DIGITAL ACCESS TO SCHOLARSHIP AT HARVARD
DASH.HARVARD.EDU

HARVARD LIBRARY
Office for Scholarly Communication

Algorithms Bridging Quantum Computation and Chemistry

Citation

McClean, Jarrod Ryan. 2015. Algorithms Bridging Quantum Computation and Chemistry. Doctoral dissertation, Harvard University, Graduate School of Arts & Sciences.

Permanent link

http://nrs.harvard.edu/urn-3:HUL.InstRepos:17467376

Terms of Use

This article was downloaded from Harvard University's DASH repository, and is made available under the terms and conditions applicable to Other Posted Material, as set forth at http://nrs.harvard.edu/urn-3:HUL.InstRepos:dash.current.terms-of-use#LAA

Share Your Story

The Harvard community has made this article openly available.
Please share how this access benefits you. Submit a story .

Accessibility

Algorithms Bridging Quantum Computation and Chemistry

a dissertation presented
by
Jarrod Ryan McClean
to
the Committee in Chemical Physics

in partial fulfillment of the requirements
for the degree of
Doctor of Philosophy
in the subject of
Chemical Physics

Harvard University
Cambridge, Massachusetts

May 2015

©2014 – Jarrod Ryan McClean
all rights reserved.

Dissertation advisor:
Professor Alán Aspuru-Guzik

Author:
Jarrod Ryan McClean

Algorithms Bridging Quantum Computation and Chemistry

Abstract

The design of new materials and chemicals derived entirely from computation has long been a goal of computational chemistry, and the governing equation whose solution would permit this dream is known. Unfortunately, the exact solution to this equation has been far too expensive and clever approximations fail in critical situations. Quantum computers offer a novel solution to this problem. In this work, we develop not only new algorithms to use quantum computers to study hard problems in chemistry, but also explore how such algorithms can help us to better understand and improve our traditional approaches.

In particular, we first introduce a new method, the variational quantum eigensolver, which is designed to maximally utilize the quantum resources available in a device to solve chemical problems. We apply this method in a real quantum photonic device in the lab to study the dissociation of the helium hydride (HeH^+) molecule. We also enhance this methodology with architecture specific optimizations on ion trap computers and show how linear-scaling techniques from traditional quantum chemistry can be used to improve the outlook of similar algorithms on quantum computers.

We then show how studying quantum algorithms such as these can be used to understand and enhance the development of classical algorithms. In particular we use a tool from adiabatic quantum computation, Feynman's Clock, to develop a new discrete time variational principle and further establish a connection between real-time

quantum dynamics and ground state eigenvalue problems. We use these tools to develop two novel parallel-in-time quantum algorithms that outperform competitive algorithms as well as offer new insights into the connection between the fermion sign problem of ground states and the dynamical sign problem of quantum dynamics.

Finally we use insights gained in the study of quantum circuits to explore a general notion of sparsity in many-body quantum systems. In particular we use developments from the field of compressed sensing to find compact representations of ground states. As an application we study electronic systems and find solutions dramatically more compact than traditional configuration interaction expansions, offering hope to extend this methodology to challenging systems in chemical and material design.

Contents

Citations to previously published work

At the time of writing, the work in Chapter 7 is submitted for publication, and the work in Chapters 2-6 has appeared, apart from minor alterations, in previous publications. The publication citations are highlighted at the beginning of the relevant chapters as title footnotes and additionally listed for reference here

Alberto Peruzzo†, Jarrod R. McClean†, Peter Shadbolt, Man-Hong Yung, Xiao-Qi Zhou, Peter J Love, Alán Aspuru-Guzik, and Jeremy L O'Brien. A new variational eigenvalue solver on a photonic quantum processor. *Nature Communications*, 5(4213):1-7, (2014).
† These authors contributed equally to this work.

Man-Hong Yung, Jorge Casanova, Antonio Mezzacapo, Jarrod R. McClean, Lucas Lamata, Alán Aspuru-Guzik, and Enrique Solano. From transistor to trapped-ion computers for quantum chemistry. *Scientific Reports*, 4(3589):1-7, (2014).

Jarrod R. McClean, Ryan Babbush, Peter J. Love, and Alán Aspuru-Guzik. Exploiting locality in quantum computation for quantum chemistry. *The Journal of Physical Chemistry Letters*, 5(24):4368-4380, (2014).

Jarrod R. McClean, John A Parkhill, and Alán Aspuru-Guzik. Feynman's clock, a new variational principle, and parallel-in-time quantum dynamics. *Proceedings of the National Academy of Sciences USA*, 110(41):E3901-E3909, (2013).

Jarrod R. McClean, Alán Aspuru-Guzik. Clock Quantum Monte Carlo: an imaginary-time method for real-time quantum dynamics. *Physical Review A* 91, 012311, (2015).

To my wife Anh, my parents, my family, and everyone else who has helped me become the person I am today.

Acknowledgments

It would be impossible for me to acknowledge all the people in my life to whom I am indebted, but I would like to begin by thanking my parents for their unwavering support. Their encouragement, love, and sacrifices are the reason I am who I am today. My amazing wife Anh has been a beacon of love and hope that lit up each day and broke through the storm clouds so often over Cambridge.

I also must thank my advisor Alán Aspuru-Guzik, who helped prove to me that imagination and hard work are the foundations that science is built on. I will always be grateful for the opportunities and wisdom he gave to me. My committee members, Eric Heller and Efthimios Kaxiras helped guide me and always had valuable insights.

Finally I wish to thank my friends and colleagues at Harvard for their support and insights. This, of course, includes all members of the Aspuru-Guzik group, past and present, whose equal I have never met in innovation and enthusiasm. To name only a few, I am especially grateful to Stéphanie Valleau, John Parkhill, Ryan Babbush, Dmitry Zubarev, James Whitfield, Thomas Markovich, Man-Hong Yung, and Peter Love. I also give special thanks to the exceptional Kebab Factory Crew and the rest of my graduate cohort.

1

Introduction

1.1 Background and motivation

To frame the themes and goals of this work, it is appropriate to begin with a famous quote by the renown physicist Paul Dirac [60],

> 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 leads to equations much too complicated to be soluble.

The physical law being referred to here is, of course, the Schrödinger equation, which governs non-relativistic quantum mechanics. The implication is that with sufficient

computational power, one could predict the whole of chemistry, ranging from small molecule synthesis to enzymatic catalysis and photosynthetic light harvesting. An efficient solution to this equation would revolutionize the computational design of drugs and materials and change the tools we have available to understand the physical world. Unfortunately, as alluded to by Dirac, the exact solution of these equations has remained practically untenable. Dirac finished his quote by saying,

> It therefore becomes desirable that approximate practical methods of applying quantum mechanics should be developed, which can lead to an explanation of the main features of complex atomic systems without too much computation.

And indeed, much progress has been made in approximate solutions of the Schrödinger equation. Many classes of systematically improvable wavefunction methods such as complete active space methods, configuration interaction, perturbation theory, coupled cluster, and density matrix renormalization group methods have been developed and used with great success [13, 17, 45, 84, 102, 107, 173, 192, 217, 249]. Simultaneously methods based entirely on the density, namely density functional theory, have continued to develop and progress to new heights of accuracy and cost effectiveness, in spite of the lack of obvious routes towards systematic improvement [62, 126, 194]. Despite these advances in approximation methods, there are still some cases where the accuracy that is feasibly attainable remains lacking.

Even today, the difficulty of accurately calculating the interactions of atoms in the simple exchange reaction $O_2 + O \rightarrow O + O_2$, has led to great confusion in interpreting its dynamics [203, 235, 258, 259]. The inaccuracies of this potential surface have led to a decade long debate including attempts to identify features such as "Reef-like" structures, and using the presence or absence of such structures to ascertain the qual-

ity of the surface. Moreover, such interpretations have led to erroneous conclusions of the basic physics, independent of the accuracy of quantitative assessment. This highlights the crucial need to make additional advances in this area. Without accurate determination of such surfaces, gaining a deeper understanding of the physical phenomena can be extremely challenging if not impossible.

It was Feynman who first suggested [70] that the difficulty arising in solving such equations might stem from the fact that we are trying to represent fundamentally quantum systems with classical ones. Consequently, a more natural solution might be to use quantum systems to represent quantum systems. This modest suggestion represents the foundation of what we know today as quantum computation, simulation, and information. The development of this field has led to many fundamental insights into the nature of quantum mechanics as well as new algorithms for problems seemingly unrelated to quantum mechanics, such as factoring large numbers [215].

Among the many developments in the general field of quantum computation, it was discovered that instances of the molecular Schrödinger might be solvable exactly (within a basis and to a fixed energetic precision), in a time that is exponentially faster than current classical algorithms as a function of the system size [6]. Since the initial proposal of the idea, there have been numerous proof of principle quantum experiments demonstrating its feasibility as well as many algorithmic extensions and variations [8, 111, 136, 166, 196, 209, 242, 261]. In fact, at present it is believed that the solution of the molecular Schrödinger equation may represent one of the first practical uses of a quantum computer that exceeds the computational capacity of modern classical supercomputers [82].

The work of this thesis began at this foundation in understanding how one might use a quantum computer to better understand and simulate chemistry at a fundamen-

tal level. From this cornerstone, we developed a new quantum algorithm capable of utilizing any quantum device to its maximum capability in studying quantum eigenvalue problems [196]. We then helped to extend this algorithm to new architectures as well as import ideas from traditional quantum chemistry to further enhance its efficacy [166, 261]. In the study of quantum algorithms, we gained new insight into the traditional methods of simulating quantum chemistry. From these insights, we were inspired to write down a new quantum variational principle connecting time dynamics to ground state eigenvalue problems [167]. This formulation facilitated not only the development of time-parallel algorithms, but a new way to stochastically sample space-time paths with a discrete form of quantum Monte Carlo [165]. In studying the performance of this algorithm, we were inspired to consider a different approach to the traditional simulation of electronic structure problems. This method uses compressed sensing techniques to find simple representations of electronic wavefunctions [164]. The continuous cross-pollination between quantum computation and classical computation has offered numerous insights into both fields, and this is not a phenomenon likely to stop soon. We have learned much in navigating this bridge between two fields, and believe there is much yet to be learned.

In this dissertation, we begin with a short introduction of the problems of quantum mechanics as well as the notation and background required for understanding this work. We then give a brief outline of each of the subsequent chapters, putting each into the context of the greater theme. This is followed by the detailed research chapters, representing the publications on these topics we have made throughout the duration of this thesis research. Finally we conclude by tying the works together and discussing the outlook for the bridge between quantum computation and quantum chemistry.

1.2 Technical background and notation

1.2.1 Quantum states and the enormity of Hilbert space

In non-relativistic quantum mechanics, the state of a physical system is represented by a vector in a complete vector space V equipped with an inner product, or Hilbert space [53]. This vector is also called the wavefunction of the physical system, and contains all possible information about that physical system. Throughout this work we will make use of the Bra-Ket notation of Dirac, where state vectors are denoted $|\Psi\rangle$, and the inner product with another state vector $|\Phi\rangle$ is denoted $\langle\Phi|\Psi\rangle$. Observable quantities are given by Hermitian operators $\hat{O}$, and the expectation value of an operator a state vector $|\Psi\rangle$, or average value expected upon repeated measurement of the observable, is given by $\langle\Psi|\,\hat{O}\,|\Psi\rangle$.

After choosing an orthonormal basis $B = \{|\phi_i\rangle\}_i$, i.e. $\langle\phi_i|\phi_j\rangle = \delta_{ij}$, for the vector space of the physical system V, one can express any state of the physical system $|\Psi\rangle$ in terms of this basis as

$$|\Psi\rangle = \sum_i c^i\,|\phi_i\rangle \tag{1.1}$$

where $c^i \in \mathbb{C}$ and one may write the matrix representation of a linear operator $\hat{O}$ as

$$O = \begin{pmatrix} \langle\phi_1|\,\hat{O}\,|\phi_1\rangle & \langle\phi_1|\,\hat{O}\,|\phi_2\rangle & \langle\phi_1|\,\hat{O}\,|\phi_1\rangle & \dots \\ \langle\phi_2|\,\hat{O}\,|\phi_1\rangle & \langle\phi_2|\,\hat{O}\,|\phi_2\rangle & \langle\phi_2|\,\hat{O}\,|\phi_3\rangle & \\ \langle\phi_3|\,\hat{O}\,|\phi_1\rangle & \langle\phi_3|\,\hat{O}\,|\phi_2\rangle & \langle\phi_3|\,\hat{O}\,|\phi_3\rangle & \\ \vdots & & & \ddots \end{pmatrix} \tag{1.2}$$

or more succinctly $[O]_{ij} = \langle \phi_i | \hat{O} | \phi_j \rangle$.

Suppose that we have a quantum system that belongs to a space V with basis $B = \{|\phi_i\rangle\}_i$ and another that belongs to a space V' with basis $B' = \{|\phi'_i\rangle\}_i$. This situation is common when considering many-particle quantum systems, where each particle has a well defined Hilbert space such as V. The composite system lives in the tensor product space $V \otimes V'$, and is spanned by the basis $B \otimes B' = \{|\phi_i\rangle \otimes |\phi'_j\rangle = |\phi_i\rangle |\phi'_j\rangle = |\phi_i \phi'_j\rangle\}_{ij}$. As such, any state of the composite system can be written in terms of this basis as

$$|\Psi\rangle = \sum_{ij} c^{ij} |\phi_i \phi'_j\rangle \tag{1.3}$$

where $c^{ij} = c$ is a complex two-index tensor completely defining the quantum state. We will sometimes call this the defining or coefficient tensor of the state. Note that c^{ij} is typically subject to the constraint of normalization and all physical observables are independent of the global phase. As such, if V has dimension M and V' has dimension M', the tensor c^{ij} contains $2MM' - 2$ real degrees of freedom.

Before generalizing to N particles, we consider an example of two qubits, which are simply two-level quantum systems. Equivalently these may be thought of as 2 spin-$\frac{1}{2}$ particles or 2 qubits. Each qubit has a standard basis, sometimes called the computational basis, given by

$$|0\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \tag{1.4}$$

$$|1\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix}. \tag{1.5}$$

In this case, $V = V'$, and any state of the two qubit system may be written as

$$|\Psi\rangle = c^{00}|00\rangle + c^{10}|10\rangle + c^{01}|01\rangle + c^{11}|11\rangle = \sum_{ij} c^{ij}|ij\rangle \tag{1.6}$$

As in the more general case, the state $|\Psi\rangle$ is, of course, defined completely by the tensor c. A property of tensors such as c that will be important throughout this work, is the concept of canonical- or separation-rank [66, 88, 92, 125, 127]. The separation-rank is the minimal r such that we may express the tensor as a sum of products of 1 index tensors. For the two-qubit example given here, this is denoted

$$c = \sum_{k}^{r} a^k v_1^k \otimes v_2^k \tag{1.7}$$

where $v_i^j \in \mathbb{C}^M$ and $a^k \in \mathbb{C}$. In the case of two subsystems (or indices in c), the separation rank is exactly equivalent to the more familiar matrix rank, however this concept will generalize naturally to an arbitrary number of indices. To make this decomposition more concrete, consider the following two states of our two-qubit system

$$|\Psi_s\rangle = \frac{1}{2}\left(|00\rangle + |01\rangle + |10\rangle + |11\rangle\right) = \sum_{ij} c_s^{ij}|ij\rangle \tag{1.8}$$

$$|\Psi_e\rangle = \frac{1}{\sqrt{2}}\left(|00\rangle + |11\rangle\right) = \sum_{ij} c_e^{ij}|ij\rangle\,. \tag{1.9}$$

In this case, we may write the defining tensor of $|\Psi_s\rangle$, or c_s as

$$c_s = \frac{1}{2}\begin{pmatrix} 1 & 1 \end{pmatrix} \otimes \begin{pmatrix} 1 & 1 \end{pmatrix} \tag{1.10}$$

Thus it has separation-rank $r = 1$, or is a separable tensor. A separable defining

tensor is the hallmark of what is more typically referred to as a separable state in quantum computation and quantum information, and has the physical property that projective measurement on any subsystem (such as a single qubit) will not affect the probability of measurement outcomes on the other subsystem [183]. A quantum system that is not separable is called entangled, and is equivalently defined as a tensor with separation rank $r > 1$. Consider for example the decomposition of the defining tensor of $|\Psi_e\rangle$, c_e,

$$c_e = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 0 \end{pmatrix} \otimes \begin{pmatrix} 1 & 0 \end{pmatrix} + \frac{1}{\sqrt{2}} \begin{pmatrix} 0 & 1 \end{pmatrix} \otimes \begin{pmatrix} 0 & 1 \end{pmatrix} \tag{1.11}$$

which has separation rank $r = 2$, and is entangled. A defining physical property of entangled states is that projective measurement on one subsystem may affect the measurement outcome probabilities for other subsystems.

Moving now to a more general system of N qubits, one finds that the quantum state lives in the vector space $\bigotimes_{i=1}^{N} V_i$ and any state may be written as

$$|\Psi\rangle = \sum_{i_1, i_2, \ldots, i_N} c^{i_1, i_2, \ldots, i_N} |i_1 i_2 \ldots i_N\rangle \tag{1.12}$$

where the defining tensor of the state c now has N indices, and for a simple 2-qubit system has on the order of 2^N complex entries. This brings us to the central problem in the simulation and study of many-body quantum systems, the enormous size of c for an N particle system. Consider a simple collection of 256 qubits. The size of the tensor c in this case is roughly $2^{256} \approx 10^{80}$ which is on the order of the number of particles in the known universe. Even if one could enumerate 1 quadrillion entries per second, it would take orders of magnitude longer than the current age of the universe

(which is estimated to be $\approx 10^{17}$ seconds) to write down an arbitrary state, much less perform meaningful computation on it.

The central goals of this work will be concerned with making progress in the study of many body quantum systems, despite the daunting size of c. A central guiding philosophy and strategy for such work, is that the universe may not prepare arbitrary quantum states, and instead the subspace of physical states (sometimes called the physical corner of Hilbert space), may be small enough to be tractable [65, 79, 199]. In classical simulation methods, one attempts to find efficient approximations to the tensor c, such as the low rank decomposition described before. In quantum simulation, one takes a novel approach, which is to forfeit detailed knowledge of the tensor c and allow a different quantum system (perhaps a quantum computer) to naturally explore this gargantuan space. It is likely that both approaches will have some strengths and some weaknesses, and only by further study will we come to understand them. Despite the strong promises of quantum computers, it is perhaps prudent to keep in mind the wisdom of renowned mathematician John von Neumann,

> Truth [...] is much too complicated to allow anything but approximations.

1.2.2 The Schrödinger Equation and (Imaginary) Time Evolution

The central equation governing the dynamics of quantum systems is the time-dependent Schrödinger equation, which in atomic units ($\hbar = 1$) is written

$$i\partial_t \left|\Psi(t)\right\rangle = H(t) \left|\Psi(t)\right\rangle \tag{1.13}$$

where $H(t)$ is a Hermitian operator called the Hamiltonian and is the generator of time dynamics. In the special case where $H(t)$ is time-independent, the time-dependence

can be separated from the equation to yield the time-independent Schrödinger equation,

$$H\left|\chi_k\right\rangle = E_k\left|\chi_k\right\rangle \tag{1.14}$$

which is an eigenvalue equation for the operator H, where $|\chi_k\rangle$ are the eigenvectors (or eigenstates) and E_k are the eigenvalues (or eigenenergies). One often orders the eigenvalues by their value, such that $E_0 < E_1 \leq ...E_k$. The eigenvectors corresponding to the few lowest eigenenergies are frequently of primary interest, as physical systems at thermal equilibrium tend to predominantly occupy the lowest state and those reasonably accessible at thermal energy scales, i.e. $E_k - E_0 < k_b T$, which is roughly 0.59 kcal/mol at room temperature.

One can use the knowledge that we are often only interested in the lowest few eigenstates to design specialized algorithms for finding and describing them. One such algorithm is imaginary time evolution [29, 52, 94, 139, 208, 238]. Consider the time-dependent Schödinger equation with a time-independent Hamiltonian after a transformation $t \rightarrow i\tau$. This transformation is sometimes called a Wick rotation into imaginary time, and the new equation is given by

$$\partial_\tau\left|\Psi(\tau)\right\rangle = -H\left|\Psi(\tau)\right\rangle \tag{1.15}$$

and has a formal solution

$$\left|\Psi(\tau)\right\rangle = e^{-H\tau}\left|\Psi(0)\right\rangle \tag{1.16}$$

To see how this equation can be used to find the lowest energy eigenstate, consider

an arbitrary initial wave function $|\Psi(0)\rangle$. Any state of the system may be formally decomposed into the eigenstates of the Hamiltonian, such that

$$|\Psi(0)\rangle = \sum_i a_i |\chi_i\rangle \tag{1.17}$$

and by using the property that for an analytic function f of an operator H with eigenstate and eigenenergy $|\chi_k\rangle$, E_k, $f(H)|\chi_k\rangle = f(E_k)|\chi_k\rangle$, we see that evolution of this state in imaginary time is given by

$$\begin{aligned}\psi(\tau) &= e^{-H\tau}|\Psi(0)\rangle \\ &= e^{-H\tau}\left(\sum_i a_i |\chi_i\rangle\right) \\ &= \sum_i a_i e^{-E_i\tau}|\chi_i\rangle\end{aligned} \tag{1.18}$$

such that in the limit $\tau \to \infty$, we find $|\Psi(\tau)\rangle \to |\chi_0\rangle$. Thus if one can find a method to evolve a state in imaginary time, eventually it will converge to the ground state, assuming the initial state was not strictly orthogonal to it. A number of methods have been developed to achieve this in practice, and for discrete quantum systems a popular approach is the repeated use of the linearized short-time propagator [269]. That is, one makes the identification

$$e^{-H\tau} = \prod_{i=1}^{N} e^{-H\frac{\tau}{N}} \approx \prod_{i=1}^{N}\left(I - H\frac{\tau}{N}\right) \tag{1.19}$$

and then finds a prescription for application of the linearized short time propagator $\left(I - H\frac{\tau}{N}\right)$. This formulation is convenient in that it only requires knowledge of how to apply the Hamiltonian to a quantum state, and results in a final state ($\tau \to \infty$)

that is totally unbiased from the linearization under some loose conditions on the time step $\frac{\tau}{N}$ relative to the spectrum of the system. This prescription is the basis for the imaginary time evolution used in both the Clock Quantum Monte Carlo [165] and NOMAGIC [164] algorithms described later in this dissertation.

1.2.3 Canonical Decompositions and Matrix Product States

In the previous section, we showed briefly for two-qubits a form of the canonical rank decomposition, sometimes called the CANDE-PARFAC or CP decomposition [66, 88, 92, 125, 127]. We will now consider the generalization to N particles, and instead of focusing on the coefficient tensor, we can write this decomposition in a more familiar way with bra-ket notation of tensor products of single particle states. Consider a quantum system with M possible single particle states, and N copies of that system. Suppose that we have a basis for the single particle system denoted by $\{|\chi_i\rangle\}_i$, and we fashion single particle states of the form

$$|\phi_i^k\rangle = \sum_j b_{ik}^j |\chi_j\rangle . \tag{1.20}$$

We can write the canonical decomposition of an N particle state $|\Psi\rangle$ of rank r as

$$|\Psi_{CP}\rangle = \sum_{k=1}^{r} a^k |\phi_1^k \phi_2^k ... \phi_N^k\rangle \tag{1.21}$$

As before, we call a state separable if $r = 1$ and entangled if $r > 1$. We will use this form of the CP decomposition in the construction of the NOMAGIC method, and also later specify it to the case of indistinguishable systems.

This particular decomposition has not received as much attention as orthogonal ex-

pansions or tensor-network decompositions in general quantum physics. For that reason, we quickly highlight the connection between the CP decomposition and matrix product states, which are the implicit ansatz of the density matrix renormalization group method [45, 249]. The connection is quite straightforward and noted abstractly in tensor references [92] but does not seem to be well recognized in physics or chemistry literature. Consider the expansion of $|\Psi_{CP}\rangle$ into its canonical tensor representation, and for simplicity of notation absorb the coefficients a^k into the state such that

$$|\Psi_{CP}\rangle = \sum_{k=1}^{r} |\phi_1^k \phi_2^k ... \phi_N^k\rangle \tag{1.22}$$

$$= \sum_{k=1}^{r} \sum_{j_1, j_2, ..., j_N}^{M} b_{1k}^{j_1} b_{2k}^{j_2} ... b_{Nk}^{j_N} \, |\chi_{j_1} \chi_{j_2} ... \chi_{j_N}\rangle \tag{1.23}$$

Define a trivial diagonal matrix B_j^i as

$$B_j^i = \begin{pmatrix} b_{j1}^i & 0 & 0 & ... & \\ 0 & b_{j2}^i & 0 & & \\ 0 & 0 & b_{j3}^i & & \\ \vdots & & & \ddots & \\ & & & & b_{jr}^i \end{pmatrix} \tag{1.24}$$

which is a square matrix of dimension r. We may rewrite the above state using these matrices as

$$|\Psi_{CP}\rangle = \sum_{j_1 j_2 ... j_M} \mathrm{Tr}\left[B_1^{j_1} B_2^{j_2} ... B_N^{j_N}\right] |\chi_{j_1} \chi_{j_2} ... \chi_{j_N}\rangle \tag{1.25}$$

which is exactly the definition of a finite matrix product state with uniform bond di-

mension r. Recall that matrix product states are invariant under a transformation by an invertible matrix X, such as $B^i_j B^k_{j+1} \rightarrow B^i_j X X^{-1} B^k_{j+1}$. Thus, states which can be decomposed with separation rank r are equivalent to matrix product states with bond dimension r that can be made totally diagonal by invertible transformations X. Moreover, this form helps to highlight the lack of geometric dependence in the CP decomposition, as although diagonal matrices that are not multiples of the identity do not commute with all matrices, they do commute with eachother.

1.2.4 Antisymmetric Systems

Until this point, we have considered the most general quantum systems possible. Now it is prudent to spend some time specializing to the case of a set of N indistinguishable, antisymmetric particles. This is, of course, crucial to the study of electrons in molecules.

The wavefunction of N electrons must be totally antisymmetric with respect to the exchange of particles. We will consider the state of the electron to be represented by single particle functions $\{|\phi_i\rangle\}_i$ expressed in a basis of M single particle functions $\{|\chi^i_j\rangle\}_{ij}$. A number of approaches can be used to enforce the desired antisymmetry in the wavefunction. For example, we may write the state with a standard tensor expansion

$$|\Psi_A\rangle = \sum_{i_1, i_2, \ldots, i_N} c^{i_1 i_2 \ldots i_N} |\chi^1_{i_1} \chi^2_{i_2} \cdots \chi^N_{i_N}\rangle \tag{1.26}$$

under the constraint that $c^{i_1 i_2 \ldots i_N}$ is totally antisymmetric under exchange of indices. This approach has the advantage that makes clear the relation between the space of N distinguishable and N antisymmetric particles. Namely, that the space of antisym-

metric particle is a subspace of the space of distinguishable particles. However, working with this construction in practical calculations can be somewhat cumbersome. One alternative to this approach is to work in a basis of antisymmetric component functions. That is, much like the CP decomposition before, we have an antisymmetric CP decomposition

$$|\Psi_A^{CP}\rangle = \sum_{k=1}^{r} a^k \mathcal{A}\left(|\phi_1^k \phi_2^k ... \phi_N^k\rangle\right) \tag{1.27}$$

where $\mathcal{A}$ is the antisymmetrization operator. These antisymmetrized tensor products are taken to be so-called Slater Determinants, and the manipulation of such objects is well studied in quantum chemistry. We will use these antisymmetric component functions in the implementation of the NOMAGIC algorithm for electronic systems later in this work.

1.2.5 Quantum computation in brief

First conceptualized by Richard Feynman [70], quantum computation is the idea of encoding and processing information with a quantum system. The number of possible physical systems one could use to encode and process this information is enormous [183], ranging from entangled photons [8] and ions [95] to superconducting circuits [58]. As such, to make progress algorithmically, it is beneficial to abstract away the physical implementation and speak in the language of qubits, the quantum counterpart of bits.

A qubit is defined as a controllable two level quantum system, with basis states $|0\rangle$

and $|1\rangle$. A convenient vector representation for these states is given by

$$|0\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \tag{1.28}$$

$$|1\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix}. \tag{1.29}$$

Ideal, reversible actions on qubits are called gates, and transformations of the qubits are unitary as dictated by evolution under the time-dependent Schrödinger equation. Unitary operators are generated by the algebra of antihermitian operators, and as such any single qubit gate can be parametrized by

$$U = \exp\left(-i \sum_i \alpha_i \sigma_i\right) \tag{1.30}$$

where α_i are real numbers and σ_i are the standard Pauli matrices that constitute a basis for 2×2 Hermitian matrices.

$$\sigma_0 = I = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \tag{1.31}$$

$$\sigma_1 = X = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \tag{1.32}$$

$$\sigma_2 = Y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \tag{1.33}$$

$$\sigma_3 = Z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \tag{1.34}$$

where in defining the Pauli matrices, we have also given their common designations when used as single qubit gates X, Y, and Z. A few other useful single qubit gates that play a prominent role in the construction of circuits are the Hadamard gate H, the T-gate, and rotation gate $R(\theta)$,

$$H = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \tag{1.35}$$

$$T = \begin{pmatrix} 1 & 0 \\ 0 & e^{i\pi/4} \end{pmatrix} \tag{1.36}$$

$$R(\theta) = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix} \tag{1.37}$$

In order to perform useful computations, it is also necessary to perform conditional actions on qubits based on the states of other qubits. The simplest of such gates (conceptually), is the controlled-NOT or CNOT gate. This gate performs a NOT (or X) gate on a target qubit based on the state of the control qubit. It has operator and matrix representations given by

$$\text{CNOT} = |0\rangle\langle 0| \otimes I + |1\rangle\langle 1| \otimes X = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix} \tag{1.38}$$

A gate set is called universal, if by successive application of the gates in the set on different qubits an arbitrary unitary on n qubits can be performed to a specified precision. The gate set $\{H, T, \text{CNOT}\}$ is one such universal set. A convenient notation

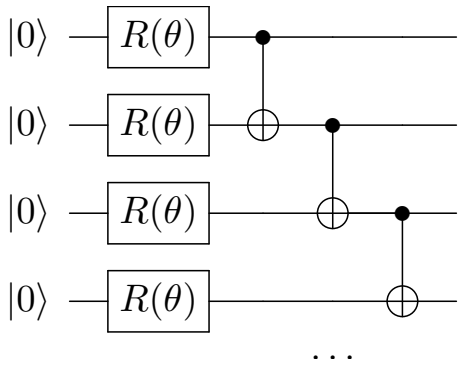

Figure 1.1: An example quantum circuit diagram. These diagrams are read from left to right, where horizontal lines denote a particular qubit, boxes represent a quantum gate, and lines shows the control relationship for multi-qubit gates. Shown here is a series of one-qubit rotations parameterized by an angle theta, $R(\theta)$, followed by CNOT gates, where the solid dot represents the control bit and the cross represents the target bit.

that can be used to express algorithms are quantum circuit diagrams, an example of which is depicted in Fig. 1.1.

1.2.6 Electronic Hamiltonians

While the aim of almost all the methods in this work is to be generally applicable to all quantum systems, electronic and chemical Hamiltonians are an application of special interest. A non-relativistic system composed of N_n nuclei and N_e electrons neglecting weak spin-spin interactions is defined by a Hamiltonian composed of the the kinetic energy and Coulomb interactions of the charged particles. In atomic units we write this Hamiltonian as

$$H = \sum_{i=1}^{N_e} \frac{-\nabla_{r_i}^2}{2} + \sum_{i=1}^{N_n} \frac{-\nabla_{R_i}^2}{2M_i} + \sum_{i,j<i}^{N_n} \frac{Z_i Z_j}{|R_i - R_j|} - \sum_{ij}^{N_n,N_e} \frac{Z_i}{|R_i - r_j|} + \sum_{i,j<i}^{N_e} \frac{1}{|r_i - r_j|} \quad (1.39)$$

where R_i, M_i, Z_i are nuclear positions, masses, and charges respectively, and the electronic positions are given by r_i. Due to the differences in mass and time scales be-

tween the electrons and nuclei, this Hamiltonian can be well-approximated (in non-dynamical problems) by assuming that the nuclei are fixed classical point charges. This approximation is known as the Born-Oppenheimer approximation, and yields a purely electronic Hamiltonian that depends only parametrically on the position of the nuclei as

$$H = \sum_{i,j<i}^{N_n} \frac{Z_i Z_j}{|R_i - R_j|} + \sum_{i=1}^{N_e} \frac{-\nabla_{r_i}^2}{2} - \sum_{ij}^{N_n,N_e} \frac{Z_i}{|R_i - r_j|} + \sum_{i,j<i}^{N_e} \frac{1}{|r_i - r_j|}. \tag{1.40}$$

It is also known, both from experimental inference and later the spin-statistics theorem of quantum field theory, that electrons must be antisymmetric with respect to exchange. One way of dealing with the antisymmetry is through explicit constraints on wavefunctions, as in previous sections where we references antisymmetric coefficient tensor constraints or antisymmetric components. Another way is to include antisymmetry in the Hamiltonian through the use of second quantization. In second quantization one makes use of the operator algebra of fermion creation and annihilation operators, $a_i^\dagger$ and a_j, to account for antisymmetry. Defining the anti-commutator $\{A, B\} := AB + BA$, these satisfy the fermion anti-commutation relations

$$\left\{a_i^\dagger, a_j^\dagger\right\} = 0 \tag{1.41}$$

$$\{a_i, a_j\} = 0 \tag{1.42}$$

$$\left\{a_i^\dagger, a_j\right\} = \delta_{ij} \tag{1.43}$$

Suppose that we discretize the electronic Hamiltonian in a single particle basis

$\{|\varphi_i\rangle\}_i$. The Hamiltonian in this basis, now including antisymmetry, may be written

$$H = \sum_{ij} h_{ij} a_i^\dagger a_j + \frac{1}{2} \sum_{ijkl} h_{ijkl} a_i^\dagger a_j^\dagger a_k a_l \tag{1.44}$$

where the coefficients of the operators are defined as integrals of the interaction terms over this single particle basis. More explicitly,

$$h_{pq} = \int d\sigma \; \varphi_p^*(\sigma) \left(-\frac{\nabla_r^2}{2} - \sum_i \frac{Z_i}{|R_i - r|} \right) \varphi_q(\sigma) \tag{1.45}$$

$$h_{pqrs} = \int d\sigma_1 \; d\sigma_2 \; \frac{\varphi_p^*(\sigma_1)\varphi_q^*(\sigma_2)\varphi_s(\sigma_1)\varphi_r(\sigma_2)}{|r_1 - r_2|} \tag{1.46}$$

where σ_i now contains the spatial and spin components of the electron, $\sigma_i = (r_i, s_i)$.

This form of the electronic Hamiltonian has been particularly useful in the development of correlated electronic structure calculations. In this work, we will use it to help formulate the electronic structure problem on quantum computers. In order to do this, we first need to map this problem, which is in the language of indistinguishable fermions, to the language of qubits, or distinguishable two-level systems. There are now at least 3 known isomorphisms that may be used to accomplish this task, the Jordan-Wigner, Parity, and Bravyi-Kitaev mappings [35, 112, 209]. In this dissertation, we primarily utilize the Jordan-Wigner transformation that is defined by

$$a_p^\dagger = (\prod\nolimits_{m<p} \sigma_m^z)\sigma_p^+ \tag{1.47}$$

$$a_p = (\prod\nolimits_{m<p} \sigma_m^z)\sigma_p^- \tag{1.48}$$

$$\sigma^\pm \equiv (\sigma^x \pm i\sigma^y)/2 \tag{1.49}$$

As a result, the electronic Hamiltonian becomes highly non-local in terms of the Pauli

operators $\{\sigma^x, \sigma^y, \sigma^z\}$ in the distinguishable qubit basis. However, an important aspect of this transformation that will be used later, is that despite this non-locality, the number of terms in the Hamiltonian is conserved up to a constant factor. Thus if there are $O(M^4)$ terms in the original fermionic Hamiltonian, where M is the number of single particle basis functions used to discretize the Hamiltonian, then there will be $O(M^4)$ terms in distinguishable qubit representation of the Hamiltonian.

1.3 Chapter outline and summaries

1.3.1 Chapters 2, 3, and 4

In the first chapters of this thesis, we will describe, and subsequently improve upon, a new quantum algorithm for the study of quantum chemistry with minimal experimental requirements. This method is called the Variational Quantum Eigensolver [196], and can be applied to general eigenvalue problems on a quantum computer, but our specific application goals at the time were focused on quantum chemistry.

Almost all quantum algorithms to date have been developed agnostic to the available hardware. That is, the approach is write down the best possible algorithm with regards to cost for attaining the result one desires, assuming that one day a quantum device will be capable of running that algorithm. However, many of the algorithms that perform optimally in the asymptotic limit of size require extraordinary resources for systems of interest.

For this reason, we proceeded with a different, co-design approach to quantum computation. We consider both the problem and the available architecture simultaneously, to achieve an optimal solution for the given hardware. This is done by separating the algorithm into components that exploit the strengths of a quantum computer

and components that can be trivially performed on a classical computer, so as not to waste expensive quantum resources.

To describe how we formulate this approach, we return to the variational principle of quantum mechanics. This principle states that for a Hermitian operator H with eigenvectors and eigenvalues $|\Psi_i\rangle, E_i$ ordered by value, that any approximate wavefunction $|\Psi_T\rangle$ (obeying necessary symmetries) satisfies

$$\frac{\langle\Psi_T| H |\Psi_T\rangle}{\langle\Psi_T|\Psi_T\rangle} \geq E_0 \tag{1.50}$$

Thus if we can create some state $|\Psi_T\rangle$ based on a set of parameters $\{\theta_i\}$, we can improve the quality of the approximation to the ground state by choosing parameters that minimize the expectation value of the energy.

The preparation of complex quantum states based on some set of parameters is an area where quantum computers excel. Indeed any parametrizable sequence of gates or repeatable sequence of quantum operations can be considered as a valid trial state preparation. We call this general idea, the "quantum hardware ansatz", and it allows one to use the available hardware to define the limits of the simulation. On a theoretical side, we have investigated the preparation of parameterizable ansatz states that we believe to be both high-quality and not efficiently prepared or sampled from on a classical computer. We believe multi-reference unitary coupled cluster [225] to be a strong candidate in this regard.

Once a state is prepared, one needs a way to evaluate the energy in order to improve the parameterization. One potential solution is to use quantum phase estimation, but this returns us to techniques which are, at the time of writing, experimentally inaccessible. Instead, one can use the linearity of quantum mechanics to write

the expectation value of the electronic Hamiltonian as

$$\langle \Psi_T | H | \Psi_T \rangle \equiv \langle H \rangle = \sum_{ij} h_{ij} \langle a_i^\dagger a_j \rangle + \sum_{ijkl} h_{ijkl} \langle a_i^\dagger a_j^\dagger a_k a_l \rangle \tag{1.51}$$

$$= \sum_{i\alpha} g_\alpha^i \langle \sigma_\alpha^i \rangle + \sum_{ij\alpha\beta} g_{\alpha\beta}^{ij} \langle \sigma_\alpha^i \sigma_\beta^j \rangle + ... \tag{1.52}$$

where the first line shows that the energy is explicitly obtainable through the two-electron reduced density matrix, weighted by the precomputed values of the electronic integrals. The second line is obtained through any of the aforementioned transformations(e.g. Jordan-Wigner or Bravyi-Kitaev) from fermions to qubits and demonstrates that the energy may be efficiently evaluated through a weighted sum of Pauli measurements on the system. Here the Greek indices denote the type of Pauli matrix (I, X, Y, Z) and the Roman index runs over the number of qubits. Recall that the number of terms in sum on the second line scales the same as the number of terms in the first with respect to the number of spin orbitals, and only a partial tomography of the quantum system is ever required.

Based on the value of the energy, a new set of parameters for the state preparation can be determined through some non-linear minimization scheme, such as Nelder-Mead simplex method or simulated annealing [73, 180]. Classical computers are well optimized to perform tasks like adding together the measurements and deciding on new parameters in a non-linear optimization, suggesting a hybrid approach.

The approach can be qualitatively outlined as

1. Prepare a quantum state on a quantum device based on an established protocol that depends on a set of parameters $\{\theta_i\}_i$.

2. Evaluate the energy, $\langle H \rangle$, by partial quantum tomography of the state and

weighted summation on a classical computer.

3. Use a classical computer to decide a new set of parameters $\{\theta_i\}_i$ that lowers the energy, and repeat until convergence.

This procedure takes advantage of a quantum computer's ability to prepare and efficiently sample select elements from complex quantum states, while offloading mundane tasks such as addition and multiplication of scalars to a classical computer. In doing so, we conserve precious quantum resources for what they do best. We call this technique the variational quantum eigensolver and this is the subject of Chapter 2, which includes an experimental implementation on a photonic quantum chip. Since its initial formulation, we have done some work in optimizations for ion traps in Chapter 3. Finally, we imported technology from classical quantum chemistry in Chapter 4, and showed how the use of the appropriate basis could offer dramatic savings, not only in this method, but all those currently being considered for quantum chemistry on quantum computers.

1.3.2 Chapters 5 and 6

The potential of quantum computers to help accelerate our simulations and understanding of chemistry are great, however this eventual promise is not the only reason to study quantum computation and quantum information in chemistry. The knowledge gained in this field has now helped to bolster both our understanding and methodology in both general physics and chemistry. In these Chapters we introduce another such knowledge transfer from quantum computation to classical simulation of quantum systems stemming from a tool known as Feynman's Clock [70, 123, 167].

The gate model of quantum computation, built from qubits and sequences of unitary gates is only one model of quantum computation currently under consideration. Another model of interest is that of adiabatic quantum computation [68]. In this model, one encodes the solution to a problem in the ground state of some Hamiltonian H_p, and prepares the ground state to this problem Hamiltonian through a slow adiabatic evolution from a simple starting Hamiltonian H_D, whose ground state is easy to prepare. By the adiabatic theorem, if the evolution of the Hamiltonian from H_D to H_p is slow enough, one remains in the ground state and the solution to the problem is found. This is often written as a simple linear schedule between the two Hamiltonians parameterized by a real number $s \in [0, 1]$ as

$$H(s) = (1 - s)H_D + sH_p. \tag{1.53}$$

In order to unify this approach with the gate model, Kitaev, building off the work of Feynman, developed the clock Hamiltonian, which encodes the result of a gate model computation into the problem Hamiltonian H_p.

To perform this mapping, one attaches an ancilla quantum register that keeps track of time, called the clock register. If one only has access to qubits then some effort must be made to keep the clock states valid, however for classical computing purposes, it suffices to use a $d-$dimensional qudit, with orthonormal basis states $\langle i|j\rangle = \delta_{ij}$. Given an initial state $|\Psi_0\rangle$ and an ordered sequence of quantum gates $\{U_i\}_i$, the Clock Hamiltonian encoding the full evolution of the quantum state under

these operations as its ground state is

$$\mathcal{H}_c = \sum_t \frac{1}{2}\left(I \otimes |t\rangle\langle t| - U_t \otimes |t+1\rangle\langle t| - U_t^\dagger \otimes |t\rangle\langle t+1| + I \otimes |t+1\rangle\langle t+1|\right) + (1 - |\Psi_0\rangle\langle\Psi_0|) \otimes |0\rangle\langle 0| \,. \tag{1.54}$$

where the final projector is penalty term that breaks the degeneracy of the ground state to the unique evolution desired. The properties of this Hamiltonian have been studied extensively, and it is known that it is frustration-free and has a spectral gap that decreases as $1/T^2$ where T is the total number of discrete time steps under consideration [33, 34, 49]. The ground state of this Hamiltonian is the history state, $|\Phi_H\rangle$, which contains the entire quantum trajectory as

$$|\Phi_H\rangle = \frac{1}{\sqrt{T}} \sum_t |\Psi_t\rangle |t\rangle \tag{1.55}$$

where $|\Psi_t\rangle$ is the state of the quantum system after the application of the first $t-1$ gates. The eigenvalue of this state is 0 by construction, though it may be adjusted through a constant shift factor if desired.

In Chapter 5, we show that this can be derived from a more general discrete time variational principle, and that recasting a dynamics problem as an eigenvalue problem in this manner can have practical computational benefits on a classical computer. In particular, we show that it makes parallel-in-time dynamics possible through use of a proper preconditioner. Moreover the proposed algorithm demonstrates an advantage over the current standard Parareal algorithm [14, 143] for the quantum dynamics problems studied.

In Chapter 6 we take this idea further and show how a method for describing corre-

lated many-body systems, namely the Full Configuration Interaction Quantum Monte Carlo (FCIQMC) [26, 27, 29, 128, 213, 220], technique can be applied to time dynamics problems with this connection. It suggests a close link between the fermion sign problem experienced in the simulation of electronic ground states and the dynamical sign problem arising in the simulation of quantum dynamics. We introduced a rotating basis methodology capable of mitigating the sign problem through approximate pre-computation and demonstrated the method's capabilities in parallel computation of quantum circuits.

1.3.3 Chapter 7

The simulation of explicit wavefunctions has pushed forward by exploiting expert knowledge of physics to identify structure and reduce the complexity of the problem to be solved. For example, the use of symmetry or knowledge that an interaction is relatively weak with regards to another in perturbation theory facilitates fast, accurate approximations. Recently a technique has emerged in the field of signal processing that attempts to exploit a different kind of structure, namely sparsity. These techniques are generally referred to as compressed sensing methodologies [193, 232].

A Hamiltonian governing many non-interacting particles may be written as the sum of the non-interacting pieces

$$H = \sum_i H_i \tag{1.56}$$

where each H_i acts only on the i'th particle. Such a Hamiltonian has a ground state

given by a rank 1 quantum state

$$|\Phi\rangle = |\phi_1\rangle |\phi_2\rangle ... |\phi_N\rangle \tag{1.57}$$

where $H_i |\phi_i\rangle = E_i |\phi_i\rangle$ and as a result $H |\Phi\rangle = (\sum_i E_i) |\Phi\rangle$. If one introduces some weak coupling between the subsystems, we conjecture that one does not expect the rank of the state to change dramatically and as such the state may be well described by a CP-decomposition with rank $r << D = M^N$. If this conjecture is true, the practical question is how to obtain this type of sparse representation without first knowing the full state. Moreover, once the technique is developed, one would like to examine what kind of systems exhibit such a decomposition.

In Chapter 7 we discuss the application of a method from the field of compressed sensing, namely orthogonal matching pursuit combined with imaginary time evolution to find exactly these types of compressed wavefunctions. The methodology we build is general in the sense that it does not require that one use a specific type of ansatz or apply it to a particular type of quantum system. As an example application we consider the electronic structure of molecules, using an antisymmetric low rank decomposition. We find solutions that are dramatically more compact than traditional orthogonal CI type expansions.

If you find that you're spending almost all your time on theory, start turning some attention to practical things; it will improve your theories. If you find that you're spending almost all your time on practice, start turning some attention to theoretical things; it will improve your practice.

Donald Knuth

2

A variational eigenvalue solver on a photonic quantum processor*

Abstract

Quantum computers promise to efficiently solve important problems that are intractable on a conventional computer. For quantum systems, where the physical dimension grows exponentially, finding the eigenvalues of certain operators is one such intractable

*Alberto Peruzzo†, **Jarrod R McClean**†, Peter Shadbolt, Man-Hong Yung, Xiao-Qi Zhou, Peter J Love, Alán Aspuru-Guzik, and Jeremy L O'Brien. A new variational eigenvalue solver on a photonic quantum processor. *Nature Communications*, 5(4213):1-7, 2014.

† These authors contributed equally to this work.

problem and remains a fundamental challenge. The quantum phase estimation algorithm efficiently finds the eigenvalue of a given eigenvector but requires fully coherent evolution. We present an alternative approach that greatly reduces the requirements for coherent evolution and we combine this method with a new approach to state preparation based on ansätze and classical optimization. We implement the algorithm by combining a highly reconfigurable photonic quantum processor with a conventional computer. We experimentally demonstrate the feasibility of this approach with an example from quantum chemistry—calculating the ground state molecular energy for He–H^+. The proposed approach drastically reduces the coherence time requirements, enhancing the potential of quantum resources available today and in the near future.

2.1 Introduction

In chemistry, the properties of atoms and molecules can be determined by solving the Schrödinger equation. However, because the dimension of the problem grows exponentially with the size of the physical system under consideration, exact treatment of these problems remains classically infeasible for compounds with more than 2–3 atoms [229]. Many approximate methods [102] have been developed to treat these systems, but efficient exact methods for large chemical problems remain out of reach for classical computers. Beyond chemistry, the solution of large eigenvalue problems [206] would have applications ranging from determining the results of internet search engines [191] to designing new materials and drugs [86].

Recent developments in the field of quantum computation offer a way forward for efficient solutions of many instances of large eigenvalue problems which are classically intractable [22, 78, 89, 98, 122, 182, 183]. Quantum approaches to finding eigen-

values have previously relied on the quantum phase estimation (QPE) algorithm. The QPE algorithm offers an exponential speedup over classical methods and requires a number of quantum operations $O(p^{-1})$ to obtain an estimate with precision p [1, 2, 6, 8, 136, 251]. In the standard formulation of QPE, one assumes the eigenvector $|\psi\rangle$ of a Hermitian operator $\mathcal{H}$ is given as input and the problem is to determine the corresponding eigenvalue λ. The time the quantum computer must remain coherent is determined by the necessity of $O(p^{-1})$ successive applications of $e^{-i\mathcal{H}t}$, each of which can require on the order of millions or billions of quantum gates for practical applications [111, 251], as compared to the tens to hundreds of gates achievable in the short term.

Here we introduce an alternative to QPE that significantly reduces the requirements for coherent evolution. We have developed a reconfigurable quantum processing unit (QPU), which efficiently calculates the expectation value of a Hamiltonian ($\mathcal{H}$), providing an exponential speedup over exact diagonalization, the only known exact solution to the problem on a traditional computer. The QPU has been experimentally implemented using integrated photonics technology with a spontaneous parametric downconversion single photon source and combined with an optimization algorithm run on a classical processing unit (CPU), which variationally computes the eigenvalues and eigenvectors of $\mathcal{H}$. By using a variational algorithm, this approach reduces the requirement for coherent evolution of the quantum state, making more efficient use of quantum resources, and may offer an alternative route to practical quantum-enhanced computation.

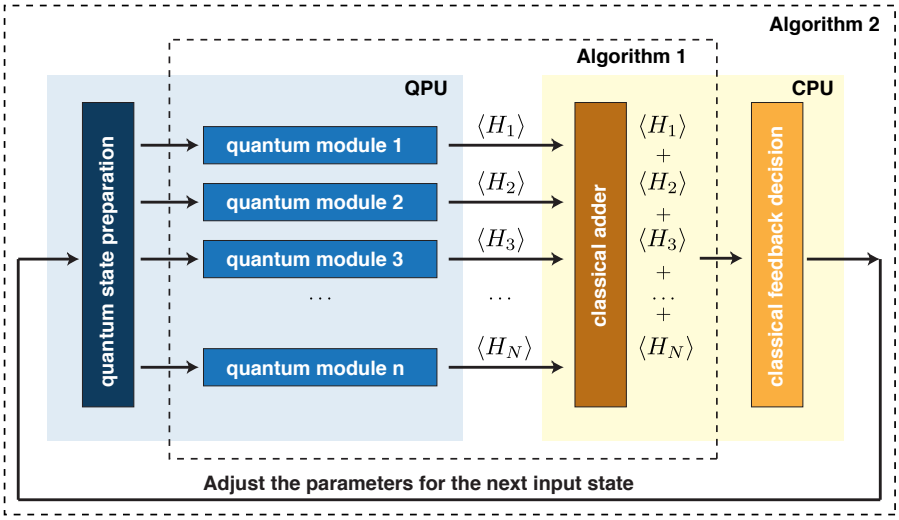

Figure 2.1: Architecture of the quantum-variational eigensolver. (Algorithm 1): Quantum states that have been previously prepared, are fed into the quantum modules which compute $\langle \mathcal{H}_i \rangle$, where $\mathcal{H}_i$ is any given term in the sum defining $\mathcal{H}$. The results are passed to the CPU which computes $\langle \mathcal{H} \rangle$. (Algorithm 2): The classical minimization algorithm, run on the CPU, takes $\langle \mathcal{H} \rangle$ and determines the new state parameters, which are then fed back to the QPU.

2.2 Results

2.2.1 Algorithm 1: Quantum expectation estimation

This algorithm computes the expectation value of a given Hamiltonian $\mathcal{H}$ for an input state $|\psi\rangle$. Any Hamiltonian may be written as

$$\mathcal{H} = \sum_{i\alpha} h^i_\alpha \sigma^i_\alpha + \sum_{ij\alpha\beta} h^{ij}_{\alpha\beta} \sigma^i_\alpha \sigma^j_\beta + \ldots \tag{2.1}$$

for real h where Roman indices identify the subsystem on which the operator acts, and Greek indices identify the Pauli operator, e.g. $\alpha = x$. Note that no assumption about the dimension or structure of the hermitian Hamiltonian is needed for this expansion to be valid. By exploiting the linearity of quantum observables, it follows

that

$$\langle \mathcal{H} \rangle = \sum_{i\alpha} h^i_\alpha \langle \sigma^i_\alpha \rangle + \sum_{ij\alpha\beta} h^{ij}_{\alpha\beta} \langle \sigma^i_\alpha \sigma^j_\beta \rangle + ... \tag{2.2}$$

We consider Hamiltonians that can be written as a number of terms which is polynomial in the size of the system. This class of Hamiltonians encompasses a wide range of physical systems, including the electronic structure Hamiltonian of quantum chemistry, the quantum Ising Model, the Heisenberg Model [147, 153], matrices that are well approximated as a sum of n-fold tensor products [188, 189], and more generally any $k-$sparse Hamiltonian without evident tensor product structure (see Supplementary Information for details). Thus the evaluation of $\langle \mathcal{H} \rangle$ reduces to the sum of a polynomial number of expectation values of simple Pauli operators for a quantum state $|\psi\rangle$, multiplied by some real constants. A quantum device can efficiently evaluate the expectation value of a tensor product of an arbitrary number of simple Pauli operators [188]. Therefore, with an n-qubit state we can efficiently evaluate the expectation value of this $2^n \times 2^n$ Hamiltonian.

One might attempt this using a classical computer by separately optimizing all reduced states corresponding to the desired terms in the Hamiltonian, but this would suffer from the N-representability problem, which is known to be intractable for both classical and quantum computers (it is in the quantum complexity class QMA-Hard [145]). The power of our approach derives from the fact that quantum hardware can store a global quantum state with exponentially fewer resources than required by classical hardware, and as a result the N-representability problem does not arise.

As the expectation value of a tensor product of an arbitrary number of Pauli operators can be measured in constant time and the spectrum of each of these operators is bounded, to obtain an estimate with precision p of an individual element with coeffi-

cient h, which is an arbitrary element from the set of constants $\{h^{ij\ldots}_{\alpha\beta\ldots}\}$, our approach incurs a cost of $O(|h|^2\ p^{-2})$ repetitions. Thus the total cost of computing the expectation value of a state $|\psi\rangle$ is bounded by $O(|h_{max}|^2\ M\ p^{-2})$, where M is the number of terms in the decomposition of the Hamiltonian and h_{max} is the coefficient with maximum norm in the decomposition of the Hamiltonian. The advantage of this approach is that the coherence time to make a single measurement after preparing the state is $O(1)$. Conversely, the disadvantage of this approach with respect to QPE is the scaling in the total number of operations as a function of the desired precision is quadratically worse ($O(p^{-2})$ vs $O(p^{-1})$). Moreover, this scaling will also reflect the number of state preparation repetitions required, whereas in QPE, the number of state preparation steps is constant. In essence, we dramatically reduce the coherence time requirement while maintaining an exponential advantage over the classical case, by adding a polynomial number of repetitions with respect to QPE.

2.2.2 Algorithm 2: Quantum variational eigensolver

The procedure outlined above replaces the long coherent evolution required by QPE by many short coherent evolutions. In both QPE and Algorithm 1 we require a good approximation to the ground state wavefunction to compute the ground state eigenvalue and we now consider this problem. Previous approaches have proposed to prepare ground states by adiabatic evolution [6], or by the quantum Metropolis algorithm [226, 262]. Unfortunately both of these require long coherent evolution. Algorithm 2 is a variational method to prepare the eigenstate and, by exploiting Algorithm 1, requires short coherent evolution. Algorithms 1 and 2 and their relationship are shown in Figure 2.1 and detailed in the Supplementary Information.

It is well known that the eigenvalue problem for an observable represented by an

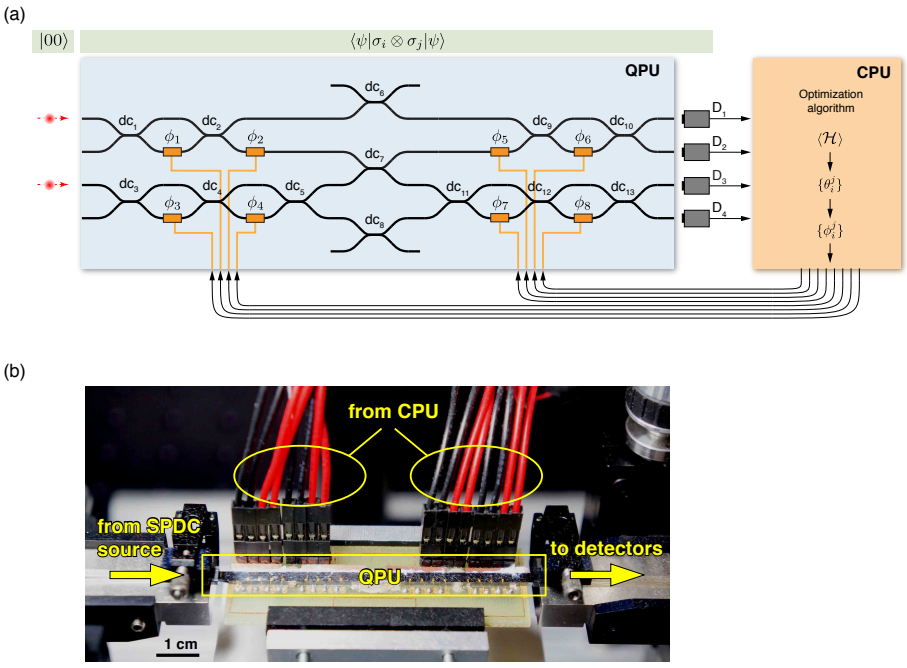

Figure 2.2: Experimental implementation of our scheme.(a) Quantum state preparation and measurement of the expectation values $\langle\psi|\sigma_i \otimes \sigma_j|\psi\rangle$ are performed using a quantum photonic chip. Photon pairs, generated using spontaneous parametric down-conversion are injected into the waveguides encoding the $|00\rangle$ state. The state $|\psi\rangle$ is prepared using thermal phase shifters ϕ_{1-8} (orange rectangles) and one CNOT gate and measured using photon detectors. Coincidence count rates from the detectors D_{1-4} are passed to the CPU running the optimization algorithm. This computes the set of parameters for the next state and writes them to the quantum device. (b) A photograph of the QPU.

operator $\mathcal{H}$ can be restated as a variational problem on the Rayleigh-Ritz quotient [204, 205], such that the eigenvector $|\psi\rangle$ corresponding to the lowest eigenvalue is the $|\psi\rangle$ that minimizes

$$\frac{\langle\psi|\,\mathcal{H}\,|\psi\rangle}{\langle\psi|\psi\rangle}. \tag{2.3}$$

By varying the experimental parameters in the preparation of $|\psi\rangle$ and computing the Rayleigh-Ritz quotient using Algorithm 1 as a subroutine in a classical minimization, one may prepare unknown eigenvectors. At the termination of the algorithm, a sim-

ple prescription for the reconstruction of the eigenvector is stored in the final set of experimental parameters that define $|\psi\rangle$.

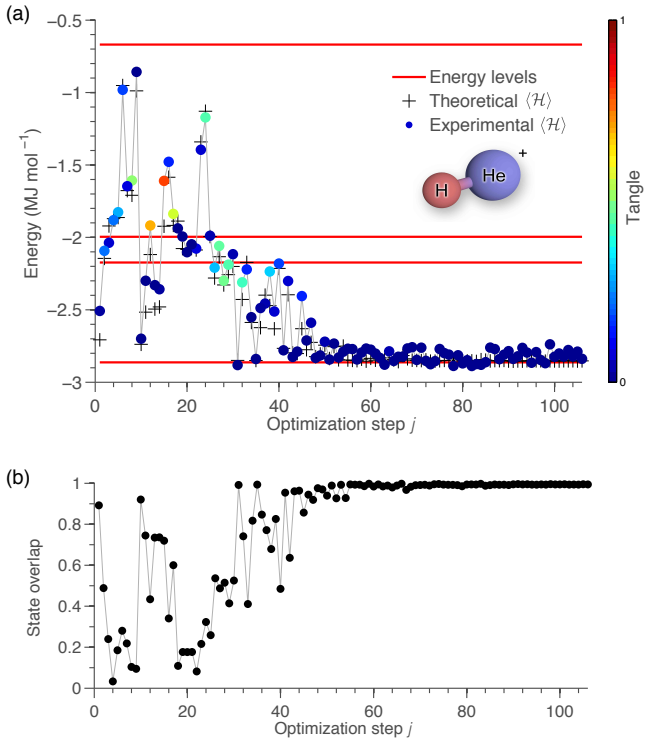

Figure 2.3: Finding the ground state of He-H$^+$ for a specific molecular separation $R = 90$ pm. (a) Experimentally computed energy $\langle \mathcal{H} \rangle$ (colored dots) as a function of the optimization step j. The color represents the tangle (degree of entanglement) of the physical state, estimated directly from the state parameters $\{\phi_i^j\}$. The red lines indicate the energy levels of $\mathcal{H}(R)$. The optimization algorithm clearly converges to the ground state of the molecule, which has small but non zero tangle. The crosses show the energy calculated at each experimental step, assuming an ideal quantum device. (b) Overlap $|\langle \psi^j | \psi^G \rangle|$ between the experimentally computed state $|\psi^j\rangle$ at each the optimization step j and the theoretical ground state of $\mathcal{H}$, $|\psi^G\rangle$. Error bars are smaller that the data points. Further details are provided in the Methods section and in the Supplementary Information.

If a quantum state is characterized by an exponentially large number of parameters, it cannot be prepared with a polynomial number of operations. The set of efficiently preparable states are therefore characterized by polynomially many parameters, and we choose a particular set of ansatz states of this type. Under these condi-

tions, a classical search algorithm on the experimental parameters which define $|\psi\rangle$, needs only explore a polynomial number of dimensions—a requirement for the search to be efficient.

One example of a quantum state parametrized by a polynomial number of parameters for which there is no known efficient classical implementation is the unitary coupled cluster ansatz [225]

$$|\Psi\rangle = e^{T-T^\dagger} |\Phi\rangle_{ref} \tag{2.4}$$

where $|\Phi\rangle_{ref}$ is some reference state, usually the Hartree Fock ground state, and T is the cluster operator for an N electron system defined by

$$T = T_1 + T_2 + T_3 + ... + T_N \tag{2.5}$$

with

$$T_1 = \sum_{pr} t_p^r \hat{a}_p^\dagger \hat{a}_r \tag{2.6}$$

$$T_2 = \sum_{pqrs} t_{pq}^{rs} \hat{a}_p^\dagger \hat{a}_q^\dagger \hat{a}_r \hat{a}_s \tag{2.7}$$

and higher order terms follow logically. It is clear that by construction the operator $(T - T^\dagger)$ is anti-hermitian, and exponentiation maps it to a unitary operator $U = e^{(T-T^\dagger)}$. For any fixed excitation level k, the reduced cluster operator is written as

$$T^{(k)} = \sum_{i=1}^{k} T_i \tag{2.8}$$

In general no efficient implementation of this ansatz has yet been developed for a classical computer, even for low order cluster operators due to the non-truncation of the

BCH series [225]. However this state may be prepared efficiently on a quantum device. The reduced anti-hermitian cluster operator $(T^{(k)} - T^{(k)\dagger})$ is the sum of a polynomial number of terms, namely it contains a number of terms $O(N^k(M-N)^k)$ where M is the number of single particle orbitals. By defining an effective Hermitian Hamiltonian $\mathcal{H} = i(T^{(k)} - T^{(k)\dagger})$ and performing the Jordan-Wigner transformation to reach a Hamiltonian that acts on the space of qubits, $\tilde{\mathcal{H}}$, we are left with a Hamiltonian which is a sum of polynomially many products of Pauli operators. The problem then reduces to the quantum simulation of this effective Hamiltonian, $\tilde{\mathcal{H}}$, which can be done in polynomial time using the procedure outlined by Ortiz et al. [188]. We note that while this state preparation procedure utilizes tools from quantum simulation, the total effective time of evolution is fixed by the expansion coefficients t^{rs}_{pq}. This is in contrast to normal difficulties encountered in quantum phase estimation, where simulations must be carried out for times that are exponential in the desired final precision.

While there is currently no known efficient classical algorithm based on these ansatz states, non-unitary coupled cluster ansatz is sometimes referred to as the "gold standard of quantum chemistry" as it is the standard of accuracy to which other methods in quantum chemistry are often compared. The unitary version of this ansatz is thought to yield superior results to even this "gold standard" [225].

2.2.3 Prototype demonstration

We have implemented the QPU using integrated quantum photonics technology [186]. Our device, shown schematically in Figure 2.2 is a reconfigurable waveguide chip that can prepare and measure arbitrary two-bit pure states using several single qubit rotations and one two-qubit entangling gate. This device operates across the full space

of possible configurations with mean statistical fidelity $F > 99\%$ [210]. The state is path-encoded using photon pairs generated via a spontaneous parametric downconversion process. State preparation and measurement in the Pauli basis is achieved by setting 8 voltage driven phase shifters and counting photon detection events with silicon single photon detectors.

The ability to prepare an arbitrary two-qubit separable or entangled state enables us to investigate 4×4 Hamiltonians. For the experimental demonstration of our algorithm we choose a problem from quantum chemistry, namely determining the bond dissociation curve of the molecule He-H^+ in a minimal basis. The full configuration

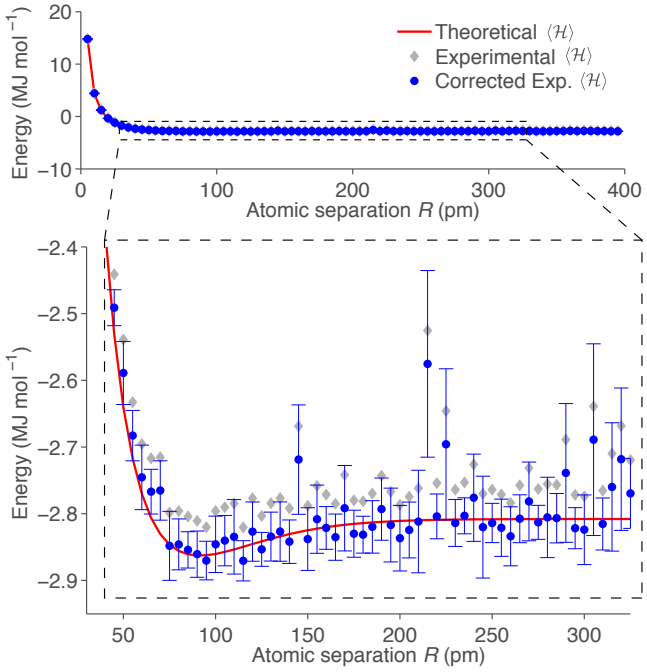

Figure 2.4: Bond dissociation curve of the He-H^+ molecule. This curve is obtained by repeated computation of the ground state energy (as shown in Figure 2.3) for several $\mathcal{H}(R)$. The magnified plot shows that after correction for the measured systematic error, the data overlap with the theoretical energy curve and importantly we can resolve the molecular separation of minimal energy. Error bars show the standard deviation of the computed energy, as described in the Methods section.

interaction Hamiltonian for this system has dimension 4, and can be written compactly as

$$\mathcal{H}(R) = \sum_{i\alpha} h^i_\alpha(R)\sigma^i_\alpha + \sum_{ij\alpha\beta} h^{ij}_{\alpha\beta}(R)\sigma^i_\alpha\sigma^j_\beta \tag{2.9}$$

The coefficients $h^i_\alpha(R)$ and $h^{ij}_{\alpha\beta}(R)$ were determined using the PSI3 computational package [56] and are tabulated in the Supplementary Table II.

In order to compute the bond dissociation of the molecule, we use Algorithm 2 to compute its ground state for a range of values of the nuclear separation R. In Figure 2.3 we report a representative optimization run for a particular nuclear separation, demonstrating the convergence of our algorithm to the ground state of $\mathcal{H}(R)$ in the presence of experimental noise. Figure 2.3a demonstrates the convergence of the average energy, while Figure 2.3b demonstrates the convergence of the overlap $|\langle\psi^j|\psi^G\rangle|$ of the current state $|\psi^j\rangle$ with the target state $|\psi^G\rangle$. The color of each entry in Figure 2.3a represents the tangle (absolute concurrence squared) of the state at that step of the algorithm. It is known that the volume of separable states is doubly-exponentially small with respect to the rest of state space [223]. Thus, the ability to traverse non-separable state space increases the number of paths by which the algorithm can converge and will be a requirement for future large-scale implementations. Moreover, it is clear that the ability to produce entangled states is a necessity for the accurate description of general quantum systems where eigenstates may be non-separable, for example the ground state of the He-H^+ Hamiltonian has small but not negligible tangle.

Repeating this procedure for several values of R, we obtain the bond dissociation curve which is reported in Figure 2.4. After the computed energies have been corrected for experimental errors, the determination of the equilibrium bond length of

the molecule was found to be $R = 92.3 \pm 0.1$ pm with a corresponding ground state electronic energy of $E =$ -2.865$\pm$0.008 MJ mol^{-1}. Full details of the correction for systematic errors and estimation of the uncertainty on E are reported in the Supplementary Information. The corresponding theoretical curve shows the numerically exact energy derived from a full configuration interaction calculation of the molecular system in the same basis. More than 96% of the experimental data are within chemical accuracy with respect to the theoretical values. At the conclusion of the optimization, we retain full knowledge of the experimental parameters, which can be used for efficient reconstruction of the state $|\psi\rangle$ in the event that additional physical or chemical properties are required.

2.3 Discussion

Algorithm 1 uses relatively few quantum resources compared to QPE. Broadly speaking, QPE requires a large number of n-qubit quantum controlled operations to be performed in series—placing considerable demands on the number of components and coherence time—while the inherent parallelism of our scheme enables a small number of n-qubit gates to be exploited many times, drastically reducing these demands. Moreover, adding control to arbitrary unitary operations in practice is difficult if not impossible for current quantum architectures (although a proposed scheme to add control to arbitrary unitary operations has recently been demonstrated [272]). To give a numerical example, the QPE circuit for a 4 x 4 Hamiltonian such as that demonstrated here would require at least 12 CNOT gates, while our method only requires one. We note that the resource saving provided by Algorithm 1 incurs a cost of polynomial repetitions of the state preparation, as compared to the single copy required

by QPE. In many cases (for example our photonic implementation), repeated preparation of a state is not significantly harder than preparation of a single copy, requiring only a polynomial overhead in time without any modification of the device.

In implementing Algorithm 2, the device prepares ansatz states that are defined by a polynomial set of parameters. This ansatz might be chosen based on knowledge of the physical system of interest (as for the unitary coupled cluster and typical quantum chemistry ansätze) thus determining the device design. However, our architecture allows for an alternative, and potentially more promising approach, where the device is first constructed based on the available resources and we define the set of states that the device can prepare as the "device ansätze". Due to the quantum nature of the device, this ansatz can be very distinct from those used in traditional quantum chemistry. With this alternative approach the physical implementation is then given by a known sequence of quantum operations with adjustable parameters—determined at the construction of the device—with a maximum depth fixed by the coherence time of the physical qubits. This approach, while approximate, provides a variationally optimal solution for the given quantum resources and may still be able to provide qualitatively correct solutions, just as approximate methods do in traditional quantum chemistry (for example Hartree Fock). The unitary coupled cluster ansatz (Eq. 2.4) provides a concrete example where our approach provides an exponential advantage over known classical techniques. For this ansätze, with as few as $40 - 50$ qubits, one expects to manipulate a state which is not efficient to simulate classically, and can provide a solution superior to the classical gold standard, non-unitary coupled cluster.

We have developed and experimentally implemented a new approach to solving the eigenvalue problem with quantum hardware. Algorithm 1 shares with QPE the need to prepare a good approximation to the ground state, but replaces a single long

coherent evolution by a number of shorter coherent calculations proportional to the number of terms in the Hamiltonian. While the effect of errors on each of these calculations is the same as in QPE, the reliance on a number of separate calculations makes the algorithm sensitive to variations in state preparation between the separate quantum calculations. This effect requires further investigation. While the local Hamiltonian problem is known to be QMA-complete [119] in its entirety, under the assumption that a good approximation to the state can be prepared, both QPE and our method can efficiently estimate the energy of the state, and it is in this setting that we compare them. In Algorithm 2, we experimentally implemented a ground state preparation procedure through a direct variational algorithm on the control parameters of the quantum hardware. The prepared state could be utilized in either Algorithm 1 or QPE if desired. Larger calculations will require a choice of ansatz, for which there are two possibilities. One could experimentally implement chemically motivated ansätze such as the unitary coupled cluster method. Alternatively one could pursue those ansätze that are most easy to implement experimentally—creating a new set of device ansätze states which would require classification in terms of their overlap with chemical ground states. Such a classification would be a good way to determine the value of a given experimental advance—for ground state problems it is best to focus limited experimental resources on those efforts that will most enhance the overlap of preparable states with chemical ground states. In addition to the above issues, which we leave to future work, an interesting avenue of research is to ask whether the conceptual approach described here could be used to address other intractable problems with quantum-enhanced computation. Examples that can be mapped to the ground state problem, and where the N-representability problem does not occur, include search engine optimisation and image recognition. It should be noted that the

approach presented here requires no control or auxiliary qubits, relying only on measurement techniques that are already well established. For example, in the two qubit case, these measurements are identical to those performed in Bell inequality experiments.

Quantum simulators with only a few tens of qubits are expected to outperform the capabilities of conventional computers, not including open questions regarding fault tolerance and errors/precision. Our scheme would allow such devices to be implemented using dramatically less resources than the current best known approach.

2.4 Methods

2.4.1 Classical optimization algorithm

For the classical optimization step of our integrated processor we implemented the Nelder-Mead (NM) algorithm [180], a simplex-based direct search (DS) method for unconstrained minimization of objective functions. Although in general NM can fail because of the deterioration of the simplex geometry or lack of sufficient decrease, the convergence of this method can be greatly improved by adopting a restarting strategy. Although other DS methods, such as the gradient descent, can perform better for smooth functions, these are not robust to the noise which makes the objective function non-smooth under experimental conditions. NM has the ability to explore neighboring valleys with better local optima and likewise this exploring feature usually allows NM to overcome non-smoothnesses. We verified that the gradient descent minimization algorithm is not able to converge to the ground state of our Hamiltonian under the experimental conditions, mainly due to the Poissonian noise associated

with our photon source and the accidental counts of the detection system, while NM converged to the global minimum in most optimization runs.

2.4.2 Mapping from the state parameters to the chip phases

The set of phases $\{\theta_i\}$, which uniquely identifies the state $|\psi\rangle$, is not equivalent to the phases which are written to the photonic circuit $\{\phi_i\}$, since the chip phases are also used to implement the desired measurement operators $\sigma_\alpha \otimes \sigma_\beta$. Therefore, knowing the desired state parameters and measurement operator we compute the appropriate values of the chip phases on the CPU at each iteration of the optimization algorithm. The algorithm for finding the phases an arbitrary two-qubit state is described in the Supplementary Information.

2.4.3 Experimental Details

Estimation of the error on $\langle \mathcal{H} \rangle$

We performed measurements of the statistical and systematic errors that affect our computation of $\langle \mathcal{H} \rangle$.

Statistical errors

Statistical errors due to the Poissonian noise associated with single photon statistics are intrinsic to the estimation of expectation values in quantum mechanics.

These errors can be arbitrarily reduced at a sublinear cost of measurement time (i.e. efficiently) since the magnitude of error is proportional to the square root of the count rate. We experimentally measured the standard deviation of an expectation value $\langle \mathcal{H}_i \rangle$ for a particular state using 50 trials. The total average coincidence rate

was $\sim$1500 s^{-1}. The standard deviation was found to be 37 KJ mol^{-1}, which is comparable with the error observed in the measurement of the ground state energy shown in Figure 2.4.

The minima of the potential energy curve was determined by a generalized least squares procedure to fit a quadratic curve to the experimental data points in the region $R = (80, 100)$ pm, as is common in the use of trust region searches for minima [55], using the inverse experimentally measured variances as weights. Covariances determined by the generalized least squares procedure were used as input to a Monte Carlo sampling procedure to determine the minimum energy and equilibrium bond distance as well as their uncertainties assuming Gaussian random error. The uncertainties reported represent standard deviations. Sampling error in the Monte Carlo procedure was 3×10^{-4} pm for the equilibrium bond distance and 3×10^{-8} MJ mol^{-1} for the energy.

In Figure 4, the large deviations from the theoretical line result from the coincidental impact of noise resulting in premature optimization termination. These points could have been rerun or eliminated using the prior knowledge of smoothness of the dissociation curve. However to accurately portray the performance of the algorithm exactly as described, with no expert interference, these points are retained.

Systematic errors

In all the measurements described above we observed a constant and reproducible small shift, $\epsilon = 50$ KJ mol^{-1}, of the expectation value with respect to the theoretical value of the energy. There are at least three effects which contribute to this systematic error.

Firstly, the down-conversion source that we use in our experiment does not pro-

duce the pure two photon state that is required for high-fidelity quantum interference. In particular, higher order photon number terms and, more significantly, photon distinguishability both degrade the performance of our entangling gate and thus the preparation of the state $|\psi\rangle$. This results in a shift of the measured energy $\langle\psi|\,\mathcal{H}\,|\psi\rangle$. Higher order terms could be effectively eliminated by use of true single photon sources (such as quantum dots or nitrogen vacancy centers in diamond), and there is no fundamental limit to the degree of indistinguishability which can be achieved through improved state engineering.

Secondly, imperfections in the implementation of the photonic circuit also reduce the fidelity with which $|\psi\rangle$ is prepared and measured. Small deviations from designed beamsplitter reflectivities and interferometer path lengths, as well as imperfections in the calibration of voltage-controlled phases shifters used to manipulate the state, all contribute to this effect. However, these are technological limitations that can be greatly improved in future realizations.

Finally, unbalanced input and output coupling efficiency also results in skewed two-photon statistics, again shifting the measured expectation value of $\langle\mathcal{H}\rangle$.

Another systematic effect that can be noted in Figure 4 is that the magnitude of the error on the experimental estimation of the ground state energy increases with R. This is due to the fact that as R increases, the first and second excited eigenstates of this Hamiltonian become degenerate, resulting in increased difficulty for the classical minimization, generating mixtures of states that increases the overall variance of the estimation.

Count rate

In our experiment the mean count rate, which directly determines the statistical error, was approximately 2000-4000 twofold events per measurement. The expectation value of a given Hamiltonian was reconstructed at each point from four two-qubit Pauli measurements. For the bond dissociation curve we measured about 100 points per optimization run. In the full dissociation curve we found the ground states of 79 Hamiltonians. The full experiment was performed in about 158 hours.

State preparation is relatively fast, requiring a few milliseconds to set the phases on the chip. However 17 seconds are required for cooling the chip, resulting in a duty-cycle of $\sim 5\%$. The purpose of this is to overcome instability of the fibre-to-chip coupling due to thermal expansion of the chip during operation. This will not be an issue in future implementations where fibres will be permanently fixed to the chip's facets. Moreover the thermal phase shifters used here will also likely be replaced by alternative technologies based on the electro-optic effect.

Brighter single photon sources will considerably reduce the measurement time.

2.5 Supplemental Information

2.5.1 Quantum eigenvector preparation algorithm

Below we detail the steps involved in implementing Algorithm 2.

1. Design a quantum circuit, controlled by a set of experimental parameters $\{\theta_i\}$, which can prepare a class of states. Using this device, prepare the initial state $|\psi^0\rangle$ and define the objective function $f(\{\theta_i^n\}) = \langle\psi(\{\theta_i^n\})|\,\mathcal{H}\,|\psi(\{\theta_i^n\})\rangle$, which efficiently maps the set of experimental parameters to the expectation value of

the Hamiltonian and is computed in this work by Algorithm 1. n denotes the current iteration of the algorithm.

2. Let $n = 0$

3. Repeat until optimization is completed

   (a) Call Algorithm 1 with $\{\theta_i\}$ as input:

      i. Using the QPU, compute $\langle\sigma_\alpha^i\rangle$, $\langle\sigma_\alpha^i\sigma_\beta^j\rangle$, $\langle\sigma_\alpha^i\sigma_\beta^j\sigma_\gamma^k\rangle$, ..., on $|\psi^n\rangle$ for all terms of $\mathcal{H}$.

      ii. Classically sum on CPU the values from the QPU with their appropriate weights, h, to obtain $f(\{\theta_i^n\})$

   (b) Feed $f(\{\theta_i^n\})$ to the classical minimization algorithm (e.g. gradient descent or Nelder-Mead Simplex method), and allow it to determine $\{\theta_i^{n+1}\}$.

2.5.2 Second Quantized Hamiltonian

When taken with the Born-Oppenheimer approximation, the Hamiltonian of an electronic system can be generally written [102] as

$$\mathcal{H}(R) = \sum_{pq} h_{pq}(R)\hat{a}_p^\dagger\hat{a}_q + \sum_{pqrs} h_{pqrs}(R)\hat{a}_p^\dagger\hat{a}_q^\dagger\hat{a}_r\hat{a}_s \tag{2.10}$$

where $\hat{a}_i^\dagger$ and $\hat{a}_j$ are the fermionic creation and annihilation operators that act on the single particle basis functions chosen to represent the electronic system and obey the canonical anti-commutation relations $\{\hat{a}_i^\dagger, \hat{a}_j\} = \delta_{ij}$ and $\{\hat{a}_i, \hat{a}_j\} = \{\hat{a}_i^\dagger, \hat{a}_j^\dagger\} = 0$. R is a vector representing the positions of the Nuclei in the system, and is fixed for

any given geometry. The constants $h_{pq}(R)$ and $h_{pqrs}(R)$ are evaluated using an initial Hartree-Fock calculation and relate the second quantized Hamiltonian to the first quantized Hamiltonian. They are calculated as

$$h_{pq} = \int dr\ \chi_p(r)^* \left(-\frac{1}{2}\nabla^2 - \sum_\alpha \frac{Z_\alpha}{|r_\alpha - r|} \right) \chi_q(r) \tag{2.11}$$

$$h_{pqrs} = \int dr_1\ dr_2\ \frac{\chi_p(r_1)^* \chi_q(r_2)^* \chi_r(r_1) \chi_s(r_2)}{|r_1 - r_2|} \tag{2.12}$$

where $\chi_p(r)$ are single particle spin orbitals, Z_α is the nuclear charge, and r_α is the nuclear position. From the definition of the Hamiltonian, it is clear that the number of terms in the Hamiltonian is $O(N^4)$ in general, where N is the number of single particle basis functions used. The map from the Fermionic algebra of the second quantized Hamiltonian to the distinguishable spin algebra of qubits is given by the Jordan-Wigner transformation [112], which for our purposes can be concisely written as

$$\hat{a}_j \to I^{\otimes j-1} \otimes \sigma_+ \otimes \sigma_z^{\otimes N-j} \tag{2.13}$$

$$\hat{a}_j^\dagger \to I^{\otimes j-1} \otimes \sigma_- \otimes \sigma_z^{\otimes N-j} \tag{2.14}$$

where σ_+ and σ_- are the Pauli spin raising and lowering operators respectively. It is clear that this transformation does not increase the number of terms present in the Hamiltonian, it merely changes their form and the spaces on which they act. Thus the requirement that the Hamiltonian is a sum of polynomially many products of Pauli operators is satisfied. As a result, the expectation value of any second quantized chemistry Hamiltonian can be efficiently measured with our scheme.

For the specific case of He-H$^+$ in a minimal, STO-3G basis [102], it turns out that

full configuration interaction (FCI) Hamiltonian has dimension four, thus a more compact representation is possible than in the general case. In this case, the FCI Hamiltonian can be written down for each geometry expanded in terms of the tensor products of two Pauli operators. Thus the Hamiltonian is given explicitly by an FCI calculation in the PSI3 computational package [56] and can be written as

$$\mathcal{H}(R) = \sum_{i\alpha} h^i_\alpha(R)\sigma^i_\alpha + \sum_{ij\alpha\beta} h^{ij}_{\alpha\beta}(R)\sigma^i_\alpha\sigma^j_\beta \tag{2.15}$$

The coefficients $h^i_\alpha(R)$ and $h^{ij}_{\alpha\beta}(R)$ are tabulated in Supplementary Table **??**.

2.5.3 Finding excited states

Frequently, one may be interested in eigenvectors and eigenvalues related to excited states (interior eigenvalues). Fortunately our scheme can be used with only minor modification to find these excited states by repeating the procedure on $\mathcal{H}_\lambda = (\mathcal{H} - \lambda)^2$. The folded spectrum method [154, 241] allows a variational method to converge to the eigenvector closest to the shift parameter λ. By scanning through a range of λ values, one can recover the eigenvectors and eigenvalues of interest. Although this operation incurs a small polynomial overhead —the number of terms in the Hamiltonian is quadratic with respect to the original Hamiltonian— this extra cost is marginal compared to the cost of solving the problem classically.

2.5.4 Application to $k-$sparse Hamiltonians

The method described in the main body of this work may be applied to general $k-$sparse Hamiltonian matrices which are row-computable even when no efficient tensor decom-

position is evident with only minor modification. A Hamiltonian $\mathcal{H}$ is referred to as $k-$sparse if there are at most k non-zero elements in each row and column of the matrix and row computable if there is an efficient algorthim for finding the locations and values of the non-zero matrix elements in each row of $\mathcal{H}$.

Let $\mathcal{H}$ be a $2^n \times 2^n$ $k-$sparse row-computable Hamiltonian. A result by Berry et al. [21] shows that $\mathcal{H}$ may be decomposed as $\mathcal{H} = \sum_{j=1}^{m} \mathcal{H}_j$ with $m = 6k^2$, $\mathcal{H}_j$ being a $1-$sparse matrix and each $\mathcal{H}_j$ may be efficiently simulated ($e^{-i\mathcal{H}_j t}$ may be acted on the qubits) by making only $O(\log^* n)$ queries to the Hamiltonian $\mathcal{H}$. Alternatively, a more recent result by Childs et al. [47] shows that it is possible to use a star decomposition of the Hamiltonian such that $m = 6k$ and each $\mathcal{H}_j$ is now a galaxy which can be efficiently simulated using $O(k + \log^* N)$ queries to the Hamiltonian. Either of these schemes may be used to implement our algorithm efficiently for a general $k-$sparse matrix, and the choice may be allowed to depend on the particular setup available. Following a prescription by Knill et al. [124], the ability to simulate $\mathcal{H}_j$ is sufficient for efficient measurement of the expectation value $\langle \mathcal{H}_j \rangle$. After determining these values, one may proceed as before in the algorithm as outlined in the main text and use them to determine new parameters for the classical minimization.

2.5.5 Computational Scaling

In this section, we demonstrate the polynomial scaling of each iteration of our algorithm with respect to system size, and contrast that with the exponential scaling of the current best-known classical algorithm for the same task. Suppose that the algorithm has progressed to an iteration j in which we have prepared a state vector $|\psi^j\rangle$ which is stored in n qubits and parameterized by the set of parameters $\{\theta_i^j\}$.

We wish to find the average value of the Hamiltonian, $\langle \mathcal{H} \rangle$ on this state. We will

assume that there are M terms comprising the Hamiltonian, and assume that M is polynomial in the size of the physical system of interest. Without loss of generality, we select a single term from the Hamiltonian, $\mathcal{H}_i$ that acts on k bits of the state, and denote the average of this term by $\langle \mathcal{H}_i \rangle = h \langle \tilde{\sigma} \rangle$ where h is a constant and $\tilde{\sigma}$ is the $k-$fold tensor product of Pauli operators acting on the system. As the expectation value of a tensor product of an arbitrary number of Pauli operators can be measured in constant time and the spectrum of each of these operators is bounded, if the desired precision on the value is given by p, we expect the cost of this estimation to be $O(|h|^2\ p^{-2})$ repetitions of the preparation and measurement procedure. Thus we estimate the cost of each function evaluation to be $O(|h_{max}|^2 M\ p^{-2})$. For most modern classical minimization algorithms (including the Nelder-Mead simplex method [180]), the cost of a single update step, scales linearly or at worst polynomially in the number of parameters included in the minimization [73]. By assumption, the number of parameters in the set $\{\theta_i^j\}$, is polynomial in the system size. Thus the total cost per iteration is roughly given by $O(n^r |h_{max}|^2\ M\ p^{-2})$ for some small constant r which is determined by the encoding of the quantum state and the classical minimization method used.

Contrasting this to the situation where the entire algorithm is performed classically, a much different scaling results. Storage of the quantum state vector $|\psi^j\rangle$ using currently known exact encodings of quantum states, requires knowing 2^n complex numbers. Moreover, given this quantum state, the computation of the expectation value $\langle \tilde{\sigma} \rangle = \langle \psi^j | \tilde{\sigma} | \psi^j \rangle$ using modern methods requires $O(2^n)$ floating point operations. Thus a single function evaluation is expected to require exponential resources in both storage and computation when performed on a classical computer. Moreover, the number of parameters which a classical minimization algorithm must manipulate

to represent this state exactly is 2^n. Thus performing even a single minimization step to determine $|\psi^{j+1}\rangle$ requires an exponential number of function evaluations, each of which carries an exponential cost. One can roughly estimate the scaling of this procedure as $O(M\ 2^{n(r+1)})$

From this coarse analysis, we conclude that our algorithm attains an exponential advantage in the cost of a single iteration over the best known classical algorithms, provided the assumptions on the Hamiltonian and quantum state are satisfied. While convergence to the final ground state must still respect the known complexity QMA-Complete complexity of this task [119], we believe this still demonstrates the value of our algorithm, especially in light of the limited quantum resource requirements.

It has today occurred to me that an amplifier using semiconductors rather than vacuum is in principle possible.

William Shockley

3

From transistor to trapped-ion computers for quantum chemistry*

Abstract

Over the last few decades, quantum chemistry has progressed through the development of computational methods based on modern digital computers. However, these methods can hardly fulfill the exponentially-growing resource requirements when applied to large quantum systems. As pointed out by Feynman, this restriction is in-

*Man-Hong Yung, Jorge Casanova, Antonio Mezzacapo, **Jarrod R McClean**, Lucas Lamata, Alán Aspuru-Guzik, and Enrique Solano. From transistor to trapped-ion computers for quantum chemistry. *Scientific Reports*, 4(3589):1-7, 2014.

trinsic to all computational models based on classical physics. Recently, the rapid advancement of trapped-ion technologies has opened new possibilities for quantum control and quantum simulations. Here, we present an efficient toolkit that exploits both the internal and motional degrees of freedom of trapped ions for solving problems in quantum chemistry, including molecular electronic structure, molecular dynamics, and vibronic coupling. We focus on applications that go beyond the capacity of classical computers, but may be realizable on state-of-the-art trapped-ion systems. These results allow us to envision a new paradigm of quantum chemistry that shifts from the current transistor to a near-future trapped-ion-based technology.

3.1 Introduction

Quantum chemistry represents one of the most successful applications of quantum mechanics. It provides an excellent platform for understanding matter from atomic to molecular scales, and involves heavy interplay of experimental and theoretical methods. In 1929, shortly after the completion of the basic structure of the quantum theory, Dirac speculated [60] that the fundamental laws for chemistry were completely known, but the application of the fundamental laws led to equations that were too complex to be solved. About ninety years later, with the help of transistor-based digital computers, the development of quantum chemistry continues to flourish, and many powerful methods, such as Hartree-Fock, configuration interaction, density functional theory, coupled-cluster, and quantum Monte Carlo, have been developed to tackle the complex equations of quantum chemistry (see e.g. [148] for a historical review). However, as the system size scales up, all of the methods known so far suffer from limitations that make them fail to maintain accuracy with a finite amount of

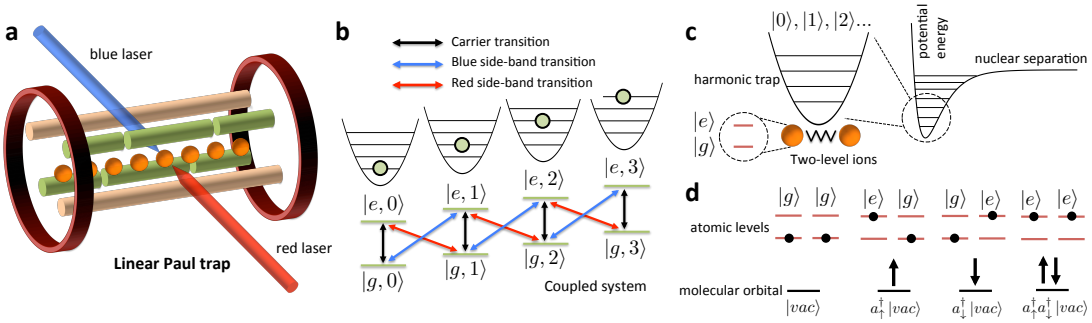

Figure 3.1: Simulating quantum chemistry with trapped ions. (a) Scheme of a trapped-ion setup for quantum simulation, which contains a linear chain of trapped ions confined by a harmonic potential, and external lasers that couple the motional and internal degrees of freedom. (b) Transitions between internal and motional degrees of freedom of the ions in the trap. (c) The normal modes of the trapped ions can simulate the vibrational degrees of freedom of molecules. (d) The internal states of two ions can simulate all four possible configurations of a molecular orbital.

resources [100]. In other words, quantum chemistry remains a hard problem to be solved by the current computer technology.

As envisioned by Feynman [70], one should be able to efficiently solve problems of quantum systems with a quantum computer. Instead of solving the complex equations, this approach, known as *quantum simulation* (see the recent reviews in Refs. [8, 117, 264]), aims to solve the problems by simulating target systems with another controllable quantum system, or qubits. Indeed, simulating many-body systems beyond classical resources will be a cornerstone of quantum computers. Quantum simulation is a very active field of study and various methods have been developed. Quantum simulation methods have been proposed for preparing specific states such as ground [2, 6, 136, 140, 200, 260] and thermal states [24, 141, 201, 226, 263, 264, 268], simulating time evolution [43, 50, 116, 135, 146, 256, 266], and the measurement of physical observables [115, 142, 161, 254].

Trapped-ion systems (see Fig. 3.1) are currently one of the most sophisticated tech-

nologies developed for quantum information processing [95]. These systems offer an unprecedented level of quantum control, which opens new possibilities for obtaining physico-chemical information about quantum chemical problems. The power of trapped ions for quantum simulation is manifested by the high-precision control over both the internal degrees of freedom of the individual ions and the phonon degrees of freedom of the collective motions of the trapped ions, and the high-fidelity initialization and measurement [95, 138]. Up to 100 quantum logic gates have been realized for six qubits with trapped ions [135], and quantum simulators involving 300 ions have been demonstrated [36].

In this work, we present an efficient toolkit for solving quantum chemistry problems based on the state-of-the-art in trapped-ion technologies. The toolkit comprises two components *i*) First, we present a hybrid quantum-classical variational optimization method, called quantum-assisted optimization, for approximating both ground-state energies and the ground-state eigenvectors for electronic problems. The optimized eigenvector can then be taken as an input for the phase estimation algorithm to project out the exact eigenstates and hence the potential-energy surfaces (see Fig. 3.2). Furthermore, we extend the application of the unitary coupled-cluster method [225]. This allows for the application of a method developed for classical numerical computations in the quantum domain. *ii*) The second main component of our toolkit is the optimized use of trapped-ion phonon degrees of freedom not only for quantum-gate construction, but also for simulating molecular vibrations, representing a mixed digital-analog quantum simulation. The phonon degrees of freedom in trapped-ion systems provide a natural platform for addressing spin-boson or fermion-boson-type problems through quantum simulation [42, 43, 81, 133, 170, 176]. It is noteworthy to mention that, contrary to the continuous of modes required for full-

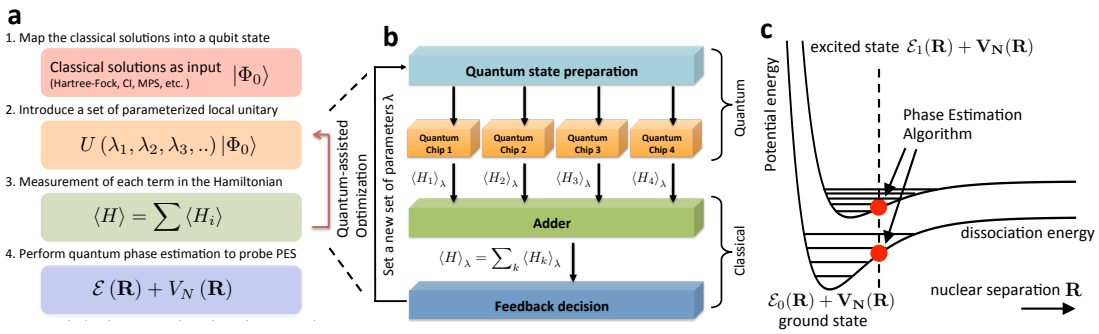

Figure 3.2: Outline of the quantum-assisted optimization method. (a) The key steps for quantum assisted optimization, which starts from classical solutions. For each new set of parameters λ's, determined by a classical optimization algorithm, the expectation value $\langle H\rangle$ is calculated. The potential energy surface is then obtained by quantum phase estimation. (b) Quantum measurements are performed for the individual terms in H, and the sum is obtained classically. (c) The same procedure is applied for each nuclear configuration $\mathbf{R}$ to probe the energy surface.

fledged quantum field theories, quantum simulations of quantum chemistry problems could reach realistic conditions for finite bosonic and fermionic mode numbers. Consequently, trapped ions can be exploited to solve dynamical problems involving linearly or non-linearly coupled oscillators, e.g., spin-boson models [137, 175], that are difficult to solve either analytically or numerically with a classical computer. Furthermore, we have also developed a novel protocol to measure correlation functions of observables in trapped ions that will be crucial for the quantum simulation of quantum chemistry.

3.2 Trapped ions for quantum chemistry

Quantum chemistry deals with the many-body problem involving electrons and nuclei. Thus, it is very well suited for being simulated with trapped-ion systems, as we will show below. The full quantum chemistry Hamiltonian, $H = T_e + V_e + T_N + V_N + V_{eN}$, is a sum of the kinetic energies of the electrons $T_e \equiv -\frac{\hbar^2}{2m}\sum_i \nabla^2_{e,i}$ and nuclei $T_N \equiv -\sum_i \frac{\hbar^2}{2M_i}\nabla^2_{N,i}$, and the electron-electron $V_e \equiv \sum_{j>i} e^2/|\mathbf{r}_i - \mathbf{r}_j|$, nuclei-nuclei $V_N \equiv$

$\sum_{j>i} Z_i Z_j e^2 / |\mathbf{R}_i - \mathbf{R}_j|$, and electron-nuclei $V_{eN} \equiv -\sum_{i,j} Z_j e^2 / |\mathbf{r}_i - \mathbf{R}_j|$ potential energies, where $\mathbf{r}$ and $\mathbf{R}$ respectively refer to the electronic and nuclear coordinates.

In many cases, it is more convenient to work on the second-quantization representation for quantum chemistry. The advantage is that one can choose a good fermionic basis set of molecular orbitals, $|p\rangle = c_p^\dagger |vac\rangle$, which can compactly capture the low-energy sector of the chemical system. This kind of second quantized fermionic Hamiltonians are efficiently simulatable in trapped ions [43]. To be more specific, we will choose first $M > N$ orbitals for an N-electron system. Denote $\phi_p(\mathbf{r}) \equiv \langle \mathbf{r}| p\rangle$ as the single-particle wavefunction corresponding to mode p. The electronic part, $H_e(\mathbf{R}) \equiv T_e + V_{eN}(\mathbf{R}) + V_e$, of the Hamiltonian H can be expressed as follows:

$$H_e(\mathbf{R}) = \sum_{pq} h_{pq} c_p^\dagger c_q + \frac{1}{2} \sum_{pqrs} h_{pqrs} c_p^\dagger c_q^\dagger c_r c_s, \tag{3.1}$$

where h_{pq} is obtained from the single-electron integral

$$h_{pq} \equiv -\int d\mathbf{r} \phi_p^*(\mathbf{r}) \left(T_e + V_{eN}\right) \phi_q(\mathbf{r}), \tag{3.2}$$

and h_{pqrs} comes from the electron-electron Coulomb interaction,

$$h_{pqrs} \equiv \int d\mathbf{r}_1 d\mathbf{r}_2 \phi_p^*(\mathbf{r}_1) \phi_q^*(\mathbf{r}_2) V_e(|\mathbf{r}_1 - \mathbf{r}_2|) \phi_r(\mathbf{r}_2) \phi_s(\mathbf{r}_1). \tag{3.3}$$

We note that the total number of terms in H_e is $O(M^4)$; typically M is of the same order as N. Therefore, the number of terms in H_e scales polynomially in N, and the integrals $\{h_{pq}, h_{pqrs}\}$ can be numerically calculated by a classical computer with polynomial resources [6].

To implement the dynamics associated with the electronic Hamiltonian in Eq. (3.1)

with a trapped-ion quantum simulator, one should take into account the fermionic nature of the operators c_p and $c_q^\dagger$. We invoke the Jordan-Wigner transformation (JWT), which is a method for mapping the occupation representation to the spin (or qubit) representation [188]. Specifically, for each fermionic mode p, an unoccupied state $|0\rangle_p$ is represented by the spin-down state $|\downarrow\rangle_p$, and an occupied state $|1\rangle_p$ is represented by the spin-up state $|\uparrow\rangle_p$. The exchange symmetry is enforced by the Jordan-Wigner transformation: $c_p^\dagger = (\prod_{m<p}\sigma_m^z)\sigma_p^+$ and $c_p = (\prod_{m<p}\sigma_m^z)\sigma_p^-$, where $\sigma^\pm \equiv (\sigma^x \pm i\sigma^y)/2$. Consequently, the electronic Hamiltonian in Eq. (3.1) becomes highly nonlocal in terms of the Pauli operators $\{\sigma^x, \sigma^y, \sigma^z\}$, i.e.,

$$H_e \xrightarrow[\text{JWT}]{} \sum_{i,j,k...\in\{x,y,z\}} g_{ijk...}\left(\sigma_1^i \otimes \sigma_2^j \otimes \sigma_3^k ...\right) \quad . \tag{3.4}$$

Nevertheless, the simulation can still be made efficient with trapped ions, as we shall discuss below.

In trapped-ion physics two metastable internal levels of an ion are typically employed as a qubit. Ions can be confined either in Penning traps or radio frequency Paul traps [138], and cooled down to form crystals. Through sideband cooling the ions motional degrees of freedom can reach the ground state of the quantum Harmonic oscillator, that can be used as a quantum bus to perform gates among the different ions. Using resonance fluorescence with a cycling transition quantum non demolition measurements of the qubit can be performed. The fidelities of state preparation, single- and two-qubit gates, and detection, are all above 99% [95].

The basic interaction of a two-level trapped ion with a single-mode laser is given by [95], $H = \hbar\Omega\sigma_+ e^{-i(\Delta t-\phi)}\exp(i\eta[ae^{-i\omega_t t} + a^\dagger e^{i\omega_t t}]) + \text{H.c.}$, where $\sigma_\pm$ are the atomic raising and lowering operators, a ($a^\dagger$) is the annihilation (creation) operator

of the considered motional mode, and Ω is the Rabi frequency associated to the laser strength. $\eta = kz_0$ is the Lamb-Dicke parameter, with k the wave vector of the laser and $z_0 = \sqrt{\hbar/(2m\omega_t)}$ the ground state width of the motional mode. ϕ is a controllable laser phase and Δ the laser-atom detuning.

In the Lamb-Dicke regime where $\eta\sqrt{\langle(a+a^\dagger)^2\rangle} \ll 1$, the basic interaction of a two-level trapped ion with a laser can be rewritten as $H = \hbar\Omega[\sigma_+ e^{-i(\Delta t-\phi)} + i\eta\sigma_+ e^{-i(\Delta t-\phi)}(ae^{-i\omega_t t} + a^\dagger e^{i\omega_t t}) + \text{H.c.}]$

By adjusting the laser detuning Δ, one can generate the three basic ion-phonon interactions, namely: the carrier interaction ($\Delta = 0$), $H_c = \hbar\Omega(\sigma_+ e^{i\phi} + \sigma_- e^{-i\phi})$, the red sideband interaction, ($\Delta = -\omega_t$), $H_r = i\hbar\eta\Omega(\sigma_+ a e^{i\phi} - \sigma_- a^\dagger e^{-i\phi})$, and the blue sideband interaction, ($\Delta = \omega_t$), $H_b = i\hbar\eta\Omega(\sigma_+ a^\dagger e^{i\phi} - \sigma_- a e^{-i\phi})$. By combining detuned red and blue sideband interactions, one obtains the Mølmer-Sørensen gate [174], which is the basic building block for our methods. With combinations of this kind of gates, one can obtain dynamics as the associated one to H_e in Eq. (3.4), that will allow one to simulate arbitrary quantum chemistry systems.

3.3 Quantum-assisted optimization

Quantum-assisted optimization [196] (see also Fig. 3.2) for obtaining ground-state energies aims to optimize the use of quantum coherence by breaking down the quantum simulation through the use of both quantum and classical processors; the quantum processor is strategically employed for expensive tasks only.

To be more specific, the first step of quantum-assisted optimization is to prepare a set of quantum states $\{|\psi_\lambda\rangle\}$ that are characterized by a set of parameters $\{\lambda\}$. After the state is prepared, the expectation value $E_\lambda \equiv \langle\psi_\lambda| H |\psi_\lambda\rangle$ of the Hamiltonian H

will be measured directly, without any quantum evolution in between. Practically, the quantum resources for the measurements can be significantly reduced when we divide the measurement of the Hamiltonian $H = \sum_i H_i$ into a polynomial number of small pieces $\langle H_i \rangle$ (cf Eq. (3.4)). These measurements can be performed in a parallel fashion, and no quantum coherence is needed to maintain between the measurements (see Fig. 3.2a and 3.2b). Then, once a data point of E_λ is obtained, the whole procedure is repeated for a new state $\{|\psi'_\lambda\rangle\}$ with another set of parameters $\{\lambda'\}$. The choice of the new parameters is determined by a classical optimization algorithm that aims to minimize E_λ (see Methods). The optimization procedure is terminated after the value of E_λ converges to some fixed value.

Finally, for electronic Hamiltonians $H_e(\mathbf{R})$, the optimized state can then be sent to a quantum circuit of phase estimation algorithm to produce a set of data point for some $\mathbf{R}$ on the potential energy surfaces (Fig. 3.2c shows the 1D case). After locating the local minima of the ground and excited states, vibronic coupling for the electronic structure can be further studied (see Supplementary Material).

The performance of quantum-assisted optimization depends crucially on (a) the choice of the variational states, and (b) efficient measurement methods. We found that the unitary coupled-cluster (UCC) states [225] are particularly suitable for being the input state for quantum-assisted optimization, where each quantum state $|\psi_\lambda\rangle$ can be prepared efficiently with standard techniques in trapped ions. Furthermore, efficient measurement methods for H_e are also available for trapped ion systems. We shall discuss these results in detail in the following sections.

3.4 Unitary coupled-cluster (UCC) ansatz

The unitary coupled-cluster (UCC) ansatz [225] assumes electronic states $|\psi\rangle$ have the following form, $|\psi\rangle = e^{T-T^\dagger}|\Phi\rangle$, where $|\Phi\rangle$ is a reference state, which can be, e.g., a Slater determinant constructed from Hartree-Fock molecular orbitals. The particle-hole excitation operator, or cluster operator T, creates a linear combination of excited Slater determinants from $|\Phi\rangle$. Usually, T is divided into subgroups based on the particle-hole rank. More precisely, $T = T_1 + T_2 + T_3 + ... + T_N$ for an N-electron system, where $T_1 = \sum_{i,a} t_i^a c_a^\dagger c_i$, $T_2 = \sum_{i,j,a,b} t_{ij}^{ab} c_a^\dagger c_b^\dagger c_j c_i$, and so on.

Here $c_a^\dagger$ creates an electron in the orbital a. The indices a, b label unoccupied orbitals in the reference state $|\Phi\rangle$, and i, j label occupied orbitals. The energy obtained from UCC, namely $E = \langle\Phi| e^{T^\dagger - T} H e^{T-T^\dagger} |\Phi\rangle$ is a variational upper bound of the exact ground-state energy.

The key challenge for implementing UCC on a classical computer is that the computational resource grows exponentially. It is because, in principle, one has to expand the expression $\tilde{H} \equiv e^{T^\dagger - T} H e^{T-T^\dagger}$ into an infinity series, using the Baker-Campbell-Hausdorff expansion. Naturally, one has to rely on approximate methods [131, 225] to truncate the series and keep track of finite numbers of terms. Therefore, in order to make good approximations by perturbative methods, i.e., assuming T is small, one implicitly assumes that the reference state $|\Phi\rangle$ is a good solution to the problem. However, in many cases, such an assumption is not valid and the use of approximate UCC breaks down. We explain below how implementing UCC on a trapped-ion quantum computer can overcome this problem.

	Simulating Quantum Chemistry	**Implementation with Trapped Ions**
Hamiltonian transformation:	The fermionic (electronic) Hamiltonian H_e is transformed into a spin Hamiltonian through the Jordan-Wigner transformation.	The spin degrees of freedom in H_e are represented by the internal degrees of freedom of the trapped ions.
	$$H_e \rightarrow \sum_{i,j,k,\cdots \in \{x,y,z\}} g_{ijk\cdots} \left(\sigma_1^i \otimes \sigma_2^j \otimes \sigma_3^k \cdots\right) \equiv \sum_{l=1}^{m} H_l$$	
Simulation of time evolution:	The time evolution operator $e^{-iH_e t}$ is split into n small-time (t/n) pieces $e^{-iH_l t/n}$ through the Suzuki-Trotter expansion.	Each individual term $e^{-iH_l t/n}$ can be simulated with trapped ions through the use of Mølmer-Sørensen gates U_{MS}. Explicitly, $e^{-iH_l t/n} = U_{MS}\left(\frac{-\pi}{2}, 0\right) U_{\sigma_z}(\phi)\, U_{MS}\left(\frac{\pi}{2}, 0\right)$.
	$$e^{-i\sum_{l=1}^{m} H_l t} \approx \left(e^{-iH_1 t/n} e^{-iH_2 t/n} \cdots e^{-iH_m t/n}\right)^n$$	
Obtaining average energy:	The average energy $\langle H_e \rangle$ of the Hamiltonian can be obtained through the sum of the individual terms $\langle H_l \rangle$, which reduces to the measurement of products of Pauli matrices.	For any prepared state $\|\psi\rangle$, average values of the products of Pauli matrices $J_{ijk\cdots} \equiv \sigma_1^i \otimes \sigma_2^j \otimes \sigma_3^k \cdots$ can be measured by first applying the pseudo time evolution operator $e^{-i(\pi/4) J_{ijk\cdots}}$ to $\|\psi\rangle$ and then measuring $\langle \sigma_1^z \rangle$.
Measuring eigenvalues:	The eigenvalues of the Hamiltonian can be obtained through the phase estimation algorithm. Good trial states can be obtained through classical computing, or the unitary coupled-cluster method.	The phase estimation algorithm can be implemented through the simulation of controlled time evolutions.
Molecular vibrations:	The inclusion of vibrational degrees of freedom is necessary for corrections on the Born-Oppenheimer picture in the electronic structure of molecules.	The vibrational degrees of freedom are represented by the quantized vibrational motion of the trapped ions.

Table 3.1: Using trapped ions to simulate quantum chemistry

3.4.1 Implementation of UCC through time evolution

We can generate the UCC state by simulating a pseudo time evolution through Suzuki-Trotter expansion on the evolution operator $e^{T-T^\dagger}$ [146]. To proceed, we consider an N-electron system with M, where $M > N$, molecular orbitals (including spins). We need totally M qubits; the reference state is the Hartree-Fock state where N orbitals are filled, and $M-N$ orbitals are empty, i.e, $|\Phi\rangle = |000...0111...1\rangle$. We also define an effective Hamiltonian $K \equiv i\left(T - T^\dagger\right)$, which means that we should prepare the state $e^{-iK}|\Phi\rangle$.

We decompose K into subgroups $K = K_1 + K_2 + K_3 + ... + K_P$, where $P \leq N$, and $K_i \equiv i(T_i - T_i^\dagger)$. We now write $e^{-iK} = \left(e^{-iK\delta}\right)^{1/\delta}$ for some dimensionless constant δ. For small δ, we have $e^{-iK\delta} \approx e^{-iK_P\delta}...e^{-iK_2\delta}e^{-iK_1\delta}$. Since each K_j contains $N^j(M-N)^j$ terms of the creation $c^\dagger$ and annihilation c operators, we will need to individually simulate each term separately, e.g., $e^{-i\left(tc_a^\dagger c_i - t^* c_i^\dagger c_a\right)}$ and $e^{-i\left(tc_a^\dagger c_b^\dagger c_j c_i - t^* c_i^\dagger c_j^\dagger c_b c_a\right)}$, which can be implemented by transforming into spin operators through Jordan-Wigner transformation. The time evolution for each term can be simulated with a quantum circuit involving many nonlocal controlled gates, which can be efficiently implemented with trapped ions as we shall see below.

3.4.2 Implementation of UCC and simulation of time evolution with trapped-ions

Our protocol for implementing the UCC ansatz requires the simulation of the small-time t/n evolution of non-local product of Pauli matrices of the form: $e^{-iH_l t/n}$, where $H_l = g_l \sigma_1^i \sigma_2^j \sigma_3^k \cdots$ for $i, j, k \in \{x, y, z\}$. Note that for any N-spin interaction, the $e^{-iH_l t/n}$ terms are equivalent to $e^{i\phi\sigma_1^z\sigma_2^x\sigma_3^x\cdots\sigma_N^x}$ through local spin rotations, which are

simple to implement on trapped ions. Such a non-local operator can be implemented using the multi-particle Mølmer-Sørensen gate [43, 176]:

$$U_{\mathrm{MS}}(\theta,\varphi) \equiv \exp\left[-i\theta(\cos\varphi S_x + \sin\varphi S_y)^2/4\right], \tag{3.5}$$

where $S_{x,y} \equiv \sum_i \sigma_i^{x,y}$ is a collective spin operator. Explicitly,

$$e^{i\phi\sigma_1^z\sigma_2^x\sigma_3^x\cdots\sigma_N^x} = U_{MS}\left(\tfrac{-\pi}{2},0\right) R_N(\phi) U_{MS}\left(\tfrac{\pi}{2},0\right) \quad . \tag{3.6}$$

Here $R_N(\phi)$ is defined as follows: for any $m \in \mathbb{N}$, $R_N(\phi) = e^{\pm i\phi\sigma_1^z}$ for $N = 4m \pm 1$, and (ii) $R_N(\phi) = e^{i\phi\sigma_1^y}$ for $N = 4m$, and (iii) $R_N(\phi) = e^{-i\phi\sigma_1^y}$ for $N = 4m-2$.

It is remarkable that the standard quantum-circuit treatment (e.g. see Ref. [251]) for implementing each $e^{-iH_l t/n}$ involves as many as $2N$ two-qubit gates for simulating N fermionic modes; in our protocol one needs only two Mølmer-Sørensen gates, which are straightforwardly implementable with current trapped-ion technology. Furthermore, the local rotation $R_N(\phi)$ can also include motional degrees of freedom of the ions for simulating arbitrary fermionic Hamiltonians coupled linearly to bosonic operators a_k and $a_k^\dagger$.

3.5 Measurement of arbitrarily-nonlocal spin operators

For any given state $|\psi\rangle$, we show how to encode expectation value of products of Pauli matrices $\langle\sigma_1^i \otimes \sigma_2^j \otimes \sigma_3^k \otimes \cdots\rangle \equiv \langle\psi|\,\sigma_1^i \otimes \sigma_2^j \otimes \sigma_3^k \otimes \cdots|\psi\rangle$, where $i,j,k \in \{x,y,z\}$, onto an expectation value of a single qubit. The idea is to first apply the unitary evolution of the form: $e^{-i\theta(\sigma_1^i \otimes \sigma_2^j \otimes \cdots)}$, which as we have seen (cf Eq. 3.6) can be generated by trapped ions efficiently, to the state $|\psi\rangle$ before the measurement. For example, defin-

ing $|\psi_\theta\rangle \equiv e^{-i\theta\left(\sigma_1^x \otimes \sigma_2^x \otimes \cdots\right)} |\psi\rangle$, we have the relation

$$\langle\psi_\theta| \sigma_1^z |\psi_\theta\rangle = \cos(2\theta) \langle\sigma_1^z\rangle + \sin(2\theta) \langle\sigma_1^y \otimes \sigma_2^x \otimes \cdots\rangle \quad , \tag{3.7}$$

which equals $\langle\psi|(\sigma_1^y \otimes \sigma_2^x \otimes ...)|\psi\rangle$ for $\theta = \pi/4$. Note that the application of this method requires the measurement of one qubit only, making this technique especially suited for trapped ion systems where the fidelity of the measurement of one qubit is 99.99% [177].

This method can be further extended to include bosonic operators in the resulting expectation values. For example, re-define $|\psi_\theta\rangle \equiv e^{-i\theta(\sigma_1^i \otimes \sigma_2^j \otimes \cdots) \otimes \left(a + a^\dagger\right)} |\psi\rangle$ and consider $\theta \to \theta \left(a + a^\dagger\right)$ in Eq. (3.7). We can obtain the desired correlation through the derivative of the single-qubit measurement:

$$\partial_\theta \langle\psi_\theta| \sigma_1^z |\psi_\theta\rangle|_{\theta=0} = -2\langle(\sigma_1^y \otimes \sigma_2^x \otimes \cdots)\left(a + a^\dagger\right)\rangle. \tag{3.8}$$

Note that the evolution operator of the form $e^{-i\theta(\sigma_1^i \otimes \sigma_2^j \otimes \cdots) \otimes (a + a^\dagger)}$ can be generated by replacing the local operation $R_N(\phi)$ in Eq. 3.6 with $e^{\pm i\phi\sigma_1^i (a + a^\dagger)}$. This technique allows us to obtain a diverse range of correlations between bosonic and internal degrees of freedom.

3.6 Probing potential energy surfaces

In the Born-Oppenheimer (BO) picture, the potential energy surface $\mathcal{E}_k(\mathbf{R}) + V_N(\mathbf{R})$ associated with each electronic eigenstate $|\phi_k\rangle$ is obtained by scanning the eigenvalues $\mathcal{E}_k(\mathbf{R})$ for each configurations of the nuclear coordinates $\{\mathbf{R}\}$. Of course, we can apply the standard quantum phase estimation algorithm [118] that allows us to ex-

tract the eigenvalues. However, this can require many ancilla qubits. In fact, locating these eigenvalues can be achieved by the phase estimation method utilizing one extra ancilla qubit [140] corresponding, in our case, to one additional ion.

This method works as follows: suppose we are given a certain quantum state $|\psi\rangle$ (which may be obtained from classical solutions with quantum-assisted optimization) and an electronic Hamiltonian $H_e(\mathbf{R})$ (cf. Eq. (3.1)). Expanding the input state, $|\psi\rangle = \sum_k \alpha_k |\phi_k\rangle$, by the eigenstate vectors $|\phi_k\rangle$ of $H_e(\mathbf{R})$, where $H_e(\mathbf{R})|\phi_k\rangle = \mathcal{E}_k(\mathbf{R})|\phi_k\rangle$, then for the input state $|0\rangle|\psi\rangle$, the quantum circuit of the quantum phase estimation produces the following output state, $\left(1/\sqrt{2}\right)\sum_k \alpha_k \left(|0\rangle + e^{-i\omega_k t}|1\rangle\right)|\phi_k\rangle$, where $\omega_k = \mathcal{E}_k/\hbar$. The corresponding reduced density matrix,

$$\frac{1}{2}\begin{pmatrix} 1 & \sum_k |\alpha_k|^2 e^{i\omega_k t} \\ \sum_k |\alpha_k|^2 e^{-i\omega_k t} & 1 \end{pmatrix}, \tag{3.9}$$

of the ancilla qubit contains the information about the weight (amplitude-square) $|\alpha_k|^2$ of the eigenvectors $|\phi_k\rangle$ in $|\psi\rangle$ and the associated eigenvalues ω_k in the off-diagonal matrix elements. All $|\alpha_k|^2$'s and ω_k's can be extracted by repeating the quantum circuit for a range of values of t and performing a (classical) Fourier transform to the measurement results. The potential energy surface is obtained by repeating the procedure for different values of the nuclear coordinates $\{\mathbf{R}\}$.

3.7 Numerical investigation

In order to show the feasibility of our protocol, we can estimate the trapped-ion resources needed to simulate, e.g., the prototypical electronic Hamiltonian $H_e = \sum h_{pq} a_p^\dagger a_q + (1/2)\sum h_{pqrs} a_p^\dagger a_q^\dagger a_r a_s$ as described in Eq. (3.1), for the specific case of the H_2 molecule

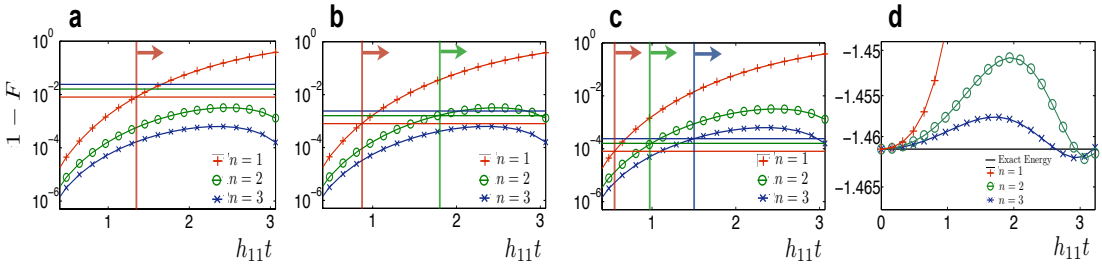

Figure 3.3: Digital error $1-F$ (curves) along with the accumulated gate error (horizontal lines) versus time in h_{11} energy units, for $n=1,2,3$ Trotter steps in each plot, considering a protocol with an error per Trotter step of $\epsilon=10^{-3}$ (a), $\epsilon=10^{-4}$ (b) and $\epsilon=10^{-5}$ (c). The initial state considered is $|\uparrow\uparrow\downarrow\downarrow\rangle$, in the qubit representation of the Hartree-Fock state in a molecular orbital basis with one electron on the first and second orbital. Vertical lines and arrows define the time domain in which the dominant part of the error is due to the digital approximation. d) Energy of the system, in h_{11} units, for the initial state $|\uparrow\uparrow\downarrow\downarrow\rangle$ for the exact dynamics, versus the digitized one. For a protocol with three Trotter steps the energy is recovered up to a negligible error.

in a minimal STO-3G basis. This is a two-electron system represented in a basis of four spin-orbitals. The hydrogen atoms were separated by 0.75 Å, near the equilibrium bond distance of the molecule. The Hamiltonian is made up of 12 terms, that include 4 local ion operations and 8 non-local interactions. Each of the non-local terms can be done as a combination of two Mølmer-Sørensen (MS) gates and local rotations, as described in Table 3.1. Therefore, to implement the dynamics, one needs 16 MS gates per Trotter step and a certain number of local rotations upon the ions. Since $\pi/2$ MS gates can be done in $\sim 50\mu$s, and local rotations can be performed in negligible times ($\sim 1\mu$s) [95, 135], the total simulation time can be assumed of about 800 μs for the $n=1$ protocol, 1.6 ms and 2.4 ms for the $n=2$ and $n=3$ protocols. Thus total simulation times are within the decoherence times for trapped-ion setups, of about 30 ms [95]. In a digital protocol performed on real quantum systems, each gate is affected by an error. Thus, increasing the number of Trotter steps leads to an accumulation of the single gate error. To implement an effective quantum simulation,

on one hand one has to increase the number of steps to reduce the error due to the digital approximation, on the other hand one is limited by the accumulation of the single gate error. We plot in Fig. 3.3a, 3.3b, 3.3c, the fidelity loss $1-|\langle\Psi_S|\Psi_E\rangle|^2$ of the simulated state $|\Psi_S\rangle$ versus the exact one $|\Psi_E\rangle$, for the hydrogen Hamiltonian, starting from the initial state with two electrons in the first two orbitals. We plot, along with the digital error, three horizontal lines representing the accumulated gate error, for $n = 1, 2, 3$ in each plot, considering a protocol with an error per Trotter step of $\epsilon = 10^{-3}$ (a), $\epsilon = 10^{-4}$ (b) and $\epsilon = 10^{-5}$ (c). To achieve a reasonable fidelity, one has to find a number of steps that fits the simulation at a specific time. The vertical lines and arrows in the figure mark the time regions in which the error starts to be dominated by the digital error. Trapped-ion two-qubit gates are predicted to achieve in the near future fidelities of 10^{-4} [121], thus making the use of these protocols feasible. In Fig. 3.3d we plot the behavior of the energy of the system for the initial state $|\uparrow\uparrow\downarrow\downarrow\rangle$ for the exact dynamics, versus the digitized one. Again, one can observe how the energy can be retrieved with a small error within a reduced number of digital steps.

3.8 Conclusions

Summarizing, we have proposed a quantum simulation toolkit for quantum chemistry with trapped ions. This paradigm in quantum simulations has several advantages: an efficient electronic simulation, the possibility of interacting electronic and vibrational degrees of freedom, and the increasing scalability provided by trapped-ion systems. This approach for solving quantum chemistry problems aims to combine the best of classical and quantum computation.

3.9 Methods

To implement the optimization with the UCC wavefunction ansatz on a trapped-ion quantum simulator, our proposal is to first employ classical algorithms to obtain approximate solutions [131, 225]. Then, we can further improve the quality of the solution by searching for the true minima with an ion trap. The idea is as follows: first we create a UCC ansatz by the Suzuki-Trotter method described in the previous section. Denote this choice of the cluster operator as $T^{(0)}$, and other choices as $T^{(k)}$ with $k = 1, 2, 3, ...$ The corresponding energy $E_0 = \langle\Phi| e^{T^{(0)\dagger} - T^{(0)}} H e^{T^{(0)} - T^{(0)\dagger}} |\Phi\rangle$ of the initial state is obtained by a classical computer.

Next, we choose another set of cluster operator $T^{(1)}$ with is a perturbation around $T^{(0)}$. Define the new probe state $|\phi_k\rangle \equiv e^{T^{(k)} - T^{(k)\dagger}} |\Phi\rangle$. Then, the expectation value of the energy $E_1 = \langle\Phi| e^{T^{(1)\dagger} - T^{(1)}} H e^{T^{(1)} - T^{(1)\dagger}} |\Phi\rangle = \langle\phi_1| H |\phi_1\rangle$ can be obtained by measuring components of the second quantized Hamiltonian,

$$\langle\phi_1| H |\phi_1\rangle = \sum\nolimits_{pqrs} \tilde{h}_{pqrs} \langle\phi_1| c_p^\dagger c_q^\dagger c_r c_s |\phi_1\rangle . \tag{3.10}$$

Recall that the coefficients $\tilde{h}_{pqrs}$ are all precomputed and known.

In order to obtain measurement results for the operators $\langle\phi_1| c_p^\dagger c_q^\dagger c_r c_s |\phi_1\rangle$, we will first convert the fermion operators into spin operators via Jordan-Wigner transformation; the same procedure is applied for creating the state $|\phi_1\rangle$. The quantum measurement for the resulting products of Pauli matrices can be achieved efficiently with trapped ions, using the method we described.

The following steps are determined through a classical optimization algorithm. There can be many choices for such an algorithm, for example gradient descent method,

Nelder-Mead method, or quasi-Newton methods. For completeness, we summarize below the application of gradient descent method to our optimization problem.

First we define the vector $\mathbf{T}^{(k)} = (t_i^{a(k)}, t_{ij}^{ab(k)}, ...)^T$ to contain all coefficients in the cluster operator $T^{(k)}$ at the k-th step. We can also write the expectation value $E\left(\mathbf{T}^{(k)}\right) \equiv \langle\phi_k| H |\phi_k\rangle$ for each step as a function of $\mathbf{T}^{(k)}$. The main idea of the gradient descent method is that $E\left(\mathbf{T}^{(k)}\right)$ decreases fastest along the direction of the negative gradient of $E\left(\mathbf{T}^{(k)}\right)$, $-\nabla E\left(\mathbf{T}^{(k)}\right)$. Therefore, the $(k+1)$-th step is determined by the following relation:

$$\mathbf{T}^{(k+1)} = \mathbf{T}^{(k)} - a_k \nabla E(\mathbf{T}^{(k)}), \tag{3.11}$$

where a_k is an adjustable parameter; it can be different for each step. To obtain values of the gradient $\nabla E\left(\mathbf{T}^{(k)}\right)$, one may use the finite-difference method to approximate the gradient. However, numerical gradient techniques are often susceptible to numerical instability. Alternatively, we can invoke the Hellman-Feynman theorem and get, e.g., $\left(\partial/\partial t_i^a\right) E(\mathbf{T}^{(k)}) = \langle\phi_k|[H, c_a^\dagger c_i] |\phi_k\rangle$, which can be obtained with a method similar to that for obtaining $E(\mathbf{T}^{(k)})$.

Finally, as a valid assumption for general cases, we assume our parametrization of UCC gives a smooth function for $E\left(\mathbf{T}^{(k)}\right)$. Thus, it follows that $E\left(\mathbf{T}^{(0)}\right) \geq E\left(\mathbf{T}^{(1)}\right) \geq E\left(\mathbf{T}^{(2)}\right) \geq \cdots$, and eventually $E\left(\mathbf{T}^{(k)}\right)$ converges to a minimum value for large k. Finally, we can also obtain the optimized UCC quantum state.

3.10 Supplementary Material

In this Supplementary Material we give further details of our proposal, including a thorough explanation of the quantum simulation of molecules involving fermionic and bosonic degrees of freedom with trapped ions, and electric dipole transition measure-

ments with a trapped-ion quantum simulator.

3.10.1 Quantum simulation

In general, quantum simulation can be divided into two classes, namely analog and digital. Analog quantum simulation requires the engineering of the Hamiltonian of a certain system to mimic the Hamiltonian of a target system. Digital quantum simulation employs a quantum computer, which decomposes the simulation process into pieces of sub-modules such as quantum logic gates. However, the use of quantum logic gates is not absolutely necessary for digital quantum simulation. For example, consider the case of trapped ions; we will see that certain simulation steps requires us to apply quantum logic gates to implement fermionic degrees of freedom, together with some quantum operations for controlling the vibronic degrees of freedom, which are analog and will implement bosonic modes.

For simulating quantum chemistry, it is possible to work in either the first-quantization representation or the second-quantization representation. This work mainly includes the latter approach, because the number of qubits required is less than that in the former approach, especially when low-energy state properties are considered. However, we note that many techniques described here are also applicable for the first-quantization approach.

3.10.2 Computational complexity of quantum chemistry

To the best of our knowledge, there is no rigorous proof showing that quantum computers are capable of solving all ground-state problems in quantum chemistry. Instead, some results indicate that some ground-state problems in physics and chem-

istry are computationally hard problems [252]. For example, the N-representability problem is known to be QMA-complete, and finding the universal functional in density functional theory is known to be QMA-hard. In spite of the negative results, quantum computers can still be valuable for solving a wide range of quantum chemistry problems. These include ground state energy computations [2, 6], as well as molecular dynamics [148].

3.10.3 Simulating electronic structure involving molecular vibrations

After the potential surface is constructed by the electronic method, we can include the effect of molecular vibrations by local expansion, e.g. near the equilibrium position, as we show below.

Electronic transitions coupled with nuclear motion

We point out that within the Born-Oppenheimer approximation, the molecular vibronic states are of the form, $\phi_n(\mathbf{r},\mathbf{R})\,\chi_{n,v}(\mathbf{R})$ where $\mathbf{r}$ and $\mathbf{R}$ respectively refers to the electronic and nuclear coordinates. The eigenfunctions $\phi_n(\mathbf{r},\mathbf{R})$ of the electronic Hamiltonian are obtained at a fixed nuclear configuration. The nuclear wavefunction $\chi_{n,v}(\mathbf{R})$, for each electronic eigenstate n, is defined through a nuclear potential surface $E_{\mathrm{el}}^{(n)}(\mathbf{R})$, which is also one of the eigenenergies of the electronic Hamiltonian.

With a quantum computer, the potential energy surface that corresponds to different electronic eigenstates can be systematically probed using the phase estimation method. We can then locate those local minima where the gradient of the energy is zero, and approximate up to second order in $\delta R_\alpha \equiv R_\alpha - R_{\alpha *}$, the deviation of the nu-

clear coordinate R_α from the equilibrium configuration $R_{\alpha*}$. The energy surface can be modeled as

$$E_{\rm el}^{(n)}(\mathbf{R}) \approx E_{\rm el}^{(n)}(\mathbf{R}_*^{(n)}) + \sum_{\alpha,\beta} D_{\alpha\beta}(\mathbf{R}_*^{(n)})\delta R_\alpha \delta R_\beta \,, \tag{3.12}$$

where $D_{\alpha\beta}(\mathbf{R}_*^{(n)}) \equiv (1/2)\,\partial^2 E_{el}^{(n)}(\mathbf{R}{=}\mathbf{R}_*^{(n)})/\partial R_\alpha \partial R_\beta$ is the Hessian matrix. With a change of coordinates for the Hessian matrices, we can always choose to work with the normal modes $\mathbf{x}^{(n)} = \{x_\alpha^{(n)}\}$ for each potential energy surface, such that

$$E_{\rm el}^{(n)}(\mathbf{x}^{(n)}) \approx E_{\rm el}^{(n)}(\mathbf{R}_*^{(n)}) + \frac{1}{2}\sum_\alpha m_\alpha \omega_\alpha^{(n)2} x_\alpha^{(n)2}. \tag{3.13}$$

Most of the important features of vibronic coupling can be captured by considering the transition between two Born-Oppenheimer electronic levels [162]. In the following, we will focus on the method of simulation of the transition between two electronic levels, labeled as $|\uparrow\rangle$ and $|\downarrow\rangle$, when perturbed by an external laser field. The Hamiltonian of the system can be written as

$$H = |\downarrow\rangle\langle\downarrow| \otimes H_G + |\uparrow\rangle\langle\uparrow| \otimes H_E, \tag{3.14}$$

where $H_G \equiv \Delta_g + H_g$ is the Hamiltonian for the nuclear motion in the electronic ground state and similarly $H_E \equiv \Delta_e + H_e$ is the nuclear Hamiltonian in the excited state. Here Δ_g and Δ_e are the energies of the two bare electronic states. In the second-quantized representation,

$$H_g = \sum_k \omega_k^{(g)} a_k^\dagger a_k \quad \text{and} \quad H_e = \sum_k \omega_k^{(e)} b_k^\dagger b_k \tag{3.15}$$

are diagonal, as viewed from their own coordinate systems. However, in general, the two sets of normal modes are related by rotation and translation, which means that a transformation of the kind $b_k = \sum_j s_{kj} a_j + \lambda_k$ is needed for unifying the representations (see Secs. 3.10.5 and 3.10.6 in this Supplementary Material).

To illustrate our method of quantum simulation with trapped ions, it is sufficient to consider one normal mode (for example, linear molecules). For this case, we assume $H_g = \omega^{(g)} a^\dagger a$, $H_e = \omega^{(e)} b^\dagger b$, and $b = a + \lambda$ where λ is a real constant. From Eq. (3.14), we need to simulate the following Hamiltonian,

$$H = H_S + \Omega\left(\sigma_z\right) a^\dagger a + \tfrac{1}{2}\lambda\omega^{(e)}\left(I + \sigma_z\right)\left(a^\dagger + a\right), \tag{3.16}$$

where the term $H_S = \frac{1}{2}\left(\Delta_g - \Delta_e\right)\sigma_z$ contains only local terms of the spin, and $\Omega\left(\sigma_z\right) = \frac{1}{2}\left(\omega^{(g)} + \omega^{(e)}\right) I + \frac{1}{2}\left(\omega^{(g)} - \omega^{(e)}\right)\sigma_z$ represents a spin-dependent frequency for the effective boson mode.

In order to examine the response of the system under external pertubations, we consider the dipole correlation function

$$C_{\mu\mu}\left(t\right) = \sum_n p_n \left\langle n, \downarrow\right| e^{iHt} \mu e^{-iHt} \mu \left|n, \downarrow\right\rangle. \tag{3.17}$$

Under the Condon approximation, assuming real electronic eigenstates, the dipole operator μ has the form,

$$\mu = \mu_{ge}\left(\left|\downarrow\right\rangle\left\langle\uparrow\right| + \left|\uparrow\right\rangle\left\langle\downarrow\right|\right) = \mu_{ge}\sigma_x. \tag{3.18}$$

Thus, the problem of simulating absorption resulting from the coupling of electronic and nuclear motion in chemistry reduces to computing expectation values of the uni-

tary operator

$$U_d = e^{iHt}\sigma_x e^{-iHt}\sigma_x, \tag{3.19}$$

and weighting the final result by $p_n\mu_{ge}^2$. The final spectrum is, of course, obtained through a Fourier transform

$$\sigma_{abs}(\omega) = \int_{-\infty}^{\infty} dt\ e^{-i\omega t} C_{\mu\mu}(t). \tag{3.20}$$

Simulation of vibronic coupling with trapped ions

The dynamics associated with the Hamiltonian in Eq. (3.16) can be generated easily with two trapped ions. As H_S commutes with the rest of the terms in Eq. (3.16), it can be eliminated via a change to an interaction picture. Considering a digital quantum simulation protocol, the remaining task is to implement the interactions $\exp[-i\Omega\left(\sigma_z\right)ta^\dagger a]$ and $\exp[-i\lambda\omega^{(e)}\left(I+\sigma_z\right)\left(a^\dagger+a\right)t/2]$ in trapped ions. The first one corresponds to the evolution associated with a detuned red sideband excitation applied to one of the ions (a dispersive Jaynes-Cummings interaction), and a rotation of its internal state in order to eliminate the residual projective term. To implement the second term we will use both ions. The term related to the operator $\sigma_z(a^\dagger + a)$ corresponds to the evolution under red and blue sideband excitations applied to one of the ions (a Jaynes-Cummings and anti Jaynes-Cummings interactions with appropriate phases). We will use the second ion to implement the term $(a^\dagger + a)$. The latter can be generated by applying again the same scheme of lasers that generates the interaction $\sigma_z(a^\dagger + a)$ where now the operator σ_z acts on the internal state of the second ion. Preparing this state in an eigenstate of σ_z one obtains the desired effective Hamiltonian. As we have shown here, one of the main appeals of a quantum simula-

tion of quantum chemistry with trapped ions is the possibility to include fermionic (electronic) as well as bosonic (vibronic) degrees of freedom, in a new kind of mixed digital-analog quantum simulator. The availability of the motional degrees of freedom in trapped ions, that straightforwardly provide the bosonic modes in an analog way, makes this system especially suited for simulating this kind of chemical problems.

3.10.4 Electric transition dipoles through weak measurement

Here we sketch the method for obtaining the transition dipole between a pair of electronic states $|g\rangle$ and $|e\rangle$. This method is similar, although not identical, to the weak measurement method using a qubit as a measurement probe. To make the presentation of our method more general, our goal is to measure the matrix element $\langle e| A |g\rangle$ for any given Hermitian matrix A. We assume that a potential energy surface between these two electronic levels is probably scanned, and the energy levels for higher excited states can be ignored. Suppose we started with a reasonable good approximation of the ground state $|g\rangle$, and we can prepare the exact ground state using the phase estimation algorithm. Then, we apply a weak perturbation λ, e.g. $e^{-i\lambda Q}$, to the ground state and obtain (to order $O(\lambda)$) the state $|i\rangle \equiv e^{-i\lambda Q} |g\rangle \approx |g\rangle + q\lambda |e\rangle$. Here λ is a small positive real number. The actual form of the Hermitian operator Q is not important, as long as $\langle e| Q |g\rangle \equiv iq \neq 0$. Note that the eigenstates are defined up a phase factor. Therefore, without loss of generality, we can assume q is a positive real number as well. In fact, the absolute value $|q|$ can be measured with repeated applications of the phase estimation algorithm.

Now, we prepare an ancilla qubit in the state $|+\rangle \equiv (|0\rangle + |1\rangle)/\sqrt{2}$, and apply a control-$U_A$, where $U_A \equiv e^{-i\lambda A}$. The resulting state becomes $(|0\rangle |i\rangle + |1\rangle U_A |i\rangle)/\sqrt{2}$. The phase estimation algorithm allows us to perform post-selection to project the

system state to $|e\rangle$. The resulting state of the ancilla qubit is $\propto \langle e|i\rangle|0\rangle + \langle e|U_A|i\rangle|1\rangle$. To the first-order expansion in λ, we have (before normalization)

$$q\lambda\,|0\rangle + (q - i\,\langle e|\,A\,|g\rangle)\,\lambda\,|1\rangle\,, \tag{3.21}$$

where we used $\langle e|i\rangle = q\lambda$, and $\langle e|\,U_A\,|i\rangle = \langle e|\,U_A\,|g\rangle + q\lambda\,\langle e|\,U_A\,|e\rangle = -i\,\langle e|\,A\,|g\rangle\,\lambda + q\lambda$. Since the value of q is known, a state tomography on the ancilla qubit state reveals the value of the matrix element $\langle e|\,A\,|g\rangle$.

Returning to the case of the electric dipole moment, it is defined as $\mu \equiv -e\sum_i \mathbf{r}_i$. In the second quantized form is $\mu = \sum_{pq} u_{pq} a_p^\dagger a_q$, where $u_{pq} \equiv -e\int \phi_p^*(\mathbf{r})\mathbf{r}\phi_q(\mathbf{r})\,d\mathbf{r}$ is nothing but the single-particle integral. The simulation of the corresponding operator $U_A \equiv e^{-i\lambda A}$, with A replaced by μ, can be performed efficiently after performing the Jordan-Wigner transformation.

3.10.5 Derivation of the spin-boson coupling

Consider the full Hamiltonian of two potential energy surfaces,

$$H = |\downarrow\rangle\langle\downarrow| \otimes H_G + |\uparrow\rangle\langle\uparrow| \otimes H_E, \tag{3.22}$$

where

$$H_G \equiv \Delta_g + H_g \tag{3.23}$$

is the Hamiltonian for the nuclear motion in the electronic ground state and similarly

$$H_E \equiv \Delta_e + H_e \tag{3.24}$$

is the nuclear Hamiltonian in the excited state. Here Δ_g and Δ_e are the zero-point energies of the two potential energy surfaces. In the second-quantized representation, we consider one normal mode for each local minimum in the potential energy surface,

$$H_g = \omega^{(g)} a^\dagger a \quad \text{and} \quad H_e = \omega^{(e)} b^\dagger b. \tag{3.25}$$

Here the two normal modes are related by a shift of a real constant λ, namely

$$b = a + \lambda. \tag{3.26}$$

Now, we will rewrite the full Hamiltonian in terms of the Pauli matrix

$$\sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} = |\uparrow\rangle \langle\uparrow| - |\downarrow\rangle \langle\downarrow| . \tag{3.27}$$

First of all, we write $H = H_{SB} + H_S$, where

$$H_{SB} = |\downarrow\rangle \langle\downarrow| \otimes \omega^{(g)} a^\dagger a + |\uparrow\rangle \langle\uparrow| \otimes \omega^{(e)} b^\dagger b, \tag{3.28}$$

and

$$\begin{aligned} H_S &\equiv |\downarrow\rangle \langle\downarrow| \Delta_g + |\uparrow\rangle \langle\uparrow| \Delta_e \\ &= \frac{1}{2} (\Delta_g + \Delta_e) I + \frac{1}{2} (\Delta_g - \Delta_e) \sigma_z. \end{aligned} \tag{3.29}$$

Next, we use Eq. (3.26) to write H_{SB} as

$$H_{SB} = \Omega (\sigma_z) \otimes a^\dagger a + \tfrac{1}{2} \lambda \omega^{(e)} (I + \sigma_z) \otimes \left(a^\dagger + a \right), \tag{3.30}$$

where the frequency of the effective mode becomes spin-dependent,

$$\begin{aligned}\Omega(\sigma_z) &\equiv |\uparrow\rangle\langle\uparrow|\omega^{(g)} + |\downarrow\rangle\langle\downarrow|\omega^{(e)} \\ &= \frac{1}{2}\left(\omega^{(g)} + \omega^{(e)}\right) I + \frac{1}{2}\left(\omega^{(g)} - \omega^{(e)}\right)\sigma_z.\end{aligned} \tag{3.31}$$

3.10.6 Multimode extension of simulating vibronic coupling

In order to extend the method of simulating vibronic coupling to the case with multiple bosonic modes, we now consider the case of Eq. 3.15. If we express the excited state modes in terms of the ground state modes such that

$$b_k = \sum_j s_{kj} a_j + \lambda_k, \tag{3.32}$$

we can write H as

$$H = H_s' + |\downarrow\rangle\langle\downarrow| \otimes H_G + |\uparrow\rangle\langle\uparrow| \otimes H_E, \tag{3.33}$$

where

$$H_G \equiv \sum_k \omega_k^{(g)} a_k^\dagger a_k, \tag{3.34}$$

and

$$H_E \equiv \sum_{kjl} \omega_k^{(e)} s_{kj} s_{lk} a_j^\dagger a_l + \sum_{kj} \omega_k^{(e)} s_{kj} \lambda_k \left(a_j^\dagger + a_j\right). \tag{3.35}$$

In the definition of H_s', the only change from H_s is given by

$$\Delta_e' = \Delta_e + \sum_k \lambda_k^2. \tag{3.36}$$

With knowledge of s_{ij} and λ_i for all modes, we can then repeat the above procedure to determine the absorption spectrum for a complicated system using a quantum

computer. The above Hamiltonian can be written in a form more familiar to quantum computation as

$$\begin{aligned} H = H'_s &+ \sum_k \Omega_k(\sigma_z) a_k^\dagger a_k \\ &+ \frac{1}{2}\sum_{kj} s_{kj}\omega_k^{(e)}\lambda_k(I+\sigma_z)(a_j^\dagger + a_j) \\ &+ \frac{1}{2}\sum_k \sum_{j\neq l} s_{kj}s_{lk}\omega_k^{(e)}(I+\sigma_z)a_j^\dagger a_l \end{aligned} \tag{3.37}$$

where we define

$$\Omega_k(\sigma_z) \equiv \frac{1}{2}(\omega_k^{(g)} + s_{kk}^2\omega_k^{(e)})I + \frac{1}{2}(\omega_k^{(g)} - s_{kk}^2\omega_k^{(e)})\sigma_z. \tag{3.38}$$

In cases where Duchinsky rotations of the normal modes can be neglected ($s_{ij} = \delta_{ij}$), this expression can be further reduced to

$$\begin{aligned} H = H'_s &+ \sum_k \Omega'_k(\sigma_z) a_k^\dagger a_k \\ &+ \frac{1}{2}\sum_k \omega_k^{(e)}\lambda_k(I+\sigma_z)(a_k^\dagger + a_k) \end{aligned} \tag{3.39}$$

with the simplification

$$\Omega'_k(\sigma_z) = \frac{1}{2}(\omega_k^{(g)} + \omega_k^{(e)})I + \frac{1}{2}(\omega_k^{(g)} - \omega_k^{(e)})\sigma_z. \tag{3.40}$$

Generality is the enemy of all art.

Konstantin Stanislavski

4

Exploiting Locality in Quantum Computation for Quantum Chemistry*

Abstract

Accurate prediction of chemical and material properties from first principles quantum chemistry is a challenging task on traditional computers. Recent developments in quantum computation offer a route towards highly accurate solutions with polynomial

*Reprinted (adapted) with permission from **Jarrod R. McClean**, Ryan Babbush, Peter J. Love, and Alán Aspuru-Guzik. Exploiting locality in quantum computation for quantum chemistry. *The Journal of Physical Chemistry Letters*, 5(24):4368-4380, 2014. Copyright (2014) American Chemical Society.

cost, however this solution still carries a large overhead. In this perspective, we aim to bring together known results about the locality of physical interactions from quantum chemistry with ideas from quantum computation. We show that the utilization of spatial locality combined with the Bravyi-Kitaev transformation offers an improvement in the scaling of known quantum algorithms for quantum chemistry and provide numerical examples to help illustrate this point. We combine these developments to improve the outlook for the future of quantum chemistry on quantum computers.

4.1 Introduction

Within chemistry, the Schrödinger equation encodes all information required to predict chemical properties ranging from reactivity in catalysis to light absorption in photovoltaics. Unfortunately the exact solution of the Schrödinger equation is thought to require exponential resources on a classical computer, due to the exponential growth of the dimensionality of the Hilbert space as a function of molecular size. This makes exact methods intractable for more than a few atoms [229].

Richard Feynman first suggested that this scaling problem might be overcome if a more natural approach was taken [70]. Specifically, instead of painstakingly encoding quantum information into a classical computer, one may be able to use a quantum system to naturally represent another quantum system and bypass the seemingly insurmountable storage requirements. This idea eventually developed into the field of quantum computation, which is now believed to hold promise for the solution of problems ranging from factoring numbers [214] to image recognition [12, 182] and protein folding [12, 195].

Initial studies by Aspuru-Guzik et. al. showed that these approaches might be par-

ticularly promising for quantum chemistry [6]. There have been many developments both in theory [111, 209, 261] and experimental realization [8, 136, 196, 242] of quantum chemistry on quantum computers. The original gate construction for quantum chemistry introduced by Whitfield et al. [251] was recently challenged as too expensive by Wecker et al. [244]. The pessimistic assessment was due mostly to the extrapolation of the Trotter error for artificial rather than realistic molecular systems, as was analyzed in detail in a followup study by many of the same authors [198]. They subsequently improved the scaling by means of several circuit enhancements [99]. The analysis of the Trotter error on realistic molecules in combination with their improvements led to a recent study where an estimate of the calculation time of Fe_2S_2 was reduced by orders of magnitude [198]. In this paper, we further reduce the scaling by exploiting the locality of physical interactions with local basis sets as has been done routinely now in quantum chemistry for two decades [41, 85]. These improvements in combination with others make quantum chemistry on a quantum computer a very attractive application for early quantum devices. We describe the scaling under two prominent measurement strategies, quantum phase estimation and Hamiltonian averaging, which is a simple subroutine of the recently introduced Variational Quantum Eigensolver approach [196].

Additionally, recent progress in accurate and scalable solutions of the Schrödinger equation on classical computers has also been significant [5, 30, 41, 85, 102, 212]. Some of these results have already appeared in the quantum computation literature in the context of in depth studies of state preparation [236, 240]. A general review of quantum simulation [38, 80] and one on quantum computation for chemistry [117] cover these topics in more depth. A collection covering several aspects of quantum information and chemistry recently appeared [114]. However many developments that

utilize fundamental physical properties of the systems being studied to enable scalability have not yet been exploited.

In this study, we hope to bring to light results from quantum chemistry as well as their scalable implementation on quantum computers. We begin by reviewing the standard electronic structure problem. Results based on the locality of physical interactions from linear scaling methods in quantum chemistry are then introduced with numerical studies to provide quantification of these effects. A discussion of the resulting impact on the most common quantum algorithms for quantum chemistry follows. We also investigate instances where a perfect oracle is not available to provide input states, demonstrating the need for advances in state preparation technology. Finally, we conclude with an outlook for the future of quantum chemistry on quantum computers.

4.2 Electronic structure problem

To frame the problem and set the notation, we first briefly introduce the electronic structure problem of quantum chemistry [102]. Given a set of nuclei with associated charges $\{Z_i\}$ and a total charge (determining the number of electrons), the physical states of the system can be completely characterized by the eigenstates $\{|\Psi_i\rangle\}$ and corresponding eigenvalues (energies) $\{E_i\}$ of the Hamiltonian H

$$\begin{aligned} H = &- \sum_i \frac{\nabla^2_{R_i}}{2M_i} - \sum_i \frac{\nabla^2_{r_i}}{2} - \sum_{i,j} \frac{Z_i}{|R_i - r_j|} \\ &+ \sum_{j>i} \frac{Z_i Z_j}{|R_i - R_j|} + \sum_{j>i} \frac{1}{|r_i - r_j|} \end{aligned} \tag{4.1}$$

where we have used atomic units, $\{R_i\}$ denote nuclear coordinates, $\{r_i\}$ electronic coordinates, $\{Z_i\}$ nuclear charges, and $\{M_i\}$ nuclear masses. Owing to the large difference in masses between the electrons and nuclei, typically the Born-Oppenheimer approximation is used to mitigate computational cost and the nuclei are treated as stationary, classical point charges with fixed positions $\{R_i\}$. Within this framework, the parametric dependence of the eigenvalues on $\{R_i\}$, denoted by $\{E(\{R_i\})_j\}$ determines almost all chemical properties, such as bond strengths, reactivity, vibrational frequencies, etc. Work has been done in the determination of these physical properties directly on a quantum computer [115].

Due to the large energy gaps between electronic levels with respect to the thermal energy scale $k_B T$, it typically suffices to study a small subset of the eigenstates corresponding to the lowest energies. Moreover, for this reason, in many molecules the lowest energy eigenstate $|\Psi_0\rangle$, or ground state, is of primary importance, and for that reason it is the focus of many methods, including those discussed here.

4.2.1 Second quantized Hamiltonian

Direct computation in a positional basis accounting for anti-symmetry in the wavefunction while using the Hamiltonian described is referred to as a first quantization approach and has been explored in the context of quantum computation [116, 243, 246]. The first quantized approach has also been realized in experiment [151]. One may also perform first quantized calculations in a basis of Slater determinants. This was introduced as a representation of the electronic wavefunction by qubits in [6] (the compact mapping) and the efficiency of time evolution in this basis was recently shown [230, 237]. The second quantized approach places the antisymmetry requirements on the operators. After choosing some orthogonal spin-orbital basis $\{\varphi_i\}$ with a

number of terms M, the second quantized Hamiltonian may be written as [102]

$$\hat{H} = \sum_{pq} h_{pq} a_p^\dagger a_q + \frac{1}{2} \sum_{pqrs} h_{pqrs} a_p^\dagger a_q^\dagger a_r a_s \tag{4.2}$$

with coefficients determined by

$$h_{pq} = \int d\sigma \; \varphi_p^*(\sigma) \left(-\frac{\nabla_r^2}{2} - \sum_i \frac{Z_i}{|R_i - r|} \right) \varphi_q(\sigma) \tag{4.3}$$

$$h_{pqrs} = \int d\sigma_1 \; d\sigma_2 \; \frac{\varphi_p^*(\sigma_1)\varphi_q^*(\sigma_2)\varphi_s(\sigma_1)\varphi_r(\sigma_2)}{|r_1 - r_2|} \tag{4.4}$$

where σ_i now contains the spatial and spin components of the electron, $\sigma_i = (r_i, s_i)$. The operators $a_p^\dagger$ and a_r obey the fermionic anti-commutation relations

$$\{a_p^\dagger, a_r\} = \delta_{p,r} \tag{4.5}$$

$$\{a_p^\dagger, a_r^\dagger\} = \{a_p, a_r\} = 0 \tag{4.6}$$

For clarity, we note that the basis functions used in quantum chemistry (such as atom-centered Gaussians) are frequently parameterized on the nuclear coordinates $\{R_i\}$, which can result in a dependence on the nuclear positions of the electronic integral terms $\{h_{pqrs}\}$. For notational simplicity the dependence of the integrals on the nuclear positions in this work will remain implied.

4.2.2 Spatial locality

It is clear by inspection that the maximum number of terms in the second-quantized Hamiltonian scales as $O(M^4)$. M can be quite large to reach chemical accuracy for systems of interest, and the number of terms present in the Hamiltonian is a domi-

nant cost factor for almost all quantum computation algorithms for chemistry. However, due to the locality of physical interactions, one might imagine that many of the terms in the Hamiltonian are negligible relative to some finite precision ϵ. While this depends on the basis, it is this observation that forms the foundation for the linear-scaling methods of electronic structure such as linear scaling density functional theory or quantum Monte Carlo [5, 7, 85, 187, 247, 253, 265]. That is, in a local basis, the number of non-negligible terms scales more like $O(M^2)$, and advanced techniques such as fast multipole methods techniques can evaluate their contribution in $O(M)$ time.

These scaling properties are common knowledge within the domain of traditional quantum chemistry, however they have not yet been exploited within the context of quantum computation. They are clearly vitally important for the correct estimate of the asymptotic scaling of any method [6, 111, 244, 251]. For that reason, we review the origin of that scaling here for the most common and readily available local basis, the Gaussian atomic orbital basis. We follow loosely the explanation presented by Helgaker, Jørgensen, and Olsen [102], and refer readers to this text for additional detail on the evaluation of molecular integrals in local basis sets. The two elements we will consider here are the cutoffs due to exponentially vanishing overlaps between Gaussians basis functions and a bound on the value of the largest integral.

By far the most common basis used in electronic structure calculations is a set of atom-centered Gaussian (either Cartesian or "Pure" spherical) functions. While the precise result can depend on the angular momentum associated with the basis function, for simplicity, consider only Gaussian S functions, which is defined by

$$|G_a\rangle = \exp\left(-ar_A^2\right) \tag{4.7}$$

where r_A is the vector from a point A which defines the center of the Gaussian. One property of Gaussian functions that turns out to be useful in the evaluation of molecular integrals is the Gaussian product rule. This rule states simply that the product of two spherical Gaussian functions may be written in terms of a single spherical Gaussian function on the line segment connecting the two centers. Consider two spherical Gaussian functions, $|G_a\rangle$ and $|G_b\rangle$ separated along the x-axis.

$$\exp\left(-ax_A^2\right)\exp\left(-bx_B^2\right) = K_{ab}^x \exp\left(-px_p^2\right) \tag{4.8}$$

where K_{ab}^x is now a constant pre-exponential factor

$$K_{ab}^x = \exp\left(-\mu X_{AB}^2\right) \tag{4.9}$$

and the total exponent p, the reduced exponent μ, and the Gaussian separation X_{AB} are given by

$$p = a + b \tag{4.10}$$

$$\mu = \frac{ab}{a+b} \tag{4.11}$$

$$X_{AB} = A_x - B_x \tag{4.12}$$

That is, the product of two spherical Gaussians is a third Gaussian centered between the original two that decays faster than the original two functions, as given by the total exponent p. The overlap integral of two spherical Gaussian S functions may be obtained through application of the Gaussian product rule after factorizing into the

three Cartesian dimensions, followed by Gaussian integration and is given by

$$S_{ab} = \langle G_a | G_b \rangle = \left(\frac{\pi}{a+b} \right)^{3/2} \exp\left(-\frac{ab}{a+b} R_{AB}^2 \right) \tag{4.13}$$

where R_{AB} is the distance between the Gaussian centers A and B. Clearly this integral decays exponentially with the square of the distance between centers, and one may determine a distance d_s such that beyond that distance, the integrals will be smaller than 10^{-k} in magnitude.

$$d_s = \sqrt{a_{\min}^{-1} \log \left[\left(\frac{\pi}{2a_{\min}} \right)^3 10^{2k} \right]} \tag{4.14}$$

where $a_{\min}$ is the minimal Gaussian exponent a (most diffuse function) in the set of Gaussian basis functions $\{|G_a\rangle\}$. While the exact decay parameters will depend on the basis set, it is generally true from this line of reasoning that there is a characteristic distance, beyond which all overlap integrals are negligible. This means that the number of interactions per basis function becomes fixed, resulting in a linear number of significant overlap integrals. As kinetic energy integrals are just fixed linear combinations of overlap integrals of higher angular momentum, the same argument holds for them as well.

For S orbitals, the two-electron Coulomb integral may be written as

$$h_{acbd} = \frac{S_{ab} S_{cd}}{R_{PQ}} \operatorname{erf}(\sqrt{\alpha} R_{PQ}) \tag{4.15}$$

where erf is the error function, P and Q are Gaussian centers formed through application of the Gaussian product rule to $|G_a\rangle |G_b\rangle$ and $|G_c\rangle |G_d\rangle$ respectively. R_{PQ} is the distance between the two Gaussian centers P and Q and α is the reduced exponent

derived from P and Q. For clarity, this may be bounded by the simpler expression

$$h_{acbd} \leq \min\left(\frac{4\alpha}{\pi} S_{ab} S_{cd}, \frac{S_{ab} S_{cd}}{R_{PQ}}\right) \tag{4.16}$$

The first of these two expressions in the min function comes from the short range bound and the latter from the long range bound of the error function. These bounds show that the integrals are determined by products of overlap terms, such that in the regime where overlap integrals scale linearly, we expect $O(M^2)$ significant two-electron terms. Moreover, as seen in the long range bound of the two-electron integral, there is some further asymptotic distance beyond which these interactions may be completely neglected, yielding an effectively linear scaling number of significant integrals. This limit can be quite large however, thus practically one expects to observe a quadratic scaling in the number of two-electron integrals (TEI).

Additionally, we note from the form of the integrals, that the maximal values the two-electron integrals will attain are determined by the basis set parameters, such as the width of the Gaussian basis functions or their angular momentum. The implication of this, is that the maximal integral magnitude for the four index two-electron integrals, $|h^{\mathrm{TEI}}_{\mathrm{max}}|$ will be independent of the molecular size for standard atom centered Gaussian basis sets, and may be treated as a constant for scaling analysis that examine cost as a function of physical system size with fixed chemical composition. The overlap and kinetic energy integrals will similarly have a maximum independent of molecular size past a very small length scale. However, the nuclear attraction integrals must also be considered.

While not typically considered a primary source of difficulty due to the relative ease of evaluation with respect to two-electron integrals, we separate the nuclear at-

traction integrals here due to the fact that the maximal norm of the elements may change as well. The nuclear attraction matrix element between S functions may be written as

$$h_{ab}^{\text{nuc}} = -\sum_i \frac{Z_i S_{ab}}{R_{Pi}} \operatorname{erf}(\sqrt{p} R_{Pi}) \tag{4.17}$$

where Z_i is the nuclear charge and R_{Pi} refers to the distance between the Gaussian center P with total exponent p formed from the product $|G_a\rangle |G_b\rangle$ to the position of the i'th nuclei. Following from the logic above, from the exponentially vanishing overlap S_{ab}, at some distance, we expect only a linear number of these integrals to be significant. However, each of the integrals considers the sum over all nuclei, which can be related linearly to the number of basis functions in atom centered Gaussian basis sets. Thus the maximal one-electron integral is not a constant, but rather can be expected to scale with the Coulomb sum over distant nuclear charges. A conservative bound can be placed on such a maximal element as follows.

The temperature and pressure a molecule reside in will typically determine the minimal allowed separation of two distinct nuclei, and will thus define a maximum nuclear density $\rho_{\max}$. Denote the maximum nuclear charge in the systems under consideration as $Z_{\max}$. The maximal density and the number of nuclei will also define a minimal radius that a sphere of charge may occupy $r_{\max}$,

$$r_{\max}^3 = \frac{3 Z_{\max} N_{\text{nuc}}}{4\pi \rho_{\max}} \tag{4.18}$$

where N_{nuc} is the number of nuclei in the system. Modeling the charge as spread uniformly within this minimal volume and using the maximum of the error function to

find a bound on the maximum for the nuclear attraction matrix element, we find

$$\begin{aligned} |h_{ab}^{\text{nuc}}| &< 4\pi\rho_{\max}S_{ab}\left|\int_0^{r_{\max}} r^2 dr \frac{1}{r}\right| \\ &= 2\pi\rho_{\max}S_{ab}r_{\max}^2 \\ &= \beta_{ab}N_{\text{nuc}}^{2/3} \end{aligned} \tag{4.19}$$

where β_{ab} is now a system size independent quantity determined only by basis set parameters at nuclei a and b, and the size dependence is bounded as $O(N_{\text{nuc}}^{2/3})$. Atom centered Gaussian basis sets will have a number of a basis functions which is a linear multiple of the number of nuclei, and as such we may now bound the maximal one-electron integral (OEI) element as

$$|h_{\max}^{\text{OEI}}| < \beta_{\max}^{\text{OEI}}M^{2/3} \tag{4.20}$$

4.2.3 Effect of truncation

The above analysis demonstrates that given some integral magnitude threshold, δ, there exists a characteristic distance d between atomic centers, beyond which integrals may be neglected. If one is interested in a total precision ϵ in the energy E_i, it is important to know how choosing δ will impact the solution, and what choice of δ allows one to retain a precision ϵ.

By specification, the discarded integrals are small with respect to the rest of the Hamiltonian (sometimes as much as 10 orders of magnitude smaller in standard calculations). As such, one expects a perturbation analysis to be accurate. Consider the new, truncated Hamiltonian $H_t = H + V$, where V is the negation of the sum of all removed terms, each of which have magnitude less than δ.

Assuming a non-degenerate spectrum for H, from perturbation theory we expect the leading order change in eigenvalue E_i to be given by

$$\Delta E_i = \langle \Psi_i | \, V \, | \Psi_i \rangle \tag{4.21}$$

if the number of terms removed from the sum is given by N_r, a worst case bound on the magnitude of this deviation follows from the spectrum of the creation and annihilation operators and is given by

$$|\Delta E_i| \leq \sum_{\{h_i : |h_i| < \delta\}} |h_i| \leq N_r \delta \tag{4.22}$$

where $\{h_i : |h_i| < \delta\}$ is simply the set of Hamiltonian elements with norm less than δ and the first inequality follows directly from the triangle inequality. We emphasize that this is a worst case bound, and generically one expects at least some cancellation between terms, such as kinetic and potential terms, when the Hamiltonian is considered as a whole. Some numerical studies of these cancellation effects have been performed [198], but additional studies are required. Regardless, under this maximal error assumption, by choosing a value

$$\delta \leq \frac{\epsilon}{N_r} \tag{4.23}$$

one retains an accuracy ϵ in the final answer with respect to the exact answer when measuring the eigenvalue of the truncated Hamiltonian H_t. Alternative, one may use the tighter bound based on the triangle inequality and remove the maximum number of elements such that the total magnitude of removed terms is less than ϵ. From the looser but simpler bound, we see a reduction of scaling from M^4 to M^2 would require

removal of the order of M^4 terms from the Hamiltonian, this constraint on δ can be rewritten in terms of M as

$$\delta \leq \frac{\epsilon}{M^4} \tag{4.24}$$

While the perturbation of the eigenvalue will have a direct influence on energy projective measurement methods such as quantum phase estimation, other methods evaluate the energy by averaging. In this case, we do not need to appeal to perturbation theory, and the δ required to achieve a desired ϵ can be found directly.

$$\langle H_t \rangle = \langle \Psi_i | H_t | \Psi_i \rangle \tag{4.25}$$

$$= E_i + \langle \Psi_i | V | \Psi_i \rangle \tag{4.26}$$

We find that under our assumption of worst case error for averaging, the result is identical to that of the first order perturbation of the eigenvalue E_i,

$$|\Delta \langle H_t \rangle| \leq \sum_{\{h_i : |h_i| < \delta\}} |h_i| \leq N_r \delta \tag{4.27}$$

In summary, we find that for both the consideration of the ground state eigenvalue and the average energy of the ground state eigenvector, there is a simple formula for the value of δ, which scales polynomially in the system size, below which one may safely truncate to be guaranteed an accuracy ϵ in the final answer. Moreover it suggests a simple strategy that one may utilize to achieve the desired accuracy. That is, sort the integrals in order of magnitude, and remove the maximum number of integrals such that the total magnitude of removed integrals is less than ϵ.

On the subject of general truncation, we note that while there may exist Hamiltonians with the same structure as the second quantized electronic structure Hamil-

tonian that have the property that removal of small elements will cause a drastic shift in the character of the ground state, this has not been seen for physical systems in quantum chemistry. Moreover, from the perturbation theory analysis given, such Hamiltonians would likely need to exhibit degenerate ground electronic states, which are not common in physical systems. In practice it is observed that removing elements on the order of $\delta = 10^{-10}$ and smaller is more than sufficient to retain both qualitative and quantitative accuracy in systems of many atoms [5, 7, 102, 253]. Moreover, the convergence with respect to this value may be tested easily for any systems under consideration.

4.2.4 Onset of favorable scaling

While the above analysis shows that locality of interactions in local basis sets provides a promise that beyond a certain length scale, the number of non-negligible integrals will scale quadratically in the number of basis functions, it does not provide good intuition for the size of that length scale in physical systems of interest. Here we provide numerical examples for chemical systems in basis sets used so far in quantum computation for quantum chemistry. The precise distance at which locality starts to reduce the number of significant integrals depends, of course, on the physical system and the basis set used. In particular, larger, more diffuse basis sets are known to exhibit these effects at comparatively larger length scales than minimal, compact basis sets. However the general scaling arguments given above hold for all systems of sufficient size.

An additional consideration which must be made for quantum computation, is that as of yet, no general technology has been developed for direct simulation in non-orthogonal basis sets. This prohibits direct simulation in the bare atomic orbital ba-

sis, however the use of Löwdin symmetric orthogonalization yields the orthogonal basis set closest to the original atomic orbital basis set in an l^2 sense [149, 163]. We find that this is sufficient for the systems we consider, but note that there have been a number of advances in orthogonal basis sets that are local in both the occupied and virtual spaces and may find utility in quantum computation [273]. Moreover, there has been recent work in the use of multiresolution wavelet basis sets that have natural sparsity and orthogonality while providing provable error bounds on the choice of basis [97]. Such a basis also allows one to avoid costly integral transformations related to orthogonality, which are known to scale as $O(M^5)$ when performed exactly. Further research is needed to explore the implications for quantum computation with these basis sets, and we focus here on the more common atom-centered Gaussian basis sets.

As a prototype system, we consider chains of hydrogen atoms separated by 1 Bohr (a_0) in the STO-3G basis set, an artificial system that can exhibit a transition to a strongly correlated wavefunction [91]. We count the total number of significant integrals for values of δ given by 10^{-15} and 10^{-7} for the symmetrically orthogonalized atomic orbital (OAO) basis and the canonical Hartree-Fock molecular orbital (MO) basis. The results are displayed in Fig. 4.1 and demonstrate that with a cutoff of $\delta = 10^{-7}$ the localized character of the OAO's allows for a savings of on the order of 6×10^6 integrals with respect to the more delocalized canonical molecular orbitals. The s in the labeling of the orbital bases simply differentiates between two possible cutoffs. These dramatic differences begin to present with atomic chains as small as 10 Å in length in this system with this basis set.

As an additional example, we consider linear alkane chains of increasing length. The results are displayed in Fig. 4.2 and again display the dramatic advantages of

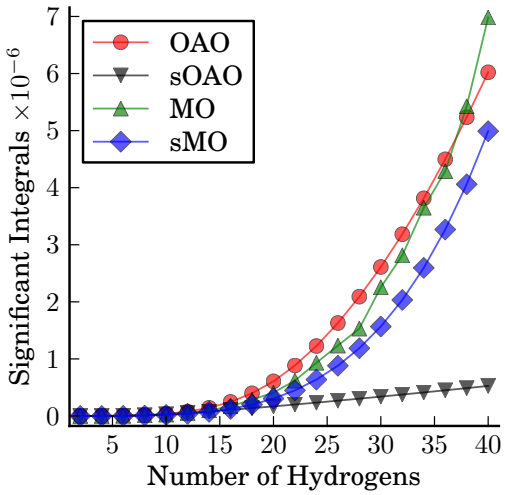

Figure 4.1: The number of significant(magnitude $> 10^{-15}$) spin-orbital integrals in the STO-3G basis set as a function of the number of hydrogens in a linear hydrogen chain with a separation of 1 a_0 for the Hartree-Fock canonical molecular orbital basis(MO) and the symmetrically orthogonalized atomic orbital basis (OAO). The sMO and sOAO, shows the same quantity with a sharper cutoff (10^{-7}) and demonstrates the advantage to localized atomic basis functions at length scales as small as 10 Å.

preserving locality in the basis set. By the point one reaches 10 carbon atoms, a savings of almost 10^8 integrals can be achieved at a truncation level of 10^{-7}.

Although localized basis sets provide a definitive scaling advantage in the medium-large size limit for molecules, one often finds that in the small molecule limit canonical molecular orbitals, the orbitals from the solution of the Hartree-Fock equations under the canonical condition, provide a more sparse representation. This is observed in the hydrogen and alkanes chains studied here for the smallest molecule sizes, and the transition for this behavior will generally be basis set dependent. For example in the alkane chains smaller than C_4H_10 studied here, such as C_3H_8, the number of significant integrals in the MO basis at a threshold of 10^{-7} is roughly 80% of that in the atomic orbital basis. The reason is that at smaller length scales, the "delocalized"

canonical molecule orbitals have similar size to the more localized atomic orbitals, but with the additional constraint of the canonical condition, a sufficient but not necessary condition for the solution of the Hartree-Fock equations that demands the Fock matrix be diagonal (as opposed to the looser variational condition of block-diagonal between the occupied and virtual spaces). A side effect of the canonical condition is that in the canonical molecular orbital basis many of the h_{pqrs} terms for distinct indices are reduced in magnitude. However, there are not enough degrees of freedom present in the orbital rotations for this effect to persist to larger length scales, and as a result local basis sets eventually become more advantageous. Moreover, it is known that at larger length scales, the canonical conditions tend to favor maximally delocalized orbitals, which can reduce the advantages of locality. These effects have been studied in some detail in the context of better orbital localizations by relaxing the canonical condition in Hartree-Fock and the so-called Least-Change Hartree-Fock method coupled with fourth-moment minimization [273].

4.3 Quantum energy estimation

Almost all algorithms designed for the study of quantum chemistry eigenstates on a quantum computer can be separated into two distinct parts: 1. state preparation and 2. energy estimation. For the purposes of analysis, it is helpful to treat the two issues separately, and in this section we make the standard assumption in doing so, that an oracle capable of producing good approximations to the desired eigenstates $|\Psi_i\rangle$ at unit cost is available. Under this assumption, energy estimation for a fixed desired precision ϵ is known to scale polynomially in the size of the system for quantum chemistry, however the exact scaling costs and tradeoffs depend on the details of the

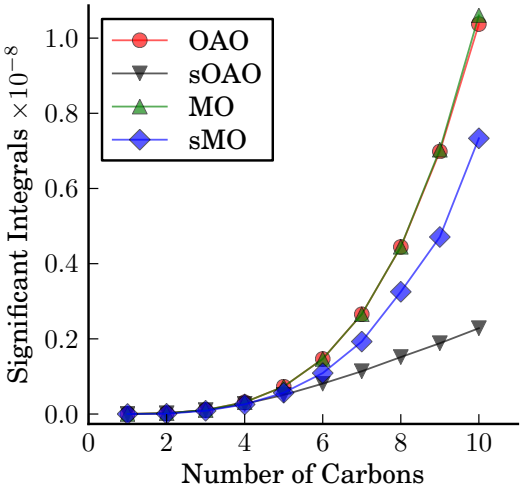

Figure 4.2: The number of significant(magnitude $> 10^{-15}$) spin-orbital integrals in the STO-3G basis set as a function of the number of carbons in a linear alkane chain for the Hartree-Fock canonical molecular orbital basis (MO) and the symmetrically orthogonalized atomic orbital basis (OAO). The sMO and sOAO shows the same quantity with a sharper cutoff (10^{-7}) and demonstrates the dramatic advantage to localized atomic basis even at this small atomic size.

method used. Here we compare the costs and benefits of two prominent methods of energy estimation used in quantum computation for chemistry: quantum phase estimation and Hamiltonian averaging.

4.3.1 Quantum phase estimation

The first method used for the energy estimation of quantum chemical states on a quantum computer was quantum phase estimation [2, 6, 122]. The method works by evolving the give quantum eigenstate $|\Psi_i\rangle$ forward under the system Hamiltonian H for a time T, and reading out the accumulated phase, which can be easily mapped to the associated eigenenergy E_i. While the basic algorithm and its variations can have many different components, the cost is universally dominated by the coherent evolution of the system.

To evolve the system under the Hamiltonian, one must find a scalable way to implement the unitary operator $U = e^{-iHT}$. The standard procedure for accomplishing this task is the use of Suzuki-Trotter splitting [222, 233], which approximates the unitary operator(at first order) as

$$\begin{aligned} U = e^{-iHT} &= \left(e^{-iH(T/m)}\right)^m \\ &= \left(e^{-i(\sum_i H_i)\Delta t}\right)^m \approx \left(\prod_i e^{-iH_i\Delta t}\right)^m \end{aligned} \tag{4.28}$$

where $\Delta t = T/m$ and H_i is a single term from the Bravyi-Kitaev transformed system Hamiltonian. Higher order Suzuki-Trotter operator splittings and their benefits have been studied in the context of quantum simulation by Berry et al. [21] and in an early arXiv version of Ref. [250], namely [250], but we largely focus on the first order formula in this work. If each of the simpler unitary operators $e^{-iH_i\Delta t}$ has a known gate decomposition, the total time evolution can be performed by chaining these sequences together.

The use of the Suzuki-Trotter splitting can be thought of as an evolution under an approximate Hamiltonian $\tilde{H}$, given by $e^{-i\tilde{H}T}$, whose eigenspectrum deviates from the original Hamiltonian by a factor depending on time-step Δt. The precise dependence of this bias depends on the order of the Suzuki-Trotter expansion used. The total resolution, ϵ, in the energies of the approximate Hamiltonian $\tilde{H}$ is determined by the total evolution time T. Thus to achieve an accuracy of ϵ in the final energy, one must utilize a time step Δt small enough that the total bias is less than ϵ and a total run time T such that the resolution is better than ϵ. If the number of gates required to implement a single timestep Δt is given by N_g, then the dominant cost of simulation

(all of which must be done coherently) is given by

$$N_c = N_g \left\lceil \frac{T}{\Delta t} \right\rceil \tag{4.29}$$

The total evolution time T required to extract an eigenvalue to chemical precision $\epsilon_{\text{chem}} = 10^{-3}$ is typically set at the Fourier limit independent of molecular size and thus can be considered a constant for scaling analysis. We then focus on the number of gates per Suzuki-Trotter time step, N_g, and the time step Δt required to achieve the desired precision.

In a first order Suzuki-Trotter splitting, the number of gates per Trotter time step is given by the number of terms in the Hamiltonian multiplied by the number of gates required to implement a single elementary term for the form $e^{-iH_i\Delta t}$. The gates per elementary term can vary based on the particular integral, however for simplicity in developing bounds we consider this as constant here. The number of terms, is known from previous analysis in this work to scale as $O(M^2)$ or in the truly macroscopic limit $O(M)$. The number of gates required to implement a single elementary term depends on the transformation used from fermionic to qubit operators. The Jordan-Wigner transformation [112] results in non-local terms that carry with them an overhead that scales as the number of qubits, which in this case will be $O(M)$. Although there have been developments in methods to use teleportation to perform these non-local operations in parallel [111] and by improving the efficiency of the circuits computing the phases in the Jordan-Wigner transformation [99], these issues can also be alleviated by choosing the Brayvi-Kitaev transformation that carries an overhead only logarithmic in the number of qubits, $O(\log M)$ [35, 209]. As a result, one expects the number of gates per Suzuki-Trotter time step N_g to scale as $O(M^2 \log M)$ or in

a truly macroscopic limit $O(M \log M)$.

To complete the cost estimate with fixed total time T, one must determine how the required time step Δt scales with the size of the system. As mentioned above, the use of the Suzuki-Trotter decomposition for the time evolution of H is equivalent to evolution under an effective Hamiltonian $\tilde{H} = H + V$, where the size of the perturbation is determined by the order of the Suzuki-Trotter formula used and the size of the timestep. Once the order of the Suzuki-Trotter expansion to be used has been determined, the requirement on the timestep is such that the effect of V on the eigenvalue of interest is less than the desired accuracy in the final answer ϵ.

This has been explored previously [99, 198], but we now examine this scaling in our context. To find V, one may expand the k'th order Suzuki-Trotter expansion of the evolution of $\tilde{H}$ into a power series as well as the power series of the evolution operator $\exp\left[-i\left(H+V\right)\Delta t\right]$, and find the leading order term V. As a first result, we demonstrate that for a k'th order propagator, the leading perturbation on the ground state eigenvalue for a non-degenerate system is $O(\Delta t)^{k+1}$.

Recall the power series expansion for the propagator

$$\exp\left[-i\left(H+V\right)\Delta t\right] = \sum_{j=0}^{\infty} \frac{(-i)^j}{j!} \left(H+V\right)^j \left(\Delta t\right)^j \tag{4.30}$$

The definition of a k'th order propagator, is one is that correct through order k in the power series expansion. As such, when this power series is expanded, V must make no contribution in the terms until $O((\Delta t)^{k+1})$. For this to be possible, it's clear that V must depend on Δt. In order for it to vanish for the first k terms, V must be proportional to $(\Delta t)^k$. Moreover, due to the alternation of terms between imaginary and real at each order in the power series with the first term being imaginary, the first possi-

ble contribution is order $(\Delta t)^k$ and imaginary. As is common in quantum chemistry, we assume a non-degenerate and real ground state, and thus the contribution to the ground state eigenvalue is well approximated by first order perturbation theory as

$$E^{(1)} = \langle \Psi_g | \, V \, | \Psi_g \rangle \tag{4.31}$$

however, as V is imaginary Hermitian and the ground state is known to be real in quantum chemistry, this expectation value must vanish. Thus the leading order perturbation to the ground state eigenvalue is at worst the real term depending on $(\Delta t)^{k+1}$.

To get a more precise representation of V for a concrete example, we now consider the first order ($k = 1$) Suzuki-Trotter expansion. As expected, the leading order imaginary error term is found to be

$$V^{(0)} = \frac{\Delta t}{2} \sum_{j<k} i \left[H_j, H_k \right] \tag{4.32}$$

whose contribution must vanish due to it being an imaginary Hermitian term. Thus we look to the leading contributing error depending on $(\Delta t)^2$, which has been obtained previously[198] from a Baker-Campbell-Hausdorff(BCH) expansion to read

$$V^{(1)} = \frac{(\Delta t)^2}{12} \sum_{i \leq j} \sum_{j} \sum_{k<j} \left[H_i \left(1 - \frac{\delta_{ij}}{2} \right), \left[H_j, H_k \right] \right] \tag{4.33}$$

Thus the leading order perturbation is given by third powers of the Hamiltonian operators. To proceed, we count the number of one- and two-electron integrals separately as $N_{\text{int}}^{\text{OEI}}$ and $N_{\text{int}}^{\text{TEI}}$ respectively. Their maximal norm elements are similarly denoted by $h_{\text{max}}^{\text{OEI}}$ and $h_{\text{max}}^{\text{TEI}}$. From this, we can draw a worst case error bound on the perturba-

tion of the eigenvalue given by

$$\begin{aligned} E^{(1)} &\leq \frac{(\Delta t)^2}{12} \sum_{i \leq j} \sum_{j} \sum_{k<j} \left| H_i \left(1 - \frac{\delta_{ij}}{2}\right), [H_j, H_k] \right| \\ &\leq \left(|h_{\max}^{\mathrm{OEI}}| N_{\mathrm{int}}^{OEI} + |h_{\max}^{\mathrm{TEI}}| N_{\mathrm{int}}^{\mathrm{TEI}}\right)^3 (\Delta t)^2 \\ &\leq \left(|\beta_{\max}^{\mathrm{OEI}}| M^{2/3} N_{\mathrm{int}}^{\mathrm{OEI}} + |h_{\max}^{\mathrm{TEI}}| N_{\mathrm{int}}^{\mathrm{TEI}}\right)^3 (\Delta t)^2 \end{aligned} \tag{4.34}$$

Where the first inequality follows from the triangle inequality and the second is a looser, but simpler bound, that may be used to elucidate the scaling behavior. Holding the looser bound to the desired precision in the final answer ϵ, this yields

$$\Delta t \leq \left[\frac{\epsilon}{\left(|\beta_{\max}^{\mathrm{OEI}}| M^{2/3} N_{\mathrm{int}}^{\mathrm{OEI}} + |h_{\max}^{\mathrm{TEI}}| N_{\mathrm{int}}^{\mathrm{TEI}}\right)^3} \right]^{1/2} \tag{4.35}$$

We emphasize that this is a worst case bound from first order perturbation theory, including no possible cancellation between Hamiltonian terms. Some preliminary work has been done numerically in establishing average cancellation between terms that shows these worst case bounds are too pessimistic [198]. Additionally, a rigorous bound not depending on perturbation theory has been previously derived [198, 244]. Continuing, we expect the total scaling under a first order Suzuki-Trotter expansion using a Bravyi-Kitaev encoding to be bounded by

$$\begin{aligned} N_c &= N_g \left\lceil \frac{T}{\Delta t} \right\rceil \leq \frac{N_g}{\epsilon \Delta t} \\ &\leq \frac{\left(|\beta_{\max}^{\mathrm{OEI}}| M^{2/3} N_{\mathrm{int}}^{\mathrm{OEI}} + |h_{\max}^{\mathrm{TEI}}| N_{\mathrm{int}}^{\mathrm{TEI}}\right)^{3/2} N_{\mathrm{int}} \log M}{\epsilon^{3/2}} \end{aligned} \tag{4.36}$$

and in the large size limit where the number of significant two-electron integrals in a

local basis set scales quadratically and the number of significant one-electron integrals scales linearly, this may be bounded by

$$N_c \leq \kappa \frac{\left(|\beta_{\max}^{\mathrm{OEI}}| M^{5/3} + |h_{\max}^{\mathrm{TEI}}| M^2\right)^{3/2} (M^2 + M)}{\epsilon^{3/2} (\log M)^{-1}} \tag{4.37}$$

where κ is a positive constant that will depend on the basis set and this expression scales as $O(M^5 \log M)$ in the number of spin-orbital basis functions.

4.3.2 Hamiltonian averaging

The quantum phase estimation algorithm has been central in almost all algorithms for energy estimation in quantum simulation. However, it has a significant practical drawback in that after state preparation, all the desired operations must be performed coherently. A different algorithm for energy estimation has recently been introduced [196, 261] that lifts all but an $O(1)$ coherence time requirement after state preparation, making it amenable to implementation on quantum devices in the near future. We briefly review this approach, which we will call Hamiltonian averaging, and bound its costs in applications for quantum chemistry.

As in quantum phase estimation, in Hamiltonian averaging one assumes the eigenstates $|\Psi_i\rangle$ are provided by some oracle. By use of either the Jordan-Wigner or Bravyi-Kitaev transformation, the Hamiltonian may be written as a sum of tensor products of Pauli operators. These transformations at worst conserve the number of independent terms in the Hamiltonian, thus we may assume for our worst case analysis the number of terms is fixed by N_{int} and the coefficients remain unchanged. From the provided copy of the state and transformed Hamiltonian, to obtain the energy one

simply performs the average

$$\langle \hat{H} \rangle = \sum_{i,j,k,... \in x,y,z} h_{ijk...} \langle \sigma_1^i \otimes \sigma_2^j \otimes \sigma_3^k ... \rangle \tag{4.38}$$

by independent Pauli measurements on the provided state $|\Psi_i\rangle$ weighted by the coefficients $h_{ijkl...}$, which are simply a relabeling of the previous two-electron integrals for convenience with the transformed operators. As $|\Psi_i\rangle$ is an eigenstate, this average will correspond to the desired eigenvalue E_i with some error related to sampling that we now quantify.

Consider an individual term

$$X_{ijkl...} = h_{ijkl...} \sigma_1^i \otimes \sigma_2^j \otimes \sigma_3^k ... \tag{4.39}$$

it is clear from the properties of qubit measurements, that the full range of values this quantity can take on is $[-h_{ijkl...}, h_{ijkl...}]$. As a result, we expect that the variance associated with this term can be bounded by

$$\mathrm{Var}\left[X_{ijkl...}\right] \leq |h_{ijkl...}|^2 \tag{4.40}$$

Considering a representative element, namely the maximum magnitude integral element $h_{\max}$, we can bound the variance of $\hat{H}$ as

$$\mathrm{Var}\left[\hat{H}\right] \leq N_{int}^2 |h_{\max}|^2 \tag{4.41}$$

The variance of the mean, which is the relevant term for our sampling error, comes

from the central limit theorem and is bounded by

$$\mathrm{Var}\left[\langle\hat{H}\rangle\right] \leq \frac{\mathrm{Var}\left[\hat{H}\right]}{N} \tag{4.42}$$

where N is the number of independent samples taken of $\langle\hat{H}\rangle$. Collecting these results, we find

$$\begin{aligned}\mathrm{Var}\left[\langle\hat{H}\rangle\right] &\leq \sum \frac{|h_{ijkl\ldots}|^2}{N} \\ &\leq \frac{\left(|\beta_{\max}^{\mathrm{OEI}}|M^{2/3}N_{\mathrm{int}}^{\mathrm{OEI}} + |h_{\max}^{\mathrm{TEI}}|N_{\mathrm{int}}^{\mathrm{TEI}}\right)^2}{N}\end{aligned} \tag{4.43}$$

Now setting the variance to the desired statistical accuracy ϵ^2 (which corresponds to a standard error of ϵ at a 68% confidence interval), we find the number of independent samples expected, N_s, is bounded by

$$N_s \leq \frac{\left(|\beta_{\max}^{\mathrm{OEI}}|M^{2/3}N_{\mathrm{int}}^{\mathrm{OEI}} + |h_{\max}^{\mathrm{TEI}}|N_{\mathrm{int}}^{\mathrm{TEI}}\right)^2}{\epsilon^2} \tag{4.44}$$

If a single independent sample of $\langle\hat{H}\rangle$ requires the measurement of each of the N_{int} quantities, then the bound on the total cost in the number of state preparations and measurements, N_m is

$$N_m \leq \frac{N_{int}\left(|\beta_{\max}^{\mathrm{OEI}}|M^{2/3}N_{\mathrm{int}}^{\mathrm{OEI}} + |h_{\max}^{\mathrm{TEI}}|N_{\mathrm{int}}^{\mathrm{TEI}}\right)^2}{\epsilon^2} \tag{4.45}$$

which if one considers the large size limit,such that the number of two-electron integrals scales quadratically and the number of one-electron integrals scales linearly, we

find

$$N_m \leq \kappa \frac{(M + M^2)\left(|\beta_{\max}^{\mathrm{OEI}}| M^{5/3} + |h_{\max}^{\mathrm{TEI}}| M^2\right)^2}{\epsilon^2} \tag{4.46}$$

where κ is a positive constant that depends upon the basis set. It is clear that this expression scales as $O(M^6)$ in the number of spin-orbital basis functions. We see from this, that under the same maximum error assumptions, Hamiltonian averaging scales only marginally worse in the number of integrals and precision as compared to quantum phase estimation performed with a first order Suzuki-Trotter expansion, but has a coherence time requirement of $O(1)$ after each state preparation. Note that each measurement is expected to require single qubit rotations that scale as either $O(M)$ for the Jordan-Wigner transformation or $O(\log M)$ for the Bravyi-Kitaev transformation. However, we assume that these trivial single qubit rotations can be performed in parallel independent of the size of the system without great difficulty, and we thus don't consider this in our cost estimate. This method is a suitable replacement for quantum phase estimation in situations where coherence time resources are limited and good approximations to the eigenstates are readily available. Additional studies are needed to quantify the precise performance of the two methods beyond worst case bounds.

4.4 Using imperfect oracles

A central assumption for successful quantum phase estimation and typically any energy evaluation scheme is access to some oracle capable of producing good approximations to the eigenstate of interest, where a "good" approximation is typically meant to imply an overlap that is polynomial in the size of the system. Additionally, a sup-

posed benefit of phase estimation over Hamiltonian averaging is that given such a good (but not perfect) guess, by projective measurement in the energy basis, in principle one may avoid any bias in the final energy related to the initial state. Here we examine this assumption in light of the Van-Vleck catastrophe [239], which we review below, and examine the consequences for measurements of the energy by QPE and Hamiltonian averaging.

The Van Vleck catastrophe [239] refers to an expected exponential decline in the quality of trial wavefunctions, as measured by overlap with the true wavefunction of a system, as a function of size. We study a simple case of the catastrophe here in order to frame the consequences for quantum computation. Consider a model quantum system consisting of a collection of N non-interacting two level subsystems with subsystem Hamiltonians given by H_i. These subsystems have ground and excited eigenstates $|\psi_g^i\rangle$ and $|\psi_e^i\rangle$ with eigenenergies $E_g < E_e$, such that the total Hamiltonian is given by

$$H = \sum_i H_i \tag{4.47}$$

and eigenstates of the total Hamiltonian are formed from tensor products of the eigenstates of the subsystems. As such the ground state of the full system is given by

$$|\Psi_g\rangle = \bigotimes_{i=0}^{N-1} |\psi_g^i\rangle \tag{4.48}$$

Now suppose we want to measure the ground state energy of the total system, but the oracle is only capable of producing trial states for each subsystem $|\psi_t^i\rangle$ such that

$\langle \psi_t^i | \psi_g^i \rangle = \Delta$, where $|\Delta| < 1$. The resulting trial state for the whole system is

$$|\Psi_t\rangle = \bigotimes_{i=0}^{N-1} |\psi_t^i\rangle \tag{4.49}$$

From normalization of the two level system, we may also write the trial state as

$$|\psi_t^i\rangle = \Delta \, |\psi_g^i\rangle + e^{-i\theta}\sqrt{1-\Delta^2} \, |\psi_e^i\rangle \tag{4.50}$$

where $\theta \in [0, 2\pi)$. Moreover, from knowledge of the gap, one can find the expected energy on each subsystem, which is given by

$$\langle \psi_t^i | \, H_i \, | \psi_t^i \rangle = \Delta^2 E_g + (1-\Delta^2) E_e \tag{4.51}$$

For the case of Hamiltonian averaging on the total system, the expected answer is given by

$$\begin{aligned} E &= \langle \Psi_t | \, H \, | \Psi_t \rangle \\ &= \sum_{i=0}^{N-1} \langle \psi_t^i | \, H_i \, | \psi_t^i \rangle \\ &= N(\Delta^2 E_g + (1-\Delta^2) E_e) \end{aligned} \tag{4.52}$$

which yields an energy bias from the true ground state, ϵ_b, given by

$$\begin{aligned} \epsilon_b &= N(\Delta^2 E_g + (1-\Delta^2) E_e) - N E_g \\ &= N(1-\Delta^2)(E_e - E_g) \\ &= N(1-\Delta^2)\omega \end{aligned} \tag{4.53}$$

where we denote the gap for each subsystem as $\omega = (E_e - E_g)$. As such, it is clear that the resulting bias is only linear in the size of the total system N.

Quantum phase estimation promises to remove this bias by projecting into the exact ground state. However, this occurs with a probability proportional to the square of the overlap of the input trial state with the target state. In this example, this is given by

$$|\langle \Psi_t | \Psi_g \rangle|^2 = |\Delta|^{2N} \tag{4.54}$$

which is exponentially small in the size of the system. That is, quantum phase estimation is capable of removing the bias exactly in this example non-interacting system, but at a cost which is exponential in the size of the system. The expected cost of removing some portion of the bias may be calculated by considering the distribution of states and corresponding energies.

Consider first the probability of measuring an energy with a bias of $\epsilon(M) = M(1 - \Delta^2)\omega$. For this to happen, it is clear that exactly M of the subsystems in the measured state are in the excited state. It is clear that this is true for $\begin{pmatrix} N \\ M \end{pmatrix}$ eigenstates, and the square of the overlap with such an eigenstate is $\left(\Delta^2\right)^{N-M}\left(1-\Delta^2\right)^{M}$ or

$$P(\epsilon(M)) = \begin{pmatrix} N \\ M \end{pmatrix} \left(\Delta^2\right)^{N-M}\left(1-\Delta^2\right)^{M} \tag{4.55}$$

which is clearly a binomial distribution. As a result, in the large N limit, this distri-

bution is well approximated by a Gaussian and we may write

$$P(\epsilon) \approx \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left[-\frac{1}{2}\left(\frac{M-\bar{N}}{\sigma}\right)^2\right] \tag{4.56}$$

$$\bar{N} = N(1-\Delta^2) \tag{4.57}$$

$$\sigma^2 = N\Delta^2(1-\Delta^2) \tag{4.58}$$

Bringing this together, we find that the probability of measuring a bias of less than $\epsilon(M)$ is given by

$$\begin{aligned} P(<\epsilon) &= \frac{1}{\sqrt{2\pi\sigma^2}} \int_0^M dM' \exp\left[-\frac{1}{2}\left(\frac{M'-\bar{N}}{\sigma}\right)^2\right] \\ &= \frac{1}{2}\left[\mathrm{erf}\left(\frac{M-\bar{N}}{\sqrt{2\sigma^2}}\right) + \mathrm{erf}\left(\frac{\bar{N}}{\sqrt{2\sigma^2}}\right)\right] \end{aligned} \tag{4.59}$$

where erf is again the error function.

Thus the expected cost in terms of number of repetitions of the full phase estimation procedure to remove a bias of at least $\epsilon(M)$ from this model system is

$$\begin{aligned} C(<\epsilon(M)) &= \frac{1}{P(<\epsilon(M))} \\ &= 2\left[\mathrm{erf}\left(\frac{M-\bar{N}}{\sqrt{2\sigma^2}}\right) + \mathrm{erf}\left(\frac{\bar{N}}{\sqrt{2\sigma^2}}\right)\right]^{-1} \end{aligned} \tag{4.60}$$

We plot the expected cost function for a range of oracle guess qualities Δ on a modest system of $N = 100$ in Fig 4.3. From this, we see that the amount of bias that can feasibly be removed depends strongly on the quality of the oracle guess. Generically, we see that for any fixed imperfect guess on the subsystem level($|\Delta| < 1$), there will be an exponential cost in phase estimation related to perfect removal of the bias.

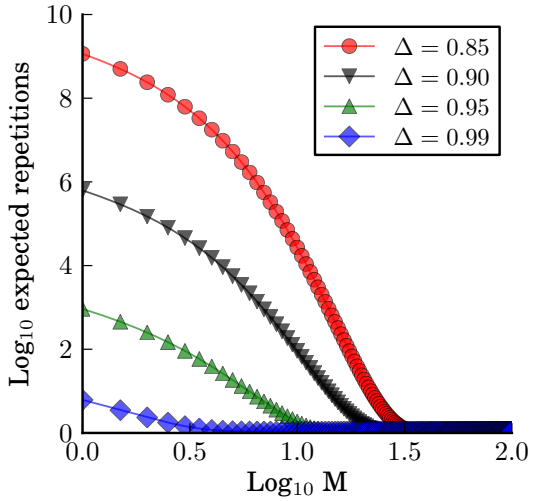

Figure 4.3: A log log plot of the expected cost in number of repetitions of measuring an energy with a bias $\epsilon(M)$ as a function of M in quantum phase estimation for different values of the oracle quality Δ. A system of $N = 100$ non-interacting subsystems is considered. A perfect, unbiased answer corresponds to $M = 0$ with expected cost $O(\Delta^{2N})$, however to aid in visualization this plot is provided only beyond $M = 1$. In general one sees that depending on the oracle quality Δ, different fractions of the bias may be removed with ease, but there is always some threshold for imperfect guesses ($|\Delta| < 1$) such that there is an exponential growth in cost.

This problem can be circumvented by improving the quality of the subsystem guesses as a function of system size. In particular, one can see that if $|\Delta|$ is improved as $(1 - 1/(2N))$ then $|\Delta|^{2N}$ is $O(1)$. However, as the subsystems in a general case could be of arbitrary size, classical determination of a subsystem state of sufficient quality may scale exponentially in the required precision and thus system size. Moreover, one would not expect the problem to be easier in general cases where interactions between subsystems are allowed. As a result, further developments in variational methods [196], quantum cooling [260], and adiabatic state preparation [6, 10, 236] will be of key importance in this area. Moreover improvements in the ansatze used to prepare the wave function such as multi-configurational self consistent field calculations(MCSCF) [236, 240] or unitary coupled cluster (UCC) [261] will be integral parts

of any practical quantum computing for quantum chemistry effort.

4.5 Adiabatic computation

A complementary solution for the problem of molecular simulation on quantum computers is that of adiabatic quantum computation. It is not known to show the same direct dependence on the overlap of the initial guess state as QPE, which may allow it to solve different problems than the quantum phase estimation or variational quantum eigensolver in practice. In [10], Babbush *et al.* show how to scalably embed the eigenspectra of molecular Hamiltonians in a programmable physical system so that the adiabatic algorithm can be applied directly. In this scheme, the molecular Hamiltonian is first written in second quantization using fermionic operators. This Hamiltonian is then mapped to a qubit Hamiltonian using the Bravyi-Kitaev transformation [35, 209]. The authors show that the more typical Jordan-Wigner transformation cannot be used to scalably reduce molecular Hamiltonians to 2-local qubit interactions as the Jordan-Wigner transformation introduces linear locality overhead which translates to an exponential requirement in the precision of the couplings when perturbative gadgets are applied. Perturbative gadgets are used to reduce the Bravyi-Kitaev transformed Hamiltonian to a 2-local programmable system with a restricted set of physical couplings. Finally, tunneling spectroscopy of a probe qubit [20] can be used to measure eigenvalues of the prepared state directly.

While the exact length of time one must adiabatically evolve is generally unknown, Babbush *et al.* argue that the excited state gap could shrink polynomially with the number of spin-orbitals when interpolating between exactly preparable noninteracting subsystems and the exact molecular Hamiltonian in which those subsystems inter-

act. This would imply that adiabatic state preparation is efficient. Their argument is based on the observation that molecular systems are typically stable in their electronic ground states and the natural processes which produce these states should be efficient to simulate with a quantum device. Subsequently, Veis and Pittner analyzed adiabatic state preparation for a set of small chemical systems and observed that for all configurations of these systems, the minimum gap occurs at the very end of the evolution when the state preparation is initialized in an eigenstate given by a CAS (complete active space) ground state [236]. The notion that the minimum gap could be bounded by the physical HOMO (highest occupied molecular orbital) - LUMO (lowest unoccupied molecular orbital) gap lends support to the hypothesis put forward by Babbush *et al.*

4.5.1 Resources for adiabatic quantum chemistry

In the adiabatic model of quantum computation, the structure of the final problem Hamiltonian (encoding the molecular eigenspectrum) determines experimental resource requirements. Since programmable many-body interactions are generally unavailable, we will assume that any experimentally viable problem Hamiltonian must be 2-local. Any 2-local Hamiltonian on n qubits can be expressed as,

$$H = \alpha \cdot \mathbf{1} + \sum_{i=1}^{n} \vec{\beta}_i \cdot \vec{\sigma}_i + \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} \vec{\gamma}_{ij} \cdot (\vec{\sigma}_i \otimes \vec{\sigma}_j) \tag{4.61}$$

where $\vec{\sigma}_i = \langle \sigma_i^x, \sigma_i^y, \sigma_i^z \rangle$ is the vector of Pauli matrices on the i^{th} qubit, $\alpha \in \mathbb{R}$ is a scalar and $\vec{\beta}_i \in \mathbb{R}^3$ and $\vec{\gamma}_{ij} \in \mathbb{R}^9$ are vectors of coefficients for each possible term.

In addition to the number of qubits, the most important resources are the number of qubit couplings and the range of field values needed to accurately implement the

Hamiltonian. Since local fields are relatively straightforward to implement, we are concerned with the number of 2-local couplings,

$$\sum_{i=1}^{n-1} \sum_{j=i+1}^{n} \mathrm{card}\left(\vec{\gamma}_{ij}\right) \tag{4.62}$$

where $\mathrm{card}\left(\vec{v}\right)$ is the number of nonzero terms in vector $\vec{v}$. Since the effective molecular electronic structure Hamiltonian is realized perturbatively, there is a tradeoff between the error in the eigenspectrum of the effective Hamiltonian, ϵ, and the strength of couplings that must be implemented experimentally. The magnitude of the perturbation is inversely related to the gadget spectral gap Δ which is directly proportional to the largest term in the Hamiltonian,

$$\max_{ij} \left\{ \left\| \vec{\gamma_{ij}}\left(\epsilon\right) \right\|_{\infty} \right\} \propto \Delta\left(\epsilon\right). \tag{4.63}$$

Thus, the smaller Δ is, the easier the Hamiltonian is to implement but the greater the error in the effective Hamiltonian. In general, there are other important resource considerations but these are typically scale invariant; for instance, the geometric locality of a graph or the set of allowed interaction terms. The Hamiltonian can be modified to fit such constraints using additional perturbative gadgets but typically at the cost of using more ancilla qubits that require greater coupling strength magnitudes.

4.5.2 Estimates of qubit and coupler scaling

The Bravyi-Kitaev transformation is crucial when embedding molecular electronic structure in 2-local spin Hamiltonians due to the fact that this approach guarantees a logarithmic upper-bound on the locality of the Hamiltonian. A loose upper-bound

(i.e. overestimation) for the number of qubits needed to gadgetize the molecular electronic Hamiltonian can be obtained by assuming that all terms have the maximum possible locality of $O(\log(M))$ where M is the number of spin-orbitals.

In general, the number of terms produced by the Bravyi-Kitaev transformation scales the same as the number of integrals in the electronic structure problem, $O\left(M^4\right)$; however, as pointed out in an earlier section, this bound can be reduced to $O\left(M^2\right)$ if a local basis is used and small integrals are truncated. Using the "bit-flip" gadgets of [113, 119] to reduce M^2 terms of locality $\log(M)$, we would need $M^2 \log(M)$ ancillae. Since the number of ancilla qubits is always more than the number of logical qubits for this problem, an upper-bound on the total number of qubits needed is $O\left(M^2 \log(M)\right)$.

The number of couplings needed will be dominated by the number of edges introduced by ancilla systems required as penalty terms by the bit-flip gadgets. Each of the $O\left(M^2\right)$ terms is associated with a different ancilla system which contains a number of qubits equal to the locality of that term. Furthermore, all qubits within an ancilla system are fully connected. Thus, if we again assume that all terms have maximum locality, an upper-bound on the number of couplers is $O\left(M^2 \log^2(M)\right)$. Based on this analysis, the adiabatic approach to quantum chemistry has rather modest qubit and coupler requirements.

4.5.3 Estimates of spectral gap scaling

In [10], Babbush *et al.* reduce the locality of interaction terms using perturbative gadgets from the "bit-flip" family, first introduced in [119] and later generalized by [113]. In the supplementary material presented in a later paper analyzing the scaling of gad-

get constructions [40], it is shown that for bit-flip gadgets, $\lambda^{k+1}/\Delta^k = O(\epsilon)$ and

$$\max_{ij} \left\{ \left\| \vec{\gamma_{ij}}(\epsilon) \right\|_\infty \right\} = O\left(\frac{\lambda^k}{\Delta^{k-1}} \right). \tag{4.64}$$

Here, λ is the perturbative parameter, Δ is the spectral gap, ϵ is the error in the eigenspectrum and $\vec{\gamma}_{ij}$ is the coefficient of the term to be reduced. Putting this together and representing the largest coupler value as γ, we find that $\Delta = \Omega\left(\epsilon^{-k}\gamma^k\right)$, where Ω is the "Big Omega" lower bound. Due to the Bravyi-Kitaev transformation, the locality of terms is bounded by, $k = O(\log(M))$; thus, $\Delta = \Omega\left(\epsilon^{-\log(M)}\gamma^{\log(M)}\right)$.

Prior analysis from this paper indicates that the maximum integral size is bounded by $\gamma \leq |\beta_{\max}^{\mathrm{OEI}}| M^{2/3}$. This gives us the bound,

$$\Delta = \Omega\left(\epsilon^{-\log(M)} \left\| \beta_{\max}^{\mathrm{OEI}} M^{2/3} \right\|^{\log(M)} \right). \tag{4.65}$$

However, Δ also depends polynomially on M^2, the number of terms present. Though known to be polynomial, it is extremely difficult to predict exactly how Δ depends on M^2 as applying gadgets to terms "in parallel" leads to "cross-gadget contamination" which contributes at high orders in the perturbative expansion of the self-energy used to analyze these gadgets [40]. Without a significantly deeper analysis, we can only conclude that,

$$\Delta = \Omega\left(\mathrm{poly}(M) \left\| \frac{\beta_{\max}^{\mathrm{OEI}} M^{2/3}}{\epsilon} \right\|^{\log(M)} \right). \tag{4.66}$$

This analysis indicates that the most significant challenge to implementing the adiabatic approach to quantum chemistry is the required range of coupler values which is certain to span *at least* several orders of magnitude for non-trivial systems.

This calls attention to an important open question in the field of Hamiltonian gad-

gets: whether there exist "exact" gadgets which can embed the ground state energy of arbitrary many-body target Hamiltonians without the use of perturbation theory. A positive answer to this conjecture would allow us to embed molecular electronic structure Hamiltonians without needing large spectral gaps. For entirely diagonal Hamiltonians, such gadgets are well known in the literature [11, 23] but fail when terms do not commute [40]. Exact reductions have also been achieved for certain Hamiltonians. For instance, "frustration-free" gadgets have been used in proofs of the QMA-Completeness of quantum satisfiability, and in restricting the necessary terms for embedding quantum circuits in Local Hamiltonian [48, 87, 178].

4.6 Conclusions

In this work, we analyzed the impact on scaling for quantum chemistry on a quantum computer that results from consideration of locality of interactions and exploitation of local basis sets. The impact of locality has been exploited to great advantage for some time in traditional algorithms for quantum chemistry, but has received relatively little attention in quantum computation thus far. From these considerations, we showed that in practical implementations of quantum phase estimation using a first order Suzuki-Trotter approximation, one expects a scaling cost on the order of $O(M^5 \log M)$ with respect to number of spin-orbitals, rather than more pessimistic estimates of $O(M^8)$-$O(M^9)$[99, 244] or $O(M^{5.5})$-$O(M^{6.5})$[198] related to the use of unphysical random integral distributions or the restriction to molecules too small to observe the effects of physical locality. We believe that the combination of the algorithmic improvements suggested by Poulin and Hastings et al [99, 198] with strategies that exploit locality presented here will result in even greater gains, and more work is

needed in this area.

We also considered the cost of Hamiltonian averaging, an alternative to quantum phase estimation with minimal coherence time requirements beyond state preparation. This method has some overhead with respect to quantum phase estimation, scaling as $O(M^6)$ in the number of spin-orbitals, but has significant practical advantages in coherence time costs, as well as the ability to make all measurements in parallel. This method can at best give the energy of the state provided when oracle guesses are imperfect, however it can easily be combined with a variational or adiabatic approach to improve the accuracy of the energy estimate. Moreover, while quantum phase estimation promises to be able to remove the bias of imperfect oracle guesses, we demonstrated how the cost of removal may strongly depend on how imperfect the guesses are.

Finally we analyzed the impact of locality on a complementary approach for quantum chemistry, namely adiabatic quantum computation. This approach does not have a known direct dependence on the quality of guess states provided by an oracle, and can in fact act as the state oracle for the other approaches discussed here.

In all cases, the consideration of physical locality greatly improves the outlook for quantum chemistry on a quantum computer, and in light of the goal of quantum chemistry to study physical systems rather than abstract constructs, it is correct to include this physical locality in any analysis pertaining to it. We believe that with these and other developments made in the area of quantum computation, quantum chemistry remains one of the most promising applications for exceeding the capabilities of current classical computers.

One day some as yet unborn scholar will recognize in the clock the machine that has tamed the wilds.

John Maxwell Coetzee

5

Feynman's clock, a new variational principle, and parallel-in-time quantum dynamics*

Abstract

We introduce a new discrete-time variational principle inspired by the quantum clock originally proposed by Feynman, and use it to write down quantum evolution as a ground state eigenvalue problem. The construction allows one to apply ground state quantum many-body theory to quantum dynamics, extending the reach of many highly

*__Jarrod R McClean__, John A Parkhill, and Alán Aspuru-Guzik. Feynman's clock, a new variational principle, and parallel-in-time quantum dynamics. *Proceedings of the National Academy of Sciences USA*, 110(41):E3901-E3909, 2013.

developed tools from this fertile research area. Moreover this formalism naturally leads to an algorithm to parallelize quantum simulation over time. We draw an explicit connection between previously known time-dependent variational principles and the new time-embedded variational principle presented. Sample calculations are presented applying the idea to a Hydrogen molecule and the spin degrees of freedom of a model inorganic compound demonstrating the parallel speedup of our method as well as its flexibility in applying ground-state methodologies. Finally, we take advantage of the unique perspective of the new variational principle to examine the error of basis approximations in quantum dynamics.

5.1 Introduction

Feynman proposed a revolutionary solution to the problem of quantum simulation: use quantum computers to simulate quantum systems. While this strategy is powerful and elegant, universal quantum computers may not be available for some time, and in fact, accurate quantum simulations may be required for their eventual construction. In this work, we will use the clock Hamiltonian originally introduced by Feynman[70, 72] for the purposes of quantum computation, to re-write the quantum dynamics problem as a ground state eigenvalue problem. We then generalize this approach, and obtain a novel variational principle for the dynamics of a quantum system and show how it allows for the natural formulation of a parallel-in-time quantum dynamics algorithm. Variational principles play a central role in the development and study of quantum dynamics[15, 57, 93, 103, 108, 120, 144, 202], and the variational principle presented here extends the arsenal of available tools by allowing one to directly apply efficient approximations from the ground state quantum many-body

problem to study dynamics.

Following trends in modern computing hardware, simulations of quantum dynamics on classical hardware must be able to make effective use of parallel processing. We will show below that the perspective of the new variational principle leads naturally to a time-parallelizable formulation of quantum dynamics. Previous approaches for recasting quantum dynamics as a time-independent problem include Floquet theory for periodic potentials[9, 59, 171] and more generally the (t, t') formalism of Peskin and Moiseyev[197]. However, the approach proposed in this manuscript differs considerably from these previous approaches. We derive our result from a different variational principle, and in our embedding the dynamics of the problem are encoded directly in its solution, as opposed to requiring the construction of another propagator. Considerable work has now been done in the migration of knowledge from classical computing to quantum computing and quantum information[6, 115, 116, 183, 240]. In this paper, we propose a novel use of an idea from quantum computation for the simulation of quantum dynamics.

The paper is organized as follows. We will first review the Feynman clock: a mapping stemming from the field of quantum computation that can be employed for converting problems in quantum evolution into ground-state problems in a larger Hilbert space. We then generalize the Feynman clock into a time-embedded discrete variational principle (TEDVP) which offers additional insight to quantum time-dynamics in a way that is complementary to existing differential variational principles [61, 75, 168]. We then apply the configuration interaction method [31, 216] from quantum chemistry to solve for approximate dynamics of a model spin system demonstrating convergence of accuracy of our proposed approach with level of the truncation. We demonstrate how this construction naturally leads to an algorithm that takes advan-

tage of parallel processing in time, and show that it performs favorably against existing algorithms for this problem. Finally we discuss metrics inspired by our approach that can be used to quantitatively understand the errors resulting from truncating the Hilbert space of many-body quantum dynamics.

5.1.1 Physical dynamics as a sequence of quantum gates

Consider a quantum system described by a time-dependent wavefunction $|\Psi(t)\rangle$. The dynamics of this system are determined by a Hermitian Hamiltonian $H(t)$ according to the time-dependent Schrödinger equation in atomic units,

$$i\partial_t |\Psi(t)\rangle = H(t) |\Psi(t)\rangle \tag{5.1}$$

A formal solution to the above equation can be generally written:

$$|\Psi(t)\rangle = \mathcal{T}\left(e^{-i\int_{t_0}^{t} dt' H(t')}\right) |\Psi(t_0)\rangle = U(t,t_0) |\Psi(t_0)\rangle \tag{5.2}$$

Where $\mathcal{T}$ is the well known time-ordering operator and $U(t,t_0)$ is a unitary operator that evolves the system from a time t_0 to a time t. These operators also obey a group composition property, such that if $t_0 < t_1 < ... < t_n < t$ then

$$\begin{aligned} |\Psi(t)\rangle = U(t,t_0) |\Psi(t_0)\rangle = \\ U(t,t_n)U(t_n,t_{n-1})...U(t_1,t_0) |\Psi(t_0)\rangle \end{aligned} \tag{5.3}$$

From the unitarity of these operators, it is of course also true that $U(t_n,t_{n-1})^\dagger = U(t_{n-1},t_n)$ where $\dagger$ indicates the adjoint. Thus far, we have treated time as a continuous variable. However, when one considers numerical calculations on a wavefunction,

it is typically necessary to discretize time.

We discretize time by keeping an ancillary quantum system, which can occupy states with integer indices ranging from $|0\rangle$ to $|T-1\rangle$ where T is the number of discrete time steps under consideration. This quantum system has orthonormal states such that

$$\langle i|j\rangle = \delta_{ij} \tag{5.4}$$

This definition allows one to encode the entire evolution of a physical system by entangling the physical wavefunction with this auxiliary quantum system representing time, known as the "time register". We define this compound state to be the history state, given by

$$|\Phi\rangle = \frac{1}{\sqrt{T}} \sum_t |\Psi_t\rangle \otimes |t\rangle \tag{5.5}$$

where subscripts are used to emphasize when we are considering a time-independent state of a system at time t. That is, we define $|\Psi_i\rangle = |\Psi(t)\rangle\,|_{t=t_i}$. From these definitions, it is immediately clear from above that the wavefunction at any time t can be recovered by projection with the time register, such that

$$|\Psi(t)\rangle\,|_{t=t_i} = \sqrt{T}\langle i|\Phi\rangle \tag{5.6}$$

Additionally, we discretize the action of our unitary operators, such that $U(t_1, t_0) = U_0$ and we embed this operator into the composite system-time Hilbert space as $(U_0 \otimes |1\rangle\langle 0|)$. While the utility of this discretization has not yet been made apparent, we will now use this discretization to transform the quantum dynamics problem into a ground state eigenvalue problem.

5.1.2 Feynman's clock

In the gate model[68, 183] of quantum computation, one begins with an initial quantum state $|\Psi_0\rangle$ state, applies a sequence of unitary operators $\{U_i\}$, known as quantum gates. By making a measurement on the final state $|\Psi_f\rangle$, one determines the result of the computation, or equivalently the result of applying the sequence of unitary operators $\{U_i\}$. The map from the sequence of unitary operators $\{U_i\}$ in the gate model, to a Hamiltonian that has the clock state as its lowest eigenvector is given by a construction called the Clock Hamiltonian[72]. Since its initial inception, much work has been done on the specific form of the clock, making it amenable to implementation on quantum computers[123]. However, for the purposes of our discussion that pertains to implementation on a classical computer, the following simple construction suffices

$$\begin{aligned}\mathcal{C} = C_0 + \frac{1}{2}\sum_t \big(I \otimes |t\rangle\langle t| - U_t \otimes |t+1\rangle\langle t| \\ -U_t^\dagger \otimes |t\rangle\langle t+1| + I \otimes |t+1\rangle\langle t+1|\big)\end{aligned} \tag{5.7}$$

where C_0 is a penalty term which can be used to specify the state of the physical system at any time. Typically, we use this to enforce the initial state, such that if the known state at time $t = 0$ is given by $|\Psi_0\rangle$, then

$$C_0 = (I - |\Psi_0\rangle\langle\Psi_0|) \otimes |0\rangle\langle 0| \tag{5.8}$$

One may verify by action of $\mathcal{H}$ on the history state defined above, $|\Phi\rangle$, that the history state is an eigenvector of this operator, with eigenvalue 0.

5.1.3 A discrete-time variational principle

We now introduce a new time-embedded discrete variational principle (TEDVP) and show how the above eigenvalue problem emerges as the special case of choosing a linear variational space. Suppose that one knows the exact form of the evolution operators and wavefunctions at times $t = 0, 1$. By the properties of unitary evolution it is clear that the following holds:

$$2 - \langle\Psi_1| U_0 |\Psi_0\rangle - \langle\Psi_0| U_0^\dagger |\Psi_1\rangle = 0 \tag{5.9}$$

From this, we can see that if the exact construction of U_i is known for all i, but the wavefunctions are only approximately known (but still normalized), it is true that

$$\sum_t \Big(2 - \langle\Psi_{t+1}| U_t |\Psi_t\rangle - \langle\Psi_t| U_t^\dagger |\Psi_{t+1}\rangle\Big) \geq 0 \tag{5.10}$$

where equality holds in the case that the wavefunction becomes exact at each discrete time point. Reintroducing the ancillary time-register, we may equivalently say that all valid time evolutions embedded into the composite system-time space as $\sum_t |\Psi_t\rangle |t\rangle$ minimize the quantity

$$\mathcal{S} = \sum_{t,t'} \langle t'| \langle\Psi_{t'}| \mathcal{H} |\Psi_t\rangle |t\rangle \tag{5.11}$$

where $\mathcal{H}$ (script font for operators denotes they act in the composite system-time space) is the operator given by

$$\begin{aligned}\mathcal{H} = \frac{1}{2}(&I \otimes |t\rangle\langle t| - U_t \otimes |t+1\rangle\langle t| \\ &- U_t^\dagger \otimes |t\rangle\langle t+1| + I \otimes |t+1\rangle\langle t+1|)\end{aligned} \tag{5.12}$$

It is then clear from the usual ground state variational principle, that the expectation value of the operator

$$\mathcal{S} = \sum_{t,t'} \langle t' | \langle \Psi_{t'} | \mathcal{H} | \Psi_t \rangle | t \rangle \tag{5.13}$$

is only minimized for exact evolutions of the physical system. This leads us immediately to a time-dependent variational principle for the discrete representation of a wavefunction given by:

$$\delta S = \delta \sum_{t,t'} \langle t' | \langle \Psi_{t'} | \mathcal{H} | \Psi_t \rangle | t \rangle = 0 \tag{5.14}$$

It is interesting to note, that this is a true variational principle in the sense that an exact quantum evolution is found at a minimum, rather than a stationary point as in some variational principles[168]. This is related to the connection between this variational principle and the McLachlan variational principle that is explored in the next section. However, to the authors knowledge, this connection has never been explicitly made until now.

To complete the connection to the clock mapping given above, we first note that this operator is Hermitian by construction and choose a linear variational space that spans the entire physical domain. To constrain the solution to have unit norm, we introduce the Lagrange multiplier λ and minimize the auxiliary functional given by

$$\mathcal{L} = \sum_{t,t'} \langle t' | \langle \Psi_{t'} | \mathcal{C} | \Psi_t \rangle | t \rangle - \lambda \left(\sum_{t,t'} \langle t' | \langle \Psi_{t'} | \Psi_t \rangle | t \rangle - 1 \right) \tag{5.15}$$

It is clear that this problem is equivalent to the exact eigenvalue problem on $\mathcal{H}$ with eigenvalue λ. The optimal trajectory is given by the ground state eigenvector of the operator $\mathcal{H}$. From this construction, we see that the clock mapping originally pro-

posed by Feynman is easily recovered as the optimal variation of the TEDVP in a complete linear basis, under the constraint of unit norm. Note that the inclusion of C_0 as a penalty term to break the degeneracy of the ground state is only a convenient construction for the eigenvalue problem. In the general TEDVP, one need not include this penalty explicitly, as degenerate allowed paths are excluded, as in other time-dependent variational principles, by fixing the initial state.

We note, as in the case of the time-independent variational principle and differential formulations of the time-dependent variational principle, the most compact solutions may be derived from variational spaces that have non-linear parameterizations. Key examples of this in chemistry include Hartree-Fock, coupled cluster theory[18, 51, 54], and multi-configurational time-dependent Hartree[19]. It is here that the generality of this new variational principle allows one to explore new solutions to the dynamics of the path without the restriction of writing the problem as an eigenvalue problem, as in the original clock construction of Feynman.

5.1.4 Connection to previous time-dependent variational principles

In the limit of an exact solution, it must be true that all valid time-dependent variational principles are satisfied. For that reason, it is important to draw the formal connection between our variational principle and previously known variational principles.

Consider only two adjacent times t and $t+1$, and the operator $\mathcal{H}$ defined in equation 5.12. Now suppose that the separation of physical times between discrete step t and $t+1$, denoted dt is small, and the physical system has an associated Hamiltonian

given by H, such that

$$U_t = e^{-iHdt} \approx I - iHdt - \frac{H^2 dt^2}{2} \tag{5.16}$$

By inserting this propagator into equation 5.14, rewriting the result in terms of $|\Psi(t)\rangle$, and dropping terms of order dt^3 we have

$$\begin{aligned}
&\delta\left(\langle\Psi(t)|\Psi(t)\rangle - \langle\Psi(t+dt)|\left(I - iHdt\right)|\Psi(t)\rangle\right. \\
&- \langle\Psi(t)|\left(I + iHdt\right)|\Psi(t+dt)\rangle + \\
&\left.\langle\Psi(t)|\, H^2 dt^2\, |\Psi(t)\rangle + \langle\Psi(t+dt)|\Psi(t+dt)\rangle\right) = 0
\end{aligned} \tag{5.17}$$

After defining the difference operator such that $\partial_t |\Psi(t)\rangle \equiv \left[|\Psi(t+dt)\rangle - |\Psi(t)\rangle\right]/dt$, we can factorize the above expression into

$$\delta \langle\Psi(t)|\left(i\partial_t - H\right)^\dagger \left(i\partial_t - H\right)|\Psi(t)\rangle = 0 \tag{5.18}$$

In the limit that $dt \to 0$, these difference operators become derivatives. Defining $\Theta = i\partial_t |\Psi(t)\rangle$ and only allowing variations of Θ, this is precisely the McLachlan variational principle [168].

$$\delta \| \Theta - H|\Psi(t)\rangle \|^2 = 0 \tag{5.19}$$

We then conclude that in the limit of infinitesimal physical time for a single time step, the TEDVP is equivalent to the McLachlan variational principle under these assumptions. Under the reasonable conditions that the variational spaces of the wavefunction and its time derivative are the same and that the parameters are complex analytic, then it is also equivalent to the Dirac-Frenkel and Lagrange variational principles[37].

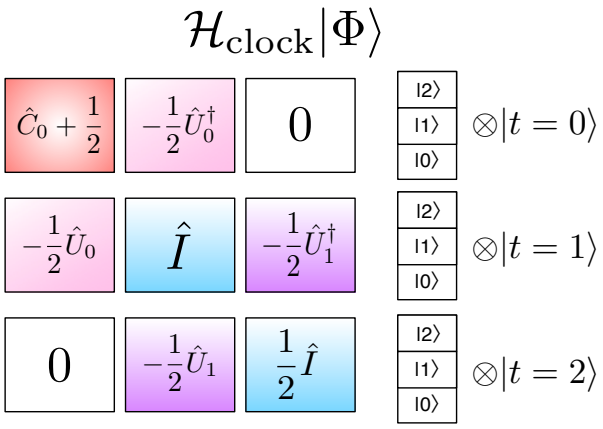

Figure 5.1: A schematic representation of the action of the clock Hamiltonian on the history state with three discrete times, and a Hilbert space of three states. Each block is a matrix with dimension of the physical system.

Moreover, as supplementary material (SI1), we provide a connection that allows other variational principles to be written as eigenvalue problems, and further discuss the merits of the integrated formalism used here.

To conclude this section, we highlight one additional difference between practical uses of the TEDVP and other variational principles: The TEDVP is independent of the method used to construct the operator U_t. In quantum information applications, this implies it is not required to know a set of generating Hamiltonians for quantum gates. Additionally, in numerical applications, one is not restricted by the choice of approximate propagator used. In cases where an analytic propagator is known for the chosen basis, it can be sampled explicitly. Suppose that the dimension of the physical system is given by N and the number of timesteps of interest is given by T. Assuming that the time register $|T\rangle$ is ordered, the resulting eigenvalue problem is block tridiagonal with dimension NT (See Fig. 5.1). This structure has been described at least once before in the context of ground-state quantum computation[172], but to the au-

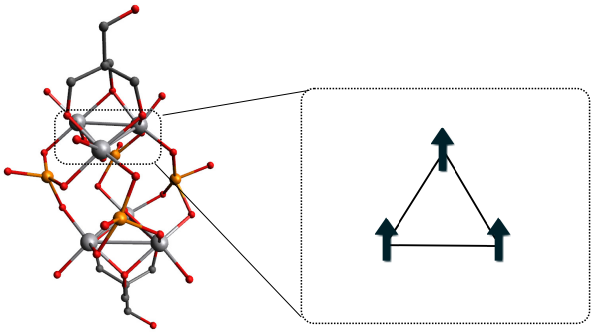

Figure 5.2: The spin triangle within the vanadium compound, used as a model system for the TEDVP. Note that coordinating sodium ions and water molecules are not depicted here. The chemical formula of this compound is given by $(CN_3H_6)_4Na_2[H_4V_6^{(IV)}O_8(PO_4)_4\ (OCH_2)_3\ CCH_2OH_2] \cdot 14H_2O$

thors knowledge, never in the context of conventional simulation of quantum systems.

5.2 Many-Body application of the TEDVP

There has been a recent rise in the interest of methods for simulating quantum spin dynamics in chemistry [105, 130]. To study the properties of the clock mapping when used to formulate approximate dynamics, we chose a simple model spin system inside an inorganic molecule [152]. Specifically, we examine the spin dynamics of the vanadium compound depicted in Fig. 5.2. By choosing the three unpaired electron spins to interact with one another by means of isotropic exchange as well as uniform static external magnetic field B_0, and a time-dependent transverse field B_1, this system can be modeled with a spin Hamiltonian

$$H = J_a(S_1 \cdot S_2 + S_1 \cdot S_3) + J_c S_2 \cdot S_3 + \mu B_0(S_1^z + S_2^z + S_3^z) + \mu B_1(S_1^x + S_2^x + S_3^x) \tag{5.20}$$

Where S_i^α is the α Pauli operator on spin i, $\mu = g\mu_b$, μ_b is the Bohr magneton, g is the measured spectroscopic splitting factor. The couplings $J_a \neq J_c$ as well as g are fitted through experimental determinations of magnetic susceptibility[152]. The fact that they are not equal is reflective of the isosceles geometry of the vanadium centers. The parameters of this model, are given by: $g = 1.95$, $J_a = 64.6 \pm 0.5K$, and $J_c = 6.9 \pm 1K$. We will allow B_0 to vary to study the properties of the clock mapping in the solution of approximate quantum dynamics.

The quantum chemistry community has decades of experience in developing and employing methods for obtaining approximate solutions of high-dimensional, ground-state eigenvector problems. By utilizing the connection we have made from dynamics to ground state problem, we will now borrow and apply the most conceptually simple approximation from quantum chemistry: configuration interaction in the space of trajectories[90], to our model problem to elucidate the properties of the clock mapping.

For the uncorrelated reference, we take the entire path of a mean-field spin evolution that is governed by the time-dependent Hartree equations, and write it as:

$$\begin{aligned} |\Psi_{MF}\rangle &= \sum_t \left(\prod_i U_t^i \left|0\right\rangle_i \right) |t\rangle = \sum_t \left(\prod_i \left|0\right\rangle_i^t \right) |t\rangle \\ &= \sum_t |\phi_t\rangle \, |t\rangle \end{aligned} \tag{5.21}$$

where $\left|0\right\rangle_i$ is the reference spin-down state for spin i, $\left|0\right\rangle_i^t$ is the reference spin-down state after rotation at time t, $|\phi_t\rangle$ is the whole product system at time t, and U_t^i is determined from the mean field Hamiltonian. The equations of motion that determine U_t^i are derived in a manner analogous to the time-dependent Hartree equations, and if

one writes the wavefunction in the physical space as

$$|\psi(t)\rangle = a(t)\prod_i U_i(t)\,|0\rangle_i = a(t)\prod_i |\phi_i\rangle = a(t)\,|\Phi(t)\rangle \tag{5.22}$$

Then the equations of motion are given by

$$a(t) = a(0) \tag{5.23}$$

$$i\dot{U}_i = (H^{(i)} - \left(\frac{f-1}{f}\right)E(t))U_i \tag{5.24}$$

Where $H^{(i)}$ is the mean field Hamiltonian for spin i formed by contracting the Hamiltonian over all other spins $j \neq i$, $E(t)$ is the expectation value of the Hamiltonian at time t, and f is the number of spins in the system.

To generate configurations, we also introduce the transformed spin excitation operator $\tilde{S}^+_{it}$, which is defined by

$$S^+_i\,|0\rangle_i = |1\rangle_i \tag{5.25}$$

$$U^{i\dagger}_t S^+_i U^{i\dagger}_t = \tilde{S}^+_{it} \tag{5.26}$$

$$\tilde{S}^+_{it}\,|0\rangle^t_i = |1\rangle^t_i \tag{5.27}$$

In analogy to traditional configuration interaction, we will define different levels of excitation (e.g. singles, doubles, ...) as the number of spin excitations at each time t that will be included in the basis in which the problem is diagonalized. For example, the basis for the configuration interaction singles (CIS) problem is defined as

$$\mathcal{B}_{\mathcal{CIS}} = \left\{\tilde{S}^+_{jt}\,|\phi_t\rangle\,|t\rangle \mid j \in [0,n], t \in [0,T)\right\} \tag{5.28}$$

Note that $\tilde{S}^+_{jt}$ for $j = n$ is simply defined to be the identity operator on all spins so that the reference configuration is naturally included. Similarly, the basis for single and double excitations (CISD) is given by

$$\mathcal{B}_{\mathcal{CISD}} = \left\{ \tilde{S}^+_{jt}\tilde{S}^+_{kt} \left|\phi_t\right\rangle \left|t\right\rangle \mid j \in [0,n], k \in [0,j), t \in [0,T) \right\} \tag{5.29}$$

Higher levels of excitation follow naturally, and it is clear that when one reaches a level of excitation equivalent to the number of spins, this method may be regarded as full configuration interaction, or exact diagonalization in the space of discretized paths.

The choice of a time-dependent reference allows the reference configuration to be nearly exact when $B_0, B_1 >> J_a, J_c$, independent of the nature of the time-dependent transverse field. This allows for the separation of the consequences of time-dependence from the effects of two versus one body spin interactions.

After a choice of orthonormal basis, the dynamics of the physical system are given by the ground state eigenvector of the projected eigenvalue problem

$$\mathcal{C}_{\mathcal{B}_i} \left|\tilde{\Phi}\right\rangle = E \left|\tilde{\Phi}\right\rangle \tag{5.30}$$

where we explicitly define the projection operator onto the basis $\mathcal{B}_i$ as $P_{\mathcal{B}_i} = \sum_{|j\rangle \in \mathcal{B}_i} |j\rangle \langle j|$ so that the projected operator is given by $\mathcal{C}_{\mathcal{B}_i} = (P_{\mathcal{B}_i} \mathcal{C} P_{\mathcal{B}_i})$

Using these constructions, we calculate the quantum dynamics of the sample system described above. For convenience, we rescale the Hamiltonian by the value of $1/\mu B_0$. To add arbitrary non-trivial time dependence to the system and mimic the interaction of the system with a transverse pulse, we take $B_1 \propto \exp(-t^2/2)\cos(mt)$. The magnitude of B_0 was taken to be $200T$ in order to model perturbative two body

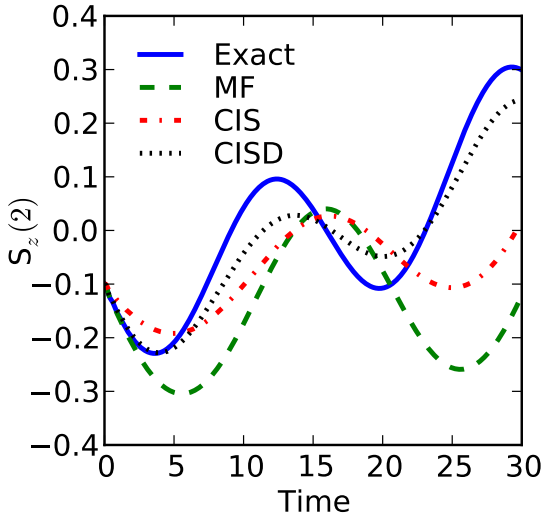

Figure 5.3: Trajectories for single-particle observables in different levels of basis set truncation starting from the reference configuration path (MF) for the Vanadium spin complex. The units of time are K^{-1}. As more levels of spin excitation are included with configuration interaction singles(CIS) and configuration interaction with singles and doubles (CISD) the trajectories converge to the exact result.

interactions in this Hamiltonian. To propagate the equations of motion and generate the propagators for the clock operator we use the enforced time-reversal symmetry exponential propagator[44] given by

$$U_t = \exp\left(-i\frac{dt}{2}H(t+dt)\right)\exp\left(-i\frac{dt}{2}H(t)\right) \tag{5.31}$$

The dynamics of some physical observables are displayed (Fig. 5.3) for the reference configuration, single excitations, double excitations, and full configuration interaction. The physical observables have been calculated with normalization at each time step. It is seen (Fig. 5.4) that as in the case of ground state electronic structure the physical observables become more accurate both qualitatively and quantitatively as the configuration space expands, converging to the exact solution with full configuration

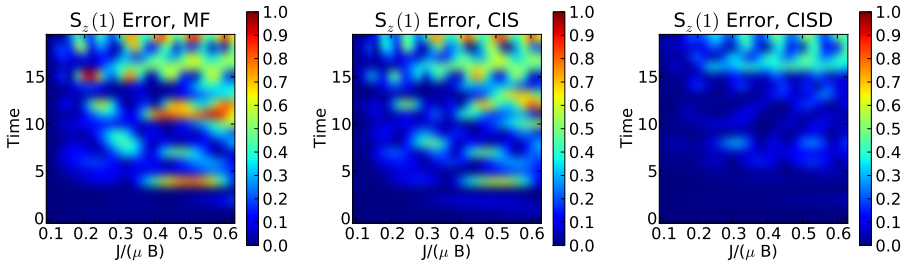

Figure 5.4: The convergence of the CI expansion *in time* is robust to strength of perturbation, and choice of initial state for the Vanadium complex. In these plots the error of the expectation value of S_z at site 1 is plotted as a function of time, and coupling constant, each propagation with a different coupling constant is begun from a different random initial product state. The expansion approaching Full CI is given by Mean Field (MF), Configuration Interaction with Single Excitations (CIS), Configuration Interaction with Single and Double Excitations (CISD), and FCI (Exact) to which the solutions are compared.

interaction. Moreover in Fig. 5.5 we plot the contributions from the reference configuration, singles space, doubles space, and triples space and observe rapidly diminishing contributions. This suggests that the time-dependent path reference used here provides an good qualitative description of the system. As a result, perturbative and other dynamically correlated methods from quantum chemistry may also be amenable to the solution of this problem.

In principle, approximate dynamics derived from this variational principle are not norm conserving, as is seen in Fig. 5.5, however this actually offers an important insight into a major source of error in quantum dynamics simulations of many-body systems, which is the truncation of the basis set as described in the last section. Simulations based on conventional variational principles that propagate within an incomplete configuration space easily preserve norm; however the trajectories of probability which should have left the simulated space are necessarily in error.

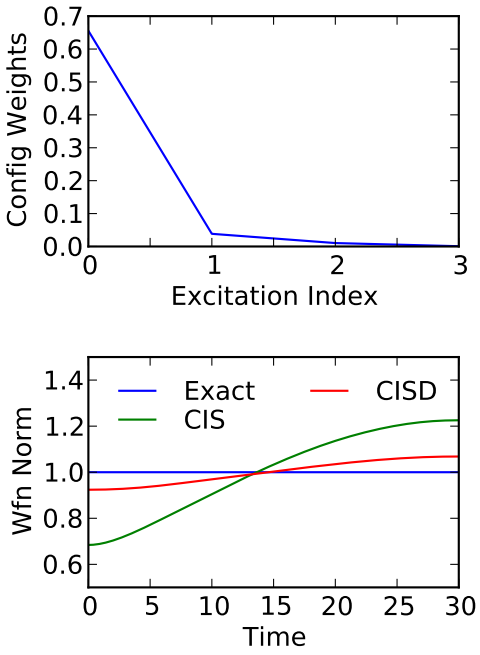

Figure 5.5: Following a simulation of the dynamics of the Vanadium spin complex, the total contribution to the ground state eigenvector from each level of excitation is plotted, where 0=MF(Mean-Field), 1=CIS (Configuration Interaction with Single Excitations), 2=CISD (Configuration interaction with Single and Double Excitations), and 3=FCI (Exact Diagonalization), and seen to decrease with excitation supporting the quality of the time-dependent reference state. The units of time are K^{-1}. The deviation of the wavefunction norm from unity resulting from projection is seen to decrease monotonically with level of excitation.

5.3 Parallel-in-time quantum dynamics

Algorithms that divide a problem up in the time domain, as opposed to spatial domain, are known as parallel-in-time algorithms. Compared to their spatial counterparts, such as traditional domain decomposition[218], these algorithms have received relatively little attention. This is perhaps due to the counterintuitive notion of solving for future times in parallel with the present. However as modern computational architectures continue to evolve towards many-core setups, exploiting all avenues of parallel speedup available will be an issue of increasing importance. The

most popular parallel-in-time algorithm in common use today is likely the *parareal* algorithm[14, 143]. The essential ingredients of parareal are the use of a coarse propagator U^c that performs an approximate time evolution in serial, and a fine propagator U^f that refines the answer and may be applied in parallel. The two propagations are combined with a predictor-corrector scheme. It has been shown to be successful with parabolic type equations[77], such as the heat equation, but it has found limited success with wave-like equations[76], like the time-dependent Schrödinger equation. We will now show how our variational principle can be used to naturally formulate a parallel-in-time algorithm, and demonstrate its improved domain of convergence with respect to the parareal algorithm for Schrödinger evolution of Hydrogen.

Starting from the TEDVP, minimization under the constraint that the initial state is fixed yields a Hermitian positive-definite, block-tridiagonal linear equation of the form

$$\mathcal{R}^f \left|\Phi\right\rangle = \left|\Lambda\right\rangle \tag{5.32}$$

where $|\Phi\rangle$(to be solved for) contains the full evolution of the system and $|\Lambda\rangle$ specifies the initial condition such that

$$\left|\Lambda\right\rangle = \begin{pmatrix} \frac{1}{2}\left|\Psi_0\right\rangle & 0 & 0 & 0 & \ldots & 0 \end{pmatrix}^T \tag{5.33}$$

The linear clock operator, $\mathcal{R}^i$, is similar to before, where now we distinguish between a clock built from a coarse operator, $\mathcal{R}^c$, from a clock built from a fine opera-

tor, $\mathcal{R}^f$, as

$$\mathcal{R}^i = \begin{pmatrix} I & -\frac{1}{2}U_0^{i\dagger} & 0 & ... & \\ -\frac{1}{2}U_0^i & I & -\frac{1}{2}U_1^{i\dagger} & & \\ 0 & -\frac{1}{2}U_1^i & I & & \\ & \vdots & \ddots & \ddots & \\ & & -\frac{1}{2}U_{T-2}^i & I & -\frac{1}{2}U_{T-1}^{i\dagger} \\ & & 0 & -\frac{1}{2}U_{T-1}^i & \frac{1}{2}I \end{pmatrix} \tag{5.34}$$

The spectrum of this operator is positive-definite and admits T distinct eigenvalues, each of which is N-fold degenerate. The conjugate gradient algorithm can be used to solve for $|\Phi\rangle$, converging at-worst in a number of steps which is equal to the number of distinct eigenvalues[67, 104]. Thus application of the conjugate gradient algorithm to this problem is able to converge requiring at most T applications of the linear clock operator $\mathcal{R}^f$. This approach on its own yields no parallel speedup. However, the use of a well-chosen preconditioner can greatly accelerate the convergence of the conjugate gradient algorithm[64].

If one uses an approximate propagation performed in serial, $\mathcal{R}^c$, which is much cheaper to perform than the exact propagation, as a preconditioning step to the conjugate gradient solve, the algorithm can converge in far fewer steps than T and a parallel-in-time speedup can be achieved. The problem being solved in this case for the clock construction is given by

$$(\mathcal{R}^c)^{-1}\mathcal{R}^f |\Phi\rangle = (\mathcal{R}^c)^{-1} |\Lambda\rangle \tag{5.35}$$

To clarify and compare with existing methods, we now introduce an example from

chemistry. The nuclear quantum dynamics of the Hydrogen molecule in its ground electronic state can be modeled by the Hamiltonian

$$H = -\frac{\hat{P}^2}{2m} + D\left(e^{-2\beta\hat{X}} - 2e^{-\beta\hat{X}}\right) \tag{5.36}$$

where $m = 918.5$, $\beta = 0.9374$, and $D = 0.164$ atomic units[158]. The initial state of our system is a gaussian wavepacket with a width corresponding to the ground state of the harmonic approximation to this potential, displaced -0.1 Å from the equilibrium position. To avoid the storage associated with the propagator of this system and mimic the performance of our algorithm on a larger system, we use the symmetric matrix-free split-operator Fourier transform method(SOFT) to construct block propagators[69]. We note that this splitting is commonly referred to as the Suzuki-Trotter formula[222, 233] in the physics literature. This method is unconditionally stable, accurate to third order in the time step dt, and may be written as

$$U_{SOFT}(t+dt,t) = e^{-iV(\hat{X})dt/2}F^{-1}e^{-i\hat{P}^2/(2m)dt}Fe^{-iV(\hat{X})dt/2} \tag{5.37}$$

Here, F and F^{-1} corresponds to the application of the fast Fourier transform (FFT) and its inverse over the wavefunction grid. The use of the FFT allows each of the exponential operators to be applied trivially to their eigenbasis and as a result the application of the propagator has a cost dominated by the FFT that scales as $O(N \log N)$, where N is the number of grid-points being used to represent $|\psi\rangle$. For our algorithm, we define a fine propagator, U^f, and a coarse propagator, U^c from the SOFT con-

struction, such that for a given number of sub-time steps k.

$$U^f = U_{SOFT}(t + kdt, t + (k-1)dt)...U_{SOFT}(t + dt, t) \tag{5.38}$$

$$U^c = U_{SOFT}(t + kdt, t) \tag{5.39}$$

We take for our problem the clock constructed from the fine-propagator and use the solution of the problem built from the coarse propagator as our preconditioner. In all cases, only the matrix free version of the propagator is used, including in the explicit solution of the coarse propagation.

One must emphasize here the importance of the matrix free aspect of the propagator. Consider, for example, an alternative scheme of parallelization where the propagators U_t are computed simultaneously by T processors and stored for application to the vector. While this scheme has apparently high parallel efficiency, the explicit calculation of a propagator with equivalent accuracy to the SOFT method can scale as $O(N^3)$ in the size of the system and require the storage of a dense matrix of size $O(N^2)$. This makes it impractical for many problems of interest and is the reason we emphasize a scalable, matrix free approach here.

From the construction of the coarse and fine propagators, with T processors, up to communication time, the cost of applying the fine clock in parallel and solving the coarse clock in serial require roughly the same amount of computational time. This is a good approximation in the usual case where the application of the propagators is far more expensive than the communication required. From this, assuming the use of T processors, we define an estimated parallel-in-time speed-up for the computational procedure given by

$$S_{clock} = \frac{T}{2(N_f + N_c)} \tag{5.40}$$

where N_f is the number of applications of the fine-propagator matrix $\mathcal{R}^f$ performed in parallel and N_c is the number of serial linear solves using the coarse-propagator matrix $\mathcal{R}^c$ used in preconditioned conjugate gradient. The factor of 2 originates from the requirement of backwards evolution not present in a standard propagation. The cost of communication overhead as well as the inner-products in the CG algorithm are neglected for this approximate study, assuming they are negligible with respect to the cost of applying the propagator.

The equivalent parallel speedup for the parareal algorithm is given by

$$S_{para} = \frac{T}{N_f + N_c} \tag{5.41}$$

where N_f and N_c are now defined for the corresponding parareal operators which are functionally identical to the clock operators without backward time evolution, and thus it lacks the same factor of 2.

As is stated above, in the solution of the linear clock without preconditioning, it is possible to converge the problem in at most T steps, independent of both the choice of physical timestep and the size of the physical system N by using a conjugate gradient method. However, with the addition of the preconditioner, the choice of timestep and total time simulated can have an effect on the total preconditioned system. This is because as the coarse (approximate) propagation deviates more severely from the exact solution, the preconditioning of the problem can become quite poor.

This problem exhibits itself in a more extreme way for the parareal algorithm, as the predictor-corrector scheme may start to diverge for very poorly preconditioned system. This has been seen before in the study of hyperbolic equations[76], and remains true in this case for the evolution of the Schrödinger equation. The construc-

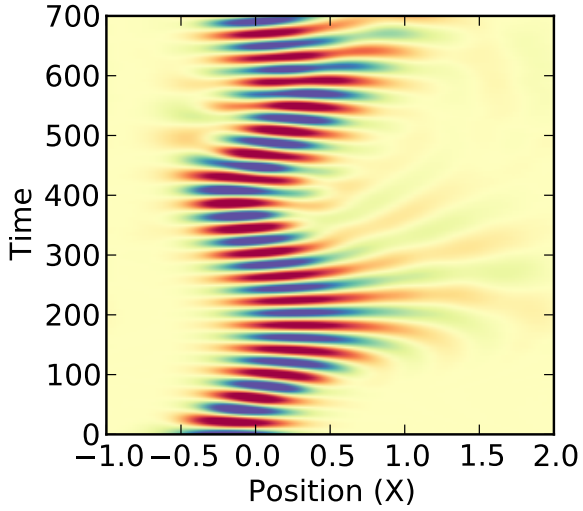

Figure 5.6: The real part of the Clock solution for the nuclear dynamics of the hydrogen molecule oscillates around an equilibrium bond length as expected, eliciting diverse interference patterns due to anharmonicity. Both time and position are given in atomic units and the color indicates the value of the real part of the waveform at that space-time point.

tion derived from the clock is more robust and is able to achieve parallel-in-time speedup for significantly longer times. This marks an improvement over the current standard for parallel-in-time evolution of the Schrödinger equation.

To give a concrete example, consider the case where we divide the evolution of the nuclear dynamics of hydrogen into pieces containing T evolution blocks, each of which is constructed from T fine evolutions for a time $dt = 0.015$ as is described above. The dynamics over which we simulate are depicted in Fig. 5.6. We average the estimated parallel speedup for 10 time blocks (which we define as the whole time in one construction of the clock) forward and the results are for the speedup are given in Fig. 5.7. In this example we see that for small T (or small total evolution times), the reduced overhead of having no backwards evolution yields an advantage for the parareal algorithm. However as the T increases, the parareal algorithm is less robust to er-

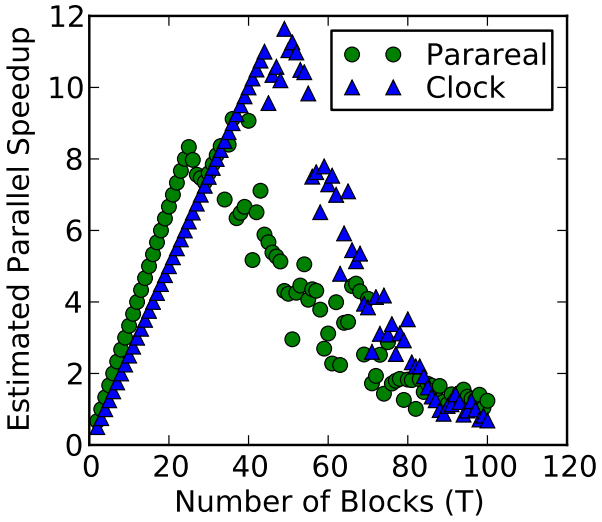

Figure 5.7: In simulating the nuclear dynamics of the Hydrogen molecule, the clock formulation not only demonstrates higher peak parallel speedup compared to parareal, but more robust speedup for longer total evolution times. This is an isoaccuracy comparison in that all points are converged to an identical level of accuracy corresponding to a fine timestep of $dt = 0.015$. The results have been averaged for 10 consecutive evolutions of the specified number of blocks T. For small times the parareal algorithm has a slight advantage due to reduced overhead, but as the total evolution time increases it is less robust to the diminishing quality of the preconditioner. The non-monotonic nature of the speedup results from the preconditioner having variable quality depending on the dynamics of the system.

rors in the coarse propagation and performance begins to degrade rapidly. In these cases, our clock construction demonstrates a clear advantage. It is a topic of current research how to facilitate even longer evolutions in the clock construction.

5.4 Norm Loss as a measure of truncation error

Conservation of the norm of a wavefunction is often considered a critical property for approximate quantum dynamics, as it is a property of the exact propagator resulting from time-reversal symmetry. However, if norm conservation is enforced in a propagation which does not span the complete Hilbert-space, one must account for the com-

ponents of the wavefunction that would have evolved into the space not included in the computation. It's not immediately clear how population density which attempts to leave the selected subspace should be folded back into the system without being able to simulate the exact dynamics. This problem is sometimes glossed over with the use of exponential propagators that are guaranteed to produce norm-conserving dynamics on a projected Hamiltonian. Some more sophisticated approaches adjust the configuration space in an attempt to mitigate the error[248].

This discrepancy is at the center of the difference between the approximate dynamics derived from the discrete variational principle here and the approximate dynamics derived from the McLachlan variational principle such as the multi-configurational time-dependent Hartree method. Mathematically, this results from the non-commutativity of the exponential map and projection operator defined above. That is

$$P_{\mathcal{B}_i(t)}\mathcal{T}\left(e^{-i\int dt' H(t')}\right)P_{\mathcal{B}_i(t)} \neq \mathcal{T}\left(\exp\left(-i\int dt' P_{\mathcal{B}_i(t')}H(t')P_{\mathcal{B}_i(t')}\right)\right) \tag{5.42}$$

in general for a Hermitian operator H. In an approximate method derived from the McLachlan or any of the other differential time-dependent variational principles, the projection is performed on the Hamiltonian. As the projection of any Hermitian operator yields another Hermitian operator, the dynamics generated from the projection are guaranteed to be unitary if a sufficiently accurate exponential propagator is used. In contrast, projection of a unitary operator, as prescribed by the TEDVP, does not always yield a unitary operator. Thus for an approximate basis, one expects norm conservation to be violated, where the degree of violation is related to the severity of the configuration space truncation error. This leads us to define a metric of trunca-

tion error which we term the instantaneous norm loss. We define this as

$$\mathcal{N}_1(t) = 1 - \| P_{\mathcal{B}_i} U_t P_{\mathcal{B}_i} |\tilde{\Psi}_t\rangle \|^2 \tag{5.43}$$

where $|\tilde{\Psi}_t\rangle$ is always assumed to be normalized, which in practice means that a renormalization is used after each time step here. However, as we proved above, in the limit of a short time step, with dynamics generated by a Hamiltonian, the TEDVP must contain essentially the same content as the McLachlan variational principle. For this reason, we propose an additional metric which is given by

$$\mathcal{N}_2(t) = \| (H(t) - P_{\mathcal{B}_i} H(t) P_{\mathcal{B}_i}) |\tilde{\Psi}_t\rangle \|^2 dt^2 = \| V(t) |\tilde{\Psi}_t\rangle \|^2 dt^2 \tag{5.44}$$

Where $H(t)$ is the physical Hamiltonian. This is motivated by appealing to the McLachlan variational principle and substituting from the exact Schrödinger equation that $i\partial_t = H |\Psi_t\rangle$ where H is the full (non-projected) Hamiltonian. By defining $V(t) = (H(t) - P_{\mathcal{B}_i} H(t) P_{\mathcal{B}_i})$, we recognize this as a perturbation theory estimate of the error resulting from the configuration basis truncation at a given point in time.

To examine the quality of these metrics and to better understand the consequences of the non-commutativity of the exponential map and projection, we return to the sample spin system. In this case, we choose a basis for the propagation space based entirely on the initial state, and do not allow it to change dynamically in time as before. We perform simulations in the space of single excitations (S) from the initial state, double excitations (SD), and in the full Hilbert space (Ex). Dynamics from the TEDVP are generated by first building the exact propagator then projecting to the desired basis set while dynamics from the McLachlan variational principle (MVP)

are modeled by projecting the Hamiltonian into the target basis set and exponentiating. A renormalization is used after each time step in the first case. Although one could perform the simulation with a timestep where timestep error is negligible, we remove this component of the calculation for this example by making the Hamiltonian time-independent. This allows direct analysis of the effect of timestep on non-commutativity deviations.

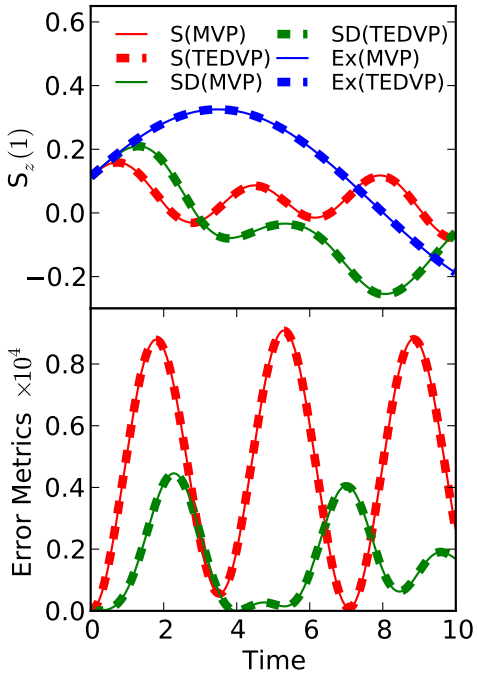

Figure 5.8: The dynamics of a spin observable in the Vanadium spin complex at a short time step, $dt = 0.01K^{-1}$, are indistinguishable when generated with equations of motion determined by the time embedded discrete variational principle (TEDVP) (dashed lines) and the McLachlan variational principle (MVP) (solid lines). Results are shown for propagations restricted to the space of single excitations from the initial state(S), double excitations (SD), and the full space (Ex). The associated error metrics, $\mathcal{N}_1(t)$ for the TEDVP (dashed) and $\mathcal{N}_2(t)$ for the MVP (solid), also yield nearly identical results, displaying peaks correlated with qualitative deviations from the exact trajectory.

In Fig. 5.8 we show the dynamics of the Vanadium spin complex for the two ap-

proximate truncation levels (S and SD) with both methods and their associated error metrics ($\mathcal{N}_1(t)$ and $\mathcal{N}_2(t)$). Deviations in the qualitative features of the observable occur after even the first peak of the proposed metrics. In this particular simulation, the configuration interaction with singles and doubles spans all but one state in the full Hilbert space. The lack of this one state results in the large qualitative errors present, associated with the first and subsequent peaks present in these error metrics. The impact of later peaks is more difficult to discern, due to error in the wavefunctions, which accumulates as the propagation proceeds.

As predicted by the connection between the TEDVP and the McLachlan variational principle, while $\mathcal{N}_1$ and $\mathcal{N}_2$ are not identical for each case, in the short time limit they yield extremely similar information, which is highlighted in Fig. 5.8 displaying the resulting longer time dynamics for a time step of $dt = 0.01$. In Fig. 5.9, however we explore the effects of a significantly larger time step and begin to discern the result of the non-commutativity discussed here. Recalling that because the Hamiltonian is time-independent in this case, the propagator used is numerically exact in both instances, so this effect is not a result of what would be traditionally called time step error, resulting from intrinsic errors of an integrator. Interestingly, it is observed that $\mathcal{N}_1(t)$ begins to decay to a nearly constant value. This occurs because the action of projection after exponentiation breaks the degeneracy of the spectrum of the unitary operator, resulting in eigenvalues with norms different than 1. As a result, repeated action and renormalization of the operator is analogous to a power method for finding the eigenvector associated with the largest eigenvalue. This effect is exacerbated by taking long time steps.

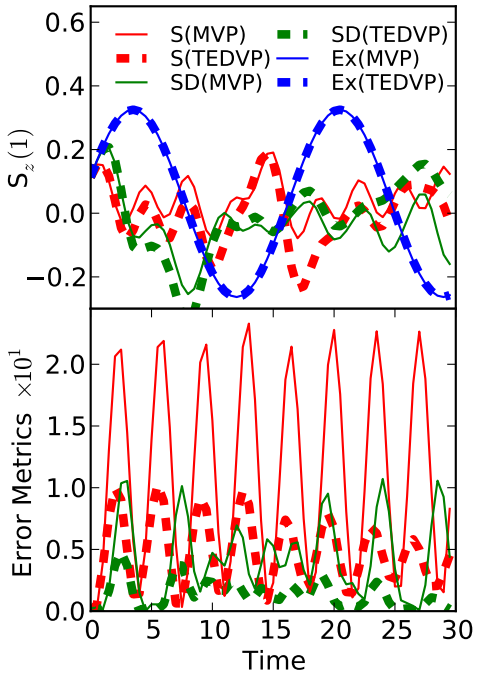

Figure 5.9: The dynamics of a spin observable in the Vanadium complex at a large time step, $dt = 0.5K^{-1}$, shows significant differences resulting from the non-commutativity of projection and exponentiation in dynamics generated by the time embedded discrete variational principle(TEDVP) (dashed lines) and the McLachlan variational principle (MVP) (solid lines). Results are shown for propagations restricted to the space of single excitations from the initial state (S), double excitations (SD), and the full space (Ex). The corresponding error metrics, $\mathcal{N}_1(t)$ for the TEDVP (dashed) and $\mathcal{N}_2(t)$ for the MVP (solid), differ considerably in this case.

5.5 Conclusions

In this manuscript, we introduce a new variational principle for time-dependent dynamics inspired by the Feynman clock originally employed for quantum computation. Unlike other previously-proposed variational principles, the proposed TEDVP approach involves the solution of an eigenvalue problem for the entire time propagation. This perspective allows for readily employing many of the powerful truncation techniques from quantum chemistry and condensed-matter physics, that have been devel-

oped for the exact diagonalization problem. We show how this formulation naturally leads to a parallel-in-time algorithm and demonstrate its improved robustness with respect to existing methods. We introduce two novel error metrics for the TEDVP that allow one to characterize the basis approximations involved. The features of the method were demonstrated by simulating the dynamics of a Hydrogen molecule and a molecular effective spin Hamiltonian. Further research directions include the use of other approximate techniques for the time dynamics such as the use of perturbation theory [102] or coupled-cluster approaches[132], and further enhancement of parallel-in-time dynamics.

5.6 Supplemental Information

5.6.1 On the general construction of eigenvalue problems from dynamics

We have presented one path for constructing eigenvalue problems from quantum dynamics problems so far, however it is instructive to illuminate precisely which part of our procedure allowed this. To do this, we will slightly generalize the Floquet-type Hamiltonians and demonstrate that the time embedding was the crucial feature that allows one to recast a dynamics problem as an eigenvalue problem. This is in addition to the choice to work in an integrated framework, which we will show simply allows for a convenient choice of basis.

Recall the definition of a Floquet-type Hamiltonian given by

$$F(t) = H(t) - i\partial_t \tag{5.45}$$

If one considers a finite time evolution for a length of time T, this operator is Hermitian in the basis of Fourier functions states given by

$$|\Psi_{nj}(t)\rangle = |\Phi_j\rangle\,|n\rangle = |\Phi_j\rangle\,e^{2\pi int/T} \tag{5.46}$$

where $|\Phi_j\rangle$ is a time-independent state of the physical system when considering the generalized inner-product

$$\langle n'|\,\langle\Phi_i|\Phi_j\rangle\,|n\rangle = \frac{\langle\Phi_i|\Phi_j\rangle}{T}\int_0^T dt' e^{-2\pi in't'/T} e^{2\pi int'/T} \tag{5.47}$$

This is the the generalized Hilbert space first introduced by Sambe[207] and generalized by Howland[106]. Because this operator is Hermitian in this basis, by noting the similarity to the Lagrange Variational principle

$$\delta L = \delta \int_0^T dt\,\langle\Psi(t)|\,F\,|\Psi(t)\rangle = 0 \tag{5.48}$$

minimization of L on this linear basis of Fourier states yields a Hermitian eigenvalue problem. Thus the time evolution can be reconstructed by solving the full time-independent eigenvalue problem in this basis, or by constructing a surrogate evolution operator as in the (t, t') formalism of Peskin and Moiseyev[197]. The use of Fourier basis states to express time dependence is natural given the form of the operator F. That is, matrix elements of the derivative operator ∂_t have a trivial analytic expression in this basis. This represents the same general idea we have been discussing here, which is to consider states in a system-time Hilbert space. However, as the solution of this variational problem will may yield a stationary point rather than a true minimum[168], ground state techniques are not appropriate for this particular operator. Moreover,

this operator is not in general Hermitian when considering arbitrary basis functions of time. For example arbitrary choices of plane waves not corresponding to the traditional Fourier basis will yield a non-Hermitian operator. The operator $F' = (1/2)(F + F^\dagger)$, which has been used in the past for the construction of approximate dynamics[211], is Hermitian, however it still suffers from a pathology that the optimal solution represents a stationary point rather than a minimum. However, the operator

$$G = F^\dagger F = (H(t) - i\partial_t)^\dagger (H(t) - i\partial_t) \tag{5.49}$$

is always Hermitian and positive semi-definite by construction. Thus one can expand the system-time Hilbert space in any linear basis of time, and the optimal path in that space will be the ground state eigenvector of the operator G utilizing the above generalized inner-product, assuming we have broken degeneracy by introducing the correct initial state. This can be viewed as an application of the McLachlan time dependent variational principle. From this, we see it is the expression systems in a system-time Hilbert space which allows for the eigenvalue problem construction. Moreover, we note that this is not limited to the use of Fourier time basis states, and that many more expressions of time dependence may be utilized to construct an eigenvalue problem within this framework.

One may ask then, what was the objective of working in the an integrated formalism, defined using unitary operators rather than differential operators. To see this, consider evaluating the system at a point in time with the Fourier construction above. One has to expand the system in all time basis functions and evaluate them at a time t. When one usually considers time, however, they are thinking of a parametric construction where a number t simply labels a specific state of a system. Embedding into

the system-time Hilbert space with this idea would be most naturally expressed as delta-functions. However matrix elements of ∂_t on this basis can be difficult to construct. In the ancillary time system framework used in the TEDVP, however, time is easily expressed as a discrete parametric variable. One might also consider the use of discretized derivatives in the operator G. However, balancing the errors in numerical derivatives and the increasing difficultly of solving the problem with the number of simultaneously stored time steps can be practically difficult.

Consideration of black holes suggests, not only that God does play dice, but that he sometimes confuses us by throwing them where they can't be seen.

Stephen Hawking

6

Clock Quantum Monte Carlo: an imaginary-time method for real-time quantum dynamics*

Abstract

In quantum information theory, there is an explicit mapping between general unitary dynamics and Hermitian ground state eigenvalue problems known as the Feynman-Kitaev Clock Hamiltonian. A prominent family of methods for the study of quantum ground states are quantum Monte Carlo methods, and recently the full configura-

*Reprinted with permission from **Jarrod R. McClean**, Alán Aspuru-Guzik. Clock Quantum Monte Carlo: an imaginary-time method for real-time quantum dynamics. *Physical Review A* 91, 012311, 2015. Copyright (2015) by the American Physical Society.

tion interaction quantum Monte Carlo (FCIQMC) method has demonstrated great promise for practical systems. We combine the Feynman-Kitaev Clock Hamiltonian with FCIQMC to formulate a new technique for the study of quantum dynamics problems. Numerical examples using quantum circuits are provided as well as a technique to further mitigate the sign problem through time-dependent basis rotations. Moreover, this method allows one to combine the parallelism of Monte Carlo techniques with the locality of time to yield an effective parallel-in-time simulation technique.

6.1 Introduction

Understanding the evolution of quantum systems is a central problem in physics and the design of emerging quantum technologies. However, exact simulations of quantum dynamics suffer from the so-called curse of dimensionality [224]. That is, the dimension of the Hilbert space grows exponentially with the size of the physical system. An effective remedy for the curse of dimensionality in some classical systems has been the use of Monte Carlo methods, which in many cases has an error with respect to number of samples that is independent of the dimension of the simulated system [25]. Unfortunately this favorable scaling is often lost in quantum systems of interest due to the emergence of the famous sign problem. In particular, it has hindered the use of Monte Carlo methods for fermionic systems, where it is sometimes called "the fermion sign problem", and for real-time dynamics of general quantum systems, where it is known as "the dynamical sign problem". The generic sign problem has been proven to belong to the computational complexity class NP-Complete [234], and recent studies of complexity have refined knowledge about the computational power of sign-

problem free (or "stoquastic") Hamiltonians [33, 34]. However, these results do not preclude the effective use of these methods on many interesting instances of physical problems.

In particular, despite the generic challenges of the sign problem, Monte Carlo methods have been used with great success in the study of electronic systems, providing a standard of accuracy in quantum chemistry and condensed matter [16, 74, 96, 184, 270]. In some of these methods, such as fixed node diffusion Monte Carlo, the use of a trial wavefunction allows one to approximately remove the complications of the sign problem at the cost of a small bias in the resulting energy. One alternative to such an approximation is the use of interacting walker methods [269], which attempt to solve the problem exactly without the bias introduced by a trial function. Recently, Booth et. al introduced an interacting walker method in the discrete basis of Slater determinants called Full Configuration Interaction Quantum Monte Carlo (FCIQMC) [29]. The sign problem in the context of this algorithm has been studied in some detail [128, 213, 220] and it has been successfully applied to both small molecular systems of chemical interest and extended bulk systems [26, 27].

The use of Monte Carlo methods to study the real-time dynamics of generic quantum systems has been comparatively less prevalent [155]. The dynamical sign problem may become more severe both with the size of the system, and duration for which it is simulated [71, 159, 228]. Despite these challenges, advances are being made in the treatment of these problems, including hybridization of Monte Carlo techniques with other methods [109, 156–158, 160].

The sign problem has been studied in the context of quantum computation, where it is known that a sufficient condition for efficient probabilistic classical simulation of the adiabatic evolution of a quantum system using Monte Carlo methods is that

the Hamiltonian governing the quantum system is sign problem free (also known as stoquastic) and frustration free [33, 34, 49]. Projector Monte Carlo algorithms have been developed specifically for this type of problem [32, 33]. Moreover, the use of tools from quantum information allows any generic unitary evolution of a quantum system to be written as the ground state eigenvalue problem of a Hermitian Hamiltonian [70, 123, 167]. In this work, we exploit this equivalence to adapt the interacting walker method introduced by Booth et. al [29] to treat the dynamical sign problem with a method designed for the fermion sign problem.

The paper is organized as follows. First, we review the time-embedded discrete variational principle [167] and derive from it the Clock Hamiltonian [70, 123, 167], which are the essential tools for writing a general unitary evolution as a ground state eigenvalue problem of a Hermitian Hamiltonian. We then review the FCIQMC method and adapt it for application to the Clock Hamiltonian. A discussion of the theoretical and practical manifestation of the dynamical sign problem in this setting follows with numerical examples from quantum computation. Finally, we introduce a general framework of basis rotations which can be used to ameliorate the sign problem and study the performance of this method when used in parallel computation.

6.2 Dynamics as a ground state problem

It has been shown that any unitary quantum evolution may be formulated as a ground state eigenvalue problem with applications to classical simulation of quantum systems [167]. We briefly review the relevant results here so that this work remains self-contained.

Consider a quantum system that is described at discrete time steps t by a normal-

ized wavefunction $|\Psi_t\rangle$. The dynamics of this system are described by a sequence of unitary operators $\{U_t\}$ such that $U_t |\Psi_t\rangle = |\Psi_{t+1}\rangle$ and $U_t^\dagger |\Psi_{t+1}\rangle = |\Psi_t\rangle$. In the case of simulating Hamiltonian dynamics, these U_t could be obtained from a Suzuki-Trotter factorization of the total evolution [222, 233]. However, we stress that explicit construction of a full unitary operator U_t is never required, only the ability to efficiently evaluate matrix elements between different physical states as detailed in a previous work [167]. From the properties of unitary evolution, the following is clear:

$$2 - \langle\Psi_{t+1}| U_t |\Psi_t\rangle - \langle\Psi_t| U_t^\dagger |\Psi_{t+1}\rangle = 0. \tag{6.1}$$

Moreover, if the wavefunctions at each point in time are only approximately known (but normalized) then

$$\sum_t \Big(2 - \langle\Psi_{t+1}| U_t |\Psi_t\rangle - \langle\Psi_t| U_t^\dagger |\Psi_{t+1}\rangle\Big) \geq 0 \tag{6.2}$$

where equality holds only in the case where the wavefunction represents an exact, valid evolution of the quantum system. To consider all moments in time simultaneously, we extend the physical Hilbert space with an ancillary quantum system to denote time. This ancillary time register takes on integer values t and is orthonormal such that $\langle t'|t\rangle = \delta_{t,t'}$. With this construction, we see that by defining

$$\begin{aligned}\mathcal{H}' = \frac{1}{2}\Big(\sum_t\ & I \otimes |t\rangle\langle t| - U_t \otimes |t+1\rangle\langle t| \\ & - U_t^\dagger \otimes |t\rangle\langle t+1| + I\,|t+1\rangle\langle t+1|\Big)\end{aligned} \tag{6.3}$$

which acts on the composite system-time Hilbert space, all valid time evolutions mini-

mize

$$\mathcal{S} = \sum_{t,t'} \langle t' | \langle \Psi_{t'} | \mathcal{H}' | \Psi_t \rangle | t \rangle . \tag{6.4}$$

Note that we have adopted the convention of script letters for operators acting in the system-time Hilbert space such as $\mathcal{H}'$ as opposed to operators only acting on the system such as U_t. The time-embedded discrete variational principle immediately follows, which simply states that this quantity is stationary under variations of the wavefunction $\delta | \Psi_t \rangle$ for all valid time evolutions, or

$$\delta \mathcal{S} = \delta \sum_{t,t'} \langle t' | \langle \Psi_{t'} | \mathcal{H}' | \Psi_t \rangle | t \rangle = 0 \tag{6.5}$$

To select a particular evolution of interest, one may introduce a penalty operator that fixes the state of the system at a given time. Typically, this might represent a known initial state, and this operator in the system-time Hilbert space is given by

$$\mathcal{C}_0 = (I - | \Psi_0 \rangle \langle \Psi_0 |) \otimes | 0 \rangle \langle 0 | . \tag{6.6}$$

The minimization of a Hermitian quadratic form constrained to have unit norm is equivalent to the eigenvalue problem for the corresponding Hamiltonian. We introduce the Lagrange multiplier λ to enforce normalization. As both $\mathcal{S}$ and C_0 are Hermitian by construction, minimization of

$$\begin{aligned} \mathcal{L} = & \sum_{t,t'} \langle t' | \langle \Psi_{t'} | \mathcal{H}' + \mathcal{C}_0 | \Psi_t \rangle | t \rangle \\ & - \lambda \left(\sum_{t,t'} \langle t' | \langle \Psi_{t'} | \Psi_t \rangle | t \rangle - 1 \right) \end{aligned} \tag{6.7}$$

is equivalent to solving for the eigenvector corresponding to the smallest eigenvalue of the Hermitian operator

$$\mathcal{H} = \mathcal{H}' + C_0 \tag{6.8}$$

which we refer to as the Clock Hamiltonian. This Hamiltonian has a unique ground state with eigenvalue 0 given by the history state,

$$|\Phi\rangle = \frac{1}{\sqrt{T}} \sum_t |\Psi_t\rangle \otimes |t\rangle \tag{6.9}$$

which encodes the entire evolution of the quantum system. Thus, the quantum dynamics of the physical system can be obtained by finding the ground state eigenvector of $\mathcal{H}$.

6.3 FCIQMC for the Clock Hamiltonian

The Full Configuration Interaction Quantum Monte Carlo (FCIQMC) method was introduced by Booth et. al as a method to accurately find the ground state for electronic structure problems in a basis of Slater determinants without appealing to the fixed node approximation to eliminate the fermion sign problem [29]. We review the basics of the theory behind this method and show how it can be adapted for the Clock Hamiltonian $\mathcal{H}$, such that it simulates the full time evolution of a quantum system.

Let $|\Phi_i\rangle$ and λ_i be the eigenvectors and corresponding eigenvalues of $\mathcal{H}$. Any vector $|\Psi\rangle$ in the system-time Hilbert space acted upon by $\mathcal{H}$ can be decomposed in terms of

the eigenvectors of $\mathcal{H}$ such that

$$|\Psi\rangle = \sum_i c_i |\Phi_i\rangle \tag{6.10}$$

It follows that for any $|\Psi\rangle$ not orthogonal to the ground state of the Clock Hamiltonian, $|\Phi_0\rangle$, that

$$\lim_{\tau\to\infty} e^{-\tau\mathcal{H}} |\Psi\rangle = \lim_{\tau\to\infty} \sum_i e^{-\tau\lambda_i} c_i |\Phi_i\rangle \propto |\Phi_0\rangle \tag{6.11}$$

Because $\mathcal{H}$ trivially commutes with itself, we may break this operator into the successive application of many operators, such that for a large number of slices N of a finite τ,

$$e^{-\tau\mathcal{H}} = \left(e^{-\frac{\tau}{N}\mathcal{H}}\right)^N \approx (1 - \delta\tau\mathcal{H})^N \tag{6.12}$$

where $\delta\tau = \tau/N$. Note that the linearized time propagator used here is both simple to implement for discrete systems as well as unbiased in the final ($\tau \to \infty$) result given some restrictions on $\delta\tau$ [231]. Thus with a prescription to stochastically apply the operator

$$\mathcal{G} = (1 - \delta\tau\mathcal{H}) \tag{6.13}$$

repeatedly to any vector in the system-time Hilbert space, we can simulate the quantum dynamics of the physical system. τ is sometimes interpreted as imaginary-time by analogy to the Wick-rotated time-dependent Schrödinger equation, however we will only refer to τ as "simulation time" here, to avoid confusion with the simultaneous presence of real and imaginary-time.

To represent a vector in the system-time Hilbert space, we introduce discrete walkers represented by $\{i, t\}$ with associated real and imaginary integer weights $\mathcal{R}(\{i, t\})$ and $\mathcal{I}(\{i, t\})$. These walkers correspond to a single system-time configuration. The in-

dices correspond to a system state $|i\rangle$ at time t with a complex integer weight defined by its real and imaginary components, $W(\{k,t\}) = \mathcal{R}(\{k,t\}) + i\mathcal{I}(\{k,t\})$. Given a collection set of these walkers, the corresponding normalized vector is given by

$$|\Psi\rangle = \frac{1}{K} \sum_{\{i,t\}} W(\{i,t\}) \, |i\rangle \otimes |t\rangle \tag{6.14}$$

where K is the normalization constant given by the sum of the absolute value of all the complex integer walker weights

$$K = \sum_{\{i,t\}} |W(\{i,t\})|. \tag{6.15}$$

We also use this notation to define matrix elements for an operator $\mathcal{O}$ between a state $|i\rangle \, |t\rangle$ and $|j\rangle \, |t'\rangle$ as

$$\mathcal{O}_{\{j,t'\},\{i,t\}} = \langle j| \, \langle t'| \, \mathcal{O} \, |i\rangle \, |t\rangle \tag{6.16}$$

To stochastically apply the operator $\mathcal{G}$ to a vector represented by a set of such walkers, the following four steps are used, adapted from the original implementation by Booth et. al:

1. Spawning: This step addresses the off-diagonal elements of $\mathcal{G}$. For each walker $\{i,t\}$, we consider $N_r = \mathcal{R}(\{i,t\})$ real parents and $N_i = \mathcal{I}(\{i,t\})$ imaginary parents, both with the correct associated sign. For each of the real parents N_i, we select a coupled state at an adjacent time and attempt to spawn a real child

and imaginary child $\{j, t'\}$ with probabilities

$$p_s^{\mathcal{R}}(\{j,t'\}|\{i,t\}) = \frac{\delta\tau \left|\mathcal{R}(\mathcal{H}_{\{j,t'\},\{i,t\}})\right|}{p_{g_t}(t',t)p_{g_s}(\{j,t'\}|\{i,t\})} \tag{6.17}$$

$$p_s^{\mathcal{I}}(\{j,t'\}|\{i,t\}) = \frac{\delta\tau \left|\mathcal{I}(\mathcal{H}_{\{j,t'\},\{i,t\}})\right|}{p_{g_t}(t',t)p_{g_s}(\{j,t'\}|\{i,t\})} \tag{6.18}$$

with corresponding signs

$$S^{\mathcal{R}} = -\text{sign}(\mathcal{R}(\mathcal{H}_{\{j,t'\},\{i,t\}})) \tag{6.19}$$

$$S^{\mathcal{I}} = -\text{sign}(\mathcal{I}(\mathcal{H}_{\{j,t'\},\{i,t\}})) \tag{6.20}$$

and for each of the imaginary parents N_i we select a coupled state at an adjacent time and attempt to spawn a real child and imaginary child $\{j, t'\}$ with probabilities

$$p_s^{\mathcal{R}}(\{j,t'\}|\{i,t\}) = \frac{\delta\tau \left|\mathcal{I}(\mathcal{H}_{\{j,t'\},\{i,t\}})\right|}{p_{g_t}(t',t)p_{g_s}(\{j,t'\}|\{i,t\})} \tag{6.21}$$

$$p_s^{\mathcal{I}}(\{j,t'\}|\{i,t\}) = \frac{\delta\tau \left|\mathcal{R}(\mathcal{H}_{\{j,t'\},\{i,t\}})\right|}{p_{g_t}(t',t)p_{g_s}(\{j,t'\}|\{i,t\})} \tag{6.22}$$

with corresponding signs

$$S^{\mathcal{R}} = \text{sign}(\mathcal{I}(\mathcal{H}_{\{j,t'\},\{i,t\}})) \tag{6.23}$$

$$S^{\mathcal{I}} = -\text{sign}(\mathcal{R}(\mathcal{H}_{\{j,t'\},\{i,t\}})) \tag{6.24}$$

where probabilities $p_s > 1$ are handling by deterministically spawning $\lfloor p_s \rfloor$ walkers and spawning an additional walker with probability $p_s - \lfloor p_s \rfloor$. $\delta\tau$ is the simulation time step and may be used to control the rate of walker spawning.

The functions $p_{g_t}(t', t)$ and $p_{g_s}(\{j, t'\}|\{i, t\})$ are the probability of suggesting a walker at the new time t' and of the particular state j respectively. For the Clock Hamiltonian, an efficient choice of the time generation function, $p_{g_t}(t', t)$ is $t' = t \pm 1$ with equal probability unless the walker is at a time boundary, in which case it should move inward with unit probability. The state generation probability $p_{g_s}(\{j, t'\}|\{i, t\})$ should be chosen based on knowledge of the structure of U_t such that connected states may reach each other. In this work we use a uniform distribution where states connected by U_t are selected randomly with equal probability, however this can be refined using knowledge of U_t.

In this case, where $\{j, t'\} \neq \{i, t\}$, the matrix elements $\mathcal{H}_{\{j,t'\},\{i,t\}}$ may be written more explicitly as

$$\mathcal{H}_{\{j,t'\},\{i,t\}} = \begin{cases} -\frac{1}{2} \langle j | U_t | i \rangle & t' = t + 1 \\ -\frac{1}{2} \langle j | U_t^\dagger | i \rangle & t' = t - 1 \\ 0 & \text{otherwise} \end{cases} \tag{6.25}$$

2. Diagonal Death/Cloning: This step addresses the application of the diagonal of $\mathcal{G}$. In this step, for each parent walker $\{i, t\}$ (not yet including child walkers spawned in the last step), calculate

$$p_d(\{i, t\}) = \delta\tau(\mathcal{H}_{\{i,t\},\{i,t\}} - S) \tag{6.26}$$

where S is a shift that is used to control the population in the simulation. Now for each real and imaginary parent N_r and N_i associated with $\{i, t\}$, if $p_d > 0$, the parent is destroyed. If $p_d < 0$, the parent is cloned with a probability $|p_d|$,

handling instances of greater than unit probabilities as in the previous step.

In the case of the Clock Hamiltonian, the diagonal matrix elements take on a simple form

$$\mathcal{H}_{\{i,t\},\{i,t\}} = \begin{cases} 1/2 + (1 - |\langle i|\Psi_0\rangle|^2) & t = 0 \\ 1/2 & t = T - 1 \\ 1 & \text{otherwise} \end{cases} \tag{6.27}$$

3. Annihilation: In this step, all previously existing and newly spawned walkers are searched, and any which match are combined such that both their real and imaginary components are summed. In the event that any walker ends up with 0 total weight, it is removed entirely from the simulation. In the case of the Clock Hamiltonian, it is advantageous to store walkers grouped by time, such that in parallel implementations the simulation can be easily split along this dimension. This will be elaborated upon later. Within each group it is advisable to use any natural ordering present on the basis states to enable binary search that can locate identical walkers in a time that is logarithmic in the number of walkers at a given time. Alternatively one can use hash tables to facilitate annihilation [28].

A single iteration of the above algorithm is cartooned in Fig. 6.1. By using this procedure, the operator G is iteratively applied until the state of walkers is equilibrated at some simulation time $\tau > \tau_{eq}$, with a number of walkers N_w. The average of

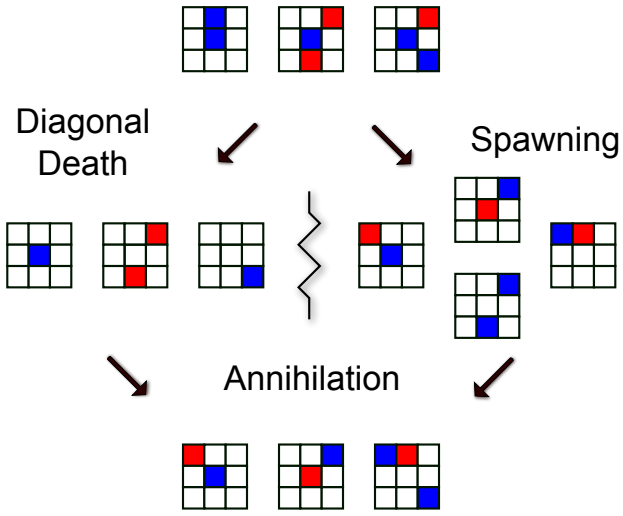

Figure 6.1: A schematic representation of a single iteration of the FCIQMC algorithm for the Clock Hamiltonian. The larger squares represent real-time, and sub-squares represent the possible quantum states occupied by either positive (blue) or negative (red) walkers. In each iteration, the set of parents spawns potential children to adjacent times, with parentage being indicated by dotted lines. Simultaneously the set of parent walkers are considered for diagonal death. Finally, the remaining set of parents and spawned children are combined, allowing walkers with opposing signs at the same state and time to annihilate.

some observable O may be estimated at simulation time τ according to

$$\langle O \rangle (\tau) = \frac{\langle \Phi(\tau) | O | \Phi(\tau) \rangle}{\langle \Phi(\tau) | \Phi(\tau) \rangle} \tag{6.28}$$

By averaging over the simulation time τ and correcting for the autocorrelation time of the quantity $\langle O \rangle$ using standard statistical procedures, the average may be converged to the desired precision. In general, however, the simulation time averaged quantity $\langle O \rangle_\tau$ may be biased due to the sign problem [128, 213, 220]. This bias may be removed to an arbitrary degree by increasing the number of walkers N_w such that

the state remains sign-coherent between steps. The number of walkers required to remove the bias to a given precision depends both on the severity of the sign problem and the amount of Hilbert space the physical problem occupies [128, 213, 220]. To this end, we define a problem-dependent number n_c such that when $N_w > n_c$, the time averaged quantity $\langle O \rangle_\tau$ is accurate to the desired precision. Because this is an NP-Complete problem, one must expect that in general, n_c is on the order of the dimension of the Hilbert space, D, that is, it grows exponentially with the size of the system and linearly with real-time. However, for many systems of interest in ground state problems it has been found that $n_c << D$ [26, 27, 213], and one might expect the same to be true for some dynamical problems. We now turn our attention to the scaling and properties of n_c for dynamical systems.

6.4 Manifestation of the sign problem

The conditions for the efficient simulation of a Hamiltonian on a classical computer have been studied in the context of quantum complexity theory. It is known that if a Hamiltonian is frustration free and has real, non-positive off-diagonal elements in a standard basis (stoquastic) that it may be probabilistically simulated to a set precision in a time that is polynomial in the size of the system [33, 34].

For practical purposes, there are limitations on the system operators one may simulate. In particular, the system operators must be the sum of a polynomial number of terms. This simply originates from the need to be able to efficiently evaluate matrix elements of a given state. The interaction of at most k particles, or $k-$local interactions, in the physical system is a sufficient but not necessary condition for this to be true. The Clock Hamiltonian construction has also been recently generalized to

open quantum systems [227], where even in this case a $2-$local construction is generally possible with the use of gadgets. Alternatively, if the Clock Hamiltonian is constructed from a sequence of unitary gates that act on at most k qubits in quantum computation, then the Clock Hamiltonian will also satisfy this requirement.

The presence of frustration in interacting systems can cause the autocorrelation time of measured observables to diverge exponentially, rendering their efficient simulation intractable even in cases where the Hamiltonian is bosonic or sign problem free [234, 245]. It has been proven generally that the Clock Hamiltonian is frustration free, with a unique ground state separated from the first excited state with a gap of $O(1/T^2)$ where T is the number of discrete time steps being considered at once.

If an operator is stoquastic (or sign problem free), then the off-diagonal elements that correspond to transitions in a Monte Carlo simulation all be non-positive. The operator $\mathcal{G}$ will contain only positive transition probabilities in this case and have a ground state corresponding to a classical probability distribution by the Perron-Frobenius theorem [33, 34]. In the context of the FCIQMC method introduced, this means that walkers will never change signs throughout the simulation, and all averages will be sign-coherent and unbiased independent of the number of walkers N_w. In the Clock Hamiltonian, the off-diagonal elements correspond to the set of unitary operators with their adjoints $\{U_t, U_t^\dagger\}$, and the penalty term $\mathcal{C}_0$. For the standard computational initial state ($|0\rangle^{\otimes N}$), the penalty term $\mathcal{C}_0$ has a fixed sign, and thus the Clock Hamiltonian is stoquastic if $\{U_t, U_t^\dagger\}$ represented in the standard basis has all real positive entries, yielding non-positive off-diagonal entries in the Clock Hamiltonian.

Given the ubiquity of $k-$local interactions in physical problems and the rigorous proof that the Clock Hamiltonian is frustration free, we will take these two conditions

as given and consider more carefully the stoquastic condition. Consider a simple example of a unitary evolution that may be simulated on a classical computer efficiently, namely reversible classical computational. All reversible classical circuits may be expressed in terms of Toffoli gate sequences, which is unitary and acts to switch a target bit conditional on the state of two control bits. In the standard computational basis it has a representation given by

$$Tof = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \end{pmatrix} \tag{6.29}$$

The Clock Hamiltonian when constructed with unitary Toffoli gates is stoquastic and $n_c \approx 1$. More explicitly, each U_t is a Toffoli gate acting on different qubits for all times t. Although a stoquastic Hamiltonian is sufficient for this to be the case, it is not a necessary condition. To see this, consider a slightly different set of unitary operators, namely the standard Pauli group gates, X_i, Y_i, and Z_i in combination with the CNOT gate. These gates have the following unitary representations in the standard

computational basis

$$X = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \tag{6.30}$$

$$Y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \tag{6.31}$$

$$Z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \tag{6.32}$$

$$CNOT = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix} \tag{6.33}$$

Considering for now only the computational basis we simulate in (a restriction we later lift), it is clear that given the complex entries and varying signs of the off-diagonals, that a Clock Hamiltonian built from this gate set will not be stoquastic if even a single Y or Z gate is used. However these gates also have the property that they map single configurations to single configurations, and as as a result no interference occurs, yielding all sign coherent averages and $n_c \approx 1$. We call this type of transformation, which is configuration number preserving, "quasi-classical", in contrast to classical which we define as configuration number preserving as well as phase preserving. Thus a stoquastic Clock Hamiltonian is a sufficient, but not necessary condition for the simulation procedure to be sign-problem free.

Consider a slightly more general local rotation R parameterized by an angle θ

$$R(\theta) = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix} \tag{6.34}$$

In this case, the value of n_c as a function of system size is more complex. These represent the real-time evolutions of local spin Hamiltonians for systems of spin-$\frac{1}{2}$ particles.

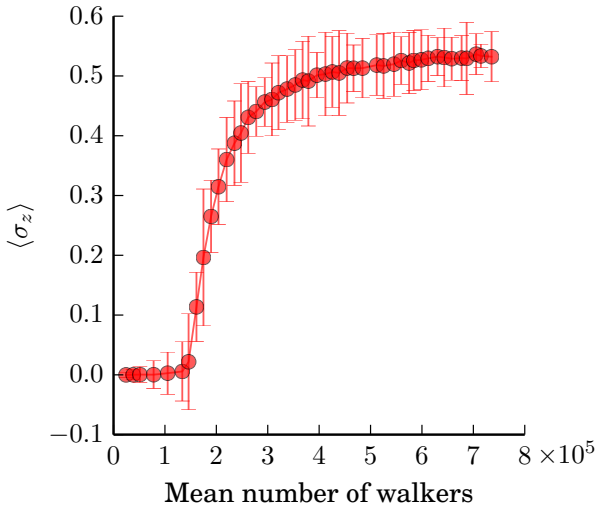

Figure 6.2: Computed expectation value for σ_z for a single qubit at the final time in the simulation as a function of the average number of walkers kept in the simulation. There are 11 total qubits in the simulation. It is apparent the system exhibits a sharp transition between a totally sign incoherent sampling where all averages become zero, and a sign coherent region where the averages begin to converge to the appropriate value.

In Fig. 6.2 we consider a single rotation $R(\theta)$ with $\theta = 5\pi/32$ applied uniformly to 11 qubits initialized to $|0\rangle^{\otimes N}$. This sequence of rotations could be applied uniformly to all qubits at once, as the individual operations trivially commute. However, maintaining the locality of the operations, that is, allowing each gate to act only on a single qubit facilitates the sampling procedure by restricting the number of connected

states for each walker to those that may be generated by local transformations on each qubit. In contrast, the application of all rotations simultaneously in principle connects each walker to 2^N states, which can make it difficult to take advantage of structure present in the specific rotations.

As the Clock Hamiltonian in this simulation is neither stoquastic nor quasi-classical, one observes a sign-incoherent region for a small number of walkers, where all averages tend towards 0, until some critical threshold $N_w > n_c$ is reached where a transition occurs to sign-coherent sampling, and the average converges to the true value. We note that some implementations of the FCIQMC algorithm have used the diagonal shift S as a proxy for convergence [29], but we did not observe a similar plateau trend here. The history state being sampled in this case is given by

$$|\Psi\rangle = \sum_t \frac{1}{\sqrt{T}} \left(R(\theta)\,|0\rangle\right)^{\otimes t} |0\rangle^{\otimes T-t} |t\rangle \tag{6.35}$$

The formal structure of this evolution is quite similar for all values of θ, however the states that result can exhibit quite different features with respect to the sign problem in sampling. In Fig. 6.3 we examine the same circuit, but include many different rotation angles. One sees that not only does the value n_c change as a function of rotation angle, but the rate of the transition is quite different as well, favoring sharper, earlier transitions for rotations that are closer to quasi-classical ($\theta = 0$).

6.5 Mitigating the sign problem

In the last section we considered the impact of sign problem as it related to local rotations (or the dynamics of distinguishable non-interacting particles). The apparent challenges in this domain are unsettling given that trivial solutions are known for this

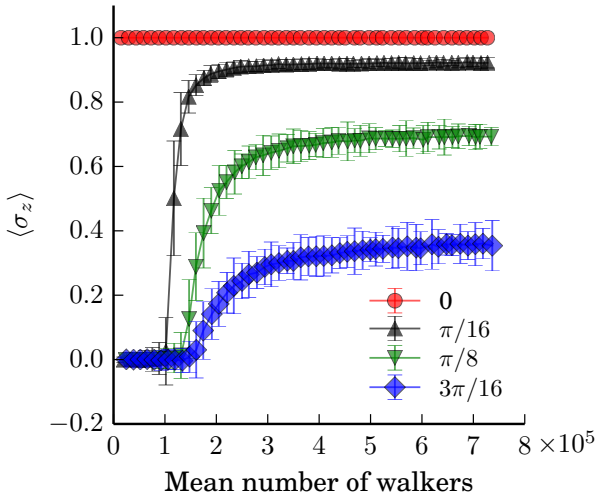

Figure 6.3: Computed expectation value for σ_z for a single qubit at the final time in the simulation as a function of the average number of walkers kept in the simulation and the rotation angle used in the simulation. The rotation angle θ is indicated by the line label. The simulation contains 11 total qubits for all rotation angles. One sees that the closer the rotation is to quasi-classical, the sharper and earlier the transition to sign coherent sampling.

problem. Here we propose a simple scheme to mitigate the sign problem to an arbitrary extent using preliminary computation.

It is known that the sign problem is generically basis dependent. To this end we propose an analogous approach to the interaction picture in quantum dynamics, where the walkers at each point in time are expressed in a new basis determined by a generic time-dependent unitary rotation given by $\{\mathcal{B}_t\}$. The evolution operators are also dressed by this change such that in the new basis, the Clock Hamiltonian is constructed from the rotated operators given by

$$U_t' = \mathcal{B}_{t+1}^\dagger U_t \mathcal{B}_t \tag{6.36}$$

Moreover, the computation of any Hermitian observable O must also take into con-

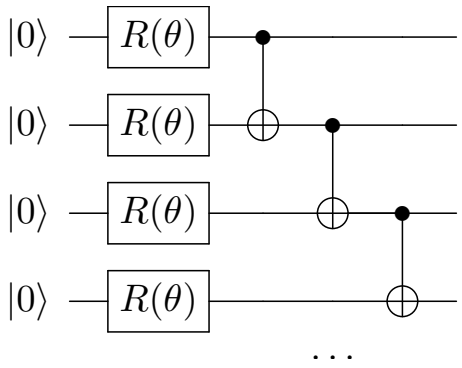

Figure 6.4: The quantum circuit diagram for the circuit used to test the efficacy of time-dependent local rotations in ameliorating the sign problem. The angle used in this case is $\theta = 0.49$. We compare the results from this circuit as a function of the number of controls that are removed from the NOT gates (crossed circle here), and whether time dependent local basis rotations are utilized. The controls are removed from the end of the circuit first.

sideration the new basis, such that

$$\langle O \rangle (\tau) = \frac{\langle \Phi(\tau)| \mathcal{B}_t^\dagger O \mathcal{B}_t |\Phi(\tau)\rangle}{\langle \Phi(\tau)|\Phi(\tau)\rangle} \tag{6.37}$$

If one finds a set of $\{\mathcal{B}_t\}$ that renders the Clock Hamiltonian stoquastic or quasi-classical, the resulting Hamiltonian may be sampled readily. One expects that in general, finding this basis must be at least as difficult as solving exactly the problem of the quantum evolution. In fact, it is easy to see that one may choose the exact evolution as the set of basis rotations, and that this renders the Clock Hamiltonian stoquastic and trivial, such that the evolution is dictated by the identity at all times. Of course the price one must pay for this is that the computation of observables requires the full evolution to be known.

However, as was seen above, it is not necessary for the Clock Hamiltonian to be rendered completely trivial. Even approaching a quasi-classical Hamiltonian in an approximate sense can greatly reduce the sampling costs. For some instances, one may

find an approximate set of rotations that make the Clock Hamiltonian nearly stoquastic or quasi-classical, and the remainder of the sign problem can be handled by maintaining a reasonable number of walkers N_w in the simulation. As an example of this procedure, we consider the simple case where $\{\mathcal{B}_t\}$ are determined entirely by the local rotations in a quantum circuit. Specifically, for local rotations, $\mathcal{B}_t = \prod_{t'<t}^{0} U_{t'}$, where $U_{t'}$ is replaced by I for two- or more qubit operations. It is clear that for circuits consisting of only local rotations, as in the previous section, this is equivalent to exact evolution and the resulting Clock Hamiltonian becomes trivial ($U'_t = I \; \forall \; t'$). To study how this works in the non-trivial case, we examine a similar circuit of local rotations, but now with a variable number of CNOT gates included. This elucidates to what extent the use of basis rotations can mitigate the sign problem when they are not an exact solution to the dynamics considered. A depiction of this circuit is given in Fig. 6.4.

In Fig. 6.5 we consider a simple quantum circuit consisting of a series of local rotations followed by NOT gates, with a variable number of the NOT gates under control. With the given rotation angle ($\theta = 0.49$), these are entangling operations not covered by the simple local basis rotation scheme we use here. However, it is seen that even for 8 controlled NOT gates, the use of local basis rotations dramatically reduces the number of walkers required to reach sign-coherent sampling, indicating this scheme can be computationally effective even for simulations containing a considerable fraction of two-qubit entangling operations. In this figure, the 4 and 8 CNOT simulations in the unrotated basis suffer similar biases due to the fact that local rotations are capable of making the sign problem difficult independent of the number of CNOT operations. The rotated basis is able to repair much of the sign problem introduced by the local rotations, but is less efficient on the 8 CNOT problem in comparison with

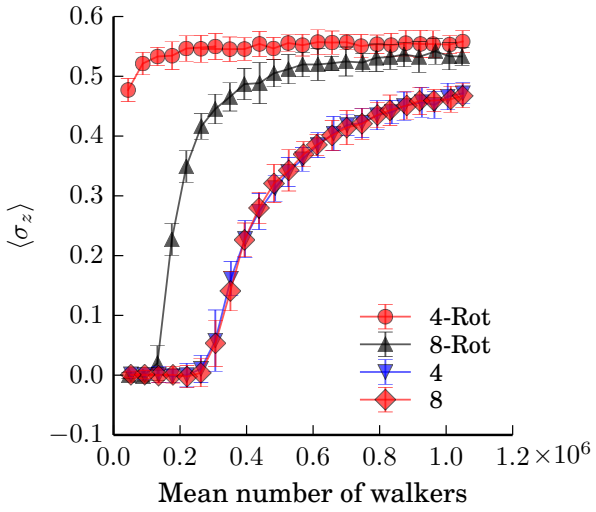

Figure 6.5: The mean value of a spin observable is plotted as a function of the mean number of walkers labeled by the number of CNOT gates both with local basis rotation(-Rot) and without. It is seen that even for a relatively high proportion of CNOT gates, the rotated basis performs far better than the non-rotated basis, requiring a lower number of walkers to reach sign-coherent sampling.

the 4 CNOT problem. Further tests of more complex quantum circuits are needed to determine the efficiency of different rotation schemes as a function of the structure of the quantum circuit.

6.6 Parallel-in-time scaling

Monte Carlo methods are often championed as the ultimate parallel algorithms, associated with the phrase "embarrassingly parallel". Given the evolution of modern computational architectures towards many-core architectures with slower clock speeds, Monte Carlo will continue to play a growing role in the numerical simulation of physics at the boundaries of our computational capabilities. Interacting walker Monte Carlo methods, can be more difficult to parallelize effectively due to the annihilation step

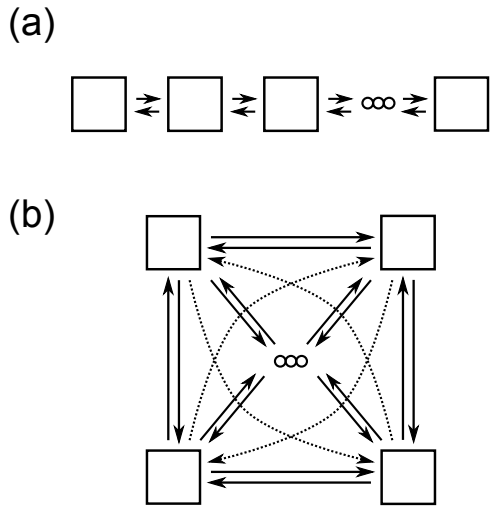

Figure 6.6: A schematic representation of the communication patterns the annihilation step of interacting walker Monte Carlo schemes, where the boxes represent different MPI processes and the ellipsis represents the rest of the processes. In the case of the Clock Hamiltonian (a), a time domain decomposition allows one to restrict communication to only nearest neighbor processes, facilitating simple, constant time communication amenable to the architecture of modern parallel computers. In the more general case (b), a clear partitioning may not be readily achievable, and all processes may need to communicate with all other processes, creating a bottleneck.

where communication of walkers is unavoidable.

In contrast to the most general interacting walker algorithm, which may require heavy communication between all processes, the FCIQMC method applied to the Clock Hamiltonian may take advantage of time-locality to create an efficient parallel-in-time algorithm using the standard method of domain decomposition in time. Using this construction, only processes containing adjacent times need to communicate their child walkers, which may be done simultaneously in a time that is constant for the number of processes involved. This remains true so long as the number of time steps under consideration is larger than the number of processes in use, which is typically

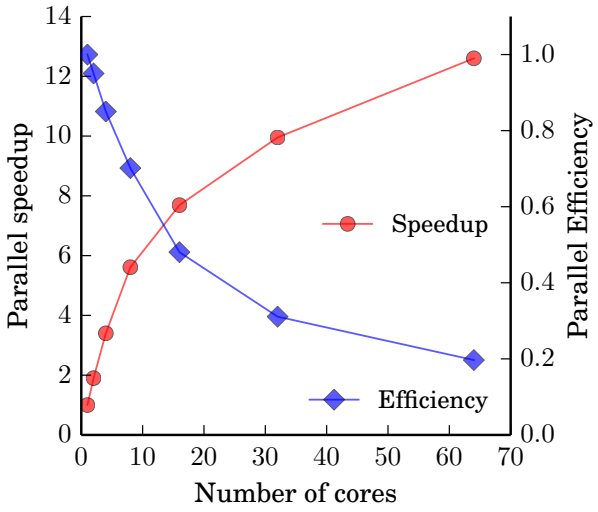

Figure 6.7: A scaling study of our method implementation with a fixed total problem size (strong scaling), showing parallel efficiencies and speedups. The simulation consisted of 11 qubits with 128 time points generated by consecutive local rotations with $\theta \approx 0.098$. The simulation maintained on average 10^6 walkers in each simulation-time step and the wall clock time was measured to the point of an equivalent number of statistical samples.

the case. In the case that the number of processes is much greater than the number of timesteps, this scheme may still be used by blocking multiple processors to each time, and utilizing an all-to-all communication pattern within each block only. The difference between these two communication patterns is highlighted in Fig. 6.6.

To demonstrate the scaling properties of this approach, we consider the scaling as a function of the number of processors for fixed total problem size, or strong scaling, with our implementation. This benchmarking is done on a standard Linux cluster composed of AMD Opteron 6376 processors. The parallel speedup with respect to single core time as a function of the number of processors is given in Fig 6.7. Here, we see that we are able to combine the parallelism of Monte Carlo with the locality of time-decomposition to achieve practical parallel efficiencies of over 95% with 2 processors and 70% with 8 using a simple MPI implementation on a commodity cluster.

6.7 Conclusions

In this work we reviewed the mapping between unitary dynamics and ground state eigenvalue problems. We then showed how the FCIQMC method, a technique originally designed to ameliorate the fermionic sign problem for ground state electronic systems, could be applied to quantum dynamics problems as a direct result of this mapping. This establishes a potential research direction for explicit connections between the fermionic and dynamical sign problems that plague quantum Monte Carlo simulations, and provides a pathway for the transfer of tools between the two domains.

The numerical consequences of the dynamical sign problem in this context were studied using a few basic quantum circuits. It was found that even local rotations can exhibit a severe sign problem depending on the form of the rotation and how different it is from a quasi-classical operation. We then introduced a general method analogous to the interaction picture in dynamics or natural orbitals in the study of eigenstates that uses basis rotations to mitigate the difficulty of the problem. The costs and benefits of different types of rotations require further research, however we showed that even local rotations can have a significant benefit for non-trivial circuits. Finally, we discussed the structure of the problem in the context of parallel-in-time dynamics, and showed high parallel efficiencies with only a basic MPI implementation on a commodity cluster.

Overall, we believe this is a promising new method for the simulation of quantum dynamics. It clarifies the bridge between dynamics and ground state problems and is capable of effectively utilizing parallel computing architectures. While we have only demonstrated it for quantum circuits, we believe it will be generally useful for the

study of quantum dynamics.

Entia non sunt multiplicanda praeter necessitatem. -

More things should not be used than are necessary.

William of Ockham

7

Compact wavefunctions from compressed imaginary time evolution*

Abstract

Simulation of quantum systems promises to deliver physical and chemical predictions for the frontiers of technology. In this work, we introduce a general and efficient black box method for many-body quantum systems using technology from compressed sensing to find compact wavefunctions without detailed knowledge of the system. No

*__Jarrod R McClean__ and Alán Aspuru-Guzik. Compact wavefunctions from compressed imaginary time evolution. *arXiv preprint arXiv:1409.7358 [physics.chem-ph]*, 2014.

knowledge is assumed in the structure of the problem other than correct particle statistics. As an application, we use this technique to compute ground state electronic wavefunctions of hydrogen fluoride and recover 98% of the basis set correlation energy or equivalently 99.996% of the total energy with 50 configurations out of a possible 10^7.

7.1 Introduction

The prediction of chemical, physical, and material properties from first principles has long been the goal of computational scientists. The Schrödinger equation contains the required information for this task, however its exact solution remains intractable for all but the smallest systems, due to the exponentially growing space in which the solutions exist. To make progress in prediction, many approximate schemes have been developed over the years that treat the problem in some small part of this exponential space. Some of the more popular methods in both chemistry and physics include Hartree-Fock, approximate density functional theory, valence bond theory, perturbation theory, coupled cluster methods, multi-configurational methods, and more recently density matrix renormalization group [13, 17, 45, 84, 102, 107, 173, 192, 217, 249].

These methods have been successful in a wide array of problems due largely to the intricate physics they compactly encode. For example, methods which are essentially exact and scale only polynomial with the size of the system have been developed for one-dimensional gapped quantum systems [134]. However such structure is not always easy to identify or even present as the size and complexity of the systems grow. For example, some biologically important transition metal compounds as well as metal

clusters lack obvious structure, and remain intractable with current methods [129, 169, 271].

The field of compressed sensing exploits a general type of structure, namely simplicity or sparsity, which has been empirically observed and is adaptive to the problem at hand. Recent developments in compressed sensing have revived the notion that Occam's razor is at work in physical systems. That is, the simplest feasible solution is often the correct one. Compressed sensing techniques have had success in quantum simulation in the context of localized wavefunctions [190] and vibrational dynamics of quantum systems [46, 257], but little has been done to exploit the possibilities for many-body eigenstates, which are critically important in the analysis and study of physical systems.

In this letter, we concisely describe a new methodology for finding compact ground state eigenfunctions for quantum systems. It is a Multicomponent Adaptive Greedy Iterative Compression (MAGIC) scheme. This method is general in that it is not restricted to a specific ansatz or type of quantum system. It operates by expanding the wavefunction with imaginary time evolution, while greedily compressing it with orthogonal matching pursuit [232]. Matching pursuit and its variants are greedy algorithms in the standard sense, that is, at each step they select a new optimal component without regard to the consequences this may have on future steps. As an example application, we choose the simplest possible ansatze for quantum chemistry, sums of non-orthogonal determinants, and demonstrate that extremely accurate solutions are possible with very compact wavefunctions. This non-orthogonal MAGIC scheme we refer to as NOMAGIC, and apply it electronic wavefunctions in quantum chemistry.

7.2 Compressed imaginary time evolution

Beginning with general quantum systems, an N-particle eigenfunction of a quantum Hamiltonian H, $|\Psi\rangle$, may be approximated by a trial function $|\tilde{\Psi}\rangle$ that is the sum of many-particle component functions $|\Phi^i\rangle$, such that

$$|\tilde{\Psi}\rangle = \sum_i^{N_c} c_i \, |\Phi^i\rangle \tag{7.1}$$

where N_c is the total number of configurations in the sum and no relation need be assumed between $|\Phi^i\rangle$ and $|\Phi^j\rangle$ for $i \neq j$. A simple example of such a component function for a general quantum system is the tensor product of N single particle functions $|\phi_j^i\rangle$

$$|\Phi^i\rangle = |\phi_0^i\rangle \, |\phi_1^i\rangle \, ... \, |\phi_{N-1}^i\rangle \tag{7.2}$$

and we will consider its anti-symmetric counterpart in applications to electronic systems. In this work, we define a state to be simple, sparse, or compact if the number of configurations, N_c, required to represent a state is much smaller than the total dimension of the Hilbert space.

One method for determining $|\tilde{\Psi}\rangle$ is a direct variational approach based on the particular parametrization of $|\Phi^i\rangle$ and choice of N_c. This approach can be plagued by issues related to the choice of initial states, difficulty of adding new states, and numerical instability of the optimization procedure if proper regularization is not applied [66, 88, 92, 125, 127].

We present an alternative technique that selects the important configurations in a black-box manner and is naturally regularized to prevent numerical instability. It

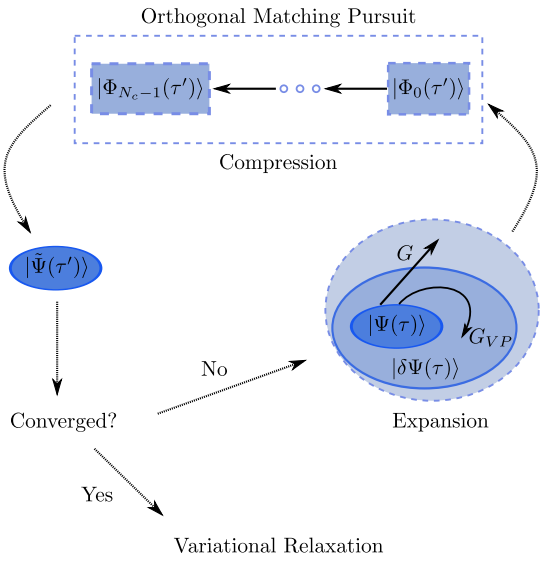

Figure 7.1: A schematic diagram of the MAGIC approach. At each iteration the wavefunction is expanded by means of the imaginary time propagator G, and subsequently compressed with orthogonal matching pursuit. The imaginary time propagator corresponding to projection into the manifold at $|\Psi(\tau)\rangle$, denoted $|\delta\Psi(\tau)\rangle$, typically prescribed by differential time dependent variational principles is given as G_{VP} and depicted to emphasize that expansion with the operator G can explore a greater part of Hilbert space. The compression is performed simultaneously with expansion in our implementation to prevent rapid growth of the wavefunction. These steps are iterated until convergence at a specified maximum number of component functions, at which point an optional variational relaxation may be performed.

is built through a combination of imaginary-time evolution and compressed sensing. Imaginary-time evolution can be concisely described as follows. Given a quantum system with a time-independent Hamiltonian H and associated eigenvectors $\{|\chi^i\rangle\}$, any state of the system $|\Omega\rangle$ may be expressed in terms of those eigenvectors as

$$|\Omega\rangle = \sum_i c_i |\chi^i\rangle \tag{7.3}$$

and the the evolution of the system for imaginary-time τ is given as

$$G\left|\Omega\right\rangle = e^{-H\tau}\left|\Omega\right\rangle = \sum_i c_i e^{-E_i\tau}\left|\chi^i\right\rangle \tag{7.4}$$

where $E_0 < E_1 \leq E_2 ... \leq E_{N-1}$ are the eigenenergies associated with $|\chi_i\rangle$. By evolving and renormalizing, eventually one is left with only the eigenvector associated with the lowest eigenvalue, or ground state. Excited states may be obtained with a number of approaches including spectral transformations (e.g. $H' = (H-\lambda)^2$ [154]), matrix deflation, or other techniques. However we will concern ourselves only with ground states in this work.

Imaginary time evolution approaches may be broadly grouped into two classes. The first class involves the explicit application of the imaginary-time propagator G to the wavefunction. This approach typically generates many configurations at every step, causing a rapid expansion in the size of the wavefunction. As a result, these methods have almost exclusively been restricted to Monte Carlo sampling procedures which attempt to assuage this explosion by stochastically sampling or selecting the most important configurations [29, 139], however the recently developed imaginary time-evolving block decimation also belongs to this class, performing truncations after expansion along a virtual bond dimension [52, 94, 208, 238].

The second class of imaginary-time approaches follow the evolution dictated by the action of G projected onto the manifold spanned by linear variations in the function at the previous time step, sometimes referred to as Galerkin or time-dependent variational methods including imaginary time MCTDH [19, 181] and DMRG in some limits [94]. While computationally convenient, it is often unclear how projection onto the original linear subspace at every time can affect evolution with respect to the ex-

act evolution. In this work, we show that the first class of explicit evolution can be used on any ansatz without configuration explosions or stochastic sampling by utilizing a technique from the field of compressed sensing, namely orthogonal matching pursuit [193, 232].

The algorithm we use is diagrammed in Fig 7.1, and proceeds iteratively as follows. The wavefunction at time $\tau = 0$, $|\Psi(\tau)\rangle$, may be any trial wavefunction that is not orthogonal to the desired eigenstate. We determine the wavefunction at time $\tau + d\tau = \tau'$ greedily, fitting one configuration $|\Phi^i(\tau')\rangle$ at a time by maximizing the functional

$$\frac{\left|\langle\Phi^i(\tau')|\, G\, |\Psi(\tau)\rangle - \sum_{j<i} c_j(\tau')\langle\Phi^i(\tau')|\Phi^j(\tau')\rangle\right|}{\sqrt{\langle\Phi^i(\tau')|\Phi^i(\tau')\rangle}} \tag{7.5}$$

with respect to the parameters that determine $|\Phi^i(\tau')\rangle$. Such that after k iterations, the wavefunction is given by

$$|\tilde{\Psi}(\tau)\rangle = \sum_i^k c_i(\tau)\, |\Phi^i(\tau)\rangle \tag{7.6}$$

The coefficients in this expansion, $c_i(\tau')$ are solved for simultaneously after each iteration by orthogonal projection, which after simplification reduces to the following linear system for the coefficient vector c

$$Sc = v \tag{7.7}$$

where $S_{ij} = \langle\Phi^i(\tau')|\Phi^j(\tau')\rangle$ and $v_i = \langle\Phi^i(\tau')|\, G\, |\Psi(\tau)\rangle$. Together, the fit and orthogonal projection step is equivalent to orthogonal matching pursuit [193, 232] applied to the signal generated by the imaginary time evolution of the state at each time step $G\,|\Psi(\tau)\rangle$. The expansion-compression procedure is advanced to the next imaginary

time step either when some accuracy convergence criteria is met, or when some preset maximum number of components N_c is reached, and the total simulation is terminated when the wavefunction converges between imaginary-time steps. We provide additional details of the numerical procedure in the supporting information for interested readers.

Note that one is free to choose a convenient form for the propagator G. In this work we use the linearized propagator $G \approx (I - d\tau(H - \lambda))$, which is both easy to implement and provably free of bias in the final result for finite single particle basis sets given some restrictions on $d\tau$ [231]. The constant shift λ can be adjusted and is taken to be the expectation value of the energy of the previous imaginary time step in our implementation.

Orthogonal matching pursuit attempts to find the sparsest solution to the problem of state reconstruction [179, 232], and thus is ideal for keeping the number of configurations minimal throughout the imaginary time evolution. However, while the solution is sparsest in the limit of total reconstruction and naturally regularized against configuration collinearity, for very severe truncations of the wavefunction, the sparsifying conditions generate a solution which is not variationally optimal for the given number of configurations. For this reason, we finish the computation with a total variational relaxation of the expectation value of the energy with respect to both coefficients and states. This retains both the benefits of imaginary time evolution in avoiding local energetic minima and of variational optimality in the final solution.

7.3 Application to chemical systems

The method we have outlined may be readily applied to any quantum system, such as spins or oscillators, however as a first application we consider ground-state electronic wavefunctions of molecules. We will take the approach that is conventional to the field of quantum chemistry, and solve the problem in a basis of Gaussian-type functions [102]. After a basis has been selected, there is a standard procedure of expanding the linear state space by excitation known as configuration interaction (CI), which can eventually yield the numerically exact solution within a basis when the full state space has been covered. This is referred to as full configuration interaction (FCI) and is the standard to which we compare. Moreover, we compare to truncated orthogonal CI methods that represent a high level of accuracy while yielding an explicit wavefunctions and requiring no additional machinery to evaluate the energy efficiently [102, 267]. Comparison to methods considering explicit correlation beyond that covered by a specific traditional Gaussian basis, such as explicitly correlated f_{12} type wavefunctions, are not yet within the scope of this work.

In the context of our approach, the indistinguishability of electrons necessitates handling of anti-symmetry. The simplest way to include anti-symmetry into the wavefunction is by utilizing anti-symmetric component tensors $|\Phi^i\rangle$. The most common anti-symmetric component function is the Slater determinant, such that we express the wavefunction as

$$|\Psi\rangle = \sum_i^{N_c} c_i \, |\Phi^i\rangle \tag{7.8}$$

where $|\Phi^i\rangle$ are Slater determinants with no fixed relations between $|\Phi^i\rangle$ and $|\Phi^j\rangle$ for

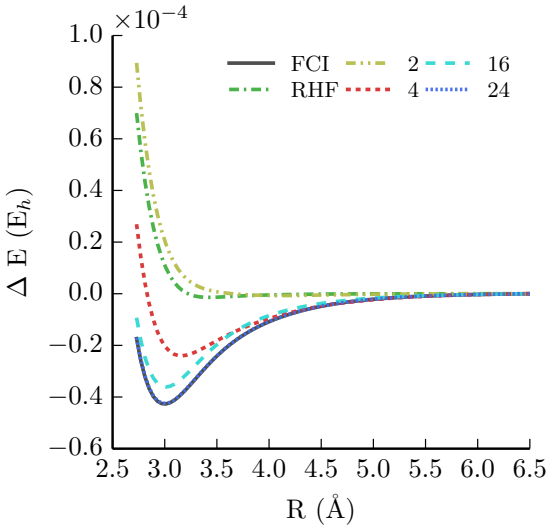

Figure 7.2: The bond dissociation curve of the helium dimer in the aug-cc-pVDZ basis showing rapid and consistent convergence in the number of non-orthogonal determinants. These represent the nuclear union curves constructed from a number of local determinants at each nuclear point given by the line label, and are sampled at a spacing of 0.04 Å. The curves have been shifted by the tail values in order to allow comparison of the features for this sensitive bond, and the 24 determinant curve is indistinguishable from the FCI solution in the graphic. At a point near the equilibrium geometry, $R = 3.01$ Å, the 24 determinant curve with the nuclear union configuration interaction technique recovers 99.9899% of the basis set correlation energy, or equivalently 99.9999% of the total energy.

$i \neq j$. While this simple form lacks extensivity [221], it is attractive for other reasons. Namely the quality of description and rate of convergence in N_c are invariant to invertible local transformations of the state (i.e. atomic orbitals vs. natural orbitals) [92], and the mathematical machinery related to the use and extension of such a wavefunction is already well developed [3, 101, 150, 219]

While the method we use for determinant selection is unique, non-orthogonal Slater determinants have been used successfully in valence bond theory [84, 107] as well as more recent symmetry breaking and projection methods [39, 110]. Unconstrained non-orthogonal Slater determinants have been utilized before, but in a purely variational context [88, 125]. Using this machinery yields explicit gradients that we utilize

in the optimization of determinants [219]. The scaling of these constructions with current algorithms is $O({N_c}^2 \max(M^2, N_e^3))$ [221] where N_c is the number of determinants and N_e is the number of electrons, however development of approximations in this area have received comparatively less attention with respect to orthogonal reference wavefunction methods and there may be ways to improve upon this scaling.

We introduce an additional enhancement for the study of chemical reactions that is greatly facilitated by the compactness of our expansions. Namely, when considering a full reaction coordinate, such as that for a bond dissociation, we perform an additional linear variational calculation in the space of components (determinants) found locally at neighbouring nuclear configurations. As a proof of principle, we include configurations from the entire curves in the following examples, but more economical truncations can be used as well. We refer to this additional step, as the nuclear union configuration interaction method and describe it in more depth in the supplemental information.

As a first application, we consider He_2 in the aug-cc-pVDZ basis [255], which is a standard atom-centered Gaussian quantum chemical basis with additional diffuse functions to better capture the weak dispersive interactions present in rare gas interactions. The helium dimer is unbound in the case of a single determinant with restricted Hartree Fock (RHF) and is not held together by a covalent bond, but rather dispersive forces and dynamical correlation only. In Fig. 7.2, we consider the dissociation of this molecule under different numbers of non-orthogonal determinants. Despite the sensitive nature of this bond, it is qualitatively captured with as few as 4 local determinants and quantitatively captured with as few as 24 determinants. The dimension of the space of this molecule is on the order of 10^4 when reduced by considerations of point group symmetry. The NOMAGIC approach does not yet utilize

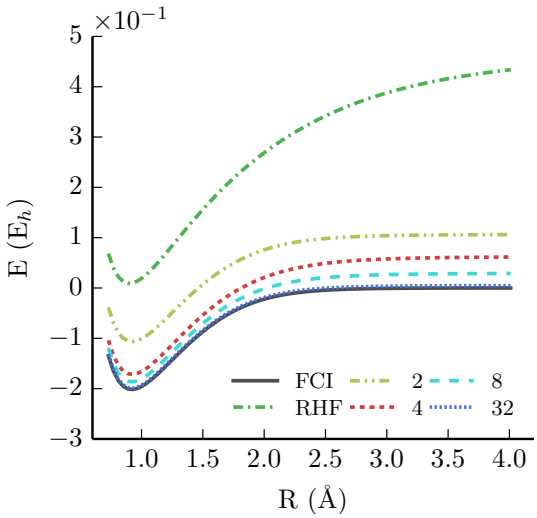

Figure 7.3: The bond dissociation curve of hydrogen fluoride in the cc-pVDZ basis showing rapid convergence in the number of non-orthogonal determinants. These are the nuclear union curves constructed from a number of local determinants at each nuclear point given by the line label, and are sampled at a spacing of 0.04 Å. The 32 determinant curve is nearly indistinguishable from the FCI curve in this graphic. At a point near the equilibrium geometry, $R = 0.93$ Å, the 32 determinant curve with the nuclear union configuration interaction technique recovers 98.6% of the basis set correlation energy, or equivalently 99.997% of the total energy.

any symmetry other than the spin symmetry enforced by the parameterization of the wavefunction.

As a second example, the dissociation of hydrogen fluoride in a cc-pVDZ basis [63] is studied. The total configuration space for this molecule is on the order of 10^7 and it involves a homolytic bond breaking of a covalent bond in the gas phase. Considering the results in Fig. 7.3, one can see that while RHF yields an unphysical dissociation solution, as few as 2 determinants are sufficient to fix the solution in a qualitative sense. Beyond this, the addition of more determinants represents a monotonically increasing degree of accuracy, with rapid convergence to a quantitative approximation by 32 determinants.

In Fig. 7.4 we select two points on the HF dissociation curve, and study the con-

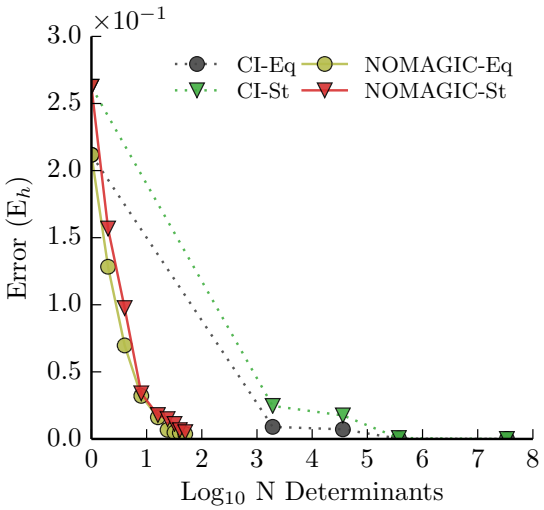

Figure 7.4: A curve of the energetic error with respect to full CI for HF bond dissociation in the cc-pVDZ basis as a function of the log of the number of determinants included for both a near-equilibrium geometry "Eq" with $R = 0.93$ Å and a stretched geometry "St" with $R = 1.73$ Å. The configuration interaction energies are generated by a standard excitation sequence from the Hartree-Fock solution, CIX(X=SD, SDT, SDTQ) followed by FCI. The number of determinants used in the full CI expansion is approximately 34 million taking into account molecular point group symmetries, or 135 million without. No symmetries other than spin are utilized in the NOMAGIC calculation.

vergence of the energy as a function of the number of determinants in the NOMAGIC method and a traditional CI expansion with the canonical Hartree-Fock orbitals. In particular, we study both a point near the equilibrium bond length ($R = 0.93$ Å) where traditional CI expansions perform relatively well and a more stretched geometry ($R = 1.73$ Å) where traditional CI expansions perform more poorly. We see that in both cases, if one considers a fixed level of accuracy in the energy, the NOMAGIC method is considerably more compact. For example, to achieve a level of accuracy superior to the CISDT expansion that uses 36021 determinants, NOMAGIC requires only 24 determinants at both geometries. That is, for equivalent accuracy, the NOMAGIC wavefunction is roughly 1500 times more compact in the space of Slater de-

terminants. By 50 determinants out of a possible 10^7 in the NOMAGIC wavefunction, we recover 98% of the basis set correlation energy or equivalently 99.996% of the total energy. To further quantify the advantage of computationally for manipulating and evaluating expectations values on compact NOMAGIC states, consider the cost of evaluating the energy of a stored CISDT state and a NOMAGIC state with N_c determinants. A NOMAGIC state with N_c determinants requires $O(N_c^2 \max(M^2, N_e^3))$ operations to evaluate the energy. In contrast, a CISDT state requires $O(M^8)$ operations in the standard case that the number of basis functions M is on the same order as the number of electrons N_e. Thus assuming $N_e \sim M$ there is a clear computational advantage in any case where $N_c < O(M^{5/2})$ provides a sufficiently accurate representation, as has been observed in all examples thus far.

7.4 Conclusions

In this work, we introduced a general method to find compact representations of quantum eigenfunctions by using imaginary-time evolution and compression. The method assumes no specific structure in the problem and can be applied to any quantum system with a variety of ansatze. The compact wavefunctions that result from this methodology not only offer potential practical advantages in terms of storage and subsequent evaluation of physical observables, but can also provide a numerical upper-bound on the minimal information necessary to identify a physical quantum state, or Kolmogorov complexity of the state. A low Kolmogorov complexity of physical quantum states would have important ramifications for the growing belief that physical states occupy a small physical "corner of Hilbert space" [65, 79, 199]. We demonstrated the method's practical success in some quantum chemical systems with

a small number of non-orthogonal Slater determinants. One practical application of this method is the creation of extremely compact trial wavefunctions for quantum Monte Carlo, which suffer little additional overhead with the use of non-orthogonal determinants and are often limited by the size of the trial wavefunction [4, 83]. Finally, we believe that extensions to this method using ansatze that contain system specific physics have the potential to be even more compact and this is subject of current research.

7.5 Supplemental Information

7.5.1 Orthogonal Matching Pursuit

In this section, we offer some additional details on the implementation of Orthogonal Matching Pursuit [232] with imaginary-time evolution in quantum systems. Given a quantum state $|\Psi(\tau')\rangle$ that one wishes to reconstruct, orthogonal matching pursuit is a greedy decomposition algorithm (i.e. one that selects the best component at each iteration without regard to future iterations) that approximates the sparse problem of finding $|\tilde{\Psi}(\tau')\rangle$ such that

$$\begin{aligned} &\min_{|\tilde{\Psi}(\tau')\rangle} || \, |\Psi(\tau')\rangle - |\tilde{\Psi}(\tau')\rangle \, ||_2^2 \\ &\text{subject to } || \, |\tilde{\Psi}(\tau')\rangle \, ||_0 < N \end{aligned} \tag{7.9}$$

This is done by considering an overcomplete dictionary $\{|\Phi^i(\tau')\rangle\}$ that can express $|\tilde{\Psi}\rangle$ as

$$|\tilde{\Psi}(\tau')\rangle = \sum_i c_i(\tau')\,|\Phi^i(\tau')\rangle \tag{7.10}$$

and at each stage selecting selecting the $|\Phi^i(\tau')\rangle$ which maximizes the overlap with the residual with respect to the target signal $|\Psi(\tau')\rangle$,

$$\max_{|\Phi^i\rangle(\tau')} \frac{|\langle\Phi^i(\tau')|\Psi(\tau')\rangle - \sum_{j<i} c_j(\tau')\langle\Phi^i(\tau')|\Phi^j(\tau')\rangle|}{\sqrt{\langle\Phi^i(\tau')|\Phi^i(\tau')\rangle}} \tag{7.11}$$

In practice for quantum systems, the dictionary $\{|\Phi^i(\tau')\rangle\}$ can be any overcomplete basis for the N-particle Hilbert space, and the location of the optimal $|\Phi^i(\tau')\rangle$ can be done with a few different methods such as discrete enumeration of all basis states, stochastic search, and direct non-linear optimization. While discrete enumeration is commonly used in the orthogonal matching pursuit literature, the high dimensional nature of quantum systems does not readily allow it. Among the remaining options, we find that direct non-linear optimization is superior to stochastic search methods for the problems we considered. Specifically, we utilized a quasi-Newton BFGS procedure with analytic gradients and inexact line search satisfying the strong Wolfe conditions [185]. The implementation closely follows that discussed in the classic text by Nocedal and Wright with a modified Cholesky regularization to protect again instabilities in the approximate Hessian.

After selection of the optimal $|\Phi^i(\tau')\rangle$, the full set of coefficients $\{c_i(\tau')\}_{i=0}^{j}$ are re-determined by orthogonal projection of the selected basis functions on the signal

$|\Psi(\tau')\rangle$. This is equivalent to solving the linear equation

$$Sc = v \tag{7.12}$$

for the coefficient vector c, where $S_{ij} = \langle\Phi^i(\tau')|\Phi^j(\tau')\rangle$, $v_i = \langle\Phi^i(\tau')|\Psi(\tau')\rangle$, and $c_i = c_i(\tau')$.

Throughout this procedure, one also has a choice of how to represent the target signal $|\Psi(\tau')\rangle$. In some cases, it is feasible to construct $|\Psi(\tau')\rangle$ explicitly from a previous time step and imaginary time propagator G, and doing so could potentially facilitate the optimization procedure by examining properties of the state. However, exact expansion of the state $|\Psi(\tau')\rangle$ using G can have many terms for even modestly sized quantum systems, negating the potential benefits of compressing the wavefunction. In practice, we found that a much better approach is to directly with $G|\Psi(\tau)\rangle$ without first expanding the wavefunction explicitly. When using the linearized propagator $G(\lambda) \approx (I - d\tau(H - \lambda))$, this means that Hamiltonian and overlap matrix elements and their derivatives are sufficient for the implementation of the procedure.

In principle, at any time step, one may continue to add elements $|\Phi^i(\tau')\rangle$ until an arbitrary convergence tolerance is reached, i.e. $||\,|\Psi(\tau')\rangle - |\tilde{\Psi}(\tau')\rangle\,||_2 < \epsilon$ for some $\epsilon > 0$. However, as only the final state in the large τ limit is of interest, and any state that is not completely orthogonal to this state will eventually converge to it, some errors in intermediate steps are permissible. Thus a more economical approach, is to terminate the addition of states $|\Phi^i\rangle$ at intermediate time steps according to some proxy, such as sufficient decrease in the energy $\tilde{E}(\tau') = \langle\tilde{\Psi}(\tau')|\,H\,|\tilde{\Psi}(\tau')\rangle$ from the previous time step.

7.5.2 NOMAGIC Algorithm

In this section we detail our implementation of the NOMAGIC procedure, including how values of algorithmic parameters may be chosen. Consider a physical system defined by the Hamiltonian H with a number of particles N, each represented on a discrete basis of M functions. The NOMAGIC algorithm begins by selecting an initial component $|\tilde{\Psi}(0)\rangle = |\Phi^0(0)\rangle$ from some approximation procedure. In this work we utilize a mean-field procedure that minimizes the energy, namely the Hartree-Fock algorithm [102], to find an initial state, however other procedures such as a tensor cross approximation may be used [127]. Using the selected initial component, the initial expectation value of energy is computed as

$$\tilde{E}(0) = \langle \tilde{\Psi}(0) | \, H \, | \tilde{\Psi}(0) \rangle \, . \tag{7.13}$$

Additionally, this initial component is also used to estimate a safe value for the imaginary timestep τ. The bound on τ that guarantees a correct final state [231] under exact propagation is given by

$$d\tau \leq \frac{2}{E_{\max} - \lambda} \tag{7.14}$$

where $E_{\max}$ is the maximum eigenvalue of H. However, one does not expect to know the eigenvalues in advance, and moreover the problem is changed by the fact that propagation may be performed to some finite, economical precision. As a result, we use a fast estimate to determine an approximate suitable timestep $d\tau$. This is done by constructing the corresponding mean-field Hamiltonian H_{mf} from the original Hamiltonian and component function by performing a partial trace on all but a tar-

get particle. In the distinguishable case, this will result in N uncoupled Hamiltonians (one for each particle) and in the indistinguishable case, a single mean-field Hamiltonian, which is the Fock matrix for electronic systems. The maximum eigenvalue of the mean-field Hamiltonian is easily found, and we define $\Delta_{mf} = E^{mf}_{\max} - E(0)$, with the constant shift set to the expectation value of the energy. From this, the value of the timestep is set to

$$d\tau = \frac{1.8}{\Delta_{mf}}. \tag{7.15}$$

We note that more efficient and adaptive schemes are possible for timestep selection in problems of this type, however these were not utilized in the current work.

Once the timestep has been selected, the algorithm propagates forward through imaginary time with the orthogonal matching pursuit algorithm, with a termination threshold based on the energy. For computational practicality, we thus set several threshold values. The maximum number of component functions allows, $N_{\text{c-max}}$, the maximum imaginary time $\tau_{\max}$, and the minimal improvement in energy ϵ_E. We also introduce the function N that counts the number of components present in a wavefunction. This procedure is detailed in Alg. 1.

Note that the update for the threshold value $\epsilon_E \leftarrow \max(\Delta/(ed\tau)), 10^{-7}$, where e is Euler's number, is based on a heuristic that if perfect evolution was achieved, the energy would decay to the ground state exponentially in imaginary time. At the termination of this imaginary time procedure, a final variational relaxation is performed on the wavefunction $|\tilde{\Psi}(\tau)\rangle$ with respect to both component functions $|\Phi_i(\tau)\rangle$ and coefficients $c_i(\tau)$ to relax the greedy constraint on the fitting procedure. It is known that direct minimization of canonical tensor decompositions can suffer from numerical

$\epsilon_E \leftarrow 0$
while $\tau < \tau_{\text{max}}$ **do**
  $\tau' \leftarrow \tau + d\tau$
  $i \leftarrow 0$
  while $i < N_{\text{c-max}}$ **do**
    Find $|\Phi^i(\tau')\rangle$ satisfying eq 7.11 on state $G(\lambda)\,|\tilde{\Psi}(\tau)\rangle$
    Determine $c_i(\tau')$ via eq 7.12
    $|\tilde{\Psi}(\tau')\rangle_i \leftarrow \sum_j^i c_j\,|\Phi^j(\tau')\rangle$
    Calculate $\Delta = (\langle\tilde{\Psi}(\tau')|_i\,H\,|\tilde{\Psi}(\tau')\rangle_i - E(\tau))$
    if $\Delta < -d\tau\epsilon_E$ **then** break
    end if
    $i \leftarrow i + 1$
  end while
  if $i = N_{\text{c-max}}$ and $\Delta > -d\tau\epsilon_E$ **then** break
  else if $i > N(|\tilde{\Psi}(\tau)\rangle)$ **then** $\epsilon_E \leftarrow \max(\Delta/(ed\tau), 10^{-7})$
  end if
  $\lambda \leftarrow \langle\tilde{\Psi}(\tau')|\,H\,|\tilde{\Psi}(\tau')\rangle$
  $\tau \leftarrow \tau + d\tau$
end while

ALGORITHM 7.1: NOMAGIC Algorithm

issues if care is not taken to constraint the length of the individual components [127]. In particular, the space of canonical rank$-k$ decompositions is not closed, however the addition of a constraint on the norm of components remedies this situation [66]. In practice, we find a loose penalty term sufficient to enforce this constraint and mitigate the potential numerical difficulties from this problem without introducing the complexities of sophisticated constrained optimizations. Specifically we variationally minimize the auxiliary functional

$$\mathcal{L} = \frac{\langle\tilde{\Psi}(\tau)|\, H\, |\tilde{\Psi}(\tau)\rangle - \gamma(\max(0, \sum_i \langle\Phi^i(\tau)|\Phi^i(\tau)\rangle - D))^2}{\langle\tilde{\Psi}(\tau)|\tilde{\Psi}(\tau)\rangle} \tag{7.16}$$

where D controls the maximum length of components and γ is the penalty parameter. In this work we choose $D = 4.0$ and $\gamma = 1.0$, however little dependence is observed in the final result on these parameters unless extreme values are taken. Note that despite the presence of the max() function, this penalty term is differentiable and introduces no substantial additional difficulty in implementation.

7.5.3 Electronic Wavefunction Parameterization

Here we detail the electronic wavefunction parametrization used in this work, as well as the expressions used for the implementation of orthogonal matching pursuit and variational relaxation in electronic systems.

In quantum chemistry, frequently one first chooses a suitable single particle spin-orbital basis for the description of the electrons, which we denote $\{|\phi_i\rangle\}$. This basis typically consists of atom-centered contracted Gaussian type functions with a spin function, and are in general non-orthogonal such that they have an overlap matrix

defined by

$$S_{ij} = \langle \phi_i | \phi_j \rangle \tag{7.17}$$

Linear combinations of these atomic orbitals are used to form molecular orbital functions

$$|\chi_m\rangle = \sum_i c_m^i \, |\phi_i\rangle \tag{7.18}$$

which have an inner product

$$\langle \chi_m | \chi_n \rangle = \sum_{i,j} c_m^{i*} c_n^j \langle \phi_i | \phi_j \rangle = \sum_i c_m^{i*} c_n^j S_{ij} \tag{7.19}$$

In our implementation, the $N-$electron component wavefunctions may be formed from the anti-symmetrized $N-$fold product of molecular orbital functions, also known as Slater determinants.

$$|\Phi^k\rangle = \mathcal{A}\left(|\chi_0^k\rangle \, |\chi_1^k\rangle \, ... \, |\chi_{N-1}^k\rangle\right) \tag{7.20}$$

where $\mathcal{A}$ is the anti-symmetrizing operator. A convenient computational representation of an anti-symmetric component function $|\Phi^k\rangle$ is given by the coefficient matrix

$$T^K = \left(c_0^K | c_1^K | ... | c_{N-1}^K\right) \tag{7.21}$$

which denotes an $M \times N$ matrix whose m'th column are the coefficients defining the m'th molecular orbital $|\chi_m^k\rangle$. This yields a convenient construction for the overlap

between two component functions

$$\langle \Phi^K | \Phi^L \rangle = M_{KL} = \det\left(V_{KL}\right) = \det\left(T^{K\dagger} S T^L\right) \tag{7.22}$$

One quantity of convenience is the so-called transition density matrix defined between determinants K and L

$$P^{KL} = T^K \left(T^{L\dagger} S T^K\right)^{-1} T^{L\dagger} \tag{7.23}$$

Hamiltonian matrix elements may be written as

$$H_{KL} = M_{KL} \left(\mathrm{Tr}\left[P^{KL}\hat{h}\right] + \frac{1}{2}\,\mathrm{Tr}\left[P^{KL}G^{KL}\right]\right) \tag{7.24}$$

where $\hat{h}$ are the single electron integrals,

$$h_{\mu\nu} = \int d\sigma\ \phi^*_\mu(\sigma) \left(-\frac{\nabla_r^2}{2} - \sum_i \frac{Z_i}{|R_i - r|}\right) \phi_\nu(\sigma) \tag{7.25}$$

$$\tag{7.26}$$

where $\sigma = (r, s)$ denotes electronic spatial and spin variables and the nuclear positions and charges are R_i and Z_i. G^{KL} is given by

$$G^{KL}_{\mu\nu} = \left(\sum_{\lambda\sigma} P^{KL}_{\lambda\sigma}(g_{\mu\nu\lambda\sigma} - g_{\mu\lambda\nu\sigma})\right) \tag{7.27}$$

with the two electron integrals $g_{\mu\nu\lambda\sigma}$

$$g_{\mu\nu\lambda\sigma} = \int d\sigma_1\ d\sigma_2\ \frac{\phi^*_\mu(\sigma_1)\phi_\nu(\sigma_1)\phi^*_\lambda(\sigma_2)\phi_\sigma(\sigma_2)}{|r_1 - r_2|} \tag{7.28}$$

From the description of orthogonal matching pursuit, we see that to utilize non-linear optimization of the component functions $|\Phi^k\rangle$ with analytic first derivatives, one needs the variations of H_{KL} and M_{KL} with respect to T^K. Allowing variations δT^K, the required expressions in the non-orthogonal spin orbital basis are as follows:

$$\delta M_{KL} = M_{KL} \operatorname{Tr}\left[S T^L (V^{KL})^{-1} \delta T^{K\dagger}\right] \tag{7.29}$$

$$\delta P^{KL} = [1 - P^{KL} S] \delta T^K (V^{KL\dagger})^{-1} T^{L\dagger} \tag{7.30}$$

$$\delta G^{KL}_{\mu\nu} = \left(\sum_{\lambda\sigma} \delta P^{KL}_{\lambda\sigma} (g_{\mu\nu\lambda\sigma} - g_{\mu\lambda\nu\sigma}) \right) \tag{7.31}$$

$$A_{KL} = \operatorname{Tr}\left[P^{KL} G^{KL}\right] \tag{7.32}$$

$$\delta A_{KL} = \operatorname{Tr}\left[(1 - S P^{KL\dagger}) G^{KL\dagger} T^L (V^{KL})^{-1} \delta T^{K\dagger}\right] \tag{7.33}$$

One must take care in implementing this expression, as it is a special case of the adjugate relations that is only strictly valid when V^{KL} is non-singular. To use this expression in evaluating cases when V^{KL} is singular, techniques developed elsewhere utilizing the singular value decomposition of V^{KL} and exact interpolation can be used [3]. Note also that numerical simplifications are possible by explicitly considering spin (α, β) and noting that $T^K = T^{K\alpha} \oplus T^{K\beta}$. These reductions of the above equations are straightforward and we do not give them here.

7.5.4 Renormalization of determinants

The form of the functional used in all optimizations formally ensures their values are independent of total normalization of the wavefunction and normalization of individual columns defining the single particle portions of the wavefunctions. While this is true in exact arithmetic, there can be practical numerical issues if these values are

allowed to become unbounded throughout the course of the simulation. For this reason, it is convenient to occasionally renormalize single particle functions as well as the total wavefunction. An efficient way to perform the renormalization at the level of a single determinant T with corresponding coefficient c defined on a non-orthogonal single particle basis with overlap matrix S is

$$\begin{aligned} Q, R &= \text{QRDecomp}(S^{1/2}T) \\ T' &= S^{-1/2}Q \\ c' &= \det(R)c \end{aligned} \tag{7.34}$$

where Q and R are the output from the well known QR decomposition of matrices, the columns of T' are orthonormal with respect to the overlap matrix S, and c' is its new coefficient in the wave function expansion. An alternative to this approach is to utilize an exponential parameterization of the coefficient space, which guarantees the preservation of normalization. The cost and benefits of using such a parameterization within this method are a subject of current research.

7.5.5 Nuclear Union Configuration Interaction

In this section we give some of the details of the nuclear union configuration interaction method used to improve the description of reaction coordinates. In the study of a set of related problems, such as set of electronic Hamiltonians differing only by the positions of the nuclei, one would like to describe each configuration with an equivalent amount of accuracy, to get the best relative features possible. In multi-reference methods, this is often done by selecting the same active space at each configuration, and rotating the orbitals and coefficients at each geometry accordingly. In the nu-

clear union configuration interaction method, we propose each reuse of the components(determinants) found locally at other geometries to give a totally identical variational space for all nuclear configurations. As the wavefunctions produced by the NOMAGIC method are especially compact, this introduces little extra overhead to the method as a whole.

Specifically, denote the component functions found at nuclear configuration R' with corresponding Hamiltonian $H(R')$ as $|\Phi^k_{R'}\rangle = |\Phi^i\rangle$ where i is now an index set variable that runs over all the component functions at all the geometries being considered. This could be a whole reaction coordinate, or simply neighboring points depending on computational restrictions or chemical/physical considerations. At each nuclear configuration R we find new coefficients $c_i(R)$ by solving

$$H(R)C = SCE \tag{7.35}$$

for its ground state eigenvector, and we define

$$H(R)_{ij} = \langle \Phi^i | \, H(R) \, | \Phi^j \rangle \tag{7.36}$$

$$S_{ij} = \langle \Phi^i | \Phi^j \rangle \tag{7.37}$$

Note that the overlap matrix may become singular, as configurations from nearby geometries are often very similar. This can be handled either through canonical orthogonalization [102] or by removing redundant configurations before attempting the diagonalization procedure. Moreover, one might expect that additional compression is possible in this space, and this is the subject of current research.

8

Conclusions

The modelling and prediction of complex chemical systems is a challenging problem, but one whose solution has far reaching consequences. Through understanding the interaction of light with matter we can understand how to better utilize renewable energy from the sun. By accurately modelling the pathways of reactions and non-covalent interactions, we can start to understand why proteins sometimes fold incorrectly, and try to inhibit this processes to prevent the onset of devastating diseases. If we could advance our understanding of how molecules interact with surfaces and metals, the design of new renewable catalysts may be within our reach. Just as classical computers changed the way we understand the classical world, so might quantum

computers change our understanding of the quantum world.

In this thesis we have explored new ways in which one might use such a quantum device to push forward the boundaries of understanding in quantum chemistry. In developing the variational quantum eigensolver, we boiled quantum algorithms down to their essential elements for attaining an advantage over classical approaches. This approach not only allows one to utilize essentially any quantum device, but also informs our understanding about how a quantum computer can outperform a classical one. We pushed this algorithm forward by finding hardware specializations for ion traps and by exploiting the natural structure of quantum chemistry problems. However, we believe this algorithm is much more general than just quantum chemistry, and there are many advantages to come for this novel approach.

We also found that the study of quantum algorithms feeds back into our understanding of classical computation for quantum systems. By importing a tool from adiabatic quantum computation, namely Feynman's Clock, we discovered a new discrete time variational principle and showed the advantage of the equivalence between quantum dynamics and eigenvalue problems. In particular, we demonstrated how this led to a new algorithm for parallel-in-time quantum dynamics. We also demonstrated how this equivalence could be used to exploit an algorithm originally designed for many-body ground state computations, FCIQMC, to perform quantum dynamics calculations. In doing so, we gained a better understanding of the sign problem in quantum dynamics and its relation to the fermion sign problem in the study of many-body ground states.

Finally we used the insights gained from our study of quantum computation to exploit the potential sparsity in rank decompositions of many-body ground states. By using a technique from the field of compressed sensing, namely orthogonal match-

ing pursuit, we built compressed representations of many-body ground states and achieved remarkable compactness. This method was not particular to a specific ansatz in its construction, but as an example we applied it with anti-symmetric canonical tensors to quantum chemistry problems, and found great compactness with respect to more traditional orthogonal configuration interaction expansions.

At the beginning of my doctoral studies, I received my first review that I would consider simply pessimistic rather than critical. It was in response to my application for a National Science Foundation graduate fellowship (which I happened to ultimately receive, but decline), and its words stuck with me for some time. I had proposed some work on the enhancement of quantum computation through large scale optimization, and the reviewer, without reference to the proposal, remarked that the intellectual merit was "average" as "quantum computation is a field doomed only to make hard problems harder". At difficult times during my graduate studies, I reflected on these words and wondered if they might be true. It is perhaps only now, looking back that I realize they could not have been more wrong.

Not only has this field advanced essentially every aspect of what we understand about our own world, but in short time of my doctoral studies, the technology has progressed past what some referred to as insurmountable hurdles to what people now regard as solvable engineering challenges. Given how much we have learned before the first true quantum computer has been built, I cannot even begin to imagine the insights that will come as we push forward. The ramifications for chemistry, physics, and all of science will astounding. It is perhaps inappropriate that this section be called a conclusion, because this is certainly only the beginning.

References

[1] D. S. Abrams and S. Lloyd, "Simulation of many-body Fermi systems on a universal quantum computer," Phys. Rev. Lett. **79**, 2586–2589 (1997).

[2] D. S. Abrams and S. Lloyd, "Quantum algorithm providing exponential speed increase for finding eigenvalues and eigenvectors," Phys. Rev. Lett. **83**, 5162–5165 (1999).

[3] C. Amovilli, *Quantum systems in chemistry and physics. trends in methods and applications*, edited by R. McWeeny, J. Maruani, Y. Smeyers, and S. Wilson, Topics in Molecular Organization and Engineering, Vol. 16 (Springer Netherlands, 1997) pp. 343–347.

[4] A. G. Anderson and W. A. Goddard, "Generalized valence bond wave functions in quantum Monte Carlo," J. Chem. Phys. **132**, 164110 (2010).

[5] E. Artacho, D. Sánchez-Portal, P. Ordejón, A. García, and J. M. Soler, "Linear-scaling ab-initio calculations for large and complex systems," Phys. Status Solidi B **215**, 809–817 (1999).

[6] A. Aspuru-Guzik, A. D. Dutoi, P. J. Love, and M. Head-Gordon, "Simulated quantum computation of molecular energies," Science **309**, 1704–1707 (2005).

[7] A. Aspuru-Guzik, R. Salomón-Ferrer, B. Austin, and W. A. Lester, "A sparse algorithm for the evaluation of the local energy in quantum Monte Carlo," J. Comput. Chem. **26**, 708–715 (2005).

[8] A. Aspuru-Guzik and P. Walther, "Photonic quantum simulators," Nat. Phys. **8**, 285–291 (2012).

[9] S. H. Autler and C. H. Townes, "Stark effect in rapidly varying fields," Phys. Rev. **100**, 703–722 (1955).

[10] R. Babbush, P. J. Love, and A. Aspuru-Guzik, "Adiabatic quantum simulation of quantum chemistry," ArXiv e-prints (2013), arXiv:1311.3967 [quant-ph] .

[11] R. Babbush, B. O'Gorman, and A. Aspuru-Guzik, "Resource Efficient Gadgets for Compiling Adiabatic Quantum Optimization Problems," Ann. Phys. (Berlin) **525**, 877–888 (2013).

[12] R. Babbush, A. Perdomo-Ortiz, B. O'Gorman, W. Macready, and A. Aspuru-Guzik, "Construction of energy functions for lattice heteropolymer models: Efficient encodings for constraint satisfaction programming and quantum annealing," in *Advances in Chemical Physics* (Wiley-Blackwell, 2014) pp. 201–244.

[13] E. Baerends, D. Ellis, and P. Ros, "Self-consistent molecular Hartree—Fock—Slater calculations I. the computational procedure," Chem. Phys. **2**, 41–51 (1973).

[14] L. Baffico, S. Bernard, Y. Maday, G. Turinici, and G. Zérah, "Parallel-in-time molecular-dynamics simulations," Phys. Rev. E **66**, 057701 (2002).

[15] R. Balian and M. Vénéroni, "Time-dependent variational principle for predicting the expectation value of an observable," Phys. Rev. Lett. **47**, 1353–1356 (1981).

[16] S. Baroni and S. Moroni, "Reptation quantum Monte Carlo: A method for unbiased ground-state averages and imaginary-time correlations," Phys. Rev. Lett. **82**, 4745–4748 (1999).

[17] R. J. Bartlett, "Many-body perturbation theory and coupled cluster theory for electron correlation in molecules," Ann. Rev. Phys. Chem. **32**, 359–401 (1981).

[18] R. J. Bartlett and M. Musiał, "Coupled-cluster theory in quantum chemistry," Rev. Mod. Phys. **79**, 291–352 (2007).

[19] M. Beck, A. Jäckle, G. Worth, and H.-D. Meyer, "The multiconfiguration time-dependent Hartree (MCTDH) method: a highly efficient algorithm for propagating wavepackets," Phys. Rep. **324**, 1 – 105 (2000).

[20] A. J. Berkley, A. J. Przybysz, T. Lanting, R. Harris, N. Dickson, F. Altomare, M. H. Amin, P. Bunyk, C. Enderud, E. Hoskinson, M. W. Johnson, E. Ladizinsky, R. Neufeld, C. Rich, A. Y. Smirnov, E. Tolkacheva, S. Uchaikin, and A. B. Wilson, "Tunneling Spectroscopy using a Probe Qubit," Phys. Rev. B **87**, 020502 (2013).

[21] D. Berry, G. Ahokas, R. Cleve, and B. Sanders, "Efficient quantum algorithms for simulating sparse hamiltonians," Comm. Math. Phys. **270**, 359–371 (2007).

[22] D. W. Berry, "High-order quantum algorithm for solving linear differential equations," ArXiv e-prints (2010), arXiv:1010.2745 [quant-ph] .

[23] J. D. Biamonte, "Non-Perturbative k-Body to Two-Body Commuting Conversion Hamiltonians and Embedding Problem Instances into Ising Spins," Phys. Rev. A **77**, 1–8 (2008).

[24] E. Bilgin and S. Boixo, "Preparing Thermal States of Quantum Systems by Dimension Reduction," Phys. Rev. Lett. **105**, 170405 (2010), arXiv:1008.4162 .

[25] K. Binder and D. W. Heermann, *Monte Carlo simulation in statistical physics: an introduction* (Springer, 2010).

[26] G. H. Booth and A. Alavi, "Approaching chemical accuracy using full configuration-interaction quantum Monte Carlo: A study of ionization potentials," J. Chem. Phys. **132**, 174104 (2010).

[27] G. H. Booth, A. Gruneis, G. Kresse, and A. Alavi, "Towards an exact description of electronic wavefunctions in real solids," Nature **493**, 365–370 (2013).

[28] G. H. Booth, S. D. Smart, and A. Alavi, "Linear-scaling and parallelizable algorithms for stochastic quantum chemistry," ArXiv e-prints (2013), arXiv:1305.6981 [physics.comp-ph] .

[29] G. H. Booth, A. J. W. Thom, and A. Alavi, "Fermion monte carlo without fixed nodes: A game of life, death, and annihilation in slater determinant space," J. Chem. Phys. **131**, 054106 (2009).

[30] D. Bowler and T. Miyazaki, "O(n) methods in electronic structure calculations," Rep. Prog. Phys. **75**, 036503 (2012).

[31] S. F. Boys, "Electronic wave functions. i. a general method of calculation for the stationary states of any molecular system," Proc. R. Soc. Lond. A Math Phys. Sci. **200**, 542–554 (1950).

[32] S. Bravyi, "Monte Carlo simulation of stoquastic Hamiltonians," ArXiv e-prints (2014), arXiv:1402.2295 [quant-ph] .

[33] S. Bravyi, D. P. Divincenzo, R. Oliveira, and B. M. Terhal, "The complexity of stoquastic local Hamiltonian problems," Quantum Info. Comput. **8**, 361–385 (2008).

[34] S. Bravyi and B. Terhal, "Complexity of stoquastic frustration-free Hamiltonians," SIAM J. Comput. **39**, 1462–1485 (2009).

[35] S. B. Bravyi and A. Y. Kitaev, "Fermionic quantum computation," Ann. Phys. **298**, 210–226 (2002).

[36] J. W. Britton, B. C. Sawyer, A. C. Keith, C.-C. J. Wang, J. K. Freericks, H. Uys, M. J. Biercuk, and J. J. Bollinger, "Engineered two-dimensional Ising interactions in a trapped-ion quantum simulator with hundreds of spins." Nature **484**, 489–92 (2012).

[37] J. Broeckhove, L. Lathouwers, E. Kesteloot, and P. V. Leuven, "On the equivalence of time-dependent variational principles," Chem. Phys. Lett. **149**, 547 – 550 (1988).

[38] I. Buluta and F. Nori, "Quantum simulators," Science **326**, 108–111 (2009).

[39] L. Bytautas, C. A. Jiménez-Hoyos, R. Rodríguez-Guzmán, and G. E. Scuseria, "Potential energy curves for Mo 2 : multi-component symmetry-projected Hartree-Fock and beyond," Mol. Phys. **112**, 1938–1946 (2014).

[40] Y. Cao, R. Babbush, J. Biamonte, and S. Kais, "Towards experimentally realizable hamiltonian gadgets," ArXiv e-prints (2013), 1311.2555 [quant-ph] .

[41] R. Car and M. Parrinello, "Unified approach for molecular dynamics and density-functional theory," Phys. Rev. Lett. **55**, 2471–2474 (1985).

[42] J. Casanova, L. Lamata, I. L. Egusquiza, R. Gerritsma, C. F. Roos, J. J. García-Ripoll, and E. Solano, "Quantum simulation of quantum field theories in trapped ions," Phys. Rev. Lett. **107**, 260501 (2011).

[43] J. Casanova, A. Mezzacapo, L. Lamata, and E. Solano, "Quantum simulation of interacting fermion lattice models in trapped ions," Phys. Rev. Lett. **108**, 190502 (2012).

[44] A. Castro, M. A. L. Marques, and A. Rubio, "Propagators for the time-dependent kohn–sham equations," J. Chem. Phys. **121**, 3425–3433 (2004).

[45] G. K.-L. Chan and S. Sharma, "The density matrix renormalization group in quantum chemistry," Ann. Rev. Phys. Chem. **62**, 465–481 (2011).

[46] X. Chen and V. S. Batista, "Matching-pursuit/split-operator-fourier-transform simulations of excited-state nonadiabatic quantum dynamics in pyrazine," J. Chem. Phys. **125**, 124313 (2006).

[47] A. Childs and R. Kothari, "Simulating sparse hamiltonians with star decompositions," in *Theory of Quantum Computation, Communication, and Cryptography*, Lecture Notes in Computer Science, Vol. 6519, edited by W. Dam, V. Kendon, and S. Severini (Springer Berlin Heidelberg, 2011) pp. 94–103.

[48] A. M. Childs, D. Gosset, and Z. Webb, "The Bose-Hubbard model is QMA-complete," ArXiv e-prints (2013), arXiv:1311.3297 [quant-ph] .

[49] A. M. Childs, D. Gosset, and Z. Webb, "Universal computation by multiparticle quantum walk," Science **339**, 791–794 (2013).

[50] A. M. Childs and R. Kothari, "Simulating sparse Hamiltonians with star decompositions," Theory of Quantum Computation Communication and Cryptography TQC 2010 Lecture Notes in Computer Science, **6519**, 94–103 (2011).

[51] J. Cizek, "On the correlation problem in atomic and molecular systems. calculation of wavefunction components in ursell-type expansion using quantum-field theoretical methods," J. Chem. Phys. **45**, 4256–4266 (1966).

[52] B. K. Clark and H. J. Changlani, "Stochastically projecting tensor networks," ArXiv e-prints (2014), arXiv:1404.2296 [cond-mat.str-el] .

[53] B. Claude Cohen-Tannoudji, *Quantum Mechanics Volume 1-2* (Hermann, 1992).

[54] F. Coester and H. Kümmel, "Short-range correlations in nuclear wave functions," Nucl. Phys. **17**, 477 – 485 (1960).

[55] A. R. Conn, N. I. Gould, and P. L. Toint, *Trust region methods*, Vol. 1 (Society for Industrial Mathematics, 1987).

[56] T. D. Crawford, C. D. Sherrill, E. F. Valeev, J. T. Fermann, R. A. King, M. L. Leininger, S. T. Brown, C. L. Janssen, E. T. Seidl, J. P. Kenny, and W. D. Allen, "Psi3: An open-source ab initio electronic structure package," J. Comp. Chem. **28**, 1610–1616 (2007).

[57] E. Deumens, A. Diz, R. Longo, and Y. Öhrn, "Time-dependent theoretical treatments of the dynamics of electrons and nuclei in molecular systems," Rev. Mod. Phys. **66**, 917–983 (1994).

[58] M. H. Devoret, A. Wallraff, and J. Martinis, "Superconducting qubits: A short review," ArXiv e-prints (2004), arXiv:0411174 [cond-mat] .

[59] D. R. Dion and J. O. Hirschfelder, "Time-dependent perturbation of a two-state quantum system by a sinusoidal field," in *Advances in Chemical Physics* (John Wiley and Sons, Inc., 2007) pp. 265–350.

[60] P. A. M. Dirac, "Quantum Mechanics of Many-Electron Systems," Proc. R. Soc. Lond. A Math Phys. Sci. **123**, 714–733 (1929).

[61] P. A. M. Dirac, "Note on exchange phenomena in the thomas atom," Math Proc. Cambridge **26**, 376–385 (1930).

[62] R. M. Dreizler and E. Engel, *Density functional theory* (Springer, 2011).

[63] T. H. Dunning, "Gaussian basis sets for use in correlated molecular calculations. i. the atoms boron through neon and hydrogen," J. Chem. Phys. **90**, 1007 (1989).

[64] S. C. Eisenstat, "Efficient implementation of a class of preconditioned conjugate gradient methods," SIAM J. Sci. Stat. Comp. **2**, 1–4 (1981).

[65] J. Eisert, M. Cramer, and M. B. Plenio, "Colloquium : Area laws for the entanglement entropy," Rev. Mod. Phys. **82**, 277–306 (2010).

[66] M. Espig and W. Hackbusch, "A regularized newton method for the efficient approximation of tensors represented in the canonical tensor format," Numerische Mathematik **122**, 489–525 (2012).

[67] V. Faber and T. Manteuffel, "Necessary and sufficient conditions for the existence of a conjugate gradient method," SIAM J. Numer. Anal. **21**, 352–362 (1984).

[68] E. Farhi, J. Goldstone, S. Gutmann, and M. Sipser, "Quantum computation by adiabatic evolution," ArXiv e-prints (2000), arXiv:0001106 [quant-ph] .

[69] M. Feit, J. F. Jr., and A. Steiger, "Solution of the Schrödinger equation by a spectral method," J. Comput. Phys. **47**, 412–433 (1982).

[70] R. Feynman, "Simulating physics with computers," Int. J. Theor. Phys. **21**, 467–488 (1982).

[71] R. Feynman and A. Hibbs, *Quantum Mechanics and Path Integrals: Emended Edition* (Dover Publications, Incorporated, 2012).

[72] R. P. Feynman, "Quantum mechanical computers," Opt. News **11**, 11–20 (1985).

[73] R. Fletcher, *Practical methods of optimization*, 2nd ed. (Wiley-Interscience, New York, NY, USA, 1987).

[74] W. M. C. Foulkes, L. Mitas, R. J. Needs, and G. Rajagopal, "Quantum Monte Carlo simulations of solids," Rev. Mod. Phys. **73**, 33–83 (2001).

[75] J. Frenkel, *Wave Mechanics* (Claredon Press, Oxford, 1934).

[76] M. J. Gander, "Analysis of the parareal algorithm applied to hyperbolic problems using characteristics," Boletín SEMA (2008).

[77] M. J. Gander and S. Vandewalle, "Analysis of the parareal time-parallel time-integration method," SIAM J. Sci. Comp. **29**, 556–578 (2007).

[78] S. Garnerone, P. Zanardi, and D. A. Lidar, "Adiabatic quantum algorithm for search engine ranking," Phys. Rev. Lett. **108**, 230506 (2012).

[79] Y. Ge and J. Eisert, "Area laws and approximations of quantum many-body states," ArXiv e-prints (2014), arXiv:1411.2995 [quant-ph] .

[80] I. M. Georgescu, S. Ashhab, and F. Nori, "Quantum simulation," Rev. Mod. Phys. **86**, 153–185 (2014).

[81] R. Gerritsma, G. Kirchmair, F. Zähringer, E. Solano, R. Blatt, and C. F. Roos, "Quantum simulation of the Dirac equation." Nature **463**, 68–71 (2010).

[82] E. Gibney, "Physics: Quantum computer quest," Nature **516**, 24–26 (2014).

[83] E. Giner, A. Scemama, and M. Caffarel, "Using perturbatively selected configuration interaction in quantum Monte Carlo calculations," Can. J. Chem. **91**, 879–885 (2013).

[84] W. A. Goddard III, T. H. Dunning Jr, W. J. Hunt, and P. J. Hay, "Generalized valence bond description of bonding in low-lying states of molecules," Acc. Chem. Res. **6**, 368–376 (1973).

[85] S. Goedecker, "Linear scaling electronic structure methods," Rev. Mod. Phys. **71**, 1085–1123 (1999).

[86] G. H. Golub and H. A. van der Vorst, "Eigenvalue computation in the 20th century," J. Comput. Appl. Math. **123**, 35–65 (2000).

[87] D. Gosset and D. Nagaj, "QuaArXiv e-printsntum 3-SAT is QMA1-complete," ArXiv e-prints (2013), arXiv:1302.0290 [quant-ph] .

[88] H. Goto, M. Kojo, A. Sasaki, and K. Hirose, "Essentially exact ground-state calculations by superpositions of nonorthogonal Slater determinants," Nanoscale Res. Lett. **8**, 1–7 (2013).

[89] R. B. Griffiths and C.-S. Niu, "Semiclassical fourier transform for quantum computation," Phys. Rev. Lett. **76**, 3228–3231 (1996).

[90] S. Habershon, "Trajectory-guided configuration interaction simulations of multi-dimensional quantum dynamics," J. Chem. Phys. **136**, 054109 (2012).

[91] J. Hachmann, W. Cardoen, and G. K.-L. Chan, "Multireference correlation in long molecules with the quadratic scaling density matrix renormalization group," J. Chem. Phys. **125**, 144101 (2006).

[92] W. Hackbusch, *Tensor spaces and numerical tensor calculus*, Vol. 42 (Springer, 2012).

[93] J. Haegeman, J. I. Cirac, T. J. Osborne, I. Pižorn, H. Verschelde, and F. Verstraete, "Time-dependent variational principle for quantum lattices," Phys. Rev. Lett. **107**, 070601 (2011).

[94] J. Haegeman, C. Lubich, I. Oseledets, B. Vandereycken, and F. Verstraete, "Unifying time evolution and optimization with matrix product states," ArXiv e-prints (2014), arXiv:1408.5056 [quant-ph] .

[95] H. Häffner, C. Roos, and R. Blatt, "Quantum computing with trapped ions," Phys. Rep. **469**, 155–203 (2008).

[96] B. Hammond, W. Lester, and P. Reynolds, *Monte Carlo Methods in Ab Initio Quantum Chemistry*, World Scientific Lecture and Course Notes in Chemistry ; Vol. 1 (World Scientific, 1994).

[97] R. J. Harrison, G. I. Fann, T. Yanai, Z. Gan, and G. Beylkin, "Multiresolution quantum chemistry: Basic theory and initial applications," J. Chem. Phys. **121**, 11587–11598 (2004).

[98] A. Harrow, A. Hassidim, and S. Lloyd, "Quantum Algorithm for Linear Systems of Equations," Phys. Rev. Lett. **103**, 150502 (2009).

[99] M. B. Hastings, D. Wecker, B. Bauer, and M. Troyer, "Improving quantum algorithms for quantum chemistry," ArXiv e-prints (2014), arXiv:1403.1539 [quant-ph] .

[100] M. Head-Gordon and E. Artacho, "Chemistry on the computer," Phys. Today **61**, 58 (2008).

[101] M. Head-Gordon, P. E. Maslen, and C. A. White, "A tensor formulation of many-electron theory in a nonorthogonal single-particle basis," J. Chem. Phys. **108**, 616–625 (1998).

[102] T. Helgaker, P. Jorgensen, and J. Olsen, *Molecular Electronic Structure Theory* (Wiley, Sussex, 2002).

[103] E. J. Heller, "Time dependent variational approach to semiclassical dynamics," J. Chem. Phys. **64**, 63–73 (1976).

[104] M. R. Hestenes and E. Stiefel, "Methods of conjugate gradients for solving linear systems," J. Res. Nat. Bur. Stand. **49** (1952).

[105] H. J. Hogben, P. J. Hore, and I. Kuprov, "Strategies for state space restriction in densely coupled spin systems with applications to spin chemistry," J. Chem. Phys. **132**, 174101 (2010).

[106] J. S. Howland, "Stationary scattering theory for time-dependent hamiltonians," Mathematische Annalen **207**, 315–335 (1974), 10.1007/BF01351346.

[107] W. J. Hunt, P. J. Hay, and W. A. Goddard, "Selfconsistent procedures for generalized valence bond wavefunctions. applications H_3, BH, H_2O, C_2H_6, and O_2," J. Chem. Phys. **57**, 738–748 (1972).

[108] R. Jackiw and A. Kerman, "Time-dependent variational principle and the effective action," Phys. Lett. A **71**, 158 – 162 (1979).

[109] V. Jadhao and N. Makri, "Iterative Monte Carlo for quantum dynamics," J. Chem. Phys. **129**, 161102 (2008).

[110] C. A. Jiménez-Hoyos, R. Rodríguez-Guzmán, and G. E. Scuseria, "Multicomponent symmetry-projected approach for molecular ground state correlations," J. Chem. Phys **139**, 204102 (2013).

[111] N. C. Jones, J. D. Whitfield, P. L. McMahon, M.-H. Yung, R. V. Meter, A. Aspuru-Guzik, and Y. Yamamoto, "Faster quantum chemistry simulation on fault-tolerant quantum computers," New J. Phys. **14**, 115023 (2012).

[112] P. Jordan and E. Wigner, "Über das paulische äquivalenzverbot," Zeitschrift für Physik **47**, 631–651 (1928).

[113] S. P. Jordan and E. Farhi, "Perturbative Gadgets at Arbitrary Orders," Phys. Rev. A **77**, 1–8 (2008).

[114] S. Kais, *Advances in Chemical Physics, Quantum Information and Computation for Chemistry*, Vol. 154 (John Wiley & Sons, 2014).

[115] I. Kassal and A. Aspuru-Guzik, "Quantum algorithm for molecular properties and geometry optimization," J. Chem. Phys. **131**, 224102 (2009).

[116] 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. U.S.A. **105**, 18681–18686 (2008).

[117] I. Kassal, J. D. Whitfield, A. Perdomo-Ortiz, M.-H. Yung, and A. Aspuru-Guzik, "Simulating chemistry using quantum computers," Ann. Rev. Phys. Chem. **62**, 185–207 (2011).

[118] P. Kaye, R. Laflamme, and M. Mosca, *An introduction to quantum computing* (Oxford University Press, USA, 2007).

[119] J. Kempe, A. Kitaev, and O. Regev, "The complexity of the local Hamiltonian problem," SIAM J. Comput. **35**, 1070–1097 (2006).

[120] A. Kerman and S. Koonin, "Hamiltonian formulation of time-dependent variational principles for the many-body system," Ann. Phys. **100**, 332 – 358 (1976).

[121] G. Kirchmair, J. Benhelm, F. Zähringer, R. Gerritsma, C. F. Roos, and R. Blatt, "Deterministic entanglement of ions in thermal states of motion," New J. Phys. **11**, 023002 (2009).

[122] A. Kitaev, "Quantum measurements and the abelian stabilizer problem," Electronic Colloquium on Computational Complexity (ECCC) **3** (1996).

[123] A. Kitaev, A. Shen, M. Vyalyi, and N. Vyalyi, *Classical and Quantum Computation*, Graduate Studies in Mathematics (American Mathematical Society, 2002).

[124] E. Knill, G. Ortiz, and R. D. Somma, "Optimal quantum measurements of expectation values of observables," Phys. Rev. A **75**, 012328 (2007).

[125] H. Koch and E. Dalgaard, "Linear superposition of optimized non-orthogonal slater determinants for singlet states," Chem. Phys. Lett. **212**, 193–200 (1993).

[126] W. Kohn, A. D. Becke, and R. G. Parr, "Density functional theory of electronic structure," J. Phys. Chem. **100**, 12974–12980 (1996).

[127] T. G. Kolda and B. W. Bader, "Tensor decompositions and applications," SIAM Rev. **51**, 455–500 (2009).

[128] M. H. Kolodrubetz, J. S. Spencer, B. K. Clark, and W. M. Foulkes, "The effect of quantization on the full configuration interaction quantum Monte Carlo sign problem," J. Chem. Phys. **138**, 024110 (2013).

[129] L. V. Kulik, B. Epel, W. Lubitz, and J. Messinger, "Electronic structure of the Mn_4O_xCa cluster in the S0 and S2 states of the oxygen-evolving complex of photosystem II based on pulse ^{55}Mn-ENDOR and EPR spectroscopy," J. Am. Chem. Soc. **129**, 13421–13435 (2007).

[130] I. Kuprov, N. Wagner-Rundell, and P. Hore, "Polynomially scaling spin dynamics simulation algorithm based on adaptive state-space restriction," J. Magn. Reson. **189**, 241 – 250 (2007).

[131] W. Kutzelnigg, "Error analysis and improvements of coupled-cluster theory," Theor. Chim. Acta **80**, 349–386 (1991).

[132] S. Kvaal, "Ab initio quantum dynamics using coupled-cluster," J. Chem. Phys. **136**, 194109 (2012).

[133] L. Lamata, J. León, T. Schätz, and E. Solano, "Dirac Equation and Quantum Relativistic Effects in a Single Trapped Ion," Physical Review Letters **98**, 253005 (2007).

[134] Z. Landau, U. Vazirani, and T. Vidick, "A polynomial-time algorithm for the ground state of 1D gapped local Hamiltonians," ArXiv e-prints (2013), arXiv:1307.5143 [quant-ph] .

[135] B. P. Lanyon, C. Hempel, D. Nigg, M. Müller, R. Gerritsma, F. Zähringer, P. Schindler, J. T. Barreiro, M. Rambach, G. Kirchmair, M. Hennrich, P. Zoller, R. Blatt, and C. F. Roos, "Universal digital quantum simulation with trapped ions." Science (New York, N.Y.) **334**, 57–61 (2011).

[136] B. P. Lanyon, J. D. Whitfield, G. G. Gillett, 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, "Towards quantum chemistry on a quantum computer," Nat. Chem. **2**, 106–111 (2010).

[137] A. Leggett, S. Chakravarty, A. Dorsey, M. Fisher, A. Garg, and W. Zwerger, "Dynamics of the dissipative two-state system," Rev. Mod. Phys. **59**, 1–85 (1987).

[138] D. Leibfried, R. Blatt, C. Monroe, and D. Wineland, "Quantum dynamics of single trapped ions," Rev. Mod. Phys. **75**, 281–324 (2003).

[139] W. A. Lester, B. Hammond, and P. J. Reynolds, *Monte Carlo methods in ab initio quantum chemistry* (World Scientific, 1994).

[140] Z. Li, M.-H. Yung, H. Chen, D. Lu, J. D. Whitfield, X. Peng, A. Aspuru-Guzik, and J. Du, "Solving quantum ground-state problems with nuclear magnetic resonance." Sci. Rep. **1**, 88 (2011).

[141] D. Lidar and O. Biham, "Simulating Ising spin glasses on a quantum computer," Phys. Rev. E. **56**, 3661–3681 (1997).

[142] D. Lidar and H. Wang, "Calculating the thermal rate constant with exponential speedup on a quantum computer," Phys. Rev. E. **59**, 2429–2438 (1999).

[143] J. Lions, Y. Maday, and G. Turinici, "A "parareal" in time discretization of pde's," Comptes Rendus de l'Academie des Sciences Series I Mathematics **332**, 661–668 (2001).

[144] B. A. Lippmann and J. Schwinger, "Variational principles for scattering processes. i," Phys Rev **79**, 469–480 (1950).

[145] Y.-K. Liu, M. Christandl, and F. Verstraete, "Quantum computational complexity of the N-representability problem: QMA complete," Phys. Rev. Lett. **98**, 110503 (2007).

[146] S. Lloyd, "Universal Quantum Simulators," Science (New York, N.Y.) **273**, 1073 (1996).

[147] S. Lloyd, "Computational capacity of the universe," Phys. Rev. Lett. **88**, 237901 (2002).

[148] P. J. Love, "Back to the Future: A roadmap for quantum simulation from vintage quantum chemistry," ArXiv e-prints (2012), arXiv:1208.5524 .

[149] P. Löwdin, "On the nonorthogonality problem connected with the use of atomic wave functions in the theory of molecules and crystals," J. Chem. Phys. **18**, 365–375 (1950).

[150] P.-O. Löwdin, "Quantum theory of many-particle systems. i. physical interpretations by means of density matrices, natural spin-orbitals, and convergence problems in the method of configurational interaction," Phys. Rev. **97**, 1474–1489 (1955).

[151] 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).

[152] M. Luban, F. Borsa, S. Bud'ko, P. Canfield, S. Jun, J. K. Jung, P. Kögerler, D. Mentrup, A. Müller, R. Modler, D. Procissi, B. J. Suh, and M. Torikachvili, "Heisenberg spin triangles in v_6-type magnetic molecules: experiment and theory," Phys. Rev. B **66**, 054407 (2002).

[153] X.-s. Ma, B. Dakic, W. Naylor, A. Zeilinger, and P. Walther, "Quantum simulation of the wavefunction to probe frustrated heisenberg spin systems," Nat. Phys. **7**, 399–405 (2011).

[154] J. K. L. MacDonald, "On the modified Ritz variation method," Phys. Rev. **46**, 828–828 (1934).

[155] C. H. Mak, "Stochastic method for real-time path integrations," Phys. Rev. Lett. **68**, 899–902 (1992).

[156] C. H. Mak and D. Chandler, "Solving the sign problem in quantum monte carlo dynamics," Phys. Rev. A **41**, 5709–5712 (1990).

[157] N. Makri, "Feynman path integration in quantum dynamics," Comp. Phys. Comm. **63**, 389 – 414 (1991).

[158] N. Makri, "On smooth Feynman propagators for real time path integrals," J. Phys. Chem. **97**, 2417–2424 (1993).

[159] N. Makri and W. H. Miller, "Monte Carlo integration with oscillatory integrands: implications for feynman path integration in real time," Chem. Phys. Lett. **139**, 10 – 14 (1987).

[160] N. Makri and W. H. Miller, "Monte Carlo path integration for the real time propagator," J. Chem. Phys. **89**, 2170–2177 (1988).

[161] C. Master, F. Yamaguchi, and Y. Yamamoto, "Efficiency of free-energy calculations of spin lattices by spectral quantum algorithms," Phys. Rev. A **67**, 032311 (2003).

[162] V. May and O. Kühn, *Charge and Energy Transfer Dynamics in Molecular Systems* (John Wiley & Sons, 2011).

[163] I. Mayer, "On Löwdin's method of symmetric orthogonalization," Int. J. Quant. Chem. **90**, 63–65 (2002).

[164] J. R. McClean and A. Aspuru-Guzik, "Compact wavefunctions from compressed imaginary time evolution," ArXiv e-prints (2014), arXiv:1409.7358 [physics.chem-ph] .

[165] J. R. McClean and A. Aspuru-Guzik, "Clock quantum Monte Carlo technique: An imaginary-time method for real-time quantum dynamics," Phys. Rev. A **91**, 012311 (2015).

[166] J. R. McClean, R. Babbush, P. J. Love, and A. Aspuru-Guzik, "Exploiting locality in quantum computation for quantum chemistry," J. Phys. Chem. Lett. **5**, 4368–4380 (2014), http://dx.doi.org/10.1021/jz501649m .

[167] J. R. McClean, J. A. Parkhill, and A. Aspuru-Guzik, "Feynman's clock, a new variational principle, and parallel-in-time quantum dynamics," Proc. Natl. Acad. Sci. U.S.A. **110**, E3901–E3909 (2013).

[168] A. McLachlan, "A variational solution of the time-dependent schrodinger equation," Mol. Phys. **8**, 39–44 (1964).

[169] J. Messinger, J. H. Robblee, U. Bergmann, C. Fernandez, P. Glatzel, H. Visser, R. M. Cinco, K. L. McFarlane, E. Bellacchio, S. A. Pizarro, S. P. Cramer, K. Sauer, M. P. Klein, and V. K. Yachandra, "Absence of Mn-centered oxidation in the S2 $\rightarrow$ S3 transition: implications for the mechanism of photosynthetic water oxidation," J. Am. Chem. Soc. **123**, 7804–7820 (2001).

[170] A. Mezzacapo, J. Casanova, L. Lamata, and E. Solano, "Digital quantum simulation of the Holstein model in trapped ions," Phys. Rev. Lett. **109**, 200501 (2012).

[171] K. F. Milfeld and R. E. Wyatt, "Study, extension, and application of floquet theory for quantum molecular systems in an oscillating field," Phys. Rev. A **27**, 72–94 (1983).

[172] A. Mizel, "Mimicking time evolution within a quantum ground state: Ground-state quantum computation, cloning, and teleportation," Phys Rev A **70**, 012304 (2004).

[173] C. Möller and M. S. Plesset, "Note on an approximation treatment for many-electron systems," Phys. Rev. **46**, 618–622 (1934).

[174] K. Mølmer and A. Sørensen, "Multiparticle entanglement of hot trapped ions," Phys. Rev. Lett. **82**, 1835–1838 (1999).

[175] S. Mostame, P. Rebentrost, A. Eisfeld, A. J. Kerman, D. I. Tsomokos, and A. Aspuru-Guzik, "Quantum simulator of an open quantum system using superconducting qubits: exciton transport in photosynthetic complexes," New J. Phys. **14**, 105013 (2012).

[176] M. Müller, K. Hammerer, Y. L. Zhou, C. F. Roos, and P. Zoller, "Simulating open quantum systems: from many-body interactions to stabilizer pumping," New Journal of Physics **13**, 085007 (2011).

[177] A. H. Myerson, D. Szwer, S. Webster, D. Allcock, M. Curtis, G. Imreh, J. Sherman, D. Stacey, A. Steane, and D. Lucas, "High-fidelity readout of trapped-ion qubits," Phys. Rev. Lett. **100**, 2–5 (2008).

[178] D. Nagaj, "Fast universal quantum computation with railroad-switch local hamiltonians," J. Math. Phys. **51**, 062201 (2010).

[179] D. Needell and R. Vershynin, "Uniform uncertainty principle and signal recovery via regularized orthogonal matching pursuit," Found. Comput. Math. **9**, 317–334 (2009).

[180] J. A. Nelder and R. Mead, "A simplex method for function minimization," Comput. J. **7**, 308–313 (1965).

[181] M. Nest, T. Klamroth, and P. Saalfrank, "The multiconfiguration time-dependent Hartree-Fock method for quantum chemical calculations," J. Chem. Phys. **122**, 124102 (2005).

[182] H. Neven, G. Rose, and W. G. Macready, "Image recognition with an adiabatic quantum computer I. Mapping to quadratic unconstrained binary optimization," ArXiv e-prints (2008), arXiv:0804.4457 [quant-ph] .

[183] M. Nielsen and I. Chuang, *Quantum Computation and Quantum Information*, Cambridge Series on Information and the Natural Sciences (Cambridge University Press, 2000).

[184] M. Nightingale and C. Umrigar, *Quantum Monte Carlo Methods in Physics and Chemistry*, Nato ASI series (Kluwer Academic Publishers, 1999).

[185] J. Nocedal and S. J. Wright, *Numerical Optimization (2nd edition)* (Springer, 2006).

[186] J. L. O'Brien, A. Furusawa, and J. Vuckovic, "Photonic quantum technologies," Nat. Photon. **3**, 687–695 (2009).

[187] C. Ochsenfeld, J. Kussmann, and D. S. Lambrecht, "Linear-scaling methods in quantum chemistry," in *Reviews in Computational Chemistry* (Wiley-Blackwell, 2007) pp. 1–82.

[188] G. Ortiz, J. E. Gubernatis, E. Knill, and R. Laflamme, "Quantum algorithms for fermionic simulations," Phys. Rev. A **64**, 022319 (2001).

[189] I. Oseledets, "Approximation of $2^d \times 2^d$ matrices using tensor decomposition," SIAM J. Matrix Anal. A. **31**, 2130–2145 (2010).

[190] V. Ozoliņš, R. Lai, R. Caflisch, and S. Osher, "Compressed modes for variational problems in mathematics and physics," Proc. Natl. Acad. Sci. U.S.A. **110**, 18368–18373 (2013).

[191] L. Page, S. Brin, R. Motwani, and T. Winograd, "The pagerank citation ranking: Bringing order to the web." Technical Report 1999-66 (Stanford InfoLab, 1999).

[192] R. G. Parr and W. Yang, *Density-functional theory of atoms and molecules*, Vol. 16 (Oxford University Press, 1989).

[193] Y. C. Pati, R. Rezaiifar, and P. Krishnaprasad, "Orthogonal matching pursuit: Recursive function approximation with applications to wavelet decomposition," in *Signals, Systems and Computers, 1993. 1993 Conference Record of The Twenty-Seventh Asilomar Conference on* (IEEE, 1993) pp. 40–44.

[194] J. P. Perdew, K. Burke, and M. Ernzerhof, "Generalized gradient approximation made simple," Phys. Rev. Lett. **77**, 3865 (1996).

[195] A. Perdomo-Ortiz, N. Dickson, M. Drew-Brook, G. Rose, and A. Aspuru-Guzik, "Finding low-energy conformations of lattice protein models by quantum annealing," Sci. Rep. **2**, 571–1–571–7 (2012).

[196] A. Peruzzo, J. McClean, P. Shadbolt, M.-H. Yung, X.-Q. Zhou, P. J. Love, A. Aspuru-Guzik, and J. L. O'Brien, "A variational eigenvalue solver on a quantum processor," Nat. Commun. **5**, 4213–1–4213–7 (2014).

[197] U. Peskin and N. Moiseyev, "The solution of the time-dependent Schrödinger equation by the (t,t') method: Theory, computational algorithm and applications," J. Chem. Phys. **99**, 4590–4596 (1993).

[198] D. Poulin, M. B. Hastings, D. Wecker, N. Wiebe, A. C. Doherty, and M. Troyer, "The Trotter step size required for accurate quantum simulation of quantum chemistry," ArXiv e-prints (2014), arXiv:1406.4920 [quant-ph] .

[199] 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**, 170501 (2011).

[200] D. Poulin and P. Wocjan, "Preparing ground states of quantum many-body systems on a quantum computer," Phys. Rev. Lett. **102**, 130503 (2009).

[201] 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).

[202] J. A. Poulsen, "A variational principle in Wigner phase-space with applications to statistical mechanics," J. Chem. Phys. **134**, 034118 (2011).

[203] T. R. Rao, G. Guillon, S. Mahapatra, and P. Honvault, "Huge quantum symmetry effect in the O + O_2 exchange reaction," J. Phys. Chem. Lett. **6**, 633–636 (2015).

[204] J. W. Rayleigh, "On finding the correction for the open end of an organ-pipe," Phil. Trans. **161**, 77 (1870).

[205] W. Ritz, "Ãber eine neue methode zur lãsung gewisser variationsprobleme der mathematischen physik," J. reine angew. Math. **135**, 1–61 (1908).

[206] Y. Saad, *Numerical methods for large eigenvalue problems*, Vol. 158 (SIAM, 1992).

[207] H. Sambe, "Steady states and quasienergies of a quantum-mechanical system in an oscillating field," Phys. Rev. A. **7**, 2203–2213 (1973).

[208] U. Schollwöck, "The density-matrix renormalization group," Rev. Mod. Phys. **77**, 259–315 (2005).

[209] J. T. Seeley, M. J. Richard, and P. J. Love, "The Bravyi-Kitaev transformation for quantum computation of electronic structure," J. Chem. Phys. **137**, 224109 (2012).

[210] P. Shadbolt, M. Verde, A. Peruzzo, A. Politi, A. Laing, M. Lobino, J. Matthews, M. Thompson, and J. O'Brien, "Generating, manipulating and measuring entanglement and mixture with a reconfigurable photonic circuit," Nat. Photon. **6**, 45–49 (2011).

[211] D. V. Shalashilin and I. Burghardt, "Gaussian-based techniques for quantum propagation from the time-dependent variational principle: Formulation in terms of trajectories of coupled classical and quantum variables," J. Chem. Phys. **129**, 084104 (2008).

[212] Y. Shao, L. F. Molnar, Y. Jung, J. Kussmann, C. Ochsenfeld, S. T. Brown, A. T. Gilbert, L. V. Slipchenko, S. V. Levchenko, D. P. O'Neill, *et al.*, "Advances in methods and algorithms in a modern quantum chemistry program package," Phys. Chem. Chem. Phys. **8**, 3172–3191 (2006).

[213] J. J. Shepherd, G. E. Scuseria, and J. S. Spencer, "The sign problem in full configuration interaction quantum Monte Carlo: Linear and sub-linear representation regimes for the exact wave function," ArXiv e-prints (2014), arXiv:1407.4800 [physics.comp-ph] .

[214] P. W. Shor, "Algorithms for quantum computation: Discrete logarithms and factoring," Proceedings 35th Annual Symposium on Foundations of Computer Science , 124–134 (1994).

[215] P. W. Shor, "Polynomial-time algorithms for prime factorization and discrete logarithms on a quantum computer," SIAM Rev. **41**, 303–332 (1999).

[216] J. C. Slater, "The theory of complex spectra," Phys. Rev. **34**, 1293–1322 (1929).

[217] J. C. Slater, "A simplification of the Hartree-Fock method," Phys. Rev. **81**, 385 (1951).

[218] B. Smith, P. Bjorstad, and W. Gropp, *Domain decomposition* (Cambridge University Press, 2004).

[219] L. Song, J. Song, Y. Mo, and W. Wu, "An efficient algorithm for energy gradients and orbital optimization in valence bond theory," J. Comp. Chem. **30**, 399–406 (2009).

[220] J. S. Spencer, N. S. Blunt, and W. M. Foulkes, "The sign problem and population dynamics in the full configuration interaction quantum Monte Carlo method," J. Chem. Phys. **136**, 054110 (2012).

[221] E. J. Sundstrom and M. Head-Gordon, "Non-orthogonal configuration interaction for the calculation of multielectron excited states," J. Chem. Phys. **140**, 114103 (2014).

[222] M. Suzuki, "General decomposition theory of ordered exponentials." Proc. Jpn. Acad. Ser. B Phys. Biol. Sci. **69**, 161–166 (1993).

[223] S. J. Szarek, "Volume of separable states is super-doubly-exponentially small in the number of qubits," Phys. Rev. A **72**, 032304 (2005).

[224] D. Tannor, *Introduction to Quantum Mechanics: A Time-Dependent Perspective* (University Science Books, 2007).

[225] A. G. Taube and R. J. Bartlett, "New perspectives on unitary coupled-cluster theory," Int. J. Quant. Chem. **106**, 3393–3401 (2006).

[226] K. Temme, T. J. Osborne, K. G. Vollbrecht, D. Poulin, and F. Verstraete, "Quantum Metropolis sampling," Nature **471**, 87–90 (2011).

[227] D. G. Tempel and A. Aspuru-Guzik, "The Kitaev-Feynman clock for open quantum systems," New J. Phys. **16**, 113066 (2014).

[228] D. Thirumalai and B. J. Berne, "On the calculation of time correlation functions in quantum systems: Path integral techniques," J. Chem. Phys. **79**, 5029–5033 (1983).

[229] L. Thogersen and J. Olsen, "A coupled cluster and full configuration interaction study of CN and CN-," Chem. Phys. Lett. **393**, 36 – 43 (2004).

[230] B. Toloui and P. J. Love, "Quantum algorithms for quantum chemistry based on the sparsity of the CI-matrix," ArXiv e-prints (2013), arXiv:1312.2579 [quant-ph] .

[231] N. Trivedi and D. M. Ceperley, "Ground-state correlations of quantum antiferromagnets: A Green-function Monte Carlo study," Phys. Rev. B **41**, 4552–4569 (1990).

[232] J. A. Tropp and A. C. Gilbert, "Signal recovery from random measurements via orthogonal matching pursuit," IEEE Trans. Inf. Theory **53**, 4655–4666 (2007).

[233] H. F. Trotter, "On the product of semi-groups of operators," Proc Am Math Soc **10**, 545–551 (1959).

[234] M. Troyer and U. J. Wiese, "Computational complexity and fundamental limitations to fermionic quantum Monte Carlo simulations," Phys. Rev. Lett. **94**, 170201 (2005).

[235] V. G. Tyuterev, R. Kochanov, A. Campargue, S. Kassi, D. Mondelain, A. Barbe, E. Starikova, M. R. D. Backer, P. G. Szalay, and S. Tashkun, "Does the reef structure at the ozone transition state towards the dissociation exist? new insight from calculations and ultrasensitive spectroscopy experiments," Phys. Rev. Lett. **113** (2014), 10.1103/physrevlett.113.143002.

[236] L. Veis and J. Pittner, "Adiabatic state preparation study of methylene," J. Chem. Phys. **140**, 214111 (2014).

[237] L. Veis, J. Višňák, T. Fleig, S. Knecht, T. Saue, L. Visscher, and J. c. v. Pittner, "Relativistic quantum chemistry on quantum computers," Phys. Rev. A **85**, 030304 (2012).

[238] G. Vidal, "Efficient simulation of one-dimensional quantum many-body systems," Phys. Rev. Lett. **93** (2004).

[239] J. H. van Vleck, "Nonorthogonality and ferromagnetism," Phys. Rev. **49**, 232–240 (1936).

[240] H. Wang, S. Kais, A. Aspuru-Guzik, and M. R. Hoffmann, "Quantum algorithm for obtaining the energy spectrum of molecular systems," Phys. Chem. Chem. Phys. **10**, 5388–5393 (2008).

[241] L.-W. Wang and A. Zunger, "Solving schrödinger's equation around a desired energy: Application to silicon quantum dots," J. Chem. Phys. **100**, 2394–2397 (1994).

[242] Y. Wang, F. Dolde, J. Biamonte, R. Babbush, V. Bergholm, S. Yang, I. Jakobi, P. Neumann, A. Aspuru-Guzik, J. D. Whitfield, and J. Wrachtrup, "Quantum simulation of helium hydride in a solid-state spin register," ArXiv e-prints (2014), arXiv:1405.2696 .

[243] N. J. Ward, I. Kassal, and A. Aspuru-Guzik, "Preparation of many-body states for quantum simulation," J. Chem. Phys. **130**, 194105 (2009).

[244] D. Wecker, B. Bauer, B. K. Clark, M. B. Hastings, and M. Troyer, "Gate-count estimates for performing quantum chemistry on small quantum computers," Phys. Rev. A **90**, 022305 (2014).

[245] T. C. Wei, M. Mosca, and A. Nayak, "Interacting boson problems can be qma hard," Phys. Rev. Lett. **104**, 040501 (2010).

[246] J. Welch, D. Greenbaum, S. Mostame, and A. Aspuru-Guzik, "Efficient quantum circuits for diagonal unitaries without ancillas," New J. Phys. **16**, 033040 (2014).

[247] H.-J. Werner and K. Pflüger, "On the selection of domains and orbital pairs in local correlation treatments," in *Annual Reports in Computational Chemistry* (Elsevier {BV}, 2006) pp. 53–80.

[248] T. Westermann and U. Manthe, "Decoherence induced by conical intersections: Complexity constrained quantum dynamics of photoexcited pyrazine," J. Chem. Phys. **137**, 22A509 (2012).

[249] S. R. White and R. L. Martin, "Ab initio quantum chemistry using the density matrix renormalization group," J. Chem. Phys. **110**, 4127–4130 (1999).

[250] J. Whitfield, J. Biamonte, and A. Aspuru-Guzik, "Quantum computing resource estimate of molecular energy simulation," ArXiv e-prints (2010), arXiv:1001.3855v1 [quant-ph] .

[251] J. D. Whitfield, J. Biamonte, and A. Aspuru-Guzik, "Simulation of electronic structure Hamiltonians using quantum computers," Mol. Phys. **109**, 735–750 (2011).

[252] J. D. Whitfield, P. J. Love, and A. Aspuru-Guzik, "Computational complexity in electronic structure," Phys. Chem. Chem. Phys. **15**, 397–411 (2013).

[253] A. Williamson, R. Q. Hood, and J. Grossman, "Linear-scaling quantum monte carlo calculations," Phys. Rev. Lett. **87**, 246406 (2001).

[254] P. Wocjan, C.-F. Chiang, D. Nagaj, and A. Abeyesinghe, "Quantum algorithm for approximating partition functions," Phys. Rev. A **80** (2009).

[255] D. E. Woon and T. H. Dunning, "Gaussian basis sets for use in correlated molecular calculations. IV. calculation of static electrical response properties," J. Chem. Phys. **100**, 2975 (1994).

[256] L.-A. Wu, M. Byrd, and D. Lidar, "Polynomial-time simulation of pairing models on a quantum computer," Phys. Rev. Lett. **89**, 1–4 (2002).

[257] Y. Wu and V. S. Batista, "Matching-pursuit for simulations of quantum processes," J. Chem. Phys. **118**, 6720–6724 (2003).

[258] A. L. V. Wyngarden, K. A. Mar, K. A. Boering, J. J. Lin, Y. T. Lee, S.-Y. Lin, H. Guo, and G. Lendvay, "Nonstatistical behavior of reactive scattering in the ^{18} O + ^{32} O_2 isotope exchange reaction," J. Am. Chem. Soc. **129**, 2866–2870 (2007).

[259] A. L. V. Wyngarden, K. A. Mar, J. Quach, A. P. Q. Nguyen, A. A. Wiegel, S.-Y. Lin, G. Lendvay, H. Guo, J. J. Lin, Y. T. Lee, and K. A. Boering, "The non-statistical dynamics of the ^{18} O + ^{32} O_2 isotope exchange reaction at two energies," J. Chem. Phys. **141**, 064311 (2014).

[260] J.-S. Xu, M.-H. Yung, X.-Y. Xu, S. Boixo, Z.-W. Zhou, C.-F. Li, A. Aspuru-Guzik, and G.-C. Guo, "Demon-like algorithmic quantum cooling and its realization with quantum optics," Nat. Photon. **8**, 113–118 (2014).

[261] M. Yung, J. Casanova, A. Mezzacapo, J. McClean, L. Lamata, A. Aspuru-Guzik, and E. Solano, "From Transistor to Trapped-Ion Computers for Quantum Chemistry," Sci. Rep. **4**, 3589–1–3589–7 (2014).

[262] M.-H. Yung and A. Aspuru-Guzik, "A quantum-quantum Metropolis algorithm." Proc. Natl. Acad. Sci. U.S.A. **109**, 754–9 (2012).

[263] M.-H. Yung, D. Nagaj, J. Whitfield, and A. Aspuru-Guzik, "Simulation of classical thermal states on a quantum computer: A transfer-matrix approach," Phys. Rev. A **82**, 060302 (2010).

[264] M.-H. Yung, J. D. Whitfield, S. Boixo, D. G. Tempel, and A. Aspuru-Guzik, "Introduction to quantum algorithms for physics and chemistry," ArXiv e-prints (2012), arXiv:1203.1331 [quant-ph] .

[265] R. Zalesny, M. G. Papadopoulos, P. G. Mezey, and J. Leszczynski, eds., *Linear-Scaling Techniques in Computational Chemistry and Physics* (Springer Netherlands, 2011).

[266] C. Zalka, "Simulating quantum systems on a quantum computer," Proc. R. Soc. A **454**, 313–322 (1998).

[267] D. Zgid, E. Gull, and G. K.-L. Chan, "Truncated configuration interaction expansions as solvers for correlated quantum impurity models and dynamical mean-field theory," Phys. Rev. B **86**, 165128 (2012).

[268] J. Zhang, M.-H. Yung, R. Laflamme, A. Aspuru-Guzik, and J. Baugh, "Digital quantum simulation of the statistical mechanics of a frustrated magnet." Nat. Comm. **3**, 880 (2012).

[269] S. Zhang and M. H. Kalos, "Exact Monte Carlo calculation for few-electron systems," Phys. Rev. Lett. **67**, 3074–3077 (1991).

[270] S. Zhang and H. Krakauer, "Quantum Monte Carlo method using phase-free random walks with slater determinants," Phys. Rev. Lett. **90**, 136401 (2003).

[271] M. Zheng and G. C. Dismukes, "Orbital configuration of the valence electrons, ligand field symmetry, and manganese oxidation states of the photosynthetic water oxidizing complex: analysis of the S2 state multiline EPR signals," Inorg. Chem. **35**, 3307–3319 (1996).

[272] X.-Q. Zhou, T. C. Ralph, P. Kalasuwan, M. Zhang, A. Peruzzo, B. P. Lanyon, and J. L. O'Brien, "Adding control to arbitrary unknown quantum operations," Nat. Comm. **2**, 1–4 (2011).

[273] M. Ziółkowski, B. Jansík, P. Jørgensen, and J. Olsen, "Maximum locality in occupied and virtual orbital spaces using a least-change strategy," J. Chem. Phys. **131**, 124112 (2009).