QUANTUM INFORMATION AND QUANTUM COMPUTING FOR CHEMICAL SYSTEMS

EDITED BY: Sabre Kais, Travis S. Humble, Karol Kowalski, Ivano Tavernelli,
Philip Walther and Jiangfeng Du
PUBLISHED IN: Frontiers in Physics and Frontiers in Chemistry

frontiers Research Topics

ISSN 1664-8714
ISBN 978-2-88971-509-1
DOI 10.3389/978-2-88971-509-1

About Frontiers

Frontiers is more than just an open-access publisher of scholarly articles: it is a pioneering approach to the world of academia, radically improving the way scholarly research is managed. The grand vision of Frontiers is a world where all people have an equal opportunity to seek, share and generate knowledge. Frontiers provides immediate and permanent online open access to all its publications, but this alone is not enough to realize our grand goals.

Frontiers Journal Series

The Frontiers Journal Series is a multi-tier and interdisciplinary set of open-access, online journals, promising a paradigm shift from the current review, selection and dissemination processes in academic publishing. All Frontiers journals are driven by researchers for researchers; therefore, they constitute a service to the scholarly community. At the same time, the Frontiers Journal Series operates on a revolutionary invention, the tiered publishing system, initially addressing specific communities of scholars, and gradually climbing up to broader public understanding, thus serving the interests of the lay society, too.

Dedication to Quality

Each Frontiers article is a landmark of the highest quality, thanks to genuinely collaborative interactions between authors and review editors, who include some of the world's best academicians. Research must be certified by peers before entering a stream of knowledge that may eventually reach the public - and shape society; therefore, Frontiers only applies the most rigorous and unbiased reviews.
Frontiers revolutionizes research publishing by freely delivering the most outstanding research, evaluated with no bias from both the academic and social point of view. By applying the most advanced information technologies, Frontiers is catapulting scholarly publishing into a new generation.

What are Frontiers Research Topics?

Frontiers Research Topics are very popular trademarks of the Frontiers Journals Series: they are collections of at least ten articles, all centered on a particular subject. With their unique mix of varied contributions from Original Research to Review Articles, Frontiers Research Topics unify the most influential researchers, the latest key findings and historical advances in a hot research area! Find out more on how to host your own Frontiers Research Topic or contribute to one as an author by contacting the Frontiers Editorial Office: frontiersin.org/about/contact

QUANTUM INFORMATION AND QUANTUM COMPUTING FOR CHEMICAL SYSTEMS

Topic Editors:
Sabre Kais, Purdue University, United States
Travis S. Humble, Oak Ridge National Laboratory (DOE), United States
Karol Kowalski, Pacific Northwest National Laboratory (DOE), United States
Ivano Tavernelli, IBM Research - Zurich, Switzerland
Philip Walther, University of Vienna, Austria
Jiangfeng Du, University of Science and Technology of China, China

Citation: Kais, S., Humble, T. S., Kowalski, K., Tavernelli, I., Walther, P., Du, J., eds. (2021). Quantum Information and Quantum Computing for Chemical Systems. Lausanne: Frontiers Media SA. doi: 10.3389/978-2-88971-509-1

Table of Contents

frontiers in Physics

EDITORIAL
published: 01 September 2021
doi: 10.3389/fphy.2021.753618

Editorial: Quantum Information and Quantum Computing for Chemical Systems

Sabre Kais[1]*, Travis Humble[2], Karol Kowalski[3], Ivano Tavernelli[4], Philip Walther[5] and Jiangfeng Du[6]

[1]Chemistry Department, Purdue University, West Lafayette, IN, United States, [2]Oak Ridge National Laboratory, Oak Ridge, TN, United States, [3]Pacific Northwest National Laboratory, Richland, WA, United States, [4]IBM Quantum, IBM Research – Zürich, Rüschlikon, Switzerland, [5]Physics Department, University of Vienna, Vienna, Austria, [6]Physics Department, University of Science and Technology of China, Hefei, China

Keywords: quantum, quantum computing, quantum chemistry, quantum information, quantum science

Editorial on the Research Topic

Quantum Information and Quantum Computing for Chemical Systems

Quantum computing has emerged as an exciting inter-disciplinary research topic that cuts across the traditional fields of physics, computer science, and engineering [1]. It is a revolutionary model of computation that has offered new insights into methods for modeling and simulation of chemical systems [2]. Applications of quantum computing to chemistry have demonstrated rapid progress on both theoretical and experimental fronts [3]. Past theoretical efforts have shown how to adapt quantum computation to a variety of problems including electronic structure and molecular dynamics. In addition, the development of quantum algorithms for quantum chemistry has been stimulated by the greater availability of more capable quantum computing devices [4]. Advances in the number and quality of qubits continues to enable remarkable proof-of-concept demonstrations working towards a milestone of quantum computational advantage [5].

However, many challenges remain along the way toward practical quantum advantages of quantum computing for computational chemistry. Several key examples are the development of new quantum algorithms for solving chemical problems thought to be intractable for classical computers, the efficient representation of those models within quantum computers, the verification and validation of chemical simulations using quantum computers, and the testing and evaluation of real-world problems on currently available devices.

The purpose of this research topic *Quantum Information and Quantum Computing for Chemical Systems* is to present the latest snapshot of leading theoretical concepts, computational methods, and experimental demonstrations of computational chemistry with quantum computers. In the opening article, Yang et al. develop new methods for measuring the time-dependent one-particle probability densities from quantum simulation circuits. They apply these methods to time-dependent density functional theory successfully using numerical simulation of the helium hydride ion with 4- and 8-orbital models. Bylaska et al. demonstrate the advantages of an efficient representation of electronic correlations based on a plane-wave basis for electronic structure calculations of molecular hydrogen. The new correlation optimized virtual orbitals are benchmarked against molecular hydrogen and found to be significantly better at capturing correlation in plane-wave calculations. Claudino et al. address the accuracy of variational quantum algorithms for modeling the exact FCI state of small molecular systems. Their numerical results find that VQE and ADAPT-VQE methods offer similar accuracies with significantly different circuit complexities.

Edited and reviewed by:
Malgorzata Biczysko,
Shanghai University, China

****Correspondence:***
Sabre Kais
kais@purdue.edu

Received: *05 August 2021*
Accepted: *13 August 2021*
Published: *01 September 2021*

Citation:
Kais S, Humble T, Kowalski K, Tavernelli I, Walther P and Du J (2021) Editorial: Quantum Information and Quantum Computing for Chemical Systems.
Front. Phys. 9:753618.
doi: 10.3389/fphy.2021.753618

Several contributions address the model of quantum computation and the extension to new paradigms for computation. This includes the contribution from Wang et al. on the use of qudits for quantum computing, which reviews techniques for building algorithms and circuits that apply to available experimental devices before presenting specifications of leading quantum algorithms using these methods. Dixit et al. present on the advantages of quantum annealing for training restricted Boltzmann machines against Markov chain Monte Carlo, which show promise of reducing the number of iterations needed to this manuscript has been authored by UT-Battelle, LLC under Contract No. DE-AC05-00OR22725 with the U.S. Department of Energy. The United States Government retains and the publisher, by accepting the article for publication, acknowledges that the United States Government retains a non-exclusive, paid-up, irrevocable, world-wide license to publish or reproduce the published form of this manuscript, or allow others to do so, for United States Government purposes. The Department of Energy will provide public access to these results of federally sponsored research in accordance with the DOE Public Access Plan (http://energy.gov/downloads/doe-public-access-plan) realize highly accurate classification. Bhatia present a model for chemical reactions using quantum finite automata using the Belousov-Zhabotinsky reaction network as a demonstration.

Applications to chemistry is one of the fastest growing areas in quantum computing and several contributions provide examples of the many different approaches under development. Komarova et al. discuss how parallelization can be used in quantum computing for simulations of vibrational dynamics. They adapt this approach to a platform of semiconducting quantum dot dimers for computing the mean position and momentum of the vibrations in parallel. Cheng et al. detail approaches to applications of quantum computing for biochemical systems focusing on methods that split the computational overhead between conventional and quantum computers. This "divide and conquer" approach offers a promising set of strategies for exploring problems in shell transition-metal and conjugated pi-electron strongly correlated systems as well as non-covalent bonding. Kardashin et al. give a demonstration of quantum machine learning for constructing a tensor network that approximates the eigenvector of a black box model. This new connection forms a bridge between the language of tensor networks and variational quantum algorithms with plans for near-term demonstrations.

This exciting collection of work highlights the ongoing development of quantum computing for chemistry. Our editorial team for this research topic is sincerely grateful to all the reviewers that offered their perspective and expertise in selecting this outstanding collection of contributions. Combining the expertise of the chemical physics community with insights from the quantum computing community continues to lead to a fresh understanding of important chemical processes and the emergence of the novel many-body formulations.

AUTHOR CONTRIBUTIONS

All authors listed have made a substantial, direct, and intellectual contribution to the work and approved it for publication.

FUNDING

TH and KK were supported by the "Embedding Quantum Computing into Many-body Frameworks for Strongly Correlated Molecular and Materials Systems" project, which is funded by the United States Department of Energy, Office of Science, Office of Basic Energy Sciences, the Division of Chemical Sciences, Geosciences, and Biosciences, and by the Quantum Science Center, a National Quantum Information Science Research Center of the U.S. Department of Energy. SK, TSH and KK also acknowledge the support from the Quantum Science Center, a National Quantum Information Science Research Center of the U.S. Department of Energy.

REFERENCES

1. National Academies of Sciences. *Engineering, and Medicine,* Quantum Computing: Progress and Prospects. Washington, DC: National Academies Press (2019).
2. Cao Y, Romero J, Olson JP, Degroote M, Johnson PD, Kieferová M, et al. Quantum Chemistry in the Age of Quantum Computing. *Chem Rev* (2019) 119(19):10856–915. doi:10.1021/acs.chemrev.8b00803
3. McArdle S, Endo S, Guzik A, Benjamin S, and Yuan X. Quantum Computational Chemistry. *Rev Mod Phys* (2020) 92(1):015003. doi:10.1103/revmodphys.92.015003
4. Bauer B, Bravyi S, Motta M, and Chan GK-L. Quantum Algorithms for Quantum Chemistry and Quantum Materials Science. *Chem Rev* (2020) 120(22):12685–717. doi:10.1021/acs.chemrev.9b00829
5. Arute F, Arya K, Bacon D, Boixo S, Chen Y, Babbush R, et al. Hartree-Fock on a Superconducting Qubit Quantum Computer. *Science* (2020) 369:1084–9. doi:10.1126/science.abb9811

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

ORIGINAL RESEARCH
published: 30 October 2020
doi: 10.3389/fphy.2020.590699

Parallel Quantum Computation of Vibrational Dynamics

Ksenia Komarova[1], *Hugo Gattuso*[2], *R. D. Levine*[1,3,4] *and F. Remacle*[1,2]*

[1]The Fritz Haber Center for Molecular Dynamics and Institute of Chemistry, The Hebrew University of Jerusalem, Jerusalem, Israel, [2]Theoretical Physical Chemistry, UR MolSys B6c, University of Liège, Liège, Belgium, [3]Department of Chemistry and Biochemistry, University of California, Los Angeles, CA, United States, [4]Department of Molecular and Medical Pharmacology, David Geffen School of Medicine, University of California, Los Angeles, CA, United States

The vibrational dynamics in a linear triatomic molecule is emulated by a quantum information processing device operating in parallel. The quantum device is an ensemble of semiconducting quantum dot dimers addressed and probed by ultrafast laser pulses in the visible frequency range at room temperature. A realistic assessment of the inherent noise due to the inevitable size dispersion of colloidal quantum dots is taken into account and limits the time available for computation. At the short times considered only the electronic states of the quantum dots respond to the excitation. A model for the electronic states quantum dot (QD) dimers is used which retains the eight lowest bands of excitonic dimer states build on the lowest and first excited states of a single QD. We show how up to $8^2 = 64$ quantum logic variables can be realistically measured and used to process information for this QD dimer electronic level structure. This is achieved by addressing the lowest and second excited electronic states of the QD's. With a narrower laser bandwidth (= longer pulse) only the lower band of excited states can be coherently addressed enabling $4^2 = 16$ logic variables. Already this is sufficient to emulate both energy transfer between the two oscillators and coherent motions in the vibrating molecule.

Keywords: quantum dots, computing by observables, lie algebra, molecular dynamics, size dispersion, 2D electronic spectroscopy, electronic coherence

Edited by:
Sabre Kais,
Purdue University, United States

Reviewed by:
Heng Fan,
Institute of Physics (CAS), China
Karol Kowalski,
Pacific Northwest National Laboratory (DOE), United States

***Correspondence:**
F. Remacle
fremacle@uliege.be

Specialty section:
This article was submitted to Quantum Computing, a section of the journal Frontiers in Physics

Received: *02 August 2020*
Accepted: *24 September 2020*
Published: *30 October 2020*

Citation:
Komarova K, Gattuso H, Levine RD and Remacle F (2020) Parallel Quantum Computation of Vibrational Dynamics. Front. Phys. 8:590699. doi: 10.3389/fphy.2020.590699

INTRODUCTION

We describe the theoretical background for an experimental setup, an ensemble of quantum dot dimers that can and has been realized in the laboratory. We show explicitly how this device is used to emulate the quantum vibrational dynamics of a linear triatomic molecule and discuss possible extensions. In 1985 Deutsch defined a quantum computer as a device that could simulate effectively an arbitrary physical system [1]. Our aim here is much more modest. We seek to describe a device that can be realized with currently available laboratory techniques. Furthermore, the device needs to provide computational answers only for a limited set of variables of the physical system. The computation is realized by mapping of the dynamics of the physical variables of this limited set using a set of observations of the time-evolution of the device. The set of possible observations of the device is the set of our N^2 logic variables. The number N is less than or equal to the number of accessible quantum states of the logic device. N^2 is larger than the number of variables of interest for the physicochemical system that is emulated.

The quantum mechanical device is characterized by a set of observables and that is why we call the logic done by the device "computing by observables" [2, 3]. With such a device one can emulate a physical system that requires up to this number for simulating a closed set of physical variables.

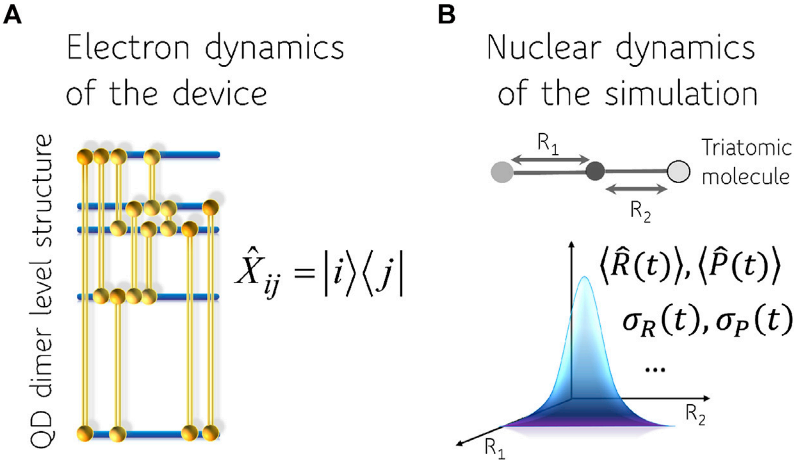


FIGURE 1 | Representation of the computing device **(A)** and the physical system **(B)**. The four bands of excited electronic states of the device that are used are indicated as well as the ten pairs i, j whose twenty corresponding coherences X_{ij} are used as logic variables. The physical system we emulate is a vibrating triatomic molecule in a non-stationary quantum mechanical state. We will simulate the mean position and momentum of each of the two coupled oscillators as well as the respective widths, σ_R and σ_P.

Increasing levels of resolution provided by the description of a physicochemical system can be characterized by an increasing numbers of variables. An example used in this paper is that of two coupled harmonic oscillators. For each oscillator there is a denumerable number of bound states. So, in principle, a fully complete characterization of each oscillator can require up to a denumerably infinite number of variables. An opposite extreme is a thermal state of the oscillator. Its complete description requires two physical variables, a normalization of the probabilities and the mean energy. For the coupled oscillators we will describe several physical levels of resolution that are intermediate between these two limits.

The variables that we deal with can be regarded as vectors in a linear space. They can, if so desired, be made orthogonal etc. As vectors, the variables seem to behave like classical variables but they are a quantum mechanical description of the system. In textbooks one is more used to a linear vector space of wave functions, what is sometimes called a Hilbert space. The observables are also defined in a space with a scalar product and a Mathematician may choose to call it a Hilbert space. But if the Hilbert space of the wave functions has N dimensions, the corresponding linear vector space of observables has N^2 dimensions. If $\{|i\rangle\}$ is a set of basis vectors in the Hilbert space of wave functions, $i = 1, 2, .., N$, a basis set for the observables can be written as $\{X_{ij} = |i\rangle\langle j|\}$, $i, j = 1, 2, .., N$. The N^2 expectation values of these are our variables. We will refer to the expectation values when $i \neq j$ as coherences. The N diagonal elements are populations. By taking linear combinations, e.g., $|i\rangle\langle j| + |j\rangle\langle i|$, we can arrange all the variables to be Hermitian and all the variables to be real.

In this paper we use two distinct linear vector spaces. One is suitable to describe the dynamics of the computing device and one the dynamics of the physical system we aim to emulate. We do not consider the device as providing an analog for the physical system. To emphasize this point we use a device that operates on the dynamics of purely electronic states pumped and read by ultrafast laser pulses. A schematic description of the device and the physical system is shown in **Figure 1**. The dynamics in the electronic Hilbert space of the quantum dot dimers is mapped to simulate the time evolving physical variables in the nuclear Hilbert space. We will therefore need a separate discussion of the logic variables provided by the computing device and those required to describe the physical system.

Quantum computing algorithms have been developed for problems of chemical interest. This includes not only electronic structure [4–13] but also quantum enhanced machine learning algorithms [8] and algorithms for dimension reduction including Principal Component Analysis, Canonical Correlation Analysis and other algebraic methods used for dimension reduction such as surprisal analysis [6, 15–17] A preliminary report on our quantum computation has been published [18].

THE COMPUTING DEVICE

Our computing device is an array of CdSe quantum dot QD dimers. Addressing the device is by a sequence of laser pulses and read-out is performed simultaneously on many dots [19–24] as in 2D electronic spectroscopy [25, 26]. Measuring over a classical ensemble of dots importantly means that we directly read the mean values and that there is no interference between measuring different observables that individually do not commute. In the short time interval during which the time and probe are performed the primary source of noise is the variability in the size of the quantum dots. This means that the two dots making the dimer are not quite identical so strictly speaking we have heterodimers. We therefore need to average the read-out over the distribution of dimer sizes. The states of different dimers are at slightly different energies so that the coherences between different pairs of states beat at slightly different frequencies. This reduces the frequency resolution or, equivalently, the time available for read-out [13] to faster than the dephasing time of the mixture of not fully identical dimers. So we do not read individual quantum states but bands of closely adjacent states belonging to different dimers. In our computations we use an experimentally realizable size of the dots and their size dispersion.

In this paper we employ rather elementary dynamics of the computing device. Using a sequence of laser pulses we bring each dimer to a multicoherent state. From then on the device is unperturbed and it is the oscillations of different coherences that are used to simulate the time dependence of the physical system. The coherence between the ground state and any electronic excited state oscillates quite fast so it dephases faster then the dephasing of the coherences between the electronically excited states. The very fast beating coherence, few fs's or less, between the ground state and an electronic excited state cannot therefore be reliably detected. So very fast oscillations cannot be simulated by our device due to the currently available dispersion in sizes of colloidal quantum dots. It is expected that a lower dispersion will be experimentally possible in the near future.

This paragraph is a sketch of the counting of electronically excited states in a CdSe dot of nanometric mean size between 2 and 5 nm in diameter and of a dimer of two dots drawn from an ensemble of dots of the same mean size. The two dots making up the dimer are of about the same but not identical size. The first electronically excited state of a dot is a band of 12 states, as follows. The spin of the hole is coupled to a p ($l = 1$) orbital of Se to give rise to two spin orbit components, 1/2 and 3/2. The s orbital on Cd contributes only a spin of 1/2. The total orbital angular momentum F of the excitonic 3/2 state has the projections $F_z = \pm 2, \pm 1, 0$ and $F_z = \pm 1, 0$ for the hole and the electron respectively, eight states all together. There are four states, ($F_z = \pm 1, 0$ and 0) for the 1/2 excitonic state. The two excitonic bands of states, (of eight and four states, respectively), can be experimentally resolved [23]. There are similarly two resolvable bands of states for the second electronically excited state of CdSe. All together, four bands of electronically excited states per dot. In the dimer, or dots with short ligands, the dots are interacting by Coulomb and exchange coupling. Thereby each one of the four bands of states of the monomer is split into two. In the visible range a dimer has eight resolvable bands of states, four bands that correspond primarily to the lower electronic state and four to the higher excited state of a monomer. Since the dimers are made of two dots that are not quite identical, all eight bands are optically active from the common dimer ground state, see **Supplementary Materials, section 1.1** for more details.

Using the counting as above the CdSe dots enable a choice of four or eight bands of states in the visible range. Addressing coherently the states of all the eight bands requires shorter laser pulses. Smaller QD have larger energy difference between their two lower excited electronic states and therefore one needs a larger energy bandwidth of the pulse or a higher carrier frequency to simultaneously address them. On the other hand, for larger dots, the energy difference between the two lowest excited states may become smaller or comparable to strength of the spin-orbit coupling, which leads to a loss of resolution between the different bands [23, 24]. Commercially available fast lasers in the visible can easily address coherently all the states of the four lowest bands. The minimal capability of our device is therefore 5^2 logic variables, the five populations of the ground and four excited states and the 5×4 coherences. Four (times two, complex values) coherences of the transitions between the ground state and the four excited states and 12 coherences between the four excited states. After the laser pulses are over, each coherence will oscillate with a fixed frequency determined by the differences in energy of the two states. The highest frequencies are for the transitions up from the ground state. These are rather fast. So in this paper we consider the coherences between pairs of excited states.

To describe the dynamics of the device we take it that initially it is in the ground electronic state. We assume that the addressing lasers are weak enough so that only one photon transitions are possible. The lasers are in the visible so that transitions between the excited states are way out of resonance and so are excluded. Indeed and as we shall discuss, the frequencies of the transitions between the excited states are in the range of molecular vibrational frequencies so that we can use the coherences between excited states to emulate the time scales relevant to the physical systems. The range of timescales is determined by the size of the quantum dot dimers as is discussed above and by the coherence width of the excitation lasers. The bandwidth of a 6 fs pulse is about 2,100 cm^{-1}.

The excitation scheme is that typically used in 2D electronic spectroscopy [25, 26]. The first fast laser pulse generates the absorption from the ground state. Next, with some delay, is a second laser pulse. One frequency axis that we will use is the Fourier transform with respect to this time delay. The system is next allowed to evolve for a time interval that is typically denoted T. After the second pulse, the system can be back in the ground state, in one of the excited states, or in a coherence between two excited states. After the time T the third laser pulse stimulates emission that is monitored in time and the second frequency axis is a Fourier transform with respect to this time. So for each value of the time T we generate a 2D frequency map. In the echo phase matching direction of emission, the populations of the excited states appear on the diagonal of the 2D maps and the coherences are on the off-diagonal at a position determined by the excitation frequency of the two excited states they connect. At different values of the time interval T, the intensities at the off-diagonal positions on the map will vary according to the time-evolution of the respective coherences, and this is how we can simulate the time dependence of the physical variables. As emphasized above, due to the size dispersion the coherence beating frequencies will have a finite spread. The ability to resolve coherences between different pairs of states depends on this spread to be limited enough. From the complementary time-dependent point of view, the coherence contributions decay in time as a Gaussian with a width that is governed by the width of the frequency distribution of the coherence [3]. On the plus side, the finite spread of the frequency of every coherence means that the computing device can simulate finite spans of values, roughly the width of the Gaussian, about different frequencies.

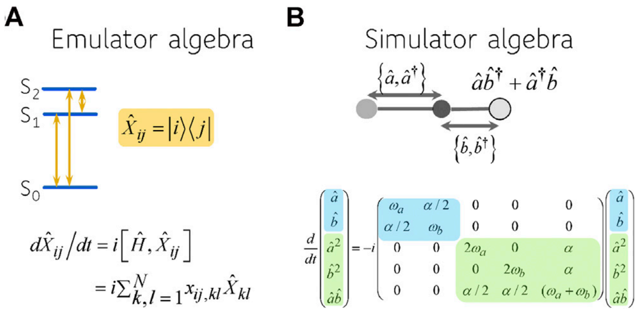


FIGURE 2 | The algebra characterizing the device **(A)** and the physical system **(B)**. The coherences $X_{ij} = |i\rangle\langle j|$ between excited levels *i* and *j* of the quantum dot dimer are our logic variables. After the laser pulses are over, the coherences oscillate in time with a frequency that is the difference between the energies of the two states that they connect. We aim to simulate the time dependence of observables of the coupled oscillators system. For example, as shown in **Eq. 7**, we need the time dependence of $\langle a^2\rangle$ to compute the variance of the bond displacement of the first oscillator.

Deterioration of the signal due to dephasing has been an issue for quantum computing from its early beginning [27, 28]. Often it is dephasing due to interaction with the environment. Here too, the limiting factor is dephasing but it is a dephasing due to an inhomogeneous broadening caused by the size dispersion of the QD's. In the short times that we probe, the electronic states of each dimer practically are not yet perturbed. The dephasing is due to the probe averaging over many dimers, dimers with slightly different frequencies [3]. We pay special care to make the dephasing realistic. The mean diameter of the CdSe QD's, 4.4 nm, that we use to compute the spectrum and the size dispersion, 5%, have been realized experimentally [23, 29]. We emphasize the considerable improvement in performance that can be achieved by an even modest reduction in the dispersion in sizes. Yet even the already accessible 5% size variability is sufficient to enable a quite powerful device. We also stress that the 2DES measurements that we rely upon are performed at room temperature, in solution [23] or on a solid state device [24].

We can emulate dynamics by measuring 2D coherence maps at a range of values for the time interval *T*. With current chemical synthetic capabilities the dephasing due to the size dispersion of the dot is the primary limitation on how long *T* can be. At longer times, say beyond 100 fs, coupling to the phonons also sets in.

For each experiment there will be a sequence of transitions leading to the same coherence between two excited states during the delay time *T* between the second and the third pulse. For weak pulses so that first order time-dependent theory and the impulsive limit transition applies at each interaction with the pulse, the two coordinates of the 2D maps are the absorption frequency from the ground state (GS) and the emission frequency to GS. When two excited states, *i* and *j*, can be reached by one photon transitions from the ground state (GS) there are four positions in the 2D maps, two diagonal ones at $(\omega_{0i}, \omega_{0i})$ and $(\omega_{0j}, \omega_{0j})$ and two off-diagonal ones $(\omega_{0i}, \omega_{0j})$ and $(\omega_{0j}, \omega_{0i})$ where ω_{0i} and ω_{0j} are the excitation frequencies of states *i* and *j* from the GS. In the rephasing phase matching (echo) direction, eight Liouvillian paths contribute to these four positions. They are spelled out explicitly in **Supplementary Figure S3** in the notation of ref. [26], see also refs. [2, 30].

Four paths fall on the diagonal of the 2D map at the excitation frequencies of the two excited states: one ground state bleaching (GSB) path for which the system is the GS during the time interval *T* and one stimulated emission paths (SE) for which the system is in excited state *i* during *T* at $(\omega_{0i}, \omega_{0i})$ and a GSB and a SE path for which the system is in excited state *j* at $(\omega_{0j}, \omega_{0j})$. At the short times below 200 fs considered here, there is no exchange of population between the GS and the two excited states during *T* and the contribution of these four paths to the 2D response is time-independent. Two time-dependent GSB paths contribute to the two off diagonal positions, $(\omega_{0i}, \omega_{0j})$ and $(\omega_{0j}, \omega_{0i})$, respectively. Two paths that are the signature of the electronic coherence between the two excited states also contribute at the same off diagonal positions. The two coherence paths are beating during *T* with the transition frequency ω_{ij} of the coherence between the excited states *i* and *j*.

The eight paths, see **Supplementary Figure S3**, contribute simultaneously to the time evolution of the density matrix and therefore to the 2D map, hence "parallel" in our title. When the system has more than two excited states, each pair of excited states contributes 8 paths to the time evolution of the density matrix. For example, there are 16 coherences between excited states for the five level system (GS + 4 excited states) discussed above, meaning that 16x8 paths contribute in parallel. Of these, only the 16x2 paths that lead to a distinct off diagonal position on the 2D maps are used to map observables of the system that is emulated. There are also 8 paths per pair of excited states in the non rephasing phase matching direction. So in principle, the larger the number of states in the band that can be accessed by one photon transition from the GS, the larger the number of observables that can be emulated.

As explained above, colloidal QD's are synthesized with a finite size dispersion which leads to a distribution of transition frequencies for each coherence, both from the GS, the ω_{0i} type and between excited states (the ω_{ij} type). At the level of the ensemble, the inherent size dispersion of the QD's leads to a Gaussian distribution in energy of the addresses $(\omega_{0i}, \omega_{0j})$ of coherences between excited states and to their Gaussian dephasing along *T* and therefore limits the number of coherence positions that can be resolved on the map. When controlled and limited to a few percent in diameter, the size dispersion can also be used to advantage. Scanning positions on the 2D maps around the address corresponding to the mean transition frequencies from the GS gives slightly different periods of the coherence along T which provides flexibility in mapping the periods of the observables of the emulated systems.

Figure 1 shows a scheme where there are four bands of excited states that are accessed by the laser pulses. The four coherences at a relatively high frequency when the ground state is one of the two states are not detected in 2D spectroscopy. There are the twelve lower frequency coherences that, as mentioned, will be used to emulate the vibrational motions of the physical system. Note that there is a spread in the frequencies and the very lower ones will be used to emulate splittings due to coupling of two vibrational modes.

THE PHYSICAL SYSTEM: LINEAR VIBRATING TRIATOMIC MOLECULE

In this paper we track the electronic dynamics in the QD dimers to emulate the vibrational motion in non-stationary states of a triatomic molecule, **Figure 2**.

The system we emulate is two coupled harmonic local vibrations, each representing a bond. The coupling is due to the motion of the central atom and it depends on its mass and its displacement from equilibrium [31]. By expanding the potential of the system and keeping only up to quadratic terms in the deviation of the bond displacements from equilibrium, the Hamiltonian for two harmonically coupled harmonic modes, denoted a and b, is, using $\hbar = 1$,

$$\hat{H} = \omega_a \hat{a}^\dagger \hat{a} + \omega_b \hat{b}^\dagger \hat{b} + (\alpha/2)\left(\hat{a}^\dagger \hat{b} + \hat{b}^\dagger \hat{a}\right) \tag{1}$$

Here we use the standard notation for the creation and annihilation operators for the two local vibrational modes. The vibrational quantum number mismatch is determined by $\langle \hat{a}^\dagger \hat{a} - \hat{b}^\dagger \hat{b} \rangle$ and the frequency mismatch is $\delta\omega = \omega_a - \omega_b$. The harmonic coupling is $\hat{a}\hat{b}^\dagger + \hat{a}^\dagger \hat{b}$ and the strength is $\alpha/2$. Due to this coupling, the creation and annihilation operators of the two oscillators are correlated.

We characterize vibrational motion of the local modes by the mean values of the bond distance, momenta and dispersion of both quantities. Computation of the values of the mean bond distance and mean momenta involves description of the evolution of both creation and annihilation operators: $\langle \hat{R}_a \rangle = \langle \hat{a} + \hat{a}^\dagger \rangle / \sqrt{2 m_a \omega_a}$. A simple application of the Heisenberg equation of motion leads to a set of equations of motion:

$$\frac{d}{dt}\begin{pmatrix} \hat{a} \\ \hat{b} \end{pmatrix} = -i \begin{pmatrix} \omega_a & \alpha/2 \\ \alpha/2 & \omega_b \end{pmatrix} \begin{pmatrix} \hat{a} \\ \hat{b} \end{pmatrix} \tag{2}$$

and an adjoint equation:

$$\frac{d}{dt}\begin{pmatrix} \hat{a}^\dagger \\ \hat{b}^\dagger \end{pmatrix} = i \begin{pmatrix} \omega_a & \alpha/2 \\ \alpha/2 & \omega_b \end{pmatrix} \begin{pmatrix} \hat{a}^\dagger \\ \hat{b}^\dagger \end{pmatrix} \tag{3}$$

One can write these as one equation for a vector of four components using a Liouvillian operator $\hat{L}$: $d\mathbf{v}/dt = i\mathbf{L}\mathbf{v}$, where $\mathbf{v}^+ = (\hat{a}, \hat{b}, \hat{a}^\dagger, \hat{b}^\dagger)$ and Liouvillian:

$$\mathbf{L} = \begin{pmatrix} -\omega_a & -\alpha/2 & 0 & 0 \\ -\alpha/2 & -\omega_b & 0 & 0 \\ 0 & 0 & \omega_a & \alpha/2 \\ 0 & 0 & \alpha/2 & \omega_b \end{pmatrix} \tag{4}$$

As is well known, one can diagonalize the Hamiltonian leading to an antisymmetric and a symmetric vibrational modes. Here this corresponds to diagonalizing the Liouvillian, **Eq. 4**, leading to the frequencies:

$$(1/2)(\omega_a + \omega_b \pm \Omega) \tag{5}$$

where $\Omega^2 = \alpha^2 + (\delta\omega)^2$.

We use as an example vibrational dynamics in the CS_2 molecule, where $\omega_a = \omega_b = \omega$ and $\Omega = \alpha$. The eigenfrequencies, **Eq. 5**, become $\omega \pm \alpha/2$ for the $\hat{a} \pm \hat{b}$ eigenvectors and similarly for their complex conjugate analogs. Taking expectation values it is useful to note that equations imply that the mean values at the time t can be computed given their initial values, for example:

$$\begin{aligned} \langle \hat{a} + \hat{b} \rangle(t) &= \langle \hat{a} + \hat{b} \rangle(0)\exp(-i(\omega + \alpha/2)t) \\ \langle \hat{a}^\dagger + \hat{b}^\dagger \rangle(t) &= \langle \hat{a}^\dagger + \hat{b}^\dagger \rangle(0)\exp(i(\omega + \alpha/2)t) \end{aligned} \tag{6}$$

The mean values of the bond displacement, $(\langle \hat{a} \rangle(t) + \langle \hat{a}^\dagger \rangle(t))/\sqrt{2m\omega}$ or the momentum $i(\langle \hat{a}^\dagger \rangle(t) - \langle \hat{a} \rangle(t)) \cdot \sqrt{m\omega}/\sqrt{2}$ can thereby be computed.

To describe the dispersion of the nuclear wave packets in coordinate space and in momentum space we need to extend our algebra. In the basis of the vibrational states the dispersion in the displacement of a local oscillator can be written as follows:

$$\begin{aligned} \sigma_R^2(t) &= \langle R^2 \rangle(t) - (\langle R \rangle(t))^2 \\ &= \frac{1}{2m\omega}\langle (\hat{a} + \hat{a}^\dagger)^2 \rangle(t) - \frac{1}{2m\omega}(\langle (\hat{a} + \hat{a}^\dagger) \rangle(t))^2 \end{aligned} \tag{7}$$

And similar for the dispersion in the momentum space:

$$\sigma_P^2 = \langle P^2 \rangle - \langle P \rangle^2 = -\frac{m\omega}{2}\langle (\hat{a}^\dagger - \hat{a})^2 \rangle + \frac{m\omega}{2}\langle (\hat{a}^\dagger - \hat{a}) \rangle^2 \tag{8}$$

Therefore we need to describe the dynamics of the $\{\hat{a}^2, (\hat{a}^\dagger)^2, \hat{a}^\dagger \hat{a}\}$ and $\{\hat{b}^2, (\hat{b}^\dagger)^2, \hat{b}^\dagger \hat{b}\}$ for the dispersion in R and P of the two coupled local bonds. Note connection between the dispersion of the wave packets in different representations to the energy transfer between the two coupled oscillators, $\langle \hat{a}^\dagger \hat{a} \rangle$ or $\langle \hat{b}^\dagger \hat{b} \rangle$ mean values. Therefore we need a hardware computing device that can emulate 10 logic variables. The computing hardware described above can emulate 16 logic variables when only the lowest four exciton bands are addressed and 64 when shorter laser pulses are used.

RESULTS

We use the device as discussed in section above to simulate the time dependence of the physical variables as revealed by the algebras for the coupled vibrations that are discussed in **section III**. Each coherence of the device is a point on the 2D frequency map generated by the 2D spectroscopy. The size dispersion of the quantum dots means that each point is actually a cloud of points at nearby frequencies. This dispersion in the frequency associated with each coherence is what enables one device to simulate different but similar physical systems. In this section we show how different coherences can simulate the time evolution of different physical variables of the system.

The results for the mean bond displacement $\langle R \rangle(t)$ and its dispersion $\sigma_R^2(t)$, **Eq. 7**, are shown in panels (a) and (c) of **Figure 3**. The computations are for a symmetric system that is not stationary in time because of an asymmetry in the initial conditions. One oscillator starts with low energy and the other, the one shown, is initially energy richer. On the right side of the figure we show relevant segments of 2D frequency maps computed at different values of the time interval T that enable the device to simulate the two functions of time $\langle R \rangle(t)$ and $\sigma_R^2(t)$.

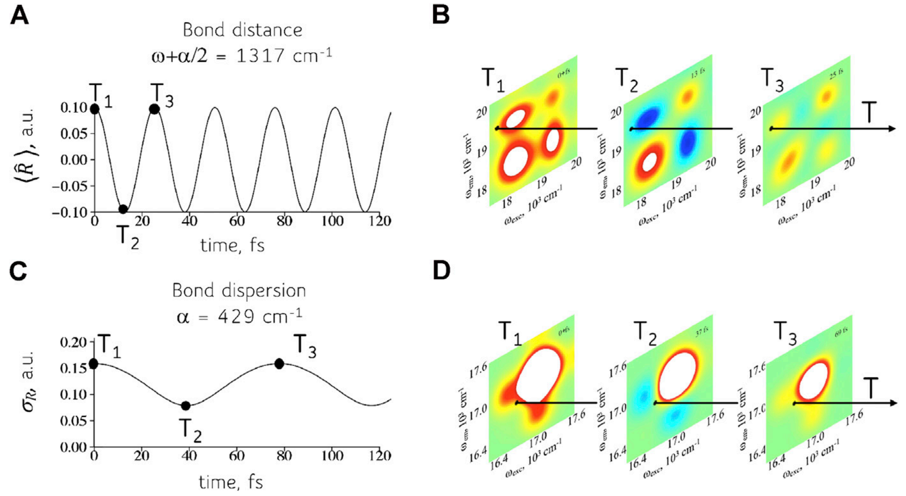


FIGURE 3 | Time-evolution of the physical variables **(A, C)** and their computation **(B, D)** using 2D frequency maps for the device of the QD dimers. The response of the quantum device, panels **(B)** and **(D)**, can simulate the time dependent observables of the physical system such as the mean bond displacement, panel **(A)**, and the dispersion in the bond displacement, panel **(C)**. The physical system shown is a symmetric triatomic molecule. Shown are the real part of the 2D maps. See **Supplementary Material, Section 1.2**, for details about the computation of the 2D maps.

Shown for each are three values of T that correspond to extreme displacement and dispersion. One can, of course, measure at intermediate values as well.

The mean bond displacement of a local oscillator, frequency ω, in a symmetric molecule such as CO_2, **Figure 3A**, is seen to be faster varying in time than the width, 3(c), of the wavepacket describing the oscillator. This reflects the difference in magnitude between the eigenfrequencies $\omega \pm \alpha/2$ and the coupling parameter α. In the absence of coupling between the oscillators, $\alpha = 0$ in the Hamiltonian, **Eq. 1**, the width of a coherent state of the harmonic potential will not vary with time. The coupling induces a slow, frequency α, energy transfer between the two oscillators as reflected in the variation of the width of the wavepacket localized on one of the two bonds.

It might seem from **Figure 3** that one has to assemble a new device for each different triatomic molecule. This is because for each one of the two frequencies $\omega + \alpha/2$ and α one would need a different electronic level structure in the dimer of the two quantum dots. The same size dispersion of the quantum dots that hitherto limited our abilities is here an advantage. The somewhat different sizes of the monomers means that there is a finite range of frequencies for each coherence. The range should not be too broad as otherwise we will lose resolution in frequency. Moreover, since the distribution of spacings of a particular coherence is, at low dispersions, Gaussian [3], we should not probe too far toward the wings. Complementary, when the range is broad we can only measure at short times before serious dephasing sets in. But the times cannot be too short because we want to measure only after the laser pulses are over. Practical implications of the distribution of spacings can be seen in **Figure 4**. Shown is a 2D frequency map for both regions used to compute the time-evolution of the bond distance (region A of **Figure 4**) and its dispersion (region B, **Figure 4**). Slightly different time-evolution of the coherences for the neighboring positions on the map enables fine-tuning of the period of the oscillations as is shown on the side panels of **Figure 4**.

A coherence is a complex valued observable so each coherence can describe two conjugate physical variables as shown in **Figure 5**. **Equation 6** shows that the expectation values, $\langle R\rangle(t) = (\langle\hat{a}\rangle(t) + \langle\hat{a}^\dagger\rangle(t))/\sqrt{2m\omega}$ and $\langle P\rangle(t) = i(\langle\hat{a}^\dagger\rangle(t) - \langle\hat{a}\rangle(t))\cdot\sqrt{m\omega}/\sqrt{2}$ of the bond displacement and of the momentum can both be computed at the same point of the map, by reading the real and imaginary values. The two values are shifted by a phase difference of $\pi/2$ as is to be expected. There is an uncertainty with the momentum, σ_P, and position, σ_R, as shown in **Eqs. 7 and 8**. Time-evolution of the uncertainty in the position, related to the width of the wavepacket in coordinate representation, is shown in **Figures 3C,D**. As discussed therein these uncertainties are time-dependent because of the coupling between the two oscillators so that they are more slowly varying and reflect the energy transfer between the two oscillators. The time-evolution of the uncertainty in the momentum representation is following the same time-dependence as σ_R therefore no additional computation is needed.

CONCLUDING REMARKS

A versatile quantum mechanical computing device that operates on a laser addressed solid array of quantum dots dimers has been discussed. Experimental data shows that such a device can operate at room temperature [23, 24]. The device was effectively used to simulate the quantal dynamics of non-stationary states of coupled vibrations. We show how the read-out of the coherences engineered in the device following interaction with a sequence of the laser pulses enables parallel

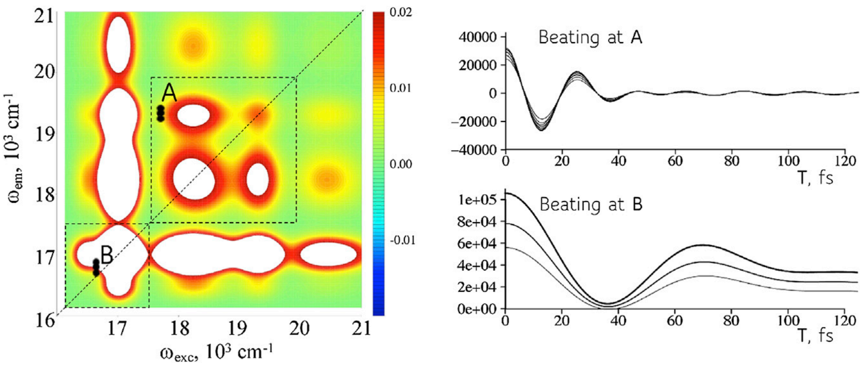


FIGURE 4 | The same device can simulate a range of frequencies. **Left:** A 2D frequency map (real part) for the device at a particular value of the time interval *T* (computed for the full frequency range of the excited states, **Supplementary Figure S4**). The dispersion in sizes means that there is a range of frequencies corresponding to each coherence. Point A. A higher frequency coherence, more off the diagonal in the map, used to compute the time-evolution of the bond distance, **Figures 3A,B**. Shown are the intensities of the coherence vs time *T* for points in the range shown on the map. Point B. Coherence at a lower frequency for the computation of the bond dispersion, **Figures 3C,D**. A wider range in the frequency is possible but reading nearer to the diagonal is experimentally more challenging.

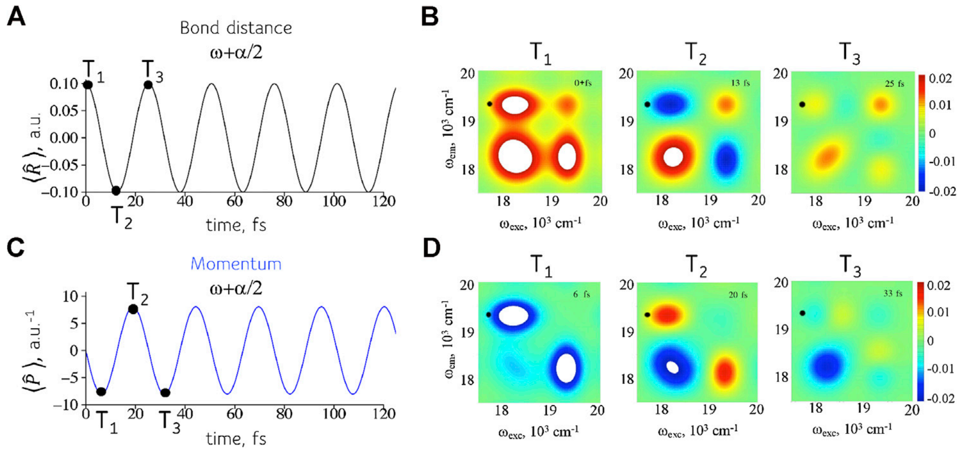


FIGURE 5 | Reading two conjugate variables at the same position on the 2D frequency maps. **Left entries:** mean position **(A)** and momentum **(C)** as given by **Eq. 6**. As expected the displacement is maximal at a turning point of the motion while the momentum is maximal when the wavepacket crosses the equilibrium position of the oscillator. **Right entries:** real **(B)** and imaginary **(D)** values of the intensities in the frequency map in the frequency range of interest. The values shown are reading at different time intervals *T*.

computation of the two conjugated variables, mean position and momentum of the local oscillators. Extension to the more detailed dynamics is shown by simulating also the dispersion in coordinate representation of the quantum wave packets of the coupled oscillators. Simulation is provided by matching the beating frequency of the coherence and the oscillation frequency of the physical variable. This mapping is in principle always possible for systems where the algebraic description of the physical variables is closed upon their commutation with Hamiltonian. Diagonalization of the respective Liouvillian operator, **Eq. 4**, determines eigenfrequencies ω_k, for example **Eq. 5**, and enables transformation to observables $A_k(t)$ that evolve in time as $A_k(0)\exp(i\omega_k t)$. It is these frequencies that need to be measured by the time evolution of the coherences of the device. The inevitable size dispersion of the quantum dots limits the span of time that is available for emulation but on the other hand allows a fine-tuning of a frequency of interest.

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

AUTHOR CONTRIBUTIONS

RDL and FR designed the research; KG, RDL and FR performed the research and wrote the ms, KG, HG and FR carried out the computations.

ACKNOWLEDGMENTS

We thank Prof. Elisabetta Collini for many incisive discussions about the experiments on the device. This work was supported by the FET open EC project COPAC #766563 and the Fonds National de la Recherche Scientifique (F.R.S-FNRS, Belgium), #T.0205.20 and #J.0012.18. Computational resources have been provided by the COPAC project and by the Consortium des Equipements de Calcul Intensif (CECI), funded by the F.R.S.-FNRS under Grant #2.5020.11.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fphy.2020.590699/full#supplementary-material

REFERENCES

1. Deutsch D. Quantum theory, the Church–Turing principle and the universal quantum computer. *Proc. R. Soc. A* (1985) **400**(1818):97–117. doi:10.1098/rspa.1985.0070.
2. Fresch B, Hiluf D, Collini E, Levine RD, Remacle F. Molecular decision trees realized by ultrafast electronic spectroscopy. *Proc Natl Acad Sci USA* (2013) **110**(43):17183–8. doi:10.1073/pnas.1314978110.
3. Gattuso H, Levine RD, Remacle F. Massively parallel classical logic via coherent dynamics of an ensemble of quantum systems with dispersion in size. *Proc Natl Acad Sci USA* (2020) **117**(35):21022. doi:10.1073/pnas.2008170117.
4. Smith A, Kim MS, Pollmann F, Knolle J. Simulating quantum many-body dynamics on a current digital quantum computer. *Npj Quantum Inf* (2019) **5**(1):106. doi:10.1038/s41534-019-0217-0.
5. Cao Y, Romero J, Olson JP, Degroote M, Johnson PD, Kieferová M, et al. Quantum chemistry in the age of quantum computing. *Chem Rev* (2019) **119**(19):10856–915. doi:10.1021/acs.chemrev.8b00803.
6. Xia R, Kais S. Quantum machine learning for electronic structure calculations. *Nat Commun* (2018) **9**(1):4195. doi:10.1038/s41467-018-06598-z.
7. Argüello-Luengo J, González-Tudela A, Shi T, Zoller P, Cirac JI. Analogue quantum chemistry simulation. *Nature* (2019) **574**(7777):215–8. doi:10.1038/s41586-019-1614-4.
8. Bian T, Murphy D, Xia R, Daskin A, Kais S. Quantum computing methods for electronic states of the water molecule. *Mol Phys* (2019) **117**(15-16):2069–82. doi:10.1080/00268976.2019.1580392.
9. Kandala A, Mezzacapo A, Temme K, Takita M, Brink M, Chow JM, et al. Hardware-efficient variational quantum eigensolver for small molecules and quantum magnets. *Nature (Lond)* (2017) **549**(7671):242–6. doi:10.1038/nature23879.
10. Hempel C, Maier C, Romero J, McClean J, Monz T, Shen H, et al. Quantum chemistry calculations on a trapped-ion quantum simulator. *Phys. Rev. X* (2018) **8**(3):031022. doi:10.1103/PhysRevX.8.031022.
11. Yung MH, Casanova J, Mezzacapo A, McClean J, Lamata L, Aspuru-Guzik A, et al. From transistor to trapped-ion computers for quantum chemistry. *Sci Rep* (2014) **4**(1):3589. doi:10.1038/srep03589.
12. Xu K, Chen J-J, Zeng Y, Zhang Y-R, Song C, Liu W, et al. Emulating many-body localization with a superconducting quantum processor. *Phys Rev Lett.* (2018) **120**(5):050507. doi:10.1103/PhysRevLett.120.050507.
13. Google. Hartree-Fock on a superconducting qubit quantum computer. *Science* (2020) **369**(6507):1084. doi:10.1126/science.abb9811.
14. Ciliberto C, Herbster M, Ialongo Alessandro D, Pontil M, Rocchetto A, Severini S, et al. Quantum machine learning: a classical perspective. *Proc R Soc A* (2018) **474**(2209):20170551. doi:10.1098/rspa.2017.0551.
15. Rebentrost P, Steffens A, Marvian I, Lloyd S. Quantum singular-value decomposition of nonsparse low-rank matrices. *Phys Rev A* (2018) **97**(1):012327. doi:10.1103/PhysRevA.97.012327.
16. Schuld M, Sinayskiy I, Petruccione F. Prediction by linear regression on a quantum computer. *Phys Rev A* (2016) **94**(2):022342. doi:10.1103/PhysRevA.94.022342.
17. Alhassid Y, Levine RD. Connection between the maximal entropy and the scattering theoretic analyses of collision processes. *Phys Rev A* (1978) **18**(1):89–116. doi:10.1103/PhysRevA.18.89.
18. Komarova K, Gattuso H, Levine RD, Remacle F. Quantum device emulates the dynamics of two coupled oscillators. *J Phys Chem Lett* (2020) **11**:6990–5. doi:10.1021/acs.jpclett.0c01880.
19. Kambhampati P. Unraveling the structure and dynamics of excitons in semiconductor quantum dots. *Acc Chem Res* (2011) **44**(1):1–13. doi:10.1021/ar1000428.
20. Caram JR, Zheng H, Dahlberg PD, Rolczynski BS, Griffin GB, Dolzhnikov DS, et al. Exploring size and state dynamics in CdSe quantum dots using two-dimensional electronic spectroscopy. *J Chem Phys* (2014) **140**(8):084701. doi:10.1063/1.4865832.
21. Turner DB, Hassan Y, Scholes GD. Exciton superposition states in CdSe nanocrystals measured using broadband two-dimensional electronic spectroscopy. *Nano Lett* (2012) **12**(2):880–6. doi:10.1021/nl2039502.
22. Cassette E, Dean JC, Scholes GD. Two-dimensional visible spectroscopy for studying colloidal semiconductor nanocrystals. *Small* (2016) **12**(16):2234–44. doi:10.1002/smll.201502733.
23. Collini E, Gattuso H, Bolzonello L, Casotto A, Volpato A, Dibenedetto CN, et al. Quantum phenomena in nanomaterials: coherent superpositions of fine structure states in CdSe nanocrystals at room temperature. *J Phys Chem C* (2019) **123**:31286–93. doi:10.1021/acs.jpcc.9b11153.
24. Collini E, Gattuso H, Kolodny Y, Bolzonello L, Volpato A, Fridman HT, et al. Room-temperature inter-dot coherent dynamics in multilayer quantum dot materials. *J Phys Chem C* (2020) https://doi.org/10.1021/acs.jpcc.1020c05572. doi:10.1021/acs.jpcc.0c05572.
25. Cho M. *Two-Dimensional Optical Spectroscopy*. Boca Raton: CRC Press (2009)
26. Mukamel S. *Principle of non linear optical spectroscopy*. Oxford: Oxford University Press (1995).
27. Nielsen MA, Chuang IL. *Quantum Computation and Quantum Information*. Cambridge: Cambridge University Press (2010).
28. Preskill J. "Fault-Tolerant Quantum Computation," in *Introduction to Quantum Computation and Information.*, eds. H-K Lo, T Spiller, S Popescu. Singapore: World Scientific (1998) 213–69.
29. Dibenedetto CN, Fanizza E, Brescia R, Kolodny Y, Remennik S, Panniello A, et al. Coupling effects in QD dimers at sub-nanometer interparticle distance. *Nano Res* (2020) **13**(4):1071–80. doi:10.1007/s12274-020-2747-3.
30. Fresch B, Cipolloni M, Yan T-M, Collini E, Levine RD, Remacle F. Parallel and multivalued logic by the two-dimensional photon-echo response of a Rhodamine–DNA complex. *J Phys Chem Lett* (2015) **6**:1714–8. doi:10.1021/acs.jpclett.5b00514.
31. Steinfeld JI. *Molecules and Radiation*. Mineola: Dover (2005)

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Physics

REVIEW
published: 10 November 2020
doi: 10.3389/fphy.2020.589504

Qudits and High-Dimensional Quantum Computing

Yuchen Wang[1,2], *Zixuan Hu*[1,2], *Barry C. Sanders*[3] *and Sabre Kais*[1,2]*

[1]Department of Chemistry, Purdue University, West Lafayette, IN, United States, [2]Department of Physics and Purdue Quantum Science and Engineering Institute, Purdue University, West Lafayette, IN, United States, [3]Institute for Quantum Science and Technology, University of Calgary, Calgary, AB, Canada

Qudit is a multi-level computational unit alternative to the conventional 2-level qubit. Compared to qubit, qudit provides a larger state space to store and process information, and thus can provide reduction of the circuit complexity, simplification of the experimental setup and enhancement of the algorithm efficiency. This review provides an overview of qudit-based quantum computing covering a variety of topics ranging from circuit building, algorithm design, to experimental methods. We first discuss the qudit gate universality and a variety of qudit gates including the pi/8 gate, the SWAP gate, and the multi-level controlled-gate. We then present the qudit version of several representative quantum algorithms including the Deutsch-Jozsa algorithm, the quantum Fourier transform, and the phase estimation algorithm. Finally we discuss various physical realizations for qudit computation such as the photonic platform, iron trap, and nuclear magnetic resonance.

Keywords: quantum information, quantum computing, qudit gates, qudit algorithm, qudit implementation

Edited by:
Marcelo Silva Sarandy,
Fluminense Federal University, Brazil

Reviewed by:
Eduardo Duzzioni,
Federal University of Santa Catarina, Brazil
Jun Jing,
Zhejiang University, China

****Correspondence:***
Sabre Kais
kais@purdue.edu

Specialty section:
This article was submitted to Quantum Computing, a section of the journal Frontiers in Physics

Received: *03 August 2020*
Accepted: *22 September 2020*
Published: *10 November 2020*

Citation:
Wang Y, Hu Z, Sanders BC and Kais S (2020) Qudits and High-Dimensional Quantum Computing. Front. Phys. 8:589504. doi: 10.3389/fphy.2020.589504

INTRODUCTION

Qudit technology, with a qudit being a quantum version of d-ary digits for $d > 2$ [23]; is emerging as an alternative to qubit for quantum computation and quantum information science. Due to its multi-level nature, qudit provides a larger state space to store and process information and the ability to do multiple control operations simultaneously [106]. These features play an important role in the reduction of the circuit complexity, the simplification of the experimental setup and the enhancement of the algorithm efficiency [100, 106, 108, 109]. The advantage of the qudit not only applies to the circuit model for quantum computers but also applies to adiabatic quantum computing devices [5, 166]; topological quantum systems [16, 37, 38] and more. The qudit-based quantum computing system can be implemented on various physical platforms such as photonic systems [60, 106]; continuous spin systems [2, 11]; ion trap [91]; nuclear magnetic resonance [48, 62] and molecular magnets [99].

Although the qudit system's advantages in various applications and potentials for future development are substantial, this system receives less attention than the conventional qubit-based quantum computing, and a comprehensive review of the qudit-based models and technologies is needed. This review article provides an overview of qudit-based quantum computing covering a variety of topics ranging from circuit building [39, 61, 71, 89, 133]; algorithm designs [2, 17, 26, 62, 79, 119, 121]; to experimental methods [2, 11, 48, 60, 62, 91, 99, 106]. In this article, high-dimensional generalizations of many widely used quantum gates are presented and the universality of the qudit gates is shown. Qudit versions of three major classes of quantum algorithms—algorithms for the oracles decision problems (e.g., the Deutsch-Jozsa algorithm [121], algorithms for the hidden non-abelian subgroup problems (e.g., the phase-estimation algorithms (PEAs) [26] and the quantum search algorithm (e.g., Grover's algorithm

[79]—are discussed and the comparison of the qudit designs vs. the qubit designs is analyzed. Finally, we introduce various physical platforms that can implement qudit computation and compare their performances with their qubit counterparts.

Our article is organized as follows. Definitions and properties of a qudit and related qudit gates are given in **Section 2**. The generalization of the universal gate set to qudit systems and several proposed sets are provided in **Section 2.1**. Then **Section 2.2** lists various examples of qudit gates and discusses the difference and possible improvement of these gates over their qubit counterparts. A discussion of the gate efficiency of synthesizing an arbitrary unitary U using geometric method is given in **Section 2.3**. The next section, **Section 3**, provides an introduction to qudit algorithms: a single-qudit algorithm that finds the parity of a permutation in **Section 3.1.1**, the Deutsch-Josza algorithm in **Section 3.1.2**, the Bernstein-Vazirani algorithm in **Section 3.1.3**, the quantum Fourier transform in **Section 3.2.1**, the PEA in **Section 3.2.2** and the quantum search algorithm in **Section 3.3**. **Section 4** is a section focused on the qudit quantum computing models other than the circuit model, which includes the measurement-based model in **Section 4.1**, the adiabatic quantum computing in **Section 4.2** and the topological quantum computing in **Section 4.3**. In **Section 5**, we provide various realizations of the qudit algorithms on physical platforms and discuss their applications. We discuss possible improvements in computational speed-up, resource saving and implementations on physical platforms. A qudit with a larger state space than a qubit can utilize the full potential of physical systems such as photon in **Section 5.1**, ion trap in **Section 5.2**, nuclear magnetic resonance in **Section 5.3** and molecular magnet in **Section 5.4**. Finally, we give a summary of the qudit systems advantages and provide our perspective for the future developments and applications of the qudit in **Section 6**.

2 QUANTUM GATES FOR QUDITS

A *qudit* is a quantum version of d-ary digits whose state can be described by a vector in the d dimensional Hilbert space $\mathscr{H}_d$ [23]. The space is spanned by a set of orthonormal basis vectors $\{|0\rangle, |1\rangle, |2\rangle, \ldots |d-1\rangle\}$. The state of a qudit has the general form

$$|\alpha\rangle = \alpha_0|0\rangle + \alpha_1|1\rangle + \alpha_2|2\rangle + \cdots + \alpha_{d-1}|d-1\rangle = \begin{pmatrix} \alpha_0 \\ \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_{d-1} \end{pmatrix} \in \mathbb{C}^d \tag{1}$$

where $|\alpha_0|^2 + |\alpha_1|^2 + |\alpha_2|^2 + \cdots + |\alpha_{d-1}|^2 = 1$. Qudit can replace qubit as the basic computational element for quantum algorithms. The state of a qudit is transformed by qudit gates.

This section gives a review of various qudit gates and their applications. **Section 2.1** provides criteria for the qudit universality and introduces several fundamental qudit gate sets. **Section 2.2** presents examples of qudit gates and illustrates their advantages compared to qubit gates. In the last section, **Section 2.3**, a quantitative discussion of the circuit efficiency is included to give a boundary of the number of elementary gates needed for decomposing an arbitrary unitary matrix.

2.1 Criteria for Universal Qudit Gates

This subsection describes the universal gates for qudit-based quantum computing and information processing. We elaborate on the criteria for universality in **Section 2.1.1** and give examples in **Section 2.1.2**.

2.1.1 Universality

In quantum simulation and computation, a set of matrices $U_k \in U(d^n)$ is called the universal quantum gate set if the product of its elements can be used to approximate any arbitrary unitary transformation U of the Hilbert space $\mathscr{H}_d^{\otimes n}$ with acceptable error measured in some appropriate norm [153]. This idea of *universality* not only applies to the qubit systems [47]; but can also be extended to the qudit logic [24, 39, 65, 102, 114, 164]. Several discussions of standards and proposals for a universal qudit gate set exist. Vlasov shows that the combination of two noncommuting single qudit gates and a two-qudit gate are enough to simulate any unitary $U \in U(d^n)$ with arbitrary precision [153]. Qudit gates can themselves be reduced to, and thus simulated by, sequences of qudit gates of lower-dimensional qudit gates [135, 137] Brylinski and Brylinski prove a set of sufficient and necessary conditions for exact qudit universality which needs some random single qudit gates complemented by one two-qudit gate that has entangled qudits [23]. Exact universality implies that any unitary gate and any quantum process can be simulated with zero error. Neither of these methods is constructive and includes a method for physical implementation. A physically workable procedure is given by Muthukrishnan and Stroud using single- and two-qudit gates to decompose an arbitrary unitary gate that operates on N qudits [118]. They use the spectral decomposition of unitary transformations and involve a gate library with a group of continuous parameter gates. Brennen et al. [21] identify criteria for exact quantum computation in qudit that relies on the QR decomposition of unitary transformations. They generate a library of gates with a fixed set of single qudit operations and "one controlled phase" gate with single parameter as the components of the universal set. Implementing the concept of a coupling graph, they proved that by connecting the nodes (equivalently logical basis states) they can show the possibility of universal computation.

2.1.2 Examples of Universal Gate Sets

An explicit and physically realizable universal set comprising one-qudit general rotation gates and two-qudit controlled extensions of rotation gates is explained in this section [108]. We first define

$$U_d(\boldsymbol{\alpha}) : \sum_{l=0}^{d-1} \alpha_l |l\rangle \mapsto |d-1\rangle, \ \boldsymbol{\alpha} := (\alpha_0, \alpha_1, \ldots, \alpha_{d-1}). \tag{2}$$

as a transformation in the d-dimension that maps any given qudit state to $|d-1\rangle$. Complex parameters of U_d may not be unique and have been addressed with probabilistic quantum search algorithm [118]. Here in this scheme, U_d can be deterministically decomposed into $d-1$ unitary transformations such that

$$U_d = X_d^{(d-1)}(a_{d-1}, b_{d-1})\cdots X_d^{(1)}(a_1, b_1),\ a_l := \alpha_l,\ b_l := \sqrt{\sum_{l=0}^{l-1} \alpha_i^2} \tag{3}$$

with

$$X_d^{(l)}(x,y) = \begin{pmatrix} 1_{l-1} & & & \\ & \frac{x}{\sqrt{|x|^2+|y|^2}} & \frac{-y}{\sqrt{|x|^2+|y|^2}} & \\ & \frac{y^*}{\sqrt{|x|^2+|y|^2}} & \frac{x^*}{\sqrt{|x|^2+|y|^2}} & \\ & & & 1_{d-l-1} \end{pmatrix}. \tag{4}$$

The d-dimensional phase gate is

$$Z_d(\theta) := \sum_{l=0}^{d-1} e^{i(1-\operatorname{sgn}(d-1-l))\theta} |l\rangle\langle l|, \tag{5}$$

which changes $|d-1\rangle$ by a phase θ and ignores the other states, and sgn represents the sign function.

Each primitive gate (such as $X_d^{(l)}$ or Z_d) has two free complex parameters to be controlled (x, y in the $X_d^{(l)}$ gate and θ in the Z_d gate). Let R_d represents either $X_d^{(l)}$ or Z_d, then the controlled-qudit gate is

$$C_2[R_d] := \begin{pmatrix} 1_{d^2-d} & \\ & R_d \end{pmatrix}, \tag{6}$$

which is a $d^2 \times d^2$ matrix that acts on two qudits. R_d acts on d substates $|d-1\rangle|0\rangle, \ldots, |d-1\rangle|d-1\rangle$, and the identity operation 1_{d^2-d} acts on the remaining substates.

Now we work on an $N = d^n$ dimensional unitary gate $U \in SU(d^n)$ operating on the n-qudit state. The sufficiency of the gates $X_d^{(l)}, Z_d$ and $C_2[R_d]$ to construct an arbitrary unitary transformation of $SU(d^n)$ is proved in three steps. The first step is the eigen-decomposition of U. By the representation theory, the unitary matrix U with N eigenvalues $\{\lambda_s\}$ and eigenstates $|E_s\rangle$ can be rewritten as

$$U = \sum_{j=1}^{N} e^{i\lambda_j} |E_j\rangle\langle E_j| = \prod_{j=1}^{N} \Upsilon_j \tag{7}$$

with eigenoperators

$$\Upsilon_j = \sum_{s=1}^{N} e^{i(1-|\operatorname{sgn}(j-s)|)\lambda_s} |E_s\rangle\langle E_s|. \tag{8}$$

Then the eigenoperators can be synthesized with two basic transformations as [118].

$$\Upsilon_j = U_{j,N}^{-1} Z_{j,N} U_{j,N}. \tag{9}$$

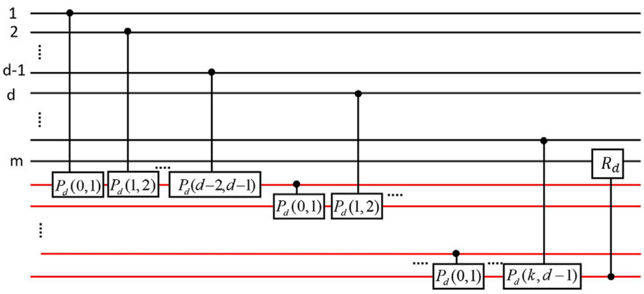


FIGURE 1 | The schematic circuit of $C_m[R_d]$ with $C_2[R_d]$ and $C_2[P_d(p,q)]$. The horizontal lines represent qudits. The auxiliary qudits initialized to $|0\rangle$ is denoted by the red lines and the black lines denoting m controlling qudits. The two-qudit controlled gates is shown as the verticle lines. $P_d(p,q)$ is the permutation of $|p\rangle$ and $|q\rangle$ state, and R_d is either $X_d^{(l)}$ or Z_d.

Here $U_{j,N}$ and $Z_{j,N}$ are the N-dimensional analogues of U_d and Z_d such that $U_{j,N}$ is applied to the jth eigenstate to produce $|N-1\rangle$ and $Z_{j,N}$ modifies the phase of $|N-1\rangle$ by the jth eigenphase λ_j, while ignoring all the other computation states. According to **Eq. 3**, $U_{j,N}$ can be decomposed with primitive gates $X_{j,N}^{(l)}(x,y)$. Thus, $X_{j,N}(x,y)$ and $Z_{j,N}$ are sufficient to decompose U.

The second step is decomposing $U_{j,N}$ and $Z_{j,N}$. In other words, $U_{j,N}$ and $Z_{j,N}$ need to be decomposed in terms of multi-qudit-controlled gates. For convenience denote $C_m[R_d]$ as

$$C_m[R_d] = \begin{pmatrix} 1_{d^m-d} & \\ & R_d \end{pmatrix}, \tag{10}$$

which acts on the d^m-dimensional computational basis of m-qudit space. It is proved in the appendix of Ref. 108 that each $U_{j,N}$ can be decomposed into some combinations of $C_m[R_d]$ and $C_m[P_d(p,q)]$ where $P_d(p,q)$ is the permutation of $|p\rangle$ and $|q\rangle$ state. The third step is using the two-qudit gates $C_2[R_d]$ and $C_2[P_d(p,q)]$ to complete the decomposition of $C_m[R_d]$. **Figure 1** shows a possible decomposition for $d > 2$. There are $r = (m-2)/(d-2)$ auxiliary qudits in the circuits (x denotes the smallest integer greater than x). The last box contains $R_d = Z_d$ or $X_d^{(l)}$. $C_m[R_d]$ is implemented with these gates combined. All of the three steps together prove that the qudit gates set

$$\Gamma_d := \left\{X_d^{(l)}, Z_d, C_2[R_d]\right\} \tag{11}$$

is universal for the quantum computation using qudit systems.

One advantage of the qudit model (compared to the qubit model) is a reduction of the number of qudits required to span the state space. To explain this, we need at least $n_1 = \log_2 N$ qubits to represent an N-dimensional system in qubits while in qudits we need $n_2 = \log_d N$ qudits. The qudit system has a reduction factor $k = n_1/n_2 = \log_2 d$. According to Muthukrishnan and Stroud's method in Ref. 118 a binary equivalent of their construction requires a number of qubit gates in the scale of $O(n_1^2 N^2)$. By analogy, the scale of the required qudit gates using the same construction is $O(n_2^2 N^2)$. So the qudit method has a $(\log_2 d)^2$ scaling advantage over the qubit case. Furthermore, in this

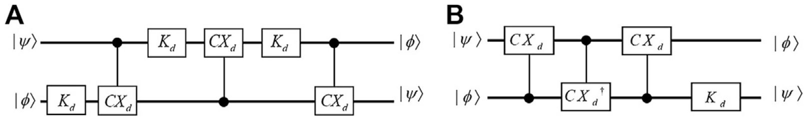


FIGURE 2 | (A) is the qudit SWAP circuit using CX_d and K_d gates [58, 131]. **(B)** is the qudit SWAP circuits with the CX_d, the $CX_d^{\dagger}$ and the K_d gates.

reviewed method, for an arbitrary unitary $U \in SU(N)$, from Eqs 7 and 8 N eigenoperators is needed and each can be decomposed with three rotations shown in **Eq. 9**. Deriving from the appendix of Ref. 108; $U_{j,N}$ can be decomposed with less than $3d^{n-1}$ multiple controlled operations. Finally, as **Figure 1** has shown, $C_m[R_d]$ needs m number of $C_2[R_d]$ and $C_2[P_d(p,q)]$. U_d can be composed with $d-1$ numbers of $X_d^{(l)}$ as in **Eq. 3**. Therefore the total number of primitive operations L in this decomposition method is

$$L \le 2N \times 3d^{n-1} \times n \times (d-1) + N \times n \le 6nd^{2n} + nd^{n}. \quad (12)$$

It is clear that there is an extra factor of n reduction in the gate requirement as the number scale of this method is $O(nN^2)$. The other advantage is these primitive qudit gates can be easily implemented with fewer free parameters [108].

For qudit quantum computing, depending on the implementation platform, other universal quantum gate sets can be considered. For example, in a recent proposal for topological quantum computing with metaplectic anyons, Cui and Wang prove a universal gates set for qutrit and *qupit* systems, for a qupit being a qudit with p dimensions and p is an prime number larger than 3 [38]. The proposed universal set is a qudit analogy of the qubit universal set and it consists several generalized qudit gates from the universal qubit set.

The generalized Hadamard gate for qudits H_d is

$$H_d|j\rangle = \frac{1}{\sqrt{d}}\sum_{i=0}^{d-1}\omega^{ij}|i\rangle, j \in \{0,1,2,\ldots,d-1\}, \quad (13)$$

where

$$\omega := \mathrm{e}^{2\pi \mathrm{i}/\mathrm{d}.} \quad (14)$$

The SUM_d gate serves as a natural generalization of the CNOT gate

$$\mathrm{SUM}_d|i,j\rangle = |i, i+j(\mathrm{mod}d)\rangle, i,j \in \{0,1,2,\ldots,d-1\}. \quad (15)$$

The Pauli σ_z, with the $\pi/8$ gate as its 4th root, can be generalized to $Q[i]$ gates for qudits,

$$Q[i]_d|j\rangle = \omega^{\delta_{ij}}|j\rangle, \quad (16)$$

with ω defined by **Eq. 14** and the related $P[i]$ gates are

$$P[i]_d|j\rangle = (-\omega^2)^{\delta_{ij}}|j\rangle,\ i,j \in \{0,1,2,\ldots,d-1\}. \quad (17)$$

In general $Q[i]_p$ is always a power of $P[i]_p$ if p is an odd prime.

The proposed gate set for the qutrit system is the sum gate SUM_3, the Hadamard gate H_3 and any gate from the set $\{P[0]_3, P[1]_3, P[2]_3\}$. As an analogue of the standard universal set for qubit $\{\mathrm{CNOT}, H, T = \pi/8 - \text{gate}\}$, the qutrit set generate the qutrit Clifford group whereas the qubit set generate the qubit Clifford group (the definition of the Clifford group can be found in **Section 2.2.1**). Whereas the rigorous proof can be found in Ref. 38; the proving process follows the idea introduced in Ref. 23 that the gate SUM_3 is imprimitive, and the Hadamard H_3 and any gate from $\{P[0]_3, P[1]_3, P[2]_3\}$ generates a dense subgroup of $SU(3)$. Similarly, the proposed gate set for the qupit system is the sum gate SUM_p, the Hadamard gate H_p and the gates $Q[i]_p$ for $i \in [p-1]$. The proof is analogous to that of the qutrit set. The Hadamard H_p and the $Q[i]$ gates are combined to form a dense subgroup of $SU(p)$ and SUM_p is shown to be imprimitive. Implementing Theorem 1.3 in Ref. 23; the set is a universal gate set. These universal gate sets for the qudit systems, with fewer numbers of gates in each set compare to that in the previous examples, have the potential to perform qudit quantum algorithms on the topological quantum computer.

2.2 Examples of Qudit Gates

In this section we introduce the qudit versions of many important quantum gates and discuss some of the gates' advantages compared to their qubit counterparts. The gates discussed are the qudit versions of the $\pi/8$ gate in **Section 2.2.1**, the SWAP gate in **Section 2.2** and the multi-level controlled gate in **Section 2.2** In **Section 2.2.3**, we also introduce how to simplify the qubit Toffoli gate by replacing one of the qubit to qudit. This gives ideas about improving the qubit circuits and gates by introducing qudits to the system.

2.2.1 Qudit Versions $\pi/8$ Gate

The qubit $\pi/8$ gate T has an important role in quantum computing and information processing. This gate has a wide range of applications because it is closely related to the Clifford group but does not belong to the group. From the Gottesman-Knill theorem [64] it is shown that the Clifford gates and Pauli measurements only do not guarantee universal quantum computation (UQC). The $\pi/8$ gate, which is non-Clifford and from the third level of the Clifford hierarchy, is the essential gate to obtaining UQC [20]. This gate can be generalized to a d dimensional qudit system, where, throughout the process, d is assumed to be a prime number greater than 2 [71].

To define the Clifford group for a d-dimensional qudit space, we first define the Pauli Z gate and Pauli X gate. The Pauli Z gate and Pauli X gate are generalized to d dimension in the matrix forms [11, 67, 124, 130].

$$X_d = \begin{pmatrix} 0 & 0 & \cdots & 0 & 1 \\ 1 & 0 & \cdots & 0 & 0 \\ 0 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & 0 \end{pmatrix}, Z_d = \begin{pmatrix} 1 & 0 & 0 & \cdots & 0 \\ 0 & \omega & 0 & \cdots & 0 \\ 0 & 0 & \omega^2 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & 0 \\ 0 & 0 & 0 & \cdots & \omega^{d-1} \end{pmatrix} \tag{18}$$

for ω the d^{th} root of unity **Eq. 14**. The function of the Z gate is adding different phase factors to each basis states and that of the X gate is shifting the basis state to the next following state. Using basis states the two gates are

$$Z_d|j\rangle := \omega^j|j\rangle\ X_d|j\rangle := |j\rangle + 1,\ j \in \{0, 1, 2, \ldots, d-1\} \tag{19}$$

In general, we define the displacement operators as products of the Pauli operators,

$$D_{(x|z)} = \tau^{xz} X_d^x Z_d^z,\ \tau := \mathrm{e}^{(d+1)\pi i/d}, \tag{20}$$

where $(x|z)$ correspond to the x and y in the exponent of τ, X and Z. This leads to the definition of the Weyl-Heisenberg group (or the generalized Pauli group) for a single qudit as [11, 67, 124, 130].

$$\mathcal{G} = \left\{\tau^c D_{\overrightarrow{\chi}} \middle| \overrightarrow{\chi} \in \mathbb{Z}_{\mathrm{d}}^2, \mathrm{c} \in \mathbb{Z}_{\mathrm{d}}\right\}\ (\mathbb{Z}_{\mathrm{d}} = \{0, 1, \ldots, \mathrm{d}-1\}), \tag{21}$$

where $\overrightarrow{\chi}$ is a two-vector with elements from $\mathbb{Z}_{\mathrm{d}}$. With these preliminary concepts defined in Eqs. 18-21, we now define the Clifford group as the following: the set of the operators that maps the Weyl-Heisenberg group onto itself under conjugation is called the *Clifford group* [124, 157];

$$\mathcal{C} = \{C \in U(d) | C\mathcal{G}C^\dagger = \mathcal{G}\}. \tag{22}$$

A recursively defined set of gates, the so-called Clifford hierarchy, was introduced by Gottesman and Chuang as

$$\mathcal{C}_{k+1} = \{U | U\mathcal{C}_1 U^\dagger \subseteq \mathcal{C}_k\}, \tag{23}$$

for $\mathcal{C}_1$ the Pauli group [66]. The sets $\mathcal{C}_{k\geq 3}$ do not form groups, although the diagonal subsets of $\mathcal{C}_3$, which is our focus here, do form a group.

The following derivations follow those in Ref. 71. The explicit formula for building a Clifford unitary gate with

$$F = \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix} \in \mathrm{SL}(2, \mathbb{Z}_{\mathrm{d}}),\ \overrightarrow{\chi} = \begin{pmatrix} x \\ z \end{pmatrix} \in \mathbb{Z}_{\mathrm{d}}^2 \tag{24}$$

is

$$C_{\left(F\middle|\overrightarrow{\chi}\right)} = D_{(x|z)} V_F, \tag{25}$$

$$V_F = \begin{cases} \dfrac{1}{\sqrt{d}} \displaystyle\sum_{j,k=0}^{d-1} \tau^{\beta^{-1}\left(\alpha k^2 - 2jk + \delta j^2\right)} |j\rangle\langle k|, & \beta \neq 0 \\ \displaystyle\sum_{k=0}^{d-1} \tau^{\alpha\gamma k^2} |\alpha k\rangle\langle k|, & \beta = 0. \end{cases} \tag{26}$$

The special case $\beta = 0$ is particularly relevant to the later derivation, and

$$\det\left(\sum_{k=0}^{d-1} \tau^{\alpha\gamma k^2} |k\rangle\langle k|\right) = \tau^{\frac{\alpha\gamma}{6}(2d-1)(d-1)d}, \\ = \begin{cases} \tau^{2\alpha\gamma}, & d = 3, \\ 1, & \forall\, d > 3, \end{cases} \tag{27}$$

can be shown. In the $d = 3$ case, we use

$$C_{\left(\begin{bmatrix} 1 & 0 \\ \gamma & 1 \end{bmatrix} \middle| \begin{bmatrix} x \\ z \end{bmatrix}\right)} \in SU(p)\, \forall\, p > 3 \tag{28}$$

and

$$\det\left(C_{\left(\begin{bmatrix} 1 & 0 \\ \gamma & 1 \end{bmatrix} \middle| \begin{bmatrix} x \\ z \end{bmatrix}\right)} \right) = \tau^{2\gamma} \text{ for } p = 3. \tag{29}$$

With all the mathematical definitions at hand, we are ready to give an explicit form of the qudit $\pi/8$ gate. We choose the qudit gate U_υ to be diagonal in the computational basis and claim that, for $d > 3$, U_υ has the form

$$U_\upsilon = U(\upsilon_0, \upsilon_1, \ldots) = \sum_{k=0}^{d-1} \omega^{\upsilon_k} |k\rangle\langle k|\, (\upsilon_k \in \mathbb{Z}_{\mathrm{d}}). \tag{30}$$

A straightforward application of **Eqs 20** and **30** yields

$$U_\upsilon D_{(x|z)} U_\upsilon^\dagger = D_{(x|z)} \sum_k \omega^{\upsilon_{k+1} - \upsilon_k} |k\rangle\langle k|. \tag{31}$$

As U_υ is to be a member of $\mathcal{C}_3$, the right hand side of **Eq. 31** must be a Clifford gate. We ignore the trivial case $U_\upsilon D_{(0|z)} U_\upsilon^\dagger = D_{(0|z)}$ and focus on the case $U_\upsilon D_{(1|0)} U_\upsilon^\dagger$ in order to derive an explicit expression for U_υ.

We define $\gamma', z', \varepsilon' \in \mathbb{Z}_{\mathrm{d}}$ such that

$$U_\upsilon D_{(1|0)} U_\upsilon^\dagger = \omega^{\varepsilon'} C_{\left(\begin{bmatrix} 1 & 0 \\ \gamma' & 1 \end{bmatrix} \middle| \begin{bmatrix} 1 \\ z' \end{bmatrix}\right)} \tag{32}$$

From **Eqs 26** and **31** we see that the right-hand side of **Eq. 32** is the most general form, and we note that $U \in SU(d)$ implies $\omega^{\gamma'} U \in SU(d)$. We rewrite the left-hand side of **Eq. 32** using **Eq. 31** and right-hand side using **Eq. 26** and obtain

$$D_{(1|0)} \sum_k \omega^{\upsilon_{k+1} - \upsilon_k} |k\rangle\langle k| = \omega^{\varepsilon'} D_{(1|z')} \sum_{k=0}^{d-1} \tau^{\gamma' k^2} |k\rangle\langle k|. \tag{33}$$

After canceling common factors of $D_{(1|0)}$, an identity between two diagonal matrices remains such that

$$\omega^{\upsilon_{k+1} - \upsilon_k} = \omega^{\varepsilon'} \tau^{z'} \omega^{kz'} \tau^{\gamma' k^2}\ (\forall k \in \mathbb{Z}_{\mathrm{d}}), \tag{34}$$

or, equivalently, using **Eq. 20**,

$$\upsilon_{k+1} - \upsilon_k = \varepsilon' + 2^{-1} z' + kz' + 2^{-1}\gamma' k^2. \tag{35}$$

From here, we derive the recursive relation

$$\upsilon_{k+1} = \upsilon_k + k\left(2^{-1}\gamma' k + z'\right) + 2^{-1} z' + \varepsilon'. \tag{36}$$

We solve for the υ_k with a boundary condition $\upsilon_0 = 0$,

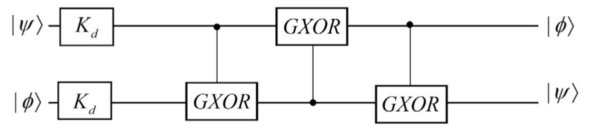


FIGURE 3 | Qudit SWAP circuits with the *GXOR* and the K_d gates [4, 155].

$$v_k = \frac{1}{12}k\{\gamma' + k[6z' + (2k-3)\gamma']\} + k\varepsilon', \tag{37}$$

where all factors are evaluated modulo d. For example, with $d = 5$, the fifth root of unity **Eq. 14** is $\omega = e^{2\pi i/5}$ and choosing $z' = 1, \gamma' = 4$ and $\varepsilon' = 0$, we obtain

$$v = (v_0, v_1, v_2, v_3, v_4) = (0, 3, 4, 2, 1) \tag{38}$$

so that

$$U_v = \begin{pmatrix} \omega^0 & 0 & 0 & 0 & 0 \\ 0 & \omega^{-2} & 0 & 0 & 0 \\ 0 & 0 & \omega^{-1} & 0 & 0 \\ 0 & 0 & 0 & \omega^2 & 0 \\ 0 & 0 & 0 & 0 & \omega^1 \end{pmatrix} \tag{39}$$

The diagonal elements of U_v are powers of ω that sum to zero modulo d and, consequently, $\det(U_v) = 1$.

For the $d = 3$ case, because of **Eq. 27** extra work is needed for solving a matrix equation similar to **Eq. 32**. We first introduce a global phase factor $e^{i\phi}$ such that

$$\det\left(e^{i\phi}\sum_{k=0}^{d-1}\tau^{\gamma k^2}|k\rangle\langle k|\right) = 1 \Rightarrow \phi = 4\pi\gamma/9. \tag{40}$$

The ninth root of unity **Eq. 14** is $\omega = e^{2\pi i/9}$ and, from **Eq. 29** we derive that

$$\det\left(\omega^{2\gamma'}C_{\left(\begin{bmatrix} 1 & 0 \\ \gamma' & 1 \end{bmatrix}\middle|\begin{bmatrix} 1 \\ z' \end{bmatrix}\right)}\right) = 1. \tag{41}$$

The qutrit version of $U_{\pi/8}$ has a more general form than in **Eq. 30**; i.e.

$$U_v = U(v_0, v_1, \ldots) = \sum_{k=0}^{2}\omega^{v_k}|k\rangle\langle k|,\ v_k \in \mathbb{Z}_9. \tag{42}$$

Then the general solution is

$$v = (0, 6z' + 2\gamma' + 3\varepsilon', 6z' + \gamma' + 6\varepsilon')\text{mod}9. \tag{43}$$

For example, choosing $z' = 1, \gamma' = 2$ and $\varepsilon' = 0$,

$$U_v = \begin{pmatrix} \omega^0 & 0 & 0 \\ 0 & \omega^1 & 0 \\ 0 & 0 & \omega^{-1} \end{pmatrix}. \tag{44}$$

The $\pi/8$ gate, with its close relation to the Clifford group, has many applications and utilities in teleportation-based UQC [66]; transversal implementation [50, 162]; learning an unknown gate [105]; or securing assisted quantum computation [28]. The generalized qudit version of the $\pi/8$ gate, U_v, is shown to be identical to the maximally robust qudit gates for qudit fault-tolerant UQC discussed in reference [150].

This gate also plays an important role in the magic-state distillation (MSD) protocols for general qudit systems, which was first established for qutrits [6] and then extended to all prime-dimensional qudits [25].

2.2.2 Qudit SWAP Gate

A SWAP gate is used to exchange the states of two qudit such that:

$$SWAP|\phi\rangle|\psi\rangle = |\psi\rangle|\phi\rangle \tag{45}$$

Various methods to achieve the SWAP gate use different variants of qudit controlled gates [4, 58, 112, 131, 155, 158, 159] as shown in **Figure 2A,B**. The most used component of the SWAP gate is a controlled-shift gate CX_d that perform the following operation:

$$CX_d|x\rangle|y\rangle = |x\rangle|x+y\rangle \tag{46}$$

with a modulo d addition. Its inverse operation is

$$CX_d^\dagger|x\rangle|y\rangle = |x\rangle|y-x\rangle \tag{47}$$

In some approaches, the operation K_d is required to complete the circuits, where

$$K_d|x\rangle = |d-x\rangle = |-x\rangle, \tag{48}$$

which outputs the modulo d complement of the input. These circuits are more complex and less intuitive then the qubit SWAP gate [58] because they are not Hermitian, i.e., $CX_d \neq CX_d^\dagger$.

One way to create a Hermitian version of the qudit CNOT uses the GXOR gate

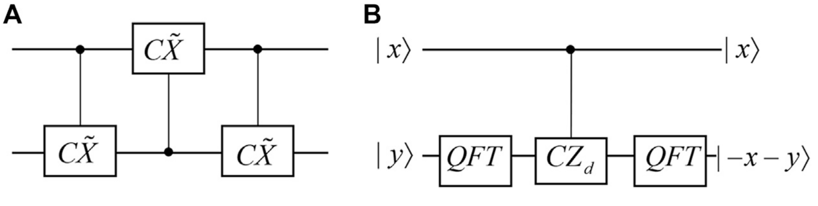


FIGURE 4 | (A) is the qudit SWAP gate with the $C\tilde{X}$ gate. **(B)** is the decomposing $C\tilde{X}$ gate. The *QFT* represents the quantum Fourier transform while CZ_d is the selective phase shift gate.

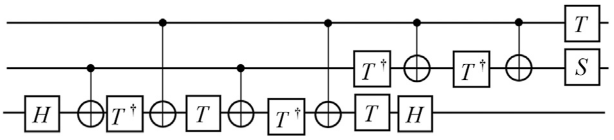


FIGURE 5 | Decomposing qubit Toffoli gate with the universal qubit gates. *H* is the Hadamard gate, *T* is the *π*/8 gate and *S* is the phase gate.

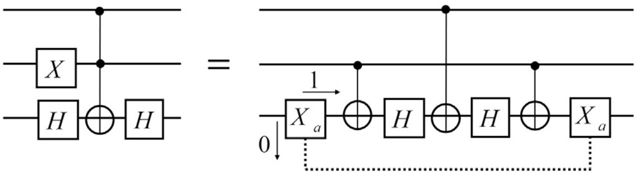


FIGURE 6 | The Simplified Toffoli gate. The first two lines represent two control qubits and the third line represents a target qutrit that has three accessible levels. The initial and final quantum states of the quantum information carrier are encoded in the $|0\rangle$ and $|1\rangle$. The *H* is the generalized Hadamard gate such that $H|0\rangle = 1/\sqrt{2}[|0\rangle+|1\rangle]$, $H|1\rangle = 1/\sqrt{2}[|0\rangle - |1\rangle]$ and $H|2\rangle = |2\rangle$. X_a gate is a qutrit gate such that $X_a|0\rangle = |2\rangle$ and $X_a|2\rangle = |0\rangle$ with $X_a|1\rangle = |1\rangle$. With the control being qubit, the target being qudit, the two qudit gate in this case is a hybrid gate.

$$GXOR|x\rangle|y\rangle = |x\rangle|x-y\rangle. \tag{49}$$

However, this SWAP gate needs to be corrected with an K_d [4] as shown in **Figure 3**. A partial SWAP gate S_p [39] works on a hybrid system where $|i\rangle$ is a qudit of dimension d_c and $|j\rangle$ is a qudit of dimension d_t

$$S_p|i\rangle\otimes|j\rangle = \begin{cases} |j\rangle\otimes|i\rangle & \text{for } i,j\in\mathbb{Z}_{\mathrm{d_p}} \\ |i\rangle\otimes|j\rangle & \text{otherwise} \end{cases} \tag{50}$$

where $d_p \le d_{\min} = \min(d_c, d_t)$.

In the rest of this section, we present a Hermitian generalization of the qudit CNOT gate with a symmetry configuration and a qudit SWAP circuit with a single type of qudit gate as shown in **Figure 4A** [61]. Compared with all the previously proposed SWAP gate for qudit, this method is easier to implement since there is only one type of gate $C\tilde{X}$ needed. To begin with, we define a gate $C\tilde{X}$ acting on *d*-level qudits $|x\rangle$ and $|y\rangle$ such that

$$C\tilde{X}|x\rangle|y\rangle = |x\rangle|-x-y\rangle, \tag{51}$$

where $|-x-y\rangle$ represents a state $|i=-x-y\rangle$ in the range $i\in\{0,\ldots,d-1\} \bmod d$. Notice that, for $d=2$, the $C\tilde{X}$ gate is equivalent to the CNOT gate. The SWAP gate for qudit can be built using three $C\tilde{X}$ gates.

$C\tilde{X}$ is generated with three steps: a qudit generalization of the *CZ* gate as CZ_d sandwiched by two quantum Fourier transform operations (QFT). The circuit illustration for the sequence of theses gate is shown in **Figure 4B**. The QFT transforms the $|E_s\rangle x$ into a uniform superposition

$$QFT|x\rangle = \frac{1}{\sqrt{d}}\sum_{k=0}^{d-1} \mathrm{e}^{\mathrm{i}2\pi xk/d}|k\rangle. \tag{52}$$

The CZ_d gate adds a phase to the target qudit depending on the state of the control qudit. Its effect on the input qudits is

$$CZ_d|x\rangle|y\rangle = \mathrm{e}^{\mathrm{i}2\pi xy/d}|x\rangle|y\rangle. \tag{53}$$

The inverse QFT undoes the Fourier transform process and the inverse of CZ_d is

$$CZ_d^\dagger|x\rangle|y\rangle = \mathrm{e}^{-\mathrm{i}2\pi xy/d}|x\rangle|y\rangle. \tag{54}$$

The full evolution of the $C\tilde{X}$ is

$$|x\rangle|y\rangle \overset{QFT_2}{\longrightarrow} \frac{1}{\sqrt{d}}\sum_{k=0}^{d-1} \mathrm{e}^{\mathrm{i}\frac{2\pi ky}{d}}|x\rangle|k\rangle \tag{55}$$

$$\overset{CZ_d}{\longrightarrow} \frac{1}{\sqrt{d}}\sum_{k=0}^{d-1} \mathrm{e}^{\mathrm{i}\frac{2\pi ky}{d}}\mathrm{e}^{\mathrm{i}\frac{2\pi xk}{d}}|x\rangle|k\rangle = \frac{1}{\sqrt{d}}\sum_{k=0}^{d-1} \mathrm{e}^{\mathrm{i}\frac{2\pi k(x+y)}{d}}|x\rangle|k\rangle \tag{56}$$

$$\overset{QFT_2}{\longrightarrow} \frac{1}{d}\sum_{l=0}^{d-1}\sum_{k=0}^{d-1} \mathrm{e}^{\mathrm{i}\frac{2\pi k(x+y)}{d}}\mathrm{e}^{\mathrm{i}\frac{2\pi kl}{d}}|x\rangle|l\rangle = |x\rangle|-x-y\rangle. \tag{57}$$

It is easy to show that $C\tilde{X}$ is its own inverse and then $C\tilde{X} = C\tilde{X}^\dagger$. For the proposed SWAP gate, both the QFT and CZ_d operations are realizable on a multilevel quantum systems. For example, there are implementations of them for multilevel atoms [118, 145]. The resulting SWAP gate provides a way to connect systems limited to the nearest-neighbour interactions. This gate provides a useful tool in the design and analysis of complex qudit circuits.

2.2.3 Simplified Qubit Toffoli Gate With a Qudit

The Toffoli gate is well known for its application to universal reversible classical computation. In the field of quantum computing, the Toffoli gate plays a central role in quantum error correction [35]; fault tolerance [43] and offers a simple universal quantum gate set combined with one qubit Hadamard gates [141]. The simplest known qubit Toffoli gate, shown in **Figure 5**, requires at least five two-qubit gates [125]. However, if the target qubit has a third level, i.e., a qutrit, the whole circuit can be achieved with three two-qubit gates [133].

A new qutrit gate X_a is introduced to the circuit that does the following: $X_a|0\rangle = |2\rangle$ and $X_a|2\rangle = |0\rangle$ with $X_a|1\rangle = |1\rangle$. The simplified circuit is shown in **Figure 6**. The two controlled gates are the CNOT gate and a control-*Z* gate, which is achieved with a CNOT gate between two Hadamard gates. The Hadamard gate here operating on the qutrit is generalized from the normal Hadamard gate operating on a qubit—it only works with the $|0\rangle$ and $|1\rangle$, such that $H|0\rangle = 1/\sqrt{2}[|0\rangle+|1\rangle]$, $H|1\rangle = 1/\sqrt{2}[|0\rangle - |1\rangle]$ and $H|2\rangle = |2\rangle$. Comparing the circuit in **Figure 6** to that in **Figure 5**, it is clear that the total number of gates is significantly reduced.

This method can be generalized to *n*-qubit-controlled Toffoli gates by utilizing a single $(n+1)$-level target carrier and using only $2n-1$ two-qubit gates [133]. In other words, the target carrier needs an extra level for each extra control qubit. Compare to the best known realization previously that requires $12n-11$

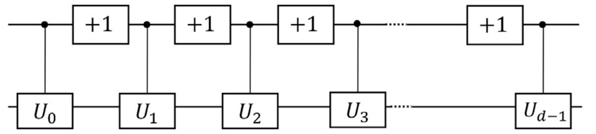


FIGURE 7 | *d*-valued Quantum Multiplexer for the second qutrit and its realization in terms of Muthukrishan-Stroud gates (the control *U* operation that only act on one specific control state). The gate labeled +1 is the shifting gate that increases the state value of the control qudit by 1 (mod *d*). Depending on the value of the top control qudit, one of U_i is applied to the second qudit, for $i \in \{0, 1, \ldots d-1\}$.

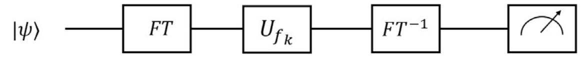


FIGURE 8 | Schematic view of the quantum circuit for the parity determining algorithm. *FT* is the Fourier transform and U_{f_k} is the gate that does one of the two permutations and the last box represents the measurement.

two-qubit gates [125]; this method offers a significant resource reduction. Furthermore, these schemes can be extended to more general quantum circuits such as the multi-qudit-controlled-unitary gate $C^n U$.

The previous method turns the target qubit into a qudit; another method simplifies the Toffoli gate by using only qudits and treating the first two levels of the qudit as qubit levels and other levels as auxiliary levels. The reduction in the complexity of Toffoli gate is accomplished by utilizing the topological relations between the dimensionality of the qudits, where higher qudit levels serve as the ancillas [89].

Suppose we have a system of n qudits denoted as Q_i, $i \in \{1, \ldots, n\}$ and each qudit has dimension $d_i \geq 2$. Qudits are initialized into pure or mix states on the first two levels, i.e., the qubit states, and zero population for the other levels, i.e., the auxiliary states. This scheme assumes the ability to perform single-qubit operations. We can apply the desirable unitary operation on the qubit states and leave the auxiliary states unchanged. We also assume that we have the ability to manipulate the auxiliary levels by a generalized inverting gate X_m

$$X_m|0\rangle = |m\rangle,\ X_m|m\rangle = |0\rangle,\ X_m|y\rangle = |y\rangle,\ \text{for}\, y \neq m, 0. \tag{58}$$

At the same time, the two-qubit *CZ* gates are applied according to certain topological connections between qudits. We introduce a set E of ordered pairs (i, j), such that $i, j \in \{1, \ldots, n\}$, $i < j$ to obtain this topology and the *CZ* gate is defined as

$$CZ|11\rangle_{Q_i,Q_j} = -|11\rangle_{Q_i,Q_j}\ CZ|xy\rangle_{Q_i,Q_j} = |xy\rangle_{Q_i,Q_j}\ \text{for}\, xy \neq 1, \tag{59}$$

with $x \in \{0, \ldots d_i - 1\}$ and $y \in \{0, \ldots d_j - 1\}$.

The set E describes an n-vertex-connected graph. Let $\tilde{E} \subseteq E$ defines an n-vertex connected *tree* (acyclic graph). The main result is: the n-qubit Toffoli gate can be achieved with less number of operations if

$$d_i \geq k_i + 1, \tag{60}$$

where d_i is the dimension of a qudit and the number k_i is the qudit's connections to other qudits within $\tilde{E}$. With this condition fulfilled, the n-qubit Toffoli gate can be realized by $2n-3$ two-qudit *CZ* gates. The detailed realization of the n-qubit Toffoli gate by the properties and special operations of the tree in topology can be found in Ref. 89. The advantage of this scheme is the scalability and the ability to implement it for the multi-qubit controlled unitary gate $C^n U$.

These $C^n U$ gates are a crucial component in the PEA which has many important applications such as the quantum simulation [8] and Shor's factoring algorithm [142]. This idea of combining qudits of different dimensions or hybrid qudit gates can also be applied to other qudit gates such as the SWAP and SUM gates as shown in Refs. [33, 39]. Thus, introducing qudits into qubit systems to create a hybrid qudit system offers the potential of improvement to quantum computation.

2.2.4 Qudit Multi-Level Controlled Gate

For a qubit controlled gate, the control qubit has only two states so it is a "do-or-don't" gate. Qudits, on the other hand, have multiple accessible states and thus a qudit-controlled gate can perform a more complicated operation [46]. The *Muthukrishan-Stroud gate* (MS gate) for a qudit applies the specified operation on the target qudit only if the control qudit is in a selected one of the *d* states, and leaves the target unchanged if the control qudit is in any other $d-1$ states. Hence, the MS gate is essentially a "do-or-don't" gate generalized to qudits and does not fully utilize the *d* states on the control qudit [118].

To fully utilize the *d* states on the control qudit, people have developed the quantum multiplexer to perform the controlled *U* operations in a qudit system as shown in **Figure 7**, where the MS gate and shifting gates are combined to apply different operations to the target depending on different states on the control states [87]. Here we discuss the *multi-value-controlled gate* (*MVCG*) for qudits, which applies a unique operation to the target qudit for each unique state of the control qudit [106].

For a *d*-dimensional qudit system, a two-qudit multi-value-controlled gate is represented by a $d^2 \times d^2$ matrix

$$MVCG = \begin{pmatrix} U_0 & 0 & 0 & \cdots & 0 \\ 0 & U_1 & 0 & \cdots & 0 \\ 0 & 0 & U_2 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & 0 \\ 0 & 0 & 0 & \cdots & U_{d-1} \end{pmatrix}, \tag{61}$$

where each U_i $(i = 0, 1, \ldots, d-1)$ is a unique unitary single-qudit operation. The U_i operation is applied to the target qudit when the control qudit is in $|E_s\rangle$ state. In the later sections, **Section 3.2.1** and **Section 3.2.2** the controlled gates are *MVCG* and improve the efficiency of the qudit algorithm. *MVCG* can be built in many physical systems and one example in a photonic system is introduced in **Section 5.1**.

2.3 Geometrically Quantifying Qudit-Gate Efficiency

In a quantum computer, each qudit can remain coherent for a limited amount of time (decoherence time). After this time, the

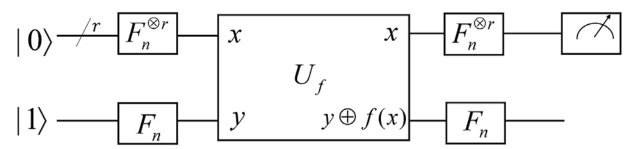


FIGURE 9 | The Deutsch-Jozsa circuit in qudit system. The F_n are the qudit Hadamard gates achieved with quantum Fourier transform.

quantum information is lost due to the outside perturbations and noises. In the computation process, quantum gates take certain amount of time to alter the states of the qudits. The decoherence time of a qudit state limits the number of quantum gates in the circuit. Therefore, we need to design more efficient algorithms and circuits. A method exists to do a general systematic evaluation of the circuit efficiency with the mathematical techniques of Riemannian geometry [126]. By reforming the quantum circuits designing problems as a geometric problem, we are able to develop new quantum algorithms or to exploring and evaluating the full potential of the quantum computers. This evaluation is able to generalized to qutrit systems, where the least amount of the gates required to synthesize any unitary operation is given [100].

To begin with, we assume that the operations done by the quantum circuit can be described by a unitary evolution U derived from the time-dependent Schrödinger equation $\mathrm{d}U/\mathrm{d}t = -\mathrm{i}HU$ with the boundary condition t_f, $U(t_\mathrm{f}) = U$. The complexity of realizing U can be characterized by a cost function $F[H(t)]$ on the Hamiltonian control $H(t)$. This allow us to define a Riemannian geometry on the space of unitary operations [122]. Finding the minimal geodesics of this Riemannian geometry is equivalent to finding the optimal control function $H(t)$ of synthesizing the desired U.

Now we transform the problem of calculating a lower bound to the gate number to finding the minimal geodesic distance between the identity operation I and U. Instead of Pauli matrices for the qubit representation of the Hamiltonian, the qutrit version of Hamiltonian is expanded in terms of the Gell-Mann matrices. Here we give an explicit form of the Gell-Mann matrices representation in d-dimension [109] which is used for qutrit (where $d = 3$) as well as other qudit systems in the later part of the section. Let e_{jk} denote the $d \times d$ matrix with a one in the (j, k) elements and 0 s elsewhere, a basis can be described as

$$u_{jk}^d = e_{jk} + e_{kj},\ 1 \le j < k \le d, \tag{62}$$

$$u_{jk}^d = \mathrm{i}\left(e_{jk} - e_{kj}\right),\ 1 \le k < j \le d, \tag{63}$$

$$u_{jj}^d = \mathrm{diag}\left(1, \ldots, 1, -j, 0_{d-2j}\right), j \in [d-1]. \tag{64}$$

Here, diag represents the diagonal matrix, 0_{d-2j} denotes the zeros of length $d - 2j$. u_{jk}^d are traceless and Hermitian and together with the identity matrix 1_d serve as the basis of the vector space of $d \times d$ Hermitian matrix. These generalized Gell-Mann matrices can be used to generate the group representation of $SU(d)$ while the other representations can be achieved by transform these matrices uniformly. To derive the bases of $SU(d^n)$, we first define $x_l = u_{jk}^d$ with $l = jd + k, l \in [d^2]$ and

$$X_l^s = \mathrm{I}^{\otimes s-1} \otimes x_l \otimes I^{\otimes d-s} \tag{65}$$

acts on the s-th qudit with x_l and leaves the other qudits unchanged. The bases of $SU(d^n)$ is constructed by $\{Y_t^{P_t}\}, t \in [n], P_t = \{i_1, \ldots, i_t\}$ with all possible $1 < i_1 < \cdots < i_k < n$, where

$$Y_t^{P_t} = \prod_{k=0}^{i} X_{j_k}^{i_k}. \tag{66}$$

Y_t^P denotes all operators with generalized Gell-Mann matrices $x_{j_1}, \ldots, x_{j_k}$ acting on t qudits at sites $P = \{i_1, \ldots, i_k\}$, respectively, and rest with identity. It is easy to prove that with the generalized Gell-Mann matrices representations,1-body and 2-body interactions can generate all 3-body interactions.

Now the Hamiltonian in terms of the Gell-Mann matrices (with the notation σ) can be written as

$$H = \sum_\sigma^{\prime} h_\sigma \sigma + \sum_\sigma^{\prime\prime} h_\sigma \sigma. \tag{67}$$

All coefficients h_σ are real and, in $\sum{}'_\sigma h_\sigma \sigma$, σ goes over all possible one- and two-body interactions whereas, in $\sum{}''_\sigma h_\sigma \sigma$, σ goes over everything else. The cost function is

$$F(H) := \sqrt{\sum_\sigma^{\prime} h_\sigma^2 \sigma + p^2 \sum_\sigma^{\prime\prime} h_\sigma^2 \sigma}, \tag{68}$$

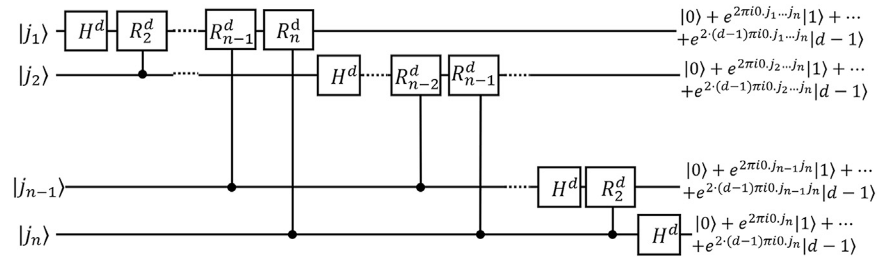


FIGURE 10 | Quantum Fourier transform in qudit system. H^d is the d-dimensional Hadamard gate and the expression of the R^d gate is shown in **Eq. 99**. Resultant states are shown to the right.

where p is a penalty cost by applying many-body terms. Now that the control cost is well defined, it is natural to form the distance in the space $SU(3^n)$ of n-qutrit unitary operators with unit determinant. We can treat the function $F(H)$ as the norm related to a Riemannian metric with a metric tensor g as:

$$g = \begin{cases} 0, & \sigma \neq \tau \\ 1, & \sigma = \tau \text{ and } \sigma \text{ is one or two body} \\ p^2, & \sigma = \tau \text{ and } \sigma \text{ is three or more body} \end{cases} . \tag{69}$$

The distance $d(I, U)$ between I and U which is the minimum curve connecting I and U equals to the minimal length solution to the geodesic equation

$$\left\langle \frac{\mathrm{d}H}{\mathrm{d}t}, K \right\rangle = \mathrm{i}\langle H, [H, K]\rangle, \tag{70}$$

where $\langle , \rangle$ denotes the inner product on the tangent space $SU(3^n)$ defined by the metric components (69), and $[,]$ denotes the matrix commutator and K is an arbitrary operator in $SU(3^n)$.

All lemmas backing up the final theorem have been proven in detail [100]; but the reasoning behind can be summarized in four parts. First let p be the three- and more-body items penalty. With large enough p, the distance $d(I, U)$ is guarantee to have a supremum that does not depend on p. Secondly, we have

$$||U - U_P|| \leq 3^n d([U])/p, \tag{71}$$

where $\| \bullet \|$ is the operator norm and U_P the corresponding unitary operator generated by the one- and two-body items projected Hamiltonian $H_P(t)$. Thirdly, given an n-qutrit unitary operator U generated by $H(t)$ with the condition $||H(t)|| \leq c$ in a time interval $[0, \Delta]$, then

$$\left|\left|U - \exp(-\mathrm{i}\overline{H})\right|\right|\Delta \leq 2\left(\mathrm{e}^{c\Delta} - 1 - c\Delta\right) = O\left(c^2\Delta^2\right), \tag{72}$$

where $\overline{H}$ is the mean Hamiltonian. Lastly, for H as an n-qutrit one- and two-body Hamiltonian, a unitary operator U_A exists that satisfies

$$\left|\left|\mathrm{e}^{\mathrm{i}H\Delta} - U_A\right|\right| \leq c_2 n^2 \Delta^3 \tag{73}$$

and can be generated with at most $c_1 n^2/\Delta$ one- and two-qutrit gates, and constants c_1 and c_2.

All these lemmas combined gives the final theorem for the qutrit system: for a unitary operator U in $SU(3^n)$, $O(n^k d(I, U)^3)$ one- and two-qutrit gates is the lower bound to synthesize a unitary U_A with the condition $||U - U_A|| \leq c$, given a constant c. It is worth mentioning that for any groups of unitaries U, which is labeled by the number of qudits n, the final theorem shows a quantum circuit exists with a polynomial of $d(I, U)$ number of gates such that it can approximates U to arbitrary accuracy. Alternatively,a polynomial-sized quantum circuit exists if and only if the distance $d(I, U)$ itself is scaling polynomially with n.

With appropriate modification, the Riemannian geometry method can be used to ascertain the circuit-complexity bound for a qudit system [109]. In this scheme, the unitary matrix $U \in SU(d^n)$ is represented by the generalized Gell-Mann matrices as defined in the earlier part of the section. The main theorem in the qudit case of the Ref. [109] is "for any small constant ε, each unitary $U_A \in SU(d^n)$ can be synthesized using $O(\varepsilon^{-2})$ one- and two-qudit gates, with error $U - U_A \leq \varepsilon$." To break up the constant ε to an explicit form, we have $\varepsilon^{-2} = N^2 d^4 n^2$, where d is the dimension of the qudit, n is the number of qudits and N is the number of the intervals that $d(I, U)$ divides into, such that a small $\delta = d(I, U)/N \leq \varepsilon$. The qudit case shows the explicit relation between the non-local quantum gate cost and the approximation error for synthesizing quantum qudit operations. In summary, for the quantum circuit model, one can decide a lower bound for the number of gates needed to synthesize U by finding the shortest geodesic curve linking I and U. This provides a good reference for the design of the quantum circuit using qudits.

3 QUANTUM ALGORITHMS USING QUDITS

A qudit, with its multi-dimensional nature, is able to store and process a larger amount of information than a qubit. Some of the algorithms described in this section can be treated as direct generalizations of their qubit counterparts and some utilize the multi-dimensional nature of the qudit at the key subroutine of the process. This section introduces examples of the well-known quantum algorithms based on qudits and divides them into two groups: algorithms for the oracle-decision problems in **Section 3.1** and algorithms for the hidden Abelian subgroup problems in **Section 3.2**. Finally, **Section 3.3** discusses how the qudit gates can improve the efficiency of the quantum search algorithm and reduce the difficulty in its physical set-up.

3.1 Qudit Oracle-Decision Algorithm

In this subsection we explore the qudit generalizations of the efficient algorithms for solving the oracle decision problems, which are quite important historically and used to demonstrate the classical-quantum complexity separation [44, 45]. The oracle decision problems is to locate the contents we want from one of the two mutually disjoint sets that is given. We start in **Section 3.1.1** with a discussion about a single-qudit algorithm that determines the parity of a permutation. In **Section 3.1.2**, the Deutsch-Jozsa algorithm in qudit system is discussed and its unique extension, the Bernstein-Vazirani algorithm is provided in **Section 3.1.3**.

3.1.1 Parity Determining Algorithm

In this section we review a single qutrit algorithm which provides a two to one speedup than the classical counterpart. This algorithm can also be generalized to work on an arbitrary d-dimensional qudit which solves the same problem of a larger computational space [62]. In quantum computing, superposition, entanglement and discord are three important parts for the power of quantum algorithms and yet the full picture behind this power is not completely clear [151].

Recent research shows that we can have a speedup in a fault tolerant quantum computation mode using the quantum contextuality [72]. The contextual nature can be explained as "a particular outcome of a measurement cannot reveal the pre-existing definite value of some underlying hidden variable" [92,

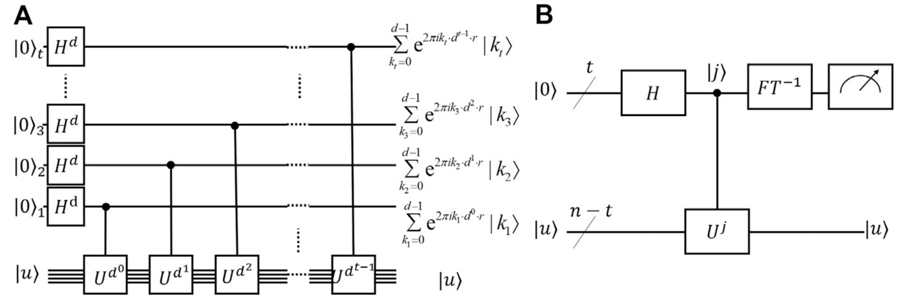


FIGURE 11 | A) The circuit for the first stage of the PEA. The qudits in the second register whose states represent $|u\rangle$ are undergoing the U operations and the generated phase factors are kicking back to the qudits in the first register, giving the results to the right. **(B)** The schematic circuit for the whole stage of PEA. After the first stage of the PEA, inverse Fourier transform (FT^{-1}) is applied to the qudits in the first register and the phase factors can be obtained by measuring the states of the first register qudits.

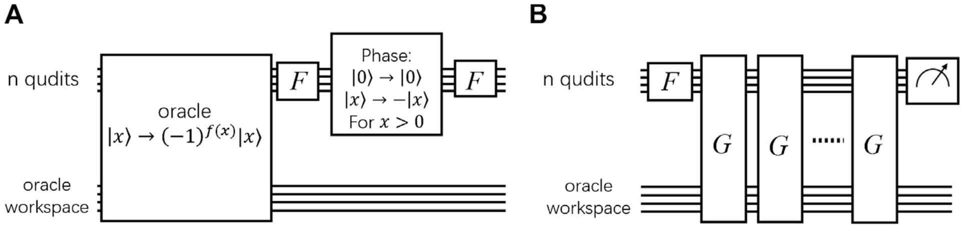


FIGURE 12 | (A) Circuit illustration for Grover iteration, ***G***, in a qudit system. The F gate is the proposed qudit gate that transforms the single-qudit state $|0_k\rangle$ into an equal weight superposition state. **(B)** Schematic circuit illustration of the qudit quantum search algorithm.

93]. In other words, the results of measurements can depend on how we made the measurement, or what combination of measurements we chose to do. For the qudit algorithm discussed below, a contextual system without any quantum entanglement is shown to solve a problem faster than the classical methods [62]. Because this qudit algorithm uses a single qudit throughout the process without utilizing any correlation of quantum or classical nature, it acts as a perfect example to study the sources of the quantum speed-up other than the quantum correlation.

The algorithm solves a black-box problems that maps d inputs to d outputs after a permutation. Consider the case of three objects where six possible permutations can be divided into two groups: *even* permutation that is a cyclic change of the elements and *odd* permutation that is an interchange between two elements. If we define a function $f(x)$ that represents the permutation on the set $x \in \{-1, 0, 1\}$, the problems become determining the parity of the bijection $f: -1, 0, 1 \to -1, 0, 1$. We use Cauchy's two-line notation to define three possible even functions f_k, namely,

$$f_1 := \begin{pmatrix} 1 & 0 & -1 \\ 1 & 0 & -1 \end{pmatrix}, f_2 := \begin{pmatrix} 1 & 0 & -1 \\ 0 & -1 & 1 \end{pmatrix},$$
$$f_3 := \begin{pmatrix} 1 & 0 & -1 \\ -1 & 1 & 0 \end{pmatrix}, \tag{74}$$

and the remaining three odd function are

$$f_4 := \begin{pmatrix} 1 & 0 & -1 \\ -1 & 0 & 1 \end{pmatrix}, f_5 := \begin{pmatrix} 1 & 0 & -1 \\ 0 & 1 & -1 \end{pmatrix},$$
$$f_6 := \begin{pmatrix} 1 & 0 & -1 \\ 1 & -1 & 0 \end{pmatrix}. \tag{75}$$

The circuit for the single qutrit algorithm in a space spanned by $\{|1\rangle, |0\rangle\, |-1\rangle\}$ is shown in **Figure 8**, where the operation U_{f_k} applies f_k to the state: $U_{f_k}(|1\rangle + |0\rangle + |-1\rangle) = |f_k(1)\rangle + |f_k(0)\rangle + |f_k(-1)\rangle)$ and FT is the single-qutrit Fourier transform

$$FT = \frac{1}{\sqrt{3}} \begin{pmatrix} \omega & 1 & \omega^{-1} \\ 1 & 1 & 1 \\ \omega^{-1} & 1 & \omega \end{pmatrix} \tag{76}$$

using ω as the cube root of unity **Eq. 14**. The process starts with state $|1\rangle$ undergoing FT and becoming $|\psi_1\rangle$ as $FT|1\rangle = |\psi_1\rangle = \omega|1\rangle + |0\rangle + \omega^{-1}|-1\rangle$. Then we obtain $|\psi_k\rangle$ by applying U_{f_k} to $|\psi_1\rangle$. It is easy to show that

$$|\psi_1\rangle = \omega^{-1}|\psi_2\rangle = \omega|\psi_3\rangle \tag{77}$$

and, similarly,

$$|\psi_4\rangle = \omega^{-1}|\psi_5\rangle = \omega|\psi_6\rangle. \tag{78}$$

Hence, application of U_{f_k} on $|\psi_1\rangle$ gives $|\psi_1\rangle$ (up to a phase factor) for an even permutation and $|\psi_4\rangle = FT|-1\rangle$ for an odd permutation. Thus, applying inverse Fourier transform FT^{-1} at the end, we measure $|1\rangle$ for even f_k and $|-1\rangle$ for odd f_k. We are able to determine the parity of f_k by a single application of f_k on a single qutrit.

Generalizing to a d–dimensional qudit system,

$$|\psi_k\rangle := \frac{1}{\sqrt{d}}\sum_{k'=1}^{d}\omega^{(k'-1)(k-1)}|k'\rangle. \tag{79}$$

In this scenario, a positive cyclic permutation maps $|\psi_2\rangle$ onto itself whereas negative permutations give $|\psi_d\rangle$. We then measure the results after applying an inverse Fourier transform to solve for the parity of the permutation. This algorithm has been implemented on the NMR system for both the qutrit [48] and quart [62] cases. It is also realized on a linear optic system [163]. Although the model problem has no significant applications and the speedup in the higher dimensional cases is not exponential, this proposed algorithm provides an elegant yet simple example for quantum computation without entanglement.

3.1.2 Qudit Deutsch-Jozsa Algorithm

Deutsch algorithm (with its origin in Ref. 44 and improved in Ref. 33 is one of the simplest examples to show the speed advantage of quantum computation. Deutsch-Jozsa algorithm is n-qubits generalization of the Deutsch algorithm. Deutsch-Jozsa algorithm can determine if a function $f(x)$ is *constant*, with constant output, or *balanced*, that gives equal instances of both outputs [125]. The process itself consists of only one evaluation of the function $f(x)$. In this algorithm, Alice sends Bob N qubits in the query register and one in the answer register where Bob applies the function to the query register qubits and stores the results in the answer register. Alice can measure the qubits in the query register to determine whether Bob's function is constant or balanced. This algorithm makes use of the superposition property of the qubit and reduces the minimum number of the function call from $2^n/2+1$ classically to only 1 with quantum algorithm. This gives another example of the advantages of quantum algorithms.

The Deutsch-Jozsa algorithm can be performed in the qudit system with a similar setup. Furthermore, with the qudit system, Deutsch-Jozsa algorithm can also find the closed expression of an affine function accurate to a constant term [53]. The *constant* and *balanced* function in the n dimensional qudit case have the following definition: "An r-qudit multi-valued function of the form

$$f:\{0,1,\ldots,n-1\}^r \longrightarrow \{0,1,\ldots,n-1\} \tag{80}$$

is *constant* when $f(x)=f(y)\,\forall x,y\in\{0,1,\ldots,n-1\}^r$ and is *balanced* when an equal number of the n^r domain values, namely n^{r-1}, is mapped to each of the n elements in the co-domain" [53].

It can be shown that all of the affine functions of r qudits

$$f(x_1,\ldots,x_r) := A_0\oplus A_1x_1\oplus\cdots\oplus A_rx_r,\ A_0,\ldots,A_r\in\mathbb{Z}_{\mathrm{n}}, \tag{81}$$

can be categorized to either constant or balanced functions [53]. If all the coefficients $A_{i\neq 0}=0$ then the function is constant. For affine function with non-zero coefficient $A_{i\neq 0}$, every element in its domain $\{0,1,\ldots,n-1\}^r$ is reducible modulo n to a unique element $m\in\{0,1,\ldots,n-1\}$. As $f(p)=f(q)$ if $p\equiv q\,(\mathrm{mod}\,n)$, each of the elements in the codomain $\{0,1,\ldots,n-1\}$ is mapped to n^{r-1} different elements in the domain. To finish the proof of the n-nary Deutsch-Jozsa algorithm, another trivial lemma is needed: Primitive n^{th} roots of unity satisfy $\sum_{k=0}^{n-1}\omega^{\alpha k}=0$ for nonzero integers α.

The circuit of the Deutsch-Jozsa algorithm in qudits is shown in **Figure 9**. This algorithm of r qudits can both distinguish whether a function U_f is balanced or constant and verify a closed expression for an affine function in U_f within a constant term which is a universal phase factor of the x-register and thus is lost during the measurement. The other coefficients of the affine function $A_1,\ldots,A_r$ are determined by measuring the state of the x-register at the output, $|A_1,\ldots,A_r\rangle$.

A detailed derivation of the circuit has been shown [53]; but the reasoning is an analogy to the qubit version of the Deutsch-Jozsa algorithm. If the function U_f is constant, the final state after the measurement is $|0\rangle^{\otimes r}|n-1\rangle$ as for $j\neq 0$ every states in the x-register have null amplitudes. Therefore, if every x-register qudit yields $|0\rangle$, it is a constant function; otherwise the function is balanced.

The Deutsch-Jozsa algorithm in the qudit system shares the same idea while enabling more applications such as determining the closed form of an affine function. Although this algorithm is mainly of theoretical interest, the n-nary version of it may have applications in image processing. It has the potential to distinguish between maps of texture in a Marquand chart since the images of which are encoded by affine functions [121]. This algorithm can also be modified to set up a secure quantum key-distribution protocol [121]. Other proposed Deutsch-Jozsa algorithms exist such as a method that makes use of the artificially allocated (subsystems) as qudits [88] and a generalized algorithm on the virtual spin representation [86].

3.1.3 Qudit Generalization of the Bernstein-Vazirani Algorithm

In **Section 3.1.2** we have discussed an application of a qudit Deutsch-Jozsa algorithm (DJA): verify a closed expression of an affine function. This application is closely related to the Bernstein-Vazirani algorithm discussed in this section. Given an input string and a function that calculates the bit-wize inner-product of the input string with an unknown string, the Bernstein-Vazirani algorithm determines the unknown string [12]. This algorithm can be treated as an extension of the Deutsch-Jozsa algorithm.

The qudit generalization of the Bernstein–Vazirani algorithm can determine a number string of integers modulo d encoded in the oracle function [95, 119]. First we introduce a positive integer d and consider the problem in modulo d throughout. Given an N-component natural number string

$$g(a) := \big(g(a_1),g(a_2),g(a_3),\ldots,g(a_N)\big),\ g(a_j)\in\{0,1,\ldots,d-1\}, \tag{82}$$

we define

$$\begin{aligned} f(x) &:= g(a)\cdot x \bmod d \\ &= g(a_1)x_1 + g(a_2)x_2 + \ldots + g(a_N)x_N \bmod d, \end{aligned} \tag{83}$$

for

$$x = (x_1, x_2, \ldots, x_N) \in \{0, 1, \ldots, d-1\}^N. \tag{84}$$

The oracle in the algorithm applies $f(x)$ to the input string x and computes the result, namely, the number string $g(a)$ encoded in the function $f(x)$.

The input state x is chosen to be $|\psi_0\rangle = |0\rangle^{\otimes N}|d-1\rangle$, where $|0\rangle^{\otimes N}$ means initialization of the N control-qudits into their $|0\rangle$ states and $|d-1\rangle$ means the target qudit is in its $d-1$ state. Quantum Fourier transforms of the pertinent input states are

$$|0\rangle \overset{QFT}{\longrightarrow} \sum_{y=0}^{d-1} \frac{|y\rangle}{\sqrt{d}} \tag{85}$$

and

$$|d-1\rangle \overset{QFT}{\longrightarrow} \sum_{y=0}^{d-1} \frac{1}{\sqrt{d}} \omega^{d-y}|y\rangle,$$

for ω a root of unity **Eq. 14**. The component-wize Fourier transform of a string encoded in the state $|x_1x_2\ldots x_N\rangle$ is

$$|x_1x_2\ldots x_N\rangle \overset{QFT}{\longrightarrow} \sum_{z\in K} \frac{\omega^{x\cdot z}|z\rangle}{\sqrt{d^N}}, \tag{86}$$

where

$$K = \{0, 1, \ldots, d-1\}^N,\ z := (z_1, z_2, \ldots, z_N). \tag{87}$$

We denote the Fourier transform of the $|d-1\rangle$ state as $|\phi\rangle$ and the input state after the Fourier transform is

$$|\psi_1\rangle = \sum_{x\in K} \frac{|x\rangle}{\sqrt{d^N}} |\phi\rangle \tag{88}$$

Now we introduce the oracle as the $O_{f(x)}$ gate such that

$$|x\rangle|j\rangle \overset{O_{f(x)}}{\longrightarrow} |x\rangle|(f(x)+j) \bmod d\rangle, \tag{89}$$

where

$$f(x) = g(a)\cdot x \bmod d. \tag{90}$$

By applying the $O_{f(x)}$ gate to $|\psi_1\rangle$ and following the formula by phase kick-back, we obtain the output state

$$O_{f(x)}|\psi_1\rangle = |\psi_2\rangle = \sum_{x\in K} \frac{\omega^{f(x)}|x\rangle}{\sqrt{d^N}} |\phi\rangle. \tag{91}$$

Finally, obtain the $|\psi_3\rangle$ which is the state after inverse Fourier transform of the first N qudits of $|\psi_2\rangle$. By measuring the first N quantum state of $|\psi_3\rangle$ we can obtain the natural number string we want that is offset up to a constant

$$g(a_1), g(a_2), g(a_3), \ldots, g(a_N) \tag{92}$$

using a single query of the oracle function.

The Bernstein-Vazirani algorithm clearly demonstrates the power of quantum computing. It outperforms the best classical algorithm in terms of speed by a factor of N [95]. The qudit generalizations of the Bernstein-Vazirani algorithm helps us comprehend the potential of the qudit systems.

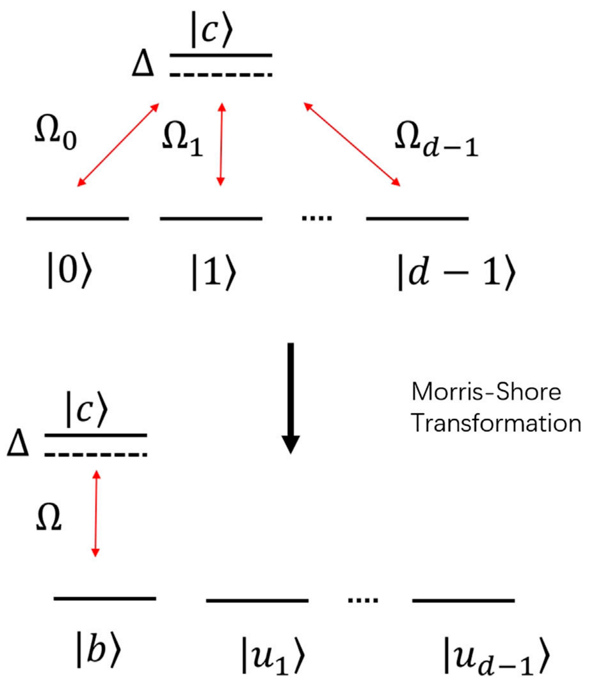


FIGURE 13 | Illustration of a qudit multipod linkage: the top is in the original basis and the bottom is in the Morris-Shore basis. Δ is a common detuning between a common (ancilla) state and other qudit states, Ω_k represents the single-photon Rabi frequencies. State $|b\rangle$ is a superposition of the qudit states weighted by the couplings Ω_k; $|u_n\rangle$ are the states that are not in the dynamics.

3.2 Qudit Algorithms for the Hidden Abelian Subgroup Problems

Many of the widely used quantum algorithms such as the discrete Fourier transform, the phase estimation and the factoring fit into the framework of the hidden subgroup problem (HSP). In this section, we review the qudit generalization of these algorithms. The qudit Fourier transform is discussed in **Section 3.2.1** and its application, the PEA is reviewed in **Section 3.2.2**. A direct application of these algorithms, Shor's factoring algorithm performed with qutrits and in metaplectic quantum architectures is also introduced **Section 3.2.2**.

3.2.1 Quantum Fourier Transform With Qudits

The quantum Fourier transform algorithm (QFT) is realizable on a qubit system [125]. QFT, as the heart of many quantum algorithms, can also be performed in a qudit system [145, 165]. In an N-dimensional system represented with n d-dimensional qudits, the QFT, $F(d, N)$, where $N = d^n$, transforms the computational basis

$$\{|0\rangle, |1\rangle, \ldots, |n-1\rangle\} \tag{93}$$

into a new basis set [26]

$$F(d,N)|j\rangle = \frac{1}{\sqrt{N}} \sum_{k=0}^{N-1} e^{2\pi i jk/N} |k\rangle. \tag{94}$$

For convenience, we write an integer j in a base-d form. If $j > 1$ then

$$j = j_1 j_2 \cdots j_n = j_1 d^{n-1} + j_2^{n-2} + \cdots + j_n d^0 \tag{95}$$

and, if $j < 1$, then

$$j = 0.j_1 j_2 \cdots j_n = j_1 d^{-1} + j_2 d^{-2} + \cdots + j_n d^{-n}. \tag{96}$$

The QFT acting on a state $|j\rangle$ can be derived and rewritten in a product form as

$$\begin{aligned} |j\rangle = |j_1 j_2 \cdots j_n\rangle &\mapsto \frac{1}{d^{n/2}} \sum_{k=0}^{d^n-1} e^{2\pi i jk/d^n} |k\rangle \\ &= \frac{1}{d^{n/2}} \sum_{k_1=0}^{d-1} \cdots \sum_{k_n=0}^{d-1} e^{2\pi i j\left(\sum_{l=1}^{n} k_l d^{-l}\right)} |k_1 k_2 \cdots k_n\rangle \\ &= \frac{1}{d^{n/2}} \sum_{k_1=0}^{d-1} \cdots \sum_{k_n=0}^{d-1} \bigotimes_{l=1}^{n} e^{2\pi i j k_l d^{-l}} |k_l\rangle \\ &= \frac{1}{d^{n/2}} \bigotimes_{l=1}^{n} \left[\sum_{k_l=0}^{d-1} e^{2\pi i j k_l d^{-l}} |k_l\rangle \right]. \end{aligned}$$

This process can be realized with the quantum circuit shown in **Figure 10**, and the fully expanded expression of the product form is shown on the right side of the figure. The generalized Hadamard gate H^d in the figure is defined as $H^d := F(d,d)$ which effects the transform

$$H^d |j_n\rangle = |0\rangle + e^{2\pi i 0.j_n}|1\rangle + \cdots + e^{2(d-1)\pi i 0.j_n}|d-1\rangle. \tag{97}$$

The matrix representation of H^d is

$$\begin{pmatrix} 1 & 1 & \cdots & 1 \\ 1 & e^{2\pi i 0.1} & \cdots & e^{2\pi i 0.(d-1)} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & e^{2(d-1)\pi i 0.1} & \cdots & e^{2(d-1)\pi i 0.(d-1)} \end{pmatrix}. \tag{98}$$

In the circuit the R_k^d gate is a phase gate that has the expression

$$R_k^d = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & e^{2\pi i/d^k} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & e^{2\pi i(d-1)/d^k} \end{pmatrix}. \tag{99}$$

The black dots in the circuit are multi-value-controlled gates that apply R_k^d to the target qudit j times for a control qudit in state $|j\rangle$. In order to complete the Fourier transform and ensure the correct sequence of $j_1 j_2 \cdots j_n$, a series of SWAP gates are applied at the end, which are not explicitly drawn in **Figure 10**.

The QFT developed in qudit system offers a crucial subroutine for many quantum algorithm using qudits. Qudit QFT offers superior approximations where the magnitude of the error decreases exponentially with d and the smaller error bounds are smaller [165]; which outperforms the binary case [34].

3.2.2 Phase-Estimation Algorithm With Qudits

With the qudit quantum Fourier transform, we are able to generalize the PEA to qudit circuits [26]. Similar to the PEA using qubit, the PEA in the qudit system is composed by two registers of qudits. The first register contains t qudits and t depends on the accuracy we want for the estimation. We assume that we can perform a unitary operation U to an arbitrary number of times using qudit gates and generate its eigenvector $|u\rangle$ and store it using the second register's qudits [17]. We want to calculate the eigenvalue of $|u\rangle$ where $U|u\rangle = e^{2\pi i r}|u\rangle$ by estimating the phase factor r.

The following derivations follow those in Ref. 26. For convenience, we rewrite the rational number r as

$$r = R/d^t = \sum_{k=0}^{t} \overline{R_l}/d^l = 0.\overline{R_1}\,\overline{R_2} \cdots \overline{R_t}. \tag{100}$$

As shown in **Figure 11A**, each qudit in the first register passes through the generalized Hadamard gate $H \equiv F(d,d)$. For the l^{th} qudit of the first register, we have

$$F(d,d)|0_l\rangle = \frac{1}{\sqrt{d}} \sum_{k_l=0}^{d-1} |k_l\rangle. \tag{101}$$

Then the l^{th} qudit is used to control the operation $U^{d^{t-l}}$ on the target qudits of the state $|u\rangle$ in the second register, which gives

$$CU^{d^{l-1}}|k\rangle \otimes |u\rangle = |k\rangle \left(U^{d^{t-l}}\right)^k |u\rangle = e^{2\pi i k d^{t-l} r}|k\rangle \otimes |u\rangle. \tag{102}$$

Note that the function of the controlled operation $CU^{d^{t-l}}$ can be considered as a "quantum multiplexer" [24, 87, 139]. After executing all the controlled operations on the qudits, the qudit system state turns out to be

$$\left(\prod_{l=1}^{t} \otimes \frac{1}{\sqrt{d}} \sum_{k_l=0}^{d-1} e^{2\pi i k_l d^{t-l} r} |k_l\rangle \right) \otimes |u\rangle. \tag{103}$$

Therefore, through a process called the "phase kick-back", the state of the first register receives the phase factor and becomes

$$|\text{Register } 1\rangle = \frac{1}{d^{t/2}} \sum_{k=0}^{d^t-1} e^{2\pi i rk} |k\rangle. \tag{104}$$

The eigenvalue r which is represented by the state $|R\rangle$ can be derived by applying the inverse QFT to the qudits in the first register:

$$F^{-1}\left(d, d^t\right)|\text{Register } 1\rangle = |R\rangle. \tag{105}$$

The whole process of PEA is shown in **Figure 11B**. To obtain the phase $r = R/d^t$ exactly, we can measure the state of the first register in the computational basis.

The PEA in qudit system provides a significant improvement in the number of the required qudits and the error rate decreases exponentially as the qudit dimension increases [129]. A long list of PEA applications includes Shor's factorization algorithm [142]; simulation of quantum systems [1]; solving linear equations [69, 128]; and quantum counting [147]. To give some examples, a quantum simulator utilizing the PEA algorithm has been used to calculate the molecular ground-state energies [8] and to obtain the energy spectra of molecular systems [13, 41, 42, 84, 154]. Recently, a method to solve the linear system using a qutrit version of the PEA has been proposed [138]. The qudit version of the PEA opens the possibility to realize all those applications that have the potential to out-perform their qubit counterparts.

Shor's quantum algorithm for prime factorization gives an important example of super-polynomial speed-up offered by a quantum algorithm over the currently-available classical algorithms for the same purpose [143]. The order-finding algorithm at the core of the factoring algorithm is a direct application of the PEA. With the previous discussion on the qudit versions of the quantum Fourier transform and phase estimation, we have the foundation to generalize Shor's factoring algorithm to the higher dimensional qudit system. Several proposals for performing Shor's algorithm on the qudit system, such as the adiabatic quantum algorithm of two qudits for factorization [166]; exist. This method makes use of a time-dependent effective Hamiltonian in the form of a sequence of rotation operators that are selected accoding to the qudit's transitions between its neighboring levels.

Another proposal carries out a computational resource analysis on two quantum ternary platforms [17]. One is the "generic" platform that uses magic state distillation for universality [25]. The other, known as a metaplectic topological quantum computer (MTQC), is a non-Abelian anyonic platform, where anyonic braiding and interferomic measurement is used to achieved the universality with a relatively low cost [37, 38]. The article discusses two different logical solutions for Shor's period-finding function on each of the two platforms: one that encodes the integers with the binary subspace of the ternary state space and optimizes the known binary arithmetic circuits; the other encodes the integer directly in the ternary space using the arithmetic circuits stemming in Ref. 16. Significant advantages for the MTQC platform are found compared to the others. In particular the MTQC platform can factorize an n-bit number with $n + 7$ logical qutrits with the price of a larger circuit-depth. To sum up the comparison, the MTQC provides significant flexibility at the period finding algorithm for the ternary quantum computers.

3.3 Quantum Search Algorithm With Qudits

The quantum search algorithm, also known as Grover's algorithm, is one of the most important quantum algorithms that illustrates the advantage of quantum computing. Grover's algorithm is able to outperform the classical search algorithm for a large database. The size of the computational space in an n-qubit system is a Hilbert space of 2^n dimensions.

Since there is a practical limit for the number of working qubits, the working Hilbert space can be expanded by increasing the dimension of each carrier of information, i.e., using qudits and qudit gates. Several schemes of Grover's quantum search with qudits have been proposed, such as one that uses the discrete Fourier transform as an alternative to the Hadamard gate [54] or another d-dimensional transformation [101] for the construction of the reflection-about-average operator (also known as the diffusion operator). In this section, an instruction on setting up Grover's algorithm in the qudit system is reviewed as well as a proposal of a new way to build a quantum gate F that can generate an equal-weight superposition state from a single qudit state [79]. With the new gate F, it is easier to realize Grover's algorithm in a physical system and improve the overall efficiency of the circuit.

Grover's algorithm solves the unstructured search problem by applying Grover's oracle iteratively as shown in **Figure 12B**. To construct the oracle, we build qudit gates to perform the oracle function $f(x)$ that acts differently on the search target s as compared to all the others. The logic behind the algorithm is to amplify the amplitude of the marked state $|s\rangle$ with the oracle function, while attenuating the amplitudes of all the other states. The marked state is amplified enough to be located in $O(\sqrt{N})$ steps for an N dimensional search space. In each step Grover's oracle is executed one time. This oracle can be broken into two parts: (1) *Oracle query*. The oracle shifts the phase of the marked state $|s\rangle$ and leaving others unchanged by doing

$$\boldsymbol{R}_s(\phi_s) = 1 + \left(e^{i\phi_s} - 1\right)|s\rangle\langle s|. \tag{106}$$

(2) *Reflection-about-average*. This operation is a reflection about a vector $|a\rangle$ with a phase ϕ_a:

$$\boldsymbol{R}_a(\phi_a) = 1 + \left(e^{i\phi_a} - 1\right)|a\rangle\langle a|. \tag{107}$$

It is constructed by applying the generalized Hadamard gate H, applying phase shift to $|0\rangle$ state and then applying H again. It is straightforward to show that

$$H^{\otimes n}\boldsymbol{R}_0(\phi_a)H^{\otimes n} = \boldsymbol{R}_a(\phi_a).$$

The two steps combined form Grover's operator $\boldsymbol{G}$, which is one execute of Grover's iteration. This process of Grover's iteration $\boldsymbol{G}$ is shown in **Figure 12A**.

Building Grover's operator in a qudit system can be simplified both algorithmically and physically. The most important improvement can be achieved by replacing the Hadamard gate H with F which drives the single-qudit state $|0_k\rangle$ into an equal weight superposition state,

$$F|0_k\rangle = \sum_{q=0}^{d-1} \xi_q |q_k\rangle, \tag{108}$$

with $\left|\xi_q\right| = d^{-1/2}$, in all qudits ($k \in \{1, 2, \ldots, n\}$). The F function can be realized by a single physical interaction in a multipod system easily. The multipod system consists of d degenerate quantum states $|0\rangle, |1\rangle, \ldots, |d-1\rangle$. A common (ancilla) state $|c\rangle$ couples these states to each other by two-photon Raman processes, as illustrated in **Figure 13**. The root-mean-square (rms) Rabi frequency as the coupling factor of the two states is

$$\Omega(t) = \sqrt{\sum_{k=0}^{d-1} |\Omega_k(t)|^2}. \tag{109}$$

Then from the two-state solution, we can calculate the dynamics of the multipod [97].

This method of building F minimizes the number and the duration of algorithmic steps and thus is fast to implement and, in addition, it also provides better protection against detrimental effects such as decoherence or imperfections. Due to its conceptual simplicity, this method has applications in numerous physical systems. Thus, it is one of the most natural and simplest realizations of Grover's algorithm in qudits.

4 ALTERNATIVE MODELS OF QUANTUM COMPUTING WITH QUDITS

The gate-based description of quantum computing is useful to establish principles of quantum computing with qudits, similar to the case for qubits. There are various approaches to quantum computing besides the gate-based model, such as the measurement-based [134]; adiabatic quantum computing [3, 55] and topological quantum computing [57]. Qudit versions of these approaches are barely explored to date, and we summarize the current status of these studies below.

4.1 Measurement-Based Qudit Computing

Measurement-based quantum computing was introduced as an alternative approach to quantum computing whereby a highly entangled state, such as a cluster state [22] or its graph-state generalization [70]; is prepared and then computation is performed by sequential single-qubit measurements in bases that are determined by a constant number of previous measurement outcomes [123, 134]. Measurement-based quantum computing is appealing in settings where preparing a highly entangled many-qubit graph state is feasible, such as parallelized controlled-phase operations [134] or cooling to the ground-state of a special Hamiltonian [123].

Measurement-based qudit quantum computing is unexplored to date. Preparatory work on generalizing graph states, implicitly including the cluster-state special case, to qudit graph states has been reported [85]. Regarding implement, qudit-based approaches have only been reported for the error-correction aspect of measurement-based qubit quantum computing [82]. In this approach, the cluster state is envisioned as comprising qudits, with the high-dimensional nature of qudits serving to encode qubits for error correction. They propose continuous-variable realizations of a qudit cluster state in a continuous-variable setting [82].

4.2 Adiabatic Qudit Computing

Adiabatic quantum computing approaches quantum computing by encoding the solution of a computational problem as the ground-state of a Hamiltonian whose description is readily obtained; the solution is obtained by preparing the ground state of a Hamiltonian whose ground-state is efficiently constructed and then evolving slowly, according to the adiabatic condition, into a close approximation of the ground state of the Hamiltonian specifying the problem [55]. The advantage of adiabatic quantum computing is evident in its natural correspondence to quantizing satisfiability problems [55]; and current efforts to exploit adiabatic quantum computing focus on quantum annealing, which is a quantum generalization of the simulated annealing metaheuristic used for non-quantum global optimization problems [40, 56, 83].

Quantum annealing is an important branch of quantum computing, particularly at the commercial level exemplified by D-Wave's early and continuing work in this domain. As D-Wave researchers themselves point out, realistic solid-state devices treated as qubits are not actually two-level systems and higher-dimensional representations of the dynamics must be considered to model and simulate realistic solid-state quantum annealers. The effect of states outside the qubit space, namely the treatment of solid-state quantum annealing as qudit dynamics, has been studied carefully with conditions established for soundness of qubit approximations [5].

In fact the qudit nature of so-called superconduting qubits, i.e., the higher-dimensional aspects of the objects serving as qubits, is not just a negative feature manifesting as leakage error; remarkable two-qubit gate performance is achieved by exploiting adiabatic evolution involving avoided crossings with higher levels [10, 110] with this exploitation for fast, high-fidelity quantum gates extendable to three-qubit gates and beyond by exploiting intermediate qudit dynamics and avoided level crossings [160, 161]. Another suggestion for exploiting qudit dynamics concerns using a degenerate two-level system with the additional freedom perhaps improving the energy gap and thus increasing success probability [156].

A dearth of studies have taken place to date into qudit-based adiabatic quantum computing. The one proposal thus far concerns a quantum adiabatic algorithm for factorization on two qudits [166]. Specifically, they consider two qudits of possibly different dimensions, thus necessitating a hybrid two-qudit gate [39]. They propose a time-dependent effective Hamiltonian to realize this two-qudit gate and its realization as radio-frequency magnetic field pulses. For this model, they simulate factorization of each of the numbers 35, 21, and 15 for two quadrupole nuclei with spins 3/2 and 1, respectively, corresponding to qudit dimensions of 4 and 3, respectively.

4.3 Topological Quantum Computing With Qudits

Topological quantum computing offers advantages over other forms of quantum computing by reducing quantum error correction overheads by exploiting topological protection. Some work has been done on topological quantum computing with qudits by proposing quantum computing with parafermions [49, 74].

Majorana fermions are expected to exhibit non-abelian statistics, which makes these exotic particles, or their quasiparticle analogue, sought after for anyonic quantum computing [90]. Majorana fermions can be generalized to $\mathbb{Z}_\mathrm{d}$ parafermions, which also exhibit non-abelian statistics and reduce to standard Majorana fermions for $d = 2$. One advantage of $d > 2$ is that parafermion braiding is an entangling operation. Importantly, encoding a qudit of dimension d in the four-parafermion fusion space enables all single-qudit Clifford gates to be generated modulo phase terms [74].

Clifford gates do not provide a universal set of gates for quantum computing. A non-Clifford gate can be achieved for parafermions encoded into parafermion zero modes by exploiting the Aharonov-Casher effect, physically implemented by move a half-fluxon around the parafermionic zero modes. Combining this non-Clifford gate with the Clifford gates achieved by parafermion braiding yields a universal gate set of non-abelian quantum computing with qudits [49].

5 IMPLEMENTATIONS OF QUDITS AND ALGORITHMS

The qubit circuit and qubit algorithm have been implemented on various physical systems such as defects in solids [27, 81, 120]; quantum dots [104, 127]; photons [113, 132]; super conducting systems [29, 31]; trapped ions [14, 15]; magnetic [7, 18, 32, 148] and non-magnetic molecules [30, 152]. For each physical representation of the qubit, only two levels of states are used to store and process quantum information. However, many quantum properties of these physical systems have more than two levels, such as the frequency of the photon [106]; energy levels of the trapped ions [91]; spin states of the nuclear magnetic resonance systems [48] and the spin state of the molecular magnetic magnets [115]. Therefore, these systems have the potential to represent qudit systems. In this section, we briefly review several physical platforms that have been used to implement qudit gates or qudit algorithms.

Although most of the systems have three or four levels available for computation, they are extensible to higher level systems and scalable to multi-qudit interactions. These pioneer implementations of qudit systems show the potential of future realization of the more powerful qudit quantum computers that have real-life applications.

5.1 Time and Frequency Bin of a photon

Photonic system is a good candidate for quantum computing because photons rarely interact with other particles and thus have a comparatively long decoherence time. In addition, photon has many quantum properties such as the orbital angular momentum [9, 52]; frequency-bin [75, 76, 96, 107] and time-bin [73, 78] that can be used to represent a qudit. Each of these properties provides an extra degrees of freedom for the manipulation and computation. Each degree of freedom usually has dimensions greater than two and thus can be used as a unique qudit. The experimental realization of arbitrary multidimensional multiphotonic transformations has been proposed with the help of ancilla state, which is achievable via the introduction of a new quantum nondemolition measurement and the exploitation of a genuine high-dimensional interferometer [60]. Experimental entanglement of high-dimensional qudits, where multiple high-purity frequency modes of the photons are in a superposition coherently, is also developed and demonstrated [96].

Here we review a single photon system that has demonstrated a proof-of-principle qutrit PEA [106]. In a photonic system, there is no deterministic way to interact two photons and thus it is hard to build a reliable controlled gate for the photonic qudits. The following photonic system bypasses this difficulty via using the two degrees of freedom on a single photon—i.e., the time-bin and frequency-bin to be the two qutrits. The frequency degree of freedom carries one qutrit as the control register and the time degree of freedom carries another qutrit as the target register. The experimental apparatus consists of the well-established techniques and fiber-optic components: continuous-wave (CW) laser source, phase modulator (PM), pulse shaper (PS), intensity modulator (IM) and chirped fiber Bragg grating (CFBG). The device is divided into three parts [106]: 1) A state preparation part that comprises a PM followed by a PS and a IM that encodes the initial state to qudits; 2) a controlled-gate part that is built with a PM sandwiched by two CFBGs to perform the control-U operation; and 3) an inverse Fourier transformation comprising a PM and then a PS to extract the phase information. Note that the controlled-gate part can perform a multi-value-controlled gate that applies different operations based on the three unique states of the control qutrit. In the PEA procedure, eigenphases can be retrieved with 98% fidelity. In addition to having long coherence lifetime, the photonic system also has a unique advantage over other common quantum devices, i.e., the ability to process and measure thousands of photons simultaneously. This allows us to generate statistical patterns quickly and infer the phase accurately whereas the normal PEA has to use additional qudits on the control register to increase accuracy.

Here we provide an example for the statistical inference of the phase based on numerical data generated by the photonic PEA experiment just described. The two unitary operations used in the experimental setup are

$$\widehat{U}_1 = \operatorname{diag}(1, \omega, \omega^2), \tag{110}$$

with ω being the cube root of unity **Eq. 14**, and

$$\widehat{U_2} = \operatorname{diag}(1, e^{i0.351\pi}, e^{i1.045\pi}). \tag{111}$$

In the experiment, photonic qutrits are sent through the control and target registers and the state of the control register qutrits is measured and counted to obtain the phase information.

Given the eigenphase ϕ of an eigenstate of the target register, the probability for the qutrit output state to fell into \ketn, where $n \in \{0, 1, 2\}$, is

$$C(n, \phi) = \frac{1}{9}\left|1 + e^{i\left(\phi - \frac{n2\pi}{3}\right)} + e^{i2\left(\phi - \frac{n2\pi}{3}\right)}\right|^2. \tag{112}$$

Now let E_0, E_1, and E_2 be the counts of the photons that fell into $|0_f\rangle$, $|1_f\rangle$, and $|2_f\rangle$. The estimated phase, denoted $\tilde{\phi}$, is the phase that has the smallest the mean-square error between the measured and theoretical results:

$$\min_{\tilde{\phi}} \sum_{n=0}^{2} \left(E_n - C\left(n, \tilde{\phi}\right)\right)^2 \tag{113}$$

The estimated phases for $\widehat{U}_1$ (110) and $\widehat{U}_2$ (111) are shown in **Table 1** [106]. The first experiment with U_1 estimates the phase of a eigenvector and gives the eigenvalue. The second experiment with U_2 estimates the phase of a state with an arbitrary value (not a fraction of π), but, by repeating the experiment, the eigenvalue can be estimated from the statistical distribution of the results.

5.2 Ion Trap

Intrinsic spin, an exclusively quantum property, has an inherently finite discrete state space which is a perfect choice for representing qubit or qudit. When a charged particle has spin, it possess a magnetic momentum and is controllable by external electromagneic pulses. This concept leads to the idea of ion trap where a set of charged ions are confined by electromagnetic field. The hyperfine (nuclear spin) state of an atom, and lowest level vibrational modes (phonons) of the trapped atoms serves as good representations of the qudits. The individual state of an atom is manipulated with laser pulse and the ions interact with each other via a shared phonon state.

The set-up of an ion trap qutrit system reviewed here can perform arbitrary single qutrit gates and a control-not gate [91]. These two kinds of gates form a universal set and thus can be combined to perform various quantum algorithms such as those discussed in **Section 3**. The electronic levels of an ion are shown in **Figure 14**. The energy levels $|0\rangle, |1\rangle, |2\rangle$ are used to store the quantum information of a qutrit. The transition between the levels are driven by the classical fields $\Omega_{03}, \Omega_{13}, \Omega_{04}$, and Ω_{24} of the Raman transitions through independent channels linked to orthogonal polarizations. We first develop a system acting as a single qutrit gate that can manipulate the energy levels of the ion via Raman transitions driven by the classical fields. The following expressions follow those in Ref. 91. For single qutrit gates, where the center-of-mass motion is excluded, we can include the spatial dependence of the Raman fields as phase factors Δ and assuming the conditions

$$\Delta \gg \Omega_{04}, \Omega_{03}, \Omega_{31}, \Omega_{42}, \tag{114}$$

the effective Hamiltonian describing the ion in this system is

$$\frac{H}{\hbar} = -\frac{|\Omega_{31}|^2}{\Delta}|1\rangle\langle 1| - \frac{|\Omega_{42}|^2}{\Delta}|2\rangle\langle 2| - \frac{|\Omega_{30}|^2 + |\Omega_{40}|^2}{\Delta}|0\rangle\langle 0| - \tag{115}$$

$$-\left[\frac{\Omega_{31}\Omega_{30}^*}{\Delta}|0\rangle\langle 1| + \frac{\Omega_{42}\Omega_{40}^*}{\Delta}|0\rangle\langle 2| + \mathrm{hc}\right]. \tag{116}$$

Knowing the Hamiltonian we are able to derive the evolution operator in the restricted three-dimensional space spanned by $\{|2\rangle, |1\rangle, |0\rangle\}$ as the following

$$U(\varphi) = \begin{pmatrix} 1 + |g|^2 C(\varphi) & g g'^{\star} C(\varphi) & -ig \sin\varphi \\ g' g^{\star} C(\varphi) & 1 + |g'|^2 C(\varphi) & -ig' \sin\varphi \\ -ig^{\star} \sin\varphi & -ig'^{\star} \sin\varphi & \cos\varphi \end{pmatrix}, \tag{117}$$

where $\varphi = \Omega t$ represents interaction time and

$$C(\varphi) = \cos\varphi - 1, \ \Omega^2 = |\kappa'|^2 + |\kappa|^2. \tag{118}$$

The notation g and g' represents

$$g := \kappa/\Omega, \ g' = \kappa'/\Omega, \ \kappa := \Omega_{42}^* \Omega_{40}/\Delta, \ \kappa' = \Omega_{31}^* \Omega_{30}/\Delta. \tag{119}$$

This evolution operator can perform all kinds of the required coherent operations that are acting on any two of the logical states. It operates on the system and works essentially as a single qutrit gate. All kinds of transitions can be realized by manipulating the κ and κ' coupling. Therefore with the proper manipulation of the parameters κ and κ' we are able to perform any arbitrary one-qutrit gate as desired.

Single qutrit gate alone is not sufficient to form a universal computational set, as we need a conditional two-qutrit gate or a two-qutirt controlled-gate to achieve universality. To define the conditional two-qutrit gate we need an auxiliary level |0'⟩ as shown in **Figure 14**. The conditional two-qutrit gate is achievable via the center-of-mass (CM) motion of ions inside the trap. The ion CM coupled to the electronic transition $|0\rangle \rightarrow |q\rangle$ is described by the Hamiltonian

$$H_{n,q} = \frac{\Omega_q \eta}{2}\left[|q\rangle_n \langle 0| a e^{-i\delta t - i\phi} + a^\dagger |0_n\rangle\langle q| e^{i\delta t + i\phi}\right]. \tag{120}$$

Here a is the annihilation operator and $a^\dagger$ is the creation operator of the CM phonons. Ω_q is the effective Rabi frequency after adiabatic elimination of upper excited levels and ϕ is the laser phase, and δ is the detuning. The Lamb-Dicke parameter is

$$\eta := \sqrt{\hbar k_\theta^2 / (2M\nu_x)}. \tag{121}$$

This Hamiltonian governs the coherent interaction between qutrits and collective CM motion. With appropriate selection of effective interaction time and laser polarizations, the CM motion coupled to electronic transitions is coherently manipulated [91].

To complete the universal quantum computation requirements, we need to develop a measurement scheme. In this scheme, von Neumann measurements distinguishing three directions $|0\rangle$, $|1\rangle$, $|2\rangle$ are made possible via the resonant interactions from $|1\rangle$ and $|2\rangle$ to states $|3\rangle$ and $|4\rangle$, respectively. The single and two-qutrit controlled gate are combined to perform various qutrit algorithms such as the quantum Fourier transform. Other variations of the ion-trap qutrit quantum computer designs use trapped ions in the presence of a magnetic field gradient [111]. The qutrit ion-trap computer provides a significant increase of the available Hilbert space while demanding only the same amount of physical resources.

5.3 Nuclear Magnetic Resonance

Nuclear magnetic resonance (NMR) is an essential tool in chemistry and involves manipulating and detecting molecules' nuclear spin states using radio-frequency electromagnetic waves [19]. Some technologies of this field are sophisticated enough to control and observe thousands of nuclei in an experiment. The NMR has the potential to scale up quantum computer to thousands of qudits [144].

In this section we review the implementation of a single-qudit algorithm that can determine the parity of a permutation on an NMR system [48]. The algorithm itself is the parity determining algorithm explained in **Section 3.1.1**. The molecule in this NMR setup is embedded in a liquid crystalline environment and the strong magnetic field is used to adjust the anisotropic molecular

TABLE 1 | Normalized photon counts and comparison of the true phase ϕ and the experimentally estimated phase ϕ' for each eigenstate of $\hat{U}_1$ (**Eq. 110**) and $\hat{U}_2$ (**Eq. 111**) [106].

$\hat{U}_1$			
Eigenstate	$\vert 0_t\rangle$	$\vert 1_t\rangle$	$\vert 2_t\rangle$
E_0	0.9948 ± 0.0004	0.0101 ± 0.0004	0.0122 ± 0.0005
E_1	0.0023 ± 0.0002	0.9805 ± 0.0009	0.0120 ± 0.0005
E_2	0.0029 ± 0.0002	0.0094 ± 0.0004	0.9758 ± 0.0010
True Phase, ϕ	0	$2\pi/3$	$4\pi/3$
Est. Phase, $\tilde{\phi}$	1.972π	0.612π	1.394π
Error, $\frac{\vert\phi-\tilde{\phi}\vert}{2\pi}$	1.4%	2.7%	3.0%
$\hat{U}_2$			
Eigenstate	$\vert 0_t\rangle$	$\vert 1_t\rangle$	$\vert 2_t\rangle$
E_0	0.878 ± 0.002	0.316 ± 0.003	0.143 ± 0.002
E_1	0.032 ± 0.001	0.530 ± 0.003	0.318 ± 0.003
E_2	0.090 ± 0.002	0.154 ± 0.002	0.539 ± 0.003
True Phase, ϕ	0	0.3511π	1.045π
Est. Phase, $\tilde{\phi}$	1.859π	0.377π	1.045π
Error, $\frac{\vert\phi-\tilde{\phi}\vert}{2\pi}$	7.1%	1.3%	0.0%

orientation. This adding a finite quadrupolar coupling term to the Hamiltonian which is as follows

$$H = -\omega_0 I_z + \Lambda\left(3I_z^2 - I^2\right), \tag{122}$$

where $\Lambda = e^2qQS/4$ is the effective value of the quadrupolar coupling [48]. The Fourier transformation is implemented by a sequence of three transition-selective pulses. A series of combinations of 180° pulses, both transition-selective and non-selective, is used to implement the permutations.

Final states of the system can be derived from a single projective measurement. Pseudopure spin states act as approximation of effect of the system on an ensemble NMR quantum computer since it is impossible to do the true projective measurements [98]. The fidelity measurement of the experiment is given as

$$F := \frac{\mathrm{tr}\left(\rho_{\mathrm{th}}^{\dagger}\rho_{\mathrm{expt}}\right)}{\sqrt{\mathrm{tr}\left(\rho_{\mathrm{th}}^{\dagger}\rho_{\mathrm{th}}\right)}\sqrt{\mathrm{tr}\left(\rho_{\mathrm{expt}}^{\dagger}\rho_{\mathrm{expt}}\right)}} \tag{123}$$

is used, where ρ_{th} and ρ_{expt} are, respectively, theoretically expected and experimentally obtained density matrices. Fidelities obtained for these proposed operations are 0.92 and above.

Another set-up of the same algorithm treats a single quart [62]. The algorithm implementation is achieved using a spinâ€“$\frac{3}{2}$ nuclei, which is commonly selected for NMR-QIP applications. In their NMR systems the four energy levels needed is made via the Zeeman splitting using a strong static magnetic field. All of the two implementations of the single-qudit algorithm show that the NMR system provides a way to realize a reliable and efficient qudit system for the quantum computing.

5.4 Molecular Magnets

Molecular quantum magnets, also called the single-molecule magnets (SMM), provides another physical representation of qudits [115]. They have phenomenal magnetic characteristics and can be manipulated via chemical means. This enables the alternation of the ligand field of the spin carriers and the interaction between the SMM with the other units. As pointed out in one of the proposals, the nuclear spin states of the molecules, which have a long life-time, are used to store the quantum information. This information is read out by the electronic states. In the mean time, the robustness of the molecule allows it to conserve its molecular, electronic and magnetic characteristics at high temperatures [116].

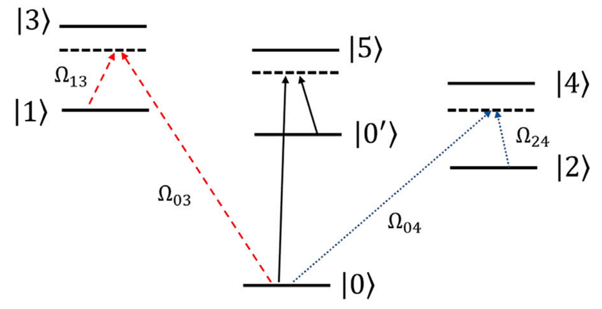


FIGURE 14 | Electronic level structure of the trapped ion. The carrier of the quantum information is the qutrit states |0⟩, |1⟩, and |2⟩. |0'⟩ is an auxiliary level used for the conditional two-qutrit gate.

As one of the SMMs, the single molecule TbPc$_2$ complex reviewed in this section possesses all necessary properties such as long lifetime and robustness. These properties are integrated as important components of a serious quantum mechanical devices, for examples, resonator [59]; molecular spin valve [149] and transistor [138, 146]. TbPc$_2$ gains its SMM properties from the strong spinâ€“orbit coupling of lanthanide ions and the ligand field [77]. Magnetic properties of TbPc$_2$ are governed by the Hamiltonian:

$$\mathscr{H} = \mathscr{H}_{\mathrm{lf}} + g_J\mu_0\mu_B J\cdot H + A_{\mathrm{hf}} I\cdot J + \left(I_z^2 - \frac{1}{3}(I+1)I\right), \tag{124}$$

where $\mathscr{H}_{\mathrm{lf}}$ is the ligand field Hamiltonian (lf), and $g_J\mu_0\mu_B J\cdot H$ represents the Zeeman energy. $A_{\mathrm{hf}} I\cdot J$ accounts for hyperfine interactions (hf) and $\left(I_z^2 - \frac{1}{3}(I+1)I\right)$ is the quadrupole term. A sweeping magnetic field associated with $m_I = \pm\frac{1}{2}$ and $\pm\frac{3}{2}$ can cause quantum tunneling of magnetization, which preserves nuclear spin while changing electronic magnetic moment. This field enables nuclear-spin measurement by suspending the TbPc$_2$ molecule on carbon nanotubes (CNT) and between gold junctions.

This measurement uses the technique of electro-migration. Initialization and manipulation of the four spin states of TbPc$_2$ can be obtained from QTM transitions driven by external ramping magnetic field. The transitions between the $\left|+\frac{1}{2}\right\rangle \leftrightarrow \left|-\frac{1}{2}\right\rangle$ states and $\left|+\frac{3}{2}\right\rangle \leftrightarrow \left|-\frac{3}{2}\right\rangle$ is achieved via applying appropriate resonate frequencies ν_{12} and ν_{23}. Relaxation and coherence times are important aspects to be analyzed for the TbPc$_2$ system, and this process is accomplished by imaging the initialized nuclear spin trajectory in real-time.

Statistical analysis of the nuclear spin coherence time makes use of the spinâ€“lattice relaxation times by fitting the data for an exponential form $(y = \exp(-t/T_1))$ and yields $T_1 \approx 17$ s for $m_I = \pm\frac{1}{2}$ and $T_1 \approx 34s$ for $m_I = \pm\frac{3}{2}$ with fidelities of $F\left(m_I = \pm\frac{1}{2}\right) \approx 93\%$ and $F\left(m_I = \pm\frac{3}{2}\right) \approx 87\%$ accordingly [115]. The TbPc$_2$ SMM can be used to execute Grover's algorithm, where the alternation of the m_I state contained in the TbPc$_2$ molecular qubit are treated by resonance frequencies [63, 99].

6 SUMMARY AND FUTURE OUTLOOK OF QUDIT SYSTEM

6.1 Summary of the Advantages of Qudit Systems Compared to Qubit Systems

Throughout the article we discuss and review many aspects of the qudit systems such as qudit gates, qudit algorithms, alternative computation models and implementations. Most gates and algorithms based on qudits have some advantages over those for qubits, such as shorter computational time, lower requirement of resources, higher availability, and the ability to solve more complex problems. The qudit system, with its high-dimensional nature, can provide more degrees of freedom and larger computational space. This section summarizes the advantages of the qudit system compared to the qubit system.

Qudit gates have the advantage of a larger working Hilbert space which reduces the number of qudits needed to represent an arbitrary unitary matrix. In our discussion of universality in **Section 2.1.2**, the qudit method proposed by Muthukrishnan and Stroud's has a $(\log_2 d)^2$ scaling advantage over the qubit case. Furthermore, Luo and Wang show that with their proposed universal computation scheme [108]; there is an extra factor of n reduction in the gate requirement, where n is the number of qudits. By introducing qudits to the construction of some well-known gates such as the Toffoli gate, the elementary gate required are reduced from $12n - 11$ gates in the qubit case to $2n - 1$ gates by introducing a single $(n+1)$-level target carrier [133] and to $2n - 3$ gates by utilizing the topological properties [89]. In our discussion of the geometrically quantified qudit-gate efficiency in **Section 2.3**, the qubit system needs $O(n^6 d(I,U)^3)$ one- and two-qubit gates to synthesize a unitary [122] while in the qutrit case the lower bound is $O(n^k d(I,U)^3)$ where k is an integer that depends on the accuracy of the approximation and can be smaller than 6 [100].

For many of the physical systems such as photons [113, 132]; super conducting systems [29, 31]; trapped ions [14, 15]; magnetic [7, 18, 32, 148] and non-magnetic molecules [30, 152] there are usually more than two available physical states available for the applications. The qudit system has a higher efficiency utilizing those extra states than the qubit system. Also using the photonic system, we can perform the multi-level controlled gate (**Section 2.2**) which can perform multiple control operations at and same time and largely reduce the number of controlled gates requirement [106].

Other than computation, the qudit also has advantages in quantum communication as it possesses a higher noise resilience than the qubit [36]. The qudit system has a higher quantum bit error rate (QBER), which is a measure of resistance to the environmental noise or eavesdropping attacks, compared to the qubit system. The higher noise tolerance of the qudits helps to increase the secret key rate as it can be shown that the secret key rate increases as the Hilbert space dimensions increase at the same noise level [140]. Notice that in practical situation, the qudit system performed on each particular physical apparatus has varied amount of advantages than the qubit and there might be cases in which the high-dimensional states have a higher transmission distance [36]. This higher noise resilience of qudits is more advantageous if the qudits are entangled. The entanglement becomes more robust by increasing the dimension of the qudits while fixing their numbers. In other words, as the noise sources act locally on every system, increasing the dimension d will reduce the number of systems and thus reduce the effect of noise resulting in the robustness increase [103]. The increasing noise level tolerance as the qudit dimension increases can be shown on an photonic OAM system as an example of its implementation [51].

In summary the qudit system possesses advantages in the circuit design, physical implementation and has the potential to outperform the qubit system in various applications.

6.2 Future Outlook of Qudit System

This review article introduces the basics of the high-dimensional qudit systems and provides details about qudit gates, qudit algorithms and implementations on various physical systems. The article serves as a summary of recent developments of qudit quantum computing and an introduction for newcomers to the field of qudit quantum computing. Furthermore we show the advantages and the potential for qudit systems to outperform qubit counterparts. Of course these advantages can come with challenges such as possibly harder-to-implement universal gates, benchmarking [80, 94, 117]; characterization of qudit gate [68, 136] and error correction connected with the complexity of the Clifford hierarchy for qudits [157].

Compared to qubit systems, qudit systems currently have received less attention in both theoretical and experimental studies. However, qudit quantum computing is becoming increasingly important as many topics and problems in this field are ripe for exploration. Extending from qubits to qudits uses in some mathematical challenges, with these mathematical problems elegant and perhaps giving new insights into quantum computing in their own right. Connections between quantum resources such as entanglement, quantum algorithms and their improvements, scaling up qudit systems both to higher dimension and to more particles, benchmarking and error correction, and the bridging between qudits and continuous-variable quantum computing [67] are examples of the fantastic research directions in this field of high-dimensional quantum computing.

AUTHOR CONTRIBUTIONS

All authors discussed the relevant materials to be added and all participated in writing the article.

ACKNOWLEDGMENTS

We would like to acknowledge the financial support by the National Science Foundation under award number 1839191-ECCS. BCS appreciates financial support from NSERC and from the Alberta government.

REFERENCES

1. Abrams DS, Lloyd S. Quantum algorithm providing exponential speed increase for finding eigenvalues and eigenvectors. *Phys Rev Lett* (1999) **83**:5162–5. doi:10.1103/PhysRevLett.83.5162
2. Adcock MRA, Høyer P, Sanders BC. Quantum computation with coherent spin states and the close Hadamard problem. *Quant Inf Process* (2016) **15**:1361–86. doi:10.1007/s11128-015-1229-0
3. Aharonov D, van Dam W, Kempe J, Landau Z, Lloyd S, Regev O. Adiabatic quantum computation is equivalent to standard quantum computation. *SIAM J Comput* (2007) **37**:166–94. doi:10.1137/S0097539705447323
4. Alber G, Delgado A, Gisin N, Jex I. Efficient bipartite quantum state purification in arbitrary dimensional Hilbert spaces. *J Phys Math Gen* (2001) **34**:8821–33. doi:10.1088/0305-4470/34/42/307
5. Amin MHS, Dickson NG, Smith P. Adiabatic quantum optimization with qudits. *Quant Inf Process* (2013) **12**:1819–29. doi:10.1007/s11128-012-0480-x
6. Anwar H, Campbell ET, Browne DE. Qutrit magic state distillation. *New J Phys* (2012) **14**:063006. doi:10.1088/1367-2630/14/6/063006
7. Aromí G, Aguilà D, Gamez P, Luis F, Roubeau O. Design of magnetic coordination complexes for quantum computing. *Chem Soc Rev* (2012) **41**: 537–46. doi:10.1039/c1cs15115k
8. Aspuru-Guzik A, Dutoi AD, Love PJ, Head-Gordon M. Simulated quantum computation of molecular energies. *Science* (2005) **309**:1704–7. doi:10.1126/science.1113479
9. Babazadeh A, Erhard M, Wang F, Malik M, Nouroozi R, Krenn M, et al. High-dimensional single-photon quantum gates: concepts and experiments. *Phys Rev Lett* (2017) **119**:180510. doi:10.1103/PhysRevLett.119.180510
10. Barends R, Kelly J, Megrant A, Veitia A, Sank D, Jeffrey E, et al. Superconducting quantum circuits at the surface code threshold for fault tolerance. *Nature* (2014) **508**:500–3. doi:10.1038/nature13171
11. Bartlett SD, de Guise H, Sanders BC. Quantum encodings in spin systems and harmonic oscillators. *Phys Rev A* (2002) **65**:052316. doi:10.1103/PhysRevA.65.052316
12. Bernstein E, Vazirani U. Quantum complexity theory. *SIAM J Comput* (1997) **26**:1411–73. doi:10.1137/S0097539796300921
13. Bian T, Murphy D, Xia R, Daskin A, Kais S. Quantum computing methods for electronic states of the water molecule. *Mol Phys* (2019) **117**:2069–82. doi:10.1080/00268976.2019.1580392
14. Blatt R, Wineland D. Entangled states of trapped atomic ions. *Nature* (2008) **453**:1008–15. doi:10.1038/nature07125
15. Bloch I. Quantum coherence and entanglement with ultracold atoms in optical lattices. *Nature* (2008) **453**:1016–22. doi:10.1038/nature07126
16. Bocharov A, Cui SX, Roetteler M, Svore KM. Improved quantum ternary arithmetic. *Quant Inf Comput* (2016) **16**:862–884. doi:10.5555/3179473.3179481
17. Bocharov A, Roetteler M, Svore KM. Factoring with qutrits: shor's algorithm on ternary and metaplectic quantum architectures. *Phys Rev A* (2017) **96**: 012306. doi:10.1103/PhysRevA.96.012306
18. Bogani L, Wernsdorfer W. Molecular spintronics using single-molecule magnets. In *Nanoscience and technology: a collection of reviews from nature journals*. Singapore: World Scientific (2010) p. 194–201.
19. Bovey FA, Mirau PA, Gutowsky H. *Nuclear magnetic resonance spectroscopy*. New York, NY: Elsevier (1988)
20. Boykin PO, Mor T, Pulver M, Roychowdhury V, Vatan F. A new universal and fault-tolerant quantum basis. *Inf Process Lett* (2000) **75**:101–7. doi:10.1016/s0020-0190(00)00084-3
21. Brennen G, O'Leary D, Bullock S. Criteria for exact qudit universality. *Phys Rev A* (2005) **71**:052318. doi:10.1103/PhysRevA.71.052318
22. Briegel HJ, Raussendorf R. Persistent entanglement in arrays of interacting particles. *Phys Rev Lett* (2001) **86**:910–3. doi:10.1103/PhysRevLett.86.910
23. Brylinski J-L, Brylinski R. Universal quantum gates In: Brylinski RK, Chen G, editors *Mathematics of quantum computation*. Boca Raton, FL Chapman and Hall/CRC (2002) p. 117–34.
24. Bullock S, O'Leary D, Brennen G. Asymptotically optimal quantum circuits ford-level systems. *Phys Rev Lett* (2005) **94**:230502. doi:10.1103/PhysRevLett.94.230502
25. Campbell ET, Anwar H, Browne DE. Magic-state distillation in all prime dimensions using quantum Reed-Muller codes. *Phys Rev X* (2012) **2**:041021. doi:10.1103/PhysRevX.2.041021
26. Cao Y, Peng S-G, Zheng C, Long G-L. Quantum fourier transform and phase estimation in qudit system. *Commun Theor Phys* (2011) **55**:790–4. doi:10.1088/0253-6102/55/5/11
27. Childress L, Gurudev Dutt MV, Taylor JM, Zibrov AS, Jelezko F, Wrachtrup J, et al. Coherent dynamics of coupled electron and nuclear spin qubits in diamond. *Science* (2006) **314**:281–5. doi:10.1126/science.1131871
28. Childs AM. Secure assisted quantum computation (2001) arXiv:quant-ph/0111046.
29. Chiorescu I, Nakamura Y, Harmans CJPM, Mooij JE. Coherent quantum dynamics of a superconducting flux qubit. *Science* (2003) **299**:1869–71. doi:10.1126/science.1081045
30. Chuang IL, Vandersypen LMK, Zhou X, Leung DW, Lloyd S. Experimental realization of a quantum algorithm. *Nature* (1998) **393**:143. doi:10.1038/30181
31. Clarke J, Wilhelm FK. Superconducting quantum bits. *Nature* (2008) **453**: 1031–42. doi:10.1038/nature07128
32. Clemente-Juan JM, Coronado E, Gaita-Ariño A. Magnetic polyoxometalates: from molecular magnetism to molecular spintronics and quantum computing. *Chem Soc Rev* (2012) **41**:7464–78. doi:10.1039/c2cs35205b
33. Cleve R, Ekert A, Macchiavello C, Mosca M. Quantum algorithms revisited. *Proc Roy Soc Lond A* (1998) **454**:339–54. doi:10.1098/rspa.1998.0164
34. Coppersmith D. An approximate Fourier transform useful in quantum factoring(2002) arXiv preprint quant-ph/0201067.
35. Cory DG, Price MD, Maas W, Knill E, Laflamme R, Zurek WH, et al. Experimental quantum error correction. *Phys Rev Lett* (1998) **81**:2152–5. doi:10.1103/PhysRevLett.81.2152
36. Cozzolino D, Da Lio B, Bacco D, Oxenløwe LK. High-dimensional quantum communication: benefits, progress, and future challenges. *Adv Quantum Tech* (2019) **2**:1900038. doi:10.1002/qute.201900038
37. Cui SX, Hong S-M, Wang Z. Universal quantum computation with weakly integral anyons. *Quant Inf Process* (2015) **14**:2687–727. doi:10.1007/s11128-015-1016-y
38. Cui SX, Wang Z. Universal quantum computation with metaplectic anyons. *J Math Phys* (2015) **56**:032202. doi:10.1063/1.4914941
39. Daboul J, Wang X, Sanders BC. Quantum gates on hybrid qudits. *J Phys Math Gen* (2003) **36**:2525–36. doi:10.1088/0305-4470/36/10/312
40. Das A, Chakrabarti BK. Colloquium: quantum annealing and analog quantum computation. *Rev Mod Phys* (2008) **80**:1061–81. doi:10.1103/RevModPhys.80.1061
41. Daskin A, Grama A, Kollias G, Kais S. Universal programmable quantum circuit schemes to emulate an operator. *J Chem Phys* (2012) **137**:234112. doi:10.1063/1.4772185
42. Daskin A, Kais S. Decomposition of unitary matrices for finding quantum circuits: application to molecular hamiltonians. *J Chem Phys* (2011) **134**: 144112. doi:10.1063/1.3575402
43. Dennis E. Toward fault-tolerant quantum computation without concatenation. *Phys Rev A* (2001) **63**:052314. doi:10.1103/PhysRevA.63.052314
44. Deutsch D. Quantum theory, the Church–Turing principle and the universal quantum computer. *J Phys* (1985) **400**:97–117. doi:10.1098/rspa.1985.0070
45. Deutsch DP, Jozsa R. Rapid solution of problems by quantum computation. *Proc Roy Soc Lond A* (1992) **439**:553–8. doi:10.1098/rspa.1992.0167
46. Di Y-M, Wei H-R. Synthesis of multivalued quantum logic circuits by elementary gates. *Phys Rev A* (2013) **87**:012325. doi:10.1103/PhysRevA.87.012325
47. DiVincenzo DP. Two-bit gates are universal for quantum computation. *Phys Rev A* (1995) **51**:1015. doi:10.1103/physreva.51.1015
48. Dogra S, Arvind, Dorai K. Determining the parity of a permutation using an experimental nmr qutrit. *Phys Lett* (2014) **378**:3452–6. doi:10.1016/j.physleta.2014.10.003
49. Dua A, Malomed B, Cheng M, Jiang L. Universal quantum computing with parafermions assisted by a half-fluxon. *Phys Rev B* (2019) **100**:144508. doi:10.1103/PhysRevB.100.144508
50. Eastin B, Knill E. Restrictions on transversal encoded quantum gate sets. *Phys Rev Lett* (2009) **102**:110502. doi:10.1103/PhysRevLett.102.110502
51. Ecker S, Bouchard F, Bulla L, Brandt F, Kohout O, Steinlechner F, et al. Overcoming noise in entanglement distribution. *Phys Rev X* (2019) **9**:041042. doi:10.1103/physrevx.9.041042

52. Erhard M, Fickler R, Krenn M, Zeilinger A. Twisted photons: new quantum perspectives in high dimensions. *Light Sci Appl* (2018) **7**:17146. doi:10.1038/lsa.2017.146
53. Fan Y. A generalization of the Deutsch-Jozsa algorithm to multi-valued quantum logic. In: 37th international symposium on multiple-valued logic (ISMVL'07); 2007 May 13-16; Oslo, Norway (2007) ,p. 12. doi:10.1109/ISMVL.2007.3
54. Fan Y. Applications of multi-valued quantum algorithms (2008) arXiv: 0809.0932. doi:10.1109/ISMVL.2007.3
55. Farhi E, Goldstone J, Gutmann S, Sipser M. Quantum computation by adiabatic evolution (2000) arXiv:quant-ph/0001106.
56. Finnila AB, Gomez MA, Sebenik C, Stenson C, Doll JD. Quantum annealing: a new method for minimizing multidimensional functions. *Chem Phys Lett* (1994) **219**:343–8. doi:10.1016/0009-2614(94)00117-0
57. Freedman MH, Larsen M, Wang Z. A modular functor which is universalÂ¶for quantum computation. *Commun Math Phys* (2002) **227**:605–22. doi:10.1007/s002200200645
58. Fujii K. Exchange gate on the qudit space and fock space. *J Opt B Quantum Semiclassical Opt* (2003) **5**:S613–S618. doi:10.1088/1464-4266/5/6/011
59. Ganzhorn M, Klyatskaya S, Ruben M, Wernsdorfer W. Strong spin-phonon coupling between a single-molecule magnet and a carbon nanotube nanoelectromechanical system. *Nat Nanotechnol* (2013) **8**:165. doi:10.1038/nnano.2012.258
60. Gao X, Erhard M, Zeilinger A, Krenn M. (2020) Computer-inspired concept for high-dimensional multipartite quantum gates. *Phys Rev Lett 125, 050501*. doi:10.1103/PhysRevLett.125.050501
61. Garcia-Escartin JC, Chamorro-Posada P. A swap gate for qudits. *Quant Inf Process* (2013) **12**:3625–31. doi:10.1007/s11128-013-0621-x
62. Gedik Z, Silva IA, Lie akmak B, Karpat G, Vidoto ELG, Soares-Pinto DO, et al. Computational speed-up with a single qudit. *Sci Rep* (2015) **5**:14671. doi:10.1038/srep14671
63. Godfrin C, Ferhat A, Ballou R, Klyatskaya S, Ruben M, Wernsdorfer W, et al. Operating quantum states in single magnetic molecules: implementation of grover's quantum algorithm. *Phys Rev Lett* (2017) **119**:187702. doi:10.1103/physrevlett.119.187702
64. Gottesman D. Theory of fault-tolerant quantum computation. *Phys Rev* (1998) **57**:127–37. doi:10.1103/PhysRevA.57.127
65. Gottesman D. Fault-tolerant computation with higher-dimensional systems. In: Williams CP, editor. Lecture notes in computer science NASA international conference on quantum computing and quantum communications; **Vol. 1509**. Berlin: Springer (1999) ,p 302–13.
66. Gottesman D, Chuang IL. Demonstrating the viability of universal quantum computation using teleportation and single-qubit operations. *Nature* (1999) **402**:390–3. doi:10.1038/46503Google Scholar
67. Gottesman D, Kitaev A, Preskill J. Encoding a qubit in an oscillator. *Phys Rev* (2001) **64**:012310. doi:10.1103/physreva.64.012310
68. Gualdi G, Licht D, Reich DM, Koch CP. Efficient Monte Carlo characterization of quantum operations for qudits. *Phys Rev A* (2014) **90**:032317. doi:10.1103/physreva.90.032317
69. Harrow AW, Hassidim A, Lloyd S. Quantum algorithm for linear systems of equations. *Phys Rev Lett* (2009) **103**:150502. doi:10.1103/PhysRevLett.103.150502
70. Hein M, Eisert J, Briegel HJ. Multiparty entanglement in graph states. *Phys Rev A* (2004) **69**:062311. doi:10.1103/PhysRevA.69.062311
71. Howard M, Vala J. Qudit versions of the qubitπ/8gate. *Phys Rev A* (2012) **86**: 022316. doi:10.1103/PhysRevA.86.022316
72. Howard M, Wallman J, Veitch V, Emerson J. Contextuality supplies the 'magic' for quantum computation. *Nature* (2014) **510**:351. doi:10.1038/nature13460
73. Humphreys PC, Metcalf BJ, Spring JB, Moore M, Jin X-M, Barbieri M, et al. Linear optical quantum computing in a single spatial mode. *Phys Rev Lett* (2013) **111**:150501. doi:10.1103/PhysRevLett.111.150501
74. Hutter A, Loss D. Quantum computing with parafermions. *Phys Rev B* (2016) **93**:125105. doi:10.1103/PhysRevB.93.125105
75. Imany P, Jaramillo-Villegas JA, Alshaykh MS, Lukens JM, Odele OD, Moore AJ, et al. High-dimensional optical quantum logic in large operational spaces. *npj Quantum Inf.* (2019) **5**. doi:10.1038/s41534-019-0173-8
76. Imany P, Jaramillo-Villegas JA, Odele OD, Han K, Leaird DE, Lukens JM, et al. 50-ghz-spaced comb of high-dimensional frequency-bin entangled photons from an on-chip silicon nitride microresonator. *Opt Express* (2018) **26**: 1825–40. doi:10.1364/oe.26.001825
77. Ishikawa N, Sugita M, Okubo T, Tanaka N, Iino T, Kaizu Y. Determination of ligand-field parameters and f-electronic structures of double-decker bis(phthalocyaninato)lanthanide complexes. *Inorg Chem* (2003) **42**:2440–6. doi:10.1021/ic026295u
78. Islam NT, Lim CCW, Cahall C, Kim J, Gauthier DJ. Provably secure and high-rate quantum key distribution with time-bin qudits. *Sci Adv* (2017) **3**: e1701491. doi:10.1126/sciadv.1701491
79. Ivanov SS, Tonchev HS, Vitanov NV. Time-efficient implementation of quantum search with qudits. *Phys Rev A* (2012) **85**:062321. doi:10.1103/PhysRevA.85.062321
80. Jafarzadeh M, Wu Y-D, Sanders YR, Sanders BC. Randomized benchmarking for qudit Clifford gates. *New J Phys* (2020) **22**:063014. doi:10.1088/1367-2630/ab8ab1
81. Jelezko F, Gaebel T, Popa I, Domhan M, Gruber A, Wrachtrup J. Observation of coherent oscillation of a single nuclear spin and realization of a two-qubit conditional quantum gate. *Phys Rev Lett* (2004) **93**:130501. doi:10.1103/PhysRevLett.93.130501
82. Joo J, Lee C-W, Kono S, Kim J. Logical measurement-based quantum computation in circuit-qed. *Sci Rep* (2019) **9**:16592. doi:10.1038/s41598-019-52866-3
83. Kadowaki T, Nishimori H. Quantum annealing in the transverse ising model. *Phys Rev E* (1998) **58**:5355–63. doi:10.1103/PhysRevE.58.5355
84. Kais S. Introduction to quantum information and computation for chemistry. *Quantum Inf Comput Chem* (2014) **154**:1–38. doi:10.1002/9781118742631.ch01
85. Keet A, Fortescue B, Markham D, Sanders BC. Quantum secret sharing with qudit graph states. *Phys Rev A* (2010) **82**:062315. doi:10.1103/PhysRevA.82.062315
86. Kessel AR, Yakovleva NM. Implementation schemes in NMR of quantum processors and the Deutsch-Jozsa algorithm by using virtual spin representation. *Phys Rev A* (2002) **66**:062322. doi:10.1103/PhysRevA.66.062322
87. Khan FS, Perkowski M. Synthesis of multi-qudit hybrid and d-valued quantum logic circuits by decomposition. *Theor Comput Sci* (2006) **367**:336–46. doi:10.1016/j.tcs.2006.09.006
88. Kiktenko EO, Fedorov AK, Strakhov AA, Man'ko VI. Single qudit realization of the Deutsch algorithm using superconducting many-level quantum circuits. *Phys Lett A* (2015) **379**:1409–13. doi:10.1016/j.physleta.2015.03.023
89. Kiktenko EO, Nikolaeva AS, Xu P, Shlyapnikov GV, Fedorov AK. Scalable quantum computing with qudits on a graph. *Phys Rev A* (2020) **101**:022304. doi:10.1103/PhysRevA.101.022304
90. Kitaev AY. Fault-tolerant quantum computation by anyons. *Ann Phys* (2003) **303**:2–30. doi:10.1016/s0003-4916(02)00018-0
91. Klimov AB, Guzmán R, Retamal JC, Saavedra C. Qutrit quantum computer with trapped ions. *Phys Rev A* (2003) **67**:062313. doi:10.1103/PhysRevA.67.062313
92. Klyachko AA, Can MA, Binicioğlu S, Shumovsky AS. Simple test for hidden variables in spin-1 systems. *Phys Rev Lett* (2008) **101**:020403. doi:10.1103/PhysRevLett.101.020403
93. Kochen S, Specker EP. *The problem of hidden variables in quantum mechanics*. Dordrecht, Netherlands: Springer Netherlands (1975) p. 293–328. doi:10.1007/978-94-010-1795-4'17
94. Kononenko M, Yurtalan M, Shi J, Lupascu A. Characterization of control in a superconducting qutrit using randomized benchmarking (2020) arXiv preprint arXiv:2009.00599.
95. Krishna R, Makwana V, Suresh AP. A generalization of Bernstein-Vazirani algorithm to qudit systems (2016) arXiv preprint arXiv:1609.03185.
96. Kues M, Reimer C, Roztocki P, Cortés LR, Sciara S, Wetzel B, et al. On-chip generation of high-dimensional entangled quantum states and their coherent control. *Nature* (2017) **546**:622–6. doi:10.1038/4650310.1038/nature22986
97. Kyoseva ES, Vitanov NV. Coherent pulsed excitation of degenerate multistate systems: exact analytic solutions. *Phys Rev A* (2006) **73**:023420. doi:10.1103/PhysRevA.73.023420
98. Lee J-S, Khitrin AK. Projective measurement in nuclear magnetic resonance. *Appl Phys Lett* (2006) **89**:074105. doi:10.1063/1.2425191
99. Leuenberger MN, Loss D. Quantum computing in molecular magnets. *Nature* (2001) **410**:789. doi:10.1038/35071024

100. Li B, Yu Z-H, Fei S-M. Geometry of quantum computation with qutrits. *Sci Rep* (2013a) **3**:2594. doi:10.1038/srep02594
101. Li HY, Wu CW, Liu WT, Chen PX, Li CZ. Fast quantum search algorithm for databases of arbitrary size and its implementation in a cavity QED system. *Phys Lett* (2011) **375**:4249–54. doi:10.1016/j.physleta.2011.10.016
102. Li W-D, Gu Y-J, Liu K, Lee Y-H, Zhang Y-Z. Efficient universal quantum computation with auxiliary Hilbert space. *Phys Rev A* (2013b) **88**:034303. doi:10.1103/PhysRevA.88.034303
103. Liu Z, Fan H. Decay of multiqudit entanglement. *Phys Rev A* (2009) **79**: 064305. doi:10.1103/physreva.79.064305
104. Loss D, DiVincenzo DP. Quantum computation with quantum dots. *Phys Rev A* (1998) **57**:120–6. doi:10.1103/PhysRevA.57.120
105. Low RA. Learning and testing algorithms for the Clifford group. *Phys Rev A* (2009) **80**:052314. doi:10.1103/PhysRevA.80.052314
106. Lu HH, Hu Z, Alshaykh MS, Moore AJ, Wang Y, Imany P, et al. Quantum phase estimation with time-frequency qudits in a single photon. *Adv Quantum Tech* (2019) **3**:1900074. doi:10.1002/qute.201900074
107. Lu H-H, Lukens JM, Peters NA, Odele OD, Leaird DE, Weiner AM, et al. Electro-optic frequency beam splitters and tritters for high-fidelity photonic quantum inf. process. *Phys Rev Lett* (2018) **120**:030502. doi:10.1103/physrevlett.120.030502
108. Luo M, Wang X. Universal quantum computation with qudits. *Sci China Phys Mech Astron* (2014) **57**:1712–7. doi:10.1007/s11433-014-5551-9
109. Luo M-X, Chen X-B, Yang Y-X, Wang X. Geometry of quantum computation with qudits. *Sci Rep* (2014) **4**:4044. doi:10.1038/srep04044
110. Martinis JM, Geller MR. Fast adiabatic qubit gates using only$\sigma_z$control. *Phys Rev A* (2014) **90**:022307. doi:10.1103/PhysRevA.90.022307
111. Hugh DM, Twamley J. Trapped-ion qutrit spin molecule quantum computer. *New J Phys* (2005) **7**:174. doi:10.1088/1367-2630/7/1/174
112. Mermin ND. From classical state swapping to quantum teleportation. *Phys Rev A* (2001) **65**:012320. doi:10.1103/PhysRevA.65.012320
113. Milburn GJ. Photons as qubits. *Phys Scripta* (2009) **T137**:014003. doi:10.1088/0031-8949/2009/t137/014003
114. Mischuck B, Mølmer K. Qudit quantum computation in the Jaynes-Cummings model. *Phys Rev A* (2013) **87**:022341. doi:10.1103/PhysRevA.87.022341
115. Moreno-Pineda E, Godfrin C, Balestro F, Wernsdorfer W, Ruben M. Molecular spin qudits for quantum algorithms. *Chem Soc Rev* (2018) **47**: 501–13. doi:10.1039/C5CS00933B
116. Moreno Pineda E, Komeda T, Katoh K, Yamashita M, Ruben M. Surface confinement of TbPc2-SMMs: structural, electronic and magnetic properties. *Dalton Trans* (2016) **45**:18417–33. doi:10.1039/C6DT03298B
117. Morvan A, Ramasesh V, Blok M, Kreikebaum J, O'Brien K, Chen L, et al. Qutrit randomized benchmarking (2020) arXiv preprint arXiv:2008.09134.
118. Muthukrishnan A, Stroud CR. Multivalued logic gates for quantum computation. *Phys Rev A* (2000) **62**:052309. doi:10.1103/PhysRevA.62.052309
119. Nagata K, Geurdes H, Patro SK, Heidari S, Farouk A, Nakamura T. Generalization of the Bernstein-Vazirani algorithm beyond qubit systems. *Quantum Stud Math Found* (2020) **7**:17–21. doi:10.1007/s40509-019-00196-4
120. Neumann P, Beck J, Steiner M, Rempp F, Fedder H, Hemmer PR, et al. Single-shot readout of a single nuclear spin. *Science* (2010) **329**:542–4. doi:10.1126/science.1189075
121. Nguyen DM, Kim S. Quantum key distribution protocol based on modified generalization of Deutsch-Jozsa algorithm in d-level quantum system. *Int J Theor Phys* (2019) **58**:71–82. doi:10.1007/s10773-018-3910-4
122. Nielsen MA. A geometric approach to quantum circuit lower bounds (2005) arXiv preprint quant-ph/0502070.
123. Nielsen MA. Cluster-state quantum computation. *Rep Math Phys* (2006) **57**: 147–61. doi:10.1016/S0034-4877(06)80014-5
124. Nielsen MA, Bremner MJ, Dodd JL, Childs AM, Dawson CM. Universal simulation of hamiltonian dynamics for quantum systems with finite-dimensional state spaces. *Phys Rev A* (2002) **66**:022317. doi:10.1103/PhysRevA.66.022317
125. Nielsen MA, Chuang IL. *Quantum computation and quantum information*. 10th ed. New York, NY: Cambridge University Press (2011)
126. Nielsen MA, Dowling MR, Gu M, Doherty AC. Quantum computation as geometry. *Science* (2006) **311**:1133–5. doi:10.1126/science.1121541
127. Nowack KC, Koppens FHL, Nazarov YV, Vandersypen LMK. Coherent control of a single electron spin with electric fields. *Science* (2007) **318**: 1430–3. doi:10.1126/science.1148092
128. Pan J, Cao Y, Yao X, Li Z, Ju C, Chen H, et al. Experimental realization of quantum algorithm for solving linear systems of equations. *Phys Rev A* (2014) **89**:022313. doi:10.1103/physreva.89.022313
129. Parasa V, Perkowski M. Quantum phase estimation using multivalued logic. In: 2011 41st IEEE international symposium on multiple-valued logic Tuusula, Finland: Institute of Electrical and Electronics Engineers (IEEE); 2011. p. 224–9. doi:10.1109/ISMVL.2011.47
130. Patera J, Zassenhaus H. The Pauli matrices in n dimensions and finest gradings of simple Lie algebras of type An−1. *J Math Phys* (1988) **29**:665–73. doi:10.1063/1.528006
131. Paz-Silva GA, Rebić S, Twamley J, Duty T. Perfect mirror transport protocol with higher dimensional quantum chains. *Phys Rev Lett* (2009) **102**:020503. doi:10.1103/PhysRevLett.102.020503
132. Prevedel R, Walther P, Tiefenbacher F, Böhi P, Kaltenbaek R, Jennewein T, et al. High-speed linear optics quantum computing using active feed-forward. *Nature* (2007) **445**:65–9. doi:10.1038/nature05346
133. Ralph TC, Resch KJ, Gilchrist A. Efficient Toffoli gates using qudits. *Phys Rev A* (2007) **75**:022313. doi:10.1103/PhysRevA.75.022313
134. Raussendorf R, Briegel HJ. A one-way quantum computer. *Phys Rev Lett* (2001) **86**:5188–91. doi:10.1103/PhysRevLett.86.5188
135. Reck M, Zeilinger A, Bernstein HJ, Bertani P. Experimental realization of any discrete unitary operator. *Phys Rev Lett* (1994) **73**:58–61. doi:10.1103/PhysRevLett.73.58
136. Reich DM, Gualdi G, Koch CP. Optimal qudit operator bases for efficient characterization of quantum gates. *J Phys Math Theor* (2014) **47**:385305. doi:10.1088/1751-8113/47/38/385305
137. Rowe DJ, Sanders BC, de Guise H. Representations of the Weyl group and Wigner functions for SU(3). *J Math Phys* (1999) **40**:3604–15. doi:10.1063/1.532911
138. Sawerwain M, Leoński W. Quantum circuits based on qutrits as a tool for solving systems of linear equations (2013) arXiv:1309.0800.
139. Shende VV, Bullock SS, Markov IL. Synthesis of quantum-logic circuits. *IEEE Trans Comput Aided Des Integrated Circ Syst* (2006) **25**:1000–10. doi:10.1109/TCAD.2005.855930
140. Sheridan L, Scarani V. Security proof for quantum key distribution using qudit systems. *Phys Rev A* (2010) **82**:030301. doi:10.1103/physreva.82.030301
141. Shi Y. Both Toffoli and controlled-NOT need little help to do universal quantum computation (2002) arXiv preprint quant-ph/0205115.
142. Shor PW. Algorithms for quantum computation: discrete logarithms and factoring. In: Proceedings 35th annual symposium on foundations of computer science 1994 Nov 20–22; Santa Fe, NM (1994) p. 124–34. doi:10.1109/SFCS.1994.365700
143. Shor PW. Polynomial-time algorithms for prime factorization and discrete logarithms on a quantum computer. *SIAM Rev* (1999) **41**:303–32. doi:10.1137/s0036144598347011
144. Slichter CP. *Principles of magnetic resonance*. **vol. 1**. Berlin/Heidelberg, Germany: Springer Science & Business Media (2013)
145. Stroud AMCR. Quantum fast fourier transform using multilevel atoms. *J Mod Optic* (2002) **49**:2115–27. doi:10.1080/09500340210123947
146. Thiele S, Balestro F, Ballou R, Klyatskaya S, Ruben M, Wernsdorfer W. Electrically driven nuclear spin resonance in single-molecule magnets. *Science* (2014) **344**:1135–8. doi:10.1126/science.1249802
147. Tonchev HS, Vitanov NV. Quantum phase estimation and quantum counting with qudits. *Phys Rev A* (2016) **94**:042307. doi:10.1103/PhysRevA.94.042307
148. Troiani F, Affronte M. Molecular spins for quantum information technologies. *Chem Soc Rev* (2011) **40**:3119–29. doi:10.1039/c0cs00158a
149. Urdampilleta M, Klyatskaya S, Cleuziou J-P, Ruben M, Wernsdorfer W. Supramolecular spin valves. *Nat Mater* (2011) **10**:502. doi:10.1038/nmat3050
150. van Dam W, Howard M. Noise thresholds for higher-dimensional systems using the discrete Wigner function. *Phys Rev A* (2011) **83**:032310. doi:10.1103/PhysRevA.83.032310
151. Van den Nest M. Universal quantum computation with little entanglement. *Phys Rev Lett* (2013) **110**:060504. doi:10.1103/PhysRevLett.110.060504

152. Vandersypen LMK, Steffen M, Breyta G, Yannoni CS, Sherwood MH, Chuang IL. Experimental realization of Shor's quantum factoring algorithm using nuclear magnetic resonance. *Nature* (2001) **414**:883. doi:10.1038/414883a
153. Vlasov AY. Noncommutative tori and universal sets of nonbinary quantum gates. *J Math Phys* (2002) **43**:2959–64. doi:10.1063/1.1476391
154. Wang H, Kais S, Aspuru-Guzik A, Hoffmann MR. Quantum algorithm for obtaining the energy spectrum of molecular systems. *Phys Chem Chem Phys* (2008) **10**:5388–93. doi:10.1039/B804804E
155. Wang X. Continuous-variable and hybrid quantum gates. *J Phys Math Gen* (2001) **34**:9577–84. doi:10.1088/0305-4470/34/44/316
156. Watabe S, Seki Y, Kawabata S. Enhancing quantum annealing performance by a degenerate two-level system. *Sci Rep* (2020) **10**:146. doi:10.1038/s41598-019-56758-4
157. Webb Z. The Clifford group forms a unitary 3-design. *Quant Inf Comput* (2016) **16**:1379–400. doi:10.5555/3179439.3179447
158. Wilmott CM. On swapping the states of two qudits. *Int J Quant Inf* (2011) **09**: 1511–7. doi:10.1142/S0219749911008143
159. Wilmott CM, Wild PR. On a generalized quantum swap gate. *Int J Quant Inf* (2012) **10**:1250034. doi:10.1142/S0219749912500347
160. Zahedinejad E, Ghosh J, Sanders BC. High-fidelity single-shot Toffoli gate via quantum control. *Phys Rev Lett* (2015) **114**:200502. doi:10.1103/PhysRevLett.114.200502
161. Zahedinejad E, Ghosh J, Sanders BC. Designing high-fidelity single-shot three-qubit gates: a machine-learning approach. *Phys Rev Appl* (2016) **6**: 054005. doi:10.1103/PhysRevApplied.6.054005
162. Zeng B, Chung H, Cross AW, Chuang IL. Local unitary versus local Clifford equivalence of stabilizer and graph states. *Phys Rev A* (2007) **75**:032325. doi:10.1103/PhysRevA.75.032325
163. Zhan X, Li J, Qin H, Bian Z-h, Xue P. Linear optical demonstration of quantum speed-up with a single qudit. *Opt Express* (2015) **23**:18422–7. doi:10.1364/OE.23.018422
164. Zhou DL, Zeng B, Xu Z, Sun CP. Quantum computation based ond-level cluster state. *Phys Rev A* (2003) **68**:062303. doi:10.1103/PhysRevA.68.062303
165. Zilic Z, Radecka K. Scaling and better approximating quantum Fourier transform by higher radices. *IEEE Trans Comput* (2007) **56**:202–7. doi:10.1109/TC.2007.35
166. Zobov VE, Ermilov AS. Implementation of a quantum adiabatic algorithm for factorization on two qudits. *J Exp Theor Phys* (2012) **114**:923–32. doi:10.1134/S106377611205007X

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

ORIGINAL RESEARCH
published: 10 November 2020
doi: 10.3389/fphy.2020.547370

A Quantum Finite Automata Approach to Modeling the Chemical Reactions

Amandeep Singh Bhatia[1] *and Shenggen Zheng*[2]*

[1]Chitkara University Institute of Engineering & Technology, Chitkara University, Punjab, India [2]Peng Cheng Laboratory, Shenzhen, China

In recent years, the modeling interest has increased significantly from molecular level to atomic and quantum levels. Computational chemistry plays a significant role in designing computational models for the operation and simulation of systems ranging from atoms and molecules to industrial processes. It is influenced by a tremendous increase in computing power and the efficiency of algorithms. The representation of chemical reactions using classical automata theory in thermodynamic terms had a great influence on computer science. The study of chemical information processing with quantum computational models is a natural goal. In this study, we have modeled chemical reactions using two-way quantum finite automata, which are halted in linear time. Additionally, classical pushdown automata can be designed for such chemical reactions with multiple stacks. It has been proven that computational versatility can be increased by combining chemical accept/reject signatures and quantum automata models.

Keywords: chemical reaction, two-way quantum finite automata, quantum finite automata, Belousov-Zhabotinsky reaction, pushdown automata, quantum chemistry

Edited by:
Jiangfeng Du,
University of Science and Technology of China, China

Reviewed by:
Son Tung Ngo,
Ton Duc Thang University, Vietnam
Xin Wang,
City University of Hong Kong, Hong Kong

***Correspondence:**
Shenggen Zheng
zhengshg@pcl.ac.cn

Specialty section:
This article was submitted to Physical Chemistry and Chemical Physics, a section of the journal Frontiers in Physics

Received: 31 March 2020
Accepted: 27 August 2020
Published: 10 November 2020

Citation:
Bhatia AS and Zheng S (2020) A Quantum Finite Automata Approach to Modeling the Chemical Reactions. Front. Phys. 8:547370. doi: 10.3389/fphy.2020.547370

1 INTRODUCTION

Recently, the connection between complex reactions and their thermodynamics has received overwhelming response among research communities. Initially, in the 1970s, Conrad [1] processed the information of molecular systems and stated that complex biochemical systems cannot be analyzed in classical computers. Till now, artificial approaches use complex biomolecules or logic gates–based reaction-diffusion systems to solve the problems [2–4]. Classical systems are not robust and incapable to describe quantum systems. Some tasks that are impossible in classical systems can be realized in quantum systems. Quantum computation is concerned with computer technology based on the principles of quantum mechanics, which describes the behavior and nature of matter and energy in quantum level [5]. Quantum computation demonstrates the computation power and other properties of the computers based on the principles of quantum mechanics.

Models of finite automata are abstract computing devices, which play a crucial role to solve computational problems in theoretical computer science. Classical automata theory is closely associated with formal language theory, where automata are ranked from simplest to most powerful depending on their language recognition power [6]. Classical automata theory has been of significant importance due to its practical real-time applications in the development of several fields. Therefore, it is the natural goal to study quantum variants of classical automata models, which play an important role in quantum information processing.

The quantum automata theory has been developed using the principles of quantum mechanics and classical automata. Quantum computational models make it possible to examine the resources needed for computations. Soon after the brainstorm of Shor's factorization quantum algorithm [7], the first models of quantum finite automata (QFAs) have been introduced. Initially, Kondacs and Watrous [8], and Moore and Crutchfield [9] proposed the concept of quantum automata separately. Since then, a variety of quantum automata models have been studied and demonstrated in various directions, such as QFAs, Latvian QFA, 1.5-way QFA, two-way QFA (2QFA), quantum sequential machine, quantum pushdown automata, quantum Turing machine, quantum multicounter machines, quantum queue automata [10], quantum multihead finite automata, QFAs with classical states (2QCFA) [11, 12], state succinctness of two-way probabilistic finite automata (2PFA), QFA, 2QFA, and 2QCFA [13–15], interactive proof systems with QFAs [16, 17], quantum finite state machines of matrix product state [18], promise problems recognition by QFA [19–22], quantum-omega automata [23] and semi-quantum two-way finite automata [24–26], time complexity advantages of QFA [27], nonuniform classes of polynomial size QFA [28, 29], QFA and linear temporal logic relationship [30], and many more since the past 2 decades [31–34]. These models are effective in determining the boundaries of various computational features and expressive power [35–37]. Quantum computers are more powerful than Turing machines and even probabilistic Turing machines. Thus, mathematical models of quantum computation can be viewed as generalizations of its physical models.

Computational biochemistry has been a rapidly evolving research area at the interface between biology, chemistry, computer science, and mathematics. It helps us to apply computational models to understand biochemical and chemical processes and their properties. A combination of chemistry and classical automata theory provides a constructive means of refining the number of objects allowed to understand the energetic cost of computation [38]. The research has been consistently grown in the field of chemical computing. There exist two ways to model complex chemical reactions: abstract devices and formal models based on multiset rewriting [39]. Complex chemical reaction networks carry out chemical processes that mimic the workings of classical automata models. Recently, Duenas-Diez and Perez-Mercader [38, 40] have designed chemical finite automata for regular languages and chemical automata with multiple stacks for context-free and context-sensitive languages. Furthermore, the thermodynamic interpretation of the acceptance/rejection of chemical automata is given. It is useful to understand the energetic cost of chemical computation. They have used the one-pot reactor (mixed container), where chemical reactions and molecular recognition takes place after several steps, without utilizing any auxiliary geometrical aid.

In classical automata theory, it is known that two-way deterministic finite automata (2DFA) can be designed for all regular languages. It has also been investigated that 2PFA can be designed for a nonregular language $L = \{a^n b^n | n \geq 1\}$ in an exponential time [38, 40]. The research has consistently evolved in the field of quantum computation and information processing. In quantum automata theory, it has been proved that 2QFA can be designed for L with one-sided bounded error and halted in linear time. Moreover, it has been demonstrated that 2QFA can be also designed for non–context-free language $L = \{a^n b^n c^n | n \geq 1\}$ [8]. Hence, 2QFA is strictly more powerful than its classical counterparts based on language recognition capability.

The field of chemistry and chemical computation plays a significant role in the evolution of computational models to mimic the behavior of systems at its atomic level. It is greatly influenced by the computing power of quantum computers. Motivated from the abovementioned facts, we have modeled chemical reactions in the form of formal languages and represented those using two-way QFAs. The main objective is to examine how chemical reactions perform chemical sequence identification equivalent to quantum automata models without involving biochemistry or any auxiliary device. The crucial advantage of this approach is that chemical reactions in the form of accept/reject signatures can be processed in linear time with one-sided bounded error (if the automata makes error only in one direction, i.e., either on "no" instances or on "yes" instances). This article is further designed as follows: Subsection is devoted to prior work. In **Section 2**, some preliminaries are given. The definition of two-way QFAs is given in **Section 3**. In **Section 4**, the chemical reactions are transcribed in formal languages and modeled using two-way QFAs approach. Summary of work is given in **Section 5**. Finally, **Section 6** is the conclusion.

1.1 Prior Work

The field of chemical computation has rich and interesting history. Various researchers have represented chemical computation using the concept of logic gates–based reaction-diffusion systems and artificial intelligence approaches. In early 1970s, Conrad [1] differentiated the information processing in molecules using digital computation. Nearly a decade later, Okamoto et al. [43] proposed the concept of a theoretical chemical diode in cyclic enzyme systems. It has been proved that it can be used to analyze the dynamic behavior of metabolic switching events in biocomputer. In 1991, Hjelmfelt et al. [44] designed neural networks and finite state machines using chemical diodes. It has been found that the execution of a universal Turing machine is possible using connecting chemical diodes. Hjelmfelt et al. constructed clocked finite state machines of binary adder, binary decoder, and stack memory and showed that finite state machines can be simulated by clocked neural networks.

In 1995, Tóth and Showalter [45] implemented AND and OR logic gates using reaction-diffusion systems, where the signals are programmed by chemical waves. It was the first empirical realization of chemical logical gates. In 1997, Magnasco [46] showed that logic gates can be constructed and executed in the chemical kinetics of homogeneous solutions. It has been proved that such constructions have computational power equivalent to Turing machine. Adamatzky and Lacy Costello [47] experimentally understood the Chemical XOR gate by following the same approach of Toth and Showlter in 2002. Further, Górecki et al. [48] constructed the chemical counters for information processing in the excitable reaction-diffusion systems.

It is one of the most promising new areas of research. Some difficulties can be caused by connecting several gates together for advanced computation. Thus, recently, researchers started focusing on native chemical computation, that is, without reaction-diffusion systems. In 1994, Adleman [49] proposed the concept of DNA computation and solved the Hamiltonian path problem by changing DNA strands. In 2009, Benenson [2] reviewed biological measurement tools for new-generation biocomputers. Prohaska et al. [3] studied protein domain using chromatin computation and introduced chromatin as a powerful machine for chemical computation and information processing. In 2012, Bryant [4] proved chromatin computer as computationally universal by using it to solve an example of combinatorial problem.

The structures of DNA and RNA are represented using the concept of classical automata theory [50, 51]. Krasinski et al. [52] represented the restricted enzyme in DNA with pushdown automata in circular mode. Khrennikov and Yurova [53] modeled the behavior of protein structures using classical automata theory and investigated the resemblance between the quantum systems and modeling behavior of proteins. Bhatia and Kumar [54] modeled ribonucleic acid (RNA) secondary structures using two-way QFAs, which are halted in linear time. Duenas-Diez and Perez-Mercader designed molecular machines for chemical reactions. The native chemical computation has been implemented beyond the scope of logic gates, that is, with chemical automata [40]. It has been demonstrated that chemical reactions transcribed in formal languages can be recognized by Turing machine without using biochemistry [38]. Recently, Bhatia and Zheng [55] modeled hairpin loop, pseudoknot, and dumbbell RNA secondary structures using 2QCFA.

2 PRELIMINARIES

In this section, some preliminaries are given. We assume that the reader is familiar with the classical automata theory and the concept of quantum computation; otherwise, reader can refer to the theory of automata [6], quantum information, and computation [5, 56]. Linear algebra is inherited from quantum mechanics to describe the field of quantum computation. It is a crucial mathematical tool and allows us to represent the quantum operations and quantum states by matrices and vectors, respectively, that obey the rules of linear algebra. The following are the notions of linear algebra used in quantum computational theory:

- Vector space (V) [56]: A vector space (V) is defined over the field F of complex numbers $\mathbb{C}$ consisting of a nonempty set of vectors, satisfying the following operations:
- Addition: If two vectors $|a\rangle$ and $|b\rangle$ belong to V, then $|a\rangle + |b\rangle \in V$.
- Multiplication by a scalar: If $|a\rangle$ belongs to V, then $\lambda|a\rangle \in V$, where $\lambda \in \mathbb{C}$.
- Dirac notation [5]: In quantum mechanics, the Dirac notation is one of the most peculiarities of linear algebra. The combination of vertical and angle bars ($|\rangle$ $\langle|$) is used to unfold quantum states. It provides an inner product of any two vectors. The bra $\langle b|$ and ket $|a\rangle$ represent the row vector and column vector, respectively.

$$|a\rangle = \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{bmatrix}, \langle b| = \begin{bmatrix} \beta_1^* & \beta_2^* & \beta_3^* \end{bmatrix}, |a\rangle\langle b| = \begin{bmatrix} \alpha_1\beta_1^* & \alpha_1\beta_2^* & \alpha_1\beta_3^* \\ \alpha_2\beta_1^* & \alpha_2\beta_2^* & \alpha_2\beta_3^* \\ \alpha_3\beta_1^* & \alpha_3\beta_2^* & \alpha_3\beta_3^* \end{bmatrix} \quad [1]$$

where β_i^* indicates the complex conjugate of complex number α_i.

- Quantum bit [34]: A quantum bit (qubit) is a unit vector defined over complex vector space $\mathbb{C}^2$. In general, it is represented as a superposition of two basis states labeled $|0\rangle$ and $|1\rangle$.

$$|\phi\rangle = \alpha|0\rangle + \beta|1\rangle \quad [2]$$

- The probability of state occurrence $|0\rangle$ is $|\alpha|^2$ and $|1\rangle$ is $|\beta|^2$. It satisfies that $|\alpha|^2 + |\beta|^2 = 1$. The two complex amplitudes (α and β) are represented by one qubit. Thus, 2^n complex amplitudes can be represented by n qubits.
- Quantum state [5]: A quantum state $|\psi\rangle$ is defined as a superposition of classical states

$$|\psi\rangle = \alpha_1|w_1\rangle + \alpha_2|w_2\rangle + \ldots + \alpha_n|w_n\rangle \quad [3]$$

where α_i s are complex amplitudes and $|w_i\rangle$s are classical states for $1 \leq i \leq n$. Therefore, a quantum state $|\psi\rangle$ can be represented as n-dimensional column vector.

$$\begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \ldots \\ \alpha_n \end{bmatrix} \quad [4]$$

- Unitary transformation: In quantum mechanics, the transformation between the quantum systems must be unitary. Consider a state $|\psi\rangle$ of quantum system at time t: $|\psi\rangle = \alpha_1|w_1\rangle + \alpha_2|w_2\rangle + \ldots + \alpha_n|w_n\rangle$ transformed into state $|\psi'\rangle$ at time t': $|\psi'\rangle = \alpha_1'|w_1\rangle + \alpha_2'|w_2\rangle + \ldots + \alpha_n'|w_n\rangle$, where complex amplitudes are associated by $|\psi'(t')\rangle = U(t'-t)|\psi(t)\rangle$, where U denotes a time reliant unitary operator, which satisfies that $(U(t'-t))^* U(t'-t) = 1$ and $\sum_{i=1}^{n}|\alpha_i|^2 = |\alpha_i'|^2 = 1$ [5].
- Hilbert space: A physical system is described by a complex vector space called Hilbert space $\mathcal{H}$ [56]. It allows us to describe the basis of the quantum system. The direct sum $|x|y\rangle : \mathcal{H}\oplus\mathcal{H} \rightarrow \mathbb{C}$ or inner product $|x|v\rangle : \mathcal{H}\otimes\mathcal{H} \rightarrow \mathbb{C}$ of two subspaces satisfies the following properties for any vectors:
- Linearity: $(\alpha\langle x| + \beta\langle y|)|z\rangle = \alpha|x|z\rangle + \beta|y|z\rangle$.
- Symmetric property: $|x|y\rangle = |y|x\rangle$.
- Positivity: $|x|x\rangle \geq 0$ and $|x|x\rangle = 0$ iff $x = 0$, where $x \in \mathcal{H}$.
- where $x, y, z \in \mathcal{H}$ and $\alpha, \beta \in \mathbb{C}$.

- Quantum finite automaton (QFA) [57]: It is defined as a quintuple $(Q, \Sigma, s_{int}, P_{acc}, U_{\sigma})$, where
- Q is a set of states,
- Σ is an input alphabet,
- Hilbert space $\mathcal{H}$ and $s_{init} \in \mathcal{H}$ is an initial vector such that $|s_{init}|^2 = 1$,
- $\mathcal{H}_{acc} \subset \mathcal{H}$ and P_{acc} is an acceptance projection operator on $\mathcal{H}_{acc}$,
- U_{σ} denotes a unitary transition matrix for each input symbol $(\sigma \in \Sigma)$.

The computation procedure of QFA consists of an input string $w = \sigma_n \sigma_2 \ldots \sigma_n$. The automaton works by reading each input symbol, and their respective unitary matrices are applied on the current state, starting with an initial state. The quantum language accepted by QFA is represented as a function $f_{QFA}(w) = |s_{init} U_w P_{acc}|^2$, where $U_w = U_{\sigma_1} U_{\sigma_2} \ldots U_{\sigma_n}$. The tape head is allowed to move only in the right direction. Finally, the probability of QFA in an acceptance state is observed: that is, indicating whether the input string is accepted or rejected by QFA. It is also called a real-time quantum finite automaton.

Based on the movement of tape head, QFA is classified as one-way QFA, 1.5-way QFA, and 2QFA. In 1.5-way QFA, the tape head is permitted to move only in the right direction or can be stationary, but it cannot move toward the left direction. It has been proved that it can be designed for non–context-free languages, if the input tape is circular [58]. In this study, we focused on the 2QFA model due to the high computational power than its classical counterparts.

3 TWO-WAY QUANTUM FINITE AUTOMATA

A quantum finite automaton (QFA) is a quantum variant of a classical finite automaton. In QFA, quantum transitions are applied by reading the input symbols from the tape [9]. Two-way quantum finite automaton (2QFA) is a quantum counterpart of a two-way deterministic finite automaton (2DFA). In 2QFA, the tape head is allowed to move either in the left direction or right direction or can be stationary. The illustration of 2DFA is shown in **Figure 1**.

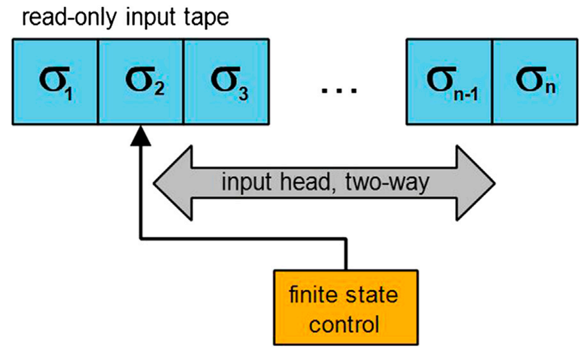


FIGURE 1 | Representation of two-way deterministic finite automata.

1. [9] A two-way quantum finite automaton is represented as sextuple $(Q, \Sigma, \delta, q_0, Q_{acc}, Q_{rej})$, where

- *Q is a finite set of states.*
- *Σ is an input alphabet.*
- *Transition function δ is defined by $\delta : Q \times \Gamma \times Q \times D \to \mathbb{C}$, where $\mathbb{C}$ is a complex number, $\Gamma = \Sigma \cup \{\#, $ and $D = \{-1, 0, +1\}$ represent the left, stationary and right direction of tape head.*
- $Q = Q_{acc} \cup Q_{rej} \cup Q_{non}$, where Q_{non}, Q_{acc}, and Q_{rej} represent the set of nonhalting, accepting, and rejecting states, respectively. The transition function must satisfy the following conditions:
- (i) Local probability and orthogonality condition:

$$\sum_{(q',d) \in Q\times D}^{\forall (q_1,\sigma_1),(q_2,\sigma_2) \in Q\times\Gamma} \overline{\delta(q_1, \sigma, q', d)}\delta(q_2, \sigma, q', d) = \begin{Bmatrix} 1 & q_1 = q_2 \\ 0 & q_1 \neq q_2 \end{Bmatrix}$$

(ii) First separability condition:

$$\sum_{q' \in Q}^{\forall (q_1,\sigma_1),(q_2,\sigma_2) \in Q\times\Gamma} \overline{\delta(q_1, \sigma_1, q', +1)}\delta(q_2, \sigma_2, q', 0)$$
$$+ \overline{\delta(q_1, \sigma_1, q', 0)}\delta(q_2, \sigma_2, q', -1)$$
$$= 0$$

(iii) Second separability condition:

$$\sum_{q' \in Q}^{\forall (q_1,\sigma_1),(q_2,\sigma_2) \in Q\times\Gamma} \overline{\delta(q_1, \sigma_1, q', +1)}\delta(q_2, \sigma_2, q', -1) = 0$$

For each $\sigma \in \Gamma$, a 2QFA is said to be simplified, if there exists a unitary linear operator V_{σ} on the inner product space such that $L_2\{Q\} \to L_2\{Q\}$. The transition function is represented as

$$\delta(q, \sigma, q', d) = \left\{ \begin{array}{c|c} q' V_{\sigma} q & \text{if } D(q') = d \\ 0 & \text{else} \end{array} \right\}, \qquad [5]$$

where $q' V_{\sigma} q$ is a coefficient of $|q'\rangle$ in $V_{\sigma}|q\rangle$.

Consider an input string w, written on the input tape enclosed with both end markers such as $\#w$. The computation of 2QFA is as follows. The tape head is above the input symbol σ, and the automaton is in any state q. Then, the state of 2QFA is changed to q' with an amplitude $\delta(q, \sigma, q', d)$ and moves the tape head one cell toward right, stationary, and left direction according to $\in \{-1, 0, +1\}$. It corresponds to a unitary evolution in the innerproduct space $\mathscr{H}_n$.

A computation of a 2QFA is a chain of superpositions $c_0, c_1, c_2, \ldots,$ where c_0 denotes an initial configuration. For any c_i, when the automaton is observed in a superposition state with an amplitude α_c, it has the form $U_{\delta}|c_i\rangle \sum_{c \in C_n} \alpha_c |c_i\rangle$,

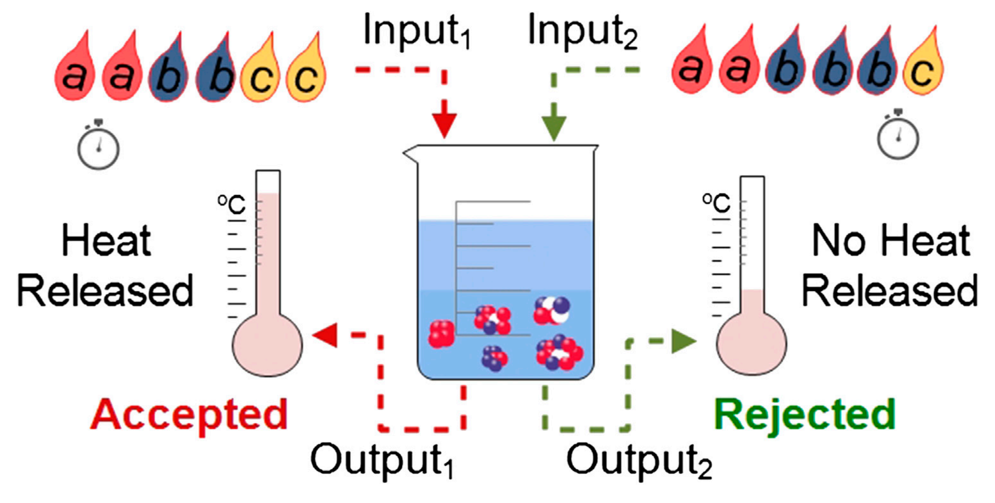


FIGURE 2 | Representation of language recognition by chemical computation. It is reproduced from [40] under the Creative Commons CCBY license.

where C_n represents the set of configurations. The probability associated with a configuration is calculated by absolute squares of amplitude. Superposition is said to be valid if the sum of the squared moduli of their probability amplitudes is unitary. In quantum theory, the time evolution is specified by unitary transformations. Each transition function δ prompts a transformation operator over the Hilbert space $\mathcal{H}_n$ in linear time.

$$U_\delta^w \left| q, j \right\rangle = \sum_{(q',d) \in Q \times D} \delta(q, w(j), q', d) \left| q', j + d \bmod |w| \right\rangle$$

for each $(q, j) \in C_{|w|}$, where $q \in Q, j \in Z_{|w|}$ and extended to $\mathcal{H}_n$ by linearity [9, 59].

4 MODELING OF CHEMICAL REACTIONS

Before we recognize the chemical reactions using two-way QFAs model, it is important to show how computational chemistry works. **Figure 2** shows the illustration of language recognition by the chemical computation model. It consists of three parts: i) a mixed container where the computation process occurs, ii) an input translator that translates the chemical aliquots into input symbols and gives those consecutively depending upon the processing time, iii) a system to monitor the response of an automaton as a chemical criterion. Finally, the chemical computation produces well-defined chemical accept/reject signatures for the input. For instance, if the number of *a*s and *b*s are equal in the input, then the chemical computation produces heat, that is, an input is said to be accepted. Otherwise, if no heat is released at the end of computation, then the input is said to be rejected by the system. The following are the construction of two-way quantum finite state machines of chemical reactions.

THEOREM 1. Two-way QFAs can recognize all regular languages.

PROOF. The proof has been shown in Ref. 9.

4.1 Chemical Reaction-1 Consisting Regular Language

For an illustrative and visual implementation, we can choose a precipitation reaction in an aqueous medium such as

$$KIO_3 + AgNO_3 \rightarrow AgIO_3\,(s) + KNO_3 \qquad [6]$$

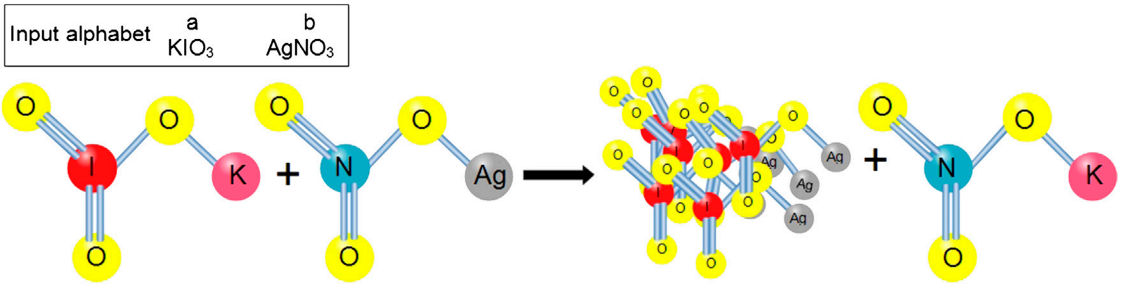


FIGURE 3 | Illustration of acid/base reaction of L_1.

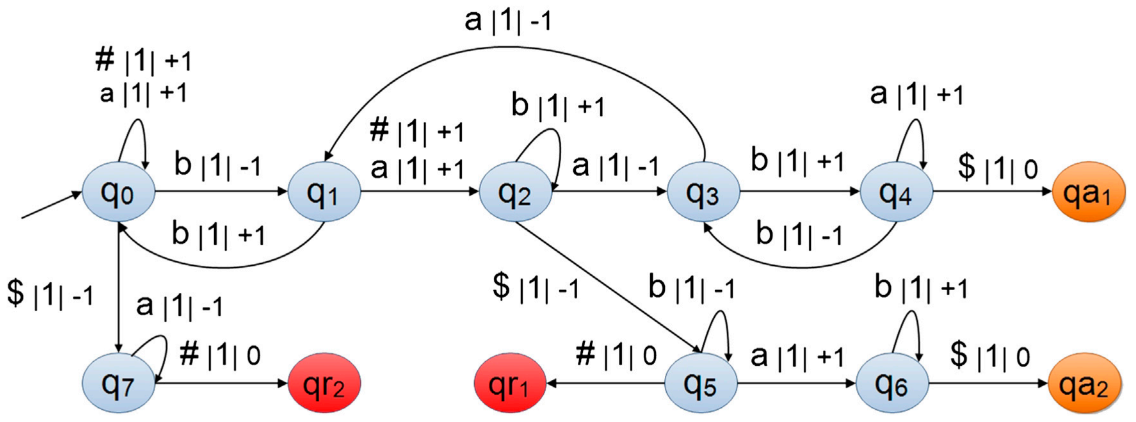


FIGURE 4 | State transition diagram of L_1.

During computation, if a white precipitate of silver iodate is observed, then the input string is said to be accepted; if the solution is clear from precipitate, the string has been rejected because there was no reaction. Therefore, we have chosen the recipes of alphabet symbols a for potassium iodate (KIO_3) and b for silver nitrate ($AgNO_3$) quantitatively. **Figure 3** shows the chemical representation of symbols a and b, the bimolecular precipitation reaction [38]. If the precipitate $\mathrm{AgIO_3}$ is not presented in the solution, then the computation is said to be rejected. For example, the input string $w = aaab$ is said to be accepted due to the presence of precipitate or, equally, heat has been determined during computation. But, the input $w = aa$ is said to be rejected due to the absence of precipitate or, precisely, heat has not been observed. The Kleene star (Σ^*) operator is a set of infinite strings of all lengths over input alphabet as well as empty string (ϵ). The language "$(a + b)^*$" means the string containing any number of "a"s or "b"s in any order or empty string. The language "$(ab)^*$" means the string containing any number of "ab"s or string of length zero. **Figure 4** shows the corresponding theoretical 2QFA state transition graph to recognize L_1.

THEOREM 2. A language $L_1 = \{(a + b)^* a (a + b)^* b (a + b)^* aa^* bb^*\}$ representing precipitation reaction in ***Eq. 6*** can be recognized by 2QFA.

PROOF. The idea of this proof is as follows. The initial state q_0 reads a right-end marker # and moves the head toward the right direction. If there is no occurrence of symbol b, then it shows no precipitate, and the input is said to be rejected by the 2QFA. Similarly, on reading the symbol b, the state q_0 is changed into q_1. If there is no occurrence of symbol a, then the state is transformed into rejecting state q_{r_1}. If the input string $w \in L_1$ contains at least one a and one b, then silver iodate is present during computation, and it is said to be recognized by 2QFA. A 2QFA for L_1 is defined as follows: $M_{2QFA} = (Q, \Sigma, q_0, Q_{acc}, Q_{rej}, \delta)$,

where $Q = \{q_0, q_1, q_2, q_3, q_4, q_5, q_6, q_7, q_{a_1}, q_{a_2}, q_{r_1}, q_{r_2}\}$, where q_0 and q_2 are used to move the head toward the $ on reading as and bs, respectively. The states q_1 and q_3 are used to confirm that the last symbol read by head is a and b, respectively.

$\Sigma = \{a, b\}, q_0$ is an initial state, $Q_{acc} = \{q_{a_1}, q_{a_2}\}$ and $Q_{rej} = \{q_{r_1}, q_{r_2}\}$.

The specification of transition functions are given in **Table 1**.

It can be noted that in 2QFA, transition matrices consist of 0 and 1, i.e., basically a two-way reversible finite automata (2RFA). Therefore, 2QFA can be designed for all the languages accepted by 2RFA. In transition matrix, each column and row have exactly only one entry 1. Hence, the dot product of any two rows is equal to zero. It is known that the language recognition power of 2RFA is an equivalent to 2DFA.

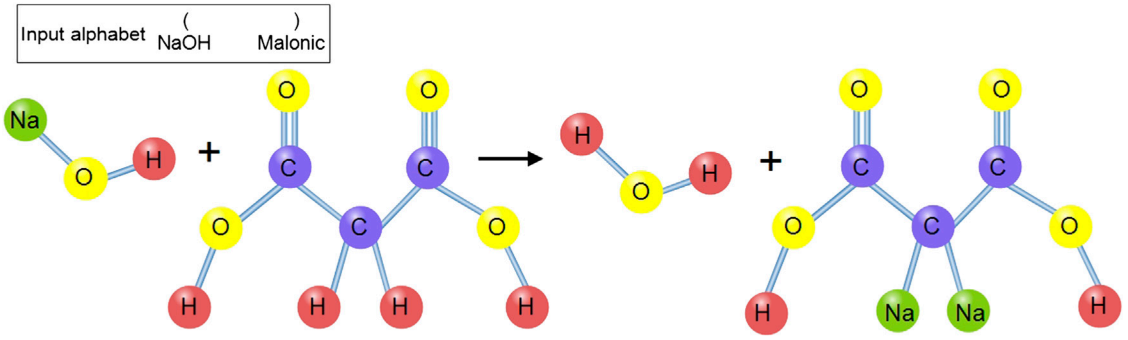


FIGURE 5 | Illustration of the acid/base reaction of L_2.

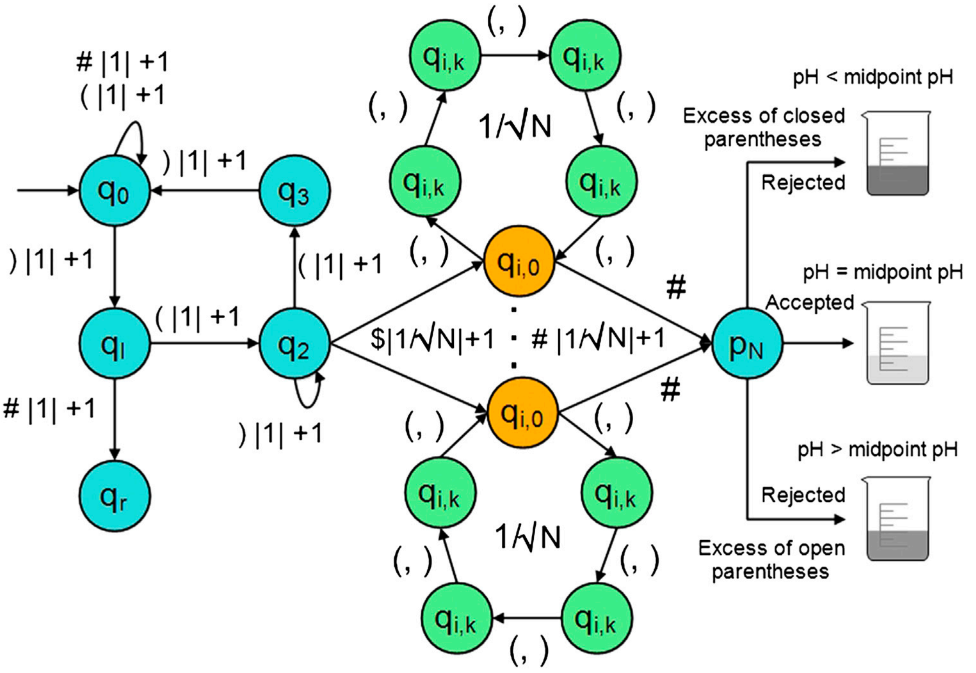


FIGURE 6 | State transition diagram of L_2.

4.2 Chemical Reaction-2 Consisting Context-Free Language

Next, we have considered the context-free language from Chomsky hierarchy satisfying the balanced chemical reaction between NaOH and malonic acid as follows:

$$H_2C_3H_2O_4 + 2NaOH \rightarrow Na_2C_3H_2O_4 + 2H_2O \qquad [7]$$

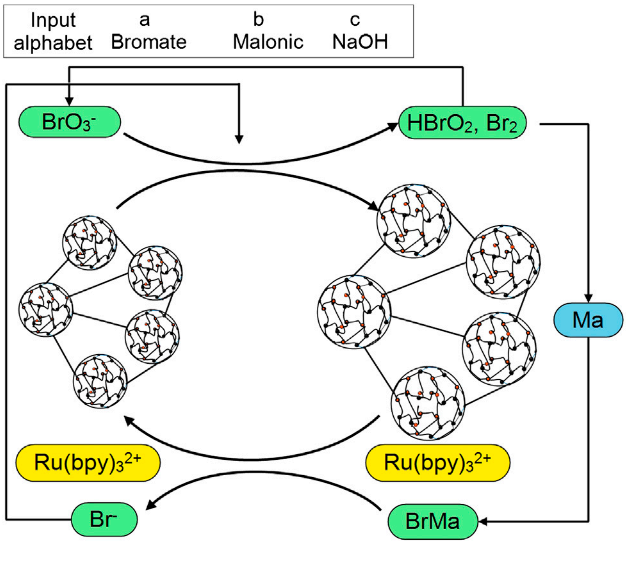


FIGURE 7 | Illustration of the acid/base reaction of L_3.

The language generated by the abovementioned chemical reaction is L_2 consisting Dyck language of all words with balanced parentheses. Figure 5 shows the acid/base reaction of L2. 2QFA is designed for L_2 as follows:

THEOREM 3. A language L_2 consisting of Dyck language of all words with balanced parentheses can be recognized by 2QFA with probability 1, otherwise rejected with probability at least $1 - \frac{1}{N}$, where N is any positive number.

PROOF. The idea of this proof is as follows. It consists of three phases. First, the initial state q_0 reads a first symbol and both heads start moving toward the right-end marker $. If the input string starts with closed parentheses, then it is said to be rejected. On reading the left-end marker #, the computation is split into N paths, denoted by $q_{1,0}, q_{2,0}, \ldots, q_{N,0}$. Each path possesses equal amplitude $\frac{1}{\sqrt{N}}$. Along the N different paths, each path moves deterministically to the right-end marker . Each computational path keeps track of the open parentheses with respect to the closed parentheses. At the end of computation, if the excess of open parentheses is observed, then it is said to be rejected. It means pH value is greater than midpoint pH, and intermediate gray tone is observed. Secondly, if there is an excess of closed parentheses, then the darkest gray tone is observed, that is, pH value is less than midpoint pH. It is said to be rejected by 2QFA with probability $1 - \frac{1}{N}$. If there is a balanced occurrence of open and closed parentheses, the input string is said to be accepted with probability 1. Hence, pH value is equal to midpoint pH, and the lightest gray tone is observed at the end of computation. A 2QFA for L_2 is defined as follows: $M_{2QFA} = (Q, \Sigma, q_0, Q_{acc}, Q_{rej}, \delta)$, where $Q = \{q_0, q_1, q_2, q_3\} \cup \{q_{i,j} \big| 1 \leq i \leq N, 0 \leq j \leq \max(i, N - i + 1)\} \cup \{p_k \big| 1 \leq k \leq N\} \cup \{s_{i,0}, w_{i,0}, r_{i,0} \big| 1 \leq i \leq N\}$

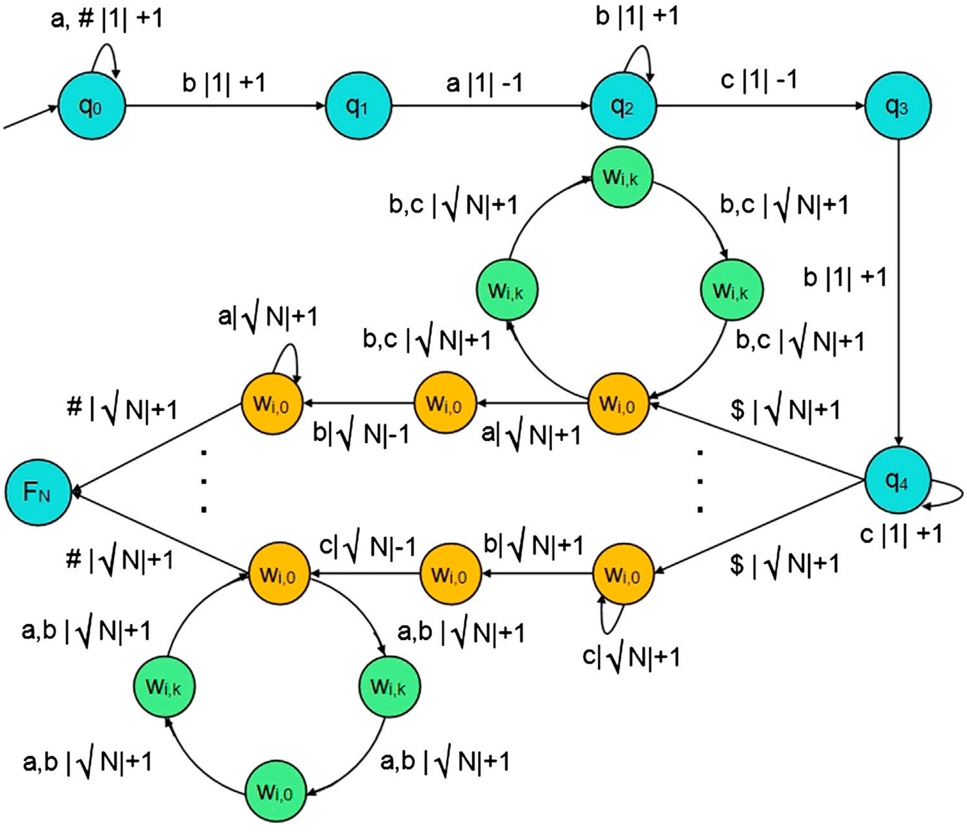


FIGURE 8 | State transition diagram of L_3.

$\cup \{q_{acc}, q_{rej}\}$, where q_1 is used to check whether the first symbol is an open parentheses or not, and q_2 and q_3 are used to traverse the input string. Figure 6 shows the state transition diagram for L2.

$\Sigma = \{(,)\}, q_0$ is an initial state, $Q_{acc} = \{p_N\}$ and $Q_{rej} = \{q_r\} \cup \{p_k | 1 \le k < N\} \cup \{r_{i,0} | 1 \le i \le N\}$.

The specification of transition functions is given in **Table 2**.

4.3 Chemical Reaction-3 Consisting Context-Sensitive Language

To implement a chemical 2QFA for context-sensitive language, we have used Belousov-Zhabotinsky (BZ) reaction network for the nonlinear oscillatory chemistry [38], which consists of temporal oscillation in the sodium bromate and malonic acid system [60] as Figure 7 shows the acid/base reaction of L3. Figure 8 shows the state transition diagram for L3.

TABLE 1 | Details of the transition functions and head function for L_1.

$V_\#\|q_0\rangle = \|q_0\rangle$	$V_a\|q_0\rangle = \|q_0\rangle$	$V_b\|q_0\rangle = \|q_1\rangle$	$V_\#\|q_1\rangle = \|q_2\rangle$
V	$V_a\|q_1\rangle = \|q_2\rangle$	$V_b\|q_1\rangle = \|q_2\rangle$	$V_b\|q_2\rangle = \|q_2\rangle$
V	$V_a\|q_7\rangle = \|q_7\rangle$	$V_a\|q_2\rangle = \|q_3\rangle$	$V_a\|q_3\rangle = \|q_1\rangle$
V	$V_a\|q_4\rangle = \|q_4\rangle$	$V_b\|q_3\rangle = \|q_4\rangle$	$V_b\|q_4\rangle = \|q_3\rangle$
$V_\#\|q_5\rangle = \|q_{r_1}\rangle$	$V_a\|q_5\rangle = \|q_6\rangle$	$V_b\|q_5\rangle = \|q_5\rangle$	$V_b\|q_6\rangle = \|q_6\rangle$
V	$V_a\|q_6\rangle = \|q_{a_2}\rangle$		

Head functions:

$D(q_0) = (+1), D(q_1) = (-1), D(q_2) = (+1), D(q_3) = (-1),$

$D(q_4) = (+1), D(q_5) = (-1), D(q_6) = (+1), D(q_7) = (-1),$

$D(q_{a_1}) = D(q_{a_2}) = (0), D(q_{r_1}) = D(q_{r_2}) = (0)$

$$3BrO_3 - +5CH_2(COOH)_2 + 3H^+ \rightarrow 3BrCH(COOH)_2 + 4CO_2 + 2HCOOH + 5H_2O \quad [8]$$

In 2019, Duenas-Diez and Perez-Mercader [38] designed chemical Turing machine for BZ reaction network. The chemical reaction is fed sequentially to the reactor as $\{(BrO_3^-)^n (MA)^n (NaOH)^n\}$, where $n > 0$. It is transcribed in formal language as $L_3 = \{a^n b^n c^n | n > 0\}$. The symbol a is interpreted as a fraction of sodium bromate, b is used for malonic acid and symbol, c is transcribed as a quantity of NaOH. It is known that L_3 is a context-sensitive language and cannot be recognized by finite automata or pushdown automata with a stack. Although it can be recognized by two-stack PDA, we have shown that L_3 can be recognized by 2QFA without using any external aid.

THEOREM 4. A language $L_3 = \{a^n b^n c^n | n > 0\}$ can be recognized by 2QFA in linear time. For a language, $L_3 = \{a^n b^n c^n | n > 0\}$, and for arbitrary N-computational paths, there exists a 2QFA such that for $w \in L_3$; it accepts w with bounded error ϵ and rejects $w \notin L_3$ with probability at least $1 - \frac{1}{N}$.

TABLE 2 | Details of the transition functions and head function for L_2.

$$V_{\#}|q_0\rangle = |q_0\rangle, V_{(}|q_0\rangle = |q_0\rangle, V_{)}|q_0\rangle = |q_1\rangle, V_{\#}|q_1\rangle = |q_r\rangle$$

$$V_{(}|q_1\rangle = |q_2\rangle, V_{)}|q_2\rangle = |q_2\rangle, V_{(}|q_2\rangle = |q_3\rangle, V_{)}|q_3\rangle = |q_0\rangle$$

$$V_{\#}|q_2\rangle = \frac{1}{\sqrt{N}}\sum_{i=1}^{N}|q_{i,0}\rangle$$

$$V_{(}|q_{i,0}\rangle = |q_{i,i}\rangle, V_{)}|q_{i,0}\rangle = |q_{i,N-i+1}\rangle, \text{for } 1 \le i \le N, V_{(}|q_{i,j}\rangle = |q_{i,j-1}\rangle, \text{for } 1 \le j \le i$$

$$V_{)}|q_{i,j}\rangle = |q_{i,j-1}\rangle, \text{for } 1 \le j \le N-i+1, 1 \le i \le N$$

$$V_{\$}|q_{i,0}\rangle = |s_{i,0}\rangle, V_{)}|s_{i,0}\rangle = |w_{i,0}\rangle, V_{(}|s_{i,0}\rangle = |r_{i,0}\rangle, \text{for } 1 \le i \le N$$

$$V_{\$}|w_{i,0}\rangle = \frac{1}{\sqrt{N}}\sum_{k=1}^{N} exp\left(\frac{2\pi i}{N}ki\right)\Big|p_k\rangle, \text{for } 1 \le i \le N$$

Head functions:

$$D(q_0) = (+1), D(q_1) = (-1), D(q_2) = (+1), D(q_3) = (-1), D(q_{i,0}) = (+1), \text{for } 1 \le i \le N$$

$$D(q_{i,j}) = (0), \text{for } 1 \le i \le N, j \ne 0, D(q_r) = (0), D(p_k) = (0), \text{for } 1 \le k \le N$$

$$D(s_{i,0}) = (-1), D(w_{i,0}) = (+1), D(r_{i,0}) = (0), \text{for } 1 \le i \le N$$

PROOF. The design of proof for BZ reaction network is as follows. It consists of two phases. First, the 2QFA traverse the input to check the form $a^+b^+c^+$. On reading the right-end marker \$, the computation is split into N paths such that $w_{1,0}, w_{2,0}, \ldots, w_{N,0}$. Second, the first path is used to check if the number of bs and cs are equal or not. The second path is used to check the initial part of an input string to identify if it is in $\{a^n b^n | n > 0\}$. On reading the right-end marker \$, the both paths are split into N different paths with an equal amplitude $\frac{1}{\sqrt{N}}$. Finally, upon reading the right-end marker #, if the number of as and bs and the number of bs and cs are equal in respective computational paths, then all paths come into N-way quantum Fourier transform (QFT) and either one acceptance state or rejection states are observed. Suppose, if the input string is not in the corrected form, then all computation paths read the # at different times. Thus, their amplitudes do not cancel each other, and the input string is said to be rejected with probability $1 - \frac{1}{N}$. Otherwise, the input string is said to be recognized by 2QFA with probability 1.

5 SUMMARY

In summary, 2QFA model can be efficiently designed for balanced chemical reaction and BZ reaction network with one-sided error bound, which are halted in linear time. **Table 3** shows the language recognition ability of different computational models. The classical 2DFA and 2PFA are known to be equal in computational power to one-way deterministic finite automata (1DFA) [24, 61]. It has been proved that 2PFAs can be designed for nonregular languages in expected polynomial time. Additionally, it has been demonstrated that the chemical PDA can be designed for aforementioned chemical reactions with multiple stacks. The recognition of languages by native chemical automata can be found in Refs. 39–41. But, we have shown that 2QFA can recognize such chemical reactions without any external aid. It has been proved that 2QFA is more powerful than classical variants because it follows the quantum superposition principle to be in more than one state at a time on the input tape. For execution, it needs at least $O(log\ n)$ quantum states to store the position of tape head, where n denotes the length of an input string.

TABLE 3 | Comparison of computational power of models.

Languages	Class	2DFA/ 2PFA	Chemical FA/PDA	2QFA
$L_1 = \{(a+b)^* a (a+b)^* b (a+b)^* aa^* bb^*\}$	RL	✓	✓	✓
A language L_2 consisting Dyck language of balanced parentheses	CFL	✗	✓ (with 1-stack PDA)	✓
$L_3 = \{a^n b^n c^n \| n > 0\}$	CSL	✗	✓ (with 2-stack PDA)	✓

RL, CFL, and CSL stand for regular languages, context-free languages, and context-sensitive languages, respectively.

6 CONCLUSION

The enhancement in many existing computational approaches provides momentum to molecular and quantum simulations at the electronic level. It helps to test new abstract approaches for considering molecules and matter. Previous attempts to model the aforementioned chemical reactions used finite automata and pushdown automata with multiple stacks. In this study, we focused on well-known languages of Chomsky hierarchy and modeled those using two-way QFAs. The crucial advantage of the quantum approach is that these chemical reactions transcribed in formal languages can be parsed in linear time, without using any external aid. We have shown that two-way quantum automata are more superior to its classical variants by using quantum transitions. To the best of our knowledge, no such modeling of chemical reactions is performed using quantum automata theory so far. For the future purpose, we will try to represent complex chemical reactions in formal languages and model those using other quantum computational models.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article; further inquiries can be directed to the corresponding author/s.

AUTHOR CONTRIBUTIONS

AB and SZ modeled the chemical reactions. AB wrote the manuscript under guidance of SZ.

ACKNOWLEDGMENTS

SZ acknowledges support in part from the National Natural Science Foundation of China (No. 61602532), the Natural Science Foundation of Guangdong Province of China (No. 2017A030313378), and the Science and Technology Program of Guangzhou City of China (No. 201707010194). This work was also supported in part by NSF-China (61772570), Pearl River S&T Nova Program of Guangzhou (201806010056), and Guangdong Natural Science Funds for Distinguished Young Scholar (2018B030306025).

REFERENCES

1. Conrad M. Information processing in molecular systems. *BioSystems* (1972) **5**: 1–14. doi:10.1016/0303-2647(72)90011-1
2. Benenson Y. Biocomputers: from test tubes to live cells. *Mol Biosyst.* (2009) **5**: 675–685. doi:10.1039/b902484k
3. Prohaska SJ, Stadler PF, Krakauer DC. Innovation in gene regulation: the case of chromatin computation. *J Theor Biol.* (2010) **265**:27–44. doi:10.1016/j.jtbi.2010.03.011
4. Bryant B. Chromatin computation. *PLoS One.* (2012) **7**:e35703. doi:10.1371/journal.pone.0035703
5. Nielsen MA, Chuang I. *Quantum computation and quantum information.* New York: AAPT (2002).
6. Hopcroft JE, Motwani R, Ullman JD. *Automata theory, languages, and computation*. **Vol. 24**. New York: Pearson addison wesley (2006). 19 p.
7. Shor PW. Polynomial-time algorithms for prime factorization and discrete logarithms on a quantum computer. *SIAM Rev.* (1999) **41**:303–332. doi:10.1137/s0036144598347011
8. Kondacs A, Watrous J. On the power of quantum finite state automata. In: Proceedings 38th annual symposium on foundations of computer science; 1997 Oct 20–22; Miami Beach, FL. IEEE (1997) p. 66–75.
9. Moore C, Crutchfield JP. Quantum automata and quantum grammars. *Theor Comput Sci.* (2000) **237**:275–306. doi:10.1016/s0304-3975(98)00191-1
10. Bhatia AS, Kumar A. On the power of quantum queue automata in real-time. arXiv preprint arXiv:1810.12095 (2018).
11. Ambainis A, Watrous J. Two-way finite automata with quantum and classical states. *Theor Comput Sci.* (2002) **287**:299–311. doi:10.1016/s0304-3975(02)00138-x
12. Zheng S, Qiu D, Li L, Gruska J. *One-way finite automata with quantum and classical states. Languages alive*. Berlin, Heidelberg: Springer (2012) 273–290 p.
13. Mereghetti C, Palano B, Pighizzini G. Note on the succinctness of deterministic, nondeterministic, probabilistic and quantum finite automata. *RAIRO Theor Inf Appl.* (2001) **35**:477–490. doi:10.1051/ita:2001106
14. Zheng S, Qiu D, Gruska J, Li L, Mateus P. State succinctness of two-way finite automata with quantum and classical states. *Theor Comput Sci.* (2013) **499**: 98–112. doi:10.1016/j.tcs.2013.06.005
15. Yakaryilmaz A, Say A. *Succinctness of two-way probabilistic and quantum finite automata*. arXiv preprint arXiv:0903.0050 (2009).
16. Nishimura H, Yamakami T. An application of quantum finite automata to interactive proof systems. *J Comput Syst Sci.* (2009) **75**:255–269. doi:10.1016/j.jcss.2008.12.001
17. Nishimura H, Yamakami T. Interactive proofs with quantum finite automata. *Theor Comput Sci.* (2015) **568**:1–18. doi:10.1016/j.tcs.2014.11.030
18. Bhatia AS, Kumar A. Quantifying matrix product state. *Quant Inf Process.* (2018) **17**:41. doi:10.1007/s11128-017-1761-1
19. Zheng S, Li L, Qiu D, Gruska J. Promise problems solved by quantum and classical finite automata. *Theor Comput Sci.* (2017) **666**:48–64. doi:10.1016/j.tcs.2016.12.025
20. Gruska J, Qiu D, Zheng S. Generalizations of the distributed deutsch–jozsa promise problem. *Math Struct Comput Sci.* (2017) **27**:311–331. doi:10.1017/s0960129515000158
21. Gruska J, Qiu D, Zheng S. Potential of quantum finite automata with exact acceptance. *Int J Found Comput Sci.* (2015) **26**:381–398. doi:10.1142/s0129054115500215
22. Gainutdinova A, Yakaryılmaz A. Unary probabilistic and quantum automata on promise problems. *Quant Inf Process.* (2018) **17**:28. doi:10.1007/s11128-017-1799-0
23. Bhatia AS, Kumar A. Quantum ω-automata over infinite words and their relationships. *Int J Theor Phys.* (2019) **58**:878–889. doi:10.1007/s10773-018-3983-0
24. Shepherdson JC. The reduction of two-way automata to one-way automata. *IBM J Res Dev.* (1959) **3**:198–200. doi:rd.32.0198/rd.32.0198
25. Zheng S, Gruska J, Qiu D. On the state complexity of semi-quantum finite automata. *RAIRO Theor Inf Appl.* (2014) **48**:187–207. doi:10.1051/ita/2014003
26. Li L, Qiu D. Lower bounds on the size of semi-quantum finite automata. *Theor Comput Sci.* (2016) **623**:75–82. doi:10.1016/j.tcs.2015.09.031
27. Zheng S, Qiu D, Gruska J. Time-space complexity advantages for quantum computing. In: International conference on theory and practice of natural computing; 2017 Dec 18-20, 2017; Prague, Czech Republic Springer (2017). p. 305–317. doi:10.1007/978-3-319-71069-3_24
28. Yamakami T. Relativizations of nonuniform quantum finite automata families. In: International conference on unconventional computation and natural computation; Springer (2019). p. 257–271. doi:10.1007/978-3-030-19311-9_20
29. Yamakami T. Nonuniform families of polynomial-size quantum finite automata and quantum logarithmic-space computation with polynomial-size advice. In: International conference on language and automata theory and applications; 2019 Jun 3-7; Tokyo Springer (2019). p. 134–145. doi:10.1007/978-3-030-13435-8_10
30. Bhatia AS, Kumar A. On relation between linear temporal logic and quantum finite automata. *J Logic Lang Inf.* (2019) **29**:109–120. doi:10.1007/s10849-019-09302-6
31. Bertoni A, Mereghetti C, Palano B. Quantum computing: 1-way quantum automata. In: International conference on developments in language theory; Springer (2003). p. 1–20. doi:10.1007/3-540-45007-6_1
32. Qiu D, Li L. An overview of quantum computation models: quantum automata. *Front Comput Sci China.* (2008) **2**:193–207. doi:10.1007/s11704-008-0022-y
33. Ambainis A, Yakaryılmaz A. Automata and quantum computing. arXiv preprint:1507.01988 (2015).
34. Bhatia AS, Kumar A. *Quantum finite automata: survey, status and research directions*. arXiv preprint arXiv:1901.07992 (2019).
35. Qiu D, Li L, Mateus P, Sernadas A. Exponentially more concise quantum recognition of non-rmm regular languages. *J Comput Syst Sci.* (2015) **81**: 359–375. doi:10.1016/j.jcss.2014.06.008
36. Singh Bhatia A. *On some aspects of quantum computational models*. [PhD thesis]. Patiala (India): Thapar Institute of Engineering and Technology (2020).
37. Say AC, Yakaryılmaz A. Quantum finite automata: a modern introduction. *Computing with new resources*. New York: Springer (2014). p. 208–222.
38. Duenas-Diez M, Perez-Mercader J. Native chemical automata and the thermodynamic interpretation of their experimental accept/reject responses. arXiv preprint:1903.03827 (2019).
39. Okubo F, Yokomori T. The computing power of determinism and reversibility in chemical reaction automata. *Reversibility and universality*. Springer (2018). p. 279–298.
40. Dueñas-Díez M, Pérez-Mercader J. How chemistry computes: language recognition by non-biochemical chemical automata. from finite automata to turing machines. *iScience* (2019) **19**:514–526. doi:10.1016/j.isci.2019.08.007
41. Freivalds R. Probabilistic two-way machines. In: International symposium on mathematical foundations of computer science; 1981 Aug 31 - Sep 4; Štrbské Pleso, Czechoslovakia Springer (1981). p. 33–45. doi:10.1007/3-540-10856-4_72
42. Dwork C., Stockmeyer L. A time complexity gap for two-way probabilistic finite-state automata. *SIAM J Comput.* (1990) **19**:1011–1023. doi:10.1137/0219069
43. Okamoto M., Sakai T., Hayashi K. Switching mechanism of a cyclic enzyme system: role as a "chemical diode". *Biosystems* (1987) **21**:1–11. doi:10.1016/0303-2647(87)90002-5
44. Hjelmfelt A, Weinberger ED, Ross J. Chemical implementation of neural networks and Turing machines. *Proc Natl Acad Sci USA.* (1991) **88**: 10983–10987. doi:10.1073/pnas.88.24.10983
45. Tóth Á, Showalter K. Logic gates in excitable media. *J Chem Phys.* (1995) **103**; 2058–2066. doi:10.1063/1.469732
46. Magnasco MO. Chemical kinetics is turing universal. *Phys Rev Lett.* (1997) **78**: 1190. doi:10.1103/physrevlett.78.1190
47. Adamatzky A, Costello BDL. Experimental logical gates in a reaction-diffusion medium: the xor gate and beyond. *Phys Rev.* (2002) **66**:046112. doi:10.1103/physreve.66.046112
48. Gorecki J, Yoshikawa K, Igarashi Y. On chemical reactors that can count. *J Phys Chem.* (2003) **107**:1664–1669. doi:10.1021/jp021041f
49. Adleman LM. Molecular computation of solutions to combinatorial problems. *Science* (1994) **266**:1021–1024. doi:10.1126/science.7973651

50. Kuppusamy L, Mahendran A. Modelling DNA and RNA secondary structures using matrix insertion–deletion systems. *Int J Appl Math Comput Sci.* (2016) **26**:245–258. doi:10.1515/amcs-2016-0017
51. Kuppusamy L, Mahendran A, Krishna SN. Matrix insertion-deletion systems for bio-molecular structures. In: International conference on distributed computing and internet technology; Feb 9-12; Bhubaneshwar, India Springer (2011). p. 301–312. doi:10.1007/978-3-642-19056-8_23
52. Krasiński T, Sakowski S, Poplawski T. Towards an autonomous multistate biomolecular devices built on DNA. In: 2014 Sixth World Congress on Nature and Biologically Inspired Computing (NaBIC 2014); 2014 Jul 30 - Aug 01; Porto, Portugal IEEE (2014). p. 23–28.
53. Khrennikov A, Yurova E. Automaton model of protein: dynamics of conformational and functional states. *Prog Biophys Mol Biol.* (2017) **130**: 2–14. doi:10.1016/j.pbiomolbio.2017.02.003
54. Bhatia AS, Kumar A. Modeling of RNA secondary structures using two-way quantum finite automata. *Chaos Solit Fractals.* (2018) **116**:332–339. doi:10.1016/j.chaos.2018.09.035
55. Bhatia AS, Zheng S. RNA-2QCFA: evolving two-way quantum finite Automata with classical states for RNA secondary structures. arXiv preprint arXiv:2007.06273 (2020).
56. Wang J. *Handbook of finite state based models and applications*. Boca Raton, FL: CRC Press (2012).
57. Dzelme-Bērziņa I. *Quantum finite automata and logic*. [PhD thesis]. Riga (Latvia):University of Latvia (2010).
58. Amano M, Iwama K. Undecidability on quantum finite automata. In: STOC '99: proceedings of the 31st annual ACM symposium on theory of computing; 1999 May 1-4; Atlanta, GA. Georgia: ACM (1999). p. 368–375.
59. Bhatia AS, Kumar A. On the power of two-way multihead quantum finite automata. *RAIRO Theor Inf Appl.* (2019) **53**:19–35. doi:10.1051/ita/2018020
60. Field RJ, Koros E, Noyes RM. Oscillations in chemical systems. ii. thorough analysis of temporal oscillation in the bromate-cerium-malonic acid system. *J Am Chem Soc.* (1972) **94**:8649–8664. doi:10.1021/ja00780a001
61. Rabin MO, Scott D. Finite automata and their decision problems. *IBM J Res Dev.* (1959) **3**:114–125. doi:10.1147/rd.32.0114

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be constructed as a potential conflict of interest.

frontiers in Chemistry

REVIEW
published: 24 November 2020
doi: 10.3389/fchem.2020.587143

Application of Quantum Computing to Biochemical Systems: A Look to the Future

Hai-Ping Cheng[1], *Erik Deumens*[1], *James K. Freericks*[2]*, *Chenglong Li*[3] and *Beverly A. Sanders*[4]

[1] Quantum Theory Project, Department of Physics, University of Florida, Gainesville, FL, United States, [2] Department of Physics, Georgetown University, Washington, DC, United States, [3] Department of Medicinal Chemistry, University of Florida, Gainesville, FL, United States, [4] Department of Computer and Information Science and Engineering, University of Florida, Gainesville, FL, United States

Chemistry is considered as one of the more promising applications to science of near-term quantum computing. Recent work in transitioning classical algorithms to a quantum computer has led to great strides in improving quantum algorithms and illustrating their quantum advantage. Because of the limitations of near-term quantum computers, the most effective strategies split the work over classical and quantum computers. There is a proven set of methods in computational chemistry and materials physics that has used this same idea of splitting a complex physical system into parts that are treated at different levels of theory to obtain solutions for the complete physical system for which a brute force solution with a single method is not feasible. These methods are variously known as embedding, multi-scale, and fragment techniques and methods. We review these methods and then propose the embedding approach as a method for describing complex biochemical systems, with the parts not only treated with different levels of theory, but computed with hybrid classical and quantum algorithms. Such strategies are critical if one wants to expand the focus to biochemical molecules that contain active regions that cannot be properly explained with traditional algorithms on classical computers. While we do not solve this problem here, we provide an overview of where the field is going to enable such problems to be tackled in the future.

Keywords: computational molecular biology, biochemistry, quantum computing, hybrid quantum-classical algorithms, quantum embedding theory

Edited by:
Sabre Kais,
Purdue University, United States

Reviewed by:
Son Tung Ngo,
Ton Duc Thang University, Vietnam
Francoise Remacle,
University of Liège, Belgium

***Correspondence:**
James K. Freericks
james.freericks@georgetown.edu

Specialty section:
This article was submitted to Physical Chemistry and Chemical Physics, a section of the journal Frontiers in Chemistry

Received: *24 July 2020*
Accepted: *15 October 2020*
Published: *24 November 2020*

Citation:
Cheng H-P, Deumens E, Freericks JK, Li C and Sanders BA (2020) Application of Quantum Computing to Biochemical Systems: A Look to the Future. Front. Chem. 8:587143. doi: 10.3389/fchem.2020.587143

1. INTRODUCTION

Biochemical systems are essential for carrying out biological functions, and their actions span extreme time and length scales. These systems consist of proteins, DNAs, RNAs, carbohydrates, or lipids (either individually or in combination) with small molecule ligands and/or with ions in aqueous or membrane environments. The functional processes can be either covalent or non-covalent, such as molecular recognition; or a combination of both, such as an enzymatic cycle. Important biological functions are, for example, stem cell maintenance, DNA repair, gene transcription and translation, signal transduction, development, learning and memory, metabolism, etc. In order to understand these elementary processes, together with experimental approaches, various computational methods have been developed at the electronic, the atomic, and more coarse-grained levels over the decades. However, full quantum calculations are intractable due to the large molecule sizes and the high demands for accuracy required for chemical applications.

The solution may lie in quantum computing: as Feynman once said, "... Nature isn't classical...if you want to make a simulation of Nature, you'd better make it quantum mechanical..." (Feynman, 1982). As a matter of fact, from the remarkable speed of enzyme-catalyzed reactions to the workings of the human brain, numerous biological puzzles are now being explored for evidence of quantum effects. Well-known examples include photosynthesis, nitrogen fixation, magnetoreception, olfaction, neuronal signal processing, protein/drug interaction, and so on. There have even been early attempts to develop quantum computing algorithms specifically for nitrogen fixation (Reiher et al., 2017).

Quantum computing is being explored to help solve a variety of problems in biochemistry and biology (Cao et al., 2019; Emani et al., 2019). In this paper, we review several approaches to allow quantum computing to be exploited to simulate biochemical systems with complicated electron correlation. We formulate a general approach of embedding to describe part of the system on classical computers and the most demanding part on a quantum computer resulting in a complete solution of the complex system with useful accuracy. This will allow quantum computers to be used for such demanding problems without the requirement that a quantum computer be available to hold and process the entire system of interest. In section 2, as motivating examples, we present three biochemical systems that are intractable with classical algorithms on classical computers due to the need to deal with complicated electron correlation. Effectively addressing them with quantum computing will lead to important scientific advances. Then we review the embedding methodologies that have been used to handle very complex systems using classical computers in section 3, which consists of dividing the system into two parts, with one part, the most computationally demanding, computed with quantum theory and the other part, considered the environment, treated with classical theory. The challenge of embedding methods is the exchange of information between the two parts. Section 4 provides a brief review of two of the most important existing quantum algorithms for chemistry, Variational Quantum Eigensolver and Quantum Phase Estimation. Section 5 presents how the idea of embedding strategies can be usefully applied to handle complex physical systems at a high level of accuracy by combining the power of quantum computers for the strongly correlated part of the system with the use of classical computers for the other parts.

In the context of computational chemistry, the distinction of quantum vs. classical has two meanings that are both relevant: The first, traditional, meaning designates the level of theory that is used to describe the chemical and biochemical systems. Because of the complexity and size of biochemical systems, treating the whole system using quantum theory is not feasible and often the systems are described using theories based on classical physics. The second, more recently introduced, meaning refers to whether the theoretical model and computational algorithms are run on quantum computers or classical computers. The promise of quantum computers is that they will eventually be sufficiently powerful to allow scientists to model complex biochemical systems accurately and efficiently with a fully quantum theoretical description. In this paper, we will make it clear which meaning is used when it is used.

2. MOTIVATING EXAMPLES

We present three important and representative biochemical systems whose properties make them attractive targets for quantum computing. The first two are open shell transition-metal and conjugated pi-electron strongly correlated systems; the last one displays extreme non-covalent intermolecular binding involving a large number of atoms. These three examples symbolize difficult cases for classical quantum chemical treatments and superior ones for quantum computing.

2.1. A Transition-Metal-Ion-Containing Enzyme: Histone Demethylase

The transcription of genetic information encoded in DNA is in part regulated by chemical modifications to histone proteins. Histone demethylases are enzymes that remove methyl (-CH3) groups from histones. The demethylase proteins alter transcriptional regulation of the genome by controlling the methylation levels that occur on DNA and/or histones and, in turn, regulate the chromatin states at specific gene loci within organisms. The big demethylase family has KDM1-6 classes (Pedersen and Helin, 2010). Defined by their mechanisms, two main classes of histone demethylases exist: a flavin adenine dinucleotide (FAD)-dependent amine oxidase, and an Fe(II) and α-ketoglutarate-dependent hydroxylase. Both operate by hydroxylation of a methyl group followed by dissociation of formaldehyde. By studying various demethylation details, improvements are possible in the understanding of how "histone code" is employed for gene on/off switching.

Figure 1 is an illustration of the JmJD2A topology, active site, and proposed catalytic mechanism which involve both transition metal ions and reaction radicals (Chen et al., 2006; Ng et al., 2007; Zheng and Huang, 2014). These are cases where the Born-Oppenheimer approximation breaks down. During the catalytic cycle, the iron metal ion has three charge states: +2, +3, and +4, and two spin states: 0 and 1/2. Oxygen has three spin states: 0, 1/2 and 1. There are at least nine catalytic steps. Considering only direct contact catalytic amino acid residues, oxygen, trimethylated quaternary amine from lysine substrate, and of course catalytic Fe ion, 151 electrons and 121 spatial orbitals must be involved to achieve accurate electronic structure and related energy calculations.

2.2. Non-Metal-Ion-Containing Enzyme: Telomerase

The study of telomerase is of tremendous significance for understanding stem cell maintenance, aging, and cancer. At each end of a chromosome, there is a region of repetitive nucleotide sequences called a telomere which protects the chromosome from deterioration or fusion with neighboring chromosomes.

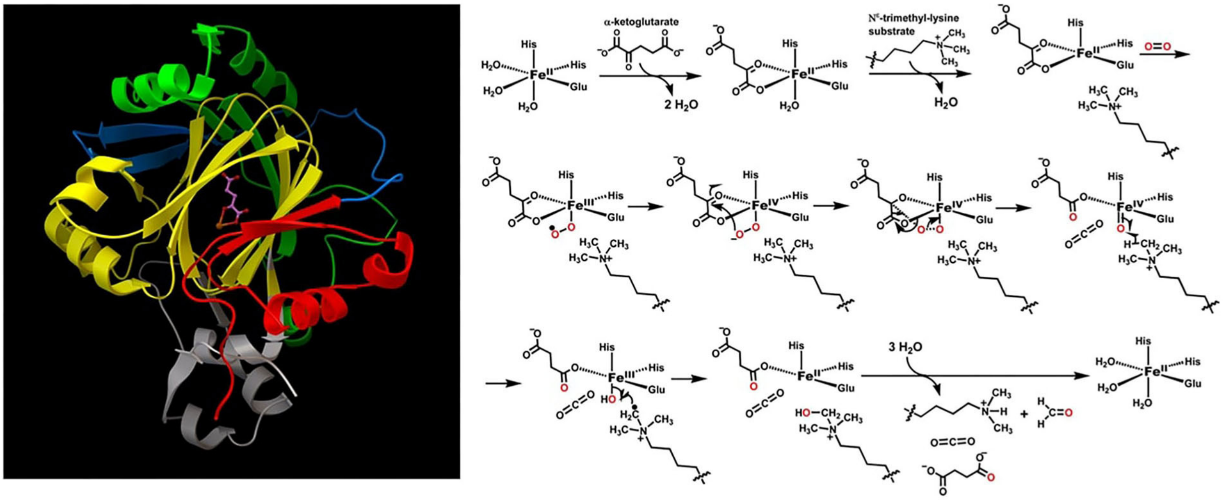


FIGURE 1 | Structure of JmJD2A. Some domains from above are highlighted: JmJ (N-terminus, red; C-terminus, yellow), Zinc finger domain (light purple), Beta-hairpin (light blue), and mixed domain linker (green). The ball-and-sticks are Fe(II) and alpha-ketoglutarate cofactors. The enzymatic reactions involve both iron redox and oxygen radical, and are thus infeasible with classical computers.

During chromosome replication, the enzymes that duplicate DNA cannot continue their duplication all the way to the end of a chromosome, so after each cell division, the telomere gets shorter. Telomeres are replenished by an enzyme, as shown in **Figure 2**, telomerase reverse transcriptase (TERT), (Cohen et al., 2007) which is the catalytic subunit of telomerase. Telomerase is active in normal stem cells and most cancer cells, but is normally absent from, or at very low levels in, most somatic (body) cells.

For vertebrates, the sequence of nucleotides in telomeres is AGGGTT (Harvey, 2014). The complementary DNA strand is TCCCAA, which also has a single-stranded TTAGGG overhang (Witzany, 2008). This sequence of TTAGGG is repeated approximately 2,500 times in humans (Sadava et al., 2011). The active telomerase is a homodimer, each monomer having telomerase reverse transcriptase (TERT), telomerase RNA, and dyskerin (Mitchell et al., 2010). Currently, there are several TERT crystal structures available; computational simulation of TERT telomere elongation is important. Snapshots of the molecular processes need be constructed and quantum computing could be used to simulate the catalytic active centers in order to better understand how these systems work, especially base fidelity preservation during the extension process. Due to its nucleobase pairing and reaction processivity, this is a case where quantum computing could make a large impact on molecular recognition.

2.3. Molecular Recognition: Biotin-(Strept)Avidin Binding

Molecular recognition, the specific interaction between multiple molecules which exhibit molecular complementarity through non-covalent binding, plays a critical role in biological interactions. Although the field is well-studied, important problems remain unsolved. For example, even though it is a classic molecular recognition issue and many studies have attempted to resolve it, the origin of strong non-covalent reversible binding of small molecule biotin to proteins avidin (Ka ∼ 10E15 M-1) and streptavidin (Ka ∼ 10E13 M-1) remains a mystery. As seen from **Figure 3**, the beta-barrel shaped avidin binds the biotin ligand with van der Waals, electrostatic, hydrogen bonding and pi-electron polarization forces; this results in a free energy of binding around −20 Kcal/mol, almost at a quasichemical bonding level.

Understanding this binding thermodynamics at the molecular level holds fundamental importance theoretically and offers key insights for molecular design. There is no doubt that in desolvation, conformational and vibrational entropy play an important role. However, the key issue here is to understand biotin-avidin intermolecular interaction, which rests on computing accurate non-bonded interaction energies. The biotin molecule (as the ligand) has 89 valence electrons and 79 frontier orbitals. By only considering direct contacting atoms from the binding amino acid residues, the active part of the molecule has a total of 379 electrons and 358 spatial orbitals. The spin state is $S = 0$ and charge state is -1. It should be obvious that such a subsystem is too large to be tackled with standard chemical methodologies for strongly correlated molecules, so quantum computing is the only option for a complete theoretical analysis. Sometimes, ligands form covalent bonding with proteins, such as anticancer drug ibrutinib binding to Bruton's tyrosine kinase (Bender et al., 2017). Quantum computing mimicking these processes not only helps fundamental understanding on molecular recognition, but also facilitates drug or materials design.

3. CLASSICAL COMPUTER EMBEDDING STRATEGIES

In this section, we review the methodology of embedding as it has been used for several decades to describe complex chemical

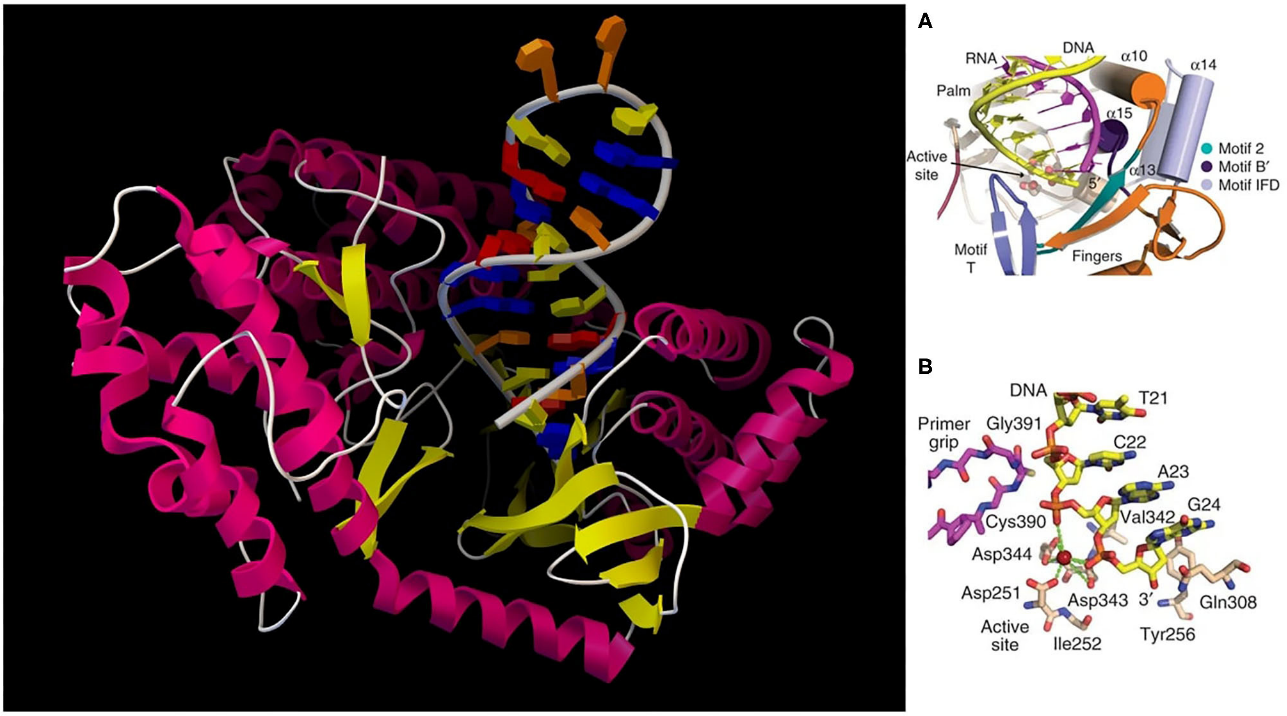


FIGURE 2 | (Left) TERT with hybrid RNA/DNA bound; (Right) **(A)** cartoon representation of the active site; **(B)** detailed active site residues and DNA substrate.

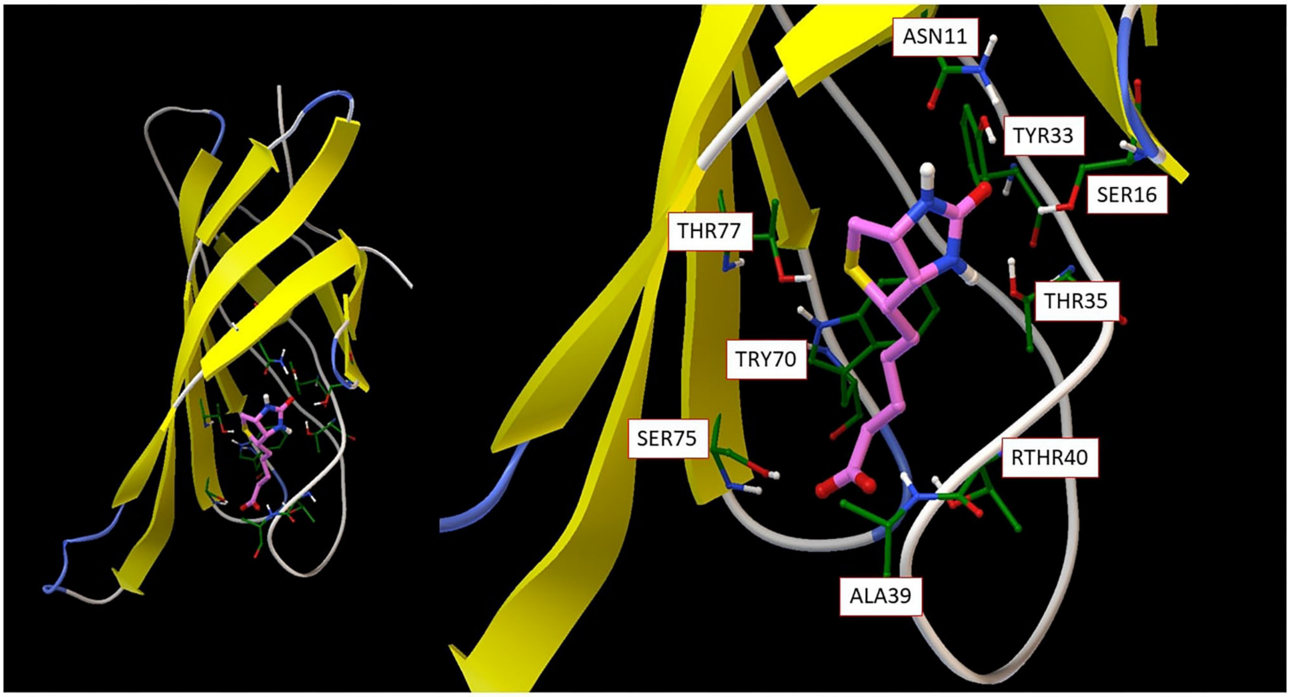


FIGURE 3 | (Left) overall avidin protein structure with biotin binding; **(Right)** detailed biotin interaction amino acid residues from avidin.

systems, including large bio-molecules in liquids and solid-state systems, by dividing them into parts that are treated with different levels of physical theory. The levels of theory range from continuum models, to classical dynamics of atoms and molecules, to full quantum-mechanical description of electronic structure and nuclear motion. These methods are then implemented in algorithms that run on classical computers. The challenge common to all these methods lies in the description of the interactions across the boundaries between the parts. There the interaction must be described with care because of the different theories being used to describe the parts on either side of the boundary.

The same methodology used to divide complex systems into parts is extended to describe some parts with algorithms that are executed on classical computers, while other parts are described by algorithms that are executed on quantum computers. The challenge of embedding methods on classical and quantum computers is the same in that the description of the interaction between the parts running on the classical and quantum computers must be handled with extreme care. In addition, because one cannot directly readout the final wavefunction from the quantum computer, hybrid algorithms must be properly designed to allow for the information from the quantum calculation to be transferred to the classical algorithm and *vice versa*.

In the method of multiscale simulation, the challenge is to describe the processes that are visible at the macroscopic level, but are fully determined by the details at some microscopic level. A paradigmatic example is the formation and propagation of cracks in materials (Gao and Klein, 1998; Rudd and Broughton, 2000; Rountree et al., 2002; Liu et al., 2004; Budarapu et al., 2014; Talebi et al., 2014). The macroscopic description is the goal, but continuum models that are effective and affordable at that scale cannot describe the basic-bond breaking process that lie at the foundation of the crack formation. Nor can molecular dynamics methods describe this process. Thus the continuum model, the molecular dynamics model, and the quantum model must all be coupled together with scale-bridging techniques used to connect them in a way that accurately preserves the physics (Hoekstra et al., 2014).

Similarly, a biomolecule can be divided into three regions: a classical region where interatomic interaction can be treated with classical force fields using standard methods (Amber http://ambermd.org/, CHARMM https://www.charmm.org/charmm/, LAMMPS https://lammps.sandia.gov/, etc); a quantum region where mean-field approximations are sufficient; and a strongly correlated region where high-level methods that treat quantum entanglement are needed, that is, techniques beyond density functional theory (DFT).

3.1. Hybrid Quantum-Classical Molecular Dynamics

We illustrate how hierarchical methods have been used in a few studies. We show a study of water-silica surface interaction, which shows that very complex amorphous systems can be handled by the method. In the amorphous water-silica interface interaction, the system is divided into two regions, the quantum and the classical. Here, the quantum region is described by effective mean-field methods, while the classical region is described by molecular dynamics using effective force laws. The two regions must be coupled together across their boundary. Various methods exist for the embedding of the quantum region (the light blue region in **Figure 4**) inside the classical region. In earlier work (Du et al., 2004), a quantum region described by DFT is embedded in a classical matrix as shown in **Figure 5**. This figure depicts a Si-O bond-breaking process on the silica surface. According to a free cluster model (Walsh et al., 2000), the calculated barrier energy of this process is E_b = 0.7–1.1 eV. However, when the cluster is embedded in a surface matrix, the calculated barrier energy E_b is equal to 0.4 eV. Including quantum effects results in a substantial decrease.

For a bio-molecule, the embedding is simpler than for amorphous materials as there are not as many bonds that connect the classical region and the quantum region. Techniques for this type of embedding are quite sophisticated (Gao and Xia, 1992; Bakowies and Thiel, 1996; Gao et al., 1998; Cui et al., 2001; Laio et al., 2002; Vreven et al., 2003; Friesner and Guallar, 2005).

The peptide hydrolysis reaction mechanism for HIV-1 protease has been studied by a hybrid Car-Parinello/classical molecular dynamics simulation (Piana et al., 2004)/Gradient-corrected BLYP density functional theory describes the reactive part of the active site and the AMBER force field describes the rest of the proteins, the solvent, and the counterions that are needed to balance the QM/MM description. The authors find that the orientation and the flexibility of the reactants, determined by the embedding protein structure, are important in determining the activation barrier for the reaction. This shows the need to include the larger structure as well as the ability of the QM/MM approach to address the problem.

As another example, a recent study (Ahsan and Senapati, 2019) on the effect of water, i.e., hydrogen bonds, in a catalytic role in epoxide ring opening in aspartate proteases using QM/MM show that the process follows a two-step mechanism with the formation of an oxyanion intermediate, which is stabilized with up to 30 kcal/mol supplied by the hydrogen bonds from the water molecules near the protein active site. This example, too, illustrates that the treatment of the whole system is crucial for a correct understanding of the biochemistry enabled by the hybrid approach.

3.2. From DFT to Strongly Correlated Systems: Quantum Embedding Theory for Molecules and for the Hubbard Model

Transition-metal molecular complexes with d and f electrons often demonstrate strong correlation effects. Active centers of many enzymes are transition-metal complexes; e.g., photosystem I and II have iron-sulfur and manganese-oxide clusters as their active centers. The large number of atoms in ligands makes high-level calculation of the whole molecule impossible. In this situation, quantum embedding is necessary. Note that quantum embedding is different from the hybrid quantum-classical simulation discussed in section 3.1. Here, we embed

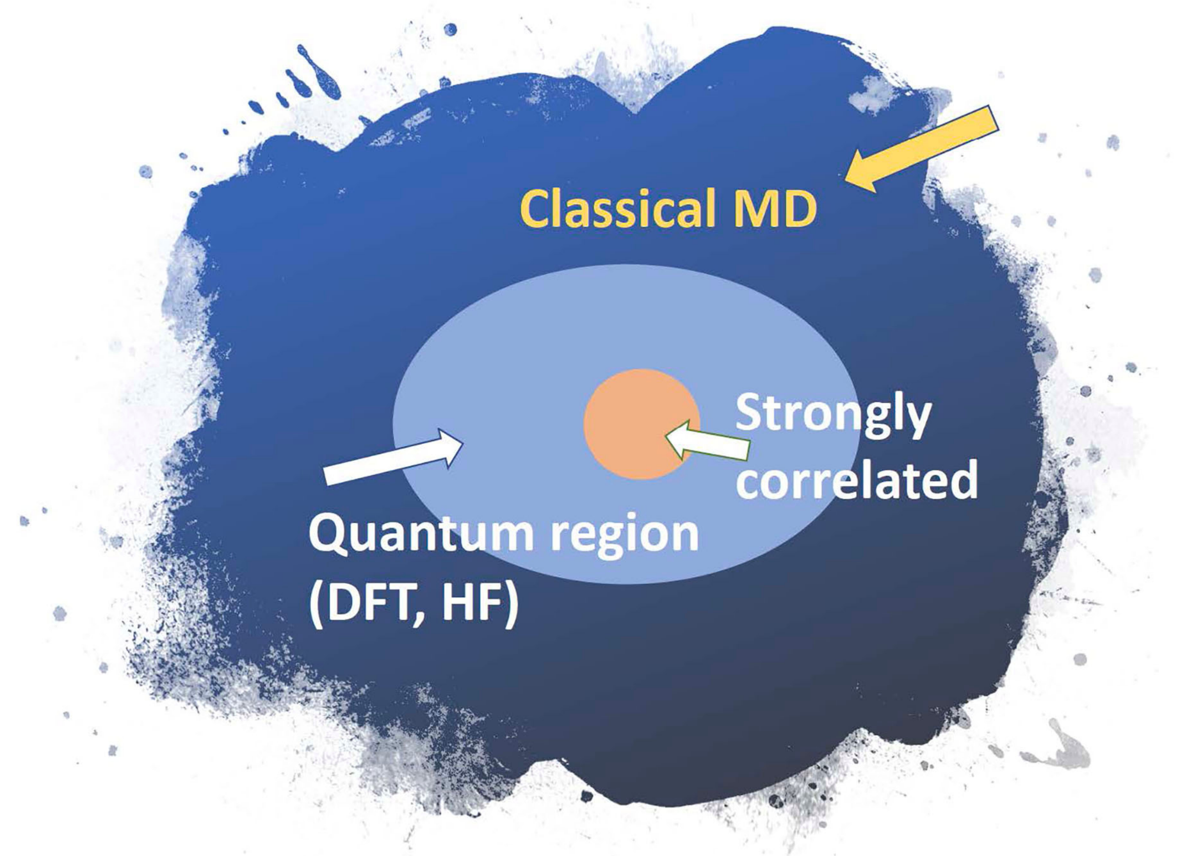


FIGURE 4 | Schematic of a hybrid simulation framework for molecules that employs a hierarchical embedding strategy.

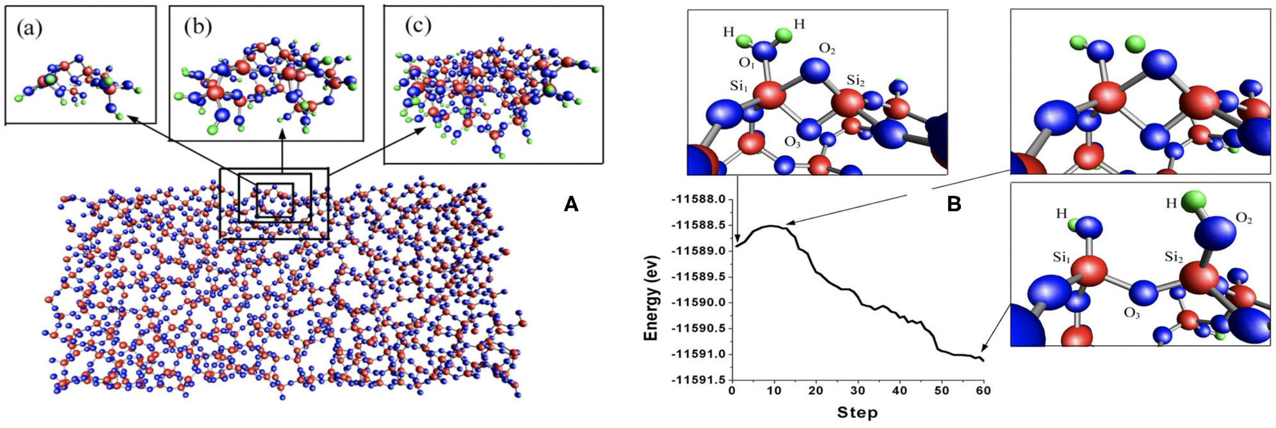


FIGURE 5 | (A) Sketch of the hybrid simulation framework as applied to amorphous glasses from Du et al. (2004). Three sizes of the quantum region were chosen (panels a–c on the left) to ensure the convergence of the reaction energy. **(B)** The right picture relates energy path with water splitting process. The energy barrier is only 0.4 eV and is zero when the reaction involves only two water molecules.

a strongly correlated subspace in a single electron space. So, we need a single-particle theory for the whole molecule and a many-body theory for the correlated subspace, which is a small but functional part of the molecule. One embedding scheme utilizes density functional theory (DFT) as the single-particle theory and does the embedding via dynamical mean-field theory (DMFT) (Georges et al., 1996). In DMFT, the correlated subspace is referred as an impurity and the impurity problem is solved by an impurity solver. We can use unitary coupled-cluster theory, which can be run on quantum computers, to create an approximation of the ground state of the active region of the molecule. Then additional qubits are employed to represent the self-consistent bath that the impurity is coupled to. Time evolution then allows the Green's function to be determined for the impurity, which can be directly measured and have its self-energy extracted after the Green's function data is transformed from the time domain to the frequency domain. The impurity self-energy is then approximated as the self-energy of the molecule in DMFT, and we can use it to obtain the molecular interacting Green's function, which in turn will be used to calculate physical properties of the molecule. Most likely, these problems will require an inhomogeneous DMFT approach, with separate impurity problems for the different atomic sites in the strongly correlated material. The challenging computation of the local Green's function for each atomic site from the local self-energy would be carried out on a classical computer.

The impurity model is defined, in part, from the on-site Coulomb interaction U-matrix. One reliable way to determine these parameters from first principles is the constrained Random Phase Approximation (cRPA) method (Aryasetiawan et al., 2004, 2006). One aims to estimate the screened Coulomb interaction for selected bands of interest, that is, within a specified energy window. For this purpose, the particle-hole polarization between all possible pairs of occupied and unoccupied states is taken into account. This approach uses the Random Phase Approximation (RPA) and directly calculates the particle-hole polarization (Petersilka et al., 1996; Aryasetiawan et al., 2004).

To make DFT+DMFT fully *ab initio*, the hopping parameters and the Coulomb interaction parameters should be provided from first principles (in the DFT part of the calculation). The Hubbard Hamiltonian can be written as

$$\hat{H}_{\text{Hubbard}} = \sum_{i,j} t_{ij}\,\hat{c}_i^{\dagger}\,\hat{c}_j + \sum_{i,\alpha,\beta,\gamma,\delta} U_{i,\alpha\beta\gamma\delta}\,\hat{c}_{i,\alpha}^{\dagger}\,\hat{c}_{i,\beta}^{\dagger}\,\hat{c}_{i,\gamma}\,\hat{c}_{i,\delta}. \qquad (1)$$

The hopping matrix t_{ij} comes from the DFT eigenenergies and provides the bath Green's function in DMFT. The Coulomb interaction $U_{i,\alpha\beta\gamma\delta}$ comes from the cRPA calculation described above, which is the only fully quantum-mechanical way to obtain the Coulomb interaction parameters. With the bath Green's function and U at hand, the effective action of the impurity problem is constructed. DMFT solves for the impurity Green's function by direct numerical sampling of the Green's function $G_{ij,\sigma}(t) = -\langle T\hat{c}_{i,\sigma}(t)\,\hat{c}_{j,\sigma}^{\dagger}(0)\rangle$. A classical computer algorithm often uses the continuous time quantum Monte Carlo algorithm (CT-QMC) (Gull et al., 2011; Zhang et al., 2019).

It has been proposed (Bauer et al., 2016) that a quantum computer algorithm could replace the CT-QMC calculation and provide the impurity Green's function $G_{ij,\sigma}(t)$, especially in cases where the classical computation suffers from the sign problem. Such a calculation embeds the impurity solver onto the quantum computer (quantum computing task), while the remainder of the DFT+DMFT iteration is carried out on classical computers. However, because describing the bath for the impurity problem is complex, it might be fruitful to instead simply solve the many-band lattice problem directly on the quantum computer. Indeed, this latter approach is more likely to be generalizable to large molecular systems. Of course, because molecules are not periodic, one will likely need to use inhomogeneous DMFT approaches if one takes the impurity problem approach.

It often is important to embed the DFT+DMFT iteration into a larger loop of charge-density (ρ^e) self-consistency (CSC) (**Figure 6**). It is known that CSC DFT+DMFT is necessary to capture charge density re-distributions even for very simple transition metal oxides like V_2O_3 under ambient conditions (Leonov et al., 2015). Similarly, in molecular calculations, charge redistribution is important for any simulation involving catalysis or other reactions. It is clear that all of these types of calculations must continue to be done on a classical computer. Only the complex strongly correlated part, involving the measurement of the Green's function will be done on the quantum computer. Because it can be directly measured, the connection between the quantum and classical calculation is simple to implement in this case.

4. BRIEF OVERVIEW OF CHEMISTRY ON QUANTUM COMPUTERS

4.1. Quantum Algorithms and Methods

One basic challenge in computational chemistry is finding a way to avoid having to explicitly maintain the full many-body wave function, because the classical resource requirements of doing so grow exponentially with the system size. In a quantum computer, on the other hand, this scaling is linear in the number of qubits used. Another classical computational challenge is propagating the wave function in time by computing $e^{-i\mathcal{H}t}\,|\Psi\rangle$. This calculation can also be efficiently implemented, in principle, on a quantum computer. This gives the promise of eventually being able to use a quantum computer to handle systems that cannot be feasibly calculated on classical computer. Such analysis can even be extended to beyond Born-Oppenheimer effects by including additional orbitals for the nuclei (Veis et al., 2016), but those approaches require significantly more resources and are likely to be applied to these large systems only far into the future. The disadvantage of working on the quantum computer is that the wave functions cannot be directly retrieved—either we resort to multiple calculations and measurements to obtain statistical knowledge of the wave function, or we settle for measuring some property of the wave function. Thus quantum algorithm design is not trivial.

In the last 20 years, significant progress has been made toward the goal of performing quantum chemistry on quantum computers (Cao et al., 2019). Most recently, Google has achieved a milestone in computational quantum chemistry by performing a Hartree-Fock calculation, the foundational algorithm in the field, on a superconducting quantum computer (Google AI Quantum and Collaborators, 2020). In this section, we describe two paradigmatic algorithms. The first is the variational quantum eigensolver (Peruzzo et al., 2014; McArdle et al., 2020), which is viewed today as the best candidate for performing chemistry

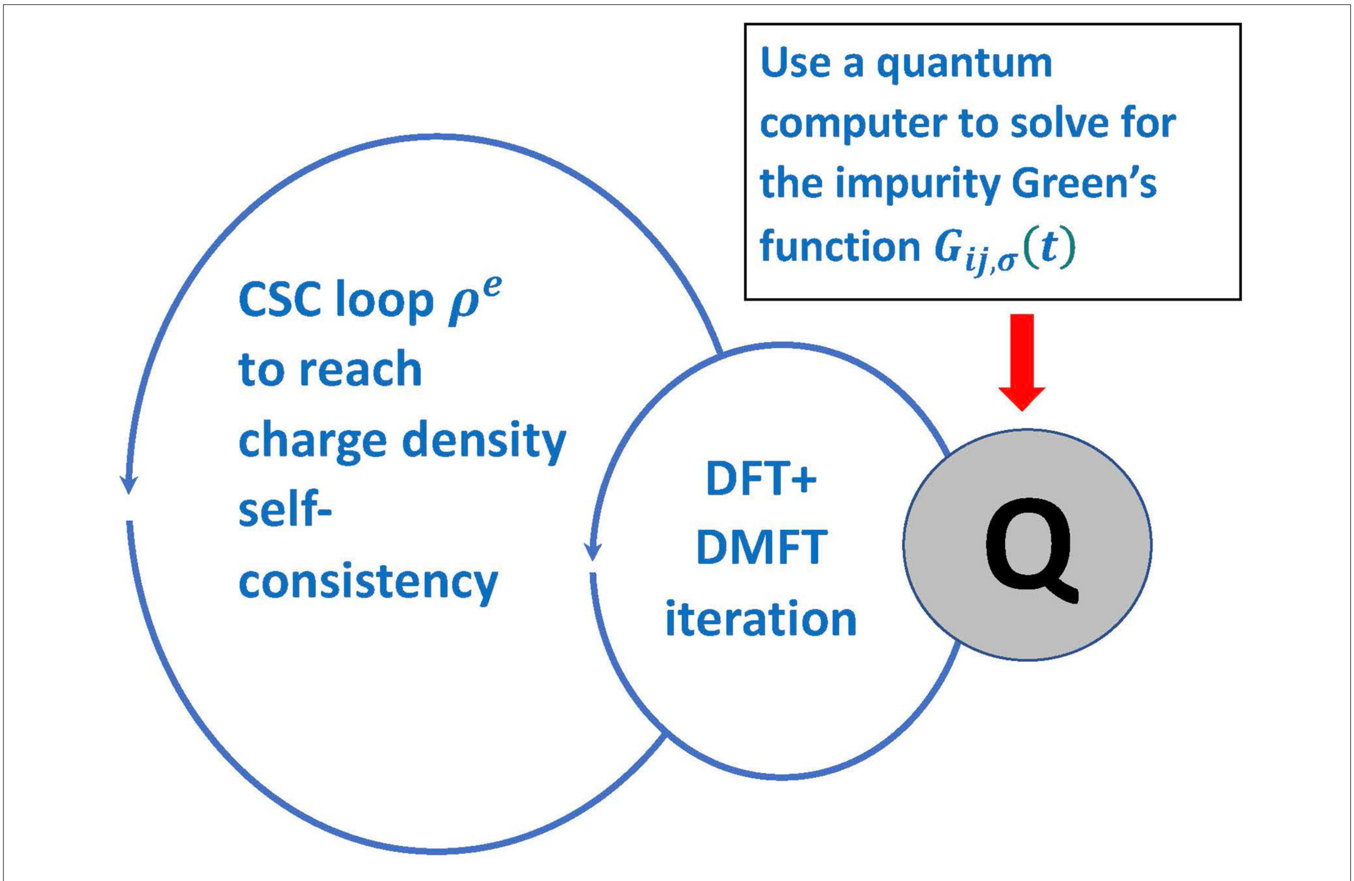


FIGURE 6 | Schematic plot of the charge density self-consistent loop in conjunction with DFT+DMFT.

on so-called noisy intermediate-scale quantum (NISQ) hardware. We also discuss the more accurate quantum phase estimation algorithm (Kitaev, 1995), which will ultimately emerge as the gold standard for quantum chemistry on a quantum computer because it can compute ground-state energies with only small systematic errors.

Of course, the first chemical systems put onto quantum computers are not going to be large biological molecules. But, with the development of the right algorithms for embedding, hierarchical structuring, and low-depth circuits, one might be able to advance biological science sooner than later. At the least, we should position ourselves to be able to try.

4.1.1. Variational Quantum Eigensolver

In the NISQ era, quantum computers will not be able to accurately execute deep circuits. They also generate results that require error/noise mitigation due to errors in state preparation and measurement, and from infidelities in quantum gate executions. Within this realm of quantum hardware, there is an algorithm that shows great promise—the variational quantum eigensolver algorithm (Peruzzo et al., 2014). This algorithm is essentially a "state preparation and then measure" algorithm leading to low-depth circuits governed primarily by the complexity of the state preparation. One starts from a single reference state (usually the Hartree-Fock state) and then creates a variational ansatz that depends on a set of variational parameters. There are several options for how to do this, discussed below. Extensions and generalizations of the algorithm to determine excited states, which are important in many biological processes, have been developed such as to maintain the low-depth characteristic so important for the ability to run on NISQ systems (Higgott et al., 2019; Nakanishi et al., 2019).

However the wavefunction has been prepared on the quantum computer, we next need to measure the expectation value of the Hamiltonian to complete the calculation. The Hamiltonian is a Hermitian operator rather than a unitary one, so it cannot be evaluated directly on the quantum computer. Instead, we break it up into a sum of its mutually commuting unitary components and evaluate the expectation value of each unitary—the total expectation value is found by accumulating the total of all of the terms. As the number of orbitals increases, the number of terms in the Hamiltonian also increases. To date, only quite simple molecules have been computed on available quantum hardware (usually with minimal bases). The first approach was hydrogen and other simple binary and tertiary molecules (Kandala et al., 2017). More recently, the more complex system H_2O in the STO-3G basis (Nam et al., 2020), has been handled.

Of course, this forms just the inner loop of the full variational calculation. One must now adjust the parameters in the variational wavefunction and repeat the whole process until the result converges with the minimum energy value. Because the data emerging from the quantum computer is noisy, this optimization problem requires complex algorithms on classical computers. The noise may even make it challenging to complete the calculation to the point where a true minimum can actually be located. The optimization problem may also suffer from "barren plateaus" which are large areas where the cost function gradient is extremely small. The sensitivity to noise can be reduced by calculating the derivative of how the energy changes when a variational parameter is changed from a matrix element measured directly on the quantum computer (Grimsley et al., 2019).

Because quantum computers have much slower clock cycles than classical computers, even with a quantum advantage for computing the results of a given measurement the quantum computations are expected to be slow. In addition, the parameters of a quantum computer often drift with time, creating additional issues associated with a changing accuracy for different expectation values over time. One may even need to correct for the drift over time or risk having data that is not accurate enough to be able to complete the outer loop of the variational cycle. Nevertheless, this approach remains the most promising approach available for now. Until we are able to perform extensive time evolution on a quantum system, it will remain the only viable strategy for quantum chemistry on NISQ era machines.

4.1.2. Phase Estimation

The quantum phase estimation (QPE) algorithm was invented by Kitaev (1995) and is closely related to the quantum Fourier transform. It provides an alternative to solving the traditional eigenvalue problem $\mathcal{H}|\psi\rangle = E|\psi\rangle$ on a quantum computer by transforming the problem to a unitary one and determining the phase λE arising from the application of $e^{i\lambda\mathcal{H}}$ to the eigenfunction as follows:

$$e^{i\lambda\mathcal{H}}|\psi\rangle = e^{i\lambda E}|\psi\rangle. \tag{2}$$

The λ parameter is introduced and a value chosen to ensure we can read the energy off without having the phase increase past 2π. We also need to measure enough binary digits in the number λE to have an accurate measure of the energy. In addition, to get the energy corresponding to a particular eigenstate with high probability, one must prepare an initial state that has high overlap with that eigenstate. This could then involve a synergy with the variational quantum eigensolver algorithms in the following way: since the variational state is an approximation, it should have a high overlap with the true ground state, allowing it to be a good choice for the initial state that is used for the phase-estimation algorithm.

There are many benefits to the phase-estimation approach. First, it will give us an accurate estimate of the ground-state energy, with the accuracy determined by how many binary digits representing the phase are computed on the quantum computer. Second, it projects onto the eigenstate it measures. This allows it to also be employed as a state-preparation protocol; measuring the ground-state energy also has the consequence of preparing the ground-state wavefunction directly on the quantum computer where it can then be employed for further quantum computations. For example, if the embedding strategy for self-energy embedding theory (described in detail in section 5) is used, one can compute the zero-temperature Green's function directly from the ground-state eigenfunction after it has been prepared by QPE.

The challenge with phase estimation is that it requires us to be able to accurately perform time evolution. This is currently beyond the scope of available hardware and most likely we will need to wait for large-scale fault-tolerant quantum computers to be available to be able to carry out such computations. Nevertheless, it is important to think through how one would work with such an algorithm now, to be ready when such hardware becomes available. Also, sparse embedding theories will allow for time evolution sooner, and possibly even on NISQ machines.

4.1.3. State Preparation

Both VQE and QPE require preparing an initial ansatz on the quantum computer. In other words, we start with an easy to initialize state, such as the Hartree-Fock state represented by $|\psi_0\rangle = |0\ldots0\rangle$, and apply operations to transform that state to a representation of the desired wave function. The complexity of this step can be non-trivial. The ADAPT-VQE approach (Grimsley et al., 2019) dynamically constructs the ansatz by iteratively choosing operators from a pool of available operators. Another approach uses a unitary coupled cluster ansatz. In standard coupled cluster, the wave function has the form

$$\psi = e^{\hat{T}}|\psi_0\rangle \tag{3}$$

where $\hat{T} = \hat{T}_1 + \hat{T}_2 + \hat{T}_3 + ...$ represents singles, doubles, triples, etc. excitations relative to the Hartree-Fock ground state reference function. In principle, one can work out as many terms as computationally feasible. In conventional coupled-cluster theory, the operator is not unitary, but can be made unitary by letting $\hat{T} \rightarrow \hat{T} - \hat{T}^{\dagger}$. Unitary coupled cluster is not practical computationally on classical computers, but is well-suited for quantum computers. A unitary singles and doubles coupled cluster approach will use a Trotterized form of the unitary coupled cluster ansatz (with only singles and doubles excitations in the exponent). Factorized forms of the unitary coupled cluster approach have also been considered, but usually, these approaches are not easily restricted to certain classes of excitations and might be better thought of within an ADAPT type methodology (Grimsley et al., 2019). There also are ansatzes that employ tensor-product-based wavefunctions (Cao et al., 2019).

5. QUANTUM COMPUTER EMBEDDING STRATEGIES

The molecular systems of interest in biological processes are complex as well as geometrically extensive. It is also known that some processes depend crucially on small differences in structure and their associated energy differences. The timescales of processes may also span several orders of magnitude, from femtoseconds for molecular vibrational changes to milliseconds for some electron transfer processes and conformation changes. The challenge is that highly accurate calculations are needed for these extreme systems.

In addition, biological processes happen at finite temperature in a liquid environment, as opposed to many chemical processes that can be understood by studying the gas phase or materials structures that are often analyzed in isolation and at absolute zero. This means that a statistical description is needed to describe the full process and to obtain accurate reaction rates. Entropy and free energy play a crucial role.

Hence, biological molecules appear to be an ideal application for the promised power of quantum computing. However, with noisy intermediate-scale quantum computers (NISQ), such calculations are currently out of reach. Even when more fully fault-tolerant quantum computers become available, it is likely that the complete statistical quantum description of realistic biomolecular systems and processes will require decomposition of the system into parts, with the parts of the molecule involving the most demanding calculations done on the quantum computer and the less expensive calculations involving the rest of the molecule done on a classical computer. This section describes several such approaches which will necessarily be hybrid quantum-classical algorithms.

5.1. Quantum Computing on Fragments

In computational chemistry for large systems, the fragment molecular orbital (FMO) method (Gordon et al., 2012; Zahariev and Gordon, 2012; Tanaka et al., 2014) was developed to solve the problem described above: namely that the system is too large to treat directly as a whole. In that case, the molecule is divided into fragments that can be chosen to (in some sense) contain atoms that interact strongly with each other, but less strongly with atoms in other fragments. First, standard methods are used to obtain an accurate description of the isolated fragments. Then, other methods, also standard, are used to describe the interaction between the fragments and the effect those interactions have on the internal structure and properties of the fragments. The result converges to a solution for the complete system of all interacting fragments with controllable accuracy. In addition, the method shows linear scaling for large systems.

To describe the biochemical systems, one can envision a similar approach. Now, however, instead of using different methodologies for different regions, one uses a classical computer to describe one region and a quantum processor for the region where the model can use the advantage offered by quantum computing. The self-consistency would typically be carried out on a classical computer. The approach is similar to the VQE method described in section 4.1.1 for quantum chemistry on quantum computers: Part of the computation is performed on the quantum computer, some information is extracted from that calculation and handed to a classical computer, which then performs the next part of the computation. That computation results in new values to be used for the next iteration on the quantum computer.

For a hybrid description of a complex biological systems, the different parts of the computation are not only different stages in an algorithm, but also describe different spatial regions of the system. Let us call the region described on the quantum computer "primary" and the region described on the classical computer, which most often surrounds the primary region in space, the "environment." We assume that the primary region fits in the quantum computer in the sense that it has sufficient qubits to represent both the quantum state in some encoding from fermions to spins (McArdle et al., 2020) and all the ancillary qubits necessary to execute the chosen algorithm.

It is necessary to choose quantum-mechanical methods to represent the states of both the primary and environment regions of the biological system so that the desired accuracy for the complete system can be achieved. It is not necessary that both regions are treated with the same method, as long as the physical description is consistent. The algorithm then inevitably requires that information is exchanged between the classical and quantum computers about the state description of the respective components. The classical computer can easily provide the necessary information to the quantum computer, which usually changes the state preparation on the quantum computer. However, as with VQE, obtaining accurate information about the state of the primary region as represented on the quantum computer can be challenging if the quantum state is complicated since there is no efficient way to directly access the entangled wavefunction stored on the qubits in the quantum processor. If a process of measurement needs to be called, then accurate calculations may require unacceptably large numbers of repeat runs of the program to obtain the required accuracy.

The general algorithm works as follows.

1. Specify a computational chemistry model for the environment region and initialize its state.
2. Specify a computational chemistry model for the primary region and prepare its state using the environment state parameters as needed.
3. Perform the algorithm to solve the computational chemistry model for the primary region on the quantum computer.
4. Extract the required information from the state of the primary region to perform the next iteration of the algorithm to converge the environment.
5. Perform the classical part of the algorithm for the primary region, using state information of the environment as needed.
6. Using information obtained by the classical part of the algorithm for the primary region, re-prepare the quantum computer for the quantum part of the algorithm for the primary region.
7. Repeat until the defined convergence criterion is met.

5.2. Sparse Green's Function Embedding Schemes

One of the challenges with accurate quantum chemistry calculations is that a large percentage of the correlation energy arises from the sum of many small contributions. This arises in part, because any standard orbital basis results in fairly full single-particle and two-particle interaction matrices. Hamiltonian evolution, or even variational methods require evaluating many, many terms. On a quantum computer, this leads to high-depth circuits, which become difficult to run on NISQ machines and may even be problematic on the expected fault-tolerant ones. One way around this problem is to transform the problem into a representation that is more sparse, or even to approximately force it into an extremely sparse representation. This is the idea behind the self-energy embedding theory (Tran et al., 2018).

Starting from an inexpensive classical calculation (such as Hartree-Fock plus MP2), one computes a representation for the self-energy of the full chemical system. Next, one determines the high-frequency moments of the self-energy. For the retarded Green's function, these moments of the self-energy are often determined by parameters in the Hamiltonian itself (the constant term is exactly determined from the Hartree-Fock approximation, the zeroth moment from the interaction, the first moment involves a few two-particle correlation functions, and so on). The strategy is then to construct an extremely sparse interaction for the effective Hamiltonian. This has the full single-particle contributions, but restricts the Coulomb interaction to on-site direct or exchange interactions only. These interaction terms are chosen to require that the low-order moments are preserved in the effective model. Then, one solves for the full self-energy of the effective model and then uses the effective self-energy as the self-energy for the full system. This approach guarantees that the low-energy moments of the final description of the molecule are exactly preserved. A self-consistency scheme is employed to update the approximation, as the moments depend on some expectation values which change as the Green's functions change with each iteration of the calculation. We also note that equality of low-order moments also implies that the two Green's functions agree exactly for short times.

The way we envision using this on a quantum computer for large molecules is to apply this approach to the strongly correlated core (or strongly correlated fragments) and one ends up with a much lower depth circuit for the time evolution because the Hamiltonian is so much sparser. This will allow more complex systems to be simulated earlier than possible with algorithms that include the full chemical complexity. The quantum computer simulates only the sparse Hamiltonian and determines the Green's function (or self-energy), which then is sent to the classical computer for the remainder of the algorithm. Even in the future, when fault-tolerant quantum computers become available, methods like the self-energy embedding theory will remain valuable as they can significantly streamline the number of operations needed to be run on the quantum computer.

6. CHALLENGES

There are a few challenges associated with the modeling of biochemical systems. The first challenge is the accuracy required to describe the structures and processes. The standard is 1 kcal/mol or 4 kJ/mol, which in atomic units used in quantum mechanics is equal to 27.2 meV or 1 mHartree. Given that the energies of large molecules relevant in biochemistry are in the thousand Hartree range, the energies need to be calculated with a precision of 6 to 8 digits, which corresponds to a single precision IEEE floating point number on a classical computer. There are a large number of integrals with weights that are small. The contribution from each integral is small, but the sum adds up to a non-negligible contribution to the total energy. Because these numbers are obtained by a large number of floating point operations in the classical part of the computation, the minimum precision needed to perform this classical part of the calculation with controlled rounding error is 15 digits, which corresponds to the double precision floating point number on classical computers. For large molecules, relevant to biology, the integral contributions are sorted and added with small numbers first to build larger numbers that can be meaningfully added together to avoid critical round-off errors. That means that the step in the hybrid quantum-classical algorithm where values must be measured from the state in the quantum processor, these results need to be obtained with the right precision. Because the standard deviation of statistical sampling with N trials goes like $1/\sqrt{N}$, the number of measurements for a given accuracy ε is $N = \varepsilon^{-2}$. A careful analysis is needed on what precision will be needed for the various terms to get acceptable accuracy for the total energies, because the required precision directly impacts the number of measurements that will be required, with a quadratic impact on total run time. Some further research to improve the algorithm for processing of the integrals will be needed to determine the minimum precision of each wave function component that must be combined with each integral to get the correct precision of the final result.

The second challenge in biochemical structure and process analysis and design is that the systems are at some finite temperature. That means a statistical description is essential. This has been taken into account for decades in the molecular dynamics simulations (Karplus and McCammon, 2002; Seabra et al., 2007; Salomon-Ferrer et al., 2013), with the method of replica exchange being one of the leading approaches (Roe et al., 2008).

However, chemical accuracy may not be sufficient. Decades of research to design drugs, enzymes, and catalysts has not been as successful as once hoped. A possible root cause is that chemical accuracy is insufficient to distinguish the competing mechanisms from each other, especially once the proper statistics at room temperature are taken into account. If 1 kcal/mo were adequate, scientists should have made more progress in identifying new mechanisms. To make the computations really insightful, it is likely that at least one and probably two or three orders of magnitude higher accuracy is required to generate new insights into drug, enzyme, and catalyst activities and reaction mechanisms.

These considerations make it clear that biochemical structures and processes are a fertile ground of problems to use and demonstrate the advantage of quantum computing over classical computing. It also shows the road to success will be difficult. But it promises to be wonderful journey!

7. CONCLUSION

This short review leaves us hopeful, but with many unanswered questions. It is clear that there are significant challenges that must be met before we can reap the benefits of quantum computers for biochemical applications. Nevertheless, due to the complexity involved in properly partitioning the sub units of these problems and then combining the results together, we need to start now to properly plan for how this will work. We can bring in ideas from a number of different areas where similar "divide and conquer" approaches have been tried and successfully completed. But the strategies that employ quantum co-processors to handle the most difficult parts of the calculations need to be properly thought out and structured so we can make rapid advances once the hardware is available. We did not map out a complete plan for how one can proceed. Instead, we described the different strategies that need to work together to achieve this goal. We are looking forward to seeing how everything comes together and how quantum computation will yield important and significant impacts on biochemistry.

AUTHOR CONTRIBUTIONS

CL wrote the biochemistry section. H-PC wrote the multi-scale simulation section. ED wrote the fragment method and its extension to quantum-classical hybrid application and the challenges. JF and BS wrote algorithms. All authors worked on the overall structure of the paper.

FUNDING

This material was based upon work supported by the National Science Foundation under Grant No. OMA-1936853. JF was also supported by the National Science Foundation under grant number CHE-1836497 and the McDevitt bequest at Georgetown. H-PC was also supported by DOE/BES DE-FG02-02ER45995. CL was also supported by the Bodor Professorship fund.

ACKNOWLEDGMENTS

The authors thank the participants of the Workshop on Applications of Quantum Computing to Biology and Chemistry held in conjunction with the Sanibel Symposium, February 21-22, 2020.

REFERENCES

Ahsan, M., and Senapati, S. (2019). Water plays a cocatalytic role in epoxide ring opening reaction in aspartate proteases: a QM/MM study. *J. Phys. Chem. B* 123, 7955–7964. doi: 10.1021/acs.jpcb.9b04575

Aryasetiawan, F., Imada, M., Georges, A., Kotliar, G., Biermann, S., and Lichtenstein, A. I. (2004). Frequency-dependent local interactions and low-energy effective models from electronic structure calculations. *Phys. Rev. B* 70:195104. doi: 10.1103/PhysRevB.70.195104

Aryasetiawan, F., Karlsson, K., Jepsen, O., and Schönberger, U. (2006). Calculations of Hubbard *u* from first-principles. *Phys. Rev. B* 74:125106. doi: 10.1103/PhysRevB.74.125106

Bakowies, D., and Thiel, W. (1996). Hybrid models for combined quantum mechanical and molecular mechanical approaches. *J. Phys. Chem*. 100, 10580–10594. doi: 10.1021/jp9536514

Bauer, B., Wecker, D., Millis, A. J., Hastings, M. B., and Troyer, M. (2016). Hybrid quantum-classical approach to correlated materials. *Phys. Rev. X* 6:031045. doi: 10.1103/PhysRevX.6.031045

Bender, A., Gardberg, A., Pereira, A., Johnson, T., Wu, Y., Grenningloh, R., et al. (2017). Ability of Bruton's tyrosine kinase inhibitors to sequester y551 and prevent phosphorylation determines potency for inhibition of fc receptor but not b-cell receptor signaling. *Mol. Pharmacol*. 91, 208–219. doi: 10.1124/mol.116.107037

Budarapu, P. R., Gracie, R., Yang, S. W., Zhuang, X. Y., and Rabczuk, T. (2014). Efficient coarse graining in multiscale modeling of fracture. *Theor. Appl. Fract. Mech*. 69, 126–143. doi: 10.1016/j.tafmec.2013.12.004

Cao, Y., Romero, J., Olson, J. P., Degroote, M., Johnson, P. D., Kieferova, M., et al. (2019). Quantum chemistry in the age of quantum computing. *Chem. Rev*. 119, 10856–10915. doi: 10.1021/acs.chemrev.8b00803

Chen, Z., Zang, J., Whetstine, J., Hong, X., Davrazou, F., Kutateladze, T., et al. (2006). Structural insights into histone demethylation by JMJD2 family members. *Cell* 125, 691–702. doi: 10.1016/j.cell.2006.04.024

Cohen, S., Graham, M., Lovrecz, G., Bache, N., Robinson, P., and Reddel, R. (2007). Protein composition of catalytically active human telomerase from immortal cells. *Science* 315, 1850–1853. doi: 10.1126/science.1138596

Cui, Q., Elstner, M., Kaxiras, E., Frauenheim, T., and Karplus, M. (2001). A QM/MM implementation of the self-consistent charge density functional tight binding (SCC-DFTB) method. *J. Phys. Chem. B* 105, 569–585. doi: 10.1021/jp0029109

Du, M.-H., Kolchin, A., and Cheng, H.-P. (2004). Hydrolysis of a two-membered silica ring on the amorphous silica surface. *J. Chem. Phys*. 120, 1044–1054. doi: 10.1063/1.1630026

Emani, P. S., Warrell, J., Anticevic, A., Bekiranov, S., Gandal, M., McConnell, M. J., et al. (2019). Quantum computing at the frontiers of biological sciences. *arXiv preprint arXiv:1911.07127v1*.

Feynman, R. (1982). Simulating physics with computers. *Int. J. Theor. Phys*. 21, 467–488. doi: 10.1007/BF02650179

Friesner, R. A., and Guallar, V. (2005). *Ab initio* quantum chemical and mixed quantum mechanics/molecular mechanics (QM/MM) methods for studying enzymatic catalysis. *Annu. Rev. Phys. Chem*. 56, 389–427. doi: 10.1146/annurev.physchem.55.091602.094410

Gao, H. J., and Klein, P. (1998). Numerical simulation of crack growth in an isotropic solid with randomized internal cohesive bonds. *J. Mech. Phys. Solids* 46, 187–218. doi: 10.1016/S0022-5096(97)00047-1

Gao, J. L., Amara, P., Alhambra, C., and Field, M. J. (1998). A generalized hybrid orbital (GHO) method for the treatment of boundary atoms in combined QM/MM calculations. *J. Phys. Chem. A* 102, 4714–4721. doi: 10.1021/jp9809890

Gao, J. L., and Xia, X. F. (1992). A priori evaluation of aqueous polarization effects through Monte Carlo QM-MM simulations. *Science* 258, 631–635. doi: 10.1126/science.1411573

Georges, A., Kotliar, G., Krauth, W., and Rozenberg, M. J. (1996). Dynamical mean-field theory of strongly correlated fermion systems and the limit of infinite dimensions. *Rev. Modern Phys*. 68, 13–125. doi: 10.1103/RevModPhys.68.13

Google AI Quantum and Collaborators (2020). Hartree-fock on a superconducting qubit quantum computer. *Science* 369, 1084–1089. doi: 10.1126/science.abb9811

Gordon, M. S., Fedorov, D. G., Pruitt, S. R., and Slipchenko, L. V. (2012). Fragmentation methods: a route to accurate calculations on large systems. *Chem. Rev*. 112, 632–672. doi: 10.1021/cr200093j

Grimsley, H. R., Economou, S. E., Barnes, E., and Mayhall, N. J. (2019). An adaptive variational algorithm for exact molecular simulations on a quantum computer. *Nat. Commun*. 10:3007. doi: 10.1038/s41467-019-10988-2

Gull, E., Millis, A. J., Lichtenstein, A. I., Rubtsov, A. N., Troyer, M., and Werner, P. (2011). Continuous-time Monte Carlo methods for quantum impurity models. *Rev. Mod. Phys*. 83, 349–404. doi: 10.1103/RevModPhys.83.349

Harvey, R. (2014). *Biochemistry, Lippincott's Illustrated Reviews*. Philadelphia, PA: Wolters Kluwer Health.

Higgott, O., Wang, D., and Brierley, S. (2019). Variational quantum computation of excited states. *Quantum* 3:156. doi: 10.22331/q-2019-07-01-156

Hoekstra, A., Chopard, B., and P., C. (2014). Multiscale modelling and simulation: a position paper. *Philos. Trans. R. Soc. A* 372:20130377. doi: 10.1098/rsta.2013.0377

Kandala, A., Mezzacapo, A., Temme, K., Takita, M., Brink, M., Chow, J. M., et al. (2017). Hardware-efficient variational quantum eigensolver for small molecules and quantum magnets. *Nature* 549, 242–246. doi: 10.1038/nature23879

Karplus, M., and McCammon, J. A. (2002). Molecular dynamics simulations of biomolecules. *Nat. Struct. Biol*. 9, 646–652. doi: 10.1038/nsb0902-646

Kitaev, A. (1995). Quantum measurements and the abelian stabilizer problem. *arXiv preprint arXiv:quant-ph/9511026*.

Laio, A., VandeVondele, J., and Rothlisberger, U. (2002). A Hamiltonian electrostatic coupling scheme for hybrid Car-Parrinello molecular dynamics simulations. *J. Chem. Phys*. 116, 6941–6947. doi: 10.1063/1.1462041

Leonov, I., Anisimov, V. I., and Vollhardt, D. (2015). Metal-insulator transition and lattice instability of paramagnetic $\mathrm{v_2o_3}$. *Phys. Rev. B* 91:195115. doi: 10.1103/PhysRevB.91.195115

Liu, B., Huang, Y., Jiang, H., Qu, S., and Hwang, K. C. (2004). The atomic-scale finite element method. *Comput. Methods Appl. Mech. Eng*. 193, 1849–1864. doi: 10.1016/j.cma.2003.12.037

McArdle, S., Endo, S., Aspuru-Guzik, A., Benjamin, S. C., and Yuan, X. (2020). Quantum computational chemistry. *Rev. Modern Phys*. 92:015003. doi: 10.1103/RevModPhys.92.015003

Mitchell, M., Gillis, A., Futahashi, M., Fujiwara, H., and Skordalakes, E. (2010). Structural basis for telomerase catalytic subunit TERT binding to RNA template and telomeric DNA. *Nat. Struct. Mol. Biol*. 17, 513–518. doi: 10.1038/nsmb.1777

Nakanishi, K. M., Mitarai, K., and Fujii, K. (2019). Subspace-search variational quantum eigensolver for excited states. *Phys. Rev. Res*. 1:033062. doi: 10.1103/PhysRevResearch.1.033062

Nam, Y., Chen, J.-S., Pisenti, N. C., Wright, K., Delaney, C., Maslov, D., et a;. (2020). Ground-state energy estimation of the water molecule on a trapped-ion quantum computer. *NPJ Quant. Inform*. 6:33. doi: 10.1038/s41534-020-0259-3

Ng, S., Kavanagh, K., McDonough, M., Butler, D., Pilka, E., Lienard, B., et al. (2007). Crystal structures of histone demethylase JMJD2A reveal basis for substrate specificity. *Nature* 448, 87–91. doi: 10.1038/nature05971

Pedersen, M., and Helin, K. (2010). Histone demethylases in development and disease. *Trends Cell Biol*. 20, 662–671. doi: 10.1016/j.tcb.2010.08.011

Peruzzo, A., McClean, J., Shadbolt, P., Yung, M.-H., Zhou, X.-Q., Love, P. J., et al. (2014). A variational eigenvalue solver on a photonic quantum processor. *Nat. Commun*. 5:4213. doi: 10.1038/ncomms5213

Petersilka, M., Gossmann, U. J., and Gross, E. K. U. (1996). Excitation energies from time-dependent density-functional theory. *Phys. Rev. Lett*. 76, 1212–1215. doi: 10.1103/PhysRevLett.76.1212

Piana, S., Bucher, D., Carloni, P., and Rothlisberger, U. (2004). Reaction mechanism of HIV-1 protease by hybrid carparrinello/classical md simulations. *J. Phys. Chem. B* 108, 11139–11149. doi: 10.1021/jp037651c

Reiher, M., Wiebe, N., Svore, K. M., Wecker, D., and Troyer, M. (2017). Elucidating reaction mechanisms on quantum computers. *Proc. Natl. Acad. Sci. U.S.A*. 114, 7555–7560. doi: 10.1073/pnas.1619152114

Roe, D., Okur, A., Simmerling, C., and WalkerX, R. (2008). *Tutorial A7: Replica Exchange*. Available online at: https://ambermd.org/tutorials/advanced/tutorial7/index.htm

Rountree, C. L., Kalia, R. K., Lidorikis, E., Nakano, A., Van Brutzel, L., and Vashishta, P. (2002). Atomistic aspects of crack propagation in brittle materials: multimillion atom molecular dynamics simulations. *Annu. Rev. Mater. Res*. 32, 377–400. doi: 10.1146/annurev.matsci.32.111201.142017

Rudd, R. E., and Broughton, J. Q. (2000). Concurrent coupling of length scales in solid state systems. *Phys. Status Solidi B Basic Res*. 217, 251–291. doi: 10.1002/(SICI)1521-3951(200001)217:1<251::AID-PSSB251>3.0.CO;2-A

Sadava, D., Hillis, D., Heller, C., and Berenbaum, M. (2011). *Life: The Science of Biology*. Sunderland, MA: Sinauer Associates Inc.

Salomon-Ferrer, R., Case, D. A., and Walker, R. C. (2013). An overview of the Amber biomolecular simulation package. *WIREs Comput. Mol. Sci*. 3, 198–210. doi: 10.1002/wcms.1121

Seabra, G. D. M., Walker, R. C., Elstner, M., Case, D. A., and Roitberg, A. E. (2007). Implementation of the SCC-DFTB method for hybrid QM/MM simulations within the Amber molecular dynamics package. *J. Phys. Chem. A* 111, 5655–5664. doi: 10.1021/jp070071l

Talebi, H., Silani, M., Bordas, S. P. A., Kerfriden, P., and Rabczuk, T. (2014). A computational library for multiscale modeling of material failure. *Comput. Mech*. 53, 1047–1071. doi: 10.1007/s00466-013-0948-2

Tanaka, S., Mochizuki, Y., Komeiji, Y., Okiyamac, Y., and Fukuzawace, K. (2014). Electron-correlated fragment-molecular-orbital calculations for biomolecular and nano systems. *Phys. Chem. Chem. Phys*. 16, 10310–10344. doi: 10.1039/C4CP00316K

Tran, L. N., Iskakov, S., and Zgid, D. (2018). Spin-unrestricted self-energy embedding theory. *J. Phys. Chem. Lett*. 9, 4444–4450. doi: 10.1021/acs.jpclett.8b01754

Veis, L., Visnak, J., Nishizawa, H., Nakai, H., and Pittner, J. (2016). Quantum chemistry beyond Born-Oppenheimer approximation on a quantum computer: a simulated phase estimation study. *Int. J. Quant. Chem*. 116, 1328–1336. doi: 10.1002/qua.25176

Vreven, T., Morokuma, K., Farkas, O., Schlegel, H. B., and Frisch, M. J. (2003). Geometry optimization with QM/MM, ONIOM, and other combined methods. I. Microiterations and constraints. *J. Comput. Chem*. 24, 760–769. doi: 10.1002/jcc.10156

Walsh, T. R., Wilson, M., and Sutton, A. P. (2000). Hydrolysis of the amorphous silica surface. II. Calculation of activation barriers and mechanisms. *J. Chem. Phys*. 113, 9191–9201. doi: 10.1063/1.1320057

Witzany, G. (2008). The viral origins of telomeres, telomerases and their important role in eukaryogenesis and genome maintenance. *Biosemiotics* 1, 191–206. doi: 10.1007/s12304-008-9018-0

Zahariev, F., and Gordon, M. S. (2012). Development of a combined quantum Monte Carlo-effective fragment molecular orbital method. *Mol. Phys*. 117, 1532–1540. doi: 10.1080/00268976.2019.1574363

Zhang, L., Staar, P., Kozhevnikov, A., Wang, Y. P., Trinastic, J., Schulthess, T., et al. (2019). DFT plus DMFT calculations of the complex band and tunneling behavior for the transition metal monoxides MnO, FeO, CoO, and NiO. *Phys. Rev. B* 100:035104. doi: 10.1103/PhysRevB.100.035104

Zheng, W., and Huang, Y., (2014). The chemistry and biology of the a-ketoglutarate-dependent histone N3-methyl-lysine demethylases. *Med. Chem. Commun.* 5:297. doi: 10.1039/c3md00325f

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Chemistry

ORIGINAL RESEARCH
published: 04 December 2020
doi: 10.3389/fchem.2020.606863

Benchmarking Adaptive Variational Quantum Eigensolvers

Daniel Claudino [1,2], *Jerimiah Wright* [1,3], *Alexander J. McCaskey* [1,2] *and Travis S. Humble* [1,3]*

[1] Quantum Computing Institute, Oak Ridge National Laboratory, Oak Ridge, TN, United States, [2] Computer Science and Mathematics, Oak Ridge National Laboratory, Oak Ridge, TN, United States, [3] Computational Sciences and Engineering, Oak Ridge National Laboratory, Oak Ridge, TN, United States

By design, the variational quantum eigensolver (VQE) strives to recover the lowest-energy eigenvalue of a given Hamiltonian by preparing quantum states guided by the variational principle. In practice, the prepared quantum state is indirectly assessed by the value of the associated energy. Novel adaptive derivative-assembled pseudo-trotter (ADAPT) ansatz approaches and recent formal advances now establish a clear connection between the theory of quantum chemistry and the quantum state ansatz used to solve the electronic structure problem. Here we benchmark the accuracy of VQE and ADAPT-VQE to calculate the electronic ground states and potential energy curves for a few selected diatomic molecules, namely H_2, NaH, and KH. Using numerical simulation, we find both methods provide good estimates of the energy and ground state, but only ADAPT-VQE proves to be robust to particularities in optimization methods. Another relevant finding is that gradient-based optimization is overall more economical and delivers superior performance than analogous simulations carried out with gradient-free optimizers. The results also identify small errors in the prepared state fidelity which show an increasing trend with molecular size.

Keywords: ADAPT-VQE, quantum computing, quantum chemistry, VQE, potential energy scan, state fidelity

OPEN ACCESS

Edited by:
Sugata Chowdhury,
National Institute of Standards and Technology (NIST), United States

Reviewed by:
Bryan Gard,
Georgia Tech Research Institute, United States
George Barron,
Virginia Tech, United States

****Correspondence:***
Travis S. Humble
humblets@ornl.gov

Specialty section:
This article was submitted to Physical Chemistry and Chemical Physics,
a section of the journal Frontiers in Chemistry

Received: *15 September 2020*
Accepted: *11 November 2020*
Published: *04 December 2020*

Citation:
Claudino D, Wright J, McCaskey AJ and Humble TS (2020) Benchmarking Adaptive Variational Quantum Eigensolvers. Front. Chem. 8:606863. doi: 10.3389/fchem.2020.606863

1. INTRODUCTION

Quantum mechanics naturally lends itself to the description of phenomena at the atomic and molecular scale, including problems of chemical interest, which has culminated in the field of research known as quantum chemistry. Despite the formal impediments to achieve exact, closed-form solutions to quantum chemistry problems, there is a wide array of possible approximations, such as coupled cluster (CC) theory (Shavitt and Bartlett, 2009), which have elevated quantum chemistry to good standing in the scientific community due to their reliability.

In practice, CC faces two main difficulties that have hindered a more widespread adoption. One is that most of the success it has garnered over the years is due to its superior performance in the weak electron correlation regime, for which single-reference (SR) CC remains unchallenged. This success is justified because many problems in chemistry, such as thermochemistry, can be adequately treated as being largely weakly correlated. Yet, many other problems of interest involve molecules and materials that do not comply with this assumption, and for these instances, SR-CC breaks down. Despite multi-reference (MR) CC being an active area of research (Jeziorski and Monkhorst, 1981), theoretical and computational challenges currently curb the applicability of MR-CC (Lyakh et al., 2012).

A second obstacle to a more extensive use of CC theory is its pronounced computational cost. Reliable SR-CC methods, such as the so-called "gold standard" of quantum chemistry, coupled cluster singles and doubles (and perturbative triples), aka CCSD(T) (Urban et al., 1985; Raghavachari et al., 1989; Watts et al., 1993), scale unfavorably with one-particle basis spanning the Hilbert space that houses the electronic wave function, which largely constrains the application of CCSD(T) to relatively small molecular systems. It is important to note that some of these limitations can be mitigated with methods such as configuration interaction (CI) in its MR formulation and the density matrix renormalization group (DMRG) which have in turn their own shortcomings, such as lack of size-extensivity and exactness contingent upon the dimensionality of the problem.

Concurrent with developments in CC theory has been the increase in performance of computing technologies, which broadens the reach of computational chemistry techniques. Presently, this trend is continuing with the adaptation of chemistry methods, including CC, to the new technology paradigm of quantum computing (Britt and Humble, 2017; Humble et al., 2019). Because of the shared foundation in quantum mechanics, one of the most immediate applications for quantum computers is quantum chemistry (McArdle et al., 2020). Recent advances have reformulated conventional problems in electronic structure for currently available quantum computing platforms (Cao et al., 2019). In particular, these efforts have led to a resurgence of the unitary coupled cluster (UCC) theory (Bartlett et al., 1989; Kutzelnigg, 1991; Taube and Bartlett, 2006; Romero et al., 2018), which can be employed in investigations where strong correlation is dominant. Quantum computing hardware appears to be well suited for building the states described by UCC, as this hardware can efficiently implement unitary operations to construct physical representations of the quantum state. Moreover, the intrinsic nature of the quantum computing logic can be exploited in order to propose new ansatze that, despite lacking a close connection to the underlying chemical intuition lent by UCC, are prone to a more efficient implementation, such as the so-called hardware efficient ansatz (Kandala et al., 2017).

It is in the context of noisy intermediate-scale quantum (NISQ) (Preskill, 2018) devices that the variational quantum eigensolver (VQE) (Peruzzo et al., 2014) has emerged as a promising method for testing the preparation and measurement of quantum states including those that represent the electronic eigenstates described by UCC (Quantum et al., 2020). Several variants of VQE are available (Parrish et al., 2019b; Chivilikhin et al., 2020), but all build on the variational principle from quantum mechanics, which constrains the quantum states that can satisfy the electronic eigenvalue problem (McClean et al., 2016). While the initial VQE proposal assumes a predefined ansatz, this constraint has been relaxed, opening the door to adaptive approaches (Grimsley et al., 2019; Ryabinkin et al., 2020), by which the preparable quantum states are driven by the problem at hand. In particular, the adaptive derivative-assembled pseudo-trotter (ADAPT) ansatz, which finds support on the recently coined "disentangled" UCC (Evangelista et al., 2019) and starts from an exact UCC representation of the electronic ground state to construct an approximate prepared state based on the dominant contributions. Early studies demonstrated this as a promising avenue for developing ansatze for specific molecules and constraints, such as highly accurate energetics or shallow circuits.

Here we benchmark adaptive VQE prescriptions, ADAPT-VQE in particular, by comparing the prepared quantum states with the conventional solutions obtained from exact diagonalization of the full configuration interaction Hamiltonian. We track the energy of the minimized expectation value as well as the fidelity of the corresponding prepared state using multiple ansatz, optimization methods, and molecular Hamiltonians. We calculate infidelity as a measure of error for the prepared quantum state relative to the expected, exact result from quantum chemistry using frozen-core Hamiltonians. Across these examples, we find that ADAPT-VQE is the more robust method due mainly to its performance with respect to optimization methods. While all methods lead to small errors as measured by the infidelity, these errors are found to grow with molecular size.

This presentation is structured as follows. In section 2, we provided an overview of the ingredients in the VQE approach relevant to our purposes, followed by a short exposition of the underpinnings of ADAPT-VQE (section 2.1) and a brief discussion on implementation of gradients and optimization in ADAPT-VQE (section 2.2). The computational details permeating the reported simulations are exposed in section 3. The main results are presented and discussed in section 4 and several conclusions are drawn in section 5.

2. VARIATIONAL QUANTUM EIGENSOLVER

This section serves to illustrate the pertinent fundamentals of the VQE algorithm and to motivate the following exposition of adaptive ansatz construction. We start by recalling the variational principle, which is at the heart of VQE, and given as

$$E \leq \min_{\Psi} \langle \Psi | \hat{H} | \Psi \rangle \tag{1}$$

where $|\Psi\rangle$ is normalized trial wave function for which Equation (1) becomes an equality when Ψ is constructed from a basis that spans the single-particle Hilbert space of all possible occupation numbers (the underlying Fock space) and the electronic Hamiltonian $\hat{H}$ for a molecular system is given as

$$\hat{H} = \sum_{pq} h_{pq} p^\dagger q + \sum_{pqrs} h_{pqrs} p^\dagger q^\dagger s r \tag{2}$$

The central problem in modern electronic structure theory is the description and quantification of the electron correlation from an un-entangled, mean-field wave function $|0\rangle$ whose preparation can be carried out in classical hardware in a timely fashion, e.g., Hartree-Fock (HF). In analogy with quantum chemistry, we can expect that there exists an operator that, once applied to $|0\rangle$, will account for the missing electron correlation. Bearing in mind that quantum computers manipulate quantum states in a well-defined Hilbert space, this configures a generic unitary operator $\hat{U}(\vec{\theta})$ whose main purpose is to build entanglement from an un-entangled reference function $|0\rangle$. The set of scalars $\vec{\theta}$ are parameters variationally varied in order to minimize the expectation value in Equation (1). With that, we recast Equation (1):

$$E \leq \min_{\vec{\theta}} \langle 0 | \hat{U}^{\dagger}(\vec{\theta}) \hat{H}_P \hat{U}(\vec{\theta}) | 0 \rangle \tag{3}$$

In order to ensure that Equation (3) meets the requirements of quantum hardware, the fermionic, second-quantized operators found in the formulation of electronic structure problem, such as those in Equation (2), are brought to a qubit (spin) representation, with the additional constraint of fermionic anti-symmetry. Our approach uses the Jordan-Wigner transformation (Jordan and Wigner, 1928), but others exist, and such a transformation yields $\hat{H}_P$ from $\hat{H}$, that is, the Hamiltonian in terms of strings of Pauli operators. Starting from the UCC ansatz, the unitary $\hat{U}(\vec{\theta})$ can be written as:

$$\hat{U}(\vec{\theta}) = \exp(\sum_k \theta_k (\hat{T}_k - \hat{T}_k^{\dagger})) = \exp(\sum_k \theta_k \tau_k) \tag{4}$$

with the $\hat{T}_k$ representing the usual cluster operators in CC theory and $\tau_k = \hat{T}_k - \hat{T}_k^{\dagger}$, ensuring the anti-Hermiticity of the operators, which is a necessary condition for their utilization in quantum computing.

Once in possession of all ingredients in Equation (3), the tasks of preparing the state $\hat{U}(\vec{\theta})|0\rangle$ and measuring the terms in $\hat{H}_P$ are delegated to the quantum hardware, and $\hat{U}(\vec{\theta})$ is varied variationally with the aid of a classical optimization routine until $\langle \hat{H}_P \rangle$ reaches its minimum, which is dependent on the chosen optimizer and is taken as a good approximation to the sought ground state energy. Due to the isomorphic property of the qubit mappings, $\langle \hat{H}_P \rangle = \langle \hat{H} \rangle$, yielding the lowest energy eigenvalue of the molecular Hamiltonian in Equation (2).

2.1. ADAPT-VQE

An important choice in the specification of the VQE method is the functional form of the ansatz $\hat{U}(\vec{\theta})$. Even for a relatively small Hilbert space, with a moderate number of cluster operators $\hat{T}_i$, the ansatz $\hat{U}(\vec{\theta})$ gives rise to a unitary that translates into multi-qubit gates and thus cannot be efficiently implemented in an actual quantum processor. Borrowing from the dynamics community, this can be alleviated by resorting to the Trotter-Suzuki decomposition, or Trotterization for short:

$$\exp(\sum_k \theta_k \tau_k) \approx \prod_k \exp(\theta_k \tau_k) \tag{5}$$

which here is limited to first-order.

The Adaptive Derivative-Assembled Pseudo-Trotter ansatz Variational Quantum Eigensolver (ADAPT-VQE) (Grimsley et al., 2019) takes advantage of Equation (5) to propose an iterative ansatz construction whereby only the perceived most relevant operator for energy lowering is added to the ansatz. A set of operators the algorithm can choose from needs to be provided, which in this work is comprised of the fermionic spin singlet adapted single and double excitations, borrowing from the usual UCCSD formulation, and subsequently mapped into the appropriate tensor products of Pauli operators via the Jordan-Wigner transformation. In principle, one could envision explicit enforcement or relaxation of other types of symmetry, and the effect of such choices on the performance of ADAPT is certainly a topic worth exploring. Moreover, the ADAPT algorithm has also recently been reported to perform well with other choices of operators, including a more economical pool of qubit operators (Tang et al., 2020), and has been applied to variational algorithms other than VQE (Zhu et al., 2020).

From a practical standpoint, at the i-th iteration of the algorithm, the energy gradient vector ($\mathcal{G}$) with respect to all $\{\theta_k\}$ in Equation (5) is computed from measurements on the circuit that prepares the state optimized in the previous iteration, represented by $|\psi_{i-1}\rangle$, with $|\psi_0\rangle = |0\rangle$. Labeling the energy at the current iteration E_i, we have:

$$\begin{aligned} \mathcal{G} &= \left(\frac{\partial E_i}{\partial \theta_1}, \ldots, \frac{\partial E_i}{\partial \theta_k}, \ldots, \frac{\partial E_i}{\partial \theta_N} \right) \\ \frac{\partial E_i}{\partial \theta_k} &= \langle \psi_{i-1} | [H, \tau_k] | \psi_{i-1} \rangle \end{aligned} \tag{6}$$

and if the norm of this vector falls below a set threshold, the algorithm is deemed converged and the ansatz-growing loop is exited. Otherwise, the operator associated with the largest absolute component of $\mathcal{G}$ is selected to increment the ansatz:

$$|\psi_i\rangle = e^{\theta_i \tau_i} |\psi_{i-1}\rangle, \quad \tau_i = \{\tau_k |\max |\langle [H, \tau_k] \rangle_{i-1}|\} \tag{7}$$

where $\langle [H, \tau_k] \rangle_{i-1}$ means this commutator was computed from observations in the circuit obtained from the previous iteration. With the selection of a new operator, the new ansatz is subject to the usual VQE routine and the corresponding energy minimum is obtained.

2.2. Gradient Estimate and Classical Optimization in ADAPT-VQE

From a quantum computing standpoint, ADAPT-VQE improves on VQE by potentially offering a more tractable circuit. However, this may come at the expense of a much larger number of measurements, as the evaluation of all $[H, \hat{A}_k]$ is performed at each iteration of the ADAPT loop, on top of the expected energy evaluations. In order to reduce the number of measurements associated with ADAPT-VQE simulations, adoption of a gradient estimate strategy can help improve the classical optimization step by reaching the sought minima with fewer calls to the hardware backend.

To motivate the discussion, we start by invoking the gradient expression as introduced in the original formulation of ADAPT-VQE:

$$\begin{aligned} \frac{\partial E}{\partial \theta_i} &= \langle \phi | \hat{H} \prod_{j=N}^{i+1} (e^{\theta_j \tau_j}) \tau_i \prod_{k=i}^{1} (e^{\theta_k \tau_k}) | 0 \rangle \\ &- \langle 0 | \prod_{k=1}^{i} (e^{-\theta_k \tau_k}) \tau_i \prod_{j=i+i}^{N} (-e^{\theta_j \tau_j}) \hat{H} | \phi \rangle \end{aligned} \tag{8}$$

where $\prod_i e^{\theta_i \tau_i} |0\rangle = |\phi\rangle$.

Equation (8) can be further simplified into a recursive formula:

$$\frac{\partial E}{\partial \theta_i} = \langle \phi | \left[\hat{H}, \prod_{j=N}^{i+1} (e^{\theta_j \tau_j}) \tau_i \prod_{j=i+1}^{N} (e^{-\theta_j \tau_j}) \right] | \phi \rangle \qquad (9)$$

Before moving further in the discussion regarding the use of gradients to support the classical optimizer, let us clarify a potential source of confusion. At a certain ADAPT-VQE iteration, the circuit previously optimized is implemented to prepare the state from which the current iteration builds upon. The gradient vector $\mathcal{G}$ is then computed upon the necessary measurements for all τ_k in the chosen operator pool (Equation 6), and the operator that has the largest commutator (in absolute value) is selected. And this the extent to which the gradient is employed at this stage. On the other hand, we now have a new ansatz, which is composed of all previously added operators that enable preparation of $|\psi_{i-1}\rangle$, along with the newly added operator from Equation (7). Each of these operators have a corresponding variational parameter, which in the following VQE step need to be re-optimized. It is in this optimization that we would employ the gradients as written in Equations (8) and (9), and whose magnitude needs to be minimized in order to signal the finding of an extremum (minimum in this case). For an operator pool containing N elements, at each ADAPT-VQE iteration, all N elements of $\mathcal{G}$ need to be evaluated, but the magnitude of this vector is not directly minimized by varying the circuit parameters, only indirectly by the addition of enough operators in the ansatz. On the other hand, for optimization purposes, at the i-th iteration, only i gradient elements are considered, and the search for the energy minimum is guided by the minimization of the magnitude of this i-th dimensional gradient vector. Finally, another crucial point worth pointing out is that the commutators in Equation (6) are equivalent to those in Equation (9) only for the operator most recently added, i.e., τ_i in Equation (7).

For the purposes of an economical quantum resource utilization, it is desirable to deploy only one circuit to be used in both energy and gradient estimates (the same circuit is implemented many times, one for each term in the Hamiltonian/gradient). Even though the recursive formula in Equation (9) could, in principle, satisfy this requirement, this commutator cannot be measured (Mitarai et al., 2018). As originally proposed, the gradient is no longer given in an expectation value form, requiring an auxiliary state to be prepared via introduction of ancilla qubits, which deviates from our requirement of saving quantum resources. For that reason, we resort to numerical finite differences as means of carrying out gradient-based optimizations in the current work.

In terms of resource estimation, for a circuit depth of $\mathcal{O}(N)$, forward or backward finite differences are akin to introducing a single $R_z(h)$, where h is the step size, leading the a circuit depth of $\mathcal{O}(N + 1)$, while the use of central differences, thus, has circuit depth of $\mathcal{O}[2(N + 1)]$, the former being used here due to its superior convergence properties. This is the cost incurred in the numerical gradient estimate for each parameter being optimized and a detailed discussion is provided in section 4.4. Such an estimate may be improved with strategies such as the quantum natural gradient (Stokes et al., 2020) or exploiting partial tomography (Parrish et al., 2019a). These ramifications are worthy of a separate study, and will not be further investigated here.

3. COMPUTATIONAL DETAILS

The quantum simulations detailed in this manuscript were performed using the VQE and ADAPT-VQE algorithms and numerical gradient strategies as implemented in the XACC hybrid quantum-classical computing framework (McCaskey et al., 2018b, 2020), with the latter algorithm leveraging a convergence criterion of $||\mathcal{G}|| \leq 10^{-2}$. We emphasize that this parameter can be of substantial impact on the results, as it controls the size of the obtained ansatz. In light of the findings in Grimsley et al. (2019), the adopted value in this paper is believed to strike a satisfactory balance between accuracy and circuit depth. The resulting circuits were simulated via the TNQVM (tensor-network quantum virtual machine) (McCaskey et al., 2018a) XACC simulation backend and employed a noiseless, matrix product state (MPS) wave function decomposition for the quantum circuit with the aid of the ITensor library (Fishman et al., 2020). XACC provides other simulation backends, as well as physical backends targeting QPUs from IBM and Rigetti. For the size of the problems studied in this work, there may not be perceived benefits from choosing TNQVM over other XACC simulation backends like Aer (Abraham et al., 2019) or QPP (Gheorghiu, 2018). TNQVM is expected to be advantageous over other simulation approaches for problems requiring more qubits (McCaskey et al., 2018a), but we leave this to future work and do not investigate it here.

The COBYLA (Powell, 1994) algorithm was used as a gradient-free optimizer, while gradient-based optimizations were carried out with the L-BFGS algorithm (Nocedal, 1980; Liu and Nocedal, 1989), with all parameters being initialized at 0 at each optimization cycle for both optimizers. Other approaches have been reported in the literature, such as random initialization (Grimsley et al., 2020), or as in the original implementation of ADAPT-VQE (Grimsley et al., 2019) where the new parameter is initialized at 0, while the previous parameters are initialized from the optimal values obtained in the previous ADAPT iteration. XACC offers both optimizers via its interface with `NLOpt` (Johnson).

The potential energy curves (PEC) of NaH and KH, were generated by imposing the frozen-core approximation, reducing the number of configurations to only those arising from one σ orbital and its σ^* counterpart, that is, a two electrons in two orbitals complete active space [CAS(2,2)] problem. The one- and two-electron integrals necessary for the construction of the Hamiltonians and the corresponding references CAS energies were obtained with PySCF (Sun et al., 2017), with all calculations performed with the STO-3G basis set (Hehre et al., 1969, 1970; Pietro et al., 1980).

The quality of the output circuits in preparing the desired states is assessed via the fidelities computed with respect to the ground state full configuration interaction (FCI) wave function. This corresponds to the lowest energy eigenvector from exact diagonalization in the 2^N Hilbert space, with orbital occupation determined by the number of electrons. In possession of the circuits from the quantum simulations, the respective state vector representation is obtained using the XACC interface to the Qiskit Aer simulator (Abraham et al., 2019).

4. RESULTS AND DISCUSSION

Typically, the quality of the state obtained from the variational optimization of the gate parameters is probed indirectly by comparison of the computed energies with trustworthy references values or the exact lowest energy eigenvalue whenever computationally feasible. Thus, we start by investigating the energy profile along the atomic displacement, and subsequently contrast these findings with the analysis of the appropriateness of the corresponding states via evaluated fidelities relative to the vector corresponding to the lowest eigenvalue in the active space.

4.1. Potential Energies Curves

We start investigating the behavior of the energy by studying the H_2 molecule. This example has been extensively approached in quantum computing, and hardly poses any difficulty, at least from the standpoint of numerical simulations, as opposed to deployment to actual hardware. However, it serves as a baseline for the following discussion, as the orbital spaces in the other molecules are reduced to an active space with the goal of resembling the H_2 molecule. Results with the VQE and ADAPT-VQE ansatze are plotted in **Figure 1**, along with FCI results.

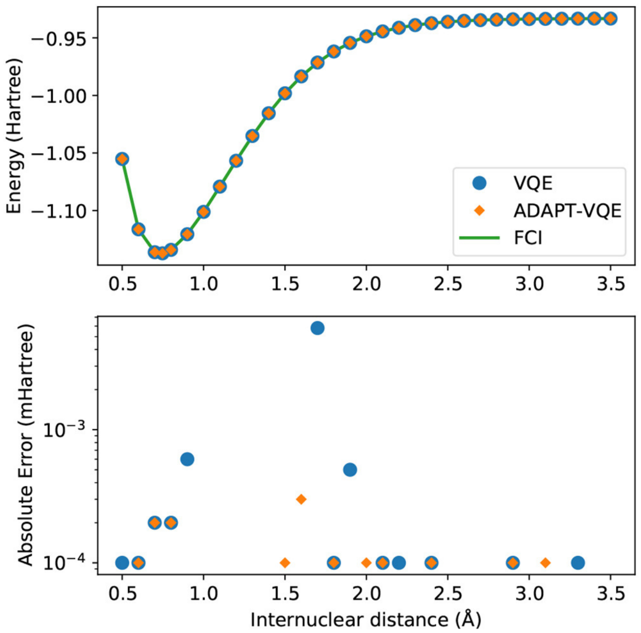


FIGURE 1 | (Top) Potential energy curves of H_2 computed with the STO-3G basis set for FCI (green solid line), VQE (blue circles), and ADAPT-VQE (orange diamonds) with the COBYLA optimizer. **(Bottom)** Absolute error in the minimized energy for VQE (blue) and ADAPT-VQE (orange) relative to the FCI reference value.

Unsurprisingly, there is a remarkable agreement between simulated and exact values, both qualitatively and quantitatively. Absolute errors from FCI are found in the sub-miliHartree range throughout the energy scan, and with either choice of ansatz, the observed errors would be inconsequential when taking into account the scale of the errors introduced by noise in the operation of quantum devices. The impression that some points are "missing" from the bottom plot of **Figure 1** is explained by these values being numerically identical to the FCI values (to seven decimal places), hence not being plotted in the logarithmic scale.

The results from the potential energy curve from simulations on the NaH molecule are presented in **Figure 2**.

Visual inspection of the top plot reveals that the choice between the two ansatze being considered here yield energies that track one another very well, but because of the energy scale of this plot, it begs a closer look. The bottom plot displays the absolute errors between VQE and ADAPT-VQE with respect to FCI. The errors here are still within chemical accuracy (<1 kcal/mol), and are unlikely to be of much relevance in the total error if such simulations are executed in a quantum computer. However, there is a clear trend of increase in the magnitude of the computed deviations when compared to the hydrogen molecule, whose results are in **Figure 1**.

In **Figure 3**, we again observe some of the patterns that follow from the analysis of **Figures 1**, **2**. The energy scale here is much too large to able to reveal relatively minor inadequacies, even though qualitative discrepancies, such as those arising from symmetry breaking or the crossing of lines of different states, would be evident had they been present. The bottom plot, exhibiting the energy differences from FCI, offers a more reliable evidence, allowing us to infer that ADAPT-VQE is overall superior, with smaller errors for the vast majority of points (the exception being 1.4 Å). Perhaps more importantly, we observe a general trend of the points from simulations with the plain VQE ansatz approaching the 1 mHartree, with the distances of 2.9 and 3.9 Å now found more than 1 kcal/mol above the respective FCI energy.

4.2. Optimization Strategies

The potential energy curves presented and discussed in section 4.1 are based upon gradient-free optimization carried out with the COBYLA optimizer. We report that analogous simulations were performed with the Nelder-Mead optimizer, which is also a gradient-free alternative, but preliminary investigations pointed to COBYLA being a superior choice, at least for the chosen molecules. To contrast the performance of gradient-free optimization in the current context, we use the L-BFGS optimizer for parameter update, as implemented in `NLOpt`, with gradient estimated via central numerical finite differences. To assess the relative performance of these two approaches as the bond in the current diatomic molecules is stretched, we plot the difference

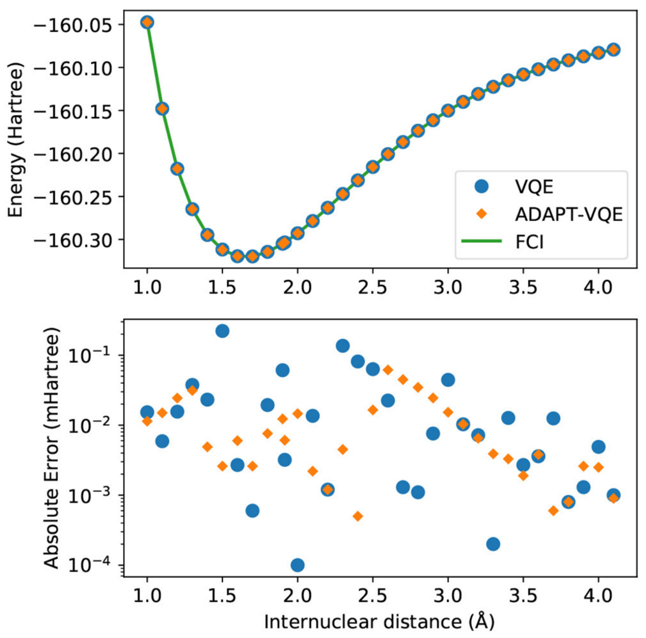


FIGURE 2 | (Top) Potential energy curves of NaH computed with the STO-3G basis set for FCI (green solid line), VQE (blue circles), and ADAPT-VQE (orange diamonds) with the COBYLA optimizer. **(Bottom)** Absolute error in the minimized energy for VQE (blue) and ADAPT-VQE (orange) relative to the FCI reference value.

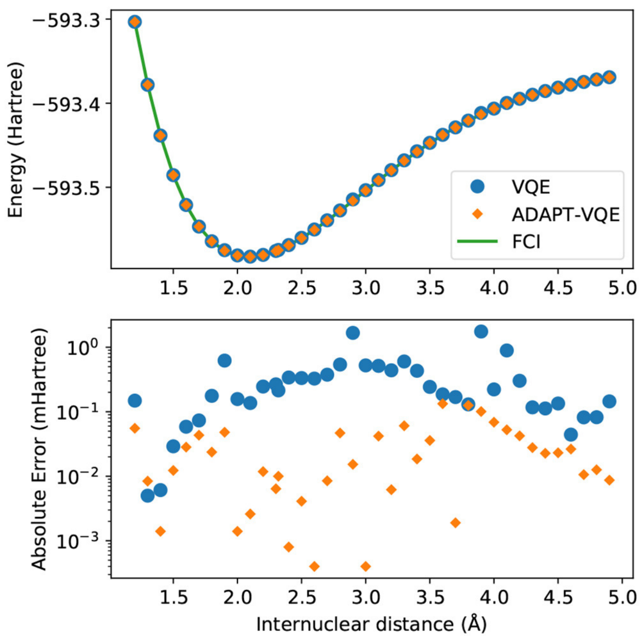


FIGURE 3 | (Top) Potential energy curves of KH computed with the STO-3G basis set for FCI (green solid line), VQE (blue circles), and ADAPT-VQE (orange diamonds) with the COBYLA optimizer. **(Bottom)** Absolute error in the minimized energy for VQE (blue) and ADAPT-VQE (orange) relative to the FCI reference value.

between energies obtained with the COBYLA optimizer and those with L-BFGS+finite differences, that is, E(COBYLA) − E(L-BFGS). That way, positive energy differences indicate there is an improvement by turning to a gradient-based optimization, while the opposite signals that the current gradient-free method reached a lower energy.

We observe compatible energies for the H_2 case, regardless of the underlying optimization strategy, for the entirety of **Figure 1**. In order to maintain consistency, we plot the energy difference between the two optimization prescriptions in a miliHartree scale, and the spike in E(COBYLA) − E(L-BFGS) in 1.7Å, when rationalized with the scale in mind, shows a deviation in the μHartree range. Due to the presence of all the many-body operators necessary for exactness (Evangelista et al., 2019), we expect and in fact observe results on par with the numerical precision imposed by the employed optimizers (10^{-6} Hartree in relative energy).

While most of the PEC for H_2 showed no major dependence on the adopted optimization procedure, according to **Figure 4**, the picture is significantly different in the case of NaH, as portrayed in **Figure 5**. Even though the values for E(COBYLA) − E(L-BFGS) are still rather small, in the sub-miliHartree range, noticeable differences are more frequent here. Albeit of μHartree in magnitude, we also observe cases where COBYLA provides a lower energy than L-BFGS, most notably for ADAPT-VQE in the 1.4 and 2.5 Å interatomic distances. On the other hand, in an overall assessment of the performance between VQE and ADAPT-VQE, the latter displays a more pronounced insensitivity with respect to the choice of optimization scheme.

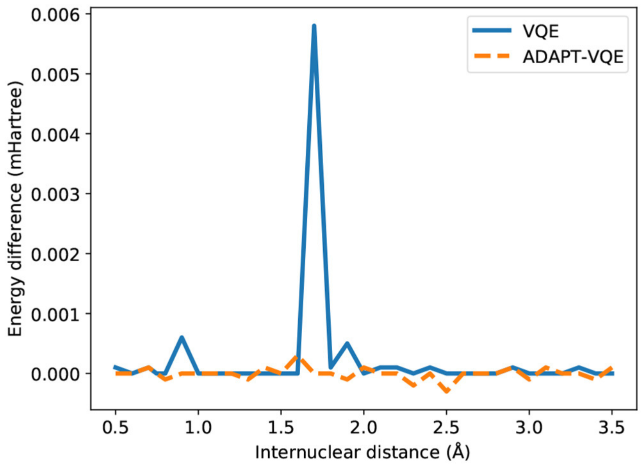


FIGURE 4 | Difference between the energies from COBYLA and L-BFGS optimization with central finite differences for the H_2 potential energy curve.

An even more drastic contrast is found from inspection of **Figure 6**, where E(COBYLA) − E(L-BFGS) are plotted for the KH molecule. Some of the qualitative assertions pointed out in **Figure 5** hold, namely that the performance of VQE is much more influenced by the choice of optimization strategy than ADAPT-VQE. Not only that, but ADAPT-VQE is

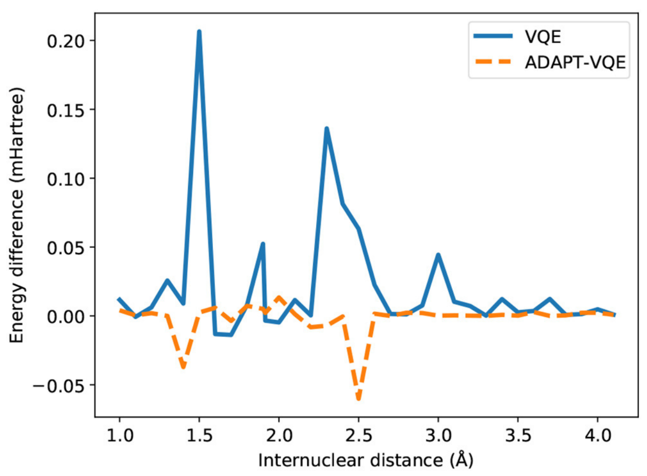


FIGURE 5 | Difference between the energies from COBYLA and L-BFGS optimization with central finite differences for the NaH potential energy curve.

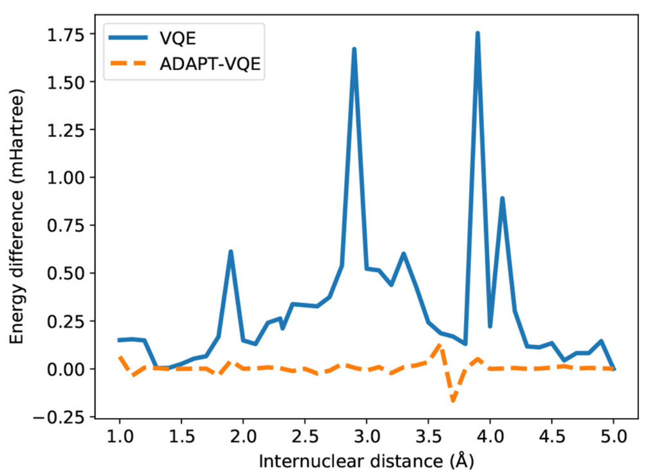


FIGURE 6 | Difference between the energies from COBYLA and L-BFGS optimization with central finite differences for the KH potential energy curve.

largely unaffected by employed optimizer, at least between the two alternatives in consideration. Here again, the differences seen for VQE correlated well with the deviations from FCI reported in **Figure 3**, further corroborating the claim that a gradient-based optimization, given the current conditions, is a more robust for approaching the lowest energy eigenvalue of molecular Hamiltonians.

4.3. State Fidelities

As previously stated, energy values can be used as valuable metric of the adequacy of a given set of variational parameters and trial state. However, the energy alone may not be indicative of the quality of the corresponding state and even acceptable energy values do not guarantee equally satisfactory values for other properties. The usual electronic Hamiltonian, as shown in Equation (2), transforms as the most symmetric irreducible representation for a given point group, therefore yielding the same energy in the case of degenerate states. Other operators, however, such as the terms in the multipole expansion of the electric potential, do not display this feature, meaning that degenerate states may yield different expectation values for such operators.

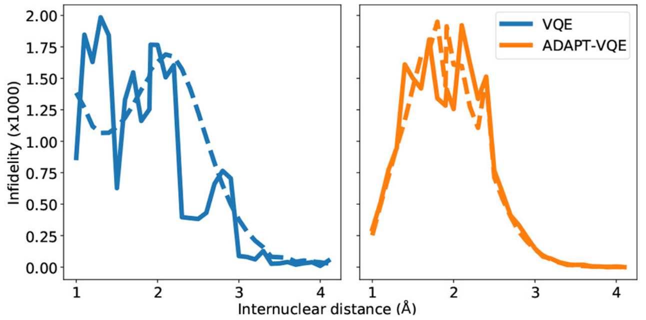


FIGURE 7 | State infidelities for VQE and ADAPT-VQE using COBYLA (solid line) and L-BFGS (dashed line) optimization with central finite differences for the NaH potential energy curve.

In order to examine the state prepared by the two circuit approaches considered here, we compute their "infidelities" with respect to the exact FCI state within the aforementioned active spaces, which is mathematically represented by $1 - |\langle \Psi_{\mathrm{FCI}}|\hat{U}(\vec{\theta})|0\rangle|$, where $\vec{\theta}$ here are the set of optimal values also utilized for the energy computations in sections 4.1 and 4.2. We acknowledge that, while this provides direct inroads in the state being output at completion of the state preparation, it cannot be experimentally realized. However, in the case of moderately sized molecules for which the exact diagonalization of the Hamiltonian is feasible, this can provide valuable insights.

The energy differences discussed in the case of the hydrogen molecule in sections 4.1 and 4.2 are quite small when considering the magnitude of the other potential sources of error that can arise in the presence of noise, either through a model or in the operation of an actual quantum device. Due to the simplicity of the electronic structure of this molecule the state prepared according to the two ansatze construction prescriptions investigated here yield infidelities that are below the numerical thresholds employed here, and certainly would be unnoticeable for realistic purposes. Because they offer little insight, we abstain from plotting the infidelity results for this molecule here.

Before delving into the particularities of each curve in **Figure 7**, we bring the reader's attention to the scale of the plots, signaling a remarkable agreement between the state prepared and the one expected (FCI). It should come as no surprise that the largest infidelities are found in the vicinity of the Coulson-Fischer point, the most demanding region in the energy landscape, and subsequently approach zero as the atoms are moved far apart. The infidelities for the VQE ansatz follow a smooth progression when employed in conjunction with the gradient-based optimizer L-BFGS, whereas the same is not true for the other combinations of ansatze and optimization. This is

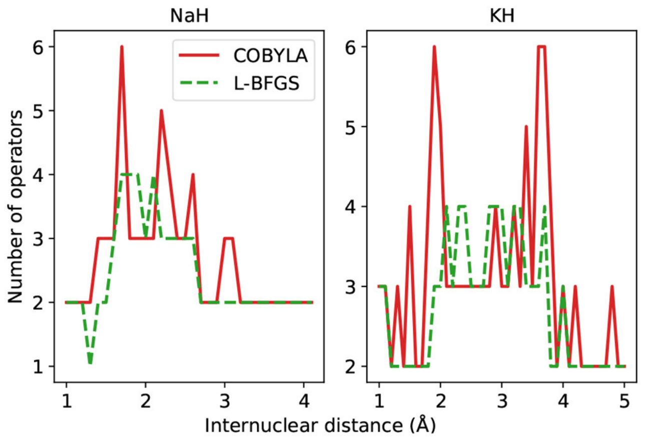


FIGURE 8 | Number of operators in the ADAPT-VQE ansatze using the COBYLA (solid line) and L-BFGS (dashed line) optimizers. The corresponding VQE ansatz has two operators.

likely a compound effect, explained by the former being a fixed circuit, where only the associate $\vec{\theta}$ changes throughout the energy scan. The latter, however, can assume a different composition, changing according to the demands of the electronic structure at each bond length. This works along with the fact that gradient-based optimization, at least in the current study, provides a tighter, more reliable solution. For the NaH and KH cases, we plot the number of operators in the ansatz in **Figure 8**.

Once again, there is a clear advantage in turning to gradient-based optimization, as it renders ansatze with fewer operators. For some internuclear distances, the ADAPT-VQE ansatze, even when optimized with L-BFGS, contain more operators than the corresponding VQE ansatz. This is not necessarily in contradiction with the some of the findings from Grimsley et al. (2019) because those results were obtained for different molecules and using different optimization implementations. Yet, we would expect that when comparing against a larger VQE problem, such as those investigated in that paper, we would see similar trends. We also speculate that another variable that can contribute to the observed behavior is the tolerance that controls how tight the optimization should be. Because we are using a default 10^{-6} threshold in relative energy as the tolerance and there is no clear connection between the quality of the energies and the respective prepared states, the absolute energies values may fall in a scale that may have a small, but non-negligible effect on the fidelities, which is also evidence of the effect it can have in the output state, further corroborated by the number of operators found in the respective ansatze, yet not enough to alter any of the main conclusions drawn from the results presented here.

Many of the main inferences from the analysis of the **Figure 7** hold for the KH molecule, whose infidelities are shown in **Figure 9**. Firstly, the infidelities, though still quite small, are about an order of magnitude larger. The smoothness and overall profile observed for the VQE UCCSD is retained, but the behavior of the ADAPT-VQE infidelities is much more erratic. Secondly, while the ADAPT-VQE ansatz for NaH around the Coulson-Fischer point is mostly the same, but the larger number of variational parameters make it more vulnerable to the optimization inconsistencies discussed above, here the large oscillations are due to ansatze of alternating operator compositions. Because the ADAPT-VQE convergence criterion depends upon a fixed numerical threshold, sometimes the ansatz at a given iteration may already be close to convergence, but still not quite below the gradient norm threshold, and upon the addition of an extra operator, the state may be improved significantly in the scale of the plots seen in this section.

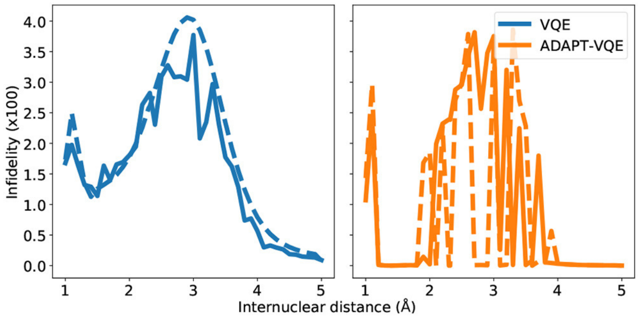


FIGURE 9 | State infidelities for VQE and ADAPT-VQE using COBYLA (solid line) and L-BFGS (dashed line) optimization with central finite differences for the KH potential energy curve.

4.4. Resource Estimation

One of the main motivations behind the present work is to serve as the baseline for following studies focusing on the investigation of the electronic structure of molecules carried out in NISQ devices. With this in mind, it is important to develop some intuition on the resource demands involved in such tasks.

First we analyze the circuit proposed by VQE and ADAPT-VQE to prepare the states whose energies and fidelities were shown in sections 4.1–4.3 in terms of total gate count and circuit depth, plotted in **Figure 10**.

Let us first compare the ADAPT-VQE results on the basis of the two optimizers. As we move from H_2 to NaH and KH, we see a more intricate picture of how these optimizers impact the final circuit. Qualitatively, L-BFGS has an overall advantage as it provides circuits that are shallower and with fewer gates. While there are a few points along the potential energy scans where the circuits generated based on L-BFGS are not as efficient as those from a COBYLA optimization, the scales of the plots are determined solely by the latter. We noticed that in several points, the simulations with the COBYLA optimizer would produce states with two instances of the same operator adjacent to each other. If the actual minimum value had been achieved in a certain iteration of ADAPT-VQE, the commutator of the same operator in the next iteration would have been zero. Because this procedure is accomplished numerically, the magnitude of this commutator is related to how close the determined minimum is from the actual one. It turns out that the default threshold in relative energy (10^{-6}) is found not to be stringent enough, which incurs a commutator whose deviation from the expected

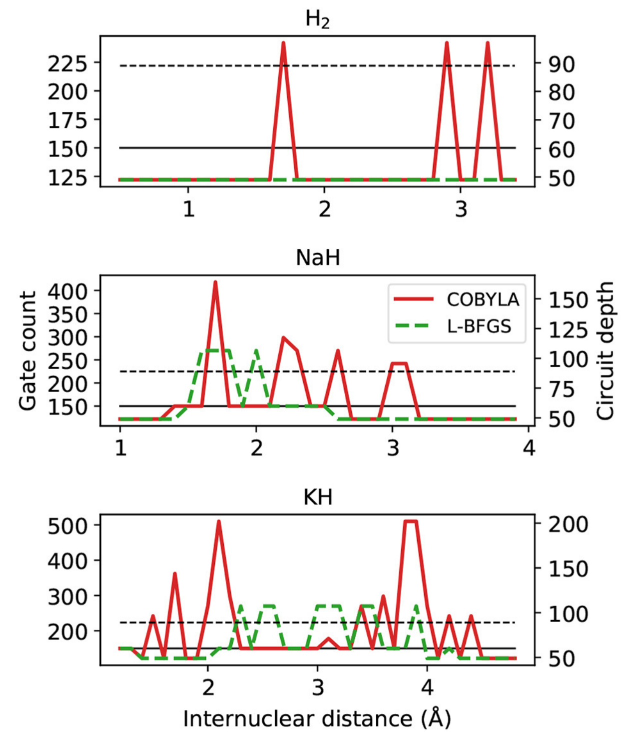


FIGURE 10 | Gate counts (left axis) and circuit depth (right axis) from ADAPT-VQE circuits optimized with the COBYLA and L-BFGS optimizers. The black solid and dashed lines are the gate counts and circuit depths from the VQE ansatz, respectively.

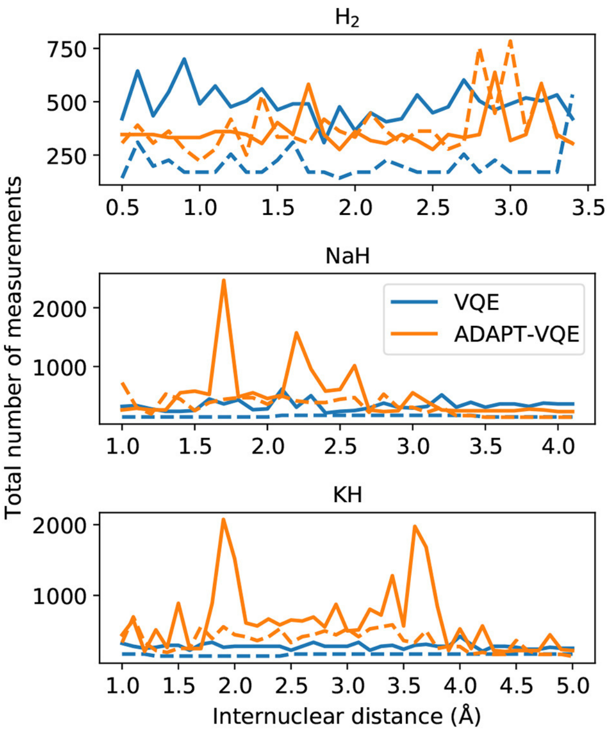


FIGURE 11 | Total number of measurements for final energy evaluation for the VQE and ADAPT-VQE ansatze using the COBYLA (solid line) and L-BFGS (dashed line) optimizers.

zero is non-negligible, resulting in the same operator being added in successive iterations. Another factor that accounts for the displayed circuit figures is the fixed gradient norm threshold in ADAPT-VQE. In some iterations, this quantity is above, but already quite close to the pre-defined 10^{-2}, and one extra iteration is performed, with only marginal energy improvement. To illustrate this, the ADAPT-VQE simulation for NaH with internuclear separation of 1.8Å converges to ansatz with three operators, with $E = -160.3146751$ Hartree and $||\mathcal{G}|| = 0.001$. Had the ADAPT cycle been stopped in the second iteration, we would have $||\mathcal{G}|| = 0.013$, with $E = -160.3146492$ Hartree, that is, the energy improvement was in the μHartree range, yet at the expense of a deeper circuit, which calls for a more flexible operator selection in ADAPT-VQE.

These resource estimation parameters in **Figure 10** are comparable between the two ansatz strategies. In general terms, the circuits optimized upon L-BFGS are more affordable than the corresponding VQE ones, while using COBYLA tends to yields circuits that are deeper and need to implement more gates. We bring attention to the fact that there is not a one-to-one correspondence between the present analysis and that in the Figures 2C,F,I in Grimsley et al. (2019). This is because the latter refers to the number of parameters/operators in the ansatz. A circuit with more parameters/operators does not readily translate into a more complex circuit, which depends on the number of qubits in a given operator and the operator locality and placement. This means these results are not at odds with what was previously reported, which were obtained for a distinct set of molecules, but can be seen as complementary.

Another important metric when estimating the necessary resources for implementation and deployment of the simulations discussed here is the number of measurements. To complement the end of the last paragraph, it is important to mention that in this context the rationalization in terms of number of operators increases in relevance. In **Figure 11**, we plot the total number of measurements to achieve the results reported in section 4.1.

As pointed out in Grimsley et al. (2019), the ansatz put forth by ADAPT-VQE offers a trade-off between circuit depth and number of measurements. We can readily confirm by visual inspection of **Figure 11** that ADAPT-VQE incurs a much larger number of measurements. These figures account for all measurements involved in computing the commutators in Equation (6), the energy evaluations at each optimization iteration, and the computations necessary to minimize the gradients when L-BFGS is employed. The measurement burden in ADAPT-VQE reported here can be partially alleviated by employing a better parameter initialization, such as starting the VQE optimization at each iteration with the previously optimized parameters and initializing just the newly added parameter at zero. This demand is also expected to be greatly relieved by resorting to a different set of operators, such as those introduced in the qubit-ADAPT-VQE variant (Tang et al., 2020), which can still span the underlying Hilbert space, yet with linear growth in the number of qubits. This approach would require much fewer commutator computations at each iteration, but would likely be of noticeable advantage for operator pools larger than those in question here. These results are also contingent upon the choice of optimizer, and there may exist better suited choices than those investigated here. Yet, we do not believe this would dramatically change the overall qualitative picture drawn in **Figure 11**.

Another key outcome from the analysis of **Figure 11** is the fact that, even though the gradient computation with L-BFGS requires more measurements per iteration, it is overall much more economical than the gradient-free optimization, represented here by COBYLA. This furthers strengthens the case for gradient-based optimization in VQE, as it not only results in smaller errors/better convergence with respect to the sought ground state, but it is also much less demanding from a resource standpoint.

5. CONCLUSION

For a broader adoption of adaptive methods for ansatz construction in the realm of quantum chemistry, and perhaps, for many-body methods in general, many aspects still needs to be explored and their underpinnings better understood. This work provides a contribution toward this goal by showing a comprehensive study of potential energy curves of a selection of molecules of the general formula XH (X = H, Na, K). Despite their simplicity, they serve to shed light on some of the mentioned characteristics, and deliver a baseline for feasible studies involving actual quantum hardware.

Even a relatively conservative gradient norm threshold of 10^{-2} in ADAPT-VQE is sufficient to provide overall better energetics than corresponding fixed ansatz approach embodied by the ordinary VQE, which is in agreement with the initial ADAPT-VQE proposal. Due to its iterative nature, ADAPT-VQE has an extra layer of tunability which can be controlled via the threshold on $||\mathcal{G}||$. This means that the errors observed with ADAPT-VQE might have been reduced had $||\mathcal{G}||$ been made tighter, which could in turn increase the depth of the circuits, and even having to cope with more necessary measurements than those of UCCSD, as suggested with $||\mathcal{G}|| = 10^{-3}$ in Figure 2i by Grimsley et al. (2019) However, upon a simple choice of gradient strategy motivated by the constraints of quantum hardware, we report that ADAPT-VQE is fairly resilient with respect to the employed optimization strategy and that encouraging improvements in performance by adopting a gradient-based approach in the search of the parameter set that minimizes the objective function can be mostly beneficial in the case of VQE. These findings call for a follow-up study on the role of optimizer in conjunction with ADAPT-VQE, extending the analysis to a larger selection of optimizers and gradient strategies.

The ongoing development of VQE methods, including ADAPT-VQE, must also address the noise that is intrinsic to the operations implemented in experimental quantum computers. The above benchmarks of infidelity and energy error place lower bounds on the expected accuracy for VQE methods using noiseless numerical simulation. However, we anticipate that the introduction of noise will substantially affect the accuracy with which the prepared ansatz state approaches the pure state expected from conventional quantum chemistry theory. However, if the state infidelity grows with increasing molecular size, as indicated by our short series of examples, then lower bounds on ansatz accuracy may become a non-trivial contribution to observed errors in experimental measurements.

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation. XACC and TNQVM are open-source and can be found at https://github.com/eclipse/xacc and https://github.com/ornl-qci/tnqvm.

AUTHOR'S NOTE

This manuscript has been authored by UT-Battelle, LLC, under Contract No. DE-AC0500OR22725 with the U.S. Department of Energy. The United States Government retains and the publisher, by accepting the article for publication, acknowledges that the United States Government retains a non-exclusive, paid-up, irrevocable, world-wide license to publish or reproduce the published form of this manuscript, or allow others to do so, for the United States Government purposes. The Department of Energy will provide public access to these results of federally sponsored research in accordance with the DOE Public Access Plan.

AUTHOR CONTRIBUTIONS

DC implemented the ADAPT-VQE algorithm, ran some of the simulations, and wrote the manuscript. JW ran some of the simulations, wrote the code to compute state fidelities, and generated the plots. AM implemented the VQE algorithm and oversaw the ADAPT-VQE implementation. TH designed the research and helped writing the manuscript.

ACKNOWLEDGMENTS

This work was supported by the Embedding Quantum Computing into Many-body Frameworks for Strongly Correlated Molecular and Materials Systems project, which was funded by the U.S. Department of Energy (DOE), Office of Science, Office of Basic Energy Sciences, the Division of Chemical Sciences, Geosciences, and Biosciences. This research used resources of the Oak Ridge Leadership Computing Facility, which is a DOE Office of Science User Facilities supported by the Oak Ridge National Laboratory under Contract DE-AC05-00OR22725. This research used resources of the Compute and Data Environment for Science (CADES) at the Oak Ridge National Laboratory, which was supported by the Office of Science of the U.S. Department of Energy under Contract No. DE-AC05-00OR22725. This work was carried out at Oak Ridge National Laboratory, managed by UT-Battelle, LLC for the U.S. Department of Energy under contract DE-AC05-00OR22725.

REFERENCES

Abraham, H., et al. (2019). Qiskit: An open-source framework for quantum computing.

Bartlett, R. J., Kucharski, S. A., and Noga, J. (1989). Alternative coupled-cluster anstze ii. The unitary coupled-cluster method. *Chem. Phys. Lett*. 155, 133–140. doi: 10.1016/S0009-2614(89)87372-5

Britt, K. A., and Humble, T. S. (2017). High-performance computing with quantum processing units. *ACM J. Emerg. Technol. Comput. Syst*. 13, 1–13. doi: 10.1145/3007651

Cao, Y., Romero, J., Olson, J. P., Degroote, M., Johnson, P. D., Kieferová, M., et al. (2019). Quantum chemistry in the age of quantum computing. *Chem. Rev*. 119, 10856–10915. doi: 10.1021/acs.chemrev.8b00803

Chivilikhin, D., Samarin, A., Ulyantsev, V., Iorsh, I., Oganov, A. R., and Kyriienko, O. (2020). MoG-VQE: Multiobjective genetic variational quantum eigensolver. *arXiv [Preprint]*. arXiv:2007.04424.

Evangelista, F. A., Chan, G. K.-L., and Scuseria, G. E. (2019). Exact parameterization of fermionic wave functions via unitary coupled cluster theory. *J. Chem. Phys*. 151:244112. doi: 10.1063/1.5133059

Fishman, M., White, S. R., and Stoudenmire, E. M. (2020). The ITensor software library for tensor network calculations. *arXiv [Preprint]*. arXiv:2007.14822.

Gheorghiu, V. (2018). Quantum++: A modern C++ quantum computing library. *PLoS ONE* 13:e0208073. doi: 10.1371/journal.pone.0208073

Grimsley, H. R., Claudino, D., Economou, S. E., Barnes, E., and Mayhall, N. J. (2020). Is the trotterized uccsd ansatz chemically well-defined? *J. Chem. Theory Comput*. 16, 1–6. doi: 10.1021/acs.jctc.9b01083

Grimsley, H. R., Economou, S. E., Barnes, E., and Mayhall, N. J. (2019). An adaptive variational algorithm for exact molecular simulations on a quantum computer. *Nat. Commun*. 10:3007. doi: 10.1038/s41467-019-10988-2

Hehre, W. J., Ditchfield, R., Stewart, R. F., and Pople, J. A. (1970). Self-consistent molecular orbital methods. IV. Use of gaussian expansions of slater-type orbitals. extension to second-row molecules. *J. Chem. Phys*. 52, 2769–2773. doi: 10.1063/1.1673374

Hehre, W. J., Stewart, R. F., and Pople, J. A. (1969). Self-consistent molecular-orbital methods. I. Use of gaussian expansions of slater-type atomic orbitals. *J. Chem. Phys*. 51, 2657–2664. doi: 10.1063/1.1672392

Humble, T. S., Thapliyal, H., Munoz-Coreas, E., Mohiyaddin, F. A., and Bennink, R. S. (2019). Quantum computing circuits and devices. *IEEE Design Test* 36, 69–94. doi: 10.1109/MDAT.2019.2907130

Jeziorski, B. and Monkhorst, H. J. (1981). Coupled-cluster method for multideterminantal reference states. *Phys. Rev. A* 24, 1668–1681. doi: 10.1103/PhysRevA.24.1668

Johnson, S. *The NLopt Nonlinear-Optimization Package.*

Jordan, P., and Wigner, E. (1928). Über das paulische Äquivalenzverbot. *Zeitsch. Phys*. 47, 631–651. doi: 10.1007/BF01331938

Kandala, A., Mezzacapo, A., Temme, K., Takita, M., Brink, M., Chow, J. M., et al. (2017). Hardware-efficient variational quantum eigensolver for small molecules and quantum magnets. *Nature* 549:242. doi: 10.1038/nature23879

Kutzelnigg, W. (1991). Error analysis and improvements of coupled-cluster theory. *Theor. Chim. Acta* 80, 349–386. doi: 10.1007/BF01117418

Liu, D., and Nocedal, J. (1989). On the limited memory BFGs method for large scale optimization. *Math. Programm*. 45, 503–528. doi: 10.1007/BF01589116

Lyakh, D. I., Musial, M., Lotrich, V. F., and Bartlett, R. J. (2012). Multireference nature of chemistry: the coupled-cluster view. *Chem. Rev*. 112, 182–243. doi: 10.1021/cr2001417

McArdle, S., Endo, S., Aspuru-Guzik, A., Benjamin, S. C., and Yuan, X. (2020). Quantum computational chemistry. *Rev. Modern Phys*. 92:015003. doi: 10.1103/RevModPhys.92.015003

McCaskey, A., Dumitrescu, E., Chen, M., Lyakh, D., and Humble, T. (2018a). Validating quantum-classical programming models with tensor network simulations. *PLoS ONE* 13:e206704. doi: 10.1371/journal.pone.0206704

McCaskey, A., Dumitrescu, E., Liakh, D., Chen, M., Feng, W., and Humble, T. (2018b). A language and hardware independent approach to quantum-classical computing. *SoftwareX* 7, 245–254. doi: 10.1016/j.softx.2018.07.007

McCaskey, A., Lyakh, D., Dumitrescu, E., Powers, S., and Humble, T. (2020). XACC: a system-level software infrastructure for heterogeneous quantum-classical computing. *Quant. Sci. Technol*. 5:024002. doi: 10.1088/2058-9565/ab6bf6

McClean, J. R., Romero, J., Babbush, R., and Aspuru-Guzik, A. (2016). The theory of variational hybrid quantum-classical algorithms. *New J. Phys*. 18:023023. doi: 10.1088/1367-2630/18/2/023023

Mitarai, K., Negoro, M., Kitagawa, M., and Fujii, K. (2018). Quantum circuit learning. *Phys. Rev. A* 98:032309. doi: 10.1103/PhysRevA.98.032309

Nocedal, J. (1980). Updating quasi-newton matrices with limited storage. *Math. Comput*. 35, 773–773. doi: 10.1090/S0025-5718-1980-0572855-7

Parrish, R., Iosue, J., Ozaeta, A., and McMahon, P. (2019a). A jacobi diagonalization and anderson acceleration algorithm for variationalquantum algorithm parameter optimization. *arXiv [Preprint]*. arXiv:1904.03206 [quant-ph].

Parrish, R. M., Hohenstein, E. G., McMahon, P. L., and Martínez, T. J. (2019b). Quantum computation of electronic transitions using a variational quantum eigensolver. *Phys. Rev. Lett*. 122:230401. doi: 10.1103/PhysRevLett.122.230401

Peruzzo, A., McClean, J., Shadbolt, P., Yung, M.-H., Zhou, X.-Q., Love, P. J., et al. (2014). A variational eigenvalue solver on a photonic quantum processor. *Nat. Commun*. 5:4213. doi: 10.1038/ncomms5213

Pietro, W. J., Levi, B. A., Hehre, W. J., and Stewart, R. F. (1980). Molecular orbital theory of the properties of inorganic and organometallic compounds. 1. STO-NG basis sets for third-row main-group elements. *Inorgan. Chem*. 19, 2225–2229. doi: 10.1021/ic50210a005

Powell, M. J. D. (1994). "A direct search optimization method that models the objective and constraint functions by linear interpolation," in *Advances in Optimization and Numerical Analysis*, eds S. Gomez and J.-P. Hennart (Dordrecht: Springer Netherlands), 51–67. doi: 10.1007/978-94-015-8330-5_4

Preskill, J. (2018). Quantum Computing in the NISQ era and beyond. *Quantum* 2:79. doi: 10.22331/q-2018-08-06-79

Quantum, G. A., et al. (2020). Hartree-fock on a superconducting qubit quantum computer. *Science* 369, 1084–1089. doi: 10.1126/science.abb9811

Raghavachari, K., Trucks, G. W., Pople, J. A., and Head-Gordon, M. (1989). A fifth-order perturbation comparison of electron correlation theories. *Chem. Phys. Lett*. 157, 479–483. doi: 10.1016/S0009-2614(89)87395-6

Romero, J., Babbush, R., McClean, J. R., Hempel, C., Love, P. J., and Aspuru-Guzik, A. (2018). Strategies for quantum computing molecular energies using the unitary coupled cluster ansatz. *Quant. Sci. Technol*. 4:014008. doi: 10.1088/2058-9565/aad3e4

Ryabinkin, I. G., Lang, R. A., Genin, S. N., and Izmaylov, A. F. (2020). Iterative qubit coupled cluster approach with efficient screening of generators. *J. Chem. Theory Comput*. 16, 1055–1063. doi: 10.1021/acs.jctc.9b01084

Shavitt, I., and Bartlett, R. J. (2009). *Many-Body Methods in Chemistry and Physics: MBPT and Coupled-Cluster Theory*. Cambridge Molecular Science. Cambridge University Press. doi: 10.1017/CBO9780511596834

Stokes, J., Izaac, J., Killoran, N., and Carleo, G. (2020). Quantum natural gradient. *Quantum* 4:269. doi: 10.22331/q-2020-05-25-269

Sun, Q., Berkelbach, T. C., Blunt, N. S., Booth, G. H., Guo, S., Li, Z., et al. (2017). PySCF: the python-based simulations of chemistry framework. *Wires Comput. Mol. Sci*. 8:e1340. doi: 10.1002/wcms.1340

Tang, H. L., Shkolnikov, V. O., Barron, G. S., Grimsley, H. R., Mayhall, N. J., Barnes, E., et al. (2020). qubit-ADAPT-VQE: An adaptive algorithm for constructing hardware-efficient ansatze on a quantum processor. *arXiv [Preprint]*. arXiv:1911.10205.

Taube, A. G., and Bartlett, R. J. (2006). New perspectives on unitary coupled-cluster theory. *Int. J. Quant. Chem*. 106, 3393–3401. doi: 10.1002/qua.21198

Urban, M., Noga, J., Cole, S. J., and Bartlett, R. J. (1985). Towards a full CCSDT model for electron correlation. *J. Chem. Phys*. 83, 4041–4046. doi: 10.1063/1.449067

Watts, J. D., Gauss, J., and Bartlett, R. J. (1993). Coupled-cluster methods with noniterative triple excitations for restricted open-shell hartree-fock and other general single determinant reference functions. Energies and analytical gradients. *J. Chem. Phys*. 98, 8718–8733. doi: 10.1063/1.464480

Zhu, L., Tang, H. L., Barron, G. S., Mayhall, N. J., Barnes, E., and Economou, S. E. (2020). An adaptive quantum approximate optimization algorithm for solving combinatorial problems on a quantum computer. *arXiv [Preprint]*. arXiv:2005.10258.

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

ORIGINAL RESEARCH
published: 01 March 2021
doi: 10.3389/fphy.2020.586374

Quantum Machine Learning Tensor Network States

Andrey Kardashin*, Alexey Uvarov and Jacob Biamonte

Skolkovo Institute of Science and Technology, Moscow, Russia

Tensor network algorithms seek to minimize correlations to compress the classical data representing quantum states. Tensor network algorithms and similar tools—called tensor network methods—form the backbone of modern numerical methods used to simulate many-body physics and have a further range of applications in machine learning. Finding and contracting tensor network states is a computational task, which may be accelerated by quantum computing. We present a quantum algorithm that returns a classical description of a rank-*r* tensor network state satisfying an area law and approximating an eigenvector given black-box access to a unitary matrix. Our work creates a bridge between several contemporary approaches, including tensor networks, the variational quantum eigensolver (VQE), quantum approximate optimization algorithm (QAOA), and quantum computation.

Keywords: quantum computing, quantum algorithms and circuits, tensor network algorithms, ground state, properties, machine learning, quantum information

Edited by:
Sabre Kais,
Purdue University, United States

Reviewed by:
Peter McMahon,
Cornell University, United States
Kathleen Hamilton,
Oak Ridge National Laboratory (DOE),
United States

****Correspondence:***
Andrey Kardashin
andrey.kardashin@skoltech.ru

Specialty section:
This article was submitted to
Quantum Computing,
a section of the journal
Frontiers in Physics

Received: *23 July 2020*
Accepted: *08 December 2020*
Published: *01 March 2021*

Citation:
Kardashin A, Uvarov A and Biamonte J
(2021) Quantum Machine Learning
Tensor Network States.
Front. Phys. 8:586374.
doi: 10.3389/fphy.2020.586374

1 INTRODUCTION

Tensor network methods provide the contemporary state of the art in the classical simulation of quantum systems. A range of numerical and analytical tools have now emerged, including tensor network algorithms, to simulate quantum systems classically; these algorithms are based in part on powerful insights related to the area law [1–9]. The area law places bounds on quantum entanglement that a many-body system can generate, which translates directly to the amount of memory required to store a given quantum state; see, e.g., [8].

The leading classical methods to simulate random circuits for quantum computational supremacy demonstration are also based on tensor network contractions. Additionally, classical machine learning has been merged with matrix product states and other tensor network methods [10–14]. How might quantum computing accelerate tensor network algorithms?

Although tensor network tools have traditionally been developed to simulate quantum systems classically, we propose a quantum algorithm to approximate an eigenvector of a unitary matrix with bounded rank tensor network states. The algorithm works given only black-box access to a unitary matrix. In general, tensor network contraction can simulate any quantum computation.

We focus on 1D chains of tensors (matrix product states) due to some associated analytical simplifications; indeed, matrix product states can be approximated classically which offers an attractive gold standard to compare the quantum algorithm against. The general framework we develop applies equally well to 2D and, e.g., sparse networks (projected entangled pair states, etc.). However, an early merger between these topics is better situated to focus on 1D.

Even in 1D, tensor networks offer certain insights into quantum algorithms. For example, the maximal degree of entanglement can often be bounded in the description of the tensor network state itself. In other words, the bond dimension (the dimension of the wires) in the tensor network acts to bound the maximal entanglement. Merging quantum computation with ideas from tensor networks provides new tools to quantify the entanglement that a given quantum circuit can generate [15, 16].

For the sake of simplicity, we work in the black-box setting and assume access to a provided unitary Q. The black-box setting does not consider the implementation of Q. Prima facie, this appears to be a limitation; in practice, however, the restriction can easily be lifted. For example, in QAOA, the problem Hamiltonian can be applied for varying times, offering a natural extension of the oracle idea by giving Q a simple time dependence [17].

In Discussion, we drop the black-box access restriction and cast the steps needed to perform a meaningful near-term demonstration of our algorithm on a quantum computer, providing a low-rank approximation to eigenvectors of the quantum computers free- (or effective) Hamiltonian. The presented algorithm falls into the class of variational quantum algorithms [18–25]. It returns a classical description, in the form of a tensor network, of an eigenvector of an operator found through an iterative classical-to-quantum optimization process.

We present a general framework to determine tensor networks using quantum processors. We focus on 1D, which enables several results related to the maximum amounts of entanglement required to demonstrate these methods. This analysis is followed by a discussion focused on applications of these techniques and what might be required for a meaningful near-term experimental demonstration.

2 METHODS

The algorithm we propose solves the following problem: *given black-box access to a unitary Q, find any eigenvector of Q.*

We work in the standard mathematical setting of quantum computing. We define n qubits arranged on a line and fix the standard canonical (computational) basis. We consider the commutative Hermitian subalgebra generated by the n-projectors:

$$P_i = |0\rangle\langle 0|_i, \tag{1}$$

where the subscript i denotes the corresponding ith qubit acted on by P_i, with the remainder of the state space acting on trivially. These form our observables.

Rank is the maximum Schmidt number (the nonzero singular values) across any of the $n-1$ stepwise partitions of the qubits on a line. Rank provides an upper bound on the bipartite entanglement that a quantum state can support; as will be seen, a rank-r state has at most $k = \log_2(r)$ ebits of entanglement. The quantum algorithm we present works by finding a sequence of maximally k ebit approximations, where the k'th approximation can be used to seed the $(k+1)$'th approximation.

An ebit is the amount of entanglement contained in a maximally entangled two-qubit (Bell) state. A quantum state with q ebits of entanglement (quantified by any entanglement measure) has the same amount of entanglement (in that measure) as q Bell states. If a task requires l ebits, it can be done with l or more Bell states, but not with fewer. Maximally entangled states in

$$\mathbb{C}^d \otimes \mathbb{C}^d \tag{2}$$

have $\log_2(d)$ ebits of entanglement. The question is then to upper bound the maximum amount of entanglement a given quantum computation can generate, providing a coarse-graining to classify quantum algorithms in terms of both the circuit depth and the maximum ebits possible. For low-depth circuits, these arguments are surprisingly relevant.

We parameterize a circuit family generating matrix product states with θ, a real vector with entries in $[0, 2\pi)$. We consider action on the initial rank-1 state $|\mathbf{0}\rangle = |0\rangle^{\otimes n}$ and define two states

$$|\psi(\theta)\rangle = U^\dagger(\theta) Q U(\theta)|\mathbf{0}\rangle \tag{3}$$

and

$$|\tilde{\psi}(\theta)\rangle = U(\theta)|\mathbf{0}\rangle, \tag{4}$$

both of yet to be specied rank.

We will construct an objective function (**Eq. 6**) to minimize and hence to recover our approximate eigenvector. The choice of this function provides a desirable degree of freedom to further tailor the algorithm to the particular quantum processor at hand. We choose

$$p_i(\theta) = \langle \psi(\theta)|P_i|\psi(\theta)\rangle \tag{5}$$

and call

$$\mathcal{L}(\theta) = \sum_{i=1}^{n} \ln p_i(\theta) \tag{6}$$

the log-likelihood function of the n-point correlator

$$\prod_{i=1}^{n} p_i(\theta). \tag{7}$$

The minimization of **Eq. 6** corresponds to maximizing the probability of measuring each qubit in $|0\rangle$. This minimization can be done using a variety of optimization and machine learning algorithms. The following summarizes the steps of the algorithm.

```
Algorithm 1: Find successive tensor network approximations of an eigenvector of $Q$.
  Choose the maximum number of ebits $k_{\max}$
  Choose the maximum number of optimization iterations $n_{\text{it}}$
  for $k \leftarrow 1$ to $k_{\max}$ do
      Construct the ansatz $U_k$ corresponding to a $k$ ebit MPS
      Set $\boldsymbol{\theta}_k$ randomly
      for $j \leftarrow 1$ to $n_{\text{it}}$ do
          Evaluate $\boldsymbol{p}(\boldsymbol{\theta}_k)$
          Evaluate $\mathcal{L}(\boldsymbol{p})$
          Update $\boldsymbol{\theta}_k$ using a classical optimizer
      end for
      Store $\mathcal{L}_k = \mathcal{L}(\boldsymbol{p})$
  end for
  return $\{\boldsymbol{\theta}_k\}_{k=1}^{k_{\max}}, \{\mathcal{L}_k\}_{k=1}^{k_{\max}}$
```

The algorithm begins with rank-1 qubit states as

$$|\tilde{\psi}(\theta)\rangle = \bigotimes_{i=1}^{n} \left(\cos\theta_1^i |0\rangle + e^{-\imath\theta_2^i} \sin\theta_1^i |1\rangle \right). \tag{8}$$

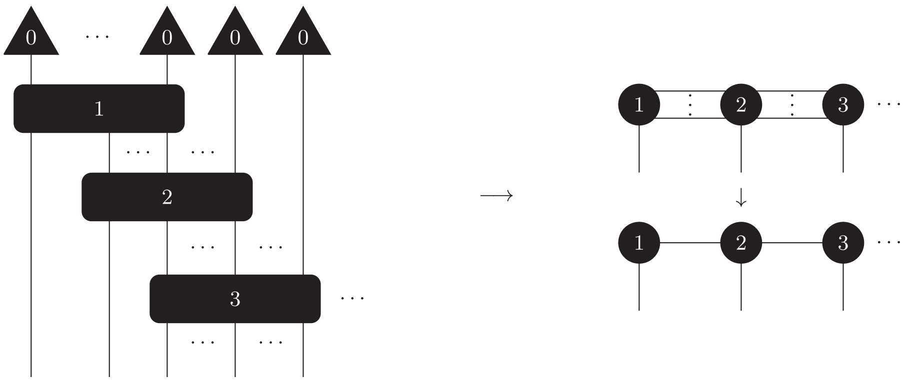


FIGURE 1 | Example of a tensor network as a quantum circuit: **(left)** quantum circuit realization of a matrix product state with open boundary conditions; **(right)** using standard graphical rewrite rules—or by manipulating equations—one readily recovers the familiar matrix product state depiction as a "train of tensors."

Minimization of the objective function **Eq. 6** returns $2n$ real numbers describing a local matrix product state. Approximations of higher rank are made by utilizing the quantum circuit structure given in **Figure 1**.

3 RESULTS

The algorithm works given only oracle access to a unitary Q. The spectrum of Q is necessarily contained on the complex unit circle and so we note immediately that

$$\begin{aligned} 1 = \max_{\phi} |\langle\phi|Q|\phi\rangle|^2 &\geq \max_{\theta} |\langle \mathbf{0}|\psi(\theta)\rangle|^2 \\ &= \max_{\theta} |\langle\tilde{\psi}(\theta)|Q|\tilde{\psi}(\theta)\rangle|^2 \end{aligned} \quad (9)$$

with equality of the left-hand side if and only if $|\phi\rangle$ is an eigenvector of Q. One advantage of the presented method is that it terminates when the measurement reaches a given value. This implies that the system is in an eigenstate. Such a certificate is not directly established using other variational quantum algorithms.

Importantly, the maximization over θ on the right-hand side of **Eq. 9** corresponds to the minimization of the log-likelihood **Eq. 6**. We will then parameterize $\tilde{\psi}(\theta_k)$, where k denotes a k ebit matrix product state of interest. Learning this matrix product state recovers an approximation to an eigenvector of Q. With a further promise on Q that all eigenvectors have a rank-p matrix product state representation, we conclude that $r < p$ implies a fundamental error in our approximation. We consider then that the r'th singular value of the state takes the value ε. It then follows that the one-norm error scales with $O(\varepsilon)$ and the two-norm error scales only with $O(\varepsilon^2)$. In general, we arrive at the monotonic sequence ordered by the following relation:

$$\begin{aligned} 1 &\geq \max_{\theta_{k+1}} \left|\langle\tilde{\psi}(\theta_{k+1})|Q|\tilde{\psi}(\theta_{k+1})\rangle\right|^2 \\ &\geq \max_{\theta_k} |\langle\tilde{\psi}(\theta_k)|Q|\tilde{\psi}(\theta_k)\rangle|^2 \end{aligned} \quad (10)$$

which is valid for $k = 1$ to $\lfloor n/2 \rfloor$ (minimum to maximum possible number of ebits).

Indeed, increasing the rank of the matrix product state approximation can improve the eigenvector approximation. Yet, it should be noted that ground state eigenvectors of physical systems are in many cases known to be well approximated with low-rank matrix product states [1–9]. This depends on the further properties of Q and is a subject of intensive study in numerical methods, further motivating the quantum algorithm we present here. We will develop our algorithm agnostic to Q, leaving a more specific near-term demonstration (in which Q is implemented); e.g., we will express any $|\tilde{\psi}(\theta)\rangle$ as a matrix product state as

$$|\tilde{\psi}(\theta)\rangle = \sum_{q,s,\ldots,n} A_q^{[\theta_q]} A_s^{[\theta_s]} \cdots A_n^{[\theta_n]} |q, s, \ldots, n\rangle \quad (11)$$

In **Eq. 11**, the rank-r of the representation is embedded into the realization of the A's. Quantum mechanics allows the deterministic generation of a class of isometries, where an isometry U that is also an endomorphism on a particular space is called unitary.

Matrix product states (**Eq. 11**) are not isometries, though correlation functions are readily calculated from them. Furthermore, matrix product states can be deterministically generated by the uniform quantum circuit given in **Figure 1**. Other isometric structures of interest include trees and the so-called Multiscale Entanglement Renormalization Ansatz (MERA) networks [3, 26–28].

Consider then a rank-r approximation to an eigenvector of Q. The blocks in **Figure 1** represent unitary maps. These circuits act on at most $\lceil \log_2(r) \rceil$ qubits. Hence, each of these blocks has at most r^2 real degrees of freedom in $[0, 2\pi)$. The general realization of these blocks using the typical basis of CNOT gates and arbitrary local unitaries can be done by a range of methods; see, e.g., [29]. A commonly used theoretical lower bound requires

$$\frac{1}{4}\left(r^2 - 3\log_2 r - 1\right) \quad (12)$$

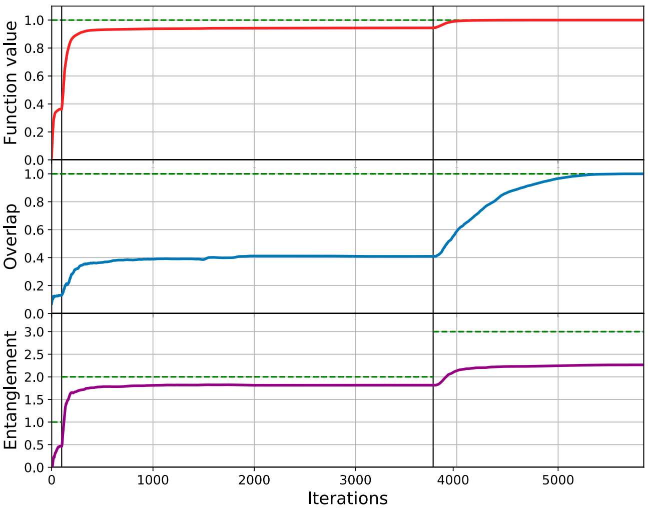


FIGURE 2 | Algorithm demonstration on randomly generated 6-qubit unitaries Q: the value of **Eq. 9 (upper)**, overlap between the variational state and the closest eigenstate of Q **(middle)**, and the von Neumann entropy of the subsystem of the first three qubits **(lower)**. The vertical solid lines indicate the iteration numbers after which k, the number of ebits that the MPS ansatz can support, increases by 1. The plot is obtained by averaging over 10 randomly generated unitaries Q.

CNOT gates, where the method in [29] requires r^2 local qubit gates and did not reach this theoretical lower bound of CNOT gates. The total number of single qubits and CNOT gates nevertheless scales as $O(r^2)$ for each block, where the number of blocks is bounded by n. Hence, the implementation complexity scales as $O(l \cdot n \cdot r^2)$, where the optimization routine terminates after l steps (perhaps in a local minimum).

Instead of preparing $|\tilde{\psi}(\boldsymbol{\theta})\rangle$ by a quantum circuit with $\boldsymbol{\theta} \in (0, 2\pi]^{\times l}$ tunable parameters as

$$|\tilde{\psi}(\boldsymbol{\theta})\rangle = \prod_l U_l |0\rangle^{\otimes n} \tag{13}$$

where U_l is adjusted by θ_l, one might adopt an alternative (heuristic) circuit realization performed by adjusting controllable Hamiltonian parameters realizing each block, subject again to the minimization of **Eq. 6**. With such an approach, one will prepare $|\tilde{\psi}(\boldsymbol{\theta})\rangle$ by tuning accessible time-dependent parameters $\theta_k(t)$ corresponding to Hermitian A^k as

$$|\tilde{\psi}\rangle = \mathcal{T}\left\{e^{-\imath \sum \theta_k(t) A^k}\right\}|0\rangle^{\otimes n}, \tag{14}$$

where $\mathcal{T}$ orders the sequence by time and superscript k indexes the kth operator A^k. Provided these sequences are localized appropriately, the matrix product structure still remains.

We then consider vertical partitions of a quantum circuit with the n qubits positioned horizontally on a line. For an m-depth quantum circuit (where m is presumably bounded above by a low-order polynomial in n), the maximum number of two-qubit gates crossed in a vertical partition is never more than m. The maximum number of ebits generated by a fully entangling two-qubit CNOT gate is never more than a single ebit. We then consider the $(n-1)$ partitions of the qubits, the maximum partition with represent to ebits is into two (ideally) equal halves, which is never more than $\lceil n/2 \rceil$. We then arrive at the general result that an m-depth quantum circuit on n qubits never applies more than

$$\min\{\lceil n/2 \rceil, m\} \tag{15}$$

ebits of entanglement. This immediately puts a lower bound of $\sim n/2$ on the two-qubit gate depth for Q to potentially drive a system into a state supporting the maximum possible ebits of entanglement.

In **Figure 2**, we demonstrate our algorithm for finding an eigenstate of randomly generated 5-qubit unitary matrices. For minimizing the function **Eq. 6**, we used the Broyden-Fletcher-Goldfarb-Shanno (BFGS) minimization method [30]. For each k ebit MPS, we place the k-layered hardware-efficient ansatz as the operators in blocks [21].

4 DISCUSSION

We now turn to the realization of Q and sketch a possible demonstration for a near-term device. Polynomial-time simulation of Hamiltonian evolution is well known to be **BQP**-hard. This provides an avenue for Q to represent a problem of significant computational interest, as simulating quantum evolution and quantum factoring are in **BQP**. We aim to bootstrap properties of the quantum processor as much as possible to reduce resources for a realization; see, for example, [21].

Let $Q(t)$ be the one-parameter unitary group generated by $\mathcal{H}$, where $\mathcal{H}$ represents a 3-SAT instance. Given access to an oracle computing

$$\langle \tilde{\psi}(\theta_1) | \mathcal{H} | \tilde{\psi}(\theta_1) \rangle, \tag{16}$$

we can minimize the overall eigenvectors, which is **NP**-hard. Hence, finding even rank-1 states can be **NP**-hard. This provides a connection between our method and QAOA [31]. Similarly, we can also use this external minimization to connect our method to VQE [20]. However, our method provides a certificate that, on proper termination, the system is indeed in such a desired eigenstate.

When $\mathcal{H}$ is a general quantum Hamiltonian, minimization of

$$\langle \tilde{\psi}(\theta_k) | \mathcal{H} | \tilde{\psi}(\theta_k) \rangle \tag{17}$$

is in turn **QMA**-hard. For example, pairing our procedure with an additional procedure (quantum phase estimation) to minimize Q, the overall eigenvectors would hence provide rank-k variational states and hence our methods provide a research direction which incorporates tensor network methods in works such as, e.g., [19–21]. It should however be noted that phase estimation adds significant experimental difficulty compared with the algorithm presented here and the algorithm is closer to VQE (with evident differences as listed above and in the main text).

For a near-term demonstration, we envision Q to be realized by bootstrapping the underlying physics of the system realizing Q, e.g., using the hardware-efficient ansatz [21]. For instance, one can realize Q as a modification of the systems free Hamiltonian using effective Hamiltonian methods (modulating local gates). This greatly reduces practical requirements on Q.

The interaction graph of the Hamiltonian generating Q can be used to define a PEPS tensor network (as it will have the same structure as the layout of the chip, it will no longer have the contractable properties of matrix product states, and yet is still of interest) [4]. The algorithm works otherwise unchanged, but the circuit acts on this interaction graph (instead of a line) to create a corresponding tensor network state (a quantum circuit in the form of, e.g., the variational ansatz). Tailored free evolution of the system Hamiltonian generates Q. Our algorithm returns a tensor network approximation of an eigenstate of Q.

The first interesting demonstrations of the quantum algorithm we have presented should realize rank-k tensor networks (matrix product state), and the corresponding tensor network can be realized with a few hundred gates for a system with a few hundred qubits.

DATA AVAILABILITY STATEMENT

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found: GitHub, https://git.io/JkjvV.

AUTHOR CONTRIBUTIONS

All authors conceived and developed the theory and design of this study and verified the methods. AK developed and deployed the code to collect numerical data. All authors contributed to interpreting the results and writing the manuscript.

ACKNOWLEDGEMENTS

AK and JB acknowledge support from agreement No. 014/20, *Leading Research Center on Quantum Computing*. AU acknowledges RFBR project No. 19-31-90159 for financial support. This manuscript has been released as a preprint as arXiv:1804.02398 (32).

REFERENCES

1. Biamonte J. Charged string tensor networks. *Proc Natl Acad Sci USA* (2017) 114(10):2447. doi:10.1073/pnas.1700736114
2. Orús R. A practical introduction to tensor networks: matrix product states and projected entangled pair states. *Ann Phys* (2014) 349:117–58. doi:10.1016/j.aop.2014.06.013
3. Vidal G. Entanglement renormalization: an introduction. In: Carr LD., editor. *Understanding quantum phase transitions*. Boca Raton, United States: Taylor and Francis (2010) p. 756.
4. Verstraete F, Murg V, Cirac JI. Matrix product states, projected entangled pair states, and variational renormalization group methods for quantum spin systems. *Adv Phys* (2008) 57:143–224. doi:10.1080/14789940801912366
5. Cirac JI, Verstraete F. Renormalization and tensor product states in spin chains and lattices. *J Phys Math Theor* (2009) 42(50):504004. doi:10.1088/1751-8113/42/50/504004
6. Schollwöck U. The density-matrix renormalization group in the age of matrix product states. *Ann Phys* (2011) 326:96–192. doi:10.1016/j.aop.2010.09.012
7. Orús R. Advances on tensor network theory: symmetries, fermions, entanglement, and holography. *Eur Phys J B* (2014) 87:280. doi:10.1140/epjb/e2014-50502-9
8. Eisert J. Entanglement and tensor network states. *Model Simulat* (2013) 3:520.
9. Evenbly G, Vidal G. Tensor network states and geometry. *J Stat Phys* (2011) 145:891–918. doi:10.1007/s10955-011-0237-4
10. Cichocki A, Lee N, Oseledets I, Phan A-H, Zhao Q, Danilo P, et al. Tensor networks for dimensionality reduction and large-scale optimization: Part 1 low-rank tensor decompositions. *Found Trends Machine Learn* (2016) 9(4–5):249–429. doi:10.1561/2200000059
11. Cichocki A, Phan A-H, Zhao Q, Lee N, Oseledets I, Sugiyama M, et al. Tensor networks for dimensionality reduction and large-scale optimization: Part 2 applications and future perspectives. *Found Trends Machine Learn* (2017) 9(6):431–673. doi:10.1561/2200000067
12. Clark SR. Unifying neural-network quantum states and correlator product states via tensor networks. *J Phys Math Theor* (2018) 51(13):135301. doi:10.1088/1751-8121/aaaaf2
13. Huggins W, Patil P, Mitchell B, Whaley KB, Miles Stoudenmire E. Towards quantum machine learning with tensor networks. *Quant Sci Technol* (2019) 4(2):024001. doi:10.1088/2058-9565/aaea94
14. Ding L, Ran S-J, Wittek P, Cheng P, García RB, Su G, et al. Machine learning by unitary tensor network of hierarchical tree structure. *New J Phys* (2017) 21:073059. doi:10.1088/1367-2630/ab31ef
15. Biamonte JD, Morales MES, Koh DE. Entanglement scaling in quantum advantage benchmarks. *Phys Rev* (2020) 101(1):012349. doi:10.1103/PhysRevA.101.012349
16. Biamonte J, Bergholm V, Lanzagorta M. Tensor network methods for invariant theory. *J Phys A Math Gener* (2013) 46:475301. doi:10.1088/1751-8113/46/47/475301
17. Morales MES, Tlyachev T, Jacob B. Variational learning of grovers quantum search algorithm. *Phys Rev* (2018) 98(6):062333. doi:10.1103/PhysRevA.98.062333
18. Yung MH, Casanova J, Mezzacapo A, McClean J, Lamata L, Aspuru-Guzik A, et al. From transistor to trapped-ion computers for quantum chemistry. *Sci Rep* (2014) 4:3589. doi:10.1038/srep03589
19. McClean J, Romero J, Ryan B, Aspuru-Guzik A. The theory of variational hybrid quantum-classical algorithms. *New J Phys* (2016) 18:023023. doi:10.1088/1367-2630/18/2/023023
20. Peruzzo A, McClean J, Shadbolt P, Yung MH, Zhou XQ, Love PJ, et al. A variational eigenvalue solver on a photonic quantum processor. *Nat Commun* (2014) 5:4213. doi:10.1038/ncomms5213
21. Kandala A, Mezzacapo A, Temme K, Takita M, Brink M, Chow JM, et al. Hardware-efficient variational quantum eigensolver for small molecules and quantum magnets. *Nature* (2017) 549:242–6. doi:10.1038/nature23879
22. Jacob B. Universal variational quantum computation. Preprint repository name [Preprint] (2019) Available from: arXiv:1903.04500 (Accessed March 11, 2019).
23. Xia R, Kais S. Quantum machine learning for electronic structure calculations. *Nat Commun* (2018) 9(1):4195. doi:10.1038/s41467-018-06598-z
24. Grant E, Benedetti M, Cao S, Hallam A, Lockhart J, Stojevic V, et al. Hierarchical quantum classifiers. *Npj Quant Inform* (2018) 4:65. doi:10.1038/s41534-018-0116-9
25. Akshay V, Philathong H, Morales MES, Biamonte JD. Reachability deficits in quantum approximate optimization. *Phys Rev Lett* (2020) 124(9):090504. doi:10.1103/PhysRevLett.124.090504
26. Entanglement GV. Entanglement renormalization. *Phys Rev Lett* (2007) 99(22):220405. doi:10.1103/PhysRevLett.99.220405
27. Giovannetti V, Montangero S, Fazio R. Quantum Multiscale entanglement renormalization ansatz channels. *Phys Rev Lett* (2008) 101(18):180503. doi:10.1103/PhysRevLett.101.180503
28. Vidal G. Class of quantum many-body states that can be efficiently simulated. *Phys Rev Lett* (2008) 101(11):110501. doi:10.1103/PhysRevLett.101.110501

29. Möttönen M, Vartiainen JJ, Bergholm V, Salomaa MM. Quantum circuits for general multiqubit gates. *Phys Rev Lett* (2004) 93(13):130502. doi:10.1103/PhysRevLett.93.130502
30. Nocedal J, Wright SJ. *Numerical optimization*. New York, NY: Springer (2006) p. 597.
31. Farhi E, Goldstone J, Gutmann S. A quantum approximate optimization algorithm. Preprint repository name [Preprint] (2014) Available from: arXiv: 1411.4028 (Accessed November 14, 2014).
32. Jacob B, Kardashin A, Uvarov A. Quantum machine learning tensor network states. Preprint repository name [Preprint] (2018) Available from: arXiv:1804.02398 (Accessed April 06, 2014).

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Chemistry

ORIGINAL RESEARCH
published: 18 March 2021
doi: 10.3389/fchem.2021.603019

Quantum Solvers for Plane-Wave Hamiltonians: Abridging Virtual Spaces Through the Optimization of Pairwise Correlations

Eric J. Bylaska[1]*, *Duo Song*[1], *Nicholas P. Bauman*[1], *Karol Kowalski*[1], *Daniel Claudino*[2,3] *and Travis S. Humble*[2,4]

[1]Physical Sciences Division, Pacific Northwest National Laboratory, Richland, WA, United States, [2]Quantum Computing Institute, Oak Ridge National Laboratory, Oak Ridge, TN, United States, [3]Computer Science and Mathematics, Oak Ridge National Laboratory, Oak Ridge, TN, United States, [4]Computational Sciences and Engineering, Oak Ridge National Laboratory, Oak Ridge, TN, United States

For many-body methods such as MCSCF and CASSCF, in which the number of one-electron orbitals is optimized and independent of the basis set used, there are no problems with using plane-wave basis sets. However, for methods currently used in quantum computing such as select configuration interaction (CI) and coupled cluster (CC) methods, it is necessary to have a virtual space that is able to capture a significant amount of electron-electron correlation in the system. The virtual orbitals in a pseudopotential plane-wave Hartree–Fock calculation, because of Coulomb repulsion, are often scattering states that interact very weakly with the filled orbitals. As a result, very little correlation energy is captured from them. The use of virtual spaces derived from the one-electron operators has also been tried, and while some correlations are captured, the amount is quite low. To overcome these limitations, we have been developing new classes of algorithms to define virtual spaces by optimizing orbitals from small pairwise CI Hamiltonians, which we term as correlation optimized virtual orbitals with the abbreviation COVOs. With these procedures, we have been able to derive virtual spaces, containing only a few orbitals, which are able to capture a significant amount of correlation. The focus in this manuscript is on using these derived basis sets to target full CI (FCI) quality results for H_2 on near-term quantum computers. However, the initial results for this approach were promising. We were able to obtain good agreement with FCI/cc-pVTZ results for this system with just 4 virtual orbitals, using both FCI and quantum simulations. The quality of the results using COVOs suggests that it may be possible to use them in other many-body approaches, including coupled cluster and Møller–Plesset perturbation theories, and open up the door to many-body calculations for pseudopotential plane-wave basis set methods.

Keywords: quantum computing, ADAPT-VQE, many-body calculations, DUCC, pseudopotential plane-wave, correlation optimized virtual orbitals, COVOs, nwchem

Edited by:
Javier Carrasco,
CIC energigune, Spain

Reviewed by:
Rajat Chaudhuri,
Indian Institute of Astrophysics, India
Masato Kobayashi,
Hokkaido University, Japan

Correspondence:
Eric J. Bylaska
eric.bylaska@pnnl.gov

Specialty section:
This article was submitted to Physical Chemistry and Chemical Physics, a section of the journal Frontiers in Chemistry

Received: *04 September 2020*
Accepted: *11 January 2021*
Published: *18 March 2021*

Citation:
Bylaska EJ, Song D, Bauman NP, Kowalski K, Claudino D and Humble TS (2021) Quantum Solvers for Plane-Wave Hamiltonians: Abridging Virtual Spaces Through the Optimization of Pairwise Correlations. Front. Chem. 9:603019. doi: 10.3389/fchem.2021.603019

INTRODUCTION

Quantum chemistry is one of the first and most successful scientific applications of digital computers (Mulliken et al., 1941; Mulliken and Rieke, 1941; Boys, 1950; Parr and Mulliken, 1950; Hall, 1951; Roothaan, 1951; Boys et al., 1956; Nesbet, 1960; Allen and Karo, 1962; Nesbet, 1963; Pople et al., 1965; Kohn and Sham, 1965; Reeves, 1966; Pulay, 1969). This success has led to a large number of research, open-source (Dupuis et al., 1989; Stanton et al., 1992; Schmidt et al., 1993; Briggs et al., 1996; Challacombe, 2000; Gygi, 2008; Giannozzi et al., 2009; Deslippe et al., 2012; Hutter et al., 2014; Gonze et al., 2016; Harrison et al., 2016; Apra et al., 2020), and commercial codes (Kresse and Furthmüller, 1996; te Velde et al., 2001; Betteridge et al., 2003; Clark et al., 2005; Werner et al., 2012; Shao et al., 2015; Frisch et al., 2016; Neese, 2018) (for a larger list of quantum chemistry software, see (Wikipedia, The Free Encyclopedia, 2020)), which are used on a regular basis by tens of thousands of scientists, engineers, and students from a variety of scientific and engineering domains. With Moore's law as a backdrop (Moore, 2006), the cycle of new machines leading to new algorithms stimulated the field for many decades, and as a consequence, a large number of quantum chemistry methods were developed along with a variety of numerical methods to solve them. However, in recent decades, the maturity and success of these codes coupled with the imminent death of Moore's law (Dubash, 2005; Rotman, 2020) that made numerical software development much more difficult and less accessible to the average scientist have resulted in the field having priorities other than just new science, such as porting and optimizing these codes to the next generation of computers (Bylaska et al., 2017a, Bylaska et al., 2017b; Richard et al., 2018; van Dam et al., 2020), standardization of methods (Crawford et al., 2017; Wilkins-Diehr and Crawford, 2018), and marketing (Goldbeck, 2017; Hocquet and Wieber, 2017).

With the advent of quantum computing, there is excitement again, and quantum chemists are beginning to rethink how they carry out quantum chemistry calculations, in particular very accurate and very expensive instances of systems containing strong electron-electron correlations. This is because it is anticipated that quantum computers with 50–100 qubits will be able to surpass classical digital computers for these types of calculations (Preskill, 2018). Quantum computing has thus emerged as an alternative avenue to the continuity of quantum chemistry in the long run (Wasielewski et al., 2020) but poses several challenges that demand careful consideration in order to eventually mature into a viable replacement for classical computers and large, highly parallelizable high-performance computing clusters.

Present quantum devices are plagued by short coherence times and vulnerability to environment interference, i.e., noise. Albeit quantum algorithms have been developed with proved exactness, such as quantum phase estimation, these are not a viable option in the present/near-term time frame. Therefore, it is desirable to limit the operation of quantum processors to a complementary concerted execution with classical counterparts, whereby each of these components is only in charge of those tasks for which it is more suitable. This has materialized into the variational quantum eigensolver (VQE) (Peruzzo et al., 2014) and other hybrid algorithms. Briefly, this class of algorithms strives to find the lowest eigenvalue of a given observable by assuming that the associated quantum state can be accurately represented by a trial wave function and whose parameters are varied according to the Rayleigh—Ritz method (variational principle), with these parameters being updated by the classical computer. The burden on the quantum processor can be further alleviated with strategies such as Trotterization, which in turn introduce other challenges (Evangelista et al., 2019; Grimsley et al., 2020) but can be successfully exploited in the construction of favorable ansatz, as long as a predefined form for the trial wave function is imposed. This is at the heart of the ADAPT-VQE (Grimsley et al., 2019).

Most high-level methods for strongly correlated systems in use today (e.g., full configuration interaction (CI), coupled cluster (CC) and Green's function (GF) approaches) are based on second-quantized Hamiltonians, which are written in terms of creation and annihilation operators for fermion orbitals. These methods are amenable to quantum computers because fermionic creation and annihilation operators can be readily mapped to qubits through the use of some established transformation, among which Jordan—Wigner (Jordan and Wigner, 1928), Bravyi—Kitaev (Bravyi and Kitaev, 2002), and binary codes (Steudtner and Wehner, 2018) stand out, where the number of qubits scales with the number of orbitals in the second-quantized Hamiltonian. In principle, converting the full many-body electronic Hamiltonian to a second-quantized form is exact and popular CC and GF approximations based on this form are very accurate. However, this conversion has a drawback in that it requires the introduction of a basis set, which, for computational cost reasons, needs to be small. Typically, these basis sets are composed of atomic-like orbitals generated with heuristics based on an atom calculation for each kind of atom in the system. An example of this type of basis set is the popular Dunning correlation consistent basis set (Dunning and Hay, 1977; Dunning, 1989) in which the atomic orbitals are optimized at the CISD (configuration interaction method with single and double excitations) level of theory. While the size of this basis set is small compared to other basis sets used in quantum chemistry, such as plane waves, it still needs to contain a large number of atomic orbitals to produce a truly accurate result.

Solving relevant chemistry problems analogously to what is classically done with MCSCF or FCI on near-term quantum computers that contain 10s to 100s of noisy qubits (Reiher et al., 2017), in which only limited numbers of operations can be performed, is a monumental challenge. One way to reduce the cost of these calculations is to develop new procedures for optimizing basis sets. In this manuscript, a new method is presented for generating a plane-wave derived correlation optimized orbital basis sets. These derived basis sets can also be used in other many-body approaches, including CC theory, and can easily be generalized to work with recently developed Filon's Integration Strategy for two-electron integrals in periodic systems (Bylaska et al., 2020). This method is different from other plane-wave derived optimized orbital basis sets (Shirley, 1996; Prendergast and Louie, 2009; Chen et al., 2011); in that, it is based on optimizing small select CI problems rather than fitting one-electron eigenvalue spectra and band structures.

The paper is organized as follows. In **Section 2**, a brief description of the second-quantized Hamiltonian and the double unitary CC downfolding method that can be used with the pseudopotential plane-wave method is given, followed by comparisons between restricted Hartree–Fock (RHF) calculations using plane-wave and Gaussian basis sets. Using this framework, CI calculations up to 20 virtual orbitals, generated from plane-wave Hartree–Fock and one-electron Hamiltonians, are shown for the H_2 molecule. The VQE quantum computing algorithms used in this work are described in **Section 3**. **Section 4** presents a new class of algorithms for generating a virtual space in which the orbitals are generated by minimizing small pairwise CI Hamiltonians, and a complete set of equations for implementing these optimizations is given in **Subsections 4.1–4.4**. Using this new type of virtual space, CI calculations up to 18 virtual orbitals for the ground state energy curve of the H_2 molecule are presented in **Section 5** followed by results using quantum computing simulations in **Section 6**, and lastly, the conclusions are given in **Section 7**.

PSEUDOPOTENTIAL PLANE-WAVE MANY-BODY HAMILTONIAN

The nonrelativistic electronic Schrödinger eigenvalue equation of quantum chemistry can be written as

$$H\left|\Psi\left(x_1, x_2, \ldots, x_{N_e}\right)\right\rangle = E\left|\Psi\left(x_1, x_2, \ldots, x_{N_e}\right)\right\rangle, \tag{1}$$

where H is the electronic structure Hamiltonian under the Born–Oppenheimer approximation and $\left|\Psi\left(x_1, x_2, \ldots, x_{N_e}\right)\right\rangle$ is the quantum mechanical wave function that is a function of the spatial and spin coordinates of the N_e electrons, $x_i = (r_i, \sigma_i)$. When solving this equation, the Pauli exclusion principle constraint of particle exchange must be enforced, in which the wave function changes sign when the coordinates of two particles, x_i and x_j, are interchanged; i.e.,

$$\begin{aligned} &\left|\Psi\left(x_1, x_2, \ldots x_i, \ldots x_j, \ldots, x_{N_e}\right)\right\rangle \\ &= -\left|\Psi\left(x_1, x_2, \ldots x_j, \ldots x_i, \ldots, x_{N_e}\right)\right\rangle. \end{aligned} \tag{2}$$

For the Born–Oppenheimer Hamiltonian, the interaction between the electrons and nuclei is described by the proper potentials $\frac{Ze}{|r_i - R_A|}$, which for plane-wave solvers can cause trouble with convergence because of the singular behavior at $|r - R_A|$. A standard way to remove this issue in plane-wave calculations is to replace these singular potentials with pseudopotentials. By making this replacement, the Hamiltonian, H, in **Eq. 1** can be written as

$$\begin{aligned} H = &-\frac{1}{2}\sum_{i=1}^{N_e}\nabla_i^2 \\ &+ \sum_{i=1}^{N_e}\sum_{A=1}^{N_A}\left(V_{local}^{(A)}\left(|r_i - R_A|\right) + \sum_{lm}\widehat{V}_{NL}^{(A),lm}\right) \\ &+ \sum_{i=1}^{N_e}\sum_{j>i}^{N_e}\frac{1}{|r_i - r_j|}, \end{aligned} \tag{3}$$

where the first term is the kinetic energy operator, the second term contains the local and nonlocal pseudopotentials, $V_{local}^{(A)}$ and $\widehat{V}_{NL}^{(A),lm}$ represent the electron-ion interactions, and the last term is the electron-electron repulsion.

Instead of writing the many-electronic Hamiltonian in the traditional Schrödinger form, as in the equations above, it is more common today to write it in an alternative representation, known as the second-quantization form, defined using the creation, $a_p^\dagger$, and annihilation, a_p, operators. The second-quantized Hamiltonian is written as

$$\begin{aligned} H &= \sum_{p=1}^{N_{basis}}\sum_{q=1}^{N_{basis}} h_{pq} a_p^\dagger a_q + \frac{1}{2}\sum_{pqrs} h_{pqrs} a_p^\dagger a_r^\dagger a_s a_q, \\ h_{pq} &= \int dx\,\phi_p^*(x)\left(-\frac{1}{2}\nabla^2\right)\phi_q(x) \\ &\quad + \int dx\,\phi_p^*(x)\left[\sum_{A=1}^{N_A}\left(V_{local}^{(A)}\left(|r - R_A|\right) + \sum_{lm}\widehat{V}_{NL}^{(A),lm}\right)\right]\phi_q(x), \\ h_{pqrs} &= \int dx_1 dx_2\,\phi_p^*(x_1)\phi_r^*(x_2)\frac{1}{|r_1 - r_2|}\phi_s(x_2)\phi_q(x_1), \end{aligned} \tag{4}$$

where $\phi_p(x)$ represent the one-electron spin-orbital basis. A nice feature about this form of the Hamiltonian is that the antisymmetry of wavefunction requirement as given in **Eq. 2** is automatically enforced through the standard fermionic anticommutation relations $\{a_p, a_q^\dagger\} = \delta_{pq}$ and $\{a_p, a_q\} = \{a_p^\dagger, a_q^\dagger\} = 0$.

In this formulation, the choice of the one-electron spin-orbital basis is nebulous and requires some care in its choosing in order to obtain accurate results with this type of Hamiltonian. Typically, in quantum chemistry, one uses the filled and virtual orbitals from a Hartree–Fock calculation. For methods that utilize linear combinations of atomic orbitals (LCAO) as the basis, the size of the basis set and subsequently generated Hartree–Fock orbitals is fairly small. However, for plane-wave solvers and other grid-based solvers, the size of the basis set is very large and the number of the one- and two-electron integrals in **Eq. 4** will become prohibitive if all possible Hartree–Fock orbitals are used.

One approach to this problem is to only include virtual orbital up to a certain energy threshold, and another related approach is to use the plane-wave derived optimized orbital basis set, e.g., the Shirley approach. While the number of these orbitals needed to accurately describe eigenvalue spectra over a range of ~100 eV is significantly smaller than the number of plane waves, it is still significantly larger than the number of orbitals generated by an LCAO method. The reason for this is that the virtual orbitals in a plane-wave Hartree–Fock calculation, because of Coulomb repulsion, are often unbound scattering states that interact very weakly with the filled orbitals. As a result, very little correlation energy is captured from them. In contrast, LCAO basis methods can only describe bound states, and hence, Hartree–Fock calculations on this basis do not generate these types of scattering states.

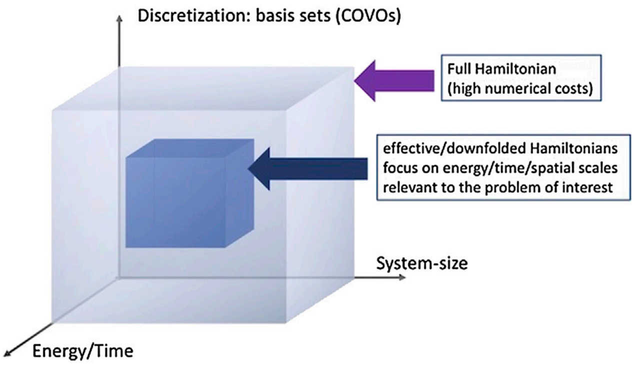


FIGURE 1 | Schematic representation of the dimensionality reduction algorithms considered in this paper: (1) discretization of the many-body problem by employing efficient single-particle basis sets (in this paper, we consider correlation optimized virtual orbitals (COVOs)) and (2) downfolding techniques based on the double unitary coupled cluster (DUCC) formalism (Bauman et al., 2019; Kowalski and Bauman, 2020); in this step, the many-body problem is rerepresented in a subspace of entire Hilbert space.

Many-Body Downfolding Techniques

One technique for reducing the dimensionality of large plane-wave calculations is to construct effective Hamiltonians that capture correlation effects of the full calculation in manageable active spaces. That way, all many-body effects are retained as opposed to simply truncating the orbital space. In Bauman et al. (2019), we introduced downfolding techniques, which utilize the double unitary CC (DUCC) ansatz for exact ground state wave function $|\Psi\rangle$,

$$|\Psi\rangle = e^{\sigma_{ext}} e^{\sigma_{\rm int}} |\Phi\rangle\,, \tag{5}$$

where $\sigma_{\rm int}$ and σ_{ext} are the general type anti-Hermitian operators

$$\sigma_{\rm int}^{\dagger} = -\sigma_{\rm int}, \tag{6}$$

$$\sigma_{ext}^{\dagger} = -\sigma_{ext}, \tag{7}$$

defined by amplitudes defining action within and outside of the predefined active space, respectively; i.e., the amplitudes defining the σ_{ext} operator must carry at least one inactive spin-orbital index whereas all amplitudes defining the $\sigma_{\rm int}$ operator carry active spin-orbital indices only. In **Eq. 5**, $|\Phi\rangle$ designates properly chosen reference function (usually chosen as a Hartree–Fock (HF) Slater determinant). The exactness of the expansion 5 has been recently discussed in Kowalski and Bauman (2020), where it was also shown that the standard UCC expansions can provide a basic approximation of the exact $\sigma_{\rm int}$ and σ_{ext} operators, i.e.,

$$\sigma_{\rm int} \simeq T_{\rm int} - T_{\rm int}^{\dagger}, \tag{8}$$

$$\sigma_{ext} \simeq T_{ext} - T_{ext}^{\dagger}, \tag{9}$$

where $T_{\rm int}$ and T_{ext} are single-reference-type internal and external cluster amplitudes (in the sense defined above).

Using ansatz in **Eq. 5** we have shown (Kowalski and Bauman, 2020) that the exact energy of the systems can be reproduced by the diagonalization of the effective (or downfolded) Hamiltonian $\overline{H}_{eff}^{(DUCC)}$ in the corresponding active space:

$$\overline{H}_{eff}^{(DUCC)} e^{\sigma_{\rm int}} |\Phi\rangle = E e^{\sigma_{\rm int}} |\Phi\rangle, \tag{10}$$

where

$$\overline{H}_{eff}^{(DUCC)} = (P + Q_{\rm int}) e^{-\sigma_{ext}} H e^{\sigma_{ext}} (P + Q_{\rm int})\,. \tag{11}$$

In **Eq. 11**, P and $Q_{\rm int}$ are the projection operators onto the reference function and all active-space excited Slater determinants (with respect to $|\Phi\rangle$).

We will discuss the utility of the downfolding techniques in the next section for the ground state calculations of H_2. This is just one of the two approaches presented in this paper for reducing the dimensionality of the quantum problem (**Figure 1**).

Results for the $^1\Sigma_g^+$ Ground State of H_2 Using Virtual Space From Hartree–Fock and One-Electron Hamiltonians

The NWChem program package (Kendall et al., 2000; Valiev et al., 2010; Bylaska et al., 2011; Bylaska, 2017; Apra et al., 2020) was used for all calculations in this study, except for the FCI calculations, which used the TINYMRCC suite by Jiří

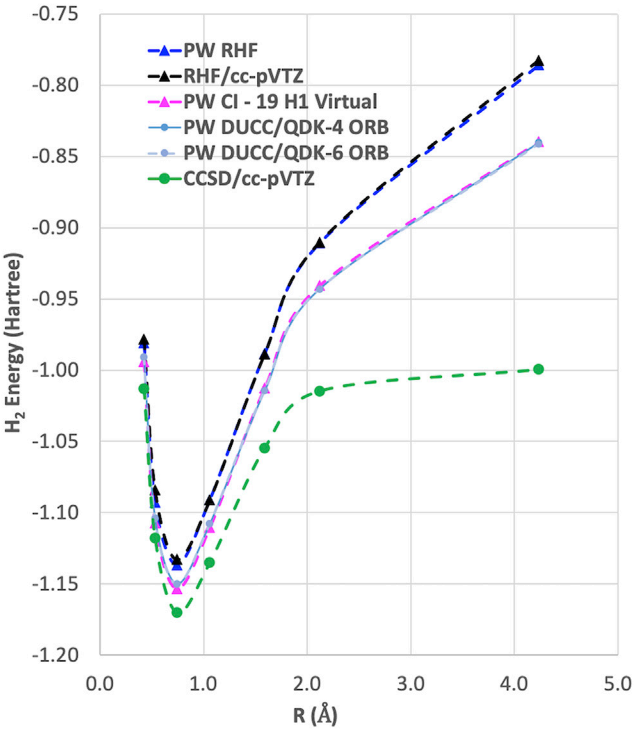


FIGURE 2 | The ground state energy curves for H_2 with RHF, CCSD, and DUCC/QDK methods using plane-wave and LCAO Gaussian basis sets. It should be noted that for the two-electron H_2 molecule CCSD gives the same answer as FCI.

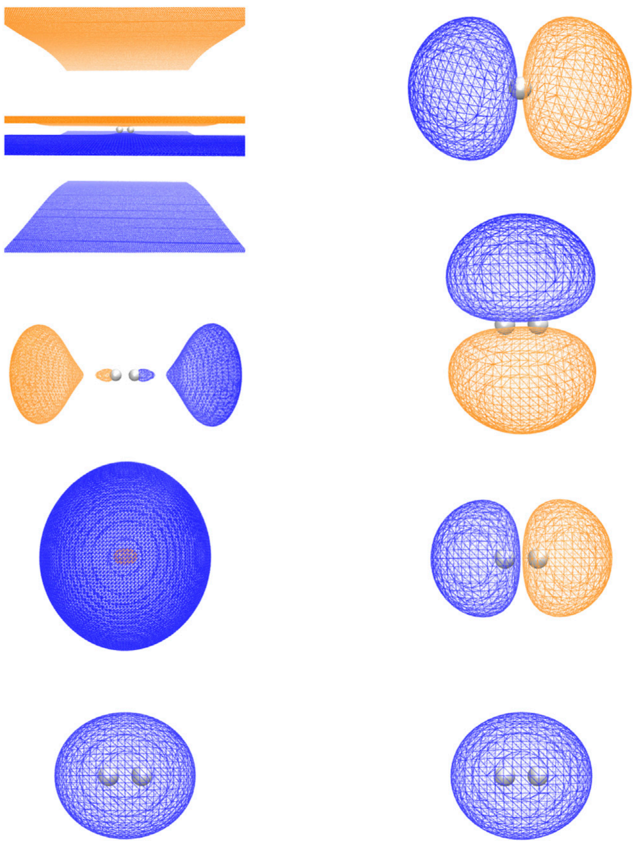


FIGURE 3 | The HOMO and first three LUMOs generated from the straight HF calculation and the H_1 Hamiltonian are shown in the left and right panels, respectively. The orbitals are displayed in the order of decreasing orbital energy from top to bottom. The isovalues of positive and negative isosurfaces are 0.007226 and −0.007226 for the RHF LUMO 1; 0.01148 and −0.01148 for the RHF LUMO 2; 0.002404 and −0.002404 for the RHF LUMO 3; and 0.03117 and −0.03117 for the others. Notice that the isovalues of RHF LUMO 1 are very close to zero, which indicates that it is a scattering state. The orientation of H_2 is rotated by 90° in the bound LUMO 3 relative to the bound LUMO 2.

Pittner. The plane-wave calculations used a simple cubic box with L = 26a_0 and cutoff energy of 100 Ry. The valence electron interactions with the atomic core are approximated with generalized norm-conserving Hamann (Hamann, 1989) pseudopotentials modified to the separable form suggested by Kleinman and Bylander (Kleinman and Bylander, 1982). The pseudopotentials used in this study were constructed using the following core radii: H: rcs = 0.8 a.u. and rcp = 0.8 a.u.; Be: rcs = 1.219 a.u., and rcp = 1.219 a.u. (vide infra). The RHF and coupled cluster singles and doubles (CCSD) LCAO calculations used the Dunning cc-pVTZ Gaussian basis set.

As can be seen in **Figure 2**, the RHF ground state energy curve of the H_2 molecule using plane-wave and LCAO Gaussian basis sets gives nearly identical results. However, when we performed plane-wave FCI calculations (not shown) for this system using up to 20 RHF virtual orbitals, the amount of correlation energy calculated was nearly zero (<1.0 e-4 Hartree). This result was not surprising since most of the virtual states were scattering states as shown in **Figure 3**.

Instead of using virtual states of the RHF Hamiltonian, virtual states were also generated using the 1-electron part of the RHF Hamiltonian, H_1 (i.e., just the kinetic energy and pseudopotential terms). As shown in **Figure 3**, the H_1 Hamiltonian generated virtual orbitals that were bound and looked like the virtual orbitals generated in an LCAO RHF calculation. Using these H_1 generated orbitals, we performed plane-wave CI calculations using 19 of these virtual orbitals. As seen in **Figure 2** and **Table 1**, a significant improvement was seen using these orbitals as they were able to capture a nonzero amount of the correlation energy; however, it was still significantly less than that found in LCAO calculations. In addition, results using the quantum phase estimation (QPE) algorithm in the Microsoft QDK package (Svore et al., 2018; Low et al., 2019) in which the number of orbitals was reduced to 4 and 6 orbitals using the DUCC method are shown. These results showed that the DUCC QDK QPE method produces total energies that are within a few milli-Hartrees of the 20 orbital FCI result with only 4 or 6 orbitals.

VARIATIONAL QUANTUM EIGENSOLVER METHODS

VQE is a method to find the quantum state that minimizes a cost function defined in operator form (Peruzzo et al., 2014; O'Malley et al., 2016). This is a hybrid computational approach in which the preparation of the quantum circuit is tuned using feedback from classical evaluations of the cost function. Reduction of a given problem to minimization, such as solving for the ground state energy (lowest energy eigenvalue) of a molecular Hamiltonian, may then rely on the variational principle to affirm that only the true ground state could satisfy the minimum energy (Kandala et al., 2017; McCaskey et al., 2019).

Formally, we may consider the problem of solving for the ground state energy, E_g, as

$$E_g = \min_{|\Psi\rangle} \langle \Psi | H | \Psi \rangle, \tag{12}$$

TABLE 1 | Total energies as a function of distance from plane-wave FCI calculations for the H_2 molecule 19 H_1 virtual orbitals.

R (H-H) (Å)	PW FCI 19 H_1 Virt.	PW QDK DUCC 4	PW QDK DUCC 6	CCSD cc-pVTZ
0.423	−0.99396	−0.99113	−0.99052	−1.01540
0.529	−1.10715	−1.10363	−1.10440	−1.12144
0.741	−1.15340	−1.14968	−1.15010	−1.17234
1.058	−1.11042	−1.10768	−1.10785	−1.13617
1.588	−1.01251	−1.01435	−1.01417	−1.05526
2.117	−0.94070	−0.94303	−0.94318	−1.01485
4.233	−0.83962	−0.84025	−0.84156	−0.99965

where H represents the second-quantized Hamiltonian of **Eq. 4** and $|\Psi\rangle$ labels the electronic configuration. Within the context of quantum computing, the fermionic representations of the Hamiltonian and state are transformed to alternate representations in terms of the spin operators. This transformation recast the molecular Hamiltonian into a representation H_P that is defined in terms of the usual Pauli spin operators. Similarly, the electronic state $|\Psi\rangle$ is represented as a variable unitary operator $\widehat{U}(\overrightarrow{\theta})$ acting on a fiducial quantum state ($|0>$). This yields the equivalent representation of **Eq. 13** as

$$E_g = \min_{\overrightarrow{\theta}} \left\langle 0 \middle| \widehat{U}^{\dagger}(\overrightarrow{\theta}) H_P \widehat{U}(\overrightarrow{\theta}) \middle| 0 \right\rangle. \tag{13}$$

The second equality is pertinent to the current context as it makes explicit the fact that 1) H_P, the Hamiltonian in terms of strings of Pauli operators, relates to H through some transformation that maps fermionic creation and annihilation operators to qubits operators, and 2) the trial wave function emerges from the action of the parameterized unitary operator $\widehat{U}(\overrightarrow{\theta})$ which builds entanglement, usually starting from an unentangled wave function, such as Hartree–Fock. For practical considerations, H is transformed into H_P with the Jordan-Wigner transformation (Jordan and Wigner, 1928), but alternatives have been reported in the literature (Bravyi and Kitaev, 2002; Setia and Whitfield, 2018; Steudtner and Wehner, 2018). The classical search for the quantum state that minimizes the energy represents the conventional computing task, while evaluation of the expectation value is performed using the quantum computer. In particular, the quantum state is prepared by executing a quantum circuit, which is expressed formally as a series of unitary operators acting on a well-defined initial state. The action of a specific sequence of unitaries is to prepare a given state that is subject to the measurements necessary to recover the desired expectation value.

In practice, the quantum state that minimizes the energy is unknown, and, therefore, a search over possible unitaries is necessary to find the form that minimizes the energy. This variational approach to circuit synthesis underlies the VQE method and an essential choice is the selection of a quantum circuit ansatz which defines the range of unitaries that may be formed to prepare and evaluate a quantum state. For electronic structure calculations, seemingly randomized unitaries may offer advantages for efficient circuit construction, but they lack much of the intuition available from theoretical chemistry (Kandala et al., 2017). Rather, ansatz circuits derived from unitary coupled cluster theory offer a convenient connection to the expected unitary forms of the minimal quantum state (Romero et al., 2017).

VQE has been applied previously to recover the electronic energy from the Hamiltonian presented in **Eq. 4**. The literature provides several examples of usage of VQE for problems of chemical interest, in terms of both simulation and implementation on actual quantum hardware. Given the current limitations faced by present quantum computers, these instances are usually accompanied by strategies that reduce the effective Hilbert space, thus leading to a decrease in the computational expense, such as the use of active spaces and natural orbitals (Verma et al., 2020), as well as downfolding techniques introduced earlier. Another route is to modify the form of the ansatz; an example of this alternative would be the so-called Trotterization, which can be used in conjunction with Hilbert space-reducing techniques.

Recently, the principle of VQE was extended to use ansatz circuits that are tailored to computational chemistry applications and specifically the unitary coupled cluster singles and doubles (UCCSD) ansatz state. Adaptive ansatz construction is attractive because it obeys the underlying complexity of the electronic structure in question, whereas a predefined form for the trial wave function in **Eq. 13** may fall short of the flexibility necessary for intricate problems. The prime example of this class of algorithms is the ADAPT-VQE, which iteratively assembles a circuit according to the expected energy gain signaled by the gradient with respect to the variational parameters.

An important consideration in the performance of both VQE and ADAPT-VQE is the depth of the ansatz circuit and the time required to construct the optimal variational circuit. For electronic structures dominated by weak correlation, ADAPT-VQE tends to be very economical, adding only operators that make a meaningful contribution toward the lowest eigenvalue in the spectrum of the Hamiltonian in **Eq. 14**. On the other hand, the usual UCCSD, by virtue of being defined ahead of time, may contain operators with little impact on the energy, but the classical optimizer will still need to perform a number of calls to the cost function in order to find their best values. Also, the gates originating from these operators, even if they are deemed unimportant because of a small associated parameter, will nevertheless be present in the circuit, adding to its depth. If high accuracy is sought, then ADAPT-VQE may require an ansatz comprised of a large number of operators, which in turn adds to the depth of the underlying circuit. More operators also mean more variational parameters, leading to an onerous optimization process. A more detailed analysis of this trade-off can be found in Grimsley et al. (2019).

ALGORITHM FOR DEFINING A VIRTUAL SPACE WITH A SMALL CI HAMILTONIAN

In this section, we present a downfolding method to define virtual orbitals for expanding the second-quantized Hamiltonian given in **Eq. 4**. These new types of orbitals are able to capture significantly more correlation energy than the virtual orbitals coming from Hartree–Fock and one-electron Hamiltonians tested in **Section 2.2**. The basis of this method is to define a set of virtual orbitals, $\{\psi_e^{(n)}(r)\}$ with $n = 1 \ldots N_{virtual}$, which we call correlation optimized virtual orbitals or COVOs for short, by optimizing a small select CI Hamiltonian with respect to a single virtual orbital and then the next virtual orbitals in sequence, subject to them being orthonormal to the filled and previously computed virtual orbitals. The algorithm to calculate these new types of orbitals can be formulated as follows:

1. Set $n = 1$.
2. Using the ground state one-electron orbital, $\psi_g(r)$ (or ground state orbitals for many-electron systems), and the virtual orbital to be optimized, $\psi_e^{(n)}(r)$, generate a CI matrix.
3. Calculate the select CI expansion coefficients by diagonalizing the CI matrix.
4. Using the CI coefficients associated with the lowest eigenvalue, calculate the gradient with respect to the $\psi_e^{(n)}(r)$ and then update with a conjugate-gradient or similar method while making sure that $\psi_e^{(n)}(r)$ is normalized and orthogonal to $\psi_g(r)$ and $\psi_e^{(m)}(r)$ for $m = 1, \ldots, n-1$.
5. If the gradient is small, then $n = n + 1$.
6. If $n \le N_{virtual}$, go to step 2; otherwise, finish.

In the case of the H_2 molecule, a small CI wave function for the 2 electron system composed of 2 one-electron orbitals, $\psi_g(r)$ and $\psi_e^{(n)}(r)$, can be written as a linear combination of 6 determinant wave functions, or just 3 determinant wave functions for just singlet (or triplet) states,

$$\begin{aligned}\Psi_i\left[\psi_g(r), \psi_e(r)\right] = c_g^{(i)}\Psi_g\left[\psi_g(r)\right] \\ + c_e^{(i)}\Psi_e\left[\psi_e(r)\right] \\ + c_m^{(i)}\Psi_m\left[\psi_g(r), \psi_e(r)\right] + \ldots.\end{aligned}$$

Using this small CI ansatz, the energies of the system can be obtained by diagonalizing the following eigenvalue equation:

$$HC_i = E_i S C_i,$$

where

$$\begin{aligned} H &= \begin{bmatrix} \langle\Psi_g|H|\Psi_g\rangle & \langle\Psi_g|H|\Psi_e\rangle & \langle\Psi_g|H|\Psi_m\rangle \\ \langle\Psi_e|H|\Psi_g\rangle & \langle\Psi_e|H|\Psi_e\rangle & \langle\Psi_e|H|\Psi_m\rangle \\ \langle\Psi_m|H|\Psi_g\rangle & \langle\Psi_m|H|\Psi_e\rangle & \langle\Psi_m|H|\Psi_m\rangle \end{bmatrix}, \\ S &= \begin{bmatrix} \langle\Psi_g|\Psi_g\rangle & \langle\Psi_g|\Psi_e\rangle & \langle\Psi_g|\Psi_m\rangle \\ \langle\Psi_e|\Psi_g\rangle & \langle\Psi_e|\Psi_e\rangle & \langle\Psi_e|\Psi_m\rangle \\ \langle\Psi_m|\Psi_g\rangle & \langle\Psi_m|\Psi_e\rangle & \langle\Psi_m|\Psi_m\rangle \end{bmatrix}, \\ C_i &= \begin{bmatrix} c_g^{(i)} \\ c_e^{(i)} \\ c_m^{(i)} \end{bmatrix}. \end{aligned} \tag{14}$$

Note that the overlap matrix, S, is the identity matrix for orthonormal ψ_g and ψ_e. The variation with respect to $\psi_e(r)$ can be simply obtained using the following formula:

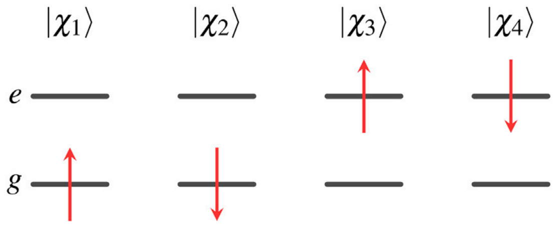
$$\begin{aligned}\frac{\delta E_i}{\delta\psi_e^*(r)} = {} & c_g^{(i)}\frac{\delta\langle\Psi_g|H|\Psi_g\rangle}{\delta\psi_e^*(r)}c_g^{(i)} + c_g^{(i)}\frac{\delta\langle\Psi_g|H|\Psi_e\rangle}{\delta\psi_e^*(r)}c_e^{(i)} \\ & + c_g^{(i)}\frac{\delta\langle\Psi_g|H|\Psi_m\rangle}{\delta\psi_e^*(r)}c_m^{(i)} + c_e^{(i)}\frac{\delta\langle\Psi_e|H|\Psi_g\rangle}{\delta\psi_e^*(r)}c_g^{(i)} \\ & + c_e^{(i)}\frac{\delta\langle\Psi_e|H|\Psi_e\rangle}{\delta\psi_e^*(r)}c_e^{(i)} + c_e^{(i)}\frac{\delta\langle\Psi_e|H|\Psi_m\rangle}{\delta\psi_e^*(r)}c_m^{(i)} \\ & + c_m^{(i)}\frac{\delta\langle\Psi_m|H|\Psi_g\rangle}{\delta\psi_e^*(r)}c_g^{(i)} + c_m^{(i)}\frac{\delta\langle\Psi_m|H|\Psi_e\rangle}{\delta\psi_e^*(r)}c_e^{(i)} \\ & + c_m^{(i)}\frac{\delta\langle\Psi_m|H|\Psi_m\rangle}{\delta\psi_e^*(r)}c_m^{(i)}.\end{aligned} \tag{15}$$

It should be noted that the above formulas can be generalized to work beyond two-electron systems by using corresponding orbitals techniques (King et al., 1967; Bylaska and Rosso, 2018). The next two **Subsections 4.1–4.4** provide formulas that can be used to generate the matrix elements in **Eq. 14** and the gradients with respect to $\psi_e^*(r)$ in **Eq. 15**.

We also note that the COVO approach proposed in this work is similar in spirit to the optimized virtual orbital space (OVOS) approach developed over 30 years ago by Adamowicz and Bartlett (Adamowicz and Bartlett, 1987; Adamowicz et al., 1988). The differences in our approach compared to this previous work is that the variational space used by COVOs is significantly bigger because plane-wave basis sets are used instead of LCAO Gaussian basis sets and that a second-order Hylleraas functional (Hylleraas, 1928; Hylleraas, 1929; Hylleraas, 1930; Hylleraas, 1964; Koga, 1992) was used to describe the correlation in the OVOS procedure rather than a small CI Hamiltonian. Other differences with the COVOs approach are that the orbitals are optimized one at a time and the cost to generate them is similar to generating regular RHF virtual orbitals (just 4 to 9 times more expensive relative to RHF). Moreover, the resulting electronic gradient is non-Hermitian, which in addition to requiring more involved optimizers can result in extended energy plateaus that occur during the initial stages of the geodesic line searches in a conjugate-gradient or quasi-Newton optimization method.

One-Electron Orbitals for Two-State Hamiltonian

The four one-electron spin orbitals of two-state Hamiltonian are



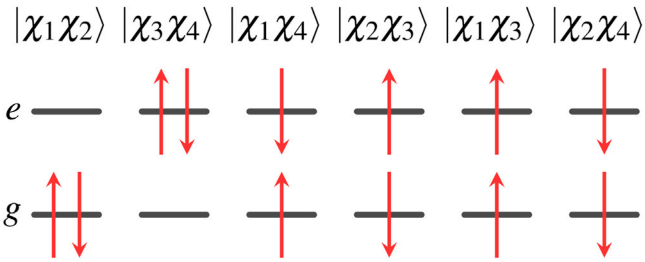
$$\begin{aligned}\chi_1(x) &= \psi_g(r)\alpha(s), \\ \chi_2(x) &= \psi_g(r)\beta(s), \\ \chi_3(x) &= \psi_e(r)\alpha(s), \\ \chi_4(x) &= \psi_e(r)\beta(s),\end{aligned}$$

where the spatial orbitals and spin functions are orthonormalized,

$$\int \psi_g^*(r)\psi_e(r)dr = \int \psi_e^*(r)\psi_g(r)dr = 0,$$
$$\int \psi_g^*(r)\psi_g(r)dr = \int \psi_e^*(r)\psi_e(r)dr = 1,$$
$$\int \alpha^*(s)\beta(s)ds = \int \beta^*(s)\alpha(s)ds = 0,$$
$$\int \alpha^*(s)\alpha(s)ds = \int \beta^*(s)\beta(s)ds = 1.$$

Two-Electron Orbitals for a Two-State Hamiltonian

For the two-state system, there are six two-electron wave functions, two of which are singlet, two are triplet, and two contain a mixture of singlet and triplet character. These wave functions can be written as



$$|\chi_1\chi_2\rangle = \Psi_g(x_1,x_2)$$
$$= \left[\psi_g(r_1)\psi_g(r_2)\right]\frac{1}{\sqrt{2}}\left[\alpha(s_1)\beta(s_2) - \alpha(s_2)\beta(s_1)\right],$$
$$|\chi_3\chi_4\rangle = \Psi_e(x_1,x_2)$$
$$= \left[\psi_e(r_1)\psi_e(r_2)\right]\frac{1}{\sqrt{2}}\left[\alpha(s_1)\beta(s_2) - \alpha(s_2)\beta(s_1)\right],$$
$$|\chi_1\chi_4\rangle = \Psi_a(x_1,x_2)$$
$$= \frac{1}{\sqrt{2}}\begin{bmatrix} \psi_g(r_1)\alpha(s_1)\psi_e(r_2)\beta(s_2) \\ -\psi_e(r_1)\beta(s_1)\psi_g(r_2)\alpha(s_2) \end{bmatrix},$$
$$|\chi_2\chi_3\rangle = \Psi_b(x_1,x_2)$$
$$= \frac{1}{\sqrt{2}}\begin{bmatrix} \psi_g(r_1)\beta(s_1)\psi_e(r_2)\alpha(s_2) \\ -\psi_e(r_1)\alpha(s_1)\psi_g(r_2)\beta(s_2) \end{bmatrix},$$
$$|\chi_1\chi_3\rangle = \Psi_u(x_1,x_2)$$
$$= \frac{1}{\sqrt{2}}\left[\psi_g(r_1)\psi_e(r_2) - \psi_e(r_1)\psi_g(r_2)\right]\left[\alpha(s_1)\alpha(s_2)\right],$$
$$|\chi_2\chi_4\rangle = \Psi_d(x_1,x_2)$$
$$= \frac{1}{\sqrt{2}}\left[\psi_g(r_1)\psi_e(r_2) - \psi_e(r_1)\psi_g(r_2)\right]\left[\beta(s_1)\beta(s_2)\right].$$

Note that Ψ_a and Ψ_b cannot be written as a product of a spatial wave function times a spin function. Moreover, these functions are not eigenfunctions of the spin operators S^2 and S_z, and as a result, these determinants contain both singlet and triplet components. However, if we take linear combinations of them, we can get two new wave functions that are separable in spatial and spin functions and at the same time being eigenfunctions of S^2 and S_z,

$$\Psi_m = \Psi_{a-b} = \frac{1}{\sqrt{2}}\left(|\chi_1\chi_4\rangle - |\chi_2\chi_3\rangle\right)$$
$$= \frac{1}{\sqrt{2}}\left[\Psi_a(x_1,x_2) - \Psi_b(x_1,x_2)\right]$$
$$= \frac{1}{\sqrt{2}}\left[\psi_g(r_1)\psi_e(r_2) + \psi_e(r_1)\psi_g(r_2)\right]$$
$$\times\frac{1}{\sqrt{2}}\left[\alpha(s_1)\beta(s_2) - \beta(s_1)\alpha(s_2)\right],$$
$$\Psi_p = \Psi_{a+b} = \frac{1}{\sqrt{2}}\left(|\chi_1\chi_4\rangle + |\chi_2\chi_3\rangle\right)$$
$$= \frac{1}{\sqrt{2}}\left[\Psi_a(x_1,x_2) + \Psi_b(x_1,x_2)\right]$$
$$= \frac{1}{\sqrt{2}}\left[\psi_g(r_1)\psi_e(r_2) - \psi_e(r_1)\psi_g(r_2)\right]$$
$$\times\frac{1}{\sqrt{2}}\left[\alpha(s_1)\beta(s_2) + \beta(s_1)\alpha(s_2)\right].$$

Matrix Elements From the One-Electron Operators

The H_1 operator for H_2 molecule is

$$H_1 = h(r_1) + h(r_2),$$

where $h(r)$ is a function/operator of the coordinate r; i.e.,

$$h(r) = -\frac{1}{2}\nabla_r^2 + \sum_{A=1}^{N_A}\left(V_{local}^{(A)}(|r - R_A|) + \sum_{lm}\widehat{V}_{NL}^{(A),lm}\right),$$
$$\langle\Psi_g|H_1|\Psi_g\rangle = 2\int \psi_g^*(r)h(r)\psi_g(r)dr,$$
$$\langle\Psi_g|H_1|\Psi_e\rangle = 0, \langle\Psi_e|H_1|\Psi_e\rangle = 2\int \psi_e^*(r)h(r)\psi_e(r)dr,$$
$$\langle\Psi_m|H_1|\Psi_m\rangle = \int \psi_g^*(r)h(r)\psi_g(r)dr + \int \psi_e^*(r)h(r)\psi_e(r)dr,$$
$$\langle\Psi_g|H_1|\Psi_m\rangle = \sqrt{2}\int \psi_g^*(r)h(r)\psi_e(r)dr,$$
$$\langle\Psi_m|H_1|\Psi_g\rangle = \sqrt{2}\int \psi_e^*(r)h(r)\psi_g(r)dr,$$
$$\langle\Psi_e|H_1|\Psi_m\rangle = \sqrt{2}\int \psi_e^*(r)h(r)\psi_g(r)dr,$$
$$\langle\Psi_m|H_1|\Psi_e\rangle = \sqrt{2}\int \psi_g^*(r)h(r)\psi_e(r)dr,$$
$$\frac{\delta\langle\Psi_g|H_1|\Psi_g\rangle}{\delta\psi_e^*(r)} = 0,$$
$$\frac{\delta\langle\Psi_g|H_1|\Psi_e\rangle}{\delta\psi_e^*(r)} = 0,$$
$$\frac{\delta\langle\Psi_e|H_1|\Psi_e\rangle}{\delta\psi_e^*(r)} = 2h(r)\psi_e(r),$$
$$\frac{\delta\langle\Psi_m|H_1|\Psi_m\rangle}{\delta\psi_e^*(r)} = h(r)\psi_e(r),$$
$$\frac{\delta\langle\Psi_g|H_1|\Psi_m\rangle}{\delta\psi_e^*(r)} = 0,$$
$$\frac{\delta\langle\Psi_m|H_1|\Psi_g\rangle}{\delta\psi_e^*(r)} = \sqrt{2}h(r)\psi_g(r),$$
$$\frac{\delta\langle\Psi_e|H_1|\Psi_m\rangle}{\delta\psi_e^*(r)} = \sqrt{2}h(r)\psi_g(r),$$
$$\frac{\delta\langle\Psi_m|H_1|\Psi_e\rangle}{\delta\psi_e^*(r)} = 0.$$

Matrix Elements From the Two-Electron Operators

The H_2 operator for H_2 molecule is

$$H_2 = \frac{1}{r_{12}} = \frac{1}{|r_1 - r_2|},$$

$$\langle \Psi_g | H_2 | \Psi_g \rangle = \iint \psi_g^*(r)\psi_g(r)\frac{1}{|r-r'|}\psi_g^*(r')\psi_g(r')drdr',$$

$$\langle \Psi_g | H_2 | \Psi_e \rangle = \iint \psi_g^*(r)\psi_e(r)\frac{1}{|r-r'|}\psi_g^*(r')\psi_e(r')drdr',$$

$$\langle \Psi_e | H_2 | \Psi_g \rangle = \iint \psi_e^*(r)\psi_g(r)\frac{1}{|r-r'|}\psi_e^*(r')\psi_g(r')drdr',$$

$$\langle \Psi_e | H_2 | \Psi_e \rangle = \iint \psi_e^*(r)\psi_e(r)\frac{1}{|r-r'|}\psi_e^*(r')\psi_e(r')drdr',$$

$$\langle \Psi_m | H_2 | \Psi_m \rangle = \left[\iint \psi_e^*(r)\psi_e(r)\frac{1}{|r-r'|}\psi_g^*(r')\psi_g(r')drdr' + \iint \psi_e^*(r)\psi_g(r)\frac{1}{|r-r'|}\psi_g^*(r')\psi_e(r')drdr'\right],$$

$$\langle \Psi_g | H_2 | \Psi_m \rangle = \sqrt{2}\iint \psi_g^*(r)\psi_g(r)\frac{1}{|r-r'|}\psi_g^*(r')\psi_e(r')drdr',$$

$$\langle \Psi_m | H_2 | \Psi_g \rangle = \sqrt{2}\iint \psi_e^*(r)\psi_g(r)\frac{1}{|r-r'|}\psi_g^*(r')\psi_g(r')drdr',$$

$$\langle \Psi_e | H_2 | \Psi_m \rangle = \sqrt{2}\iint \psi_e^*(r)\psi_e(r)\frac{1}{|r-r'|}\psi_e^*(r')\psi_g(r')drdr',$$

$$\langle \Psi_m | H_2 | \Psi_e \rangle = \sqrt{2}\iint \psi_g^*(r)\psi_e(r)\frac{1}{|r-r'|}\psi_e^*(r')\psi_e(r')drdr',$$

$$\frac{\delta\langle \Psi_g | H_2 | \Psi_g \rangle}{\delta\psi_e^*(r)} = 0,$$

$$\frac{\delta\langle \Psi_g | H_2 | \Psi_e \rangle}{\delta\psi_e^*(r)} = 0,$$

$$\frac{\delta\langle \Psi_e | H_2 | \Psi_g \rangle}{\delta\psi_e^*(r)} = 2\left[\int \frac{\psi_e^*(r')\psi_g(r')}{|r-r'|}dr'\right]\psi_g(r),$$

$$\frac{\delta\langle \Psi_e | H_2 | \Psi_e \rangle}{\delta\psi_e^*(r)} = 2\left[\int \frac{\psi_e^*(r')\psi_e(r')}{|r-r'|}dr'\right]\psi_e(r),$$

$$\frac{\delta\langle \Psi_m | H_2 | \Psi_m \rangle}{\delta\psi_e^*(r)} = \left[\int \frac{\psi_g^*(r')\psi_g(r')}{|r-r'|}dr'\right]\psi_e(r) + \left[\int \frac{\psi_g^*(r')\psi_e(r')}{|r-r'|}dr'\right]\psi_g(r),$$

$$\frac{\delta\langle \Psi_g | H_2 | \Psi_m \rangle}{\delta\psi_e^*(r)} = 0,$$

$$\frac{\delta\langle \Psi_m | H_2 | \Psi_g \rangle}{\delta\psi_e^*(r)} = \sqrt{2}\left[\int \frac{\psi_g^*(r')\psi_g(r')}{|r-r'|}dr'\right]\psi_g(r),$$

$$\frac{\delta\langle \Psi_e | H_2 | \Psi_m \rangle}{\delta\psi_e^*(r)} = \sqrt{2}\left\{\left[\int \frac{\psi_e^*(r')\psi_g(r')}{|r-r'|}dr'\right]\psi_e(r) + \left[\int \frac{\psi_e^*(r')\psi_e(r')}{|r-r'|}dr'\right]\psi_g(r)\right\},$$

$$\frac{\delta\langle \Psi_m | H_2 | \Psi_e \rangle}{\delta\psi_e^*(r)} = \sqrt{2}\left[\int \frac{\psi_g^*(r')\psi_e(r')}{|r-r'|}dr'\right]\psi_e(r).$$

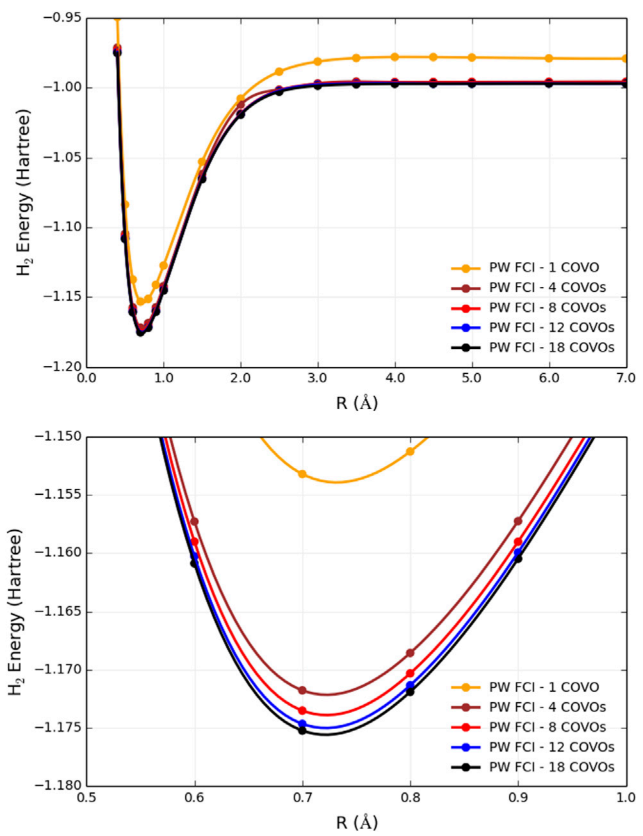


FIGURE 4 | Plots of total energies as a function of distance from plane-wave FCI calculations for the H_2 molecule with 1, 4, 8, 12, and 18 correlation optimized virtual orbitals. The top plot shows energy from R = 0.4 Å to R = 7.0 Å, and the bottom plot zooms in near the energy minima.

RESULTS FOR $^1\Sigma_g^+$ GROUND STATE OF THE H_2 MOLECULE USING CORRELATION OPTIMIZED VIRTUAL ORBITALS (COVOS)

The results for PW FCI calculations of H_2 with 1, 4, 8, 12, and 18 COVOs are shown in **Figure 4** and **Table 2**. The average difference error for the 1, 4, 8, and 12 COVOs calculations from the 18 COVOs calculation is 11.8 kcal/mol, 1.4 kcal/mol, 0.9 kcal/mol, and 0.3 kcal/mol, respectively. While the error is significant for 1 virtual, the difference is quite small by 4 virtual orbitals, and the error steadily decreases as the number of virtual orbitals increases. The error seen in the 4 optimized virtual orbitals' calculations is similar to the 1.6 kcal/mol error seen in the DUCC calculations for the 19 H_1 virtual orbitals' calculations in **Section 2.2**. Another measure of the error is the extensivity error. The energy for large R should be the same as the energy of twice the energy of an isolated H atom. For the pseudopotential plane-wave method being used, the energy of 2 H atoms is −0.997765 Hartrees (E(1H) = −0.498825 Hartrees). This difference at R = 7 Å is found to be 11.6, 1.2, 1.1, 0.5, and 0.4 kcal/mol for 1, 4, 8, 12, and 18 optimized virtual orbital calculations, respectively.

TABLE 2 | Total energies as a function of distance for the H_2 molecule from plane-wave FCI calculations with 1, 4, 8, 12, and 18 COVOs and ADAPT-VQE simulations with 4 COVOs. Nonparallelity errors (NPE) are evaluated with respect to the calculations with the largest virtual orbital space (PW FCI 18 COVOs) and reported in milli-Hartree.

R (H-H)	PW FCI	PW FCI	PW VQE	PW FCI	PW FCI	PW FCI
(Å)	1 COVO	4 COVOs	4 COVOs	8 COVOs	12 COVOs	18 COVOs
0.60	−1.13749	−1.15729	−1.15728	−1.15902	−1.16028	−1.16089
0.70	−1.15321	−1.17179	−1.17178	−1.17353	−1.17467	−1.17525
0.80	−1.15128	−1.16858	−1.16857	−1.17033	−1.17136	−1.17192
0.90	−1.14124	−1.15726	−1.15724	−1.15903	−1.15995	−1.16049
1.00	−1.12742	−1.14216	−1.14213	−1.14399	−1.14478	−1.14533
1.50	−1.05311	−1.06195	−1.06195	−1.06473	−1.06516	−1.06564
2.00	−1.00793	−1.01225	−1.01220	−1.01868	−1.01916	−1.01945
2.50	−0.98862	−1.00150	−1.00150	−1.00195	−1.00228	−1.00301
3.00	−0.98137	−0.99704	−0.99701	−0.99737	−0.99789	−0.99872
3.50	−0.97883	−0.99573	−0.99570	−0.99629	−0.99698	−0.99766
4.00	−0.97810	−0.99613	−0.99611	−0.99614	−0.99693	−0.99736
4.50	−0.97817	−0.99609	−0.99608	−0.99609	−0.99722	−0.99729
5.00	−0.97845	−0.99604	−0.99598	−0.99603	−0.99716	−0.99727
6.00	−0.97906	−0.99597	−0.99596	−0.99597	−0.99705	−0.99719
7.00	−0.97928	−0.99596	−0.99596	−0.99596	−0.99703	−0.99717
NPE	11.88	6.00	6.04	1.10	0.76	—

With only 4 optimized virtual orbitals, the correlation energy at the minimum was found to be −0.035 Hartrees, which is comparable to the −0.039 Hartrees found with CCSD/cc-pVTZ. The correlation energy decreases to −0.037, then −0.038, and finally −0.039 Hartrees as the number of optimized virtual orbitals increases to 8, 12, and 18 orbitals, respectively. These results showed that by 18 COVOs the same amount of correlation energy was recovered as with the cc-pVTZ LCAO basis set calculation. The results also showed that there was systematic convergence toward the benchmark LCAO result as the number of COVOs was increased, and with just 4 COVOs, a significant portion of the correlation energy was recovered.

QUANTUM SIMULATIONS OF THE $^1\Sigma_g^+$ GROUND STATE OF THE H_2 MOLECULE USING COVOS

The previous section provides indisputable evidence for the performance of the proposed virtual orbitals for correlation energy recovery. Besides the possible ramifications in quantum chemistry carried out with classical computers, one immediate application is in the realm of quantum simulations. Because the present quantum hardware has not fully matured, hybrid algorithms that leverage classical resources and restrict the workload delegated to quantum computers, namely, state preparation and measurements of highly entangled states, are essential to meaningful quantum computations. The COVOs meet this requirement by decreasing the dimensionality of the problem, i.e., by enabling simulations with fewer qubits.

In order to probe the performance of COVOs in quantum simulations, we use the Hamiltonian with 4 COVOs and simulate the $^1\Sigma_g^+$ ground state of H_2 in the same bond distances shown in **Table 2** and **Figure 4**. The ansatz circuit for the simulations is generated according to the ADAPT-VQE algorithm as implemented in the XACC (McCaskey et al., 2018b, McCaskey et al., 2020) framework for hybrid quantum computing using the tensor network quantum virtual machine (TNQVM) as the noiseless simulator backend (McCaskey et al., 2018a). In the present study, the ADAPT-VQE cycle is repeated until the norm of the gradient vector falls below 1e-2 and we use an operator pool containing all spin-adapted single and double excitation operators (one- and two-body rotations). A detailed account of ADAPT-VQE is exposed elsewhere (Grimsley et al., 2019).

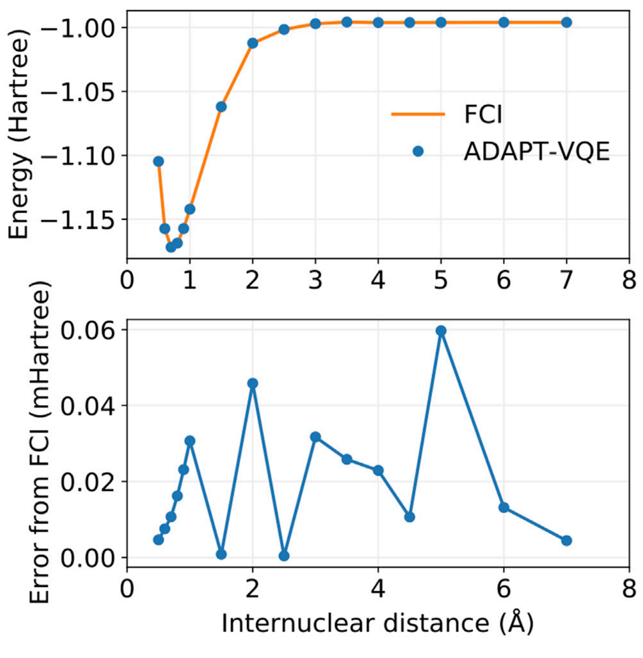


FIGURE 5 | Potential energy curves for FCI and ADAPT-VQE **(top)** and the deviations in ADAPT-VQE energies with respect to FCI **(bottom)**.

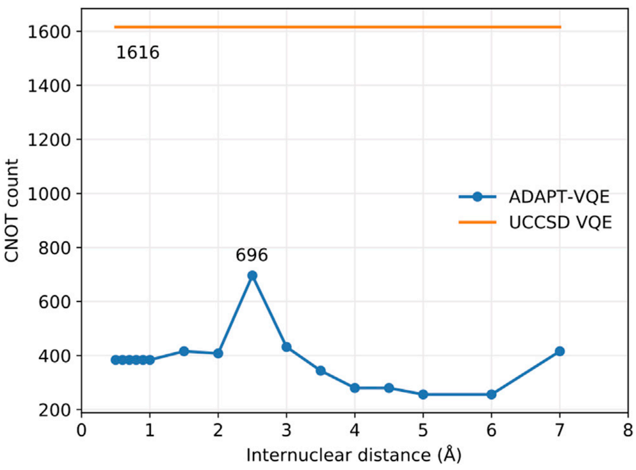


FIGURE 6 | CNOT gate count for ADAPT-VQE and ordinary VQE with both singlet-adapted singles and doubles operators.

Optimization of the parameterized circuit is conducted with the COBYLA (Powell, 1994) optimizer as implemented in the NLOpt package (Johnson, 2007). Results for the simulated potential energy curve are plotted in **Figure 5**.

It is evident from **Figure 5** that ADAPT-VQE can generate a circuit capable of reproducing the FCI results in the current active space. These simulations deliver a smooth, continuous potential energy curve that tracks the FCI values strikingly well. The deviations from the corresponding FCI energies are all found below 1e-4 Hartree. This means that not only do these simulations deliver results that are well below the conventional chemical accuracy mark, but also more importantly in the current context is that this error is inconsequential compared to the effect of noise in case of deployment on actual quantum hardware.

It is remarkable that the results degrade little throughout the energy scan, which attests to the aptness and flexibility of ADAPT-VQE in determining an ansatz according to the complexity of the underlying electronic structure. The ansatz in the vicinity of the equilibrium bond length 0.5–1.0 Å is comprised solely by pair excitations as would be expected given a restricted HF reference, which means no determinant obtained via one-body rotations can lower the energy below that of HF. As we approach the Coulson-Fischer point (Coulson and Fischer, 1949), single excitations start to become part of the ansatz, which signals the inadequacy of a restricted reference wave function and that inclusion of these operators enables the ansatz to remain in the ${}^1\Sigma_g^+$ potential energy curve, which means that this flexibility may come at the expense of deeper circuits. Because one-qubit gates tend to be executed in a short timescale and are fairly insensitive to noise, we can use the number of CNOTs present in the circuit as indicative of the complexity in its implementation, which we provide in **Figure 6**, showing that the ansatzes generated from ADAPT-VQE are much more affordable than those obtained by ordinary UCCSD VQE simulations.

Along these lines, once the operator composition of the ansatz is defined, by virtue of introducing more parameters, we are likely to experience a more arduous optimization of the corresponding parameterized gates. This has a compound effect with the circuit depth since more measurements are needed, each of which requires the circuit to be implemented and measurements to take place. **Figure 7** gives a profile of the optimization performance along the potential energy scan.

It should come as no surprise that the optimization is more difficult in the regime of stronger correlation. This region also demands a more complex ansatz, as the top plot in **Figure 7** shows that only in this vicinity (1.5–3.5 Å) we observe ansatzes with more than four operators. Interestingly, the number of objective function calls does not show large deviations for ansatzes with 1–3 operators, regardless of where they are found in the potential energy curve, which is further corroborated by the relatively small error bars in the corresponding columns of the bottom plot. This observation does not hold as more parameters/operators are introduced in the ansatz in order to accommodate a more complex electronic structure. Thus, with four parameters, not only are more calls to the objective function needed, but also there is a more pronounced standard deviation. Ansatzes with five or more operators can only be found in the (1.5–3.5 Å), as we can see that the calls to the objective function coming from them dominate the overall number of optimization cycles. Due to the scarce occurrence of these ansatzes in the current energy scan, the corresponding statistical information that can be derived from these instances is not as reliable. All in all, this plot is valuable in lending additional insight into the resources required to perform these simulations. It is important to mention

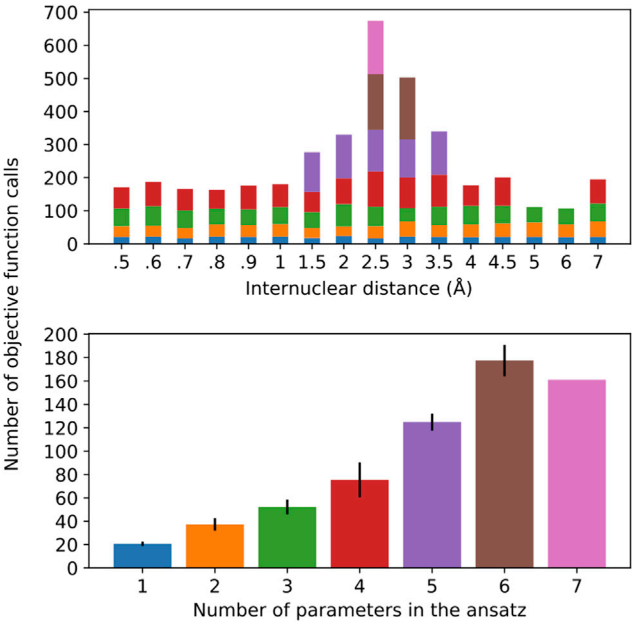


FIGURE 7 | Number of objective function calls as a function of the H-H distance for the different ansatz compositions **(top)** and the average number of objective function calls per ansatz size, with error bars representing one standard deviation **(bottom)**. The bar colors on the top plot represent the ansatz sizes in the bottom plot.

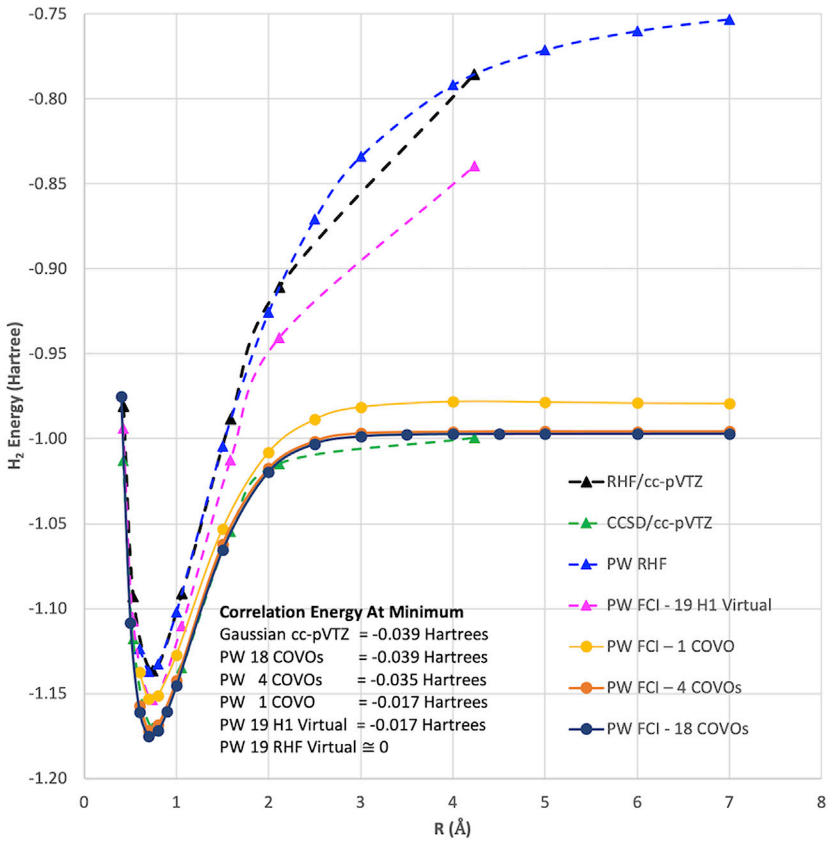


FIGURE 8 | Summary of various plane-wave and Gaussian basis set RHF and FCI calculations for the $^1\Sigma_g^+$ ground state of H_2 molecule. FCI calculations (not shown) using up to 20 RHF virtual orbitals produced only a negligible amount of correlation energy (<1.0 e-4 Hartree, i.e., visually the same as RHF results). It should be noted that for the two-electron H_2 molecule CCSD gives the same answer as FCI.

that, for every new ansatz, the variational parameters are initialized at zero. Alternatively, the parameters corresponding to the previously optimizer ansatzes could be initialized at their optimal values and the new parameter would be introduced in the ansatz, which would accelerate convergence. Moreover, the convergence profile likely displays pronounced dependence on the chosen optimizer, which is not pursued here.

CONCLUSION

In summary, we have developed a new approach for defining virtual spaces with a pseudopotential plane-wave code for use in many-body methods described by second-quantized Hamiltonians. The method is based on optimizing the virtual orbitals to minimize a small select CI Hamiltonian (i.e., COVOs) that contains configurations containing filled RHF orbitals and the one virtual orbital to be optimized. Subsequent virtual orbitals are optimized in the same way, but with the added constraint of being orthogonal to the previously calculated filled and virtual orbitals. The method was applied to the simple, but nontrivial, H_2 molecule. As summarized in **Figure 8**, these new types of virtual orbitals were significantly better at capturing correlation in plane-wave calculations than from virtual spaces from Hartree–Fock and one-electron Hamiltonian, and moreover, we were able to obtain good agreement with Gaussian cc-pVTZ basis set results

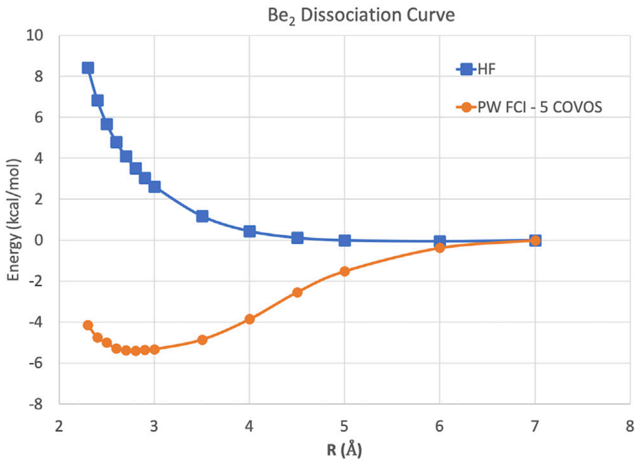


FIGURE 9 | Potential energy curves in kcal/mol for the Be_2 dimer using plane-wave Hartree–Fock and FCI with 5 COVOs calculations.

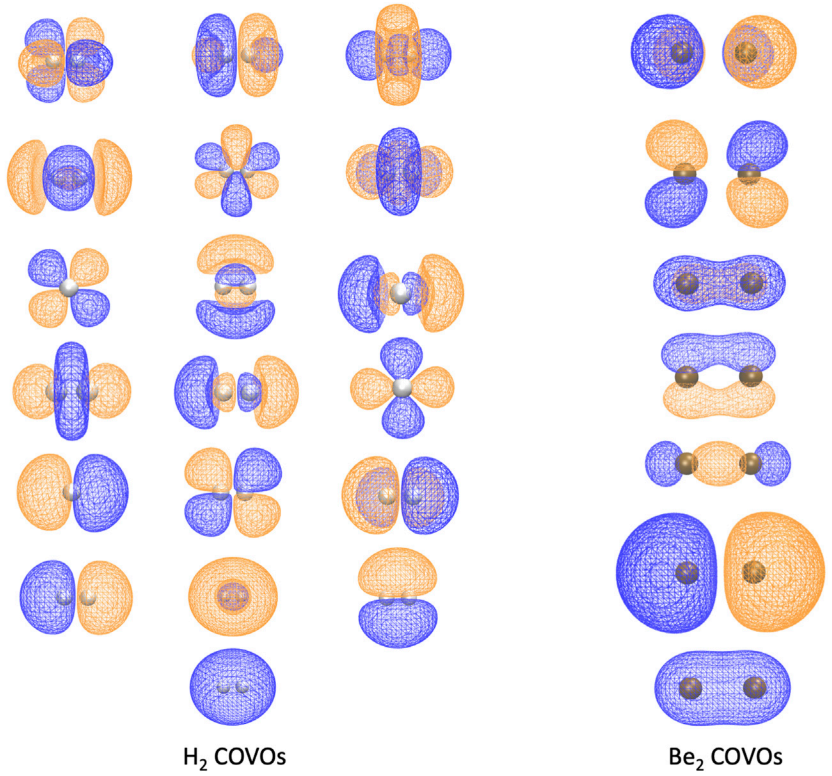


FIGURE 10 | The 1 filled RHF orbital and 18 COVOs for the H_2 molecule and 2 filled RHF orbitals and 5 COVOs for the Be_2 molecule are shown in the left and right panels, respectively. The orbitals are displayed in the order of increasing orbital energy from left to right and bottom to top. The distance between two atoms at which the energy achieves its minimum is 0.7 Å for H_2 and 2.8 Å for Be_2. The positive and negative isosurfaces are colored in blue and orange, respectively. Notice that, for H_2, the shapes of COVOs 1, 3, and 4 are the same as those of LUMOs 1, 2, and 3 from the H_1 Hamiltonian, respectively, while COVO 2 has a similar shape to LUMO 1 from the straight HF in **Figure 3** although the former is a bound state while the latter is essentially a scattering state.

with just 4 virtual orbitals for the H_2 molecule. Subsequent calculations showed that the correlation energy converged steadily as more virtual orbitals were included in the calculation. With 18 virtual orbitals, the correlation energies were found to be converged to less than 0.5 kcal/mol. The robustness of the proposed basis sets is corroborated by its ready applicability to quantum simulations, which in the case of ADAPT-VQE show remarkable agreement with the classical, exact diagonalization result (FCI) in the same basis set (4 COVOs).

Because this study is focused on how one might carry out plane-wave CI calculations on near-term quantum computers in the next few years, we have only shown results for the H_2 dimer. However, we are optimistic that these correlation optimized virtual orbitals open up the door to many-body calculations using pseudopotential plane-wave calculations, including coupled cluster, Møller–Plesset, and Green's function theories as well as other FCI-approaching methods for quantum computers. We hope in future studies to more thoroughly test the effectiveness of the COVOs procedure on larger and more complicated molecules and materials. To lend credence to this assertion, we show the promising results for Be_2 dimer with a small number of COVOs in **Figure 9**. Also as shown in **Figure 10**, the shapes of the COVOs end up being similar to what is found for the virtual orbitals from LCAO calculations. This suggests that new classes of LCAO basis sets might be able to be generated using a simple rotation of the filled orbitals and COVOs. Future work will focus on using this approach on larger molecular and periodic systems. With the validation granted by our quantum simulations, further studies are called for, including the further reduction of dimension by employing active-space DUCC downfolded Hamiltonians, OVOS, and natural orbitals, as well as work in conjunction with VQE methods.

DATA AVAILABILITY STATEMENT

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found below: https://github.com/ebylaska/PWH2-Data/archive/master.zip.

AUTHOR CONTRIBUTIONS

All authors contributed to the writing and editing of the manuscript. EB and DS developed and implemented in the NWChem the algorithm for defining a virtual space with small CI Hamiltonians. KK wrote a part of the

NWChem-DUCC interface and codeveloped the original form of the DUCC formalism. NB wrote a part of the NWChem-DUCC interface, codeveloped the original form of the DUCC formalism, carried out FCI and QPE calculations, and prepared formatted integrals for VQE calculations. DC wrote the algorithm for and performed the VQE quantum simulations.

FUNDING

This material is based upon work supported by the United States Department of Energy (DOE), Office of Science, Office of Basic Energy Sciences, Chemical Sciences, Geosciences, and Biosciences Division through its "Embedding Quantum Computing into Many-body Frameworks for Strongly Correlated Molecular and Materials Systems" project at Pacific Northwest National Laboratory (PNNL). This work was also supported by the Quantum Science Center (QSC), a National Quantum Information Science Research Center of the United States Department of Energy (DOE). We also would like to thank the DOE BES Chem CCS, DOE BES Geochemistry, and DOE Advanced Scientific Computing Research (ASCR) ECP NWChemEx programs for their support of software development for high-performance computers and computer time needed to carry out the work. PNNL is operated for the United States Department of Energy by the Battelle Memorial Institute under Contract DE-AC06-76RLO-1830. This research was also partially supported, thru their support of software development for high-performance computers, by DOE BES Chem CCS and DOE BES Geosciences programs, as well as the Exascale Computing Project (17-SC-20-SC), a collaborative effort of the United States Department of Energy Office of Science and the National Nuclear Security Administration. Calculations have been performed using computational resources at the Pacific Northwest National Laboratory (PNNL). This research used resources of the National Energy Research Scientific Computing Center (NERSC), a User Facility supported by the Office of Science of the United States DOE under Contract No. DE-AC02-05CH11231, and the Argonne ALCF computing center through their early science program. This research also used resources of the Oak Ridge Leadership Computing Facility, which is a DOE Office of Science User Facility supported under Contract DE-AC05-00OR22725. This manuscript has been authored in part by UT-Battelle, LLC under Contract No. DE-AC05-00OR22725 with the United States Department of Energy. The United States Government retains and the publisher, by accepting the article for publication, acknowledges that the United States Government retains a nonexclusive, paid-up, irrevocable, worldwide license to publish or reproduce the published form of this manuscript or allow others to do so, for United States Government purposes. The Department of Energy will provide public access to these results of federally sponsored research in accordance with the DOE Public Access Plan.

ACKNOWLEDGMENTS

We would like to thank the NWChem project team and the people that have helped the progress of the NWChem software over the years. DC would like to thank Alexander McCaskey for discussions and help with the software engineering of the quantum algorithm used in this paper.

REFERENCES

Adamowicz, L., and Bartlett, R. J. (1987). Optimized virtual orbital space for high-level correlated calculations. *J. Chem. Phys.* 86, 6314–6324. doi:10.1063/1.452468

Adamowicz, L., Bartlett, R. J., and Sadlej, A. J. (1988). Optimized virtual orbital space for high-level correlated calculations. II. Electric properties. *J. Chem. Phys.* 88, 5749–5758. doi:10.1063/1.454721

Allen, L. C., and Karo, A. M. (1962). Electronic structure of simple Molecules1. *J. Phys. Chem.* 66, 2329–2331. doi:10.1021/j100818a011

Aprà, E., Bylaska, E. J., De Jong, W. A., Govind, N., Kowalski, K., Straatsma, T. P., et al. (2020). NWChem: past, present, and future. *J. Chem. Phys.* 152, 184102. doi:10.1063/5.0004997

Bauman, N. P., Bylaska, E. J., Krishnamoorthy, S., Low, G. H., Wiebe, N., Granade, C. E., et al. (2019). Downfolding of many-body Hamiltonians using active-space models: extension of the sub-system embedding sub-algebras approach to unitary coupled cluster formalisms. *J. Chem. Phys.* 151, 014107. doi:10.1063/1.5094643

Betteridge, P. W., Carruthers, J. R., Cooper, R. I., Prout, K., and Watkin, D. J. (2003). CRYSTALSversion 12: software for guided crystal structure analysis. *J. Appl. Crystallogr.* 36, 1487. doi:10.1107/S0021889803021800

Boys, S. F., Cook, G. B., Reeves, C. M., and Shavitt, I. (1956). Automatic fundamental calculations of molecular structure. *Nature* 178, 1207–1209. doi:10.1038/1781207a0

Boys, S. F. (1950). 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. doi:10.1098/rspa.1950.0036

Bravyi, S. B., and Kitaev, A. Y. (2002). Fermionic quantum computation. *Ann. Phys.* 298, 210–226. doi:10.1006/aphy.2002.6254

Briggs, E., Sullivan, D., and Bernholc, J. (1996). Real-space multigrid-based approach to large-scale electronic structure calculations. *Phys. Rev. B Condens. Matter* 54, 14362. doi:10.1103/PhysRevB.54.14362

Bylaska, E. J., and Rosso, K. (2018). Corresponding orbitals derived from periodic bloch states for electron transfer calculations of transition metal oxides. *J. Chem. Theor. Comput.* 14, 4416–4426. doi:10.1021/acs.jctc.7b01180

Bylaska, E. J., Aprà, E., Kowalski, K., Jacquelin, M., De Jong, W. A., Vishnu, A., et al. (2017a). "Transitioning NWChem to the next generation of manycore machines," in *Exascale scientific applications: scalability and performance portability*. Editors T. P. Straatsma, K. B. Antypas, and T. J. J. Timothy Williams (Boca Raton, FL: CRC Press), 165. doi:10.1201/b21930-8

Bylaska, E. J., Jacquelin, M., de Jong, W. A., Hammond, J. R., and Klemm, M. (2017b). "Performance evaluation of NWChem ab-initio molecular dynamics (AIMD) simulations on the Intel® xeon Phi™ processor," in *High performance computing*. Editors J. M. Kunkel, R. Yokota, M. Taufer, and J. Shalf (Cham, Switzerland: Springer International Publishing), 404–418. doi:10.1007/978-3-319-67630-2_30

Bylaska, E. J. (2017). "Plane-wave DFT methods for chemistry," in *Annuual Reports in Computational Chemistry*. Editor D. A. Dixon (Amsterdam, Netherlands: Elsevier), 13, 185–228. doi:10.1016/bs.arcc.2017.06.006

Bylaska, E. J., Waters, K., Hermes, E. D., Zádor, J., and Rosso, K. M. (2020). A filon-like integration strategy for calculating exact exchange in periodic boundary conditions: a plane-wave dft implementation. *Mater. Theory.* 4, 1–31. doi:10.1186/s41313-020-00019-9

Bylaska, E., Tsemekhman, K., Govind, N., and Valiev, M. (2011). "Large-scale plane-wave-based density functional theory: formalism, parallelization, and applications," in *Computational methods for large systems: electronic structure* *approaches for biotechnology and nanotechnology*. Editor J. R. Reimers (Hoboken, NJ: Wiley Online Library), 77–116. doi:10.1002/9780470930779.ch3

Challacombe, M. (2000). Linear scaling computation of the fock matrix. v. hierarchical cubature for numerical integration of the exchange-correlation matrix. *J. Chem. Phys.* 113, 10037–10043. doi:10.1063/1.1316012

Chen, M., Guo, G., and He, L. (2011). Electronic structure interpolation via atomic orbitals. *J. Phys. Condens. Matter* 23, 325501. doi:10.1088/0953-8984/23/32/325501

Clark, S. J., Segall, M. D., Pickard, C. J., Hasnip, P. J., Probert, M. I. J., Refson, K., et al. (2005). First principles methods using castep. *Z. für Kristallogr. - Cryst. Mater.* 220, 567–570. doi:10.1524/zkri.220.5.567.65075

Coulson, C. A., and Fischer, I. (1949). Notes on the molecular orbital treatment of the hydrogen molecule. *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science.* 40, 386. doi:10.1080/14786444908521726

Crawford, T. D., Clementi, C., Harrison, R., Head-Gordon, T., Jha, S., Krylov, A., et al. (2017). The molecular sciences software institute. *Abstr. Pap. Am. Chem. Soc.* Vol. 254. WASHINGTON, DC: AMER CHEMICAL SOC 1155 16TH ST, NW.

Deslippe, J., Samsonidze, G., Strubbe, D. A., Jain, M., Cohen, M. L., and Louie, S. G. (2012). Berkeleygw: a massively parallel computer package for the calculation of the quasiparticle and optical properties of materials and nanostructures. *Comput. Phys. Commun.* 183, 1269–1289. doi:10.1016/j.cpc.2011.12.006

Dubash, M. (2005). Moore's law is dead, says gordon moore. *Techworld. com* 13.

Dunning, T. H. (1989). Gaussian basis sets for use in correlated molecular calculations. i. the atoms boron through neon and hydrogen. *J. Chem. Phys.* 90, 1007–1023. doi:10.1063/1.456153

Dunning, T. H., and Hay, P. J. (1977). Gaussian basis sets for molecular calculations. *Methods of electronic structure theory 3*, 1–27. doi:10.1007/978-1-4757-0887-5_1

Dupuis, M., Watts, J. D., Villar, H. O., and Hurst, G. J. B. (1989). The general atomic and molecular electronic structure system hondo: version 7.0. *Comput. Phys. Commun.* 52, 415–425. doi:10.1016/0010-4655(89)90116-1

Evangelista, F. A., Chan, G. K., and Scuseria, G. E. (2019). Exact parameterization of fermionic wave functions via unitary coupled cluster theory. *J. Chem. Phys.* 151, 244112. doi:10.1063/1.5133059

Frisch, M. J., Trucks, G. W., Schlegel, H. B., Scuseria, G. E., Robb, M. A., Cheeseman, J. R., et al. (2016). *Gaussian 16 revision C.01.* Wallingford CT: Gaussian Inc.

Giannozzi, P., Baroni, S., Bonini, N., Calandra, M., Car, R., Cavazzoni, C., et al. (2009). Quantum espresso: a modular and open-source software project for quantum simulations of materials. *J. Phys. Condens. Matter* 21, 395502. doi:10.1088/0953-8984/21/39/395502

Goldbeck, G. (2017). The scientific software industry: a general overview. Goldbeck Consulting, Technical Report.

Gonze, X., Jollet, F., Araujo, F. A., Adams, D., Amadon, B., Applencourt, T., et al. (2016). Recent developments in the abinit software package. *Comput. Phys. Commun.* 205, 106–131. doi:10.1016/j.cpc.2016.04.003

Grimsley, H. R., Claudino, D., Economou, S. E., Barnes, E., and Mayhall, N. J. (2020). Is the trotterized uccsd ansatz chemically well-defined? *J. Chem. Theor. Comput.* 16, 1–6. doi:10.1021/acs.jctc.9b01083

Grimsley, H. R., Economou, S. E., Barnes, E., and Mayhall, N. J. (2019). An adaptive variational algorithm for exact molecular simulations on a quantum computer. *Nat. Commun.* 10, 3007. doi:10.1038/s41467-019-10988-2

Gygi, F. (2008). Architecture of qbox: a scalable first-principles molecular dynamics code. *IBM J. Res. Dev.* 52, 137–144. doi:10.1147/rd.521.0137

Hall, G. (1951). The molecular orbital theory of chemical valency viii. a method of calculating ionization potentials. *Proc. R. Soc. Lond. A Math. Phys. Sci.* 205, 541–552. doi:10.1098/rspa.1951.0048

Hamann, D. R. (1989). Generalized norm-conserving pseudopotentials. *Phys. Rev. B Condens. Matter* 40, 2980. doi:10.1103/physrevb.40.2980

Harrison, R. J., Beylkin, G., Bischoff, F. A., Calvin, J. A., Fann, G. I., Fosso-Tande, J., et al. (2016). Madness: a multiresolution, adaptive numerical environment for scientific simulation. *SIAM J. Sci. Comput.* 38, S123–S142. doi:10.1137/15M1026171

Hocquet, A., and Wieber, F. (2017). "only the initiates will have the secrets revealed": computational chemists and the openness of scientific software. *IEEE Ann. Hist. Comput.* 39, 40–58. doi:10.1109/MAHC.2018.1221048

Hutter, J., Iannuzzi, M., Schiffmann, F., and VandeVondele, J. (2014). cp2k: atomistic simulations of condensed matter systems. *WIREs Comput. Mol. Sci.* 4, 15–25. doi:10.1002/wcms.1159

Hylleraas, E. A. (1929). Neue berechnung der energie des heliums im grundzustande, sowie des tiefsten terms von ortho-helium. *Z. Phys.* 54, 347–366. doi:10.1007/bf01375457

Hylleraas, E. A. (1964). The schrödinger two-electron atomic problem. *Adv. Quant. Chem.* 1, 1–33. doi:10.1016/s0065-3276(08)60373-1

Hylleraas, E. A. (1930). Über den Grundterm der Zweielektronenprobleme von H?, He, Li+, Be++ usw. *Z. Phys.* 65, 209–225. doi:10.1007/bf01397032

Hylleraas, E. A. (1928). Über den Grundzustand des Heliumatoms. *Z. Phys.* 48, 469–494. doi:10.1007/bf01340013

Johnson, S. (2007). The NLopt nonlinear-optimization package. Available at: http://github.com/stevengj/nlopt.doi:10.1364/np.2007.ntub1

Jordan, P., and Wigner, E. (1928). Über das Paulische äquivalenzverbot. *Z. Phys.* 47, 631–651. doi:10.1007/BF01331938

Kandala, A., Mezzacapo, A., Temme, K., Takita, M., Brink, M., Chow, J. M., et al. (2017). Hardware-efficient variational quantum eigensolver for small molecules and quantum magnets. *Nature* 549, 242–246. doi:10.1038/nature23879

Kendall, R. A., Aprà, E., Bernholdt, D. E., Bylaska, E. J., Dupuis, M., Fann, G. I., et al. (2000). High performance computational chemistry: an overview of nwchem a distributed parallel application. *Comput. Phys. Commun.* 128, 260–283. doi:10.1016/S0010-4655(00)00065-5

King, H. F., Stanton, R. E., Kim, H., Wyatt, R. E., and Parr, R. G. (1967). Corresponding orbitals and the nonorthogonality problem in molecular quantum mechanics. *J. Chem. Phys.* 47, 1936–1941. doi:10.1063/1.1712221

Kleinman, L., and Bylander, D. M. (1982). Efficacious form for model pseudopotentials. *Phys. Rev. Lett.* 48, 1425. doi:10.1103/physrevlett.48.1425

Koga, T. (1992). Hylleraas wave functions revisited. *J. Chem. Phys.* 96, 1276–1279. doi:10.1063/1.462164

Kohn, W., and Sham, L. J. (1965). Self-consistent equations including exchange and correlation effects. *Phys. Rev.* 140, A1133. doi:10.1103/PhysRev.140.A1133

Kowalski, K., and Bauman, N. P. (2020). Publisher's Note: "Sub-system quantum dynamics using coupled cluster downfolding techniques" [J. Chem. Phys. 152, 244127 (2020)]. *J. Chem. Phys.* 153, 049902. doi:10.1063/5.0008436 doi:10.1063/5.0021137

Kresse, G., and Furthmüller, J. (1996). Efficiency of ab-initio total energy calculations for metals and semiconductors using a plane-wave basis set. *Comput. Mater. Sci.* 6, 15–50. doi:10.1016/0927-0256(96)00008-0

Low, G. H., Bauman, N. P., Granade, C. E., Peng, B., Wiebe, N., Bylaska, E. J., et al. (2019). Q# and nwchem: tools for scalable quantum chemistry on quantum computers. Preprint: arXiv:1904.01131.

McCaskey, A., Dumitrescu, E., Chen, M., Lyakh, D., and Humble, T. (2018a). Validating quantum-classical programming models with tensor network simulations. *PloS One* 13, e0206704–19. doi:10.1371/journal.pone.0206704

McCaskey, A. J., Dumitrescu, E. F., Liakh, D., Chen, M., Feng, W., and Humble, T. S. (2018b). A language and hardware independent approach to quantum-classical computing. *Software* 7, 245–254. doi:10.1016/j.softx.2018.07.007

McCaskey, A. J., Lyakh, D. I., Dumitrescu, E. F., Powers, S. S., and Humble, T. S. (2020). XACC: a system-level software infrastructure for heterogeneous quantum-classical computing. *Quantum Sci. Technol.* 5, 024002. doi:10.1088/2058-9565/ab6bf6

McCaskey, A. J., Parks, Z. P., Jakowski, J., Moore, S. V., Morris, T. D., Humble, T. S., et al. (2019). Quantum chemistry as a benchmark for near-term quantum computers. *Npj Quantum Inf.* 5, 1–8. doi:10.1038/s41534-019-0209-0

Moore, G. E. (2006). Cramming more components onto integrated circuits, Reprinted from Electronics, volume 38, number 8, April 19, 1965, pp.114 ff. *IEEE Solid-State Circuits Soc. Newsl.* 11, 33–35. doi:10.1109/N-SSC.2006.4785860

Mulliken, R. S., Rieke, C. A., and Brown, W. G. (1941). Hyperconjugation. *J. Am. Chem. Soc.* 63, 41–56. doi:10.1021/ja01846a008

Mulliken, R. S., and Rieke, C. A. (1941). Improved computations on conjugation and hyperconjugation. *J. Am. Chem. Soc.* 63, 1770–1771. doi:10.1021/ja01851a514

Neese, F. (2018). Software update: the orca program system, version 4.0. *WIREs Comput Mol Sci.* 8, e1327. doi:10.1002/wcms.1327

Nesbet, R. K. (1963). Computer programs for electronic wave-function calculations. *Rev. Mod. Phys.* 35, 552. doi:10.1103/RevModPhys.35.552

Nesbet, R. K. (1960). Ground state electronic wave function of methane. *J. Chem. Phys.* 32, 1114–1122. doi:10.1063/1.1730858

O'Malley, P. J. J., Babbush, R., Kivlichan, I. D., Romero, J., McClean, J. R., Barends, R., et al. (2016). Scalable quantum simulation of molecular energies. *Phys. Rev. X.* 6, 031007. doi:10.1103/PhysRevX.6.031007

Parr, R. G., and Mulliken, R. S. (1950). LCAOSelf-consistent field calculation of the π-electron energy levels of cis- andtrans−1,3-butadiene. *J. Chem. Phys.* 18, 1338–1346. doi:10.1063/1.1747474

Peruzzo, A., McClean, J., Shadbolt, P., Yung, M. H., Zhou, X. Q., Love, P. J., et al. (2014). A variational eigenvalue solver on a photonic quantum processor. *Nat. Commun.* 5, 4213. doi:10.1038/ncomms5213

Pople, J. A., Santry, D. P., and Segal, G. A. (1965). Approximate self-consistent molecular orbital theory. I. Invariant procedures. *J. Chem. Phys.* 43, S129–S135. doi:10.1063/1.1701475

Powell, M. J. D. (1994). *A direct search optimization method that models the objective and constraint functions by linear interpolation*. Dordrecht, Netherlands: Springer Netherlands, 51–67. doi:10.1007/978-94-015-8330-5_4

Prendergast, D., and Louie, S. G. (2009). Bloch-state-based interpolation: an efficient generalization of the shirley approach to interpolating electronic structure. *Phys. Rev. B* 80, 235126. doi:10.1103/PhysRevB.80.235126

Preskill, J. (2018). Quantum computing in the nisq era and beyond. *Quantum.* 2, 79. doi:10.22331/q-2018-08-06-79

Pulay, P. (1969). Ab initiocalculation of force constants and equilibrium geometries in polyatomic molecules. *Mol. Phys.* 17, 197–204. doi:10.1080/00268976900100941

Reeves, C. M. (1966). An algorithm for generating projective reduction formulas for matrix elements of many-electron wavefunctions. *Commun. ACM* 9, 276–279. doi:10.1145/365278.365540

Reiher, M., Wiebe, N., Svore, K. M., Wecker, D., and Troyer, M. (2017). Elucidating reaction mechanisms on quantum computers. *Proc. Natl. Acad. Sci. U.S.A.* 114, 7555–7560. doi:10.1073/pnas.1619152114

Richard, R. M., Bertoni, C., Boschen, J. S., Keipert, K., Pritchard, B., Valeev, E. F., et al. (2019). Developing a computational chemistry framework for the exascale era. *Comput. Sci. Eng.* 21, 48–58. doi:10.1109/MCSE.2018.2884921

Romero, J., Babbush, R., McClean, J. R., Hempel, C., Love, P. J., and Aspuru-Guzik, A. (2017). Strategies for quantum computing molecular energies using the unitary coupled cluster ansatz. *Quantum Sci. Technol.* 4, 014008. doi:10.1088/2058-9565/aad3e4

Roothaan, C. C. J. (1951). New developments in molecular orbital theory. *Rev. Mod. Phys.* 23, 69. doi:10.1103/RevModPhys.23.69

Rotman, D. (2020). We're not prepared for the end of moore's law. *MIT Technology Review.*

Schmidt, M. W., Baldridge, K. K., Boatz, J. A., Elbert, S. T., Gordon, M. S., Jensen, J. H., et al. (1993). General atomic and molecular electronic structure system. *J. Comput. Chem.* 14, 1347–1363. doi:10.1002/jcc.540141112

Setia, K., and Whitfield, J. D. (2018). Bravyi-kitaev superfast simulation of electronic structure on a quantum computer. *J. Chem. Phys.* 148, 164104. doi:10.1063/1.5019371

Shao, Y., Gan, Z., Epifanovsky, E., Gilbert, A. T., Wormit, M., Kussmann, J., et al. (2015). Advances in molecular quantum chemistry contained in the q-chem 4 program package. *Mol. Phys.* 113, 184–215. doi:10.1080/00268976.2014.952696

Shirley, E. L. (1996). Optimal basis sets for detailed brillouin-zone integrations. *Phys. Rev. B Condens. Matter.* 54, 16464. doi:10.1103/PhysRevB.54.16464

Stanton, J. F., Gauss, J. r., Watts, J. D., Lauderdale, W. J., and Bartlett, R. J. (1992). The aces ii program system. *Int. J. Quant. Chem.* 44, 879–894. doi:10.1002/qua.560440876

Steudtner, M., and Wehner, S. (2018). Lowering qubit requirements for quantum simulations of fermionic systems. Preprint: arXiv:1712.07067v2.

Svore, K., Roetteler, M., Geller, A., Troyer, M., Azariah, J., Granade, C., et al. (2018). "Q# enabling scalable quantum computing and development with a high-level dsl," in Proceedings of the Real World Domain Specific Languages Workshop 2018, Vienna, Austria, February 2018. 1–10. doi:10.1145/3183895.3183901

te Velde, G., Bickelhaupt, F. M., Baerends, E. J., Fonseca Guerra, C., van Gisbergen, S. J. A., Snijders, J. G., et al. (2001). Chemistry with adf. *J. Comput. Chem.* 22, 931–967. doi:10.1002/jcc.1056

Valiev, M., Bylaska, E. J., Govind, N., Kowalski, K., Straatsma, T. P., Van Dam, H. J. J., et al. (2010). NWChem: a comprehensive and scalable open-source solution for large scale molecular simulations. *Comput. Phys. Commun.* 181, 1477–1489. doi:10.1016/j.cpc.2010.04.018

van Dam, H., Apra, E., Bair, R., Boschen, J., Bylaska, E., De Jong, W., et al. (2020). NWChemEx–computational chemistry for the exascale era. *Bull. Am. Phys. Soc.* 65.

Verma, P., Huntington, L., Coons, M., Kawashima, Y., Yamazaki, T., and Zaribafiyan, A. (2020). Scaling up electronic structure calculations on quantum computers: the frozen natural orbital based method of increments. Preprint: arXiv:2002.07901.

Wasielewski, M. R., Forbes, M. D. E., Frank, N. L., Kowalski, K., Scholes, G. D., Yuen-Zhou, J., et al. (2020). Exploiting chemistry and molecular systems for quantum information science. *Nat. Rev. Chem.* 4, 490–15. doi:10.1038/s41570-020-0200-5

Werner, H.-J., Knowles, P. J., Knizia, G., Manby, F. R., and Schütz, M. (2012). Molpro: a general-purpose quantum chemistry program package. *WIREs Comput. Mol. Sci.* 2, 242–253. doi:10.1002/wcms.82

Wikipedia, The Free Encyclopedia (2020). *List of quantum chemistry and solid-state physics software*. San Francisco: Wikimedia Foundation Inc.

Wilkins-Diehr, N., and Crawford, T. D. (2018). NSF's inaugural software institutes: the science gateways community institute and the molecular sciences software institute. *Comput. Sci. Eng.* 20, 26–38. doi:10.1109/MCSE.2018.05329813

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Physics

ORIGINAL RESEARCH
published: 19 March 2021
doi: 10.3389/fphy.2021.546538

A Comparison of Three Ways to Measure Time-Dependent Densities With Quantum Simulators

Jun Yang, James Brown and James Daniel Whitfield*

Department of Physics and Astronomy, Dartmouth College, Hanover, NH, United States

Quantum algorithms are touted as a way around some classically intractable problems such as the simulation of quantum mechanics. At the end of all quantum algorithms is a quantum measurement whereby classical data is extracted and utilized. In fact, many of the modern hybrid-classical approaches are essentially quantum measurements of states with short quantum circuit descriptions. Here, we compare and examine three methods of extracting the time-dependent one-particle probability density from a quantum simulation: direct Z-measurement, Bayesian phase estimation, and harmonic inversion. We have tested these methods in the context of the potential inversion problem of time-dependent density functional theory. Our test results suggest that direct measurement is the preferable method. We also highlight areas where the other two methods may be useful and report on tests using Rigetti's quantum virtual device. This study provides a starting point for imminent applications of quantum computing.

Keywords: quantum measurement, TDDFT, Bayesian inference, harmonic inversion, potential inversion

Edited by:
Travis S. Humble,
Oak Ridge National Laboratory (DOE),
United States

Reviewed by:
Marcos César de Oliveira,
State University of Campinas, Brazil
Pavel Dub,
Los Alamos National Laboratory
(DOE), United States
Titus Morris,
Oak Ridge National Laboratory (DOE),
United States

***Correspondence:**
James Daniel Whitfield
james.d.whitfield@dartmouth.edu

Specialty section:
This article was submitted to
Quantum Engineering and Technology,
a section of the journal
Frontiers in Physics

Received: *28 March 2020*
Accepted: *22 February 2021*
Published: *19 March 2021*

Citation:
Yang J, Brown J and Whitfield JD
(2021) A Comparison of Three Ways
to Measure Time-Dependent Densities
With Quantum Simulators.
Front. Phys. 9:546538.
doi: 10.3389/fphy.2021.546538

1. INTRODUCTION

The real time simulation of quantum systems on a classical computer is a difficult problem even for a supercomputer due to the fact that the Hilbert space grows exponentially with the system size [1]. A universal quantum computer is believed to be the solution of the difficulty, where it is known that a wide class of physical systems can be simulated efficiently on a quantum computer [1–6]. But running a practically meaningful quantum algorithm may require a large amount of qubits, e.g., factoring 2,048 bit RSA integers may take up to 20 millions qubits [7], which is far beyond the capacity of the current best 53-qubit quantum computer [8]. So the current quantum technology works best when paired with classical algorithms. We have been studying the application of such a hybrid algorithm in quantum chemistry. The primary example is the time-dependent density functional theory (TDDFT) [9]. To utilize quantum technology in classical algorithms, quantum measurement are necessary, here we measured the density operator on Rigetti's quantum device and then utilized the density to perform the potential inversion within the framework of TDDFT.

Density functional theory (DFT) is a powerful tool in modeling condensed matter systems [10]. In the framework of DFT, a non-interacting system with a self-consistently determined potential is constructed to replace the interacting system. The additional potential term in the non-interacting system is known as the Kohn-Sham potential. Such a system with non-interacting particles is the Kohn-Sham(K-S) system. In the K-S system, the calculation of an exchange-correlation term is required. However, the exact form of the exchange-correlation potential is not yet known. This term is usually obtained with some approximation methods [11–14], machine learning methods [15, 16]. In fact, the utilization of a quantum computer can help generate an accurate exchange-correlation potential. This idea is mentioned in the article [17], where a hybrid method of generating exchange-correlation potential for classical DFT calculation is proposed.

The time dependent counterpart of DFT, time-dependent density functional theory (TDDFT) is widely used in finding the dynamics of the system when a time-dependent potential is present. Similar to DFT, TDDFT uses the time dependent K-S system where a time-dependent K-S potential is required. We call the task of constructing such a K-S potential when given the time-evolution of the on-site probability density, the K-S potential inversion problem. In article [18], a scheme of solving the K-S potential inversion problem utilizing a quantum computer was proposed. We have recently returned to this proposal with improved numerical methods for inverting the potential [19]. To obtain the K-S potential, we need to get the density of the time evolved many-particle system using a quantum computer.

In this paper, we will present three different methods of measuring the density operator on a quantum computer and compare the performance of the methods.

An outline for the remainder of the article is as follows: first, we discuss the phase estimation approach to measurement. Then we describe the circuit implementation for measuring the on-site fermionic density. Qubit descriptions for the fermionic operator are explained in the next part followed by the illustration of a two-electron test. Finally, three schemes for extracting the density are tested numerically and compared.

2. METHODS

2.1. Phase Estimation

Quantum phase estimation [20, 21] plays an important role in the quantum algorithm zoo [22], it is a key sub module of many quantum algorithms [23–25]. It is also an important procedure to measure the on-site density operator in our work.

We'll next describe the general picture of doing the measurement of an arbitrary operator and how quantum phase estimation plays a role in our work. To implement the measurement of an arbitrary observable, we will consider the circuit as shown in **Figure 1**. The circuit has two parts, the part before the dashed line is for evolving the initial state at time t under a fixed fermionic Hamiltonian of chemical interests.

The system of the most chemical interests is the interacting electron system. The Hamiltonian of a many-body interacting system is given by

$$H=\sum_{i}^{N}\left[-\frac{\nabla_i^2}{2}+V_{ext}(\mathbf{r}_i)+\frac{1}{2}\sum_{j}^{N}\frac{1}{|\mathbf{r}_i-\mathbf{r}_j|}\right] \tag{1}$$

where $V_{ext}(\mathbf{r}_i)$ is the external potential energy consists of the interaction between the electrons and the external field.

The second quantized form of the above many-body Hamiltonian is given by

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

where the fermionic operators $\{a_p, a_p^\dagger\}$ satisfy $a_q^\dagger a_p + a_p a_q^\dagger = \delta_{pq}, a_p a_q = -a_q a_p$ and $a_p^\dagger a_q^\dagger = -a_q^\dagger a_p^\dagger$. Given the basis set $\{\chi_p(\mathbf{r})\}$, the coefficients h_{pq}, h_{pqrs} are given by

$$h_{pq}=\int d\mathbf{r}\chi_p^\star(\mathbf{r})\left(-\frac{1}{2}\nabla^2+V_{ext}(\mathbf{r})\right)\chi_q(\mathbf{r}) \tag{3}$$

$$h_{pqrs}=\int d\mathbf{r}_1 d\mathbf{r}_2\frac{\chi_p^\star(\mathbf{r}_1)\chi_q^\star(\mathbf{r}_2)\chi_r(\mathbf{r}_2)\chi_s(\mathbf{r}_1)}{|\mathbf{r}_2-\mathbf{r}_1|} \tag{4}$$

The latter half is a phase estimation circuit where $U_O(\tau)=e^{-iO\tau}$ where O is the observable to be measured.

For a general state $|\psi\rangle$, we can expand it in the eigenspace of the operator O. To be precise, for a general state $|\psi\rangle=\sum_k c_k|k\rangle$, where O_k and $|k\rangle$ are the eigenvalue and eigenvector of the operator O. Thus, the probability of measuring zero on the top register is given by

$$\begin{aligned}P(0|\tau,t)&=\frac{1}{2}+\frac{1}{4}[\langle\psi(t,\tau)|\,\psi(t)\rangle+\langle\psi(t)|\,\psi(t,\tau)\rangle]\\&=\sum_k|c_k(t)|^2\cos^2\left(\frac{O_k\tau}{2}\right)\\&=\frac{1}{2}+\frac{1}{4}\sum_k|c_k(t)|^2\left(e^{iO_k\tau}+e^{-iO_k\tau}\right)\end{aligned} \tag{5}$$

where $c_k(t)=\langle k|U(t)|\psi\rangle$.

In this article, we only consider $O=n_j=a_j^\dagger a_j$ in order to measure the local on-site density at site j. This is because the inverse potential is determined by the on-site density and its first and second order derivatives [18, 19].

The eigenvalues of $n_j=a_j^\dagger a_j$ are 0 and 1, so the wave function after the unitary evolution $U(t)$ is given by $|\psi(t)\rangle = c_0(t)\,|\psi_{n_j=0}\rangle + c_1(t)\,|\psi_{n_j=1}\rangle$. Thus, the expectation value of the density is given by,

$$\langle n_j(t)\rangle=\langle\psi(t)|a_j^\dagger a_j|\psi(t)\rangle=|c_1(t)|^2 \tag{6}$$

2.2. Qubit Encoding

To implement the evolution and phase estimation algorithm on a quantum computer, we need to encode the Hamiltonian into qubits. A standard way is to use Jordan-Wigner (JW) transformation, which encodes a fermionic system of M orbitals into M qubits.

$$a_p=\frac{1}{2}\left(X_p+iY_p\right)Z_1Z_2\dots Z_{p-1} \tag{7}$$

$$a_p^\dagger=\frac{1}{2}\left(X_p-iY_p\right)Z_1Z_2\dots Z_{p-1} \tag{8}$$

With the above transformation, the fermionic Hamiltonian can be encoded into qubit representation. Thus the Hamiltonian can be written as $H=\sum_i h_i$, where all the h_i's are tensor product of Pauli operators.

There is not an easy way to construct arbitrary unitary operators on a quantum computer [26]. A pragmatic way to simulate the propagator $U(t)=e^{-iHt}$ is applying the Trotter decomposition.

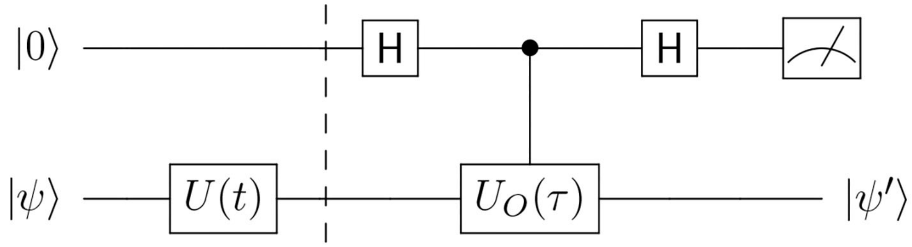


FIGURE 1 | The circuit for measuring the density matrix. The half before the dashed line is used for evolving the state to time *t*, the half after is used for doing the measurement of an observable at time *t* via phase estimation.

$$U(t) = e^{-iHt} \approx \left(e^{-ih_1 t/N} e^{-ih_2 t/N} \dots e^{-ih_n t/N}\right)^N \tag{9}$$

Each individual term in the decomposition above can be simulated efficiently on a quantum computer [27].

3. RESULTS AND DISCUSSION

We tested our methods on two different Hamiltonians, one is the 4-orbital HeH^+ model, the other is the 8-orbital HeH^+. For the 4-orbital model, the basis set used to examine the HeH^+ molecule is that given in [28] which results in four spin orbitals. The interatomic distance is 1.401 Bohr. The basis functions are orthogonalized and then transformed such that the one-body Coulomb matrix is diagonal. This transformation was chosen so that a corresponding scalar time-dependent Kohn-Sham potential could be calculated using for this system using the method of [19].

For the larger case, we used HeH^+ at the same geometry but in the 6-31G basis set [29]. This resulted in twice the number of basis functions as the minimal example. The integrals in the 6–31 basis were computed using Pis4 [30].

The basis functions are orthogonalized and then transformed such that the one-body Coulomb matrix is diagonal. This transformation was chosen so that a corresponding time-dependent Kohn-Sham potential could be calculated using for this system using the method of [19]. In both models, the initial state at $t = 0$ places two electrons in the first two modes of opposite spin. This state is obtained by employing two X-gates to prepare $|\psi(0)\rangle = |1100\rangle$ in the 4-orbital model and $|\psi(0)\rangle = |11000000\rangle$ in the 8-orbital model.

Using Rigetti's quantum virtual machine [31], we then evolve the system under its Hamiltonian for times less than three atomic units. The propagation is implemented via the first-order Trotterization with Trotter step equal to three. To reduce the Trotter error in evolution, either a shorter Trotter step or a higher order Trotter approximation must be used [32]. This means more quantum gates are needed, making it harder to be implemented on a near term device. Additional sources of error are associated with finite sampling from the binomial distribution and the error associated with the inference steps. To make the virtual machine slightly closer to a real quantum computer, in all methods below, measurement noise was added into the system, giving 1% probability of flipping the qubit. It should be noted that the quantum noise found on the actual device was much higher, so we will not present the results from the actual quantum device.

In **Figures 2A,C,D,F**, we used 3,000 quantum measurement samples per time-point. For harmonic inversion, a total of 120,000 quantum measurements occur for extracting the density at each time point. This is because there were 40 equally spaced τ-points and 3,000 quantum measurements were used per fixed τ. A comparison between the measuring results, the exact solution of the original Hamiltonian and the exact solution of the Trotterized Hamiltonian are compared in **Figure 2**. In each subfigure, the dark green dots are the result from measurement result, the red dashed line is the solution of the Trotterized Hamiltonian, the black solid line is the solution of the original Hamiltonian.

3.1. Method 1: Z-Basis Measurement

In the first method, we rely on the fact that the Jordan-Wigner transformation of the on-site density operator has a simple form $a_p^\dagger a_p = (1 - Z_p)/2$. Thus, we can directly measure the local density operator by measuring Z_p without passing in the phase estimation circuit after the dotted line in **Figure 1**.

For an arbitrary wave function $|\psi(t)\rangle = c_0(t)\,|\psi_{n_p=0}\rangle + c_1(t)\,|\psi_{n_p=1}\rangle$, where ψ_{n_p} denotes the state projected into the subspace where the p-th qubit is in state n_p. Given the fact that $\langle Z_p(t)\rangle = |c_0(t)|^2 - |c_1(t)|^2$ and $|c_0(t)|^2 + |c_1(t)|^2 = 1$, both amplitudes $|c_0(t)|^2$ and $|c_1(t)|^2$ can be obtained from the measurement.

By repeating the measurement at each time step in the time range $0 \le t \le 3$, we obtain the expectation value of the density. The results based on 15 equally spaced time-points with 3,000 measurements at each fixed time are shown in **Figures 2A,D**. The exact time evolution of the density is also shown in the figure for comparison along with error bars of 2σ reflective of the $N = 3,000$ sample variance of the binomial distribution.

The simplicity of this measurement approach reduces the classical runtime to the lowest of the three methods compared, and the convergence of the error bars is faster than the Bayesian measurement discussed later.

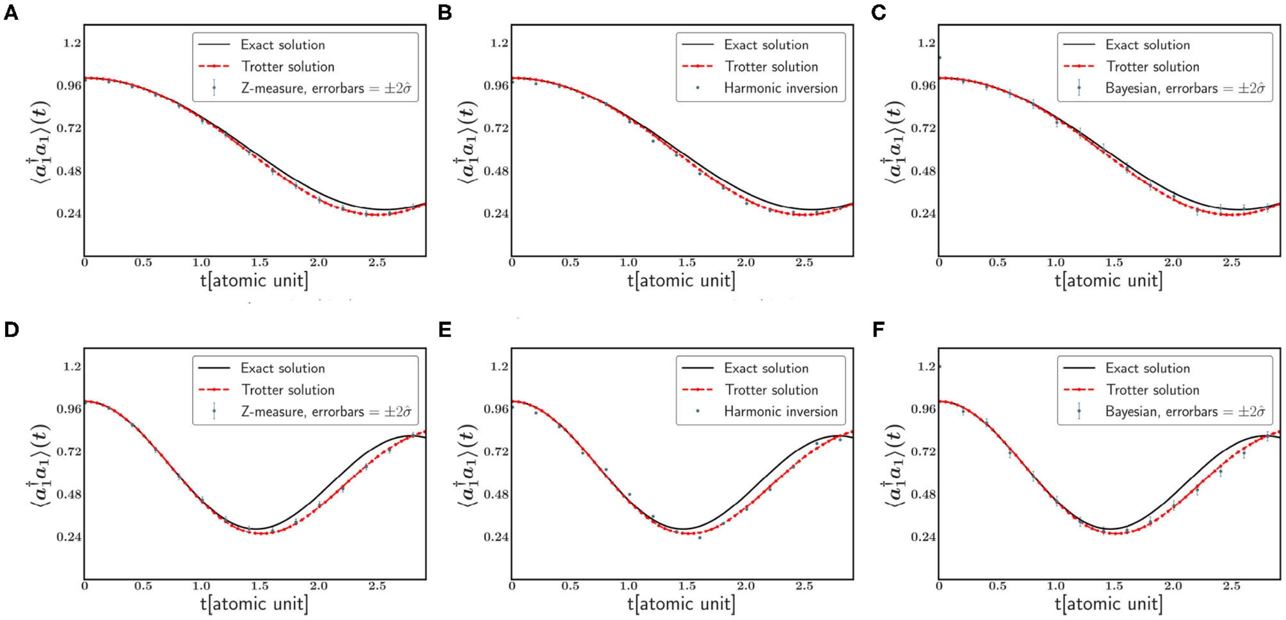


FIGURE 2 | The expectation value of $a_1^\dagger a_1$ is measured via **(A,D)** direct Z-basis measurement, **(B,E)** harmonic inversion, and **(C,F)** Bayesian inference. The upper panel are the results of 4-orbital model, the lower panel are the results of the 8-orbital model. All of the data points are plotted against the exact Trotter solution depicted in as a continuous line. In **(C,F)**, the first point shows a large deviation from the Trotter solution typical of the behavior of Bayesian inference whenever the exact density's value is close to one. Error bars shown in **(A,C,D,F)** are two standard deviations about the mean. The harmonic inversion does not have error bars because the error comes from two sources: from the sampling error at different time τ and from the reconstruction of the density using harmonic inversion.

3.2. Method 2: Harmonic Inversion

Harmonic inversion is a technique of extracting the amplitudes A_j, frequencies f_j, phases ϕ_j, and exponential decay constants α_j out of a signal,

$$f(\tau) = \sum_j A_j e^{-i(2\pi f_j \tau - \phi_j) - \alpha_j \tau} \tag{10}$$

which is evenly sampled [33, 34]. The signal reconstructed from harmonic inversion has the same form as the probability $P(0|\tau)$ except for the decaying term which is negligible when the decoherence is not considered. By comparing the form of the reconstructed signal with the probability, we can obtain the density from the reconstructed signal.

The results of density measurement through harmonic inversion are shown in **Figures 2B,E**. Each point in **Figures 2B,E** was computed through harmonic inversion using the HarmInv package [35]. Because the local density operator $a_p^\dagger a_p$ only has eigenvalues zero and one, the measurement outcome has a simple form

$$P(0|\tau, t) = A_0(t) + A_1(t)\left(e^{-i2\pi f\tau} + e^{i2\pi f\tau}\right) \tag{11}$$

where $A_0(t) = \frac{1}{2}(2 - |c_1(t)|^2)$, $A_1(t) = |c_1(t)|^2/4$, and $f = 1/2\pi$.

3.3. Method 3: Bayesian Inference

Bayesian inference can be used to estimate the density as well. As a powerful tool of making inferences, Bayesian inference has wide applications. We applied Bayesian inference to infer the unknown parameters in a quantum system which, in our case, is the on-site density. The density estimation was implemented via sequential Monte Carlo (SMC) [36]. This method requires the most communication between the classical and quantum processors since the SMC suggests each τ-point based on the previous outcomes. The Bayesian experimental design is based on the implementation found in the QInfer package [37]. Bayesian inference gives the probability distribution of a parameter over the parameter space. The final decision is made according to the posterior probability $P(\theta|d_1, d_2 \ldots d_N)$, where θ is the parameter we want to estimate, d_i's are the outcome of each measurement. In the present application, $\theta \equiv \langle n_j(t) \rangle$.

Recall the Bayesian rule, the posterior probability is updated by carrying out experiments sequentially,

$$P(\theta|d_1, d_2, \ldots d_N) \propto \prod_{i=1}^{N} P(d_i|\theta)P(\theta) \tag{12}$$

where $P(\theta)$ is the prior probability, $P(d_i|\theta)$ is the likelihood function.

The likelihood function contains the information about the parameters before conducting any experiments. Since we know nothing before the experiment, we can initialize the prior with a uniform distribution over the parameter space. For the phase estimation circuit of **Figure 1**, the likelihood function is given by

$$P(d|\,\langle n_j(t)\rangle\,;\tau) = \frac{1}{2} + \frac{(-1)^d}{4}\,\langle \psi(t)|\{U_O(\tau) + U_O^\dagger(\tau)\}|\psi(t)\rangle \tag{13}$$

where $U_O(\tau) = \exp(-i\tau a_j^\dagger a_j)$ and $d = 0$ or 1. Note, when $d = 0$ we recover Equation (5).

With this we can rewrite the likelihood function as

$$P(d|\,\langle n_j(t)\rangle\,;\tau) = \delta_{d,0} + \frac{(-1)^d}{2}(\cos\tau - 1)\,\langle n_j(t)\rangle \tag{14}$$

This can be compared with Equation (5) in the case that $d = 0$.

The results of Bayesian inference are shown in **Figures 2C,F**. Bayesian inference has good performance within a wide range of the time domain except at the boundary of the estimate domain e.g., when the density is one or zero. This is based on numerical evidence since the majority of the points at or near the boundary of the estimation domain needed to be discarded when cleaning the data as discussed below.

Unlike harmonic inversion, τ in the phase estimation circuit is not required to be evenly spaced. Another advantage of Bayesian inference is that we do not need to know the exact form of the function to be estimated *a priori*. Bayesian inference could also be applied to estimate more general parameters.

3.3.1. Comparison

To quantify the accuracy of these density extraction methods, we employ the L_1 norm to measure the deviation from the Trotter solution. For discrete data points, the deviation is given by the loss function on the density at the first site: $L = \sum_k^N |\tilde{n}_1(t_k) - n_1(t_k))|/N$, where $\tilde{n}_1(t_k)$ is the outcome of the measurement at time t_k, $n_1(t_k)$ is the solution of the Trotterized Hamiltonian.

Figure 3 shows how the loss function scales with the number of trials for each of the three approaches. The convergence rate for determining the bias of a coin would be 0.5 but here additional measurement error has been introduced into the model which prevents $L = 0$ situation even with an infinite number of samples. Further, in our implementation, the Bayesian and harmonic inversion techniques sometimes reported anomalously poor estimates of the density at a given time. A single fluctuation of this type along the time trace of the density entirely dominates the loss function. For the sake of comparison, we did not include the data points that are 5σ away from the exact solution in all three methods. This led to more stable results when the number of trials is small. Another benefit of filtering the data is that for the Bayesian inference, estimates close to the boundary of the domain are subject to large fluctuations giving poor estimates. So we can exclude the wrong data points by setting a 5σ window. Although the discarding procedure is *ad hoc* and requires knowing the exact answer, we have tested our data at various levels of cutoff finding that at any fixed cutoff harmonic inversion had the most points discarded and consistently displayed marginally faster convergence rates.

Figure 3A shows the scaling of loss function of the 4-orbital model. The slope of the fitting lines are -0.4961, -0.3747, and -0.4103, respectively. Points 5σ away from the exact density under the Trotter approximation are not used for calculating the loss function. This resulted in 73.87, 87.26, and 90.07% of points used in the plotted data, respectively.

Figure 3B shows the scaling of loss function of the 8-orbital model. The slope of the fitting lines are -0.4346, -0.4595, and -0.4184, respectively. Points 5σ away from the exact density under the Trotter approximation are not used for calculating the loss function. This resulted in 80.53, 84.40, and 93.20% of points used in the plotted data, respectively.

Harmonic inversion measures 40 times more than the other two methods, so the actual data and fitting line should be shifted to the right by 40 times the number of measurements showing in the figure.

Regardless of the possible improvement in convergence, it should be reminded that the harmonic inversion technique uses many quantum computer queries to estimate $P(0|\tau, t)$ at variable τ before inferring the density at a fixed time t. In comparing the three methods, all require time evolution of the system wave function to time t. In the harmonic inversion and Bayesian estimation techniques, additional gates are needed for the τ propagation under the observable for density. The difference between queries in harmonic inversion and Bayesian inference is the selection of the τ parameter in $U_O(\tau)$.

While the convergence rates are all approximately the same, it is clear that the Z-basis measurement has the best performance in terms of the number of queries of the quantum computer. In the case considered here, the direct Z measurements are convenient for the Jordan-Wigner encoding. In other circumstances with different fermion-to-spin transforms, the direct measurement technique may not be as fruitful. For existing and near-term quantum devices, the constraints of low circuit depth suggests direct measurement of the Z operators as the best path forward when using a Jordan-Winger transformed qubit Hamiltonian.

The runtime of these three methods also varies. Since direct Z-measurements are the simplest from an inference point of view, the classical computation time is also the least. Bayesian inference requires many steps for the sequential Monte Carlo to converge [36]. Consequently, this method used the longest amount of classical computational time. Although harmonic inversion uses 40 times more measurement per time-point, it is interesting to note that it only took an intermediate amount of classical processing time.

4. CONCLUSIONS

We have tested three different methods of measuring the on-site density operator for a toy model inspired by TDDFT. We were able to conclude that direct Z measurements obtains the best estimates of the on-site density for a given number of quantum computer queries. This is based on the use of the Jordan-Wigner transform and simulated measurement noise. Of course, we could have considered other fermion-to-spin transforms which lead to different encodings of the $a_i^\dagger a_i$.

For improving our noise models, we can do no better than testing our circuits on current and future quantum devices. We

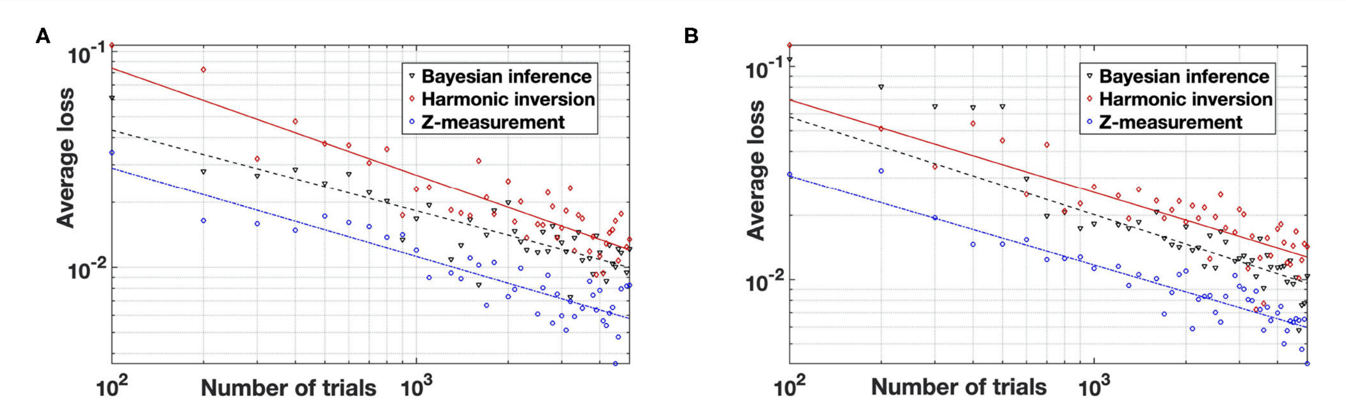


FIGURE 3 | Loss function vs. number of measurements of the quantum computer (trials) in both 4-orbital model and 8-orbital model. Red diamonds, black triangles, and blue circles are drawn from harmonic inversion, Bayesian inference and *Z*-basis measurement respectively. **(A)** 2 Electron 4 orbital case. **(B)** 2 Electron 8 orbital case.

tested our circuits on Rigetti's quantum device but found that the loss function depends heavily on which qubits are used as well as the permutation of qubit labels within the circuit. Time evolution under the full Hamiltonian did not return any signal even when using only one first-order Trotter step. We therefore resorted to using a truncated Hamiltonian which included the one-body Hamiltonian and only the Coulomb-like (h_{ijji}) terms of the two-body Hamiltonian. After encoding and exponentiation, this Hamiltonian results in 66 universal gates and compiled non-deterministically using the PyQuil package PyQuil to approximately 200 allowable gates on the Rigetti device. Due to decoherence, only a weak signal was present where amplitudes recovered were between three and twenty percent of the exact solution. The recovered amplitude depended mostly on qubit selection but also changed run-to-run. The frequency and sinusoidal shape of the signal was recovered more reliably. In our present study, the eigenenergies were not interesting but we suspect that problems that depend on the frequencies may be more successfully calculated on the current Rigetti device.

We plan to continue our inquiry into the TDDFT potential inversion problem using existing and forthcoming quantum technology. Tasks that avoid QMA-hard [38] state preparation problems will continue to be of interest to those looking for new applied areas of quantum computation.

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

AUTHOR CONTRIBUTIONS

JY and JW designed the model and the computational framework. JY and JB performed the calculations. JY, JB, and JW wrote the manuscript with input from all authors.

FUNDING

JY, JB, and JW were supported by the U.S. Department of Energy, Office of Science, Office of Advanced Scientific Computing Research, under the Quantum Computing Application Teams program (Award 1979657). JW was also supported by the NSF (Grant 1820747) and additional funding from the DOE (Award A053685). Calculations were performed using Dartmouth's Discovery Linux HPC Cluster.

REFERENCES

1. Georgescu IM, Ashhab S, Nori F. Quantum simulation. *Rev Mod Phys*. (2014) 86:153–85. doi: 10.1103/RevModPhys.86.153
2. Lloyd S. Universal quantum simulators. *Science*. (1996) 273:1073–8. doi: 10.1126/science.273.5278.1073
3. Gilyén A, Su Y, Low GH, Wiebe N. Quantum singular value transformation and beyond: exponential improvements for quantum matrix arithmetics. In: *Proceedings of the 51st Annual ACM SIGACT Symposium on Theory of Computing, STOC 2019*. New York, NY: Association for Computing Machinery (2019). p. 193–204.
4. Cao Y, Romero J, Olson JP, Degroote M, Johnson PD, Kieferová M, et al. Quantum chemistry in the age of quantum computing. *Chem Rev*. (2019) 119:10856–915. doi: 10.1021/acs.chemrev.8b00803
5. Berry DW, Ahokas G, Cleve R, Sanders BC. Efficient quantum algorithms for simulating sparse Hamiltonians. *Commun Math Phys*. (2007) 270:359–71. doi: 10.1007/s00220-006-0150-x
6. Nielsen MA, Bremner MJ, Dodd JL, Childs AM, Dawson CM. Universal simulation of Hamiltonian dynamics for quantum systems with finite-dimensional state spaces. *Phys Rev A*. (2002) 66:022317. doi: 10.1103/PhysRevA.66.022317
7. Gidney C, Eker M. How to factor 2048 bit RSA integers in 8 hours using 20 million noisy qubits. *arXiv [Preprint]*. (2019) arXiv:1905.09749.
8. Arute F, Arya K, Babbush R, Bacon D, Bardin JC, Barends R, et al. Quantum supremacy using a programmable superconducting processor. *Nature*. (2019) 574:505–10. doi: 10.1038/s41586-019-1666-5
9. Runge E, Gross EKU. Density-functional theory for time-dependent systems. *Phys Rev Lett*. (1984) 52:997–1000. doi: 10.1103/PhysRevLett.52.997

10. Hohenberg P, Kohn W. Inhomogeneous electron gas. *Phys Rev*. (1964) 136:B864–71. doi: 10.1103/PhysRev.136.B864
11. Slater JC. A Simplification of the Hartree-Fock Method. *Phys Rev*. (1951) 81:385–90. doi: 10.1103/PhysRev.81.385
12. Perdew JP, Burke K, Ernzerhof M. Generalized gradient approximation made simple. *Phys Rev Lett*. (1996) 77:3865–8. doi: 10.1103/PhysRevLett.77.3865
13. von Barth U, Hedin L. A local exchange-correlation potential for the spin polarized case. I. *J Phys C*. (1972) 5:1629–42. doi: 10.1088/0022-3719/5/13/012
14. Rajagopal AK, Callaway J. Inhomogeneous electron gas. *Phys Rev B*. (1973) 7:1912–9. doi: 10.1103/PhysRevB.7.1912
15. Snyder JC, Rupp M, Hansen K, Müller KR, Burke K. Finding density functionals with machine learning. *Phys Rev Lett*. (2012) 108:253002. doi: 10.1103/PhysRevLett.108.253002
16. Nagai R, Akashi R, Sugino O. Completing density functional theory by machine learning hidden messages from molecules. *NPJ Comput Mater*. (2020) 6:43. doi: 10.1038/s41524-020-0310-0
17. Hatcher R, Kittl JA, Bowen C. A Method to calculate correlation for density functional theory on a quantum processor. *arXiv [Preprint]:1903.05550*. (2019).
18. Whitfield JD, Yung MH, Tempel DG, Boixo S, Aspuru-Guzik A. Computational complexity of time-dependent density functional theory. *New J Phys*. (2014) 16:083035. doi: 10.1088/1367-2630/16/8/083035
19. Brown J, Yang J, Whitfield JD. Solver for the electronic V-representation problem of time-dependent density functional theory. *J Chem Theory Comput*. (2020) 16:6014–26. doi: 10.1021/acs.jctc.9b00583
20. Kitaev AY. Quantum measurements and the Abelian stabilizer problem. *arXiv [Preprint]*. (1995) arXiv:quant-ph/9511026.
21. Cleve R, Ekert A, Macchiavello C, Mosca M. Quantum algorithms revisited. *Proc R Soc Lond Ser A Math Phys Eng Sci*. (1998) 454:339–54. doi: 10.1098/rspa.1998.0164
22. Jordan S. *Quantum Algorithm Zoo*. (2019). Available online at: https://quantumalgorithmzoo.org/
23. Harrow AW, Hassidim A, Lloyd S. Quantum Algorithm for linear systems of equations. *Phys Rev Lett*. (2009) 103:150502. doi: 10.1103/PhysRevLett.103.150502
24. Shor PW. Polynomial-time algorithms for prime factorization and discrete logarithms on a quantum computer. *SIAM J Comput*. (1997) 26:1484–509. doi: 10.1137/S0097539795293172
25. Shor P. Algorithms for quantum computation: Discrete logarithms and factoring. In: *Proceedings of the 35th Annual IEEE Symposium on Foundations of Computer Science* (1994). p. 124–34.
26. Kim J, Lee JS, Lee S. Implementing unitary operators in quantum computation. *Phys Rev A*. (2000) 61:032312. doi: 10.1103/PhysRevA.61.032312
27. Whitfield JD, Biamonte J, Aspuru-Guzik A. Simulation of electronic structure Hamiltonians using quantum computers. *Mol Phys*. (2011) 109:735–50. doi: 10.1080/00268976.2011.552441
28. Szabo A, Ostlund N. Modern quantum chemistry: introduction to advanced electronic structure theory. In: *Dover Books on Chemistry*. Dover Publications (1996).
29. Ditchfield R, Hehre WJ, Pople JA. Self-Consistent Molecular-Orbital Methods. IX. An extended Gaussian-type basis for molecular-orbital studies of organic molecules. *J Chem Phys*. (1971) 54:724–8. doi: 10.1063/1.1674902
30. Smith DGA, Burns LA, Simmonett AC, Parrish RM, Schieber MC, Galvelis R, et al. PSI4 1.4: Open-source software for high-throughput quantum chemistry. *J Chem Phys*. (2020) 152:184108. doi: 10.26434/chemrxiv.11930031.v1
31. Smith RS, Curtis MJ, Zeng WJ. A Practical quantum instruction set architecture. *arXiv [Preprint]:160803355*. (2016).
32. Hatano N, Suzuki M. *Finding Exponential Product Formulas of Higher Orders*. Lecture Notes in Physics (2005). p. 37–68. doi: 10.1007/11526216_2
33. Mandelshtam VA, Taylor HS. Harmonic inversion of time signals and its applications. *J Chem Phys*. (1997) 107:6756–69. doi: 10.1063/1.475324
34. Mandelshtam VA. FDM: the filter diagonalization method for data processing in NMR experiments. *Prog Nucl Mag Res Sp*. (2001) 38:159–96. doi: 10.1016/S0079-6565(00)00032-7
35. Johnson SG. *Harmonic Inversion of Time Signals by the Filter Diagonalization Method*. GitHub (2005).
36. Granade CE, Ferrie C, Wiebe N, Cory DG. Robust online Hamiltonian learning. *New J Phys*. (2012) 14:103013. doi: 10.1088/1367-2630/14/10/103013
37. Granade C, Ferrie C, Hincks I, Casagrande S, Alexander T, Gross J, et al. QInfer: Statistical inference software for quantum applications. *Quantum*. (2017) 1:5. doi: 10.22331/q-2017-04-25-5
38. Gharibian S, Huang Y, Landau Z, Shin SW. Quantum Hamiltonian complexity. *Found Trends Theor Comput Sci*. (2015) 10:159–282. doi: 10.1561/0400000066

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Physics

ORIGINAL RESEARCH
published: 29 June 2021
doi: 10.3389/fphy.2021.589626

Training Restricted Boltzmann Machines With a D-Wave Quantum Annealer

Vivek Dixit[1,2], *Raja Selvarajan*[1,2], *Muhammad A. Alam*[3,4], *Travis S. Humble*[5] *and Sabre Kais*[1,2,4]*

[1]Department of Chemistry, Purdue University, West Lafayette, IN, United States, [2]Department of Physics and Astronomy, Purdue University, West Lafayette, IN, United States, [3]Department of Electrical and Computer Engineering, Purdue University, West Lafayette, IN, United States, [4]Birck Nanotechnology Center, Purdue University, West Lafayette, IN, United States, [5]Quantum Computing Institute, Oak Ridge National Laboratory, Oak Ridge, TN, United States

Edited by:
Jacob D. Biamonte,
Skolkovo Institute of Science and Technology, Russia

Reviewed by:
Marcos César de Oliveira,
State University of Campinas, Brazil
Soumik Adhikary,
Skolkovo Institute of Science and Technology, Russia

****Correspondence:***
Sabre Kais
kais@purdue.edu

Specialty section:
This article was submitted to Quantum Engineering and Technology, a section of the journal Frontiers in Physics

Received: *04 August 2020*
Accepted: *17 June 2021*
Published: *29 June 2021*

Citation:
Dixit V, Selvarajan R, Alam MA, Humble TS and Kais S (2021) Training Restricted Boltzmann Machines With a D-Wave Quantum Annealer. Front. Phys. 9:589626. doi: 10.3389/fphy.2021.589626

Restricted Boltzmann Machine (RBM) is an energy-based, undirected graphical model. It is commonly used for unsupervised and supervised machine learning. Typically, RBM is trained using contrastive divergence (CD). However, training with CD is slow and does not estimate the exact gradient of the log-likelihood cost function. In this work, the model expectation of gradient learning for RBM has been calculated using a quantum annealer (D-Wave 2000Q), where obtaining samples is faster than Markov chain Monte Carlo (MCMC) used in CD. Training and classification results of RBM trained using quantum annealing are compared with the CD-based method. The performance of the two approaches is compared with respect to the classification accuracies, image reconstruction, and log-likelihood results. The classification accuracy results indicate comparable performances of the two methods. Image reconstruction and log-likelihood results show improved performance of the CD-based method. It is shown that the samples obtained from quantum annealer can be used to train an RBM on a 64-bit "bars and stripes" dataset with classification performance similar to an RBM trained with CD. Though training based on CD showed improved learning performance, training using a quantum annealer could be useful as it eliminates computationally expensive MCMC steps of CD.

Keywords: bars and stripes, quantum annealing, classification, image reconstruction, log-likelihood, machine learning, D-wave, RBM (restricted Boltzmann machine)

1 INTRODUCTION

Quantum computing holds promise for a revolution in the field of science, engineering, and industry. Most of the R&D work related to quantum computing is focused on gate based approach [1–3], an alternative to this is the adiabatic quantum computing (AQC) [4–7]. In AQC, a system of qubits starts with a simple Hamiltonian whose ground state is known. Gradually, the initial Hamiltonian evolves into a final Hamiltonian. The final Hamiltonian is designed in such a way that its ground state corresponds to the solution to the problem of interest. According to the quantum adiabatic theorem, a quantum system that begins in the non-degenerate ground state of a time-dependent Hamiltonian will remain in the instantaneous ground state provided the Hamiltonian changes sufficiently slowly [8–11]. It has been shown theoretically that an AQC machine can give solutions that are very difficult to find using classical methods [12].

D-Wave's quantum annealer has been investigated by several researchers for machine learning and optimization problems. Mott *et al.* [13] used D-Wave to classify Higgs-boson-decay signals vs. background. They showed that the quantum annealing-based classifiers perform comparably to the state-of-the-art machine learning methods. Das *et al.* has used a D-Wave for clustering applications [14]. Mniszewski *et al.* [15] found that the results for graph partitioning using D-Wave systems are comparable to commonly used methods. Alexandrov *et al.* [16] used a D-Wave for matrix factorization. Lidar *et al.* [17] used a D-Wave for the classification of DNA sequences according to their binding affinities. Kais *et al.* have used D-Wave's quantum annealer for prime factorization and electronic structure calculation of molecular systems [18, 19].

RBM is a widely used machine learning technique for unsupervised and supervised tasks. However, its training is time consuming due to the calculation of model-dependent term in gradient learning. RBMs are usually trained using a method known as Contrastive Divergence (CD). CD uses Markov chain Monte Carlo (MCMC) which requires a long equilibration time. Further, the CD does not follow the gradient of the log-likelihood [20] and is not guaranteed to give correct results. Therefore, better sampling methods can have a positive impact on RBM learning. Among other works related to the topic, Adachi *et al.* [21] used quantum annealing for training RBMs, which were further used as layers of a two-layered deep neural network and post-trained by the back-propagation algorithm. The authors conclude that the hybrid approach results in faster training, although the relative effectiveness of RBM trained using a quantum-annealer vs. contrastive divergence has not been documented. Benedetti *et al.* [22] used a D-Wave quantum annealer to train an RBM on a 16-bit bars and stripes dataset. To train the RBM effectively an instance dependent temperature was calculated during each iteration. Caldeira *et al.* [23] used a QA-trained RBM for galaxy morphology image classification. Principal component analysis was used to compress the original dataset. They also explored the use of temperature estimation and examined the effect of noise by comparing the results from an older machine and a newer lower-noise version. Sleeman *et al.* [24] investigated a hybrid system that combined a classical deep neural network autoencoder with a QA-based RBM. Two datasets, the MNIST and the MNIST Fashion datasets, were used in this study. Image classification and image reconstruction were investigated. Winci *et al.* [25] developed a quantum-classical hybrid algorithm for a variational autoencoder (VAE). A D-Wave quantum annealer was used as a Boltzmann sampler for training the VAE. Dymtro *et al.* [26] performed a benchmarking study, to show that for harder problems Boltzmann machines trained using quantum annealing gives better gradients as compared to CD. Lorenzo *et al.* [27] used RBM trained with reverse annealing to carry out semantic learning that achieved good scores on reconstruction tasks. Koshka *et al.* [28] showed D-Wave quantum annealing performed better than classical simulated annealing for RBM training when the number of local valleys on the energy landscape was large. Dumoulin *et al.* [29] assessed the effect of various parameters like limited connectivity, and noise in weights and biases of RBM on its performance. Koshka *et al.* explored the energy landscape of an RBM embedded onto a D-Wave machine, which was trained with CD [30–33]. Dixit *et al.* [34] used a QA-trained RBM to balance the ISCX cybersecurity dataset, and training an intrusion detection classifier. There has been growing interest in quantum machine learning including Boltzmann machines [35–38], however, training quantum machine learning models on a moderate or large dataset is challenging. An RBM with 64 visible and 64 hidden units can be trained using a quantum annealer which is very difficult to do using existing gate based approaches.

In this work, our objective is to train an RBM using quantum annealing (QA) *via* samples obtained from the D-Wave 2000Q quantum annealer and compare its performance with an RBM trained with CD. The model-dependent term in the gradient of log-likelihood has been estimated by using samples drawn from a quantum annealer. Trained models are compared with respect to classification accuracy, image reconstruction, and log-likelihood values. To carry out this study, the bars and stripes (BAS) dataset has been used.

2 METHODS

2.1 Restricted Boltzmann Machine

A Restricted Boltzmann Machine is an energy-based model, inspired by the Boltzmann distribution of energies for the Ising model of spins. An RBM models the underlying probability distribution of a dataset and can be used for machine learning applications. However, an efficient method of RBM training is still not discovered. An RBM is comprised of two layers of binary variables known as visible and hidden layers. The variables or units in the visible and hidden layers are denoted as $\{v_1, v_2, \ldots, v_n\}$ and $\{h_1, h_2, \ldots, h_m\}$, respectively. The variables in one layer interact with the variables in the other layer, however, interactions between the variables in the same layer are not permitted. The energy of the model is given by:

$$E = -\sum_i b_i v_i - \sum_j c_j h_j - \sum_{i,j} v_i w_{ij} h_j, \tag{1}$$

where b_i and c_j are bias terms; w_{ij} represents the strength of the interaction between variables v_i and h_j. Let us represent the variables in the visible layer collectively by a vector: $v \in \{0,1\}^n$, similarly for the hidden layer: $h \in \{0,1\}^m$. Using this representation **Eq. 1** can be written as:

$$E(v,h) = -b^T v - c^T h - h^T W v, \tag{2}$$

where b and c are bias vectors at the visible and hidden layer, respectively; W is a weight matrix composed of w_{ij} elements. The probability that the model assigns to the configuration $\{v, h\}$ is:

$$P(v,h) = \frac{1}{Z} e^{-E(v,h)}, \qquad Z = \sum_v \sum_h e^{-E(v,h)}, \tag{3}$$

where Z is the partition function. Substituting value of $E(v,h)$, from **Eq. 2**, we get:

$$Z = \sum_v \sum_h e^{b^T v + c^T h + h^T \cdot W \cdot v} = \sum_h e^{c^T h} \sum_v e^{b^T v + h^T \cdot W \cdot v} \tag{4}$$

$$Z = \sum_h e^{c^T h} \sum_v e^{\left(b^T + h^T \cdot W\right) v} = \sum_h e^{c^T h} \sum_v e^{s \cdot v}, \tag{5}$$

where s is:

$$s = b^T + h^T W = [s_1, s_2, \ldots, s_n]; \tag{6}$$

n is the number of variables in the visible layer. Now, Z can be written as:

$$Z = \sum_h e^{c^T h} \prod_{j=1}^{n} \left(1 + e^{s_j}\right) \tag{7}$$

From **Eq. 7**, we notice that the calculation of Z involves summation over 2^m configuration, where m is the number of variables in the hidden layer. On the contrary, we need 2^{m+n} configurations to evaluate Z using **Eq. 3**.

2.2 Maximization of the Log-likelihood Cost Function

The partition function, Z, is hard to evaluate. The joint probability, $P(v, h)$, being a function of Z is also hard. Due to the bipartite graph structure of the RBM, the conditional distributions $P(h|v)$ and $P(v|h)$ are simple to compute,

$$P(h|v) = \frac{P(v,h)}{P(v)} \tag{8}$$

where $P(v)$ is given by the following expression:

$$P(v) = \frac{\sum_h e^{-E(v,h)}}{Z}. \tag{9}$$

Substituting values from **Eq. 3** and **Eq. 9** into **Eq. 8** gives:

$$P(h|v) = \frac{exp\left\{\sum_j c_j h_j + \sum_j \left(v^T W\right)_j h_j\right\}}{Z'}, \tag{10}$$

where

$$Z^{'} = \sum_h exp\left(c^T h + h^T W v\right). \tag{11}$$

$$P(h|v) = \frac{1}{Z'} \prod_j exp\left\{c_j h_j + \left(v^T W\right)_j h_j\right\} \tag{12}$$

Let's denote

$$\tilde{P}\left(h_j|v\right) = exp\left\{c_j h_j + \left(v^T W\right)_j h_j\right\} \tag{13}$$

Now, the probability to find an individual variable in the hidden layer $h_j = 1$ is:

$$P\left(h_j = 1|v\right) = \frac{\tilde{P}\left(h_j = 1|v\right)}{\tilde{P}\left(h_j = 0|v\right) + \tilde{P}\left(h_j = 1|v\right)} = \frac{exp\left\{c_j + \left(v^T W\right)_j\right\}}{1 + exp\left\{c_j + \left(v^T W\right)_j\right\}} \tag{14}$$

Thus, the individual hidden activation probability is given by:

$$P\left(h_j = 1|v\right) = \sigma\left(c_j + \left(v^T W\right)_j\right), \tag{15}$$

Where σ is the logistic function. Similarly, the activation probability of a visible variable conditioned on a hidden vector h is given by:

$$P\left(v_i = 1|h\right) = \sigma\left(b_i + \left(h^T W\right)_i\right). \tag{16}$$

An RBM is trained by maximizing the likelihood of the training data. The log-likelihood is given by:

$$l(W, b, c) = \sum_{t=1}^{N} \log P\left(v^{(t)}\right) = \sum_{t=1}^{N} \log \sum_h P\left(v^{(t)}, h\right) \tag{17}$$

$$l(W, b, c) = \sum_{t=1}^{N} \log \sum_h e^{-E\left(v^{(t)}, h\right)} - N \cdot \log \sum_{v,h} e^{-E(v,h)}. \tag{18}$$

Where $v^{(t)}$ is a sample from the training dataset. Denote $\theta = \{W, b, c\}$. The gradient of the log-likelihood is given by:

$$\begin{aligned}\nabla_\theta l(\theta) = \sum_{t=1}^{N} \frac{\sum_h e^{-E\left(v^{(t)},h\right)} \nabla_\theta\left(-E\left(v^{(t)}, h\right)\right)}{\sum_h e^{-E\left(v^{(t)},h\right)}} - N \\ \cdot \frac{\sum_{v,h} e^{-E(v,h)} \nabla_\theta\left(-E(v,h)\right)}{\sum_{v,h} e^{-E(v,h)}}\end{aligned} \tag{19}$$

$$\begin{aligned}\nabla_\theta l(\theta) = \sum_{t=1}^{N} \left\langle \nabla_\theta\left(-E\left(v^{(t)}, h\right)\right)\right\rangle_{P\left(h|v^{(t)}\right)} - N \\ \cdot \left\langle \nabla_\theta\left(-E(v,h)\right)\right\rangle_{P(v,h)},\end{aligned} \tag{20}$$

Where $\langle \cdot \rangle_{P(v,h)}$ is the expectation value with respect to the distribution $P(v, h)$. The gradient with respect to θ can also be expressed in terms of its components:

$$\nabla_w l = \frac{1}{N} \sum_{t=1}^{N} \left\langle v^{(t)} \cdot h^{(t)} \right\rangle_{P\left(h|v^{(t)}\right)} - \langle v \cdot h \rangle_{P(v,h)} \tag{21}$$

$$\nabla_b l = \frac{1}{N} \sum_{t=1}^{N} \left\langle v^{(t)} \right\rangle_{P\left(h|v^{(t)}\right)} - \langle v \rangle_{P(v,h)} \tag{22}$$

$$\nabla_c l = \frac{1}{N} \sum_{t=1}^{N} \left\langle h^{(t)} \right\rangle_{P\left(h|v^{(t)}\right)} - \langle h \rangle_{P(v,h)} \tag{23}$$

The first term in **Eq. 20** is the expectation value of $\nabla_\theta\left(-E\left(v^{(t)}, h\right)\right)$ with respect to the Boltzmann distribution, $v^{(t)}$ is a row vector from the training dataset with N records, and h is a hidden vector. Given $v^{(t)}$, h can be calculated via **Eq. 15**.

The second term in **Eq. 20** is a model-dependent term, the expectation value of $\nabla_\theta\left(-E(v, h)\right)$, v and h can be any possible binary vectors. This term is difficult to evaluate as it requires all possible combinations of v and h. Generally, this term is estimated using contrastive divergence, where one uses many cycles of Gibbs sampling to transform the training data into data drawn from the proposed distribution. We used **Eq. 15** and **Eq. 16** to sample from hidden and visible layers repeatedly. Once we have the gradient of log-likelihood (**Eq. 18**), weights and biases can be estimated using gradient ascent optimization:

$$\theta_j^{new} = \theta_j^{old} + \epsilon \cdot \nabla_{\theta_j} l\left(\theta_j\right) \tag{24}$$

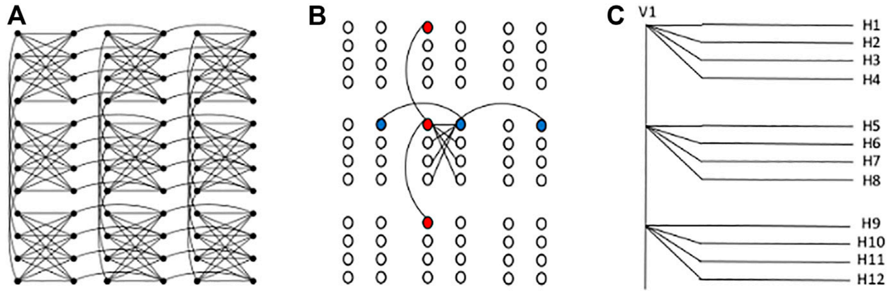


FIGURE 1 | (A) C3 Chimera graph of qubits. **(B)** Three vertical (red) and horizontal (blue) qubits are chained to form a visible and a hidden unit. **(C)** Connectivity of a visible unit (V1) with 12 hidden units (H1 to H12). Here, each unit is formed by ferromagnetic couplings between three qubits.

where ϵ is the learning rate.

Alternatively, the second term can be calculated using samples drawn from the D-Wave quantum annealer, which is a faster procedure than MCMC.

2.3 D-Wave Hamiltonian and Arrangement of Qubits

The Hamiltonian for a D-Wave system of qubits can be represented as:

$$h_{Ising} = -\frac{A(s)}{2}\left(\sum_i \hat{\sigma}_x^i\right) + \frac{B(s)}{2}\left(\sum_i h_i \hat{\sigma}_z^i + \sum_{(i>j)} J_{ij}\hat{\sigma}_z^i \hat{\sigma}_z^j\right) \tag{25}$$

where $\hat{\sigma}_{x,z}^{(i)}$ are Pauli matrices operating on i^{th} qubit. h_i and J_{ij} are the qubit biases and coupling strengths. s is called the anneal fraction. $A(s)$ and $B(s)$ are known as anneal functions. At $s = 0$, $A(s) \gg B(s)$, while $A(s) \ll B(s)$ for $s = 1$. As we increase s from 0 to 1, anneal functions change gradually to meet these boundary conditions. In the standard quantum annealing (QA) protocol, s changes from 0 to 1. The network of qubits starts in a global superposition over all possible classical states and after $s = 1$, the system is measured in a single classical state.

The arrangement of qubits on the D-Wave 2000Q quantum annealer forms a C16 Chimera graph with 16×16 unit cells (2048 qubits are mapped into 16×16 matrices of unit cells; each unit cell has eight qubits). **Figure 1A** shows a C3 Chimera graph with 3×3 unit cells. Within each unit cell, there are two sets of four qubits that are connected in a bipartite fashion. As shown in the figure, each qubit in a unit cell is connected to four qubits of the same unit cell and two qubits of other unit cells. Thus, each qubit can be connected to a maximum of six qubits. This connectivity can be enhanced by forming strong ferromagnetic couplings between the qubits, which forces coupled qubits to stay in the same state.

2.4 Restricted Boltzmann Machine Embedding Onto the D-Wave QPU

Mapping an AQC algorithm on specific hardware is nontrivial and requires creative mapping. Several algorithms can be used to map a graph to the physical qubits on an adiabatic quantum computer [39, 40]. However, it is nontrivial to find a simple embedding when the graph size is large. Taking into consideration the arrangement of qubits on the 2000Q processor, we found a simple embedding that utilizes most of the working qubits. In the present study, we investigated RBMs in two configurations, one with 64 visible units and 64 hidden units, another with 64 visible units and 20 hidden units. Here, we will discuss the embedding of the RBM with 64 units in both layers. Each unit of the RBM is connected to 64 other units, but in the D-Wave each qubit only connects to six other qubits. To enhance the connectivity, qubits can be coupled together or cloned by setting $J_{i,j} = -1$. This forces the two qubits to stay in the same state. In our embedding, one unit of RBM is formed by connecting 16 qubits. The D-Wave processor has qubits arranged in 16×16 matrices of unit cells. Each unit cell has two sets of four qubits arranged in a bipartite fashion. Each qubit in the left column of the unit cell can be connected to one qubit of the unit cell just above it and one just below it. There are 16 unit cells along one side, so a chain of 16 qubits can be formed. This chain forms one visible unit of the RBM. **Figure 1B** shows the procedure to couple three qubits to form a chain that represents a visible unit. The qubits that are connected together to form a vertical chain forming a visible unit are shown in red. Since there are four qubits in the left column of the unit cell, four chains can be formed resulting in four visible units of RBM. The four qubits that form the right column of the unit cell can be connected to form horizontal chains as shown in **Figure 1B**. These horizontal chains form the hidden units. There are 16 unit cells along the horizontal direction in a C16 Chimera graph, therefore each horizontal chain is also composed of 16 qubits. Utilizing the arrangement of qubits of the D-Wave QPU, 64 vertical and 64 horizontal chains can be formed representing the 64 visible and 64 hidden units of RBM. **Figure 1C** shows the scheme that we used to connect one visible unit (V1) to the hidden units. In this fashion, one can embed an RBM with 64 visible and 64 hidden units on a C16 Chimera graph. Of course, care must be taken for any inaccessible qubits to form a further restricted RBM. In our experiments, we found that the absence of a few qubits does not affect the performance of the resulting network.

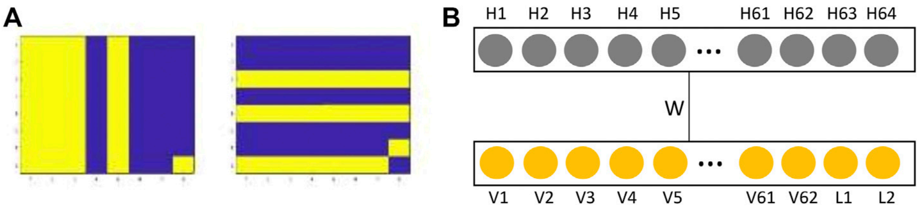


FIGURE 2 | (A) Example of the bars **(left)** and stripes **(right)** patterns of size 8 × 8. A blue cell represents a 0 while a yellow cell a 1. Last two bits (bottom right) are labels; bar = 01, stripe = 10. **(B)** RBM for classification. Box with yellow units is the visible layer. The hidden layer is shown by a box with grey units. 8 × 8 bits of an image is rearranged into a record of size 1 × 64 bits which is applied to the visible layer. Units L1 and L2 are the labels.

2.5 Classification and Image Reconstruction

Each record of the bars and stripes dataset is made up of 64 bits. The last two bits are for labeling the pattern: 01 for a bar and 10 for a stripes pattern–**Figure 2A**. If the last two bits are 00 or 11, the prediction by RBM is incorrect. Once we obtained the weights and biases of the RBM from the training step, RBM can be used for classification or image reconstruction. To predict the class of a test record we apply its first 62 bits at the visible layer **Figure 2B**. We randomly input either zero or one for the last two classifying bits (L1 and L2). We then run 50 Gibbs cycles, keeping the 62 visible units clamped at the values of the test record. At the end of 50 Gibbs cycles label units, 63 and 64 are read. $L1 = 0$ and $L2 = 1$ indicates a bar pattern, while $L1 = 1$ and $L2 = 0$ suggests a stripe pattern. For the problem of image reconstruction, the goal is to predict the missing part of an image. A similar procedure can be applied for image reconstruction where a trained RBM is used to predict the values of the missing units. In this case, we clamp the visible units where values are given, and run 50 Gibbs cycles, at the end, we sample from the units where values have to be predicted.

3 RESULTS AND DISCUSSION

In the present work, we have used the bars and stripes (BAS) dataset. An example of a bar and a stripe pattern is shown in **Figure 2A**. BAS is a popular dataset for RBM training, it has been used by several researchers [22, 41–43]. This is a binary dataset consisting of records of 64 bits in length, with the last 2 bits representing the label of the record: 01 for a bar and 10 for a stripe pattern. Our dataset is comprised of 512 unique records. The number of unique samples used for training is 400, with the remaining 112 samples were held for testing. Classification of bars and stripes, image reconstruction, and the log-likelihood values are used to compare the performances of trained RBMs.

3.1 Restricted Boltzmann Machine Training

Equation 20 has been used to train the RBMs. The first term in this equation is a data-dependent term which can be exactly calculated using the conditional probabilities $P(h|v)$ and $P(v|h)$ given by **Eq. 15** and **Eq. 16**. The second term is a model-dependent, which requires expectation value over all possible hidden h and visible v vectors, which is clearly intractable. Typically, the model-dependent term is approximately estimated using a method known as contrastive divergence (CD). In this approach, samples needed to calculate the model-dependent term are obtained by running the Gibbs chain starting from a sample from the training data (**Figure 3**). If n Gibbs steps are performed, the method is known as CD-n. It is shown by Hinton that $n = 1$ could be sufficient for convergence (CD-1) [20]. In CD-1, first, a data sample is applied at the visible layer, then **Eq. 15** is used to generate a corresponding hidden vector at the hidden layer. Now, this hidden vector is used to generate a new visible vector using **Eq. 16**, which is in turn used to generate a new hidden vector. These new visible and hidden vectors are used to calculate the model-dependent term. This process is repeated for each record in the dataset. A detailed description of RBM training using CD is given in a review article by Hinton [44]. The model-dependent term can also be calculated using samples (v and h) obtained from an RBM mapped on the D-Wave. From **Eq. 19** we notice that in the second term the expectation value should be calculated with respect to $e^{-E(v,h)}$ distribution, while samples from the D-Wave follow a distribution of $e^{-\frac{E(v,h)}{kT}}$. It should be noted as the RBM training starts, the weights and biases are random, and samples from the D-Wave are not expected to have a Boltzmann distribution, however, as the training progresses the underlying probability distribution moves toward the Boltzmann distribution. Following the approach used by Adachi *et al.* [21], we used a hyperparameter, S, such that for the model-dependent term, we sample from $e^{\frac{-E(v,h)}{SkT}}$ distribution. Here, S is a hyperparameter, which is determined by the calculation of the classification accuracy for various values of S. The optimal condition corresponds to the case when $SkT = 1$. A different approach was taken by Benedetti *et al.* [22]. They calculated effective temperature during each epoch. Their approach is difficult to apply in the present case of 64 bits record length BAS dataset. The BAS dataset that they used was comprised of just 16-bit records. A complex dataset leads to a complicated distribution, which makes training with a dynamical effective temperature difficult.

In order to train an RBM using D-Wave, model parameters (w_{ij}, b_i and c_j) were initialized with random values, the first term of **Eq. 20** was calculated exactly using these weights and biases, and the training dataset. The weights and biases were then used to embed the RBM onto the D-Wave QPU, and quantum annealing

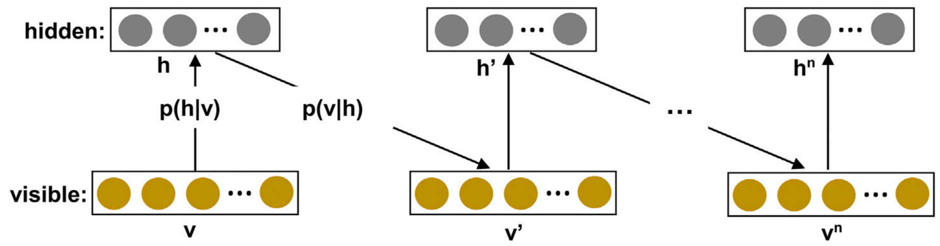


FIGURE 3 | Illustration of the contrastive divergence algorithm for training a Restricted Boltzmann Machine (RBM). If sampling stops at n_{th} Gibbs step of the Markov Chain, the procedure is known as CD-n.

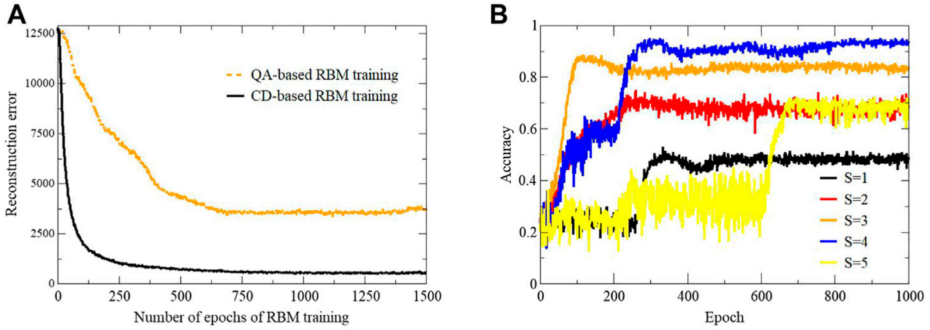


FIGURE 4 | (A) Reconstruction error versus epoch of the RBM training. **(B)** The plot of the classification accuracy versus epoch for different values of the hyperparameter *S*.

was performed. Once annealing was complete, the D-Wave returned low energy solutions. Based on the mapping of the RBM, visible v and hidden h vectors were obtained from the solutions returned from the D-Wave. These v and h samples were used to calculate the model-dependent expectation value which in turn gives the gradient of log-likelihood. The gradient was further used to calculate new weights and biases (**Eq. 24**). The whole process was repeated until some convergence criterion was achieved. One of the several ways to monitor the progress of model learning during the RBM training is by estimating the reconstruction error for each training epochs. The reconstruction error is defined as:

$$\text{Reconstruction error} = \sum_{i=1}^{n}\sum_{t=1}^{N}\left(v^{(t)} - v'\right)^2 \tag{26}$$

where $v^{(t)}$ is a data record and v' is the reconstructed visible vector (**Figure 3**). N and n are the number of records in the training dataset and the number of units in the visible layer, respectively. The plot of reconstruction error versus epoch is presented in the left panel of **Figure 4**. An optimal value of the empirical parameter S is important for a correct sampling of v and h vectors. The effect of change in S on the classification accuracy is shown in the right-side panel of **Figure 4**, a plot between accuracy and epoch. The term epoch means a full cycle of iterations, with each training pattern participating only once. Accuracy is defined as:

$$\text{Accuracy} = \frac{\text{Number of correct predictions}}{\text{Total number of predictions}} \tag{27}$$

The classification accuracy is maximum for $S = 4$. The performance of the model during the training process can be visualized by plotting classification accuracy with epochs. **Figure 5** shows the plot of classification accuracy vs. epochs for bars (left) and stripes (right) patterns. This calculation was performed on the test dataset. As the number of epoch increases from 0 to 400, the classification accuracy increases after that it stays constant. Based on these results, we conclude that the performance of QA-trained RBM is similar to CD-trained RBM. However, from **Figure 5** we notice that there are higher fluctuations in the classification accuracy with CD-1 based training.

3.2 Image Reconstruction

For classification tasks, both training methods (QA and CD-1) showed similar results. Classification task requires the prediction of target labels (only 2 bits) based on the features in the dataset. An input data record is applied at the visible layer and the target labels are reconstructed. A more difficult task would be the reconstruction of not just the target labels, but also some other bits of the record. We call this task - image reconstruction. Here, we take a 64-bit record from the test dataset, corrupt some of its bits, and then apply this modified test record to the visible layer of a trained RBM. We follow the

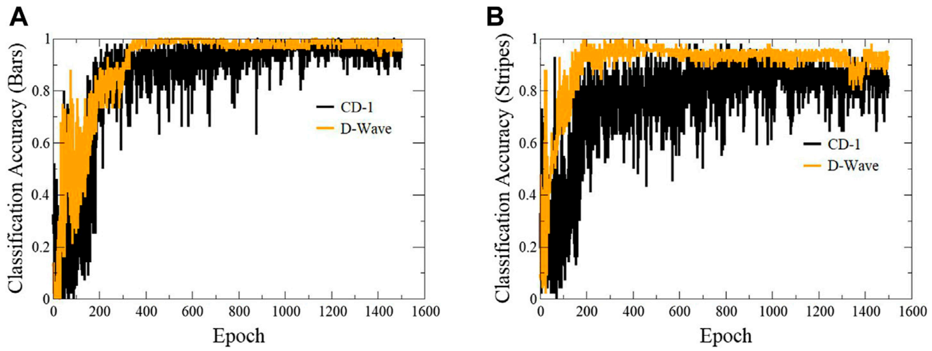


FIGURE 5 | Plots showing classification accuracy of individual classes with epoch. The gradient of log-likelihood was calculated using samples generated via contrastive divergence and D-Wave's quantum annealer. A comparison of classification accuracy for both methods is presented for bars **(A)** and stripes **(B)** patterns.

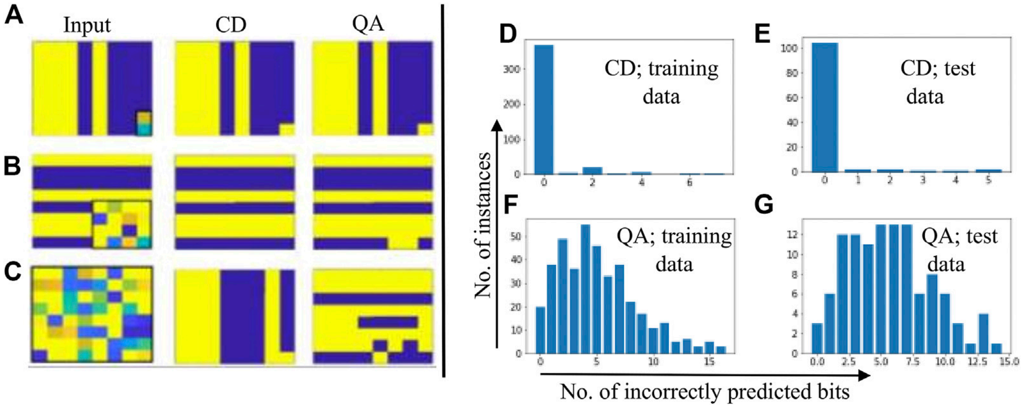


FIGURE 6 | Original data was first corrupted, then reconstructed. **Left panel:** The images in the left column are input data fed to the RBMs; output obtained from the CD and QA-trained RBMs is shown in the middle and right column, respectively. The bits of the input images that were corrupted are enclosed in black boxes. **Right panel:** Histograms showing the distribution of incorrectly predicted bits versus the number of instances; 16 bits of the records were corrupted and then reconstructed using trained RBMs.

procedure explained earlier for the image reconstruction. The results of image reconstruction are presented in **Figure 6**. In **Figure 6A**, only target labels were corrupted/reconstructed. We notice that both training methods correctly reproduced the classifying labels. In the second case, **Figure 6B**, 16 bits of the original data record were corrupted. The RBM trained using CD-1 correctly predicted all the bit, while two bits were incorrectly predicted by the RBM trained with QA. In the third case, **Figure 6C** completely random 64-bit input vector (all bits corrupted) was fed to both RBMs. In the case of CD-1, the output is a bar pattern, whereas QA trained RBM resulted in a stripes pattern with many bits incorrectly predicted. **Figure 6B** shows a particular case where 16 bits of a record were corrupted and then reconstructed, histograms in the right panel of **Figure 6** show the results when 16 bits of all the records of the dataset were corrupted and then reconstructed. These histograms show the plots between the number of instances of the dataset versus the number of incorrectly predicted bits. **Figure 6D** and **Figure 6E** show the results where CD-trained RBM was used for image reconstruction. **Figure 6D** shows the case where the records of the training data were corrupted and fed to a CD-trained RBM for performing the reconstruction. The histogram shows that in over 350 cases all the bits were correctly predicted. In the cases where some bits were not correctly predicted, the number of incorrectly predicted bits was less than or equal to eight. **Figure 6E** shows a similar histogram for the test dataset which had 112 records. Around 100 records were correctly predicted. **Figure 6F** and **Figure 6G** show the histogram for the cases where QA-based trained RBM was used for image reconstruction. In **Figure 6F** where records from the training dataset were used, for most of the instances around 4 bits were incorrectly predicted. **Figure 6G** shows the results for the case where the records from the test dataset were used for image reconstruction. In this case, most of the reconstructed images show about 6–7 bits incorrectly predicted. From these plots, it is clear that the CD-trained RBM performed better than the QA-trained RBMs.

3.3 Log-likelihood Comparison

The classification accuracy results indicated similar performances of both methods (CD-1 and quantum annealing). However,

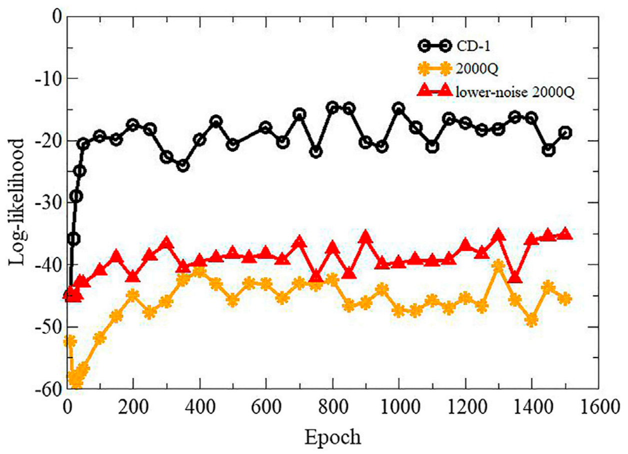


FIGURE 7 | Plot showing variation in log-likelihood with training epochs. Label CD-1 represents CD-based RBM training, while label 2000Q indicates QA-based RBM training employing the D-Wave 2000Q.

image reconstruction suggests the improved performance of CD-1. In order to further compare and quantify the performances of these two methods, the log-likelihood of training data was calculated. Several researchers have used "log-likelihood" in order to compare different RBM models [45, 46]. The log-likelihood has been computed using **Eq. 18**. It involves the computation of the partition function Z. If the number of units in the hidden layer is not too large, Z can be exactly calculated using **Eq. 7**. To calculate the log-likelihood, the number of hidden units was set to 20. The log-likelihood of both models was computed at various epochs. The results are presented in **Figure 7**. From the figure, we notice that the log-likelihood is higher for RBM trained using CD-1 compared to quantum annealing. A lower value of the log-likelihood for the D-Wave trained model could be attributed to a restricted range of allowed values for the bias field h and the coupling coefficients J. Another reason could be an instance (each set of h and J) dependent temperature variation during the RBM training [22], which disturbs the learning of embedded RBM. **Figure 7** also compares the log-likelihood values calculated using the new lower-noise D-Wave 2000Q processor and an earlier 2000Q processor. D-Wave's lower-noise machine shows slightly improved log-likelihood values over the entire training range.

3.4 Feature Reduction

Next, we compared the performance of a QA and CD-1 trained RBM for features reduction applications. Both RBMs that were used in this experiment had 64 visible and 20 hidden units. The BAS dataset was comprised of 62 binary features and two target labels. The target labels were removed and random 0 and 1s were added before feeding the dataset to an RBM for feature reduction. Input data vector from the training data was applied to the visible layer of the RBM. **Equation 15** was then used to obtain the hidden variables. The compressed feature vector was sampled from 20 hidden units of the hidden layer. This procedure was used to extract features corresponding to 400 records of the training data and 112 records of the test data. Several classifiers were used to compare the effectiveness of feature reduction of QA and CD-1 trained RBMs. The results are presented in **Table 1**. Both feature reduction methods resulted in high classification accuracy, with classifiers trained on data obtained from CD-trained RBM showing slightly improved classification accuracy than the classifiers trained on data from a QA-trained RBM.

Currently, quantum computers are in their developmental stage. Therefore, the comparison of quantum annealing based RBM training with a mature classical approach like CD is uneven. However, such comparison is important as results from a CD trained RBM provides a reference with respect to which performance of QA based RBM training can be assessed. Though a better performance of CD was expected, QA performed satisfactorily. For classification and feature reduction tasks both QA and CD performed comparably. For the image reconstruction task, CD performed better than QA. The lower performance of QA-based training could be attributed to two main reasons. First, we obtain samples from the D-Wave assuming that it operates at a fixed temperature. However, this is an approximation, and one should calculate an effective temperature during each epoch. This mismatch degrades RBM's learning during the training. Efforts have been made towards developing methods to estimate this instance-dependent temperature for small datasets [22, 23], none of which have been shown to be efficient for bigger datasets. Another reason for the lower performance of QA-trained RBM could be hardware limitations like limited connectivity, lower coherence time, noise, etc. These limitations will be removed as quantum annealing technology matures. D-Wave's new machine, advantage, has higher connectivity, qubits and lower noise compared to the older machine. In **Figure 7** we have shown that the new 2000Q machine performs better than the older machine. We believe as technology evolves and a new algorithm for calculation of effective temperature is developed QA based RBM training will improve and will be able to deal with larger datasets. Quantum annealing offers a fundamentally different approach to estimate model-dependent term of the gradient of log-likelihood compared to CD and PCD. Depending on the complexity of a dataset, the CD might need hundreds of Gibbs cycles to reach the

TABLE 1 | Comparison of classifiers trained on compressed BAS dataset obtained from CD-1 and QA trained RBMs. The classification accuracy was estimated on the test data comprised of 56 bars and 56 stripes instances. The label "No. of bars" (No. of stripes) indicates the number of correctly predicted bar (stripe) instances. The labels CD-1 and QA indicate that the feature reduction was performed using contrastive divergence and quantum annealing, respectively.

Classifier	Accuracy		No. of bars		No. of stripes	
	CD-1	QA	CD-1	QA	CD-1	QA
SVM	0.98	0.88	54	42	56	56
MLP	0.98	0.98	55	55	56	55
KNN	1.00	0.99	56	55	56	56
Decision tree	1.00	0.89	56	46	56	54
Gaussian process	0.99	1.00	55	56	56	56
Ada boost	0.82	0.79	50	43	42	45

equilibrium to finally give one sample, while using a QA-based approach one can obtain 10,000 samples almost instantaneously.

4 CONCLUSION

In this work, we present an embedding that can be used to embed an RBM with 64 visible and 64 hidden units. We trained an RBM by calculating the model-dependent term of the gradient of the log-likelihood using samples obtained from the D-Wave quantum annealer. The trained RBM was embedded onto the D-Wave QPU for classification and image reconstruction. We also showed that a new lower-noise quantum processor gives improved results. The performance of the RBM was compared with an RBM trained with a commonly used method called contrastive divergence (CD-1). Though both methods resulted in comparable classification accuracy, CD-1 training resulted in better image reconstruction and log-likelihood values. RBM training using the samples from a quantum annealer removes the need for time-consuming MCMC steps during training and classification procedures. These computationally expensive MCMC steps are an essential part of training and classification with CD-1. QA-based RBM learning could be improved by the calculation of an instance-dependent temperature and incorporating this temperature in the RBM training procedure though better methods to compute the effective temperature of the D-Wave machine on large datasets is still an open problem.

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusion of this article will be made available by the authors, without undue reservation.

AUTHOR CONTRIBUTIONS

SK, TH, and MA designed the work, VD and RS performed the calculations. All authors discussed the results and wrote the paper.

FUNDING

We are grateful for the support from Integrated Data Science Initiative Grants (IDSI F.90000303), Purdue University. SK would like to acknowledges funding by the U.S. Department of Energy (Office of Basic Energy Sciences) under Award No. DE-SC0019215. TH would like to acknowledge funding by the U.S. Department of Energy (Office of Basic Energy Science) under Award No. ERKCG12. This manuscript has been authored by UT-Battelle, LLC under Contract No. DE-AC05-00OR22725 with the U.S. Department of Energy.

ACKNOWLEDGMENTS

The United States Government retains and the publisher, by accepting the article for publication, acknowledges that the United States Government retains a non-exclusive, paid-up, irrevocable, worldwide license to publish or reproduce the published form of this manuscript, or allow others to do so, for United States Government purposes. The Department of Energy will provide public access to these results of federally sponsored research in accordance with the DOE Public Access Plan. This manuscript has been released as a pre-print at arXiv: 2005.03247 (cs.LG) [47].

REFERENCES

1. Rieffel E, and Polak W. *Quantum Computing: A Gentle Introduction*. Cambridge: MIT Press (2014).
2. Nielsen MA, and Chuang IL. *Quantum Computation and Quantum Information: 10th Anniversary Edition*. Cambridge: Cambridge University Press (2010). doi:10.1017/CBO9780511976667
3. Kais S. *Quantum Information and Computation for Chemistry*. Hoboken, New Jersey: John Wiley & Sons (2014).
4. Farhi E, Goldstone J, Gutmann S, and Sipser M. Quantum Computation by Adiabatic Evolution. *arXiv preprint quant-ph/0001106* (2000).
5. Aharonov D, van Dam W, Kempe J, Landau Z, Lloyd S, and Regev O. Adiabatic Quantum Computation Is Equivalent to Standard Quantum Computation. *SIAM J Comput* (2007) 37:166–94. doi:10.1137/S0097539705447323
6. Kadowaki T, and Nishimori H. Quantum Annealing in the Transverse Ising Model. *Phys Rev E* (1998) 58:5355–63. doi:10.1103/PhysRevE.58.5355
7. Finnila AB, Gomez MA, Sebenik C, Stenson C, and Doll JD. Quantum Annealing: A New Method for Minimizing Multidimensional Functions. *Chem Phys Lett* (1994) 219:343–8. doi:10.1016/0009-2614(94)00117-0
8. Santoro GE, and Tosatti E. Optimization Using Quantum Mechanics: Quantum Annealing through Adiabatic Evolution. *J Phys A: Math Gen* (2006) 39:R393–R431. doi:10.1088/0305-4470/39/36/r01
9. Das A, and Chakrabarti BK. Colloquium: Quantum Annealing and Analog Quantum Computation. *Rev Mod Phys* (2008) 80:1061–81. doi:10.1103/RevModPhys.80.1061
10. Albash T, and Lidar DA. Adiabatic Quantum Computation. *Rev Mod Phys* (2018) 90:015002. doi:10.1103/RevModPhys.90.015002
11. Boixo S, Rønnow TF, Isakov SV, Wang Z, Wecker D, Lidar DA, et al. Evidence for Quantum Annealing with More Than One Hundred Qubits. *Nat Phys* (2014) 10:218–24. doi:10.1038/nphys2900
12. Biamonte JD, and Love PJ. Realizable Hamiltonians for Universal Adiabatic Quantum Computers. *Phys Rev A* (2008) 78:012352. doi:10.1103/physreva.78.012352
13. Mott A, Job J, Vlimant J-R, Lidar D, and Spiropulu M. Solving a Higgs Optimization Problem with Quantum Annealing for Machine Learning. *Nature* (2017) 550:375–9. doi:10.1038/nature24047
14. Das S, Wildridge AJ, Vaidya SB, and Jung A. Track Clustering with a Quantum Annealer for Primary Vertex Reconstruction at Hadron Colliders. *arXiv preprint arXiv:1903.08879* (2019).
15. Ushijima-Mwesigwa H, Negre CFA, and Mniszewski SM. Graph Partitioning Using Quantum Annealing on the D-Wave System. In: Proceedings of the Second International Workshop on Post Moores Era Supercomputing, PMES'17. Denver, CO: ACM (2017). p. 22–9.
16. O'Malley D, Vesselinov VV, Alexandrov BS, and Alexandrov LB. Nonnegative/binary Matrix Factorization with a D-Wave Quantum Annealer. *PLOS ONE* (2018) 13:1–12. doi:10.1371/journal.pone.0206653
17. Li RY, Di Felice R, Rohs R, and Lidar DA. Quantum Annealing versus Classical Machine Learning Applied to a Simplified Computational Biology Problem. *Npj Quan Inf* (2018) 4:14. doi:10.1038/s41534-018-0060-8
18. Jiang S, Britt KA, McCaskey AJ, Humble TS, and Kais S. Quantum Annealing for Prime Factorization. *Sci Rep* (2018) 8:17667. doi:10.1038/s41598-018-36058-z

19. Xia R, Bian T, and Kais S. Electronic Structure Calculations and the Ising Hamiltonian. *J Phys Chem B* (2018) 122:3384–95. doi:10.1021/acs.jpcb.7b10371
20. Hinton GE. Training Products of Experts by Minimizing Contrastive Divergence. *Neural Comput* (2002) 14:1771–800. doi:10.1162/089976602760128018
21. Adachi SH, and Henderson MP. Application of Quantum Annealing to Training of Deep Neural Networks. *arXiv preprint arXiv:1510.06356* (2015).
22. Benedetti M, Realpe-Gómez J, Biswas R, and Perdomo-Ortiz A. Estimation of Effective Temperatures in Quantum Annealers for Sampling Applications: A Case Study with Possible Applications in Deep Learning. *Phys Rev A* (2016) 94: 022308. doi:10.1103/PhysRevA.94.022308
23. Caldeira J, Job J, Adachi SH, Nord B, and Perdue GN. Restricted Boltzmann Machines for Galaxy Morphology Classification with a Quantum Annealer. *arXiv preprint arXiv:1911.06259* (2019).
24. Sleeman J, Dorband J, and Halem M. *A Hybrid Quantum Enabled Rbm Advantage: Convolutional Autoencoders for Quantum Image Compression and Generative Learning* (2020), April 27–May 9, 2020, California: Society of Photo-Optical Instrumentation Engineers (SPIE).
25. Winci W, Buffoni L, Sadeghi H, Khoshaman A, Andriyash E, and Amin MH. A Path towards Quantum Advantage in Training Deep Generative Models with Quantum Annealers. *Mach Learn Sci Technol* (2020) 1:045028. doi:10.1088/2632-2153/aba220
26. Korenkevych D, Xue Y, Bian Z, Chudak F, Macready WG, and Rolfe J. *Benchmarking Quantum Hardware for Training of Fully Visible Boltzmann Machines* (2016).
27. Rocutto L, Destri C, and Prati E. Quantum Semantic Learning by Reverse Annealing of an Adiabatic Quantum Computer. *Adv Quan Tech* (2021) 4: 2000133. doi:10.1002/qute.202000133
28. Koshka Y, and Novotny MA. Comparison of D-Wave Quantum Annealing and Classical Simulated Annealing for Local Minima Determination. *IEEE J Sel Areas Inf Theor* (2020) 1:515–25. doi:10.1109/JSAIT.2020.3014192
29. Dumoulin V, IanGoodfellow JAC, and Bengio Y. On the Challenges of Physical Implementations of Rbms. In: AAAI Publications, Twenty-Eighth AAAI Conference on Artificial Intelligence (2014) Quebec, Canada, July 27–31, 2014.
30. Koshka Y, and Novotny MA. Comparison of Use of a 2000 Qubit D-Wave Quantum Annealer and Mcmc for Sampling, Image Reconstruction, and Classification. *IEEE Trans Emerg Top Comput Intell* (2021) 5:119–29. doi:10.1109/TETCI.2018.2871466
31. Koshka Y, and Novotny MA. 2000 Qubit D-Wave Quantum Computer Replacing Mcmc for Rbm Image Reconstruction and Classification. In: 2018 International Joint Conference on Neural Networks (IJCNN); July 8–13, 2018, Rio de Janeiro, Brazil (2018). p. 1–8. doi:10.1109/IJCNN.2018.8489746
32. Koshka Y, Perera D, Hall S, and Novotny MA. Determination of the Lowest-Energy States for the Model Distribution of Trained Restricted Boltzmann Machines Using a 1000 Qubit D-Wave 2x Quantum Computer. *Neural Comput* (2017) 29:1815–37. doi:10.1162/neco_a_00974
33. Koshka Y, Perera D, Hall S, and Novotny MA. Empirical Investigation of the Low Temperature Energy Function of the Restricted Boltzmann Machine Using a 1000 Qubit D-Wave 2X. In: 2016 International Joint Conference on Neural Networks (IJCNN); July 24–29, 2016, Vancouver, BC(2016). p. 1948–54. doi:10.1109/IJCNN.2016.7727438
34. Dixit V, Selvarajan R, Aldwairi T, Koshka Y, Novotny MA, and Humble TS. Training a Quantum Annealing Based Restricted Boltzmann Machine on Cybersecurity Data. In: IEEE Transactions on Emerging Topics in Computational Intelligence (2021).
35. Amin MH, Andriyash E, Rolfe J, Kulchytskyy B, and Melko R. Quantum Boltzmann Machine. *Phys Rev X* (2018) 8:021050. doi:10.1103/PhysRevX.8.021050
36. Lloyd S, Mohseni M, and Rebentrost P. Quantum Algorithms for Supervised and Unsupervised Machine Learning. *arXiv preprint arXiv: 1307.0411* (2013).
37. Rebentrost P, Mohseni M, and Lloyd S. Quantum Support Vector Machine for Big Data Classification. *Phys Rev Lett* (2014) 113:130503. doi:10.1103/PhysRevLett.113.130503
38. Wiebe N, Kapoor A, and Svore KM. Quantum Deep Learning. *arXiv preprint arXiv:1412.3489* (2014).
39. Goodrich TD, Sullivan BD, and Humble TS. Optimizing Adiabatic Quantum Program Compilation Using a Graph-Theoretic Framework. *Quan Inf Process* (2018) 17:1–26. doi:10.1007/s11128-018-1863-4
40. Date P, Patton R, Schuman C, and Potok T. Efficiently Embedding Qubo Problems on Adiabatic Quantum Computers. *Quan Inf Process* (2019) 18:117. doi:10.1007/s11128-019-2236-3
41. Krause O, Fischer A, and Igel C. Population-contrastive-divergence: Does Consistency Help with Rbm Training?. *Pattern Recognition Lett* (2018) 102: 1–7. doi:10.1016/j.patrec.2017.11.022
42. Schulz H, Müller A, and Behnke S. Investigating Convergence of Restricted Boltzmann Machine Learning. In: NIPS 2010 Workshop on Deep Learning and Unsupervised Feature Learning; December 2010, Whistler, Canada, 1 (2010). p. 6–1.
43. Upadhya V, and Sastry P. Learning Rbm with a Dc Programming Approach. *arXiv preprint arXiv:1709.07149* (2017).
44. Hinton GE. *A Practical Guide to Training Restricted Boltzmann Machines*. Berlin, Heidelberg: Springer (2012). p. 599–619. doi:10.1007/978-3-642-35289-8_32
45. Tieleman T. Training Restricted Boltzmann Machines Using Approximations to the Likelihood Gradient. In: Proceedings of the 25th International Conference on Machine Learning, ICML '08. New York, NY, USA: ACM (2008). p. 1064–71. doi:10.1145/1390156.1390290
46. Cho K, Raiko T, and Ilin A. Parallel Tempering Is Efficient for Learning Restricted Boltzmann Machines. In: The 2010 International Joint Conference on Neural Networks (IJCNN); July 18–23, 2010, Barcelona, Spain (2010). p. 1–8. doi:10.1109/IJCNN.2010.5596837
47. Dixit V, Selvarajan R, Alam MA, Humble TS, and Kais S. Training and Classification Using a Restricted Boltzmann Machine on the D-Wave 2000Q. *arXiv preprint arXiv:2005.03247* (2020).

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Advantages of publishing in Frontiers

OPEN ACCESS
Articles are free to read
for greatest visibility
and readership

FAST PUBLICATION
Around 90 days
from submission
to decision

HIGH QUALITY PEER-REVIEW
Rigorous, collaborative,
and constructive
peer-review

TRANSPARENT PEER-REVIEW
Editors and reviewers
acknowledged by name
on published articles

Frontiers
Avenue du Tribunal-Fédéral 34
1005 Lausanne | Switzerland

Visit us: www.frontiersin.org
Contact us: frontiersin.org/about/contact

REPRODUCIBILITY OF RESEARCH
Support open data
and methods to enhance
research reproducibility

DIGITAL PUBLISHING
Articles designed
for optimal readership
across devices

FOLLOW US
@frontiersin

IMPACT METRICS
Advanced article metrics
track visibility across
digital media

EXTENSIVE PROMOTION
Marketing
and promotion
of impactful research

LOOP RESEARCH NETWORK
Our network
increases your
article's readership