Kvantni algoritmi

Delić, Karlo

SVEUČILIŠTE U ZAGREBU

PRIRODOSLOVNO–MATEMATIČKI FAKULTET

MATEMATIČKI ODSJEK

Karlo Delić

QUANTUM ALGORITHMS

Diplomski rad

Voditelj rada:
Matija Kazalicki

Zagreb, veljača, 2024.

Ovaj diplomski rad obranjen je dana _____ pred ispitnim povjerenstvom u sastavu:

1. _____, predsjednik
2. _____, član
3. _____, član

Povjerenstvo je rad ocijenilo ocjenom _____.

Potpisi članova povjerenstva:

1. _____
2. _____
3. _____

Я хотел бы выразить самую глубокую благодарность Татьяне А. Беспаловой за возможность погружения в мир исследований квантовых вычислений и постоянное техническое руководство в течение этого пути.

Great thanks to Guido Pupillo, Francesco Tacchino and Ivano Tavernelli for suggesting the research problem to tackle in this and subsequent works, and to Marie Curie project MoQS for bringing us all together. I am further thankful to my advisor Matija Kazalicki for his patience and illuminating discussions on quantum computing.

Majci Suzani, ocu Denisu i bratu Lovri, za sve ove godine.

Contents

Contents **iii**

0 Catalogue **1**
0.1 Note on notation 1

1 Introduction and Motivation **3**
1.1 Thesis outline 5
1.2 Intermezzo (for the less physically inclined) 6

2 Fermi-Hubbard Model **21**
2.1 Symmetries of the model 23
2.2 One spin theorem and one pseudospin theorem 25
2.3 half filling zero spin projection sector basis 26
2.4 Exactly solvable limiting cases 28
2.5 General analytic approach 37

3 VQE **46**
3.1 NISQ regime 47
3.2 Variational Quantum Eigensolver 48
3.3 Model implementation 55
3.4 Building the ansatz 55
3.5 Ansatz characterization 61

4 VQE Results **64**
4.1 General considerations 64
4.2 Depth analysis 65
4.3 Limiting cases 66
4.4 Full set of parameters 67
4.5 Noisy calculations 71
4.6 Conclusion and future prospects 73

Bibliography **75**

0

Catalogue

Let us begin with a short prelude including notational choices, conventions and some general remarks. This part should be skimmed on a first reading and then used as a reference while reading the rest.

The complete codebase for the thesis and an additional paper [1] is written in Python, relying on a popular quantum computing package Qiskit. All of the scripts referenced in the main text are contained in a public repository

`https://github.com/dels1997/vqe`

and are available for testing, calculations and visualizations.

For those with very little physics knowledge, we provide an entire Section 1.2 which includes an overview of the most pertinent concepts, along with simple examples.

0.1 Note on notation

The notational choices are usually the most common ones, however due to the thesis covering a variety of disciplines, it is prudent to include them in a single place for readability and disambiguation later on. Thus, here are some important remarks to this end:

- We use j and l to index spin sites, q to index momenta and σ to index spin.
- Indices are placed on operators as follows: direction index (such as x, y or z) is a superscript, the site or momentum as a first index in subscript and spin as a second index in subscript.
- Vectors in braket notation are denoted as $|\psi\rangle$, and in explicit form as matrices with round brackets, such as
$$|\uparrow\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix} .$$

- Operators are denoted by upright bold symbols, and in explicit form as matrices with square brackets, e.g.
$$\mathbf{s}^z = \frac{1}{2}\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}.$$
Exception is the identity operator, denoted by the conventional blackboard unity symbol $\mathbb{1}$.

- Operators related to a particular site are denoted by noncapital letters, such as $\mathbf{s}^z$, while the global operators, i.e. pertaining to the entire system (lattice) are denoted by capital letters, such as $\mathbf{S}^z$.

- Vectors besides quantum statevectors are also denoted by upright bold symbols, for example $\mathbf{x}$ and $\boldsymbol{\theta}$.

- $\mathbf{x}$ is used to denote a general position vector.

- Products of states or spaces are usually omitted for convenience and readability, so that, e.g. the product state $|\psi_1\rangle \otimes |\psi_2\rangle$ may be written as $|\psi_1\psi_2\rangle$.

- Term "lattice" is used to refer to a physical system that includes multiple spin sites.

- Spin sites in the considered Fermi-Hubbard model (2.3), as well as the discrete Fourier transform (1.11) are 1-indexed.

- The number of spin sites in the model is denoted by N, resulting in a number of qubits $2N$.

- Statevector $|\widetilde{0}\rangle$ is used to denote a number occupation state in a 36×36 subspace 2.2, $|0\rangle$ to denote either a state of qubit/electron of spin down or an empty orbital in second quantization notation; and $|\mathbf{0}\rangle$ to denote a product of all $2N$ such states.

1

Introduction and Motivation

Nature isn't classical, dammit, and if you want to make a simulation of nature, you'd better make it quantum mechanical, and by golly it's a wonderful problem, because it doesn't look so easy.

– Richard P. Feynman

During the 20th century, both physics and computation theory underwent major breakthroughs and radical change. Physicists were starting to get worried about the classical frameworks such as electromagnetism predicting infinities and other unphysical results, most famously the ultraviolet catastrophe and the expected instability of the atoms [2]. After some time and several steps away from existing theories, such as Planck's solution to the UV issue with energy quantization [3] and Einstein's description of the photoelectric effect [4], that earned each of them a Nobel Prize, it became more and more clear that a paradigm shift is needed, and one came along in form of quantum physics by Heisenberg and Schrödinger in terms of matrix and wave mechanics, which laid the cornerstone of modern quantum physics [5, 6].

In scope of wave mechanics, the fundamental Schrödinger's equation provides a powerful and universal tool to predict the evolution of quantum systems, emphasizing the wave-like nature of particles and their dynamic interplay. In Heisenberg's matrix mechanics, physically observable quantities are represented by matrices, resulting in the groundbreaking uncertainty principle [7], revealing an inherent limit to the precision with which certain sets of properties of physical objects can be known simultaneously. Naturally, both would go on to be awarded the Nobel Prizes, Heisenberg in 1932 "for the creation of quantum mechanics" and Schrödinger a year later "for the discovery of new productive forms of atomic theory".

Impact of quantum mechanics is immeasurable in everyday life, and as one of its prominent consequences, we can highlight the development of the transistor technology using

materials exhibiting a quantum effect of semiconductivity—bunching transistor-like elements into (micro)chips would go on to become the foundation of almost all of the modern electronic devices and change the world forever. Expectedly, this resulted in another Nobel Prize awarded jointly to Bardeen, Brattain and Shockley [8, 9].

Compared to the world of physics, there were no significant issues with computation theory at the turn of the century and during the following period. A revolution unfolded nevertheless when Alan Turing wrote a paper in 1936 [10], almost single-handedly establishing the theory of computation. Particularly, due to formalization of the idea of the algorithm, he was able to demonstrate for the second time (after Church) that there exist problems that are not computable, even in principle, as well as introduce a computing machine that carries his name and would later take place at the top of the computing hierarchy. Ideas presented by Turing in this paper and during the subsequent years have had an enormous effect on modern computation and programming, ranging from the development of different computation models to the area of artificial intelligence. Besides this, his successful efforts to break the Enigma code during the 1940s likely helped the world come to a peaceful resolution of a global conflict sooner than seemed possible.

For all his efforts, and particularly for being such a formative figure in the field of modern computation theory, eponymous Turing Award was established in 1966 and is often regarded as the Nobel Prize of computer science.

Reflecting upon the course of the 20^{th} with the current knowledge of quantum physics at disposal and bearing in mind the two revolutions we have outlined, it seems peculiar that the first formal attempts at exploiting quantum phenomena for the purposes of simulation and computation arrived in the 1960s, and fruitful ones for that matter as late as 1980s. Of course, there is a plethora of reasons behind this, of which we can emphasize both the fact that the applications of quantum mechanics were not immediately apparent (and are still not even today), but also the unavailability of necessary technology, particularly in terms of computing power. Notwithstanding, in 1981 Feynman spoke of the idea of simulating natural physical systems using a sort of a quantum computer [11], encapsulated succinctly in the opening quotation of this chapter, and serious strides towards a formulation of the field of quantum computing were starting to be made. Among the first of the pioneering results, in 1985 Deutsch thought of a universal quantum Turing machine [12], analogous to the classical one conceptualized by Turing himself. During the rest of 1980s and the beginning of 1990s, several solutions of somewhat contrived problems with quantum computers followed, and demonstrated that this model of computation is capable of solving them in exponentially less time than its classical counterpart would be able to. However, the field starting gaining serious traction in 1994, when Shor constructed a polynomial quantum algorithm for integer factorization [13], an immensely important problem (still) believed to have no classical solution of polynomial complexity.

In the following decades, a significant amount of progress has been made, both in

terms of theoretical development and technological advance, and quantum computers with hundreds of qubits have been employed and ran, albeit under highly curated and exacting environmental conditions [14, 15, 16]. This is because the same phenomenon enabling the potential spectacular advantage of quantum computers—entanglement—is also preventing reaping the benefits of this potential. Namely, when entanglement occurs with the undesired degrees of freedom outside of the system, i.e. the environment, decoherence takes place and the inner entanglement that could have been exploited for useful computation is lost. Thus, one of the main challenges of quantum computing is to quickly build entanglement between the qubits before the errors or decoherence overly corrupt the quantum state and make it behave like a classical one, resulting in loss of its advantageous quantum properties. This issue is insurmountable at the moment and it prohibits the construction of robust and scalable quantum computers, especially on the scales needed to implement the famous algorithms such as Shor's and Grover's [17], as high as several thousands or tens of thousands of qubits.

For these reasons, a field—believed to be able to surpass the conventional computers in the near future—has emerged within quantum computing that deals with algorithms and devices appropriate for systems with intermediate numbers of qubits, usually in the hundreds—denoted the Noisy Intermediate Scale Quantum (NISQ) [18, 19]. Among the algorithms in its scope, the one on which we will focus in this thesis is the Variational Quantum Eigensolver (VQE), used for finding the eigenstates, energies and other properties of physical models. The model which we choose to explore is the Fermi-Hubbard model with interactions expanded beyond the nearest neighbours, and we will be particularly interested in its ground state energy, a defining property governing the existence of different phases of real-world systems and transitions between them. Likewise, ability to calculate this energy exactly is invaluable in understanding of the existing materials with peculiar properties and development of new ones, as well as possible applications in building the quantum computing hardware.

The quest for efficient solutions to the Fermi-Hubbard model and other strongly correlated quantum systems takes us into the realm of algorithms based on symmetries, group theory and quantum properties of matter—this thesis embarks on a two-fold exploration to solve the FH model of moderate size. The first approach leverages the power of analytic results and iterative block diagonalization based on symmetries and quantum nature of the model, while the second embraces the modern VQE methodology, relying on intuition and results built in the first.

1.1 Thesis outline

The rest of this introductory chapter is devoted to Intermezzo section, in which an intuitive overview of fundamental phenomena is given for readers without physical and quantum

computing background. In the second chapter, the Fermi-Hubbard model is introduced in detail and expanded with interactions between the next nearest neighbours. After this, we use several analytic results for handling model's limiting regimes, and then we carry out the process of complexity reduction using the symmetries (both quantum and classical) of the model, culminating in several exact results regarding the ground states and the corresponding energies of the model. All the calculations are complemented by implementations in Python. In the third chapter, the formalism of VQE is motivated and developed, followed by progressive construction of the quantum circuit to be used as the solution for the model. The fourth and final chapter provides a series of results regarding the constructed solution. Specifically, it is shown to work across a wide range of parameters in which the model exhibits qualitatively different behaviour, as well as in the presence of a particular type of noise. Ultimately, we present a brief review of the obtained results, accompanied by a succinct examination of potential avenues for further research.

1.2 Intermezzo (for the less physically inclined)

In this section, we provide an overview of the most pertinent topics that may not be known to an average mathematician or computer scientist, most notably in quantum physics, second quantization and quantum computing.

Quantum physics

When working on a formal treatment of quantum mechanics, its formalization is highly nontrivial and requires subtle knowledge of operator theory, functional analysis, measure theory, etc. [20] However, for the purposes of current work, most of the gory details are unnecessary, once the basic postulates and elements are outlined. Particularly, for quantum computing, there are few technical requirements beyond linear algebra. On the other hand, for the quantum model underlying our analysis, the Fermi-Hubbard model, second quantization formalism will be needed, and as such will be explained here in an intuitive manner.

Let us begin with the postulates of quantum mechanics[1].

Postulate 1 *An isolated quantum system is described by a complex Hilbert space $\mathcal{H}$. The state of the system is determined in full by its corresponding statevector, i.e. a unit vector in the system's state space $\mathcal{H}$. The Hilbert space of a composite system is the tensor product of Hilbert spaces of the single component systems.*

Before expanding on this postulate, let us briefly introduce the useful notation which is used in quantum physics almost exclusively—*Dirac braket notation*. In it, vectors are

[1] As mentioned, in a significantly reduced and simplified form.

denoted by $|\psi\rangle$, while the inner product of two vectors is denoted by $\langle |\psi_1\rangle , |\psi_2\rangle\rangle := \langle \psi_1|\psi_2\rangle$, from which its name actually stems, the first element is a *bra* and the second is a *ket*, together forming a bra(c)ket. The usefulness of this notation is more apparent with its simple demarcation of the outer product $|\psi_1\rangle \langle \psi_2|$, which is a Kronecker product of one vector with a Hermitian adjoint of the other, and generates a matrix of dimension equal to product of their individual dimensions. Of course, this expression immediately raises an issue, because in definition of notation for the inner product, the bra $\langle \psi_1|$ seems to be only a notational convenience, while in definition of outer product, it seems that it is an actual object, and we referred to it as a Hermitian adjoint, which is mostly correct, at least in operational sense. The underlying rigor here is that for each ket (vector) $|\psi\rangle$, there exists a dual vector (bra) $\langle \psi_1|$ which is an element of the dual space, i.e. a space of all linear forms on the underlying Hilbert space, as guaranteed by the Riesz lemma [21]. To make consequential sense of this, the lemma simply implies that for every ket, we are able to write a corresponding bra, and that in technical terms, this corresponds to applying the Hermitian conjugation to the vector and transposing it.

Now returning to the first postulate 1, we can already see the shift from the intuitive and experiential world of classical physics: if the states $|\psi_1\rangle$, $|\psi_2\rangle$, ... are valid statevectors, then any normalized linear combination of them is as well. In other words, the system can simultaneously be described by a linear combination of several states, or in other words be in a superposition of several different states at once—famously observed in phenomena such as the wave/particle duality—which is completely opposed to the everyday physical experience. The normalization condition here is important[2] for the following reason: the square of the absolute value of the coefficient of each state in a linear combination is interpreted as the probability that, upon measurement, system is found in that particular state. Evidently, all of the probabilities ought to sum to unity, resulting in said normalization condition of a generic quantum state.

Postulate 2 *Physical quantitites that can be measured or observed in experiments are denoted quantum mechanical observables. They are described by the set of self-adjoint*[3] *operators on $\mathcal{H}$, and their values correspond to eigenvalues of such operators.*

Although the measurement process in quantum mechanics is complex and not even fully understood at the moment, we will soon explain what it corresponds to in rough

[2] And actually the states are not described by vectors, but by rays in a quotient space in which all the vectors of same direction but different norms are identified.

[3] In literature, the terms symmetric, Hermitian and self-adjoint are sometimes used interchangeably (especially the latter two), however this is problematic even in some simple examples, but especially when dealing with operators on curved spacetimes, as in quantum gravity [22]. There is also a fundamental issue due to the Hellinger-Toeplitz theorem from which it follows that an unbounded symmetric operator (such as energy or position) cannot be defined everywhere, which is contrary to how we usually think of many natural operators [21].

terms. At this moment, let us simply imagine that we are observing the system to find out some of its properties. Evidently, there are properties such as energy of the system, which we will soon define to be an eigenvalue of a particular operator (Hamiltonian), that have obvious physical interpretation and should thus be real-valued. Since the self-adjoint operators are those with real spectra (as per different spectral theorems) [21], then it is convenient and intuitive to define observables as was done in this postulate.

From a technical standpoint, there can be several identical eigenvalues, i.e. multiple eigenstates corresponding to a particular eigenvalue. This is a phenomenon known as *degeneracy*, and it simply implies that there is not only a single eigenvector associated to an eigenvalue, but rather a higher-dimensional subspace of the Hilbert space, spanned by associated eigenvectors. As expected, all linear combinations that are parts of this space are eigenvectors with the same eigenvalue as well, seen easily from eigenvalue definition.

Postulate 3 *State of the time-independent system is determined by the time-independent Schrödinger's equation for statevector* $|\psi\rangle$*:*

$$\mathbf{H}(\mathbf{x})\,|\psi(\mathbf{x})\rangle = E\,|\psi(\mathbf{x})\rangle :\Longleftrightarrow \left[-\frac{\hbar^2}{2m}\nabla_{\mathbf{x}}^2 + V(\mathbf{x})\right]|\psi(\mathbf{x})\rangle = E\,|\psi(\mathbf{x})\rangle\ , \tag{1.1}$$

where $V(\mathbf{x})$ *is the potential energy of the sytem, while E denotes the total energy of the system as an eigenstate of the Hamiltonian operator* $\mathbf{H}$*.*

We first note that the aforementioned principle of superposition, i.e. the fact that a normalized linear combination of states is also a possible state of the system is evidently embedded into (1.1) since it is a linear partial differential equation. In a fully rigorous treatment, it is a matter of choice of whether to postulate the Schrödinger's equation or not and we will not dwell on this issue, however its plausibility can be seen somewhat immediately if one recalls that the momentum operator in classical physics acts as a generator of motion. Then, quite crudely, as per de Broglie wave-particle duality relation [23], momentum and wave number are proportional, so the statevector may contain an $\exp[\imath\mathbf{p}\cdot\mathbf{x}/\hbar]$ term. The action of $\nabla_{\mathbf{x}}^2$ will then yield a factor of $(\imath\mathbf{p})^2 = -\mathbf{p}^2$, and the first term in equation (1.1) turns to $\mathbf{p}^2/2m$, of identical form as the classical kinetic energy. Similarly, the potential term acting on an eigenstate will add a well-defined value to its total energy, corresponding to the potential energy.

There are several issues which render the Schrödinger's equation (1.1) difficult to solve. First of all, it is hard to find the exact form of Hamiltonian $\mathbf{H}$ that will describe a particular system well, for example due to lack of access to small systems in which the potential $V(\mathbf{x})$ can have an arbitrarily complicated form. Furthermore, since one is usually interested in composite systems, for which the corresponding Hilbert spaces are equal to products of individual Hilbert spaces as per postulate 1, the total system size will scale exponentially and, in general, yield an exponential number of equations to solve simultaneously. Sometimes,

this number can be reduced based on either approximations, or symmetry reductions, as we will explore in Section 2.5, but these methods are not applicable for an arbitrary system.

In general, we wish to solve the Schrödinger's equation (1.1) for eigenvalues and eigenenergies since then the other properties of physical models can be calculated and useful, experimentally verifiable predictions can be made. In present work, we will focus on the ground state energy calculations since, as was mentioned in introduction, they define the system behaviour in different phases and allow for understanding of existing materials and conceptualization of new ones.

Examples in quantum physics

To try and actualize the somewhat abstract notions put forward by the postulates of quantum mechanics in the previous subsection, let us consider some illustrative examples.

When discussing the concept of a statevector, a simple system to work with is a spin-$1/2$ particle such as electron, with states of spin projection up, denoted $|\uparrow\rangle$ and spin projection down, denoted $|\downarrow\rangle$ (we will refer to these as spin up and down, and usually omit the "projection"). If we focus only on the spin degree of freedom, evidently the Hilbert space is of dimension two, and an obvious basis (among the infinitely many) is of the two states: spin up and spin down. Thus, these states correspond to vectors

$$|\uparrow\rangle := \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad |\downarrow\rangle := \begin{pmatrix} 0 \\ 1 \end{pmatrix}. \tag{1.2}$$

Now the first postulate 1 and the following explanation, considered in scope of this example, imply that the system can, counterintuitively, be in a linear combination of these two states, such as

$$\sqrt{\frac{2}{3}}\,|\uparrow\rangle + \frac{1}{\sqrt{3}}\,|\downarrow\rangle = \frac{1}{\sqrt{3}}\begin{pmatrix} \sqrt{2} \\ 1 \end{pmatrix}.$$

This means that if we were to observe the system in some way, perhaps by acting on it with a magnetic field and then detecting in which way it is deflected[4], we would detect spin up with probability of $2/3$ and spin down with probability of $1/3$.

Regarding the second postulate 2, what would be some common examples of observables? Well, one was already mentioned in third postulate 3 and it is the Hamiltonian of the system, whose eigenvalues correspond to energies of different states in which the system could be found. In other words, diagonalizing the Hamiltonian produces the knowledge of the complete energy spectrum of the system. In current example, we have introduced states of spin projection up and down—naturally, we would like to assign an observable to this property as well, since, as we have noted, it can be actually observed in experiments.

[4] As was done in the seminal Stern-Gerlach experiment [24].

First of all, this projection can be taken along any spatial direction, but the common choice is the z axis, with spin operator denoted by $\mathbf{s}^z$. Exemplified by the Schrödinger's equation as en eigenvalue equation for the Hamiltonian, in case of spin we would like to construct an operator whose eigenvectors are $|\uparrow\rangle$ and $|\downarrow\rangle$, with eigenvalues $+1/2$ and $-1/2$[5], respectively. However, due to the simple form of these vectors (since they constitute a basis) given by (1.2), from the described requirements

$$\mathbf{s}^z|\uparrow\rangle = +\frac{1}{2}|\uparrow\rangle = \frac{1}{2}\begin{pmatrix}1\\0\end{pmatrix}, \quad \mathbf{s}^z|\downarrow\rangle = -\frac{1}{2}|\downarrow\rangle = -\frac{1}{2}\begin{pmatrix}0\\1\end{pmatrix},$$

one can immediately observe that the correct matrix form of the spin z projection operator is

$$\mathbf{s}^z = \frac{1}{2}\begin{bmatrix}1 & 0\\0 & -1\end{bmatrix}. \tag{1.3}$$

For further use, let us introduce the *Pauli matrices*, which simply correspond to three spin projection matrices of spin-$1/2$, up to a constant factor:

$$\sigma^x := \begin{bmatrix}0 & 1\\1 & 0\end{bmatrix}; \quad \sigma^y := \begin{bmatrix}0 & -\imath\\\imath & 0\end{bmatrix}; \quad \sigma^z := \begin{bmatrix}1 & 0\\0 & -1\end{bmatrix}. \tag{1.4}$$

Operator $\mathbf{s}^z$ can also be used to illustrate the concept of a *good quantum number*. Evidently, since $|\uparrow\rangle$ and $|\downarrow\rangle$ are eigenstates of the $\mathbf{s}^z$ operator, they are left invariant upon action of said operator. For them, the value of the spin projection along the z axis is well-defined, as the corresponding eigenvalue. On the other hand, if we were to take some arbitary linear combination of them $|\psi\rangle = \alpha|\uparrow\rangle + \beta|\downarrow\rangle$, how would it behave under action of $\mathbf{s}^z$? This is found from (1.3)

$$\mathbf{s}^z|\psi\rangle = \frac{1}{2}\begin{bmatrix}1 & 0\\0 & -1\end{bmatrix}\begin{pmatrix}\alpha\\\beta\end{pmatrix} = \frac{1}{2}\begin{pmatrix}\alpha\\-\beta\end{pmatrix} \not\propto |\psi\rangle$$

and we see that for a general state, action of $\mathbf{s}^z$ will not leave it invariant up to a factor—for such states that are not eigenstates we say that spin projection along z is not a good quantum number; its value is not well-defined.

The concept of quantum number will be immensely important in both of the following parts of the thesis. During the iterative diagonalization of the Hamiltonian, in certain subspaces, we will construct linear combinations that have good quantum numbers and then be able to reduce the Hamiltonian into blocks.

After introducing the notion of a good quantum number, it is important to point out another important subtlety regarding the different observables being well-defined at the

[5]Factors of $1/2$ are not important to present discussion, but are necessary—this has to do with the algebras of different spin operators [25].

same time. Let us consider two operators **a** and **b** that commute. Then, if a state $|\psi\rangle$ is an eigenstate of **a**, then it is an eigenstate of **b**, which can easily be seen from

$$\begin{aligned}&\mathbf{a}\,|\psi\rangle = \lambda\,|\psi\rangle \wedge \mathbf{ab} = \mathbf{ba} \implies \\ &\implies \mathbf{a}\left[\mathbf{b}\,|\psi\rangle\right] = \mathbf{ab}\,|\psi\rangle = \mathbf{ba}\,|\psi\rangle = \mathbf{b}\lambda\,|\psi\rangle = \lambda\left[\mathbf{b}\,|\psi\rangle\right]\ ,\end{aligned}$$

implying that $\mathbf{b}\,|\psi\rangle$ is a multiple of $|\psi\rangle$, as desired[6]. This result has profound implications in quantum mechanics. First of all, by its converse, we find that two observables do not have well-defined values simultaneously if they do not commute. This is encoded in the celebrated *Heisenberg uncertainty principle* [7], and its most famous implication is that the position and momentum of a particle cannot be known simultaneously, and to emphasize, this is not due to some perturbations induced by measurements or similar effects—this is nature's limit. On the other hand, if the observables' operators commute, then they have a common set of eigenstates and one can, in principle, know their values simultaneously after a measurement. This fact will be of immense importance later on when we take different operators that commute with the Hamiltonian and then find Hamiltonian's eigenstates by finding the eigenstates of those other simpler operators first.

An additional fundamental concept to explore in quantum mechanics is *entanglement*—a uniquely quantum phenomenon that has no classical counterpart and is also responsible for potential quantum advantage one may find in quantum computation. As an example, let us recall from postulate 1 that the Hilbert space of a composite system is found as a product of individual spaces. Thus, we can write the following state of two spin-$1/2$ particles:

$$\frac{1}{\sqrt{2}}\left[|\uparrow\rangle \otimes |\uparrow\rangle + |\uparrow\rangle \otimes |\downarrow\rangle\right] = \underbrace{|\uparrow\rangle}_{\mathcal{H}_1} \otimes \underbrace{\frac{1}{\sqrt{2}}\left[|\uparrow\rangle + |\downarrow\rangle\right]}_{\mathcal{H}_2}\ . \tag{1.5}$$

For this state, we see how it can be factorized trivially, so that the statevectors belonging to two constituents are separated. An important implication here is that if we were to change the state of one system, the state of the other would not be affected.

As an opposite example, we can consider a Bell state (important in quantum computing applications):

$$\frac{1}{\sqrt{2}}\left[|\uparrow\rangle \otimes |\downarrow\rangle + |\downarrow\rangle \otimes |\uparrow\rangle\right]\ .$$

As can trivially be shown by contradiction, this state cannot be separated into a product of states from two constituent subsystems, as (1.5) can. In other words, if we were to measure in some way the state of the first particle (e.g. by using a magnetic field, as mentioned

[6] As usual, there is subtlety in this result when degeneracies are introduced, however the conclusion and implications are unchanged.

before) and find that the first subsystem (first particle) is in state $|\uparrow\rangle$, then we immediately know that the complete system is in state $|\uparrow\rangle \otimes |\downarrow\rangle$ and thus the second subsystem is in state $|\downarrow\rangle$, as can be confirmed by a similar measurement. It is important to note that this holds (and was thoroughly verified experimentally) even if the two subsystems are separated enough so that communication between them is not possible between the observations. Thus, we seemingly find a "spooky action at a distance", as it was demarcated by Einstein [26], because it naively implies the transfer of information at a speed faster than that of light, in contradiction with the theory of relativity. Of course, no such contradiction actually occurs, but it took around 30 years for it to be explained away in a somewhat satisfying manner when John Bell wrote one of the most elegant and profound papers of the 20th century [27][7].

The concept of entanglement is evidently outside of regular physical experience, but how can it actually be exploited for quantum advantage? There are many examples, including quantum teleportation, quantum zero-knowledge games, etc. [31], but let us simply observe that once the first particle is measured, knowledge is immediately obtained about the second, which would obviously not be possible if the particles were considered in realm of classical physics, and this purely a consequence of entanglement, i.e. the system state not being factorizable.

Let us now turn our attention to an important restriction on quantum systems—the *Pauli exclusion principle* for fermions [32]. It has been empirically observed that the fundamental particles are indistinguishable, for example there is no way to differentiate between two electrons [33]. For this reason, if we are considering a product state of two electrons, e.g. (omitting the state product symbol as announced in 0.1) $|\psi_1\rangle |\psi_2\rangle$, we immediately see that it is not physically realizable, since upon the exchange of electrons, we would get a state $|\psi_2\rangle |\psi_1\rangle$, that based on the individual electron states, can be different from the initial one. Thus, we ought to find the states that truly encapsulate the particle indistinguishability. For example, let us say that some composite state $|\Psi\rangle$ of two particles acquires a phase factor λ upon particle exchange. Then doing the particle exchange twice, the acquired phase factor is equal to λ^2, however the state is returned to its initial configuration since the particles were swapped twice. Then, a simple relation $|\Psi\rangle = \lambda^2 |\Psi\rangle$ holds, from which we find[8] that $\lambda = \pm 1$. The $\lambda = +1$ case describes a class of particles called *bosons*, which we will not consider in this work. The particles of interest to us are called *fermions* and correspond to $\lambda = -1$. Bosonic and fermionic states are simply produced via symmetrization and antisymmetrization of statevectors, respectively. For example, the initial two-particle state can

[7] Nobel Prize in physics was awarded in 2022 for experimental confirmation of violation of Bell's inequalities to John F. Clauser, Alain Aspect and Anton Zeilinger [28, 29, 30].

[8] In two dimensions, there actually exists an additional class of solutions, giving rise to quasiparticles called anyons.

be symmetrized to yield

$$\frac{1}{\sqrt{2}}\left[|\psi_1\rangle|\psi_2\rangle + |\psi_2\rangle|\psi_1\rangle\right] \overset{1\leftrightarrow 2}{\rightsquigarrow} \frac{1}{\sqrt{2}}\left[|\psi_2\rangle|\psi_1\rangle + |\psi_1\rangle|\psi_2\rangle\right] = \frac{1}{\sqrt{2}}\left[|\psi_1\rangle|\psi_2\rangle + |\psi_2\rangle|\psi_1\rangle\right]$$

or antisymmetrized to yield

$$\frac{1}{\sqrt{2}}\left[|\psi_1\rangle|\psi_2\rangle - |\psi_2\rangle|\psi_1\rangle\right] \overset{1\leftrightarrow 2}{\rightsquigarrow} \frac{1}{\sqrt{2}}\left[|\psi_2\rangle|\psi_1\rangle - |\psi_1\rangle|\psi_2\rangle\right] = -\frac{1}{\sqrt{2}}\left[|\psi_1\rangle|\psi_2\rangle - |\psi_2\rangle|\psi_1\rangle\right] . \tag{1.6}$$

The first of these states (bosonic) can easily be seen to acquire no additional factor upon exchange $1 \leftrightarrow 2$, while the second (fermionic) acquires a factor of -1. A direct consequence of this is the following. Let us suppose that two particles are both in the same state, denoted $|\psi\rangle$. If they are of bosonic nature, a simple symmetrized state is

$$|\psi\rangle|\psi\rangle$$

since it is evidently invariant to particle exchange. However, when we try to construct an antisymmetric state, such as the one in (1.6), we find that

$$\frac{1}{\sqrt{2}}\left[|\psi\rangle|\psi\rangle - |\psi\rangle|\psi\rangle\right] = 0 ,$$

or in other words the expression vanishes. A profound implication of this result is the fact that no two fermions can occupy the same state in physical systems—already denoted the Pauli exclusion principle. For the model we will consider later on, it will mean that the electrons on a given site of the lattice will have to be in different states—since they are particles of spin $1/2$, there are only two states available and thus each spin site can accomodate at most two electrons.

Second quantization

An aforementioned issue of dealing with systems of multiple constituents is present not only in the quantum realm. For example the problem of three bodies interacting solely through gravitational force in classical physics has no known closed-form solution. For this reason, the formalism of second quantization that allows handling the many-body quantum systems is a remarkable toolkit, and it is used in several branches of physics, such as quantum field theory, condensed matter theory, nuclear physics, etc. [34]

In second quantization, the focus shifts from the statevectors of individual particles to the creation and annihilation operators that operate on states with variable particle num-

bers. For this purpose, the total Hilbert space of the system is decomposed into an antisymmetrized direct sum of Hilbert spaces of subsystems, denoted the Fock space[9,10].

Compared to the statevectors in regular quantum mechanics (historically also known as first quantization), the key elements of second quantization formalism are the aforementioned creation and annihilation operators, which act on quantum states. In this way, we get the quantum states in the occupation number representation—instead of explicitly specifying the properties, such as position, of each particle, we use occupation numbers to denote how many particles occupy each quantum state. This representation simplifies the mathematical formulation, making it especially convenient for many-body systems.

Let us illustrate the behaviour of these operators with simple examples. In solid state physics, we often consider lattices with several sites on which there can exist particles. In present work, these particles are fermions of spin $1/2$ and as per the exclusion principle, only two of them can exist on each spin site: one of spin up $(+1/2)$ and one of spin down $(-1/2)$. Thus, we can write the statevector of a system with 4 sites and one particle of spin up on site 2 as follows:

$$\left[|0_{1\uparrow}\rangle \oplus |0_{1\downarrow}\rangle \oplus |1_{2\uparrow}\rangle \oplus |0_{2\downarrow}\rangle \oplus |0_{3\uparrow}\rangle \oplus |0_{3\downarrow}\rangle \oplus |0_{4\uparrow}\rangle \oplus |0_{4\downarrow}\rangle\right]_{\text{antisym.}} =: |0_{1\uparrow}1_{2\uparrow}0_{3\uparrow}0_{4\uparrow}0_{1\downarrow}0_{2\downarrow}0_{3\downarrow}0_{4\downarrow}\rangle \ ,$$

i.e. 0 or 1 denote the (non)existence of a particle in a particular state, or in other words a particle of a particular spin on a particular site. The second expression is consistent with compact notation of composite states, and also the fermionic nature of constituents is incorporated into this expression via antisymmetrization, analogous to (1.6). This statevector can also be written using the *creation operator* $\mathbf{c}^\dagger_{2\uparrow}$:

$$|0_{1\uparrow}1_{2\uparrow}0_{3\uparrow}0_{4\uparrow}0_{1\downarrow}0_{2\downarrow}0_{3\downarrow}0_{4\downarrow}\rangle = \mathbf{c}^\dagger_{2\uparrow}\,|0_{1\uparrow}0_{2\uparrow}0_{3\uparrow}0_{4\uparrow}0_{1\downarrow}0_{2\downarrow}0_{3\downarrow}0_{4\downarrow}\rangle \ ,$$

with two operator subscript indices corresponding, in order, to spin site and particle spin. The *annihilation operator* $\mathbf{c}_{2\uparrow}$ has an inverse effect (and is actually a Hermitian adjoint of the creation operator, $(\mathbf{c}^\dagger)^\dagger = \mathbf{c}$):

$$\mathbf{c}_{2\uparrow}\,|0_{1\uparrow}1_{2\uparrow}0_{3\uparrow}0_{4\uparrow}0_{1\downarrow}0_{2\downarrow}0_{3\downarrow}0_{4\downarrow}\rangle = |0_{1\uparrow}0_{2\uparrow}0_{3\uparrow}0_{4\uparrow}0_{1\downarrow}0_{2\downarrow}0_{3\downarrow}0_{4\downarrow}\rangle \ .$$

[9] Roughly, the difference of sums and products in this context is that sums combine Hilbert spaces in a way that keeps their individual structures separate (think spaces with different numbers of particles, evidently disjoint), while the products combine spaces in a way that encompasses all possible combinations from individual spaces.

[10] When considering discrete solid state models such as the Fermi-Hubbard model, we usually need not mess with statevectors encoding spatial dependence of quantum fields exactly due to the discrete nature of the system, allowing us to deal only with the creation and annihilation operators. Furthermore, since the particles at hand are fermions, and thus subject to the Pauli exclusion principle, not only are the system states simpler, but the calculations are simpler as well since both creation and annihilation operators are nilpotent of order 2. For practical purposes, this reduces to switch from the usual quantum states in initial Hilbert space, to occupation number states in Fock space.

The intuitive notion of these operators is evident, they simply create/destroy particles in particular states. From a more technical perspective, if we define an *anticommutator* of two operators $\mathbf{a}$ and $\mathbf{b}$ via $\{\mathbf{a},\mathbf{b}\} \coloneqq \mathbf{ab} + \mathbf{ba}$, the fermionic creation and annihilation operators satisfy the anticommutation relations[11]:

$$\forall\, j,l,\sigma,\sigma' \colon \quad \left\{\mathbf{c}_{j\sigma}, \mathbf{c}^\dagger_{l\sigma'}\right\} = \delta_{jl}\delta_{\sigma\sigma'}\,, \quad \left\{\mathbf{c}_{j\sigma}, \mathbf{c}_{l\sigma'}\right\} = 0\,, \quad \left\{\mathbf{c}^\dagger_{j\sigma}, \mathbf{c}^\dagger_{l\sigma'}\right\} = 0\,. \tag{1.7}$$

The antisymmetric nature of fermions (1.6) discussed in terms of regular quantum mechanics is embedded into the anticommutation relations. For example, by the third identity in (1.7), we find that swapping the two particles induces a minus sign in a quantum state:

$$|1_{1\uparrow}1_{1\downarrow}\rangle = \mathbf{c}^\dagger_{1\uparrow}\mathbf{c}^\dagger_{1\downarrow}|0_{1\uparrow}0_{1\downarrow}\rangle \overset{(1.7)}{=} -\mathbf{c}^\dagger_{1\downarrow}\mathbf{c}^\dagger_{1\uparrow} = -|1_{1\downarrow}1_{1\uparrow}\rangle\,.$$

Similarly, using $l = j$ and $\sigma' = \sigma$ in third equation in (1.7) yields $2\left(\mathbf{c}^\dagger_{j\sigma}\right)^2 = 0 \implies \left(\mathbf{c}^\dagger_{j\sigma}\right)^2 = 0$, directly expressing the fact that no two fermions in the same quantum state can occupy a single site, known as the previously introduced Pauli exclusion principle. With some additional care, it can be shown rigorously that second quantization indeed preserves all of the properties of fermions [21].

Based on the intuitive notion of creation and annihilation operators, the Hamiltonian in second quantization takes a remarkably different form compared to its first quantization counterpart. It involves summations over all subsets of particles, accounting for their interactions. Additionally, terms involving the creation and annihilation operators emerge naturally, allowing for the representation of processes involving variable particle numbers[12].

To further elucidate the use of second quantization operators, let us consider specific terms that are usually found in solid state Hamiltonians and see how they act on different quantum states. This will be useful later on when we define the Fermi-Hubbard Hamiltonian (2.1).

For example, the term $\mathbf{c}^\dagger_{j\sigma}\mathbf{c}_{l\sigma}$. What is the effect of this term on number occupation states? Well, since the first operator to act is the rightmost one, i.e. $\mathbf{c}_{l\sigma}$, it will annihilate every state in which there is no particle of spin σ on site l. Thus, only the states in which such an excitation exists will yield a nontrivial contribution to the final state upon the action of this term. The next term to act is the creation operator $\mathbf{c}^\dagger_{j\sigma}$, and it will excite a particle of spin σ on site j. Consequently, it will annihilate states in which such a particle is already excited due to the Pauli exclusion principle. Therefore, a general state that will remain as a nonzero contribution after the action of the operator $\mathbf{c}^\dagger_{j\sigma}\mathbf{c}_{l\sigma}$ will be of form

[11] And construct an algebra similar to Clifford.

[12] The terms "first" and "second" quantization have drifted from their original intended meaning, which was essentially related to the fact that it was thought that the already quantized wavefunctions ought to be quantized once more, hence the term "second". As a thumb rule, if the phase-space variables trajectories, we call the process of quantisation "first", and if the phase-space variables are fields, we call it "second" [35].

$|\ldots, 1_{l\sigma}, \ldots, 0_{j\sigma}, \ldots\rangle$. The action is carried out easily (up to a possible sign change due to fermionic anticommutation relations (1.7)):

$$\mathbf{c}^\dagger_{j\sigma}\mathbf{c}_{l\sigma}\,|\ldots, 1_{l\sigma}, \ldots, 0_{j\sigma}, \ldots\rangle \sim \mathbf{c}^\dagger_{j\sigma}\,|\ldots, 0_{l\sigma}, \ldots, 0_{j\sigma}, \ldots\rangle \sim |\ldots, 0_{l\sigma}, \ldots, 1_{j\sigma}, \ldots\rangle\ . \tag{1.8}$$

From this example, we can immediately see that the effect of this operator was a "hopping" of the particle from site l to site j, and that no particles on other sites were created or annihilated.

Another important operator is a special case of the previous one, obtained by setting $j = l$, i.e. $\mathbf{c}^\dagger_{j\sigma}\mathbf{c}_{j\sigma}$. Evidently, since the rightmost term acts the first, the only nonvanishing terms after its action will be those in which there exists a particle of spin σ at site j, all others will vanish. The action of this operator on such a state can also be found easily:

$$\mathbf{c}^\dagger_{j\sigma}\mathbf{c}_{j\sigma}\,|\ldots, 1_{j\sigma}, \ldots\rangle = \mathbf{c}^\dagger_{j\sigma}\,|\ldots, 0_{j\sigma}, \ldots\rangle = |\ldots, 1_{j\sigma}, \ldots\rangle\ . \tag{1.9}$$

From this, we find that the state in which there exists a particle of spin σ at site j was left invariant, while we previously noted that the remaining states will be annihilated. These two operations can be interpreted as multiplying by 1 and multiplying by 0, respectively. It turns out that this result holds in general, and even for particles that are not fermions, i.e. this operator counts the particles of a certain spin on a certain site and is thus denoted the *number operator* and demarcated by

$$\mathbf{n}_{j\sigma} \coloneqq \mathbf{c}^\dagger_{j\sigma}\mathbf{c}_{j\sigma}\ . \tag{1.10}$$

Finally, as an important method often used when presented with a Hamiltonian of model on a discrete lattice in second quantization form, we highlight the application of the discrete Fourier transform to the creation and annihilation operators. As we will see later on in Subsection 2.4, carrying out DFT allows to turn from many operators corresponding to different sites, to the operators common to all sites, summed over different wave numbers, as can also be demonstrated to have further applications [36].

To apply this method, we define the discrete Fourier transform in accordance with the notational choices presented in 0.1:

$$\begin{aligned}
\mathbf{c}_{j\sigma} &= \frac{1}{\sqrt{N}}\sum_{q\in\Gamma} e^{-\imath q j}\mathbf{c}_{q\sigma} \Longleftrightarrow \mathbf{c}_{q\sigma} = \frac{1}{\sqrt{N}}\sum_{j=1}^{N} e^{\imath jq}\mathbf{c}_{q\sigma}\,. \\
\mathbf{c}^\dagger_{j\sigma} &= \frac{1}{\sqrt{N}}\sum_{q\in\Gamma} e^{\imath q j}\mathbf{c}^\dagger_{q\sigma} \iff \mathbf{c}^\dagger_{q\sigma} = \frac{1}{\sqrt{N}}\sum_{j=1}^{N} e^{-\imath jq}\mathbf{c}^\dagger_{q\sigma}\,,
\end{aligned} \tag{1.11}$$

where the momentum q takes values in set

$$\Gamma \coloneqq \left\{\frac{2\pi j}{N}\right\}_{j\in[1\cdot\cdot N]}\ . \tag{1.12}$$

As a simple consequence of these definitions, we include the following useful identity that can easily be derived by direct calculation, or by simply noting that a polygon centered at the origin of the complex plane trivially has a centre of mass at the origin:

$$\sum_{j=1}^{N} e^{\imath jq} = N\delta_{q,0}\,. \tag{1.13}$$

Total spin and pseudospin (projection) operators

Tying together the ideas of operators in first quantization and treatment of multi-particle systems with second quantization, let us introduce operators that correspond to properties of entire composite systems of multiple particles.

An example of a spin projection operator along the z axis for a single particle was given in (1.3). How could we construct an operator that contains the information on the complete spin projection along the z axis, i.e. for the entire particle system? Well, in (1.10) we constructed an operator counting the number of particles of a particular spin on a particular site. Thus, we can simply sum $+1/2$ for all the particles of spin up and $-1/2$ for all the particles of spin down and define the total spin projection along the z axis operator as follows:

$$\mathbf{S}^z \coloneqq \sum_j \left(+\frac{1}{2}\right)\mathbf{n}_{j\uparrow} + \sum_j \left(-\frac{1}{2}\right)\mathbf{n}_{j\downarrow} = \frac{1}{2}\sum_j \left(\mathbf{n}_{j\uparrow} - \mathbf{n}_{j\downarrow}\right)\,. \tag{1.14}$$

Analogously to the fermionic creation and annihilation operators defined by (1.7), one can define the spin raising (lowering) operators $\mathbf{S}^+$ ($\mathbf{S}^-$) as destroying (creating) a particle of spin down (up), followed by creating (destroying) a particle of spin up (down):

$$\begin{aligned}\mathbf{S}^+ &\coloneqq \sum_j \mathbf{c}^\dagger_{j\uparrow}\mathbf{c}_{j\downarrow}\,,\\ \mathbf{S}^- &\coloneqq \sum_j \mathbf{c}^\dagger_{j\downarrow}\mathbf{c}_{j\uparrow} = \mathbf{S}^{+\dagger}\,.\end{aligned}$$

Now the total spin operator (or rather its square) is defined in expected way, however it can be rewritten for implementational and computational convenience:

$$\begin{aligned}\mathbf{S}^2 &\coloneqq \sum_{jl}\left[\mathbf{s}^x_j\mathbf{s}^x_l + \mathbf{s}^y_j\mathbf{s}^y_l + \mathbf{s}^z_j\mathbf{s}^z_l\right] =\\ &= \frac{1}{2}\left[\mathbf{S}^+\mathbf{S}^- + \mathbf{S}^-\mathbf{S}^+\right] + \left(\mathbf{S}^z\right)^2\,.\end{aligned} \tag{1.15}$$

Later on, after the Fermi-Hubbard model Hamiltonian (2.1) is introduced, we will note that the total spin projection along the z axis $\mathbf{S}^z$ and the square of the total spin operator $\mathbf{S}^2$ both commute with it, implying that they admit a common set of eigenvectors, which will be useful for Hamiltonian diagonalization later on in Section 2.5.

A class of operators analogous to the spin operators are the *pseudospin* operators, defined analogously:

$$\begin{aligned}
\mathbf{R}^z &:= \frac{1}{2}\sum_j (\mathbf{n}_{j\uparrow} + \mathbf{n}_{j\downarrow} - \mathbb{1}) \,, \\
\mathbf{R}^+ &:= \sum_j (-1)^j \mathbf{c}^\dagger_{j\uparrow}\mathbf{c}^\dagger_{j\downarrow} \,, \\
\mathbf{R}^- &:= \sum_j (-1)^j \mathbf{c}_{j\downarrow}\mathbf{c}_{j\uparrow} = \mathbf{R}^{+\dagger} \,, \\
\mathbf{R}^2 &:= \frac{1}{2}\left[\mathbf{R}^+\mathbf{R}^- + \mathbf{R}^-\mathbf{R}^+\right] + (\mathbf{R}^z)^2 \,.
\end{aligned} \tag{1.16}$$

Intuitively, we can see that, e.g. the first pseudospin operator quantifies how far the state is from having the same number of particles as there are spin sites (later introduced as half filling), since the summation of the first two terms is equal to the total number of particles (since each particle can be either of spin up or spin down) and the third term is equal to half of the total particle number (once summed over all spin sites), all multiplied by a factor of $1/2$.

Similarly to $\mathbf{S}^z$, the pseudospin projection operator $\mathbf{R}^z$ contains only the numbering operator terms and thus commutes with the model Hamiltonian (2.3) as well. However, the more detailed discussion of commutation of $\mathbf{R}^2$ is deferred until the following chapter since it depends on the nature of the interactions in the system 2.1.

Why bother with quantum computing?

The idea of superposition found in postulate 1 can be applied to construct a quantum version of a bit—a *qubit*. As a classical bit can either be in state 0 or 1, a qu(antum)bit can be in a linear combination of basis states $|0\rangle$ and $|1\rangle$, which is physically realizable via a two state system, for example the aforementioned spin-$1/2$ particle.

First examples of quantum algorithms, such as those by Deutsch and Josza were somewhat contrived and of no apparent use, however still demonstrated indisputably that there exist problems that are exponentially costly via classical computation, but polynomially when solved using its quantum counterpart. During the rest of the 80s and the first half of 90s, several similar examples were constructed, but, as was mentioned in introduction, quantum computing really came into the light in 1994 when Peter Shor constructed efficient quantum algorithms for prime factorization and discrete logarithm problems[13]. This was a remarkable stepping stone in quantum computing for at least two reasons. From the complexity theory perspective, these algorithms are believed to be among those "in-between" problems that are not in P and are not NP-complete; in other words without a

[13] Just in August of 2023, a more efficient version of Shor's algorithm was found [37].

polynomial solution, while some of the most efficient implementations of Shor's algorithm theoretically run with $\mathcal{O}(\log n^3)$ complexity. Secondly, currently used encryption methods such as RSA rely on the fact that the factorization of large numbers is intractable, implying that an actual useful implementation of Shor's algorithm would allow for encryption breaking by quick factorization of the large numbers used in cryptographic keys. After this, several other potentially useful algorithms, such as the Grover's unstructured search algorithm were constructed—thus, it is presumed that the potential implementation of quantum computers in the forthcoming age could be harnessed for great computational advantage.

Examples of quantum computing

Finally, let us illustrate quantum counterparts of logical gates and include several examples, such as the ones that will be used later on when building actual quantum circuits. These circuits are, somewhat intuitively, also represented by linear operators since we want to use them to manipulate qubits, which are described by quantum states. A singular condition imposed on operators that represent quantum gates is *unitarity*—this is easy to see since unitary operations preserve the inner product, and thus the total sum of probabilites for all the states included in a linear combination will remain a unity, as it should.

As a basic example, we can consider the NOT gate, usually denoted by X which, in accordance with the classical NOT gate, acts as a bit flip. In other words, it will transform state $|0\rangle$ to $|1\rangle$ and vice versa, with obvious matrix form:

$$X = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} . \tag{1.17}$$

This gate evidently acts on a single qubit and has a form of a σ^x Pauli matrix (1.4). When applied to a qubit in state $\alpha|0\rangle + \beta|1\rangle$, this gate will produce the following state:

$$X[\alpha|0\rangle + \beta|1\rangle] = \alpha|1\rangle + \beta|0\rangle \ . \tag{1.18}$$

The gates Y and Z are also defined as respective Pauli matrices, and these three gates can be used to generate rotations in qubit space. Specifically, if we imagine a unit sphere and statevector $|0\rangle$ at an upper point on the z axis, and statevector $|1\rangle$ at a lower point on the z axis, then the rotations by an angle θ around different axes can be generated by these three matrices:

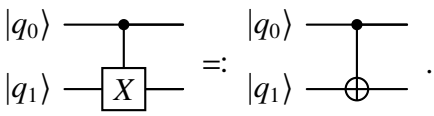
$$\begin{aligned}
\mathbf{R}_x(\theta) &:= \exp[\imath X\frac{\theta}{2}] = \begin{bmatrix} \cos\frac{\theta}{2} & -\imath\sin\frac{\theta}{2} \\ -\imath\sin\frac{\theta}{2} & \cos\frac{\theta}{2} \end{bmatrix} ; \\
\mathbf{R}_y(\theta) &:= \exp[\imath Y\frac{\theta}{2}] = \begin{bmatrix} \cos\frac{\theta}{2} & -\sin\frac{\theta}{2} \\ \sin\frac{\theta}{2} & \cos\frac{\theta}{2} \end{bmatrix} ; \\
\mathbf{R}_z(\theta) &:= \exp[\imath Z\frac{\theta}{2}] = \begin{bmatrix} e^{-\imath\frac{\theta}{2}} & 0 \\ 0 & e^{\imath\frac{\theta}{2}} \end{bmatrix} .
\end{aligned} \tag{1.19}$$

These resulting rotations have a somewhat similar form t othe usual rotations in 2D plane, and they will in general mix states $|0\rangle$ and $|1\rangle$[14].

A simple way of expanding these single-qubit gates to act on two qubits is via controls. For example, we can introduce the controlled qubit gate CX, constructed by the following prescription. Let this gate act on two qubits. If the first qubit is in state $|0\rangle$, then leave the second qubit unchanged. If the first qubit is in state $|1\rangle$, act on the second qubit with gate X. The first qubit (control qubit) is left unchanged in both cases. As an illustration, we can easily find the matrix representation of this gate:

$$
\begin{aligned}
|0\rangle\langle 0| \otimes \mathbb{1} + |1\rangle\langle 1| \otimes X &= \begin{pmatrix} 1 & 0 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} \otimes \mathbb{1} + \begin{pmatrix} 0 & 1 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \end{pmatrix} \otimes X = \\
&= \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \otimes \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \otimes \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} = \\
&= \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} + \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix} = \\
&= \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix}
\end{aligned}
$$

and further write it as a quantum circuit, with the upper wire corresponding to the first qubit and the lower wire to the second qubit:



(1.20)

The second circuit represents a usual way of denoting the controlled-X gate, which will also be used in this work.

In a similar way, one can define controlled versions of different gates, especially of the quantum rotations (1.19), as we will see during the construction of a special quantum circuit in Section 3.4.

[14]The different exponential and imaginary terms crop up due to the rotation group $SU(2)$ of a spin-$1/2$ system being a double cover of the classical rotation group $SO(3)$.

2

Fermi-Hubbard Model

The Fermi-Hubbard model was independently proposed by Martin Gutzwiller [38], Junjiro Kanamori [39] and John Hubbard [40] in 1963 to explore the behaviour of interacting electron systems. Initially, it was used to describe the phenomenon of itinerant ferromagnetism of transition metals, but its relevance and applicability would go on to reach far beyond that original scope. Today, it finds application in describing various materials, such as metal oxides and ultracold atomic gases, and displays intriguing phases with behaviours such as Mott insulation, (anti)ferromagnetism and superfluidity [41].

The model Hamiltonian is of a deceivingly simple form and although it has been a subject of intense research for more than half of a decade, no general analytic solution for it is known at the time. In simplest terms, the issues with the problem solvability mostly stem from the interaction term U in (2.1) since it induces strong correlations between the electrons and does not allow to decouple the Hamiltonian terms so they can be solved separately as smaller problems (as we will have done in Subsection 2.4 for the noninteracting regime). For these reasons, the FH model is a focal point for both theoretical and experimental studies in modern research, with several numerical methods developed for tackling it. As will be seen in the following chapter, we will also take a somewhat different approach.

A general form of the Fermi-Hubbard Hamiltonian is:

$$\mathbf{H} = -\sum_{\langle j,l\rangle\sigma} t_{jl}\mathbf{c}^{\dagger}_{j\sigma}\mathbf{c}_{l\sigma} + U\sum_{j}\mathbf{n}_{j\uparrow}\mathbf{n}_{j\downarrow}\,, \tag{2.1}$$

where $\langle j,l\rangle$ denotes iteration over all neighbouring[1] sites, while σ denotes iteration over all (in fermionic case both) spins. As outlined in Subsection 1.2, we can immediately understand the particular terms intuitively. First, the terms with factors $-t_{jl}$, described by (1.8), correspond to particles of spin σ hopping between sites j and l, and the total

[1] This can in general mean first, second, third, etc. neighbours.

effect of this transition is to change (most likely lower) the energy of the system by exactly t_{jl}. On the other hand, the U term sums over products of form $\mathbf{n}_{j\uparrow}\mathbf{n}_{j\downarrow}$, which will evidently be nonzero if and only if there exist both a particle of spin up and a particle of spin down on the jth site. Thus, it corresponds to the onsite interaction, e.g. a repulsion ($U > 0$)[2,3] of electrons at the same site and is taken to be constant (site-independent), partly because the different particles of the same kind are indistinguishable.

As an obvious omission, we see that there is no chemical potential term[4] present in (2.1). This is due to the fact that we will exclusively focus on the half-filled regime, for several reasons. First of all, the subspace at half filling and vanishing total spin projection is of the largest dimension, as can be seen by a simple combinatorial argument (cf. 2.3), and thus the trickiest to handle with our algorithms, as was already observed in terms of the solution complexity [43]—in other words, if the approaches are powerful enough for this configuration, it is expected that they will handle the remaining ones with ease. Second, at half filling (with some additional constraints), there hold several important analytic results, such as Lieb's and Shen's theorems 1 and 2, which allow for educated approach to reducing the complexity of our algorithms. Third, and perhaps most importantly, the model at half filling, especially with next nearest neighbour interactions added exhibits a rich phase diagram with several phases in which it displays qualitatively different properties, such as a metallic, insulating and antiferromagnetic behaviour [44].

For the sake of concreteness, let us denote the hopping terms in (2.1) more explicitly and include the next nearest neighbour hopping, since, as mentioned in previous paragraph, this results in a more complex phase diagram. Specifying the nearest neighbour (NN) and next nearest neighbour (NNN) interaction strengths t and t', respectively, (2.1) can now be written more explicitly as:

$$\mathbf{H} = -\sum_{j=1}^{N}\sum_{\sigma}\left(t\mathbf{c}^{\dagger}_{j\sigma}\mathbf{c}_{(j+1)\sigma} + t'\mathbf{c}^{\dagger}_{j\sigma}\mathbf{c}_{(j+2)\sigma} + \text{H.c.}\right) + U\sum_{j=1}^{N}\mathbf{n}_{j\uparrow}\mathbf{n}_{j\downarrow}\,, \tag{2.2}$$

where we have assumed that the hopping terms depend only on the distance between the two sites and not the sites themselves, due to the high degree of symmetry of the Hamiltonian—it is not changed by interchanging the sites. Furthermore, H.c. stands for Hermitian conjugate of the expression preceding the term, and is necessary here compared to (2.1) because otherwise, e.g. hopping between sites j and $j+1$ would be unidirectional, and in this way we make sure that hopping in the other direction is possible as well.

As a technical detail, we note that in case of the interaction range being equal to half the system size, it will be counted twice. This is the case with the $N = 4$ system we will

[2]This term is usually taken to be positive since it stems from the repulsive Coulomb interaction between electrons.

[3]Once the solution in the repulsive case is found, one can simply handle the $U \rightsquigarrow -U$ case using the Shiba map [42] which exchanges the spin and pseudospin.

[4]This term regulates the number of particles in the system.

focus on and one can readily see this fact by summing explicitly in equation (2.2). To account for this, we restrict our attention to the particular Hamiltonian of $N = 4$ case and divide the t' term by 2 to account for this overcounting. Thus, the most explicit form of the Hamiltonian is:

$$\mathbf{H} = -\sum_{j=1}^{4}\sum_{\sigma}\left(t\mathbf{c}^{\dagger}_{j\sigma}\mathbf{c}_{(j+1)\sigma} + \frac{t'}{2}\mathbf{c}^{\dagger}_{j\sigma}\mathbf{c}_{(j+2)\sigma} + \text{H.c.}\right) + U\sum_{j=1}^{4}\mathbf{n}_{j\uparrow}\mathbf{n}_{j\downarrow} \tag{2.3}$$

and it can be visualized in Figure 2.1.

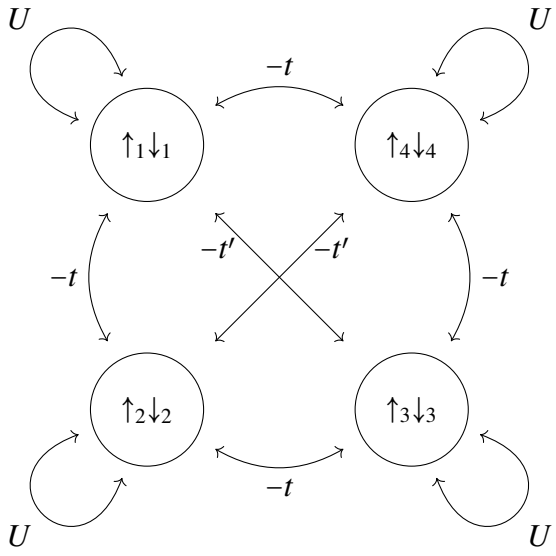


Figure 2.1: Visualization of the model we focus on in this work, with nearest neighbour hopping t, next nearest neighbour hopping t' and onsite interaction U.

2.1 Symmetries of the model

Spin symmetry

Both the total spin operator $\mathbf{S}^2$ and its projection $\mathbf{S}^z$ commute with the Hamiltonian. This can be easily seen for $\mathbf{S}^z$ since in (2.2) the hopping terms destroy and create particles of the same spin, and the onsite interaction terms simply count the spins, without creating or destroying any particles. To prove commutation of the total spin operator, it is more convenient to use its second form in (1.15). In it we see that the third term commutes with

$\mathbf{H}$ since it comprises of $\mathbf{S}^z$ only, while for the first term, it can be shown that both $\mathbf{S}^+$ and $\mathbf{S}^-$ commute with $\mathbf{H}$ separately and thus the expression as whole does as well. We recall that these commutation properties imply that $\mathbf{H}$, $\mathbf{S}^z$ and $\mathbf{S}^2$ all admit a common basis of eigenstates, and as was announced, this fact will be used to diagonalize $\mathbf{H}$ more easily in Section 2.5.

Pseudospin symmetry

As for the spin projection along the z axis, we see that $\mathbf{J}^z$, as defined in (1.16) consist of number operators only and since the FH Hamiltonian preserves the particle number, $\mathbf{J}^z$ commutes with it. On the other hand, it can be shown that the total pseudospin operator $\mathbf{J}^2$ commutes with $\mathbf{H}$ only if the following condition is satisfied by the hopping interactions (it commutes with the onsite interaction term regardless) [45]:

$$\forall\, j, l\colon\ t_{jl}\left[(-1)^j + (-1)^l\right] = 0\ .$$

In our consideration, we see that the nearest neighbour hopping term t, which is between adjacent sites, satisfies this condition since $t\left[(-1)^j + (-1)^{j+1}\right] = 0$. However, the next nearest neighbour hopping reaches two sites apart and $t'\left[(-1)^j + (-1)^{j+2}\right] \neq 0$. Thus, the total pseudospin operator $\mathbf{J}^2$ will commute with $\mathbf{H}$ if and only if we exclude this next nearest neighbour hopping term. In this case once again, we will be able to leverage the much more easily found eigenstates of $\mathbf{J}^z$ and $\mathbf{J}^2$ to diagonalize $\mathbf{H}$.

C_{4v} symmetry group

Evidently, the model we are analysing, depicted in Figure 2.1 possesses symmetries of a square. Thus, our Hamiltonian possesses them as well. Operations that leave the Hamiltonian invariant are rotations by $\pi/2$ (C_4), π ($C_2 = C_4^2$) and $3\pi/2$ (C_4^3), along with four reflections: two along the vertical planes that contain the square diagonals ($\sigma_{\mathrm{v}}^{(13)}$ and $\sigma_{\mathrm{v}}^{(24)}$) and two along the planes bisecting the angles of axes invariant to C_2 rotations ($\sigma_{\mathrm{d}}^{(13)}$ and $\sigma_{\mathrm{d}}^{(24)}$). If we add the identity operation (E), we see that the corresponding symmetry group is equal to $C_{4v} = D_4$.

This symmetry group has five irreducible representations (irreps), and they are given in Table 2.1—these will be used during the complexity reduction process in Subsection 2.5, while the group operators that commute with the Hamiltonian will be used directly in the final step of Subsection 2.5.

C_{4v}	E	$2C_4$	C_2	$2\sigma_v$	$2\sigma_d$	quad. fun.
A_1	1	1	1	1	1	x^2+y^2, z^2
A_2	1	1	1	−1	−1	-
B_1	1	−1	1	1	−1	x^2-y^2
B_2	−1	1	1	−1	1	xy
E	2	0	−2	0	0	(xz, yz)

Table 2.1: Character table of point group C_{4v} with five irreps and quadratic functions.

2.2 One spin theorem and one pseudospin theorem

In this section, we outline two historically important theorems that help restrict the space of different states needed to explore to find ground states of the Fermi-Hubbard model on a bipartite lattice at half filling, as will be seen in Section 2.5.

Theorem 1 (Lieb's theorem—repulsive case)**.** *Let the onsite interaction of the Fermi-Hubbard model be constant and positive ($U_i \coloneqq U > 0$). Let there be an even number of sites, with the electron number equal to it, i.e. the half filling regime. Finally, let the hopping interaction graph be connected bipartite with sublattices* $\mathbf{A}$ *and* $\mathbf{B}$ *and* $|\mathbf{B}| \geq |\mathbf{A}|$*, meaning that the only nonvanishing hopping terms are between the two sublattices. Then:*

a) Ground state of the model is unique (apart from the trivial $(2\mathbf{S}+1)$*-fold degeneracy.);*

b) Spin of the ground state is equal to

$$\mathbf{S} = \frac{1}{2}\left(|\mathbf{B}| - |\mathbf{A}|\right).$$

Proof of this theorem is somewhat involved and thus omitted since it relies on several tools outside the thesis scope. Let us now analyze how our model (2.3) falls within the theorem's conditions. First of all, when $t' = 0$, then the first site interacts with third site only (and vice versa), while the second site interacts with fourth site only (and vice versa). Thus, the model lattice is bipartite with $A = \{1,3\}$ and $B = \{2,3\}$, with $|B| \geq |A|$ evidently satisfied. As we are considering half filling, the number of electrons is equal to the number of sites. Finally, we have restricted our attention to the repulsive constant interaction case and have assumed in (2.3) that $U > 0$. Thus, all of the assumptions of theorem 1 are satisfied and we can conclude that the total spin of the ground state is equal to

$$\mathbf{S} = \frac{1}{2}(2-2) = 0.$$

Since the total spin of the ground state is zero, the total spin projection of the ground state must be zero as well. This fact will be of importance later on and will serve as a filter

when we choose the subspaces with vanishing $\mathbf{S}^2$ and $\mathbf{S}^z$ during the process of complexity reduction of the Hamiltonian 2.5.

Theorem 2 (Shen's theorem—repulsive case)**.** *Let the onsite interaction of the Fermi-Hubbard model be constant and positive ($U_i := U > 0$). Let there be an even number of sites, with the electron number equal to it, i.e. the half filling regime. Finally, let the hopping interaction graph be connected and bipartite with sublattices* **A** *and* **B** *and* $|\mathbf{B}| \geq |\mathbf{A}|$*, meaning that the only nonvanishing hopping terms are between the two sublattices. Then:*

a) Ground state of the model is unique;

b) Pseudospin of the ground state is equal to

$$\mathbf{R} = 0\,.$$

Similarly to before, one can use this theorem during complexity reduction of the Hamiltonian to focus only on the subspaces with $\mathbf{R}^2 = 0$ and $\mathbf{R}^z = 0$. Of course, the latter condition is satisfied trivially at half filling because the $\mathbf{R}^z$ quantifies how far away a particular state is from half filling (1.16), and of course this property is unchanged by the Hamiltonian (2.3) since it commutes with $\mathbf{R}^z$.

2.3 half filling zero spin projection sector basis

Having outlined two theorems on the spin and pseudospin nature of the system at particular conditions we are considering, let us more explicitly determine the set of vectors that span the complete subspace of the half-filled zero spin projection configurations. We do so because states with definite values of $\mathbf{S}^z$ and $\mathbf{R}^z$ are simple to construct and number, compared to those with definite values of **S** and **R**.

In a fermionic system with N sites, there are $2N$ possible particles, N of spin up and N of spin down. Then half-filing (or Shen's theorem 2 since the only possible projection of $\mathbf{R} = 0$ along the z axis is zero) implies that there exist exactly N particles in the system in ground state, while total spin projection zero $\mathbf{S}^z = 0$ implies that half of those, i.e. $N/2$ are spin up and half are spin down. Considering, e.g. the N possible spin up particles, we can choose $N/2$ of them in $\binom{N}{N/2}$ ways since their ordering is irrelevant. The same holds for the spin down particles, yielding a total of $\binom{N}{N/2}^2$ possibilities, which in $N = 4$ case reduces to:

$$\left.\binom{N}{N/2}^2\right|_{N=4} = \binom{4}{2}^2 = 36\,.$$

Thus, with very little calculation, the initial state space of dimension $2^8 = 256$ can be reduced to size of 36, with all the states found easily and given explicitly (with inverse lexicographic ordering) in Table 2.2.

State index	Statevector	State index	Statevector
$\|\tilde{0}\rangle$	$\|1100\rangle\|1100\rangle$	$\|\widetilde{18}\rangle$	$\|0110\rangle\|1100\rangle$
$\|\tilde{1}\rangle$	$\|1100\rangle\|1010\rangle$	$\|\widetilde{19}\rangle$	$\|0110\rangle\|1010\rangle$
$\|\tilde{2}\rangle$	$\|1100\rangle\|1001\rangle$	$\|\widetilde{20}\rangle$	$\|0110\rangle\|1001\rangle$
$\|\tilde{3}\rangle$	$\|1100\rangle\|0110\rangle$	$\|\widetilde{21}\rangle$	$\|0110\rangle\|0110\rangle$
$\|\tilde{4}\rangle$	$\|1100\rangle\|0101\rangle$	$\|\widetilde{22}\rangle$	$\|0110\rangle\|0101\rangle$
$\|\tilde{5}\rangle$	$\|1100\rangle\|0011\rangle$	$\|\widetilde{23}\rangle$	$\|0110\rangle\|0011\rangle$
$\|\tilde{6}\rangle$	$\|1010\rangle\|1100\rangle$	$\|\widetilde{24}\rangle$	$\|0101\rangle\|1100\rangle$
$\|\tilde{7}\rangle$	$\|1010\rangle\|1010\rangle$	$\|\widetilde{25}\rangle$	$\|0101\rangle\|1010\rangle$
$\|\tilde{8}\rangle$	$\|1010\rangle\|1001\rangle$	$\|\widetilde{26}\rangle$	$\|0101\rangle\|1001\rangle$
$\|\tilde{9}\rangle$	$\|1010\rangle\|0110\rangle$	$\|\widetilde{27}\rangle$	$\|0101\rangle\|0110\rangle$
$\|\widetilde{10}\rangle$	$\|1010\rangle\|0101\rangle$	$\|\widetilde{28}\rangle$	$\|0101\rangle\|0101\rangle$
$\|\widetilde{11}\rangle$	$\|1010\rangle\|0011\rangle$	$\|\widetilde{29}\rangle$	$\|0101\rangle\|0011\rangle$
$\|\widetilde{12}\rangle$	$\|1001\rangle\|1100\rangle$	$\|\widetilde{30}\rangle$	$\|0011\rangle\|1100\rangle$
$\|\widetilde{13}\rangle$	$\|1001\rangle\|1010\rangle$	$\|\widetilde{31}\rangle$	$\|0011\rangle\|1010\rangle$
$\|\widetilde{14}\rangle$	$\|1001\rangle\|1001\rangle$	$\|\widetilde{32}\rangle$	$\|0011\rangle\|1001\rangle$
$\|\widetilde{15}\rangle$	$\|1001\rangle\|0110\rangle$	$\|\widetilde{33}\rangle$	$\|0011\rangle\|0110\rangle$
$\|\widetilde{16}\rangle$	$\|1001\rangle\|0101\rangle$	$\|\widetilde{34}\rangle$	$\|0011\rangle\|0101\rangle$
$\|\widetilde{17}\rangle$	$\|1001\rangle\|0011\rangle$	$\|\widetilde{35}\rangle$	$\|0011\rangle\|0011\rangle$

Table 2.2: Base states of half filling zero spin projection sector for $N = 4$. First block of four states corresponds to up spins on each of the four sites, the second to down spins.

For example, state $|\tilde{8}\rangle$ corresponds to two particles on the first site, one spin up particle on the third site and one spin down particle on the fourth site:

$$|1010\rangle\,|1001\rangle = \mathbf{c}^{\dagger}_{1\uparrow}\mathbf{c}^{\dagger}_{3\uparrow}\mathbf{c}^{\dagger}_{1\downarrow}\mathbf{c}^{\dagger}_{4\downarrow}\,|0\rangle\,.$$

Of course, we will use Lieb's and Shen's theorems to further reduce this subspace into smaller nonoverlapping blocks, however this process will be more involved than a simple combinatorial construction, as shown in Section 2.5. Before this, we will handle some explicitly solvable limiting cases to gain a better understanding of the model and to obtain results that can be compared to later calculations.

Finally, let us emphasize that although these two theorems are only applicable in the limiting $t' = 0$ case, we find that even for $t' \neq 0$ at half filling, the spin projection sector

containing the ground state is again a zero spin projection sector $\mathbf{S}^z = 0$ [1], and thus we focus our attention to it in both cases.

2.4 Exactly solvable limiting cases

As explained in Intermezzo 1.2, we wish to solve the Fermi-Hubbard model by finding its exact eigenstates and the corresponding eigenenergies. Having found a convenient system basis explicitly, let us consider limiting behaviour of the system with $N = 4$ sites at half filling.

$t = t' = 0$ exact regime

Intuitively, this case corresponds to no possibility of hopping between different sites. Effectively, this decouples the different sites and the Hamiltonian is simply a sum of one-site terms

$$\mathbf{H} = \sum_{j=1}^{4} U \mathbf{n}_{j\uparrow} \mathbf{n}_{j\downarrow} =: \sum_{j=1}^{4} \mathbf{H}_j \,.$$

This implies that each local Hamiltonian can be diagonalized in subspace of its corresponding site, but keeping the total number of spin up and spin down particles consistent with half filling when combining these into the total system state. We do this in the single site occupation number basis, with elements $\{|0\rangle, |1\rangle, |2\rangle\}$ based on how many particles are present at said site, immediately observing that since $U > 0$, the lowest eigenvalue of 0 is obtained for $|0\rangle$ or one of the two vectors in subspace of $|1\rangle$. Thus, since the only possible contributions to the energy are the positive U terms, the favourable configurations are those in which there are no particles on the same site, i.e. with zero doubly occupied sites.

To see this in a different algebraic way, we recall that the Hamiltonian term on a particular site connects only the states for which the occupations on that particular site are identical in the total basis 2.2, which can be seen as follows. To recall, in equation (2.3), only the U term is nonzero. This term counts the number of doubly occupied sites, so the only way for it to not vanish is for some sites to be doubly occupied. However, even in that case, this contribution will vanish for different base states. For example, let us consider states $|\widetilde{0}\rangle$ and $|\widetilde{1}\rangle$ from Table 2.2. Although the $j = 1$ contribution will yield a constant factor of 1 for the first site, the states of the remaining three sites represent different basis

states of three possible particles and thus do not overlap, yielding a total of zero. Explicitly:

$$\begin{aligned}\langle\tilde{0}|\mathbf{H}|\tilde{1}\rangle &= \langle\tilde{0}|U\mathbf{n}_{1\uparrow}\mathbf{n}_{1\downarrow}|\tilde{1}\rangle \\ &= U\,\langle 1100|\,\langle 1100|\,\mathbf{c}^{\dagger}_{1\uparrow}\mathbf{c}_{1\uparrow}\mathbf{c}^{\dagger}_{1\downarrow}\mathbf{c}_{1\downarrow}\,|1100\rangle\,|1010\rangle \\ &= U\,\langle 1100|\,\mathbf{c}^{\dagger}_{1\uparrow}\mathbf{c}_{1\uparrow}\,|1100\rangle\,\langle 1100|\,\mathbf{c}^{\dagger}_{1\downarrow}\mathbf{c}_{1\downarrow}\,|1010\rangle \\ &= U\cdot 1\cdot\underbrace{\langle 1100|1010\rangle}_{0} \\ &= 0\,.\end{aligned}$$

Thus, there remain only (some of) the diagonal terms in the Hamiltonian, particularly those connecting states that have zero, one or two doubly occupied sites, yielding Hamitlonian values of 0, U and $2U$, respectively. Regarding the zero doubly occupied sites states, there will be $\binom{4}{2} = 6$ of them since once we pick the two occupied spin up sites, the two spin down are determined. Consulting 2.2, we see that these six states are $|\tilde{5}\rangle$, $|\widetilde{10}\rangle$, $|\widetilde{15}\rangle$, $|\widetilde{20}\rangle$, $|\widetilde{25}\rangle$, $|\widetilde{30}\rangle$. Similarly, we can choose the two sites to be doubly occupied in 6 ways and those states are $|\tilde{0}\rangle$, $|\tilde{7}\rangle$, $|\widetilde{14}\rangle$, $|\widetilde{21}\rangle$, $|\widetilde{28}\rangle$, $|\widetilde{35}\rangle$. Evidently, the only remaining possibility is one doubly occupied site and all the remaining $36 - 6 - 6 = 24$ states will be of that kind. Combinatorially, there are 4 ways to choose the single doubly occupied site and then $\binom{3}{2} = 3$ to choose the two singly occupied sites, along with an additional factor of 2 corresponding to ordering of their two spins (which are necessarily different since the two spins on the doubly occupied site are different as well). Therefore, the matrix is diagonal with the values (in order) given in Table 2.3. The ground state energy is 0 and the ground state manifold is spanned by the six vectors with zero doubly occupied sites.

$2U$	U	U	U	U	0	U	$2U$	U	U	0	U	U	U	$2U$	0	U	U
U	U	0	$2U$	U	U	U	0	U	U	$2U$	U	0	U	U	U	U	$2U$

Table 2.3: Diagonal elements of Hamiltonian at half filling zero spin projection sector for $N = 4$ and $t, t' = 0$.

Hamiltonian in the momentum space

We now want to shift our attention to the opposite regime of vanishing U. Once the hopping terms t and/or t' are nonzero and the onsite interaction U is zero, the FH Hamiltonian is simple to handle in momentum space representation. Let us transform it for this purpose, of course using (1.11)—conveniently, each of the three terms in (2.3) can be handled separately.

For the t term, we find (omitting the summation bounds for readability):

$$
\begin{aligned}
&-t \sum_{j\sigma} \left(\mathbf{c}^\dagger_{j\sigma} \mathbf{c}_{(j+1)\sigma} + \mathbf{c}^\dagger_{(j+1)\sigma} \mathbf{c}_{j\sigma} \right) \rightsquigarrow \\
&\quad \rightsquigarrow -t \sum_{j\sigma} \left[\left(\frac{1}{N} \sum_{q_1} e^{\imath q_1} \mathbf{c}^\dagger_{q_1\sigma} \sum_{q_2} e^{-\imath q_2 (j+1)} \mathbf{c}_{q_2\sigma} \right) + \left(\frac{1}{N} \sum_{q_1} e^{\imath q_1 (j+1)} \mathbf{c}^\dagger_{q_1\sigma} \sum_{q_2} e^{-\imath q_2} \mathbf{c}_{q_2\sigma} \right) \right] = \\
&\quad = -\frac{t}{N} \sum_{j\sigma} \left[\sum_{q_1 q_2} e^{\imath j (q_1 - q_2)} e^{-\imath q_2} \mathbf{c}^\dagger_{q_1\sigma} \mathbf{c}_{q_2\sigma} + \sum_{q_1 q_2} e^{\imath j (q_1 - q_2)} e^{\imath q_1} \mathbf{c}^\dagger_{q_1\sigma} \mathbf{c}_{q_2\sigma} \right] = \\
&\quad = -\frac{t}{N} \sum_{q_1 q_2 \sigma} \underbrace{\sum_{j} e^{\imath j (q_1 - q_2)}}_{(1.13)} \mathbf{c}^\dagger_{q_1\sigma} \mathbf{c}_{q_2\sigma} \left[e^{-\imath q_2} + e^{\imath q_1} \right] = \\
&\quad = -\frac{t}{\cancel{N}} \sum_{q_1 q_2 \sigma} \cancel{N} \delta_{q_1 q_2} \mathbf{c}^\dagger_{q_1\sigma} \mathbf{c}_{q_2\sigma} \left[e^{-\imath q_2} + e^{\imath q_1} \right] = \\
&\quad = -t \sum_{q_1\sigma} \mathbf{c}^\dagger_{q_1\sigma} \mathbf{c}_{q_1\sigma} \underbrace{\left[e^{\imath q_1} + e^{-\imath q_2} \right]}_{2\cos q_1} = \{ q_1 \rightsquigarrow q \} \\
&\quad = \sum_{q\sigma} \left(-2t \cos q \right) \mathbf{c}^\dagger_{q\sigma} \mathbf{c}_{q\sigma} \,,
\end{aligned}
\tag{2.4}
$$

where we have used the sum over j to get Kronecker-delta and then summed it over q_2.

Similarly, for the t' term in (2.3), the process is analogous, however with two changes. First, the distance between different terms is now two sites, so the $\cos q$ term from (2.4) becomes $\cos 2q$. Second, range of t' (2) is equal to half of the size of the system ($4/2 = 2$) so there is an additional corrective factor of 2 to divide with, yielding:

$$
-\frac{t'}{2} \sum_{j\sigma} \left(\mathbf{c}^\dagger_{j\sigma} \mathbf{c}_{(j+2)\sigma} + \mathbf{c}^\dagger_{(j+2)\sigma} \mathbf{c}_{j\sigma} \right) \rightsquigarrow \sum_{q\sigma} \left(-t' \cos 2q \right) \mathbf{c}^\dagger_{q\sigma} \mathbf{c}_{q\sigma} \,. \tag{2.5}
$$

For the final U term in (2.3), we find:

$$
\begin{aligned}
&U \sum_j \mathbf{n}_{j\uparrow} \mathbf{n}_{j\downarrow} = U \sum_j \mathbf{c}^\dagger_{j\uparrow} \mathbf{c}_{j\uparrow} \mathbf{c}^\dagger_{j\downarrow} \mathbf{c}_{j\downarrow} \rightsquigarrow \\
&\quad \rightsquigarrow U \sum_j \frac{1}{N^{4/2}} \sum_{q_1} e^{\imath q_1 j} \mathbf{c}^\dagger_{q_1\uparrow} \sum_{q_2} e^{-\imath q_2 j} \mathbf{c}_{q_2\uparrow} \sum_{q_3} e^{\imath q_3 j} \mathbf{c}^\dagger_{q_3\downarrow} \sum_{q_4} e^{-\imath q_4 j} \mathbf{c}_{q_4\downarrow} = \\
&\quad = \frac{U}{N^2} \sum_{q_1 q_2 q_3 q_4} \underbrace{\sum_j e^{\imath j (q_1 - q_2 + q_3 - q_4)}}_{(1.13)} \mathbf{c}^\dagger_{q_1\uparrow} \mathbf{c}_{q_2\uparrow} \mathbf{c}^\dagger_{q_3\downarrow} \mathbf{c}_{q_4\downarrow} = \\
&\quad = \frac{U}{N} \sum_{q_1 q_2 q_3} \mathbf{c}^\dagger_{q_1\uparrow} \mathbf{c}_{q_2\uparrow} \mathbf{c}^\dagger_{q_3\downarrow} \mathbf{c}_{(q_1 - q_2 + q_3)\downarrow}
\end{aligned}
\tag{2.6}
$$

and it is evidently of the most complex form. In comparison, the other two terms (2.4) and (2.5) are block diagonal, with each 2×2 block corresponding to a particular momentum q.

Now that the momentum representation is obtained, we can handle the limiting cases with zero and small U.

$U = t' = 0$ exact regime

Evidently, this case corresponds to no onsite interaction and no hopping to nonadjacent sites. Effectively, this removes the diagonal terms of the Hamiltonian, but results in a more complicated structure than 2.3, as can be seen in 36×36 scheme in Figure 2.2, with white cells corresponding to zero terms. For example, the upper rightmost orange line of interactions corresponds to basis states from 2.2 that are 24 states apart. Specifically, the first interaction that is a part of that line, i.e. the one in the first row is between states $|\tilde{0}\rangle = |1100\rangle\,|1100\rangle$ and $|\widetilde{24}\rangle = |0101\rangle\,|1100\rangle$ from which we can see that said term corresponds to spin up particle on site 4 hopping to site 1.

The issue with this matrix is evident. Even for such a small system of $N = 4$ sites, the Hamiltonian is still nontrivial to solve analytically: there are 192 nonzero terms in its matrix. However, the $U = 0$ condition turns this problem into a system of noninteracting fermions, and there is a more clever way to solve it instead of brute force diagonalization. Expectedly, this is due to the block-diagonal nature of the noninteracting Hamiltonian, in this regime including only the (2.4) term:

$$\mathbf{H} = \sum_{q\sigma} \left(-2t\cos q\right) \mathbf{c}^{\dagger}_{q\sigma}\mathbf{c}_{q\sigma}\,. \tag{2.7}$$

Of course for $N = 4$, there will be only four possible excitations in the momentum space given by (1.12)[5]:

$$\begin{aligned} q &= 0 && \rightsquigarrow -2t\,,\\ q &= \frac{2\pi}{4} = \frac{\pi}{2} && \rightsquigarrow 0\,,\\ q &= \frac{4\pi}{4} = \pi && \rightsquigarrow 2t\,,\\ q &= \frac{6\pi}{4} = \frac{3\pi}{2} && \rightsquigarrow 0\,. \end{aligned}$$

It is from these last four equations that we find that excitations with momentum $q = 0$ are of the lowest energy and there will be only four excitations (out of possible eight, two for

[5]In momentum space representation, excitations are momentum modes, i.e. different than regular particles, however due to the definition of DFT (1.11), their number is the same. This can be seen easily by using the transform on the total number operator $\sum_{j\sigma}\mathbf{n}_{j\sigma} = \sum_{j\sigma}\mathbf{c}^{\dagger}_{j\sigma}\mathbf{c}_{j\sigma}$, which simply turns it into summation of the same form, only over the N momentum modes instead of spin sites $\sum_{q\sigma}\mathbf{c}^{\dagger}_{q\sigma}\mathbf{c}_{q\sigma}$.

each mode/site) at half filling, so any ground state will contain two of (one of each spin) the $q = 0$ modes and then two with any of momenta with second lowest energy, e.g. $q = \pi/2$ or $q = 3\pi/2$. Thus, we find that the total energy of the ground state of the system is equal to $-2t - 2t + 0 + 0 = -4t$. *A priori* there are six such states in total since two excitations ought to be of momentum $q = 0$ and then the remaining two particles ought to be in one of the four states with momenta $q \in \{\pi/2, 3\pi/2\}$, with $\binom{4}{2} = 6$ distinct possibilities, for a total degeneracy of the ground state equal to $1 \cdot 6 = 6$. This result does not violate the uniqueness in Lieb's theorem 1 because the assumption $U > 0$ is not satisfied.

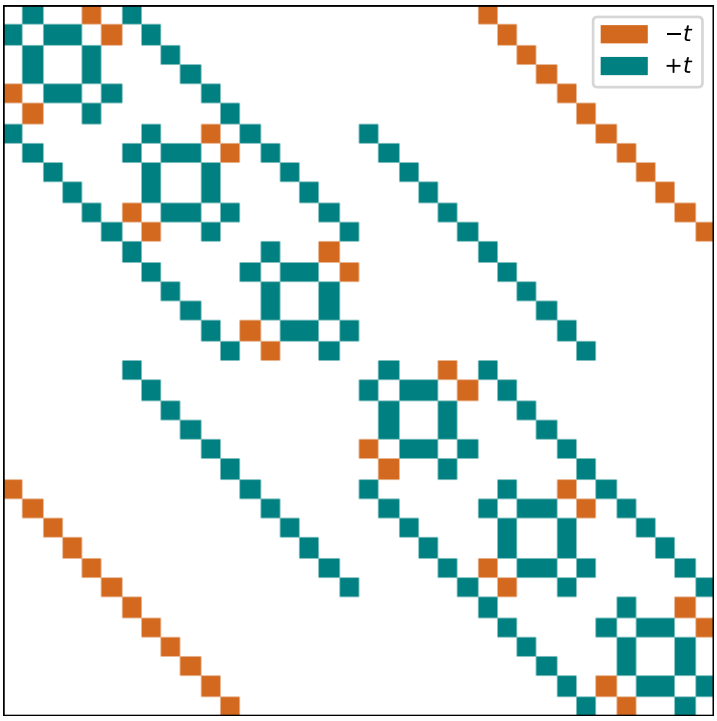


Figure 2.2: Graphical representation of the Hamiltonian terms in the $U = t' = 0$ case, on image of dimension 36×36, with white cells corresponding to zero terms.

$U = 0$ exact regime

This regime corresponds to one similar to the previous subsection, only with more intricate hopping interactions—in this case, one can find 288 nonzero terms in the Hamiltonian in Figure 2.3, rendering it even more difficult to solve via direct numerical diagonalization than Figure 2.2. The $\pm t$ terms are equal to before, while the extra matrix elements

correspond to the $\pm t'$ hopping[6]. For example, the upper rightmost red line corresponds to interactions of basis states that are 18 states apart. For example, in the second row that interaction connects the states $|\tilde{1}\rangle = |1100\rangle\,|1010\rangle$ and $|\widetilde{19}\rangle = |0110\rangle\,|1010\rangle$ from which we can see that it causes the hopping of spin up particle on site 3 to site 1.

In this regime, the Hamiltonian in the momentum space is a sum of (2.4) and (2.5):

$$\mathbf{H} = \sum_{q\sigma} \left(-2t\cos q - t'\cos 2q\right) \mathbf{c}^{\dagger}_{q\sigma}\mathbf{c}_{q\sigma}\,. \tag{2.8}$$

As (2.7), this Hamiltonian is block-diagonal as well, with the same q range for $N = 4$, determined by (1.12), so the energies of all four possible excitations are:

$$\begin{aligned}
q &= 0 && \rightsquigarrow -2t - t'\,,\\
q &= \frac{2\pi}{4} = \frac{\pi}{2} && \rightsquigarrow t'\,,\\
q &= \frac{4\pi}{4} = \pi && \rightsquigarrow 2t - t'\,,\\
q &= \frac{6\pi}{4} = \frac{3\pi}{2} && \rightsquigarrow t'\,.
\end{aligned}$$

To determine the ground states, we proceed analogously to the previous section. However, in this case the degeneracies and the populated momenta depend on the relative sizes of t and t'. If $t' < t$, then the additional populated momenta are as before and the ground state energy is $(-2t - t') + (-2t - t') + t' + t' = -4t$, again with degeneracy of four. On the other hand, if $t' > t$, the populated momentum will be $q = \pi$, with two particles, yielding a total ground state energy $(-2t - t') + (-2t - t') + (2t - t') + (2t - t') = -4t'$. Evidently, since there ought to be two particles in each of the two momenta, there is only one ground state configuration, i.e. there is no degeneracy in this case. Finally, if $t = t'$, then there will once again be two particles with momentum $q = 0$ and each with energy $-2t - t' = -3t$. However, now the remaining three momenta are all of the same energy $t' = 2t - t' = t'$. Keeping in mind that the remaining two particles ought to be of opposite spin, we see that, e.g. spin up particle can have any of the three momenta and that the same holds for the spin down particle, yielding a total ground state degeneracy of $3 \cdot 3 = 9$. The total energy in this case is $-3t - 3t + t + t = -4t$.

Of course, since t' corresponds to hoppings that are further apart than those corresponding to t, it is physically expected that it is of smaller absolute value than t.

[6]Note that there cannot be any overlap between the t and t' interactions, i.e. they cannot both be included in the same matrix element.

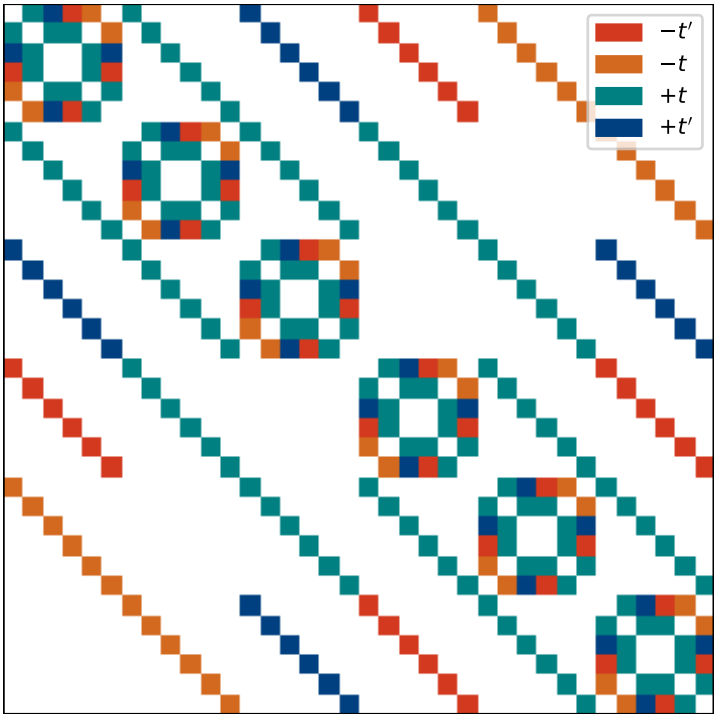


Figure 2.3: Graphical representation of the Hamiltonian terms in the $U = 0$, $t' \neq 0$ case, on image of dimension 36×36, with white cells corresponding to zero terms.

$U \to 0^+$ limit regime

It can be checked that near $U = 0$, out of the four ground states in the more physically reasonable $|t'| < |t|$ case, there is only one that persists as a true ground state as the small U interaction is turned on (when $t' = 0$, this fact is consistent with Lieb's theorem 1 implication that the ground state is unique for $U > 0$). This is a linear combination of states with $\pm\pi/2$ modes populated, of course along with the minimal 0 modes:

$$|\Psi\rangle = \frac{1}{\sqrt{2}}\left[\mathbf{c}^\dagger_{\pi/2\uparrow}\mathbf{c}^\dagger_{\pi/2\downarrow} - \mathbf{c}^\dagger_{-\pi/2\uparrow}\mathbf{c}^\dagger_{-\pi/2\downarrow}\right]\mathbf{c}^\dagger_{0\uparrow}\mathbf{c}^\dagger_{0\downarrow}|0\rangle \tag{2.9}$$

Of course, due to this state being a part of the ground state manifold in both the noninteracting and weakly interacting case, there is a continuous way of transferring to nonzero (but still small) U, and thus the effect of tuning up U is perturbative in nature. It is straightforward to apply the standard quantum perturbation theory, e.g. for the first order correction

we calculate the expectation value of(2.6) [7]:

$$\langle\Psi|\frac{U}{4}\sum_{q_1q_2q_3}\mathbf{c}^\dagger_{q_1\uparrow}\mathbf{c}_{q_2\uparrow}\mathbf{c}^\dagger_{q_3\downarrow}\mathbf{c}_{(q_1-q_2+q_3)\downarrow}|\Psi\rangle\ .$$

Although there are many terms in previous expectation value, we can easily observe that most of them vanish. Evidently, in state $|\Psi\rangle$ there are always both spin up and spin down particles of same momenta. Thus, no terms which transfer the state into a configuration that is not of this kind can have nonzero contribution. Consequently, the momenta in pairs will be either $(0,0)$, $(\pi/2,\pi/2)$ or $(-\pi/2,-\pi/2)$. As before, due to the exclusion principle, only two of the momenta can be of the same nonzero mode. Bearing all of this in mind, we see that the only nonvanishing terms are for the choices of momenta given in Table 2.4.

For example, we can calculate the expected value for the first row of said table:

$$\begin{aligned}
\langle\Psi|\frac{U}{4}\mathbf{c}^\dagger_{0\uparrow}\mathbf{c}_{0\uparrow}\mathbf{c}^\dagger_{0\downarrow}\mathbf{c}_{0\downarrow}|\Psi\rangle &\overset{(2.9)}{=}\frac{U}{4}\langle\Psi|\,\mathbf{c}^\dagger_{0\uparrow}\mathbf{c}_{0\uparrow}\mathbf{c}^\dagger_{0\downarrow}\mathbf{c}_{0\downarrow}\frac{1}{\sqrt{2}}\left[\mathbf{c}^\dagger_{\pi/2\uparrow}\mathbf{c}^\dagger_{\pi/2\downarrow}-\mathbf{c}^\dagger_{-\pi/2\uparrow}\mathbf{c}^\dagger_{-\pi/2\downarrow}\right]\mathbf{c}^\dagger_{0\uparrow}\mathbf{c}^\dagger_{0\downarrow}\,|0\rangle = \\
&=\frac{U}{4\sqrt{2}}\langle\Psi|\,\mathbf{c}^\dagger_{0\uparrow}\mathbf{c}_{0\uparrow}\mathbf{c}^\dagger_{0\downarrow}\mathbf{c}_{0\downarrow}\left[\mathbf{c}^\dagger_{\pi/2\uparrow}\mathbf{c}^\dagger_{\pi/2\downarrow}-\mathbf{c}^\dagger_{-\pi/2\uparrow}\mathbf{c}^\dagger_{-\pi/2\downarrow}\right]\mathbf{c}^\dagger_{0\uparrow}\mathbf{c}^\dagger_{0\downarrow}\,|0\rangle = \\
&=-\frac{U}{4\sqrt{2}}\langle\Psi|\,\mathbf{c}^\dagger_{0\uparrow}\mathbf{c}_{0\uparrow}\mathbf{c}^\dagger_{0\downarrow}\left[\mathbf{c}^\dagger_{\pi/2\uparrow}\mathbf{c}^\dagger_{\pi/2\downarrow}-\mathbf{c}^\dagger_{-\pi/2\uparrow}\mathbf{c}^\dagger_{-\pi/2\downarrow}\right]\mathbf{c}^\dagger_{0\uparrow}\mathbf{c}_{0\downarrow}\mathbf{c}^\dagger_{0\downarrow}\,|0\rangle = \\
&=\frac{U}{4\sqrt{2}}\langle\Psi|\,\mathbf{c}^\dagger_{0\uparrow}\mathbf{c}^\dagger_{0\downarrow}\left[\mathbf{c}^\dagger_{\pi/2\uparrow}\mathbf{c}^\dagger_{\pi/2\downarrow}-\mathbf{c}^\dagger_{-\pi/2\uparrow}\mathbf{c}^\dagger_{-\pi/2\downarrow}\right]\underbrace{\mathbf{c}_{0\uparrow}\mathbf{c}^\dagger_{0\uparrow}}_{\mathbb{1}-\mathbf{n}_{0\uparrow}}\underbrace{\mathbf{c}_{0\downarrow}\mathbf{c}^\dagger_{0\downarrow}}_{\mathbb{1}-\mathbf{n}_{0\downarrow}}|0\rangle = \\
&=\frac{U}{4\sqrt{2}}\langle\Psi|\,\mathbf{c}^\dagger_{0\uparrow}\mathbf{c}^\dagger_{0\downarrow}\left[\mathbf{c}^\dagger_{\pi/2\uparrow}\mathbf{c}^\dagger_{\pi/2\downarrow}-\mathbf{c}^\dagger_{-\pi/2\uparrow}\mathbf{c}^\dagger_{-\pi/2\downarrow}\right]|0\rangle = \\
&=\frac{U}{4\sqrt{2}}\langle\Psi|\,\mathbf{c}^\dagger_{0\uparrow}\mathbf{c}^\dagger_{0\downarrow}\left[\mathbf{c}^\dagger_{\pi/2\uparrow}\mathbf{c}^\dagger_{\pi/2\downarrow}-\mathbf{c}^\dagger_{-\pi/2\uparrow}\mathbf{c}^\dagger_{-\pi/2\downarrow}\right]|0\rangle = \\
&=\frac{U}{4}\langle\Psi|\,\frac{1}{\sqrt{2}}\left[\mathbf{c}^\dagger_{\pi/2\uparrow}\mathbf{c}^\dagger_{\pi/2\downarrow}-\mathbf{c}^\dagger_{-\pi/2\uparrow}\mathbf{c}^\dagger_{-\pi/2\downarrow}\right]\mathbf{c}^\dagger_{0\uparrow}\mathbf{c}^\dagger_{0\downarrow}\,|0\rangle = \\
&=\frac{U}{4}\langle\Psi|\Psi\rangle = \\
&=\frac{U}{4}\,,
\end{aligned}$$

where we have repeatedly used fermionic anticommutation relation (1.7) and the fact that the number of particles of any kind in vacuum state is zero.

[7] Interestingly, this state transforms according to the B_1 irreducible representation from Table 2.1 of C_{4v}. It can be demonstrated that no noninteracting model on four sites with periodic boundary conditions can have this property; i.e. this is an intrinsically new effect of strong correlations [46, 47].

The remaining four terms in Table 2.4 each contribute half of $U/4$ and thus the total first order correction is $3U/4$. The second order calculation is a bit more tedious to carry out, so we simply report that the additional correction term is $-(13/128)U^2$. The visualization of this perturbative result is given in Figure 2.4. Expectedly, as the onsite interaction strength increases, the unperturbed case $U = 0$ that corresponds to a system of free fermions is wildly off in energy calculation. On the other hand, the first order correction is precise for smaller values of U, but second order correction is sufficiently precise for even outside the regime in which U is considered a perturbation.

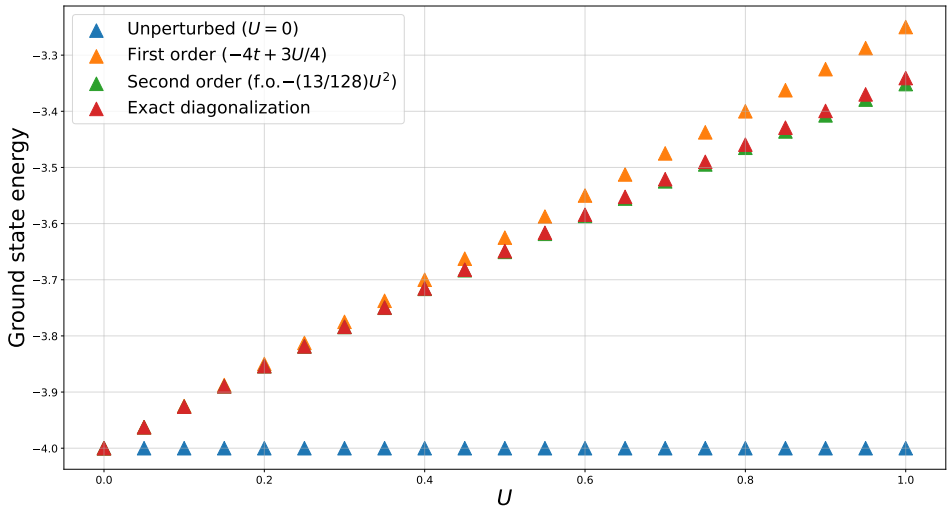


Figure 2.4: Comparison of the ground state energy of the Fermi-Hubbard model with $t' = 0$ found by perturbation theory and its exact value.

q_1	q_2	q_3	$q_4 (= q_1 - q_2 + q_3)$
0	0	0	0
0	0	$\frac{\pi}{2}$	$\frac{\pi}{2}$
0	0	$-\frac{\pi}{2}$	$-\frac{\pi}{2}$
$\frac{\pi}{2}$	$\frac{\pi}{2}$	0	0
$-\frac{\pi}{2}$	$-\frac{\pi}{2}$	0	0

Table 2.4: Nonvanishing terms in expectation value of the onsite interaction Hamiltonian in state (2.9).

2.5 General analytic approach

In scope of previous section, we have seen how different limiting cases of our model can be tackled to find their exact solutions. However, interesting physical effects are observed within systems that include all of the contributions in (2.1) [44], so it is of interest to explore approaches that do not rely on limiting behaviour and associated approximations. Of course, in the thermodynamic limit $N \to \infty$ at half filling, it is conveniently possible to use Lieb and Wu's exact solution via Bethe ansatz [48], all reduced to one integral expression. However in finite systems, solving BA equations is highly nontrivial since one usually ought to resort to numerical solutions of a system of nonlinear equations.

On the other hand, regardless of the strengths of different terms in the Hamiltonian (2.3), if we fix the number of sites to 4 and retain the model structure, its symmetries will not change with changing interactions. We now explore the powerful machinery resting on different symmetries of the particular model we are handling.

Complexity reduction via lattice symmetry

Exploiting the C_{4v} symmetry to reduce the complexity of our model via representation theory is useful for at least three reasons. Firstly, it allows us to find solutions to the model in a computationally feasible way, with values that can be compared to solutions found using VQE later on. Second, since the only requirement for the method to be applicable is the underlying symmetry group, we can use it in a completely analogous form for different physical models on the same lattice. Third, the method can be expanded to different lattice structures, by changing the symmetry group and carrying out the same steps.

As before, we restrict our analysis to the half filling with zero total spin projection for the moment, i.e. the $\mathbf{S}^z = \mathbf{R}^z = 0$ subspace of dimension 36, with basis given in Table 2.2. How to use group theory to further separate the Hamiltonian in this block into smaller blocks?

We recall that the five irreps of group C_{4v} are given in Table 2.1. Let us denote a generic operation from this group by R and its ith irrep by Γ_i. Then the algorithmic procedure for reducing the Hamiltonian to block form is as follows [49][8].

First, since we have restricted our consideration to a particular subspace of size 36×36, let us denote the states from Table 2.2 by $\left\{|\tilde{j}\rangle\right\}_{j\in[1\cdot\cdot36]}$. Then, we can easily find matrix elements for every operation R from in the $\left\{|\tilde{j}\rangle\right\}$ basis, simply by acting on the basis states. We denote these matrices by $\tilde{R}_{mn}$ and then construct the corresponding projectors

$$P^{\Gamma_i}_{mn} = \sum_R \mathrm{tr}\left(\Gamma^i_R\right) R_{mn}\,,$$

[8] There are many examples of succesfully using group theory to solve variants of Hubbard model exactly [50, 51, 52, 53].

where $\mathrm{tr}\left(\Gamma_R^i\right)$ is the entry in character table corresponding to operator R and ith irrep.

Once all of the projectors are found, we apply them to the initial basis states $\left\{|\tilde{j}\rangle\right\}$ to construct linear combinations that possess definite symmetries and transform according to different representations:

$$|m,i\rangle = \sum_j P^i_{mj}\,|\tilde{j}\rangle \ ,$$

with i corresponding to the irrep index and m to one of 36 states when projecting onto that representation. In this way, the 36×36 relevant subspace is reduced into blocks, the largest of which is of size 5×5. The code associated with this calculation is implemented in full and can be found in the repository, particularly the `group_theory_diagonalization.py` script and the results can be compared to [54].

Explicitly diagonalizing all of these matrices, we find energies that agree with the exact ones for several values of parameters U, t and t'. However, there is a similar approach to Hamiltonian reduction using symmetries beyond the C_{4v} group, so we turn our attention to that approach, since it will result in slightly smaller final block sizes.

Complexity reduction via general symmetries

As in the previous section, the accompanying code to this one can be found in the repository, in `general_symmetry_reduction.py` script. In it, all of the steps about to be described are implemented from scratch and can be used to find any of the intermediate block-diagonal forms of the Hamiltonian. Furthermore, this approach is useful for several reasons as well. Besides the expected ability to check the solutions found with VQE, we can also apply this approach to different lattices and physical models and reduce their complexity, if they are of same symmetries, regardless of size.

Let us note that at a certain point in the following calculation, we will set $t' = 0$ since in this case, the Hamiltonian commutes with the total pseudospin operator (as discussed in Subsection 2.1) and allows us to continue the reduction process up until to analytically diagonalizable blocks. To reduce to blocks of similar size when $t' = 0$, we would have to construct more complex operators, which are not necessary for present discussion and illustrate no computational differences, and are thus omitted.

To proceed with reduction, let us note that since we are able to explicitly write the spin projection operator $\mathbf{S}^z$, as was done in (1.14), we can write it in matrix form for the four site system we are considering. First, it is easy to see that possible eigenvalues of this operator are in range from -2 to $+2$ in steps of $1/2$—there are nine of them. This is due to the fact that there can be at most four particles of same spin, one for each of four sites, due to the Pauli exclusion principle. Evidently, since $\mathbf{S}^z$ commutes with $\mathbf{H}$, i.e. Hamiltonian preserves the spin projection, there can be no mixing between subspaces with different $\mathbf{S}^z$ eigenvalues, because it would imply that $\mathbf{H}$ changed the total spin projection. Thus, we

wish to diagonalize this operator and use its eigenvectors as a new basis in which the initial Hamiltonian will be block diagonal (albeit with somewhat large blocks initially) with each block corresponding to a single eigenvalue of $\mathbf{S}^z$. Particularly, let us denote the eigenvalues of $\mathbf{S}^z$ by $\{v_i\}_{i\in[1\cdot\cdot256]}$, but with eigenstates ordered by their eigenvalue in increasing order. Then, all we need to do is transform the Hamiltonian into this basis via

$$\mathbf{H} \rightsquigarrow \begin{bmatrix} v_1^{*\top} & v_2^{*\top} & \dots & v_{256}^{*\top} \end{bmatrix}^{-1} \mathbf{H} \begin{bmatrix} v_1^{\top} & v_2^{\top} & \dots & v_{256}^{\top} \end{bmatrix}. \tag{2.10}$$

In this way, the basis states with same $\mathbf{S}^z$ eigenvalue are adjacent (because the eigenvectors are ordered by their eigenvalues), so the Hamiltonian is expectedly separated into nine blocks, one for each eigenvalue of $\mathbf{S}^z$. Each block size corresponds to the number of states with a particular eigenvalue and the blocks are of dimension $n \times n$, with $n \in (1, 8, 28, 56, 70, 56, 28, 8, 1)$ corresponding to spin projection values $\mathbf{S}^z \in (-2, -3/2, -1, -1/2, 0, 1/2, 1, 3/2, 2)$, respectively, as can be seen in code explicitly. At this moment, we simply recall Lieb's theorem 1 and focus on the block of size 70×70, corresponding to the value $\mathbf{S}^z = 0$ since we know the ground state is contained in that subspace[9]. Thus, we go on to reduce this particular block further.

The next operators to use are, in order, $\mathbf{R}^z$, $\mathbf{S}^2$, $\mathbf{R}^2$ and the translation operator $\mathbf{T}$, but before proceeding, we note that have we started with $\mathbf{R}^z$ instead of $\mathbf{S}^z$, we would get the same eigenvalues as for $\mathbf{S}^z$ and therefore block sizes as for $\mathbf{S}^z$, which is expected since the total spin projection is analogous to the number of particles, however centered around half filling as zero. Let us note that we could have carried the reductions out in this different order, both of these and the subsequent operators, but it is most convenient to begin with these two for several reasons. First, as mentioned before, $\mathbf{S}^z$ and $\mathbf{R}^z$ can be determined for any given state in basis such as 2.2 (but for the total Hilbert space) immediately by visual inspection. Second, when starting with operators such as $\mathbf{S}^2$ or $\mathbf{R}^2$, the largest block in both cases is of size 96×96 (and the second largest of size 81×81), which increases the initial complexity—with our schema, the current maximal block is always of the minimal possible size. Of course, if we use the same set of operators, the final block sizes ought to be identical.

Some of the previous attempts of symmetry reduction were slightly less efficient, however still carried out until all of the solutions were found analytically. For example, Villet and Steeb used the spin symmetry, followed by a C_{4v} symmetry, resulting in the largest block in the final Hamiltonian of size 7×7 [55] (evidently larger than we will find). Noce and Cuoco used both the spin and pseudospin operators, along with the translation operator, resulting in a maximum block size of 4×4 [56]—this is a part of the approach we will take as well since the blocks of this size can be handled easily. Finally, Schumann used the operators constructed from C_{4v} group operators, as well as more complex operators found

[9] As we have already noted, since this block is the most complex, finding the solutions for it will be sufficient to solve the complete system.

by Grosse [57] and was able to reduce the Hamiltonian into blocks of dimension 1×1 [45], however we will stop at 3×3 since it is not computationally necessary to go further, and the calculations are somewhat more tedious after this point[10].

To continue the described procedure with $\mathbf{R}^z$, we recall that it commutes with $\mathbf{H}$ regardless of the NNN hopping strength t'. Following this, we easily find its eigenvalues for all 70 states in $\mathbf{S}^z = 0$ subspace and then order the basis so that these states are adjacent. From technical point of view, we transform the Hamiltonian back to initial basis, diagonalize $\mathbf{R}^z$ in the desired subspace and change the order of basis elements in said subspace, and then transform the Hamiltonian back again (we are doing these Hamiltonian transformations repeatedly so that we are able to check the expected block sizes and Hamiltonian elements at each step). This is easy to do in this case since the eigenstates of $\mathbf{R}^z$ are the same as those of $\mathbf{S}^z$, so a simple reordering of the basis states is sufficient. In this way, the 70×70 subspace is split into five blocks of sizes $n \times n$ and $n \in (1, 16, 36, 16, 1)$, with corresponding eigenvalues of $\mathbf{R}^z$ equal to $\mathbf{R}^z \in (-2, -1, 0, 1, 2)$, respectively. The block of interest, i.e. $\mathbf{R}^z = 0$ is, expectedly, the one with size 36×36, which is the dimension we have found before for half filling and zero total spin projection via combinatorial arguments in 2.3.

The next operator that commutes with $\mathbf{H}$ we will use is $\mathbf{S}^2$. First, we emphasize once more that the basis states of set $\{\mathbf{S}^z, \mathbf{R}^z\}$ we have naturally chosen before (because their eigenvalues are visible by inspection) are not eigenstates of $\mathbf{S}^2$. This is of course related to the fact that single particle spin states usually combine in nontrivial ways to construct particles of well-defined total spin (e.g. singlets, doublets, triplets, etc.). However, due to commutation with the Hamiltonian, they still span the same vector space, so we simply ought to take the previously obtained 36 basis states of the $\mathbf{S}^z = \mathbf{R}^z = 0$ subspace and diagonalize $\mathbf{S}^2$ in it. Thus, we project the $\mathbf{S}^2$ matrix onto this subspace, exactly as we did to $\mathbf{H}$ in (2.10) and diagonalize the simple 36×36 spin matrix that is obtained. It is important to note that this matrix is significantly simpler to diagonalize than the Hamiltonian projection to this subspace would be since it is sparser than the Hamiltonian and all of its terms are constant, so this diagonalization ought to be done only once and it's results are then applicable to different Hamiltonians. In this way, we find the 36 states with well-defined total spin[11], and we first note that the only eigenvalues we get are 0 ($\mathbf{S}^2 = 0$), 1 ($\mathbf{S}^2 = 2$) and 2 ($\mathbf{S}^2 = 6$), which is in agreement with the fact that we restricted the analysis to space in which spin projection is zero, which would not be possible for total spin $1/2$ or $3/2$. After we transform the subbasis so that it is an eigenbasis of $\mathbf{S}^2$ and order the basis elements in order of increasing eigenvalue, we find that the 36×36 block is split into three blocks:

[10] However, the aforementioned papers should certainly be analyzed if one wishes to obtain a complete understanding of the four-site model in a physical sense. For present work which focuses more on mathematical and computational perspective of the model, this is of no direct consequence.

[11] Note that these states also have well-defined $\mathbf{S}^z$ and $\mathbf{R}^z$ because the complete subspace corresponds to a single value of each of these operators.

one of size 20 (eigenvalue $\mathbf{S} = 0$), one of size 15 (eigenvalue $\mathbf{S} = 1$) and one of size 1 (eigenvalue $\mathbf{S} = 2$). Evidently, this last block is trivially diagonalizable since it is already diagonal with a single element and we find an eigenvalue of Hamiltonian with energy 0. Regarding the other two remaining blocks, as per Lieb's theorem 1, the relevant one is of dimension 20×20 with $\mathbf{S} = 0$.

At this moment, we would like to carry out the same procedure for operator $\mathbf{R}^2$. However, to be able to do so, the $t' = 0$ condition ought to be imposed, as it is necessary for the operator to commute with the Hamiltonian, as discussed in Subsection 2.1. The diagonalization with respect to $\mathbf{R}^2$ is carried out for the 20×20 block, and it is split into three: one of size 10×10 (eigenvalue $\mathbf{R} = 0$), one of size 9×9 (eigenvalue $\mathbf{R} = 1$) and one of size 1×1 (eigenvalue $\mathbf{R} = 2$)[12]. Once again, the final block is trivially diagonal with a single element, corresponding to energy eigenvalue of $2U$.

Finally, we arrive at the translation operator $\mathbf{T}$. At this moment, we can recall the symmetry group of the Hamiltonian, which is C_{4v}. Here, translating the system by one site equals exactly to rotation by $\pi/2$, i.e. operation of C_4, so the operator $\mathbf{T}$ is a representation of C_4. Similarly, $\mathbf{T}^2$ is a representation of C_2, $\mathbf{T}^3$ is a representation of C_4^3, while the operator $\mathbf{T}^4$ corresponds to identity, which is easily noted since shifting the system by one site four times returns it to the initial state. Evidently, due to idempotency of order 4, the eigenvalues of $\mathbf{T}$ are $\pm 1, \pm\imath$.

Applying the same procedure as before, we find that the block of size 10×10 (with eigenvalue $\mathbf{R} = 0$) is split into four blocks: one of size 3×3 (eigenvalue $\mathbf{T} = -1$), one of size 2×2 (eigenvalue $\mathbf{T} = -\imath$), one of size 2×2 (eigenvalue $\mathbf{T} = \imath$) and one of size 3×3 (eigenvalue $\mathbf{T} = 1$).

For completeness, let us mention the splitting of the remaining blocks: the block of size 9×9 is split into five blocks: one of size 3×3 (eigenvalue $\mathbf{T} = -1$), one of size 2×2 (eigenvalue $\mathbf{T} = -\imath$), one of size 2×2 (eigenvalue $\mathbf{T} = \imath$) and two of size 1×1 (eigenvalue $\mathbf{T} = 1$). The block of size 9×9 is split into five blocks as well: two of size 1×1 (eigenvalue $\mathbf{T} = -1$), one of size 2×2 (eigenvalue $\mathbf{T} = -\imath$), one of size 2×2 (eigenvalue $\mathbf{T} = \imath$) and one of size 3×3 (eigenvalue $\mathbf{T} = 1$). Finally, the block of size 6×6 is split into four blocks: one of size 1×1 (eigenvalue $\mathbf{T} = -1$), one of size 2×2 (eigenvalue $\mathbf{T} = -\imath$), one of size 2×2 (eigenvalue $\mathbf{T} = \imath$) and one of size 1×1 (eigenvalue $\mathbf{T} = 1$).

In this way, and returning to the relevant $\mathbf{S}^z = \mathbf{R}^z = \mathbf{S}^2 = \mathbf{R}^2 = 0$ subspace, we find that the Hamiltonian has been reduced into blocks that are at most of size 3×3, as can be seen in the first 10×10 subspace of the matrix in Figure 2.5.

Before we write the remaining four blocks of the 10×10 susbspace explicitly, let us note that we have not explicitly used the reflection operators of symmetry group C_{4v}. This is due to the fact that within the final four blocks, there is no distinguishing by eigenvalues

[12] For completeness, let us note that the 15×15 block is split into two blocks, one of size 9×9 (eigenvalue $\mathbf{R} = 0$) and one of size 6 (eigenvalue $\mathbf{R} = 1$).

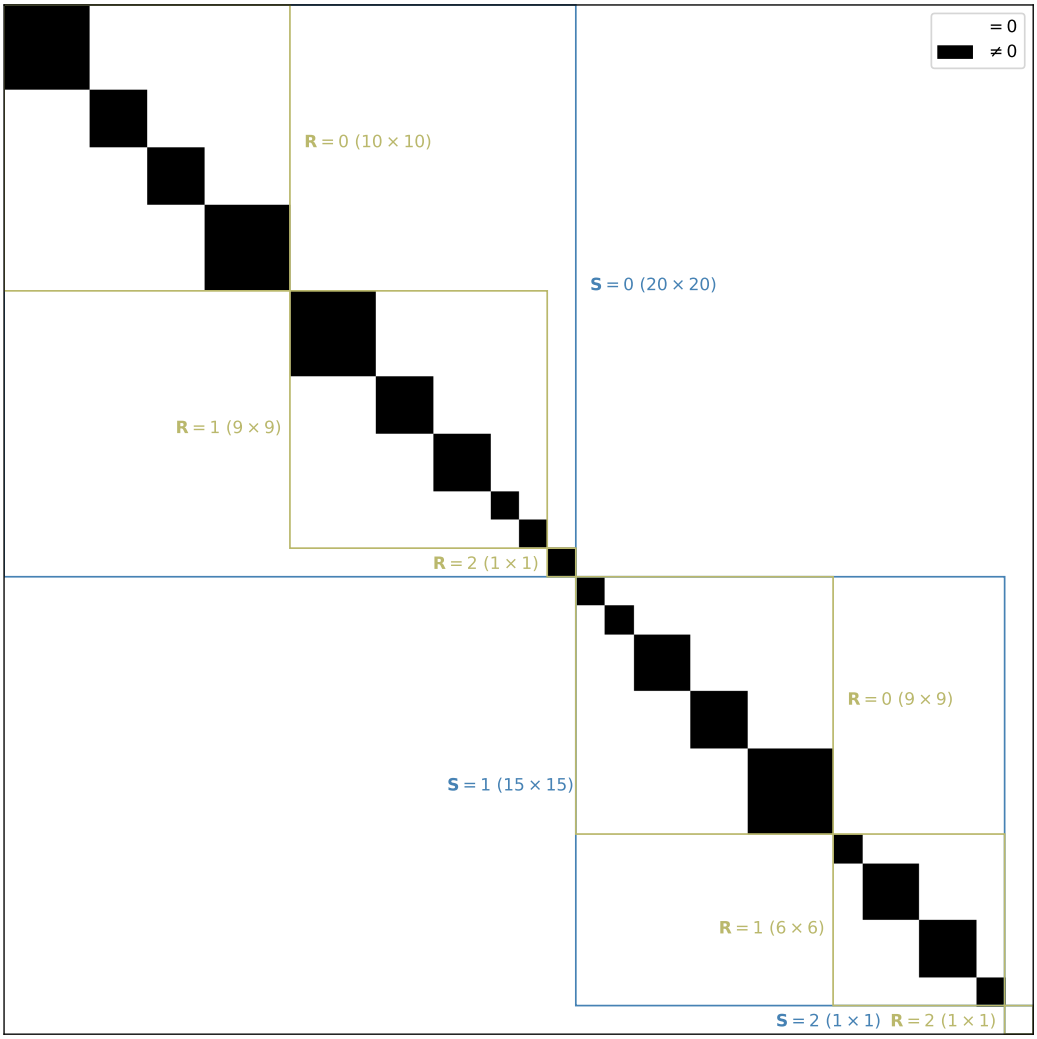


Figure 2.5: Symbolic visualization of the 36×36 block of Hamiltonian, with nonzero terms in black.

of the reflection operators. In other words, in each of the blocks, the eigenvalues of all reflection operators are the same, as can be seen explicitly by running the accompanying code. Also, as per Lieb's theorem 1, the uniqueness of the ground state in the 36×36 block is checked explicitly via numerical diagonalization of the entire block and is found to hold. An analogous conclusion holds for Shen's theorem 2.

Thus, the final form of the remaining four matrices is as follows:

$$\begin{aligned}
\mathbf{H}_{\mathbf{T}=-\mathbb{1}} &= \begin{bmatrix} 0 & 0 & -\sqrt{12}t \\ 0 & 2U & 2t \\ -\sqrt{12}t & 2t & U \end{bmatrix}, \quad \mathbf{H}_{\mathbf{T}=-\imath\mathbb{1}} = \begin{bmatrix} U & (1+\imath)\sqrt{2}t \\ (1-\imath)\sqrt{2}t & U \end{bmatrix}, \\
\mathbf{H}_{\mathbf{T}=\mathbb{1}} &= \begin{bmatrix} 2U & 0 & \sqrt{12}t \\ 0 & 0 & -2t \\ \sqrt{12}t & -2t & U \end{bmatrix}, \qquad \mathbf{H}_{\mathbf{T}=\imath\mathbb{1}} = \begin{bmatrix} U & (1-\imath)\sqrt{2}t \\ (1+\imath)\sqrt{2}t & U \end{bmatrix},
\end{aligned} \tag{2.11}$$

in agreement with the results provided in [45].

The two 2×2 matrices in (2.11) are diagonalized easily after solving quadratic equations, and we find that each of them has eigenvalues $U \pm 2t$[13]. Neither of values correspond to the ground state energy. On the other hand, the remaining two matrices, corresponding to $\mathbf{T} = \pm 1$ produce tedious expressions when diagonalized symbolically. These can be found in [45] and we see analytically, as expected from theorems 1 and 2, that the ground state is unique. Similarly, expanding the explicit expression for the ground state energy up to second order in U (weak interaction limit), we find the the terms to agree with ones obtained via second order perturbation theory, and depicted in Figure 2.4.

At this moment, we do an additional consistency check and explicitly diagonalize all of the remaining blocks in the 36×36 subspace (we have reduced the other blocks inside this subspace even if they do not correspond to the ground state, as already shown in Figure 2.5) and find their explicit energies with parameter choices $U = 5.0$ and $t = 1.0$, obtaining the explicit numerical values given in Table 2.5, which are expectedly in complete agreement with **TABLE VII** in [45], with vanishing chemical potential. In said table, each row corresponds to a block in Figure 2.5, in the same order.

At last, we emphasize that this approach is useful even if it is applicable in part only. Evidently, in this case we have placed certain restriction on the system, such as half filling and $t' = 0$, however even outside of these restrictions, the Hamiltonian is reduced into

[13] After exact diagonalization of these simple matrices, we find that there still remains a twofold energy degeneracy in these blocks. While this is not relevant to our present work since these states are not ground states, as mentioned before, the invariant constructed by Grosse in [57] and used by Schumann in [45] is enough to completely determine all of the states of the system in terms of different quantum numbers, resulting in a complete characterisation of the system spectrum. Of course, this invariant is not needed in our present case since we are focusing on the ground state and we find it to be unique, as is consistent with Lieb's and Shen's theorems 1 and 2.

$\mathbf{S}^2 = 0$	$\mathbf{R}^2 = 0$	-11.84428877	0.84428877	-4.0
		-3.0	-7.0	-
		-7.0	-3.0	-
		-10.84428877	-6.0	1.84428877
	$\mathbf{R}^2 = 1$	1.5136025	-3.16571209	-8.34789042
		-5.70156212	0.70156212	-
		0.70156212	-5.70156212	-
		-5.0	-	-
		-5.0	-	-
	$\mathbf{R}^2 = 2$	0.0	-	-
$\mathbf{S}^2 = 1$	$\mathbf{R}^2 = 0$	-5.0	-	-
		-5.0	-	-
		-4.29843788	-10.70156212	-
		-4.29843788	-10.70156212	-
		-1.65210958	-6.83428791	-11.5136025
	$\mathbf{R}^2 = 1$	-5.0	-	-
		-7.0	-3.0	-
		-3.0	-7.0	-
		-5.0	-	-
$\mathbf{S}^2 = 2$	$\mathbf{R}^2 = 2$	-10.0	-	-

Table 2.5: Energies of all 36 eigenstates corresponding to the $\mathbf{S}^z = \mathbf{R}^z = 0$ block, with parameter values $U = 5.0$, $t = 1.0$ and $t' = 0$. The ground state energy is given by the highlighted value.

blocks (recall that the half filling block is the largest out of all possible fillings), albeit not as small as the ones that we have found[14]. This has powerful implications for real life applications, since evidently the number of terms in Hamiltonian scales as $\mathcal{O}(N^2)$ (quadratically) with N being the space dimension, and for the usual diagonalization algorithms, such as Singular Value Decomposition (SVD), the time complexity is $\mathcal{O}(N^3)$. Thus, the complexity of a usual operation such as diagonalization can be reduced significantly by simple separation of Hamiltonian into blocks based on eigenvalues of particular operators, or by exploiting the symmetries in general.

To conclude, we have successfully used the symmetries of the model, both those that are quantum in nature (such as spin and pseudospin) and those that are geometric (such as rotations) to reduce the initially intractable Hamiltonian to easily diagonalizable blocks of

[14]Furthermore, this approach is applicable to the four-site Hubbard model with attractive onsite interaction ($U < 0$) as well, simply by applying the aforementioned Shiba map [42] which exchanges the spin and pseudospin.

maximal size 3×3.

3

VQE

Quantum mechanics is just completely strange and counterintuitive. We can't believe that things can be here [in one place] and there [in another place] at the same time. And yet that's a fundamental piece of quantum mechanics. So then the question is, life is dealing us weird lemons, can we make some weird lemonade from this?

– Seth Lloyd

Reaching further into the quantum realm, we explore the Variational Quantum Eigensolver (VQE) as one among the most promising algorithms realizable on NISQ devices in more detail. VQE was originally developed by Peruzzo et al. in 2014 [58], while its theoretical foundation was formalized by McClean et al. in 2016 [59].

A high-level goal of VQE is to compute an upper bound for the ground-state energy of a given Hamiltonian, which one can recall from Intermezzo 1.2 to be a necessary step in computing the properties of different molecules and materials. As was also emphasized previously, to be able to do so is of fundamental importance quantum chemistry [60, 61] and condensed matter physics [62] and by extension, possibly extremely useful for drug discovery [63], material science [64], chemical engineering [65], etc.

Why shall one seek quantum algorithms such as VQE next to the many available classical ones? After all, the methods in field of classical computational chemistry are grounded in nearly a century of research and they provide efficient methods to approximate properties of molecular systems. Although this is correct, the issue with such methods arises for increasingly larger systems because the computational cost becomes intractably expensive when highly accurate calculations are desired, reducing the their applicability. As a particular example of this, we can highlight the models involving strong electron correlations, such as the Fermi-Hubbard model explored in this work. Interactions resulting in such correlations generally require computational resources that scale exponentially with system size, rendering the classical methods insufficient [66].

As a constraint of similar type, and as was explored in detail in Section 2.5, reducing the model iteratively while relying on many different symmetries is only viable for a narrow spectrum of systems. Furthermore, even if one were able to use the full range of the described toolkit and obtain reductions such as the demonstrated ones, the complexity of this method would still increase exponentially—algorithms like matrix multiplication or diagonalization are polynomial in Hilbert space size, but the issue is that the space size increases exponentially with the number of particles, since it is a product of all the subspaces. Thus, the paradigm shift that VQE allows for is to replace the exponentially costly algorithms—both analytic, as the reductions we explored, and others, as those of computational chemistry—with quantum algorithms of polynomial complexity.

3.1 NISQ regime

As was outlined in the introductory Chapter 1, the main issue preventing applications of quantum computing is decoherence, mostly due to entanglement with the environment, resulting in loss of the useful inner entanglement of the system. This difficulty has propelled a particular subset of quantum devices—NISQ devices—to the industrial and academic forefronts, with a mix of hope and belief that they may outperform conventional computers in the near future.

We recall that NISQ regime corresponds to quantum computers with a limited number of qubits. It is important to emphasize the most significant consequence of the limited qubit number—it is not possible to implement robust error correction schemes because the standard quantum fault-tolerance techniques introduce significant overheads and are thus unsuitable in this regime [67]. Therefore, for such a device to be able to reach advantage over their classical counterparts in applications, the number of quantum gates acting on them ought to remain of small enough size, as more operations and time it takes to carry them out result in higher probability of introducing errors into the system, causing undesired decoherence. Furthermore, the algorithms ran on these devices are of hybrid nature, since the underlying systems and their manipulations are implemented using qubits and quantum gates, while the parameter optimization is carried out using classical optimizers.

Due to these somewhat stringent restrictions, the applications of quantum algorithms on NISQ devices are limited. However, the excitement about development in this regime has been built on its successes. Initially, NISQ computers have been constructed and shown to outperform classical brute force simulations on problems designed to fit quantum devices' capabilities [68, 69]. Similarly, algorithms such as QAOA and VQE, the latter of which is considered in this work, have been successful in solving different quantum physics and chemistry problems, and have potential applications in fields such as material science, biology and cryptography [19]. As more engineers and academics continue to demonstrate the prowess of NISQ devices and corresponding algorithms in tackling real-world problems,

venturing into research within this regime becomes not only a natural choice, but also a necessity to fully comprehend and harness the potential of quantum technologies.

3.2 Variational Quantum Eigensolver

In this section, we will outline some of the building blocks of VQE, as well as some technical details necessary for the complete pipeline of our approach.

Variational principle

To explain the first term in VQE acronym, we consider the variational principle, a simple but powerful physical method, built on a Rayleigh-Ritz principle [70, 71]. We can understand it intuitively as follows.

Let a physical system be described by Hamiltonian H, acting on Hilbert space $\mathcal{H}$. Then, due to the spectral theorem for Hermitian/self-adjoint operators [21], there exists an orthonormal basis for $\mathcal{H}$ consisting of eigenvectors of $\mathbf{H}$. Let us take one such basis $\{|\psi_\lambda\rangle\}_{\lambda\in\mathrm{Spec}(\mathbf{H})}$[1], satisfying $\mathbf{H}|\psi_\lambda\rangle = E_\lambda|\psi_\lambda\rangle$ for all $\lambda \in \mathrm{Spec}(\mathbf{H})$. Further, let us suppose that the spectrum of $\mathbf{H}$ is bounded from below, with the greatest lower bound E_0—ground state energy.

An arbitrary state $|\psi\rangle$ in $\mathcal{H}$ can be expanded in the chosen basis as $|\psi\rangle = \sum_\lambda a_\lambda|\psi_\lambda\rangle \coloneqq \sum_\lambda \langle\psi_\lambda|\psi\rangle|\psi_\lambda\rangle$. Thus, for the expectation value of H in this arbitrary state $|\psi\rangle$, one can

[1] There are subtleties in formalizing this approach when the spectrum of the model is continuous, but the implications are the same.

write:

$$
\begin{aligned}
\langle\psi|\mathbf{H}|\psi\rangle = \langle\psi|\mathbb{1}H\mathbb{1}|\psi\rangle &= \sum_{\lambda_1,\lambda_2\in\mathrm{Spec}(\mathbf{H})} \langle\psi|\Big(|\psi_{\lambda_1}\rangle\langle\psi_{\lambda_1}|\Big)\mathbf{H}\Big(|\psi_{\lambda_2}\rangle\langle\psi_{\lambda_2}|\Big)|\psi\rangle \\
&= \sum_{\lambda_1,\lambda_2\in\mathrm{Spec}(\mathbf{H})} \langle\psi|\psi_{\lambda_1}\rangle \underbrace{\langle\psi_{\lambda_1}|\mathbf{H}|\psi_{\lambda_2}\rangle}_{\langle\psi_{\lambda_1}|E_{\lambda_2}|\psi_{\lambda_2}\rangle} \langle\psi_{\lambda_2}|\psi\rangle \\
&= \sum_{\lambda_1,\lambda_2\in\mathrm{Spec}(\mathbf{H})} \langle\psi|\psi_{\lambda_1}\rangle \Big(E_{\lambda_2} \underbrace{\langle\psi_{\lambda_1}|\psi_{\lambda_2}\rangle}_{\delta_{\lambda_1,\lambda_2}}\Big) \langle\psi_{\lambda_2}|\psi\rangle \\
&= \sum_{\lambda_1\in\mathrm{Spec}(\mathbf{H})} E_{\lambda_1}\langle\psi|\psi_{\lambda_1}\rangle\langle\psi_{\lambda_1}|\psi\rangle = \\
&= \{\lambda_1 \rightsquigarrow \lambda\} \\
&= \sum_{\lambda\in\mathrm{Spec}(\mathbf{H})} E_{\lambda}\langle\psi|\psi_{\lambda}\rangle\langle\psi_{\lambda}|\psi\rangle \\
&= \sum_{\lambda\in\mathrm{Spec}(\mathbf{H})} E_{\lambda}\langle\psi|\psi_{\lambda}\rangle\langle\psi_{\lambda}|\psi\rangle = \\
&= \{\forall\lambda\colon E_\lambda \geq E_0 \text{ by definition of } E_0\} \\
&\geq \sum_{\lambda\in\mathrm{Spec}(\mathbf{H})} E_0\langle\psi|\psi_{\lambda}\rangle\langle\psi_{\lambda}|\psi\rangle \\
&= E_0 \sum_{\lambda\in\mathrm{Spec}(\mathbf{H})} \langle\psi|\psi_{\lambda}\rangle\langle\psi_{\lambda}|\psi\rangle \\
&= E_0 \langle\psi| \underbrace{\left(\sum_{\lambda\in\mathrm{Spec}(\mathbf{H})} |\psi_{\lambda}\rangle\langle\psi_{\lambda}|\right)}_{\mathbb{1}} |\psi\rangle \\
&= E_0\langle\psi|\psi\rangle \\
&= E_0\,,
\end{aligned}
$$

implying that for Hamiltonian $\mathbf{H}$ with ground state energy E_0 and all states $|\psi\rangle \in \mathcal{H}$ holds a relation:

$$E_0 \leq \langle\psi|\mathbf{H}|\psi\rangle\ .$$

Thus, whatever state $|\psi\rangle$ we use to find the expectation value of $\mathbf{H}$, we will never underestimate the correct ground state energy. This implies that the state $|\psi\rangle$ can be parameterized as

$$|\psi(\theta_1,\theta_2,\ldots,\theta_n)\rangle \coloneqq |\psi(\boldsymbol{\theta})\rangle \tag{3.1}$$

and then its value minimized with respect to these parameters, encoded in a vector $\boldsymbol{\theta}$ of length n, without the possibility of it being less in value than the correct one. To this end, one can use the different optimization procedures to explore the hyperparameter space

spanned by values $\{\theta_i\}_{i\in[1\cdot\cdot n]}$ and find those that produce minimal expectation values of the Hamiltonian.

Jordan-Wigner transformation

Wishing to analyze a model such as the Fermi-Hubbard model using a quantum computer, we run into an immediate conundrum. In quantum computing theory and applications, the objects being manipulated are comprised of qubits, and FH model is comprised of fermions, which satisfy different physical properties, such as the anticommutation relations (1.7) that are not satisfied by qubit operators.

Thus, as one pursues applications of quantum computing in real physical systems, it is necessary to encode them in some way—fermions ought to be mapped to the qubit domain. As it turns out, Jordan and Wigner [72] demonstrated such a transformation exists; an isomorphism between fermionic space and qubit space, demarcated *Jordan-Wigner (JW) transformation*.

Originally, the JW transformation used spin-$1/2$ operators[2] to explicitly construct fermionic ladder operators. Thus, since individual qubits are equivalent to spin-$1/2$ systems, an inverse of JW transformation can be used to map fermions to qubits, i.e. simulate fermions on a quantum computer [73].

To gain an intuitive understanding of the JW transformation, let us consider a space of a single fermion, with the basis states $|0\rangle$ and $|1\rangle$, corresponding to whether the orbital is vacant or occupied by a fermion, respectively, and analogous to the spin-$1/2$ system in which the two states corresponded to spin projection being up or down. Then the creation operator will convert the vacant state into an occupied one, and null the occupied one, as per the Pauli exclusion principle. This implies that a simple matrix form for the operator is

$$\mathbf{c}^\dagger \rightsquigarrow \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} = \frac{\sigma^x - \imath\sigma^y}{2}, \tag{3.2}$$

rewritten in terms of Pauli matrices (1.4), which we recall to correspond to spin-$1/2$ matrices, without the factor of $1/2$. As one can easily see, operators $\mathbf{c}^\dagger$ and $\mathbf{c}$ found in this way do not satisfy the fermionic anticommutation relations (1.7). However, this can be corrected for by adding the factors of σ^z Pauli matrices of all spins before a particular site:

$$\mathbf{c}_j^\dagger \rightsquigarrow \left(\bigotimes_{l=1}^{j-1} \sigma_l^z\right) \otimes \frac{\sigma_j^x - \imath\sigma_j^y}{2}. \tag{3.3}$$

[2]The need for the JW transformation arises since the spin operators are neither fermionic nor bosonic [36]. Thus, in order to construct a consistent solution to the problem, they are mapped to either the fully fermionic, or fully bosonic operators, in our case the former, via JW transformation.

In this way, when the spins on two sites are exchanged, there is an additional product of a string of σ^z matrices. Evidently, since all Pauli matrices square to the identity, when these operators are multiplied, the only ones that remain are those between the two sites, exactly accounting for the necessary phase factor of ± 1, since we can simply imagine that we are exchanging the particles in steps, where each step exchanges the two neighbouring particles, and each σ^z matrix $\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$ adds a -1 factor to the occupied states, while leaving the vacant states invariant.

Now the inverse JW transformation can also be used on the Hamiltonian of the system to convert it to qubit space, and once this is done, the subsequent calculations are all done using qubits and wavefunctions in the corresponding Hilbert space.

An important point can be observed from JW formula (3.3). Recalling the form of z Pauli matrix σ^z from (1.4), we see that a string of these operators does not mix qubit states $|0\rangle$ and $|1\rangle$, only adds possible relative phase factors. On the other hand, as per (3.2), the term $(\sigma^x - \imath\sigma^y)/2$ evaluates to a matrix $\begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}$, and thus in qubit space, it will annihilate spin up $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$ and flip spin down $\begin{pmatrix} 0 \\ 1 \end{pmatrix}$ to a spin up. In other words, the physical identification between fermionic and qubit spaces via JW transformation will be that, in a fermionic one-particle space of dimension 2, two possible states correspond to a particle of a given spin on a given site either existing on that site or not, while in the qubit case they will correspond to a qubit existing, however either being in a spin down or spin up state.

State preparation and problem formulation

Once the mapping of the fermionic physical Hamiltonian to the qubit space has been established, we wish to use the variational principle in this setting—due to the principle and its implications being described and motivated in Subsection 3.2, a natural way presents itself immediately. Namely, one could construct a quantum circuit with several parameterized gates, thereby effectively constructing an operator $\mathbf{U}(\theta)$ to prepare a parameterized state such as (3.1) if one acts on a systems of qubits with it. As an initial state, it is customary to set all $2N$ qubits in the $|0\rangle$ state and then act on them using this operator, which will in general encode the parameter dependence based on its structure. Thus, the state of $2N$ qubits from (3.1) can be written as

$$|\psi(\theta\rangle := \mathbf{U}(\theta)\,|0\rangle^{\otimes 2N} := \mathbf{U}(\theta)\,|\mathbf{0}\rangle\,, \tag{3.4}$$

where we have demarcated the product state of all $2N$ qubits in state $|0\rangle$ by $|\mathbf{0}\rangle$.

In this way, one is able to formulate a generic VQE problem

$$E_{\mathrm{VQE}} = \min_{\theta} \langle \mathbf{0}|\, \mathbf{U}^\dagger(\theta)\mathbf{H}\mathbf{U}(\theta)\,|\mathbf{0}\rangle$$

as an optimization over θ of the expectation value of $\mathbf{H}$ in state $\mathbf{U}(\theta)|\mathbf{0}\rangle$. Since the structure of this state is decided through construction of $\mathbf{U}(\theta)$, it is conventionally referred to as the **ansatz**.

Shot noise

In real quantum devices, there are several types of noise: readout noise, depolarizing noise, amplitude and phase damping noise, and so on [74]. In this work, we will focus on the shot noise. Generally, to achieve a precision of ε on the expectation value of an operator such as a Hamiltonian, we are required to perform $\mathcal{O}(1/\varepsilon^2)$ repetitions (usually denoted as *shots*) of the circuit execution, each completed with a measurement at the end [14, 59], according to the central limit theorem and properties of Gaussian distribution. Thus, this type of noise stems from the fact that in real world, we can do only finitely many measurements, compared to infinitely many in an ideal case[3]. In other words, it allows for testing whether or not the solution we construct will also work well once the number of simulated measurements is finite, and thus serves as an important criterion on the applicability of our solutions.

Optimization procedure

Having described the generic form of the VQE problem (3.2) and its ansatz as a parameterized trial solution in (3.2), one ought to actually optimize its parameters and solve the problem in full.

It is known that optimization of the parameterized quantum ansätze is NP-hard in general [77], implying that not all problems can be solved via the optimization process in reasonable timeframes. Thus, to render algorithms of this kind viable for applications, it is essential to construct and rely on efficient optimization strategies, i.e. the ones that provide reliable approximations of exact solutions in feasible times and using reasonable amounts of resources. In practice, this is done via choice of the optimizer and its parameters and, as expected, this choice affects the precision and efficiency of the problem-solving significantly.

Let us denote the objective function we wish to minimize, i.e. the expectation value in (3.2) by

$$\mathfrak{F}(\theta) \coloneqq \langle \mathbf{0}| \mathbf{U}^\dagger(\theta)\mathbf{H}\mathbf{U}(\theta)|\mathbf{0}\rangle \ .$$

Then, the optimizer is used to minimize $\mathfrak{F}(\theta)$, yielding the ground state energy which is close to the exact one, of course depending on whether or not the ansatz's and optimizer's

[3] Besides this, in VQE appplications, it is often needed to evaluate sets of noncommuting operators, compared to a single operator we are evaluating in this work (Hamiltonian) and this significantly increasing the needed number of shots [75, 76].

structure and properties are chosen well enough. We will deal with the ansatz construction in detail in Section 3.4, so let us consider the different optimizers now.

First of all, the importance of the proper optimizer choice and tuning cannot be overstated. Some optimizers are faster than the others, some require less function calls during evaluation, some converge more quickly and some are more robust in noisy calculations.

In this work, we will consider three optimizers—SLSQP, COBYLA and SPSA. First reason for our optimizer choice was their demonstrable success and robustness across different physical models, for example in [78], it was shown that COBYLA and SPSA outperform some older optimizers, while SLSQP was shown to have a great performance in [79]. Furthermore, all three are based on qualitatively different underlying optimization methods and thus allow to analyze the properties and differences of such methods.

SLSQP

Sequential **L**east **SQ**uares **P**rogramming (SLSQP) is a gradient-based optimizer that iteratively solves quadratic subproblems that are constructed as approximations of our original nonlinear problem (3.2) [80]. It is ideal for problems described by functions (and possibly constraints) that are twice continuously differentiable, as the noiseless VQE optimization is. From this, and based on previous results such as [79], we expect the optimizer to work well in the noiseless case, however to not be robust once the noise is introduced.

COBYLA

Constrained **O**ptimization **BY** **L**inear **A**pproximation (COBYLA) is an example of a gradient-free optimizer, i.e. one that is not based on gradient descent [81]. It works by connecting $n + 1$ points in the parameter space (n being the number of parameters) to define a hyperplane on which the local slope of the cost function is calculated, followed by replacement of the highest cost function value point by stepping from the lowest cost point in the slope direction:

$$\theta_{\max} \rightsquigarrow \theta_{\min} + \delta \nabla \mathfrak{F}(\theta_{\min}) ,$$

where δ represents the step size parameter, i.e. decides on how far-reaching the point updates are. This parameter is decreased when the algorithm detects that is has stopped minimizing the value of the objective function $\mathfrak{F}$, thus allowing the convergence. Although COBYLA works well even when the optimization function is nondifferentiable, it assumes that an exact measurement is carried out in each iteration, and is thus not suitable for calculations including shot noise. However, it is extremely fast and allows for quick iteration and solution checking in the noiseless case.

SPSA

Compared to direct measurements employed by the previous two optimizers, **S**imultaneous **P**erturbation **S**tochastic **A**pproximation (SPSA) carries out a random sampling instead [82]. Specifically, it chooses a small random vector δ during each iteration of the traversal of the parameter space and approximates the gradient in the following way:

$$\frac{\mathfrak{F}(\theta + \alpha\delta) - \mathfrak{F}(\theta - \alpha\delta)}{2\alpha\delta}, \tag{3.5}$$

i.e. the objective function is perturbed in the direction of a small random vector δ, with update magnitude determined by the learning rate α.

Evidently, only two measurements are required per iteration, making it computationally feasible for direct applications as well as the other two optimizers. Furthermore, it is also well-behaved under different types of noise due to several reasons, of which one can emphasize that in the process, all optimization variables are perturbed simultaneously, averaging the noise contributions and resulting in a more stable gradient estimate. Thus, it is better suited for noisy calculations case compared to the previous two optimizers [83, 84, 85], and is expected to work better once the shot noise is introduced.

Barren plateaus

An important obstacle in modern machine learning, and especially the neural networks is the emergence of vanishing gradients, an issue that hurts the learning process since it prevents updates to the model parameters. Unsurprisingly, it was shown [59, 86] that this issue is very much present in quantum variational algorithms as well, in certain cases even exponential vanishing of the gradient in terms of the number of qubits in the system. However, despite the analogy with the classical problem, in quantum pipelines the problem is aggravated further for two main reasons. First, the process of measurement is inherently stohastic on quantum devices and so is the gradient estimation. When the gradients are vanishing, i.e. approaching zero, it becomes increasingly difficult to determine their sign, and failure to do so results in a random optimization process, requiring an immense amount of measurements to counteract the effect. Second, the problem is dependent on many factors, such as the number of qubits, ansatz expressibility, entanglement degree, nonlocality, etc., which makes it harder to effectively mitigate the issue [14].

During the ansatz construction, we will address some of the methods to alleviate the barren plateau issue.

3.3 Model implementation

For all of the following implementations and subsequent calculations, we use the well-known Python package for quantum computing—Qiskit—which has in recent years evolved to contain a plethora of problem solving toolkits, particularly most of what we will need in our VQE pipeline.

We implement the model in two distinct ways, to ensure consistency and correctness of our results. First one is using the `LatticeModelProblem` class from `qiskit_nature.second_q.problems` in which we specify the onsite interaction directly and the hopping terms via weights of edges in the interaction graph. This model is solved explicitly using the `NumPyMinimumEigensolver`, and the particle number is simply passed to the `LMP` class as an auxiliary operator that it calculates concurrently with other properties of each eigenstate and then the states are filtered based on the value of this operator, which we have set to correspond to half filling. The second approach uses the standard `FermionicOp` class in which the fermionic creation and annihilation operators are specified explicitly. In this case, we use the same solver, however with an additional Hamiltonian term to force the desired particle numbers. In other words, the total Hamiltonian is equal to the physical one, plus the term of form $\lambda\left[(\mathbf{N}_{\uparrow} - N/2)^2 + (\mathbf{N}_{\downarrow} - N/2)^2\right]$, evidently equal to zero at half filling with total spin projection zero and when we tune up λ to be large, this term will be much larger than the lower excitation energies. Thus, it plays a role analogous to a custom loss function, punishing heavily the states that do not satisfy the requirements. Regarding the mapping from fermion space to qubit space, we use Qiskit's `JordanWignerMapper`.

Implementations of both of these models and the functions that solve them, along with implementations of all other auxiliary second quantization operators can be found in `model_utils.py` script.

3.4 Building the ansatz

When trying out the common ansätze for the problem at hand, such as the Hardware-efficient ansatz (HEA) and Unitary Coupled Cluster (UCC) ansatz, we quickly find that they are not sufficient. First of all, they require either a large amount of parameters and circuit depth, which is not conducive to later calculations with noise, but they also show issues with barren plateaus, making it impossible for the optimizers to arrive at the global minimum [14, 87, 88]. An intuitive reasoning behind this problem is that these ansätze are simply too expressive, i.e. based on the values of their parameters, they can take on values in most of the available Hilbert space. For this reason, the optimization procedure is not restricted well-enough to the actually physically correct subspace of the total space, and thus it is much trickier to arrive at the correct solution. To mitigate this effect, we can

construct a custom ansatz and embed certain restrictions into it. For example, we could, in the beginning, raise exactly half of the possible excitations (to correspond to the half-filled regime, i.e. total pseudospin projection zero) and also raise particles of so that the total spin projection is as desired and then only use the gates that preserve both the total particle number and the total spin projection[4].

For completeness, we now explicitly build the ansatz for $N = 4$ spin sites. We approach building our ansatz in steps, while relying heavily on the underlying physics and symmetries. Since we are working at half filling, out of 8 possible particles, we raise exactly 4 of them. Regarding their spin orientations, as mentioned in 2.3, it was found in [1] that the spin projection sector containing the ground state(s) is, as in the $t' = 0$ case, the $\mathbf{S}^z = 0$ sector. Thus, we solve the case in which two of these particles are of spin up and two are spin down, yielding a total spin projection zero. Note that this yields an identical restriction of the Hilbert space as in the $t' = 0$ regime—in that case Lieb's theorem 1 was used, and in this case explicit calculations. Of course, in principle, one would use the constructed solution to find lowest energy states in all sectors (both of spin and pseudospin) and then take those with lowest energy among all of them, however, as was already pointed out, this sector is of the largest complexity and thus a solution working for it is presumed to work for others as well, with change in the initial state based on the particles that are excited [43]. As an important clarification regarding the total spin of the particles, we note that here we are talking somewhat interchangeably about flipping qubits and raising fermions of a particular spin. This is convenient to do, and also completely consistent, because as was explained at the end of Subsection 3.2, identification between a fermionic state of a particular spin on a particular site being vacated (occupied) and qubit being of spin down (up) can be done directly.

Thus, we begin our ansatz circuit by flipping spins of qubits 0, 2, 5 and 7 to up value using instances of quantum gates X, which in terms of fermions corresponds to four excited particles, 0 and 2 of spin up and 5 and 7 of spin down.

After this, since the only constraint placed on quantum gates is unitarity, we see that they will in general not conserve the particle number (evidently X is one such gate). If we were to employ gates that do not conserve the particle number, applying them would throw the resulting statevector outside the constrained space into a significantly larger complete Hilbert space, and make it harder for VQE to converge to a correct solution. Thus, we restrict the gates we use to particle number conserving ones. Regarding the two qubit gates of this kind, we first find that in general they will have five independent real parameters that we are able to vary. This can be seen in a number occupation basis of two particles as follows.

[4] As an additional, but nonobvious issue, HEA and UCC, due to their lack of preservation of symmetries such as the total particle number or the total spin, can result in kinks in the potential energy surfaces, i.e. lack of smoothness, which is not directly relevant to our current work, but still represents an issue for applications in quantum chemistry [88].

The first and the last 1×1 sectors, corresponding in order to zero or two particles cannot mix with any other sectors. This results in these blocks being diagonal, i.e. no offdiagonal elements in first and last row and column, yielding a total of two free parameters—each unitary matrix of dimension 1×1 is determined by $1^2 = 1$ parameter. Regarding the middle 2×2 sector, which corresponds to exactly one particle being excited—we see that rotating between these two states conserves the particle number. Thus, we can write this matrix as a standard rotation, resulting in an additional parameter, along with phases for each of the four terms, which are constrained by unitarity. In other words, since the dimensions of the submatrix is 2×2 and the total matrix is block diagonal, this submatrix ought to be unitary as well. This implies that there are $2^2 = 4$ free parameters parameters to choose from. Finally, one of the parameters ought to be discarded due to the global phase irrelevance. Thus, there are in total five free parameters of a general two qubits particle number conserving gate. Using such a gate directly in our ansatz would yield a too large number of parameters in the resulting circuit and would do so unnecessarily since precise solutions are found with simple, one-parameter gates. Therefore, we will focus on the three gates given in Figure 3.1, with $\mathcal{R}_{x,y,z}(\theta)$ being standard 2×2 rotation matrices around x, y and z axes in spin space, respectively, as given in (1.19). It is easy to see that these custom gates conserve the particle number and also that they are symmetric in two qubits (invariant under their exchange)—this is why they are visualized in a manifestly symmetric manner in Figure 3.1.

Let us demonstrate this for the $\mathcal{R}_x^{\mathrm{s}}$ gate. Most generally, we can write the statevector for the qubit on the first wire as $\alpha_1 |0\rangle + \beta_1 |1\rangle$ and on the second as $\alpha_2 |0\rangle + \beta_2 |1\rangle$. Then the total system state after the action of the first CX gate in $\mathcal{R}_x^{\mathrm{s}}$ will be (when the first qubit it in $|0\rangle$ state, the second qubit is not altered):

$$\begin{aligned}[\alpha_1 |0\rangle + \beta_1 |1\rangle] \otimes [\alpha_2 |0\rangle + \beta_2 |1\rangle] &= \alpha_1 |0\rangle \otimes [\alpha_2 |0\rangle + \beta_2 |1\rangle] + \beta_1 |1\rangle \otimes [\alpha_2 |0\rangle + \beta_2 |1\rangle] \rightsquigarrow \\ &\rightsquigarrow \alpha_1 |0\rangle \otimes [\alpha_2 |0\rangle + \beta_2 |1\rangle] + \beta_1 |1\rangle \otimes [\alpha_2 |1\rangle + \beta_2 |0\rangle] \ . \end{aligned} \tag{3.6}$$

What follows is the action of the controlled rotation $\mathcal{R}_x(\theta)$ (we use the rotation matrix definition from (1.19)):

$$\begin{aligned} &\alpha_1 |0\rangle \otimes [\alpha_2 |0\rangle + \beta_2 |1\rangle] + \beta_1 |1\rangle \otimes [\alpha_2 |1\rangle + \beta_2 |0\rangle] = \\ &= \alpha_1\alpha_2 |00\rangle + \alpha_1\beta_2 |01\rangle + \beta_1\alpha_2 |11\rangle + \beta_1\beta_2 |10\rangle \rightsquigarrow \\ &\rightsquigarrow \alpha_1\alpha_2 |00\rangle + \alpha_1\beta_2 \left[\cos\frac{\theta}{2} |0\rangle - \imath \sin\frac{\theta}{2} |1\rangle\right] |1\rangle + \\ &+ \beta_1\alpha_2 \left[-\imath \sin\frac{\theta}{2} |0\rangle + \cos\frac{\theta}{2} |1\rangle\right] |1\rangle + \beta_1\beta_2 |10\rangle \ . \end{aligned}$$

Finally, after the second CX gate, the resulting state is:

$$\alpha_1\alpha_2|00\rangle + \alpha_1\beta_2\left[\cos\frac{\theta}{2}|0\rangle - \imath\sin\frac{\theta}{2}|1\rangle\right]|1\rangle +$$
$$+\beta_1\alpha_2\left[-\imath\sin\frac{\theta}{2}|0\rangle + \cos\frac{\theta}{2}|1\rangle\right]|1\rangle + \beta_1\beta_2|10\rangle =$$
$$\alpha_1\alpha_2|00\rangle + \alpha_1\beta_2\cos\frac{\theta}{2}|01\rangle - \imath\alpha_1\beta_2\sin\frac{\theta}{2}|11\rangle -$$
$$-\imath\beta_1\alpha_2\sin\frac{\theta}{2}|01\rangle + \beta_1\alpha_2\cos\frac{\theta}{2}|11\rangle + \beta_1\beta_2|10\rangle =$$
$$\alpha_1\alpha_2|00\rangle + \left[\alpha_1\beta_2\cos\frac{\theta}{2} - \imath\beta_1\alpha_2\sin\frac{\theta}{2}\right]|01\rangle + \left[-\imath\alpha_1\beta_2\sin\frac{\theta}{2} + \beta_1\alpha_2\cos\frac{\theta}{2}\right]|11\rangle - \beta_1\beta_2|10\rangle \rightsquigarrow$$
$$\rightsquigarrow\alpha_1\alpha_2|00\rangle + \left[\alpha_1\beta_2\cos\frac{\theta}{2} - \imath\beta_1\alpha_2\sin\frac{\theta}{2}\right]|01\rangle + \left[-\imath\alpha_1\beta_2\sin\frac{\theta}{2} + \beta_1\alpha_2\cos\frac{\theta}{2}\right]|10\rangle - \beta_1\beta_2|11\rangle \ .$$

In this final expression, we immediately see that the first and the last terms are symmetric upon simultaneous exchange of the qubits and indices $1 \leftrightarrow 2$. Furthermore, the same exchange transforms the two middle terms into one another, and thus we have shown that the expression as a whole is symmetric in two qubits. The same holds for the remaining two gates in Figure 3.1. Further still, when writing out the four terms in the initial qubit state (3.6), we note that the states with coefficients $\alpha_1\alpha_2$ and $\beta_1\beta_2$ remain the same after the transformation, i.e. $|00\rangle$ and $|11\rangle$. The states $|01\rangle$ and $|10\rangle$ are mixed, but they have the same total spin projection and thus this mixing does not change it. In other words, our symmetric gates also preserve the total spin projection of the system. The same is true for the total particle number. If we recall the intuition behind the JW transformation, the states with zero or two particles are not mixed with any others, while the two states that each have one particle are mixed, again resulting in conservation. Thus, since our Hamiltonian commutes with the total spin and pseudospin projection operators and preserves these quantum numbers, by constructing the gates that do so as well, we can imbue our ansatz with the properties of the system, while also ensuring that no unphysical solutions are produced when it acts on the initial qubit state.

Regarding the ordering of qubits in the circuit, first four wires are of qubits with spin up and the remaining four of spin down, and with this fixed, we can proceed inserting the gates into the ansatz. We do so inspired by the efficient exact solution of the noninteracting problem in terms of controlled rotation gates [89] and the ansatz in [43], while relying on the physical significance of terms in Hamiltonian (2.3). Evidently, the hopping terms such as the t ones (nearest neighbour hopping) correspond to nonvanishing amplitudes between the states which only differ by position of that one particle between the two sites. Thus, we include the gates for this hopping type as four gates in each spin block, so that spins of the same orientation differing by one site are entangled, resulting in Figure 3.2, where $\mathcal{R}_y^{\mathrm{s}}$ gates are used since they mix the different spin orientations.

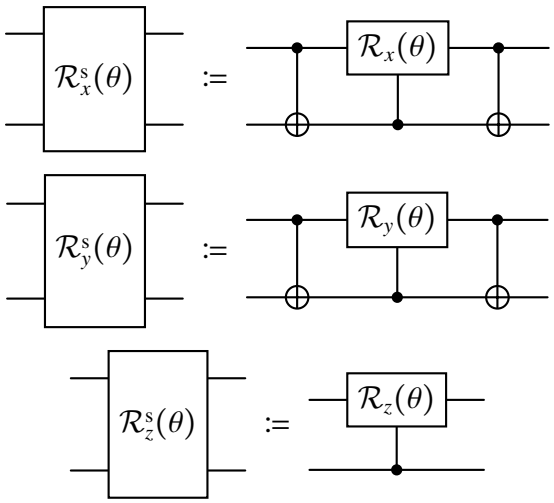


Figure 3.1: Number of function calls during optimization procedure for SLSQP optimizer and different values of model parameters U and t'.

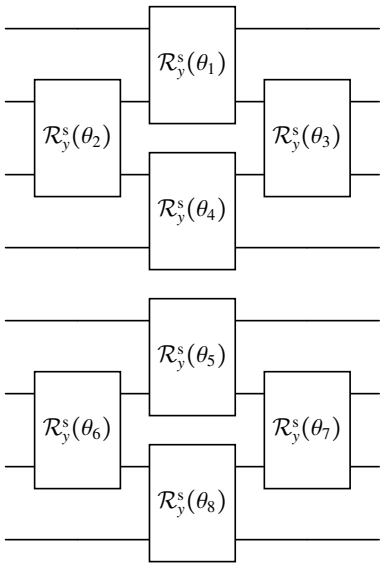


Figure 3.2: First set of entangling gates, corresponding to hopping terms of the Fermi-Hubbard model.

Following this, we add a layer of gates associated with onsite interaction terms in (2.3). To this end, we entangle the qubits of spin up with those of spin down on the same sites, however we do not rotate between those states, but rather add relative phase factors to them. This is due to the fact that the onsite term, as opposed to the hopping terms, does not alter the statevectors, it simply numbers the particles in them. Thus, we use the $\mathcal{R}_z^{\mathrm{s}}$ gates, as shown in Figure 3.3, however all four gates with the same parameter (again, we want no induced phases between different states since onsite interaction does not alter them).

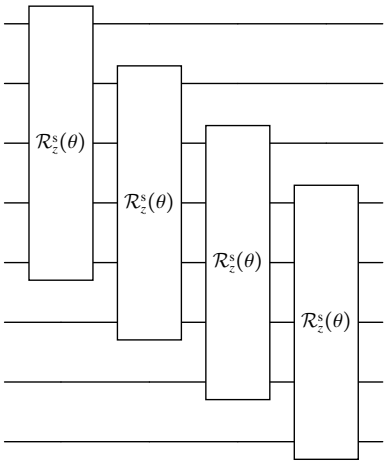


Figure 3.3: Second set of entangling gates, corresponding to onsite interaction term of the Fermi-Hubbard model.

Finally, we add another layer of gates emulating hopping terms of the Hamiltonian, albeit with two modifications. First, we include the NNN hopping terms as well by using four gates that entangle next nearest sites. Second, we add two gates that correspond to periodic boundary conditions and rotational invariance of the model by connecting the last qubit in each of the two blocks to first qubit in said block. As in the first set of gates that correspond to hopping terms, we want rotations that mix the states, so we use $\mathcal{R}_x^{\mathrm{s}}$ gates. The resulting circuit is visualized in figure 3.4.

Finally, the complete ansatz we use, along with the layer repetition and initially excited particles, is given in Figure 3.5. As each layer has 21 parameters, the total ansatz has $L \cdot 21 = 21L$ of them, particularly $3 \cdot 21 = 63$ parameters for the three-layer ansatz used in the noiseless case and $4 \cdot 21 = 84$ parameters for the four-layer ansatz used in the noisy case, as will be explained later.

Furthermore, the ordering of qubit spins in first block of four spins up and then a block of four spins down is now seen to yield slightly more readable circuits compared to alternating (interleaved) qubit ordering.

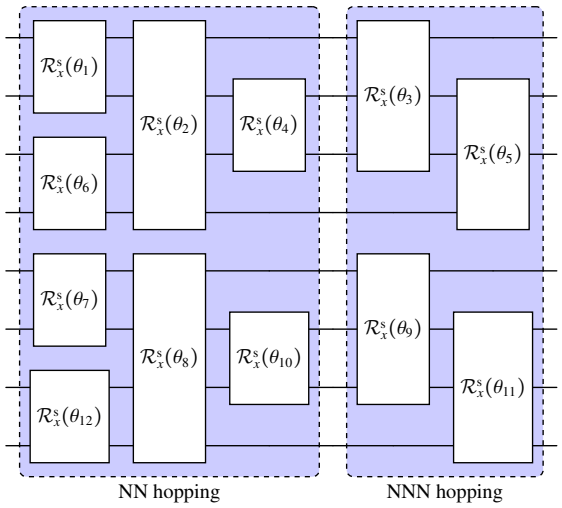


Figure 3.4: Third set of entangling gates, corresponding to hopping terms of the Fermi-Hubbard model.

3.5 Ansatz characterization

Let us give a short overview of metrics used to characterize the different ansästze. An obvious one is the number of parameters which in the case of our ansatz 3.5 depends on the number of layers L linearly, as $21L$. Furthermore, for larger system size, i.e. increased N, the general prescription outlined in 3.4 evidently scales linearly with N, which is easy to see for each of its three constituents separately. Thus, the complete scaling of ansatz complexity—in terms of both the circuit complexity and the number of parameters—is of order $\mathcal{O}(NL)$.

A less obvious metric to use is the ansatz *expressivity*[5], which corresponds to the span of states that a particular ansatz can reach as its parameters are varied. Consequently, expressivity determines the theoretical bound on accuracy that the ansatz can attain in approximating the true ground (or another) state of the model, given that all the parameters are fully optimized. However, from another perspective, high expressivity can make the spanned Hilbert space too large and render the optimizer unable to find an optimal parameter set—highly expressive ansätze are prone to barren plateaus [90]. For this reason, in the previous section we focused on imposing the symmetry requirements into the ansatz to ensure that the spanned space is significantly smaller than the total Hilbert space. The definition of expressivity is based in measure theory and is thus ommited, but its implementation can be found in the `utils.py`.

Another metric we will consider is the *entangling capability*, which quantifies the capa-

[5] Also denoted *expressibility* in literature.

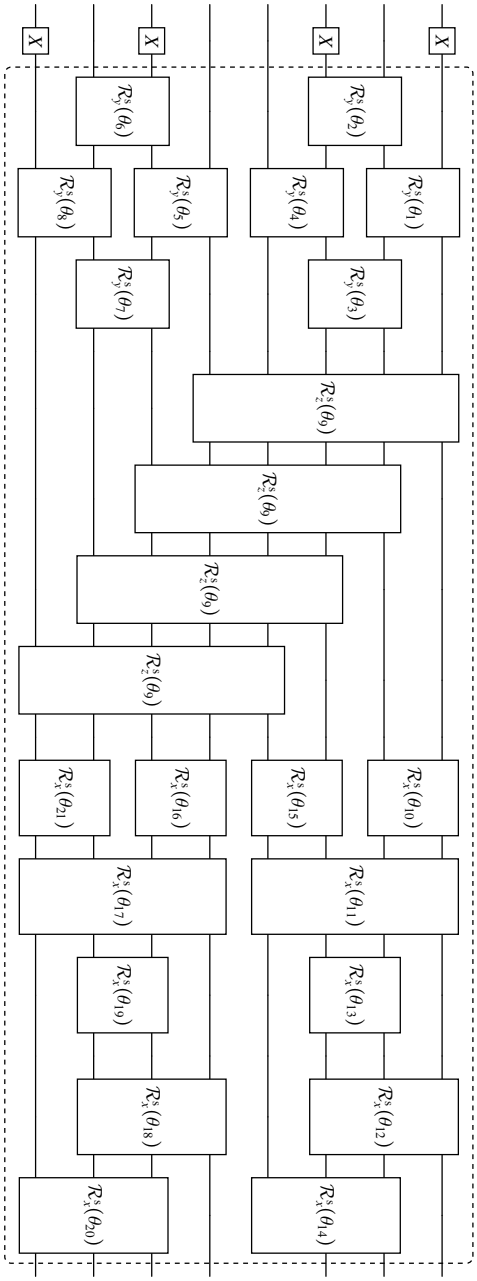


Figure 3.5: Complete ansatz used in VQE calculations.

bility of the ansatz to generate highly entangled stated, which may be needed depending on how entangled the exact state is. A measure proposed in [91] is the *Meyer-Wallach (MW) entanglement measure* [92], used among many others for its scalability and ease of computation. As before, its definition is given in the `utils.py` script and thus omitted here—its high-level interpretation is somewhat simple and it will be exemplified for different ansatz depths later on.

4

VQE Results

However beautiful the strategy, you should occasionally look at the results.

– Winston S. Churchill

We will focus our analysis on parameter values of up to 11.0, both for U and t', mainly because there are no new phases outside of this range, and thus no qualitatively different behaviour [44]. Of course, the proposed solution has been checked for several random high parameter values, such as $U = 100.0$ or $t' = 100.0$ and with the other parameter varied, and expectedly, found to be easily solvable by the constructed ansatz because in those cases. This is because the rest of the interactions are usually negligible (compared to the greatest one or two) so the solutions take simple forms, for example those described in limiting cases in Section 2.4. Furthermore, in real materials (like cuprates) extremely large values of U (as well as t and t') are not observed (but rather moderate ones of order $U = 10$) [93], and even if they were, since the onsite interaction is significantly larger than the hopping terms t and t', perturbative analysis is usually sufficiently precise (as we have shown in Figure 2.4), while in the cases of negligible U, we find that the problem reduces to that of noninteracting fermions and is once again easily solvable (as we have shown in Section 2.4).

4.1 General considerations

Before we outline the different results, let us mention several general observations made during the optimization process.

Global seed choice of the algorithms was generally found not to impact the results much. It was observed that sometimes it affected the initial rate of convergence, presumably due to a different initial position within the complex hyperparameter landscape from

which the convergence was slightly more easy/difficult initially, but the final values of convergence were seldom altered much. Similarly, in cases in which the ansatz was either not capable enough (in terms of producing an entangled state), or an optimizer was not able to find a path to a global minimum, seed altering again made no qualitative difference.

Regarding the parameter initializations, among trying out several more sophisticated ones such as beta distribution [94], it was found that the best results are obtained either by setting all parameters to zero, or by setting no initial values, and letting VQE decide on them by calculating them based on the ansatz form, or just producing random ones. However, the experiments with different initializations were not exhaustive because the simplest choices were sufficient for precise solutions, so these and similar schemes [95] could prove valuable if we were to encounter systems that are more difficult to solve, either due to larger size or more complex interactions.

The main script used to run the simulations is `run_and_log.py` that allows for a high degree of modification—such as choosing the ansatz and its parameters, optimizer and its parameters, number of shots, model interaction strengths, the qubit ordering, etc.

As a technical observation, during the actual simulations, we will be specifying the numbers of fermions of different spins, so the number forcing operator described in 3.3 will be excluded from the Hamiltonian since it is unnecessary with these specifications, and was found to complicate the search space and have a nonnegligible detrimental effect on the results.

4.2 Depth analysis

Before deciding on ansatz depth, i.e. how many layers should it contain, we carry out a simple analysis of ansätze with different numbers of layers, as shown in Figure 4.1. First of all, we find that an extremely close value to the exact ground state energy for $U = 4.0$ and $t' = 0.5$ is obtained, somewhat dramatically, already with 3 layers, as seen in the upper left plot. As a consistency check, we also calculate the expressivity of the ansätze [96] and we find that all of them are essentially of the same expressivity, implying that the state space is correctly reduced to only a small subspace of the total Hilbert space and that the ansatz gates preserve this property. From an optimization perspective, this prevents the optimizer from getting lost by exploring possibilities that are known not to yield the correct solution. These values are shown in the lower left plot. From the lower right plot, we see that the entangling capability first increases steeply with increasing layers, but stagnates around $4 - 5$ layers. However, the entangling capability at 3 layers is enough for the SLSQP optimizer to find the correct ground state energies extremely precisely, as will soon be seen. In the noisy case, 4 layers will be sufficient.

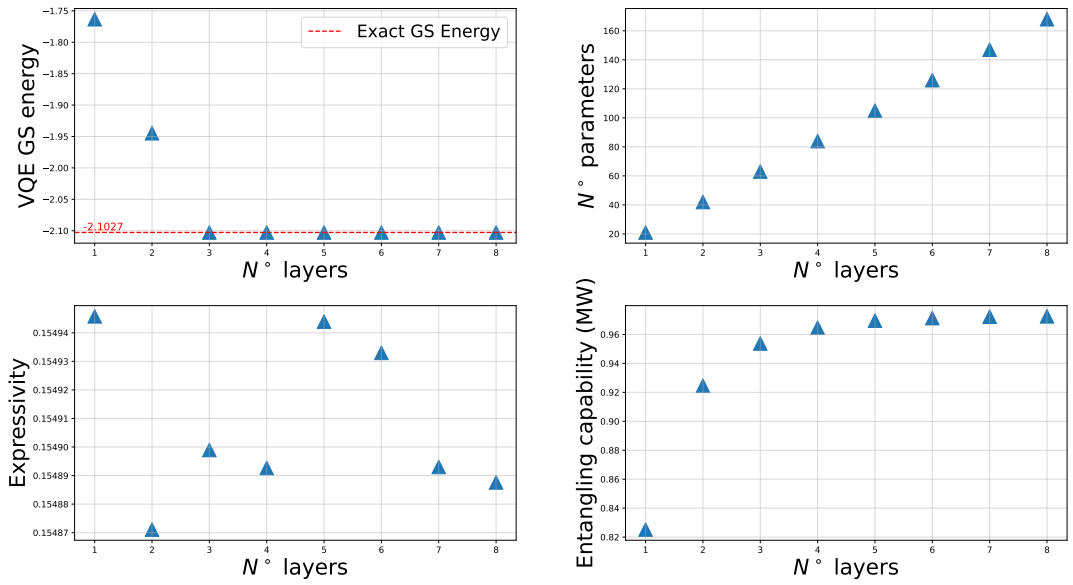


Figure 4.1: Plots of four metrics for the ansatz 3.5 with different layer numbers. **Upper left**: VQE ground state energies energies for $U = 4.0$ and $t' = 0.5$. **Upper right**: number of parameters, equal to number of layers multiplied by 21. **Lower left**: ansatz expressivity. **Lower right**: entangling capability of the ansatz, quantified by the Meyer-Wallach entanglement measure.

4.3 Limiting cases

We start the analysis with the 3-layer ansatz and limiting cases we have explored during the analytic approaches 2.4. We use the SLSQP optimizer, since it is expected to work best in the noiseless case (as will be confirmed empirically later on).

First of all, we can consider the noninteracting $U = 0$ regime. Of course, as a consequence of the fact that in this case the Hamiltonian can be separated into blocks, as we have already done analytically in (2.8), the solution should be easy for our ansatz to find as well—this is a useful preliminary test in the sense that if we wish for it to be robust across all parameter values, it should easily handle this noninteracting case. For the purpose of numerical stability and compatibility with Qiskit's model builder functions, we set the onsite interaction parameter to a negligibly small value $U = 10^{-6}$ and the extremely precise results are given for different values of t' in Figure 4.2—the errors are of order of magnitude $10^{-6} - 10^{-7}$.

As a side note, in this case, there also exists an efficient (with number of gates quadratic in number of qubits) implementation of the exact ground state solution in terms of two-qubit rotation gates, as described and implemented in [89, 97] and provided in

`noninteracting_solution.py` script. The actual optimal circuits can be visualized by running said script—these were mentioned to be an inspiration for the gates included in our ansatz.

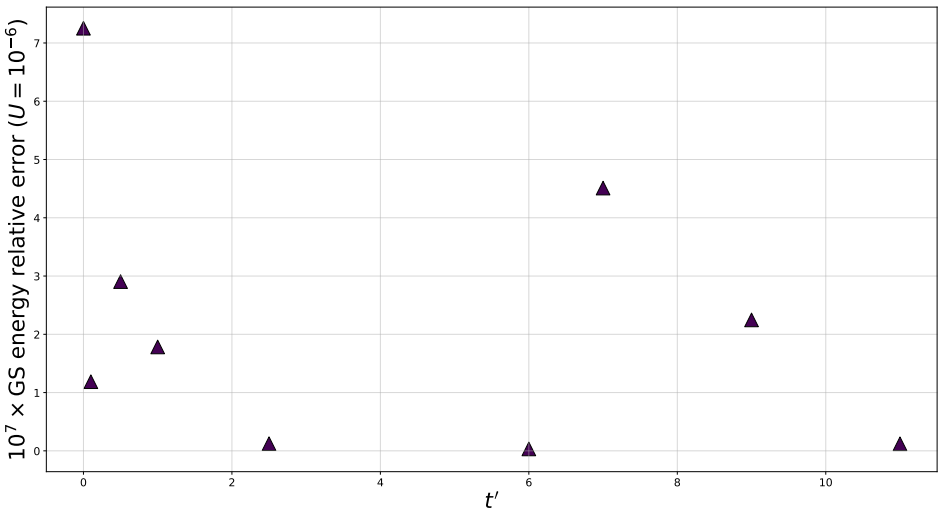


Figure 4.2: Relative error of the VQE ground state energies found using ansatz 3.5 with depth of 3 layers and SLSQP optimizer, for negligible onsite interaction strength $U = 10^{-6}$.

As an opposite regime, we consider the strongly interacting limit including only the onsite interaction U, i.e. with $t = 0$, $t' = 10^{-6}$, with results shown in Figure 4.3—once again the results are precise with relative errors of the VQE ground state energies of order of magnitude $10^{-2} - 10^{-3}$, however less precise than in the noninteracting case, which is expected since the former describes the system of essentially free fermions[1].

Similar calculations are carried out using COBYLA and SPSA optimizers, to the same end result—obtained values are precise to the same order of magnitude for most of the parameter values.

4.4 Full set of parameters

After the limiting cases were shown to work well, we consider the intermediate model parameter value pairs since these produce the most interesting behaviour in practice.

[1] Although the relative error values seem to increase toward stronger interaction U values, after the range depicted in Figure 4.3, they are found to plateau—this is consistent with the ease of solving the problem for even larger values of U we have considered separately beforehand.

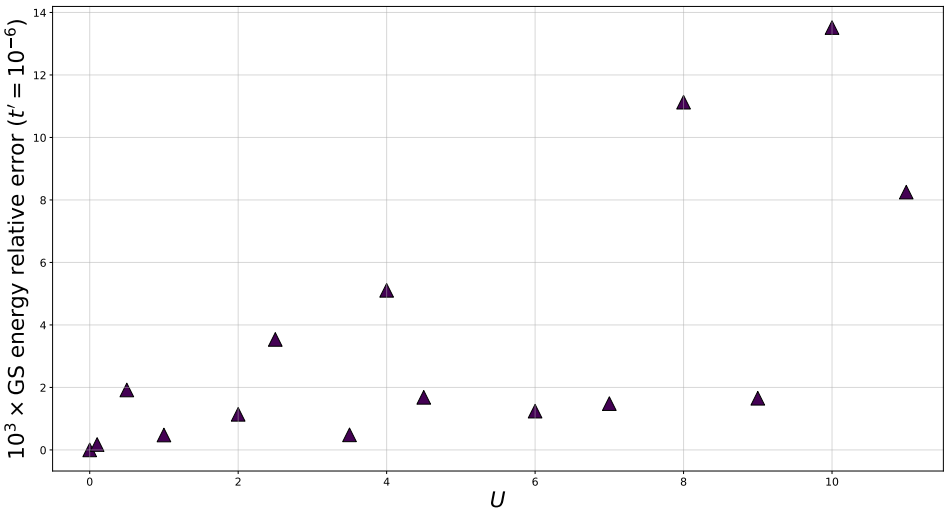


Figure 4.3: Relative error of the VQE ground state energies found using ansatz 3.5 with depth of 3 layers and SLSQP optimizer, for negligible hopping interaction strengths $t = 0$, $t' = 10^{-6}$.

First, we compare convergence speed (defined via the number of function calls during the optimization process) and relative errors on a somewhat reduced set of pairs of parameter values (around 100 of them). After the preliminary results, we run the optimizer shown to be the best on an extended set of parameters (around 200), to find how well our solution works with better coverage of the model parameter space. Regarding the convergence speed, we include only the numbers corresponding to cases in which the solution of up to a 2.5% relative error is found—otherwise this quantifier does not make much sense, it could just be a barren plateau, to which the optimizers can converge pretty fast in general and would thus skew the results.

The results of the preliminary ~100 calculations can be seen in Table 4.1. We find that our solution works well for a wide array of parameter value pairs for all three optimizers used, but that it works extremely well with SLSQP, as expected [79]. Similarly, we find that COBYLA converges the fastest—this was mentioned as its useful property that allows for fast iteration in all stages of ansätze building and evaluation; while SPSA converges the slowest, as is expected since its optimization procedure is carried out via stochastic perturbations and is thus usually not as well-directed as the other two—this can also be seen in Figure 4.4.

These calculations are done without (meta)optimization of the optimizers' hyperpa-

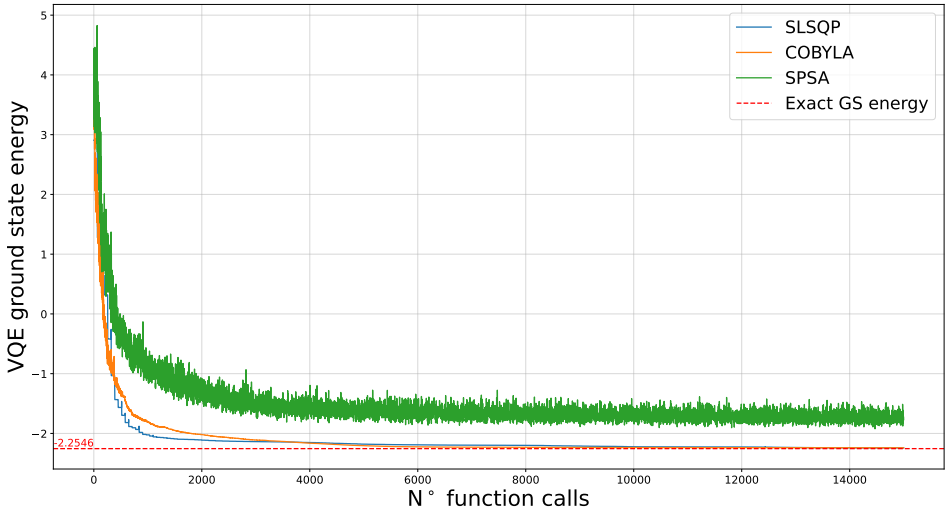


Figure 4.4: Speed of convergence comparison of SLSQP, COBYLA and SPSA for ansatz 3.5 with $L = 3$ layers and model parameter values $U = 4.0$ and $t' = 0.5$.

rameters, i.e. with all hyperparameters set to default values, as a way to avoid overengineering our solution that then may not be robust once the noise is introduced or the model size/interaction complexity is increased. However, when simple changes such as a slight ansatz depth increase or optimization of the learning rate of COBYLA and SPSA are employed, their average relative errors are also seen to drop under 1%. Since SLSQP is manifestly sufficient for the noiseless case, we will omit COBYLA's performance results after (meta)optimization and ansatz depth increase. However, we will carry out the improvement process for SPSA and use it to solve a model configuration using once the shot noise is introduced in the following section.

	COBYLA	SLSQP	SPSA
Avg. $N°$ function calls	1893	4222	12791
Avg. relative error	5.31%	0.16%	6.17%

Table 4.1: Number of function calls and relative errors of the ansatz 3.5 with $L = 3$ layers and different optimizers, averaged over parameter value pairs explored in Figure 4.5. Here the number of function calls is calculated as one needed for the ground state energy to converge within 2.5% of the exact result for the first time.

But before that, as was announced, we run SLSQP with the full set of ~200 model

parameter value pairs, with the results for the relative error values shown in Figure 4.5. We find that VQE ground state energies in most parts of the phase space are under 0.25% relative error, but even in the most demanding part, i.e. small t' and large/moderately large U, they are under 2% relative error. This is somewhat consistent with phase transition at $t = t'$ and qualitatively different behaviour on sides of this threshold, antiferromagnetic below it and metallic, insulating and one with incommensurate phase above it as U is increased [44, 98][2].

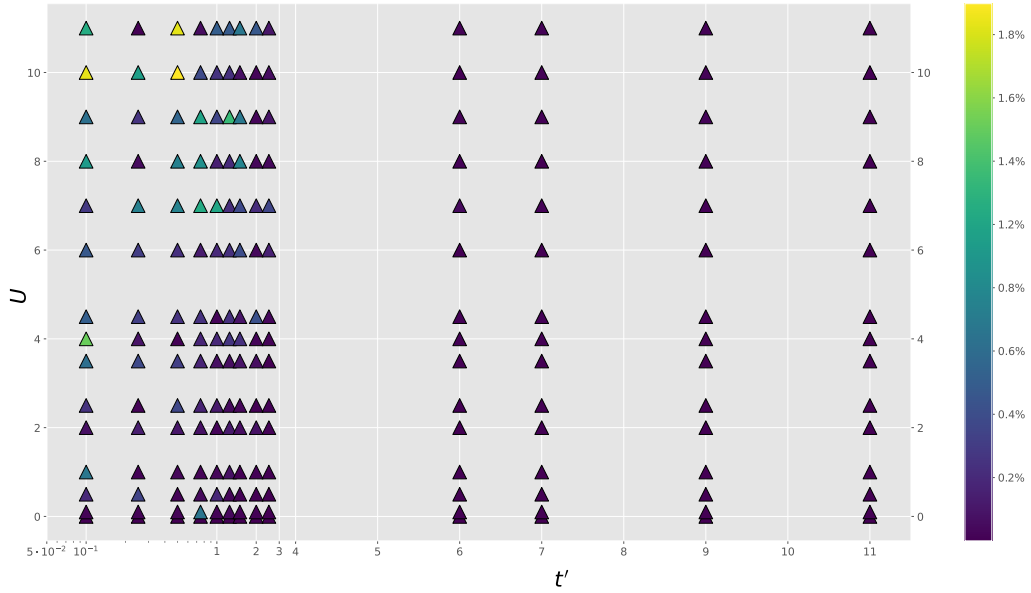


Figure 4.5: Relative error of ground state energy obtained by VQE and the exact value found by numerical diagonalization for SLSQP optimizer and different values of model parameters U and t'.

Similarly, we find from Figure 4.6 that the number of function calls using SLSQP is usually under $5 \cdot 10^3$, except in the aforementioned regime of small t' and large/moderately large U (in which it is still reasonably small and less than 10^4), completely consistent with the conclusions following Figure 4.5. A direct implication of this low number of function calls needed, i.e. fast convergence, is that the optimization process can be run on personal computing devices in short time (~10 minutes) and is thus easy to handle, improve and use in regular applications.

[2]But not particularly precisely, since an infeasibly dense array of parameter value pairs would be necessary to explore the behaviour around the phase transition exactly, which is not necessary for our purpose of building a correct solver, however a significant shift in relative error values happens in a moderately wide area around this range.

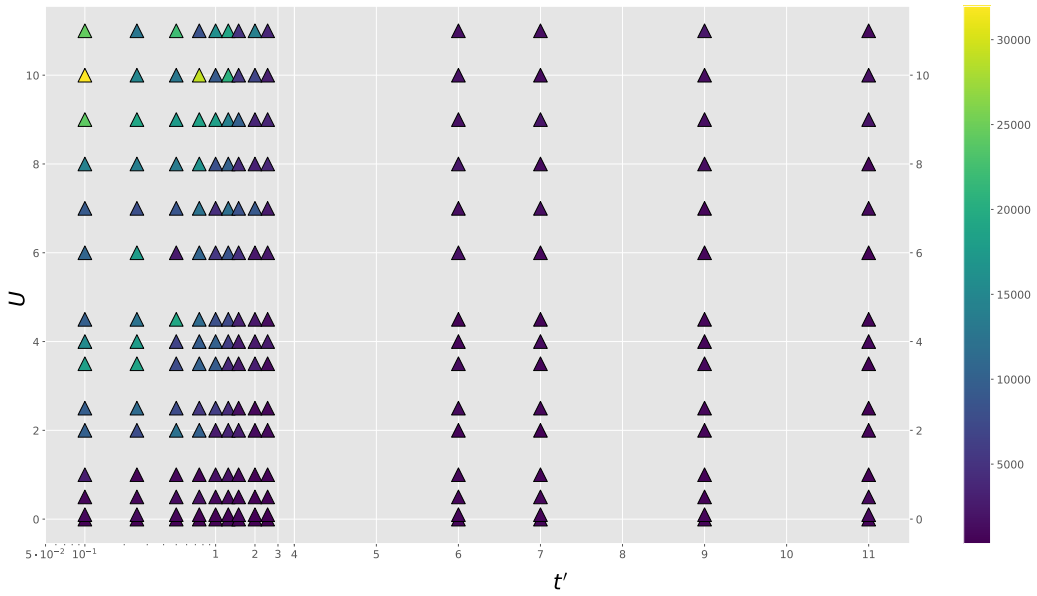


Figure 4.6: Number of function calls during optimization procedure for SLSQP optimizer and different values of model parameters U and t', as a quantifier of the convergence speed/rate of the optimization process.

As an additional note, we see that in VQE approach we can handle the different cases that are not susceptible to analytic results such as Lieb's 1 and Shen's 2 theorems, or the perturbative analysis and different symmetries outlined in Chapter 2. Thus, although the approach relies on system properties such as symmetries to some extent, we find it flexible and less constrained in its application scope.

4.5 Noisy calculations

Even before running the noisy simulations, COBYLA is expected to work poorly with shot noise included since it is based on an assumption that in each evaluation, the sampling is done for an infinite amount of shots, i.e. that the measurement is exact. Similarly, SLSQP is intended for optimization of twice differentiable objective functions and is thus also expected not to be resistant to noise introduction. On the other hand, due to its stochastic nature, SPSA is expected to work well in the noisy case. As announced in previous section, the 5% error reported in Table 4.1 should be lowered if we intend to reach extremely precise results such as those we did with SLSQP. To this end, we make several optimizations to our approach. First, we consider a set of different parameter values for moderate and strong onsite interactions and on each of them use Optuna, an efficient and precise hyperparameter

optimization framework to find the optimal values of two SPSA hyperparameters: learning rate and perturbation. Learning rate determines how aggresively will the parameter values be updated in each iteration (α in (3.5)), while perturbation determines the magnitude of the small random vector ($\mathbf{c}$ in (3.5)) in direction of which the update will be carried out. If these values are not set, then they are calibrated during a preliminary calculation phase, however using Optuna, we always find their near-optimal values in less than 100 trials, resulting in relative error of the VQE ground state energy less than 0.25% each time (this makes sense since based on the extremely precise results obtained with SLSQP, shown in Figure 4.5, the ansatz is evidently expressive and capable enough to produce the correct solution: the main issue is how well the optimization process is carried out). Based on the optimal hyperparameter values for several interaction strengths, we finally choose learning rate $\alpha = 0.1$ and perturbation $|\mathbf{c}| = 0.001$. Finally, in agreement with the lower right plot in Figure 4.1, there is an increase in entangling capability when increasing the ansatz depth from 3 to 4 layers, and more precise results are indeed found with SPSA for this depth (increasing further results in worse performance, as expected after a certain point since the number of ansatz parameters and thus the size of the parameter space is increased, so it is more difficult to find paths to a global minimum). Finally, due to the inherently stochastic nature of SPSA, each simulation is ran 5 times and the minimal value taken—this usually does not affect the performance with shot numbers for which the optimizer finds the correct solution easily, however for marginal shot numbers, VQE ground state energies can differ significantly across different runs with all hyperparameters and initial values held constant. Employing these three changes, we find that the SPSA is also able to produce an average relative error of less than 1%, so we move on to the noisy case with this setup directly[3].

The noisy calculations are carried out with SPSA for a limited set of parameter pair values, however it was found that an increasingly large number of shots was needed in the more demanding regime of small t' and large/moderately large U and thus the complete set of computations such as the one that was done in Figure 4.5 is outside of the computational scope of this thesis. However, the solution was shown to produce extremely precise values at a number of shots of order 10^6 or more[4] for all the considered parameter values in a set of pair products $(U, t') \in (1.0, 2.0, 4.0, 8.0) \times (0.1, 0.5, 1.0)$, with one such result shown in Figure 4.7, compared to the completely inaccurate SLSQP (even more shots up to 10^{10} was tried for SLSQP, to no avail).

[3]In many cases, pretraining the circuit in weak interaction regime was found to be of help, however mostly for small and moderate values of U.

[4]Of course, this number of shots is somewhat large, but using the improvements we described, as well as larger upper bound on the number of iterations of the optimizer, it can be lowered—the point is that it exists and thus renders the solution applicable in the noisy case.

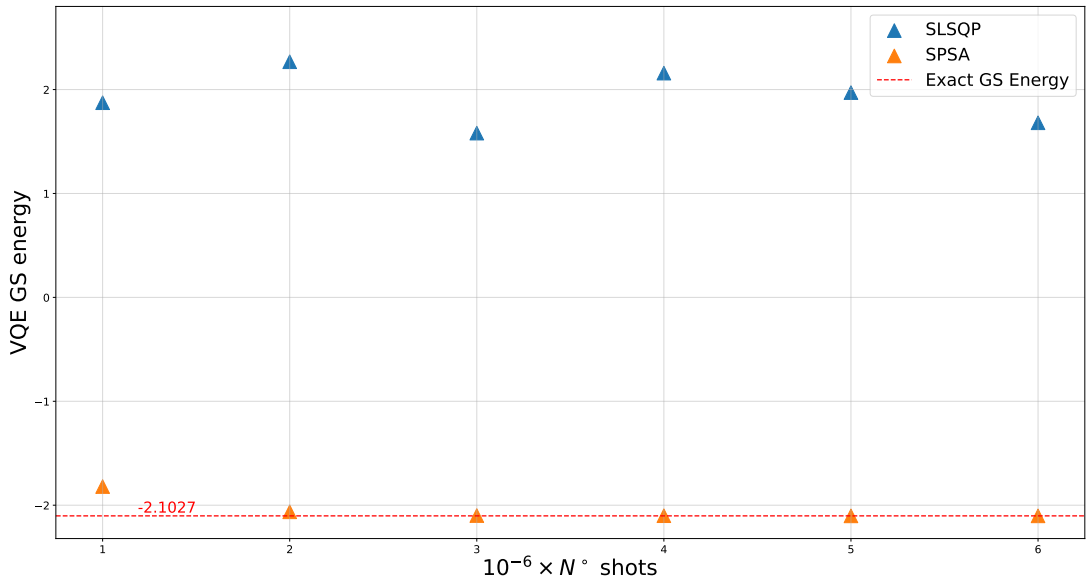


Figure 4.7: VQE ground state energies found using ansatz with $L = 4$ layers and optimizers SLSQP and SPSA with model parameters $U = 4.0$ and $t' = 0.1$.

4.6 Conclusion and future prospects

Following the procedure described in previous chapter, we have constructed a simple but powerful, efficient and robust ansatz for the Fermi-Hubbard model beyond the nearest neighbour interaction relying on previous research, quantum and geometric symmetries of the model and the physical interpretation of the terms in its Hamiltonian, while also demonstrating its ability to produce precise results once the shot noise is introduced into the system. Besides the calculations outlined in this work, further noiseless calculations for $N = 6$ demonstrate that it yields precise results for that system size as well, with noisy calculations to be considered next [1].

As some additional followups to this work, an interesting approach seems to be the employment of ansätze with adaptive structure since it has been shown [14, 99] that they can also help alleviate barren plateaus, so a combination of the current solution and adaptability is presumed to increase precision and robustness of our solution.

Similarly, methods for initial ansatz parameter selection based on machine learning models have shown promise in practice and could thus be used on the constructed ansatz directly as an additional preliminary step [100].

Finally, it is important to mention an important desire to construct a broader and more general VQE toolkit that can be applied to other solid state models beyond the Fermi-Hubbard Hamiltonian, and considerations outlined in this work are expected to be helpful

in this regard.

Bibliography

[1] T. A. Bespalova, K. Delić, G. Pupillo, F. Tacchino & I. Tavernelli. Unpublished. 2024.

[2] P. Ehrenfest. *Welche Züge der Lichtquantenhypothese spielen in der Theorie der Wärmestrahlung eine wesentliche Rolle?*. Ann. Phys. (Leipzig) (**36**). 1911.

[3] M. Planck. *Über irreversible Strahlungsvorgänge*. Ann. Physik (**1**). 1900.

[4] A. Einstein. *On a Heuristic Point of View Concerning the Production and Transformation of Light*. Ann. Phys. (**17**). 1905.

[5] W. Heisenberg. *Über quantentheoretische Umdeutung kinematischer und mechanischer Beziehungen*. Z. Phys. (**33**). 1925.

[6] E. Schrödinger. *An Undulatory Theory of the Mechanics of Atoms and Molecules*. Phys. Rev. (**28**). 1926.

[7] W. Heisenberg. *Über den anschaulichen Inhalt der quantentheoretischen Kinematik und Mechanik*. Z. Phys. (**43**). 1927.

[8] J. Bardeen & W. H. Brattain. *The Transistor, A Semi-Conductor Triode*. Phys. Rev. (**74**). 1948.

[9] W. Shockley. *The Theory of p-n Junctions in Semiconductors and p-n Junction Transistors*. The Bell System Technical Journal (**28**). 1949.

[10] A. M. Turing. *On Computable Numbers, with an Application to the Entscheidungsproblem*. Proc. London Math. Soc. (**s2-42**). 1937.

[11] R. P. Feynman. *Simulating physics with computers*. Int. J. Theor. Phys. (**21**). 1982.

[12] D. Deutsch. *Quantum theory, the Church–Turing principle and the universal quantum computer*. Proc. R. Soc. Lond. A (**400**). 1985.

[13] P. W. Shor. *Polynomial-Time Algorithms for Prime Factorization and Discrete Logarithms on a Quantum Computer*. SICOMP (**26**). 1997.

[14] J. Tilly, H. Chen, S. Cao, D. Picozzi, K. Setia, Y. Li, E. Grant, L. Wossnig, I. Rungger, G. H. Booth & J. Tennyson. *The Variational Quantum Eigensolver: A review of methods and best practices*. Phys. Rep. (**986**). 2022.

[15] P. Jurcevic, A. Javadi-Abhari, L. S. Bishop, I. Lauer, D. F. Bogorin, M. Brink, et al. *Demonstration of quantum volume 64 on a superconducting quantum computing system*. QST (**6**). 2021.

[16] S. Ebadi, T. T. Wang, H. Levine, A. Keesling, G. Semeghini, A. Omran, et al. *Quantum phases of matter on a 256-atom programmable quantum simulator*. Nature (**595**). 2021.

[17] L. K. Grover. *A fast quantum mechanical algorithm for database search*. arXiv:quant-ph/9605043. 1997.

[18] J. Preskill. *Quantum Computing in the NISQ era and beyond*. Quantum (**2**). 2018.

[19] M. Brooks. *Beyond quantum supremacy: the hunt for useful quantum computers*. Nature (**574**). 2019.

[20] S. J. Gustafson, I. M. Sigal. *Mathematical Concepts of Quantum Mechanics*. Springer Berlin, Heidelberg. 2011.

[21] M. Reed & B. Simon, *Methods of Mathematical Physics, Vol. 2: Fourier Analysis, Self-Adjointness*. Academic Press, Waltham. 1972.

[22] T. Jurić, *Observables in Quantum Mechanics and the Importance of Self-Adjointness*. Universe (**8**). 2022.

[23] L. de Broglie. *XXXV. A tentative theory of light quanta*. Lond. Edinb. Dublin philos. mag. j. sci. (**47**). 1924.

[24] W. Gerlach & O. Stern. *Der experimentelle Nachweis der Richtungsquantelung im Magnetfeld*. Z. Phys. (**9**). 1922.

[25] S. S. M. Wong. *Introductory Nuclear Physics: 2nd edition*. Wiley-VCH. 1999.

[26] A. Einstein, B. Podolsky & N. Rosen. *Can Quantum-Mechanical Description of Physical Reality Be Considered Complete?*. Phys. Rev. (**47**). 1935.

[27] J. S. Bell. *On the Einstein Podolsky Rosen paradox*. Phys. Phys. Fiz. (**1**). 1964.

[28] S. J. Freedman & J. F. Clauser. *Experimental Test of Local Hidden-Variable Theories*. Phys. Rev. Lett. (**28**). 1972.

[29] A. Aspect, P. Grangier & G. Roger. *Experimental Tests of Realistic Local Theories via Bell's Theorem*. Phys. Rev. Lett. (**47**). 1981.

[30] D. Bouwmeester, J.-W. Pan, M. Daniell, H. Weinfurter & A. Zeilinger. *Observation of Three-Photon Greenberger-Horne-Zeilinger Entanglement*. Phys. Rev. Lett. (**82**). 1999.

[31] M. A. Nielsen, I. L. Chuang. *Quantum Computation and Quantum Information: 10th Anniversary Edition*. Cambridge: Cambridge University Press. 2010.

[32] W. Pauli. *Über den Zusammenhang des Abschlusses der Elektronengruppen im Atom mit der Komplexstruktur der Spektren*. Z. Physik (**31**). 1925.

[33] C. K. Hong, Z. Y. Ou & L. Mandel. *Measurement of subpicosecond time intervals between two photons by interference*. Phys. Rev. Lett. (**59**). 1987.

[34] G. D. Mahan. *Many-Particle Physics*. Springer New York, NY. 2000.

[35] `https://physics.stackexchange.com/questions/330428/first-quantization-vs-second-quantization`

[36] K. Delić. *Towards quantum glasses: Disorder in topological frustrated quantum systems*. Master thesis. 2021.

[37] O. Regev. *An Efficient Quantum Factoring Algorithm*. arXiv:2308.06572. 2023.

[38] M. C. Gutzwiller. *Effect of Correlation on the Ferromagnetism of Transition Metals*. Phys. Rev. Lett. (**10**). 1963.

[39] J. Kanamori. *Electron Correlation and Ferromagnetism of Transition Metals*. Prog. Theor. Phys. (**30**). 1963

[40] J. Hubbard. *Electron correlations in narrow energy bands*. Proc. R. Soc. A: Math. Phys. Eng. Sci. (**276**). 1963.

[41] *The Hubbard model at half a century*. Nature Phys. (**9**). 2013.

[42] H. Shiba & P. A. Pincus. *Thermodynamic Properties of the One-Dimensional Half-Filled-Band Hubbard Model*. Phys. Rev. B (**5**). 1972.

[43] S. Stanišić, J. L. Bosse, F. M. Gambetta, et al. *Observing ground-state properties of the Fermi-Hubbard model using a scalable algorithm on a quantum computer*. Nat. Commun. (**13**). 2022.

[44] S. Daul & R. M. Noack. *Phase diagram of the half-filled Hubbard chain with next-nearest-neighbor hopping*. Phys. Rev. B (**61**). 2000.

[45] R. Schumann. *Thermodynamics of a 4-site Hubbard model by analytic diagonalization*. Ann. Phys. (**514**). 2001.

[46] D. P. Arovas, E. Berg, S. A. Kivelson & S. Raghu. *The Hubbard Model*. Annu. Rev. Condens. Matter Phys. (**13**). 2022.

[47] H. Yao and S. A. Kivelson. *Fragile Mott Insulators*. Phys. Rev. Lett. (**105**). 2010.

[48] E. H. Lieb & F. Y. Wu. *Absence of Mott Transition in an Exact Solution of the Short-Range, One-Band Model in One Dimension*. Phys. Rev. Lett. (**20**). 1968.

[49] F. A. Cotton. *Chemical Applications of Group Theory, 3rd Edition*. Wiley-Interscience. 1990.

[50] C. Chen, A. Reich & L. M. Falicov. *Surface properties of a heavy-fermion system: An exact many-body solution to a periodic-cluster Hubbard model*. Phys. Rev. B (**38**). 1988.

[51] J. K. Freericks, L. M. Falicov & D. S. Rokhsar. *Exact solutions of frustrated ordinary and chiral eight-site Hubbard models*. Phys. Rev. B (**44**). 1991.

[52] A. Reich & L. M. Falicov. *Heavy-fermion system: An exact many-body solution to a periodic-cluster Hubbard model*. Phys. Rev. B (**37**). 1988.

[53] O. J. Heilmann, E. H. Lieb. *Violation of the Noncrossing Rule: the Hubbard Hamiltonian for Benzene*. Ann. N. Y. Acad. Sci. (**172**). 1971.

[54] J. Yepez. *Lecture notes: Fermi-Hubbard model*. `https://www.phys.hawaii.edu/~yepez/Spring2013/lectures/Lecture8_Hubbard_Model_Notes.pdf`. 2013.

[55] C. M. Villet & W.-H. Steeb. *Four-Point Hubbard Model and Conserved Quantities*. J. Phys. Soc. Japan (**59**). 1990.

[56] C. Noce & M. Cuoco. *Exact-diagonalization method for correlated-electron models*. Phys. Rev. B (**54**). 1996.

[57] H. Grosse. *The symmetry of the Hubbard model*. Lett. Math. Phys. (**18**). 1989.

[58] A. Peruzzo, J. McClean, P. Shadbolt, et al. *A variational eigenvalue solver on a photonic quantum processor*. Nat. Commun. (**5**). 2014.

[59] J. R. McClean, J. Romero, R. Babbush & A. Aspuru-Guzik. *The theory of variational hybrid quantum-classical algorithms*. New J. Phys. (**18**). 2016.

[60] P. Deglmann, A. Schäfer & C. Lennartz. *Application of quantum calculations in the chemical industry—An overview*. Quantum Chem. (**115**). 2015.

[61] B. J. Williams-Noonan, E. Yuriev & D. K. Chalmers. *Free Energy Methods in Drug Design: Prospects of "Alchemical Perturbation" in Medicinal Chemistry Miniperspective*. J. Med. Chem. (**61**). 2018.

[62] M. A. Continentino. *Key Methods and Concepts in Condensed Matter Physics*. IOP Publishing. 2021.

[63] Y. Cao, J. Romero & A. Aspuru-Guzik. *Potential of quantum computing for drug discovery*. IBM J. Res. Dev. (**62**). 2018.

[64] V. Lordi, J. M. Nichol. *Advances and opportunities in materials science for scalable quantum computing*. MRS Bulletin (**46**). 2021.

[65] Y. Cao, et al. *Quantum Chemistry in the Age of Quantum Computing*. Chem. Rev. (**119**). 2019.

[66] Y. Zhou, E. M. Stoudenmire & X. Waintal. *What Limits the Simulation of Quantum Computers?*. Phys. Rev. X (**10**). 2020.

[67] A. Montanaro & S. Stanišić. *Error mitigation by training with fermionic linear optics*. arXiv:2102.02120. 2021.

[68] F. Arute, K. Arya, R. Babbush, et al. *Quantum supremacy using a programmable superconducting processor*. Nature (**574**). 2019.

[69] H.-S. Zhong et al. *Quantum computational advantage using photons*. Science (**370**). 2020.

[70] J. Rayleigh. *In finding the correction for the open end of an organ-pipe*. Phil. Trans. (**161**). 1870.

[71] W. Ritz. *Über eine neue Methode zur Lösung gewisser Variationsprobleme der mathematischen Physik*. J. für die Reine und Angew. Math. (**135**). 1909.

[72] P. Jordan, E. Wigner. *Über das Paulische Äquivalenzverbot*. Z. Physik (**47**). 1928.

[73] A. Aspuru-Guzik, A. D. Dutoi, P. J. Love & M. Head-Gordon. *Simulated Quantum Computation of Molecular Energies*. Science (**309**). 2005.

[74] M. Oliv, A. Matic, T. Messerer & J. M. Lorenz. *Evaluating the impact of noise on the performance of the Variational Quantum Eigensolver*. arXiv:2209.12803. 2022.

[75] A. Arrasmith, L. Cincio, R. D. Somma & P. J. Coles. *Operator Sampling for Shot-frugal Optimization in Variational Algorithms*. arXiv:2004.06252. 2020.

[76] D. Wecker, M. B. Hastings & M. Troyer. *Progress towards practical quantum variational algorithms*. Phys. Rev. A (**92**). 2015.

[77] L. Bittel & M. Kliesch. *Training Variational Quantum Algorithms Is NP-Hard*. Phys. Rev. Lett. (**127**). 2021.

[78] I. Miháliková, et al. *Best-Practice Aspects of Quantum-Computer Calculations: A Case Study of the Hydrogen Molecule*. Molecules (**27**). 2022.

[79] A. G. Rattew, S. Hu, M. Pistoia, R. Chen & S. Wood. *A Domain-agnostic, Noise-resistant, Hardware-efficient Evolutionary Variational Quantum Eigensolver*. arXiv:1910.09694. 2020.

[80] D. Kraft. *A Software Package for Sequential Quadratic Programming*. Wiss. Berichtswesen d. DFVLR. 1988.

[81] M. J. D. Powell. *A View of Algorithms for Optimization without Derivatives 1*. Math. Today (Southend-on-Sea) (**43.5**). 2007.

[82] J. C. Spall. *An overview of the simultaneous perturbation method for efficient optimization*. Johns Hopkins APL Tech. Dig. (**19**). 1998.

[83] S. Finck & H.-G. Beyer. *Performance analysis of the simultaneous perturbation stochastic approximation algorithm on the noisy sphere model*. Theor. Comput. Sci. (**419**). 2012.

[84] T. Durrani & G. Morison, *SPSA for noisy non-stationary blind source separation*. 2003 IEEE International Conference on Acoustics, Speech, and Signal Processing. 2003.

[85] L. Wang, J. Zhu & J. C. Spall. *Mixed Simultaneous Perturbation Stochastic Approximation for Gradient-Free Optimization with Noisy Measurements*. 2018 Annual American Control Conference (ACC). 2018.

[86] J. R. McClean, S. Boixo, V. N. Smelyanskiy, R. Babbush & H. Neven. *Barren plateaus in quantum neural network training landscapes*. Nat. Commun. (**9**). 2018.

[87] A. Kandala, A. Mezzacapo, K. Temme, et al. *Hardware-efficient variational quantum eigensolver for small molecules and quantum magnets*. Nature (**549**). 2017.

[88] I. G. Ryabinkin, S. N. Genin & A. F. Izmaylov. *Constrained Variational Quantum Eigensolver: Quantum Computer Search Engine in the Fock Space*. J. Chem. Theory Comput. (**15**). 2019.

[89] Z. Jiang, et al. *Quantum algorithms to simulate many-body physics of correlated fermions*. Phys. Rev. Appl. (**9**). 2018.

[90] Z. Holmes, K. Sharma, M. Cerezo & P. J. Coles. *Connecting Ansatz Expressibility to Gradient Magnitudes and Barren Plateaus*. PRX Quantum (**3**). 2022.

[91] S. Sim, P. D. Johnson & A. Aspuru-Guzik. *Expressibility and Entangling Capability of Parameterized Quantum Circuits for Hybrid Quantum-Classical Algorithms*. Adv. Quantum Technol. (**1**). 2019.

[92] D. A. Meyer & N. R. Wallach. *Global entanglement in multiparticle systems*. J. Math. Phys. (**43**). 2002.

[93] K. Okunishi. *Filling dependence of the zigzag Hubbard ladder for the quasi-one-dimensional superconductor*. $Pr_2Ba_4Cu_7O_{15-\delta}$. Phys. Rev. B (**75**). 2007.

[94] A. Kulshrestha & I. Safro. *BEINIT: Avoiding Barren Plateaus in Variational Quantum Algorithms*. arXiv:2204.13751. 2022.

[95] E. Grant, L. Wossnig, M. Ostaszewski & M. Benedetti. *An initialization strategy for addressing barren plateaus in parametrized quantum circuits*. Quantum (**3**). 2019.

[96] `https://obliviateandsurrender.github.io/blogs/expr.html`

[97] `https://qiskit.org/ecosystem/nature/tutorials/11_quadratic_hamiltonian_and_slater_determinants.html`

[98] L. F. Tocchio, F. Becca & C. Gros. *Interaction-induced Fermi-surface renormalization in the t_1-t_2 Hubbard model close to the Mott-Hubbard transition*. Phys. Rev. B (**81**). 2010.

[99] H. R. Grimsley, S. E. Economou, E. Barnes & N. J. Mayhall. *An adaptive variational algorithm for exact molecular simulations on a quantum computer*. Nat. Commun. (**10**). 2019.

[100] F. Sauvage, S. Sim, A. A. Kunitsa, W. A. Simon, M. Mauri & A. Perdomo-Ortiz. *FLIP: A flexible initializer for arbitrarily-sized parametrized quantum circuits*. arXiv:2103.08572. 2021.

Sažetak

U ovom radu istražujemo algoritamske pristupe primjenjive na kvantne sustave, posebno Fermi-Hubbardov model. Uz povijesni osvrt i motivaciju, dan je i kratki pregled potrebnih fizikalnih i računalnih koncepata. Nakon što je model predstavljen i neki od njegovih rubnih slučajeva analitički riješeni, pokazujemo kako koristiti klasične i kvantne simetrije da bismo njegov Hamiltonijan reducirali na lako rješive blokove. Intuicija i razumijevanje izgrađeni u ovom dijelu omogućavaju nam da postupno izgradimo ansatz za kvantni varijacijski algoritam. Nakon konstrukcije sklopa, provodimo simulacije koje otkrivaju da predloženo rješenje radi precizno s različitim optimizatorima, izvršava se u kratkom vremenu te je robusno na uvođenje šuma koji potječe od konačnog broja mjerenja. Konačno, ističemo nekoliko prirodnih smjerova za buduća istraživanja.

Summary

In this work, we explore algorithmic approaches applicable to quantum systems, specifically the Fermi-Hubbard model. After outlining the motivation to do so, we follow with a brief overview of necessary physical and computing concepts. Once the model is introduced and some of its limiting cases are solved analytically, we demonstrate how to use classical and quantum symmetries to reduce its Hamiltonian into easily solvable blocks. The intuition and understanding built doing so allow us to gradually build the quantum ansatz for the Variational Quantum Eigensolver. After constructing the circuit, we conduct simulations, revealing that the proposed solution works precisely with various optimizers and runs in a short time, all while being robust even when the shot noise is introduced. Finally, we outline some natural directions for future research.

Résumé

I was born in the beautiful city of Split on the 30th of October 1997. I attended both elementary and high school there, cultivating love for science and technology from a young age.

Throughout my life, I have participated in multiple competitions in physics, mathematics and computer science. Most notable results are the honourable mention at the International Physics Olympiad and medals from state competitions in high school; along with victories in the Lumen Data Science and Stem Games competitions in college.

Looking ahead, I aspire to continue my professional development in both academia and industry.