^{\Delta\tau \hat{H}_1} \dots \mathrm{e}^{\Delta\tau \hat{H}_{L-1}} \left[\mathrm{e}^{\Delta\tau \hat{H}_L} c_y \, \mathrm{e}^{-\Delta\tau \hat{H}_L}\right] \mathrm{e}^{-\Delta\tau \hat{H}_{L-1}} \dots \mathrm{e}^{-\Delta\tau \hat{H}_1}. \tag{H.36}$$

We can then use Eq. (H.29) to push the matrix $\mathrm{e}^{-\Delta\tau H_L}$ to the left of the operators:

$$\hat{B}^{-1}_{\vec{s}}(\tau_1, \tau_2) c_y \hat{B}_{\vec{s}}(\tau_1, \tau_2) = \mathrm{e}^{\Delta\tau \hat{H}_1} \dots \mathrm{e}^{\Delta\tau \hat{H}_{L-1}} \left[\mathrm{e}^{-\Delta\tau H} \vec{c}\right]_y \mathrm{e}^{-\Delta\tau \hat{H}_{L-1}} \dots \mathrm{e}^{-\Delta\tau \hat{H}_1} \tag{H.37}$$

$$= \sum_i \left(\mathrm{e}^{-\Delta\tau H_L}\right)_{yi} \mathrm{e}^{\Delta\tau \hat{H}_1} \dots \mathrm{e}^{\Delta\tau \hat{H}_{L-2}} \left[\mathrm{e}^{\Delta\tau \hat{H}_{L-1}} c_i \mathrm{e}^{-\Delta\tau \hat{H}_{L-1}}\right] \mathrm{e}^{-\Delta\tau \hat{H}_{L-2}} \dots \mathrm{e}^{-\Delta\tau \hat{H}_1} \tag{H.38}$$

$$= \sum_{ij} \left(\mathrm{e}^{-\Delta\tau H_L}\right)_{yi} \left(\mathrm{e}^{-\Delta\tau H_{L-1}}\right)_{ij} \mathrm{e}^{\Delta\tau \hat{H}_1} \dots \left[\mathrm{e}^{\Delta\tau \hat{H}_{L-2}} c_j \mathrm{e}^{-\Delta\tau \hat{H}_{L-2}}\right] \dots \mathrm{e}^{-\Delta\tau \hat{H}_1}. \tag{H.39}$$

By repeating the same trick many times, we eventually obtain

$$\widehat{B}_{\vec{S}}^{-1}(\tau_1,\tau_2)c_y\widehat{B}_{\vec{S}}(\tau_1,\tau_2) = \sum_{ij\ldots mn} \left(\mathrm{e}^{-\Delta\tau H_L}\right)_{yi} \left(\mathrm{e}^{-\Delta\tau H_{L-1}}\right)_{ij} \ldots \left(\mathrm{e}^{-\Delta\tau H_1}\right)_{mn} c_n \,. \tag{H.40}$$

By recognizing multiplication, we can simplify the RHS to

$$\widehat{B}_{\vec{S}}^{-1}(\tau_1,\tau_2)c_y\widehat{B}_{\vec{S}}(\tau_1,\tau_2) = \left[\left(\mathrm{e}^{-\Delta\tau H_L}\right)\ldots\left(\mathrm{e}^{-\Delta\tau H_1}\right)\,\vec{c}\,\right]_y \,. \tag{H.41}$$

Using the definition of the propagator matrix in Eq. (H.34), the above equation can be written more compactly as

$$\widehat{B}_{\vec{S}}^{-1}(\tau_1,\tau_2)c_y\widehat{B}_{\vec{S}}(\tau_1,\tau_2) = \left[B_{\vec{S}}(\tau_1,\tau_2)\,\vec{c}\,(\tau_2)\right]_y \,, \tag{H.42}$$

where $B_{\vec{S}}(\tau_1,\tau_2)$ is the matrix of the operator $\widehat{B}_{\vec{S}}(\tau_1,\tau_2)$. By starting from Eq. (H.31) and following a similar procedure,[2] we can obtain:

$$\widehat{B}_{\vec{S}}(\tau_1,\tau_2)c_y^\dagger\widehat{B}_{\vec{S}}^{-1}(\tau_1,\tau_2) = \left[\vec{c}^{\,\dagger}(\tau_1)B_{\vec{S}}(\tau_1,\tau_2)\right]_y \tag{H.43}$$

where the imaginary time ordering is $\tau_1 > \tau_2$.

From Eq. (H.3), the unequal-time Green's function for the ground state over a particular path $\vec{S}$ in auxiliary-field space is

$$G(\tau_1,\tau_2)_{\ell j} = \frac{\langle\phi_{\mathrm{T}}|\widehat{B}_{\vec{S}}(\Theta,\tau_1)\,c_\ell\,\widehat{B}_{\vec{S}}(\tau_1,\tau_2)c_j^\dagger\,\widehat{B}_{\vec{S}}(\tau_2,0)|\phi\rangle}{\langle\phi_{\mathrm{T}}|\widehat{B}_{\vec{S}}(\Theta,0)|\phi\rangle} \,. \tag{H.44}$$

where $\Theta > \tau_1 > \tau_2$. We note that because these propagators use the same auxiliary-field path $\vec{S}$, we have

$$\widehat{B}_{\vec{S}}(\Theta,\tau_1) = \widehat{B}_{\vec{S}}(\Theta,\tau_2)\widehat{B}_{\vec{S}}^{-1}(\tau_1,\tau_2) \,, \tag{H.45}$$

which says that going forward from τ_1 to Θ is the same as going backward from τ_1 to τ_2 and then from τ_2 forward to Θ. Thus

$$G(\tau_1,\tau_2)_{\ell j} = \frac{\langle\phi_{\mathrm{T}}|\widehat{B}_{\vec{S}}(\Theta,\tau_2)\widehat{B}_{\vec{S}}^{-1}(\tau_1,\tau_2)\,c_\ell\,\widehat{B}_{\vec{S}}(\tau_1,\tau_2)c_j^\dagger\,\widehat{B}_{\vec{S}}(\tau_2,0)|\phi\rangle}{\langle\phi_{\mathrm{T}}|\widehat{B}_{\vec{S}}(\Theta,0)|\phi\rangle} \,. \tag{H.46}$$

Using Eq. (H.42), we have

$$G(\tau_1,\tau_2)_{\ell j} = \frac{\langle\phi_{\mathrm{T}}|\widehat{B}_{\vec{S}}(\Theta,\tau_2)\,[B(\tau_1,\tau_2)\,\vec{c}\,]_\ell\,c_j^\dagger\widehat{B}_{\vec{S}}(\tau_2,0)|\phi\rangle}{\langle\phi_{\mathrm{T}}|\widehat{B}_{\vec{S}}(\Theta,0)|\phi\rangle} \tag{H.47}$$

$$= \sum_x [B(\tau_1,\tau_2)]_{\ell y}\,\frac{\langle\phi_{\mathrm{T}}|\widehat{B}_{\vec{S}}(\Theta,\tau_2)c_y\,c_j^\dagger\widehat{B}_{\vec{S}}(\tau_2,0)|\phi\rangle}{\langle\phi_{\mathrm{T}}|\widehat{B}_{\vec{S}}(\Theta,0)|\phi\rangle} \tag{H.48}$$

$$= \sum_x [B(\tau_1,\tau_2)]_{\ell y}\,\frac{\langle\phi_{\mathrm{L}}|c_y\,c_j^\dagger|\phi_{\mathrm{R}}\rangle}{\langle\phi_{\mathrm{L}}|\phi_{\mathrm{R}}\rangle} \tag{H.49}$$

[2] This time we push the matrices to the right of the operators instead of to the left.

where we have defined

$$|\phi_{\rm R}\rangle = \widehat{B}_{\vec{S}}(\tau_2, 0)\,|\phi\rangle \tag{H.50}$$

$$\langle\phi_{\rm L}| = \langle\phi_{\rm T}|\,\widehat{B}_{\vec{S}}(\Theta, \tau_2)\,. \tag{H.51}$$

The latter is obtained by

$$|\phi_{\rm L}\rangle = \widehat{B}^{-1}_{\vec{S}}(\Theta, \tau_2)\,|\phi_{\rm T}\rangle\ . \tag{H.52}$$

We recognize

$$G(\tau_2)_{yj} = \frac{\langle\phi_{\rm L}|c_y\,c_j^\dagger|\phi_{\rm R}\rangle}{\langle\phi_{\rm L}|\phi_{\rm T}\rangle} \tag{H.53}$$

as the equal-time Green's function for two Slater determinants $|\phi_{\rm L}\rangle$ and $|\phi_{\rm R}\rangle$ which we know how to calculate from Appendix F. We can then write

$$G(\tau_1, \tau_2)_{\ell j} = \sum_x [B(\tau_1, \tau_2)]_{\ell y}\, G(\tau_2)_{yj}\,. \tag{H.54}$$

If we think of $G(\tau_1, \tau_2)_{\ell j}$ as the (ℓ, j)-th element of a square matrix $G(\tau_1, \tau_2)$ then Eq. (H.54) is equivalent to

$$G(\tau_1, \tau_2) = B(\tau_1, \tau_2)G(\tau_2) \tag{H.55}$$

Alternatively, if we use Eq. (H.43), we will arrive at a different expression:

$$G(\tau_1, \tau_2)_{\ell j} = \frac{\langle\phi_{\rm T}|\widehat{B}_{\vec{S}}(\Theta, \tau_1)\,c_\ell\,\widehat{B}_{\vec{S}}(\tau_1, \tau_2)c_j^\dagger\,\widehat{B}_{\vec{S}}(\tau_2, 0)|\phi\rangle}{\langle\phi_{\rm T}|\widehat{B}_{\vec{S}}(\Theta, 0)|\phi\rangle} \tag{H.56}$$

$$= \frac{\langle\phi_{\rm T}|\widehat{B}_{\vec{S}}(\Theta, \tau_1)\,c_\ell\,\widehat{B}_{\vec{S}}(\tau_1, \tau_2)c_j^\dagger\,\widehat{B}^{-1}_{\vec{S}}(\tau_1, \tau_2)\widehat{B}_{\vec{S}}(\tau_1, 0)|\phi\rangle}{\langle\phi_{\rm T}|\widehat{B}_{\vec{S}}(\Theta, 0)|\phi\rangle} \tag{H.57}$$

$$= \frac{\langle\phi_{\rm T}|\widehat{B}_{\vec{S}}(\Theta, \tau_1)\,c_\ell\,[\,\overrightarrow{c}^\dagger B(\tau_1, \tau_2)]_j\widehat{B}_{\vec{S}}(\tau_1, 0)|\phi\rangle}{\langle\phi_{\rm T}|\widehat{B}_{\vec{S}}(\Theta, 0)|\phi\rangle} \tag{H.58}$$

$$= \sum_y B(\tau_1, \tau_2)_{yj}\,\frac{\langle\phi_{\rm T}|\widehat{B}_{\vec{S}}(\Theta, \tau_1)\,c_\ell\,c_y^\dagger\widehat{B}_{\vec{S}}(\tau_1, 0)|\phi\rangle}{\langle\phi_{\rm T}|\widehat{B}_{\vec{S}}(\Theta, 0)|\phi\rangle} \tag{H.59}$$

$$= \sum_y B(\tau_1, \tau_2)_{yj}\,\frac{\langle\phi_{\rm L}|c_\ell c_y^\dagger|\phi_{\rm R}\rangle}{\langle\phi_{\rm L}|\phi_{\rm R}\rangle} \tag{H.60}$$

where

$$\langle\phi_{\rm L}| = \langle\phi_{\rm T}|\,\widehat{B}_{\vec{S}}(\Theta, \tau_1) \tag{H.61}$$

$$|\phi_{\rm R}\rangle = \widehat{B}_{\vec{S}}(\tau_1, 0)\,|\phi\rangle\ . \tag{H.62}$$

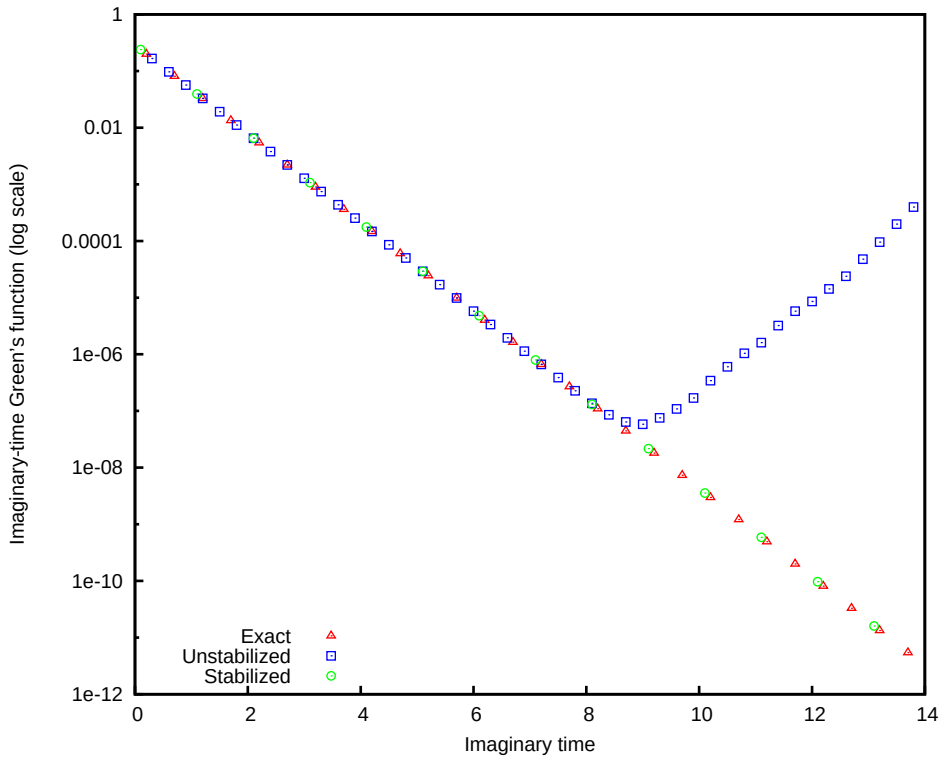


Figure H.1: Comparisons of the accuracy of the calculation of the on-site unequal-time Green's function i.e. $\langle\Psi_0|c_{i\uparrow}(\tau)c^\dagger_{i\uparrow}(0)|\Psi_0\rangle$ for a non-interacting one-dimensional 5-site lattice with 3 spin-$\uparrow$ and 3 spin-$\downarrow$ electrons using three different methods: exact, stabilized and unstabilized calculations. The vertical axis is in logarithmic scale.

Thus

$$G(\tau_1,\tau_2)_{\ell j} = \sum_y G(\tau_1)_{ly} B_{yj}(\tau_1,\tau_2) \tag{H.63}$$

$$G(\tau_1,\tau_2) = G(\tau_1) B(\tau_1,\tau_2) \tag{H.64}$$

where $B(\tau_1,\tau_2)$ is just a product of matrices that has been defined in Eq. (H.34). Eqs. (H.55) and (H.64) are the desired results of this appendix: they express the unequal-time Green's function as a matrix product of one-body propagators and the equal-time Green's function for two Slater determinants.

However, as will be discussed in Appendix J, if we naively carry out Eqs. (H.55) and (H.64) as written, the repeated multiplication of the short-time propagators to form $B(\tau_1,\tau_2)$ will lead to numerical instabilities. In Fig. H.1 we show a comparison of the accuracy of three different ways of calculating the on-site unequal-time Green's function i.e. $\langle\Psi_0|c_{i\uparrow}(\tau)c^\dagger_{i\uparrow}(0)|\Psi_0\rangle$ vs the imaginary time τ for a non-interacting one-dimensional 5-site lattice with 3 spin-$\uparrow$ and 3 spin-$\downarrow$ electrons. The three methods agree well up to $\tau = 8$. The unstabilized calculation described thus far gives results that diverge exponentially from the exact calculation while the stabilized calculation (which we will describe below) agrees very well with the exact calculation. Thus, we need a way to stabilize this procedure.

First we show that the equal-time Green's function is a projector.

$$G(\tau)^2 = G(\tau)G(\tau) \tag{H.65}$$
$$= \left[I - Q(P^\dagger Q)^{-1}P^\dagger\right]\left[I - Q(P^\dagger Q)^{-1}P^\dagger\right] \tag{H.66}$$
$$= I - Q(P^\dagger Q)^{-1}P^\dagger - Q(P^\dagger Q)^{-1}P^\dagger + \left[Q(P^\dagger Q)^{-1}P^\dagger\right]\left[Q(P^\dagger Q)^{-1}P^\dagger\right] \tag{H.67}$$
$$= I - 2Q(P^\dagger Q)^{-1}P^\dagger + Q(P^\dagger Q)^{-1}\cancel{(P^\dagger Q)(P^\dagger Q)^{-1}}P^\dagger \tag{H.68}$$
$$= I - 2Q(P^\dagger Q)^{-1}P^\dagger + Q(P^\dagger Q)^{-1}P^\dagger \tag{H.69}$$
$$= I - Q(P^\dagger Q)^{-1}P^\dagger \tag{H.70}$$
$$G(\tau)^2 = G(\tau)\,. \tag{H.71}$$

Thus the equal-time Green's function $G(\tau)$ is a projector. Next, we show a temporal decomposition of the unequal-time Green's function. Suppose $\tau_1 > \tau_2 > \tau_3$ then

$$G(\tau_1,\tau_3) = B_{\vec{S}}(\tau_1,\tau_3)G(\tau_3) \qquad \text{from Eq. (H.55)} \tag{H.72}$$
$$= B_{\vec{S}}(\tau_1,\tau_3)G^2(\tau_3) \qquad \text{from Eq. (H.71)} \tag{H.73}$$
$$= \left[B_{\vec{S}}(\tau_1,\tau_3)G(\tau_3)\right]G(\tau_3) \tag{H.74}$$
$$= G(\tau_1,\tau_3)G(\tau_3) \qquad \text{Eq. (H.55) again} \tag{H.75}$$
$$= G(\tau_1,\tau_3)B^{-1}_{\vec{S}}(\tau_2,\tau_3)[B_{\vec{S}}(\tau_2,\tau_3)G(\tau_3)] \qquad \text{insert forward and backward propagators} \tag{H.76}$$
$$= G(\tau_1,\tau_3)B^{-1}_{\vec{S}}(\tau_2,\tau_3)G(\tau_2,\tau_3) \qquad \text{Eq. (H.55) applied to the last 2 terms} \tag{H.77}$$
$$= [G(\tau_1)B_{\vec{S}}(\tau_1,\tau_3)]B^{-1}_{\vec{S}}(\tau_2,\tau_3)G(\tau_2,\tau_3) \qquad \text{Eq. (H.64) applied to } G(\tau_1,\tau_3) \tag{H.78}$$
$$= [G(\tau_1)B_{\vec{S}}(\tau_1,\tau_2)]G(\tau_2,\tau_3) \qquad \text{combine the two propagators} \tag{H.79}$$
$$= G(\tau_1,\tau_2)G(\tau_2,\tau_3)\,. \qquad \text{Eq. (H.64) applied to the first 2 terms} \tag{H.80}$$

Using this composition property, we can break up a large τ into a set of N smaller intervals of length $\varepsilon = (\tau_1 - \tau_2)/N$ so that

$$G(\tau_1,\tau_2) = \prod_{n=0}^{N-1} G(\tau_2 + (n+1)\varepsilon, \tau_2 + n\varepsilon)\,. \tag{H.81}$$

Applying Eq. (H.55) to each of the "small-time" Green's functions $G(\tau_2 + (n+1)\varepsilon, \tau_2 + n\varepsilon)$, we can write Eq. (H.81) more explicitly as

$$G(\tau_1,\tau_2) = B(\tau_1,\tau_1-\varepsilon)G(\tau_1-\varepsilon)\ldots B(\tau_2+2\varepsilon,\tau_2+\varepsilon)G(\tau_2+\varepsilon)B(\tau_2+\varepsilon,\tau_2)G(\tau_2)\,. \tag{H.82}$$

Because the equal-time Green's functions are always well-behaved because of the numerical stabilization discussed in Appendix J, we can ensure that $G(\tau_1,\tau_2)$ is also well-behaved by choosing ε to be small enough such that $B(\tau+\varepsilon,\tau)$ is well-behaved.

Appendix I

Back-propagation

This appendix derives the back-propagation procedure and estimator as stated in Section 3.10.2 and has adapted materials from Purwanto [45].

The back-propagated estimator in Eq. (3.80) is

$$\langle \widehat{A} \rangle_{\rm bp} = \frac{\langle \phi_{\rm T} | {\rm e}^{-\tau_{\rm bp}\hat{H}} \widehat{A} | \Psi_0 \rangle}{\langle \phi_{\rm T} | {\rm e}^{-\tau_{\rm bp}\hat{H}} | \Psi_0 \rangle} \tag{I.1}$$

$$= \lim_{\tau\to\infty} \frac{\langle \Psi_{\rm BP} | \widehat{A} | \Psi^{(\tau)} \rangle}{\langle \Psi_{\rm BP} | \Psi^{(\tau)} \rangle} \tag{I.2}$$

where

$$|\Psi^{(\tau)}\rangle = {\rm e}^{-\tau\widehat{H}} \, |\phi_{\rm T}\rangle \tag{I.3}$$

is the regular "forward" walk at time τ and

$$\langle \Psi_{\rm BP}| = \langle \phi_{\rm T}| \, {\rm e}^{-\tau_{\rm bp}\hat{H}} \tag{I.4}$$

is the back-propagated wave function that we need an efficient way to calculate. Note that to reduce clutter, we will freely drop the normalization factor ${\rm e}^{E_{\rm T}}$ from the propagators.

The back-propagation technique originates by carefully looking at the denominator of Eq. (I.1). At any imaginary time τ, the ensemble of random walkers that have been generated represents ${\rm e}^{-\tau\hat{H}} \, |\phi_{\rm T}\rangle$. If we let the random walk proceed for another imaginary time interval $\tau_{\rm bp}$ then we will get a representation of ${\rm e}^{-\tau_{\rm bp}\widehat{H}} {\rm e}^{-\tau\widehat{H}} \, |\phi_{\rm T}\rangle$. However, the Hermiticity of ${\rm e}^{-\tau_{\rm bp}\widehat{H}}$ allows us to apply it to the *left* on $\langle\phi_{\rm T}|$:

$$\langle \phi_{\rm T} | {\rm e}^{-\tau_{\rm bp}\hat{H}} {\rm e}^{-\tau\hat{H}} | \phi_{\rm T} \rangle = \begin{cases} \langle \phi_{\rm T}| \left[{\rm e}^{-\tau_{\rm bp}\hat{H}} {\rm e}^{-\tau\hat{H}} \, |\phi_{\rm T}\rangle \right] & \text{if } {\rm e}^{-\tau_{\rm bp}\hat{H}} \text{ is applied to the right} \\ \left[\langle \phi_{\rm T}| \, {\rm e}^{-\tau_{\rm bp}\hat{H}} \right] \left[{\rm e}^{-\tau\hat{H}} \, |\phi_{\rm T}\rangle \right] & \text{if } {\rm e}^{-\tau_{\rm bp}\hat{H}} \text{ is applied to the left} \, . \end{cases} \tag{I.5}$$

These two interpretations can give exactly the same result using the same auxiliary field paths in ${\rm e}^{-\tau_{\rm bp}\hat{H}}$ if the walkers generated by the operation $\langle \phi_{\rm T}| \, {\rm e}^{-\tau_{\rm bp}\hat{H}}$ must be matched *one-to-one* with the walkers representing ${\rm e}^{-\tau\hat{H}} \, |\phi_{\rm T}\rangle$. The operators ${\rm e}^{-\tau_{\rm bp}\hat{H}}$ and ${\rm e}^{-\tau\hat{H}}$ belong to different imaginary time segments and therefor have different paths.

First we will define some necessary notation. Consider a walker $|\phi_{\rm j}^{(\tau)}\rangle$ at imaginary time τ. An application of $\widehat{B}_{\rm K/2}\widehat{B}_{\rm V}(\vec{x}_{\rm j})\widehat{B}_{\rm K/2}$ evolve the walker into the next time step $|\phi_{\rm j}^{(\tau+\Delta\tau)}\rangle$. We have included a subscript j in $\vec{x}$ to emphasize that these configurations are different

for each walker. Because these operators together depend on the configurations at time τ which are stochastically sampled, these operators are also stochastic. Let's define the "total" propagator at imaginary time τ to be

$$\widehat{B}_{\mathrm{j}}^{(\tau)} \equiv \widehat{B}_{\mathrm{K}/2}\widehat{B}_{\mathrm{V}}(\vec{x}_{\mathrm{j}})\widehat{B}_{\mathrm{K}/2} \tag{I.6}$$

and the time-ordered product of these stochastic operators that transforms the walker $|\phi_{\mathrm{j}}^{(\tau)}\rangle$ into the walker $|\phi_{\mathrm{j}}^{(\tau')}\rangle$ to be

$$\widehat{B}_{\mathrm{j}}^{(\tau':\tau)} \equiv \widehat{B}_{\mathrm{j}}^{(\tau'-\Delta\tau)}\widehat{B}_{\mathrm{j}}^{(\tau'-2\Delta\tau)}\dots\widehat{B}_{\mathrm{j}}^{(\tau+\Delta\tau)}\widehat{B}_{\mathrm{j}}^{(\tau)} \tag{I.7}$$

where we have assumed $\tau' > \tau$. Because of the one-to-one correspondence between $\vec{x}_{\mathrm{j}}$ and $\widehat{B}_{\mathrm{j}}^{(\tau)}$, Eq. (I.7) also defines a "path" in auxiliary field space and relates the two walkers by:

$$|\phi_{\mathrm{j}}^{(\tau')}\rangle = \widehat{B}_{\mathrm{j}}^{(\tau':\tau)}\,|\phi_{\mathrm{j}}^{(\tau)}\rangle\ . \tag{I.8}$$

Similarly for the weight, we can define

$$\mathcal{F}^{(\tau)} \equiv \mathrm{e}^{\Delta\tau[E_{\mathrm{T}}-U(N_\uparrow+N_\downarrow)/2]} \tag{I.9}$$

to be the factor that multiplies the weight at every iteration (even though it doesn't actually depend on τ) and

$$\mathcal{F}^{(\tau':\tau)} \equiv \mathcal{F}^{(\tau'-\Delta\tau)}\mathcal{F}^{(\tau'-2\Delta\tau)}\dots\mathcal{F}^{(\tau+\Delta\tau)}\mathcal{F}^{(\tau)}\ . \tag{I.10}$$

A single application of the propagator in Eq. (I.6) updates the weight according to Eq. (E.35)

$$w_{\mathrm{j}}^{(\tau+\Delta\tau)} = \frac{\langle\phi_{\mathrm{T}}|\phi_{\mathrm{j}}^{(\tau+\Delta\tau)}\rangle}{\langle\phi_{\mathrm{T}}|\phi_{\mathrm{j}}^{(\tau)}\rangle}\,\mathcal{F}^{(\tau)}w_{\mathrm{j}}^{(\tau)}\ . \tag{I.11}$$

It is straightforward to show that

$$w_{\mathrm{j}}^{(\tau')} = \frac{\langle\phi_{\mathrm{T}}|\phi_{\mathrm{j}}^{(\tau')}\rangle}{\langle\phi_{\mathrm{T}}|\phi_{\mathrm{j}}^{(\tau)}\rangle}\mathcal{F}^{(\tau':\tau)}w_{\mathrm{j}}^{(\tau)}\ . \tag{I.12}$$

Suppose our population of walkers represent the wave function at time τ as

$$|\Psi^{(\tau)}\rangle = \sum_{\mathrm{j}} w_{\mathrm{j}}^{(\tau)}\frac{|\phi_{\mathrm{j}}^{(\tau)}\rangle}{\langle\phi_{\mathrm{T}}|\phi_{\mathrm{j}}^{(\tau)}\rangle}\ . \tag{I.13}$$

From now on τ will be a constant. Now if we propagate these walkers further to time $\tau' \equiv \tau + \tau_{\mathrm{bp}}$, the wave function will become

$$|\Psi^{(\tau')}\rangle = \mathrm{e}^{-\tau_{\mathrm{bp}}\widehat{H}}\,|\Psi^{(\tau)}\rangle = \sum_{\mathrm{j}} w_{\mathrm{j}}^{(\tau')}\frac{|\phi_{\mathrm{j}}^{(\tau')}\rangle}{\langle\phi_{\mathrm{T}}|\phi_{\mathrm{j}}^{(\tau')}\rangle}\ . \tag{I.14}$$

Now we will show that we can indeed reuse the same propagation paths (in the sense of Eq. (I.7)) that took us from τ to τ' to obtain $|\phi_{\mathrm{BP}}\rangle$ in Eq. (I.4).

Let's denote the walkers in the back-propagated population by $|\eta_{\mathrm{j}}^{(\theta)}\rangle$ and their weights $u_{\mathrm{j}}^{(\tau)}$. The imaginary time θ ranges from 0 to τ_{bp}. The subscript j of a walker indicates the path that is back-traversed by that walker i.e. the walker $|\phi_{\mathrm{j}}^{(\theta)}\rangle$ back-traverses the auxiliary field path $\widehat{B}^{(\tau':\tau)}$ that propagated the walker $|\phi_{\mathrm{j}}^{(\tau)}\rangle$ into $|\phi_{\mathrm{j}}^{(\tau')}\rangle$.

The backward walk is also importance-sampled. However, it has two compromises compared to the normal "forward" walk:

- The "guiding" wave function in the back-traversal of $|\phi_{\mathrm{j}}^{(\theta)}\rangle$ is $|\phi_{\mathrm{j}}^{(\tau)}\rangle$. This is clearly less optimal than the trial wave function $|\phi_{\mathrm{T}}\rangle$ because $|\phi_{\mathrm{T}}\rangle$ is our best guess for the ground state (and usually obtained with some preliminary calculations) while $|\phi_{\mathrm{j}}^{(\tau)}\rangle$ is just an arbitrary walker. Nevertheless, because the statistical distribution of $|\phi_{\mathrm{T}}\rangle$ represents the ground-state wave function, the chance that $|\phi_{\mathrm{T}}\rangle$ is radically different from the ground state wave function is also low. This also implies that each walker in the backward walk has a different guiding wave function. However, using $|\phi_{\mathrm{j}}^{(\tau)}\rangle$ as the guiding wave function is not wrong because in the Monte Carlo representation of a wave function in Eq. (3.55), the guiding wave function will cancel out.
- The sampling of auxiliary fields may not be optimal. What this means is that a configuration $\vec{x}_{\mathrm{j}}$ that increases the overlap with the guiding wave function $|\phi_{\mathrm{T}}\rangle$ on the forward walk might not lead to a larger overlap with the guiding wave function $|\phi_{\mathrm{j}}^{(\tau)}\rangle$ on the backward walk.

According to Purwanto and Zhang [29, p. 056702-8], in practice, these two compromises have not been a major problem.

As in the forward walk, the initial population are all identical to the trial wave function i.e. for all j:

$$|\eta_{\mathrm{j}}^{(\theta=0)}\rangle = |\phi_{\mathrm{T}}\rangle \tag{I.15}$$

$$u_{\mathrm{j}}^{(\theta=0)} = \langle\phi_{\mathrm{j}}^{(\tau)}|\phi_{\mathrm{T}}\rangle \ . \tag{I.16}$$

Note the interesting role reversal of the walker $|\phi_{\mathrm{j}}^{(\tau)}\rangle$ and and the trial wave function $|\phi_{\mathrm{T}}\rangle$ in the weight compared to Eq. (3.54). For concreteness, we'll show the first few steps of the backward walk for a walker. Note that τ is fixed and θ is changing.

Step 1 takes the walker from $\theta = 0$ to $\theta = \Delta\tau$ and uses the last time step of the forward path in Eq. (I.7)

$$|\eta_{\mathrm{j}}^{(\Delta\tau)}\rangle = \left[\widehat{B}_{\mathrm{j}}^{(\tau'-\Delta\tau)}\right]^{\dagger} |\phi_{\mathrm{T}}\rangle \tag{I.17}$$

$$u_{\mathrm{j}}^{(\Delta\tau)} = \frac{\langle\phi_{\mathrm{j}}^{(\tau)}| \left[\widehat{B}_{\mathrm{j}}^{(\tau'-\Delta\tau)}\right]^{\dagger} |\phi_{\mathrm{T}}\rangle}{\langle\phi_{\mathrm{j}}^{(\tau)}|\phi_{\mathrm{T}}\rangle} \left[\mathcal{F}^{(\tau'-\Delta\tau)}\right]^{*} \langle\phi_{\mathrm{j}}^{(\tau)}|\phi_{\mathrm{T}}\rangle \tag{I.18}$$

where * denotes complex conjugation and † Hermitian conjugation. Recall that propagation by $\left[\widehat{B}_{\mathrm{j}}(\tau' - \Delta\tau)\right]^{\dagger}$ does not require sampling of auxiliary fields because we are simply reusing the already-sampled auxiliary field configuration $\vec{x}_{\mathrm{j}}^{(\tau'-\Delta\tau)}$at time $\tau' - \Delta\tau$. This saves a lot of computational power because calculating the overlaps in the sampling process requires evaluating the ratio of determinants of matrices for each lattice site.

Step 2 takes the walker from $\theta = \Delta\tau$ to $\theta = 2\Delta\tau$ and uses the second-to-last time step of the forward path in Eq. (I.7)

$$|\eta_{\mathrm{j}}^{(2\Delta\tau)}\rangle = \left[\widehat{B}_{\mathrm{j}}^{(\tau'-2\Delta\tau)}\right]^{\dagger} |\eta_{\mathrm{j}}^{(\Delta\tau)}\rangle \tag{I.19}$$

$$= \left[\widehat{B}_{\mathrm{j}}^{(\tau'-2\Delta\tau)}\right]^{\dagger} \left[\widehat{B}_{\mathrm{j}}^{(\tau'-\Delta\tau)}\right]^{\dagger} |\phi_{\mathrm{T}}\rangle \ , \tag{I.20}$$

or with the notation of Eq. (I.7):

$$|\eta_{\mathrm{j}}^{(2\Delta\tau)}\rangle = \left[\widehat{B}_{\mathrm{j}}^{(\tau':\tau'-2\Delta\tau)}\right]^{\dagger} |\phi_{\mathrm{T}}\rangle \ . \tag{I.21}$$

Similarly, it's straightforward to show that the weight evolves as

$$u_{\mathrm{j}}^{(2\Delta\tau)} = \frac{\langle\phi_{\mathrm{j}}^{(\tau)}| \left[\widehat{B}_{\mathrm{j}}^{(\tau':\tau'-2\Delta\tau)}\right]^{\dagger} |\phi_{\mathrm{T}}\rangle}{\langle\phi_{\mathrm{j}}^{(\tau)}|\phi_{\mathrm{T}}\rangle} \left[\mathcal{F}^{(\tau':\tau'-2\Delta\tau)}\right]^{*} \langle\phi_{\mathrm{j}}^{(\tau)}|\phi_{\mathrm{T}}\rangle \tag{I.22}$$

If this propagation continues all the way to $\theta = \tau_{\mathrm{bp}}$, the walker becomes

$$|\eta_{\mathrm{j}}^{(\tau_{\mathrm{bp}})}\rangle = \left[\widehat{B}_{\mathrm{j}}^{(\tau':\tau)}\right]^{\dagger} |\phi_{\mathrm{T}}\rangle \ . \tag{I.23}$$

(Recall that $\tau = \tau' - \tau_{\mathrm{bp}}$). The weight becomes

$$u_{\mathrm{j}}^{(\tau_{\mathrm{bp}})} = \frac{\langle\phi_{\mathrm{j}}^{(\tau)}| \left[\widehat{B}_{\mathrm{j}}^{(\tau':\tau'-\tau_{\mathrm{bp}})}\right]^{\dagger} |\phi_{\mathrm{T}}\rangle}{\langle\phi_{\mathrm{j}}^{(\tau)}|\phi_{\mathrm{T}}\rangle} \left[\mathcal{F}^{(\tau':\tau'-\tau_{\mathrm{bp}})}\right]^{*} . \langle\phi_{\mathrm{j}}^{(\tau)}|\phi_{\mathrm{T}}\rangle \ . \tag{I.24}$$

Collect the RHS together, taking complex conjugates as necessary:

$$u_{\mathrm{j}}^{(\tau_{\mathrm{bp}})} = \left[\frac{\langle\phi_{\mathrm{T}}|\widehat{B}_{\mathrm{j}}^{(\tau':\tau'-\tau_{\mathrm{bp}})}|\phi_{\mathrm{j}}^{(\tau)}\rangle \, \mathcal{F}^{(\tau':\tau'-\tau_{\mathrm{bp}})}}{\langle\phi_{\mathrm{T}}|\phi_{\mathrm{j}}^{(\tau)}\rangle} \langle\phi_{\mathrm{T}}|\phi_{\mathrm{j}}^{(\tau)}\rangle \right]^{*} . \tag{I.25}$$

Multiply and divide by $\langle\phi_{\mathrm{T}}|\phi_{\mathrm{j}}^{(\tau)}\rangle$, recall that $\tau = \tau' - \tau_{\mathrm{bp}}$ and rearrange:

$$u_{\mathrm{j}}^{(\tau_{\mathrm{bp}})} = \Bigg[\underbrace{\frac{\langle\phi_{\mathrm{T}}|\widehat{B}_{\mathrm{j}}^{(\tau':\tau)}|\phi_{\mathrm{j}}^{(\tau)}\rangle \, \mathcal{F}^{(\tau':\tau)}}{\langle\phi_{\mathrm{T}}|\phi_{\mathrm{j}}^{(\tau)}\rangle} \langle\phi_{\mathrm{T}}|\phi_{\mathrm{j}}^{(\tau)}\rangle}_{w_{\mathrm{j}}^{(\tau')}} \frac{\langle\phi_{\mathrm{T}}|\phi_{\mathrm{j}}^{(\tau)}\rangle}{\langle\phi_{\mathrm{T}}|\phi_{\mathrm{j}}^{(\tau)}\rangle} \Bigg]^{*} . \tag{I.26}$$

We have used Eq. (I.12) to identify the weight $w_{\mathrm{j}}^{(\tau')}$ of a forward-walk random walker at time τ'. We also recognize $w_{\mathrm{j}}^{(\tau)} = \langle\phi_{\mathrm{T}}|\phi_{\mathrm{j}}^{(\tau)}\rangle$ in the denominator and leave $\langle\phi_{\mathrm{T}}|\phi_{\mathrm{j}}^{(\tau)}\rangle$ in the numerator alone. Thus

$$u_{\mathrm{j}}^{(\tau_{\mathrm{bp}})} = \left[\frac{w_{\mathrm{j}}^{(\tau')}}{w_{\mathrm{j}}^{(\tau)}} \langle\phi_{\mathrm{T}}|\phi_{\mathrm{j}}^{(\tau)}\rangle\right]^{*} \ . \tag{I.27}$$

A good thing about Eq. (I.27) is that it does not involve any weights at intermediate times θ.

Together, Eqs. (I.23) and (I.27) are what we need to get a formula for an observable $\widehat{A}$ in back-propagation. Because both the forward and backward walk have importance sampling, both the left- and right-hand wave functions have the form in Eq. (3.55). The left-hand wave function is

$$\sum_{\mathrm{j}} w_{\mathrm{j}}^{(\tau)} \frac{|\phi_{\mathrm{j}}^{(\tau)}\rangle}{\langle\phi_{\mathrm{T}}|\phi_{\mathrm{j}}^{(\tau)}\rangle} \tag{I.28}$$

and the right-hand wave function is

$$\sum_{\mathrm{j}} u_{\mathrm{j}}^{(\tau_{\mathrm{bp}})} \frac{|\eta_{\mathrm{j}}^{(\tau_{\mathrm{bp}})}\rangle}{\langle\phi_{\mathrm{j}}^{(\tau)}|\eta_{\mathrm{j}}^{(\tau_{\mathrm{bp}})}\rangle} \,. \tag{I.29}$$

Now we need to match each walker in the right-hand wave function with its counterpart in the left-hand walk in Eq. (I.2).

The denominator of Eq. (I.2) is

$$\begin{aligned}
\text{denominator} &= \langle\Psi_{\mathrm{BP}}|\Psi^{(\tau)}\rangle && \text{(I.30)}\\
&= \sum_{\mathrm{j}} \left[u_{\mathrm{j}}^{(\tau_{\mathrm{bp}})}\right]^* w_{\mathrm{j}}^{(\tau)} \frac{\langle\eta_{\mathrm{j}}^{(\tau_{\mathrm{bp}})}|\phi_{\mathrm{j}}^{(\tau)}\rangle}{\left[\langle\phi_{\mathrm{j}}^{(\tau)}|\eta_{\mathrm{j}}^{(\tau_{\mathrm{bp}})}\rangle\right]^* \langle\phi_{\mathrm{T}}|\phi_{\mathrm{j}}^{(\tau)}\rangle} && \text{(I.31)}\\
&= \sum_{\mathrm{j}} \left[\frac{w_{\mathrm{j}}^{(\tau')}}{\cancel{w_{\mathrm{j}}^{(\tau)}}\cancel{\langle\phi_{\mathrm{T}}|\phi_{\mathrm{j}}^{(\tau)}\rangle}}\right] \cancel{w_{\mathrm{j}}^{(\tau)}} \frac{\cancel{\langle\eta_{\mathrm{j}}^{(\tau_{\mathrm{bp}})}|\phi_{\mathrm{j}}^{(\tau)}\rangle}}{\cancel{\langle\eta_{\mathrm{j}}^{(\tau_{\mathrm{bp}})}|\phi_{\mathrm{j}}^{(\tau)}\rangle}\cancel{\langle\phi_{\mathrm{T}}|\phi_{\mathrm{j}}^{(\tau)}\rangle}} \,. && \text{(I.32)}
\end{aligned}$$

After cancellations, we are left with simply

$$\text{denominator} = \sum_{\mathrm{j}} w_{\mathrm{j}}^{(\tau')} \,. \tag{I.33}$$

The numerator of Eq. (I.2) is evaluated in a similar way:

$$\begin{aligned}
\text{numerator} &= \langle\Psi_{\mathrm{BP}}|\widehat{A}|\Psi^{(\tau)}\rangle && \text{(I.34)}\\
&= \sum_{\mathrm{j}} \left[u_{\mathrm{j}}^{(\tau_{\mathrm{bp}})}\right]^* w_{\mathrm{j}}^{(\tau)} \frac{\langle\eta_{\mathrm{j}}^{(\tau_{\mathrm{bp}})}|\widehat{A}|\phi_{\mathrm{j}}^{(\tau)}\rangle}{\left[\langle\phi_{\mathrm{j}}^{(\tau)}|\eta_{\mathrm{j}}^{(\tau_{\mathrm{bp}})}\rangle\right]^* \langle\phi_{\mathrm{T}}|\phi_{\mathrm{j}}^{(\tau)}\rangle} \,. && \text{(I.35)}
\end{aligned}$$

After cancellations, we have

$$\text{numerator} = \sum_{\mathrm{j}} w_{\mathrm{j}}^{(\tau')} \frac{\langle\eta_{\mathrm{j}}^{(\tau_{\mathrm{bp}})}|\widehat{A}|\phi_{\mathrm{j}}^{(\tau)}\rangle}{\langle\eta_{\mathrm{j}}^{(\tau_{\mathrm{bp}})}|\phi_{\mathrm{j}}^{(\tau)}\rangle} \tag{I.36}$$

Putting Eqs. (I.33) and (I.36) together, we have the formula for the back-propagated estimator

$$\langle\widehat{A}\rangle_{\mathrm{bp}} = \frac{\sum_{\mathrm{j}} w_{\mathrm{j}}^{(\tau')} \frac{\langle\eta_{\mathrm{j}}^{(\tau_{\mathrm{bp}})}|\widehat{A}|\phi_{\mathrm{j}}^{(\tau)}\rangle}{\langle\eta_{\mathrm{j}}^{(\tau_{\mathrm{bp}})}|\phi_{\mathrm{j}}^{(\tau)}\rangle}}{\sum_{\mathrm{j}} w_{\mathrm{j}}^{(\tau')}} \,. \tag{I.37}$$

To summarize, in this appendix, we have derived a procedure to calculate the expectation of a physical observable using b1ack-propagation.

Appendix J

Stabilization of matrix multiplication

In this appendix, we discuss how a loss of numerical accuracy can occur in a "naive" CPMC implementation and show a procedure to counter this loss of accuracy. We have adapted materials from Assaad [42, pp.120-1,127] and Zhang [46, pp. 16-7].

In Appendix F, we showed that the one-particle equal-time Green's function between two Slater determinants $|P\rangle$ and $|Q\rangle$ is

$$G_{ij} = \langle c_i c_j^\dagger \rangle = \frac{\langle P|c_i c_j^\dagger|Q\rangle}{\langle P|Q\rangle} \tag{J.1}$$

$$= I - \left[Q(P^\dagger Q)^{-1} P^\dagger\right] . \tag{J.2}$$

In a CPMC algorithm, we want to calculate $\langle c_i c_j^\dagger \rangle$ for the ground state wave function i.e.

$$\frac{\langle \psi_0|c_i c_j^\dagger|\psi_0\rangle}{\langle \psi_0|\psi_0\rangle} \approx \frac{\langle \psi_T|e^{-\tau_L \hat{H}} c_i c_j^\dagger e^{-\tau_R \hat{H}}|\psi_T\rangle}{\langle \psi_0|e^{-\tau_L \hat{H}} e^{-\tau_R \hat{H}}|\psi_0\rangle} \tag{J.3}$$

so we identify $|P\rangle = e^{-\tau_L \hat{H}}|\psi_T\rangle$ and $|Q\rangle = e^{-\tau_R \hat{H}}|\psi_T\rangle$ in Eq. (J.2). On finite-precision machines the repeated propagation of $|\psi_T\rangle$ by $e^{-\tau_L \hat{H}}$ will cause numerical instabilities such that round-off errors dominate such that $|\phi_j^{(\tau)}\rangle$ represents an unfaithful propagation of $|\phi_j^{(0)}\rangle$. This instability is directly related to the collapse to a bosonic ground state. That is, if we let the propagation continue without any stabilization procedure, all single-particle orbitals in each walker would become the same.

To alleviate this problem, we periodically perform a modified Gram-Schmidt reorthonormalization on the walkers. At some appropriate step in the random walk when reorthonormalization is to be carried out, we decompose the matrix P of a walker into $P = U_1 D_1 V_1$ where U_1 is an orthonormal matrix, D_1 is a diagonal matrix and V_1 is an upper triangular matrix. Similarly, we decompose Q as $Q = U_2 D_2 V_2$. It turns out that the one-particle Green's

function is invariant even if we discard the D's and V's matrices because

$$Q(P^\dagger Q)^{-1}P^\dagger = (U_2 D_2 V_2)\left[(U_1 D_1 V_1)^\dagger (U_2 D_2 V_2)\right]^{-1}(U_1 D_1 V_1)^\dagger \tag{J.4}$$

$$= (U_2 D_2 V_2)\left[V_1^\dagger D_1^\dagger U_1^\dagger U_2 D_2 V_2\right]^{-1}(V_1^\dagger D_1^\dagger U_1^\dagger) \tag{J.5}$$

$$= U_2 (D_2 V_2)\left[(V_1^\dagger D_1^\dagger)(U_1^\dagger U_2)(D_2 V_2)\right]^{-1}(V_1^\dagger D_1^\dagger) U_1^\dagger \tag{J.6}$$

$$= U_2 \cancel{(D_2 V_2)}\cancel{(D_2 V_2)^{-1}}\,(U_1^\dagger U_2)^{-1}\,\cancel{(V_1^\dagger D_1^\dagger)^{-1}}\cancel{(V_1^\dagger D_1^\dagger)}\,U_1^\dagger \tag{J.7}$$

$$Q(P^\dagger Q)^{-1}P^\dagger = U_2 [U_1^\dagger U_2]^{-1} U_1^\dagger . \tag{J.8}$$

Thus the numerical instabilities disappear from the problem of computing the equal-time Green's function. Since all physical observables relies on the equal-time Green's function, the above stabilization procedure leaves the physical results invariant.

Appendix K

Two-body operators

This appendix shows that any two-body operator can be expressed as a sum of squares of one-body operators. It is not essential to the implementation of this thesis but shows that the Hubbard-Stratonovich transformation can be used to handle any two-body interactions. It expands on the materials in Zhang [47, p.10] and Motta *et al.* [24, p. 9].

The general form of an operator in second-quantization is

$$\widehat{H} = \sum_{ij} A_{ij} c_i^\dagger c_j + \sum_{ijk\ell} V_{ijk\ell}\, c_i^\dagger c_j^\dagger c_\ell c_k \,, \tag{K.1}$$

where the second term is the most general form of a two-body operator:

$$\widehat{V} = \sum_{ijk\ell} V_{ijk\ell}\, c_i^\dagger c_j^\dagger c_\ell c_k \,. \tag{K.2}$$

We first introduce two new indices $\alpha = (i, \ell)$ and $\beta = (k, j)$ and letting $\mathcal{V}_{\alpha\beta} = \mathcal{V}_{(i,\ell),(k,j)} = V_{ijk\ell}$. Note that this is simply a different scheme of labeling the indices. If there are M single-particle basis states (labeled by i,ℓ,j and k), there are M^2 possible values of α (and M^2 values of β). Note that each of the indices i,ℓ,j and k ranges over both space and spin coordinates e.g. on an N-site lattice with two possible spins on each lattice site, $M = 2N$ hence $M^2 = 4N^2$.

We first prove that the matrix $\mathcal{V}$ is Hermitian, that is

$$\mathcal{V}^*_{\alpha\beta} = \mathcal{V}_{\beta\alpha} \,. \tag{K.3}$$

From the way α and β are defined, this is equivalent to showing that

$$V^*_{ijk\ell} = V_{k\ell ij} \,. \tag{K.4}$$

Note that in this notation when converting from $\mathcal{V}$ to V, the first of the two subscripts of $\mathcal{V}$ forms the two "outer" subscripts of V (in the correct order) while the second of the two subscripts of $\mathcal{V}$ forms the inner subscripts of V (in reversed order). Thus $\mathcal{V}_{\beta\alpha}$ becomes $V_{k\ell ij}$.

Eq. (K.4) follows immediately from the Hermiticity of $\widehat{V} = \widehat{H} - \widehat{K}$ (which is in turn Hermitian because $\widehat{H}$ and $\widehat{K}$ are both Hermitian). Since $\widehat{V}$ is Hermitian, we have

$$\widehat{V}^\dagger = \widehat{V} \,. \tag{K.5}$$

The LHS of Eq. (K.5) is

$$\widehat{V}^\dagger = \left(\sum_{ijk\ell} V_{ijk\ell} c_i^\dagger c_j^\dagger c_\ell c_k\right)^\dagger \tag{K.6}$$

$$= \sum_{ijk\ell} V_{ijk\ell}^* c_k^\dagger c_\ell^\dagger c_j c_i \qquad V_{ijk\ell} \text{ is a scalar} \tag{K.7}$$

while the RHS is

$$\widehat{V} = \sum_{ijk\ell} V_{ijk\ell} c_i^\dagger c_j^\dagger c_\ell c_k \tag{K.8}$$

$$= \sum_{ijk\ell} V_{k\ell ij} c_k^\dagger c_\ell^\dagger c_j c_i \quad \text{rename } i \to k, j \to \ell, k \to i, \ell \to j \ . \tag{K.9}$$

By "comparing coefficients" of Eqs. (K.7) and (K.9), we have proved Eq. (K.4), hence $\mathcal{V}$ is Hermitian. Thus we can diagonalize $\mathcal{V}$ as

$$\mathcal{V}_{\alpha\beta} = R^\dagger \Lambda R \tag{K.10}$$

$$= \sum_{\gamma=1}^{M^2} (R^\dagger)_{\alpha\gamma} \lambda_\gamma R_{\gamma\beta} \tag{K.11}$$

where Λ is the diagonal matrix containing the eigenvalues λ_γ. As a side note, because $\mathcal{V}$ is Hermitian, the λ_γ's are real.

$\widehat{V}$ can then be rewritten

$$\widehat{V} = \sum_{ijkl} V_{ijk\ell} c_i^\dagger c_j^\dagger c_\ell c_k \tag{K.12}$$

$$= \sum_{ijkl} V_{ijk\ell} c_i^\dagger (\delta_{j\ell} - c_\ell c_j^\dagger) c_k \qquad \{c_j^\dagger, c_\ell\} = \delta_{j\ell} \tag{K.13}$$

$$= -\sum_{ijkl} V_{ijk\ell} c_i^\dagger c_\ell c_j^\dagger c_k + \sum_{ijk\ell} V_{ijk\ell} c_i^\dagger c_k \delta_{j\ell} \tag{K.14}$$

$$= -\sum_{(i,\ell)} \sum_{(k,j)} (c_i^\dagger c_\ell)(c_j^\dagger c_k) V_{ijk\ell} + \sum_{ijk} V_{ijkj} c_i^\dagger c_k \, . \tag{K.15}$$

Let $p_\alpha = c_i^\dagger c_\ell$, $q_\beta = c_j^\dagger c_k$, $\alpha = (i,\ell)$ and $\beta = (k,j)$, we have

$$\widehat{V} = -\sum_{\alpha=1}^{M^2} \sum_{\beta=1}^{M^2} p_\alpha q_\beta \mathcal{V}_{\alpha\beta} + \sum_{ik} \left(\sum_j V_{ijkj}\right) c_i^\dagger c_k \tag{K.16}$$

$$= -\sum_{\alpha,\beta=1}^{M^2} p_\alpha q_\beta \left(\sum_{\gamma=1}^{M^2} (R^\dagger)_{\alpha\gamma} \lambda_\gamma R_{\gamma\beta}\right) + \sum_{ik} \left(\sum_j V_{ijkj}\right) c_i^\dagger c_k \qquad \text{using Eq. (K.11)} \, . \tag{K.17}$$

Note that p_α and q_β are Hermitian conjugates of each other[1]

$$\widehat{V} = -\sum_{\gamma}^{M^2} \left(\sum_{\alpha}^{M^2} q_\alpha^\dagger (R^\dagger)_{\alpha\gamma}\right) \lambda \left(\sum_{\beta}^{M^2} R_{\gamma\beta} q_\beta\right) + \sum_{ik} \left(\sum_j V_{ijkj}\right) c_i^\dagger c_k \, . \tag{K.18}$$

Define

$$\widehat{\rho}_\gamma = \sum_\beta^{M^2} R_{\gamma\beta} q_\beta = \sum_{kj} R_{\gamma(k,j)} c_j^\dagger c_k \tag{K.19}$$

and

$$\widehat{\rho}_0 = \sum_{ik} \left(\sum_j V_{ijkj} \right) c_i^\dagger c_k \,. \tag{K.20}$$

Note that both $\widehat{\rho}_\gamma$ and $\widehat{\rho}_0$ are one-body operators. Then we can rewrite Eq. (K.18):

$$\widehat{V} = -\frac{1}{2} \sum_\gamma^{M^2} \lambda_\gamma \{\widehat{\rho}_\gamma, \widehat{\rho}_\gamma^\dagger\} + \widehat{\rho}_0 \tag{K.21}$$

where $\{,\}$ denotes the anticommutator. Since

$$\{\widehat{\rho}_\gamma, \widehat{\rho}_\gamma^\dagger\} = \frac{1}{2} \left[\left(\widehat{\rho}_\gamma + \widehat{\rho}_\gamma^\dagger \right)^2 - \left(\widehat{\rho}_\gamma - \widehat{\rho}_\gamma^\dagger \right)^2 \right] , \tag{K.22}$$

we have succeeded in writing the two-body operator $\widehat{V}$ as a sum of squares of one-body operators. The operator $\widehat{\rho}_0$ is a linear one-body operator and can be combined into the one-body part of the Hamiltonian.

Thus the general second-quantized operator in Eq. (K.1) now can be expressed entirely in terms of one-body operators

$$\widehat{H} = \widehat{H}_0 - \frac{1}{4} \sum_{\gamma=1}^{M^2} \lambda_\gamma \left[\left(\widehat{\rho}_\gamma + \widehat{\rho}_\gamma^\dagger \right)^2 - \left(\widehat{\rho}_\gamma - \widehat{\rho}_\gamma^\dagger \right)^2 \right] \tag{K.23}$$

[1] From the definition of p_α and q_β in Eq. (K.15) and the resulting expression in Eq. (K.16), we are essentially making the association

$$\sum_\alpha p_\alpha = \sum_{(i,\ell)} c_i^\dagger c_\ell$$

and

$$\sum_\beta q_\beta = \sum_{(k,j)} c_j^\dagger c_k$$

Taking the Hermitian conjugate of this equation gives

$$\begin{aligned} \left(\sum_\beta q_\beta \right)^\dagger &= \left(\sum_{(k,j)} c_j^\dagger c_k \right)^\dagger \\ &= \sum_{(k,j)} c_k^\dagger c_j \\ &= \sum_{(i,\ell)} c_i^\dagger c_\ell \qquad \text{rename } i \to k,\ \ell \to j \\ &= \sum_\alpha p_\alpha \end{aligned}$$

This is what we meant by saying q_β and p_α are Hermitian conjugates. This also means that the α and β indices are the same.

where the λ_γ's are real and $\widehat{H}_0$ is a (linear) one-body operator:

$$\widehat{H}_0 = \sum_{ik} \left[A_{ik} + \sum_j V_{ijkj} \right] c_i^\dagger c_j \,, \tag{K.24}$$

which is easy to propagate and requires no Monte Carlo sampling. The sum in Eq. (K.23) contains the squares of one-body operators $\widehat{\rho}_\gamma$:

$$\widehat{\rho}_\gamma = \sum_{kj} R_{\gamma(k,j)} c_j^\dagger c_k \,. \tag{K.25}$$

From Appendix A, we know that every square of an operator produces an auxiliary field under the Hubbard-Stratonovich transformation. Since there are $2M^2$ squares of one-body operators in Eq. (K.23) (two for each term of the γ-summation), the worst case scenario (i.e. a completely generic two-body operator) will require sampling $2M^2$ auxiliary fields per random walk step. As noted on page 109, for a N-site lattice with spin-1/2 particles, $M = 2N$ i.e. the worst case scenario requires sampling $8N^2$ auxiliary fields per random walk step.

However, this number is often much smaller in practice. As we have seen in the Hirsch spin transformation in Eq. (A.15), we only need to sample N auxiliary fields per random walk step.

Bibliography

[1] M. Bajdich and L. Mitas, Acta Phys. Slovaca **59**, 81 (2009).

[2] S. A. Jafari, Iranian J. Phys. Research **8**, 113 (2008).

[3] A. Szabo and N. S. Ostlund, *Modern Quantum Chemistry: Introduction to Advanced Electronic Structure Theory* (Dover Publications, Mineola, New York, 1989).

[4] G. Kotliar and D. Vollhardt, Phys. Today **57**, 53 (2004).

[5] S. Zhang, in *Emergent Phenomena in Correlated Matter*, Vol. 3, edited by E. Pavarini, E. Koch, and U. Schollwöck, Lecture Notes of the Autumn School Correlated Electrons (2013) Chap. 15.

[6] S. Zhang, J. Carlson, and J. E. Gubernatis, Phys. Rev. B **55**, 7464 (1997).

[7] J. Hubbard, Proc. R. Soc. A **276**, 238 (1963).

[8] F. Essler, H. Frahm, F. Gohmann, A. Klumper, and V. Korepin, *The One-Dimensional Hubbard Model* (Cambridge University Press, 2005).

[9] E. H. Lieb and F. Y. Wu, Phys. Rev. Lett. **20**, 1445 (1968).

[10] The Society for Industrial and Applied Mathematics, *The Best of the 20th Century: Editors Name Top 10 Algorithms*, `https://www.siam.org/pdf/news/637.pdf`, May 2000.

[11] N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, and E. Teller, J. Chem. Phys **21**, 1087 (1953).

[12] W. L. McMillan, Phys. Rev. A **138**, 442 (1965).

[13] J. B. Anderson, The Journal of Chemical Physics **63**, 1499 (1975).

[14] R. Blankenbecler, D. J. Scalapino, and R. L. Sugar, Phys. Rev. D **24**, 2278 (1981).

[15] D. J. Scalapino and R. L. Sugar, Phys. Rev. B **24**, 4295 (1981).

[16] G. Sugiyama and S. Koonin, Ann. Phys. **168**, 1 (1986).

[17] J. Hirsch, Phys. Rev. B **28**, 4059 (1983).

[18] J. Hirsch, Phys. Rev. B **31**, 4403 (1985).

[19] S. Fahy and D. Hamann, Phys. Rev. B **43**, 765 (1991).

[20] S. B. Fahy and D. R. Hamann, Phys. Rev. Lett. **65**, 3437 (1990).

[21] S. Zhang, Phys. Rev. Lett. **83**, 2777 (1999).

[22] S. Zhang and H. Krakauer, Phys. Rev. Lett. **90**, 136401 (2003).

[23] M. Feldbacher and F. F. Assaad, Phys. Rev. B **63**, 073105 (2001).

[24] M. Motta, D. E. Galli, S. Moroni, and E. Vitali, J. Chem. Phys **140**, 024107 (2014).

[25] H. Nguyen, H. Shi, J. Xu, and S. Zhang, Comput. Phys. Commun. **185**, 3344 (2014).

[26] M. H. Kalos and P. A. Whitlock, *Monte Carlo Methods*, 2nd Edition (Wiley-VCH, 2008).

[27] J. W. Negele and H. Orland, *Quantum Many-Particle Systems*, 1st Edition, Advanced Book Classics (Westview Press, 1998).

[28] Ashcroft and Mermin, *Solid State Physics* (Holt, Rinehart and Winston, New York, 1976).

[29] W. Purwanto and S. Zhang, Phys. Rev. E **70**, 056702 (2004).

[30] D. J. Thouless, Nucl. Phys. **21**, 225 (1960).

[31] D. R. Hamann and S. B. Fahy, Phys. Rev. B **41**, 11352 (1990).

[32] I. Kosztin, B. Faber, and K. Schulten, Am. J. Phys. **64**, 633 (1996).

[33] W. von der Linden, Phys. Rep. **220**, 53 (1992).

[34] Wikipedia, *Sherman–Morrison formula — Wikipedia, The Free Encyclopedia*, [Online; accessed 29-March-2014], 2013.

[35] S. Zhang and M. H. Kalos, Phys. Rev. Lett. **67**, 3074 (1991).

[36] R. R. dos Santos, Braz. J. Phys. **33**, 36 (2003).

[37] M. Calandra Buonaura and S. Sorella, Phys. Rev. B **57**, 11446 (1998).

[38] M. Jarrell and J. Gubernatis, Physics Reports **269**, 133 (1996).

[39] E. Vitali, M. Rossi, L. Reatto, and D. E. Galli, Phys. Rev. B **82**, 174510 (2010).

[40] S. H. Simon, *The Oxford Solid State Basics* (Oxford University Press, 2013).

[41] D. J. Griffiths, *Introduction to Quantum Mechanics* (Pearson Prentice Hall, 2004).

[42] F. F. Assaad, in *Quantum Simulations of Complex Many-Body Systems: From Theory to Algorithms*, Vol. 10 NIC Series, edited by J. Grotendorst, D. Marx, and A. Muramatsu, NIC Series (John von Neumann Institute for Computing, Julich, 2002).

[43] S. Rombouts, K. Heyde, and N. Jachowicz, Phys. Lett. A **242**, 271 (1998).

[44] N. J. Higham, *Function of Matrices: Theory and Computation* (SIAM, Philadelphia, 2008).

[45] W. Purwanto, "Quantum Monte Carlo Method for Boson Ground States: Application to Trapped Bosons with Attractive and Repulsive Interactions," PhD Thesis (The College of William and Mary, 2005).

[46] S. Zhang, in *Quantum Monte Carlo Methods in Physics and Chemistry*, Vol. 525, edited by M. Nightingale and C. Umrigar, NATO Science Series C: Mathematical and Physical Sciences (Kluwer Academic, Dordrecht, 1999), p. 399.

[47] S. Zhang, in *Theoretical Methods for Strongly Correlated Electrons*, edited by D. Senechal, A.-M. Tremblay, and C. Bourbonnais, CRM Series in Mathematical Physics (Springer, 2004) Chap. 2, p. 39.