Optimization algorithms benchmarking and intra-architecture gate mapping studies for the variational quantum eigensolver

Doctoral thesis for the award of the doctoral degree
of the Faculty of Mathematics and Natural Sciences
of the University of Cologne

submitted by
Fotios Gkritsis

accepted in the year 2026

Abstract

The advancements in classical computational components such as CPUs and GPUs have had a reverberating impact on the democratization of computational power available, with the field of computational chemistry reaping significant benefits from the propulsion of the limits of systems studied. The accurate prediction of chemical properties and reaction dynamics through computational simulation is now a centerpiece of modern chemical research, significantly accelerating discoveries research areas of materials science and drug development, as well as industrially relevant processes such as the emission of aldehydes from soft foams, flame retardants and pyrolysis and the carbonylation of epoxides. These achievements of computational chemistry still are overshadowed in the face of fundamental bottlenecks where classical methods remain constrained by the exponential scaling when treating large and/or strongly correlated quantum many-body systems. This innate computational complexity often constitutes chemically accurate calculations based on first principles impractical, or plain impossible, for systems beyond modest sizes.

Quantum computing offers a potential pathway around these limitations, originally formally formulated by David Deutsch as a tool for simulating quantum systems, with the first quantum algorithms indicating speed-ups compared to classical counterparts arising from Peter Shor and Lov Grover for the prime factorization of an integer number and unstructured search respectively. Relying on fundamental theorems by Claude Shannon, it has been demonstrated that through the use of error-correcting codes fault-tolerant quantum computers are able to safeguard information against noise and decoherence, strongly indicating towards exponential speedups in solving central computational chemistry problems, such as determining ground state energies or reaction pathways of chemical systems. Substantial strides in experimental quantum computing over the past decade, including proof-of-principle demonstrations of quantum error correction, mean to bring full-scale fault-tolerant quantum computing within the grasp of humanity. The breakthroughs needed to attain scalable, error-corrected quantum computation with capabilities exceeding those of classical methods in real-world scenarios is likely still some years away. While striving to address the engineering challenges standing in the way of achieving fault tolerant quantum computing, substantial effort has been focused toward developing and benchmarking hybrid algorithms tailored for the Noisy Intermediate- Scale Quantum (NISQ) devices, available now to researchers. These hybrid quantum-classical algorithms, with the most notable being the Variational Quantum Eigensolver (VQE), allow the probing of quantum computing's utility in computational chemistry even in the face of significant hardware noise and limited coherence times.

The work presented in this thesis introduces several key additions the VQE framework, geared especially towards to NISQ hardware, where the mindful use of available resources is imperative. First, the implementation of the Conjugate Model Search optimizer is described, an optimization method that integrates simple models that are surrogate to the original function, combined with conjugate gradient optimization to navigate effectively parameter landscapes typical of variational algorithms, prone to noise. Second, this thesis contains the proposal and extensive study of a hyperparameter tuning procedure employing genetic algorithms, allowing for comprehensive simulation benchmarks that augment the robustness and efficiency of VQE implementations with respect to convergence time through real wall-time models of quantum processors.

Continuing, a bridge is presented between quantum simulation and quantum chemistry, through qubit-to-fermion mappings of existing quantum gates, further enabling the direct and accurate modeling of chemical systems on trapped atom quantum hardware with the added arsenal provided by techniques designed for qubit architectures. Additionally, experimental data contained in this thesis include detailed benchmarking data collected from state-of-the-art trapped ion quantum processors available at the Johannes Gutenberg Universität Mainz, demonstrating the actual viability and performance of various methodologies for extracting results. Lastly, the state of VQE is examined, as it served as a springboard for a plethora of advancements in the development of quantum computing hardware, as well as the direction the field of simulation of computational quantum chemistry via quantum devices is headed towards in the coming years.

Contents

1 Chemistry Overview

Following the development of quantum mechanics and the formulation of the Schrödinger equation, the fundamental mechanisms of the interaction of nuclei and electrons have been understood and the computational intractability of those interactions have been encapsulated in the electronic structure problem. Despite the equations governing the nature of chemistry being known, the exponential scaling of computational intractability of the resulting quantum many-body problem poses significant hurdles to the calculation of properties of molecular systems.

The task of finding computationally tractable models for the intractable original problem is almost akin to an art-form, as striking a balance between easing restrictions on the requirement for computational resources and still retaining the ability to predict chemical properties with reasonable accuracy remains a key driver of innovation in the field of quantum chemistry, both classically and in the context of quantum computing [1]. Ab-initio chemistry methods that only take physical constants as inputs have flourished with the available computational hardware from mean-field methods such as Hartree-Fock [2] initially limited to approximations omitting the treatment of electron correlations, to more sophisticated methods being able to calculate a wide range of properties including absolute and relative energies of given molecules, electronic charge density distributions, dipoles and higher multipole moments, vibrational frequencies or other spectroscopic quantities with high accuracy.

These methods enable the ability to make predictions as a complement to experimental studies into these systems by being competitive in the accuracy of their predictions. Algorithms such as Density Functional Theory (DFT) [3] and the Density Matrix Renormalization Group (DMRG) [4] have made possible the simulation of large scale chemically relevant systems, they are not without flaws, as they struggle addressing phenomena such as intermolecular interactions and transition states of strongly correlated systems.

In the first section of this thesis, the formulation of the electronic structure problem, along with common approximations and treatments in an effort to find solutions, will be briefly presented. Due to the central role that the field of quantum chemistry plays in the race towards quantum advantage, it is important to understand the underlying methods that have accumulated and their impact in shaping the theory for simulating these systems using quantum hardware. For a more comprehensive overview overview of the foundations of quantum chemistry, one can refer to [5, 6, 7].

1.1 A brief overview of quantum chemistry fundamentals

One of the central aims of quantum chemistry is to accurately depict the the behavior of electrons in atoms and molecules, in an effort to elucidate the workings of chemical bonding and molecular interactions. This can be achieved by solving approximately the non-relativistic time-independent Schrödinger equation [5]:

$$\hat{\mathcal{H}}|\Phi\rangle = \mathcal{E}|\Phi\rangle \tag{1}$$

with $\hat{\mathcal{H}}$ being the Hamiltonian operator describing the system of interest consisting of M-many nuclei and N-many electrons whose representation is done through the position vectors $\boldsymbol{R_A}$ and $\boldsymbol{r_i}$ respectively. By denoting the distances between the i-th electron and the A-th nucleus as $\boldsymbol{r_{iA}} = |\boldsymbol{r_{iA}}| = |\boldsymbol{r_{iA}} - \boldsymbol{R_A}|$; between two electrons i and j as $\boldsymbol{r_{ij}} = |\boldsymbol{r_i} - \boldsymbol{r_j}|$ and between two nuclei A and B, as $\boldsymbol{R_{AB}} = |\boldsymbol{R_A} - \boldsymbol{R_B}|$, as well as Z_A and M_A the atomic number and mass of nucleus A respectively, one can write down the Hamiltonian operator as:

$$\begin{aligned}\hat{\mathcal{H}} = &-\sum_{i=1}^{N}\frac{1}{2}\nabla_i^2 - \sum_{A=1}^{M}\frac{1}{2M_A}\nabla_A^2 - \sum_{i=1}^{2}\sum_{A=1}^{M}\frac{Z_A}{\boldsymbol{r_{iA}}} \\ &+ \sum_{i=1}^{N}\sum_{j>i}^{N}\frac{1}{\boldsymbol{r_{ij}}} + \sum_{A=1}^{M}\sum_{B>A}^{M}\frac{Z_A Z_B}{\boldsymbol{R_{AB}}}\end{aligned} \tag{2}$$

In an effort to alleviate some of the complexity of the Hamiltonian operator $\mathcal{H}$, by utilizing the orders of magnitude scale difference in mass between the electrons and the nuclei in a molecule, to approximate the movement of electrons as one through fixed nuclei [6]. This approximation, called the **Born-Oppenheimer approximation**, allows for the elimination of kinetic energy term of

the nuclei in the Hamiltonian operator $\hat{\mathcal{H}}$ and the approximation of the repulsion between them as constant.

In atomic units, the electronic Hamiltonian for depicting the behavior of a system with N electrons and M point charges after the Born-Oppenheimer approximation can be described as [5]:

$$\mathcal{H}_{elec} = -\sum_{i=1}^{N} \frac{1}{2}\nabla_i^2 - \sum_{i=1}^{N}\sum_{A=1}^{M} \frac{Z_A}{r_{iA}} + \sum_{i=1}^{N}\sum_{j>1}^{N} \frac{1}{r_{ij}} \tag{3}$$

The solution of a Shrödinger equation involving the Hamiltonian in Eq. 3 is an electronic wavefunction $\psi(\boldsymbol{r})$, also commonly referred to as *spatial orbitals*. However, for the accurate depiction of the physical behavior of an electron, the inclusion of the *Pauli exclusion principle* is necessary via spin as an additional degree of freedom, by defining spin functions $\alpha(w), \beta(w)$ in common chemistry notation, denoting spin-up($\uparrow$) and spin down($\downarrow$) respectively. These spin functions are only required to be complete and orthonormal and by adopting the compact formalism of $\mathbf{x} \mapsto \{\boldsymbol{r}, w\}$, one can then introduce *spin orbitals* $\boldsymbol{\chi}(\mathbf{x})$. As a consequence, the separable wavefunction can be written as the product of spin orbitals as:

$$|\Phi^{\mathrm{HP}}(\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_N)\rangle \equiv \boldsymbol{\chi_1}(\mathbf{x}_1)\boldsymbol{\chi_2}(\mathbf{x}_2)\ldots\boldsymbol{\chi_N}(\mathbf{x}_N) \tag{4}$$

called a *Hartree product*. In order for an electronic wavefunction to be considered a valid fermionic wavefunction, they must fulfill the antisymmetric property under the exchange of any two electrons of coordinates $\boldsymbol{\chi_i}$ and $\boldsymbol{\chi_j}$:

$$|\Phi(\ldots, \boldsymbol{\chi_i}, \ldots, \boldsymbol{\chi_j}, \ldots, \boldsymbol{\chi_N})\rangle = -|\Phi(\ldots, \boldsymbol{\chi_j}, \ldots, \boldsymbol{\chi_i}, \ldots, \boldsymbol{\chi_N})\rangle \tag{5}$$

Hartree products alone do not obey the antisymmetric principle described in Eq.5, but an appropriate linear combination of them does, through the use of *Slater determinants* [5]:

$$|\Phi^{\mathrm{SD}}(\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_N)\rangle = \frac{1}{\sqrt{N!}} \begin{vmatrix} \boldsymbol{\chi_1}(\mathbf{x}_1) & \boldsymbol{\chi_2}(\mathbf{x}_1) & \ldots & \boldsymbol{\chi_N}(\mathbf{x}_1) \\ \boldsymbol{\chi_1}(\mathbf{x}_2) & \boldsymbol{\chi_2}(\mathbf{x}_2) & \ldots & \boldsymbol{\chi_N}(\mathbf{x}_2) \\ \vdots & \vdots & \ddots & \vdots \\ \boldsymbol{\chi_1}(\mathbf{x}_N) & \boldsymbol{\chi_2}(\mathbf{x}_N) & \ldots & \boldsymbol{\chi_N}(\mathbf{x}_N) \end{vmatrix} \tag{6}$$

Swapping the $\boldsymbol{\chi}$ coordinates between any two electrons i, j is equivalent to swapping two rows of the Slater determinant, which incurs a change in the sign of the determinant, as well as the fact that two electrons occupying the same spin orbital would have as a consequence two columns being equal, resulting in the determinant to be zero, consolidate Slater determinants as valid constructs of fermionic wavefunctions under the antisymmetric principle detailed in Eq. 5 and the Pauli exclusion principle [5].

A convenient shorthand notation for a normalized Slater determinant containing only the Trace of the determinant of N electrons occupying N spin orbitals $(\boldsymbol{\chi_i}, \boldsymbol{\chi_j}, \ldots, \boldsymbol{\chi_k})$ as well as omitting the normalization constant $(N!)^{-1/2}$ is usually introduced as:

$$|\Phi^{\mathrm{SD}}(\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_N)\rangle = |\boldsymbol{\chi_i}(\mathbf{x_1}), \boldsymbol{\chi_j}(\mathbf{x_2}), \ldots, \boldsymbol{\chi_k}(\mathbf{x_N})\rangle \tag{7}$$

The notation can be further simplified by the consistent choice of electron label ordering to be $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_N$ as:

$$|\Phi^{\mathrm{SD}}(\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_N)\rangle = |\boldsymbol{\chi_i}, \boldsymbol{\chi_j}, \ldots, \boldsymbol{\chi_k}\rangle \tag{8}$$

This notation is especially useful when considering another representation of Slater determinants via the use of fermionic creation and annihilation operators $a_i^\dagger$ and a_i which create or destroy an electron in the spin orbital $\boldsymbol{\chi}_i$ respectively. By assuming a fermionic vacuum state $|0\rangle$ which contains zero electrons and defining the fermionic operator relations:

Fermionic operators

$$\hat{a}_i|0\rangle = 0 \quad \{\hat{a}_i, \hat{a}_j\} = \{\hat{a}_i^\dagger, \hat{a}_j^\dagger\}, = 0 \tag{9}$$

$$\langle 0|0\rangle = 1 \quad \{\hat{a}_i, \hat{a}_j^\dagger\} = \delta_{i,j}. \tag{10}$$

where $\{,\}$ denotes anti-commutation of operators. For these relations to hold, one needs to consistently choose an orthonormal set of spatial orbitals and spin functions, and by extension, spin

orbitals, as $\langle \boldsymbol{\chi}_i | \boldsymbol{\chi}_j \rangle = \delta_{ij}$. Upon treating electrons in pairs such as the Bardeen-cooper-Scheifer theory as a model of superconductivity [8], or the Doubly Occupied Configuration Interaction (DOCI) theory as an approximation to the electronic structure simulation problem, the excitations become of hard-core bosonic nature [9]:

Bosonic operators

$$[\hat{b}_p, \hat{b}_q^\dagger] = [\hat{b}_p^\dagger, \hat{b}_q^\dagger] = [\hat{b}_p, \hat{b}_q] = 0 \quad (p \neq q) \tag{11}$$

$$\{\hat{b}_p^\dagger, \hat{b}_q^\dagger\} = \{\hat{b}_p, \hat{b}_q\} = 0 \tag{12}$$

$$\{\hat{b}_p, \hat{b}_q^\dagger\} = 1 \tag{13}$$

Equipped with these operator relations, the Hamiltonian $\hat{\mathcal{H}}_{elec}$ in Eq. 3 can now be rewritten in second quantization form as [10]:

$$\hat{\mathcal{H}}_{elec} = \sum_{A>B} \frac{Z_A Z_B}{\boldsymbol{R}_{AB}} + \sum_{pq} h_{pq} \hat{a}_p^\dagger \hat{a}_q + \sum_{pqrs} V_{pqrs} \hat{a}_p^\dagger \hat{a}_q^\dagger \hat{a}_r \hat{a}_s \tag{14}$$

The matrices h_{pq} and V_{pqrs} represent one and two-body electron integrals respectively, these matrix elements of these integrals are defined as:

$$h_{ij} = \int d\boldsymbol{r} \boldsymbol{\chi}_i^*(\boldsymbol{r}) h_i \boldsymbol{\chi}_j(\boldsymbol{r}) = \int d\boldsymbol{r} \boldsymbol{\chi}_i^*(\boldsymbol{r}) \left(-\nabla^2(r) + \sum_A \frac{Z_A}{r_{iA}} \right) \boldsymbol{\chi}_j(\boldsymbol{r}) \tag{15}$$

$$V_{ijkl} = \frac{1}{2} \int \int d\boldsymbol{r}_1 d\boldsymbol{r}_2 \boldsymbol{\chi}_i^*(\boldsymbol{r}_1) \boldsymbol{\chi}_j^*(\boldsymbol{r}_2) \frac{1}{r_{ij}} \boldsymbol{\chi}_k^*(\boldsymbol{r}_2) \boldsymbol{\chi}_l^*(\boldsymbol{r}_1) \tag{16}$$

Furthermore, in the realm of quantum chemistry, the following notation regarding these electronic integrals has been introduced when dealing with spin orbitals [5]:

$$\begin{aligned}
[i|h|j] &= \langle i|h|j \rangle = \int d\mathbf{x}_1 \boldsymbol{\chi}_i^*(\mathbf{x}_1) h(\mathbf{r}_1) \boldsymbol{\chi}_j(\mathbf{x}_1) \\
[ik|jl] &= \langle ij|kl \rangle = \langle \boldsymbol{\chi}_i \boldsymbol{\chi}_j | \boldsymbol{\chi}_k \boldsymbol{\chi}_l \rangle = \int d\mathbf{x}_1 \mathbf{x}_2 \boldsymbol{\chi}_{\boldsymbol{i}}^*(\mathbf{x_1}) \boldsymbol{\chi}_{\boldsymbol{j}}^*(\mathbf{x_2}) \frac{1}{rij} \boldsymbol{\chi}_{\boldsymbol{k}}^*(\mathbf{x_1}) \boldsymbol{\chi}_{\boldsymbol{l}}^*(\mathbf{x_2}) \\
\langle ij||kl \rangle &= \langle ij||kl \rangle - \langle ij||lk \rangle = \int d\mathbf{x}_1 \mathbf{x}_2 \boldsymbol{\chi}_{\boldsymbol{i}}^*(\mathbf{x_1}) \boldsymbol{\chi}_{\boldsymbol{j}}^*(\mathbf{x_2}) \frac{1}{rij} (1 - \mathcal{P}_{12}) \boldsymbol{\chi}_{\boldsymbol{k}}^*(\mathbf{x_1}) \boldsymbol{\chi}_{\boldsymbol{l}}^*(\mathbf{x_2})
\end{aligned} \tag{17}$$

In this formalism, $\langle ij||kl \rangle$ describes an antisymmetrized two-electron integral as it proves quite handy in various derivations while the operator $\mathcal{P}$ ensures the antisymmetrization of the normal two-electron integral by accounting for the exchange of the coordinates of electrons one and two. The notation of $[ij|kl] = \langle ik|jl \rangle$ is also commonly referred to as **chemists notation**, which although is not used in this thesis, it is valuable for reading literature heavily geared towards chemistry rather than physics. In contrast, for spatial orbitals another set of brackets is used to underline the difference in the sets of orbitals used [5]:

$$(i|h|j) = h_{ij} = (\psi_i|h|\psi_j) = \int d\mathbf{r}_1 \psi_i^*(\mathbf{r}_1) h(\mathbf{r}_1) \psi_j(\mathbf{r}_1) \tag{18}$$

$$(ij|kl) = (\psi_i \psi_j | \psi_k \psi_l) = \int d\mathbf{r}_1 d\mathbf{r}_2 \psi_i^*(\mathbf{r}_1) \psi_j(\mathbf{r}_1) \frac{1}{r_{12}} \psi_k^*(\mathbf{r}_2) \psi_l(\mathbf{r}_2) \tag{19}$$

These are especially useful for defining the *Coulomb integral* $J_{ij} \equiv (ii|jj)$ pertaining to the electronic repulsion and the *exchange integral* $K_{ij} \equiv (ij|ij)$ which accounts for the application of the Pauli exclusion principle. Moreover, owing to the basis set $\{\boldsymbol{\chi}_i\}$ of spin orbitals being in general not orthonormal, thus having non vanishing overlap with one another, as is frequent throughout optimization procedures, the overlap matrix S_{ij} of spin orbitals is defined as [5]:

$$S_{ij} = \langle \boldsymbol{\chi_i} | \boldsymbol{\chi}_j \rangle \tag{20}$$

1.1.1 The Hartree-Fock approximation and method

The Hartree-Fock approximation, equivalently dubbed the molecular orbital approximation, is a vital pillar in the study of quantum chemistry. Most computational methods of quantum chemistry build upon the foundational blocks provided by this theory. Through the second quantization of the electronic structure Hamiltonian $\hat{H}$ presented in Eq. 14, the task of finding its ground state energy as well as other excited state energies, has been divided in two fronts; firstly finding a set of orbitals (and by extension Slater determinants) for the given molecule and afterwards solving the eigenvalue problem. An initial set of suitable orbitals can be retrieved via the Hartree-Fock formalism while the solutions deriving from the second quantization of the electronic structure Hamiltonian can be found using a quantum computer and variational eigensolvers [11] or via error correcting algorithms in fault tolerant quantum computers.

A systematic way of calculating the molecular energy under the regime of only mean-field electron correlations, also called the Hartree-Fock energy E_{HF}, is by choosing a set of molecular orbitals, and by extension a single Slater determinant that represents the minimum of the energy of the electronic structure Hamiltonian. A linear combination $|\varphi\rangle = \sum_i c_i|\boldsymbol{\chi}_i\rangle$ of spin orbitals with coefficient weights c_i as a variational trial wave function is employed to find that minimum energy [5]:

$$E_{\mathrm{HF}} = \langle\varphi|\hat{H}|\varphi\rangle = \sum_{ij} c_i^* c_j \langle\boldsymbol{\chi}_i|\hat{H}|\boldsymbol{\chi}_j\rangle \tag{21}$$

imposing the constraint that the wavefunction remains normalized [5]

$$\langle\varphi|\varphi\rangle - 1 = \sum_{ij} c_i^* c_j \langle\boldsymbol{\chi}_i|\hat{H}|\boldsymbol{\chi}_j\rangle - 1 = 0 \tag{22}$$

A common way to tackle the process of finding the expectation value of the Hamiltonian operator H is through the use of Lagrange multipliers E as [5]:

$$\begin{aligned}\mathcal{L} &= \langle\varphi|\hat{H}|\varphi\rangle - E(\langle\varphi|\varphi\rangle - 1) \\ &= \sum_{ij} c_i^* c_j \langle\boldsymbol{\chi}_i|\hat{H}|\boldsymbol{\chi_j}\rangle - E\left(\sum_{ij} c_i^* c_j \langle\boldsymbol{\chi}_i|\boldsymbol{\chi_j}\rangle - 1\right)\end{aligned} \tag{23}$$

Therefore the minimization of E_{HF} leads to the Hartree-Fock equations:

$$\hat{f}|\varphi\rangle = S_{ij}E|\varphi\rangle \tag{24}$$

where $\hat{f}$ is the Fock operator, defined as:

$$\hat{f} = \sum_i h_i + \sum_k^{N/2} J_k - K_k \tag{25}$$

These Hartree-Fock equations are solved with the *self consistent field* (SCF) procedure, by making an initial guess for the orbitals $\boldsymbol{\chi}_i$ to construct the Fock operator using the currently selected orbitals, and then solving iteratively the eigenvalue problem to update the orbitals contributing to the integral operators and orbital energies until convergence.

Although the Hartree-Fock method serves as a good starting point for describing molecular electronic structure, it fails to describe correlations between electrons, which occurs from electron-electron interactions. This inability has played a catalytic role in the development of post-Hartree-Fock methods, such as Configuration Interaction (CI) [7] and Coupled Cluster (CC) [12] which account for interactions due to such correlations. Calculations in computational chemistry are usually allowed to run iteratively until the resolution of energies up to ΔE of 1.6×10^{-3} Hartree (1 Hartree is equal to $4.3598 \times 10^{-18} J$ in SI units). This energy difference is selected as a standard because it is the realm where one can perform energy reaction predictions at standard conditions and reaction rates $\propto e^{-\Delta E/k_B T}$ with temperature T and Boltzmann constant k_B with at least one order of magnitude precision [13, 14].

1.2 Post Hartree-Fock methods

Though the Hartree-Fock method provides a considerable starting point in the effort of describing the behavior of electrons in a molecular setting, as described in Section 1.1.1, mean-field approaches fail to depict the correlation dynamics of electrons. This hurdle poses the need for post-Hartree-Fock methods such as Configuration Interaction and Coupled Cluster. By using the results of the Hartree-Fock along with achieving a balance the number of orbitals that are considered when modeling the system, one is able to efficiently tackle the task of finding the energy of a molecular state.

1.2.1 Truncations of the full orbital space

In order to make post-Hartree-Fock methods computationally tractable with sacrificing little accuracy, one chooses only a subspace of the full orbital space to be *active* as the most relevant to the problem to be simulated, resulting in the concept of **active spaces**. The curation of these active spaces enable the treatment of electron correlation in the most significant regions of the molecular orbital space, especially those orbitals involved in the reactivity of excited-state properties.

The spatial orbitals contained in complete active spaces are divided in core, active and virtual orbitals, denoting orbitals that always hold two electrons, partially occupied orbitals or orbitals always containing zero electrons respectively. Consequently, these active spaces are defined by two parameters, the number of active electrons in the computation and the number of active spatial orbitals that one considers. This flexibility allows one to expand the space to include orbitals associated with the bonds of interest, or shrink the space to omit bonds that do not play an active role in the process of interest.

Although advances have been made in the field of automated algorithms for providing assistance in the process of active space selection, it is often the case that chemical intuition as well as preliminary calculations are required to pinpoint orbitals that exhibit partial occupancy. The active space selection process remains a crucial step that can influence the accuracy and efficiency of calculations in a major way, as orbitals outside the constructed active space may still hold significant information about the behavior of the process, making larger active spaces or other approaches necessary [15].

1.2.2 Single reference Truncated Configuration Interaction

Expansion around the Hartree-Fock determinant found through the SCF procedure to construct an exact wave function for any state of the system to form a complete set of N-electron wavefunctions. Since these wavefunctions can be adequately described by defining a *configuration* of spin orbitals forming these N-electron determinants, this procedure is called *configuration interaction* (CI). When the wavefunction is produced from the infinite set of determinants deriving from N electrons, the resulting wavefunction is described as:

$$|\Psi^{\mathrm{CI}}\rangle = (1 + C_1 + C_2 + C_3 + \cdots + C_n)|\Psi^{\mathrm{HF}}\rangle \tag{26}$$

The excitation operator for the promotion of an electron to a higher orbital is denoted as C_i and the prefactors are called CI coefficients. In practice, of course, one cannot have access to an infinite basis set of spin orbitals, therefore choosing to work with a finite set of $\{\boldsymbol{\chi}_i | i = 1, 2, \ldots, 2K\}$. Despite the fact that these $\binom{2K}{N}$ do not form a complete N-electron basis, the diagonalization of the (also finite) Hamiltonian matrix yields exact solutions within the subspace spanned by these $2K$ orbitals, with this procedure being known as *Full Configuration Interaction* or Full CI as shorthand [5, 16].

Therefore by employing the use of active spaces one can tackle larger problems, albeit with lower accuracy, and the selection of the orbitals included in the active space is pivotal to the successful description of electronic excitations and strong correlation effects [17]. A Complete Active Space Configuration Interaction (CASCI) method constructs a CI wavefunction within the predefined active space, which contains orbital that are relevant to the chemical process one is interested in. Upon the decision of the active space, all the possible Slater determinants are generated and then by solving the Configuration interaction Hamiltonian matrix to obtain the CI coefficients, one arrives to the corresponding energies [16].

1.2.3 Single reference Truncated Coupled Cluster

The Coupled Cluster family of methods present a different approach, by using an exponential ansatz for the wavefunction, as a means of accurately describing weakly systems with dynamic electron correlation [18]. Starting from the Hartree-Fock state, and $\hat{T}$ being defined as the cluster operator, with T_n portraying the n-body excitation operators and t coefficients within the $\hat{T}_i$ cluster operator, the wavefunction prepared by Coupled Cluster methods is expanded as:

$$|\Psi^{\mathrm{CC}}\rangle = \exp(\hat{T})|\Psi^{\mathrm{HF}}\rangle = \exp(\hat{T}_1 + \hat{T}_2 + \cdots + \hat{T}_n)|\Psi^{\mathrm{HF}}\rangle \tag{27}$$

The first two orders of the cluster operator $\hat{T}_n$ are going to be described explicitly in this thesis as they are the more frequently used, as $\hat{T}_1$ and $\hat{T}_2$ with the index sets I_{occ} and I_{virt} used to describe the indices containing the occupied and virtual orbitals respectively [5].

$$\begin{aligned} \hat{T}_1 &\equiv \sum_{pq} t_{pq}\hat{a}_p^\dagger \hat{a}_q \qquad\qquad p \in I_{virt}, q \in I_{occ} \\ \hat{T}_2 &\equiv \sum_{pqrs} t_{pqrs}\hat{a}_p^\dagger \hat{a}_q^\dagger \hat{a}_r \hat{a}_s \quad p, q \in I_{virt}, r, s \in I_{occ} \end{aligned} \tag{28}$$

When truncating the cluster operator $\hat{T}$ at different levels, one can obtain different levels of the Coupled Cluster theory, as CC Singles includes only single excitations via the $\hat{T}_1$ cluster operator. The more widely used truncations of Coupled Cluster theory are CC Singles and Doubles (CCSD), CC Singles, Doubles and Triples (CCSDT) and CC Singles Doubles with perturbative corrections (CCSD(T)).

The CCSD approach is attractive for many quantum chemistry applications as it manages to strike a delicate compromise between computational scaling and accuracy. Owing to its inclusion of single and double excitations via the the $\hat{T}_1$ and $\hat{T}_2$ cluster operators, it manages to capture a significant portion of the dynamic correlation effects for most chemically relevant systems. With a scaling of computational effort of $\mathcal{O}(N^6)$ on the number of orbitals N [19], it finds success in the prediction of a plethora of molecular properties [20].

By including triple excitations in CCSDT, one can significantly enhance the accuracy for systems presenting strong correlation effects or multi-reference characteristics. However the computational cost of CCSDT scaling as $\mathcal{O}(N^8)$ [21] is prohibitive for most applications, bar small or medium-sized systems [20]. This led to the development of hybrid approaches between CCSD and CCSDT, namely CCSD(T) which is CCSD with perturbative corrections to approximate the effect of triple excitations without having to bear the full computational brunt of storing and calculating the coefficients of the triply excited cluster operator. The CCSD(T) is regarded as the "gold standard" as it manages to capture wavefunctions with low approximation error. However, its scaling of $\mathcal{O}(N^7)$ [22] also renders it prohibitive for large scale simulations through classical means.

1.3 Fermion-to-Qubit mappings

Creation and annihilation operators constitute a handy way for the representation of fermionic states. However, most commonly available quantum computing platforms such as superconducting qubits operate with bosons, giving rise to the need for a mapping from fermionic algebra, to bosonic algebra, respecting the anti-commutation relations of the original fermionic behavior. This disparity can be bridged with the help of mappings that allow one to go back and forth in these relations. One most prominent choice throughout the literature for quantum computing is the Jordan-Wigner (JW) mapping [23].

The Jordan-Wigner mapping translates the fermionic creation and annihilation operators into qubit operators, obeying the Lie algebra of Pauli matrices $\boldsymbol{X}, \boldsymbol{Y}$ and $\boldsymbol{Z}$. Simple qubit creation and annihilation operators for a qubit p can be defined as:

$$\begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} = \frac{1}{2}(\boldsymbol{X_p} - i\boldsymbol{Y_p}), \quad \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} = \frac{1}{2}(\boldsymbol{X_p} + i\boldsymbol{Y_p}) \tag{29}$$

However, these operators do not respect the anti-commutation relations of fermionic operators, which are essential in describing the Pauli exclusion principle of fermionic systems. With the introduction of a product of Pauli $\boldsymbol{Z}$ operators called Wigner strings to these transformations, one is able to reinstate the antisymmetric behavior of fermionic wavefunctions. Under the Jordan-Wigner transformation, the fermionic creation and annihilation operators for a fermion with $n \in (0, N]$, where N is the number of total fermions in the system, would be transformed to corresponding qubit operators as follows:

Jordan-Wigner mapping

$$\hat{a}_n \mapsto \frac{1}{2}(\boldsymbol{X_n} - i\boldsymbol{Y_n}) \prod_{n=1}^{n-1} Z_n \tag{30}$$

$$\hat{a}_n^\dagger \mapsto \frac{1}{2}(\boldsymbol{X_n} + i\boldsymbol{Y_n}) \prod_{n=1}^{n-1} Z_n \tag{31}$$

There exist other fermion-to-qubit mappings, such as the Brabyi-Kitayev[24] mapping which make use of the parity transform by combining it with the Jordan-Wigner transform to encode shorter representations of fermionic operators, with the ability to choose a more efficient encoder, realized by the classical data structure of Fenwick trees[25] and variants of the Bravyi-Kitaev transformation yield different results when the total number of fermionic modes $N \bmod 2 \neq 0$. This more involved construction of the qubit operations results in making the Jordan-Wigner transform the usual choice for encoding fermions to qubits.

2 Chapter I: Quantum Simulation on Noisy Intermediate Scale Quantum devices

Quantum computing, initially championed by Richard Feynman as an intuitive way of examining systems in nature, has made advanced considerably from a theoretical concept into a field with notable practical potential. Deutsch and Josza's groundbreaking work demonstrated the quantum advantage of performing certain calculations in fewer steps than classical computers [26]. Subsequent work by Shor and Grover further solidified this advantage, showcasing exponential speedups in factoring and polynomial speedups in unsorted searches, respectively [27, 28].

One of the primary applications of quantum computing is that of Hamiltonian simulation, which involves efficiently implementing time evolution $e^{i\hat{H}t}$, where $\hat{H}$ is the Hamiltonian of the system to be simulated, and estimating its eigenvectors and eigenvalues using Quantum Phase Estimation (QPE). The application of this technique has led to the development of a plethora of quantum algorithms for quantum chemistry [10], paving the way for potential breakthroughs in materials science and drug discovery [29, 30, 31]. To address the resource limitations of early quantum algorithms, researchers have focused on reducing quantum resource requirements and introducing innovative techniques like double factorization [32], qubitization [33] and quantum Krylov diagonalization algorithms [34]. Through advancements such as these, the cost of implementing quantum algorithms has significantly lowered.

While quantum computing offers the promise of exponential speedups over classical alternatives through such algorithms, the realization of fault-tolerant quantum computers remains a critical challenge for their successful execution. Hardware progress has been made with technologies such as ion-traps [35] and superconducting qubits [36], and initial demonstrations of scalable error correction using the surface code have been achieved [37, 36]. However, the development of fully error-corrected quantum machines capable of outperforming classical algorithms remains several years away.

2.1 Variational Quantum Eigensolver and its components

Although efforts towards fault tolerant quantum computers have been intensified in the recent years via demonstrating error suppression through state certification [38] and surface code increase [37], as well as progress in the construction of logical qubits as well as demonstrating logical memory beyond the break-even regime [36], algorithms of interest such as Quantum Fourier Transforms (QFT) and Quantum Phase Estimation (QPE) are out of reach for the quantum hardware available.

The Variational Quantum Eigensolver (VQE) algorithm was created as an effort to make use of quantum computers of limited size that do not have the necessary resources to run algorithms such as QFT or QPE [39, 40]. The study of such systems commonly referred to as Noisy Intermediate Noisy Scale Quantum (NISQ) era devices provides important insights into the effects of noise into circuit execution and the furthering of development of error mitigation techniques, which are crucial pieces towards quantum error corrected systems. The VQE algorithm constitutes an attractive avenue for NISQ systems due to their hybrid quantum-classical nature, as they can offload a significant portion of the computation in classical computers.

The quantum part of the quantum-classical nature of the VQE algorithm arises from utilizing the quantum computer to construct a trial wavefunction $|\psi(\boldsymbol{\theta})\rangle$ with a parametrized quantum circuit with parametrization parameters $\boldsymbol{\theta} = (\theta_1, \theta_2, \ldots, \theta_n)$, where n is the number of parametrized gates in the circuit, serving as an ansatz. The Hamiltonian $\hat{H}$ in quantum chemistry is usually the second quantization of the electronic structure Hamiltonian as described in Section 1.1, but approaches using the first quantization have also been formulated [41] and the task is usually to obtain the lowest energy state of $\hat{H}$ via this parametrized quantum circuit. Upon deciding on a molecule, a molecular structure (geometry) and the active molecular orbitals, while relegating the rest of the orbitals to a mean-field approximation, it is common to use a Hartree-Fock method to construct a wavefunction in order to act as an initial state for the parametrization of the quantum circuit.

By using a mapping to translate a fermionic system to qubits such as the mapping detailed in Section 1.3, one is now in a position to extract expectation values of that Hamiltonian. That process can be done via imaginary time evolution [42] or through the use of decompositions to jointly measurable terms $\hat{O}_m$ which then with a number of rotations into the appropriate measurement basis can be used to form a cost function of the system to act as an oracle.

The classical part of the VQE algorithm stems from the use of a classical optimizer to guide the procedure to the choice of $\boldsymbol{\theta}^*$ to retrieve the lowest energy configuration of the system. This is achieved by making use of the samples obtained by the quantum computer to construct the energy, or other relevant to the optimization procedure quantities, at various configurations of $\boldsymbol{\theta}$ to iteratively find the ground state. By leveraging the variational principle and under the premise that the ground state or an adequately close state that one is aiming is indeed representable by the ansatz used, this procedure gives the lowest eigenvalue of the Hamiltonian that also represents the ground state of the system.

$$E(\boldsymbol{\theta}) = \frac{\langle\psi(\boldsymbol{\theta})|\hat{H}|\psi(\boldsymbol{\theta})\rangle}{\langle\psi(\boldsymbol{\theta})|\psi(\boldsymbol{\theta})\rangle} = \frac{\langle\psi|\hat{H}|\psi\rangle}{\langle\psi|\psi\rangle}, \quad \boldsymbol{\theta}^* = \underset{\boldsymbol{\theta}}{\operatorname{argmin}} E(\boldsymbol{\theta}) \tag{32}$$

An additional advantage to the modular nature of the VQE algorithm is that it lends itself to improving upon each component both irrespectively and in tandem with one another, meaning that advancements in parameter optimization procedures for example, can be studied and occur independently of advancements in ansätze used, but at the same time, advancements in any area can make previously unfeasible sizes, realistic experimental goals, as well as establishing experimental pipelines geared towards bringing all components together [11].

The main components of the VQE algorithm will be discussed in the following sections, the inherent limitations in each as well as possible avenues to alleviate some of these shortcomings are an active topic of research, yielding a significant body of literature. Some of the more succinct overviews on the topic can be found in [43, 44].

2.1.1 Ansätze

Upon deciding on the mapping of the molecular orbitals whose behavior is to be simulated with the help of the quantum computer, the efficient utilization of the available quantum resources is of paramount importance for the feasibility of algorithms and a successful optimization, as limited quantum hardware quality and physical qubit counts permeate the quantum computing

landscape [40]. For that reason, the choice of ansatz is central to the success of Variational Quantum Algorithms, owing to the fact that the ansätze used must be ideally expressive enough to capture the dynamics one is trying to simulate, as well as managing the cost in terms of quantum resources needed to prepare a target state from a reference state. The expressivity of an ansatz is understood to be the ability of the ansatz to explore the Hilbert space where the initial state and the target state are in, and its cost is measured in gates that comprise each operation of the ansatz.

Most quantum algorithms tackling chemistry problems use the Hartree-Fock state $|\psi_{\text{ref}}\rangle = |\psi_{\text{HF}}\rangle = |00\ldots011\ldots1\rangle$ due to its ease in preparing it, namely by a single layer of Pauli rotations [11], as a cornerstone for the construction of more involved wavefunctions with parametrized unitary rotations as:

$$|\psi(\boldsymbol{\theta})\rangle = \hat{U}(\boldsymbol{\theta})|\psi_{\text{ref}}\rangle = \prod_i \hat{U}_i(\theta_i)|\psi_{\text{ref}}\rangle \tag{33}$$

The Unitary Coupled Cluster Singles and Doubles (UCCSD) ansatz is one of the most well known examples of an approach to VQE algorithms that is chemically motivated in its construction of the variational wavefunction [39]. Stemming from classical Coupled Cluster theory, the UCCSD ansatz uses anti-hermitian exponential operators, substituting those described in Section 1.2 to systematically construct a wavefunction that includes electron excitations up to the second order from a reference state, often leading to rapid convergence for weakly correlated systems [45]. Even though any unitary can be decomposed into one- and two-qubit rotations, the decompositions required to implement these exponential operators are not compatible with NISQ computer hardware, so one *trotterizes* these anti-hermitian rotations, through the Trotter-Suzuki formula [45, 46].

$$e^{\hat{H}t} = e^{\sum_{j=1}^{m} \hat{O}_j t/r} = \left(\prod_{j=1}^{m} e^{\hat{O}_j t/r} \right)^r + \mathcal{O}(m^2 t^2/r) \tag{34}$$

This representation becomes exact for the limit that $r \to \infty$, with UCCSD taking the first order of the expansion into account when constructing an ansatz. However, after applying the Jordan-Wigner transformation, the circuit representation one ends up with contains long CNOT ladders that are not compatible with the capabilities of current NISQ devices [47].

On the other extreme of the cost-expressivity pendulum, there exist Hardware-Efficient Ansatz (HEA) approaches, which aim to align themselves more directly with the native operations and connectivity architecture of a given quantum processor, in an attempt to reduce circuit depth [31, 48]. These approaches, instead of being motivated from a specific electronic structure model such as the second quantized Hamiltonian, it relies on repetitive tilings of the parametrized single qubit rotation and entangling gates that are natively available in the hardware platform used, yielding very densely parametrized ansätze. As the available gate sets of any platform are by construction universal, it is possible to construct any arbitrary wavefunction given enough depth.

However, because the hardware efficient ansatz approach is devoid of chemical or physical insights, it is susceptible to parametrization of non-physical parts of the trial wavefunction, making the training of the parametrization parameters of the ansatz prone to the *barren plateaus* problem, in which the optimization landscape is characterized by very small changes to the cost function and its gradient, making the convergence to the global minimum instead of one of the many local minima very challenging. In comparison to UCCSD, in a HEA approach, since it does not preserve any structure of the electron-electron correlations, the VQE algorithm is likely to require more circuit layers to achieve an accuracy that is in the same league as UCCSD for certain molecular problems [48].

In an effort to strike a balance between accuracy and cost of implementation, additional approaches of ansätze have been explored, namely those that try to construct building blocks in such a way that their parametrization is restricted to only relevant parts of the Hilbert space one attempts to explore. As a result, when studying systems where select physical quantities remain constant throughout the systems' dynamics, classes of ansätze that preserve these quantities are extremely useful. One such example in quantum chemistry is the *Quantum Number Preserving* (QNP) ansatz as one both reasonably implementable, as well as physically motivated [49]. The motivating insight of that class of ansatz, is that systems that have well-defined quantum numbers that are conserved such as the total particle number in electronic structure models or total spin magnetization in lattice models can use these ansätze to have dramatically reduced search spaces during optimization.

This is achieved through the use of local parametrized unitary gates that preserve the target symmetry, while remaining expressive enough to depict complex correlations within the allowed subspace. More specifically, these gate blocks, when tiled in a brick-layer structure, can represent arbitrary states given sufficient circuit depth. These gate blocks act on the qubits in a way that allows for entanglement between them to occur, but does not break the quantum number to be preserved. Owing to the structure of the ansatz, one can simply add more layers to the ansatz to construct a trial wavefunction with the required correlation while retaining a high degree of tunability, as each local gate block is parametrized. This avenue separates itself from completely unrestrained ansätze, where the trial wavefunctions that are constructed can be in non-physical parts of the Hilbert space, that are disallowed by the symmetries of the problem one is interested in. More information about this class of ansatz can be found in [49], as well as Sections 3.3 and 3.4.1.

2.1.2 Measurement Schemes

The process of mapping the second quantized Hamiltonian described in Eq. 14 into qubit operators yields $\mathcal{O}(N^4)$ Pauli terms to construct an estimator for the energy of the state. As part of the VQE algorithm, many such measurements of these terms is required, otherwise referred to as experiment repetitions or shots. These distinct shots, along with the variances incurred by these individual Pauli terms comprise the total cost in quantum resources when estimating the energy to the desired precision. Measurement schemes in VQE algorithms are centered around the task of mapping the initial problem of measuring a large set of Pauli terms that scale rapidly with system size, to one that is more in line with the current hardware available, by attempting to reign in circuit depth, measurement counts and statistical uncertainty introduced by the estimators.

Decomposition-based methods aim to decompose the Hamiltonian into a sum of observables $\hat{P}_k$ with coefficients c_k as $\hat{H} = \sum_k c_k \hat{P}_k$. As the number of Pauli terms increases rapidly for interesting problem sizes, a common approach is to measure sets of these terms jointly in the same shot in a shared basis [31]. By performing a basis rotation over the prepared wavefunction, a configuration of m different partitions $G_1, G_2, \ldots, G_m$ of Pauli terms can be obtained that are measurable simultaneously. To estimate the energy of the Hamiltonian, one then estimates each disjoint partition of Pauli terms and adds their contribution to the energy as

$$\langle\psi|\hat{H}|\psi\rangle = \sum_{i=1}^{m} \sum_{\hat{P}_k \in G_i} c_k \langle\psi|\hat{P}_k|\psi\rangle, \tag{35}$$

This procedure therefore requires only m distinct measurement configurations as well as enough measurement repetitions to reach the desired accuracy in energy estimation. It is worth noting that some freedom is allowed into the allocation of a given shot budget over these m distinct terms, as some terms contribute much less to the energy and have much better variance characteristics than others, requiring much fewer shots to measure them in a given precision and allowing the allocation of the extra shots to the rest of the Pauli terms [50]. Additional grouping strategies have been considered such as fermionic basis rotations [51] and graph coloring methods [52] to minimize the number of distinct partitions m needed, achieving significant reductions.

As the task of measuring one- and two-electron integrals becomes exceedingly hard as the number of terms grows, alternative schemes such as *single and double factorization* schemes, which reformulate the two-electron integrals into low-rank decompositions [53, 54, 55] have emerged. In single factorizations, the repulsion two-body electron integrals are decomposed approximately [53] into rank-1 objects, with double factorization approaches factorizing the rank-1 objects of single factorization even further [54, 55]. These approximations allows for a reduction in the effective terms needed to accurately represent the Hamiltonian. By allowing some truncation in the $N(N+1)/2$ original terms makes the solution approximate, but enables even further reduction in the overheads accrued in measurement.

Other novel approaches for measurement revolve around the use of *classical shadows* as estimators, it is possible to estimate $\mathcal{O}(N)$ observables using only $\mathcal{O}(\log(N))$ measurements [56]. In lieu of measuring Pauli terms or groups of Pauli terms directly, one can obtain a classical shadow representation ρ of the quantum state by applying a random unitary transformation U that transforms $\rho \to U\rho U^\dagger$ and classically storing the outcome in a bit-string $|b\rangle$.

$$\mathbb{E}\left[U^\dagger|b\rangle\langle b|U\right] = \mathcal{M}(\rho) \Rightarrow \rho_{\text{shadow}} = \mathbb{E}\left[\mathcal{M}^{-1}(U^\dagger|b\rangle\langle b|U)\right] \tag{36}$$

After gathering a multitude of randomized measurement outcomes M as shadows, it is possible to construct estimates for a plethora of observables $\hat{O}_l$ and through classical shadow theory there exist geometrical arguments for the statistical guarantees on accuracy [56], only relying a non-zero overlap of these observables $\hat{O}_l$ with the measurement basis to be able to be estimated with the random measurement taken.

A variety of different protocols using shadows has been developed recently for estimating the overlaps between shadow measurements and fermionic Gaussian states [57] using the group of Gaussian Unitaries. Lastly, the latest approaches utilize particle number symmetry to construct estimators for k-Reduced Density Matrices scaling with $\mathcal{O}(k^2\eta)$, where η is the number of particles in the fermionic wavefunction [58]. A common idiosyncrasy of these approaches is that in current hardware, the randomized measurements required by these protocols is very demanding as loading a new quantum circuit representing a random unitary to measure at every iteration takes orders of magnitude longer than taking repeated samples from a specific circuit, owing to the control electronics of the quantum computer [59].

2.1.3 Optimization of parametrized quantum circuit

The process of finding the optimal parameters of a parametrized quantum circuit to represent the desired energy state is the centerpiece of what constitutes VQE a hybrid quantum-classical algorithm, as it allows for the classical compute to alleviate some of the brunt of the optimization procedure. The choice of optimizer depends heavily on the hardware characteristics of the quantum device such as repetition rate, loading and cooling times, as well as the intricacies of the cost function one is trying to optimize, including but not limited, to whether the cost function is differentiable, auto-differentiable, its gradient and curvature in parameter space.

A significant source of error in these optimization routines that aim to find parameters $\boldsymbol{\theta} = (\theta_1, \theta_2, \ldots, \theta_n)$, where d is the number of real, tunable parameters in the quantum circuit such that $E(\boldsymbol{\theta}) = \langle\psi(\boldsymbol{\theta})|\hat{H}|\psi(\boldsymbol{\theta})\rangle$ is minimized comes from the samples themselves, as most quantum setups are prone to noise in over- or under-rotations of the input parameters or readout errors. These optimizer routines often rely on gradients of the cost function constructed for estimating the energy of $\hat{H}$, or higher order derivatives such as Hessians [60]. This also makes shot noise arising from the statistical nature of finite quantum measurements as well as the barren plateaus that cost functions exhibit structurally additional challenges to be considered, as some optimizers might be appropriate when simulating hardware on a classical computer, but not suitable when working with real hardware and inversely.

The nature of optimization procedures as well as the unique characteristics of each procedure, make the selection of the optimization routine a key factor in the convergence of Variational Quantum Algorithms. The following section contains a small choice of optimizers that are frequently used across hybrid quantum-classical optimization loops, denoting their basic principles as well as a few advantages and disadvantages related to each, as the is currently is no optimizer that constitutes a panacea for iteratively updating the parameters in quantum circuits.

2.2 Optimization algorithms in hybrid quantum-classical procedures

Variational quantum algorithms have been prominent candidates for application on quantum devices by leveraging classical optimizers to optimize quantum circuits, mitigating the need for extensive quantum resources and enhancing resilience against coherent noise, through their requirement for shorter quantum circuits and by providing protection against noise through the ability of optimizing away the systemic gate under/over rotations, meaning the gate does not perform exactly the operation it is expected to due to errors in calibration and noise. The classical optimization component of these variational quantum algorithms often favors gradient-free optimizers, often yielding better performance in terms of robustness against noise, but with the increasing complexity of VQA optimization problems, which are, in the worst case, NP-hard [61], more sophisticated approaches are required. Optimization algorithms that require gradients or higher order derivatives can have better performance, but may fall prey to the barren plateau problem caused by the exponential dimension of the Hilbert space when simulating with more and more qubits [62], as well as the resources required for calculating those higher order derivatives can be excessively costly.

The choice of optimization algorithm of the parametrized ansatz is as crucial as the choice of the rest of other components of the VQE procedure, as it is relying on a multitude of factors. Such

factors encompass the quantum computing platform used, its characteristics and the tolerance for noise as well as hyperparameter sensitivity of the optimization algorithm in reliably converging the cost function for a plethora of cases to avoid overfitting.

One crucial insight of quantum cost functions resides in the fact that the cost functions which are derived from quantum gates are trigonometric polynomials, where the number of eigenvalues of the generators of the quantum gates is equivalent to the order of these polynomials, thus paving the way for generalized parameter shift rules to be formulated [63]. Parameter shift rules of quantum gates has enabled the potential for gradient based algorithms to be used extensively when applicable, allowing for a more diverse ecosystem of algorithms to be at the arsenal of hybrid quantum-classical optimization procedures, as this allows the omittion finite-difference derivatives which are unreliable in NISQ devices due to noise, either from measurement shot-noise or miscalibration of the quantum circuit [64, 65, 49].

2.2.1 L-BFGS-B

The Limited-memory Broyden-Fletcher-Goldfarb-Shanno algorithm extended for Bound constraints (L-BFGS-B) [60, 66, 67] is a quasi-Newton optimization method, geared towards handling bound-constrained problems, meaning it is suited for finding the minimum of a function which is constrained to upper and lower bounds on its variables. This attribute of the L-BFGS-B algorithm The feature of using limited memory by storing vectors representing the inverse Hessian instead of the $n \times n$ approximation in memory, where n are the number of bounded plus free variables of the problem, makes it a popular choice in terms of efficiency, even for large scale problems and makes the L-BFGS-B algorithm a prime candidate for zero-noise optimizations of parametrized quantum circuits, as most quantum gates are periodic to some $k \in \mathbb{N}$ multiple of 2π.

Upon the k-th iteration and having stored in memory m updates, the algorithm assumes an estimate of the optimal value $\mathbf{x_0}$, where $\mathbf{x_0}$ is the set of initial optimal values parametrizing the cost function to be minimized by the algorithm, and their corresponding derivatives $\mathbf{g_k} := \nabla \mathbf{f}(\mathbf{x_k})$. These derivatives play an central role in the formation of the estimate for the Hessian matrix of $\mathbf{f}(\mathbf{x})$ [68]. The saved updates m for the positions and their gradients are stored as $\boldsymbol{s_k} = \boldsymbol{x_{k+1}} - \boldsymbol{x_k}$ and $\boldsymbol{y_k} = \boldsymbol{g_{k+1}} - \boldsymbol{g_k}$ respectively, with $\rho_k = (\boldsymbol{y_k^\top s_k})^{-1}$ being the initial construction of the inverse Hessian at the k-th iteration and $\top$ symbol for denoting the transpose of the $\boldsymbol{y_k}$ vector. The BFGS recursion for the inverse Hessian is implemented as:

$$H_{k+1} = (I - \rho_k \boldsymbol{s_k y_k^\top}) H_k (I - \rho_k \boldsymbol{y_k s_k^\top}) + \rho_k \boldsymbol{s_k s_k^\top}$$

With these tools in hand for the position updates s_k and gradient updates y_k, one can calculate the descent direction for the k-th iteration with $k > 0$, by defining $q = g_k$ and then performing the following recursive procedure [68]:

$q = \boldsymbol{g_k}$
for $i = k-1, k-2, \ldots, k-m$ **do**
  $\alpha_i = \rho_i \boldsymbol{s_i^\top} q$
  $q = q - \alpha_i \boldsymbol{y_i}$
end for
$\gamma_k = \frac{\boldsymbol{s_{k-1}^\top y_{k-1}}}{\boldsymbol{y_{k-1}^\top y_{k-1}}}$
$H_k^0 = \gamma_k I$
$z = H_k^0 q$
for $i = k-m, k-m+1, \ldots, k-1$ **do**
  $\beta = \rho_i \boldsymbol{y_i^\top} z$
  $z = z + s_i(\alpha_i - \beta_i)$
end for

This procedure yields the search direction for maximization problems, whereas for minimization problems, one can simply take $z = -z$.

The L-BFGS-B algorithm is widely adopted for its efficiency, reliability, frugality of resources required and its versatility, as it can be used in a plethora of different optimization problems. However, it is not without shortcomings, as it is prone to converging to local minima due to the vanishing gradients of the parameter landscape and its sensitivity to initial conditions make its application a non-trivial task. In most real-world applications and use cases, the L-BFGS-B algorithm would not be well-suited, as its reliance on gradient calculations would both pose an

exorbitant cost in resources needed for calculating the quantities dictated by the algorithm, as well as require a vast number of repetitions to ensure noise does not obfuscate the results of these gradients.

2.2.2 Rotosolve

When examining parametrized quantum circuits and their intricacies, it is possible to come up with more efficient methods for performing that specific optimization task. The Rotosolve optimization algorithm [69] leverages the fact that, given an objective function expressed by Hermitian operator M and the candidate solution to the optimization problem described by that objective function can approximated by a parametrized quantum circuit $\hat{U} = \hat{U}_D \dots \hat{U}_1$, most quantum circuits can be decomposed to either fixed gates such as a Controlled-Z or parametrized gates of the form $\hat{U}_d = exp(-i\frac{\theta_d}{2}H_d)$, with $\theta_d \in (-\pi, \pi]$ and H_d being Hermitian and Unitary matrices. For a particularly interesting case where the generator has eigenvalues ± 1, the following gives the optimal angle of each θ_d as a closed form expression:

$$\begin{aligned}\theta_d^* &= \underset{\theta_d}{\operatorname{argmin}} \langle H \rangle_{\theta_d} \\ &= -\frac{\pi}{2} - \arctan2\left(2\langle H \rangle_{\theta_d=0} - \langle H \rangle_{\theta_d=\frac{\pi}{2}} - \langle H \rangle_{\theta_d=-\frac{\pi}{2}}, \langle H \rangle_{\theta_d=\frac{\pi}{2}} - \langle H \rangle_{\theta_d=-\frac{\pi}{2}}\right)\end{aligned}$$

Genereralizations of these formulas can be found for any gate satisfying parameter shift rules [49]. One of the greatest advantages of the Rotosolve optimization algorithm is its tailored nature to finding the optimal angles in rotation gates within very few objective function evaluations, which can be predetermined at the start of each iteration of the optimization procedure, alleviating the cost of re-programming the control electronics of the quantum computer and submitting constantly new circuits [70]. Although Rotosolve performs exceptionally well in the absence of noise when evaluating the objective function, that stands to be its main limitation as it struggles to find optimal solutions in noisy parameter landscapes. Additionally, its reliance on having at least 3 evaluations for each free parameter in the optimization problem is a liability when dealing with ansätze that have a very large number of parameters in their parametrization.

These unique characteristics have sparked interest for further improvements of the Rotosolve optimizer, with notable results revolving around the commonly used Unitary Coupled Cluster (UCC) ansatz for simulating quantum chemistry problems, replacing the usual $\boldsymbol{\theta} = 0$ parametrization that would yield the Hartree-Fock state, with parameters generated by the second order Møller-Plesset (MP2) perturbation theory for initial parametrization of the UCC ansatz [71]. Moreover, an extension of Rotosolve is *RotoSelect* [69], which also performs optimization on the ansatz itself, in an effort to reduce the generators and rotations needed, while still being ansatz-agnostic like Rotosolve.

2.2.3 SPSA

The Simultaneous Perturbation Stochastic Approximation (SPSA) algorithm [72] is a stochastic gradient-free optimization algorithm that is regarded as well-suited for problems with noisy objective functions or when the gradient is difficult or expensive to compute. Unlike gradient-based optimization methods, SPSA estimates the gradient using a perturbation scheme, making it a notable candidate for the optimization task of parametrized quantum circuits with a high number of circuit parameters.

At each iteration, the current parameter vector is perturbed in two directions, and the objective function is evaluated at both perturbed points. By comparing the function values at these points, SPSA estimates the gradient direction with the help of gain sequences a_k and c_k, controlled by the hyperparameters a, c, A, α and γ [73]. These hyperparameters denote the gradient scaling parameter, the scaling parameter, the gradient stability constant, the gradient scaling exponent and the scaling exponent respectively. This estimated gradient is then used to update the parameter vector, moving it towards a region of lower objective function values. The perturbation scheme employed in SPSA ensures that the gradient estimate remains unbiased, even in the presence of noise.

Considerable effort has directed towards leveraging the inherent benefits of SPSA such as foregoing the need for explicit gradient measurements, with other techniques such as the Quantum Natural Gradient (QNG) [74], [75], requiring very few quantum circuits to estimate the gradient by constructing approximations of the Fubini matrix, or adaptive momentum for adjusting the effect of the calculated gradient [76].

While SPSA and the relevant variants of the algorithm offer several advantages, they also have limitations. One of its primary advantages is its classical computational efficiency, especially for large-scale problems, as by avoiding the explicit computation of the gradient, SPSA can significantly reduce computational costs. Additionally, SPSA is relatively robust to noise due to the random choice for each component of the perturbation vector. However, SPSA may converge slower than gradient-based methods for smooth, noiseless problems. Furthermore, the small perturbation required by the algorithm is also its main crux in current noisy hardware, as the choice of perturbation size can affect the algorithm's convergence rate and accuracy.

Lastly, the SPSA algorithm is highly sensitive to the selection of hyperparameters for the coefficients a, c, A, α and γ, making the process of finding hyperparameters that would perform well in a given problem a considerable challenge. The author proposes a method through which this critical process of finding good hyperparameters with which to perform the optimization of the parameters governing a quantum circuit can be made systematic, through careful consideration of the quantum resources available and their effect on the optimizer.

2.2.4 Surrogate models of cost functions for optimizers

In the ideal scenario of noiseless experimental conditions or in a FTQC regime, gradient-based optimizers like L-BFGS-B optimization algorithm described in Section 2.2.1 are able to consistently converge to optimal solutions. However, the introduction of experimental or sampling noise significantly complicates the optimization landscape, rendering many energy estimators unstable. To address this challenge, previous research has explored stochastic-based [77, 78] and model-based optimization approaches [11, 79]. In NISQ computing systems, it is of vital importance to tackle the task of evaluating the parametrized quantum circuit ansatz efficiently. As such, in an effort to utilize the concept of gradient descent without extracting numerical gradients from the quantum device, one can use surrogate models whose gradients can be computed more readily.

Surrogate models in hybrid quantum-classical optimization algorithms are approximations of the true objective function based on sampled quantum data. They provide a computationally inexpensive way to estimate the objective function's behavior, enabling more efficient optimization strategies. Common surrogate model choices are polynomial models of quadratic order or higher to fit the sampled data [79, 11, 59] and Gaussian processes [80] providing a probabilistic approach by defining a distribution over functions and their ability to quantify uncertainty.

In [79], [59] additional benefits of surrogate models pertaining to circuit batching and its effects on wall clock time are detailed. The ideas presented in [59] a step further, by performing a line search to find the optimal step in the conjugate gradient direction, while still retaining the advantages of utilizing previously sampled neighboring points to construct surrogate function models, as well as examining the potential use of the second order coefficients when fitting a quadratic function. The main surrogate model that will be examined in this thesis is that of a quadratic function when tasked with approximating a Hamiltonian near its ground state.

Preconditioning and its effects on optimization procedures

Using the quadratic model fitted to the samples taken by a quantum device using least squares linear regression fitting with polynomial features, one can use the gradient of that quadratic model as a guide for making a gradient descent algorithm. However, when fitting a quadratic model, one can maximize the knowledge gained by the second-order terms of the quadratic model, accounting for the curvature of the cost function landscape which often contains barren plateaus. Therefore, the Hessian matrix H can be used as a preconditioner matrix M to the standard gradient descent process as:

$$\boldsymbol{\theta}_{k+1} = \boldsymbol{\theta} - \alpha M^{-1} \boldsymbol{\nabla} C(\boldsymbol{\theta}_k) \tag{37}$$

where $\boldsymbol{\theta}_i$ is the point around which the cost function $C(\boldsymbol{\theta})$ is approximated with the help of a quadratic model and α the step size, with $\alpha \in \mathbb{R}$, which is typically taken to be equal to the first order gradient.

A benefit of using the Hessian as a preconditioner matrix is the improvement of the convergence rate of the algorithm by reducing the condition number of the matrix it is transforming [81]. Another consequence of transforming the system is the clustering of its eigenvalues more tightly, improving the convergence rate of iterative methods. The use of Hessians as preconditioners for gradients incurs additional considerations when designing an optimization algorithm, such as the computational overhead of the Hessians construction, storing and application to the gradient vectors, with some algorithms such as the L-BFGS-B algorithm described in Section 2.2.1 employing intricate strategies of circumventing those challenges.

2.2.5 Conjugate Model Gradient Descent algorithms

The Conjugate Gradient Descent (CGD) algorithm is an iterative method for solving large-scale optimization problems [82], particularly those involving large, sparse, symmetric positive-definite matrices. Originally developed for solving linear systems of equations, the algorithm has been extended to tackle nonlinear optimization problems as well. It is well equipped for dealing with high-dimensional spaces where direct methods, such as direct diagonalization of the Hamiltonian operator $\hat{H}$ to find its eigenvalues, become computationally prohibitive. As this optimization problem of finding ground states of Hamiltonian operators describing quantum systems in quantum chemistry can be approximated by a quadratic function, one can take advantage of the CGD optimization algorithm to construct an optimizer that is tailored to quantum circuits simulating

such Hamiltonian operators through the use of surrogate models. At its core, the CGD algorithm seeks to minimize a quadratic function of the form:

$$f(\boldsymbol{x}) = \frac{1}{2}\boldsymbol{x}^\top A\boldsymbol{x} - \boldsymbol{b}^\top \boldsymbol{x} + c$$

where A is a symmetric positive-definite matrix, $\boldsymbol{b}$ is a vector and c is a scalar constant. The method generates a sequence of approximations x_k that converge to the true solution $\boldsymbol{x}^*$ by moving along conjugate directions. It is this property of the constructed search directions directions that ensures that each new search is orthogonal with respect to the symmetric positive-definite matrix A that yields improved convergence efficiency.

Armed with an initial guess $\boldsymbol{x_0}$ and the initial residual $\boldsymbol{r_0} = \boldsymbol{b} - A\boldsymbol{x_0}$, the initial search direction $\boldsymbol{d_0} = \boldsymbol{r_0}$, and k_{max} the maximum number of iterations allotted to the optimization algorithm, an overview of the CGD algorithm steps are as follow [83]:

for $k = 0, 1, \ldots, k_{max}$ **do**
  Compute step size: $a_k = \frac{\boldsymbol{r_k}^\top \boldsymbol{r_k}}{\boldsymbol{d_k}^\top A \boldsymbol{d_k}}$
  Update the solution vector: $\boldsymbol{x}_{k+1} = \boldsymbol{x}_k + a_k \boldsymbol{d}_k$
  Compute the new residual: $\boldsymbol{r}_{k+1} = \boldsymbol{r}_k - a_k A \boldsymbol{d}_k$
  Compute the conjugate direction coefficient: $\beta_{k+1} = \frac{\boldsymbol{r}_{k+1}^\top \boldsymbol{r}_{k+1}}{\boldsymbol{r}_k^\top \boldsymbol{r}_k}$
  Update the direction: $\boldsymbol{d}_{k+1} = \boldsymbol{r}_{k+1} + \beta_k \boldsymbol{d}_k$
  Check convergence criteria; else increment k by 1.
end for

It is worth noting that the conjugate direction coefficient β has multiple ways of being calculated [84, 85, 83, 86], whose impact is situational on the properties of the cost function to be minimized.

This sets up the stage for utilizing the work on surrogate models for approximating the cost function, making the most out of the samples for the expectation value taken by the quantum device. By combining the use of a surrogate model cost function fitted to the samples taken as a proxy for the true cost function landscape, the conjugate gradients offer a plethora of advantages, such as performance guarantees and known error bounds from the target vector per iteration, compared to solely relying on the first-order coefficients extracted from the fitted model [79], culminating to the creation of the Conjugate Model Gradient (CMG) algorithm by the author [59], as outlined in Section 3.3.

Performance guarantees of Conjugate Gradient Descent

The use of conjugate gradients as directions of descent for optimization of parametrized quantum circuits stems from the performance guarantees that algorithm can provide, assuming the matrix A of dimension $n \times n$ is symmetric positive-definite and the estimator function f is convex. More specifically, performance guarantees about the **convergence rate** of the method where the method is proven to converge at most n iterations in exact arithmetic but convergence in floating-point arithmetic where physically meaningful convergence criteria are introduced, convergence is typically achieve in far fewer steps [87].

Claims about the **error bounds** after k iterations can also be made, such that the error in the norm of A after k iterations satisfies [83, 87]:

$$\|\boldsymbol{x}_k - \boldsymbol{x}^*\|_A \leq 2\left(\frac{\sqrt{\kappa}-1}{\sqrt{\kappa}+1}\right)^k \|\boldsymbol{x}_0 - \boldsymbol{x}^*\|_A \tag{38}$$

where $\kappa = \frac{\lambda_{max}}{\lambda_{min}}$ is the condition number of matrix A and $\lambda_{max}, \lambda_{min}$ are the maximum and minimum eigenvalues of A. This highlights further the fact that the CGD method converges faster when the condition number κ is lower, meaning the matrix A is better conditioned. In scenarios where this is not the case, preconditioning techniques can be employed to ameliorate the convergence rate, but not with guaranteed results [88].

Conjugate Model Search

In an iterative optimization algorithm like Conjugate Model Gradient optimization, the step size plays a critical role in the convergence speed of the optimization procedure, its stability as proper step size selection can help in the prevention of overshooting minima in the optimization landscape or slow down the convergence if the step sizes taken are too small. Motivated by the good performance characteristics of L-BFGS-B, the choice in step size for a Conjugate Model Gradient algorithm can be adjusted through the same exact or inexact line search methods used in the line search optimization of the L-BFGS-B algorithm such as Wolfe or backtracking line search [89, 90].

In Alg. 1, an algorithm is outlined which helps tackle these issues via line searching for the appropriate step size along the descent direction dictated by the Conjugate Gradient method, following the same convention of [79] for reparametrizing the sample number k as $k = \eta(m+1)(m+2)/2$, where m is the number of parameters to be optimized. The benefits of improved robustness when the surrogate model constructed by the quantum samples taken from the quantum device does not depict the cost function landscape precisely do not come without drawbacks, such as increased quantum resources as the region that needs to be explored by the line search optimization might be outside of the previously explored and sampled parameter space. This, in practice, can mean an increase in batches of quantum circuit executions, which is often one of the most demanding resources with respect to wall-time when running hybrid quantum-classical algorithms, as will be discussed in greater detail in Section 2.2.6. The way that the optimal step size α_m for these optimization algorithms is decided, when f is the oracle function and $\boldsymbol{V}$ is the gradient (either given or constructed) is described by:

$$\alpha_m = \arg\min_{\alpha} f(\boldsymbol{x} + \alpha \boldsymbol{V}) \tag{39}$$

Another important aspect of optimization algorithms is their sensitivity to the selection of hyperparameters which have a direct impact on their performance, be it in the form of convergence rate or accuracy when converged. By permitting more freedom in the choice of the step size hyperparameter and allowing the optimization procedure to find an optimal step size along the conjugate gradient descent direction, Conjugate Model Search (CMS) can self-regulate it and achieve more consistent results.

Preconditioned Conjugate Model Search

In Alg. 2, the algorithm for an optimization procedure is presented where the second order information of the quadratic model fitted in one of the steps of the Conjugate Model Search algorithm is utilized as a preconditioner for faster convergence of the optimizer. This is done by using the coefficients of the second order polynomial terms of the quadratic model in Alg. 1 to construct a matrix that approximates the Hessian of the gradient via the process detailed in Eq. 37.

Since this poses no quantum overhead to the optimization algorithm, as it does not require more samples from the quantum device, it is convenient to use preconditioning of the gradient by using the approximate Hessian matrix constructed from the coefficients of the surrogate model that was fitted during the sampling stage. As a consequence, instead of relying on second order derivative observables, in cases where the cost function can be locally approximated by a quadratic form, the Preconditioned CMS optimization routine can offer a compromise between computational tractability and search directions informed through higher-order derivatives of the cost function.

This method enables the CMS algorithm to have more curvature awareness of the cost function, and accelerating convergence in some cases. In practice, it is a common strategy to adopt schemes for controlling the updates of the approximate Hessian prior to applying it to the gradient, as it can accumulate errors over iterations of the routine, such as the periodic recalculation of the approximate Hessian after a fixed number of updates or enforcing a trust region for its condition number.

However, it is worth noting that after extensive numerical experimentation from the author, this approach does not seem to yield reliable numerically stable convergence for larger problems. The most likely culprit of those instabilities seems to come from the approximate Hessian matrix used, as some second order coefficients tend to be very small, leading to a very large condition number κ for the matrix, as the condition number for the approximate Hessian often ended up outside the bounds of trust regions.

Algorithm 1 Conjugate Model Search

Input: Initial point $\boldsymbol{x_0}$ of dimension N, learning rate γ, sample radius δ, maximum iterations n, number of new samples η, rate decay exponent α, stability constant A, sample radius decay exponent ξ, tolerance ε, oracle for function f.

Initialize lists L, L′
Initialize a list G
Initialize a list H
Let $k \leftarrow \eta(N+1)(N+2)/2$
Let $\boldsymbol{x} \leftarrow \boldsymbol{x_0}$
for m in 0...n **do**
    Let $\delta' \leftarrow \delta/(m+1)^{\xi}$
    Sample k points uniformly at random from the δ'-neighborhood of x to generate a set S
    for each x' in $S \cup \{\boldsymbol{x}\}$ **do**
        Add $(x', f(x'))$ to L
    end for
    Clear list L'
    for each tuple (x', y') in L **do**
        if $|x' - x| < \delta$ **then**
            Add (x', y') in L'
        end if
    end for
    Fit $f(x) = \boldsymbol{x}^T A\boldsymbol{x} + b\boldsymbol{x} + c \sim y$ to the points (x', y') in L' using least squares linear regression.
    Let $\boldsymbol{g_m}$ be the gradient of f at $\boldsymbol{x}$ (i.e. $\boldsymbol{g_m} = b$).
    if $|\boldsymbol{g_m}| < \varepsilon$ **then**
        return $\boldsymbol{x}$
    end if
    if $m = 0$ **then**
        Let $h_0 \leftarrow \boldsymbol{g_0}$
    else
        $\beta_m \leftarrow \boldsymbol{g_m^T g_m / g_{m-1}^T g_{m-1}}$
        $h_m \leftarrow g_m + \beta_m h_{m-1}$
    end if
    $\gamma' = \gamma/(m+1+A)^{\alpha}$
    Let $\boldsymbol{V} \leftarrow -h_m * \gamma'$
    Perform line search around $\boldsymbol{x}$ in the direction of $\boldsymbol{V}$, obtaining optimal step size α
    Add $\boldsymbol{g_m}$ to the list G
    Add h_m to the list H
    Let $\boldsymbol{x} \leftarrow \boldsymbol{x} - \alpha \cdot \gamma' \cdot h_m$
    Let $m \leftarrow m + 1$
end for
return $\boldsymbol{x}$

Algorithm 2 Preconditioned Conjugate Model Search

Input: Initial point $\boldsymbol{x_0}$ of dimension N, learning rate γ, sample radius δ, maximum iterations n, number of new samples η, rate decay exponent α, stability constant A, sample radius decay exponent ξ, tolerance ε, oracle for function f.

Initialize lists L, L′
Initialize a list D
Initialize a list H
Let $k \leftarrow \eta(N+1)(N+2)/2$
Let $\boldsymbol{x} \leftarrow \boldsymbol{x_0}$
for m in 0...n **do**
    Let $\delta' \leftarrow \delta/(m+1)^{\xi}$
    Sample k points uniformly at random from the δ'-neighborhood of $\boldsymbol{x}$ to generate a set S
    for each x' in $S \cup \{\boldsymbol{x}\}$ **do**
        Add $(x', f(x'))$ to L
    end for
    Clear list L'
    for each tuple (x', y') in L **do**
        if $|x' - x| < \delta$ **then**
            Add (x', y') in L'
        end if
    end for
    Fit $f(x) = \boldsymbol{x}^T A\boldsymbol{x} + b\boldsymbol{x} + c \sim y$ to the points (x, y) in L' using least squares linear regression.
    Let p be the cardinality of the set of parameters representing point $\boldsymbol{x}$
    Let M be an $p \times p$ matrix, containing the second order coefficients of the least squares linear regression.
    Let $\boldsymbol{g_m}$ be the gradient of f at x (i.e. $\boldsymbol{g_m} = b$).
    Let $\boldsymbol{d_m} \leftarrow M^{-1}\boldsymbol{g_m}$ be the preconditioned gradient
    if $|\boldsymbol{d_m}| < \varepsilon$ **then**
        return $\boldsymbol{x}$
    end if
    if $m = 0$ **then**
        Let $h_0 \leftarrow \boldsymbol{d_0}$
    else
        $\beta_m \leftarrow \boldsymbol{d_m^T d_m / d_{m-1}^T d_{m-1}}$
        $h_m \leftarrow \boldsymbol{d_m} + \beta_m h_{m-1}$
    end if
    $\gamma' = \gamma/(m+1+A)^{\alpha}$
    Let $\boldsymbol{V} \leftarrow -h_m * \gamma'$
    Perform line search around $\boldsymbol{x}$ in the direction of $\boldsymbol{V}$, obtaining optimal step size α
    Add $\boldsymbol{d_m}$ to the list D
    Add h_m to the list H
    Let $\boldsymbol{x} \leftarrow \boldsymbol{x} - \alpha \cdot \gamma' \cdot h_m$
    Let $m \leftarrow m + 1$
end for
return $\boldsymbol{x}$

2.2.6 Wall-time cost models for quantum computing devices

Computational time taken by an algorithmic procedure is of vital importance, as these times have to be constrained within the decoherence time of the quantum computer they are employing. For NISQ devices, taking into account the computational time is an absolute necessity, because as powerful error mitigation techniques might be, they cannot make up for the degradation of the quality of results taken past the decoherence time of the quantum device [91, 92].

When factoring into consideration the wall-time constraint imposed by the hardware of quantum devices, the need of measuring the performance characteristics of each optimizing algorithm and their potential for use in hybrid-quantum optimization becomes imperative. Through the use of time cost models, it is possible to make resource estimates for various problem sizes as well as different quantum devices that come with their own innate characteristics, as they vary wildly across different devices and architectures and play a central role in the feasibility and success of a variational optimization approach in NISQ devices.

Wall-time cost model of a superconducting qubit device

Most superconducting qubit experimental arrays are controlled via digital electronics that program the sequence of pulses that drive the superconducting chips to perform operations. The programming of those Field Programmable Gate Arrays (FPGAs) is a very demanding aspect of quantum hardware engineering but its esoteric knowledge is outside the scope of this thesis. What is important to this thesis is the knowledge that in comparison to taking repeated shots of the same circuit with the same circuit parameters, loading a new circuit layout, be it through different circuit parameters at a given circuit or a different circuit layout altogether, are changes that carry significant impact to the performance of the quantum device.

Moreover, it is important to note that the way those changes are submitted to the quantum device matters, as requesting a completely unrelated circuit layout from the quantum devices differs orders of magnitude in terms of wall-clock time, rather than swapping from one circuit to another that when those have sent as a batch of circuits to the FPGA [59].

In this thesis, the wall-clock time model for a Google Sycamore chip that is described in [59] is used as a measure of how long the optimization procedure takes when accounting for the parameters that extend beyond just counting epoch iterations of the respective algorithms. More specifically, the parameters for characterizing optimizer performance are the number of batches of circuits that are sent to the controlling device of the quantum computer, the number of distinct circuits one needs to execute in order to acquire the expectation values needed, as well as the number of the experiment repetitions, denoted below as a, b and c respectively:

$$t_{\text{wall}} = a * 1s + b * 0.042s + c * 5 * 10^{-5}s. \tag{40}$$

Wall-time cost model of a trapped atomic ions qubit device

Another popular platform for quantum computing is trapped atomic ions, which typically exhibit encouraging qubit counts and gate operation fidelities over their superconducting counterparts, but current hardware allows for experiment repetition rates that are dramatically slower [93] than superconduting qubits. In this thesis, in collaboration with the Johannes Gutenberg Universität Mainz, the author performed benchmarking for a time-cost of their quantum computer setup. A simple circuit comprising of layers of a parametrized single qubit Y rotation (RY) gate and a parametrized IsingZZ, denoted I_{ZZ}, gate per qubit was used to measure the impact of different shots, qubit counts as well as layerings of the same circuit to probe the effects of circuit depth. A graphical representation of such a circuit for 4 qubits and 2 layers of such an array of gates is depicted in Fig. 1.

Circuits of that repeated structure can be used to probe the behavior of devices when crucial quantities such as circuit depth and the number of qubits in the circuit, as well as the shots of a given circuit. The gates used for these circuits have been chosen from the native gate set supported by the device, to yield a best-case scenario of the performance, from a calibration perspective. This resulted in the data of Fig 2 to be generated, when averaged over 10 experimental runs.

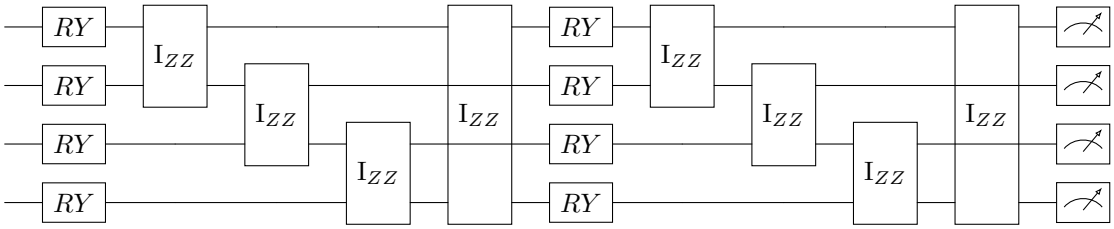


Figure 1: Benchmarking circuit layering of RY gates and IsingZZ gates for investigation of the Johannes Gutenberg Universität Mainz quantum device behavior when submitted with circuits of different size and depth, as well as the effect of the variation of shot repetitions on wall-time performance.

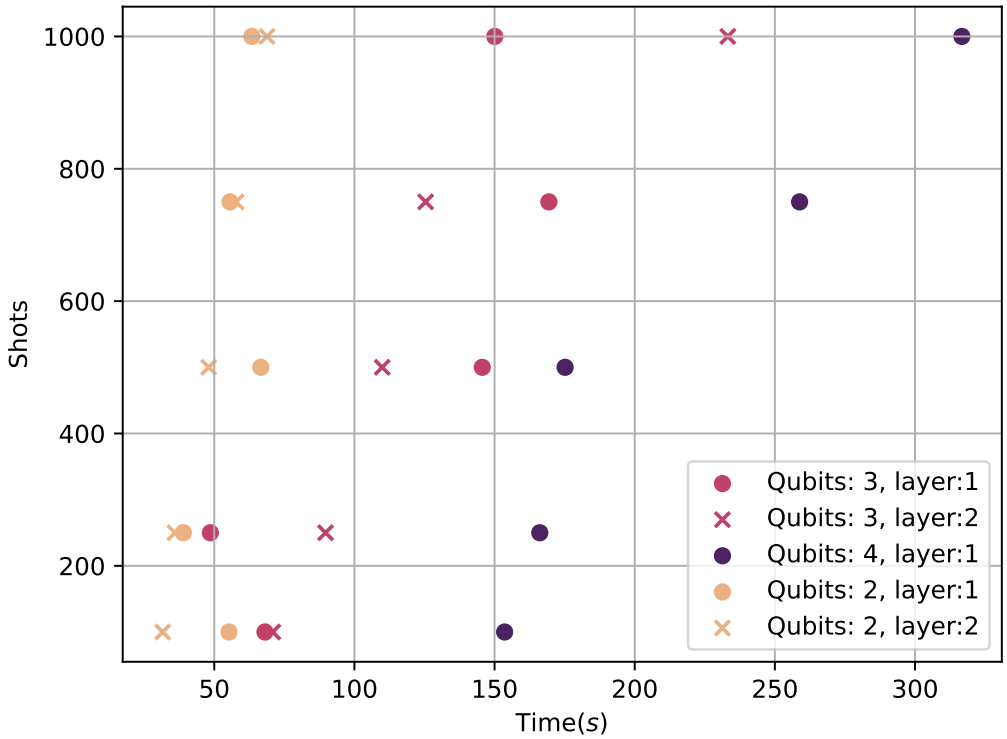


Figure 2: Wall time in seconds for executing a circuit with structure described by Fig.1 for a given number of shots.

Based on the dataset of benchmarking wall-time required by the quantum device shown in Fig. 2, a wall-time cost model that would describe the devices behavior is expected to display the effect of adding more experimental repetitions on wall-time and the toll of ion shuttling should be evident both via increasing the circuit size, thus requiring more operations per experimental repetition from the quantum hardware and even more prominently via the increase in number of qubits, as that dramatically increases the number of shuttling operations required to implement the given circuit.

Owing to the intricacies of each quantum device, there is a plethora of factors influencing the choice of optimization algorithm when running hybrid quantum-classical algorithms with the help of those devices. Due to that, the subsequent choice of hyperparameters and therefore, the performance of each optimization algorithm play a paramount role in the success and feasibility of such algorithms. In Section 2.3, a systematic way for finding suitable hyperparameters is proposed, based on the problem and the quantum device at hand.

2.3 Genetic hyperparameter tuning

As outlined in Section 2.2.6, the choice of appropriate hyperparameters plays an instrumental role in hybrid quantum-classical procedures. As such, the careful selection of hyperparameters is paramount in an optimizers' performance. These parameters, which govern an algorithm's behavior, can significantly influence convergence speed, solution quality, generalizability meaning to be readily applicable to a variety of optimization tasks, and computational efficiency. An optimizer with well-tuned hyperparameters can exhibit accelerated convergence, improved solutions quality and optimized resource utilization. Techniques like grid search, random search, Bayesian optimization, and evolutionary algorithms can be used to identify optimal hyperparameter configurations effectively, leading to superior outcomes in various optimization tasks compared to using them without any regard to the selection of their hyperparameters. To that end, a systematic way of comparing the performance of different sets of hyperparameters for a given optimizers, as well as the performance of various optimizers for a given problem, the author proposed a methodology for tackling that objective with a central point being the assignment of performance metrics to optimization paths.

2.3.1 Performance metric of a set of hyperparameters for a given optimizer and problem

The introduction of a performance metric of the hyperparameter set is necessary in the effort of evaluating how suitable they are for the optimization task they are trying to perform. The author of this thesis proposes the following formula as a **performance metric** p as a numeric criterion of hyperparameter set performance, with t being the wall-time required for the completion of the optimization algorithm, c is the cost function value that the optimization procedure returned, o is the optimal value of the optimization procedure of the cost function could possibly converge, while a and b are constants for influencing the balance between speed and distance of the cost function value returned by optimization procedure from optimal cost respectively:

$$p = at + \mathrm{H}(c)e^{b(c-(o+|\varepsilon|))} \tag{41}$$

where $\mathrm{H}(c)$ denotes a step function to act as an activator or a nullifier, as an optimization might converge numerically below or above the optimization target, with

$$\mathrm{H}(c) = \begin{cases} 1 & c-(o+|\epsilon|) \geq 0 \\ 0 & c-(o+|\epsilon|) < 0 \end{cases} \tag{42}$$

This formulation encapsulates both the hardware intricacies of the quantum device used in the hybrid algorithm, as well as allowing for tuning based on the resolution one needs for that particular optimization. For example one might not need that many experiment repetitions (and therefore wall-time) for a given accuracy of converging to the solution, as is the case with Hamiltonians with a ladder-like spectrum, like the Hubbard-type Hamiltonian described in Section 2.3.4. In Section 2.3.5 it is shown that hyperparameters that are acquired through this route are able to scale onto larger problems and perform better than their out-of-the-box hyperparameter counterparts, providing a path for a reasonable balance between hyperparameter optimization and computational wall-time.

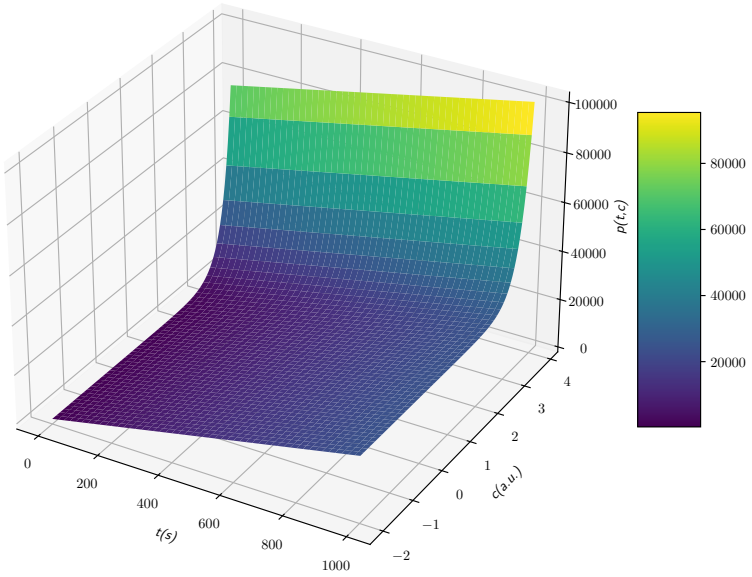


Figure 3: Figure showing the behavior of the performance metric function described in Eq. 41 for $a = 25$, $b = 3$ and $\epsilon = 0.1$ with $o = 0$ for simplicity. This graph shows the behavior that the performance metric function penalizes any increase in wall-time, but prioritizes more heavily the accuracy to a given tolerance ϵ.

2.3.2 Genetic Hyperparameter Tuning optimization algorithm

Following, we present the algorithm for generating and tuning hyperparameters of optimizers to improve towards a given performance metric which is calculated by equation 41 for a specified problem.

Continuing, it is necessary to choose a systematic way with which each generation of sets of hyperparameters is created, as different hyperparameters might have various intuitive ways of being optimized due to what they are influencing in the optimization algorithm, such as the number of shots used in an expectation value estimation or the number of different samples needed to construct a surrogate model for the cost function as shows in Section 2.2.4. A selection of different possible strategies for generating sets of hyperparameters is the following:

- Linear scaling: Generally used for hyperparameters with predictable desired behavior, such as the number of repetition shots in a quantum circuit where striving for lower is desirable.
- Uniform random selection from an interval: Useful for hyperparameters which have values in a known, closed interval.
- Random selection from normal distribution: Most commonly used for hyperparameters which one already has a good estimate of what a suitable value is, but with the help of standard deviation, one can skew the selection of the value.
- Genetic mutation: Taking reference from genetics, select a size rescaling mutation m_s, and set a mutation rate m_r, with $m_r \in [0, 1]$. Then, select a random number r, with $r \in [0, 1]$, if r is above the mutation rate m_r, keep the generated hyperparameter value $m_s h_p$, else keep the original hyperparameter value h_p.
- Physical model of randomness: Can be any source of physical randomness such as atmospheric noise, physically fair dice, etc.

Upon choosing a performance metric formula and a generation strategy for each hyperparameter, as well as guidance for each generation strategy, an algorithmic procedure can be executed to find sets of hyperparameters that minimize the performance metric suggested in Eq. 41. The cornerstone of that procedure is the iterative nature through which the sets of hyperparameters are assigned a performance metric based on the constraints set besides the initial problem and the quantum device, such as the accuracy one needs to reach in evaluating the cost function. This allows for hyperparameters that have major impact on the performance metric to steer the selection of hyperparameter sets accordingly. Moreover, it is important to average over a number of independent optimization runs and to take into account the variance of these runs as:

$$p_{averaged} = \sqrt{\bar{p}^2 + \tilde{p}} \tag{43}$$

where $\bar{p}$ denotes the mean over the performance metrics obtained in the Genetic Hyperparameter procedure, and $\tilde{p}$ denotes the variance of these performance metrics. The optimization path an optimization algorithm takes through parameter space is rarely deterministic, so such measures are needed to mitigate any outliers due to the stochasticity of the procedure and arrive consistently to reasonable results, and by extension, the acceptance of a new set of hyperparameters as the best performing.

Algorithm 3 Genetic Hyperparameter Tuning Optimization

Input: Initial set of hyperparameters x_0, target energy t, cost tolerance ε, number of averaging runs n, set of guidance rules G containing a mutation rule g_i for each hyperparameter i in set x_0, number of generations M

Initialize a list b representing the best performing hyperparameters found
for i in $0 \dots n$ **do**
    Perform optimization with given hyperparameters x_0
    Assign performance metric p_i to the set of hyperparameters to each optimization run
end for
Average performance metrics p_i over n optimization runs to get averaged performance metric p_0
Let $x \leftarrow x_0$
for m in $0 \dots$ M **do**
    Apply guidance rule g_i to each hyperparameter i in x, to get a set of transformed hyperparameters x_m
    for r in $0 \dots$ n **do**
        Perform optimization with given hyperparameters x_m
        Assign performance metric p_r to the set of hyperparameters to each optimization run
    end for
    Average performance metrics p_r over n optimization runs to get averaged performance metric p
    if $p \leq b$ **then**
        $b \leftarrow x_m$
    end if
end for
return b

2.3.3 Unitary paired-Coupled Cluster Doubles Ansatz

Careful consideration is needed regarding the use of the limited quantum resources provided by NISQ devices, with sources of inspiration being available in fields relevant to the computation one wants to perform on these devices. The Unitary Paired Coupled Cluster Doubles (UpCCD) ansatz draws intuition from the pair-electron (seniority-zero) approximation, in which all molecular orbitals are treated as doubly occupied or completely unoccupied. The benefit of this approximation arises from the restriction of the Hilbert space that the Hamiltonian spans and therefore reducing the computational complexity in comparison to unrestricted coupled cluster methods.

Even with the approximations employed by the UpCCD ansatz, it is expressive enough to replicate results of the Doubly Occupied Configuration Interaction (DOCI), which constitutes an exact solution for the electronic structure problem within the zero-seniority subspace. The effectiveness of UpCCD has been demonstrated to be very high in systems that are dominated by static correlation, such as bond-breaking processes, as opposed to systems that rely more heavily on dynamic correlation, with correlations beyond the zero-seniority approximation proving useful for the description of short range electron-electron interactions.

The main appeal of this approximation comes fom the insight that most chemical systems exhibit mostly paired-electron configurations, especially in their ground states [9]. This allows for the trade-off between accuracy stemming from dynamic electron correlations compared to more general approaches like Full CI and the ability to use more detailed basis sets for the description of the orbitals or higher degree of discretization afforded by halving the number of qubits needed to represent the system, thus improving the overall performance of the simulation of the electronic structure dynamics.

In this paired-electron approximation, the usual second quantized electronic structure Hamiltonian of Eq. 14 takes the form of the following restricted Hamiltonian:

$$\hat{H}_r = C + \sum_{p,q} u_{pq} \hat{b}_p^\dagger \hat{b}_q + \sum_{p \neq q} w_{pq} \hat{b}_p^\dagger \hat{b}_p \hat{b}_q^\dagger \hat{b}_q \tag{44}$$

where C is a constant energy offset, with the Hamiltonian matrix elements of u_{pq} and w_{pq} being related to the one- and two-body electron integral coefficients h_{pq} and V_{pqrs} of the second quantized electronic structure Hamiltonian and $\hat{b}_p$ represents the electron-pair annihilation operator in mode p. Owing to the electron pair symmetry, these coefficient mappings are greatly simplified, with the indices $\{p, q, r, s\}$ spanning 1 to $2N$ spin orbitals and even/odd indices denoting alpha/beta electrons and the indices $i, j \in [1, N]$ molecular orbitals, as [9]:

$$\begin{aligned} u_{ii} &= 2h_{ii} + V_{iiii} \\ u_{ij} &= V_{iijj} \quad (i \neq j) \\ w_{ij} &= 2V_{ijji} - V_{ijij} \quad (i \neq j) \end{aligned} \tag{45}$$

Formally, the UpCCD ansatz is derived from the coupled cluster operator by restricting it to only include pair excitations between doubly occupied orbitals. A general formula for the unitary operator in the UpCCD ansatz in first-order Trotterization [46] is given by the approximation [9]:

$$\hat{U} = \exp\left(\sum_{pq} \frac{t_{pq}}{2} (X_p Y_q - Y_p X_q)\right) \tag{46}$$

$$\approx \prod_{pq} \exp(\frac{t_{pq}}{2} (X_p Y_q - Y_p X_q)) \tag{47}$$

where t_{pq} are CCSD amplitudes that can be recovered with the help of software packages like PySCF [94, 95]. The circuit that is needed to represent such an operator on a quantum device has a gate depth that scales linearly with the number of orbitals $\mathcal{O}(N)$, as opposed to the often prohibitive $\mathcal{O}(N^4)$ scaling of usual Unitary Coupled Cluster methods [9].

2.3.4 Examples of Hamiltonians used in Variational Quantum Algorithms

This section of the thesis covers Hamiltonians of systems whose energy landscapes are non trivial, but their simulation for the purpose of finding their ground state lies within the grasp of capabilities of currently available quantum hardware.

Conrotatory ring opening of cyclobutane

The ring-opening reaction of cyclobutene as a process for the production of 1,3-butadiene, commonly known as *electrocyclic* reactions, that are described by the Woodward-Hoffman rules [96]. This type of ring opening reaction involves the conrotatory motion of electrons in the four-member ring of carbons, yielding the severing of a single bond and the formation of two new double bonds between carbon atoms. These rules favor the contoratory motion for four-carbon systems like cyclobutene with the addition of energy in the form of heat, making the production of 1,3-butadiene the lowest-energy pathway for this reaction, while preserving orbital symmetry.

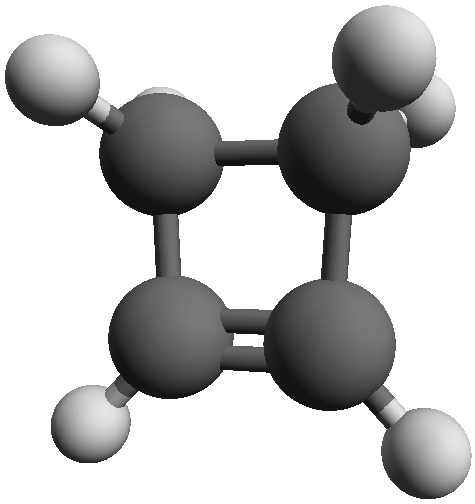


Figure 4: A ball-and-stick depiction of cyclobutene.

The Hamiltonian describing the system is the second quantized electronic structure Hamiltonian:

$$\hat{\mathcal{H}}_{elec} = \sum_{pq} h_{pq} \hat{a}_p^\dagger \hat{a}_q + \frac{1}{2} \sum_{pqrs} V_{pqrs} \hat{a}_p^\dagger \hat{a}_q^\dagger \hat{a}_r \hat{a}_s \tag{48}$$

where h_{pq} are the one-electron integrals of kinetic and potential energy terms, V_{pqrs} denote the two-electron integrals of electron-electron repulsions with $\hat{a}^\dagger$ and $\hat{a}$ describing fermionic creation and annihilation operators.

Through the Jordan-Wigner transformation, the creation and annihilation operators are mapped to qubit operations as:

$$\hat{a}_p^\dagger \mapsto \frac{1}{2}(X - iY) \quad \text{and} \quad \hat{a}_q \mapsto \frac{1}{2}(X + iY) \tag{49}$$

When restricted to a seniority-zero projection, meaning that all electron pairs remain spin-paired throughout the calculation, the product terms that arise are only of the nature of XX, YY, ZZ or Z [59]. This, along with the appropriate active space selection, helps significantly when considering the use of the UpCCD ansatz to capture the electron correlation described by this electronic structure Hamiltonian, owing to the ability of jointly measuring a lot of the terms of the Hamiltonian and by extension the number of batches of circuits required to measure the energy as described in Section 3.3. This makes the introduction of pair operators that obey the same $SU(2)$ algebra as the fermionic creation and annihilation operators worthwhile for giving a compact representation:

$$\begin{gathered} \hat{P}_p^\dagger = \hat{a}_{p\alpha}^\dagger \hat{a}_{p\beta}^\dagger, \qquad \hat{N} = \sum_{p,\sigma} \hat{a}_{p\sigma}^\dagger \hat{a}_{p\sigma} \\ [\hat{P}_p, \hat{P}_q^\dagger] = (1 - \hat{N}_p)\delta_{pq}, \quad [\hat{N}_p, \hat{P}_q] = -2\hat{P}_q \delta_{pq} \end{gathered} \tag{50}$$

with p, q and α, β denoting orbital and spin indices respectively, σ denoting spin (spin-up or spin-down) and δ_{ij} the Kronechker delta function [59]. After applying the projectors $\hat{P}_{S_0}$ to project the $\hat{\mathcal{H}}_{\text{elec}}$ to the seniority-zero subspace, the second quantized electronic structure Hamiltonian becomes [59]:

$$\begin{aligned}\hat{\mathcal{H}}_{S_0} &= \hat{P}_{S_0}\hat{\mathcal{H}}_{\text{elec}}\hat{P}_{S_0} \\ &= \sum_p h_{pp}\hat{N}_p + \frac{1}{4}\sum_{p\neq q}(2V_{pqpq} - V_{pqqp})\hat{N}_p\hat{N}_q + \sum_{pq} V_{ppqq}\hat{P}_p^\dagger\hat{P}_q\end{aligned} \tag{51}$$

These types of projections are very useful as they massively reduce the Hilbert space the Hamiltonian spans, making the optimization of quantum circuits exploring them much easier, as seen in Section 3.3.

Hubbard type model Hamiltonians

The reduced Bardeen-Cooper-Schrieffer Hamiltonian which can be solved exactly by the Richardson - Gaudin equations up until a critical value of G, called g_c, at which mean field approximations break number symmetry and give rise to non number-preserving solutions [97] and can be written in second quantized form as:

$$\hat{H} = \sum_p \epsilon_p(\hat{n}_p + \hat{n}_{\bar{p}}) - g\sum_{p\neq q} a_p^\dagger a_{\bar{p}}^\dagger a_{\bar{q}} a_q \tag{52}$$

where $\hat{n}_p \coloneqq a_p^\dagger a_p$ is the particle number operator and $a_p^\dagger$ is the fermionic creation operator that creates a spin up fermion in orbital p and $a_{\bar{p}}^\dagger$ creates a spin down fermion, and ϵ_p is the spin-independent self-energy of a particle in orbital p, and g is the interaction strength between particle pairs. For $g > 0$ the interaction is repulsive and for $g < 0$ it is attractive ($g_c < 0$ in this sign convention). This differs from the convention used in [98]

Since the two-body interaction is done in pairs, it is useful to define the corresponding fermionic pair creation operator $\hat{P}_p^\dagger$ in orbital p, along with the number operator of pairs $\hat{N}_p$. These operators are explicitly defined via

$$\hat{P}_p^\dagger \coloneqq a_p^\dagger a_{\bar{p}}^\dagger \tag{53}$$

$$\hat{N}_p \coloneqq a_p^\dagger a_p + a_{\bar{p}}^\dagger a_{\bar{p}} \tag{54}$$

and satisfy the commutation relations

$$[\hat{P}_p, \hat{P}_q^\dagger] = \delta_{pq}(1 - \hat{N}_p) \tag{55}$$

$$[\hat{N}_p, \hat{P}_q] = 2\delta_{pq}\hat{P}_q. \tag{56}$$

This enables an alternative representation of these operators in terms of the Pauli operators X, Y, and Z:

$$\hat{N}_p \mapsto (1 - Z_p) \tag{57}$$

$$\hat{P}_p^\dagger \mapsto \frac{1}{2}(X_p - iY_p) \tag{58}$$

$$\hat{P}_p \mapsto \frac{1}{2}(X_p + iY_p) \tag{59}$$

Applying this mapping to the Hamiltonian described in Eq. 52, one arrives at

$$\hat{H} = \sum_p \epsilon_p(\hat{n}_p + \hat{n}_{\bar{p}}) - g\sum_{p\neq q} a_p^\dagger a_{\bar{p}}^\dagger a_{\bar{q}} a_q \tag{60}$$

$$= \sum_p \epsilon_p N_p - \frac{g}{2}\sum_{p\neq q}\left(P_p^\dagger P_q + P_q^\dagger P_p\right) \tag{61}$$

$$= \sum_p \epsilon_p(1 - Z_p) - \frac{g}{2}\sum_{p\neq q}(X_pX_q + Y_pY_q) \tag{62}$$

$$\hat{H} = \sum_{p=1}^{M}(\epsilon_p - \frac{g}{2})(1 - Z_p) - \frac{g}{2}\sum_{p>q}^{M}(X_pX_q + Y_pY_q), \tag{63}$$

where M is the total number of spatial orbitals. In order to compare Hamiltonians with similar structure, only the case where $\epsilon_p = p$ is considered, akin to the DOCI approach of the conrotatory ring reaction path for cyclobutane.

Notice that the mapping described in Eq. 57 to Eq. 59 differs from the standard Jordan-Wigner mapping in that under the latter $\hat{N}_p \mapsto 1/2(Z_p + Z_{\bar{p}} - 2)$, but it ensures that the orbitals represented by the first qubits have the lowest self-energy and thus leads to a canonical ordering of orbitals.

2.3.5 Numerical results

In order to utilize the algorithmic procedure in Alg. 3 and the performance metric described by Eq. 41, it is important to take into account performance parameters that extend beyond counting of iterations of the respective optimization algorithms. As shown in Eq. 40, the number of batches, distinct parametrized circuits within a batch and the number of shots for each parametrized circuit directly correlate to the wall-time it takes for the quantum device to yield results.

Thus, it is crucial to identify how each optimization algorithm handles these queries to oracle functions, more specifically the number of batches that are submitted at each optimization algorithm iteration, as the number of shots and number of distinct circuits is known by definition when deciding a measurement decomposition scheme for the expectation value estimators and these numbers do not change throughout the optimization algorithm. These help build a more objective view of the performance of each procedure with respect to the resources it uses.

In more detail, for a given state of the optimization procedure, the number of batches that are required for each optimization algorithm can be calculated as:

- Rotosolve: Number of parameters in the parametrized quantum circuit, times the number of iterations of the optimization algorithm.
- SPSA: Number of iterations of the optimization algorithm.
- ModelGrad: Number of distinct batches of circuits submitted to the oracle function f per iteration of the optimization algorithm.
- ConjModelGrad: Number of distinct batches of circuits submitted to the oracle function f per iteration of the optimization algorithm.
- ConjModelSearch: Number of distinct batches of circuits submitted to the oracle function f per iteration of the optimization algorithm.

Through the assignment of a performance metric for a given set of hyperparameters, the genetic procedure and the selection of hyperparameters can be implemented. Numerical experiments were carried out by using the VQE algorithm to calculate the ground state of a reduced BCS Hamiltonian as shown in Section 2.3.4 of 4 qubits across 10000 generations.

Owing to the stochastic nature of the VQE, to ensure that the hyperparameters consistently yield optimization paths that truly converge and not to be misled by an outlier in admitting a set of hyperparameters when the VQE optimization procedure happened to converge faster, steps were taken to prevent such an occurrence.

More specifically, for each generation of hyperparameters, the VQE optimization procedure was carried out 10 times in parallel and then by taking the averaged performance metric of these runs while also taking into consideration the root median square (RMS) error as shown in Eq. 43. Only after a set of hyperparameters that produces a $p_{averaged}$ that is lower (or higher, depending on whether one minimizes or maximized the cost function) that the previously best $p_{averaged}$, the new set of hyperparameters is saved as the current best set found through the genetic hyperparameter tuning procedure.

In the numerical experiments that were carried out with the help of a high performance computing (HPC) cluster of Covestro Deutschland AG, the optimizers that were examined were optimizers commonly used throughout the literature: SPSA [72], Rotosolve [69], Model Gradient Descent [79], Conjugate Model Gradient Descent [59] and finally the Conjugate Model Search developed by the author and presented in this thesis. In Fig. 5, the results a genetic hyperparameter tuning procedure are illustrated in the form of the evolution of all the hyperparameters involved in the aforementioned optimizers.

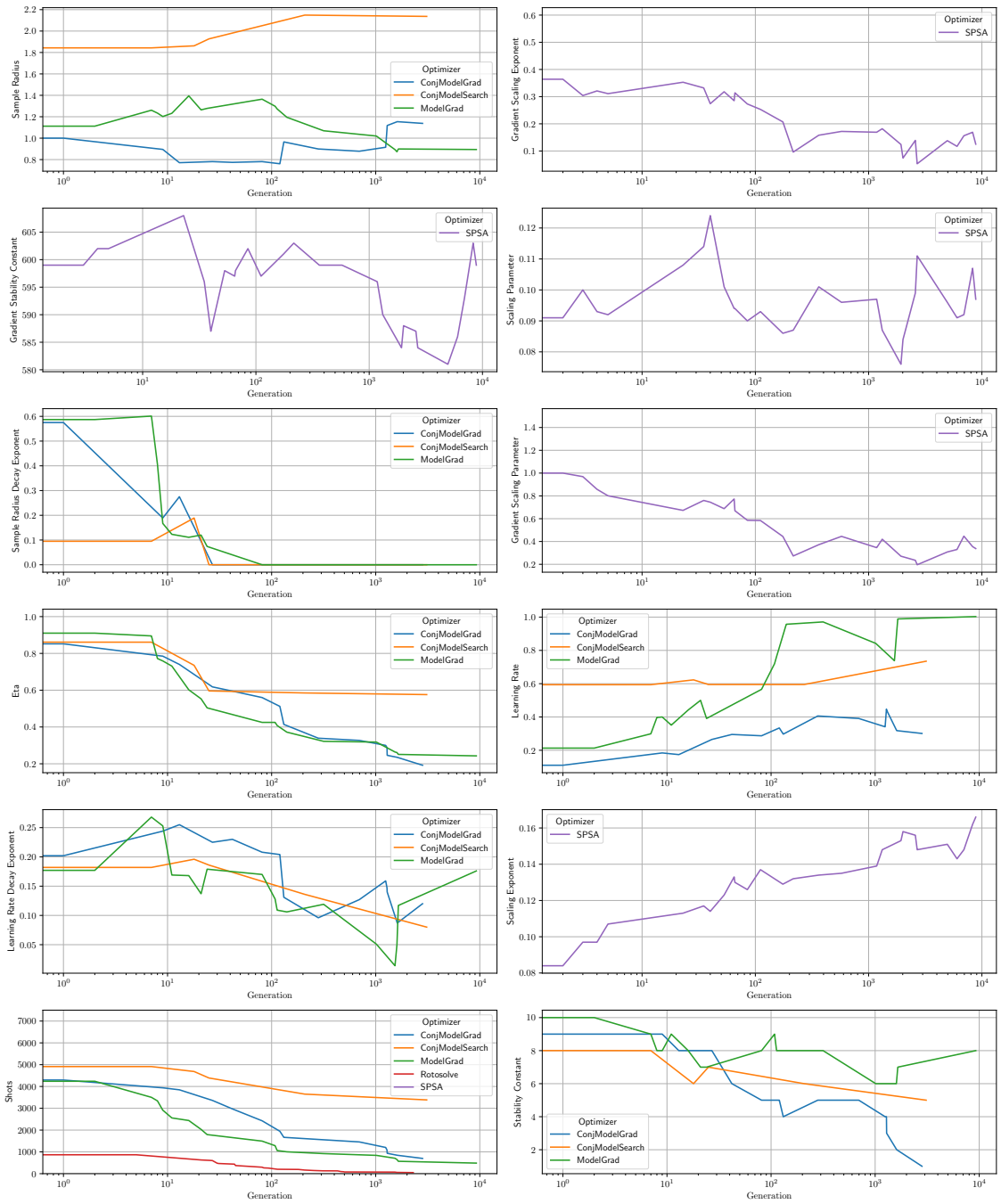


Figure 5: Evolution of the set of best performing hyperparameter for each applicable optimizer over the span of 10000 generations.

The number of shots allotted to the quantum device is also treated as a hyperparameter, as this has a direct influence on the real-life performance of the optimizer. Moreover, in the surrogate model based optimizers, a hyperparameter η is used, as a single lever to direct the number of samples k taken with the function oracle to construct the surrogate model, where m is the dimension of parameter space of the quantum circuit. With the help of the η hyperparameter, the number of samples is reformed to be $k = \eta(m+1)(m+2)/2$.

In Table 1, a tabulation including any limits to the minimum or maximum numeric values of the hyperparameters is presented. These constraints can either emerge empirically through manual tuning or as constraints with physical meaning, such as the allocation of shots in a quantum device.

As shorthand notation in the guidance used for constructing new hyperparameters, $L(x)$ denotes a linear scaling between the old numeric value of the hyperparameter and the new one to be used proportional to x, while $\mathcal{N}(\mu, \sigma)$ is used to denote a normal distribution where $\mu = 0$ is used to convey that no change is performed to the numeric value of the hyperparameter and is used as-is as a mean for the distribution. With that guidance, a new numerical value for the hyperparameter is chosen as one drawn from a distribution with the characteristics described in the guidance.

SPSA	Initial	Final	Guidance
Scaling parameter c	0.11	0.097	$0 \leq \mathcal{N}(0, 0.05) \leq 1$
Scaling exponent γ	0.101	0.166	$0 \leq \mathcal{N}(0, 0.01) \leq 1/6$
Gradient scaling parameter a	1.5	0.338	$0 \leq \mathcal{N}(0, 0.1) \leq 1$
Gradient scaling exponent α	0.6	0.125	$0 \leq \mathcal{N}(0, 0.1) \leq 1$
Gradient stability constant A	600	599	$10 \leq \mathcal{N}(0, 5)$
Shots	2×10^6	108138	$400 \leq L(0.8)$
ModelGrad			
Sample radius δ	1.0	0.894	$0.1 \leq \mathcal{N}(-0.2, 0.5)$
Learning rate γ	0.15	1.003	$0 \leq \mathcal{N}(0, 0.1)$
Number of samples η	1.0	0.243	$0.1 \leq \mathcal{N}(-0.2, 0.2)$
Stability constant A	10	8	$0 \leq \mathcal{N}(0, 10)$
Sample radius decay exponent ξ	0.5	0.0	$0 \leq \mathcal{N}(0, 0.5)$
Learning rate decay exponent α	0.2	0.176	$0.1 \leq \mathcal{N}(0, 0.1)$
Shots	5000	485	$400 \leq L(0.8)$
Conjugate ModelGrad			
Sample radius δ	1.0	1.138	$0.1 \leq \mathcal{N}(-0.2, 0.5)$
Learning rate γ	0.15	0.301	$0 \leq \mathcal{N}(0, 0.1)$
Number of samples η	1.0	0.192	$0.1 \leq \mathcal{N}(-0.2, 0.2)$
Stability constant A	10	1	$0 \leq \mathcal{N}(0, 10)$
Sample radius decay exponent ξ	0.5	0.0	$0 \leq \mathcal{N}(0, 0.5)$
Learning rate decay exponent α	0.2	0.12	$0.1 \leq \mathcal{N}(0, 0.1)$
Shots	5000	700	$400 \leq L(0.8)$
Conjugate ModelSearch			
Sample radius δ	2.0	2.137	$0.1 \leq \mathcal{N}(0, 0.05)$
Learning rate γ	0.5	0.735	$0 \leq \mathcal{N}(0, 0.1)$
Number of samples η	1.0	0.576	$0.1 \leq \mathcal{N}(-0.2, 0.2)$
Stability constant A	10	5	$0 \leq \mathcal{N}(0, 10)$
Sample radius decay exponent ξ	0.5	0.0	$0 \leq \mathcal{N}(0, 0.5)$
Learning rate decay exponent α	0.2	0.08	$0.1 \leq \mathcal{N}(0, 0.1)$
Shots	5000	3384	$400 \leq L(0.8)$
Rotosolve			
Shots	1000	50	$50 \leq L(0.7)$

Table 1: Tabulation of initial and finalized hyperparameters of various optimizers for a given system over 10000 generations of the genetic hyperparameter tuning process.

Through the genetic hyperparameter tuning procedure, the various hyperparameters that were generated for each optimizer were assigned a performance metric over the course of convergence of the VQE algorithm with the methodology detailed in Section 2.3.1. For each set generated, the VQE algorithm is run 10 times in parallel with the same hyperparameters and starting circuit parameters in the VQE optimization, each run with its own performance perfomance metric and then averaged as discussed in Section 2.3.1. The visualization of the performance metrics for the sets of hyperparameters of each optimizer that were accepted as optimal over the course of a 10000 generation genetic hyperparameter tuning procedure is shown in Fig. 6.

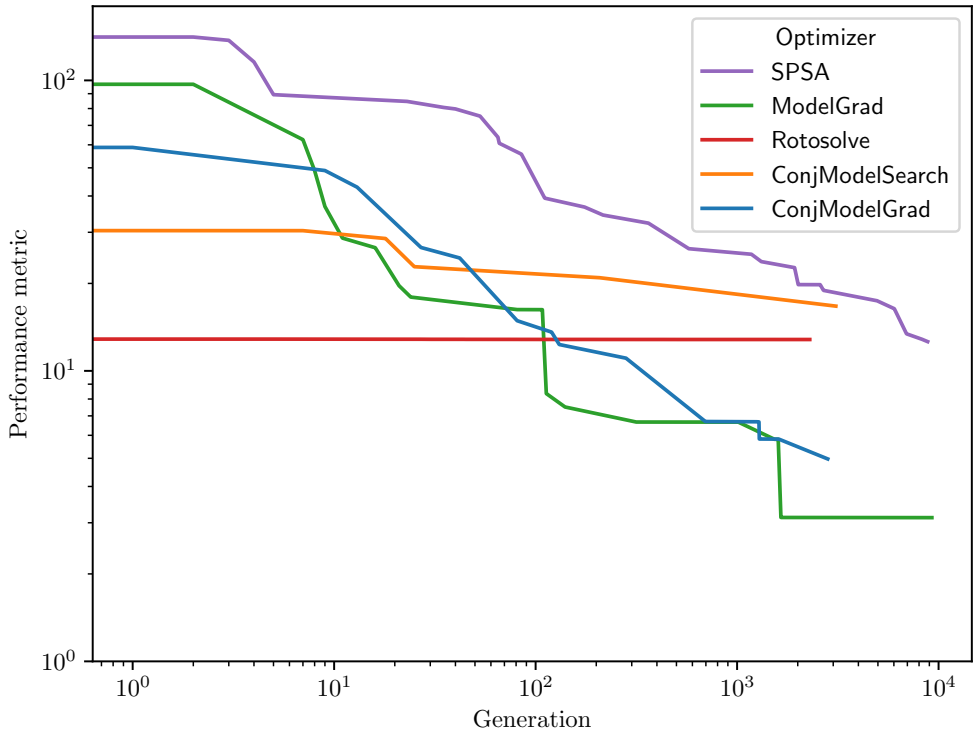


Figure 6: Evolution of the performance metric of sets of hyperparameters for various optimizers over the course of 10000 generations of the genetic hyperparameter tunis procedure.

An important point of view can also be seen through the lens of the optimization paths that the optimizers take, with each accepted generation yielding better results with respect to the number of optimization epochs or conversely, wall-time in Fig. 7. This translates to tangible performance boosts with respect to wall-time from the selection of suitable hyperparameters, showing improvements up to 2 orders of magnitude for some optimizers, in comparison to the original set of hyperparameters that were found heuristically.

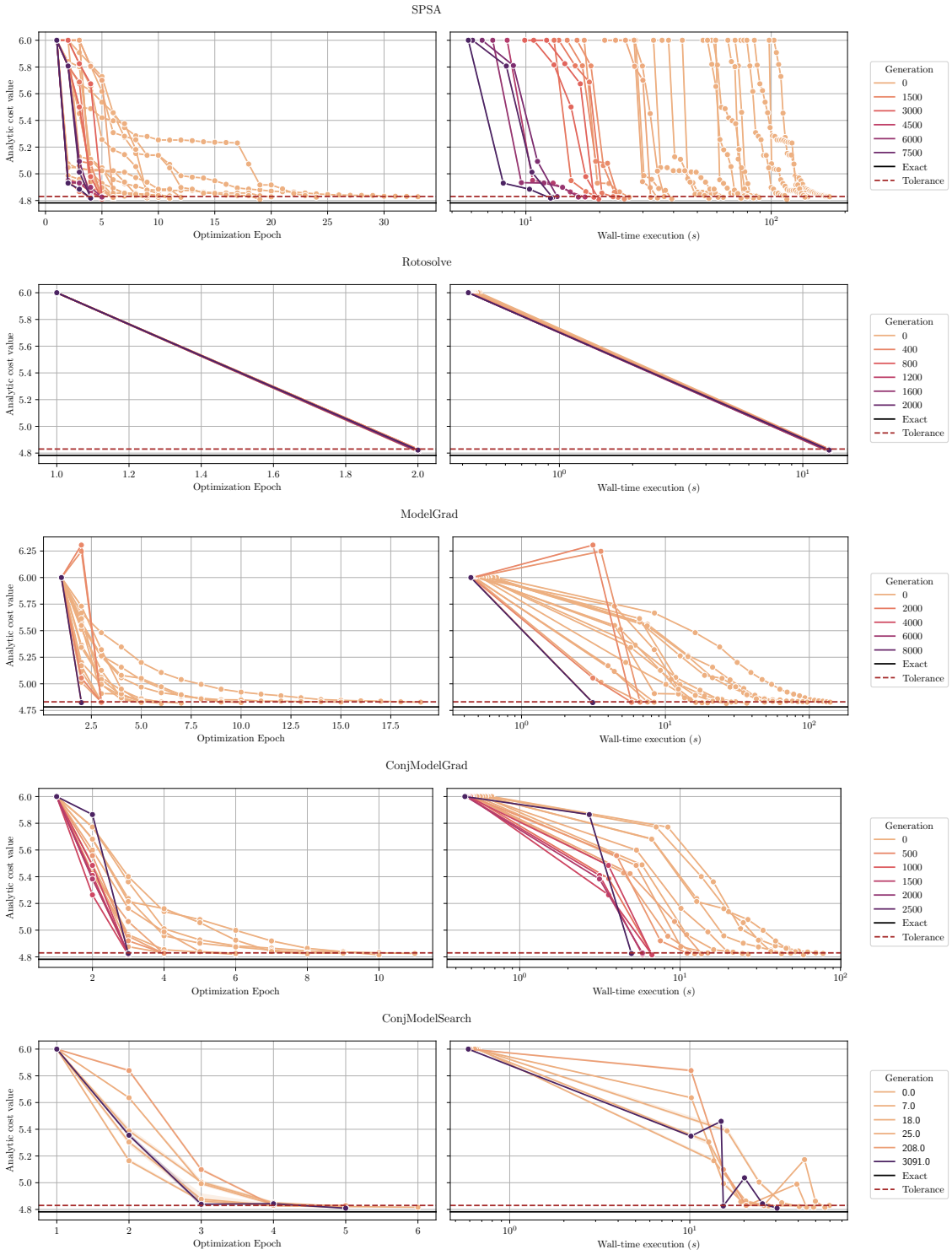
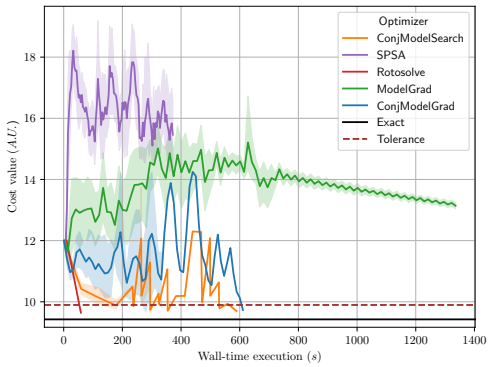
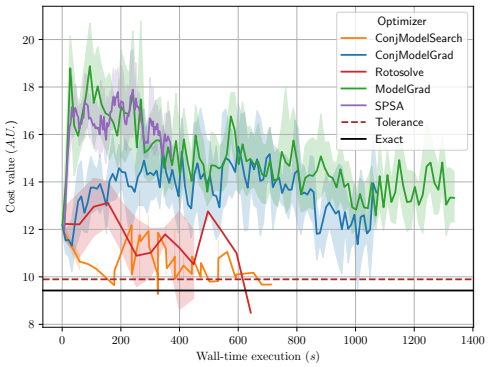


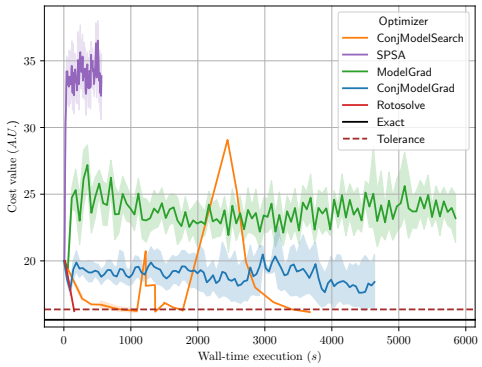
Figure 7: Analytic values of optimizers over the course of optimization across generations. Every dot marker represents the value of the cost function evaluated at the circuit parameters that the optimizer using the current set of hyperparameters of that generation on each optimization epoch, resulting in the need for fewer epochs needed, as generation yields better sets of hyperparameters.



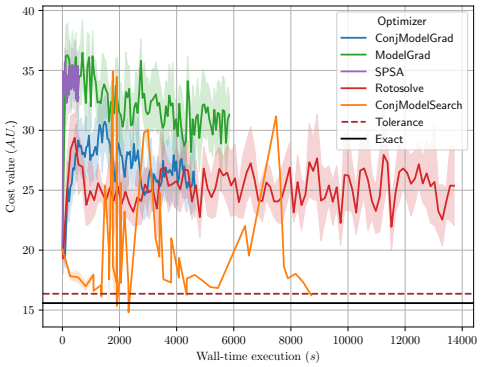
(a) System size 6



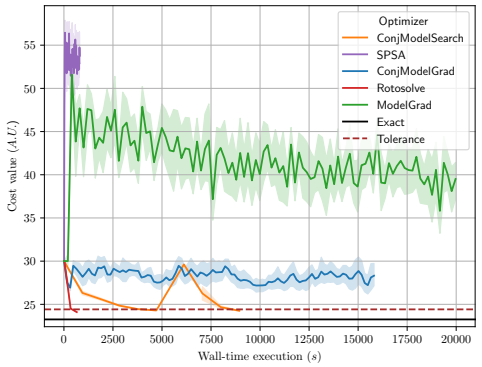
(b) System size 6 with fake noise



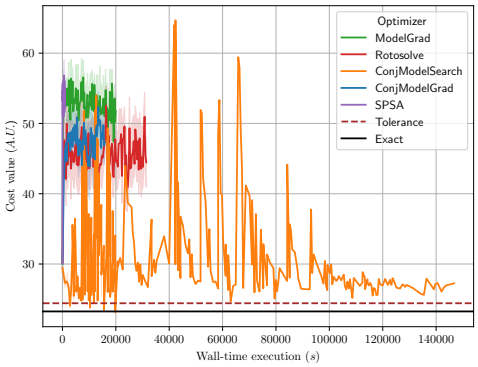
(c) System size 8



(d) System size 8 with fake noise



(e) System size 10



(f) System size 10 with fake noise

Figure 8: Scalability of the optimal set of hyperparameters for a reduced BCS Hamiltonian of system size 4 starting from the half-filled state, averaged over 10 VQE optimizations, of various optimizers across system sizes of 6, 8 and 10 qubits, in the absence and the presence of the shot noise proxy. The VQE optimization target lies 5% above the lowest energy of the Hamiltonian which can be retrieved through exact diagonalization of the sparse operator describing the Reduced BCS Hamiltonian with the help of the OpenFermion package [99].

Another aspect that these numerical experiments aimed to examine is the scalability of the hyperparameters, as well as the effect of noise in their capacity to converge with these hyperparameters. In the experiments carried out for testing scalability, the noiseless estimations of the expectation values as well as expectation values with a proxy of shot noise, described as fake noise for clarity, were also used.

To make the calculation of an expectation value computationally efficient, a value was drawn from a normal distribution whose mean is the analytic expectation value, i.e. the expectation value in the absence of noise and variance is the variance of the estimator used to produce the expectation value. This approximation was used as a means to imitate sampling noise.

In Fig. 8, the sets of hyperparameters that yielded the best performance metrics for a problem of smaller system size (size 4) were used to for VQE algorithms of different system sizes of 6, 8 and 10 qubits, as well as the absence or not of this proxy of noise. As before, to ensure that the optimization procedures are not heavily affected by the stochastic nature of VQE optimizations, 10 parallel instances of the same optimization for each optimizer was performed, with solid lines depicting the average and light shades depicting the spread of the optimization paths.

Most reused hyperparameter sets cause the optimizers to be trapped in local minima in the parameter space, causing them in some cases to even exhaust their resources in areas where the cost function is even higher than the starting point, with little means of escaping that plateau. A notable exception to this pattern of behavior is Conjugate Model Search, that manages to utilize the set of hyperparameters that was tuned for smaller system sizes to converge systems of larger size and manages to do so also in the presence of fake noise when drawing the expectation values required for its algorithm. From the numerical experiments conducted, it is shown that Conjugate Model Search can handily outperform previous state-of-the-art surrogate model optimizers in three critical aspects: it manages to always converge, faster and with far smaller spread across different optimization paths, i.e. more reliably. While Rotosolve is triumphant in noiseless regimes, in noisy environments Rotosolve tends to underperform, while Conjugate Model Search manages to converge on much larger systems than those it was originally trained in.

This behavior is especially important as a thorough genetic hyperparameter tuning procedure can be very computationally demanding, where the ability of Conjugate Model Search to able to use successfully a set of hyperparameters that have been optimized to perform well on a smaller problem of size that is computationally of reasonable size, means that finding new sets of hyperparameters for the same problems of larger size is not imperative, as is the case for other surrogate model based optimizers. By the set of hyperparameters and the optimizer performing well in this context is meant with respect to RMS error across optimization runs (converge reliably) and the wall-time model (converge fast) that best describes the architecture considered, as each quantum devices comes with its own idiosyncrasies.

These numerical experiments on hyperparameter tuning shed further light on the the trajectory of the VQE algorithm as a prospective avenue for achieving quantum advantage. Through the introduction of comprehensive ways of comparing the performance of various sets of hyperparameters of a given optimizer, along with a systematic way of benchmarking performance between different optimizers and the implementation of a new optimization algorithm, the VQE framework seems like it can still be a reasonable way of utilizing NISQ devices for the purposes of providing insights to the quantum devices and the efficient translation of interesting physical or chemical problems onto quantum computers, as well as increasing the quantum-ready capabilities of computational clusters. However, even equipped with these tools of optimization, the prospect of VQE becoming a viable path towards quantum advantage seems uncertain, as the requirements in system sizes, experiment repetitions and error rates are still challenges that remain ever present.

3 Chapter II: Quantum Computing on Hardware

However powerful the algorithms that have been currently been developed may be, they are still bound by the limitations presented by the available hardware today, both in computation size in terms of qubits, as well as runtime. These limitations regarding runtime arise from the inherent imperfections of the operations in quantum devices, such as for example gate errors due to miscalibration, limited coherence times of the physical qubits or readout errors of the measurements taken from these devices [100].

Significant advancements have been made towards the construction of fault tolerant devices [36], but Quantum Error Correction (QEC) algorithms remain at a nascent stage of adoption in terms of their support from current day hardware as they require a large number of qubits and very low error rates. As an alternative approach while hardware developments continue to unfold, **Quantum Error Mitigation** (QEM) schemes have been proposed. The goal of these QEM techniques, unlike QEC, is not to supress all error exhibited by the quantum devices, but rather apply protocols and routines to partially compensate for the accrued errors during quantum computations.

Through the combination of post-processing tactics with strategic design of quantum circuits and calibration routines, quantum error mitigation methods constitute a practical route to extract as much of the performance of NISQ devices as possible. While only a stand-in for full fault-tolerant protocols, these approaches have already expanded the scope and size of feasible quantum simulations and algorithms on current quantum and classical hardware, they help incrementally forge a path toward more robust and sizable quantum computations by leveraging redundancy either in space (qubit count) or in time (circuit depth).

3.1 Common error mitigation techniques

As QEM schemes aim to act as a segue between the capabilities of current hardware and the high fidelity constraints required by QEC schemes in order to run effectively, it is useful to understand the types of **noise** that destabilize quantum measurements in the first place. The type of noise present in quantum computers can be categorized in two main types, coherent and incoherent noise.

Assuming that these quantum noise channels are able to be represented by a unitary operation, these quantum channels $\mathcal{E}$, when applied to systems with density matrix ρ have the properties of having $0 \leq \mathrm{Tr}\{\mathcal{E}(\rho)\} \leq 1$ when they occur, being convex-linear maps, and finally, completely positive meaning that for any quantum operation A in the ensemble of operations $\mathcal{E}$, $\mathcal{E}(A) \geq 0$ [42].

Coherent noise often emerges from control inaccuracies of the apparatus driving the quantum device, or miscalibrations of the system upon setup. This can be manifested as consistent over- or under- rotations of quantum gates by a small angle γ over the original parameter θ, thus the quantum gates perform the operation $\hat{U}(\theta') \mapsto \hat{U}(\theta \pm \gamma)$. A characteristic that sets this type of noise apart from its counterpart is the fact that these preserve the trace $\mathrm{Tr}(\rho)$ of a quantum system. These errors can cause the fidelity algorithms such as the VQE to degrade as they shift the minimum energy to be found at the same energy as an unperturbed setting, which is why compensating for them is important as they are not self averaging, leading to potential biases in the measurement outcomes from the quantum device [40].

As an antithesis to its coherent counterpart, **incoherent noise** describes quantum channels of noise such as amplitude damping, phase damping or dephasing the originally pure state ρ into a stochastic mix of ρ_{noisy} states. Incoherent noise processes cause stochastic errors by eroding the phase stored in the qubit registers and their ability to stay coherence of long periods of time. This is often evident in experiments as degradation of the prepared state, as deeper circuits or longer execution times observe larger accumulation of incoherent noise, making it increasingly challenging to extract measurements with the necessary fidelity [40].

These types of noise are described by density matrices of size $2^n \times 2^n$ for n sized systems, therefore density matrix simulators are finding it computationally demanding to store and manipulate matrices of that size efficiently. These factors, along with the fact that on real devices, these noise channels interplay between different qubits due to crosstalk of the quantum modes, makes the study of QEM and QEC techniques ever more crucial to the advancement of quantum computing towards achieving practical utility [42].

Another avenue for achieving the goal of building estimators whose bias is reduced under common noise models is by techniques such as zero-noise extrapolation, where one artificially

inputs noise at different degrees and type into the circuit to draw an extrapolation of the expected performance if there were no noise [101, 102]. By accounting for the noise models taken into consideration, one can add gates to the circuit to combat the effects that the noise channels had on the results of the circuit measurements [103, 104].

3.1.1 Post-selection

A common but effective quantum error mitigation technique is known as **post-selection**, where the concept of symmetry preservation or physical constraints are employed to discard measurements that do not obey them [105]. Another way to view post-selection is as a purification process of the noisy state by performing projections that can only obey the symmetries S in that Hamiltonian, biasing the measurement outcomes towards more meaningful outcomes. Post-selection is often referred to as a zero-cost error mitigation method as it does not take a toll on quantum resources such as qubit counts or circuit depth, but requires a higher number of experiment repetitions which is a drawback of a varying degree depending on the platform with which the quantum device is implemented.

The outline of the post-selection procedure here closely follows the presentation in [105]. Symmetries S are defined as quantities that commute with the Hamiltonian of interest $[\hat{H}, S] = 0$, as well as any decompositions of fast-forwardable terms $[\hat{O}_i, S] = 0$ that one might choose for that Hamiltonian. The actual measurement process of the Hamiltonian comprises of a set of projector-valued measurement (PVM) operators $\hat{M}_i$ for each observed value i, with probability p_i, thus adhering to the Born rule [105]:

$$p_i = \langle\psi|\hat{M}_i|\psi\rangle \tag{64}$$

By constructing these PVM operators in such a way that $\hat{M}_s$ acts as $\hat{M}_s|\psi\rangle = |\psi\rangle$, where S denotes the target eigenspace of the symmetry to be explored in the absence of noise, the density matrix derived from this symmetry projection can be written as [105]:

$$\rho_{\mathrm{s}} = \frac{\hat{M}_{\mathrm{s}}\rho\hat{M}_{\mathrm{s}}}{\mathrm{Tr}[\hat{M}_{\mathrm{s}}\rho]} \tag{65}$$

To solidify this as an error mitigation technique, it is sufficient to calculate the overlap of the projected density matrix ρ_s with the pure state $|\psi\rangle$ than the pre-selected density matrix ρ [105]:

$$\mathrm{Tr}[\rho_s|\psi\rangle\langle\psi|] = \frac{\mathrm{Tr}[\rho|\psi\rangle\langle\psi|]}{\mathrm{Tr}[\hat{M}_s\rho]} \geq \mathrm{Tr}[\rho|\psi\rangle\langle\psi|] \tag{66}$$

During the simulation of fermions, and especially in the case of quantum chemistry, symmetries such as the number operator $\hat{N} = \sum_i \hat{a}_i^\dagger \hat{a}_i$ of the fermions to be simulated, or the number of alpha and beta electrons are conserved throughout the calculation and therefore proving very useful for use in such a setting of a post-selection protocol.

3.1.2 Echo verification

Historically, Quantum Phase Estimation (QPE) has been a central source of inspiration as a starting points for a variety of quantum algorithms whose goal is to estimate eigenvalues of unitary operators or Hermitian Hamiltonians via decomposition in fast-forwardable terms [42]. An overview following closely the pedagogical examples contained in [106] touching on the single control phase estimation follows. The way that this is achieved has been by measuring an ancilla qubit in the plus state $|+\rangle_a$ and added to the circuit implementing system of interest, with the system being described by the action of a unitary operator on the zero state of a system $\hat{U}_p(\theta)|0\rangle_s = |\psi\rangle_s$ to prepare the state $\frac{1}{\sqrt{2}}(|0\rangle_a|\psi\rangle_s + |1\rangle_a e^{i\hat{O}t}|\psi\rangle_s)$ as:

This type of circuit can be utilized for the extraction of expectation values by separating the unitary $\hat{U}_p(\theta)$ acting on the zero state of the system as $\hat{U}_p(\theta) = \hat{U}_a(\theta)\hat{S}_{\mathrm{prep}}$, where the $\hat{S}_{\mathrm{prep}}$ is a state preparation operation, transforming the all zero state of the system into the reference state $|\psi_{\mathrm{ref}}\rangle$ for the ansatz that remains unchanged, and $\hat{U}_a(\theta)$ denotes the parametrized ansatz to be applied to the circuit. The reference state $|\psi\rangle_s$ contains the eigenstates $|E_j\rangle_s$ of the observable $\hat{O}$ as a linear superposition $|\psi\rangle_s = \sum_j m_j|E_j\rangle_s$, making the wavefunction $|\varphi(t)\rangle$ prepared [106]:

$$|\varphi(t)\rangle = \sum_j \frac{1}{\sqrt{2}}(|0\rangle_a + e^{iE_j t}|1\rangle_a)|E_j\rangle_s \tag{67}$$

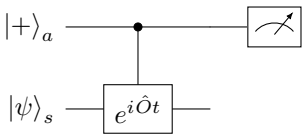


Figure 9: A circuit performing a SWAP test for encoding the phase in an ancilla qubit.

In order to extract from the ancilla qubit the eigenvalues E_j and their probabilities M_j, where $M_j = |m_j|^2$, and construct the full expectation value of the Hamiltonian $\langle \hat{H} \rangle = \sum_j M_j E_j$, the trace over the system register s can be utilized [106]:

$$\begin{aligned} \rho_a(t) &= \mathrm{Tr}_s[|\varphi(t)\rangle\langle\varphi(t)|] \\ &= \frac{1}{2}\begin{pmatrix} 1 & g(t) \\ g^*(t) & 1 \end{pmatrix} \end{aligned} \tag{68}$$

where $g(t) = \sum_j M_j e^{iE_j t}$. The function $g(t)$ can be used as a phase function to estimate the expectation values of H, with a common technique being through its derivative [106]:

$$g'(t) = \sum_j i M_j E_j e^{iE_j t} \quad \rightarrow \quad \mathrm{Im}(g'(0)) = \sum_j M_j E_j = \langle \hat{H} \rangle \tag{69}$$

This process relegates the procedure of finding expectation values via the Hadamard test to retrieving them via the phase function of the controlled ancilla qubit, with systematic ways of finding the optimal, with respect to the estimation of expectation values, use of the phase function and its gradient [107] in a protocol called **echo verification**. These protocols can be extended to include verification as depicted in Fig. 10, the process of returning the system to its original state to check for errors in evolution of the system by undoing the action of the original $\hat{U}_p(\theta)$ operator through the application of its conjugate, $\hat{U}_p^\dagger(\theta)$ and thus applying $\hat{U}^\dagger$ and $S_{\mathrm{prep}}^\dagger$ to undo the computation, creating redundancy in circuit depth.

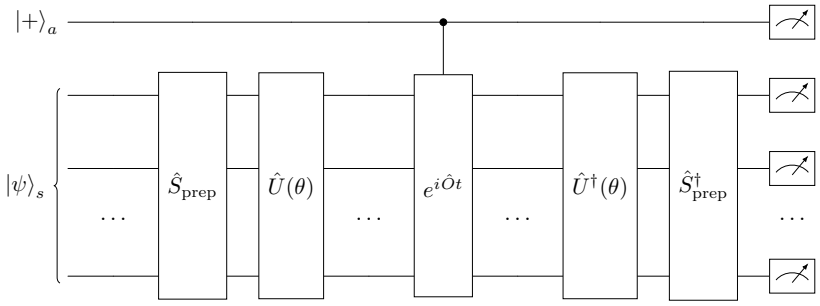


Figure 10: Circuit depicting the protocol of ancilla-controlled echo verification.

The density matrix of the ancilla qubit ρ_a can be decomposed to two parts; one part $\rho_a^{(\mathrm{v})}$ that passes the verification and another $\rho_a^{(\mathrm{f})}$ that fails to pass the verification protocol. The phase function $g(t)$ remains unchanged by the process of undoing the computation since the part of the density matrix of the ancilla $\rho_a^{(\mathrm{f})} = |1\rangle_a\langle 1|_a$ contributes on average 0 to the trace, owing to not passing the verification criterion always projects the ancilla in the $|1\rangle_s$ state [106].

$$\mathrm{Tr}[\rho_a|0\rangle_a\langle 1|_a] = \mathrm{Tr}\left[\rho_a^{(\mathrm{v})}|0\rangle_a\langle 1|_a + \underbrace{\rho_a^{(\mathrm{f})}|0\rangle_a\langle 1|_a}_{0}\right] = g(t) \tag{70}$$

A decomposition of the phase function to account for the event of noise in the system is the following [106]:

$$g_{\mathrm{noisy}}(t) = p_{\mathrm{ne}}(t)g(t) + \mathcal{O}(2^{-N}p_{\mathrm{err}}(t)) \tag{71}$$

where N is the system register size used for verification and $p_{\mathrm{ne}}(t)$, $p_{\mathrm{err}}(t)$ denote the probabilities of no errors and errors occurring respectively, with $p_{\mathrm{ne}}(t)$ also being referred to as the fidelity of

the measurement process. As it can be shown that by rescaling the amplitudes M_j to account for the time dependence in the noise occurrence, $p_{\text{ne}}(t) = p_{\text{ne}}$, the corrected expectation value of an observable $\hat{O}$ becomes:

Echo verification

$$\langle \hat{O} \rangle_{\text{corrected}} = \frac{\text{Im}(g'(0))}{p_{\text{ne}}} \tag{72}$$

Variants of this protocol have been developed that are freed from the need for ancilla qubits [106]. When suitable reference states are available, for the encoding of the phase function in the phase difference between the phase reference state $|\psi_O\rangle$ and the original reference state $|\psi_{ref}\rangle_s$ of the system. The condition for this to be possible is that the reference state for an observable $\hat{O}|\psi_O\rangle = E_O|\psi_O\rangle$ must also be an eigenstate of the unitary ansatz $\hat{U}(\theta)$ applied after the state preparation step as $\hat{U}|\psi_O\rangle = E|\psi_O\rangle$, with the protocol being referred to as **control-free echo verification** [106].

When performing simulations for chemistry calculations, the fermionic vacuum state $|0\ldots0\rangle$ is always an eligible state for this protocol when used in conjunction with Quantum Number Preserving ansätze since the aforementioned conditions are met. In Fig. 11 the circuit that would execute this protocol is depicted. Upon applying a Hadamard gate on a qubit (top-most qubit in Fig. 11, labeled #1 for convenience) and the subsequent application of a state preparation conditioned on the selected qubit #1, one retrieved the superposition between the reference state and the fermionic vacuum as $1/\sqrt{2}(|0\ldots0\rangle_s + |\psi_{\text{ref}}\rangle_s)$. Lastly, time evolving under the rotation caused by the observable and applying the uncomputation of $C\hat{S}^\dagger_{\text{prep}}\hat{U}^\dagger_{\text{QNP}}(\theta)$ produces the following state in qubit #1, allowing for the rest of the system register to be used for verification of the measurements:

$$|\varphi(t)\rangle = \frac{1}{\sqrt{2}}(e^{iE_O}|0\rangle_1 + e^{i\hat{O}t}|1\rangle_1) \tag{73}$$

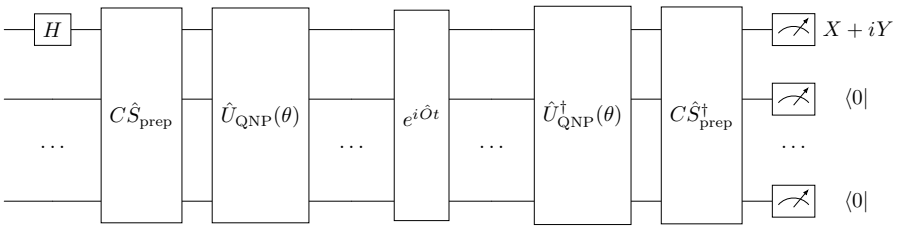


Figure 11: Circuit depicting the protocol of Control-free echo verification. This is done by preparing a superposition of an appropriate phase reference state and a starting state for the calculation via a Hadamard gate H, where the only changes permitted are incurred by the ansatz $\hat{U}_{\text{QNP}}(\theta)$ and the imaginary time evolution $e^{i\hat{O}t}$.

3.1.3 Virtual distillation

By exploring error mitigation techniques that utilize redundancies in space [103] as a means of estimating expectation values more accurately, the technique of virtual distillation [108], or otherwise known as exponential error suppression [109] came about. In the **virtual distillation** error mitigation protocol, one tries to estimate the expectation value of an observable through the purified density matrix from the measurements taken from multiple copies of the noisy system, where the redundancy in space comes into play as this approach requires multiple N sized registers for each copy [101].

Due to the imperfect nature of current generation NISQ devices, the states prepared are imperfect. An avenue to account for this is by separating the density matrix of the dominant eigenvector and those states $|\psi_\varepsilon\rangle$ that have been subjected to noise as ρ_ε and the action of a *derangement operator* D to ensure that only symmetric states can contribute to the expectation value measurement process [109]. In a minimal example of 2 copies and SWAP acting as the derangement operator shown in Fig. 12, by measuring the ancilla qubit in the $|+\rangle$ state, one can extract the expectation value of the observable $\hat{O}$.

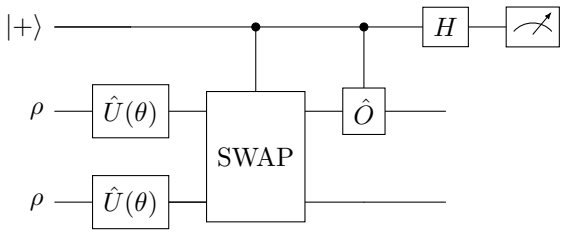


Figure 12: Representation of a virtual distillation circuit.

This protocol is shown to work for arbitrary $\lambda > 0$ and $\lambda > (1-\lambda)p_\varepsilon$ as [109], meaning that the state one is looking for, needs to have non-zero overlap with the dominant state and the other eigenvalues that appear in the measurement process. With $\lambda \leq 1$, the number of qubits N of each register and $\sum_i p_\varepsilon = 1$ representing a probability distribution, the spectral decomposition of a density matrix prepared by a quantum computer can be expressed as:

$$\begin{aligned} \rho_{\text{noisy}} &= \lambda|\psi\rangle\langle\psi| + (1-\lambda)\rho_\varepsilon \\ &= \lambda|\psi\rangle\langle\psi| + (1-\lambda)\sum_{i=2}^{2^N} p_\varepsilon|\psi_\varepsilon\rangle\langle\psi_\varepsilon| \end{aligned} \tag{74}$$

The goal of virtual distillation is to retrieve the expectation value of $\langle\psi|\hat{O}|\psi\rangle$ while keeping any contributions to the expectation value from erroneous states $\langle\psi_\varepsilon|\hat{O}|\psi_\varepsilon\rangle$ as suppressed as possible. By preparing multiple copies of the ρ_{noisy} system, in an ideal case where no noise channels have affected the state preparation, one would have prepared the state $|\psi,\ldots,\psi\rangle$ and measuring the designated (denoted as $|+\rangle$ in Fig. 12) qubit, would yield $\langle\psi,\ldots,\psi|\hat{O}\psi,\ldots,\psi\rangle = \langle\psi|\hat{O}|\psi\rangle$ [109].

By admitting the action of noise channels into one of the copies of the state prepared, then the prepared state becomes $|\psi_\varepsilon,\psi,\ldots,\psi\rangle$, contributing to the estimator:

$$\langle\psi_\varepsilon,\ldots,\psi,\psi|\hat{O}\psi_\varepsilon,\psi,\ldots\psi\rangle = \langle\psi_\varepsilon|\hat{O}|\psi_\varepsilon\rangle \tag{75}$$

However, when the order of registers is permuted, for example through the action of $\hat{O}$SWAP as in Fig. 12, this error would be canceled due to the orthonormal structure of the eigenvectors of a density matrix [109]:

$$\langle\psi_\varepsilon,\ldots,\psi,\psi|\hat{O}\psi,\psi,\ldots\psi_\varepsilon\rangle = \langle\psi_\varepsilon|\hat{O}|\psi\rangle\langle\psi_\varepsilon|\psi\rangle = 0 \tag{76}$$

Any such operation that permutes *all* registers *with all* can be called a derangement operation. As the way that these permutation maps of each site to another is equivalent to another, the cyclical permutation of indices is usually chosen for simplicity, but other permutations of these exist, for example round-robin strategies [109, 59].

Virtual Distillation

$$\langle\hat{O}\rangle_{\text{corrected}} = \frac{\text{Tr}(\hat{O}\rho^M)}{\text{Tr}(\rho^M)} \tag{77}$$

Through the action of the derangement operation on M copies of the system, the exponential suppression in errors scaling with the number of copies M of the original density matrix ρ can suppress $(1-\lambda)^M$ of errors in contributions to the estimator as:

$$\text{Tr}(\hat{O}\rho^M) = \lambda^M\langle\psi|\hat{O}|\psi\rangle + (1-\lambda)^M\sum_{k=2}^{2^N} p_\varepsilon^M\langle\psi_\varepsilon|\hat{O}|\psi_\varepsilon\rangle \tag{78}$$

Another advantage of the virtual distillation protocol is that its inherent weakness towards coherent errors across all registers is that it can be readily remedied when combined with other error mitigation techniques in an attempt to address those as well. Post-selection protocols can be

utilized to emphasize the dominant eigenvector further, with the ability to be complemented by zero-noise extrapolation [109, 108].

By utilizing a plethora of SWAP tests, the task of finding an expectation value through this protocol can be realized by many different routes. One notable such route is the compact unitary SWAP test of dual state purification [110], owing to its equivalence to the controlled version of Echo Verification as discussed in Section 3.1.2, and retaining the ability to be performed without the use of an ancilla qubit.

By formulating the density matrix ρ under the effects of a noise process as $\rho = U^\dagger(\boldsymbol{\theta})_\epsilon|0\rangle\langle 0|(U(\boldsymbol{\theta})_\epsilon)^\dagger$ and its dual $\bar{\rho} = U^\dagger(\boldsymbol{\theta})_\epsilon|0\rangle\langle 0|(U^\dagger(\boldsymbol{\theta})_\epsilon)^\dagger_\epsilon$, the inability to implement the complex conjugate due to $(U(\boldsymbol{\theta})_\epsilon)^\dagger_\epsilon \neq U^\dagger(\theta)_\epsilon$ is explicitly underlined [110]. Armed with these definitions, the estimator for the state-verified observable for control-free echo verification can be attained similarly to Eq. 77 by using dual state purification [110]:

$$\langle \hat{O} \rangle_{\text{corrected}} = \frac{\text{Tr}(\hat{O}\rho\bar{\rho})}{\text{Tr}(\rho\bar{\rho})} \tag{79}$$

3.2 Application of optimizers

With the crucial components for the successful implementation of the Variational Quantum Eigensolver algorithm in an experimental setting presented, the importance for a choice of suitable optimizers is evident. Through significant but still finite, error suppression provided by error mitigation techniques to the oracles of the cost function to be minimized are posed to fare better in the simulation of systems with quantum computers. In Section 3.3, a comprehensive study of implementations that explore the effect of different error mitigation schemes have on expectation values drawn from real superconducting qubit hardware.

With these constraints in mind, the pursuit for optimizers that are able to tolerate these imperfections in the function oracle and still achieve convergence is well underway. Through this pursuit, the optimizer of Conjugate Model Gradient Descent was created. When paired with ansätze such as the Quantum Number Preserving ansatz and measurement methods that make use of the symmetries innate to the Hamiltonian, a significant help to the task of optimizers can be given, as for the Conjugate Model Gradient Descent optimizer the number of calls to the oracle function scales with the number of parameters gates n roughly as $\mathcal{O}(n^2)$ when used in conjunction with the QNP ansatz [59].

Ansätze influence the number of parameters in the quantum circuit to be optimized in order to find the lowest energy of the cost function, and measurement methods directly impact the number of observables through the decomposition of the Hamiltonian into jointly measurable term groups, with each term group requiring a distinct circuit due to bases rotations for them to be measured. Both of these factors directly dictate how much time a hybrid algorithm runs in terms of wall-time, as longer run times give more chances for error channels to act on the computation, assuming that the rest of quantum and classical resources such as number of qubits and computer memory do not become a bottleneck.

Though Conjugate Model Gradient Descent was a considerable step forward from other surrogate model based optimizers at the time by constructing and performing optimization along the conjugate gradient direction with no additional expenditure of quantum resources with better hyperparameter scalability characteristics than its predecessor Model Gradient Descent, the scaling of the number of calls to the oracle function along with the size of discrepancy of the converged energy from the theoretical target as the optimization goal for modest system sizes posed significant obstacles towards the use of Conjugate Model Gradient Descent as the optimizer of choice for the largest VQE convergence of chemical energies at the time. These impediments acted as motivation for the conceptualization of Conjugate Model Search by the author as detailed in Alg. 1, in an effort to make better use of each call to the function oracle as that is by far the largest bottleneck in the procedure and require as few calls to the function oracle as possible. These efforts resulted in the development of Conjugate Model Search, which is shown to have superior performance to the Conjugate Model Gradient Descent with respect to convergence speed, the consistency of the quality of the results obtained, as well as its tolerance to forms of noise.

3.3 Purification-based quantum error mitigation of pair-correlated electron simulation

Bibliographic information

O'Brien, T. E., Anselmetti, G., **Gkritsis, F.**, Elfving, V. E., Polla, S., Huggins, W. J., ..., Rubin, N. C. (2023, December). Purification-based quantum error mitigation of pair-correlated electron simulations. Nat. Phys. 19, 1787–1792 (2023). Retrieved 2024-09-02, from: `https://www.nature.com/articles/s41567-023-02240-y` DOI: 10.1038/s41567-023-02240-y

Summary

We examine various error mitigation techniques based on purification methods applied to a superconducting quantum computer, specifically Google's Sycamore chip. Due to the significant computational expense of simulating quantum circuits with noise compared to ideal state simulators, and the inherent approximation of analytically implemented noise channels, evaluating error mitigation techniques on real hardware is crucial. Our research focuses on achieving accurate results, up to chemical accuracy, in quantum chemistry, particularly in the Doubly Occupied Configuration Interaction (DOCI) approximation and the Richardson-Gaudin Model, a model for superconductivity. We demonstrate a notable enhancement, two orders of magnitude, compared to post-selection methods, and observe favorable scalability with increasing system size, deriving scaling laws. This study represented one of the most extensive and intricate Variational Quantum Eigensolver (VQE) experiments conducted on quantum hardware to that date.

Contributions

The author contributed to this work: (1) Conceptualization and implementation of the Conjugate Model Gradient optimizer (2) Hyperparameter Tuning for the Conjugate Model Gradient optimizer (3) Implementation of Richardson Gaudin models in OpenFermion Python package [99] (4) Conversion of reduced Bradley-Cooper-Schieffer Hamiltonians to a corresponding electronic structure Hamiltonian (5) Provided quantum run time estimates with real wall-time models of the Google Sycamore chip for the experiment (6) Numerical simulations of the experiment, including investigation of pre-initializations of the circuit parameters and optimal layouts, and helped writing the manuscript.

[DOI: https://doi.org/10.1038/s41567-023-02240-y]

3.4 Quantum computing on ultracold fermions in optical superlaticces

The dynamics of ultracold fermions can lend themselves as a prime candidate as a means of quantum simulations. As discussed in Section 2.1.1, a class of ansätze can require preservation of symmetries of quantum numbers such as the spin-squared $\hat{S}^2$ or the number of spin-up ($\hat{N}_\alpha$) and spin-down ($\hat{N}_\beta$) electrons when tasked with the simulation of correlated ground and excited states of systems in quantum chemistry. The benefits and ease of implementation of these Quantum Number Preserving (QNP) ansätze can vary depending on the type of platform of quantum computing used for the purpose of quantum simulation [49]. In order to make the most of the fundamental advantages offered by the neutral atom architecture of quantum computers, such as the intrinsic respect of fermionic exchange symmetry a mapping is needed to utilize the fermionic exchange statistics to dynamics of systems found in quantum chemistry. Such a mapping can be done through the use of optical superlattices, where the tesselation of gates is implemented easily, while simultaneously reducing the depth, and therefore sources of error, for the relatively deep decompositions of building block elements $\hat{Q}$ of the QNP ansatz compared to when implemented with fermion-to-qubit mappings [111].

As described in Section 2.1.2, measurement methods are central points of interest for the efficient implementation of quantum simulations. That is even more prevalent for quantum computing technologies where experiment repetitions are expensive in a wall-time sense, such as neutral atoms with a measurement cycle taking roughly 1 second. The large number of terms in a molecular Hamiltonians such as the one presented in Eq. 14, along with the shot requirements of estimating each term to the desired accuracy made the need for the development of measurement methods that mitigate the increase of complicated terms from the two-electron interactions. Through the use of factorization techniques, the energy estimator containing up to $\mathcal{O}(n^4)$ terms in the typical formulation of the second quantized electronic structure Hamiltonian where n is the number of molecular orbitals, that energy estimator can be approximated through the use of $\mathcal{O}(n^2)$ basis transformations $\hat{U}_l$, where l is the number of leaves of the approximation, by the just on-diagonal elements [54, 55].

Moreover, a prominent factorization technique, Regularized Compressed Double Factorization (RC-DF) has been shown to be extendable to incorporate the use of Fluid Fermionic Fragments (FFF) [32]. The $\hat{U}_l$ basis transformations along with the degree-one $\hat{p}^{(1)}$ and degree-two $\hat{p}_l^{(2)}$ polynomials of number operators can be combined with the coefficients derived from the FFF procedure to accommodate terms that are proportional to these number operators seen in the degree-one and degree-two polynomials, yielding significant savings in the number of experimental repetitions.

3.4.1 Quantum number preserving gate fabric ansätze

The **quantum number preserving** (QNP) ansatz through the preservation of symmetry at the gate level, can help pave the way for variational ansätze towards designs that are both expressive and efficient enough to depict the dynamics of strongly correlated fermionic systems. The effect that this embedding of preservation achieves is mitigation against symmetry contamination in the trial wavefunction, meaning wavefunctions that by construction do not satisfy the parameters of the studied system are not permissible, and under ideal conditions, ensuring that the variational state produced by the ansatz resides within the irreducible representation of the targeted Hilbert space [49].

The two central elements composing a QNP ansatz are the Pair-Exchange QNP_{PX} and the Orbital Rotation gate QNP_{OR}, constructed with the help of the Givens rotation gate $\hat{G}$ [49]. The Givens rotation gate has entries as described below, with $c \coloneqq \cos(\lambda/2)$ and $s \coloneqq \sin(\lambda/2)$ as the elements of the matrix defining the gate responsible for regulating the rotation between the 1-particle states:

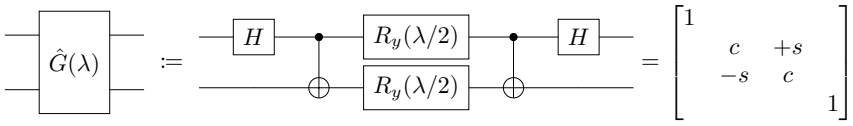


Figure 13: Decomposition of a Givens rotation gate and its generator matrix.

The Givens rotation gate described in Fig. 13 as a cornerstone for the implementation of more complex gates, $\mathrm{QNP}_{\mathrm{OR}}$ and $\mathrm{QNP}_{\mathrm{PX}}$ to construct the $\hat{Q}$ block elements of a QNP ansatz [112]. The $\mathrm{QNP}_{\mathrm{OR}}$ Orbital Rotation gate acts as a rotation to both spin orbitals α and β identically, by rotating the basis of the spatial orbitals describing the electrons. The $\mathrm{QNP}_{\mathrm{PX}}$ Pair Exchange gate implements the diagonal pair Givens rotation of orbital pairs [49]. Moreover, the design of the $\mathrm{QNP}_{\mathrm{OR}}$ and $\mathrm{QNP}_{\mathrm{PX}}$ gates depicted in Fig. 14, accommodates to the nearest-neighbor architecture of most prominent quantum computing platforms, while simultaneously canceling out some strings of Pauli Z strings incurred by the Jordan-Wigner mapping [49].

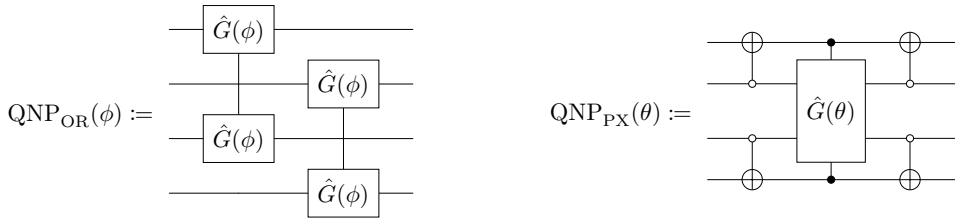


Figure 14: Quantum number preserving Orbital rotation and Pair Exchange gates decompositions into parametrized Givens rotation gates under Jordan-Wigner decompositionfor the depiction of spin-1/2 systems with interleaved ordering of spin orbitals.

Another way to examine the result of these $\hat{G}$ blocks of the QNP ansatz proposed in [113] to construct a trial wavefunction out of the initial $|\Psi^{\mathrm{HF}}\rangle$ state is through the help of spin-adapted one-body operators $\hat{\kappa}_{pq}^{(1)}$ and paired two-body operators $\hat{\kappa}_{pq}^{(2)}$, as they can be used to build arbitrary wavefunctions from unitary fermionic operators [113, 114]. The construction of the single excitation operator $\hat{E}_{pq}$ can be written with the help of the usual creation and annihilation operators, with $\hat{a_p}^\dagger$ and $\hat{a}_p^\dagger$ representing the creation of an electron of spin up ad spin down respectively at spatial orbital labeled p as:

$$\hat{E}_{pq} = \hat{a_p}^\dagger \hat{a_q} + \hat{a}_p^\dagger \hat{a}_q \tag{80}$$

By expressing the spin-adapted one-body operators and paired two-body operators $\hat{\kappa}_{pq}^{(1)}$ and $\hat{\kappa}_{pq}^{(2)}$ respectively, one can construct the building blocks of the QNP ansatz:

$$\hat{\kappa}_{pq}^{(1)} = \hat{E}_{pq} - \hat{E}_{qp} \quad \text{and} \quad \hat{\kappa}_{pq}^{(2)} = \hat{E}_{pq}^2 - \hat{E}_{qp}^2 \tag{81}$$

Following the formalism presented in [113], the spin-adapted one-body and paired two-body operators can be used to illustrate the operators prepared by the QNP ansatz:

$$\hat{U}_{pq}^{(m)} = \exp(\theta_{pq,1}^{(m)} \hat{\kappa}_{pq}^{(1)}) \exp(\theta_{pq,2}^{(m)} \hat{\kappa}_{pq}^{(2)}) \tag{82}$$

The QNP approach typically uses the Hartree-Fock state $|\Psi^{\mathrm{HF}}\rangle$ as an initial starting guess for a wavefunction, as it is computationally straightforward to prepare and satisfies the criterion of quantum number preservation [11]. As shown in Section 1.1.1, the Hartree-Fock state represents a single Slater determinant, thus the QNP ansatz requires multiple layers in order to prepare circuits whose depth provide enough expressivity for the representation of correlation.

Recent efforts into furthering the ideas of the QNP ansatz include the **tiled Unitary Product State** (tUPS) by modifying the structure of the wavefunction prepared as [113]:

$$|\Psi^{\mathrm{tUPS}}\rangle = \prod_{m=1}^{L} \left(\prod_{p=1}^{A} \hat{U}_{2p+1,2p}^{(m)} \prod_{p=1}^{B} \hat{U}_{2p,2p-1}^{(m)} \right) |\Psi^{\mathrm{HF}}\rangle \tag{83}$$

where $A = (N-2)/2$ or $(N-1)/2$ and $B = N/2$ or $(N-1)/2$ for even or odd number of spatial orbitals N respectively. The operators $\hat{U}_{pq}^{(m)}$ prepared by the tUPS ansatz have a form very similar to the QNP approach, containing an additional one-body operator to allow for more freedom in axis rotations in the two-electron subspace spanned by the $\hat{U}_{pq}^{(m)}$ operators [113].

$$\hat{U}_{pq}^{(m)} = \exp(\theta_{pq,1}^{(m)} \hat{\kappa}_{pq}^{(1)}) \exp(\theta_{pq,2}^{(m)} \hat{\kappa}_{pq}^{(2)}) \exp(\theta_{pq,3}^{(m)} \hat{\kappa}_{pq}^{(1)}) \tag{84}$$

This avenue requires more free parameters per building block, and approaches beyond the HF state as an initial state might require more sophisticated error mitigation strategies to prepare the Perfect Pairing or Orbital Optimized states prior to the ansatz application. The Perfect Pairing approximation draws inspiration from valence bond theory, by interleaving the qubit registers depicting occupied and empty spatial orbitals to achieve maximal entanglement in the shallowest circuit possible in terms of fabric layers [113]. Orbital Optimization on the other hand, relies on the transformation of the molecular integrals used to construct the state in the first place through the help of a classical computer and the optimization of those integrals via the orbital gradient and the two-body density matrix which is acquired for the energy estimator [113].

3.4.2 Tensor factorization-based measurement methods

Conventional strategies for measuring the energy of a Hamiltonian such as the grouping of Pauli terms that are jointly measurable are able to be applied generally, but often yield an unsustainable number of distinct number of measurement bases and quantum circuit shots to measure them within a specified accuracy. In the special case of an electronic structure molecular Hamiltonian, the Hamiltonian has symmetric features that one is able to leverage in order to compress and through **tensor factorization techniques**, reformulate the Hamiltonian in such a way that it retains the original information while reducing the quantum resources needed for measurement. As these techniques focus heavily on the electron integrals involved in a second quantized electronic structure Hamiltonian, a reformulation of Eq. 14 to further highlight the role of electron integrals is shown below [32]:

$$\hat{H} = E_c + \sum_{pq}(p|\hat{h}_c|q)\hat{E}_{pq} + \frac{1}{2}\sum_{pqrs}(pq|rs)(\hat{E}_{pq}\hat{E}_{rs} - \delta_{qr}\hat{E}_{ps}) \tag{85}$$

where $\hat{E}_{pq} \coloneqq \hat{p}^\dagger\hat{q} + \hat{\bar{p}}^\dagger\hat{\bar{q}}$ is the singlet excitation operator, ϕ the spacial molecular orbitals, with the symmetric one-electron integrals $(p|\hat{h}_c|q)$ and the 8-fold symmetric two-electron integrals $(pq|rs)$ being expressed as:

$$\begin{aligned}(p|\hat{h}_c|q) &= \int \phi_p^*(\boldsymbol{r})\left(-\frac{1}{2}\nabla^2(\boldsymbol{r}) - \sum_m \frac{Z_m}{\boldsymbol{r} - r_m}\right)\phi_q(\boldsymbol{r})d\boldsymbol{r} \\ (pq|rs) &= \iint \phi_p^*(\boldsymbol{r}_1)\phi_q(\boldsymbol{r}_2)\frac{1}{r_{12}}\phi_r^*(\boldsymbol{r}_1)\phi_s(\boldsymbol{r}_2)d\boldsymbol{r}_1 d\boldsymbol{r}_2\end{aligned} \tag{86}$$

Geared towards the tackling of such challenges, tensor factorization techniques offer a way of lowering the quantum resources needed. Depending on the type of compression and the symmetries exhibited by the Hamiltonian, there are varying tensor factorization schemes, such as Tensor Hypercontraction [115], explicit double factorization and compressed versions of double factorization of the Hamiltonian [54, 55].

When considering an explicit double factorized representation [54], the effective tensor $\mathcal{F}_{pq}$ contracted against the one-body operators E_{pq} is shown to be [32]:

$$\mathcal{F}_{pq} \coloneqq (p|\hat{h}_c|q) - \frac{1}{2}\sum_r (pr|qr) + \sum_r (pq|rr) = \sum_k U_{pq}^{\varnothing}\mathcal{F}_k^{\varnothing}U_{qk}^{\varnothing} \tag{87}$$

and two-electron integral tensor can be doubly factorized when $(pq|rs)$ is real and 8-fold symmetric as [32]:

$$(pq|rs) = \sum_{t=1}^{n_t} V_{pq}^t g_t V_{rs}^t = \sum_{t=1}^{n_t}\sum_{kl=1}^{n} U_{pq}^t U_{qk}^t Z_{kl}^t U_{rl}^t U_{sl}^t \tag{88}$$

where the $^{\varnothing}$ notation denotes a single expansion index, as the eigendecomposition of $\mathcal{F}_{pq}$ into tensors $U_{pk}^{\varnothing}$ and U_{qk} is analytical and Z_{kl}^t are coefficient tensors containing the compressed two-electron information [55].

The Hamiltonian under Regularized Compressed Double Factorization (RC-DF) can be written

explicitly with $n_t \leq n(n+1)/2$ where n the number of spatial orbitals, as [32]:

$$\begin{aligned}\hat{H} = \mathcal{E} - \frac{1}{2}\sum_{k} \mathcal{F}_k^{\varnothing} U^{\varnothing} \left(\hat{Z}_k + \hat{Z}_{\bar{k}}\right) (U^{\varnothing})^\dagger \\ + \frac{1}{2}\sum_{t=1}^{n_t}\sum_{k,l=1}^{n} Z_{kl}^t U^t \left(\hat{Z}_k\hat{Z}_l - \delta_{kl} + \hat{Z}_k\hat{Z}_{\bar{l}} + \hat{Z}_{\bar{k}}\hat{Z}_l + \hat{Z}_{\bar{k}}\hat{Z}_{\bar{l}} - \delta_{\bar{k}\bar{l}}\right) (U^t)^\dagger\end{aligned} \tag{89}$$

where $U^{\varnothing}, U^t$ denote orbital rotations for every $t \in [1, n_t]$, $\hat{Z}_k, \hat{Z}_{\bar{k}}$ are Pauli operators that are applied on qubits $2k$ and $2k+1, k \in [0, n-1]$ and the offset $\mathcal{E}$ is:

$$\mathcal{E} = E_c + \sum_p (p|\hat{h}_c|p) + \frac{1}{2}\sum_{pq}(pp|qq) - \frac{1}{4}\sum_{pq}(pq|pq) \tag{90}$$

A central advantage of the double factorization of electron integrals is the fact that it is possible to truncate the number of terms t that participate in the sum (also referred to as leafs), which if left untreated, causes the variance of the energy estimator to fluctuate [32]. In order to ameliorate these fluctuations, in the cost function that optimizes for the compressed representation of the Hamiltonian [55], the RC-DF procedure adds a regularization term, resulting in a nested optimization; one for retrieving the orbital rotations' generators $U_{pq}^t \coloneqq \exp(X_{pq}^t)$ and another for the coefficients Z_{kl}^t.

The error Δ_{pqrs} up to which one compresses the two-electron integral tensor is given by the Frobenius norm expressed below [32]:

$$\Delta_{pqrs} \coloneqq \frac{1}{2}\left\| \sum_{t=1}^{n_t}\sum_{kl=1}^{n} U_{pq}^t U_{qk}^t Z_{kl}^t U_{rl}^t U_{sl}^t - (pq|rs) \right\|_{\mathcal{F}}^2 \tag{91}$$

The cost function of the RC-DF procedure for arriving at the optimal compression of the Hamiltonian is the following [32]:

$$C(X, Z) \coloneqq \frac{1}{2}\|\Delta_{pqrs}\|_{\mathcal{F}}^2 + \sum_{tlk} \rho_{tkl} |Z_k^t l|^\gamma \tag{92}$$

where $\mathcal{F}$ denotes a Frobenius norm, $\rho_{tkl} \geq 0$ the regularization tensor of weights penalizing large $|Z_{kl}^t|$ matrices, with a uniform $\rho_{tlk} = \rho$ being a usual penalty choice and $\gamma = 1$ or $\gamma = 2$ when performing L1 and L2 regularization respectively [32].

Both the explicit (X-DF) and the truncated sum of terms (C-DF/RC-DF) in the Hamiltonian of Eq. 89 tensor factorizations can be taken one step further through a technique called Fluid Fermionic Fragments (FFF) [116] in their efforts to reduce experimental repetitions, by building energy estimators with lower variance. The FFF technique achieves this by reorganizing quadratic creation and annihilation operators termed *fluid fermionic fragments* over different parts of the energy estimators with some fragment coefficient. Through this combination of techniques, when compared to Pauli grouping schemes, the energy estimator exhibits a lower mean square error in the energy estimator, allowing for fewer shots to be spent, thus speeding up the optimization process [32].

Finally, another adoption advantage of the RC-DF method over tensor hypercontraction or other double factorization schemes is its compatibility with error corrected quantum algorithms via qubitization, owing to its reduced overall depth due to the low lambda factor and thus, run time of phase estimation routines [32].

3.5 Simulating Chemistry with Fermionic Optical Superlattices

Bibliographic information

Gkritsis, F., Dux, D., Zhang, J., Naman, J., Gogolin, C., Preiss, P.M. (2025, January). Simulating Chemistry with Fermionic Optical Superlattices. PRX Quantum 6, 010318 Retrieved 2025-01-25, from: `https://journals.aps.org/prxquantum/abstract/10.1103/PRXQuantum.6.010318`
DOI: 10.1103/PRXQuantum.6.010318

Summary

We show that quantum number preserving Ansätze for variational optimization in quantum chemistry find an elegant mapping to ultracold fermions in optical superlattices. Using native Hubbard dynamics, trial ground states for arbitrary molecular Hamiltonians can be prepared and their molecular energies measured in the lattice. The scheme requires local control over interactions and chemical potentials and global control over tunneling dynamics, but foregoes the need for optical tweezers, shuttling operations, or long-range interactions. We describe a complete compilation pipeline from the molecular Hamiltonian to the sequence of lattice operations, thus providing a concrete link between quantum simulation and chemistry. Our work enables the application of recent quantum algorithmic techniques, such as Double Factorization and quantum Tailored Coupled Cluster, to present-day fermionic optical lattice systems with significant improvements in the required number of experimental repetitions. We provide detailed quantum resource estimates for small non-trivial hardware experiments.

Contributions

The author contributed to this work: (1) Made substantial contributions in deriving the explicit decomposition rules for the $\hat{U}^t$ and $\hat{U}^{pt}$ gates for various decomposition schemes (2) Derived, implemented and tested numerically the merging rules of sequential $\hat{U}^t$ to enable reduction of circuit depth (3) Implemented the $\hat{U}^t$ and $\hat{U}^{pt}$ gates as fermionic and as qubit gates, to allow testing for different simulators (4) Created simulation pipelines for the decompilation of high-level gates to operations native to ultracold fermions, merging of gates and addition of gate parameter noise for the study of sensitivity to noise (5) Proposed and prepared numerical benchmarks for test cases (6) Setting up code infrastructure for carrying out the numerical and error analysis of the test cases, (7) wrote significant parts the manuscript.

[DOI: https://doi.org/10.1103/PRXQuantum.6.010318]

3.6 Quantum computing on trapped ion hardware

Trapped-ion quantum information processing is an architecture that is actively pursued for the realization of a scalable quantum computer, with the achievement of various milestones toward fault tolerant quantum information processing [117]. The experimental realization of a trapped-ion architecture can be done through the use of segmented Paul traps to create trapping regions that act as hosts for the ions used [118]. By adjusting the potential of the trap, reconfigurations of the ion crystals is possible to create dynamic registers capable of storage and process of quantum information, while simultaneously providing effective all-to-all connectivity, in stark contrast of most superconducting platforms available now.

That structural advantage, coupled with the high gate fidelities of a trapped-ion quantum processor allows for the exploration of deeper algorithms with higher wall-time execution times, paving the way for the mapping of electrons of larger molecules through fermion-to-qubit mapping as such the Jordan-Wigner or the Bravyi-Kitaev transformations [119, 35]. Owing to the limits of the classical methods in their capacity to perform simulations with regards to the number of spatial orbitals simulated, despite having reached the required raw physical qubit counts, platforms such as trapped ions need to address two main challenges in order to grow past these classical limits and provide quantum advantage.

A pertinent challenge that limits the capabilities of trapped atom quantum computing devices is the noise incurred by imperfect gates and measurements alike. This leads to the limitation of the gate depth of the algorithms that are able to be executed and the accuracy of the results obtained from these experiments. Recent progress has been made to construct algorithms for quantum chemistry simulations that make use of the implementation of block-invariant symmetry-shifted Tensor Hypercontraction (BLISS-THC) [120] and quantum diagonalization algorithms based on subspaces constructed by quantum Krylov states [121]. However, even with these advancements, the noise levels of currently available hardware render quantum advantage beyond reach.

Another difficult engineering challenge impeding the growth of architectures based on neutral atms and trapped ions alike, is the challenge of the so-called clock speed, used to describe the time for a single elementary logic operator of these quantum devices to be executed, which is orders of magnitude slower compared both to classical computers, as well as superconducting platforms. In order for variational algorithms to be a viable strategy for tackling molecular electronic structure problems, there is still much work needed to push the field forward, since Variational Quantum Eigensolver algorithms was regarded as a trade-off between the long coherence times needed by non-variational algorithms like phase estimation and the number of repeated executions required for the shorter circuits of VQE but this belief has been recalibrated towards the utilization of VQE algorithms to construct quantum-ready pipelines for quantum chemistry simulations as well as provide insights into the performance characteristics of a quantum processor.

3.6.1 Infrastructure for interface with a trapped ion quantum processor

One challenge in the implementation of hybrid computing algorithms is the implementation of a programming interface enabling their execution, connecting the classical high performance computers with a trapped-ion quantum processor. For the quantum processor to be seamlessly integrated in the quantum chemistry workflow in such a way that there is no need for external experimental control, it is necessary to provide automated communication of the quantum algorithms to be executed in the form of quantum circuits, along with translation in optimized hardware commands and back communication of the measurement results. External experimental control would be far too time-consuming and error-prone, debilitating any prospect of scaling experiments to involve the number of qubits required for reasonable quantum chemistry simulations.

To make use of the high connectivity and high gate fidelities of a trapped-ion quantum computer provided by the Johannes Gutenberg Universität Mainz. for the existing as well as upcoming new, improved trapped ion quantum computing setups with segmented Paul traps, a PennyLane plugin device in the Python programming language was developed for the purpose of remote testing, control and experiment execution of circuits [122].

In collaboration with the Johannes Gutenberg Universität Mainz, the PennyLane device [122] was implemented to act as an interface with both the back-end simulator of the trapped ion device, as well the trapped ion quantum processor itself. This allowed for the implementation and deployment of experiments, as well as diagnosing problems of the operations performed by the device from the author of the thesis.

The operations of Orbital Rotation and Single Excitation were studied in an effort to construct the fundamental components of a QNP ansatz for the Orbital Rotation and the Pair Exchange gates through operations that are native to the quantum processor which consist of the operator set of RX, RY, RZ, SWAP and IsingZZ, denoted in this thesis as I_{ZZ} in circuits as a shorthand notation. The definition of the IsingZZ gate is given by the definition in Eq. 93:

$$\mathrm{IsingZZ}(\phi) = \exp\left(-i\frac{\phi}{2}(Z \otimes Z)\right) = \begin{bmatrix} e^{-i\phi/2} & 0 & 0 & 0 \\ 0 & e^{i\phi/2} & 0 & 0 \\ 0 & 0 & e^{i\phi/2} & 0 \\ 0 & 0 & 0 & e^{-i\phi/2} \end{bmatrix} \tag{93}$$

With all the native components defined, the construction of more complex building blocks is now possible. The decomposition for the Orbital Rotation that was used in all subsequent experiments is depicted in Fig. 15:

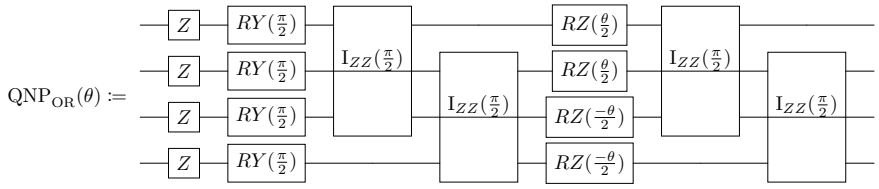


Figure 15: Decomposition of the Orbital Rotation gate into operations native to the Johannes Gutenberg Universität Mainz processor.

Due to the size of the available quantum registers in the quantum device reliably available for control, the Single Excitation was used as a stepping stone towards the implementation of the Double Excitation operation needed to construct the QNP_{PX}. The decomposition of the Single Excitation operation denoted as $U_{SE}(\theta)$ in circuits that was used is shown in Fig. 16:

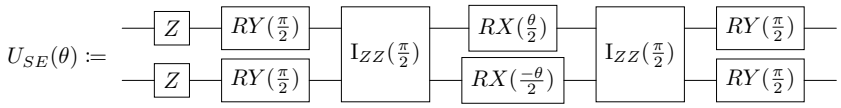


Figure 16: Decomposition of the Single Excitation gate into operations native to the Johannes Gutenberg Universität Mainz processor

This systematic way of interfacing with the quantum device and its simulator enabled the identification of issues in the correct calibration of each gate via unit tests, by crosschecking the results against basic, minimal circuits of size which is easily simulated by classical computers. Moreover, the infrastructure developed allowed for the seamless benchmarking of local simulators, as well simulators accessed via the application program interface (API) of the trapped ion device stack and the quantum device.

The importance of efficiency in the use of quantum resources is capital in all current generation quantum computing platforms and quantum processors based on trapped ions are no exception. The software interface that was developed, along with the function of communication of circuits and results between a high performance cluster and the quantum device or its simulator, was also developed with the goal of reducing gate counts and maximizing fidelity. By taking into consideration the calibration of the gates as well as laser intensity and pulse duration needed to perform each operation, it is feasible to commute and cancel out redundant operations and substitute those that would have inferior fidelity with others that have exactly the same effect on the ions [123].

The role of parallelization of operations remains central in the trapped ion architecture by examining the operations that can be performed in parallel without incurring increased shuttling or register reconfiguration operations. By pinpointing and batching operations that can be performed simultaneously, it is possible to not only improve execution speed but also reduces the cumulative error that might otherwise result from excessive sequential operations. In trapped-ion architectures, owing to the fact that the process of mapping qubits onto ions is a combinatorial problem that scales exponentially with the number of qubits required, heuristic approaches are utilized based

on the initial interactions between qubits, working in tandem with dynamic scheduling algorithms that determine optimal register reconfigurations. These dynamic heuristics move and swap ions based on the next gate operations, effectively reducing the overhead associated with physical reconfiguration of the ion register.

Continuing, upon building these circuits, a crucial piece is the use of symbolic parameters. As iterative hybrid algorithms maintain the same circuit structure with only gate parameter differing between different stages of the computation, the symbolic parameters act as placeholders for the compilation processes. Upon the compilation of the circuit, the actual numerical values to be used are substituted into the circuit, ensuring this way that the topology of the circuit remains the same throughout the experiment, allowing for more consistent behavior and drastically reducing the overhead that would be incurred from a full recompilation of the same circuit for every different set of parameters update.

3.6.2 Improvements on the Echo Verification protocol

Dirichlet method of evaluating expectation value

Instead of the usual (direct) measurement of the phase function at $\pi/2$ time [106] for the implementation of the exact unitary one is trying to measure as in Eq. 72, a new method was explored based on Dirichlet kernels when working with trapped ions. This stemmed from the need to measure the derivative of the phase function at $t = 0$, taking advantage of parameter shift rules to efficiently calculate the derivative at the desired time. In this thesis, only the case where the matrices have ladder-like spectra that have equi-distant eigenvalues are examined, with the possibility of extension of this approach to cases where this condition is not imposed exists in other work [63].

Since the observables that are usually involved in NISQ circuits are unitary, upon the diagonalization of the matrix describing the observables, the number of non-degenerate eigenvalues R, as well as their spacing l is obtained. Primed with that information, the formula for retrieving the set of appropriate times $\{t_{R_i}\}$ to measure the imaginary part of the phase function is described as:

$$t_{R_i} = \frac{(2(R_i+1)-1)\pi}{lR}, \quad \text{for } R_i \in 1, 2, \ldots, R \tag{94}$$

In order for the procedure to yield an expectation value, the imaginary part of the measured phase function at the appropriate times is then multiplied by a coefficient c_{R_i}, where the set of coefficients $\{c_{R_i}\}$ can constructed as:

$$c_{R_i} = \frac{(-1)^{R_i} l}{4R \sin^2(\frac{(2(R_i+1)-1)\pi}{4R})}, \quad \text{for } R_i \in 1, 2, \ldots, R \tag{95}$$

Time mirroring

To mitigate further the effects of errors in the measurement process, an additional technique was developed called *time mirroring*, by using the antisymmetric nature with respect to time evolution of the imaginary part of the phase function. Whether using the Dirichlet or the direct method of evaluating the expectation value via the phase function, a set of times for the phase function to be measured, as well as the coefficients for these measurements in the calculation of the expectation value is determined. These times are in the interval $t_{R_i} \in [0, \pi]$ and time mirroring doubles the size of the set $\{t_{R_i}\}$ by also including every time $-t_{R_i}$ of the original set.

Additionally, to account properly for the contribution of each time when constructing the expectation value, the coefficients assigned to these times also need to be modified. From the original set of coefficients $\{c_{R_i}\}$ that are assigned to each time t_{R_i}, the coefficient is then halved, with $c_{R_i}/2$ assigned to the $+t_{R_i}$ time, and $c_{R_i}/2$ assigned to the $-t_{R_i}$ time of the phase function. The cardinality of $\{t_M\}$ and $\{c_m\}$, meaning the number of elements contained in these sets is denoted by $|\cdot|$, with $|\{t_M\}| = 2|\{t_{R_i}\}|$ and $|\{c_M\}| = 2|\{c_{R_i}\}|$.

This makes the set of times $\{t_M\}$ and coefficients $\{c_M\}$ of time mirroring:

$$\begin{aligned} \{t_M\} &= \{-t_{R_i}\} \cup \{t_{R_i}\} \\ \{c_m\} &= \{\frac{c_{R_i}}{2}\} \end{aligned} \tag{96}$$

This allows the echo verification method to be more robust in incoherent errors of the quantum device since the points which the phase function is evaluated are more, however, it is not able to account for coherent errors of the quantum device, either in the calibration of the gates or coherent measurement errors.

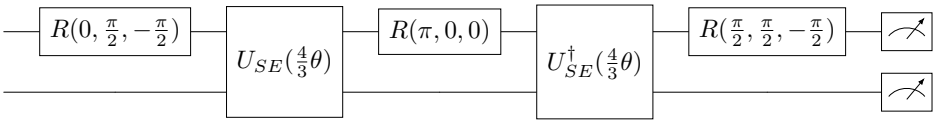
3.6.3 Experimental results from a trapped ion device

These efforts in building the necessary infrastructure as well as the developments to make the echo verification method more robust, have culminated in the first connection tests and execution of small quantum circuits to be successfully performed on the trapped ion device of Johannes Gutenberg Universität Mainz, with experimental data collected by the author of the thesis. The received quantum circuit was translated into the native gate set and then optimized, in order to significantly reduce the total number of gate operations using a quantum circuit compiler. The current generation of hardware from Johannes Gutenberg Universität Mainz is currently limited to operate with multiple single ions and two-ion crystals, leading to increased shuttling overhead resulting in increased time required for an experiment of size larger than 2 qubits or the outright loss of ions during shuttling. A shuttling compiler has been developed to find the optimal qubit-to-ion mapping and then add the qubit position and shuttling operation information in between a gate operation sequence [123].

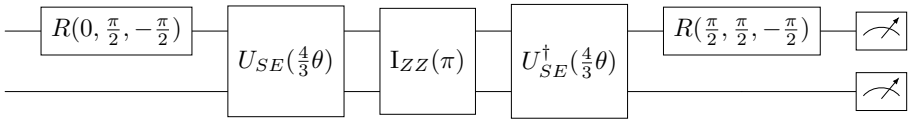
The single qubit rotations were implemented with the use of a general rotation gate $R(\phi, \theta, \omega)$, which is easily decomposable to basic RY and RZ gates with appropriate angles that are native to the device apparatus of the Johannes Gutenberg Universität Mainz to create the desired single qubit rotation as follows:

$$R(\phi,\theta,\omega) = RZ(\omega)RY(\theta)RZ(\phi) = \begin{bmatrix} e^{-i(\phi+\omega)/2}\cos(\theta/2) & -e^{-i(\phi-\omega)/2}\sin(\theta/2) \\ e^{-i(\phi-\omega)/2}\sin(\theta/2) & e^{i(\phi+\omega)/2}\cos(\theta/2) \end{bmatrix} \tag{97}$$

With the current generation hardware, experiments were performed to measure the expectation value of different observables using the Echo Verification method with the added improvements discussed in Section 3.6.2. This tested the performance of the device with respect to both the integrity of the end-to-end pipeline that was developed, as well as the quality of the results obtained by comparing them to the values obtained from circuits that are easily simulatable classically. The circuits that were used were 2 ion qubit systems using Z_0 and Z_0Z_1 for observables as shown below, with decomposition in Fig. 16 for U_{SE} used experimentally, but omitted here for circuit brevity:



(a)



(b)

Figure 17: Circuits used for the experiments with the Johannes Gutenberg Universität Mainz trapped ion device, depicting the circuits for 2 qubits with observables Z_0 and Z_0Z_1 in (a) and (b) respectively, with the omission of the depiction of the U_{SE} decomposition.

Using the circuits depicted in Fig. 17, the expectation values of Pauli Z on the first qubit were measured, labeled 0, as well as Pauli Z on the first and second qubit, labeled 0 and 1 respectively. A variety of measurement method protocols were used, to establish the performance of echo verification based measurement methods, keeping in mind that these methods become more potent as more qubits are available for verification. With that motivation in mind, these experiments aimed to pinpoint the transition where compared to directly measuring the observable term-wise, echo verification measurement protocols would yield expectation values that are closer

to ideal benchmarks. These benchmarks are taken to be expectation values that one can retrieve from classically simulating the circuit, as they are analytically computed.

The expectation values for the observable Z_0 on 2-qubit trapped ion hardware from Johannes Gutenberg Universität Mainz with the different means these expectation values were calculated over a number of points in parameter space that parametrized the U_{SE} gate described in the circuit of Fig. 17a are illustrated in Fig. 18:

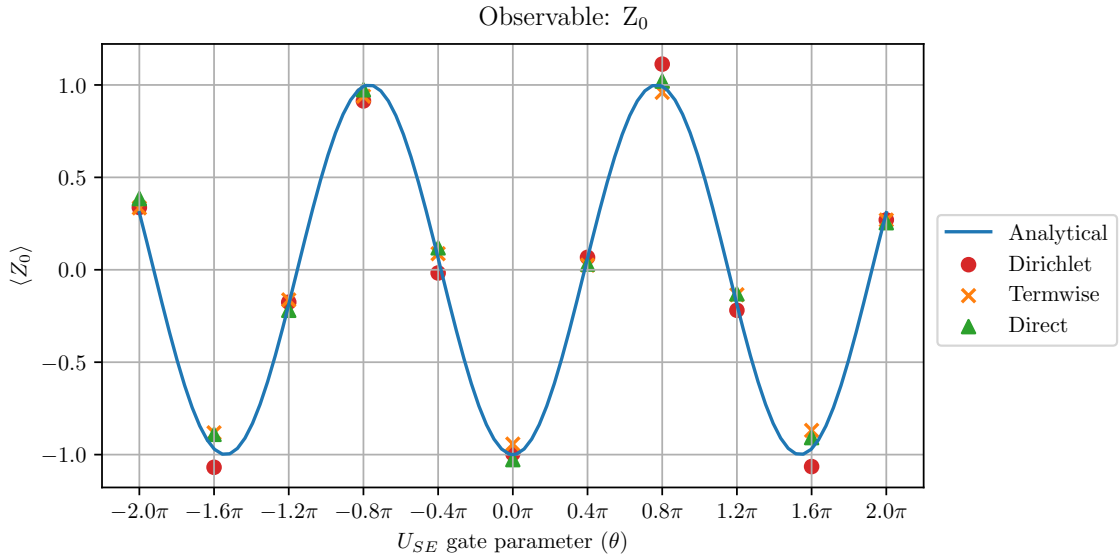


Figure 18: Expectation values for the 2 qubit circuit of Fig. 17a, for the observable Z_0, using various methods for the calculation of the expectation values.

Using the same circuit and configuration in terms of experiment repetitions (4000 shots), the procedure was repeated with Z_0Z_1 as the observable. The expectation values of Z_0Z_1 for a 2-qubit system are depicted in Fig. 19:

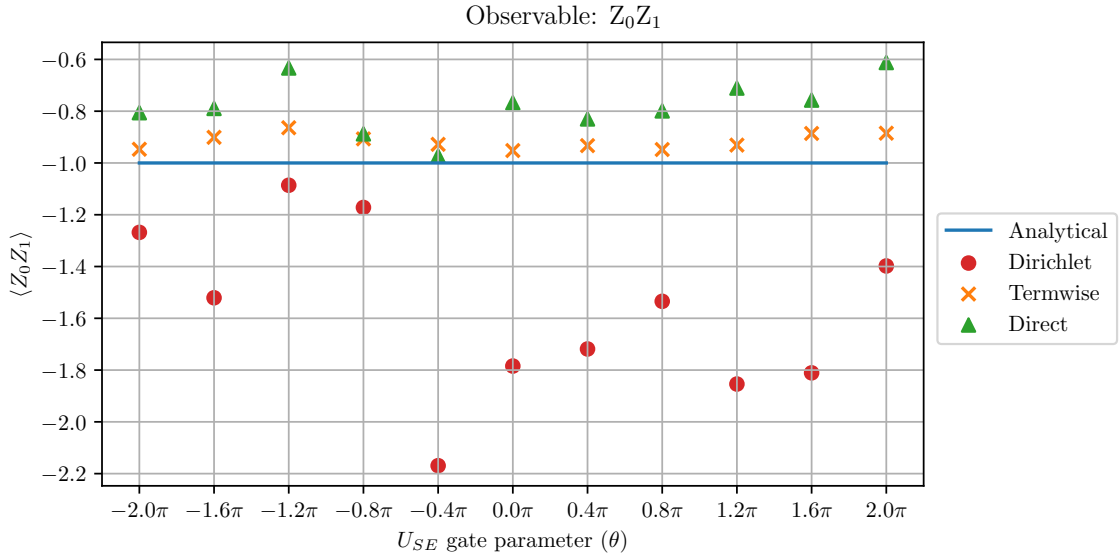


Figure 19: Expectation values for the 2 qubit circuit of Fig. 17b, for the observable Z_0Z_1, using various methods for the calculation of the expectation values.

The reasons behind the subpar performance of the Dirichlet measurement method on the estimation value of the observable Z_0Z_1 are not clear. Instances of such behavior from the quantum device were observed when experiments of magnetic nature were being conducted near the experi-

mental apparatus or heavy ion loss had occurred but its definitive root cause was not fully identified and resolved in the duration of this work and is part of the ongoing efforts for improvement of the quantum device of the Johannes Gutenberg Universität Mainz trapped ion devices.

For the echo verification protocol, the 4000 shots that were allotted to the quantum device, half of them (2000) were used for the calculation of fidelity. As in Eq. 72, the fidelity is the measure of the probability of the experimental setup to perform accurately the needed operations for returning to the appropriate initial state, thus the slope of the imaginary part of the phase function $\mathrm{Im}(g(t))$is evaluated at $t = 0$ as $\mathrm{Im}(g'(0))$ is rescaled by the fidelity in that point of measurement. The results of the phase functions and by extension, the expectation values, is heavily influenced by the fidelity of the device and are shown in Fig 20 below:

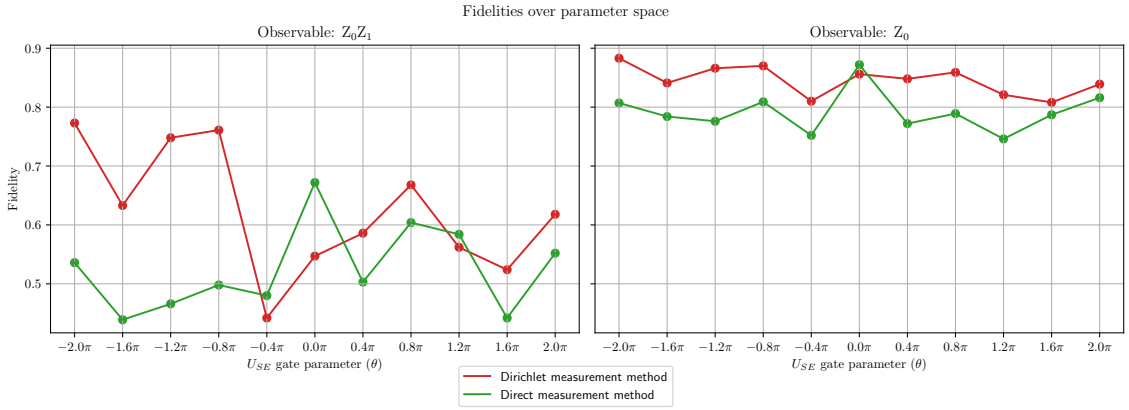


Figure 20: Fidelity metric for Echo Verification measurement method for 2 qubit experiments over the parameter space of the U_{SE} gate used for different observables.

The rest of the 2000 shots were used for the calculation of the expectation values through the evaluation of the phase function at appropriate times according to the echo verification protocol used. The phase functions that were measured and produced the expectation values in Fig. 18 and Fig. 19, showing the times that each phase function was measured to construct the expectation value at that point by taking the slope of the phase function at time $t = 0$:

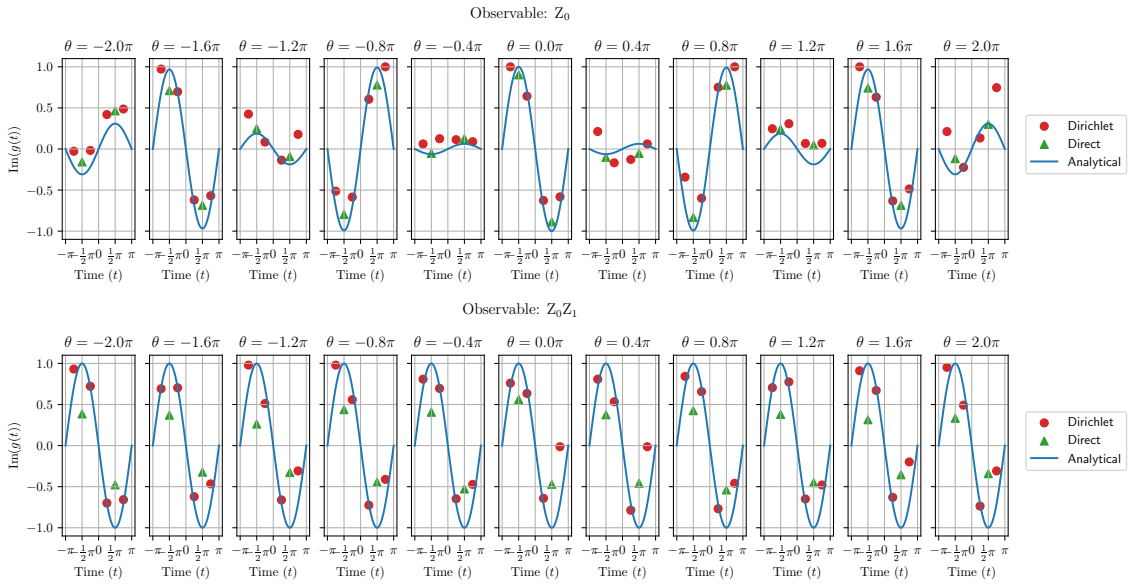


Figure 21: Phase function measurements for the calculation of expectation values for 2 qubits with the direct and Dirichlet measurement protocols for different parameters in the ansatz.

Going one step further, the same procedure was performed for 3 qubit systems with the circuit depicted in Fig. 22 with 6000 total shots for both observables on the same qubits. Although

seemingly small, this 3-qubit circuit was selected as a stepping stone towards the realization of a Double Excitation gate, which is critical for the construction of QNP fabric gate structures [49], while also allowing for the impact of ion shuttling and ion loss to become pronounced.

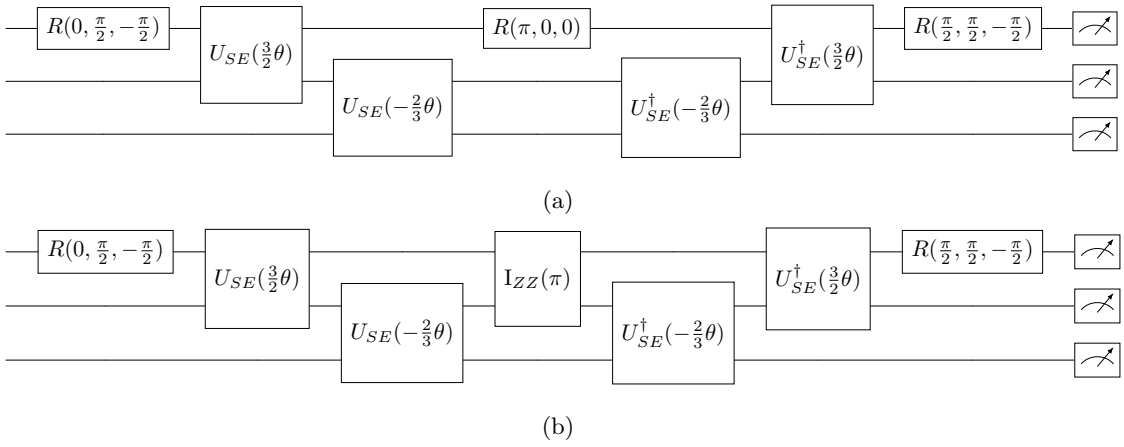


Figure 22: Circuits used for the experiments with the Johannes Gutenberg Universität Mainz trapped ion device, depicting the circuits for 3 qubits with observables Z_0 and Z_0Z_1 in (a) and (b) respectively, with the omission of the depiction of the U_{SE} decomposition.

These measurements provided the following expectation values for the observable Z_0 on a 3-qubit circuit using the term-wise, direct and Dirichlet methods for calculating the expectation values over different values θ parametrizing the U_{SE} gates of the circuit is shown in Fig. 23:

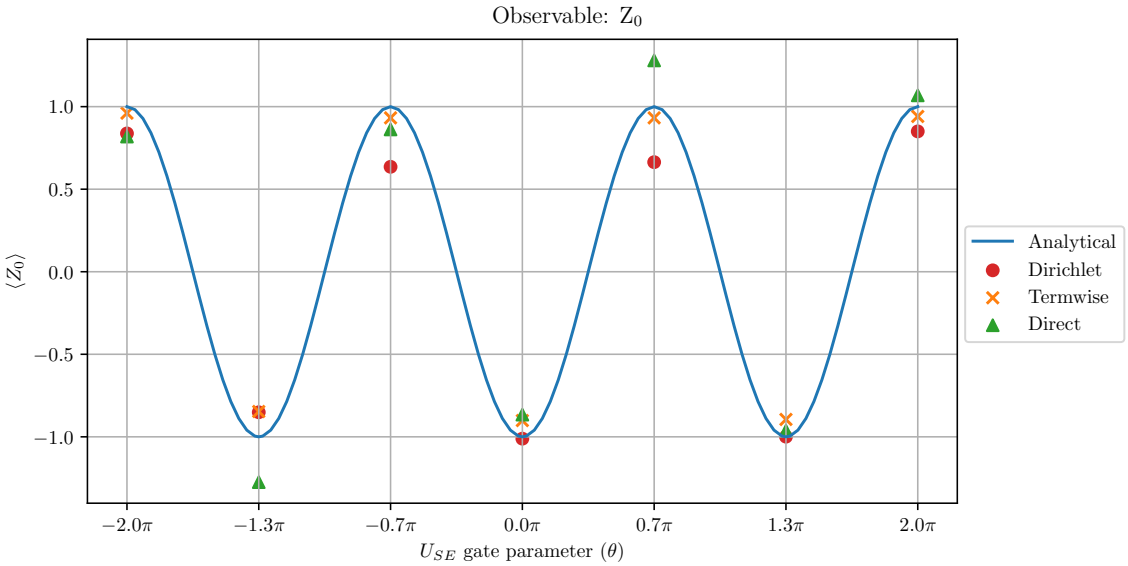


Figure 23: Expectation values for the 3 qubit circuit of Fig. 22a, for the observable Z_0, using various methods for the calculation of the expectation values.

As before, the same circuit and configuration was utilized, with the same overall shot budget of 6000 shots for the quantum device, with measurements of the expectation value over the same values for the θ parameter of the U_{SE} gates, as in the previous Z_0 case. The expectation values that were constructed using measurements from the device for the Z_0Z_1 observable in 3 qubits are illustrated in Fig. 24:

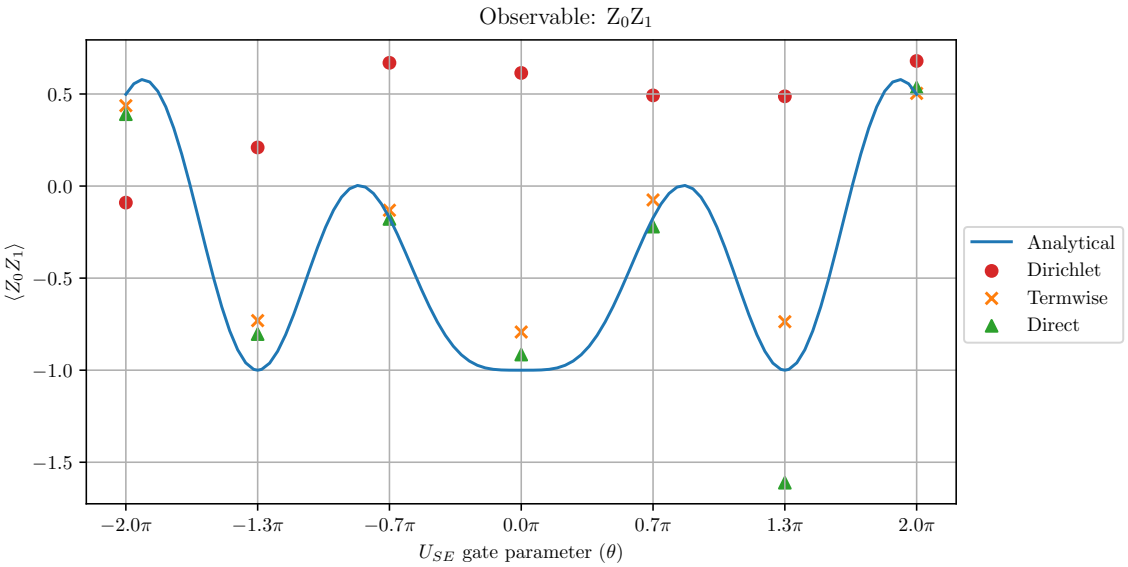


Figure 24: Expectation values for the 3 qubit circuit of Fig. 22b, for the observable Z_0Z_1, using various methods for the calculation of the expectation values.

Calculating the expectation value of Z_0Z_1 with the Dirichlet protocol proved challenging, as the performance of the device was not stable enough to yield expectation values that were close to those of an analytic device. The way the data these expectation values was collected measurement protocol was by grouping the evaluations for each measurement method separately for one another and then requesting the evaluation of the symbolically identical circuit at different parameters, as swapping between measurement methods would mean the recompilation of the circuit each time, greatly increasing the execution time.

Using the Dirichlet measurement protocol, as well as the direct measurement protocol, the fidelity for both protocols of each expectation value, for which 3000 shots were invested for each point over the parameter space for these experiments, are shown below in Fig. 25:

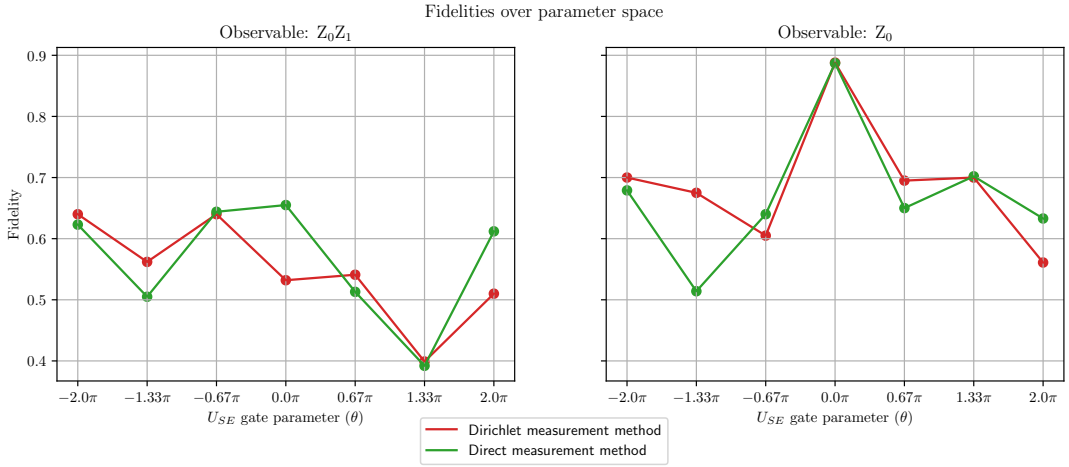


Figure 25: Fidelity metric for Echo Verification measurement method for 3 qubit experiments over the parameter space of the U_{SE} gates used for different observables.

The rest of the 3000 shots that were allotted to the trapped ion quantum device were invested into the measurement of the phase function at various points in time, depending on the protocol. The phase functions on which the expectation values in Fig. 23 and Fig. 24 are derived from are shown in Fig. 26:

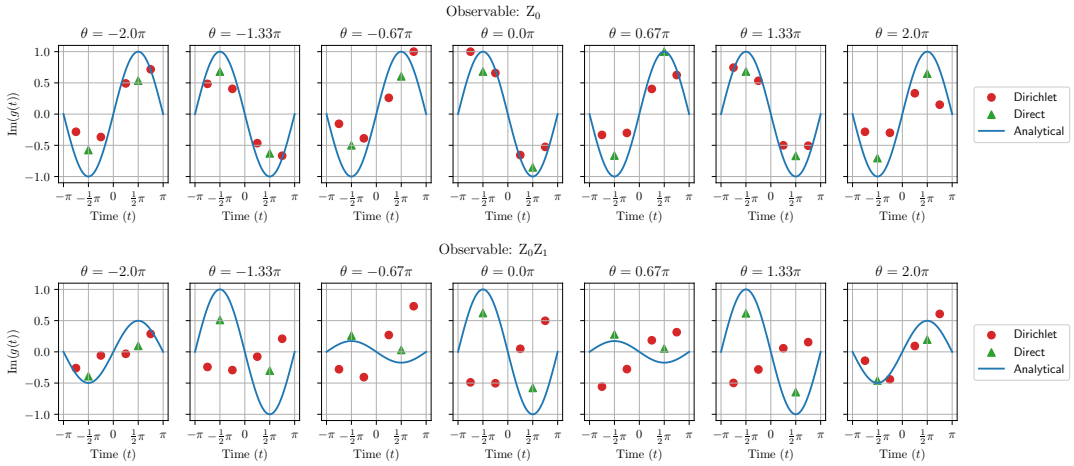


Figure 26: Phase function measurements for the calculation of expectation values for 3 qubits with the direct and Dirichlet measurement protocols for different parameters in the ansatz.

While the fidelity for each point in parameter space was comparable between the two measurement protocols for both observables selected, the results for the phase functions of the Z_0Z_1 observable and by extension the expectation values derived from them, had significant discrepancies from both the direct counterpart, as well as the analytical value expected.

The implementation of this end-to-end process for seamless circuit design and execution by the trapped ion device allows for the easy adaptation of larger setups thanks to the compiler software developed [123], allowing for smooth integration into quantum chemistry simulations workflows.

The outlook for next-gen setups from the Johannes Gutenberg Universität Mainz is to create trapped ion crystals that allow for the individual addressing of up to ten ions, since the current single and two ion crystals lead to significant shuttling overhead. The next generation of devices will enable not only the reduction of shuttling overhead, reducing execution duration times, but also an increase in the number of ions available for the fermionic electrons of larger molecules to be mapped onto them, leading to the simulation chemically more interesting systems, closer to the regime where classical methods encounter size limitations with the prospect of some day surpassing them.

4 Discussion

The research presented in this thesis introduces a number of significant developments of chemistry simulations within the framework of the Variational Quantum Eigensolver (VQE), particularly pertaining to its performance and applicability on noisy intermediate scale quantum (NISQ) devices. More specifically, new optimization strategies are presented and thoroughly tested numerically, including the Conjugate Model Search (CMS) optimizer and the genetic hyperparameter tuning procedure for optimizers used in hybrid quantum-classical algorithms. Two new additions to the echo verification measurement protocol are proposed and explored, as they were then used to extract expectation values from real trapped ion hardware on the Johannes Gutenberg Universität Mainz platform through the developed software stack. A complete compilation pipeline is provided for the mapping of molecular Hamiltonians to optical lattice operations is detailed through rules for decomposition and merging of fundamental component gates.

Since its proposal as a means of estimating energies of molecular structures and proving an excellent testbed for a plethora of aspects of quantum computing, the VQE framework currently seems to have reached a peak in terms of its capabilities, as no king rules forever [124]. Considerable headway has been made through the use of purification error mitigation techniques, both algorithmically and experimentally, yielding accurate results and successfully accounting for cert ain types of noise. The ever-increasing Hilbert space of experiments involving larger qubit counts acts beneficially for the purpose of filtering errors that would cause measurements which would not satisfy verification criteria. Unfortunately, this purification is achieved at the expense of very high number of experiment repetitions as shown in Section 3.3, leading to unsustainable requirements on the computational time required to reach the necessary accuracy for noteworthy system sizes. Latest developments in the field seek to ameliorate this by providing gates with very high fidelity, without significant compromise on gate operation time [125, 126].

In the pursuit of accurately estimating expectation values from quantum apparatus for use in quantum algorithms, research efforts have committed to exploring slower platforms in terms of experiment repetition time, but with each individual repetition aspiring to have greater accuracy than superconducting qubits. Trapped atoms offer the advantage of high accuracy native gates and coherence times, and with efforts such as those in Section 3.5, the booming field of quantum simulation can offer invaluable insights through the leveraging of techniques currently present in quantum algorithm literature, with the caveat of the very slow repetition times in comparison to superconducting qubits. Trapped ions offer similar benefits to trapped atoms, while also not being restricted to just neighboring qubit interactions, with its own limitation of possible ion loss and overheads incurred by ion shuttling. Moreover, advancements in error correcting code protocols have been far from stagnant, as efforts are made to achieve such protocols in other platforms other than superconducting qubits, for example by using Gottesman-Kitaev-Preskill (GKP) qubits to allow the use of arbitrary error correcting codes [127].

Ultimately, the accumulated experience and intuition developed through this work point to the rapid, accurate control of quantum gates and the optimization routines to control them, as well as the innate barren plateaus of energy gradients as drivers for VQE implementations finding better use as benchmarking and infrastructure tools, rather than realistic path forwards towards achieving quantum advantage or at least reach the break-even regime of rivaling current state of the art classical computational chemistry methods, even with the help of pre-initialization of parameters. Moreover, as interest is being diverted from NISQ devices and variational algorithms, it is being redirected towards a significant barrier for achieving fault tolerant quantum computing, namely the engineering challenges concerning the individual components that make up the quantum apparatus. Recently, a push for the upgrading of these components is observed, with significant advancements to cryogenic equipment for platforms that rely on dilution refrigerators such as superconducting qubits [128]. Also, significant headway has been made towards the interconnection of smaller quantum devices to allow for scaling of the available qubit counts to cater towards relevant problems, with clear benefits to quantum chemistry and quantum simulation [129]. As more novel ways of utilizing quantum computers emerge, such as generators of true randomness, it is unclear whether quantum chemistry will be the pioneering application of quantum computing that will demonstrate quantum advantage over classical methods.

5 Availability of primary data

Owing to the nature of a collaboration between an academic body and an industrial company, some parts of the work included remain confidential. Considerable effort has been made to provide the most amount of data possible of the ideas contained in this thesis.

Data from the Purification-experiment in collaboration with Google Quantum AI in Section 3.3 is available at `https://doi.org/10.5281/zenodo.7225821`.

Data from the Fermionic VQE in collaboration with Max Planck Institute of Quantum Optics in Section 3.5 is available at `https://doi.org/10.5281/zenodo.12665177`.

The implementation of Richardson Gaudin models, as well as the code for converting reduced Bradley-Cooper-Schieffer Hamiltonians into corresponding electronic structure Hamiltonian integrals has been integrated by the author into the OpenFermion Python Package at: `https://github.com/quantumlib/OpenFermion/pull/770`.

Data from the benchmarking of the behavior of the trapped ion device with respect to qubit sizes and circuit layers from Johannes Gutenberg Universität Mainz can be found here: `https://github.com/fgkritsis/PhD_notebooks/tree/main/Mainz_data/mainz_device_benchmarking`

Data from the phase functions, expectation values and fidelities from the trapped ion quantum devices from Johannes Gutenberg Universität Mainz can be found at: `https://github.com/fgkritsis/PhD_notebooks/tree/main/Mainz_data/mainz_phase_fn_data`

Data from the hyperparameter tuning of optimizers over 10000 generations can be found at: `https://github.com/fgkritsis/PhD_notebooks/tree/main/Optimization_data/optimizer_hp_tuning_data/10000gens`

Data from the comparison of performance of the various optimizers examined in noiseless and noisy settings over different system sizes can be found here: `https://github.com/fgkritsis/PhD_notebooks/tree/main/Optimization_data/optimizer_hp_comparison_data`

6 List of publications

O'Brien, T. E., Anselmetti, G., **Gkritsis, F.**, Elfving, V. E., Polla, S., Huggins, W. J., ..., Rubin, N. C. (2023, December). Purification-based quantum error mitigation of pair-correlated electron simulations. Nat. Phys. 19, 1787–1792 (2023). Retrieved 2024-09-02, from: `https://www.nature.com/articles/s41567-023-02240-y` DOI: 10.1038/s41567-023-02240-y

Gkritsis, F., Dux, D., Zhang, J., Naman, J., Gogolin, C., Preiss, P.M. (2025, January). Simulating Chemistry with Fermionic Optical Superlattices. PRX Quantum 6, 010318. Retrieved 2025-01-25, from: `https://journals.aps.org/prxquantum/abstract/10.1103/PRXQuantum.6.010318`

References

[1] J. C. Slater. "Atomic Shielding Constants". In: *Phys. Rev.* 36 (1 July 1930), pp. 57–64. DOI: 10.1103/PhysRev.36.57. URL: https://link.aps.org/doi/10.1103/PhysRev.36.57.

[2] J. C. Slater. "A Simplification of the Hartree-Fock Method". In: *Phys. Rev.* 81 (3 Feb. 1951), pp. 385–390. DOI: 10.1103/PhysRev.81.385. URL: https://link.aps.org/doi/10.1103/PhysRev.81.385.

[3] O. Anatole von Lilienfeld, Ivano Tavernelli, Ursula Rothlisberger, and Daniel Sebastiani. "Optimization of Effective Atom Centered Potentials for London Dispersion Forces in Density Functional Theory". In: *Phys. Rev. Lett.* 93 (15 Oct. 2004), p. 153004. DOI: 10.1103/PhysRevLett.93.153004. URL: https://link.aps.org/doi/10.1103/PhysRevLett.93.153004.

[4] Alberto Baiardi and Markus Reiher. "The density matrix renormalization group in chemistry and molecular physics: Recent developments and new challenges". In: *The Journal of Chemical Physics* 152.4 (Jan. 2020), p. 040903. ISSN: 0021-9606. DOI: 10.1063/1.5129672. eprint: https://pubs.aip.org/aip/jcp/article-pdf/doi/10.1063/1.5129672/19996154/040903_1_1.5129672.pdf. URL: https://doi.org/10.1063/1.5129672.

[5] A. Szabo and N.S. Ostlund. *Modern Quantum Chemistry: Introduction to Advanced Electronic Structure Theory*. Dover Books on Chemistry. Dover Publications, 1996. ISBN: 9780486691862. URL: https://books.google.fr/books?id=k-DcCgAAQBAJ.

[6] G.H.F. Diercksen, B.T. Sutcliffe, and A. Veillard. *Computational Techniques in Quantum Chemistry and Molecular Physics: Proceedings of the NATO Advanced Study Institute held at Ramsau, Germany, 4–21 September, 1974*. Nato Science Series C: Springer Netherlands, 2012. ISBN: 9789401018159. URL: https://books.google.ch/books?id=g_SpCAAAQBAJ.

[7] C. David Sherrill and Henry F. Schaefer. "The Configuration Interaction Method: Advances in Highly Correlated Approaches". In: *Advances in Quantum Chemistry*. Ed. by Per-Olov Löwdin, John R. Sabin, Michael C. Zerner, and Erkki Brändas. Vol. 34. Academic Press, 1999, pp. 143–269. DOI: https://doi.org/10.1016/S0065-3276(08)60532-8. URL: https://www.sciencedirect.com/science/article/pii/S0065327608605328.

[8] J. Bardeen, L. N. Cooper, and J. R. Schrieffer. "Theory of Superconductivity". In: *Phys. Rev.* 108 (5 Dec. 1957), pp. 1175–1204. DOI: 10.1103/PhysRev.108.1175. URL: https://link.aps.org/doi/10.1103/PhysRev.108.1175.

[9] Vincent E. Elfving, Marta Millaruelo, José A. Gámez, and Christian Gogolin. "Simulating quantum chemistry in the seniority-zero space on qubit-based quantum computers". In: *Phys. Rev. A* 103 (3 Mar. 2021), p. 032605. DOI: 10.1103/PhysRevA.103.032605. URL: https://link.aps.org/doi/10.1103/PhysRevA.103.032605.

[10] 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. "Quantum Chemistry in the Age of Quantum Computing". In: *Chemical Reviews* 119.19 (Aug. 2019), pp. 10856–10915. ISSN: 1520-6890. DOI: 10.1021/acs.chemrev.8b00803. URL: http://dx.doi.org/10.1021/acs.chemrev.8b00803.

[11] Google AI Quantum, Collaborators*†, Frank Arute, Kunal Arya, Ryan Babbush, Dave Bacon, Joseph C. Bardin, Rami Barends, Sergio Boixo, Michael Broughton, Bob B. Buckley, David A. Buell, Brian Burkett, Nicholas Bushnell, Yu Chen, Zijun Chen, Benjamin Chiaro, Roberto Collins, William Courtney, Sean Demura, Andrew Dunsworth, Edward Farhi, Austin Fowler, Brooks Foxen, Craig Gidney, Marissa Giustina, Rob Graff, Steve Habegger, Matthew P. Harrigan, Alan Ho, Sabrina Hong, Trent Huang, William J. Huggins, Lev Ioffe, Sergei V. Isakov, Evan Jeffrey, Zhang Jiang, Cody Jones, Dvir Kafri, Kostyantyn Kechedzhi, Julian Kelly, Seon Kim, Paul V. Klimov, Alexander Korotkov, Fedor Kostritsa, David Landhuis, Pavel Laptev, Mike Lindmark, Erik Lucero, Orion Martin, John M. Martinis, Jarrod R. McClean, Matt McEwen, Anthony Megrant, Xiao Mi, Masoud Mohseni, Wojciech Mruczkiewicz, Josh Mutus, Ofer Naaman, Matthew Neeley, Charles Neill, Hartmut Neven, Murphy Yuezhen Niu, Thomas E. O'Brien, Eric Ostby, Andre Petukhov, Harald Putterman, Chris Quintana, Pedram Roushan, Nicholas C. Rubin, Daniel Sank, Kevin J. Satzinger, Vadim Smelyanskiy, Doug Strain, Kevin J. Sung, Marco Szalay, Tyler Y.

Takeshita, Amit Vainsencher, Theodore White, Nathan Wiebe, Z. Jamie Yao, Ping Yeh, and Adam Zalcman. "Hartree-Fock on a superconducting qubit quantum computer". In: *Science* 369.6507 (Aug. 2020), pp. 1084–1089. DOI: 10.1126/science.abb9811. URL: https://doi.org/10.1126/science.abb9811.

[12] Isaiah Shavitt and Rodney J. Bartlett. *Many-Body Methods in Chemistry and Physics: MBPT and Coupled-Cluster Theory*. Cambridge Molecular Science. Cambridge: Cambridge University Press, 2009. ISBN: 9780521818322. DOI: 10.1017/CBO9780511596834. URL: https://doi.org/10.1017/CBO9780511596834.

[13] Henry Eyring. "The Activated Complex in Chemical Reactions". In: *The Journal of Chemical Physics* 3.2 (Feb. 1935), pp. 107–115. ISSN: 0021-9606. DOI: 10.1063/1.1749604. URL: https://doi.org/10.1063/1.1749604.

[14] David Feller, Kirk A. Peterson, and Branko Ruscic. "Improved accuracy benchmarks of small molecules using correlation consistent basis sets". In: *Theoretical Chemistry Accounts* 133.1 (Nov. 2013), p. 1407. ISSN: 1432-2234. DOI: 10.1007/s00214-013-1407-z. URL: https://doi.org/10.1007/s00214-013-1407-z.

[15] Jie J. Bao and Donald G. Truhlar. "Automatic Active Space Selection for Calculating Electronic Excitation Energies Based on High-Spin Unrestricted Hartree–Fock Orbitals". In: *Journal of Chemical Theory and Computation* 15.10 (Oct. 2019), pp. 5308–5318. ISSN: 1549-9618. DOI: 10.1021/acs.jctc.9b00535. URL: https://doi.org/10.1021/acs.jctc.9b00535.

[16] P. J. Knowles and N. C. Handy. "A new determinant-based full configuration interaction method". In: *Chemical Physics Letters* 111.4 (Nov. 1984), pp. 315–321. ISSN: 0009-2614. URL: https://www.sciencedirect.com/science/article/pii/000926148485513X.

[17] Per-Åke Malmqvist and Björn O. Roos. "The CASSCF state interaction method". In: *Chemical Physics Letters* 155.2 (Feb. 1989), pp. 189–194. ISSN: 0009-2614. URL: https://www.sciencedirect.com/science/article/pii/0009261489853473.

[18] Rodney J. Bartlett and Monika Musiał. "Coupled-cluster theory in quantum chemistry". In: *Rev. Mod. Phys.* 79 (1 Feb. 2007), pp. 291–352. DOI: 10.1103/RevModPhys.79.291. URL: https://link.aps.org/doi/10.1103/RevModPhys.79.291.

[19] John F. Stanton and Rodney J. Bartlett. "The equation of motion coupled-cluster method. A systematic biorthogonal approach to molecular excitation energies, transition probabilities, and excited state properties". In: *The Journal of Chemical Physics* 98.9 (May 1993), pp. 7029–7039. ISSN: 0021-9606. DOI: 10.1063/1.464746. URL: https://doi.org/10.1063/1.464746.

[20] Rodney J. Bartlett and Monika Musiał. "Coupled-cluster theory in quantum chemistry". In: *Rev. Mod. Phys.* 79 (1 Feb. 2007), pp. 291–352. DOI: 10.1103/RevModPhys.79.291. URL: https://link.aps.org/doi/10.1103/RevModPhys.79.291.

[21] T. Crawford and Henry III. "An Introduction to Coupled Cluster Theory for Computational Chemists". In: *Rev Comp Chem*. Vol. 14. Jan. 2007, pp. 33–136. ISBN: 9780470125915. DOI: 10.1002/9780470125915.ch2. URL: https://doi.org/10.1002/9780470125915.ch2.

[22] Jiří Čížek. "On the Correlation Problem in Atomic and Molecular Systems. Calculation of Wavefunction Components in Ursell-Type Expansion Using Quantum-Field Theoretical Methods". In: *The Journal of Chemical Physics* 45.11 (Dec. 1966), pp. 4256–4266. ISSN: 0021-9606. DOI: 10.1063/1.1727484. URL: https://doi.org/10.1063/1.1727484.

[23] P. Jordan and E. Wigner. "Über das Paulische Äquivalenzverbot". In: *Zeitschrift für Physik* 47.9 (Sept. 1928), pp. 631–651. ISSN: 0044-3328. DOI: 10.1007/BF01331938. URL: https://doi.org/10.1007/BF01331938.

[24] S. B. Bravyi and A. Y. Kitaev. "Fermionic Quantum Computation". In: *Annals of Physics* 298.1 (2002), pp. 210–226. ISSN: 0003-4916. DOI: https://doi.org/10.1006/aphy.2002.6254. URL: https://www.sciencedirect.com/science/article/pii/S0003491602962548.

[25] Vojt ěch Havlí ček, Matthias Troyer, and James D. Whitfield. "Operator locality in the quantum simulation of fermionic models". In: *Phys. Rev. A* 95 (3 Mar. 2017), p. 032332. DOI: 10.1103/PhysRevA.95.032332. URL: https://link.aps.org/doi/10.1103/PhysRevA.95.032332.

[26] 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 (1992), pp. 553–558. URL: https://api.semanticscholar.org/CorpusID:121702767.

[27] 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. DOI: 10.1137/S0097539795293172. URL: https://doi.org/10.1137/S0097539795293172.

[28] Lov K. Grover. "From Schrödinger's equation to the quantum search algorithm". In: *Pramana* 56.2 (Feb. 2001), pp. 333–348. ISSN: 0973-7111. DOI: 10.1007/s12043-001-0128-3. URL: https://doi.org/10.1007/s12043-001-0128-3.

[29] Daniel A. Lidar and Haobin Wang. "Calculating the thermal rate constant with exponential speedup on a quantum computer". In: *Phys. Rev. E* 59 (2 Feb. 1999), pp. 2429–2438. DOI: 10.1103/PhysRevE.59.2429. URL: https://link.aps.org/doi/10.1103/PhysRevE.59.2429.

[30] A. Aspuru-Guzik, A. D. Dutoi, P. J. Love, and M. Head-Gordon. "Simulated Quantum Computation of Molecular Energies". In: *Science* 309.5741 (2005), pp. 1704–1707. DOI: 10.1126/science.1113479. eprint: https://www.science.org/doi/pdf/10.1126/science.1113479. URL: https://www.science.org/doi/abs/10.1126/science.1113479.

[31] 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". In: *Nature* 549.7671 (Sept. 2017), pp. 242–246. ISSN: 1476-4687. DOI: 10.1038/nature23879. URL: http://dx.doi.org/10.1038/nature23879.

[32] Oumarou Oumarou, Maximilian Scheurer, Robert M. Parrish, Edward G. Hohenstein, and Christian Gogolin. "Accelerating Quantum Computations of Chemistry Through Regularized Compressed Double Factorization". In: *Quantum* 8 (June 2024), p. 1371. ISSN: 2521-327X. DOI: 10.22331/q-2024-06-13-1371. URL: https://doi.org/10.22331/q-2024-06-13-1371.

[33] D. W. Berry, C. Gidney, M. Motta, J. R. McClean, and R. Babbush. "Qubitization of arbitrary basis quantum chemistry leveraging sparsity and low rank factorization". In: *Quantum* 3 (2019), p. 208.

[34] N. Yoshioka, M. Amico, W. Kirby, P. Jurcevic, A. Dutt, B. Fuller, S. Garion, H. Haas, I. Hamamura, A. Ivrii, R. Majumdar, Z. Minev, M. Motta, B. Pokharel, P. Rivero, K. Sharma, C. J. Wood, A. Javadi-Abhari, and A. Mezzacapo. *Diagonalization of large many-body Hamiltonians on a quantum processor*. 2024. arXiv: 2407.14431 [quant-ph]. URL: https://arxiv.org/abs/2407.14431.

[35] S. A. Moses, C. H. Baldwin, M. S. Allman, R. Ancona, L. Ascarrunz, C. Barnes, J. Bartolotta, B. Bjork, P. Blanchard, M. Bohn, J. G. Bohnet, N. C. Brown, N. Q. Burdick, W. C. Burton, S. L. Campbell, J. P. Campora, C. Carron, J. Chambers, J. W. Chan, Y. H. Chen, A. Chernoguzov, E. Chertkov, J. Colina, J. P. Curtis, R. Daniel, M. DeCross, D. Deen, C. Delaney, J. M. Dreiling, C. T. Ertsgaard, J. Esposito, B. Estey, M. Fabrikant, C. Figgatt, C. Foltz, M. Foss-Feig, D. Francois, J. P. Gaebler, T. M. Gatterman, C. N. Gilbreth, J. Giles, E. Glynn, A. Hall, A. M. Hankin, A. Hansen, D. Hayes, B. Higashi, I. M. Hoffman, B. Horning, J. J. Hout, R. Jacobs, J. Johansen, L. Jones, J. Karcz, T. Klein, P. Lauria, P. Lee, D. Liefer, S. T. Lu, D. Lucchetti, C. Lytle, A. Malm, M. Matheny, B. Mathewson, K. Mayer, D. B. Miller, M. Mills, B. Neyenhuis, L. Nugent, S. Olson, J. Parks, G. N. Price, Z. Price, M. Pugh, A. Ransford, A. P. Reed, C. Roman, M. Rowe, C. Ryan-Anderson, S. Sanders, J. Sedlacek, P. Shevchuk, P. Siegfried, T. Skripka, B. Spaun, R. T. Sprenkle, R. P. Stutz, M. Swallows, R. I. Tobey, A. Tran, T. Tran, E. Vogt, C. Volin, J. Walker, A. M. Zolot, and J. M. Pino. "A Race-Track Trapped-Ion Quantum Processor". In: *Phys. Rev. X* 13 (4 Dec. 2023), p. 041052. DOI: 10.1103/PhysRevX.13.041052. URL: https://link.aps.org/doi/10.1103/PhysRevX.13.041052.

[36] Rajeev Acharya et al. "Quantum error correction below the surface code threshold". In: *Nature* (Dec. 2024). ISSN: 1476-4687. DOI: 10.1038/s41586-024-08449-y. URL: https://doi.org/10.1038/s41586-024-08449-y.

[37] Rajeev Acharya et al. “Suppressing quantum errors by scaling a surface code logical qubit”. In: *Nature* 614.7949 (Feb. 2023), pp. 676–681. ISSN: 1476-4687. DOI: 10.1038/s41586-022-05434-1. URL: https://doi.org/10.1038/s41586-022-05434-1.

[38] Hsin-Yuan Huang, John Preskill, and Mehdi Soleimanifar. *Certifying almost all quantum states with few single-qubit measurements*. 2024. arXiv: 2404.07281 [quant-ph]. URL: https://arxiv.org/abs/2404.07281.

[39] Alberto Peruzzo, Jarrod McClean, Peter Shadbolt, Man-Hong Yung, Xiao-Qi Zhou, Peter J. Love, Alán Aspuru-Guzik, and Jeremy L. O’Brien. “A variational eigenvalue solver on a photonic quantum processor”. In: *Nature Communications* 5.1 (July 2014), p. 4213. ISSN: 2041-1723. DOI: 10.1038/ncomms5213. URL: https://doi.org/10.1038/ncomms5213.

[40] Jarrod R McClean, Jonathan Romero, Ryan Babbush, and Alán Aspuru-Guzik. “The theory of variational hybrid quantum-classical algorithms”. In: *New Journal of Physics* 18.2 (Feb. 2016), pp. 023–023. ISSN: 1367-2630. DOI: 10.1088/1367-2630/18/2/023023. URL: http://dx.doi.org/10.1088/1367-2630/18/2/023023.

[41] Yuan Su, Dominic W. Berry, Nathan Wiebe, Nicholas Rubin, and Ryan Babbush. “Fault-Tolerant Quantum Simulations of Chemistry in First Quantization”. In: *PRX Quantum* 2 (4 Nov. 2021), p. 040332. DOI: 10.1103/PRXQuantum.2.040332. URL: https://link.aps.org/doi/10.1103/PRXQuantum.2.040332.

[42] M.A. Nielsen and I.L. Chuang. *Quantum Computation and Quantum Information: 10th Anniversary Edition*. Cambridge University Press, 2010. ISBN: 9781107002173. URL: https://books.google.fr/books?id=j2ULnwEACAAJ.

[43] M. Cerezo, Andrew Arrasmith, Ryan Babbush, Simon C. Benjamin, Suguru Endo, Keisuke Fujii, Jarrod R. McClean, Kosuke Mitarai, Xiao Yuan, Lukasz Cincio, and Patrick J. Coles. “Variational quantum algorithms”. In: *Nature Reviews Physics* 3.9 (Sept. 2021), pp. 625–644. ISSN: 2522-5820. DOI: 10.1038/s42254-021-00348-9. URL: https://doi.org/10.1038/s42254-021-00348-9.

[44] Jules Tilly, Hongxiang Chen, Shuxiang Cao, Dario Picozzi, Kanav Setia, Ying Li, Edward Grant, Leonard Wossnig, Ivan Rungger, George H. Booth, and Jonathan Tennyson. “The Variational Quantum Eigensolver: A review of methods and best practices”. In: *Physics Reports* 986 (Nov. 2022), pp. 1–128. ISSN: 0370-1573. URL: https://www.sciencedirect.com/science/article/pii/S0370157322003118.

[45] Cornelius Hempel, Christine Maier, Jonathan Romero, Jarrod McClean, Thomas Monz, Heng Shen, Petar Jurcevic, Ben P. Lanyon, Peter Love, Ryan Babbush, Alán Aspuru-Guzik, Rainer Blatt, and Christian F. Roos. “Quantum Chemistry Calculations on a Trapped-Ion Quantum Simulator”. In: *Phys. Rev. X* 8 (3 July 2018), p. 031022. DOI: 10.1103/PhysRevX.8.031022. URL: https://link.aps.org/doi/10.1103/PhysRevX.8.031022.

[46] Seth Lloyd. “Universal Quantum Simulators”. In: *Science* 273.5278 (Aug. 1996), pp. 1073–1078. DOI: 10.1126/science.273.5278.1073. URL: https://doi.org/10.1126/science.273.5278.1073.

[47] Harper R. Grimsley, Sophia E. Economou, Edwin Barnes, and Nicholas J. Mayhall. “An adaptive variational algorithm for exact molecular simulations on a quantum computer”. In: *Nature Communications* 10.1 (July 2019), p. 3007. ISSN: 2041-1723. DOI: 10.1038/s41467-019-10988-2. URL: https://doi.org/10.1038/s41467-019-10988-2.

[48] Panagiotis Kl. Barkoutsos, Jerome F. Gonthier, Igor Sokolov, Nikolaj Moll, Gian Salis, Andreas Fuhrer, Marc Ganzhorn, Daniel J. Egger, Matthias Troyer, Antonio Mezzacapo, Stefan Filipp, and Ivano Tavernelli. “Quantum algorithms for electronic structure calculations: Particle-hole Hamiltonian and optimized wave-function expansions”. In: *Phys. Rev. A* 98 (2 Aug. 2018), p. 022322. DOI: 10.1103/PhysRevA.98.022322. URL: https://link.aps.org/doi/10.1103/PhysRevA.98.022322.

[49] G.-L. R Anselmetti, D. Wierichs, C. Gogolin, and R. M. Parrish. “Local, expressive, quantum-number-preserving VQE ansätze for fermionic systems”. In: *New Journal of Physics* 23.11 (Nov. 2021), p. 113010. ISSN: 1367-2630. DOI: 10.1088/1367-2630/ac2cb3. URL: http://dx.doi.org/10.1088/1367-2630/ac2cb3.

[50] Linghua Zhu, Senwei Liang, Chao Yang, and Xiaosong Li. "Optimizing Shot Assignment in Variational Quantum Eigensolver Measurement". In: *Journal of Chemical Theory and Computation* 20.6 (Mar. 2024), pp. 2390–2403. ISSN: 1549-9618. DOI: 10.1021/acs.jctc.3c01113. URL: https://doi.org/10.1021/acs.jctc.3c01113.

[51] William J. Huggins, Jarrod R. McClean, Nicholas C. Rubin, Zhang Jiang, Nathan Wiebe, K. Birgitta Whaley, and Ryan Babbush. "Efficient and noise resilient measurements for quantum chemistry on near-term quantum computers". In: *npj Quantum Information* 7.1 (Feb. 2021), p. 23. ISSN: 2056-6387. DOI: 10.1038/s41534-020-00341-7. URL: https://doi.org/10.1038/s41534-020-00341-7.

[52] Vladyslav Verteletskyi, Tzu-Ching Yen, and Artur F. Izmaylov. "Measurement optimization in the variational quantum eigensolver using a minimum clique cover". In: *The Journal of Chemical Physics* 152.12 (Mar. 2020), p. 124114. ISSN: 0021-9606. DOI: 10.1063/1.5141458. URL: https://doi.org/10.1063/1.5141458.

[53] Mario Motta, Erika Ye, Jarrod R. McClean, Zhendong Li, Austin J. Minnich, Ryan Babbush, and Garnet Kin-Lic Chan. "Low rank representations for quantum simulation of electronic structure". In: *npj Quantum Information* 7.1 (May 2021), p. 83. ISSN: 2056-6387. DOI: 10.1038/s41534-021-00416-z. URL: https://doi.org/10.1038/s41534-021-00416-z.

[54] Vera von Burg, Guang Hao Low, Thomas Häner, Damian S. Steiger, Markus Reiher, Martin Roetteler, and Matthias Troyer. "Quantum computing enhanced computational catalysis". In: *Phys. Rev. Res.* 3 (3 July 2021), p. 033055. DOI: 10.1103/PhysRevResearch.3.033055. URL: https://link.aps.org/doi/10.1103/PhysRevResearch.3.033055.

[55] Jeffrey Cohn, Mario Motta, and Robert M. Parrish. "Quantum Filter Diagonalization with Compressed Double-Factorized Hamiltonians". In: *PRX Quantum* 2 (4 Dec. 2021), p. 040352. DOI: 10.1103/PRXQuantum.2.040352. URL: https://link.aps.org/doi/10.1103/PRXQuantum.2.040352.

[56] Hsin-Yuan Huang, Richard Kueng, and John Preskill. "Predicting many properties of a quantum system from very few measurements". In: *Nature Physics* 16.10 (Oct. 2020), pp. 1050–1057. ISSN: 1745-2481. DOI: 10.1038/s41567-020-0932-7. URL: https://doi.org/10.1038/s41567-020-0932-7.

[57] Kianna Wan, William J. Huggins, Joonho Lee, and Ryan Babbush. "Matchgate Shadows for Fermionic Quantum Simulation". In: *Communications in Mathematical Physics* 404.2 (Dec. 2023), pp. 629–700. ISSN: 1432-0916. DOI: 10.1007/s00220-023-04844-0. URL: https://doi.org/10.1007/s00220-023-04844-0.

[58] Guang Hao Low. *Classical shadows of fermions with particle number symmetry*. 2024. arXiv: 2208.08964 [quant-ph]. URL: https://arxiv.org/abs/2208.08964.

[59] T. E. O'Brien et al. "Purification-based quantum error mitigation of pair-correlated electron simulations". In: *Nature Physics* 19.12 (Dec. 2023), pp. 1787–1792. ISSN: 1745-2481. DOI: 10.1038/s41567-023-02240-y. URL: https://doi.org/10.1038/s41567-023-02240-y.

[60] Dong C. Liu and Jorge Nocedal. "On the limited memory BFGS method for large scale optimization". In: *Mathematical Programming* 45.1 (Aug. 1989), pp. 503–528. ISSN: 1436-4646. DOI: 10.1007/BF01589116. URL: https://doi.org/10.1007/BF01589116.

[61] L. Bittel and M. Kliesch. "Training Variational Quantum Algorithms Is NP-Hard". In: *Phys. Rev. Lett.* 127 (12 Sept. 2021), p. 120502. DOI: 10.1103/PhysRevLett.127.120502. URL: https://link.aps.org/doi/10.1103/PhysRevLett.127.120502.

[62] J. R. McClean, S. Boixo, V. N. Smelyanskiy, R. Babbush, and H. Neven. "Barren plateaus in quantum neural network training landscapes". In: *Nature Communications* 9.1 (Nov. 2018), p. 4812. ISSN: 2041-1723. DOI: 10.1038/s41467-018-07090-4. URL: https://doi.org/10.1038/s41467-018-07090-4.

[63] D. Wierichs, J. Izaac, C. Wang, and C. Yen-Yu Lin. "General parameter-shift rules for quantum gradients". In: *Quantum* 6 (Mar. 2022), p. 677. ISSN: 2521-327X. DOI: 10.22331/q-2022-03-30-677. URL: https://doi.org/10.22331/q-2022-03-30-677.

[64] J. S. Kottmann, A. Anand, and A. Aspuru-Guzik. "A feasible approach for automatically differentiable unitary coupled-cluster on quantum computers". In: *Chem. Sci.* 12 (10 2021), pp. 3497–3508. DOI: 10.1039/D0SC06627C. URL: http://dx.doi.org/10.1039/D0SC06627C.

[65] Andrea Mari, Thomas R. Bromley, and Nathan Killoran. "Estimating the gradient and higher-order derivatives on quantum hardware". In: *Phys. Rev. A* 103 (1 Jan. 2021), p. 012405. DOI: 10.1103/PhysRevA.103.012405. URL: https://link.aps.org/doi/10.1103/PhysRevA.103.012405.

[66] Richard H. Byrd, Peihuang Lu, Jorge Nocedal, and Ciyou Zhu. "A Limited Memory Algorithm for Bound Constrained Optimization". In: *SIAM Journal on Scientific Computing* 16.5 (1995), pp. 1190–1208. DOI: 10.1137/0916069. eprint: https://doi.org/10.1137/0916069. URL: https://doi.org/10.1137/0916069.

[67] Ciyou Zhu, Richard H. Byrd, Peihuang Lu, and Jorge Nocedal. "Algorithm 778: L-BFGS-B: Fortran subroutines for large-scale bound-constrained optimization". In: *ACM Trans. Math. Softw.* 23.4 (Dec. 1997), pp. 550–560. ISSN: 0098-3500. DOI: 10.1145/279232.279236. URL: https://doi.org/10.1145/279232.279236.

[68] Hermann Matthies and Gilbert Strang. "The solution of nonlinear finite element equations". In: *International Journal for Numerical Methods in Engineering* 14.11 (1979), pp. 1613–1626. DOI: https://doi.org/10.1002/nme.1620141104. eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/nme.1620141104. URL: https://onlinelibrary.wiley.com/doi/abs/10.1002/nme.1620141104.

[69] M. Ostaszewski, E. Grant, and M. Benedetti. "Structure optimization for parameterized quantum circuits". In: *Quantum* 5 (Jan. 2021), p. 391. ISSN: 2521-327X. DOI: 10.22331/q-2021-01-28-391. URL: https://doi.org/10.22331/q-2021-01-28-391.

[70] L. Geck, A. Kruth, H. Bluhm, S. van Waasen, and S. Heinen. "Control electronics for semiconductor spin qubits". In: *Quantum Science and Technology* 5.1 (Dec. 2019), p. 015004. ISSN: 2058-9565. DOI: 10.1088/2058-9565/ab5e07. URL: http://dx.doi.org/10.1088/2058-9565/ab5e07.

[71] Weitang Li, Yufei Ge, Shi-Xin Zhang, Yu-Qin Chen, and Shengyu Zhang. "Efficient and Robust Parameter Optimization of the Unitary Coupled-Cluster Ansatz". In: *Journal of Chemical Theory and Computation* 20.9 (May 2024), pp. 3683–3696. ISSN: 1549-9618. DOI: 10.1021/acs.jctc.4c00155. URL: https://doi.org/10.1021/acs.jctc.4c00155.

[72] James C. Spall. "A Stochastic Approximation Technique for Generating Maximum Likelihood Parameter Estimates". In: *1987 American Control Conference*. 1987, pp. 1161–1167. DOI: 10.23919/ACC.1987.4789489.

[73] J.C. Spall. "Implementation of the simultaneous perturbation algorithm for stochastic optimization". In: *IEEE Transactions on Aerospace and Electronic Systems* 34.3 (1998), pp. 817–823. DOI: 10.1109/7.705889.

[74] James Stokes, Josh Izaac, Nathan Killoran, and Giuseppe Carleo. "Quantum Natural Gradient". In: *Quantum* 4 (May 2020), p. 269. ISSN: 2521-327X. DOI: 10.22331/q-2020-05-25-269. URL: https://doi.org/10.22331/q-2020-05-25-269.

[75] Julien Gacon, Christa Zoufal, Giuseppe Carleo, and Stefan Woerner. "Simultaneous Perturbation Stochastic Approximation of the Quantum Fisher Information". In: *Quantum* 5 (Oct. 2021), p. 567. ISSN: 2521-327X. DOI: 10.22331/q-2021-10-20-567. URL: https://doi.org/10.22331/q-2021-10-20-567.

[76] Marco Wiedmann, Marc Hölle, Maniraman Periyasamy, Nico Meyer, Christian Ufrecht, Daniel D. Scherer, Axel Plinge, and Christopher Mutschler. "An Empirical Comparison of Optimizers for Quantum Machine Learning with SPSA-Based Gradients". In: *2023 IEEE International Conference on Quantum Computing and Engineering (QCE)*. IEEE, Sept. 2023. DOI: 10.1109/qce57702.2023.00058. URL: http://dx.doi.org/10.1109/QCE57702.2023.00058.

[77] Ryan Sweke, Frederik Wilde, Johannes Meyer, Maria Schuld, Paul K. Faehrmann, Barthélémy Meynard-Piganeau, and Jens Eisert. "Stochastic gradient descent for hybrid quantum-classical optimization". In: *Quantum* 4 (Aug. 2020), p. 314. ISSN: 2521-327X. DOI: 10.22331/q-2020-08-31-314. URL: https://doi.org/10.22331/q-2020-08-31-314.

[78] S. Stanisic, Jan L. Bosse, F. M. Gambetta, R. A. Santos, W. Mruczkiewicz, T. E. O'Brien, E Ostby, and A Montanaro. "Observing ground-state properties of the Fermi-Hubbard model using a scalable algorithm on a quantum computer". In: *Nature Communications* 13.1 (Oct. 2022), p. 5743. ISSN: 2041-1723. DOI: 10.1038/s41467-022-33335-4. URL: https://doi.org/10.1038/s41467-022-33335-4.

[79] Kevin J Sung, Jiahao Yao, Matthew P Harrigan, Nicholas C Rubin, Zhang Jiang, Lin Lin, Ryan Babbush, and Jarrod R McClean. "Using models to improve optimizers for variational quantum algorithms". In: *Quantum Science and Technology* 5.4 (Sept. 2020), p. 044008. DOI: 10.1088/2058-9565/abb6d9. URL: https://doi.org/10.1088/2058-9565/abb6d9.

[80] Jasper Snoek, Hugo Larochelle, and Ryan P. Adams. *Practical Bayesian Optimization of Machine Learning Algorithms*. 2012. arXiv: 1206.2944 [stat.ML]. URL: https://arxiv.org/abs/1206.2944.

[81] Yousef Saad. *Iterative Methods for Sparse Linear Systems*. SIAM, Jan. 2003. DOI: 10.1137/1.9780898718003.ch4.

[82] Magnus R. Hestenes and Eduard Stiefel. "Methods of Conjugate Gradients for Solving Linear systems". In: *Journal of Research of the National Bureau of Standards* 49.6 (Dec. 1952), pp. 409–436. DOI: 10.6028\%2Fjres.049.044. URL: https://doi.org/10.6028%5C%2Fjres.049.044.

[83] Magnus R. Hestenes and Eduard Stiefel. "Methods of conjugate gradients for solving linear systems". In: *Journal of research of the National Bureau of Standards* 49 (1952), pp. 409–435. URL: https://api.semanticscholar.org/CorpusID:2207234.

[84] R. Fletcher and C. M. Reeves. "Function minimization by conjugate gradients". In: *The Computer Journal* 7.2 (Jan. 1964), pp. 149–154. ISSN: 0010-4620. DOI: 10.1093/comjnl/7.2.149. URL: https://doi.org/10.1093/comjnl/7.2.149.

[85] E. Polak and G. Ribiere. "Note sur la convergence de méthodes de directions conjuguées". In: *R.I.R.O.* 3.16 (1969), pp. 35–43. DOI: 10.1051/m2an/196903R100351. URL: https://doi.org/10.1051/m2an/196903R100351.

[86] Y. H. Dai and Y. Yuan. "A Nonlinear Conjugate Gradient Method with a Strong Global Convergence Property". In: *SIAM Journal on Optimization* 10.1 (1999), pp. 177–182. DOI: 10.1137/S1052623497318992. eprint: https://doi.org/10.1137/S1052623497318992. URL: https://doi.org/10.1137/S1052623497318992.

[87] Boris Polyak. *Introduction to Optimization*. Optimization Software, July 2020.

[88] Lukas Exl, Johann Fischbacher, Alexander Kovacs, Harald Oezelt, Markus Gusenbauer, and Thomas Schrefl. "Preconditioned nonlinear conjugate gradient method for micromagnetic energy minimization". In: *Computer Physics Communications* 235 (2019), pp. 179–186. ISSN: 0010-4655. DOI: https://doi.org/10.1016/j.cpc.2018.09.004. URL: https://www.sciencedirect.com/science/article/pii/S0010465518303187.

[89] Philip Wolfe. "Convergence Conditions for Ascent Methods". In: *SIAM Review* 11.2 (1969), pp. 226–235. DOI: 10.1137/1011036. eprint: https://doi.org/10.1137/1011036. URL: https://doi.org/10.1137/1011036.

[90] Tuyen Trung Truong and Hang-Tuan Nguyen. "Backtracking Gradient Descent Method and Some Applications in Large Scale Optimisation. Part 2: Algorithms and Experiments". In: *Applied Mathematics & Optimization* 84.3 (Dec. 2021), pp. 2557–2586. ISSN: 1432-0606. DOI: 10.1007/s00245-020-09718-8. URL: https://doi.org/10.1007/s00245-020-09718-8.

[91] Daniel A. Lidar and K. Birgitta Whaley. "Decoherence-Free Subspaces and Subsystems". In: *Irreversible Quantum Dynamics*. Berlin, Heidelberg: Springer Berlin Heidelberg, 2003, pp. 83–120. ISBN: 978-3-540-44874-7. DOI: 10.1007/3-540-44874-8_5. URL: https://doi.org/10.1007/3-540-44874-8_5.

[92] D. Bacon. *Decoherence, Control, and Symmetry in Quantum Computers*. 2003. arXiv: quant-ph/0305025 [quant-ph]. URL: https://arxiv.org/abs/quant-ph/0305025.

[93] A. K. Ratcliffe, R. L. Taylor, J. J. Hope, and A. R. R. Carvalho. "Scaling Trapped Ion Quantum Computers Using Fast Gates and Microtraps". In: *Physical Review Letters* 120.22 (May 2018). ISSN: 1079-7114. DOI: 10.1103/physrevlett.120.220501. URL: http://dx.doi.org/10.1103/PhysRevLett.120.220501.

[94] Qiming Sun, Timothy C. Berkelbach, Nick S. Blunt, George H. Booth, Sheng Guo, Zhendong Li, Junzi Liu, James D. McClain, Elvira R. Sayfutyarova, Sandeep Sharma, Sebastian Wouters, and Garnet Kin-Lic Chan. "PySCF: the Python-based simulations of chemistry framework". In: *WIREs Computational Molecular Science* 8.1 (Jan. 2018), e1340. ISSN: 1759-0876. DOI: 10.1002/wcms.1340. URL: https://doi.org/10.1002/wcms.1340.

[95] Qiming Sun, Xing Zhang, Samragni Banerjee, Peng Bao, Marc Barbry, Nick S. Blunt, Nikolay A. Bogdanov, George H. Booth, Jia Chen, Zhi-Hao Cui, Janus J. Eriksen, Yang Gao, Sheng Guo, Jan Hermann, Matthew R. Hermes, Kevin Koh, Peter Koval, Susi Lehtola, Zhendong Li, Junzi Liu, Narbe Mardirossian, James D. McClain, Mario Motta, Bastien Mussard, Hung Q. Pham, Artem Pulkin, Wirawan Purwanto, Paul J. Robinson, Enrico Ronca, Elvira R. Sayfutyarova, Maximilian Scheurer, Henry F. Schurkus, James E. T. Smith, Chong Sun, Shi-Ning Sun, Shiv Upadhyay, Lucas K. Wagner, Xiao Wang, Alec White, James Daniel Whitfield, Mark J. Williamson, Sebastian Wouters, Jun Yang, Jason M. Yu, Tianyu Zhu, Timothy C. Berkelbach, Sandeep Sharma, Alexander Yu. Sokolov, and Garnet Kin-Lic Chan. "Recent developments in the PySCF program package". In: *The Journal of Chemical Physics* 153.2 (July 2020), p. 024109. ISSN: 0021-9606. DOI: 10.1063/5.0006074. URL: https://doi.org/10.1063/5.0006074.

[96] R. B. Woodward and Roald Hoffmann. "Stereochemistry of Electrocyclic Reactions". In: *Journal of the American Chemical Society* 87.2 (Jan. 1965), pp. 395–397. ISSN: 0002-7863. DOI: 10.1021/ja01080a054. URL: https://doi.org/10.1021/ja01080a054.

[97] R.W. Richardson. "A restricted class of exact eigenstates of the pairing-force Hamiltonian". In: *Physics Letters* 3.6 (1963), pp. 277–279. ISSN: 0031-9163. DOI: https://doi.org/10.1016/0031-9163(63)90259-2. URL: https://www.sciencedirect.com/science/article/pii/0031916363902592.

[98] Thomas M. Henderson and Gustavo E. Scuseria. "Correlating the antisymmetrized geminal power wave function". In: *The Journal of Chemical Physics* 153.8 (Aug. 2020), p. 084111. ISSN: 0021-9606. DOI: 10.1063/5.0021144. eprint: https://pubs.aip.org/aip/jcp/article-pdf/doi/10.1063/5.0021144/15578147/084111_1_online.pdf. URL: https://doi.org/10.1063/5.0021144.

[99] Jarrod R. McClean, Kevin J. Sung, Ian D. Kivlichan, Yudong Cao, Chengyu Dai, E. Schuyler Fried, Craig Gidney, Brendan Gimby, Pranav Gokhale, Thomas Häner, Tarini Hardikar, Vojtěch Havlíček, Oscar Higgott, Cupjin Huang, Josh Izaac, Zhang Jiang, Xinle Liu, Sam McArdle, Matthew Neeley, Thomas O'Brien, Bryan O'Gorman, Isil Ozfidan, Maxwell D. Radin, Jhonathan Romero, Nicholas Rubin, Nicolas P. D. Sawaya, Kanav Setia, Sukin Sim, Damian S. Steiger, Mark Steudtner, Qiming Sun, Wei Sun, Daochen Wang, Fang Zhang, and Ryan Babbush. *OpenFermion: The Electronic Structure Package for Quantum Computers*. 2019. arXiv: 1710.07629 [quant-ph]. URL: https://arxiv.org/abs/1710.07629.

[100] John Preskill. "Quantum Computing in the NISQ era and beyond". In: *Quantum* 2 (Aug. 2018), p. 79. ISSN: 2521-327X. DOI: 10.22331/q-2018-08-06-79. URL: https://doi.org/10.22331/q-2018-08-06-79.

[101] Zhenyu Cai, Ryan Babbush, Simon C. Benjamin, Suguru Endo, William J. Huggins, Ying Li, Jarrod R. McClean, and Thomas E. O'Brien. "Quantum error mitigation". In: *Rev. Mod. Phys.* 95 (4 Dec. 2023), p. 045005. DOI: 10.1103/RevModPhys.95.045005. URL: https://link.aps.org/doi/10.1103/RevModPhys.95.045005.

[102] Suguru Endo, Zhenyu Cai, Simon C. Benjamin, and Xiao Yuan. "Hybrid Quantum-Classical Algorithms and Quantum Error Mitigation". In: *Journal of the Physical Society of Japan* 90.3 (Mar. 2021), p. 032001. ISSN: 0031-9015. DOI: 10.7566/JPSJ.90.032001. URL: https://doi.org/10.7566/JPSJ.90.032001.

[103] Kristan Temme, Sergey Bravyi, and Jay M. Gambetta. "Error Mitigation for Short-Depth Quantum Circuits". In: *Phys. Rev. Lett.* 119 (18 Nov. 2017), p. 180509. DOI: 10.1103/PhysRevLett.119.180509. URL: https://link.aps.org/doi/10.1103/PhysRevLett.119.180509.

[104] Dayue Qin, Xiaosi Xu, and Ying Li. "An overview of quantum error mitigation formulas". In: *Chinese Physics B* 31.9 (Aug. 2022), p. 090306. ISSN: 1674-1056. DOI: 10.1088/1674-1056/ac7b1e. URL: https://doi.org/10.1088/1674-1056/ac7b1e.

[105] X. Bonet-Monroig, R. Sagastizabal, M. Singh, and T. E. O'Brien. "Low-cost error mitigation by symmetry verification". In: *Phys. Rev. A* 98 (6 Dec. 2018), p. 062339. DOI: 10.1103/PhysRevA.98.062339. URL: https://link.aps.org/doi/10.1103/PhysRevA.98.062339.

[106] Thomas E. O'Brien, Stefano Polla, Nicholas C. Rubin, William J. Huggins, Sam McArdle, Sergio Boixo, Jarrod R. McClean, and Ryan Babbush. "Error Mitigation via Verified Phase Estimation". In: *PRX Quantum* 2 (2 May 2021), p. 020317. DOI: 10.1103/PRXQuantum.2.020317. URL: https://link.aps.org/doi/10.1103/PRXQuantum.2.020317.

[107] Stefano Polla, Gian-Luca R. Anselmetti, and Thomas E. O'Brien. "Optimizing the information extracted by a single qubit measurement". In: *Phys. Rev. A* 108 (1 July 2023), p. 012403. DOI: 10.1103/PhysRevA.108.012403. URL: https://link.aps.org/doi/10.1103/PhysRevA.108.012403.

[108] William J. Huggins, Sam McArdle, Thomas E. O'Brien, Joonho Lee, Nicholas C. Rubin, Sergio Boixo, K. Birgitta Whaley, Ryan Babbush, and Jarrod R. McClean. "Virtual Distillation for Quantum Error Mitigation". In: *Phys. Rev. X* 11 (4 Nov. 2021), p. 041036. DOI: 10.1103/PhysRevX.11.041036. URL: https://link.aps.org/doi/10.1103/PhysRevX.11.041036.

[109] Bálint Koczor. "Exponential Error Suppression for Near-Term Quantum Devices". In: *Phys. Rev. X* 11 (3 Sept. 2021), p. 031057. DOI: 10.1103/PhysRevX.11.031057. URL: https://link.aps.org/doi/10.1103/PhysRevX.11.031057.

[110] Mingxia Huo and Ying Li. "Dual-state purification for practical quantum error mitigation". In: *Phys. Rev. A* 105 (2 Feb. 2022), p. 022427. DOI: 10.1103/PhysRevA.105.022427. URL: https://link.aps.org/doi/10.1103/PhysRevA.105.022427.

[111] Fotios Gkritsis, Daniel Dux, Jin Zhang, Naman Jain, Christian Gogolin, and Philipp M. Preiss. "Simulating Chemistry with Fermionic Optical Superlattices". In: *PRX Quantum* 6 (1 Jan. 2025), p. 010318. DOI: 10.1103/PRXQuantum.6.010318. URL: https://link.aps.org/doi/10.1103/PRXQuantum.6.010318.

[112] Juan Miguel Arrazola, Olivia Di Matteo, Nicolás Quesada, Soran Jahangiri, Alain Delgado, and Nathan Killoran. "Universal quantum circuits for quantum chemistry". In: *Quantum* 6 (June 2022), p. 742. ISSN: 2521-327X. DOI: 10.22331/q-2022-06-20-742. URL: https://doi.org/10.22331/q-2022-06-20-742.

[113] Hugh G. A. Burton. *Accurate and gate-efficient quantum ansätze for electronic states without adaptive optimisation*. 2024. arXiv: 2312.09761 [physics.chem-ph]. URL: https://arxiv.org/abs/2312.09761.

[114] Hugh G. A. Burton, Daniel Marti-Dafcik, David P. Tew, and David J. Wales. "Exact electronic states with shallow quantum circuits from global optimisation". In: *npj Quantum Information* 9.1 (July 2023), p. 75. ISSN: 2056-6387. DOI: 10.1038/s41534-023-00744-2. URL: https://doi.org/10.1038/s41534-023-00744-2.

[115] Edward G. Hohenstein, Robert M. Parrish, and Todd J. Martínez. "Tensor hypercontraction density fitting. I. Quartic scaling second- and third-order Møller-Plesset perturbation theory". In: *The Journal of Chemical Physics* 137.4 (July 2012), p. 044103. ISSN: 0021-9606. DOI: 10.1063/1.4732310. URL: https://doi.org/10.1063/1.4732310.

[116] Seonghoon Choi, Ignacio Loaiza, and Artur F. Izmaylov. "Fluid fermionic fragments for optimizing quantum measurements of electronic Hamiltonians in the variational quantum eigensolver". In: *Quantum* 7 (Jan. 2023), p. 889. ISSN: 2521-327X. DOI: 10.22331/q-2023-01-03-889. URL: https://doi.org/10.22331/q-2023-01-03-889.

[117] Lukas Postler, Sascha Heuβen, Ivan Pogorelov, Manuel Rispler, Thomas Feldker, Michael Meth, Christian D. Marciniak, Roman Stricker, Martin Ringbauer, Rainer Blatt, Philipp Schindler, Markus Müller, and Thomas Monz. "Demonstration of fault-tolerant universal quantum gate operations". In: *Nature* 605.7911 (May 2022), pp. 675–680. ISSN: 1476-4687. DOI: 10.1038/s41586-022-04721-1. URL: https://doi.org/10.1038/s41586-022-04721-1.

[118] Wolfgang Paul. "Electromagnetic Traps for Charged and Neutral Particles (Nobel Lecture)". In: *Angewandte Chemie International Edition in English* 29.7 (July 1990), pp. 739–748. ISSN: 0570-0833. DOI: 10.1002/anie.199007391. URL: https://doi.org/10.1002/anie.199007391.

[119] Laird Egan, Dripto M. Debroy, Crystal Noel, Andrew Risinger, Daiwei Zhu, Debopriyo Biswas, Michael Newman, Muyuan Li, Kenneth R. Brown, Marko Cetina, and Christopher Monroe. "Fault-tolerant control of an error-corrected qubit". In: *Nature* 598.7880 (Oct. 2021), pp. 281–286. ISSN: 1476-4687. DOI: 10.1038/s41586-021-03928-y. URL: https://doi.org/10.1038/s41586-021-03928-y.

[120] Athena Caesura, Cristian L. Cortes, William Pol, Sukin Sim, Mark Steudtner, Gian-Luca R. Anselmetti, Matthias Degroote, Nikolaj Moll, Raffaele Santagati, Michael Streif, and Christofer S. Tautermann. *Faster quantum chemistry simulations on a quantum computer with improved tensor factorization and active volume compilation*. 2025. arXiv: 2501.06165 [quant-ph]. URL: https://arxiv.org/abs/2501.06165.

[121] Jeffery Yu, Javier Robledo Moreno, Joseph T. Iosue, Luke Bertels, Daniel Claudino, Bryce Fuller, Peter Groszkowski, Travis S. Humble, Petar Jurcevic, William Kirby, Thomas A. Maier, Mario Motta, Bibek Pokharel, Alireza Seif, Amir Shehata, Kevin J. Sung, Minh C. Tran, Vinay Tripathi, Antonio Mezzacapo, and Kunal Sharma. *Quantum-Centric Algorithm for Sample-Based Krylov Diagonalization*. 2025. arXiv: 2501.09702 [quant-ph]. URL: https://arxiv.org/abs/2501.09702.

[122] Ville Bergholm, Josh Izaac, Maria Schuld, Christian Gogolin, Shahnawaz Ahmed, Vishnu Ajith, M. Sohaib Alam, Guillermo Alonso-Linaje, B. AkashNarayanan, Ali Asadi, Juan Miguel Arrazola, Utkarsh Azad, Sam Banning, Carsten Blank, Thomas R Bromley, Benjamin A. Cordier, Jack Ceroni, Alain Delgado, Olivia Di Matteo, Amintor Dusko, Tanya Garg, Diego Guala, Anthony Hayes, Ryan Hill, Aroosa Ijaz, Theodor Isacsson, David Ittah, Soran Jahangiri, Prateek Jain, Edward Jiang, Ankit Khandelwal, Korbinian Kottmann, Robert A. Lang, Christina Lee, Thomas Loke, Angus Lowe, Keri McKiernan, Johannes Jakob Meyer, J. A. Montañez-Barrera, Romain Moyard, Zeyue Niu, Lee James O'Riordan, Steven Oud, Ashish Panigrahi, Chae-Yeun Park, Daniel Polatajko, Nicolás Quesada, Chase Roberts, Nahum Sá, Isidor Schoch, Borun Shi, Shuli Shu, Sukin Sim, Arshpreet Singh, Ingrid Strandberg, Jay Soni, Antal Száva, Slimane Thabet, Rodrigo A. Vargas-Hernández, Trevor Vincent, Nicola Vitucci, Maurice Weber, David Wierichs, Roeland Wiersema, Moritz Willmann, Vincent Wong, Shaoming Zhang, and Nathan Killoran. *PennyLane: Automatic differentiation of hybrid quantum-classical computations*. 2018. URL: https://arxiv.org/abs/1811.04968.

[123] Fabian Kreppel, Christian Melzer, Diego Olvera Millán, Janis Wagner, Janine Hilder, Ulrich Poschinger, Ferdinand Schmidt-Kaler, and André Brinkmann. "Quantum Circuit Compiler for a Shuttling-Based Trapped-Ion Quantum Computer". In: *Quantum* 7 (Nov. 2023), p. 1176. ISSN: 2521-327X. DOI: 10.22331/q-2023-11-08-1176. URL: https://doi.org/10.22331/q-2023-11-08-1176.

[124] Blizzard Entertainment. *World of Warcraft: Wrath of the Lich King*. (PC). Blizzard Entertainment. Scene: Fall of the Lich King. 2008.

[125] David A. Rower, Leon Ding, Helin Zhang, Max Hays, Junyoung An, Patrick M. Harrington, Ilan T. Rosen, Jeffrey M. Gertler, Thomas M. Hazard, Bethany M. Niedzielski, Mollie E. Schwartz, Simon Gustavsson, Kyle Serniak, Jeffrey A. Grover, and William D. Oliver. "Suppressing Counter-Rotating Errors for Fast Single-Qubit Gates with Fluxonium". In: *PRX Quantum* 5 (4 Dec. 2024), p. 040342. DOI: 10.1103/PRXQuantum.5.040342. URL: https://link.aps.org/doi/10.1103/PRXQuantum.5.040342.

[126] Simone D. Fasciati, Boris Shteynas, Giulio Campanaro, Mustafa Bakr, Shuxiang Cao, Vivek Chidambaram, James Wills, and Peter J. Leek. *Complementing the transmon by integrating a geometric shunt inductor*. 2024. arXiv: 2410.10416 [quant-ph]. URL: https://arxiv.org/abs/2410.10416.

[127] Blayney W. Walshe, Ben Q. Baragiola, Hugo Ferretti, José Gefaell, Michael Vasmer, Ryohei Weil, Takaya Matsuura, Thomas Jaeken, Giacomo Pantaleoni, Zhihua Han, Timo Hillmann, Nicolas C. Menicucci, Ilan Tzitrin, and Rafael N. Alexander. "Linear-Optical Quantum Computation with Arbitrary Error-Correcting Codes". In: *Phys. Rev. Lett.* 134 (10 Mar. 2025), p. 100602. DOI: 10.1103/PhysRevLett.134.100602. URL: https://link.aps.org/doi/10.1103/PhysRevLett.134.100602.

[128] Scott A. Manifold, George B. Long, and Jonathan J. Burnett. *Thermal Capacity Mapping of Cryogenic Platforms for Quantum Computers*. 2025. arXiv: 2503.10775 [quant-ph]. URL: https://arxiv.org/abs/2503.10775.

[129] Aziza Almanakly, Beatriz Yankelevich, Max Hays, Bharath Kannan, Réouven Assouly, Alex Greene, Michael Gingras, Bethany M. Niedzielski, Hannah Stickler, Mollie E. Schwartz, Kyle Serniak, Joel Î-j. Wang, Terry P. Orlando, Simon Gustavsson, Jeffrey A. Grover, and William D. Oliver. "Deterministic remote entanglement using a chiral quantum interconnect". In: *Nature Physics* (Mar. 2025). ISSN: 1745-2481. DOI: 10.1038/s41567-025-02811-1. URL: https://doi.org/10.1038/s41567-025-02811-1.