At the intersection of quantum computing and quantum chemistry

A dissertation presented

by

James Daniel Whitfield

to

The Department of Chemistry and Chemical Biology

in partial fulfillment of the requirements

for the degree of

Doctor of Philosophy

in the subject of

Chemical Physics

Harvard University

Cambridge, Massachusetts

November 2011

Thesis advisor   Author
Alán Aspuru-Guzik   **James Daniel Whitfield**

At the intersection of quantum computing and quantum chemistry

Abstract

Quantum chemistry is concerned with solving the Schrödinger equation for chemically relevant systems. This is typically done by making useful and systematic approximations which form the basis for model chemistries. A primary contribution of this dissertation, is taking concrete steps toward establishing a new model chemistry based on quantum computation. Electronic energies of the system can be efficiently obtained to fixed accuracy using quantum algorithms exploiting the ability of quantum computers to efficiently simulate the time evolution of quantum systems. The quantum circuits for simulation of arbitrary electronic Hamiltonians are given using quantum bits associated with single particle basis functions. This model chemistry is applied to hydrogen molecule as a case study where all necessary quantum circuits are clearly laid out. Furthermore, our collaboration to experimentally realize a simplified version of the hydrogen molecule quantum circuit is also included in this thesis. Finally, alternatives to the gate model of quantum computation are pursued by exploring models based on the quantum adiabatic theorem and on the generalization of random walks.

Citations to Previously Published Work

Chapter 2 is a largely unmodified contribution to an undergraduate textbook:

> *Mathematical modeling II: Quantum mechanics and spectroscopy*
> T. L. Story with chapter on electronic structure contributed by J. D. Whitfield
> Educational Publisher. To appear Fall 2011.

Parts of the review article,

> *Simulating chemistry with quantum computers*
> I. Kassal*, J. D. Whitfield* A. Perdomo-Ortiz, M.-H. Yung, and A. Aspuru-Guzik
> Annual Reviews of Physical Chemistry. Volume 62, Pages 185-207 (2011)

appear in many places throughout the thesis. Many of the major themes of the thesis are found in,

> *Simulation of electronic structure Hamiltonians using quantum computers*
> J. D. Whitfield, J. D. Biamonte, and A. Aspuru-Guzik
> Molecular Physics. Volume 109, 735 (2011)

and large portions of this publication have been incorporated into Chapters 5, 6, and 8. The description of the experimental simulation of the hydrogen molecule in Chapter 8 is largely based on

> *Towards quantum chemistry on a quantum computer*
> B. P. Lanyon, J. D. Whitfield, G. G. Gillet, M. E. Goggin, M. P. Almeida, I. Kassal, J. D. Biamonte, M. Mohseni, B. J. Powell, M. Barbieri, A. Aspuru-Guzik, and A. G. White
> Nature Chemistry. Volume 2, 106 (2010)

Chapters 9 and 10 apart from minor modifications appear respectively in,

> *Adiabatic quantum simulators*
> J. D. Biamonte, V. Bergholm, J. D. Whitfield, J. Fitzsimons, and A. Aspuru-Guzik
> AIP Advances. Volume 1, 022126 (2011)

> *Quantum stochastic walks: A generalization of classical random walks and quantum walks*
> J. D. Whitfield, C. A. Rodríguez-Rosario, and A. Aspuru-Guzik.
> Physical Review A. Volume 81, 022323 (2010)

Other work I was involved in but not explicitly included in the thesis are the articles

Simulation of classical thermal states on a quantum computer: A transfer-matrix approach
M.-H. Yung, D. Nagaj, J. D. Whitfield, and A. Aspuru-Guzik
Physical Review A. Volume 82, 060302(R) (2010)

Solving quantum ground-state problems with nuclear magnetic resonance
Z. Li, M.-H. Yung, H. Chen, D. Lu, J. D. Whitfield, X. Peng, A. Aspuru-Guzik, and J. Du
Accepted to Scientific Reports. Preprint available at arXiv:1106.0440

Contents

Acknowledgments

First, I want to acknowledge the people who have guided me thus far. Foremost, my advisor, Alán Aspuru-Guzik, for teaching me how to approach scientific problems, to pool resources, and to remain positive through it all. I cannot adequately express my thanks for the past 5 years. I'd like to thank, my Dad, James Whitfield, Jr. who stood behind me from the beginning and watched me put my best foot forward; my high school freshman mathematics teacher, Mike McRae, for teaching me how to enjoy mathematics; Professor Troy L. Story for believing in me all these years; Professor William A. Lester for the introduction to academic research; and Professors Eric J. Heller and Scott Aaronson for serving in my advising committee and challenging me think deeply about why we do science.

Second, I want to thank those who ventured the trenches with me. From Alán's group (or should I say research division) I learned immensely from discussions with César Rodriguez-Rosario, Man-Hong Yung, Ville Bergholm, David Tempel, Sergio Boixo, Jarrod McClean, Ivan Kassal, Patrick Rebentrost, Alejandro Perdomo-Ortiz, Mark Watson, Kenta Hongo, Joel Yung-Zhou, Sangwoo Shim, Roberto (Robo) Olivares-Amaya, Leslie Vogt, Johannes Hachmann, Roel Sánchez-Carrera, Ali Najmaie, Şule Atahan-Evrenk, Dmitrij Rappoport, Semion Saikin and visiting professors Carlos Amador-Bedolla, Salvador Venegas-Andraca, and especially Peter Love. Jacob Biamonte, Ben Lanyon, and Cody Nathan Jones have also contributed to my scientific development in their own ways.

Third, my friends outside the laboratory have made my overall experience one I won't regret. The long list includes, The W. E. B. DuBois Society, Harvard Minority Scientists Group, Toastmasters, Kunal Surana, Dunster House and all the Dunsterites, Professor and Mrs. Porter, Edwin Homan, Brian Roland and Alexander Shalack among many others.

Others who I wish to thank are the agencies that funded my research: The Center for Excitonics, The Army Research Office and Harvard's Graduate Prize Fellowship. Thanks to Prince Georges County, Maryland, forever home, for a wonderful place to grow up.

I have to thank my biggest fan, my mom, Wendelin Whitfield, for being a fantastic supporter for all of these years. Thank to my sister, Jamie, for being there for me and proofreading so many of my papers and essays over the years. Thank you to my whole family for everything. Lastly, but, far from least, I have to thank my beautiful and supportive wife, Leslie Whitfield, for standing beside me through the whole process.

Dedication

Dedicated all the things seen and unseen for the space and time to look around.

Chapter 1

Introduction

You might not realize it, but the introduction should introduce the concepts, backgrouand[sic], and goals of the dissertation.

The Elements of Theses(pg 2), P. S. Disdainful 1991, retrieved from ctan.org:uctext.tex

1.1 Overview

A primary goal of this dissertation is to explore computational chemistry using a new model chemistry based on quantum simulation [15]. In the context of this thesis, quantum simulation takes on a very specific meaning whereby quantum simulation is *the mapping of the dynamics of the quantum system under study to another, more controllable one.* In particular, I will focus on instances where the controllable quantum system is a quantum computer. This model chemistry is unlike the vast majority of modern approaches which rely on conventional computational devices.

The motivation for connecting model chemistry and quantum computation is the inherent difficulty of solving the Schrödinger equation for large systems. In chemistry, the most important task is to calculate the energy of molecules to within chemical accuracy $(k_B T \sim 1 \text{ kJ mol}^{-1})$, which is required to predict reaction rates, equilibrium geometries, transition states and many other properties. The difficulty of solving this problem exactly has led to the development of many first-principles and semi-empirical model chemistries as discussed in Chapter 2. Tremendous methodological progress and the exponential growth of the power of classical computers have allowed for spectacular computational successes, such as density functional theory (DFT) calculations of entire proteins or nanostructures [94, 45]. Nevertheless, there are numerous important instances where approximate methods fail. For example, DFT is unreliable for systems involving excited-state charge transfer [68], conical intersections [134] which are important in green fluorescent protein and rhodopsin, or strong electronic correlations, such as those in high-temperature superconductors [129]. Although some of these effects are captured by post-Hartree-Fock wave function methods, there are weaknesses associated with these methods as well. For example, coupled-cluster methods fail for bond dissociation [222].

The price paid by all known model chemistries is either the truncation of the Hilbert space or an uncontrolled approximation. Many model chemistries developed for

computational chemistry are forced to rely on intelligently truncating the Hilbert space using methods such as the density matrix renormalization group [44]. Because the full Hilbert space is not treated, there are always states that cannot be accurately described using a given truncation. The hope is that these states are rare, in some way unphysical or uninteresting. DFT introduces a model chemistry based on the single particle density but requires an uncontrolled approximation involving the exchange-correlation functional discussed in the next chapter. Using sampling techniques grouped under the umbrella term quantum Monte Carlo [131], the 'dimensional curse' can be circumvented. However, treating fermions involves the uncontrolled fixed-node approximation. Moreover, these sampling techniques are usually limited to imaginary time propagation.

Theoretical computer science suggests that these limitations are not mere shortcomings of the algorithm designers and programmers but could stem from the inherent difficultly of simulating quantum systems. Extensions of computer science and information processing using quantum mechanics, see Chapter 3, has led to new ways of understanding the limitations of computational power. Interestingly, this new perspective helps us understand previously characterized model chemistries in a new light. In Chapter 4, I will review results from quantum computational complexity literature on common model chemistries including Hartree-Fock and density functional theory.

In 1982, Richard Feynman [75] suggested that an avenue for the tractable simulation of quantum systems would be to build quantum mechanical simulators. Quantum simulation is the idea of using quantum computational devices for more efficient simulation [75, 139]. The most versatile quantum simulator is a universal quantum computer—a device that uses quantum systems themselves to store and process data. Further developing a model chemistry based on the Feynman's proposal is the focus of this thesis.

Any model chemistry has several stages of development [173]: formulation, implementation, verification/validation, and prediction. In this thesis, I will describe progress through the first two stages, laying the ground work for further development until quantum simulation becomes routine enough to allow for validation across well characterized molecular data. Once validation has been accomplished, the model chemistry developed in this thesis will be useful for the most practical stage: prediction.

The formulation of the model chemistry will be done in Chapters: 5-7. Chapter 5 concerns the measurement technique, phase estimation, which is at the heart of the quantum simulator's advantage. Specifically, due to the properties of quantum measurement, the collapse of an approximate wave function to the exact wave function avoids many of the difficulties associated with optimization of variational parameters found in many of the current model chemistries. Next, Chapter 6 details how to time-propagate an approximate wave function so that the energy can be retrieved during the measurement scheme. Last, our description of the model chemistry concludes with Chapter 7, which describes state preparation techniques.

The second stage, implementation, has been achieved for the hydrogen molecule both theoretically and experimentally. In Chapter 8, I describe the application of the model chemistry developed to the hydrogen molecule. First, I present the theoretical model of hydrogen exemplifying the ideas in Chapters 6 and, then, I describe our experimental collaboration that implemented hydrogen molecule using quantum optics and computed the

energy at various bond lengths along the dissociation curve. This was the first experimental demonstration of quantum simulation for chemistry.

The next two chapters explore different approaches to computational control. In Chapter 9, an alternative scheme for quantum simulation based on a 'passive' model of computational control, adiabatic quantum computation, is presented. Finally, a more abstract study of quantum Markov chains is described in Chapter 10 in the context of random walks. Finally, conclusions and directions for further research are given in the final chapter.

Chapter 2

Overview of electronic structure methods

"The underlying physical laws necessary for the mathematical theory of a large part of physics and the whole of chemistry are thus completely known, and the difficulty is only that the exact application of these laws lead to equations much too complicated to be soluble."

–P.A.M. Dirac
Proc. R. Soc. London A. vol. 33 p 714 (1929)

2.1 Introduction

Since their introduction, computers have been used heavily in chemistry. Except in standard cases, such as the particle in a box, the harmonic oscillator, or the hydrogen atom, analytic solutions are not known and computational methods have been developed. Due to the exponential growth (known as Moore's law) of computing power since the introduction of semiconductor technology, computers are now routinely employed in chemistry for both qualitative and quantitative results.[1] Computable quantities include equilibrium geometries, transition state complexes, dipole moments, energy differences and potential energy surfaces.

Although computers have tremendously aided our understanding of chemical systems and their processes, the computational cost of the numerical methods for solving Schrödinger's equation grows rapidly with increases in the size of the system and the quality of its description. In this chapter, several avenues to obtaining chemical information using a classical computer are explored, providing a backdrop upon which model chemistry using quantum computers will be laid.

The material found in this chapter is standard and further information can be found in various textbooks including [60, 209, 117, 89].

---

[1] This exponential scaling will be difficult to continue past the middle of this century due to realistic constraints on how small transistors can be e.g. making transistors smaller than atoms is not likely. This limitation is one of the motivating factors in pursuing quantum computing for chemistry.

Quantity	Atomic Units	SI units
Reduced Planck's constant	$1\ \hbar$	1.055×10^{-36} J s
Mass of electron	$1\ m_e$	9.109×10^{-31} kg
Proton charge	$1\ e$	1.602×10^{-19} C
Bohr radius	$1\ a_0$	5.292×10^{-11} m
Hartree Energy	$1\ E_h$	4.360×10^{-18} J

Table 2.1: Atomic units and the corresponding SI values.

2.2 Electronic structure: Hamiltonians and wave functions

2.2.1 Atomic units

When working with atoms and molecules, the unit system known as atomic units is particularly convenient and will be used throughout the thesis. In these units, the elementary charge on a proton, the mass of an electron, Planck's reduced constant, and the electrostatic force constant found in Coulomb's law, $\frac{1}{4\pi\epsilon_0}$, are all set to unity. In Table 2.2.1, the corresponding SI values are given. Most of these constants appear in the solution of the hydrogen atom. For instance, the atomic unit of length, the Bohr radius $a_0 = \frac{\hbar^2}{m_e e^2}$, is the most likely distance from the nucleus to find the electron in the ground state of the hydrogen atom (ignoring reduced mass corrections) and theionization energy of the hydrogen atom is half an atomic unit.

2.2.2 The Born-Oppenheimer approximation

The full Hamiltonian of a molecular system (without relativistic corrections) in atomic units is:

$$
\begin{aligned}
H_{full} \;=\; & \sum_i \left[-\frac{1}{2}\nabla_i^2 + \sum_{j:i>j} \frac{1}{r_{ij}} - \sum_A \frac{Z_A}{r_{iA}} \right] \\
& + \sum_A \left[-\frac{1}{2M_A}\nabla_A^2 + \sum_{B:B>A} \frac{Z_A Z_B}{r_{AB}} \right]
\end{aligned} \tag{2.1}
$$

In Eq. (2.1), lowercase indices label electrons, capital indices label nuclei. Z_A and M_A are the nuclear charge and mass of nucleus A, respectively. The positive numbers r_{AB}, r_{iA}, and r_{ij} are the Euclidean distances between nuclei A and B, between electron i and nucleus A, and between electrons i and j, respectively.

To simplify the problem, the robust Born-Oppenheimer approximation is invoked. The Born-Oppenheimer approximation assumes that the wave function of a molecule can

be expressed as a product of the nuclear wave function and electronic wave function (parameterized by the nuclear coordinates) due to the difference in mass between electrons and nuclei. This approximation allows the time-independent Schrödinger equation for the electronic wave function to be solved separately from the Schrödinger equation for the nuclei. Since the nuclei are fixed while solving the electronic problem, this is sometimes called 'the clamped nuclei approximation.' The justification of the approximation comes from the separation of the time scales between the nuclei and the electrons since the typical mass of nuclei is several thousand times larger than the electron mass. Writing the electronic wave function with $\mathbf{R}_n$ as parameters assumes that the electrons respond instantaneously to changes in nuclear coordinates and is sometimes called the 'adiabatic approximation' to reflect this assumption.

Fixing the nuclei's location allows the electronic structure problem to be isolated from the nuclear motion. The electronic wave function $\Psi(\mathbf{x_e}; \mathbf{R_n})$ is governed by the electronic Hamiltonian

$$H_{elec}(\mathbf{R}_n) = -\frac{1}{2}\sum_i \nabla_i^2 - \sum_{i,A} \frac{Z_A}{r_{iA}} + \sum_{i>j} \frac{1}{r_{ij}} = T_e + V_{en} + V_{ee} \tag{2.2}$$

where the nuclear configuration, $\mathbf{R}_n$, has become a parameter for the electronic problem through the static potential due to the nuclei, V_{en}. The eigenvalues of H_{elec} are the electronic energies.

Once the kth electronic energy, $E_{el,k}(\mathbf{R}_n)$, has been solved at many values of $\mathbf{R}_n$, the nuclear wave function can then be solved using the potential energy surface abbreviated PES. The PES

$$U_k(\mathbf{R}_n) = E_{el,k}(\mathbf{R}_n) + V_{nn}(\mathbf{R}_n)$$

is the average potential generated by the electrons. The PES is a function of all the coordinates of the nuclei but due to the Born-Oppenheimer approximation, it has no dependence on the individual electronic coordinates since they have been averaged over. Solving $H_{nuc} = T_n + U_k(\mathbf{R}_n)$ provides the nuclear energy levels which are superimposed on each electronic energy level k. Given access to the full PES, reaction rates, equilibrium geometries, and transition states are among the properties that can be calculated.

Since the PES requires the electronic energy at many points, the focus of this chapter is solving for the electronic energies $E_{el,k}(\mathbf{R}_n)$; this is the electronic structure problem. Typically, the ground state energy is the focus as excited electronic states are usually not populated at thermal equilibrium; however in photochemistry, excited states often become important as optical photons cause transitions to excited states.

2.2.3   Wave functions in electronic structure

Wave functions and molecular orbitals

Linear combinations of atomic orbitals are used to specify molecular orbitals. In Section 2.4, the algorithm for computing the expansion coefficients is given. Each molecular orbital is a single-electron wave function and N of these orbitals must be combined to obtain N-electron wave functions. Since orbitals are single electron wave functions, neither

molecular orbitals nor atomic orbitals describe the actual system if it contains more than 1 electron; it is the N-body wave functions that serves this purpose. However, the molecular orbital picture of chemistry is pervasive because, as discussed later, the molecular orbitals are used to construct good approximations to the N-electron wave function.

The N-electron wave function is required to be antisymmetric; if two electrons are exchanged, the wave function should change sign from positive to negative. J. C. Slater noted that the determinant is antisymmetric because if two rows are interchanged, the determinant changes sign. Thus, the required anti-symmetry of the N-electron wave function can be accomplished by identifying rows with electronic coordinates x_k and columns with orbitals u_j. N-electron wave functions of this form are called Slater determinants.

$$\Psi(x_1, x_1, ..., x_N) = \frac{1}{\sqrt{N!}} \begin{vmatrix} u_i(x_1) & u_j(x_1) & \cdots & u_m(x_1) \\ u_i(x_2) & u_j(x_2) & \cdots & u_m(x_2) \\ . & & \cdots & . \\ . & & \cdots & . \\ . & & \cdots & . \\ u_i(x_N) & u_j(x_N) & \cdots & u_m(x_N) \end{vmatrix} \tag{2.3}$$

In Eq. (2.3), the pre-factor, $1/\sqrt{N!}$, is required to normalize the wave function, and the dependence on the nuclear coordinates has been suppressed. Introducing new terminology and notation, if orbitals $\{u_i, u_j, ..., u_m\}$ are selected from a set of M orbitals and combined using a determinant, it will be denoted $SD[u_i u_j ... u_m]$ and the orbitals $\{u_i u_j ... u_m\}$ are said to be the occupied orbitals and the remaining unused orbitals are called virtual or unoccupied orbitals. Notice that the ith electron coordinate, x_i, appears in every orbital; this is a consequence of the fact that the ith electron cannot *really* be labeled since all electrons are indistinguishable.

Due to the intrinsic spin of electrons, orbitals may be spin orbitals or spatial orbitals. Spatial orbitals only depend on the Cartesian position of the electron, r, while spin orbitals depend on both spin and position. Since this chapter only concerns non-relativistic Hamiltonians, the value of the electron's spin will be either $m_s = 1/2$ or $m_s = -1/2$ corresponding to spin up and spin down respectively. Throughout the chapter, spin-orbitals are denoted $u(x)$ with $x = (r, m_s)$ and spatial orbitals such as the 1s orbital of hydrogen will be denoted $\phi(r)$.

The spin-orbital includes one of two spin functions commonly denoted α or β. The spin functions $\alpha(m_s)$, $\beta(m_s)$ are the eigenfunction of S_z for spin-1/2 particles. It is important to note that the two spin functions are orthonormal: $\langle \alpha | \beta \rangle = \sum_{m_s=-1/2}^{1/2} \alpha^*(m_s)\beta(m_s) = 0$ and $\langle \alpha | \alpha \rangle = \langle \beta | \beta \rangle = 1$. Letting σ stand for a generic spin function (either α or β), the spin-orbital $u(x)$ can be written as $\phi(r)\sigma(m_s)$.

Examples of spatial orbitals include plane waves and the single-electron solutions to hydrogen's Schrödinger equation. In solid state physics, plane waves are best for describing the periodic structure of solids, while in chemistry, the atomic orbitals of hydrogen provide a starting point for discussion of molecular orbitals.

Single-electron basis sets

The hydrogen wave functions have the form $\psi_{nlm}(\mathrm{r}, \theta, \phi) = \mathrm{R}_{nl}(\mathrm{r}) \mathrm{Y}_{ml}(\theta, \phi)$. The radial part of the hydrogenic wave function, $\mathrm{R}_{nl}(\mathrm{r})$, is a polynomial in r, the distance from the origin, with a nodal structure that is mundane for the hydrogen atom but is not necessarily relevant for other atomic and molecular systems. Simplifying this polynomial leads to Slater Type Orbitals (STO) given as

$$\phi_{STO}(\mathrm{r}_A, \theta, \phi; \; n, m, l, \zeta) = N \mathrm{r}_A^{n-1} \exp(-\zeta \, \mathrm{r}_A) \, Y_{ml}(\theta, \phi) \tag{2.4}$$

Here N is a normalization constant, and the spherical coordinate system $(\mathrm{r}_A, \theta, \phi)$ is defined around atom A. The parameter, ζ, is called the orbital exponent. The orbital exponent is related to the nuclear charge and its deviation from the nuclear charge can be attributed to screening effects. Large values of orbital exponents lead to small radial extent, and small values lead to large radial extent.

The nearly universally used Gaussian Type Orbitals (GTO) were introduced by S. F. Boys in 1950. Similar functions are encountered in probability theory called normal distributions. Gaussian type basis functions replace $\exp(-\zeta \mathrm{r})$ with $\exp(-\alpha \mathrm{r}^2)$ giving spherical Gaussians the form

$$\phi_{sGTO}(\mathrm{r}, \theta, \phi; \; n, m, l, \alpha) = N \mathrm{r}_A^{n-1} \exp(-\alpha \, \mathrm{r}^2) \, Y_{ml}(\theta, \phi)$$

and Cartesian Gaussians having the form

$$\phi_{GTO}(r; i, j, k, \alpha) = N \, X_A^i Y_A^j Z_A^k \exp(-\alpha \, |\mathbf{R}_A - r|^2)$$

Again, N is the normalization constant and the coordinate $\mathbf{R}_A = (X_A, Y_A, \; Z_A)$ specifies the center of the GTO and $r = (x, y, z)$ represents the Cartesian coordinate where the GTO is being evaluated. Cartesian Gaussians are mainly used in practice and the abbreviation GTO will be reserved for the Cartesian Gaussians. The Gaussian orbital exponent, α, is similar to the Slater orbital exponent in that it controls the radial extent of the function.

The integers i, j, and k are not quantum numbers, but their sum indicates the angular momentum of the GTO. Letting

$$L = i + j + k \tag{2.5}$$

when $L=0$, the GTO is s-type, $L=1$ is p-type, $L=2$ is d-type and so on. Notice that for d-type orbitals with $i + j + k = 2$ there are six possible GTOs: X_A^2, Y_A^2, Z_A^2 , $X_A Y_A$, $X_A Z_A, Y_A Z_A$, but there are only five d-type orbitals. Using appropriate linear combinations of the six possible GTOs, five orthogonal orbitals with two quanta of angular momentum can be found while the 6th has spherical symmetry. For larger angular momentum, the number of Cartesian Gaussians is also greater than the number of spherical harmonics and a similar transformation can be done. Different conventions are in use regarding the inclusion of the redundant GTOs in the single-electron basis set.

In comparing STO and GTO, one on hand, the STO has the correct cusp as the distance from the center goes to zero, and, as the distance increases, it decays similar to the hydrogenic functions. On the other hand, Gaussians have the incorrect behavior at small r
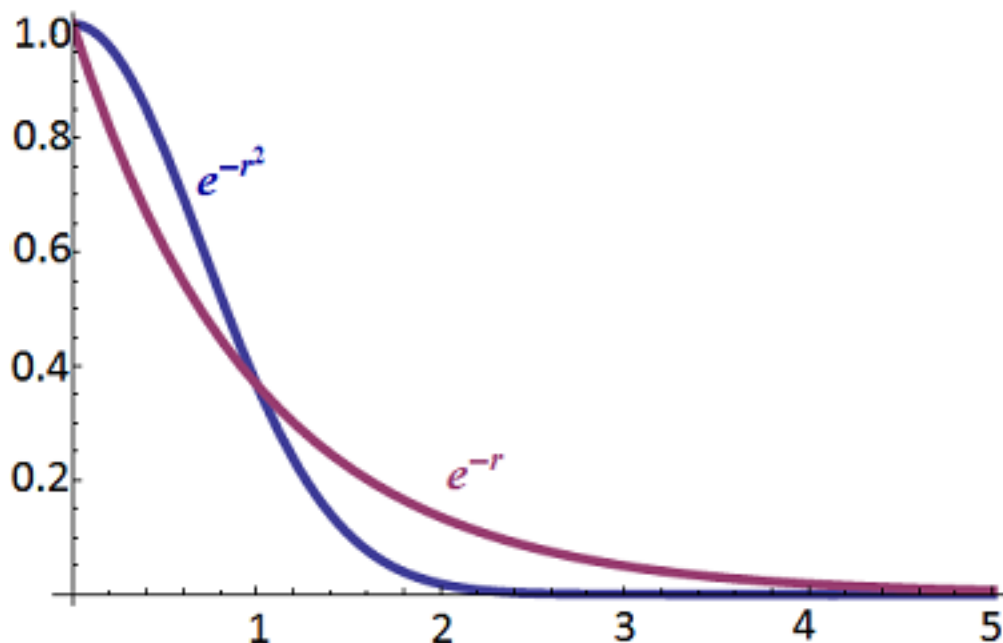
Figure 2.1: Comparison of GTO and STO orbitals.

and decay too rapidly at large r. See Fig. 2.1 for an illustration. The purpose of introducing the Gaussian functions is more for practical consideration. Their mathematical properties allow for much faster evaluation of the integrals that will be introduced shortly.

The incorrect cusp and long range decay of the GTO can be fixed by using a linear combination of Gaussian functions each with a unique orbital exponent to form a contracted Gaussian (CGTO) basis function

$$\phi_{CGTO}(r) = \sum_k d_k \phi_{GTO}(r; \alpha_k).$$

The GTOs in the contraction are called primitives. It is important to note that the orbital exponents, $\{\alpha_k\}$, and contraction coefficients, $\{d_k\}$, of the CGTO and the orbital exponents of the STOs and GTOs are not manipulated when optimizing the N-electron wave function. The contraction coefficients and the orbital exponents have been computed for many different computational purposes (geometry optimizations, thermochemistry, etc.) and are freely available from online databases.

The number of basis functions used in a calculation dictates the computational difficulty of the task, thus basis set selection is an important decision before embarking on a computation. Because the computational cost scales with the number of functions in the basis set, an infinite basis set cannot be used. However, the basis set limit is the value that the computational method would have returned if an infinite basis was used. This can be estimated using a basis set extrapolation to the infinite basis set limit. Next, basis set classification schemes are discussed.

The STO, GTO, and CGTO offer an array of options. The idea of model chemistries was pioneered by John A. Pople, who shared the 1998 Nobel Prize in chemistry (with W. Kohn), for his contributions to electronic structure. A model chemistry is a uniform level of calculation whose strengths and weaknesses can be assessed by comparison to experiment. By applying a model chemistry to a large number of molecular systems, trends in the errors and breakdowns in the approximation are discovered providing insights into when and how to use that level of theory. The establishment of practical model chemistries requires that the techniques and the basis sets be specified in standardized ways. The notation for basis sets is discussed during the remainder of this subsection.

A minimal basis is a basis where one function (STO, CGTO, or GTO) is used per atomic orbital. Using lithium as an example, the atomic orbitals are 1s, 2s, $2p_x$, $2p_y$, $2p_z$ therefore, a minimal basis requires five basis functions. Going beyond the minimal basis, using physical insights, the valence electron orbitals are the most important for correctly modeling chemistry. In some cases, when the core electrons are too costly to simulate or moving at relativistic speeds, effective core potentials (ECP) (also known as pseudopotentials) can be used to replace the core electrons with an effective potential so only the valence electrons are treated explicitly using non-relativistic equations.

A split valence basis set uses two or more functions for each valence atomic orbital. When two STOs with different orbital exponents, ζ, are used to represent each valence atomic orbital, the basis is called double zeta.[2] Triple zeta and quadruple zeta have 3 and 4 unique orbital exponents, respectively.

Earlier, it was pointed out that some of the deficiencies of the GTO can be overcome using contracted Gaussians which are linear combinations of primitive Gaussians. When m primitive GTOs are numerically fit to each STO of a minimal basis set, the Pople notation for this basis set is STO-mG. The more Gaussians used, the more accurately the STO is reproduced. For split valence basis sets with a contraction of m Gaussians for each core atomic orbitals and two CGTOs for each valence function with i and j GTOs used in each respective contraction, the notation m-ijG is used.

For instance, using the 6-31G basis, lithium's 1s orbital is described by a single CGTO orbital composed of 6 primitive GTO orbitals. For 2s and 2p (with three components) there are two orbitals: one CGTO with 3 primitives and a second uncontracted GTO.

For a triple zeta basis, the notation used is m-ijkG, e.g. 6-311G, where the first number indicates how many primitives to use for the core and how many primitives to use for each of the three basis functions for each valence orbital.

For particular purposes, the basis set may need to be augmented with special functions. First, when considering the polarization of an orbital that occurs when atoms bond or when an electric field is applied, polarization functions should be included. Think of the example of the minimal basis hydrogen atom with its sole 1s orbital being distorted by interaction with some external potential field. Since all s functions are spherically symmetric, using multiple s functions to describe the hydrogen's valence orbital cannot describe the distortion, whereas a p orbital allows for asymmetric charge distortion. In this case, the

---

[2] ζ is the greek letter 'zeta.'

p orbitals are called polarization functions. For more flexibility, polarization functions of higher angular momentum may also be included. When describing the basis using Pople notation, '*' denotes the inclusion of polarization functions for all atoms besides hydrogen and '**' indicates that all atoms including hydrogen have polarization functions e.g. 6-31G* and 6-31G**.

When describing dispersion forces, anions, and other situations where electrons will likely be found far from the nucleus, diffuse functions are used to augment the basis set. Diffuse functions have small orbital exponents and much larger radial extents. The Pople basis notation uses '+' to indicate when diffuse functions are included for all atoms except hydrogen and '++' when hydrogen also has a diffusion function.

To facilitate basis set extrapolation, Thom H. Dunning introduced the correlation consistent polarized valence X zeta basis with X= double, triple, etc. abbreviated cc-pVxZ with x=D,T,Q,5... The term 'correlation consistent' indicates that the basis is optimized for correlated methods (beyond Hartree-Fock) that are discussed from Subsection 2.4.5 forward. 'Polarized' means that polarization functions are included and the 'VxZ' indicates the number of functions used per valence atomic orbital. Like the other split valence basis sets, the core orbitals are described with a single function.

2.2.4   Matrix elements of the electronic Hamiltonian

Moving the beyond single electron picture of the previous subsection, this subsection evaluates the matrix elements of the N-electron Hamiltonian in terms of the single-electron orbital basis set chosen to describe the system. The Hamiltonian can be separated into two terms to simplify the derivation with the single electron Hamiltonian, $H^{(1)} = T_e + V_{ne}$, acting on only one electron of the system and $H^{(2)} = V_{ee}$ acting on two electrons. The kinetic energy of an electron depends only on that particular electron and the electron-nuclear potential also depends only on a single electron since the nuclear configuration has been fixed. This is contrasted with the electron-electron potential which couples electrons i and j through the $r_{ij} = |r_i - r_j|$ distance.

To evaluate the matrix elements of Hamiltonian, defining the one- and two-electron integrals will simplify notation. One-electron integrals are defined as

$$h_{ij} = \langle u_i | (T_e + V_{en}) | u_j \rangle = \int dx_1 \, u_i^*(x_1) \left(-\frac{1}{2}\nabla_1^2 - \sum_A \frac{Z_A}{r_{1A}} \right) u_j(x_1) \qquad (2.6)$$

Note that the spin has been included in the integration. The spin is simple to integrate over, and this is often done when the same spatial part of the wave function is used for both the α and β components. Similarly,

$$h_{ijkl} = \int dx_1 dx_2 \, \frac{u_i^*(x_1)u_j^*(x_2)u_k(x_2)u_l(x_1)}{r_{12}} = \langle u_i | \langle u_j | V_{ee} | u_k \rangle | u_k \rangle \qquad (2.7)$$

The symmetry $h_{ijkl} = h_{jilk}$ follows since relabeling electrons does not change the value of the integral. Additionally, when the spin-orbitals are real, $h_{ijkl} = h_{ljki} = h_{lkji} = h_{ikjl}$. Be forewarned that indexing of these integrals in this chapter is not standard but will

be used later in Chapter 8. Typically, one encounters the following two notations for the two-electron integrals

$$[ij|kl] = \int dx_1 dx_2 \, \frac{u_i^*(x_1)u_j(x_1)u_k^*(x_2)u_l(x_2)}{r_{12}}$$

which is called the "chemists' notation" and

$$\langle ij|kl \rangle = \int dx_1 dx_2 \, \frac{u_i^*(x_1)u_j^*(x_2)u_k(x_1)u_l(x_2)}{r_{12}}$$

which is the "physicists' notation."

The Slater-Condon rules

The rules for evaluating the matrix elements are called the Slater-Condon rules. The Slater-Condon rules for evaluating matrix elements of the Hamiltonian imply that the Hamiltonian is necessarily sparse since couplings between a state and any triple excitation from it vanishes. The Slater-Condon rules are given in Tables 2.2 and 2.3. The remainder of this section explains a method for deriving these rules.

To evaluate the matrix elements, the second quantized representation of quantum mechanics is useful and will be used several times throughout the thesis. Similar to the raising and lowering operators of harmonic oscillators that create and annihilate vibrational quanta, electron creation operators, $a_j^\dagger$, and annihilation, a_j, operators place or remove an electron from an orthonormal orbital u_j. Wave function anti-symmetry is enforced by the anti-commutation relations for fermions.

$$
\begin{align}
[a_i, a_j]_+ &= a_i a_j + a_j a_i = 0 \tag{2.8} \\
[a_j^\dagger, a_i]_+ &= a_j^\dagger a_i + a_i a_j^\dagger = \delta_{ij} \tag{2.9}
\end{align}
$$

Here $[A,B]_+ = AB + BA$ is called the anti-commutator and δ_{ij} is the Kronecker delta function function that is one when $i = j$ and zero if $i \neq j$. Note that the Pauli exclusion principal follows directly from Eq. (2.8). The raising and lowering operators act on the vacuum state, $|vac\rangle$, which contains no particles. Finally, there is a rule that $a_m|vac\rangle = 0$ since there are no electrons to remove from the vacuum.

Using these operators, the Slater determinant, $SD[u_\nu u_\mu ... u_\eta]$, is written $a_\nu^\dagger a_\mu^\dagger ... a_\eta^\dagger |vac\rangle$ and any one- or two-electron operator is written as

$$
\begin{align}
A(x) &= \sum_{ij} \left(\int u_i^*(x)A(x)u_j(x) \, dx \right) a_i^\dagger a_j = \sum_{ij} A_{ij} a_i^\dagger a_j \tag{2.10} \\
A(x_1, x_2) &= \frac{1}{2} \sum_{ijkl} \left(\int u_i^*(x_1)u_j^*(x_2)A(x_1,x_2)u_k(x_2)u_l(x_1) \, dx_1 dx_2 \right) a_i^\dagger a_j^\dagger a_k a_l \\
&= \frac{1}{2} \sum_{ijkl} A_{ijkl} a_i^\dagger a_j^\dagger a_k a_l \tag{2.11}
\end{align}
$$

Now, the evaluation of the matrix elements is a matter of manipulating the commutation relations.

First, observe if orbital l is not occupied then

$$a_l \left(a_\nu^\dagger a_\mu^\dagger ... a_\eta^\dagger \right) |vac\rangle = (-1)^N \left(a_\nu^\dagger a_\mu^\dagger ... a_\eta^\dagger \right) a_l |vac\rangle = 0$$

Similarly, if orbital j is not occupied then

$$\langle vac| \left(a_\nu a_\mu ... a_\eta \right) a_j^\dagger = (-1)^N \langle vac| a_j^\dagger \left(a_\nu a_\mu ... a_\eta \right) = (-1)^N \left(a_j |vac\rangle \right)^\dagger \left(a_\nu a_\mu ... a_\eta \right) = 0.$$

Therefore, when considering the expectation value of the two electron operator, only terms which act on occupied orbitals can be non-zero.

Consider operator $a_i^\dagger a_j^\dagger a_k a_l$ and the expectation value of $\langle \psi | H^{(2)} | \psi \rangle$. Repeatedly using Eqs. (2.8) and (2.9) simplifies $a_i^\dagger a_j^\dagger a_k a_l$ as

$$
\begin{aligned}
a_i^\dagger a_j^\dagger a_k a_l &= a_i^\dagger \left(\delta_{jk} - a_k a_j^\dagger \right) a_l \\
&= \delta_{jk} \left(\delta_{il} - a_l a_i^\dagger \right) - \left(\delta_{ik} - a_k a_i^\dagger \right) a_j^\dagger a_l \\
&= \delta_{jk} \delta_{il} - \delta_{ik} \left(\delta_{jl} - a_l a_j^\dagger \right) - a_l a_i^\dagger \delta_{jk} + a_k a_i^\dagger a_j^\dagger a_l \\
&= \delta_{jk} \delta_{il} - \delta_{ik} \delta_{jl} + a_l a_j^\dagger \delta_{ik} - a_l a_i^\dagger \delta_{jk} + a_k a_i^\dagger a_j^\dagger a_l
\end{aligned}
$$

As orbitals l and k are necessarily occupied, the raising operator will cause the state to vanish, $\langle \Psi | a_l = \langle \Psi | a_k = 0$, meaning the last three terms are zero. Finally, as desired,

$$
\begin{aligned}
\langle \Psi | H^{(2)} | \Psi \rangle &= \sum_{i>j} \sum_{k>l} \langle \Psi | h_{ijkl} a_i^\dagger a_j^\dagger a_k a_l | \Psi \rangle \\
&= \sum_{i>j} \sum_{k>l} h_{ijkl} \left[\left(\delta_{jk} \delta_{il} - \delta_{ik} \delta_{jl} \right) \langle \Psi | \Psi \rangle + 0 - 0 + 0 \right] \\
&= \sum_{i>j} \sum_{k>l} h_{ijkl} \left(\delta_{jk} \delta_{il} - \delta_{ik} \delta_{jl} \right) = \frac{1}{2} \sum_{ij} \left(h_{ijji} - h_{ijij} \right) \quad (2.12)
\end{aligned}
$$

To derive the remaining matrix elements, a similar procedure can be employed using the excitation operators $a_p^\dagger a_m$ and $a_p^\dagger a_q^\dagger a_n a_m$ when the occupied orbitals differ between the left and right state.

2.3 Full configuration interaction

Full configuration interaction (FCI) is both the simplest and the most computationally intensive method discussed in this chapter. The essence of FCI algorithm is to solve the secular equation and obtain the eigenstates for the N-electron Hamiltonian Eq. (2.2). The energies obtained at the basis set limit, E_{FCI}, are the exact non-relativistic Born-Oppenheimer energies of the system.

The Hamiltonian commutes with operators $\mathbf{S}^2$ and S_z so eigenstates of the Hamiltonian must also be eigenstates of these operators. Slater determinants are not necessarily eigenfunctions of the total spin operator, $\mathbf{S}^2$. This is addressed using linear combinations

$$\begin{array}{c}\hline \text{Case 1: Same spin orbitals}\\[2mm]\end{array}$$

Case 1: Same spin orbitals

$$|L\rangle = |\cdots u_m u_n \cdots\rangle$$
$$|R\rangle = |\cdots u_m u_n \cdots\rangle \qquad \langle L|H^{(1)}|R\rangle = \sum_i h_{ii}$$

Case 2: Differ by one spin orbital

$$|L\rangle = |\cdots u_m u_n \cdots\rangle$$
$$|R\rangle = |\cdots u_p u_n \cdots\rangle \qquad \langle L|H^{(1)}|R\rangle = h_{mp}$$

Case 3: Differ by two or more spin orbitals

$$|L\rangle = |\cdots u_m u_n \cdots\rangle$$
$$|R\rangle = |\cdots u_p u_q \cdots\rangle \qquad \langle L|H^{(1)}|R\rangle = 0$$

Table 2.2: Slater-Condon rules for one-electron operators. The integral defined in Eq. (2.4) uses the operators $T_e + V_{en}$; however any one-electron operator could have been used.

Case 1: Identical orbitals

$$|L\rangle = |\cdots u_m u_n \cdots\rangle$$
$$|R\rangle = |\cdots u_m u_n \cdots\rangle \qquad \langle L|H^{(2)}|R\rangle = \tfrac{1}{2}\sum_{ij}(h_{ijji} - h_{ijij})$$

Case 2: Differ by one orbital

$$|L\rangle = |\cdots u_m u_n \cdots\rangle$$
$$|R\rangle = |\cdots u_p u_n \cdots\rangle \qquad \langle L|H^{(2)}|R\rangle = \sum_j (h_{mjjp} - h_{mjpj})$$

Case 3: Differ by two orbitals

$$|L\rangle = |\cdots u_m u_n \cdots\rangle$$
$$|R\rangle = |\cdots u_p u_q \cdots\rangle \qquad \langle L|H^{(2)}|R\rangle = h_{nmpq} - h_{nmqp}$$

Case 4: Differ by three or more orbitals

$$|L\rangle = |\cdots u_m u_n u_l \cdots\rangle$$
$$|R\rangle = |\cdots u_p u_q u_r \cdots\rangle \qquad \langle L|H^{(2)}|R\rangle = 0$$

Table 2.3: Slater-Condon rules for two-electron operators. The integrals h_{ijkl} are defined using V_{ee}, but any other two-electron operator has the same form.

of Slater determinants to create a spin-adapted basis. The states in the spin-adapted basis are called configuration state functions or configurations and are eigenfunctions of $\mathbf{S}^2$ and S_z. The Hamiltonian couples configurations according to the Slater-Condon rules leading to the name configuration interaction.

The FCI expansion is typically written using a reference state $|\mathrm{Ref}\rangle$ and all possible configurations generated as excitations from that reference state. In explaining the algorithm, Slater determinants are used rather than configuration state functions to simplify notation. Let the operator T_m^μ be defined such that an electron from occupied orbital m is placed in unoccupied orbital μ.

For example, $T_m^\mu SD[u_i...u_m...u_k] = SD[u_i...u_\mu...u_k]$. Similarly, a double excitation operator replacing orbitals m and n with orbitals μ and η is written $T_{mn}^{\mu\eta}$. Using this notation, the wave function, $|\mathrm{FCI}\rangle$ is

$$|\mathrm{FCI}\rangle = c_0|\mathrm{Ref}\rangle + \sum_{k<m} c_m^k T_m^k |\mathrm{Ref}\rangle + \sum_{k<m,\, j<n} c_{mn}^{kj} T_{mn}^{kj} |\mathrm{Ref}\rangle + \dots \qquad (2.13)$$

The FCI ground state wave function includes contributions from all orders of excitations and the coefficients of the expansion, $\{c_m^k,\ c_{mn}^{kj},\ \dots\}$, are optimized until the lowest energy is found. This variational method is only applicable to the ground state (or the lowest energy state of a given symmetry).

Alternatively, direct matrix diagonalization can be considered. The roots of the secular equation

$$\det[H - \lambda I] = 0$$

are the eigenvalues of the Hamiltonian (electronic energies). The eigenvectors of the Hamiltonian are the ground state and excited states of the electronic system for the given nuclear configuration.

In either formulation of the FCI method, the drawback is the same: the rapid increase in the problem size with increasing basis set sizes. The underlying difficulty of the electronic structure problem can be understood by considering the scaling of the problem with the number of basis functions, M, and the number of electrons, N. There are

$$\begin{pmatrix} M \\ N \end{pmatrix} = \frac{M!}{N!(M-N)!} \qquad (2.14)$$

possible Slater determinants. The number of configurations is less than the number of Slater determinants due to symmetry considerations which reduce the number of states that need be considered. While these simplifications help, the problem size grows rapidly nonetheless.

The scaling of a computational problem describes how much more time is expected as the inputs get larger. If an algorithm scales quadratically, then, if on input of size k it took 150 elementary operations, then, on an input that is six times larger the algorithm would take 36 times more operations. These notions are defined more precisely in Chapter 4.

If $M \geq cN$ for any $c>1$ then Eq. (2.14) scales as $(c\text{-}1)^N$. Straightforward matrix diagonalization, e.g. using Gaussian elimination, requires computational effort that scales with the cube of the matrix size. As the matrix size is growing with the number of configurations, the computational effort will scale exponentially with the number of electrons for a sufficiently large basis set.

Instead of considering all excitations from the reference state, approximate configuration interaction schemes are often employed. For instance, optimization may only include configurations that differ by single excitations from the reference state (CIS), by singles and double excitations (CISD), and so on. This greatly reduces the computational cost; however, it introduces an undesirable feature.

From thermodynamics, the energy of a system should be extensive (if the size of the system doubles, the energy should double). Since the Hamiltonian respects this property, its eigenfunctions and eigenvalues, as obtained by FCI, also respect this property. Unfortunately, for truncated CI, the size extensive property is violated.

2.4 Hartree-Fock theory

In this section, the algorithmic details of Hartree-Fock will be emphasized, providing a gateway for the remainder of the chapter. The Hartree-Fock ansatz is a mean-field solution. mean-field solutions attempt to capture the interactions between particles in an average way so that a single particle only experiences this average potential instead of interacting explicitly with the other particles. The problem is greatly simplified from a N-electron problem to an essentially single-electron problem. Similar mean-field approximations are also used in statistical mechanics.

The single electron problem is solved to figure out where the electrons are likely to be found. From that information, the mean-field is recomputed. The single electron problem is solved again using the new mean-field potential, which, in turn, changes the potential which is recalculated to solve the individual electron problem, etc. This process continues until the average potential and the solutions to the individual electron problem converge from one iteration to the next. For this reason, mean-field approximations are known as self-consistent field (SCF) methods.

The Hartree-Fock SCF procedure generates the mean-field approximation for an N-electron system after a nuclear configuration has been fixed. Using the same spatial orbitals for both the α and β spin functions leads to the restricted Hartree-Fock formalism. Unrestricted Hartree-Fock allows the spatial components for the two spin functions to be different, which is appropriate when considering open shell systems and systems in magnetic fields.

Before the SCF procedure begins, a basis set, e.g. 6-31G, of spatial functions, $\{\phi_i(r),\ i{=}1,\ 2,\dots(M/2)\}$, is selected. When considering restricted Hartree-Fock formalism, the M spin-functions are

$$\{u_{2i}(x) = \phi_i(r)\alpha(m_s), u_{2i+1}(x) = \phi_i(r)\beta(m_s)\}$$

Also, before beginning the SCF procedure, the overlap matrix S with matrix elements, $S_{ij} = \langle\phi_i|\phi_j\rangle = \int \phi_i^*(r)\phi_j(r)dr$, is computed along with the one-electron and two-electron integrals from Eqs. (2.6) and (2.7) using the basis set functions. When dealing with systems where the number of integrals is very large, other parts of the algorithm may take longer than the computation of each integral. In this case, direct or "on the fly" methods are used where integrals are computed on demand; otherwise the integrals are stored in a lookup table.

These inputs are used to calculate an orthogonal set of spin-orbitals called the molecular orbitals denoted

$$\{\psi_a(x) \ : \ a = 1, 2, \ldots, M\} \tag{2.15}$$

Since each molecular orbital is a linear combination of the basis set functions, when the basis sets consists of atomic orbitals this method is known as the linear combination of atomic orbitals (LCAO). This forms the basis for the general chemistry description of hybridized bonding orbitals such as sp^2. Again, notice that the orbitals are single electron functions, and the occupied orbitals are combined in a single Slater determinant to form the N-electron wave function.

2.4.1 The Hartree-Fock equations

In the Hartree-Fock approximation, each electron is subjected to an average potential generated by the other electrons, reducing the N-electron problem to a single electron problem. The single electron operator of primary concern in the Hartree-Fock procedure is the Fock operator, denoted $\mathbf{F}$. The Fock operator is the sum of the single electron Hamiltonian, $H^{(1)}$, and the mean-field, V^{HF}. The mean-field V^{HF} is the sum of two operators: the Coulomb operator and the exchange operator. They are most easily defined by their action on spin-orbitals.

The Coulomb operator, $J_b(x_1)$, acts on molecular spin-orbital $\psi_a(x_1)$ as

$$J_b(x_1)\psi_a(x_1) = \left(\int dx_2 \frac{|\psi_b(x_2)|^2}{r_{12}} \right) \psi_a(x_1) \tag{2.16}$$

Multiplying on the left by $\psi_a^*(x_1)$ and integrating over x_1 yields

$$h_{abba} = J_{ab} = \int dx_1 dx_2 \frac{|\psi_b(x_2)|^2 |\psi_a(x_1)|^2}{r_{12}} \tag{2.17}$$

J_{ab} is called the Coulomb integral and it is the Coulomb potential generated by the interaction of two electric charge densities described by $|\psi_b(x_2)|^2$ and by $|\psi_a(x_1)|^2$. To evaluate this integral, techniques from classical electrodynamics, such as the multipole expansions, can be applied.

The exchange operator, $K_b(x_1)$ operates on $\psi_a(x_1)$ as

$$K_b(x_1)\psi_a(x_1) = \left(\int dx_2 \frac{\psi_b^*(x_2)\psi_a(x_2)}{r_{12}} \right) \psi_b(x_1) \tag{2.18}$$

The orbitals, ψ_a and ψ_b, "exchange" electron coordinates, hence the name of the operator. This operator is not a multiplicative operator, like the Coulomb operator. Instead, it is nonlocal because the value of $K_b(x_1)$ depends on the values of $\psi_a(x_1)$ at all possible coordinates and not just in a local neighborhood of x_1. The corresponding exchange integral, K_{ab}, is

$$h_{abab} = K_{ab} = \int dx_1 dx_2 \frac{[\psi_b^*(x_2)\psi_a(x_2)][\psi_a^*(x_1)\psi_b(x_1)]}{r_{12}} \tag{2.19}$$

The exchange integral has no classical analog, and if the spin of $\psi_a(x_1)$ and $\psi_b(x_1)$ are different, the integral vanishes.

Using definitions from Eqs. (2.15) and (2.17), the Fock operator can be written as

$$\mathbf{F}(x_1) = H^{(1)} + V^{HF} = \left(-\frac{1}{2}\nabla^2 + \sum_A \frac{Z_A}{r_{xA}}\right) + \left(\sum_b^M (J_b - K_b)\right) \qquad (2.20)$$

And in matrix form as

$$\mathbf{F}_{mn} = \int \psi_m(x_1)\mathbf{F}(x_1)\psi_n(x_1)dx_1 = h_{mn} + \sum_b (h_{mbbn} - h_{mbnb})$$

And the Hartree-Fock equations are then

$$\mathbf{F}\psi_i = \varepsilon_i\psi_i \text{ with } i = 1, 2, \ldots, M \qquad (2.21)$$

In the Hartree-Fock equations, Eq. (2.21), the set $\{\varepsilon_k\}$ are the orbital energies. From Eqs. (2.16), (2.18), and (2.20) the ith orbital energy is

$$\varepsilon_i = h_{ii} + \sum_{b=1}^M (J_{ib} - K_{ib}) \qquad (2.22)$$

where the indices are over the molecular orbitals. It should be emphasized that the sum of the occupied orbital energies is not the Hartree-Fock energy, which from the Slater-Condon rules is

$$
\begin{aligned}
E_{HF} &= \langle \Psi_{HF}|H|\Psi_{HF}\rangle = \sum_{i=1}^N h_{ii} + \frac{1}{2}\sum_{i=1,j=1}^{M,M}(J_{ij} - K_{ij}) \\
&= \sum_{i=1}^N \varepsilon_i - \frac{1}{2}\sum_{i=1,j=1}^{M,M}(J_{ij} - K_{ij})
\end{aligned}
\qquad (2.23)
$$

The molecular orbitals can be ordered by their orbital energy and the first N are occupied in the Hartree-Fock wave function. The molecular orbital with the highest energy is called the highest occupied molecular orbital (HOMO), and the first unoccupied molecular orbital is called the lowest unoccupied molecular orbital (LUMO). The unoccupied orbitals are called the virtual orbitals, and this is determined by the size of the basis set.

2.4.2 Derivation of the Hartree-Fock equations

Of course, Eq. (2.20) was just asserted. The derivation of the Hartree-Fock equations can be performed by variationally minimizing the energy of a single configuration state function with respect to the coefficients of the orbitals within the configuration

$$E_{HF} = \min_{\Psi \in SD_1} \langle \Psi|H|\Psi\rangle \text{ implies } |\Psi\rangle = |HF\rangle$$

Since the desired molecular orbitals should be orthogonal, this constraint can be enforced using the technique of Lagrange multipliers where the functional (a functional is a "function" that takes functions as inputs)

$$L[\{\psi_i\}] = E[\{\psi_i\}] - \sum_{ij} (\langle \psi_i \mid \psi_j \rangle - \varepsilon_{ij}) \tag{2.24}$$

is minimized through the calculus of variation. The full derivation will not be given, but in Eq. (2.24) the Lagrange multipliers $\{\varepsilon_{ij}\}$ appear instead of the orbital energies found in Eq. (2.21).

The set of Lagrange multipliers form a matrix, and the eigenvalues of this matrix are the orbital energies. The unitary rotation of the orbitals, $\psi_i'(x) = \sum_j \psi_j(x) U_{ji}$ does not affect the energy of the Slater determinant, nor alter the Coulomb or exchange operators in Eqs. (2.16) and (2.18). Unitary rotations must satisfy $\sum_k U_{ki}^* U_{kj} = \delta_{ij}$.

When selecting the unitary transformation that diagonalizes the matrix formed from the Lagrange multipliers

$$\varepsilon_m \delta_{mn} = \sum_{jk} U_{jm}^* \varepsilon_{jk} U_{kn} \tag{2.25}$$

the Hartree-Fock equations given in Eq. (2.21) are recovered. These equations are called the canonical Hartree-Fock equations, and these orbitals are the canonical orbitals. Other rotations of the orbitals find use in various situations, but rotations of the orbitals do not affect the Hartree-Fock wave function. In fact, when expanding the FCI wave function, the orbitals that minimize the number of CSF in the expansion are usually not the molecular orbitals. Also, for some applications, orbitals that are more localized may be preferred, and using the appropriate rotation of the orbitals, this can be accomplished.

2.4.3 Roothann matrix equations

The procedure for deriving the Hartree-Fock equation is Lagrange minimization; however, Clemens C. J. Roothaan (1951) reformulated the technique using linear algebra, and most modern implementations use the Roothann equations. As stated in the introduction to the algorithm, a known basis set is used to evaluate the single electron Hamiltonian and calculate the overlap matrix, S. A transformation matrix, C, is used to transform the given basis set orbitals into the molecular orbital basis as

$$\psi_j(x) = \sum_i u_i(x) C_{ij} \tag{2.26}$$

Expanding the Hartree-Fock Eqs. (2.21) using the basis gives

$$\mathbf{F}(x)\psi_i(x) = \varepsilon_i \psi_i(x)$$

$$\mathbf{F}(x) \left(\sum_m u_m(x) C_{mi} \right) = \varepsilon_i \left(\sum_m u_m(x) C_{mi} \right)$$

Multiplying $u_n^*(x)$ on the left and integrating

$$\sum_m C_{mi} \int u_n^*(x)\mathbf{F}(x)u_m(x)dx \;=\; \sum_m \left(\int u_n^*(x)u_m(x)dx \right) C_{mi}\varepsilon_i$$

$$\sum_m \mathbf{F}_{nm}C_{mi} \;=\; \sum_m S_{nm}C_{mi}\varepsilon_i$$

The final matrix equation is called the Roothaan equation and is written in final matrix form as

$$\mathbf{F}C = SC\varepsilon \tag{2.27}$$

This is almost an eigenvalue equation that can be solved using the secular equation. However, the overlap matrix S alters the procedure. The way to proceed is to define a new matrix X such that $XS^\dagger X = \mathbf{1}$. Note that $A^\dagger$ is the conjugate transpose of A where $(A^\dagger)_{ij} = (A^*)_{ji}$. Because S is an overlap matrix its eigenvalues are all non-negative and $X = S^{-1/2}$ exists and is unique. However, when S is singular (or nearly singular) due to (near) redundancy in the basis, more advanced techniques are used to remove linear dependencies. The matrix X transforms the basis to an orthogonal set via $C_x = X^{-1}C$ so that

$$\mathbf{F}XX^{-1}C = SXX^{-1}C\varepsilon$$

Multiply on the left by the conjugate transpose of X yields

$$(X^\dagger \mathbf{F}X)(X^{-1}C) = (X^\dagger SX)(X^{-1}C)\varepsilon$$

Finally, redefining variables yields the following eigenvalue equation

$$\mathbf{F}_x C_x = C_x\varepsilon \tag{2.28}$$

The eigenvalues of $\mathbf{F}_x$ are computed and then used to recompute $C = XC_x$ and finally, using the new coefficient matrix, the mean-field, V^{HF}, is recomputed with the new molecular orbitals defined in Eq. (2.26). In Fig. 2.2, the key steps of the Hartree-Fock algorithm are depicted in a flow chart.

2.4.4 Semi-empirical methods

In sharp contrast to the exponential scaling of the computational time, the run time of Hartree-Fock is quartic, in the basis set size. Usually, the majority of the time is dominated by the calculation of the four index integrals, Eq. (2.7), in the selected basis set. This scaling can often be improved using techniques from classical electrodynamics to quickly evaluate Coulomb integrals between distant orbitals resulting in cubic scaling with basis set size.

If this cost is too demanding for a particular application, semi-empirical methods are employed where simpler model Hamiltonians with adjustable parameters are used. Semi-empirical ideas are contrasted with *ab initio* (Latin: from the beginning) techniques such as Hartree-Fock and FCI where properties are calculated from first principals. Semi-empirical methods, as the name suggests, use experimental parameters to improve the accuracy and speed of computational methods.

Pick basis set

Calculate **S**, $\{h_{nm}\},\{h_{ijkl}\}$

Calculate **X** by diagonalizing **S**

Guess initial values for **C**

Compute V^{HF} from **C**, $\{h_{ijkl}\}$

Compare new **C=XC$_x$** to previous **C**. *If converged, s*top to compute properties.

Form the Fock matrix: $F_{nm}=h_{nm}+V^{HF}_{nm}$

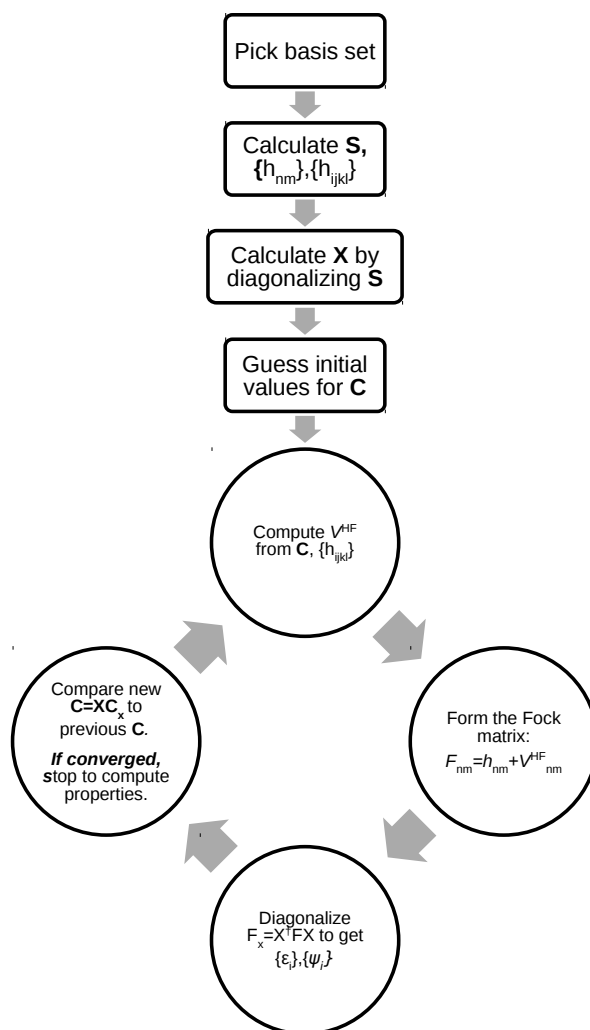Diagonalize $F_x=X^{T}FX$ to get $\{\varepsilon_i\},\{\psi_i\}$

Figure 2.2: **Flow chart of Hartree-Fock algorithm.** The initialization steps are in the boxes above and the self-consistent loop is repeated until convergence. When the coefficient matrix has stopped changing, the molecular orbitals and the Hartree-Fock wave function are computed along with any properties of interest.

The Hückel model, encountered in standard quantum chemistry, is an example of a semi-empirical Hamiltonian. The extended Hückel model is applied to systems besides conjugated systems, but retains the same form of the secular equation as the original Hückel model. Namely,

$$\det[H - ES] = 0$$

Essentially all semi-empirical methods assume only the valence electrons are important. The diagonal elements of the extended Hückel model, H_{mm}, are estimated from the ionization energy of the atomic system but can be adjusted for accuracy when the atomic ionization potential does not suit the molecule. The off-diagonal elements are more difficult to estimate but the following formula is often used.

$$H_{mn} = C \; S_{mn} \frac{(H_{nn} + H_{mm})}{2} \tag{2.29}$$

In Eq. (2.29), C is a numerical constant that is usually set to 1.75 but can be adjusted to improve experimental agreement, and S_{mn} is the overlap integral between the orbitals.

Finally, there is a set of semi-empirical approximations that are used to speed up Hartree-Fock calculations by approximating the integrals or ignoring them altogether. The most commonly used are zero differential overlap (ZDO), complete neglect of differential overlap (CNDO), and intermediate neglect of differential overlap (INDO). Products of the form $u_i^*(x_1)u_j(x_1)$ are called differential overlaps. If, for instance, in the ZDO approximation all differential overlaps are set to zero, then Hartree-Fock equations are solved using the assumption $h_{ijkl} = \gamma_{ij}\delta_{il}\delta_{jk}$ for the two-electron integrals. The various acronyms indicate how often the differential overlap is ignored and how the integral values γ_{ij} are approximated.

2.4.5 Correlation energy

The difference between the exact energy full configuration interaction energy and the HF energy is called correlation energy

$$E_{corr} = E_{FCI} - E_{HF} \tag{2.30}$$

and was defined by Per-Olov Löwdin (1955). The values in Eq. (2.30) usually refer to the basis set converged values. If a specific basis is used, Eq. (2.30) refers to the basis set correlation energy. The Hartree-Fock energy is typically less than 1% error. Unfortunately, this last percent of energy is responsible for much of chemistry as the energy of a bond or the barrier to rotation may only be a fraction of the total energy. For instance, the energy of the carbon atom is about 36.75 E_h whereas the energy of a carbon-hydrogen bond is about 0.18 E_h. This is half a percent of the carbon's energy. For larger molecules the total energy will be larger, so the energy of the carbon-hydrogen bond will represent an even smaller percentage.

Types of correlation energy

Correlation energy comes in two types, dynamical correlation and non-dynamical correlation energy. Dynamical correlation comes from the fact that electrons should be
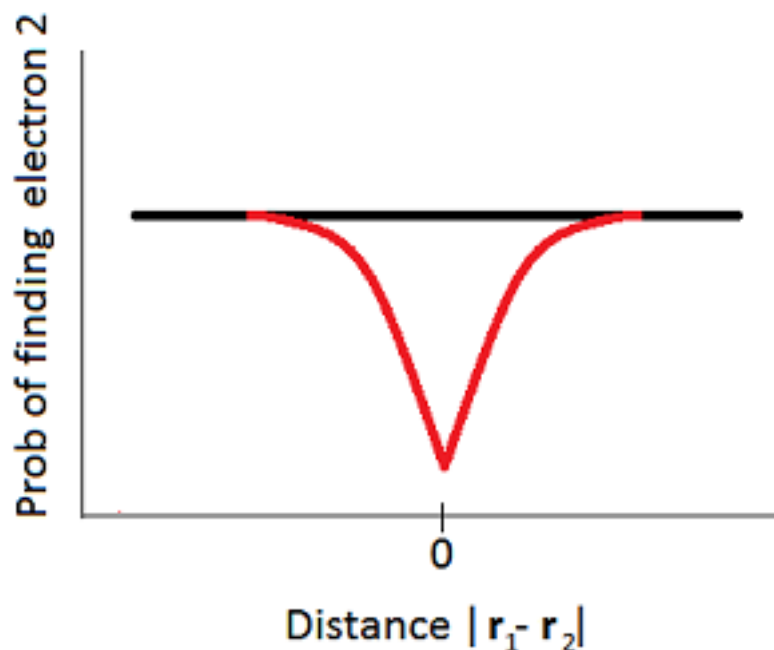
Figure 2.3: **The pair correlation function** of two opposite spin electrons at positions r_1 and r_2. In black is the Hartree-Fock pair correlation function, and, in red is the actual pair correlation function due to the Coulomb repulsion between charged particles.

correlated such that they avoid one another due to the Coulomb interaction, but, because of the mean-field approximation, they are not explicitly correlated. Instead, in the mean-field approximation, the electrons only see an averaged $N-1$ electron cloud instead of $N-1$ individual electrons. Depicted in Fig. 2.3 is the pair correlation function for electrons of opposite spin indicating the probability of finding a second electron some distance from the first electron. The black line is the pair correlation function expected from Hartree-Fock. The red cusp is due to the instantaneous Coulomb repulsion between electrons. If the electrons had the same spin, the pair correlation function would reflect the Pauli exclusion principal since antisymmetry is strictly enforced by the use of Slater determinants. The correlation due to the antisymmetry is called Fermi correlation and is not included in the definition of correlation energy. The inclusion of the explicit Coulomb interaction also modifies the pair correlation function of electrons possessing the same spin.

   Static (non-dynamical) correlation occurs when the problem is not well described by a single configuration. When multiple configurations are degenerate or near-degenerate, the Hartree-Fock solution will be in error due to the single configuration assumption. Examples when this situation occurs include bond breaking, excited state studies, and transition metal chemistry.

2.4.6 Multireference approaches

To capture the multireference nature of the problem, techniques have been developed to address the shortcomings of Hartree-Fock. The most straightforward approach is to include multiple determinants when performing the SCF iterations. Multi-configuration self consistent field (MCSCF) do exactly that. In these schemes, the SCF iterations optimize the orbital expansion and the configuration coefficients at the same time. The SCF procedure with multiple determinants requires much more technical expertise because of spurious minima in the coefficient space. Also, care must be used when selecting which orbitals to include. Popular methods for performing the MCSCF calculation are the Complete Active Space SCF (CASSCF) where some set of orbitals is chosen (by the user) and all excitations using a fixed number of electrons are permitted. The inactive space consists of other orbitals that may be occupied or unoccupied initially and remain so throughout the calculation. The principles behind the MCSCF method are essentially those of Hartree-Fock and FCI, however MCSCF methods are used for quantitative applications because of the freedom to limit the problem in intelligent ways.

2.5 Møller-Plesset many-body perturbation theory

When the Hartree-Fock wave function is a good description, perturbation theory is a straightforward way to improve the accuracy of the results. In 1934, Christian Møller and Milton S. Plesset, introduced a perturbation theory for many-body quantum mechanics and their perturbation techniques fall under the heading of Møller-Plesset perturbation theory.

The unperturbed Hamiltonian and the perturbation must be defined first. The Hartree-Fock wave function is not an eigenstate of the electronic Hamiltonian, however, it is an eigenstate of the Hartree-Fock Hamiltonian

$$H^{HF} = H^{(0)} = \sum_i \mathbf{F}(x_i) \tag{2.31}$$

Taking the difference between this Hamiltonian and the full electronic Hamiltonian gives the perturbation as

$$H^{(1)} = H - H^{(0)} = \sum_i \left(\sum_{j:\, i>j} \frac{1}{r_{ij}} - V^{HF}(x_i) \right) \tag{2.32}$$

The Hartree-Fock wave function is the ground state of H^{HF} with eigenvalue

$$E^{(0)} = \sum_i \varepsilon_i$$

As emphasized earlier, $E^{(0)}$ is not the Hartree-Fock energy as given in Eq. (2.23). The first order correction the energy is

$$E^{(1)} = \left\langle HF|H^{(1)}|HF \right\rangle \tag{2.33}$$

Applying the two-electron Slater-Condon rules to $(r_{ij})^{-1}$ and V^{HF} in Eq. (2.33) gives

$$E^{(1)} = \left\langle HF \left| \sum_{i,j<i} \frac{1}{r_{ij}} \right| HF \right\rangle - \left\langle HF \left| \sum_i V^{HF}(x_i) \right| HF \right\rangle \tag{2.34}$$

$$= \frac{1}{2} \sum_{ij} (J_{ij} - K_{ij}) - \sum_i \left[\sum_b (J_{ib} - K_{ib}) \right] \tag{2.35}$$

$$= -\frac{1}{2} \sum_j (J_{ij} - K_{ij}) \tag{2.36}$$

Adding the first order correction to the zeroth order energy yields the Hartree-Fock energy. Møller-Plesset theory goes on to calculate further corrections to the wave function and the energy, and the methods are named according to the order of the perturbation theory as MP2, MP3, ..., MPn,...

Continuing the discussion, consider the next correction, MP2. The second order correction is

$$E^{(2)} = \sum_{J>0} \frac{H_{0J}^{(1)} H_{J0}^{(1)}}{E^{(0)} - E_J^{(0)}} \tag{2.37}$$

The matrix elements to be computed are

$$H_{0J}^{(1)} = \left\langle HF|H^{(1)}|\Psi_j \right\rangle = \left\langle HF \left| V_{ee} \right| \Psi_J \right\rangle - \left\langle HF \left| \sum_i V^{HF}(x_i) \right| \Psi_J \right\rangle$$

Considering the first term, according to the Slater-Condon rules only single and double excitations need to be considered. The rules for the singly and doubly excited configurations

$$\langle HF|V_{ee}|T_m^p HF\rangle = \sum_j (h_{mjjp} - h_{mjpj})$$

$$\langle HF|V_{ee}|T_{mn}^{pq} HF\rangle = h_{mnqp} - h_{mnpq}$$

For the second term, from the Slater-Condon rules, only single excitations must be considered. From Eqs. (2.17), (2.19), and (2.20)

$$\left\langle HF \left| \sum_i V^{HF}(x_i) \right| T_m^p HF \right\rangle = \sum_b h_{mbbp} - h_{mbpb}$$

Putting it all together,

$$H_{0J}^{(1)} = \sum_j (h_{mjjp} - h_{mjpj}) + (h_{pqmn} - h_{pqnm}) - \sum_b (h_{mbbp} - h_{mbpb})$$

$$= (h_{pqmn} - h_{pqnm}) \tag{2.38}$$

Repeating the calculation for $H_{J0}^{(1)}$ gives the complex conjugate of Eq. (2.38). Note that the single excitation contributions canceled out.

Now consider the denominator, $E^{(0)} - E_j^{(0)}$. Since the numerator is only non-zero for double excitations,

$$
\begin{aligned}
E^{(0)} - E_J^{(0)} &= \left\langle HF \left| H^{(0)} \right| HF \right\rangle - \left\langle HF \left| (T_{mn}^{pq})^\dagger H^{(0)} T_{mn}^{pq} \right| HF \right\rangle & (2.39) \\
&= \varepsilon_n + \varepsilon_m - \varepsilon_p - \varepsilon_q & (2.40)
\end{aligned}
$$

Finally, the expression for the second order energy correction is

$$
E^{(2)} = \sum_{p<q}^{virt} \sum_{m<n}^{occ} \frac{|h_{pqmn} - h_{pqnm}|^2}{\varepsilon_n + \varepsilon_m - \varepsilon_p - \varepsilon_q} \tag{2.41}
$$

Higher order corrections to the energy and the wave function can be derived in much the same way.

Perturbation theory is useful for improving upon the Hartree-Fock solutions because it includes contributions from many other determinants to reproduce the dynamical correlations. That said, MPn does not always converge to the exact solution nor does higher order perturbation theory always improve the wave function. Instead, in some cases, higher order perturbation corrections will have errors greater than the Hartree-Fock solution. In most cases, MP2 is a relatively inexpensive way to recover around 85 to 95% of the correlation energy and higher orders of Møller-Plesset theory are mostly used for methodology testing rather than practical applications.

2.6   Coupled-cluster methods

Coupled-cluster methods are a class of pertubative techniques that look very different from the perturbation techniques discussed in the previous section. The origin of the coupled-cluster method was in the field of nuclear physics, but, it was initially ignored until its importance in chemistry was recognized. Coupled-cluster method focuses on the physical nature of the excitation rather than the configurations emphasized in configuration interaction.

The coupled-cluster ansatz for the wave function is given by

$$
|CC\rangle = \left[\prod_\mu \left(1 + t_\mu T_\mu \right) \right] |HF\rangle \tag{2.42}
$$

where $\{t_\mu\}$ are the coefficients parameterizing the wave function and T_μ is shorthand for some excitation operator $T_{m_1 m_2 \ldots m_k}^{\nu_1 \nu_2 \ldots \nu_k}$. Notice that this wave function is composed of a product over all excitations unlike the CI expansions discussed so far, which were sums over excitations.

Eq. (2.42) is simplified by considering that the square of the excitation operator $\hat{\tau}_\mu^2 = 0$. This follows using the second quantized representation and recalling that excitations

are from occupied orbtials to virtual orbitals. Thus,

$$
\begin{aligned}
|CC\rangle &= \left[\prod_{\mu}\left(\mathbf{1}+t_{\mu}T_{\mu}\right)\right]|HF\rangle \\
&= \left[\prod_{\mu}\left(\mathbf{1}+t_{\mu}T_{\mu}+\frac{(t_{\mu}T_{\mu})^2}{2!}+\frac{(t_{\mu}T_{\mu})^3}{3!}+\dots\right)\right]|HF\rangle \\
&= \left[\prod_{\mu}\exp\left(t_{\mu}T_{\mu}\right)\right]|HF\rangle \\
&= e^{\mathbf{T}}|HF\rangle
\end{aligned}
\tag{2.43}
$$

with $\mathbf{T}=\sum_{\mu}t_{\mu}T_{\mu}$ as the cluster operator. This is called the coupled-cluster exponential ansatz.

The cluster operator is broken up to reflect the hierarchy of excitation levels.

$$
\mathbf{T}=\mathbf{T}_1+\mathbf{T}_2+\mathbf{T}_3+\dots+\mathbf{T}_{\mathrm{N}}
\tag{2.44}
$$

Each $\mathbf{T}_n$ includes all excitations of nth order. For example, the first two excitation hierarchies are

$$
\mathbf{T}_1 = \sum_{m}^{occ}\sum_{p}^{virt}t_m^p\,a_p^{\dagger}a_m = \sum_{m}^{occ}\sum_{p}^{virt}t_m^p T_m^p
\tag{2.45}
$$

$$
\mathbf{T}_2 = \sum t_{mn}^{pq}a_p^{\dagger}a_m a_q^{\dagger}a_n = \sum t_{mn}^{pq}T_{mn}^{pq}
\tag{2.46}
$$

The exponential wave function requires different approaches than the ones seen so far in this chapter. Instead of variational minimization, the projected Schrödinger equation is used. The idea is to project the coupled-cluster wave function onto a particular excitation. For instance, the coupled-cluster energy is obtained by a projection onto the Hartree-Fock state

$$
E = \langle HF\,|\mathbf{H}|\,CC\rangle
\tag{2.47}
$$

and for excited configuration $|\mu\rangle$, the projected coupled-cluster equation

$$
\langle\mu\,|\mathbf{H}|\,CC\rangle = E\,\langle\mu|CC\rangle
\tag{2.48}
$$

gives a set of equations for the coefficients of the parameterization. The previous two equations are used to calculate the coupled-cluster energy. However, this equation, in contrast to

$$
E = \frac{\langle CC\,|\mathbf{H}|\,CC\rangle}{\langle CC|CC\rangle}
$$

is not necessarily variational. Usually, the variational principle for coupled-cluster breaks down at the fourth and higher orders of perturbation theory, so coupled-cluster theory is termed quasivariational. These equations are always extensive for any level of truncation, unlike CI, which is only extensive for HF and FCI.

The coupled-cluster ansatz is typically truncated at some level. Otherwise, the coupled-cluster wave function is the same as the FCI wave function. Often, just singles and double excitations are kept and this method is then called CCSD. Similarly, CCSDT would keep single, double, and triple excitations. It is possible to include the triple excitations perturbatively when performing a CCSD calculation. This is abbreviated as CCSD(T) and is considered a high quality level of theory. Unfortunately, the scaling of CCSD(T) is on the order of the seventh power of the number of basis functions.

It should be noted that even the truncated coupled-cluster wave functions include excitations at all orders due to the exponential nature of the wave function. For example, the 6th order excitations include contributions, not only from $T_{m_1...m_6}^{\nu_1...\nu_6}$ but also products of a quadruple excitation and a double excitation, two triple excitations, three double excitation, and so on. The disconnected amplitudes allow the wave function to recover dynamic correlation in a systematic and efficient way.

2.7 Density functional theory

The wave function contains the totality of the information that quantum mechanics can provide about a system, and, so far this chapter has been concerned with methods to obtain improved wave functions. In this section, a different approach is pursued. While the wave function of N electrons is a function of $3N$ spatial coordinates, density functional theory is predicated on the fact that the 3 spatial dimensional ground state (probability) density also contains the same information.

Density functional theory (DFT) methods and applications are numerous and of significant practical value due to their speed and reasonable accuracy. Theoretically, the foundations of DFT are exact, but to make any progress beyond trivial examples, approximations must be made. Unlike the wave function based methods, the methods to improve the accuracy are not systematic. In this section, an introduction to the ideas, implementation, and applications of DFT are given. More detailed expositions are found in the monographs listed in the beginning of the chapter.

2.7.1 Densities in quantum mechanics

'Density' in quantum mechanics has several meanings. It may refer to the density matrix, the reduced density matrix, or the single-particle reduced probability density. In DFT, the density refers to the last of these notions, but it is helpful to examine the other two as they have important applications to electronic structure and the general study of quantum mechanics as exemplified by the work in Chapter 10.

First, the density matrix is the quantum mechanical equivalent of the probability distribution function. This connection is further elaborated in the next chapter. It is a matrix formed by taking the outer product of the wave function, ψ.

$$\rho(x_1, x_1', x_2, x_2', ..., x_N, x_N') = \psi^*(x_1, x_2, ..., x_N)\psi(x_1', x_2', ..., x_N') = |\psi\rangle\langle\psi| \qquad (2.49)$$

Three important facts summarize the properties of density matrices:

1. Other valid density matrices can be obtained via convex combinations,

$$\rho = p_1 \rho_1 + p_2 \rho_2 + ... + p_M \rho_M$$

   with $\sum_{k=1}^{M} p_k = 1$ and each p_k is greater than or equal to zero.

2. If the wave functions used in each outer-product are normalized, then

$$\int dx_1 ... dx_N \rho(x_1, x_1, x_2, x_2, ..., x_N, x_N) = 1$$

3. Considering only the diagonal of the density matrix, where for each k, $x_k = x'_k$ gives the probability distribution of the N electrons that determines measurement outcomes.

So far, it appears that no improvement has been made as the density matrix has twice as many coordinates as the wave function. The reduced density matrix remedies this problem by removing electrons from consideration. Similar to marginal probability distributions, coordinates of the density matrix are removed by integration over the variables to be ignored. Turning to the definition, the kth order reduced density matrix (k-RDM) is

$$\rho^{(k)}(x_1, x'_1, x_2, x'_2, ..., x_k, x'_k) = \int dx_N dx_{N-1} ... dx_{N-k}$$

$$\rho(x_1, x'_1, x_2, x'_2, ..., x_{N-k}, x_{N-k}, ..., x_{N-1}, x_{N-1}, x_N, x_N)$$

$$(2.50)$$

The most important types of RDMs are the first and second order RDM. This is because the electronic Hamiltonian contains only one- and two-electron terms. While the 2-RDM contains sufficient information to characterize the energy, obtaining valid 2-RDMs is not straightforward. Unconstrained minimization of the energy corresponding to a 2-RDM can tend to negative infinity because not all seemingly valid 2-RDMs correspond to N-electron wave functions as in Eq. (2.50). Distinguishing which 2-RDMs came from N-electron wave functions is known as N-representability problem, and for 2-RDMs, the exact constraints are not known. This problem is discussed further in Chapter 4. For the 1-RDM, N-representability simply requires positivity and the normalization constraint, but two-electron properties are not directly accessible.

DFT is based on the even simpler, one-electron reduced *probability* distribution. Usually in the context of DFT, the term 'density' refers to the probability distribution given by

$$n(x) = \rho^{(1)}(x, x) = \int dx_2 dx_3 ... dx_N |\psi(x, x_2, ..., x_N)|^2 \qquad (2.51)$$

The density has only 3 spatial coordinates instead of the $3N$ spatial coordinates of the wave function, but, surprisingly, for electronic Hamiltonians, the density characterizes all properties. This fact is proved in the next section along with two other theorems that provide the theoretical basis for DFT. The density in quantum chemistry is usually normalized to N instead of unity as in Eq. (2.51).

For the rest of this section, the spin will be ignored so that $n(x) \rightarrow n(r)$. Generalization that utilize information on the spin can be found elsewhere.

2.7.2 Theoretical basis for DFT: three theorems

Density functional theory is based on three important theorems. The first two theorems provide the basis for DFT, and the third provides a method for practical implementation. The landmark paper by P. Hohenberg and W. Kohn in 1964 showed (1) densities are uniquely determined by the external potential and (2) that the density satisfies a variational theorem. The third theorem to be discussed comes from a 1965 paper by W. Kohn and L. Sham, which introduces a fictitious non-interacting electronic system that reproduces the same density as the interacting system. For laying the foundation of DFT, Walter Kohn shared the 1998 Nobel Prize in chemistry. This section continues with the statement and proof of the first Hohenberg-Kohn theorem.

Theorem 1. *(First Hohenberg-Kohn theorem) The ground state probability density is uniquely determined by the external potential (up to an arbitrary energy shift).*

Proof. The proof given here will be for non-degenerate ground states, but the theorem holds more generally. Suppose, to get a contradiction, two different external potentials give rise to the same density. Let Ψ_1 be the unique ground state of the first Hamiltonian and Ψ_2 be the ground state of the second and both giving rise to the same density, $n(r)$.

$$E_1 = \langle \Psi_1 | H_1 | \Psi_1 \rangle = \langle \Psi_1 | (T + V_{ee}) | \Psi_1 \rangle + \int n(r) V_{en}^{(1)}(r) dr \qquad (2.52)$$

is the ground state energy of the first Hamiltonian. Similarly,

$$E_2 = \langle \Psi_2 | H_2 | \Psi_2 \rangle = \langle \Psi_2 | (T + V_{ee}) | \Psi_2 \rangle + \int n(r) V_{en}^{(2)}(r) dr \qquad (2.53)$$

The variational principle and the assumption that H_1 has a unique ground state imply

$$E_1 < E_1' = \langle \Psi_2 | (T + V_{ee}) | \Psi_2 \rangle + \int n(r) V_{en}^{(1)}(r) dr = E_2 + \int n(r) (V_{en}^{(1)}(r) - V_{en}^{(2)}(r)) dr \quad (2.54)$$

Also

$$E_2 \leq E_2' = E_1 + \int n(r) (V_{en}^{(2)}(r) - V_{en}^{(1)}(r)) dr \qquad (2.55)$$

The case of equality is allowed for since the ground state H_2 may be degenerate. Adding Eqs. (2.54) and (2.55), yields

$$E_1 + E_2 < E_1 + E_2$$

which is a contradiction. Notice that the difference in the potentials appears in both Eqs. (2.54) and (2.55) so potentials that differ by a constant shift are equivalent. Thus, one is forced to accept that the density is in one-to-one correspondence with the external potential (up to a constant shift of the potential). Notice that we assumed that H_1 has a non-degenerate ground state. □

As a corollary of the theorem, the density is also, *in theory*, in one-to-one correspondence with all properties of the system including the ground state energy and all

excited state energies, electric, and magnetic response properties, and any other property of the system. Unfortunately, this theorem only proves that such a correspondence exists but does not provide a method for extracting such properties. The extraction of properties from the density require some sort of mapping from the density to a number or perhaps a few numbers which correspond to the desired properties. These mappings are called functionals. More generally, a functional is any 'function of a function' in that it assigns numerical values to input functions the same way that ordinary functions assign numeral values to input numbers. Common practice is to write the input function in square brackets i.e. $A[n]$. In the proof, one of the most important, and, elusive, functionals was encountered. In all electronic Hamiltonians, the electrons have kinetic energy and electrons interact via the Couloumb term; only the nuclear configuration changes from system to system. The universal functional encompasses the part of Hamiltonian that does not change from electronic system to electronic system

$$F[n] = \min_{\psi \to n(r)} \langle \Psi | (T + V_{ee}) | \Psi \rangle = T[n] + E_{ee}[n] \tag{2.56}$$

The universal functional is not likely to be solved as discussed in Chapter 4, but, in the next section, a scheme for approximating the universal functional is given. Before further discussing the universal functional, the variational principle for densities and the Kohn-Sham system are introduced.

Theorem 2. *(Second Hohenberg-Kohn theorem) The density corresponding to the lowest energy, corresponds to the density obtained from the exact wave function.*

Proof. If $\tilde{n}(x)$ is an approximate density, then its corresponding potential $\tilde{V}_{ext}$ determines a corresponding ground state wave function $\tilde{\Psi}$. Comparing $\tilde{\Psi}$ and the actual ground state, Ψ, allows the variational principle to be used in completing the proof. $\qquad \square$

Notice the point of the second theorem is to avoid the energetic catastrophes of the 2-RDM approaches. However, it raises a different representability problem instead of the aforementioned N-representability. Assuming that there is a corresponding potential $\tilde{V}_{ext}$ for each density $\tilde{n}$ presents the V-representability problem. An example of a density that is not V-representable was first constructed using the degenerate ground states of systems. The constraints for V-representable densities are not known, but V-representability can largely be circumvented using the Levy-Lieb constrained minimization search that will not be discussed further in this chapter. This method only requires N-representability and can also be used to remove the assumption of non-degeneracy found in the proof of the first Hohenberg-Kohn theorem.

Finally, attention turns to the final, and, in many ways, the most practical of the DFT theorems discussed in this section the Kohn-Sham theorem (1965). The Kohn-Sham theorem aligns the methodologies of DFT with that of Hartree-Fock and gives a practical method for implementing DFT that is similar to the techniques discussed in Section 2.4.

Theorem 3. *(Kohn-Sham Theorem) Given any electronic system, there exists a non-interacting system ($V_{ee} = 0$) that reproduces the same density. That is, there exists a*

mapping from the external potential, V_{ext}, of the original system to another potential, V^{KS}, for the non-interacting system with an effective one-electron Hamiltonian

$$\left(-\frac{1}{2}\nabla^2 + V^{KS}[n](r)\right)\phi_j(r) = \epsilon_j\phi_j(r) \qquad (2.57)$$

with the density (normalized to N) given by

$$n(r) = \sum_j |\phi_j(r)|^2 \qquad (2.58)$$

Proof. The uniqueness of the V^{KS} potential is guaranteed by the first Hohenberg-Kohn theorem with $V_{ee} = 0$. The existence of the Kohn-Sham potential is guaranteed for all V-representable densities again with $V_{ee} = 0$. □

Thus far, no approximations have been made, so all is still exact. However, no practical applications can yet be made. The usefulness of the Kohn-Sham system for proceeding is explored next.

2.7.3   The Kohn-Sham system

The non-interacting system, while formally exact for all V-representable densities, is not without complications which will be the focus of this subsection. The total energy functional for the Kohn-Sham system can again be partitioned into the universal part and the system-dependent term, $F[n] + \int n(r)V_{ext}(r)dr$. Since the system is non-interacting, the universal functional can be written as follows

$$F[n] = (T_s + J)[n] + E_{xc}[n] \qquad (2.59)$$

The first term is the non-interacting approximation to the kinetic energy given by

$$T_s[n] = -\frac{1}{2}\sum_j \langle\phi_j|\nabla^2|\phi_j\rangle$$

This term is always less than or equal to the kinetic energy of the interacting system but captures the majority of the kinetic energy of the interacting system. This approximation of the kinetic energy is at heart of the success of the Kohn-Sham approach. The earlier Thomas-Fermi theory (1926/1928) that was also based on the density was largely ignored because it poorly approximated the kinetic energy leading to a theoretical model that did not even predict the binding of molecules. This is in contrast to more successful Hartree-Fock theory that was developed around the same time. In the Kohn-Sham approach, the kinetic energy is approximated by the same term found in Hartree-Fock theory.

The second term is the energy of the classical Coulomb interaction. Recall in Eq. (2.16), the same expression was encountered in Hartree-Fock theory. As a result the potential associated with this functional

$$V_H[n](r) = \int dr' \frac{n(r')}{|r - r'|} \qquad (2.60)$$

is called the Hartree potential.

The final term, $E_{xc}[n]$, is the exchange-correlation functional and is a junkyard of sorts that is supposed to encapsulate the non-classical electron-electron interactions, e.g. the exchange term in Hartree-Fock theory from Eq. (2.18), any missing kinetic energy and account for the self-interaction of electrons. The exchange-correlation functional is where approximations must be made and the ideas behind these approximations will be discussed in Section 2.7.4. The exchange-correlation functional is related to the exchange-correlation potential, V_{xc}, through a functional derivative

$$V_{xc}[n](r) = \frac{\delta E_{xc}[n]}{\delta n(r)}(r) \qquad (2.61)$$

The functional derivative is defined in the calculus of variations and is the extension of the derivative of functions to functionals.

Returning to Eq. (2.57) with Eqs. (2.60) and (2.61) in tow, the Kohn-Sham potential in the one-electron Hamiltonian is given by

$$V^{KS}[n](r) = V_H[n](r) + V_{xc}[n](r) + V_{ext}(r) \qquad (2.62)$$

The Kohn-Sham procedure

Before discussing approximations to the exchange-correlation term, this subsection turns to the procedure and the caveats of the Kohn-Sham approach. Again, the idea is the non-interacting system, described by a single determinant, is used to reproduce the density of the interacting system. Since the electrons do not interact, the problem becomes a single-electron problem similar to Hartree-Fock. Just as in the Roothaan equations, the Kohn-Sham system is usually solved using basis sets. The basis sets used are often optimized for use in DFT but share the same fundamental ideology as those discussed earlier.

The Kohn-Sham potential, like the mean-field potential in Hartree-Fock, depends on its eigenfunctions. This means the Kohn-Sham system must be solved iteratively like the Hartree-Fock equations. For each iteration, the density is updated according to Eq. (2.58) and the Kohn-Sham potential is updated until the potential and density converge to self-consistency.

Unlike Hartree-Fock, the Kohn-Sham single determinant wave function has no physical meaning except (1) to reproduce the density according to Eq. (2.58) and (2) the energy of the highest occupied orbital is (minus) the ionization energy of the system. Also worth noting, the highest occupied molecular orbital energy of the anion is the electron affinity. The correspondence between the ionization energy and the energy of the highest occupied Kohn-Sham orbital holds only for the exact V^{KS}, not for the approximate versions discussed next.

2.7.4 Approximations to the exchange-correlation functional

Currently, there are nearly 200 approximations to this functional in the literature. Needless to say, all of them cannot be discussed. Instead, classes of approximations are grouped to allow the reader an orientation to the large variety of exchange-correlation

approximations. While properties of the exact exchange-correlation potential and functional are known, e.g. $V_{xc}(r) = -\frac{1}{|r|}$ as $r \to \infty$, systematically improving the functionals is not straightforward; hence the large number of approximate functionals.

The earliest and simplest approximation is the Local Density Approximation (LDA). Here, the exchange-correlation functional from the uniform electron gas is applied to other systems. This gives a local functional whose value at a point depends only on the density at that particular point.

$$V_{xc} = V_{xc}[n]$$

The LDA worked excellently in solid state physics applications where metals are reasonably approximated by uniform electron gas. However, for chemistry, molecules do not have nearly uniform electron densities.

This prompted the introduction of functionals that depend on more than just the local value of the density. The generalized gradient approximations (GGA) are the set of functionals that depend on the density and its first gradient.

$$V_{xc}^{GGA} = V_{xc}[n, \nabla n]$$

Continuing along this path, meta-GGA, were introduced where the potential is a functional of the density, its first and second gradients and the non-interacting kinetic energy functional,

$$V_{xc}^{mGGA}[n, \nabla n, \nabla^2 n, T_s]$$

Yet, another approach is to use orbital functionals where V_{xc} is a functional of the orbitals. This leads to a more complicated implementation and is an active area of research.

Finally, the most widely used type of orbitals are the hybrid functionals. These functionals utilize linear combinations of GGAs and the (exact) exchange contribution derived from Hartree-Fock, see Eq. (2.18). The foremost among them is the B3LYP (pronounced 'B-three-lip') functional which is a linear combination of the LDA, the exact exchange functional, the B88 functional (Becke 1988), and the LYP functional (Lee, Yang and Parr). In B3LYP, there are 3 fitting parameters that were optimized using a training set (G2 reference set) of 50 small molecules that have been well characterized experimentally.

Perhaps, it may seem a bit odd that the exact exchange is not solely used, leaving only the correlation aspects to be approximated. This is because when using the exact exchange term, the approximate correlation functional cannot correctly compensate the terms in the potential it is responsible for. By mixing the exact exchange with approximate exchange, often errors in the correlation can be cancelled by errors in the exchange approximation. In fact, when scrutinizing the success of B3LYP, it is often found that the close agreement between the B3LYP results and the experimental results comes down to a cancellation of large errors in the exchange approximation with large errors in the correlation approximation.

A final note on the approximate functionals regards their evaluation during the self consistent Kohn-Sham calculations. Typically, the exchange-correlation potential must be evaluated numerically, unlike the other terms in the Kohn-Sham potential and in the Hartree-Fock potential, which can be evaluated using the orbitals from the basis set. The integrals involving the density and the exchange-correlation potential are typically evaluated numerically on a grid, although grid-free methods exist.

2.8   Electronic structure calculations

Standard calculations include equilibrium geometries, vibrational frequencies, bond strengths, electron ionization and affinities, dipole moments, polarizabilities, infrared intensities, NMR chemical shifts, transitions states, reaction barriers, free energies, and many other properties. The accuracy of these properties is determined by the basis set used and the method employed. Typically, quantitative accuracy can often be obtained using CCSD(T) with the cc-pVTZ basis set, however for larger systems, often this cannot be done in a feasible amount of time. DFT is often employed for larger systems despite the inability to objectively ascertain the accuracy. Due to the large DFT literature and the numerous functionals, often the accuracy can be inferred using prior experience. MCSCF and CASSCF calculations are often used when chemical intuition provides insight into which orbitals and configurations are expected to dominate the ground state wave function. However, the limitation on the size of the system and its description is what leads us to the explorations pursued in the remainder of the thesis. The hope is that model chemistries on quantum computers can provide new approaches to tackling large systems and large basis sets, and I will present our first steps toward establishing this paradigm.

Chapter 3

Quantum computation and quantum information

"Information is physical."

- R. Landauer [125]

3.1 Introduction

The idea of mapping the dynamics of a quantum system of interest onto the dynamics of a controllable quantum system was first proposed in 1982 by Feynman [75] and developed further in 1996 by Lloyd [139]. Such a quantum computer would be able to obtain information inaccessible with classical computers. Consequently, quantum simulation promises to be a powerful new tool for quantum chemistry.

In order to approach these ideas maturely in the later chapters, basic concepts in quantum computation must be introduced. The purpose of this chapter is provide an introduction to quantum computation as an orientation to the results presented through the thesis.

In the next section, the notion of information as a physical construct is given as motivation for quantum information. Quantum information is discussed and compared with its classical counterpart in Section 3.3. The transformation of input information into useful quantities is done via quantum computation which is discussed in Section 3.4. Experimental progress toward the realization of quantum computers is briefly discussed at the end of the chapter.

3.2 Information is physical

The notion that information is a physical construct may not be immediately obvious but, it is so critical to this thesis that it must be addressed before continuing. The notion that information is physical has been embraced as the founding tenet of quantum information and quantum computation. Once one has accepted that information is a physical construct, it is easy to fathom the use of quantum objects as information carriers leading
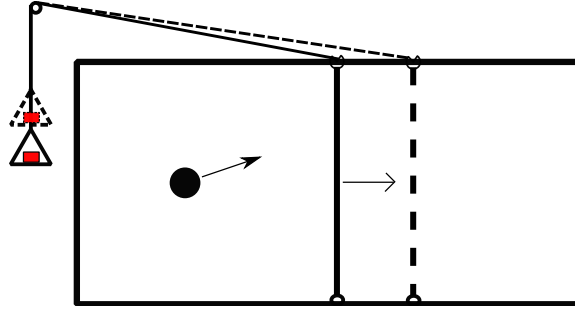
Figure 3.1: **Szilard's engine** is a simplification of the Maxwell's demon paradox. In Szilard's engine there is only one particle and the demon will insert a movable partition into the box. The demon measures which side the particle is on and attaches a pulley to extract work as the particle pushes the wall to the other side. If the volume increase occurs adiabatically, the maximum amount of work is extracted is $kT \log 2$.

directly to quantum information. To motivate the physical nature of information, consider the famous quote by J. C. Maxwell, which started the longstanding debate on the possibility of violating the second law of thermodynamics which states that entropy of closed systems increase [143]

> If we conceive of a being whose faculties are so sharpened that he can follow every molecule in its course, such a being, whose attributes are as essentially finite as our own, would be able to do what is impossible to us. . . A being, who can see the individual molecules, opens and closes this hole, so as to allow only the swifter molecules to pass from A to B, and only the slower molecules to pass from B to A. He will thus, without expenditure of work, raise the temperature of B and lower that of A, in contradiction to the second law of thermodynamics.

An interesting simplification of Maxwell's Demon was given by L. Szilard where only one molecule is involved. If the demon places a movable wall attached to a pulley, see Fig. 3.1, the maximum amount of work that can be extracted as the single particle pushes the piston is $kT \log 2$.[1] This relation is the same as Shannon's for maximal entropy of a single bit of information discovered nearly 20 years later [191].

In the 1980's, R. Landauer and C. Bennett realized that the demon was not violating the second law of thermodynamics because of the entropy expended in erasing memory [23]. Even if the demon does not erase its memory, the memory acts as an entropy sink and absorbs the entropy that the demon has taken from the environment as illustrated in Fig. 3.2. Thus, the second law of thermodynamics is argued to hold even for the demon by subjecting his information to the laws of physics: "Information is physical."

---

[1] k is a numerical constant which depends on the base of the log and the units used for measuring the work and temperature. For the natural log, energy units of Joules and temperature in Kelvin, $k = k_B = 1.3806488 \times 10^{-23} J/K$ is the Boltzmann constant.

Demon's memory

S=0
perfectly
ordered

S>0
disordered

000000
000000
000000

010001
001010
110001
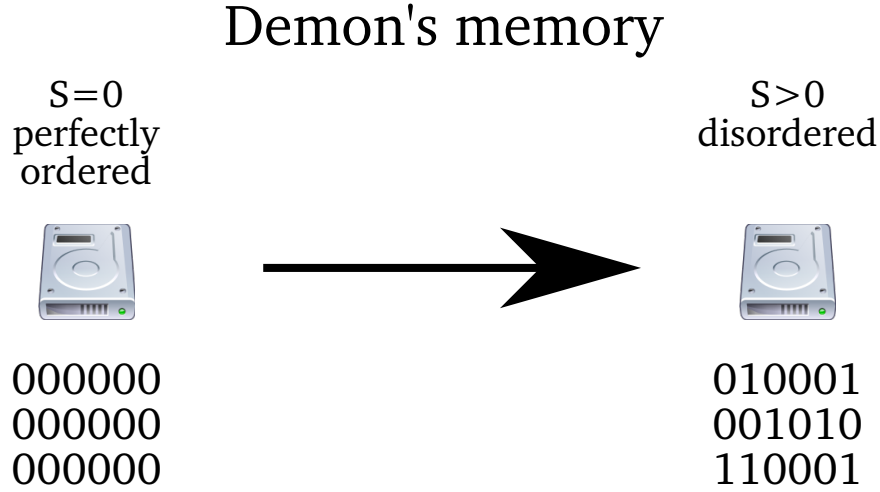
Figure 3.2: **The memory of the demon** begins in a completely ordered state with zero entropy and as the demon performs measurements, the entropy of his memory increases to some finite value depicted on the right. This is one possible solution for why Maxwell's Demon does not violate the second law of thermodynamics and a motivation for regarding information as a physical entity.

3.3 Quantum information

Assuming information is physical, the consequences of using quantum physics to describe information and its transformations leads to very different notions.

3.3.1 Differences between classical and quantum information

Quantum information is described by a wave function instead of a probability distribution as in the case of classical information. From a mathematical point of view, the key differences between the wave function and the classical probability vector are the normalization constraint (L_2 versus L_1) and the field of the elements used in the vectors (real versus complex numbers). That is,

$$\begin{array}{lll} \text{Classical:} & \sum_k p_k = 1 & p_k \in \mathcal{R} \\ \text{Quantum:} & \sqrt{\sum_k |\langle k|\psi\rangle|^2} = 1 & \langle k|\psi\rangle \in \mathcal{C} \end{array}$$

The elements of the Hilbert space[2] are, in Dirac notation, kets which we write as $|\cdot\rangle$. The ket represents a state without specifying which basis it is expanded in. Often a set of kets are defined as eigenstates of an operator and the nth eigenstate is written as $|n\rangle$. Elements

---

[2]A Hilbert space is a (possibly infinite dimensional) vector space where the limit point of any sequence of vectors is in the Hilbert space i.e. a compact space. Additionally, a Hilbert space must possess a properly defined inner product. Informally, a Hilbert space is a set of objects (kets, vectors, functions, etc.) where the angles between and the lengths of the objects are well defined. When a space only possesses the notion of length it is called a Banach space.

	Classical	Quantum
Normalization	$\sum_k p_k = 1$	$Tr[\rho] = 1$
Positivity ($\forall k$)	$p_k \in \mathcal{R}^+$	$\langle k|\rho|k\rangle \geq 0$
Expectation values	$\sum_k p_k f(k)$	$Tr[\rho F]$ where $F = F^\dagger$
Marginals	$\sum_k p_{km} = p_m$	$Tr_B[\rho^{AB}]$
Combining systems	$\Omega_1 \times \Omega_2$	$\mathcal{H}^A \otimes \mathcal{H}^B$
Entropy	$-\sum_k p_k \log p_k$	$-Tr[\rho \log \rho]$

Table 3.1: **Comparison of quantum and classical information.** The normalization and positivity in both the quantum and classical case allow for a probabilistic interpretation. The quantum positivity requirement is basis independent. Combining sample spaces of two probability distributions occurs using the Cartesian product while the tensor product is used for combining quantum systems. The entropy in the classical and quantum case align when the probabilities are the eigenvalues of the density matrix. This occurs when the density matrix is diagonal.

of the dual space (the input on the left of the inner product) is called the bra and written as $\langle\cdot|$. The inner product of a bra, $\langle b|$ and a ket, $|k\rangle$ is written as $\langle b|k\rangle$ which looks like a 'braket'. The outer product of the two elements is then $|k\rangle\langle b|$ and is an element of a matrix.

In the course of discussing the differences between quantum and classical information, the similarities are most easily seen using the quantum density matrix ρ introduced in the previous chapter, see Table 3.1. Given a wave function $|\psi\rangle$, $\rho = |\psi\rangle\langle\psi|$; however, given a set of wave functions, $\{|\psi_k\rangle\}$, with probabilities $\{p_k\}$, $\rho = \sum_k p_k|\psi_k\rangle\langle\psi_k|$. Classical probability distributions can be 'lifted' to quantum objects by using an orthogonal set of wave functions. In this way, diagonal density matrices are classical and the quantum mechanical attributes are contained in the coherences, or off-diagonal elements of the density matrix. The interface between the classical and quantum probability vectors will be more fully explored in the context of Markov process in Chapter 10.

Classical information is predicated on the notion of a bit which is either a '0' or a '1'. Similarly, the basic unit of quantum information is the quantum bit, or qubit—a two level quantum system. Any two level quantum mechanical system can be considered a qubit, and there is a wide variety of experimental implementations ranging from superconductors, to nuclear spins, to polarized light, which are discussed later in Section 3.5. A qubit can exist in an arbitrary superposition state $\alpha|0\rangle + \beta|1\rangle$ where α and β are complex numbers such that $|\alpha|^2 + |\beta|^2 = 1$, see Fig. 3.3.

The states of n qubits are elements of an exponentially large, 2^n-dimensional, Hilbert space, spanned by a basis of the form $|x_1\rangle \cdots |x_n\rangle \equiv |x_1 \ldots x_n\rangle$, where each $|x_i\rangle$ is $|0\rangle$ or $|1\rangle$. The states $|0\rangle$ and $|1\rangle$ label the upper and lower eigenstates of σ^z. The choice of σ^z as the computational basis is arbitrary but conventional. Matrix representations of operators and states are written in this basis unless otherwise stated.

Since two-level systems can be described as spin-half particles, the relevant (Pauli) matrices are

$$\sigma^x = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \qquad \sigma^y = \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix} \qquad \sigma^z = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$$

Notice that we have defined the Pauli matrices with eigenvalues ± 1 instead of $\pm \hbar/2$. Together with the two-dimensional identity matrix, $\mathbf{1}$, the Pauli matrices form an operator basis for two-level systems. The eigenvectors of σ^z are labeled as $|0\rangle$ and $|1\rangle$ corresponding to the eigenvalues $+1$ and -1 respectively.

Similar to the Cartesian product for sample spaces, the Hilbert space of two quantum systems are combined using the tensor product. For finite Hilbert spaces, the matrix representation of the tensor product is the Kronecker product. That is, for a $n \times m$ matrix A and a $p \times q$ matrix B, A tensor B is given by a $np \times mq$ matrix as follows

$$A \otimes B = \begin{bmatrix} a_{11}B & a_{12}B & \cdots & a_{1n}B \\ a_{21}B & a_{22}B & \cdots & a_{2n}B \\ \vdots & \vdots & & \vdots \\ a_{m1}B & a_{m2}B & \cdots & a_{mn}B \end{bmatrix}$$

3.3.2 Entanglement

The ability of quantum states to exist in superpositions of basis states is a precursor to a uniquely quantum mechanical phenomenon called entanglement: joint states of a combined system that cannot be written as a (tensor) product of states of the individual systems — the individual systems have no well-defined state on their own but their joint state can be perfectly defined, hence they are 'entangled.' As an example of an entangled state, the two-qubit Bell state $|\Phi^+\rangle = (|00\rangle + |11\rangle)/\sqrt{2}$ cannot be written as a product state $|\phi_1\rangle \otimes |\phi_2\rangle$. Note, just as superpositions depend on the choice of basis states, entanglement is dependent on the partitioning of the Hilbert space i.e. where the division between system A and system B is placed can affect measures of entanglement between the subsystems.

For a more quantitative description, the Schmidt decomposition is used. The Schmidt decomposition is essentially a restatement of the singular value decomposition (SVD) discussed in Appendix C. Given a pure state, $|\psi\rangle$, that is partitioned into two subsystems: $\mathcal{H}_A$ and $\mathcal{H}_B$, we define the Schmidt decomposition as

$$|\psi\rangle = \sum_j^{\chi} \lambda_j |a_j\rangle |b_j\rangle$$

The set $\{a_j\}$ ($\{b_j\}$) spans $\mathcal{H}_A$ ($\mathcal{H}_B$) and is an orthonormal set. The set of positive scalars, $\{\lambda_j\}$, are called the Schmidt coefficients, $\sum_j \lambda_j^2 = 1$, and the number of non-zero Schmidt coefficients, χ, is known as the Schmidt rank. If $\chi = 1$, the state is a product state and is unentangled with respect to the partitioning of the Hilbert space. In the example of the Bell state, $|\Phi^+\rangle$, the Schmidt rank is 2.

A measure of bipartite entanglement can be defined using the reduced states.

$$\rho_A \quad = \quad Tr_B[|\psi\rangle\langle\psi|] = \sum_b \langle b| \left[\sum_{jk}^{\chi} \lambda_j^2 |a_j\rangle\langle a_k| \otimes |b_j\rangle\langle b_k| \right] |b\rangle \tag{3.1}$$

$$= \quad \sum_j^{\chi} \lambda_j^2 |a_j\rangle\langle a_j| \tag{3.2}$$

Similarly, $\rho_B = \sum_k \lambda_k^2 |b_k\rangle\langle b_k|$. The entropy of entanglement is now defined in the following equations

$$E[\Psi] \quad = \quad -Tr[\rho_A \log \rho_A] = -Tr[\rho_B \log \rho_B] \tag{3.3}$$

$$= \quad -\sum_j \lambda_j^2 \log \lambda_j^2 \tag{3.4}$$

The entropy of entanglement is invariant under local operations, continuous, additive and always bounded by $\log(\chi)$. This is the unique measure of entanglement for bipartite systems with these properties; however, there is a large field surrounding entanglement witnesses and measures of entanglement beyond the bipartite case. See Ref. [10] for a review.

When the Schmidt rank is much lower than the dimension of the Hilbert space, truncations of the Hilbert space can be made, allowing classical computers to represent the state efficiently [221]. This realization has resulted in a host of algorithms that truncate the Hilbert based on entanglement [189, 44].

3.4   Quantum computation

Quantum computing requires the ability to prepare a large number of qubits into some initial state, manipulate the state in an arbitrary way, (for example, using quantum logic gates) and read out the final state, much like the requirements of classical computing. However, there are fundamental differences between quantum and classical computers discussed now.

3.4.1   Differences between quantum and classical computation

The linearity of quantum theory implies that a quantum computer can execute classical computations in superposition. For example, if the input state contains all possible input values $|x\rangle$ of a function $f(x)$, the function can be computed using a unitary operation U_f as

$$\sum_x a_x |x\rangle|0\rangle \xrightarrow{U_f} \sum_x a_x |x\rangle|f(x)\rangle \tag{3.5}$$

With a single call to U_f, the quantum computer produces a state that contains information about all the possible outputs of $f(x)$.

Nevertheless, quantum computation has several limitations. First, the no-cloning theorem [155, 104] states that an unknown quantum state cannot be copied perfectly. More

importantly, the information of a general quantum state cannot be read out with a single projective measurement, because that would collapse a superposition into one of its components. Therefore, although the state in Eq. (3.5) contains information about all possible outputs, that information is not immediately accessible. Instead, a quantum algorithm has to be designed in a way that makes it easy to measure a global property of f, without making it necessary to compute all the individual $f(x)$. A second limitation is that the computation performed must be reversible. Since unitary operations always have an inverse (namely their hermitian conjugate), this implies that computations performed by a quantum computer are reversible. Classically, Bennett showed that all classical circuits could be made reversible using extra bits (called ancilla bits) as scratch bits without incurring significant overhead [22]. A third limitation, is measurement of the quantum state necessarily restricts extractable information. This prevents using a quantum computer to "explore all possible paths" using superposition since measurement will ruin the superposition. If one were able to explore all possible paths, this would be equivalent to simulating a non-deterministic computer, as described in Chapter 4. This brute force search for a solution cannot be performed arbitrarily fast as discussed in Appendix D. The limited readout means algorithms that provide quantum advantage must leave the answer in a quantum state that can be easily measured. Measurement schemes are discussed further in Chapter 5.

3.4.2 Models of quantum computation

There are several models of quantum computation. Most work in quantum simulation has been done in the circuit and adiabatic models. Although the two are known to be computationally equivalent—any computation that can be performed in one model can performed in the other in a comparable amount of time [6, 107, 149]—different problems are solved more naturally in different models. We discuss the two models in turn, but note that other models hold promise for the development of future simulation algorithms, including topological quantum computing [112, 153], one-way quantum computing [181, 182], and quantum walks [106].

The circuit model

A cornerstone of quantum computation is the generalization of the classical circuit model, composed of classical bits and logical gates, to a quantum circuit model [63, 18, 155]. A quantum circuit is a multi-qubit unitary transformation U which maps a set of initial states to some final states. Usually, a unitary gate is decomposed into elementary gates that involve a few (one or two) qubits each.

In classical computing, the NAND gate is universal [158], meaning that any logical circuit can be constructed using NAND gates only. Similarly, in quantum computing, there are sets of unitary operations that form universal gate sets. A quantum computer that can implement such a set is called universal, and can perform any unitary transformation U to an arbitrary accuracy. It turns out that the set containing all single-qubit gates in addition to any two-qubit entangling gate is universal [104], see Fig. 3.3. An entangling gate is realized by any physical interaction that can generate entanglement between qubits.

The single qubit Bloch sphere rotations, R_x, R_y, and R_z gates can generate any

other single qubit gate where R_n is defined as $\exp[-i\sigma^n\theta/2]$ for real angle θ. The Hadamard gate defined as

$$\mathsf{H} = \frac{1}{\sqrt{2}} \left[\begin{array}{cc} 1 & 1 \\ 1 & -1 \end{array} \right]$$

is especially important, as it will be seen throughout this thesis. The two-qubit CNOT gate is an example of an entangling gate. It leaves one qubit space unchanged and acts with σ^x on the second qubit when the first qubit is in the state $|1\rangle$. The gate is given as $\mathsf{CNOT} = |1\rangle\langle 1| \otimes \sigma^x + |0\rangle\langle 0| \otimes \mathbf{1}$.

The CNOT gate is also an example of a gate that is conditionally applied depending on the state of the control qubit. By "conditional," it means that if the control qubit is in the logical 1 state, then the gate is applied; if the control qubit is in the state 0, then the gate is not applied. Controlled gates will play an important role in Chapter 5. Examples of experimental implementations of quantum gates have been reviewed [124], and some of the experiments relevant to quantum simulation are discussed in Section 3.5.

Besides the elementary gates, an important quantum transformation is the quantum Fourier transform (QFT). It transforms any quantum state $|\varphi\rangle = \sum_x \varphi(x)|x\rangle$ into its Fourier representation

$$U_{QFT}|\varphi\rangle = \sum_{k=0}^{N-1} \tilde{\varphi}(k)|k\rangle \tag{3.6}$$

where $\tilde{\varphi}(k) = (1/\sqrt{N})\sum_{x=0}^{N-1} \varphi(x)e^{2\pi ikx/N}$ are the discrete Fourier coefficients of $\varphi(x)$. The QFT can be efficiently implemented using a quantum circuit [155]: for n qubits, the number of elementary gates required is quadratic in the number of qubits. For comparison, the classical fast Fourier transform requires a number of gates that scales like $n2^n$. The exponentially faster QFT is used to in Chapter 5 for the measurement of observables and in Chapter 6 for the simulation of quantum dynamics. Note that the Hadamard transform is the binary QFT.

A perfect qubit is decoupled from the environment such that the Schrödinger equations (SE) governs the evolution and only unitary operations, e.g. quantum circuit elements, are allowed. Non-unitary effects are typically the result of the environment and are grouped into the term decoherence. Usually, the effects of the environment are assumed negligible and left as a problem for the experimentalist. This restriction on the operations manifest itself in the circuit model of quantum computing by requiring that all circuit elements are unitary with the sole exception of measurement.

Adiabatic quantum computation

An alternative to the gate model is the adiabatic model of quantum computation introduced to the field of quantum information and quantum computation by Ref. [72]. In this model, the quantum computer remains in its ground state throughout the computation. The Hamiltonian $H(t)$ of the computer is changed slowly from a simple initial Hamiltonian H_i to a final Hamiltonian H_f whose ground state encodes the solution to the computational problem. The adiabatic theorem states that if the variation of the Hamiltonian is sufficiently slow, the easy to prepare ground state of H_i will be transformed continuously into the ground
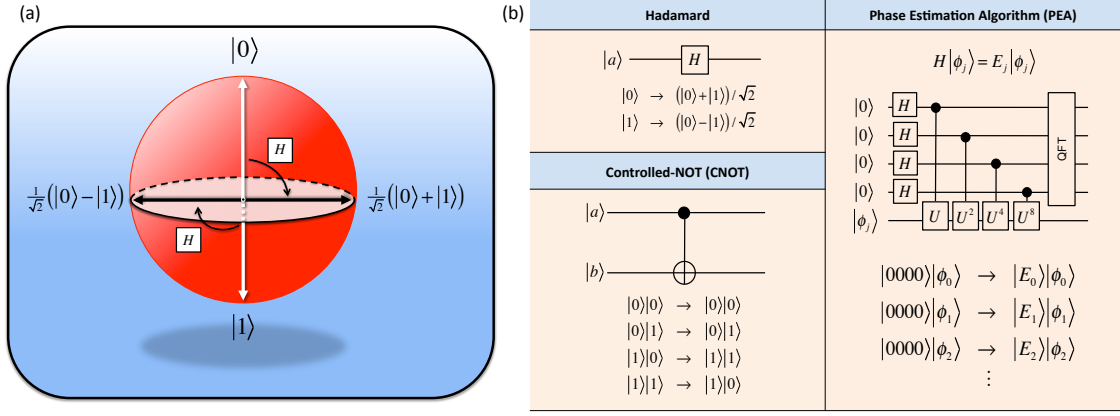
Figure 3.3: **The qubit, elementary gates, and the phase estimation algorithm.** (a) The quantum state of a qubit can be represented in a Bloch sphere. (b) The action of the Hadamard gate, H, on a qubit is shown in (a). The CNOT gate, together with single qubit gates, form a universal gate set for quantum computation. The quantum circuit for the phase estimation algorithm, see Chapter 5, is shown on the right. Here $U |\phi_k\rangle = e^{2\pi i E_k} |\phi_k\rangle$ and QFT is the quantum Fourier transform. The eigenvalues in this case have four-digit accuracy.

state of H_f. It is desirable to complete the evolution as quickly as possible; the maximum rate of change is mostly determined by the energy gap between the ground and first excited states during the evolution by the adiabatic theorem [146]. The adiabatic approximation to bound the evolution speed has been a topic of literature discussion [11, 216, 215, 33]. The applications of adiabatic quantum computation to simulation include preparing quantum states of interest and solving optimization problems such as protein folding [170].

3.5   Experimental progress

Practical requirements for qubits and their manipulation was originally outlined by DiVincenzo [66], and experimental progress toward satisfying the DiVincenzo criteria were reviewed in Ref. [124]. Since this thesis focuses mainly on theory, a brief resume of recent work is given with an emphasis on simulation. Key challenges in building a quantum computer are to isolate qubits in physical systems, prepare arbitrary states, implement arbitrary qubit evolution, read out qubit states, overcome noise, and to do all this on a large scale; that is, with a large number of qubits. Today, there is a large ongoing international effort to build a quantum computer using a wide range of different physical systems. Significant progress has been made with ionic [188, 130], photonic [159], superconducting [171], and solid state [167] systems, for example. Experimental quantum simulation has rapidly progressed [42, 39] since the early simulation of quantum oscillators. Here we review chemical applications of available quantum computational devices.

Quantum optics

On an optical quantum computer, various degrees of freedom of single photons, such as polarization or path, are used to encode quantum information [160, 119]. The optical approach offers a path to building a large-scale quantum computer provided single photon sources and high quality detectors [115, 119]. It has also provided a test bed in which to explore the next steps in experimental quantum computing. Today, the available photonic quantum computer technology is reaching the stage where the first small-scale quantum algorithms can be explored [128, 127].

This architecture was used for the first quantum simulation of a molecular system, a minimal-basis model of the hydrogen molecule H_2, described in Chapter 8. Qubits were encoded in photon polarization, while two-qubit gates were implemented probabilistically using linear-optical elements and projective measurement. The minimal-basis description of H_2 used two spin-orbitals per atom. Since the FCI Hamiltonian is block-diagonal with 2×2 blocks, two qubits sufficed for the experiment: one for storing the system wave function and one for the readout of the PEA. The PEA was implemented iteratively, extracting one bit of the value of the energy at a time. Twenty bits of the energy were obtained, and the answer was exact within the basis set.

Nuclear magnetic resonance

Nuclear spins can serve as qubits, being addressed and read out using an NMR spectrometer [19]. The first experimental quantum simulation, of a harmonic oscillator, was performed using NMR [198]. The platform has since been used to simulate a number of model systems [154, 40, 233, 168], leading up to the recent simulation of H_2 [70] which is similar to the simulation described in Chapter 8. The experiment used ^{13}C-labeled chloroform, in which the carbon and hydrogen nuclear spins formed two qubits. The NMR experiment achieved a higher precision than the optical simulation using a modified measurement scheme. An important contribution of Ref. [70] was the experimental realization of adiabatic state preparation. Adiabatic state preparation and their experimental realization will be discussed further in Chapter 7. The most recent experimental effort using NMR was the simulation of laser driven isomerization of malonaldehydes [141].

Superconducting systems

The circulating current (clockwise or counterclockwise) flowing in a micrometer-sized loop of a superconductor can be used as a qubit [234, 133]. Examples of applications based on superconducting qubits include the tailor-made generation of harmonic oscillator states [92] and the implementation of the Deutsch-Jozsa and Grover quantum algorithms [64]. Recently, the free-energy function for the four-amino-acid peptide assisted by a chaperone protein has been experimentally realized [169]. A microprocessor consisting of an array of coupled superconductor qubits has been used to implement time-dependent various Hamiltonians with $H_i \propto \sum_i \sigma_i^x$ as the initial Hamiltonian [101, 97, 83]. The quantum hardware operating at a temperature of 20 mK found the correct solution with a probability of 78%. Characterization of this device is currently underway [85, 25, 84, 126]. This technology will continue to play a key role in the search for a full fledged quantum computer

as evidenced by the first sale of a commercial superconducting quantum computing device to Lockheed-Martin [145]. Quantum simulation of atomic collision with super conducting qubits was recently proposed [178] but awaits experimental implementation.

Trapped ions

Qubits can also be encoded in the electronic states of cold trapped ions, offering one of the most controllable systems available today [56, 96, 31]. This platform has already produced sophisticated simulations of physical systems [76, 110, 77], but chemical applications are still to come.

Chapter 4

Computational complexity in quantum chemistry

"There's no Moore's law for dilution refrigerators."

–Peter Love, regarding asymptotic analysis of quantum algorithms implemented with superconducting qubits.

4.1 Introduction

Quantum chemistry is concerned with solving the Schrödinger equation for chemically relevant systems. By solving a differential or eigenvalue equation, the properties of the system and the dynamics of the state are obtained. Examples of properties include: equilibrium geometries, the dissociation energy of molecules, and the vibrational frequency of bonds. The difficulty is the apparent exponential growth of the problem with the number of electrons as well as the quality of the description of the system. For practical applications, the accuracy required is typically orders of magnitude smaller than the total energy of the system. As a concrete example, an atom of carbon has total energy of about 1000 eV while the bond of carbon to hydrogen is only 5 eV.

Solving the eigenvalue equation can be done using Gaussian elimination and a matrix of size N requires on the order of N^3 operations. When N is exponential in the problem size, this algorithm is no longer considered efficient and a host of sparse matrix algorithms and iterative techniques have been developed to address shortcomings of the straight-forward approaches, see Chapter 2 for examples. Similarly, computational chemistry has bifurcated into a host of methods for diagonalizing the Hamiltonian as well as calculating other properties of interest.

To understand the computational advantages of quantum algorithms for chemical simulation, I will discuss some aspects of computational complexity theory. This chapter focuses on the total number of operations necessary to complete a computational task. As a first problem, the Hartree-Fock procedure typically quoted as scaling less than $O(M^4)$, will be shown to be NP-complete. Following that, the attention turns to complexity results

for reduced density matrices, demonstrating the QMA-completeness of N-representability and the QMA-hardness of the universal functional of density functional theory.

4.2   Preliminary notions

The definitions of the key complexity classes P (Polynomial), BQP (Quantum Polynomial, the quantum version of P), NP (Non-deterministic Polynomial), QMA (Quantum Merlin Aurthur, the quantum version of NP), and QMA(k) (Quantum Merlin Aurthur with k provers) are the ones relevant for the remainder of this paper. Before defining them, a brief introduction to the idea behind computational complexity is given and a more thorough discussion can be found in any textbook on computer science e.g. [195, 14].

A proper measure of the *time complexity* of an algorithm is how many basic operations (or how much time) it takes to solve problems of increasing size. Conventionally, a computational problem is described as easy or tractable if there exists an efficient algorithm for solving it, one that scales polynomially with input size (for an input of size n, as $O(n^k)$ for some fixed k).[1] Otherwise, the problem is considered intractable. In some situations, this distinction is rough one: for reasonable problem sizes, an "inefficient" algorithm scaling exponentially as $O(1.0001^n)$ would be faster than an "efficient" $O(n^{1000})$ algorithm. Nevertheless, this convention has proved useful. From a theoretical perspective, the division allows for considerable progress to be made without considering the minutiae of the specific system or implementation.

From a practical standpoint, Moore's "law" states the density of transistors in classical computers doubles every two years. This exponential growth is expected to cease sometime this century. Thus, for a fixed computational time, if the computer cannot solve an instance using a polynomial-time algorithm, one need not wait long as the exponential growth of computing power will reasonably quickly overtake the fixed polynomially large cost. However, if the algorithm runs in exponential-time, one may be forced to wait several lifetimes in order for an instance to become soluble.

The time complexity can be characterized using different though equivalent computational models. In the context of computer science, *equivalent* means that one model can simulate the other with polynomial overhead which respects the boundary between efficient and inefficient algorithms alluded to in the previous paragraph. Turing machines and circuits are two typical models discussed in the context of computational complexity. Additionally, the automata is another important model but will not be discussed here. In computer science, any set of finite strings is called a *language* and if the computational device *accepts* (returns True or outputs 1) on all strings contained in a language then the device is said to *recognize* the language. If the machine is also required to halt on all inputs, then it is said to *decide* the language. The difference being, a Turing machine may recognize a language but fail to halt on some instances i.e. run forever. Alternatively, languages can be thought of as *decision problems* where the strings in the language are the accept

---

[1]The notation, $f(x) = O(g(x))$, implies that f is dominated by g at asymptotically large values of x. To indicated that f is asymptotically larger than g, we write $f(x) = \Omega(g(x))$. If f is dominated by and dominates function g, we write $f(x) = \Theta(g(x))$. This notation is used throughout the thesis.

instances.

A Turing machine is specified by six items: the set of states that the machine can be in Q, the alphabet for the inputs Σ, the tape alphabet Γ which contains alphabet Σ, the transition rule δ, the reject state, and an accept state. In this model, the transition rule, δ, takes the current machine state and the value of the tape, then decides what the value on the tape should be changed to, what state the machine should transition to, and whether to move left or right. The machine either halts when it enters an accept or reject state and returns the appropriate value, or it loops forever. That is, the Turing machine partitions the set of input strings into accept, reject, or undecided groups.

While the notion of the Turing machine may sound (and in many ways is) contrived, it was conceived by Alan Turing in 1936 to answer David Hilbert's (of 'Hilbert space' fame) question about deciding if a polynomial has roots which are integers in a "finite number of operations." This gives rise to the modern notion of the algorithm which was first defined using Turing machines (and the equivalent Alonzo Church's lambda calculus). The *Church-Turing*[2] thesis says an algorithm is defined as any procedure a Turing machine can perform. For the record, deciding if the roots are integral is not decidable (although recognizable) using a Turing machine.

The circuit model of computation has been more useful when generalizing to the quantum setting. The time complexity characterizes the number of circuit elements (or gates) used to compute a boolean value given an input. So long as a universal set of gates are used, the time complexity will only differ by a polynomial factor. The key caveat to consider when using this model is the notion of uniformity. To properly define an algorithm in the circuit model, a way of generating circuits for all input sizes must be given concisely. A circuit family for an algorithm is *uniform* if there is a polynomial-time algorithm that specifies the circuit given the input size.

4.2.1 Computational problems and problem instances

Computational problems have a very specific meaning in computer science: computational problems are collections of problem instances that a computer may be employed to solve. Each question is called a problem instance and the answer is called the solution to that instance.

As an example, the computational problem of deciding if a molecule has a dipole moment, Dipole, is a collection of questions; each instance about a different molecule. The questions 'Does BCl_3 have a dipole moment?', and 'Does NH_3 have a dipole moment?' are instances of this computational problem. The string $x=$"BCl_3" is not in the language Dipole since "BCl_3" does not have dipole moment while "NH_3" is in the language since it has a dipole moment.

As a second example, consider Sudoku puzzles. In the newspaper, instances of Sudoku puzzles are printed, but, it is the collection of all Sudoku puzzles that is the computational problem Sudoku. Since Sudoku puzzles are not decision problems, they do not

---

[2]The naming convention in computer science is always alphabetical, hence 'Aa-' last names are always first authors.

have a corresponding language. However, VALID SUDOKU is the language of all nine by nine matrices that are valid puzzles.

4.2.2 Reductions

If the decision problem A can be solved when an oracle for B is available, then A is Turing reducible to B. A stronger requirement is that problem instance of type A be convertible into problem instance of type B. This is known as the *Karp reduction*. Reducibility is a way of formalizing the relationship between problems. Essentially, it asks, "If I can solve one problem, what other problems can I solve using the same resources?"

Turing reduction. Also called 'Cook reduction.' A language A is polynomial time Turing reducible to language B, if a polynomial time machine with access to an oracle for B can decide membership in the language A in polynomial time.

Karp reduction. Also called 'many-one reduction.' A language A is Karp reducible to language B if a polynomial time computable function, f, exists such that for every $x \in B$, it is the case that $f(x) \in A$.

For a distinction between the two types of reductions, consider trying to determine if the heat capacity at constant pressure, C_p, of a substance is less than some critical value (language A). Now, suppose an oracle, probably in the form of a bomb calorimeter, has been provided for deciding if the heat capacity at constant volume, C_V, of an given substance is above or below a critical value. Call language $B(cv_0)$ the set of strings labeling substances with C_v below the critical value, cv_0. Then by adding the number of moles times the ideal gas constant to cv_0, one can determine membership in language A via the formula $C_p = C_V + nR$. Thus, A is Karp reducible to B.

Suppose instead, an oracle for evaluating if at fixed pressure the internal energy, U, of the given substance at a given temperature is less than some critical value (language C) is given. Then by evaluating the internal energy at two different temperatures, the heat capacity at constant pressure can be determined. Because the reduction has to use the oracle more than once, A is Turing reducible to C.

4.2.3 Promise problems

Promise problems will play a key role throughout the remainder of the text and it is easy to understand their importance. Reconsider the previous example, where language A is the string corresponding to substances with C_p less than some critical value. Imagine you have a terrible lab instructor, who gives you a substance that has heat capacity *nearly* identical to the critical value that decides membership in A. It may take a large number of repetitions of the experimental protocol to be certain that the substance belongs or does not belong to language A.

A reasonable lab instructor would announce at the beginning of lab that all the substances he is handing out for the experiment are promised to be at least one Joule per Kelvin away from the critical value. Given this promise, the student will be able to complete

their lab in a reasonable amount of time. That is the computational problem of deciding if the substance given in A should take one lab period. Without this promise, it may take one student three or more lab periods (should they be so diligent).

In many ways, promise problems are just guarantees on the format of the inputs to the program. Formally,

Promise problem. A promise problem is defined by languages, L_{accept} and L_{reject} where $L_{accept} \cap L_{reject} = \emptyset$. The promise is that $x \in L_{accept} \cup L_{reject}$.

Note that promise problems are defined using two languages, whereas the decision problems are defined using one language (and its complement).

4.2.4 Relevant polynomial time complexity classes

The class of all decision problems, those with a yes-or-no answer, that are easy (tractable) for classical computers (classical Turing machines) is called P.

Polynomial time complexity class. P is the class of languages that are decidable in polynomial time on a deterministic Turing machine. That is a language L is in P if and only if there exists a deterministic Turing machine M that runs for polynomial time on all inputs and behaves as follows:

- if $x \in L$, M accepts.

- if $x \notin L$, M rejects.

If the Turing machine is no longer deterministic, i.e. it has a random number generator, then the class of languages is BPP, defined as:

Bounded-error probabilistic polynomial time complexity class. BPP is the class of languages that are decidable in polynomial time on a probabilistic Turing machine. That is, a language L is in BPP if and only if there exists a deterministic Turing machine M that has access to a random number generator running for polynomial time on all inputs and behaves as follows:

- if $x \in L$, M accepts with probability greater than 2/3.

- if $x \notin L$, M rejects with probability greater than 2/3.

Moving beyond probability to quantum mechanics, the class BQP can now be defined where the probabilistic Turing machine is replaced by a quantum Turing machine [62, 26] i.e. a quantum computer.

Bounded-error quantum polynomial time complexity class. BQP is the class of languages that are decidable in polynomial time on a quantum Turing machine. That is, a language L is in BQP if and only if there exists a quantum Turing machine M running in polynomial time on all inputs and behaves as follows:

- if $x \in L$, M accepts with probability greater than 2/3.

- if $x \notin L$, M rejects with probability greater than 2/3.

4.2.5 Relevant non-deterministic complexity classes

So far, the discussion has centered on complexity classes containing problems that are solvable in polynomial time given access to various resources (a computer, a random number generator and a controllable quantum system). Now, attention turns to the class of problems that can be computed non-deterministically.

The original notion of non-determinism is a Turing machine whose transition rule maps the machine state and the tape symbol to any number of possible outputs. In essence, this is a machine that can clone itself at will to pursue all options at once. In reality, there appears to be no such machines; however, NP is the class of problems that could be solved in polynomial time by such a machine. Whether a deterministic polynomial time algorithm can be used to simulate such a machine is a restatement of the famously open question of $P \overset{?}{=} NP$.

Rather than resorting to imaginary machines, the non-deterministic classes can be defined by considering verification of solutions obtained in some non-deterministic way. The class of all problems whose answer can be verified in polynomial time with a Turing machine corresponds to the class NP. For example, even though it is not known how to factor numbers efficiently, factoring is in NP because we can check the proposed answer efficiently by multiplication. Note that $P \subseteq NP$ because any easy problem can be verified easily.

Non-deterministic polynomial time complexity class. NP is the class of languages, L that have polynomial-time Turing machine verifiers. That is, for a given input, x, there exists a polynomial time algorithm M so that:

- if $x \in L$, then there exists y such that $M(x, y) = 1$.

- if $x \notin L$, then for all y, $M(x, y) = 0$.

The string y, is called the *certificate* or the *proof* of x being in L.

Giving the verifier access to a probabilistic machine leads to the MA complexity class.

Merlin-Arthur complexity class. MA is the class of languages that have a polynomial time probabilistic verifiers. That is, given input x there exists a polynomial time algorithm R that has access to a random number generator which behaves as follows:
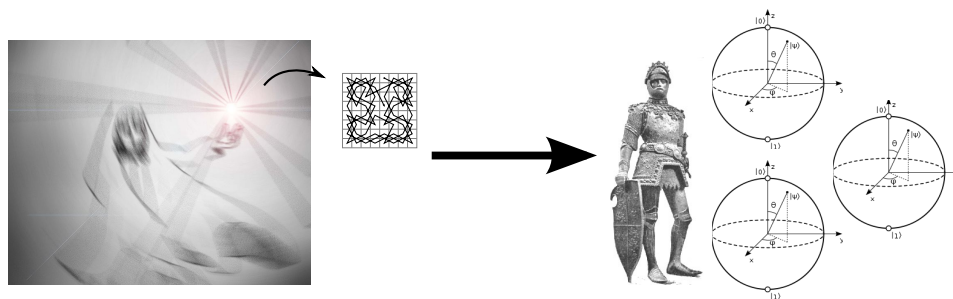
Figure 4.1: **Merlin-Arthur complexity classes.** In the Merlin-Arthur games, Merlin is a wizard who gets his answers magically (hence non-deterministically) but he is a incompetent wizard who is prone to errors. Arthur is seeking the solution to some problem and when Merlin gives him a valid proof, Arthur should usually accept (completeness) and when Merlin gives him an proof of something incorrect, he should be able to spot the inaccuracy (soundness). The idea of QMA is depicted, where Merlin has produced a knight's tour of a chess board (a Knight must visit every square on a chess board once and only once) and King Arthur has access to qubits to verify the tour. Figure composed from images found on Wikipedia under Creative Commons License.

- if $x \in L$, then there exists y such that $R(x, y)$ accepts with probability greater than 2/3.

- if $x \notin L$, then for all y, $R(x, y)$ rejects with probability greater than 2/3.

Finally, if Arthur has access to a quantum computer, the complexity class is QMA, see Fig. 4.1.

Quantum Merlin-Arthur complexity class. QMA is the class of languages that have a polynomial time quantum verifiers. That is, given input x and a polynomial time algorithm U which behaves as follows:

- if $x \in L$, then there exists $|y\rangle$ such that final state $U|x, y\rangle$ is accepted with probability greater than 2/3.

- if $x \notin L$, then for all $|y\rangle$, final state $U|x, y\rangle$ is rejected with probability greater than 2/3.

The conjectured relation among the complexity classes mentioned above are illustrated in Fig. 4.2.

4.2.6 Completeness and hardness

The goals of classifying computational problems in relationship to complexity classes motivates introduction of the terms *hard* and *complete*. The classification of a computational problems as hard for a complexity class means if an algorithm can solve this
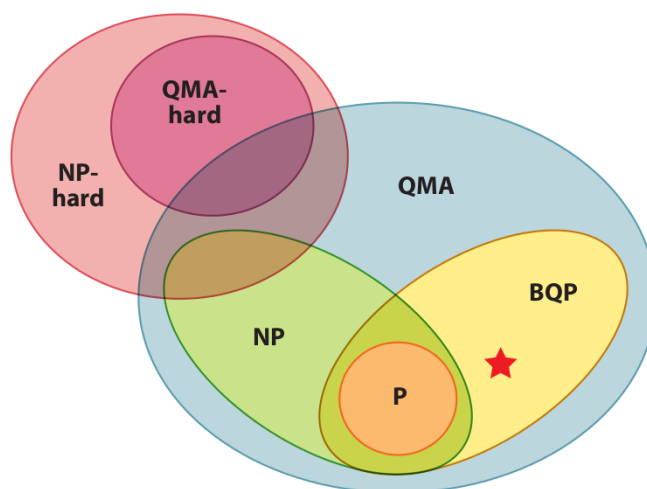
Figure 4.2: **Computational complexity classes.** Shown are the conjectured relations among the computational complexity classes discussed in this chapter. The star indicates where the conceptual emphasis of this thesis lies.

problem, then all problems in the class can be solved efficiently. Complete problems for a class are, in some sense, the hardest problems in that class; these problems are both in the class and hard for the class. More precisely,

Hard. A language B is hard for a complexity class $\mathbf{C}$ if every language A in $\mathbf{C}$ is polynomial time reducible to B.

Complete. A language B is complete for a complexity class $\mathbf{C}$ if every language A in $\mathbf{C}$ is polynomial time reducible to B *and* B is in $\mathbf{C}$.

4.3   Computational chemistry problems

4.3.1   Quantum chemistry

Before continuing, relevant information from Chapter 2 is recalled. For quantum chemistry, the annihilation, $\{a_k\}$, and creation operators, $\{a_k^\dagger\}$, correspond respectively to removing and adding an electron into a particular single particle wave function, $\{\phi_k(\mathbf{x})\}$. The single particle wave functions is called orbitals, and the set of single particle wave functions are called the basis set. The number of functions in the basis set will be denoted M throughout this chapter. Anti-symmetry of the N-electron wave function is enforced by

the canonical anti-commutation relations

$$[a_p,\ a_q]_+ = a_p a_q + a_q a_p = 0 \qquad\qquad [a_p,\ a_q^\dagger]_+ = \delta_{pq}\mathbf{1}$$

The vacuum state, $|vac\rangle$, is the normalized state that the creation/annihilation operators act on and it represents the system with no particles.

4.3.2 Computational problems

k**-LOCAL HAMILTONIAN.** Given n qubits and a collection of $poly(n)$ Hamiltonian terms $\{H_m :\ 0 \le \|H_m\| \le 1, H_m$ affects $\le k$ qubits$\}$, is the ground state energy (lowest eigenvalue) of $H = \sum H_m$ less than a or are all energies greater than b, promised that $b - a \ge 1/poly(n)$?

Ref. [113] showed that 5-LOCAL HAMILTONIAN is QMA-complete and it was subsequently reduced to 2-LOCAL HAMILTONIAN [107] and to 2-LOCAL SPATIAL HAMILTONIAN [162] where the couplings are restricted to a 2D lattice.

Since chemists are not interested in qubit systems, a more interesting problem is the FERMIONIC LOCAL HAMILTONIAN. This was defined and shown to be QMA-complete in [142]. The proof is given after the definition.

FERMIONIC LOCAL HAMILTONIAN. Given an N fermion system and M single particle orbitals with $N \le M \le poly(N)$, does a two-fermion Hamiltonian A with norm between 0 and 1 have ground state energy less than a or greater than b promised that $b - a \ge 1/poly(N)$?

To show that FERMIONIC LOCAL HAMILTONIAN is in QMA, note the quantum algorithm phase estimation can be used to verify a proof state given by Merlin. The probability for the accept and reject cases follow from the error bounds on phase estimation. See Chapter 5 for a full discussion of the phase estimation algorithm. Next, we show that FERMIONIC LOCAL HAMILTONIAN is QMA-hard.

Lemma 4. *FERMIONIC LOCAL HAMILTONIAN is QMA-hard.*

Proof. By mapping qubit Hamiltonians onto fermionic Hamiltonians it is seen that instances of FERMIONIC LOCAL HAMILTONIAN can embed instances of 2-LOCAL HAMILTONIAN and thus is a Karp reduction. This is done using a construction identifying each qubit with a pair of fermionic modes analogous to appending spin functions to a set of spatial orbitals. For n qubits, let $M = 2n$ and for each qubit, say $|q_i\rangle = \alpha|0_i\rangle + \beta|1_i\rangle$, write $\alpha\ a_{2i}^\dagger + \beta\ a_{2i+1}^\dagger$. Next, the Pauli matrices map to quadratic terms. For example, the Pauli sigma X operator becomes

$$\sigma_i^x = |0_i\rangle\langle 1_i| + |1_i\rangle\langle 0_i| \leftrightarrow a_{2i}^\dagger a_{2i+1} + a_{2i+1}^\dagger a_{2i}$$

and similarly for σ^y and σ^z.

Finally, the double occupancy must be penalized and this can be done by adding a penalty to each pair of qubits, $P_i = C(2a_{2i}^\dagger a_{2i} - \mathbf{1})(2a_{2i+1}^\dagger a_{2i+1} - \mathbf{1})$. The penalty $P = \sum_i P_i$ is also quadratic and commutes with the fermionic Pauli matrices. The constant C is selected so that the ground state has the correct occupancy. $\square$

Note in the proof that FERMIONIC LOCAL HAMILTONIAN is QMA-hard, the mapping used was a Karp reduction, since there is a mapping (using spin-orbitals) that transforms members of the 2-LOCAL HAMILTONIAN language to the FERMIONIC LOCAL HAMILTONIAN language. A similar Karp reduction is used later in the proof of Theorem 5, but all the other reductions will be Turing reductions where an oracle for problem A is required to decide problem B. FERMIONIC LOCAL HAMILTONIAN captures most of what chemists are interested in, however one can impose more structure. Namely,

ELECTRONIC STRUCTURE. Given an N electron system, M orbitals, a configuration of static nuclei, and a value of energy, E_T: Is the ground state energy of

$$H_{elec} = T_e + V_{ee} + V_{en} = \sum_{ij} T_{ij} a_i^\dagger a_j + V_{ijkl} a_i^\dagger a_j^\dagger a_k a_l + V_{ij}^{eN} a_i^\dagger a_j$$

less than $E_T - \delta/2$ or greater than $E_T + \delta/2$ promised that $\delta > 1/poly(N)$?

Notice that this is slightly different than asking for fixed error ϵ. The promise of a gap that scales with the system size follows from considering the classical time-frequency uncertainty from Fourier transforms. Since the scheme that Arthur uses with his quantum computer for the Hamiltonian problems is based on phase estimation, see Chapter 5, the uncertainty in the frequency is determined by the length of propagation time. Thus, if his measurement is to be performed in an amount of time that scales polynomially with the input size, the frequency can only be determined to inverse polynomial accuracy. Note that this argument is completely classical and based solely on the Fourier uncertainty. Theoretically, measurements of the energy can be done in arbitrarily short amounts of time. See Ref. [7] for an example.

4.4 Hartree-Fock is NP-complete

Hartree-Fock is one of the most important quantum chemistry techniques as it typically recovers over 99% of the non-relativistic Born-Oppenhemier energy, it provides the mathematical setting for the chemist's notion of molecular orbitals, and it is the starting point for many other, more accurate techniques. See Chapter 2 for more information on the Hartree-Fock method. The implementation of Hartree-Fock requires evaluating and manipulating $O(M^4)$ two-electron integrals. However, since Hartree-Fock solves a nonlinear eigenvalue equation through an iterative method, see Fig. 2.2, convergence is the key concern that prevents the actual scaling from being polynomial. The computational complexity result, proving that the unqualified scaling cannot be polynomial (unless P= NP), was provided in an unpublished appendix of [190] available only on the arXiv.

For the purposes of this chapter, the Hartree-Fock procedure can be succinctly explained as the minimization of the energy of an N-electron system given M basis functions with the restriction to a single configuration[3], that is

$$|HF\rangle = b_1^\dagger b_2^\dagger \cdots b_N^\dagger |vac\rangle$$

---

[3]The actual ground state wave function is some superposition of configurations: $\Psi_0 =$

for some set $b_i^\dagger = \sum_j C_{ij} a_j^\dagger$ related to the original set of orbitals via unitary transformation. The problem HARTREE-FOCK will be defined analogous to the other Hamiltonian problems with a promise given on the gap.

HARTREE-FOCK. Given an N-electron system and M single-particle basis functions, decide if the Hartree-Fock energy $E_{HF} = \langle HF|H|HF \rangle$ is less than a or greater than b promised that $b - a \geq 1/poly(N)$.

Theorem 5. *HARTREE-FOCK is NP-complete [190]*

Proof. Showing HARTREE-FOCK is in NP is straightforward but the proof that it is NP-hard relies on showing that $|HF\rangle$ is exact for an NP-hard spin glass Hamiltonian that has been mapped to fermions using the pair construction above.

To prove inclusion in the complexity class NP, an algorithm for *verifying* the energy in polynomial time must be given. If the coefficient matrix C describing the correct orbital rotation is given, then energy is calculated easily using Condon-Slater rules, see Tables 2.2 and 2.3. Thus, given the answer, the accept or reject conditions are easily verified.

To prove that any other NP problem can be mapped to the Hartree-Fock problem (NP-hardness), first, a quick review of Ising models. The Hamiltonian of Ising models is $H_{ising} = J_{ij} S_i S_j$ with $J_{ij} \in \{+1, 0, -1\}$ and $S_i \in \{+1, -1\}$. Ref. [17] showed that although $1 \times L \times L$ Ising lattice can be solved by Onsager's solution, $2 \times L \times L$ lattice is NP-complete. Stronger results show all non-planar graphs (even in two-dimensions) are NP-complete [95].

Next, using the fermion pair construction as in Lemma 4, the following equation can be derived

$$S_i S_j = \sigma_i^z \sigma_j^z = \sum_{p,q=0,1} (-1)^{p+q} n_{2i+p} n_{2j+q}$$

Since a penalty Hamiltonian is added to prevent double occupancy, a promised gap of $1/poly(N)$ is required for HARTREE-FOCK. Finally, the satisfying assignment of the spin glass is some bit string $|s_1 s_2 \cdots s_n\rangle$ for $n = 2L^2$ and $s_i = \pm 1$. The correct and exact Hartree-Fock solution for the n-electron wave function should assign $b_i^\dagger = a_{2i}^\dagger$ when $s_i = 0$ and $b_i^\dagger = a_{2i+1}^\dagger$ when $s_i = 1$. This is a Karp reduction since there is a mapping from members of one language to another. □

4.5 Reduced wave function approaches

Many computational chemistry algorithms seek to minimize the energy or other properties by manipulating the wave function. To quote Coulson [59], "wave functions tell us more than we need to know... All the necessary information required for energy and calculating properties of molecules is embodied in the 1st and 2nd order density matrices."

$\sum_K C_K (a_{K_1}^\dagger a_{K_2}^\dagger \cdots a_{K_N}^\dagger)|vac\rangle$ with complex valued C_K constrained such that $\sum_K |C_K|^2 = 1$. Optimizing over C_K is called Full Configuration Interaction and the known algorithms have exponential complexity and are currently limited to very small systems.

k^{th} **order reduced density matrix.** (k-RDM) is defined as a partial trace over all but k of the fermions similar to the marginals of a multipartite probability function

$$\rho^{(k)} = Tr_{k+1,k+2,\ldots N}[\rho^{(N)}]$$

In fact, the Hohenberg-Kohn theorem, see Section 2.7, proves that the probability density n (the diagonal of the 1-RDM with $n_i \equiv \rho_{ii}^{(1)}$) obtained from the ground state wave function of electrons in a molecular system is in one-to-one correspondence with the external potential e.g. from the nuclei. Therefore, all properties of the system are determined by the probability density. Unfortunately, the proof is not constructive, so most properties of the system are practically inaccessible. On the other hand, using the 2-RDM, it is easy to evaluate properties of the electronic system as the Hamiltonian only contains two-electron terms. However, problems arise when determining if the 2-RDM is valid or invalid. Unconstrained optimization of the energy using all 2-RDMs can lead to ground state energies of negative infinity if the validity of the 2-RDMs is not determined. This is known as the *N-representability* problem. Formally,

N-REPRESENTABILITY. Given N fermions and M single particle basis functions and error tolerance $\beta \geq 1/poly(N)$, decide if (i) a given 2-RDM, $\rho^{(2)}$, is equal to the 2-RDM of an N-fermion state, $\mu^{(N)}$, or (ii) if ρ is bounded away from all other 2-RDM consistent with an N-fermion state by

$$\|Tr_{3,\cdots,N}(\mu^{(N)}) - \rho\|_{tr} \geq \beta$$

Indeterminate cases do not matter.

Next, attention turns to the computational problem that arises when using the one-electron probability density: UNIVERSAL FUNCTIONAL. Before defining the problem of UNIVERSAL FUNCTIONAL, recall how it is used. In the ELECTRONIC STRUCTURE problem, the Hamiltonian has a specific form: $(T_e + V_{ee}) + V_{en}$. For different problem instances V_{en} varies but the $T_e + V_{ee}$ terms are always the same. The universal functional of DFT, introduced in Chapter 2, takes as input the probability density and returns the lowest possible energy of $T_e + V_{ee}$.

UNIVERSAL FUNCTIONAL. Given N electrons and an probability density $n(x) = \rho^{(1)}(x, x)$, approximate
$$F[n] = \min_{\mu^{(N)} \to n} Tr[(T_e + V_{ee})\mu^{(N)}]$$
to an polynomial accuracy in the number of electrons.

Notice this problem is hard for the ELECTRONIC STRUCTURE language as V_{en} can be computed efficiently and accurately since it is a single electron operator, $Tr[\rho^{(1)}V_{en}]$. In the next subsection, the proof that N-REPRESENTABILITY is QMA-complete is sketched. Details of the intricacies of geometric optimization are skipped and interested readers are referred to the original proof [142].

4.5.1 N-REPRESENTABILITY

Theorem 6. *N-REPRESENTABILITY is QMA-complete.*

Proof. The idea of the proof is that a membership oracle for 2-RDM solves FERMIONIC LOCAL HAMILTONIAN (hardness) and Merlin could send a polynomial number of copies of the actual state using the proper qubit encoding of a consistent $\mu^{(N)}$ (inclusion in the class). With an oracle for N-REPRESENTABILITY, the space of valid 2-RDMs can be characterized as a vector space of dimension l with $l \leq poly(N)$ and $l \leq poly(M)$ accomplished by using a complete set of observables on the space of pairs of fermions. Since N-REPRESENTABILITY only considers fixed number of electrons, only two-electron operators need to be considered as $a_i^\dagger a_j = a_i^\dagger \left(\sum_k^M \frac{a_k^\dagger a_k}{N-1} \right) a_j$. Next, a convex program (shallow-cut ellipsoid algorithm) minimizes the expectation value of the energy. Technical considerations for the convex program are found in [142]. Because the oracle is used many times for the convex program, this is not a Karp reduction.

Next, demonstrating the that N-REPRESENTABILITY is in QMA merely requires asking Merlin to send $\mu^{(N)}$. To encode the fermionic state in qubits, the Jordan-Wigner transformation discussed in Subsection 6.4.2 is used. Polynomial copies of the proof are sent to the verifier, who first checks the value of $\sum a_k^\dagger a_k = \sum_k |1\rangle\langle 1|_k$, which should be N. Then, Arthur picks observables randomly from a set of two-electron operator similar to the Pauli matricies to test the expectation value of the reduced state comparing with the outcomes of similar measurement on the proof state until convinced. Merlin might try to cheat by sending entangled copies of $\mu^{(N)}$, but it can be proved that Arthur cannot be fooled based on the Markov inequality.[4] See Ref. [4] for a proof of this fact.   □

4.5.2   UNIVERSAL FUNCTIONAL

Theorem 7. *UNIVERSAL FUNCTIONAL is QMA-hard.*

Proof. Although N-REPRESENTABILITY proves that FERMIONIC LOCAL HAMILTONIAN is QMA-hard, it does not demonstrate that the more specialized case of ELECTRONIC STRUCTURE is QMA-hard. This was done by Ref. [190] en-route to showing that UNIVERSAL FUNCTIONAL is QMA-hard.

The proof begins with the QMA-complete 2-LOCAL SPATIAL HAMILTONIAN

$$H_{2loc} = \sum_{\langle ij \rangle} c_{ij} \tau_i \otimes \nu_j$$

with τ_i and ν_j sigma matrices at sites i and j, and c_{ij} non-zero iff i and j are neighbors on a 2D lattice. From H_{2loc}, a series of reductions to a special case of ELECTRONIC STRUCTURE is achieved. Along the way, a number of common models of magnetism are shown to be QMA-complete.

---

[4]Given any random variable X, Markov's inequality states $\langle |X| \rangle \geq aP(|X| \geq a)$ with $|X|$ the absolute value of X and $\langle X \rangle$ defining the expectation value of X.

A key technical contribution of the proof is the use of perturbative gadgetry, see Section 9.4 for an example, to reduce H_{2loc} to a 2D Heisenberg Hamiltonian with local field,

$$H_{heis} = J \sum_{ij} \sum_{k \in \{x,y,z\}} \sigma_i^k \sigma_j^k + \sum_i \sum_{k \in \{x,y,z\}} B_i^k \sigma_i^k$$

Next, the reduction of the Heisenberg model to the Hubbard model

$$H_{hub} = -t \sum a_{i,s}^\dagger a_{j,s} + U \sum_i n_{i,\uparrow} n_{i,\downarrow} - \sum_i \sigma_i \cdot B_i$$

was previously known [58]. Using second order perturbation theory, at half filling the *ground state* of the Hubbard model can reproduce the ground state of arbitrary Heisenberg Hamiltonians. The intuition is essentially the same as in the proof of Lemma 4 whereby each electron occupies one of two sites corresponding to a spin pointing up or pointing down. The final reduction shows the Hubbard model with local fields arises as a special case of ELECTRONIC STRUCTURE for carefully chosen V_{en}. Although the Hubbard model is a phenomenological model, in Ref. [190], the authors show that the Hubbard model with local field arises with an appropriate external electrostatic potential with $1/poly(N)$ corrections.

As a result of this string of reductions, given an oracle for UNIVERSAL FUNCTIONAL, ELECTRONIC STRUCTURE can be solved in polynomial time using convex optimization since the density matrices form a convex set and the functional is convex. Since every Hubbard Model is an instance of ELECTRONIC STRUCTURE, UNIVERSAL FUNCTIONAL is QMA-hard. □

4.5.3 Extensions to pure states

The pure state generalization of both the N-REPRESENTABILITY and UNIVERSAL FUNCTIONAL have interesting consequences. First, consider PURE N-REPRESENTABILITY, where the question is now: "Did the 2-RDM come from a pure state?" (instead of a mixed state as in N-REPRESENTABILITY). The verifier, Arthur, can now be cheated by a prover who sends entangled states. However, if multiple unentangled provers are used the soundness property remains intact, thus PURE N-REPRESENTABILITY is QMA(k). In Ref. [86], it was shown that QMA(k)=QMA(2).

If the two non-communicating Merlins send the state that k Merlins would have sent, Arthur could be cheated if they entangle states that should not be entangled. Using the pairwise swap test (a quantum algorithmic version of the Hong-Ou-Mandel effect) on each of the supposedly unentangled states reveals if the given state is not a product state (or nearly so). If the state from the first Merlin passes the test, the state from the second Merlin can be used in the verifying circuit. See Ref. [86] for more details.

Next, consider the optimization over pure states given the universal function, PURE UNIVERSAL FUNCTIONAL. The optimization over pure states is harder since the states are no longer convex like the set of density matrices. However, given the oracle for PURE UNIVERSAL FUNCTIONAL, the optimization is still in NP causing

$$\text{NP}^{\text{ UNIVERSAL FUNCTIONAL}} = \text{QMA}$$

4.6   Other topics

This literature review focused on the difficulty of three approaches to electronic structure. There are several other results worth mentioning. While proving that #BQP= #P, Ref. [38] showed that calculating the density of states is complete for #P, the class of counting problems. Troyer and Wiese [217] showed that the sign problem of Quantum Monte Carlo is NP-hard, while [180] briefly touched on the complexity class of Bohmian mechanics.

Also, it is worth mentioning that a number of computational complexity results exist for quantum information based wave function ansatzs such matrix product states and its generalizations. However, these results were not reviewed as they would warrant a separate, more thorough consideration.

Chapter 5

Measurements and the phase estimation algorithm

> We dance and round in a ring and suppose,
> But the Secret sits in the middle and knows.

–Poem by Robert Frost

Molecular energies are the eigenvalues of an associated time-independent Hamiltonian H and can be efficiently obtained to fixed accuracy, using a quantum algorithm with three distinct steps [15]: encoding a molecular wave function into qubits; simulating its time evolution using quantum logic gates; and extracting the approximate energy using the phase estimation algorithm [111, 1]. The latter is a general-purpose quantum algorithm for evaluating the eigenvalues of arbitrary Hermitian or unitary operators. The algorithm estimates the phase ϕ accumulated by a molecular eigenstate $|\Psi\rangle$ under the action of the time-evolution operator, $U = e^{-iHt}$; that is

$$e^{-iHt}|\Psi\rangle = e^{-iEt}|\Psi\rangle = e^{-i2\pi\phi}|\Psi\rangle \tag{5.1}$$

where E is the energy eigenvalue of $|\Psi\rangle$. Therefore, estimating the phase for each eigenstate amounts to estimating the eigenvalues of the Hamiltonian. The details of the phase estimation are the concern of this chapter and the other steps for obtaining the molecular energies will be described in subsequent chapters.

In principle, an individual measurement can be carried out in any basis. However, as experimental measurement techniques usually address individual qubits, a method is needed to carry out more complicated measurements. In particular, to measure an observable A, one would like to carry out a measurement in its eigenbasis $\{|e_k\rangle\}$. This is achieved by the phase estimation algorithm (PEA) [111, 2]

$$\frac{1}{\sqrt{N}} \sum_t |t\rangle|\psi\rangle \xrightarrow{C-U} \frac{1}{\sqrt{N}} \sum_{k,t} c_k e^{-iA_k t}|t\rangle|e_k\rangle \xrightarrow{QFT} \sum_k c_k |A_k\rangle|e_k\rangle \tag{5.2}$$

where $c_k = \langle e_k|\psi\rangle$ and A_k are the eigenvalues of A; $C-U$ is the unitary $U = \exp(-iAt)$ controlled by the ancilla qubits, which are initialized in the state $\frac{1}{\sqrt{N}}\sum_{t=2}^{N}|t\rangle$. When

measuring the ancilla, the eigenvalue A_k will be measured with probability $|c_k|^2$ and, if the eigenstates are non-degenerate, the wave function will collapse to the eigenvector $|e_k\rangle$.[1] For the PEA to be efficient, it must be possible to simulate the propagator e^{-iAt} efficiently.

To obtain ground state energies, PEA relies on the assumption that the input wave function has signification overlap with the ground state. Since each qubit represents the occupancy of molecular orbitals in the N-electron wave function, the HF guess for the ground state $|\psi_0^{HF}\rangle$ requires no superposition of states and is thus straightforward to prepare. For $|\langle\psi_0^{HF}|\psi_0^{FCI}\rangle| = 1 - \epsilon$, where ϵ is small, the phase estimation algorithm can be applied to retrieve an estimate of the ground state energy. Simultaneously, the state of the system will collapse to $|\psi_0^{FCI}\rangle$ when measured in the H^{FCI} basis (via PEA) with high probability, see Section 5.2 below. If the Hartree-Fock guess is insufficient, more sophisticated state preparation procedures exist, as discussed in Chapter 7.

The following section discusses to the mathematical particulars of phase estimation, Section 5.2 focuses on error analysis, and Section 5.3 reviews various applications of quantum measurement to quantum chemistry. The next chapter will be concerned with the simulation of dynamics.

5.1 The phase estimation algorithm

In this section, I describe how to obtain molecular energies given the time evolution of the molecular Hamiltonian described above. The key idea is to Fourier transform the oscillating phase, $\langle\psi_n(0)|\psi_n(t)\rangle = \exp(-iE_nt)$, to obtain the electronic energy. The eigenenergy is converted into relative phases. The relative phase can then be measured using the quantum phase estimation algorithm (PEA). As the phase is measured, the input state partially collapses to the set of states consistent with the measurements obtained up to that point.

To determine an eigenvalue associated with an eigenstate, consider the phase of an eigenstate of the Hamiltonian, H, evolving dependent on a register qubit being in state $|1\rangle$.

$$|0\rangle|\psi_n\rangle + e^{-iHt}|1\rangle|\psi_n\rangle = |0\rangle|\psi_n\rangle + e^{-iE_nt}|1\rangle|\psi_n\rangle \qquad (5.3)$$

By letting $E_n = 2\pi(\phi - K)/t$ where $0 \leq \phi < 1$ and K is an unobservable integer, the unknown eigenvalue becomes encoded in the relative phase of the register qubit as $|0\rangle + e^{-2\pi i(\phi-K)}|1\rangle$ [166, 15, 111]. Then, ϕ is estimated using the phase estimation technique detailed in next subsection.

Once ϕ has been estimated, it must be inverted to obtain the energy. Given bounds for the energy eigenvalue, $[E_{min}, E_{max})$, the time of evolution is selected as $t = 2\pi/(E_{max} - E_{min}) = 2\pi/\omega$ and an energy shift of E_s is used to ensure $K = (E_s - E_{min})/\omega$ is an integer. The energy shift, E_s, is affected by a gate acting on the register qubit which applies a phase to the qubit if it is in state $|1\rangle$ and does nothing otherwise. Using these parameters the measured value of ϕ corresponds to the value of the energy, $E_\phi = \omega(\phi - K) + E_s$.

---

[1]Other methods for eigenvalue measurement include pairing adiabatic quantum evolution with Kitaev's original scheme as described in Chapter 9 and applications of the Hellmann-Feynman theorem [161].

This analysis follows from straightforward signal processing based on the Shannon-Nyquist sampling theorem where the unobservable integer, K, is responsible for aliasing.

The PEA can be accomplished in several different ways depending on the technology available and in the recent experimental implementations, including ours described in Chapter 8, the iterative phase estimation algorithm was used. This scheme is described next. The original reference for the iterative scheme is Ref. [166] which builds from Kitaev's exposition on quantum measurements [111]. Other useful expositions are found in Refs. [214, 67, 232, 116].

5.1.1   Iterative phase estimation

The goal of the phase estimation algorithm is to obtain the best L bit estimate of a phase ϕ. The readout will be values of 0 or 1 from the fractional binary expansion of ϕ

$$\phi = 0.j_1 j_2 \ldots j_L = \left(\frac{j_1}{2^1}\right) + \left(\frac{j_2}{2^2}\right) + \cdots + \left(\frac{j_L}{2^L}\right) \tag{5.4}$$

Each j_k is zero or one. The relative phase generated by the unitary propagator $U(t_0) = \exp(-iHt_0)$ applied as in Eq. (5.3) is defined as $\exp(-2\pi i\phi)$. Performing $U(2^k t_0)$ results in a relative phase $\exp(-2\pi i(2^k \phi))$. Observing that $2^k \phi = j_1 \ldots j_k j_k.j_{k+1} \ldots j_L$ we can find the binary expansion of length L deterministically.

First, $U(2^{L-1} t_0)$ is implemented resulting in relative phase $\exp(-2\pi i\, j_1 \ldots j_{L-1}.j_L) = \exp(-2\pi i\, j_L/2)$. Since j_L is 0 or 1, the relative phase is either $+1$ or -1, respectively. The Hadamard transform described in Subsection 3.4.2 distinguishes $(|0\rangle + |1\rangle)/\sqrt{2}$ from $(|0\rangle - |1\rangle)/\sqrt{2}$, allowing a projective measurement in the computational basis to identify j_L.

To obtain j_{L-1} and more significant bits, the iterative method uses the gate S_k to deterministically obtain j_k. Since data is read from the least significant digit first, the counter-rotation S_k is a function of previously obtained bits. The form of the S_k gate is

$$\mathsf{S}_\mathsf{k} = \begin{bmatrix} 1 & 0 \\ 0 & \Phi_k \end{bmatrix} \quad \text{with} \quad \Phi_k = \exp\left[2\pi i \sum_{l=2}^{L-k+1} \frac{j_{k+l-1}}{2^l}\right] \tag{5.5}$$

This gate removes the $L - k - 1$ least significant digits so the state of quantum computer becomes $(|0\rangle + \exp[-i\pi j_k]|1\rangle)|\psi_n\rangle$ where j_k is zero or one. Finally, effecting the Hadamard transformation leads to $|j_k\rangle|\psi_n\rangle$ and measurement of the register in the $\{|0\rangle, |1\rangle\}$ basis yields the value of j_k. When the binary expansion of ϕ is length L, the measurements are deterministic, otherwise the remainder causes some randomness but does not significantly affect the results. A complete error analysis can be found in Section 5.2.

At the end of the kth iteration, a measurement of the logical state of the top qubit is performed, yielding the result 0 or 1. This represents an estimate of the kth bit of the binary expansion of ϕ. Because each iteration provides one bit, the total number of iterations (L) determines the precision with which the phase is estimated. An important feature is that least significant bits are estimated first i.e. k is iterated backwards from L to 1 and that measured bits are used to improve the estimation of more significant bits.

As first noted in Ref. [40], in the phase estimation algorithm, the number of uses of $U(t_0)$ for a fixed time t_0 scales exponentially with the number of bits desired from ϕ.
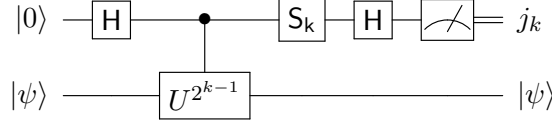
Figure 5.1: **Iterative phase estimation circuit for obtaining the kth bit.** The phase is represented using a binary decimal expansion $\phi = 0.j_1 j_2 \cdots j_L$. To gain precision in the algorithm, one adjusts the parameterized gate $\mathsf{S_k}$ according to the values of all previous obtained bits. U is a representation of the propagator of the system being simulated. Acting on the eigenstate with this operator advances the phase of the state allowing bit j_k to be read out deterministically.

This is a consequence of the Fourier uncertainty principal; the more information required in the frequency domain, the longer the propagation time. When the unitary time-propagator is decomposed into gates, this means an exponential increase in the number of gates is required for an exponential increase in the precision of the measurement. In total, for L bits of precision, the state must be prepared once and the unitary propagator $U(t_0)$ is used

$$\sum_{k=0}^{L-1} 2^k = 2^L - 1$$

times. Unless otherwise noted $U = U(t_0)$.

5.1.2 Estimating overlaps

Suppose one is interested in obtaining the expectation value of some positive semi-definite self-adjoint operator A with spectral radius $r(A)$. Following the analysis in the start of this section, the time of evolution to encompass the range of eigenvalues is $t = 2\pi/\omega$ with $\omega = r(A)$. In this way, any eigenvalue of A corresponds uniquely to a value of the phase $0 \leq \phi < 1$. If the operator is unbounded or the spectral radius cannot be estimated, a bound on the spectral support of the input state will suffice with ω equal to the difference between the highest and lowest eigenvalues of the support.

If input state ψ is an eigenstate, then the analysis performed in the previous section follows through directly as the phase will encode the eigenvalue through $\lambda_\phi = \omega\phi$.

Estimating the expectation value of a unitary operation can be done using amplitude estimation algorithm. The algorithm described originally in Ref. [34] combines the PEA and Grover search to obtain the amplitude of an operator. Appendix D contains a review of Grover search. The overarching idea is to convert amplitudes into phases and then apply phase estimation. This is done by performing phase estimation on the Grover operator given in Eq. (D.9). Rotations of the form, Eq. (D.15), have eigenvalues, $\lambda_\pm = e^{\pm\theta}$, which can be phase estimated given state w or r, see Fig. D.3 for a visual representation of these states. From λ_+ or λ_-, the value of θ can be established. Since $2\phi = \theta$ as in Eq. (D.19), the (absolute) overlap between s and w can be determined.

Estimating overlaps, as described in [116], assumes access to a unitary preparation procedure for $|s\rangle$. Then define $|w\rangle = U|s\rangle$ for some circuit U. In this way, $\langle w|s\rangle = \langle s|U|s\rangle$. Therefore, from the estimate for θ the overlap is calculated using

$$\sin\frac{\theta}{2} = |\langle w|s\rangle| \tag{5.6}$$

5.2 Error analysis

In this section, the analysis of the precision of the measurement is considered disregarding the accuracy of the evolution and assuming that an eigenstate is given. The analysis presented here is based on Refs. [67, 116]. Suppose the phase $0 \leq \phi < 1$ is only being estimated to L bits and the complete binary expansion of ϕ is expressed as

$$\phi = \sum_{k=1}^{\infty} \frac{j_k}{2^{k+1}}$$

with j_k equal to zero or one. If the phase is only measured to L bits

$$
\begin{aligned}
\phi - \tilde{\phi} &= \sum_{k=1}^{\infty} \frac{j_k}{2^k} - \sum_{k=1}^{L} \frac{j_k}{2^k} \\
&= \sum_{k=L+1}^{\infty} \frac{j_k}{2^k} = \sum_{k'=1}^{\infty} \frac{j_{k'+L}}{2^{k'+L}} \\
&= \frac{1}{2^L} \sum_{k'=1}^{\infty} \frac{j_{k'+L}}{2^{k'}} \equiv \frac{\delta}{2^L}
\end{aligned}
\tag{5.7}
$$

Here $0 \leq \delta < 1$.

The process of estimating ϕ begins with phase-kickback using the controlled version of the time-propagator

$$\frac{1}{\sqrt{2}}[|0\rangle|\psi_n\rangle + |1\rangle U^{2^{L-1}}|\psi_n\rangle] = \frac{1}{\sqrt{2}}[|0\rangle|\psi_n\rangle + (e^{-2\pi i\phi})^{2^{L-1}}|1\rangle|\psi_n\rangle]$$

The resulting relative phase can be simplified as $2^{L-1}\phi = j_1 \ldots j_{L-1}.j_L + \delta/2 = K + (j_L + \delta)/2$ with K some irrelevant integer. This leads to $\exp[-2\pi i(2^{L-1}\phi)] = \exp[-2\pi i(j_L + \delta)/2]$. The Hadamard transformation applied to the first qubit leads to the state

$$\frac{1}{\sqrt{2}}[|0\rangle|\psi_n\rangle + c|1\rangle|\psi_n\rangle] \xrightarrow{\text{H}} \frac{1}{2}[(1+c)|0\rangle + (1-c)|1\rangle]|\psi_n\rangle \tag{5.8}$$

with $c \equiv \exp[-2\pi i(j_L + \delta)/2]$. The probability of measuring zero is

$$
\begin{aligned}
P(M_L = 0) &= \frac{1}{4}|c+1|^2 = \frac{2 + c + c^*}{4} \\
&= (2 + e^{-2\pi i(j_L+\delta)/2} + e^{2\pi i(j_L+\delta)/2})/4 \\
&= \frac{2 + 2\cos[-2\pi(j_L+\delta)/2]}{4} \\
&= \frac{1 + \cos[-\pi(j_L+\delta)]}{2}
\end{aligned}
\tag{5.9}
$$

Considering the possible priors

$$P(M_L = 0|j_L = 0) = \frac{1 + \cos[-\pi\delta]}{2} = \cos^2[\pi\delta/2],$$

and

$$P(M_L = 0|j_L = 1) = \frac{1 + \cos[-\pi\delta - \pi]}{2} = \frac{1 - \cos[-\pi\delta]}{2} = \sin^2[\pi\delta/2].$$

Thus, the probability of getting the correct answer is

$$P(M_L = j_L) = P(M_L = 0|j_L = 0) = P(M_L = 1|j_L = 1) = \cos^2[\pi\delta/2] \tag{5.10}$$

Subsequent bits If the Lth bit was correct, counter-rotation yields

$$0.j_1 j_2 \cdots j_{L-1} j_L + \frac{\delta}{2^L} \tag{5.11}$$

$$\xrightarrow{U^{2^{L-2}}} \quad K.j_{L-1}j_L + \frac{\delta}{2^L}2^{L-2} = K.j_{L-1}j_L + \frac{\delta}{4} \tag{5.12}$$

$$\xrightarrow{S_{L-1}} \quad K.j_{L-1}0 + \frac{\delta}{4} \tag{5.13}$$

Similarly, if the first m-bits were correct, then

$$0.j_1 j_2 \cdots j_{L-m}j_{L-(m-1)} \cdots j_L + \frac{\delta}{2^L} \tag{5.14}$$

$$\xrightarrow{U^{2^{L-m-1}}} \quad K.j_{L-m}j_{L-(m-1)} \cdots j_L + \frac{\delta}{2^L}2^{L-m-1} \tag{5.15}$$

$$\xrightarrow{S_{L-m}} \quad K.j_{L-m}0 + \frac{\delta}{2^{m+1}} \tag{5.16}$$

The relative phase factor becomes $c = \exp[-2\pi i(j_{L-m}/2 - \delta/2^{m+1})]$. Inserting this into Eq. (5.9) leads to

$$P(M_{L-m} = 0|S_{L-m}) = \frac{1}{4}|1 + c|^2 = \frac{2 + c + c^*}{4} \tag{5.17}$$

$$= \frac{1}{4}\left(2 + 2\cos\left[\frac{-2\pi(j_{L-m} - \delta 2^{-m})}{2}\right]\right) \tag{5.18}$$

where the prior S_{L-m} denotes the previous bits correctly obtained. Next

$$P(M_{L-m} = 0|j_{L-m} = 0, S_{L-m}) = (1 + \cos[\pi\delta 2^{-m}])/2 = \cos^2[\pi\delta 2^{-m-1}]$$

and similarly

$$P(M_{L-m} = 1|j_{L-m} = 1, S_{L-m}) = \frac{1 - \cos[-\pi - \pi\delta/2^m]}{2} = \frac{1 + \cos[-\pi\delta/2^m]}{2} = \cos^2[\pi\delta 2^{-m-1}]$$

Finally, the probability to correctly ascertain the bit, given the prior S_{L-m} is

$$P(M_{L-m} = j_{L-m}|S_{L-m}) = \cos^2[\pi\delta 2^{-m-1}] \tag{5.19}$$

Probability of getting all bits correct Using the definition of conditional probability, probability of correctly obtaining the phase is

$$P(M_L = j_L)P(M_{L-1} = j_{L-1}|S_{L-1})P(M_{L-2} = j_{L-2}|S_{L-2})\cdots$$

$$= P(M_L = j_L)\frac{P(M_{L-1} = j_{L-1}, M_L = j_L)}{P(M_L = j_L)}\frac{P(M_{L-2} = j_{L-2}, M_{L-1} = j_{L-1}, M_L = j_L)}{P(M_{L-1} = j_{L-1}, M_L = j_L)}\cdots$$

$$= P(M_1 = j_1, M_2 = j_2, \cdots, M_L = j_L)$$

From Eqs. (5.10) and (5.19), the probability to get all bits correct is

$$P(M_1 = j_1, M_2 = j_2, \cdots, M_L = j_L) \tag{5.20}$$

$$= \cos^2\left(\frac{\pi\delta}{2^1}\right)\cos^2\left(\frac{\pi\delta}{2^2}\right)\cdots\cos^2\left(\frac{\pi\delta}{2^L}\right) \tag{5.21}$$

$$= \frac{\sin^2(\delta\pi)}{4^L \sin^2(\pi\delta/2^L)} \tag{5.22}$$

Rounding up Now, consider rounding up. Instead of obtaining $\tilde{\phi}$, the desired estimate would be $\tilde{\phi}_r = \tilde{\phi} + 2^{-L-1}$. Following Eq. (5.7), the error obtained using this estimate is $|\phi - \tilde{\phi}_r| = (1-\delta)2^{-L-1}$. When $\delta > 1/2$, the error is lower than that obtained in Eq. (5.7).

Using Eq. (5.10), with δ replaced by $1 - \delta$, obtaining the first bit of $\tilde{\phi}_r$ has probability $\cos^2[\pi(1-\delta)/2]$. Similarly, for the subsequent bits, the probability is $P(M_{L-m} = j_{L-m}^{(r)}|S_{L-1}^{(r)}) = \cos^2[\pi(1-\delta)2^{-m-1}]$. Finally

$$P(\{M_k\} = \{j_k^{(r)}\}) = \prod_{k=1}^{L}\cos^2\left(\frac{\pi(1-\delta)}{2^k}\right) = \frac{\sin^2(\pi(1-\delta))}{4^L \sin^2(\pi(1-\delta)/2^L)} \tag{5.23}$$

The limit of Eq. (5.22) is

$$\lim_{L\to\infty} P(M_\phi = J_\phi) = \lim_{L\to\infty}\frac{\sin^2(\delta\pi)}{4^L \sin^2(\pi\delta/2^L)}$$

$$= \frac{\lim_{L\to\infty}\sin^2(\delta\pi)}{\lim_{L\to\infty}(2^L \sin(\pi\delta/2^L))^2}$$

$$= \left(\frac{\sin(\pi\delta)}{\pi\delta}\right)^2 \tag{5.24}$$

and the limit of Eq. (5.23) is

$$\lim_{L\to\infty} P(M_\phi = J_{\tilde{\phi}_r}) = \left(\frac{\sin(\pi - \pi\delta)}{\pi(1-\delta)}\right)^2 = \left(\frac{\sin(\pi\delta)}{\pi(1-\delta)}\right)^2 \tag{5.25}$$

If $0 < \delta \leq 1/2$, then Eq. (5.24) is preferred and the success probability is lowest at $\delta = 1/2$ thus the success probability is bounded as $P(M_\phi = J_\phi) \geq 4/\pi^2$. If $1/2 < \delta < 1$, then when Eq. (5.25) is used the lowest success probability also occurs at $\delta = 1/2$ and $P(M_{\tilde{\phi}} = J_{\tilde{\phi}}) \geq 4/\pi^2$.

Success probability Finally, the success probability of finding an estimate of ϕ to L bits implies $\tilde{\phi}$ and $\tilde{\phi}_r$ are acceptable answers. The probability of each event is $4/\pi^2$ and the probability that either happens is $8/\pi^2$. Therefore, the error probability is lower bounded by

$$1 - P_{\tilde{\phi}} - P_{\tilde{\phi}_r} < 1 - 8/\pi^2 \tag{5.26}$$

5.3 Other measurement schemes

For the purposes of this thesis, the phase estimation algorithm is the most important technique, as such it will be applied in Chapter 8 to the hydrogen molecule to illustrate the salient features of the model quantum chemistry. In Chapter 9, another measurement model is described in the context of adiabatic quantum computing. To provide a sense of the breadth of the literature, the remainder of this chapter describes other techniques for measurement.

For quantum simulations of chemistry to be interesting, information about the resulting state must be extracted in the end; however, full characterization (quantum state tomography) generally requires resources that scale exponentially with the number of qubits. This is because a measurement projects a state into one consistent with the measurement outcome. Because only a limited amount of information can be extracted efficiently, one needs a specialized measurement scheme to extract the desired observables, such as dipole moments and correlation functions.

For example, if we are interested in molecular energies, the observable is the Hamiltonian H, and we need to simulate the actual dynamics e^{-iHt}, see Chapter 6. Note that the PEA is closely related to classical algorithms for preparing eigenstates by Fourier analysis of a propagating system [90, 211]. As noted earlier, the dynamics must be simulated for longer to achieve a higher precision in the eigenvalues. More precisely, for a final accuracy of ϵ, the PEA must run for a time $O(1/\epsilon)$ [155, 40]. The simulation of dynamics is considered in Chapter 6.

Because quantum measurement is inherently random, repeating a measurement on multiple copies of the same system helps to determine expectation values of observables. The central limit theorem implies that measuring N copies of a state results in a precision that scales as $1/\sqrt{N}$ (the standard quantum limit, SQL). For example, repeating the PEA gives an SQL estimate of the coefficients $|c_k|^2$; these can be used to calculate the expectation value $\langle A \rangle = \sum_k |c_k|^2 A_k$, also to the SQL. When entanglement is available, one can achieve precision scaling as $1/N$—this is the Heisenberg or quantum metrology limit [78]. An algorithm for the expectation values of observables was proposed in Ref. [116] that gets arbitrarily close to the Heisenberg limit.

The first algorithm for measuring a chemical observable was Lidar and Wang's calculation of the thermal rate constant by simulating a reaction in first quantization and using the PEA to obtain the energy spectrum and the eigenstates [137]. These values were used to calculate the rate constant on a classical computer by integrating the flux-flux correlation function. Ref. [103] improved on this method with a more direct approach to the rate constant that efficiently obtains the product branching ratios given different reactant states—if the initial state is a thermal state, see Section 7.3), this gives the rate constant

directly. Furthermore, this method was used to obtain the entire state-to-state scattering matrix. A method for reaction rates using a dedicated quantum simulator in which artificial molecules are experimentally manipulated has also been proposed [196].

More generally, correlation functions provide information about a system's transport and spectroscopic properties. On a quantum computer, the correlation function of any two observables can be estimated efficiently if their pseudo-dynamics can each be simulated efficiently [163, 202]. The method does not suffer from the dynamic sign problem that plagues classical Monte Carlo methods for sampling correlation functions. An alternative approach to the measurement of correlation functions uses techniques of linear-response theory [213].

Molecular properties such as the dipole moment or the static polarizability are also of chemical interest. They are derivatives of the molecular energy with respect to an external parameter, such as the electric field. In Ref. [102], it was shown how to calculate molecular properties by combining PEA and the quantum gradient algorithm of Ref. [99]. The algorithm is insensitive to the dimensionality of the derivatives, an obstacle to classical computers. For example, the molecular gradient and Hessian can be computed—and used to optimize the geometry—with a number of energy evaluations independent of system size.

Chapter 6

Simulation of chemical Hamiltonians

Doing is a quantum leap from imagining.

– Barabara Sher

6.1   Introduction

In the previous chapter, I discussed the measurement scheme of the phase estimation algorithm where the time evolving phase is Fourier transformed to extract the energy of the system. The ability of quantum computers to efficiently simulate the time evolution of quantum systems is leveraged by phase estimation. The prescription for simulating the time evolution of chemical systems on a quantum computer is detailed in this chapter including my work describing the quantum circuits for each non-commuting term in the Hamiltonian.

6.2   Suzuki-Trotter formulas and the unitary propagator

Theoretical quantum simulation relies heavily on the Trotter decomposition to handle non-commuting terms in the Hamiltonian when mimicking the unitary time propagator of the system to be simulated. While there are many ways to exponentiate a matrix [151], the Trotter formula and its generalizations are a class of symmetric approximations that are appropriate for quantum computation since (ignoring environmental coupling) the gates must be unitary. To approximate evolution under a Hamiltonian $H = \sum_{i=1}^{m} H_i$ consisting of m non-commuting but local terms $\{H_i\}_{i=1}^{m}$, one applies the sequence of unitary gates $\{e^{-iH_i t/n}\}_{i=1}^{m}$ a total of n times. As the number of repetitions, n, tends to infinity, the approximation error caused by the non-commutativity vanishes and the approximation converges to the exact result [139]. If each time slice requires a constant number of gates independent of the parameter t/n, then reducing the approximation error by repeating the sequence n times can become expensive for high accuracy applications.

Using a quantum computer for obtaining molecular energies requires efficiently approximating the unitary propagator, $\exp(-iHt)$, for a time sufficiently long to resolve the

Fourier frequency to a desired precision. The present section continues by describing the Trotter decomposition for non-commuting Hamiltonians and generalizations which provide a lower bound on the simulation time.

6.2.1 Trotter decomposition

Using the second quantized representation allows for a straightforward decomposition of the exponential map of each term of the Hamiltonian. However, the terms of this decomposition do not always commute. The goal of the Trotter decomposition is to approximate the time evolution operator of a set of non-commuting operators. The operators are exponentiated individually for small time steps and the procedure is repeated such that their product provides a reasonable approximation to the exponentiation of the sum. Using this approximation, the construction of the time propagator can be efficiently carried out on a quantum computer provided that the Hamiltonian can be decomposed into a sum of local Hamiltonians [139]. The first-order Trotter decomposition is given by

$$e^{-iHt} = \left(e^{-ih_1 \Delta t} e^{-ih_2 \Delta t} \cdots e^{-ih_N \Delta t} \right)^{\frac{t}{\Delta t}} + O(t \Delta t) \tag{6.1}$$

where $t/\Delta t$ is the Trotter number [87]. As the Trotter number tends to infinity, or equivalently $\Delta t \to 0$, the error in the approximation vanishes. If $t/\Delta t$ is not an integer, the remainder is simulated as another Trotter time slice.

Frequently, the Trotter decomposition is applied to the non-commuting kinetic and potential energy operators of the Hamiltonian. This happens often enough to warrant a special name: split-operator technique [74, 121] and it will be discussed further in the context of quantum simulation with first quantization.

There exist higher order approximations (Suzuki-Trotter formulas) which reduce the error even further. For instance, the second-order approximation is given by

$$e^{-iHt} \approx \left(\left(e^{-ih_1 \frac{\Delta t}{2}} \cdots e^{-ih_{N-1} \frac{\Delta t}{2}} \right) e^{-ih_N \Delta t} \left(e^{-ih_{N-1} \frac{\Delta t}{2}} \cdots e^{-ih_1 \frac{\Delta t}{2}} \right) \right)^{\frac{t}{\Delta t}} + O(t(\Delta t)^2) \tag{6.2}$$

Higher order approximations take increasingly more complicated forms [87] and were first studied in the context of black box quantum simulation of sparse Hamiltonians by Berry *et al.* [27]. The authors considered Hamiltonians composed of m efficiently implementable terms and showed that the number of exponentials cannot scale better than linear in the time as measured by units rescaled by the maximum energy of the system. Rescaling the time by the maximum energy of the system is necessary otherwise the difficulty of simulating H and cH would the same when c is a real constant. Multiplying time by the constant $c > 1$ should increase the length of simulation time. To remove arbitrary constants, the simulation of the system scales with $t/||H||$.

Generalizations to time-dependent Hamiltonians has been discussed in Refs. [174, 228] but require promises on the smoothness of the time-dependence to guarantee efficiency. Also, Ref. [228] is intended to serve as a technical review of quantum simulation.

6.2.2 Suzuki-Trotter formulas

Briefly, we review the use of Suzuki-Trotter formulas in quantum simulation for time-independent sparse Hamiltonians, providing an introduction to the literature. Lloyd's well cited 1996 paper [139] confirmed Feynman's suggestion that quantum simulation was possible by introducing Trotterization into the, then emerging, field of quantum simulation. Furthermore, he asserted that physical Hamiltonians are typically k-local to further build the case for quantum simulation.

Refs. [5] and [48] began working on simulating black-box sparse Hamiltonians showing that they, too were efficiently simulatable. Sparsity requires that the number of elements per row of a Hamiltonian be bounded by some polynomial of n when the dimension of the Hilbert space is $D = 2^n$. It is also required that each matrix element can be retrieved efficiently. Ref. [5] used a coloring scheme to decompose the Hamiltonian into a sum of 2×2 block diagonal matrices. This coloring scheme has been updated in several Refs.: [27, 28, 54]. The coloring scheme and blackbox simulation will not be discussed further.

Ref. [27] was the first to approach the general problem of simulating non-commuting Hamiltonians by using higher order Suzuki-Trotter formulas. Ref. [165] has returned to this issue and their contributions will be discussed in turn. The important results of Ref. [27] are (1) the use of higher order Trotter-Suzuki decomposition to bound the number of non-commuting exponentials, N_{exp}, necessary to affect a simulation for some amount of time t, (2) proof of a no-go theorem for sub-linear black-box simulation and (3) improvements upon the coloring scheme of [5] and [48] for black box simulation of sparse Hamiltonians. The simulations in this thesis are concerned with the first two results and they will be explained in more detail after describing the Suzuki-Trotter formulas.

M. Suzuki has studied and extended the Trotter formula essentially continuously since 1990 and this work was reviewed in [87]. The recursive formulas introduced by Suzuki define a fractal pattern where a combination of forward and backward propagation leads to an improved approximation to the desired exponential. Suzuki defines a recursive formula for higher order Trotter formulas in following recursive way beginning with the split operator formula, $e^{Ax/2}e^{Bx}e^{Ax/2}$, for m operators

$$S_2(x) = \left(\prod_{k=1}^{m} e^{h_k x}\right)\left(\prod_{k=m}^{1} e^{h_k x}\right) \tag{6.3}$$

$$S_4(x) = S_2(z_2 x)^2 S_2((1 - 4z_2)x)S_2(z_2 x)^2 \tag{6.4}$$

$$\vdots \qquad \vdots$$

$$S_{2k}(x) = S_k(z_k x)^2 S_k((1 - 4z_k)x)S_k(z_k x)^2 \tag{6.5}$$

The values of z_j are selected so that S_{2j} is correct through $2j^{th}$ order and it can be shown [87] that $z_i = (4 - 4^{1/(2i-1)})^{-1}$. If there are m non-commuting Hamiltonians, then the first order approximation takes $m = N_{exp}$ and for the split operator formula, S_2, the number of exponentials is $2m - 1$. In general, $2(m - 1)5^{k-1} + 1$ exponentials are used for the S_{2k} approximant.

For the kth order Suzuki-Trotter, with m Hamiltonians in the sum, and error

tolerance given by ε, Ref. [27] gives the bound on the number of exponentials as

$$N_{exp} \leq m5^{2k}(m||H||t)^{1+(1/2k)}\varepsilon^{-1/(2k)} \tag{6.6}$$

More recently, [165] present an improvement by noting that the relative ratio of Hamiltonian norms is also important. The idea being if some of the Hamiltonians have very small weight, then their exponentials can be effectively ignored. The improvement they presented was

$$N_{exp} \leq 2(2m-1)5^{k-1}||H_1||t\left(\frac{4emt||H_2||}{\varepsilon}\right)^{1/(2k)}\frac{4me}{3}\left(\frac{5}{3}\right)^{k-1} \tag{6.7}$$

where the sub-Hamiltonians are ordered such that $||H_1|| \geq ||H_2|| \geq \ldots \geq ||H_m||$.

The optimal order of Trotter decomposition, k^*, is determined by selecting the best compromise between time-step length and an decomposition using more exponentials. In Ref. [27]

$$k^* = \text{round}\left[\frac{1}{2}\sqrt{\log_5\frac{m||H||t}{\varepsilon}+1}\right] \tag{6.8}$$

and in Ref. [165] as

$$k^* = \max\left\{\text{round}\left[\sqrt{\frac{1}{2}\log_{25/3}\frac{4em||H_2||t}{\varepsilon}}\right],1\right\} \tag{6.9}$$

In Ref. [27] it was shown for unstructured sparse Hamiltonians $N_{exp} \geq ||H||t$. The proof by contradiction, relied on the lower bounds to quantum mechanical problems based on the polynomial method [20]. These bounds could be violated given sub-linear simulation time. In a departure from the methods discussed so far, Ref. [50] used quantum walk based approaches to push the scaling closer to linear in the reweighed time.

6.3 First quantization

The first quantization method, due to Zalka [236, 229] and improved in Ref. [103] simulates particles governed by the Schrödinger equation on a grid in real space. For a single particle in one dimension, space is discretized into 2^n points, which, when represented using n qubits, range from $|0\ldots0\rangle$ to $|1\ldots1\rangle$. The particle's wave function can be expanded in this position representation as $|\psi\rangle = \sum_{x=0}^{2^n-1} a_x|x\rangle$. The Hamiltonian to be simulated is

$$H = T + V = \frac{p^2}{2m} + V(x) \tag{6.10}$$

and the corresponding time-propagator can be implemented using the quantum version of the split-operator method discussed at the beginning of this chapter

$$U(\delta t) = e^{-i(T+V)\delta t} \approx U_{\text{QFT}}e^{-iT\delta t}U_{\text{QFT}}^\dagger e^{-iV\delta t} \tag{6.11}$$

The operators $e^{-iV\delta t}$ and $e^{-iT\delta t}$ are diagonal in the position and momentum representations, respectively. A diagonal operator can be easily implemented because it amounts to adding

a phase $e^{-iV(x)\delta t}$ to each basis state $|x\rangle$. Furthermore, it is easy on a quantum computer to switch between the position and momentum representations of a wave function using the efficient quantum Fourier transform. Finally, the scheme can be easily generalized to many particles in three dimensions: a system of B particles requires $3Bn$ qubits, n for each degree of freedom.

The first quantization method can be applied to many problems. The earliest applications established that as few as 10-15 qubits would be needed for a proof-of-principle demonstration of single-particle dynamics [207] (later improved to 6-10 [21]). The method could also be used to faithfully study the chaotic dynamics of the kicked rotor model [132]. The first chemical application was the proposal of a method for the calculation of the thermal rate constant [137], see Section 5.3.

Ref. [103] investigated the applicability of the first quantization method to the simulation of chemical dynamics. The simplest approach is to consider all the nuclei and electrons explicitly, in which case the exact non-relativistic molecular Hamiltonian reads

$$H = \sum_i \frac{p_i^2}{2m_i} + \sum_{i<j} \frac{q_i q_j}{r_{ij}} \tag{6.12}$$

where $r_{ij} \equiv |\mathbf{r}_i - \mathbf{r}_j|$ is the distance between particles (nuclei and electrons) i and j, which carry charges q_i and q_j, respectively. As before, the split-operator method can be used to separate the unitaries that are diagonal in the position and momentum bases. H preserves permutational symmetry, meaning that if the initial state is properly (anti-)symmetrized, see below, it will stay so throughout the simulation.

The Born-Oppenheimer approximation (BOA) was discussed in Subsection 2.2.2 and assumes that the wave function of a molecule can be expressed as a product of the nuclear wave function and electronic wave function (parameterized by the nuclear coordinates).

As the BOA has been widely used in quantum chemistry, it might seem extravagant to explicitly simulate all the nuclei and electrons. Nevertheless, the exact simulation is, in fact, faster than using the BOA for reactions with more than approximately four atoms [103]. The reason for this is the need to evaluate the potential $V(\mathbf{r}_1, \ldots, \mathbf{r}_B)$ on the fly on the quantum computer. In the exact case, the potential is simply the pairwise Coulomb interaction; by contrast, evaluating the complicated, many-body potential energy surfaces that are supplied by the BOA is a much more daunting task, even considering that one can use nuclear time steps that are approximately a thousand times longer. That is, exact simulation minimizes arithmetic, which is the bottleneck of the quantum computation; by contrast, the bottleneck on classical computers is the prohibitive scaling of the Hilbert space size, which is alleviated by the BOA.

The first quantization approach to quantum simulation suffers from the fact that even the simplest simulations proposed in [103] might require dozens of qubits and millions of quantum gates. Nevertheless, it has advantages that would make it useful if large quantum computers are built. Most importantly, because the Coulomb interaction is pairwise, simulating a system of B particles requires $O(B^2)$ gates, a significant asymptotic improvement over the second quantized scaling of $O(M^5)$, where M is the size of the basis set.
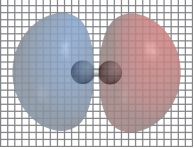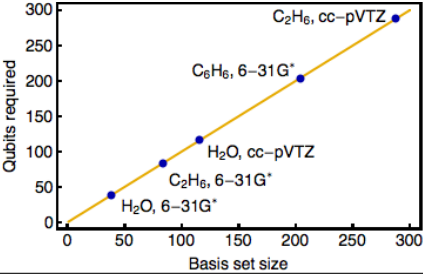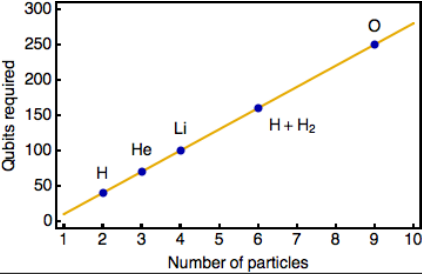
	Second quantized	First quantized
wave function encoding	Fock state in a given basis: $\|\psi\rangle = \|0100\rangle$ (shown over levels $\chi_1\ \chi_2\ \chi_3\ \chi_4$)	On a grid of 2^n sites per dimension: $\|\psi\rangle = \sum_{\mathbf{x}} a_{\mathbf{x}}\|\mathbf{x}\rangle$
Qubits required to represent the wave function	One per basis state (spin-orbitals) 	$3n$ per particle (nuclei & electrons)
Molecular Hamiltonian	$\sum_{pq} h_{pq} a_p^\dagger a_q + \dfrac{1}{2}\sum_{pqrs} h_{pqrs} a_p^\dagger a_q^\dagger a_r a_s$ Coefficients pre-computed classically	$\sum_i \dfrac{p_i^2}{2m_i} + \sum_{i<j} \dfrac{q_i q_j}{r_{ij}}$ Interaction calculated on the fly
Quantum gates required for simulation	$O(M^5)$ with number of basis states	$O(B^2)$ with number of particles
Advantages	• Compact wave function representation (requires fewer qubits) • Takes advantage of classical electronic-structure theory to improve performance • Already experimentally implemented	• Better asymptotic scaling (requires fewer gates) • Treats dynamics better • Can be used for computing reaction rates or state-to-state transition amplitudes

Table 6.1: Comparison of second and first quantization approaches to quantum simulation.

6.4   Second quantization

6.4.1   The electronic Hamiltonian in second quantization

Recapping the discussion from Chapter 2, the molecular electronic Hamiltonian in second quantized form is given by

$$H = \sum_{p,q} h_{pq} a_p^\dagger a_q + \frac{1}{2} \sum_{p,q,r,s} h_{pqrs} a_p^\dagger a_q^\dagger a_r a_s \qquad (6.13)$$

where the sum is over the single particle basis set. The annihilation $\{a_j\}$ and creation operators $\{a_j^\dagger\}$ obey the fermionic anti-commutation relations, see Table 6.2

$$[a_p,\ a_q]_+ = 0 \qquad\qquad\qquad [a_p,\ a_q^\dagger]_+ = \delta_{pq}\mathbf{1} \qquad (6.14)$$

where the anti-commutator of operators A and B is defined as $[A, B]_+ \equiv AB + BA$ and $\mathbf{1}$ is the identity operator. The annihilation (creation) operators correspond to a set of orbitals, $\{\chi_i\}$, where each orbital is a single-particle wave function composed of a spin and a spatial function, denoted σ_i and φ_i, respectively. As the Hamiltonian considered here commutes with the electron spin operators, σ_i is restricted to be one of two orthogonal functions of a spin variable $m_s = \pm 1/2$ that are denoted $\alpha(m_s)$ and $\beta(m_s)$.

Although any basis can be used, the molecular orbitals are particularly convenient for state preparation reasons so that the Hartree-Fock state can be easily prepared. Even if the Hartree-Fock calculation does not converge, or does so at an extremely slow rate, the computed integrals still parameterize the evolution; however, the basis for the creation and annihilation operators would just be the atomic orbital basis.

The integrals $\{h_{pq}\}$ and $\{h_{pqrs}\}$ appearing in Eq. (6.13) are evaluated using a known set of basis functions (the basis set) during the Hartree-Fock procedure. Ideally, the number of basis functions used would be infinite, but in practice only a finite number of basis functions are used. By selecting Gaussian single-particle basis functions, these integrals are efficiently computable.

See Chapter 2 for more information on basis sets, the molecular orbitals and the Hartree-Fock algorithm that is used to compute the transformation between them. Chapter 4 discusses the computational complexity of Hartree-Fock.

6.4.2   Representing the molecular Hamiltonian with qubits

For experimental addressability, the qubits must, in general, be distinguishable. However, the electrons of the molecular system are indistinguishable. In other words, whereas the operators $a_i^\dagger$ and a_i obey the canonical fermionic anticommutation relations, $[a_i, a_j^\dagger]_+ = \delta_{ij}$, the qubit operators that exchange $|0\rangle$ and $|1\rangle$ do not obey this relation.

The Jordan-Wigner transform is used to circumvent this issue by expressing fermionic operators in terms of the Pauli spin operators $\{\sigma^x, \sigma^y, \sigma^z, \mathbf{1}\}$ that correspond to the algebra of distinguishable spin 1/2 particles first pointed out in Ref. [98] and derived in Appendix B. This allows for a convenient implementation on a quantum computer [201, 163].

The Jordan-Wigner transform for N fermionic modes to N spins is given by

$$a_j \Leftrightarrow \mathbf{1}^{\otimes j-1} \otimes \sigma^+ \otimes \sigma^{z \otimes N-j} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}^{\otimes j-1} \otimes \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \otimes \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}^{\otimes N-j} \qquad (6.15\text{a})$$

$$a_j^\dagger \Leftrightarrow \mathbf{1}^{\otimes j-1} \otimes \sigma^- \otimes \sigma^{z \otimes N-j} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}^{\otimes j-1} \otimes \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} \otimes \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}^{\otimes N-j} \qquad (6.15\text{b})$$

where $\sigma^+ \equiv \frac{\sigma^x + i\sigma^y}{2} = |0\rangle\langle 1|$ and $\sigma^- \equiv \frac{\sigma^x - i\sigma^y}{2} = |1\rangle\langle 0|$. The qubit state $|0\ldots0\rangle$ corresponds to the vacuum state and the string of σ^z operators, preserve the commutation relations in Eq. (6.14) since σ^z and $\sigma^\pm$ anti-commute. The spin variable representation of relevant operators after the Jordan-Wigner transformation is given in Table 6.3.

6.4.3 Quantum circuits for quantum simulation

Quantum circuit primitives

Each term of Eq. (6.13) can be exponentiated after performing the Jordan-Wigner transformation to Pauli spin matrices using only the universal gate set containing the CNOT gate and arbitrary single qubit rotations. We outline the procedure for generating quantum circuits for chemical systems and summarize the results in Table 6.4. The construction of quantum circuits for general fermionic Hamiltonians is also discussed in Ref. [163, 201, 164].

To understand the exponential map of the product of Pauli spin matrices, first consider the exponential map of two σ^z operators. To create the unitary gate $\exp[-i\frac{\theta}{2}(\sigma^z \otimes \sigma^z)]$, the CNOT gate can be used to first entangle two qubits, then the R_z gate is applied, and is followed by a second CNOT gate.
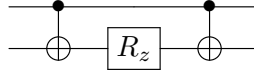
This construction can be generalized to more qubits by using additional CNOT gates. For example, the circuit for the three-body operator involving three qubits, $\exp[-i\frac{\theta}{2}(\sigma^z \otimes \sigma^z \otimes \sigma^z)]$, is simulated by the following quantum circuit
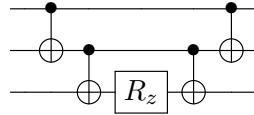
As seen from the three-qubit example above, this construction can be readily extended for n-fold products of σ^z operators.

Construction of different Pauli matrix tensor products

If one requires a different tensor product of Pauli matrices besides the product of σ^z, a change of basis can be accomplished using the appropriate unitary transformation: Hadamard transformation (denoted H) changes between σ^x basis and σ^z basis, and Y $=$

$R_x(-\pi/2) = \exp(i\sigma^x \pi/4)$ transforms from σ^z basis to σ^y basis and $\mathsf{Y}^\dagger$ from σ^y to σ^z. Circuits of this form are the basis for the construction of the molecular unitary propagator, as illustrated in Table 6.4, where the circuit representations are given.

To produce the state indicated in Eq. (5.3), the evolution of the eigenstate must be dependent on the register qubit requiring that the construction of the unitary evolution operator described earlier in this subsection be modified. The constructions listed in Table 6.4 only need to be slightly modified; since the underlying constructions rely on R_z gates, changing these rotations into controlled R_z rotations ($|1\rangle\langle 1| \otimes R_z + |0\rangle\langle 0| \otimes \mathbf{1}$) is sufficient to make the entire unitary dependent on the readout qubit's state.

In summary, we have reviewed the applications of Suzuki-Trotter formulas, discussed first quantized simulations and detailed the steps necessary for second quantized simulation. The goal is to demonstrate how quantum computers can simulate quantum systems in polynomial time. This allows the measurement scheme discussed in the previous chapter to be used to extract the energy. The last issue to consider is the initial state to be used. Before continuing on to discuss the application of quantum simulation to the hydrogen molecule, the next chapter discusses the initial state to be used in the quantum simulation.

Second quantization		
Creation operator	$a_i^\dagger|j_1,...,0_i,...,j_n\rangle$ $= \Gamma_i^{\mathbf{j}}|j_1,...,1_i,...,j_n\rangle$ $a_i^\dagger|j_1,...,1_i,...,j_n\rangle = 0$	with $\Gamma_i^{\mathbf{j}} = \prod_{k=1}^{i-1}(-1)^{j_k}$
Annihilation operator	$a_i|j_1,...,1_i,...,j_n\rangle$ $= \Gamma_i^{\mathbf{j}}|j_1,...,0_i,...,j_n\rangle$ $a_i|j_1,...,0_i,...,j_n\rangle = 0$	
Commutation relations	$\{a_i^\dagger, a_j\} = \delta_{ij}\mathbf{1}$ $\{a_i, a_j\} = 0$	

Fock space		
Basis vectors, configuration state functions	$|\mathbf{j}\rangle = |j_1, j_2, ..., j_n\rangle$ $= \prod_{p=1}^N \left(a_p^\dagger\right)^{j_p}|vac\rangle$	where $j_i = 0,1$
Inner product	$\langle \mathbf{j}|\mathbf{k}\rangle = \prod_{p=1}^n \delta_{j_p,k_p}$	
Vacuum state	$\langle vac|vac\rangle = 1$ $a_i|vac\rangle = 0$	
Single electron operator, $\hat{A}(x)$	$\hat{A}_{ij} a_i^\dagger a_j$	$\hat{A}_{ij} = \int \chi_i(x_1)\hat{A}(x_1)\chi_j(x_1)dx_1$ is a single electron operator and $\{\chi_k\}$ corresponds to $\{a_k\}$

Table 6.2: **An overview of second quantization for fermionic particles.** For quantum chemistry, the annihilation and creation operators correspond to removing (adding) an electron into a particular orbital, $\{\chi_k\}$. The anti-symmetry is enforced by the canonical commutation relations, and the N-electron wave function is expanded over the configuration state functions of the Fock space. See Chapter 2 for more information.

Description	Second Quantization	Pauli representation
Number Operator	$h_{pp}a_p^\dagger a_p$	$\frac{h_{pp}}{2}(\mathbf{1}_p - \sigma_p^z)$
Excitation Operator	$h_{pq}a_p^\dagger a_q + h_{qp}a_q^\dagger a_p$	$\frac{1}{2}\left(\bigotimes_{k=q+1}^{p-1}\sigma_k^z\right)\left(\begin{array}{c}\Re\{h_{pq}\}(\sigma_q^x\sigma_p^x + \sigma_q^y\sigma_p^y) \\ +\Im\{h_{pq}\}(\sigma_q^y\sigma_p^x - \sigma_q^x\sigma_p^y)\end{array}\right)$
Coulomb Operator	$h_{pqqp}a_p^\dagger a_q^\dagger a_q a_p$	$\frac{h_{pqqp}}{4}\left(\mathbf{1} - \sigma_p^z - \sigma_q^z + \sigma_p^z\sigma_q^z\right)$
Number with[a] Excitation Operator	$h_{pqqr}a_p^\dagger a_q^\dagger a_q a_r + h_{rqqp}a_r^\dagger a_q^\dagger a_q a_p$	$\left(\bigotimes_{k=r+1}^{p-1}\sigma_k^z\right)\left[\left(\begin{array}{c}\Re\{h_{pqqr}\}(\sigma_r^x\sigma_p^x + \sigma_r^y\sigma_p^y) \\ +\Im\{h_{pqqr}\}(\sigma_r^y\sigma_p^x - \sigma_r^x\sigma_p^y)\end{array}\right) - \sigma_q^z\left(\begin{array}{c}\Re\{h_{pqqr}\}(\sigma_r^x\sigma_p^x + \sigma_r^y\sigma_p^y) \\ +\Im\{h_{pqqr}\}(\sigma_r^y\sigma_p^x - \sigma_r^x\sigma_p^y)\end{array}\right)\right]$
Double Excitation Operator	$h_{pqrs}a_p^\dagger a_q^\dagger a_r a_s + h_{srqp}a_s^\dagger a_r^\dagger a_q a_p$	$\left(\bigotimes_{k=s+1}^{r-1}\sigma_k^z\right)\left(\bigotimes_{k=q+1}^{p-1}\sigma_k^z\right)\left[\frac{\Re\{h_{pqrs}\}}{8}\left(\begin{array}{c}\sigma_s^x\sigma_r^x\sigma_q^x\sigma_p^x - \sigma_s^x\sigma_r^x\sigma_q^y\sigma_p^y \\ +\sigma_s^x\sigma_r^y\sigma_q^x\sigma_p^y + \sigma_s^y\sigma_r^x\sigma_q^x\sigma_p^y \\ +\sigma_s^y\sigma_r^x\sigma_q^y\sigma_p^x - \sigma_s^y\sigma_r^y\sigma_q^x\sigma_p^x \\ +\sigma_s^x\sigma_r^y\sigma_q^y\sigma_p^x + \sigma_s^y\sigma_r^y\sigma_q^y\sigma_p^y\end{array}\right) + \frac{\Im\{h_{pqrs}\}}{8}\left(\begin{array}{c}\sigma_s^y\sigma_r^x\sigma_q^x\sigma_p^x + \sigma_s^x\sigma_r^y\sigma_q^x\sigma_p^x \\ +\sigma_s^x\sigma_r^x\sigma_q^y\sigma_p^x - \sigma_s^x\sigma_r^x\sigma_q^x\sigma_p^y \\ -\sigma_s^y\sigma_r^y\sigma_q^y\sigma_p^x + \sigma_s^x\sigma_r^y\sigma_q^y\sigma_p^y \\ +\sigma_s^y\sigma_r^x\sigma_q^y\sigma_p^y + \sigma_s^y\sigma_r^y\sigma_q^x\sigma_p^y\end{array}\right)\right]$

Table 6.3: **The spin variable representation of second quantized Hermitian operators** after performing the Jordan-Wigner transformation to yield tensor products of Pauli sigma matrices (spin variables) that have the proper anti-commutation relations. See Eqs. (6.15b) and (6.15a) for the form of the transformation used. The subscripts p, q, r, s label the qubit as well as the molecular spin orbital that corresponds to that qubit. It is assumed that $p > q > r > s$ for all cases listed. The pre-factors h_{ij} and h_{ijkl} are one- and two-electron integrals given in Eqs. (2.6) and (2.7). In our algorithm, these are calculated using a classical computer. Rarely is this the limiting step since the basis sets are typically chosen for ease of integration; moreover, their computation time is low enough that computing them on the fly is often feasible for large computations. In this table, we provide for the case that the one- and two-electron integrals are complex. The integrals, h_{pqqp} and h_{pqpq}, are referred to as the Coulomb and exchange integrals, respectively. The exchange operator has the same representation (with opposite sign) as the Coulomb integral due to the commutation relations. The circuit representation of the exponentiation of each of these terms is given in Table 6.4.

---

[a] The spin variable representation of this operator depends if q is an orbital in the range p to r or if it is outside this range.

Second quantized operators		Circuit
Number operator	$h_{pp}a_p^\dagger a_p$	$\mathsf{T}(\theta)$
Excitation operator	$h_{pq}(a_p^\dagger a_q + a_q^\dagger a_p)$	
Coulomb and exchange operators	$h_{pqqp}a_p^\dagger a_q^\dagger a_q a_p$	
Number-excitation[1] operator	h_{pqqr} $\;\;(a_p^\dagger a_q^\dagger a_q a_r + a_r^\dagger a_q^\dagger a_q a_p)$	where $\mathsf{M}=\{\mathsf{H},\mathsf{Y}\}$
Double excitation operator	h_{pqrs} $\;\;(a_p^\dagger a_q^\dagger a_r a_s + a_s^\dagger a_r^\dagger a_q a_p)$	where $(\mathsf{M}_1,\mathsf{M}_2,\mathsf{M}_3,\mathsf{M}_4)=$ $\{(\mathsf{H},\mathsf{H},\mathsf{H},\mathsf{H}),(\mathsf{Y},\mathsf{Y},\mathsf{Y},\mathsf{Y}),$ $(\mathsf{H},\mathsf{Y},\mathsf{H},\mathsf{Y}),(\mathsf{Y},\mathsf{H},\mathsf{Y},\mathsf{H}),$ $(\mathsf{Y},\mathsf{Y},\mathsf{H},\mathsf{H}),(\mathsf{H},\mathsf{H},\mathsf{Y},\mathsf{H}),$ $(\mathsf{Y},\mathsf{H},\mathsf{H},\mathsf{Y}),(\mathsf{H},\mathsf{Y},\mathsf{Y},\mathsf{H})\}$
Notation:		

Table 6.4: **The quantum circuits corresponding to evolution of second quantized operators**. Here, p, q, r, and s are orbital indices corresponding to qubits such that qubit state $|1\rangle$ indicates an occupied orbital and $|0\rangle$ indicates unoccupied. It is assumed that the orbital indices satisfy $p > q > r > s$. These circuits were found by performing the Jordan-Wigner transformation given in Eqs. (6.15b) and (6.15a) and then exponentiating the obtained Pauli spin variables as exemplified in Subsection 6.4.3. In each circuit, $\theta = \theta(h)$ where h is the integral preceding the operator. Gate $\mathsf{T}(\theta)$ is defined by $\mathsf{T}|0\rangle = |0\rangle$ and $\mathsf{T}|1\rangle = \exp(-i\theta)|1\rangle$, $\mathsf{G}(\phi)$ is the global phase gate given by $\exp(-i\phi)\mathbf{1}$, the change-of-basis gate Y is defined as $R_x(-\pi/2)$, and H refers to the Hadamard gate. For the number-excitation operator, both $\mathsf{M} = \mathsf{Y}$ and $\mathsf{M} = \mathsf{H}$ must be implemented in succession. Similarly, for the double excitation operator each of the 8 quadruplets must be implemented in succession. The global phase gate is included due to the phase-estimation procedure, see Chapter 5. Phase estimation requires controlled versions of these operators which is accomplished by changing all gates with θ-dependence into controlled gates dependent on register qubits.

Chapter 7

Playing Merlin: ideas and difficulties

> Proper preparation prevents pitifully poor performance.

–Military adage

In Chapter 6, techniques for implementing quantum dynamics on a quantum computer were presented. The focus of the present chapter is on preparing initial states. In this chapter, I will discuss the preparation of integrable initial states, our algorithm for preparing thermal states, and techniques for ground states. The chapter concludes with a discussion of the limitations on arbitrary state preparation.

7.1 General methods

Zalka's original 1998 paper [236] on quantum simulation contained a general state-preparation scheme, rediscovered several times [82, 105, 114] and improved in Ref. [197]. Zalka's scheme builds the quantum state one qubit at a time by performing a rotation (dependent on the previous qubits) that redistributes the wave function amplitude as desired. Gaussian wavepackets or molecular orbitals can be constructed efficiently [225, 114]. An (anti-)symmetrization algorithm for first quantized algorithms was first presented in [1] and extensions of this algorithm can be used to prepare Slater determinants necessary for chemical simulation [225]. Furthermore, in [225] a procedure for translating states prepared in second quantization language into first quantized wave functions and vice versa was given. Our extension of Zalka's method to prepare thermal states [235] is summarized in Section 7.3.

7.2 Ground states

A large part of quantum chemistry is concerned with the calculation of ground-state properties of molecules, making it desirable to prepare such states on a quantum

computer. In Chapter 5, PEA was presented which is used to measure a quantum state in the eigenbasis of a Hermitian operator. This suggests a method for preparing a ground state: measuring in the eigenbasis of the Hamiltonian will project a state $|\psi\rangle$ to the ground state $|g\rangle$ with probability $|\langle\psi|g\rangle|^2$.

The problem, therefore, is to prepare a state close to the ground state, from which we can project the ground-state component. Choosing a random state $|\psi_{\text{rand}}\rangle$ is bound to fail, because the overlap is expected to be exponentially small in the number of qubits n: $\langle\psi_{\text{rand}}|g\rangle \sim 2^{-n}$. This means that one would have to repeat the PEA exponentially many times before chancing upon the ground state.

First, consider the Hartree-Fock wave function as an approximation to the FCI wave function. This is the output of an classical algorithm returning, usually in polynomial time, a computational basis state where the qubits which correspond to occupied molecular orbitals are in the state $|1\rangle$ with the remaining qubits in state $|0\rangle$. This state can easily be prepared from the state $|00...0\rangle$ with the σ^x gate applied to the occupied orbitals.

If the molecules are being considered at equilibrium geometries, the Hartree-Fock guess may have sufficient overlap for the phase estimation algorithm to collapse the state to the exact ground state with sufficiently high probability. The overlap can be further improved by choosing a more sophisticated classical approximation method such as a multi-configuration self-consistent field (MCSCF) wave function [224, 223]. In Ref. [218], the effects of initial states for the phase estimated quantum simulation CH_2 molecules were studied for a variety of geometries and eigenstates using initial guesses obtained via multi-configuration approaches.

7.2.1 Adiabatic state preparation

Even if the classical guess for the initial state is poor, the initial state can be improved using adiabatic state preparation (ASP) [72, 231, 15]. Suppose the classical guess is the Hartree-Fock state, $|\psi_0^{HF}\rangle$, then using knowledge of Hamiltonians H_{HF} and H_{FCI}, a state close to the actual ground state is recovered via adiabatic evolution. Consider a smooth one-parameter family of adiabatic path Hamiltonians

$$H(s) = (1-s)H_{HF} + sH_{FCI} \tag{7.1}$$

for monotonic $s \in [0,1]$. This was the adiabatic path originally suggested in Ref. [15]. Other paths may be used as in Ref. [170], but in this chapter we restrict our attention to evolution of the form in Eq. (7.1).

Let the instantaneous energies of $H(s)$ be given by the sequence

$$E_0(s) < E_1(s) \leq \cdots \leq E_{N-1}(s) \tag{7.2}$$

then the adiabatic state preparation procedure is efficient when the total run time, T, satisfies the following

$$T \gg \min_{0 \leq s \leq 1} (E_1(s) - E_0(s))^{-2} \tag{7.3}$$

according to the adiabatic approximation [146]. It should be noted that the conditions for the adiabatic approximation are not always satisfied and have recently been discussed in the literature: [11, 216, 215, 33].

If the adiabatic evolution induced transitions into higher energy states, the unnormalized state of the system is $|\psi_0^{ASP}\rangle = |\psi_0^{FCI}\rangle + \lambda|k\rangle$, where $|k\rangle$ is orthogonal to $|\psi_0^{FCI}\rangle$. While the error in the wave function is linear in $\lambda < 1$, the overestimate of the energy in the expectation value $\langle H_0^{FCI}\rangle$ is only quadratic. So long as this error is small, the phase estimation algorithm can then be applied as the system will collapse to $|\phi_0^{FCI}\rangle$ with probability $|\langle\phi_0^{FCI}|\psi_0^{ASP}\rangle|^2$.

Modified versions of the adiabatic procedure exist using dephasing in the energy basis to achieve faster evolution [32, 230]. The key idea is to perform phase estimation and ignore the measurement outcome. By frequently performing PEA, the state will be re-projected into the ground state with high probability. This idea was initially presented in Ref. [5] and the basic principal is the quantum Zeno effect first discussed in Ref. [148].

Again, consider the minimal basis hydrogen molecule as an example system. Here adiabatic state preparation has been applied at various bond by others both numerically [15] and experimentally [70] using NMR technology. As the bond length increases, the Hartree-Fock state has decreasing overlap with the exact state, reaching 0.5 at large separations. The ASP method allowed the preparation of the FCI ground state with a high fidelity in both theory and in experiment.

Procedures similar to ASP have been proposed to study low-energy states of some toy models in physics [161] and superconductivity [231]. In addition, it is also possible to encode a thermal state into the ground state of a Hamiltonian [199], offering a way to prepare a thermal state, a problem further discussed in the following section.

7.3   Thermal state preparation

Thermal states are significant because they can be used to solve many problems, ranging from statistical mechanics to the calculation of thermal rate constants. Classical algorithms typically rely on Markov chain Monte Carlo methods, which sample from the Gibbs density matrix, $\rho = e^{-\beta H}/Z$, where Z is the partition function. The challenge is that it is generally impossible to sample from the eigenstates $|e_k\rangle$ of a certain Hamiltonian H if they are not pre-determined (which is often more challenging).

Zalka's state preparation algorithm [236] is applicable to preparing the coherent encoding of thermal states (CETS) $|\psi_{\text{th}}\rangle$

$$|\psi_{\text{th}}\rangle = \sum_k \sqrt{e^{-\beta E_k}/Z}|e_k\rangle|e_k\rangle \tag{7.4}$$

which is equivalent to the Gibbs density matrix, $\rho_{\text{th}} = \sum_k e^{-\beta E_k}/Z|e_k\rangle\langle e_k|$, if one register is traced out. If the eigenstates and eigenvalues are known, it is possible to construct the CETS directly. Combining ideas from belief-propagation [147] and quantum amplitude amplification [104], we were able to construct the CETS of classical Hamiltonians with a quadratic quantum speedup [235].

Other methods can be envisioned. For example, assuming the PEA can be efficiently implemented, one can prepare the thermal state of any classical or quantum Hamiltonian from a Markov chain constructed by repeating an open system evolution [213]. A

limitation of this approach is that the Metropolis step can make too many transitions between states of very different energies, sometimes leading to a slow convergence rate of the resulting Markov chain. This issue was addressed by building up the Markov chain by applying random local unitary operations [212]. The resulting operation is a Metropolis-type sampling for quantum states; although the underlying Markov chain is classical in nature, performing it on a quantum computer provides the benefit of being able to use the PEA without explicitly solving the eigenvalue equations. However, quantum computers can implement Markov chains corresponding to thermal states of classical Hamiltonians with a quadratic speed-up [210, 200, 230, 175, 176, 46].

Lastly, a thermal state can be prepared by modeling the physical interaction with a heat bath [236, 213]. However, the computational cost of these methods is not well understood.

7.3.1 QMA-hardness and future prospects

Unfortunately, the procedures for ground- and thermal-state preparation outlined above are not fully scalable to larger systems. A simple way to see this is to imagine a system composed of N identical, non-interacting molecules. Even if one molecule can be prepared with a ground-state overlap of $1 - \epsilon$ by any method, the fidelity of the N-molecule state will be exponentially small, $(1 - \epsilon)^N$. This argument was explained in the Nobel lecture of W. Kohn [118]. ASP would fail when the energy gap became so small that the Hamiltonian would have to be varied exponentially slowly.

More generally, there are broad classes of Hamiltonians for which finding the ground-state energy (and therefore also a thermal state) is known to be QMA-hard, see Chapter 4 for examples. Nevertheless, the scaling of the ground- and thermal-state energy problems for chemical systems on a quantum computer is an open question. It is possible that algorithms can be found that are not efficient for all QMA-hard Hamiltonians, but, nevertheless succeed for chemical problems.

In fact, there are physical systems such as spin glasses in nature which may never settle into their ground states. However, a host of realistic systems (e.g. insulators and molecular systems) can, on physical grounds, be expected to retain a large energy gap and should thus be amenable to quantum simulation algorithms which rely on adiabatic state preparation.

Chapter 8

Quantum simulation of hydrogen molecule: theory and experiment

> Observation always involves theory.

–Edwin Hubble

> There is something irreversible about acquiring knowledge; and the simulation of the search for it differs in a most profound way from reality.

–J. Robert Oppenheimer, referring to the distinction between theory and practice

The scalable second quantized algorithm for simulating hydrogen is used to illustrate the ideas indicated in the previous two chapters. Following that, an experimental implementation of a simplified version of the second quantized algorithm is presented.

We take the standard approach to quantum-chemical calculations by solving an approximate Hamiltonian created by using the Born-Oppenheimer approximation (where the electronic Hamiltonian is parameterized by the nuclear configuration) and choosing a suitable truncated basis set in which to describe the non-relativistic electronic system. Typical sets consist of a finite number of single-electron atomic orbitals, which are combined to form antisymmetric multi-electron states (configurations), see Chapter 2 for more background and terminology. Calculating the eigenvalues of the electronic Hamiltonian using all configurations gives the exact energy in the basis set and is referred to as full configuration interaction (FCI). For N orbitals and m electrons there are $\binom{N}{m} \approx N^m/m!$ ways to allocate the electrons among the orbitals. This exponential growth is the handicap of FCI calculations on classical computers.

8.1   Quantum simulation of the hydrogen molecule

To illustrate the algorithmic details of a scalable simulation of quantum systems, the hydrogen molecule in a minimal basis is used as an instructive example. The minimal

Integrals	Value (a.u.)
$h_{11} = h_{22}$	-1.252477
$h_{33} = h_{44}$	-0.475934
$h_{1221} = h_{2112}$	0.674493
$h_{3443} = h_{4334}$	0.697397
$h_{1331} = h_{1441} = h_{2332} = h_{4224}$ $= h_{3113} = h_{3223} = h_{4114} = h_{4224}$	0.663472
$h_{1313} = h_{2424} = h_{3241} = h_{1423} = h_{1243}$	0.181287

Table 8.1: The one-electron and two-electron integrals defined in Eqs. (2.6) and (2.7) are evaluated using the molecular spatial orbitals obtained from a restricted Hartree-Fock calculation at an internuclear distance of 1.401000 atomic units ($7.414 \cdot 10^{-11}$ m) [152].

basis is the minimum number of spatial-functions needed to describe the system and in the case of H_2, one spatial-function is needed per atom denoted φ_{H1} and φ_{H2}. In this simple case, the Hartree-Fock procedure is not necessary as the molecular spatial-orbitals are determined by symmetry and are given by $\varphi_g = \varphi_{H1} + \varphi_{H2}$ and $\varphi_u = \varphi_{H1} - \varphi_{H2}$. These two spatial functions correspond to four spin-orbitals once the spin-functions α and β corresponding to spin up and down respectively have been included.

$$|\chi_1\rangle = |\varphi_g\rangle|\alpha\rangle, \qquad |\chi_2\rangle = |\varphi_g\rangle|\beta\rangle, \qquad |\chi_3\rangle = |\varphi_u\rangle|\alpha\rangle, \qquad |\chi_4\rangle = |\varphi_u\rangle|\beta\rangle \qquad (8.1)$$

The form of the spatial function is determined by the basis set used. The STO-3G basis is a common minimal basis that approximates a single electron spatial Slater type orbitals (STO), with a contraction of three real Gaussian functions, see Subsection 2.2.3. Using this orbital basis, the spatial integrals of the Hamiltonian were evaluated for H_2 at bond distance 1.401000 atomic units ($7.414 \cdot 10^{-11}$ m), see Table 8.1. Considering H from Eq. (6.13) as $H^{(1)} + H^{(2)}$ we have

$$H^{(1)} = h_{11}a_1^\dagger a_1 + h_{22}a_2^\dagger a_2 + h_{33}a_3^\dagger a_3 + h_{44}a_4^\dagger a_4 \qquad (8.2)$$

The following circuit applies the single-electron propagator for a time t
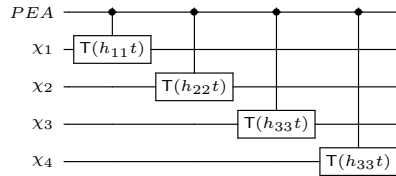
The gate T is defined as

$$\boxed{\mathsf{T}(\theta)} = \begin{bmatrix} 1 & \\ & e^{-i\theta} \end{bmatrix} \qquad (8.3)$$

The two-electron Hamiltonian can also be expanded. As electrons are indistinguishable

$$
\begin{aligned}
h_{pqrs} &= \int \mathrm{d}\mathbf{x}_1 \, \mathrm{d}\mathbf{x}_2 \frac{\chi_p(\mathbf{x}_1)\chi_q(\mathbf{x}_2)\chi_r(\mathbf{x}_2)\chi_s(\mathbf{x}_1)}{r_{12}} \\
&= \int \mathrm{d}\mathbf{x}_2 \, \mathrm{d}\mathbf{x}_1 \frac{\chi_p(\mathbf{x}_2)\chi_q(\mathbf{x}_1)\chi_r(\mathbf{x}_1)\chi_s(\mathbf{x}_2)}{r_{12}} = h_{qpsr}
\end{aligned}
$$

and $a_p^\dagger a_q^\dagger a_r a_s = a_q^\dagger a_p^\dagger a_s a_r$, the two-electron Hamiltonian can be simplified as

$$
\begin{aligned}
H^{(2)} &= h_{1221} a_1^\dagger a_2^\dagger a_2 a_1 + h_{3443} a_3^\dagger a_4^\dagger a_4 a_3 + h_{1441} a_1^\dagger a_4^\dagger a_4 a_1 + h_{2332} a_2^\dagger a_3^\dagger a_3 a_2 \\
&\quad + (h_{1331} - h_{1313}) \, a_1^\dagger a_3^\dagger a_3 a_1 + (h_{2442} - h_{2424}) \, a_2^\dagger a_4^\dagger a_4 a_2 \\
&\quad + \Re(h_{1423})(a_1^\dagger a_4^\dagger a_2 a_3 + a_3^\dagger a_2^\dagger a_4 a_1) + \Re(h_{1243})(a_1^\dagger a_2^\dagger a_4 a_3 + a_3^\dagger a_4^\dagger a_2 a_1) \\
&\quad + \Im(h_{1423})(a_1^\dagger a_4^\dagger a_2 a_3 - a_3^\dagger a_2^\dagger a_4 a_1) + \Im(h_{1243})(a_1^\dagger a_2^\dagger a_4 a_3 - a_3^\dagger a_4^\dagger a_2 a_1)
\end{aligned}
$$

The first six terms

$$
\begin{aligned}
& h_{1221} a_1^\dagger a_2^\dagger a_2 a_1 + h_{3443} a_3^\dagger a_4^\dagger a_4 a_3 + h_{1441} a_1^\dagger a_4^\dagger a_4 a_1 \\
+ \quad & h_{2332} a_2^\dagger a_3^\dagger a_3 a_2 + (h_{1331} - h_{1313}) \, a_1^\dagger a_3^\dagger a_3 a_1 + (h_{2442} - h_{2424}) \, a_2^\dagger a_4^\dagger a_4 a_2
\end{aligned}
$$

can be simulated using the system Hamiltonian that employs only commuting two-local terms described in Section 6.2. Notice after the Jordan-Wigner transform of the relevant operator we have

$$
\begin{aligned}
\sum_{p<q}(h_{pqqp} - h_{pqpq}\delta_{\sigma_p\sigma_q}) a_p^\dagger a_q^\dagger a_q a_p &= \left(\frac{1}{4}\sum_{p<q}(h_{pqqp} - h_{pqpq}\delta_{\sigma_p\sigma_q}) \right) \mathbf{1} \\
&\quad - \frac{1}{4}\sum_q \left(\sum_{p\neq q}(h_{pqqp} - h_{pqpq}\delta_{\sigma_p\sigma_q}) \right) \sigma_q^z \\
&\quad + \sum_{p<q} \frac{(h_{pqqp} - h_{pqpq}\delta_{\sigma_p\sigma_q})}{4} \sigma_p^z \sigma_q^z \qquad (8.4)
\end{aligned}
$$

The factor of $1/2$ is accounted for because the indistinguishably of the electrons reduces the summation in Eq. (6.13).

Following Eq. (8.4), let $\Theta \equiv (1/4)\sum_{p<q}(h_{pqqp}-h_{pqpq}\delta_{\sigma_p\sigma_q})$ and let $\theta_p \equiv \sum_{q:p\neq q}(h_{pqqp}-h_{pqpq}\delta_{\sigma_p\sigma_q})$. The following circuit illustrates the one and two-local interactions required to implement Eq. (8.4)
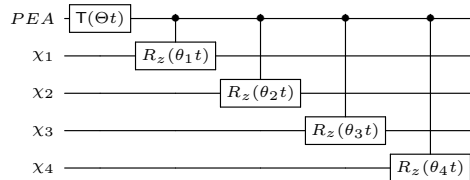
Defining $\eta_{pq} \equiv \frac{1}{4}(h_{pqqp} - h_{pqpq}\delta_{\sigma_p \sigma_q})$, the three local interactions can be depicted as
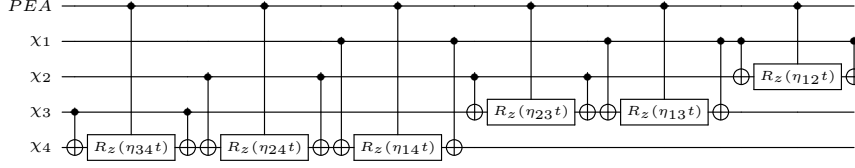
Each term commutes, and thus can be realized in any order.

The remaining terms are strictly real, leaving

$$h_{1423}(a_1^\dagger a_4^\dagger a_2 a_3 + a_3^\dagger a_2^\dagger a_4 a_1) + h_{1243}(a_1^\dagger a_2^\dagger a_4 a_3 + a_3^\dagger a_4^\dagger a_2 a_1)$$

Consider the general term $a_p^\dagger a_q^\dagger a_r a_s + a_s^\dagger a_r^\dagger a_q a_p$. Due to the anti-commutation rules, all sets of operators corresponding to a set of four distinct spin-orbitals, (p, q, r, s), are simulated using the same circuit. This is due to the fact that the Jordan-Wigner of each operator generates same set of operators (namely, the eight combinations involving an even number of σ^x and σ^y operators). However, depending on if σ^+ or σ^- is used each term of spin operators will have a different sign. If we define

$$h^{(1)} \;=\; (h_{pqrs}\delta_{\sigma_p \sigma_s}\delta_{\sigma_q \sigma_r} - h_{qprs}\delta_{\sigma_p \sigma_r}\delta_{\sigma_q \sigma_s}) \tag{8.5}$$

$$h^{(2)} \;=\; (h_{psqr}\delta_{\sigma_p \sigma_r}\delta_{\sigma_q \sigma_s} - h_{spqr}\delta_{\sigma_p \sigma_q}\delta_{\sigma_r \sigma_s}) \tag{8.6}$$

$$h^{(3)} \;=\; (h_{prsq}\delta_{\sigma_p \sigma_q}\delta_{\sigma_r \sigma_s} - h_{prqs}\delta_{\sigma_p \sigma_s}\delta_{\sigma_q \sigma_r}) \tag{8.7}$$

then

$$\frac{1}{2}\sum h_{pqrs}(a_p^\dagger a_q^\dagger a_r a_s + a_s^\dagger a_r^\dagger a_q a_p) \;=\; \frac{1}{8}\sum \left(\bigotimes_{k=p+1}^{q-1}\bigotimes_{k=r+1}^{s-1}\sigma_k^z\right) \tag{8.8}$$

$$\begin{pmatrix} (\sigma_p^x\sigma_q^x\sigma_r^x\sigma_s^x + \sigma_p^y\sigma_q^y\sigma_r^y\sigma_s^y)(-h^{(1)} - h^{(2)} + h^{(3)}) \\ +(\sigma_p^x\sigma_q^x\sigma_r^y\sigma_s^y + \sigma_p^y\sigma_q^y\sigma_r^x\sigma_s^x)(+h^{(1)} - h^{(2)} + h^{(3)}) \\ +(\sigma_p^y\sigma_q^x\sigma_r^y\sigma_s^x + \sigma_p^x\sigma_q^y\sigma_r^x\sigma_s^y)(-h^{(1)} - h^{(2)} - h^{(3)}) \\ +(\sigma_p^y\sigma_q^x\sigma_r^x\sigma_s^y + \sigma_p^x\sigma_q^y\sigma_r^y\sigma_s^x)(-h^{(1)} + h^{(2)} + h^{(3)}) \end{pmatrix}$$

Considering the application to hydrogen molecule, observe that $h^{(1)} = -h^{(2)}$ and $h^{(3)} = 0$, indicating that only the terms $\{\sigma^x\sigma^x\sigma^y\sigma^y, \sigma^y\sigma^y\sigma^x\sigma^x, \sigma^y\sigma^x\sigma^x\sigma^y, \sigma^x\sigma^y\sigma^y\sigma^x\}$ must be considered. The resulting quantum circuit is illustrated in Table A1.

To assess the Trotter error, we simulated this circuit using a classical computer using the first-order Trotter decomposition. The pseudo-code for the H_2 simulation is given in Appendix A and the results are summarized in Fig. 3. Although the gates increase with the Trotter number, reducing the Trotter error of the dynamics is only practical if the measurement is precise enough to detect such errors. Thus, in practice, there is a balance between the Trotter number selected and the number of bits to be obtained by the measurement procedure. Finally, as noted in Section 6.2.1, higher-order Trotter decompositions also allow for more precise simulation with larger time steps.

8.2   Resource count for hydrogen molecule

In order to illustrate this algorithm, we performed numerical simulations for H_2 in the same minimal basis (STO-3G) employed in our experiment. Unlike our experimental mapping, the logical states of each register qubit are now identified with the fermionic occupancy of the four single-electron spin-orbitals (i.e., $|1\rangle$ = occupied, $|0\rangle$ = unoccupied). Therefore, the calculation requires a total of five qubits, taking into consideration the single control qubit required for the IPEA. If quantum error correction is needed, the number of qubits will increase according to the scheme used [155]. Fig. 8.1 shows the error in the ground state energy as a function of the Trotter step. The ground state energies of the approximate unitary propagators were obtained via direct diagonalization on a classical computer. A precision of $\pm 10^{-4} E_h$ is achieved at a Trotter number of 6, which corresponds to 522 gates. Note that this gate count is to construct U^1 and includes both one- and two-qubit operations. This estimate does not take into consideration error correction for the qubits and it uses a continuous set of gates. In the path to large scale implementations, both will be serious considerations and will increase the complexity of the algorithm and the number of qubits necessary [155, 57]. The unitary matrix must be raised to various powers to perform phase estimation. If one tries to maintain a fixed accuracy of 13 bits, approximately 8.5×10^6 gates must be used for the IPEA estimation procedure. Note that this can be achieved by repeating the 522 gates required for U many times. This cost does not include the resources associated with preparing an approximate eigenstate, see Chapter 7.

Although the estimates just given exceed the capabilities of current quantum computers, these resource requirements grow only *polynomially* with the size of the system. Consequently, for large enough chemical systems, quantum computers with around 100 qubits are predicted to outperform classical computational devices for the first-principles calculation of chemical properties.

8.3   Experimental implementation of molecular hydrogen simulation

In the remainder of this chapter, we discuss the first experimental realization of a quantum algorithm for quantum chemical calculations. The experimental setup is a linear-optical quantum computing architecture and we simulate the hydrogen molecule in a minimal basis. We perform a key algorithmic step—the iterative phase estimation algorithm [137, 15, 40, 57]—in full, achieving a high level of precision and robustness to error. We implement other algorithmic steps with assistance from a classical computer.

As described in Subsection 8.4.1, the Hamiltonian for the H_2 molecule in a minimal basis is block diagonalized in the symmetry adapted basis, with 2×2 sub-matrices $H^{(1,6)}$ and $H^{(3,4)}$. We map the configurations spanning each sub-space to the qubit computational basis. Because the subspaces are two-dimensional, one qubit suffices to represent the wave function. The corresponding time-evolution operators, $U^{(p,q)} = e^{-iH^{(p,q)}t/\hbar}$, where $(p,q) = (1,6)$ or $(3,4)$, are therefore one-qubit operators. Finding the eigenvalues of each separately, using a phase estimation algorithm, amounts to performing FCI. For the purpose of our
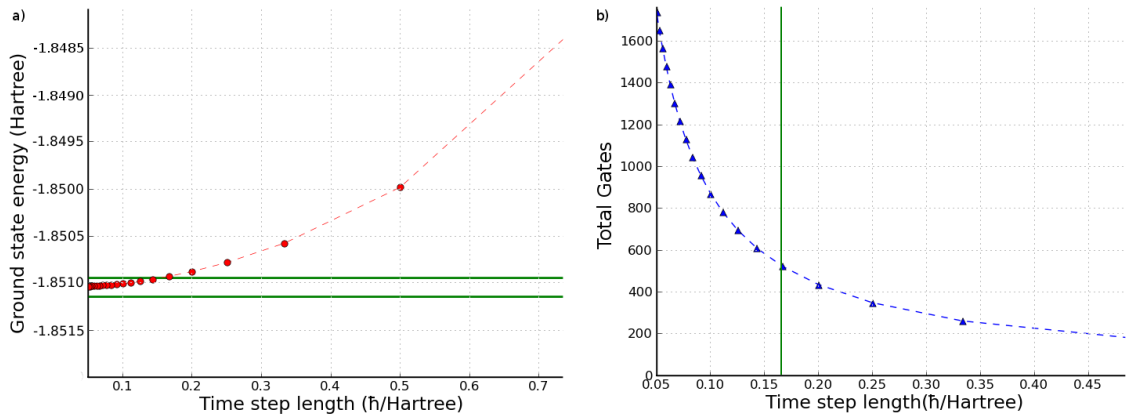
Figure 8.1: **Trotter error analysis and resource count for hydrogen molecule using a scalable quantum simulation algorithm**. **(a)** Plot of ground state energy of hydrogen molecule as a function of the length of the time step. As the time step length decreases, the accuracy of the approximation increases in accordance with Eq. (6.1). The total time of propagation, t, was unity and this time was split into time steps, dt. The circles are at integer values of the Trotter number, $T_n \equiv t/dt$. Green horizontal lines indicate the bounds for $\pm 10^{-4} E_h$ precision. **(b)** Gates for a single construction of the approximate unitary as a function of time step. As the time step decreases, more gates must be used to construct the propagator. The triangles indicate integer values of the Trotter number and the green vertical line corresponds to the same threshold from graph (**a**). Perfect gate operations are assumed.

demonstration, we encode exact eigenstates, obtained via a preliminary calculation on a classical computer. In Section 8.5 we show that the experimental algorithm is in fact robust to imperfect eigenstate encoding.

We implement the IPEA in a photonic architecture, encoding qubits into the polarization of single photons generated in pairs through the process of spontaneous parametric downconversion. An experimental schematic is shown in Fig. 8.2c. In our encoding, the logical state $|0\rangle$ ($|1\rangle$) corresponds to the horizontal, $|H\rangle$ (vertical, $|V\rangle$), polarization. Single qubit gates can be easily implemented using birefringant wave plates to perform arbitrary polarization rotations. Multi-qubit gates are possible by combining linear-optical elements (phase shifters and beam splitters) with projective measurement, as originally demonstrated in 2001 [159, 115]. Such gates are high-quality, well-characterized, and have several in-principle paths to scalable optical quantum computing [119]. The key device of our experiment is a photonic two-qubit controlled-unitary quantum logic gate, recently developed in our laboratory [127]. This device exploits multiple photonic degrees of freedom to extend the range of possible operations. We note that our implementation of the IPEA is the first, in any physical system, to use quantum logic gates which can generate the non-classical correlations required for efficient quantum computing. We acknowledge a previous implementation in a liquid-state ensemble NMR architecture [232], which is arguably an in-principle non-scalable approach [35]. Another implementation in ions [47] was semi-classical in the sense that it combined single-qubit measurement and feedback in place of entangling gates. We now provide a simple conceptual explanation of our experimental method.

We implement the iterative phase estimation algorithm, see Chapter 5) (IPEA), which minimizes the number of qubits and quantum logic gates required. Fig. 8.2a shows the IPEA at iteration k. In the case where the phase to be calculated has an exact expansion to m bits, a perfect implementation of m iterations of the algorithm (that is, without experimental error) will deterministically extract the exact phase. This is not generally the case and it leads to a non-zero probability of intrinsic algorithmic error: Let m bits of a binary expansion of ϕ be $\tilde{\phi} = 0.j_1 j_2...j_m$, such that $\phi = \tilde{\phi} + \delta 2^{-m}$ where δ is a remainder $0 \le \delta < 1$. A perfect implementation of the algorithm achieves an precision of $\pm 2^{-m}$ with error probability $\epsilon \le 1 - 8/\pi^2 \approx 0.19$, which is independent of m (the bound is met for $\delta = 0.5$), see Section 5.2 for derivation.

In the case of perfect algorithm implementation, this error can always be eliminated by simply repeating each IPEA iteration multiple (n) times, yielding n possible values for the corresponding bit. A majority vote is performed by selecting the mode of the set of these samples—that is, choosing the result that is observed most often, see Section 8.4.1. This technique only requires repeating the algorithm a fixed number of times which does not affect its overall efficiency.

Each iteration of our IPEA implementation requires two qubits, and therefore two photons, to be injected into the optical circuit shown in Fig. 8.2c; one in each optical input mode. We call these photons the 'control' and 'register', after the qubits that they encode (Fig. 8.2a). Each optical mode then passes through a polarizing beam splitter set to prepare (transmit) the fiducial initial state $|H\rangle$. The required IPEA input states are then prepared using wave plates: the register polarization is rotated into a state that represents (encodes)
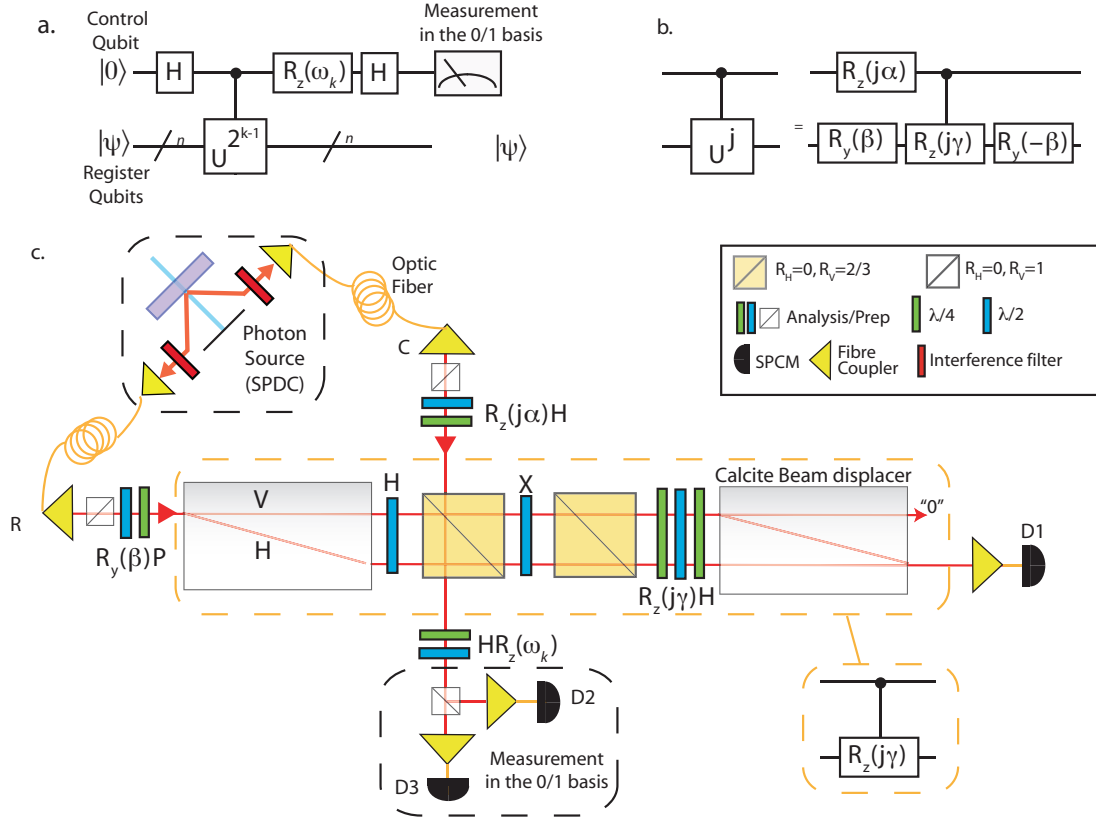
Figure 8.2: **The quantum algorithm for calculating energies of many-body quantum systems and our experimental implementation with linear optics. a.** The iterative phase estimation algorithm, see Chapter 5, at iteration k represented using quantum circuit notation. To produce an m-bit approximation to the phase of the eigenstate, ϕ, the algorithm is iterated m times. Each iteration obtains one bit of ϕ (ϕ_k); starting from the least significant (ϕ_m), k is iterated backwards from m to 1. The angle ω_k depends on all previously measured bits as $\omega_k = -2\pi b$, where b, in the binary expansion of ϕ, is $b = 0.0\phi_{k+1}\phi_{k+2}...\phi_m$ and $\omega_m = 0$. H is the Hadamard gate. **b.** Quantum circuit model representation of our gate network for a two-qubit controlled-U^j gate, as discussed in Subsection 8.4.1. **c.** Two-qubit optical implementation of the IPEA. Photon pairs are generated by spontaneous parametric down-conversion (SPDC), coupled into a single-mode optical fiber and launched into free space optical modes C (control) and R (register). The operation of the optical controlled-R_z gate is described by Lanyon et al. [127]. Coincident detection events (3.1 ns window) between single photon counting modules (SPCMs) D1 and D3 (D2 and D3) herald a successful run of the circuit and result 0 (1) for ϕ_k. Waveplates are labeled with their corresponding operations. PPBS: partially polarizing beam splitter. X is the standard Pauli X gate ($0 \leftrightarrow 1$) and $R_{\hat{n}}$ is the rotation of the qubit about the $\hat{n}$ axis of the Bloch sphere. $\lambda/2$ ($\lambda/4$) is a half (quarter) wave plate.

an eigenstate of our molecular Hamiltonian (H_{ij}) and the control photon is rotated into the superposition state $(|H\rangle + |V\rangle)/\sqrt{2}$. Both modes then pass through an optical network that implements a controlled-$U^{2^{k-1}}$ gate: if the control is horizontally polarized, then nothing is changed (the mathematical 'identity' operation is implemented). If the control is vertically polarized, then the operation $U_{ij}^{2^{k-1}}$ is implemented on the register polarization—that is, the evolution of the register polarization state is equivalent to (simulates) the evolution of a molecular eigenstate due to the molecular Hamiltonian, for a specific time interval. In this case, the joint polarization state will pick up a phase determined by k and the energy of the encoded eigenstate. Note that this non-linear process (that is, conditional dynamics at the single photon level) occurs only if a single photon exits each port, hence the circuit is non-deterministic (we note that this can be overcome and efficient scalable linear optical quantum computing is possible [115, 156]). Finally, the control polarization is measured in the H/V basis, yielding the result H (0) or V (1). This result is the estimate for one bit of the binary expansion of the phase. This same circuit (one iteration of the IPEA) is repeated with many photon pairs and a majority vote is used to determine the estimated bit value.

The circuit is changed in two ways for the next iteration: the molecule is simulated for a different time period (that is, k is changed), and previous results are fed back (the $R_z(\omega_k)$ gate is set to an angle determined by all previously measured bits, see Fig. 8.2). Both changes are implemented by changing the settings of wave plates.

To resolve the energy differences relevant to chemical processes, absolute molecular energies must be computed to an accuracy greater than $k_B T \sim 1$ kJ mol^{-1}. Furthermore, because we choose to encode the molecular Hamiltonian in atomic units, the phase ϕ is read out in Hartrees. Therefore, to achieve a precision of 1 kJ mol^{-1}, the IPEA must be able to resolve the phase to 14 bits, as $(2\pi \times 2^{-14})E_h$ corresponds to 1.01 kJ mol^{-1}. To demonstrate that the necessary precision can be achieved, we carried out the IPEA to 20 bits, using $n = 31$ samples per bit. Fig. 8.3 shows our results: H$_2$ energies calculated over a range of internuclear separations, reconstructing the potential energy surfaces. Each point is calculated with a 20-bit iterative phase estimation algorithm, using $n = 31$ samples per bit (repetitions of each iteration). In every case, the algorithm successfully returned the energy to within the target precision of $\pm(2^{-20} \times 2\pi)E_h \approx 10^{-5}E_h$, in the minimal basis. For example, the ground state energy obtained at the equilibrium bond length, 1.3886 atomic units is $-0.20399 \pm 0.00001\, E_h$, which agrees exactly with the results obtained on a classical computer to an uncertainty in the least significant bit.

8.4 Discussion

Achieving such high precision will become a far more significant challenge for large-scale implementations. Owing to the small-scale of our demonstration, we are able to implement each power of $U^{(p,q)}$ directly, by re-encoding the same number of gates, see Section 8.4.1. Therefore, the probability of error introduced by gate imperfections remains a constant for each bit. This is the main algorithmic feature that allows for the high precision obtained. However, following current proposals, the circuit network for U will not generally have the same form as U^j for larger implementations, as detailed in Section 8.1. For each
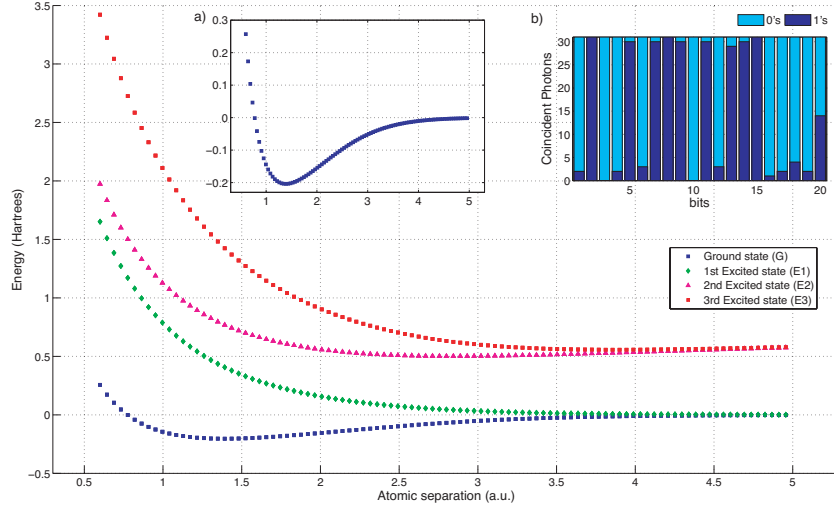
Figure 8.3: **Experimental quantum simulation results.** Each point is calculated using a 20-bit IPEA and employing $n = 31$ samples per bit. Every case was successful, achieving the target precision of $10^{-5}E_h$. Curve G (E3) is the low (high) eigenvalue of $H^{(1,6)}$. Curve E1 is a triply degenerate spin-triplet state, corresponding to the lower eigenvalue of $H^{(3,4)}$ as well as the eigenvalues of $H^{(2)}$ and $H^{(5)}$. Curve E2 is the higher (singlet) eigenvalue of $H^{(3,4)}$. Measured phases are converted to energies via $E = 2\pi\phi E_h + (e^2/R) - E_\infty$; the second term accounts for the proton-proton Coulomb energy at atomic separation R and E_∞ is the ground state energy of two hydrogen atoms at infinite separation. **Inset a**: Curve G rescaled to highlight the bound state. **Inset b**: Example of raw data for the ground state energy obtained at the equilibrium bond length, 1.3886 a.u. The measured binary phase is $\phi = 0.01001011101011100000$ which corresponds to the exact value, in our minimal basis, to a precision of $\pm 2^{-20}E_h$. Note that the exact value has a remainder of $\delta \approx 0.5$ after a 20 bit expansion, hence the low contrast in the measured 20th bit.

additional digit of precision sought, the gate requirements of the algorithm are roughly doubled, thereby amplifying any gate error.

8.4.1 Model Hamiltonian used for experimental implementation

We use the minimal STO-3G basis [88] for H_2, consisting of one $|1s\rangle$-type atomic orbital per atom. These two functions are combined to form the bonding (antibonding) molecular orbitals [209]. The four corresponding single-electron molecular spin-orbitals are combined antisymmetrically to form the six two-electron configurations $\{\Phi_i\}_{i=1}^{6}$ that form the basis for our simulation. Owing to symmetry, the Hamiltonian is block-diagonal in this basis, with blocks acting on each of the four subspaces spanned by $\{|\Phi_1\rangle, |\Phi_6\rangle\}$, $\{|\Phi_2\rangle\}, \{|\Phi_3\rangle, |\Phi_4\rangle\}$, and $\{|\Phi_5\rangle\}$, see below. Therefore, finding the eigenvalues of the two 2×2 sub-matrices in the Hamiltonian, $H^{(1,6)}$ and $H^{(3,4)}$, amounts to performing the FCI. Estimating the eigenvalues of 2×2 matrices is the simplest problem for the IPEA.

We employ a time step of $t = 1\ \hbar/E_h$, chosen so that $0 \leq Et/2\pi\hbar \leq 1$. For our proof-of-principle demonstration, all necessary molecular integrals are evaluated classically (see below) using the Hartree-Fock procedure [209]. We use these integrals to calculate the matrix elements of H and U, then directly decompose each $U^{(p,q)}$ operator into a logic gate network. We decompose the $U^{(p,q)}$ operators into a global phase and a series of rotations of the one-qubit Hilbert space

$$U = e^{i\alpha}R_y(\beta)R_z(\gamma)R_y(-\beta) \tag{8.9}$$

where α, β, and γ, are real angles. U^j is achieved by keeping β constant and replacing angles α and γ with $j\alpha$ and $j\gamma$, respectively. Our decomposition of the controlled-U^j is shown in Fig. 8.2b.

Minimal basis and symmetries in the electronic Hamiltonian of the hydrogen molecule

The two $|1s\rangle$-type atomic orbitals are combined to form the bonding and antibonding molecular orbitals [209], $|g\rangle$ and $|u\rangle$. Taking into account electron spin, the single-electron molecular spin-orbitals are denoted, $|g\uparrow\rangle$, $|g\downarrow\rangle$, $|u\uparrow\rangle$ and $|u\downarrow\rangle$, where $|\uparrow\rangle$ and $|\downarrow\rangle$ are the electron spin eigenstates. These are combined antisymmetrically to form Slater determinants, denoted $|\Psi|$. Each row of a Slater determinant corresponds to an electron and each column to an orbital [209]. The six two-electron configurations that form the basis for our simulation are: $|\Phi_1\rangle = |g\uparrow, g\downarrow\rangle \equiv (|g\uparrow, g\downarrow\rangle - |g\downarrow, g\uparrow\rangle)/\sqrt{(2)}$, $|\Phi_2\rangle = |g\uparrow, u\uparrow|$, $|\Phi_3\rangle = |g\uparrow, u\downarrow|$, $|\Phi_4\rangle = |g\downarrow, u\uparrow|$, $|\Phi_5\rangle = |g\downarrow, u\downarrow|$, and $|\Phi_6\rangle = |u\uparrow, u\downarrow|$. Due to symmetry, the Hamiltonian is block-diagonal in this basis, with blocks acting on each of the four subspaces spanned by $\{|\Phi_1\rangle, |\Phi_6\rangle\}$, $\{|\Phi_2\rangle\}$, $\{|\Phi_3\rangle, |\Phi_4\rangle\}$, and $\{|\Phi_5\rangle\}$. Configurations from different subspaces not mixed by the Hamiltonian. In particular, $|\Phi_1\rangle$ and $|\Phi_6\rangle$ mix only with each other because they have g (even) symmetry while the rest have u (odd) symmetry. Of the remaining states only $|\Phi_3\rangle$ and $|\Phi_4\rangle$ mix because they have the same total z-projection of the spin, $m_S = 0$. $|\Phi_2\rangle$ and $|\Phi_5\rangle$ have, respectively, $m_S = 1$ and $m_S = -1$. Therefore, the Hamiltonian is block-diagonal with four subspaces spanned by $\{|\Phi_1\rangle, |\Phi_6\rangle\}$, $\{|\Phi_2\rangle\}$, $\{|\Phi_3\rangle, |\Phi_4\rangle\}$, and $\{|\Phi_5\rangle\}$. There are no approximations involved here, and finding

the eigenvalues of the two 2×2 sub-matrices in the Hamiltonian ($H^{(1,6)}$ and $H^{(3,4)}$) amounts to performing an exact calculation (FCI) in the minimal basis. One should also note that it follows from the requirement that the wave functions are spin eigenstates, that the eigenstates of the subspace $\{|\Phi_3\rangle, |\Phi_4\rangle\}$ will be $(|\Phi_3\rangle \pm |\Phi_4\rangle)/\sqrt{2}$. Additionally, there will be a three-fold degeneracy of the triplet state with angular momentum $S = 1$. That is, the states $|\Phi_2\rangle$, $|\Phi_5\rangle$, and $(|\Phi_3\rangle + |\Phi_4\rangle)/\sqrt{2}$ are degenerate.

8.4.2 Details of classical computational methods

Restricted Hartree-Fock calculations were carried out on a classical computer using the STO-3G basis [88]. The software used was the PyQuante quantum chemistry package version 1.6. The molecular integrals from the Hartree-Fock procedure are used to evaluate the matrix elements of the Hamiltonians $H^{(1,6)}$ and $H^{(3,4)}$, described earlier.

8.4.3 Classical error correction techniques

When running the IPEA, the probability of correctly identifying any individual bit with a single sample ($n = 1$) is reduced from unity by both theoretical (inexact phase expansion to k bits) and experimental factors (such as imperfect gates). However, as long as it remains above 0.5, repeated sampling and a majority vote will reduce the probability of error exponentially with n, in accordance with the Chernoff bound [155]. This technique allows for a significant increase in success probability, at the expense of repeating the experiment a fixed number of times. We note that this simple classical error correction technique can only play a small role when it comes to dealing with errors in large-scale implementations. Here, the numerous errors in large quantum logic circuits will make achieving a bit success probability over 0.5 a significant challenge. This will need to be addressed with quantum error correction techniques [155, 57].

8.4.4 Count rates

We operate with a low-brightness optical source (spontaneous parametric down-conversion pumping power 50 mW) to reduce the effects of unwanted multi-photon-pair emissions (which cannot be distinguished by our non-photon-number-resolving detectors and introduce error into the circuit operation). This yields approximately 15 coincident detection events per second at the output of our optical circuit. Therefore, each iteration can be repeated 15 times a second. Reconfiguring the circuit for different iterations takes approximately 7 seconds, largely due to the finite time required to rotate standard computer-controlled waveplate mounts. Therefore, obtaining a 20-bit estimation of a phase takes approximately 3 minutes, when using $n = 31$ samples to determine the logical state of each bit (as was employed to achieve the results shown in Fig. 8.2). Note that approximately 95% of this time is spent rotating waveplates. In future implementations, this time could be reduced significantly using integrated-photonics such as qubit manipulation using an electrooptically-controlled waveguide Mach-Zehnder interferometer [135].

8.5   Additional experimental results

We use the estimation of the ground state energy at the equilibrium bond length, 1.3886 a_0 (where a_0 is the Bohr radius) to study the effect of varying a range of experimental parameters on the IPEA success probability, where we define the latter as the probability of correctly obtaining the phase to an precision of 2^{-m}. Fig. 8.5a shows results measured over a range of n, the number of samples used to determine each bit.

The probability of correctly identifying any individual bit with a single sample ($n = 1$) is reduced from unity by both theoretical (δ) and experimental factors (such as imperfect gates). However, as long as it remains above 0.5, repeated sampling and a majority vote will improve the probability of correct identification. The data show that this is achieved and the error probability decreases exponentially with n. This technique allows for a significant increase in success probability, at the expense of repeating the experiment a fixed number of times. We note that this simple classical error correction technique can only play a small role when it comes to dealing with errors in large-scale implementations. Here, the numerous errors in large quantum logic circuits will make achieving a bit success probability over 0.5 a significant challenge, that must be met with quantum error correction techniques [155, 57].

Fig. 8.5b shows the algorithm success probability measured as a function of the number of extracted bits (phase precision). By employing $n = 101$ samples per bit, we achieve near perfect algorithm success probability up to 47 bits (yielding an energy precision of $\approx 10^{-13}E_h$), where this limit is imposed only by the machine-level precision used for the classical preprocessing of the Hamiltonians. It is insightful to understand how achieving such high precision will become a far more significant challenge for large-scale implementations: due to the small-scale of our demonstration, we are able to implement each power of $U^{(i,j)}$ directly. By re-encoding the same number of gates. Therefore, the probability of error introduced by gate imperfections remains a constant for each bit (and, in our implementation, under 50%). This is the main algorithmic feature that allows the high precision obtained in this experiment. However, as expounded earlier, this will not be possible for larger implementations. In general, U will not have the same form as U^n. For each additional digit of precision sought, the gate requirements of the algorithm are roughly doubled, thereby amplifying any gate error.

Fig. 8.5c shows the algorithm success probability measured as a function of the fidelity F (see caption) between the encoded register state and the ground state. The results show that our implementation is robust for $F > 0.5$. Because the probability of correctly obtaining each bit in a single measurement ($n = 1$) is greater than 0.5 in this regime, multiple sampling ($n > 1$) enables the success probability to be amplified arbitrarily close to unity. This is a general feature that will hold for large-scale implementations. However, for $F < 0.5$, the measured success probabilities are very low.

If the register state output after each iteration is used as the input of the next, then the problem with low eigenstate fidelities can be overcome as the measurement of the control qubit collapses the wave function. Any pure encoded register state can be written in the eigenstate basis as $|G\rangle = \sum_i \alpha_i |\lambda_i\rangle$, where $|\alpha_i|^2$ is the fidelity of $|G\rangle$ with eigenstate $|\lambda_i\rangle$. Successful measurement of the mth bit associated with $|\lambda_i\rangle$ will cause the register wave function to collapse into a state with a greater fidelity with $|\lambda_i\rangle$—those eigenstates with a low

probability of returning the measured bit value will be diminished from the superposition. As more bits are successfully measured, the register state will rapidly collapse to $|\lambda_i\rangle$. In this way, the algorithm will return all the bits associated with $|\lambda_i\rangle$ with probability at least $|\alpha_i|^2(1-\epsilon)$. With current technology, correct operation of our optical circuit requires destructive measurement of both the control and register qubits after each IPEA iteration. Therefore, in our experiment the register state must be re-prepared for each iteration.

Obtaining IPEA success probabilities

Denoting the first m binary bits of a phase ϕ as $\tilde{\phi} = 0.\phi_1\phi_2...\phi_m$, there is, in general, a remainder $0 \le \delta < 1$, such that $\phi = \tilde{\phi} + \delta 2^{-m}$. To achieve an accuracy of $\pm 2^{-m}$, the IPEA success probability is the sum of the probabilities for obtaining $\tilde{\phi}$ and $\tilde{\phi} + 2^{-m}$. This can be estimated experimentally, for a given phase, by simply repeating the algorithm a large number of times and dividing the number of acceptable results by the total. An estimate with an error less than 10% would require over 100 algorithm repetitions. We calculate the result shown in Fig. 8.5c in this way. However, using this technique to obtain Fig. 8.4b-c, and Fig. 8.5 (described below), would take a long time—the 20 points shown in each would require more than 100 hours of waveplate rotation time alone. Instead, to obtain these results we force the appropriate feedforward trajectory $(R(\omega_k))$ for each accepted phase value and use $n = 301$ samples to estimate the 0/1 probabilities for each bit. Using the standard binomial cumulative distribution function, it is then possible to calculate the majority vote success probability for each bit of each accepted value for a given n (1 and 101 in the figures). The probability for obtaining an accepted phase value is then the product of the majority vote success probabilities for each bit, and the total algorithm success probability is the sum of the probabilities for obtaining each accepted phase. The error bars represent a 68% confidence interval and are obtained from a direct Monte-Carlo simulation of the above process.

Note that forcing the correct feedforward in this way, and taking many samples to estimate the 0/1 probabilities for each bit, simply allows us to accurately estimate the probability that the algorithm will return the correct phase by itself, without forcing the correct feedforward.

Experimental model

A simple computational model of our experiment produced the lines shown in Fig. 8.5. This model allows for two experimental imperfections, which are described below, but otherwise assumes perfect optic element operation. The model consists of a series of operators, representing optical elements and noise sources, acting on a vector space representing both photonic polarization and longitudinal spatial mode [127]. Firstly the model allows for photon distinguishability, quantified by an imperfect relative non-classical interference visibility of 0.93 (ideal 1), which reduces the quality of our two-qubit logic gate. Secondly, the model allows for phase damping of the control qubit, described by the operation elements [155]

$$\begin{pmatrix} 1 & 0 \\ 0 & \sqrt{1-\gamma} \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} 0 & 0 \\ 0 & \sqrt{\gamma} \end{pmatrix} \tag{8.10}$$
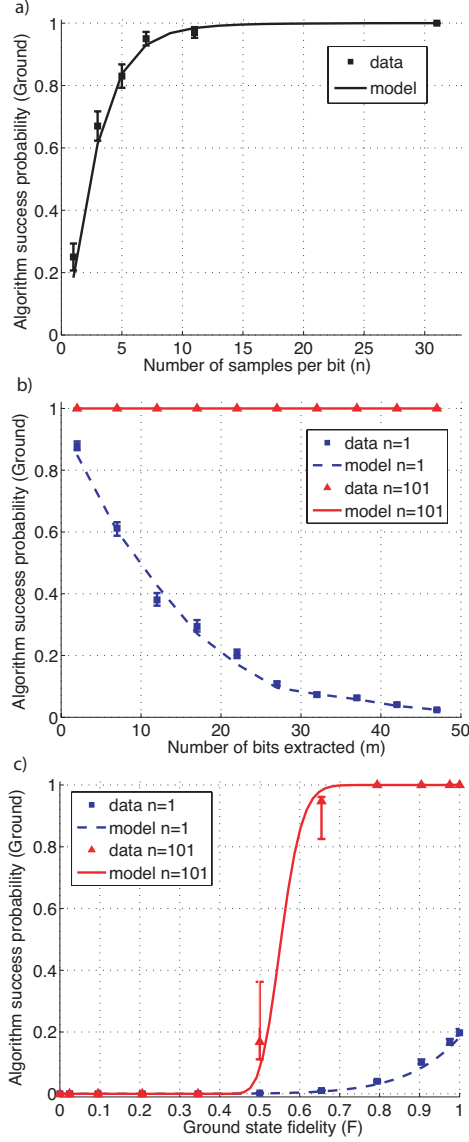
Figure 8.4: **IPEA success probability measured over a range of parameters.** Probabilities for obtaining the ground state energy, at the equilibrium bond length 1.3886 a_0, as a function of: **(a)** the number of times each bit is sampled (n); **(b)** the number of extracted bits (m); **(c)** the fidelity between the encoded register state and the ground state (F). The standard fidelity [155] between a measured mixed ρ and ideal pure $|\Psi\rangle$ state is $F = \langle\Psi|\rho|\Psi\rangle$. **(a)** & **(b)** employ a ground state fidelity of $F \approx 1$. **(a)** & **(c)** employ a 20-bit IPEA. All lines are calculated using a model that allows for experimental imperfections. This model, as well as the technique used to calculate success probabilities and error bars, are detailed in the text.

Our model employs $\gamma = 0.06$ (ideal 0), which corresponds to $\approx 3\%$ dephasing. These experimental imperfections are attributed to a combination of residual higher-order photon pair emissions from our optical source and circuit alignment drift during long measurement sets.

Chapter 9

Adiabatic quantum simulators

> Festina lente

–Latin adage, translation: *make haste slowly.*

> Harraka harraka hiena barraka

–Swahili adage, translation: *hurry hurry blessing has no blessing.*

9.1 Introduction

In this thesis, the exploration of quantum simulation for chemistry has been pursued using the quantum circuit model. In this chapter, I will describe our work on another model of quantum simulation using a very different construction. Constructing a practical method of quantum simulation is a significant challenge. Gate-model based simulation methods (with quantum error correction) can provide a scalable solution, but are well out of reach of present-day experiments, except for small systems.

On the other hand, some experimental implementations seem to be well suited to operate as adiabatic processors where the computation is done passively by changing control parameters slowly. For these setups, moving away from the gate model may offer a less resource-intensive, and consequently a more feasible solution for simulating medium-sized quantum systems. This chapter addresses the problem by presenting a hybrid model of quantum simulation, consisting of an adiabatically controlled simulation register coupled to a single gate-model readout qubit. Our scheme can simulate the constant observables of arbitrary spin graph Hamiltonians. It allows the measurement of the expectation value of any constant k-local observable using a $k+1$-local measurement Hamiltonian. Fault tolerance of the adiabatic model of quantum computing is a topic of growing interest. Its robustness in the presence of noise has been studied in [53, 13]. We only consider the simulation protocol here, and do not study the adiabatic error correction that would be required for a practical large scale implementation.

In most quantum computing architectures, the natural interactions are two-local. However, under certain conditions k-local interactions can be well approximated using techniques such as perturbative Hamiltonian gadgets [107, 162, 30, 29] or average Hamiltonian
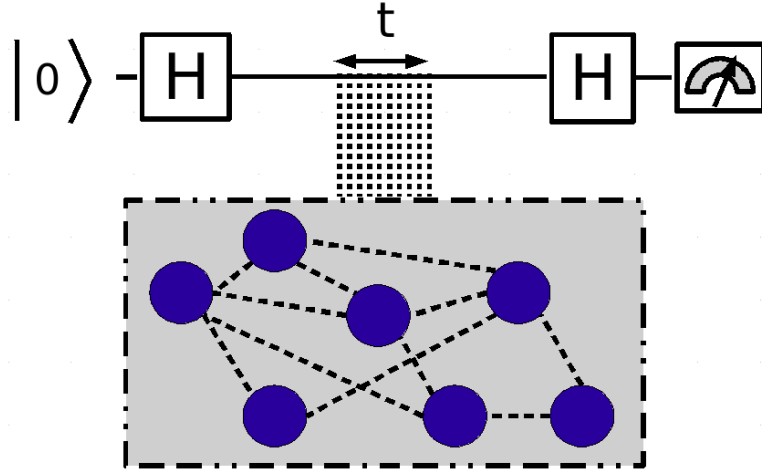
Figure 9.1: **A depiction of the AQS scheme.** The upper line represents the readout qubit which is the only qubit where quantum gates are applied. The grayed box represents the system which is coupled to the readout qubit for time t. After repeating this measurement with various times the energy of the system is obtained.

methods [227] — this provides even the possibility of utilizing gate model fault tolerance to protect the slower adiabatic evolution, an approach that seems promising. We note that Ref. [36] considered the mapping of a given n-qubit target Hamiltonian with k-local interactions onto a simulator Hamiltonian with two-local interactions.

In this chapter, we will begin by giving an overview of the simulator design, including initialization, adiabatic evolution, and measurement. The following section investigates the performance of the method in simulating a small system in the presence of noise. Finally, we will present our conclusions followed by further details of the methods, including details of the numeric simulations.

9.2 Simulator overview

We consider the simulation of systems represented by finite collections of spins acted on by a time-independent Hamiltonian described by a graph, $G = (V, E)$ — e.g. , the Heisenberg and Ising models. Each graph vertex, $v \in V$, corresponds to a spin acted on by a local Hamiltonian L_v, and each edge $e \in E$ to a two-local interaction K_e between the involved vertices. The target Hamiltonian is given by

$$H_T = \sum_{v \in V} L_v + \sum_{e \in E} K_e \tag{9.1}$$

The simulator, see Fig. 9.1, consists of an adiabatically controlled simulation register S with Hamiltonian H_S, and a probe register, P, which will be acted on by gate

operations and measured projectively. We will engineer the probe P such that it behaves as a controllable two-level system with the orthonormal basis $\{|p_0\rangle, |p_1\rangle\}$. Without loss of generality, the probe Hamiltonian can be expressed as $H_P = \delta|p_1\rangle\langle p_1|$, where δ is the spectral gap between the probe's ground and first excited states.

9.2.1 Initialization

We first set the simulation register Hamiltonian H_S to $H_{S,I}$ and prepare S and P in their respective ground states. The Hamiltonian $H_{S,I}$ has a simple ground state which can be (i) computed classically and (ii) prepared experimentally in polynomial time — such as a classical approximation to the ground state of the simulated system. The simulator Hamiltonian is thus initially given by

$$H_0 = H_{S,I} \otimes \mathbf{1}_P + \mathbf{1}_S \otimes H_P \tag{9.2}$$

9.2.2 Adiabatic evolution

According to the adiabatic theorem [146], a quantum system prepared in an energy eigenstate will remain near the corresponding instantaneous eigenstate of the time-dependent Hamiltonian governing the evolution if there are no level crossings and the Hamiltonian varies slowly enough. By adjusting the simulator parameters, we adiabatically change H_S from $H_{S,I}$ to $H_{S,T}$, the fully interacting Hamiltonian of the system to be simulated.

Let us denote the ground state of $H_{S,T}$ as $|s_0\rangle$. At the end of a successful adiabatic evolution, register P is still in its ground state $|p_0\rangle$, and register S is in (a good approximation to) the ground state of the simulated system, $|s_0\rangle$. Hence, the simulator is now in the ground state $|g\rangle = |s_0\rangle \otimes |p_0\rangle$ of its instantaneous Hamiltonian

$$H_1 = H_{S,T} \otimes \mathbf{1}_P + \mathbf{1}_S \otimes H_P \tag{9.3}$$

9.2.3 Measurement

The measurement procedure begins by bringing S and P adiabatically into interaction. The simulator Hamiltonian becomes

$$H_2 = H_1 + \underbrace{A \otimes |p_1\rangle\langle p_1|}_{H_{SP}} \tag{9.4}$$

where the operator A corresponds to an observable of the simulated system that is a constant of motion, that is, it commutes with $H_{S,T}$. Hence, the total energy itself can always be measured by choosing $A = H_{S,T}$. Other such observables depend on the particular system and often can be analytically constructed given the Hamiltonian.

Let us use $a_{\min}$ to denote the lowest eigenvalue of A. If $a_{\min} + \delta > 0$, then $|g\rangle$ is also the ground state of H_2 and, in the absence of noise, the transitions from the ground state are perfectly suppressed during the adiabatic evolution. Assuming that A can be decomposed into a sum of two-local operators, the interaction term, H_{SP}, involves

three-local interactions. These terms can be implemented using either Hamiltonian gadget techniques or average Hamiltonians, see Sections 9.4 and 9.5 for details.

After the adiabatic evolution, at time $t = 0$, we apply a Hadamard gate to the measurement probe which puts it into a superposition of its two lowest states. This is no longer an eigenstate of H_2, and the system will evolve as

$$|\psi(t)\rangle = \frac{1}{\sqrt{2}}|s_0\rangle \otimes (|p_0\rangle + e^{-i\omega t}|p_1\rangle) \tag{9.5}$$

where $\omega = (a_0 + \delta)/\hbar$, and $a_0 = \langle s_0|A|s_0\rangle$ is the expectation value we wish to measure. We have thus encoded the quantity, a_0, a property of the ground state of $H_{S,T}$, into the time dependence of the probe P.

After a time t, we again apply a Hadamard gate to the probe, resulting in the state

$$|\psi(t)\rangle = |s_0\rangle \otimes (\cos(\omega t/2)|p_0\rangle + i \sin(\omega t/2)|p_1\rangle) \tag{9.6}$$

and then measure the probe. The probability of finding it in the state $|p_0\rangle$ is

$$P_0(t) = \frac{1}{2}(1 + \cos(\omega t)) = \cos^2(\omega t/2) \tag{9.7}$$

If we have non-demolition measurements, see e.g. [194], at our disposal, then measuring the probe does not disturb the state of the simulator which can be reused for another measurement.

One repeats the measurement with different values of t until sufficient statistics have been accumulated to reconstruct ω and hence a_0 — this is reminiscent of Ramsey spectroscopy [179] and hence should seem natural to experimentalists. In essence, we have performed the phase estimation algorithm without the corrective gate in Eq. (5.5), using the interaction Hamiltonian H_{SP} instead of a controlled unitary.

If the ground state subspace of $H_{S,T}$ is degenerate and overlaps more than one eigenspace of A, or the simulation register S has excitations to higher energy states at the beginning of the measurement phase, the probability of finding the probe in the state $|p_0\rangle$ is given by a superposition of harmonic modes. For example, for a thermalized state with inverse temperature β, we obtain

$$P_0(t) = \frac{1}{2}\left(1 + \frac{1}{\sum_{xy} e^{-\beta E_x}} \sum_{kl} e^{-\beta E_k} \cos(\omega_{k,l} t)\right) \tag{9.8}$$

where the first summation index runs over the energy eigenstates and the second over the eigenstates of A in which energy has value E_k, and $\omega_{k,l} = (a_{k,l} + \delta)/\hbar$.

9.3 Effects of noise

Any large scale implementation of the proposed method would most likely require adiabatic error correction [100, 136, 12]. However, for small systems this might not be necessary indicating that our method is feasible for immediate experimental technologies.

For this reason, we will examine how robust a small scale implementation of our protocol would be.

To assess the effects of noise on the simulation method, we performed a numerical simulation of the simplest nontrivial implementation of the hybrid simulator, consisting of two simulator qubits and one probe qubit, with a randomly chosen H_T. Each qubit was coupled to its own bosonic heat bath with an Ohmic spectral density using the Born-Markov approximation [37]. The qubit-bath couplings were chosen such that the resulting single-qubit decoherence times T_1 and T_2 are compatible with recent superconducting flux qubit experiments with fully tunable couplings, e.g. see [157]. The noise model is described further in Section 9.6. The observable A was chosen to be $H_{S,T}$, the simulated Hamiltonian itself, which means that the ground state energy was being measured.

We simulated measurements on 40 evenly distributed values of the time delay t. At each t_i we performed 50 measurements, and averaged the results to get an estimate of $P_0(t_i)$. An exponentially decaying scaled cosine function was then fitted to this set of data points to obtain an estimate for ω and thus for s_0. The fit was done using the MATLAB LSQNONLIN algorithm, guided only by fixed order of magnitude initial guesses for the parameter values.
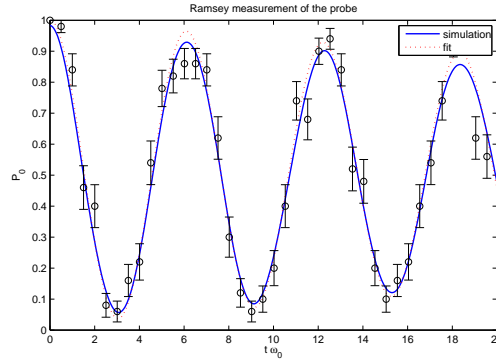
Figure 9.2: **Measurement procedure under Markovian noise.** The continuous curve represents the probability of finding the probe in the state $|p_0\rangle$, the circles averaged measurement results and the dotted line a least squares fit to them. $\hbar\omega_0 = h \cdot 25$ MHz is the energy scale of the Hamiltonians involved.

The results of the simulation are presented in Fig. 9.2. The noise, together with the slight nonadiabaticity of the evolution, cause excitations out of the ground state which result in a signal consisting of multiple harmonic modes. However, the ground state mode still dominates and with a realistic level of noise and relatively modest statistics we are able to reconstruct ω to a relative precision of better than 0.01. This includes both the uncertainty due to finite statistics, and the errors introduced by the environment through decoherence and the Lamb-Stark shift.

In an experimental implementation, there may also be other noise sources not considered here related to the measurement process itself, as well as systematic (hardware)

errors in the quantum processor (qubits, couplers, etc.). Nonetheless, our results indicate that a simple experimental implementation of the simulation scheme could be possible using existing hardware. For future implementations, there exist fault tolerant constructions for adiabatic quantum computing [100, 136, 12, 16, 61].

The presented simulation scheme differs significantly from existing gate-model methods. Instead of a series of coherent gate operations, it uses an adiabatic control sequence which may require less complicated control hardware. At the measurement stage we require single-qubit gate operations and measurements, but only on the probe qubit. These operations should be relatively simple to implement compared to a full Trotter decomposition of the simulated Hamiltonian. Without error correction, our method has limited scalability, but it might outperform gate-model simulators in some small to medium sized problems. A simple experimental implementation could be feasible with present-day technology. In order to simulate a system of n qubits with a two-local Hamiltonian described by the graph G, ideally our scheme requires one probe qubit for the readout and n simulation qubits. Additionally, if Hamiltonian gadgets are used to implement three-local interactions, one ancilla qubit is required for each two-local term in the simulated observable (represented by an edge in the corresponding graph). In practice, the number of ancillas may be slightly higher if more involved types of gadgets are used to implement the three-local interactions. The total number of qubits required thus scales as $O(n)$ for sparse graphs and $O(n^2)$ for maximally connected ones.

9.4   Gadgets

One possible way to effect the 3-local Hamiltonians we require in our construction is to approximate them with 2-local interactions.

Perturbative gadget

This Gadget Hamiltonian construction was proposed in [107] and has since been used and extended by others [162, 30]. These papers contain further background details including notation.

We label the working qubits 1–3. The effective 3-local interaction is well approximated on qubits 1, 2, and 3 (within ϵ) in a low energy subspace — constructed by adding a penalty Hamiltonian to an ancillary (mediator) qubit, m. The mediator qubit, m, doubles the size of the state space and the penalty Hamiltonian splits the Hilbert space into low and high energy subspaces, separated by an energy gap Δ (which is inversely proportional to a polynomial in ϵ — in our case reduced to $\Delta \geq \epsilon^{-3}$). The details of the gadget we develop for use in this work follow.

We will apply a Hamiltonian H_p to the mediator qubit m as well as a Hamiltonian V to qubits 1, 2, and 3 and m. The Hamiltonian $H_p + V$ has a ground state that is ϵ-close to the desired operator $JA_1 \otimes A_2 \otimes A_3 \otimes |0\rangle\langle 0|$.

$$H_p \equiv \Delta |1\rangle\langle 1|_m \tag{9.9}$$

$$V(J, \Delta(\varepsilon)) \equiv \mathbf{y} + \Delta^{1/3} |0\rangle\langle 0|_m - \Delta^{1/3} A_1 \otimes A_2 + \frac{\Delta^{2/3}}{\sqrt{2}} (A_2 - A_1) \otimes \sigma^x + J A_3 \otimes (\mathbf{1} - \Delta^{2/3} |1\rangle\langle 1|_m)$$

(9.10)

where $\mathbf{y}$ is some Hamiltonian already acting on qubits $1 - 3$ as well as a possible larger Hilbert space. Note the gadget above assumes $A_i^2 = \mathbf{1}$, $\forall i = 1, 2, 3$.

The Dyson series can be rewritten as the *self-energy expansion* and under the appropriate conditions is known to provide a series approximation to the low-lying eigenspace of an operator. To verify that the Hamiltonian $H_p + V$ gives the desired approximation, one relies on expansion of the self-energy potential to 4th order (here the higher order terms give rise to effective interactions greater than second order) given as

$$\Sigma_-(z) = \langle 0|V|0\rangle + \frac{\langle 0|V|1\rangle\langle 1|V|0\rangle}{z - \Delta} + \frac{\langle 0|V|1\rangle\langle 1|V|1\rangle\langle 1|V|0\rangle}{(z - \Delta)^2} + \mathcal{O}\left(\frac{\|V\|^4}{\Delta^3}\right)$$

(9.11)

where the operator is written in the $\{|0\rangle, |1\rangle\}$ basis of m. One considers the range $|z| \leq 2|J| + \epsilon$ and notes that $\|\Sigma_-(z) - J A_1 \otimes A_2 \otimes A_3\| = \mathcal{O}(\epsilon)$ and applies the Gadget Theorem [107].

Before concluding this section, we note that care must be taken when adiabatically evolving gadgets. Ref. [6] contains a proof that the linear path Hamiltonian is universal for adiabatic quantum computation. Universality (and hence a non-exponentially contracting gap) remains when the locality of the construction is reduced using perturbative gadgets [162].

9.4.1 Exact diagonal gadget

Modern experimental implementations of adiabatic quantum computers typically are limited to only being able to couple spins with one type of coupling e.g. $\sigma^z \otimes \sigma^z$. In such a case, the standard gadget Hamiltonian approach will not work as these gadgets require multiple types of couplers [29]. We will now provide a new type of gadget Hamiltonian which creates an effective, exact $\sigma^z \otimes \sigma^z \otimes \sigma^z$ interaction in the low energy subspace using just $\sigma^z \otimes \sigma^z$ and local terms.

Let us first assume that we have access to a penalty function $H_{\text{AND}}(x_*, x_1, x_2)$, where $x_i \in \{0, 1\}$ such that $H_{\text{AND}} = 0$ any time $x_* = x_1 x_2$ and is greater than some large constant Δ for all $x_* \neq x_1 x_2$. By solving a system of constraints, such a penalty function is possible to write as a sum of two-local terms

$$H_{\text{AND}}(x_*, x_1, x_2) = \Delta(3x_* + x_1 x_2 - 2x_* x_1 - 2x_* x_2)$$

(9.12)

The Boolean variables $\{x_i\}$ can be represented on qubits using the correspondence $x_i \simeq \hat{x}_i = |1\rangle\langle 1|_i$. Furthermore, we have $\sigma^z = \mathbf{1} - 2|1\rangle\langle 1|$ as usual. This gives us

$$H = \sigma_1^z \sigma_2^z \sigma_3^z = \mathbf{1} - 2\hat{x}_1 - 2\hat{x}_2 - 2\hat{x}_3 + 4\hat{x}_1 \hat{x}_2 + 4\hat{x}_1 \hat{x}_3 + 4\hat{x}_2 \hat{x}_3 - 8\hat{x}_1 \hat{x}_2 \hat{x}_3$$
$$\hat{=} \mathbf{1} - 2\hat{x}_1 - 2\hat{x}_2 - 2\hat{x}_3 + 4\hat{x}_1 \hat{x}_2 + 4\hat{x}_1 \hat{x}_3 + 4\hat{x}_2 \hat{x}_3 - 8\hat{x}_* \hat{x}_3 + H_{\text{AND}}(x_*, x_1, x_2)$$

(9.13)

where (9.13) holds in the low energy subspace after the introduction of an ancilla qubit $*$.

By writing the Boolean operators in terms of Pauli matrices, we obtain

$$
\begin{aligned}
H \;=\; & \left(\frac{3}{4}\Delta - 1\right)\mathbf{1} + \left(2 - \frac{\Delta}{2}\right)\sigma_*^z + \left(\frac{\Delta}{4} - 1\right)(\sigma_1^z + \sigma_2^z) + \\
& \sigma_3^z - \frac{\Delta}{2}\sigma_*^z(\sigma_1^z + \sigma_2^z) - 2\sigma_*^z\sigma_3^z \\
& + \left(\frac{\Delta}{4} + 1\right)\sigma_1^z\sigma_2^z + (\sigma_1^z + \sigma_2^z)\sigma_3^z
\end{aligned}
$$

We have hence provided a method which allows one to create diagonal k-local couplings using one- and two-local couplings. This method opens the door to simulate a wider range of Hamiltonians using current and next generation quantum computing technology. In addition, it should also prove useful in future investigations into the fault tolerance of the proposed protocol, in a way similar to the comparison in [40].

9.5 Average Hamiltonian method

An alternate method of generating the special Hamiltonians we require is to make use of the time average of a series of simpler generating Hamiltonians [227]. It has long been known that by regularly switching between a set of fixed Hamiltonians $\{H_i\}$, it is possible to approximate time evolution under any other Hamiltonian, H, provided that H is contained within the algebra generated by $\{H_i\}$. This fact lies at the heart of both average Hamiltonian theory and geometric control theory. Over the years, many methods have been developed for making the approximations accurate to high order; however, here we will focus only on approximations correct to first order.

In order to construct an average Hamiltonian, we will make use of a first order approximation to the Baker-Campbell-Hausdorff formula

$$
\log(e^A e^B) \approx A + B + \frac{1}{2}[A, B] \tag{9.14}
$$

From this we obtain formulas for approximating the exponential of both the sum and the Lie bracket of A and B to first order.

$$
\begin{aligned}
e^{A+B} &\approx e^A e^B \tag{9.15} \\
e^{[A,B]} &\approx e^A e^B e^{-A} e^{-B} \tag{9.16}
\end{aligned}
$$

These equations can be related to time evolution under some Hamiltonians, H_i, by replacing A and B with operators of the form $iH_i t/\hbar$. Clearly, by applying these rules recursively, it is possible to generate any Hamiltonian in the Lie algebra generated by the initial set of static Hamiltonians. We note that the combination of any pairwise entangling Hamiltonian, such as an Ising Hamiltonian together with tunable local Z and X fields, is sufficient to generate the full Lie algebra $\mathfrak{su}(2^N)$ for a N qubit system, and so can be used to approximate an arbitrary Hamiltonian.

Although the time-varying Hamiltonian means that the system does not have a ground state in the normal sense, the average energy of the system is minimized when the

Table 9.1: Noise model parameters

T_1	$\sim N(1.0, 0.1)$ μs
T_ϕ	$\sim N(1.3, 0.1)$ μs
T_2^{-1}	$= \frac{1}{2}T_1^{-1} + T_\phi^{-1}$
T	20 mK
ω_c	20 ω_0
ω_0	$2\pi \cdot 25$ MHz

system is within $O(t)$ of the ground state of the target average Hamiltonian. As a result of this, if the time scale for switching between Hamiltonians is small compared to the time scale for the adiabatic evolution of the system, the system will behave as if it was experiencing the average Hamiltonian.

9.6 Noise model

The noise model used in our MATLAB simulation consists of a separate bosonic heat bath coupled to each of the qubits. The baths have Ohmic spectral densities

$$J(\omega) = \hbar^2 \omega \Theta(\omega) \Theta(\omega_c - \omega) \tag{9.17}$$

where the cutoff, ω_c, was chosen to be above every transition frequency in the system, and are assumed to be uncorrelated. Each bath is coupled to its qubit through an interaction operator of the form

$$D = \lambda(\cos(\alpha)\sigma^z + \sin(\alpha)\sigma^x) \tag{9.18}$$

Using the Born-Markov approximation, we obtain an evolution equation which is of the Lindblad form [37]. Denoting the level splitting of a qubit by $\hbar\Delta$, we obtain the following uncoupled single-qubit decoherence times

$$T_1^{-1} = \lambda^2 \sin^2(\alpha) 2\pi\Delta \coth(\beta\hbar\Delta/2) \tag{9.19}$$

$$T_2^{-1} = \frac{1}{2}T_1^{-1} + \lambda^2 \cos^2(\alpha) 4\pi/(\hbar\beta) \tag{9.20}$$

where $\beta = \frac{1}{k_B T}$. Given T_1, T_2, T, and Δ, we can solve the bath coupling parameters λ and α separately for each qubit, and then use the same noise model in the fully interacting case. The values used for the parameters are presented in Table 9.1. The single-qubit decoherence times, T_1 and T_2, were chosen to be compatible with recent coupled-qubit experiments with fully tunable couplings such as [157]. The bath temperature, T, and the energy scale, $\hbar\omega_0$ of H_T, the Hamiltonian to be simulated, were likewise chosen to match the temperatures and coupling strengths found in contemporary superconducting qubit experiments. To simulate manufacturing uncertainties, the actual values of the T_1 and T_ϕ parameters for the individual qubits were drawn from a Gaussian distribution with a small standard deviation.

9.7 Measurement theorems

The pre-measurement system Hamiltonian is

$$H_1 = H_{S,T} \otimes \mathbf{1} + \mathbf{1} \otimes H_P \tag{9.21}$$

The operators H_S and H_P can be expanded in terms of their eigenvalues and eigenstates using the (possibly degenerate) spectral decomposition

$$H_S = \sum_{kj} s_k |s_{k,j}\rangle\langle s_{k,j}|, \quad s_0 < s_1 < \dots \tag{9.22}$$

and correspondingly for H_P.

Let the states of systems S and P begin in their ground state subspaces, the full normalized state $|g\rangle$ of the simulator belonging to the ground state subspace of the non-interacting Hamiltonian H_1

$$|g\rangle \in \text{span}\{|s_{0,k}\rangle \otimes |p_{0,l}\rangle\}_{k,l} \tag{9.23}$$

Registers S and P are brought adiabatically into interaction with each other. The Hamiltonian becomes

$$H_2 = H_1 + \underbrace{A \otimes (\mathbf{1}_P - \Pi_{p_0})}_{H_{SP}} \tag{9.24}$$

where the operator A corresponds to an observable of the simulated system that commutes with H_S, and

$$\Pi_{p_0} = \sum_m |p_{0,m}\rangle\langle p_{0,m}| \tag{9.25}$$

is the projector to the ground state subspace of H_P. Because A and H_S commute, they have shared eigenstates

$$H_S |s_{k,j}\rangle = s_k |s_{k,j}\rangle \tag{9.26}$$

$$A |s_{k,j}\rangle = a_{k,j} |s_{k,j}\rangle \tag{9.27}$$

A ground state lemma

We will now show that Hamiltonians, H_1 and H_2, have the same ground state subspace given that $a_{\min} + p_1 - p_0 > 0$ where $a_{\min}$ is A's lowest eigenvalue.

Lemma 8. *Let H_1 and H_2 be the finite dimensional Hamiltonians defined previously in Eqs. (9.21) and (9.24), and $a_{min} + p_1 - p_0 > 0$. Now, iff $H_1|\star\rangle = \lambda|\star\rangle$, where λ is the smallest eigenvalue of H_1, then $H_2|\star\rangle = \kappa|\star\rangle$, where κ is the smallest eigenvalue of H_2.*

Proof. Firstly, we have

$$\langle g|H_1|g\rangle = s_0 + p_0$$
$$\langle g|H_{SP}|g\rangle = 0$$

Now, expanding an arbitrary normalized state $|\phi\rangle$ in the eigenbases of H_S and H_P,

$$|\phi\rangle = \sum_{xyij} c_{xyij}|s_{x,i}\rangle|p_{y,j}\rangle$$

we get

$$\langle\phi|H_1|\phi\rangle = \sum_{xyij} |c_{xyij}|^2(s_x + p_y)$$

$$\geq \sum_{xyij} |c_{xyij}|^2(s_0 + p_0) = s_0 + p_0$$

$$\langle\phi|H_{SP}|\phi\rangle = \langle\phi|\sum_{xyij} c_{xyij}A|s_{x,i}\rangle(\mathbf{1}_P - \Pi_{p_0})|p_{y,j}\rangle$$

$$= \langle\phi|\sum_{xij}\sum_{y\geq 1} c_{xyij}a_{x,i}|s_{x,i}\rangle|p_{y,j}\rangle$$

$$= \sum_{xij}\sum_{y\geq 1}|c_{xyij}|^2 a_{x,i}$$

$$\geq a_{\min}\sum_{xij}\sum_{y\geq 1}|c_{xyij}|^2 \geq 0 \quad\text{if}\quad a_{\min} \geq 0.$$

Hence, if A's lowest eigenvalue $a_{\min} \geq 0$ then $|g\rangle$ is the ground state of $H_2 = H_1 + H_{SP}$ as well. If A is not nonnegative, we can perform the transformation

$$H_S' = H_S + a_{\min}\mathbf{1}$$
$$H_P' = H_P - a_{\min}\Pi_{p_0}$$
$$A' = A - a_{\min}\mathbf{1}$$

This leaves H_2 invariant, but makes A' nonnegative. As long as $a_{\min} + p_1 - p_0 > 0$, then $\{|p_{0,m}\rangle\}_m$ still span the ground state subspace of H_P' and the above analysis remains valid. $\qquad\square$

Ground state degeneracy and excitations

If the ground state subspace of $H_{S,T}$ is degenerate and overlaps more than one eigenspace of A, or the simulation register S has excitations to higher energy states at the beginning of the measurement phase, we need a more involved analysis of the measurement procedure. Assume the pre-measurement-phase state of the simulator is given by

$$\rho_0 = \left(\sum_{klmn} c_{klmn}|s_{k,l}\rangle\langle s_{m,n}|\right) \otimes \begin{pmatrix} a & 0 \\ 0 & 1-a \end{pmatrix} \tag{9.28}$$

All three terms in H_2 commute given that $[H_S,\ A] = 0$.

Hence,

$$e^{qH_2} = (e^{qH_S} \otimes \mathbf{1}_P)(\mathbf{1}_S \otimes e^{qH_P})(e^{qH_{SP}})$$

$$= (e^{qH_S} \otimes e^{qH_P})(e^{qA} \otimes (\mathbf{1}_P - \Pi_{p_0}) + \mathbf{1}_S \otimes \Pi_{p_0}) \tag{9.29}$$

As a result of the measurement procedure right before the actual measurement the state is given by

$$
\begin{aligned}
\rho_1(t) \;=\;& \mathsf{H}_p e^{-itH_2/\hbar}\mathsf{H}_p\rho_0\mathsf{H}_p e^{itH_2/\hbar}\mathsf{H}_p \\[4pt]
=\;& \mathsf{H}_p e^{-itH_{SP}/\hbar}\left(\sum_{klmn} c_{klmn}e^{-it(s_k-s_m)/\hbar}|s_{k,l}\rangle\langle s_{m,n}|\right) \\[4pt]
& \otimes\frac{1}{2}\begin{pmatrix} 1 & e^{it\delta/\hbar}(2a-1) \\ e^{-it\delta/\hbar}(2a-1) & 1 \end{pmatrix} e^{itH_{SP}/\hbar}\mathsf{H}_p \\[4pt]
=\;& \sum_{klmn}\left(c_{klmn}e^{-it(s_k-s_m)/\hbar}|s_{k,l}\rangle\langle s_{m,n}|\right) \\[4pt]
& \otimes\frac{1}{2}\mathsf{H}_p\begin{pmatrix} 1 & e^{it\omega_{m,n}}(2a-1) \\ e^{-it\omega_{k,l}}(2a-1) & e^{-it(a_{k,l}-a_{m,n})/\hbar} \end{pmatrix}\mathsf{H}_p,
\end{aligned}
\tag{9.30}
$$

where $\omega_{k,l}=(a_{k,l}+\delta)/\hbar$ and H_p is the Hadamard gate operating on the probe. Projecting the probe to the ground state subspace, we get $=$

$$
\begin{aligned}
& \Pi_{p_0}\rho_1(t)\Pi_{p_0} \\
=\;& \sum_{klmn}\left(c_{klmn}e^{-it(s_k-s_m)/\hbar}|s_{k,l}\rangle\langle s_{m,n}|\right)\otimes\frac{1}{4}(1+e^{-it(a_{k,l}-a_{m,n})/\hbar}+2\cos(\omega_{m,n}t)(2a-1))|0\rangle\langle 0|
\end{aligned}
$$

Thus, the probability of finding the probe in the ground state subspace is given by a superposition of harmonic modes

$$
P_0(t)=Tr(\Pi_{p_0}\rho_1(t)\Pi_{p_0})=\frac{1}{2}\left(1+\sum_{kl}c_{klkl}\cos(\omega_{k,l}t)(2a-1)\right)
\tag{9.31}
$$

Chapter 10

Quantum stochastic walks

"Not all those who wander are lost."

–J. R. R. Tolkien

10.1   Introduction

This chapter is a departure from the simulation theme of the thesis and opts to focus on understanding the interplay between the classical and quantum worlds. In Chapter 3, fundamental topics in quantum information were introduced surrounding the nature of quantum information. Throughout the thesis, manipulation of the quantum state has been considered using unitary transforms where the state remains (or should) remain in a pure state. In this chapter, in lieu of a discussion of error correction, we consider the interplay between quantum and classical stochastic processes. Quantum mechanically, these stochastic processes cause the system to exchange energy (relaxation) or lose phase information (dephasing) which causes the loss of quantum information. Instead of considering arbitrary noise processes, this chapter focuses on quantum stochastic processes related to classical Markov chains.

Many classical algorithms, such as most Markov-chain Monte Carlo algorithms, are based on classical random walks (CRWs), a probabilistic motion through the vertices of a graph. The quantum walk (QW) model is a unitary analog of the CRW that is used to study and develop quantum algorithms [3, 73, 219, 120, 9]. The quantum-mechanical nature of the QW yields different distributions for the position of the walker, as a QW allows for superposition and interference effects. Algorithms based on QWs exhibit that an exponential speedup over their classical counterparts have been developed [52, 226, 51]. Additionally, QWs have inspired the development of an intuitive approach to quantum algorithm design [192], some based on scattering theory [71]. Finally, QWs have recently been shown to be universal for quantum computation [49].

The QW model can be partitioned into two classes: those that have a discrete-time parameter and those with a continuous-time parameter. There are important differences between the discrete and continuous models of QWs. The continuous quantum walks perform unitary evolution on the vertex space of a graph [73]. However, discrete quantum walks

must introduce an auxiliary space [8, 144]. The connection between the discrete-time QW and the continuous-time QW in one dimension was first described using the proper formal limits [203]. The connection was further elucidated by showing that discrete quantum walks could efficiently simulate continuous-time quantum walks by properly controlling simulation errors [50].

Decoherence describes the transition of quantum-mechanical density matrices to classical probability distributions [208, 122, 138, 79, 185]. The transition from the QW into the classical regime for discrete-time QW has been studied by introducing decoherence to specific models of the discrete-time QW [41, 108, 186, 140]. Decoherence has also been been studied as nonunitary effects on continuous-time QW in the context of quantum transport, such as environmentally assisted energy transfer in photosynthetic complexes [183, 150, 172, 43, 184] and quantum state transfer in superconducting qubits [206, 204]. For the purposes of experimental implementation, the vertices of the graph in a walk can be implemented using a qubit per vertex (an inefficient or unary mapping) or by employing a quantum state per vertex (the binary or efficient mapping). The choice of mapping impacts the simulation efficiency and their robustness under decoherence [91, 69, 205]. The previous proposed approaches for exploring decoherence in quantum walks have *added* environmental effects to a QW based on computational or physical models such as pure dephasing [184] but have not considered walks where the environmental effects are constructed axiomatically from the underlying graph. One of the main advantages of the approach considered in this work is that it allows us to study all Markovian decoherence mechanisms on the same footing.

In this chapter, we define the continuous-time quantum stochastic walk (QSW) using a set of axioms that incorporate unitary and nonunitary effects. For the remainder of the chapter, we focus on the continuous-time QW but will sketch the corresponding procedure for the discrete-time QW. A CRW is a type of classical stochastic processes. From the point of view of the theory of open quantum systems, the generalization of a classical stochastic process to the quantum regime is known to be a quantum stochastic process [208, 122, 138, 79, 185, 150] which is the most general type of evolution of a density matrix, and not simply the Hamiltonian process proposed by the QW approach. The main goal of this article is to introduce a set of axioms that allow for the construction of a quantum stochastic process constrained by a graph. We call all the walks that follow these axioms QSWs. We will show that the family of QSWs includes both the CRW and the QW as limiting cases. The QSW can yield new distributions that are not found either in the CRW or the QW. The connection between the three types of walks discussed in this chapter is summarized in Fig. 10.1. The QSW provides the fundamental tools to study the quantum-to-classical transition of walks, as well as new methods of control by the application of nonunitary operations on a quantum system.

10.2   Axiomatic quantum walks

In this section, we describe an axiomatic approach to constructing stochastic dynamics that correspond to the classical random walk (CRW) and then generalize the stochastic process to the quantum case leading to what we refer to as the quantum stochastic walk (QSW). The CRW describes probabilistic motion over a graph. The dynamics of the proba-
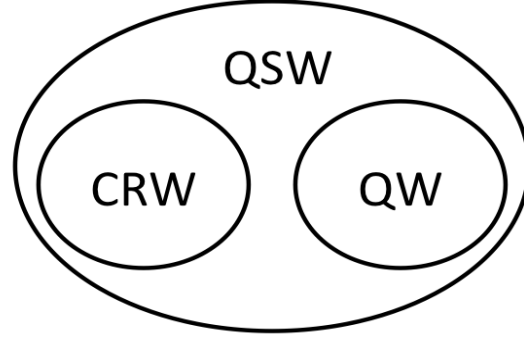
Figure 10.1: **QSW is a quantum stochastic process** defined by a set of axioms that describe the connectivity of an underlying graph. The family of possible QSWs encompasses the CRW and the QW as limiting cases. QSWs can be used to generate walks that continuously interpolate between QW and CRW and, therefore, can be employed to analyze algorithmic classical-to-quantum transitions. The family of QSWs also describes more general quantum stochastic processes with a behavior that cannot be recovered by the QW and CRW approaches.

bility distribution function is given by the master equation for a classical stochastic process

$$\frac{d}{dt}p_a = \sum_b M_b^a \, p_b \tag{10.1}$$

where the vector element p_b is the probability of being found at the vertex b of the graph. The matrix M is the generator for the evolution; its structure is constrained by axioms derived from the connectivity of the graph. For example, if vertices a and b are connected, $M_b^a = -\gamma$; if they are not, $M_b^a = 0$, and $M_a^a = d_a\gamma$, where d_a is the degree of vertex a.

In analogy to the CRW, the QW [73] has been defined so that the probability vector element p_a is replaced with $\langle a|\psi\rangle$, which evolves according to the Schrödinger equation

$$\frac{d}{dt}\langle a|\psi\rangle = -i\sum_b \langle a|H|b\rangle\langle b|\psi\rangle \tag{10.2}$$

where H is a Hamiltonian to be defined based on axioms coming from the graph. A choice of this paragraph is $\langle a|H|b\rangle = M_b^a$. This unitary evolution effectively rotates populations into coherences and back.[1] The QW fails to completely capture the stochastic nature of CRW. The random aspect of the QW comes solely from the probabilistic nature of measurements performed on the wave function.

---

[1]The Hermitian form of the Hamiltonian eliminates directed graphs from consideration, motivating work on pseudo-Hermitian Hamiltonians [187].

Since a classical stochastic process can be generalized to the quantum regime by means of a quantum stochastic process, a CRW should be generalized to a QSW derived from the graph.[2] For the generalization, we identify the probability vector with elements p_a with a density matrix with elements $\rho_{a\alpha}$ and generalize the evolution to a quantum stochastic process, $\frac{d}{dt}\rho = \mathcal{M}[\rho]$, where $\mathcal{M}$ is a superoperator [208, 122, 79]. To make this evolution look similar to Eq. (10.1), we write the density matrix in terms of its indices, $\rho = \sum_{a,\alpha} \rho_{a\alpha} |a\rangle\langle\alpha|$, and the quantum stochastic master equation becomes

$$\frac{d}{dt}\rho_{a\alpha} = \sum_{b,\beta} \mathcal{M}^{a\alpha}_{b\beta}\ \rho_{b\beta} \tag{10.3}$$

with the tensor $\mathcal{M}^{a\alpha}_{b\beta} = \langle a|\, \mathcal{M}\big[\,|b\rangle\langle\beta|\,\big]\,|\alpha\rangle$. This correspondence was first pointed out by Mohseni *et al.* in the context of energy transfer [150].

For a quantum stochastic process to be related to a walk, the superoperator $\mathcal{M}$ must be related to the underlying graph. The connectivity of the vertices will impose conditions on the transition rates of $\mathcal{M}$. Since the quantum stochastic process is more general than both the classical stochastic process and the Schrödinger equation, the correspondence of the connectivity of the graph to the rules imposed on $\mathcal{M}$ should include and go beyond the connectivity axioms for each of those. For a vertex m connected to vertices that include vertex n, we define processes $|m\rangle \leftrightarrows |n\rangle$ which occur at some rate that can evolve $|m\rangle$ to and from $|n\rangle$. Transition rates for vertices that are not connected are defined to be zero. We employ these connectivity rules as the main principle for defining valid QSWs from a given graph. To further explore the connection from the QSW to the CRW and QW as well as more general behaviors, we discuss the different limiting cases.

For the classical case, the allowed transitions come from incoherent population hopping of the form $|m\rangle\langle m| \leftrightarrows |n\rangle\langle n|$ and, for completeness, $|m\rangle\langle m| \leftrightarrows |m\rangle\langle m|$. These conditions constrain $\mathcal{M}$ to operate only on the diagonal elements of ρ, as in Eq. (10.1). In other words, the QSW transition tensor must have the property $\sum_{\alpha,\beta} \delta_{a\alpha} \mathcal{M}^{a\alpha}_{b\beta} \delta_{b\beta} = M^a_b$.

The QSW should also recover the QW, where evolution among vertices happens through coherences developed by a Hamiltonian. These transitions include terms of the form, $|m\rangle\langle m| \leftrightarrows |m\rangle\langle n|$, that exchange populations to coherences with the connected vertices. For completeness, we also consider transitions of the form $|m\rangle\langle n| \leftrightarrows |m\rangle\langle n|$. If m is also connected to another vertex l, additional transitions can happen. One transition exchanges populations into coherences among the two vertices connected to m, $|m\rangle\langle m| \leftrightarrows |l\rangle\langle n|$; the other exchanges coherences between m and a connected vertex into coherence among two vertices connected to m, $|m\rangle\langle n| \leftrightarrows |l\rangle\langle n|$. These conditions allow for the recovery of Eq. (10.2). By including the conjugates of these, we have now exhausted all the possibilities of the transitions that can happen following the connectivity of vertex m. These rules can be applied to vertices with any number of connections, and serve as the basis for the correspondence of the graph to the QSW.

To complete the generalization, we need an equation of motion for a quantum stochastic process. We choose the Kossakowski-Lindblad master equation [122, 138, 79],

---

[2]Just like a classical stochastic process includes, but is not limited, to bistochastic processes, a quantum stochastic processes includes but is not limited to unitary processes [208].

Matrix elements' connectivity	Transition rate
$\lvert m\rangle\langle m\rvert \leftrightarrows \lvert m\rangle\langle m\rvert$	$\mathcal{L}^{mm}_{mm} = \langle m\lvert L_k\rvert m\rangle\langle m\lvert L_k^\dagger\rvert m\rangle - \langle m\lvert L_k^\dagger L_k\rvert m\rangle$
$\lvert m\rangle\langle m\rvert \leftrightarrows \lvert n\rangle\langle n\rvert$	$\mathcal{L}^{nn}_{mm} = \langle n\lvert L_k\rvert m\rangle\langle m\lvert L_k^\dagger\rvert n\rangle$
$\lvert m\rangle\langle m\rvert \leftrightarrows \lvert m\rangle\langle n\rvert$	$\mathcal{L}^{mn}_{mm} = \langle m\lvert L_k\rvert m\rangle\langle m\lvert L_k^\dagger\rvert n\rangle + i\langle m\lvert H\rvert n\rangle - \frac{1}{2}\langle m\lvert L_k^\dagger L_k\rvert n\rangle$
$\lvert m\rangle\langle n\rvert \leftrightarrows \lvert m\rangle\langle n\rvert$	$\mathcal{L}^{mn}_{mn} = \begin{aligned}&\langle m\lvert L_k\rvert m\rangle\langle n\lvert L_k^\dagger\rvert n\rangle - i\langle m\lvert H\rvert m\rangle + i\langle n\lvert H\rvert n\rangle\\&-\tfrac{1}{2}\langle m\lvert L_k^\dagger L_k\rvert m\rangle - \tfrac{1}{2}\langle n\lvert L_k^\dagger L_k\rvert n\rangle\end{aligned}$
$\lvert m\rangle\langle n\rvert \leftrightarrows \lvert l\rangle\langle n\rvert$	$\mathcal{L}^{ln}_{mn} = \langle l\lvert L_k\rvert m\rangle\langle n\lvert L_k^\dagger\rvert n\rangle - i\langle l\lvert H\rvert m\rangle - \frac{1}{2}\langle l\lvert L_k^\dagger L_k\rvert m\rangle$
$\lvert m\rangle\langle m\rvert \leftrightarrows \lvert l\rangle\langle n\rvert$	$\mathcal{L}^{ln}_{mm} = \langle l\lvert L_k\rvert m\rangle\langle m\lvert L_k^\dagger\rvert n\rangle$

Table 10.1: **Axioms for the quantum stochastic walk for processes that connect vertex $\lvert m\rangle$ to its neighbors.** Sum over k, and conjugate elements are implied. Axioms in the first and second rows correspond to the classical random walk. Axioms in the third, fourth, and fifth rows contain the quantum walk using H and additional evolution due to $\{L_k\}$ terms. The axiom in the sixth column comes only from the quantum stochastic walk, having no equivalent in the classical random walk or quantum walk.

$\frac{d}{dt}\rho = \mathcal{L}[\rho] = \sum_k -i[H,\rho] - \frac{1}{2}L_k^\dagger L_k \rho - \frac{1}{2}\rho L_k^\dagger L_k + L_k \rho L_k^\dagger$, which evolves the density matrix as a quantum stochastic process inspired by environmental effects under the Markov approximation. The Hamiltonian term describes the coherent evolution (Schrödinger equation) while the rest describe a stochastic evolution. We now set $\mathcal{M} \to \mathcal{L}$ in Eq. (10.3), where

$$\mathcal{L}^{a\alpha}_{b\beta} = \sum_k \delta_{\alpha\beta}\langle a\rvert\left(-iH - \frac{1}{2}L_k^\dagger L_k\right)\lvert b\rangle \qquad (10.4)$$
$$+ \delta_{ab}\langle\beta\rvert\left(iH - \frac{1}{2}L_k^\dagger L_k\right)\lvert\alpha\rangle + \langle a\rvert L_k\rvert b\rangle\langle\beta\rvert L_k^\dagger\rvert\alpha\rangle.$$

The connectivity conditions between a vertex m and some connected vertices n and l and the corresponding non-zero transition rates according to Eq. (10.4) can be summarized in Table 10.1.[3]

The axioms from Table 10.1 capture the behavior sketched in Fig. 10.1. To recover the CRW, it suffices to consider axioms in rows 1 and 2. These axioms are a classical subset of the transition rates of the QSW. However, the QW is obtained by making all the rates of each element of $\{L_k\}$ zero. In this case, only the subset of axioms from rows 3, 4, and 5 from Table 10.1 are relevant, corresponding to the Hamiltonian, and the QW is recovered. If the rates of $\{L_k\}$ are nonzero, these axioms contain behavior beyond the QW. Finally, the sixth row has properties that have no equivalent in either the CRW or the QW; it is a type of transition that appears exclusively in the QSW and leads to different distributions. The choice of H or $\{L_k\}$ is not uniquely determined by the axioms.

---

[3]For compactness, we assume that all connections are equally weighted; the general case can be obtained by having coefficients that depend on the vertices.

For example, a CRW is equivalent to a QSW with the choice of each element of $\{L_k\}$ to be associated to M in the following manner: First, since connections are defined between two vertices, it is easier to write the index k in terms of two indices $k \equiv (\kappa, \kappa')$. Using this notation, we enforce the graph by choosing $L_{(\kappa,\kappa')} = M_\kappa^{\kappa'} |\kappa\rangle\langle\kappa'|$. The Hamiltonian rates of H are set to zero and it is easy to show that coherences vanish exponentially fast. This choice of $\{L_k\}$ ensures all the transition rates are zero except the axiom in the second row of Table 10.1, thereby recovering the CRW. An interpolation between the CRW and the QW conditions can be used to obtain the classical-to-quantum transition of any graph by changing the relative rates of the unitary and nonunitary processes.

10.3 Interplay between different kinds of walks

The classical to quantum transition has been extensively studied in the discrete time case but the study of the same transition for continuous time quantum walks has not received equal attention [108]. The use of dephasing and relaxation has been motivated by physical models [183, 109]. However, we are interested in a unifying mathematical framework for classical and quantum transition using the Kossakowski-Lindblad master equation. Here, we hope to provide another view of the transition between the classical and quantum regimes motivated by our axiomatic approach to quantum walks.

Our approach differs from other approaches in the literature in that we do not limit our attention to dephasing and relaxation as the sole decoherence mechanisms. For example, the mechanism considered in Ref. [108] was a dephasing model that causes off-diagonal coherences of the density operator to decay, thereby making the density operator into a classical probability vector. However, the classical probability vector produced is not the probability vector obtained with the CRW. Although the quantum aspects of the evolution are no longer present, the resulting probability vector does not equal the CRW for any value of the decoherence. In Fig. 10.2, we have plotted the Euclidean distance between between the classical and decoherent quantum evolution as a function of the dephasing strength.

The QSW provides a general framework to study the *direct* transition between the CRW and the QW. Using set $\{L_{(\kappa,\kappa')} = M_\kappa^{\kappa'} |\kappa\rangle\langle\kappa'|\}$ we obtain the classical evolution using only the environment i.e. no Hamiltonian evolution. This quantum to classical transition is different from the pure dephasing model as the Hamiltonian is not required to recover the classical dynamics. The idea of allowing the environment to generate dynamics is similar to the recent ideas of quantum state engineering driven by dissipation [220, 123, 65]. Now, using a single parameter $\omega \in [0, 1]$, we can interpolate between the QW and the CRW (or any other environmentally driven dynamics):

$$\mathcal{L}_\omega[\rho] = \quad (\omega - 1)i\,[H, \rho]$$
$$+\omega \sum_k \left(L_k \rho L_k^\dagger - \frac{1}{2}[L_k^\dagger L_k, \rho]_+ \right) \tag{10.5}$$

To illustrate this interpolation, we study two examples. First, we consider the walk on an infinite line, then we consider glued tree of depth three [52, 51]. The glued tree
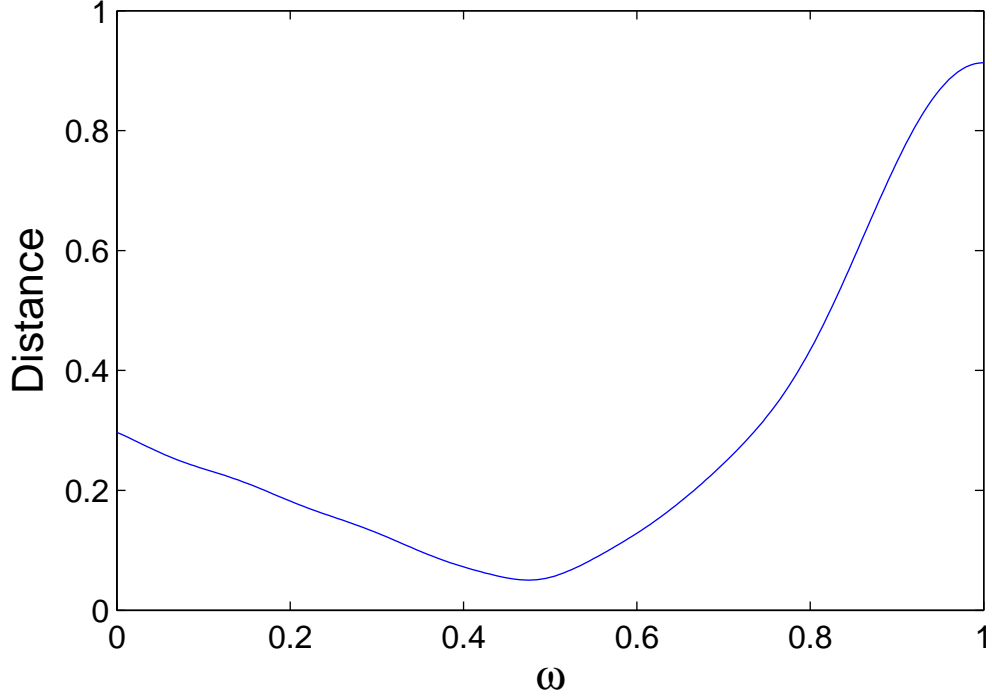
Figure 10.2: **The distance between the classical and purely dephased quantum walk** on the line is plotted as a function of the dephasing strength ω at fixed time $t = 5$ (ω and t in atomic units). The Euclidean distance metric between two probability mass functions, $p^{(1)}$, $p^{(2)}$ is defined by $d(p^{(1)}, p^{(2)}) = \sqrt{\sum_i (p_i^{(1)} - p_i^{(2)})^2}$ and the abscissa parameter, ω, is defined in Eq. (10.5) with the set $\{L_k\} = \{|k\rangle\langle k|\}$ as the pure dephasing operators. It is emphasized that pure dephasing does not lead to the classical random walk.

is an interesting case for the QSW as the QW exhibits an exponential transport speedup over the CRW and the QSW allows us to observe the change from classical to quantum transport.

10.3.1 Infinite line

The walk on an infinite line is defined for a graph with each vertex j being connected to its nearest neighbors. The conditions for the dynamics of the walk on the line can be obtained from Table 10.1 by making $j = m$, $j - 1 = l$ and $j + 1 = n$. In Fig. 10.3a, we consider the quantum to classical transition for the walk on a line at a fixed time of $t = 5$.

To obtain behavior that is different from both the CRW and the QW, we specify a set $\{L_k\}$ with only one member that is $L = M = \sum_{\kappa,\kappa'} M_\kappa^{\kappa'} |\kappa\rangle\langle\kappa'|$. This choice makes the sixth axiom nonzero, a type of transition that cannot be interpreted either as the CRW or the QW. We illustrate the resulting distribution, and compare it to CRW and the QW, in Fig. 10.3. Although the first two moments of the distribution generated using $L = M$ agree with the CRW, coherences are maintained indicating that not all coherent walks spread

with super linear variances.

10.3.2 Glued trees

The set of glued trees $\{G_n\}$ are a series of graphs originally defined in [52] and by embedding the graph in an oracle, the problem can be shown to provide an exponentially separation between all classical algorithms (including the CRW) and QW [51]. The first three graphs of the glued tree series are shown in Fig. 10.6. A classical walker moves quickly toward the center of the graph, but takes time exponential in the depth of the tree to find the rightmost vertex. In the case of quantum propagation, the graph can be reduced to a walk on a line [52, 106], and, thus, the quantum propagation remains a polynomial function of the depth of the graph. We studied the effect of decoherence on this graph for continuous time propagation, to be contrasted to the discrete propagation study in [109]. We have plotted the probability of finding the walker at the final state at time t for a glued tree of depth three in Fig. 10.4a. The probability distribution of a quantum walk never converges but instead gives oscillatory behavior due to the wave-like nature of quantum evolution [3]. However, the time averaged probability, $\bar{P}(T) = (1/T) \int_0^T dt P(t) = (1/T) \sum_t P(t)$, converges and is plotted in Fig. 10.4b. Notice that the transport properties of mixed quantum-classical walks is worst for small values of the interpolation parameter, ω, see Fig. 10.5.

This is in contrast to the results of studies of transport through biological systems where the couplings between sites are irregular where small amounts of decoherence helps satisfy certain resonance conditions allowing faster transport [183, 150, 172, 43, 184]. The regularity of the glued tree is responsible for rapid transport in the quantum case [52] and here weak decoherence retards the transport below that of the CRW. It will be interesting to investigate if this behavior will be found in other regular graphs which demonstrate rapid quantum transport such as the hypercube.

10.4 Discussion and conclusion

Although in this article we focused on continuous-time walks, parallel arguments holds for discrete-time walks. A discrete-time CRW evolves each time step by a Markov chain with a stochastic matrix $\mathcal{S}$ following $p_a = \sum_b \mathcal{S}_b^a p_b$. The quantum analogue is to use a quantum stochastic map $\mathcal{B}$ [208] by means of $\rho'_{a\alpha} = \sum_{b\beta} \mathcal{B}_{b\beta}^{a\alpha} \rho_{b\beta}$, or equivalently as a superoperator, $\rho' = \mathcal{B}[\rho] \equiv \sum_k \mathcal{C}_k \rho \mathcal{C}_k^\dagger$. A similar set of axioms to Table 10.1 can be computed by using $\mathcal{B}_{b\beta}^{a\alpha} = \sum_k \langle a | \mathcal{C}_k | b \rangle \langle \beta | \mathcal{C}_k^\dagger | \alpha \rangle$.[4] The connection between the discrete-time QW and the continuous-time QW has been studied in Refs. [203, 50].

In conclusion, we introduced an axiomatic approach to define the quantum stochastic walk, which behaves according to a quantum stochastic process as constrained by a particular graph. This walk recovers the classical random walk and the quantum walk as special cases, as well as the transition between them. The quantum stochastic walk allows

---

[4]There is no need for an explicit coin, as the stochastic map is the most general evolution of a density matrix, including random processes. However, if desired, a coin could be implemented as an environment from the open quantum system interpretation of the map.

for the construction of new types of walks that result in different probability distributions. As an example, we studied the walk on a line and glued tree. The quantum stochastic walk provides a general framework to study quantum walks under decoherence.[5] Since the quantum stochastic walk contains the quantum walk, it is therefore also universal for quantum computation. Its quantum to classical transition can also be used to examine decoherence phenomena in quantum computation and classical-to-quantum phase transitions. New quantum stochastic distributions might suggest new kinds of quantum algorithms or even new classes of quantum algorithms based on this model.

---

[5]Reexamination of previous work that considered environmental effects from physical motivations might suggest that, if interpreted as a quantum stochastic walk, the environment is changing the effective graph.
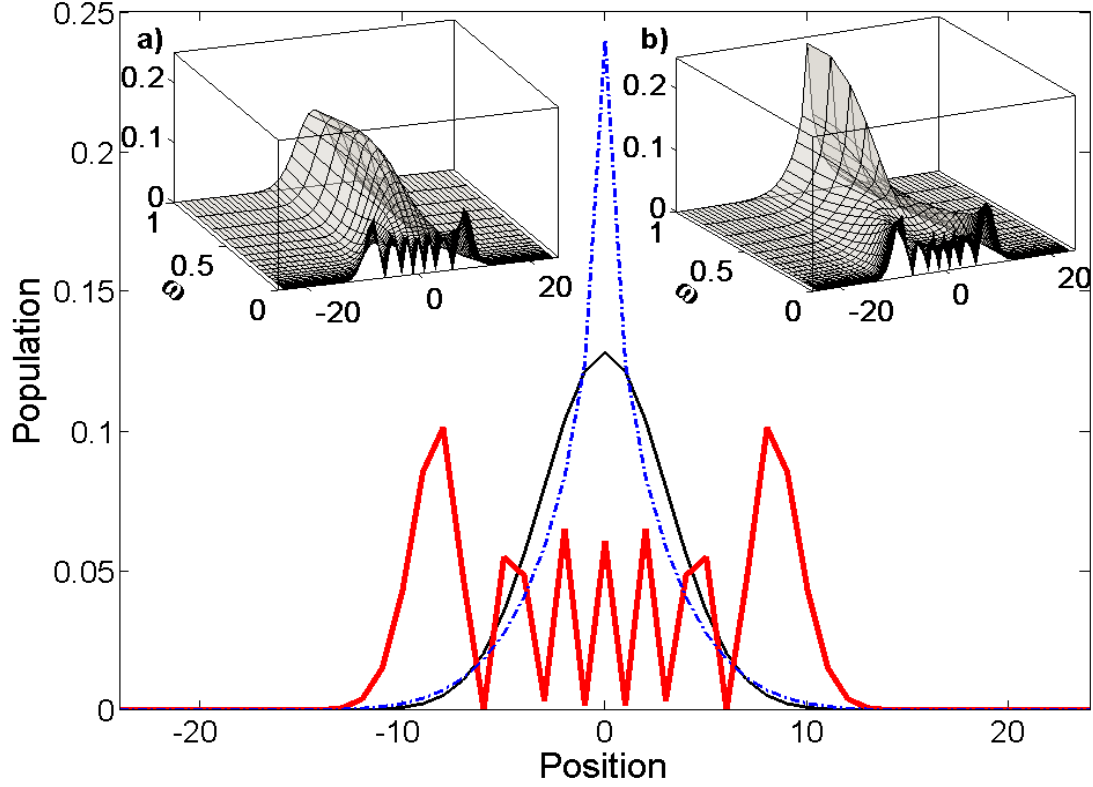
Figure 10.3: **Probability distributions of a quantum walker** starting at position 0 of an infinite line defined by M are shown at time $t = 5$ (atomic units). The heavy line (red) represents the quantum walk, the thin line (black) represents the classical random walk, and the dotted line (blue) represents a choice of a quantum stochastic walk due to a single operator $L = \sum_{\kappa,\kappa'} M_\kappa^{\kappa'} |\kappa\rangle\langle\kappa'|$ whose distribution cannot be obtained from the classical or quantum walks. **a)** The transition of the quantum walk to the classical random walk is shown. The diagonal elements of the density matrix (Population) are plotted as a function of the vertex of the walker (Position) and a coupling parameter ω. When $\omega = 0$, the evolution is due only to the Hamiltonian, and when $\omega = 1$, the Hamiltonian dynamics are not present with the environmental dynamics responsible for dynamics. **b)** The transition of the quantum walk to the quantum stochastic walk is shown.
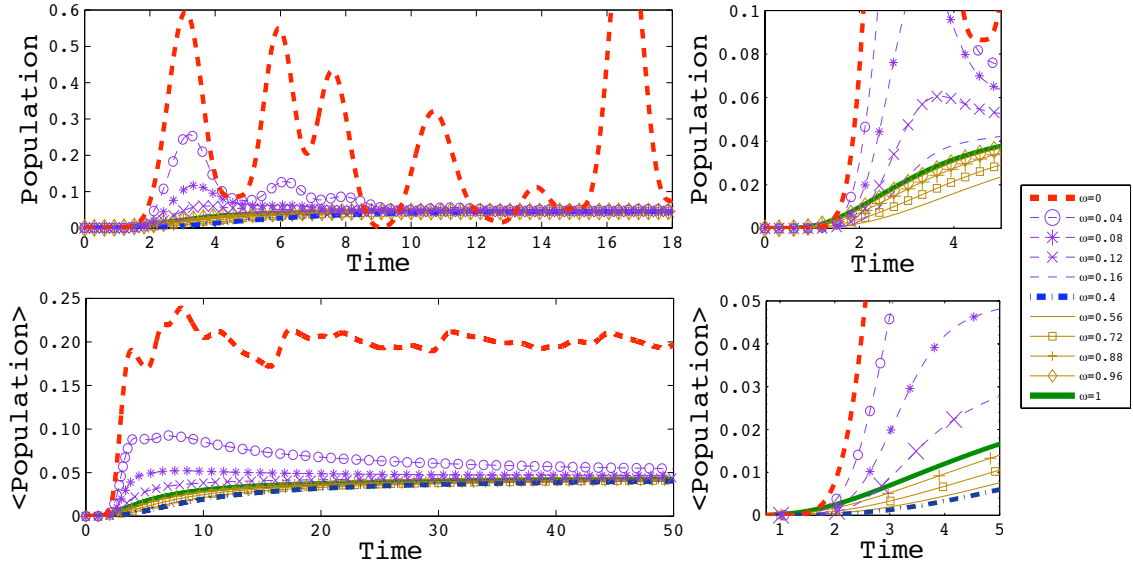
Figure 10.4: **Transport probability through glued tree** of depth three as a function of time (atomic units) and classical-quantum mixing parameter ω. In both figures, the line with $\omega = 1$ corresponds to the classical random walk (CRW), the line with $\omega = 0$ corresponds to the quantum walk (QW). The lines with $\omega = (0.16, 0.12, 0.08, 0.04)$ correspond to mixed quantum-classical walks that have performance worse than the CRW, the lines with $\omega = (0.96, 0.88, 0.72, 0.56)$ are the mixed classical-quantum walks that perform better than the CRW, and the $\omega = 0.4$ line is the worst performing mixed walk. **a)** The probability of being at the exit site as a function of time is shown for various mixing parameters of the classical and quantum walks. **b)** The time average of the probability of being at the exit site. For quantum walks, no stationary distribution exists due to unitary evolution, however, the time average of the probability distribution of a quantum walk converges [3].
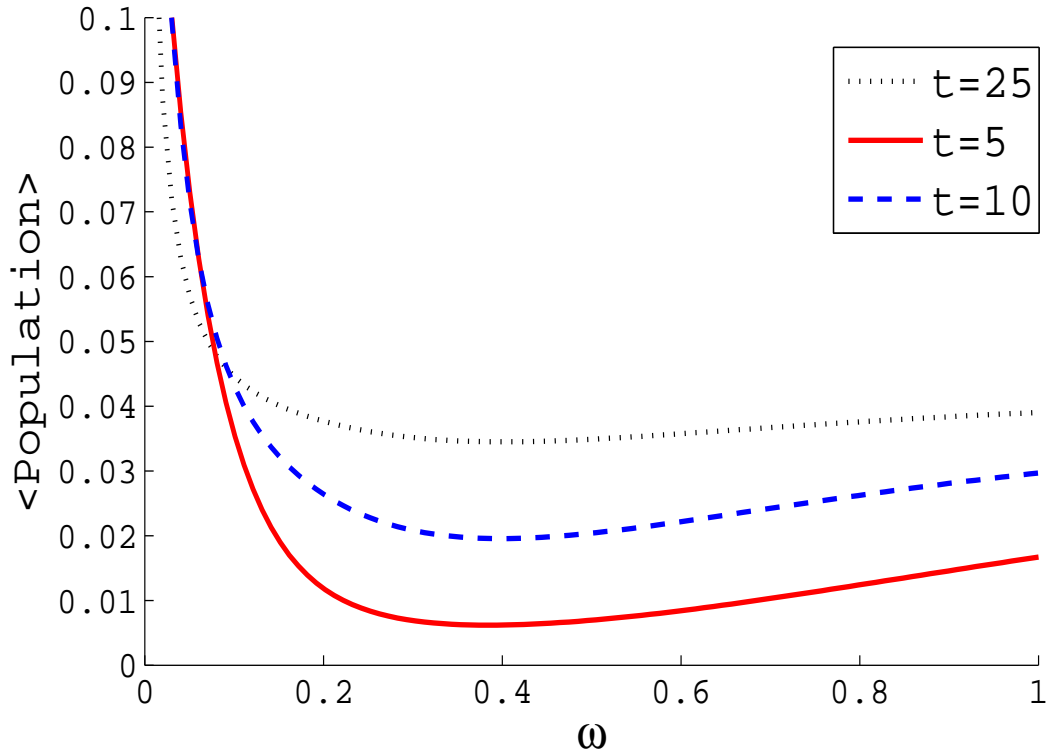
Figure 10.5: **Plot of the time averaged population for transport through the** G_3 **tree** at three fixed times, $t = \{5, 10, 25\}$, as a function of the decoherence strength $\omega \in [0, 1]$ (atomic units). When $\omega = 1$ the classical walk is obtained, and when $\omega = 0$, the unitary walk is recovered. Notice that the probability is not a strictly decreasing function of the decoherence.
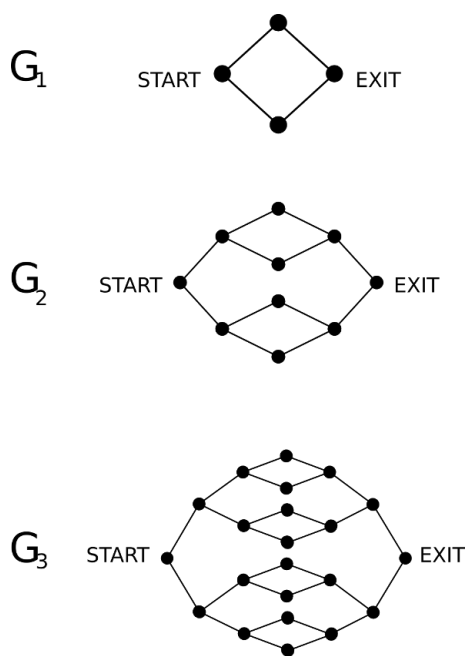
Figure 10.6: **Examples of glued trees of depth one, two and three.** The glued tree of depth n consists of two binary trees of depth n connected at the leaves.

Chapter 11

Future directions

"Imagine that you are explaining your ideas to your former smart, but ignorant, self at the beginning of your studies"

–R. P. Feynman, On writing theses

Quantum computation is an emerging field with researchers encompassing physics, engineering, and computer science and, since circa 2005, chemistry. The goal of building a quantum computer including which architecture, circumventing decoherence, and experimental controls have been pursued by experimentalists worldwide. The algorithms and applications of quantum computation was recognized nearly immediately once quantum mechanics was applied to information theory. The applications to cryptography and basic research into computation has also flourished.

It was my goal to extend the early work [15] on quantum computation for chemistry into a research program that has detailed the necessary steps, helped construct the first implementation and explored alternative ways to parameterize the evolution. This is only the beginning. Subsequent studies, both theoretically and experimentally, are needed to understand systematic errors, ideal areas of applications and practical adaptations to experimental hardware. These outstanding tasks are the future of the field. Experimental groups across America, China, Australia and Europe have been working to realize the next generation of quantum chemistry calculations, but what remains is a systematic simulation of many molecules to calibrate expectations and get a sense of when a quantum computer helps and when it does not.

Other than furthering the model chemistry on a quantum computer, there are other reasons to explore quantum computation and quantum information in the context of chemistry. Quantum mechanics is a profound theory that can, at times, be extremely difficult or counter-intuitive. However, chemists have been exceedingly pedantic and pragmatic about their approach to quantum mechanics. Arrow pushing, steric hindrance and molecular orbitals are standard fare for sophomores and freshmen in chemistry and the lifeline for predicting mechanisms and likely reactions. These conceptual tools lie at the surface layer of quantum mechanics, while largely respecting the underlying physics. Perhaps one of the goals of studying computational chemistry and quantum information, should be to provide a richer set of tools at this interface.

Bibliography

[1] D. Abrams and S. Lloyd. Simulation of many-body Fermi systems on a universal quantum computer. *Phys. Rev. Lett.*, 79:2586, 1997. [62, 83]

[2] D. Abrams and S. Lloyd. Quantum algorithm providing exponential speed increase for finding eigenvalues and eigenvectors. *Phys. Rev. Lett.*, 83(24):5162–5165, 13 December 1999. [62]

[3] D. Aharonov, A. Ambainis, J. Kempe, and U. Vazirani. Quantum walks on graphs. In *Proc. 33rd ACM Symposium on Theory of Computing*, page 50, 2001. [115, 122, 125]

[4] D. Aharonov and O. Regev. A lattice problem in quantum NP. *arxiv:quant-ph/0307330*, 2003. [59]

[5] D. Aharonov and A. Ta-Shma. Adiabatic quantum state generation. *SIAM J. Comput.*, 37:47–82, 2008. Also see arxiv:quant-ph/0301023. [73, 85]

[6] D. Aharonov, W. van Dam, J. Kempe, Z. Landau, S. Lloyd, and O. Regev. Adiabatic quantum computation is equivalent to standard quantum computation. *SIAM Journal of Computing*, 37:166, 2007. [42, 109]

[7] Y. Aharonov and D. Bohm. Time in the quantum theory and the uncertainty relation for time and energy. *Phys. Rev.*, 22:1649, 1961. [56]

[8] Y Aharonov, L. Davidovich, and N. Zagury. Quantum random walks. *Phys. Rev. A*, 48:1687, 1993. [116]

[9] A. Ambainis. Quantum random walks – new method for designing quantum algorithms. *Lecture Notes in Computer Science*, 4910:1–4, 2008. [115]

[10] L. Amico, R. Fazio, A. Osterloh, and V. Vedral. Entanglement in many-body systems. *Rev. Mod. Phys.*, 80:517, 2008. [41]

[11] M. H. S. Amin. Consistency of the adiabatic theorem. *Phys. Rev. Lett.*, 102:220401, 2009. [44, 84]

[12] M. H. S. Amin, D. V. Averin, and J. A. Nesteroff. Decoherence in adiabatic quantum computation. *Phys. Rev. A*, 79:022107, 2009. [106, 108]

[13] M. H. S. Amin, P. J. Love, and C. J. S. Truncik. Thermally assisted adiabatic quantum computation. *Phys. Rev. Lett.*, 100:060503, 2008. [103]

[14] S. Arora and B. Barak. *Computational Complexity: A Modern Approach*. Cambridge University Press, 2009. [48]

[15] A. Aspuru-Guzik, A. D. Dutoi, P. Love, and M. Head-Gordon. Simulated quantum computation of molecular energies. *Science*, 309:1704, 2005. [1, 62, 63, 84, 85, 91, 128]

[16] D. Bacon and S. T. Flammia. Adiabatic gate teleportation. *Phys. Rev. Lett.*, 103(12):120504, Sep 2009. [108]

[17] F Barahona. On the computational complexity of Ising spin glass models. *J. Phys. A: Math. Gen.*, 15:3241, 1982. [57]

[18] A. Barenco, C. H. Bennett, R. Cleve, D. DiVincenzo, N. Margolus, P. Shor, T. Sleator, J. Smolin, and H. Weinfurter. Elementary gates of quantum computation. *Phys. Rev. A*, 5(52):3457, 1995. [42]

[19] J. Baugh, J. Chamilliard, C. M Chandrashekar, M. Ditty, A. Hubbard, R. Laflamme, M. Laforest, D. Maslov, O. Moussa, C. Negrevergne, M. Silva, S. Simmons, C. A Ryan, D. G Cory, J. S Hodges, and C. Ramanathan. Quantum information processing using nuclear and electron magnetic resonance: Review and prospects. *Physics in Canada*, 63:197–206, 2007. [45]

[20] R. Beals, H. Buhrman, R. Cleve, M. Mosca, and R. de Wolf. Quantum lower bounds by polynomials. In *Proceedings of FOCS' 98*, pages 352–361, 1998. [74]

[21] G. Benenti and G. Strini. Quantum simulation of the single-particle Schrödinger equation. *Am. J. Phys.*, 76:657–662, 2008. [75]

[22] C. H. Bennett. Logical reversibility of computation. *IBM J. Res. Develop.*, 17, 1973. [42, 158]

[23] C. H. Bennett. Notes on Landauer's principle, reversible computation and Maxwell's demon. *Studies Hist. Phil. Mod. Phys.*, 34:501–510, 2003. [37]

[24] C. H. Bennett, E. Bernstein, G. Brassard, and U. Vazirani. Strengths and weaknesses of quantum computing. *SIAM J. Computing*, 26(5):1510–1524, 1997. [161]

[25] A. J. Berkley, M. W. Johnson, P. Bunyk, R. Harris, J. Johansson, T. Lanting, E. Ladizinsky, E. Tolkacheva, M. H. S Amin, and G. Rose. A scalable readout system for a superconducting adiabatic quantum optimization system. *Supercond. Sci. Technol.*, 23:105014, 2010. [45]

[26] E. Bernstein and U. Vazirani. Quantum complexity theory. *SIAM J. Computing*, 26(5):1411–1473, 1997. [51]

[27] D. W. Berry, G. Ahokas, R. Cleve, and B. C. Sanders. Efficient quantum algorithms for simulating sparse Hamiltonians. *Commun. Math. Phys.*, 270:359–371, 2007. [72, 73, 74]

[28] D. W. Berry and A. M. Childs. Black-box Hamiltonian simulations and unitary implementation. *arXiv:0910.4157*, 2009. [73]

[29] J. D. Biamonte. Nonperturbative k-body to two-body commuting conversion Hamiltonians and embedding problem instances into Ising spins. *Phys. Rev. A*, 77(5):052331, 2008. [103, 109]

[30] J. D. Biamonte and P. J. Love. Realizable Hamiltonians for universal adiabatic quantum computers. *Phys. Rev. A*, 78:012352, 2008. [103, 108]

[31] R. Blatt and D. Wineland. Entangled states of trapped atomic ions. *Nature*, 453:1008, 2008. [46]

[32] S. Boixo, E. Knill, and R. D. Somma. Quantum state preparation by phase randomization. *Quant. Inf. Comp.*, 9:0833, 2009. [85]

[33] S. Boixo and R. D. Somma. Necessary condition for quantum adiabatic approximation. *Phys. Rev. A*, 81:032308, 2010. [44, 84]

[34] G. Brassard, P. Hoyer, M. Mosca, and A. Tapp. Quantum amplitude amplification and estimation. *arxiv:quant-ph/0005055*, 2000. [65]

[35] S. L. Braunstein, C. M. Caves, R. Jozsa, N. Linden, S. Popescu, and R. Schack. Separability of very noisy mixed states and implications for NMR quantum computing. *Phys. Rev. Lett.*, 83:1054–1057, 1999. [93]

[36] S. Bravyi, D. P. DiVincenzo, D. Loss, and B. M. Terhal. Quantum simulation of many-body Hamiltonians using perturbation theory with bounded-strength interactions. *Phys. Rev. Lett.*, 101(7):070503, 2008. [104]

[37] H.-P. Breuer and F. Petruccione. *The theory of open quantum systems*. Oxford University Press, New York 2002. [107, 111]

[38] B. Brown, S. T. Flammia, and N. Schuch. Computational difficulty of computing the density of states. *Phys. Rev. Lett.*, 107:040501, 2011. [61]

[39] K. L. Brown, W. J. Munro, and V. M. Kendon. Using quantum computers for quantum simulation. *Entropy*, 12:2268, 2010. [44]

[40] K. R. Brown, R. J. Clark, and I. L. Chuang. Limitations of quantum simulation examined by simulating a pairing Hamiltonian using nuclear magnetic resonance. *Phys. Rev. Lett.*, 97:050504, 2006. [45, 64, 69, 91, 110]

[41] T. Brun, H. Carteret, and A. Ambainis. Quantum to classical transition for random walks. *Phys. Rev. Lett.*, 91:130602, 2003. [116]

[42] I. Buluta and F. Nori. Quantum simulators. *Science*, 326:108, 2009. [44]

[43] F. Caruso, A. Chin, A. Datta, S. Huelga, and M. Plenio. Fundamental mechanisms of noise supported energy transfer in biological systems. *arXiv:0901.4454*, 2009. [116, 122]

[44] G. K-L. Chan, J. J. Dorando, D. Ghosh, J. Hachmann, E. Neuscamman, H. Wang, and T. Yanai. An introduction to the density matrix renormalization group ansatz in quantum chemistry. *Prog. Theor. Chem. and Phys.*, 18:49, 2008. [2, 41]

[45] J. R. Chelikowsky, Y. Saad, T.-L. Chan, M. L. Tiago, A. T. Zayak, and Y. Zhou. Pseudopotentials on grids: Application to the electronic, optical, and vibrational properties of silicon nanocrystals. *J. Comput. Theor. Nanosci.*, 6:1247–1261(15), 2009. [1]

[46] C.F. Chiang and P. Wocjan. Quantum algorithm for preparing thermal Gibbs states-detailed analysis. *NATO Sci. for Peace and Secuity: Info. and Comm. Security*, 26:138–147, 2010. Also arXiv:1001.1130. [86]

[47] J. Chiaverini, J. Britton, D. Leibfried, E. Knill, M. D. Barrett, R. B. Blakestad, W. M. Itano, J. D. Jost, C. Langer, R. Ozeri, T. Schaetz, and D. J. Wineland. Implementation of the semiclassical quantum Fourier transform in a scalable system. *Science*, 308:997, 2005. [93]

[48] A. M. Childs. *Quantum information processing in continuous time*. Ph.D., MIT, Cambridge, MA, 2004. [73]

[49] A. M. Childs. Universal computation by quantum walk. *Phys. Rev. Lett.*, 102:180501, 2009. [115]

[50] A. M. Childs. On the relationship between continuous- and discrete-time quantum walk. *Comm. Math. Phys.*, 294:581, 2010. [74, 116, 122]

[51] A. M. Childs, R. Cleve, E. Deotto, and E. Farhi. Exponential algorithmic speedup by quantum walk. In *Proc. 36th ACM Symposium on Theory of Computing*, page 59, 2003. [115, 120, 122]

[52] A. M. Childs, E. Farhi, and S. Gutmann. An example of the difference between quantum and classical random walks. *Quantum Information Processing*, 1:35, 2002. [115, 120, 122]

[53] A. M. Childs, E. Farhi, and J. Preskill. Robustness of adiabatic quantum computation. *Phys. Rev. A*, 65:012322, 2001. [103]

[54] A. M. Childs and R. Kothari. Simulating sparse Hamiltonians with star decompositions. In W. van Dam, V. Kendon, and S. Severini, editors, *Theory of Quantum Computation, Communication, and Cryptography*, volume 6519 of *Lecture Notes in Computer Science*, pages 94–103. Springer Berlin / Heidelberg, 2011. [73]

[55] A. M. Childs, A. J. Landahl, and P. A. Parrilo. Quantum algorithm for ordered search problem via semidefinite programming. *Phys. Rev. A*, 75:032335, 2007. [157]

[56] J. I. Cirac and P. Zoller. Quantum computations with cold trapped ions. *Phys. Rev. Lett.*, 74:4091, 1995. [46]

[57] C. R. Clark, T. S. Metodi, S. D. Gasster, and K. R. Brown. Resource requirements for fault-tolerant quantum simulation: The ground state of the transverse Ising model. *Phys. Rev. A*, 79:062314, 2009. [91, 98, 99]

[58] C. L. Cleveland and R. Medina A. Obtaining a Heisenberg Hamiltonian from the Hubbard model. *Am. J. Phys.*, 44:44, 1976. [60]

[59] C. A. Coulson. Present state of molecular structure calculations. *Rev. Mod. Phys.*, 32:170, 1960. [57]

[60] C. J. Cramer. *Essentials of Computational Chemistry: Theories and models*. Wiley, 2004. [4]

[61] I. J. Crosson, D. Bacon, and K. R. Brown. Making classical ground-state spin computing fault-tolerant. *Phys. Rev. E*, 82(3):031106, September 2010. [108]

[62] D. Deutsch. Quantum theory, the Church-Turing principle and the universal quantum computer. *Proc. R. Soc. A*, 400(1818):97–117, July 1985. [51, 155]

[63] D. Deutsch. Quantum computational networks. *Proc. Roy. Soc. A*, 425:73, 1989. [42]

[64] L. DiCarlo, J. M. Chow, J. M. Gambetta, Lev S. Bishop, B. R. Johnson, D. I. Schuster, J. Majer, A. Blais, L. Frunzio, S. M. Girvin, and R. J. Schoelkopf. Demonstration of two-qubit algorithms with a superconducting quantum processor. *Nature*, 460:240–244, 2009. [45]

[65] S. Diehl, A. Micheli, A. Kantian, B. Kraus, H. P. Büchler, and P. Zoller. Quantum states and phases in driven open quantum systems with cold atoms. *Nature Phys.*, 4(11):878–883, 2008. [120]

[66] D. P. DiVincenzo. The physical implementation of quantum computation. *Fortschr. Phys.*, 48:771, 2000. Also arXiv:quant-ph/0002077. [44]

[67] M. Dobšíček, G. Johansson, V.S. Shumeiko, and G. Wendin. Arbitrary accuracy iterative quantum phase estimation algorithm using a single ancillary qubit: A two-qubit benchmark. *Phys. Rev. A*, 76:030306(R), 2007. [64, 66]

[68] A. Dreuw and M. Head-Gordon. Failure of time-dependent density functional theory for long-range charge-transfer excited states: The zincbacteriochlorin-bacterlochlorin and bacteriochlorophyll-spheroidene complexes. *J. Am. Chem. Soc.*, 126:4007–4016, 2004. [1]

[69] M. Drezgich, A. Hines, M. Sarovar, and S. Sastry. Complete characterization of mixing time for the continuous quantum walk on the hypercube with Markovian decoherence model. *Quantum Inf. & Comp.*, 9:856, 2009. [116]

[70] J. Du, N. Xu, X. Peng, P. Wang, S. Wu, and D. Lu. NMR implementation of a molecular hydrogen quantum simulation with adiabatic state preparation. *Phys. Rev. Lett*, 104:030502, 2010. [45, 85]

[71] E. Farhi, J. Goldstone, and S. Gutmann. A quantum algorithm for the Hamiltonian NAND tree. *Theory of Computing*, 4:169–190, 2008. [115]

[72] E. Farhi, J. Goldstone, S. Gutmann, J. Lapan, A. Lundgren, and D. Preda. A quantum adiabatic evolution algorithm applied to random instances of an NP-complete problem. *Science*, 292:472, 2001. [43, 84]

[73] E. Farhi and S. Gutmann. Quantum computation and decision trees. *Phys. Rev. A*, 58:915, 1998. [115, 117]

[74] M.D. Feit, Jr. J.A. Fleck, and A. Steiger. Solution of the Schrödinger equation by a spectral method. *J. Comp. Phys.*, 47:412, 1982. [72]

[75] R. Feynman. Simulating physics with computers. *Int. J. Theoretical Phys.*, 21:467, 1982. [2, 36]

[76] A. Friedenauer, H. Schmitz, J. T. Glueckert, D. Porras, and T. Schaetz. Simulating a quantum magnet with trapped ions. *Nature Phys.*, 4:757, 2008. [46]

[77] R. Gerritsma, G. Kirchmair, F. Zähringer, E. Solano, R. Blatt, and C. F. Roos. Quantum simulation of the Dirac equation. *Nature*, 463:68, 2010. [46]

[78] V. Giovannetti, S. Lloyd, and L. Maccone. Quantum-enhanced measurements: Beating the standard quantum limit. *Science*, 306:1330, 2004. [69]

[79] V. Gorini, A. Kossakokowski, and E C G. Sudarshan. Completely positive dynamical semigroups of n-level systems. *J. Math. Phys*, 17:821, 1976. [116, 118]

[80] L. K. Grover. A fast quantum mechanical algorithm for database search. *Proc. 28th ACM Symp. on Theory of Comp. (STOC '96)*, page 212, 1996. Also see arxiv:quant-ph/9605043. [157]

[81] L. K. Grover. Quantum mechanics helps in searching for a needle in a haystack. *Phys. Rev. Lett.*, 79:325, 1997. [157]

[82] Lov Grover and Terry Rudolph. Creating superpositions that correspond to efficiently integrable probability distributions. *arXiv:quant-ph/0208112*, 2002. [83]

[83] R. Harris, J. Johansson, A. J. Berkley, M. W. Johnson, T. Lanting, Siyuan Han, P. Bunyk, E. Ladizinsky, T. Oh, I. Perminov, E. Tolkacheva, S. Uchaikin, E. M. Chapple, C. Enderud, C. Rich, M. Thom, J. Wang, B. Wilson, and G. Rose. Experimental demonstration of a robust and scalable flux qubit. *Phys. Rev. B*, 81:134510, 2010. [45]

[84] R. Harris, M. W Johnson, T. Lanting, A. J Berkley, J. Johansson, P. Bunyk, E. Tolka-cheva, E. Ladizinsky, N. Ladizinsky, T. Oh, F. Cioata, I. Perminov, P. Spear, C. En-derud, C. Rich, S. Uchaikin, M. C Thom, E. M Chapple, J. Wang, B. Wilson, M. H. S. Amin, N. Dickson, K. Karimi, B. Macready, C. J. S. Truncik, and G. Rose. Exper-imental investigation of an eight qubit unit cell in a superconducting optimization processor. *Phys. Rev. B*, 82:024511, 2010. [45]

[85] R. Harris, T. Lanting, A. J. Berkley, J. Johansson, M. W. Johnson, P. Bunyk, E. Ladizinsky, N. Ladizinsky, T. Oh, and S. Han. Compound Josephson-junction coupler for flux qubits with minimal crosstalk. *Phys. Rev. B*, 80:052506, 2009. [45]

[86] A. Harrow and A. Montanaro. An efficient test for product states, with applications to quantum Merlin-Arthur game. *Proc. of the 2010 IEEE 51st Annual Symp. on Found. of Comp. Sci. (FOCS '10)*, pages 633–642, 2010. [60]

[87] N. Hatano and M. Suzuki. Finding exponential product formulas of higher orders. In A. Das and B.K. Chakrabarti, editors, *Quantum Annealing and Other Optimization Methods*, Lectures Notes in Physics, pages 37–68. Springer, 2005. [72, 73]

[88] W. J. Hehre, R. F. Stewart, and J. A. Pople. Self-consistent molecular orbital methods I. Use of Gaussian expansions of Slater-type atomic orbitals. *J. Chem. Phys.*, 51:2657, 1969. [97, 98]

[89] T. Helgaker, P. Jorgensen, and J. Olsen. *Molecular Electronic-Structure Theory*. John Wiley and Sons, 2000. [4]

[90] E.J. Heller. The semiclassical way to molecular spectroscopy. *Acc. Chem. Res.*, 14:368–375, 1981. [69]

[91] A. Hines and P. Stamp. Decoherence in quantum walks and quantum computers. *arXiv:0711.1555*, 2007. [116]

[92] Max Hofheinz, H. Wang, M. Ansmann, Radoslaw C. Bialczak, Erik Lucero, M. Neeley, A. D. O'Connell, D. Sank, J. Wenner, John M. Martinis, and A. N. Cleland. Syn-thesizing arbitrary quantum states in a superconducting resonator. *Nature*, 459:546, 2009. [45]

[93] R. A. Horn and C. R. Johnson. *Matrix Analysis*. Cambridge University Press, 2005. [160]

[94] L. Hung and E. A. Carter. Accurate simulations of metals at the mesoscale: Explicit treatment of 1 million atoms with quantum mechanics. *Chem. Phys. Lett.*, 475:163, 2009. [1]

[95] S. Istrail. Statistical mechanics, three-dimensionality and NP-completeness. *Proc. 32nd ACM Symp. on Theory of Comp. (STOC '00)*, page 87, 2000. [57]

[96] M. Johanning, A. F. Varón, and C. Wunderlich. Quantum simulations with cold trapped ions. *J. Phys. B.*, 42:154009, 2009. [46]

[97] M. W. Johnson, P. Bunyk, F. Maibaum, E. Tolkacheva, A. J. Berkley, E. M. Chapple, R. Harris, J. Johansson, T. Lanting, I. Perminov, E. Ladizinsky, T. Oh, and G. Rose. A scalable control system for a superconducting adiabatic quantum optimization processor. *Supercond. Sci. Tech.*, 23:065004, 2010. [45]

[98] P. Jordan and E. Wigner. Über das Paulische äquivalenzverbot. *Z. Phys. A.*, 47:631, 1928. [77]

[99] S. P. Jordan. Fast quantum algorithm for numerical gradient estimation. *Phys. Rev. Lett.*, 95:050501, 2005. [70]

[100] Stephen P. Jordan, Edward Farhi, and Peter W. Shor. Error correcting codes for adiabatic quantum computation. *Phys. Rev. A*, 74:052322, 2006. [106, 108]

[101] W. M. Kaminsky, S. Lloyd, and T. P. Orlando. Scalable superconducting architecture for adiabatic quantum computation. *arXiv:quant-ph/0403090*, 2004. [45]

[102] I. Kassal and A. Aspuru-Guzik. Quantum algorithm for molecular properties and geometry optimization. *J. Chem. Phys.*, 131:224102, 2009. [70]

[103] I. Kassal, S. P. Jordan, P. J. Love, M. Mohseni, and A. Aspuru-Guzik. Polynomial-time quantum algorithm for the simulation of chemical dynamics. *Proc. Natl. Acad. Sci.*, 105:18681, 2008. [69, 74, 75]

[104] P. Kaye, R. Laflamme, and M. Mosca. *An Introduction to Quantum Computing.* Oxford University Press, 2007. [41, 42, 85]

[105] P. Kaye and M. Mosca. Quantum networks for generating arbitrary quantum states. In *International Conf. on Quantum Information*, Rochester, NY, 2004. [83]

[106] J. Kempe. Quantum random walks - an introductory overview. *arXiv:quant-ph/030308*, 2003. [42, 122]

[107] J. Kempe, A. Kitaev, and O. Regev. The complexity of the local Hamiltonian problem. *SIAM J. Computing*, 35(5):1070–1097, 2006. [42, 55, 103, 108, 109]

[108] V. Kendon. Decoherence in quantum walks – a review. *Mathematical Structures in Computer Science*, 17:1169–1220, 2007. [116, 120]

[109] V. Kendon and B. Tregenna. Decoherence can be useful in quantum walks. *Phys. Rev. A*, 67:042315, 2003. [120, 122]

[110] K. Kim, M.-S. Chang, S. Korenblit, R. Islam, E. E. Edwards, J. K. Freericks, G.-D. Lin, L.-M. Duan, and C. Monroe. Quantum simulation of frustrated Ising spins with trapped ions. *Nature*, 465:590–593, 2010. [46]

[111] A. Kitaev. Quantum measurements and the abelian stabilizer problem. *arXiv:quant-ph/9511026*, 1995. [62, 63, 64]

[112] A. Kitaev. Fault-tolerant quantum computation by anyons. *Ann. Phys.*, 303:2, 2003. [42]

[113] A. Kitaev, A. Shen, and M. Vyalyi. *Classical and quantum computation*, volume 47. American Mathematics Society, Graduate Studies in Mathematics, 2002. [55]

[114] A. Kitaev and W. A. Webb. Wavefunction preparation and resampling using a quantum computer. *arXiv:0801.0342*, 2008. [83]

[115] E. Knill, R. Laflamme, and G. J. Milburn. A scheme for efficient quantum computation with linear optics. *Nature*, 409:46–52, 2001. [45, 93, 95]

[116] E. Knill, G. Ortiz, and R. Somma. Optimal quantum measurements of expectation values of observables. *Phys. Rev. A*, 75:012328, 2007. [64, 66, 69]

[117] W. Koch and M. C. Holthausen. *A Chemist's guide to density functional theory*. Wiley-VCH, 2001. [4]

[118] W. Kohn. Nobel lecture: Electronic structure of matter—wave functions and density functionals. *Rev. Mod. Phys.*, 71:1253–1266, 1999. [86]

[119] P. Kok, W. J. Munro, K. Nemoto, T. C. Ralph, J. P. Dowling, and G. J. Milburn. Linear optical quantum computing with photonic qubits. *Rev. Mod. Phys.*, 79:135, 2007. [45, 93]

[120] N. Konno. Quantum walks. In U. Franz and M. Schürmann, editors, *Quantum Potential Theory, Lecture Notes in Mathematics*, page 309. Springer-Verlag, 2008. [115]

[121] D. Kosloff and R. Kosloff. A Fourier method solution for the time dependent Schrödinger equation as a tool in molecular dynamics. *J. Comput. Phys.*, 51:35, 1983. [72]

[122] A. Kossakowski. On quantum statistical mechanics of non-Hamiltonian systems. *Rep. Math. Phys*, 3:247, 1972. [116, 118]

[123] B. Kraus, H. P. Büchler, S. Diehl, A. Kantian, A. Micheli, and P. Zoller. Preparation of entangled states by quantum Markov processes. *Phys. Rev. A*, 78(4):042307, 2008. [120]

[124] T. D. Ladd, F. Jelezko, R. Laflamme, Y. Nakamura, C. Monroe, and J. L. O'Brien. Quantum computers. *Nature*, 464(7285):45–53, 2010. [43, 44]

[125] R. Landauer. Information is physical. *Phys. Today*, 44:23–29, 1991. [36]

[126] T. Lanting, R. Harris, J. Johansson, M. H. S. Amin, A. J. Berkley, S. Gildert, M. W. Johnson, P. Bunyk, E. Tolkacheva, E. Ladizinsky, N. Ladizinsky, T. Oh, I. Perminov, E. M Chapple, C. Enderud, C. Rich, B. Wilson, M. C Thom, S. Uchaikin, and G. Rose. Observation of co-tunneling in pairs of coupled flux qubits. *Phys. Rev. B*, 82:060512, 2010. [45]

[127] B. P. Lanyon, M. Barbieri, M. P. Almeida, T. Jennewein, T. C. Ralph, K. J. Resch, Geoff J. Pryde, J. L. O'Brien, A. Gilchrist, and A. G. White. Simplifying quantum logic using higher-dimensional Hilbert spaces. *Nature Phys.*, 5:134–140, 2009. [45, 93, 94, 100]

[128] B. P. Lanyon, T. J. Weinhold, N. K. Langford, M. Barbieri, D. F. V. James, A. Gilchrist, and A. G. White. Experimental demonstration of a compiled version of Shor's algorithm with quantum entanglement. *Phys. Rev. Lett.*, 99:250505, 2007. [45]

[129] P. A. Lee, N. Nagaosa, and X.-G. Wen. Doping a Mott insulator: Physics of high-temperature superconductivity. *Rev. Mod. Phys.*, 78:17, 2006. [1]

[130] D. Leibfried, B. DeMarco, V. Meyer, D. Lucas, M. Barrett, J. Britton, W. M. Itano, B. Jelenkovic, C. Langer, T. Rosenband, and D. J. Wineland. Experimental demonstration of a robust, high-fidelity geometric two ion-qubit phase gate. *Nature*, 422:412–415, 2003. [44]

[131] W. A. Lester Jr., L. Mitas, and B. Hammond. Quantum Monte Carlo for atoms, molecules and solids. *Chem. Phys. Lett.*, 478:1–10, 2009. [2]

[132] B. Lévi, B. Georgeot, and D. L. Shepelyansky. Quantum computing of quantum chaos in the kicked rotator model. *Phys. Rev. E*, 67:046220, 2003. [75]

[133] B. G. Levi. Superconducting qubit systems come of age. *Phys. Today*, 62:14, 2009. [45]

[134] B. G. Levine and T. J. Martínez. Isomerization through conical intersections. *Ann. Rev. Phys. Chem.*, 58:613, 2007. [1]

[135] Yang Liao, Jian Xu, Ya Cheng, Zenghui Zhou, Fei He, Haiyi Sun, Juan Song, Xinshun Wang, Zhizhan Xu, Koji Sugioka, and Katsumi Midorikawa. Electro-optic integration of embedded electrodes and waveguides in LiNbO$_3$ using a femtosecond laser. *Opt. Lett.*, 33:2281, 2008. [98]

[136] D. A. Lidar. Towards fault tolerant adiabatic quantum computation. *Phys. Rev. Lett.*, 100(17):160506, 2008. [106, 108]

[137] D.A. Lidar and H. Wang. Calculating the thermal rate constant with exponential speedup on a quantum computer. *Phys. Rev. E*, 59:2429, 1999. [69, 75, 91]

[138] G. Lindblad. On the generators of quantum dynamical semigroups. *Commun. Math. Phys*, 48:119, 1975. [116, 118]

[139] Seth Lloyd. Universal quantum simulators. *Science*, 273:1073–8, Aug 1996. [2, 36, 71, 72, 73]

[140] P. Love and B. Boghosian. From Dirac to diffusion: decoherence in quantum lattice gases. *Quantum Information Processing*, 4:335, 2005. [116]

[141] D. Lu, N. Xu, R. Xu, H. Chen, J. Gong, X. Peng, and J. Du. Simulation of chemical isomerization reaction dynamics on a nmr quantum simulator. *Phys. Rev. Lett.*, 107:020501, 2011. [45]

[142] Y.-K. Lui, M. Christandl, and F. Verstraete. N-representability is QMA-complete. *Phys. Rev. Lett.*, 98:110503, 2007. [55, 58, 59]

[143] J. C. Maxwell. *Theory of Heat.* Dover Publications, 2001. [37]

[144] D. A. Mayer. From quantum cellular automata to quantum lattice gases. *J. Stat. Phys.*, 85:551, 1996. [116]

[145] Z. Merali. First sale for quantum computing. *Nature*, 474:18, 2011. [46]

[146] A. Messiah. *Quantum Mechanics.* Dover Publications, 1999. [44, 84, 105]

[147] M. Mezard and A. Montanari. *Information, Physics, and Computation.* Oxford University Press, 2009. [85]

[148] B. Misra and E. C. G. Sudarshan. The Zeno's paradox in quantum theory. *J. Math. Phys.*, 18:756, 1977. [85]

[149] A. Mizel, D. A. Lidar, and M. Mitchell. Simple proof of equivalence between adiabatic quantum computation and the circuit model. *Phys. Rev. Lett.*, 99:070502, 2007. [42]

[150] M. Mohseni, P. Rebentrost, S. Lloyd, and A. Aspuru-Guzik. Environment-assisted quantum walks in energy transfer of photosynthetic complexes. *Journal of Chemical Physics*, 129(174106), 2008. [116, 118, 122]

[151] C. Moler and C. Van Loan. Nineteen dubious ways to compute the exponential of a matrix, twenty-five years later. *SIAM Review*, 45:1–46, 2003. [71]

[152] R. P. Muller. Python quantum chemistry (PyQuante) program, version 1.6, (2007). [88]

[153] C. Nayak, S. H. Simon, A. Stern, M. Freedman, and S. Das Sarma. Non-abelian anyons and topological quantum computation. *Rev. Mod. Phys.*, 80:1083, 2008. [42]

[154] C. Negrevergne, R. Somma, G. Ortiz, E. Knill, and R. Laflamme. Liquid-state NMR simulations of quantum many-body problems. *Phys. Rev. A*, 71:032344, 2005. [45]

[155] M. Nielsen and I. Chuang. *Quantum Computation and Quantum Information.* Cambridge University Press, 2001. [41, 42, 43, 69, 91, 98, 99, 100, 101]

[156] Michael A. Nielsen. Optical quantum computation using cluster states. *Phys. Rev. Lett.*, 93:040503, 2004. [95]

[157] A. O. Niskanen, K. Harrabi, F. Yoshihara, Y. Nakamura, S. Lloyd, and J. S. Tsai. Quantum coherent tunable coupling of superconducting qubits. *Science*, 316(5825):723–726, 2007. [107, 111]

[158] L. Null and J. Lobur. *The essentials of computer organization and architecture.* Jones & Bartlett Publishers, 2003. [42]

[159] J. L. O'Brien, G. J. Pryde, A. G. White, T. C. Ralph, and D. Branning. Demonstration of an all-optical quantum controlled-NOT gate. *Nature*, 426:264, 2003. [44, 93]

[160] Jeremy L. O'Brien. Optical quantum computing. *Science*, 318:1567, 2007. [45]

[161] S. Oh. Quantum computational method of finding the ground state energy and expectation values. *Phys. Rev. A*, 77:012326, 2008. [63, 85]

[162] R. Oliveira and B. Terhal. The complexity of quantum spin systems on a two-dimensional square lattice. *Quant. Inf. Comp.*, 8:0900, 2008. [55, 103, 108, 109]

[163] G. Ortiz, J. Gubernatis, E. Knill, and R. Laflamme. Quantum algorithms for fermionic simulations. *Phys. Rev. A*, 64:022319, 2001. [70, 77, 78]

[164] E. Ovrum and M. Hjorth-Jensen. Quantum computation algorithm for many-body studies. *arXiv:0705.1928*, 2007. [78]

[165] A. Papageorgiou and C. Zhang. On the efficiency of quantum algorithms for Hamiltonian simulation. *ArXiv e-prints*, May 2010. [73, 74]

[166] S. Parker and M. Plenio. Efficient factorization with a single pure qubit and log N mixed qubits. *Phys. Rev. Lett.*, 85:3049, 2000. [63, 64]

[167] Yu. A. Pashkin, T. Yamamoto, O. Astafiev, Y. Nakamura, D. V. Averin, and J. S. Tsai. Quantum oscillations in two coupled charge qubits. *Nature*, 421:823, 2003. [44]

[168] X. Peng, J. Du, and D. Suter. Quantum phase transition of ground-state entanglement in a Heisenberg spin chain simulated in an NMR quantum computer. *Phys. Rev. A*, 71:012307, 2005. [45]

[169] A. Perdomo-Ortiz, M. Drew-Brook, N. Dickson, G. Rose, and A. Aspuru-Guzik. Experimental realization of a 8-qubit quantum-adiabatic algorithm for a lattice protein model: Towards optimization on a quantum computer. *In preparation*, 2011. [45]

[170] A. Perdomo-Ortiz, S. E. Venegas-Andraca, and A. Aspuru-Guzik. Sombrero adiabatic quantum computation. *Quantum Information Processing*, pages 1–20, 2010. [44, 84]

[171] J. H. Plantenberg, P. C. de Groot, C. J. P. M. Harmans, and J. E. Mooij. Demonstration of controlled-NOT quantum gates on a pair of superconducting quantum bits. *Nature*, 447:836–839, 2007. [44]

[172] M. Plenio and S. Huelga. Dephasing-assisted transport: quantum networks and biomolecules. *New J. Phys.*, 10:113019, 2008. [116, 122]

[173] J. Pople. Quantum chemical models. *Nobel Lecture*, 1998. [2]

[174] D. Poulin, A. Qarry, R. Somma, and F. Verstraete. Quantum simulation of time-dependent hamiltonians and the convenient illusion of Hilbert space. *Phys. Rev. Lett.*, 106(17):170501, 2011. [72]

[175] D. Poulin and P. Wocjan. Preparing ground states of quantum many-body systems on a quantum computer. *Phys. Rev. Lett.*, 102:130503, 2009. [86]

[176] D. Poulin and P. Wocjan. Sampling from the thermal quantum Gibbs state and evaluating partition functions with a quantum computer. *Phys. Rev. Lett.*, 103:220502, 2009. [86]

[177] J. Preskill. Lecture notes for quantum computation. Available at . [161]

[178] E. J. Pritchett, C. Benjamin, A. Galiautdinov, M. R. Geller, A. T. Sornborger, P. C. Stancil, and J. M. Martinis. Quantum simulation of molecular collisions with superconducting qubits. *arXiv:1008.0701*, 2010. [46]

[179] N. Ramsey. *Molecular Beams*. Oxford University Press, 1963. [106]

[180] V. A. Rassolov and S. Garashchuk. Computational complexity in quantum chemistry. *Chem. Phys. Lett.*, 464:262–264, 2008. [61]

[181] R. Raussendorf and H. J. Briegel. A one-way quantum computer. *Phys. Rev. Lett.*, 86:5188, 2001. [42]

[182] R. Raussendorf, D. E. Browne, and H. J. Briegel. Measurement-based quantum computation on cluster states. *Phys. Rev. A*, 68:022312, 2003. [42]

[183] P. Rebentrost, M. Mohseni, and A. Aspuru-Guzik. Role of quantum coherence and environmental fluctuations in chromophoric energy transport. *J. Phys. Chem. B*, 113:9942, 2009. [116, 120, 122]

[184] P. Rebentrost, M. Mohseni, I. Kassal, S. Lloyd, and A. Aspuru-Guzik. Environment-assisted quantum transport. *New Journal of Physics*, 11:033003, 2009. [116, 122]

[185] C. A. Rodriguez-Rosario and E. C. G. Sudarshan. Non-Markovian open quantum systems. *arXiv:0803.1183*, 2008. [116]

[186] A. Romanelli, R. Siri, G. Abal, A. Auyuanet, and R. Donangelo. Decoherence in the quantum walk on the line. *Physica A*, 347:137 – 152, 2005. [116]

[187] S. Salimi and A. Sorouri. Pseudo-Hermitian continuous-time quantum walks. *arXiv:0904.4026*, 2009. [117]

[188] Ferdinand Schmidt-Kaler, Hartmut Haffner, Mark Riebe, Stephan Gulde, Gavin P. T. Lancaster, Thomas Deuschle, Christoph Becher, Christian F. Roos, Jurgen Eschner, and Rainer Blatt. Realization of the Cirac-Zoller controlled-NOT quantum gate. *Nature*, 422:408–411, 2003. [44]

[189] U. Schollwöck. The density-matrix renormalization group in the age of matrix product states. *Ann. Phys.*, 326:96–192, 2011. [41]

[190] Norbert Schuch and Frank Verstraete. Computational Complexity of interacting electrons and fundamental limitations of density functional theory. *Nature Phys.*, 5:732, 2009. Also see appendix of arxiv:0712.0483. [56, 57, 59, 60]

[191] C. E. Shannon. A mathematical theory of communication. *Bell Sys. Tech. J.*, 27:379–423,623–656, 1948. [37]

[192] N. Shenvi, J. Kempe, and K.B. Whaley. Quantum random-walk search algorithm. *Phys. Rev. A*, 67:052307, 2003. [115]

[193] P. W. Shor. Polynomial-time algorithms for prime factorization and discrete logarithms on a quantum computer. *SIAM J. Sci. Statist. Comput.*, 26:1484, 1997. [157]

[194] I. Siddiqi, R. Vijay, F. Pierre, C. M. Wilson, L. Frunzio, M. Metcalfe, C. Rigetti, R. J. Schoelkopf, M. H. Devoret, D. Vion, and D. Esteve. Direct observation of dynamical bifurcation between two driven oscillation states of a Josephson junction. *Phys. Rev. Lett.*, 94(2):027005, 2005. [106]

[195] M. Sipser. *Introduction to the Theory of Computation.* PWS Publishing Company, 1997. [48]

[196] A.Y. Smirnov, S. Savel'ev, L.G. Mourokh, and F. Nori. Modelling chemical reactions using semiconductor quantum dots. *Europhys. Lett.*, 80:67008, 2007. [70]

[197] A. N. Soklakov and R. Schack. Efficient state preparation for a register of quantum bits. *Phys. Rev. A*, 73:012307, 2006. [83]

[198] S. Somaroo, C. H. Tseng, T. F. Havel, R. Laflamme, and D. G. Cory. Quantum simulations on a quantum computer. *Phys. Rev. Lett.*, 82:5381–5384, 1999. [45]

[199] R. Somma, C. Batista, and G. Ortiz. Quantum approach to classical statistical mechanics. *Phys. Rev. Lett.*, 99:1–4, 2007. [85]

[200] R. Somma, S. Boixo, H. Barnum, and E. Knill. Quantum simulations of classical annealing processes. *Phys. Rev. Lett.*, 101:130504, 2008. [86]

[201] R. Somma, G. Ortiz, J. E. Gubernatis, E. Knill, and R. Laflamme. Simulating physical phenomena by quantum networks. *Phys. Rev. A*, 65:042323, 2002. [77, 78]

[202] R. Somma, G. Ortiz, E. Knill, and J. Gubernatis. Quantum simulations of physics problems. *Int. J. Quantum Inf.*, 1:189, 2003. [70]

[203] F. Strauch. Connecting the discrete- and continuous-time quantum walks. *Phys. Rev. A*, 74:030301, 2006. [116, 122]

[204] F. Strauch. Reexamination of decoherence in quantum walks on the hypercube. *Phys. Rev. A*, 79:032319, 2009. [116]

[205] F. Strauch. Unary vs binary walks talk. *Private Communication*, 2009. [116]

[206] F. Strauch and C. Williams. Theoretical analysis of perfect quantum state transfer with superconducting qubits. *Phys. Rev. B*, 78:094516, 2008. [116]

[207] G. Strini. Error sensitivity of a quantum simulator. I: A first example. *Fortsch. Phys.*, 50:171–183, 2002. [75]

[208] E. C. G. Sudarshan, P. M. Mathews, and Jayaseetha Rau. Stochastic dynamics of quantum-mechanical systems. *Phys. Rev.*, 121(3):920–924, Feb 1961. [116, 118, 122]

[209] A. Szabo and N. Ostlund. *Modern Quantum Chemistry: Introduction to Advanced Electronic Structure Theory.* Dover Publications, 1996. [4, 97]

[210] M. Szegedy. Quantum speed-up of Markov chain based algorithms. In *Proc. 45th Ann. IEEE Symp. Found. Comput. Sci.*, pages 32–41, 2004. [86]

[211] D. J. Tannor. *Introduction to Quantum Mechanics: A Time-Dependent Perspective.* University Science Books, 2006. [69]

[212] K. Temme, T. J. Osborne, K. G. Vollbrecht, D. Poulin, and F. Verstraete. Quantum Metropolis sampling. *Nature*, 471:87, 2011. [86]

[213] B. M. Terhal and D. P. DiVincenzo. Problem of equilibration and the computation of correlation functions on a quantum computer. *Phys. Rev. A*, 61(2), 2000. [70, 85, 86]

[214] A. Tomita and K. Nakamura. Measured quantum Fourier transform of 1024 qubits on fiber optics. *Int. J. Quantum Information*, 2:119, 2004. [64]

[215] D. M. Tong. Quantitative condition is necessary in guaranteeing the validity of the adiabatic approximation. *Phys. Rev. Lett.*, 104:120401, 2010. [44, 84]

[216] D. M. Tong, K. Singh, L. C. Kwek, and C. H. Oh. Sufficiency criterion for the validity of the adiabatic approximation. *Phys. Rev. Lett.*, 98:150402, 2007. [44, 84]

[217] M. Troyer and U.-J. Wiese. Computational complexity and fundamental limitations to fermionic quantum Monte Carlo simulations. *Phys. Rev. Lett.*, 94:170201, 2005. [61]

[218] L. Veis and J. Pittner. Quantum computing applied to calculations of molecular energies: CH_2 benchmark. *J. Chem. Phys.*, 133:194106, 2010. [84]

[219] S. Venegas-Andraca. *Quantum Walks for Computer Scientists.* Morgan and Claypool, 2008. [115]

[220] Frank Verstraete, Michael M. Wolf, and J. Ignacio Cirac. Quantum computation and quantum-state engineering driven by dissipation. *Nature Phys.*, 5(9):633–636, 2009. [120]

[221] G Vidal. Efficient classical simulation of slightly entangled quantum computations. *Phys. Rev. Lett.*, 91:147902, 2003. [41]

[222] T. Van Voorhis and M. Head-Gordon. Benchmark variational coupled cluster doubles results. *J. Chem. Phys.*, 113:8873, 2000. [1]

[223] H. Wang, S. Ashhab, and F. Nori. An efficient quantum algorithm for preparing states of molecular systems. *Phys. Rev. A*, 79:042335, 2009. [84]

[224] H. Wang, S. Kais, A. Aspuru-Guzik, and M. Hoffmann. Quantum algorithm for obtaining the energy spectrum of molecular systems. *Phys. Chem. Chem. Phys.*, 10:5388, 2008. [84]

[225] N. Ward, I. Kassal, and A. Aspuru-Guzik. Preparation of many-body states for quantum simulation. *J. Chem. Phys.*, 130:194105, 2009. [83]

[226] J. Watrous. Quantum simulations of classical random walks and undirected graph connectivity. *J. Comput. Syst. Sci*, 62:376, 2001. [115]

[227] J.S. Waugh, L.M. Huber, and U. Haeberlen. Approach to high-resolution NMR in solids. *Phys. Rev. Lett.*, 20(5):180–182, 1968. [104, 110]

[228] N. Wiebe, D. W. Berry, P. Hoyer, and B. C. Sanders. Simulating quantum dynamics on a quantum computer. *arXiv:1101.3489*, 2011. [72]

[229] S. Wiesner. Simulations of many-body quantum systems by a quantum computer. *arXiv:quant-ph/9603028*, 1996. [74]

[230] P. Wocjan and A. Abeyesinghe. Speedup via quantum sampling. *Phys. Rev. A*, 78:042336, 2008. [85, 86]

[231] L.-A. Wu, M. S. Byrd, and D. A. Lidar. Polynomial-time simulation of pairing models on a quantum computer. *Phys. Rev. Lett.*, 89:057904, 2002. [84, 85]

[232] L. Xiu-Mei, L. Jun, and S. Xian-Ping. Experimental realizations of arbitrary accuracy iterative phase estimation algorithms on a ensemble quantum computers. *Chinese Phys. Lett.*, 24:3316, 2007. [64, 93]

[233] X. Yang, A. M. Wang, F. Xu, and J. Du. Experimental simulation of a pairing Hamiltonian on an NMR quantum computer. *Chem. Phys. Lett.*, 422:20–24, 2006. [45]

[234] J. Q. You and F. Nori. Superconducting circuits and quantum information. *Physics Today*, 58:42, 2005. [45]

[235] M.-H. Yung, D. Nagaj, J. D. Whitfield, and A. Aspuru-Guzik. Simulation of classical thermal states on a quantum computer: A renormalization group approach. *Phys. Rev. A*, 82:060302(R), 2010. Also see Chapter 5. [83, 85]

[236] C. Zalka. Simulating quantum systems on a quantum computer. *Proc. R. Soc. Lond. A*, 454:313, 1998. [74, 83, 85, 86]

Appendix A

Quantum circuit for hydrogen molecule propagator

The time evolution of input states is measured to extract eigenvalues of the system. A state preparation procedure must be used if a particular state is desired e.g. the ground state. Time evolution is implemented in such a way that a measurable relative phase is generated on a register qubit. This register qubit is then measured to obtain a bit of information about the phase. The time evolution is then effected for wice as long, and again the register is measured to obtain the next digit. This process is repeated until all digits are read out and, then, the phase can be converted to an energy eigenvalue following the prescription of Section 5.1.

In this appendix, we explicitly spell out the gates that would be necessary for effecting a single time step of the evolution generated by the hydrogen molecule's Hamiltonian. The time step idea is to alternatively apply each of the non-commuting unitary operators for a fraction of the total time and repeat the time steps until the total time has been reached. As the length of the time steps goes to zero, the error of the approximation vanishes. For constructing long time evolution, the time steps should be composed following the Trotter prescription from Section 6.2.1.

The decomposition of the propagator of H_2 Hamiltonian is given using standard quantum gates. The entire unitary operator is applied if the register qubit is in state $|1\rangle$. Qubits are named: Register, Q1, Q2, Q3, Q4. The integrals, h_{ij} and h_{ijkl}, are parametrized by the nuclear configuration and are given in Table 8.1 for the for the equilibrium geometry. Additionally, we define $n_{ij} = (h_{ijji} - h_{jiji})/4$. Note only h_{1313} and h_{2424} are the only non-zero terms due to spin orthogonality. Following the main text, let $\Theta \equiv (1/4) \sum_{p<q} (h_{pqqp} - h_{pqpq} \delta_{\sigma_p \sigma_q})$ and $\theta_p \equiv \sum_{q:p \neq q} (h_{pqqp} - h_{pqpq} \delta_{\sigma_p \sigma_q})$.

Gate	Target qubit	Control qubit	Parameter
Hadamard	Register		
SINGLE ELECTRON OPERATORS			
cPhase	Q1	Register	$h_{11}t$
cPhase	Q2	Register	$h_{22}t$
cPhase	Q3	Register	$h_{33}t$
cPhase	Q4	Register	$h_{44}t$
TWO ELECTRON OPERATORS: NUMBER-NUMBER OPERATOR			
Phase	Register		Θt
cR_z	Q1	Register	$-\theta_1 t$
cR_z	Q2	Register	$-\theta_2 t$
cR_z	Q3	Register	$-\theta_3 t$
cR_z	Q4	Register	$-\theta_4 t$
CNOT	Q4	Q3	
cR_z	Q4	Register	$2n_{34}t$
CNOT	Q4	Q3	
CNOT	Q4	Q2	
cR_z	Q4	Register	$2n_{24}t$
CNOT	Q4	Q2	
CNOT	Q4	Q1	
cR_z	Q4	Register	$2n_{14}t$
CNOT	Q4	Q1	
CNOT	Q3	Q2	
cR_z	Q3	Register	$2n_{23}t$
CNOT	Q3	Q2	
CNOT	Q3	Q1	
cR_z	Q3	Register	$2n_{13}t$
CNOT	Q3	Q1	
CNOT	Q2	Q1	
cR_z	Q2	Register	$2n_{12}t$
CNOT	Q2	Q1	

Gate	Target qubit	Control qubit	Parameter
Two Electron Operators: excitation-excitation operator			
XXYY			
Hadamard	Q1		
Hadamard	Q2		
R_x	Q3		$-\pi/2$
R_x	Q4		$-\pi/2$
CNOT	Q2	Q1	
CNOT	Q3	Q2	
CNOT	Q4	Q3	
cR_z	Q4	Register	$-t(h_{1423} + h_{1243})/4$
CNOT	Q4	Q3	
CNOT	Q3	Q2	
CNOT	Q2	Q1	
R_x	Q4		$\pi/2$
R_x	Q3		$\pi/2$
Hadamard	Q2		
Hadamard	Q1		
YYXX			
R_x	Q1		$-\pi/2$
R_x	Q2		$-\pi/2$
Hadamard	Q3		
Hadamard	Q4		
CNOT	Q2	Q1	
CNOT	Q3	Q2	
CNOT	Q4	Q3	
cR_z	Q4	Register	$-t(h_{1423} + h_{1243})/4$
CNOT	Q4	Q3	
CNOT	Q3	Q2	
CNOT	Q2	Q1	
Hadamard	Q4		
Hadamard	Q3		
R_x	Q2		$\pi/2$
R_x	Q1		$\pi/2$

Gate	Target qubit	Control qubit	Parameter
XYYX			
Hadamard	Q1		
R_x	Q2		$-\pi/2$
R_x	Q3		$-\pi/2$
Hadamard	Q4		
CNOT	Q2	Q1	
CNOT	Q3	Q2	
CNOT	Q4	Q3	
cR_z	Q4	Register	$t(h_{1423} + h_{1243})/4$
CNOT	Q4	Q3	
CNOT	Q3	Q2	
CNOT	Q2	Q1	
Hadamard	Q4		
Hadamard	Q3		
R_x	Q2		$\pi/2$
R_x	Q1		$\pi/2$
YXXY			
R_x	Q1		$\pi/2$
Hadamard	Q2		
Hadamard	Q3		
R_x	Q4		$\pi/2$
CNOT	Q2	Q1	
CNOT	Q3	Q2	
CNOT	Q4	Q3	
cR_z	Q4	Register	$t(h_{1423} + h_{1243})/4$
CNOT	Q4	Q3	
CNOT	Q3	Q2	
CNOT	Q2	Q1	
R_x	Q1		$-\pi/2$
Hadamard	Q2		
Hadamard	Q3		
R_x	Q4		$-\pi/2$

Table A1: A circuit that implements the unitary propagator associated with the two-electron operators of the H_2 Hamiltonian in the minimal basis. The circuit is derived from the two non-vanishing real integrals of the two-electron interaction and the corresponding operators. In the circuit, $\theta \equiv h_{1423} + h_{1243} = 0.36257$. In Eq. (8.9) two of the four pre-factors vanish due to the spin orthogonality however, in general, there would be eight terms to simulate.

Appendix B

Derivation of the Jordan-Wigner transformation

In this brief appendix, the Jordan-Wigner transformation used to map fermions onto qubits is derived from the fermionic algebra. This transformation is also used to solve various spin Hamiltonians by mapping them onto non-interacting fermions.

By evaluating the matrix elements of the creation/annihilation operators, a matrix representation can be found using Pauli spin matrices. The rules are given in Table 6.2. From there

$$a_p|\mathbf{k}\rangle = \delta_{k_p,1}\Gamma_p^{\mathbf{k}}|k_1,...,0_p,...,k_n\rangle \tag{B.1}$$

$$\langle\mathbf{m}|a_p = \delta_{m_p,0}\Gamma_p^{\mathbf{m}}\langle m_1,...,1_p,...,m_n| \tag{B.2}$$

Combining these equations leads to

$$
\begin{aligned}
\langle\mathbf{m}|a_p|\mathbf{k}\rangle &= \delta_{k_p,1}\Gamma_p^{\mathbf{k}}\langle m_1,...,m_p,...,m_n|k_1,...,0_p,...,k_n\rangle \\
&= \delta_{m_p,0}\Gamma_p^{\mathbf{m}}\langle m_1,...,1_p,...,m_n|k_1,...,k_p,...,k_n\rangle \\
&= \delta_{k_p,1}\Gamma_p^{\mathbf{k}}\left(\delta_{m_p,0}\prod_{i=1,i\neq p}^{n}\delta_{m_i,k_i}\right) = \delta_{m_p,0}\Gamma_p^{\mathbf{m}}\left(\delta_{k_p,1}\prod_{i=1,i\neq p}^{n}\delta_{m_i,k_i}\right) \\
&= \delta_{k_p,1}\left[\prod_{j=1}^{p-1}(-1)^{k_j}\right]\left[\prod_{j=p+1}^{n}(1)^{k_j}\right]\left(\delta_{m_p,0}\prod_{i=1,i\neq p}^{n}\delta_{m_i,k_i}\right) \\
&= \left[\prod_{j=1}^{p-1}(-1)^{k_j}\delta_{m_j,k_j}\right]\delta_{m_p,0}\delta_{k_p,1}\left[\prod_{j=p+1}^{n}(1)^{k_j}\delta_{m_j,k_j}\right] \\
&= \prod_{j=1}^{p-1}\left[\langle m_j|0\rangle\langle 0|k_j\rangle - \langle m_j|1\rangle\langle 1|k_j\rangle\right]\langle m_p|0\rangle\langle 1|k_p\rangle \prod_{j=p+1}^{n}\left[\langle m_j|0\rangle\langle 0|k_j\rangle + \langle m_j|1\rangle\langle 1|k_j\rangle\right] \\
&= \prod_{j=1}^{p-1}\left[\langle m_j|\sigma^z|k_j\rangle\right]\langle m_p|\sigma^+|k_p\rangle \prod_{j=p+1}^{n}\left[\langle m_j|\sigma_0|k_j\rangle\right]
\end{aligned}
$$

Now supposing that $|\mathbf{m}\rangle = \bigotimes_{j=1}^{n} |m_j\rangle$ and similarly for $|\mathbf{k}\rangle$, the transformation commonly known as the Jordan-Wigner transform is derived:

$$\langle \mathbf{m}|a_p|\mathbf{k}\rangle = \langle \mathbf{m}| \left[\bigotimes_{j=1}^{p-1} \sigma_j^z \otimes \sigma_p^+ \bigotimes_{j=p+1}^{n} \sigma_j^0 \right] |\mathbf{k}\rangle = \langle \mathbf{m}| \left[(\sigma^z)^{\otimes p-1} \otimes (\sigma^+)_p \otimes (\sigma^0)^{\otimes n-p} \right] |\mathbf{k}\rangle$$

$$(\text{B.3})$$

Appendix C

Singular value decomposition

The singular value decomposition is similar to the eigenvalue decomposition; however, it can be applied to matrices that are not square. Singular values are frequently used to characterize norms of matrices. For instance, the operator norm is the largest singular value and the trace norm is the sum of the singular values.

The singular value decomposition of T says that

$$T = V\Sigma W^\dagger \tag{C.1}$$

with the following spatial decomposition: $[n \times m] = [n \times n][n \times m][m \times m]$. Here $[n \times m]$ denotes the space of matrices with n columns and m rows. V and W are unitary and Σ is diagonal. The diagonal elements of Σ are the singular values and, they are uniquely specified by T.

The singular values of matrix T are the eigenvalues of $\sqrt{T^\dagger T}$. Since $T^\dagger T$ is positive semi-definite, the square root function is well defined. Additionally, since $T^\dagger T$ is normal ($[T, T^\dagger] = 0$) we have $T^\dagger T = V\Lambda V^\dagger$ with Λ a diagonal matrix and V a unitary matrix. We define the singular matrix as $\Sigma \equiv \sqrt{\Lambda}$.

For arbitrary matrix, T, we now give constructions for the input and output unitary matrices, V and W, which will prove that the SVD exists for all matrices. If T is non-singular, we can invert the diagonal singular value matrix, Σ^{-1} by taking the reciprocal of the singular values. Letting $W = T^\dagger V\Sigma^{-1}$, W is unitary and $T = V\Sigma W^\dagger$.

When T is singular, the inverse of Σ is no longer available as some of the singular values are zero. Instead, the pseudo-inverse is taken where only the inverse of non-zero diagonal elements is used. As before, the output matrix W is defined using the pseudo-inverse. To satisfy the unitary constraint, arbitrary orthonormal vectors are included as columns of W. Thus, the singular value decomposition exists for all matrices regardless of shape and the matrices need not satisfy special properties.

The vector form of the SVD can be easily seen as $TW = V\Sigma$ and $V^\dagger T = \Sigma W^\dagger$ imply

$$T|w_i\rangle = \sigma_i|v_i\rangle \tag{C.2}$$
$$\langle v_i|T = \langle w_i|\sigma_i \tag{C.3}$$

The vectors $\{|v_i\rangle\}$ are called the left singular vectors and $\{|w_i\rangle\}$, the right singular vectors.

C.1 Normal matrices

If matrix A is normal ($A^\dagger A = A A^\dagger$), then the spectral decomposition exist for A that is $A = S D S^\dagger$ for some S unitary and D diagonal. The singular value decomposition is

$$
A = S D S^\dagger = S \begin{bmatrix} |\lambda_1| e^{i\varphi_1} & & & \\ & |\lambda_2| e^{i\varphi_2} & & \\ & & \ddots & \\ & & & |\lambda_n| e^{i\varphi_n} \end{bmatrix} S^\dagger \tag{C.4}
$$

$$
= S \begin{bmatrix} |\lambda_1| & & & \\ & |\lambda_2| & & \\ & & \ddots & \\ & & & |\lambda_n| \end{bmatrix} \begin{bmatrix} e^{i\varphi_1} & & & \\ & e^{i\varphi_2} & & \\ & & \ddots & \\ & & & e^{i\varphi_n} \end{bmatrix} S^\dagger \tag{C.5}
$$

$$
= S \Sigma D S^\dagger \tag{C.6}
$$

$$
= S \Sigma (S D^\dagger)^\dagger = V \Sigma W^\dagger \tag{C.7}
$$

C.2 Schmidt decomposition

The Schmidt decomposition discussed in Subsection 3.3.2 is essentially a restatement of the SVD we have discussed so far. We now construct the Schmidt decomposition using the SVD. Given an arbitrary orthonormal basis for H_A and H_B denoted as $\{|i\rangle\}$ and $\{|j\rangle\}$ respectively, we wish to ascertain the Schmidt vectors and the Schmidt coefficients. Since we can span $H_A \otimes H_B$ by the tensor product of the bases, we can write the state ψ as, $|\psi\rangle = \sum_{ij} T_{ij} |i\rangle |j\rangle$, where T is $[n \times m]$. Here n is the dimension of H_A and m is the dimension of H_B. Applying the SVD to matrix T yields $T = V \Sigma W^\dagger$. Evaluating the $(i,j)^{th}$ matrix element and inserting resolutions of the identity we get,

$$
T_{ij} = \sum_{l=1}^{m} \sum_{k=1}^{n} V_{ik} \Sigma_{kl} W_{lj}^\dagger.
$$

Inserting this into the expression for ψ

$$
\begin{aligned}
|\psi\rangle &= \sum_{ij} T_{ij} |i\rangle |j\rangle \\
&= \sum_{k=1}^{n} \sigma_k \delta_{kl} \left(\sum_i V_{ik} |i\rangle \right) \left(\sum_j W_{lj}^\dagger |j\rangle \right) \\
&= \sum_{k=1}^{\chi} \sigma_k |a_k\rangle |b_k\rangle
\end{aligned}
$$

with $|a_j\rangle \equiv \sum_{k=1}^{n} V_{ik} |i\rangle$ and $|b_j\rangle \equiv \sum_{k=1}^{m} W_{jl}^* |j\rangle$. The singular values are the Schmidt coefficients and the number of non-zero singular values is the Schmidt number.

C.3   Canonical angles

Canonical angles determine how far apart two spaces are from one another. It is a useful measure since it is unitarily invariant. This implies that the bases chosen to describe the two spaces do not matter. We will only consider spaces of equal dimension but it is not difficult to generalize to non-equal dimensions.

From the bases $\{|i\rangle\}$ and $\{|j\rangle\}$ for spaces $\mathcal{H}_A$ and $\mathcal{H}_B$, we define the unitary matrices A and B such that the columns of A (B) are the vectors $\{|i\rangle\}(\{|j\rangle\})$. The singular values of $C = A^\dagger B$ can be bounded by unity due to the submultiplicative of matrix norms ($||AB|| \leq ||A|| \, ||B||$). Since A and B are unitary, their singular values (the absolute value of the eigenvalues) are 1. Therefore, $||C|| \leq ||A^\dagger|| \, ||B|| = 1 \cdot 1 = 1$ where we have used the operator norm (largest singular value).

Thus, all the singular values of C are between zero and one, so for each singular value σ_i we can define an angle as $\cos\theta_i = \sigma_i$. This angle θ_i is called the canonical angle.

Appendix D

Deutsch's algorithm and Grover quantum search

D.1 Deutsch's algorithm

To illustrate how superposition can be used for computational advantage, the simple Deutsch algorithm [62] is presented. Deutsch's algorithm is a straightforward application of quantum computing that provides a speedup over classical computation. The problem it solves can be posed as follows: given a function $f : \{0,1\} \to \{0,1\}$ we want to know if the function is balanced or unbalanced. If $f(x)$ is balanced, then $f(0) = f(1)$. If $f(0) \neq f(1)$, then $f(x)$ is unbalanced. The function, f, is implemented in Deutsch's algorithm using an oracle.

An oracle is black box whose workings are not important but evaluates a useful function; in our case f. For example, consider the problem of factoring a large number, say N. The input to the oracle could be considered as two numbers and the oracle would simply perform multiplication and return one if the product is N and zero otherwise. Some problems may have a very complicated oracle that is 'expensive' to access. Any problem that requires the use of the oracle is optimized as to reduce the number of times one would have to visit. Characterizing the difficulty of the computation based on the number of queries is called query complexity. While many insights can be gained using this model, it will only be considered in this appendix, while the more general setting of time complexity was discussed in Chapter 4.

For Deutsch's algorithm, if we label one qubit *top* and the other *bottom* then we require an oracle that evaluates $f(x)$ by performing the following unitary transform: $|x\rangle_{top}|y\rangle_{bottom} \to |x\rangle_{top}|y \oplus f(x)\rangle_{bottom}$. In this context, $\oplus$ denotes addition modulus 2. Note that $x \to f(x)$ is not unitary if f is not bijective, therefore a single qubit oracle is disallowed in this context since we are considering balance functions. In Fig. D.1, the circuit for Deutsch's algorithm is drawn and we will examine the wave function at each point indicated in the diagram.

The algorithm is initialized with the state of the two qubits as $|0\rangle_t|1\rangle_b$. The Hadamard gates transform between the σ^z basis (the computational basis) and the σ^x basis and is written in matrix form in Fig. 3.3. The eigenstates of σ^x are $(|0\rangle \pm |1\rangle)/\sqrt{2}$, and, we
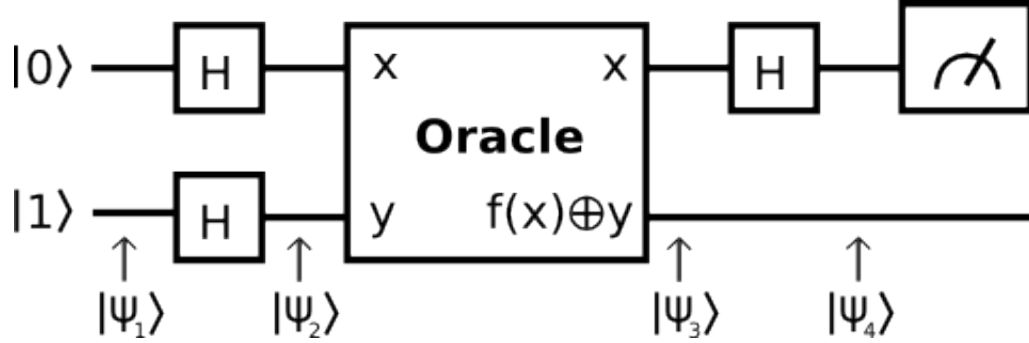
Figure D.1: Deutsch's algorithm is used to determine if a function computed by the oracle is balance or unbalanced. This algorithm requires only one query due to quantum parallelism. In the classical case, two calls to the oracle are always required.

will denote them $|X\pm\rangle$. Now, $|\psi_2\rangle = |X+\rangle|X-\rangle$. Observe the effect of the oracle on $|\psi_2\rangle$

$$
\begin{aligned}
|\psi_3\rangle &= U_o|\psi_2\rangle = U_o\frac{(|0\rangle + |1\rangle)_t}{\sqrt{2}}\frac{(|0\rangle - |1\rangle)_b}{\sqrt{2}} \\
&= \frac{1}{2}U_o(|00\rangle - |01\rangle + |10\rangle - |11\rangle)_{tb} \\
&= \frac{1}{2}(|0\rangle_t|f(0)\oplus 0\rangle_b - |0\rangle_t|f(0)\oplus 1\rangle_b + |1\rangle_t|f(1)\oplus 0\rangle_b - |1\rangle_t|f(1)\oplus 1\rangle_b) \\
&= \frac{|0\rangle_t}{2}(\,|f(0)\rangle - |\neg f(0)\rangle\,)_b + \frac{|1\rangle_t}{2}(\,|f(1)\rangle - |\neg f(1)\rangle\,)_b
\end{aligned}
\tag{D.1}
$$

Here, $\neg x$ is the negation of x e.g. if $x = 1$ then $\neg x = 0$.

Suppose the function is balanced and let $c = f(0) = f(1)$. In this case we have $|\psi_3\rangle = (1/2)(|0\rangle + |1\rangle)_t(|c\rangle - |\neg c\rangle)_b$. Considering each value of c shows that we can write

$$|\psi_3\rangle = (-1)^c|X+\rangle|X-\rangle$$

In the unbalanced case, we have that $|f(0)\rangle - |\neg f(0)\rangle = |\neg f(1)\rangle - |f(1)\rangle$. We define c' as $f(0)$ and immediately get $|f(0)\rangle - |\neg f(0)\rangle = (-1)^{c'}|-\rangle$. Substituting this into Eq. (D.1),

$$|\psi_3\rangle = \frac{|0\rangle_t - |1\rangle_t}{2}(-1)^{c'}|-\rangle_b = (-1)^{c'}|X+\rangle_t|X-\rangle_b$$

Applying the final Hadamard gate to the top qubit reveals

$$
|\psi_4\rangle = \left\{
\begin{array}{ll}
(-1)^{c''}|0\rangle|X-\rangle & \text{balanced} \\
(-1)^{c''}|1\rangle|X-\rangle & \text{unbalanced}
\end{array}
\right\}
$$

The value of c'' is not important as the expectation value of an observable is all that can be measured. Therefore global phases can be ignored, but *not* relative phases. If the first qubit is measured in the σ^z basis as zero, then the function is balanced; if it is one then, the function is unbalanced.

Thus, this algorithm only takes one call to the oracle, whereas the classical algorithm must make two calls. This is a rather simple example, but the importance of Deutsch's algorithm was not necessarily the practical application of quantum computing, but, to show an explicit separation between classical and quantum computation. Since the introduction of this algorithm, there have been various extensions of it in many different directions. Unfortunately, deciding if a function is balance or not is not of particular practical importance, so, now, attention turns to Grover quantum search, which has more practical applications.

D.2 Grover search

Grover quantum search algorithm, along with Shor's factoring algorithm [193], is one of the most important applications of quantum information science to date. Grover developed this algorithm in 1996 [80, 81], and, remarkably it was optimal for the task it set out to solve. This task was to identify a marked state using a quantum computer.

Database search, classical results, and oracles

The database search problem is to find a marked state among $N = 2^n$ unsorted items. The canonical example is that of finding a particular name when given a phone number and a phone book.

Classically, if we want to find items from a set, W, it will take $O(N)$ time to search if W only has one element. To formalize these ideas, suppose for database D we have function $f : D \rightarrow \{0, 1\}$ defined as

$$f(x) = \begin{cases} 1, & x \in W \\ 0, & x \notin W \end{cases} \tag{D.2}$$

In this section, we only discuss the case where one item is marked. To find marked item s, we will have to apply $f(x)$ to each $x \in \{0, 1, \cdots, N-1\}$. On average, this will take $N/2$ tries and the worst case it will take $N - 1$ tries. In other words, as the database increases in size, the time required to search it will scale linearly with the size of the database. The advantage of the quantum algorithm is only in the case of unsorted databases. If a database can be sorted according to some enumeration, the quantum speed-up is only a constant [55]. If the classical algorithm is allowed to spend $O(N \log N)$ time to sort the database, search algorithms such as binary search can be performed in $O(\log(N))$ queries and in some cases better.

The Grover quantum search algorithm uses discrete steps to evolve the uniform superposition of all states to the desired state in approximately $\sqrt{N}$ steps. Beginning the quantum search at the uniform superposition is different from starting from the uniform probability distribution due to the coherences between the states as explained in Section 3.3. In the classical search, one must randomly guess which item to query. The quantum algorithm proceeds by using the oracle, to change the sign of marked items as explained in the next subsection. A second operator reflects each amplitude across the mean of the wave function amplitude and is examined in Subsection D.2.2. The combination of the two operations is called a Grover operation, see Subsection D.2.3, and we will see that this operator

acts non-trivially in the two-dimensional space which is spanned by the initial state and the answer state. As the composition of two reflections is a rotation, the Grover operator has a quaint geometrical interpretation, which we will see in Subsection D.2.3.

D.2.1 Phase shift oracle

Using an oracle that computes function f from Eq. (D.2), we would like to create an operator that changes the sign of marked items. First, we give the form of a unitary oracle, second, we show how we can pick input states to create the desired operation. See Section 3.4 for any unfamiliar notation.

Now, suppose we had an oracle U_o that operates as $U_o|x\rangle|y\rangle \rightarrow |x\rangle|y \oplus f(x)\rangle$. The reason for the oracle to be implemented as a two qubit operation is that we wish it to be unitary. Each unitary operation is invertible, so each input must map to a unique output i.e. the operation must be injective (one-to-one). For functions that are not injective, using extra space to keep the input allows the operation to be reversible [22]. In our case, U_o calculates $f(x)$ corresponding to Eq. (D.2).

Note that if we let the second qubit, $|y\rangle$, be the state $|X-\rangle = (|1\rangle - |0\rangle)/\sqrt{2}$, then oracle changes the sign of an input vector if it is a marked state. Observe

$$U_o \left(\frac{|x\rangle|0\rangle - |x\rangle|1\rangle}{\sqrt{2}} \right) = \left(\frac{|x\rangle|f(x)\rangle - |x\rangle|1 \oplus f(x)\rangle}{\sqrt{2}} \right)$$

Temporarily ignoring normalization, either $f(x) = 0$ and the output state is $|x\rangle(|0\rangle - |1\rangle)$ or $f(x) = 1$ and the output state is $|x\rangle(|0\rangle - |1\rangle)$. Therefore, we write

$$U_o|x\rangle|X-\rangle = (-1)^{f(x)}|x\rangle|X-\rangle.$$

If the second qubit is always in state $|X-\rangle$, then the sign of answer state, $|w\rangle$, changes and the other $N - 1$ basis vectors are left unchanged. Thus,

$$U_o = -R_w = \mathbf{1} - 2|w\rangle\langle w| \tag{D.3}$$

Here, R_w is a reflection about the $|w\rangle$ state.

D.2.2 Reflection about the mean

In this subsection, the reflection about the mean is explained by, first, providing the definition and then providing the motivation for this definition. The diffusion operator, U_{mean}, is defined as

$$U_{mean} = R_s = 2|s\rangle\langle s| - \mathbf{1} \tag{D.4}$$

The state $|s\rangle$ is selected as a superposition over all database states including those of the target space. We pick

$$|s\rangle = \mathsf{H}^{\otimes n}|0 \cdots 0\rangle = \frac{1}{\sqrt{N}} \sum_{x=0}^{N-1} |x\rangle \tag{D.5}$$

Here, $N = 2^n$ is the normalization constant assuming there are n qubits in the array.

The operator U_{mean} with s defined as in Eq. (D.5) is called an inversion about the mean. First, define the mean of $|v\rangle$ as

$$\bar{v} = \sum_{k=0}^{N-1} \frac{\langle k|v\rangle}{N}$$

Starting from $|s\rangle\langle s|$ and inserting the resolution of the identity in the same basis as the answer state w, $\mathbf{1} = \sum |i\rangle\langle i|$,

$$|s\rangle\langle s|v\rangle = \sum_{i=0}^{N-1} |s\rangle\langle s|i\rangle\langle i|v\rangle = \sum_i \left(\frac{1}{\sqrt{N}}\sum_{j=0}^{N-1}|j\rangle\right)\left(\frac{1}{\sqrt{N}}\sum_{k=0}^{N-1}\langle k|\right)|i\rangle\langle i|v\rangle \quad \text{(D.6)}$$

$$= \frac{1}{N}\sum_{ijk}|j\rangle\langle k|i\rangle\langle i|v\rangle = \frac{1}{N}\sum_{ijk}\delta_{ki}\langle i|v\rangle|j\rangle = \sum_j \frac{\sum_i\langle i|v\rangle}{N}|j\rangle \quad \text{(D.7)}$$

$$= \sum_j \bar{v}|j\rangle \quad \text{(D.8)}$$

Putting it all together, U_{mean} performs the following operation: $U_{mean}|v\rangle \rightarrow \sum_j (2\bar{v} - v_j)|j\rangle$ with $v_j \equiv \langle j|v\rangle$. Looking at Fig. D.2, this operation has the form of an inversion.
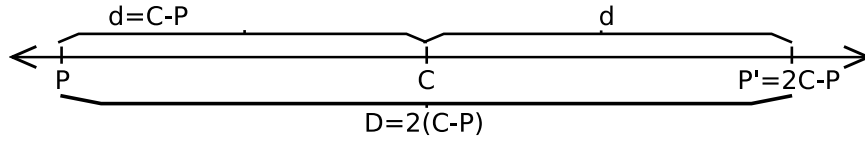

Figure D.2: **The reflection of point P about point C** results in point P'. The distance from P to C is $d = C - P$ and the distance from P to P' is $D = 2d = 2(C - P)$. If follows that $P' = P + 2(C - P) = 2C - P$. In equation (D.8), the average amplitude $\bar{v}$ becomes the center C and v_j can be identified as P.

D.2.3 The Grover operator and its two-dimensional operation

The Grover operator is given by

$$G = U_{mean}U_o = -R_s R_w \quad \text{(D.9)}$$

The Grover search works by evolving the state, $|s\rangle$, toward the answer state, $|w\rangle$, using only a two-dimensional subspace. The states that have no overlap with $|s\rangle$ and $|w\rangle$ are eigenstates of G with eigenvalue -1. We begin by, using Eqs. (D.3) and (D.4) to expand G as

$$G = 2|s\rangle\langle s| + 2|w\rangle\langle w| - 4\langle s|w\rangle|s\rangle\langle w| - \mathbf{1} \quad \text{(D.10)}$$

If $|t\rangle$ has no overlap with w or s then $G|t\rangle = -|t\rangle$. Although $|s\rangle$ is a superposition of all states in the database basis, the vectors $|w\rangle$ and $|s\rangle$ cannot span the entire space. An example of such a vector is any superposition of an even number of states with alternating sign. For instance, $|t\rangle = \mathsf{H}^{\otimes n}|1\cdots 1\rangle = (1/\sqrt{N})\sum(-1)^x|x\rangle$ is orthogonal to $|s\rangle$. If the

superposition $|t\rangle$ does not include the answer state, then t is in the $N-2$ dimensional space orthogonal to $\text{span}\{|w\rangle, |s\rangle\}$. In the case that $|t\rangle$ includes the answer state, beginning the search from $|t\rangle$ would perform equally well as $|s\rangle$.

The trajectory generated by repeated application of G on $|s\rangle$ explores this two-dimensional space of $\text{span}\{|w\rangle, |s\rangle\}$. To use an orthonormal basis, we construct a normalized state $|r\rangle$ orthogonal to $|w\rangle$ using a method like Gram-Schmidt decomposition [93].

$$|r\rangle = \sqrt{\frac{N}{N-1}}|s\rangle - \sqrt{\frac{1}{N-1}}|w\rangle = \sqrt{\frac{1}{N-1}}\sum_{x \neq w}|x\rangle \tag{D.11}$$

By construction, $\text{span}\{|w\rangle, |r\rangle\} = \text{span}\{|w\rangle, |s\rangle\}$. Now, rewrite the initial superposition in this basis as $|s\rangle = \sqrt{(N-1)/N}|r\rangle + N^{-1/2}|w\rangle$. If we consider the effect of G on $|r\rangle$ using Eqs. (D.10) and (D.11), then

$$G|r\rangle = (2|s\rangle\langle s| + 2|w\rangle\langle w| - 4\langle s|w\rangle|s\rangle\langle w| - \mathbf{1})|r\rangle \tag{D.12}$$

$$= \left(1 - \frac{2}{N}\right)|w\rangle + \frac{2\sqrt{N-1}}{N}|r\rangle \tag{D.13}$$

$$= \cos\theta|w\rangle + \sin\theta|r\rangle \tag{D.14}$$

Assuming that G is real, we know that G is unitary if and only if its two-dimensional matrix representation on the w, r space takes the following form

$$G = T(\theta) = \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix} \tag{D.15}$$

with θ defined through Eq. (D.14), assuming a matrix representation of arbitrary state $|v\rangle$ is given by

$$|v\rangle = \begin{bmatrix} \langle w|v\rangle \\ \langle r|v\rangle \end{bmatrix}$$

To characterize initial state s, let ϕ_r be the angle between $|s\rangle$ and $|r\rangle$. By inserting this angle of rotation into a rotation matrix, we can rotate a vector beginning in state $|r\rangle$ to state $|s\rangle$, and the following equation is satisfied

$$\begin{bmatrix} \cos\phi_r & \sin\phi_r \\ -\sin\phi_r & \cos\phi_r \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} \sin\phi_r \\ \cos\phi_r \end{bmatrix} = \begin{bmatrix} \langle w|s\rangle \\ \langle r|s\rangle \end{bmatrix}. \tag{D.16}$$

From Eq. (D.11), we derive that

$$|s\rangle = \sqrt{\frac{N-1}{N}}|r\rangle + \sqrt{\frac{1}{N}}|w\rangle = \sqrt{\frac{1}{N}}\begin{bmatrix} 1 \\ \sqrt{N-1} \end{bmatrix} \tag{D.17}$$

$$= \cos\phi_r|r\rangle + \sin\phi_r|w\rangle \tag{D.18}$$

and it follows $\cos\phi_r = \sqrt{(N-1)/N}$ and $\sin\phi_r = \sqrt{1/N}$.

Using geometric arguments, we can show that

$$2\phi_r = \theta \tag{D.19}$$

One proof using the double angle formula and Eq. (D.14) is given by

$$
\begin{aligned}
\sin 2\phi_r &= 2(\sin \phi_r)(\cos \phi_r) && \text{(D.20)} \\
&= 2\sqrt{\frac{N-1}{N}}\sqrt{\frac{1}{N}} && \text{(D.21)} \\
&= \sin \theta && \text{(D.22)}
\end{aligned}
$$

A summary of these geometric results is given in Fig. D.3.

Thus, after k iterations, the initial state with angle ϕ_r has been rotated to a vector characterized by angle $\Phi_k = k(2\phi_r) + \phi_r$. When Φ_k is close to $\pi/2$, we have accomplished the goal of creating high overlap with the answer. So, if N is large, then the state $|s\rangle$ is approximately $|r\rangle$. Then, we will need the following to be approximately satisfied.

$$
\begin{aligned}
\Phi_k = \frac{\pi}{2} &= (2k+1)\phi_r && \text{(D.23)} \\
\frac{\pi}{2\phi_r} &= 2k+1 && \text{(D.24)} \\
k &= \frac{\pi}{4\phi_r} - \frac{1}{2} && \text{(D.25)}
\end{aligned}
$$

From Eqs. (D.17) and (D.18) we have that $\sin \phi_r = 1/\sqrt{N}$. Using the small angle approximation, we say $\phi_r \approx 1/\sqrt{N}$. Substituting into Eq. (D.25) we have

$$
k = \frac{\pi}{4}\sqrt{N} - \frac{1}{2} \approx \frac{\pi}{4}\sqrt{N}. \tag{D.26}
$$

Hence, the algorithm obtains the answer with high probability after a number of iterations that is scales as $\sqrt{N}$.

D.2.4 Lower bound for Grover search

Briefly, note that Grover search is optimal with respect to the number of oracle calls it makes [177, 24]. Historically, its proof is important because of the technique that was used in the proof: the hybrid method.

Grover algorithm, which relies on the oracle to find a marked item in time $O(\sqrt{N})$ can be proved optimal by changing a critical subset of oracle outputs such that the errors cannot be detected in time $\Omega(\sqrt{N})$ [24]. This lower bound proof technique is a precursor to the widely used adversary method. The idea is an adversary makes modification of the true oracle on a subset of critical answers to trick you. Any quantum algorithms that access the oracle less than some lower bound is not sensitive to critical changes of the oracle. For the Grover search problem, if the oracle is accessed less than $O(\sqrt{N})$ times, then for any algorithm (Grover or otherwise), there is an oracle for which the algorithm will fail to distinguish the desired object from any other with high probability [24].
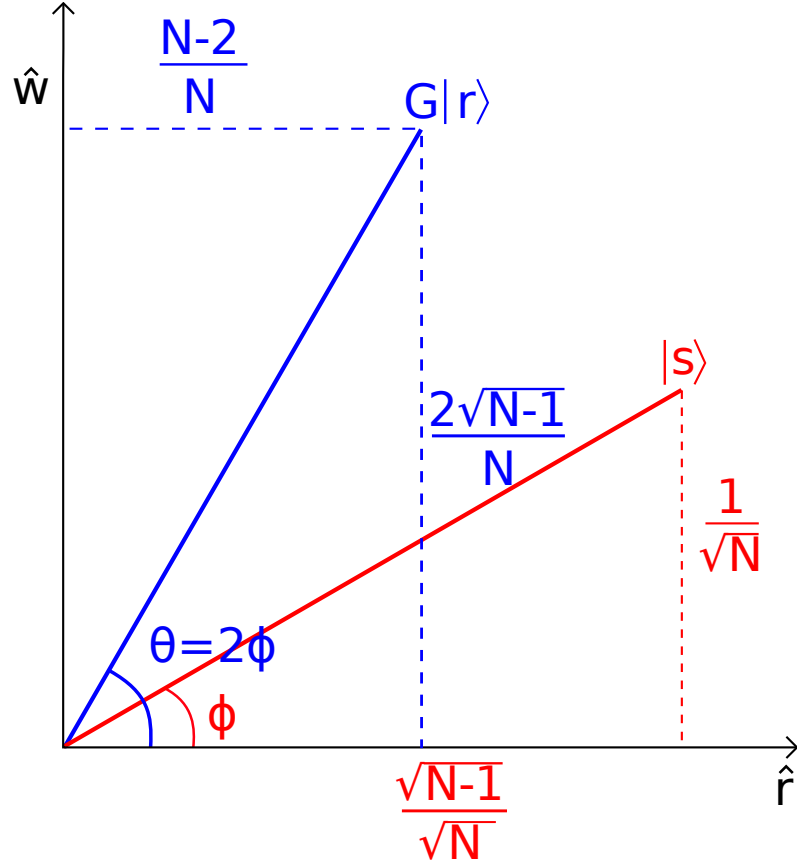
Figure D.3: **Grover search in the plane of the initial and target states.** To summarize Eqs. (D.14), (D.17), (D.18), and (D.22), we depict the states $G|r\rangle$ and $|s\rangle$. The answer state $|w\rangle$ is parallel to the ordinate axis. The abscissa is the orthogonal state $|r\rangle$ defined in Eq. (D.11). The uniform superposition of all states is $|s\rangle$ and is characterized in Eqs. (D.18) and (D.17). If ϕ_r is the angle between the $\hat{r}$ axis and $|s\rangle$ then the Grover operator G performs a rotation of $2\phi_r$. After $O(\sqrt{N})$ applications of the G to the state $|s\rangle$, there is high probability that measurement will result in marked state w.