THESIS FOR THE DEGREE OF DOCTOR OF PHILOSOPHY

For a Fistful of Qubits: Computational Quantum Chemistry on Near-Term Quantum Computers

Mårten Skogh

Department of Chemistry and Chemical Engineering
Chalmers University of Technology
Gothenburg, Sweden 2023

For a Fistful of Qubits:
Computational Quantum Chemistry on Near-Term Quantum Computers
Mårten Skogh
ISBN 978-91-7905-951-4

Doktorsavhandlingar vid Chalmers tekniska högskola.
Ny serie nr 5417
ISSN 0346-718X

Department of Chemistry and Chemical Engineering
Chalmers University of Technology
SE-412 96 Gothenburg, Sweden
Telephone +46 (0)31 772 1000
www.chalmers.se

Cover: *Roll of the qubits*, illustrating the stochastic nature of quantum computation, and the intersection of quantum computing and computational chemistry.

Printed by Chalmers digitaltryck
Gothenburg, Sweden 2023

For a Fistful of Qubits:
Computational Quantum Chemistry on Near-Term Quantum Computers

MÅRTEN SKOGH

Department of Chemistry and Chemical Engineering
Chalmers University of Technology

Abstract

Quantum computing has been touted as a great new frontier of computation, pushing the limits of what we consider within reach of computation. While not all problems are expected to be efficiently solved by a quantum computer, quantum chemistry is among those where many have speculated on near-term quantum advantage. A leading approach to near-term advantage comes in the form of variational quantum algorithms. In particular, variations of the Variational Quantum Eigensolver (VQE) algorithm form popular choices for chemistry on noisy quantum hardware.

This thesis dives into the topic of quantum computing in the noisy-intermediate scale era using variational quantum algorithms, the VQE in particular. Leveraging both classical simulations as well existing quantum computers, challenges of near-term quantum computing are explored. A parameter transfer approach is tested, aimed at helping speedup optimization variational parameters; an error mitigation strategy requiring close to no overhead is developed to reduce errors; and to help gauge the quality of quantum calculations beyond the point of quantum advantage, the topology of the electron density is analyzed. In addition, the application of near-term quantum computers to non-Born–Oppenheimer problems is explored, both for static and dynamic cases. The extension to the non-Born–Oppenheimer opens for new qubit reduction schemes which are analyzed.

Exploration of the limits of near-term quantum hardware and algorithms forms a common thread among the topics investigated. While quantum advantage still remains out of grasp for current generations of quantum computers, hope for near-term advantage remains. By pushing the boundaries, useful quantum computing might come one step closer.

Keywords: computational chemistry, quantum computation, variational quantum algorithms, quantum error mitigation, nonadiabatic processes

"I wish to God these calculations had been executed by steam."

— Charles Babbage

List of Appended Papers

This thesis is based on the appended papers listed below:

I **Accelerating variational quantum eigensolver convergence using parameter transfer**
Mårten Skogh, Oskar Leinonen, Phalgun Lolur, Martin Rahm
Electronic Structure, 2023, 5 (3)

II **Reference-State Error Mitigation: A Strategy for High Accuracy Quantum Computation of Chemistry**
Phalgun Lolur, Mårten Skogh, Werner Dobrautz, Christopher Warren, Janka Biznárová, Amr Osman, Giovanna Tancredi, Göran Wendin, Jonas Bylander, Martin Rahm
Journal of Chemical Theory and Computation, 2023, 19 (3), 783-789

III **The Electron Density: A Fidelity Witness for Quantum Computation**
Mårten Skogh, Phalgun Lolur, Werner Dobrautz, Christopher Warren, Janka Biznárová, Amr Osman, Giovanna Tancredi, Göran Wendin, Jonas Bylander, Martin Rahm
Submitted to *Chemical Science*

IV **A quantum computing implementation of nuclearelectronic orbital (NEO) theory: Toward an exact pre-Born–Oppenheimer formulation of molecular quantum systems**
Arseny Kovyrshin, Mårten Skogh, Anders Broo, Stefano Mensa, Emre Sahin, Jason Crain, Ivano Tavernelli
The Journal of Chemical Physics, 2023, 158 (21), 214119

V **Nonadiabatic Nuclear–Electron Dynamics: A Quantum Computing Approach**
Arseny Kovyrshin, Mårten Skogh, Lars Tornberg, Anders Broo, Stefano Mensa, Emre Sahin, Benjamin C. B. Symons, Jason Crain, Ivano Tavernelli
The Journal of Physical Chemistry Letters, 2023, 14 (31), 7065-7072

Contribution Report

The author's contributions to the appended papers are summarized below:

I Designed quantum circuits together with P.L. and C.W. Analyzed the data together with P.L. and M.R. Developed the mitigation strategy together with P.L, M.R., and W.D. Wrote the paper together with my co-authors.

II Designed and planned the study together with M.R. and P.L. Carried out simulations together with O.L. Analyzed the results and wrote the paper with my co-authors.

III Performed simulations of quantum circuits. Analyzed the data together with P.L., W.D., and M.R. Wrote the manuscript together with my co-authors.

IV Performed quantum computational simulations together with A.K. Analyzed the data together with A.K., J.C., and I.T. Assisted in writing the paper together with my co-authors.

V Performed dynamics simulations together with A.K. and L.T. Analyzed results together with A.K., L.T., J.C., and I.T. Assisted in writing the paper together with my co-authors.

Publications Not Included in this Thesis

The following publications are not included in the thesis:

Benchmarking the variational quantum eigensolver through simulation of the ground state energy of prebiotic molecules on high-performance computers
Phalgun Lolur, Martin Rahm, Mårten Skogh, Laura García-Álvarez, Göran Wendin
AIP Conference Proceedings 2362 (1)

Toward Accurate Post-Born-Oppenheimer Molecular Simulations on Quantum Computers: An Adaptive Variational Eigensolver with Nuclear-Electronic Frozen Natural Orbitals
Anton Nykänen, Aaron Miller, Walter Talarico, Stefan Knecht, Arseny Kovyrshin, Mårten Skogh, Lars Tornberg, Anders Broo, Stefano Mensa, Benjamin C. B. Symons, Emre Sahin, Jason Crain, Ivano Tavernelli, Fabijan Pavoševi
arXiv preprint

Acronyms

1-RDM one-particle Reduced Density Matrix

ADAPT-VQE Adaptive Derivative-Assembled Pseudo-Trotter ansatz Variational Quantum Eigensolver

AO Atomic Orbital

BCP Bond Critical Point

BO Born–Oppenheimer

CASCI Complete Active Space Configurational Interaction

CCSD Coupled-Cluster including Single and Double excitations

CDR Clifford Data Regression

HF Hartree–Fock

JW Jordan–Wigner

LCAO Linear Combination of Atomic Orbitals

MO Molecular Orbital

MP2 second order Møller–Plesset perturbation theory

NEO Nuclear-Electronic Orbitals

NEO-UCC Nuclear-Electronic Orbitals Unitary Coupled Cluster

NEOCASCI Nuclear-Electronic Orbital Complete Active Space Configurational Interaction

NISQ Noisy Intermediate-Scale Quantum

NNA Non-Nuclear Attractor

PES Potential Energy Surface

QITE Quantum Imaginary Time Evolution

QTAIM Quantum Theory of Atoms in Molecules

REM Reference Error Mitigation

SPAM State Preparation And Measurement

SPSA Simultaneous Perturbation Stochastic Approximation

STO Slater-type orbital

STO-nG Slater-type orbital approximated as n Gaussian functions

UCC Unitary Coupled-Cluster

UCCSD Unitary Coupled-Cluster including Single and Double excitations

VarQITE Variational Quantum Imaginary Time Evolution

VarQTE Variational Quantum Time Evolution

VQE Variational Quantum Eigensolver

Acknowledgements

After four years and some change, this adventure comes to an end. Having walked a long road, paved by those before me, I have hopefully put down a brick of my own for others to walk further.

In contrast to the many calculations performed within it, the work that has gone into this thesis was not carried out in a vacuum, without any environmental interaction. On the contrary, this thesis stands as a testament to the power of collaboration. Many thanks are owed, some of which will be given below.

What is a Ph.D. student without his or her supervisor? I wish to thank Martin Rahm for his supervision, for our discussions, and for the many cray fish parties. On a similar note, I owe a dept of gratitude to my industrial supervisor, Anders Broo, for giving me this opportunity, for making me feel part of the the group, and the for many insightful discussions. I am also grateful to my examiner, Göran Johansson, who recommended me to pursue this Ph.D.

To all my colleagues and collaborators, at Chalmers, AstraZeneca and elsewhere, I want to express my gratitude for all the discussions, insights, laughs, and good times. To all my friends outside of research: thanks for putting up with cancelled plans and late arrivals due to crashing code and manuscript deadlines.

I would like to especially thank: Phalgun and Werner, whom I have worked side by side with at Chalmers; Oskar and Felix, whom I had the pleasure of supervising during their master's theses; Marco, Hilda, Ruslan, and Christoph who gave invaluable feedback on this thesis; Lars, for all the pi(e), and for patiently explaining things to me when I found myself clueless; and to Arseny, whose herculean efforts in our collaborations cannot be understated, and who has encouraged and inspired me throughout our work together. I cannot thank you all enough.

I am thankful to my parents who have always given me all the support I could ever wish for, who raised me to become who I am today. Tack, morsan och farsan.

Hanna, with you the bad parts have not been as bad and the good ones have become even better. I am not sure how or why you have put up with me writing this thesis, but I am forever grateful.

Mårten Skogh, Göteborg, October 2023

Contents

CHAPTER 1

Introduction

Chemistry is fundamental to all life on Earth, and humans are no exception. On the contrary, we might be even more dependent on chemistry than any other species that has ever called this planet home. "Better living through chemistry" might as well have been the slogan for humanity[1]. Whether it be the proteins in our bodies or the medicines we design to interact with these, the combustion of fuels for energy or the redox reactions that power our smartphones; chemistry forms a crucial foundation for our modern way of life.

The way by which useful chemicals were found was long dominated by chance discoveries. Someone mixed two or more things (usually involving substantial heating), the result showed interesting properties and was passed onto other proto-chemists [2]. While it to this day remains important to mix elements and molecules together to see what comes out, we have evolved our selection processes for these substances through a thorough understanding of the underlying principles guiding the formation of new substances. While this understanding can help the experimental chemist to mix substances based on so-called *chemical intuition* [3], it often requires fairly complex experiments that take time and resources. Modern computers offer chemists a powerful complement to experiments that can be both faster and cheaper [4]. Calculations can also offer new insight into the quantum world of chemical reactions by extracting properties that are challenging to obtain in the lab.

By turning the theoretical work of many generations of chemists into computer code, the possibilities for new and exciting predictions have opened up. As we strive for molecules and materials with ever better tailored properties, the demand for more computation follows. This hunger for ever more powerful calculations has largely been satiated by the simultaneous increase in computational power, eloquently predicted by Gordon Moore to roughly double every two years; the now famous *Moore's law* [5]. While Moore's law has more or less held true for well over half a century, its end has time and time again been predicted to be just around the corner. Even though these predictions have so far not turned out true, there are physical limits to the exponential growth of "cramming"[2] transistors onto silicon. Sooner or later, the smallest components will reach the size of atoms and the current methods for controlling the 1's and 0's will no longer work [6–8]. At this scale, quantum effects dominate

[1] "Better Things for Better Living... Through Chemistry" was a slogan adopted by the company Du Pont in 1935 [1]. It is often quoted as simply "Better living through chemistry".

[2] *Cramming* is the actual word used by Gordon Moore in 1965 in his article *"Cramming more components onto integrated circuits"*.

the properties of matter, leading to quantum effects that are detrimental to the type of computation we today rely on. Thus, the quantum realm puts a limit on the theoretical power of our current model of computation. Perhaps paradoxically, the same quantum effects might act to allow a new type of computer, vastly more powerful for certain calculations.

The initial ideas of using quantum effects in computing were put forward in the early 1980's [9, 10][3]. Since then, the development of algorithms and hardware have both taken off. Already at the time of writing, a plethora of quantum devices and software are available for researchers, engineers, and hobbyists alike to get their hands wet in this burgeoning field. Nonetheless, we have yet to see a *useful* quantum computation take place on a quantum computer, although terms such as *quantum supremacy* and *quantum advantage* have been used from time to time [11, 12].

This thesis should be seen as a child of its time, a time of much hype and speculation, of quantum apostles and classical skeptics. It has been a work originating in the space between academia and industry, in the overlap of physics, chemistry, and computer science. Working with the near-term applications in focus, trying to push the abilities of existing and soon-to-come hardware. To this end, an array of possible avenues have been pursued, and the structure of the thesis has been made to put the various topics explored into a larger context.

The chapters have been written to accommodate readers coming from different backgrounds, going into most of the theoretical background needed to understand the research presented later on. Chapter 2 gives a background to the fundamental quantum physics, chemistry, and computation common to all of the subsequent chapters and all appended papers. The following chapters dive into topics encompassing the research presented in papers II–III. Chapter 3 focuses on the problem of optimizing variational quantum algorithms, and how one can improve their convergence. Chapter 4 takes a look at the perhaps most challenging part of quantum computing: the noise. Chapter 5 raises the question how one can gauge the quality of a quantum simulation as we move towards a situation where a classical calculation can no longer be used for validation. Chapter 6 discusses the relaxation of the Born–Oppenheimer (BO) approximation, and the simulation of quantum dynamics on quantum computers. Chapter 7 provides a short summary of the appended papers. Finally, Chapter 8 offers concluding remarks and an outlook on the future of the field.

1.1 Classical, Quantum, and Conventional

Writing on this topic has proven an interesting exercise in phrasing oneself correctly. I often find myself writing phrases such as "classical quantum chemistry", which on a second read through baffles even the author. I will, therefore, lay forth some conventions below, which I hope will ease the reader's experience, and spare myself from further baffling.

Classical Used as to distinguish from the idea of "modern" physics and chemistry, i.e., theories based on quantum or relativity theory. Classical will also distinguish between the ideas of classical and quantum computing.

[3]However, the real breakthrough of quantum computing came with the discovery of factorization on quantum computers, discovered by Peter Shor in 1994, referred to as *Shor's algorithm*. This breakthrough threatened to break the RSA encryption scheme, which is heavily dependent on the use of large prime factors and the cost of factorization.

Quantum A word that will be used to both refer to the idea of quantum *chemistry* as well as quantum *computing*.

Conventional Using a phrase such as "classical quantum chemistry" has proven confusing to many a reader. To mitigate some confusion, the word *conventional* chemistry calculations will be used to refer to classical computing applied to quantum chemistry problems.

1.2 Calculation, simulation, and emulation

An additional duo of words with similar meaning are formed by *calculation*, *simulation*, and *emulation*. In this thesis, a calculation will refer to the execution of an algorithm to yield a resulting value. Simulation, on the other hand, is the act of mimicking a physical system on a machine, quantum or classical. The term emulation will be used to refer explicitly to the use of a classical computer to simulate (emulate) a quantum computer.

Calculation The execution of an algorithm in order to obtain a specific value.

Simulation The explicit attempt to imitate the behavior of a physical system using a computer of any sort.

Emulation Use of a classical computer to simulate a quantum computer.

CHAPTER 2

Quantum, Computing, and Chemistry

Quantum computing is often seen as heralding a new era in computational chemistry, allowing for solutions to classically intractable problems [13]. While this is true in the fault-tolerant regime [14], the current generation of Noisy Intermediate-Scale Quantum (NISQ) computers do not have the same theoretical underpinning [15]. Contemporary quantum computation has more in common with the advent of classical computing than the smartphones and high-performance computational clusters we today take for granted.

The century or so that classical computers have been available has given us an exceptional set of tools and tricks to help deal with the otherwise extremely hard problem of simulating quantum systems. While these developments have been of utmost importance to chemistry, and science in general, they also offer stiff competition to any up-and-coming computational method wishing to prove itself. Exponential speedup is theoretically guaranteed for certain quantum algorithms and problems [16, 17], see for example Shor's algorithm for integer factorization [18, 19]. However, the cases where this speedup exists often come with caveats. For example, the quantum phase estimation algorithm for extracting ground state energies requires an initial overlap with the ground state solution that might not be trivial to achieve [14].

This chapter will go through most of the underlying theory, laying the foundation for later chapters.

2.1 Quantum Mechanics and Quantum Chemistry

Fundamental to modern theoretical chemistry, quantum mechanics has been hailed as the most accurate theory of all time [20]. This accuracy also applies to the properties and behaviors of atoms and molecules. Although the equations needed to model these behaviors often are deceivingly concise, any numerical evaluation is, for all but the smallest systems, infeasible. The inability to accurately simulate quantum mechanical systems was well-known early on in the development of quantum theory. Paul Dirac famously stated [21]

> “The underlying physical laws necessary for the mathematical theory of a large part of physics and the whole of chemistry are thus completely known, and the difficulty is only that the exact application of these laws leads to equations much too complicated to be solvable. It, therefore, becomes desirable that approximate practical methods of applying quantum mechanics should be developed, which can lead to an explanation of the main features of complex atomic systems without too much computation.” „

As mentioned in the previous chapter, these initial problems have been offset to some degree by the introduction of computers. Of equal, or perhaps even greater, importance have been many of the theoretical developments and approximations implemented to reduce the computational load.

2.1.1 Dirac Notation and Second Quantization

As we add more particles to our calculations, we soon start to see that the work required to solve the equations becomes increasingly more demanding. We call the resulting problem a *many-body* (or n-body) problem, and these, except for a select number of cases, do not have a closed-form solution.

One insight that greatly aids in the treatment of many quantum mechanical problems is the concept of *indistinguishability*. indistinguishability tells that quantum particles with equal quantum numbers are fundamentally impossible to distinguish. That is to say, all electrons with the same spin are indistinguishable, the same goes for protons and neutrons, or any elementary particle. If all particles are the same, then there is no need to keep track of individual particles. Instead, we can look at *occupation* of quantum states. In fact, *occupation number representation* is a phrase commonly used to refer to this view of quantum mechanics, another common name is *second quantization*. In this thesis and in all the appended papers, the second quantization approach is utilized.

Fundamental to the occupation representation is the idea of a basis set, a set of functions that can be occupied. The basis set is a set of functions, $\{\phi_1, \dots, \phi_l, \dots, \phi_N\}$, that forms a basis for the Hilbert space where the wave function is defined. We will make use of braket notation to describe our quantum states, where the state vector is given by the *ket*, $|\cdot\rangle$, defined as a sum of weighted basis states

$$\Psi(\boldsymbol{r}) \rightarrow |\Psi\rangle = \sum_k c_k \, |n_1, \dots, n_l, \dots, n_N\rangle \tag{2.1}$$

where c_k is the weight of each basis state. The probability interpretation of the wave function requires that

$$\sum_k |c_k|^2 = 1. \tag{2.2}$$

The corresponding *bra*, $\langle \cdot |$, defines a covector to $|\cdot\rangle$ as

$$\langle\Psi| = |\Psi\rangle^\dagger . \tag{2.3}$$

The ket, $|n_1, \dots, n_l, \dots, n_N\rangle$, collects information of the occupation of the basis functions ϕ_l as n_l. Since we will be primarily concerned with the treatment of fermions, and fermions

must fulfill the exclusion principle, the occupations must be either 0 or 1, i.e., $n_l \in \{0, 1\}$. This binary representation makes for a convenient shorthand notation (which also applies to qubits)

$$|n_1, \ldots, n_l, \ldots, n_N\rangle = |m\rangle\,, \qquad m = \sum_l n_l 2^{(l-1)}. \tag{2.4}$$

Another important concept in second quantization comes in the form of ladder operators, or creation and annihilation operators, $a_l^\dagger$ and a_l. These operators act on a state and the occupation, n_l, according to

$$a_l^\dagger\,|0_l\rangle = |1_l\rangle\,, \quad a_l^\dagger\,|1_l\rangle = 0 \tag{2.5}$$

$$a_l\,|1_l\rangle = |1_l\rangle\,, \quad a_l\,|0_l\rangle = 0 \tag{2.6}$$

Two terms that will be used from time to time are Hilbert and Fock space. The Hilbert space is a linear (vector) space with a norm defined by an inner product [22]. The Fock space, in turn, is a Hilbert space spanned by the occupation basis, $\{|k\rangle\}$ [23]. A common term used in chemistry is active space, which seems to, more or less, correspond to the Fock space.

2.1.2 Chemical Basis Sets and Orbitals

As mentioned, in the second quantized picture we are interested in describing our atoms and molecules in terms of basis functions, ϕ_l, that approximate the true molecular wave function. The conventional way of constructing a molecular basis is to start with a set of Atomic Orbitals (AOs) which are then linearly combined to form Molecular Orbitals (MOs). Both the AOs and MOs form a basis for the same Hilbert space.

According to Kato's cusp condition, the electron density must have *cusps*[1] at the location of the atomic nuclei [24], a feature that also has to be present in the underlying orbitals. A type of orbital that satisfies Kato's condition is the Slater-type orbital (STO) [25]. However, the overlap of STOs is computationally costly to calculate accurately. It is worth noting that there is no expected quantum advantage for these integral evaluations. The radial part of the STO is given by

$$R(r) = C\, r^{n-1} e^{-\zeta r} \tag{2.7}$$

where n corresponds to the natural quantum number, C is a normalization constant, and ζ relates to the nuclear charge.

A common compromise is to approximate the expensive STOs with the help of a number of Gaussian functions centered on the atomic nuclei, STOs. These approximated STOs are referred to as Slater-type orbital approximated as n Gaussian functions (STO-nG), where n is the number of Gaussian functions used in the approximation. The STO-nG is an example of a minimal basis set, where minimal refers to the use of a single basis function to describe each orbital occupied in the atom [23]. For modern conventional calculations, minimal basis sets are no longer seen as valid for most types of calculations, and a basis such as a correlation-consistent basis set (cc-pVnZ) [26] is more commonly used [13]. These larger basis sets introduce more basis functions to describe the AOs, allowing for more flexibility, but also increasing the size of the corresponding Hilbert space.

As mentioned, a common approach to go from AOs to MOs is the Linear Combination of Atomic Orbitals (LCAO) which are combined to form MOs [27]. The MOs form the

[1] A cusp is a point on a curve at which, if you move along the curve, you would have to abruptly reverse direction.

basis set we then use to form our state $|\Psi\rangle$. Finding the linear combinations is often done using the Hartree–Fock (HF) method, which optimizes the orbitals to give a single Slater determinant minimal solution [23].

In many sizeable chemistry calculations, a substantial number of the MOs either become close to fully occupied or completely unoccupied. To reduce computational load, one can limit the number of orbitals under consideration, the *active space*. A common assumption is to fill the core MOs and *freeze* these, and similarly remove the MOs with the highest HF energy. Active spaces are often represented as a tuple (e, o). Here, e is the number of electrons and o is the number of spatial orbitals in the active space.

While a minimal basis set is today far from being considered accurate, it is widely used in quantum computing, owing to the low number of qubits required to represent it. Some argue that minimal basis sets should not be used in quantum computing as they have long been superseded by more accurate basis sets. On the other hand, one could argue that with the limited number of qubits available, a minimal basis set is more suited to the problems being considered. Using a large basis will, in many cases, require reducing the active space substantially.

In addition to the basis set approaches described above, other methods exist to create an approximate basis for the wave function, such as the basis set-free approach of Kottmann et al. [28].

2.1.3 The Molecular Hamiltonian

To accurately describe chemistry, a quantum mechanical approach is necessary. What ultimately dictates the behavior of atoms and molecules is the molecular Hamiltonian, $\hat{H}_{mol}$. The molecular Hamiltonian describes the kinetic and potential energy of all interacting particles (nuclei and electrons) in the molecule, where one considers one- and two-body terms

$$\hat{H}_{mol} = \hat{H}_A + \hat{H}_{AB} = \sum_{pq} h_{pq}\, a_p^\dagger a_q + \sum_{ijkl} h_{pqrs}\, a_p^\dagger a_q^\dagger a_r a_s \tag{2.8}$$

Often the molecular Hamiltonian is further divided into nuclear and electronic terms, as well as all cross terms resulting from their interaction

$$\hat{H}_{mol} = \hat{H}_N + \hat{H}_{NN} + \hat{H}_{Ne} + \hat{H}_{ee} + \hat{H}_e. \tag{2.9}$$

Here, $\hat{H}_N$ and $\hat{H}_e$ correspond to nuclear and electron kinetic energies, and $\hat{H}_{NN}$, $\hat{H}_{Ne}$, $\hat{H}_{ee}$ to nuclear-nuclear, nuclear-electron, and electron-electron interactions, respectively.

The coefficient h_{pq} and h_{pqrs} in Eq. (2.8) are integrals on the form

$$h_{pq} = -\int \phi_p^*(\boldsymbol{r}) \frac{1}{2} \nabla^2 \phi_q(\boldsymbol{r})\, \mathrm{d}\boldsymbol{r} \tag{2.10}$$

$$h_{pqrs} = \int \phi_p^*(\boldsymbol{r}) \phi_q^*(\boldsymbol{r}) \frac{Z_A Z_B}{|\boldsymbol{r_A} - \boldsymbol{r_B}|} \phi_r(\boldsymbol{r}) \phi_s(\boldsymbol{r})\, \mathrm{d}\boldsymbol{r}, \tag{2.11}$$

where $\phi_\gamma, \gamma \in \{p, q, r, s\}$ are spin orbitals.

A common approach to simplify and reduce the computational cost of molecular calculations is to impose the so-called Born–Oppenheimer (BO) approximation, named after physicists Max Born and J. Robert Oppenheimer. The BO approximation, in essence, results from the separation of time scales that the nuclei and electrons move at, resulting from

the large difference in mass. Consequently, the electrons are able to almost instantaneously adapt to any movement of the nuclei. As will be discussed in Chapter 6 there are cases where this approximation breaks down. By applying the BO approximation, we reduce our molecular problem to that of the *electronic structure problem*, which only concerns itself with the task of optimally constructing the electron wave function, interacting with an effective potential of the nuclei. Eq. 2.9 then simplifies to

$$\hat{H}_{mol} = \hat{H}'_e + \hat{H}_{ee} + V_{NN}. \tag{2.12}$$

where $\hat{H}'_e$ combines the kinetic energy H_e and the nuclear potential interaction with the electrons. The nuclear-nuclear interaction is demoted from an operator to a classical potential V_{NN}, losing its "hat" ($\hat{}$) in the process. The BO approximation will be discussed in more detail in Chapter 6, where the effects of relaxing the BO approximation are also discussed.

2.2 Quantum Computation

Quantum computation forms an extension of the classical theory of computation, opening up new possibilities, and bringing some seemingly unsolvable problems closer to a solution. While the quantum computer can, in theory, do what a regular computer can, the idea that quantum computers will replace its classical counterparts in the near term is far from realistic. Nevertheless, the idea of applying quantum computers to quantum mechanical problems, such as the electronic structure problem in Section 2.1.3, is still a leading candidate for useful quantum computation [13, 29].

2.2.1 Complexity in Computation

As the motivation behind quantum computation primarily comes from the promise of solving certain problems faster than would be possible classically. What is fast and what is slow is often determined through the evaluation of algorithmic complexity. Complexity in the context of algorithms usually refers to one of two things: *time complexity* or *space complexity*. Time complexity concerns itself with the study of the scaling behavior of algorithms as the problem size grows, while space complexity similarly describes the amount of memory required for different problem sizes.

The time complexity of an algorithm, $T(n)$, is often hard to determine exactly. Instead one opts to find an upper limit to the time complexity for large problem ($n \to \infty$) instances, $\mathcal{O}(g(n))$, where the function $g(n)$ provides an asymptotic upper bound, i.e. $\lim_{n\to\infty} T(n) \leq g(n)$, and the lower complexity limit, $\Omega(h(n))$, such that $\lim_{n\to\infty} T(n) \geq h(n)$[2]. These are usually referred to as the "worst" and "best" case scaling, respectively. Often, most of the focus is given to the upper bound, $\mathcal{O}(n)$ as the worst case is the limiting factor [16, 30].

Quantum advantage, or quantum supremacy, is defined as problems where a quantum computer could solve a previously exponential problem in polynomial time. That is, taking

[2]This is a bit hand-wavy, proper definitions are
$\mathcal{O}(g(n)) \implies \limsup_{n\to\infty} \frac{|T(n)|}{g(n)} < \infty$
$\Omega(h(n)) \implies \liminf_{n\to\infty} \frac{|T(n)|}{g(n)} > 0$

a problem from $\mathcal{O}(e^n)$ on a classical computer, to $\mathcal{O}(n^c)$ on a quantum machine. Other advantages exist as well, such as a quadratic speedup for Grover's algorithm, $\mathcal{O}(n) \rightarrow \mathcal{O}(\sqrt{n})$ [31].

In addition to the worst-case analysis of the algorithm, the speed of the underlying hardware also plays and important role when trying to predict a cross-over point, i.e., a point where a quantum computer will outperform a classical computer. Often the term quantum resources is used to refer to the qubit and gate requirements needed to implement certain algorithms, forming an additional measure of complexity.

2.2.2 Quantum Information

Quantum computing encompasses and extends upon classical computing, meaning that any computation that can be performed on a classical computer can also be implemented on a quantum computer [16]. As a consequence, most of the classical information theory developed since the 1920s by the likes of Alan Turing [32] and Claude Shannon [33], also applies to quantum computers. For this reason, the field of quantum computation forms an overlap of computer science and quantum physics, with terminology and concepts deriving from both.

The fundamental unit of information for quantum computing is the *qubit* (quantum binary digit, or quantum bit). The qubit, like its classical counterpart, has two orthogonal states $|0\rangle$ and $|1\rangle$. However, unlike its classical counterpart, all combinations

$$|\Psi\rangle = a\,|0\rangle + b\,|1\rangle \tag{2.13}$$

are also valid qubit values, under the one restriction that

$$|a|^2 + |b|^2 = 1. \tag{2.14}$$

Note that a and b are allowed to be complex values.

The state of a single qubit can be visualized using the *Bloch sphere*, named after physicist Felix Bloch [34]. To see how this visualization works, we will re-express equation 2.13 in a polar form. Since a and b are complex numbers, we can rewrite these as $a = \sin(\theta/2), b = \cos(\theta/2)e^{i\phi}$ which gives

$$|\Psi\rangle = \sin\frac{\theta}{2}\,|0\rangle + \cos\frac{\theta}{2}e^{i\phi}\,|1\rangle\,. \tag{2.15}$$

The polar form of Eq. 2.15 naturally maps to a sphere of radius 1, see Figure 2.1. A measurement of a state on the Bloch sphere corresponds to projecting the state onto a line passing through the origin. One often assumes that this line is the z-axis, commonly referred to as the computational basis.

Assuming that one measures in the computational basis, the act of measuring will always yield the result $|0\rangle$ *or* $|1\rangle$, i.e., a single measurement collapses the state and yields one bit of information.

While the Bloch sphere and the state vectors are commonly seen in quantum computation, these representations primarily deal with pure states. As we introduce noise and interactions with the environment, we need another complement to the pure state representation. For so-called mixed states, an alternative representation comes in the form of density matrices, ρ, defined from an ensemble of states $|\psi_l\rangle$ where each state is associated with a probability p_l [16, 22]

$$\rho = \sum_l p_l\,|\psi_l\rangle\,\langle\psi_l|\,. \tag{2.16}$$

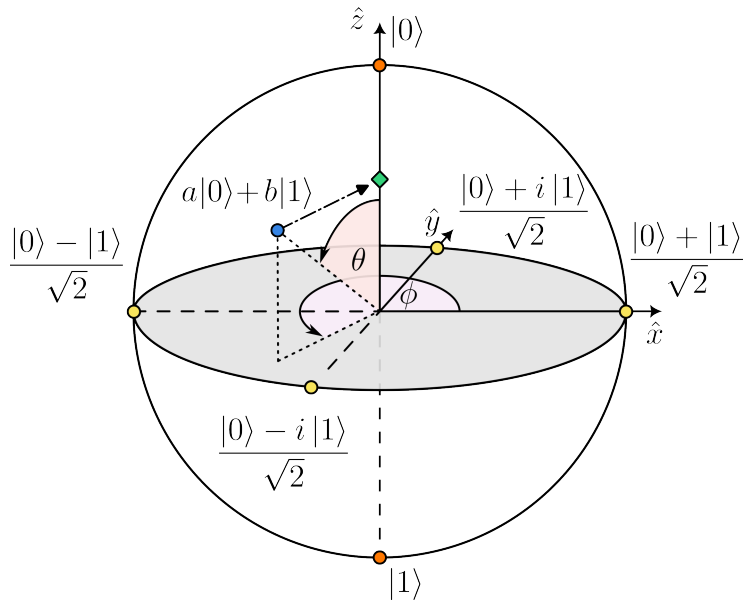


Figure 2.1: The Bloch sphere. The two states on the poles of the Bloch sphere form the computational basis, $|0\rangle$ and $|1\rangle$. Points on the equator (yellow) correspond to equal superpositions, $|a| = |b| = \frac{1}{\sqrt{2}}$. Single measurements of the qubit project the qubit to one of the measurement eigenstates. Several averaged measurements of a state $a\,|0\rangle + |1\rangle$ (blue dot) corresponds to a projection onto the measurement axis (green diamond), the $\hat{z}$-axis in this case.

The density matrix representation can also be used to represent pure states. In the case of pure states, the density matrix is can be defined as

$$\rho_{|\Psi\rangle} = |\Psi\rangle \langle\Psi| . \tag{2.17}$$

2.2.3 What Quantum Computing Can and Cannot Do

Quantum computing is often described to the public as a *panacea* to all computational ailments. Everything from helping to solve climate change [35], to faster gaming [36], financial trading [37] and faster than light communication [38] are often cited as lying within the grasp of future quantum computers. While some of the things being touted are true, many things are not. The impact of quantum computing will most likely not be as far-reaching as that of the classical computer, simply due to the remarkable efficiency of the classical computer. There are only a few things that a classical computer cannot do that a quantum computer might be able to do better. However, those things happen to be of great importance.

2.3 Quantum Algorithms for Quantum Chemistry

The new quantum revolution offers two major avenues for simulating quantum mechanical systems. The first is to construct another quantum mechanical system that behaves like the system we are targeting, but for some reason, is easier to control, manipulate, or study. Using an approach like this is referred to as *quantum simulation* [39]. The second option is to

perform calculations using a quantum computer. While this thesis is primarily concerned with the latter, one can argue that the current generation of quantum computers lands somewhere between these two; not being genuine simulators, while simultaneously falling short of being true universal computational machines.

The descriptions herein will be limited to gate-based quantum computing and make heavy use of the circuit model for quantum computing [40, 41]. While the gate-based circuit model is the most common abstraction, there are other types of quantum computational models that can be utilized, such as measurement based [42], continuous variable [43], and adiabatic quantum computing [44] to mention a few.

2.3.1 The Variational Quantum Eigensolver

The Variational Quantum Eigensolver (VQE) is a so-called variational quantum algorithm, designed to solve quantum mechanical eigenvalue problems on NISQ hardware [45]. By leveraging a parameterized ansatz the VQE allows for finding minimum energies, even in noise energy landscapes. Often referred to as a *hybrid* quantum algorithm, the VQE divides its computational workload between both the classical and quantum computer. The VQE algorithm is outlined in Figure 2.2. For all the appended papers, the VQE has been the variational algorithm of choice for identifying ground state solutions.

In a traditional VQE calculation, the algorithm as a whole can be viewed as trying to transform an initial guess $|\psi_{init}\rangle$ into the problem ground state $|\psi_{GS}\rangle$ such that properties of the ground state can be evaluated. To our help we have a parameterized unitary ansatz $U(\boldsymbol{\theta})$ dependent on the variational parameters $\boldsymbol{\theta}$. In this light, the VQE algorithm can be seen as the classical minimization problem

$$\min_{\boldsymbol{\theta}} \left\langle \psi_{init} \left| \hat{U}^{\dagger}(\boldsymbol{\theta}) \hat{H} \hat{U}(\boldsymbol{\theta}) \right| \psi_{init} \right\rangle = \min_{\boldsymbol{\theta}} \left\langle \psi(\boldsymbol{\theta}) \left| \hat{H} \right| \psi(\boldsymbol{\theta}) \right\rangle = \min_{\boldsymbol{\theta}} \langle E(\boldsymbol{\theta}) \rangle \tag{2.18}$$

where the quantum computer is a black box that, based on the parameters it is fed, evaluates the energies of the trial states. While the VQE is a specific instance of a variational algorithm, most variational algorithms utilize the optimization loop of Figure 2.2.

Looking at the time complexity of VQE, it is often dependent on the ansatz chosen and measurement scheme used. For example, the Unitary Coupled-Cluster including Single and Double excitations (UCCSD) ansatz will scale as $\mathcal{O}(n^2N^4)$ with the number of orbitals, n, and electrons, N, [46] much like conventional Coupled-Cluster including Single and Double excitations (CCSD) [47]. Where the VQE does offer an advantage is in its ability to represent the exponential state of an arbitrary wavefunction, something that is prohibitively expensive on a classical machine.

2.3.2 Problem Encoding

The first part of any calculation on a quantum computer is to decide how to encode the problem onto qubits. For the electronic structure problem, we seek to use qubits to represent spin orbital occupations. A common mapping (also called mapping) the Jordan–Wigner (JW) fermionic encoding [48, 49]. This is a straightforward mapping whereby each qubit q_i represents one basis function f_i in the basis set. In the case of one function per orbital, each

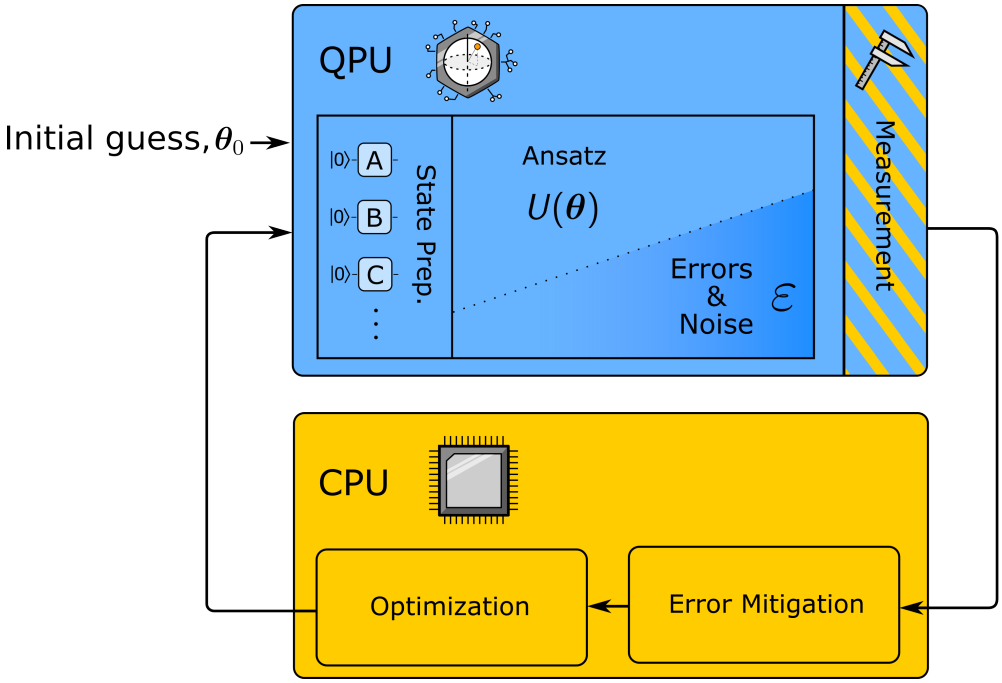


Figure 2.2: A overview of the VQE algorithm. A Quantum Processing Unit (QPU) and a Classical Processing Unit (CPU) work in tandem to minimize the measured energy. As the state preparation and ansatz are applied, quantum noise and errors affect the prepared state. Measurement yields the energy which can subsequently be improved by error mitigation. A classical optimizer is then used to update the ansatz parameters $\boldsymbol{\theta}$. The figure is adapted from Paper I.

qubit represents the occupation of an orbital. In short the JW mapping looks like

$$\underbrace{|f_1, f_2, \dots, f_n\rangle}_{\text{Fermionic state}} \xrightarrow{\text{JW}} \underbrace{|j_1, j_2, \dots, j_n\rangle}_{\text{Qubit state}}. \tag{2.19}$$

The next step in the JW mapping is to map also the quantum mechanical operators. In most cases, it is enough to map the fermionic creation and annihilation operators $a^\dagger$ and a.

The JW transformation of $a_r^\dagger$ and a_r are defined as

$$a_r^\dagger = (X_r + iY_r)Z_r^{\rightarrow} \tag{2.20}$$
$$a_r = (X_r - iY_r)Z_r^{\rightarrow} \tag{2.21}$$

where X, Y and Z are the Pauli matrices (also called the Pauli gates in quantum computing)

$$X = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \quad Y = \begin{bmatrix} 0 & i \\ -i & 0 \end{bmatrix}, \quad Z = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}. \tag{2.22}$$

The $Z_r^{\rightarrow}$ operator makes sure that the resulting encoding is anti-symmetric with respect to the exchange of any two states. It applies the Pauli Z gate to all qubits with an index less than r,

$$Z_r^{\rightarrow} = \prod_{i=1}^{r} Z_i. \tag{2.23}$$

The JW encoding offers a straightforward mapping. However, the $Z_r^{\rightarrow}$ operation in Eq. (2.23) maps a fermionic operator acting on a single spin orbital i, to a qubit operator on several qubits. In the worst case, a single $a^\dagger$ or a operator maps to an operation on all qubits.

Other encodings are available for fermionic problems, such as the parity and the Bravyi–Kitaev encodings [50, 51]. In particular, the parity encoding has found use for electronic structure problems [52]

$$\underbrace{|\overbrace{f_{1\uparrow},\dots,f_{n_\alpha\uparrow}}^{\text{Spin-up}},\overbrace{f_{1\downarrow}\dots,f_{n_\beta\downarrow}}^{\text{Spin-down}}\rangle}_{\text{Fermionic state}} \xrightarrow{\text{parity}} \underbrace{\left|p_1,p_2,\dots,p_{n_\alpha},\dots,p_{n_\alpha+n_\beta}\right\rangle}_{\text{Qubit state}}, \tag{2.24}$$

where so-called spin-block order has been used in the ordering of fermionic occupations, i.e., all spin-up orbitals are ordered in one continuous block, followed by a similarly ordered block of spin-down occupations.

The reason for the adoption of parity encoding primarily lies in the ease of qubit tapering, i.e., reduction of the number of qubits, corresponding to particle symmetries such as total particle number and total spin. The parity encoding, as the name suggests, encodes the parity p of the occupation, not the occupation itself. Specifically, each qubit q_i stores the parity p_i of all spin orbitals f_j for all $j \leq i$. We can write this as

$$p_i = \left(\sum_{j=0}^{i} f_j\right) \bmod 2. \tag{2.25}$$

To understand how the parity encoding helps to reduce the number of qubits, one can look at Eq. (2.24) and see that for a system with fixed number of electrons ($n_\alpha + n_\beta =$ constant) the total parity, i.e., $p_{n_\alpha+n_\beta}$ will never change since the sum of all occupations will be constant. Similarly, for a system with fixed spin ($n_\alpha =$ constant, $n_\beta =$ constant) the parity of the first spin blocks, p_{n_α}, will remain unchanged.

An alternative parity mapping can also be made where the spin-blocks are encoded independently as

$$p_{i,\eta} = \left(\sum_{j=0}^{i} f_{j,\eta}\right) \bmod 2. \tag{2.26}$$

where $\eta \in \uparrow, \downarrow$.

Since the two qubits that correspond to these conserved parities will maintain a constant value, they can be removed from consideration, a process called tapering [52]. The parity mapping paired with tapering is something we have made use of in most of the appended papers. In Paper IV, we extend the tapering to also include nuclei, specifically protons, which is further discussed in Chapter 6. Additional symmetries can also be identified through more sophisticated analysis [53]. Further discussion on qubit reduction schemes can be found in Chapter 6 and Paper IV.

2.3.3 Wave Function Ansatz

The *ansatz* (plural *ansätze*) in the quantum computational sense is a unitary operation, U, that takes an initial state $|\psi_0\rangle$ to a trial state $|\psi_t\rangle$ as

$$U|\psi_0\rangle = |\psi_t\rangle\,. \tag{2.27}$$

For NISQ algorithms, one often chooses a parameterized ansatz circuit, $U(\theta)$, such that the parameters can be optimized according to Eq. (2.18).

There are several ways of constructing an ansatz, but one usually divides the ansätze into *chemically-inspired* and hardware-efficient ansätze. The hardware-efficient ansätze generally are constructed in a layered manner, where each layer performs some combination of entanglement between all qubits, and parameterized rotations of the qubits. The arguably most common chemically-inspired ansatz is the Unitary Coupled-Cluster (UCC). The UCC ansatz is defined as the unitary operator

$$U_{\mathbf{UCC}}(\boldsymbol{\theta}) = e^{T-T^\dagger}. \tag{2.28}$$

The operator T is, in turn, defined as a sum of k-particle operators

$$T = \sum_k T_k \tag{2.29}$$

where

$$T_1 = \sum_{pq} \theta_q^p a_p^\dagger a_q \tag{2.30}$$

$$T_2 = \sum_{pqrs} \theta_{rs}^{pq} a_p^\dagger a_q^\dagger a_r a_s \tag{2.31}$$

$$\vdots$$

Compared to the conventional, non-unitary coupled-cluster operator, the UCC operator is theoretically the more accurate operator [54]. The conventional coupled-cluster operator does not preserve the norm of the wave function, and does require normalization, a step that is exponentially hard on a classical computer. Many additional developments have modified the UCC approach in order to reduce the number of gates required for implementation. Among these, the Adaptive Derivative-Assembled Pseudo-Trotter ansatz Variational Quantum Eigensolver (ADAPT-VQE) [55], and those approaches based on it, e.g, [56, 57], are perhaps the most promising for near-term applications.

Another ansatz with its roots in conventional computational methods is the class of Quantum Imaginary Time Evolution (QITE) ansätze, primarily the Variational Quantum Imaginary Time Evolution (VarQITE) which is well suited for NISQ devices. While imaginary time evolution has not been used to solve any ground state problem in this thesis, real time evolution is used in Paper V and discussed in Chapter 6.

Hardware-efficient ansätze come in many shapes and sizes. The focus is often high expressivity while maintaining a low circuit depth. Rather than translating a conventional method (such as coupled-cluster) to the quantum computer, hardware-efficient algorithms use repeated patterns of rotations and entanglement to construct an ansatz that can potentially generate a large number of states. This design draws parallels to the construction of neural networks, where many repeated layers are used to approximate an unknown function.

2.3.4 Suzuki–Trotter Decomposition

Even after the encoding, it is common to find that the operator one wishes to implement is not a native operation on a quantum computer. An example of this can be an operator

$$U = e^{A+B} \tag{2.32}$$

where A and B are both Pauli operators. Note that operators of the form e^A *are* native to many types of quantum hardware and can be implemented without approximation. As such, we would like to express our non-native operation as a series of native operations. For scalar-valued exponentiation, this is not a problem since $e^{a+b} = e^a e^b$. However, unless e^A and e^B commute, $e^{A+B} \neq e^A e^B$. One can, of course, ignore this error and approximate the sum in the exponent as the product $e^{A+B} \approx e^A e^B$ regardless. One can improve on this approximation by applying a small part $\frac{A}{n}$ and $\frac{B}{n}$ repeated n times. In fact, if one was to let $n \to \infty$, the result is exact

$$e^{A+B} = \lim_{n\to\infty} (e^{\frac{A}{n}} e^{\frac{B}{n}})^n. \tag{2.33}$$

This approximation is known as the first-order Suzuki–Trotter decomposition, also some times referred to as Trotterization.

Unless otherwise specified, the first-order Suzuki–Trotter decomposition with $n = 1$ will be used in this thesis. The Trotterization error that can be introduced is discussed in Chapter 6 and showcased in Paper V.

CHAPTER 3

VQE Optimization and Convergence

A major challenge for all variational quantum algorithms is that of large-scale, non-convex optimization [58]. The large optimization landscapes[1] resulting from variational quantum algorithms are often riddled by local minima [65], and are prone to a phenomena known as *barren plateaus* [62, 66–68] The two common variants of ansätze, chemically-inspired and hardware-efficient, both present benefits and challenges during optimization. While chemically inspired ansätze often produce a more convex optimization landscape, with lower risk of local minima, they require deep circuits; their hardware-efficient counterparts, in contrast, produce compact circuits that in turn can leave most optimizers trapped in sub-optimal solutions [69]. As quantum circuits will unavoidably need to grow in order to provide any computational advantage, the question of how to efficiently optimize them remains an important one.

In Papers I and IV, we implement two related warm-start techniques to overcome some of these challenges. In what follows, the challenge of optimization will be introduced and the learnings of Papers I and IV will be put into context.

3.1 Cost Landscapes

For a general optimization problem, a cost (or loss) function is what one seeks to minimize, $C(\boldsymbol{\theta})$[2]. Minimization of $C(\boldsymbol{\theta})$ implies finding the parameters $\boldsymbol{\theta} = [\theta_1, \dots, \theta_n]^T$ that minimize the value of $C(\boldsymbol{\theta})$. More concisely, this process can be expressed as

$$\min_{\boldsymbol{\theta}} C(\boldsymbol{\theta}) = C(\boldsymbol{\theta}^*) = C^* \tag{3.1}$$

where the asterisk is used to indicate the parameters that give the globally optimal cost, i.e., the exact solution $\boldsymbol{\theta}^*$ and the global minimum C^*. For many problems in physics and

[1] While optimization problems encountered in the wild can be extremely high-dimensional — the open-source large language model Llama from Meta has 65 *billion* parameters (6.5×10^{10}) [59], and GPT-4 from OpenAI is rumored to have over one *trillion* (10^{12}) parameters [60] — one often borrows terms from geography, (an inherently low-dimensional field), talking about landscapes [61], plateaus [62], and basins [63, 64] formed by the cost function.

[2] Whether one wants to minimize a cost or maximize some gain, both problems can be seen as equivalent by changing the sign of the cost function, $C(\boldsymbol{\theta}) \to -C(\boldsymbol{\theta})$

chemistry, the cost that one seeks to minimize is the energy, E. The goal is often finding the ground state energy, E_0. Expressing the ground state as a minimization problem rests on the variational principle [22],

$$\frac{\left\langle \psi(\boldsymbol{\theta}) \middle| \hat{H} \middle| \psi(\boldsymbol{\theta}) \right\rangle}{\langle \psi(\boldsymbol{\theta}) | \psi(\boldsymbol{\theta}) \rangle} = E(\boldsymbol{\theta}) \geq E_0. \tag{3.2}$$

This principle states that any trial wave function only ever gives an upper bound to the ground state. The trial wave function can never give an energy lower than the ground state energy. In other words, the lower the energy, the closer we get to the correct ground state energy.

Gradient descent (also called steepest descent) algorithms are common approaches to solving optimization problems like that of Eq. (3.2) [70, 71]. The basis for gradient descent methods is straightforward: follow the negative gradient. The steepest descent algorithm is based on iteration, generating new parameters $\boldsymbol{\theta_{i+1}}$ as

$$\boldsymbol{\theta}_{i+1} = \boldsymbol{\theta}_i - \alpha \nabla C(\boldsymbol{\theta}_i), \tag{3.3}$$

where $\alpha \geq 0$ is the step length. Most iterative methods require that an initial guess, $\boldsymbol{\theta}_0$ is provided at the start of optimization. As will be shown in the following section, even though gradient descent is a very powerful method, it struggles in non-convex cost landscapes.

3.1.1 Local Minima and Basins

While optimization of variational quantum algorithms can encompass many parameters, large-scale optimization is nothing new, and a plethora of methods and strategies are available to those in need [72–75]. Nevertheless, the stochastic nature of hybrid quantum algorithms has proven a challenge. The complex optimization landscape and its many local minima is in part to blame.

Local minima are a common feature of many real world optimization problems. A local minimum is the smallest solution in some neighborhood of the cost landscape, in contrast to the global minimum, which is the minimum for the entire landscape. The local minimum is often surrounded by a basin. The basin is a region of the cost landscape wherein, if one follows the negative gradient, one always ends up at the corresponding local minimum. These basins thus trap many gradient-based methods.

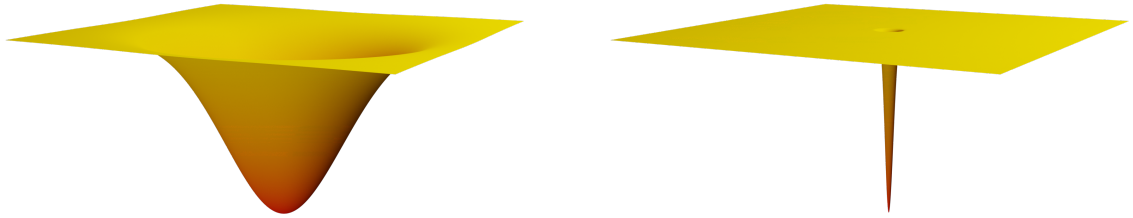
An optimization landscape with many local minima is thus challenging to traverse as conventional gradient descent methods easily get stuck in the local minima. In fact, if the provided initial guess, $\boldsymbol{\theta}_0$, does not lie within the basin of the global minimum, the optimal solution cannot be found. The importance of the initial guess should thus not be underestimated.

A correct choice of optimizer is often crucial when trying to avoid local minima. Owing to the prevalence of local minima, many optimizers have been developed to help overcome the challenges. Examples include stochastic descent optimization algorithms [70, 76–78], particle swarm optimizers [79, 80], and gradient-free approaches [81]. An optimizer that has seen frequent use in the appended papers is the Simultaneous Perturbation Stochastic Approximation (SPSA) [82].

3.1.2 Barren Plateaus

An additional problem that often arises in variational quantum algorithms is that of a vanishing gradient, $\nabla\langle\hat{O}\rangle \approx 0$. This phenomenon is often referred to as a barren plateau (Figure 3.1). Here, for some operator $\hat{O}$, there is no gradient to follow during optimization. As noise is added to the flat landscape, the true gradient is almost impossible to evaluate.

The causes for barren plateaus are plenty: circuit expressiveness and parameter count [67, 68], noise [66, 83], the locality of the cost function [84, 85], and entanglement [86]. Recent efforts have tried to create a unified theory for barren plateaus, creating a framework that accounts for the mentioned causes [87, 88].



(a) Non-barren landscape. **(b)** Barren landscape.

Figure 3.1: Two-dimensional cost landscapes. A non-barren (a) landscape maintains a non-zero gradient throughout most of its domain. A barren (b) landscape includes vast barren plateaus where the gradient is close to zero. The problem of barren plateaus worsens with the dimensionality of the problem, like finding a needle in an d-dimensional haystack.

While barren plateaus are common in variational algorithms, there are several strategies to try to avoid them. One such strategy is to find a proper starting point.

3.2 The Importance of a Good Guess

In contrast to classical neural networks where a random initialization of parameters is a common and effective initialization strategy [89], random initialization of variational quantum circuits often leads to issues with local minima or barren plateaus [90]. In response, several initialization strategies have been proposed [91–94].

Instead of random initialization, optimizations in both conventional computational chemistry and the VQE often utilize an initial guess motivated by physical or chemical understanding. Commonly, the guess provides an approximate solution to the optimization problem at hand. This methodology, initialization from an approximate solution, is referred to as performing a *warm-start*. In wave function methods, warm-starting is often performed by providing the Hartree–Fock solution.

A properly chosen initial guess can often help avoid many of the issues associated with both local minima and barren plateaus [95]. At the limit of the perfect initialization, i.e., initializing with the exact parameters, one ultimately avoids the problem of optimization altogether, leaving only that of confirming convergence.

3.3 Parameter Transfer

In Paper I, we analyze a warm-start approach for the problem of Potential Energy Surface (PES) sampling. The method we apply, which we refer to as *parameter transfer*, is an approach derived from conventional computational quantum chemistry. While the problem of PES sampling is specific, the idea of parameter transfer has broader applicability. In Paper IV, we extend the parameter transfer approach to the transfer of parameters between different levels of theory, from a BO calculation to a Nuclear-Electronic Orbitals (NEO) calculation. The parameter transfer is by no means limited to these two cases.

Parameter transfer is the straightforward approach of taking a solution from one calculation and using it as the initial guess for a subsequent calculation, as illustrated in Figure 3.2. This is a tried and tested approach, widely used in many disparate fields, including conventional quantum chemistry, quantum optimization [92, 96–98], and machine learning [99, 100]. The simplicity of the approach makes it attractive to study also for the VQE algorithm.

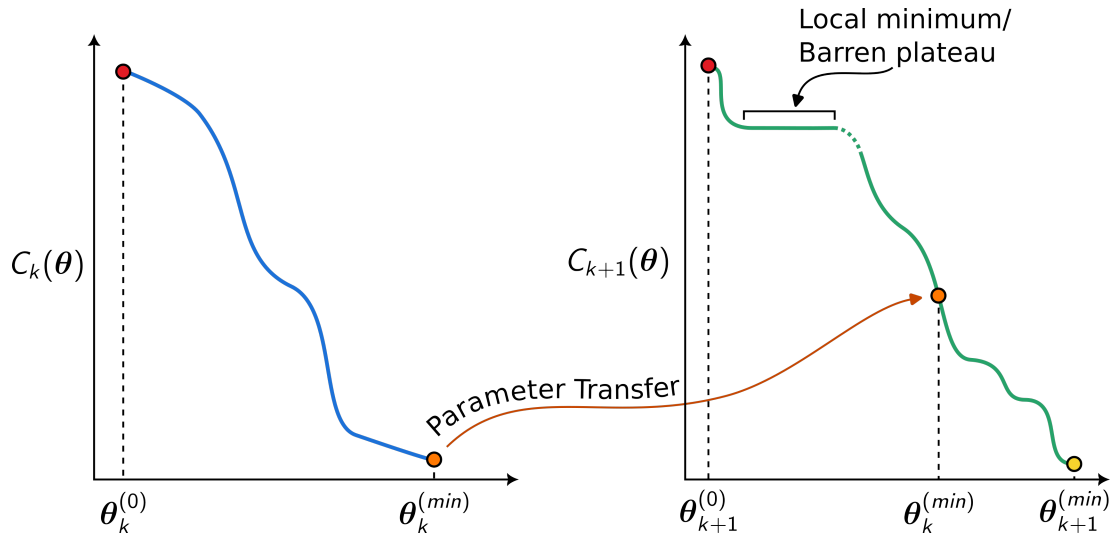


Figure 3.2: Schematic representation of the parameter transfer approach allowing for faster convergence. An initial calculation k is performed and converged for a cost function C_k. The optimized solution, $\theta_k^{(min)}$, is passed as an initial guess to the next optimization for configuration C_{k+1}. Parameter transfer can help overcome both local minima and barren plateaus.

To evaluate the performance of warm-start approaches, in Paper I we introduce the measure of *speedup*, S. Speedup compares the number of iterations needed to converge VQE calculations starting from the HF guess, N_{HF}, as compared to using a parameter transfer initial guess, N_{PT}.

For the PES sampling of Paper I, we define the speedup with respect to the source geometry, $\boldsymbol{R_S}$, and target geometry, $\boldsymbol{R}$. These different geometries result in different Hamiltonians and cost landscapes. The difference between the two geometries is given by the vector $\boldsymbol{d} = \boldsymbol{R} - \boldsymbol{R_S}$ and the speedup is defined as

$$S(\boldsymbol{R}, \boldsymbol{d}) = \frac{N_{HF}(\boldsymbol{R})}{N_{PT}(\boldsymbol{R}, \boldsymbol{d})} - 1. \tag{3.4}$$

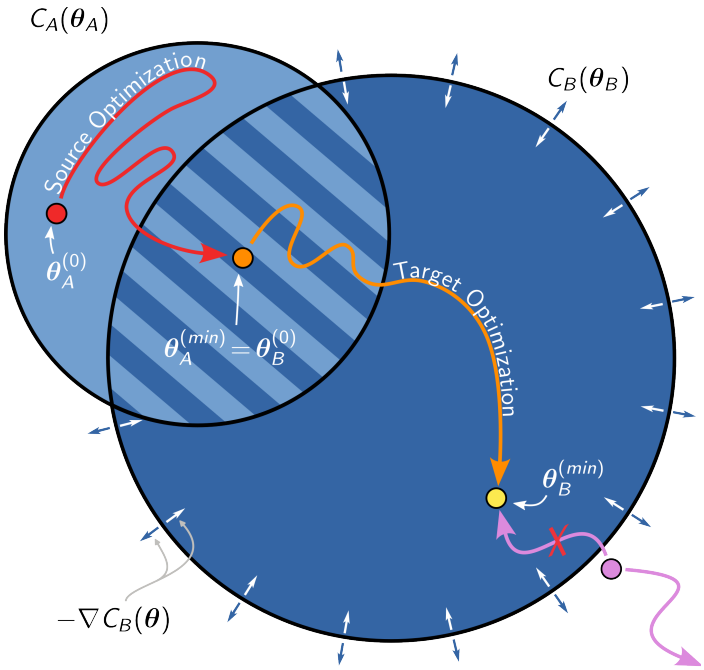


Figure 3.3: Illustration of parameter transfer overcoming a local minimum in the target cost function C_B. The local minimum is overcome by solving an initial problem C_A with a solution θ_A^{min} lying within the basin of C_B. Using θ_A^{min} as an initial guess allows for a steepest descent search for θ_B^{min}. Starting at any point outside the basin of C_B (purple) would not allow for steepest descent optimization, even though the point might lie closer.

While the speedup at a given position can be interesting, a perhaps more useful measure is the average speedup across the PES

$$\bar{S}(\boldsymbol{d}) = \frac{1}{n} \sum_{(\boldsymbol{R},\boldsymbol{d})} S(\boldsymbol{R}, \boldsymbol{d}), \tag{3.5}$$

where the notation $(\boldsymbol{R}, \boldsymbol{d})$ is used to denote the set of vectors $\boldsymbol{R}$ and $\boldsymbol{d}$ with a constant transfer distance $d = |\boldsymbol{d}|$. In effect, averaging all speedups for a given transfer distance d. One can conclude from Eq. (3.4) that the speedup is limited by the convergence of the HF starting point (or any other reference method one chooses). We find that this upper limit scales as ap^b where p is the number of the parameters. The values of coefficient and exponent are found to be $a = 2.5$ and $b = 0.25$ (Figure 3.4).

The findings of Paper I indicate that the speedup offered by the transfer of optimal parameters in the PES drops of quickly with the distance d. For a fixed $d = 0.1$ Å we find the average speedup up to be $\bar{S}_{0.1} = 0.7$. The average speedup for the four convergence curves presented in Paper IV gives $S = 1.0$, reaching $S > 3$ for the largest ansatz (42 parameters, UCCSDT).

The effectiveness of parameter transfer is also affected by other problem circumstances, such as the choice of optimizer, and whether or not the performed calculation is affected by noise. Figure 3.5 highlights these specific factors, with Figure 3.5a showing the ideal behavior of Eq. (3.5), and Figures 3.5b and 3.5c showing the impact of noise in calculations. On the one hand, for the more noise-resilient SPSA optimizer, we see an increased speedup with the use of parameter transfer (3.5b) in the noisy setting. On the other, for the COBYLA

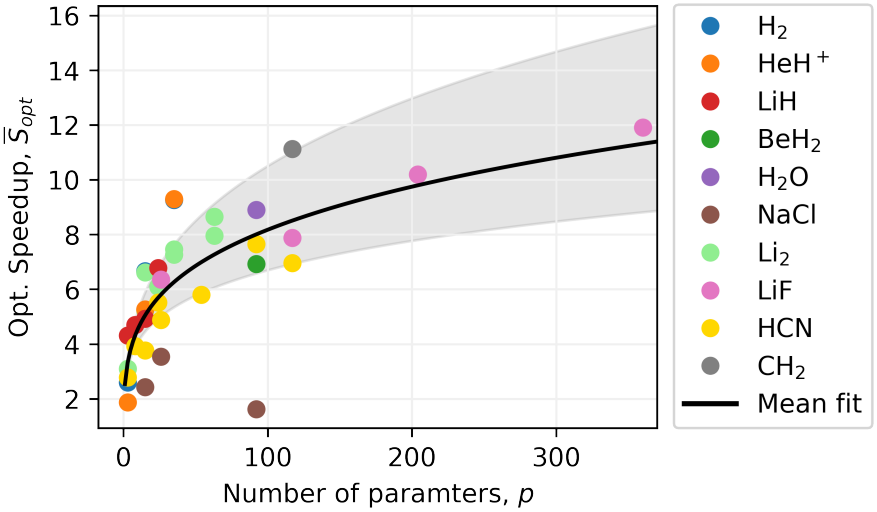


Figure 3.4: Optimal speedup, achieved for $\boldsymbol{d} = 0$ (i.e., starting from an already converged solution), as a function of the number of ansatz parameters, n. A function $f(x) = ap^b$ is fitted, leaving out the highest parameter points (LiF), where $a = 2.5403$ and $b = 0.2538$. The grey area are constructed as $ap^{b\pm\beta}$ ($\beta = 0.1775$), such that 50 % of all points fall within the area.

optimizer, less suited to the noisy environment, we see barely any effect of parameter transfer. An interesting feature of Figures 3.5b and 3.5c is the linear dependence on the transfer distance.

An unexpected consequence of the Z_2-symmetry tapering implemented in Paper IV was a sign change among some of the Hamiltonian terms and, by extension, the signs of some of the UCCSD parameters. The sign change made direct parameter transfer incorrect. However, by identifying the operators where this sign change occurred, we could account for this in the parameter transfer. Alternatively, one can change the signs of the Hamiltonian terms to match directly.

Common to all warm start methods is the idea that one can solve a related, similar problem, and reuse the solution. This line of reasoning naturally begs the question *what makes two problems similar*. While this question has come up during this work, no satisfying answer has been reached.

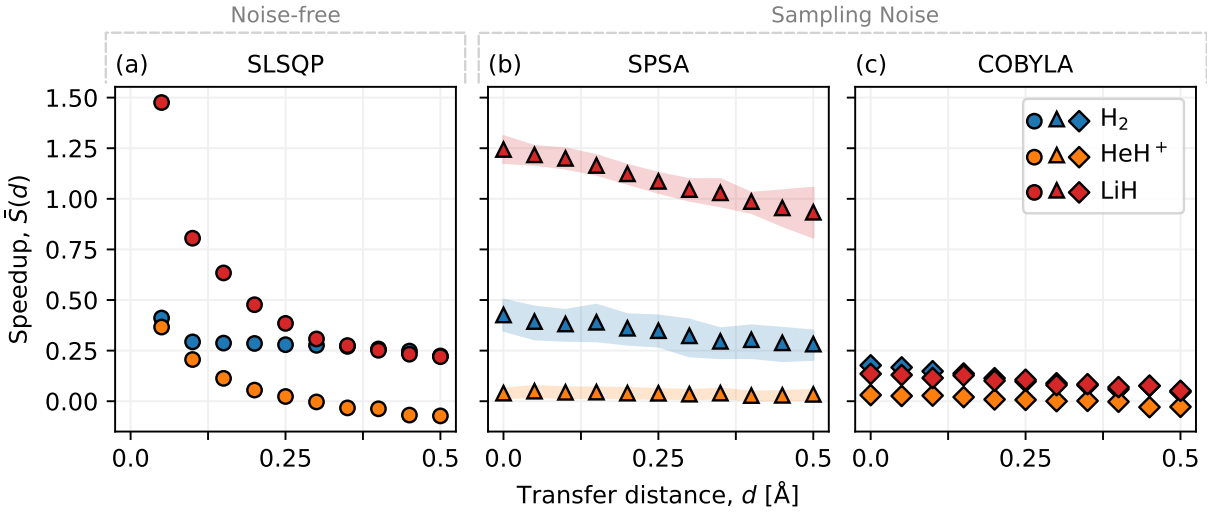


Figure 3.5: Effects of warm-start approach for three optimizers (a) SLSQP, (b) SPSA, and (c) COBYLA. Noise-free simulations (a) as well as simulations including sampling noise (b)-(c) are also included. Warm-starting affects the different optimizers differently. Additionally, the inclusion of sampling noise results in an almost linear speedup. The figure is reproduced from Paper I.

CHAPTER 4

Errors and Noise — Improvise, Adapt, and Overcome

One of the defining characteristics of current and near-term quantum computers is the prevalence of noise and errors [15, 101, 102]. Reducing the effects of noise is thus a necessity for any near-term algorithms. While variational algorithms can deal with some of the errors that noise introduces, on account of their variational flexibility [103], additional efforts often have to be applied.

Noise and errors are nothing unique to quantum computing, an inherent part of classical computing and communication is to deal with noise and errors. Many of the techniques used in quantum error correction and mitigation derive from classical methods, see for example Richardson extrapolation [104] and the Hamming error correcting codes [105]. In Paper II we introduce a low-overhead error mitigation scheme, specifically tailored for quantum chemistry calculations on NISQ hardware.

4.1 Quantum Errors and Noise

In classical digital computation there is only one type of error, the *bit flip*. The bit flip is the unwanted change of a 1 to a 0, or vice versa 0 to 1. In quantum computing a single qubit can experience any erroneous rotation of the Bloch sphere. Analogous to the classical case, bit flips can occur ($|0\rangle \rightarrow |1\rangle$ and $|1\rangle \rightarrow |0\rangle$), but also sign flips ($|0\rangle + |1\rangle \rightarrow |0\rangle - |1\rangle$ and $|0\rangle - |1\rangle \rightarrow |0\rangle + |1\rangle$)[1]. Errors can come in a number of flavors, they apply any degree of rotation, not just π rotations corresponding to flips, and they can exhibit varying degrees of correlation, amongst other properties. An exhaustive account for types of errors that can effect a quantum computer is not in the scope of this thesis.

The causes of the noise and errors are multiple, however, common sources are applying incorrect gates [106], State Preparation And Measurement (SPAM) errors [107, 108], and interactions with the environment (such as the occasional cosmic ray [109, 110]), just to name a few.

[1] Note that a change $|0\rangle \rightarrow -|0\rangle$ or $|1\rangle \rightarrow -|1\rangle$ corresponds to global phase changes, and does not affect any observable measurements.

A clear example of these type of errors arising from hardware is seen in the results of Papers II and III. In Paper II circuits for both H_2 and HeH^+ where implemented on the three qubit Särimner superconducting qubit quantum computer. Only two of the three qubits where used though. The H_2 circuit had a single parameterized gate and one two-qubit gate (Figure 4.1). The HeH^+ circuit contained three parameterized gates with four two-qubit entangling gates.

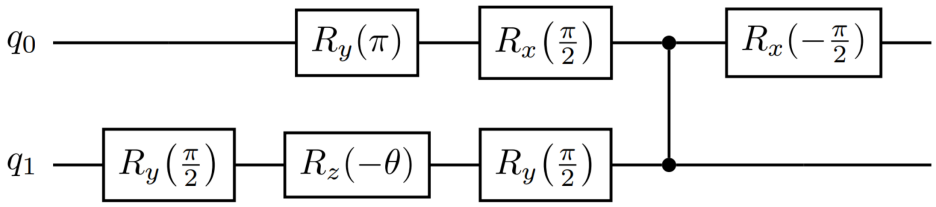


Figure 4.1: Quantum circuit used to simulate H_2 on the Särimner three-qubit device. Note the single parameterized gate. Qubits q_1 and q_2 are initialized to the $|q_1, q_2\rangle = |00\rangle$ state.

Even for the short H_2 circuit employed in Paper II, the effects of noise are substantial, especially around the ground state solution (Figure 4.2). The larger error in and around the minima (and maxima) can be motivated by imagining performing the same sweep of θ with increased noise. As noise increases, the curve will flatten out. At the point of "infinite" noise the measured expectation value loses all dependence on θ and becomes a flat line, assuming sufficient sampling. As will be shown in the next section, the state that one reaches in the limit of high noise can be modelled as a maximally-mixed state.

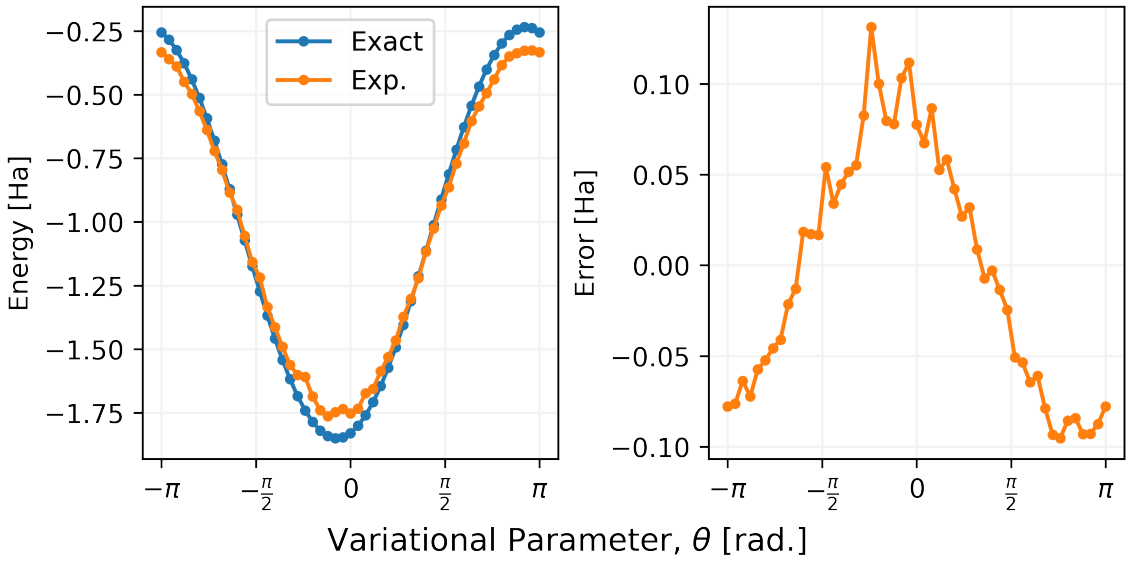


Figure 4.2: Evaluation of the H_2 Hamiltonian on the Chalmers Särimner three-qubit quantum computer, using a one-parameter hardware-efficient ansatz (Figure 4.1). Substantial errors are measured as a result of noise in the quantum calculation. The exact solution, calculated through noise-free simulation, is also shown. Each point in the graph corresponds to 5000 individual readouts of the Hamiltonian terms. Substantial errors are found in the region surrounding the ground state solution.

For the longer three-parameter (θ_A, θ_B, and θ_C) circuit, errors are, as one might expect,

even higher (Figure 4.3). Like the case for H_2, the errors are most prominent at the location of minima and maxima, with a similar flattening of the landscape occurring. A peculiar pattern can also be seen where, for $\theta_{A,B} \approx -\pi$, the error increases along a row in the data. While we do not have a satisfactory explanation for this particular feature, the error highlights the sporadic nature of noise in quantum devices.

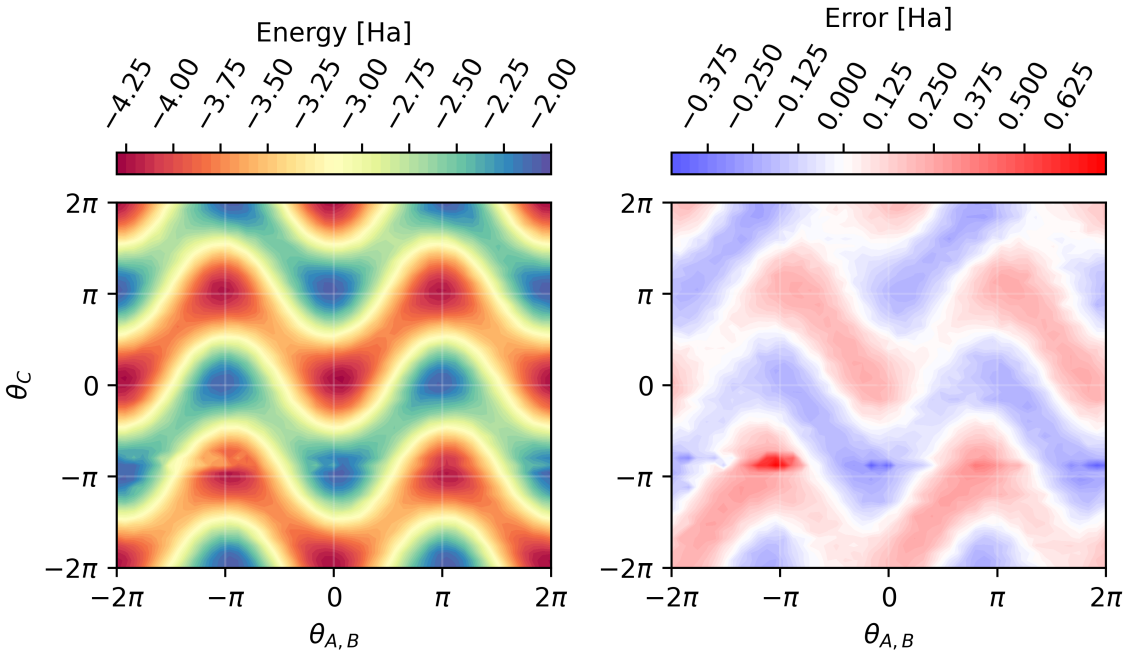


Figure 4.3: Results for calculation on the Chalmers Särimner three-qubit quantum computer, using a three-parameter UCCSD ansatz (two of the parameters are equal, $\theta_A = \theta_B$). Substantial errors are measured as a result of noise in the quantum calculation. Each point corresponds to 5000 individual readouts of the quantum state. Substantial errors are found around the ground state solution.

While the sources of errors presented so far are all due to physical effects in hardware, inaccuracies can also arise due to theoretical inaccuracies and approximations. An example of this is Trotterization where an operator cannot be implemented as-is, and instead needs to be approximated as a series of native operations. We see this error in Paper V, where Trotterization errors contribute to a substantial deviation in the nuclear-electronic entanglement during time evolution, see Figure 6.7.

The idea of errors in calculations has also brought up the discussion of accuracy as opposed to precision. In conventional computational chemistry, the concept of *chemical accuracy* exists as a well-defined measure of error with respect to experiment [111]. In particular this value is often defined as an error of no more than 1.6 mHa (1 kcal/mol) [112]. In several quantum computational papers targeting chemical calculations, chemical accuracy has been used to refer to errors with respect to the error-free calculation at *the same level of theory*. High-profile examples include [52] and [103]. As we discuss in Paper II, this is incorrect use of the term, and we propose a new term *chemical precision* to represent the error with and without noise, at the same level of theory.

4.2 Modelling Quantum Noise

Similar to how the simulation of atoms and molecules is beneficial to the understanding of those systems, simulations of quantum computers offer much insight into the behavior of the physical machines. To accurately model a physical quantum computer one also needs to include the effects of noise, see for example [113]. Accurately simulating the noise on a quantum device is an even harder problem than simulating the logical quantum computer. The extra complexity follows as the logical quantum computer is confined to its pure states, clean from any environmental effects. The noisy quantum computer on the other hand needs to account for its environment, which might very well need a quantum treatment in and of itself. For a system with a Hilbert space of dimension d, all quantum error operations can be described as an operator-sum containing no more than d^2 elements [16].

As accurately modelling the noise experienced by a quantum computer is computationally expensive in and of itself, one often opts to approximate the effects of noise. Perhaps the simplest choice of so called noise model is to approximate the noise as a depolarizing noise channel. This type of noise model is what was used to model quantum noise in Papers II, I, and III. The depolarizing noise channel can be modelled in different ways, the way used in Papers II, I, and III is to introduce it as a probability p_d that after each circuit gate another gate $P \in \{X, Y, Z\}$ is applied. This random application of gates effectively maps the resulting state onto the maximally-mixed state

$$\rho \to \lambda(1-\lambda)\rho + \frac{\lambda}{d}\boldsymbol{I} \tag{4.1}$$

where λ is a parameter corresponding to the amount of noise and d is the dimension of the density matrix ρ. Specifically, λ relates to p_d as

$$p_d = \frac{\lambda}{4}. \tag{4.2}$$

To fulfill complete positivity, λ also needs to satisfy

$$0 < \lambda < 1 + \frac{1}{d^2 - 1}. \tag{4.3}$$

The depolarizing channel can be viewed as projection onto the center of the Bloch sphere, ultimately projecting the entire Bloch sphere to the origin, resulting in a uniform mixed state. While the depolarizing noise model is easy and efficient to implement, it greatly simplifies the noise processes and the resulting errors.

A more realistic result can be achieved by explicitly modelling aspects such as decoherence and dephasing, as well as by introducing more accurate gate errors. Additionally, SPAM errors can be included by also simulating the errors when measuring. These more realistic models are not necessarily much more costly than the depolarizing noise model [114].

4.3 Mitigation and Correction

When it comes to addressing the errors caused by quantum noise, two regimes exist. Firstly, there is the idea of *error correction*. This is the ideal type of correction where errors can be

completely negated, retaining the correct quantum state for a set of *logical qubibts*. For error correction, the logical qubits are encoded in several physical qubits. Secondly, there is the idea of *error mitigation*, whereby *some* error is corrected in an observable, or a collection of observables. However, the state itself is often left untouched. As error correction seeks to correct the quantum state, it is inherently a quantum operation that has to be performed before measurement. Error mitigation on the other hand is often a correction performed after measurement, i.e., after the state has been collapsed.

One of the major differences when correcting quantum errors, as compared to the classical case, is the care one must take as to not cause a collapse of any superposition needed for calculations. In classical physics this is not a problem, we can measure the value of each bit without changing their value. For the qubit this is no longer true. By measuring we force the state of the qubit into one of its computational eigenstates $|0\rangle$ or $|1\rangle$. Thus by measuring a superposition, $|\psi\rangle = a\,|0\rangle + b\,|1\rangle$, all information contained in the phase between the states is lost, affecting any subsequent measurement.

Most variational algorithms have some degree of inherent error mitigation as they are free to adapt their parameters to shifts in the cost landscape due to noise [103]. This effect is however often far from enough. In response to this, several different error mitigation techniques have been proposed. To list a few: zero-noise extrapolation [115], probabilistic error cancellation [116], subspace expansion [117, 118], and many more, see [119–121].

Additionally, a mitigation method employed in Paper II was measurement (or readout) error mitigation [122, 123], whereby a confusion matrix $\boldsymbol{K}$ is constructed by state tomography. The entries of the confusion matrix, $K_{n,m}$, represent the probability of measuring the state $|m\rangle$ after preparing state $|n\rangle$. The confusion matrix is then used to fit a model of the SPAM errors. The model can be used for mitigation. In Paper II a least squares fitting was used together with Bayesian unfolding [124]. As the measurement error mitigation uses tomography, it is limited to mitigating a small number of qubits.

A pitfall of some error mitigation techniques is the risk of *over-correcting*, i.e., giving a value lower than what should be possible in the original problem. This is especially true for extrapolation techniques. It should be noted that over-correction can violate the variational principle, Eq. 3.2. Although, the reduced final error is often worth the risk.

4.3.1 Using a Reference

Another approach to error mitigation is to compare to already known values. While most of the cost landscape is hard to classically simulate, a subset of points can be efficiently simulated by a classical computer per the Gottesmann–Knill theorem [125]. Circuits made solely from gates in the Clifford group all fall in this category of being efficiently simulabel. As such, one can *exactly*[2] calculate what the quantum computer should yield at those points.

Perhaps the most straightforward implementation of this class of mitigation techniques comes in the form of the Reference Error Mitigation (REM), introduced in Paper II. Figure 4.4 shows the underlying idea in the one-parameter case. Here, a single Clifford state is used as a reference. By making the corresponding measurement on the quantum computer, the error at the reference point can then be exactly determined. This error can then be subtracted across the cost landscape.

[2] Your are only limited by the numerical accuracy of your computer, but in terms of current state-of-the-art quantum computers, you are pretty much exact.

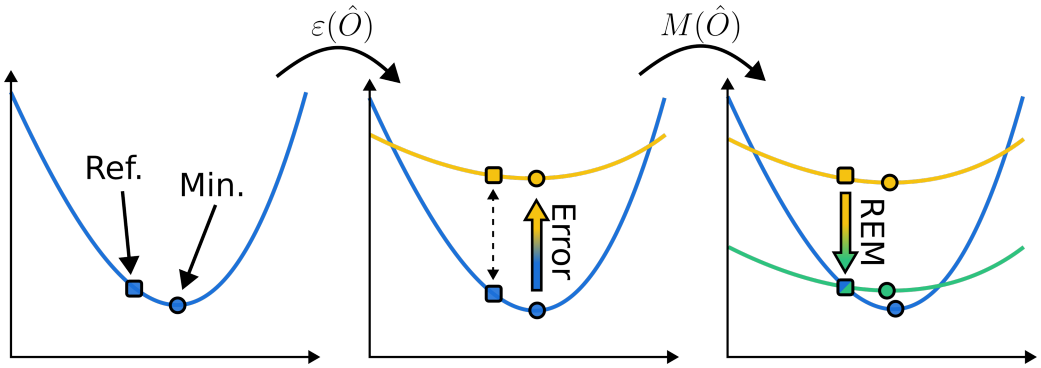


Figure 4.4: Effects of noise and subsequent correction using a reference value. A theoretically ideal optimization landscape (blue) with a true minimum (circle) and a reference point (square). The exact curve is distorted due to noise effects, $\varepsilon(\hat{O})$, measurements instead find the noise-induced minimum and reference (yellow). Through estimation of the error at the reference point (dashed line), a correction $M(\hat{O})$ (green) can be applied to the noisy results to mitigate errors around the reference.

The effects of applying the REM mitigation is substantial for cases where the reference lies reasonably close to the solution. The HF solution ($\theta = 0$) provides such a reference for H_2 at equilibrium bond distance, see Figure 4.5. The improvement in measured energy close to the reference point is clear to see, and REM outperforms the more complex measurement error mitigation in and around the minimum.

There exists two natural extensions to the REM approach: making use of more Clifford points, and/or using a higher order correction. One example is Clifford Data Regression (CDR) [126] where a large number of Clifford points are used together with a regression algorithm to find an error model that can be used to mitigate errors. However, a number of additional Clifford based methods have been proposed [127–130]. These will collectively be referred to as learning based methods. Learning based methods can use a number of models to fit their measured test sets, and many more approaches will likely appear.

One can additionally view the measurement error mitigation, presented in the previous section, as a Clifford method as it prepares Clifford states and performs fitting. The difference being the exponential number of Clifford states required to perform full state tomography. In Figure 4.6, the relation between the different approaches and their resources requirements are sketched out.

A clear benefit of higher order mitigation strategies, such as the measurement error mitigation, is their ability to correct values across the the energy landscape. They also change the shape of the landscape, potentially aiding in better gradient evaluation and by extension easier optimization.

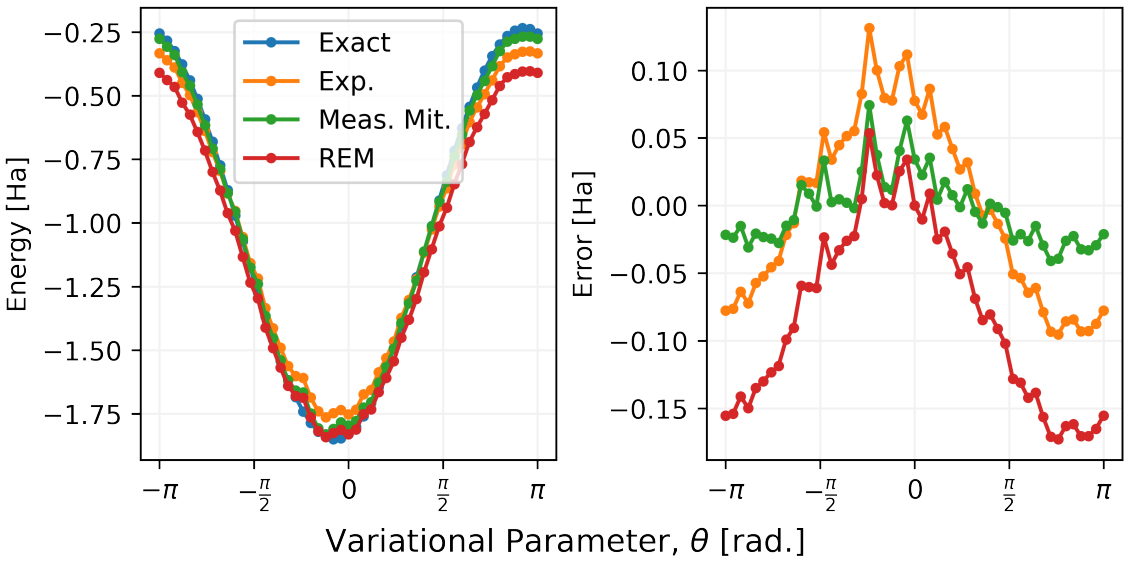


Figure 4.5: Effects of REM and measurement mitigation on experimental data from three-qubit Särimner quantum computer. The application of REM and measurement error mitigation both reduce errors in the area around the minimum, with REM offering a better correction.

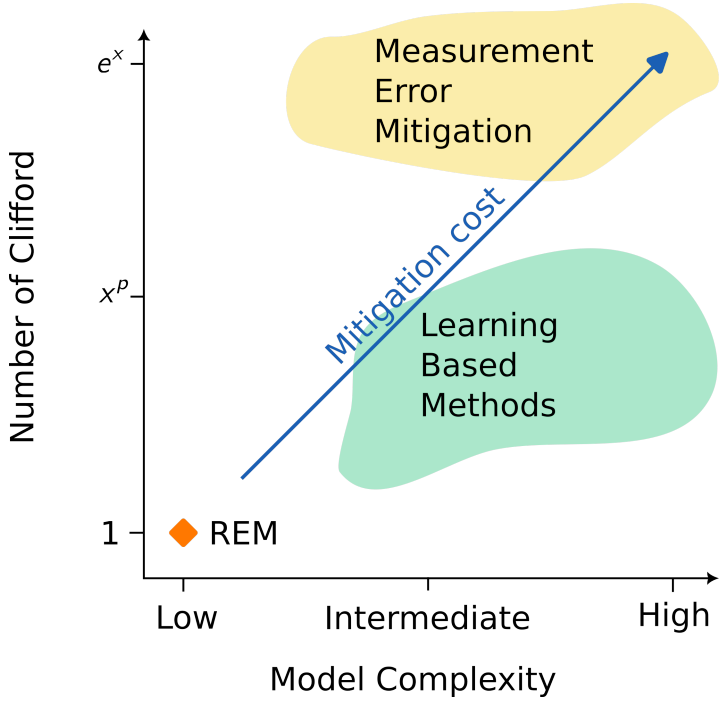


Figure 4.6: Schematic of different Clifford based methods and their required resources. Learning based methods can require a polynomial number of states to train on, while readout error mitigation requires measurement of an exponential number of Clifford states. Combined with the complexity of model used, the cost of mitigation increases with each extra Clifford state used. REM uses a single Clifford state and a constant complexity mitigation scheme, making it low cost.

CHAPTER 5

Approximating Calculation Quality

While quantum computers are currently becoming evermore powerful, classical computers are still able to outperform them for most tasks, including quantum chemical calculations. As long as this holds true, we can easily validate the output of quantum calculations by performing the corresponding classical calculation. However, if at some point in the future a quantum computer is able to outperform all their classical counterparts, how are we to know that the quantum computer is giving us a correct answer? The conventional wisdom for most of science has been to compare to experiments. As Richard Feynman put it [131]

> “It doesn’t make any difference how beautiful your guess is, it doesn’t make any difference how smart you are, who made the guess, or what his name is. If it disagrees with experiment, it’s wrong. That’s all there is to it.”

A drawback of experiments is that they are relatively expensive and often slow, resulting in a long feedback cycle for any poor computational chemist trying to predict the results of a yet to be performed experiment. It would be much more convenient if the the computational results could be pre-validated using either already available experimental data, or by comparing to some other, cheaper computational method.

In Paper III we explore using the electron density and its topology as such a validation measure.

5.1 Quantum Fidelity

If you are given two states ρ_A and ρ_B, and asked to determine *how* similar these two states are, how would you determine this? What is the measure of “likeness” for quantum states? A common answer is the Uhlmann-Josza fidelity [132], more commonly called quantum fidelity, or just fidelity. For two quantum states ρ_A and ρ_B, quantum fidelity defined as[1]

$$F(\rho_A, \rho_B) = \left(\mathrm{Tr} \left[\sqrt{\sqrt{\rho_A}\rho_B\sqrt{\rho_A}} \right] \right)^2 \leq 1, \tag{5.1}$$

[1] Some authors instead define the fidelity as $\mathrm{Tr}\left[\sqrt{\sqrt{\rho_A}\rho_B\sqrt{\rho_A}}\right]$, i.e., the square root of the value presented here. Perhaps most notably this is the definition used in [16].

where the $F = 1$ if and only if $\rho_A = \rho_B$ up to a global phase. For two pure states the fidelity simplifies to

$$F(|\Psi_A\rangle, |\Psi_B\rangle) = |\langle\Psi_A|\Psi_B\rangle|^2 \leq 1. \tag{5.2}$$

However, one can imagine other measures of fidelity (or approximate fidelity as will be discussed in the next section). To help facilitate development of fidelity measures, Richard Josza proposed four (here presented as five) axioms that such a measure $\mathcal{F}$ should satisfy [132, 133]

(J1a) $\mathcal{F}(\rho_A, \rho_B) \in [0, 1]$
(J1b) $\mathcal{F}(\rho_A, \rho_B) = 1$ if and only if $\rho_A = \rho_B$
(J2) $\mathcal{F}(\rho_A, \rho_B) = \mathcal{F}(\rho_B, \rho_A)$
(J3) $\mathcal{F}(|\psi_A\rangle\langle\psi_A|, \rho_B) = \langle\psi_A|\rho_B|\psi_A\rangle = \mathrm{Tr}[\rho_A\rho_B]$ if ρ_A is a pure state
(J4) $\mathcal{F}(U\rho_A U^\dagger, U\rho_B U^\dagger) = \mathcal{F}(\rho_B, \rho_A)$ for all unitaries U.

The Uhlmann-Josza fidelity satisfies all of the axioms. Other measures have been proposed as well, see [132].

However, problems arise when trying to actually calculate the fidelity. Firstly, the states in question are exponentially large meaning that you would require exponential storage. On top of the size of the states you would also have to perform an exponential number of calculations. Secondly, you would already need to have the correct solution in order to compare the fidelity. As for how to perform the actual measurement, two options exist. Either the states are compared on the quantum hardware directly, or the states are extracted from the quantum device the comparison is made on a classical computer.

The idea of extracting a quantum state from a quantum computer requires performing something known as full state tomography. A complete measurement of the quantum state [16, 134, 135]. Given that a quantum state requires an exponential number of complex coefficients, and that a single measurement is limited to only provide a single classical bit of information, quantum tomography is infeasible for all but the smallest states.

5.2 Approximating Fidelity

An alternative to measuring the actual fidelity is to measure a smaller set of operators that approximate the fidelity. While the Hamiltonian can be such an operator, it is often used in the variational setting. Minimizing the energy does not correspond to approximating the fidelity as one does not get a measure of the states closeness to the correct state. However, using extrapolation techniques [136] could possible qualify the Hamiltonian as an approximation for the fidelity.

Assuming ρ_t to be a target state that one wishes to approximate, and ρ_p a state that has been prepared. One then seeks to find an observable, $\hat{O}$, that can approximate $F(\rho_t, \rho_p)$. In other words

$$F(\rho_t, \rho_p) \approx \mathrm{Tr}[\hat{O}\rho_p]. \tag{5.3}$$

Ideally this observable is a so-called *fidelity witness* [137], W, an observable that fulfills two criteria. Firstly, the expectation value $\mathrm{Tr}[W\rho_p] = 1$ if and only if $\rho_p = \rho_t$, and secondly,

$\mathrm{Tr}[W\rho_p] \leq F(\rho_t, \rho_p)$. In other words, the expectation value of the witness is equal to 1 only when we have prepared state, and for all other states, the witness measurement will be a lower bound to the exact fidelity.

In Paper III we use the term fidelity witness in less strict sense to refer to any measure that can be used to gauge the quality of a calculation. Below, the same convention will be used.

5.3 Electron Density as a Fidelity Witness

Through the Hellmann-Feynman theorem, the electron density uniquely defines the forces acting on a molecule in its ground state [138]. Thus a wave function that gives the correct electron density must also have full fidelity ($F = 1$) with the ground state. The electron density is additionally an experimental observable, to varying degrees, through transmission electron microscopy [139], atomic force microscopy [140], sweeping tunneling microscopy [141], and, the perhaps the most accurate experimental method, X-ray diffraction crystallography [142]. The uniqueness of the electron density, combined with being experimentally observable, makes for an interesting candidate as fidelity witness.

The electron density[2] $\varrho(\boldsymbol{r})$ is not a density in the sense that there would necessarily be x number of electrons in a given volume with an electron density x. Instead, the electron density describes the *probability* of finding an electron at a given point in space, the probability density.

In our second quantized picture, the electron density takes the form of a weighted sum over all spin orbital products

$$\varrho(\boldsymbol{r}) = \sum_{ij} D_{ij}\, \phi_i^\dagger(\boldsymbol{r})\phi_j(\boldsymbol{r}), \tag{5.4}$$

where the coefficients D_{ij} are elements of the one-particle Reduced Density Matrix (1-RDM), $\boldsymbol{D}$. The elements of the 1-RDM are in turn defined as

$$D_{ij} = \left\langle \Psi \left| a_i^\dagger a_j \right| \Psi \right\rangle . \tag{5.5}$$

While the conventional density matrix ρ scales exponentially with the number of spin orbitals being considered, $\mathcal{O}(e^{2N})$, the 1-RDM scales quadratically, $\mathcal{O}(N^2)$, making it computationally efficient to compute.

An additional benefit of the 1-RDM comes from its many properties and constraints [23], making it a useful resource for error mitigation.An example of such a property is the trace of the 1-RDM, $\mathrm{Tr}[D]$, which needs to equal the number of the electrons in the system

$$\mathrm{Tr}[\boldsymbol{D}] = \sum_i n_i = N, \tag{5.6}$$

where n_i is the number of electrons in spin orbital i,

$$n_i = \left\langle \Psi \left| a_i^\dagger a_i \right| \Psi \right\rangle = D_{ii}. \tag{5.7}$$

[2] In an effort to reduce the risk for confusion ϱ is used for electron density, not to be confused with ρ used for density matrices.

As stated, the many properties of the electron density, not least of which are being experimentally observable and efficient to compute, makes it an alluring fidelity witness. The question that remains is how one would define the measure to go with the witness. One approach to using the electron density in this capacity is to simply check the difference at each point in space

$$F_\varrho(\varrho, \varrho_W) = 1 - \frac{1}{2N} \int |\varrho(\boldsymbol{r}) - \varrho_W(\boldsymbol{r})| \mathrm{d}\boldsymbol{r}. \tag{5.8}$$

Here, the integral is normalized with respect to the number of electrons, N. The normalized integral itself in Eq. (5.8) is a measure of infidelity, $(F - 1)$.

The integral in Eq. (5.8) is not necessarily easy to solve, and measuring *every* point in space might prove experimentally challenging. An alternative approach is to select specific points of the electron density where the comparison is made

$$F'_\varrho(\varrho, \varrho_W) = 1 - \frac{1}{C} \sum_K |\varrho(\boldsymbol{r}) - \varrho_W(\boldsymbol{r})|. \tag{5.9}$$

Here, we sum the differences at K separate points. However, reducing the number of points under consideration also makes the normalization more complicated. While for the entire space the normalization parameter is twice the number of electrons, finding a similar constant for a discrete set of points can prove difficult. Another question that arises as we move to the discrete case is how to choose these points. An intriguing choice is given by the critical points of the electron density, which can be identified through Bader analysis [143].

5.3.1 Topology of the Electron Density

In Paper III the critical points of the electron density are studied as a possible foundation for a fidelity witness. The general approach we present is illustrated in Figure 5.1. The process can be described as occurring in five steps: first a potential wave function is identified through some method (e.g., VQE). The 1-RDM is then constructed by measuring all elements D_{pq}. The 1-RDM is subsequently used to generate the electron density. Penultimately, topological features, such as critical points are calculated, and, finally, compared to the identified features of an experimental, or otherwise obtained, electron density.

An approach for analyzing the electron density is provided by Quantum Theory of Atoms in Molecules (QTAIM) [144]. Front and center to QTAIM stands the electron density, and through QTAIM analysis, critical points and atomic basins in the electron density are identified. Critical points are those points of the electron density where the gradient, $\nabla D(\boldsymbol{r})$, is equal to zero or undefined, $\nabla D(\boldsymbol{r}) = 0$. An individual critical point can further be analyzed by its the Laplacian

$$\nabla^2 D(\boldsymbol{r}) = \begin{bmatrix} \frac{\partial^2 \varrho}{\partial x^2} & \frac{\partial^2 \varrho}{\partial xy} & \frac{\partial^2 \varrho}{\partial xz} \\ \frac{\partial^2 \varrho}{\partial yx} & \frac{\partial^2 \varrho}{\partial y^2} & \frac{\partial^2 \varrho}{\partial yz} \\ \frac{\partial^2 \varrho}{\partial zx} & \frac{\partial^2 \varrho}{\partial zy} & \frac{\partial^2 \varrho}{\partial z^2} \end{bmatrix}. \tag{5.10}$$

Critical points can then be classified based on the rank, η, and the signature, κ, of the Laplacian. These two values are often grouped as a tuple (η, κ), defining the class of the critical point. Table 5.1 lists the four different classes and their names in QTAIM.

Attractors and Bond Critical Points (BCPs) were the main focus of Paper III, nevertheless, the methodology extends to treatments of all four classes. Attractors in QTAIM come

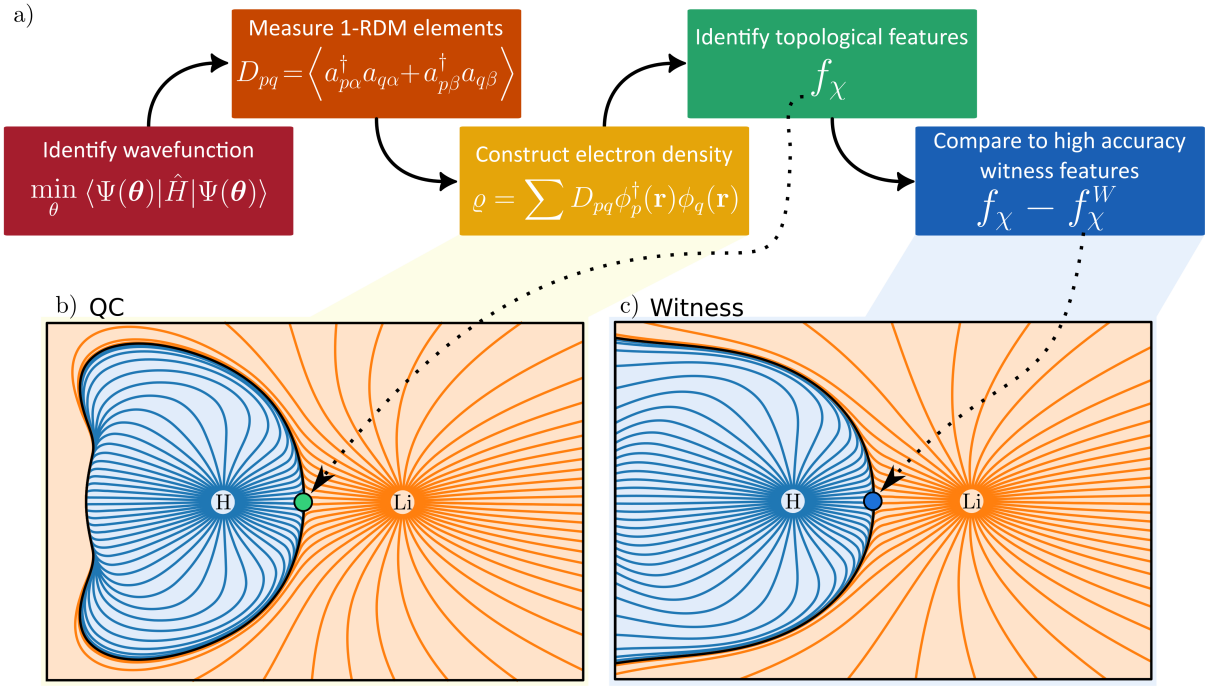


Figure 5.1: a) Outline of the methodology presented in Paper III. The shown process is divided into five steps starting with (1) using a quantum algorithm to find the wave function of interest; (2) measuring the 1-RDM; (3) from the 1-RDM constructing the electron density; (4) identify relevant topological features of the electron density; and (5) compare the identified features to a high-accuracy witness. b,c) Streamplots with overlaid Bond Critical Points (BCPs) for a calculation on quantum hardware and a simulated witness. Figure is reproduced from Paper III.

in two flavors: nuclear attractors, and Non-Nuclear Attractors (NNAs). The NNAs are a curiosity in that they are local concentrations of electrons that do not coincide with any nuclei. NNAs are in fact referred to as (pseudo) atoms in QTAIM [145]. The minima along the ridges that connect neighboring attractors are called BCPs. BCPs correspond to saddle points in the electron density.

A notable result in Paper III is the dependence on the active space for the proposed witness. In particular, the symmetry of the active space is of importance. The most explicit example of this is the results for HCN (see Table 5.2). The active space used for HCN consisted solely of π-orbitals, and since the bond axis of HCN perfectly coincides with the nodal line of all π-orbitals, the electron density along the bond axis is not changed by the quantum calculation. The denisty instead remains at the value reached by the preceding HF calculation. One thus has to take care when selecting points in the electron density.

Even though we in Paper III did not test other methods for selecting these points, an alternative approach can be to select points with respect to the active space being used. An approach could be to subtract the electron density of the frozen orbitals, ϱ_{HF}, from the total electron density to yield the active space electron density

$$\varrho_{\text{as}} = \varrho - \varrho_{HF}. \tag{5.11}$$

Such a selection would decrease the influence the choice of active space has. Although, such

Table 5.1: The four different classes of critical points, and their respective names used in QTAIM.

Class	Description	QTAIM name
(3,-3)	Local maximum	Attractor
(3,-1)	Saddle point	Bond Critical Point (BCP)
(3, 1)	Saddle point	Ring critical point
(3, 3)	Local minimum	Cage critical point

Table 5.2: Electron density and the Laplacian of the electron density at critical points along the molecular bonds. The reference used is conventional CCSD/aug-cc-pVTZ calculations. The critical points of the HCN electron density do not change during the quantum calculation. The table displays data from Paper III.

		$\varrho(\boldsymbol{r})$ [Bohr^{-3}]			$\nabla^2\varrho(\boldsymbol{r})$ [Bohr^{-5}]		
Mol	CP	Noise-free	Noisy	Witness	Noise-free	Noisy	Witness
H2[a]	bond	0.2524	0.2506	0.2684	-0.7818	-0.7542	-1.2445
LiH[b]	bond	0.0431	0.0399	0.0394	0.1437	0.1177	0.1553
Li2[c]	NNA	0.0158	0.0134	0.0129	-0.0139	-0.0102	-0.0143
Li2[c]	bond	0.0154	0.0132	0.0121	0.0017	-0.0025	0.0101
Li2[c]	bond	0.0154	0.0132	0.0121	0.0022	-0.0019	0.0101
H-CN[c]	bond	0.3934	0.3934	0.4831	1.3149	1.2480	-0.3586
HC-N[c]	bond	0.2635	0.2635	0.2990	-0.7468	-0.7346	-1.3241

[a] Calculated on Chalmers Särimner quantum computer.
[b] Calculated on IBM's ibm_quito quantum computer.
[c] Simulated on classical hardware.

a subtraction could complicate comparisons with experimental values, as a similar transformation needs to be applied to the experimental electron density.

An outstanding challenge with the use of critical points is the normalization issue of Eq. (5.9). Nevertheless, normalized or unnormalized, the electron density can still prove a useful tool for gauging the quality of quantum calculations.

CHAPTER 6

Beyond Born-Oppenheimer

Introduced in section 2.1.3, the Born–Oppenheimer (BO) approximation is a routine approximation in most quantum chemistry calculations. However, for light nuclei the BO approximation often breaks down. This is especially true for molecules containing hydrogen atoms [146]. As just so happens, hydrogen is *the* most abundant element in the universe [147], and also part of what is perhaps the most important molecule for life, *water* (H_2O). As such, going beyond Born-Oppenheimer approximation, can prove crucial for applications in fields such as bio-chemistry, pharmaceuticals, and green technology like hydrogen fuel cells [148–150]. In Papers IV and V we investigated the implementation of non-BO methods for near-term quantum computers. The calculations presented in this section were all performed as noise-free emulations.

6.1 Born-Oppenheimer in More Detail

The advent of the BO came in a time with very limited computational power. Presented in 1927 [151], still the early days of quantum mechanics, the BO approximation allowed physicists to break the relatively big problems of atoms and molecules into two smaller problems: that of the electrons, and that of the nuclei. This split can reduce the size of the Hilbert space by many orders of magnitude.

Restating the BO approximation, the molecular wave function $|\Psi\rangle$ can be separated into two parts: one describing the nuclei, $|\psi_N\rangle$, and one describing the electrons, $|\psi_e\rangle$. Under the BO approximation, $|\Psi\rangle$ can be written as a product state $|\Psi\rangle = |\psi_N\rangle \otimes |\psi_e\rangle$, i.e., a state without entanglement between nuclei and electrons. The product state allows for the two sub-systems (the nuclear and electronic) to be solved almost independently. The remaining interaction will be simpler than what is captured in the full wave function description, often being a purely classical potential. Normally, one considers the electrons as interacting with the potential formed by the protons, or vice versa, the protons moving in the potential of the electrons.

Central to the BO approximation is the separation of scales. In particular the large difference in mass between the nuclei and the electron. Specifically, while the rest mass of the proton (the hydrogen nuclei) is 1.672×10^{-27} kg, the electron's rest mass is a mere $m_e =$

9.109×10^{-31} kg[1] [152]. The ratio between the two is thus close to 1836 ($\frac{m_p}{m_e}$). The much smaller mass of the electron allows it to quickly react to nuclear motion.

Another approximation that is often used in conjunction with the BO approximation is the idea of *clamped-nuclei*, which states that the position of the nuclei is fixed [153]. In effect the nuclei are approximated as point-like classical charges. The separable state and clamped nuclei approximations are often taken as one and the same, however, this generally not the case.

6.2 Nuclear-Electronic Orbitals

The effectiveness of the BO approximation has been proven time and again, and many methods exist to account for some of the nuclear-electronic effects without having to abandon the BO approximation. Examples include semi-classical approximations [154], Ehrenfest dynamics [155, 156], surface hopping [157], and imaginary time path integral based methods [158]. Nonadiabatic effects near surface crossings and conical intersections are perhaps the clearest examples of non-BO effects in chemistry [159]. Many processes in chemistry are governed to some extent by these nonadiabatic effects. Examples include radiationless decay [160], energy and charge transfer [161], and thermodynamic and spectroscopic phenomena [162]. Further indication to the importance of a proper quantum treatment of nuclei is given by quantum nuclear and kinetic isotope effects [163, 164], where the primary nuclear quantum effects are tunneling, zero point energy, and delocalization. These effects can be substantial in a wide variety of cases, ranging from ultrafast spectroscopy [165], to enzymatic reactions [149], and the deceivingly simple-looking hydrogen molecule [166].

At the limit of including all nuclei, the wave function reads as follows

$$|\Psi\rangle = |\phi_e, \phi_n\rangle \,, \tag{6.1}$$

where the the state can no longer be seen as a product state, $|\phi_e, \phi_n\rangle \neq |\phi_e\rangle \otimes |\phi_n\rangle$. A middle ground between the BO approximation and the fully quantum treatment can be reached by adopting the NEO approach [167–170]. The NEO framework allows to include some of the nuclear quantum effects with the largest de Broigle wavelength. The NEO wavefunction can be written as

$$|\Psi\rangle = \left|\psi_{n,\mathrm{BO}}\right\rangle \otimes |\psi_e, \psi_n\rangle \,, \tag{6.2}$$

where $\left|\psi_{n,\mathrm{BO}}\right\rangle$ represent nuclei treated under the BO approximation, while $|\psi_e, \psi_n\rangle$ contains nuclei and electrons treated as part of a non-separable state. Figure 6.1 illustrates the three approaches, going from a fully BO picture, through NEO, to a fully quantum treatment.

For both Paper IV and V the NEO approach was utilized to include some protons into the wave function, with other nuclei being treated as classical point charges. As only protons were additionally included, all particles in the wave function are fermions. In general, transitioning to an electron-nuclear wave function, also requires treating fermion-bosonic systems, something that the NEO approach is fully capable of treating. The Hamiltonian

[1] As I find the precision with which experimentalists are able to measure these masses fascinating, here are the masses with even more decimals: $m_p = 1.672\,621\,923\,69(51) \times 10^{-27}$ kg, $m_e = 9.109\,383\,701\,5(28) \times 10^{-31}$ kg. Truly amazing!

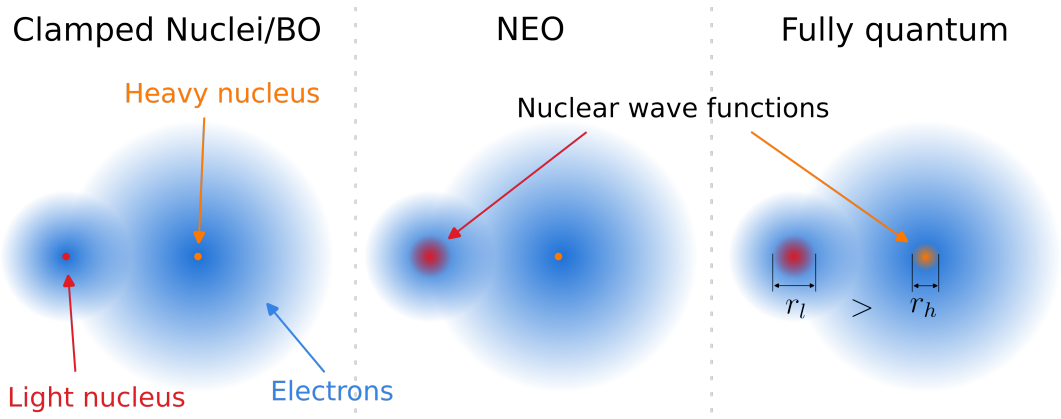


Figure 6.1: Schematic illustration of the transition from a BO approximation, through the NEO representation, to a fully quantum treatment. In the clamped nuclei picture (left), the nuclei are treated as charged single points in space. The NEO representation (middle) allows to treat some nuclei (often the lightest) quantum mechanically by introducing nuclear basis functions. Finally, all nuclei can be treated as part of the wave function (right). The width of the nuclear probability distribution are larger for lighter nuclei, r_l, than for heavier, r_h.

used in papers IV and V takes the shape

$$\begin{aligned}\hat{H}_{NEO} = &\sum_{ij} h_{ij} a_i^\dagger a_j + \sum_{ijkl} h_{ijkl} a_i^\dagger a_j^\dagger a_k a_l \\ &+ \sum_{IJ} h_{IJ} a_I^\dagger a_J + \sum_{IJKL} h_{IJKL} a_I^\dagger a_J^\dagger a_K a_L \\ &+ \sum_{ijIJ} h_{ijIJ} a_i^\dagger a_j a_I^\dagger a_J \\ &+ \sum_{AB} \frac{Z_A Z_B}{|R_A - R_B|}\end{aligned} \tag{6.3}$$

where, indices i, j run over electron spin orbitals, I, J run over protonic spin orbitals, and A, B index classical nuclei. In addition to the electronic one- and two-particle integrals h_{ij}, h_{ijkl}, the n second quantized Hamiltonian contains the protonic one- and two-particle integrals h_{IJ}, h_{IJKL}, as well as the nuclear-electron h_{ijIJ} integral defined as per Eqs. (2.10) and (2.11). The additional terms account for the nuclear kinetic energy and the interaction with the classical nuclei (), as well as the nuclear-nuclear () and nuclear-electron () interactions.

We extend the UCC ansatz, to the NEO framework, which we will refer to the NEO-UCC ansatz

$$U_{\text{NEO-UCC}} = e^{T - T^\dagger} \tag{6.4}$$

where the cluster operator, T, is now defined as to also include the active nuclear orbitals. The cluster operator reads as

$$T = \sum_{(o)} T^{(o)} = T^{(1)} + T^{(2)} + \ldots \tag{6.5}$$

where the excitation order o accounts for both protonic and electronic excitations. Specifically, it sums the electron excitation order, e, together with the proton excitation order, p, as $o = e + p$.

$$T = \sum_o T^{(o)} = \sum_o \sum_{e=1}^{o} T^{(e,o-e)}. \tag{6.6}$$

Explicitly expanding these operators in terms of fermionic creation and annihilation operators gives

$$T^{(1)} = T^{(1,0)} + T^{(0,1)} = \sum_{ij} \theta_j^i \, a_i^\dagger a_j + \sum_{IJ} \theta_J^I a_I^\dagger a_J \tag{6.7}$$

$$\begin{aligned} T^{(2)} &= T^{(2,0)} + T^{(1,1)} + T^{(0,2)} \\ &= \sum_{ijkl} \theta_{lk}^{ij} \, a_i^\dagger a_j^\dagger a_k a_l + \sum_{ijIJ} \theta_{jJ}^{iI} \, a_i^\dagger a_j a_I^\dagger a_J + \sum_{IJKL} \theta_{LK}^{IJ} \, a_I^\dagger a_J^\dagger a_K a_L \\ &\vdots \end{aligned} \tag{6.8}$$

It is clear from this definition that the NEO-UCC forms a superset to the conventional electronic UCC, with additional terms for the protonic and the mixed excitaions. It is important to note that the crucial terms, those taking us beyond the BO effects, are the mixed terms, i.e., where $e \neq 0, p \neq 0$. The optimal parameters $\boldsymbol{\theta}$ can be found using the variational principle, see Eq. (2.18).

As it stands, Eqs. (6.7) and (6.8) represent the generalized NEO-UCC operators. The generalized UCC operator allows for multideterminantal reference states [171]. However, since most ansätze act on a single-determinant HF state, the excitations are normally restricted as to go from initially occupied orbitals to unoccupied orbitals. This is also the case in Papers IV and V. While the generalized operators do offer higher expressivity and potentially higher accuracy [54], it suffers from the large number of gates needed for implementation.

In Paper IV we found that including mixed doubles $T^{(1,1)}$ and triples $T^{(2,1)}$ for H_2 improves the energy by roughly 1 mHa. Also in Paper IV, we found the proton tunneling barrier in malonaldehyde is widely overestimated in the BO using Complete Active Space Configurational Interaction (CASCI) (11.962 mHa) as compared to the inclusion of nuclear degrees of freedom in the corresponding Nuclear-Electronic Orbital Complete Active Space Configurational Interaction (NEOCASCI) (5.011 mHa). Both of these calculations highlight the importance of non-BO interactions when striving for chemical accuracy.

6.3 Qubit Mapping

With the tools to represent, prepare, and measure the NEO wave function, what remains is to map the states and operators to qubit representation. As discussed in section 2.3.2, several encodings exist that can map a fermionic problem to a problem represented on qubits. As the Jordan-Wigner offers perhaps the most straightforward and pedagogical mapping, albeit not the most efficient, it is a good starting point.

Extending the JW encoding to the combined NEO wave function can be done in several ways. A convenient starting point is to order the orbitals by particle type and spin. That is,

a ket describing a Slater determinant would thus be encoded as

$$|\Psi\rangle = \big|\psi_p, \psi_e\big\rangle = |\overbrace{f^p_{n^p_\alpha+n^p_\beta}, \cdots, f^p_{n^p_\alpha+1}}^{\alpha-\text{protons}}, \underbrace{f^p_{n^p_\alpha}, \cdots, f^p_1}_{\beta-\text{protons}}, \overbrace{f^e_{n^e_\alpha+n^e_\beta}, \cdots, f^e_{n^e_{\alpha+1}}}^{\alpha-\text{electrons}}, \underbrace{f^e_{n^e_\alpha}, \cdots, f^e_1}_{\beta-\text{electrons}}\rangle. \quad (6.9)$$

Using the JW encoding, the mapping is simply one-to-one, that is, each qubit directly encodes the occupations of a fermionic orbital, thus giving the corresponding qubit state

$$|\Psi\rangle = |\overbrace{q^p_{n^p_\alpha+n^p_\beta}, \cdots, q^p_{n^p_\alpha+1}}^{\alpha-\text{protons}}, \underbrace{q^p_{n^p_\alpha}, \cdots, q^p_1}_{\beta-\text{protons}}, \overbrace{q^e_{n^e_\alpha+n^e_\beta}, \cdots, q^e_{n^e_{\alpha+1}}}^{\alpha-\text{electrons}}, \underbrace{q^e_{n^e_\alpha}, \cdots, q^e_1}_{\beta-\text{electrons}}\rangle. \quad (6.10)$$

The fermionic operators are mapped using the parity encoding, according to Section 2.3.2. This allowed tapering of four qubits, corresponding to the four classes of distinguishable particles: α- and β-spin electrons and protons. We further identified Z_2-symmetries corresponding to point group symmetries, and spin polarization, making tapering of additional qubits possible. The suggested tapering schemes halves the number of qubits needed to represent the H_2 molecule, as shown in Table 6.1.

Table 6.1: Resource estimations for the H_2 molecule implemented within the NEO framework. The three different tapering schemes are applied in succession. $|\{\hat{H}\}|$ indicates the number of Pauli terms in the encoded Hamiltonian after application of the different tapering schemes. Table is reproduced from Paper IV.

		# Two-qubit Gates		
Reduction	# Qubits	TwoLocal	NEO-UCCSD	$\lvert\{\hat{H}\}\rvert$
None	16	968	2546	861
+Four-qubit	12	558	2204	828
+PG symmetry	11	465	1472	825
+Spin projection	8	255	1202	428

6.4 Including Dynamics

The NEO framework, so far only applied to static problems, can be extended to the time-dependent case. The evolution of a quantum system over time is something that we will refer to as quantum dynamics.

The time-evolution of a wave function with a given Hamiltonian $\hat{H}$ from a starting time t_0 to some time t_f reads as follows

$$\big|\Psi(t_f)\big\rangle = \mathcal{U}(t_f, t_0)\,|\Psi(t_0)\rangle\,, \quad (6.11)$$

where the time-evolution operator is defined as

$$\mathcal{U}(t_f, t_0) = \mathcal{T}\exp\left[-i\int_{t_0}^{t_f} \hat{H}(t)t\;\mathrm{d}t\right]. \quad (6.12)$$

Here, $\mathcal{T}$ is the time-ordering operator which ensures that any product that could arise in the exponent is ordered such that operators acting at a time t_1 are to the right of operators acting at a time t_2, where $t_1 < t_2$.

The integral in Eq. 6.12 can be approximated as a sum

$$-i \int_{t_0}^{t_f} \hat{H}(t) t \, \mathrm{d}t \approx -i \sum_k \hat{H}(t_j) \delta t_j \tag{6.13}$$

which in turn can be Trotterized to give an iterative form

$$|\psi_{k+1}\rangle = e^{-iH_j \Delta t_j} |\psi_j\rangle = U(t_j + \Delta t, t_j) |\psi_j\rangle \, . \tag{6.14}$$

Classical implementations of the time evolution often employ numeric differential equation solver, as finding an analytical expression for Eq. (6.12) may prove difficult. The Euler forward method gives perhaps the simplest example,

$$|\psi_{k+1}\rangle = |\psi_j\rangle - i\Delta t H(t_j) |\psi_j\rangle \tag{6.15}$$

whereby one approximates the evolution for the duration Δt as being $H(t_j)$.

More elaborate methods exist as well. A common type of solver for these types of differential equations is the Runge–Kutta family of methods (which includes Euler forward) [172]. In Paper V we use fourth-order Runge–Kutta as a reference to benchmark our quantum implementation.

6.4.1 Time-Dependent Hamiltonian

We setup a model system for proton transfer in the double-well potential present in malonaldehyde (Figure 6.2). The PES was identified by performing a series of single-point calculations moving the proton between the two oxygen atoms, using second order Møller–Plesset perturbation theory (MP2) in the non-BO setting. The remaining atoms in the molecular scaffold were frozen, and taken from MP2 structure optimization in C_s symmetry.

Having identified three stationary structures where the proton is located on left and right minima (C_s equilibrium), respectively, and top of the barrier (C_{2v} transition state), we used them to create three NEO Hamiltonians: $\hat{H}_L$, $\hat{H}_M$, and $\hat{H}_R$. These Hamiltonians, based on the HF orbitals, sharing the same active space for the protons, while the electronic active space differs. Specifically, the electronic active space is composed of HF orbitals where the classical proton is located in the left, right, and middle positions in the BO picture. The time-dependent Hamiltonian is expanded as a linear combination of three different Hamiltonians with time-dependent coefficients

$$\hat{H}(t) = \alpha(t)\hat{H}_L + \beta(t)\hat{H}_M + \gamma(t)\hat{H}_R, \tag{6.16}$$

where the coefficients fulfill $\alpha(t) + \beta(t) + \gamma(t) = 1$. An illustration of the evolution of the coefficients is given in Figure 6.4.

The resulting time-dependent Hamiltonian models the coupled proton-electron dynamics, showcasing nuclear quantum effects beyond the BO approximation.

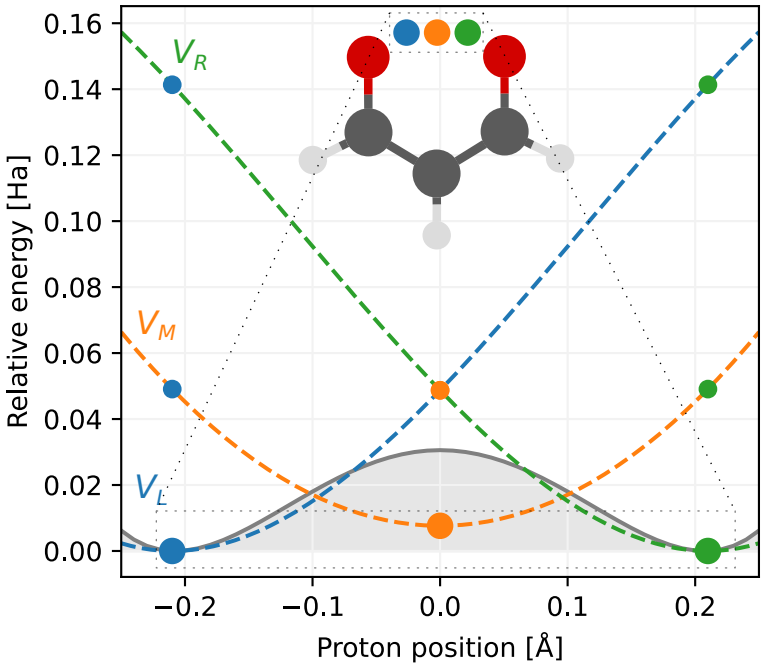


Figure 6.2: Illustration of the malonaldehyde molecule and the three proton centers (blue, orange, and green). The gray curve represents the underlying PES and is not to scale. The electronic potentials (V_L, V_M, V_R) associated with the different stationary points are sketched by fitting a third-order polynomial.

6.4.2 Quantum Algorithms for Time Evolution

Several options exist for simulating quantum time evolution on quantum computers. For adiabatic evolution, one can use the VQE to find the ground state at all given times t, effectively giving the time evolution. For nonadiabatic evolution, the evolution will no longer follow the ground state energy. An alternative approach is to encode all the steps taken in a process such as Eq. (6.14) into the quantum circuit. In such an approach, each time evolution step, $e^{iH_i\delta t_i}$, Suzuki–Trotter decomposition. In Paper V this Trotterized approach is chosen, implementing a first-order Trotterization, results of which will be discussed in the next section. The explicit Suzuki–Trotter decomposition used was

$$\mathcal{U}_1(t_f, t_0) \approx \prod_{t_j=t_0}^{t_{N-1}} U(t_j + \Delta t, t_j) = \prod_{t_j=t_0}^{t_{N-1}} \prod_{k=1}^{K} e^{-ih_k(t_j+\Delta t/2)P_k\Delta t}, \tag{6.17}$$

where P_k are the Pauli terms in the encoded Hamiltonian.

An alternative approach to the Trotterized scheme shown in Figure 6.5 is Variational Quantum Time Evolution (VarQTE) [173], where the time evolution can be approximated by a much shorter circuit. A challenge with the VarQTE approach is the need to approximate the fidelity of a state, a problem that will be further explored in Chapter 5.

In Paper V we employ the Trotterization scheme using Suzuki first-order decomposition, for both the adiabatic and the nonadiabatic regimes, Figure 6.6. Comparing to a classical calculation utilizing fourth-order Runge–Kutta, the calculations on the emulated quantum computer perform well, closely following the Runge–Kutta reference.

The difference in rate of the protonic versus electronic dynamics is clear to see when

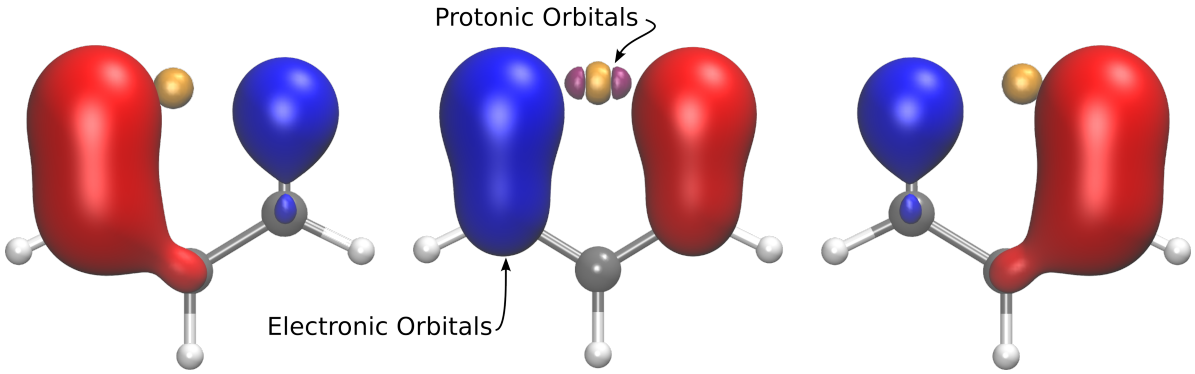


Figure 6.3: Part of the electronic and protonic active space used for setting up $\hat{H}_L$, $\hat{H}_M$, and $\hat{H}_R$, illustrating the delocalization of electrons and the top-most proton.

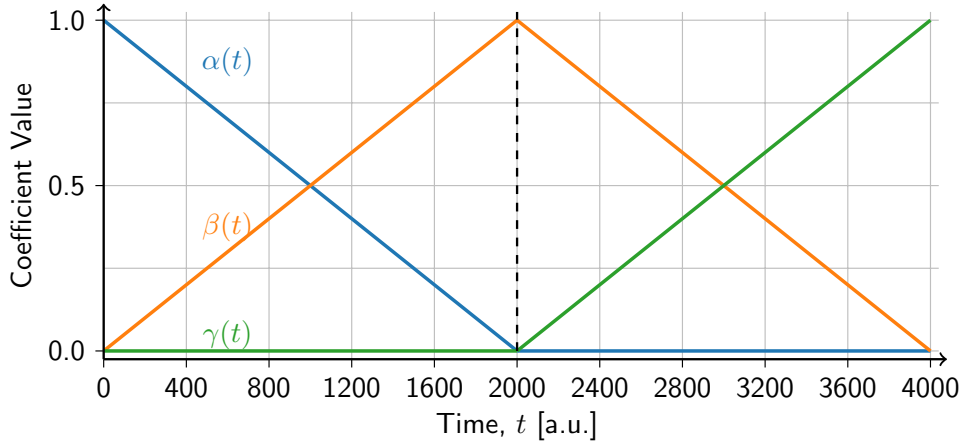


Figure 6.4: Mixing coefficients of the three Hamiltonians used to simulate time evolution in Paper V. The effective Hamiltonian is constructed as $\hat{H} = \alpha(t)\hat{H}_L + \beta(t)\hat{H}_M + \gamma(t)\hat{H}_R$, with $\alpha(t) + \beta(t) + \gamma(t) = 1$. Reproduced from Paper V.

comparing the protonic and electronic occupations (second and third panels from the bottom in Figure 6.6). The protonic occupation slowly evolves, with some oscillations arising due to slight deviations from the instantaneous ground state. The electrons in contrast show high frequency oscillations through out, albeit with lower amplitudes.

Comparing the adiabatic and nonadiabatic cases, the adiabatic calculation shows significantly less oscillations in the protonic occupations, owing to the slower driving of the system. The same cannot be said for the nonadiabatic case, where the occupations oscillate significantly. Similar behaviors can be seen in the energies of the two systems, where the adiabatic energy smoothly changes for the three Hamiltonians. An interesting feature of the nonadiabatic evolution is decreased amplitude in the oscillations of the electronic occupation.

An important point to make in regards to Figure 6.6 is the ability of the adiabatic evolution of to reach the ground states of each Hamiltonian during time-evolution. Reaching the ground states verifies that the evolution is indeed adiabatic.

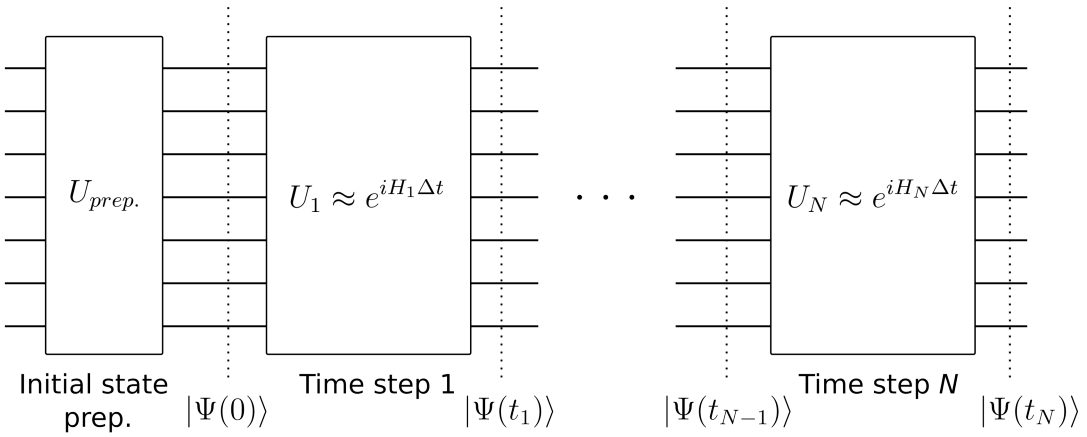


Figure 6.5: Trotterized time evolution circuit. An initial state, $|\Psi_0\rangle$, is prepared by means of UCC or some other stationary state solver, $U_{prep.}$. $|\Psi_0\rangle$ is then evolved according to the Trotterized approximation $U_k \approx \exp[iH_k\Delta t]$.

6.4.3 Nuclear-Electron entanglement

The NEO approach allows for evaluation of the entanglement between nuclei and electrons. One measure of entanglement is the von Neumann entropy (also called entanglement entropy)[2], defined as [175, 176] By bipartioning the system into electronic and nuclear subspaces, the entanglement between the two subspaces can be measured.

$$s = -\operatorname{Tr}\left[\rho_e \ln \rho_e\right] = -\operatorname{Tr}\left[\rho_p \ln \rho_p\right] \tag{6.18}$$

The resulting entanglement entropy can be seen in Figure 6.7.

A substantial deviation from the expected entropy can be seen in the adiabatic case. This error is due to size of the time steps, Δt, differing between the two methods leading to a noticeable Trotterization error. This can be resolved by either decreasing the time step, or increasing the order of the Trotterization. Looking at the exact values, the adiabatic evolution effectively returns to the initial entanglement, as is to be expected from the symmetry of the system. This is in agreement with the recovery of the Hamiltonian ground states shown in Figure 6.6. In the nonadiabatic case the entanglement instead continues to increase, reaching a maximum as the simulation ends. The asymmetry and higher entanglement is indicative of nonadiabatic behavior induced by nonadiabatic evolution.

[2]Compare to the classical Shannon entropy $s = -\sum_i p_i \ln p_i$. If the von Neumann entropy is expressed in terms of eigenvectors $|\phi_l\rangle$ and eigenvalues λ_l, then $\rho = \sum \lambda_l |\phi_l\rangle |\phi_l\rangle$ and the two measures align, $s = \sum_l \lambda_l \ln \lambda_l$ [174].

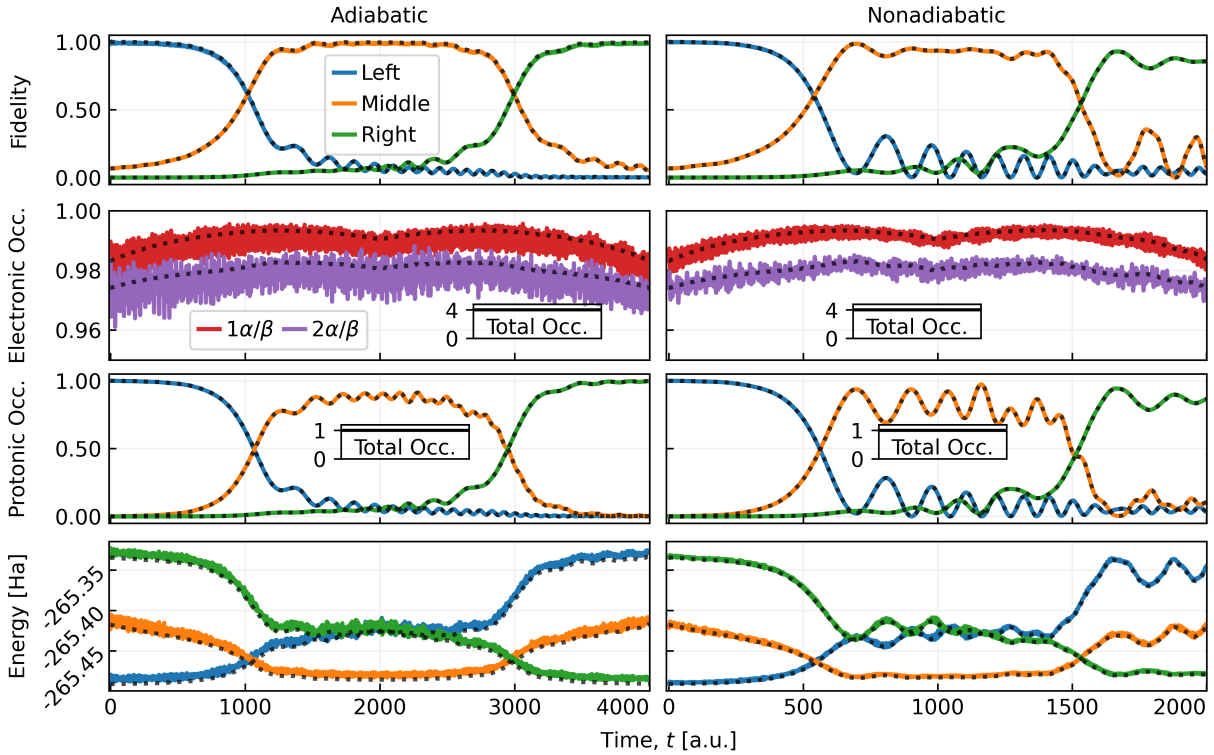


Figure 6.6: Adiabatic (left) and nonadiabatic (right) time evolution performed by Suzuki–Trotter decomposition. The gray, dashed line is a near-exact reference calculated using fourth-order Runge–Kutta. Reproduced from Paper V.

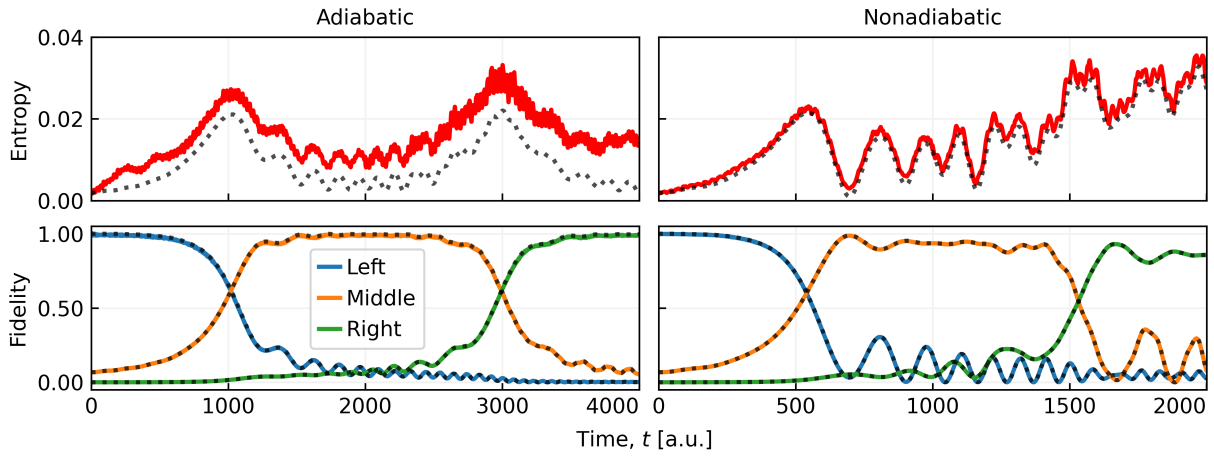


Figure 6.7: Fidelity and proton-electron entanglement during adiabatic (left) and nonadiabatic (right) time evolution. The dashed lines are a near-exact reference calculated using fourth-order Runge–Kutta. The entanglement shows a substantial error in the adiabatic domain, owing to larger time step and insufficient Trotterization used in the time evolution. Figure is reproduced from Paper V.

CHAPTER 7

Summary of Papers

Paper I

Accelerating Variational Quantum Eigensolver Convergence using Parameter Transfer

A common computational task in chemistry is the calculation of a Potential Energy Surface (PES). This involves determining the energy of a molecular system as a function of the atomic positions of the molecule, and in turn often comes down to somehow sampling this surface with a number of points. The time to solution for many VQE calculations is often prohibitively expensive, both on simulators and real quantum hardware. It is therefore important that these times are reduced. One way to improve upon this is to borrow ideas from conventional calculations and try to relate points on the PES such that one can reuse converged parameters from one calculation to warm-start the next.

This paper explores this methodology on simulated VQE calculations with a varying degree of separation between the points being studied as well as for systems with a different degree of fidelity with the Hartree-Fock state. We find that the reduction of optimization iterations required scales with the number of parameters in the circuit, and that the speedup this incurs also scales inversely with the fidelity of the trial state with the Hartree-Fock state. For a set transfer distance of 0.1 Å (Figure 7.1), the achieved speedup across PESs varies greatly, with an average of of 73 % across the test set.

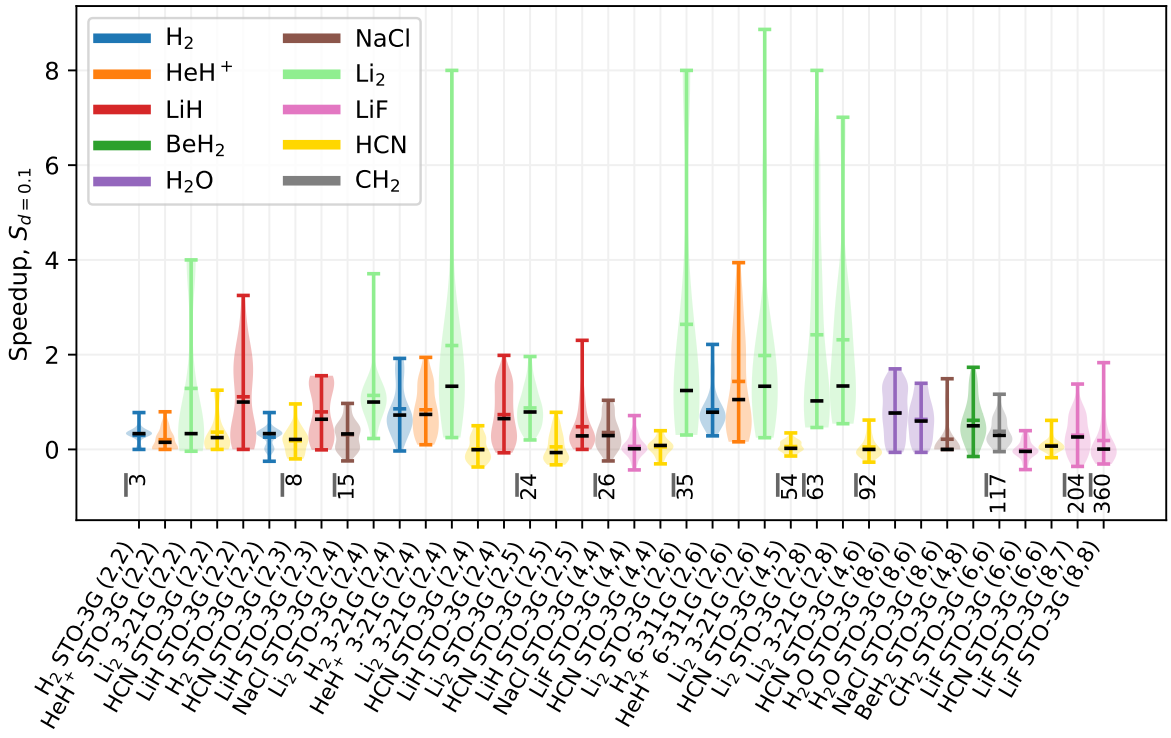


Figure 7.1: Speedup for an assortment of small molecules at a transfer distance d of 0.1 Å. Molecules are sorted according to the number of parameters in the respective VQE ansatz. The width of each violin is in proportion to the number of points along the PES that achieve the given speedup. The figure is reproduced from Paper I.

Paper II

Reference-State Error Mitigation: A Strategy for High Accuracy Quantum Computation of Chemistry

Perhaps the most substantial hindrance to quantum computing today is the amount of noise experienced during computation. For those doing calculations on NISQ hardware there is seldom a result that does not require substantial post-processing through several noise-mitigation steps. There exists several methods to accomplish this, however, many require substantial amounts of pre- or post-processing in the form of additional operator measurements, noise extrapolation, or other types of mitigations.

In this paper we present a simple and almost zero-overhead mitigation technique. The method which we name Reference Error Mitigation (REM) can drastically increase the accuracy for many types of chemical NISQ calculations. By leveraging an initial state in the VQE optimization procedure, often chosen to be the Hartree-Fock state, we can correct for a substantial amount of the noise in the Fock space close to our initial guess. Figure 7.2 shows the effects of REM along the PES of H_2 and HeH^+.

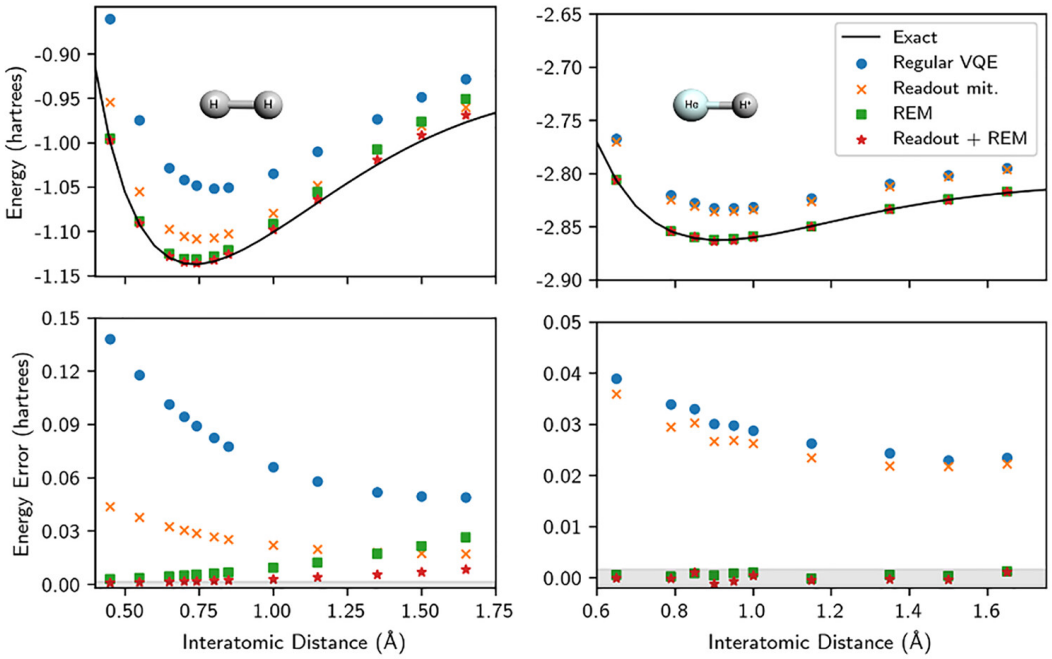


Figure 7.2: Mitigation effects of on the PESs for H_2 and HeH^+. In addition, readout error mitigation is employed. The lower figures show the error, with the shaded region corresponding to 1.6 mHa. The figure is reproduced from Paper II.

Paper III

The Electron Density: A Fidelity Witness for Quantum Computation

As quantum computers become more powerful, there might come a time where the calculations they perform are no longer classically tractable to exactly verify. To help verify the results of future quantum computers, we implement a fidelity surrogate based on the topological features of the electron density. We find that the effectiveness of the surrogate is heavily dependent on the symmetry and properties of the studied molecule, and the active space under consideration.

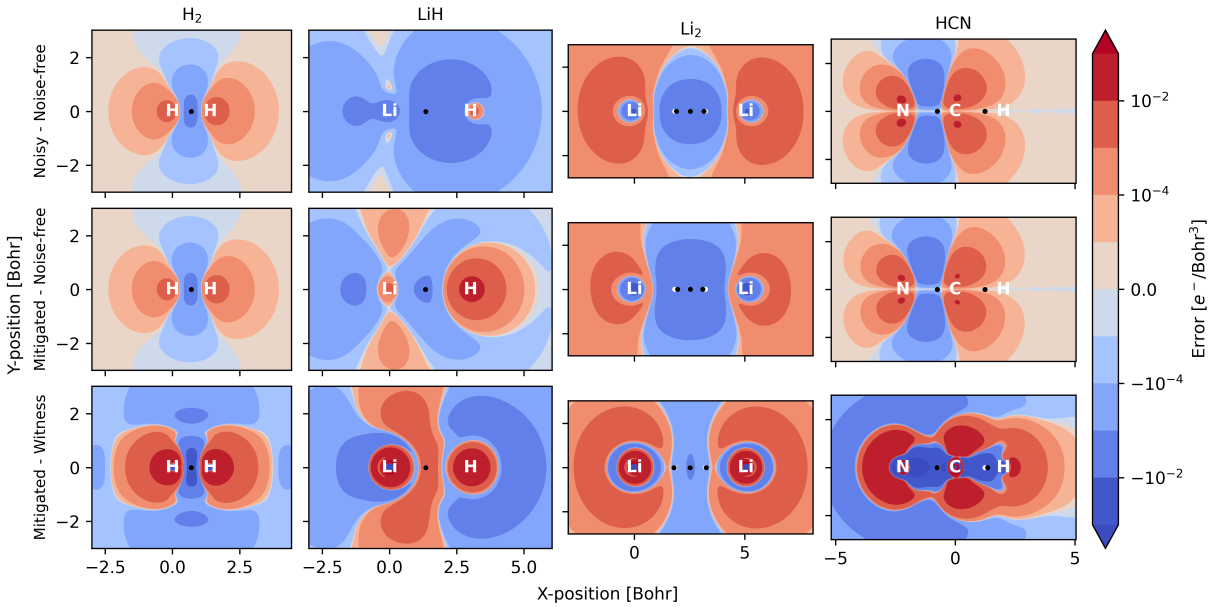


Figure 7.3: Differences in the electron densities between (top) noisy and noise-free (middle) mitigated and noise-free, and (bottom) mitigated and near-exact witness for H_2, LiH, Li_2, and HCN. The figure is reproduced from Paper III.

Paper IV

A quantum computing implementation of nuclearelectronic orbital (NEO) theory: Toward an exact pre-Born–Oppenheimer formulation of molecular quantum systems

Quantum computing brings the promise of larger and more complex chemical simulations, and offers to open up calculations that have previously often been neglected. One such avenue is to relax the restrictions imposed by the BO approximation.

In this paper we explore how such an relaxation could be implemented on near-term quantum hardware. We utilize the NEO framework to include nuclear orbitals in a VQE calculation, showing how the qubit and gate requirements in the circuits can be reduced substantially by looking at the problem symmetries. We additionally implement a parameter transfer strategy, similar to that analyzed in Paper I.

We implement a NEO-UCC algorithm for two molecules: H_2 and malonaldehyde. The results of including varying orders of excitaions in the NEO-UCC for H_2 is shown in Table 7.1

Table 7.1: Ground state energies for H_2, calculated using BO and NEO approaches. The errors decrease as higher order excitations are included. The effect of pure nuclear single and double excitations ($S^{(0,1)}D^{(0,2)}$) is minimal. Table is reproduced from Paper IV.

	Method	Energy/Ha	Error/Ha
	NEOFCI	-1.066121	
	NEOHF	-1.041487	0.024634
NEOUCC	$S^{(0,1)}D^{(0,2)}$	-1.041487	0.024634
	$S^{(1,0)}D^{(2,0)}$	-1.065040	0.001082
	$S^{(1,0)}D^{(2,0)}D^{(1,1)}$	-1.066037	0.000084
	$S^{(1,0)}D^{(2,0)}T^{(2,1)}$	-1.065063	0.001059
	$S^{(1,0)}D^{(2,0)}D^{(1,1)}T^{(2,1)}$	-1.066117	0.000004
	$SD^{(2,0)}D^{(0,2)}$	-1.065049	0.001073
	$SDT^{(2,1)}$	-1.066121	0.000001
	SDT	-1.066121	0.000001

Paper V

Nonadiabatic Nuclear–Electron Dynamics: A Quantum Computing Approach

Quantum dynamics offers an interesting avenue for quantum computing, simulating the time evolution of chemical systems. Time accurate time evolution is computationally expensive to classically perform and is a such a candidate for quantum simulation.

Building on the methodology we developed in Paper IV, we implement a trotterized scheme to perform quantum dynamics and proton transfer in the malonaldehyde molecule. The dynamics is performed both in the adiabatic and nonadiabatic regimes. The effect of trotterization is also studied, and we find that errors owing to poor approximation in the trotterized operator incurs substantial errors in the measured nuclear-electron entanglement.

CHAPTER 8

Conclusion and Outlook

The goal of this thesis has been to investigate how near-term quantum computers can be used to study chemical concepts. Tangential topics such as error mitigation and optimization were discovered and incorporated along the way.

The parameter transfer strategy presented in Paper I builds upon methods from conventional chemistry. The method proved effective and somewhat versatile, with similar approaches being implemented successfully for both PES sampling and for transferring parameters between levels of theory. Application and development of parameter transfer strategies will most likely happen naturally as related problems are found. As quantum computers become more powerful and harder to classically emulate, one can imagine a time where small calculations are carried out in advance, validated with a classical machine, and then used to warm-start a larger quantum calculation.

An additional topic that came up during the work on transferring parameters was the idea of having a measure of similarity between problems. This type of measure would then help predict when and where parameter transfer could be useful, and perhaps explain why. Such a measure would likely draw parallels to the fidelity axioms of Josza, listed in Chapter 5. Such a measure could indicate how much speedup should be expect from a given transfer. Having a value on expected speedup could in turn lead to a sort of hyper optimization where an optimal set of calculations can be planned out such that one feeds into the next. Much like the use of HF to initialize post-HF methods. Further development could help overcome one of the challenges facing the fault-tolerant quantum phase estimation algorithm: the requirement of an initial overlap with the ground state.

As quantum computers become more powerful, and hopefully at some point surpass our ability to simulate them classically, we will enter a period where access to physical quantum computers will become crucial. The current process of validating quantum calculations consist to a large part in comparing to noise- and error-free calculations, performed on classical computers. This type of validation will become infeasible for calculations involving more than 30 or 40 qubits, and outright impossible as we approach 100 qubits [11]. The method proposed in Paper III is an attempt to look ahead, to help gauge the quality of yet to be realized calculations. Future work in this direction should focus on creating a proper measure, such that the values generated can be used to also quantitatively to gauge quality of calculations.

The REM method proved surprisingly efficient at reducing errors in the ground state energy. The ease of use and the low cost also makes it a good choice for an initial mitigation,

before perhaps more powerful, but costlier, methods are implemented. REM also works well when applied together with some other mitigation strategies, such as readout mitigation. A possible extension to the method presented herein, could come in the form of higher level extrapolation and inclusion of more points. A draw back of the current method is that it leaves the shape of the cost landscape unchanged. An higher-order extrapolation technique would additionally allow for corrections of the shape. The points used for extrapolation do not need to be new Clifford states. An option is to use the same reference state but change the noise level, taking inspiration from methods such as Richardson extrapolation.

For future calculations, effort should also be put towards incorporating more realistic noise models. Explicit depolarizing and dephasing noise, as well as readout errors should be used. With experimental data, such as that presented herein, an accurate noise model could most likely be constructed. Having an accurate noise model would in turn allow to get more value out of experiments as one can be better prepared.

Taking a step beyond the BO approximation proved both challenging and rewarding. As the calculations naturally require more resources, much effort was put into reducing these requirements. The non-BO calculations performed in this thesis demonstrated a possible approach to time-evolution on near-term quantum computers. However, the Trotterization method employed produces quantum circuits that are far too deep for current quantum computers. Going forward, it is thus crucial to reduce the circuit depth. If the circuit depth can be sufficiently reduced, quantum time-evolution and NEO dynamics can prove interesting cases for physical NISQ computers.

As the work in this thesis was in part performed at, and as part of, AstraZeneca, many of the discussions where had on the topic of industrial relevance. What needs do the pharmaceutical industry have that are answered by quantum computing? The ideas of non-BO and quantum time evolution in Papers IV and V came from my colleagues at AstraZeneca. Both methods exist today, however, they are often seen as too computationally expensive for most use cases. Quantum computers would hopefully lower the cost of such calculations, making them relevant for industrial use.

When quantum computers are presented to industry, and the public at large, it is often the increased accuracy that is presented, offering to perform previously impossible calculations. However, precision and accuracy are not necessarily what is requested by industry. Value in quantum computing could also be found in performing calculations with current day accuracy, but at reduced cost. This would effectively allow for more parallelism when, e.g., screening large amount of molecules and conformers. This is a particular interesting application for NISQ computers, as they hopefully bridge the gap to fault-tolerant quantum computing. Unfortunately, the crossover point where quantum computers start to outperform classical ones seem to be far into the future [177].

In general, the idea of any computational advantage in the NISQ era is controversial. Quantum computers boasting hundreds of qubits have been developed, yet the largest calculations consist of closer to 20 qubits, and some even using HF as the level of theory [103]. Nevertheless, the fact that classically simulating more than tens of qubits soon becomes impossible still instills hope for quantum usefulness [15].

Bibliography

[1] Andrew M. Shanken. "Better Living: Toward a Cultural History of a Business Slogan". In: *Enterprise & Society* 7.3 (2006), pp. 485–519. ISSN: 1467-2227, 1467-2235. DOI: 10.1093/es/khl001.

[2] J. R. (James Riddick) Partington. *A short history of chemistry*. London, Macmillan; New York, St. Martin's Press, 1960. 442 pp.

[3] Venkatesan S. Thimmakondu et al. "Why an integrated approach between search algorithms and chemical intuition is necessary?" In: *Physical Chemistry Chemical Physics* 24.19 (2022), pp. 11680–11686. DOI: 10.1039/D2CP00315E.

[4] Theresa Sperger, Italo A. Sanhueza, and Franziska Schoenebeck. "Computation and Experiment: A Powerful Combination to Understand and Predict Reactivities". In: *Accounts of Chemical Research* 49.6 (2016), pp. 1311–1319. ISSN: 0001-4842. DOI: 10.1021/acs.accounts.6b00068.

[5] Gordon E. Moore. "Cramming more components onto integrated circuits". In: *Electronics* 38.8 (1965), pp. 114 ff.

[6] R.R. Schaller. "Moore's law: past, present and future". In: *IEEE Spectrum* 34.6 (1997), pp. 52–59. ISSN: 0018-9235. DOI: 10.1109/6.591665.

[7] Scott E. Thompson and Srivatsan Parthasarathy. "Moore's law: the future of Si microelectronics". In: *Materials Today* 9.6 (2006), pp. 20–25. ISSN: 1369-7021. DOI: 10.1016/S1369-7021(06)71539-5.

[8] Greg Yeric. "Moore's law at 50: Are we planning for retirement?" In: *2015 IEEE International Electron Devices Meeting (IEDM)*. 2015 IEEE International Electron Devices Meeting (IEDM). Washington, DC, USA: IEEE, 2015, pp. 1.1.1–1.1.8. ISBN: 978-1-4673-9894-7. DOI: 10.1109/IEDM.2015.7409607.

[9] Richard P. Feynman. "Simulating physics with computers". In: *International Journal of Theoretical Physics* 21.6 (1982), pp. 467–488. ISSN: 0020-7748, 1572-9575. DOI: 10.1007/BF02650179.

[10] Paul Benioff. "The computer as a physical system: A microscopic quantum mechanical Hamiltonian model of computers as represented by Turing machines". In: *Journal of Statistical Physics* 22.5 (1980), pp. 563–591. ISSN: 0022-4715, 1572-9613. DOI: 10.1007/BF01011339.

[11] Frank Arute et al. "Quantum supremacy using a programmable superconducting processor". In: *Nature* 574.7779 (2019), pp. 505–510. ISSN: 0028-0836, 1476-4687. DOI: 10.1038/s41586-019-1666-5.

[12] Leonardo Novo, Juani Bermejo-Vega, and Raúl García-Patrón. "Quantum advantage from energy measurements of many-body quantum systems". In: *Quantum* 5 (2021), p. 465. ISSN: 2521-327X. DOI: 10.22331/q-2021-06-02-465. arXiv: 1912.06608.

[13] Sam McArdle et al. "Quantum computational chemistry". In: *Reviews of Modern Physics* 92.1 (2020), p. 015003. ISSN: 0034-6861, 1539-0756. DOI: 10.1103/RevModPhys.92.015003. arXiv: 1808.10402.

[14] A. Yu Kitaev. *Quantum measurements and the Abelian Stabilizer Problem*. 1995. DOI: 10.48550/arXiv.quant-ph/9511026. arXiv: quant-ph/9511026.

[15] John Preskill. "Quantum Computing in the NISQ era and beyond". In: *Quantum* 2 (2018), p. 79. ISSN: 2521-327X. DOI: 10.22331/q-2018-08-06-79. arXiv: 1801.00862.

[16] Michael A. Nielsen and Isaac L. Chuang. *Quantum computation and quantum information*. 10th anniversary ed. Cambridge ; New York: Cambridge University Press, 2010. 676 pp. ISBN: 978-1-107-00217-3.

[17] David Deutsch and Richard Jozsa. "Rapid solution of problems by quantum computation". In: *Proceedings of the Royal Society of London. Series A: Mathematical and Physical Sciences* 439.1907 (1992), pp. 553–558. DOI: 10.1098/rspa.1992.0167.

[18] P.W. Shor. "Algorithms for quantum computation: discrete logarithms and factoring". In: *Proceedings 35th Annual Symposium on Foundations of Computer Science*. Proceedings 35th Annual Symposium on Foundations of Computer Science. 1994, pp. 124–134. DOI: 10.1109/SFCS.1994.365700.

[19] Peter W. Shor. "Polynomial-Time Algorithms for Prime Factorization and Discrete Logarithms on a Quantum Computer". In: *SIAM Journal on Computing* 26.5 (1997), pp. 1484–1509. ISSN: 0097-5397, 1095-7111. DOI: 10.1137/S0097539795293172. arXiv: quant-ph/9508027.

[20] Tim Maudlin. "Philosophy of Physics: Quantum Theory". In: *Philosophy of Physics*. Princeton University Press, 2019. ISBN: 978-0-691-19067-9. DOI: 10.1515/9780691190679.

[21] Paul Adrien Maurice Dirac and Ralph Howard Fowler. "Quantum mechanics of many-electron systems". In: *Proceedings of the Royal Society of London. Series A, Containing Papers of a Mathematical and Physical Character* 123.792 (1929), pp. 714–733. DOI: 10.1098/rspa.1929.0094.

[22] Jun J. Sakurai and Jim Napolitano. *Modern quantum mechanics*. Second edition. Cambridge: Cambridge University Press, 2017. 550 pp. ISBN: 978-1-108-42241-3. DOI: 10.1017/9781108499996.

[23] Trygve Helgaker, Poul Jørgensen, and Jeppe Olsen. *Molecular electronic-structure theory*. Repr. as paperback ed. Chichester Weinheim: Wiley, 2012. 908 pp. ISBN: 978-1-118-53147-1.

[24] Tosio Kato. "On the eigenfunctions of many-particle systems in quantum mechanics". In: *Communications on Pure and Applied Mathematics* 10.2 (1957), pp. 151–177. ISSN: 1097-0312. DOI: 10.1002/cpa.3160100201.

[25] J. C. Slater. "Atomic Shielding Constants". In: *Physical Review* 36.1 (1930), pp. 57–64. ISSN: 0031-899X. DOI: 10.1103/PhysRev.36.57.

[26] Thom H. Dunning Jr. "Gaussian basis sets for use in correlated molecular calculations. I. The atoms boron through neon and hydrogen". In: *The Journal of Chemical Physics* 90.2 (1989), pp. 1007–1023. ISSN: 0021-9606. DOI: 10.1063/1.456153.

[27] B. H. Bransden and C. J. Joachain. *Physics of atoms and molecules*. 2nd ed. Harlow, England ; New York: Prentice Hall, 2003. 1114 pp. ISBN: 978-0-582-35692-4.

[28] Jakob S. Kottmann et al. "Reducing Qubit Requirements while Maintaining Numerical Precision for the Variational Quantum Eigensolver: A Basis-Set-Free Approach". In: *The Journal of Physical Chemistry Letters* 12.1 (2021), pp. 663–673. ISSN: 1948-7185, 1948-7185. DOI: 10.1021/acs.jpclett.0c03410.

[29] Yudong Cao et al. "Quantum Chemistry in the Age of Quantum Computing". In: *Chemical Reviews* 119.19 (2019), pp. 10856–10915. ISSN: 0009-2665, 1520-6890. DOI: 10.1021/acs.chemrev.8b00803.

[30] Colin P. Williams. *Explorations in quantum computing*. 2. ed. Texts in computer science. London Heidelberg: Springer, 2011. 717 pp. ISBN: 978-1-84628-886-9.

[31] Lov K. Grover. "A fast quantum mechanical algorithm for database search". In: *Proceedings of the twenty-eighth annual ACM symposium on Theory of computing - STOC '96*. the twenty-eighth annual ACM symposium. Philadelphia, Pennsylvania, United States: ACM Press, 1996, pp. 212–219. ISBN: 978-0-89791-785-8. DOI: 10.1145/237814.237866.

[32] Alan Mathison Turing. *The Essential Turing*. Oxford University Press, 2004. 622 pp. ISBN: 978-0-19-825079-1.

[33] Claude Elwood Shannon and Aaron D. Wyner. *Collected papers*. Ed. by Neil J. A. Sloane. Piscataway, NJ: IEEE Press, 1993. 924 pp. ISBN: 978-0-7803-0434-5.

[34] F. Bloch. "Nuclear Induction". In: *Physical Review* 70.7 (1946), pp. 460–474. ISSN: 0031-899X. DOI: 10.1103/PhysRev.70.460.

[35] Expert Panel. "Council Post: 15 Significant Ways Quantum Computing Could Soon Impact Society". In: *Forbes* (2023).

[36] "The Rise of Quantum Computing in Gaming". In: *PC Tech* (2023).

[37] Ali El Kaafarani. "Council Post: Four Ways Quantum Computing Could Change The World". In: *Forbes* (2021).

[38] Dan Hurley. "The Quantum Internet Will Blow Your Mind. Here's What It Will Look Like". In: *Discover Magazine* (2020).

[39] I. M. Georgescu, S. Ashhab, and Franco Nori. "Quantum simulation". In: *Reviews of Modern Physics* 86.1 (2014), pp. 153–185. ISSN: 0034-6861, 1539-0756. DOI: 10.1103/RevModPhys.86.153.

[40] David Elieser Deutsch and Roger Penrose. "Quantum computational networks". In: *Proceedings of the Royal Society of London. A. Mathematical and Physical Sciences* 425.1868 (1997), pp. 73–90. DOI: 10.1098/rspa.1989.0099.

[41] David P. DiVincenzo. "The Physical Implementation of Quantum Computation". In: *Fortschritte der Physik* 48.9 (2000), pp. 771–783. ISSN: 1521-3978. DOI: 10.1002/1521-3978(200009)48:9/11<771::AID-PROP771>3.0.CO;2-E.

[42] H. J. Briegel et al. “Measurement-based quantum computation”. In: *Nature Physics* 5.1 (2009), pp. 19–26. ISSN: 1745-2481. DOI: 10.1038/nphys1157.

[43] Seth Lloyd and Samuel L Braunstein. “Quantum Computation over Continuous Variables”. In: *PHYSICAL REVIEW LETTERS* 82.8 (1999).

[44] Tameem Albash and Daniel A. Lidar. “Adiabatic quantum computation”. In: *Reviews of Modern Physics* 90.1 (2018), p. 015002. ISSN: 0034-6861, 1539-0756. DOI: 10.1103/RevModPhys.90.015002.

[45] Alberto Peruzzo et al. “A variational eigenvalue solver on a quantum processor”. In: *Nature Communications* 5.1 (2014), p. 4213. ISSN: 2041-1723. DOI: 10.1038/ncomms5213. arXiv: 1304.3061.

[46] Vincent E. Elfving et al. “Simulating quantum chemistry in the seniority-zero space on qubit-based quantum computers”. In: *Physical Review A* 103.3 (2021), p. 032605. DOI: 10.1103/PhysRevA.103.032605.

[47] Abhinav Anand et al. “A quantum computing view on unitary coupled cluster theory”. In: *Chemical Society Reviews* 51.5 (2022), pp. 1659–1684. ISSN: 1460-4744. DOI: 10.1039/D1CS00932J.

[48] P. Jordan and E. Wigner. “Über das Paulische Äquivalenzverbot”. In: *Zeitschrift für Physik* 47.9 (1928), pp. 631–651. ISSN: 0044-3328. DOI: 10.1007/BF01331938.

[49] R. Somma et al. “Simulating physical phenomena by quantum networks”. In: *Physical Review A* 65.4 (2002), p. 042323. ISSN: 1050-2947, 1094-1622. DOI: 10.1103/PhysRevA.65.042323.

[50] Jacob T. Seeley, Martin J. Richard, and Peter J. Love. “The Bravyi-Kitaev transformation for quantum computation of electronic structure”. In: *The Journal of Chemical Physics* 137.22 (2012), p. 224109. ISSN: 0021-9606, 1089-7690. DOI: 10.1063/1.4768229. arXiv: 1208.5986.

[51] Andrew Tranter et al. “A Comparison of the Bravyi–Kitaev and Jordan–Wigner Transformations for the Quantum Simulation of Quantum Chemistry”. In: *Journal of Chemical Theory and Computation* 14.11 (2018), pp. 5617–5630. ISSN: 1549-9618. DOI: 10.1021/acs.jctc.8b00450.

[52] Abhinav Kandala et al. “Hardware-efficient Variational Quantum Eigensolver for Small Molecules and Quantum Magnets”. In: *Nature* 549.7671 (2017), pp. 242–246. ISSN: 0028-0836, 1476-4687. DOI: 10.1038/nature23879. arXiv: 1704.05018.

[53] Sergey Bravyi et al. “Tapering off qubits to simulate fermionic Hamiltonians”. In: *arXiv:1701.08213 [quant-ph]* (2017). arXiv: 1701.08213.

[54] Andreas Köhn and Jeppe Olsen. “Capabilities and limits of the unitary coupled-cluster approach with generalized two-body cluster operators”. In: *The Journal of Chemical Physics* 157.12 (2022), p. 124110. ISSN: 0021-9606. DOI: 10.1063/5.0104815.

[55] Harper R. Grimsley et al. “An adaptive variational algorithm for exact molecular simulations on a quantum computer”. In: *Nature Communications* 10.1 (2019), p. 3007. ISSN: 2041-1723. DOI: 10.1038/s41467-019-10988-2.

[56] Ho Lun Tang et al. "Qubit-ADAPT-VQE: An Adaptive Algorithm for Constructing Hardware-Efficient Ansätze on a Quantum Processor". In: *PRX Quantum* 2.2 (2021), p. 020310. ISSN: 2691-3399. DOI: 10.1103/PRXQuantum.2.020310.

[57] Panagiotis G. Anastasiou et al. *TETRIS-ADAPT-VQE: An adaptive algorithm that yields shallower, denser circuit ans\"atze*. 2022. DOI: 10.48550/arXiv.2209.10562. arXiv: 2209.10562[quant-ph].

[58] Eric R. Anschuetz and Bobak T. Kiani. "Quantum variational algorithms are swamped with traps". In: *Nature Communications* 13.1 (2022), p. 7760. ISSN: 2041-1723. DOI: 10.1038/s41467-022-35364-5.

[59] Hugo Touvron et al. *LLaMA: Open and Efficient Foundation Language Models*. 2023. arXiv: 2302.13971[cs].

[60] Reed Albergotti. "The secret history of Elon Musk, Sam Altman, and OpenAI". In: *Semafor* (2023).

[61] Xiaozhen Ge, Rebing Wu, and Herschel Rabitz. "Optimization Landscape of Quantum Control Systems". In: *Complex System Modeling and Simulation* 1.2 (2021), pp. 77–90. ISSN: 2096-9929. DOI: 10.23919/CSMS.2021.0014.

[62] Andrew Arrasmith et al. "Effect of barren plateaus on gradient-free optimization". In: *Quantum* 5 (2021), p. 558. DOI: 10.22331/q-2021-10-05-558.

[63] Gustavo G. Rondina and Juarez L. F. Da Silva. "Revised Basin-Hopping Monte Carlo Algorithm for Structure Optimization of Clusters and Nanoparticles". In: *Journal of Chemical Information and Modeling* 53.9 (2013), pp. 2282–2298. ISSN: 1549-9596. DOI: 10.1021/ci400224z.

[64] David J. Wales and Jonathan P. K. Doye. "Global Optimization by Basin-Hopping and the Lowest Energy Structures of Lennard-Jones Clusters Containing up to 110 Atoms". In: *The Journal of Physical Chemistry A* 101.28 (1997), pp. 5111–5116. ISSN: 1089-5639. DOI: 10.1021/jp970984n.

[65] Lennart Bittel and Martin Kliesch. "Training Variational Quantum Algorithms Is NP-Hard". In: *Physical Review Letters* 127.12 (2021), p. 120502. ISSN: 0031-9007, 1079-7114. DOI: 10.1103/PhysRevLett.127.120502.

[66] Samson Wang et al. "Noise-induced barren plateaus in variational quantum algorithms". In: *Nature Communications* 12.1 (2021), p. 6961. ISSN: 2041-1723. DOI: 10.1038/s41467-021-27045-6.

[67] Jarrod R. McClean et al. "Barren plateaus in quantum neural network training landscapes". In: *Nature Communications* 9.1 (2018), p. 4812. ISSN: 2041-1723. DOI: 10.1038/s41467-018-07090-4.

[68] Zoë Holmes et al. "Connecting Ansatz Expressibility to Gradient Magnitudes and Barren Plateaus". In: *PRX Quantum* 3.1 (2022), p. 010313. DOI: 10.1103/PRXQuantum.3.010313.

[69] Lorenzo Leone et al. *On the practical usefulness of the Hardware Efficient Ansatz*. 2022. arXiv: 2211.01477[quant-ph].

[70] David Newton, Farzad Yousefian, and Raghu Pasupathy. "Stochastic Gradient Descent: Recent Trends". In: *Recent Advances in Optimization and Modeling of Contemporary Problems*. INFORMS TutORials in Operations Research. INFORMS, 2018, pp. 193–220. ISBN: 978-0-9906153-2-3. DOI: 10.1287/educ.2018.0191.

[71] Juan C. Meza. "Steepest descent". In: *WIREs Computational Statistics* 2.6 (2010), pp. 719–722. ISSN: 1939-0068. DOI: 10.1002/wics.117.

[72] William W. Hager, D. W. Hearn, and Panos M. Pardalos. *Large Scale Optimization: State of the Art*. Springer Science & Business Media, 2013. 470 pp. ISBN: 978-1-4613-3632-7.

[73] Ran Cheng and Yaochu Jin. "A Competitive Swarm Optimizer for Large Scale Optimization". In: *IEEE Transactions on Cybernetics* 45.2 (2015), pp. 191–204. ISSN: 2168-2275. DOI: 10.1109/TCYB.2014.2322602.

[74] R. H. Byrd et al. "A Stochastic Quasi-Newton Method for Large-Scale Optimization". In: *SIAM Journal on Optimization* 26.2 (2016), pp. 1008–1031. ISSN: 1052-6234. DOI: 10.1137/140954362.

[75] Diederik P. Kingma and Jimmy Ba. *Adam: A Method for Stochastic Optimization*. 2017. DOI: 10.48550/arXiv.1412.6980. arXiv: 1412.6980[cs].

[76] Ryan Sweke et al. "Stochastic gradient descent for hybrid quantum-classical optimization". In: *arXiv:1910.01155 [quant-ph]* (2020). arXiv: 1910.01155.

[77] Léon Bottou. "Stochastic Gradient Descent Tricks". In: *Neural Networks: Tricks of the Trade: Second Edition*. Ed. by Grégoire Montavon, Geneviève B. Orr, and Klaus-Robert Müller. Lecture Notes in Computer Science. Berlin, Heidelberg: Springer, 2012, pp. 421–436. ISBN: 978-3-642-35289-8. DOI: 10.1007/978-3-642-35289-8_25.

[78] J. A. Gidi et al. "Stochastic optimization algorithms for quantum applications". In: *arXiv:2203.06044 [quant-ph]* (2022). arXiv: 2203.06044.

[79] J. Kennedy and R. Eberhart. "Particle swarm optimization". In: *Proceedings of ICNN'95 - International Conference on Neural Networks*. Proceedings of ICNN'95 - International Conference on Neural Networks. Vol. 4. 1995, 1942–1948 vol.4. DOI: 10.1109/ICNN.1995.488968.

[80] Riccardo Poli, James Kennedy, and Tim Blackwell. "Particle swarm optimization: An overview". In: *Swarm Intelligence* 1.1 (2007), pp. 33–57. ISSN: 1935-3820. DOI: 10.1007/s11721-007-0002-0.

[81] A. R. Conn, Katya Scheinberg, and Luis N. Vicente. *Introduction to derivative-free optimization*. MPS-SIAM series on optimization 8. Philadelphia: Society for Industrial and Applied Mathematics/Mathematical Programming Society, 2009. 277 pp. ISBN: 978-0-89871-668-9.

[82] J.C. Spall. "Multivariate stochastic approximation using a simultaneous perturbation gradient approximation". In: *IEEE Transactions on Automatic Control* 37.3 (1992), pp. 332–341. ISSN: 1558-2523. DOI: 10.1109/9.119632.

[83] Daniel Stilck França and Raul García-Patrón. "Limitations of optimization algorithms on noisy quantum devices". In: *Nature Physics* 17.11 (2021), pp. 1221–1227. ISSN: 1745-2481. DOI: 10.1038/s41567-021-01356-3.

[84] A. V. Uvarov and J. D. Biamonte. "On barren plateaus and cost function locality in variational quantum algorithms". In: *Journal of Physics A: Mathematical and Theoretical* 54.24 (2021), p. 245301. ISSN: 1751-8121. DOI: 10.1088/1751-8121/abfac7.

[85] M. Cerezo et al. "Cost function dependent barren plateaus in shallow parametrized quantum circuits". In: *Nature Communications* 12.1 (2021), p. 1791. ISSN: 2041-1723. DOI: 10.1038/s41467-021-21728-w.

[86] Carlos Ortiz Marrero, Mária Kieferová, and Nathan Wiebe. "Entanglement-Induced Barren Plateaus". In: *PRX Quantum* 2.4 (2021), p. 040316. DOI: 10.1103/PRXQuantum.2.040316.

[87] Michael Ragone et al. *A Unified Theory of Barren Plateaus for Deep Parametrized Quantum Circuits*. 2023. arXiv: 2309.09342[quant-ph].

[88] Enrico Fontana et al. *The Adjoint Is All You Need: Characterizing Barren Plateaus in Quantum Ans\"atze*. 2023. DOI: 10.48550/arXiv.2309.07902. arXiv: 2309.07902[quant-ph].

[89] Meenal V. Narkhede, Prashant P. Bartakke, and Mukul S. Sutaone. "A review on weight initialization strategies for neural networks". In: *Artificial Intelligence Review* 55.1 (2021), pp. 291–322. ISSN: 1573-7462. DOI: 10.1007/s10462-021-10033-z.

[90] Leo Zhou et al. "Quantum Approximate Optimization Algorithm: Performance, Mechanism, and Implementation on Near-Term Devices". In: *Physical Review X* 10.2 (2020), p. 021067. DOI: 10.1103/PhysRevX.10.021067.

[91] Edward Grant et al. "An initialization strategy for addressing barren plateaus in parametrized quantum circuits". In: *Quantum* 3 (2019), p. 214. DOI: 10.22331/q-2019-12-09-214.

[92] Huan-Yu Liu et al. "Mitigating barren plateaus with transfer-learning-inspired parameter initializations". In: *New Journal of Physics* 25.1 (2023), p. 013039. ISSN: 1367-2630. DOI: 10.1088/1367-2630/acb58e.

[93] Wen Wei Ho and Timothy H. Hsieh. "Efficient variational simulation of non-trivial quantum states". In: *SciPost Physics* 6.3 (2019), p. 029. ISSN: 2542-4653. DOI: 10.21468/SciPostPhys.6.3.029.

[94] Tyler Volkoff and Patrick J. Coles. "Large gradients via correlation in random parameterized quantum circuits". In: *Quantum Science and Technology* 6.2 (2021), p. 025008. ISSN: 2058-9565. DOI: 10.1088/2058-9565/abd891.

[95] M. Cerezo et al. "Variational quantum algorithms". In: *Nature Reviews Physics* 3.9 (2021), pp. 625–644. ISSN: 2522-5820. DOI: 10.1038/s42254-021-00348-9.

[96] Alexey Galda et al. "Transferability of optimal QAOA parameters between random graphs". In: *arXiv:2106.07531 [quant-ph]* (2021). arXiv: 2106.07531.

[97] Ruslan Shaydulin et al. "Parameter Transfer for Quantum Approximate Optimization of Weighted MaxCut". In: *arXiv:2201.11785 [quant-ph]* (2022). arXiv: 2201.11785.

[98] Antonio A. Mele et al. "Avoiding barren plateaus via transferability of smooth solutions in a Hamiltonian variational ansatz". In: *Physical Review A* 106.6 (2022), p. L060401. ISSN: 2469-9926, 2469-9934. DOI: 10.1103/PhysRevA.106.L060401.

[99] Karl Weiss, Taghi M. Khoshgoftaar, and DingDing Wang. "A survey of transfer learning". In: *Journal of Big Data* 3.1 (2016), p. 9. ISSN: 2196-1115. DOI: 10.1186/s40537-016-0043-6.

[100] Fuzhen Zhuang et al. "A Comprehensive Survey on Transfer Learning". In: *Proceedings of the IEEE* 109.1 (2021), pp. 43–76. ISSN: 0018-9219, 1558-2256. DOI: 10.1109/JPROC.2020.3004555.

[101] A. A. Clerk et al. "Introduction to quantum noise, measurement, and amplification". In: *Reviews of Modern Physics* 82.2 (2010), pp. 1155–1208. DOI: 10.1103/RevModPhys.82.1155.

[102] Sitan Chen et al. "The complexity of NISQ". In: *Nature Communications* 14.1 (2023), p. 6001. ISSN: 2041-1723. DOI: 10.1038/s41467-023-41217-6.

[103] Google AI Quantum and Collaborators*† et al. "Hartree-Fock on a superconducting qubit quantum computer". In: *Science* 369.6507 (2020), pp. 1084–1089. ISSN: 0036-8075, 1095-9203. DOI: 10.1126/science.abb9811.

[104] Lewis Fry Richardson and Richard Tetley Glazebrook. "IX. The approximate arithmetical solution by finite differences of physical problems involving differential equations, with an application to the stresses in a masonry dam". In: *Philosophical Transactions of the Royal Society of London. Series A, Containing Papers of a Mathematical or Physical Character* 210.459 (1911), pp. 307–357. DOI: 10.1098/rsta.1911.0009.

[105] R. W. Hamming. "Error detecting and error correcting codes". In: *The Bell System Technical Journal* 29.2 (1950), pp. 147–160. ISSN: 0005-8580. DOI: 10.1002/j.1538-7305.1950.tb00463.x.

[106] Easwar Magesan et al. "Efficient Measurement of Quantum Gate Error by Interleaved Randomized Benchmarking". In: *Physical Review Letters* 109.8 (2012), p. 080505. DOI: 10.1103/PhysRevLett.109.080505.

[107] Kieu Quang Tuan, Hung Q. Nguyen, and Le Bin Ho. "Direct state measurements under state-preparation-and-measurement errors". In: *Quantum Information Processing* 20.6 (2021), p. 197. ISSN: 1573-1332. DOI: 10.1007/s11128-021-03144-7.

[108] Christopher Jackson and S. J. van Enk. "Detecting correlated errors in state-preparation-and-measurement tomography". In: *Physical Review A* 92.4 (2015), p. 042312. DOI: 10.1103/PhysRevA.92.042312.

[109] Matt McEwen et al. "Resolving catastrophic error bursts from cosmic rays in large arrays of superconducting qubits". In: *Nature Physics* 18.1 (2022), pp. 107–111. ISSN: 1745-2481. DOI: 10.1038/s41567-021-01432-8.

[110] Yasunari Suzuki et al. "Q3DE: A fault-tolerant quantum computer architecture for multi-bit burst errors by cosmic rays". In: *2022 55th IEEE/ACM International Symposium on Microarchitecture (MICRO)*. 2022 55th IEEE/ACM International Symposium on Microarchitecture (MICRO). 2022, pp. 1110–1125. DOI: 10.1109/MICRO56248.2022.00079.

[111] Trygve Helgaker et al. "A priori calculation of molecular properties to chemical accuracy". In: *Journal of Physical Organic Chemistry* 17.11 (2004), pp. 913–933. ISSN: 1099-1395. DOI: 10.1002/poc.841.

[112] Kirk A. Peterson, David Feller, and David A. Dixon. "Chemical accuracy in ab initio thermochemistry and spectroscopy: current strategies and future challenges". In: *Theoretical Chemistry Accounts* 131.1 (2012), p. 1079. ISSN: 1432-2234. DOI: 10.1007/s00214-011-1079-5.

[113] Lars Tornberg. *Superconducting qubits: measurement, entanglement, and noise*. Doktorsavhandlingar vid Chalmers Tekniska Högskola N.S., 2960. Göteborg: Chalmers Univ. of Technology, 2009. 175 pp. ISBN: 978-91-7385-279-1.

[114] Easwar Magesan et al. "Modeling quantum noise for efficient testing of fault-tolerant circuits". In: *Physical Review A* 87.1 (2013), p. 012324. ISSN: 1050-2947, 1094-1622. DOI: 10.1103/PhysRevA.87.012324.

[115] Ying Li and Simon C. Benjamin. "Efficient Variational Quantum Simulator Incorporating Active Error Minimization". In: *Physical Review X* 7.2 (2017), p. 021050. DOI: 10.1103/PhysRevX.7.021050.

[116] Kristan Temme, Sergey Bravyi, and Jay M. Gambetta. "Error Mitigation for Short-Depth Quantum Circuits". In: *Physical Review Letters* 119.18 (2017), p. 180509. DOI: 10.1103/PhysRevLett.119.180509.

[117] Jarrod R. McClean et al. "Decoding quantum errors with subspace expansions". In: *Nature Communications* 11.1 (2020), p. 636. ISSN: 2041-1723. DOI: 10.1038/s41467-020-14341-w.

[118] Tyler Takeshita et al. "Increasing the Representation Accuracy of Quantum Simulations of Chemistry without Extra Quantum Resources". In: *Physical Review X* 10.1 (2020), p. 011004. DOI: 10.1103/PhysRevX.10.011004.

[119] Zhenyu Cai et al. *Quantum Error Mitigation*. 2022. arXiv: 2210.00921 [quant-ph].

[120] Suguru Endo, Simon C. Benjamin, and Ying Li. "Practical Quantum Error Mitigation for Near-Future Applications". In: *Physical Review X* 8.3 (2018), p. 031027. DOI: 10.1103/PhysRevX.8.031027.

[121] Ryuji Takagi et al. "Fundamental limits of quantum error mitigation". In: *npj Quantum Information* 8.1 (2022), pp. 1–11. ISSN: 2056-6387. DOI: 10.1038/s41534-022-00618-z.

[122] Seth T. Merkel et al. "Self-consistent quantum process tomography". In: *Physical Review A* 87.6 (2013), p. 062119. ISSN: 1050-2947, 1094-1622. DOI: 10.1103/PhysRevA.87.062119.

[123] Junan Lin et al. "Independent state and measurement characterization for quantum computers". In: *Physical Review Research* 3.3 (2021), p. 033285. DOI: 10.1103/PhysRevResearch.3.033285.

[124] G. D'Agostini. "A multidimensional unfolding method based on Bayes' theorem". In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 362.2 (1995), pp. 487–498. ISSN: 0168-9002. DOI: 10.1016/0168-9002(95)00274-X.

[125] Daniel Gottesman. *The Heisenberg Representation of Quantum Computers*. 1998. DOI: 10.48550/arXiv.quant-ph/9807006. arXiv: quant-ph/9807006.

[126] Piotr Czarnik et al. "Error mitigation with Clifford quantum-circuit data". In: *Quantum* 5 (2021), p. 592. DOI: 10.22331/q-2021-11-26-592.

[127] Armands Strikis et al. "Learning-Based Quantum Error Mitigation". In: *PRX Quantum* 2.4 (2021), p. 040330. ISSN: 2691-3399. DOI: 10.1103/PRXQuantum.2.040330.

[128] Frank Arute et al. *Observation of separated dynamics of charge and spin in the Fermi-Hubbard model*. 2020. DOI: 10.48550/arXiv.2010.07965. arXiv: 2010.07965[quant-ph].

[129] Ashley Montanaro and Stasja Stanisic. *Error mitigation by training with fermionic linear optics*. 2021. DOI: 10.48550/arXiv.2102.02120. arXiv: 2102.02120[quant-ph].

[130] Stasja Stanisic et al. "Observing ground-state properties of the Fermi-Hubbard model using a scalable algorithm on a quantum computer". In: *Nature Communications* 13.1 (2022), p. 5743. ISSN: 2041-1723. DOI: 10.1038/s41467-022-33335-4.

[131] Richard P. Feynman. "Seeking New Laws (The Character of Physical Law)". Lecture. Lecture. Cornell University, 1964.

[132] Yeong-Cherng Liang et al. "Quantum fidelity measures for mixed states". In: *Reports on Progress in Physics* 82.7 (2019), p. 076001. ISSN: 0034-4885, 1361-6633. DOI: 10.1088/1361-6633/ab1ca4.

[133] Richard Jozsa. "Fidelity for Mixed Quantum States". In: *Journal of Modern Optics* 41.12 (1994), pp. 2315–2323. DOI: 10.1080/09500349414552171.

[134] Shahnawaz Ahmed et al. "Quantum State Tomography with Conditional Generative Adversarial Networks". In: *arXiv:2008.03240 [quant-ph]* (2020). arXiv: 2008.03240.

[135] G. Mauro D'Ariano, Matteo G.A. Paris, and Massimiliano F. Sacchi. "Quantum Tomography". In: *Advances in Imaging and Electron Physics*. Vol. 128. Elsevier, 2003, pp. 205–308. ISBN: 978-0-12-014770-0. DOI: 10.1016/S1076-5670(03)80065-4.

[136] Trygve Helgaker, Wim Klopper, and David P. Tew. "Quantitative quantum chemistry". In: *Molecular Physics* 106.16 (2008), pp. 2107–2143. ISSN: 0026-8976. DOI: 10.1080/00268970802258591.

[137] M. Gluza et al. "Fidelity Witnesses for Fermionic Quantum Simulations". In: *Physical Review Letters* 120.19 (2018), p. 190501. ISSN: 0031-9007, 1079-7114. DOI: 10.1103/PhysRevLett.120.190501.

[138] R. P. Feynman. "Forces in Molecules". In: *Physical Review* 56.4 (1939), pp. 340–343. ISSN: 0031-899X. DOI: 10.1103/PhysRev.56.340.

[139] Jannik C. Meyer et al. "Experimental analysis of charge redistribution due to chemical bonding by high-resolution transmission electron microscopy". In: *Nature Materials* 10.3 (2011), pp. 209–215. ISSN: 1476-4660. DOI: 10.1038/nmat2941.

[140] Leo Gross et al. "The Chemical Structure of a Molecule Resolved by Atomic Force Microscopy". In: *Science* 325.5944 (2009), pp. 1110–1114. DOI: 10.1126/science.1176210.

[141] P. L. Giunta and S. P. Kelty. "Direct observation of graphite layer edge states by scanning tunneling microscopy". In: *The Journal of Chemical Physics* 114.4 (2001), pp. 1807–1812. ISSN: 0021-9606. DOI: 10.1063/1.1334349.

[142] N. K. Hansen and P. Coppens. "Testing aspherical atom refinements on small-molecule data sets". In: *Acta Crystallographica Section A: Crystal Physics, Diffraction, Theoretical and General Crystallography* 34.6 (1978), pp. 909–921. ISSN: 0567-7394. DOI: 10.1107/S0567739478001886.

[143] R F W Bader. "Atoms in molecules". In: *Accounts of Chemical Research* 18 (1985).

[144] R.F.W. Bader and R.F. Bader. *Atoms in Molecules: A Quantum Theory*. International series of monographs on chemistry. Clarendon Press, 1990. ISBN: 978-0-19-855168-3.

[145] W.L. Cao et al. "On the presence of non-nuclear attractors in the charge distributions of Li and Na clusters". In: *Chemical Physics Letters* 141.5 (1987), pp. 380–385. ISSN: 00092614. DOI: 10.1016/0009-2614(87)85044-3.

[146] Sharon Hammes-Schiffer. "Theoretical perspectives on non-Born–Oppenheimer effects in chemistry". In: *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences* 380.2223 (2022), p. 20200377. DOI: 10.1098/rsta.2020.0377.

[147] Jennifer A. Johnson. "Populating the periodic table: Nucleosynthesis of the elements". In: *Science* 363.6426 (2019), pp. 474–478. DOI: 10.1126/science.aau9540.

[148] Fabijan Pavoševi , Tanner Culpitt, and Sharon Hammes-Schiffer. "Multicomponent Quantum Chemistry: Integrating Electronic and Nuclear Quantum Effects via the Nuclear–Electronic Orbital Method". In: *Chemical Reviews* 120.9 (2020), pp. 4222–4253. ISSN: 0009-2665. DOI: 10.1021/acs.chemrev.9b00798.

[149] David R. Weinberg et al. "Proton-Coupled Electron Transfer". In: *Chemical Reviews* 112.7 (2012), pp. 4016–4093. ISSN: 0009-2665. DOI: 10.1021/cr200177j.

[150] Jeffrey J. Warren, Tristan A. Tronic, and James M. Mayer. "Thermochemistry of Proton-Coupled Electron Transfer Reagents and its Implications". In: *Chemical Reviews* 110.12 (2010), pp. 6961–7001. ISSN: 0009-2665. DOI: 10.1021/cr100085k.

[151] M. Born and R. Oppenheimer. "Zur Quantentheorie der Molekeln". In: *Annalen der Physik* 389.20 (1927), pp. 457–484. ISSN: 00033804, 15213889. DOI: 10.1002/andp.19273892002.

[152] Eite Tiesinga et al. "CODATA Recommended Values of the Fundamental Physical Constants: 2018*". In: *Journal of Physical and Chemical Reference Data* 50.3 (2021), p. 033105. ISSN: 0047-2689. DOI: 10.1063/5.0064853.

[153] Brian Sutcliffe. "To what question is the clamped-nuclei electronic potential the answer?" In: *Theoretical Chemistry Accounts* 127.3 (2010), pp. 121–131. ISSN: 1432-2234. DOI: 10.1007/s00214-009-0594-0.

[154] William H. Miller. "Classical S Matrix: Numerical Application to Inelastic Collisions". In: *The Journal of Chemical Physics* 53.9 (1970), pp. 3578–3587. ISSN: 0021-9606, 1089-7690. DOI: 10.1063/1.1674535.

[155] Ivano Tavernelli *, Ute F. Röhrig, and Ursula Rothlisberger. "Molecular dynamics in electronically excited states using time-dependent density functional theory". In: *Molecular Physics* 103.6 (2005), pp. 963–981. ISSN: 0026-8976. DOI: 10.1080/00268970512331339378.

[156] Luning Zhao et al. "Nuclear–electronic orbital Ehrenfest dynamics". In: *The Journal of Chemical Physics* 153.22 (2020), p. 224111. ISSN: 0021-9606. DOI: 10.1063/5.0031019.

[157] John C. Tully. "Molecular dynamics with electronic transitions". In: *The Journal of Chemical Physics* 93.2 (1990), pp. 1061–1071. ISSN: 0021-9606. DOI: 10.1063/1.459170.

[158] Tyler D. Hone, Peter J. Rossky, and Gregory A. Voth. "A comparative study of imaginary time path integral based methods for quantum dynamics". In: *The Journal of Chemical Physics* 124.15 (2006), p. 154103. ISSN: 0021-9606. DOI: 10.1063/1.2186636.

[159] David R. Yarkony. "Nonadiabatic Quantum Chemistry—Past, Present, and Future". In: *Chemical Reviews* 112.1 (2012), pp. 481–498. ISSN: 0009-2665. DOI: 10.1021/cr2001299.

[160] Roland Mitri et al. "Ab initio nonadiabatic dynamics study of ultrafast radiationless decay over conical intersections illustrated on the Na3F cluster". In: *The Journal of Chemical Physics* 125.2 (2006), p. 024303. ISSN: 0021-9606. DOI: 10.1063/1.2209233.

[161] Volkhard May and Oliver Kühn. *Charge and energy transfer dynamics in molecular systems*. Fourth edition. Weinheim, Germany: Wiley-VCH, 2023. 522 pp. ISBN: 978-3-527-69627-7.

[162] Jeffrey R. Reimers et al. "Non-adiabatic effects in thermochemistry, spectroscopy and kinetics: the general importance of all three Born–Oppenheimer breakdown corrections". In: *Physical Chemistry Chemical Physics* 17.38 (2015), pp. 24641–24665. DOI: 10.1039/C5CP02238J.

[163] Konstantin Karandashev et al. "Kinetic isotope effects and how to describe them". In: *Structural Dynamics* 4.6 (2017), p. 061501. ISSN: 2329-7778. DOI: 10.1063/1.4996339.

[164] Alexandra Vardi-Kilshtain, Neta Nitoker, and Dan Thomas Major. "Nuclear quantum effects and kinetic isotope effects in enzyme reactions". In: *Archives of Biochemistry and Biophysics* 582 (2015), pp. 18–27. ISSN: 1096-0384. DOI: 10.1016/j.abb.2015.03.001.

[165] Markus Kowalewski et al. "Catching Conical Intersections in the Act: Monitoring Transient Electronic Coherences by Attosecond Stimulated X-Ray Raman Signals". In: *Physical Review Letters* 115.19 (2015), p. 193003. ISSN: 0031-9007, 1079-7114. DOI: 10.1103/PhysRevLett.115.193003.

[166] Eszter Saly, Dávid Ferenc, and Edit Mátyus. "Pre-Born–Oppenheimer energies, leading-order relativistic and QED corrections for electronically excited states of molecular hydrogen". In: *Molecular Physics* 121.17 (2023), e2163714. ISSN: 0026-8976. DOI: 10.1080/00268976.2022.2163714.

[167] Tao E. Li and Sharon Hammes-Schiffer. "Nuclear-Electronic Orbital Quantum Dynamics of Plasmon-Driven H2 Photodissociation". In: *Journal of the American Chemical Society* 145.33 (2023), pp. 18210–18214. ISSN: 0002-7863. DOI: 10.1021/jacs.3c04927.

[168] Joseph A. Dickinson, Qi Yu, and Sharon Hammes-Schiffer. "Generalized Nuclear-Electronic Orbital Multistate Density Functional Theory for Multiple Proton Transfer Processes". In: *The Journal of Physical Chemistry Letters* 14.26 (2023), pp. 6170–6178. DOI: 10.1021/acs.jpclett.3c01422.

[169] Tao E. Li and Sharon Hammes-Schiffer. "Electronic Born–Oppenheimer approximation in nuclear-electronic orbital dynamics". In: *The Journal of Chemical Physics* 158.11 (2023), p. 114118. ISSN: 0021-9606. DOI: 10.1063/5.0142007.

[170] Simon P. Webb, Tzvetelin Iordanov, and Sharon Hammes-Schiffer. "Multiconfigurational nuclear-electronic orbital approach: Incorporation of nuclear quantum effects in electronic structure calculations". In: *The Journal of Chemical Physics* 117.9 (2002), pp. 4106–4118. ISSN: 0021-9606, 1089-7690. DOI: 10.1063/1.1494980.

[171] Joonho Lee et al. "Generalized Unitary Coupled Cluster Wave functions for Quantum Computation". In: *Journal of Chemical Theory and Computation* 15.1 (2019), pp. 311–324. ISSN: 1549-9618. DOI: 10.1021/acs.jctc.8b01004.

[172] Jean Christophe Tremblay and Tucker Carrington Jr. "Using preconditioned adaptive step size Runge-Kutta methods for solving the time-dependent Schrödinger equation". In: *The Journal of Chemical Physics* 121.23 (2004), pp. 11535–11541. ISSN: 0021-9606. DOI: 10.1063/1.1814103.

[173] Marcello Benedetti, Mattia Fiorentini, and Michael Lubasch. "Hardware-efficient variational quantum algorithms for time evolution". In: *Physical Review Research* 3.3 (2021), p. 033083. ISSN: 2643-1564. DOI: 10.1103/PhysRevResearch.3.033083.

[174] Ingemar Bengtsson and Karol Zyczkowski. *Geometry of Quantum States: An Introduction to Quantum Entanglement*. Cambridge: Cambridge University Press, 2006. DOI: 10.1017/CBO9780511535048.

[175] John von Neumann. "Thermodynamik quantummechanischer gesamheiten". In: *Gott. Nach.* 1 (1927), p. 273.

[176] John von Neumann. *Mathematical Foundations of Quantum Mechanics: New Edition*. Princeton University Press, 2018. 324 pp. ISBN: 978-0-691-17856-1.

[177] V. E. Elfving et al. *How will quantum computers provide an industrially relevant computational advantage in quantum chemistry?* 2020. DOI: 10.48550/arXiv.2009.12472. arXiv: 2009.12472[physics,physics:quant-ph].

Paper I

IOP Publishing Electron. Struct. 5 (2023) 035002 https://doi.org/10.1088/2516-1075/ace86d

Electronic Structure

PAPER

Accelerating variational quantum eigensolver convergence using parameter transfer

Mårten Skogh[1,2], Oskar Leinonen[1,3], Phalgun Lolur[1,4] and Martin Rahm[1,*]

[1] Department of Chemistry and Chemical Engineering, Chalmers University of Technology, SE-412 96 Gothenburg, Sweden
[2] Data Science & Modelling, Pharmaceutical Science, R&D, AstraZeneca, Gothenburg, Sweden
[3] Present address: Department of Physics, University of Oslo, N-0316 Oslo, Norway
[4] Present address: National Quantum Computing Centre (NQCC), Rutherford Appleton Laboratory, Harwell Campus, Didcot, Oxfordshire, OX11 0QX, United Kingdom
* Author to whom any correspondence should be addressed.

E-mail: martin.rahm@chalmers.se

Keywords: quantum computation, potential energy surfaces, quantum chemistry

Supplementary material for this article is available online

OPEN ACCESS

RECEIVED
29 March 2023

REVISED
9 July 2023

ACCEPTED FOR PUBLICATION
18 July 2023

PUBLISHED
4 August 2023

Abstract

One impediment to the useful application of variational quantum algorithms in quantum chemistry is slow convergence with large numbers of classical optimization parameters. In this work, we evaluate a quantum computational warm-start approach for potential energy surface calculations. Our approach, which is inspired by conventional computational methods, is evaluated using simulations of the variational quantum eigensolver. Significant speedup is demonstrated relative to calculations that rely on a Hartree–Fock initial state, both for ideal and sampled simulations. The general approach of transferring parameters between similar problems is promising for accelerating current and near-term quantum chemistry calculations on quantum hardware, and is likely applicable beyond the tested algorithm and use case.

1. Introduction

Quantum computing has been proposed as a possible next step in the field of computational chemistry, especially when applied to highly correlated systems [1]. Theoretical proof of potential quantum speedup exists for problems involving highly correlated systems [2, 3]. However, these results rely on fault-tolerant quantum computers. While quantum computing has come a long way, fault-tolerant computing is yet to be realized [4–6]. While the technology matures, other avenues to useful quantum computing are being explored. The period leading up to fault-tolerant quantum computing is often referred to as the Noisy Intermediate-Scale Quantum (NISQ) era [7]. Quantum devices in this regime are susceptible to many types of errors and noise [8, 9].

Despite these impediments, algorithms have been designed to leverage existing NISQ hardware for quantum chemistry calculations [1, 3, 10]. Several such approaches rely on so-called Variational Quantum Algorithms (VQAs) [11]. These variational algorithms, a subset of which will be described in more detail below, utilize a parameterized quantum circuit to generate trial states for which a cost (or loss) can be evaluated through measurement on the NISQ device. By coupling the quantum computer to a classical optimization algorithm, the circuit parameters can be optimized such that the cost is minimized. Even though the addition of a classical optimization step allows for some level of noise suppression [11, 12], several works [9, 12–15] have highlighted the challenges of optimizing the high dimensional parameter spaces associated with these types of quantum circuits. While there is no formal proof that VQAs can out-perform classical implementations, they are nonetheless useful in pushing the limits of near-term quantum computation [16, 17].

For chemistry applications, the (currently) most common algorithm is the Variational Quantum Eigensolver (VQE), which has been used to perform several quantum chemistry calculations on NISQ hardware [18–22]. The VQE algorithm is briefly summarized below, and in more detail in, e.g. [3, 19, 23, 24].

The number of parameters and energy evaluations required for VQE calculations can reach into the tens or hundreds of thousands, even for small molecules [12]. The large number of parameters makes calculations slow, and in some instances practically unfeasible. For example, more than 16 000 variational parameters and 69 qubits are required to describe a single hydrogen cyanide molecule using unitary coupled cluster theory including single and double excitations (UCCSD) and a standard 6−31 + G(d) basis set. More advanced versions of the VQE algorithm exist that reduce the number of parameters. However, the fundamental challenge of high dimensional optimization remains as the size of the simulated system increases.

In this work, we outline and evaluate a general strategy for accelerating convergence of VQE-type [23–25] algorithms through the use of a parameter transfer (PT) approach. We provide proof-of-concept calculations that demonstrate speedup of VQE-based evaluation of potential energy surfaces (PESs) of molecules, both for ideal simulations and in the presence of sampling noise.

2. The VQE algorithm

The objective of any VQA is to find the parameters $\boldsymbol{\theta}^{\min}$ that minimize the value of some cost (loss) function $C(\boldsymbol{\theta})$. How one defines the cost function varies. Nevertheless, some aspects of the cost function are common to all implementations. The type of cost functions that we will discuss in this work can be seen as functions of (1) a parameterized circuit, $U(\boldsymbol{\theta})$, and (2) a set of measurable operators (observables), $\{\hat{O}_k\}$, that sums to the cost.

The VQE is a particular implementation of a VQA that aims to identify a set of parameters that minimize the expectation value of a Hamiltonian [26]. The VQE algorithm can consequently be formulated as a minimization problem

$$\min_{\boldsymbol{\theta}} E(\boldsymbol{\theta}) = \min_{\boldsymbol{\theta}} \langle \Phi(\boldsymbol{\theta}) | \hat{H} | \Phi(\boldsymbol{\theta}) \rangle, \tag{1}$$

where the cost is the energy $E(\boldsymbol{\theta})$. Because we rely on the Born–Oppenheimer approximation, energy minimization is performed for fixed nuclear coordinates $\boldsymbol{R}$ with the corresponding Hamiltonian $\hat{H}(\boldsymbol{R})$. When formulated in second quantization this electronic Hamiltonian reads as,

$$\hat{H} = E_{NN} + \sum h_{ij}\, \hat{a}_i^+ \hat{a}_j + \tfrac{1}{2} \sum h_{ijkl}\, \hat{a}_i^+ \hat{a}_j^+ \hat{a}_k \hat{a}_l \ . \tag{2}$$

In equation (2), h_{ij} and h_{ijkl} are one- and two-electron integrals, and E_{NN} is the classical nuclear-nuclear repulsion energy, $\hat{a}_m^+$ and $\hat{a}_m$ are the fermionic creation and annihilation operators for the spin orbital m [27, 28]. As we simulate physical systems, the energy minimization of equation (1) is commonly subjected to constraints, such as particle and spin conservation.

The implementation of VQE can be broken down into three main components: quantum circuit execution, measurement, and classical optimization of parameters with respect to the expectation value (figure 1). The quantum circuit itself can be represented by two unitary operators: the state preparation, U_{prep}, and the ansatz, $U(\boldsymbol{\theta})$. The state preparation circuit acts on the all-zero state to create some initial state $|\Phi_{\text{init}}\rangle = U_{\text{prep}}|0\ldots0\rangle$. This initial state is then modified through the application of the ansatz to produce a trial state $|\Phi(\boldsymbol{\theta})\rangle$, that is $U(\boldsymbol{\theta})|\Phi_{\text{init}}\rangle = |\Phi(\boldsymbol{\theta})\rangle$. Measurements of relevant Pauli operators are subsequently used to reconstruct the expectation value that, in turn, is used to guess a new set of parameters to be used in the ansatz. This process is repeated until some convergence criterion is met. While the VQE algorithm may appear straightforward, each step presents challenges that are critical for the quality of the calculation. A more detailed description of how the electronic structure problem is mapped to a quantum circuit ansatz is provided in the supporting information (SI).

A crucial part of the VQE algorithm, which can make or break any attempt at convergence, is providing a suitable initial guess (state), $|\Phi_{\text{init}}\rangle$, for the wavefunction (*state preparation* in figure 1). In this work, we explore the general approach of accelerating convergence by transferring parameters from a previously converged calculation or otherwise known solution.

3. Parameter Transfer (PT)

Central to the idea of PT is the existence of an underlying similarity between two problems. In many cases, some form of commonality is apparent. For example, in chemistry, molecular structures might differ by a minor chemical substitution, or be related by symmetry or conformation. An important class of problems also related to similarity is quantum mechanical problems treated at different levels of theory or accuracy. The level of theory might, for instance, be increased, by utilizing larger basis sets, more flexible wavefunctions, or by limiting approximations in the Hamiltonian. Others, including one of us (MS), have shown that PT can be used to accelerate post-Born–Oppenheimer VQE calculations, by relying on related

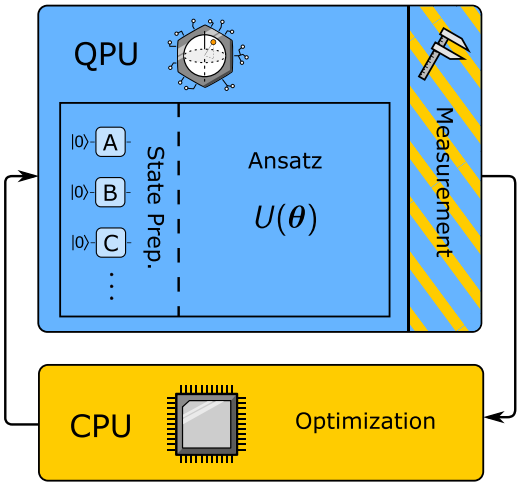


Figure 1. Schematic of a Variational Quantum Algorithm. The procedure can be viewed as consisting of 3 major steps: circuit execution, measurement, and classical optimization. The steps are repeated until convergence. The circuit execution can be further sub-divided into state preparation and description of the ansatz.

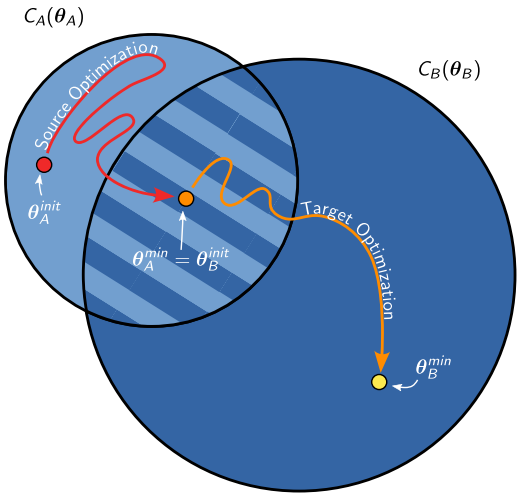


Figure 2. A schematic illustration of the PT approach. Two VQA problems—source A and target B—overlap in a region where problems can be considered similar. A set of converged parameters, $\boldsymbol{\theta}_A^{\min}$, are transferred from the source to be used as initial parameters, $\boldsymbol{\theta}_B^{\text{init}}$, in the optimization of the target problem.

Born–Oppenheimer approximation-based calculations [29]. In what follows, we will first briefly discuss PT in the context of a general cost function, $C(\boldsymbol{\theta})$. We will then address the case of molecular ground state energies (cf $E(\boldsymbol{\theta})$ in equation (1)).

Providing a general definition or measure of similarity is outside the scope of this work. We here view similarity in terms of relations between the shape of the landscapes of two cost functions, $C_A(\boldsymbol{\theta}_A)$ and $C_B(\boldsymbol{\theta}_B)$. Regions around minima of cost functions, $\boldsymbol{\theta}_A^{\min}$ and $\boldsymbol{\theta}_B^{\min}$, are considered similar if the two minima offer approximate solutions to the other, i.e. if $C_B\left(\boldsymbol{\theta}_A^{\min}\right) \approx C_B\left(\boldsymbol{\theta}_B^{\min}\right)$ and $C_A\left(\boldsymbol{\theta}_B^{\min}\right) \approx C_A\left(\boldsymbol{\theta}_A^{\min}\right)$. Note that the two cost functions need not have similar values, nor the same number of parameters. To understand this seeming dichotomy, we can consider a simple example: the separate minimizations in x of $y = x^2$ and $y = x^4 + c$. Both problems share the same solution in $x^{\min} = 0$ while differing in cost at an arbitrary value x, since c is unbound. For the purpose of optimization, it is important that the two convex regions surrounding each minimum overlap (figure 2). Given that an overlap exist it is possible to start at $\boldsymbol{\theta}_A^{\min}$ and follow the steepest decent to reach $\boldsymbol{\theta}_B^{\min}$.

In the optimization problems considered herein, there is an underlying order to the calculations; one is performed before the other. We use *source* to mean an initial calculation. *Target*, in contrast, refers to the calculation to which the parameters are transferred and that we wish to accelerate.

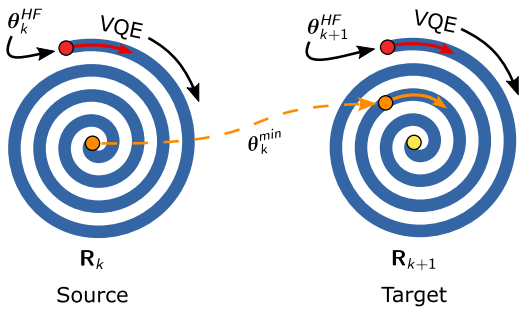


Figure 3. Schematic view of two VQE calculations being performed in series, where spirals represent gradual convergence towards a ground state. The converged parameters, $\boldsymbol{\theta}_k^{\min}$, for a geometry $\boldsymbol{R}_k$ are passed to the next VQE calculation for geometry $\boldsymbol{R}_{k+1}$. PT thereby allows for shortcutting the optimization process, as compared to an otherwise default initial guess.

The underlying principle of PT is common practice in conventional quantum chemistry. For example, PT can be useful for speeding up convergence of self-consistent field optimization in PES sampling, either by transferring parameters between wavefunctions of neighboring geometries or different levels of theory. PES, in turn, form the basis for evaluating equilibria and transition state geometries, reaction rates, spectroscopic constants, and other properties of molecules. PT is also related to the field of transfer learning in machine learning [30, 31]. The implementation of PT in quantum computation has thus far been limited. Promising speedup for the Quantum Approximate Optimization Algorithm [32–37] and for VQA models [38] have been demonstrated with similar approaches. IBM has implemented extrapolation (or bootstrapping) methods in their Qiskit framework for the purpose of accelerating PES sampling of molecules [39]. However, while some methods are available, testing and characterization of these methods is lacking. Overall, little is known about the efficiency of PT when applied to the VQE, or in the presence of noise.

Here, we combine PT with VQE and apply it to the problem of PES sampling of molecules. Figure 3 illustrates how PT connects, and ideally accelerates, consecutive variational VQE optimizations on a PES, by providing improved initial guesses. We emphasize that this combination of PT with VQE is but a small first step towards extending PT to quantum computation of chemistry more generally. Some of us (MS) have already built upon this first step and explore PT between levels of theory in another work [29].

The initial geometry, $\boldsymbol{R}_0$, in a PT-VQE PES evaluation is, naturally, performed without any transfer (as there is no previous solution to transfer from). The subsequent calculation of a neighboring geometry, $\boldsymbol{R}_1$, is then initiated using parameters $\boldsymbol{\theta}_0^{\min}$, optimized for $\boldsymbol{R}_0$. Extending to the general case of neighboring geometries $\boldsymbol{R}_k$, and $\boldsymbol{R}_{k+1}$, we utilize a converged source ground state, $|\Phi\left(\boldsymbol{\theta}_{\boldsymbol{R}_k}^{\min}\right)\rangle$, as the initial guess for the target ground state $|\Phi\left(\boldsymbol{\theta}_{\boldsymbol{R}_{k+1}}^{GS}\right)\rangle$. The difference in geometry (the *transfer step*) is then $\boldsymbol{d} = \boldsymbol{R}_{k+1} - \boldsymbol{R}_k$ and the *transfer distance* is $d = ||R_{k+1} - R_k||$. Here we intentionally leave the norm unspecified for the general case since what constitutes *distance* may vary between systems. We will apply PT to accelerate exploration of relatively simple one-dimensional PES sampling using VQE. For this reason, we refer to Euclidean distances (2-norm).

4. Quantifying speedup

To quantify the efficacy of PT-accelerated VQE, we compare it to the de facto standard method for chemically inspired ansätze: Hartree–Fock (HF), a mean-field description of electron interactions [28]. We define *speedup* for a given target geometry $\boldsymbol{R}$ and transfer step $\boldsymbol{d}$ as

$$S = S(\boldsymbol{R}, \boldsymbol{d}) = \frac{N_{\mathrm{HF}}(\boldsymbol{R}, \boldsymbol{d})}{N_{\mathrm{PT}}(\boldsymbol{R}, \boldsymbol{d})} - 1, \tag{3}$$

where, N_{HF} and N_{PT} are the number of times energy expectation values, $\langle E \rangle$, needs to be evaluated for convergence to be reached with a HF and a PT initial guess, respectively. In our following evaluation of PT, we will additionally present speedup averaged over all n considered combinations $(\boldsymbol{R}, \boldsymbol{d})$ where $\|\boldsymbol{d}\| = d$,

$$\bar{S} = \bar{S}(d) = \frac{1}{n} \sum_{(\boldsymbol{R}, \boldsymbol{d})} S(\boldsymbol{R}, \boldsymbol{d}) \quad . \tag{4}$$

Some optimization algorithms require additional evaluations as part of a setup routine, as discussed in the supporting information.

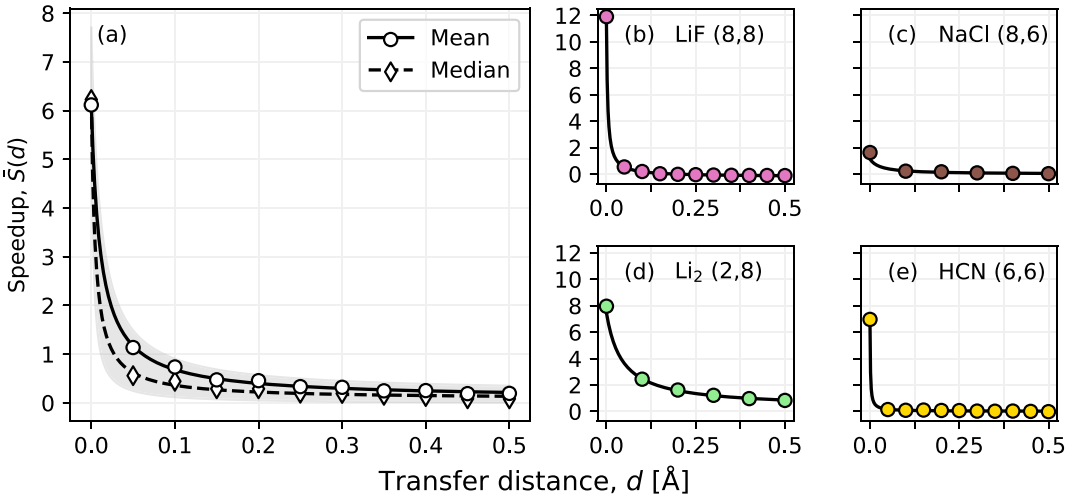


Figure 4. (a) Equation (4) averaged over all our tested systems. The grey area corresponds to the interquartile range (middle 50% of values). Black lines are fitted rational functions, details are given in the supporting information. (b)–(e) Mean speedup in STO-3G basis set simulations of LiF (8,8), NaCl (6,6), Li_2 (2,8), and HCN (6,6), respectively. Values within parentheses denote the size of active spaces as (electrons, orbitals).

5. Methods

All simulations were performed using the VQE algorithm implemented in Qiskit 0.39.3 [39] combined with Gaussian 16, version B.01-AVX2 [40]. All calculations relied on Unitary Coupled Cluster Singles and Doubles (UCCSD) theory, while basis sets and active spaces were varied as indicated in figure 4. We denote the number of active orbitals and electrons using the notation (electrons, orbitals). Active spaces were determined using Qiskit's *ActiveSpaceTransformer*, limiting the active space to those orbitals closest to the Fermi level. Molecular Hamiltonians were mapped to simulated qubits using parity encoding. All simulations implemented two-qubit reduction, made possible by considering alpha and beta spin parity conservation [41]. The resulting quantum circuits thus range in size from 2 to 14 qubits in the noise-free simulations (details are provided in the SI).

We relied on Qiskit's Sequential Least Squares Programming (SLSQP) [42] optimization algorithm for exact simulations. Constrained Optimization By Linear Approximation (COBYLA) [43] and the Simultaneous Perturbation Stochastic Approximation (SPSA) [44] optimization algorithms were both evaluated for sampling-based simulations. Convergence of the SPSA algorithm was determined over the last 100 energy evaluations, by comparing $\langle E \rangle$ averaged over the first and last sets of 50 values. The calculation was considered converged if the difference between the two sets were below a threshold of 0.0001 Ha.

Sampling and noise: Each combination of target and transfer geometry were run with the QASM simulator in Qiskit, which performs sampling from an ideal statevector simulation (as opposed to exact statevector). As a practical compromise between speed of simulation and sampling precision when using the SPSA optimizer we took 32 786 (2^{15}) samples for each calculation. When using the COBYLA optimizer 1048 576 (2^{20}) samples were instead used, owing to the algorithm's lower noise-resilience. These choices retain a substantial contribution to stochastic noise in the simulations, and is similar in magnitude to what can be expected in current and near-term hardware experiments [20]. To evaluate average performance, multiple calculations were performed for both the HF starting guess and the PT implementation, the results of which were averaged. Speedup for each point $(\boldsymbol{R}, \boldsymbol{d})$ was evaluated as an average over all possible combinations of 10–15 separate calculations initialized with HF and PT. To limit computational cost associated with these simulations, the considered circuit sizes were reduced to include only 2 (H_2 and HeH^+) and 4 qubits (LiH).

6. Results and discussion

6.1. PT in noise-free VQE

To evaluate PT, we first look at noise-free (statevector) simulations. While not being faithful models of NISQ hardware, noise-free simulations can provide valuable insight into the limits of quantum algorithms. By removing the effects of noise, algorithmic errors can be separated from those caused by imperfect hardware

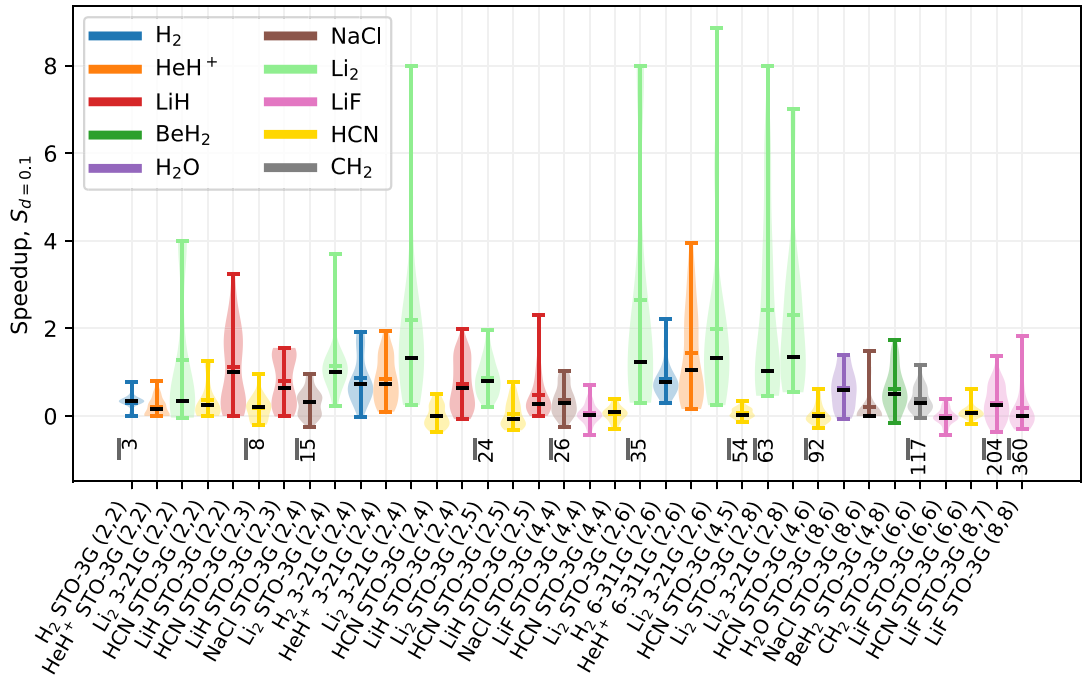


Figure 5. Speedup quantified for a transfer distance $d = 0.1$ Å across the test set of molecular problems. Results are sorted by number of parameters in the variational ansatz shown at the bottom of the figure, increasing from left to right. Mean values of each distribution is indicated by a colored horizontal line between extrema (compare figure 4), while black horizontal lines mark the medians. Widths of colored areas scale with the number of samples of a given PES with a given speedup.

and choices of peripherals such as optimizers. We will later return to assess the effects of sampling noise in a subset of examples.

Our test set consist of 36 different quantum chemistry problems: PES of 10 molecules treated with different basis sets and active spaces, which are detailed below. For each case, we consider the dissociation of a single atom from a molecule, making a one-dimensional PES. In triatomics, such as BeH_2, H_2O and HCN, a single H atom is dissociated, while all bond angles remain fixed.

The molecules we study cover a range of electronic structures in terms of electron correlation, including both archetypical ionic dimers (HeH^+ and NaCl) and covalently bound molecules (H_2 and H_2O). Our test set was selected to represent a variety of parameter set sizes, quantum circuits and Hamiltonians of relevance to current and near-term hardware experiments, while still allowing for classical simulation.

The speedup quantified for a given calculation will inevitably be linked to the choice of optimization algorithm. Some such algorithms will function well for noiseless simulations, whereas others are more suitable in the presence of noise. For noise-free optimization, we here rely on the SLSQP optimizer. SLSQP is a representative optimizer for convex optimization in the absence of noise, commonly used in statevector simulations of VQE [45–48]. As we later implement sampling noise, we will subsequently use the SPSA algorithm, which is better adapted for optimization in the presence of noise. We stress that it is outside the scope of this work to provide benchmarking of optimization algorithms (such benchmarking has been done by others [49–51]). Nevertheless, we expect the general conclusion of our work–useful speedup with PT–to hold for *reasonable* choices of optimizers.

One important limit to PT is the transfer distance d. As d approaches zero in the absence of noise the transfer perfectly corresponds to the already converged solution. Figure 4(a) shows an expected decrease in average and median speedup over our test set as a function of increasing transfer distance. Figures 4(b)–(e) serve to illustrate the variety in speedup between calculations in our test set. We attribute these differences to the suitability of the HF initial guess: there is less benefit to PT in systems that are well-described by a HF state. In some cases, such as the H dissociation from HCN, the speedup is marginal or non-existent, except for very short transfer distances.

For a transfer distance of 0.1 Å, which is a practical choice for PES sampling, we note an average speedup of 70% (figure 4(a)). Note that the choice of a fixed transfer distance for all systems is far from optimal, and 0.1 Å is here merely chosen to facilitate comparison. In figure 5 we look further at the speedup corresponding to this specific transfer distance, but additionally averaged over entire bond dissociation curves, from

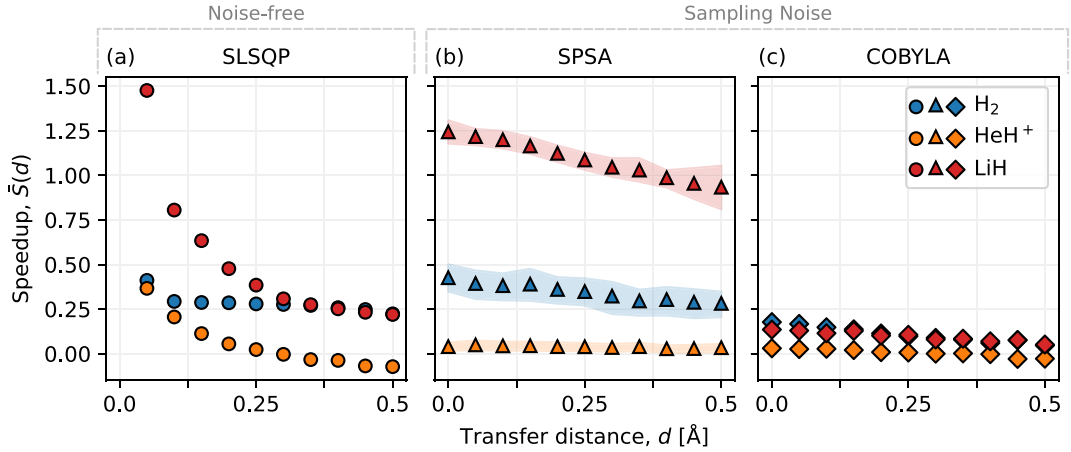


Figure 6. Mean speedup ($\bar{S}(d)$) as a function of transfer distance d in: (a) exact (statevector) simulation combined with SLSQP optimizer. Peak values at $d = 0$ are not shown. (b) Sampling-based simulations with SPSA optimizer. Colored shaded areas show interquartile range over repeated calculations. (c) Sampling-based simulations utilizing the COBYLA optimizer. The COBYLA optimizer performs poorly in the noisy setting, converging prematurely, resulting in similar performance for both HF and PT initial states.

compressed bonds to equilibrated geometries and dissociation. Results in figure 5 correspond to calculations with increasing number of circuit parameters from left to right. An overall trend is not apparent, except for some molecules (like Li_2) where an improvement with the number of parameters may be present. The PT approach appears to perform well under noiseless conditions even for calculations exceeding 100 parameters.

6.2. Effects of measurement sampling

Even with fault-tolerant quantum computing, which is yet to be realized, one needs repeated sampling of expectation values of the quantum system. Such sampling from a noise-less state introduces a stochastic variability, or sampling noise. In this section, we reevaluate three of our test set calculations, H_2 (2,2), HeH^+ (2,2), and LiH (2,3), in the presence of sampling noise. Due to the presence of sampling noise, SLSQP is no longer suitable (nor practically feasible) and we use the SPSA and COBYLA optimization algorithms (figure 6).

The most notable difference in the performance of PT when sampling noise is introduced is a markedly different dependence of speedup on transfer distance. Instead of a rapid decline, as in figure 4, speedup appear to decrease linearly with increasing transfer distance. We attribute this difference to the inherently larger challenge of optimization in the presence of noise. Figure 6(c) serves to highlight the importance of choosing optimizers well suited for the task at hand. Even though convergence is reached with the COBYLA optimizer, the number of evaluations at convergence *decreases* with increasing sampling noise, implying that figure 6(c) reflects a poor choice in optimizer, not a failure of PT.

The large effect of even relatively low noise levels in combination with improper choices of optimizers, highlights a challenge that NISQ algorithms face when being compared to conventional methods. We further discuss optimizer suitability in the supporting information. Overall, our evaluation suggests the potential for substantial PT-driven speedup both without and in the presence of sampling noise.

7. Conclusions

In this work, we have outlined and provided a proof-of-concept for PT, an approach for accelerating consecutive VQE calculations. In our test set, PT provides an average speedup of 70%, as long as a suitable optimization algorithm is used, compared to when a simulation is initialized by a standard HF initial state. Speedup is apparent both in ideal and sampled statevector simulations, provided that source and target geometries are reasonably (∼0.1 Å) close. Results can nevertheless vary substantially between molecular systems, as expected. The PT approach does not appear sensitive to problem complexity, quantified in terms of number of quantum circuit parameters. Simulations subjected to sampling noise also appear less sensitive to the choice of transfer distance. Combined, these results lead us to conclude that PT is an attractive approach for accelerating VQE convergence even for larger more correlated systems. While this study is limited to PT within the context of PES sampling, we also stress that PT can be considered a general principle extendable to other types of calculations. In particular, calculations involving transfer of parameters between different levels of theory is a promising avenue, as shown in [29].

Data availability statement

All data that support the findings of this study are included within the article (and any supplementary files).

Acknowledgments

This research relied on computational resources provided by the Swedish National Infrastructure for Computing (SNIC) at C3SE, NSC and PDC partially funded by the Swedish research council through Grant Agreement No. 2018-05973 and by the National Academic Infrastructure for Supercomputing in Sweden (NAISS) at C3SE and NSC partially funded by the Swedish research council through Grant Agreement No. 2022-06725.

Funding

This research has been supported by funding from the Wallenberg Center for Quantum Technology (WACQT) and from the EU Flagship on Quantum Technology H2020-FETFLAG-2018-03 Project 820363 OpenSuperQ.

ORCID iDs

Mårten Skogh https://orcid.org/0000-0002-6604-3359
Oskar Leinonen https://orcid.org/0009-0004-0149-1531
Phalgun Lolur https://orcid.org/0000-0002-3409-3415
Martin Rahm https://orcid.org/0000-0001-7645-5923

References

[1] Cao Y *et al* 2019 Quantum chemistry in the age of quantum computing *Chem. Rev.* **119** 10856–915
[2] Nielsen M A and Chuang I L 2010 *Quantum Computation and Quantum Information* (Cambridge University Press)
[3] McArdle S, Endo S, Aspuru-Guzik A, Benjamin S and Yuan X 2020 Quantum computational chemistry *Rev. Mod. Phys.* **92** 015003
[4] Aharonov D and Ben-Or M 2008 Fault-tolerant quantum computation with constant error rate *SIAM J. Comput.* **38** 1207–82
[5] Knill E, Laflamme R and Zurek W H 1998 Resilient quantum computation *Science* **279** 342–5
[6] Kitaev A Y 2003 Fault-tolerant quantum computation by anyons *Ann. Phys., NY* **303** 2–30
[7] Preskill J 2018 Quantum computing in the NISQ era and beyond *Quantum* **2** 79
[8] Georgopoulos K, Emary C and Zuliani P 2021 Modeling and simulating the noisy behavior of near-term quantum computers *Phys. Rev.* A **104** 062432
[9] Wang S, Fontana E, Cerezo M, Sharma K, Sone A, Cincio L and Coles P J 2021 Noise-induced barren plateaus in variational quantum algorithms *Nat. Commun.* **12** 6961
[10] Elfving V E, Broer B W, Webber M, Gavartin J, Halls M D, Lorton K P and Bochevarov A 2020 How will quantum computers provide an industrially relevant computational advantage in quantum chemistry? (arXiv:2009.12472)
[11] Cerezo M *et al* 2021 Variational quantum algorithms *Nat. Rev. Phys.* **3** 625–44
[12] Lolur P, Rahm M, Skogh M, García-Álvarez L and Wendin G 2021 Benchmarking the variational quantum eigensolver through simulation of the ground state energy of prebiotic molecules on high-performance computers *AIP Conf. Proc.* **2362** 030005
[13] Bittel L and Kliesch M 2021 Training variational quantum algorithms is NP-hard *Phys. Rev. Lett.* **127** 120502
[14] Cerezo M and Coles P J 2021 Higher order derivatives of quantum neural networks with barren plateaus *Quantum Sci. Technol.* **6** 035006
[15] McClean J R, Boixo S, Smelyanskiy V N, Babbush R and Neven H 2018 Barren plateaus in quantum neural network training landscapes *Nat. Commun.* **9** 4812
[16] McCaskey A J, Parks Z P, Jakowski J, Moore S V, Morris T D, Humble T S and Pooser R C 2019 Quantum chemistry as a benchmark for near-term quantum computers *npj Quantum Inf.* **5** 99
[17] Lubinski T, Johri S, Varosy P, Coleman J, Zhao L, Necaise J, Baldwin C H, Mayer K and Proctor T 2021 Application-oriented performance benchmarks for quantum computing (arXiv:2110.03137)
[18] Arute F *et al* Google AI Quantum and Collaborators 2020 Hartree-Fock on a superconducting qubit quantum computer *Science* **369** 1084–9
[19] Peruzzo A, McClean J, Shadbolt P, Yung M-H, Zhou X-Q, Love P J, Aspuru-Guzik A and O'Brien J L 2014 A variational eigenvalue solver on a photonic quantum processor *Nat. Commun.* **5** 4213
[20] Kandala A, Mezzacapo A, Temme K, Takita M, Brink M, Chow J M and Gambetta J M 2017 Hardware-efficient variational quantum eigensolver for small molecules and quantum magnets *Nature* **549** 242–6
[21] Colless J I, Ramasesh V V, Dahlen D, Blok M S, Kimchi-Schwartz M E, McClean J R, Carter J, de Jong W A and Siddiqi I 2018 Computation of molecular spectra on a quantum processor with an error-resilient algorithm *Phys. Rev.* X **8** 011021
[22] Hempel C *et al* 2018 Quantum chemistry calculations on a trapped-ion quantum simulator *Phys. Rev.* X **8** 031022
[23] 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
[24] Tang H L, Shkolnikov V O, Barron G S, Grimsley H R, Mayhall N J, Barnes E and Economou S E 2021 Qubit-ADAPT-VQE: an adaptive algorithm for constructing hardware-efficient ansätze on a quantum processor *PRX Quantum* **2** 020310
[25] Cervera-Lierta A, Kottmann J S and Aspuru-Guzik A 2021 The meta-variational quantum eigensolver (Meta-VQE): learning energy profiles of parameterized Hamiltonians for quantum simulation *PRX Quantum* **2** 020329

[26] Yuan X, Endo S, Zhao Q, Li Y and Benjamin S 2019 Theory of variational quantum simulation *Quantum* **3** 191
[27] Szabo A and Ostlund N S 1996 *Modern Quantum Chemistry: Introduction to Advanced Electronic Structure Theory* (Dover Publications)
[28] Helgaker T, Jorgensen P and Olsen J 2014 *Molecular Electronic-Structure Theory* (Wiley)
[29] Kovyrshin A, Skogh M, Broo A, Mensa S, Sahin E, Crain J and Tavernelli I 2023 A quantum computing implementation of nuclearelectronic orbital (NEO) theory: toward an exact pre-born–oppenheimer formulation of molecular quantum systems *J. Chem. Phys.* **158** 214119
[30] Weiss K, Khoshgoftaar T M and Wang D 2016 A survey of transfer learning *J. Big Data* **3** 9
[31] Zhuang F, Qi Z, Duan K, Xi D, Zhu Y, Zhu H, Xiong H and He Q 2021 a comprehensive survey on transfer learning *Proc. IEEE* **109** 43–76
[32] Shaydulin R, Lotshaw P C, Larson J, Ostrowski J and Humble T S 2022 Parameter transfer for quantum approximate optimization of weighted maxcut (arXiv:2201.11785)
[33] Jing H, Wang Y and Li Y 2022 Data-driven quantum approximate optimization algorithm for cyber-physical power systems (arXiv:2204.00738)
[34] Galda A, Liu X, Lykov D, Alexeev Y and Safro I 2021 Transferability of optimal QAOA parameters between random graphs (arXiv:2106.07531)
[35] Egger D J, Marecek J and Woerner S 2021 Warm-starting quantum optimization *Quantum* **5** 479
[36] Tate R, Farhadi M, Herold C, Mohler G and Gupta S 2022 Bridging classical and quantum with SDP initialized warm-starts for QAOA (arXiv:2010.14021)
[37] Okada K N, Nishi H, Kosugi T and Matsushita Y 2022 Systematic study on the dependence of the warm-start quantum approximate optimization algorithm on approximate solutions (arXiv:2209.02942)
[38] Liu H-Y, Sun T-P, Y-c W, Han Y-J and Guo G-P 2023 Mitigating barren plateaus with transfer-learning-inspired parameter initializations *New J. Phys.* **25** 013039
[39] Aleksandrowicz G *et al* 2019 Qiskit: an open-source framework for quantum computing (https://doi.org/10.5281/zenodo.2562111)
[40] Hartree D R and Hartree W 1935 Self-consistent field, with exchange, for beryllium *Proc. R. Soc.* A **150** 9–33
[41] Frisch M J *et al* 2016 Gaussian 16 Gaussian Inc.
[42] Kraft D 1988 A software package for sequential quadratic programming 88-28 DLR German Aerospace Center – Institute for Flight Mechanics, Koln, Germany
[43] Powell M J D 1998 Direct search algorithms for optimization calculations *Acta Numer.* **7** 287–336
[44] Spall J C 1998 Implementation of the simultaneous perturbation algorithm for stochastic optimization *IEEE Trans. Aerosp. Electron. Syst.* **34** 817–23
[45] Gao Q, Jones G O, Motta M, Sugawara M, Watanabe H C, Kobayashi T, Watanabe E, Ohnishi Y, Nakamura H and Yamamoto N 2021 Applications of quantum computing for investigations of electronic transitions in phenylsulfonyl-carbazole TADF emitters *npj Comput. Mater.* **7** 70
[46] Rattew A G, Hu S, Pistoia M, Chen R and Wood S 2020 A domain-agnostic, noise-resistant, hardware-efficient evolutionary variational quantum eigensolver (arXiv:1910.09694)
[47] Choudhary K 2021 Quantum computation for predicting electron and phonon properties of solids *J. Phys.: Condens. Matter* **33** 385501
[48] Yalouz S, Koridon E, Senjean B, Lasorne B, Buda F and Visscher L 2022 Analytical nonadiabatic couplings and gradients within the state-averaged orbital-optimized variational quantum eigensolver *J. Chem. Theory Comput.* **18** 776–94
[49] Lockwood O 2022 An empirical review of optimization techniques for quantum variational circuits (arXiv:2202.01389)
[50] 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 *Quantum Sci. Technol.* **4** 014008
[51] Lavrijsen W, Tudor A, Müller J, Iancu C and de Jong W 2020 Classical optimizers for noisy intermediate-scale quantum devices *2020 IEEE Int. Conf. on Quantum Computing and Engineering (QCE)* pp 267–77

Supporting Information for:

Accelerating Variational Quantum Eigensolver Convergence using Parameter Transfer

Mårten Skogh[1,2], Oskar Leinonen[1], Phalgun Lolur[1], and Martin Rahm*[1]
[1]*Department of Chemistry and Chemical Engineering, Chalmers University of Technology, SE-412 96 Gothenburg, Sweden*
[2]*Data Science & Modelling, Pharmaceutical Science, R&D, AstraZeneca, Gothenburg, Sweden*
Email: martin.rahm@chalmers.se

Source code and data available through the following link: https://doi.org/10.5878/fvza-z272

Problem Details

Figure S1 shows the quantum circuit for one of our test cases, the two-qubit quantum circuit corresponding to the calculation of H_2 with an STO-3G basis set. Other circuits are available through the supplementary data link provided above. In Figure S1, the first gate, $\boldsymbol{U}$, acting on $\boldsymbol{q_0}$ comprises the state preparation, constructing a HF state. The remaining gates are part of the variational ansatz, which includes four parameterized gates. Since the ansatz used is the UCCSD and the basis set is minimal, the three parameters of the parameterized gates correspond to the two possible single excitations (t[0] and t[1]) and the double excitation (t[2]).

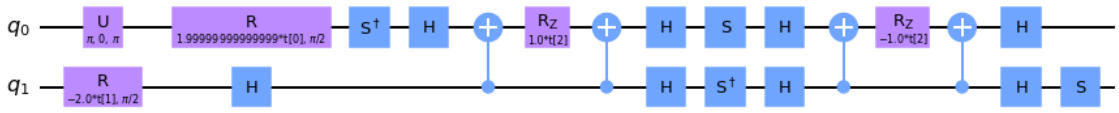


Figure S1: Quantum circuit for the VQE calculations of H_2 at the UCCSD/STO-3G level of theory. The circuit contains three parameters, t[0], t[1], and t[2] and two qubits $\boldsymbol{q_0}$ and $\boldsymbol{q_1}$. The VQE problem maps the electronic structure problem to the minimization given by Eq. (1).

To see the correspondence of the variational parameters to the UCC ansatz, it is beneficial to look at the definition of the UCCSD operator

$$\boldsymbol{U(t) = e^{T-T^+}}, \tag{S1}$$

where, in the case of H_2 in minimal basis, the cluster operator, $\boldsymbol{T}$, is defined as

$$\boldsymbol{T = t_0\, a_1^+ a_0 + \; t_1\, a_3^+ a_2 + t_2\, a_1^+ a_3^+ a_2 a_0}\,, \tag{S2}$$

where our three $\boldsymbol{t = [t_0, t_1, t_2]}$ parameters appear. Eq. S1 illustrates the relationship between the circuit parameters and the electronic structure problem. Application of the quantum circuit on the prepared state thus equals the unitary operation

$$\boldsymbol{U(t)|\Phi_0\rangle = |\Phi(t)\rangle}\,. \tag{S3}$$

By measuring the state after application of the parameterized circuit, the energy can be calculated as the expectation value of the Hamiltonian

$$\boldsymbol{E = \langle \Phi(t)|\hat{H}|\Phi(t)\rangle}\,. \tag{S4}$$

In Eq. (S4), the Hamiltonian is expressed a sum of Pauli terms weighted by coefficients $\boldsymbol{c_i}$,

$$\hat{\boldsymbol{H}} = \sum \boldsymbol{c_i}\, \boldsymbol{\sigma_{n-1}} \otimes \ldots \otimes \boldsymbol{\sigma_0}\ . \qquad (\boldsymbol{S5})$$

The VQE solution (the ground state energy) to the electronic structure problem can thus be represented as the variational minimization given in Eq, (1).

Effects of Optimizer Setup on Speedup

How optimization is implemented varies greatly between optimization algorithms. In Figure S2 we give a generalized picture of optimization that we find relevant to PT. We wish to highlight the initial step of Figure S2, the *setup*. The setup is typically a one-time event, performed to calibrate optimizer parameters, or perform some minimal number of evaluations before checking convergence. For situations where convergence is reached following relatively few iterations, the setup stage might account for a substantial part of the total number of evaluations. We use N_s to denote the number of evaluations required before performing a first optimization iteration. This distinction is brought up because our definition of speedup is somewhat sensitive to N_s in those cases where convergence is rapid. The value of N_s naturally varies between different choices of optimization algorithms, as well as their respective configuration.

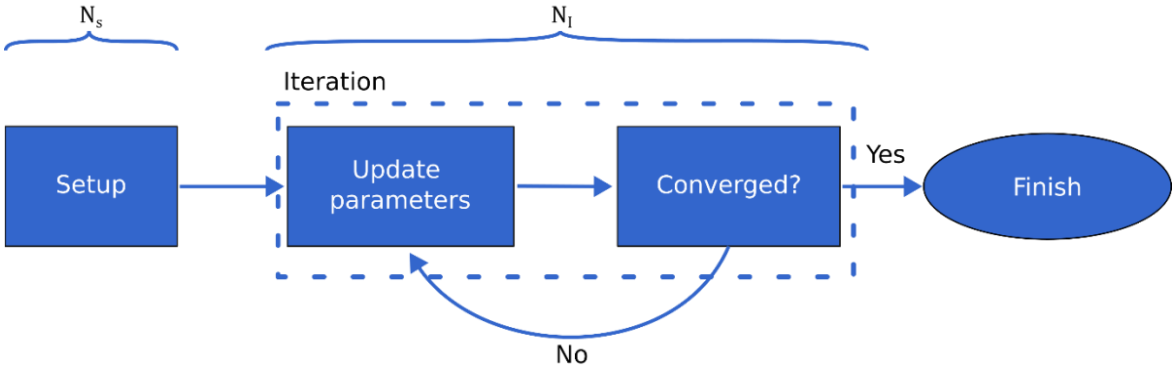


Figure S2: Schematic outline of a generic optimization procedure. A setup procedure is first performed to calibrate the algorithm and/or before convergence can be evaluated. The optimizer then iterates until convergence. The total number of function evaluations performed during optimization is the sum of those performed as part of the setup, N_s, and those performed as repeated iterations, N_I.

In practice, problems containing few (in our case quantum circuit) parameters will see the largest effect of the setup stage on evaluated speedup S, due to their generally rapid convergence. We can quantify the importance of this effect by letting $N_{HF} = N_s + N_I^{(HF)}$ and $N_{PT} = N_s + N_I^{(PT)}$, where $N_I^{(HF)}$ and $N_I^{(PT)}$ are the number of evaluations performed while iterating, starting from a HF and PT guess respectively. Speedup S can then be expressed in terms of $N_I^{(HF)}$ and $N_I^{(PT)}$,

$$S(\boldsymbol{R}, \boldsymbol{d}) = \frac{N_{HF}}{N_{PT}} - 1 = \frac{N_s + N_I^{(HF)}}{N_s + N_I^{(PT)}} - 1. \qquad (S6)$$

Eq. (S6) shows us that if both options (default and PT initialized) for calculation converge fast, such that $N_I^{(HF)} \ll N_s$ and $N_I^{(PT)} \ll N_s$, even optimal parameter transfer, $N_I^{(PT)} = 1$, will only correspond a marginal speedup.

$$S(\boldsymbol{R}, \boldsymbol{d}) \approx \frac{N_s}{N_s} - 1 = 0\,. \qquad (S7)$$

Analytical Approximation and Analysis of $\bar{S}(\boldsymbol{d})$

To ease further study of PT it is helpful to identify an analytical representation of the underlying function, $\bar{S}(\boldsymbol{d})$. Figure 3 and S2 reveals $\bar{S}(\boldsymbol{d})$ to be seemingly continuous and smooth for $\boldsymbol{d} \neq \mathbf{0}$. We expect the same to hold for other systems provided that no degenerate (or close to degenerate) states are present. The dependence of speedup on transfer distance, $\bar{S}(\boldsymbol{d})$, can for our test set be accurately captured by a polynomial fraction of the form

$$\bar{S}(\boldsymbol{d}) = \frac{a_1\|\boldsymbol{d}\| + b_1}{a_2\|\boldsymbol{d}\| + b_2} \,. \tag{S8}$$

At its maximum, $\boldsymbol{d} = \mathbf{0}$, Eq. (S7) simplifies to $\bar{S}(\mathbf{0}) = \bar{S}_{peak} = b_1/b_2$. By assuming $b_1 = \bar{S}_{peak}$ and $b_2 = 1$, Eq. (S8) reduces to,

$$\bar{S}(\boldsymbol{d}) = \frac{a_1\|\boldsymbol{d}\| + S_{peak}}{a_2\|\boldsymbol{d}\| + 1} \,. \tag{S9}$$

Eq. (S9) is the form we (successfully) use to fit the dependence of speedup to *d* throughout this work (Figure S3). This fitting is done by least squares fitting of a_1 and a_2 so to minimize the error with the calculated mean speedup, $\bar{S}(\boldsymbol{d})$.

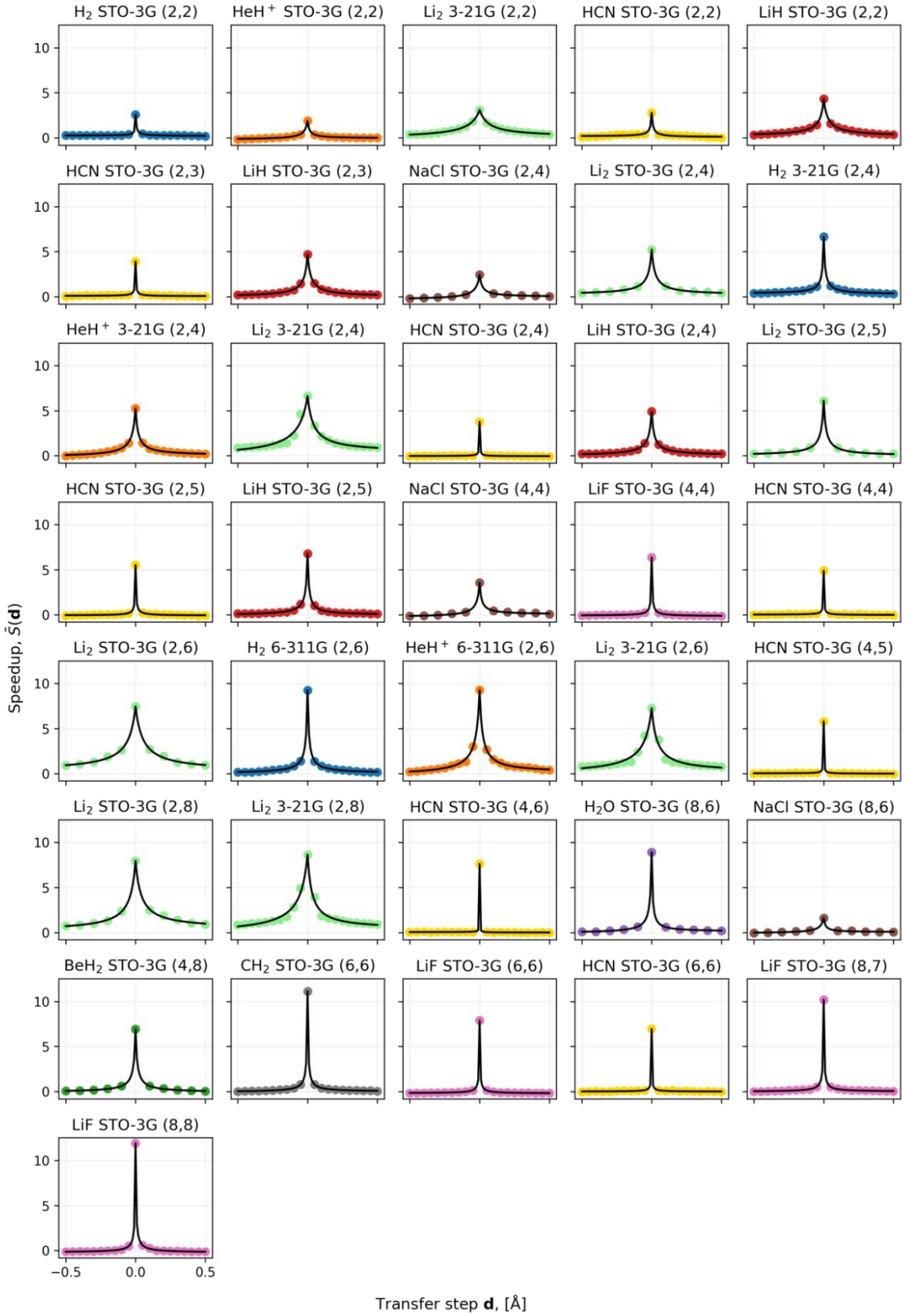


Figure S3: Averaged speedup, $\bar{S}(\boldsymbol{d})$, as a function of transfer step, $\boldsymbol{d}$, for all investigated systems. A rational function given by Eq. (S8) is fitted to the simulated data (black line). Since the transfer step is a vector, its sign can vary depending on the relative position of source and target. The sharpness of the peak, in the one-dimensional case determined by the left and right derivatives at $\boldsymbol{d} = \mathbf{0}$, is indicative of the sensitivity to transfer distance.

Data Summary and Error Analysis

Figure S4 shows the PT speedup calculated for each combination of $\boldsymbol{R}$ and $\boldsymbol{d}$ in all tested systems. For most systems a clear ridge of high speedup can be seen for $\boldsymbol{d} = \mathbf{0}$. Certain systems also show "steady state" behavior as the bond broken region is reached, whereby a constant speedup is maintained for most $\boldsymbol{d}$. Figure S5 shows the corresponding difference in the converged energy when implementing PT as compared to starting optimization from a HF guess. The difference is for most instances within or close to the convergence limit 10^{-6} Ha, implying that both starting guesses in manage to find the ground state. A notable exception to this is CH_2 (6,6), where PT, for some a small set of $\boldsymbol{R}$ and $\boldsymbol{d}$, manages to substantially converge below the converged HF guess (-0.03 Ha).

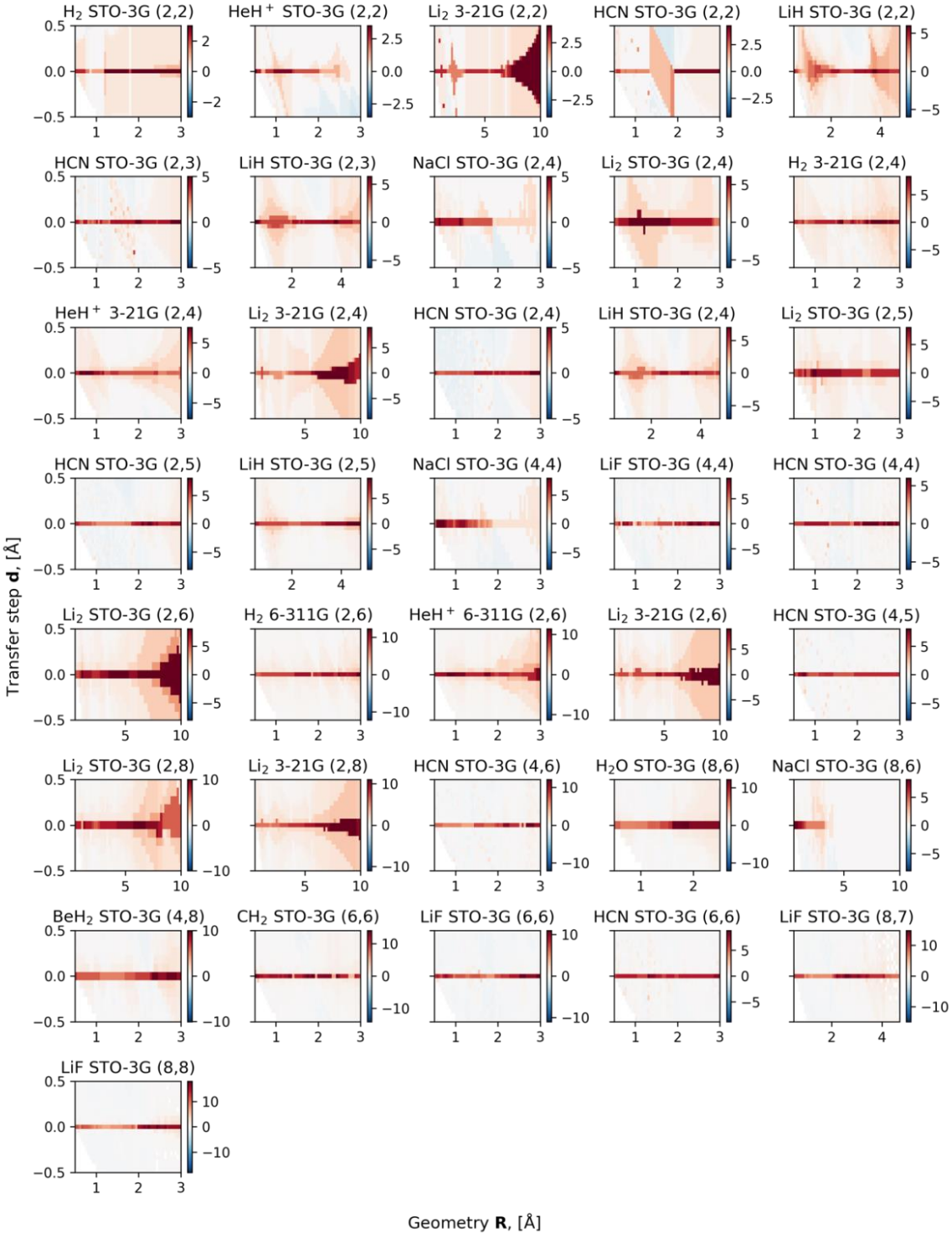


Figure S4: $S(\boldsymbol{R}, \boldsymbol{d})$ for all tested systems and combinations of source and target geometries. The y-axis corresponds to the source geometry, while the x-axis corresponds to transfer step $\boldsymbol{d}$, as defined in the main text. Red and blue colors correspond to high and low speedup, respectively. A ridge of high speedup is consistently shown near $\boldsymbol{d} = \mathbf{0}$. White (not to be confused with light grey) indicates a lack of data, typically seen for source geometries where nuclei are closer than 0.5 Å due to limitations in Gaussian 16. Some regions marked in white also correspond to lack of data due to failed convergence of the HF reference calculation.

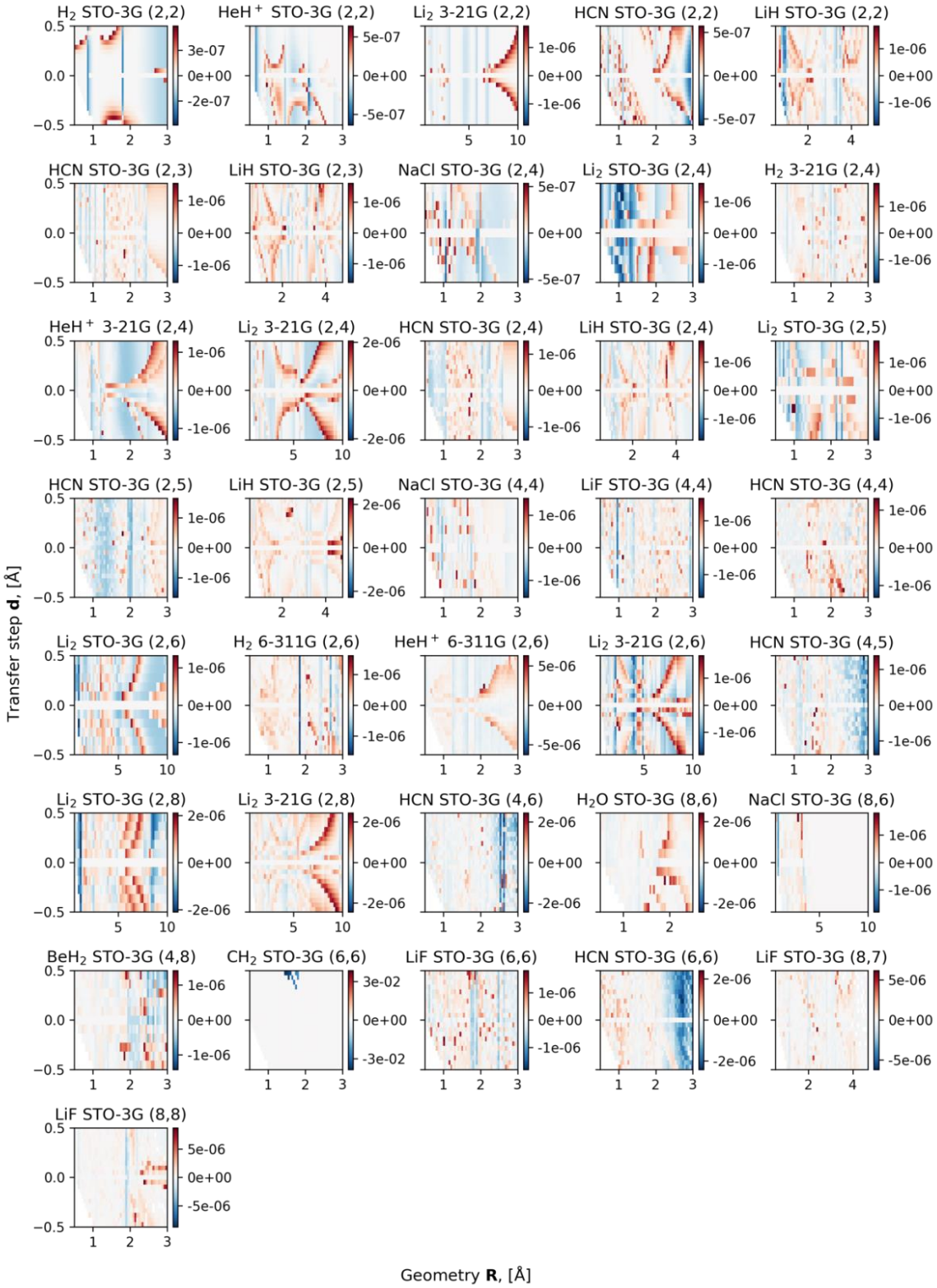


Figure S5: Difference in converged energy (in Ha) when starting from a PT guess as compared to starting from a HF guess. The color of each data point gives the difference between the two starting strategies, with blue indicating convergence of PT to a lower energy and red vice versa. The convergence limit used was 10^{-6} Ha. White (not to be confused with light grey) indicates a lack of data, typically seen for source geometries where nuclei are closer than 0.5 Å due to limitations in Gaussian 16. Some regions marked in white also correspond to lack of data due to failed convergence of the HF reference calculation.

Figure S6 shows the errors with respect to the exact diagonalization of the Hamiltonian when PT is used in conjunction with sampling noise. There is seemingly little connection between error and speedup, indicating that PT does not accelerate the convergence by sacrificing precision.

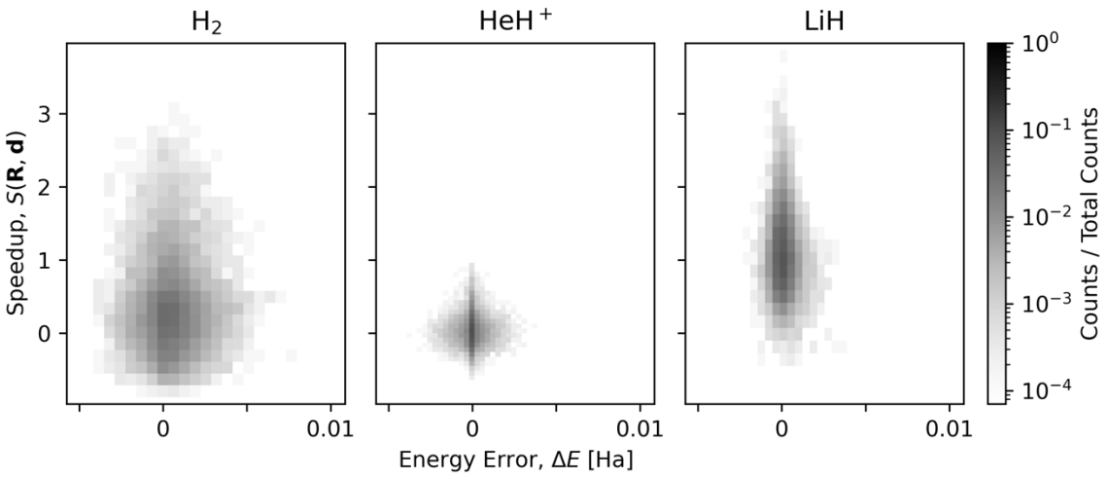


Figure S6: Correlation between speedup, $S(\boldsymbol{R}, \boldsymbol{d})$, and energy error. Color represents the proportion of calculations in each bin, the darker the more calculations with the given speedup and error. A convergence limit of 10^{-4} Ha was used for all systems.

Suitability of COBYLA in Noisy Conditions

Noise adds a substantial challenge to many optimizers. An example of a less suitable optimizer in a noisy setting is COBYLA. While this optimizer manages to reach convergence (Figure S7), reducing the sampling noise by increasing the number of samples *increases* the number of evaluations needed for convergence. This relation between sample size and evaluations indicates that the fast convergence is caused by the noisy environment. We consider this a "false" convergence, not truly finding the ground state, instead relying on noise to pass the convergence test. What further corroborates this interpretation is that all tested sample sizes fail to reach the chosen convergence limit of 10^{-4} Ha.

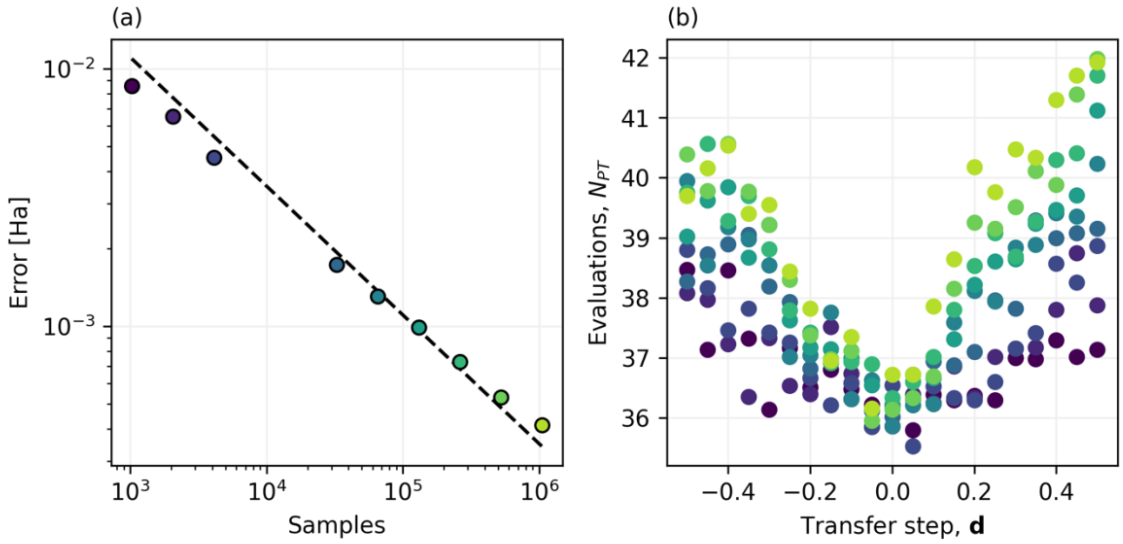


Figure S7: Calculations performed with the COBYLA optimizer using varying number of samples (between 2^{10} and 2^{20}). The number of samples used for each calculation is indicated by the color of each point. Blue indicates a low number of samples; green indicates a high number of samples. (a) The error decreases (almost) as the standard deviation of a normal distribution $\sigma/\sqrt{n}$ (black line). (b) Average number of function evaluations before PT convergence, N_{PT} for a given transfer step $\boldsymbol{d}$. The number of evaluations needed to converge *increases* with the number of samples.

Quantum Circuit Details

The focus of our work is to study the effects of increased parameter counts, beyond what is currently feasible on NISQ hardware, in combination with PT. The circuits we study therefore contain a varying number of gates, parameters, and qubits. Details of these quantum circuits are provided in Table S1. Graphical representations of a selection of the smaller and medium sized circuits are available through the SND link, provided at the first page of this document.

I

Table S1: Quantum circuit details. The number of active electrons and orbitals determine the circuit size and number of parameters.

Active Space	Single-qubit gates[a]	Two-qubit gates[b]	Parameterized gates	Parameters	Qubits
(2, 2)	13	4	4	3	2
(2, 3)	322	172	40	8	4
(2, 4)	883	560	84	15	6
(2, 5)	1836	1276	144	24	8
(2, 6)	3277	2416	220	35	10
(2, 7)	5302	4076	312	48	12
(2, 8)	8007	6352	420	63	14
(4, 3)	322	172	40	8	4
(4, 4)	1618	1096	152	26	6
(4, 5)	4442	3284	360	54	8
(4, 6)	9074	7080	640	92	10
(4, 7)	15978	12924	1000	140	12
(4, 8)	25570	21232	1440	198	14
(6, 4)	883	560	84	15	6
(6, 5)	4444	3284	360	54	8
(6, 6)	11669	9200	828	117	10
(6, 7)	23526	19268	1488	204	12
(6, 8)	40975	34448	2340	315	14
(8, 5)	1836	1276	144	24	8
(8, 6)	9076	7080	640	92	10
(8, 7)	23524	19268	1488	204	12
(8, 8)	46916	39568	2688	360	14

[a]: Excluding parameterized single-qubit gates. The single-qubit gates used are H, S, $S^{\dagger}$, and X.
[b]: Only the CX (CNOT) gate was used.

Paper II

JCTC Journal of Chemical Theory and Computation

pubs.acs.org/JCTC Article

Reference-State Error Mitigation: A Strategy for High Accuracy Quantum Computation of Chemistry

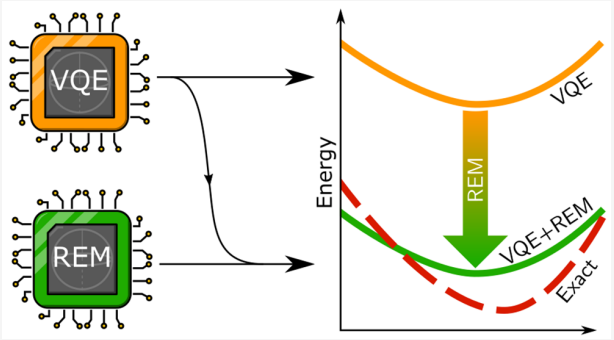
Phalgun Lolur,[‖] Mårten Skogh,[‖] Werner Dobrautz, Christopher Warren, Janka Biznárová, Amr Osman, Giovanna Tancredi, Göran Wendin, Jonas Bylander, and Martin Rahm*

Cite This: *J. Chem. Theory Comput.* 2023, 19, 783−789

Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information

ABSTRACT: Decoherence and gate errors severely limit the capabilities of state-of-the-art quantum computers. This work introduces a strategy for reference-state error mitigation (REM) of quantum chemistry that can be straightforwardly implemented on current and near-term devices. REM can be applied alongside existing mitigation procedures, while requiring minimal post-processing and only one or no additional measurements. The approach is agnostic to the underlying quantum mechanical ansatz and is designed for the variational quantum eigensolver. Up to two orders-of-magnitude improvement in the computational accuracy of ground state energies of small molecules (H_2, HeH^+, and LiH) is demonstrated on superconducting quantum hardware. Simulations of noisy circuits with a depth exceeding 1000 two-qubit gates are used to demonstrate the scalability of the method.



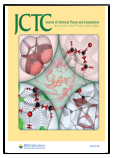
INTRODUCTION

Quantum computers hold a potential for solving problems that are intractable on current and future computers.[1,2] Quantum chemistry is one of the research areas where *quantum advantage* is expected in the near future.[3−6] One of the major challenges in realizing practical quantum computation of chemistry is the sensitivity of quantum devices to noise. Errors due to noise can be caused by several factors such as spontaneous emission, control and measurement imperfection, and unwanted coupling with the environment.[7] Whereas reliable error correction is expected in future quantum computers, such *fault-tolerant* machines will put high demands on both quality and number of physical qubits.[4] Increasingly robust hybrid algorithms[2,8−10] are being designed for quantum chemistry on near-term, noisy intermediate-scale quantum (NISQ) devices.[11] Unfortunately, noise causes such algorithms to produce results, such as energies of molecules, that are of relatively low quality, even as they rely on shallow quantum circuits.[12−15] We will return to discuss how one can define quality in terms of accuracy and precision in this context.

The general challenge of noise in quantum hardware has motivated the development of several methods for *error mitigation*:[16] readout/measurement error mitigation,[17] zero noise/Richardson extrapolation,[18,19] Clifford data regression,[20] training by fermionic linear optics (TFLO),[21] rescaling as per Arute et al.,[22] probabilistic error cancellation,[23] quantum subspace expansion,[24] postselection,[25] McWeeny purification,[26] virtual state distillation,[27] and symmetry verification[28] are some examples of techniques exploited to improve the quality of measurements of encoded Hamiltonians through pre- or postprocessing. Some of these techniques have been shown to offer improvements when computing energies of small molecules with variational algorithms.[26,29] A combination of mitigation strategies is often a good approach to minimize errors.

In this study, we report on a chemistry-inspired error-mitigation strategy that can be combined with any variant of the variational quantum eigensolver[9,30] (VQE). Our approach, reference-state error mitigation (REM), relies on postprocessing that can be readily performed on a classical computer. The method is applicable across a wide range of noise intensities and is low-cost in that it requires an overhead of at most one additional VQE energy evaluation. REM can readily be employed together with other error mitigation methods, and throughout this work we additionally use readout mitigation, which corrects for hardware-specific nonideal correlation between prepared and measured states.[17,22] Readout mitigation works by performing an initial calibration against a subset of

Received: August 4, 2022
Published: January 27, 2023



II

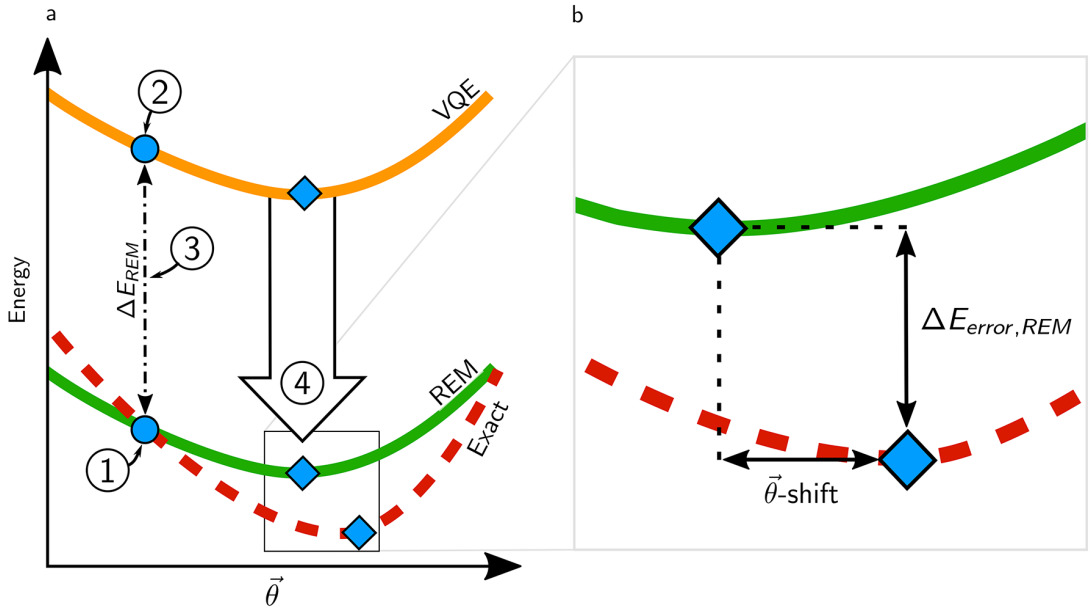


Figure 1. (a) One-dimensional representation of the electronic energy E as a function of quantum circuit parameters $\vec{\theta}$. The REM approach can be explained as a four-step process: (1) A computationally tractable reference solution (such as Hartree−Fock) is computed on a classical computer. (2) A quantum measurement of the VQE surface (orange) is made using parameters corresponding to the reference solution. (3) The difference in energy calculated for the reference solution on classical and quantum hardware defines an error, ΔE_{REM}. (4) The error estimate is assumed to be systematic and is used to correct the VQE surface in the proximity of the reference coordinate. The resulting REM corrected surface and the exact (noise-free) solution are represented by solid green and dashed red lines, respectively. Blue dots represent the coordinates of the reference calculations, while diamonds indicate minima of the different energy landscapes. (b) The inset shows the remaining error after application of REM. A possible difference in the location of the minima is indicated by $\vec{\theta}$-shift (b). The difference in energy between the VQE and REM minima is indicated by $\Delta E_{error,REM}$.

Clifford gates, whereby known states are prepared and measured; see the Supporting Information (SI).

The evaluation of *computational accuracy* in this work should not be confused with *chemical accuracy*.[31] We here use the term *computational accuracy* specifically when comparing results of a quantum calculation with the exact solution *at that same level of theory*. Computational accuracy then, in the context of VQE calculations, refers to how accurate a given VQE problem is solved with respect to the given Hamiltonian and ansatz. This accuracy can only be quantified so long as it is possible to solve the problem without noise, e.g., using conventional quantum chemistry (which we can still do; discussing the limits of conventional quantum chemistry methods is outside the scope of this work). Practical implementations of VQE on real hardware currently suffer from drastic deficiencies in level of theory, basis set size,[32] proper consideration of the physical environment (e.g., solvent effects), and dynamical effects. These limitations keep quantum computation (including our own) from accurately predicting real chemical processes. On the other hand, *chemical accuracy* is the correct term for what is required to make realistic predictions and is commonly defined as an error of 1 kcal/mol (~1.6 millihartree) from the exact solution.[9,15,26,33] We encourage the community to use the appropriate terminology. The hunt for chemical accuracy in the NISQ era is far from over.

METHODS

The goal of the VQE algorithm is to minimize the electronic energy with respect to a set of quantum circuit parameters, i.e.,

$$E_{VQE} = \min_{\vec{\theta}} E(\vec{\theta}) = \min_{\vec{\theta}} \langle \Psi(\vec{\theta}) | \hat{H} | \Psi(\vec{\theta}) \rangle \quad (1)$$

where $\vec{\theta} = [\theta_1, \theta_2, ..., \theta_n]$, $\hat{H}$ is the molecular Hamiltonian, and $|\Psi(\vec{\theta})\rangle$ represents the parametrized trial state generated by the VQE circuit. The VQE energy, $E_{VQE}(\vec{\theta})$, can therefore be thought of as living on an n-dimensional surface in parameter space. A one-dimensional representation of such a surface is shown in orange at the top of Figure 1. The $E_{VQE}(\vec{\theta})$ surface is associated with some degree of systematic and random noise that can only be partially removed by accounting for state-preparation and measurement errors through readout mitigation.

As the name suggests, the REM method rests on an appropriate choice of a reference wave function, or reference state. We recommend the reference state is (a) chemically motivated, i.e., likely physically similar to the sought state, and (b) fast (or at least viable) to evaluate using a classical computer. These properties are often also desired in the choice of initial guess for VQE, thus the initial point for optimization is commonly a good choice of reference state. The Hartree−Fock state is a practical example of an often-suitable reference wave function that is based on a computationally efficient mean-field description of the electronic potential. We rely on Hartree−Fock as the reference state in this work, but we will investigate different reference choices, better suited to more strongly correlated problems, in future work.

For the REM method to be practically useful, the cost associated with calculating the reference state has to be *lower* than the actual VQE calculation. One way to evaluate the relative cost of classical as compared to quantum computation is to compare their respective computational complexity. A

https://doi.org/10.1021/acs.jctc.2c00807
J. Chem. Theory Comput. 2023, 19, 783−789

Table 1. Total Ground State Energies of Molecules at Experimental Equilibrium Distances[a], without and with the Application of REM[b]

Molecule[a]	$E_{exact}(\vec{\theta}_{min})$	$E_{VQE}(\vec{\theta}_{min,VQE})$	E_{REM}	$\Delta E_{error,VQE}$	$\Delta E_{error,REM}$
H_2[c]	−1.1373	−1.1085(60)	−1.1355	0.029(6)	0.002(8)
HeH^+[d]	−2.8542	−2.825(4)	−2.853(6)	0.029(4)	0.003(6)
LiH[d]	−7.8787	−7.599(33)	−7.852(62)	0.280(33)	0.029(62)
LiH[e]	−7.8811	−7.360(4)	−7.871(7)	0.521(4)	0.011(7)
BeH_2[e]	−15.5895	−13.987(5)	−15.563(10)	1.602(5)	0.0263(10)

[a]Calculations refer to experimental bond distances from the National Institute of Standards and Technology (NIST).[36] Details on measurement and confidence bounds are provided in the SI. [b]Readout mitigation has been applied for all VQE calculations. All energies are given in hartree. Associated standard deviations are provided for all errors. [c]Run on Chalmers Särimner with 5000 samples as a single point measurement without optimization. The sample size of 5000 is motivated by our previous experience with the device. The computation on the Särimner device was performed as a complete sweep of the single variational parameter in the circuit, and not as a VQE optimization. Therefore, the energy variance must be approximated through other means, which we describe in the SI. [d]Run on ibmq_quito with 8192 samples, the maximum allowed. [e]Simulated results using a noise model from ibmq_athens.

lower bound estimate of the cost of VQE is the number of measurements required to evaluate the energy. In an ideal scenario, i.e., without noise, such measurements approximately scale as $O(n^4)$, where n is the number of basis functions. In contrast, the cost of conventional Hartree−Fock calculations have a practical scaling between $O(n^2)$ and $O(n^4)$.[34]

Once the parametrized reference state $|\Psi(\vec{\theta}_{ref})\rangle$ is prepared, a determination of the resulting energy error ΔE_{REM} at the reference parameters can be made,

$$\Delta E_{REM} = E_{VQE}(\vec{\theta}_{ref}) - E_{exact}(\vec{\theta}_{ref}) \tag{2}$$

where $E_{exact}(\vec{\theta}_{ref})$ is the exact solution (up to numerical precision) for the reference state, evaluated on a classical computer. $E_{VQE}(\vec{\theta}_{ref})$ refers to the energy evaluated from measurements on a quantum computer at the reference parameter value, $\vec{\theta}_{ref}$. The idea of approaching error mitigation by comparing noisy measurements to noise-free tractable classical calculations shares some commonality with mitigation techniques such as Clifford data regression, TFLO, and the rescaling technique of Arute et al.[22] However, our method is distinct from these methods in that REM only relies on the use of *a single* conventional calculation to generate a reference state. Because REM incorporates mitigation through the choice of a chemically and physically motivated initial guess for the VQE algorithm it does not require training on a large number of measured expectation values, as Clifford data regression and TFLO do. REM also considers the effects of both circuit depth and composition, which are often overlooked by other methods.

It is important to make the distinction between the initial state in the VQE calculation and the reference state, which need not be the same. The reference state can either be a part of the VQE optimization or be prepared and measured separately from the variational procedure. Provided that the reference state is also used as an initial guess for the VQE algorithm, it is possible to perform REM without incurring *any* additional measurement cost.

The exact energy at any arbitrary coordinate, $E_{exact}(\vec{\theta})$, can be expressed as

$$E_{exact}(\vec{\theta}) = E_{VQE}(\vec{\theta}) - \Delta E_{REM} - \Delta E_p(\vec{\theta}) \tag{3}$$

where $\Delta E_p(\vec{\theta})$ includes any parameter-dependence of noise present and $\Delta E_p(\vec{\theta}_{ref}) = 0$. The underlying assumption of the REM method is that such parameter dependence of the noise is negligible close to the reference geometry, i.e., $\lim_{\Delta\vec{\theta}\to 0} \Delta E_p(\vec{\theta}) = 0$ where $\Delta\vec{\theta} = |\vec{\theta} - \vec{\theta}_{ref}|$. In other words, the effectiveness of the REM approach can be assumed dependent on the Euclidean distance of the reference state to the exact solution $|\vec{\theta}_{exact} - \vec{\theta}_{ref}|$, given that both are in the same convex region of the energy surface. When this approximation fails, noise can shift features in the energy surface, such as the optimal coordinates identified using the VQE algorithm, $\vec{\theta}_{min,VQE}$, away from the true minimum, $\vec{\theta}_{min,exact}$ (Figure 1b), resulting in a $\vec{\theta}$-shift. When evaluating our method, we will not quantify $\Delta E_p(\vec{\theta})$ but instead compare energies obtained for the two minima on the exact and the VQE surface,

$$\Delta E_{error,VQE} = E_{VQE}(\vec{\theta}_{min,VQE}) - E_{exact}(\vec{\theta}_{min,exact}) \tag{4}$$

In eq 4, $E_{exact}(\vec{\theta}_{min,exact})$ is the exact solution obtained by an ideal noise-free VQE optimization and $E_{VQE}(\vec{\theta}_{min,VQE})$ is the energy of a converged noisy VQE optimization. The error remaining after applying REM to a converged noisy VQE optimization is

$$\Delta E_{error,REM} = \Delta E_{REM} - \Delta E_{error,VQE} \tag{5}$$

RESULTS AND DISCUSSION

To assess the reliability of REM, we have implemented it for the ground state energy computation of small molecules on two current NISQ devices, the ibmq_quito of IBMQ and the Särimner device of Chalmers University. Details of hardware, circuits, measurements and estimates on confidence bounds are provided in the SI. Table 1 shows an amalgamation of our measurement results for the ground state energy of the hydrogen molecule (H_2), helium hydride (HeH^+), and lithium hydride (LiH). The ansatze[3] used for the H_2 and HeH^+ molecules are chemistry-inspired and based on unitary coupled cluster theory,[35] whereas a hardware-efficient ansatz is used for LiH. Table 1 also includes results of simulations of LiH and beryllium hydride (BeH_2). The latter circuits are substantially larger than what is feasible on current devices as they would incur insurmountable errors due to noise. Combined, this test set ranges from a two-qubit circuit with just 1 two-qubit gate for H_2, to a six-qubit circuit with 1096 two-qubit gates for BeH_2 (Table 2). Our simulations of BeH_2 contains 26 variational parameters (Table 2).

Table 1 shows how the application of REM reduces the error by up to 2 orders of magnitude compared to regular VQE when used together with readout mitigation. Without readout mitigation, VQE errors are substantially larger (Tables S3 and

https://doi.org/10.1021/acs.jctc.2c00807
J. Chem. Theory Comput. 2023, 19, 783−789

II

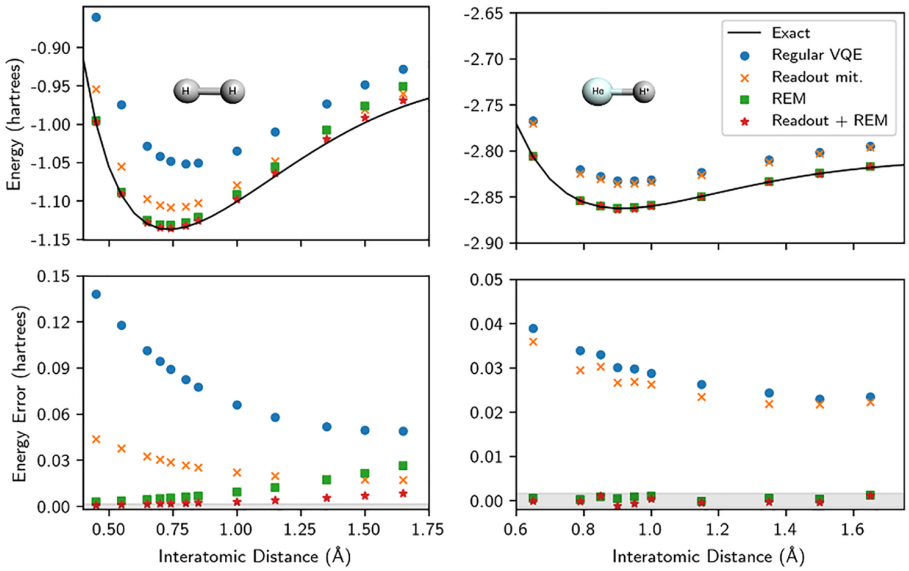


Figure 2. Top: Potential energy surfaces for the dissociation of H_2 and HeH^+. Exact noise-free solutions from state-vector simulations are represented by black lines. Regular VQE energies obtained using a quantum computer are shown as blue dots. Results following readout mitigation and REM are shown as orange crosses and green squares, respectively. The combination of both readout mitigation and REM is shown as red stars. Measurements for H_2 and HeH^+ were performed on Chalmers Särimner and ibmq_quito, respectively. Bottom: error of the different approaches relative to the exact solution in the given minimal basis set. The gray region corresponds to an error of 1.6 millihartree (1 kcal/mol) with respect to the corresponding noise-free calculations.

S6). The examples in Table 1 are sorted by increasing circuit depth (see also Table 2 for details), which indicate both the robustness and scalability of the approach. The remaining error after mitigation is consistently on the order of millihartree, and the magnitude of the REM correction grows with the complexity of the quantum circuit (cf. H_2 vs BeH_2 in Table 1).

In principle, unsuitable choices of reference states combined with significant parameter dependence of noise, $\Delta E_p(\vec{\theta}) \not\approx 0$, might result in overcorrection of the measured VQE energy, taking the energy below the true minimum, as can be seen in some results for the dissociation curve of HeH^+ (Figures 2 and 3). Nevertheless, we note that REM consistently improves the measured energies, even at relatively high noise levels (Figure 3) and drastically improves the computational accuracy for all calculations summarized in Table 1.

Testing the Limits of REM. Table 1 demonstrates the effectiveness of REM when applied to molecules in their equilibrium geometry. The test set is small, in practice limited to what is feasible to run on current NISQ hardware. These geometries also represent situations where the degree of electron correlation is relatively low.

To investigate how the REM method performs out of equilibrium, we show in Figure 2 the bond dissociation of H_2 and HeH^+. For HeH^+, which dissociates to He and an isolated proton—a state well described by a single-reference Hartree–Fock description—the REM method provides highly accurate results across the entire binding curve. For H_2 on the other hand, the effectiveness of the current implementation of REM decreases in regions where the Hartree–Fock state offers a poor description, such as the stretched H_2 bond. Nevertheless, the method consistently provides a substantial improvement across the potential energy surface. A more suitable description of the partially broken bond of H_2 should ideally account for the static correlation arising due to near degeneracy of multiple states; i.e., it would require a multireference (MR) or open-shell (OS) reference. We will investigate the use of such MR/OS states and an adaptive choice of the most suitable single reference (SR) state within the REM framework in an upcoming study.

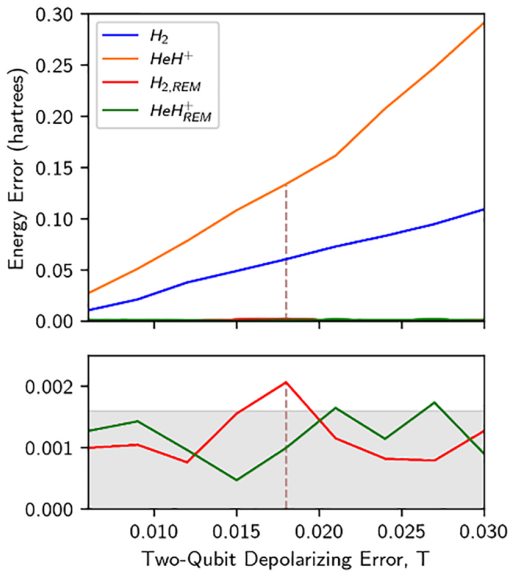


Figure 3. Absolute errors as a function of increasing depolarizing errors for ground state calculations of H_2 and HeH^+. The two-qubit depolarizing error of the Chalmers Särimner device is indicated by a vertical line for reference. Bottom inset: Energy errors after application of REM are largely independent of the noise level, and the computational accuracy is consistently close to or below 1.6 millihartree (1 kcal/mol) for these molecules.

Figure 2 also illustrates the effect of readout mitigation,[17] which we use per default in all measurements and that we

https://doi.org/10.1021/acs.jctc.2c00807
J. Chem. Theory Comput. 2023, 19, 783–789

recommend together with REM. Other mitigation strategies may, in principle, also be combined with REM.

The robustness of REM was further evaluated by performing simulations of H_2 and HeH^+ while varying the noise level. Noise was introduced in these simulations by modeling imperfect gate-fidelities as single (S) and two-qubit (T) depolarizing errors, connected through a linear relationship, $S = 0.1\ T$ (see the SI). This kind of noise modeling enables straightforward comparison with error rates on physical quantum devices (Figure 3). REM is shown to be effective despite the steady increase in single- and two-qubit depolarizing error rates.

CONCLUSIONS

In this work we demonstrate an error mitigation strategy applicable to quantum chemical computations on NISQ devices. The REM method relies on accurately determining the error in energy due to hardware and environmental noise for a reference wave function that can be feasibly evaluated on a classical computer. The underlying assumption of REM is a negligible dependence of noise on circuit parameters in the vicinity of this reference wave function. In this work Hartree–Fock references are used, which describes the physics of the molecular states well enough for REM to perform effectively. The REM method is shown to drastically improve the computational accuracy at which total energies of molecules can be computed using current quantum hardware. REM is well suited for calculations with significant amounts of noise and improve calculated energies in all herein tested cases.

The performance of REM is dependent on the quality of the supplied reference state. A Hartree–Fock (mean-field) solution is expected to provide a sufficient reference for molecules that do not exhibit large multireference character. We will investigate the use of references based on multi-reference and open-shell states, better suited for more strongly correlated problems, in an upcoming study.

In our herein studied problems, the computational accuracy is improved by up to 2 orders of magnitude after application of REM. However, in the presence of substantial quantum circuit parameter dependence of noise it cannot be ruled out that REM may underestimate the true energy. The nonvariational nature of error mitigation strategies remains a problem to be solved. No single mitigation technique will completely resolve the issue of noise, and REM is no exception. One strength of REM is its ability to be combined with other error mitigation techniques without incurring additional cost. REM does not incur meaningful additional classical or quantum computational overhead and can be used to reduce errors on near-term devices by orders of magnitude when running VQE calculations. Because error rates vary both between NISQ devices and between circuits, it is not currently productive to rely on error cancellation, i.e., systematic errors inherent in quantum chemical levels of theory, when evaluating relative energies of chemical transformations. By enabling more precise evaluations of molecular total energies, REM moves us toward meaningful relative comparisons and toward chemical accuracy.

APPENDIX: CIRCUIT COMPLEXITY DETAILS

Table 2 is presented here.

Table 2. Comparison of Circuit Complexities for H_2, HeH^+, LiH, and BeH_2 Listed in Order of Increasing Complexity[a]

Molecule	Ansatz	Qubits	Depth	Two-Qubit Gates	Parameters
H_2	Simplified UCCD	2	5	1	1
HeH^+	UCCSD	2	14	4	2
LiH	Hardware-efficient	4	9	6	8
LiH	UCCSD	4	275	172	8
BeH_2	UCCSD	6	1480	1096	26

[a]Details are provided in the SI.

ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acs.jctc.2c00807.

Details of the Chalmers device; description of read out error mitigation; details for H_2, HeH^+, LiH, and BeH_2 calculations, including quantum circuits and Hamiltonians; simulation details for LiH and BeH_2; calibration details for ibmq_quito device; and description of depolarizing noise model used for simulations (PDF)

AUTHOR INFORMATION

Corresponding Author

Martin Rahm − *Department of Chemistry and Chemical Engineering, Chalmers University of Technology, SE-412 96 Gothenburg, Sweden;* orcid.org/0000-0001-7645-5923; Email: martin.rahm@chalmers.se

Authors

Phalgun Lolur − *Department of Chemistry and Chemical Engineering, Chalmers University of Technology, SE-412 96 Gothenburg, Sweden;* Present Address: (P.L.) National Quantum Computing Centre (NQCC), Science and Technology Facilities Council (STFC), UK Research and Innovation (UKRI), Rutherford Appleton Laboratory, Harwell Campus, Didcot, Oxfordshire OX11 0QX, United Kingdom; orcid.org/0000-0002-3409-3415

Mårten Skogh − *Department of Chemistry and Chemical Engineering, Chalmers University of Technology, SE-412 96 Gothenburg, Sweden; Data Science & Modelling, Pharmaceutical Science, R&D, AstraZeneca, SE-431 83 Mölndal, Gothenburg, Sweden*

Werner Dobrautz − *Department of Chemistry and Chemical Engineering, Chalmers University of Technology, SE-412 96 Gothenburg, Sweden;* orcid.org/0000-0001-6479-1874

Christopher Warren − *Department of Microtechnology and Nanoscience MC2, Chalmers University of Technology, SE-412 96 Gothenburg, Sweden*

Janka Biznárová − *Department of Microtechnology and Nanoscience MC2, Chalmers University of Technology, SE-412 96 Gothenburg, Sweden*

Amr Osman − *Department of Microtechnology and Nanoscience MC2, Chalmers University of Technology, SE-412 96 Gothenburg, Sweden*

Giovanna Tancredi − *Department of Microtechnology and Nanoscience MC2, Chalmers University of Technology, SE-412 96 Gothenburg, Sweden*

https://doi.org/10.1021/acs.jctc.2c00807
J. Chem. Theory Comput. 2023, 19, 783−789

Göran Wendin − *Department of Microtechnology and Nanoscience MC2, Chalmers University of Technology, SE-412 96 Gothenburg, Sweden*
Jonas Bylander − *Department of Microtechnology and Nanoscience MC2, Chalmers University of Technology, SE-412 96 Gothenburg, Sweden*

Complete contact information is available at: https://pubs.acs.org/10.1021/acs.jctc.2c00807

II

Author Contributions

Conceptualization: P.L., M.S., M.R. Funding acquisition: G.W., J.By. Investigation (Theoretical): P.L., M.S. Investigation (Experimental): C.W., G.T. Device Fabrication: J.Bi., A.O. Supervision: M.R., J.By., G.W. Writing: M.R., P.L., M.S., W.D., C.W., J.By.

Author Contributions

‖P.L. and M.S. contributed equally to this work.

Funding

This research has been supported by funding from the Wallenberg Center for Quantum Technology (WACQT) and from the EU Flagship on Quantum Technology H2020-FETFLAG-2018-03 Project 820363 OpenSuperQ. W.D. acknowledges funding from the European Union's Horizon Europe research and innovation program under the Marie Skłodowska-Curie grant agreement No. 101062864. This research relied on computational resources provided by the Swedish National Infrastructure for Computing (SNIC) at C3SE, NSC, and PDC partially funded by the Swedish research council through Grant Agreement No. 2018-05973.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

We also acknowledge experimental assistance from Christian Križan and Andreas Bengtsson as well as insights from Jorge Fernández Pendás at Chalmers University of Technology. The Chalmers device was made at Myfab Chalmers; it was packaged in a holder and printed circuit board with original designs shared by the Quantum Device Lab at ETH Zürich.

■ ABBREVIATIONS

BeH_2, beryllium hydride; H_2, hydrogen molecule; HeH^+, helium hydride cation; LiH, lithium hydride; NISQ, noisy intermediate-scale quantum; NIST, National Institute for Standards and Technology; REM, reference-state error mitigation; TFLO, training by Fermionic linear optics; VQE, variational quantum eigensolver; MR, multireference; OS, open-shell; SR, single-reference.

■ REFERENCES

(1) Feynman, R. P. Simulating Physics with Computers. *International Journal of Theoretical Physics* **1982**, *21* (6), 467−488.
(2) Abrams, D. S.; Lloyd, S. Quantum Algorithm Providing Exponential Speed Increase for Finding Eigenvalues and Eigenvectors. *Phys. Rev. Lett.* **1999**, *83* (24), 5162−5165.
(3) McArdle, S.; Endo, S.; Aspuru-Guzik, A.; Benjamin, S. C.; Yuan, X. Quantum Computational Chemistry. *Rev. Mod. Phys.* **2020**, *92* (1), No. 015003.
(4) Cao, Y.; Romero, J.; Olson, J. P.; Degroote, M.; Johnson, P. D.; Kieferová, M.; Kivlichan, I. D.; Menke, T.; Peropadre, B.; Sawaya, N. P. D.; Sim, S.; Veis, L.; Aspuru-Guzik, A. Quantum Chemistry in the Age of Quantum Computing. *Chem. Rev.* **2019**, *119* (19), 10856−10915.
(5) Lanyon, B. P.; Whitfield, J. D.; Gillett, G. G.; Goggin, M. E.; Almeida, M. P.; Kassal, I.; Biamonte, J. D.; Mohseni, M.; Powell, B. J.; Barbieri, M.; Aspuru-Guzik, A.; White, A. G. Towards Quantum Chemistry on a Quantum Computer. *Nat. Chem.* **2010**, *2* (2), 106−111.
(6) Cerezo, M.; Arrasmith, A.; Babbush, R.; Benjamin, S. C.; Endo, S.; Fujii, K.; McClean, J. R.; Mitarai, K.; Yuan, X.; Cincio, L.; Coles, P. J. Variational Quantum Algorithms. *Nature Reviews Physics* **2021**, *3* (9), 625−644.
(7) Unruh, W. G. Maintaining Coherence in Quantum Computers. *Phys. Rev. A (Coll Park)* **1995**, *51* (2), 992−997.
(8) Hadfield, S.; Wang, Z.; O'Gorman, B.; Rieffel, E. G.; Venturelli, D.; Biswas, R. From the Quantum Approximate Optimization Algorithm to a Quantum Alternating Operator Ansatz. *Algorithms* **2019**, *12* (2), 34.
(9) Peruzzo, A.; McClean, J.; Shadbolt, P.; Yung, M.-H.; Zhou, X.-Q.; Love, P. J.; Aspuru-Guzik, A.; O'Brien, J. L. A Variational Eigenvalue Solver on a Quantum Processor. *Nat. Commun.* **2014**, *5* (1), 4213.
(10) Farhi, E.; Goldstone, J.; Gutmann, S. A Quantum Approximate Optimization Algorithm. *arXiv*, 1411.4028 [quant-ph], November 14, 2014, https://arxiv.org/abs/1411.4028 (accessed 2023-01-10).
(11) Preskill, J. Quantum Computing in the NISQ Era and Beyond. *Quantum* **2018**, *2* (July), 79.
(12) Hempel, C.; Maier, C.; Romero, J.; McClean, J.; Monz, T.; Shen, H.; Jurcevic, P.; Lanyon, B. P.; Love, P.; Babbush, R.; Aspuru-Guzik, A.; Blatt, R.; Roos, C. F. Quantum Chemistry Calculations on a Trapped-Ion Quantum Simulator. *Phys. Rev. X* **2018**, *8* (3), 031022.
(13) Sawaya, N. P. D.; Smelyanskiy, M.; McClean, J. R.; Aspuru-Guzik, A. Error Sensitivity to Environmental Noise in Quantum Circuits for Chemical State Preparation. *J. Chem. Theory Comput* **2016**, *12* (7), 3097−3108.
(14) Piveteau, C.; Sutter, D.; Bravyi, S.; Gambetta, J. M.; Temme, K. Error Mitigation for Universal Gates on Encoded Qubits. *Phys. Rev. Lett.* **2021**, *127* (20), 200505.
(15) Kandala, A.; Mezzacapo, A.; Temme, K.; Takita, M.; Brink, M.; Chow, J. M.; Gambetta, J. M. Hardware-Efficient Variational Quantum Eigensolver for Small Molecules and Quantum Magnets. *Nature* **2017**, *549* (7671), 242−246.
(16) Suzuki, Y.; Endo, S.; Fujii, K.; Tokunaga, Y. Quantum Error Mitigation as a Universal Error Reduction Technique: Applications from the NISQ to the Fault-Tolerant Quantum Computing Eras. *PRX Quantum* **2022**, *3* (1), No. 010345.
(17) Bravyi, S.; Sheldon, S.; Kandala, A.; McKay, D. C.; Gambetta, J. M. Mitigating Measurement Errors in Multiqubit Experiments. *Phys. Rev. A (Coll Park)* **2021**, *103* (4), 042605.
(18) Li, Y.; Benjamin, S. C. Efficient Variational Quantum Simulator Incorporating Active Error Minimization. *Phys. Rev. X* **2017**, *7* (2), 021050.
(19) Giurgica-Tiron, T.; Hindy, Y.; Larose, R.; Mari, A.; Zeng, W. J. Digital Zero Noise Extrapolation for Quantum Error Mitigation. *Proceedings - IEEE International Conference on Quantum Computing and Engineering, QCE 2020*; IEEE: 2020; pp 306−316.
(20) Czarnik, P.; Arrasmith, A.; Coles, P. J.; Cincio, L. Error Mitigation with Clifford Quantum-Circuit Data. *Quantum* **2021**, *5*, 592.
(21) Montanaro, A.; Stanisic, S. Error Mitigation by Training with Fermionic Linear Optics. *arXiv*, 2102.02120 [quant-ph], February 2021, http://arxiv.org/abs/2102.02120 (accessed 2023-01-09).
(22) Arute, F.; Arya, K.; Babbush, R.; Bacon, D.; Bardin, J. C.; Barends, R.; Bengtsson, A.; Boixo, S.; Broughton, M.; Buckley, B. B.; Buell, D. A.; Burkett, B.; Bushnell, N.; Chen, Y.; Chen, Z.; Chen, Y.-A.; Chiaro, B.; Collins, R.; Cotton, S. J.; Courtney, W.; Demura, S.; Derk, A.; Dunsworth, A.; Eppens, D.; Eckl, T.; Erickson, C.; Farhi, E.; Fowler, A.; Foxen, B.; Gidney, C.; Giustina, M.; Graff, R.; Gross, J. A.; Habegger, S.; Harrigan, M. P.; Ho, A.; Hong, S.; Huang, T.; Huggins, W.; Ioffe, L. B.; Isakov, S. v; Jeffrey, E.; Jiang, Z.; Jones, C.; Kafri, D.; Kechedzhi, K.; Kelly, J.; Kim, S.; Klimov, P. v; Korotkov, A. N.; Kostritsa, F.; Landhuis, D.; Laptev, P.; Lindmark, M.; Lucero, E.;

Marthaler, M.; Martin, O.; Martinis, J. M.; Marusczyk, A.; McArdle, S.; McClean, J. R.; McCourt, T.; McEwen, M.; Megrant, A.; Mejuto-Zaera, C.; Mi, X.; Mohseni, M.; Mruczkiewicz, W.; Mutus, J.; Naaman, O.; Neeley, M.; Neill, C.; Neven, H.; Newman, M.; Niu, M. Y.; O'Brien, T. E.; Ostby, E.; Pató, B.; Petukhov, A.; Putterman, H.; Quintana, C.; Reiner, J.-M.; Roushan, P.; Rubin, N. C.; Sank, D.; Satzinger, K. J.; Smelyanskiy, V.; Strain, D.; Sung, K. J.; Schmitteckert, P.; Szalay, M.; Tubman, N. M.; Vainsencher, A.; White, T.; Vogt, N.; Yao, Z. J.; Yeh, P.; Zalcman, A.; Zanker, S. Observation of Separated Dynamics of Charge and Spin in the Fermi-Hubbard Model. *arXiv*, 2010.07965 [quant-ph], October 2020, http://arxiv.org/abs/2010.07965 (accessed 2023-01-10).

(23) Temme, K.; Bravyi, S.; Gambetta, J. M. Error Mitigation for Short-Depth Quantum Circuits. *Phys. Rev. Lett.* **2017**, *119* (18), 180509.

(24) McClean, J. R.; Kimchi-Schwartz, M. E.; Carter, J.; de Jong, W. A. Hybrid Quantum-Classical Hierarchy for Mitigation of Decoherence and Determination of Excited States. *Phys. Rev. A (Coll Park)* **2017**, *95* (4), 042308.

(25) Aliferis, P.; Gottesman, D.; Preskill, J. Accuracy Threshold for Postselected Quantum Computation. *arXiv*, quant-ph/0703264, September 2007, http://arxiv.org/abs/quant-ph/0703264 (accessed 2023-01-09).

(26) Arute, F.; Arya, K.; Babbush, R.; Bacon, D.; Bardin, J. C.; Barends, R.; Boixo, S.; Broughton, M.; Buckley, B. B.; Buell, D. A.; Burkett, B.; Bushnell, N.; Chen, Y.; Chen, Z.; Chiaro, B.; Collins, R.; Courtney, W.; Demura, S.; Dunsworth, A.; Farhi, E.; Fowler, A.; Foxen, B.; Gidney, C.; Giustina, M.; Graff, R.; Habegger, S.; Harrigan, M. P.; Ho, A.; Hong, S.; Huang, T.; Huggins, W. J.; Ioffe, L.; Isakov, S. v; Jeffrey, E.; Jiang, Z.; Jones, C.; Kafri, D.; Kechedzhi, K.; Kelly, J.; Kim, S.; Klimov, P. v; Korotkov, A.; Kostritsa, F.; Landhuis, D.; Laptev, P.; Lindmark, M.; Lucero, E.; Martin, O.; Martinis, J. M.; McClean, J. R.; McEwen, M.; Megrant, A.; Mi, X.; Mohseni, M.; Mruczkiewicz, W.; Mutus, J.; Naaman, O.; Neeley, M.; Neill, C.; Neven, H.; Niu, M. Y.; O'Brien, T. E.; Ostby, E.; Petukhov, A.; Putterman, H.; Quintana, C.; Roushan, P.; Rubin, N. C.; Sank, D.; Satzinger, K. J.; Smelyanskiy, V.; Strain, D.; Sung, K. J.; Szalay, M.; Takeshita, T. Y.; Vainsencher, A.; White, T.; Wiebe, N.; Yao, Z. J.; Yeh, P.; Zalcman, A. Hartree-Fock on a Superconducting Qubit Quantum Computer. *Science* **2020**, *369* (6507), 1084−1089.

(27) Huggins, W. J.; McArdle, S.; O'Brien, T. E.; Lee, J.; Rubin, N. C.; Boixo, S.; Whaley, K. B.; Babbush, R.; McClean, J. R. Virtual Distillation for Quantum Error Mitigation. *Phys. Rev. X* **2021**, *11* (4), 041036.

(28) Bonet-Monroig, X.; Sagastizabal, R.; Singh, M.; O'Brien, T. E. Low-Cost Error Mitigation by Symmetry Verification. *Phys. Rev. A* **2018**, *98* (6), 062339.

(29) McCaskey, A. J.; Parks, Z. P.; Jakowski, J.; Moore, S. v.; Morris, T. D.; Humble, T. S.; Pooser, R. C. Quantum Chemistry as a Benchmark for Near-Term Quantum Computers. *npj Quantum Inf* **2019**, *5* (1), 99.

(30) McClean, J. R.; Romero, J.; Babbush, R.; Aspuru-Guzik, A. The Theory of Variational Hybrid Quantum-Classical Algorithms. *New J. Phys.* **2016**, *18* (2), 023023.

(31) Hoffmann, R.; Schleyer, P. V. R.; Schaefer, H. F. Predicting Molecules - More Realism, Please! *Angewandte Chemie - International Edition* **2008**, *47* (38), 7164−7167.

(32) Elfving, V. E.; Broer, B. W.; Webber, M.; Gavartin, J.; Halls, M. D.; Lorton, K. P.; Bochevarov, A. How Will Quantum Computers Provide an Industrially Relevant Computational Advantage in Quantum Chemistry? *arXiv*, 2009.12472, September 2020, http://arxiv.org/abs/2009.12472 (accessed 2023-01-10).

(33) O'Malley, P. J. J.; Babbush, R.; Kivlichan, I. D.; Romero, J.; McClean, J. R.; Barends, R.; Kelly, J.; Roushan, P.; Tranter, A.; Ding, N.; Campbell, B.; Chen, Y.; Chen, Z.; Chiaro, B.; Dunsworth, A.; Fowler, A. G.; Jeffrey, E.; Lucero, E.; Megrant, A.; Mutus, J. Y.; Neeley, M.; Neill, C.; Quintana, C.; Sank, D.; Vainsencher, A.; Wenner, J.; White, T. C.; Coveney, P. V.; Love, P. J.; Neven, H.; Aspuru-Guzik, A.; Martinis, J. M. Scalable Quantum Simulation of Molecular Energies. *Phys. Rev. X* **2016**, *6* (3), 031007.

(34) Köppl, C.; Werner, H.-J. Parallel and Low-Order Scaling Implementation of Hartree−Fock Exchange Using Local Density Fitting. *J. Chem. Theory Comput* **2016**, *12* (7), 3122−3134.

(35) Taube, A. G.; Bartlett, R. J. New Perspectives on Unitary Coupled-Cluster Theory. *Int. J. Quantum Chem.* **2006**, *106* (15), 3393−3401.

(36) Johnson, R. D., III *NIST Computational Chemistry Comparison and Benchmark Database*. NIST Standard Reference Database Number 101. http://cccbdb.nist.gov/ (accessed 2022-11-22).

II

https://doi.org/10.1021/acs.jctc.2c00807
J. Chem. Theory Comput. 2023, 19, 783−789

Supporting Information

Reference-State Error Mitigation: A Strategy for High Accuracy Quantum Computation of Chemistry

Phalgun Lolur[1], Mårten Skogh[1,2], Werner Dobrautz[1], Christopher Warren[3], Janka Biznárová[3], Amr Osman[3], Giovanna Tancredi[3], Göran Wendin[3], Jonas Bylander[3] and Martin Rahm[1,*]

1. Department of Chemistry and Chemical Engineering, Chalmers University of Technology, SE-412 96 Gothenburg, Sweden
2. Data Science & Modelling, Pharmaceutical Science, R&D, AstraZeneca, Gothenburg, SE-431 83 Mölndal, Sweden
3. Department of Microtechnology and Nanoscience MC2, Chalmers University of Technology, SE-412 96 Gothenburg, Sweden

* Corresponding author: martin.rahm@chalmers.se

II

Table of Content

Notation

The following notation is used in the SI:

$E_{exact}(\vec{\theta}_{ref})$ – Energy of the reference state in the absence of noise

$E_{VQE}(\vec{\theta}_{ref})$ – Energy of the reference state from VQE

$E_{VQE*}(\vec{\theta}_{ref})$ – Energy of the reference state from VQE with readout mitigation

$E_{exact}(\vec{\theta}_{min})$ – Energy of the state of interest in the absence of noise

$E_{VQE}(\vec{\theta}_{min,VQE})$ – Energy of the state of interest from VQE

$E_{VQE*}(\vec{\theta}_{min,VQE})$ – Energy of the state of interest from VQE with readout mitigation

E_{REM} – Energy of the state of interest from VQE with REM

E_{REM*} – Energy of the state of interest from VQE with readout mitigation and REM

$\Delta E_{error,VQE}$ – Error of the VQE energy with respect to $E_{exact}(\vec{\theta}_{\mathrm{ref}})$

$\Delta E_{error,VQE*}$ – Error of the readout mitigated VQE energy with respect to $E_{\mathrm{exact}}(\vec{\theta}_{\mathrm{ref}})$

$\Delta E_{error,REM}$ – Error of the REM mitigated VQE energy with respect to $E_{\mathrm{exact}}(\vec{\theta}_{\mathrm{ref}})$

$\Delta E_{error,REM*}$ – Error of the REM+readout mitigated VQE energy with respect to $E_{exact}(\vec{\theta}_{ref})$

r – Interatomic distance (Å)

Chalmers Device Details

We executed the quantum algorithm at Chalmers on a superconducting three-qubit quantum processor named Särimner, of which we only use two qubits (Q_0, Q_1). This device is shown in Figure S1 and consists of three transmon qubits,[1] coupled using a single tunable coupler, C_1. Single-qubit gates are implemented using on-chip drive lines to individually control each qubit with microwave pulses. Each qubit can be individually measured using its readout resonator with readout performed simultaneously using frequency multiplexed pulses on the common readout feedline. Two-qubit gates are activated via an AC flux-pulse applied to the coupler to modulate its frequency.[2] The coupler is itself a frequency-tunable transmon qubit, however, it only serves to mediate the interaction between pairs of qubits and itself never enters the computational space during operation. A full list of parameters of the device can be found in Table S1.

The AC flux-pulse which activates the coupling takes the form $\Phi(t) = \Phi_b + \Omega(t)\cos(\omega_d t)$, where Φ_b is the DC flux bias of the drive, $\Omega(t)$ is the envelope of the pulse which consists of a 25 ns cosine rise and fall and a 310 ns flat top, and ω_d is the carrier frequency. The carrier frequency is chosen such that it drives a controlled-Z (CZ) transition between Q_0 and Q_1. This is on resonance with the transition $|20\rangle \leftrightarrow |11\rangle$ and occurs at $\omega_d = |\omega_0 + \eta_0 - \omega_1|$. A full oscillation between these states brings about a conditional phase on the $|11\rangle$ state which can be calibrated to implement a CZ gate.

The parametric gate is useful in this higher connectivity architecture as the interaction between pairs can be selectively chosen. Different gates are activated between qubits when the frequencies between transitions for pairs of qubits are off-resonant from one another as they are in this device. The third qubit can also be detuned such that all transitions are far off-resonant and there is no risk of a frequency collision when operating the device with just Q_0 and Q_1.

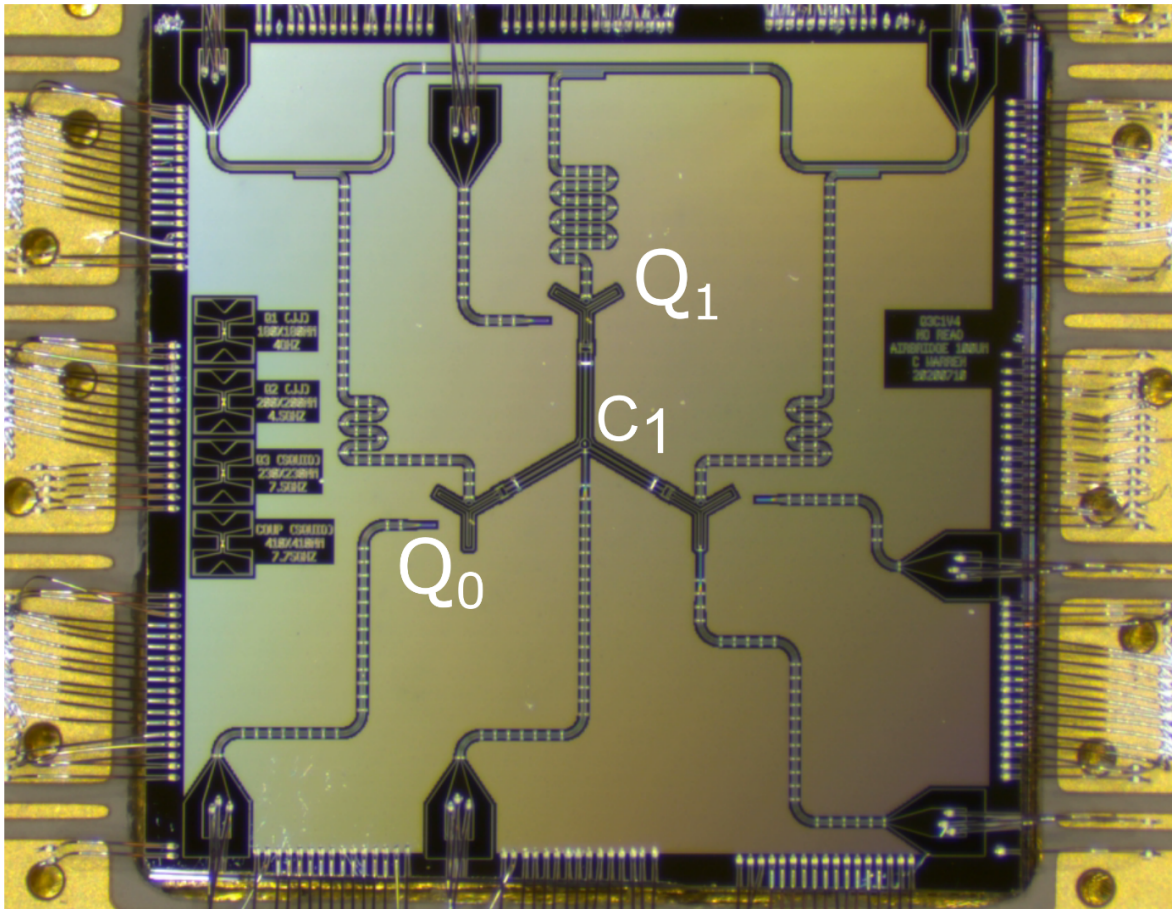


Figure S1: Microscope image of the bonded 3-qubit Särimner device from Chalmers. Only Q_0 and Q_1 are used in this work. The coupler, C_1, consists of a flux-tunable transmon qubit and serves to mediate interactions between pairs. Each qubit has individual readout and drive lines which are used to control the device. The readout lines are coupled to a shared feedline which is used for multiplexed readout.

Table S1: Experimental parameters for the 3-qubit Chalmers device Särimner. We report our average gate error achieved through randomized benchmarking for single qubit gates, as our gate set differs from that of IBM, as well as interleaved randomized benchmarking for our 2-qubit CZ gate.

Qubit	T_1 [μs]	T_2 [μs]	Qubit Frequency, ω_i [GHz]	Anharmonicity, η_i [GHz]	Single Qubit Gate Error	CZ-Gate Error
Q_0	35.98	38.74	3.799	-0.1885	4.9e-4	1.8e-2
Q_1	36.24	39.34	4.383	-0.1837	5.2e-4	1.8e-2

II

The single- and two-qubit gate fidelities during the execution of the quantum algorithm were 99.95% and 98.2%. The device design, fabrication methods, measurement setup, gate implementation on hardware, and tune-up methods are described in detail elsewhere.[3]

Readout Error Mitigation

Readout error mitigation is performed by constructing a calibration of the confusion matrix C. The entries of this matrix, $C_{i,j}$, are the probabilities of preparing the state $|i\rangle$ and then measuring the state $|j\rangle$, i.e., $C_{i,j} = P(\,j \mid i\,)$. The matrix C can then be used to correct a set of Pauli string measurements $\vec{m} = [P(0), \ldots, P(j), \ldots, P(n)]^{\mathrm{T}}$ by either multiplying by the inverse confusion matrix, $C^{-1}\vec{m} = \vec{m}'$, performing a least-squares fit to reconstruct the most likely outcome, or by a procedure known as 'Bayesian Unfolding'.[4]

In this work, we mitigate our results by implementing a least-square fit to a quadratic cost function, where $\lambda(\vec{x}) = (\vec{m} - C\vec{x})^2$ with the constraint that the sum of the resulting vector must be 1 and each element itself is bounded in the interval [0,1]. This avoids some issues resulting from matrix inversion arising from small off-diagonal elements, which can lead to unphysical results. The Sequential Least Squares Programming[5] (SLSQP) optimizer was used to find the $\vec{x}$ that minimize the cost for each set of measured Pauli strings. The confusion matrix of the Chalmers Särimner device is reported in Figure S2.

Both options for readout mitigation are also available in Qiskit and can be straightforwardly implemented for calculations on real devices and simulations.

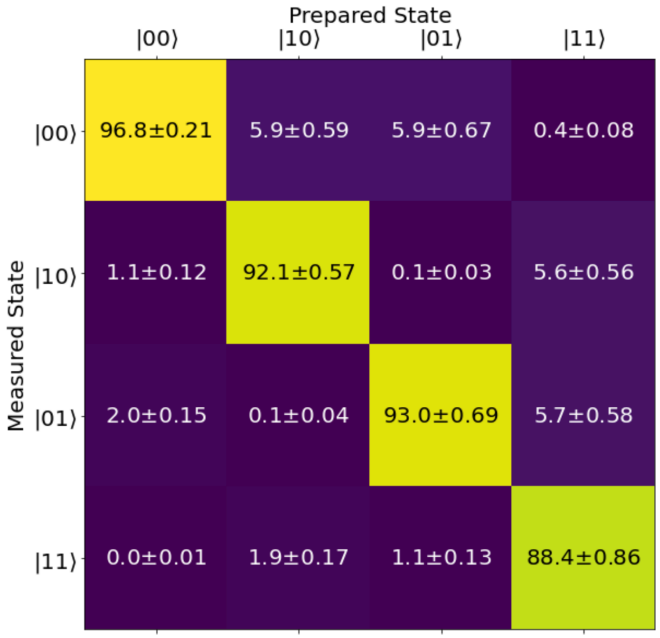


Figure S2. Confusion matrix for the Chalmers Särimner device. The matrix consists of probabilities of measuring a state given a specific preparation. The confusion matrix is measured before each run of the VQE algorithm with 1000 shots and repeated 100 times to give an estimate for fluctuations in the readout.

General Computational Details

For all calculations described below, PySCF[6] was used to generate the initial Hartree-Fock state, utilizing the STO-3G basis set. The circuit ansatz was constructed using a parameterized wavefunction based on unitary coupled cluster theory[7] as implemented in Qiskit 0.21.[8] Additionally, parity mapping[7] was used to map the fermionic spin-orbital occupation to qubits. Two-particle reduction was used to reduce the problem by two qubits in all cases, corresponding to alpha and beta spin parity conservation. For the optimization in the VQE algorithm the COBYLA optimizer was used. Unless specifically stated, default Qiskit 0.21 settings and parameters were used.

II

Hydrogen – H_2

Ansatz, circuit and computation details

The H_2 wavefunction can be represented with four qubits, where each qubit corresponds to one molecular spin orbital in minimal STO-3G basis. Since single excitations do not contribute to the final ground state energy of H_2, only double excitations were included in the ansatz which reduced the complexity further, to a single parameter, θ. The resulting circuit, compiled to Qiskit's native U1, U2 and U3 gates, is depicted in Figure S3.

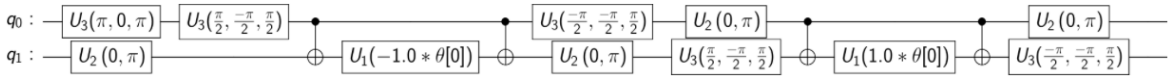


Figure S3: Quantum circuit for H_2 implementing double excitation of the unitary coupled-cluster operator with respect to the Hartree-Fock reference state. The circuit is compiled to the native gates of the IBMQ-Quito device.

The circuit shown in in Figure S3 was then shortened further by removing a repeated entangling step. A single entangling step was found sufficient to explore the Hilbert space of the problem. The resulting circuit was transpiled to gates native to the Särimner device (Figure S4) and is similar to the one used by Kandala et. Al.[9] The Särimner gate set consists of the set of single-qubit gates, $\left\{R_x(\pm\pi), R_y(\pm\pi), R_x\left(\pm\frac{\pi}{2}\right), R_y\left(\pm\frac{\pi}{2}\right), R_z(\theta)\right\}$, and two-qubit gate set {CZ}. Both circuits return the same energy up to the eighth decimal point of a hartree numerically, justifying our circuit design. Since the circuit contains only one parameter, it was varied in the interval [-π, π] to obtain energy as a function of the variational parameter at several geometries. The obtained energies were fit using the lmfit package[10] into a cosine function, A $\cos(\theta - \alpha)$, where A and α are fit parameters, to give the minima at different geometries (Figure S5). All experimental calculations were run on the Chalmers Särimner device with 5000 shots. The specific number of shots were chosen based on previous experience with the device. The details of the Hamiltonian operator to be minimized are given in Table S2 for different geometries of H_2. The exact solutions for the given circuit and Hamiltonians were calculated using QuTiP.[11]

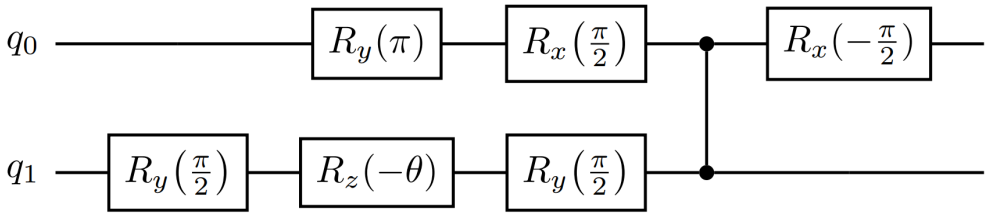


Figure S4: A compact quantum circuit for H_2 implementing double excitation of the coupled-cluster operator with respect to the Hartree-Fock reference state. The circuit is compiled to the native gates of Chalmers Särimner device.

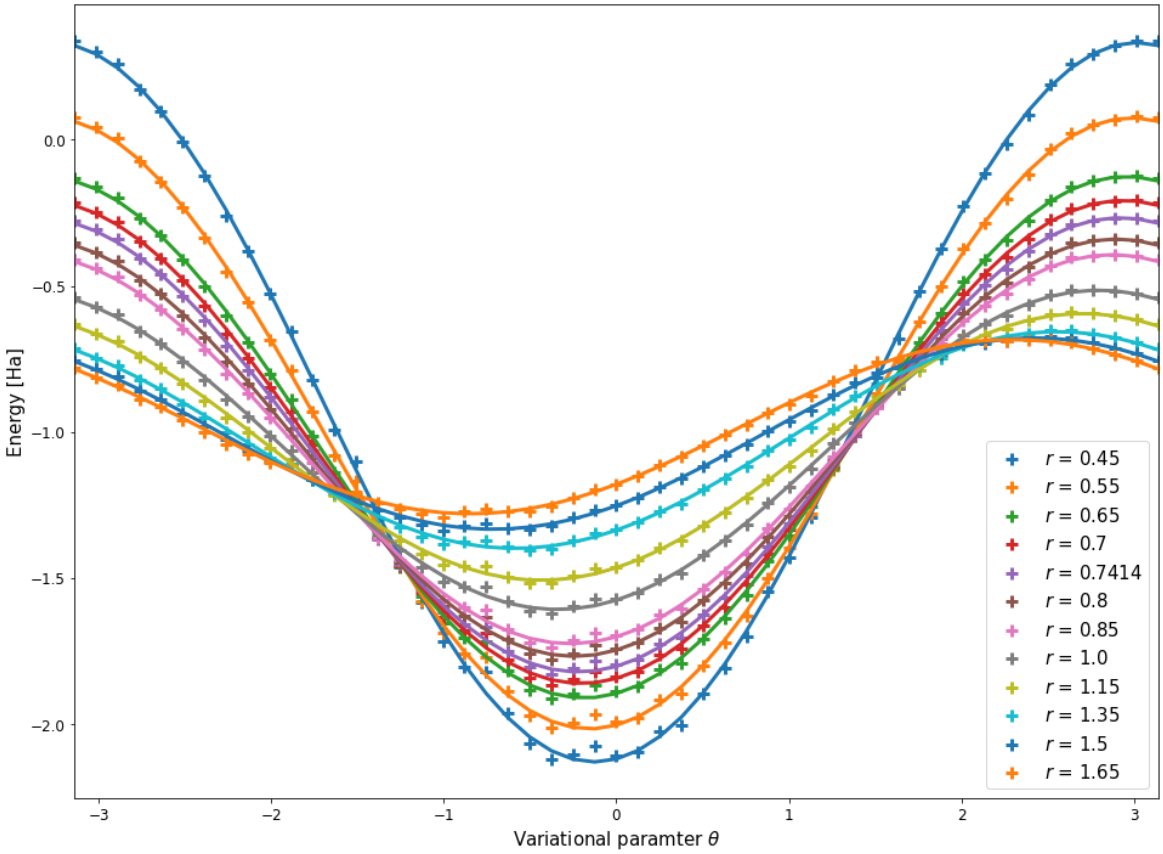


Figure S5: Measurement results for a single sweep of θ between $-\pi$ and π. The same measurement results have been used with different electronic Hamiltonians to generate the energies using QuTiP[11] for the different geometries. A cosine function of the form, $A\cos(\theta - \alpha)$, is used to generate the fits for the data.

H_2 Hamiltonian

Table S2: The electronic Hamiltonian for H_2 expressed as weighted Pauli string operators after parity mapping for various geometry.

r	II	IZ	ZI	ZZ	XX
0.45	-0.908	0.634	-0.634	-0.013	0.167
0.55	-0.981	0.536	-0.536	-0.012	0.171
0.65	-1.028	0.455	-0.455	-0.012	0.176
0.70	-1.044	0.420	-0.420	-0.012	0.179
0.7414	-1.054	0.394	-0.394	-0.011	0.181
0.80	-1.063	0.360	-0.360	-0.011	0.185
0.85	-1.068	0.334	-0.334	-0.010	0.188
1.00	-1.069	0.268	-0.268	-0.009	0.197
1.15	-1.058	0.215	-0.215	-0.007	0.206
1.35	-1.033	0.161	-0.161	-0.005	0.220
1.50	-1.010	0.129	-0.129	-0.004	0.230
1.65	-0.985	0.103	-0.103	-0.003	0.239

Data

The total energy of a system is the sum of its electronic and nuclear repulsion energies at a given geometry for H_2. The optimal parameters are reported in Table S3. The nuclear energies are reported in Table S4, and the electronic energies are reported in Table S5.

Table S3: Optimized parameters, before and after readout mitigation, that minimize the Hamiltonian operator on the Chalmer Särimner device for calculations of H_2.

r	θ (Uncorrected)	θ (Readout mit.)
0.45	-0.1186	-0.1272
0.55	-0.1437	-0.1540
0.65	-0.1737	-0.1861
0.70	-0.1906	-0.2042
0.7414	-0.2056	-0.2202
0.80	-0.2284	-0.2445
0.85	-0.2495	-0.2669
1.00	-0.3220	-0.3438
1.15	-0.4106	-0.4372
1.35	-0.5553	-0.5876
1.50	-0.6802	-0.7153
1.65	-0.8121	-0.8477

Table S4: Nuclear repulsion energies (V_{NN}) of H_2 at various geometries, r. All V_{NN} energies in hartrees.

r	V_{NN}
0.45	1.1759
0.55	0.9621
0.65	0.8141
0.70	0.7560
0.7414	0.7138
0.80	0.6615
0.85	0.6226
1.00	0.5292
1.15	0.4602
1.35	0.3920
1.50	0.3528
1.65	0.3207

II

Table S5: Exact, regular VQE and mitigated electronic energies and errors of H_2 at various geometries, r. All energies are in hartrees.

r [Å]	$E_{exact}(\vec{\theta}_{ref})$	$E_{VQE}(\vec{\theta}_{ref})$	$E_{VQE*}(\vec{\theta}_{ref})$	$E_{exact}(\vec{\theta}_{min}$	$E_{VQE}(\vec{\theta}_{min,VQE})$	$E_{VQE*}(\vec{\theta}_{min,VQE}$	E_{REM}	E_{REM*}	$\Delta E_{error,VQE}$	$\Delta E_{error,VQE*}$	$\Delta E_{error,REM}$	$\Delta E_{error,REM*}$
0.45	-0.9875	-0.8524	-0.9446	-0.9984	-0.8604	-0.9546	-0.9955	-0.9975	0.1380	0.0438	0.0029	0.0010
0.55	-1.0791	-0.9649	-1.0426	-1.0926	-0.9749	-1.0550	-1.0890	-1.0914	0.1178	0.0376	0.0036	0.0012
0.65	-1.1130	-1.0162	-1.0820	-1.1299	-1.0287	-1.0974	-1.1254	-1.1284	0.1013	0.0325	0.0045	0.0015
0.70	-1.1173	-1.0281	-1.0886	-1.1362	-1.0419	-1.1058	-1.1312	-1.1345	0.0943	0.0304	0.0050	0.0017
0.7414	-1.1167	-1.0331	-1.0897	-1.1373	-1.0482	-1.1085	-1.1318	-1.1355	0.0891	0.0288	0.0055	0.0018
0.80	-1.1109	-1.0345	-1.0861	-1.1341	-1.0516	-1.1073	-1.1280	-1.1321	0.0825	0.0268	0.0062	0.0020
0.85	-1.1025	-1.0317	-1.0794	-1.1284	-1.0507	-1.1030	-1.1215	-1.1261	0.0776	0.0253	0.0068	0.0023
1.00	-1.0661	-1.0093	-1.0473	-1.1012	-1.0352	-1.0793	-1.0919	-1.0981	0.0660	0.0219	0.0092	0.0030
1.15	-1.0210	-0.9752	-1.0054	-1.0679	-1.0099	-1.0483	-1.0557	-1.0639	0.0580	0.0196	0.0122	0.0040
1.35	-0.9572	-0.9227	-0.9449	-1.0251	-0.9733	-1.0072	-1.0078	-1.0195	0.0517	0.0179	0.0172	0.0056
1.50	-0.9109	-0.8830	-0.9005	-0.9981	-0.9486	-0.9808	-0.9765	-0.9912	0.0495	0.0173	0.0216	0.0070
1.65	-0.8678	-0.8452	-0.8590	-0.9771	-0.9283	-0.9599	-0.9508	-0.9688	0.0489	0.0172	0.0263	0.0084

Helium hydride – HeH^+

Ansatz, circuit and computation details

Similar to H_2, the HeH^+ wavefunction can be represented with two qubits using parity mapping[7] in a minimal basis. The circuit consists of three parameters – two single excitation parameters, *θ[0]* and *θ[1]*, and a double excitation parameter, *θ[2]*. These parameters are optimized using the VQE algorithm implemented in Qiskit and run on the IBMQ-Quito device with 8192 shots, the maximum number of allowed shots we had access to on IBMQ. The details of the Hamiltonian operator to be minimized are given in Table S6 for different geometries of HeH^+.

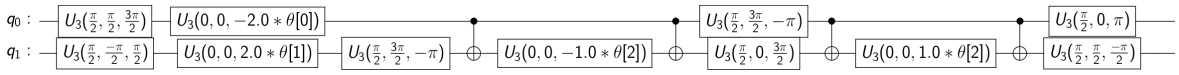


Figure S6: Quantum circuit for HeH^+ implementing the UCCSD operator with respect to the Hartree-Fock reference state. The circuit is compiled to the native gates of IBMQ-Quito device.

HeH^+ Hamiltonian

The details of the Hamiltonian operator to be minimized are given in Table S6 for different geometries of HeH^+.

Table S6: The electronic Hamiltonian expressed as weighted Pauli string operators after parity mapping for various geometry of HeH^+.

r	II	IZ	ZI	ZZ	ZX	XZ	IX	XI	XX
0.65	-3.229	0.635	-0.635	-0.074	-0.094	0.094	0.094	0.094	0.157
0.7899	-3.161	0.560	-0.560	-0.097	-0.106	0.106	0.106	0.106	0.144
0.85	-3.129	0.538	-0.538	-0.108	-0.111	0.111	0.111	0.111	0.137
0.90	-3.101	0.523	-0.523	-0.118	-0.114	0.114	0.114	0.114	0.131
0.95	-3.073	0.512	-0.512	-0.128	-0.117	0.117	0.117	0.117	0.124
1.00	-3.045	0.503	-0.503	-0.139	-0.119	0.119	0.119	0.119	0.117
1.15	-2.962	0.488	-0.488	-0.173	-0.122	0.122	0.122	0.122	0.095
1.35	-2.857	0.488	-0.488	-0.217	-0.115	0.115	0.115	0.115	0.066
1.5	-2.785	0.495	-0.495	-0.247	-0.104	0.104	0.104	0.104	0.047
1.65	-2.721	0.506	-0.506	-0.273	-0.090	0.090	0.090	0.090	0.032

Data

The total energy of a system is the sum of its electronic and nuclear repulsion energies at a given geometry for H_2. The optimal parameters are reported in Table S7. The nuclear energies are reported in Table S8, and the electronic energies are reported in Table S9.

Table S7: The optimized parameter values that minimize the Hamiltonian operator on the IBMQ-Quito device for various geometry of HeH^+. The angles are in radians.

r	$\theta[0]$	$\theta[1]$	$\theta[2]$
0.65	0.011	0.008	-0.061
0.7899	0.014	0.016	-0.067
0.85	0.013	0.010	-0.065
0.90	0.012	0.013	-0.063
0.95	0.017	0.015	-0.065
1.00	0.021	0.021	-0.063
1.15	0.017	0.017	-0.053
1.35	0.012	0.012	-0.036
1.5	0.009	0.003	-0.025
1.65	0.008	0.005	-0.018

Table S8: Nuclear repulsion energies (V_{NN}) of HeH^+ at various geometries, r. All V_{NN} energies in hartrees.

r	V_{NN}
0.65	1.6282
0.7899	1.3399
0.85	1.2451
0.90	1.1759
0.95	1.1141
1.00	1.0584
1.15	0.9203
1.35	0.7840
1.5	0.7056
1.65	0.6414

II

Table S9: The exact, uncorrected and mitigated electronic energies and errors of HeH^+ at various geometries, r. All energies are in hartrees.

r [Å]	$E_{exact}(\vec{\theta}_{ref})$	$E_{VQE}(\vec{\theta}_{ref})$	$E_{VQE*}(\vec{\theta}_{ref})$	$E_{exact}(\vec{\theta}_{min})$	$E_{VQE}(\vec{\theta}_{min,VQE})$	$E_{VQE*}(\vec{\theta}_{min,VQE})$	E_{REM}	E_{REM*}	$\Delta E_{error,VQE}$	$\Delta E_{error,VQE*}$	$\Delta E_{error,REM}$	$\Delta E_{error,REM*}$
0.65	-2.7964	-2.7580	-2.7604	-2.8062	-2.7673	-2.7703	-2.8057	-2.8063	0.0389	0.0359	0.0005	-0.0001
0.7899	-2.8447	-2.8110	-2.8150	-2.8542	-2.8203	-2.8247	-2.8540	-2.8544	0.0338	0.0294	0.0002	-0.0002
0.85	-2.8517	-2.8195	-2.8225	-2.8608	-2.8278	-2.8305	-2.8600	-2.8597	0.0330	0.0302	0.0008	0.0010
0.90	-2.8540	-2.8244	-2.8261	-2.8626	-2.8326	-2.8359	-2.8622	-2.8638	0.0300	0.0267	0.0004	-0.0012
0.95	-2.8542	-2.8253	-2.8267	-2.8622	-2.8324	-2.8353	-2.8614	-2.8629	0.0298	0.0269	0.0008	-0.0007
1.00	-2.8529	-2.8252	-2.8270	-2.8602	-2.8315	-2.8339	-2.8592	-2.8598	0.0287	0.0263	0.0010	0.0004
1.15	-2.8445	-2.8181	-2.8206	-2.8495	-2.8233	-2.8261	-2.8497	-2.8500	0.0262	0.0235	-0.0002	-0.0004
1.35	-2.8314	-2.8076	-2.8093	-2.8339	-2.8095	-2.8120	-2.8333	-2.8341	0.0243	0.0219	0.0005	-0.0003
1.5	-2.8234	-2.8008	-2.8013	-2.8247	-2.8017	-2.8029	-2.8244	-2.8251	0.0230	0.0218	0.0003	-0.0004

Lithium hydride – LiH

Ansatz, circuit and computation details

A LiH wavefunction can be represented with twelve qubits where each qubit corresponds to one molecular spin orbital in a minimal STO-3G basis. Symmetries in parity mapping[7] was used to further reduce the problem size by two qubits to ten qubits. As shown by Kandala *et al*[9], removing the orbitals that do not participate in bonding can bring down the problem size by four qubits. Coupled with the frozen core approximation, the final LiH circuit can be represented by just four qubits in and around equilibrium geometry. A hardware efficient ansatz, inspired by Qiskit's two-local circuit class is used to construct a compact circuit representing LiH. It consists of alternating rotating layers of entanglement layers and is chosen to utilize the connectivity of IBMQ-Quito's connectivity as shown in Figure S7.

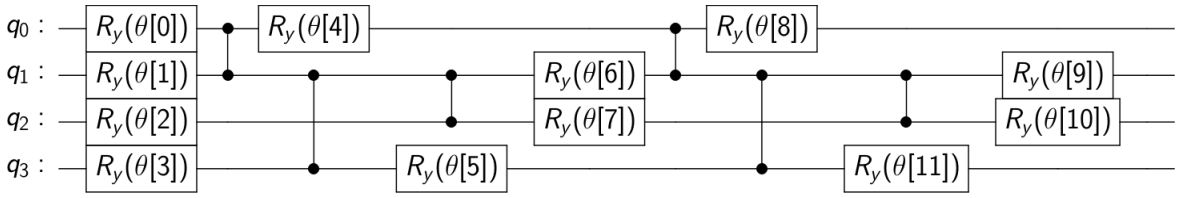


Figure S7: Quantum circuit for LiH utilizing a hardware-efficient ansatz.

LiH Hamiltonian details at 1.5949 Å

Table S10: The electronic Hamiltonian of LiH can be expressed as weighted Pauli string operators with the following coefficients:

IIII	-0.207	ZXIZ	0.012	XZIZ	-0.013	YYXZ	0.008
IIIZ	-0.094	IXZX	-0.003	XIZX	-0.002	XXXZ	-0.008
IIZX	-0.003	ZXZX	-0.003	XZZX	0.002	YYXI	0.008
IIIX	0.003	IXIX	0.003	XIIX	0.002	XXXI	-0.008
IIXX	-0.001	ZXIX	0.003	XZIX	-0.002	ZZZZ	0.084
IIYY	0.001	IXXX	-0.009	XIXX	-0.008	ZZXZ	-0.009
IIZZ	-0.212	ZXXX	-0.009	XZXX	0.008	ZZXI	-0.009
IIXZ	0.019	IXYY	0.009	XIYY	0.008	XIZZ	-0.009
IIXI	0.019	ZXYY	0.009	XZYY	-0.008	XZZZ	0.009
IIZI	0.359	YYIZ	0.032	ZIIZ	0.114	XIXZ	0.007
IZII	0.094	XXIZ	-0.032	ZIZX	-0.011	XZXZ	-0.007
ZXII	0.003	YYZX	-0.009	ZIIX	0.011	XIXI	0.007
IXII	0.003	XXZX	0.009	ZIXX	-0.034	XZXI	-0.007
XXII	-0.001	YYIX	0.009	ZIYY	0.034	ZIZZ	0.060
YYII	0.001	XXIX	-0.009	IZZZ	-0.056	ZIXZ	0.011
ZZII	-0.212	YYXX	-0.031	IZXZ	-0.013	ZIXI	0.011
XZII	-0.019	XXXX	0.031	IZXI	-0.013	IZZI	0.114
XIII	0.019	YYYY	0.031	IXZZ	-0.002	IXZI	-0.011
ZIII	-0.359	XXYY	-0.031	ZXZZ	-0.002	ZXZI	-0.011
IZIZ	-0.122	ZZIZ	0.056	IXXZ	0.002	YYZI	-0.034
IZZX	0.012	ZZZX	0.002	ZXXZ	0.002	XXZI	0.034
IZIX	-0.012	ZZIX	-0.002	IXXI	0.002	ZZZI	-0.060
IZXX	0.032	ZZXX	0.003	ZXXI	0.002	XIZI	-0.011
IZYY	-0.032	ZZYY	-0.003	YYZZ	-0.003	XZZI	0.011
IXIZ	0.012	XIIZ	0.013	XXZZ	0.003	ZIZI	-0.113

Data

Frozen core energy: -7.7983328 hartrees

Nuclear repulsion energy: 0.99538004 hartrees

Table S11: The exact, uncorrected, and mitigated electronic energies and errors of LiH at 1.5949 Å. All energies are in hartrees.

$E_{exact}(\vec{\theta}_{ref})$	$E_{VQE}(\vec{\theta}_{ref})$	$E_{VQE*}(\vec{\theta}_{ref})$	$E_{exact}(\vec{\theta}_{min})$	$E_{VQE}(\vec{\theta}_{min,VQE})$	$E_{VQE*}(\vec{\theta}_{min,VQE})$	E_{REM}	E_{REM*}	$\Delta E_{error,VQE}$	$\Delta E_{error,VQE*}$	$\Delta E_{error,REM}$	$\Delta E_{error,REM}$
-7.8620	-7.6064	-7.6071	-7.8787	-7.6071	-7.6102	-7.8627	-7.8651	0.2717	0.2686	0.0160	0.0136

Optimal Parameters:

Table S12: The optimized parameter values that minimize the Hamiltonian operator on the IBMQ-Quito device for LiH at 1.5949 Å. The angles are in radians.

$\theta[0]$	$\theta[1]$	$\theta[2]$	$\theta[3]$	$\theta[4]$	$\theta[5]$	$\theta[6]$	$\theta[7]$	$\theta[8]$	$\theta[9]$	$\theta[10]$	$\theta[11]$
3.8987	-6.5469	-1.2442	-5.0653	1.5509	2.0379	3.1205	-4.7523	2.3617	6.2591	-5.9394	3.2559

Simulation Details for Lithium Hydride (LiH) and Beryllium Hydride (BeH_2)

As shown in the previous section, LiH wavefunction can be represented with four qubits. Making similar approximations as shown by Kandala *et al*[9], the BeH_2 wavefunction can be represented by six qubits when using the frozen core approximation and orbital reduction. Qiskit's UCCSD module is used to construct the problem ansatz. The circuits are too large to be represented here but the details of the circuit can be found in Table A1. A noise-model from IBMQ-Athens has been added to the simulations to replicate real-world noisy behavior. All simulations have been run using 20,000 shots and repeated 5 times to ensure a high number of samples. The sampling noise is expressed in terms of the standard deviations of our errors. Readout mitigation has been applied for all the simulations.

Table S13: Total ground state energies of molecules(simulated) at experimental equilibrium distances, without and with the application of REM. Bond distances have been obtained from the National Institute of Standards and Technology (NIST). A noise model from IBMQ-Athens has been added to all the simulations. Readout mitigation has been applied for all the VQE calculations. All energies are given in Hartree. The sampling error of the simulations is represented as the standard deviation.

Molecule	$E_{exact}(\vec{\theta}_{min})$	$E_{VQE} * (\vec{\theta}_{min,VQE})$	E_{REM*}	$\Delta E_{error,VQE} *$	$\Delta E_{error,REM*}$
LiH	-7.8811	-7.3599	-7.8705	0.5213 ± 0.003	0.0106 ± 0.002
BeH_2	-15.5895	-13.9873	-15.5632	1.6021 ± 0.005	0.0263 ± 0.007

IBM-Quito Device calibration and connectivity details

Table S14: IBM-Quito's calibration details, as imported from IBM Quantum Services. The device reports a quantum volume of 32.

Qubit	T1 (μs)	T2 (μs)	Frequency (GHz)	Anharmonicity (GHz)	Readout assignment error	Prob meas0 prep1	Prob meas1 prep0	Readout length (ns)	ID error	Single-qubit Pauli-X error	CNOT error	Gate time (ns)
Q0	85.96	95.63	5.300	-0.33148	0.0300	0.0468	0.0132	5351.111	$2.93 \cdot 10^{-4}$	$2.93 \cdot 10^{-4}$	0_1:$6.276 \cdot 10^{-3}$	0_1:234.667
Q1	125.84	128.97	5.081	-0.31925	0.0134	0.0208	0.0060	5351.111	$2.92 \cdot 10^{-4}$	$2.92 \cdot 10^{-4}$	1_3:$1.079 \cdot 10^{-2}$; 1_2:$7.076 \cdot 10^{-3}$; 1_0:$6.276 \cdot 10^{-3}$	1_3:334.222; 1_2:298.667; 1_0:270.222
Q2	81.75	123.32	5.322	-0.33232	0.0237	0.0358	0.0116	5351.111	$2.55 \cdot 10^{-4}$	$2.55 \cdot 10^{-4}$	2_1:$7.076 \cdot 10^{-3}$	2_1:263.111
Q3	96.32	10.29	5.164	-0.33508	0.0281	0.0454	0.0108	5351.111	$3.13 \cdot 10^{-4}$	$3.13 \cdot 10^{-4}$	3_4:$1.557 \cdot 10^{-3}$; 3_1:$1.079 \cdot 10^{-2}$	3_4:277.333; 3_1:369.778
Q4	148.13	162.06	5.052	-0.31926	0.0208	0.0328	0.0088	5351.111	$2.74 \cdot 10^{-4}$	$2.74 \cdot 10^{-4}$	4_3:$1.557 \cdot 10^{-2}$	4_3:312.889

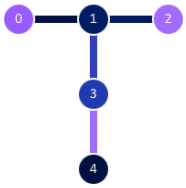


Figure S8: Connectivity representation of the IBMQ-Quito[12] device.

Depolarizing Noise Model

Depolarizing noise channels are a common way to model decoherence and gate errors in quantum devices.[13,14] After observing that our two-qubit (T) gate-error is approximately an order of magnitude larger than our single-qubit (S) gate error, a noise model was constructed that provides S as a linear function of T, $S = 0.1\,T$. This noise model was used to evaluate the effect of REM, applied to a range of depolarizing errors. The depolarizing noise was modelled in Qiskit as a probability p' of applying each of the Pauli gates X, Y, or Z after running a single-qubit gate, and as probability p'' of applying any combination of $P_i \otimes P_j$ where $P_i, P_j \in \{I, X, Y, Z\}$ after running a two-qubit gate.

II

Uncertainty due to Sampling

The uncertainty associated with REM is due to the limited number of samples that can be measured for each point on the noisy energy surface. The resulting spread in the result is characterized by its variance, $\mathrm{Var}(E(\vec{\theta}))$. Assuming that the measurements are normally distributed, the variance of the REM results are inherently larger than the unmitigated ones, as we need to sample two independent distributions, both for the VQE and the reference calculation. Thus, we need to add the variances to get the resulting REM variance

$$\mathrm{Var}\left(E'(\vec{\theta})\right) = \mathrm{Var}\left(E(\vec{\theta})\right) + \mathrm{Var}\left(E(\vec{\theta}_{ref})\right). \tag{S1}$$

This type of trade-off between the accuracy and the variance is common among mitigation strategies, see for example Cai et al.[15]

To understand the variance that appears in the VQE calculation we can view it as a random process $P(\vec{\theta})$ which generates a random variable X_E corresponding to the minimum energy found by the algorithm, based on some initial set of parameters $\vec{\theta}$. The distribution formed by $P(\vec{\theta})$ is in general not known and is hard to approximate with classical means. The most direct way of learning about $P(\vec{\theta})$ is through repeated sampling of several VQE calculations for the same initial set of parameters. For all calculations presented herein the standard deviations presented are calculated from five repeated runs of the VQE algorithm. The exception being the calculation performed on Chalmers' Särimner device.

The Särimner calculation was not performed as a VQE calculation, but it was instead implemented as an exhaustive sweep of the variational parameter. Thus, the errors in the H_2 calculation are estimated as the sum of weighted Bernoulli samples, where each measurement of a Pauli string, P_i, can be seen as sampling a random variable $X_i \in [-1, 1]$, giving either the +1 eigenstate with probability p_i, or the -1 eigenstate with probability $q_i = 1 - p_i$. The variance of the mean of such sampling is $\mathrm{Var}(\bar{X}_i) = \frac{p_i q_i}{n_i}$ where n_i is the number of samples (for all our measurements $n_i = 5000$). We further assume the measurement of each Pauli string to be independent, which allows for the variances of commuting Pauli strings to be summed into an effective variance $\mathrm{Var}(\bar{X}_s) = \sum_r \mathrm{Var}(\bar{X}_r)$, where r is the index over commuting Pauli strings and s is the index over the resulting non-commuting sets of Pauli strings. The resulting non-commuting effective variances are weighted with the square of their respective Hamiltonian coefficients and summed into the total variance, $\mathrm{Var}(\bar{X}_E) = \sum_s h_s^2 \mathrm{Var}(\bar{X}_s)$. Here the square of the coefficient comes from the definition of variance, $\mathrm{Var}(X) = \mathrm{E}[X^2] - \mathrm{E}[X]^2$. Finally, the standard deviation is obtained by calculating $\sqrt{\mathrm{Var}(\bar{X}_E)}$.

References

(1) Koch, J.; Yu, T. M.; Gambetta, J.; Houck, A. A.; Schuster, D. I.; Majer, J.; Blais, A.; Devoret, M. H.; Girvin, S. M.; Schoelkopf, R. J. Charge-Insensitive Qubit Design Derived from the Cooper Pair Box. *Phys Rev A* **2007**, *76* (4), 042319.

(2) McKay, D. C.; Filipp, S.; Mezzacapo, A.; Magesan, E.; Chow, J. M.; Gambetta, J. M. Universal Gate for Fixed-Frequency Qubits via a Tunable Bus. *Phys Rev Appl* **2016**, *6* (6), 064007.

(3) Bengtsson, A.; Vikstål, P.; Warren, C.; Svensson, M.; Gu, X.; Kockum, A. F.; Krantz, P.; KriŽan, C.; Shiri, D.; Svensson, I. M.; Tancredi, G.; Johansson, G.; Delsing, P.; Ferrini, G.; Bylander, J. Improved Success Probability with Greater Circuit Depth for the Quantum Approximate Optimization Algorithm. *Phys Rev Appl* **2020**, *14* (3), 1.

(4) D'Agostini, G. A Multidimensional Unfolding Method Based on Bayes' Theorem. *Nucl Instrum Methods Phys Res A* **1995**, *362* (2–3), 487–498.

(5) Nocedal, J.; Wright, S. J. *Numerical Optimization. Springer. ISBN 978-0-387-30303-1*; Springer US, 2006.

(6) Sun, Q.; Berkelbach, T. C.; Blunt, N. S.; Booth, G. H.; Guo, S.; Li, Z.; Liu, J.; McClain, J. D.; Sayfutyarova, E. R.; Sharma, S.; Wouters, S.; Chan, G. K. L. PySCF: The Python-Based Simulations of Chemistry Framework. *Wiley Interdiscip Rev Comput Mol Sci* **2018**, *8* (1).

(7) McArdle, S.; Endo, S.; Aspuru-Guzik, A.; Benjamin, S. C.; Yuan, X. Quantum Computational Chemistry. *Rev Mod Phys* **2020**, *92* (1), 015003.

(8) ANIS, M. D. S.; Abraham, H.; AduOffei; Agarwal, R.; Agliardi, G.; Aharoni, M.; Akhalwaya, I. Y.; Aleksandrowicz, G.; Alexander, T.; Amy, M.; Anagolum, S.; Arbel, E.; Asfaw, A.; Athalye, A.; Avkhadiev, A.; Azaustre, C.; BHOLE, P.; Banerjee, A.; Banerjee, S.; Bang, W.; Bansal, A.; Barkoutsos, P.; Barnawal, A.; Barron, G.; Barron, G. S.; Bello, L.; Ben-Haim, Y.; Bevenius, D.; Bhatnagar, D.; Bhobe, A.; Bianchini, P.; Bishop, L. S.; Blank, C.; Bolos, S.; Bopardikar, S.; Bosch, S.; Brandhofer, S.; Brandon; Bravyi, S.; Bronn, N.; Bryce-Fuller; Bucher, D.; Burov, A.; Cabrera, F.; Calpin, P.; Capelluto, L.; Carballo, J.; Carrascal, G.; Carriker, A.; Carvalho, I.; Chen, A.; Chen, C.-F.; Chen, E.; Chen, J. (Chris); Chen, R.; Chevallier, F.; Chinda, K.; Cholarajan, R.; Chow, J. M.; Churchill, S.; Claus, C.; Clauss, C.; Clothier, C.; Cocking, R.; Cocuzzo, R.; Connor, J.; Correa, F.; Cross, A. J.; Cross, A. W.; Cross, S.; Cruz-Benito, J.; Culver, C.; Córcoles-Gonzales, A. D.; D, N.; Dague, S.; Dandachi, T. el; Dangwal, A. N.; Daniel, J.; Daniels, M.; Dartiailh, M.; Davila, A. R.; Debouni, F.; Dekusar, A.; Deshmukh, A.; Deshpande, M.; Ding, D.; Doi, J.; Dow, E. M.; Drechsler, E.; Dumitrescu, E.; Dumon, K.; Duran, I.; EL-Safty, K.; Eastman, E.; Eberle, G.; Ebrahimi, A.; Eendebak, P.; Egger, D.; Emilio; Espiricueta, A.; Everitt, M.; Facoetti, D.; Farida; Fernández, P. M.; Ferracin, S.; Ferrari, D.; Ferrera, A. H.; Fouilland, R.; Frisch, A.; Fuhrer, A.; Fuller, B.; GEORGE, M.; Gacon, J.; Gago, B. G.; Gambella, C.; Gambetta, J. M.; Gammanpila, A.; Garcia, L.; Garg, T.; Garion, S.; Gates, T.; Gil, L.; Gilliam, A.; Giridharan, A.; Gomez-Mosquera, J.; Gonzalo; de la Puente González, S.; Gorzinski, J.; Gould, I.; Greenberg, D.; Grinko, D.; Guan, W.; Gunnels, J. A.; Gupta, H.; Gupta, N.; Günther, J. M.; Haglund, M.; Haide, I.; Hamamura, I.; Hamido, O. C.; Harkins, F.; Hasan, A.; Havlicek, V.; Hellmers, J.; Herok, Ł.; Hillmich, S.; Horii, H.; Howington, C.; Hu, S.; Hu, W.; Huang, J.; Huisman, R.; Imai, H.; Imamichi, T.; Ishizaki, K.; Ishwor; Iten, R.; Itoko, T.; Ivrii, A.; Javadi, A.; Javadi-Abhari, A.; Javed, W.; Jianhua, Q.; Jivrajani, M.; Johns, K.; Johnstun, S.; Jonathan-Shoemaker; JosDenmark; JoshDumo; Judge, J.; Kachmann, T.; Kale, A.; Kanazawa, N.; Kane, J.;

Kang-Bae; Kapila, A.; Karazeev, A.; Kassebaum, P.; Kelso, J.; Kelso, S.; Khanderao, V.; King, S.; Kobayashi, Y.; Kovi11Day; Kovyrshin, A.; Krishnakumar, R.; Krishnan, V.; Krsulich, K.; Kumkar, P.; Kus, G.; LaRose, R.; Lacal, E.; Lambert, R.; Landa, H.; Lapeyre, J.; Latone, J.; Lawrence, S.; Lee, C.; Li, G.; Lishman, J.; Liu, D.; Liu, P.; Maeng, Y.; Maheshkar, S.; Majmudar, K.; Malyshev, A.; Mandouh, M. el; Manela, J.; Manjula; Marecek, J.; Marques, M.; Marwaha, K.; Maslov, D.; PawełMaszota; Mathews, D.; Matsuo, A.; Mazhandu, F.; McClure, D.; McElaney, M.; McGarry, C.; McKay, D.; McPherson, D.; Meesala, S.; Meirom, D.; Mendell, C.; Metcalfe, T.; Mevissen, M.; Meyer, A.; Mezzacapo, A.; Midha, R.; Miller, D.; Minev, Z.; Mitchell, A.; Moll, N.; Montanez, A.; Monteiro, G.; Mooring, M. D.; Morales, R.; Moran, N.; Morcuende, D.; Mostafa, S.; Motta, M.; Moyard, R.; Murali, P.; Müggenburg, J.; Nadlinger, D.; Nakanishi, K.; Nannicini, G.; Nation, P.; Navarro, E.; Naveh, Y.; Neagle, S. W.; Neuweiler, P.; Ngoueya, A.; Nicander, J.; Nick-Singstock; Niroula, P.; Norlen, H.; NuoWenLei; O'Riordan, L. J.; Ogunbayo, O.; Ollitrault, P.; Onodera, T.; Otaolea, R.; Oud, S.; Padilha, D.; Paik, H.; Pal, S.; Pang, Y.; Panigrahi, A.; Pascuzzi, V. R.; Perriello, S.; Peterson, E.; Phan, A.; Piro, F.; Pistoia, M.; Piveteau, C.; Plewa, J.; Pocreau, P.; Pozas-Kerstjens, A.; RafałPracht; Prokop, M.; Prutyanov, V.; Puri, S.; Puzzuoli, D.; Pérez, J.; Quant02; Quintiii; R, I.; Rahman, R. I.; Raja, A.; Rajeev, R.; Ramagiri, N.; Rao, A.; Raymond, R.; Reardon-Smith, O.; Redondo, R. M.-C.; Reuter, M.; Rice, J.; Riedemann, M.; Rietesh; Risinger, D.; Rocca, M. la; Rodr\'\iguez, D. M.; RohithKarur; Rosand, B.; Rossmannek, M.; Ryu, M.; SAPV, T.; Sa, N. R. C.; Saha, A.; Saki, A. A.-; Sanand, S.; Sandberg, M.; Sandesara, H.; Sapra, R.; Sargsyan, H.; Sarkar, A.; Sathaye, N.; Schmitt, B.; Schnabel, C.; Schoenfeld, Z.; Scholten, T. L.; Schoute, E.; Schulterbrandt, M.; Schwarm, J.; Seaward, J.; Sergi; Sertage, I. F.; Setia, K.; Shah, F.; Shammah, N.; Sharma, R.; Shi, Y.; Shoemaker, J.; Silva, A.; Simonetto, A.; Singh, D.; Singh, P.; Singkanipa, P.; Siraichi, Y.; Siri; Sistos, J.; Sitdikov, I.; Sivarajah, S.; Sletfjerding, M. B.; Smolin, J. A.; Soeken, M.; Sokolov, I. O.; Sokolov, I.; Soloviev, V. P.; SooluThomas; Starfish; Steenken, D.; Stypulkoski, M.; Suau, A.; Sun, S.; Sung, K. J.; Suwama, M.; Słowik, O.; Takahashi, H.; Takawale, T.; Tavernelli, I.; Taylor, C.; Taylour, P.; Thomas, S.; Tillet, M.; Tod, M.; Tomasik, M.; de la Torre, E.; Toural, J. L. S.; Trabing, K.; Treinish, M.; Trenev, D.; TrishaPe; Truger, F.; Tsilimigkounakis, G.; Tulsi, D.; Turner, W.; Vaknin, Y.; Valcarce, C. R.; Varchon, F.; Vartak, A.; Vazquez, A. C.; Vijaywargiya, P.; Villar, V.; Vishnu, B.; Vogt-Lee, D.; Vuillot, C.; Weaver, J.; Weidenfeller, J.; Wieczorek, R.; Wildstrom, J. A.; Wilson, J.; Winston, E.; WinterSoldier; Woehr, J. J.; Woerner, S.; Woo, R.; Wood, C. J.; Wood, R.; Wood, S.; Wootton, J.; Wright, M.; Xing, L.; Yang, B.; Yeralin, D.; Yonekura, R.; Yonge-Mallo, D.; Young, R.; Yu, J.; Yu, L.; Zachow, C.; Zdanski, L.; Zhang, H.; Zoufal, C.; aeddins-ibm; alexzhang13; b63; bartek-bartlomiej; bcamorrison; brandhsn; catornow; charmerDark; deeplokhande; dekel.meirom; dime10; dlasecki; ehchen; fanizzamarco; fs1132429; gadial; galeinston; georgezhou20; georgios-ts; gruu; hhorii; hykavitha; itoko; jessica-angel7; jliu45; jscott2; klinvill; krutik2966; ma5x; michelle4654; msuwama; ntgiwsvp; ordmoj; sagar pahwa; pritamsinha2304; ryancocuzzo; saswati-qiskit; septembrr; sethmerkel; shaashwat; sternparky; strickroman; tigerjack; tsura-crisaldo; vadebayo49; welien; willhbang; wmurphy-collabstar; yang.luh; Čepulkovskis, M. Qiskit: An Open-Source Framework for Quantum Computing. 2021.

(9) Kandala, A.; Mezzacapo, A.; Temme, K.; Takita, M.; Brink, M.; Chow, J. M.; Gambetta, J. M. Hardware-Efficient Variational Quantum Eigensolver for Small Molecules and Quantum Magnets. *Nature* **2017**, *549* (7671), 242–246.

(10) Newville, M.; Otten, R.; Nelson, A.; Ingargiola, A.; Stensitzki, T.; Allan, D.; Fox, A.; Carter, F.; Michał; Osborn, R.; Pustakhod, D.; lneuhaus; Weigand, S.; Glenn; Deil, C.; Mark; Hansen, A. L. R.; Pasquevich, G.; Foks, L.; Zobrist, N.; Frost, O.; Beelen, A.; Stuermer; azelcer; Hannum, A.; Polloreno, A.; Nielsen, J. H.; Caldwell, S.; Almarza, A.; Persaud, A. Lmfit/Lmfit-Py: 1.0.3. **2021**.

(11) Johansson, J. R.; Nation, P. D.; Nori, F. QuTiP: An Open-Source Python Framework for the Dynamics of Open Quantum Systems. *Comput Phys Commun* **2012**, *183* (8), 1760–1772.

(12) *IBM Quantum*. https://quantum-computing.ibm.com/ (accessed 2021-12-11).

(13) Georgopoulos, K.; Emary, C.; Zuliani, P. Modelling and Simulating the Noisy Behaviour of Near-Term Quantum Computers. **2021**, 1–13.

(14) Vovrosh, J.; Khosla, K. E.; Greenaway, S.; Self, C.; Kim, M.; Knolle, J. Simple Mitigation of Global Depolarizing Errors in Quantum Simulations. **2021**, 1–10.

(15) Cai, Z.; Babbush, R.; Benjamin, S. C.; Endo, S.; Huggins, W. J.; Li, Y.; McClean, J. R.; O'Brien, T. E. Quantum Error Mitigation. *arXiv:2210.00921 [quant-ph]. arXiv.org ePrint archive*. October 3, 2022. http://arxiv.org/abs/2210.00921 (accessed 2023-01-10).

Paper III

The Electron Density: A Fidelity Witness for Quantum Computation

Mårten Skogh[a,b], Phalgun Lolur[a], Werner Dobrautz[a], Christopher Warren[c], Janka Biznárová[c], Amr Osman[c], Giovanna Tancredi[c], Jonas Bylander[c], and Martin Rahm[a,†]

There is currently no combination of quantum hardware and algorithms that can provide an advantage over conventional calculations of molecules or materials. However, if or when such a point is reached, new strategies will be needed to verify predictions made using quantum devices. We propose that the electron density, obtained through experimental or computational means, can serve as a robust benchmark for validating the accuracy of quantum computation of chemistry. An initial exploration into topological features of electron densities, facilitated by quantum computation, is presented here as a proof of concept. Additionally, we examine the effects of constraining and symmetrizing measured one-particle reduced density matrices on noise-driven errors in the electron density distribution. We emphasize the potential benefits and future need for high-quality electron densities derived from diffraction experiments for validating classically intractable quantum computations of materials.

III

Introduction

In this study, we demonstrate the evaluation of electron densities of molecules using quantum computation. We also suggest that the electron density can be a potent validation tool of future quantum calculations, which may prove intractable to solve with conventional quantum chemistry. The study of electron densities is central to several fields of chemistry, physics, and materials science. The Hohenberg-Kohn theorem stipulates that the electron density uniquely defines ground state properties of electronic systems.[1] Through the Hellman-Feynman theorem[2], electron densities provide information on the forces acting within molecules.[3,4] As one of the most information-rich observables in physical sciences,[5–10] the density lays the foundation for density functional theory (DFT), a formalism for predicting properties of many-electron systems.[11] As experiments are the arbiter of truth, the buck often stops with the electron density.

Electron densities can, importantly, be reconstructed from the refinement of X-ray diffraction and scattering data,[9] using, e.g., multipolar models,[5–8,10] X-ray constrained wave functions,[12] or the maximum entropy method.[13] One motivation for our work is that experimentally determined electron densities can prove useful for testing the accuracy of future quantum computations of materials; simulations of which may be unfeasible with conventional computers. Today, it is often preferable, cheaper, and faster to obtain information on electron distribution through conventional quantum mechanical calculations, e.g., by solving the Schrödinger equation at some level of approximation.[14] To obtain highly accurate computational results (energies, densities, or other properties), *ab initio* quantum mechanical methods are the most reliable. Unfortunately, the electronic structure problem scales exponentially with system size.[2]

In 1980 and 1981, Benioff[15], Manin[16] and Feynman[17] pointed out that quantum computers may offer a way forward, enabling larger and more reliable simulations of quantum systems.[18] In recent years, several quantum devices have been developed and applied to compute energies of small molecules (see e.g.,19–21). No such calculation has, however, yet exceeded the accuracy or speed of a conventional quantum chemical calculation. So, just how far away are we from useful quantum computation of chemistry? To quantifiably answer that question, it is necessary to perform benchmarking and validate the results of quantum computation in some manner. Thus far, such efforts have almost exclusively focused on a comparison against total energies obtained using state-of-the-art (and near exact) quantum chemistry.[22–25] However, whereas accurate energies for smaller molecules are available, such comparisons will not be possible if future quantum calculations of more complex systems are made possible. We emphasize at the onset that whether the crossing of such a technological threshold is made in the current noisy intermediate-scale quantum (NISQ) era or if it will require full fault tolerance quantum devices does not affect our main message or conclusions. Here, we suggest that the subtle variability of electron densities, accessible either computationally or experimentally, can act as a potent benchmark[26] for the quantum computation of materials. We draw the colloquial analogy to *fidelity witnesses* in the title,[27] by which we mean experimentally accessible observables whose values (here in terms of topological features) help to quantify the fidelity of a quantum calculation.

In what follows, we demonstrate calculations of electron densities (and their topological features) of molecular hydrogen (H_2) and lithium hydride (LiH) using quantum computers. These molecules are archetypical examples of fundamentally different chemical bonds (covalent and ionic). Simulation of quantum hardware is also employed to derive the electron density of larger molecules, the lithium dimer (Li_2) and hydrogen cyanide

[a] *Department of Chemistry and Chemical Engineering, Chalmers University of Technology, Gothenburg, Sweden*
[b] *Data Science & Modelling, Pharmaceutical Science, R&D, AstraZeneca, Gothenburg, Sweden.*
[c] *Department of Microtechnology and Nanoscience MC2, Chalmers University of Technology, Gothenburg, Sweden.*
† martin.rahm@chalmers.se
* Electronic supplementary information (ESI) available.

(HCN). The quantum volume[28] and noise level of the devices we use are insufficient to demonstrate any advantage over classical implementations (details of the hardware are provided in the Methods section and Supporting Information).[29] However, they suffice for our goal – a first proof-of-principle evaluation of electron densities using quantum devices. By comparing aspects of the electron density topology in these molecules, we showcase a different way to benchmark the quality of quantum hardware calculations of chemistry. To do so, we make use of the quantum theory of atoms in molecules (QTAIM),[30–32] a well-established framework for performing topological analysis of electron densities, which can provide clues into atomic properties,[7] chemical bonding,[8,10] lattice energies,[33] chemical bond strengths, and reactivity.[34–37]

III

Topological Analysis of Electron Densities

Within the Born-Oppenheimer formalism, the average one-electron density for a system of *N* electrons can be expressed as[38]

$$\rho(r) = N \int \Psi_{el}^{*}(r, r_2, \dots, r_N; R)\, \Psi_{el}(r, r_2, \dots, r_N; R) dr_2, \dots, dr_N\,, \quad (1)$$

where r, R, and Ψ_{el} denote the electronic and nuclear coordinates and the electronic wave function, respectively. The electronic density is a real-valued scalar field, lending itself to topological analysis and the extraction of topological features f_χ. One way to analyze such topology is to study *critical points* (CPs), i.e., locations in the density where the gradient vanishes, $\nabla\rho = \mathbf{0}$, or is undefined.[39] The critical points form a concise set of features that yield insight into the molecular structure.

To characterize and distinguish between critical points, we make use of the Hessian and its trace, the Laplacian (or curvature) of the density,

$$H = \begin{bmatrix} \frac{\partial^2\rho}{\partial x^2} & \frac{\partial^2\rho}{\partial xy} & \frac{\partial^2\rho}{\partial xz} \\ \frac{\partial^2\rho}{\partial yx} & \frac{\partial^2\rho}{\partial y^2} & \frac{\partial^2\rho}{\partial yz} \\ \frac{\partial^2\rho}{\partial zx} & \frac{\partial^2\rho}{\partial zy} & \frac{\partial^2\rho}{\partial z^2} \end{bmatrix}, \quad (2)$$

$$\nabla^2\rho = \nabla\cdot\nabla\rho = \frac{\partial^2\rho}{\partial^2 x} + \frac{\partial^2\rho}{\partial^2 y} + \frac{\partial^2\rho}{\partial^2 z}\,. \quad (3)$$

For our analysis and discussion, it suffices to distinguish between critical points using so-called signatures, κ, defined as the sum of the signs of the three eigenvalues of the Hessian.[40] The critical points in molecules we will analyze are of two forms: κ = -3 indicates local maxima, such as positions of nuclei, where all curvatures are negative. A *non-nuclear attractor* (NNA) is a rare example of a CP located at off-nuclei positions. We will discuss one NNA in Li_2. In contrast, κ = -1 corresponds to saddle points in the electron density, i.e., positions where the curvature in one direction is positive. The latter topological feature is commonly referred to as *bond-critical points* (BCPs) because they are often (but not always) found between neighboring atoms that are chemically bound. A BCP is a point of lowest electron density along a path of highest electron density; the lowest point along a ridge connecting maxima. Because the density uniquely defines them, BCPs constitute suitable points of comparison between levels of theory and experiment. BCPs are furthermore useful for characterizing bonds in different ways.[41] For example, the sign of the Laplacian (Eq. (3)) at a BCP is an indication of local depletion (if positive) or concentration (if negative) of the electron density relative to its surroundings. A negative sign of $\nabla^2\rho$ indicates a covalent bond, while a positive sign hints at an ionic (or closed-shell) type of interaction.

Topological analysis of electron densities (viz. QTAIM) also offers a way of *defining atoms* within molecules.[40] Within QTAIM, atoms are identified with *basins* (Figures 1b and 1c), non-overlapping regions of space within which all gradient trajectories of the electron density terminate at the same local maximum (i.e., a $\kappa = -3$ critical point). Note, therefore, that basins need not strictly be centered around nuclei but can be associated with NNAs. Neighboring basins are separated by zero-flux surfaces where $\nabla\rho(\boldsymbol{r})\cdot\boldsymbol{n} = 0$, and where $\boldsymbol{n}$ is the normal to the surface at $\boldsymbol{r}$. We will use such basins to ascribe partial charge to atoms inside molecules, and we suggest such measures, along with ρ and $\nabla^2\rho$ at BCPs, as examples of potent density-based quantum computational fidelity witnesses.

Quantum Computation of Electron Densities

For quantum computation, the electronic structure problem is conveniently expressed within second quantization.[42] Within this formalism, the electron density can be defined as

$$\rho(r) = \sum_{p,q=1}^{n} D_{pq} \sum_{\sigma\in\{\alpha,\beta\}} \phi_{p\sigma}^{*}(r)\phi_{q\sigma}(r)\,, \quad (3)$$

where $\phi_{p\sigma}^{*}(r)$ and $\phi_{q\sigma}(r)$ correspond to the n spin orbitals and p and σ denote spatial and spin indices, respectively. D_{pq} is an entry in the one-particle reduced density matrix (1-RDM).[43] In the following, we use p and q as spatial variables and σ and τ to index spin, in turn, denoted as α or β (i.e., $\sigma, \tau \in \{\alpha, \beta\}$). The 1-RDM can be expressed as,

$$D_{pq} = \langle a_{p\alpha}^{\dagger} a_{q\alpha}\rangle + \left\langle a_{p\beta}^{\dagger} a_{q\beta}\right\rangle, \quad (4)$$

where,

$$\langle a_{p\sigma}^{\dagger} a_{q\tau}\rangle = \langle\Psi| a_{p\sigma}^{\dagger} a_{q\tau}|\Psi\rangle\,. \quad (5)$$

In Eq. (4), $a_{p\sigma}^{\dagger}$ and $a_{q\sigma}$ are the fermionic ladder (creation and annihilation) operators of electrons in spin orbital $\phi_{q\sigma}(r)$ while $|\Psi\rangle$ represents the wave function, defined as a linear combination of Slater determinants, $|\psi_i\rangle$,

$$|\Psi\rangle = \sum_i c_i|\psi_i\rangle\,. \quad (6)$$

In this work, 1-RDMs are constructed following measurements of parametrized quantum circuits (ansätze), representing the ground state of molecules of interest. Measurement of the 1-RDM scales as $O(n^2)$, rendering a density-based fidelity witness approach computationally efficient. We will return to discuss the sensitivity of D_{pq} to noise on quantum devices and how the outcome of calculations can be affected by enforcing physically motivated constraints on the 1-RDM.

Several algorithms[18,44,45] can be utilized to prepare the ground state solution, $|\Psi_{\mathrm{GS}}\rangle$, of a molecule by encoding the fermionic chemistry problem[18,46,47] onto a quantum computer. We rely on the variational quantum eigensolver[48] (VQE) algorithm in this study, as it has been thoroughly used with the current generation of quantum devices. The VQE algorithm, outlined elsewhere,[21,22] leverages both quantum and classical computation to iteratively optimize a parameterized quantum circuit $U(\boldsymbol{\theta})$. It minimizes the expectation value of the Hamiltonian, $\min_{\boldsymbol{\theta}} \langle\Psi_0|U^{\dagger}(\boldsymbol{\theta})\hat{H}U(\boldsymbol{\theta})|\Psi_0\rangle$, where $|\Psi_0\rangle$ is the initial reference state, usually the Hartree-Fock configuration. The ground state problem can be reduced to the electronic Hamiltonian, defined in second quantization as

$$\hat{H} = \sum_{p,q} h_{pq} \sum_{\sigma} a_{p\sigma}^{\dagger} a_{q\sigma} + \frac{1}{2} \sum_{pqrs} V_{pqrs} \sum_{\sigma\tau} a_{p\sigma}^{\dagger} a_{q\tau}^{\dagger} a_{s\tau} a_{r\sigma} \ , \quad (7)$$

where h_{pq} and V_{pqrs} are the one- and two-electron integrals. We stress that our specific choices of algorithms and encoding procedures are not essential for the general case of calculating electron densities with a quantum computer. What is necessary is a) identification of a state of interest (in our case, the ground state) by some quantum algorithm or quantum simulation[49], and b) reconstruction of the 1-RDM through measurements of $\langle a_p^{\dagger} a_q \rangle$, following Eq. (4).

Effects of Noise

Noise is perhaps the single most defining characteristic of current quantum computing and the NISQ era.[29] NISQ algorithms are often hybrid in nature, dividing the computational load between both quantum and conventional hardware. In our case, the orbitals $\phi_{p\sigma}(r)$ and $\phi_{q\tau}(r)$ of Eq. (3) are precisely known functions represented on a conventional computer and assumed errorless. The quantum computer, in turn, stores the orbital occupations and phase. It is in the quantum computer where noise enters as uncertainty and measurement errors. As our interest lies with the 1-RDM, we will focus on the specific effects of noise on the elements D_{pq}.

By defining elementwise errors as $\varepsilon_{pq} = D_{pq} - D_{pq}^{NF}$, where D_{pq}^{NF} is a noise- and error-free reference value, we can divide the effects of noise on the off-diagonal elements into two categories: diagonally symmetric ($\varepsilon_{ij} = \varepsilon_{ji}$) and asymmetric ($\varepsilon_{ij} \neq \varepsilon_{ji}$) errors.

To see how different kinds of noise may affect off-diagonal elements, we first look at a Jordan-Wigner mapping of our fermionic creation and annihilation operators:

$$a_{p\sigma}^{\dagger} = (X_k - iY_k) \otimes Z^{k\rightarrow} \quad (8)$$

$$a_{p\sigma} = (X_k + iY_k) \otimes Z^{k\rightarrow}. \quad (9)$$

We here use k to index our qubits, where each qubit k maps to a unique spin orbital, $\psi_{p\sigma}$. X and Y represent the corresponding Pauli gates, whereas $Z^{k\rightarrow}$ is the application of Pauli Z gates to all qubits $k-1, k-2, \ldots, 1$. Note that Eqs. (8) and (9) are identical except for the sign (phase) of the Y-gate. In practice, this similarity means that the measurements of $\langle a_{p\sigma}^{\dagger} a_{q\tau} \rangle$ and $\langle a_{q\tau}^{\dagger} a_{q\sigma} \rangle$ will both perform the same measurements of Pauli strings $X_k \otimes Z^{k\rightarrow}$ and $Y_k \otimes Z^{k\rightarrow}$, with the phase introduced as a classical coefficient. Any measured diagonal asymmetries in the 1-RDM should, therefore, be solely due to insufficient sampling of the noisy state. Thus, provided independent measurements, and time-independent noise for a given set of circuit parameters, $\boldsymbol{\theta}$, we can expect to sample the same noisy quantum state $\mathrm{P}(\boldsymbol{\theta})$ with every measurement. Here, we use independent measurements to mean single evaluations of the same quantum circuit that are not affected by previous evaluations. We phrase the expectation value in density matrix representation (not to be confused with a *reduced density matrix*), where $\mathrm{P}(\boldsymbol{\theta})$ is the parameterized density matrix. In this representation, the expectation value of an operator $\hat{O}$ is given as $\langle \hat{O} \rangle = \mathrm{Tr}[\mathrm{P}\hat{O}]$. In other words, provided that the above assumptions hold, any pair of real-valued off-diagonal elements must be equal, $\mathrm{Tr}\left[\mathrm{P}(\boldsymbol{\theta})\, a_i^{\dagger} a_j\right] = \mathrm{Tr}\left[\mathrm{P}(\boldsymbol{\theta})\, a_j^{\dagger} a_i\right]$.

While noise is an unavoidable part of contemporary quantum computing, symmetries and properties in the studied system can often be used to gauge and combat errors. As such, it is of interest to study how the 1-RDM and derived properties thereof are affected by noise. And conversely, how enforcing known symmetries and properties of the 1-RDM mitigates the effect of noise. We have opted to study two important properties of the 1-RDM: its Hermiticity, $D_{pq} = D_{qp}^{*}$, and that its trace equals the number of electrons of the studied molecular system[52], $\mathrm{Tr}(\boldsymbol{D}) = n_{el}$.[50]

Because the 1-RDM is Hermitian, we can, assuming real-valued entries, enforce the desired transpose ($D_{pq} = D_{qp}$) symmetry

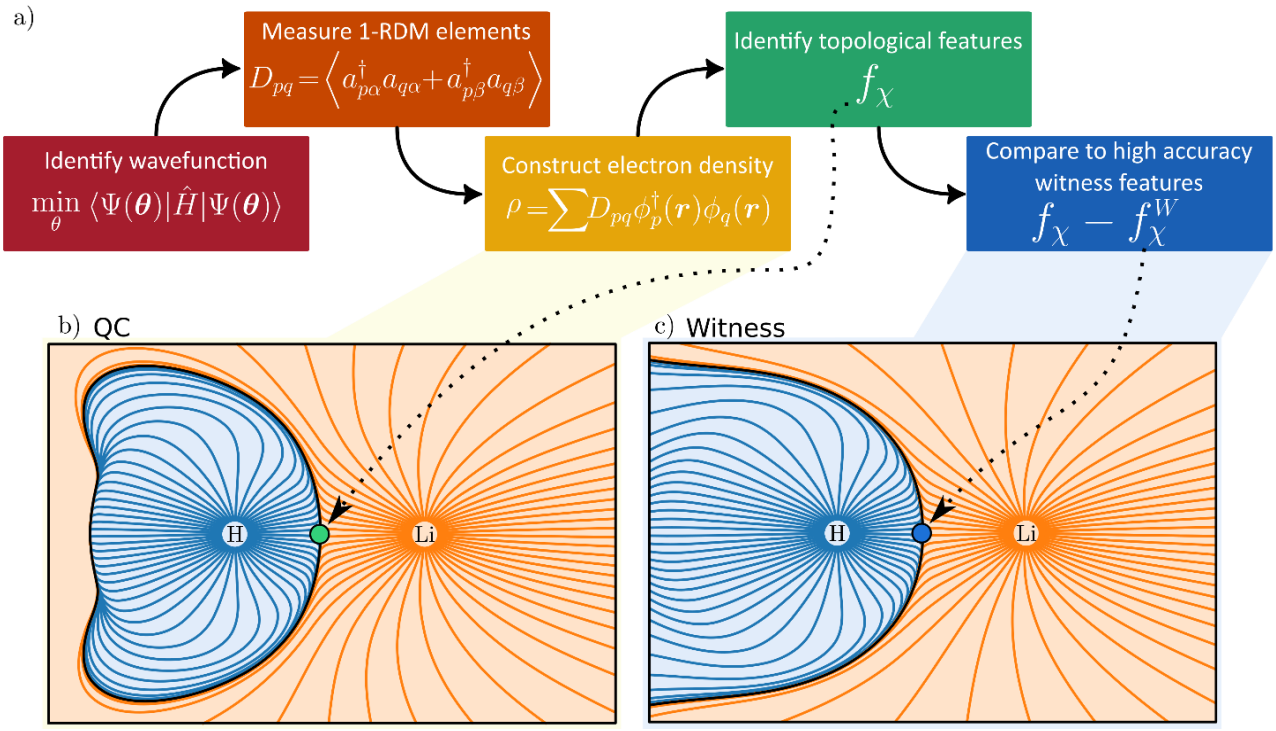


Figure 1 a) Five-step procedure for electron density-based witness estimation. A wavefunction is initially converged using a quantum algorithm. The corresponding one-particle reduced density matrix (1-RDM, $\boldsymbol{D}$) is subsequently measured from the converged wavefunction. The electron density, ρ, is constructed from the measured 1-RDM. Topological analysis of the constructed electron density generates topological features f_χ, such as critical points. Finally, the quality of the quantum calculation can be evaluated by comparing topological features to a known (possibly experimentally determined) witness. b) Gradient field lines of the electron density and atomic basins in LiH derived from quantum computation (QC). c) Gradient field lines and atomic basins for the near-exact witness, derived from conventional quantum chemistry calculations at the CCSD/aug-cc-pVTZ level of theory. Green and blue circles indicate BCPs in data from the quantum calculation and witness, respectively.

on the measured 1-RDMs by averaging the corresponding off-diagonal elements as

$$D'_{pq} = D'_{qp} = \frac{D_{pq} + D_{qp}}{2} . \quad (11)$$

We also ensure particle conservation by normalizing the sum of all diagonal elements, $\mathrm{Tr}(\boldsymbol{D})$, to equal the total number of electrons, n_{el},

$$D^{mit}_{pq} = D'_{pq} \frac{n_{el}}{\mathrm{Tr}(\boldsymbol{D})} . \quad (12)$$

Diagonal and off-diagonal elements are expected to be affected by similar degrees of noise and are subject to the same rescaling. We refer to the combined result of Eqs. (12) and (13) as a noise-mitigated 1-RDM. Our mitigation approach is not unique in relying on 1- or 2-RDMs to reduce errors, and similar techniques have been implemented by others.[51–53] Thus, we emphasize that our aim is not to propose a novel mitigation strategy but rather to study its effect on a noisy 1-RDM and the resulting topology of the electron density. As our focus lies in evaluating topological qualities of the electron densities in the presence of noise and errors, we have opted for a conceptually straightforward mitigation strategy. As will be noted in the results and discussion section, some noise effects on the measured number of electrons can already be avoided by using certain fermion-to-qubit encodings and qubit tapering. In particular, the use parity encoding[46] allows us to preserve the *parity of* the number of particles of each spin species (α and β). For the case of lone α and β electrons, such as the singlet H_2, the parity conservation coincides with particle conservation. Consequently, performing the parity transformation in such cases will effectively protect against any error in the particle or projected spin numbers.

Results and Discussion

To evaluate the viability of the electron density's topological properties as a form of fidelity witness, we perform quantum chemistry calculations on both quantum and conventional computers. The quantum calculations rely on hardware-efficient ansätze[54] and minimal or small basis sets to expand the molecular wavefunction (see the Methods section). We use high-quality electron densities from conventional coupled cluster calculations at the CCSD/aug-cc-pVTZ level of theory as reference data. We note that reference densities can, in principle (up to a feasible limit), be obtained in many ways, including other costly *ab initio* methods, more affordable DFT functionals, diffraction data, or even wavefunction-fitting experiments.[55]

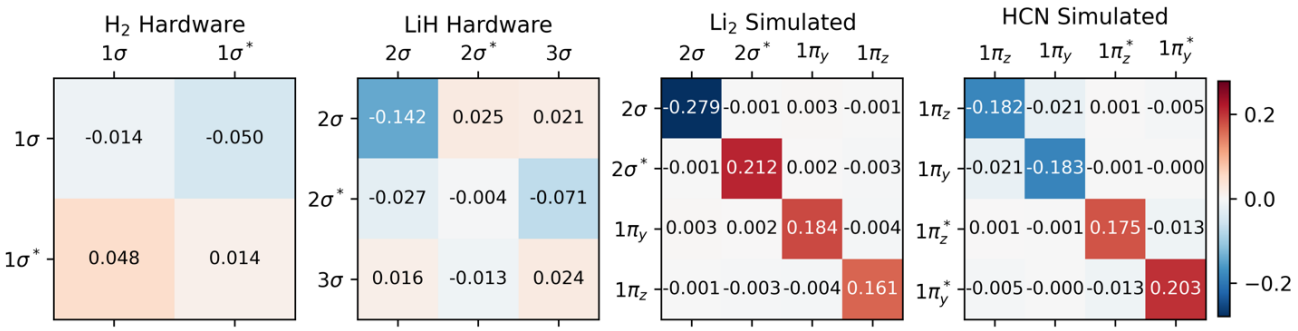


Figure 2 Errors in 1-RDMs, comparing quantum devices (hardware and simulated) to exact (statevector) calculations. Color and intensity indicate the sign and magnitude of the errors, respectively. Hardware data (for H_2 and LiH) demonstrate higher asymmetry than simulations (of Li_2, HCN) with depolarizing noise models. Only 1-RDM elements that correspond to the active space are shown. All bonds are aligned with the x-axis.

Central to our work are analyses of the significant contribution of noise on the quality of electron densities obtained with current quantum calculations. To that end, we not only compare with noiseless quantum simulations and near-exact conventional calculations, but we also implement and evaluate the efficacy of the above-described mitigation strategy for topological properties. To make our test set feasible on available quantum hardware, we rely on small basis sets (STO-3G for H_2, LiH and Li_2, and 6-31G for HCN). A frozen (1s) core approximation is used for atoms heavier than H, together with relatively small active spaces: H_2 (2,2), LiH (2,3), Li_2 (2,4), and HCN (4,4). The canonical molecular orbitals used to define these active spaces are provided in the SI (Figures S5-8). Two kinds of encodings are used to map the fermionic spin orbitals to qubits: parity encoding with two-qubit reduction for H_2 and LiH, and Jordan-Wigner encoding for Li_2 and HCN. While quantum results from H_2 and LiH are obtained from real quantum hardware, results for Li_2 and HCN are from simulations of quantum hardware that include a noise model. Further computational details are available in the Methods section and the SI.

Figure 2 demonstrates that significant errors can be present in both the diagonal and off-diagonal elements of the 1-RDMs. Similar conclusions have been drawn by Arute *et al.*[52] and Smart *et al.*[56] Note that an error along the diagonal is especially detrimental, as this – unless there is fortuitous error cancellation – can result in an incorrect (and unphysical) total number of electrons (Table 1). Even when the number of electrons is conserved by means of the encoding method, as in the case of H_2, errors are still apparent along the diagonal of the 1-RDM. Table 1 summarizes how over- and underestimation of the true number of electrons is a) present for larger molecules and b) can be effectively corrected by rescaling the 1-RDM.
We want to reiterate that errors in the off-diagonal elements come in two flavors: symmetric and asymmetric with respect to the diagonal. Note that the latter type of error only arises in the data from real quantum devices. The found asymmetry agrees well with the reduced number of measurement samples used in these calculations relative to the 10^6 samples applied in our quantum simulations. Physical hardware can also experience noise levels fluctuating over time, breaking our previous assumption of time-invariant noise.

Table 1: Number of electrons obtained from the trace of the 1-RDMs. For H_2, the noisy calculation retains all electrons due to parity encoding and qubit tapering. Our calculations of larger molecules deviate noticeably from the correct number of electrons due to noise, effects that can be corrected by error mitigation. Estimated error bounds are given as a standard deviation based on the number of samples used in the measurements. The electron number for H_2 is not affected by the number of samples.

	Number of electrons		
Molecule	Noise-free	Noisy	Mitigated
H_2	2.00	2.00[(a)]	2.00
LiH	4.00	3.88[(b)] ± 0.011	4.00
Li_2	6.00	6.28[(c)] ± 0.003	6.00
HCN	14.00	14.01[(c)] ± 0.003	14.00

[(a)] Chalmers Särimner device. [(b)] ibmq_quito device. [(c)] Simulation using a depolarizing noise model.

Moving to topological analysis of our electron densities, we look first at atomic partial charges (Table 2). The partial charges attributed to each topological atom are evaluated by partitioning space into atomic QTAIM basins (Figures 1b and 1c). Table 2 shows atomic partial charges derived from such topological atoms and how noise can affect the quantification of this important chemical concept. Also shown in Table 2 are the results following the application of our error mitigation strategy of rescaling and symmetrizing the 1-RDM. Whereas the latter approach generally improves results, its occasional failures can impart valuable lessons on how to improve chemically informed error mitigation.

Table 2: Atomic partial charges derived from a topological QTAIM analysis of the electron density. Rescaling and symmetrization of 1-RDM results generally improve computed partial charges, with the clear exceptions of the NNA of Li_2.

		Partial atomic charge		
Molecule	Atom	Noise-free	Noisy	Mitigated
H_2	H	0.00	0.00[(a)]	0.00
	H	0.00	0.00[(a)]	0.00
LiH	Li	0.84	0.85[(b)]	0.85
	H	-0.84	-0.73[(b)]	-0.85
Li_2	Li	0.32	0.07[(c)]	0.16
	NNA	-0.65	-0.42[(c)]	-0.32
	Li	0.32	0.07[(c)]	0.16
HCN	C	0.89	0.74[(c)]	0.75
	N	-1.07	-0.94[(c)]	-0.93
	H	0.18	0.18[(c)]	0.18

[(a)] Chalmers Särimner device. [(b)] ibmq_quito device. [(c)] Simulation using a depolarizing noise model.

One case of failure of our mitigation strategy lies with the NNA of Li_2, a rare (but well-known) feature in the electron density where a local maximum is present between the two nuclei (Figure 3). One explanation for the reduced accuracy of the NNA's partial charge is a substantial over-occupation due to noise of the valence π orbitals and the $2\sigma^*$ orbital (Figure S4), all of which have but a small overlap with the NNA basin. In contrast, the valence 2σ orbital exhibits a significant reduction in population due to noise and overlaps prominently with the NNA basin. In other words, it is the incorrect relative filling due to noise of the 1π and $2\sigma^*$ orbitals over the 2σ that effectively removes electrons from the inter-nuclear NNA basin in favor of the nuclear basins. Because the rescaling aspect of our error mitigation strategy (viz. Eq. (12)) acts on all orbitals proportionally to their occupation, the 2σ is scaled down to a larger extent, further exaggerating the error in this case.

One example where our approach to error mitigation does little to correct topological features is HCN. The main reason for this is that the used active space is small and only includes the two bonding π and two anti-bonding π^* orbitals. In contrast to our example of Li_2, the π orbitals in HCN are polarized due to the difference in electronegativity of N and C. The π^* orbitals are more pronounced on C and, consequently, have a larger overlap with the identified C basin. In contrast, the π orbitals have a larger N contribution and overlap slightly more with the N basin. Because the net effect of noise is an overpopulation of the antibonding π^* orbitals (Figure S4), the charge distribution is slightly skewed in favor of C (Table 2). A second reason for the negligible effect of our mitigation scheme in the case of HCN is that our calculations are based on sufficiently many, 10^6, simulated measurements, which effectively removes asymmetric errors in the 1-RDM. The main advantage of our error mitigation approach in this example is ensuring the correct number of electrons (Table 1).

Figure 3 illustrates the effect of noise as electron density difference maps in planes across the different molecules, while Table 3 quantifies noise (and mitigation) at selected critical points in the density. Because our calculations rely on a hybrid-quantum-classical algorithm, where orbitals are handled on a classical computer, electron densities are always constrained by the symmetries inherent to those orbitals. Such symmetries can also protect against noise along certain mirror planes. For example, Table 3 shows how the electron density at BCPs reported for HCN are all invariant to noise. This robustness is

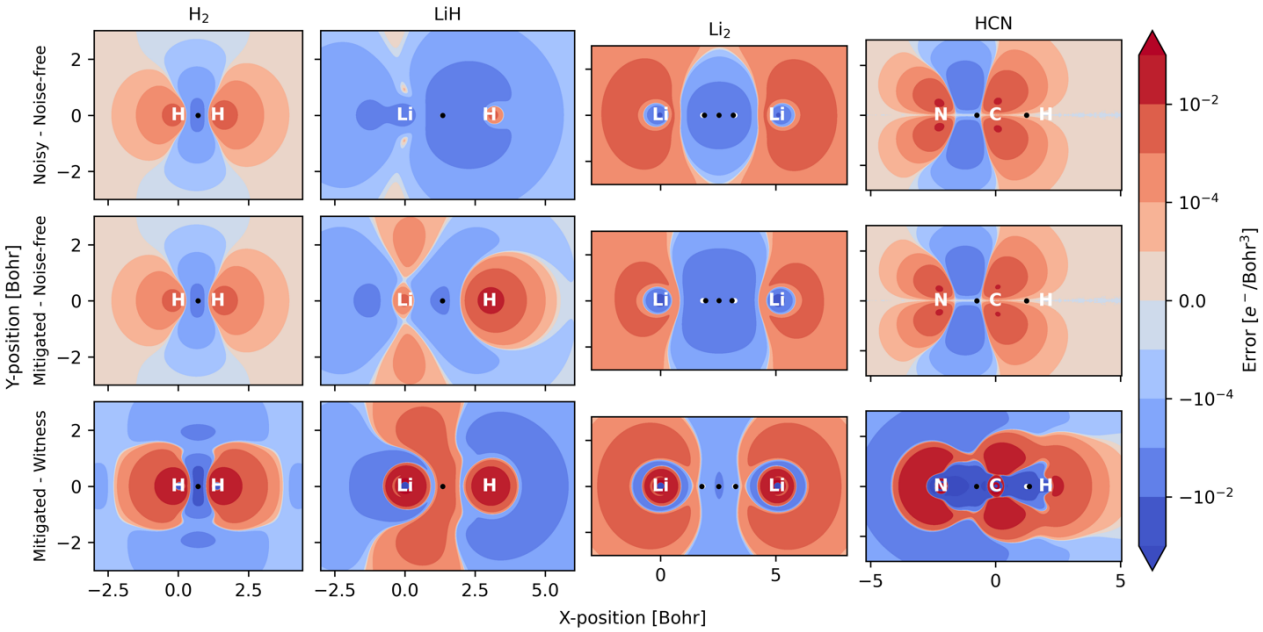


Figure 3 Electron density difference between noisy and noise-free (top), mitigated and noise-free (middle), and between mitigated and witness (bottom) results. The increased occupation of anti-bonding orbitals due to noise can be seen in all studied systems, resulting in a decreased electron density between nuclei.

not general but a consequence of the small (4e, 4o) active space used. Because the π and π^*orbitals have nodal planes along the bond axis, it is only in the Laplacian of $\rho(r)$ that one can distinguish between noisy, mitigated, and noise-free results for these points within HCN. For Li_2, the situation is reversed, and the actual worsening of results following adaptation of our error mitigation is attributed to the rescaling of the 2σ occupation, which substantially affects CPs on the boundary of the NNA basin.

Our examples serve to illustrate the need to account for *noise imbalance* in future NISQ-computation of chemistry that targets high fidelity and accuracy. What we mean by noise imbalance is that some observables are more protected against errors than others. For example, in the context of electron densities, symmetries of orbitals in the chosen active space, the encoding, the ansatz, etc., can all affect relative noise when comparing local properties at different points in space. We note that multiple approaches exist that could potentially mitigate noise imbalance (see, for instance, Refs. 57–60 and references therein).

Methods

Calculations on H_2 were performed with a hardware-efficient circuit on the Särimner device of Chalmers, employing 50,000 samples. Calculations for LiH were performed on ibmq_quito utilizing a hardware-efficient two-local ansatz using 8192 samples.[61] To reduce the effects of noise, readout mitigation was applied to all hardware calculations. Calculations of H_2 and LiH made use of parity encoding to reduce the required number of qubits and circuit depths.

Calculations of Li_2 and HCN were performed on simulated devices using two consecutive layers of the ExcitationPreserving hardware-efficient ansatz available in Qiskit version 0.42.1.[61] Simulations utilized a depolarizing noise model and were performed over 10^6 samples. All calculations were additionally simulated without noise for comparison.

Our near-exact witness data was obtained with conventional quantum chemistry methods, using PySCF[62] at the CCSD/aug-cc-pVTZ level of theory. Topological analyses of electron densities were performed with Critic2.[63] Additional computational and hardware details are provided in the SI.

III

Table 3: Values of the electron density $\rho(\boldsymbol{r})$ and its Laplacian $\nabla^2\rho(\boldsymbol{r})$ at critical points in the density of a selection of molecules, comparing noise-free, noisy, and error mitigated data against a near-exact witness.

		$\rho(\boldsymbol{r})$				$\nabla^2\rho(\boldsymbol{r})$			
Mol	CP	Noise-free	Noisy	Mitigated	Witness[a]	Noise-free	Noisy	Mitigated	Witness[a]
H_2	bond	0.2524	0.2506	0.2506	0.2684	-0.7818	-0.7542	-0.7542	-1.2445
LiH	bond	0.0431	0.0399	0.0419	0.0394	0.1437	0.1177	0.1513	0.1553
Li_2	NNA	0.0158	0.0134	0.0118	0.0129	-0.0139	-0.0102	-0.0085	-0.0143
Li_2	bond	0.0154	0.0132	0.0116	0.0121	0.0017	-0.0025	-0.0005	0.0101
Li_2	bond	0.0154	0.0132	0.0116	0.0121	0.0022	-0.0019	-0.0001	0.0101
H-CN	bond	0.3934	0.3934	0.3934	0.4831	1.3149	1.2480	1.2385	-0.3586
HC-N	bond	0.2635	0.2635	0.2635	0.2990	-0.7468	-0.7346	-0.7321	-1.3241

[a] CCSD/aug-cc-pVTZ level of theory

Conclusions

Whereas tests of quantum computational accuracy are straightforward for small computational problems, for which conventional computation can be referenced, this will not always be so. Due to the rapid advances in quantum computational hardware, we may eventually face situations when it is not easy to validate whether a given problem has been solved to our satisfaction. The motivation behind this work is a proof-of-concept for utilizing electron densities as fidelity witnesses, future-proof benchmarks, for the quality of the quantum computation of materials and molecules. To that end, we have demonstrated the first topological analyses of electron densities inside molecules carried out with the assistance of quantum computers. Our work focuses on measuring one-particle reduced density matrices on real and simulated quantum hardware, how error mitigation can be applied to these entities, and how the resulting electron densities differ from near-exact reference calculations. The molecular systems studied herein are small enough to be treatable by current quantum hardware and small enough for comparisons with high-quality electron densities derived from conventional computations. Our examples are chosen to demonstrate the sometimes-detrimental effects of noise being imbalanced both spatially and with respect to 1-RDM entries. Therefore, we suggest that noise imbalance should be considered when designing or selecting error mitigation techniques, ansätze, encoding, and active spaces. Our work emphasizes the growing potential and need for high-quality (quantum) crystallography experiments,[64,65] i.e., experimental determination of electron densities near the diffraction limit.

Author Contributions

Conceptualization: MR
Investigation (Theoretical): MS, PL, WD, MR
Investigation (Experimental): CW, GT, JBy
Device Fabrication: JBi, AO
Writing: MR, MS, WD, PL

Conflicts of interest

The authors report no conflicts of interest.

Acknowledgements

This research relied on computational resources provided by the Swedish National Infrastructure for Computing (SNIC) at C3SE, NSC and PDC partially funded by the Swedish research council through grant agreement no 2018-05973 and by the National Academic Infrastructure for Supercomputing in Sweden (NAISS) at C3SE and NSC partially funded by the Swedish research council through grant agreement no. 2022-06725. This research has been supported by funding from the Wallenberg Center for Quantum Technology (WACQT), the EU Flagship on Quantum Technology H2020-FETFLAG-2018-03 project 820363 OpenSuperQ, and the HORIZON-CL4-2022-QUANTUM-01-SGA project 101113946 OpenSuperQPlus100. WD acknowledges funding from the European Union's Horizon Europe research and innovation programme under the Marie Skłodowska-Curie grant agreement No. 101062864.

Notes and references

1 P. Hohenberg and W. Kohn, *Physical Review*, 1964, **136**, B864–B871.
2 F. Jensen, *Introduction to computational chemistry*, John Wiley & Sons, Chichester, West Sussex, 3rd edn., 2017.
3 H. Hellmann, *Zeitschrift für Physik*, 1933, **85**, 180–190.
4 R. P. Feynman, *Physical Review*, 1939, **56**, 340–343.
5 N. K. Hansen and P. Coppens, *Acta Crystallographica Section A*, 1978, **34**, 909–921.
6 F. L. Hirshfeld, *Theor Chim Acta*, 1977, **44**, 129–138.
7 A. Fischer, D. Tiana, W. Scherer, K. Batke, G. Eickerling, H. Svendsen, N. Bindzus and B. B. Iversen, *Journal of Physical Chemistry A*, 2011, **115**, 13061–13071.
8 M. Stokkebro Schmøkel, L. Bjerg, J. Overgaard, F. Krebs Larsen, G. K. Hellerup Madsen, K. Sugimoto, M. Takata and B. Brummerstedt Iversen, *Angewandte Chemie International Edition*, 2013, **52**, 1503–1506.
9 C. Gatti and P. Macchi, *Modern Charge-Density Analysis*, Springer Netherlands, Dordrecht, 1st edn., 2012.
10 R. F. Stewart, J. Bentley and B. Goodman, *J Chem Phys*, 1975, **63**, 3786–3793.
11 F. Jensen, *Introduction to computational chemistry*, John wiley & sons, 2017.
12 D. Jayatilaka and D. J. Grimwood, *Acta Crystallogr A*, 2001, **57**, 76–86.
13 D. M. Collins, *Nature*, 1982, **298**, 49–51.
14 P. MacChi, *Crystallogr Rev*, 2013, **19**, 58–101.
15 P. Benioff, *J Stat Phys*, 1980, **22**, 563–591.
16 Y. Manin, *Mathematics as Metaphor: Selected Essays of Yuri I. Manin*, American Mathematical Society, Providence, Rhode Island, 2007.
17 R. P. Feynman, *International Journal of Theoretical Physics*, 1982, **21**, 467–488.
18 S. McArdle, S. Endo, A. Aspuru-Guzik, S. C. Benjamin and X. Yuan, *Rev Mod Phys*, 2020, **92**, 015003.
19 A. Kandala, A. Mezzacapo, K. Temme, M. Takita, M. Brink, J. M. Chow and J. M. Gambetta, *Nature*, 2017, **549**, 242–246.
20 C. Hempel, C. Maier, J. Romero, J. McClean, T. Monz, H. Shen, P. Jurcevic, B. P. Lanyon, P. Love, R. Babbush, A. Aspuru-Guzik, R. Blatt and C. F. Roos, *Phys Rev X*, 2018, **8**, 31022.

21 A. Peruzzo, J. McClean, P. Shadbolt, M.-H. Yung, X.-Q. Zhou, P. J. Love, A. Aspuru-Guzik and J. L. O'Brien, *Nat Commun*, 2014, **5**, 4213.

22 S. McArdle, S. Endo, A. Aspuru-Guzik, S. Benjamin and X. Yuan, *Rev Mod Phys*, 2020, **92**, 15003.

23 Y. Cao, J. Romero, J. P. Olson, M. Degroote, P. D. Johnson, M. Kieferová, I. D. Kivlichan, T. Menke, B. Peropadre, N. P. D. Sawaya, S. Sim, L. Veis and A. Aspuru-Guzik, *Chem Rev*, 2019, **119**, 10856–10915.

24 P. Lolur, M. Skogh, W. Dobrautz, C. Warren, J. Biznárová, A. Osman, G. Tancredi, G. Wendin, J. Bylander and M. Rahm, *J Chem Theory Comput*, 2023, **19**, 783–789.

25 W. Dobrautz, I. O. Sokolov, K. Liao, P. L. Ríos, M. Rahm, A. Alavi and I. Tavernelli, arXiv:2303.02007, 2023

26 J. Eisert, D. Hangleiter, N. Walk, I. Roth, D. Markham, R. Parekh, U. Chabaud and E. Kashefi, *Nature Reviews Physics*, 2020, **2**, 382–390.

27 M. Gluza, M. Kliesch, J. Eisert and L. Aolita, *Phys Rev Lett*, 2018, **120**, 190501.

28 A. W. Cross, L. S. Bishop, S. Sheldon, P. D. Nation and J. M. Gambetta, *Phys Rev A (Coll Park)*, 2019, **100**, 032328, DOI:10.1103/PhysRevA.100.032328.

29 J. Preskill, *Quantum*, 2018, **2**, 79.

30 R. F. W. Bader, *Acc Chem Res*, 1985, **18**, 9–15.

31 R. F. W. Bader, *Chem Rev*, 1991, **91**, 893–928.

32 R. F. W. Bader, T. A. Keith, K. M. Gough and K. E. Laidig, *Mol Phys*, 1992, **75**, 1167–1189.

33 A. v. Shishkina, V. v. Zhurov, A. I. Stash, M. v. Vener, A. A. Pinkerton and V. G. Tsirelson, *Cryst Growth Des*, 2013, **13**, 816–828.

34 Z. J. D. Dunitz, O. J. B. Goodenough, K. P. Hemmerich, E. J.A. Ibers, G. C. K.Jorgensen, M. J. B. Neilands, Berkeley D. Reinen and O. R.J.P. Williams, *Structure and Bonding*, 1979.

35 P. Macchi, *Acta Crystallogr B Struct Sci Cryst Eng Mater*, 2017, **73**, 330–336.

36 T. S. Koritsanszky and P. Coppens, *Chem Rev*, 2001, **101**, 1583–1627.

37 P. L. Ayers, R. J. Boyd, P. Bultinck, M. Caffarel, R. Carbó-Dorca, M. Causá, J. Cioslowski, J. Contreras-Garcia, D. L. Cooper, P. Coppens, C. Gatti, S. Grabowsky, P. Lazzeretti, P. Macchi, Á. Martín Pendás, P. L. A. Popelier, K. Ruedenberg, H. Rzepa, A. Savin, A. Sax, W. H. E. Schwarz, S. Shahbazian, B. Silvi, M. Solà and V. Tsirelson, *Comput Theor Chem*, 2015, **1053**, 2–16.

38 R. McWeeny, *Methods of molecular quantum mechanics*, Academic Press, London; San Diego, 1989.

39 P. L. A. Popelier, in *The Chemical Bond*, Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim, Germany, 2014, pp. 271–308.

40 Richard Bader, *Atoms in Molecules: A Quantum Theory*, Clarendon Press, 1990.

41 W. Nakanishi, S. Hayashi and K. Narahara, *Journal of Physical Chemistry A*, 2008, **112**, 13593–13599.

42 A. Szabo and N. S. Ostlund, *Modern Quantum Chemistry: Introduction to Advanced Electronic Structure Theory*, Dover Publications Inc. Mineola, First., 1996.

43 T. Helgaker, P. Jørgensen and J. Olsen, *Molecular Electronic-Structure Theory*, 2000.

44 M. H. Yung, J. D. Whitfield, S. Boixo, D. G. Tempel and A. Aspuru-Guzik, *Adv Chem Phys*, 2014, **154**, 67–106.

45 J. D. Whitfield, J. Biamonte and A. Aspuru-Guzik, *Mol Phys*, 2011, **109**, 735–750.

46 J. T. Seeley, M. J. Richard and P. J. Love, *Journal of Chemical Physics*, 2012, **137**, 22, 224109, DOI:10.1063/1.4768229.

47 Y. Cao, J. Romero, J. P. Olson, M. Degroote, P. D. Johnson, M. Kieferová, I. D. Kivlichan, T. Menke, B. Peropadre, N. P. D. Sawaya, S. Sim, L. Veis and A. Aspuru-Guzik, *Chem Rev*, 2019, **119**, 10856–10915.

48 A. Peruzzo, J. McClean, P. Shadbolt, M. H. Yung, X. Q. Zhou, P. J. Love, A. Aspuru-Guzik and J. L. O'Brien, *Nat Commun*, 2014, **5**, 4213, DOI:10.1038/ncomms5213.

49 T. H. Johnson, S. R. Clark and D. Jaksch, *EPJ Quantum Technol*, 2014, **1**, 10.

50 T. Helgaker, P. Jorgensen and J. Olsen, *Molecular Electronic-Structure Theory*, 2014.

51 J. Tilly, P. V. Sriluckshmy, A. Patel, E. Fontana, I. Rungger, E. Grant, R. Anderson, J. Tennyson and G. H. Booth, *Phys Rev Res*, 2021, **3**, 033230.

52 Google AI Quantum, Collaborators*†, F. Arute, K. Arya, R. Babbush, D. Bacon, J. C. Bardin, R. Barends, S. Boixo, M. Broughton, B. B. Buckley, D. A. Buell, B. Burkett, N. Bushnell, Y. Chen, Z. Chen, B. Chiaro, R. Collins, W. Courtney, S. Demura, A. Dunsworth, E. Farhi, A. Fowler, B. Foxen, C. Gidney, M. Giustina, R. Graff, S. Habegger, M. P. Harrigan, A. Ho, S. Hong, T. Huang, W. J. Huggins, L. Ioffe, S. V Isakov, E. Jeffrey, Z. Jiang, C. Jones, D. Kafri, K. Kechedzhi, J. Kelly, S. Kim, P. V Klimov, A. Korotkov, F. Kostritsa, D. Landhuis, P. Laptev, M. Lindmark, E. Lucero, O. Martin, J. M. Martinis, J. R. McClean, M. McEwen, A. Megrant, X. Mi, M. Mohseni, W. Mruczkiewicz, J. Mutus, O. Naaman, M. Neeley, C. Neill, H. Neven, M. Y. Niu, T. E. O'Brien, E. Ostby, A. Petukhov, H. Putterman, C. Quintana, P. Roushan, N. C. Rubin, D. Sank, K. J. Satzinger, V. Smelyanskiy, D. Strain, K. J. Sung, M. Szalay, T. Y. Takeshita, A. Vainsencher, T. White, N. Wiebe, Z. J. Yao, P. Yeh and A. Zalcman, *Science (1979)*, 2020, **369**, 1084–1089.

53 S. E. Smart and D. A. Mazziotti, *Phys Rev A (Coll Park)*, 2019, **100**, 022517.

54 A. Kandala, A. Mezzacapo, K. Temme, M. Takita, M. Brink, J. M. Chow and J. M. Gambetta, *Nature*, 2017, **549**, 242–246.

55 M. Woińska, D. Jayatilaka, B. Dittrich, R. Flaig, P. Luger, K. Woźniak, P. M. Dominiak and S. Grabowsky, *ChemPhysChem*, 2017, **18**, 3290–3291.

56 S. E. Smart, J.-N. Boyn and D. A. Mazziotti, *Phys Rev A (Coll Park)*, 2022, **105**, 022405.

57 A. Kandala, K. Temme, A. D. Córcoles, A. Mezzacapo, J. M. Chow and J. M. Gambetta, *Nature*, 2019, **567**, 491–495.

58 K. Temme, S. Bravyi and J. M. Gambetta, *Phys Rev Lett*, 2017, **119**, 180509.

59 D. Bultrini, M. H. Gordon, P. Czarnik, A. Arrasmith, M. Cerezo, P. J. Coles and L. Cincio, *Quantum*, 2023, **7**, 1034.

60 Z. Cai, R. Babbush, S. C. Benjamin, S. Endo, W. J. Huggins, Y. Li, J. R. McClean and T. E. O'Brien, arXiv, 2002, preprint, arXiv:2210.00921v2.

61 Qiskit contributors, Qiskit: An Open-source Framework for Quantum Computing, 2023, DOI: 10.5281/zenodo.2573505.

62 Q. Sun, T. C. Berkelbach, N. S. Blunt, G. H. Booth, S. Guo, Z. Li, J. Liu, J. D. McClain, E. R. Sayfutyarova, S. Sharma, S. Wouters and G. K. Chan, *WIREs Computational Molecular Science*, PySCF, DOI:10.1002/wcms.1340.

63 A. Otero-de-la-Roza, E. R. Johnson and V. Luaña, *Comput Phys Commun*, 2014, **185**, 1007–1018.

64 A. Genoni and P. Macchi, *Crystals (Basel)*, 2020, **10**, 473.

65 S. Grabowsky, A. Genoni and H.-B. Bürgi, *Chem Sci*, 2017, **8**, 4159–4176.

III

Supporting information to:

The Electron Density: A Fidelity Witness for Quantum Computation

1. Mårten Skogh[1,2], Phalgun Lolur[1], Werner Dobrautz[1], Christopher Warren[3], Janka Biznárová[3], Amr Osman[3], Giovanna Tancredi[3], Jonas Bylander[3], and Martin Rahm*[1]
2. Department of Chemistry and Chemical Engineering, Chalmers University of Technology, Gothenburg, Sweden
3. Data Science & Modelling, Pharmaceutical Science, R&D, AstraZeneca, Gothenburg, Sweden
4. Department of Microtechnology and Nanoscience MC2, Chalmers University of Technology, Gothenburg, Sweden

III

Table of Contents

Data Availability

Data and code will be made available through an online repository following (and during) peer-review of the manuscript.

Calculation Details

Hardware Calculations

Qiskit[1] was used throughout this work to generate all problem Hamiltonians, as well as to map the fermionic states onto qubits. For H_2 and LiH, parity encoding was used, which allowed two-qubit tapering to reduce the total number of qubits in the calculations. Using a minimal STO-3G basis, the resulting states comprise 2 and 4 qubits, respectively. The LiH active space was reduced by freezing the Li 1s core orbital and removing the $2p_y$ and $2p_z$ orbitals (the p orbitals perpendicular to the bond axis). This removal of orbitals is motivated by the work of Kandala *et al.*[2] To increase the accuracy of our hardware calculations, we implemented hardware-efficient circuits. Figures S1 and S2 show the full circuits used for H_2 and LiH, respectively. Additionally, the two calculations were implemented on two different quantum computers. The H_2 problem on Chalmers' three-qubit Särimner device, and the LiH problem on IBM's five-qubit ibmq_quito chip.

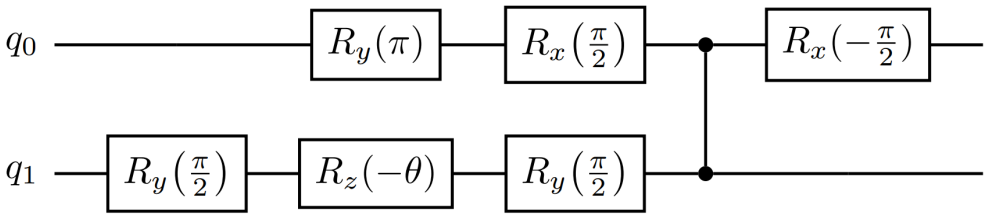


Figure S1: Hardware-efficient circuit used in the H_2 calculation.

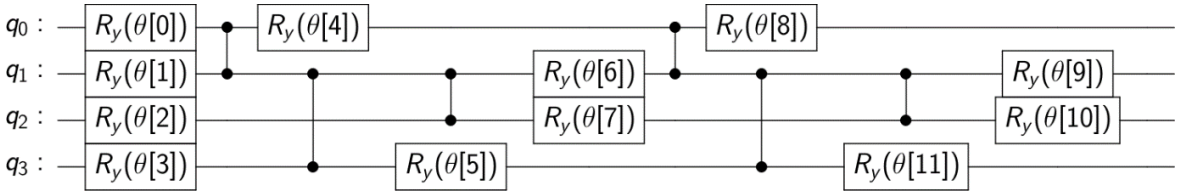


Figure S2: Hardware-efficient circuit used in the LiH calculation.

Simulated Quantum Calculations

Both our noisy simulations, for Li_2 and HCN, were implemented in Qiskit using a double-layer excitation preserving ansatz. A depolarizing noise model, including one- and two-qubit errors, was used to introduce noise effects, details of which are provided below. Due to technical issues with Qiskit, we could not utilize parity encoding for Li_2 and HCN and instead opted for a Jordan-Wigner encoding in these simulations. Noise-free statevector simulations were performed for all systems, including H_2 and LiH, to provide reference values and estimate the effects of noise. Both noise-free and noisy simulations used the COBYLA[3] optimizer during the VQE convergence.

Conventional Quantum Chemistry Calculations

In addition to our quantum computer calculations, we calculate our witness densities using PySCF, a Python package for conventional quantum chemistry[4]. These calculations are performed at the CCSD/aug-cc-pVTZ level of theory.

Topological Analysis

From the 1-RDMs calculated by either quantum computation, simulation, or conventional calculation, a cube file (Gaussian cube file format) was generated using PySCF. The cube files were subsequently used for the topological analysis with the software Critic2.[5]

Hardware Details

The Särimner device used for H_2 calculations is a three-qubit chip. Only two of the three qubits were used in encoding the H_2 state. The qubits are superconducting transmon qubits, and all three qubits are coupled through a single tunable coupler. Measured hardware characteristics are given in Table S1.

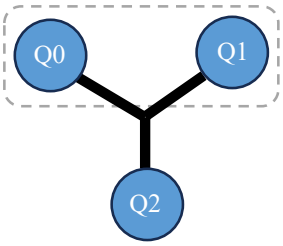


Figure S4: Coupling map for qubits Q0, Q1, and Q2 on Chalmers' Särimner quantum processor. The qubits within the dashed area were used for the calculation of H_2.

The ibmq_quito device is also a superconducting transmon qubit based processor that hosts five qubits, of which four were used in our calculations of LiH (Q0, Q1, Q2, Q3). The chip has a T-shaped coupling map, illustrated in Figure S3.

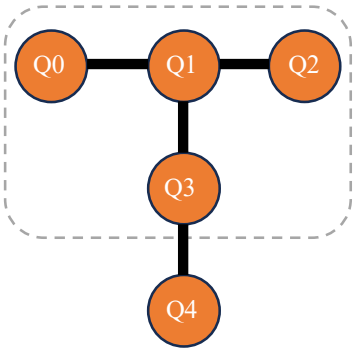


Figure S3: Coupling map for qubits Q0 through Q4 on the ibmq_quito. The qubits within the dashed area were used for the calculation of LiH.

Further details for both quantum processors can be found in the supporting information of Ref.[6]

Simulated Noise Model

To incorporate noise in our quantum hardware simulations, we implement a depolarizing noise model. The model applies random one- and two-qubit gate errors with probabilities 0.001 and 0.002, respectively. One can view the depolarizing channel as random, unintentional applications of additional gates after each intended gate in a circuit. The one-qubit depolarizing channel applies a random single-qubit gate, commonly one of the Pauli gates $\{X, Y, Z\}$, with probability p_1 after each gate in the circuit. Similarly, for two-qubit errors, a two-qubit gate is chosen randomly from a set of Pauli products $\{IX, IY, \ldots, ZY, ZZ\}$ with probability p_2. While the depolarizing model is computationally efficient, it should be noted that it gives a simplified description pf the noise processes in a physical quantum computer. More elaborate models exist, they however require more classical resources to simulate.

Estimating Sampling Uncertainty

For our hardware calculations, we have estimated the sampling uncertainty by providing upper and lower bounds equal to the standard deviation of the measured mean values. For the simulated calculations, this standard deviation is provided by Qiskit as part of the measurement procedure. For our calculations on physical hardware, we have estimated the standard deviation of the mean, s, as that of a Bernoulli distributed variable, $s_i(\bar{X}_i) = \sqrt{\frac{q_i p_i}{n_i}}$. Here, we define p_i as the probability of measuring the +1 eigenvalue and complementary $q_i = 1 - p_i$ as the

probability of measuring the -1 eigenvalue. We also use n_i to denote the number of samples used for measurement. For an operator $\hat{O}$ written as a sum of Pauli operators $\hat{P}_k$, the effective standard deviation can be calculated as $s(\langle\hat{O}\rangle)^2 = \sum_k s(\langle\hat{P}_k\rangle)^2$.

Additional Data

In this section, we present additional data and results from our calculations for the interested reader.

One-Particle Reduced Density Matrices (1-RDM)

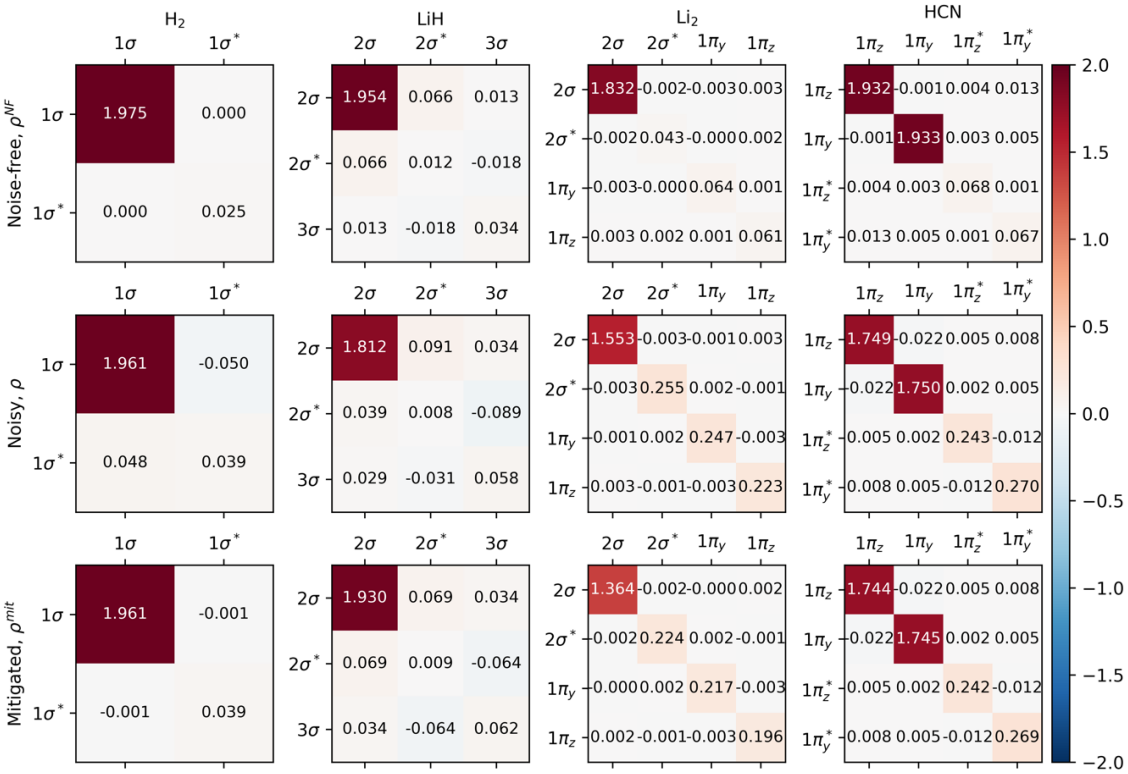


Figure S4: 1-RDMs of all non-witness calculations.

Natural Orbital Occupations

We calculate the natural orbital occupation by diagonalizing the 1-RDMs of the active spaces presented in Figure S4. The orbitals outside these active spaces correspond to Hartree-Fock orbitals, which are either fully occupied (2.0) or empty (0.0). The orbitals are then sorted, and the occupations are listed in Tables S1, S2, S3, and S4 for H_2, LiH, Li_2, and HCN, respectively, and the active space orbitals are indicated in orange.

Table S1: Natural orbital occupation after diagonalizing the H_2 1-RDMs in Figure S4. Active space orbitals are indicated in orange.

	Occupation		
Orbital #	Noise-free	Noisy	Mitigated
1	1.975	1.960	1.961
2	0.025	0.040	0.039

Table S2: Natural orbital occupation after diagonalizing the LiH 1-RDMs in Figure S4. Note the negative occupation of orbital 4 due to noise. Active space orbitals are indicated in orange.

	Occupation		
Orbital #	Noise-free	Noisy	Mitigated
1	2.000	2.000	2.000

2	1.956	1.814	1.933
3	0.044	0.092	0.104
4	0.000	-0.028	-0.038
5	0.000	0.000	0.000
6	0.000	0.000	0.000

Table S3: Natural orbital occupation after diagonalizing the Li_2 1-RDMs in Figure S4. Active space orbitals are indicated in orange.

	Occupation		
Orbital #	Noise-free	Noisy	Mitigated
1	2.000	2.000	2.000
2	2.000	2.000	2.000
3	1.832	1.553	1.364
4	0.064	0.255	0.224
5	0.061	0.247	0.217
6	0.043	0.222	0.195
7	0.000	0.000	0.000
8	0.000	0.000	0.000
9	0.000	0.000	0.000
10	0.000	0.000	0.000

Table S4: Natural orbital occupation after diagonalizing the HCN 1-RDMs in Figure S4. Active space orbitals are indicated in orange.

	Occupation		
Orbital #	Noise-free	Noisy	Mitigated
1	2.000	2.000	2.000
2	2.000	2.000	2.000
3	2.000	2.000	2.000
4	2.000	2.000	2.000
5	2.000	2.000	2.000
6	1.933	1.772	1.766
7	1.931	1.727	1.723
8	0.068	0.275	0.274
9	0.067	0.238	0.237
10	0.000	0.000	0.000
11	0.000	0.000	0.000

Molecular Orbitals

The orbitals used in the active spaces for H_2, LiH, Li_2, and HCN are presented in Figures S5, S6, S7, and S8, respectively. All bonds are aligned along the x-axis.

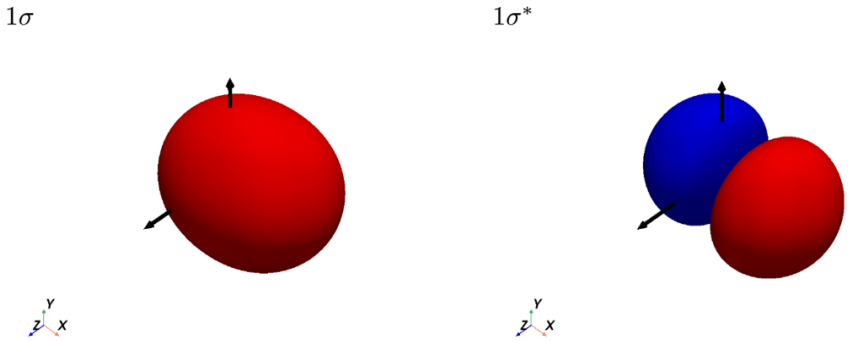


Figure S5: Canonical molecular orbitals in the active space for H2. Iso surfaces are shown for the values ±0.05.

III

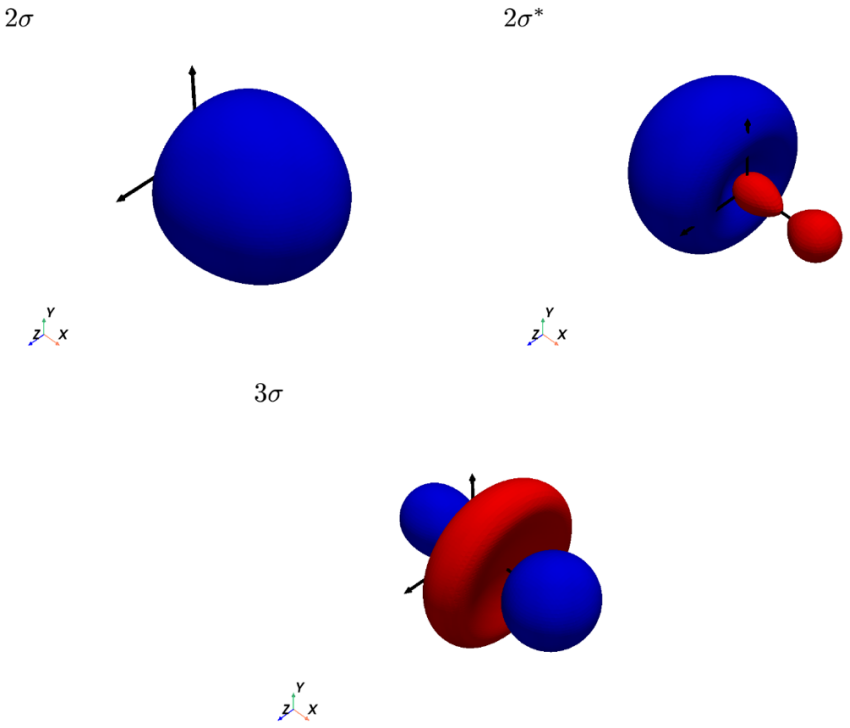


Figure S6: Canonical molecular orbitals in the active space for LiH. Iso surfaces are shown for the values ±0.05.

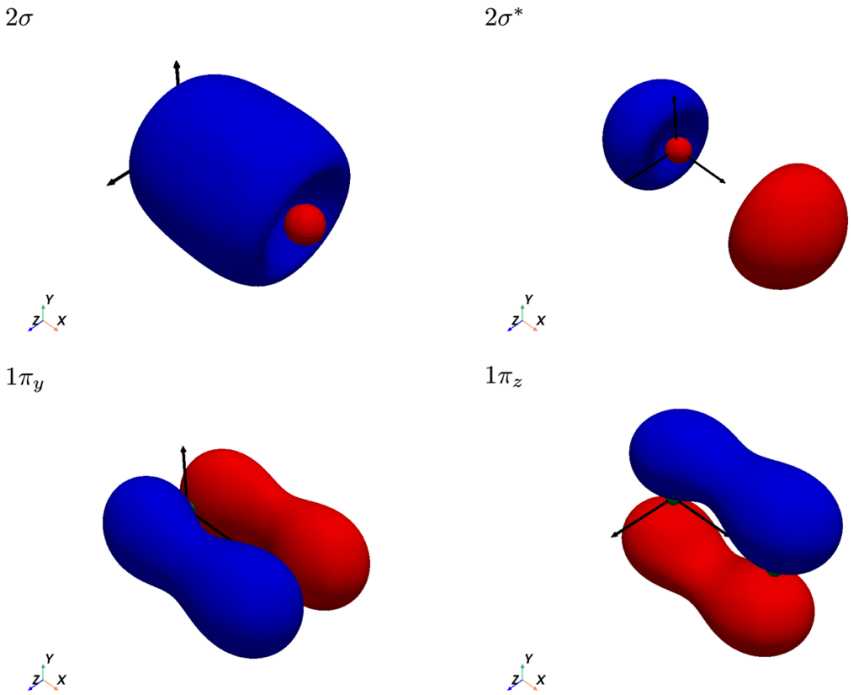


Figure S7: Canonical molecular orbitals in the active space for Li_2. Iso surfaces are shown for the values ±0.05.

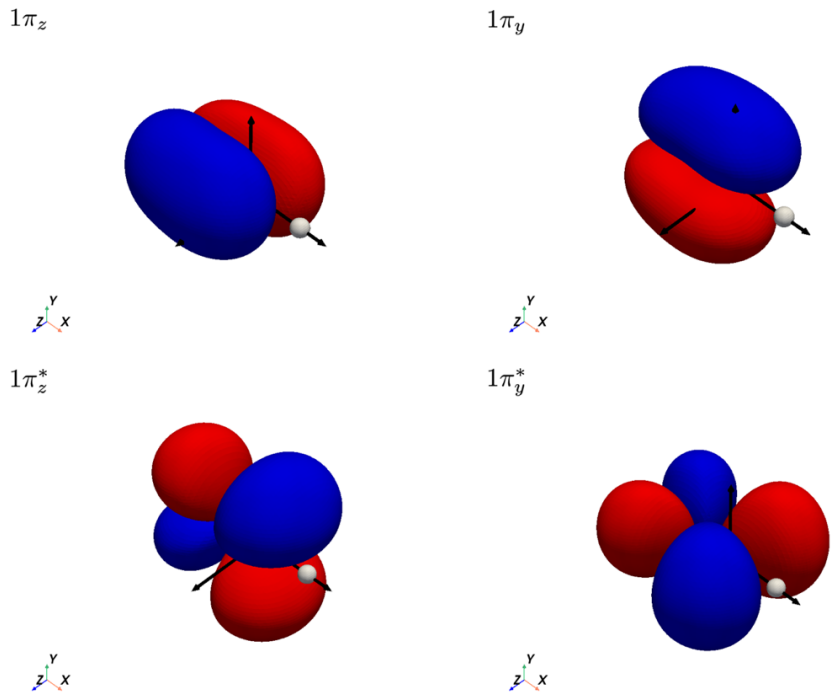


Figure S8: Canonical molecular orbitals in the active space for HCN. Iso surfaces are shown for the values ±0.05.

Additional Li_2 and HCN Density Difference Plots

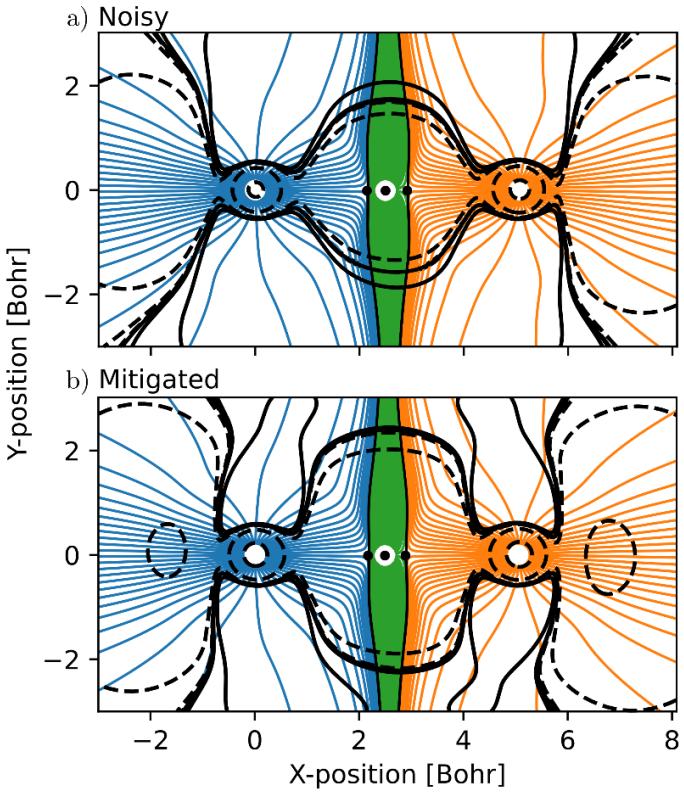


Figure S9: Topological analysis of Li_2. Contour plot compared to noise-free calculations for noisy (a) and mitigated (b) simulations. The contours are overlain streamlines showing the atomic basins for the Li_2 nuclei and NNA. Solid lines indicate a higher electron density when compared to the noise-free results; dashed lines indicate

a lower electron density. The contour lines are separated on a logarithmic scale. Lithium basins are shown in blue and orange, the NNA basin is indicated by green.

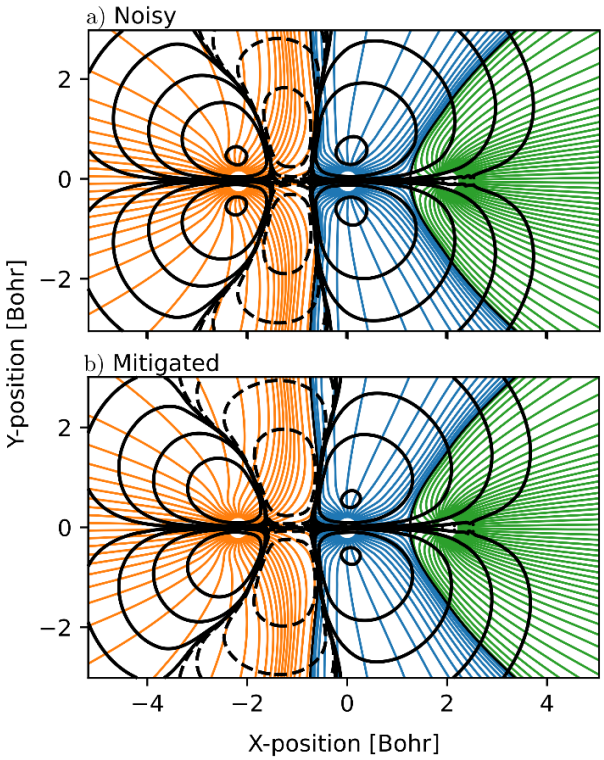


Figure S10: Topological analysis of HCN. Contour plot compared to noise-free calculations for noisy (a) and mitigated (b) simulations. The contours are overlain streamlines showing the atomic basins for the HCN nuclei. Solid lines indicate a higher electron density when compared to the noise-free results; dashed lines indicate a lower electron density. The contour lines are separated on a logarithmic scale. The basins of N, C, and H are indicated by orange, blue, and green, respectively.

Grid-Based Electron Integration

To validate our cube files, the integrated number of electrons was calculated for a large volume. The integrated value is in good agreement with the value derived from the trace of the 1-RDM (Table S1).

Table S5: Total electron numbers for the tested systems calculated using summation density integration and the trace of the 1-RDM.

Molecule	Integrated density[(d)]			Trace of 1-RDM		
H_2[(a)]	2.00	2.00	2.00	2.00	2.00	2.00
LiH[(b)]	4.00	3.88	4.00	4.00	3.88	4.00
Li_2[(c)]	6.00	6.28	6.00	6.00	6.28	6.00
HCN[(c)]	14.00	14.01	14.00	14.00	14.01	14.00

[(a)] Chalmers Särimner device. [(b)] ibmq_quito device. [(c)] Simulation using a depolarizing noise model. [(d)] A margin of 10 a.u. from the nearest nuclei was used when constructing the cube file.

References

1. Qiskit contributors. Qiskit: An Open-source Framework for Quantum Computing. Preprint at https://doi.org/10.5281/zenodo.2573505 (2023).
2. Kandala, A. *et al.* Hardware-efficient variational quantum eigensolver for small molecules and quantum magnets. *Nature* **549**, 242–246 (2017).
3. Powell, M. J. D. A Direct Search Optimization Method That Models the Objective and Constraint Functions by Linear Interpolation. in *Advances in Optimization and Numerical Analysis* 51–67 (Springer Netherlands, 1994). doi:10.1007/978-94-015-8330-5_4.
4. Sun, Q. *et al.* PySCF: the Python-based simulations of chemistry framework. *WIREs Computational Molecular Science* **8**, (2018).
5. Otero-de-la-Roza, A., Johnson, E. R. & Luaña, V. Critic2: A program for real-space analysis of quantum chemical interactions in solids. *Comput Phys Commun* **185**, 1007–1018 (2014).
6. Lolur, P. *et al.* Reference-State Error Mitigation: A Strategy for High Accuracy Quantum Computation of Chemistry. *J Chem Theory Comput* **19**, 783–789 (2023).

Paper IV

The Journal of Chemical Physics ARTICLE scitation.org/journal/jcp

A quantum computing implementation of nuclearelectronic orbital (NEO) theory: Toward an exact pre-Born–Oppenheimer formulation of molecular quantum systems

Cite as: J. Chem. Phys. 158, 214119 (2023); doi: 10.1063/5.0150291
Submitted: 13 March 2023 • Accepted: 16 May 2023 •
Published Online: 5 June 2023

Arseny Kovyrshin,[1,a] Mårten Skogh,[1,2] Anders Broo,[1] Stefano Mensa,[3] Emre Sahin,[3] Jason Crain,[4,5] and Ivano Tavernelli[6,b]

AFFILIATIONS

[1] Data Science and Modelling, Pharmaceutical Sciences, R & D, AstraZeneca Gothenburg, Pepparedsleden 1, Molndal SE-431 83, Sweden
[2] Department of Chemistry and Chemical Engineering, Chalmers University of Technology, Gothenburg, Sweden
[3] The Hartree Centre, STFC, Sci-Tech Daresbury, Warrington WA4 4AD, United Kingdom
[4] IBM Research Europe, Hartree Centre STFC Laboratory, Sci-Tech Daresbury, Warrington WA4 4AD, United Kingdom
[5] Department of Biochemistry, University of Oxford, Oxford OX1 3QU, United Kingdom
[6] IBM Quantum, IBM Research Europe – Zurich, Säumerstrasse 4, 8803 Rüschlikon, Switzerland

[a] **Electronic mail:** arseny.kovyrshin@astrazeneca.com
[b] **Author to whom correspondence should be addressed:** ita@zurich.ibm.com

ABSTRACT

Nuclear quantum phenomena beyond the Born–Oppenheimer approximation are known to play an important role in a growing number of chemical and biological processes. While there exists no unique consensus on a rigorous and efficient implementation of coupled electron–nuclear quantum dynamics, it is recognized that these problems scale exponentially with system size on classical processors and, therefore, may benefit from quantum computing implementations. Here, we introduce a methodology for the efficient quantum treatment of the electron–nuclear problem on near-term quantum computers, based upon the Nuclear–Electronic Orbital (NEO) approach. We generalize the electronic two-qubit tapering scheme to include nuclei by exploiting symmetries inherent in the NEO framework, thereby reducing the Hamiltonian dimension, number of qubits, gates, and measurements needed for calculations. We also develop parameter transfer and initialization techniques, which improve convergence behavior relative to conventional initialization. These techniques are applied to H_2 and malonaldehyde for which results agree with NEO full configuration interaction and NEO complete active space configuration interaction benchmarks for ground state energy to within 10^{-6} hartree and entanglement entropy to within 10^{-4}. These implementations therefore significantly reduce resource requirements for full quantum simulations of molecules on near-term quantum devices while maintaining high accuracy.

I. INTRODUCTION

Chemical simulations at the atomic scale have advanced to the point where they are now capable of describing a wide range of phenomena, including those occurring in complex systems and biological contexts. Much of this progress is due to advances in the power and efficiency of methods that can capture the quantum mechanical nature of electrons. Common to most practical implementations of these molecular dynamics schemes are two main assumptions: (a) that the atomic nuclei behave as classical particles and (b) that the Born–Oppenheimer (BO) approximation holds wherein the electrons are assumed to move sufficiently fast that they respond adiabatically to the nuclear motion.[1] Concerning the first point, several approximated solutions that can capture, at least

IV

partially, nuclear quantum effects have been proposed, including semi-classical approximations,[2] Ehrenfest dynamics,[3,4] correlated-trajectory based solutions,[5] quantum dynamics on pre-computed surfaces [such as multiple spawning[6] and (Multi-Configuration Time-Dependent Hartree) MCTDH[7]], and *a posteriori* quantum corrections for molecular vibrations and zero point energies modeled on the harmonic oscillator.[8]

It is, however, now widely acknowledged that in systems containing light atoms, nuclear quantum effects are not negligible and can make appreciable contributions to processes including proton delocalization and tunneling, occurring, e.g., in biological systems,[9] including enzymatic catalysis,[10] in tautomeric transitions, and in determining the relative stability of crystal polymorphs.[11] The observation of an isotope-dependence in reaction rate[12] is a further particularly striking signature of nuclear quantum effects. Even in a familiar substance, such as liquid and solid water, the presence of light atoms involved in hydrogen bonding[13] suggests that nuclear quantum effects can give important contributions to its static and dynamical properties.[14,15] Finally, phenomena such as charge transfer reactions occurring, for instance, in photo-chemistry (e.g., in the light harvesting complexes) may involve situations where the electronic potential energy surfaces (PESs) become fully degenerate. At these so-called conical intersections, nuclear motion couples to more than a single PES and the adiabatic approximation breaks down, opening new radiationless decay pathways. These examples provide a strong motivation for the extension of the current theoretical frameworks beyond the BO and classical nuclei approximations. It is worth mentioning that the combined electron–nuclear quantum dynamics is not the only approach for the investigation of these processes. Perturbation theory in the small parameter $\kappa = (m/M)$ (where m and M are the masses of the electrons and nuclei, respectively) can be used to construct effective vibrational Hamiltonians that can account for static and dynamical, nonadiabatic effects. An efficient implementation of vibronic states calculations in the quantum computing framework was investigated in Ref. 16.

Currently, several methodologies have been developed that give a full quantum description of chemical systems. They are often referred to as the Nuclear–Electronic Orbital (NEO) approach,[17] pre-Born–Oppenheimer (pBO) quantum theory,[18,19] or Nuclear Orbital plus Molecular Orbital (NO+MO) theory.[20] In these formulations, nuclear tunneling and isotope effects arise naturally. Additionally, phenomena involving the breakdown of the BO approximation are also captured. However, the computational cost of exact nuclear–electronic structure methods implemented on classical computers increases exponentially with the system size. In contrast, quantum processors have the potential to reduce this cost scaling to polynomial-time—underpinning the concept of quantum advantage.[21–23]

Quantum computing is emerging as a new computational paradigm for the efficient solution of quantum mechanical problems, which can have a tremendous impact in different domains, such as many-body physics,[24–34] high energy physics,[35–45] quantum chemistry,[46–57] material design,[58] and biology and medicine.[59–62] The first quantum algorithm for treatment of both nuclei and electrons in a fully quantum simulation setting was proposed by Kassal *et al.*[63] and applied to the calculation of the quantum dynamics for the hydrogen molecule. Later, Veis *et al.*[64] adopted the NO + MO approach for constructing molecular Hamiltonians and proposed a Quantum Phase Estimation (QPE) algorithm for computing the corresponding ground state wave function. Finally, Ollitrault *et al.*[65] presented a quantum algorithm for the simulation of nonadiabatic electron–nuclear dynamics, including excited states.

The NEO approach is particularly versatile and stands as a compromise between the BO approximation and methodologies such as pBO. In the NEO approach, light mass nuclei are treated using orbital techniques in the same way as are the electrons. Furthermore, the BO separation between electronic and nuclear wave functions is—in general—not necessary in this case. On the contrary, the heavy nuclei (usually everything heavier than hydrogen) are described by classical point charges, determining the geometry of the molecular scaffold. Pavošević and Hammes-Schiffer[66] recently proposed a quantum computing implementation of the NEO approach, employing the Variational Quantum Eigensolver (VQE) algorithm for the optimization of the hydrogen molecule and positronium hydride ground state wave functions, where only one proton was treated at the quantum level.

In this study, we investigate the potential of the NEO approach for chemical systems and focus on the construction of the corresponding qubit Hamiltonian and wave function *Ansätze* with the aim of accessing larger system sizes than have been previously investigated. Specifically, we explore and implement improved wave function parameterizations, operational mappings of second quantized molecular Hamiltonians to the qubit space, and efficient parameter initialization schemes for several electron–nuclear wave function *Ansätze*. It is particularly important to devise qubit tapering schemes, such as two-qubit reduction, the exploitation of molecular point group (PG) symmetries, and the application of projectors to efficiently sample the relevant sectors of the molecular Fock space.

After a detailed description of the NEO Hamiltonian and of the corresponding wave function in the quantum computing setting, we tested our algorithm on the evaluation of the ground state energies of the hydrogen and malonaldehyde molecules using exact—noise-free—state vector VQE simulators. In the case of malonaldehyde, our approach was able to predict—despite the quite small size of the nuclear basis set and the imposed rigidity of the molecular scaffold—proton transfer barriers in reasonable good agreement with the available references. Exploiting the full quantum nature of the molecular wave function, we also proposed methods for the evaluation of additional interesting quantum features, such as the entanglement entropy of the subsystems (electronic and nuclear) from the full NEO wave function.

II. THEORY

A. Nuclear–electronic orbital approach

Relative to the standard BO approximation, with point charges representing nuclei, the Hamiltonian in the NEO approach includes additional kinetic and potential energy operators for selected light nuclei. This approach incorporates the quantum effects of light nuclei and additionally introduces a separation between heavy and light nuclear motion. We denote the collective spatial coordinates of all quantum objects as $\boldsymbol{r}$ and additionally introduce superscripts where required to distinguish between nuclei, $\boldsymbol{r}^n$, and electrons, $\boldsymbol{r}^e$. We refer to both quantum nuclear and electronic coordinates combined as $\boldsymbol{r} = \{\boldsymbol{r}^e, \boldsymbol{r}^n\}$. The coordinates for classical nuclei will be

denoted as $\boldsymbol{R}$. Quantities associated with electrons will carry lower case indices i,j, while upper case indices, I,J, are used for nuclear quantities, regardless of whether they are classical or not. The total wave function, $\Psi(\boldsymbol{r};\ \boldsymbol{R})$, will still retain a parametric dependence on classical nuclear coordinates, $\boldsymbol{R}$. Thus, the Schrödinger equation reads as follows (in atomic units):

$$\left[-\sum_{i}^{N^e}\frac{1}{2}\nabla_i^2-\sum_{I}^{N^n}\frac{1}{2M_I}\nabla_I^2+\sum_{i<j}^{N^e}\frac{1}{|\boldsymbol{r}_i-\boldsymbol{r}_j|}+\sum_{I<J}^{N^n}\frac{1}{|\boldsymbol{r}_I-\boldsymbol{r}_J|}-\sum_{i,I}^{N^e,N^n}\frac{Z_I}{|\boldsymbol{r}_i-\boldsymbol{r}_I|}+\sum_{I,A}^{N^n,N^c}\frac{Z_IZ_A}{|\boldsymbol{r}_I-\boldsymbol{R}_A|}-\sum_{i,A}^{N^e,N^c}\frac{Z_A}{|\boldsymbol{r}_i-\boldsymbol{R}_A|}+\sum_{A<B}^{N^c}\frac{Z_AZ_B}{|\boldsymbol{R}_A-\boldsymbol{R}_B|}\right]\Psi(\boldsymbol{r};\boldsymbol{R})=E\Psi(\boldsymbol{r};\boldsymbol{R}),\tag{1}$$

where Z_I are nuclear charges, M_I are nuclear masses, and N^e, N^n, and N^c are the total number of electrons, quantum nuclei, and classical nuclei, respectively.

The NEO approach solves Eq. (1) using both electronic,

$$\phi_i(\boldsymbol{x}_i)\equiv\phi_i(\boldsymbol{r}_i,s_i),\tag{2}$$

and nuclear,

$$\phi_I(\boldsymbol{x}_I)\equiv\phi_I(\boldsymbol{r}_I,s_I),\tag{3}$$

spin orbitals constructed separately from nuclear and electronic basis sets, originally in Gaussian basis sets.[17] Note that the spin coordinates s_i in Eqs. (2) and (3) are combined together with $\boldsymbol{r}_i$ into a new variable $\boldsymbol{x}_i$ for brevity. Based on these spin orbitals, one can form antisymmetric products, Slater determinants, for electrons,

$$\Phi_\mu(\boldsymbol{x}^e)=\hat{S}^-\prod_i\phi_i(\boldsymbol{x}_i^e),\tag{4}$$

and symmetric or antisymmetric spin orbital products for bosonic or fermionic nuclei, correspondingly,

$$\Phi_\nu(\boldsymbol{x}^n)=\hat{S}^\pm\prod_I\phi_I(\boldsymbol{x}_I^n),\tag{5}$$

where $\hat{S}^{+/-}$ are symmetrizer/antisymmetrizer operators. These constitute the multi-particle basis set for the Nuclear–Electronic Orbital Full Configuration Interaction (NEOFCI) approach,[17] which expresses the nuclear–electronic wave function as follows:

$$\Psi(\boldsymbol{x};\boldsymbol{R})=\sum_{\mu\nu}^{C^e,C^n}C_{\mu\nu}\Phi_\mu(\boldsymbol{x}^e)\Phi_\nu(\boldsymbol{x}^n),\tag{6}$$

where C^e and C^n are the total number of electronic and nuclear states [Eqs. (4) and (5)]. We emphasize that in Eq. (6), only one type of nuclei is considered. In the most general case, each type of nuclei will have its own set of spin orbitals and the corresponding symmetric or antisymmetric product, which, in turn, enters the sum in Eq. (6). As in this work, we will only consider protons and electrons as quantum particles, only masses $M_I=M=1874m_e$ and charges $Z_I=1e$ will be considered in the quantum part of the Hamiltonian. Within such a framework, one gets the following second quantization representation for the Hamiltonian:[17]

$$\begin{aligned}\hat{H}=&-\sum_{ij}\left[\int\phi_i^*(\boldsymbol{x})\frac{1}{2}\nabla^2\phi_j(\boldsymbol{x})d\boldsymbol{x}\right]\hat{a}_i^\dagger\hat{a}_j-\sum_{IJ}\left[\int\phi_I^*(\boldsymbol{x})\frac{1}{2M}\nabla^2\phi_J(\boldsymbol{x})d\boldsymbol{x}\right]\hat{a}_I^\dagger\hat{a}_J\\&+\frac{1}{2}\sum_{ijkl}\left[\int\phi_i^*(\boldsymbol{x}_1)\phi_k^*(\boldsymbol{x}_2)\frac{1}{|\boldsymbol{r}_1-\boldsymbol{r}_2|}\phi_l(\boldsymbol{x}_2)\phi_j(\boldsymbol{x}_1)d\boldsymbol{x}_1d\boldsymbol{x}_2\right]\hat{a}_i^\dagger\hat{a}_k^\dagger\hat{a}_l\hat{a}_j\\&+\frac{1}{2}\sum_{IJKL}\left[\int\psi_I^*(\boldsymbol{x}_1)\psi_K^*(\boldsymbol{x}_2)\frac{1}{|\boldsymbol{r}_1-\boldsymbol{r}_2|}\psi_L(\boldsymbol{x}_2)\psi_J(\boldsymbol{x}_1)d\boldsymbol{x}_1d\boldsymbol{x}_2\right]\hat{a}_I^\dagger\hat{a}_K^\dagger\hat{a}_L\hat{a}_J\\&-\sum_{ijKL}\left[\int\phi_i^*(\boldsymbol{x}_1)\psi_K^*(\boldsymbol{x}_2)\frac{1}{|\boldsymbol{r}_1-\boldsymbol{r}_2|}\psi_L(\boldsymbol{x}_2)\phi_j(\boldsymbol{x}_1)d\boldsymbol{x}_1d\boldsymbol{x}_2\right]\hat{a}_i^\dagger\hat{a}_K^\dagger\hat{a}_L\hat{a}_j\\&+\sum_{IJ,A}\left[\int\phi_I^*(\boldsymbol{x})\frac{Z_A}{|\boldsymbol{r}-\boldsymbol{R}_A|}\phi_J(\boldsymbol{x})d\boldsymbol{x}\right]\hat{a}_I^\dagger\hat{a}_J-\sum_{ij,A}\left[\int\phi_i^*(\boldsymbol{x})\frac{Z_A}{|\boldsymbol{r}-\boldsymbol{R}_A|}\phi_j(\boldsymbol{x})d\boldsymbol{x}\right]\hat{a}_i^\dagger\hat{a}_j+\frac{1}{2}\sum_{AB}\frac{Z_AZ_B}{|\boldsymbol{R}_A-\boldsymbol{R}_B|}.\end{aligned}\tag{7}$$

Note that in addition to anti-commutation relations between indistinguishable fermions,

$$\left[\hat{a}_i^\dagger,\hat{a}_j\right]^+=\delta_{ij},\tag{8}$$

$$\left[\hat{a}_i^\dagger,\hat{a}_j^\dagger\right]^+=\left[\hat{a}_i,\hat{a}_j\right]^+=0,\tag{9}$$

$$\left[\hat{a}_I^\dagger,\hat{a}_J\right]^+=\delta_{IJ},\tag{10}$$

$$\left[\hat{a}_I^\dagger,\hat{a}_J^\dagger\right]^+=\left[\hat{a}_I,\hat{a}_J\right]^+=0,\tag{11}$$

distinguishable fermions do commute (between protons and electrons), implying

$$\left[\hat{a}_I^\dagger,\hat{a}_j\right]=0,\tag{12}$$

$$\left[\hat{a}_I^\dagger,\hat{a}_j^\dagger\right]=\left[\hat{a}_I,\hat{a}_j\right]=0.\tag{13}$$

IV

B. Quantum computing *Ansatz*

The VQE is a quantum variational algorithm that uses a classical optimization routine to find the parameters of a quantum state that minimizes the expectation value of a given molecular Hamiltonian.[67] Some of the key advantages of the VQE algorithm include its ability to be implemented on near-term quantum computers, which have a limited number of qubits and are prone to errors, and its potential to scale to larger systems, hence becoming a tool for practical applications in the near future.[68–71]

Mathematically, the VQE optimization algorithm can be expressed as follows: Let $\hat{H}$ be the Hamiltonian of the system and let $|\Psi(\boldsymbol{\theta})\rangle$ be the trial state, which is a function of parameters $\boldsymbol{\theta}$. The goal is to find the values of $\boldsymbol{\theta}$ that minimize the expectation value of the Hamiltonian, given by

$$E(\boldsymbol{\theta}) = \langle \Psi(\boldsymbol{\theta})|\hat{H}|\Psi(\boldsymbol{\theta})\rangle. \tag{14}$$

To find the optimal values of $\boldsymbol{\theta}$, we can use a classical optimization algorithm to iteratively update the values of $\boldsymbol{\theta}$. Then, the quantum part of the algorithm is used to apply the quantum circuit to the quantum state of the system and measure the energy of the resulting state until the minimum value of $E(\boldsymbol{\theta})$ is found.

For the VQE optimization of the system wave function within a quantum computing setting, it is desirable to design a flexible, expressive, and well-behaved *Ansatz*, which can deliver good accuracy with the fewest number of parameters. While there are a vast number of different *Ansätze* available, in this work, we focus on the Unitary Coupled Cluster (UCC) and hardware-efficient parameterizations. The main advantage of the latter arises from the shallow circuit depth required for its implementation—an important prerequisite for near-term (non-fault tolerant) hardware.[72] Unfortunately, for chemical systems, there is no systematic, practical approach for the initialization and optimization of the variational parameters when using hardware-efficient *Ansätze*. Moreover, these often show limited expressivity, while the optimization process suffers from the presence of barren plateaus.[73] Here, we will explore several variants of the TwoLocal hardware-efficient *Ansatz*[74] for the NEO wave function, details of which will be discussed in Sec. IV B.

In contrast to hardware-efficient *Ansätze*, the Unitary Coupled Cluster (UCC) *Ansatz* has a very well-tailored parameter space and naturally approximates wave functions for chemical systems. This well-established quantum chemistry *Ansatz* was first introduced for VQE by Peruzzo *et al.*[67] Although the resulting quantum circuits are, in general, deeper than the hardware-efficient ones, the optimization of the UCC circuit is less prone to barren plateaus and lack of particle, total spin, and spin projection conservation. This becomes especially important when one uses multiple centers for protons and electrons, leading hardware-efficient *Ansatz* optimizations to the wrong results with an incorrect number of electrons and nuclei.

While a hardware-efficient *Ansatz* does not require any special adaptation for the NEO wave function, we must introduce tailored modifications to the UCC *Ansatz* to extend it from the purely electronic to the NEO case. Here, we refer to the modified forms as NEOUCC. The parameterized unitary operators $\hat{U}_{CC}(\boldsymbol{t})$,

$$\hat{U}_{CC}(\boldsymbol{t}) = e^{\hat{T}(\boldsymbol{t}) - \hat{T}^\dagger(\boldsymbol{t})}, \tag{15}$$

are at the core of the UCC *Ansatz*. Acting upon some initial state $|\psi_{init}\rangle$ (typically the Hartree–Fock solution), it produces a trial wave function of the form

$$|\psi(\boldsymbol{t})\rangle = \hat{U}_{CC}(\boldsymbol{t})|\psi_{init}\rangle. \tag{16}$$

In general, the operator $\hat{T}$ is given as a sum of cluster operators of increasing order, i.e., the number of particles being acted on, o,

$$\hat{T} = \sum_o \hat{T}^{(o)} = \hat{T}^{(1)} + \hat{T}^{(2)} + \cdots. \tag{17}$$

In other words, $o = 1$ corresponds to one-particle excitations, $o = 2$ corresponds to two-particle excitations, and so on. In second quantization, these are expressed as

$$\hat{T}^{(1)} = \sum_{pq} t_q^p \, \hat{a}_p^\dagger \hat{a}_q, \tag{18}$$

$$\hat{T}^{(2)} = \sum_{pqrs} t_{rs}^{pq} \, \hat{a}_p^\dagger \hat{a}_q^\dagger \hat{a}_r \hat{a}_s. \tag{19}$$

Here, p, q, r, s index all spin orbitals, both protonic and electronic, and t represents coefficients to be optimized. Labeling the initially occupied spin orbitals with i, j, k, l and the initially unoccupied spin orbitals with a, b, c, d (analogous for protons using upper case: I, J, K, L and A, B, C, D), we can limit the cluster operators to a subset that conforms to chosen occupations of the initial state,

$$t_{ij\ldots IJ\ldots}^{ab\ldots AB\ldots} \prod_{\gamma,\lambda} \hat{a}_\gamma^\dagger \hat{a}_\lambda \prod_{\Gamma,\Lambda} \hat{a}_\Gamma^\dagger \hat{a}_\Lambda, \tag{20}$$

where $\gamma \subseteq \{a, b, \ldots\}$, $\lambda \subseteq \{i, j, \ldots\}$, $\Gamma \subseteq \{A, B, \ldots\}$, and $\Lambda \subseteq \{I, J, \ldots\}$. That is, annihilation (creation) operators will only act on occupied (unoccupied) spin orbitals, indexed as $a, b, c, d, \ldots$ $(i, j, k, l, \ldots)$ for electronic spin orbitals and as $A, B, C, D, \ldots$ $(I, J, K, L, \ldots)$ for protonic spin orbitals. Since we now obtain a mix of electron and proton operators, we introduce a dual superscript $\hat{T}^{(e,p)}$, where the first index, e, is the order of electronic excitation in each operator and p is similarly the order of the protonic excitations, that is,

$$\hat{T} = \sum_{e,p} \hat{T}^{(e,p)}, \tag{21}$$

where $\hat{T}^{(e,p)}$ is, e.g.,

$$\hat{T}^{(1,0)} = \sum_{ia} t_i^a \, \hat{a}_a^\dagger \hat{a}_i, \tag{22}$$

$$\hat{T}^{(0,1)} = \sum_{IA} t_I^A \, \hat{a}_A^\dagger \hat{a}_I, \tag{23}$$

$$\hat{T}^{(2,0)} = \sum_{abij} t_{ij}^{ab} \, \hat{a}_a^\dagger \hat{a}_b^\dagger \hat{a}_i \hat{a}_j, \tag{24}$$

$$\hat{T}^{(0,2)} = \sum_{ABIJ} t_{IJ}^{AB} \, \hat{a}_A^\dagger \hat{a}_B^\dagger \hat{a}_I \hat{a}_J, \tag{25}$$

$$\hat{T}^{(1,1)} = \sum_{aBiJ} t_{iJ}^{aB} \, \hat{a}_a^\dagger \hat{a}_i \hat{a}_B^\dagger \hat{a}_J. \tag{26}$$

The sum in Eq. (21) is normally truncated to some fixed order of e and p. Electronic structure VQE calculations are often limited to one- and two-particle cluster operators, while higher orders are ignored. This approximation is commonly referred to as Unitary Coupled Cluster Singles and Doubles (UCCSD). We will extend this notation to indicate when different orders are used for electrons, protons, and mixed terms, respectively. For example, an *Ansatz* using singles and doubles for electrons and nuclei, while using a mixed double consisting of single electronic and single nuclear excitations, a mixed triple consisting of double electronic and single nuclei operators, will be indicated as

$$\mathrm{UCCSDT}^{(2,1)} \Rightarrow \overbrace{\hat{T}^{(1,0)} + \hat{T}^{(0,1)}}^{\mathrm{S}} + \underbrace{\hat{T}^{(2,0)} + \hat{T}^{(1,1)} + \hat{T}^{(0,2)}}_{\mathrm{D}} + \hat{T}^{(2,1)}. \tag{27}$$

Note that the lack of superscript implies the use of all o-particle operators, i.e., $\mathrm{S} \Rightarrow \hat{T}^{(1,0)} + \hat{T}^{(0,1)}$, $\mathrm{D} \Rightarrow \hat{T}^{(2,0)} + \hat{T}^{(1,1)} + \hat{T}^{(0,2)}$, and $\mathrm{T} \Rightarrow \hat{T}^{(3,0)} + \hat{T}^{(2,1)} + \hat{T}^{(1,2)} + \hat{T}^{(0,3)}$. For systems with fewer electrons or protons than the excitation order, $n_e < e + p$ or $n_n < e + p$, only mixed operators are considered. For example, in the case of the hydrogen molecule, the possible triple and quadruple excitation operators are $\mathrm{T} \Rightarrow \hat{T}^{(2,1)} + \hat{T}^{(1,2)}$ and $\mathrm{Q} \Rightarrow \hat{T}^{(2,2)}$, respectively.

C. *Ansatz* initialization

The initial guess for parameters in any variational quantum algorithm is in many cases important to the performance of the calculation and can be crucial. Having a proper initial guess can reduce the number of evaluations needed to reach convergence, while also allowing convergence to the correct minimum (or maximum). As such, it proves wise to choose initial values for the *Ansatz* parameters with care. For hardware-efficient *Ansätze*, this is especially vital, where otherwise initial parameters often are chosen randomly, making convergence to a global minimum difficult.

The convergence problem arises due to several factors, e.g., the presence of local minima in the optimization landscape and the presence of barren plateaus.[73,75] As these problems compound with the number of variational parameters, it can be useful to solve an approximate problem in order to find a set of starting parameters closer to those of the true ground state. In the case of NEO calculations, one such approximate solution can be taken from the conventional electronic structure calculations under the BO approximation and with nuclei represented as classical point charges. Thus, we suggest initializing the parameters for the electronic part of the NEO calculation with those from the converged electronic structure calculation, inspired by our previous work.[76] For the case of the NEOUCC *Ansatz*, the circuit operators corresponding to electronic, nuclear, and mixed excitations can be separated according to Fig. 1. As $U_{T^{(e,0)}}(\boldsymbol{\theta}^e)$ operators are the same for NEO and electronic cases, one can initialize their parameters, $\boldsymbol{\theta}^e$, with those from the electronic calculation and set all other parameters to zero, $\boldsymbol{\theta}^{0,n} = \boldsymbol{\theta}^{e,n} = \mathbf{0}$.

For a hardware-efficient *Ansatz*, the parameter transfer becomes slightly more complicated as it is difficult to assign given parameters to specific nuclear or electronic parts. Therefore, special

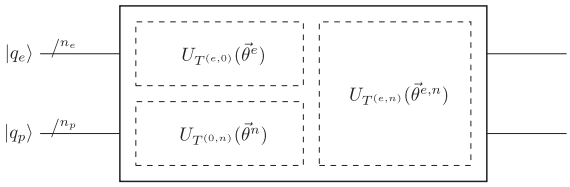


FIG. 1. Schematic circuit used for the initialization of the electronic parameters of the NEOUCC *Ansatz*. The conventional electronic operators can be considered as a subset of the NEO operators. n_e and n_p are the numbers of qubits representing electron and proton spin orbitals, respectively.

IV

care should be taken such that the parameter transfer only affects qubits associated with electronic orbitals. A generalized construction of the TwoLocal *Ansatz* can be created from sets of parameterized single qubit rotational gates separated by entangling gates applied between qubits. Both the electronic and NEO problems can be represented in this manner (see Fig. 2). In general, the number of layers needed to accurately describe the electronic problem, $d_{\mathrm{El.}}$, is smaller or equal to d_{NEO}, the number of layers needed to describe the NEO problem. With this in mind and a proper selection of entangling scheme, one can find a transfer protocol between electronic and NEO parameters. In Sec. IV B, we present in detail the two variants of parameter initialization employing CZ entangling layers. CZ entangling layers are beneficial in that they allow for the limited influence of electronic parameters on nuclear qubits if the nuclear parameters are unset (equal to zero).

The suggested initialization schemes assume that the electronic parameters could also be relaxed during the wave function optimization. It would be especially important for cases where multiple BO potential energy surfaces intersect. In addition, one should assure that the presence of additional nuclear orbitals in the NEO representation preserves the point group symmetry of the system. Nonetheless, even if the point group symmetry is the same, the corresponding $\mathbb{Z}_2$-symmetry reduction might introduce a difference in the sign pattern between the off-diagonal elements of the NEO and electronic qubit Hamiltonians. This stems from the difference in sign between eigenvalues of $\mathbb{Z}_2$-symmetry for the

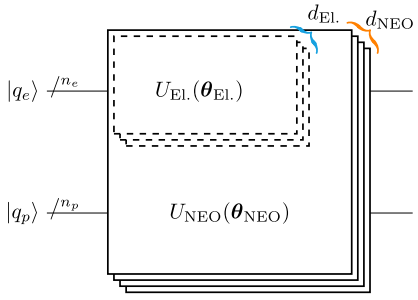


FIG. 2. Schematic circuit used for the initialization of the electronic parameters in the NEO hardware-efficient *Ansatz*. The electronic and NEO hardware-efficient circuits have $d_{\mathrm{El.}}$ and d_{NEO} layers, respectively. n_e and n_p are the numbers of qubits representing electron and proton spin orbitals, respectively.

NEO and electronic states and must be taken into account during the parameter transfer. We describe the procedure in detail in Sec. IV B.

D. Entanglement entropy measurement

The measurement of the entanglement entropy (or von Neumann entropy) can provide additional insights into the correlation between the two particle subsystems, namely, electrons and nuclei. This is particularly interesting when studying systems composed of distinguishable particles and will provide a direct measurement of the entanglement between the two subsystems. Here, we will introduce the subsystem entropy in the NEO framework as

$$s_i = -\operatorname{Tr}\left(\hat{\rho}^i \ln \hat{\rho}^i\right), \tag{28}$$

where $i \in \{e, n\}$, $\hat{\rho}^e = \mathrm{Tr}_n \hat{\rho}^{e,n}$, and $\hat{\rho}^n = \mathrm{Tr}_e \hat{\rho}^{e,n}$ are the corresponding reduced density matrices and $\hat{\rho}^{e,n}$ is the full density matrix. Alternatively, one can also monitor single-orbital entropy, as shown in Appendix A.

III. COMPUTATIONAL DETAILS

The NEO framework was implemented in a locally modified version of Qiskit.[74] All quantum computing calculations were performed as state vector simulations using variations of the TwoLocal *Ansatz* as well as NEOUCC *Ansätze*. The former employed parameterized R_y and R_z single rotation gates together with CZ entangling gates. The various entanglement patterns used in this work are summarized in Secs. IV B and IV C. All the classical algorithm calculations were carried out using the GAMESS-US software package.[77] In order to prepare the test systems and reference results, the Nuclear Electronic Orbitals Hartree–Fock (NEOHF), NEOFCI, and Nuclear–Electronic Orbital Complete Active Space Configuration Interaction (NEOCASCI)[17] calculations were performed. The system Hamiltonians were expanded in the orbitals obtained from the NEOHF calculations, and the corresponding NEO wave functions were optimized with the VQE algorithm in Qiskit. The SciPy[78] implementations of the Constrained Optimization By Linear Approximation (COBYLA),[79] Conjugate Gradient (CG),[80,81] and Sequential Least SQuares Programming (SLSQP)[82] methods were used for the optimization of the circuit parameters. The fermion-to-qubit transformation used to map the electronic and nuclear operators together with the employed qubit reduction techniques is described in Sec. III A. The additional qubit space reduction obtained by exploiting the NEO Hamiltonian's point group symmetries is discussed in Sec. III B.

A. Fermionic mapping and four-qubit reduction

The fermionic parity mapping[70] scheme is used in this work for both electronic and nuclear second-quantized operators. Note that we are only treating nuclei consisting of single protons quantum mechanically, and as such, only fermionic representations are necessary. We assume that the fermionic occupation number basis vectors are written in what we will refer to as particle-spin block representation,

$$|f\rangle = |\underbrace{f^p_{n^p_\beta+n^p_\alpha}, \ldots, f^p_{1+n^p_\alpha}}_{\beta\text{−spin prot.}}, \overbrace{f^p_{n^p_\alpha}, \ldots, f^p_1}^{\alpha\text{− spin prot.}}, \underbrace{f^e_{n^e_\beta+n^e_\alpha}, \ldots, f^e_{1+n^e_\alpha}}_{\beta\text{−spin elec.}}, \overbrace{f^e_{n^e_\alpha}, \ldots, f^e_1}^{\alpha\text{− spin elec.}}\rangle,$$
$$f_i^{e/p} \in \{0, 1\}, \tag{29}$$

where f_i^e and f_i^p are the spin orbital occupation numbers for electrons and protons, respectively, and $n^e_\alpha, n^e_\beta, n^p_\alpha, n^p_\beta$ are the number of α and β spin orbitals for the two types of particle. With parity encoding, instead of storing the occupation of individual spin orbitals, the qubits store information about the parity of the fermionic state as follows:

$$p_i^{e/p} = \bigoplus_{j=1}^{i} f_i^{e/p}, \tag{30}$$

where $\oplus$ represents addition modulo 2. A change to the parity encoding thus yields the parity state

$$|p\rangle = \underbrace{p^p_{n^p_\beta+n^p_\alpha}, \ldots, p^p_{1+n^p_\alpha}}_{\beta\text{−spin prot.}}, \overbrace{p^p_{n^p_\alpha}, \ldots, p^p_1}^{\alpha\text{− spin prot.}}, \underbrace{p^e_{n^e_\beta+n^e_\alpha}, \ldots, p^e_{1+n^e_\alpha}}_{\beta\text{−spin elec.}}, \overbrace{p^e_{n^e_\alpha}, \ldots, p^e_1}^{\alpha\text{− spin elec.}}\rangle,$$
$$p_i^{e/p} \in \{0, 1\}, \tag{31}$$

thus mapping the local state of each spin orbital to the state of several qubits. Note that the qubit mapping for each type of particle is performed within the corresponding subspace such that the parity of electronic qubits is not affected by the state of nuclear qubits, and vice versa.

The fermionic creation and annihilation operators in the parity basis have the following representations:

$$a_i^\dagger \equiv \frac{1}{2}\left[X^{\leftarrow}_{i+1} \otimes \left(X_i \otimes Z_{i-1} - iY_i\right)\right], \tag{32}$$

$$a_i \equiv \frac{1}{2}\left[X^{\leftarrow}_{i+1} \otimes \left(X_i \otimes Z_{i-1} + iY_i\right)\right]. \tag{33}$$

Here, we introduce $X^{\leftarrow}_{i+1}$ as an operator that acts on all qubits of index $i+1$ and higher within the electronic or nuclear subspaces, accounting for fermionic antisymmetry under the exchange of identical particles. For the edge case of $i = 1$, the expression simplifies to $\frac{1}{2}\left[X^{\leftarrow}_{i+1} \otimes \left(X_i + iY_i\right)\right]$, and similarly for $i = n^{e/p}_\alpha + n^{e/p}_\beta$, we get $\frac{1}{2}\left(X_i \otimes Z_{i-1} + iY_i\right)$.

The utilization of the parity encoding simplifies the identification of the discrete $\mathbb{Z}_2$-symmetries associated with the particle number operators within the α and β subspaces[83] for both electrons and nuclei. As we are dealing with a Hamiltonian based on the nonrelativistic Schrödinger equation, Eq. (1), the number of particles within the α and β subspaces are conserved, and therefore, the total parity numbers, such as $p^p_{n^p_\beta+n^p_\alpha}$, $p^p_{n^p_\alpha}$, $p^e_{n^e_\beta+n^e_\alpha}$, and $p^e_{n^e_\alpha}$, are constant. The corresponding four qubits can thus be treated classically, and their contribution to the expectation value of the Hamiltonian can be calculated beforehand, employing the "tapering" procedure described in Ref. 83.

B. Qubit reduction from Hamiltonian symmetries

If a molecular system obeys some point group symmetry, the number of qubits can be further reduced using methods introduced by Bravyi *et al.*[83] This approach allows for additional qubit tapering by identifying a $\mathbb{Z}_2$-symmetry of the Hamiltonian corresponding to the specific molecular point group symmetry.

In addition, when the system wave function can be restricted to specific subspaces of the total Hilbert space, one can further impose an additional $\mathbb{Z}_2$-symmetry, which confines the solutions in the region of interest. This may, for example, happen in the case of the hydrogen molecule if one is interested in the orthohydrogen nuclear isomer, the dominant nuclear spin isomer (75%) at room temperature. Thus, one can just consider the maximally spin-polarized triplet state for the protons. This allows for either alpha or beta spin orbitals to be ignored. To reduce our active space to the selected set of nuclear alpha or beta orbitals, we introduce (here for the case of orthohydrogen) a projector to states for which beta spin orbitals are unoccupied,

$$\hat{P} = \bigotimes_{J}^{\beta\,\text{nuc.}} \frac{1}{2}(I + Z) \bigotimes_{I}^{\alpha\,\text{nuc.}} I \bigotimes_{i}^{\text{elec.}} I, \tag{34}$$

and produce the projected Hamiltonian

$$\hat{H}_P = \hat{P}\hat{H}\hat{P}. \tag{35}$$

In this case, one can find as many $\mathbb{Z}_2$-symmetries as the number of beta spin orbitals, and the qubits corresponding to these spin orbitals can be therefore tapered off. This operation can also be performed in the Hamiltonian construction procedure by simply removing terms containing creation/annihilation operators acting on these spin orbitals.

IV. RESULTS

Our simulations focused on the evaluation of the NEO ground state energies and entanglement entropies for H_2 and malonaldehyde. We first introduce a detailed description of the two systems and describe the conditions under which the conventional quantum chemistry calculations were performed. We then describe (in Sec. IV B) the procedures for the initialization of the quantum computing calculations, as well as the transformations employed in the mapping of the fermionic electronic structure problem in the qubit space (see Secs. III A and III B). In Sec. IV C, we will report the results for the hydrogen molecule, discussing the performance of the different wavefunction *Ansätze*, including the one inspired by UCC and the more quantum native, hardware efficient, TwoLocal *Ansatz*. Finally, in Sec. IV D, we will present and discuss the results for the second and more challenging system, malonaldehyde, which is a prototypical system used to investigate quantum effects in molecular proton transfer reactions. In this case, the delocalization of the proton wave function between the donor and acceptor moieties makes the choice of the proton basis set and wave function *Ansatz* more challenging than in the case of H_2 molecule.

In both cases, we will also discuss the level of entanglement between the two quantum subsystems, the electronic and nuclear components, providing additional insights into the nature of the proton transfer process.

A. Test systems and classical reference calculations

We chose H_2 for testing the developed methodology, as the exact reference results are readily accessible and all possible types of interaction between particles, such as electron–electron, proton–proton, and electron–proton, are present. The NEO orbitals and corresponding integrals were prepared with the NEOHF method. For ease of calculation, the rather small split-valence double-ζ Gaussian basis set in 6-31 contraction scheme (6-31G)[84] basis set was used for the electronic orbitals, and the nuclear orbitals were constructed using the smallest basis set available in GAMESS-US,[77] split-valence double-ζ nuclear basis set composed of two uncontracted Cartesian S functions (DZSNB).[17] For the showcase in Sec. IV B, the electronic orbitals were constructed in single-ζ (minimal) basis set contracted from six Gaussian primitives (STO-6G). The basis set functions for electrons and nuclei were centered at the positions corresponding to the experimental ground state structure[85] with an inter-nuclear distance equal to 0.7414 Å. We must stress that, in general, one must use the floating orbital approach,[86–88] where the centers of the basis set functions are optimized to best fit the final wave function according to the variational principle. However, as it entails a more elaborate implementation, we prefer to postpone this step and perform all tests with fixed orbital centers, similar to the studies performed in Refs. 89–92. Note that our H_2 Hamiltonian, similar to the work in Ref. 93, contains contributions from global translational and rotational motion. As these contributions increase the total energy, they preferably must be eliminated.[90,94,95] We emphasize that our aim is not in delivering ultimate accuracy results but rather to establish the resource-efficient methodology for the quantum treatment of the electron–nuclear problem on near-term quantum computers. Based on the NEOHF orbitals, the reference total energy has been evaluated with the NEOFCI[17] method (see Table II). The same structure was used for the electronic Hartree–Fock (HF) calculation with a 6-31G basis set giving four orbitals for electrons. Consequently, these orbitals were used in Full Configuration Interaction (FCI) calculation for obtaining the reference energy. The reference energies for electronic and NEO cases are presented in Tables I and II, respectively.

TABLE I. H_2 ground state energies calculated with quantum computing *Ansätze* (UCCSD and NEOUCCS$^{(1,0)}$D$^{(2,0)}$) and classical algorithms (FCI and "exact diagonalization"). Two Hamiltonians are used: the electronic Hamiltonian ($\hat{H}_{\text{El.}}$) and the projected NEO Hamiltonian ($P_e\hat{H}_{\text{NEO}}P_e$). The UCCSD electronic parameters reproducing the FCI wave function provide a good initial guess for the NEOUCCS$^{(1,0)}$D$^{(2,0)}$ *Ansatz* (applied on $P_e\hat{H}_{\text{NEO}}P_e$) with an error of 8.35×10^{-5} hartree. Further optimization of parameters ("relaxed") recovers the ground state energy of $P_e\hat{H}_{\text{NEO}}P_e$ ("exact diagonalization").

Hamiltonian	Method	Parameters	Energy (hartree)
$\hat{H}_{\text{El.}}$	FCI		−1.151 683
	UCCSD	Electronic	−1.151 683
$P_e\hat{H}_{\text{NEO}}P_e$	Exact diagonalization		−1.065 040
	NEOUCCS$^{(1,0)}$D$^{(2,0)}$	Electronic	−1.064 956
		Relaxed	−1.065 040

TABLE II. Ground state energies for the H_2 NEO Hamiltonian obtained with different variants of the NEOUCC quantum computing *Ansatz* and the COBYLA optimizer. In general, a good agreement with NEOFCI is reached for all variants; however, one should emphasize the importance of the $T^{(2,1)}$ excitations in delivering excellent agreement with the reference. The calculations employ both "ordinary" and "advanced" initialization procedures, as described in Sec. IV B. Although the "advanced" initialization scheme improves slightly the accuracy, we expect more significant improvements in the case of a larger number of parameters.

	Method			
	Order	Initialization	Energy (hartree)	Iterations
NEOUCC	SD	Advanced	−1.066 039	856
		Ordinary	−1.066 039	1341
	$SDT^{(1,2)}$	Advanced	−1.066 039	1093
		Ordinary	−1.066 039	1674
	$SDT^{(2,1)}$	Advanced	−1.066 120	1377
		Ordinary	−1.066 121	2179
	SDT	Advanced	−1.066 121	1074
		Ordinary	−1.066 120	3945
	NEOFCI		−1.066 121	

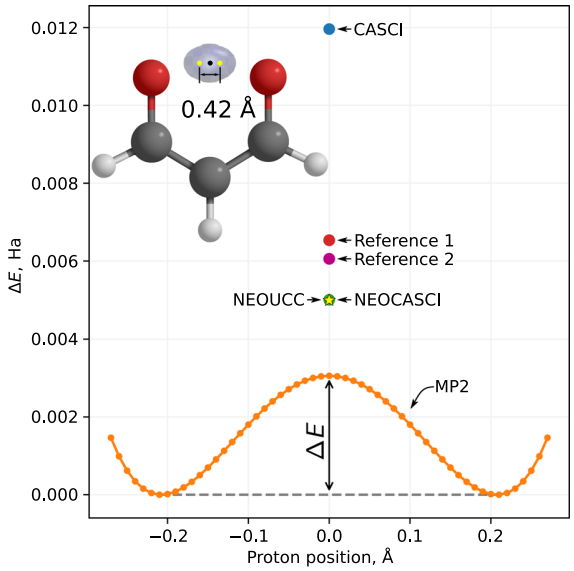


FIG. 3. Potential energy surface for the proton translocation process in malonaldehyde. The curve in orange corresponds to a MP2 calculation, in which the moving proton (as all other atoms) is described as a classical point charge and it is driven from one minimum to the other using a fixed Hamiltonian of C_{2V} symmetry, i.e., with all atoms (but the moving proton) fixed in their original position. This curve is mainly given as guidance to the eyes and does not correspond to the exact potential energy profile. The malonaldehyde structure together with the first protonic NEOHF orbital (isosurface value 10^{-4}) is given in the upper inset. The distance between the equilibrium positions is found to be 0.42 Å. Both NEO-CASCI (classical algorithm, green dot) and NEOUCC (quantum algorithm, yellow star) methods incorporate quantum proton effects and provide reasonable barrier values, while the classical description of the proton with CASCI (blue dot) leads to a considerable overestimation of the barrier. All calculations are performed using the MP2 geometries. Proton positions computed with the NEO approaches correspond to the expectation values of their position operators. The values labeled with "Reference 1" (red dot) and "Reference 2" (purple dot) are taken from Refs. 96 and 97, respectively.

For the second test system, we chose malonaldehyde. This molecule is an extensively studied benchmark system for quantum effects involving intramolecular hydrogen bonds. The key structural feature is the hydrogen bond O–H–O for which there are two possible asymmetric configurations. This leads to a double-well potential with the states connected by a proton barrier. The experimental ground state structure[98] has C_s point group symmetry, and the proton transfer occurs at a symmetric transition state structure in C_{2v} symmetry.[17,99] To obtain a reasonable starting structure for the description of the proton transfer process, we first optimized the malonaldehyde structure with the second-order Møller–Plesset (MP2)[100] in C_{2v} symmetry and the 6-31G basis set. The final structure corresponds to the transition state (TS) geometry.[17,99] After fixing all the coordinates but one, the search for minima was performed in C_s symmetry. This revealed two possible equilibrium locations for the proton separated from each other by 0.42 Å as well as the location of the barrier's maximum. In the inset of Fig. 3, we marked the equilibrium positions with yellow dots; they are symmetrically arranged with respect to the position of the barrier maximum marked with a black dot. These coordinates were used as the centers for the nuclear basis functions in three different NEOHF calculations performed with DZSNB and 6-31G basis sets for the nuclei and the electrons, respectively. Two of these NEOHF calculations were performed in the C_{2v} symmetry corresponding to the TS geometry; in the first case, two basis functions were placed in correspondence of the equilibrium positions (yellow points), while in the second case, two more basis functions were added at the barrier's maximum (black point). The former will be referred to as the "2-center C_{2v}" calculation (consisting of four nuclear orbitals), while the latter, the "3-center C_{2v}" calculation, makes use of six nuclear orbitals. By enabling a full relaxation of the three-center structure, we can obtain a good estimation of the barrier height (as the minimum energy in the C_{2v} symmetry). The third NEOHF calculation corresponds to the relaxed "1-center C_s" configuration and is described by two nuclear orbitals placed at one of the equilibrium locations (yellow points). The localization of the proton in one of the two equivalent minima (bonded geometries) is a consequence of the relaxation of the molecular scaffold from the C_{2v} to the asymmetric C_s geometry. In all cases, only one proton was treated quantum mechanically, while all other nuclei were considered as classical point charges. We also kept the same number of electronic orbitals associated with the "quantum" proton, placing them in correspondence of all three centers(yellow and black), with the aim of achieving an equal description of the electronic wave function in the "1-center C_s" and "3-center C_{2v}" setups. Due to the presence of the classical nuclei, the NEO Hamiltonian does not reflect the global rotational and translational degrees of freedom.[101] For each of the three sets of NEOHF orbitals ("1-center C_s," "2-center C_{2v}," and "3-center C_{2v}"), we performed NEOCASCI calculations to obtain the corresponding reference energies (see Table VII). To keep the electronic active space compact, in all calculations, the 17 low-lying core orbitals for

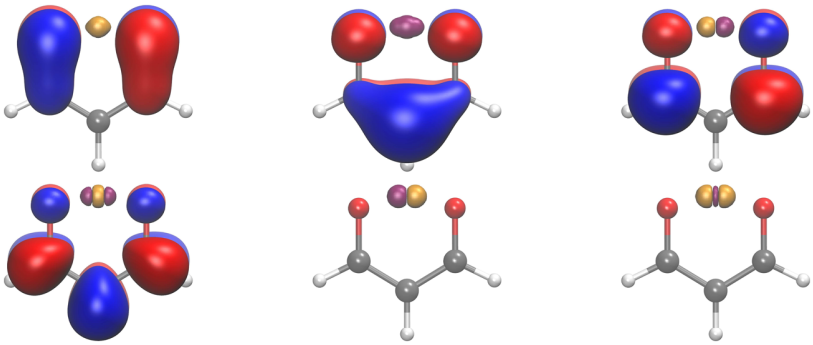


FIG. 4. Electronic (red/blue) and nuclear (purple/orange) orbitals (isosurface value is 0.05) included in the active space of NEOCASCI calculation for malonaldehyde in the "3-center C_{2v}" setup. The orbitals have been prepared with a NEOHF calculation using a DZSNB and a 6-31G basis sets for the nuclei and the electrons, respectively.

electrons were considered fully occupied and only four orbitals hosting four electrons were considered active. Concerning the nuclear active space, we used all two orbitals, four orbitals, and six orbitals for "1-center C_s," "2-center C_{2v}," and "3-center C_{2v}" setups, respectively. The results for the reference NEOCASCI calculations are reported in Table VII, while the electronic and nuclear NEOHF orbitals used in the active space for the "3-center C_{2v}" setting are shown in Fig. 4.

B. Initialization of the electronic parameters in NEO *Ansatz*

As discussed in Sec. II C, the choice of initial parameters in any VQE *Ansatz* can have a significant influence on the convergence rate as well as the accuracy of a calculation. Here, we show that the optimized wave function of the electronic subsystem offers a good starting point to the NEO algorithm. As already mentioned earlier, the transfer of parameters between the electronic and NEO *Ansätze* might need additional modifications if a point group symmetry is manifest. To this end, one needs to consider the different sign patterns occurring in the NEO and electronic Hamiltonians after applying the $\mathbb{Z}_2$-symmetry reduction corresponding to a certain point group symmetry. We illustrate this procedure on the example of the hydrogen molecule using NEOHF orbitals constructed with minimal size STO-6G and DZSNB basis sets for electrons and nuclei, respectively, as explained in Sec. IV A. In this way, we can deal with only four orbitals for the nuclei and two orbitals for the electrons. This setup results in a NEO qubit Hamiltonian spanning 12 spin orbitals (eight nuclear spin orbitals and four electronic spin orbitals), while the electronic Hamiltonian spans just four spin orbitals. The $\mathbb{Z}_2$-symmetry arising from the D_{2h} point group of the electronic qubit Hamiltonian in parity mapping is given by the following operator:

$$\mathbb{Z}_2^{\mathrm{El.}} = IZIZ. \tag{36}$$

Applying tapering to the electronic Hamiltonian leads to a projection into the symmetry subspace corresponding to the HF state with $\mathbb{Z}_2^{\mathrm{El.}}$ eigenvalue equal to -1. Thus, after two-qubit reduction[83] and symmetry tapering, the resulting electronic Hamiltonian will take the following form:

$$\hat{H}_{\mathrm{El.}} = -0.322\,833 \times I - 0.803\,007 \times Z - 0.180\,939 \times X. \tag{37}$$

For the NEO qubit Hamiltonian in parity encoding, the corresponding $\mathbb{Z}_2$-symmetry is

$$\mathbb{Z}_2^{\mathrm{NEO}} = IZZZIZZZIZIZ. \tag{38}$$

However, in this case, the symmetry operator eigenvalue is $+1$. Thus, the eigenvalue with a different sign will be incorporated into the Hamiltonian during the symmetry tapering. To reveal this effect, we further perform four-qubit reduction according to Sec. III A and project the NEO Hamiltonian onto the electronic subspace. The corresponding projector, P_e, is similar to the one introduced in Sec. III B except that it additionally projects the NEO qubit Hamiltonian onto states where the nuclear alpha spin orbitals occupied according to the lowest NEOHF occupation. Thus, the resulting NEO qubit Hamiltonian,

$$\hat{P}_e\hat{H}_{\mathrm{NEO}}\hat{P}_e = -0.217\,313 \times I - 0.816\,590 \times Z + 0.181\,576 \times X, \tag{39}$$

is limited to the states where only the energetically lowest triplet NEOHF occupation for protons is allowed. One can clearly recognize in Eq. (39) that the off-diagonal term acquires a change of sign when compared to the electronic case, Eq. (37). This difference in the sign signature is also present for the parameters in the electronic and NEO *Ansätze*. One way to solve this issue would be to perform a signature similarity transformation of the electronic Hamiltonian. For the specific case discussed here, this will correspond to

$$Z\hat{H}_{\mathrm{El.}}Z = -0.322\,833 \times I - 0.803\,007 \times Z + 0.180\,939 \times X. \tag{40}$$

Since such a transformation is isospectral,[102] one can use the transformed electronic Hamiltonian to find the parameters of the electronic *Ansatz* and then apply these parameters as an initial guess for the NEO *Ansatz*. In this work, this "signature similarity transformation" is performed by comparing and fixing the signs of the Pauli words of the projected NEO Hamiltonian and of the electronic Hamiltonian according to the expressions in Eqs. (37) and (39).

The similar weights (coefficients) obtained for the projected NEO and in the electronic Hamiltonians, Eqs. (39) and (40), attests to the efficacy of the proposed parameter initialization methodology. However, before transferring electronic parameters to the NEO *Ansatz*, it could be beneficial to additionally optimize them using the projected NEO Hamiltonian $\hat{P}_e\hat{H}_{\mathrm{NEO}}\hat{P}_e$ [Eq. (39)]. More

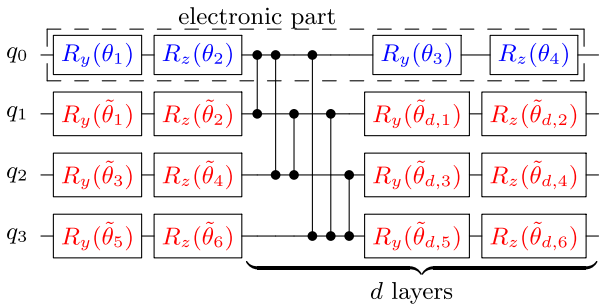


FIG. 5. The "expanded" variant for the initialization of the electronic parameters in hardware-efficient *Ansatz*. The circuit is designed for the case of four-qubit NEO Hamiltonian, corresponding to H_2 using minimal basis sets, STO-6G for electrons and DZSNB for nuclei. We use an "all-to-all" entangling scheme with CZ gates applied between all qubits. The electronic subsystem is delimited by a dashed box. Parameters in blue are initialized according to the results of the electronic subsystem optimization. All parameters in red: $\tilde{\theta}_i$ $(i = 1, 6)$ and $\tilde{\theta}_{d,i}$ $(i = 1, 6)$ are initialized to 0, with the exception of $\tilde{\theta}_{d_{\max},1}$, which is set to π ($d_{\max}$ is the maximum number of repetitions). While for this specific case, one entangling layer, $d = 1$, would be sufficient, for larger systems, more layers will be required.

details will be given in Sec. IV C. In the following, we limit ourselves to the case of the initialization of the electronic parameters for the hardware-efficient *Ansatz* for H_2 in the minimal basis set. After performing symmetry tapering and four-qubit reduction and restricting to nuclear triplet states according to Secs. III A and III B, one finally obtains a four-qubit NEO Hamiltonian [See Appendix D, Eq. (D1)]. The first qubit (of the four-qubit register) corresponds to the electronic state, while the other three qubits describe the nuclear state (see Fig. 5). One can initialize electronic parameters

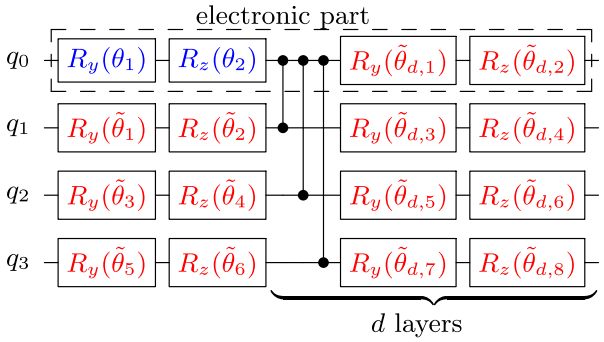


FIG. 6. The "stacked" variant for the initialization of the electronic parameters in hardware-efficient *Ansatz*. The circuit is designed for the case of the four-qubit NEO Hamiltonian, corresponding to H_2 using minimal basis sets, STO-6G for electrons and DZSNB for nuclei. The electronic and nuclear parameters are shown in blue and red, respectively. In this case, entangling CZ gates are only applied between electronic and nuclear subsystem qubits. The electronic subsystem is delimited by a dashed box. Parameters in blue are initialized according to the results of the electronic subsystem optimization. All parameters in red: $\tilde{\theta}_i$ $(i = 1, 6)$ and $\tilde{\theta}_{d,i}$ $(i = 1, 8)$ are initialized to 0, with the exception of $\tilde{\theta}_1$, which is set to π. The choice of an even number of stacked entangling layers, d, is preferable as it avoids phase flip induced by the CZ gates.

simply by transferring the wave function of the electronic subsystem into the NEO *Ansatz*, as shown in Fig. 5. All nuclear parameters must then be initialized to zero except for the second R_y gate on qubit q_1. In this specific case, it is set to π resulting in the lowest energy triplet nuclear occupation. This procedure, denoted as "expanded," can be easily generalized to larger system sizes. A second more economic variant only requires the additions of entangling gates between the electronic and the nuclear qubit subregisters, as shown in Fig. 6. This variant, referred to as "stacked," can be also generalized to larger systems and is advantageous in cases with weak entanglement between the protons. It is worth mentioning that for the electronic parameters (in blue color in Fig. 6) to remain as close as possible to their initial values (i.e., for avoiding the phase flip induced by the CZ gate), the number of entangling layers is required to be even.

C. Quantum simulation: The hydrogen molecule

Initially, we have performed VQE optimization of the electronic UCCSD *Ansatz* for the ground state of the H_2 molecule employing an electronic Hamiltonian with classical nuclei, $\hat{H}_{\mathrm{El}}$. The final energy agrees well with the reference value obtained from FCI calculations (see Table I). To verify the validity of the electronic UCCSD parameters for the NEOUCCSD *Ansatz*, we employ the projected NEO Hamiltonian, $P_e\hat{H}_{\mathrm{NEO}}P_e$, in a manner similar to Eq. (39) in Sec. IV B. After projection and tapering, the five-qubit NEO Hamiltonian considers only states where the nuclear part corresponds to the lowest energy triplet nuclear NEOHF occupation and has the same size as the electronic qubit Hamiltonian. The corresponding ground state energy of 1.065 040 hartree has been calculated by exact diagonalization of $P_e\hat{H}_{\mathrm{NEO}}P_e$. Initializing the electronic parameters in $\mathrm{NEOUCCS}^{(1,0)}\mathrm{D}^{(2,0)}$ (see notation in Sec. II B) to the values from the electronic UCCSD *Ansatz*, we have evaluated the corresponding expectation value, $\langle P_e\hat{H}_{\mathrm{NEO}}P_e\rangle$. One can see from Table I that the resulting expectation value approximates the ground state reference energy for $P_e\hat{H}_{\mathrm{NEO}}P_e$ within an error of less than 10^{-4} hartree. By letting the electronic parameters further relax in optimization employing the $\mathrm{NEOUCCS}^{(1,0)}\mathrm{D}^{(2,0)}$ *Ansatz* and $P_e\hat{H}_{\mathrm{NEO}}P_e$, one removes this error, recovering the exact reference energy (see Table I). Having resolved the reduced problem based on the $P_e\hat{H}_{\mathrm{NEO}}P_e$ Hamiltonian, we proceed to the full eight-qubit NEO Hamiltonian. After symmetry tapering, four-qubit reduction, and projection to the nuclear triplet state according to Secs. III B and III A, the electronic state is spanned by five qubits, and three qubits are used for the nuclear state. The results for the VQE energy optimizations employing various types of NEOUCC parameterizations with advanced and ordinary parameter initialization are presented in Table II. For the advanced initialization procedure, we employ relaxed electronic parameters obtained with $\mathrm{NEOUCCS}^{(1,0)}\mathrm{D}^{(2,0)}$ and $P_e\hat{H}_{\mathrm{NEO}}P_e$ as discussed above. All variants of the NEOUCC parameterization in Table II contain double and single excitations for nuclei and electrons, which will be indicated as

$$\mathrm{SD} \equiv \mathrm{S}^{(1,0)}\mathrm{S}^{(0,1)}\mathrm{D}^{(0,2)}\mathrm{D}^{(1,1)}\mathrm{D}^{(2,0)}. \tag{41}$$

If mixed electronic-nuclear double excitations are added, then the resulting $\mathrm{NEOUCCDD}^{(1,1)}$ *Ansatz* deliver the results with an error

below 10^{-4} hartree. The addition of the mixed triple, $T^{(1,2)}$, excitations in the NEOUCCSDD$^{(1,1)}$T$^{(1,2)}$ *Ansatz* does not decrease the error further. If instead one adds the mixed triple excitation with double electronic excitation, $T^{(2,1)}$, the error decreases below 5×10^{-6} hartree. Including both variants of mixed triple excitation brings the error below 10^{-6} hartree. The insignificance of proton–proton correlation as compared to electron–electron correlation is well in line with previous research.[92] It is clear that the $D^{(1,1)}$ and the $T^{(2,1)}$ operators are the primary contributors to the lowering of the energy, while the $T^{(1,2)}$ operator has minimal effect. The difference in final NEOUCC energy between optimizations with advanced initialization (set to relaxed electronic parameters) and ordinary initialization (set to the NEOHF state) fall close to the convergence tolerance, $\epsilon \leq 10^{-6}$ hartree. For NEOUCC *Ansätze*, the advanced initialization scheme improves accuracy for the cases with a large number of variational parameters, such as NEOUCCSDT. However, the main benefit lies in a faster convergence rate, in certain cases reaching a fourfold speedup (see Table II). Additional analysis for convergence of the NEOUCC *Ansatz* using different initialization schemes can be found in Appendix B.

Table III presents the assessment for the contribution of each type of excitation operator to the correlation energy in the NEO wave function. The correlation energy within the chosen basis set amounts to 0.024 634 hartree. If one uses double and single excitation operators for the protons only, then energy cannot decrease lower than NEOHF. Adding the single and double operators, NEOUCCSD, for electrons brings the energy nearly to the same value as for the case when only electronic operators are included (see Table III). This can be attributed to the low correlation between protons. However, the mixed electron–nuclear operators make a substantial contribution to the correlation energy. Specifically, addition of mixed double excitation, $D^{(1,1)}$, brings error below 10^{-4} hartree. If one also includes mixed triple operators, $T^{(2,1)}$, the error falls further below 5×10^{-6} hartree. The ultimate accuracy of 5×10^{-7} hartree is reached if all excitation operators up to triple-order are included.

We supplement our study of the correlation nature in the H_2 molecule with the measure of entanglement entropy. Specifically, we evaluate the entanglement between electronic and nuclear subsystems, according to Sec. II D. We find that the entanglement for the H_2 molecule at equilibrium based on the NEOFCI wave function amounts to 0.0069. The same value can be accurately reproduced with the NEOUCC *Ansatz* (see Table IV) when quadruple excitation operators are included. To monitor the change in entanglement upon dissociation, we also evaluated NEOFCI entropies for distances of 1.9582 and 3.1751 Å. Note that these quantities correspond to the separation between the centers hosting the nuclear and electronic orbitals and are taken as an approximated measure of the inter-atomic distance. Table IV shows how the entanglement between electron and proton subsystems (according to NEOFCI) strengthens to 0.0103 during the "bond extension" and then slightly declines to 0.0097 after "dissociation." The corresponding NEOUCC results are in good agreement with the reference calculations, demonstrating the reliability of the NEOUCC approach.

Although a UCC type *Ansatz* can be very accurate, in general, they cannot yet be efficiently implemented on current quantum hardware. Thus, in the following, we introduce an optimal strategy for the application of the TwoLocal *Ansatz* with the NEO Hamiltonian, employing the parameter initialization procedures discussed in Sec. IV B. A comparison of resource requirements between different implementations can be found in Table X of Appendix C. Similar to the NEOUCC case, we first optimize the TwoLocal *Ansatz* for the electronic Hamiltonian and then relax electronic parameters further, employing the $P_e\hat{H}_{\mathrm{NEO}}P_e$ Hamiltonian. For both Hamiltonians, we used the TwoLocal *Ansatz* with eight entangling layers (see Figs. 2 and 5). As can be seen from Table V, the TwoLocal *Ansatz* performs similarly to the NEOUCC *Ansatz* in terms of accuracy, confirming the validity of the advanced parameter initialization. However, the eight-qubit NEO Hamiltonian represents a difficult optimization case for the TwoLocal *Ansatz* with eight entangling layers when the ordinary parameter initialization is used. In fact, the initialization of the circuit parameters is mostly set randomly, failing often to deliver energies below the NEOHF energy. One can see from Table VI that for the best optimization case,

TABLE III. The ground state energies for the H_2 NEO Hamiltonian obtained with different variants of the NEOUCC quantum computing *Ansatz* and classical computational methods (NEOHF and NEOFCI). Increasing the excitation order in the NEOUCC *Ansatz* lowers the error in the calculation. The mixed $D^{(1,1)}$ nuclear–electron operators have a larger contribution to the correlation than T operators, while the mixed triple operators with higher electronic excitation order, $T^{(2,1)}$, play a more important role than $T^{(1,2)}$. The significance of pure nuclear operators is minimal.

	Method	Energy (hartree)	Error (hartree)
	NEOFCI	−1.066 121	
	NEOHF	−1.041 487	0.024 634
NEOUCC	$S^{(0,1)}D^{(0,2)}$	−1.041 487	0.024 634
	$S^{(1,0)}D^{(2,0)}$	−1.065 040	0.001 082
	$S^{(1,0)}D^{(2,0)}D^{(1,1)}$	−1.066 037	0.000 084
	$S^{(1,0)}D^{(2,0)}T^{(2,1)}$	−1.065 063	0.001 059
	$S^{(1,0)}D^{(2,0)}D^{(1,1)}T^{(2,1)}$	−1.066 117	0.000 004
	$SD^{(2,0)}D^{(0,2)}$	−1.065 049	0.001 073
	$SDT^{(2,1)}$	−1.066 121	0.000 001
	SDT	−1.066 121	0.000 001

TABLE IV. Von Neumann entropy between the electronic and nuclear subspaces (Sec. II D) at three points along the H_2 dissociation curve. The entropy increases with the extension of the bond length, converging to 0.0097 upon dissociation. Although the NEOUCCSD quantum computing *Ansatz* fails to predict this trend, increasing the order of excitations, one can accurately reproduce the reference value obtained with the classical NEOFCI approach.

		Separation		
	Method	0.7414Å	1.9582Å	3.1751Å
	NEOFCI	0.0069	0.0103	0.0097
NEOUCC	SD	0.0062	0.0050	0.0037
	SDT	0.0068	0.0104	0.0098
	SDTQ	0.0068	0.0104	0.0098

IV

TABLE V. H_2 ground state energies calculated with the TwoLocal and two classical approaches: FCI and "exact diagonalization." Two Hamiltonians are used: the electronic Hamiltonian ($\hat{H}_{\text{El.}}$) and the projected NEO Hamiltonian ($P_e\hat{H}_{\text{NEO}}P_e$). The TwoLocal electronic parameters reproducing the FCI wave function provide a good initial guess for the NEO calculation with the Hamiltonian $P_e\hat{H}_{\text{NEO}}P_e$, leading to an error of 8.35×10^{-5} hartree. Further optimization of parameters ("relaxed") reduces the error to 3.83×10^{-7} hartree compared to the ground state energy of the $P_e\hat{H}_{\text{NEO}}P_e$ ("exact diagonalization").

Hamiltonian	Method	Parameters	Energy (hartree)
$\hat{H}_{\text{El.}}$	FCI		−1.151 683
	TwoLocal	Electronic	−1.151 683
$P_e\hat{H}_{\text{NEO}}P_e$	Exact diagonalization		−1.065 040
	TwoLocal	Electronic	−1.064 956
	TwoLocal	Relaxed	−1.065 039

TABLE VI. Ground state energies for the H_2 NEO Hamiltonian obtained with various implementations of the TwoLocal quantum *Ansatz* and with the classical NEOFCI approach. The "advanced" initialization scheme for electronic parameters performs significantly better than the "ordinary" one for both CG and SLSQP optimizers. The "stacked" initialization with 14 entangling layers delivers the most accurate result with an error of 2×10^{-4} hartree using the CG optimizer.

			Energy (hartree)	
	Method	Layers	SLSQP	CG
TwoLocal	Ordinary		−1.013 923	−1.014 310
	Expanded		−1.065 054	−1.065 172
	Stacked	$d = 3$	−1.065 175	−1.065 633
		$d = 4$	−1.065 056	−1.065 381
		$d = 10$	−1.065 055	−1.065 663
		$d = 11$	−1.064 975	−1.065 814
		$d = 12$	−1.065 055	−1.065 720
		$d = 13$	−1.064 939	−1.065 691
		$d = 14$	−1.065 054	−1.065 889
		$d = 20$	−1.065 056	−1.065 666
	NEOFCI		−1.066 121	

the error is still larger than for five-qubit TwoLocal *Ansatz* applied to $P_e\hat{H}_{\text{NEO}}P_e$.

When setting the initial parameters of the TwoLocal *Ansatz* to the values of the relaxed electronic structure calculation according to the "expanded" variant with eight entangling layers (see Figs. 2 and 5 and Table VI), the energy decreases below the value from the ordinary initialization, giving an error of 10^{-3} hartree. The "stacked" variant of the electronic parameter initialization (Fig. 6) performs better (see Table VI), giving an error of about 2×10^{-4} hartree when 14 stacked layers are used. It should also be noted that CG optimizer performed best in the case of TwoLocal *Ansatz*, while SLSQP and COBYLA deliver similar results.

D. Quantum simulation: Malonaldehyde

We start our malonaldehyde investigation with the "2-center C_{2v}" setup, which corresponds to the situation in which the proton is equally shared between the two oxygen atoms (see Fig. 3). A similar setup was already proposed in Refs. 17 and 103 for the evaluation of double well splitting. However, our aim is different and consists mainly in providing benchmark results (more specifically, ground state energies) for the validation of the quantum NEO algorithm. Using the same transformations as in Sec. IV C, we converted the second quantized Hamiltonian based on 16 spin orbitals to an eight-qubit Hamiltonian, five qubits spanning the electronic subspace, and three qubits spanning the nuclear one. Increasing excitation order from double to quadruple in the NEOUCC *Ansatz*, we could recover the NEOCASCI reference results and reach 10^{-6} hartree accuracy, as shown in Table VII. The results for the entanglement entropy are presented in Table VII and are also in very good agreement with the reference values.

For evaluation of the proton transfer barrier, we used the "3-center C_{2v}" (top of the barrier) and "1-center C_s" (bottom of the potential, see Fig. 3) setups, as described in Sec. IV A. Based on the NEOCASCI calculations at these two setups, the barrier is estimated to be 0.005 011 hartree. This is rather close to the values obtained in Ref. 96 [full dimensional calculations with near basis-set-limit frozen-core CCSD(T)] and Ref. 97 (*in silico* transient x-ray absorption spectroscopy) of 0.0065 and 0.0061 hartree, respectively. The results concerning the TS barriers (summarized in Fig. 3) are reported for completeness. Since they are performed with different levels of approximation (classical vs quantum proton), they cannot be directly compared and any potential correspondence may be considered fortuitous. On the other hand, using a fully classical approach, such as Complete Active Space Configuration Interaction (CASCI) for the evaluation of the barrier, we observed a severe overestimation of the barrier, which reaches a value of 0.011 96 hartree (see Fig. 3 and Table VII). Note that in Fig. 3 we also report the full potential energy surface computed with MP2 and a classical proton. However, at this small basis set size for the expansion of the electronic wave function, we can only expect a qualitative description of the process, which we only report as a guide to the eye. Within the quantum computing context, the qubit Hamiltonian for the "3-center C_{2V}" setup is spanned by ten qubits, five qubits for electronic subspace and five qubits for nuclear subspace. The "1-center C_S" qubit Hamiltonian is instead spanned by seven qubits, six qubits for electronic subspace and one qubit for nuclear subspace. The energies for both systems estimated with various NEOUCC schemes are presented in Table VII. Already with a rather modest NEOUCCSD *Ansatz*, we observe very good agreement with the NEOCASCI reference, with a deviation smaller than 10^{-4} hartree. The error decreases upon inclusion of higher order excitation operators in NEOUCC and gets below 2×10^{-6} hartree for NEOUCCSDTQ. These calculations confirm the quality and the potential of the NEOUCC *Ansatz* for the calculation of quantum nuclei effects and prove its fast convergence toward the exact solution.

In Table VII, we summarize the values for the barrier energies together with the ones for the von Neumann entropy, s_n [Eq. (28)], associated with the entanglement between the nuclear and electronic subsystems. In addition, in this case, the NEOUCC results are in

TABLE VII. Energies and von Neumann entropies, s_n [Eq. (28)], of malonaldehyde obtained with different NEOUCC parameterizations and NEO Hamiltonians based on "3-center C_{2v}," "2-center C_{2v}," and "1-center C_s" setups. The barrier height, ΔE, is evaluated based on the difference between energies corresponding to "3-center C_{2v}" and "1-center C_s" setups. We observe good agreement with barrier height obtained in Refs. 96 and 97 for all variants of NEOUCC, while the electronic CASCI method severely overestimates it. However, since these results were obtained with different levels of approximation (classical vs quantum proton), they cannot be directly compared, and any potential correspondence may be considered fortuitous.

		"1-center C_s"		"2-center C_{2v}"		"3-center C_{2v}"		
	Method	Energy (hartree)	Entropy	Energy (hartree)	Entropy	Energy (hartree)	Entropy	ΔE (hartree)
	Reference 96							0.006 534
	Reference 97							0.006 056
	CASCI	−265.528 820				−265.516 859		0.011 962
	NEOCASCI	−265.490 948	0.0000	−265.452 062	0.0866	−265.485 937	0.0044	0.005 011
	SD	−265.490 685	0.0000	−265.451 123	0.0616	−265.485 761	0.0031	0.004 924
NEOUCC	SDT	−265.490 943	0.0000	−265.452 051	0.0862	−265.485 936	0.0044	0.005 007
	SDTQ	−265.490 947	0.0000	−265.452 062	0.0866	−265.485 939	0.0045	0.005 009

very good agreement with the NEOCASCI references for all settings: "1-center C_s," "2-center C_{2v}," and "3-center C_{2v}." Assuming that the estimate for the C_s system is not very sensitive to the basis set size, our calculations support the picture in which the level of electron–nuclear entanglement increases as the proton is transferred to the top of the barrier (however, this observation cannot be conclusive in view of the small basis set used for the electronic and nuclear problem, as well as of the rigidity imposed to the molecular scaffold). Further investigations are needed to shed light on the convergence of the entanglement entropy between electronic and nuclear subsystems as a function of the nuclear basis set size.

V. DISCUSSION AND CONCLUSIONS

In this work, we developed and demonstrated a quantum algorithm for the treatment of quantum nuclear degrees of freedom on near-term quantum computers, based on the NEO approach.[17] This approach allows the description of both components of molecular systems, i.e., electrons and nuclei, at the same footing, namely, at a quantum level, enabling the correct characterization of important nuclear quantum effects, such as proton transfer and non-adiabatic effects. From the quantum computing perspective, a quantum treatment of the full molecular wave function gives us the opportunity to extend the potential exponential quantum advantage already investigated for the electronic subsystem to the multi-component electron–nuclear wave function. In addition, this method will enable a rigorous estimation of the level of entanglement between the electronic and nuclear subsystems, providing a new and interesting tool for the understanding of post-BO effects. In order to make our approach more suited to near-term quantum computing, we implemented an embedding scheme, which allows the restriction of the nuclear quantum wave function to a selected number of atoms, while keeping the rest of the nuclei described at a classical level (i.e., as point charges).

The ability of NEOUCC *Ansätze* to deliver the ground state energy in noiseless VQE calculations was proven in two main applications: the hydrogen molecule and malonaldehyde. In the latter case, only the shared proton between the two carbonyl groups was "quantized," while all other nuclei in the molecules were treated classically. Our results agree with conventional NEOFCI and NEOCASCI calculations to within an error of about 10^{-6} hartree.

Our NEOUCC calculations revealed that apart from the high contribution to the correlation energy from interactions between electrons, a substantial part of correlation energy is attributed to the electron–nuclear interaction. Specifically, double electron–nuclear excitations are required for reducing the error below 10^{-4} hartree, and only after the addition of triplet $T^{(2,1)}$ excitations, one can reach an accuracy of about 10^{-6} hartree. The importance of nuclear–electron quantum correlations is also demonstrated by the significant amount of nuclear–electron von Neumann entropy, which increases at bond breaking.

Moreover, we investigated the possibility of utilizing hardware-efficient *Ansätze* as more compact representations of the NEO wave functions. We showed that the parameter initialization scheme plays an important role in the optimization process. In particular, to reach an accuracy below 5×10^{-4} hartree, we introduced several ways of initialization of the NEO electronic wave function parameters from the converged electronic calculation. Such an efficient initialization procedure is not restricted to the use of hardware-efficient *Ansätze* and can also significantly increase convergence in the case of NEOUCC wave functions. In addition, the same scheme can also be employed beyond the typical cases for which the BO separation is employed, for instance, in the case of weakly interacting molecular fragments.

Furthermore, we introduced a scheme for reducing the dimensionality of the effective Hamiltonian through the exploitation of those symmetries inherent in the NEO approach. Specifically, we have generalized the two-qubit reduction, already employed for the electronic wave function, extending it to the nuclear subsystem. The dimensionality of the nuclear Hilbert space is further reduced by leveraging the point group symmetry inherent in the molecular system. As this last step might have undesired consequences on the initialization of the electronic parameter, we have introduced and discussed a procedure for resolving the issue. In the last step, when dealing with nuclei of the same spin polarization, we can additionally project out unpopulated spin orbitals from the

Hamiltonian and remove corresponding qubits. This was demonstrated in this work for both the orthohydrogen spin isomer and malonaldehyde with a single proton described as a quantum particle. Thus, for the qubit Hamiltonian in the case of orthohydrogen, we succeeded in reducing the number of qubits from 16 to 8. Similarly, for the malonaldehyde instead of using a 20-qubit Hamiltonian, we demonstrated the possibility to safely reduce the system to an effective ten-qubit Hamiltonian. Reduction in the number of qubits leads to a significant decrease in the number of two-qubit entangling gates in the *Ansatz* as well as the total number of terms in the Hamiltonian required for measurement in VQE, making the entire algorithm more suited for near-term hardware calculations.

IV

In conclusion, we have proposed and implemented a resource-efficient quantum algorithm for the simulation of molecular systems, where both electrons and nuclei are treated quantum mechanically within the NEO approach. Although the execution of the proposed NEO/VQE algorithm on state-of-the-art quantum hardware for systems of the size of malonaldehyde (with one quantum hydrogen atom) and orthohydrogen (with two quantum hydrogen atoms) is still premature, the developed machinery allowed us to benchmark our approach using VQE state vector simulations. By means of our applications, we could show the reliability and versatility of the NEOUCC *Ansatz*, which is capable of capturing nuclear–electron wave function energies and subsystem entropies with controlled accuracy. In particular, we demonstrated the possibility of evaluating the proton transfer barrier in malonaldehyde with increasing accuracy by incorporating quantum nuclear effects with the proposed NEOUCC approach. Building on these very promising initial results, we aim in the future to further generalize this framework by including sampling techniques and the possibility of investigating quantum electron–proton dynamics, opening up—thanks to the favorable scaling of the proposed quantum algorithm [$\mathcal{O}(N_b^e + N_b^n)^4$, where N_b^e and N_b^n are the numbers of basis set functions for electrons and nuclei, respectively]—new avenues in the study of nuclear quantum effects in catalysis.

ACKNOWLEDGMENTS

This work was supported by the Hartree National Center for Digital Innovation, a collaboration between STFC and IBM. The authors thank Lars Tornberg for the helpful discussions. This research was supported by funding from Horizon 2020 via NEASQC (Grant No. 951821), the Wallenberg Center for Quantum Technology (WACQT), and from the NCCR MARVEL, a National Center of Competence in Research, funded by the Swiss National Science Foundation (Grant No. 205602). IBM, the IBM logo, and ibm.com are trademarks of International Business Machines Corp., registered in many jurisdictions worldwide. Other product and service names might be trademarks of IBM or other companies. The current list of IBM trademarks is available at https://www.ibm.com/legal/copytrade.

AUTHOR DECLARATIONS

Conflict of Interest

The authors have no conflicts to disclose.

Author Contributions

Arseny Kovyrshin: Conceptualization (equal); Formal analysis (equal); Investigation (equal); Methodology (equal); Validation (equal); Visualization (equal); Writing – original draft (equal); Writing – review & editing (equal). **Mårten Skogh**: Formal analysis (equal); Investigation (equal); Methodology (equal); Visualization (equal); Writing – original draft (equal); Writing – review & editing (equal). **Anders Broo**: Funding acquisition (equal); Project administration (equal); Writing – review & editing (equal). **Stefano Mensa**: Investigation (equal); Supervision (equal); Validation (equal); Visualization (equal); Writing – original draft (supporting); Writing – review & editing (equal). **Emre Sahin**: Data curation (equal); Investigation (equal); Visualization (equal); Writing –

TABLE VIII. The single orbital entropies, so_i, and quantum information (total, electronic, and nuclear) for H_2 (studied in Sec. IV C) evaluated with NEOUCCSDT.

	Orbital	so_i
Electronic	1	0.0793
	2	0.0630
	3	0.0221
	4	0.0397
	QI(electronic)	0.2041
Nuclear	5	0.0035
	6	0.0035
	7	0.0035
	8	0.0035
	QI(nuclear)	(≈6.4%) 0.0139
	QI(total)	0.2179

TABLE IX. The single orbital entropies, so_i, and quantum information (total, electronic, and nuclear) for malonaldehyde ("2-center C_{2v}" setup) evaluated with NEOUCCSDTQ.

	Orbital	so_i
Electronic	1	0.0896
	2	0.1955
	3	0.2049
	4	0.0806
	QI(electronic)	0.5705
Nuclear	5	0.0867
	6	0.0866
	7	0.0000
	8	0.0001
	QI(nuclear)	(≈23%) 0.1734
	QI(total)	0.7440

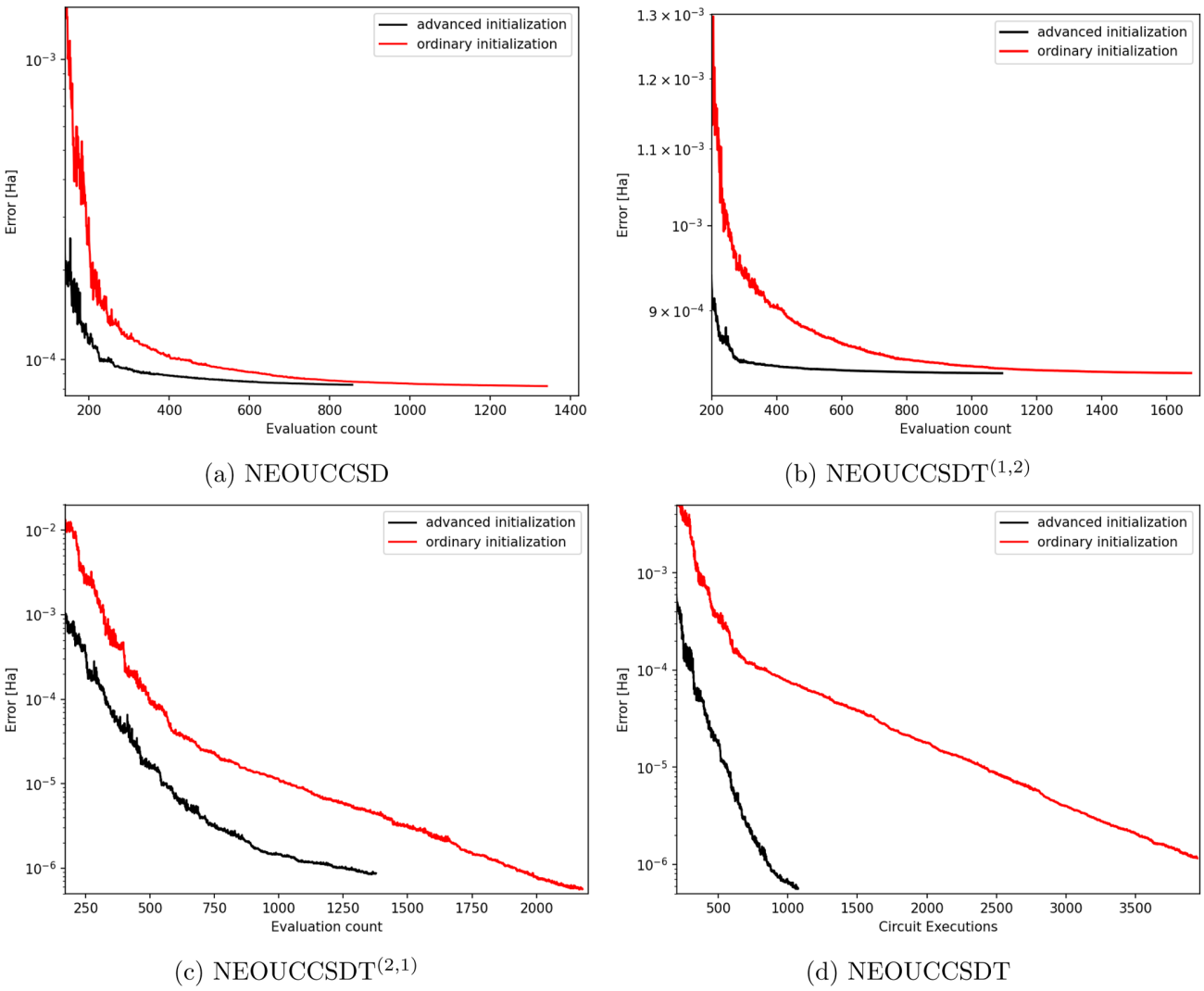


FIG. 7. Energy convergence plots for the H_2 molecule with different NEOUCC *Ansätze* and two initialization schemes. The classical optimization of the circuit parameters is done employing the COBYLA optimizer. The "advanced" initialization scheme introduced in Sec. II C (red lines) clearly outperforms the "ordinary" initialization procedure (black lines) for what concerns the speed of convergence, especially for a large number of parameters (with SDT excitations).

original draft (supporting); Writing – review & editing (equal). **Jason Crain**: Investigation (equal); Supervision (supporting); Writing – original draft (supporting); Writing – review & editing (equal). **Ivano Tavernelli**: Conceptualization (equal); Investigation (equal); Methodology (equal); Supervision (equal); Validation (equal); Writing – original draft (equal); Writing – review & editing (equal).

DATA AVAILABILITY

The data that support the findings of this study are available from the corresponding author upon reasonable request.

APPENDIX A: SINGLE ORBITAL ENTROPY AND QUANTUM INFORMATION

Single orbital entropies[104–106] are calculated as

$$so_i = -\sum_{\alpha=1}^{4} \omega_{i,\alpha} \ln(\omega_{i,\alpha}), \tag{A1}$$

where $\omega_{i,\alpha}$ are the eigenvalues of the one-orbital reduced density matrix,[105,107] the index i runs over molecular orbitals, and α runs over the four possible occupations: $--, -\uparrow, \downarrow-, \downarrow\uparrow$. The sum over

all single orbital entropies, the total quantum information,[105] is given by

$$QI = \sum_i so_i. \tag{A2}$$

The results for H_2 (studied in Sec. IV C) and malonaldehyde in "2-center C_{2v}" setting (studied in Sec. IV D) are shown in Tables VIII and IX.

APPENDIX B: PARAMETER INITIALIZATION AND CONVERGENCE

Figure 7 shows the different convergence profiles for the optimizations using "advanced" and "ordinary" initialization schemes with the NEOUCC *Ansätze* and the COBYLA optimizer. As expected, we notice a faster convergence when the parameters precomputed for the electronic subsystem are transferred to the NEOUCC *Ansatz* (Fig. 7). In addition, the initial energy values for the "advanced" initialization are lower than the ones using the "ordinary" initialization. In addition, our calculations show that as the complexity (and therefore the accuracy) of the *Ansatz* increases, the effect of "advanced" initialization of the parameters becomes more important. Specifically, the convergence rate decreases upon the addition of triple excitations in going from NEOUCCSD [Fig. 7(a)] to NEOUCCSDT$^{(2,1)}$ and NEOUCCSDT$^{(1,2)}$ [Figs. 7(c) and 7(b), respectively]. Even though the number of electronic parameters is the same in both initialization procedures, the benefit of the "advanced" procedure is indisputable, especially for the NEOUCCSDT *Ansatz* where the convergence rate is increased more than threefold.

APPENDIX C: RESOURCE ESTIMATION

Both the number of qubits and the depth of the quantum circuit play a crucial role in determining the performance of quantum algorithms in near-term quantum computers. It is therefore important to analyze the hardware requirements needed to implement our quantum algorithm. As for the quantum gates, the number of entangling gates can strongly impact the circuit noise. In Table X, we summarize the number of qubits, two-qubit entangling gates, and terms in the qubit Hamiltonian needed for the implementation of a VQE calculation for H_2, as reported in Sec. IV C. Specifically, these numbers refer to the final counts obtained after applying three different reduction schemes sequentially: (i) four-qubit tapering, (ii) point group (PG) symmetry, and (iii) projection onto a selected spin state. As wave function *Ansätze*, we employ TwoLocal (8 layers for the electronic part and 14 nuclear–electronic entangling layers) and NEOUCCSD, since they showed similar accuracy (see Sec. IV C). With the proposed schemes, we achieved a significant reduction—compared to the standard procedure—in the number of qubits and Hamiltonian terms, as well as in the number of entangling gates, for both *Ansätze*.

TABLE X. Estimated resource requirements for the H_2 molecule (see Sec. IV C) using various reduction schemes. Shown are the number of qubits, two-qubit gates in TwoLocal as well as in NEOUCCSD *Ansätze*, and the number of Pauli terms ($|\{H\}|$) in the encoded Hamiltonian. All values are calculated using parity encoding. For the TwoLocal *Ansatz*, eight electron-to-electron entangling layers are applied to the electronic subsystem together with 14 electron-to-proton entangling layers.

		No. of two-qubit gates		
Reduction	No. of qubits	TwoLocal	NEOUCCSD	$\|\{H\}\|$
None	16	968	2546	861
+Four-qubit	12	558	2204	828
+PG symmetry	11	465	1472	825
+Spin projection	8	255	1202	428

APPENDIX D: NEO QUBIT HAMILTONIAN FOR H_2 IN MINIMAL BASIS SET

$$\begin{aligned}\hat{H}_{\mathrm{NEO}} = {} & 1.697\,878 \times IIII - 0.373\,036 \times ZIII + 0.356\,877 \times IZII \\ & - 0.373\,036 \times ZZII - 0.337\,003 \times IIZI + 1.105\,641 \times ZIZI \\ & - 0.337\,002 \times IZZI + 1.094\,360 \times ZZZI - 0.416\,401 \times IIIZ \\ & + 0.004\,053 \times ZIIZ + 0.004\,053 \times ZZIZ - 0.004\,053 \times IIZZ \\ & - 0.004\,053 \times IZZZ + 0.416\,401 \times ZZZZ + 0.174\,434 \times XIII \\ & - 0.174\,434 \times XZII - 0.006\,416 \times IXII + 0.006\,416 \times ZXZI \\ & - 0.004\,740 \times YYII - 0.004\,740 \times XXZI + 0.004\,740 \times YYIZ \\ & + 0.004\,740 \times XXZZ - 0.186\,061 \times IIXI + 0.186\,061 \times IZXI \\ & - 0.748\,764 \times XIXI + 0.748\,764 \times XZXI + 0.004\,740 \times ZXXI \\ & + 0.004\,740 \times IYYI - 0.004\,740 \times ZXXZ - 0.004\,740 \times IYYZ \\ & - 0.006\,416 \times XXXI - 0.006\,416 \times YXYI + 0.181\,576 \times IIIX \\ & + 0.174\,434 \times XIIX - 0.174\,434 \times XZIX - 0.006\,416 \times IXIX \\ & + 0.006\,416 \times ZXZX - 0.186\,061 \times IIXX + 0.186\,061 \times IZXX \\ & - 0.006\,416 \times XXXX - 0.006\,416 \times YXYX. \end{aligned} \tag{D1}$$

REFERENCES

[1] T. E. Markland and M. Ceriotti, "Nuclear quantum effects enter the mainstream," Nat. Rev. Chem. **2**(3), 0109 (2018).
[2] W. H. Miller, "Classical S matrix: Numerical application to inelastic collisions," J. Chem. Phys. **53**, 3578 (1970).
[3] I. Tavernelli, U. F. Röhrig, and U. Rothlisberger, "Molecular dynamics in electronically excited states using time-dependent density functional theory," Mol. Phys. **103**(6–8), 963–981 (2005).
[4] I. Tavernelli, "Electronic density response of liquid water using time-dependent density functional theory," Phys. Rev. B **73**(9), 094204 (2006).
[5] S. K. Min, F. Agostini, I. Tavernelli, and E. K. U. Gross, "Ab initio nonadiabatic dynamics with coupled trajectories: A rigorous approach to quantum (de)coherence," J. Phys. Chem. Lett. **8**(13), 3048–3055 (2017).
[6] T. J. Martínez, "Insights for light-driven molecular devices from *ab initio* multiple spawning excited-state dynamics of organic and biological chromophores," Acc. Chem. Res. **39**(2), 119–126 (2006).
[7] H.-D. Meyer, U. Manthe, and L. S. Cederbaum, "The multi-configurational time-dependent Hartree approach," Chem. Phys. Lett. **165**(1), 73–78 (1990).
[8] O. Christiansen, "Møller–Plesset perturbation theory for vibrational wave functions," J. Chem. Phys. **119**(12), 5773–5781 (2003).
[9] L. Pereyaslavets, I. Kurnikov, G. Kamath, O. Butin, A. Illarionov, I. Leontyev, M. Olevanov, M. Levitt, R. D. Kornberg, and B. Fain, "On the importance of accounting for nuclear quantum effects in ab initio calibrated force fields in biological simulations," Proc. Natl. Acad. Sci. U. S. A. **115**(36), 8878–8882 (2018).
[10] D. J. Heyes, M. Sakuma, S. P. de Visser, and N. S. Scrutton, "Nuclear quantum tunneling in the light-activated enzyme protochlorophyllide oxidoreductase," J. Biol. Chem. **284**(6), 3762–3767 (2009).

[11]M. Rossi, P. Gasparotto, and M. Ceriotti, "Anharmonic and quantum fluctuations in molecular crystals: A first-principles study of the stability of paracetamol," Phys. Rev. Lett. **117**(11), 115702 (2016).
[12]A. Vardi-Kilshtain, N. Nitoker, and D. T. Major, "Nuclear quantum effects and kinetic isotope effects in enzyme reactions," Arch. Biochem. Biophys. **582**, 18–27 (2015).
[13]S. Raugei and M. L. Klein, "Nuclear quantum effects and hydrogen bonding in liquids," J. Am. Chem. Soc. **125**(30), 8992–8993 (2003).
[14]K. H. Kim, H. Pathak, A. Späh, F. Perakis, D. Mariedahl, J. A. Sellberg, T. Katayama, Y. Harada, H. Ogasawara, L. G. M. Pettersson, and A. Nilsson, "Temperature-independent nuclear quantum effects on the structure of water," Phys. Rev. Lett. **119**(7), 075502 (2017).
[15]B. Pamuk, J. M. Soler, R. Ramírez, C. P. Herrero, P. W. Stephens, P. B. Allen, and M.-V. Fernández-Serra, "Anomalous nuclear quantum effects in ice," Phys. Rev. Lett. **108**(19), 193003 (2012).
[16]P. J. Ollitrault, A. Baiardi, M. Reiher, and I. Tavernelli, "Hardware efficient quantum algorithms for vibrational structure calculations," Chem. Sci. **11**, 6842–6855 (2020).
[17]S. P. Webb, T. Iordanov, and S. Hammes-Schiffer, "Multiconfigurational nuclear-electronic orbital approach: Incorporation of nuclear quantum effects in electronic structure calculations," J. Chem. Phys. **117**, 4106–4118 (2002).
[18]B. Simmen, E. Mátyus, and M. Reiher, "Elimination of the translational kinetic energy contamination in pre-Born–Oppenheimer calculations," Mol. Phys. **111**(14–15), 2086–2092 (2013).
[19]A. Muolo, A. Baiardi, R. Feldmann, and M. Reiher, "Nuclear-electronic all-particle density matrix renormalization group," J. Chem. Phys. **152**(20), 204103 (2020).
[20]H. Nakai, "Simultaneous determination of nuclear and electronic wave functions without Born–Oppenheimer approximation: Ab initio NO + MO/HF theory," Int. J. Quantum Chem. **86**(6), 511–517 (2002).
[21]S. Lee, J. Lee, H. Zhai, *et al.*, "Evaluating the evidence for exponential quantum advantage in ground-state quantum chemistry," Nat. Commun. **14**, 1952 (2023).
[22]M. Kjaergaard, M. E. Schwartz, J. Braumüller, P. Krantz, J. I.-J. Wang, S. Gustavsson, and W. D. Oliver, "Superconducting qubits: Current state of play," Annu. Rev. Condens. Matter Phys. **11**, 369–395 (2020).
[23]H.-L. Huang, D. Wu, D. Fan, and X. Zhu, "Superconducting quantum computing: A review," Sci. China Inf. Sci. **63**(8), 180501 (2020).
[24]R. Somma, G. Ortiz, J. E. Gubernatis, E. Knill, and R. Laflamme, "Simulating physical phenomena by quantum networks," Phys. Rev. A **65**, 042323 (2002).
[25]D. Wecker, M. B. Hastings, N. Wiebe, B. K. Clark, C. Nayak, and M. Troyer, "Solving strongly correlated electron models on a quantum computer," Phys. Rev. A **92**, 062318 (2015).
[26]B. Bauer, D. Wecker, A. J. Millis, M. B. Hastings, and M. Troyer, "Hybrid quantum-classical approach to correlated materials," Phys. Rev. X **6**, 031045 (2016).
[27]A. Smith, M. S. Kim, F. Pollmann, and J. Knolle, "Simulating quantum many-body dynamics on a current digital quantum computer," npj Quantum Inf. **5**(1), 106 (2019).
[28]C. Cade, L. Mineh, A. Montanaro, and S. Stanisic, "Strategies for solving the Fermi-Hubbard model on near-term quantum computers," Phys. Rev. B **102**, 235122 (2020).
[29]A. Chiesa, F. Tacchino, M. Grossi, P. Santini, I. Tavernelli, D. Gerace, and S. Carretta, "Quantum hardware simulating four-dimensional inelastic neutron scattering," Nat. Phys. **15**, 455–459 (2019).
[30]F. Tacchino, A. Chiesa, S. Carretta, and D. Gerace, "Quantum computers as universal quantum simulators: State-of-the-art and perspectives," Adv. Quantum Technol. **3**(3), 1900052 (2020).
[31]P. Suchsland, P. K. Barkoutsos, I. Tavernelli, M. H. Fischer, and T. Neupert, "Simulating a ring-like Hubbard system with a quantum computer," Phys. Rev. Res. **4**, 013165 (2021); arXiv:2104.06428.
[32]A. Miessen, P. J. Ollitrault, and I. Tavernelli, "Quantum algorithms for quantum dynamics: A performance study on the spin-boson model," Phys. Rev. Res. **3**, 043212 (2021).
[33]F. Libbi, J. Rizzo, F. Tacchino, N. Marzari, and I. Tavernelli, "Effective calculation of the Green's function in the time domain on near-term quantum processors," Phys. Rev. Res. **4**, 043038 (2022).
[34]J. Rizzo, F. Libbi, F. Tacchino, P. J. Ollitrault, N. Marzari, and I. Tavernelli, "One-particle Green's functions from the quantum equation of motion algorithm," Phys. Rev. Res. **4**, 043011 (2022).
[35]E. A. Martinez, C. A. Muschik, P. Schindler, D. Nigg, A. Erhard, M. Heyl, P. Hauke, M. Dalmonte, T. Monz, P. Zoller, and R. Blatt, "Real-time dynamics of lattice gauge theories with a few-qubit quantum computer," Nature **534**(7608), 516–519 (2016).
[36]N. Klco, E. F. Dumitrescu, A. J. McCaskey, T. D. Morris, R. C. Pooser, M. Sanz, E. Solano, P. Lougovski, and M. J. Savage, "Quantum-classical computation of Schwinger model dynamics using quantum computers," Phys. Rev. A **98**(3), 032331 (2018).
[37]A. Roggero and J. Carlson, "Dynamic linear response quantum algorithm," Phys. Rev. C **100**, 034610 (2019).
[38]S. V. Mathis, G. Mazzola, and I. Tavernelli, "Toward scalable simulations of lattice gauge theories on quantum computers," Phys. Rev. D **102**, 094501 (2020).
[39]G. Mazzola, S. V. Mathis, G. Mazzola, and I. Tavernelli, "Gauge-invariant quantum circuits for $U(1)$ and Yang-Mills lattice gauge theories," Phys. Rev. Res. **3**, 043209 (2021).
[40]S. L. Wu, S. Sun, W. Guan, C. Zhou, J. Chan, C. L. Cheng, T. Pham, Y. Qian, A. Z. Wang, R. Zhang, M. Livny, J. Glick, P. K. Barkoutsos, S. Woerner, I. Tavernelli, F. Carminati, A. Di Meglio, A. C. Y. Li, J. Lykken, P. Spentzouris, S. Y.-C. Chen, S. Yoo, and T.-C. Wei, "Application of quantum machine learning using the quantum kernel algorithm on high energy physics analysis at the LHC," Phys. Rev. Res. **3**, 033221 (2021).
[41]G. Clemente, A. Crippa, and K. Jansen, "Strategies for the determination of the running coupling of (2 + 1)-dimensional QED with quantum computing," Phys. Rev. D **106**, 114511 (2022).
[42]W. Guan, G. Perdue, A. Pesah, M. Schuld, K. Terashi, S. Vallecorsa, and J.-R. Vlimant, "Quantum machine learning in high energy physics," Mach. Learn.: Sci. Technol. **2**, 011003 (2021).
[43]J. Schuhmacher, L. Boggia, V. Belis, E. Puljak, M. Grossi, M. Pierini, S. Vallecorsa, F. Tacchino, P. Barkoutsos, and I. Tavernelli, "Unravelling physics beyond the standard model with classical and quantum anomaly detection," arXiv:2301.10787 (2023).
[44]K. Anna Woźniak, V. Belis, E. Puljak, P. Barkoutsos, G. Dissertori, M. Grossi, M. Pierini, F. Reiter, I. Tavernelli, and S. Vallecorsa, "Quantum anomaly detection in the latent space of proton collision events at the LHC," arXiv:2301.10780 (2023).
[45]V. Belis, S. González-Castillo, C. Reissel, S. Vallecorsa, E. F. Combarro, G. Dissertori, and F. Reiter, "Higgs analysis with quantum classifiers," EPJ Web Conf. **251**, 03070 (2021).
[46]P. J. J. O'Malley, R. Babbush, I. D. Kivlichan, J. Romero, J. R. McClean, R. Barends, J. Kelly, P. Roushan, A. Tranter *et al.*, "Scalable quantum simulation of molecular energies," Phys. Rev. X **6**, 031007 (2016).
[47]A. Kandala, A. Mezzacapo, K. Temme, M. Takita, M. Brink, J. M. Chow, and J. M. Gambetta, "Hardware-efficient variational quantum eigensolver for small molecules and quantum magnets," Nature **549**(7671), 242–246 (2017).
[48]P. K. Barkoutsos, J. F. Gonthier, I. Sokolov, N. Moll, G. Salis, A. Fuhrer, M. Ganzhorn, D. J. Egger, M. Troyer, A. Mezzacapo, S. Filipp, and I. Tavernelli, "Quantum algorithms for electronic structure calculations: Particle-hole Hamiltonian and optimized wave-function expansions," Phys. Rev. A **98**(2), 022322 (2018).
[49]Y. Cao, J. Romero, J. P. Olson, M. Degroote, P. D. Johnson, M. Kieferová, I. D. Kivlichan, T. Menke, B. Peropadre, N. P. D. Sawaya *et al.*, "Quantum chemistry in the age of quantum computing," Chem. Rev. **119**(19), 10856–10915 (2019).
[50]S. McArdle, S. Endo, A. Aspuru-Guzik, S. C. Benjamin, and X. Yuan, "Quantum computational chemistry," Rev. Mod. Phys. **92**(1), 015003 (2020).
[51]I. O. Sokolov, P. K. Barkoutsos, P. J. Ollitrault, D. Greenberg, J. Rice, M. Pistoia, and I. Tavernelli, "Quantum orbital-optimized unitary coupled cluster methods in the strongly correlated regime: Can quantum algorithms outperform their classical equivalents?," J. Chem. Phys. **152**(12), 124107 (2020).
[52]P. J. Ollitrault, A. Kandala, C.-F. Chen, P. K. Barkoutsos, A. Mezzacapo, M. Pistoia, S. Sheldon, S. Woerner, J. M. Gambetta, and I. Tavernelli, "Quantum equation of motion for computing molecular excitation energies on a noisy quantum processor," Phys. Rev. Res. **2**(4), 043140 (2020).

IV

[53]I. O. Sokolov, P. K. Barkoutsos, L. Moeller, P. Suchsland, G. Mazzola, and I. Tavernelli, "Microcanonical and finite-temperature *ab initio* molecular dynamics simulations on quantum computers," Phys. Rev. Res. **3**, 013125 (2021).
[54]O. Kiss, F. Tacchino, S. Vallecorsa, and I. Tavernelli, "Quantum neural networks force fields generation," Mach. Learn.: Sci. Technol. **3**(3), 035004 (2022).
[55]P. K. Barkoutsos, F. Gkritsis, P. J. Ollitrault, I. O. Sokolov, S. Woerner, and I. Tavernelli, "Quantum algorithm for alchemical optimization in material design," Chem. Sci. **12**(12), 4345–4352 (2021).
[56]P. J. Ollitrault, A. Miessen, and I. Tavernelli, "Molecular quantum dynamics: A quantum computing perspective," Acc. Chem. Res. **54**(23), 4229–4238 (2021).
[57]M. Rossmannek, P. K. Barkoutsos, P. J. Ollitrault, and I. Tavernelli, "Quantum HF/DFT-embedding algorithms for electronic structure calculations: Scaling up to complex molecular systems," J. Chem. Phys. **154**(11), 114105 (2021).
[58]H. Ma, M. Govoni, and G. Galli, "Quantum simulations of material on near-term quantum computers," Npj Comput. Mater. **6**, 85 (2020).
[59]A. Robert, P. K. Barkoutsos, S. Woerner, and I. Tavernelli, "Resource-efficient quantum algorithm for protein folding," npj Quantum Inf. **7**(1), 38 (2021).
[60]S. Mensa, E. Sahin, F. Tacchino, P. K. Barkoutsos, and I. Tavernelli, "Quantum machine learning framework for virtual screening in drug discovery: A prospective quantum advantage," arXiv:2204.04017 (2022).
[61]A. Baiardi, M. Christandl, and M. Reiher, "Quantum computing for molecular biology," arXiv:2212.12220 (2022).
[62]S. Maniscalco, E.-M. Borrelli, D. Cavalcanti, C. Foti, A. Glos, M. Goldsmith, S. Knecht, K. Korhonen, J. Malmi, A. Nykänen *et al.*, "Quantum network medicine: Rethinking medicine with network science and quantum algorithms," arXiv:2206.12405 (2022).
[63]I. Kassal, S. P. Jordan, P. J. Love, M. Mohseni, and A. Aspuru-Guzik, "Polynomial-time quantum algorithm for the simulation of chemical dynamics," Proc. Natl. Acad. Sci. U. S. A. **105**(48), 18681–18686 (2008).
[64]L. Veis, J. Višňák, H. Nishizawa, H. Nakai, and J. Pittner, "Quantum chemistry beyond Born–Oppenheimer approximation on a quantum computer: A simulated phase estimation study," Int. J. Quantum Chem. **116**, 1328–1336 (2016).
[65]P. J. Ollitrault, G. Mazzola, and I. Tavernelli, "Nonadiabatic molecular quantum dynamics with quantum computers," Phys. Rev. Lett. **125**, 260511 (2020).
[66]F. Pavošević and S. Hammes-Schiffer, "Multicomponent unitary coupled cluster and equation-of-motion for quantum computation," J. Chem. Theory Comput. **17**(6), 3252–3258 (2021).
[67]A. Peruzzo, J. McClean, P. Shadbolt, M.-H. Yung, X.-Q. Zhou, P. J. Love, A. Aspuru-Guzik, and J. L. O'Brien, "A variational eigenvalue solver on a photonic quantum processor," Nat. Commun. **5**, 4213 (2014).
[68]N. Moll, P. Barkoutsos, L. S. Bishop, J. M. Chow, A. Cross, D. J. Egger, S. Filipp, A. Fuhrer, J. M. Gambetta, M. Ganzhorn, A. Kandala, A. Mezzacapo, P. Müller, W. Riess, G. Salis, J. Smolin, I. Tavernelli, and K. Temme, "Quantum optimization using variational algorithms on near-term quantum devices," Quantum Sci.Technol. **3**, 030503 (2018).
[69]A. Choquette, A. Di Paolo, P. K. Barkoutsos, D. Sénéchal, I. Tavernelli, and A. Blais, "Quantum-optimal-control-inspired ansatz for variational quantum algorithms," Phys. Rev. Res **3**, 023092 (2021).
[70]J. T. Seeley, M. J. Richard, and P. J. Love, "The Bravyi-Kitaev transformation for quantum computation of electronic structure," J. Chem. Phys. **137**(22), 224109 (2012).
[71]P. Suchsland, F. Tacchino, M. H. Fischer, T. Neupert, P. Kl. Barkoutsos, and I. Tavernelli, "Algorithmic error mitigation scheme for current quantum processors," Quantum **5**, 492 (2021).
[72]J. Preskill, "Quantum computing in the NISQ era and beyond," Quantum **2**, 79 (2018).
[73]J. R. McClean, S. Boixo, V. N. Smelyanskiy, R. Babbush, and H. Neven, "Barren plateaus in quantum neural network training landscapes," Nat. Commun. **9**(1), 4812 (2018).
[74]Qiskit, Qiskit: An open-source framework for quantum computing (2021); http://www.qiskit.org.
[75]S. Wang, E. Fontana, M. Cerezo, K. Sharma, A. Sone, L. Cincio, and P. J. Coles, "Noise-induced barren plateaus in variational quantum algorithms," Nat. Commun. **12**(1), 6961 (2021).
[76]M. Skogh, O. Leinonen, P. Lolur, and M. Rahm, "Accelerating variational quantum eigensolver convergence using parameter transfer" (unpublished) (2023).
[77]G. M. J. Barca, C. Bertoni, L. Carrington, D. Datta, N. De Silva, J. E. Deustua, D. G. Fedorov, J. R. Gour, A. O. Gunina, E. Guidez, T. Harville, S. Irle, J. Ivanic, K. Kowalski, S. S. Leang, H. Li, W. Li, J. J. Lutz, I. Magoulas, J. Mato, V. Mironov, H. Nakata, B. Q. Pham, P. Piecuch, D. Poole, S. R. Pruitt, A. P. Rendell, L. B. Roskop, K. Ruedenberg, T. Sattasathuchana, M. W. Schmidt, J. Shen, L. Slipchenko, M. Sosonkina, V. Sundriyal, A. Tiwari, J. L. Galvez Vallejo, B. Westheimer, M. Włoch, P. Xu, F. Zahariev, and M. S. Gordon, "Recent developments in the general atomic and molecular electronic structure system," J. Chem. Phys. **152**(15), 154102 (2020).
[78]P. Virtanen, R. Gommers, T. E. Oliphant, M. Haberland, *et al.*, "SciPy 1.0: Fundamental algorithms for scientific computing in Python," Nat. Methods **17**(3), 261–272 (2020).
[79]M. J. D. Powell, "A direct search optimization method that models the objective and constraint functions by linear interpolation," in *Advances in Optimization and Numerical Analysis* (Springer, 1994), pp. 51–67.
[80]M. R. Hestenes and E. Stiefel, "Methods of conjugate gradients for solving linear systems," J. Res. Natl. Bur. Stand. **49**(6), 409 (1952).
[81]J. L. Nazareth, "Conjugate gradient method," Wiley Interdiscip. Rev.: Comput. Stat. **1**(3), 348–353 (2009).
[82]D. Kraft, "Algorithm 733: TOMP–Fortran modules for optimal control calculations," ACM Trans. Math. Software **20**(3), 262–281 (1994).
[83]S. Bravyi, J. M. Gambetta, A. Mezzacapo, and K. Temme, "Tapering off qubits to simulate fermionic Hamiltonians," arXiv:1701.08213 (2017).
[84]R. Ditchfield, W. J. Hehre, and J. A. Pople, "Self-consistent molecular-orbital methods. IX. An extended Gaussian-type basis for molecular-orbital studies of organic molecules," J. Chem. Phys. **54**(2), 724–728 (1971).
[85]K. P. Huber and G. Herzberg, *Molecular Spectra and Molecular Structure. IV. Constants of Diatomic Molecules* (National Institute of Standards and Technology, 1979).
[86]T. Iordanov and S. Hammes-Schiffer, "Vibrational analysis for the nuclear-electronic orbital method," J. Chem. Phys. **118**(21), 9489–9496 (2003).
[87]A. C. Hurley, "The computation of floating functions and their use in force constant calculations," J. Comput. Chem. **9**(1), 75–79 (1988).
[88]M. Tachikawa, K. Taneda, and K. Mori, "Simultaneous optimization of GTF exponents and their centers with fully variational treatment of Hartree–Fock molecular orbital calculation," Int. J. Quantum Chem. **75**(4–5), 497–510 (1999).
[89]H. Nakai, K. Sodeyama, and M. Hoshino, "Non-Born–Oppenheimer theory for simultaneous determination of vibrational and electronic excited states: Ab initio NO + MO/CIS theory," Chem. Phys. Lett. **345**(1–2), 118–124 (2001).
[90]H. Nakai, M. Hoshino, K. Miyamoto, and S. Hyodo, "Elimination of translational and rotational motions in nuclear orbital plus molecular orbital theory," J. Chem. Phys. **122**(16), 164101 (2005).
[91]K. R. Brorsen, "Quantifying multireference character in multicomponent systems with heat-bath configuration interaction," J. Chem. Theory Comput. **16**(4), 2379–2388 (2020).
[92]F. Pavošević, T. Culpitt, and S. Hammes-Schiffer, "Multicomponent quantum chemistry: Integrating electronic and nuclear quantum effects via the nuclear-electronic orbital method," Chem. Rev. **120**(9), 4222–4253 (2020).
[93]M. Tachikawa and Y. Osamura, "Isotope effect of hydrogen and lithium hydride molecules. Application of the dynamic extended molecular orbital method and energy component analysis," Theor. Chem. Acc. **104**(1), 29–39 (2000).
[94]W. Kołos and L. Wolniewicz, "Nonadiabatic theory for diatomic molecules and its application to the hydrogen molecule," Rev. Mod. Phys. **35**, 473–483 (1963).
[95]A. Muolo, E. Mátyus, and M. Reiher, "H_3^+ as a five-body problem described with explicitly correlated Gaussian basis sets," J. Chem. Phys. **151**(15), 154110 (2019).
[96]Y. Wang, B. J. Braams, J. M. Bowman, S. Carter, and D. P. Tew, "Full-dimensional quantum calculations of ground-state tunneling splitting of malonaldehyde using an accurate *ab initio* potential energy surface," J. Chem. Phys. **128**(22), 224314 (2008).
[97]N. H. List, A. L. Dempwolff, A. Dreuw, P. Norman, and T. J. Martínez, "Probing competing relaxation pathways in malonaldehyde with transient X-ray absorption spectroscopy," Chem. Sci. **11**, 4180–4193 (2020).

[98]S. L. Baughcum, R. W. Duerst, W. F. Rowe, Z. Smith, and E. B. Wilson, "Microwave spectroscopic study of malonaldehyde (3-hydroxy-2-propenal). 2. Structure, dipole moment, and tunneling," J. Am. Chem. Soc. **103**(21), 6296–6303 (1981).
[99]E. M. Fluder and J. R. De la Vega, "Intramolecular hydrogen tunneling in malonaldehyde," J. Am. Chem. Soc. **100**(17), 5265–5267 (1977).
[100]C. Møller and M. S. Plesset, "Note on an approximation treatment for many-electron systems," Phys. Rev. **46**(7), 618–622 (1934).
[101]Y. Yang, P. E. Schneider, T. Culpitt, F. Pavošević, and S. Hammes-Schiffer, "Molecular vibrational frequencies within the nuclear-electronic orbital framework," J. Phys. Chem. Lett. **10**(6), 1167–1172 (2019).
[102]Z. B. Charles, M. Farber, C. R. Johnson, and L. Kennedy-Shaffer, "The relation between the diagonal entries and the eigenvalues of a symmetric matrix, based upon the sign pattern of its off-diagonal entries," Linear Algebra Appl. **438**, 1427–1445 (2013).
[103]Q. Yu and S. Hammes-Schiffer, "Nuclear-electronic orbital multistate density functional theory," J. Phys. Chem. Lett. **11**(23), 10106–10113 (2020).
[104]Ö. Legeza, F. Gebhard, and J. Rissler, "Entanglement production by independent quantum channels," Phys. Rev. B **74**, 195112 (2006).
[105]K. Boguslawski, P. Tecmer, Ö. Legeza, and M. Reiher, "Entanglement measures for single- and multireference correlation effects," J. Phys. Chem. Lett. **3**, 3129–3135 (2012).
[106]R. Feldmann, A. Muolo, A. Baiardi, and M. Reiher, "Quantum proton effects from density matrix renormalization group calculations," J. Chem. Theory Comput. **18**(1), 234–250 (2022).
[107]K. Boguslawski and P. Tecmer, "Orbital entanglement in quantum chemistry," Int. J. Quantum Chem. **115**, 1289–1295 (2015).

Paper V

THE JOURNAL OF PHYSICAL CHEMISTRY LETTERS
A JOURNAL OF THE AMERICAN CHEMICAL SOCIETY

pubs.acs.org/JPCL Letter

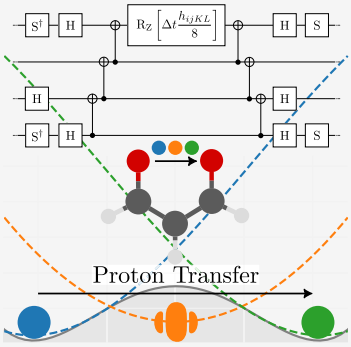
Nonadiabatic Nuclear–Electron Dynamics: A Quantum Computing Approach

Arseny Kovyrshin,* Mårten Skogh, Lars Tornberg, Anders Broo, Stefano Mensa, Emre Sahin, Benjamin C. B. Symons, Jason Crain, and Ivano Tavernelli

Cite This: J. Phys. Chem. Lett. 2023, 14, 7065–7072 Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information

ABSTRACT: Coupled quantum electron–nuclear dynamics is often associated with the Born–Huang expansion of the molecular wave function and the appearance of nonadiabatic effects as a perturbation. On the other hand, native multicomponent representations of electrons and nuclei also exist, which do not rely on any a priori approximation. However, their implementation is hampered by prohibitive scaling. Consequently, quantum computers offer a unique opportunity for extending their use to larger systems. Here, we propose a quantum algorithm for simulating the time-evolution of molecular systems and apply it to proton transfer dynamics in malonaldehyde, described as a rigid scaffold. The proposed quantum algorithm can be easily generalized to include the explicit dynamics of the classically described molecular scaffold. We show how entanglement between electronic and nuclear degrees of freedom can persist over long times if electrons do not follow the nuclear displacement adiabatically. The proposed quantum algorithm may become a valid candidate for the study of such phenomena when sufficiently powerful quantum computers become available.



Most strategies for practical simulation of materials at the molecular scale make two fundamental assumptions: first, that atomic nuclei behave classically and, second, that electronic and nuclear dynamics are adiabatically separable (the Born–Oppenheimer (BO) approximation) such that the system can be described by a product of stationary eigenstates. Both can fail under certain important circumstances. In particular, the BO approximation requires that time-dependent perturbations preserve the instantaneous electronic ground state, a condition satisfied only if the perturbation is sufficiently slow and the energy separation between the ground state and other low-lying levels is sufficiently large. If, however, the time scales of electronic and nuclear degrees of freedom become closer or the energy gap separating potential energy surfaces becomes small (as it does near avoided crossings or conical intersections) nonadiabatic effects arise. Such situations occur when, for example, electron dynamics is driven by ultrafast laser pulses or excited state reaction pathways traverse quasi-degenerate electronic levels. Also, despite the small de Broglie wavelengths of nuclei relative to electrons, nuclear delocalization still occurs over length scales that are important in chemical processes. In these regimes, the classical nuclei approximation also fails.

Consequently, reactions involving light elements can be influenced significantly by quantum effects, e.g., tunneling, whereby the reaction takes a route through a classically forbidden region in configuration space ("through" the energy barrier) in contrast to conventional transition state theory.[1,2] In some cases, nonclassical processes can even supersede traditional kinetics, driving reactions exclusively toward a product for which the classical path would have a higher barrier.[3] In biological systems, quantum phenomena involving light nuclei are known to influence catalysis and enzymatic activity[4,5] with implications, e.g., in the design of novel inhibitors.[6] In particular, the observed rates of certain enzyme-catalyzed reactions cannot be accounted for without introducing corrections for proton tunnelling.[7,8] Also, in DNA chemistry, there is evidence pointing toward proton tunnelling playing a role in the generation of abnormal base pairs (tautomers), which may be implicated in a particular class of mutations.[9]

Significant and fundamental practical challenges limit the extent to which nonclassical features of both electrons and nuclei can be incorporated together into high-fidelity molecular simulations. It is well-known that methods to solve the full Schrödinger equation exactly on a classical computer exhibit prohibitive scaling. In fact, the dimension of the Hilbert space increases exponentially with the system size while the complexity of determining solutions scales factorially with the number of basis functions.[10] This is further

Received: June 9, 2023
Accepted: July 21, 2023
Published: August 1, 2023

https://doi.org/10.1021/acs.jpclett.3c01589
J. Phys. Chem. Lett. 2023, 14, 7065–7072

exacerbated when the descriptions are extended to the treatment of coupled electron−nuclear dynamics.

However, quantum computation has opened new prospects for improving the scaling behavior in this important class of problems.[11−16] In a recent publication[17] we introduced a new quantum computing algorithm, based upon the nuclear-electronic orbital (NEO) framework,[18] for the efficient treatment of quantum electron−nuclear effects on near-term quantum computers. In particular, we demonstrated how the NEO approach can be applied to the study of multicomponent quantum mechanical systems composed of electrons and a selected set of light nuclei (protons) in order to go beyond the Born−Oppenheimer approximation.[19,20] This first approach was, however, fully static, providing only a minimal energy proton transfer path along a predefined reaction coordinate.

In this work, we extend the framework to study the reaction dynamics of chemical systems with important electron−nuclear quantum effects. To this end, we consider malonaldehyde as a model system for proton transfer involving intramolecular hydrogen bonds.[21] The key structural feature is the "O−H···O" hydrogen bond (see Figure 1) for which there

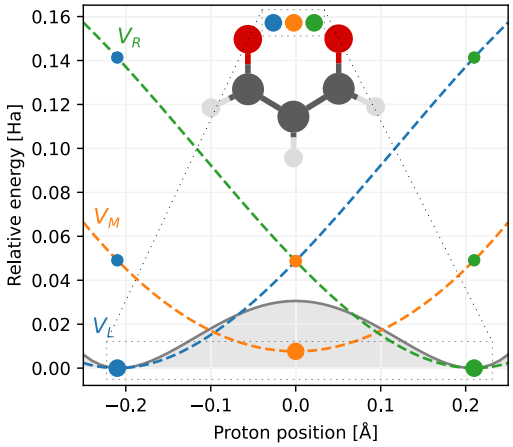


Figure 1. A schematic image of the potentials (V_R, V_M, V_L) associated with the three reference Hamiltonians $\hat{H}_R$, $\hat{H}_M$, $\hat{H}_L$ defined in the main text and used throughout this work. The inset malonaldehyde shows the transition state structure with "Left", "Middle", and "Right" settings for a proton. Apart from the discrete points shown, the potential energy curves are not exact and merely serve to provide the reader with a general guide. The curves were generated by fitting a third-order polynomial to the three points corresponding to each Hamiltonian. The gray shaded area is the scaled-up electronic potential energy calculated with MP2 and aims to illustrate the total potential, and in particular the reaction barrier.

are two possible asymmetric configurations leading to a double-well potential with a proton barrier separating the isomers. This system has previously been investigated using the NEO framework coupled with density functional theory on classical processors.[22] Here, the aim is to demonstrate real-time evolution of the proton transfer process using a quantum computing implementation designed to exploit quantum speed-up. As a first proof-of-principle concept, we consider a rigid molecular scaffold with the electronic configuration optimized for the three protonation states (protons on the right, middle, and left positions). However, the dynamics of the molecular scaffold can easily be introduced by means of ab initio or classical (force field based) forces.

In the NEO approach[18] we make use of Slater determinants comprised of both electronic and protonic spin orbitals. These are expanded with separate nuclear and electronic Gaussian basis sets[18] and optimized using the nuclear electronic orbitals Hartree−Fock (NEOHF) method. Note that by analogy to the α and β electrons in the unrestricted Hartree−Fock (UHF) approach, electrons and protons in NEOHF only interact through the Coulomb potential. We can then construct the corresponding multiparticle basis set for the nuclear-electronic orbital full configuration interaction (NEOFCI) approach,[18] leading to the following second quantization representation of the electron−nuclear Hamiltonian:[18]

$$\begin{aligned}\hat{H} = &\sum_{ij} h_{ij}\hat{a}_i^\dagger\hat{a}_j + \sum_{IJ} h_{IJ}\hat{a}_I^\dagger\hat{a}_J + \frac{1}{2}\sum_{ijkl} h_{ijkl}\hat{a}_i^\dagger\hat{a}_k^\dagger\hat{a}_l\hat{a}_j \\ &+ \frac{1}{2}\sum_{IJKL} h_{IJKL}\hat{a}_I^\dagger\hat{a}_K^\dagger\hat{a}_L\hat{a}_J - \sum_{ijKL} h_{ijKL}\hat{a}_i^\dagger\hat{a}_K^\dagger\hat{a}_L\hat{a}_j \\ &+ \sum_{IJ,A} h_{IJ,A}\hat{a}_I^\dagger\hat{a}_J - \sum_{ij,A} h_{ij,A}\hat{a}_i^\dagger\hat{a}_j + \frac{1}{2}\sum_{AB}\frac{Z_A Z_B}{|\boldsymbol{R}_A - \boldsymbol{R}_B|}\end{aligned} \tag{1}$$

Here h_{pq} and $h_{pq,A}$ are one-particle integrals involving kinetic energy and interactions with classical nuclei, respectively; h_{pqrs} are two-particle integrals responsible for interactions between quantum particles; $\hat{a}^\dagger$ and $\hat{a}$ are the Fermionic creation and annihilation operators; Z is nuclear charge; and last, $\boldsymbol{R}$ is the nuclear coordinate for classical, point-like nuclei (for details see ref 17). In eq 1 upper case indices label protonic spin orbitals, while lower case indices are used for electronic spin orbitals. Indices A and B denote classical nuclei, while I, J, K, and L are reserved for nuclei treated quantum mechanically. Note that creation and annihilation operators must obey anticommutation relations for indistinguishable Fermions (proton−proton or electron−electron) and commutation relations between distinguishable Fermions (proton−electron).

As we aim to study chemical reaction dynamics, the NEO wave function acquires time dependence through the time-dependent configuration interaction (CI) coefficients, $C_{\mu\nu}(t)$. Accordingly, the time-dependent nuclear-electronic wave function is given by

$$|\Psi(t)\rangle = \sum_{\mu\nu} C_{\mu\nu}(t)|\Phi_\mu^e\rangle|\Phi_\nu^n\rangle \tag{2}$$

where $|\Phi_\mu^e\rangle$ and $|\Phi_\nu^n\rangle$ are electronic and nuclear configurations, respectively. It is important to stress that neither the Slater determinants nor the molecular orbitals are allowed to change during the dynamics, but only the CI coefficients $C_{\mu\nu}(t)$ evolve in time according to the time-dependent Schrödinger equation for the molecular Hamiltonian in eq 1. Thus, the evolution of the wave function can be expressed by the equation of motion for CI coefficients

$$\boldsymbol{HC} = i\frac{\partial}{\partial t}\boldsymbol{C} \tag{3}$$

where matrix elements of $\boldsymbol{H}$ are defined through

$$H_{\kappa\lambda,\mu\nu} = \langle\Phi_\kappa^e|\langle\Phi_\lambda^n|\hat{H}|\Phi_\mu^e\rangle|\Phi_\nu^n\rangle \tag{4}$$

On a quantum computer, the initial NEO wave function, $|\Psi(t_0)\rangle$, will be efficiently approximated using the nuclear-

https://doi.org/10.1021/acs.jpclett.3c01589
J. Phys. Chem. Lett. 2023, 14, 7065−7072

Table 1. Energies and von Neumann Entanglement Entropies, eq 5, for Malonaldehyde Obtained with NEOCASCI and NEOUCCSDT Using $\hat{H}_L$, $\hat{H}_M$, and $\hat{H}_R$[a]

	$\hat{H}_L/\hat{H}_R$		$\hat{H}_M$		
Method	Energy (Ha)	Entropy	Energy (Ha)	Entropy	ΔE (Ha)
NEOCASCI[17]	−265.490948	0.0000	−265.485937	0.0044	0.005011
NEOCASCI	−265.491028	0.0020	−265.485912	0.0038	0.005116
NEOUCCSDT[17]	−265.490943	0.0000	−265.485936	0.0044	0.005007
NEOUCCSDT	−265.491023	0.0019	−265.485909	0.0038	0.005115

[a]The barrier height, ΔE, is evaluated as the difference in energy between the L and M setups. We observe good agreement with the barrier height reported in ref 17 for NEOUCCSDT.

electronic orbitals unitary coupled cluster singles, doubles, and triples (NEOUCCSDT) ansatz.[17]

As a model for the demonstration of our quantum algorithm for electron−proton quantum dynamics, we consider the proton transfer process in malonaldehyde using the setup discussed in our previous publication.[17] As shown in Figure 1, the proton can be localized at two possible asymmetric equilibrium positions separated from each other by 0.42 Å (marked with blue and green spheres) as well as at the peak of the barrier in the middle (orange sphere). The two asymmetric structures are characterized by a different relaxation of the electronic orbitals and are associated with the formation of two different OH bonds: one with the oxygen on the right (green sphere) and the other with the oxygen on the left (blue sphere) in the inset of Figure 1. Finally, the symmetric setting has the proton shared equally between the two oxygen atoms. These setups define three different Hamiltonians: $\hat{H}_R^0$ (R for "Right"), $\hat{H}_L^0$ (L for "Left"), and $\hat{H}_M^0$ (M for "Middle"), respectively. The nuclear and electronic orbitals were optimized with NEOHF calculations using the transition state molecular scaffold of malonaldehyde[17] and imposing C_s symmetry. For protons, a split-valence double-ζ nuclear basis set composed of 2 uncontracted Cartesian S functions (DZSNB) was used. Electronic orbitals were represented using a split-valence double-ζ Gaussian basis set in the 6-31 contraction scheme (6-31G). In the asymmetric cases described by $\hat{H}_R^0$ and $\hat{H}_L^0$, the NEOHF calculations were performed with two nuclear basis functions at the equilibrium position (green and blue spheres respectively). Whereas, in the symmetric case, the two nuclear basis functions were located at the barrier's maximum (orange sphere). In all cases, only the transferred proton was treated quantum mechanically (described by two nuclear orbitals placed at one of the locations), while all other nuclei were considered as classical point charges at fixed positions. We also kept the same set of electronic basis functions in all three cases, putting electronic basis functions at all proton locations (R, L, and M).

For each of the three sets of electronic NEOHF orbitals (associated with $\hat{H}_R^0$, $\hat{H}_M^0$, and $\hat{H}_L^0$) we performed nuclear-electronic orbital complete active space configuration interaction (NEOCASCI) calculations to obtain the corresponding reference energies (see Table 1). To keep the electronic active space compact, in all calculations the 17 low-lying core orbitals for electrons were considered fully occupied and only 4 orbitals hosting 4 electrons were considered active. The nuclear active space consisted of the lowest-energy nuclear NEOHF orbitals from the $\hat{H}_R^0$, $\hat{H}_M^0$, and $\hat{H}_L^0$ setups. Thus, the nuclear active space for all three NEOCASCI calculations remains the same and consists of 3 orbitals. The corresponding Hamiltonians will be denoted as $\hat{H}_R$, $\hat{H}_M$, and $\hat{H}_L$. Electronic and nuclear NEOHF orbitals used in the active space for all three settings are shown in Figure 2. Based on the NEOCASCI calculations at these three setups, the barrier is estimated to be 0.0051 hartree. This is rather close to the value obtained in ref 17, 0.0050 hartree.

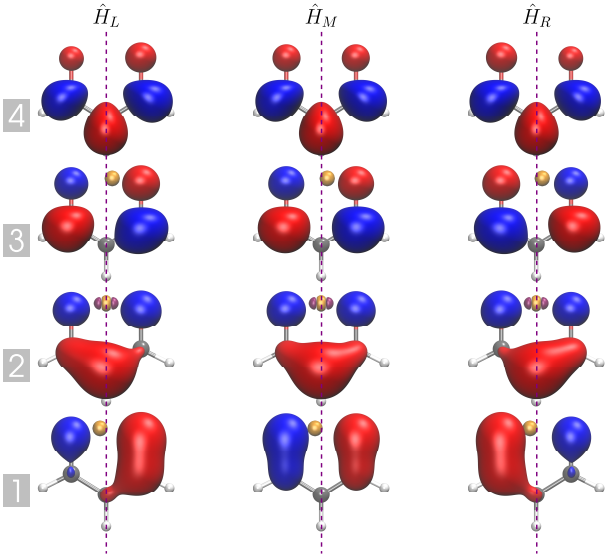


Figure 2. Electronic (red/blue) and nuclear (orange/purple) orbitals (isosurface value is 0.05) included in the active space of NEOCASCI calculation for malonaldehyde in the "Left" ($\hat{H}_L$), "Middle" ($\hat{H}_M$), and "Right" ($\hat{H}_R$) setups. The orbitals were prepared with a NEOHF calculation using the DZSNB and 6-31G basis sets for the nuclei and the electrons, respectively.

To enable calculations on a quantum computer, we mapped the second-quantized Hamiltonians ($\hat{H}_R$, $\hat{H}_M$, and $\hat{H}_L$) spanning 14 spin orbitals to corresponding 8-qubit Hamiltonians, with 6 qubits spanning the electronic subspace and 2 qubits spanning the nuclear one. The parity Fermion-to-qubit transformation for electronic and nuclear operators was used together with the qubit reduction techniques as described in ref 17. The energy values for all systems estimated with the Variational Quantum Eigensolver (VQE) and the NEO-UCCSDT analogue are presented in Table 1. The NEO-UCCSDT results are in agreement with the NEOCASCI references for all settings, and the value for the energy barrier agrees with the reference value to within 1×10^{-6} Hartree, which confirms the excellent quality of the present setup for studying proton dynamics. We also report the von Neumann entropy, s, which captures the entanglement between the nuclear and electronic subsystems. This is obtained according to

$$s = -\mathrm{Tr}(\hat{\rho}_e \ln \hat{\rho}_e) = -\mathrm{Tr}(\hat{\rho}_n \ln \hat{\rho}_n) \quad (5)$$

https://doi.org/10.1021/acs.jpclett.3c01589
J. Phys. Chem. Lett. 2023, 14, 7065−7072

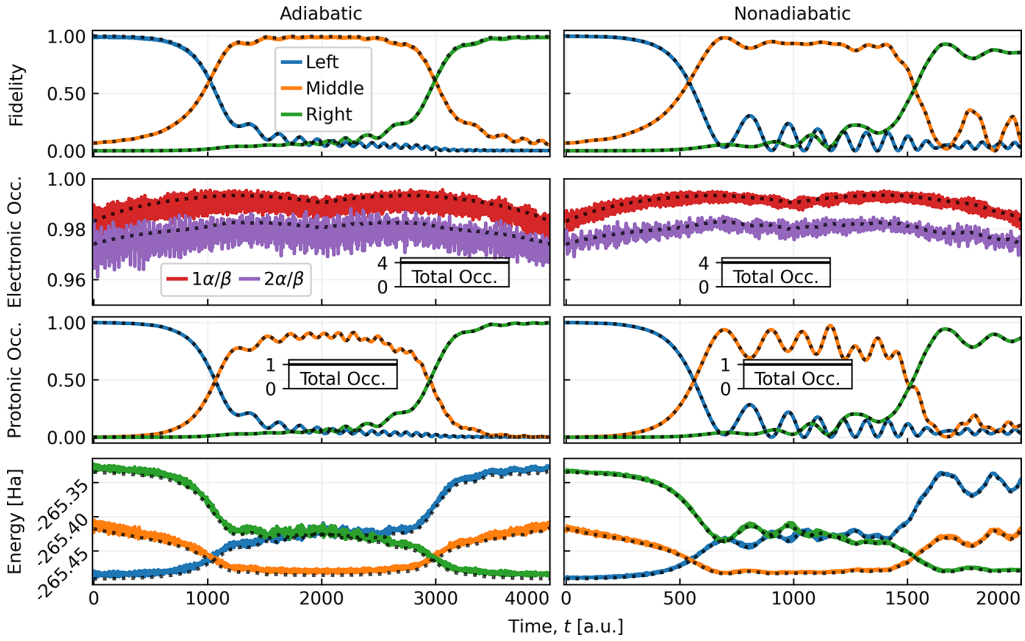


Figure 3. (left panel) Adiabatic evolution performed for time t_f = 4000 au with time step Δt = 1.0 au. (right panel) Dynamics performed in nonadiabatic regime for time t_f = 2000 au with Δt = 0.5 au. The black dotted lines show the reference results, generated by RK4. Comparing our results to the reference, it is evident there is only a minor loss of accuracy.

where $\hat{\rho}_e = \mathrm{Tr}_n[\hat{\rho}_{e,n}]$ and $\hat{\rho}_n = \mathrm{Tr}_e[\hat{\rho}_{e,n}]$ are the reduced density matrices corresponding to electrons and nuclei respectively, and $\hat{\rho}_{e,n}$ is the full density matrix. In contrast to our previous study,[17] the proton−electron entanglement in the $\hat{H}_L$ and $\hat{H}_R$ ground states is no longer zero but amounts to 0.0020 and 0.0019 for NEOUCCSDT and NEOCASCI, respectively. This is mainly due to the extension of nuclear active space for $\hat{H}_L$ and $\hat{H}_R$ (compared to ref 17), which includes orbitals at the top of, as well as on both sides of the barrier separating the two minima. Our calculations still show increasing electron−nuclear entanglement as the proton approaches the top of the barrier.

After preparing the ground state wave function for a given nuclear configuration, we can propagate it from the initial time t_0 to the final time t_f by means of the time-evolution operator $\mathcal{U}(t_f, t_0)$

$$|\Psi(t_f)\rangle = \mathcal{U}(t_f, t_0)|\Psi(t_0)\rangle \tag{6}$$

The form of $\mathcal{U}(t_f, t_0)$ depends on the Hamiltonian describing the system. For a general closed system, described by a time-dependent Hamiltonian, the exact time evolution is given by the Dyson series

$$\mathcal{U}(t_f, t_0) = \mathcal{T}\exp\left[-i\int_{t_0}^{t_f} \hat{H}(t)\mathrm{d}t\right] \tag{7}$$

where $\mathcal{T}$ is the time-ordering operator. Expressing the time-dependent Hamiltonian, $\hat{H}(t)$, as a sum of weighted Pauli strings, P_k, with time-dependent weights $h_k(t)$

$$\hat{H}(t) = \sum_{k=1}^{K} h_k(t)P_k \tag{8}$$

we can approximate time-ordered Dyson series with the first-order decomposition formula given by Suzuki.[23] Specifically, breaking the evolution into a series of N discrete steps of size $\Delta t = t_f/N$, eq 7 can be approximated as follows:

$$\mathcal{U}(t_f, t_0) \approx \prod_{t_j=t_0}^{t_{N-1}} U(t_j + \Delta t, t_j) = \prod_{t_j=t_0}^{t_{N-1}} \prod_{k=1}^{K} \mathrm{e}^{-ih_k(t_j+\Delta t/2)P_k\Delta t} \tag{9}$$

The corresponding quantum circuit is shown schematically in Figure S1.

It should be noted that, with such an approach, the number of time steps required to achieve a desired time evolution typically results in circuits that are too deep to be implemented on noisy, near-term hardware.[24] On the other hand, other noise-resilient time-propagation algorithms such as Variational Real-Time Evolution (VRTE)[25,26] are also available.

In this work, we consider time evolution to be driven by a linear combination of the three Hamiltonians described above: $\hat{H}_L$, $\hat{H}_M$, and $\hat{H}_R$. These Hamiltonians share the same pure nuclear contributions, specifically, the second, sixth, and eighth terms in eq 1 (the fourth term is absent in our setup as only one proton is considered). The pure electronic and mixed electron−nuclear terms vary due to the different electronic active spaces employed (see Figure 2). Due to the large potential energy barrier separating the minima of $\hat{H}_L$ and $\hat{H}_R$, we cannot expect a spontaneous transition of the shared proton from one minimum to the other. To capture such dynamics, we opted for a parametrized time-dependent Hamiltonian, made of a normalized linear combination of the three original Hamiltonians ($\hat{H}_L$, $\hat{H}_M$, and $\hat{H}_R$)

$$\hat{H}(t) = \alpha(t)\hat{H}_L + \beta(t)\hat{H}_M + \gamma(t)\hat{H}_R \tag{10}$$

where normalization requires that $\alpha(t) + \beta(t) + \gamma(t) = 1$. In the applications, we will use the following parametrization

https://doi.org/10.1021/acs.jpclett.3c01589
J. Phys. Chem. Lett. 2023, 14, 7065−7072

$$\hat{H}(t) = \begin{cases} \left(1 - \frac{2t}{t_f}\right)\hat{H}_L + \frac{2t}{t_f}\hat{H}_M & \text{for } t < \frac{t_f}{2} \\ \left(1 - \frac{2t - t_f}{t_f}\right)\hat{H}_M + \frac{2t - t_f}{t_f}\hat{H}_R & \text{for } t \geq \frac{t_f}{2} \end{cases} \tag{11}$$

After preparing the ground state of the Hamiltonian $\hat{H}_L$ ($\alpha(t_0) = 1$, $\beta(t_0) = \gamma(t_0) = 0$), $|\Psi_L\rangle$, we will "adiabatically" drive the system toward the ground state of $\hat{H}_R$ (namely, $|\Psi_R\rangle$) by updating the Hamiltonian's parameters in eq 10 following eq 11 (profiles in Figure S2) and evolving the state of the system using eq 9. Given a final time, in this case t_f = 4000 au, this protocol guarantees a smooth transition in the Hamiltonian space from $\hat{H}_L$ at $t = 0$ to $\hat{H}_R$ at $t = t_f$ passing through $\hat{H}_M$. Note that the time-dependent Hamiltonian only accounts for the dynamics of the degrees of freedom not explicitly included in the quantum description of the electronic and protonic wave functions implemented in the quantum circuit. Despite this approximation, the setup can account for the main quantum correlation effects between the two subsystems.

The progress of the system dynamics in the time interval $[t_0 = 0, t_f = 4000]$ is monitored by measuring the expectation values of the energy and of the proton and electron occupation numbers. The energies associated with the three Hamiltonians

$$E_X = \langle\Psi(t)|\hat{H}_X|\Psi(t)\rangle, \quad X \in \{L, M, R\} \tag{12}$$

are used to examine the time-dependent expectation value $E(t)$ associated with the Hamiltonian in eq 10, which is shown in Figure 3 (lower panel). In addition, we also monitor the progress of the electron transfer dynamics by measuring the expectation values of the electron and the proton occupation numbers in their corresponding spin−orbitals using the operators $\hat{N}_i = \hat{a}_i^\dagger \hat{a}_i$ and $\hat{N}_I = \hat{a}_I^\dagger \hat{a}_I$, respectively. The occupation numbers are given by

$$n_{I/i} = \langle\Psi(t)|\hat{N}_{I/i}|\Psi(t)\rangle \tag{13}$$

where the index i labels the electronic spin−orbitals shown in Figure 2 (red/blue), while I labels the nuclear spin-orbitals (orange/purple). Of particular interest are the occupations of the three different nuclear spin-orbitals n_L, n_M, n_R (i.e., eq 13 with $I = L, M, R$), which reach their highest values for the ground state of the corresponding Hamiltonians, $\hat{H}_L$, $\hat{H}_M$, $\hat{H}_R$. Their time evolution is shown in the second panel from the bottom of Figure 3. Similarly, we also monitored the occupations of the first two α and β electronic spin orbitals (see the second panel from the top of Figure 3).

Of particular interest is also the measure of the state fidelity, $\mathcal{F}$, which is computed as the squared absolute overlap between the time-dependent state and the ground states of the reference Hamiltonians Ψ_L, Ψ_M, and Ψ_R

$$\mathcal{F}_X(|\Psi(t)\rangle, |\Psi_X\rangle) = |\langle\Psi(t)|\Psi_X\rangle|^2, \quad X \in \{L, M, R\} \tag{14}$$

$\mathcal{F}_X$ takes a value in the range $[0, 1]$, where 1 corresponds to a maximum absolute overlap and therefore to identical (up to a global phase) states. In particular, we will use the fidelities $\mathcal{F}_M$ and $\mathcal{F}_R$ to assess the adiabaticity of our time-evolution protocol.

In the following, we present two time evolution settings that we named for convenience "slow" (or adiabatic) and "fast" (or nonadiabatic). In the first case, the entire Hamiltonian evolution (eq 10) is completed in 4000 time steps with step size Δt = 1.0 au, which (as we will see below) is sufficient to closely approximate adiabtic evolution of the proton−electron state (see Figure 3). In the second case, we double the speed at which we perform the dynamics, acting in a nonadiabatic, "fast", regime (see right panels of Figure 3). This is realized by halving the step size, Δt = 0.5 au, and keeping the same total number of steps, 4000. For all different time-evolution settings, we also performed classical reference calculations using the fourth-order Runge−Kutta (RK4)[27] integrator (see Figure 3). In the "slow", adiabatic case, the change over time of parameters α, β, and γ results in a one-way transfer of the proton from the oxygen in the left position (L) to the one on the right (R) (see Figure 1). The transfer is clearly evidenced by the evolution of the nuclear occupation numbers in Figure 3 (second panel from the bottom) and the fidelity plot in Figure 3 (top panel). The occupation numbers of the nuclear spin-orbitals evolve from a set of values (n_L = 1, n_M = 0, n_R = 0) at t = 0 au, transitioning to an intermediate regime dominated by $n_M \sim 1$ around time 2000 au, and finally converging toward the values (n_L = 0, n_M = 0, n_R = 1) at the end of the simulation. The time evolution of the system energy (see the bottom left panel in Figure 3) inversely mirrors the changes just described for the occupation numbers. At the beginning of the simulation, the expectation values for $\hat{H}_M$ (E_M) and $\hat{H}_R$ (E_R) greatly exceed the corresponding ground state energies as the starting NEOUCCSDT state is optimized for the $\hat{H}_L$ ground state (see Figure 3). Then, as the simulation progresses toward the middle position (dominated by $\hat{H}_M$), E_L slowly grows while E_M and E_R decrease. At the halfway mark ($t \sim$ 2000 au), the E_L and E_R values are nearly equal, while E_M closely approaches the ground state energy value of $\hat{H}_M$. This means that our protocol is capable of driving the system adiabatically from the initial state (with the proton localized to the left) to the middle configuration with the equally shared proton. This is also supported by the values of occupations in the middle panel of Figure 3. A similar description, in reverse order, applies to the second half of the dynamics leading to a final state with a large overlap with the ground state of the Hamiltonian $\hat{H}_R$. Concerning the electronic structure, we also observe changes in the occupation of the spin orbitals involved in the calculations (the second panel from the top of Figure 3). Specifically, the occupation numbers of the first two lower-lying spin orbitals increase slightly during the proton transfer process, reaching a maximum at the transition barrier of the L-M and M-R half-reactions. Note that the occupations refer to both α and β spin orbitals because we are in the restricted picture. It is worth stressing that for both protons and electrons, the total number of particles (integrated over all basis functions) remains conserved throughout the full reaction path and amounts to 1 and 4, respectively. Constant particle numbers confirm the absence of particle leakage, which is a signature of faithful implementation of the unitary time-evolution protocol. The high-frequency oscillations observed in the nuclear occupancy numbers correspond to the nonadiabatic part of the coherent time evolution, i.e., slight deviations from the instantaneous ground state of $\hat{H}(t)$. These can be suppressed if we further decrease the rate of change of the Hamiltonian coefficients α, β, and γ, thus approaching the adiabatic limit. This can be seen in Figure S3, where the adiabatic speed (change in α, β, and γ per Δt = 1.0 au) was five times slower (t_f = 20000 au).

V

https://doi.org/10.1021/acs.jpclett.3c01589
J. Phys. Chem. Lett. 2023, 14, 7065−7072

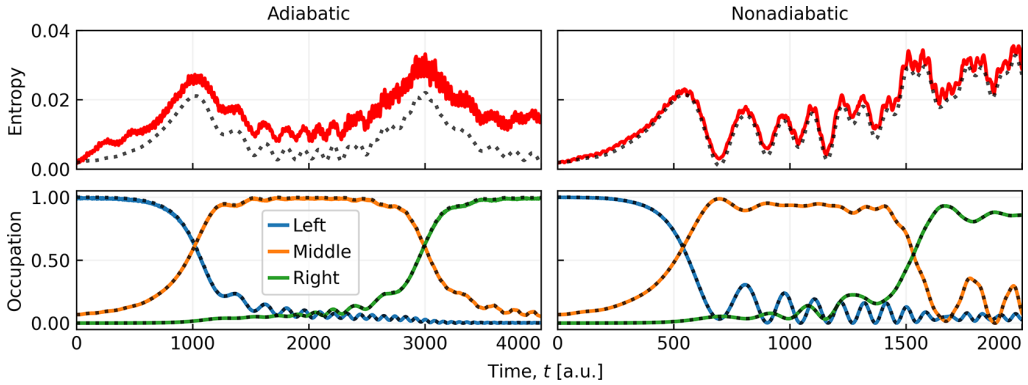


Figure 4. Nuclear−electron entanglement entropy (top) and the fidelities with respect to reference ground states in the "Left", "Middle", and "Right" setups (bottom) as a function of time for the evolution in the (left panel) adiabatic regime for t_f = 4000 au with Δt = 1.0 au; (right panel) nonadiabatic regime for time t_f = 2000 au with Δt = 0.5 au. The reference values, generated using RK4, are shown as black dotted lines. While the fidelities are in good agreement with reference values, the entanglement entropy (it is orders of magnitude smaller than the fidelity) acquires a significant error for the dynamics with Δt = 1.0 au. This error decreases as the resolution of the simulation improves to Δt = 0.5 au.

Of particular interest is the analysis of the entanglement entropy evaluated for the two subsystems, eq 5, along the reaction path. For details, see ref 28. In the adiabatic regime, the exact entropy profile shows a bimodal, roughly symmetric, shape with the two peaks at the L-M and M-R transition points. As before, the oscillations are caused by nonadiabatic effects and disappear if we further slow down the dynamics (see Figure S4). Note that while the exact profile (dotted line in top left panel of Figure 4) fully recovers the initial value, in the first-order Suzuki[23] approach a residual entanglement entropy remains after full transfer of the proton (red curve, Figure 4). This deviation is attributed to the first-order Suzuki approximation error and can be reduced by decreasing the step size to Δt = 0.5 au, as shown in Figure 4 (top right panel). The fidelities with respect to the ground state wave functions corresponding to the different components of the Hamiltonian (eq 10) are reported in Figure 3 (top panel) and confirm the quality of the simulations (see also the fully adiabatic profiles in Figure S3). The loss of accuracy in short (t_f = 4000, see left panels of Figures 3 and 4) and long (t_f = 20000, see Figures S3 and S4) adiabatic regimes are comparable, proving that the evolution operator is stable over a long-time simulation and does not accumulate significant errors. Finally, it is worth noting that entropy sharply rises at the moments when the evolving state has equal fidelity first for $\mathcal{F}_L$ and $\mathcal{F}_M$ and then later for $\mathcal{F}_M$ and $\mathcal{F}_R$. At these points in the evolution, there is thus a maximum amount of entanglement between the protonic and electronic degrees of freedom along the adiabatic trajectory.

In the "fast" regime, the physics is quite different. During the entire proton transfer process, the entropy of the system keeps growing (with large fluctuations) and reaches its maximum value at the end of the simulation. In contrast to the slow regime, we observe an accumulation of entanglement due to excitations to higher energy levels along the nonadiabatic trajectory. Both the exact (dashed line in Figure 4) and the first-order Suzuki[23] dynamics confirm this behavior. It is worth mentioning that a classical treatment of the transferred proton will miss important quantum effects, including delocalization and tunneling, as well as the possibility to describe the entanglement with the electronic state, which is the main subject of this study. The entanglement will persist unless a decoherence channel is introduced. To monitor this process, one would need to include bath degrees of freedom (i.e., the dynamics of the molecular scaffold and of the solvent) and simulate the dynamics as an open quantum system. The dynamic motion of the molecular scaffold can be added to the driving Hamiltonian via one-particle terms responsible for interaction with classical nuclei in eq 1 (the sixth and seventh terms). The extra degrees of freedom can then be propagated classically using ab initio forces or a classical force field.

In this Letter, we presented a quantum computing algorithm for modeling electron−nuclear coupled dynamics in molecular systems formulated in the second quantization framework. In classical computation setups, nuclear and electronic degrees of freedom exponentially increase with the system size, leading to unfavorable scaling in both memory and execution time. On the other hand, quantum computers have the potential to solve these same problems in polynomial time and using polynomial memory. The quantum algorithm was tested with classical simulations, leading to very accurate results and interesting insights about the proton transfer process in a realistic model of malonaldehyde. In particular, we showed that when the dynamics is sufficiently fast, such that the electrons cannot follow adiabatically, entanglement between electronic and nuclear degrees of freedom is generated and persists over time. The classical emulation of the quantum algorithm will quickly become unfeasible, as the number of electronic and nuclear basis functions (and therefore qubits) exceeds about 20−30 qubits. From this point on, only quantum computers will be able to perform the coupled electron−nuclear quantum dynamics on such large scales. However, the noise in state-of-the-art quantum computers currently hampers demonstrations of long-time adiabatic quantum simulations of complex systems with more than 10−20 qubits due to the required circuit depth. Further developments are therefore needed for a more efficient implementation of the proposed electron−nuclear quantum dynamics scheme before reaching the fault-tolerant regime. In particular, we are planning to investigate more efficient encoding schemes for the time-evolution operator in conjunction with error mitigation schemes,[29] as well as the possibility of applying variational time-evolution algorithms.

https://doi.org/10.1021/acs.jpclett.3c01589
J. Phys. Chem. Lett. 2023, 14, 7065−7072

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acs.jpclett.3c01589.

Discussion of the algorithm's computational scaling; illustration of the quantum circuit for dynamics; diagram of the time-evolution of the Hamiltonian components; figure of the adiabatic time evolution of the coupled electron−nuclear dynamics; figure of the entropy and state fidelity computed for the adiabatic coupled electron−nuclear dynamics (PDF)

■ AUTHOR INFORMATION

Corresponding Author

Arseny Kovyrshin − *Data Science and Modelling, Pharmaceutical Sciences, R&D, AstraZeneca Gothenburg, Molndal SE-431 83, Sweden;* orcid.org/0009-0009-1136-4428; Email: arseny.kovyrshin@astrazeneca.com

Authors

Mårten Skogh − *Data Science and Modelling, Pharmaceutical Sciences, R&D, AstraZeneca Gothenburg, Molndal SE-431 83, Sweden; Department of Chemistry and Chemical Engineering, Chalmers University of Technology, 412 96 Gothenburg, Sweden*

Lars Tornberg − *Data Science and Modelling, Pharmaceutical Sciences, R&D, AstraZeneca Gothenburg, Molndal SE-431 83, Sweden*

Anders Broo − *Data Science and Modelling, Pharmaceutical Sciences, R&D, AstraZeneca Gothenburg, Molndal SE-431 83, Sweden*

Stefano Mensa − *The Hartree Centre, STFC, Sci-Tech Daresbury, Warrington WA4 4AD, United Kingdom*

Emre Sahin − *The Hartree Centre, STFC, Sci-Tech Daresbury, Warrington WA4 4AD, United Kingdom*

Benjamin C. B. Symons − *The Hartree Centre, STFC, Sci-Tech Daresbury, Warrington WA4 4AD, United Kingdom*

Jason Crain − *IBM Research Europe, Hartree Centre STFC Laboratory, Sci-Tech Daresbury, Warrington WA4 4AD, United Kingdom; Department of Biochemistry, University of Oxford, Oxford OX1 3QU, U.K.;* orcid.org/0000-0001-8672-9158

Ivano Tavernelli − *IBM Quantum, IBM Research Europe−Zurich, 8803 Rüschlikon, Switzerland;* orcid.org/0000-0001-5690-1981

Complete contact information is available at: https://pubs.acs.org/10.1021/acs.jpclett.3c01589

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

This work was supported by the Hartree National Centre for Digital Innovation, a collaboration between STFC and IBM. This research was supported by funding from Horizon 2020 via the NEASQC project (Grant Number 951821); the Wallenberg Center for Quantum Technology (WACQT); and the NCCR MARVEL, a National Centre of Competence in Research, funded by the Swiss National Science Foundation (Grant Number 205602). IBM, the IBM logo, and ibm.com are trademarks of International Business Machines Corp., registered in many jurisdictions worldwide. Other product and service names might be trademarks of IBM or other companies. The current list of IBM trademarks is available at https://www.ibm.com/legal/copytrade.

■ REFERENCES

(1) Wild, R.; Wester, R. Rate of Quantum-tunnelling Reaction Revealed. *Measurement* **2023**, *10*, 3.

(2) Schreiner, P. R.; Reisenauer, H. P.; Pickard, F. C., IV; Simmonett, A. C.; Allen, W. D.; Mátyus, E.; Császár, A. G. Capture of Hydroxymethylene and Its Fast Disappearance Through Tunnelling. *Nature* **2008**, *453*, 906−909.

(3) Schreiner, P. R.; Reisenauer, H. P.; Ley, D.; Gerbig, D.; Wu, C.-H.; Allen, W. D. Methylhydroxycarbene: Tunneling Control of a Chemical Reaction. *Science* **2011**, *332*, 1300−1303.

(4) Klinman, J. P.; Kohen, A. Hydrogen Tunneling Links Protein Dynamics to Enzyme Catalysis. *Annual review of biochemistry* **2013**, *82*, 471−496.

(5) Cha, Y.; Murray, C. J.; Klinman, J. P. Hydrogen Tunneling in Enzyme Reactions. *Science* **1989**, *243*, 1325−1330.

(6) Nagel, Z. D.; Klinman, J. P. A 21st Century Revisionist's View at a Turning Point in Enzymology. *Nat. Chem. Biol.* **2009**, *5*, 543−550.

(7) Bañacký, P. Dynamics of Proton Transfer and Enzymatic Activity. *Biophys. Chem.* **1981**, *13*, 39−47.

(8) Basran, J.; Sutcliffe, M. J.; Scrutton, N. S. Enzymatic H-transfer Requires Vibration-driven Extreme Tunneling. *Biochemistry* **1999**, *38*, 3218−3222.

(9) Slocombe, L.; Sacchi, M.; Al-Khalili, J. An Open Quantum Systems Approach to Proton Tunnelling in DNA. *Communications Physics* **2022**, *5*, 109.

(10) Verma, P.; Huntington, L.; Coons, M. P.; Kawashima, Y.; Yamazaki, T.; Zaribafiyan, A. Scaling Up Electronic Structure Calculations on Quantum Computers: The Frozen Natural Orbital Based Method of Increments. *J. Chem. Phys.* **2021**, *155*, 034110.

(11) Aspuru-Guzik, A.; Dutoi, A. D.; Love, P. J.; Head-Gordon, M. Simulated Quantum Computation of Molecular Energies. *Science* **2005**, *309*, 1704−1707.

(12) Kassal, I.; Jordan, S. P.; Love, P. J.; Mohseni, M.; Aspuru-Guzik, A. Polynomial-time Quantum Algorithm for The Simulation of Chemical Dynamics. *Proc. Natl. Acad. Sci. U. S. A.* **2008**, *105*, 18681−18686.

(13) Kassal, I.; Whitfield, J. D.; Perdomo-Ortiz, A.; Yung, M.-H.; Aspuru-Guzik, A. Simulating Chemistry Using Quantum Computers. *Annu. Rev. Phys. Chem.* **2011**, *62*, 185−207.

(14) Cao, Y.; Romero, J.; Olson, J. P.; Degroote, M.; Johnson, P. D.; Kieferová, M.; Kivlichan, I. D.; Menke, T.; Peropadre, B.; Sawaya, N. P.; et al. Quantum Chemistry in The Age of Quantum Computing. *Chem. Rev.* **2019**, *119*, 10856−10915.

(15) Ollitrault, P. J.; Jandura, S.; Miessen, A.; Burghardt, I.; Martinazzo, R.; Tacchino, F.; Tavernelli, I. Quantum Algorithms for Grid-based Variational Time Evolution. *arXiv* **2022**, 2203.02521.

(16) Miessen, A.; Ollitrault, P. J.; Tacchino, F.; Tavernelli, I. Quantum Algorithms for Quantum Dynamics. *Nature Computational Science* **2023**, *3*, 25−37.

(17) Kovyrshin, A.; Skogh, M.; Broo, A.; Mensa, S.; Sahin, E.; Crain, J.; Tavernelli, I. A quantum computing implementation of nuclearelectronic orbital (NEO) theory: Toward an exact pre-Born-Oppenheimer formulation of molecular quantum systems. *J. Chem. Phys.* **2023**, *158*, 214119.

(18) Webb, S. P.; Hammes-Schiffer, S.; Iordanov, T. Multiconfigurational nuclear-electronic orbital approach: Incorporation of nuclear quantum effects in electronic structure calculations. *J. Chem. Phys.* **2002**, *117*, 4106−4118.

(19) Zhao, L.; Tao, Z.; Pavošević, F.; Wildman, A.; Hammes-Schiffer, S.; Li, X. Real-time Time-dependent Nuclear- Electronic Orbital Approach: Dynamics Beyond The Born−Oppenheimer Approximation. *J. Phys. Chem. Lett.* **2020**, *11*, 4052−4058.

(20) Pavošević, F.; Culpitt, T.; Hammes-Schiffer, S. Multicomponent Quantum Chemistry: Integrating Electronic and Nuclear

V

Quantum Effects Via The Nuclear−electronic Orbital Method. *Chem. Rev.* **2020**, *120*, 4222−4253.
(21) Coutinho-Neto, M. D.; Viel, A.; Manthe, U. The Ground State Tunneling Splitting of Malonaldehyde: Accurate Full Dimensional Quantum Dynamics Calculations. *J. Chem. Phys.* **2004**, *121*, 9207−9210.
(22) Yu, Q.; Roy, S.; Hammes-Schiffer, S. Nonadiabatic Dynamics of Hydrogen Tunneling with Nuclear-Electronic Orbital Multistate Density Functional Theory. *J. Chem. Theory Comput.* **2022**, *18*, 7132−7141.
(23) Suzuki, M. General Decomposition Theory of Ordered Exponentials. *Proceedings of the Japan Academy, Series B* **1993**, *69*, 161−166.
(24) Reiner, J.-M.; Zanker, S.; Schwenk, I.; Leppäkangas, J.; Wilhelm-Mauch, F.; Schön, G.; Marthaler, M. Effects of Gate Errors in Digital Quantum Simulations of Fermionic Systems. *Quantum Science and Technology* **2018**, *3*, 045008.
(25) Yuan, X.; Endo, S.; Zhao, Q.; Li, Y.; Benjamin, S. C. Theory of Variational Quantum Simulation. *Quantum* **2019**, *3*, 191.
(26) McArdle, S.; Jones, T.; Endo, S.; Li, Y.; Benjamin, S.; Yuan, X. Variational Ansatz-based Quantum Simulation of Imaginary Time Evolution. *Npj Quantum Inf.* **2019**, *5*, 75.
(27) Press, W.; Teukolsky, S.; Vetterling, W.; Flannery, B. *Numerical Recipes: The Art of Scientific Computing*, 3rd ed.; Cambridge University Press: New York, NY, 2007.
(28) For the entanglement entropy to show the symmetric pattern for the bond breaking and dissociation with the right and left oxygen atoms, one needs to align the phase signature for the orbitals of the "Middle" and "Right" setups. We perform the similarity transformation for the "Middle" Hamiltonian and wave function at the midpoint of the simulation—equivalent to flipping the phase signature on orbitals of the middle setup (see Figure 2).
(29) van den Berg, E.; Minev, Z. K.; Kandala, A.; Temme, K. Probabilistic Error Cancellation with Sparse Pauli-Lindblad Models on Noisy Quantum Processors. *arXiv* **2022**, 2201.09866.

https://doi.org/10.1021/acs.jpclett.3c01589
J. Phys. Chem. Lett. 2023, 14, 7065−7072

Supporting Information: Nonadiabatic Nuclear-Electron Dynamics: A Quantum Computing Approach

Arseny Kovyrshin,[*,†] Mårten Skogh,[*,†,‡] Lars Tornberg,[*,†] Anders Broo,[*,†] Stefano Mensa,[*,¶] Emre Sahin,[¶] Benjamin C. B. Symons,[*,¶] Jason Crain,[*,§] and Ivano Tavernelli[*,⊥]

† *Data Science and Modelling, Pharmaceutical Sciences, R&D, AstraZeneca Gothenburg, Pepparedsleden 1, Molndal SE-431 83, Sweden*

‡ *Department of Chemistry and Chemical Engineering, Chalmers University of Technology, Gothenburg, Sweden*

¶*The Hartree Centre, STFC, Sci-Tech Daresbury, Warrington, WA4 4AD, United Kingdom*

§*IBM Research Europe, Hartree Centre STFC Laboratory, Sci-Tech Daresbury, Warrington WA4 4AD, United Kingdom*

‖*Department of Biochemistry, University of Oxford, Oxford, OX1 3QU, UK*

⊥*IBM Quantum, IBM Research Europe – Zurich, Säumerstrasse 4, 8803 Rüschlikon, Switzerland*

E-mail: arseny.kovyrshin@astrazeneca.com; marten.skogh@astrazeneca.com; lars.tornberg@astrazeneca.com; anders.broo@astrazeneca.com; stefano.mensa@stfc.ac.uk; benjamin.symons@stfc.ac.uk; jason.crain@ibm.com; ita@zurich.ibm.com

Discussion on the algorithm's computational scaling

The number of qubits will scale linearly with the size of the basis set used to describe the electronic and the nuclear degrees of freedom. Concerning the number of parameters and the circuit depth required for the implementation of the time propagation algorithm, these depend crucially on the wavefunction Ansatz implemented. In general, it can be shown that a polynomial number of variational parameters suffice to achieve the desired accuracy. In the case of a variational time-propagation algorithm such as varQTE[1] the circuit depth remains constant for the entire propagation time. On the contrary, in a Trotter-like implementation of the time evolution operator the circuit depth will increase linearly with the simulation time[1].

Illustration of the quantum circuit for dynamics

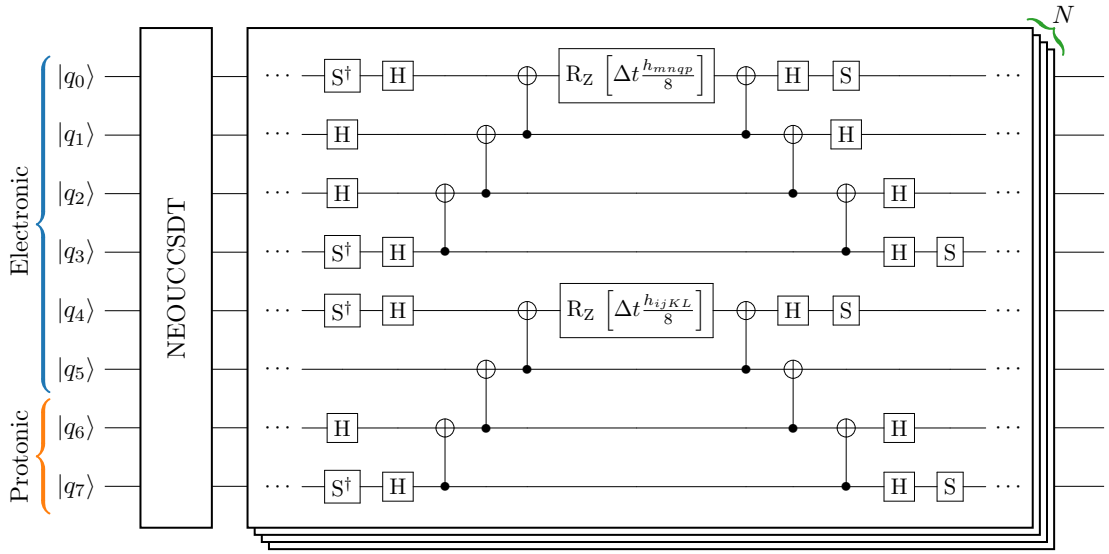


Figure S1: Schematic drawing of the time evolution operator, Eq. (9) implemented as a quantum circuit with a series of N time steps of size Δt on 8 qubits.

Diagram of the time-evolution of the Hamiltonian components

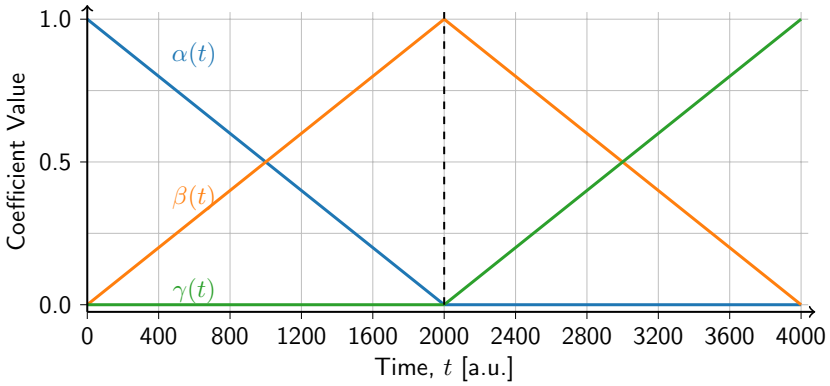


Figure S2: Adiabatic mixing of the Hamiltonians. The different Hamiltonians are mixed in pairs. Initial mixing is performed between $\hat{H}_L$ and $\hat{H}_M$ through $\alpha(t)\hat{H}_L + \beta(t)\hat{H}_M$, with the subsequent mixing of $\hat{H}_M$ and $\hat{H}_R$ as $\beta(t)\hat{H}_M + \gamma(t)\hat{H}_R$.

Figure of the adiabatic time evolution of the coupled electron-nuclear dynamics

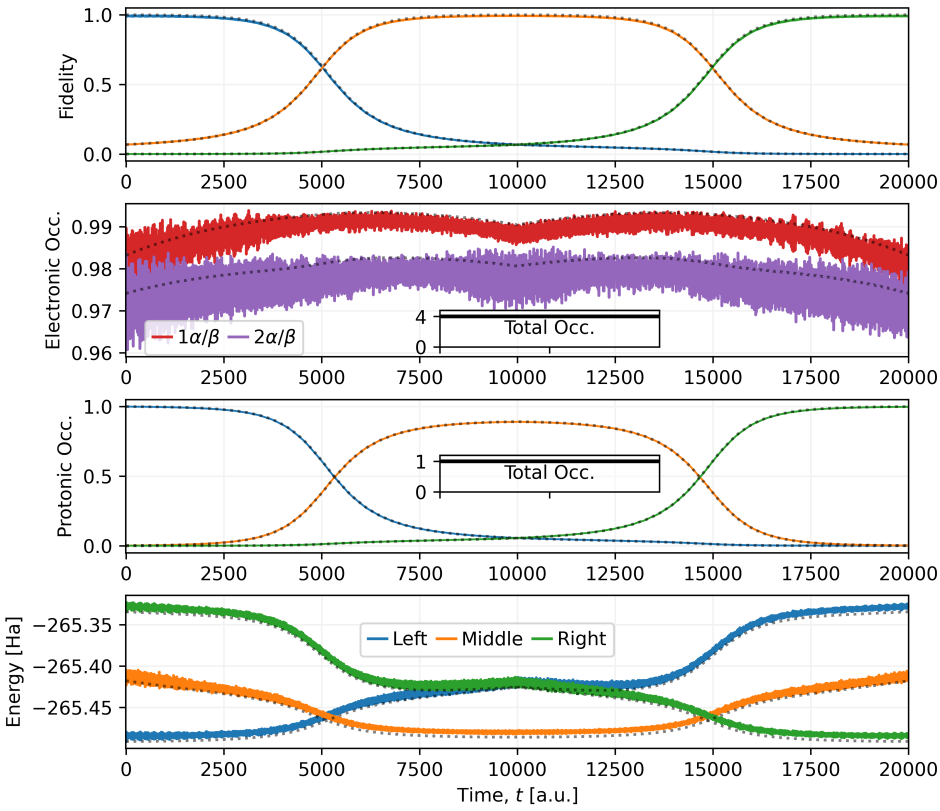


Figure S3: Adiabatic evolution performed for time $t_f = 20000$ a.u. with $\Delta t = 1.0$ a.u.. There are slight deviations from the reference results which are shown as black dotted lines. The reference evolution uses the RK4 method.

Figure of the entropy and state fidelity computed for the adiabatic coupled electron-nuclear dynamics

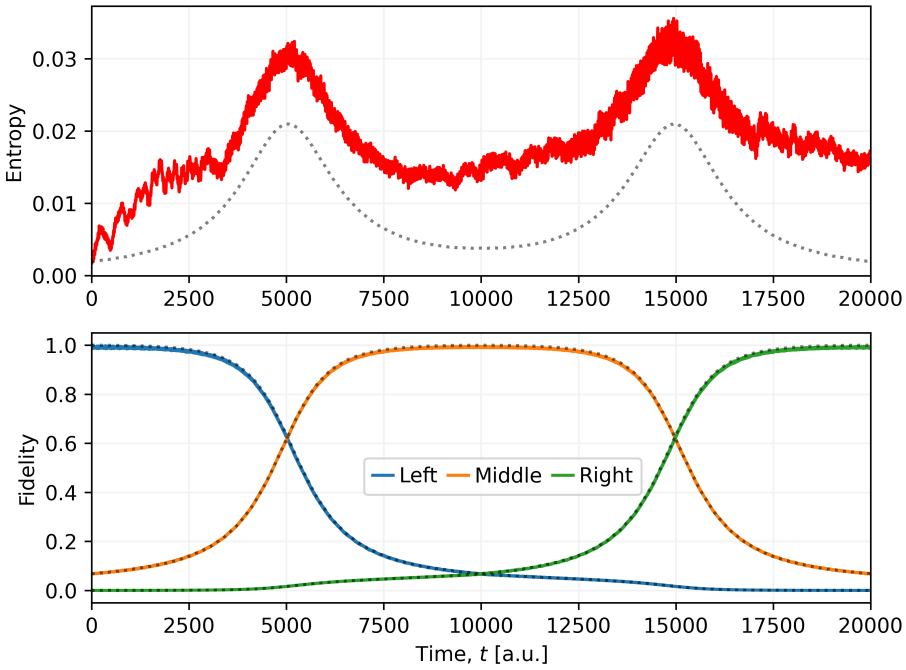


Figure S4: Nuclear-electron entanglement entropy (top) and the fidelities with respect to reference ground states in the 'Left', 'Middle', 'Right' setups (bottom) as a function of time in the adiabatic evolution for $t_f = 20000$ a.u. with $\Delta t = 1$ a.u.. The reference values, generated using RK4, are shown as black dotted lines. The values for fidelities and entanglement entropy show similar results to the shorter evolution with $t_f = 4000$ a.u. proving that the evolution operator is stable over a long-time simulation and does not accumulate errors.

References

(1) Miessen, A.; Ollitrault, P. J.; Tacchino, F.; Tavernelli, I. Quantum Algorithms for Quantum Dynamics. *Nature Computational Science* **2023**, *3*, 25–37.

V